

Université du Québec à Montréal (UQAM)
Faculté des sciences

ACT3035– Examen Intra (R)
Laboratoire d'actuariat

Enseignant: Noureddine Meraihi

2019/02/20

Nom: _____

Code permanent: _____

Signature: _____

Cet examen contient 6 pages (incluant la page couverture) et 5 questions sur un total de 50points.
Bon succès à tous!

Distribution des points

| Question | Points | Score |
|----------|--------|-------|
| 1 | 10 | |
| 2 | 10 | |
| 3 | 10 | |
| 4 | 10 | |
| 5 | 10 | |
| Total: | 50 | |

Instructions

- L'examen commence à 9:00 pour une durée de 180 minutes;
- Prévoyez un 5 minutes pour la remise de votre examen;
- Chaque minute de retard vous coûtera 5% de cet examen;
- Vous avez le droit de consulter vos notes de cours personnelles;
- Vous avez le droit de consulter les notes de cours du blog;
- Vous avez le droit de consulter l'aide de RStudio;
- Il est strictement interdit de faire des recherches sur le web (tous vos logs sur les ordis de l'UQAM sont sauvegardés);
- Il est strictement interdit d'utiliser un quelconque moyen de communication pendant l'examen (logs sauvegardés);
- Pour toutes les questions, le terme **df** désigne *data frame*
- N'oubliez pas de sauvegarder aussi souvent que possible (Ctrl+s)!
- Le nom de votre fichier doit contenir votre code permanent, par exemple: MERN12345678.r
- Inscrivez votre code permanent sur le questionnaire, remettez-le au surveillant.
- L'examen compte pour 50% de la note finale;
- vous n'êtes pas obligés d'imprimer les notes du cours

Avec la disponibilité des données en ligne aujourd'hui, il n'est pas rare de voir les compagnies d'assurance utiliser ces données afin de compléter ou d'améliorer leur modèle actuariel. Nous profitons alors d'utiliser ces données mêmes pour poser des questions d'examen.

Les questions #1 et #2 sont basées sur les données de la ville de Montréal décrivant les actes criminels enregistrés par le Service de police de la Ville de Montréal (SPVM). On peut par exemple utiliser ces données pour la tarification automobile.

La question #3 est quant à elle basée sur les données du marché immobilier de la région de Montréal. On peut par exemple utiliser ces données afin mettre à jour la valeur immobilière du portefeuille d'assurance ou étudier le marché immobilier pour des investisseurs potentiels.

1. (10 points) Compléter le code ci-dessous, qui est fourni dans le script de réponse (donc vous n'avez pas besoin de le retaper, seulement le compléter) afin d'obtenir les résultats ¹ du tableau (1).

```
1 Sys.setlocale(locale="fr_FR.UTF-8")
2 lien <- "http://donnees.ville.montreal.qc.ca/dataset/5829b5b0-ea6f-476f-be94-bc2b8797769a/resource/c6f482bf-bf0f-4960-8b2f-9982c211addd/download/interventionscitoyendo.csv"
3 df <- read.csv(..., stringsAsFactors = FALSE)
4 df$CATEGORIE[df$... %in% "Infractions entraînant la mort"] <- "InfractionsEntrainantLaMort"
5 df$CATEGORIE[df$... %in% "M\%xe9fait"] <- "Mefait"
6 df$CATEGORIE[df$... %in% "Vol dans / sur v\%xe9hicule \xe0 moteur"] <- "VolDansSurVehiculeMoteur"
7 df$CATEGORIE[df$... %in% "Vol de v\%xe9hicule \xe0 moteur"] <- "VolDeVehiculeAMoteur"
8 df$CATEGORIE[df$... %in% "Vols qualifi\%xe9s"] <- "VolQualifie"
9 ... #fonction qui compte le nombre d'occurrence dans une variable
```

| Var1 | Freq |
|-----------------------------|-------|
| InfractionsEntrainantLaMort | 111 |
| Introduction | 36991 |
| Mefait | 30006 |
| VolDansSurVehiculeMoteur | 40198 |
| VolDeVehiculeAMoteur | 18095 |
| VolQualifie | 7676 |

Tableau 1: Résultat attendu de la question 1-b

2. (10 points) Dans la base de données de la question #1, nous possédons les coordonnées géographiques (longitude et latitude) de chaque acte criminel commis sur la carte de Montréal. Il est possible de tracer ces points sur une carte en utilisant la librairie **leaflet**. On vous demande de faire un échantillonnage de 1000 observations des coordonnées géographiques afin de tracer le graphique illustré à la figure (1) :

NB: Vous n'avez qu'à compléter le code ci-dessous.

```
1 ... %>%
2 ...
3 dplyr::sample_n(...) %>%
4 leaflet() %>%
```

¹Les résultats peuvent être différents que ceux affichés, car les données sont mis à jours sur le site de la ville de Montréal

```

5 addTiles() %>%
6 setView(-73.7,45.54, zoom = 10) %>%
7 addCircleMarkers(radius = 1) %>%
8 addCircleMarkers(stroke=FALSE)

```

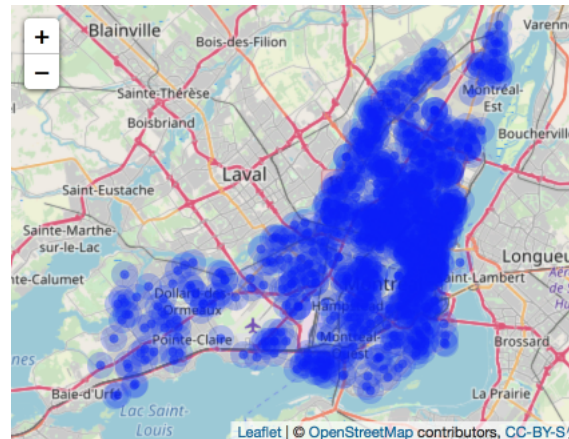


Figure 1: Question 2

3. (10 points) Les données `maison_montreal.csv` décrivent les propriétés vendues/achetées dans l'île de Montréal. Nous avons plusieurs variables qui décrivent chaque propriété. Nous nous intéressons plus particulièrement au prix d'achat moyen (moyenne du `BuyPrice`) par année de construction (`ConstructionYear`) dans la ville d'Outremont.

Après avoir choisi les variables d'intérêt, vous pouvez compléter le code ci-dessous afin d'obtenir le graphique illustré à la figure (2).

```

1 ggplot(data=..., aes(x=..., y=...)) +
2   geom_smooth(color="grey40") +
3   geom_point(color="red", size=.5) +
4   ylab("Prix moyen d'achat") +
5   xlab("Année de construction")

```

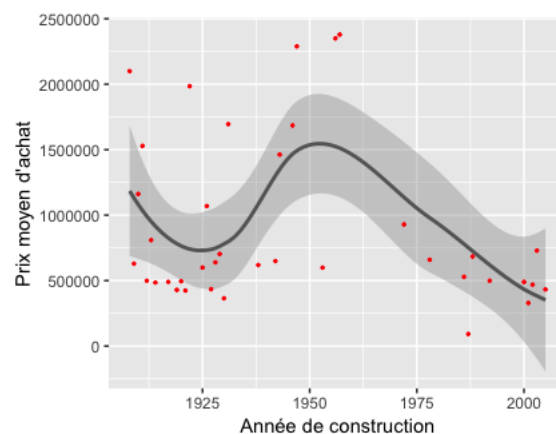


Figure 2: Question 3

4. (10 points) Écrivez une fonction appelée `CompInterSection` qui prends deux arguments (des vecteurs). Cette fonction vous retourne le complément de l'intersection des deux vecteurs que vous avez entré en argument. $A \cap B^c$

si $A = \{1, 3, 5, 7, 10\}$ et $B = \{1, 5, 10, 11, 13\}$, alors $A \cap B^c = \{3, 7\}$

```
1 > CompInterSection(a,b)
2 [1] 3 7
```

NB: N'utilisez pas les fonction intégrée des librairies, programmez votre propre fonction.

5. (10 points) La distance de Manhattan, appelée aussi taxi-distance, est la distance entre deux points parcourue par un taxi lorsqu'il se déplace dans une ville où les rues sont agencées selon un réseau ou quadrillage. Un taxi-chemin est le trajet fait par un taxi lorsqu'il se déplace d'un nœud du réseau à un autre en utilisant les déplacements horizontaux et verticaux du réseau.

Entre deux points A et B, de coordonnées respectives (X_A, Y_A) et (X_B, Y_B) , la distance de Manhattan est définie par :

$$d(A, B) = |X_B - X_A| + |Y_B - Y_A|$$

Écrivez une fonction appelée `Manhatan(A,B)` qui calcule cette distance Manhatan comme défini précédemment.

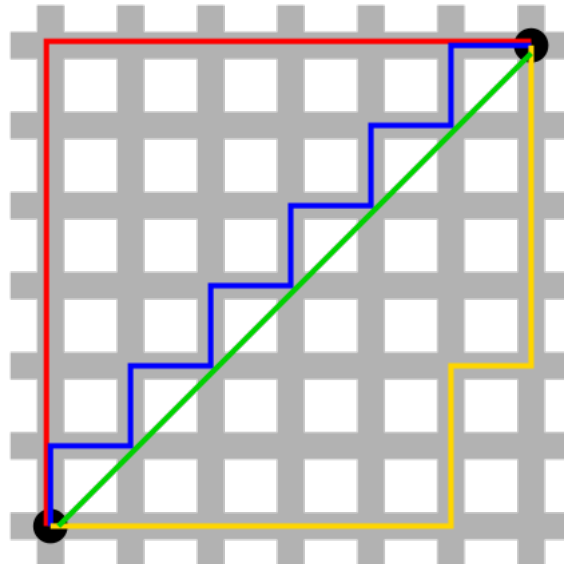


Figure 3: Distance Manhatan = 12 (rouge, bleu et jaune)

Fin de l'examen

- N'oubliez pas d'identifier votre fichier `MERN12345678.R` que vous remettez avec votre code permanent comme nom du fichier.
- Inscrivez votre code permanent sur le questionnaire, remettez-le au surveillant.
- Déposez votre examen final en suivant le lien dans le gabarit de réponse.
- Vous serez avisé par courriel quand les notes seront disponibles.