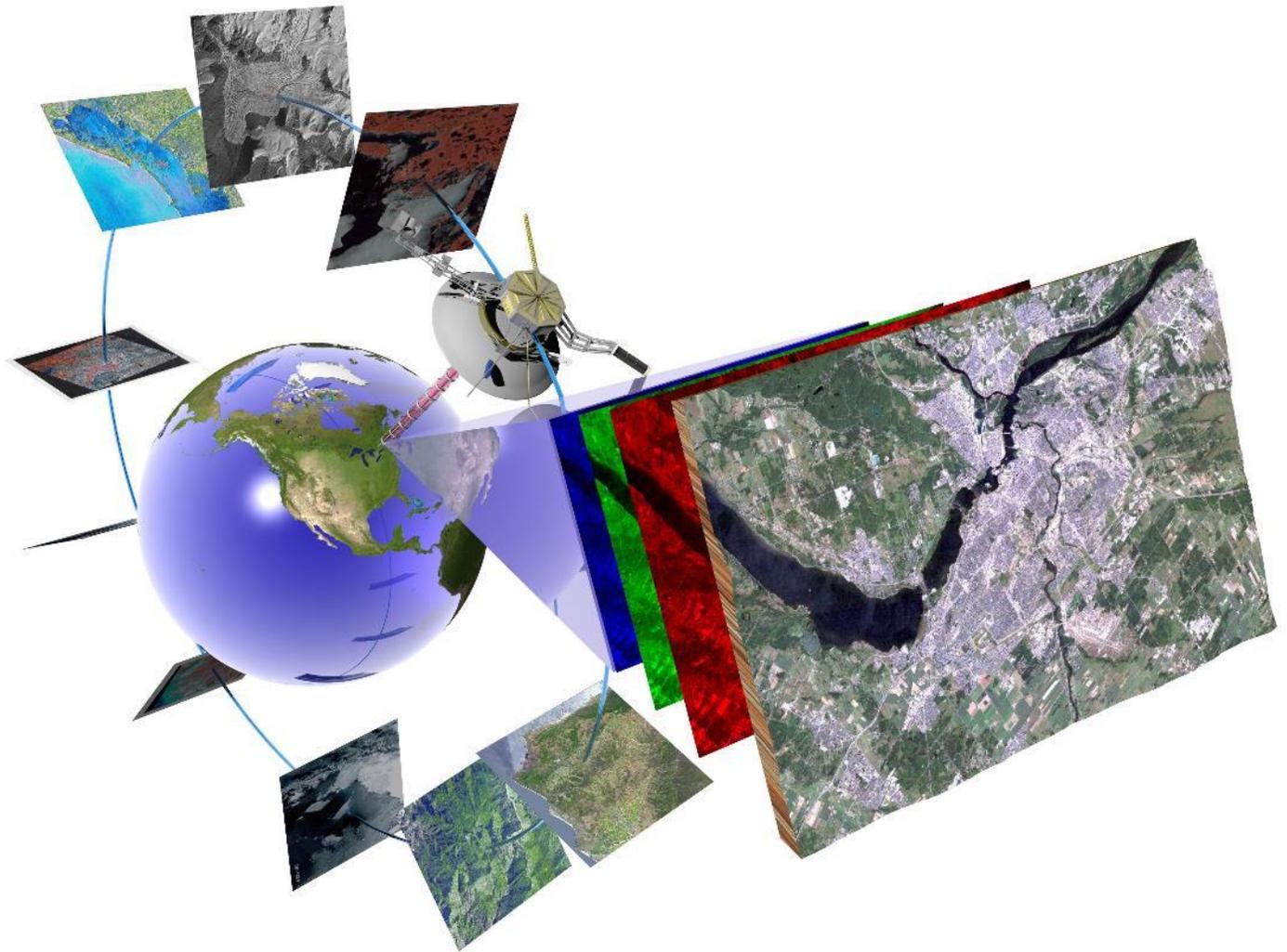


Notes des labos pour ***SIG et la Terre numérique***



Ces notes ont été élaborées par Michael Sawada, Anders Knudby, Galen Richardson et Claudia Sauro, à l'Université d'Ottawa. Elles sont mises gratuitement à la disposition du public en tant que ressource éducative ouverte sous licence CC BY 4.0.

Contents

Introduction	4
Comment écrire un bon rapport	1
Comment faire un bon graphique	1
Comment faire un bon tableau	3
Comment faire une bonne carte	5
Comment faire référence à des documents publiés dans votre travail	9
Laboratoire n°1 : Conception cartographique, symbologie et problème d'agrégation spatiale	1
Section 1: Conception cartographique	2
Question 1: Quel est l'objectif des cartes créées dans ce laboratoire et comment puis-je accéder aux données?.....	3
Question 2: Comment créer une carte et symboliser une classe d'entités?.....	5
Question 3: Quelles sont les façons courantes de symboliser une classe d'entités polygonales?.....	5
Question 4: Que sont les intervalles et comment affectent-elles l'apparence de ma carte?	9
Question 5: Comment créer une mise en page de carte et définir l'échelle de la carte?	10
Question 6: Comment concevoir des éléments cartographiques esthétiques?.....	12
Question 7: Comment exporter une carte?.....	14
Question 8: Comment étiqueter les éléments de la carte?.....	14
Question 9: Comment modifier l'emplacement des étiquettes?.....	15
Section 2: Problème d'agrégation spatiale.....	18
Question 10: Qu'est-ce que le Problème d'agrégation spatiale?.....	18
Question 11: Comment créer différentes zones pour visualiser le problème d'agrégation spatiale?	23
Question 12: Comment symboliser deux classes d'entités à l'aide de la même symbologie?.....	26
Questions à remettre	27
Laboratoire n°2 : Systèmes de coordonnées et transformations	29
Q1: Pourquoi la projection des données est-elle importante dans les SIG ?	30
Q2: Comment convertir les degrés, minutes et secondes en degrés décimaux?.....	31
Q3: Quelle est la distance réelle d'un degré de longitude à l'équateur?.....	34
Q4: Comment calculer la distance réelle entre deux points en utilisant la longitude et la latitude?	35
Q5: Comment calculer l'incertitude de localisation d'une couche numérique?.....	36
Q6: Quelle est l'incertitude de localisation pour une représentation raster d'une couche géospatiale numérique?	40
Q7: Si ArcGIS Pro ne reconnaît pas le système de coordonnées d'une couche, comment puis-je indiquer à ArcGIS le système de coordonnées de la couche?	42
Q8: Comment convertir entre NAD27 et NAD83 au Canada?.....	48
Q9: Comment puis-je spécifier la conversion NAD27 et NAD83 après avoir ajouté les données à ArcGIS Pro?	53
Q10: Comment obtenir les coordonnées NAD27 et NAD83 pour chaque point dans un ensemble de points?	56

Nom de l'étudiant: _____ # d'étudiant: _____

Q11: Comment effectuer un changement permanent en NAD83 (ou tout autre système de coordonnées) pour une couche géospatiale numérique?	65
Questions à remettre	67
Laboratoire n°3: Travailler avec des bases de données géographiques.....	72
Q1: Comment télécharger des données à partir du Scholars GeoPortal?	73
Q2: Où puis-je obtenir les données correspondant aux secteurs de recensement?	77
Q3: Comment peut-on relier la couche de données géospatiales Census_tracts à la table de données d'attributs contenant les informations sur la population ?.....	89
Q4: Que puis-je faire maintenant que j'ai établi une classe de relation entre ma table de population de recensement et ma classe d'entités de recensement ?	92
Q5: Comment puis-je cartographier la population dans les secteurs de recensement à l'aide de mon tableau de population?	95
Q6: Je veux calculer la densité de population ; comment puis-je le faire?.....	98
Questions à remettre	103
Laboratoire #4 : Analyse spatiale avec vecteur	107
Q1: Comment effectuer une requête spatiale?.....	108
Q2: Quel est le lien entre les requêtes SQL et les requêtes spatiales (« Select layer by location »)?.....	111
Q3: Comment créer une requête SQL avec un champ de type date?	113
Q4: Comment entreprendre des mesures centrophiques pour un ensemble de points?....	116
Q5: Comment visualiser la variabilité spatiale de l'intensité des points à l'intérieur d'une région?	120
Q6: Comment calculer les agressions par unité de surface dans une région?	127
Q7: Comment créer une superposition d'union et pourquoi?.....	133
Questions à remettre	141
Laboratoire n°5 : Télédétection et SIG raster	143
Q1: Comment puis-je visualiser différentes combinaisons de bandes d'une image télédétectée?	144
Q2: Comment classifier une image télédétectée?	147
Q3: Comment créer un ensemble de données d'entraînement correspondant à mes classes d'information?.....	167
Q4: Comment créer un fichier de signature à partir de mes polygones d'entraînement?.....	172
Q5: Comment puis-je classer mon image en utilisant la classification d'image supervisée? ...	177
Q6: Comment puis-je nettoyer ma couche raster classifiée?	179
Q7: Comment généraliser mon raster classifié "MLCLASS_RE_21_filtered" pour que la taille des cellules soit la même que celle de la couche "MLCLASS_L7_21" de Q2?	182
Q8: Comment puis-je vérifier l'exactitude d'un résultat classifié?	187
Q9: Comment puis-je créer un raster contenant uniquement les forêts à partir de chacun de mes résultats classifiés, et déterminer où les forêts ont diminué?	194
Questions à remettre	196

Nom de l'étudiant: _____ # d'étudiant: _____

Introduction

Ce manuel de laboratoire est utilisé à l'Université d'Ottawa pour le cours intitulé SIG et la Terre numérique (code de cours GEG2320). Il peut être utilisé librement pour enseigner différents cours, et être modifié et distribué librement, avec une attribution appropriée conformément à la licence [CC BY 4.0](#).

Le document contient une section d'introduction avec des directives générales sur la façon de produire des figures, des tableaux et des cartes de qualité. Les étudiants sont tenus de se référer à ces directives et de les respecter lorsqu'ils remettent leurs réponses pour chaque laboratoire. Ces lignes directrices générales sont suivies de cinq exercices de laboratoire distincts, chacun consistant en i) des étapes guidées démontrant une analyse de données effectuée dans ArcGIS, suivies de ii) une série de questions à remettre.

Comment écrire un bon rapport

Comment faire un bon graphique

Les valeurs par défaut créées par Excel et d'autres tableurs produisent généralement des graphiques de mauvaise qualité. Comparez les deux graphiques de la Figure 1 et de la Figure 2. Dans la plupart des disciplines, les graphiques devraient ressembler à la Figure 2. Notez que les graphiques ne comportent pas de titre ; la description appropriée de la figure (légende de la figure) est fournie sous le graphique.

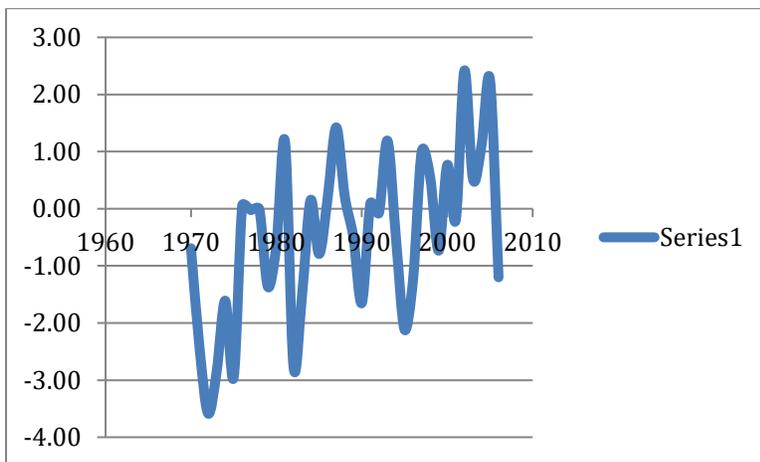


Figure 1. Température annuelle moyenne de l'air à Whitehorse (Territoire du Yukon) pour la période 1970-2006. Le graphique a été réalisé en utilisant les paramètres par défaut d'Excel et enfreint un certain nombre de règles relatives à la réalisation de graphiques.

Nom de l'étudiant: _____ # d'étudiant: _____

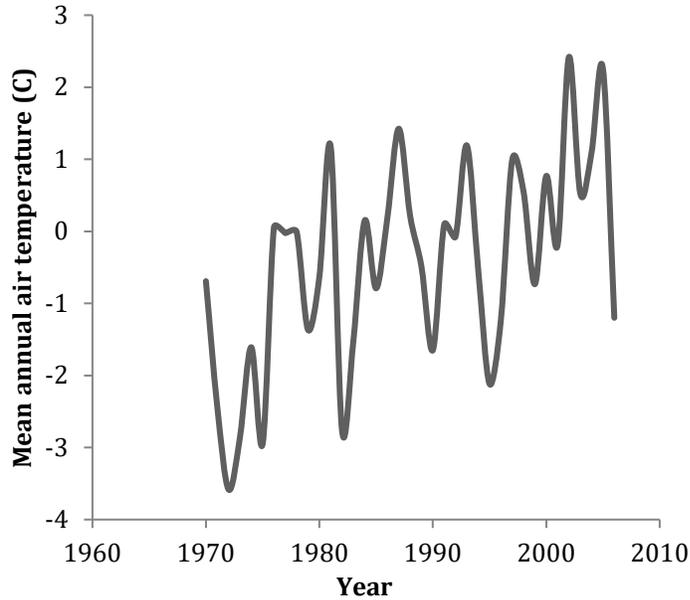


Figure 2. Température annuelle moyenne de l'air à Whitehorse (Territoire du Yukon) pour la période 1970-2006. Le graphique a été ajusté pour des raisons esthétiques : suppression des lignes de grille, intersection de l'axe des x et de l'axe des y à la valeur minimale, ligne d'une épaisseur de 1 pt, titres des axes des x et des y, réduction de la taille du graphique.

Voici d'autres exemples de figures créées avec R (un logiciel de statistiques ; vous apprendrez probablement ce logiciel dans des cours plus avancés) (Figure 3).

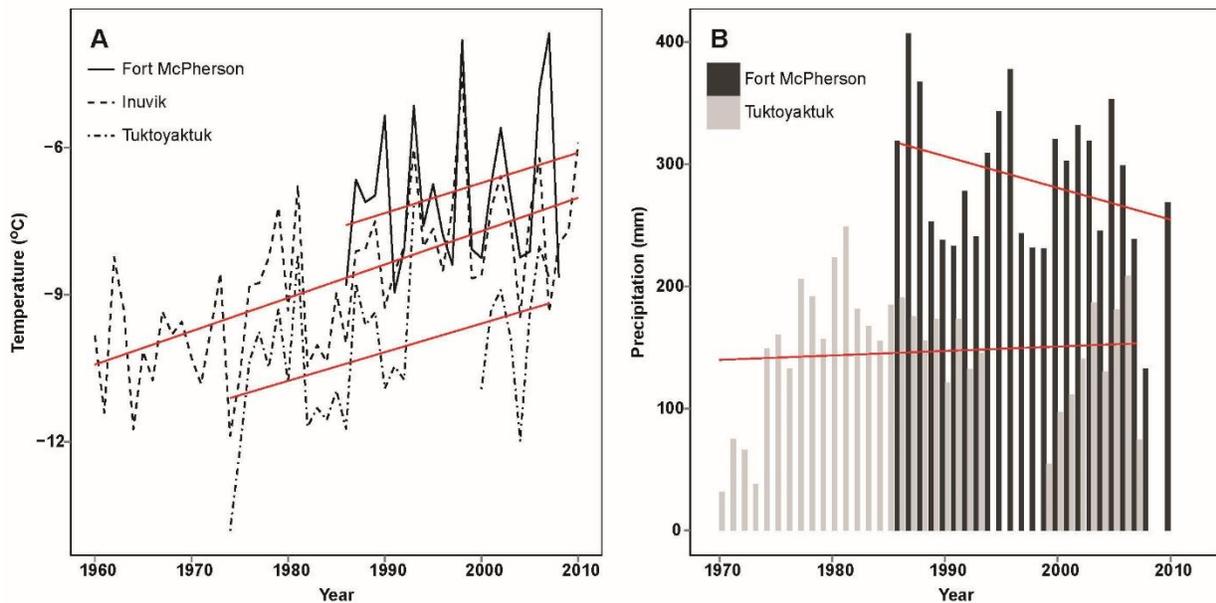


Figure 3. A) Température annuelle moyenne de l'air de 1960 à 2010 pour Fort McPherson, Inuvik et Tuktoyaktuk, T.N.-O., Canada. Les lignes rouges sont des lignes de régression linéaire pour chaque station. B) Précipitations totales de 1970 à 2010 pour Fort McPherson et Tuktoyaktuk, T.N.-O., Canada

Conseils pour faire un bon graphique:

Nom de l'étudiant: _____ # d'étudiant: _____

- Vous pouvez probablement utiliser des graphiques "linéaires" ou "en nuage de points x-y" (ces termes sont ceux d'Excel et ne sont pas d'usage courant). Les graphiques "linéaires" supposent que les points sont régulièrement espacés (c'est-à-dire quotidiennement ou annuellement) et offrent moins d'options ; les graphiques "x-y scatter" sont mieux adaptés à un espacement irrégulier ou à des données manquantes. Si des données sont manquantes, par exemple si l'année a été omise par inadvertance, les graphiques linéaires, par exemple, peuvent poser problème. Jouez avec les différentes options de graphiques et voyez ce qui fonctionne le mieux. Dans certaines situations, l'utilisation d'un graphique linéaire alors qu'il faudrait utiliser un graphique en nuage de points entraînera des problèmes ; les graphiques linéaires ou en nuage de points sont appropriés dans des situations différentes. Apprenez dans quelles circonstances vous devriez utiliser l'un ou l'autre.
- Utilisez une légende - sous le graphique ou la série de graphiques - plutôt qu'un titre pour chaque graphique.
- Si vous avez plusieurs séries de données, utilisez une légende, ou vous pouvez étiqueter les lignes sur le graphique.
- Désactivez la couleur d'arrière-plan et supprimez les lignes de la grille. Celles-ci encombrant le graphique.
- Sur les échelles des axes, utilisez des points décimaux et un espacement appropriés. Par exemple, il est courant d'indiquer les années de la décennie (1900, 1910, ...) plutôt que des années arbitraires ou intermédiaires. Vous devriez éviter d'utiliser des décimales sur les étiquettes des axes, mais cela dépend de l'ensemble de données particulier.
- Veillez à ce que la police de caractères soit suffisamment grande pour être lue sur une copie imprimée. En règle générale, la taille finale du texte doit être au minimum de 6 points.
- Dans les graphiques, il est préférable d'utiliser un texte sans empattement. **Arial est une police sans empattement. Times New Roman est une police avec empattement.** Vous devez généralement utiliser la même police pour tous les graphiques. Évitez toutes les polices fantaisistes disponibles. Alors qu'une police avec empattement peut être plus facile à lire dans un paragraphe, dans une figure ou une carte, elle donne l'impression d'avoir été tapée à la machine et d'être moins professionnelle. Il est également préférable de s'en tenir aux polices standard ; l'utilisation d'une police rare ou inhabituelle peut parfois poser des problèmes lorsque vous changez d'ordinateur ou de logiciel.
- N'utilisez pas de diagrammes circulaires.

Comment faire un bon tableau

Les tableaux 1 et 2 montrent deux tableaux différents représentant les mêmes données. Tableau 1 est un exemple de tableau qui ne répond pas aux normes attendues dans les rapports et les essais scientifiques. Tableau 2 montre le format que vous devez utiliser pour présenter un tableau dans votre travail.

Nom de l'étudiant: _____ # d'étudiant: _____

Tableau 1

Rang	Pays d'étude	Population de chaque pays	Pourcentage de la population mondiale
1	La Chine	1,371,380,000	18.94289
2	L'Inde	1,275,260,000	17.6
3	Les États-Unis	321,564,000	4.431
4	L'Indonésie	255,770,000	3.5223
5	Le Brésil	204,714,000	2.82

Ce tableau montre le nombre de personnes vivant en Chine, en Inde, aux États-Unis et en Indonésie, qui sont les pays les plus peuplés. Les données pour le Brésil sont basées sur la participation volontaire de ses citoyens au recensement annuel.

Tableau 2

Tableau 1. Top 5 des pays les plus peuplés du monde et pourcentage relatif par rapport à la population mondiale totale.

Rang	Pays	Population	% de la population mondiale
1	La Chine	1,371,380,000	18.9
2	L'Inde	1,275,260,000	17.6
3	Les États-Unis	321,564,000	4.4
4	L'Indonésie	255,770,000	3.5
5	Le Brésil	204,714,000	2.8

Note : Les données pour le Brésil sont basées sur la participation volontaire au recensement annuel. Les données pour le Brésil sont basées sur la participation volontaire au recensement annuel. Source : Office des statistiques internationales : Office des statistiques internationales

Les deux tableaux suivants sont des exemples de variations du tableau de base. Le tableau 3 montre comment diviser une colonne en deux colonnes différentes à l'aide de lignes horizontales. Le tableau 4 montre comment délimiter les lignes en utilisant une légère nuance de gris. Vous ne reviendrez à l'ombrage que lorsque votre tableau contiendra de nombreuses colonnes et lignes qui rendront sa lecture difficile.

Tableau 3. Répartition altitudinale des fleurs et des arbres endémiques et indigènes de Stony Hill, Québec, Canada.

Altitude (m)	Fleurs		Arbres	
	Endémiques	Indigènes	Endémiques	Indigènes
200	0	32	0	11
400	0	30	0	8
600	1	28	0	7
800	2	20	1	8
1000	1	21	0	3
1200	4	15	2	3

Nom de l'étudiant: _____ # d'étudiant: _____

Tableau 4. Estimation des prises accessoires des chalutiers de pêche. Les données sont exprimées en tonnes et représentent le total par mois pour la période 2010-2014.

	Ayu	Brill	Dory	Flier	Hake	Jack	Ling	Porgy	Rudd	Sole	Walu
Jan.	12	8	45	56	85	94	82	96	98	32	36
Fev.	56	32	56	11	18	62	66	35	33	39	95
Mar.	98	11	47	89	85	8	56	6	5	12	39
Apr.	5	8	74	21	11	9	85	23	25	96	36
Mai	6	12	62	35	33	39	21	95	6	17	46
Juin	11	18	54	9	7	32	23	12	95	21	63
Juil.	3	68	14	21	8	15	45	65	61	77	11
Août	98	45	56	56	11	18	86	62	35	33	39
Sept.	32	0	43	57	39	28	47	66	72	80	11
Oct.	64	0	11	18	62	35	12	33	39	95	6
Nov.	26	34	33	78	95	23	5	12	8	9	2
Dec.	12	48	26	54	85	8	2	6	5	12	39

Conseils pour faire un bon tableau:

- La légende du tableau (description du tableau) se trouve toujours au-dessus du tableau, elle est en italique, alignée à gauche et commence par "Tableau 1".
- La légende du tableau est claire, concise et informative.
- Les titres des colonnes sont clairs, concis et informatifs.
- Les symboles, signes, abréviations ou acronymes sont explicites ou expliqués dans une note.
- Le nombre de décimales est cohérent.
- Il n'y a pas de lignes de séparation entre les colonnes, l'espace entre elles est suffisant.
- Il n'y a pas de caractères gras, italiques ou soulignés inutiles. Ceux-ci peuvent être utilisés pour mettre en évidence des valeurs individuelles, mais il faut alors une note explicative.
- Les lignes horizontales sont réduites au minimum ; elles sont utilisées pour séparer les en-têtes de colonne et à la fin du tableau. Ne pas utiliser de lignes verticales.
- La police de caractères est lisible et cohérente (toujours utiliser une police sans empattement) ; ne jamais utiliser un texte plus petit que 6 points.
- Si nécessaire, des informations supplémentaires sont fournies dans une note située sous le tableau et dans une police plus petite.
- La source des données est indiquée sous le tableau.
- Le tableau est intelligible en soi.
- Les tableaux doivent être placés le plus près possible de l'endroit où ils sont mentionnés pour la première fois dans le texte.

Comment faire une bonne carte

Les Figure 4 et Figure 5 présentent les mêmes données sous la forme d'une carte choroplèthe thématique. L'objectif de la carte est de montrer la répartition de la population en Écosse. La

Nom de l'étudiant: _____ # d'étudiant: _____

Figure 4 est incorrecte et ne répond pas aux normes attendues dans les rapports et les essais scientifiques. La Figure 5 répond aux normes minimales d'une carte.

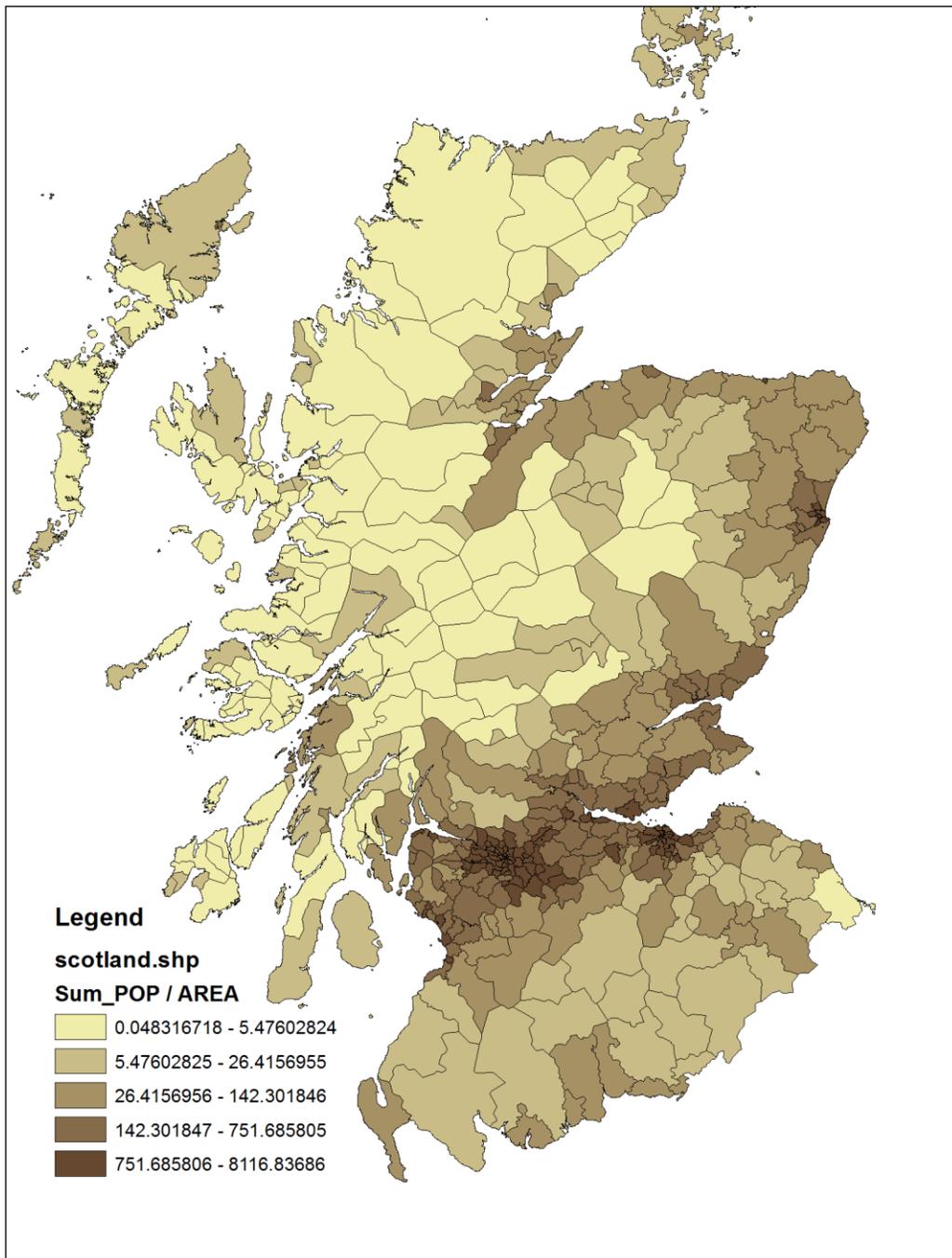


Figure 4. Densité de la population en Écosse.

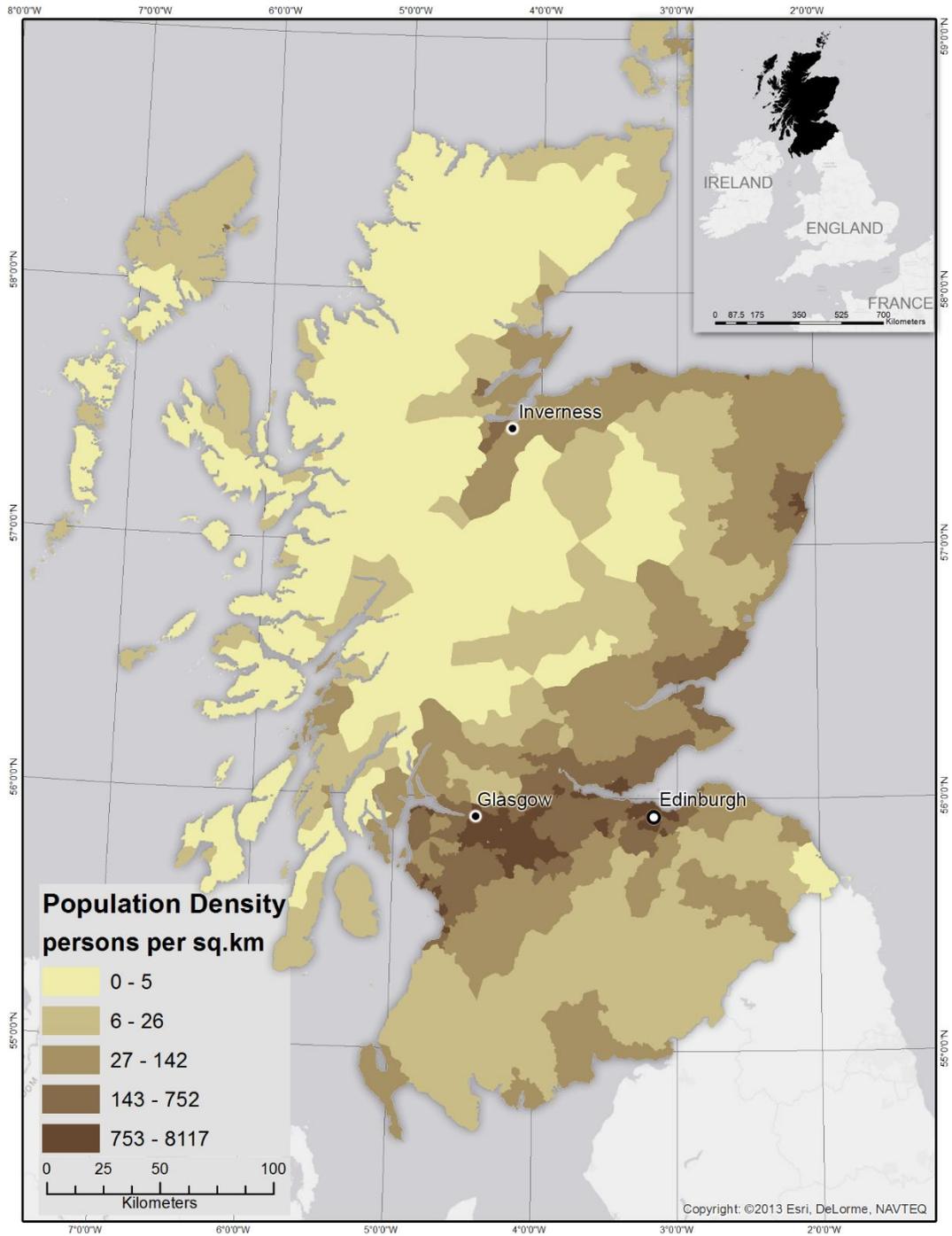


Figure 5. Écosse.

La Figure 4 est une carte inacceptable pour de nombreuses raisons:

- La carte de la figure 4 utilise des contours de polygones alors qu'ils n'ajoutent rien à l'objectif de la carte, qui est de montrer la densité de population - les contours de polygones sont un paramètre par défaut dans la plupart des systèmes SIG. Pourquoi les polygones sont-ils délimités ? La carte concerne-t-elle les divisions administratives en Écosse ou la répartition de la population en Écosse ? Si les limites des unités de

Nom de l'étudiant: _____ # d'étudiant: _____

données ne sont pas nécessaires pour transmettre le message thématique, elles ne devraient jamais être incluses, comme le montre la figure 5.

- La figure 4 ne présente aucun contexte géographique. Si quelqu'un n'est pas déjà familiarisé avec l'Écosse, il n'aura aucune idée de la situation géographique de l'Écosse. L'inclusion d'une carte incrustée, comme le montre la figure 5, permet de remédier à ce problème et d'inclure les pays environnants de couleur homogène.
- Dans la figure 4, comment une personne ne connaissant pas l'Écosse peut-elle expliquer où la population est élevée/faible ou pourquoi elle est élevée/faible dans certaines régions ? Quelques grandes villes incluses et étiquetées sur la carte aideraient à résoudre ce problème, comme cela a été fait dans la figure 5.
- La figure 4 ne fait pas référence au Nord. Si une carte NE représente PAS un continent entier, une référence au nord est nécessaire. La référence à la direction du nord peut être faite en utilisant soit une flèche du nord, soit, comme dans la figure 5, un graticule.
- La figure 4 ne permet pas de déterminer les coordonnées d'un point quelconque de la carte. L'utilisation d'un graticule étiqueté, comme le montre la figure 5, permet un positionnement plus précis.
- La figure 4 ne permet pas de déterminer les coordonnées d'un lieu sur la carte. L'inclusion d'un graticule étiqueté, comme le montre la figure 5, permet un positionnement plus précis.
- La figure 4 ne permet pas de déterminer la taille relative des éléments. La figure 5, en revanche, comporte une barre d'échelle claire.
- La légende de la figure 4 est totalement incorrecte. Tout d'abord, il est évident que l'élément situé en bas à gauche de la carte est une légende, parfois appelée clé de la carte. Elle n'a pas besoin d'un titre indiquant "Légende". De plus, qu'est-ce que Scotland.shp dans la légende ? Qu'est-ce que Sum_POP / AREA et quels sont les chiffres, quelles sont les unités ? L'étiquetage de la légende est incorrect. Les nombres figurant dans la légende amènent le lecteur à croire que tout ce qui est cartographié est exact à 9 décimales près. Peu de choses sont exactes à la neuvième décimale. Comparez avec la légende de la figure 5. Une légende doit indiquer clairement la variable représentée, les unités de mesure doivent être clairement indiquées et les données doivent être présentées à un niveau de précision raisonnable.
- La légende de la figure 4 contient des informations redondantes. Une légende ne doit jamais répéter le titre ou le contenu d'une carte. La légende d'une carte doit être suffisante pour que la personne qui lit la carte sache ce qu'elle présente. Dans la figure 5, vous pouvez voir que la légende de la carte indique simplement "Écosse" parce que c'est la seule information nécessaire à la lecture de la carte, qu'elle a déjà un contexte régional et que la légende montre clairement ce que la carte représente.

Il existe bien d'autres règles en cartographie, mais la comparaison ci-dessus devrait vous aider à comprendre la composition d'une carte au niveau le plus élémentaire. En général, les règles suivantes doivent être appliquées à toutes les cartes que vous produisez :

- Symbolisez votre carte en fonction de son objectif. Tout symbole superflu doit être éliminé.

Nom de l'étudiant: _____ # d'étudiant: _____

- Chaque carte doit comporter un titre ou une légende, cette dernière étant préférable, une indication du nord sous la forme d'une flèche ou d'un réticule, une légende correctement composée, des étiquettes d'axes (x, y ou longitude, latitude) et une barre d'échelle.
- Les cartes thématiques comportent un nombre minimum de symboles contextuels que le lecteur peut utiliser pour ancrer les descriptions de la carte.
- Les cartes choroplèthes ne sont utilisées que pour les variables non dimensionnelles ou les variables de densité. Elles ne sont jamais utilisées pour présenter des variables de comptage.
- Utiliser des symboles proportionnels pour présenter les variables de comptage sur les cartes.

Comment faire référence à des documents publiés dans votre travail

Dans de nombreuses disciplines, la manière habituelle de référencer les documents que vous avez consultés est la méthode « author-date » ou le format APA. Les notes de bas de page ne sont généralement pas utilisées (bien qu'elles soient courantes en géographie humaine) et les citations sont évitées. À la fin de chaque travail, vous devez référencer tous les documents (à l'exception des notes de cours ou de lecture) que vous avez utilisés.

Chaque revue académique a son propre format pour référencer les travaux publiés. Dans ce cours, vous devez suivre le format APA : les références dans le texte doivent être données entre parenthèses avec un espace séparant le(s) auteur(s) et la date : e.g. *Les processus de lavage de pente n'ont pas été étudiés de manière approfondie (Clark 1988) mais certains travaux ont été effectués sur l'île de Banks (Lewkowicz 1983)*. Les auteurs multiples sont désignés par (Ex et Why 1986) ou, s'il y a plus de deux auteurs, par (Ex et al. 1986). Veuillez respecter ces conventions ! Rien ne contrarie plus un auteur que d'être cité incorrectement (par exemple, si les auteurs sont Monroe et Wicander (1993), citez le livre comme Wicander et Monroe (1993) ou, pire encore, Monroe (1993)).

À la fin de votre travail, il doit y avoir une section "Références". Dans cette section, les références sont classées par ordre alphabétique d'auteur. Si plusieurs références proviennent du même auteur, elles sont classées par ordre chronologique, de la plus ancienne à la plus récente. Vous ne devez pas citer des documents que vous n'avez pas consultés (par exemple, ne lisez pas un texte qui fait référence à un autre ouvrage et ne citez pas ce dernier comme si vous l'aviez lu). S'il est essentiel que vous citiez le second ouvrage et que vous ne pouvez pas obtenir l'original, il est permis de le citer comme Cruden (1974) en French (1976)).

Le format d'un article de revue est le suivant:

Mackay, J.R. 1971. The origin of massive icy beds in permafrost, western arctic coast. Canadian Journal of Earth Sciences, **8**: 397-422.

Format for a book:

Nom de l'étudiant: _____ # d'étudiant: _____

French, H.M. 1996. The periglacial environment. 2nd Edition. Longman, London.

Format d'un chapitre de livre:

Van Vliet-Lanöe, B. 1985. Frost effects in soils. *In* Soils and Quaternary landscape evolution. Edited by J. Boardman. Wiley, Chichester, pp. 117-158.

Format d'un site internet:

Indiquez l'adresse du site web et la date à laquelle vous l'avez consulté comme suit :

www.nrc.ca/cgi-bin/cisti/, October 15, 2005.

Si le travail que vous souhaitez citer n'entre dans aucune de ces catégories, vous pouvez vérifier le format approprié en consultant un volume récent de la revue Canadian Journal of Earth Sciences (<http://www.nrcresearchpress.com/journal/cjes>) ou le style de référencement APA.

Nom de l'étudiant: _____ # d'étudiant: _____

Laboratoire n°1 : Conception cartographique, symbologie et problème d'agrégation spatiale

Objectif : Ce laboratoire vous présentera les mises en page cartographiques et vous guidera dans la création de cartes publiables et dans la compréhension des méthodes de représentation des données par le biais de cartes. Les exercices renforceront ces concepts à travers la création de quatre mises en page de cartes qui présentent les meilleures pratiques cartographiques.

Résultats d'apprentissage: A la fin de ce laboratoire, les étudiants seront capables de:

- Créer des mises en page cartographiques qui respectent la hiérarchie visuelle de la carte et utilisent les éléments cartographiques de manière appropriée.
- Étiqueter des caractéristiques de points et de polygones en utilisant différentes stratégies de placement d'étiquettes et d'équations d'étiquetage.
- Comprendre les différents types de séparations et la symbologie principale pour les éléments polygonaux.
- Expliquer le problème d'agrégation spatiale.

Procédure:

- Lisez les exemples de questions dans les sections 1 et 2.
- Répétez les exemples de questions dans ArcGIS Pro.
- Répondez aux questions d'évaluation.

Données/matériel:

- Liens vers le catalogue de données ouvertes de la ville de Surrey.

Documentation complémentaire:

- Aide ArcGIS
- Source de données: <https://data.surrey.ca/>
- Lecture: Slocum, T.A., McMaster, R.B., Kessler, F.C., & Howard, H.H. (2022). Thematic Cartography and Geovisualization, Fourth Edition (4th ed.). CRC Press. <https://doi.org/10.1201/9781003150527>
- Lecture: Fotheringham, A. S., & Wong, D. W. S. (1991). The Modifiable Areal Unit Problem in Multivariate Statistical Analysis. Environment and Planning A: Economy and Space, 23(7), 1025–1044. <https://doi.org/10.1068/a231025>
- Glossaire de Statistique Canada: <https://www150.statcan.gc.ca/n1/pub/92-195-x/2021001/geo/ct-sr/ct-sr-eng.htm>

Devoir:

- Répondre aux questions (#1-6)
- Le total des points disponibles pour ce travail est de XX points, ce qui représente XX% de votre note finale.
- Tous les travaux pratiques doivent être réalisés et remis individuellement.

Format:

Les étudiants remettront les questions avec les réponses complètes.

Les tableaux et les figures doivent être produits selon les normes du Département de géographie, environnement et géomatique de l'Université d'Ottawa.

Pour des directives sur le formatage et la présentation appropriés des tableaux et des figures, veuillez consulter la **section d'introduction de ce manuel de laboratoire.**

Section 1: Conception cartographique

La conception cartographique est un processus itératif au cours duquel un cartographe conceptualise et génère des cartes en fonction des besoins du public cible (Slocum, 2009). Les principales étapes du processus cartographique sont les suivantes:

1. Déterminer l'objectif de la carte,
2. Collecter des données qui peuvent être utilisées pour illustrer le message de la carte,
3. Concevoir et réaliser la mise en page de la carte,
4. Déterminer si la mise en page de la carte permet de présenter efficacement l'objectif de la carte et d'améliorer la qualité de la carte.

Les mises en page cartographiques créées par les cartographes sont par nature percutantes et permettent d'expliquer clairement un phénomène géographique à l'aide de cartes et d'éléments cartographiques.

L'une des principales forces qui guident la prise de décision en matière de cartographie est le concept de **hiérarchie visuelle**. La hiérarchie visuelle est le principe qui consiste d'organiser les éléments de manière à mettre en valeur les caractéristiques les plus importantes de la carte en ajustant le contraste visuel, la taille, l'emplacement et l'équilibre des éléments de la carte. Pour évaluer une carte, il est plus facile de se demander "vers quoi mes yeux sont attirés". La figure 1 montre une carte (à gauche) et une version avec un classement de l'importance visuelle attribuée aux éléments de la carte (à droite).

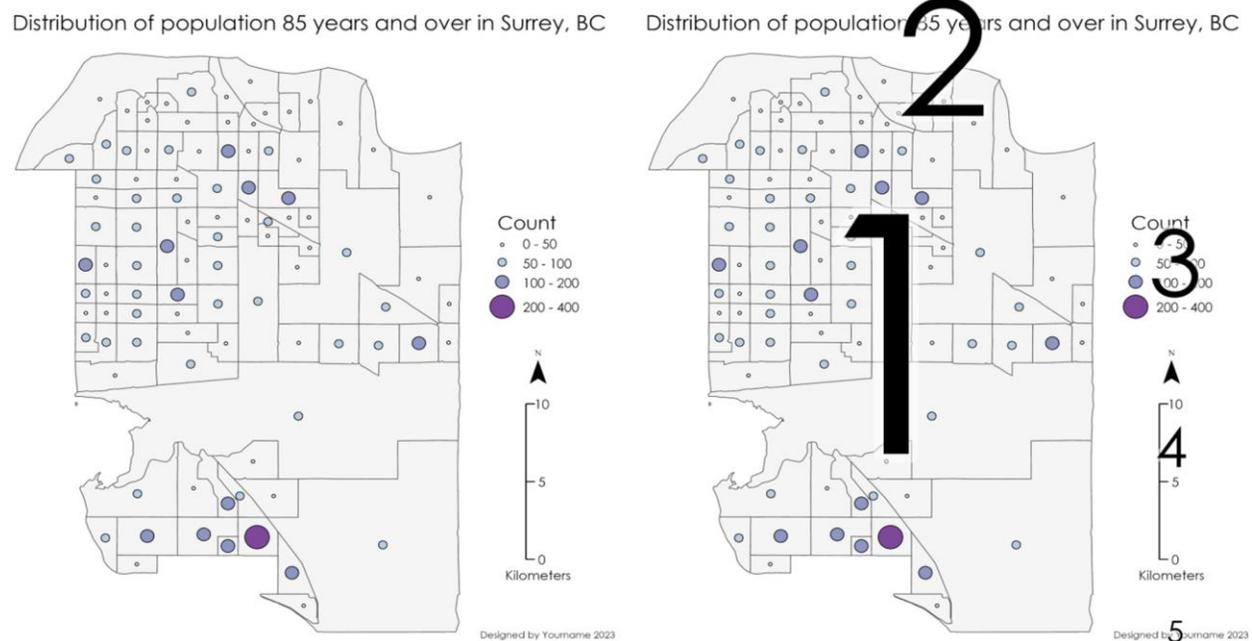


Figure 1. Deux cartes de Surrey, en Colombie-Britannique, ont été créées pour illustrer la hiérarchie visuelle (de 1 à 5). N'oubliez jamais que les données géographiques sont l'objet principal et qu'elles doivent occuper la plus grande partie de l'espace sur une carte.

La carte de Surrey, en Colombie-Britannique, occupe la majeure partie de la surface de cette mise en page et constitue la première chose qui attire l'attention de l'observateur.

L'emplacement et la taille de police relativement grande du titre en font le deuxième élément le

Nom de l'étudiant: _____ # d'étudiant: _____

plus important visuellement dans cette mise en page. La taille de la légende est légèrement supérieure à celle de la flèche du nord et de la barre d'échelle, tandis que le nom et la date, situés dans le coin inférieur droit, constituent l'élément le moins important sur le plan visuel. Une hiérarchie visuelle appropriée conduit le lecteur de la carte au titre, puis aux autres éléments de la carte. La meilleure façon d'y parvenir est d'utiliser efficacement la taille et le contraste visuel.

Avant de créer des cartes pour l'exercice et le devoir, tenez compte des conseils suivants:

1. Faites un croquis avant de commencer la mise en page de la carte.
2. Tenez compte de l'importance relative de chaque élément de la carte lorsque vous décidez de la mise en forme.
3. Remplir efficacement les espaces vides avec des éléments cartographiques.
4. Exporter une image provisoire de votre carte et examiner la hors d'ArcGIS ; cela peut vous aider à remarquer les problèmes de mise en forme qui doivent être résolus.
5. Il existe de nombreuses sources d'aide en dehors de ce cours (ArcGIS Help, Google) qui peuvent vous aider à résoudre les problèmes cartographiques que vous pourriez rencontrer.

De plus, plus un objet ou une partie d'objet est foncé, plus il aura un poids visuel important. Il convient donc d'être prudent et de tenir compte de ce facteur lors de l'utilisation d'entourages de cartes.

Question 1: Quel est l'objectif des cartes créées dans ce laboratoire et comment puis-je accéder aux données?

L'objectif des cartes créées dans ce laboratoire est de s'entraîner à créer des mises en page de cartes qui respectent les meilleurs principes cartographiques et affichent des informations vectorielles sur la ville de Surrey, en Colombie-Britannique. Les données utilisées dans ce laboratoire sont accessibles sur le site de la ville de Surrey Open Data. <https://data.surrey.ca/>

- Les jeux de données de ce site Web qui sont nécessaires pour ce laboratoire sont les suivants:
- Sites patrimoniaux <https://data.surrey.ca/dataset/heritage-sites>
- Arbres importants <https://data.surrey.ca/dataset/important-trees>
- Pistes cyclables <https://data.surrey.ca/dataset/bike-routes>
- Sites d'écosystèmes <https://data.surrey.ca/dataset/ecosystem-sites>
- Recensement 2016 du Surrey (« *2016 Surrey Census* »)
 - Télécharger : Niveau du secteur de recensement (« *Census Tract Level* ») et niveau de l'aire de diffusion (« *Dissemination Area Level* »).

Sur le site web Surrey Open Data, visitez la page web et téléchargez chaque fichier au format géodatabase.

Nom de l'étudiant: _____ # d'étudiant: _____

Data and Resources

Important Trees
Available in CSV format

Important Trees
Available in JSON format

Important Trees
Available in KML format

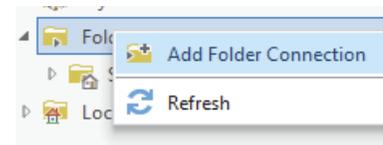
Important Trees
Available in File Geodatabase format

Important Trees
Specification at: <http://resources.arcgis.com/en/help/rest/apiref/>

Une fois que vous avez téléchargé toutes les données (ou téléchargé le fichier de données à partir du référentiel du laboratoire), décompressez le fichier zip et copiez tous les dossiers qui se terminent par .gdb dans un seul dossier nommé "Lab Input Data" (données de laboratoire). Assurez-vous que le dossier "Lab Input Data" est enregistré dans le même dossier que votre travail, afin qu'il soit facile à trouver.

Name	Date modified	Type
bike_routes.gdb	2023-03-13 2:17 PM	File folder
Census_2016_Surrey_DisseminationAreas_UTM.gdb	2018-03-28 11:18 AM	File folder
Census2016_Surrey_CensusTract_UTM.gdb	2018-03-28 11:18 AM	File folder
ecosystem_sites.gdb	2023-03-13 2:17 PM	File folder
heritage_sites.gdb	2023-03-13 2:17 PM	File folder
important_trees.gdb	2023-03-13 2:15 PM	File folder

Pour commencer à travailler sur ce laboratoire, créez un nouveau projet dans ArcGIS Pro. La première chose à faire est d'ouvrir le panneau du catalogue où vous pouvez accéder à vos données. Dans la barre de menu du haut, cliquez sur « view », puis sur « catalog pane ». Un nouveau menu appelé « Catalog » devrait apparaître sur le côté gauche de votre écran. Cliquez sur l'icône "dossiers" dans votre catalogue et vous verrez un dossier d'accueil existant avec une géodatabase à l'intérieur. Par défaut, tous les résultats des outils ArcGIS Pro que vous exécutez seront enregistrés dans cette géodatabase, c'est pourquoi elle est appelée « Default Geodatabase » (Géodatabase par défaut).



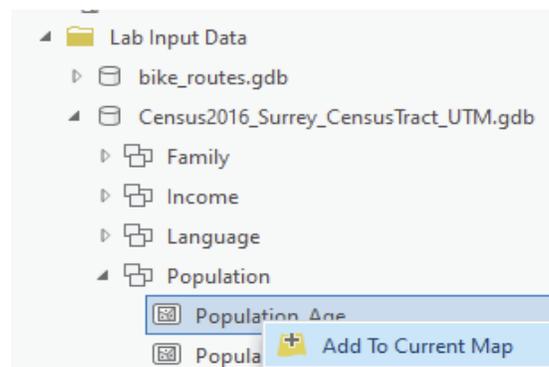
Cliquez avec le bouton droit de la souris sur l'icône "dossiers" dans le menu du catalogue pour créer une connexion au dossier *Lab Input Data*. Une fois que vous avez sélectionné le dossier contenant vos données (*Lab Input Data*), cliquez sur « Ok » et il sera visible dans la section des dossiers du menu du catalogue. Si vous rencontrez des problèmes avec les connexions de dossiers, veuillez-vous référer à ce site web <https://pro.arcgis.com/fr/pro-app/latest/help/projects/connect-to-a-folder.htm>

Conseil : Si vous voyez votre dossier dans l'explorateur Windows, vous pouvez glisser-déposer le dossier dans le volet catalogue au lieu d'utiliser la fonction « Add Folder Connection » (ajouter une connexion de dossier) pour créer une connexion de données.

Nom de l'étudiant: _____ # d'étudiant: _____

Question 2: Comment créer une carte et symboliser une classe d'entités?

Pour les questions 2 et 3, nous allons créer une carte, appliquer une symbologie de base, puis créer une mise en page de carte. Dans le dossier *Lab Input Data* de votre panneau de catalogue, ouvrez la géodatabase appelée **Census2016_Surrey_CensusTract_UTM.gdb**, puis ouvrez le jeu de données sur la population, cliquez avec le bouton droit de la souris sur **Population_Age** et cliquez sur « *Add to Current map* » (Ajouter à la carte actuelle). La carte doit être zoomée sur l'étendue de la classe d'entités récemment ajoutée. Dans le volet de contenu à droite de la carte, sous « *Drawing Order* » (ordre de dessin), décochez les cartes de base **World Topographic Map** et **World hillshade**. Cliquez ensuite avec le bouton droit de la souris sur **Population_Age** et cliquez sur « *Symbology* ». L'onglet Symbologie devrait apparaître à droite de votre carte. Pour modifier la façon dont **Population_Age** est affiché sur votre carte, sous « *Primary Symbology* » (symbologie primaire), cliquez sur le menu déroulant, qui devrait être sur « *Single Symbol* » (symbole unique) par défaut, pour sélectionner la symbologie que vous souhaitez utiliser pour cette classe d'éléments. Si vous rencontrez des problèmes avec la symbolisation d'une classe d'entités, veuillez-vous référer à ce site web: <https://pro.arcgis.com/fr/pro-app/latest/help/mapping/layer-properties/symbolize-feature-layers.htm>



Question 3: Quelles sont les façons courantes de symboliser une classe d'entités polygonales?

La première question que vous devez vous poser lorsque vous symbolisez des données dans ArcGIS Pro est la suivante : « Quel type de données suis-je en train de symboliser ? » Les données catégorielles (noms et lieux) doivent être symbolisées à l'aide de valeurs uniques, une couleur unique étant attribuée à chaque catégorie unique. Pour les données quantitatives (population et pourcentages), il existe plusieurs façons de symboliser les données, telles que Choroplèthe (couleurs graduées), symboles gradués, densité de points et symboles proportionnels (Figure 2).

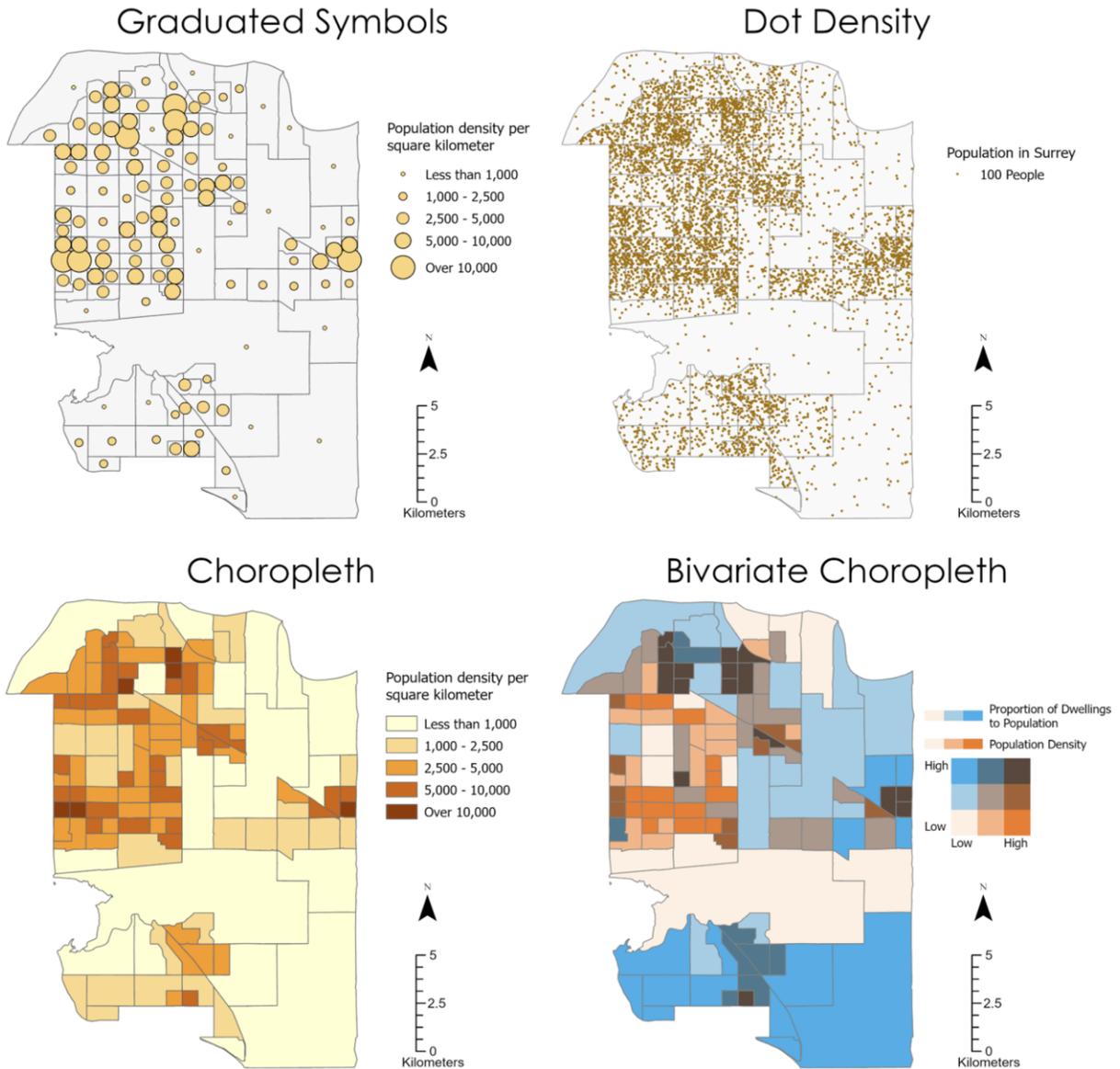


Figure 2. Comparaison de différents types de symbolisation quantitative à Surrey, en Colombie-Britannique.

La méthode la plus répandue pour symboliser une classe d'entités polygonales consiste à utiliser une symbologie choroplèthe (couleurs graduées). Même si les cartes choroplèthes sont faciles à réaliser et à interpréter, elles présentent également plusieurs faiblesses dont il faut être conscient:

1. Par définition, une carte choroplèthe suppose que les données sont réparties uniformément dans chaque polygone, ce qui est rarement le cas. Par exemple, dans les cartes de population présentées ci-dessus, il serait déraisonnable de supposer que la population est uniformément répartie dans chaque polygone ; au contraire, les gens ont tendance à vivre dans des quartiers résidentiels, tandis que d'autres zones à l'intérieur de chaque polygone sont utilisées principalement pour le transport, les achats, les loisirs, etc. et ont donc une densité de population plus faible. Examiner la figure 3:

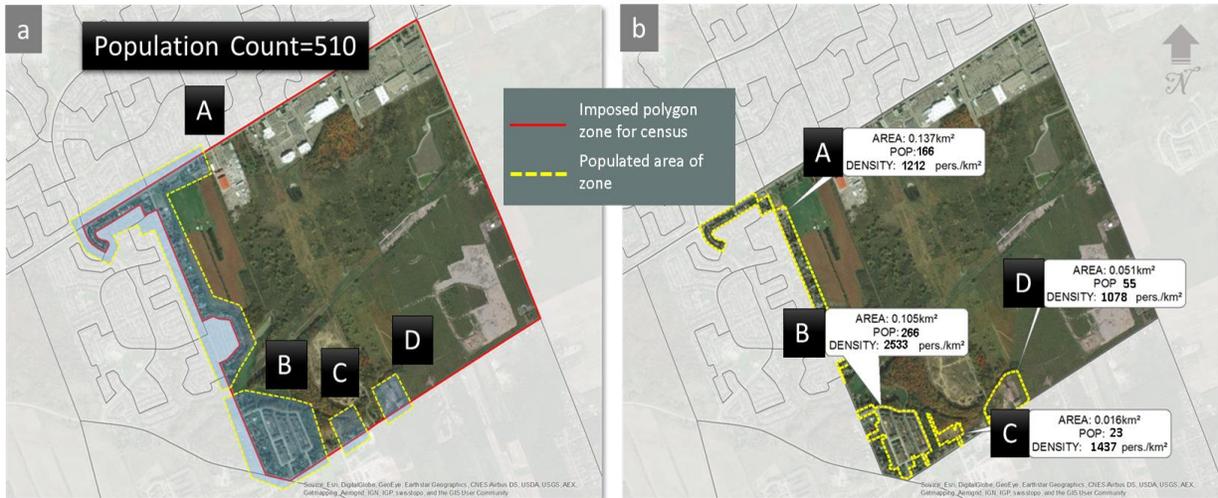


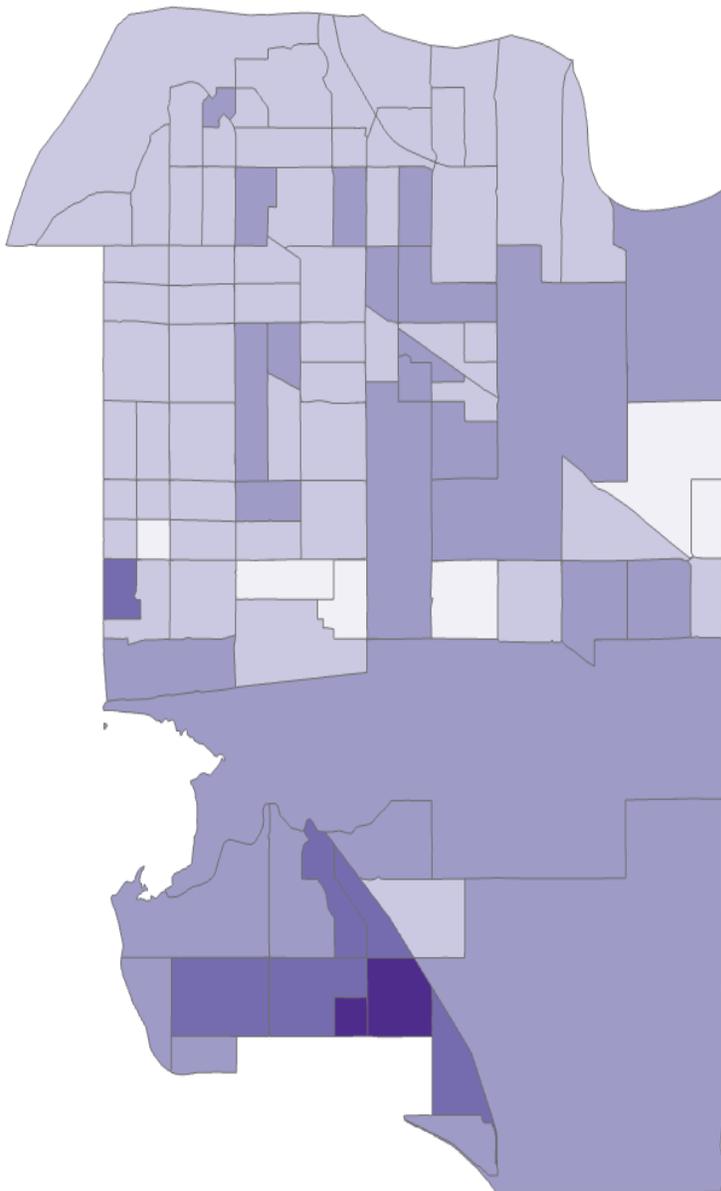
Figure 3: a) Une seule zone de diffusion (contour rouge) avec une population totale de 510 personnes. La partie peuplée de la zone est représentée en jaune et étiquetée A-D. Le reste de la zone est constitué de zones commerciales, agricoles ou industrielles où la population ne peut pas résider. b) Population de 510 personnes répartie entre les parties peuplées de la zone, proportionnellement à la superficie de chaque partie peuplée. Remarquez que l'hypothèse d'homogénéité à l'intérieur des parties peuplées de la sous-zone est toujours vraie puisque la densité de population reste constante.

Dans le monde réel, ce polygone entier est attribué à 510 personnes (figure 3a), alors qu'en réalité, les gens n'habitent que les zones entourées par les lignes jaunes pointillées. De plus, ces petites zones ne sont même pas reliées géographiquement les unes aux autres, et chacune d'entre elles a son propre nombre d'habitants et sa propre densité. Vous vous demandez maintenant où s'arrêter. Si nous faisons un zoom sur l'une des zones jaunes, nous verrions également d'autres zones où les gens ne peuvent pas vivre et nous pourrions les exclure du calcul de la densité de population afin d'obtenir des densités de population encore plus précises. Devrions-nous alors exclure les garages ou les places de stationnement puisque les gens ne peuvent pas y vivre ? Qu'en est-il des arrière-cours ? Les cours avant ? On pourrait continuer ainsi à l'infini, *ad infinitum, ad nauseam*. Au final, aucune carte ne correspondra à la réalité parce que la réalité est infiniment complexe. Ici, nous pouvons simplement comprendre que nous pouvons mieux approcher la réalité si nous comprenons qu'aucune quantité n'est naturellement distribuée uniformément dans l'espace, à moins que - vous en connaissez une ?

2. Pour produire des comparaisons significatives entre des zones dans un choroplèthe, les données qui vont être cartographiées doivent être normalisées, par exemple en cartographiant la densité de la population au lieu de la population totale dans chaque zone. Ceci peut être réalisé en affichant les données normalisées à l'aide du menu déroulant de normalisation dans le panneau de symbologie. Les symboles gradués, les symboles proportionnels et les cartes de densité de points ne nécessitent pas de données normalisées et peuvent être utilisés pour représenter des données de comptage telles que la population.

Nom de l'étudiant: _____ # d'étudiant: _____

Pour votre première carte, sélectionnez la symbologie primaire « *Graduated Colors* », le champ (« *Field* ») « *Average age* », la normalisation "<None>", la méthode « *Manual Interval* » et le nombre de classes "5". Définissez les valeurs supérieures (« *Upper value* ») à 35, 40, 45, 50 et 55 et mettez à jour les étiquettes (« *Label* »). Utilisez cette carte pour pratiquer les éléments cartographiques des questions 5, 6 et 7.



Primary symbology

Graduated Colors

Field: Average age

Normalization: <None>

Method: Manual Interval

Classes: 5

Color scheme: [Color gradient]

Classes Histogram Scales

Symbol	Upper value	Label
[Lightest purple]	≤ 35	30 - 35
[Light purple]	≤ 40	35 - 40
[Medium-light purple]	≤ 45	40 - 45
[Medium purple]	≤ 50	45 - 50
[Darkest purple]	≤ 55	50 - 55

Question 4: Que sont les intervalles et comment affectent-elles l'apparence de ma carte?

Les intervalles de classe, appelées « *Method* » (méthode) dans l'onglet de symbologie, vous permettent de visualiser des plages quantitatives de données. Les méthodes d'intervalles de classe comprennent les Seuils naturels (Jenks) (« *Natural Breaks (Jenks)* »), l'intervalle égal (« *Equal Interval* »), l'intervalle manuel (« *Manual Interval* »), l'écart type (« *Standard Deviation* »), le quantile et l'intervalle géométrique (« *Geometrical Interval* »). La carte ci-dessous (figure 4) présente quatre des méthodes d'intervalle de classe les plus courantes. De plus amples informations sur les méthodes d'intervalle de classe peuvent être trouvées ici :

<https://pro.arcgis.com/fr/pro-app/latest/help/mapping/layer-properties/data-classification-methods.htm>

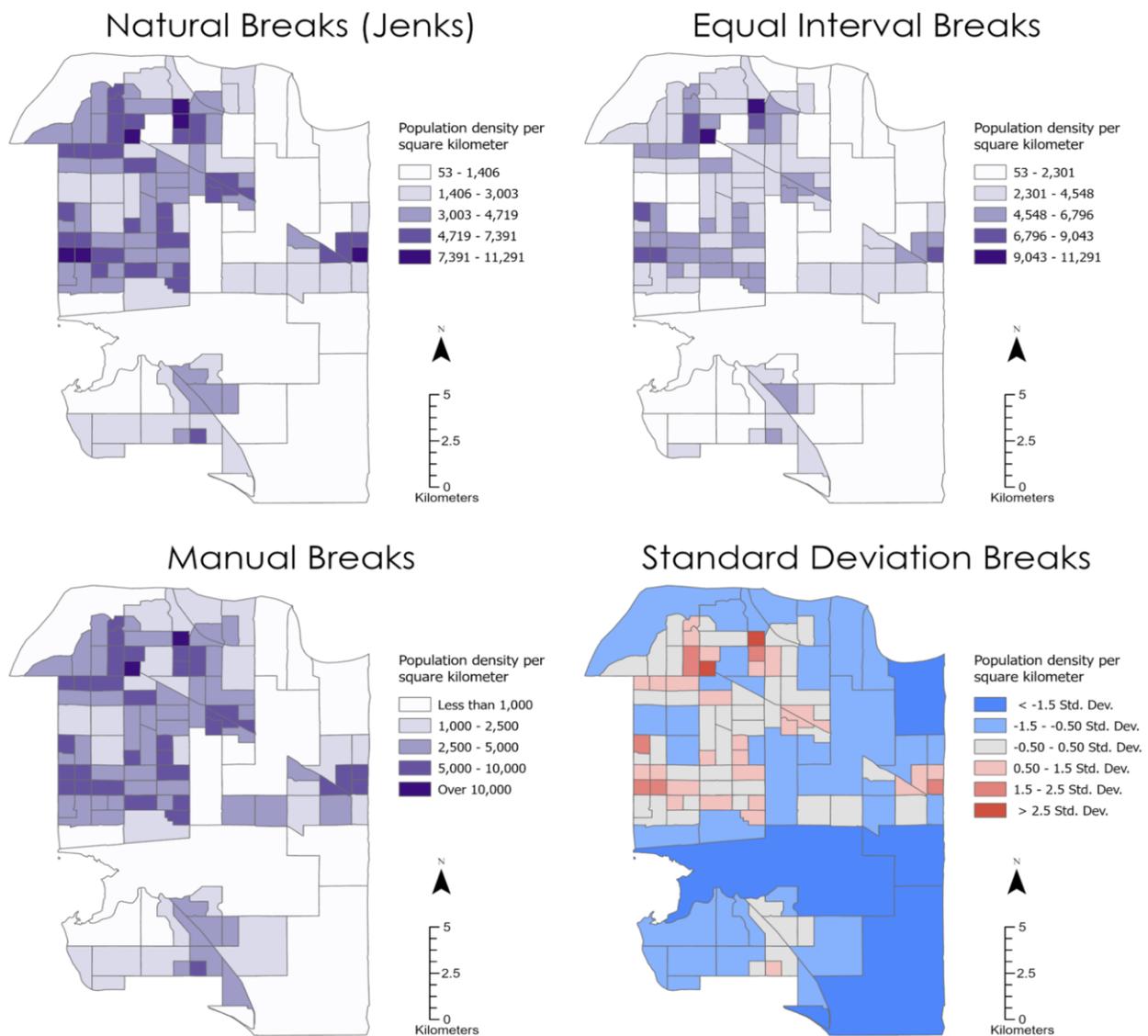
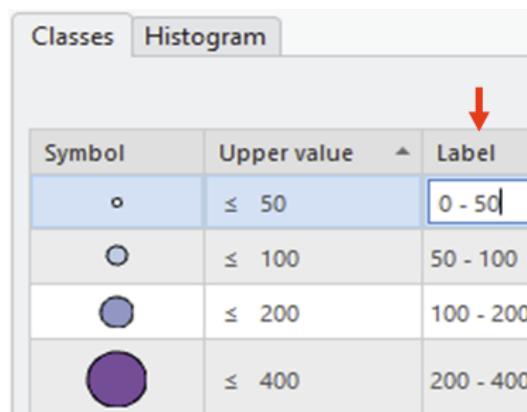


Figure 4. Les méthodes les plus courantes utilisées pour les intervalles de classe visualisées à l'aide des données démographiques de Surrey, BC.

Nom de l'étudiant: _____ # d'étudiant: _____

La détermination de la méthode de classification appropriée est une décision cartographique qui peut avoir un impact sur le message que la carte transmet. Les intervalles de classe déterminent également en partie le type d'échelle de couleurs qui peut être utilisé et qui, à son tour, dépend en fin de compte du type de données cartographiées. Une règle de base est que les intervalles manuels doivent être utilisés la plupart du temps si vous utilisez des valeurs numériques pour décrire vos données, telles que la densité de population par kilomètre carré. Il est beaucoup plus facile pour les lecteurs de comprendre et de visualiser des nombres tels que 1 000 habitants par kilomètre carré, plutôt que 1 461. Si vous décrivez vos données dans des catégories telles que faible, moyenne et élevée, des intervalles naturels ou des intervalles égaux peuvent être plus appropriés. Vous pouvez également ajuster ce qui est affiché sur la carte à côté de chaque intervalle en modifiant la colonne d'étiquettes dans le panneau de symbologie et en changeant la palette de couleurs, soit en modifiant la couleur de chaque intervalle, soit en sélectionnant une palette de couleurs dans le menu déroulant.



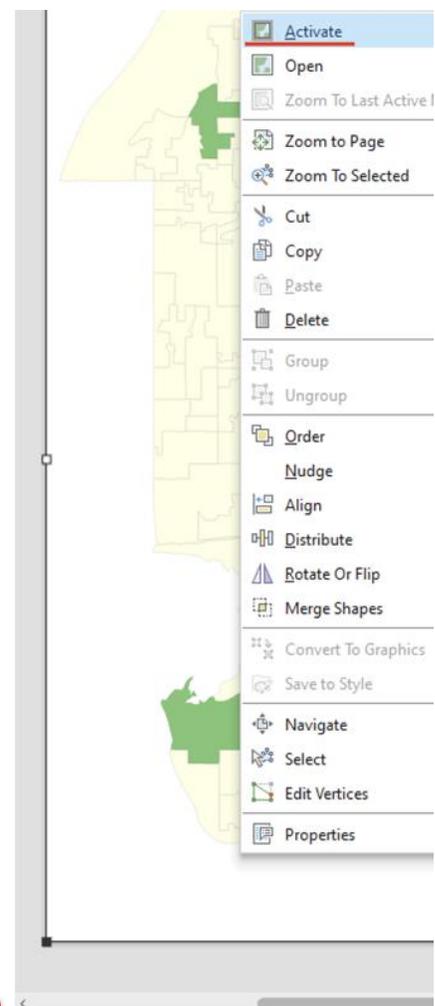
Symbol	Upper value	Label
○	≤ 50	0 - 50
○	≤ 100	50 - 100
○	≤ 200	100 - 200
○	≤ 400	200 - 400

Question 5: Comment créer une mise en page de carte et définir l'échelle de la carte?

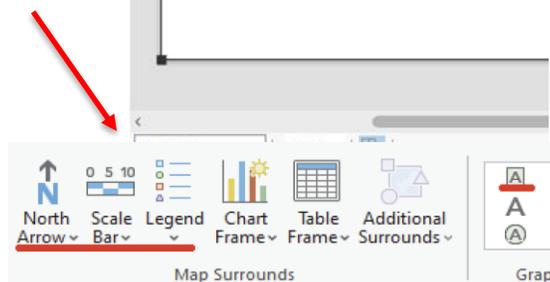
La première étape de la création d'une carte consiste à visualiser l'emplacement de la carte et des éléments cartographiques. Il est recommandé de faire un croquis rapide sur une feuille de papier ou de passer un peu de temps à réfléchir à l'emplacement de la carte et des éléments cartographiques (titre, légende, flèche du nord, barre d'échelle). L'étape suivante consiste à décider de la taille et de l'orientation de la page en fonction de votre croquis cartographique. Si votre carte occupe une page entière, il est recommandé de choisir une page de format lettre. Toutefois, si vous créez une carte pour un usage numérique ou si vous produisez une carte qui sera intégrée à un rapport, il est recommandé de définir un format de page personnalisé. Le format carré (20 cm x 20 cm) est un format de page courant qui permet de diviser facilement la page pour réaliser des mises en page cartographiques. Les mises en page les plus efficaces sont créées lorsque les cartographes examinent la carte qu'ils essaient de réaliser ainsi que les éléments de la carte, puis choisissent une taille de page appropriée. Par exemple, puisque Surrey, BC a plus de hauteur que de largeur, votre choix de taille de page doit refléter ce fait. Notez que vous pouvez toujours modifier la taille de votre page après avoir commencé à travailler sur votre mise en page cartographique.

Nom de l'étudiant: _____ # d'étudiant: _____

Pour créer une mise en page, il suffit de cliquer sur le bouton « *New Layout* » (Nouvelle mise en page) dans l'onglet « *Insert* » (Insertion) du groupe « *Project* » (Projet). Après avoir commencé une mise en page dans ArcGIS Pro avec une page vierge, cliquez sur « *Insert* », puis sur « *Map frame* » et sélectionnez la carte que vous souhaitez insérer dans votre mise en page. Considérez le cadre de la carte comme le contour de l'emplacement de votre carte sur la page. Dessinez un cadre de carte, puis cliquez dessus avec le bouton droit de la souris et cliquez sur « *Activate* » (activer). Une fois le cadre de carte activé, vous pouvez déplacer la carte autour du cadre et effectuer un zoom avant. L'échelle de la carte (soulignée par une ligne rouge en bas à gauche de l'image à droite) peut être réglée manuellement pour augmenter ou diminuer la taille de la carte. Une fois que vous avez placé votre carte de manière appropriée, cliquez sur le bouton « *Layout* » (mise en page) dans la barre de menu supérieure et cliquez sur « *Close Activation* » (fermer l'activation). Cela empêchera toute modification subséquente de votre carte à l'intérieur du cadre de la carte. Vous trouverez ici des conseils et des ressources supplémentaires sur la mise en place d'un cadre de carte et sur la façon de supprimer la bordure du cadre de carte : <https://pro.arcgis.com/fr/pro-app/latest/help/layouts/add-and-modify-map-frames.htm>



Après avoir placé votre carte, allez dans le ruban du menu supérieur, trouvez le groupe « *Map Surrounds* » et sélectionnez une Flèche du Nord, une barre d'échelle et une légende appropriée. Dans le sous-menu « *Graphics and text* » (Graphiques et texte) du ruban du menu supérieur, sélectionnez « rectangle » et dessinez un cadre pour le titre de la carte. Il est préférable de



dessiner un texte en rectangle plutôt qu'une zone de texte linéaire, car cela vous offre davantage d'options si vous souhaitez ajouter un arrière-plan à votre zone de texte. Gardez à l'esprit qu'il ne faut pas mettre de titre sur une carte qui n'est pas une carte autonome, c'est-à-dire que si vous avez une question qui ne demande qu'une carte, la carte doit avoir un titre, si vous avez une question qui demande une carte et l'interprétation de certaines cartes, vous devez omettre le titre et utiliser à la place une légende de figure numérotée. Cette procédure est expliquée dans l'avant-propos avec des exemples.

Des ressources supplémentaires pour les mises en page de cartes sont disponibles sur cette page et dans les sections suivantes de l'aide ArcGIS: <https://pro.arcgis.com/fr/pro-app/latest/help/layouts/layouts-in-arcgis-pro.htm>

Question 6: Comment concevoir des éléments cartographiques esthétiques?

Cette section présente brièvement quelques conseils pour améliorer les éléments de votre carte et leur donner un air plus professionnel.

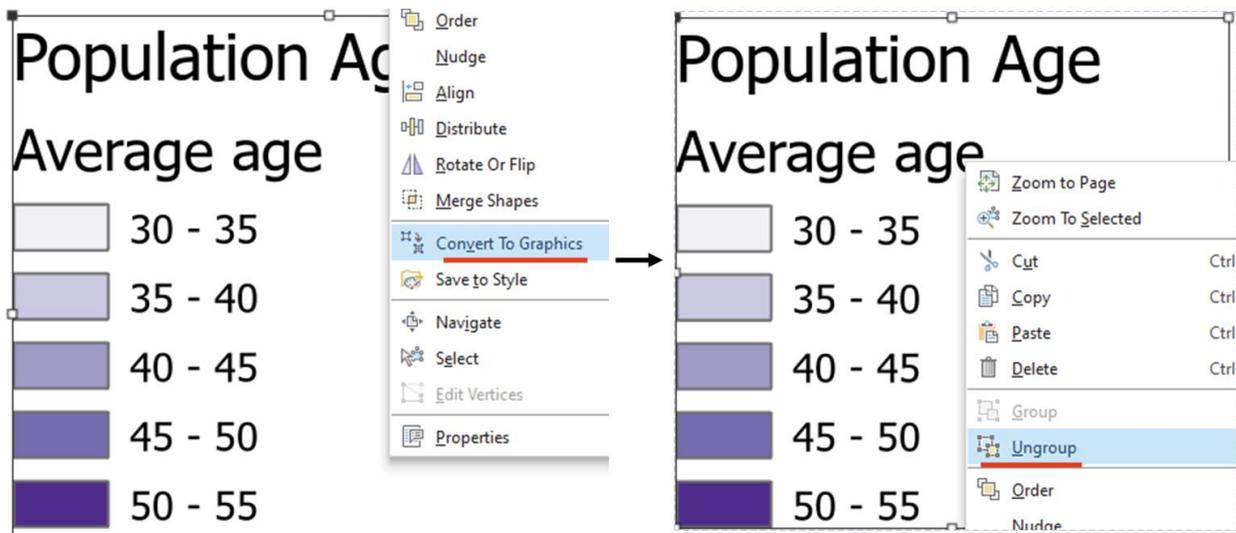
1. Légende

Il est recommandé d'utiliser les propriétés de la légende pour modifier ses caractéristiques. Tout comme l'ordre de dessin d'une carte, vous pouvez activer et désactiver les couches de la légende pour décider des symboles que vous souhaitez y voir figurer. Consultez l'aide d'Esri pour plus de conseils sur la mise en forme des légendes à l'aide des propriétés de légende:

<https://pro.arcgis.com/fr/pro-app/latest/help/layouts/work-with-a-legend.htm>

Il faut noter qu'il est parfois beaucoup plus facile de travailler sur une légende lorsqu'il s'agit d'un graphique dans ArcGIS Pro. Cette opération ne doit être effectuée que si vous êtes fixé sur la palette de couleurs et que vous n'avez pas l'intention d'apporter d'autres modifications au contenu de la légende (comme l'ajout d'une autre couche).

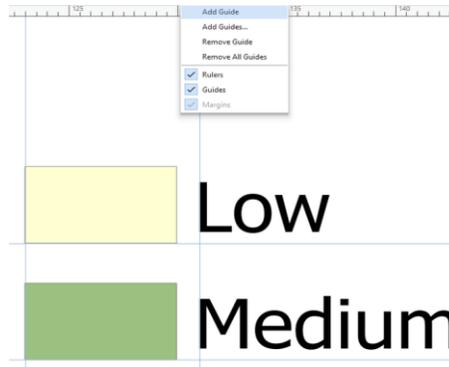
Si vous prévoyez créer plusieurs cartes et d'utiliser la mise en page comme modèle, il est recommandé de mettre à jour la légende dans les propriétés de la légende. Sinon, lorsque la légende est sous forme de graphique, il est plus facile de la peaufiner et de lui donner un aspect plus professionnel. Pour convertir une légende en graphique, cliquez avec le bouton droit de la souris sur la légende, puis cliquez sur « *Convert To Graphics* » (Convertir en graphique). Ensuite, si vous cliquez avec le bouton droit de la souris et que vous cliquez sur « *ungroup* » (dégrouper), vous pouvez facilement déplacer les différentes parties de la légende, telles que les symboles et le texte (vous allez peut-être devoir dégroupier le graphique deux fois en fonction de la complexité de votre légende).



Vous pouvez utiliser des guides pour aligner les éléments de la carte, les symboles et le texte de la légende afin d'améliorer la cohérence de l'espace et de la mise en page. Une fois

Nom de l'étudiant: _____ # d'étudiant: _____

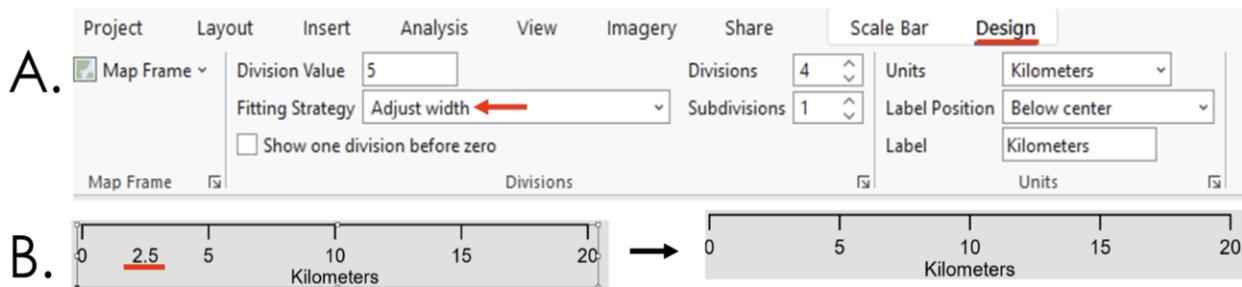
que vous avez modifié les éléments de la légende, sélectionnez l'ensemble du texte et des symboles, puis cliquez sur le bouton droit de la souris et cliquez sur « *Group* » pour les grouper afin de pouvoir à nouveau déplacer la légende sur la carte en tant qu'élément unique.



L'image ci-dessus montre comment les lignes directrices peuvent être utilisées pour formater une légende, avec un espacement cohérent entre le texte et le symbole.

2. Barre d'échelle

La barre d'échelle par défaut n'est généralement pas la plus facile à interpréter, il faut donc s'attendre à devoir modifier la barre d'échelle sur la plupart de vos cartes. Une fois que vous avez placé une barre d'échelle, cliquez dessus une fois et le menu de configuration devrait apparaître dans le ruban du menu supérieur. Ouvrez le menu de configuration pour trouver des outils utiles pour ajuster votre barre d'échelle. La stratégie d'ajustement est cruciale car elle détermine la manière dont la barre d'échelle est ajustée - en général, « *Adjust width* » (Ajuster la largeur) est le format le plus facile à mettre en forme. Modifiez ensuite la valeur de la division, les divisions et les subdivisions pour modifier l'affichage de la légende. Assurez-vous que vous utilisez les unités de distance appropriées ; le système métrique est la règle, sauf indication contraire. Si vous souhaitez affiner votre barre d'échelle, vous pouvez la convertir en graphique (similaire à la légende). Vous pouvez ainsi supprimer les valeurs de division inutiles ou inverser facilement la barre d'échelle à la verticale pour les mises en page en mode portrait. Attention: *une fois la barre d'échelle convertie en graphique, la modification de la taille du graphique entraîne une distorsion de l'échelle ; il ne faut jamais faire cela.*



3. Flèche du Nord

Les flèches du nord sont relativement simples à implémenter dans ArcGIS Pro. En règle générale, vous souhaitez une simple flèche du nord, dont la taille doit être suffisante pour que

Nom de l'étudiant: _____ # d'étudiant: _____

l'on comprenne de quoi il s'agit, mais qui ne doit pas être trop grande. La flèche du nord est un élément de la carte qui a relativement moins d'importance que d'autres et qui doit être présent sur la carte sans occuper trop d'espace. Il faut la chercher plutôt que de la voir immédiatement.

Question 7: Comment exporter une carte?

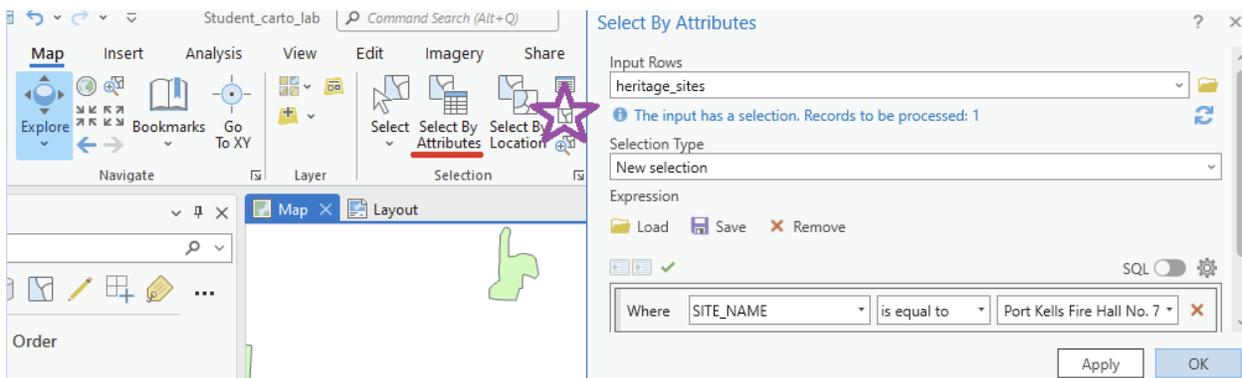
Une fois que vous avez terminé de travailler sur la mise en page de votre carte, cliquez sur « *Share* » (Partager) dans la barre de menu du haut, puis sur « *Export Layout* » (Exporter la mise en page). Vous serez invité à définir un type de fichier. Il est recommandé d'utiliser le format PNG en raison de la taille réduite du fichier. Réglez le DPI à au moins 300 et définissez le nom et le chemin du fichier. Vous pouvez utiliser la fonction « *clip to graphics* » (découper les graphiques) pour supprimer les espaces vides dans votre carte exportée.



Question 8: Comment étiqueter les éléments de la carte?

L'étiquetage des entités cartographiques est important pour fournir un contexte et des informations sur les entités que vous cartographiez. Cette section explique comment étiqueter les cartes, comment placer les étiquettes et comment utiliser les formules d'étiquetage pour les entités de points et de polygones.

Pour commencer, placez les sites patrimoniaux, les pistes cyclables et les sites écosystémiques (« **heritage_sites**, **bike_routes**, **ecosystem_sites** ») sur votre carte dans ArcGIS Pro. Ils seront situés dans leurs géodatabases respectives dans le dossier **Lab Input Data**. Dans le menu du haut, sélectionnez votre carte « *Map* », cliquez sur « *Select By Attributes* », utilisez les lignes d'entrée (« *Input Rows* ») "**heritage_sites**" et la clause « *Where SITE_NAME is equal to Port Kells Fire Hall No. 7* » (voir l'image ci-dessous) :

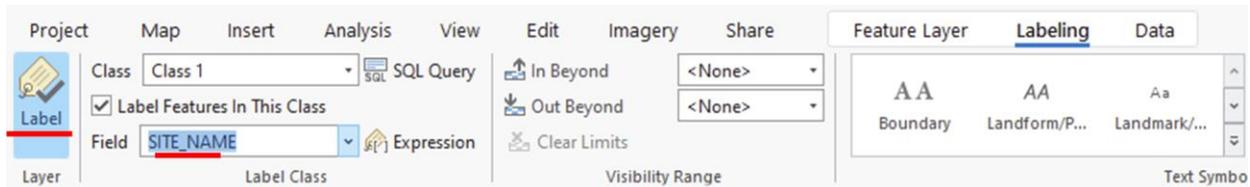


Cliquez sur « *Ok* » et une marque bleu clair apparaît dans le coin nord-ouest de Surrey, BC. Zoomez sur cette région puisque nous allons cartographier la région de Port Kells de Surrey dans cette partie du laboratoire. Ensuite, dans la barre de menu du haut, cliquez sur le bouton

Nom de l'étudiant: _____ # d'étudiant: _____

de désélection (étoile mauve dans l'image ci-dessus) pour désélectionner la caserne de pompiers.

Pour activer les étiquettes pour les sites patrimoniaux, cliquez une fois sur "heritage_sites" dans le volet de contenu (« *Contents* »), puis dans la barre de menu supérieure, cliquez sur « *Labeling* » (Étiquetage) et ensuite sur « *Label* » (Étiquette). Assurez-vous que le champ est défini par SITE_NAME pour afficher les noms des sites patrimoniaux.

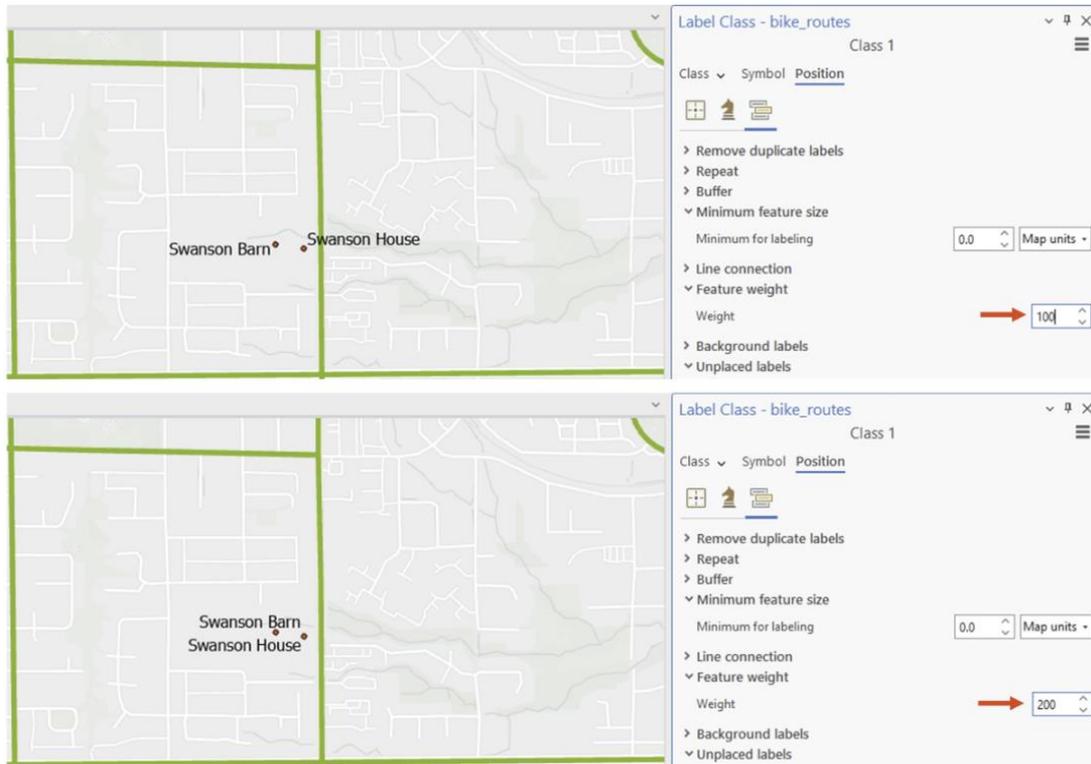


Plus d'informations sur l'étiquetage sont disponibles ici: <https://pro.arcgis.com/fr/pro-app/latest/help/mapping/text/labeling-basics.htm>

Question 9: Comment modifier l'emplacement des étiquettes?

Pour cette question, nous nous concentrerons sur la modification de l'emplacement des étiquettes d'un élément de point. Il existe deux façons principales d'ajuster l'emplacement des étiquettes d'entités de points, en ajustant les poids des entités et en changeant la préférence de placement.

Les propriétés de poids (« *Feature Weight* ») des éléments peuvent indiquer l'importance relative de certains éléments de la carte par rapport à d'autres. Vous pouvez modifier les poids des éléments en cliquant avec le bouton droit de la souris sur un élément de la carte dans le volet de contenu (« *Contents* ») et en cliquant sur « *Labeling properties* » (Propriétés d'étiquetage). Dans le nouveau volet des classes d'étiquettes, cliquez sur « *Position* », ensuite « *Feature Weight* », et vous pourrez modifier l'importance de l'élément. Dans l'exemple ci-dessous, le poids de l'élément des pistes cyclables a été augmenté afin d'éviter que les noms de patrimoine ne se chevauchent avec les pistes cyclables. Pour plus d'informations sur les poids des éléments, voir le site suivant: <https://pro.arcgis.com/en/pro-app/latest/help/mapping/text/prevent-labels-from-overlapping-certain-features.htm>



Une autre méthode de placement des étiquettes, qui peut être utilisée indépendamment ou avec les poids des éléments, consiste à définir la préférence de positionnement dans les propriétés de l'étiquette (« *Label Properties* »). Dans la figure 5a, les deux étiquettes sont un peu trompeuses - il n'est pas facile de voir quel est le point Swanson Barn et quel est le point Swanson House. En changeant la position des étiquettes (en inversant les positions 2 et 4), les étiquettes peuvent être placées plus clairement dans la figure 5b.



Figure 5: Options d'étiquetage des cartes dans ArcGIS Pro

Si les caractéristiques sont très complexes et que vous terminez la mise en page d'une carte, il est parfois préférable de modifier manuellement la position des étiquettes. Dans ce cas, l'outil

Nom de l'étudiant: _____ # d'étudiant: _____

de conversion des étiquettes en graphiques (« *Convert Labels to Graphics* ») peut être utilisé.
<https://pro.arcgis.com/fr/pro-app/latest/tool-reference/cartography/convert-labels-to-graphics.htm>

Section 2: Problème d'agrégation spatiale

Question 10: Qu'est-ce que le Problème d'agrégation spatiale?

Le problème d'agrégation spatiale (en anglais, « *modifiable areal unit problem (MAUP)* ») décrit un effet par lequel les données rapportées dans des polygones spatiaux affectent l'interprétation cartographique des modèles ainsi que les résultats analytiques des données collectées à l'intérieur de ces mêmes polygones (Fotheringham et al., 1991). Un exemple extrême de MAUP est le gerrymandering aux États-Unis, qui consiste de modifier les limites spécifiques d'un électorat afin d'altérer les résultats d'une élection. La plupart des cas de MAUP sont de nature plus subtile, mais il est néanmoins important de prendre en compte la manière dont vos données spatiales polygonales sont définies (par exemple, la manière dont les polygones sont constitués ou dont les limites sont déterminées, appelée zonage et échelle (taille), et la manière dont les changements dans la géométrie des polygones peuvent affecter les résultats de toute analyse spatiale, y compris la manière dont les données apparaissent sur une carte. Le MAUP est présent chaque fois que vous analysez des données liées à des entités polygonales ou à des images (raster). Lorsque vous cartographiez quelque chose dans une région et que vous utilisez des polygones, ceux-ci divisent le plus souvent la région entière en un ensemble de polygones qui ne se chevauchent pas et que l'on appelle une tessellation parce qu'ils remplissent tout l'espace.

Prenons l'exemple de la localisation d'une population dans l'espace, par exemple des personnes pique-niquant dans un champ plat, sans arbres et avec de l'herbe courte (un champ homogène) (figure 6).



Figure 6. Personnes dans le champ de Krüöppen le jour de la fête nationale ; à gauche : photos de pique-niqueurs ; à droite : Quadrant carré superposé à une partie du champ de pique-nique. Image créée à l'aide de DALL-E et d'Adobe Photoshop.

Nous superposons ensuite un quadrant (un grand carré) à une partie aléatoire de cette aire de pique-nique (figure 6). Ensuite, nous divisons ce quadrant en 10 polygones et comptons le nombre de personnes à l'intérieur de chacun de ces 10 polygones. Lorsque nous comptons des points dans des polygones, cela s'appelle une opération de point dans un polygone (« *point-in-polygon operation* »). Ensuite, nous créons un autre ensemble de 10 polygones de tailles (échelles) et de formes (zonations) différentes et répétons la même opération de point dans un polygone. Ensuite, pour chacun de nos polygones, nous divisons les nombres de personnes par la surface des polygones pour obtenir le nombre de personnes par polygone. Il s'agit d'une valeur de densité, que nous pouvons représenter sur une carte choroplèthe en variant continuellement la couleur de l'échelle/clé de la carte. Le résultat est illustré par la figure 7.

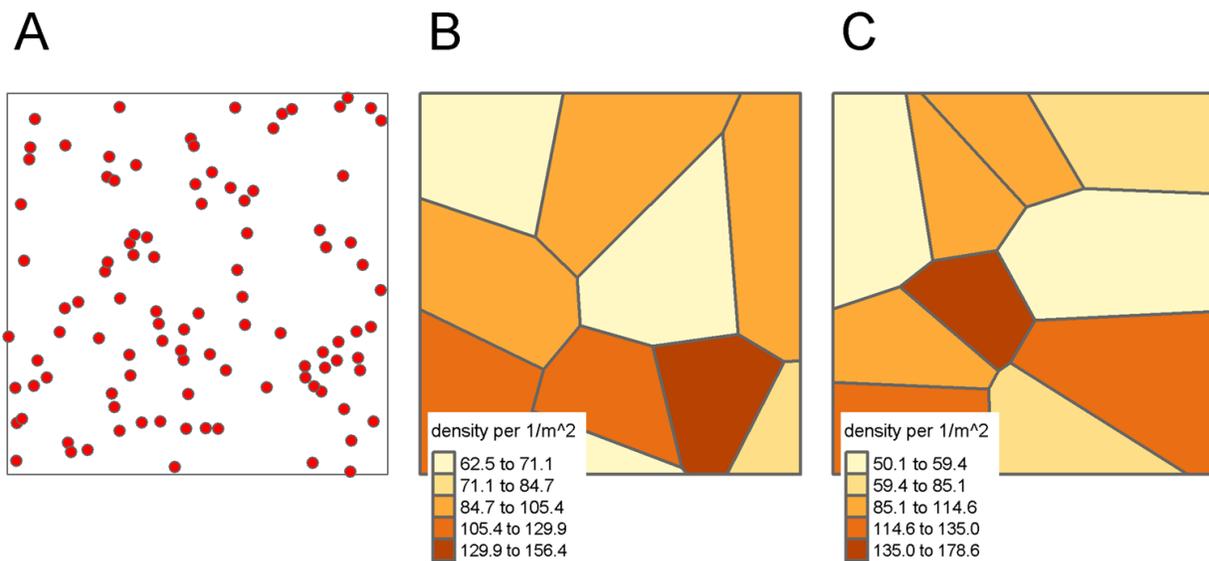


Figure 7: A) la distribution de la population dans une région où chaque point rouge représente une chose (personne, objet, etc.) ; B) une tessellation de la région avec des polygones montrant la densité de la population dans chaque région sur la base des points montrés en A ; C) une tessellation différente de la région avec le même nombre de polygones que dans B mais avec des densités différentes et en utilisant la même population que dans A.

La figure 7 montre clairement que deux tessellations polygonales différentes du même espace fournissent des cartes très différentes de la densité des pique-niqueurs. Si l'on examine uniquement les valeurs des densités minimale et maximale dans les figures 7b et 7c, on constate que les intervalles sont différents. Si l'on vous demandait de décrire l'endroit où la densité de pique-niqueurs est la plus élevée, vous concluriez, d'après la figure 7b, qu'ils se trouvent en bas à droite du champ, alors que si vous utilisiez la figure 7c, vous concluriez qu'ils se trouvent près du centre de la région. Bien que les pique-niqueurs restent immobiles comme le montre la figure 7a, les conclusions tirées sont sensiblement différentes en raison de l'utilisation de deux tessellations ou zonations polygonales distinctes. Ces zonations, tout en contenant la même quantité de polygones (10 sur chaque carte), diffèrent en termes de limites et d'échelle. Cette différence de partition spatiale a donné lieu à diverses valeurs et interprétations. Ces variations sont une conséquence directe du problème d'agrégation spatiale (MAUP).

Nom de l'étudiant: _____ # d'étudiant: _____

Le problème d'agrégation spatiale (MAUP) garantit essentiellement que des résultats distincts apparaissent lorsque la même analyse (comme l'analyse des points dans les polygones décrite dans la figure 7) est effectuée sur un ensemble de données identique, étant donné que l'ensemble de données est représenté par des polygones qui diffèrent uniquement par l'emplacement de leurs limites (zonation) et par leur superficie (échelle). En d'autres termes, l'échelle géographique de l'analyse des données et les limites définies pour la zone peuvent influencer de manière significative le résultat de l'analyse. Les deux composantes du MAUP sont donc les suivantes :

- Problème de zonage: Cela implique l'émergence de résultats différents lors de l'utilisation des mêmes données sous-jacentes lorsque la nature arbitraire des limites utilisées pour résumer ces données est ce que l'on appelle des unités spatiales imposées (Bonham-Carter 1993) - même si la zone/échelle physique de ces unités est maintenue constante, les limites elles-mêmes peuvent être modifiées et, par conséquent, les résultats de toute analyse utilisant les données résumées dans ces polygones peuvent également changer. Ainsi, la façon dont vous définissez les polygones que vous utilisez pour l'analyse influence directement le résultat de toute mesure.
- Problème d'échelle: Il s'agit du fait que les résultats des analyses statistiques seront différents en fonction de l'échelle spatiale de l'analyse. Par exemple, l'analyse des données au niveau d'un quartier (une grande unité) peut donner des résultats différents de l'analyse au niveau des secteurs de recensement individuels (des unités plus petites). Similairement, une carte géologique peut représenter une formation géologique importante, telle qu'un bassin sédimentaire ou une intrusion ignée. Cette formation plus vaste peut englober des unités plus petites, imbriquées les unes dans les autres, de divers types de roches. Si vous entreprenez une analyse, comme l'évaluation de paramètres hydrogéologiques tels que la pression par unité géologique, vous constaterez que les résultats varient en fonction de la taille des unités géologiques que vous examinez. En fait, l'échelle des unités géologiques utilisées dans votre analyse peut influencer de manière significative le résultat. À mesure que l'échelle change (de la petite échelle, comme un pays, à la grande échelle, comme un quartier, ou d'une formation géologique à un type de roche), l'analyse des données révèle des tendances ou des relations différentes.

En réalité, ces deux aspects du MAUP ne sont pas séparables pour un ensemble de données polygonales particulier. Ces problèmes mettent en évidence les défis de l'analyse spatiale et l'importance de comprendre l'influence de l'échelle spatiale et des limites zonales dans l'analyse des données géographiques. Il est important de tenir compte du MAUP lors de la conception des études, de l'interprétation des résultats et de la prise de décisions politiques basées sur des données spatiales, car ce qui peut être vrai pour une zonation et une échelle peut ne pas l'être pour une autre. Pour voir comment il s'agit d'une loi de la géographie, voir la boîte 1.

Pour comprendre quels types d'analyses peuvent être affectés par le MAUP, considérons la corrélation entre deux variables géographiques distribuées aléatoirement dans l'espace. Il suffit

Nom de l'étudiant: _____ # d'étudiant: _____

de savoir que la corrélation est une mesure statistique qui indique le degré de relation linéaire entre deux variables. La corrélation est échelonnée entre +1 et -1, où +1 est une corrélation positive parfaite et indique que lorsque la valeur d'une variable augmente, celle de l'autre augmente également. En contrepartie, une valeur de -1 correspond à une corrélation négative parfaite (l'augmentation d'une variable entraîne la diminution de l'autre). Une valeur proche de 0 signifie qu'il n'y a pas de relation entre les deux variables.

Dans notre cas, nous considérerons deux variables qui sont les emplacements des pique-niqueurs au cours des deux années suivantes (il peut s'agir de deux modèles de points quelconques, par exemple l'évaluation des emplacements d'échantillonnage des propriétés géotechniques à partir de données de forage sur deux ans pour voir si les échantillons sont aléatoires et non biaisés). En tout état de cause, nous émettons l'hypothèse qu'il n'existe aucune relation entre les emplacements des pique-niqueurs au cours des années suivantes, car le terrain du parc est homogène, il n'y a que de l'herbe qui a la même hauteur partout et il n'y a pas de différence sous-jacente dans la perméabilité de la géologie superficielle car la matrice du sol est la même partout (donc pas de zones mouillées), les emplacements des pique-niqueurs au cours d'une année ne devraient pas avoir de corrélation avec les emplacements des pique-niqueurs au cours de l'année suivante. En d'autres termes, un endroit est aussi bon qu'un autre sur le terrain et les pique-niqueurs devraient donc être répartis au hasard sur le terrain. Si deux variables géographiques ou spatiales sont distribuées de manière aléatoire, il ne devrait pas y avoir de corrélation entre elles, même si les données sous-jacentes sont zonées et analysées au niveau écologique. Par exemple, considérons les emplacements des pique-niqueurs au cours de deux années suivantes du jour de la fête nationale (Figure 8):

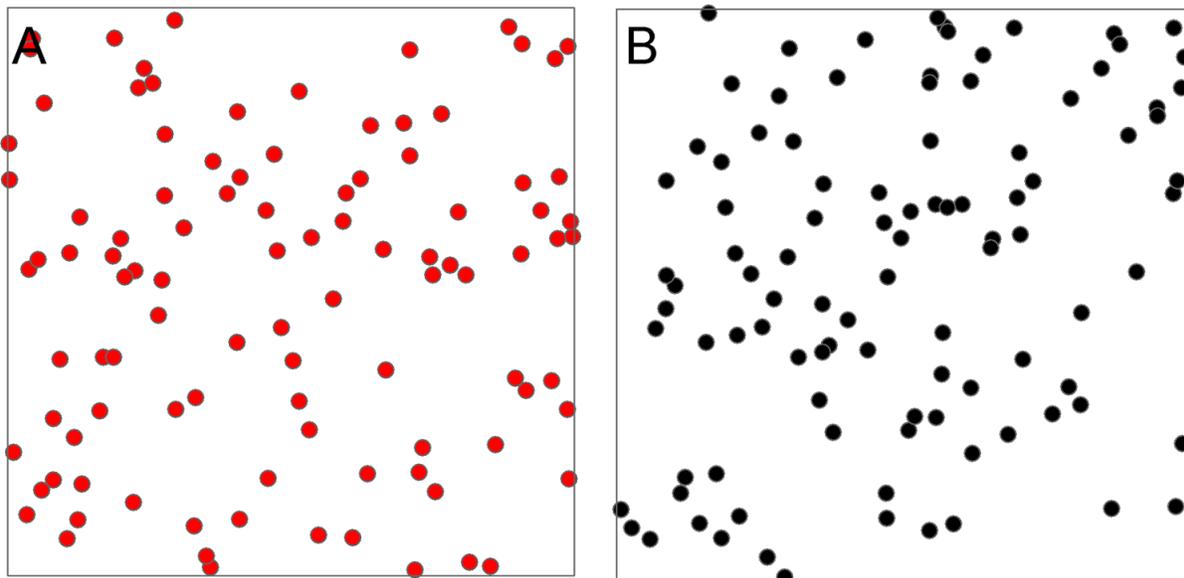


Figure 8. A) emplacements des pique-niqueurs en 2021 ; B) 2022

Ici encore, nous choisirons de comparer les densités de pique-niqueurs entre les deux années en tessellant l'espace contenant les emplacements de chaque année à l'aide de 10 polygones,

puis en corrélant les valeurs des 10 polygones entre eux. Il ne devrait pas y avoir de corrélation dans les valeurs de densité entre les deux années puisque les pique-niqueurs sont répartis au hasard chaque année. En fait, la corrélation est proche de 0, ce qui signifie qu'il n'y a pas de corrélation si l'on utilise des mesures spécifiquement adaptées à l'analyse de motifs ponctuels bivariés. Par conséquent, il n'existe en réalité aucune relation entre les emplacements des pique-niqueurs en 2021 et 2022. Toutefois, si nous n'avons pas accès aux données brutes et que les mêmes données étaient présentées dans un ensemble de polygones, la situation changerait. Selon la façon dont les polygones sont créés, le nombre de pique-niqueurs dans chaque polygone représenté comme une valeur de densité pour les deux années pourrait facilement montrer une forte corrélation positive, une corrélation nulle ou même une forte corrélation négative (Figure 9).

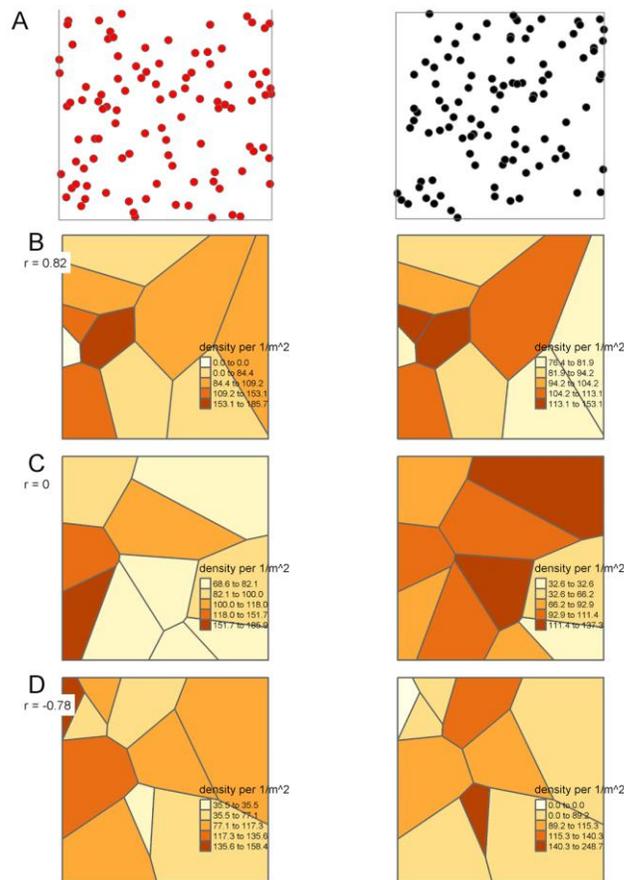


Figure 9: A) emplacements des pique-niqueurs en 2021 (à gauche) et 2022 (à droite) ; B) dix polygones qui fourniraient une forte corrélation positive ; C) pour les mêmes données, un ensemble de polygones qui ne montrerait aucune corrélation ; D) pour les mêmes données, un ensemble de polygones qui montrerait une forte corrélation négative. Rappelez-vous que la seule chose qui a changé est la configuration des polygones (taille des zones et limites des zones pour chaque zone ou polygone). Notez également que les échelles de densité varient pour chaque carte dans chaque colonne, comme expliqué précédemment. L'emplacement du point ne change jamais. Le ' r ' représente le R de Pearson, le coefficient de corrélation et, bien qu'il ne soit pas approprié pour les données géographiques, il est utilisé ici pour illustrer l'effet du MAUP sans utiliser de mesures économétriques/statistiques spatiales plus complexes.

Nom de l'étudiant: _____ # d'étudiant: _____

Ce que nous voyons dans la figure 9 est l'aspect analytique du MAUP qui illustre que si nous changeons la zonation et l'échelle de 10 polygones dans le quadrant, nous obtiendrons des corrélations différentes pour les mêmes données sous-jacentes. En fait, nous pouvons modifier au hasard les limites et les tailles des 10 polygones jusqu'à ce que nous obtenions une relation qui nous convienne. Gardez à l'esprit que nous faisons cela pour illustrer la MAUP et qu'en réalité, vous ne changeriez JAMAIS les polygones simplement pour obtenir des valeurs qui correspondent à vos idées préconçues. Cela s'appelle de la **fraude académique**.

Lorsque vous travaillez avec des données qui sont présentées comme des attributs de polygones et que ces polygones ne sont pas des "unités naturelles" telles qu'un lac ou une île, ce que vous pourriez interpréter ou déclarer comme un résultat analytique pourrait simplement être dû à la configuration des polygones. Il s'agit d'un problème majeur dans la science des données spatiales et l'analyse géographique. Vous devez en être pleinement conscient et faire part de cette limitation lorsque vous analysez des données provenant de géométries polygonales.

Question 11: Comment créer différentes zones pour visualiser le problème d'agrégation spatiale?

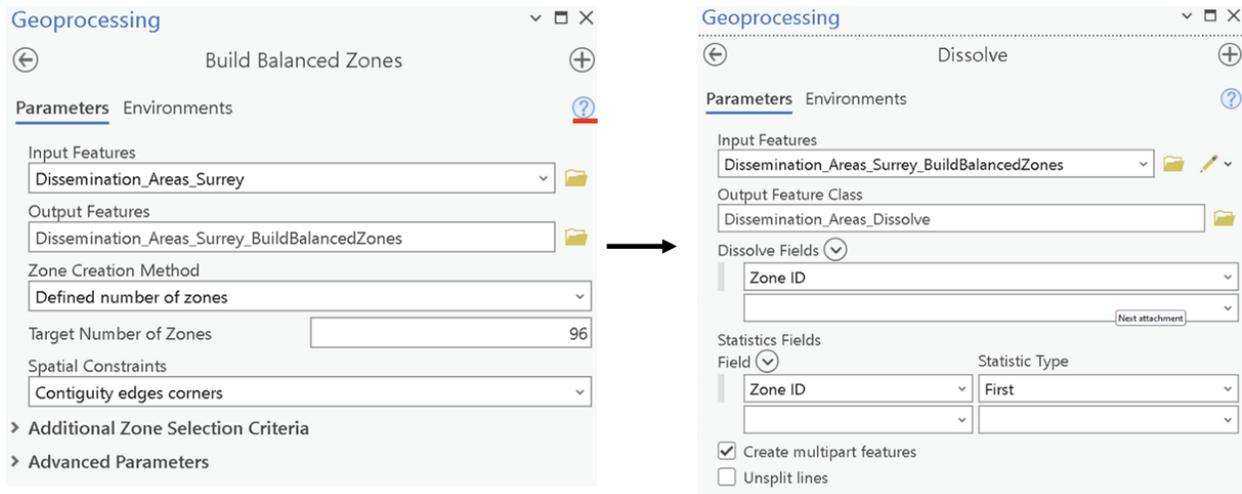
Les secteurs de recensement (SR) sont utilisés par Statistique Canada pour diviser le Canada en zones géographiques composées d'environ 7 500 personnes qui suivent des caractéristiques physiques reconnaissables telles que des rivières ou des routes. Les secteurs de recensement sont utilisés à Surrey, en Colombie-Britannique, pour diviser la ville en zones utilisées à des fins statistiques. Cependant, la définition des SR peut influencer la manière dont les données de recensement en question sont représentées. Dans cette question, nous allons créer des SR synthétiques et évaluer comment différentes limites de zones peuvent modifier la densité des sites patrimoniaux par SR. Cet exemple de MAUP montrera comment la division des zones dans une carte peut modifier le message de la carte.

La première étape consistera à créer un nouvel ensemble de SR synthétiques dont les limites diffèrent de celles définies par Statistique Canada. Les SR sont créés à partir du conglomérat des aires de diffusion (AD), qui constituent la plus petite unité géographique utilisée dans le recensement. Nous allons créer un nouvel ensemble de SR synthétiques en utilisant les outils de géotraitement d'ArcGIS et les aires de diffusion. La première étape est de localiser la classe d'entités des aires de diffusion, qui se trouve dans la géodatabase

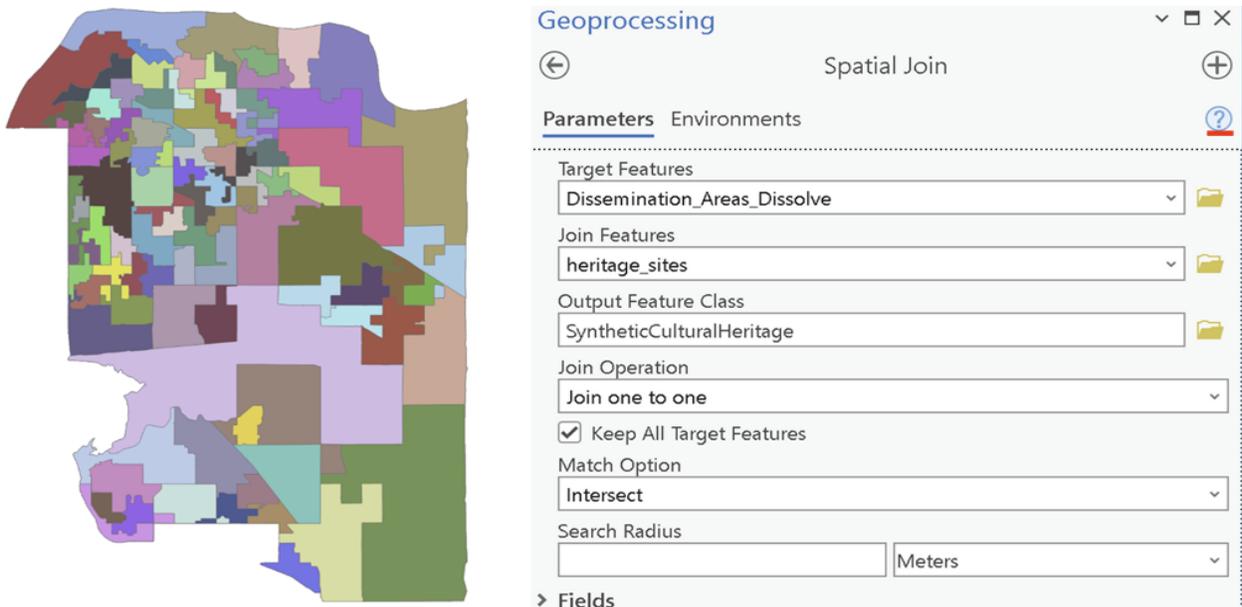
Census_2016_Surrey_DisseminationAreas_UTM.gdb, à l'intérieur du jeu de données d'entités **Surrey_Without_Data** (dossier), appelé "**Dissemination_Areas_Surrey**". Ajoutez-la à votre carte. Ensuite, dans la barre de menu du haut, cliquez sur « *Analysis* », puis sur « *Tools* » (outils), et le panneau de « *geoprocessing* » (géotraitement) s'affichera dans la partie gauche de votre écran. Dans la barre de recherche "find tool", recherchez "Build Balanced Zones" et cliquez sur l'outil. Choisissez comme input feature "**Dissemination_Areas_Surrey**", « *Zone creation method* » doit être « *Defined number of zones* » (Nombre de zones défini) et définissez le nombre cible de zones comme étant **96**, puisque c'est le nombre de SR qu'il y a dans la ville de Surrey BC. Cliquez ensuite sur « *Run* » (Exécuter). Si vous rencontrez des

Nom de l'étudiant: _____ # d'étudiant: _____

problèmes avec l'outil, cliquez sur le point d'interrogation en haut à droite pour accéder à la page d'aide d'ESRI.



Une fois que les zones équilibrées ont été générées, nous devons utiliser l'outil de dissolution pour agréger les AD dans leurs SR synthétiques. Dans la barre de recherche de l'outil de recherche, entrez « *Dissolve* » et cliquez sur l'outil. L'élément d'entrée doit être **Dissemination_Areas_Surrey_BuildBalancedZones**, et l'outil de dissolution doit être configuré à "Zone ID". Cliquez ensuite sur Run pour exécuter l'outil. Notez que vous pouvez recevoir un avis disant que « *The Number of Zones requested was not reached* » (le nombre de zones demandées n'a pas été atteint) - ignorez cet avertissement et continuez à utiliser la classe d'entités. Tous les résultats des outils de géotraitement sont enregistrés par défaut dans la géodatabase du dossier d'origine dans lequel votre projet est enregistré. Votre résultat de l'outil de dissolution devrait ressembler à la carte ci-dessous, mais avec une symbologie différente pour les zones.



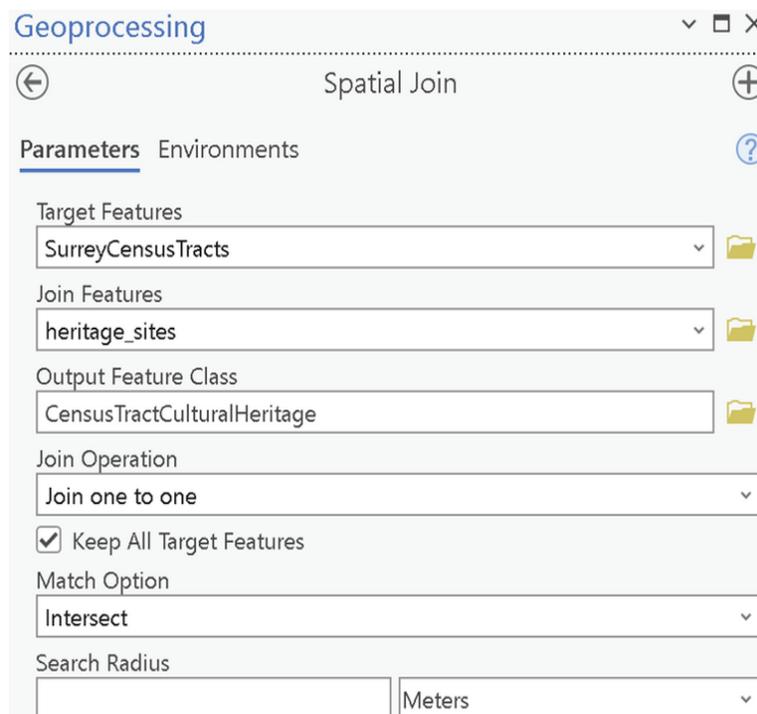
Nom de l'étudiant: _____ # d'étudiant: _____

L'étape suivante est d'ajouter les données de points des sites patrimoniaux aux SR synthétiques à l'aide de « *Spatial Join* ». Un champ intitulé "**Join_count**" sera créé et nous permettra de savoir combien de sites patrimoniaux sont situés dans chaque SR synthétique. Dans la barre de recherche, entrez « *Spatial Join* » et cliquez sur l'outil. « *Spatial Join* » (jointure spatiale) est un outil complexe, il est donc recommandé de cliquer sur le point d'interrogation pour mieux comprendre le fonctionnement de l'outil avant de continuer. Utilisez la jointure spatiale avec les paramètres suivants:

- « Target Features »: Dissemination_Area_Dissolve
- « Join Features »: heritage_sites (*note: si vous n'avez pas encore ajouté heritage_sites à votre carte, cliquez sur l'icône de dossier à côté de Join Features et accédez à la classe d'entités heritage_sites dans heritage_sites.gdb*).
- « Join operation »: « Join one to one »
- « Keep all target features »: Coché
- « Match Option »: « Intersect »

Cliquez ensuite sur **Run**.

Une fois que vous avez terminé la jointure spatiale pour les SR synthétiques, ouvrez **SurreyCensusTracts** dans votre carte à partir du fichier **Census2016_Surrey_CensusTract_UTM.gdb**. Effectuez une jointure spatiale similaire en joignant les **heritage_sites** à la classe d'entités **SurreyCensusTracts**.



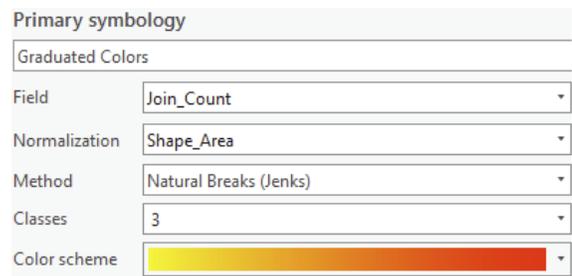
Nom de l'étudiant: _____ # d'étudiant: _____

Question 12: Comment symboliser deux classes d'entités à l'aide de la même symbologie?

Cette question utilise un outil qui permet à l'utilisateur de symboliser une classe d'entités en se basant sur la symbolisation d'une autre classe d'entités. Il est important d'avoir les mêmes intervalles et une symbolisation similaire lorsque l'on compare les résultats entre les éléments de la carte, c'est donc un outil très important à savoir utiliser.

Une fois que vous avez réalisé les deux jointures spatiales, il est temps de les symboliser en utilisant les mêmes intervalles afin que nous puissions faire des comparaisons entre les cartes. Nous allons comparer la densité des sites patrimoniaux par zone. Tout d'abord, symbolisons la classe de caractéristiques **SyntheticCulturalHeritage**. Cliquez avec le bouton droit de la souris sur **PatrimoineCulturelSynthétique** et sélectionnez « *Symbology* », puis symbolisez l'élément avec:

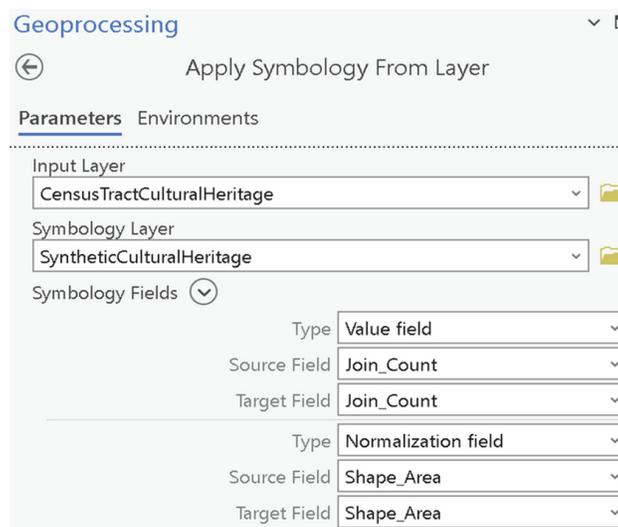
- « Primary symbology »: « *Graduated Colors* »
- « Field »: **Join_count**
- « Normalization »: **Shape_Area**
- « Method »: « *Natural Breaks (Jenks)* »
- « Classes »: **3**



Primary symbology	
Graduated Colors	
Field	Join_Count
Normalization	Shape_Area
Method	Natural Breaks (Jenks)
Classes	3
Color scheme	

Ensuite, étiqueter les trois classes, **faible**, **moyenne** et **élevée**.

Une fois que vous avez terminé la symbolisation du **SyntheticCulturalHeritage**, utilisez l'outil de géotraitement « *Apply Symbology From Layer* » pour appliquer une symbologie similaire au **CensusTractCulturalHeritage**. Choisissez la couche **CensusTractCulturalHeritage** pour « *Input Layer* » et la couche **SyntheticCulturalHeritage** pour « *Symbology Layer* », puis exécutez l'outil.



Geoprocessing			
Apply Symbology From Layer			
Parameters		Environments	

Input Layer	CensusTractCulturalHeritage		
Symbology Layer	SyntheticCulturalHeritage		
Symbology Fields			
Type	Value field		
Source Field	Join_Count		
Target Field	Join_Count		

Type	Normalization field		
Source Field	Shape_Area		
Target Field	Shape_Area		

Vous devriez maintenant avoir deux couches avec une symbologie identique qui peuvent être utilisées pour la comparaison.

Questions à remettre

1. [5 points] Créez une carte qui respecte les critères ci-dessous:

Il ne doit pas y avoir de carte de base (préchargée).

Les sites historiques suivants sont étiquetés sur la carte: « Bulman's Garage », « Port Kells Post Office », « Port Kells Fire Hall No. 7 », « Port Kells Elementary School », « Church Manse », and « Port Kells Park ».

Inclure les pistes cyclables et les sites écosystémiques dans cette carte.

Les sites patrimoniaux doivent être étiquetés et leur nom de site ne doit pas toucher ou traverser les pistes cyclables.

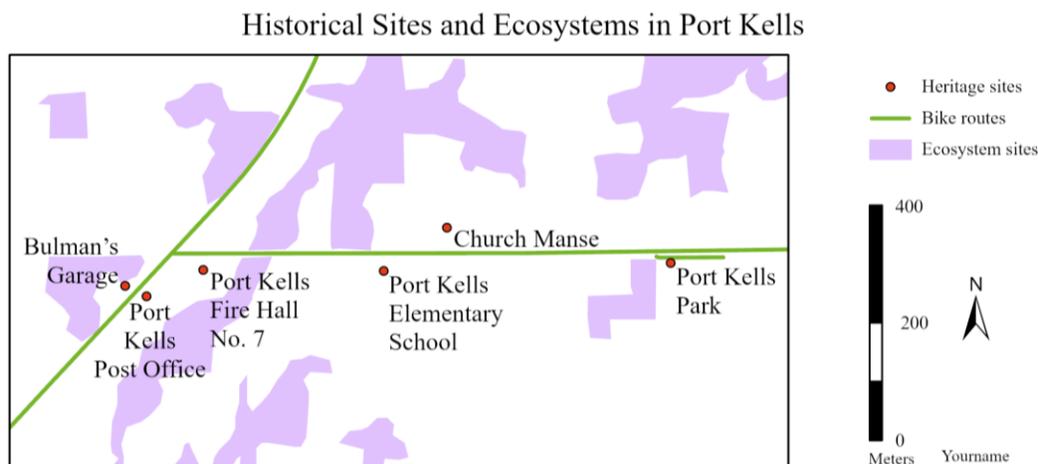
L'échelle de la carte doit être de 1:10 000.

Les meilleures pratiques cartographiques doivent être utilisées lors de la création de cette carte.

Vous devez utiliser la police **Arial**, ainsi qu'une barre d'échelle et une flèche du nord différentes de celles de la carte ci-dessous.

La couleur de la symbologie des sites écosystémiques, des pistes cyclables et des sites patrimoniaux doit être différente de celle de la carte ci-dessous.

Votre carte doit ressembler à celle présentée ci-dessous, mais il n'est pas nécessaire qu'elle soit identique :



2. [3 points] Dans un paragraphe court (150 mots maximum), expliquez comment la carte que vous avez soumise pour la question 1 utilise la hiérarchie visuelle. Pourquoi est-il important d'établir une hiérarchie visuelle?

3. [8 points] Comme dans les exemples de questions 10 et 11 ci-dessus, créez une nouvelle classe d'entités de recensement synthétique. Ensuite, joindre spatialement **important_trees** (situé dans **important_trees.gdb**) au secteur de recensement synthétique et à une classe d'entités de secteur de recensement. En utilisant les meilleures pratiques cartographiques, créez deux cartes avec une symbologie et une étendue identique et avec des éléments similaires. Veillez à ce que chaque carte ait un titre approprié

4. [2 points] Expliquez le problème d'agrégation spatiale (MAUP) et comment il se manifeste dans les deux cartes que vous avez créées pour la question 3.

Nom de l'étudiant: _____ # d'étudiant: _____

5. [1 point] Si je devais créer une carte montrant le nombre de chiens présents dans chaque secteur de recensement de Surrey (C.-B.), quelle symbologie devrais-je utiliser ? Justifiez votre réponse (100 mots maximum).

6. [1 point] Si je devais créer une carte montrant la densité de chiens dans chaque secteur de recensement de Surrey (C.-B.), est-ce qu'une carte choroplèthe serait appropriée ? Justifiez votre réponse (100 mots maximum).

Nom de l'étudiant: _____ # d'étudiant: _____

Laboratoire n°2 : Systèmes de coordonnées et transformations

Objectif : L'objectif de ce laboratoire est de présenter les systèmes de coordonnées géographiques et les transformations. L'exercice suivant vous permettra de définir un système de coordonnées et de calculer l'incertitude de la position dans les couches géospatiales numériques.

Résultats d'apprentissage : À la fin de ce laboratoire, vous serez en mesure de:

- Identifier les principaux composants de l'interface utilisateur d'ArcGIS Desktop ArcGIS Pro.
- Définir un système de coordonnées géographiques.
- Transformer les systèmes de coordonnées géographiques.
- Calculer l'incertitude de localisation des couches géospatiales numériques.

Procédure:

- Complétez le groupe 1, tutoriels 1-5, à partir de l'ESRI Academy.
- Lisez et travaillez sur les exemples.
- Répondez aux questions à la fin du laboratoire.

Données/matériel:

- Toutes les données sont fournies dans **lab2.zip**.

Devoir:

- Répondez aux questions (#1-9).
- Le total des points disponibles pour ce travail est de 27 points, ce qui représente 10% de votre note finale.
- Tous les travaux pratiques doivent être réalisés et remis individuellement.

Format:

Les étudiants remettront les questions avec les réponses complétées.

Les tableaux et les figures doivent être produits selon les normes du Département de géographie, environnement et géomatique de l'Université d'Ottawa.

Pour des directives sur le formatage et la présentation appropriés des tableaux et des figures, veuillez consulter la [section d'introduction](#) de ce manuel de laboratoire.

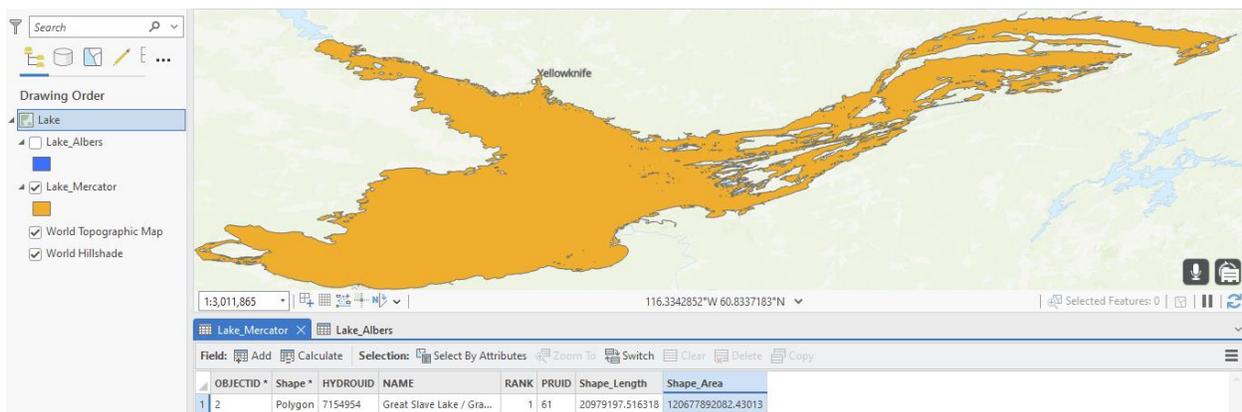
Nom de l'étudiant: _____ # d'étudiant: _____

Q1: Pourquoi la projection des données est-elle importante dans les SIG ?

La projection, qui consiste à convertir un modèle ellipsoïdal de la terre sur une carte plate, est un élément essentiel pour travailler avec des données géospatiales. Les projections peuvent être utilisées pour préserver les caractéristiques suivantes des entités terrestres : forme, superficie, direction et distance. Cependant, il n'existe pas de projection unique capable de préserver toutes ces caractéristiques d'entités terrestres sur un dispositif de visualisation bidimensionnel tel qu'un moniteur ou un papier plat. Vous devez donc choisir la projection appropriée pour la tâche à accomplir, en fonction de l'échelle (ville, pays, continent, monde) et de l'emplacement de ce que vous cartographiez.

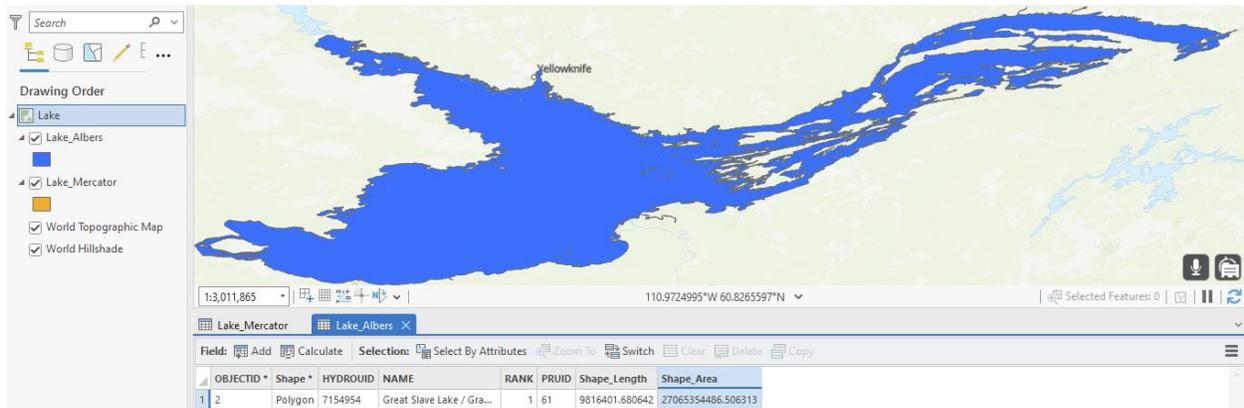
La projection la plus couramment utilisée pour les cartes du monde est la projection de Mercator (créée par Gerardus Mercator en 1569). La projection de Mercator a été conçue pour la navigation et préserve relativement bien les formes. De plus, chaque ligne droite sur une carte de Mercator est une ligne de rhumb, c'est-à-dire une ligne de direction constante, qu'un navire pourrait utiliser pour naviguer et se rendre d'un point à un autre. La projection de Mercator, plus que toute autre, a été l'un des outils essentiels de la colonisation mondiale.

Cependant, la projection de Mercator n'est pas adaptée à la plupart des cartes du Canada, car elle crée d'importantes distorsions des objets spatiaux. En effet, à mesure que l'on s'éloigne de l'équateur, que ce soit vers le nord ou vers le sud, la superficie et la distance augmentent. La façon dont la projection de Mercator déforme le Grand lac des Esclaves dans les Territoires du Nord-Ouest en est un exemple. En utilisant le système de coordonnées de Mercator, la taille calculée du lac est d'environ 120 678 km².



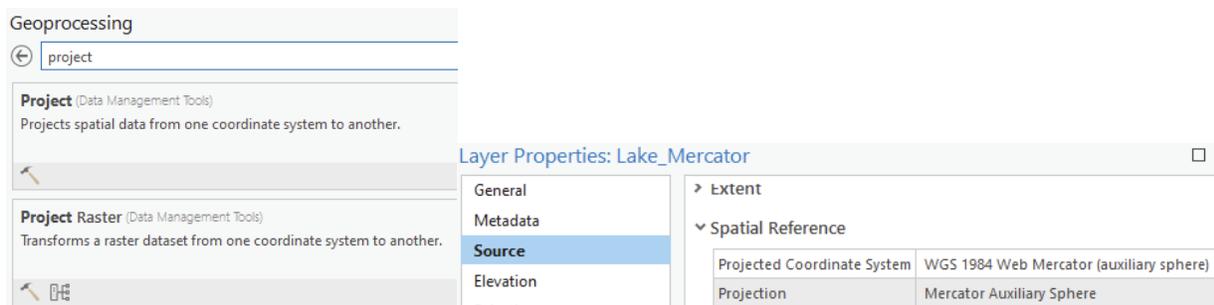
Toutefois, si l'on utilise une projection créée spécifiquement pour le Canada et destinée à préserver à la fois la forme et, dans une certaine mesure, la superficie, l'Albers Equal Area, la taille du lac est d'environ 27 065 km². Ce chiffre est nettement plus proche de la taille réelle du lac, qui est estimée à 27,200 km².

Nom de l'étudiant: _____ # d'étudiant: _____



Il est essentiel d'utiliser la bonne projection lorsque vous travaillez avec des données géospaciales. Si vous envisagez d'effectuer des calculs de superficie ou des analyses spatiales à partir des données, il est nécessaire de projeter vos données dans une projection qui préserve la superficie. Les projections telles que la projection conique à aire égale de Canada Albers (« *Canada Albers Equal Area Conic* ») sont les plus appropriées pour les besoins de projets portant uniquement sur des aires et, de même, la projection conique conforme de Lambert (« *Lambert Conformal Conic* ») est appropriée pour les cartes à petite échelle du Canada.

Les outils *Project* et *Project Raster* d'ArcGIS Pro peuvent être utilisés pour convertir vos données dans la projection appropriée. Pour déterminer la projection d'une couche, cliquez dessus avec le bouton droit de la souris, cliquez sur « *Propriétés* » (propriétés), cliquez sur « *Source* » (source), puis faites défiler vers le bas et ouvrez la section « *Spatial Reference* » (références spatiales). Seules les couches dont la projection est déjà définie peuvent être reprojetées dans une autre projection.



Q2: Comment convertir les degrés, minutes et secondes en degrés décimaux?

Si l'on vous donne des coordonnées telles que 79°22'34" W, la première chose à faire est de remarquer que ces coordonnées font référence à un lieu situé à l'OUEST du méridien d'origine. Par conséquent, vous devrez ajouter un signe (-) au résultat après l'avoir converti en degrés décimaux. Il s'agit là d'une simple convention : les longitudes à l'est du méridien d'origine sont

Nom de l'étudiant: _____ # d'étudiant: _____

considérées comme positives et celles à l'ouest sont considérées comme négatives. Pour convertir 79°22'34" W en degrés décimaux, suivez les équations suivantes :

$$dd = \text{deg} + \frac{\text{min}}{60} + \frac{\text{sec}}{60^2}$$

$$\text{e.g., } Long = 79^{\circ}22'34"W$$

$$\begin{aligned} dd &= 79 + \frac{22}{60} + \frac{34}{60^2} \\ &= 79 + .3666 + .00944 \\ &= -79.37611W \end{aligned}$$

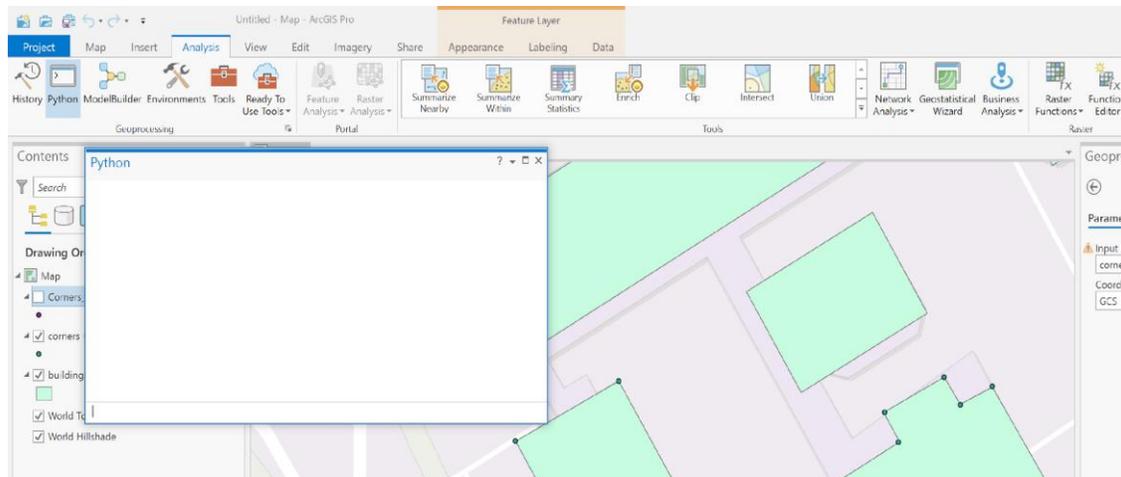
Remarquez qu'à la fin, un signe (-) est ajouté. Ce signe ne signifie pas "négatif", mais indique plutôt à quelqu'un ou à quelque chose que cette coordonnée se trouve à l'ouest, et non à l'est, de Greenwich, en Angleterre.

Pour inverser le processus, il faut enlever la partie -79 du degré décimal de la coordonnée, ce qui donne des minutes décimales. Vous multipliez les minutes décimales par 60 et enlevez la partie entière de ce qui reste. Ensuite, vous multipliez la partie décimale restante par 60 pour obtenir le nombre de secondes, comme ceci:

$$\begin{aligned} &-79.37611 \text{ } dd \\ &= 79^{\circ} \text{ et } .37611 \times 60' \\ &= 79^{\circ} \text{ et } 22.5666 \text{ } dm \\ &= 79^{\circ} \text{ et } 22' \text{ et } .5666 * 60'' \\ &= 79^{\circ} \text{ et } 22' \text{ et } 33.996'' \\ &= 79^{\circ}22'33.996'' \end{aligned}$$

Attendez, pourquoi ce résultat est-il différent de 79°22'34" W ? Notez que vous devriez idéalement faire cela sur une calculatrice ou un ordinateur, sinon vous obtiendriez un résultat différent pour les secondes en raison de la précision numérique (le nombre de décimales que vous conservez par rapport à un ordinateur), comme cela s'est produit ici. La méconnaissance de la précision numérique peut conduire à des erreurs plus importantes ! Heureusement, ces calculs sont faciles à effectuer dans la fenêtre Python d'ArcGIS. Ouvrez la fenêtre Python à partir du ruban « *Analysis* » (Analyse) d'ArcGIS Pro:

Nom de l'étudiant: _____ # d'étudiant: _____



Tapez le texte suivant:

$$dd = 79 + 22.0/60.0 + 34.0/60.0**2$$

Cette instruction enregistre le résultat dans une variable appelée `dd` dans la fenêtre Python. Pour voir ce qui est enregistré, tapez le nom `dd` comme ceci, et appuyez sur Entrée:

```
Python
dd=79+22/60+34/60**2
dd
79.37611111111111
```

Vous pouvez voir qu'il y a beaucoup de décimales, en fait les 1 se répètent pendant longtemps. Il faut donc reconvertir `dd` en degrés, minutes et secondes. Tout d'abord, nous enlevons le 79, en le notant sur un morceau de papier, pour obtenir les minutes décimales (DM),

```
Python
m=dd-79
m
0.37611111111111006
```

Dans `m`, vous avez la partie décimale des minutes, nous allons donc la multiplier par 60 et enregistrer le résultat dans une variable appelée `mins`,

Nom de l'étudiant: _____ # d'étudiant: _____

```
Python
mins = m * 60
mins
22.566666666666038
```

Notez également 22 minutes, là où vous avez noté 79 degrés précédemment. Enfin, supprimez les 22 minutes:

```
Python
s = mins - 22
s
0.5666666666660376
```

et multiplie le résultat enregistré dans `s` par 60 pour obtenir le nombre de seconds.

```
Python
s*60
33.999999999962256
```

Le résultat peut donc être arrondi à 34". On voit bien que, même si 79°22'34" W peut être exprimé comme la somme de nombres rationnels, l'ordinateur ne peut pas retenir suffisamment de décimales pour garantir un résultat parfait de 34", et le résultat peut donc être irrationnel. Cela doit être considéré comme une source d'incertitude lors de la conversion de DMS en DD.

Q3: Quelle est la distance réelle d'un degré de longitude à l'équateur?

La distance d'un degré de longitude, par exemple si vous naviguez de la longitude 0° à 1° Ouest exactement le long d'une ligne de latitude, en supposant que la Terre est une sphère, est donnée par la formule suivante:

$$d = 111 \text{ km} \times \cos(\text{latitude})$$

Nom de l'étudiant: _____ # d'étudiant: _____

La latitude est généralement exprimée en radians plutôt qu'en degrés, soit $rad = deg \times \frac{\pi}{180}$.
Donc, par exemple, pour une latitude de 0° :

$$\begin{aligned}d &= 111 \text{ km} \times \cos\left(0 \times \frac{\pi}{180}\right) \\ &= 111 \text{ km} \times 1 \\ &= 111 \text{ km}\end{aligned}$$

Il y a donc 111 km dans 1 degré de longitude - à l'équateur.

Q4: Comment calculer la distance réelle entre deux points en utilisant la longitude et la latitude?

Vous pouvez calculer la distance entre deux points dont vous connaissez la longitude et la latitude en utilisant ce que l'on appelle la formule de Haversine pour la distance orthodromique (« *great-circle distance* »). Cette formule suppose que la terre est une sphère, et l'équation suivante fournit une approximation de la véritable distance géodésique (distance sur un ellipsoïde):

$$D_{A \rightarrow B} = r \cos^{-1} \left[\sin \phi_A \sin \phi_B + \cos \phi_A \cos \phi_B \cos(\beta) \right]$$

où, $D_{A \rightarrow B}$ = distance du point A au B
 r = rayon de la terre en km ou m
 ϕ_A = latitude du point A
 λ_A = longitude du point A
 ϕ_B = latitude du point B
 λ_B = longitude du point B
 $\beta = \lambda_A - \lambda_B$

L'équation ci-dessus suppose que toutes les mesures angulaires sont exprimées en radians.
Ainsi, par exemple, si nous avons les deux points suivants:

Point 1		Point 2	
LAT	LONG	LAT	LONG
46.62006	-84.3039	46.62009	-84.3038

Nom de l'étudiant: _____ # d'étudiant: _____

$$D_{A \rightarrow B} = r \cos^{-1} [\sin \phi_A \sin \phi_B + \cos \phi_A \cos \phi_B \cos(\beta)]$$

where, $D_{A \rightarrow B}$ = distance from point A to B and,

$$r = 6378 \text{ km}$$
$$\phi_A = 46.620061 \times \frac{\pi}{180} = 0.813673562$$
$$\lambda_A = -84.303857 \times \frac{\pi}{180} = -1.471379877$$
$$\phi_B = 46.62009 \times \frac{\pi}{180} = 0.813674138$$
$$\lambda_B = -84.30383 \times \frac{\pi}{180} = -1.471379406$$
$$\beta = -1.471379877 - (-1.471379406) = -4.71239E - 07$$
$$= 6378 \cos^{-1} \left[\begin{array}{l} \sin(0.813673562) \sin(0.813674138) + \\ \cos(0.813673562) \cos(0.813674138) \cos(-4.71239E - 07) \end{array} \right]$$
$$= 0.00426 \text{ km} = 4.26 \text{ m}$$

Donc, la distance approximative entre ces deux points pour un modèle sphérique de la terre, d'un rayon de 6378 km, est d'un peu plus de 4 m. Il est toujours préférable d'utiliser un ordinateur pour effectuer de tels calculs, car il faut conserver de nombreuses décimales. Comme dans l'exemple ci-dessus, vous pouvez facilement effectuer ce calcul dans la fenêtre python d'ArcGIS (attention à l'utilisation des parenthèses !):

```
Python
>>> Lat1 = math.radians(46.620061)
>>> Long1 = math.radians(-84.303857)
>>> Lat2 = math.radians(46.620094)
>>> Long2 = math.radians(-84.30383)
>>> 6378*math.acos(math.sin(Lat1)*math.sin(Lat2)+math.cos(Lat1)*math.cos(Lat2)*math.cos(Long1 - Long2))
0.004211873988023198
>>> |
```

Dans l'exemple ci-dessus, les unités sont en km parce que le rayon de la terre était en km. Vous pouvez convertir en mètres en multipliant le résultat par 1000. Vous pouvez également effectuer cette opération dans Excel en utilisant les fonctions RADIANS() et ACOS(), COS() et SIN() dans une formule de feuille de calcul.

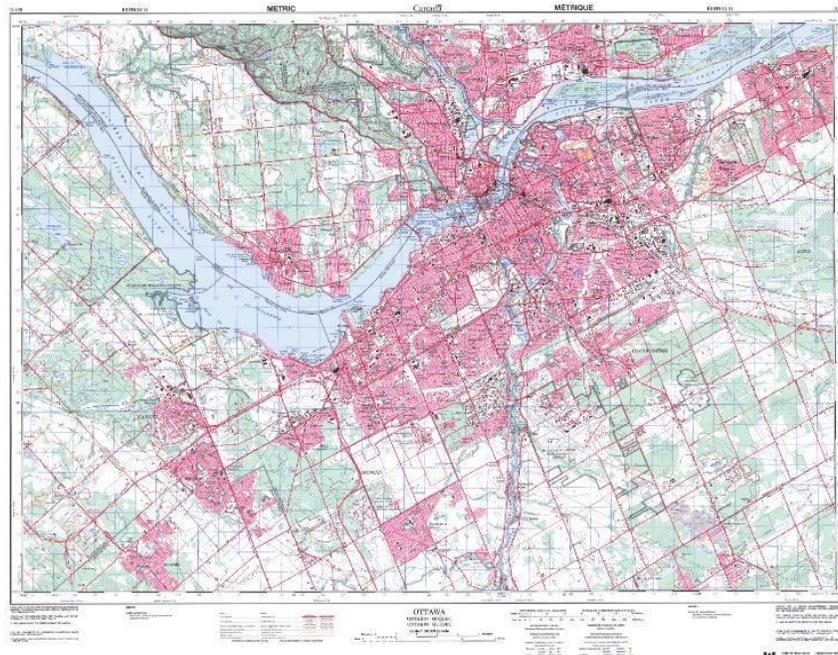
Q5: Comment calculer l'incertitude de localisation d'une couche numérique?

Une carte sur papier est produite à partir de photographies aériennes en utilisant des techniques photogrammétriques. Un photogrammétriste examine deux photos aériennes superposées à l'aide d'un stéréoscope et utilise une série de techniques et de connaissances spécialisées pour transférer (tracer) ces caractéristiques sur une feuille d'acétate transparente. La hauteur de vol et l'objectif de l'appareil photo, ainsi que d'autres aspects de la photo aérienne, déterminent l'échelle et, à son tour, la plus grande échelle cartographique que le cartographe peut produire à partir de ces photos. Vous pouvez voir comment un stéréoscope fonctionne avec des photos aériennes en cliquant ici: <https://youtu.be/mZkmXEFw7BE>. La carte

Nom de l'étudiant: _____ # d'étudiant: _____

produite par le cartographe est ensuite numérisée manuellement ou numériquement à l'aide d'algorithmes humains ou mécaniques afin de créer des couches de données géospatiales numériques utilisées dans les processus géomatiques au sein d'un SIG.

En fonction de ce processus, une couche de données géospatiales numériques provenant d'une carte papier est associée à une **échelle de base**. L'**échelle de base** est l'échelle de la carte sur laquelle la représentation numérique éventuelle est basée. Ces représentations numériques deviennent ensuite des couches thématiques telles que les routes, les zones bâties, les rivières, les zones forestières, etc. Par exemple, une carte de la région d'Ottawa à l'échelle 1:50 000 (NTS O31G5) se présente comme ci-dessous :

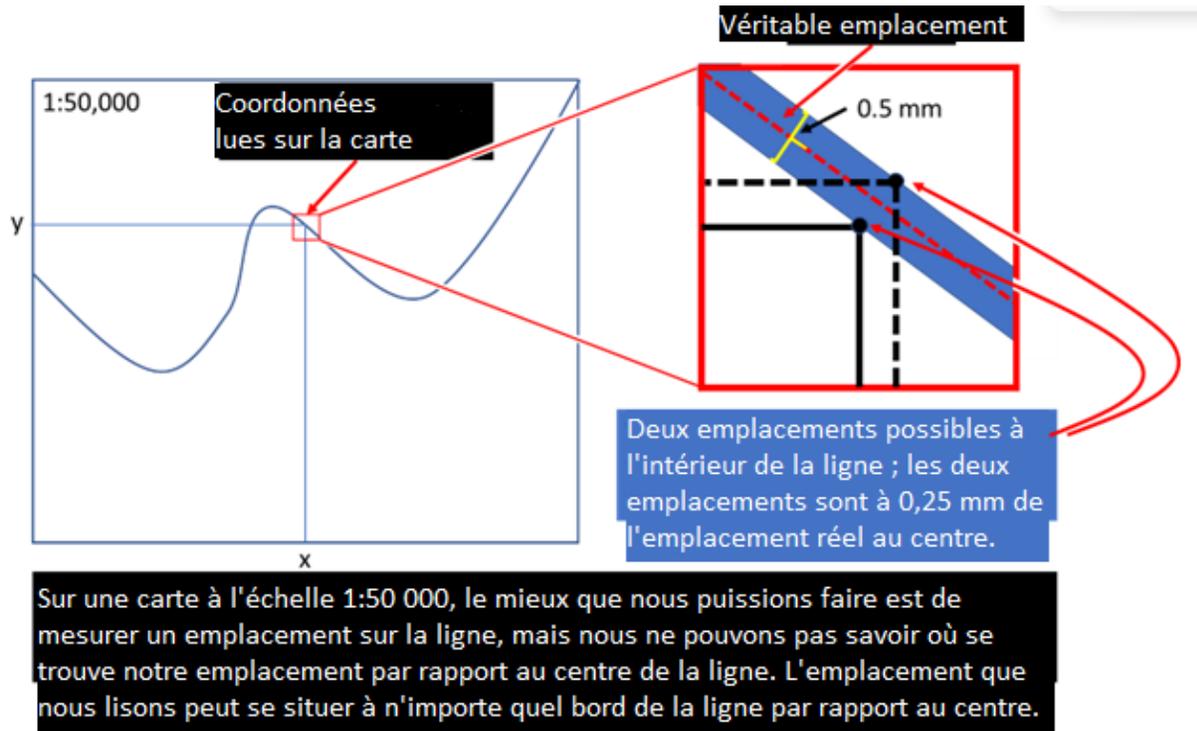


L'élément le plus petit d'une carte qui peut être vu à une distance d'observation normale est de 0,5 mm. Ainsi, pour une carte à l'échelle 1:50 000, cette ligne de 0,5 mm (ou le diamètre de 0,5 mm d'un point) couvre une largeur de $0,5 \text{ mm} \times 50\,000 = 25\,000 \text{ mm}$, ou 25 m, dans le monde réel.

En supposant que le cartographe n'a pas été biaisé en traçant les objets sur les photos aériennes lors de la représentation cartographique, nous pouvons dire qu'une position (une paire de coordonnées) lue à partir d'un élément sur cette carte sera au maximum à $\frac{1}{2}$ de la largeur de la ligne de 0,5 mm de sa véritable position. Pourquoi ? Parce que nous devons supposer que la véritable position d'un élément sur une carte se trouve au centre de la ligne de 0,5 mm. $0,5 \times 25 \text{ m} = 12,5 \text{ m}$, de sorte que l'incertitude d'une position lue à partir d'un élément sur une carte à l'échelle 1:50 000 serait d'environ 12,5 m. Il est important de garder à l'esprit une telle incertitude positionnelle lorsque vous utilisez des données géospatiales. Si vous utilisez une carte pour organiser un pique-nique, une erreur de 12,5 m dans le lieu de rendez-vous n'est pas très grave. Mais si vous faites atterrir un avion de nuit, une erreur de 12,5 m sur

Nom de l'étudiant: _____ # d'étudiant: _____

le côté peut vous placer en dehors de la piste d'atterrissage, ou si vous construisez un pont qui relie la route de l'autre côté...



Un concept similaire est celui de l'élément résoluble minimal, qui est basé sur l'échelle de la carte. L'élément résoluble minimal correspond à la distance de séparation minimale entre deux lignes de manière à ce qu'elles puissent être distinguées en tant qu'entités individuelles sur une carte. Cette valeur est basée uniquement sur la valeur de 0,5 mm et représente, dans le cas d'une carte à l'échelle 1:50 000, la valeur de 25 m.

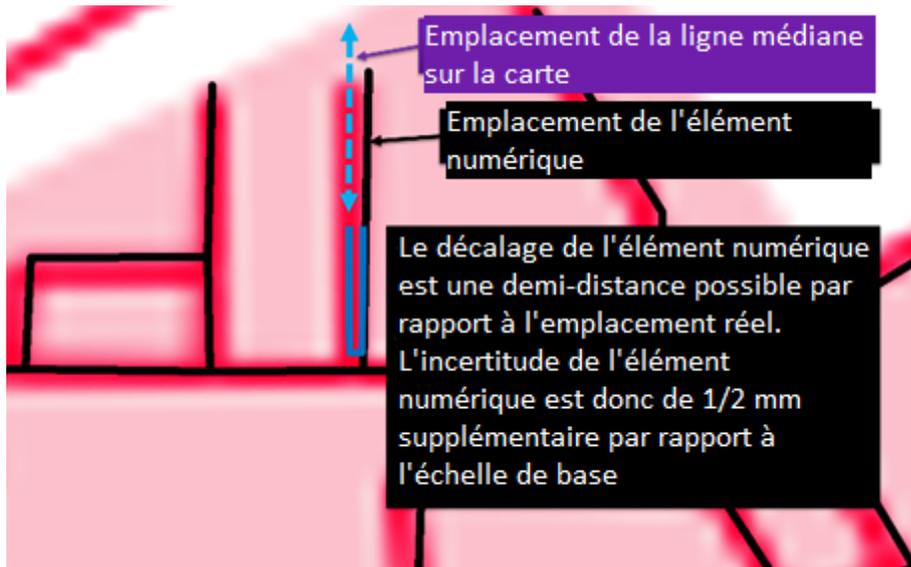
Alors, pour calculer l'incertitude de la position d'une couche cartographique géospatiale due à l'échelle de base, il suffit de procéder comme suit:

$$\begin{aligned} \delta &= 0.25 \text{ mm} \times \text{échelle dénominateur} \\ \text{échelle de la carte} &= \frac{1}{50000} \\ \delta &= 0.25 \text{ mm} \times 50000 \\ &= 12500 \text{ mm} \end{aligned}$$

Cependant, une couche géospatiale numérique provenant d'une carte papier comporte une incertitude supplémentaire, car l'entité (humaine ou IA) qui a créé la représentation numérique a dû extraire des caractéristiques de la carte papier en les traçant dans un système SIG. Ce processus s'appelle la **numérisation** et ajoute lui-même de l'incertitude. Une machine ou une personne ne peut pas tracer parfaitement les caractéristiques d'une carte papier. Un tracé parfait se produirait lorsque la personne ou la machine tracerait toujours les éléments sur la

Nom de l'étudiant: _____ # d'étudiant: _____

ligne centrale exacte de l'élément représenté sur la carte, mais cela n'arrive pas très souvent, même en utilisant des algorithmes de machine, à cause de choses simples comme la distorsion de l'image lorsqu'une carte est photographiée (numérisée), l'étirement ou le rétrécissement du papier dû à l'humidité et/ou à l'intérieur d'un scanner, etc. La figure ci-dessous montre un exemple où l'emplacement d'un élément numérique (ligne noire) est situé à l'intérieur, mais pas au centre, de l'élément sous-jacent (ligne rouge épaisse).



En supposant que la représentation numérique respecte la largeur de trait de 0,5 mm lors du traçage de la carte, les coordonnées d'un lieu sur la représentation numérique (lignes noires dans la figure 7) auront une incertitude par rapport à la carte de base égale à celle de l'échelle de base de la carte, plus précisément,

$$\begin{aligned} \delta &= 0.25 \text{ mm} \times \text{échelle dénominateur} \\ \text{échelle de la carte} &= \frac{1}{50000} \\ \delta &= 0.25 \text{ mm} \times 50000 \\ &= 12500 \text{ mm} \end{aligned}$$

L'incertitude totale est calculée à l'aide de formules de propagation des erreurs qui supposent qu'il n'y a pas de **corrélations** entre les erreurs de la personne qui a réalisé la carte originale et les erreurs commises par l'entité qui l'a transformée en une représentation numérique. L'erreur totale est donc additive, et est donnée par:

Nom de l'étudiant: _____ # d'étudiant: _____

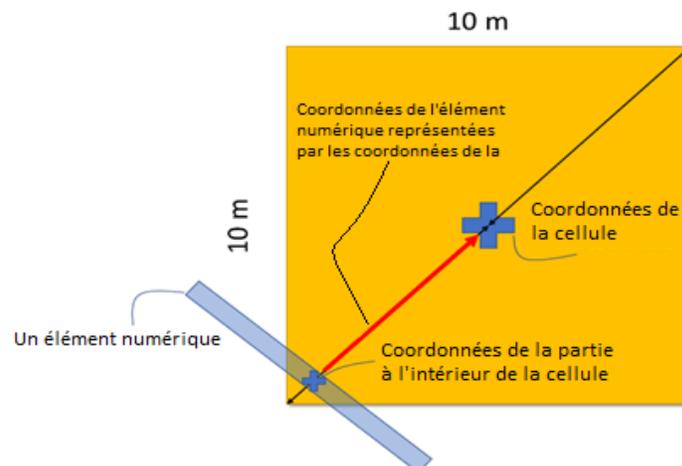
$$\begin{aligned} \text{total } \delta &= \sqrt{\text{erreur d'échelle de base}^2 + \text{erreur numérique}^2} \\ &= \sqrt{12500^2 + 12500^2} \\ &= \sqrt{156250000 + 156250000} \\ &= \sqrt{312500000} \\ &= 17677.67 \\ &= 17.68 \text{ m} \end{aligned}$$

Cette formule est dérivée d'une équation différentielle en série de Taylor. Vous pouvez consulter le document intitulé *UncertaintyinGISProcesses.pdf* pour la dérivation des différentes formules.

Enfin, si vous décidez d'imprimer ou d'afficher une carte à une échelle plus petite que l'échelle de base, vous devrez ajouter un autre terme à l'équation ci-dessus. N'oubliez pas non plus qu'il ne s'agit que d'une approximation et qu'il existe probablement d'autres sources d'incertitude dans le processus d'élaboration de la carte qui ne sont pas prises en compte par l'échelle seulement.

Q6: Quelle est l'incertitude de localisation pour une représentation raster d'une couche géospatiale numérique?

Chaque couche de données raster a une résolution spatiale, qui correspond à la taille d'une cellule en unités de surface dans le monde réel. Par exemple, un raster avec une résolution spatiale de 10 m est composé de cellules de 10 m x 10 m, soit 100 m², dans le monde réel. Chaque couche de données raster a une résolution spatiale, qui correspond à la taille d'une cellule en unités de surface dans le monde réel. Par exemple, un raster avec une résolution spatiale de 10 m est composé de cellules de 10 m x 10 m, soit 100 m², dans le monde réel. En supposant que la cellule du raster est carrée, ce qui est le cas de la plupart d'entre elles, la longueur et la largeur de la cellule sont identiques. De plus, la coordonnée d'une cellule raster est référencée au centre de la cellule.



Nom de l'étudiant: _____ # d'étudiant: _____

Tout élément ou partie d'un élément qui a été encodé (transformé d'une représentation vectorielle en une représentation raster) dans une représentation raster peut avoir été au maximum à $\frac{1}{2}$ de la diagonale de la cellule du centre de la cellule pendant l'encodage, tout en restant à l'intérieur de la cellule. Dans ce cas, la cellule serait utilisée pour représenter l'élément ou une partie de celui-ci. Par conséquent, l'incertitude positionnelle générale concernant la position coordonnée d'un élément ou d'une partie d'élément, due à la représentation par une grille avec des cellules carrées, où chaque cellule a, par exemple, une superficie de 100 m^2 , peut être donnée par la formule suivante:

$$\begin{aligned}\delta_{raster} &= \frac{\sqrt{2A}}{2} \\ A &= l \times w \\ l &= 10m \\ w &= 10m \\ &= \frac{\sqrt{2(100)}}{2} \\ &= \frac{\sqrt{200}}{2} \\ &= \frac{14.14214}{2} \\ &= 7.071068\end{aligned}$$

L'incertitude de la position donnée par l'équation ci-dessus s'ajoute à toutes les autres incertitudes si votre couche raster a été créée à partir d'une couche géospatiale numérique vectorielle. Donc, si vous avez une couche raster de routes avec une résolution spatiale de 25 m, et que celles-ci proviennent d'une couche vectorielle de routes, qui, à son tour, a été **numérisée** à partir d'une carte papier à une échelle de 1:50 000, alors l'erreur de position totale serait donnée par:

Nom de l'étudiant: _____ # d'étudiant: _____

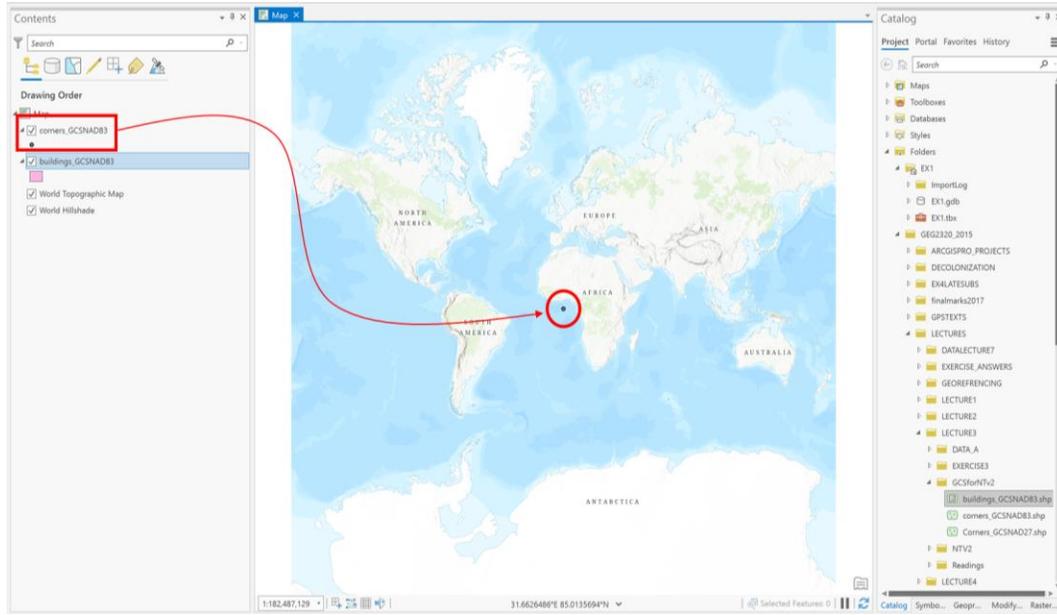
$$\begin{aligned} \text{taille de cellule} &= 25 \text{ m} \times 25 \text{ m} \\ \text{superficie de la cellule} &= 625 \text{ m}^2 \\ \text{erreur de raster} &= \frac{\sqrt{2(625)}}{2} = 17.68 \text{ m} \\ \therefore \text{total } \delta &= \sqrt{\text{erreur d'échelle de base}^2 + \text{erreur numérique}^2 + \text{erreur de raster}^2} \\ &= \sqrt{12500^2 + 12500^2 + 17680^2} \\ &= \sqrt{156250000 + 156250000 + 312500000} \\ &= \sqrt{625000000} \\ &= 25000 \\ &= 25 \text{ m} \end{aligned}$$

Nous incluons donc l'erreur de position du raster dans l'erreur totale. Par conséquent, par définition, les objets ou les entités des données raster ont une plus grande incertitude de localisation que leur représentation vectorielle correspondante - si le raster a été créé à partir de données vectorielles. Notez que si vous prenez un raster et que vous le convertissez en vecteur, la précision n'est pas améliorée puisque l'erreur de position du raster est toujours présente.

Q7: Si ArcGIS Pro ne reconnaît pas le système de coordonnées d'une couche, comment puis-je indiquer à ArcGIS le système de coordonnées de la couche?

Lorsqu'une couche géospatiale ne contient pas d'informations sur le système de coordonnées et qu'elle est ajoutée à ArcGIS Pro, un avertissement peut s'afficher et, souvent, la couche de données spatiales n'apparaît pas à l'endroit prévu. La figure ci-dessous représente une situation typique, où les données se sont retrouvées près de 0° latitude et 0° longitude. Si vous ajoutez des données à votre projet et que vous ne les trouvez pas, il s'agit souvent d'un bon endroit pour les rechercher!

Nom de l'étudiant: _____ # d'étudiant: _____



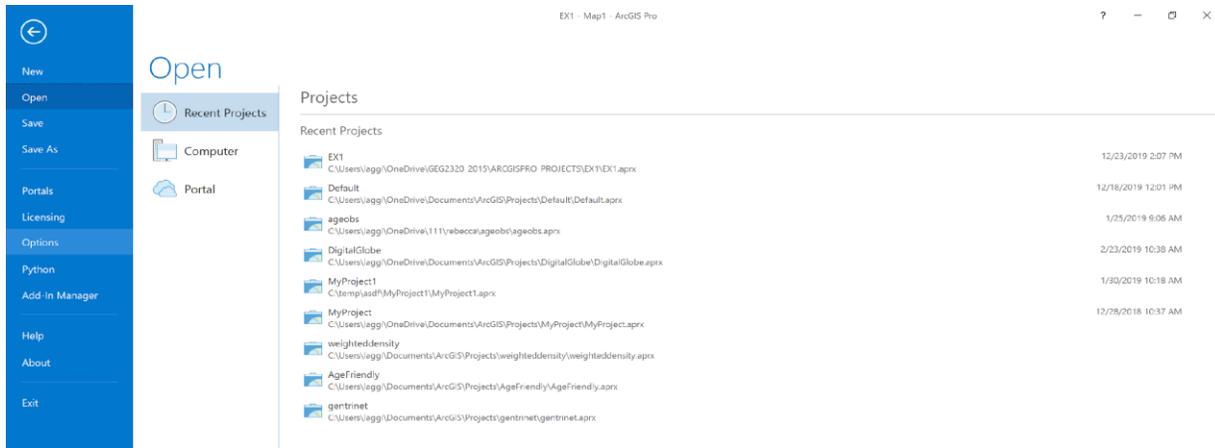
ArcGIS Pro affichera toujours la couche, mais elle ne s'alignera pas sur les autres couches dont les systèmes de coordonnées sont connus.

Dans le dossier GCSforNTV2, vous avez les couches suivantes, avec les systèmes de coordonnées suivants:

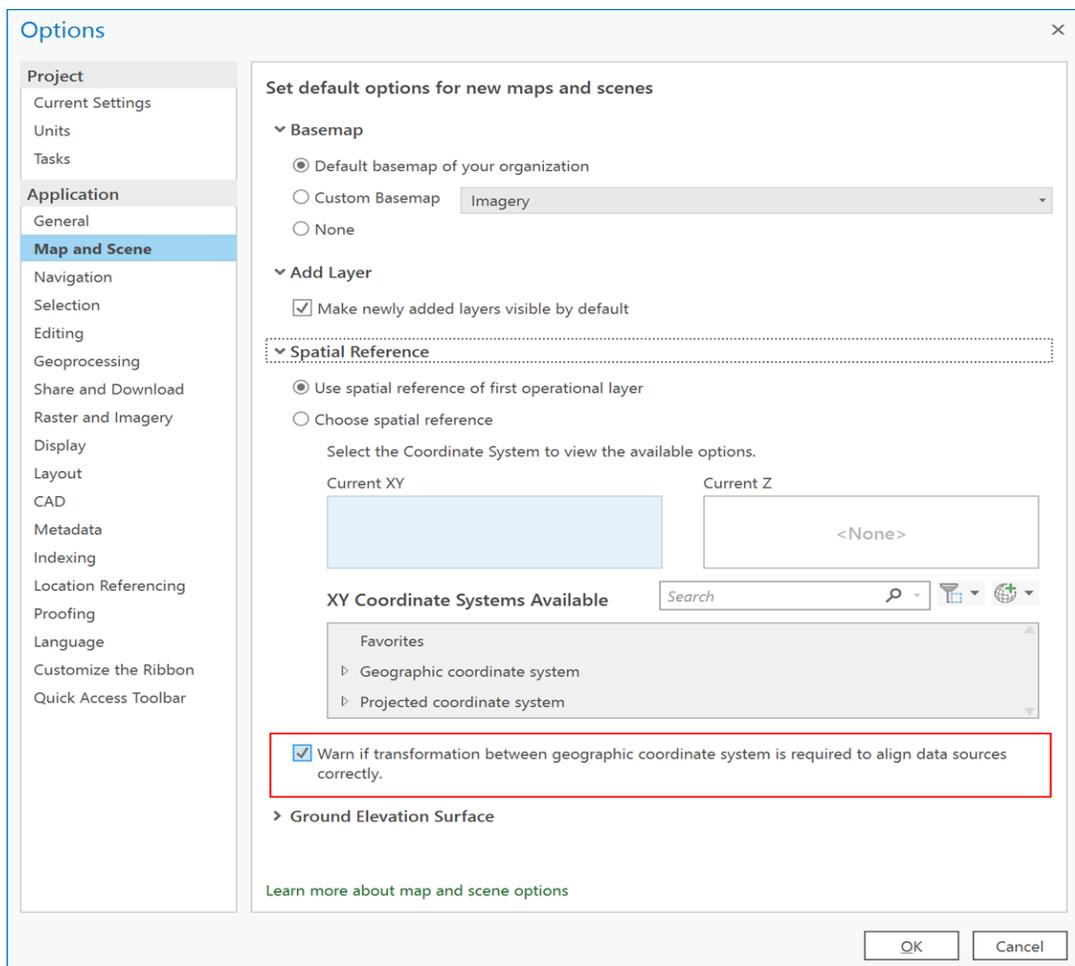
<i>Layer</i>	<i>Coordinate System</i>
corners_GCSNAD83	North American Datum of 1983
corners_GCSNAD27	North American Datum of 1927
buildings_GCSNAD83	North American Datum of 1983

1. Ajouter la couche corners_GCSNAD83 à ArcGIS Pro dans un nouveau projet.
2. Cliquez sur le bouton « *Full Extent* » dans l'onglet Map et vous verrez un point au centre de la carte près de 0,0 dans les coordonnées cartésiennes du système de coordonnées par défaut "WGS 1984 Web Mercator Auxiliary Sphere" d'ArcGIS Pro, comme dans la figure ci-dessus. Le petit point au milieu de la carte contient des points qui représentent les coins du bâtiment Simard Hall, à Ottawa ! Afin de voir ces points au bon endroit, nous devons définir le système de coordonnées de la couche et toutes les transformations géographiques nécessaires pour amener les coordonnées de la couche aux coordonnées de la carte afin que les choses s'alignent correctement. Pour cela, suivez les instructions suivantes.
3. Allez dans Project -> Options:

Nom de l'étudiant: _____ # d'étudiant: _____

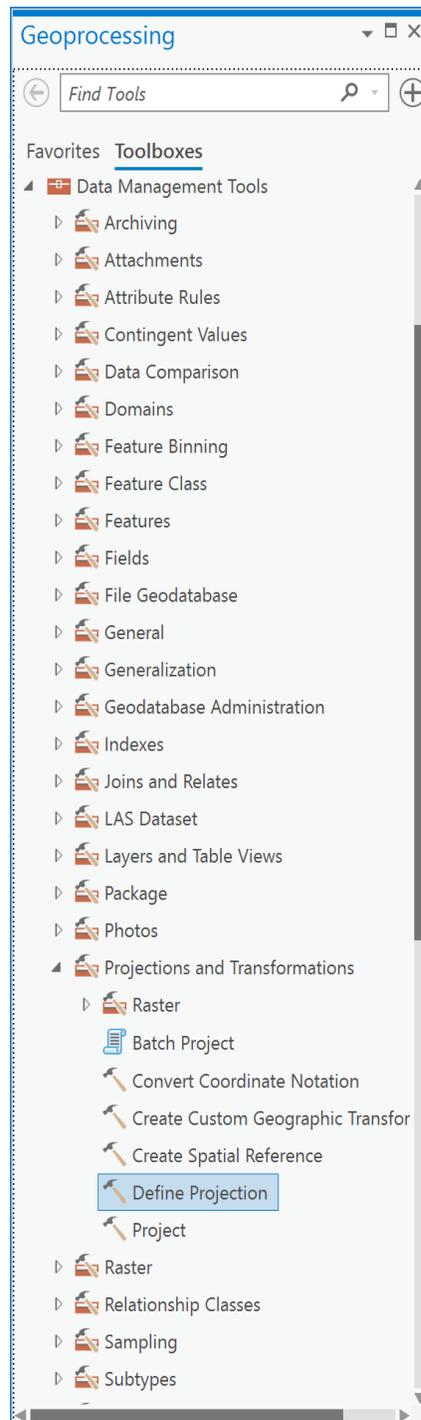


4. Dans la boîte de dialogue Options, choisissez l'élément « *Map and Scene* », faites défiler vers le bas et cochez la case "*Warn if transformation between geographic coordinate systems is required to align data sources correctly*" (Avertir si une transformation entre les systèmes de coordonnées géographiques est nécessaire pour aligner correctement les sources de données).



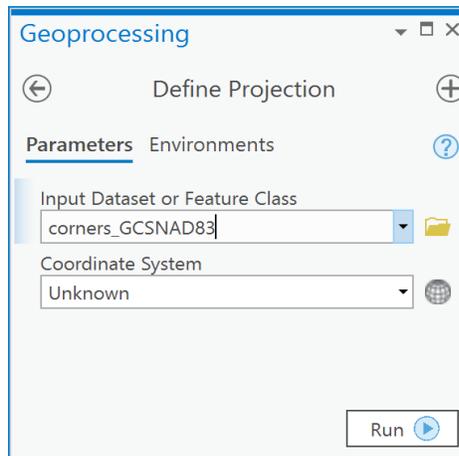
Nom de l'étudiant: _____ # d'étudiant: _____

5. Retournez à votre carte.
6. Allez dans le volet "Geoprocessing" sous "Toolboxes".
7. Pour spécifier le système de coordonnées des couches dans ArcGIS Pro, vous utilisez l'outil « *Define Projection* » (Définir la projection) sous « *Data Management* » (Gestion des données) -> « *Projections and Transformations* » (Projections et transformations):

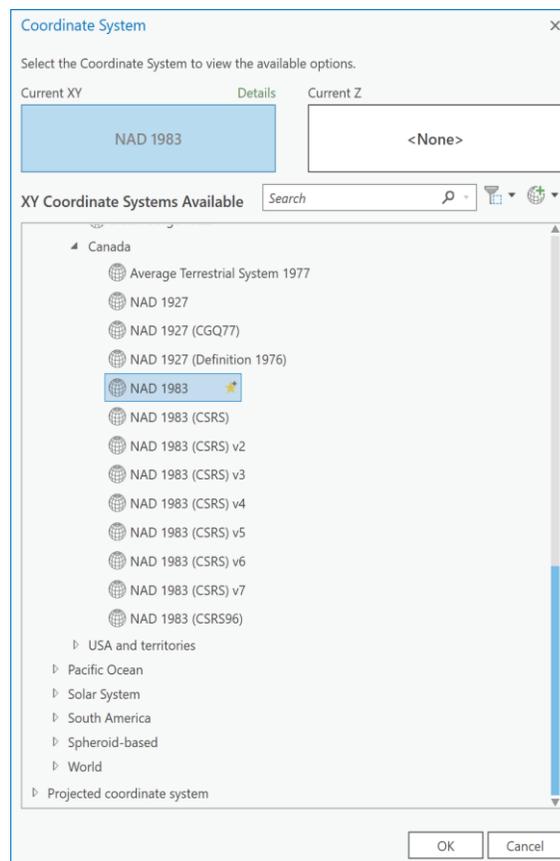


Nom de l'étudiant: _____ # d'étudiant: _____

- Dans l'outil « *Define Projection* » (Définir la projection), choisissez la couche **corners_GCSNAD83** comme "*Input Dataset or Feature Class*" (Jeu de données d'entrée ou classe d'entités):

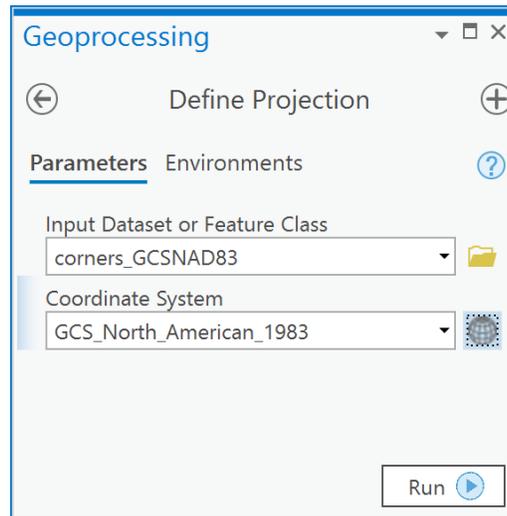


- A droite de la zone de saisie « *Coordinate System* », cliquez sur le bouton qui ressemble à un globe terrestre.
- Maintenant, naviguez vers le dossier « *Geographic Coordinate Systems* » (Systèmes de coordonnées géographiques) -> *North America* (Amérique du Nord) -> *Canada* et choisissez "**NAD 1983**" et cliquez sur OK.

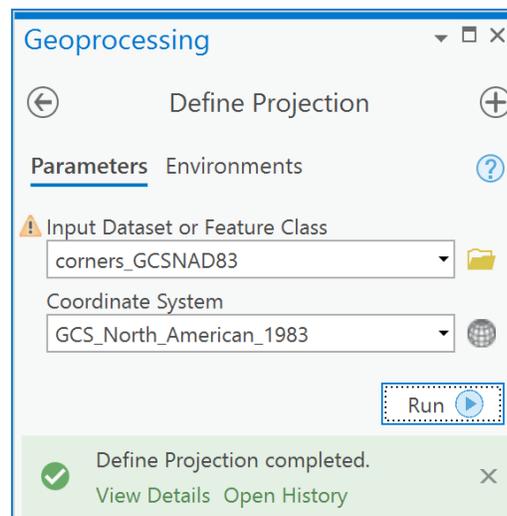


Nom de l'étudiant: _____ # d'étudiant: _____

11. Votre boîte de dialogue doit maintenant ressembler à ceci, si c'est le cas, cliquez sur Run:



12. Une fois l'opération terminée, l'écran suivant s'affiche:



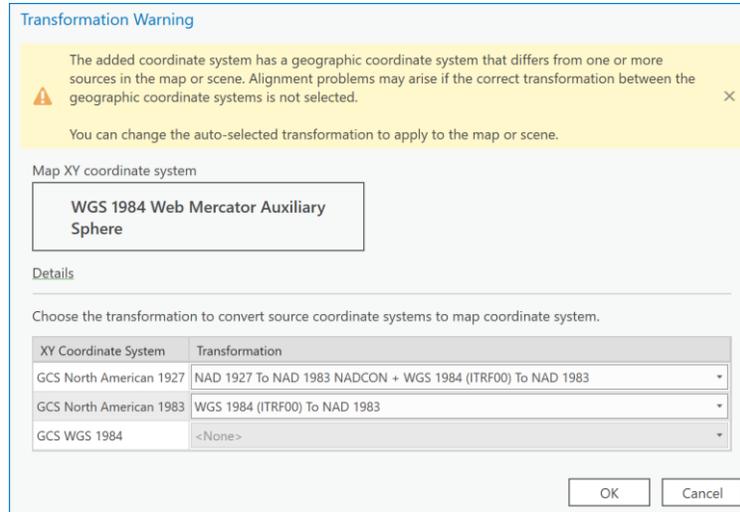
13. Cela signifie que le système de coordonnées de votre couche est défini comme étant le NAD83. La définition de la projection indique simplement à ArcGIS quel est le système de coordonnées géographiques (SCG) de la couche. L'étudiant observateur se rendra compte que sans les informations du tableau 1, il n'y aurait aucun moyen de savoir avec certitude quel est le SCG d'origine et, par conséquent, il n'y aurait aucun moyen d'utiliser en toute confiance l'outil « Define Projection » (Définir la projection).

14. Répétez maintenant cette procédure pour les deux autres couches énumérées dans le tableau, en notant que **corners_GCSNAD27** est dans le système de référence nord-américain de 1927.

Nom de l'étudiant: _____ # d'étudiant: _____

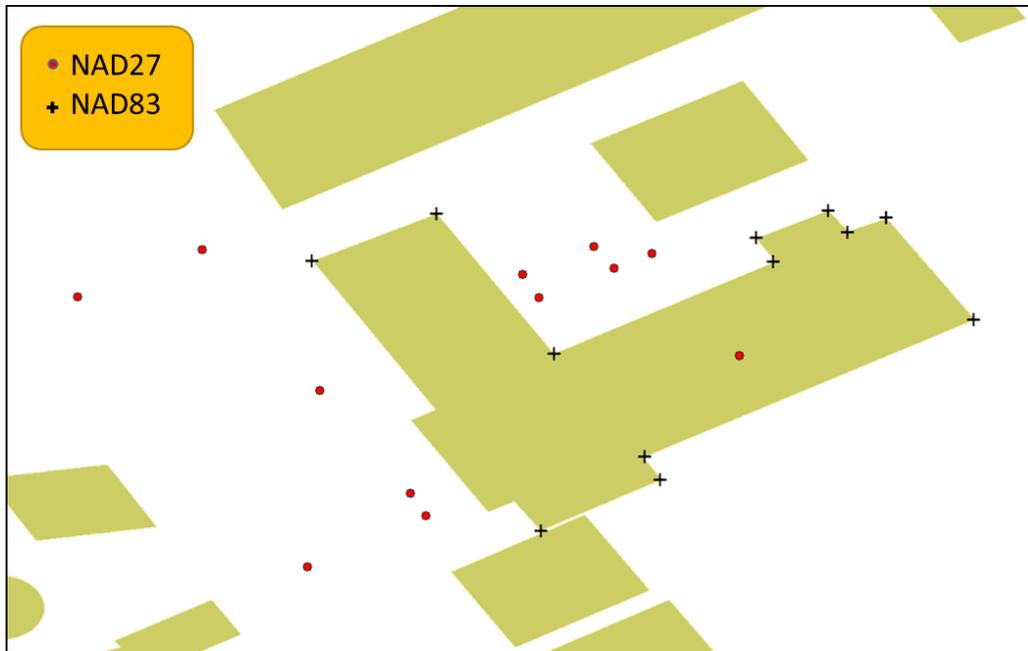
Q8: Comment convertir entre NAD27 et NAD83 au Canada?

Si vous ajoutez une couche à ArcGIS Pro qui contient un système de coordonnées géographiques différent de celui défini pour la carte, l'avertissement suivant s'affiche:

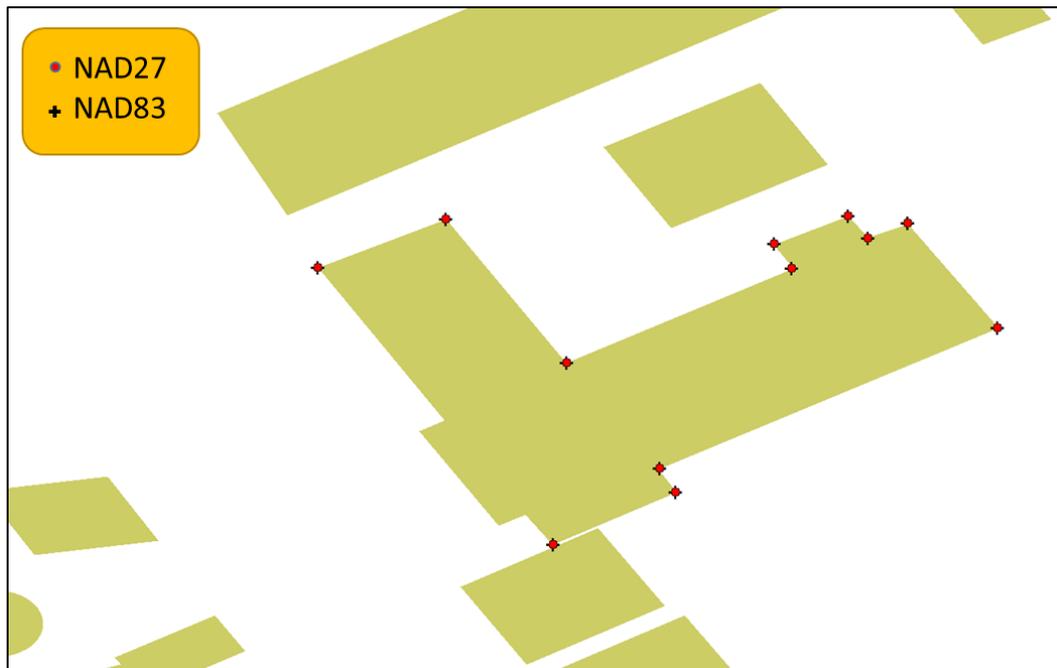


Cet avertissement signifie que vous essayez d'afficher deux systèmes de coordonnées géographiques différents dans le même espace. ArcGIS Pro vous permettra de le faire si vous cliquez sur le bouton "Annuler". Cependant, cela signifie que les emplacements d'objets ou d'entités communs ne seront pas alignés entre la couche que vous ajoutez et d'autres couches déjà présentes dans votre cadre de données. Par exemple, dans la figure ci-dessous, les angles des bâtiments de Simard Hall sont représentés par des "+" noirs dans le système de coordonnées NAD83 et par des points rouges dans le système de coordonnées NAD27. Les deux sont décalés car les données représentées en points rouges n'ont pas été transformées correctement dans le système de coordonnées NAD83.

Nom de l'étudiant: _____ # d'étudiant: _____



Toutefois, si vous spécifiez la transformation géographique correcte pour les couches, les données s'aligneront correctement, comme le montre l'illustration ci-dessous.

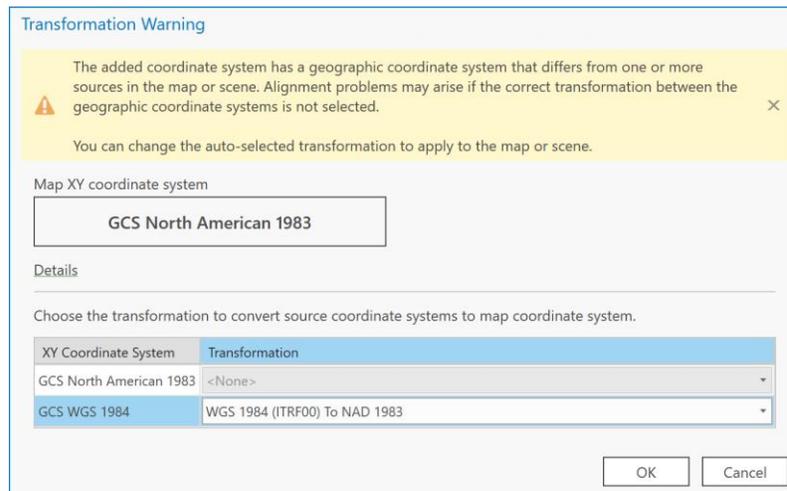


Pour voir ces effets en action,

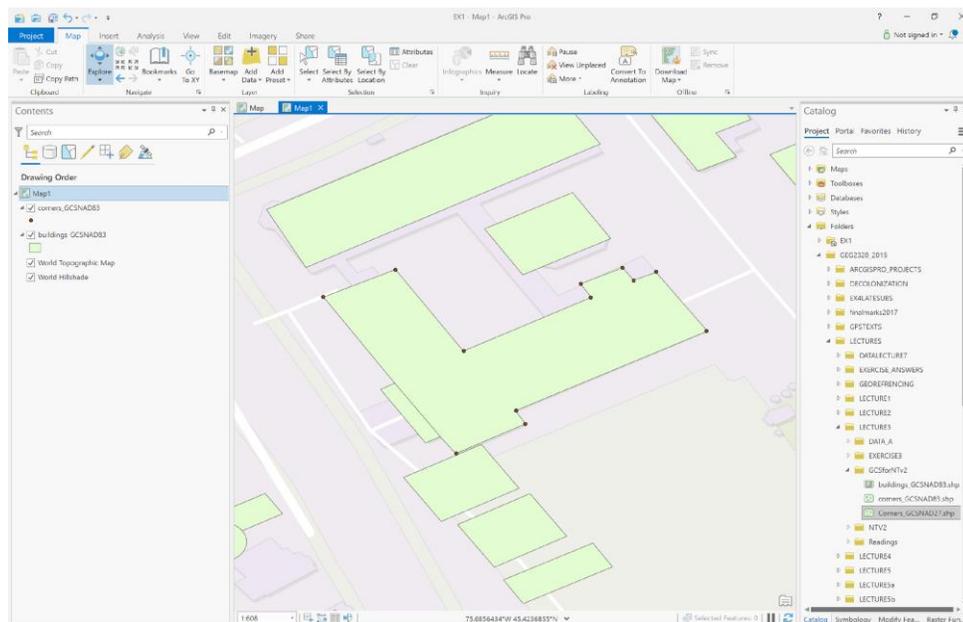
1. Cliquez sur l'onglet "*Insert*" puis sur le bouton "**New Map**" pour créer une nouvelle carte.
2. Maintenant, à partir du dossier GCSforNTv2, ajoutez la couche appelée buildings_GCSNAD83. Vous verrez la boîte de dialogue d'avertissement, cliquez simplement sur OK. Vous obtenez cet avertissement parce que le système de

Nom de l'étudiant: _____ # d'étudiant: _____

longitude/latitude par défaut appelé système de coordonnées géographiques (SCG) pour la carte de base dans ArcGIS Pro est WGS 1984 (utilisant l'ellipsoïde WGS 84 pour les coordonnées horizontales), alors que la couche que vous avez ajoutée est dans le système de référence nord-américain de 1983 (utilisant l'ellipsoïde GRS 1980 pour les coordonnées horizontales). Ces deux systèmes utilisent des modèles ellipsoïdaux différents pour leur système de coordonnées 3D long/lat. Cliquez sur OK.



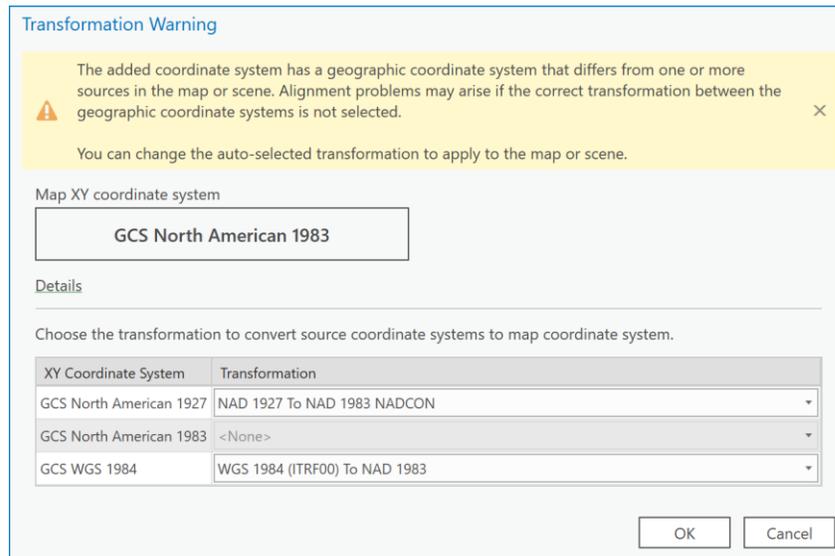
3. Ajoutez maintenant corners_GCSNAD83. Il n'y aura pas de dialogue d'avertissement cette fois-ci.
4. Zoomer sur l'étendue de la couche corners_GCSNAD83. Vous verrez que la couche de points appelée corners_GCSNAD83 s'aligne sur les coins du bâtiment Simard Hall:



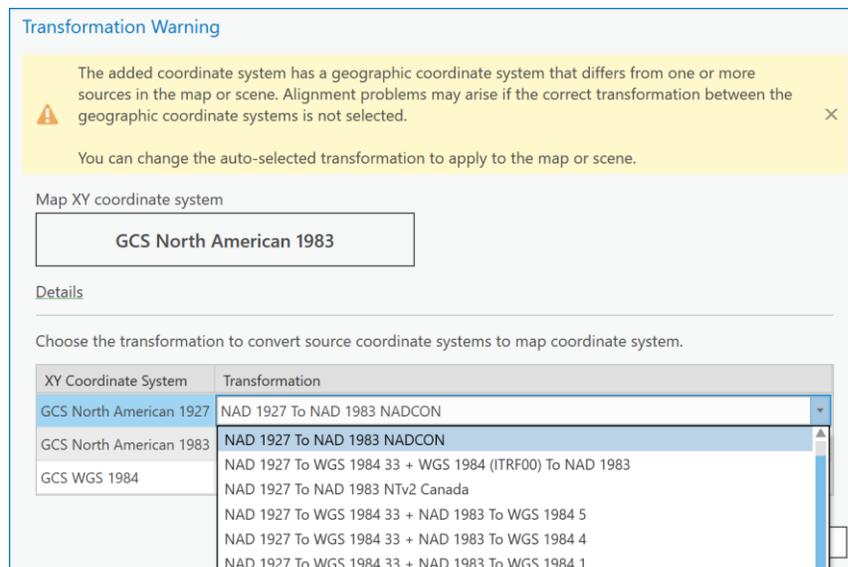
5. Ajoutez ensuite la couche Corners_GCSNAD27. Vous verrez une boîte de dialogue d'avertissement qui spécifie que le SCG de la couche ajoutée, soit Corners_GCSNAD27, est différent du SCG de la carte. En effet, le système de

Nom de l'étudiant: _____ # d'étudiant: _____

coordonnées géographiques de la couche Corners_GCSNAD27 correspond au Système de référence nord-américain de 1927 (NAD 27), alors que le système de coordonnées de la carte a été défini par le système de coordonnées de la première couche ajoutée, la couche Buildings_GCSNAD83, qui correspond au Système de référence nord-américain de 1983 (NAD 83). Cette boîte de dialogue vous demande donc comment transformer le NAD 27 en NAD 83:

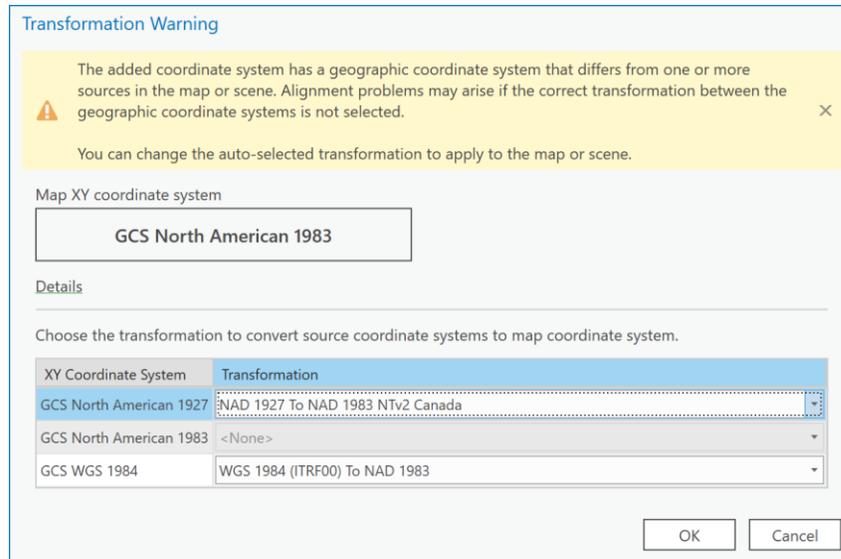


6. Cliquez sur la liste déroulante à côté de "GCS North American 1927":

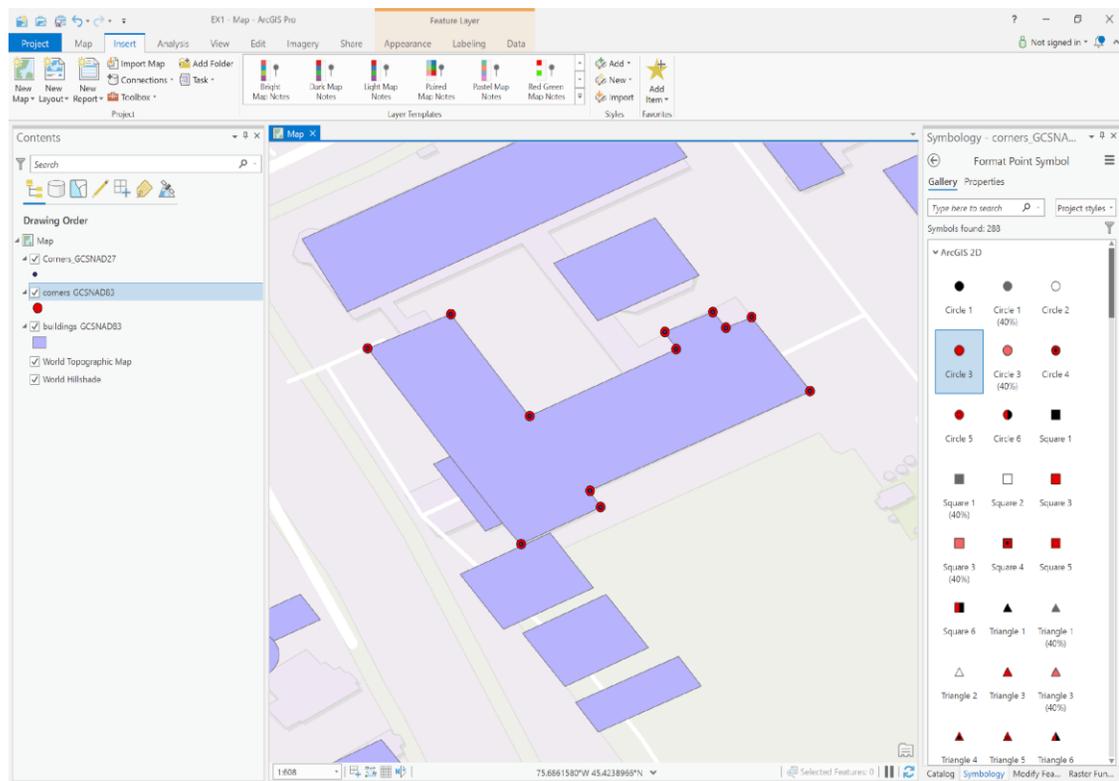


7. Choisir la transformation "NAD 1927 to NAD 1983 NTV2 Canada".

Nom de l'étudiant: _____ # d'étudiant: _____



8. Vous voyez maintenant que toutes les données s'alignent:



Notez également que ce processus ne donne qu'une transformation d'affichage ou une transformation à la volée ; les données sous-jacentes dans les couches géospatiales numériques ne changent pas, ArcGIS Pro a simplement calculé de nouvelles coordonnées de sorte que toutes les entités dans Corners_GCSNAD27 sont affichées en coordonnées NAD83.

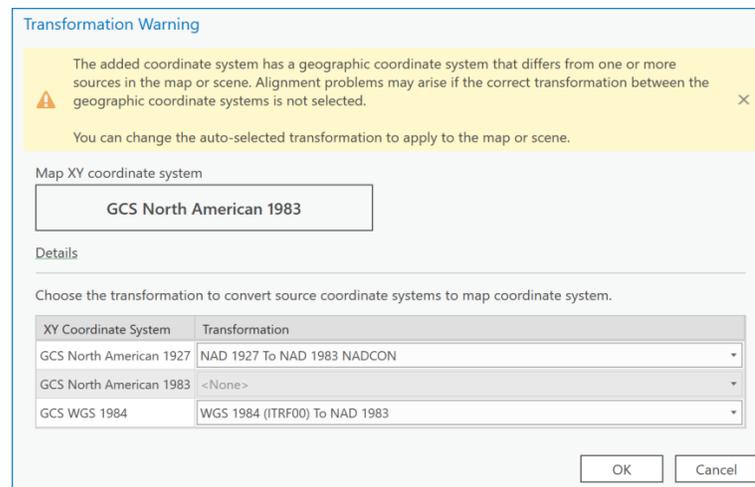
Nom de l'étudiant: _____ # d'étudiant: _____

Les entités enregistrées dans Corners_GCSNAD27 sont toujours enregistrées en NAD27, cette technique ne les modifie pas.

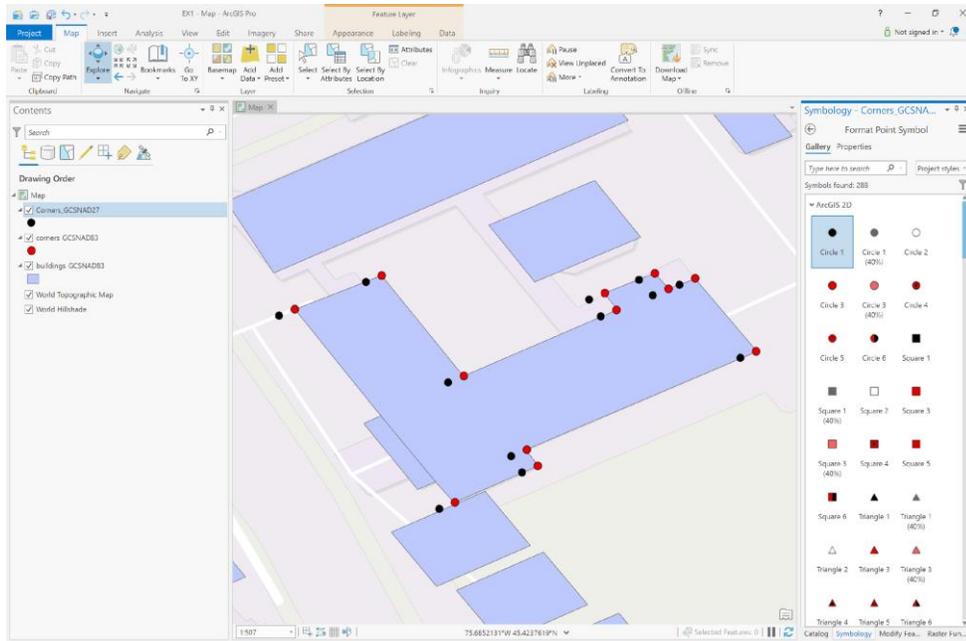
Q9: Comment puis-je spécifier la conversion NAD27 et NAD83 après avoir ajouté les données à ArcGIS Pro?

Vous pouvez spécifier n'importe quelle transformation pour les couches d'une carte dans ArcGIS Pro en ouvrant la boîte de dialogue "Map properties" et en cliquant sur l'élément "Transformations...".

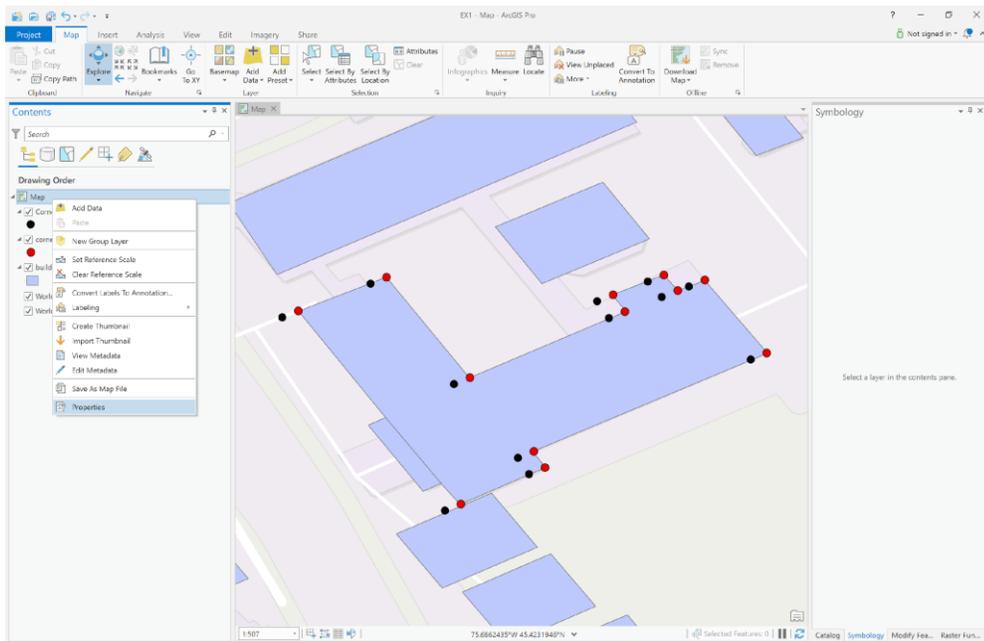
1. Répéter les étapes 1 à 4 de l'exemple Q7.
2. La boîte de dialogue suivante s'affiche, cliquez sur "OK"
3. Votre carte ressemblera à ce qui suit. Lorsque vous avez cliqué sur OK, vous avez automatiquement choisi de convertir le NAD27 GCS en NAD83 en utilisant la transformation NADCON. Cette transformation n'est valable que pour les 48 états américains inférieurs. Le NADCON n'est pas défini pour le Canada. Bien que NADCON soit mieux que rien, vous pouvez voir que la couche 'Corners_GCSNAD27' n'est pas alignée avec les deux autres couches. Nous devons donc spécifier la transformation correcte pour le Canada, qui est la transformation NTV2. Ceci est facile à corriger dans l'étape suivante.



Nom de l'étudiant: _____ # d'étudiant: _____

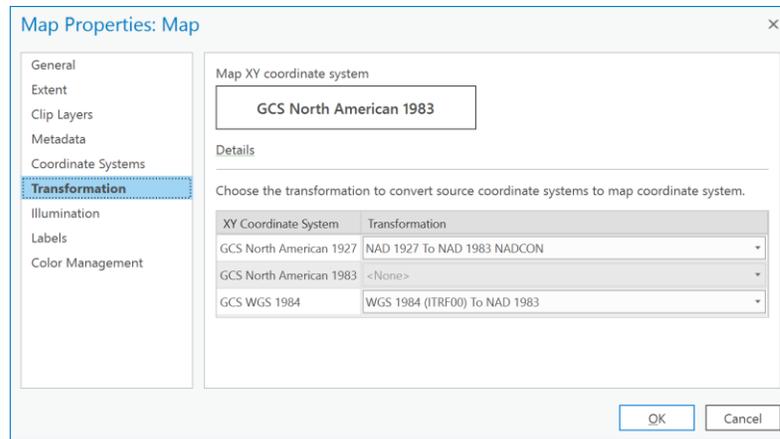


4. Cliquez avec le bouton droit de la souris sur la carte et choisissez « *Propriétés* »

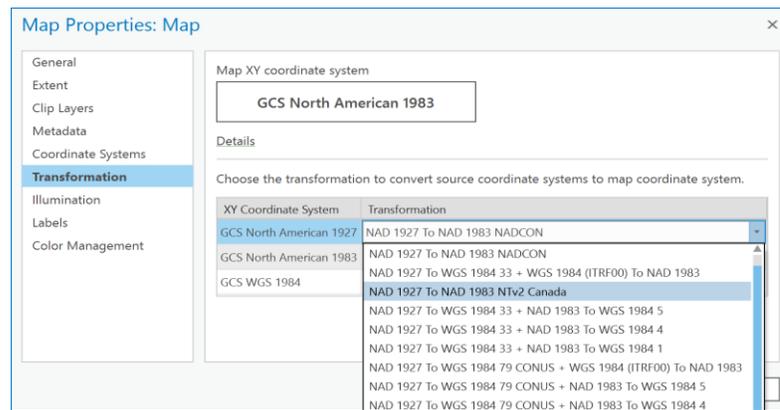


5. Dans la boîte de dialogue Propriétés de la carte, choisissez l'élément *Transformation*:

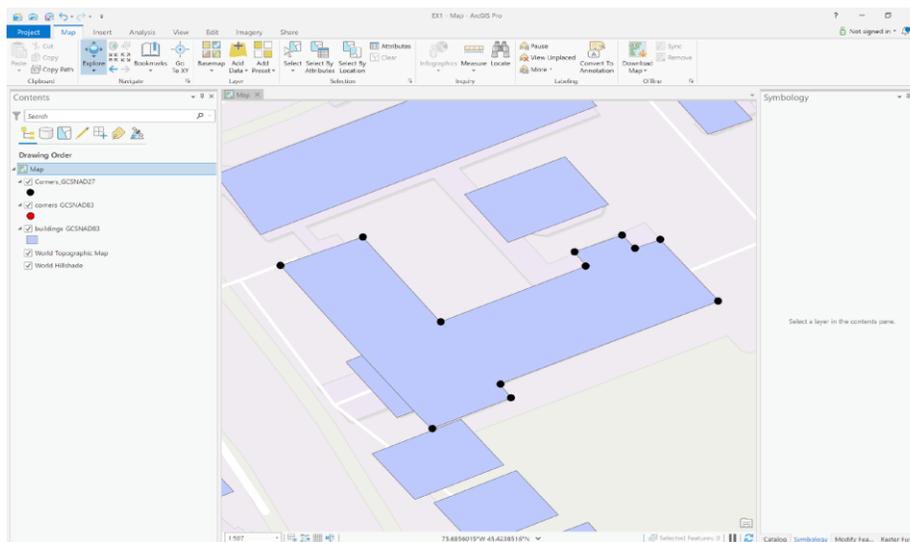
Nom de l'étudiant: _____ # d'étudiant: _____



6. Cliquez sur le menu déroulant à côté de "GCS North American 1927" et choisissez la transformation "NAD 1927 to NAD 1983 NTv2 Canada", puis cliquez sur OK:



7. Maintenant, tout est aligné:



Nom de l'étudiant: _____ # d'étudiant: _____

Q10: Comment obtenir les coordonnées NAD27 et NAD83 pour chaque point dans un ensemble de points?

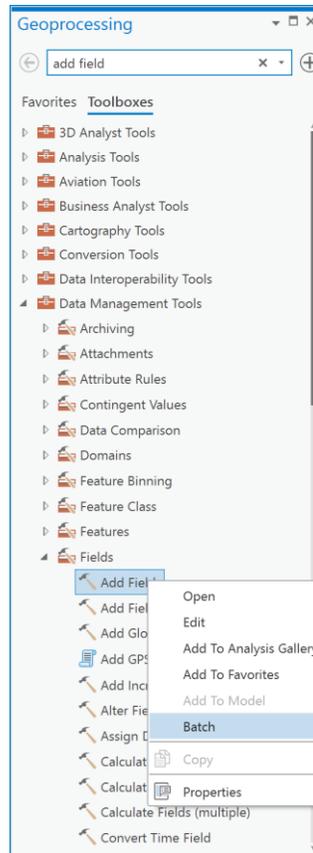
Pour obtenir les coordonnées d'un ensemble de points dans un système de coordonnées prédéfini, il existe trois méthodes:

1. Utilisez l'outil d'identification (« *identify tool* »), cliquez sur un point et examinez les coordonnées.
2. Utilisez l'outil « *Add XY Coordinates* » (Ajouter des coordonnées XY) dans « *Data Management Tools* » -> « *Features* » (le plus simple).
3. Utilisez l'outil « *Add Field* » dans la boîte à outils « *Geoprocessing Fields* » et calculez la géométrie dans ces champs dans le système de coordonnées original de la (des) couche(s).

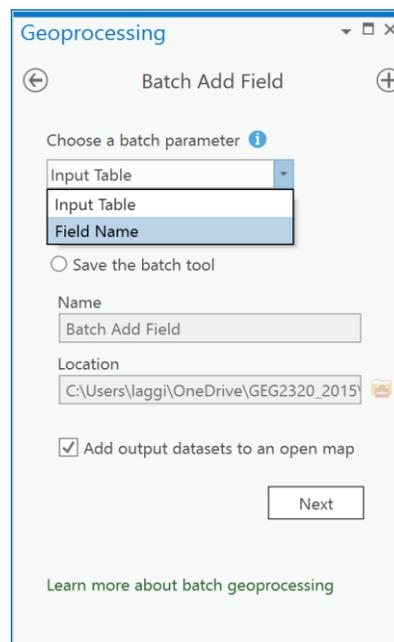
Les deuxième et troisième approches vous permettent d'acquérir les coordonnées de plus d'un point à la fois avec une grande précision, de sorte que vous utiliserez probablement l'une ou l'autre de ces méthodes la plupart du temps. Pour développer des compétences SIG supplémentaires, cet exemple utilisera la troisième méthode, qui prend un peu plus de temps mais utilise davantage d'outils utiles dans d'autres contextes.

1. Créer une nouvelle carte.
2. Ajouter la couche *Corners_GCSNAD27* à ArcGIS Pro. Cliquez sur OK dans toutes les boîtes de dialogue qui s'affichent. Il ne s'agit pas ici d'alignement mais plutôt d'ajout de coordonnées dans le système de coordonnées original, la transformation que nous choisissons n'a donc pas d'importance.
3. Allez dans « *Data Management Tools* » -> « *Fields* » et cliquez avec le bouton droit de la souris sur l'outil « *Add Field* » et choisissez « *Batch* »

Nom de l'étudiant: _____ # d'étudiant: _____

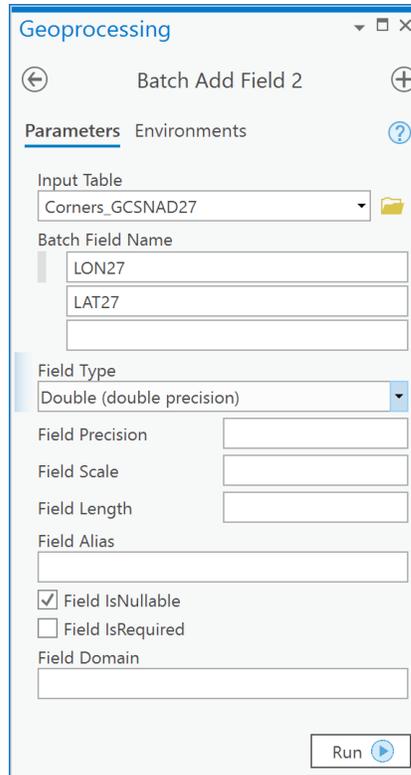


4. La boîte de dialogue de traitement par lots s'affiche. Dans la liste déroulante « Choose a batch parameter », choisissez « *Field Name* » et cliquez sur Next:



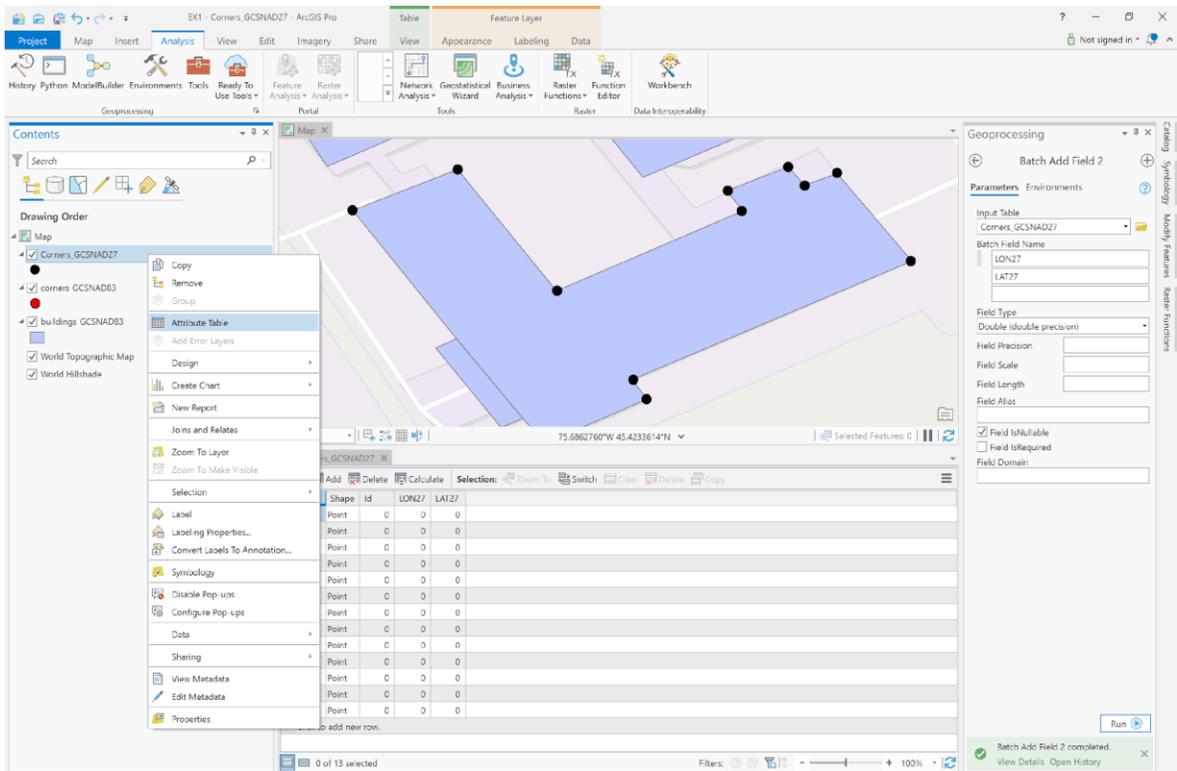
Nom de l'étudiant: _____ # d'étudiant: _____

5. La boîte de dialogue Paramètres s'affiche. Dans cette boîte de dialogue, choisissez "Corners_GCSNAD27" dans la liste déroulante sous "Input table". Sous "Batch Field Name", saisissez LON27 et LAT27 comme vous le voyez ci-dessous. Enfin, sous "Field type" (type de champ), choisissez "Double" et cliquez sur Run (exécuter):

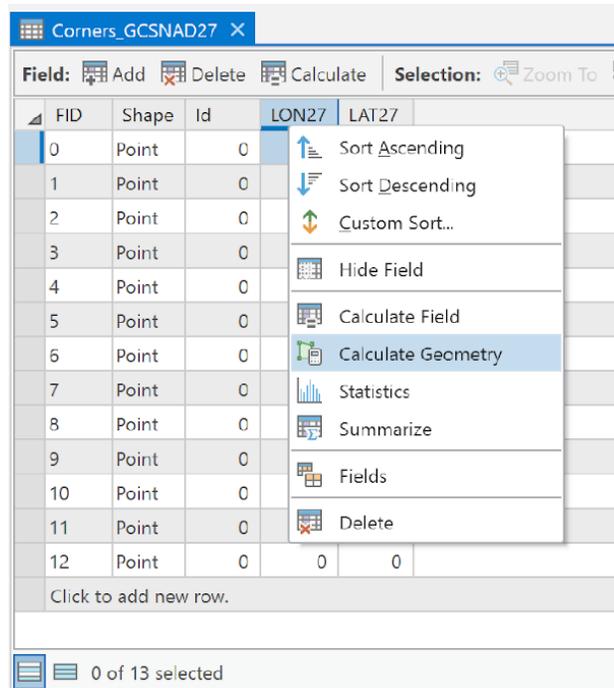


6. Faites maintenant un clic droit sur la couche 'Corners_GCSNAD27' et choisissez « *Attribute Table* ». Vous verrez qu'il y a maintenant deux nouveaux attributs, l'un appelé LON27 et l'autre appelé LAT27.

Nom de l'étudiant: _____ # d'étudiant: _____



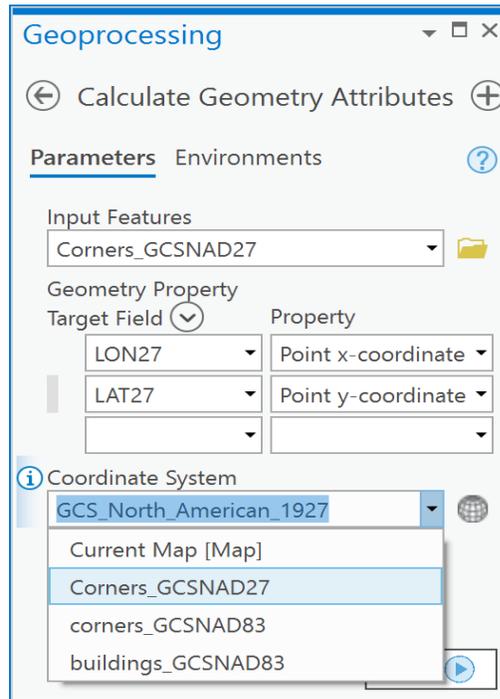
7. Cliquez avec le bouton droit de la souris sur l'attribut LON27 et choisissez « *Calculate geometry* »:



8. Sous « *Geometry Property-Target Field* » (Propriété de la géométrie - Champ cible), sélectionnez LON27 puis LAT27. À côté de LON27, dans le menu déroulant des

Nom de l'étudiant: _____ # d'étudiant: _____

propriétés, choisissez « *Point x-coordinate* » et à côté de LAT27, choisissez « *Point y-coordinate* ». Sous « *Coordinate System* » (système de coordonnées), choisissez *Corners_GCSNAD27* et le système de coordonnées spécifiera *GCS_North_American_1927* comme système de coordonnées pour calculer les valeurs GCS. Votre boîte de dialogue doit ressembler à l'exemple ci-dessous, si c'est le cas, cliquez sur Run:



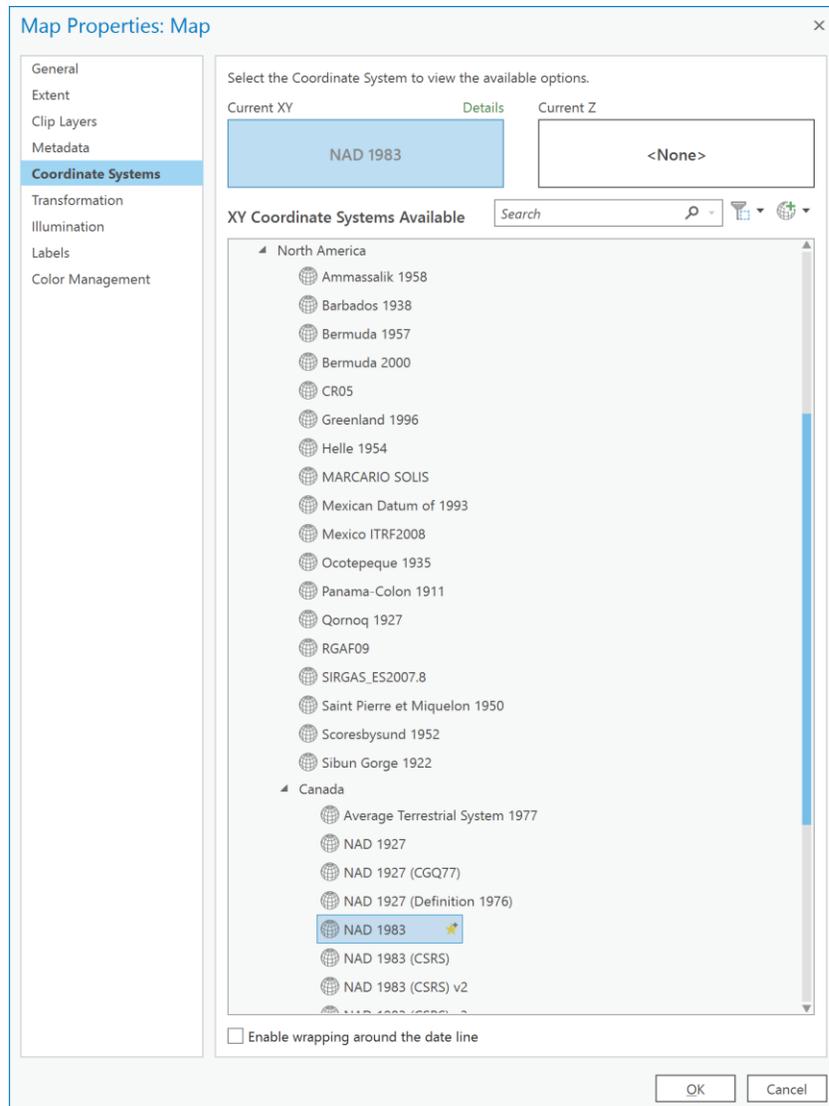
9. Vous devriez maintenant voir les éléments suivants dans le volet Attributs:

The screenshot shows the Attribute Table for 'Corners_GCSNAD27' in ArcGIS Pro. The table has columns for FID, Shape, Id, LON27, and LAT27. The data shows 13 rows of point features with their respective coordinates.

FID	Shape	Id	LON27	LAT27
0	Point	0	-75.685685	45.423514
1	Point	0	-75.685685	45.423514
2	Point	0	-75.685562	45.423371
3	Point	0	-75.68649	45.423454
4	Point	0	-75.686315	45.42352
5	Point	0	-75.686168	45.423075
6	Point	0	-75.686001	45.423147
7	Point	0	-75.68615	45.423323
8	Point	0	-75.685843	45.423452
9	Point	0	-75.685866	45.423485
10	Point	0	-75.685766	45.423524
11	Point	0	-75.685738	45.423494
12	Point	0	-75.686023	45.423178

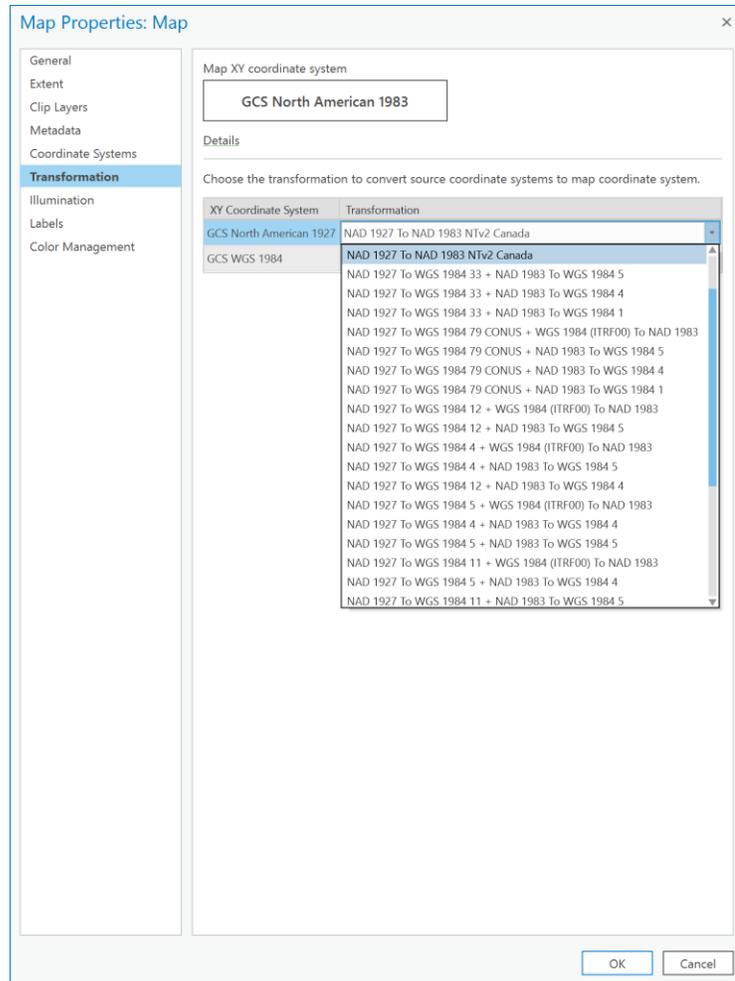
Nom de l'étudiant: _____ # d'étudiant: _____

10. Vous avez maintenant calculé les coordonnées de chaque coin de bâtiment en longitude et latitude NAD 27.
11. Changez maintenant le système de coordonnées de la carte affichée en NAD 83 en cliquant avec le bouton droit de la souris sur la carte dans la table des matières et en choisissant « *Propriétés* » (Propriétés). Dans la boîte de dialogue Propriétés, choisissez l'élément « *Coordinate System* » (Système de coordonnées). Choisissez ensuite « *Geographic Coordinate System* » (Système de coordonnées géographiques) -> « *North America* » (Amérique du Nord) -> Canada -> NAD 1983:



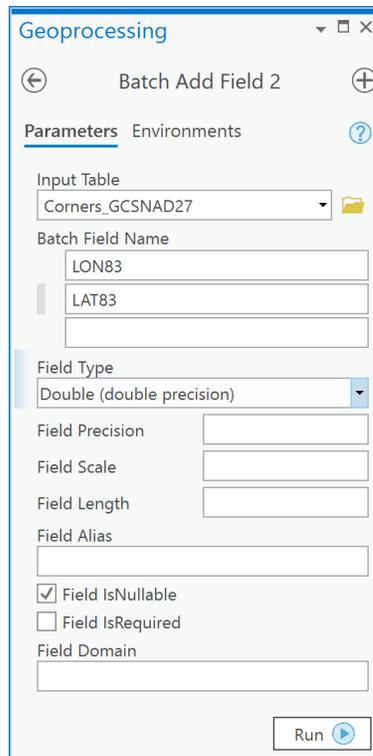
12. Passez maintenant à l'élément Transformation et choisissez "NAD 1927 to NAD 1983 NTV2 Canada" et cliquez sur click OK.

Nom de l'étudiant: _____ # d'étudiant: _____

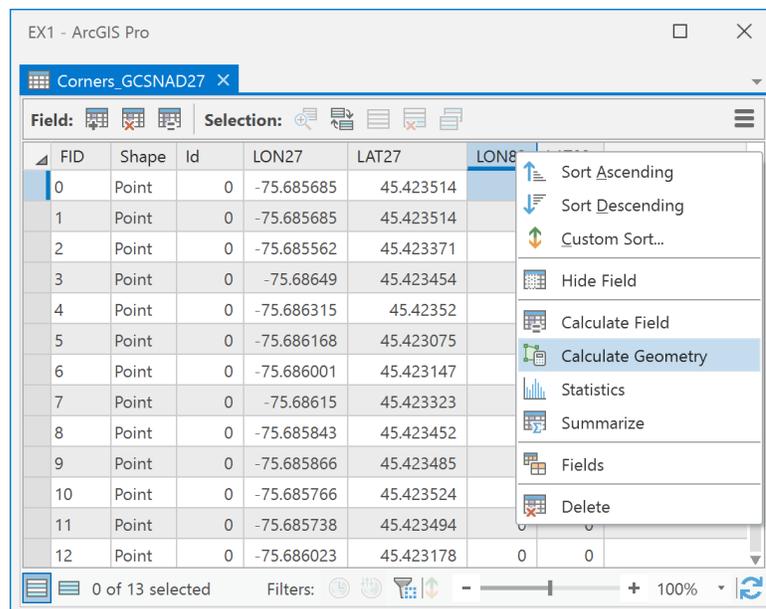


13. Répétez ensuite les étapes 3 et 4. La boîte de dialogue *Parameters* s'affiche. Dans cette boîte de dialogue, choisissez 'Corners_GCSNAD27' dans la liste déroulante sous « *Input table* ». Sous « *Batch Field Name* », saisissez LON83 et LAT83 comme indiqué ci-dessous. Enfin, sous « *Field type* », choisissez 'Double' et cliquez sur Run:

Nom de l'étudiant: _____ # d'étudiant: _____

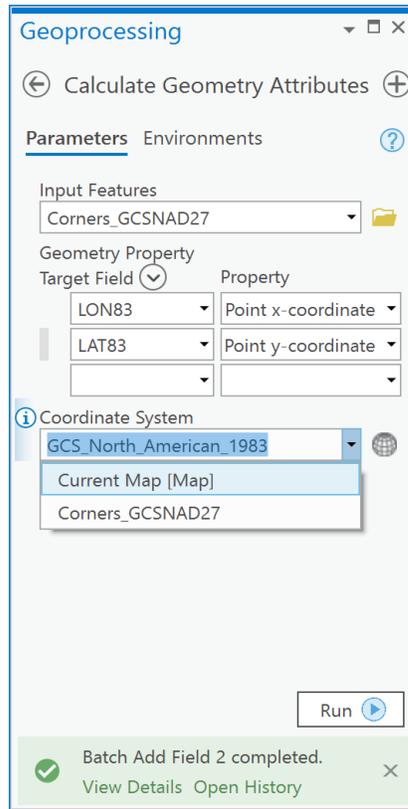


14. Ensuite, dans le volet Attributs, cliquez avec le bouton droit de la souris sur le champ LON83 et choisissez « *Calculate Geometry* » (Calculer la géométrie):



15. Dans la boîte de dialogue « *Calculate Geometry Attributes* » (Calculer les attributs géométriques), remplissez les paramètres suivants pour calculer les coordonnées en NAD83. Cliquez sur Run (Exécuter) une fois que votre boîte de dialogue ressemble à l'illustration ci-dessous:

Nom de l'étudiant: _____ # d'étudiant: _____



16. Vous constaterez maintenant que les coordonnées NAD27 et les coordonnées NAD83 diffèrent dans les 3 à 5èmes décimales:

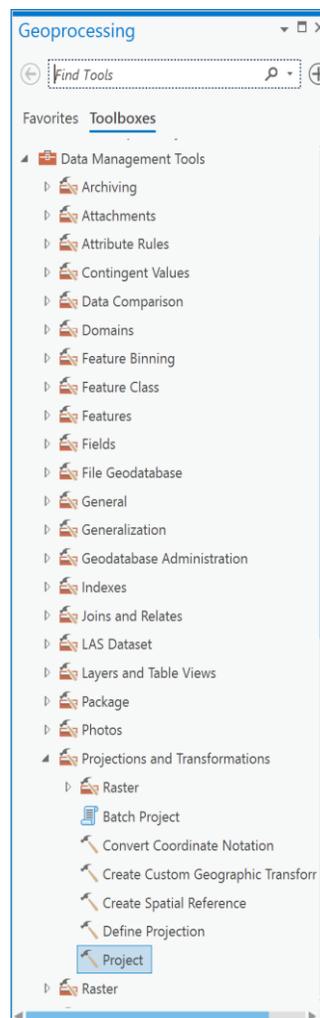
FID	Shape	Id	LON27	LAT27	LON83	LAT83
0	Point	0	-75.685685	45.423514	-75.685356	45.423566
1	Point	0	-75.685685	45.423514	-75.685356	45.423566
2	Point	0	-75.685562	45.423371	-75.685234	45.423423
3	Point	0	-75.68649	45.423454	-75.686162	45.423505
4	Point	0	-75.686315	45.42352	-75.685987	45.423571
5	Point	0	-75.686168	45.423075	-75.68584	45.423127
6	Point	0	-75.686001	45.423147	-75.685673	45.423198
7	Point	0	-75.68615	45.423323	-75.685822	45.423374
8	Point	0	-75.685843	45.423452	-75.685515	45.423504
9	Point	0	-75.685866	45.423485	-75.685538	45.423537
10	Point	0	-75.685766	45.423524	-75.685438	45.423575
11	Point	0	-75.685738	45.423494	-75.68541	45.423545
12	Point	0	-75.686023	45.423178	-75.685695	45.42323

Nom de l'étudiant: _____ # d'étudiant: _____

Q11: Comment effectuer un changement permanent en NAD83 (ou tout autre système de coordonnées) pour une couche géospatiale numérique?

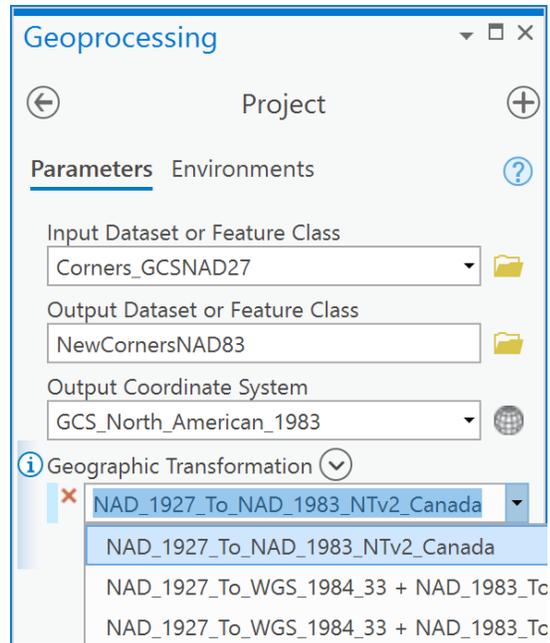
Comme indiqué précédemment, une transformation d'affichage n'est pas une modification permanente des coordonnées enregistrées dans un fichier. Pour modifier les coordonnées de manière permanente, il existe deux méthodes:

1. Exporter la couche actuelle de la table des matières à l'aide de l'outil de géotraitement « *Feature Class to Feature Class* » (également sous « *Conversion Tools* » -> « *To Geodatabase* ») et spécifier la transformation géographique dans les paramètres de l'environnement de l'outil « *Feature Class to Feature Class* ».
2. Utilisez l'outil de la boîte à outils appelé « *Project* » sous « *Data Management* » (Gestion des données) -> Projections et transformations. Cette méthode est souvent préférée car elle est spécialement conçue pour changer les systèmes de coordonnées des données spatiales vectorielles.
3. Dans le volet « *Geoprocessing* » (Géotraitement), ouvrez l'outil « *Data Management Tools* » (Outils de gestion des données) -> Projections et transformations -> « *Project* »:



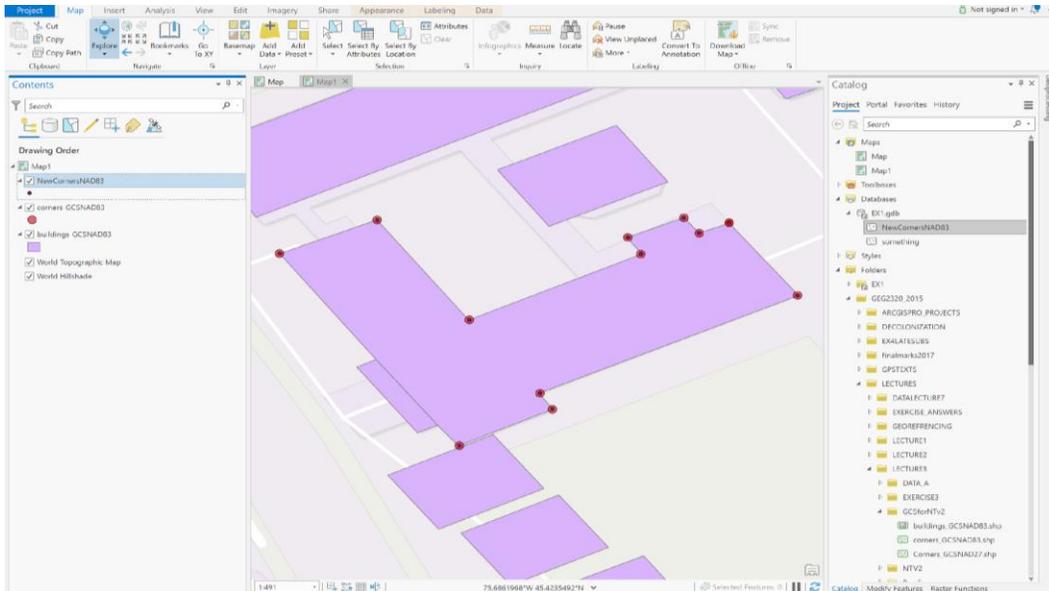
Nom de l'étudiant: _____ # d'étudiant: _____

4. Dans la boîte de dialogue de l'outil « *Project* », sous « *Input Dataset of Feature Class* » (Jeu de données d'entrée ou classe d'entités), choisissez la couche "Corners_GCSNAD27". Sous « *Output Dataset or Feature Class* » (Jeu de données ou classe d'entités de sortie), nommez votre sortie "NewCornersNAD83". Sous « *Output Coordinate System* » (Système de coordonnées de sortie), choisissez GCS_North_American_1983. Enfin, sous « *Geographic Transformation* » (Transformation géographique), choisissez "NAD 1927 to NAD 1983 NTV2 Canada" et cliquez sur Run:



5. Nous allons maintenant voir si tout a fonctionné comme prévu. Créez une nouvelle carte. Ajoutez 'Buildings_GCSNAD83', 'Corners_GCSNAD83' et 'NewCornersNAD83' à la carte. Ils devraient tous s'aligner sans spécifier de transformation:

Nom de l'étudiant: _____ # d'étudiant: _____



Questions à remettre

1. [2 point] Indiquer la distance réelle d'un degré de longitude aux latitudes suivantes en supposant que le modèle de la Terre est une sphère (voir Q3 et Q4):

<i>Latitude</i>	<i>Distance (km)</i>
5	
40	
56	
65	
90	

Que remarquez-vous à propos de la distance réelle d'un degré de longitude, lorsque la latitude se rapproche de l'équateur? Pourquoi?

2. [2 points] Calculer les équivalents en degrés décimaux des coordonnées suivantes avec 5 décimales (voir Q2).

<i>Latitude</i>	<i>Longitude</i>	<i>Latitude dd</i>	<i>Longitude dd</i>
45°6'37" N	75°40'12" W		
45°12'12" N	75°54'46" W		
45°15'15" N	75°26'2" W		
45°19'27" N	76°2'52" W		
45°17'34" N	75°54'22" W		
45°18'30" N	75°35'6" W		

Nom de l'étudiant: _____ # d'étudiant: _____

45°14'44" N	75°53'60" W
45°20'31" N	76°2'51" W
45°20'33" N	75°43'25" W
45°21'2" N	75°45'53" W
45°21'11" N	75°21'10" W
45°25'42" N	75°47'23" W
45°22'14" N	75°51'25" W
45°25'25" N	75°57'20" W
45°23'0" N	76°51'39" W
45°55'55" N	75°54'28" W
45°25'54" N	75°39'41" W
45°18'1" N	75°29'26" W
45°21'19" N	75°32'58" W
45°29'1" N	76°46'4" W

3. [2 points] Calculer l'incertitude de localisation à l'échelle de base pour les échelles cartographiques suivantes (voir Q5):

<i>échelle</i>	<i>incertitude (m)</i>
1 :500	
1 :20000	
1 :50000	
1 :300000	
1 :4000000	

4. [2 point] Pourquoi l'incertitude augmente-t-elle avec une réduction de l'échelle de la carte?
5. [2 points] Calculez l'incertitude de localisation pour les couches géospatiales numériques vectorielles suivantes qui ont été numérisées à partir de cartes aux échelles de base indiquées (voir Q5) :

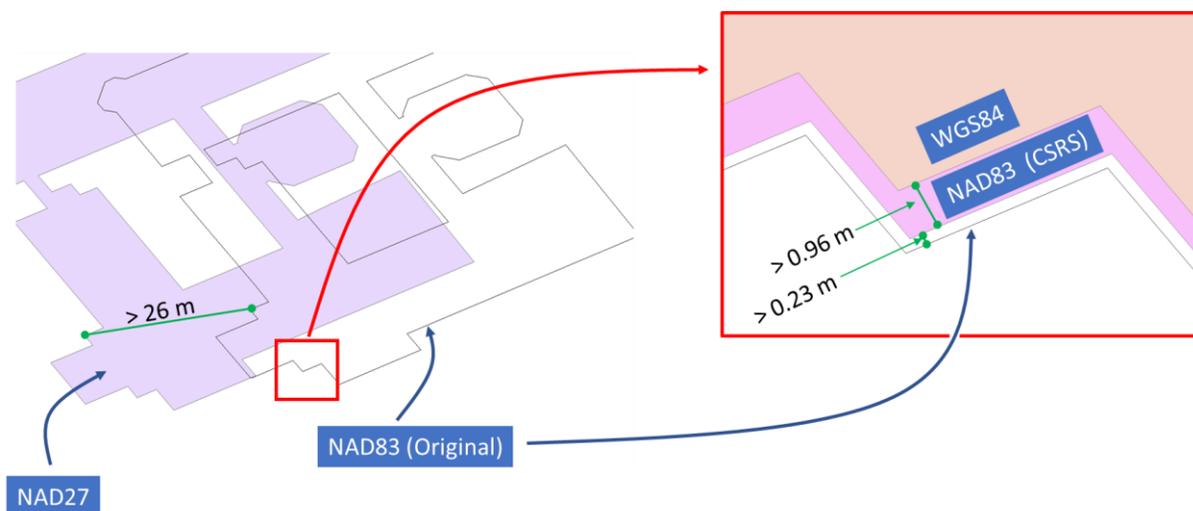
<i>échelle</i>	<i>incertitude (m)</i>
1 :5000	
1 :20000	
1 :50000	
1 :300000	
1 :4000000	

6. [3 points] Calculez l'incertitude de la localisation pour les couches raster suivantes (*qui ont toutes des cellules carrées et ont été encodées à partir de couches géospatiales numériques vectorielles*), compte tenu des informations fournies dans le tableau (voir Q6):

- <https://resources.esri.ca/getting-technical/your-guide-to-canadian-datum-transformations-and-arcgis-online-part-3>
 - Quels sont les systèmes de coordonnées utilisés pour la plupart des données canadiennes?
 - Quel est le système de coordonnées par défaut pour ArcGIS Pro?
 - Quelle transformation est nécessaire pour convertir les coordonnées de NAD83 SCRS en WGS84 au Canada?
 - Quelle transformation est nécessaire pour convertir les coordonnées de NAD83 Original à NAD83 CSRS en Ontario?
 - À Ottawa (voir la figure sur la page suivante), quelles sont les différences en mètres entre:
 - NAD27 et NAD83 (original).
 - NAD83 (original) et NAD83 CSRS.
 - NAD83 CSRS et WGS84.

Notes: Vous pouvez trouver les fichiers de transformation NTv2 de l'Ontario ici:

<https://www.arcgis.com/home/item.html?id=7d26e16ba42040d7a4ae66a4e02d16a0>. Ces fichiers sont nécessaires lorsque l'on travaille avec des municipalités qui n'ont pas adopté les systèmes de coordonnées NAD83 ou WGS84. Pour télécharger les fichiers spécifiques aux provinces pour la transformation entre le NAD83 et le NAD83 CSRS, voir : <https://webapp.csrscs.nrcan-rncan.gc.ca/geod/data-donnees/transformations.php?locale=fr> - vous devez créer un compte et vous connecter pour les télécharger. Vous voudrez peut-être en savoir plus sur le changement de 2022 aux États-Unis pour les données qui s'y trouvent et vous pouvez le lire ici: <https://geospatial.blogs.com/geospatial/2018/08/new-2022-state-plane-coordinate-system-will-change-locations-by-up-to-15-meters.html> (rappel, ceci peut être traduit à l'aide d'extensions de navigateur).



9. [4 points] Créez une carte (par exemple, une mise en page) dans le système de coordonnées « NAD 1983 Modified Transverse Mercator Zone 9 ». Aucune des couches

Nom de l'étudiant: _____ # d'étudiant: _____

que vous utiliserez n'a un système de coordonnées/projection prédéfini. Votre carte doit contenir les couches suivantes que vous trouverez dans lab2.zip:

Couches	Métadonnées des systèmes de coordonnées
buildings_1_NAD83.shp	North American Datum of 1983
buildings_2_NAD27.shp	North American Datum of 1927
buildings_3_NAD27_UTM18N.shp	Universal Transverse Mercator Zone 18 N based on NAD27 geographic coordinates
buildings_4_NAD83_CANADA_ALBERS.shp	Canada Albers Equal Area Conic based on NAD83 geographic coordinates

Laboratoire n°3: Travailler avec des bases de données géographiques

Objectif : L'objectif de ce laboratoire est de vous permettre d'acquérir de l'expérience dans la création d'une base de données géographiques.

Résultats d'apprentissage: A la fin de ce laboratoire, vous serez capable de:

- Créer une géodatabase.
- Importer des données dans la géodatabase.
- Créer une classe de relations dans la géodatabase.
- Répondre à des questions spatiales en utilisant des requêtes ad-hoc spatiales.
- Démontrer le calcul de champs.
- Démontrer la création d'une carte.

Procédure:

- Lisez et travaillez sur les exemples.
- Répondez aux questions à la fin du laboratoire.

Données/matériel:

- Toutes les données sont fournies dans lab3.zip. Vous pouvez utiliser ces données si vous n'avez pas accès au *Scholars Geoportal*.

Devoir:

- Répondez aux questions (#1-6).
- Le total des points disponibles pour ce travail est de 27 points, ce qui représente 12.5% de votre note finale.
- Tous les travaux pratiques doivent être réalisés et remis individuellement.

Format:

Les étudiants remettront les questions avec les réponses complétées.

Les tableaux et les figures doivent être produits selon les normes du Département de géographie, environnement et géomatique de l'Université d'Ottawa. Pour des directives sur le formatage et la présentation appropriés des tableaux et des figures, veuillez consulter la **section d'introduction de ce manuel de laboratoire.**

Nom de l'étudiant: _____ # d'étudiant: _____

Q1: Comment télécharger des données à partir du Scholars GeoPortal?

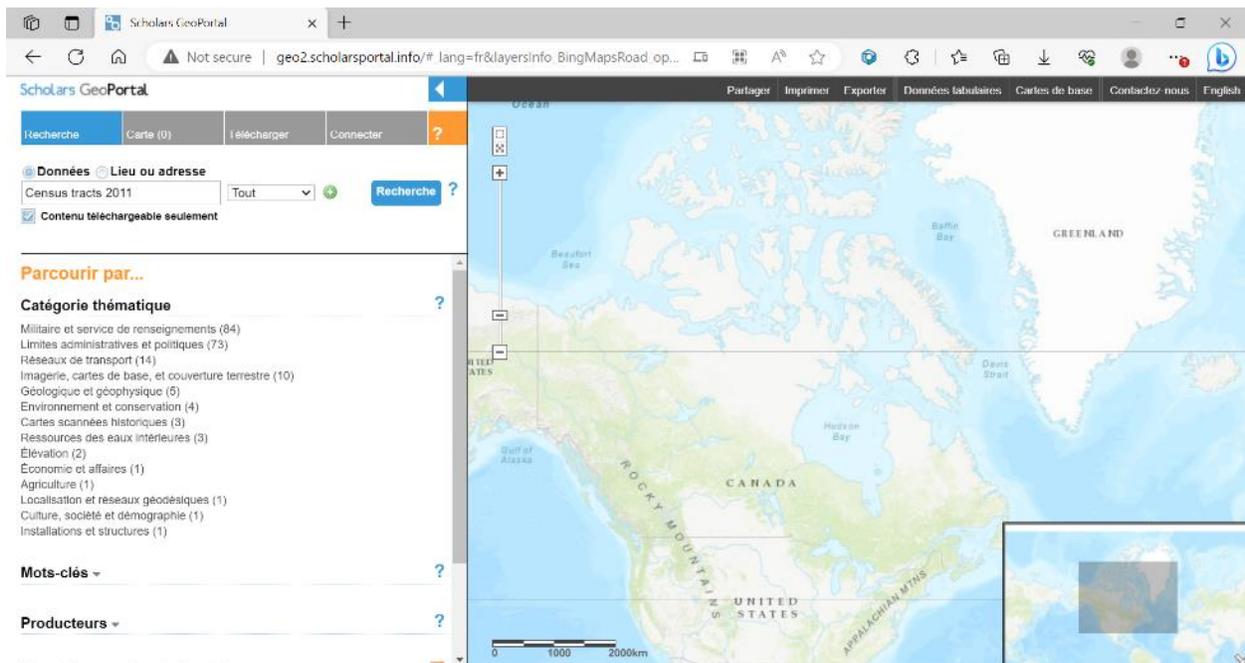
Le Scholars GeoPortal est une infrastructure matérielle et logicielle qui permet aux étudiants et aux chercheurs de l'Ontario d'accéder à des ensembles de données géospatiales à grande échelle. Il s'agit notamment de données vectorielles terrestres (eau, caractéristiques culturelles, etc.), de données de recensement et de certaines images raster.

Vous ne pouvez télécharger certaines données du Scholars GeoPortal que si vous êtes authentifié par une université de l'Ontario, soit en accédant aux données sur le campus, soit en utilisant un VPN. Si c'est le cas, vous devriez suivre les instructions ci-dessous pour télécharger les données de cet exercice. Apprendre à télécharger des données à partir du Scholars Geoportal est très utile.

Note: si vous ne pouvez pas accéder au Scholars GeoPortal, vous pouvez trouver tous les mêmes ensembles de données dans le fichier appelé lab3.zip, il vous suffit de les trier en fonction des noms de fichiers.

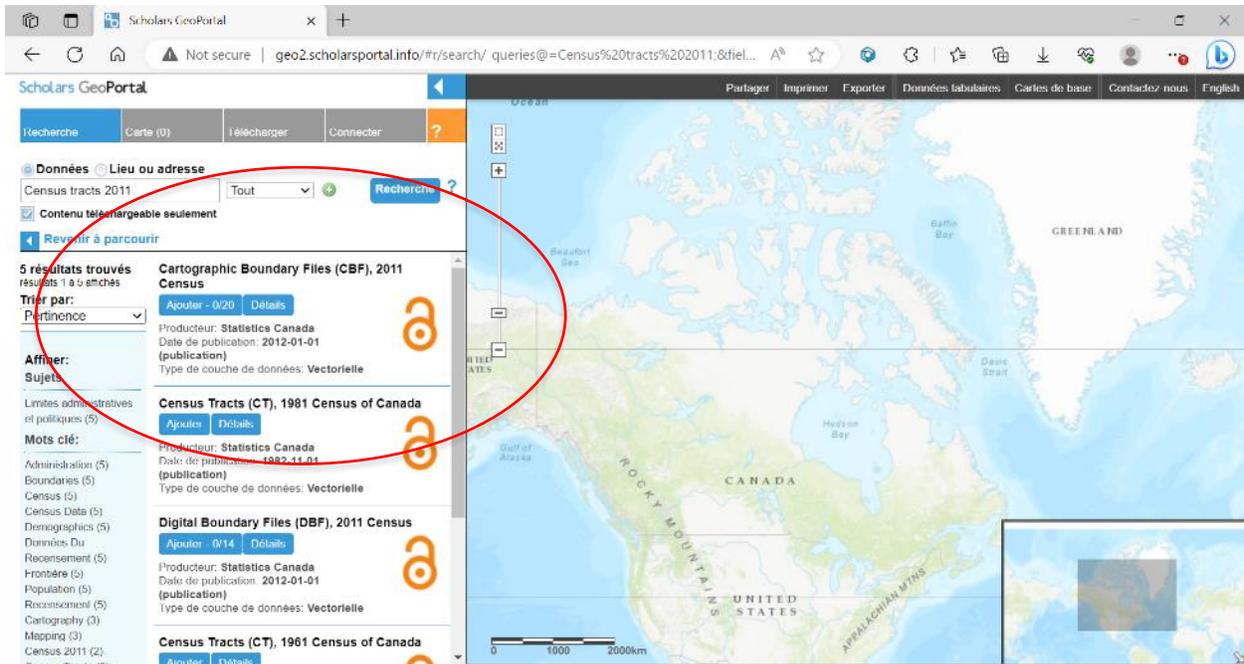
Pour télécharger les secteurs de recensement d'Ottawa-Gatineau dont vous avez besoin pour cet exercice:

1. Aller sur <http://geo2.scholarsportal.info/>
2. Dans la barre de recherche, tapez "Census tracts 2011" et cliquez sur le bouton Recherche:

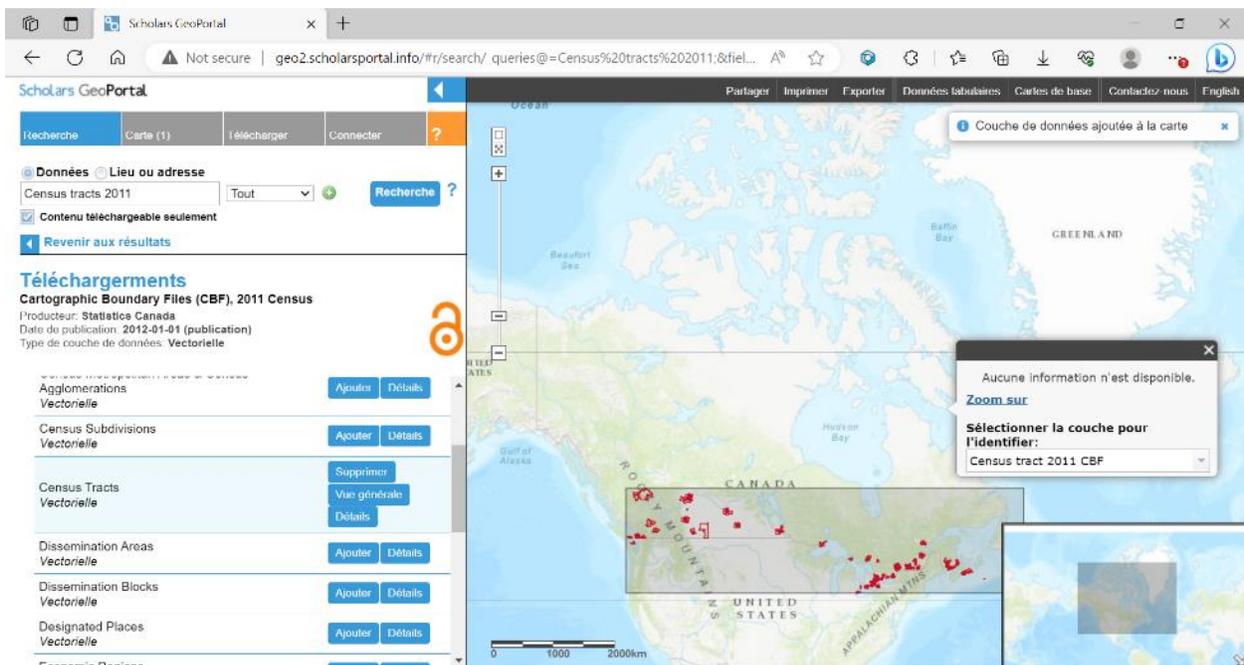


Nom de l'étudiant: _____ # d'étudiant: _____

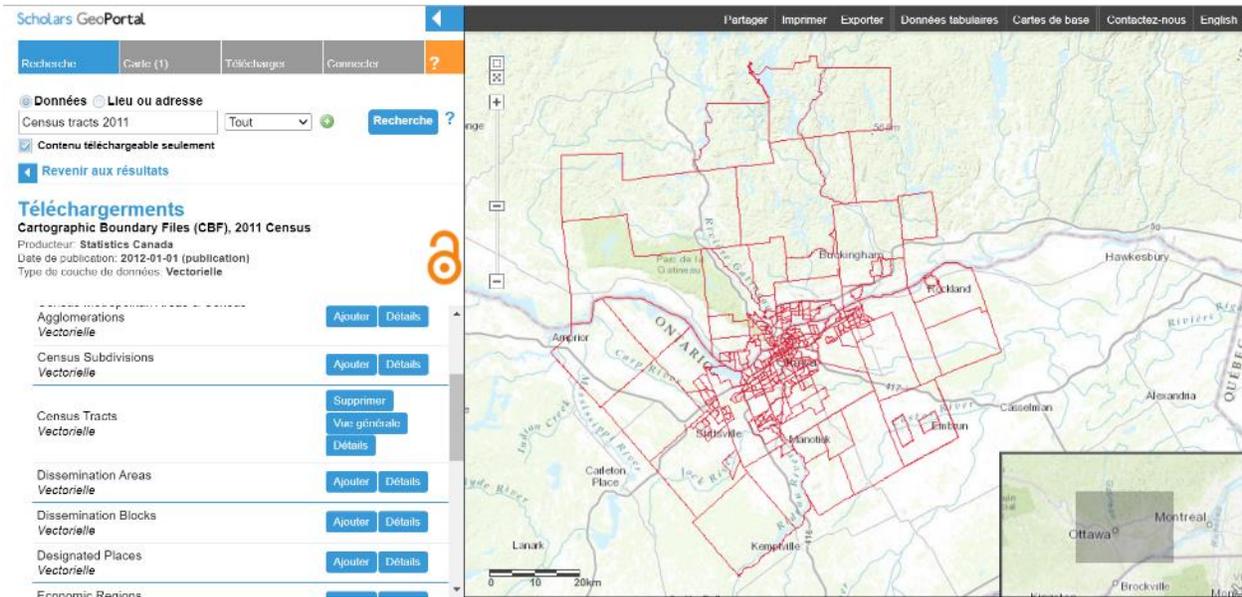
3. Lorsque la recherche est terminée, la liste suivante s'affiche (notez que l'ordre de la liste peut différer de celui que vous voyez ci-dessous). Trouvez l'élément "Cartographic Boundary Files (CBF), 2011 Census":



4. Pour les "Cartographic Boundary Files (CBF), 2011 Census", cliquez sur le bouton Ajouter - 0/20 et vous verrez une autre liste. Dans cette liste, recherchez l'élément intitulé "Census Tracts Vector" et cliquez sur le bouton Ajouter:



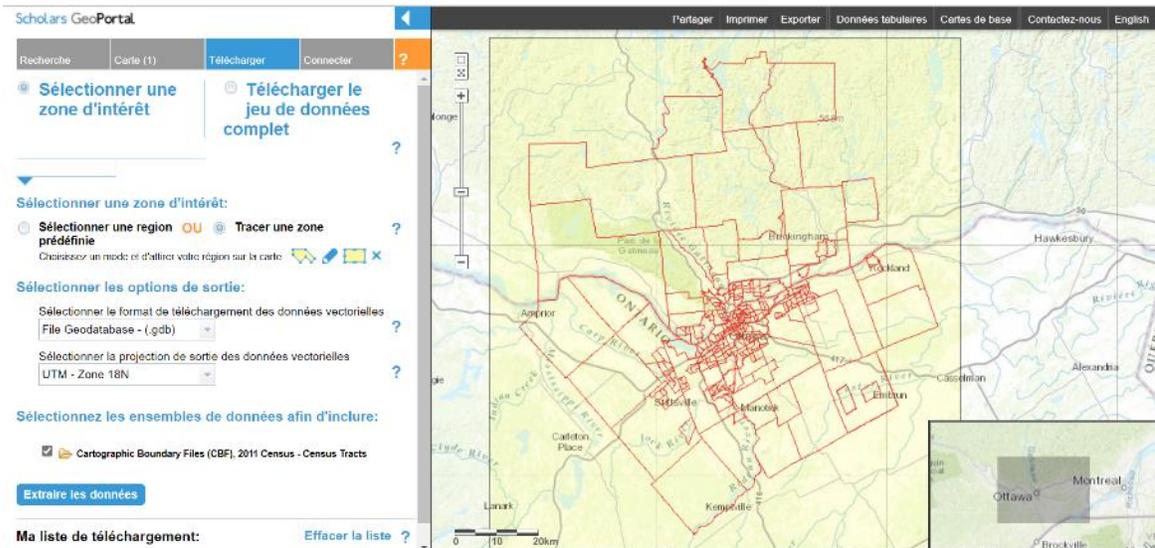
5. Vous pouvez zoomer sur la région d'Ottawa à l'aide des outils de zoom:



6. Cliquez sur l'onglet Télécharger en haut de la page.

7. Procédez comme suit:

- a. Sélectionnez "Sélectionner une zone d'intérêt".
- b. Sélectionnez " Tracer une zone ".
- c. Cliquez sur l'outil rectangle.
- d. Cliquez et faites glisser une boîte autour de la région d'Ottawa-Gatineau de manière à ce que toutes les lignes rouges soient entièrement contenues dans la boîte que vous dessinez.
- e. Dans la section "Sélectionner les options de sortie", choisissez « File Geodatabase - (.gdb) » dans le menu "Sélectionner le format de téléchargement des données vectorielles".
- f. Dans la section "Sélectionner la projection de sortie des données vectorielles" choisissez UTM - Zone 18N.



8. Cliquez sur le bouton "Extraire les données". Lorsque le téléchargement est terminé, vous verrez un fichier .zip à télécharger dans la section "Ma liste de téléchargement".

Ma liste de téléchargement:

[Effacer la liste ?](#)

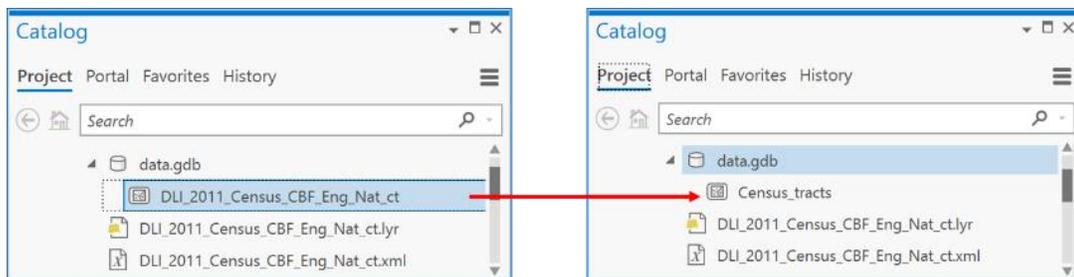
Cartographic Boundary Files (CBF), 2011 Census - Census Tracts:

[Télécharger tous - environ 261.5 KB](#)

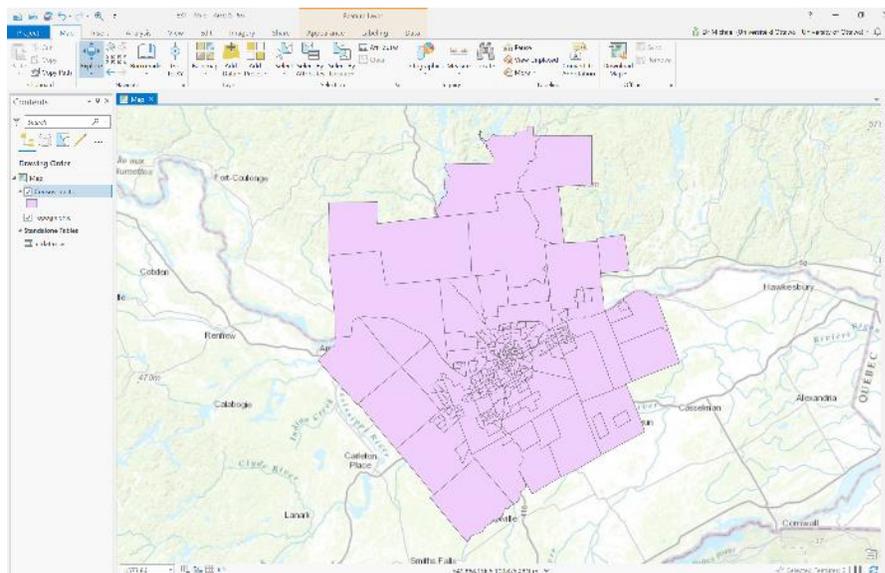
DLI_2011_Census_CBF_Eng_Nat_ct.zip



9. Téléchargez le fichier en cliquant sur le lien dont le texte est en gras et souligné, et décompressez le fichier vers un emplacement dont vous vous souviendrez.
10. Ouvrez ArcGIS Pro et ajoutez le dossier que vous avez décompressé à l'étape précédente.
11. Dans la géodatabase **data.gdb** de ce dossier, vous trouverez une seule classe d'entités qui s'appellera quelque chose comme "DLI_2011_Census....". Dans le volet « *Catalog* », renommez cette classe d'entités "Census_tracts" (cliquez avec le bouton droit de la souris sur la classe d'entités et choisissez « *Rename* »).



12. Ajoutez maintenant "Census_tracts" à une nouvelle carte:



13. Ouvrez la table d'attributs. Dans la table d'attributs, vous pouvez voir que les colonnes ou les champs n'identifient que chacun des secteurs de recensement. En particulier, le champ **CTUID** est important car il contient un identifiant unique pour chaque secteur de recensement.

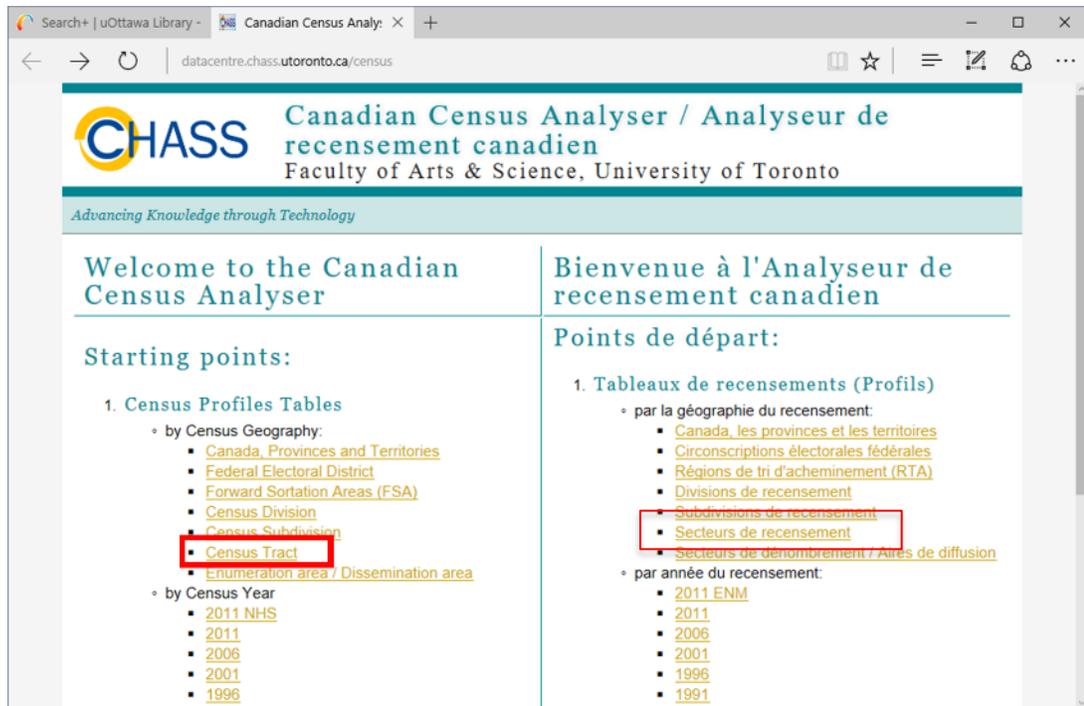
OBJECTID	Shape	CTUID	CTNAME	CMAUID	CMANAME	CMATYPE	CMAPUID	PRUID	PRNAME	Shape_Length	Shape_Area
1	Polygon	5050151.08	0151.08	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	21050.810924	12904044.31956
2	Polygon	5050103.00	0103.00	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	4134.350028	839357.60299
3	Polygon	5050810.01	0810.01	505	Ottawa - Gatineau (p...	B	24505	24	Quebec / Québec	7613.873489	3032294.61721
4	Polygon	5050170.04	0170.04	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	4650.77598	1341659.908304
5	Polygon	5050133.00	0133.00	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	6520.270482	2399590.038124
6	Polygon	5050007.02	0007.02	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	3569.580512	421410.052837
7	Polygon	5050610.02	0610.02	505	Ottawa - Gatineau (p...	B	24505	24	Quebec / Québec	5216.891258	1594179.441413
8	Polygon	5050127.00	0127.00	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	23139.23255	19679560.736875
9	Polygon	5050160.03	0160.03	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	17437.421004	11343443.586152
10	Polygon	5050125.05	0125.05	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	6718.400844	1982854.998591
11	Polygon	5050056.00	0056.00	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	3820.730813	593577.826658
12	Polygon	5050141.08	0141.08	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	14383.195416	8713545.528
13	Polygon	5050904.00	0904.00	505	Ottawa - Gatineau (p...	B	24505	24	Quebec / Québec	77605.085772	167303842.655632

Q2: Où puis-je obtenir les données correspondant aux secteurs de recensement?

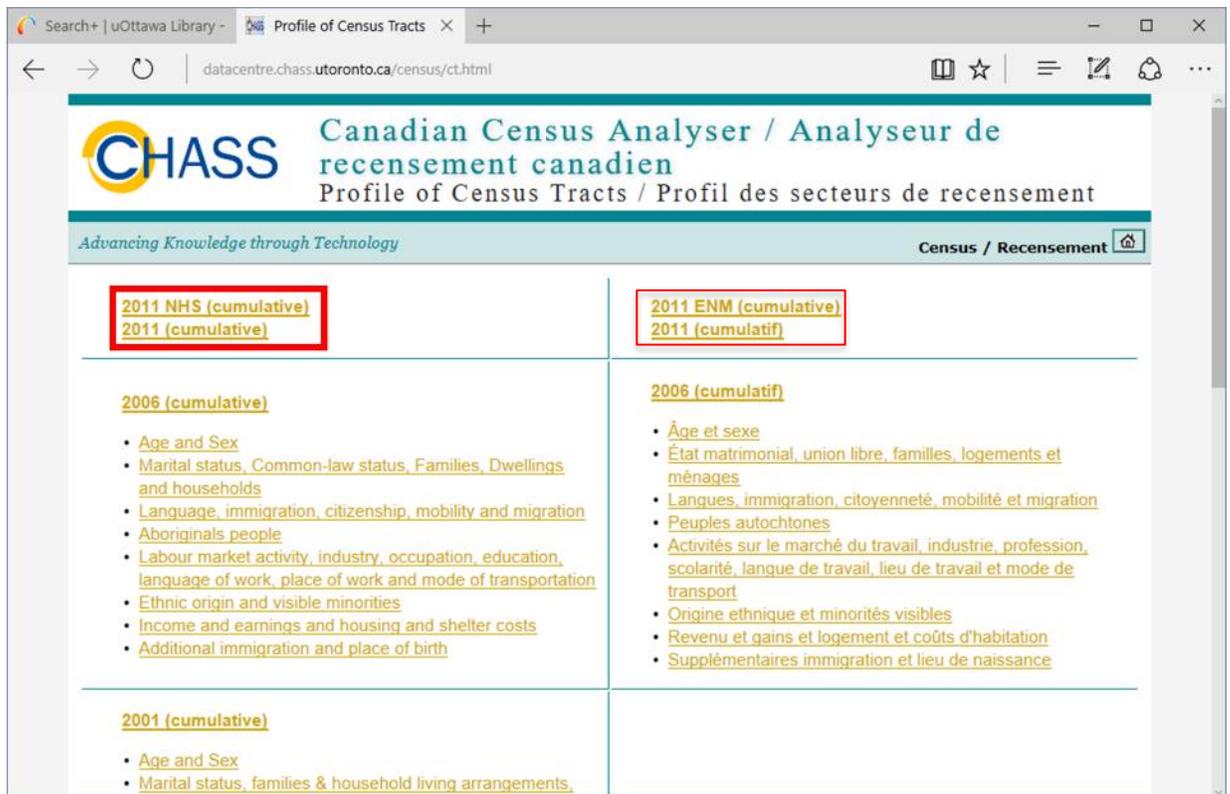
Les données de chaque secteur de recensement sont accessibles par l'entremise de l'outil Web de l'Analyseur du recensement canadien, qui est disponible à partir du campus ou lorsque vous êtes authentifié par le VPN de l'Université d'Ottawa. L'analyseur de recensement n'est pas une source de données, mais plutôt une interface ou un programme Web. La source de données est toujours l'institution, l'organisation ou la personne qui a créé les données, dans ce cas-ci Statistique Canada.

Note: Comme pour les données du Scholars GeoPortal, si vous ne pouvez pas accéder aux données de l'Analyseur du recensement canadien, vous pouvez trouver les données dont vous avez besoin dans le fichier lab3.zip.

1. Aller sur <http://datacentre.chass.utoronto.ca/census/>
2. Ensuite, sous "Tableaux de recensement (Profils)", choisissez "Secteur de recensement" :



3. Une liste de dates s'affiche, et nous nous intéressons aux données du recensement de 2011. Il existe deux types de données de recensement 2011:

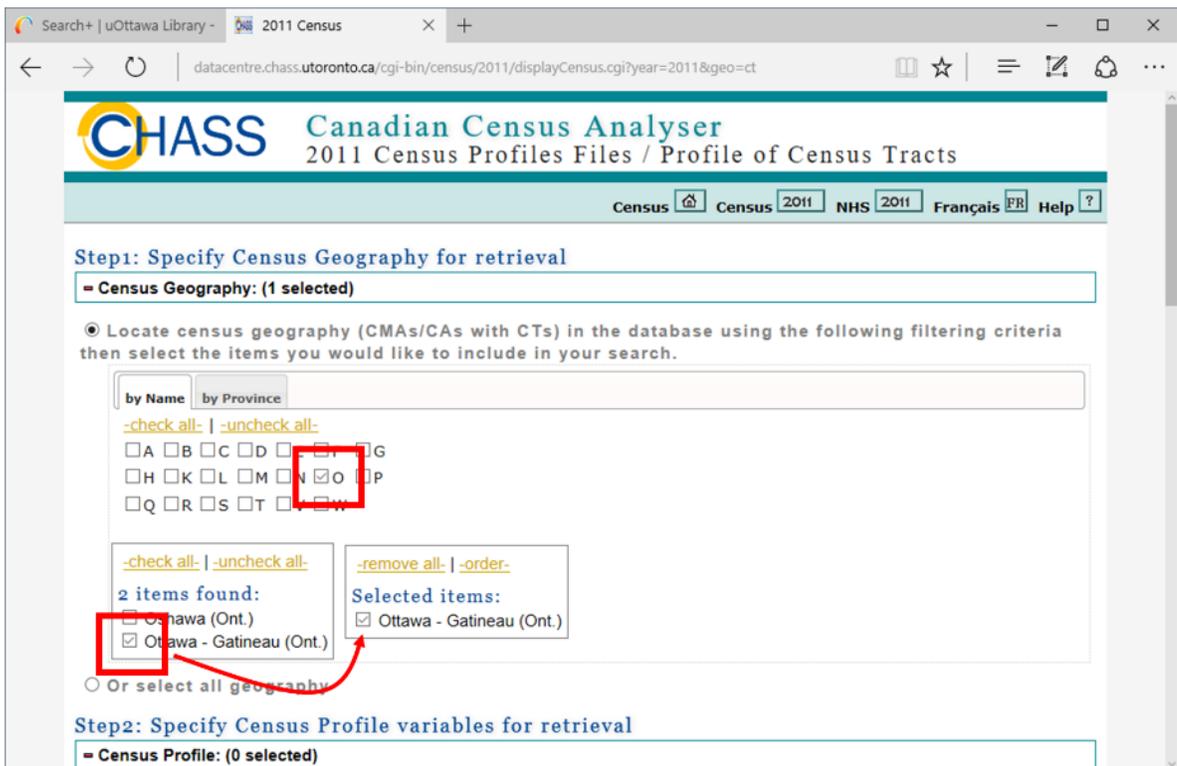


Nom de l'étudiant: _____ # d'étudiant: _____

- a. 2011 ENM (cumulative): les données du recensement complet qui contiennent plus de 1,000 variables déclarées par environ 20 % de la population et qui contiennent des détails sur le travail, les transports, etc.
- b. 2011 (cumulatif): le recensement complet auquel tout le monde participe et qui contient des données sur la langue, l'âge, le sexe, l'état matrimonial et la répartition par âge et par population.

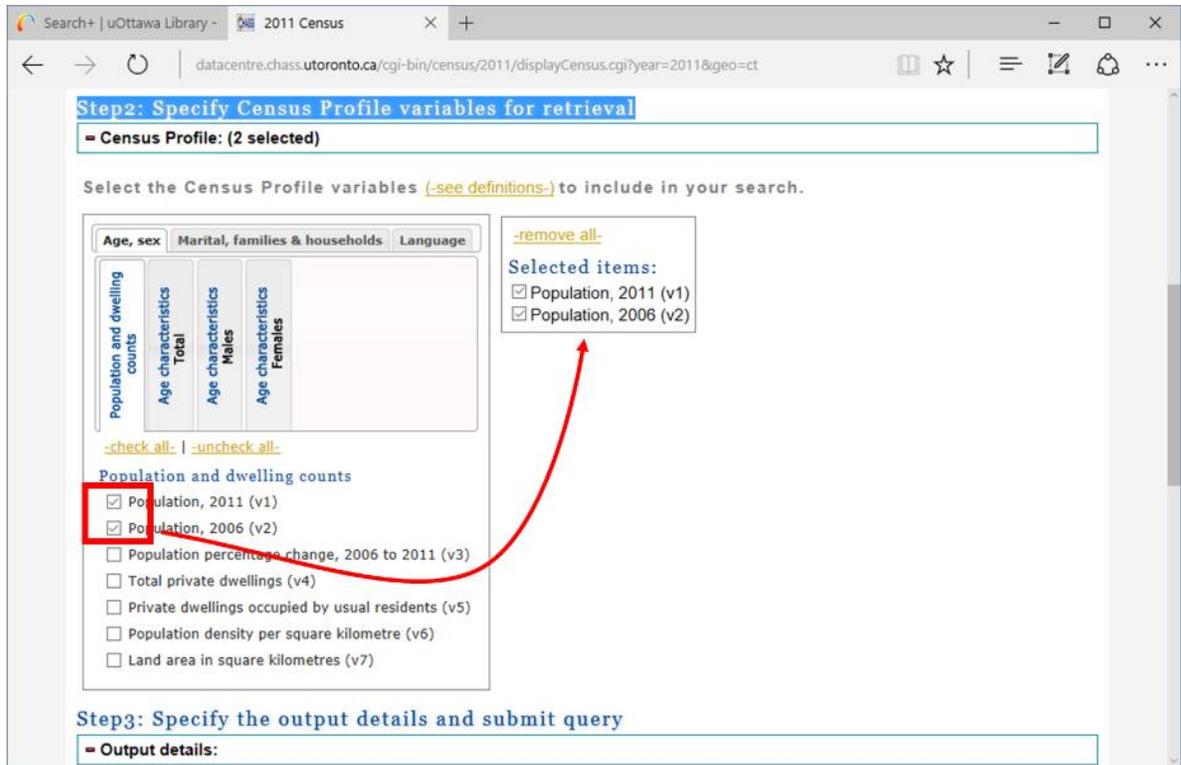
Les données sont disponibles au niveau du secteur de recensement pour les deux types de données de recensement.

4. Cliquez sur "2011 (cumulatif)", puis, sous "Étape 1 : Spécifier la géographie du recensement à extraire", cochez la case "O" et, en dessous, la case "Ottawa-Gatineau (Ont.)". Vous devriez voir ce qui suit:

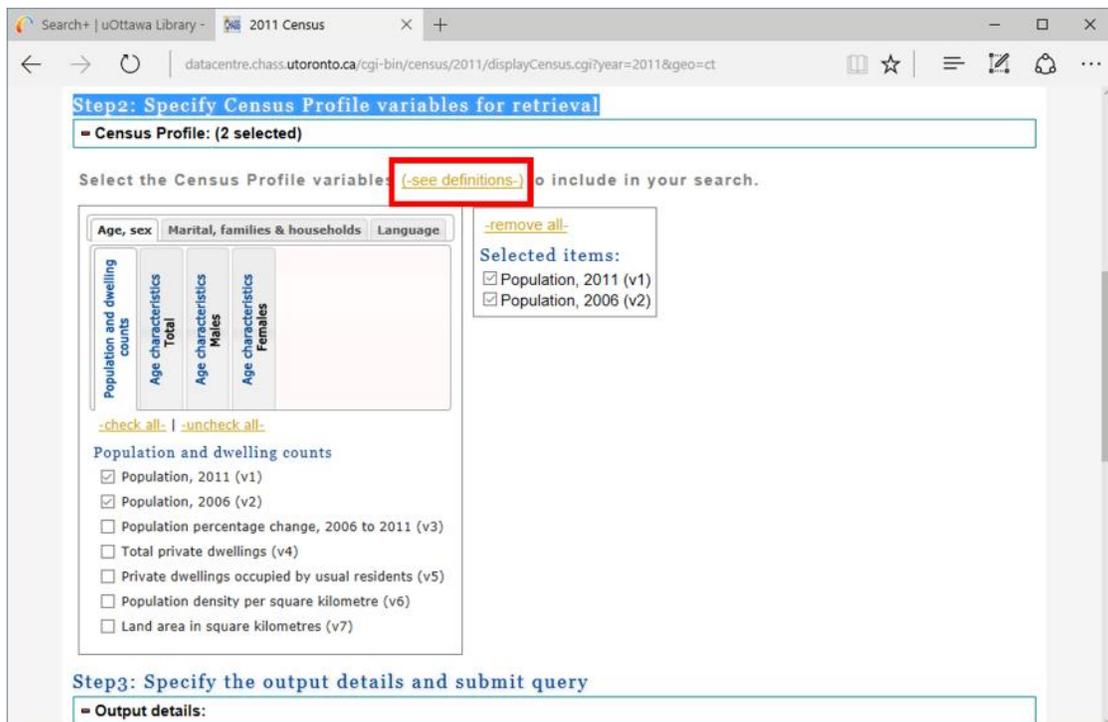


5. Maintenant, sous "Étape 2 : Spécifier les variables du profil de recensement à extraire", cochez les cases Population, 2011 (v1) et Population, 2006 (v2) pour spécifier que vous voulez télécharger ces deux variables de recensement:

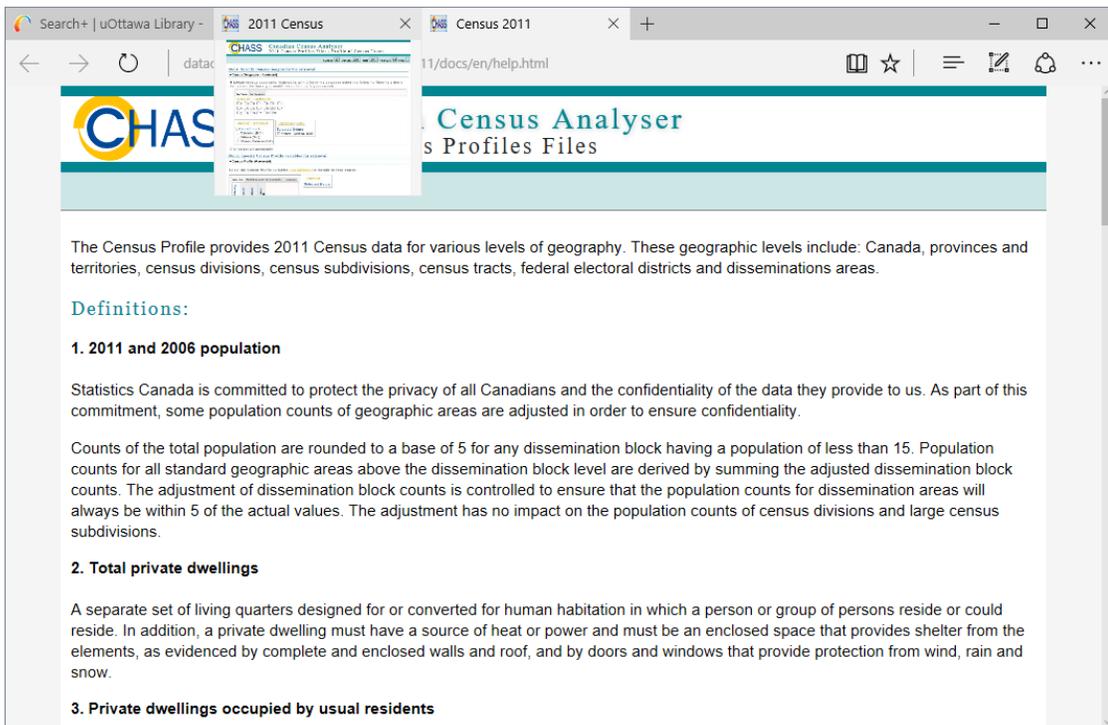
Nom de l'étudiant: _____ # d'étudiant: _____



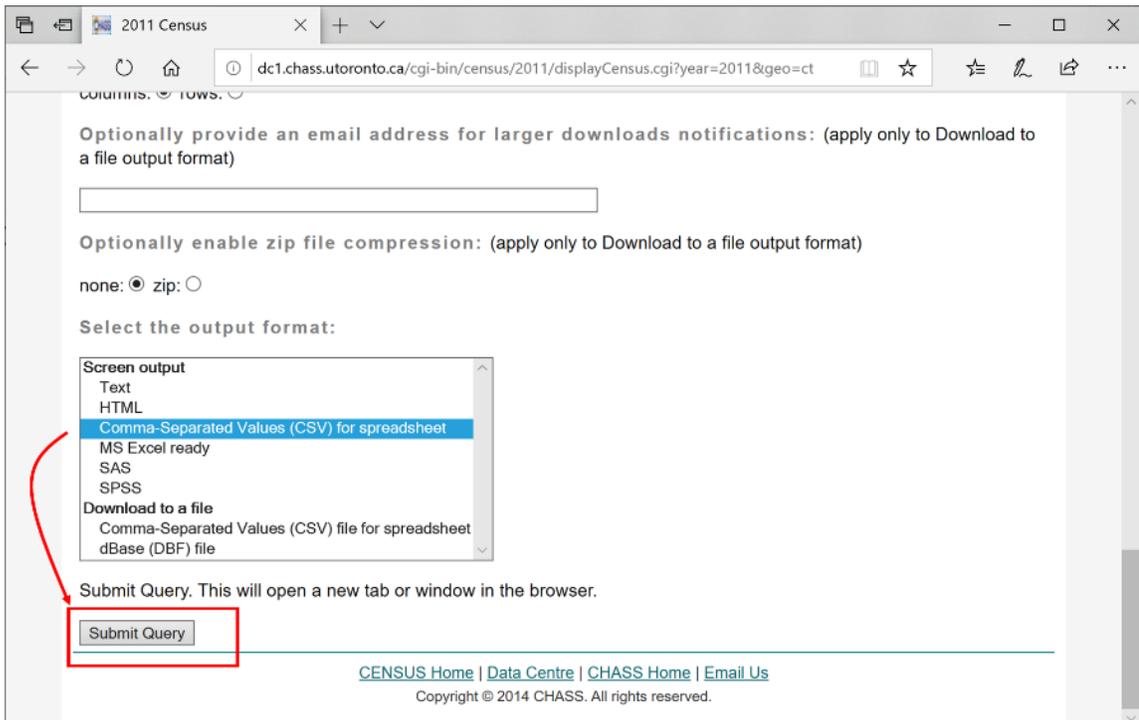
6. Ensuite, vous devez savoir ce que vous téléchargez, c'est-à-dire ce que signifient ces variables et comment elles sont définies. Cliquez donc sur le lien (-voir les définitions-):



7. Lisez les informations sur la population de 2011 et de 2006:

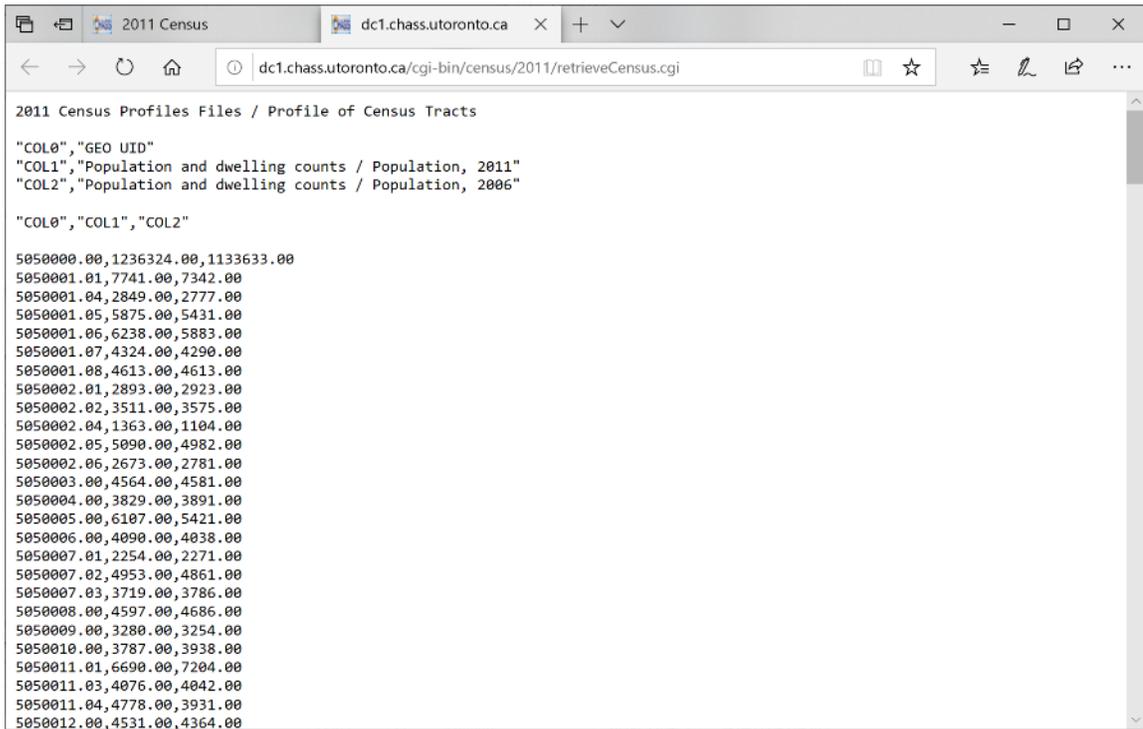


8. Retournez à l'onglet Recensement 2011 dans votre navigateur. Sous "Étape 3 : Spécifier les détails de la sortie et soumettre la requête", sous "Sélectionner le format de sortie", choisissez "Valeurs séparées par des virgules (CSV) pour tableur" et cliquez sur le bouton Soumettre la requête:



Nom de l'étudiant: _____ # d'étudiant: _____

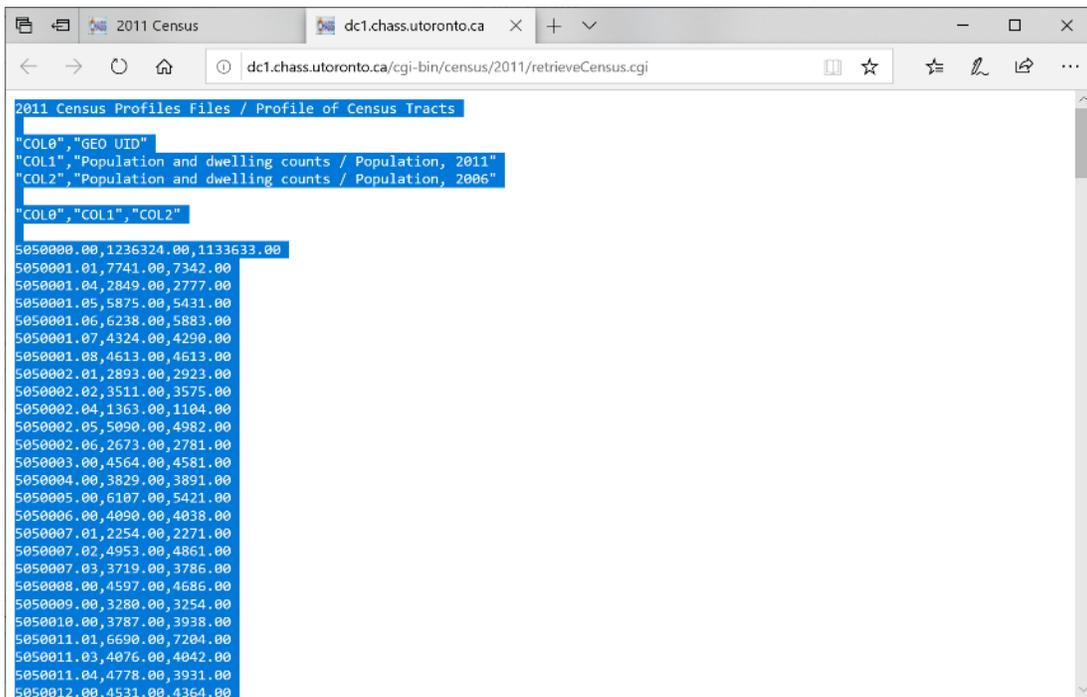
9. Vous verrez maintenant une page web avec les données qui ressemble à ceci:



```
2011 Census Profiles Files / Profile of Census Tracts
"COL0", "GEO UID"
"COL1", "Population and dwelling counts / Population, 2011"
"COL2", "Population and dwelling counts / Population, 2006"

"COL0", "COL1", "COL2"
5050000.00,1236324.00,1133633.00
5050001.01,7741.00,7342.00
5050001.04,2849.00,2777.00
5050001.05,5875.00,5431.00
5050001.06,6238.00,5883.00
5050001.07,4324.00,4290.00
5050001.08,4613.00,4613.00
5050002.01,2893.00,2923.00
5050002.02,3511.00,3575.00
5050002.04,1363.00,1104.00
5050002.05,5090.00,4982.00
5050002.06,2673.00,2781.00
5050003.00,4564.00,4581.00
5050004.00,3829.00,3891.00
5050005.00,6107.00,5421.00
5050006.00,4090.00,4038.00
5050007.01,2254.00,2271.00
5050007.02,4953.00,4861.00
5050007.03,3719.00,3786.00
5050008.00,4597.00,4686.00
5050009.00,3280.00,3254.00
5050010.00,3787.00,3938.00
5050011.01,6690.00,7204.00
5050011.03,4076.00,4042.00
5050011.04,4778.00,3931.00
5050012.00,4531.00,4364.00
```

10. Sélectionnez tout le texte de la page web en utilisant la séquence de touches Ctrl+A sur le clavier, suivie de Ctrl+C pour copier les données:



```
2011 Census Profiles Files / Profile of Census Tracts
"COL0", "GEO UID"
"COL1", "Population and dwelling counts / Population, 2011"
"COL2", "Population and dwelling counts / Population, 2006"

"COL0", "COL1", "COL2"
5050000.00,1236324.00,1133633.00
5050001.01,7741.00,7342.00
5050001.04,2849.00,2777.00
5050001.05,5875.00,5431.00
5050001.06,6238.00,5883.00
5050001.07,4324.00,4290.00
5050001.08,4613.00,4613.00
5050002.01,2893.00,2923.00
5050002.02,3511.00,3575.00
5050002.04,1363.00,1104.00
5050002.05,5090.00,4982.00
5050002.06,2673.00,2781.00
5050003.00,4564.00,4581.00
5050004.00,3829.00,3891.00
5050005.00,6107.00,5421.00
5050006.00,4090.00,4038.00
5050007.01,2254.00,2271.00
5050007.02,4953.00,4861.00
5050007.03,3719.00,3786.00
5050008.00,4597.00,4686.00
5050009.00,3280.00,3254.00
5050010.00,3787.00,3938.00
5050011.01,6690.00,7204.00
5050011.03,4076.00,4042.00
5050011.04,4778.00,3931.00
5050012.00,4531.00,4364.00
```

11. Ouvrez le Bloc-notes sur votre bureau et utilisez la séquence de touches Ctrl+V pour coller les données dans le Bloc-notes:

```
*ctdata.txt - Notepad
File Edit Format View Help
2011 Census Profiles Files / Profile of Census Tracts

"COL0", "GEO UID"
"COL1", "Population and dwelling counts / Population, 2011"
"COL2", "Population and dwelling counts / Population, 2006"

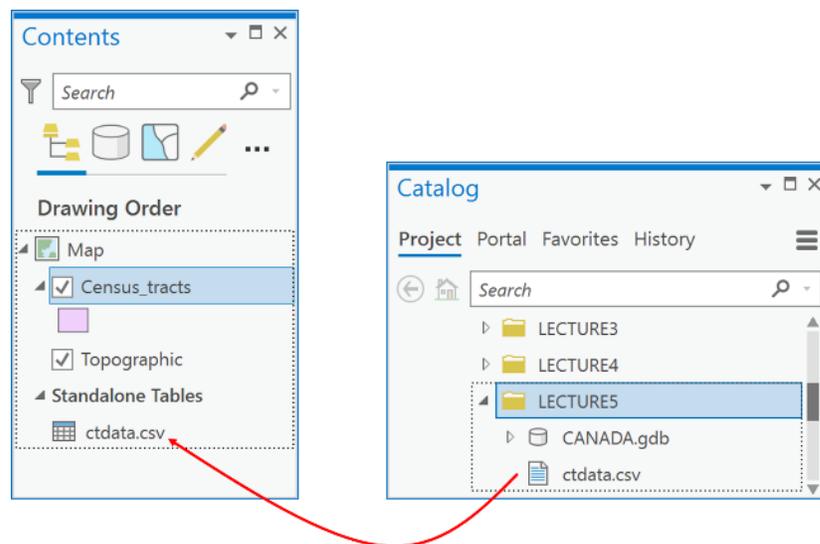
"COL0", "COL1", "COL2"

5050000.00,1236324.00,1133633.00
5050001.01,7741.00,7342.00
5050001.04,2849.00,2777.00
5050001.05,5875.00,5431.00
5050001.06,6238.00,5883.00
5050001.07,4324.00,4290.00
5050001.08,4613.00,4613.00
5050002.01,2893.00,2923.00
5050002.02,3511.00,3575.00
5050002.04,1363.00,1104.00
5050002.05,5090.00,4982.00
5050002.06,2673.00,2781.00
5050003.00,4564.00,4581.00
5050004.00,3829.00,3891.00
5050005.00,6107.00,5421.00
5050006.00,4090.00,4038.00
5050007.01,2254.00,2271.00
5050007.02,4053.00,4851.00
```

12. Vous devez maintenant modifier les données afin qu'elles puissent être ouvertes en tant que fichier ".csv" dans ArcGIS Pro. Pour cela, faites en sorte que vos données ressemblent à ce qui suit. Il est important qu'il n'y ait pas de lignes vides entre l'en-tête (la première ligne qui contient les étiquettes des colonnes) et les données. Enregistrez ensuite le fichier sous le nom "**ctdata.csv**". Utilisez l'explorateur Windows pour vérifier que vous n'avez pas accidentellement utilisé une autre extension pour le fichier (il peut facilement se retrouver sous la forme *ctdata.csv.txt*, ce qui ne fonctionnera pas !)

```
ctdata.csv - Notepad
File Edit Format View Help
"CTUID", "POP2011", "POP2006"
5050000.00,1236324.00,1133633.00
5050001.01,7741.00,7342.00
5050001.04,2849.00,2777.00
5050001.05,5875.00,5431.00
5050001.06,6238.00,5883.00
5050001.07,4324.00,4290.00
5050001.08,4613.00,4613.00
5050002.01,2893.00,2923.00
5050002.02,3511.00,3575.00
5050002.04,1363.00,1104.00
5050002.05,5090.00,4982.00
5050002.06,2673.00,2781.00
5050003.00,4564.00,4581.00
5050004.00,3829.00,3891.00
5050005.00,6107.00,5421.00
5050006.00,4090.00,4038.00
5050007.01,2254.00,2271.00
5050007.02,4953.00,4861.00
5050007.03,3719.00,3786.00
5050008.00,4597.00,4686.00
5050009.00,3280.00,3254.00
5050010.00,3787.00,3938.00
5050011.01,6600.00,7204.00
```

13. Retournez maintenant à ArcGIS Pro. Ajoutez le dossier dans lequel vous avez enregistré le fichier .csv et ajoutez le fichier .csv au contenu du projet:



14. Ouvrez et examinez cette table d'attributs autonome ainsi que la table de la classe d'entités Census_tracts que vous avez téléchargée, renommée et ajoutée à ArcGIS Pro à la question 1 ci-dessus. Vous constaterez que les deux tables comportent un champ CTUID, mais que les valeurs des champs sont dans des formats différents (voir la figure ci-dessus). Plus précisément, le fichier ctdata.csv a interprété le CTUID comme un champ numérique, alors que le fichier Census_tracts interprète le champ CTUID comme du texte. Les champs ne peuvent être joints que s'ils a) sont tous les deux des types de

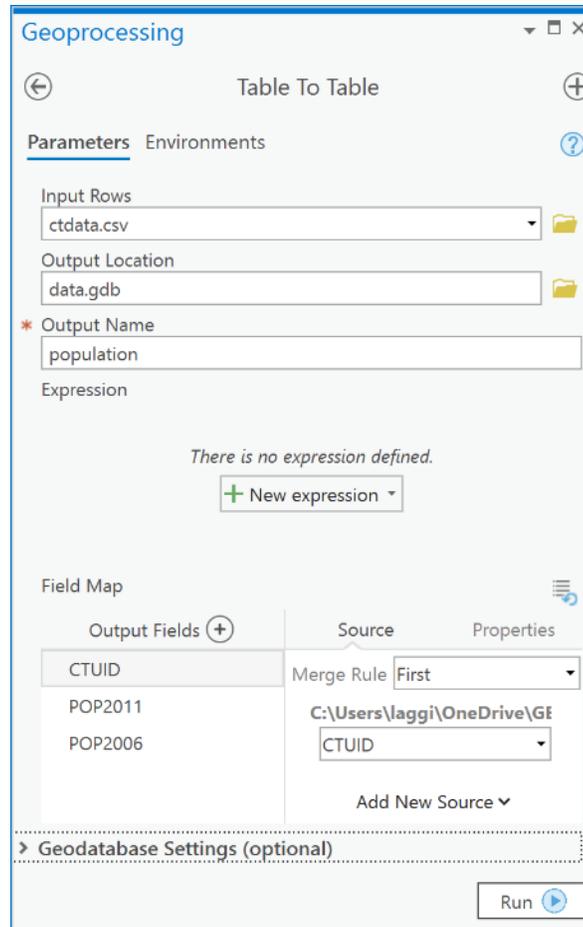
données de type texte ou numérique, et b) s'ils contiennent des valeurs communes qui sont identiques pour les mêmes caractéristiques. Dans ce cas, le format d'un CTUID doit être une valeur textuelle de la forme "#####.##" sans exception, comme "5050151.08". La confusion vient du fait que ces champs de texte utilisent des nombres, mais ces nombres n'ont pas de signification numérique, ce sont juste des étiquettes uniques pour chaque secteur de recensement. Par conséquent, nous devons manipuler les données du fichier ctdata.csv avant de pouvoir accéder aux champs de population pour les cartographier ou les analyser plus en détail.

Champ commun mais formats différents

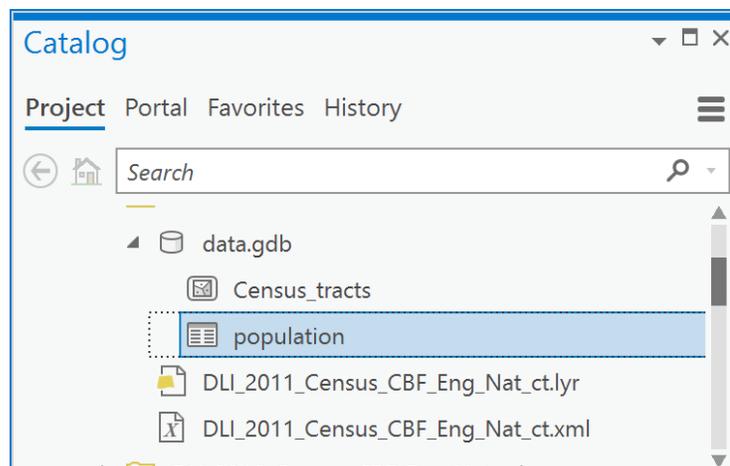
Cadré à droite = numérique

Cadré à gauche = texte

- Ouvrez l'outil « *Table to Table* », et sélectionnez votre fichier **ctdata.csv** comme « *Input Rows* », et la géodatabase **data.gdb** de Q1 comme « *Output Location* ». Appelez le résultat "population" en tant que nouvelle table. Cliquez ensuite sur Run.



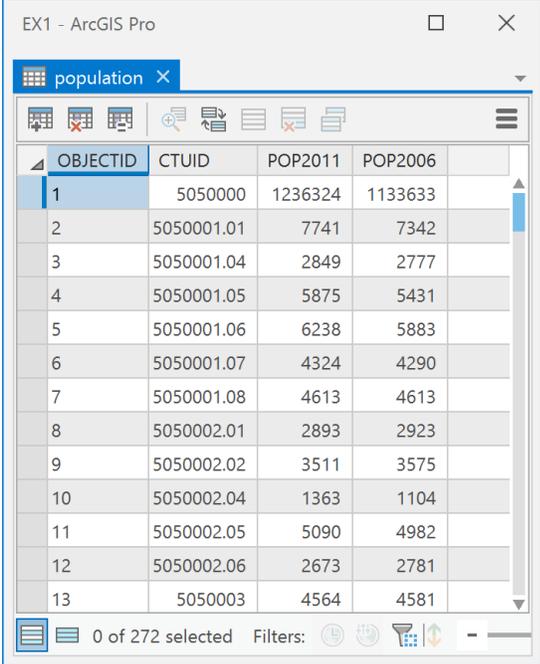
16. Vous devriez maintenant avoir un tableau de population dans votre **data.gdb**.



17. Ouvrez le tableau 'population' et vous verrez qu'un nouvel attribut appelé OBJECTID a été ajouté. Cet attribut est automatiquement ajouté à toutes les tables d'une géodatabase. Vous remarquerez que la première ligne contient des populations très importantes ; c'est parce qu'il s'agit de la population totale pour chaque année. Vous pouvez supprimer cette ligne ou la laisser, cela n'affectera pas le reste de l'exercice.

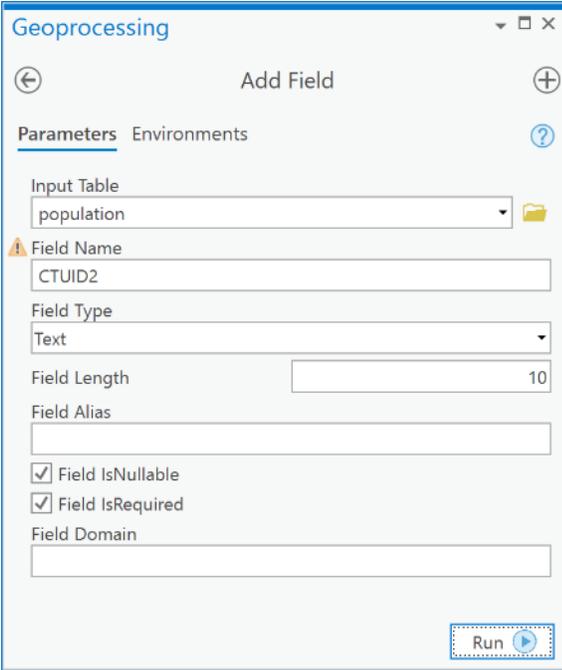
Nom de l'étudiant: _____ # d'étudiant: _____

Plus important encore, vous remarquerez que le CTUID est toujours au format numérique et qu'il manque à de nombreuses valeurs la partie ".##" (telle que ".00") des CTUIDs:



OBJECTID	CTUID	POP2011	POP2006
1	5050000	1236324	1133633
2	5050001.01	7741	7342
3	5050001.04	2849	2777
4	5050001.05	5875	5431
5	5050001.06	6238	5883
6	5050001.07	4324	4290
7	5050001.08	4613	4613
8	5050002.01	2893	2923
9	5050002.02	3511	3575
10	5050002.04	1363	1104
11	5050002.05	5090	4982
12	5050002.06	2673	2781
13	5050003	4564	4581

18. Ajoutez un nouveau champ (« *Data Management Tools* » (Outils de gestion des données) -> « *Fields* » (Champs) -> « *Add Field* » (Ajouter un champ)) à la table "population", appelé CTUID2. Il doit être de type "Texte" et sa longueur doit être de 10 (« *Field Length* »). Exécutez l'opération.



Geoprocessing

Add Field

Parameters Environments

Input Table
population

Field Name
CTUID2

Field Type
Text

Field Length
10

Field Alias

Field IsNullable

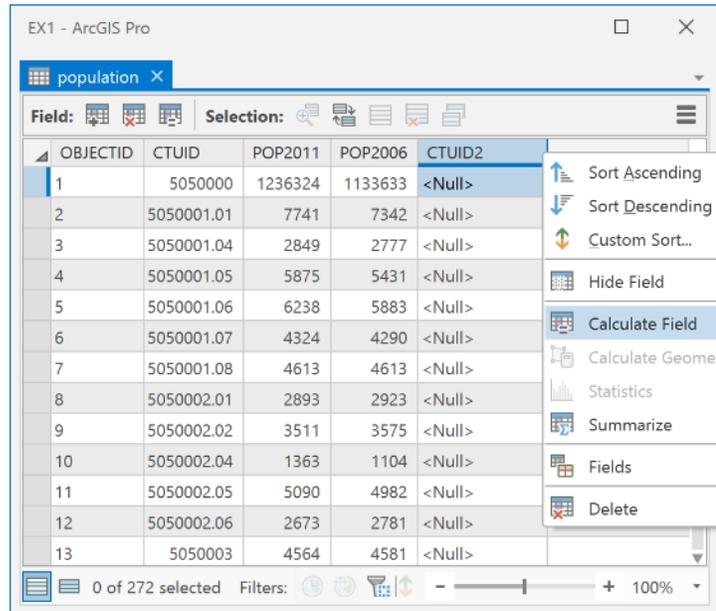
Field IsRequired

Field Domain

Run

Nom de l'étudiant: _____ # d'étudiant: _____

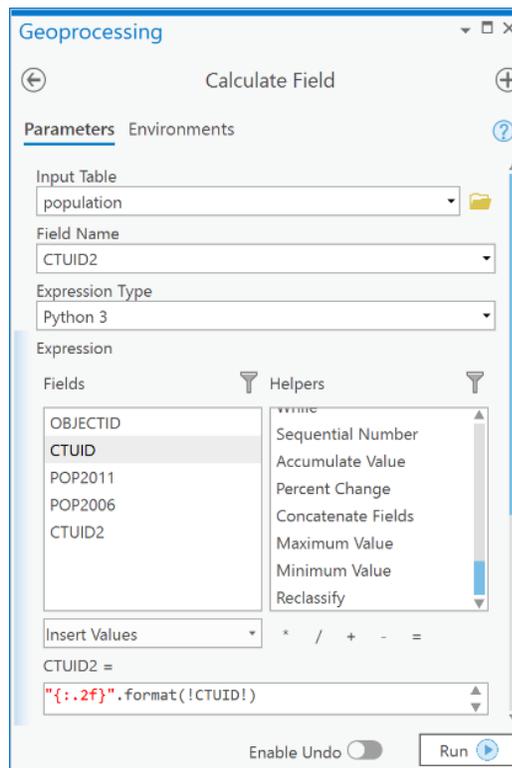
19. Dans la table de population, vous devriez maintenant voir un champ CTUID2. Cliquez avec le bouton droit de la souris sur ce champ et choisissez « *Calculate field* » (Calculer le champ):



20. Dans l'outil Calculate Field, tapez, sous CTUID =, la commande suivante:

"{: .2f} ".format(!CTUID!)

21. Cliquez ensuite sur Run:



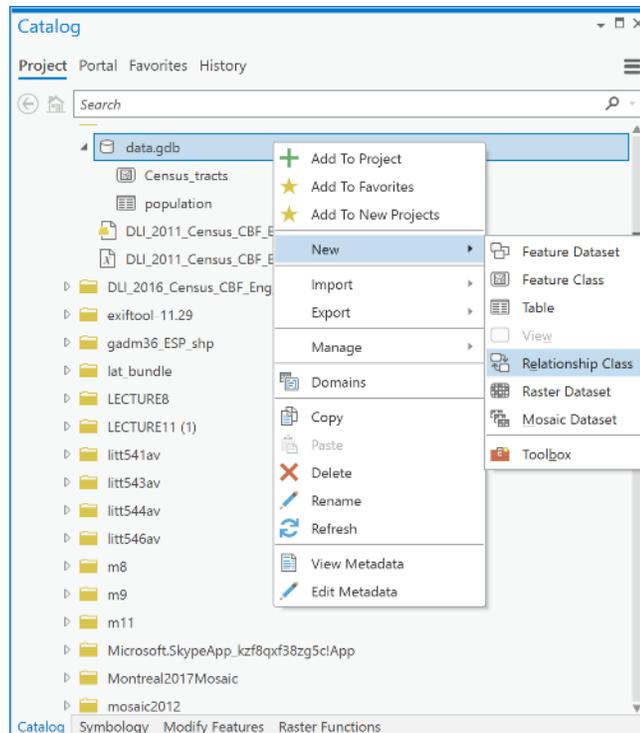
22. Vous verrez maintenant que votre tableau de population contient les CTUID correctement formatés dans le champ CTUID2 et que les valeurs sont cadrées à gauche, ce qui indique visuellement qu'il s'agit d'un champ de texte.

OBJECTID	CTUID	POP2011	POP2006	CTUID2
1	5050000	1236324	1133633	5050000.00
2	5050001.01	7741	7342	5050001.01
3	5050001.04	2849	2777	5050001.04
4	5050001.05	5875	5431	5050001.05
5	5050001.06	6238	5883	5050001.06
6	5050001.07	4324	4290	5050001.07
7	5050001.08	4613	4613	5050001.08
8	5050002.01	2893	2923	5050002.01
9	5050002.02	3511	3575	5050002.02
10	5050002.04	1363	1104	5050002.04
11	5050002.05	5090	4982	5050002.05
12	5050002.06	2673	2781	5050002.06
13	5050003	4564	4581	5050003.00
14	5050004	3829	3891	5050004.00
15	5050005	6107	5421	5050005.00
16	5050006	4090	4038	5050006.00
17	5050007.01	2254	2271	5050007.01
18	5050007.02	4953	4861	5050007.02
19	5050007.03	3719	3786	5050007.03
20	5050008	4597	4686	5050008.00
21	5050009	3280	3254	5050009.00
22	5050010	3787	3938	5050010.00

Q3: Comment peut-on relier la couche de données géospatiales Census_tracts à la table de données d'attributs contenant les informations sur la population ?

Une façon d'accéder aux champs de la table de population à partir de votre couche de secteurs de recensement est d'indiquer à ArcGIS comment faire correspondre les valeurs des champs CTUID de la couche de secteurs de recensement aux valeurs CTUID de la table de population. Pour cela, vous pouvez créer une **classe de relation** 1 à 1 dans la géodatabase, puis l'utiliser comme base pour interroger et sélectionner les données de recensement de la table de population dans la table d'attributs de Census_tracts.

1. Dans ArcCatalog, cliquez avec le bouton droit de la souris sur la géodatabase **data.gdb** et choisissez « *New* » (Nouveau) -> « *Relationship Class* » (Classe de relations):

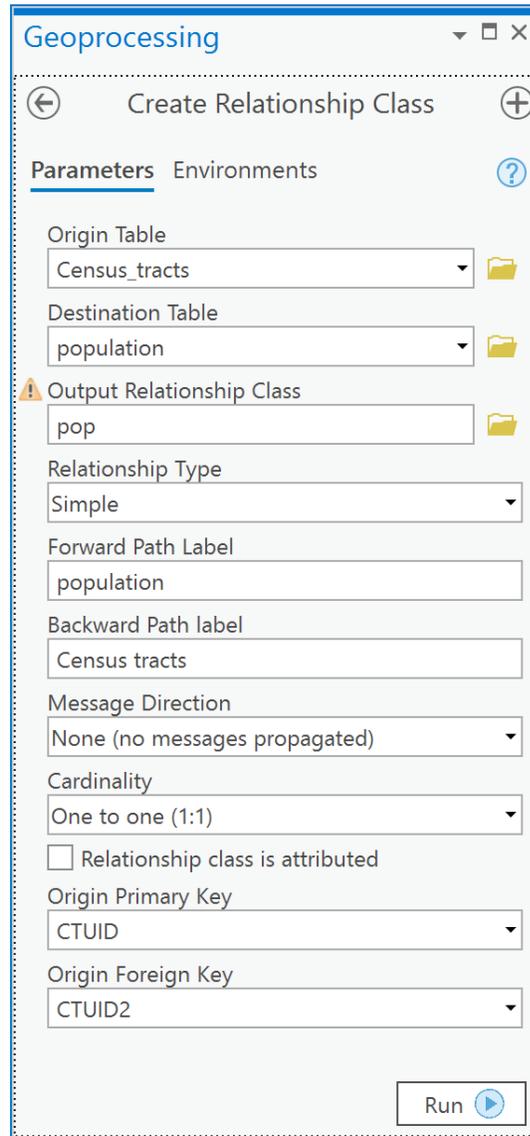


2. Dans la boîte de dialogue « *Create Relationship Class* » (Créer une classe de relations), définissez les éléments suivants:
 - **Origin Table:** Choisissez *Census_tracts*.
 - **Destination Table:** Choisissez *population*. Vous avez maintenant spécifié que la table *Census_tracts* contient les données sources de cette relation et que ces données sont liées à la table *population*.
 - **Output Relationship Class:** C'est le nom de la géodatabase sous lequel cette relation sera visible, appelez-le *pop*.
 - **Relationship Type:** Simple.
 - **Forward path label:** Tapez "population", qui spécifie l'étiquette lorsque vous naviguez de la classe d'origine (*Census_tracts*) à la classe de destination (*population*). Nous l'appelons donc "population" pour nous rappeler que la relation va de *Census_tracts* à *population*.
 - **Backward path label:** Tapez 'Census tracts'.
 - **Message Direction:** Choisissez '*None (no messages propagated)*', puisque cette classe de relation ne sera pas utilisée pour la modification d'éléments. Si elle était utilisée pour la modification, la suppression d'un élément dans une classe d'éléments enverrait un message à l'autre classe pour l'informer qu'il y a eu suppression. De telles opérations sont critiques et font partie de ce que l'on appelle l'intégrité référentielle : recherchez "referential integrity" dans le fichier d'aide d'ArcGIS Pro pour comprendre ce type de relation.
 - **Cardinality:** Ceci spécifie le type de relation, dans notre cas, chaque secteur de recensement peut avoir au maximum un seul enregistrement dans la table d'attributs de la *population*, nous choisissons donc une relation « *One to one (1:1)* ».
 - **Relationship class is attributed:** Nous n'avons pas besoin d'enregistrer de nouveaux attributs pour cette relation, c'est pourquoi la case « *Relationship class is attributed* » n'est pas cochée.

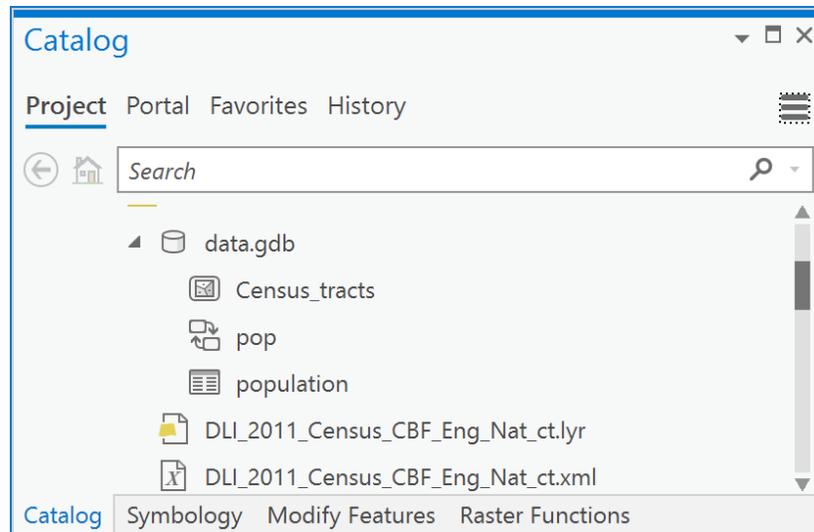
Nom de l'étudiant: _____ # d'étudiant: _____

- **Origin Primary Key:** Spécifier le champ commun qui contient un identifiant unique pour chaque unité spatiale dans la couche Census_tracts (origine), dans ce cas, il s'agit du champ CTUID.
- **Origin Foreign Key:** Spécifier le champ commun dans la table de population, donc sélectionner CTUID2 dans la liste déroulante.

3. Cliquez Run.



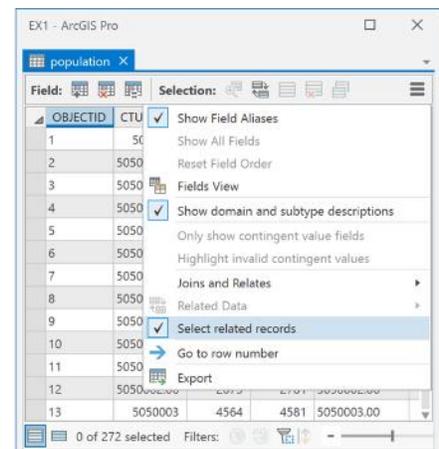
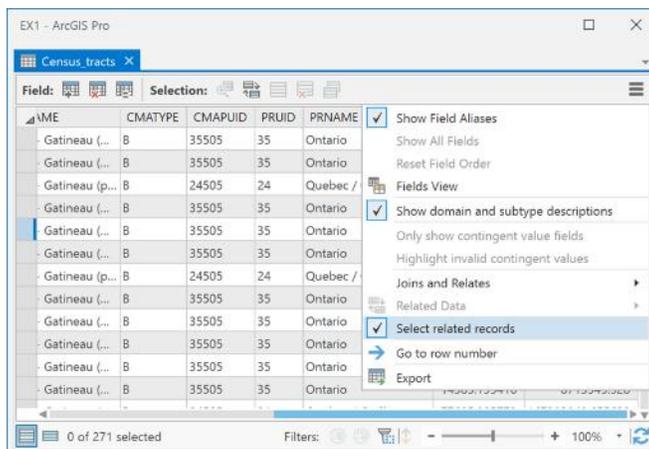
4. Examinez maintenant la géodatabase data.gdb et vous devriez voir l'ajout de l'élément de classe de relation appelé 'pop':



Q4: Que puis-je faire maintenant que j'ai établi une classe de relation entre ma table de population de recensement et ma classe d'entités de recensement ?

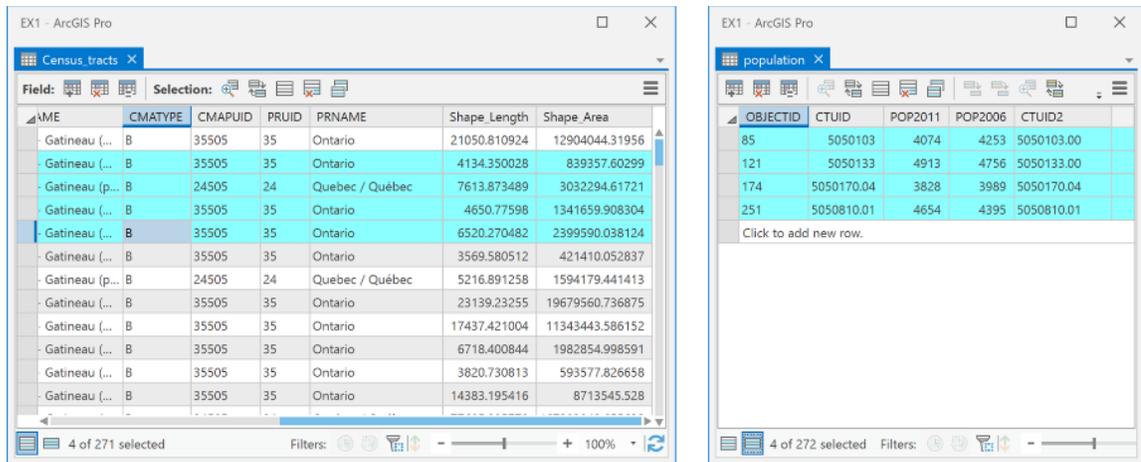
La classe de relation est semblable à une jointure de table, mais vous ne voyez pas les données de la table liée (population) dans la table Census_tracts. Il existe des limites à ce que l'on peut faire avec les classes de relations. Par exemple, vous pouvez effectuer des requêtes à l'aide de tables liées en SQL, mais vous ne pouvez pas utiliser les champs des tables participant à une classe de relation pour les calculs de champs dans la table d'origine. Voyons ce que nous pouvons faire:

1. Placez les tableaux **Census_tracts** et **population** côte à côte. Dans le menu, choisissez « *Select related records* » (sélectionner les enregistrements liés) dans les deux tableaux:

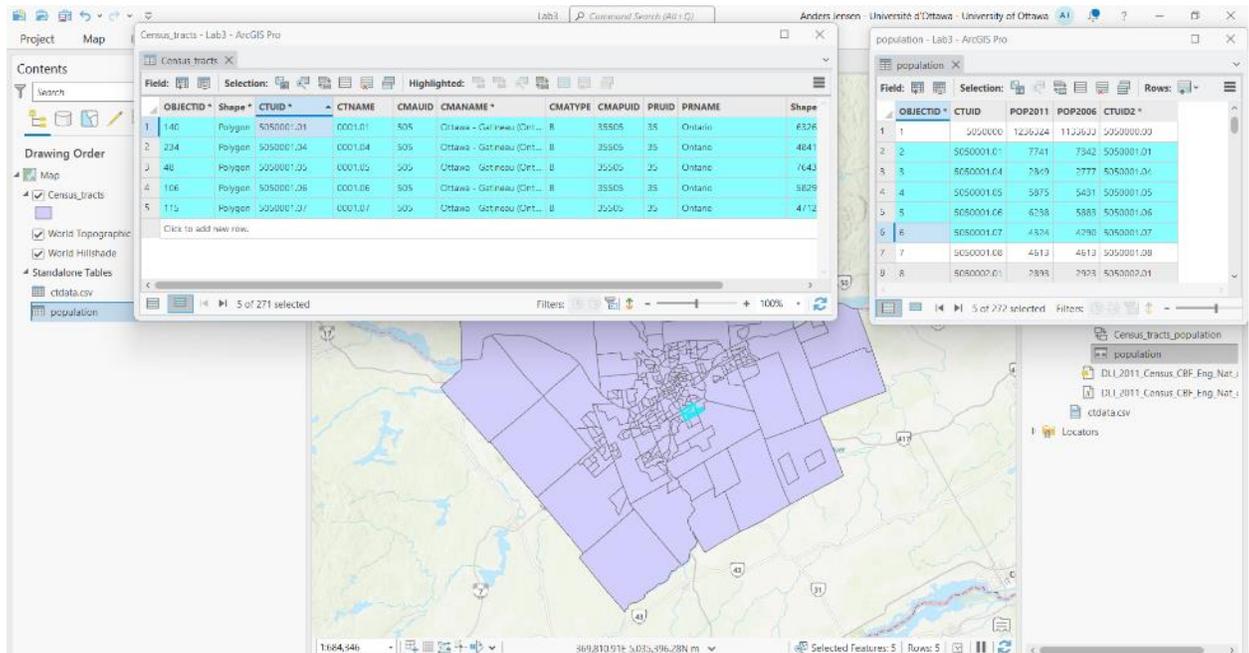


Nom de l'étudiant: _____ # d'étudiant: _____

2. Sélectionnez ensuite quelques rangées dans le tableau **Census_tracts** et examinez le tableau **population**. Vous constaterez que les données correspondantes sont sélectionnées dans le tableau de la population!

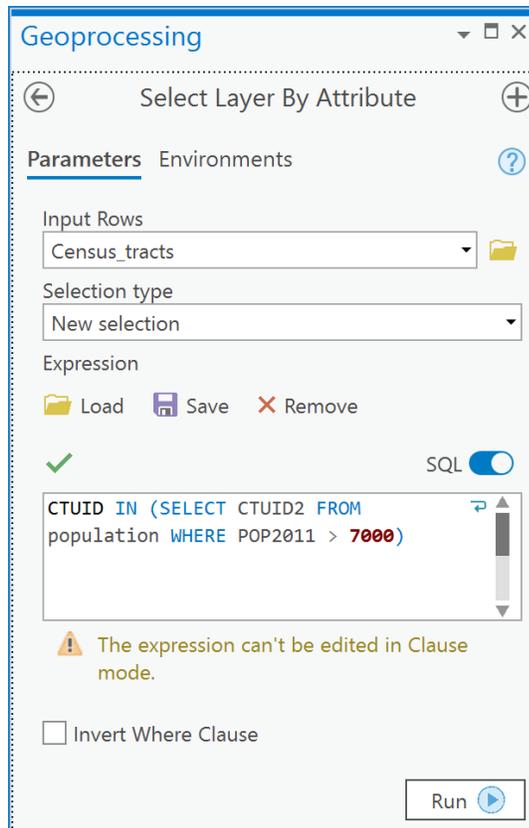


3. Pour obtenir l'effet inverse, sélectionnez quelques rangées dans le tableau population, puis choisissez dans le menu « *Related Data* » (Données connexes) -> **Census_tracts** - Census tracts. Vous verrez maintenant que ces caractéristiques seront sélectionnées dans le tableau **Census_tracts** et sur la carte.:



4. La classe de relation ajoute ce lien entre les deux tables d'attributs qui ont des valeurs communes dans les champs CTUID et CTUID2. Vous pouvez donc faire une requête dans la table **Census_tracts** à l'aide d'instructions SQL telles que celle-ci:

CTUID IN (SELECT CTUID2 FROM population WHERE POP2011 > 7000)

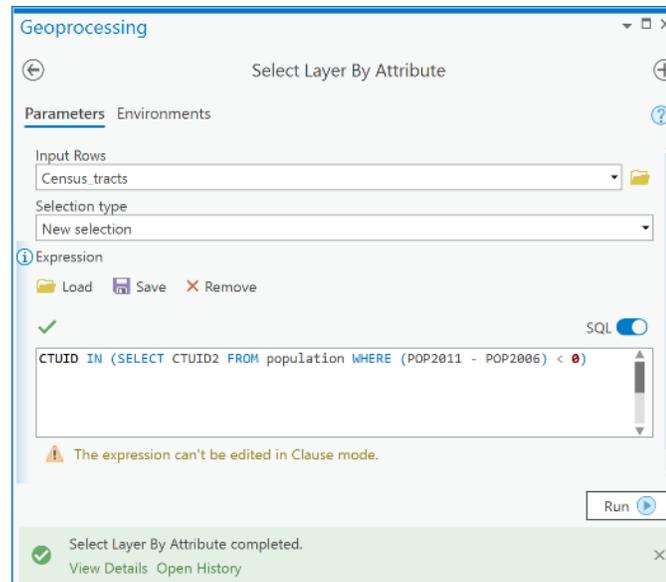


Vous verrez apparaître tous les secteurs de recensement sélectionnés dont la population est supérieure à 7000 personnes.

Note: L'utilisation de données de recensement qui sont présentées dans des zones polygonales telles que les secteurs de recensement comporte une mise en garde : on suppose implicitement que le nombre d'éléments, tels que les personnes, est uniformément réparti sur le territoire de chaque secteur. En réalité, ce n'est pas le cas, car les gens ne peuvent pas vivre partout dans un secteur de recensement. Par exemple, les gens ne vivent pas dans les forêts ou dans les rues, car ils vivent dans des maisons. La cartographie dasymétrique permet d'améliorer la précision des cartes en ne montrant que les données où les gens peuvent vivre plutôt que partout. Nous abordons ce sujet dans un cours plus avancé.

5. Vous pouvez également utiliser des requêtes SQL plus complexes qui évaluent des expressions arithmétiques sur les champs de la table liée. Vous souhaitez peut-être sélectionner tous les secteurs de recensement dont la population a diminué entre 2006 et 2011:

`CTUID IN (SELECT CTUID2 FROM population WHERE (POP2011 - POP2006) < 0)`

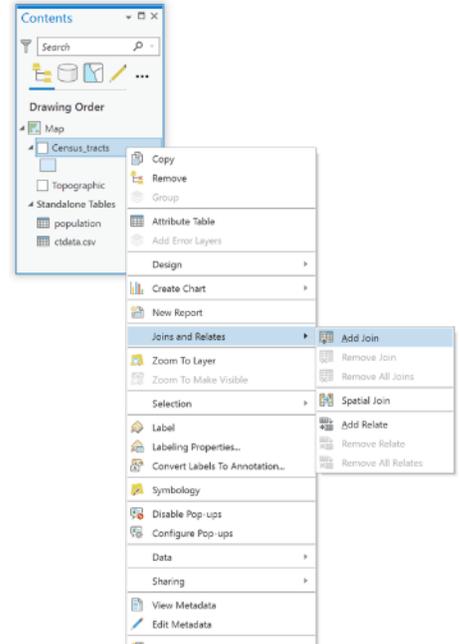
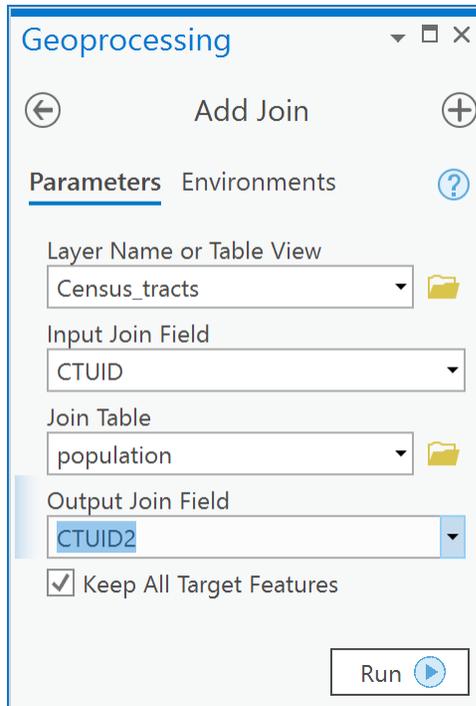


Q5: Comment puis-je cartographier la population dans les secteurs de recensement à l'aide de mon tableau de population?

Les classes de relation sont limitées lorsqu'il s'agit de symboliser des champs dans des tables liées. Nous ne pouvons donc pas mapper la population dans la couche **Census_tracts** à l'aide des classes de relation. Nous devons donc nous appuyer sur des jointures de tables pour ajouter les attributs de la table **population** à la couche **Census_tracts**. Vous pouvez rapidement effectuer une jointure ad hoc de la table de population à la couche des secteurs de recensement afin d'avoir un accès direct aux champs POP2011 et POP2006, de sorte qu'ils soient visibles dans la couche des secteurs de recensement comme s'ils faisaient partie de la table d'attributs de la couche des secteurs de recensement - mais sachez qu'il ne s'agit que d'une jointure virtuelle lorsque vous utilisez l'outil « *Add Join* » (Ajouter une jointure).

1. Cliquez avec le bouton droit de la souris sur la couche des secteurs de recensement dans le volet Contenu et choisissez « *Joins and Relates* » (jointures et relations) -> « *Add Join* » (ajouter une jointure):

- In the Joins & Relates dialog box, under 'Layer Name or Table View', specify Census_tracts. Under 'Input Join Field' choose CTUID. Under 'Join Table' choose population and finally under 'Output Join Field' choose CTUID2. Click OK.



- Ouvrez la table d'attributs **Census_tracts**, faites défiler vers la droite et vous devriez voir les champs POP2011 et POP2006 ajoutés à la fin de la liste des champs de la table d'attributs:

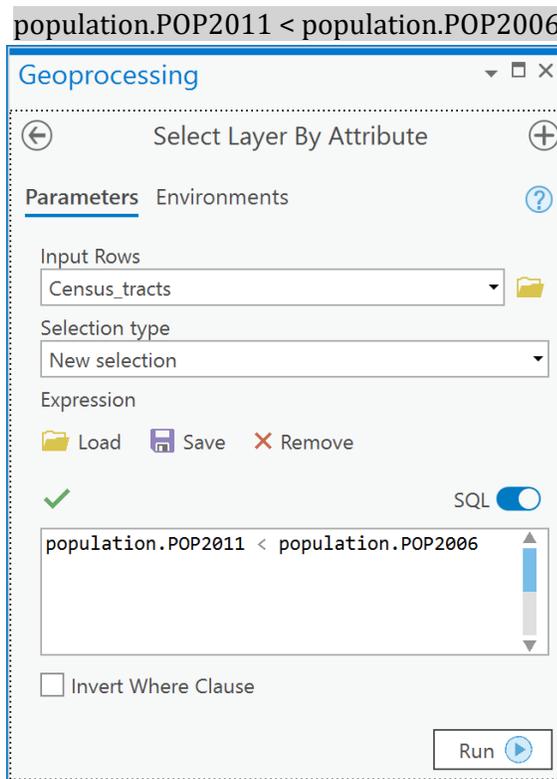
Colonnes jointes du tableau de population

OBJECTID	Shape	CTUID	CTNAME	CMAUID	CMANAME	CMATYPE	CMAPUID	PRUID	PRNAME	Shape_Length	Shape_Area	OBJECTID2	CTUID2	POP2011	POP2006	CTUID2
1	Polygon	5050151.08	0151.08	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	21050.810924	12904044.3195	155	5050151.08	5571	3395	5050151.08
2	Polygon	5050103.00	0103.00	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	4134.350028	839357.6029	85	5050103	4074	4253	5050103.00
3	Polygon	5050810.01	0810.01	505	Ottawa - Gatineau (p...	B	24505	24	Quebec / Québec	7613.873489	3032294.6172	251	5050810.01	4654	4395	5050810.01
4	Polygon	5050170.04	0170.04	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	4650.77598	1341659.90830	174	5050170.04	3828	3989	5050170.04
5	Polygon	5050133.00	0133.00	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	6520.270482	2399590.03812	121	5050133	4913	4756	5050133.00
6	Polygon	5050007.02	0007.02	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	3569.580512	421410.05283	18	5050007.02	4953	4861	5050007.02
7	Polygon	5050610.32	0610.32	505	Ottawa - Gatineau (p...	B	24505	24	Quebec / Québec	5216.891258	1594179.44141	229	5050610.32	2153	2237	5050610.32
8	Polygon	5050127.00	0127.00	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	23139.232255	19679560.73687	115	5050127	3902	3548	5050127.00
9	Polygon	5050160.03	0160.03	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	17437.421004	11343443.58615	158	5050160.03	4869	3215	5050160.03
10	Polygon	5050125.05	0125.05	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	6718.400844	1982854.99859	106	5050125.05	5195	5497	5050125.05
11	Polygon	5050056.00	0056.00	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	3820.730813	593577.82665	74	5050056	3906	3700	5050056.00
12	Polygon	5050141.08	0141.08	505	Ottawa - Gatineau (...)	B	35505	35	Ontario	14383.195416	8713545.52	142	5050141.08	8632	8391	5050141.08
13	Polygon	5050904.00	0904.00	505	Ottawa - Gatineau (p...	B	24505	24	Quebec / Québec	77605.085772	167303842.65563	271	5050904	677	676	5050904.00

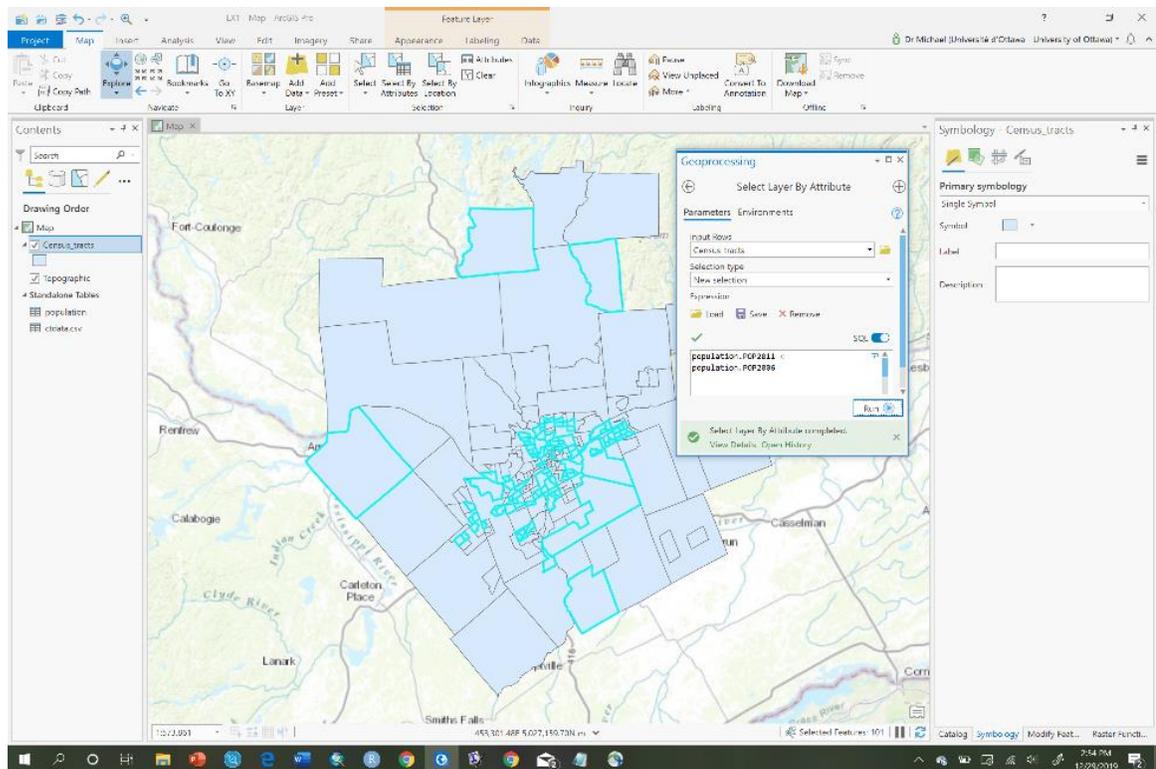
- Vous pouvez maintenant effectuer des requêtes, cartographier ou calculer des champs en utilisant ces enregistrements ajoutés dans la table d'attributs Census_tracts. Par exemple, en utilisant l'outil « *Select by Attributes* » (Sélection par attributs) dans l'onglet « *Map* » (Carte) pour ouvrir l'outil « *Select Layer by Attribute* » (Sélection de la couche par attributs) et, par exemple, montrer quels sont les secteurs de recensement dont la

Nom de l'étudiant: _____ # d'étudiant: _____

population a diminué entre 2011 et 2006 (notez qu'il s'agit simplement d'une autre façon de faire ce que nous avons fait à la Q5). Vous pouvez effectuer une requête pour telle que:



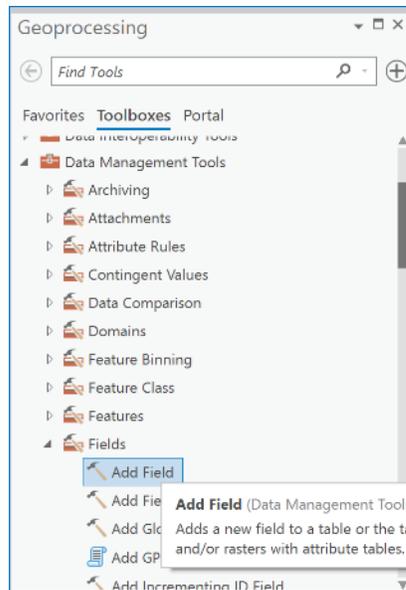
4. Vous verrez maintenant la sélection effectuée dans la table d'attributs des secteurs de recensement et la carte.



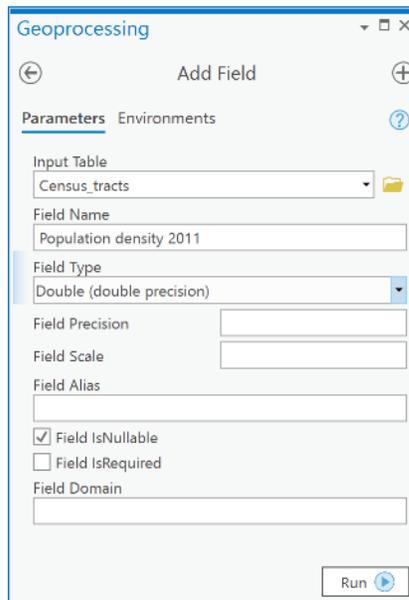
Q6: Je veux calculer la densité de population ; comment puis-je le faire?

Pour calculer des valeurs à l'aide de champs existants, vous devez ajouter un nouveau champ à votre table d'attributs. Par exemple, si je veux calculer la densité de la population en 2011 dans chaque secteur de recensement, j'ajouterai un champ de type Double (parce que la densité est un ratio et que les ratios seront probablement des nombres réels) à la table d'attributs, puis je calculerai la densité de la population en tant que nombre de personnes par kilomètre carré,

1. Allez dans « *Data Management Tools* » (Outils de gestion des données) -> « *Field* » (Champ)-> « *Add Field* » (Ajouter un champ),

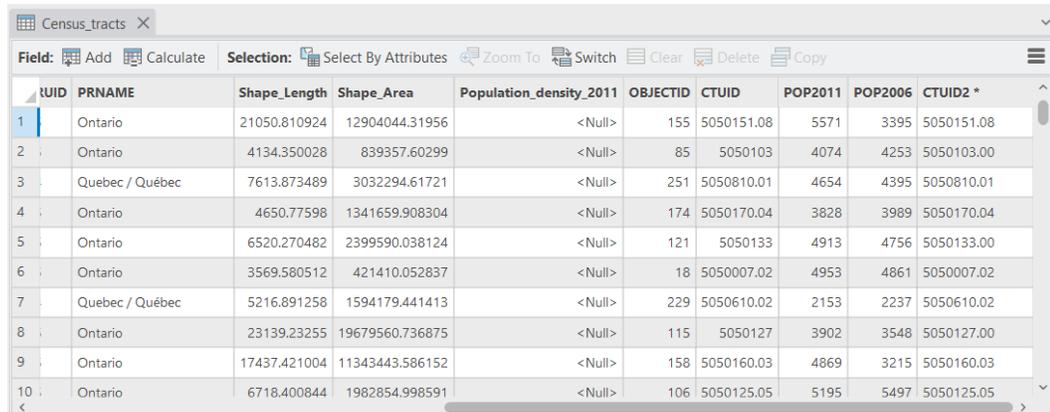


2. Dans la boîte de dialogue « *Add Field* » (Ajouter un champ), choisissez votre couche de secteurs de recensement dans le menu déroulant « *Input Table* » (Tableau d'entrée), puis, sous « *Field Name* » (Nom du champ), tapez POPULATION DENSITY 2011 et choisissez « *Field Type* » (Type de champ) comme *Float* ou *Double*, et cliquez sur Run:



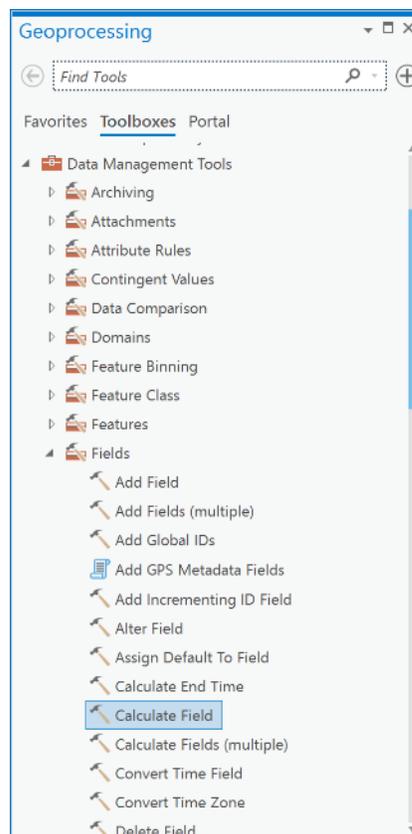
3. Examinez votre table d'attributs des secteurs de recensement, vous devriez voir un nouveau champ appelé POPULATION_DENSITY_2011 avec toutes les valeurs <Null>. Remarquez qu'ArcGIS a remplacé les espaces du nom que vous avez saisi par des caractères de soulignement "_". Ceci s'explique par le fait que les espaces ne sont pas autorisés dans les noms de champs des bases de données.

Nom de l'étudiant: _____ # d'étudiant: _____

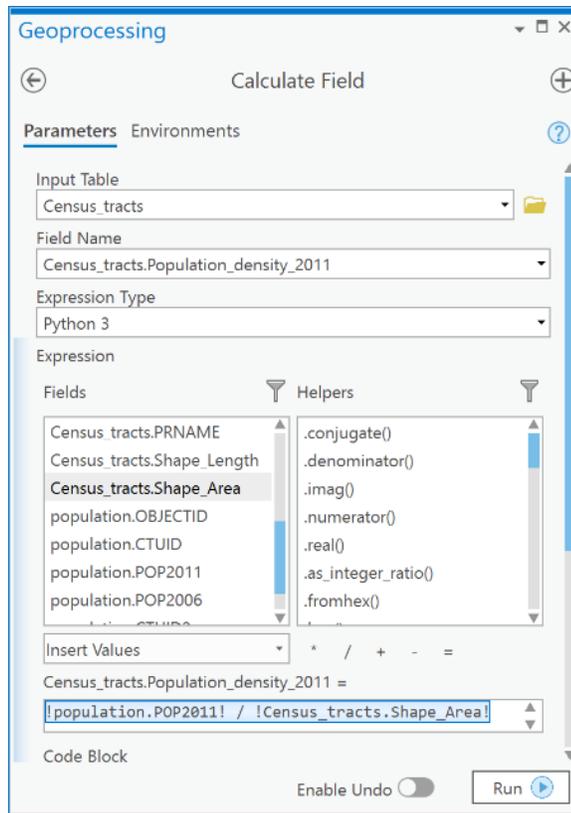


UID	PRNAME	Shape_Length	Shape_Area	Population_density_2011	OBJECTID	CTUID	POP2011	POP2006	CTUID2 *
1	Ontario	21050.810924	12904044.31956	<Null>	155	5050151.08	5571	3395	5050151.08
2	Ontario	4134.350028	839357.60299	<Null>	85	5050103	4074	4253	5050103.00
3	Quebec / Québec	7613.873489	3032294.61721	<Null>	251	5050810.01	4654	4395	5050810.01
4	Ontario	4650.77598	1341659.908304	<Null>	174	5050170.04	3828	3989	5050170.04
5	Ontario	6520.270482	2399590.038124	<Null>	121	5050133	4913	4756	5050133.00
6	Ontario	3569.580512	421410.052837	<Null>	18	5050007.02	4953	4861	5050007.02
7	Quebec / Québec	5216.891258	1594179.441413	<Null>	229	5050610.02	2153	2237	5050610.02
8	Ontario	23139.23255	19679560.736875	<Null>	115	5050127	3902	3548	5050127.00
9	Ontario	17437.421004	11343443.586152	<Null>	158	5050160.03	4869	3215	5050160.03
10	Ontario	6718.400844	1982854.998591	<Null>	106	5050125.05	5195	5497	5050125.05

4. Ensuite, assurez-vous que rien n'est sélectionné dans votre table d'attributs Census_tracts en cliquant sur le bouton « Clear » (Effacer) dans le groupe « Selection » de l'onglet « Map » (Carte) d'ArcGIS Pro.
5. Ensuite, ouvrez l'outil « Calculate Field » (Calculer le champ) dans ArcToolbox sous « Data Management Tools » (Outils de gestion des données) -> « Fields » (Champs):



6. Dans la boîte de dialogue « Calculate Field », (i) choisissez votre couche de secteurs de recensement sous « Input Table », (ii) sous « Field Name », choisissez le champ Population_density_2011, (iii) entrez l'expression: **!population.POP2011!
/ !Census_tracts.Shape_Area!**, et (iv) cliquez sur Run.



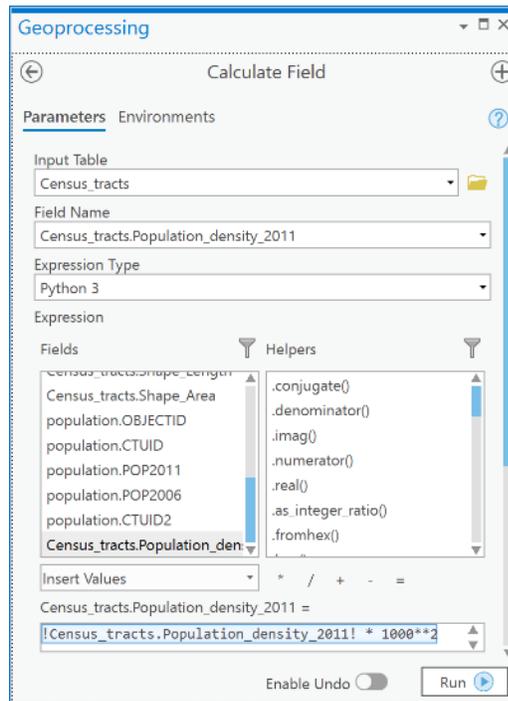
7. Ouvrez votre table d'attributs des secteurs de recensement, supprimez toutes les jointures et examinez la densité de la population:

	OBJECTID	CTUID	POP2011	POP2006	CTUID2	Population density 2011
56	155	5050151.08	5571	3395	5050151.08	0.000432
99	85	5050103	4074	4253	5050103.00	0.004854
21	251	5050810.01	4654	4395	5050810.01	0.001535
04	174	5050170.04	3828	3989	5050170.04	0.002853
24	121	5050133	4913	4756	5050133.00	0.002047
37	18	5050007.02	4953	4861	5050007.02	0.011753
13	229	5050610.02	2153	2237	5050610.02	0.001351
75	115	5050127	3902	3548	5050127.00	0.000198
52	158	5050160.03	4869	3215	5050160.03	0.000429
91	106	5050125.05	5195	5497	5050125.05	0.00262
58	74	5050056	3906	3700	5050056.00	0.00658
28	142	5050141.08	8632	8391	5050141.08	0.000991

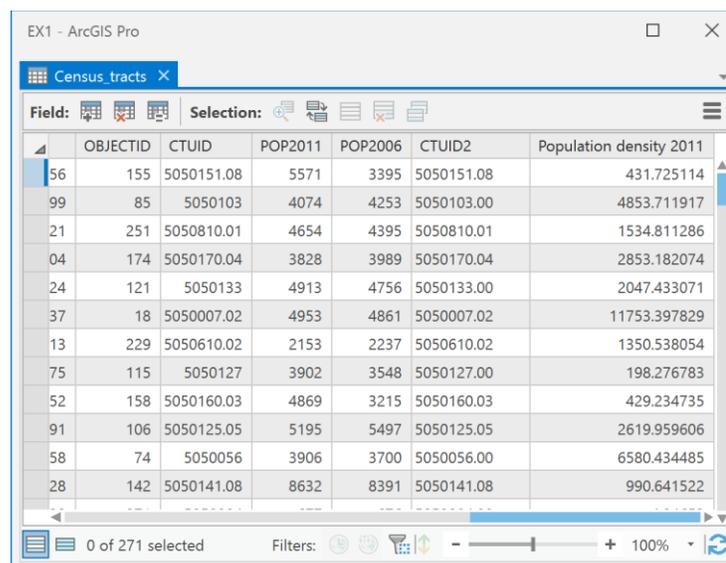
8. Les valeurs de la densité de population sont exprimées en nombre de personnes par mètre carré, ce qui n'est pas une façon normale de représenter la densité de population ; en général, les valeurs sont exprimées en nombre de personnes par kilomètre carré.

Nom de l'étudiant: _____ # d'étudiant: _____

Pour changer cela, nous allons multiplier ce champ par 1000^2 , ou 1 000 000, parce qu'il y a 1000^2 m² dans un 1 km². Dans la boîte de dialogue « *Calculate Field* », (i) choisissez votre couche de secteurs de recensement sous « *Input Table* », (ii) sous « *Field Name* », choisissez le champ **Census_tracts.Population_density_2011**, (iii) entrez l'expression **!Census_tracts.Population_density_2011! * 1000**2**, et (iv) cliquez sur Run.

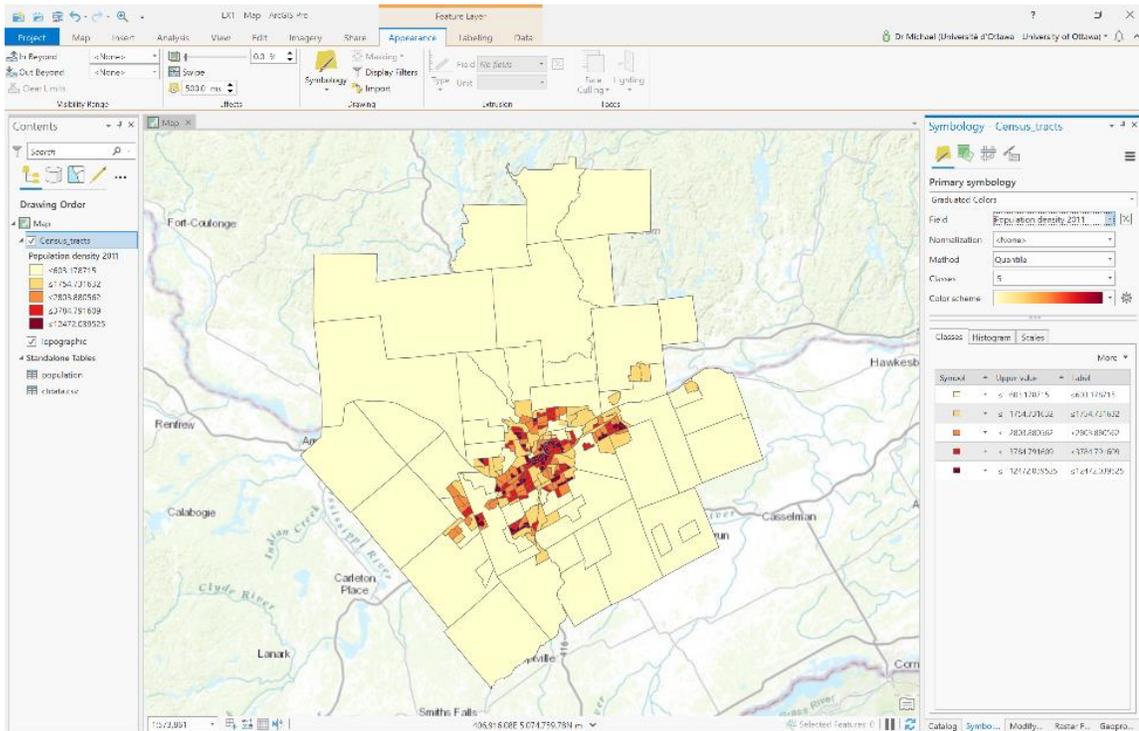


9. Examinez maintenant le tableau des attributs des secteurs de recensement et vous verrez que les valeurs de la densité de population sont maintenant plus raisonnables (et interprétables) en tant que personnes par kilomètre carré:



OBJECTID	CTUID	POP2011	POP2006	CTUID2	Population density 2011
56	155 5050151.08	5571	3395	5050151.08	431.725114
99	85 5050103	4074	4253	5050103.00	4853.711917
21	251 5050810.01	4654	4395	5050810.01	1534.811286
04	174 5050170.04	3828	3989	5050170.04	2853.182074
24	121 5050133	4913	4756	5050133.00	2047.433071
37	18 5050007.02	4953	4861	5050007.02	11753.397829
13	229 5050610.02	2153	2237	5050610.02	1350.538054
75	115 5050127	3902	3548	5050127.00	198.276783
52	158 5050160.03	4869	3215	5050160.03	429.234735
91	106 5050125.05	5195	5497	5050125.05	2619.959606
58	74 5050056	3906	3700	5050056.00	6580.434485
28	142 5050141.08	8632	8391	5050141.08	990.641522

10. Maintenant, symbolisez la couche des secteurs de recensement par la densité de la population à l'aide d'une carte à couleurs graduées. Cliquez avec le bouton droit de la souris sur la couche des secteurs de recensement dans la table des matières et choisissez Symbology. Sélectionnez « *Graduated Colors* » comme « *Primary Symbology* », puis définissez le « *Field* » comme Population_Density_2011, et sous « *Method* » sélectionnez une classification Quantile. Le résultat devrait ressembler à ceci:



Questions à remettre

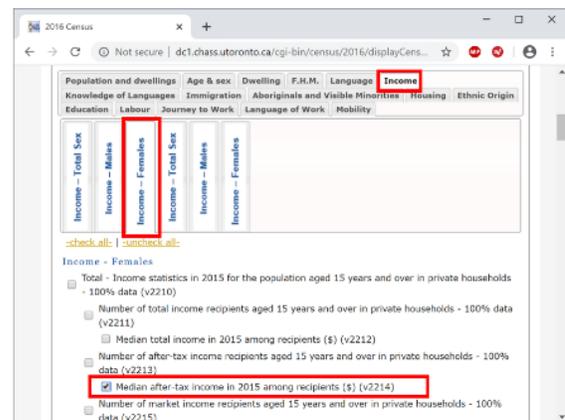
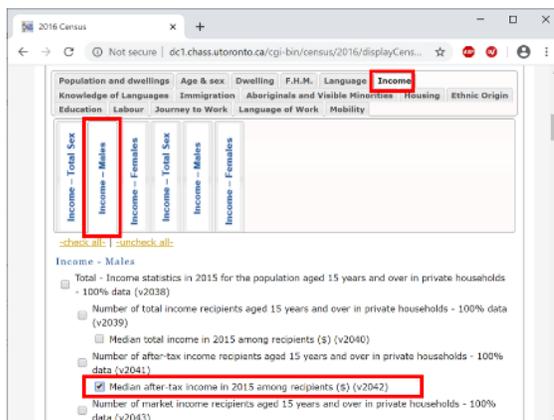
Note: Certaines de ces questions nécessitent une explication, alors lisez attentivement. Certaines questions requièrent également une carte comme élément de réponse et celle-ci doit être présentée correctement. Veillez donc revoir le cours de formation ESRI "Creating a Map Layout" (Création d'une présentation cartographique). Les points seront déduits en cas de présentation incorrecte (n'oubliez pas non plus que les barres d'échelle doivent être métriques).

Note: Pour certaines questions, vous devez télécharger des données à partir du Scholars GeoPortal ou de l'Analyseur de recensement canadien. Si vous n'avez pas accès à ces ressources, vous pouvez trouver les fichiers de données pertinents dans lab3.zip.

1. [2 points] Comment une classe de relation peut-elle assurer l'intégrité référentielle dans la géodatabase?
2. [2 points] Dans Q6, vous avez calculé la densité de population dans la couche Census_tracts. Dans quel système de coordonnées la surface de chaque secteur de recensement a-t-elle été calculée ? Comment le savez-vous? Le système de

coordonnées que vous avez choisi est-il un bon choix pour le calcul de la superficie ? Pourquoi ou pourquoi pas?

3. [1 points] Pensez-vous que la carte de la densité de la population est spatialement précise en ce qui concerne l'endroit où les gens sont regroupés ? Pourquoi ou pourquoi pas?
4. [9 points] Téléchargez les secteurs de recensement de 2016 pour Ottawa-Gatineau. Téléchargez ensuite les variables suivantes pour la région d'Ottawa-Gatineau à partir des tableaux de 2016 de l'Analyseur du recensement canadien (voir Q2) (notez que pour les questions 5 et 6 ci-dessous, vous devez également télécharger d'autres variables, vous pouvez préférer les télécharger toutes en même temps, c'est vous qui décidez):



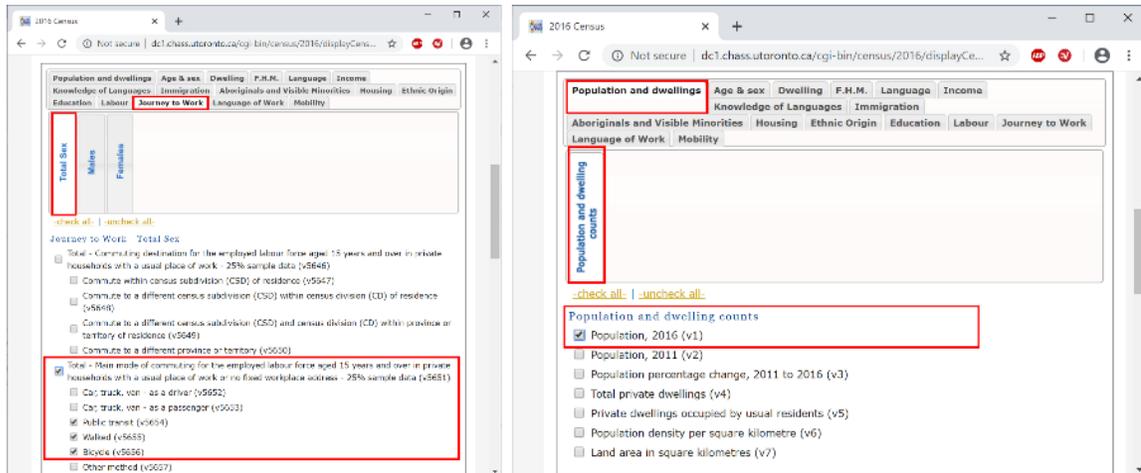
Joindre votre table de données à votre couche de secteurs de recensement pour 2016 (voir Q5 pour cette procédure). Répondez aux questions suivantes [1 point chacune], seule la dernière (i) nécessite la création d'une carte:

- a. Dans *combien* de secteurs de recensement les hommes gagnent-ils moins que le revenu médian après impôt de 35.000\$? (vous avez besoin d'une requête pour répondre à cette question, voir Q5 pour des exemples de requêtes)
- b. Dans *combien* de secteurs de recensement les femmes gagnent-elles moins que le revenu médian après impôts de 35.000\$
- c. Dans *combien* de secteurs de recensement les hommes ont-ils un revenu médian après impôts supérieur à celui des femmes?
- d. Dans *combien* de secteurs de recensement les femmes ont-elles un revenu médian après impôt supérieur à celui des hommes?
- e. Dans *combien* de secteurs de recensement les femmes ET les hommes ont-ils un revenu médian après impôt inférieur à 35 000 \$?
- f. Dans *combien* de secteurs de recensement les hommes ont-ils un revenu médian après impôt compris entre 35 000 \$ et 45 000 \$?
- g. Dans *combien* de secteurs de recensement les femmes ont-elles un revenu médian après impôt compris entre 35 000 \$ et 45 000 \$?
- h. Dans combien de secteurs de recensement les femmes ET les hommes ont-ils un revenu médian après impôt compris entre 35 000 \$ et 45 000 \$?
- i. Décrivez la distribution spatiale du ratio du revenu médian masculin/féminin après impôt (créez un nouveau champ et calculez le ratio ou utilisez le revenu

Nom de l'étudiant: _____ # d'étudiant: _____

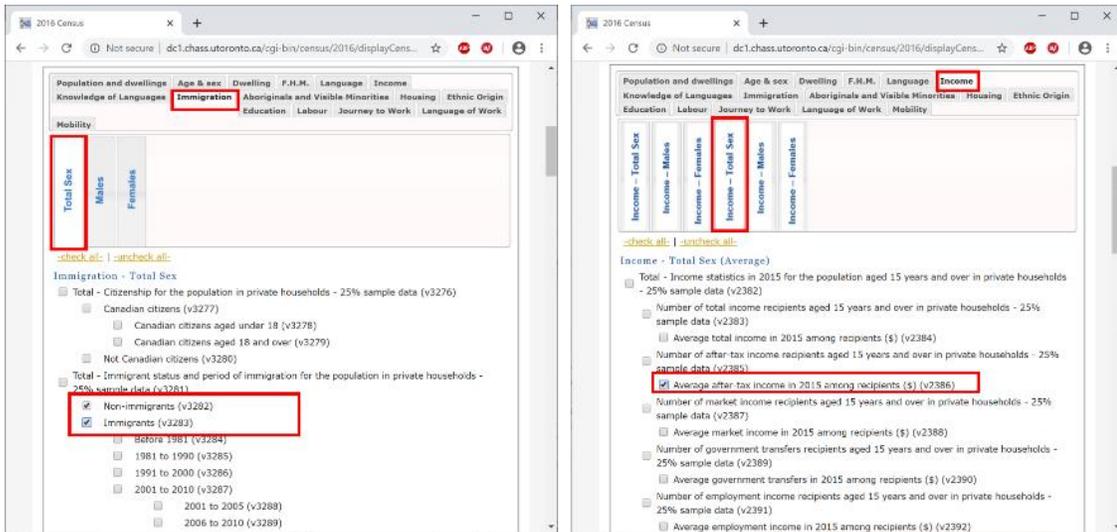
médian féminin comme champ de normalisation dans la symbologie pour les couleurs graduelles). Pour cette question, présentez une carte en tant que figure.

5. [6 points] Téléchargez les variables suivantes à partir des tableaux de 2016 de l'Analyseur du recensement canadien (voir Q2):



Joindre votre table de données à votre couche de secteurs de recensement pour 2016 (voir Q5 pour cette procédure). Répondez aux questions suivantes [1 point chaque]:

- Dans combien de secteurs de recensement le vélo est-il plus utilisé que la marche comme mode de transport pour se rendre au travail?
 - Dans combien de secteurs de recensement la marche est-elle plus utilisé que le vélo comme mode de transport pour se rendre au travail?
 - Quel est le revenu moyen après impôt des secteurs de recensement où davantage de personnes se rendent au travail à pied qu'à vélo?
 - Quel est le revenu moyen après impôt des secteurs de recensement où davantage de personnes se rendent au travail à vélo qu'à pied?
 - Créez une carte montrant le ratio vélo/marche et expliquez la tendance que vous observez. (Vous avez besoin d'une carte et d'une explication).
 - Créez une carte montrant la proportion de la population totale qui se rend au travail à vélo et expliquez la tendance que vous observez.
6. [7 points] Téléchargez les variables suivantes à partir des tableaux de 2016 de l'Analyseur du recensement canadien (voir Q2):



Joignez votre table de données à votre couche de secteurs de recenseurs pour 2016 (voir Q5 pour cette procédure). Répondez aux questions suivantes [1 point chaque]:

- Dans combien de secteurs de recensement y a-t-il plus d'immigrés que de non-immigrés?
- Dans combien de secteurs de recensement le ratio immigrés/non-immigrés est-il supérieur à 0,25?
- Dans combien de secteurs de recensement le ratio immigrés/non-immigrés est-il supérieur à 0,5?
- Dans combien de secteurs de recensement le ratio immigrés/non-immigrés est-il supérieur à 0.4?
- Quel est le revenu moyen des secteurs de recensement dont le ratio de population immigrée/non-immigrée est supérieur à 0.4?
- Quel est le revenu moyen des secteurs de recensement dont le ratio de population immigrée/non-immigrée est inférieur à 0.4?
- Quelle est la tendance du ratio entre la population immigrée et la population non immigrée ?

Laboratoire #4 : Analyse spatiale avec vecteur

Objectif : L'objectif de ce laboratoire est de vous fournir une introduction aux concepts de base de l'analyse spatiale et aux capacités de la science des systèmes d'information géographique.

Résultats d'apprentissage: À la fin de ce laboratoire, les élèves seront capables de:

- Exécuter une requête spatiale.
- Différencier une requête spatiale d'une requête attributaire.
- Effectuer une requête basée sur la date.
- Évaluer les mesures centrographiques d'un ensemble de points.
- Créer et expliquer une estimation de la densité du noyau (KDE) d'un ensemble de points.
- Utiliser les opérations point-dans-polygone pour calculer la densité.
- Opérationnaliser la superposition d'unions (fonction Agréger).

Procédure:

- Lisez et travaillez sur les exemples.
- Répondez aux questions à la fin du laboratoire.

Données/matériel:

- Fourni avec le plan du laboratoire

Devoir:

- Répondez aux questions (#1-5).
- Le total des points disponibles pour ce travail est de 15 points, soit 12,5% de la note totale du cours.
- Tous les travaux pratiques doivent être réalisés et remis individuellement.

Format:

Les étudiants remettront les questions avec les réponses complétées.

Les tableaux et les figures doivent être produits selon les normes du Département de géographie, environnement et géomatique de l'Université d'Ottawa. Pour des directives sur le formatage et la présentation appropriés des tableaux et des figures, veuillez consulter la [section d'introduction](#) de ce manuel de laboratoire.

Cette section présente des exemples de choses que vous devrez savoir pour répondre aux questions que vous remettrez pour obtenir des notes. Votre tâche consiste à examiner et à répéter les exemples de cette section, puis à déterminer lesquels sont pertinents pour vous aider à répondre aux questions qui se trouvent à la fin de l'exercice.

Q1: Comment effectuer une requête spatiale?

Vous avez appris à connaître le langage de requête structuré (SQL) et à utiliser des requêtes pour répondre à des questions sur les données d'attribut à l'aide de la fonction « *Select Layer By Attribute* » (Sélectionner une couche par attribut) dans ArcGIS. Une requête spatiale utilise la boîte de dialogue « *Select Layer By Location* » (Sélectionner une couche par localisation) et généralement deux couches spatiales ou plus dans le même espace. Par exemple, supposons que nous ayons deux couches de points : une couche appelée **Crime**¹, qui montre les emplacements des crimes et la seconde couche appelée **Schools**, qui montre l'emplacement des écoles (Figure 1):

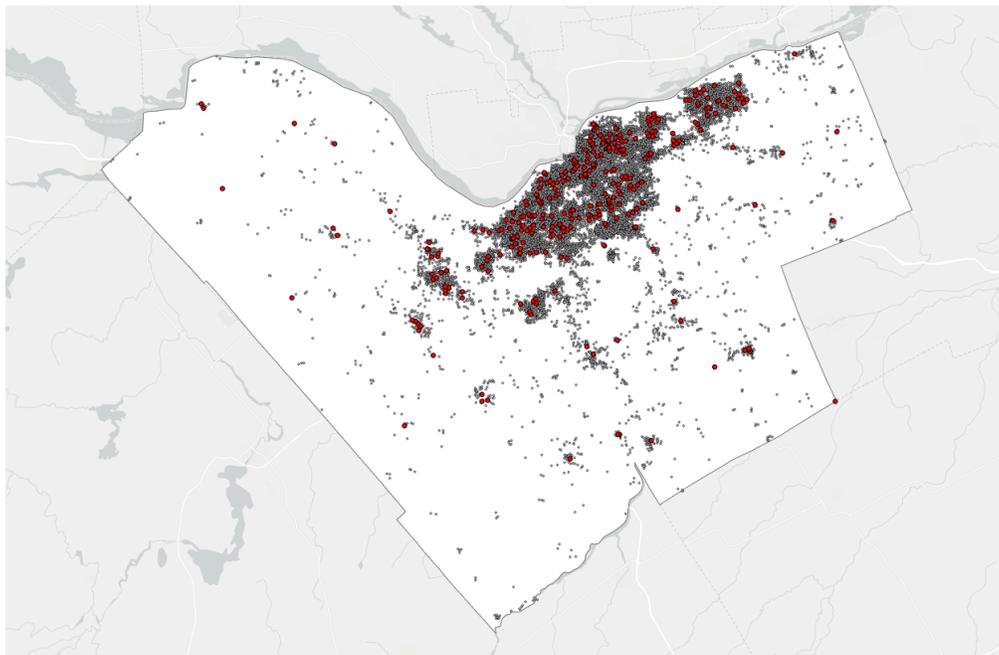
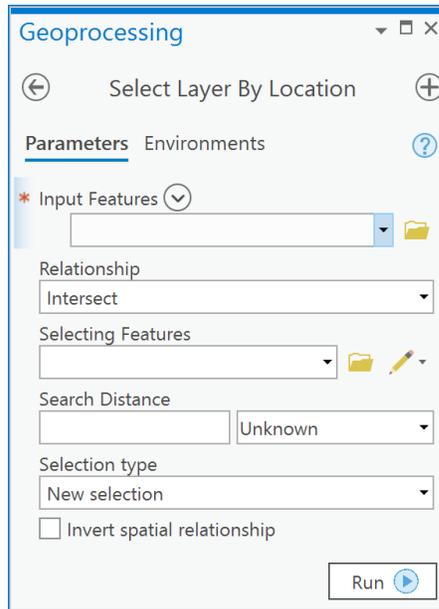


Figure 6: Deux couches à utiliser pour les requêtes spatiales et les requêtes d'attributs, représentées par des points de couleur différente.

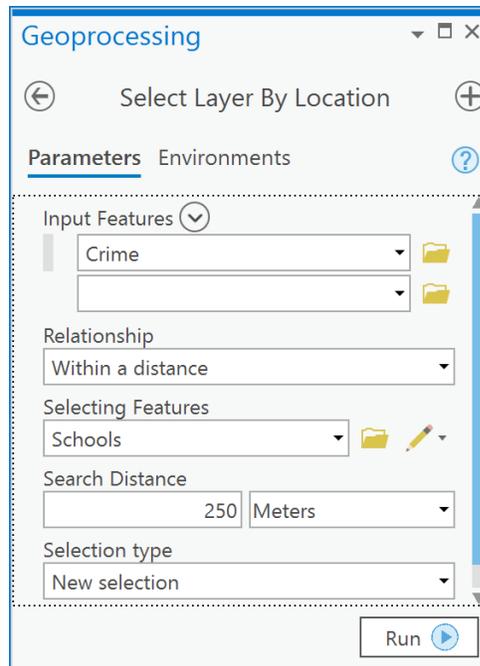
Nous pourrions vouloir sélectionner tous les crimes commis dans un rayon de 250 m autour des écoles d'Ottawa. La manière la plus efficace d'effectuer une telle tâche est d'utiliser l'outil « *Select Layer By Location* » (Sélectionner la couche par localisation) pour sélectionner tous les points de criminalité situés à moins de 250 m des écoles.

¹ La couche Crime est une couche synthétique représentant les appels de service (lorsque quelqu'un appelle le 911) et ne représente pas les crimes réels commis aux endroits spécifiés. À petite échelle, le jeu de données est représentatif des tendances réelles de la criminalité à Ottawa au cours de la période 1998-2001. Il a été créé en utilisant des surfaces de probabilité dérivées des appels de service réels au cours de la période spécifiée. Il n'est pas précis au niveau local.

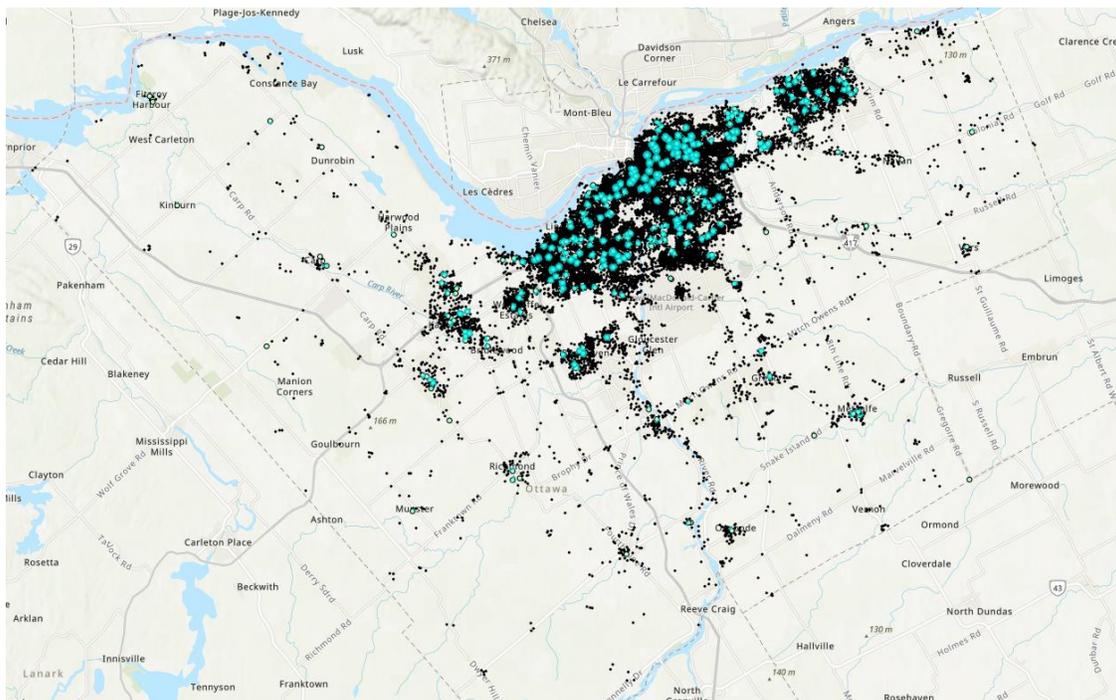
1. Ajouter **Crime** et **Schools** à ArcGIS Pro (à partir du jeu de données Ottawa dans la géodatabase OTTAWADATA dans le dossier Lecture7).
2. Dans l'onglet « *Map* » (Carte), trouvez le groupe "Selection" et cliquez sur « *Select Layer By Location* » (Sélectionner une couche par emplacement) pour ouvrir l'outil:



3. Dans la boîte de dialogue, les caractéristiques d'entrée (« *Input Features* ») sont la cible de la sélection et ici, comme nous voulons sélectionner tous les crimes commis dans un rayon de 250 m autour d'une école, la classe de données **Crime** doit être sélectionnée sous « *Input Features* ». Sous « *Relationship* », choisissez « *Within a distance* » (Parmi une distance). Sous « *Selecting Features* » (Sélection de données), choisissez **Schools**. Entrez 250 dans la zone de saisie « *Search Distance* » (Distance de recherche). Cliquez sur « *Run* » (Exécuter).



4. Le résultat devrait ressembler à ceci:



Vous venez d'entreprendre une requête spatiale. Vous avez utilisé une relation spatiale - "A une certaine distance" (Within a distance) - entre deux classes d'entités dans le même domaine spatial pour sélectionner des entités dans l'une des classes d'entités participantes. L'opération « *Within a distance* » « crée des zones tampon à l'aide de la distance de zone

tampon autour des entités source et renvoie toutes les entités qui intersectent les zones de tampon ». (Aide ESRI).

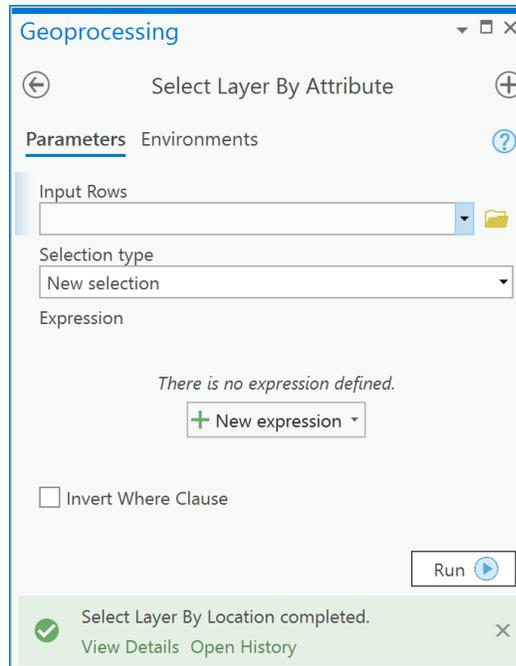
Q2: Quel est le lien entre les requêtes SQL et les requêtes spatiales (« Select layer by location »)?

L'utilisation combinée de *Select Layer By Location* et de *Select Layer By Attribute* est une combinaison puissante. Par exemple, vous pourriez répondre à une question telle que : "Combien d'agressions se sont produites dans un rayon de 250 m autour d'une école à Ottawa ? Combien d'agressions se sont produites à moins de 250 m d'une école à Ottawa ?" Pour répondre à cette question, il faut combiner une requête spatiale et une requête SQL:

1. Répéter l'exemple de la question Q1 pour sélectionner tous les crimes commis dans un rayon de 250 m autour de chaque école.
2. Ouvrez la table d'attributs de la couche **Crime**.
3. Vous remarquerez que 6251 points de criminalité (sur 25454) ont été sélectionnés. Vous remarquerez également qu'il existe d'autres attributs tels que « *SUMMARY* » (résumé) qui fournit des informations sur le type d'infraction pour chaque point:

	OBJECTID *	Shape *	PRIORITY	SUMMARY	OCCDATE
1	1	Point	3	Weapons	1999-05-05
2	2	Point	3	Weapons	1999-04-17
3	3	Point	1	Weapons	1999-03-26
4	4	Point	1	Weapons	1999-08-11
5	5	Point	1	Weapons	1999-07-28
6	6	Point	1	Weapons	1998-10-17

4. Donc, parmi les 6251 entités sélectionnées résultant de la requête spatiale de l'exemple de la question Q1, vous voulez maintenant sélectionner uniquement celles qui sont des « *Assault* » (agression). Pour effectuer cette sélection, gardez la table **Crime** ouverte et allez dans l'onglet *Map* (Carte), trouvez le groupe "Selection" et choisissez « *Select Layer By Attribute* » (Sélectionner la couche par attribut).



- Examinez la boîte de dialogue « *Select Layer By Attribute* » ci-dessous et faites en sorte que la vôtre soit identique, l'essentiel est ici d'indiquer à ArcGIS que vous souhaitez « **Select subset from the current selection** » (sélectionner un sous-ensemble à partir de la sélection actuelle) afin que seuls les points qui répondent au critère d'avoir un type « *SUMMARY* » égal à « *Assault* » (Agression) soient pris en compte:

The feature class on which to make the selection

The selection method. Here you must tell ArcGIS that you want to select from within the currently selected features only. In other words, you want to select from the features that are already selected from the spatial query.

The SQL SELECT statement WHERE clause

- Cliquez sur « *Run* ».

7. En examinant la table d'attributs de **Crime**, vous pouvez voir qu'il y a 909 agressions à moins de 250 m d'une école à Ottawa. Cela représente environ 14 % ($909/6232 * 100$) du total des crimes commis à moins de 250 m des écoles.

	OBJECTID *	Shape *	PRIORITY	SUMMARY	OCCDATE
1	18883	Point	1	Assault	1999-12-04
2	18884	Point	1	Assault	1999-04-10
3	18885	Point	1	Assault	1999-08-03
4	18886	Point	1	Assault	1998-08-03
5	18887	Point	1	Assault	1999-10-22
6	18888	Point	1	Assault	1998-12-03

Q3: Comment créer une requête SQL avec un champ de type date?

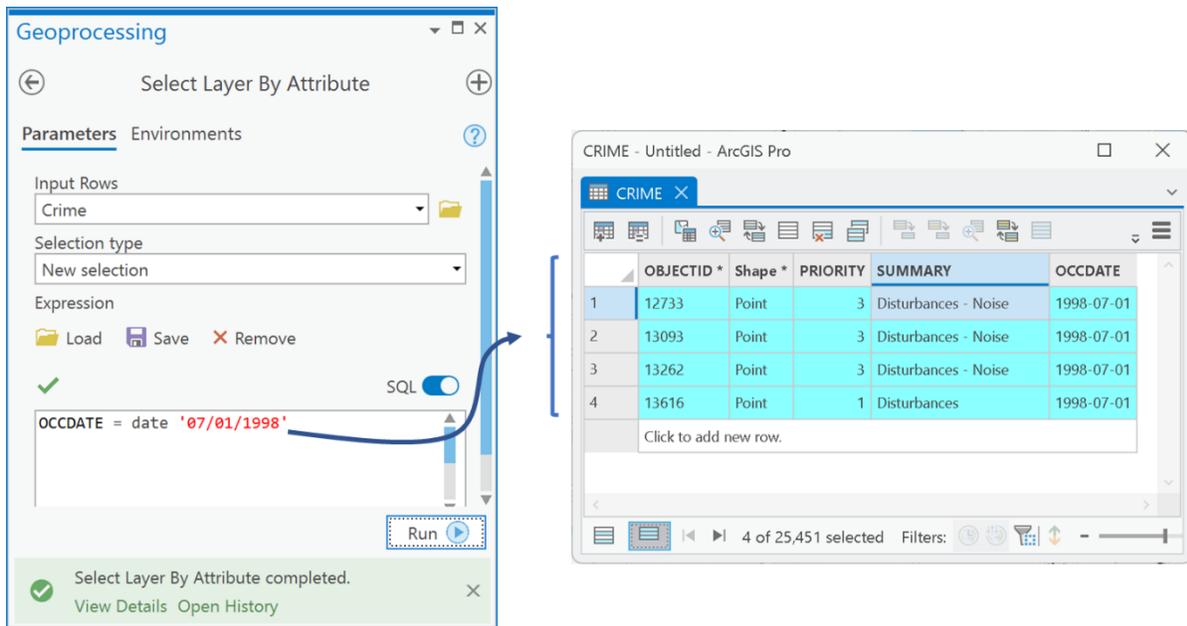
Le temps et la date sont des attributs implicites des données spatiales. Pourquoi ? Parce que toutes les données spatiales sont collectées à une date et une heure déterminée. Certaines données spatiales, comme la couche **Crime**, contiennent des informations sur la date à laquelle une observation a eu lieu. Dans le cas de la couche **Crime**, il existe un champ *OCCDATE* qui nous indique la date à laquelle une observation de point donné s'est produite. Le format de la date est le suivant :

mois/jour/année

Par exemple, un point dont l'OCCDATE est 7/1/1998 signifie que l'événement en question s'est produit le 1er juillet 1998. Pour interroger un champ Date dans une géodatabase ou un fichier de forme (*shapefile*), vous devez faire précéder la valeur de la date du mot-clé **'date'** (sans les apostrophes). Par exemple, si je voulais connaître tous les crimes survenus le 1er juillet 1998, j'entrerais ce qui suit dans la boîte de dialogue « *Select Layer By Attribute* » (Sélection de la couche par attribut):

OCCDATE = date '07/01/1998'

Cela permet de sélectionner tous les événements (et toutes les caractéristiques) survenus le 1er juillet 1998. Par exemple, faites la requête ci-dessus sur la couche **Crime** en utilisant la boîte de dialogue « *Select Layer By Attribute* » (Sélectionner une couche par attribut):



Vous pouvez également rechercher toutes les dates comprises dans une plage de dates en utilisant des opérateurs relationnels tels que `<`, `>`, `<=`, `>=`. Par exemple, si vous souhaitez connaître tous les crimes commis au cours du mois de juillet 1998, vous pouvez utiliser la requête SQL suivante:

`OCCDATE >= date '07/01/1998' AND OCCDATE < date '08/01/1998'`

Ceci sélectionnerait toutes les dates du mois de juillet puisque le symbole `"<"` est utilisé pour s'assurer qu'aucune date postérieure ou incluse au 1er août n'est sélectionnée.

CRIME - Untitled - ArcGIS Pro

	OBJECTID *	Shape *	PRIORITY	SUMMARY	OCCDATE
1	105	Point	2	Weapons	1998-07-19
2	126	Point	2	Traffic Stop	1998-07-29
3	173	Point	2	Traffic Stop	1998-07-18
4	189	Point	2	Traffic Stop	1998-07-18
5	261	Point	2	Traffic Stop	1998-07-28
6	518	Point	2	Traffic Stop	1998-07-30

428 of 25,451 selected

Vous pouvez également utiliser des conditions de date avec des conditions basées sur d'autres champs. Supposons que vous souhaitez obtenir toutes les agressions commises au cours du mois de janvier 1999:

`OCCDATE >= date '01/01/1999' AND OCCDATE < date '02/01/1999' AND SUMMARY = 'Assault'`

CRIME - Untitled - ArcGIS Pro

	OBJECTID *	Shape *	PRIORITY	SUMMARY	OCCDATE
1	18993	Point	2	Assault	1999-01-02
2	19001	Point	2	Assault	1999-01-21
3	19016	Point	1	Assault	1999-01-01
4	19039	Point	2	Assault	1999-01-31
5	19054	Point	1	Assault	1999-01-22
6	19084	Point	1	Assault	1999-01-23

42 of 25,451 selected

Q4: Comment entreprendre des mesures centrographiques pour un ensemble de points?

Dans la lecture 6, vous avez été initié aux mesures centrographiques des ensembles de points. Ces mesures comprennent le centre moyen, la distance standard (DS) et l'ellipse d'écart type. Ces mesures centrographiques sont souvent utilisées pour comparer les distributions de phénomènes ponctuels à différents moments dans une zone ou un domaine spatial (figure 2). La figure 2a illustre un déplacement vers le sud-est du centre moyen des crimes d'agression à Ottawa entre janvier et juin 1999. En outre, les cercles de distance standard pour ces deux mois indiquent que la dispersion globale des agressions est restée la même, comme l'indique la taille des cercles DS (figure 2a). Toutefois, les ellipses d'écart-type pour les deux mois indiquent que, pour janvier 1999, la dispersion directionnelle était légèrement orientée N-S et que l'orientation est devenue plus NNE-SSW en juillet 1999 (Figure 7b).

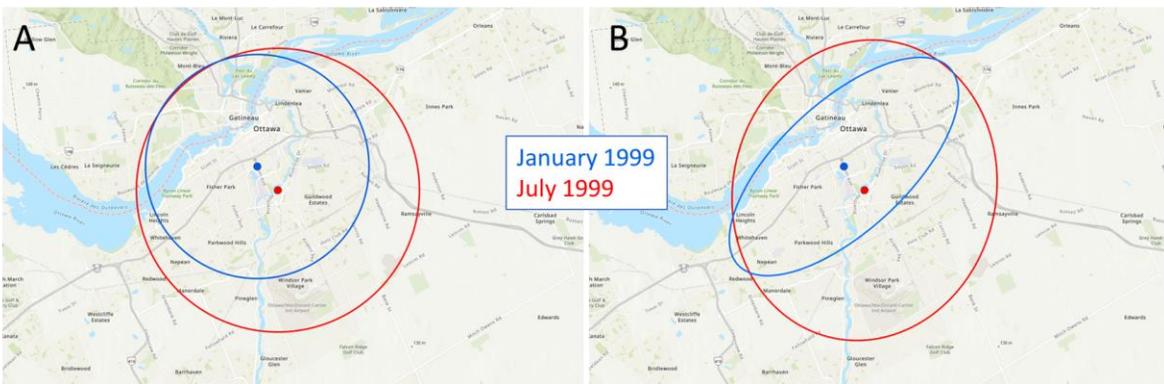
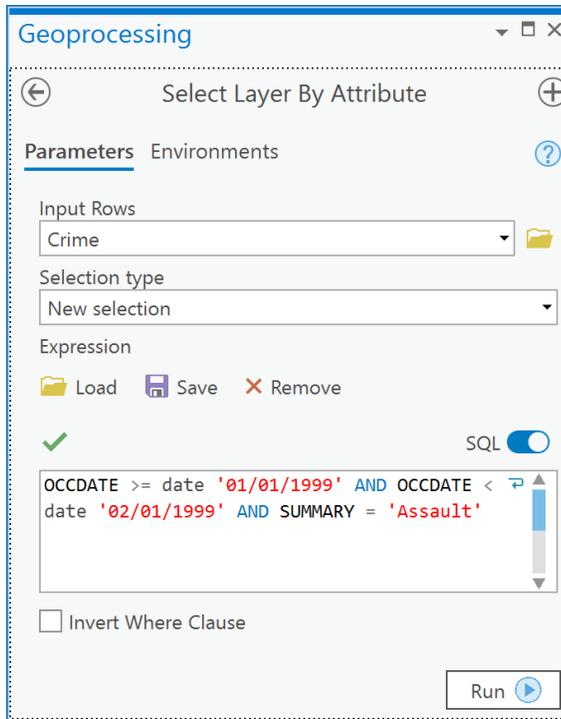


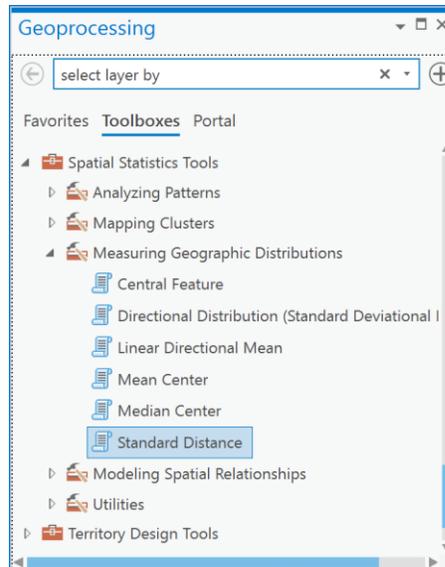
Figure 7: A) Cercles de distance standard et B) ellipses d'écart type pour les agressions à Ottawa en 1999.

Supposons que vous souhaitez comparer la dispersion des agressions à Ottawa entre janvier 1999 et février 1999. Pour effectuer cette analyse, vous devez:

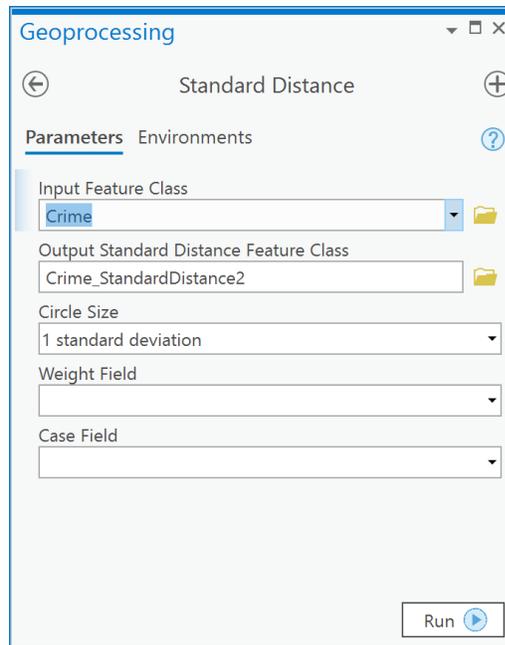
1. Utiliser « *Select Layer By Attribute* » pour chercher les agressions commises en janvier 1999:



2. Allez dans « *Spatial Statistics tools* » (Outils de statistiques spatiales) -> « *Measuring Geographic Distributions* » (Mesure des distributions géographiques) -> « *Standard Distance* » (Distance standard),



3. Dans le menu déroulant « *Input Feature Class* », choisissez **Crime** et cliquez sur « *Run* ».



4. Le résultat est un fichier de forme polygonale qui montre un cercle avec une distance standard qui représente la dispersion des agressions en janvier. Comme nous avons utilisé une requête pour sélectionner un sous-ensemble de caractéristiques dans l'ensemble des données criminelles, seuls les éléments sélectionnés ont été utilisés pour calculer la distance standard. C'est le comportement par défaut de presque tous les outils, car lorsqu'il y a une sélection dans une couche dans ArcGIS et que cette couche est utilisée dans un outil, l'outil n'opère que sur les caractéristiques sélectionnées.
5. Répéter ensuite les étapes 1 à 3 pour le mois de février 1999.
6. Le résultat (figure 3) montrera deux cercles de distance standard sur votre carte, un pour les agressions de janvier 1999 et un pour les agressions de février 1999:

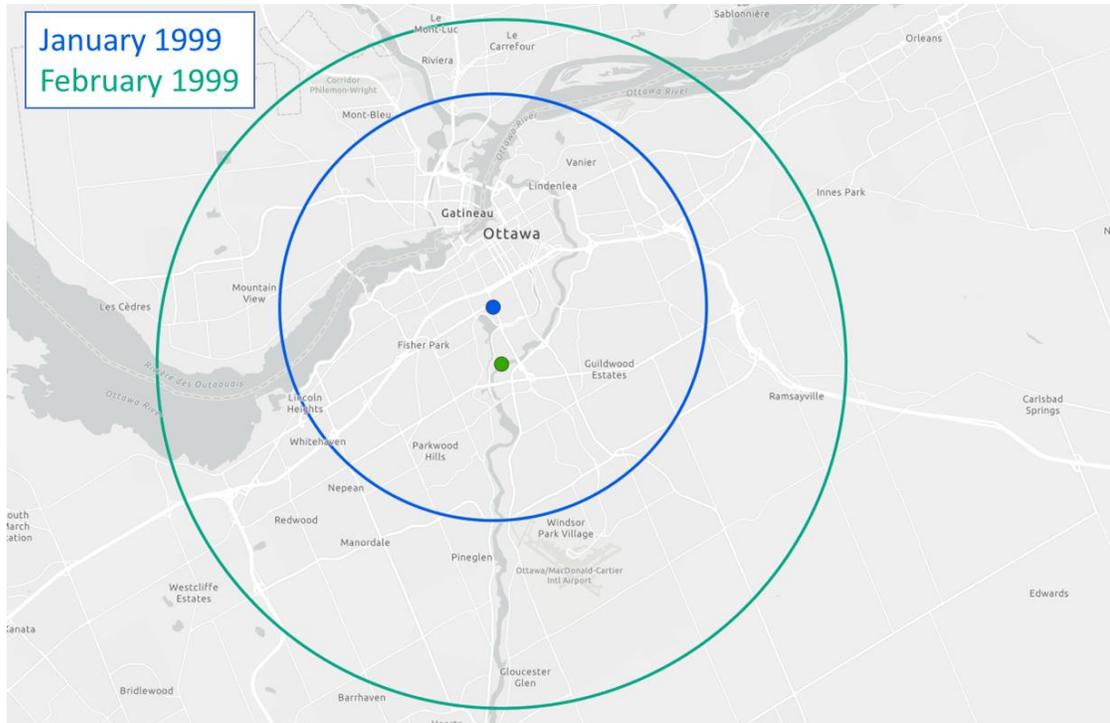


Figure 8: Cercles de distance standard comparant la dispersion des agressions en janvier et février 1999 à l'intérieur d'Ottawa.

Les agressions sont plus *dispersées* en février qu'en janvier 1999, comme le montre le cercle de distance standard plus grand en février (figure 3). Comme les centres de chaque cercle ne coïncident pas, il semble y avoir un déplacement du *centre moyen* d'activité vers l'est en février.

Nous aurions pu utiliser le volet Python pour entreprendre l'analyse ci-dessus, par exemple pour obtenir toutes les mesures centrographiques - centre moyen, distance standard et ellipse d'écart type - pour les crimes d'agression de janvier,

```
# Créer des variables de type chaîne pour contenir
# la clause SQL « where » ainsi que
# ainsi que les noms des couches d'entrée et de
# couches de sortie
input_lyr = 'crime'
out_mean_center = 'january_mean_center'
out_sd = 'january_sd'
out_sde = 'january_sde'
query_string = "OCCDATE >= date '01/01/1999' AND OCCDATE < date
'02/01/1999' AND SUMMARY = 'Assault'"

# Sélectionner les crimes dans la plage de dates
SelectLayerByAttribute(input_lyr, where_clause=query_string)

# Créer une couche avec un point pour le centre moyen
MeanCenter(input_lyr, out_mean_center)
```

Nom de l'étudiant: _____ # d'étudiant: _____

```
# Créer une couche avec un cercle pour la distance standard
StandardDistance(input_lyr, out_sd)
```

```
# Créer une couche avec une ellipse d'écart-type
DirectionalDistribution(input_lyr, out_sde)
```

L'avantage du code ci-dessus est que lorsque nous voulons réexécuter les instructions python pour le mois de février, tout ce que nous devons faire est de changer les variables de type texte en haut, principalement en remplaçant « *january* » par « *february* » et en modifiant les mois dans la clause SQL « *where* ».:

```
input_lyr = 'crime'
out_mean_center = 'february_mean_center'
out_sd = 'february_sd'
out_sde = 'february_sde'
query_string = "OCCDATE >= date '02/01/1999' AND OCCDATE < date
'03/01/1999' AND SUMMARY = 'Assault'"

SelectLayerByAttribute(input_lyr, where_clause=query_string)

MeanCenter(input_lyr, out_mean_center)

StandardDistance(input_lyr, out_sd)

DirectionalDistribution(input_lyr, out_sde)
```

Q5: Comment visualiser la variabilité spatiale de l'intensité des points à l'intérieur d'une région?

Une lecture vous a présenté l'estimation de la densité du noyau (« *KDE - kernel density estimate* »). Un *KDE* montre la variation de l'intensité, appelée λ , dans un domaine spatial. Un *KDE* est représenté par une surface de densité de noyau créée à partir d'un ensemble de points. Le *KDE* est créé en plaçant une fonction mathématique 3D, appelée noyau, dont la valeur maximale est centrée sur l'emplacement de chaque point d'un motif. Ensuite, les valeurs z (valeurs de hauteur) de toutes les fonctions mathématiques 3D (les noyaux) sont additionnées (figure 4) à chaque emplacement dans l'espace pour créer la surface de le *KDE*. Ce processus aboutit à une couche matricielle qui présente des valeurs élevées là où il y a beaucoup de points et des valeurs plus faibles là où il y a moins de points (figure 4).

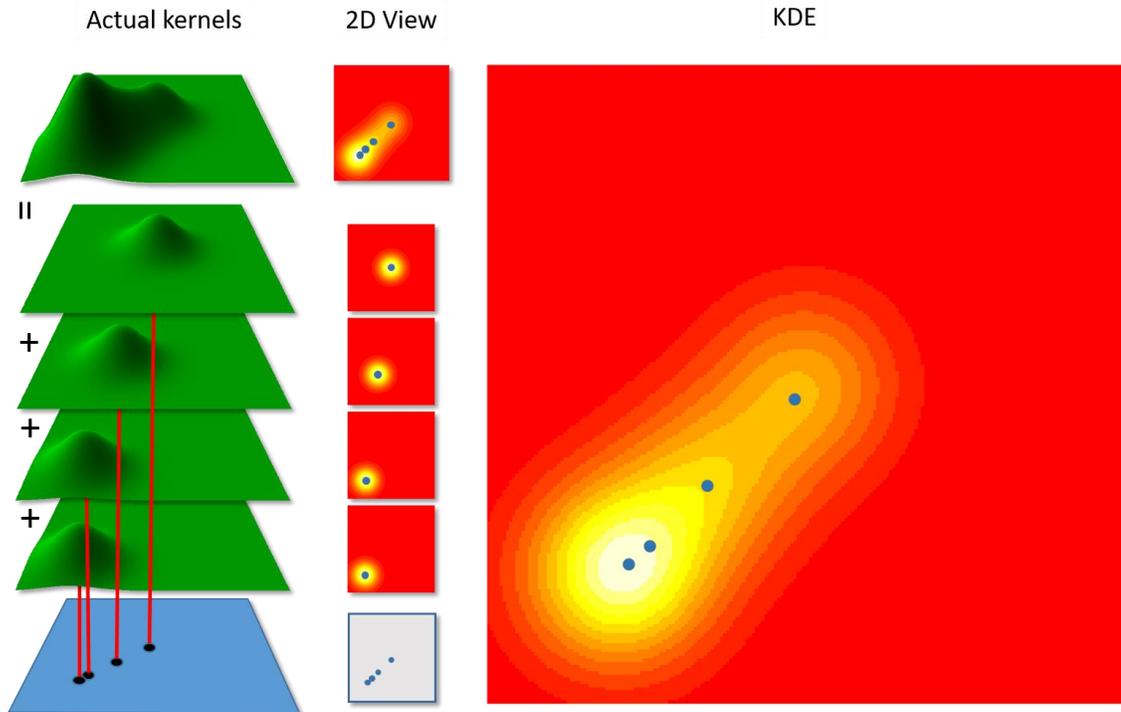
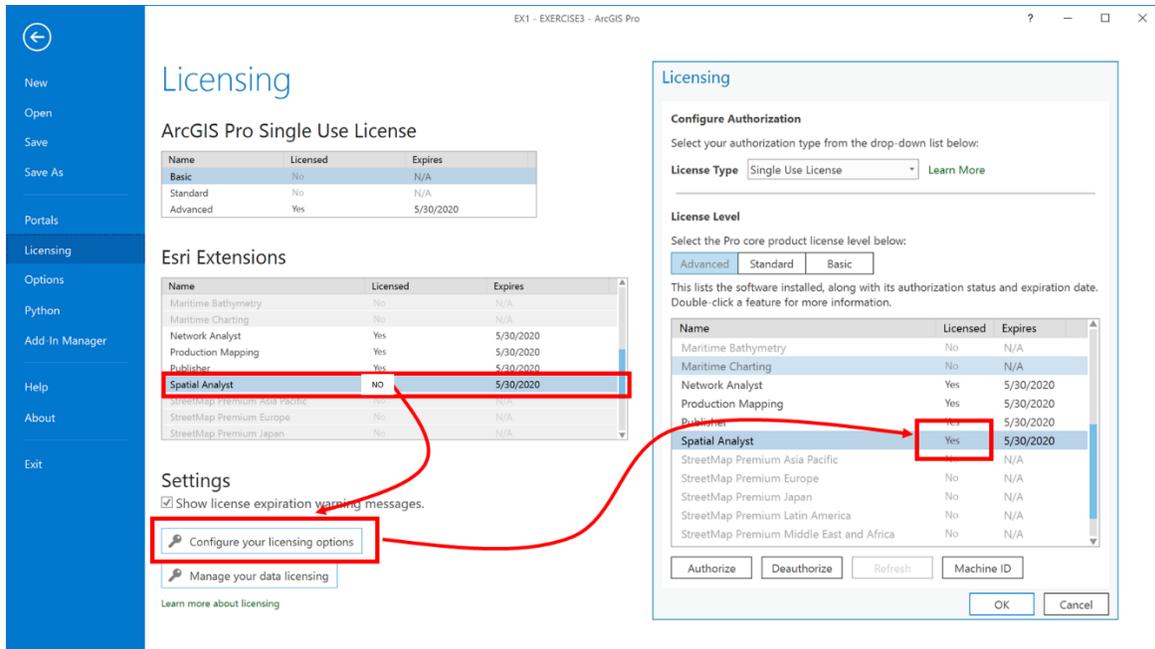


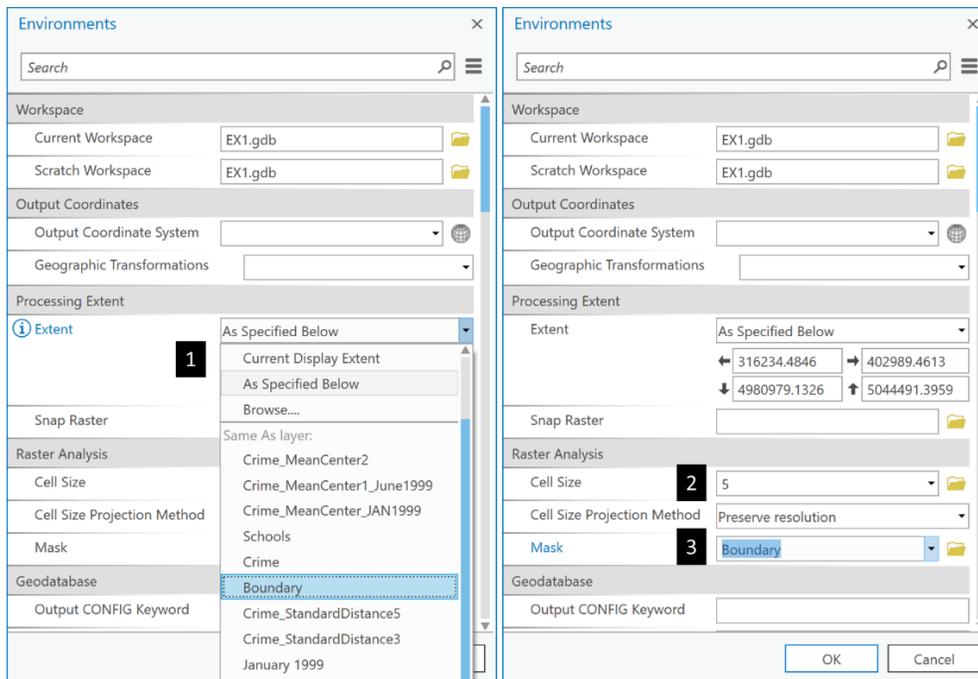
Figure 9: La création d'un KDE. Les noyaux gaussiens 3D (à gauche) placés au-dessus de chaque point d'un motif ponctuel sont additionnés à chaque endroit de la région étudiée.

Par exemple, vous pouvez créer un KDE pour les agressions de janvier 1999 et le comparer au KDE pour les agressions de février 1999:

1. Vous devez vous assurer que vous chargez l'extension « *Spatial Analyst* » dans ArcGIS. Pour le faire, cliquez sur l'onglet « *Project* » et cherchez « *Spatial Analyst* » sous « *Esri Extensions* » - il se peut que vous deviez faire défiler la page. S'il y a une case à cocher, cochez-la et revenez à l'onglet *Map*. S'il n'y a pas de case à cocher et que le mot « *No* » (non) apparaît, cliquez sur le bouton « *Configure your licensing options* » (Configurez vos options de licence). Vous verrez une liste et trouverez à nouveau celle qui indique « *Spatial Analyst* ». Vous devriez alors voir une case à cocher. Cochez-la et cliquez sur **OK**, puis revenez à votre onglet de carte (« *Map* »). S'il n'y a pas de case à cocher, choisissez la ligne qui contient « *Spatial Analyst* » et cliquez sur le bouton « *Authorize* ». Une fois que vous avez terminé, cliquez sur **OK** et revenez à votre onglet « *Map* ». Le vôtre sera similaire ou identique à ce que vous voyez ci-dessous :

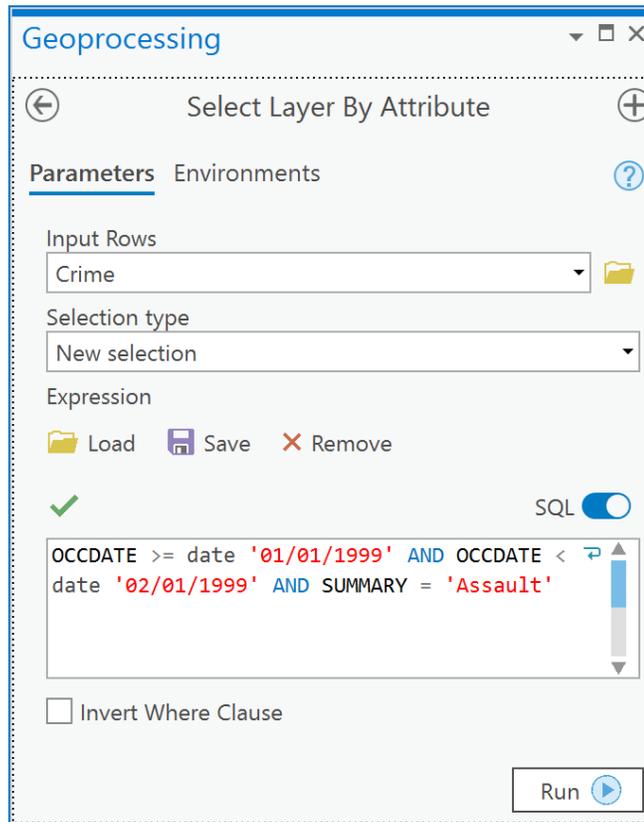


- Ajoutez maintenant à votre carte la couche appelée « **Boundary** » du jeu de données d'entités OTTAWADATA.
- Vous devez définir un certain nombre de paramètres clés de l'environnement pour pouvoir utiliser les outils de l'analyste spatial dans le volet de géotraitement. Plus précisément, vous devez définir l'étendue du traitement, la taille des cellules et le masque d'analyse dans les paramètres de l'environnement d'ArcGIS Pro. Allez dans l'onglet « *Analysis* » et dans le groupe « *Geoprocessing* » cliquez sur le bouton « *Environments* » et définissez vos paramètres comme vous le voyez dans 1,2,3 ci-dessous :

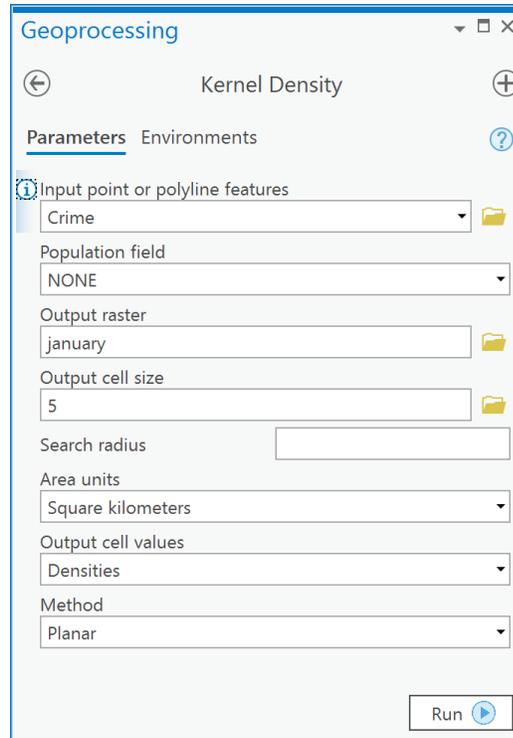


Nom de l'étudiant: _____ # d'étudiant: _____

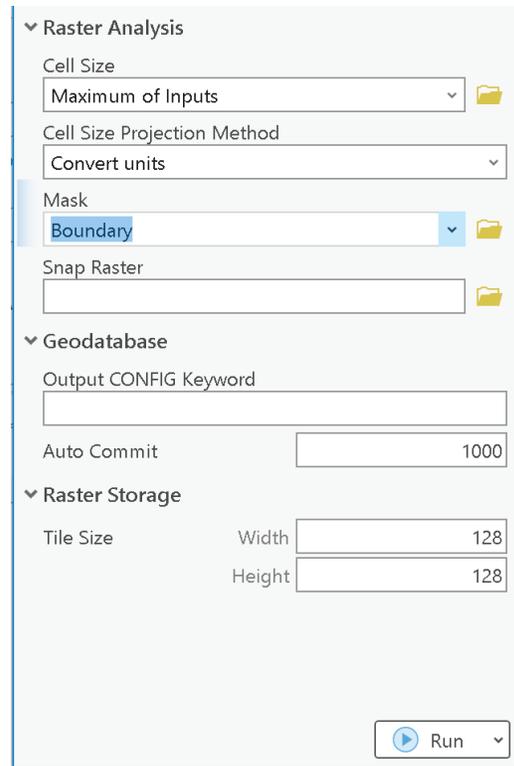
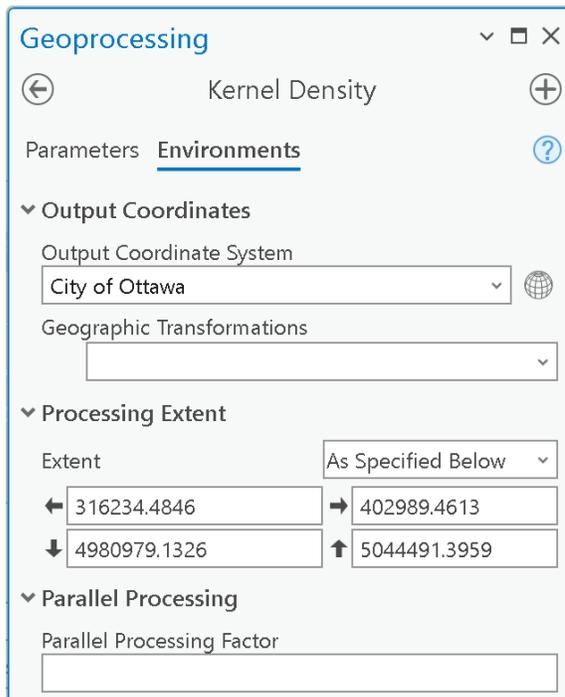
4. Sélectionnez maintenant tous les points de criminalité qui sont des agressions en janvier 1999, comme vous l'avez fait dans l'EX Q4 ci-dessus:



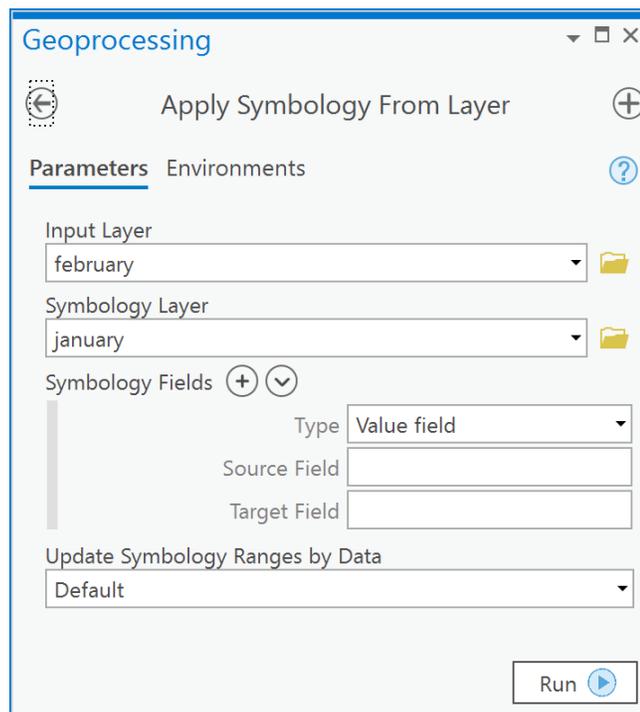
5. Allez ensuite dans « *Spatial Analyst Tools* » (Outils d'analyse spatiale) -> « *Density* » (Densité) -> « *Kernel Density* » (Densité de noyau). Sous « *Input point or polyline features* », choisissez **Crime**. Sous « *Output raster* » appelez-le '*january*'. Ici, le « *Search radius (optional)* » (rayon de recherche (facultatif)) est le paramètre de largeur de bande. ArcGIS calcule une valeur par défaut raisonnable pour la plupart des cas en utilisant la règle de Silverman (voir le fichier d'aide). Assurez-vous que tous les autres paramètres sont tels que vous les voyez ci-dessous et cliquez sur « *Run* » (Exécuter).



6. Définissez les paramètres d'environnement critiques pour le niveau de l'outil en sélectionnant l'onglet « *Environments* » et en remplissant les champs suivants:



7. Répétez les étapes 3 à 6 pour le mois de février 1999 en appelant la sortie *KDE* 'february'.
8. Pour comparer deux ou plusieurs estimations KDE, les échelles z doivent être les mêmes sur les deux cartes, sinon une valeur de couleur telle que le bleu le plus foncé sur une carte ne correspondrait pas à la même magnitude de densité sur l'autre - de même avec n'importe quelle couleur sur l'échelle z. La façon la plus simple de créer deux cartes avec la même échelle z est d'utiliser l'outil de géotraitement "Apply Symbology from Layer" (Appliquer la symbologie d'une couche) pour importer la symbologie d'une couche existante pour laquelle vous avez déjà créé une classification. Dans le cas des agressions, j'importe à partir de la couche qui a la densité la plus élevée, qui dans notre cas est celle de janvier, et j'importe donc la symbologie de janvier dans celle de février :



9. Vos résultats (figure 5) vous permettent de comparer les surfaces d'intensité des deux mois.

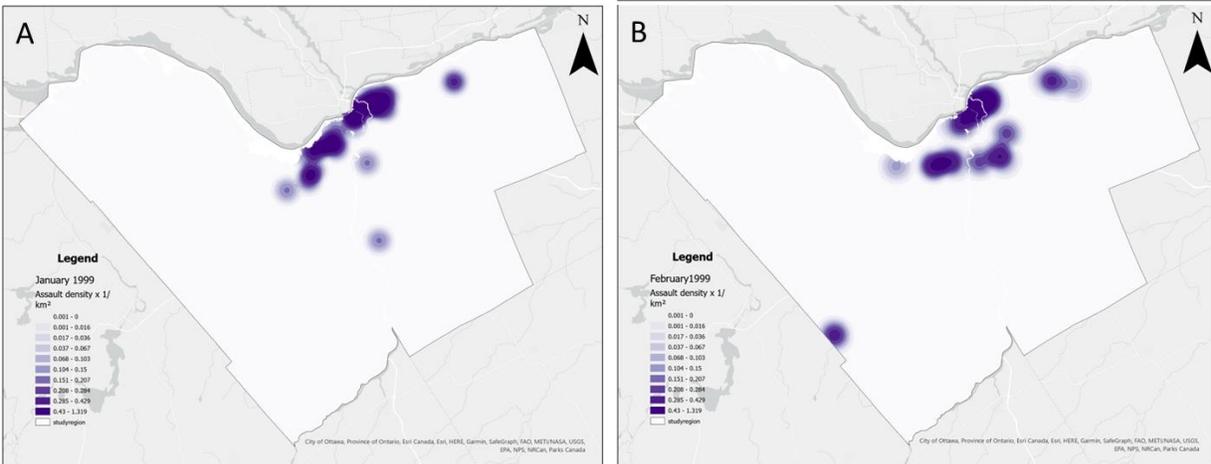


Figure 10: Kernel density estimates for assaults in A) January 1999; B) February 1999

On peut constater que l'intensité est plus élevée entre la rivière Rideau et le canal Rideau en janvier 1999 qu'en février 1999 (figure 5). De plus, les valeurs d'intensité sont plus étalées en février 1999 (figure 5). Ces observations sont cohérentes avec les observations de la distance standard dans l'EX Q4, l'intensité plus élevée au même endroit en janvier 1999 conduit à une distance standard plus petite, alors que l'intensité plus diffuse en février 1999 conduit à une distance standard plus grande.

Une autre façon, plus efficace, de comparer deux cartes de la même variable est de calculer la différence entre elles. Pensez-y. Lorsque vous comparez deux cartes quantitatives, vous vous intéressez à leurs différences et à leurs similitudes. Une carte des différences peut aider à identifier rapidement les incongruités entre deux périodes. Nous pouvons facilement calculer la différence entre les densités des agressions entre janvier et février à l'aide de l'outil Python d'ArcGIS Pro.

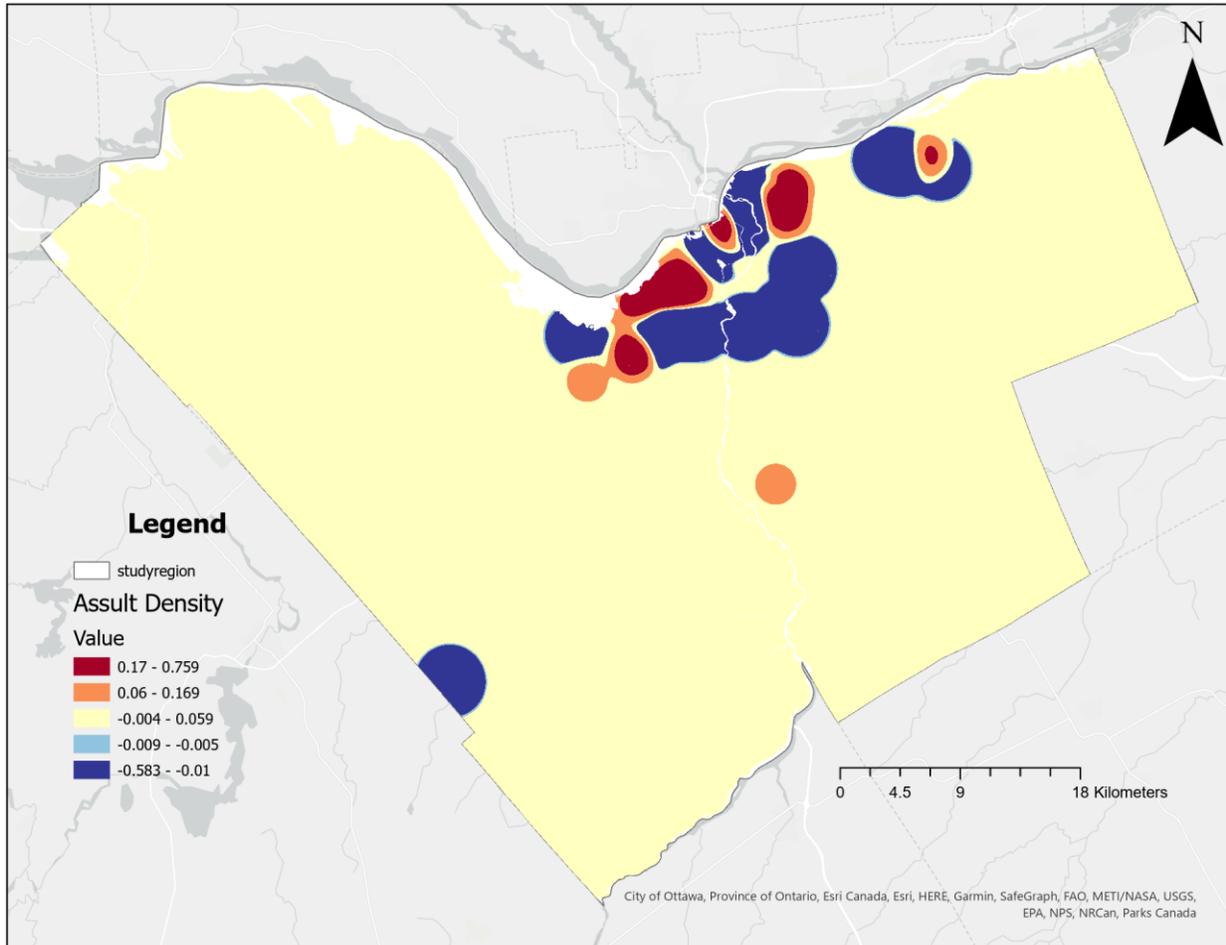
1. Allez dans l'onglet « *Analysis* » (Analyse) et cliquez sur le bouton Python.
2. Entrez les lignes suivantes, à l'exception de celles précédées d'un # qui sont des commentaires à lire pour comprendre la signification des instructions python.

```
# Importer la boîte à outils de l'analyste spatial
# pour pouvoir appeler les outils par leur nom
from arcpy.sa import *

# Créer des variables raster python à partir de
# les rasters de janvier et février
jan = Raster('january')
feb = Raster('february')

# Calculez la différence entre janvier et février
# et créez une nouvelle couche appelée 'diff'.
diff = jan - feb
```

3. Le résultat sera une carte des différences. Lorsque la carte présente des valeurs positives, le mois de janvier a connu une intensité d'agressions plus élevée que le mois de février. Lorsque les valeurs sont inférieures à zéro, le mois de février a connu une intensité d'agressions plus élevée :



Q6: Comment calculer les agressions par unité de surface dans une région?

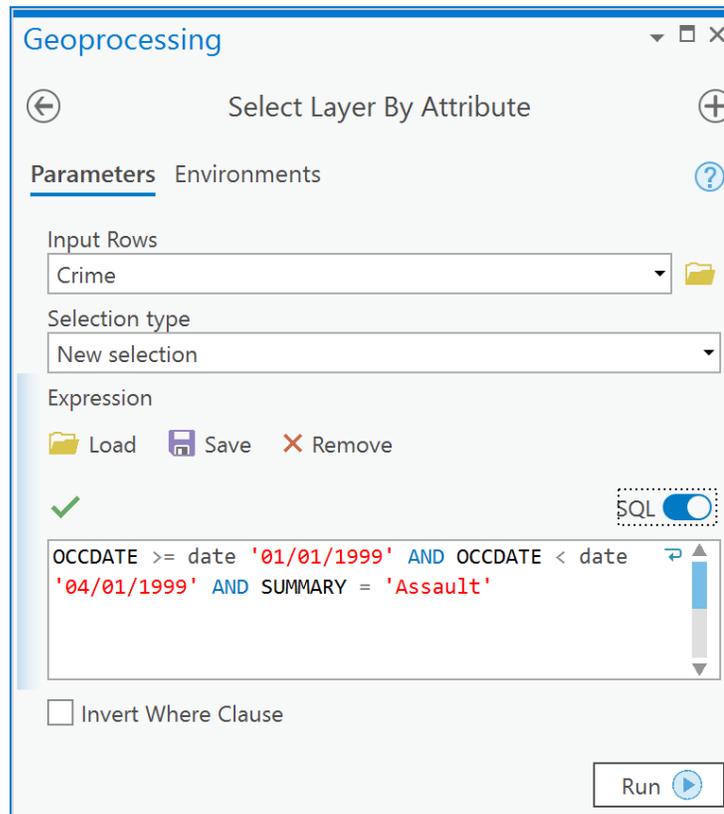
Dans la lecture 6, nous vous avons présenté certaines relations entre points et polygones. Une relation couramment utilisée consiste à examiner la densité des points à l'intérieur des polygones. Pour cela, il faut compter chaque point à l'intérieur de chaque polygone, puis diviser ce nombre par la surface du polygone afin d'obtenir une mesure des points par unité de surface. Ce type d'opération est facilement réalisable dans les SIG, car l'opération " points dans polygone " est courante. Pour compter le nombre de points dans un nombre quelconque de polygones au sein du même domaine spatial, il suffit d'effectuer une jointure spatiale à l'aide de l'outil « *Spatial Join* » (jointure spatiale) dans ArcToolbox afin d'effectuer des jointures spatiales lorsque vous souhaitez uniquement utiliser les données sélectionnées à partir d'une requête dans le cadre de l'opération de jointure spatiale.

Nom de l'étudiant: _____ # d'étudiant: _____

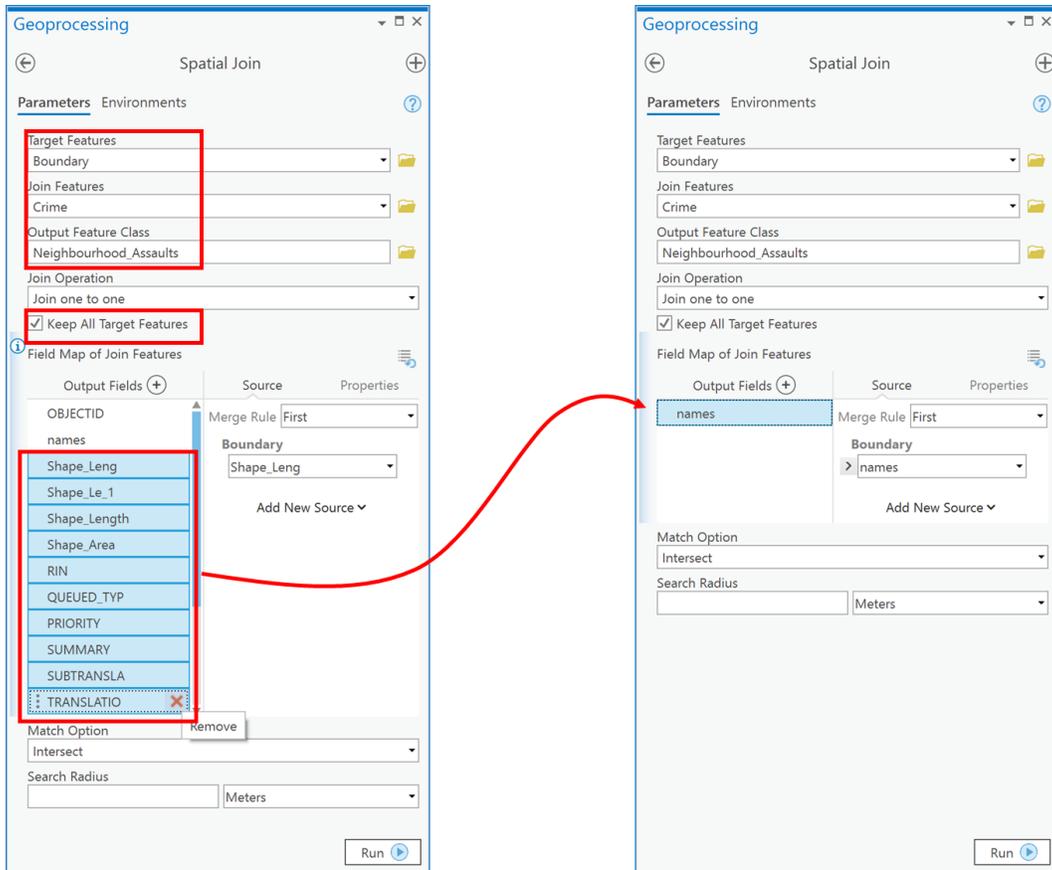
Supposons que vous souhaitez calculer, dans les quartiers d'Ottawa, la densité des agressions de janvier à mars 1999 (inclus) et la comparer à la dispersion des agressions de juin à août 1999 (inclus) :

1. Créer une requête sur la couche **Crime** pour les agressions entre janvier 1999 et mars 1999,

OCCDATE >= date '01/01/1999' AND OCCDATE < date '04/01/1999' AND SUMMARY = 'Assault'



2. Cliquez sur « *Analysis Tools* » (Outils d'analyse) -> « *Overlay* » (Superposition) -> « *Spatial Join* » (Jointure spatiale). Les « *Target Features* » seront la couche polygonale **Boundary** qui contient les quartiers, les « *Join Features* » sont la couche **Crime** qui contient les enregistrements sélectionnés à partir de la requête de l'étape 1 ci-dessus. La « *Output Feature Class* » (classe de caractéristiques de sortie) doit s'appeler **Neighbourhood_Assaults** (agressions dans les quartiers). Cochez la case « *Keep All Target Features* » (Conserver toutes les caractéristiques cibles). Sous « *Field Map of Join Features* » (carte des champs des entités jointes), supprimez tous les champs à l'exception de "names" (noms) qui ne conservera que le nom du quartier dans la sortie. Cet outil créera automatiquement un champ "Join_Count" dans la classe d'entités de sortie qui contient le nombre de points trouvés dans chaque polygone de quartier dans **Boundary**. Avant de cliquer sur « *Run* », assurez-vous que les options de la boîte de dialogue « *Spatial Join* » sont les mêmes que celles ci-dessous :



3. Vous avez maintenant une nouvelle classe d'entités dans votre table des matières. Ouvrez la table d'attributs de la nouvelle classe d'entités et vous verrez le tableau suivant avec un nouveau champ appelé « *Join_Count* » qui contient le nombre d'agressions dans chaque quartier. Le champ « *Shape_Area* » contient la surface de chaque quartier en unités cartographiques (m²) du système de coordonnées actuel (qui est le même système de coordonnées par défaut que la couche cible) :

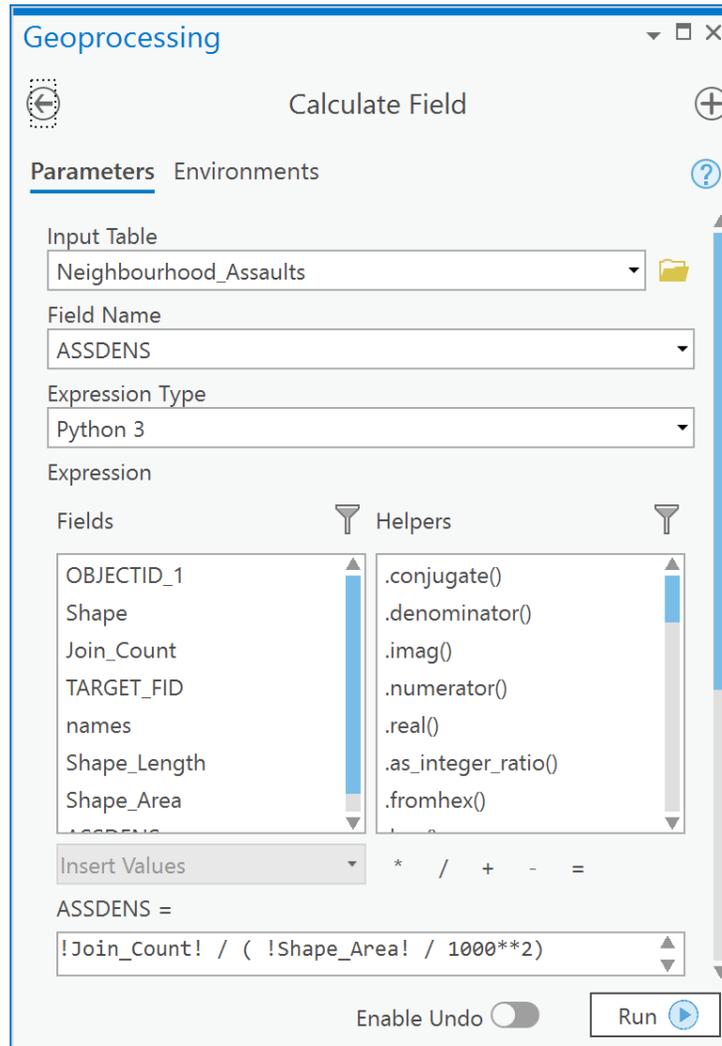
Boundary_SpatialJoin - Exercise4_formerly3 - ArcGIS Pro

Field:	Selection:	OBJECTID_1 *	Shape *	Join_Count	TARGET_FID	OBJECTID	names	Shape_Length	Shape_Area
1		1	Polygon	0	1	1	Barrhaven	8540.50976	4441790.840001
2		2	Polygon	6	2	2	Bayshore	5723.440175	1337789.372141
3		3	Polygon	0	3	3	Beacon Hill South - Car...	7291.193688	2232023.589849
4		4	Polygon	0	4	4	Beaverbrook	7316.337812	2543752.143629
5		5	Polygon	0	5	5	Bells Corners East	7872.663162	3354774.643387
6		6	Polygon	1	6	6	Bells Corners West	5248.885678	1370699.736583

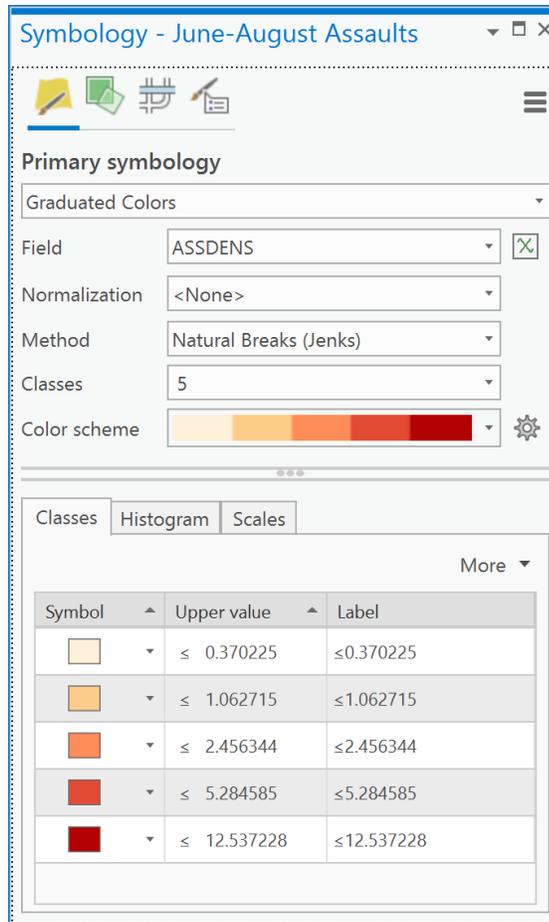
0 of 108 selected | Filters: 100%

4. Ensuite, ajoutez un nouveau champ à la table, appelé **ASSDENS** pour densité d'agression, et donnez-lui le type double (revoyez l'EXERCICE 2 si vous ne vous souvenez pas comment procéder).
5. Maintenant, calculez dans le champ ASSDENS la densité d'agression en personnes par kilomètre carré à l'aide d'une expression telle que:

$$\text{!Join_Count!} / (\text{!Shape_Area!} / 1000^{**2})$$



6. Créez ensuite une carte choroplèthe de la densité des agressions à l'aide du champ **ASSDENS**. Cliquez sur la couche **Neighbourhood_Assaults** dans le volet « Contents » (Contenu). Ensuite, dans le ruban « Feature Layer » (Couche d'objets), choisissez l'onglet « Appearance » (Apparence). Dans le groupe « Drawing » (Dessin), cliquez sur le bouton « Symbology » (Symbologie) et choisissez « Graduated Colors » (Couleurs graduées). Dans le volet « Symbology » (Symbologie), sous « Field » (Champ), choisissez **ASSDENS** en choisissant « Graduated colors » (Couleurs graduées) dans l'onglet « Symbology » (Symbologie), par ex.:



7. Répéter les étapes 1 à 6 pour la période juin-août 1999 (inclus).
8. Les résultats (figure 6) permettent de comparer la densité par quartier entre les deux périodes de temps.

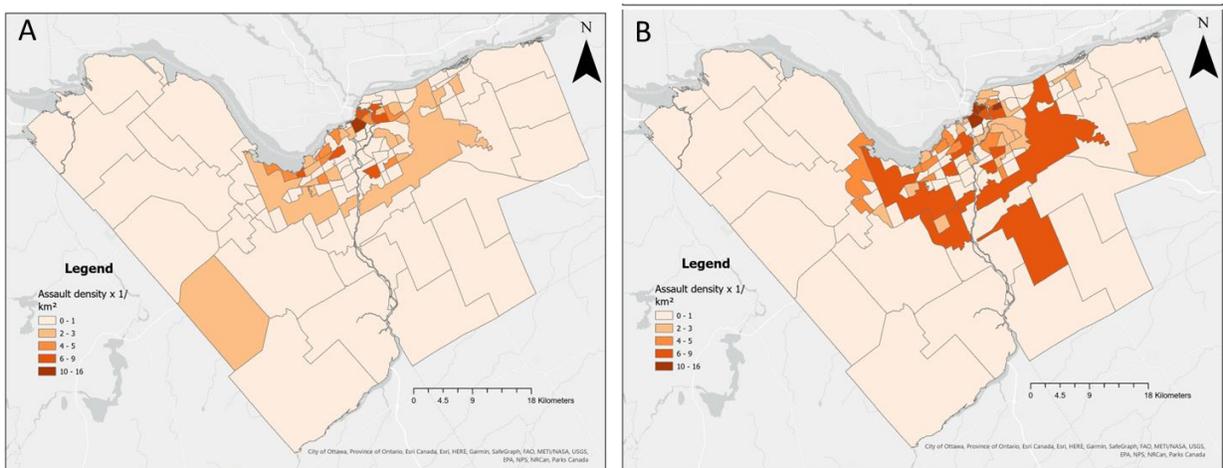


Figure 11: Agressions dans les quartiers d'Ottawa en A) janvier-mars 1999 ; B) juin-août 1999

Remarquez à nouveau que lorsque vous comparez deux cartes de densité, l'échelle z des agressions par kilomètre carré est la même pour les deux cartes (figure 6), sinon la même

Nom de l'étudiant: _____ # d'étudiant: _____

couleur sur chaque carte représenterait une valeur différente, ce qui conduirait à des comparaisons erronées. Dans l'ensemble, les tendances observées au cours des deux saisons sont assez similaires, mais la densité des agressions est plus élevée dans un plus grand nombre de quartiers du centre d'Ottawa au cours de l'été qu'au cours de l'hiver. Le profil de l'été présente également une plus grande similitude spatiale dans les quartiers centraux.

Cet exemple comporte un certain nombre d'étapes qui peuvent également être réalisées dans le volet Python. L'avantage est que pour effectuer la même analyse pour une autre fenêtre temporelle, il suffit de modifier les valeurs dans les déclarations surlignées en jaune:

```
# Importer les fonctions "management" et "analysis"
from arcpy.management import *
from arcpy.analysis import *

input_lyr = 'crime'
query_string = "OCCDATE >= date '01/01/1999' AND OCCDATE < date
'04/01/1999' AND SUMMARY = 'Assault'"
nbrhoods = 'Boundary'
spatial_join_output = 'Neighbourhood_Assaults'
density_field_name = 'ASSDENS'
field_to_keep = 'names'

# Premièrement, sélectionnez les agressions à l'aide
# de la clause SQL 'where'
# contenue dans la chaîne de requête 'query_string'
SelectLayerByAttribute(input_lyr, where_clause=query_string)

# Créer un objet de cartographie des champs. C'est complexe
# mais l'idée est de s'assurer que seuls les
# noms des voisins sont transférés dans la jointure spatiale.
fm = arcpy.FieldMap()
fms = arcpy.FieldMappings()
fm.addInputField(nbrhoods, field_to_keep)
of = fm.outputField
of.name = field_to_keep
fm.outputField = of
fm.mergeRule = 'first'
fms.addFieldMap(fm)

# Joindre spatialement les points de criminalité sélectionnés
# aux quartiers et transférer
# uniquement le champ des noms vers la sortie
# à l'aide de la variable fms (fieldmapping)
SpatialJoin(nbrhoods,
            input_lyr,
```

```

        spatial_join_output,
        'JOIN_ONE_TO_ONE',
        'KEEP_ALL',
        field_mapping=fms
    );

# Ajouter un champ appelé "ASSDENS" à la couche
# jointe spatialement
AddField(spatial_join_output,
         density_field_name,
         'DOUBLE'
    );

# Calculez maintenant le champ comme le nombre
# d'agressions par kilomètre carré
CalculateField(spatial_join_output,
              density_field_name,
              '!Join_Count! / ( !Shape_Area! / 1000**2) '
    );

```

Q7: Comment créer une superposition d'union et pourquoi?

NOTE: Créez une nouvelle carte ArcGIS Pro avant de commencer cet exemple. De plus, si vous avez gardé ArcGIS Pro ouvert depuis le(s) exemple(s) précédent(s), assurez-vous de réinitialiser l'option Environnements de géotraitement (« *Geoprocessing Environments* ») -> Étendue de traitement (« *Processing Extent* ») à l'option "Default" (par défaut). **Si vous ne le faites pas**, l'étendue spatiale d'Ottawa s'appliquera à toutes les sorties analytiques et vous ne verrez pas les résultats à l'échelle de l'Amérique du Nord.

Dans la lecture 6, vous avez découvert les superpositions d'intersection et d'union. Une superposition d'union est utilisée lorsque vous souhaitez déterminer l'union géométrique d'un nombre quelconque de polygones. La sortie d'une opération d'union "contient des polygones représentant l'union géométrique de toutes les entrées ainsi que tous les champs de toutes les classes d'entités d'entrée" (ESRI). Ainsi, une union combine à la fois les polygones et leurs attributs (figure 7). Dans la figure 7, il y a deux couches : une couche contient un polygone orange avec un attribut ORANGE et un polygone vert avec un attribut GREEN (vert). La seconde couche est un triangle jaune avec l'attribut YELLOW (jaune). Lorsque le polygone jaune est uni à la couche des polygones orange et vert, les attributs ORANGE et YELLOW sont combinés là où le polygone jaune a croisé le polygone orange, et un nouveau polygone est créé avec les deux attributs ORANGE et YELLOW dans deux champs distincts. En examinant la table d'attributs fusionnée, nous pouvons voir lesquels des nouveaux polygones générés se sont croisés, car ils possèdent des valeurs dans les deux attributs du nouveau polygone. Le même processus s'applique à l'intersection des polygones jaune et vert. Là où il n'y a pas eu

d'intersection, les nouveaux polygones générés ont simplement un attribut ORANGE ou GREEN.

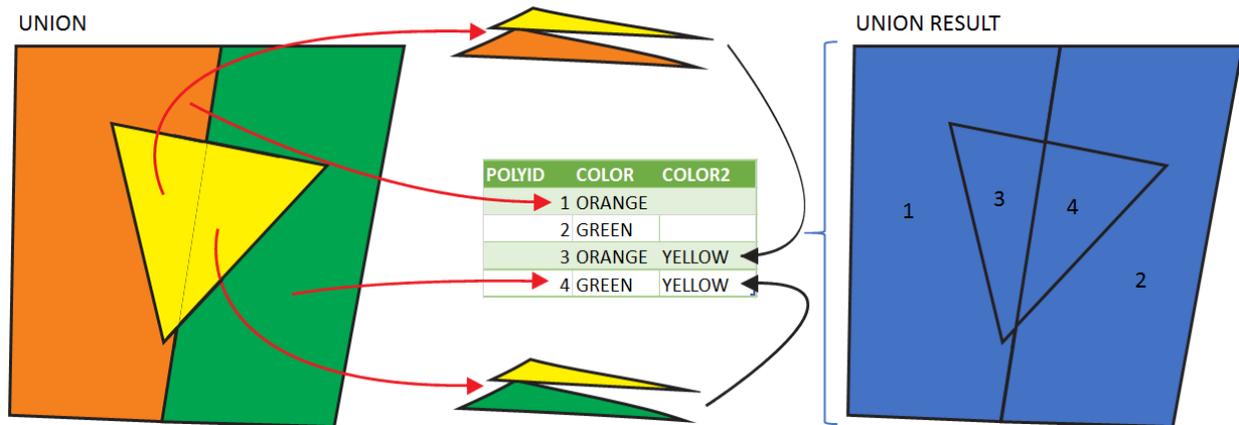


Figure 12: Exemple d'union géométrique

Si vous étudiez la richesse en espèces d'arbres (le nombre d'espèces différentes dans une zone donnée) en Amérique du Nord, vous pouvez superposer un ensemble de polygones représentant les limites de l'aire de répartition de différentes espèces.

Considérons trois limites d'aire de répartition (figure 8) pour trois espèces de noyer (*Jugulans* spp.) en Amérique du Nord (https://www.fs.fed.us/nrs/atlas/littlefia/species_table.html), notamment le noyer noir (« *Black walnut* »), le noyer blanc (« *White walnut* ») et le petit noyer (« *Little walnut* »). Elles sont ajoutées à ArcGIS Pro à partir de l'ensemble de données **TREES** dans la géodatabase OTTAWADATA.gdb du dossier Lecture7/Data. Chaque couche de noyer contient un ensemble de polygones. Dans la table d'attributs de chaque couche, un champ appelé CODE spécifie la valeur '1' pour chacun des polygones de chaque couche (Figure 8). En plaçant ces trois couches dans un processus de superposition d'union, on obtient une nouvelle couche de polygones (figure 8). La nouvelle couche de polygones fusionnée comporte un ensemble de nouveaux polygones et une nouvelle table d'attributs. La nouvelle table d'attributs contient l'ensemble combiné des attributs de toutes les tables d'attributs des couches d'entrée. Plus précisément, vous pouvez maintenant voir la combinaison des champs CODE (figure 8). La valeur de CODE était " 1 " dans chacune des couches d'entrée pour chaque polygone d'entrée et donc, pour un polygone de sortie donné dans la couche fusionnée, une valeur de " 1 " dans l'un des trois champs CODE fusionnés (CODE, CODE1 ou CODE2) signifie que, pour le polygone fusionné en question, l'une des trois couches d'entrée contenait une espèce d'arbre à cet endroit donné (si vous aviez entrepris une superposition d'intersection des trois couches d'entrée, les seuls polygones renvoyés seraient ceux dans lesquels les trois couches d'entrée contenaient un polygone). Un nouveau champ appelé "SUMCODE" peut être ajouté à la table fusionnée. Calculé comme la somme des champs CODE individuels fusionnés, il représente le nombre d'espèces de noyers qui se superposent dans chaque polygone fusionné.

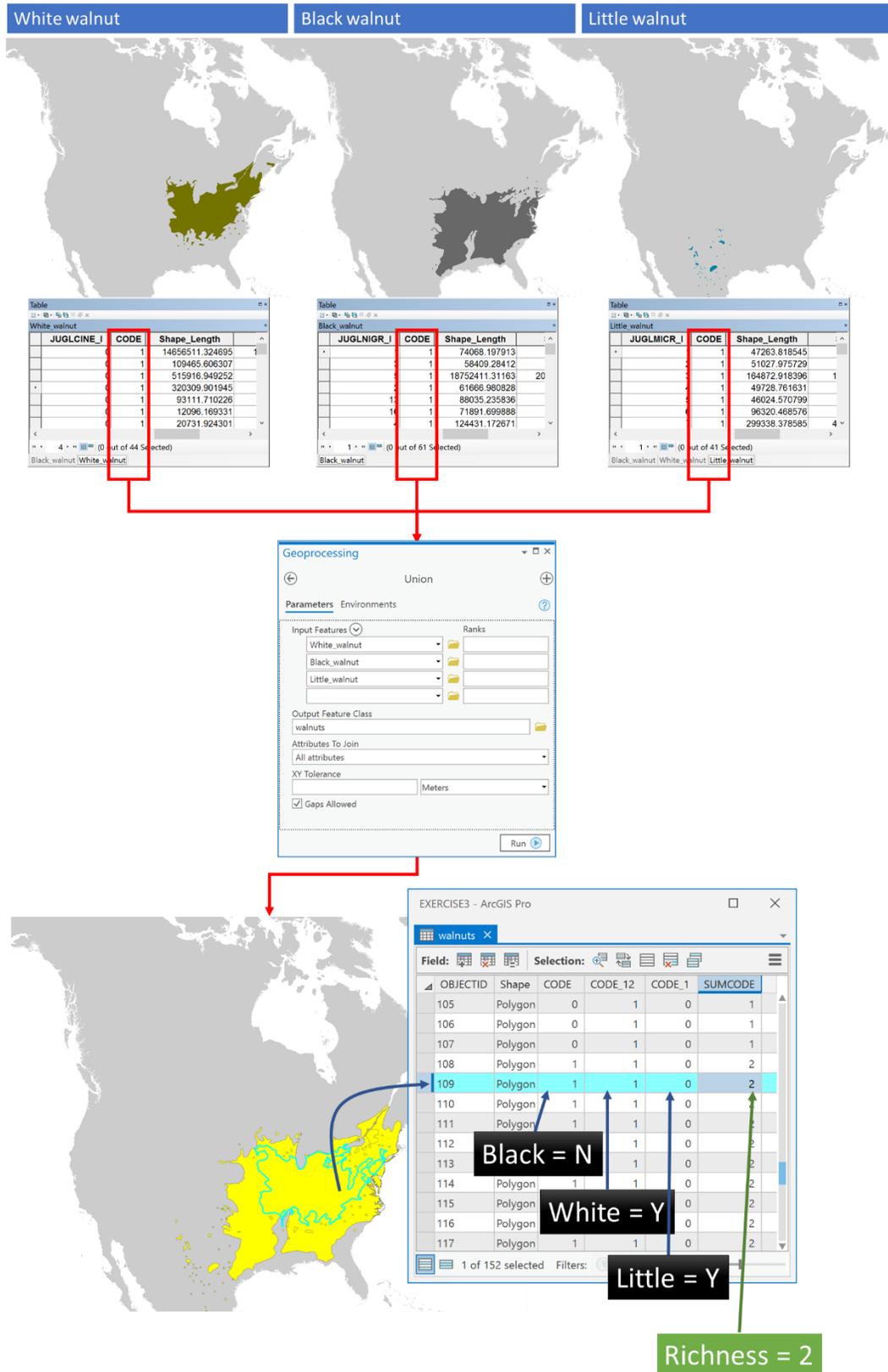
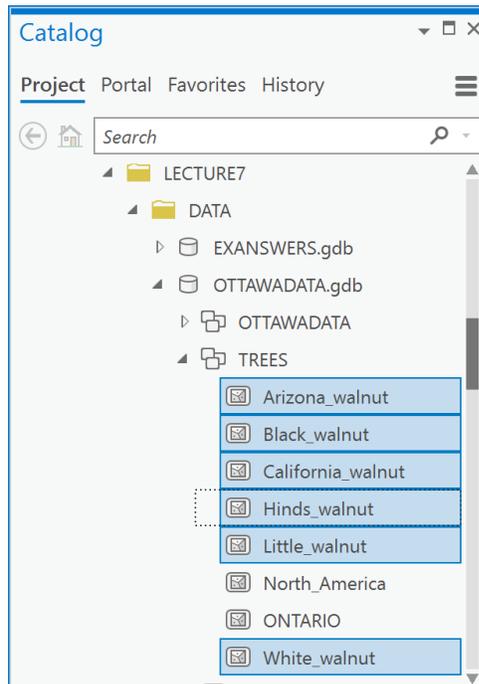


Figure 13: Exemple du processus de superposition d'union combinant la géométrie et les attributs dans le résultat pour trois espèces de noyer (*Jugulans spp.*)

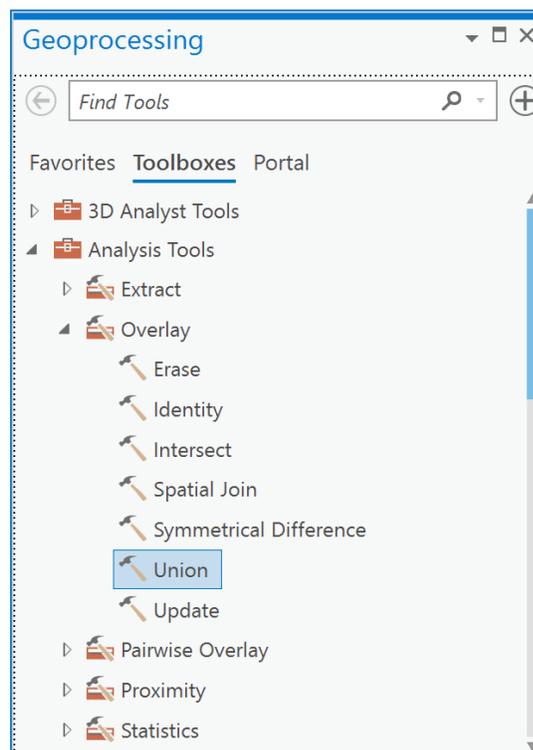
Nom de l'étudiant: _____ # d'étudiant: _____

Pour essayer ce processus par vous-même:

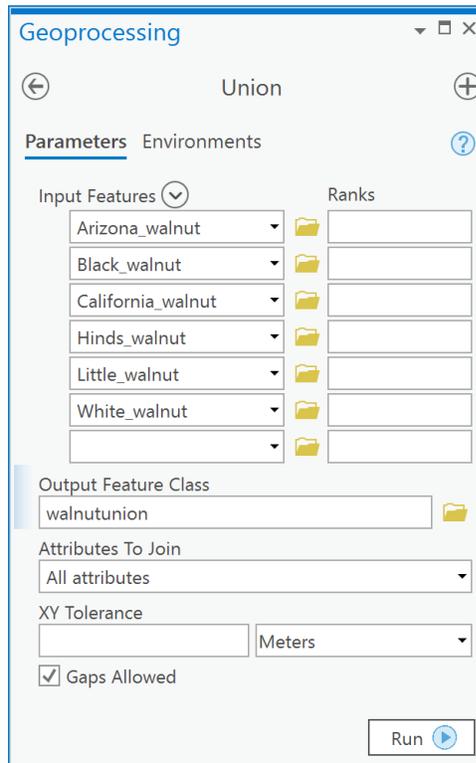
1. Ajouter les couches suivantes à ArcGIS Pro:



2. Choisissez « *Analysis Tools* » (Outils d'analyse) -> « *Union* » dans le volet « *Geoprocessing* » (Géotraitement).



3. Dans la boîte de dialogue Union, ajoutez les six couches de données et appelez le résultat 'walnutunion' (union de noyers)



4. Ouvrez la table d'attributs de la couche résultante (Notez que vous verrez beaucoup plus de champs que ceci parce que j'ai, dans la vue des champs, désactivé tous les champs sauf les champs CODE)

EXERCISE3 - ArcGIS Pro

walnutunion

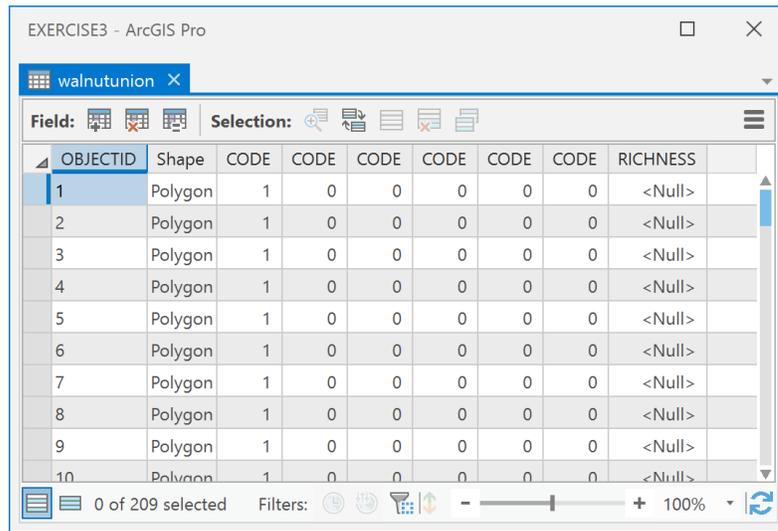
Field: Selection:

OBJECTID	Shape	CODE	CODE	CODE	CODE	CODE	CODE	
1	Polygon	1	0	0	0	0	0	
2	Polygon	1	0	0	0	0	0	
3	Polygon	1	0	0	0	0	0	
4	Polygon	1	0	0	0	0	0	
5	Polygon	1	0	0	0	0	0	
6	Polygon	1	0	0	0	0	0	
7	Polygon	1	0	0	0	0	0	
8	Polygon	1	0	0	0	0	0	
9	Polygon	1	0	0	0	0	0	
10	Polygon	1	0	0	0	0	0	

0 of 209 selected Filters: 100

Nom de l'étudiant: _____ # d'étudiant: _____

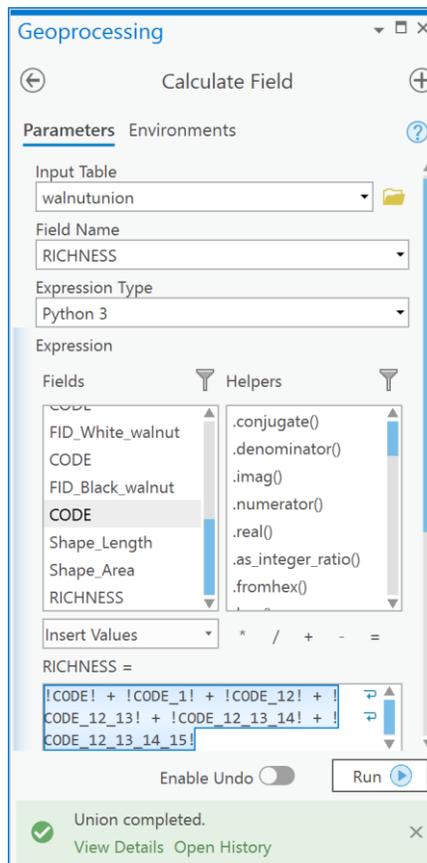
5. Ajouter un nouveau champ appelé « *RICHNESS* » de type « *Long Integer* ».



OBJECTID	Shape	CODE	CODE	CODE	CODE	CODE	CODE	RICHNESS
1	Polygon	1	0	0	0	0	0	<Null>
2	Polygon	1	0	0	0	0	0	<Null>
3	Polygon	1	0	0	0	0	0	<Null>
4	Polygon	1	0	0	0	0	0	<Null>
5	Polygon	1	0	0	0	0	0	<Null>
6	Polygon	1	0	0	0	0	0	<Null>
7	Polygon	1	0	0	0	0	0	<Null>
8	Polygon	1	0	0	0	0	0	<Null>
9	Polygon	1	0	0	0	0	0	<Null>
10	Polygon	1	0	0	0	0	0	<Null>

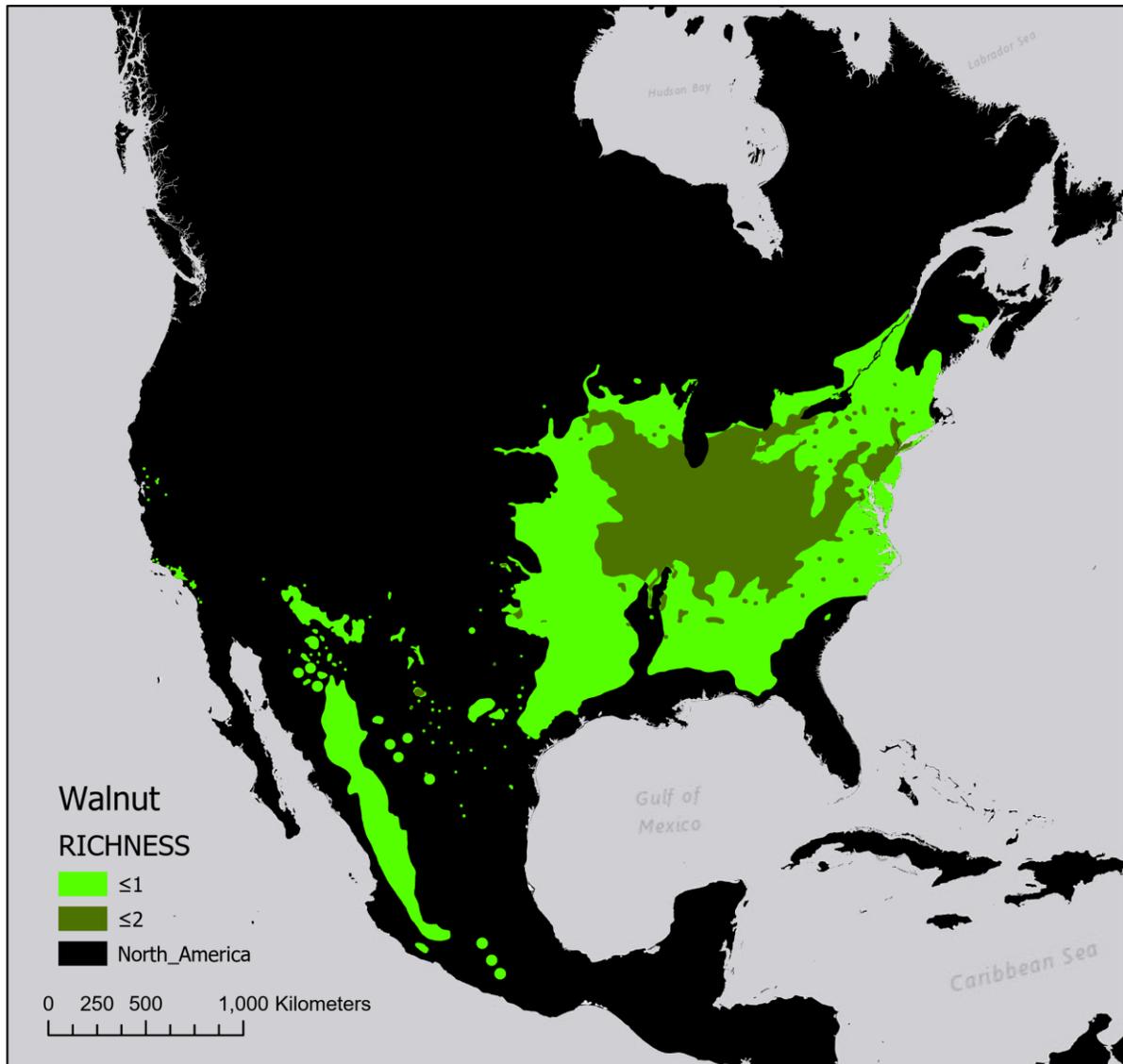
6. Calculez le nouveau champ comme la somme de chacun des champs CODE (CODE, CODE_1, CODE_12). Utilisez une expression Python telle que:

`!CODE! + !CODE_1! + !CODE_12! + !CODE_12_13! + !CODE_12_13_14! + !CODE_12_13_14_15!`



Dans la calculatrice de champs : N'oubliez pas qu'une table attributaire ne peut pas avoir deux champs du même nom dans la même table. Ainsi, lorsque le processus d'union se produit, les champs originaux appelés "CODE" dans les couches d'entrée du processus d'union sont automatiquement renommés par ArcGIS en ajoutant un "_1" ou un "_12", etc....

7. Créer une carte choroplèthe à l'aide d'une symbologie quantitative,



Comme dans nos autres exemples, nous pourrions également formuler un ensemble d'instructions Python pour accomplir cette tâche. Cet exemple est un peu plus complexe et il se peut que vous ne compreniez pas toutes les instructions, ce qui n'est pas grave. L'avantage d'un tel ensemble d'instructions Python est qu'il vous suffit de modifier les instructions surlignées pour créer une colonne de richesse en espèces pour n'importe quel nombre d'arbres.

Nom de l'étudiant: _____ # d'étudiant: _____

```
# Importer les fonctions "management" et "analysis".
from arcpy.management import *
from arcpy.analysis import *

# Définir le nom de la classe d'entités de sortie
# et le nom du champ de la richesse en espèces
out_fc="walnuts"
richnessfld="RICHNESS"

# Créer une liste de couches à unir
input_features=['White_walnut',
                'Black_walnut',
                'Little_walnut',
                'California_walnut',
                'Arizona_walnut',
                'Hinds_walnut'];

# Union toutes les couches d'arbres en gardant
# tous les attributs car nous
# avons besoin du champ CODE de chaque couche.
u1ayer=Union(input_features,out_fc,"ALL")

# Créer une liste de noms de champs car
# nous ne savons pas combien de couches
# seront unies, ni quel sera le nom de tous
# les champs CODE après le
# processus d'union. Nous savons seulement
# qu'ils commenceront par 'CODE'.
fldlist=["!{}!+".format(fld.name) for fld in
arcpy.ListFields(u1ayer,"CODE*")]

# Créer l'expression dans la calculatrice de champ
# qui ressemblera à '!CODE!+ !CODE_1!+...'
calcexpr=' '.join([s for s in fldlist])
calcexpr=calcexpr[0:-1]

# Ajouter un champ appelé "RICHNESS" à la couche unie
AddField(u1ayer,
        richnessfld,
        'LONG'
        );

# Calculez maintenant le champ RICHNESS
# en utilisant la chaîne d'expression
```

Nom de l'étudiant: _____ # d'étudiant: _____

```
# stockée dans calcexpr
CalculateField(ulayer,
               richnessfld,
               calcexpr
               );
```

Questions à remettre

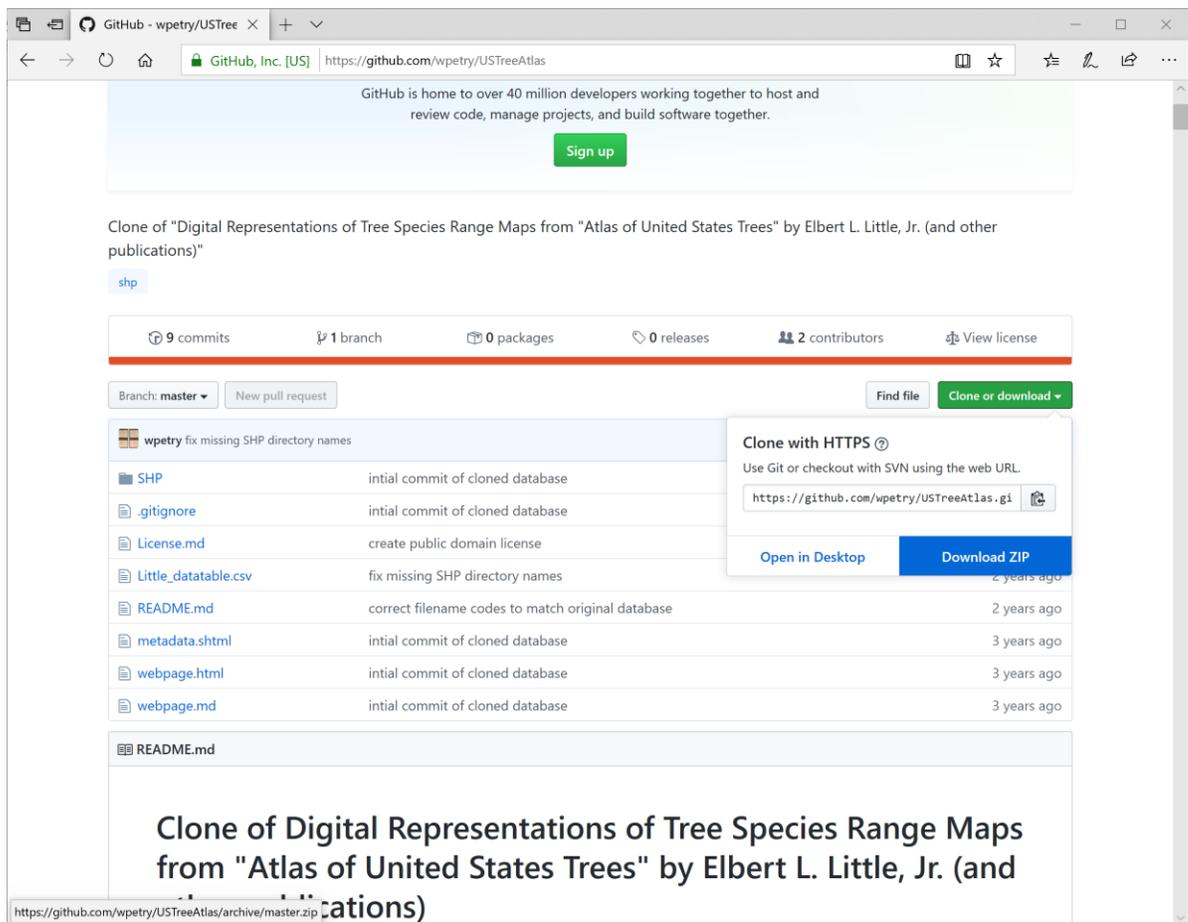
Chacune des questions nécessite une explication - lisez donc attentivement. Chaque question nécessite une carte comme élément de réponse, et il doit s'agir d'une carte correctement présentée (exclure le graticule est acceptable mais pas nécessaire). Veuillez lire la section 2.3 du manuel de style qui se trouve au début de l'exercice 1 ou 2.

1. [2 points] En utilisant des cercles de distance standard (« standard distance circles ») et des centres moyens (« mean centres »), comparez la distribution des accidents (« Accident » dans le champ SUMMARY de l'ensemble de données sur les caractéristiques de la criminalité) entre l'hiver et l'été à Ottawa en 1999. Définissez l'hiver comme la période de janvier à mars et l'été comme la période de juin à août. Présentez une carte des résultats et expliquez ce que ces résultats vous apprennent sur la répartition des accidents entre les saisons.
2. [2 points] Pour les mêmes périodes d'hiver et d'été (Q1), créez des estimations de KDE pour les accidents au cours de chaque saison. Comparez les cartes pour chaque saison.
3. [5 points] Dans le cadre de votre travail à la ville d'Ottawa, on vous demande de déterminer les zones de la ville qui devraient faire l'objet d'injections sélectives de TreeAzin, un insecticide systémique qui tue les larves de l'agrile du frêne. Le programme d'injection est sélectif et vise à "protéger les frênes dans les quartiers où le pourcentage de couverture forestière des espèces de frênes est élevé" (<http://ottawa.ca/en/residents/water-and-environment/plants-and-animals/invasive-species>). Allez sur le site Web des données ouvertes de la ville d'Ottawa et recherchez "**Tree Inventory**" (Inventaire des arbres). Téléchargez et décompressez le fichier shapefile. Importez ce fichier dans votre géodatabase OTTAWADATA.gdb (dans le dossier Lecture7). Calculer la densité des frênes (**arbres/km²**) dans les quartiers d'Ottawa. Utilisez la classe d'entités « *neighborhoods* » dans l'ensemble de données OTTAWADATA.gdb/OTTAWADATA. Lors du calcul de la densité, assurez-vous de n'inclure que ***Fraxinus*** spp. Lors de la création de votre requête, utilisez une instruction de requête telle que :

```
"SPECIES" LIKE 'Fraxinus%'
```

pour sélectionner uniquement des espèces de frênes. Présenter une carte de la densité des frênes dans les quartiers d'Ottawa. Expliquez la tendance de la densité des frênes à Ottawa. Sur la base du schéma observé, quelle recommandation donneriez-vous au directeur de la ville concernant le programme d'injection?

4. [2 marks] Présentez un KDE pour les mêmes frênes que dans (Q3) et comparez les résultats avec (Q3). Quelles sont les similitudes et les différences que vous remarquez ? Pourquoi y a-t-il des similitudes et des différences en fonction des méthodes utilisées pour calculer la densité par rapport à l'intensité ? Vos recommandations changeraient-elles ou resteraient-elles les mêmes ? Pourquoi ? (*Remarque : il peut être utile de superposer les limites des quartiers sans remplissage coloré au-dessus du KDE*).
5. [4 marks] Téléchargez les fichiers de forme pour toutes les espèces de *Fraxinus* trouvées sur <https://github.com/wpetry/USTreeAtlas>. Il y a 16 espèces au total. Vous pouvez télécharger l'ensemble du répertoire sous la forme d'un fichier zip, puis le décompresser afin d'accéder facilement aux dossiers de chaque espèce - plutôt que de télécharger chaque espèce individuellement. Pour le faire, il vous suffit de cliquer sur le bouton vert et de choisir « *Download Zip* ».



À l'aide de l'outil de superposition d'union (« *union overlay tool* »), créez une carte qui montre la richesse en espèces de *Fraxinus* en Amérique du Nord (tout dossier qui commence par "Frax..." dans le répertoire github). Notez que vous devrez utiliser l'outil « *Define Projection* » pour chacun des fichiers que vous téléchargez. Les informations sur le système de coordonnées se trouvent sur la page GitHub (les métadonnées). Présentez une carte de la richesse en espèces de *Fraxinus* dans le même système de coordonnées que l'ensemble de données **TREES** dans le fichier OTTAWADATA.gdb. Décrivez le motif de la richesse en *Fraxinus* spp. en Amérique du Nord.

Laboratoire n°5 : Télédétection et SIG raster

Objectif : L'objectif de ce laboratoire est de vous présenter un processus de classification de base par télédétection et d'évaluer la détection des changements dans les régions forestières à l'aide d'un SIG de type raster.

Résultats de l'apprentissage : À la fin de ce laboratoire, vous serez capable de :

- Effectuer une amélioration spectrale en utilisant des combinaisons de bandes et de ratios.
- Entraîner un classificateur d'images.
- Expliquer un fichier de signatures.
- Réaliser une classification supervisée en utilisant des images télédéteectées à haute et moyenne résolution.
- Nettoyer et généraliser une couche raster thématique.
- Évaluer la précision d'un résultat classifié.

Procédure:

- Lisez et travaillez sur les exemples.
- Répondez aux questions à la fin du laboratoire.

Données/matériels:

- Toutes les données sont fournies dans **lab5.zip**.

Travail:

- Répondez aux questions (#1-8) dans les espaces fournis.
- Le total des points disponibles pour ce travail est de 17 points, soit 15% de la note totale du cours.
- Tous les travaux pratiques doivent être réalisés et remis individuellement.

Format:

Les étudiants remettront les questions avec les réponses complètes.

Les tableaux et les figures doivent être produits selon les normes du Département de géographie, environnement et géomatique de l'Université d'Ottawa.
Pour des directives sur le formatage et la présentation appropriés des tableaux et des figures, veuillez consulter la [section d'introduction](#) de ce manuel de laboratoire.

Nom de l'étudiant: _____ # d'étudiant: _____

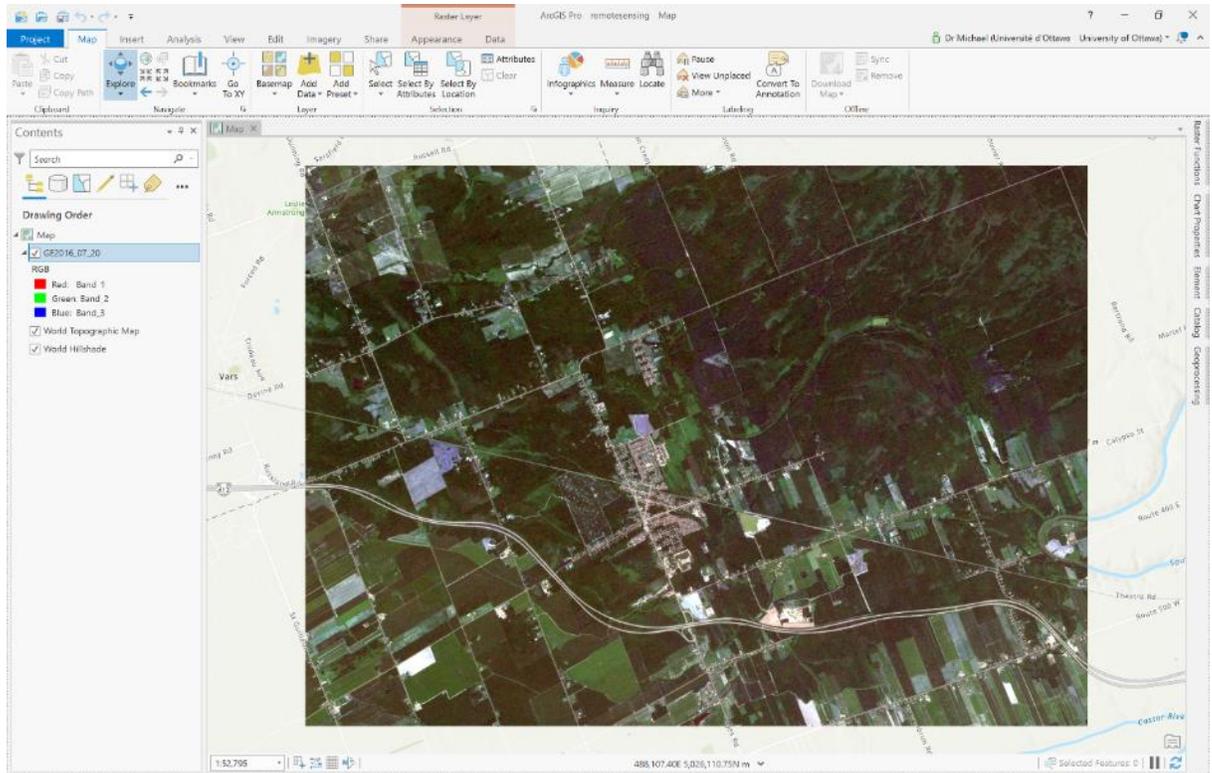
La disparition des forêts est un problème qui ne cesse de s'aggraver dans le monde d'aujourd'hui (voir <https://earthenginepartners.appspot.com/science-2013-global-forest>). Dans les questions de l'exemple, vous apprendrez à classer les forêts et les zones non forestières sur deux images satellites datant respectivement de 2001 et de 2016, puis à déterminer la quantité de forêt perdue ou gagnée entre-temps dans une petite région du sud-est d'Ottawa. Pour obtenir une estimation de la perte de forêt, dans ce laboratoire, vous ferez de la détection de changement post-classification en utilisant une image Landsat 7 ETM+ de 2001 à une résolution de 30 m, et une image RapidEye de 2016 à une résolution de 5 m. Une détection de changement post-classification vous permet d'éviter de nombreuses étapes de prétraitement nécessaires pour d'autres méthodes de détection de changement, et vous permet également d'utiliser directement les nombres numériques dans les images sans, par exemple, entreprendre des corrections radiométriques absolues ou des corrections de la réflectance du sol pour les différentes images utilisées.

Cette section présente des exemples de choses que vous devrez savoir pour répondre aux questions que vous remettrez pour obtenir vos notes. Votre tâche est de revoir et de répéter les exemples de cette section, et les couches dérivées que vous produisez dans cette section sont utilisées directement dans les questions à la fin de l'exercice. D'autres questions sont pertinentes pour vous aider à répondre aux questions à la fin de l'exercice.

Q1: Comment puis-je visualiser différentes combinaisons de bandes d'une image télédéteectée?

Diverses combinaisons de bandes peuvent vous aider à identifier différentes caractéristiques dans les images de télédétection.

1. Ajoutez l'image appelée "RE2016_07_20" du fichier IMAGERY.GDB dans le fichier lab5.zip à ArcGIS.

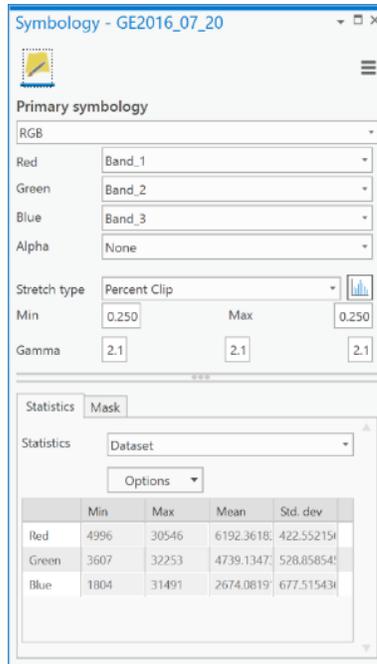


- Il s'agit d'une section d'une image multispectrale RapidEye acquise le 20 juillet 2016. Cette constellation de satellites est utilisée pour la surveillance et la recherche agricole et phénologique en raison du temps de revisite quotidien (hors du nadir, 5,5 jours au nadir) dû à une constellation de 5 satellites. Le capteur est donc utile pour surveiller le stress des cultures ou l'apparition de maladies dans de vastes régions. Le capteur RapidEye est multispectral, avec les bandes suivantes:

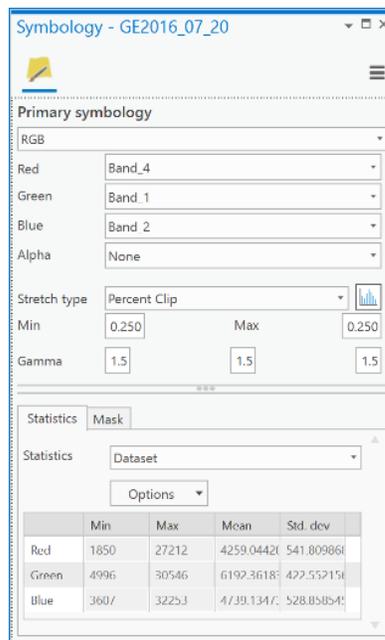
- 440 - 510 nm (bleu)
- 520 - 590 nm (vert)
- 630 - 685 nm (rouge)
- 690 - 730 nm (bord rouge)
- 760 - 850 nm (proche IR)

Cette image n'est pas corrigée atmosphériquement. Dans la plupart des cas, l'utilisation de cette image pour évaluer le stress des plantes ou la phénologie nécessiterait une correction atmosphérique, comme vous pourrez l'apprendre dans un autre cours, mais nous ne le ferons pas ici. De plus, l'image que vous utilisez ici est orthorectifiée, mais uniquement à partir d'une élévation constante, ce qui n'est pas très précis.

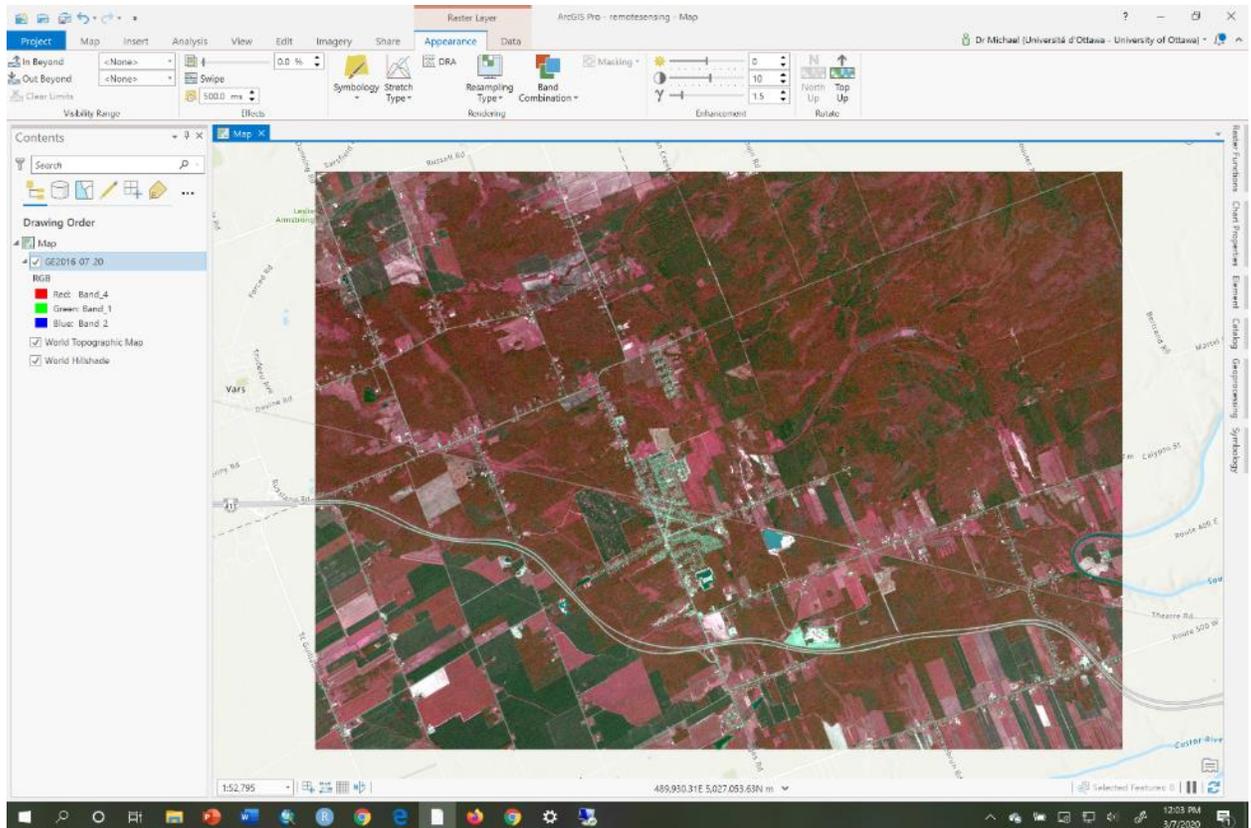
- Pour visualiser d'autres combinaisons de bandes de l'image, cliquez avec le bouton droit de la souris sur la couche "RE2016_07_20" dans le volet de contenu (« *Contents* »), puis sélectionnez symbologie (« *Symbology* »):



4. Par défaut, ArcGIS affiche la combinaison de bandes de couleurs visibles 3, 2, 1 pour le rouge, le vert et le bleu. Vous pouvez changer cette combinaison en Fausses couleurs infrarouges (« *False Color Infrared* »), par exemple, en choisissant 4,1,2 (Band_4 pour la bande rouge, Band_3 pour la bande verte et Band_2 pour la bande bleue) dans la colonne Band de l'onglet Symbology:



5. La carte résultante est un "composite fausses couleurs" et indique en rouge les endroits où la végétation est présente.



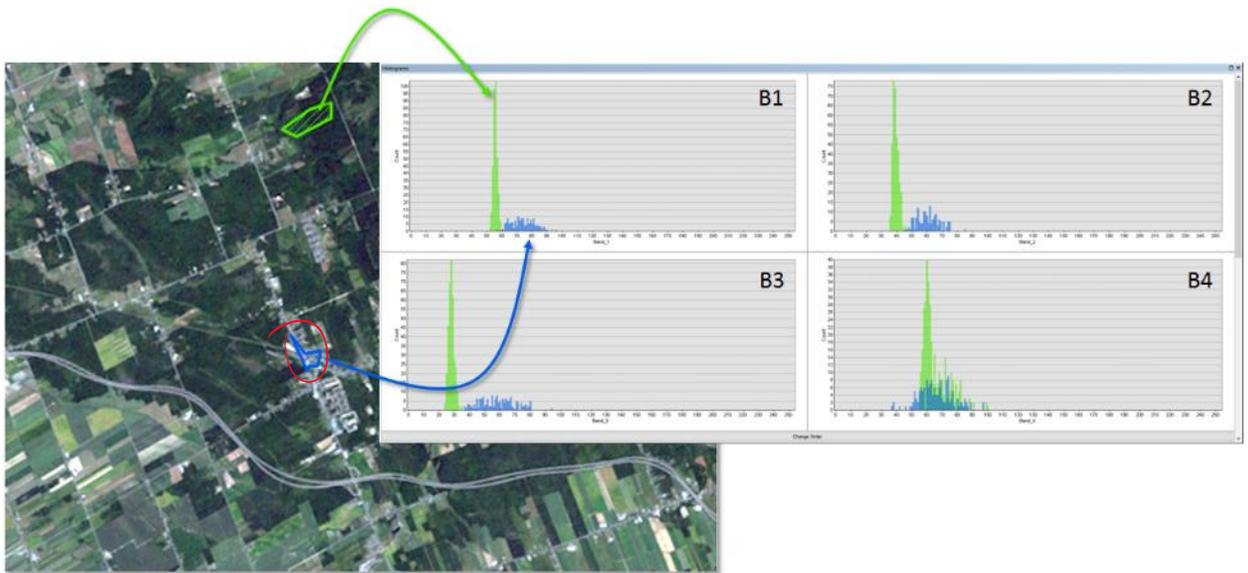
6. Vous pouvez essayer d'autres combinaisons de bandes pour créer différents composites de fausses couleurs, dont certains peuvent visuellement mettre en valeur différents éléments du paysage imagé.

Q2: Comment classifier une image télédéteectée?

Dans ce laboratoire, nous allons effectuer ce que l'on appelle une "classification supervisée", dans laquelle l'analyste d'images (vous) fournit à l'ordinateur des exemples de pixels contenant différentes caractéristiques du paysage à traiter en tant que classes d'information. Une classe d'information est une catégorie nominale que vous souhaitez extraire d'une image télédéteectée. Par exemple, si vous souhaitez extraire les zones forestières d'une image, vous pouvez créer une classe d'information appelée **Forêt** et une autre appelée **Non-Forêt**. Vous pouvez ensuite trouver des exemples de pixels appartenant à la classe d'information définie par l'utilisateur appelée **Forêt** et un autre ensemble de pixels pour la classe appelée **Non-Forêt**. Vous pouvez avoir de nombreuses autres classes d'information en fonction du type d'information que vous souhaitez extraire d'une image. Par exemple, les classes d'information peuvent inclure "eau", "terres cultivées", "résidentiel", "commercial", "forêt dense", "forêt ouverte", etc. En fonction du type d'image obtenue par télédéteecté et de sa résolution spatiale et spectrale, il est possible d'extraire de nombreux types d'informations.

Pour effectuer une classification supervisée, vous avez besoin de deux choses:

1. Données d'apprentissage : Les données d'apprentissage sont constituées de pixels de l'image appartenant à différentes classes d'informations que vous souhaitez extraire de l'image pour créer une couche raster thématique. Par exemple, les pixels appartenant à la catégorie **Forêt** et les pixels appartenant à la catégorie **Non-Forêt** sont deux exemples de classes d'information. Examinez la figure ci-dessous. Le polygone vert est un polygone d'entraînement qui englobe un ensemble de pixels (toutes bandes confondues) que l'analyste a identifiés comme appartenant à la classe d'information appelée Forêt dans l'image. La fréquence absolue (le nombre de pixels) des valeurs de pixels "**Forêt**" avec différents valeurs numériques (VN), à l'intérieur du polygone vert, est représentée sous forme d'histogrammes verts pour chaque bande. De même, le polygone bleu contient des pixels non forestiers et constitue un site d'entraînement contenant des pixels non forestiers dans toutes les bandes de l'image. La fréquence absolue des VN dans la catégorie d'information **Non-Forest** pour chaque bande est représentée sous forme d'histogrammes bleus. Pour que la classification soit précise, les pixels d'entraînement de chaque classe d'information doivent avoir séparé au moins un des histogrammes, c'est-à-dire que les histogrammes de chaque classe d'entraînement doivent se chevaucher le moins possible. Ce point est abordé en détail plus loin. Dans ce cas, la séparabilité est observée dans les bandes B1, B2, B3, mais pas dans la bande B4.



2. Un fichier de signature : Un fichier de signature porte l'extension **".gsg"** et contient un résumé statistique des différentes valeurs de pixels dans chacune des classes d'information spécifiées dans l'ensemble de données d'apprentissage :

```

I7_2001_08_25t.gsg - Notepad
File Edit Format View Help
# Signatures Produced by ClassSig from
# Class-Grid __1000001
# and Stack __1000000

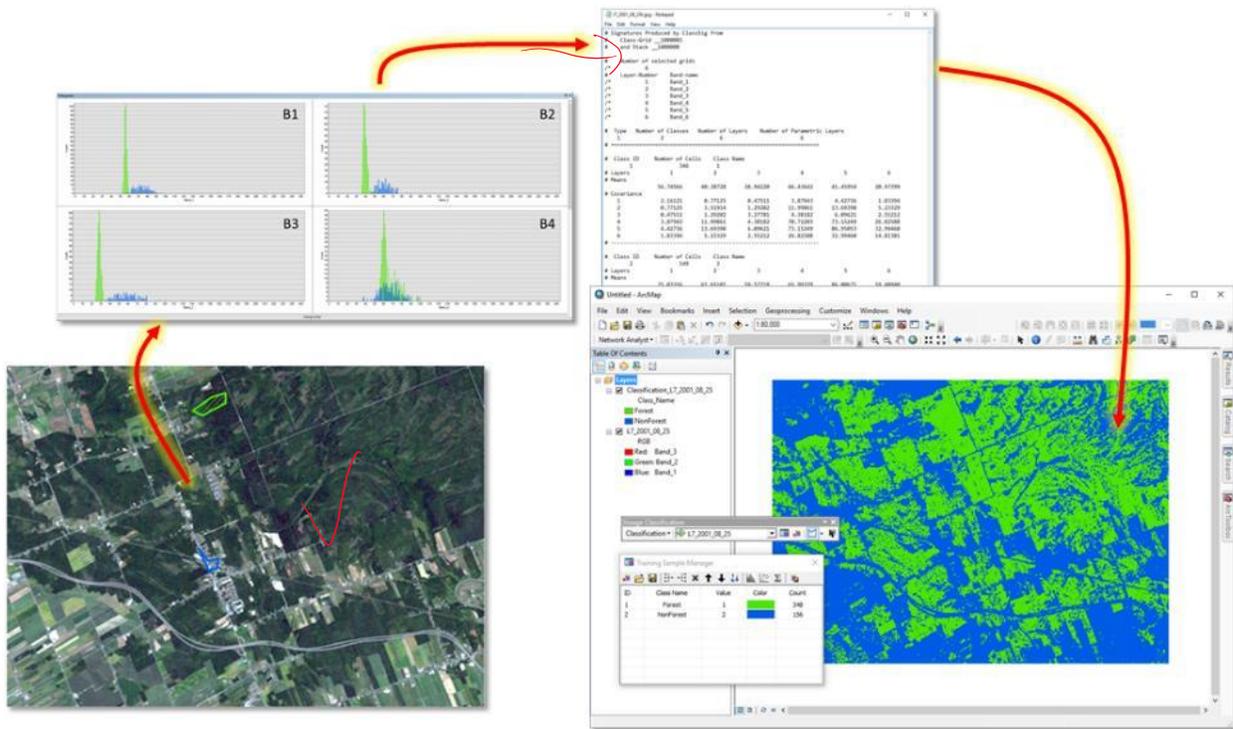
# Number of selected grids
/*
6
# Layer-Number Band-name
/*
1 Band_1
/*
2 Band_2
/*
3 Band_3
/*
4 Band_4
/*
5 Band_5
/*
6 Band_6

# Type Number of Classes Number of Layers Number of Parametric Layers
1 2 6 6
# -----
# Class ID Number of Cells Class Name
1 346 1
# Layers 1 2 3 4 5 6
# Means 56.74566 40.38728 28.94220 66.43642 41.45954 20.97399
# Covariance
1 2.16121 0.77125 0.47511 3.87943 4.42736 1.83394
2 0.77125 3.51914 1.29202 11.99861 13.69398 5.23329
3 0.47511 1.29202 3.27781 4.38182 6.09621 2.55212
4 3.87943 11.99861 4.38182 70.72203 73.15249 26.82588
5 4.42736 13.69398 6.09621 73.15249 86.95053 32.99460
6 1.83394 5.23329 2.55212 26.82588 32.99460 14.81381
# -----
# Class ID Number of Cells Class Name
2 149 2
# Layers 1 2 3 4 5 6
# Means 75.03356 61.65101 59.57718 65.99329 86.00671 59.40940
# Covariance
1 56.16779 50.66044 78.50077 -58.31058 7.27680 58.92536
2 50.66044 51.12062 77.10820 -46.60371 17.95506 61.78573
3 78.50077 77.10820 128.92137 -81.39475 32.30015 104.02562
4 -58.31058 -46.60371 -81.39475 119.56077 46.36491 -45.51075
5 7.27680 17.95506 32.30015 46.36491 147.54725 84.61886
6 58.92536 61.78573 104.02562 -45.51075 84.61886 122.89207

```

Le fichier de signature ci-dessus spécifie les valeurs moyennes des histogrammes et leur variabilité (covariance) au sein de chaque bande d'image. Le fichier de signature est un résumé statistique des histogrammes de chaque bande, pour chaque classe d'information. Par exemple, le fichier contient la moyenne et la covariance des VN pour les pixels appartenant à la classe Forêt, et de même pour Non-Forêt, pour chaque bande. Ces informations permettent à un classificateur automatique de placer chaque pixel de l'image dans sa classe respective.

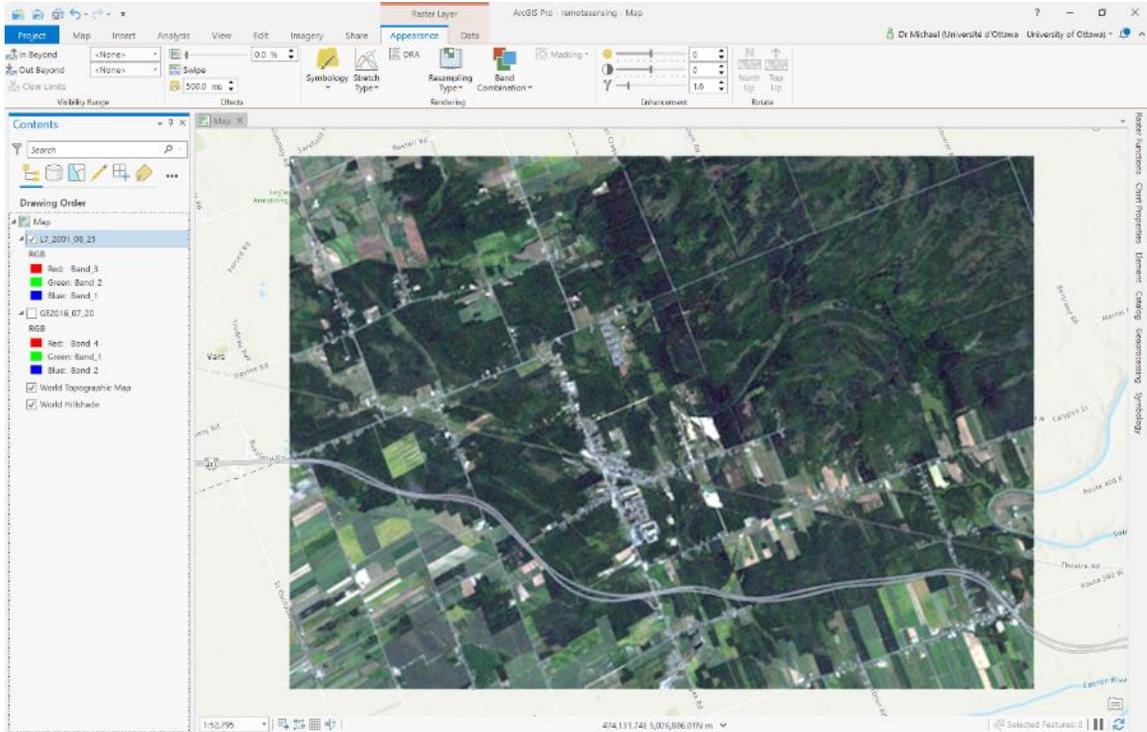
À l'aide du fichier de signatures créé à partir des histogrammes d'entraînement, une classification supervisée attribue tous les pixels de l'image aux classes d'information spécifiées dans le fichier de signatures. Le classificateur examinera votre fichier de signatures et attribuera chaque pixel de votre image à l'une des catégories d'informations que vous avez définies dans l'ensemble de données d'apprentissage. Le résultat du classificateur est une couche raster thématique qui peut ensuite être utilisée dans l'algèbre de cartes ou dans d'autres processus de modélisation raster:



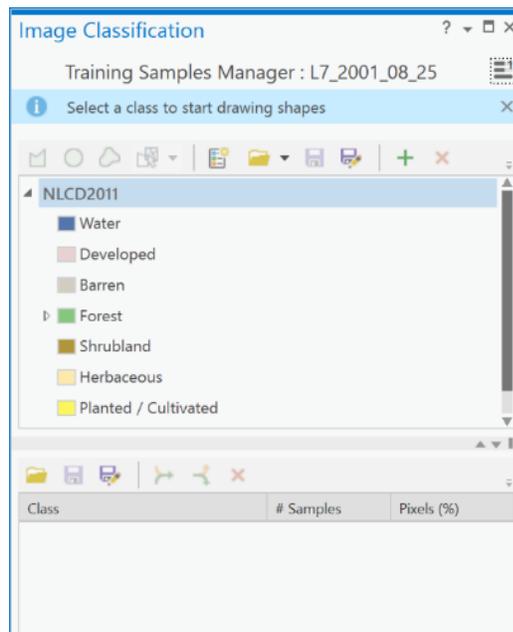
Les classificateurs sont des algorithmes d'apprentissage automatique qui apprennent, à partir de vos données d'apprentissage (dans le fichier de signature), à classer n'importe quel pixel de l'image, y compris ceux qui ne font pas partie de votre ensemble d'apprentissage. Il existe de nombreux types de classificateurs qui présentent des avantages et des désavantages différents ; ils portent des noms tels que Maximum de vraisemblance, Réseaux neuronaux, Machines à vecteurs de support, K-means, Forêt aléatoire et Iso-Cluster. Quel que soit le classificateur utilisé, la qualité des données d'apprentissage est primordiale et a le plus grand effet sur la précision de la couche cartographique résultante.

Ok, commençons. Nous allons d'abord travailler avec une image Landsat 7 du 25 août 2001 appelée "L7_2001_08_25" dans le fichier IMAGERY.GDB:

3. Ajoutez le fichier "L7_2001_08_25" à ArcGIS et veillez à le sélectionner dans le volet *Contents*:



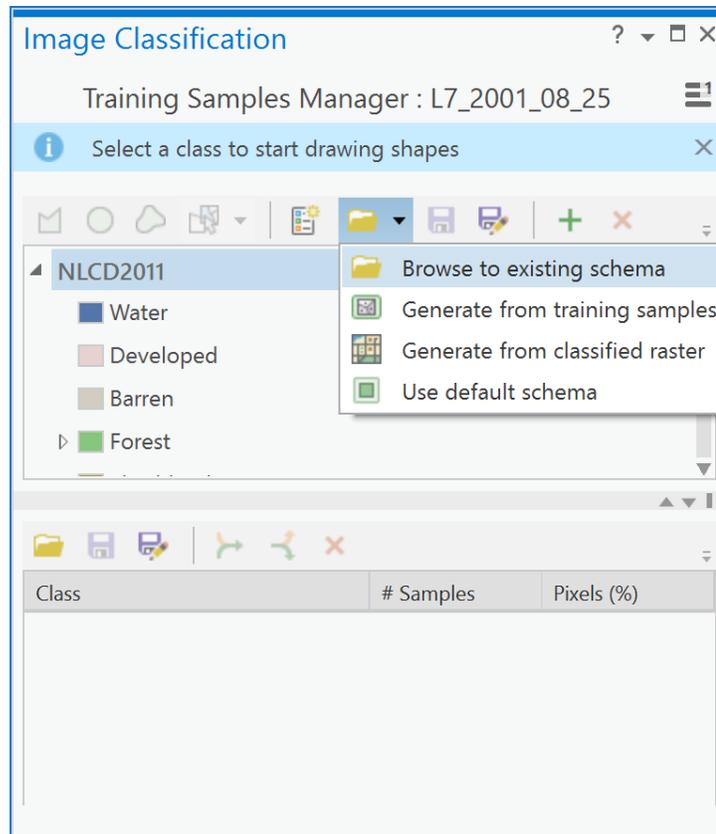
4. Cliquez sur l'onglet Imagerie (« *Imagery* »), puis sur le bouton Outils de classification (« *Classification Tools* ») et choisissez le gestionnaire d'échantillons d'entraînement (« *Training Samples Manager* »). Il se peut qu'il y ait un ensemble de classes d'information par défaut, ne vous inquiétez pas, nous nous en occuperons dans les étapes suivantes.



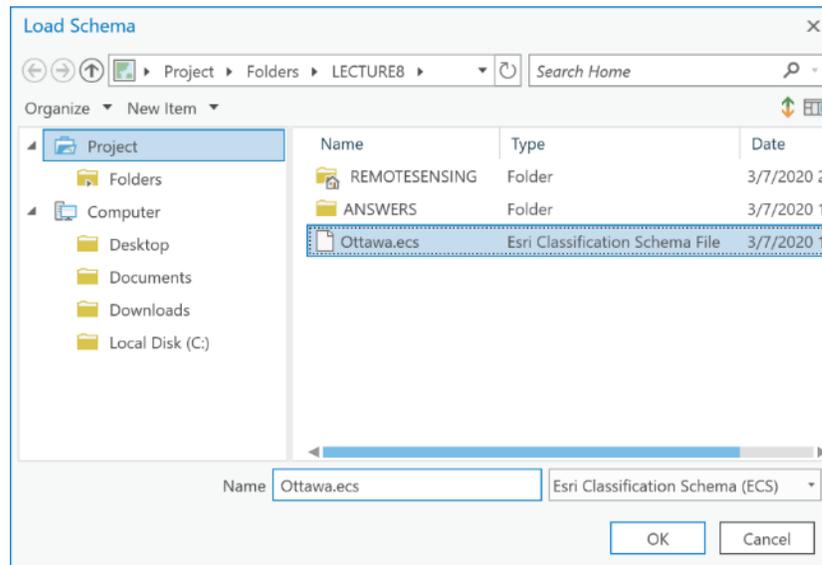
5. Voici la description du gestionnaire d'échantillons d'entraînement (« *Training Samples Manager* »).

Outil	Fonction
	Créer un échantillon d'entraînement en dessinant un point sur les pixels ou les objets dans le raster.
	Créer un échantillon d'entraînement en traçant un polygone autour des pixels ou objets dans le raster.
	Créer un échantillon d'entraînement en traçant un cercle autour des pixels ou objets dans le raster.
	Créer un échantillon d'entraînement en traçant une forme à main levée autour des pixels ou objets dans le raster.
	Créer un échantillon d'entraînement en sélectionnant un segment à partir d'une couche segmentée. Cette option est seulement disponible s'il existe une couche segmentée dans la fenêtre Contents (Contenu) . Activer Segment Picker (Sélecteur de segment) en mettant en surbrillance la couche segmentée dans la fenêtre Contents (Contenu) , puis sélectionner la couche dans la liste déroulante Segment Picker (Sélecteur de segment) .
	Créer une structure de classification. Cliquer avec le bouton droit sur le titre New Schema (Nouvelle structure) et cliquez sur Add New Class (Ajouter une nouvelle classe) .
	Sélectionner une option relative à la structure de classification. <ul style="list-style-type: none"> • Accéder à une structure existante. • Générer une structure à partir d'une classe d'entités d'échantillons d'entraînement existante. • Générer une structure à partir d'un raster classé existant. • Utiliser la structure par défaut, 2011 National Land Cover Database.
	Enregistrer les modifications apportées à la structure.
	Enregistrer une copie de la structure.
	Ajouter une catégorie de classe à la structure. Sélectionner le nom de la structure pour créer une classe parent au niveau le plus élevé. Sélectionner le nom d'une classe existante pour créer une sous-classe.
	Supprimer la classe sélectionnée ou la catégorie de sous-classe de la structure.

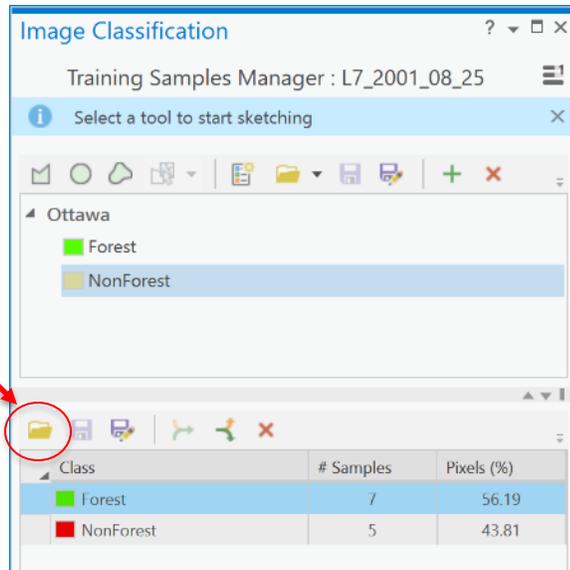
6. En utilisant le gestionnaire d'échantillons d'entraînement, vous ouvrirez un jeu d'entraînement préexistant de polygones dans une classe d'entités appelée "L7_2001_08_25_Training" dans la géodatabase IMAGERY.GDB. Cliquez d'abord sur le bouton Schéma de classification (« *Classification schema* ») et choisissez Parcourir vers un schéma existant (« *Browse to an existing schema* »):



7. Naviguez jusqu'à l'endroit où vous avez dézippé le fichier lab5.zip, choisissez le fichier "Ottawa.ecs" et cliquez sur OK:

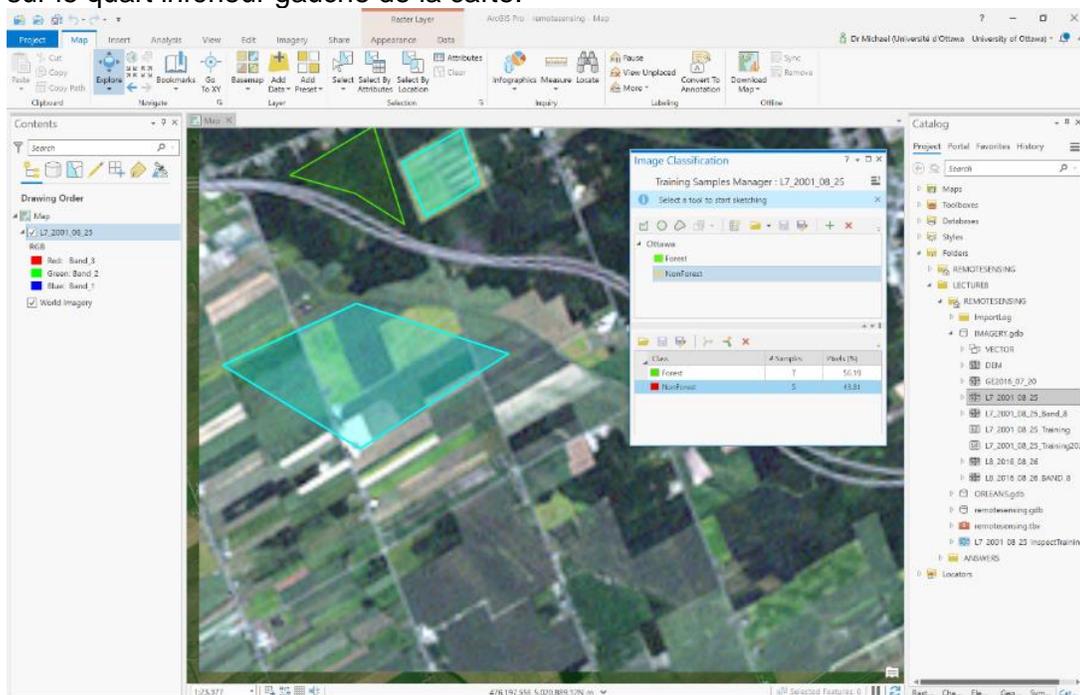


8. Cliquez maintenant sur le bouton Charger des échantillons d'entraînement (« Load training samples ») dans le demi-volet du bas, naviguez jusqu'à la classe d'objets "L7_2001_08_25_Training" et choisissez-la dans IMAGERY.GBD. Votre volet Gestionnaire d'échantillons d'entraînement devrait maintenant ressembler à ceci:

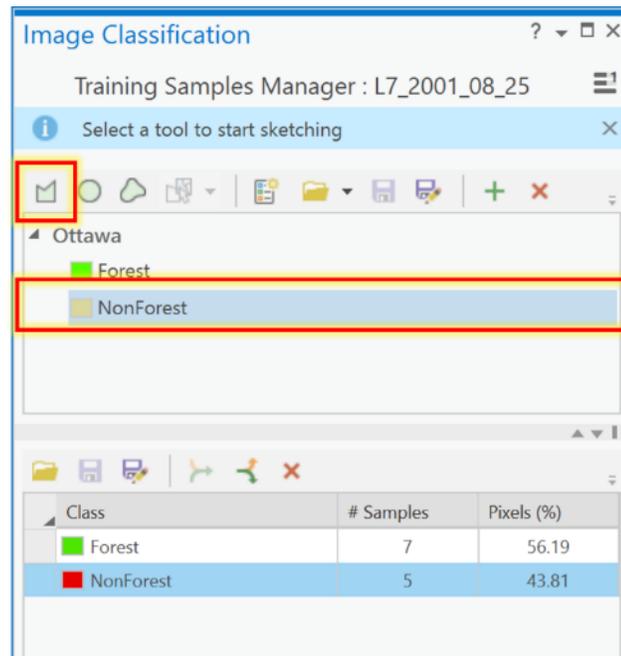


Vous pouvez voir que la classe **Forêt** correspond à la couleur verte et que les sites d'entraînement sont représentés par des polygones verts sur la carte. De même, pour la classe d'information **Non-Forest**, les polygones sont représentés en rouge. Ce qui est important ici, c'est que les deux classes d'information, **Forêt** et **Non-Forêt**, aient approximativement le même pourcentage de pixels dans chaque classe. À l'heure actuelle, la classe **Forêt** contient 56 % des pixels et la classe **Non-Forêt** 44 %. Ce n'est pas si mal, mais nous allons ajouter un nouveau polygone à la classe **Non-Forêt** pour augmenter son pourcentage.

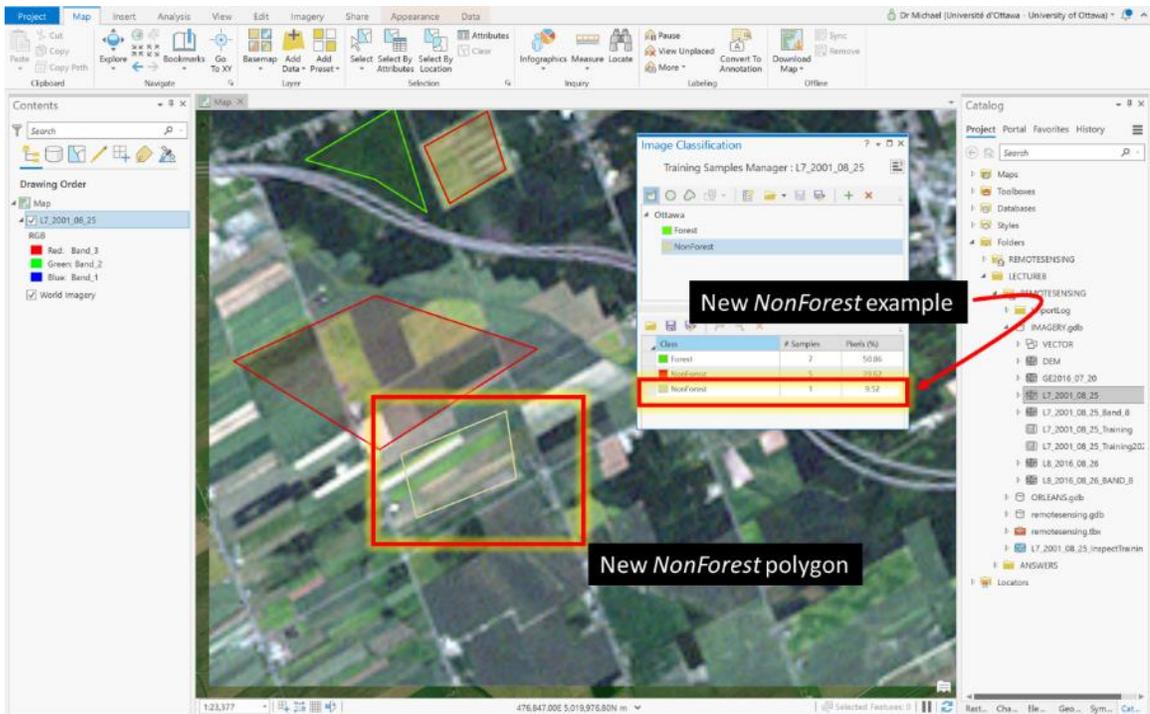
9. Pour ajouter un autre exemple de Non-Forêt aux échantillons d'entraînement, zoomez sur le quart inférieur gauche de la carte:



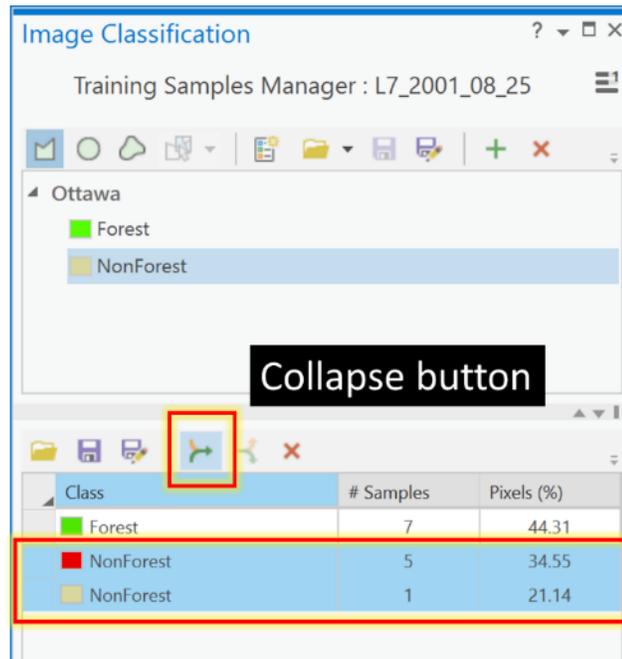
10. Ensuite, dans le Gestionnaire d'échantillons d'entraînement, dans le demi-volet du haut, sous le schéma appelé Ottawa, cliquez sur la classe d'information **NonForest** (non forestier). Vous verrez que l'outil polygone de la barre d'outils devient actif (il s'agit de l'outil le plus à gauche; des versions plus récentes d'ArcGIS peuvent avoir un bouton Rectangle en premier, ce qui fait que le bouton Polygone est le deuxième. Veillez à sélectionner l'outil approprié.).



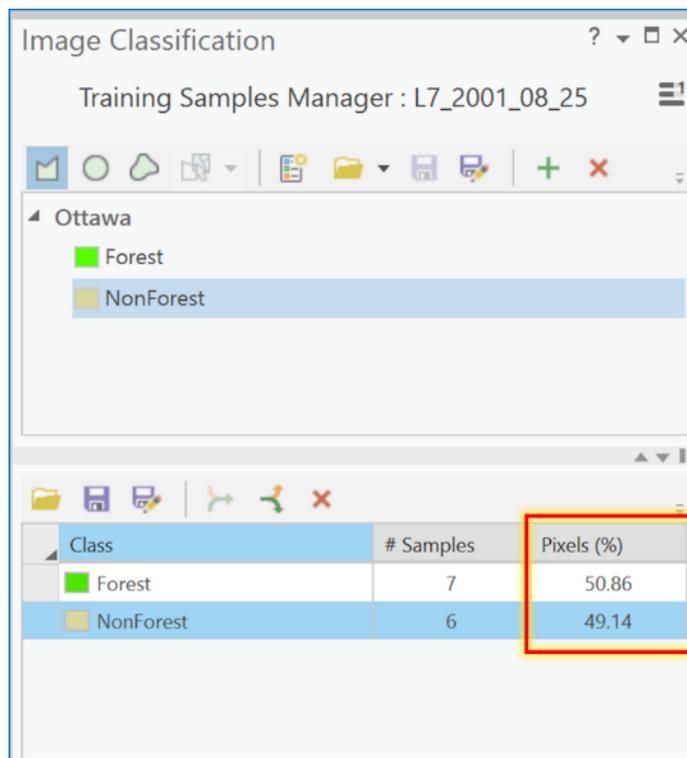
11. Cliquez maintenant sur l'outil polygone et dessinez un polygone dans n'importe quelle zone non boisée de l'image satellite. **Ne choisissez pas d'autre outil de dessin que l'outil polygone pour dessiner votre polygone:**



12. Examinez le volet Gestionnaire d'échantillons d'entraînement, le demi-volet du bas, et vous verrez maintenant les polygones de l'exemple non forestier d'origine en rouge et le nouvel exemple non forestier que vous venez de créer. Appuyez sur la touche Ctrl et cliquez sur les deux exemples non forestiers, puis cliquez sur le bouton Réduire (« Collapse ») pour fusionner les deux exemples non forestiers en un seul ensemble:



13. Le volet Gestionnaire d'échantillons d'entraînement devrait ressembler maintenant à ceci:

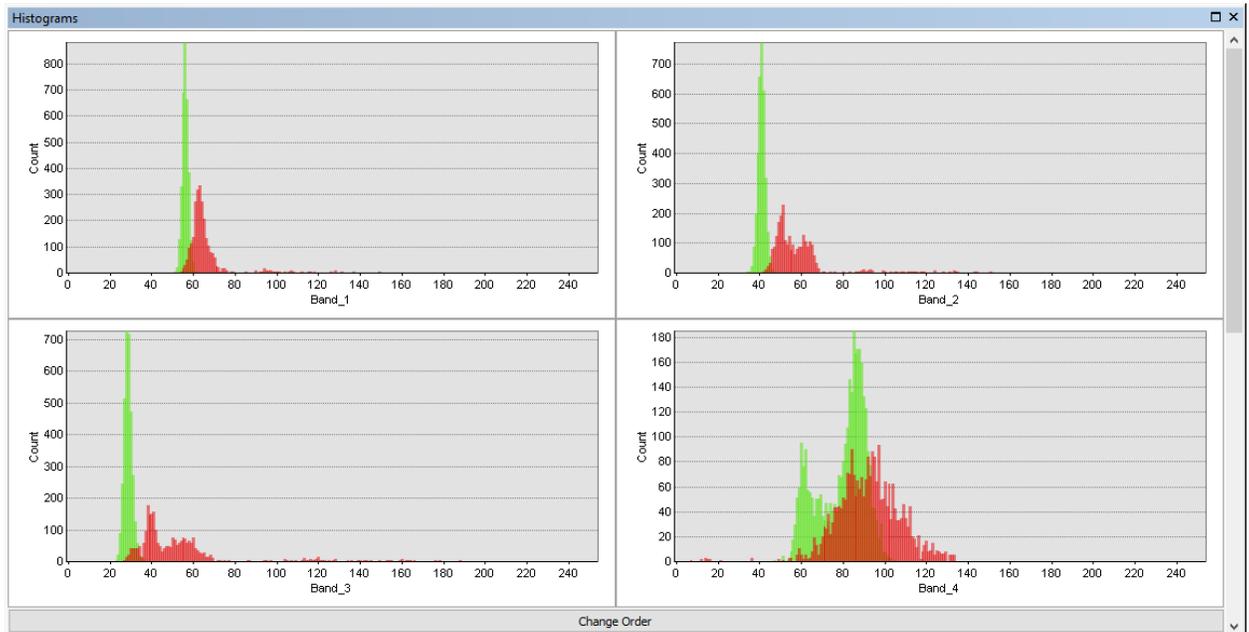


14. Vérifiez que les deux classes d'information ont maintenant environ 50 % de pixels chacune. Si le polygone que vous avez ajouté était trop grand, la classe **Non- Forêt** aura maintenant beaucoup plus que 50 % des pixels, ce qui signifie que vous n'avez pas réussi à équilibrer la taille des échantillons d'entraînement des deux classes. Dans ce cas, vous pouvez soit ajouter un polygone à la classe **Forêt** (et ainsi de suite, jusqu'à ce que vous parveniez à un résultat correct), ou vous pouvez faire ceci:

- Sélectionnez la classe **NonForest** dans la moitié du bas du gestionnaire d'échantillons d'entraînement et cliquez sur le bouton "*Expand*" (à côté du bouton "*Collapse*").
- Vous verrez maintenant plusieurs lignes étiquetées **NonForest**, chacune correspondant à un polygone individuel.
- Supprimez la dernière - celle que vous venez de créer.
- Ajoutez-en un nouveau (plus petit).
- Sélectionnez tous les polygones **NonForest** et réduisez-les ("*Collapse*").
- Vérifiez que la répartition entre les deux classes est plus ou moins égale.

15. Examinez attentivement le volet Gestionnaire d'échantillons d'entraînement, qui fournit de nombreuses informations importantes sur les échantillons d'entraînement pour chaque classe d'information. Une chose qui n'est pas montrée est que **Forest** (la première classe listée) reçoit une valeur de pixel de 1 et **NonForest** (la classe suivante listée) reçoit une valeur de pixel de 2 (et ainsi de suite, si vous avez plus de classes). Ces valeurs deviendront les valeurs des cellules pour toute classification de l'image. La colonne "Pixels %" indique combien de pixels se trouvent dans chaque classe d'apprentissage. Dans le cas ci-dessus, 50,86 % des pixels d'entraînement se trouvent dans la classe d'information **Forêt** et 49,14 % dans l'ensemble d'entraînement **Non-Forêt**. Nous avons donc obtenu une proportion presque égale, ce qui est acceptable et meilleur que notre point de départ.
16. Ensuite, nous devons enregistrer nos nouveaux polygones d'entraînement dans une classe d'objet. Pour cela, cliquez sur le bouton "Save" dans le volet du bas du Gestionnaire d'échantillons d'entraînement, naviguez jusqu'au fichier IMAGERY.GDB et enregistrez le nouveau fichier d'entraînement sous le nom "L7_2001_08_25_Training2023". Nous n'avons pas besoin de sauvegarder le schéma "Ottawa" dans le demi-volet du haut parce qu'il n'a pas changé. Si nous avons ajouté une nouvelle classe d'information, nous aurions également dû enregistrer le nouveau schéma.
17. L'étape suivante est fondamentale car elle nous permettra d'explorer le degré de "pureté" de nos classes d'information sur la base des pixels que nous avons choisis pour chaque classe. Nous voulons savoir quelles sont les bandes qui présentent la plus grande *séparabilité* entre les forêts et les zones non forestières. La **séparabilité spectrale** fait référence à la distance spectrale entre les valeurs des pixels qui composent chaque classe d'information dans un ensemble de données d'échantillons d'entraînement. En fin de compte, nous voulons savoir quelles bandes de notre image de télédétection seraient les meilleures pour classer notre image en **Forêt** et **Non-Forêt**. Les bandes les plus performantes permettront de distinguer clairement nos deux classes d'information et il n'y aura que peu ou pas de chevauchement entre les valeurs des pixels correspondant à la catégorie **Forêt** et **Non-Forêt** dans ces bandes.

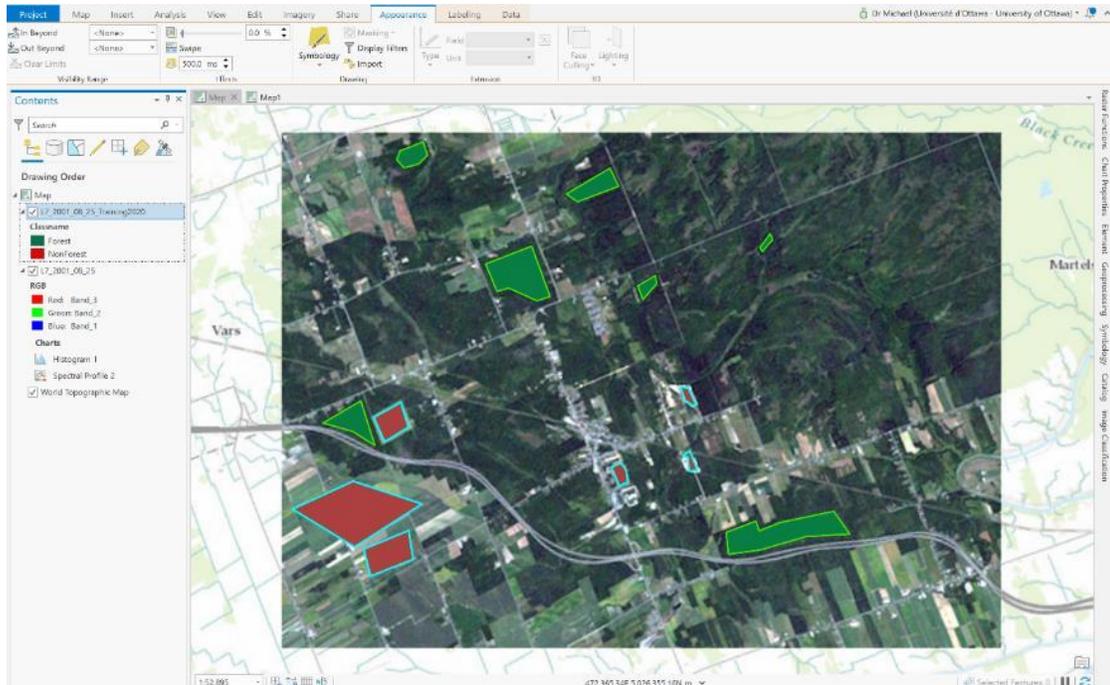
Pour comprendre la séparabilité spectrale, regardez la figure ci-dessous pour voir les histogrammes pour quatre bandes et deux classes d'information différenciées par les couleurs rouge et verte. Chaque histogramme montre la fréquence des valeurs numériques pour chaque classe, classe rouge ou classe verte, sur l'axe des ordonnées (y) et l'étendue des valeurs numériques associées à la classe rouge ou verte sur l'axe des abscisses (x). Par exemple, dans la Bande_1, nous voyons que la classe verte se situe entre 52 et 60 et la classe rouge entre 55 et 150. Ainsi, bien que les deux histogrammes soient clairement différents, il y a également un certain chevauchement autour des valeurs de 55-60, indiquées en rouge foncé. Toutefois, les pics des histogrammes rouge et vert ne se chevauchent pas, ce qui est une bonne chose. En général, une "bonne" bande pour la classification est une bande dont les histogrammes rouge et vert ne se chevauchent pas du tout. S'il n'y a pas de chevauchement entre les classes d'information dans une bande, les classes rouge et verte sont distinctes et complètement séparées. Par conséquent, une telle bande permettrait de classer l'image dans les classes d'information rouge et verte. Ici, aucune des bandes de l'image ne présente une séparabilité complète. La bande présentant le plus faible chevauchement est la bande_2. Pourquoi ? Parce qu'il y a peu de chevauchement entre les histogrammes rouge et vert. Inversement, la bande la plus défavorable est la bande_4, où le chevauchement est considérable et où la bande_4 n'aiderait donc pas à classer l'image correspondante dans les bandes rouge et verte.



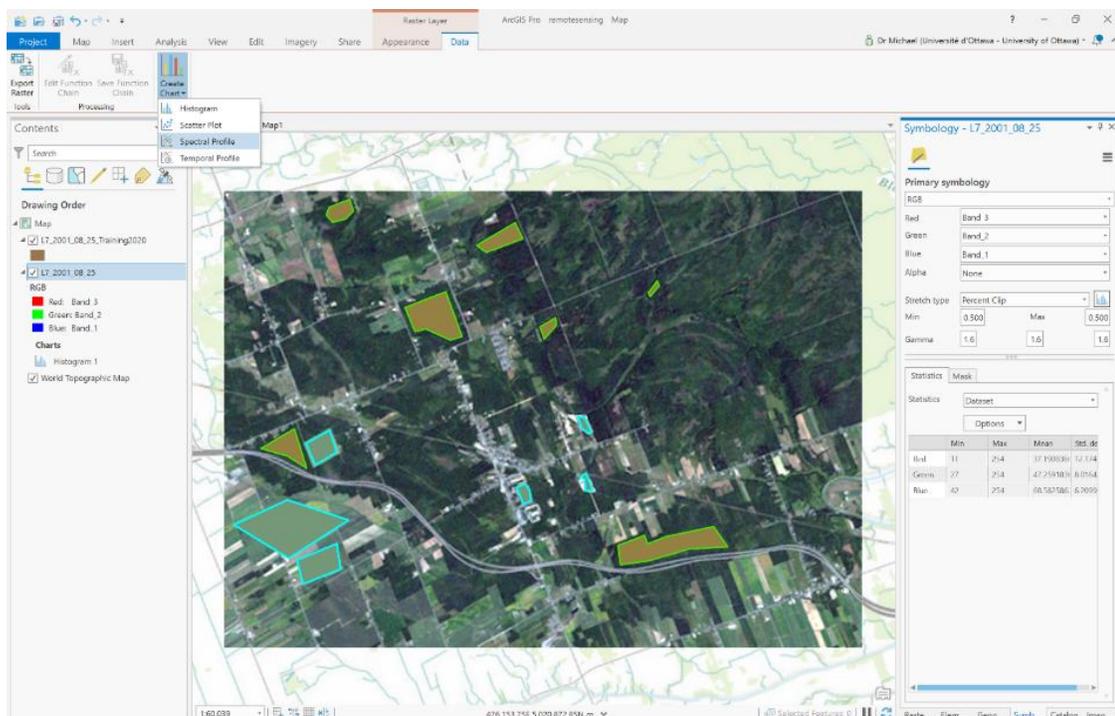
Il n'est pas nécessaire de générer des histogrammes pour chaque bande afin d'examiner la séparabilité. La façon la plus courante d'examiner la séparabilité spectrale entre les différentes bandes est de générer un graphique des statistiques de chaque histogramme pour chaque bande dans un profil spectral.

Nom de l'étudiant: _____ # d'étudiant: _____

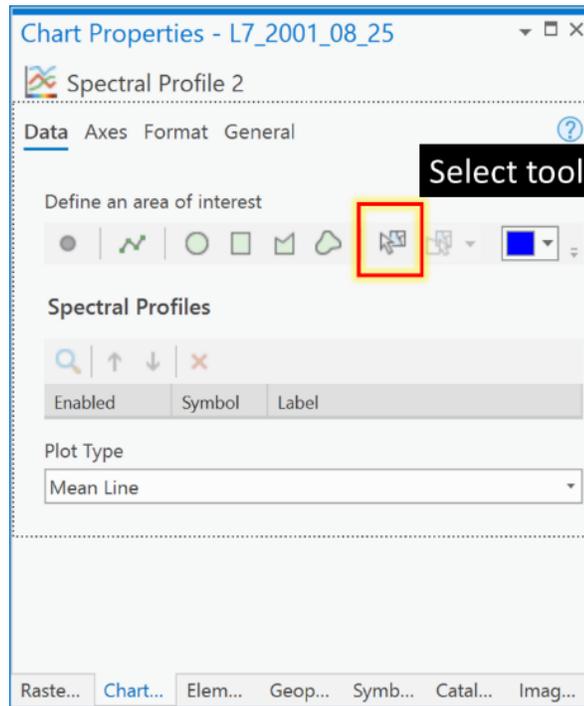
18. Pour créer un graphique spectral, assurez-vous d'abord que vous avez ajouté votre couche "L7_2001_08_25_Training2023" au panneau de contenu, allez dans la symbologie, et symbolisez-la en utilisant des valeurs uniques avec le champ « *ClassName* », et mettez **Forêt** en vert et **Non-Forêt** en rouge.



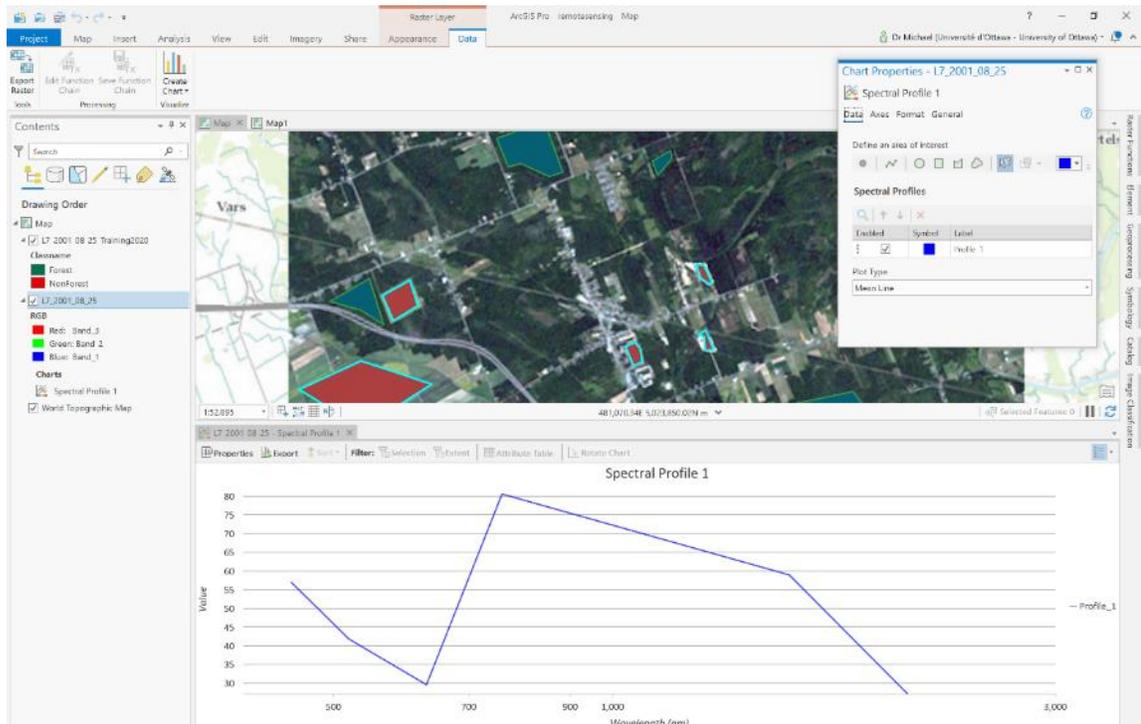
19. Ensuite, choisissez la couche "L7_2001_08_25" dans le volet de contenu. Cliquez ensuite sur l'onglet Données (« *Data* »). Cliquez maintenant sur Créer un graphique (« *Create Chart* ») et choisissez Profil spectral (« *Spectral Profile* »):



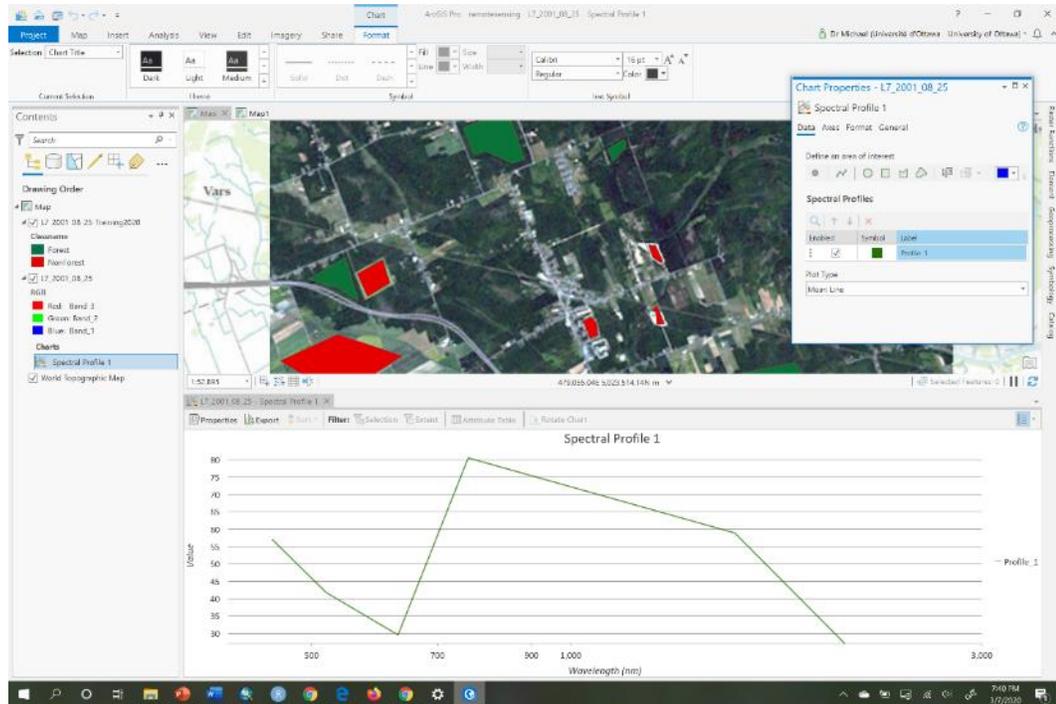
20. Ensuite, dans le volet Propriétés du graphique (« *Chart Properties* »), cliquez sur le bouton Sélecteur d'entités (« *Feature selector* »):



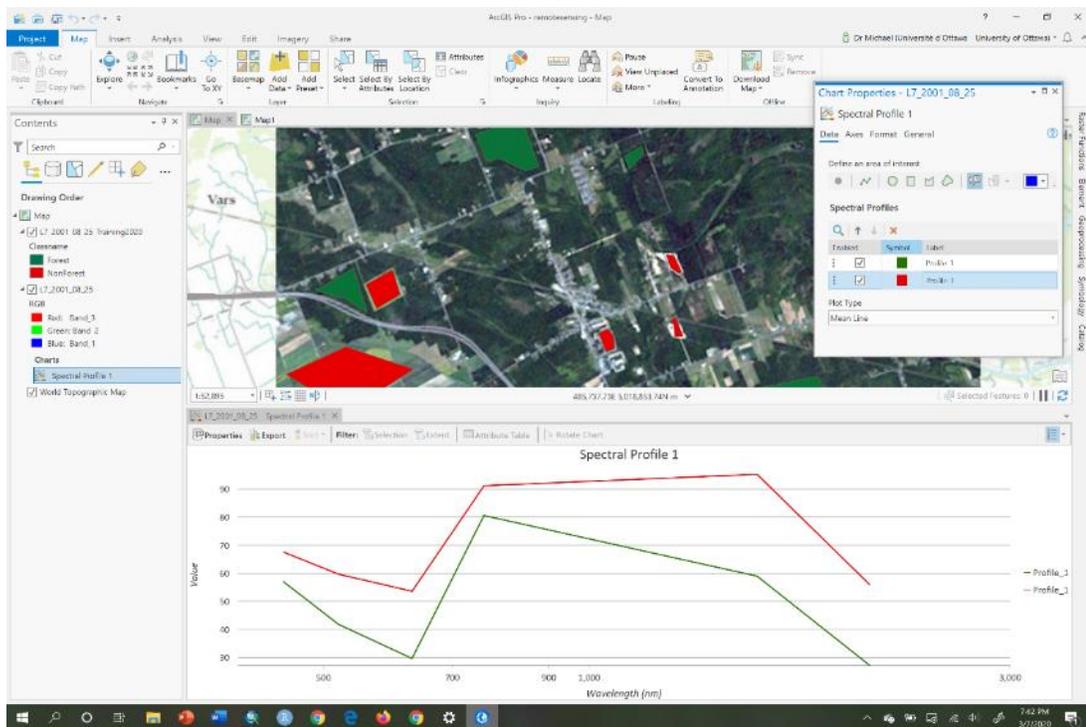
21. Cliquez maintenant sur l'un des polygones de **forêt** dans le panneau de la carte. Vous verrez un graphique créé avec une seule ligne appelée *Profile_1* qui représente la moyenne des VN sur toutes les bandes pour cette classe d'information appelée **Forêt**.



22. Maintenant, dans le volet Propriétés du graphique (« *Chart Properties* »), changez le label de cette ligne de moyenne spectrale de "Profil_1" à **Forêt** (si cela fonctionne!) et changez la couleur en vert et vous verrez votre ligne devenir verte:



23. Répétez maintenant les étapes 21-22 mais pour les polygones d'entraînement **NonForest** et rendez la nouvelle ligne rouge pour correspondre à **Non-Forêt** et renommez le label si vous le pouvez dans votre version d'ArcGIS Pro:

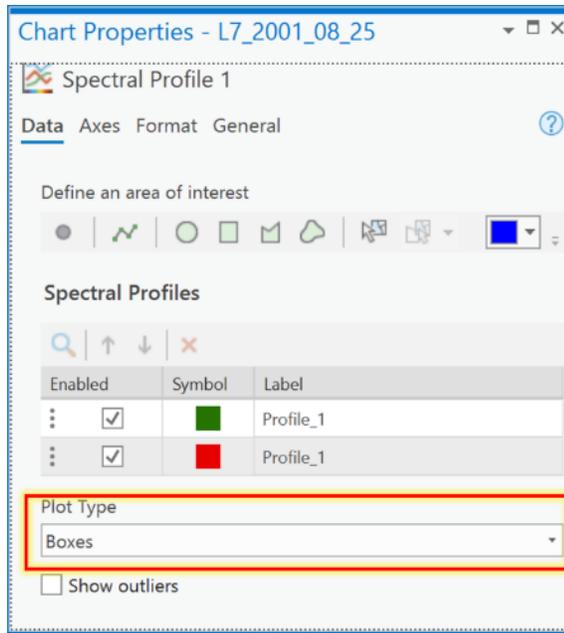


24. Examinez maintenant le graphique *Spectral Profile 1*:



Ce graphique montre la longueur d'onde sur l'axe des x plutôt que les noms des bandes parce que les métadonnées de l'image contiennent des informations sur la longueur d'onde de chaque bande. Vous pouvez déduire les bandes en observant où les lignes ont leurs "cassures". Par exemple, Band_1 se trouve à l'extrême gauche, Band_2 juste après 500 nm et Band_3 juste avant 700 nm, etc. Nous constatons que, dans chaque bande, il existe une différence entre les valeurs moyennes de chaque classe d'information. C'est peut-être dans la bande 4, juste après 700 nm, que la séparabilité est la plus faible, car c'est là que les lignes sont les plus proches. Nous pouvons vérifier cela à l'aide d'un diagramme statistique plus détaillé appelé diagramme en « boîtes et moustaches ».

25. Pour créer le graphique en boîte et en moustache de la séparabilité spectrale pour chaque bande, il suffit de modifier le type de graphique dans le volet Propriétés du graphique (« *Chart Properties* »). Dans le volet, changez le type de tracé (« *Plot Type* ») de moyenne (« *Mean* ») à boîtes (« *Boxes* »):

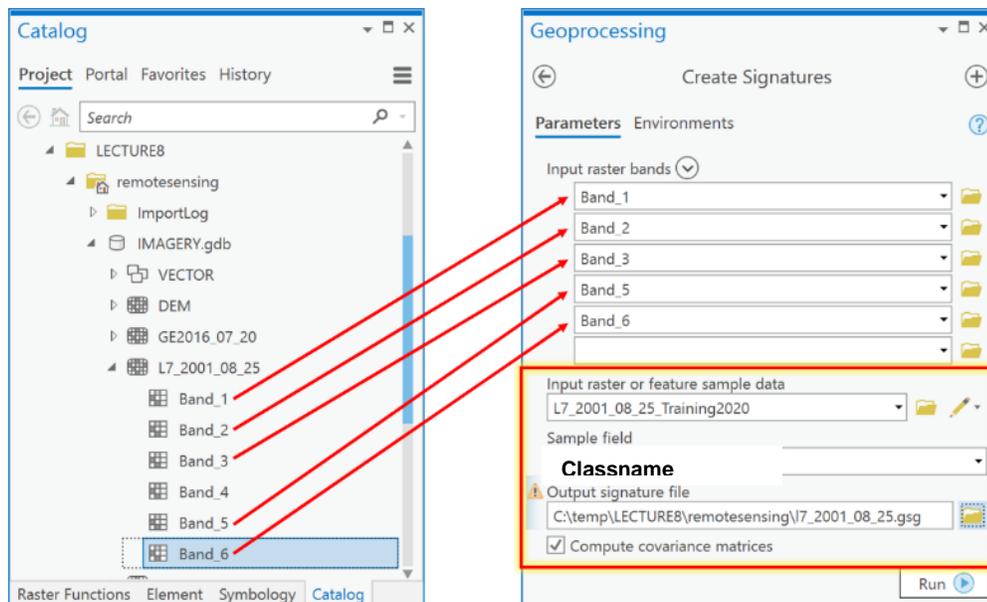


26. Examinez maintenant le diagramme en « boîtes et moustaches ».

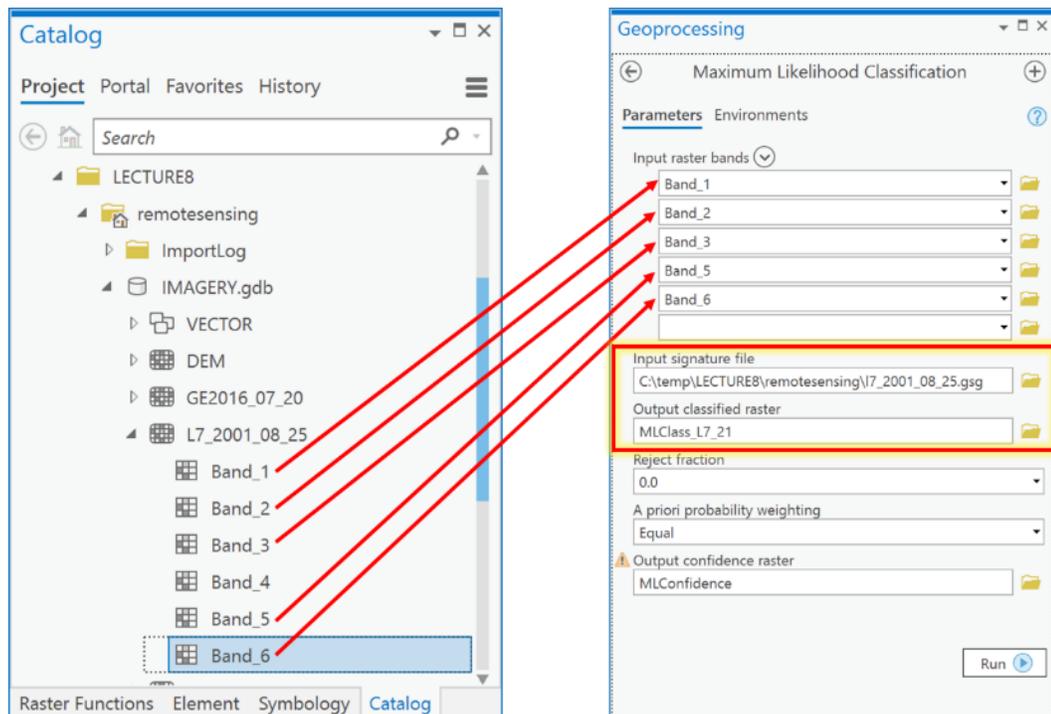


Dans ce diagramme, il y a en fait un diagramme en boîte et moustache pour chaque bande, chacun montrant les VN **forestiers** (vert) et **non forestiers** (rouge). Le haut et le bas des boîtes colorées sont les quartiles supérieurs et inférieurs de l'histogramme des VN pour la classe d'information. Les moustaches sont les deux lignes à l'extérieur de la boîte qui s'étendent jusqu'aux VN minimum et maximum pour la classe d'information. Il s'agit donc d'un résumé visuel des histogrammes pour chaque classe d'information, dans lequel nous pouvons maintenant examiner quelles bandes présentent la meilleure séparabilité. Il y a beaucoup plus d'informations ici que dans le graphique du profil spectral. Ce qui importe le plus ici, c'est que les boîtes ne se chevauchent pas, et ce pour toutes les bandes à l'exception de la Bande_4, qui est celle qui était également la plus proche dans le graphique du profil spectral. Nous pouvons donc envisager d'exclure la bande_4 de toute classification des **forêts** et des zones **non forestières**, mais nous devrions probablement conserver toutes les autres bandes.

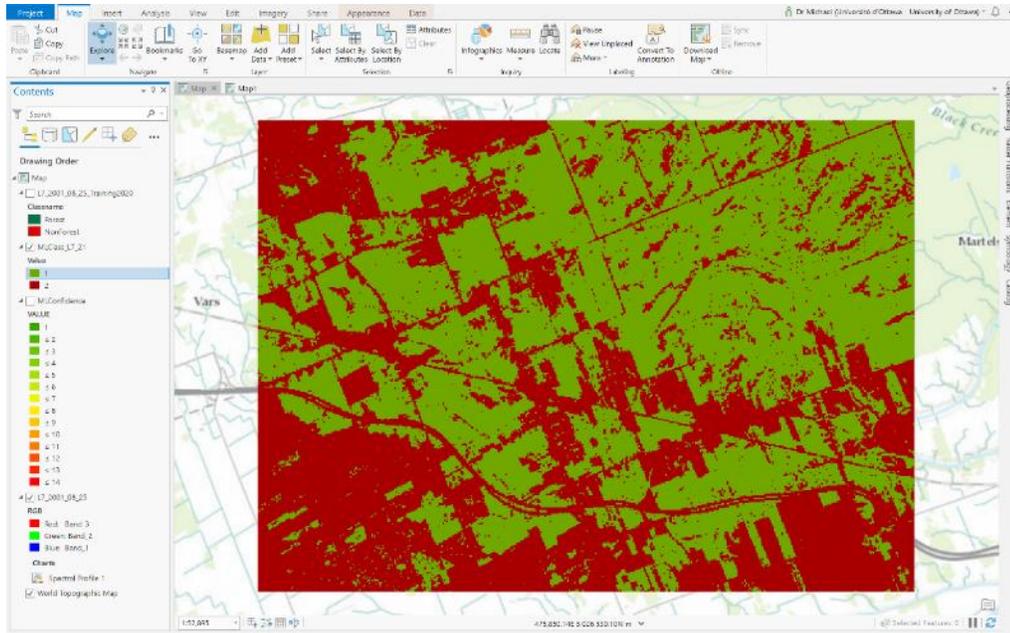
27. Créez un fichier de signatures à partir des sites d'entraînement que nous pourrions utiliser pour la classification. Pour créer le fichier, allez dans l'onglet Analyse (« *Analysis* »), cliquez sur Outils (« *Tools* ») et recherchez l'outil Créer des signatures (« *Create Signatures* ») qui se trouve dans la boîte à outils *Spatial Analysis*. Désactivez l'outil *Create Signatures* afin de pouvoir le voir en même temps que votre volet Catalogue. Ensuite, dans le volet *Catalog*, naviguez jusqu'à IMAGERY.GDB et agrandissez l'image "L7_2001_08_25" de manière à voir toutes les bandes individuelles (Band_1, Band_2,...,Band_6). Faites glisser toutes les bandes à l'exception de la Bande_4 (puisque nous savons qu'elle a une mauvaise séparabilité) dans "Input raster bands", sous "Input raster or feature sample data" choisissez la classe objet "L7_2001_08_25_Training2023", sous "Sample field" choisissez **Classname**, et sauvegardez le fichier dans le dossier où vous avez vos données, et appelez-le "l7_2001_08_25.gsg". Cliquez sur "Run" (Exécuter).



28. Utilisez maintenant le fichier de signatures pour classer l'image dans les deux classes d'information Forêt et Non-Forêt. Pour cela, recherchez maintenant l'outil d'analyse spatiale "Maximum Likelihood Classification" (Classification par maximum de vraisemblance). Désactivez encore l'outil afin de pouvoir le voir en même temps que votre volet Catalogue. Ensuite, dans le volet Catalogue, naviguez jusqu'à IMAGERY.GDB et agrandissez l'image "L7_2001_08_25" de manière à voir toutes les bandes individuelles (Bande_1, Bande_2, ..., Bande_6). Faites glisser toutes les bandes à l'exception de la Bande_4 dans "Input raster bands", sous "Input signature file", choisissez le fichier de signature créé à l'étape précédente appelé "I7_2001_08_25.gsg", sous "Output classified raster" appelez-le "MLCLASS_L7_21", et enregistrez-le dans IMAGERY.GDB. Laissez toutes les autres valeurs par défaut et cliquez sur *Run*.



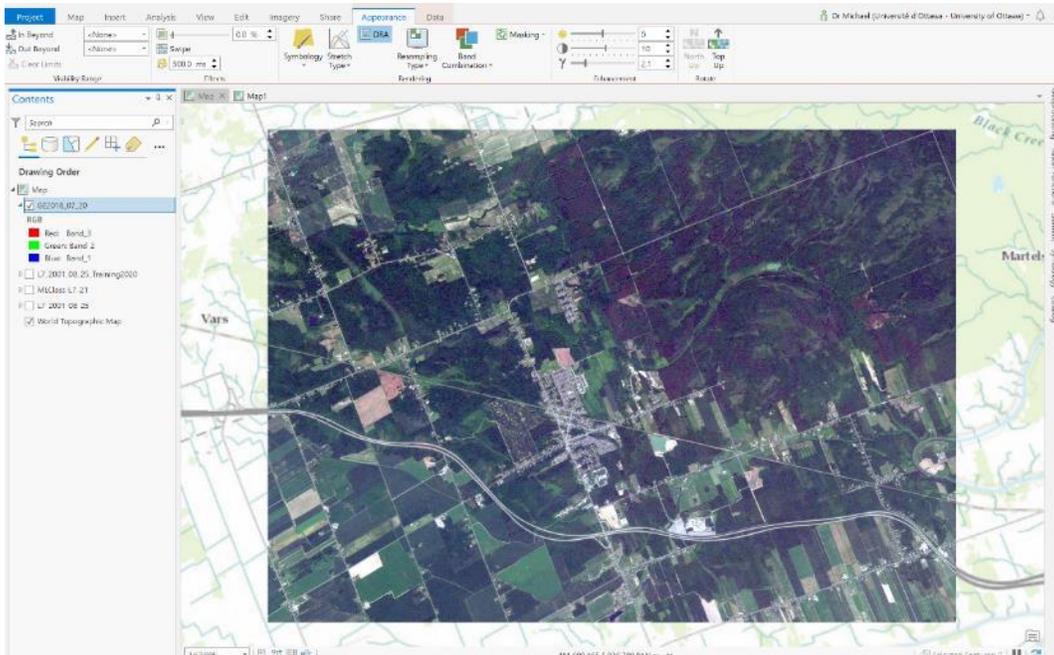
29. Examinez le raster classifié résultant (notez que vous pouvez changer la valeur 1 - Forêt en vert et la valeur 2 - Non-Forêt en rouge). Nous avons maintenant converti une image de télédétection en une couche raster thématique.



Q3: Comment créer un ensemble de données d'entraînement correspondant à mes classes d'information?

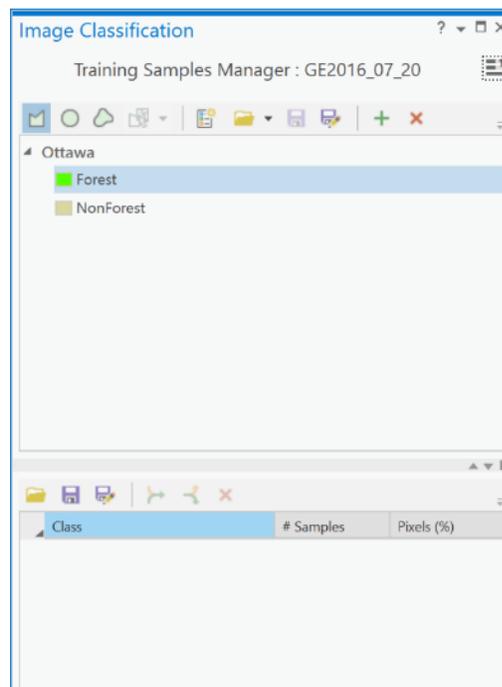
Les données d'entraînement sont constituées de pixels appartenant à différentes classes d'information. Par exemple, certains pixels appartiennent à la classe appelée **Forêt**, et un autre ensemble de pixels à la classe appelée **Non-Forêt**.

1. Si elle n'est pas déjà présente, ajoutez l'image "RE2016_07_20" à ArcGIS à partir de IMAGERY.GDB.

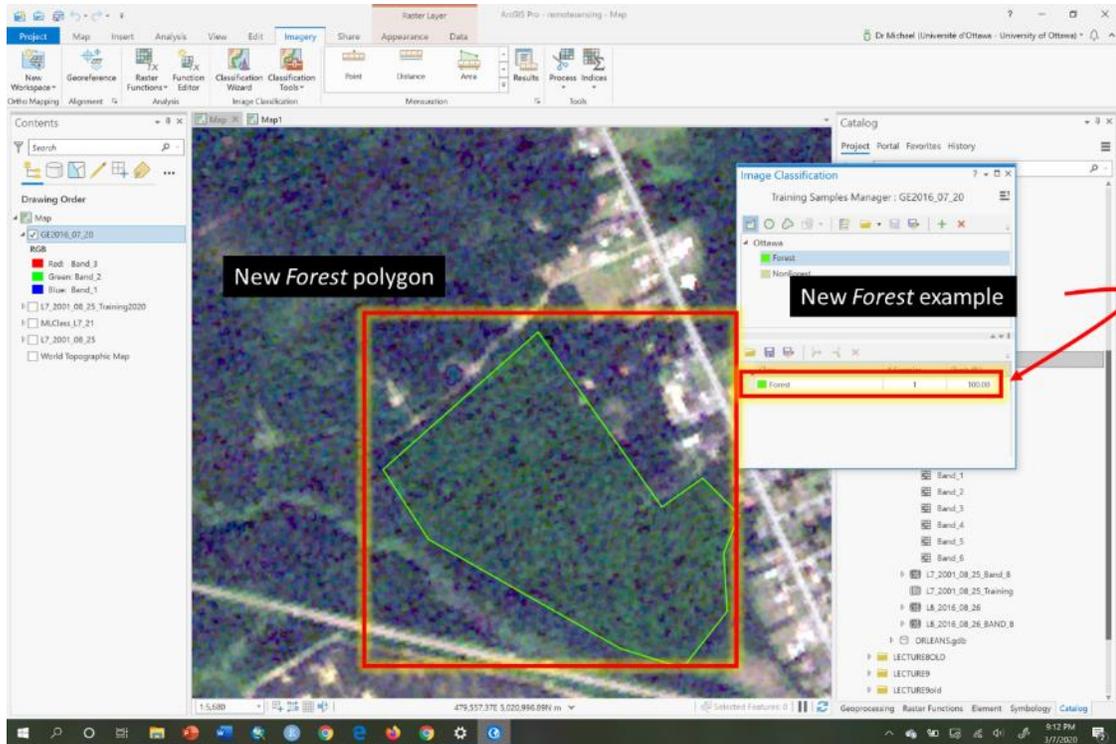


Nom de l'étudiant: _____ # d'étudiant: _____

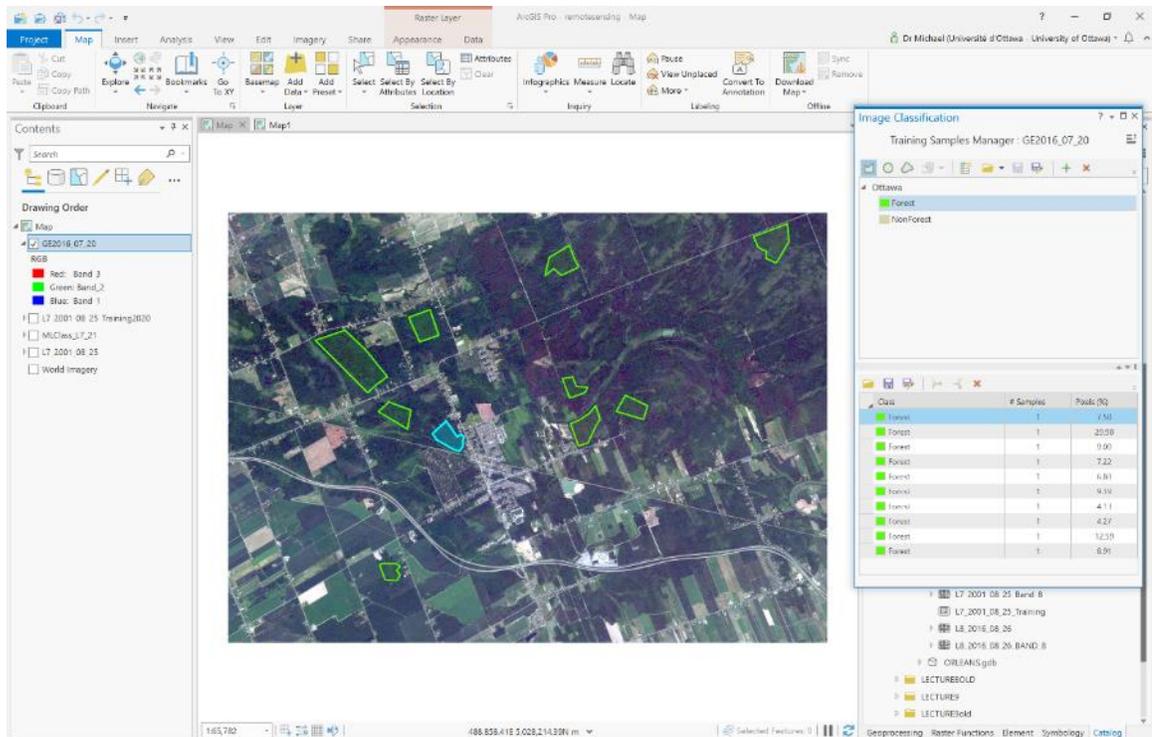
2. Cliquez sur l'onglet Imagerie (« *Imagery* »), puis sur le bouton Outils de classification (« *Classification Tools* ») et choisissez le gestionnaire d'échantillons d'entraînement (« *Training Samples Manager* »).
3. En utilisant le gestionnaire d'échantillons d'entraînement, ouvrez le schéma Ottawa.ecs, comme vous l'avez fait précédemment.
4. Nous allons maintenant créer un nouvel ensemble de polygones d'entraînement pour les classes d'information Forêt et Non-Forêt, à utiliser avec l'image RapidEye. Notez que nous ne pouvons pas simplement réutiliser les polygones que nous avons utilisés pour l'image Landsat, car les zones à l'intérieur de ces polygones peuvent avoir changé de Forêt à Non-Forêt, ou vice versa, dans la période entre 2001 (lorsque l'image Landsat a été acquise) et 2016 (lorsque l'image RapidEye a été acquise). À cette fin, dans le Gestionnaire d'échantillons d'entraînement, dans le demi-volet du haut, sous le schéma appelé Ottawa, cliquez sur la classe d'information Forêt. Vous verrez que l'outil polygone de la barre d'outils devient actif.



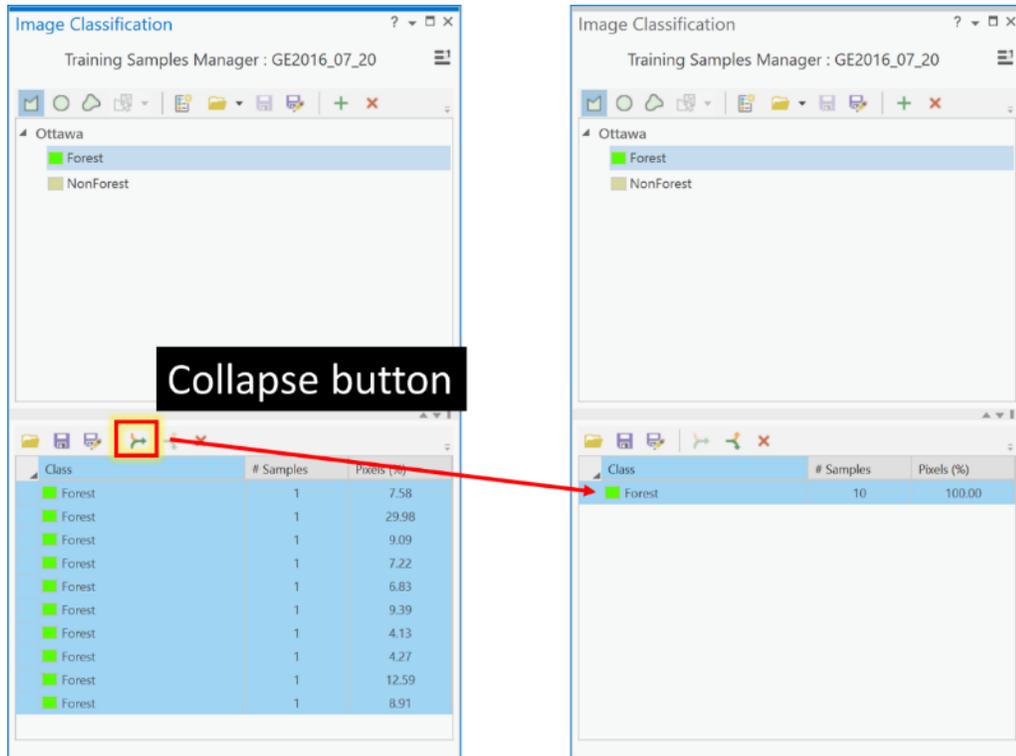
5. Cliquez maintenant sur l'outil polygone et dessinez un polygone dans n'importe quelle zone boisée de l'image satellite.



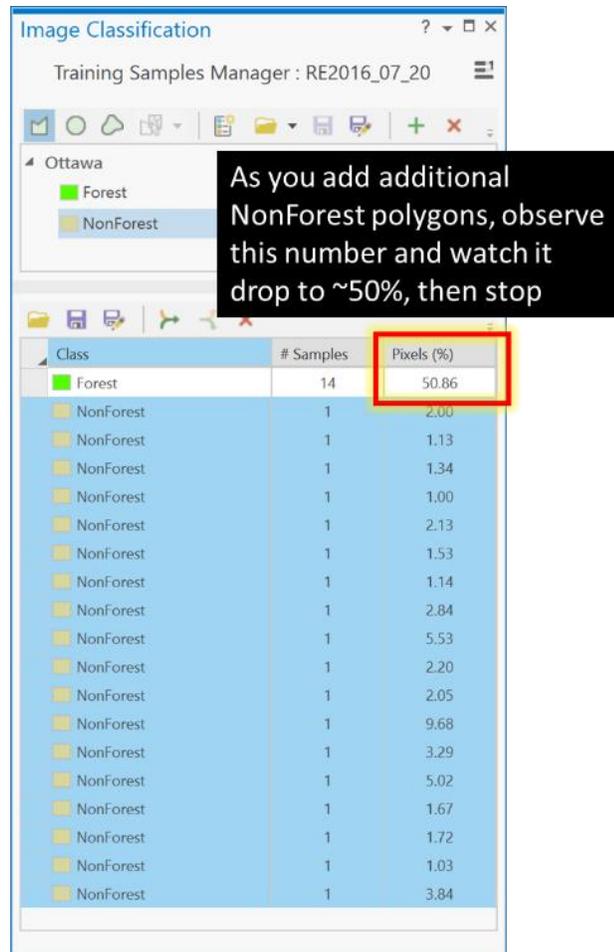
6. Répétez l'étape précédente neuf fois dans différentes zones forestières. Vous obtiendrez ainsi quelque chose qui ressemble à ceci:



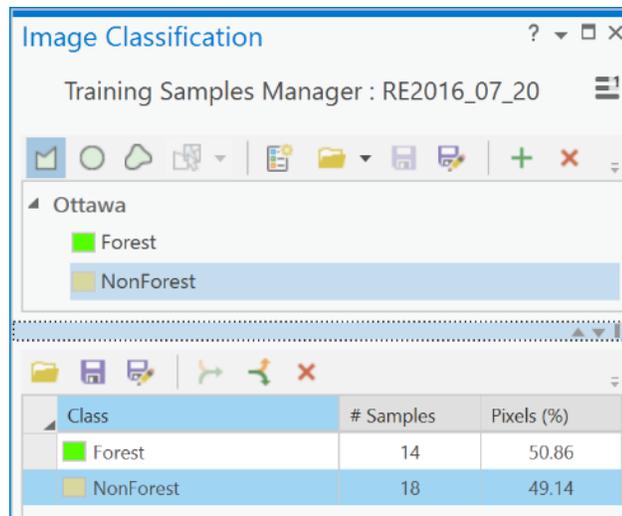
- Examinez le volet Gestionnaire d'échantillons d'entraînement, le demi-volet du bas, et vous verrez maintenant les 10+ polygones de **forêt** énumérés. Tenez la touche *Shift* (Maj) enfoncée et cliquez sur le premier et le dernier exemple de polygone forestier pour sélectionner tous les exemples de **forêt**. Cliquez ensuite sur le bouton Réduire (« *Collapse* ») pour fusionner les exemples de **forêt** en un seul groupe:



- Ensuite, dans le panneau Gestionnaire d'échantillons d'entraînement, cliquez sur **Non-Forest** et répétez les étapes 5 à 7 pour la classe d'information **Non-Forest**. Au final, vous devriez obtenir un panneau Gestionnaire d'échantillons d'entraînement qui ressemble à l'illustration ci-dessous, avec un pourcentage presque égal de pixels dans chaque classe d'information. Pour vous assurer que vous créez des classes égales, le nombre de polygones n'est pas ce que vous surveillez au fur et à mesure que vous ajoutez des exemples supplémentaires de **Non-Forest**, mais vous observez le % de pixels de la classe **Forest** et essayez de l'amener à environ 50 %. Fusionnez ensuite tous les polygones non forestiers à ce stade:



9. Le volet Gestionnaire d'échantillons d'entraînement ressemble maintenant à ceci:

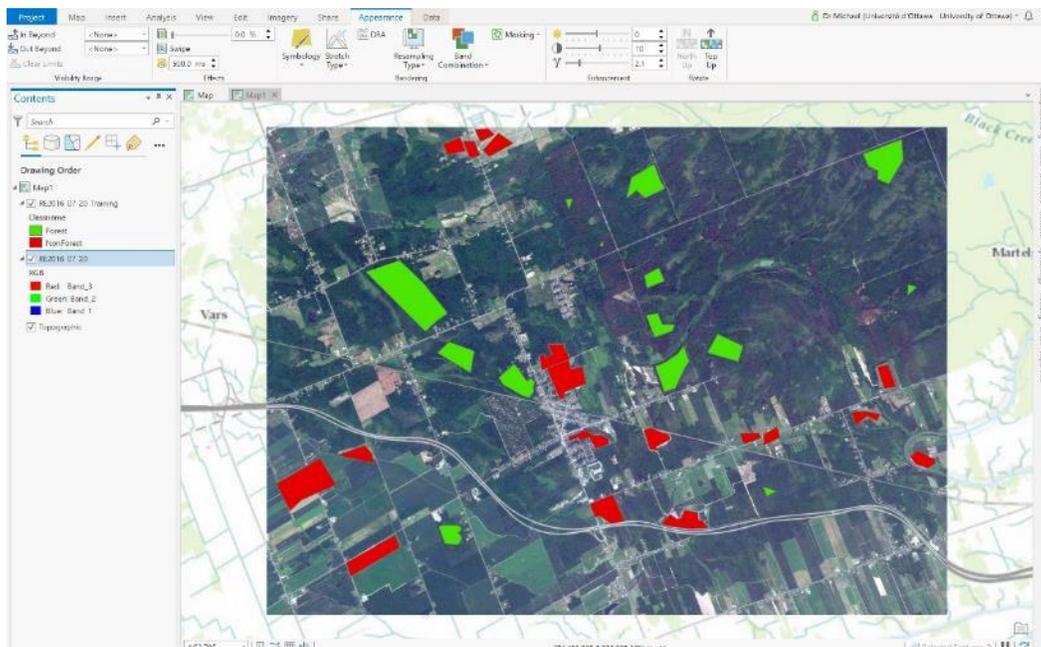


10. Ensuite, nous devons enregistrer nos nouveaux polygones d'entraînement dans une classe d'objet. Pour cela, cliquez sur le bouton "Enregistrer" dans le volet du bas du gestionnaire d'échantillons d'entraînement, naviguez jusqu'à IMAGERY.GDB et enregistrez le nouveau fichier d'entraînement sous le nom "RE2016_07_20_training".

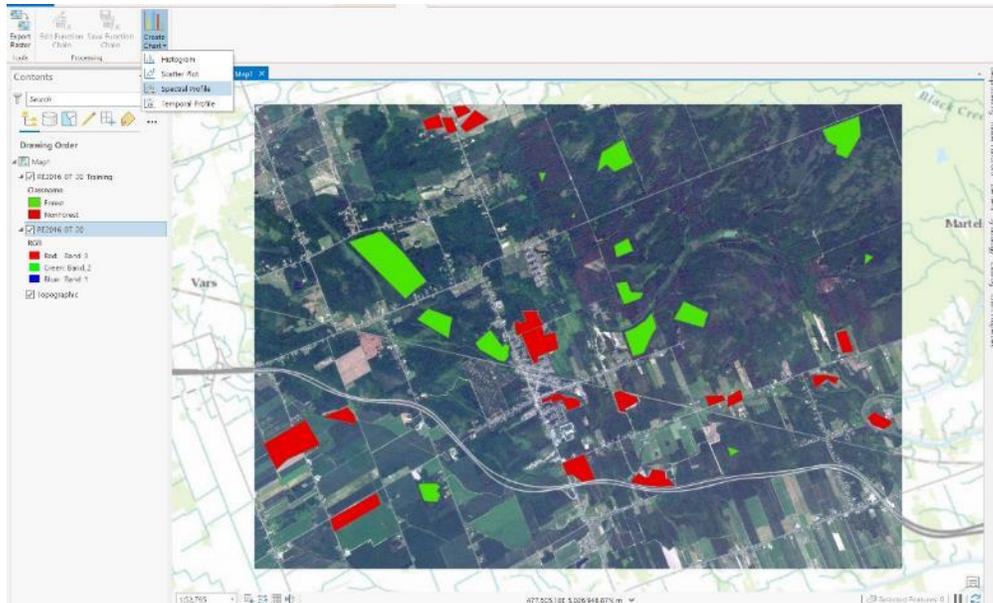
Q4: Comment créer un fichier de signature à partir de mes polygones d'entraînement?

Pour créer un fichier de signature, vous avez besoin d'un ensemble de données de polygones d'entraînement et d'une image multibande que vous souhaitez utiliser pour extraire les informations des pixels de chacune des classes d'informations d'entraînement. Vous devez ensuite évaluer à nouveau la séparabilité avant de créer le fichier de signature, car vous utilisez une nouvelle image.

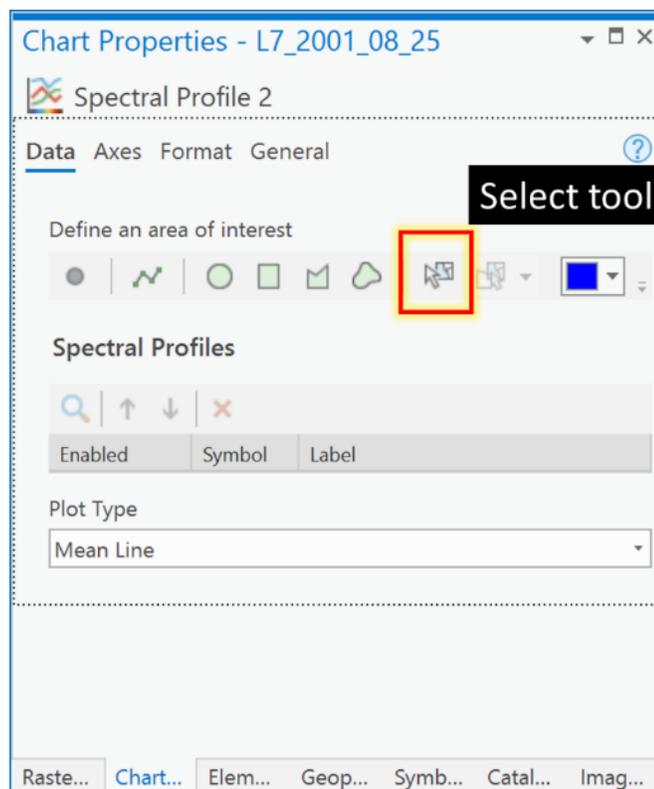
1. Ajoutez la classe d'éléments "RE2016_07_20_training" de la Q3 ci-dessus au volet Contenu et symbolisez-la à l'aide de valeurs uniques et du champ **Classname**, et donnez à la classe "**Forest**" la couleur verte et à la classe "**NonForest**" la couleur rouge.



2. Ensuite, choisissez la couche "RE2016_07_20" dans le volet de contenu. Cliquez ensuite sur l'onglet Données (« Data »). Cliquez maintenant sur Créer un graphique (« Create Chart ») et choisissez Profil spectral (« Spectral Profile »):

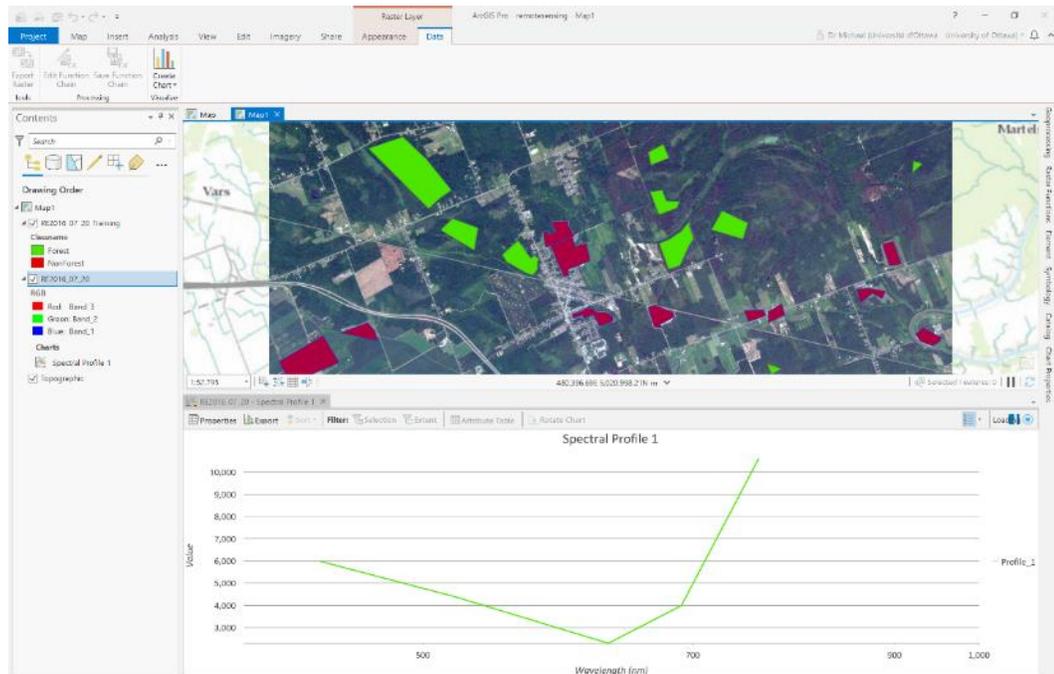


3. Ensuite, dans le volet Propriétés du graphique (« *Chart Properties* »), cliquez sur le bouton Sélecteur d'objet (« *Feature Selector* »):

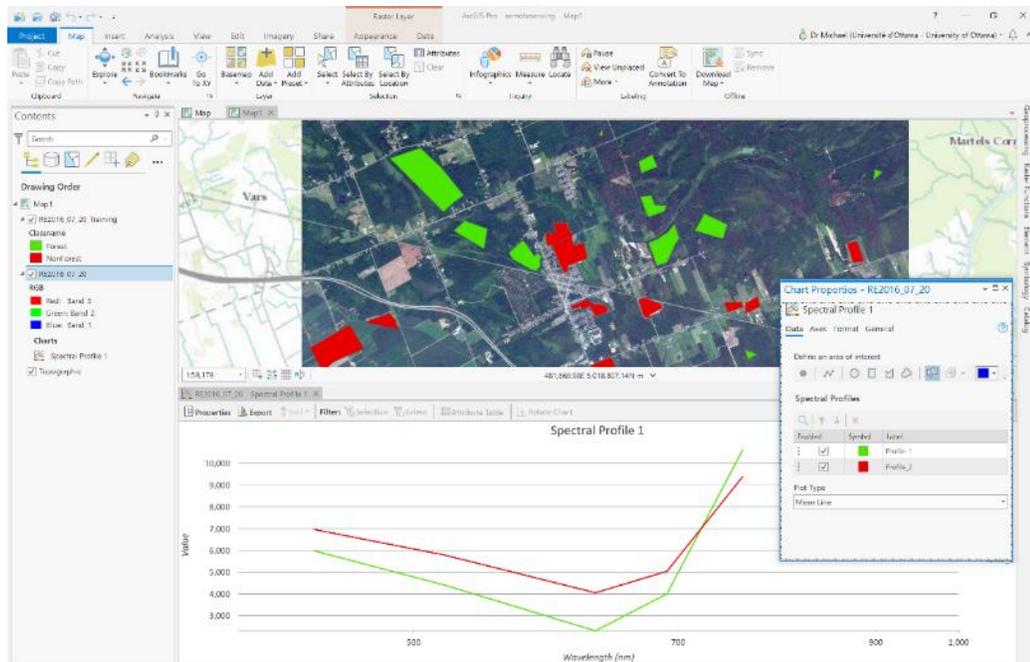


Nom de l'étudiant: _____ # d'étudiant: _____

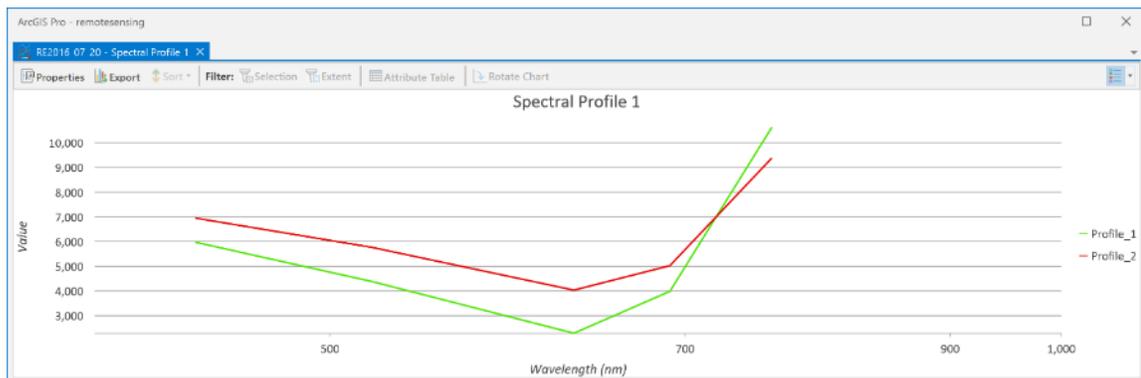
4. Cliquez maintenant sur l'un des polygones de forêt dans le volet de la carte. Vous verrez un graphique créé avec une seule ligne appelée "Profile_1" qui représente la moyenne des VN sur toutes les bandes pour cette classe d'information appelée **Forest**. Dans le volet Propriétés du graphique, changez l'étiquette de cette ligne de moyenne spectrale de "Profil_1" à **Forêt** (si cela fonctionne !) et changez la couleur en vert et vous verrez votre ligne devenir verte:



5. Répétez les étapes 3-4 pour les polygones d'entraînement **NonForest**, et faites en sorte que la nouvelle ligne soit rouge pour correspondre à **NonForest**. Renommez le label si vous le pouvez dans votre version d'ArcGIS Pro:

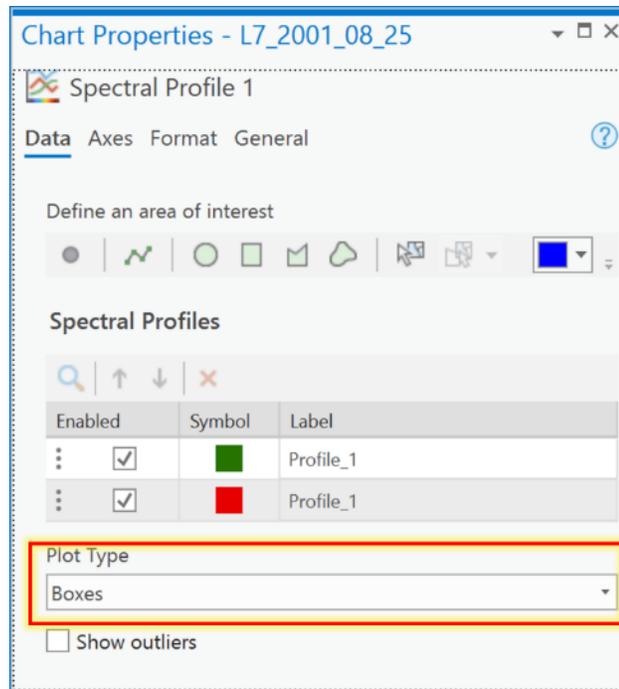


6. Examinez maintenant le graphique du profil spectral:



Tout comme dans Q2, ce graphique montre la longueur d'onde sur l'axe des x plutôt que les noms des bandes, parce que les métadonnées de l'image contiennent des informations sur la longueur d'onde de chaque bande. Les VN moyens montrent des différences constantes, mais il est difficile de juger de la séparabilité à partir des seules moyennes. Nous devons donc examiner les diagrammes en « boîtes et moustaches » pour mieux choisir les bandes dont les boîtes ne se chevauchent pas et qui présentent donc une bonne séparabilité.

7. Pour créer le graphique en boîte et en moustache de la séparabilité spectrale pour chaque bande, il suffit de modifier le type de graphique dans le volet Propriétés du graphique (« *Chart Properties* »). Dans le volet, changez le type de tracé (« *Plot Type* ») de moyenne (« *Mean* ») à boîtes (« *Boxes* »):

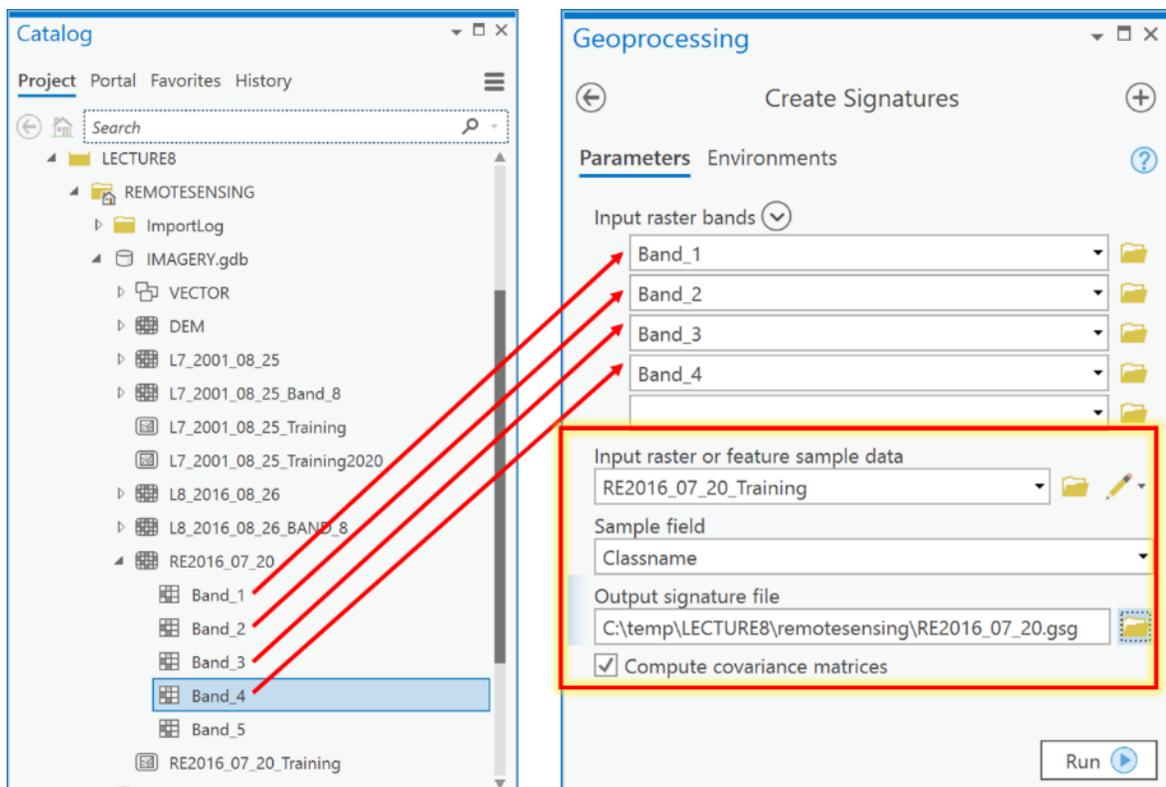


8. Examinez maintenant le diagramme en « boîtes et moustaches ».



La plupart des bandes présentent une bonne séparabilité, à l'exception de la Bande_5 où les cases se chevauchent. La largeur de bande spectrale de la Bande_5 sur ce capteur RapidEye est de 760-850 nm, ce qui correspond à la partie proche de l'infrarouge du spectre. C'est dans cette même partie du spectre que l'image "L7_2001_08_25" a montré une faible séparabilité entre forêt et non forêt (dans la bande 4, qui couvre 750-900 nm pour le capteur Landsat). Ainsi, pour notre image actuelle, la bande 5 sera exclue de toute classification des **forêts** et des zones **non forestières**.

- Créez un fichier de signatures à partir des sites d'entraînement que nous pourrions utiliser pour la classification. Pour créer le fichier, allez dans l'onglet Analyse (« *Analysis* »), cliquez sur Outils (« *Tools* ») et recherchez l'outil Créer des signatures (« *Create Signatures* ») qui se trouve dans la boîte à outils *Spatial Analysis*. Désactivez l'outil *Create Signatures* afin de pouvoir le voir en même temps que votre volet Catalogue. Ensuite, dans le volet *Catalog*, naviguez jusqu'à IMAGERY.GDB et agrandissez l'image " RE2016_07_20" de manière à voir toutes les bandes individuelles (Band_1, Band_2,...,Band_5). Faites glisser toutes les bandes à l'exception de la Bande_5 (puisque nous savons qu'elle a une mauvaise séparabilité) dans "Input raster bands", sous "Input raster or feature sample data" choisissez la classe objet " RE2016_07_20_training ", sous "Sample field" choisissez **Classname**, et sauvegardez le fichier dans le dossier où vous avez vos données, et appelez-le " RE2016_07_20.gsg". Cliquez sur "Run" (Exécuter).



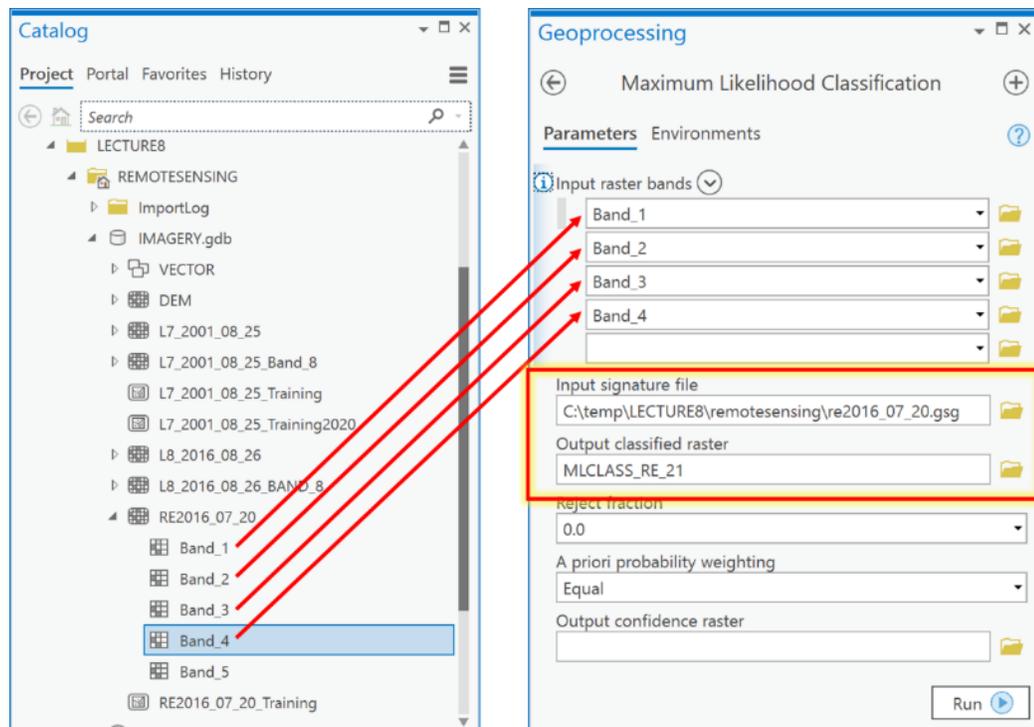
- Vous avez maintenant un fichier de signatures que vous pouvez utiliser pour la classification des images.

Q5: Comment puis-je classer mon image en utilisant la classification d'image supervisée?

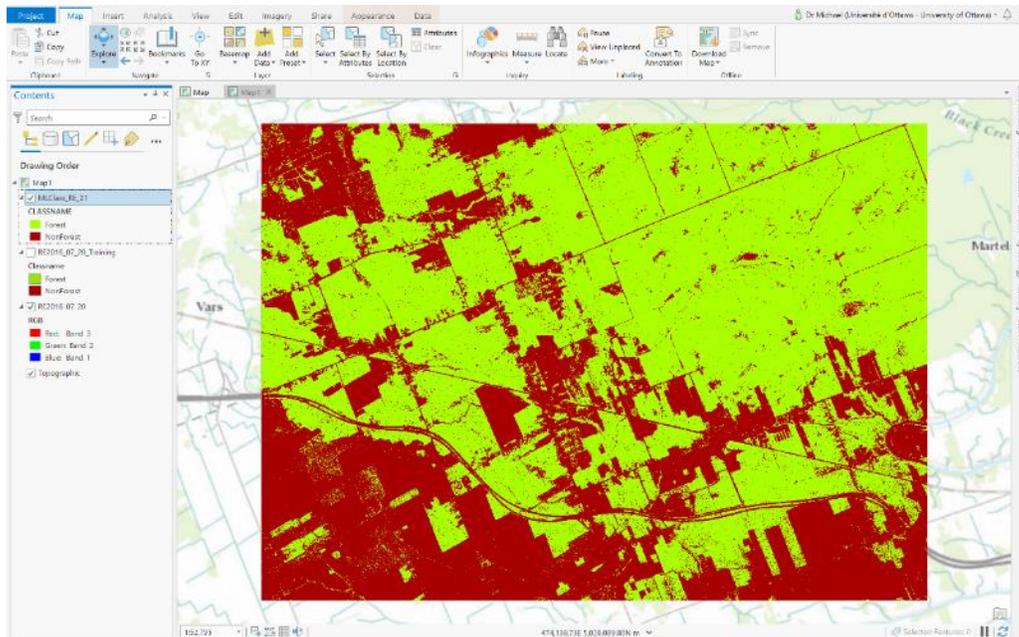
Pour classer une image, vous devez avoir répondu aux questions ci-dessus car vous avez besoin du fichier de signature "RE2016_07_20.gsg" pour l'image RapidEye "RE2016_07_20" afin de procéder à une classification supervisée.

Maintenant que vous avez obtenu un fichier de signatures à partir de l'exemple précédent, vous pouvez procéder à une classification supervisée:

1. Utilisez maintenant le fichier de signatures pour classer l'image dans les deux classes d'information **Forêt** et **Non-Forêt**. Pour cela, recherchez maintenant l'outil d'analyse spatiale "Maximum Likelihood Classification" (Classification par maximum de vraisemblance). Désactivez encore l'outil afin de pouvoir le voir en même temps que votre volet Catalogue.
2. Ensuite, dans le volet Catalogue, naviguez jusqu'à IMAGERY.GDB et agrandissez l'image "RE2016_07_20" de manière à voir toutes les bandes individuelles (Bande_1, Bande_2, ..., Bande_5). Faites glisser toutes les bandes à l'exception de la Bande_5 dans "Input raster bands", sous "Input signature file", choisissez le fichier de signature créé durant l'exemple précédente appelé "RE2016_07_20.gsg", sous "Output classified raster" appelez-le "MLCLASS_RE_21", et enregistrez-le dans IMAGERY.GDB. Laissez toutes les autres valeurs par défaut et cliquez sur *Run* (Exécuter):

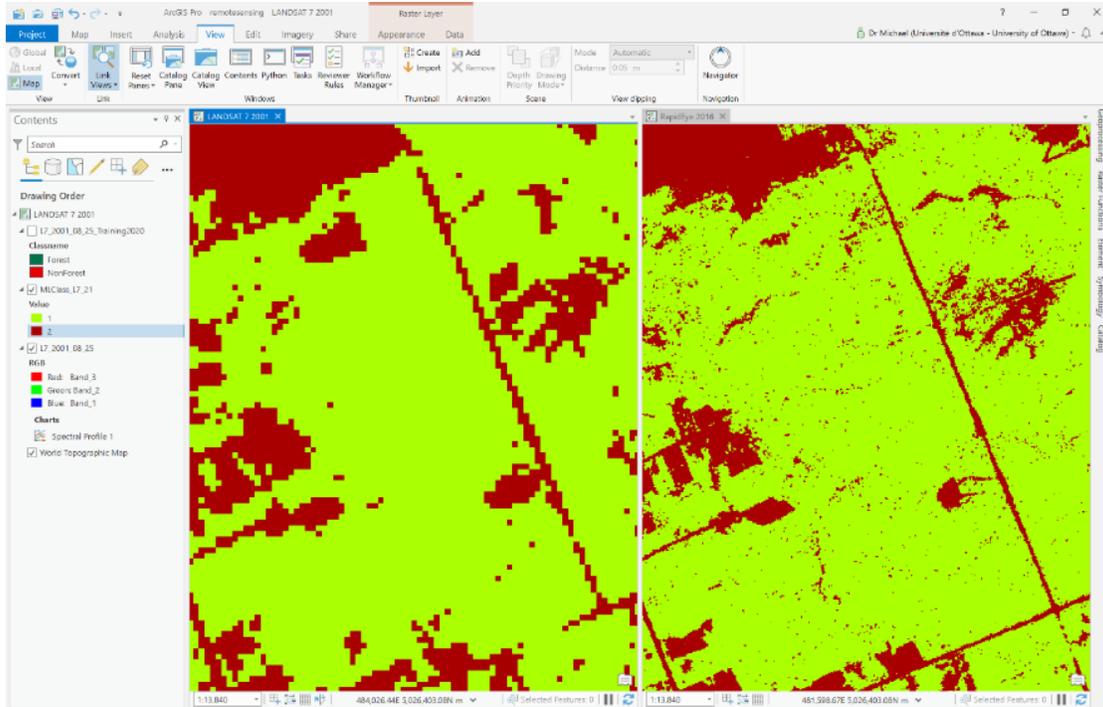


3. Examinez la couche raster classée résultante (notez que vous pouvez changer la valeur 1 - **Forest** en vert et la valeur 2 - **NonForest** en rouge). Nous avons maintenant converti l'image en une couche raster thématique.

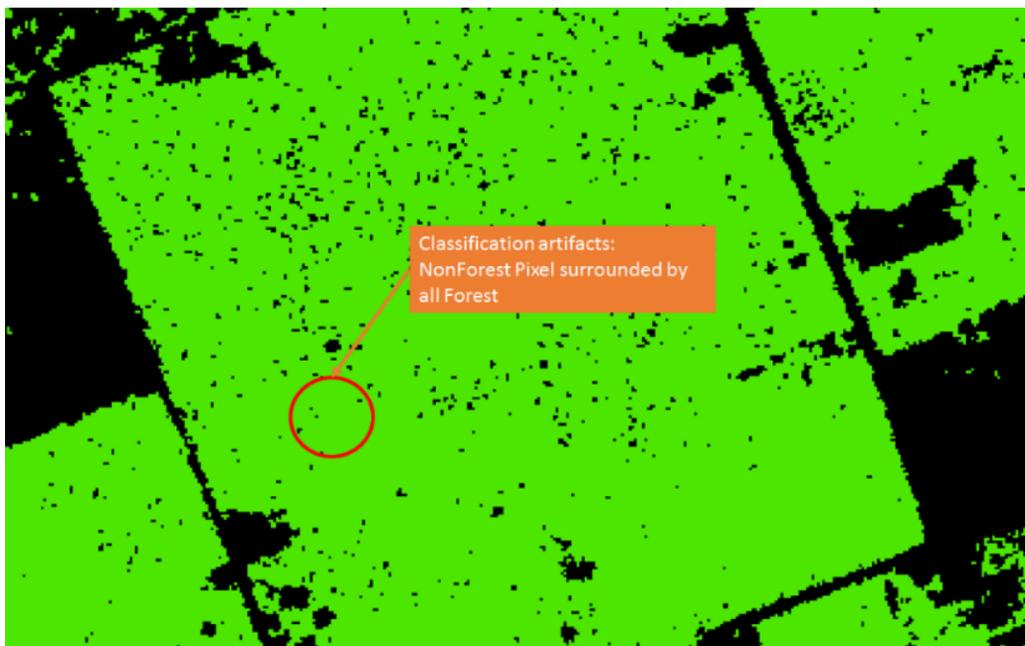


Q6: Comment puis-je nettoyer ma couche raster classifiée?

Dans le résultat classifié, vous remarquerez de nombreux artefacts du processus de classification; ceux-ci sont particulièrement évidents lorsque des pixels de **forêt** sont entourés de pixels **non forestiers** ou vice-versa. En règle générale, plus la résolution de l'image est élevée (plus les pixels sont petits), plus il y a de pixels mal classifiés. Étant donné que la couche Landsat 7 avait une résolution de 30 m et que la couche RapidEye avait une résolution de 5 m, il n'est pas surprenant qu'il y ait plus de pixels mal classifiés dans la carte réalisée à partir de l'image RapidEye. Par exemple, en regardant les résultats de classification de Landsat 7 et de RapidEye côte à côte (comme indiqué ci-dessous), nous pouvons voir que la **forêt** de Landsat 7 est plus verte et contient moins de pixels mal classifiés que celle de RapidEye au même endroit.

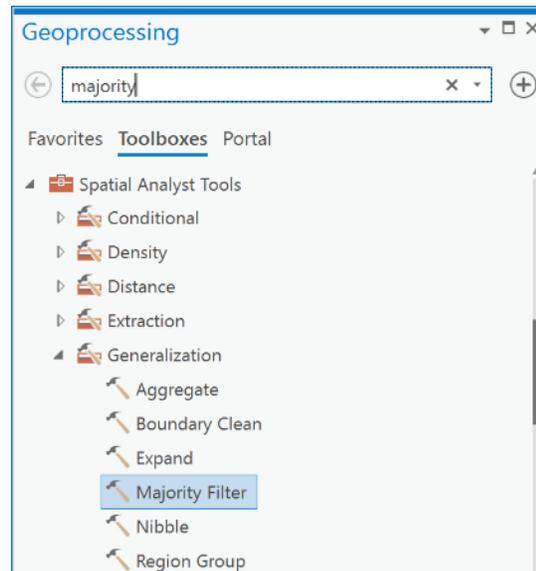


Il est peu probable que des pixels non forestiers entourés de forêts représentent une perte de forêt sur ce pixel dans la majorité des cas. Cela dépend en partie de la définition de ce qu'est une "forêt" ?!? Bien qu'il ne s'agisse pas d'une définition formelle, beaucoup considèrent qu'une forêt est une zone dominée par des arbres - mais cela ne signifie pas qu'il doit y avoir des arbres partout ! Les forêts peuvent avoir des clairières sans arbres, mais ces clairières font toujours partie de la forêt. Avec l'imagerie RapidEye, de telles clairières, ou des zones qui ne contiennent qu'un petit arbre, ou un arbre malsain, peuvent avoir été mal classifiées comme non forestières (voir les exemples dans la figure ci-dessous).

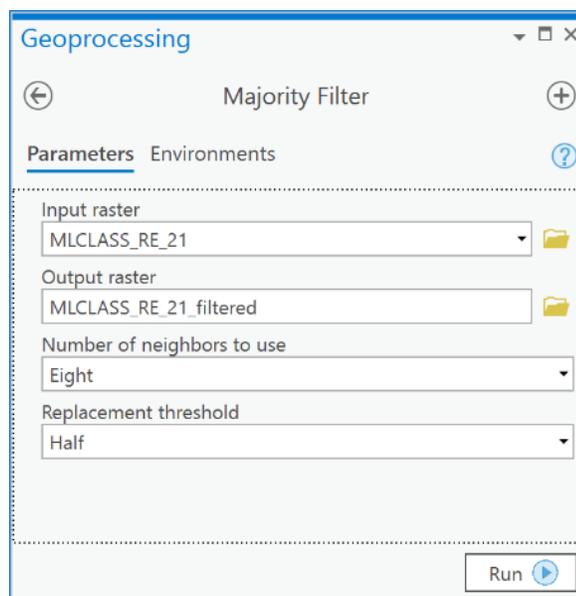


Nous pouvons filtrer certains des pixels erronés à l'aide d'un filtre majoritaire:

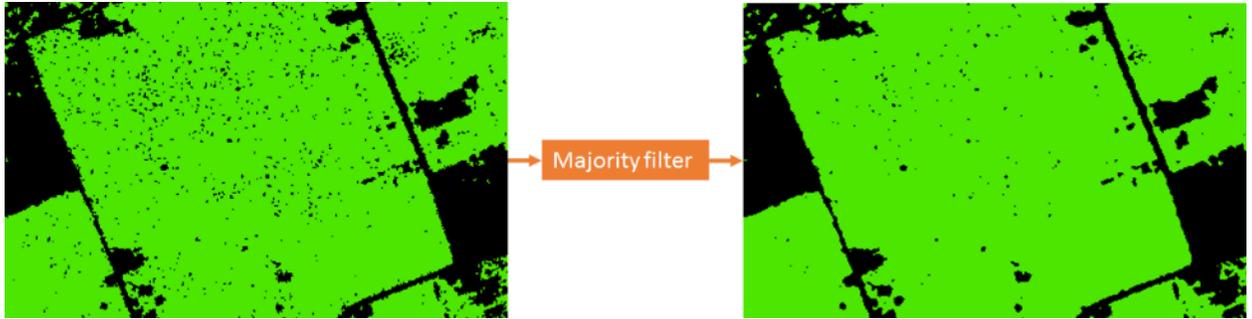
1. Ouvrez l'outil Filtre de majorité (« *Majority Filter* ») sous Outils (« *Tools* ») d'analyse spatiale -> Généralisation (« *Generalization* »):



2. Dans le menu déroulant « Input raster », sélectionnez votre couche raster classifiée "MLCLASS_RE_21". Appelez la sortie "MLCLASS_RE_21_filtered" et enregistrez-la dans IMAGERY.gdb. Choisissez le nombre de voisins à utiliser (optionnel) comme huit voisins, et cliquez sur *Run*.



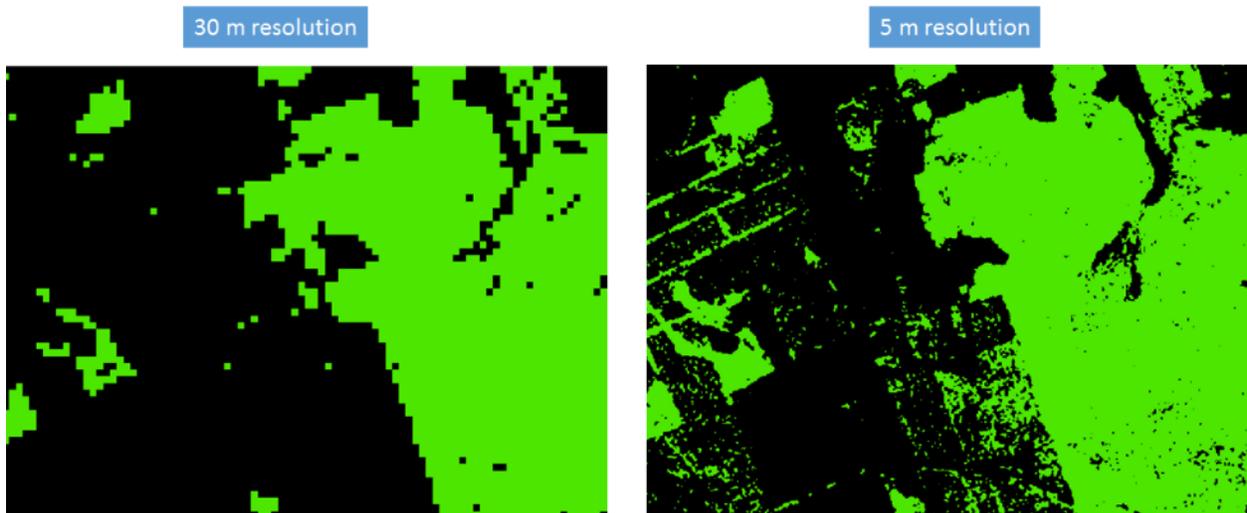
3. Le résultat est une couche classifiée plus propre avec moins de pixels mal classifiés/abhorrés:



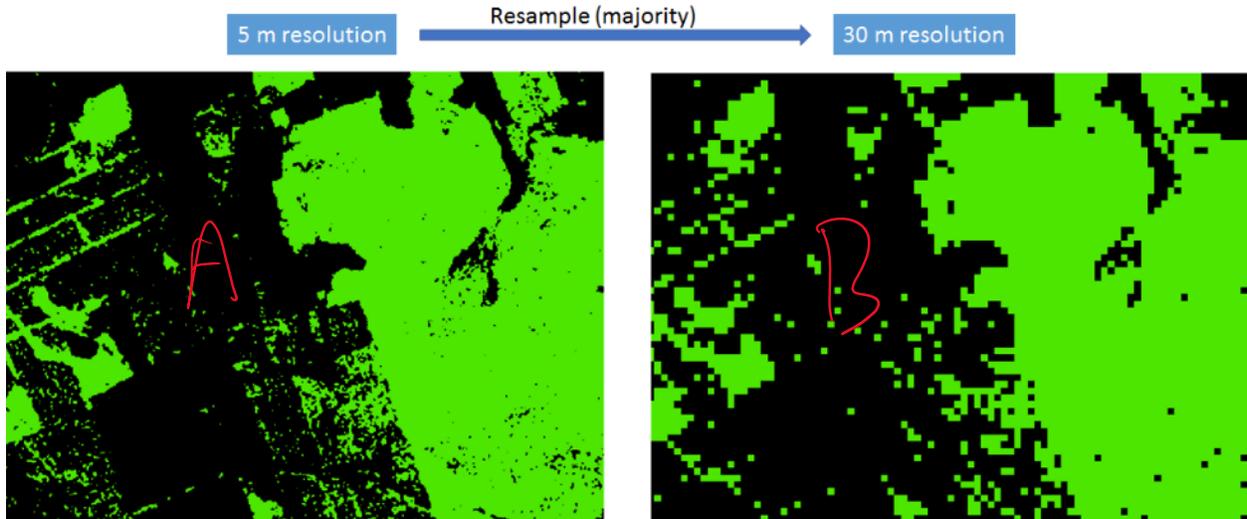
Il est parfois nécessaire d'exécuter plusieurs fois le filtre de la majorité (ou des outils similaires) pour obtenir le résultat au niveau de généralisation requis pour un objectif particulier.

Q7: Comment généraliser mon raster classifié "MLCLASS_RE_21_filtered" pour que la taille des cellules soit la même que celle de la couche "MLCLASS_L7_21" de Q2?

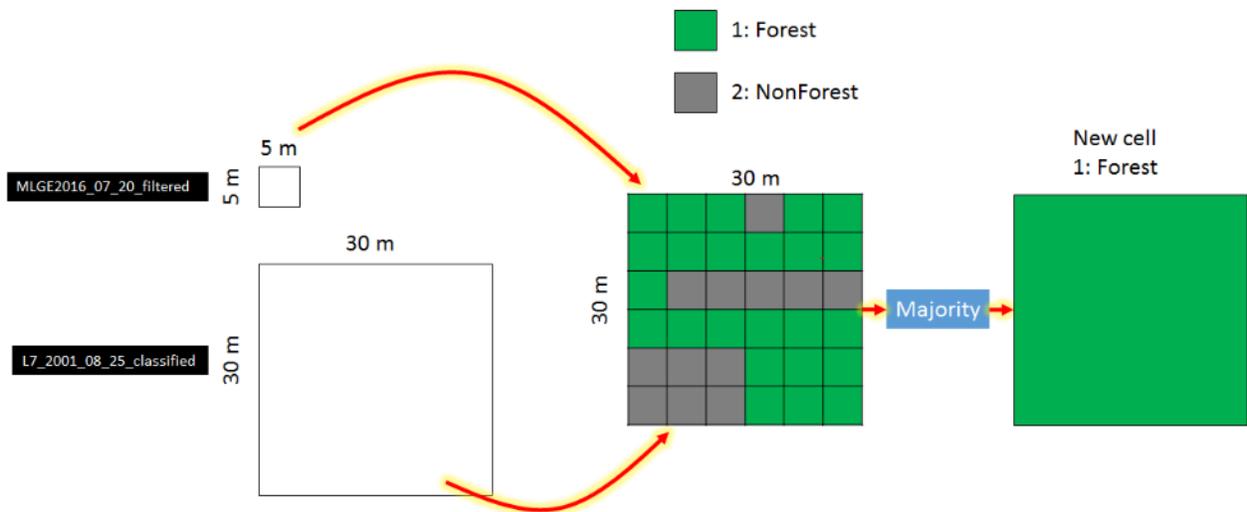
La scène Landsat 7 qui a été classifiée dans Q2 est dans la couche appelée "MLCLASS_L7_21" et a une résolution spatiale de 30 mètres (c'est-à-dire que chaque cellule est de 30 x 30 mètres). Votre couche "MLCLASS_RE_21_filtered" a la même résolution que l'image RapidEye qui est de 5 m (voir comparaison ci-dessous).



Comme il y a 36 cellules dans la couche "MLCLASS_RE_21_filtered" pour chaque cellule dans "MLCLASS_L7_21", vous augmenterez la taille des cellules de "MLCLASS_RE_21_filtered" pour qu'elle corresponde à la taille des cellules de "MLCLASS_L7_21".



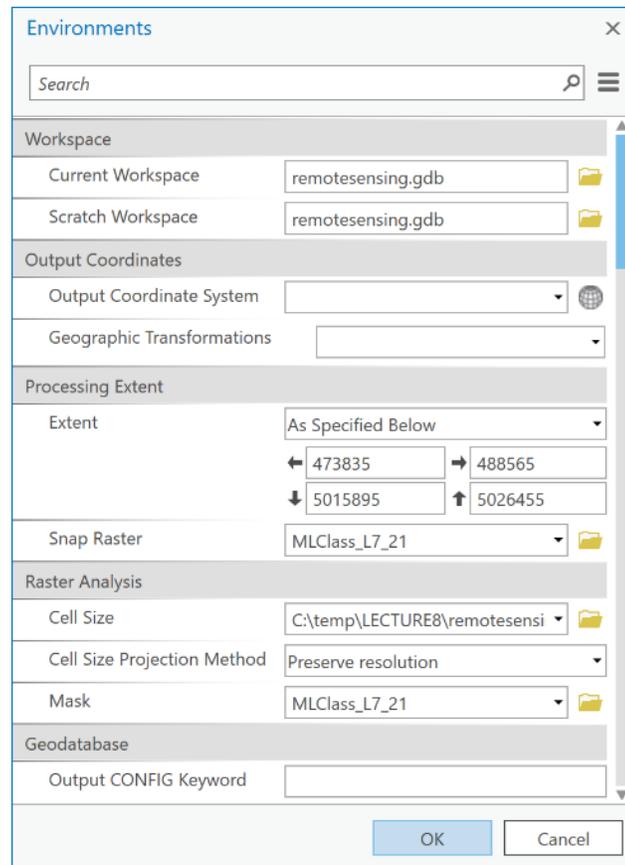
Vous devez donc indiquer à ArcGIS comment résumer les cellules "MLCLASS_RE_21_filtered" lorsque vous augmentez la taille des cellules de cette couche. En d'autres mots, dans chaque nouvelle cellule de 30 x 30 m, ArcGIS doit regarder les 36 cellules plus petites qui la composent et doit décider comment calculer la valeur (Forêt ou Non-Forêt) à l'intérieur de la nouvelle taille de cellule. Comme il s'agit de données nominales (Forêt et Non-Forêt), ArcGIS prend par défaut la valeur de la cellule de 5 x 5 m la plus proche du centre de la nouvelle cellule de 30 x 30 m. Cependant, avec des données nominales, une fonction d'agrégation plus appropriée serait que la nouvelle cellule de 30 x 30 m prenne une valeur correspondant à la majorité des valeurs qu'elle contient. Par exemple, s'il y a 24 cellules forestières et 12 cellules non forestières dans la nouvelle cellule de 30 m, la nouvelle cellule à 30 m doit être une cellule forestière, comme dans l'exemple ci-dessous.



Lorsque vous généralisez votre couche raster à la même taille de cellule que le résultat de la classification Landsat 7, vous devez également vous assurer que les cellules de la couche généralisée s'alignent correctement sur les cellules de la couche Landsat 7. Pour ce faire, vous devez définir le paramètre « *Snap Raster* » dans les paramètres de l'environnement global ainsi que le paramètre « *Processing Extent* »:

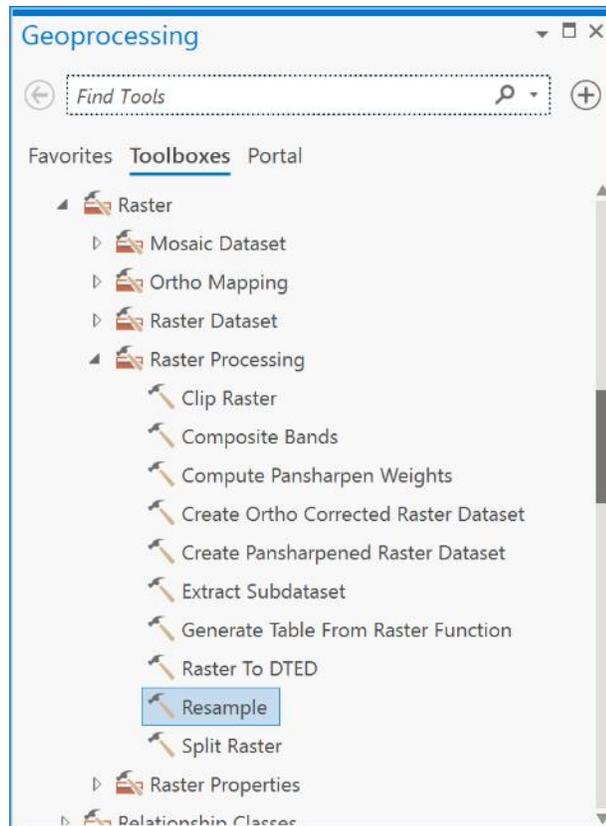
Nom de l'étudiant: _____ # d'étudiant: _____

1. Dans la boîte de dialogue « *Environment Settings* », sous « *Processing Extent* », choisissez "MLClass_L7_21" et sous « *Snap Raster* », choisissez également "MLClass_L7_21". Ensuite, sous « *Raster Analysis* », « *Cell Size* », choisissez "MLClass_L7_21" et sous « *Mask* », choisissez également "MLClass_L7_21", puis cliquez sur OK:

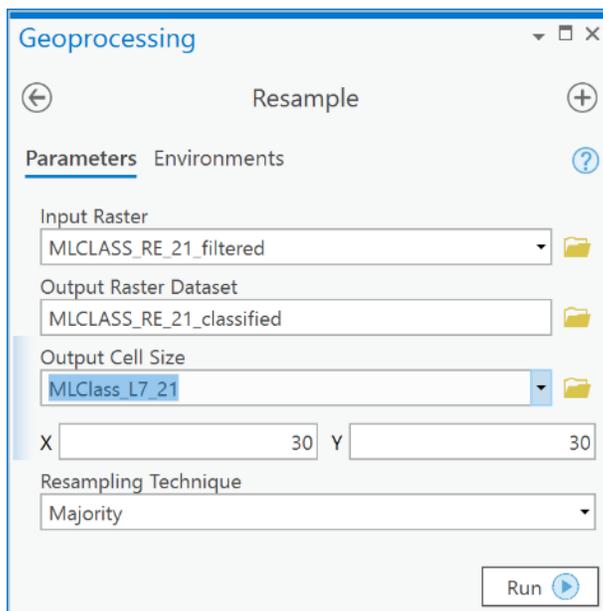


En complétant l'étape 2, vous vous êtes assuré que tout processus de traitement dans l'analyse raster sera conforme à la taille des cellules et à l'étendue de la couche "MLCLASS_L7_21". N'oubliez pas que ces paramètres seront conservés jusqu'à ce que vous fermiez et rouvriez ArcGIS ou que vous les modifiez à nouveau. Ainsi, si vous décidez d'effectuer d'autres analyses, avec d'autres couches, dont l'étendue est différente de "MLCLASS_L7_21", les résultats ne seront produits qu'à l'intérieur de l'étendue spécifiée. Il se peut donc que vous deviez effacer la propriété d'étendue ou la ramener à sa valeur par défaut lorsque vous aurez terminé les analyses raster.

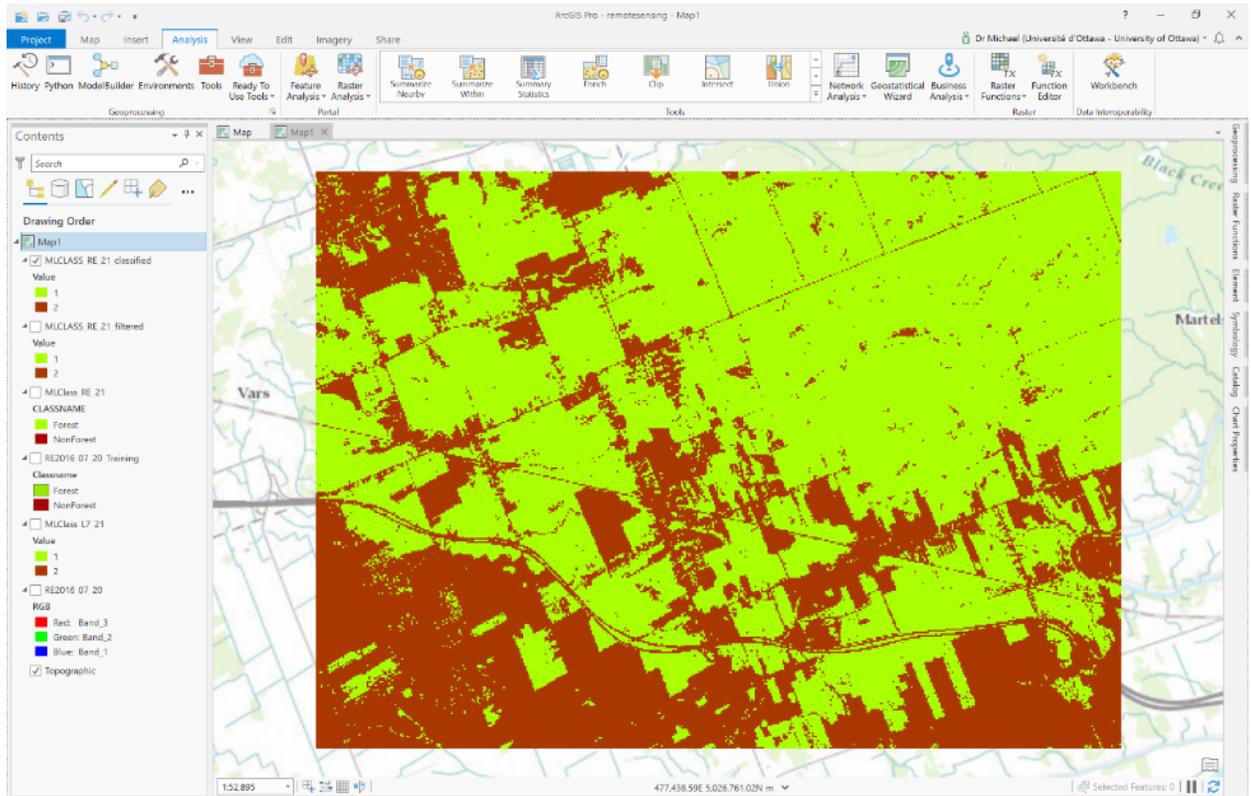
2. Pour convertir le produit RapidEye à une résolution de 30 m, vous utilisez l'outil « *Resample* » sous « *Data Management* » -> « *Raster* » -> « *Raster Processing* ». Ouvrez cet outil:



3. Dans la boîte de dialogue « *Resample* », votre raster d'entrée (« *Input Raster* ») doit être spécifié comme "MLCLASS_RE_21_filtered", le raster de sortie (« *Output Raster Dataset* ») doit être appelé "MLCLASS_RE_21_classified" et la taille de cellule de sortie (optionnelle) doit être spécifiée "MLclass_L7_21". **Plus important encore**, la technique de rééchantillonnage facultative (« *Resampling Technique* ») doit être réglée sur MAJORITÉ (« *Majority* »). Cliquez ensuite sur OK.



4. Votre résultat agrégé ressemblera à ceci:

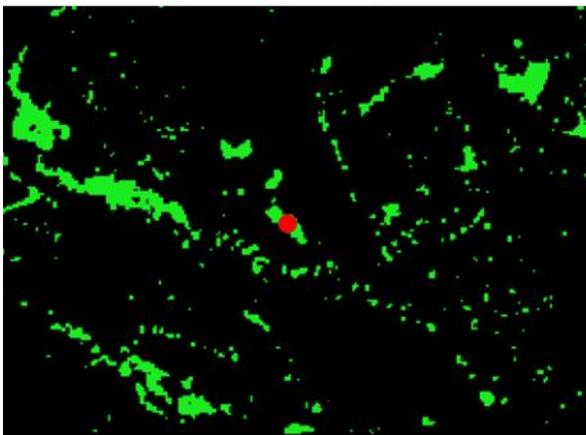


Q8: Comment puis-je vérifier l'exactitude d'un résultat classifié?

La précision d'une classification est testée en comparant les pixels classifiés à ce que l'on appelle les données de référence. Les données de vérité terrain (VT) sont des lieux qui n'ont pas été utilisés dans la classification mais dont on sait qu'ils appartiennent à l'une des classes d'information que vous avez définies, c'est-à-dire qu'ils appartiennent à la catégorie Forêt ou Non-Forêt. En d'autres termes, les données VT constituent un "ensemble de données de test" et sont indépendantes des données utilisées pour créer le fichier de signatures qui a entraîné le classificateur. Les données VT peuvent être acquises en allant sur le terrain avec une unité GPS et en observant plusieurs endroits pour détecter la présence d'une classe d'information particulière. Toutefois, sauf nécessité absolue, d'autres méthodes de collecte de données VT indépendantes sont généralement utilisées, car il est souvent difficile, lent et coûteux d'acquérir des données VT sur place. Ainsi, vous pouvez généralement collecter des exemples de données VT en observant directement les caractéristiques de l'image pré-classifiée et en les étiquetant comme l'une des classes d'information, dans notre cas comme Forêt ou Non-Forêt. Pour évaluer la précision d'une classification, vous devez collecter ces données en plusieurs points choisis au hasard et calculer une matrice de confusion qui vous permettra de quantifier le nombre de fois où la classification est correcte et le nombre de fois où elle ne l'est pas.

Note : Le terme "vérité de terrain" est mal vu par certains, car il implique que ces données sont intrinsèquement "vraies". Ce n'est évidemment pas le cas, car ces données sont produites par des humains, et les humains font des erreurs. Néanmoins, nous utilisons ce terme ici parce qu'il est courant et que vous le rencontrerez à l'avenir si vous travaillez davantage avec l'imagerie satellitaire. D'autres termes utilisés pour la même chose sont "données de validation" ou "données de test", mais ces termes eux-mêmes peuvent avoir des significations différentes pour différentes personnes, et ne sont donc pas idéaux non plus. Nous nous en tiendrons à "vérité terrain" pour cet exercice !

Classified result



Multispectral image

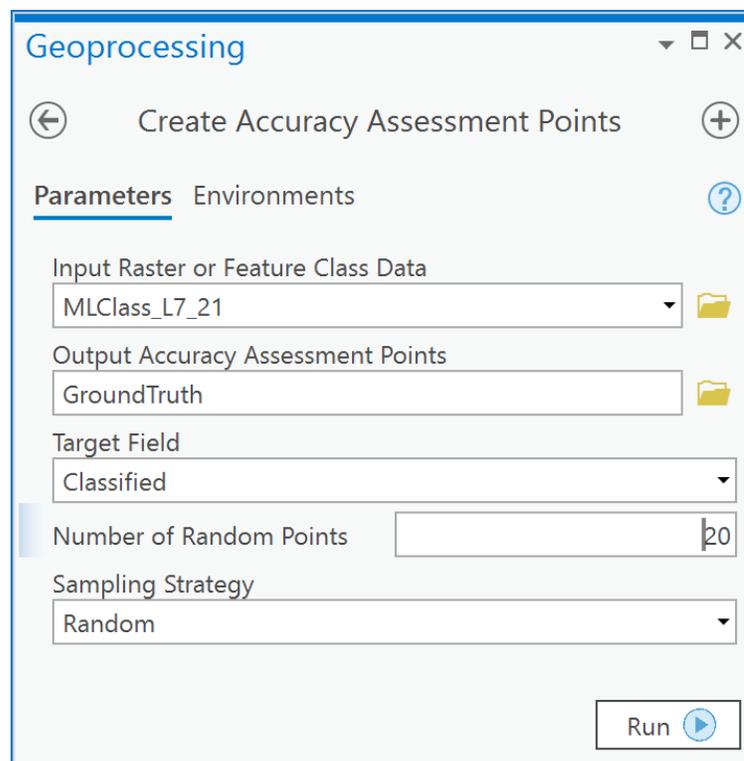


Exemple de classification erronée. Le pixel classé comme forêt (à gauche) n'est manifestement pas une forêt (à droite), comme le montre un point de vérité terrain (croix rouge, à droite) identifié manuellement à partir d'une image multispectrale.

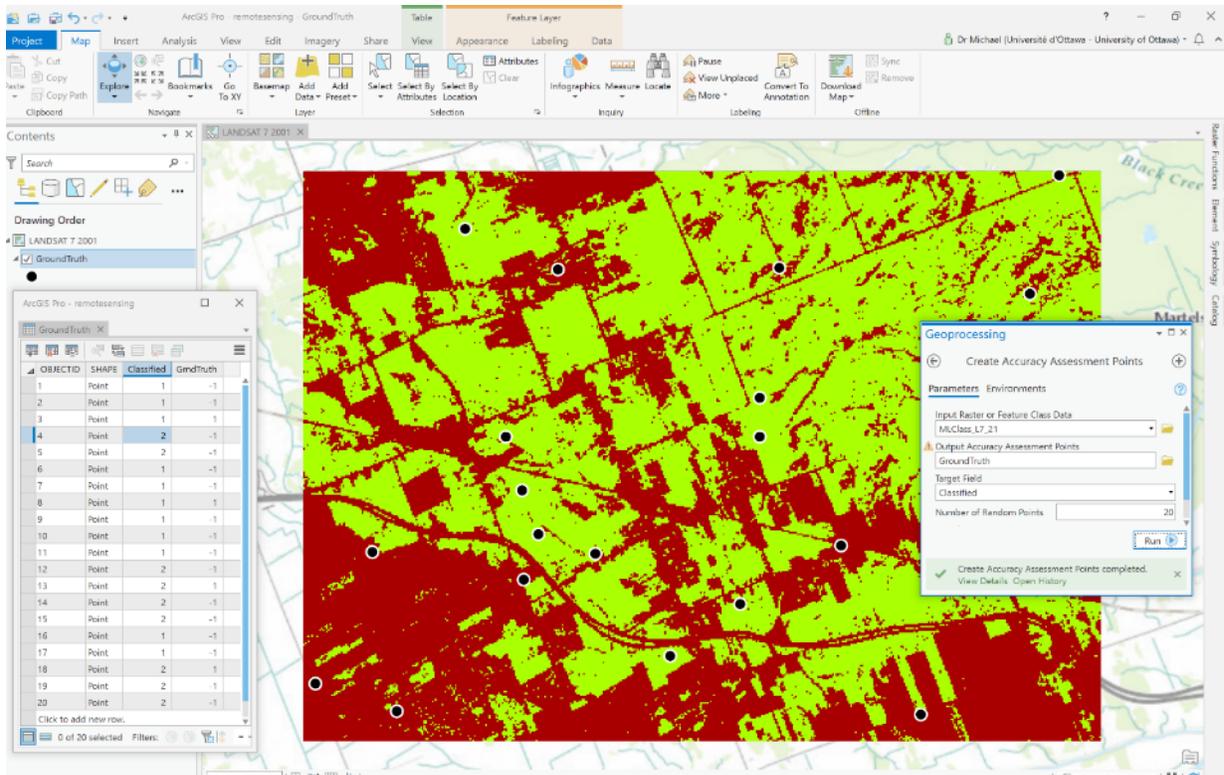
Vous calculerez la précision du résultat "MLCLASS_L7_21", puisque tout le monde aura le même résultat à partir de la question 2 ci-dessus.

1. Allez dans la boîte à outils, trouvez « *Spatial Analyst* » -> « *Segmentation and Classification* » et choisissez l'outil « **Create Accuracy Assessment Points** » (Créer des points d'évaluation de précision).
2. Dans l'outil, sous « *Input raster or feature class data* », choisissez la couche "MLCLASS_L7_21" et sous « *Output Accuracy Assessment Points* », appelez le résultat "GroundTruth" et assurez-vous de l'enregistrer dans la géodatabase IMAGERY.gdb. Sous « *Target Field* », choisissez « *Classified* ». Sous « *Number of Random Points* » (nombre de points aléatoires), choisissez 20, et enfin sous « *Sampling Strategy* » (stratégie d'échantillonnage), choisissez « *Random* » (aléatoire).

Cela créera 20 points aléatoires (où les coordonnées x et y de chaque point sont choisies, dans l'étendue de l'image, à partir de la fonction de densité de probabilité uniforme, indépendamment et avec une probabilité égale) dans chacune des classes "Forêt" et "Non-Forêt". Pour chaque point aléatoire, la valeur du pixel de la couche "MLCLASS_L7_21" sera extraite à l'emplacement du point dans la couche "MLCLASS_L7_21". Ces valeurs seront stockées dans la table « *GroundTruth* » dans une colonne appelée « *Classified* » contenant les valeurs de pixel classées pour chacun des points.



3. Le résultat sera 20 points situés de manière aléatoire dans l'image:

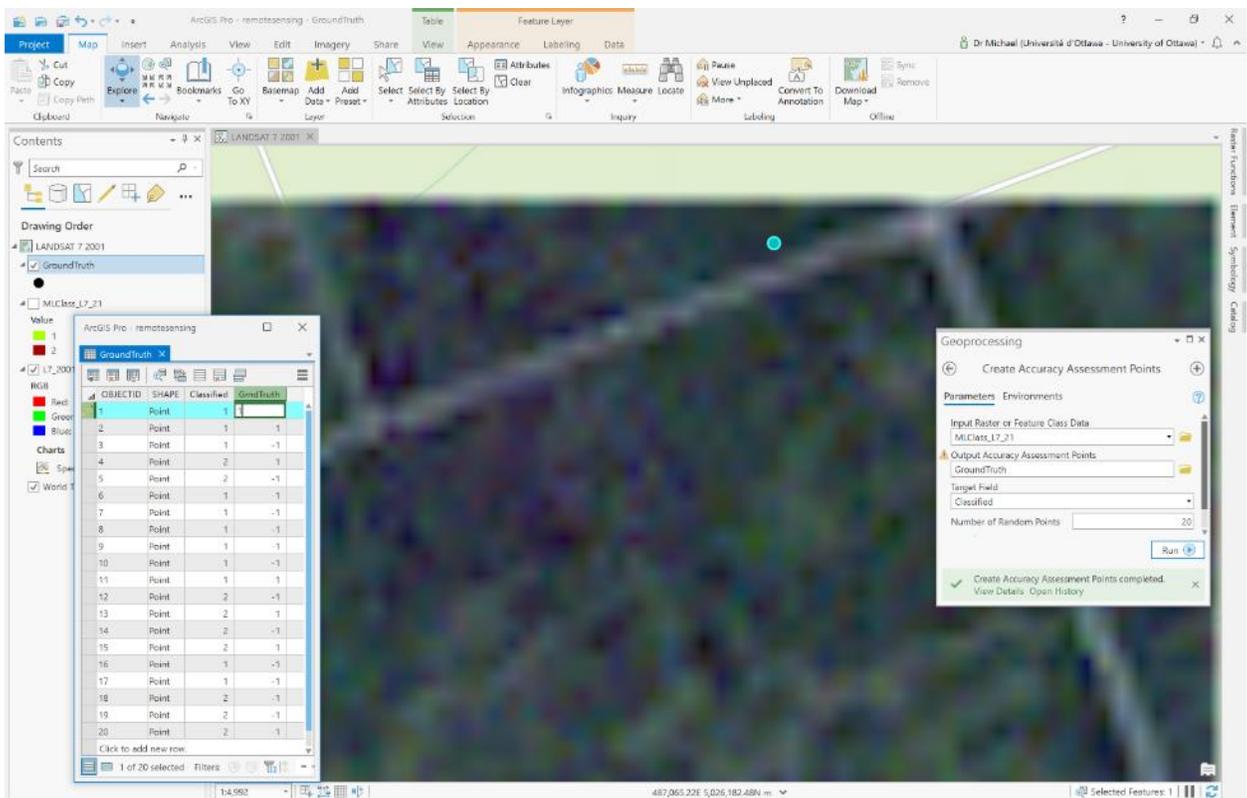


4. Ouvrez maintenant le tableau appelé « *GroundTruth* ». Il devrait ressembler à ceci:

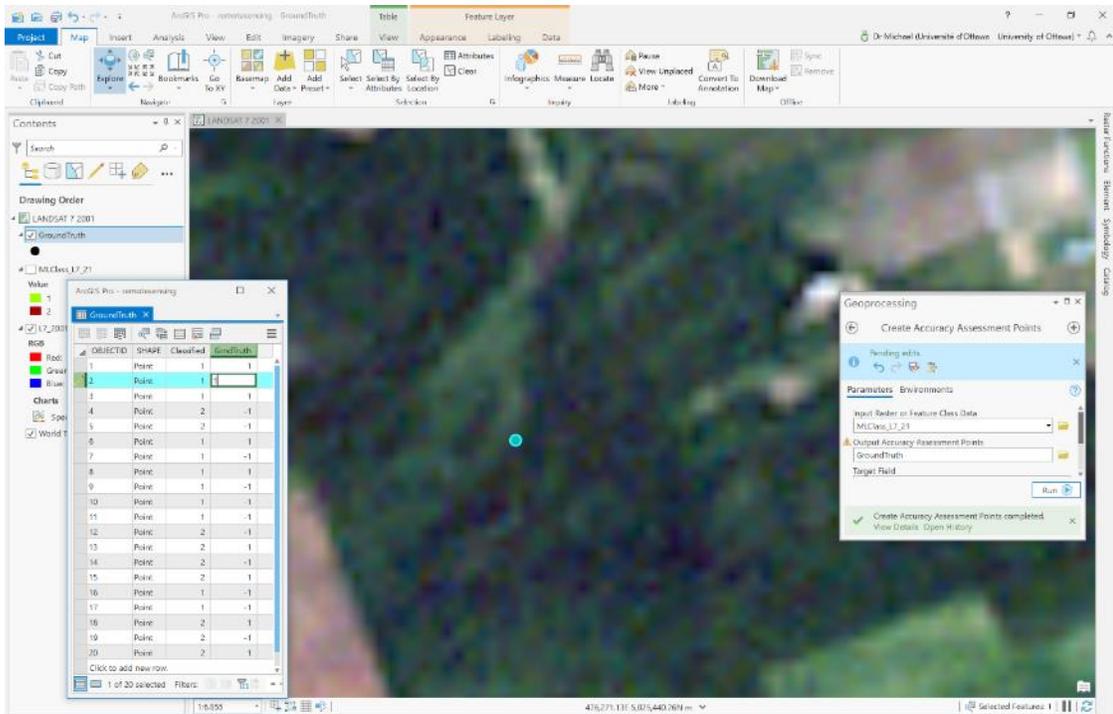
The screenshot shows the 'GroundTruth' table in ArcGIS Pro. The table has four columns: 'OBJECTID', 'SHAPE', 'Classified', and 'GrndTruth'. The data is as follows:

OBJECTID	SHAPE	Classified	GrndTruth
1	Point	1	-1
2	Point	1	-1
3	Point	1	-1
4	Point	2	-1
5	Point	2	-1
6	Point	1	-1
7	Point	1	-1
8	Point	1	-1
9	Point	1	-1
10	Point	1	-1
11	Point	1	-1
12	Point	2	-1
13	Point	2	-1
14	Point	2	-1
15	Point	2	-1
16	Point	1	-1
17	Point	1	-1
18	Point	2	-1
19	Point	2	-1
20	Point	2	-1

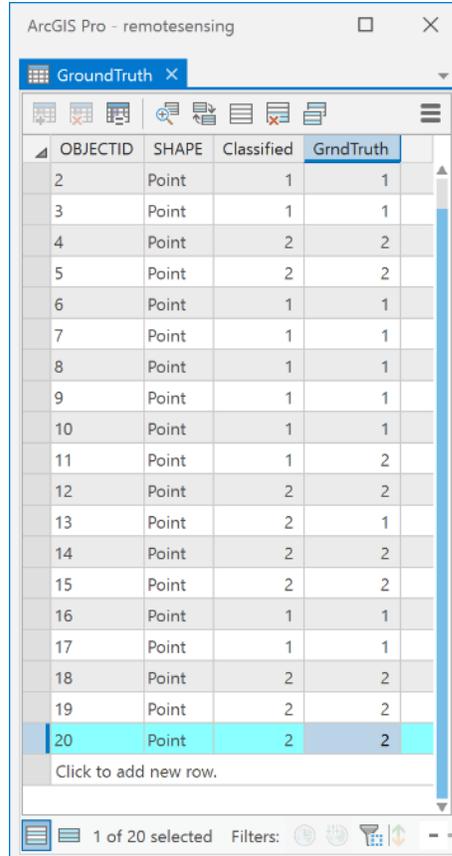
- Chaque point aléatoire est associé à la valeur de la classe d'information Forêt = 1 ou Non-Forêt = 2, en fonction de la classe dans laquelle se trouve le point. Ces données figureront dans la colonne « Classified » du tableau d'attributs « GroundTruth » (voir le tableau ci-dessus).
- Nous devons maintenant remplir la colonne **GrndTruth** du tableau, dont les valeurs sont actuellement de -1. Les valeurs de -1 nous indiquent que la "vraie" classe pour ce point n'a pas été observée et entrée dans le tableau. Nous devons donc remplir ces "vraies" valeurs de classe, en nous rappelant que la valeur 1 signifie Forêt et la valeur 2 signifie Non-Forêt. Pour cela, assurez-vous que les seules couches visibles sur la carte sont "L7_2001_08_25" et « GroundTruth ». Sélectionnez maintenant la première ligne du tableau "GroundTruth", afin d'identifier le premier point. Pour la ligne actuellement sélectionnée, cliquez avec le bouton droit de la souris et choisissez « *Zoom To* ». Zoomez ensuite un peu plus près jusqu'à ce que vous puissiez voir clairement dans quelle classe d'information se trouve le point sélectionné. Saisissez ensuite un 1 ou un 2 dans la colonne **GrndTruth**, en fonction de votre interprétation de l'image à ce point. Dans l'exemple ci-dessous, vous pouvez voir que la forêt se trouve sous le point sélectionné, vous devez donc saisir une valeur de 1 dans la colonne GrndTruth pour cette ligne double-cliquant sur la cellule où vous souhaitez inscrire le 1.



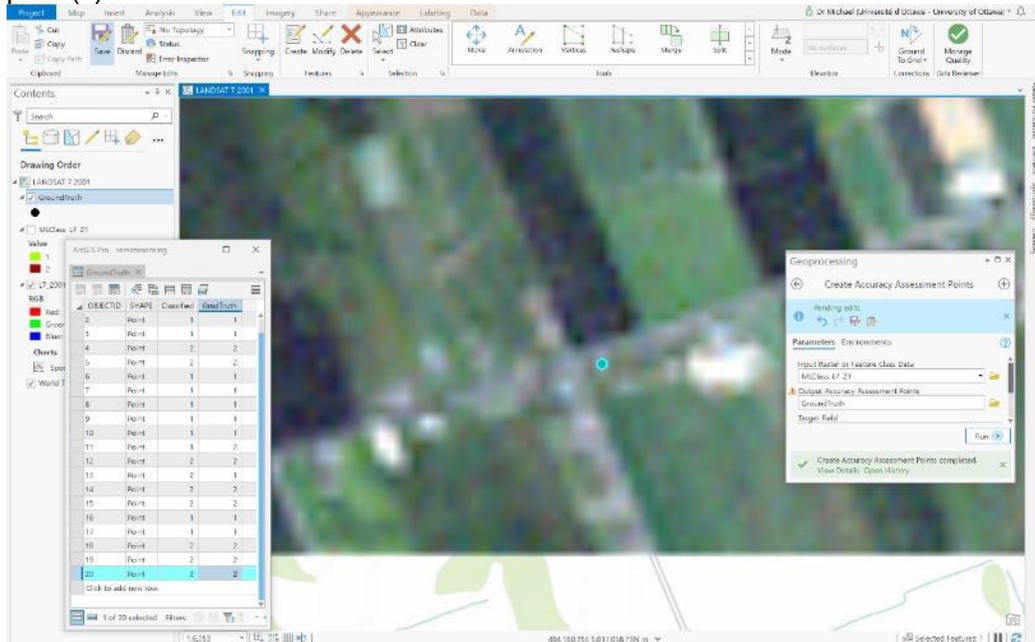
- Sélectionnez maintenant la deuxième rangée, "Zoom To", et examinez s'il y a de la forêt ou non à cet endroit et entrez une valeur de 1 ou 2, respectivement, dans la colonne "GrndTruth" du tableau "GroundTruth". Dans l'exemple ci-dessous, il s'agit à nouveau d'une forêt, vous devez donc entrer une valeur de 1:



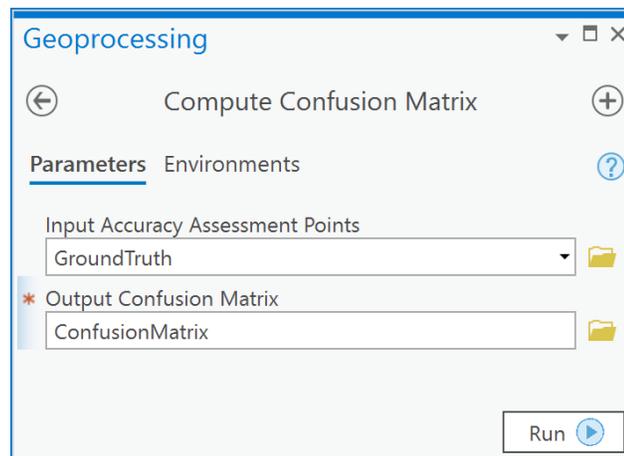
8. Répétez ces étapes pour tous les points du tableau "GroundTruth", afin de créer un tableau complet et prêt pour l'évaluation de la précision, par exemple:



9. Allez maintenant dans l'onglet « *Edit* » et cliquez sur « *Save* » pour sauvegarder les modifications apportées au tableau "GroundTruth". Cliquez ensuite sur le bouton « *Clear* » dans le groupe « *Selection* » du même onglet pour désélectionner le/les point(s):



10. Votre tableau "GroundTruth" contient maintenant une colonne appelée **Classified** qui contient le résultat du classificateur à chaque point et une colonne appelée **GrndTruth** qui contient la classe d'information observée à chaque point. Maintenant, vous voulez calculer combien de fois **Forest** (valeur 1) correspond à **NonForest** (valeur 2), combien de fois **Forest** (valeur 1) correspond à **Forest** (valeur 1), combien de fois **NonForest** (valeur 2) correspond à **NonForest** (valeur 2) et enfin combien de fois **NonForest** (valeur 2) correspond à **Forest** (valeur 1). Comme nous ne traitons que deux classes dans cet exemple, vous pourriez facilement le faire manuellement, mais ArcGIS dispose d'un outil qui le fait pour nous, appelé « *Compute Confusion Matrix* », qui se trouve sous « *Spatial Analyst* » -> « *Segmentation and Classification* » dans ArcToolbox. Ouvrez cet outil et, sous « *Input Accuracy Assessment Points* », choisissez votre table "GroundTruth". Appelez le résultat ConfusionMatrix et sauvegardez-le dans votre fichier IMAGERY.gdb :



11. Ouvrez le tableau ConfusionMatrix:

OBJECTID	ClassValue	C_1	C_2	Total	U_Accuracy	Kappa
1	C_1	10	1	11	0.909091	0
2	C_2	1	8	9	0.888889	0
3	Total	11	9	20	0	0
4	P_Accuracy	0.909091	0.888889	0	0.9	0
5	Kappa	0	0	0	0	0.79798

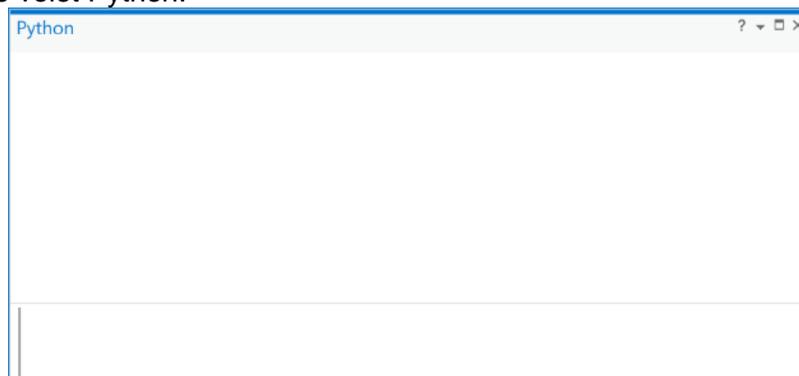
Dans le tableau, C_1 signifie la valeur de la classe 1 (Forêt), et C_2 signifie la valeur de la classe 2 (Non-Forêt). Les valeurs indiquées dans les cellules où une colonne et une rangée du même nom se croisent indiquent le nombre de fois où les deux classes correspondantes ont été correctement classées. Par exemple, dans le tableau illustré ci-dessus, il est arrivé 10 fois qu'un C_1 corresponde à un C_1 (où la colonne C_1 croise la rangée C_1) et 1 fois qu'un C_1 corresponde à un résultat classifié C_2. Ainsi, 10 points de Forest ont été correctement classifiés et 1 a été incorrectement classifié. Similairement, il y a eu 8 fois où un C_2 a été correctement classifié comme un C_2, et une fois où il a été classifié comme un C_1. Ainsi, 8 fois sur 9, NonForest a été correctement classifié. La valeur de Kappa est une mesure, allant de -1 à 1, qui indique dans quelle mesure les performances du classificateur sont supérieures à celles d'un classificateur aléatoire. Ici, elle est de 0,797, ce qui est considéré comme un résultat relativement bon. Vous pouvez consulter le fichier d'aide de l'outil « *Compute Confusion Matrix* » pour plus d'informations.

12. Maintenant, répétez toutes les étapes pour cette question avec votre couche "MLCLASS_RE_21_classified" en utilisant l'image RE2016_07_20 comme couche de vérité terrain.

Q9: Comment puis-je créer un raster contenant uniquement les forêts à partir de chacun de mes résultats classifiés, et déterminer où les forêts ont diminué?

Puisque vous avez maintenant deux rasters, tous deux contenant des classes d'information appelées **Forest** et **NonForest**, vous pouvez utiliser une simple expression d'algèbre de carte pour créer une couche raster booléenne où la valeur de 1 signifie **Forest** et 0 signifie **NonForest**. Vous pouvez le faire de plusieurs façons, ici nous allons essayer d'utiliser le langage de programmation Python qui est intégré dans ArcGIS :

1. Ouvrez le volet Python:



2. Tapez les trois lignes suivantes et appuyez sur la touche Entrée (notez que vous n'avez pas besoin de taper les lignes en vert qui sont précédées du signe #, puisqu'il ne s'agit que de commentaires pour vous indiquer ce que fait chaque ligne):

```
# accéder aux fonctions de Spatial Analyst
from arcpy.sa import *

# Créer une variable raster à partir des résultats de la classification
filtrée
x = Raster("MLCLASS_RE_21_classified")

# Créer un raster booléen en utilisant l'opérateur relationnel pour les
égalités logiques
forest_2016 = x == 1
```

La première ligne indique à ArcGIS de vous permettre d'accéder aux fonctions de Spatial Analyst dans le volet python. La deuxième ligne utilise la fonction Raster() pour créer un objet raster qui peut être utilisé dans une expression d'algèbre de carte Python. La troisième ligne est une expression d'algèbre de carte relationnelle utilisant des égalités logiques (==). Cette expression examine chaque cellule de la couche raster appelée x, et si la cellule a une valeur de 1 (**Forêt**), alors la nouvelle valeur de la cellule raster aura également une valeur de 1. Si la valeur de la cellule est différente de 1, la valeur du raster de la cellule en question sera égale à zéro.

Nom de l'étudiant: _____ # d'étudiant: _____

- Le résultat sera une couche matricielle appelée "forest_2016" dans le volet Contenu (« Contents »), mais avec des valeurs de pixel de 1 pour les cellules qui sont des forêts et de 0 sinon.
- Modifiez maintenant ces lignes pour créer un raster binaire pour l'image Landsat classifiée:

```
# accéder aux fonctions de Spatial Analyst  
from arcpy.sa import *
```

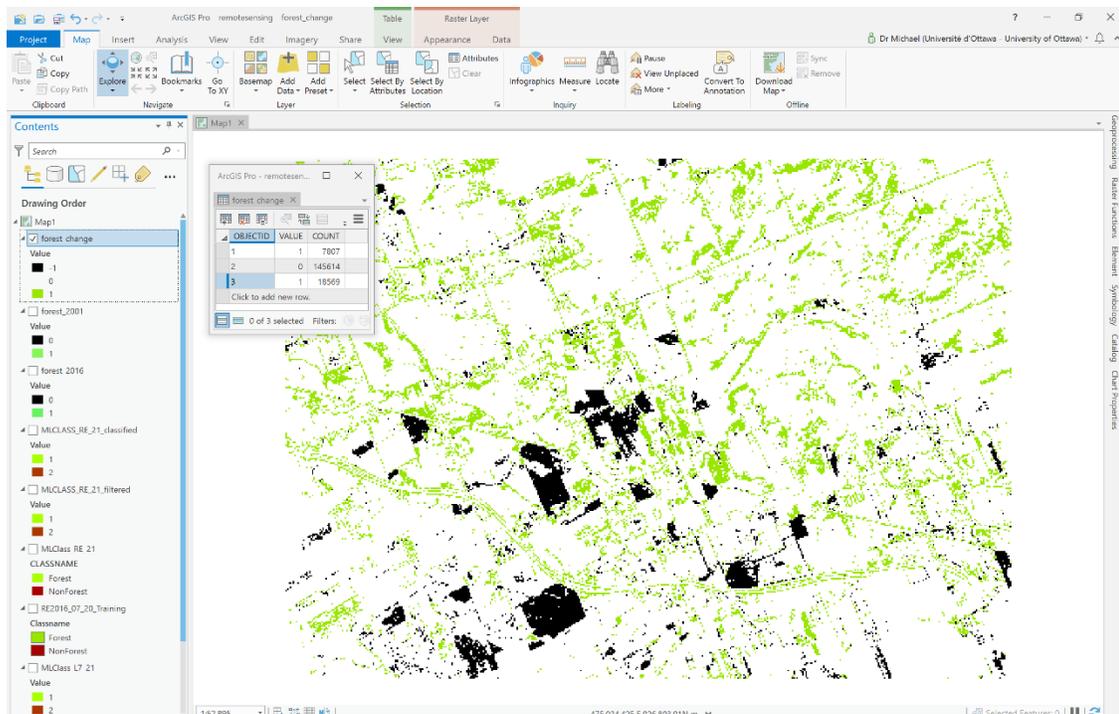
```
# Créer une variable raster à partir des résultats de la classification filtrée  
y = Raster("MLClass_L7_21")
```

```
# Créer un calque booléen en utilisant l'opérateur relationnel pour les égalités logiques  
forest_2001 = y == 1
```

- Maintenant, pour déterminer la quantité de forêt qui a été perdue ou gagnée, nous pouvons soustraire la forêt_2001 de la forêt_2016:

```
# Soustraire 2016 de 2001 pour voir où la forêt a gagné et/ou perdu du terrain  
forest_change = forest_2016 - forest_2001
```

- La carte obtenue ressemblera à ceci (elle ne sera pas exactement la même car les données d'entraînement pour le RapidEye sont différentes):



Ici, la table d'attributs nous indique dans le champ « Count » le nombre total de pixels pour chaque valeur.

Questions à remettre

1. [1 point] Quelle est l'incertitude positionnelle due à la résolution des cellules dans la couche "MLCLASS_L7_21" ?
2. [3 points] Dans Q2, vous avez créé une couche raster thématique appelée "MLCLASS_L7_21", et dans Q3-Q7, vous avez créé une couche appelée "MLCLASS_RE_21_classified". Dans Q8, vous avez ensuite créé des matrices de confusion pour les deux couches. Présentez une carte pour chaque résultat et chaque matrice de confusion. Expliquez quel résultat est le plus précis, comment vous le savez et quelle en est, selon vous, la raison.
3. [3 points] Dans le raster que vous avez créé à la fin de la Q9, vous aviez des valeurs de cellule de -1, 0 et 1. Décrivez ce que ces valeurs représentent dans le paysage, par exemple, que signifie une valeur de cellule de -1, que signifie une valeur de cellule de 0 et que signifie une valeur de cellule de +1 ? Présentez la carte avec votre explication. Quelle est votre conclusion générale sur la perte de forêts dans la région?
4. [3 points] Dans la couche différenciée que vous avez créée à la fin de la Q9, vous pourriez être intéressé par les grandes régions contiguës de cellules avec des valeurs qui indiquent une perte de forêt (conversion de forêt à non-forêt). Où se trouvent ces grandes régions contiguës ? Utilisez l'outil « *Region Group* » sous « *Spatial Analyst* » -> « *Generalization* » avec votre résultat de Q9 comme couche d'entrée. Lisez le fichier d'aide de l'outil « *Region Group* » en cliquant sur le bouton Aide de l'outil. Expliquez ce que fait l'outil « *Region Group* ». Maintenant, ouvrez la table d'attributs du raster résultant du « *Region Group* », cliquez avec le bouton droit de la souris sur la colonne "Count" et triez par ordre décroissant (« *Sort Descending* »). Examinez le champ "Link". Comment les valeurs du champ "Link" se relient-elles aux valeurs des cellules dans le résultat de la Q9 ? Fournir un tableau des surfaces en unités cartographiques pour les 10 plus grands groupes contigus de cellules qui sont passées de **Forêt** à **Non-Forêt** entre 2001 et 2016. Présentez également une carte montrant ces régions.
5. [2 points] Quelle est la taille moyenne des régions (taille des parcelles) en mètres carrés pour les régions qui sont passées de la catégorie Forêt à la catégorie Non-Forêt entre 2001 et 2016 ? Pour répondre à cette question, vous devrez effectuer une requête sur la table d'attributs de la couche des régions groupées (de #4) et faire un clic droit sur la colonne "Count" et choisir "Statistics" pour obtenir la moyenne, pour répondre à cette question.
6. [2 points] Sélectionnez uniquement les régions qui sont passées de la catégorie **Forêt** à la catégorie **Non-Forêt** entre 2001 et 2016 et dont la superficie est supérieure ou égale à 3,000 mètres carrés. Quelle est la pente moyenne de chaque région ? Pour répondre à cette question, vous devrez utiliser l'outil « *Spatial Analyst* » -> « *Surface* » -> « *Slope* » avec les données DEM (il s'agit d'un Modèle Numérique de Terrain) trouvées dans IMAGERY.gdb, ainsi qu'une requête d'attribut sur la couche des régions groupées (de #4) et « *Spatial Analyst* » -> « *Zonal Statistics as Table* », pour répondre à cette question. Présentez un tableau montrant la pente moyenne pour chaque région sélectionnée.

Nom de l'étudiant: _____ # d'étudiant: _____

7. [1 point] Examinez la couche appelée L8_2016_08_26, qui est une image Landsat 8 OLI. Utilisez votre outil d'identification pour cliquer sur quelques endroits différents et notez la VN à chaque fois. Comparez ces VN avec les VN pour des caractéristiques similaires dans la couche "L7_2001_08_25". Pourquoi les VN de l'image Landsat 8 sont-ils plus élevés que ceux de l'image Landsat 7 pour des éléments similaires ?
8. [2 points] Créez et présentez une carte NDVI (Indice de Végétation par Différence Normalisée) en utilisant l'image L8_2016_08_26. Consultez vos notes de cours ou l'Internet pour déterminer les bandes appropriées à utiliser pour calculer le NDVI avec les données Landsat 8, car elles ne sont pas les mêmes que pour Landsat 7. Qu'est-ce que cette carte vous apprend ? Pourquoi ?