

Projet

Votre sous-titre

Vos noms

26 mars 2025

1 Introduction

Ce document est un modèle pour votre projet. Évitez des changements dans le préambule du document. Votre document doit avoir une introduction, une section d’exploration, une section d’analyse, et une discussion/conclusion finale. Les références sont également nécessaires : vous devez citer `tidyverse`, par exemple, ainsi que les œuvres consultées pour l’analyse statistique ou pour votre introduction.

Voici la première section. Notez l’étiquette entre `{ }` dans le code (faites attention aux espaces). Un mot **en gras** et un mot en *italique*. Selon Barnier 2023, la statistique est très utile! Consultez [cette page](#) pour savoir plus des citations. La note moyenne du fichier `villes2.csv` est de 70.26. Consultez la Section 2 pour apprendre comment générer des figures et des tableaux à partir de votre code R.

Vous pouvez également créer des listes numérotées ou des listes de points, bien que ce type d’élément ne soit pas commun dans un article académique :

- point 1
 - point 2
-
1. premier point
 2. deuxième point

2 Figures et tableaux

2.1 Figures

Il est également facile d’inclure un bloc de code dans votre document. Examinez la Figure 1. Dans le préambule de la figure, vous décidez quelle sera sa taille ainsi que sa légende.

L'argument `label` permet d'avoir des références croisées dans le texte, une caractéristique importante dans un document académique. Vous devez créer plusieurs figures **dans votre script**. Toutefois, il faut choisir quelle(s) figure(s) vous considérez suffisamment importante(s) pour le document final, ce qui dépendra de votre analyse, c'est-à-dire les variables choisies, etc. Une figure peut afficher plus de deux variables, naturellement, mais, si vous décidez d'inclure plusieurs variables dans votre modèle, il faudra avoir les figures cohérentes. Veuillez noter qu'il ne s'agit pas d'une opinion (bien que la statistique soit subjective). Vous êtes libres pour choisir les variables de votre analyse, mais il y a des analyses plus ou moins correctes selon ce choix, naturellement.

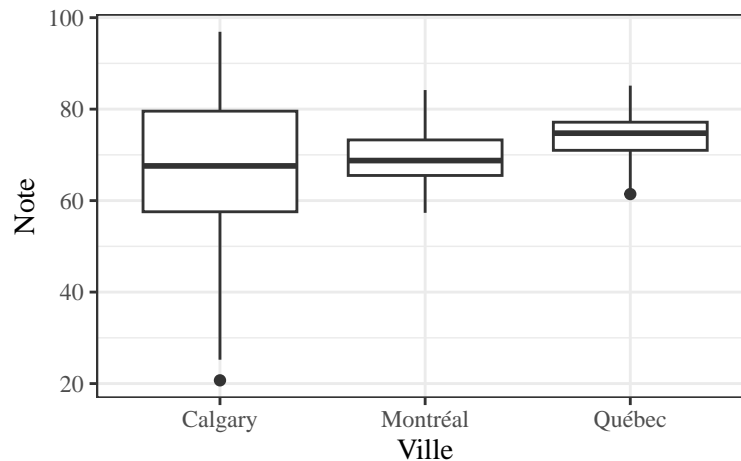


Figure 1: Une figure `ggplot2`

Dans chaque bloc de code, vous êtes libre à créer n'importe quelle structure en utilisant votre code R. Il faut pourtant respecter la linéarité de l'analyse : pour utiliser un objet, il faut le charger *avant* de son utilisation, exactement comme vous faites dans un script R traditionnel.

2.2 Tableaux

Tableau 1: Mon tableau

ville	M	ET
Calgary	67.01	16.63
Montréal	69.58	5.67
Québec	74.17	5.51

Vous pouvez utiliser la fonction `kable()` de l'extension `knitr` pour imprimer un tableau mis en forme dans votre PDF. Consultez le code ci-dessous pour le Tableau 1. Naturellement, vous

pouvez imprimer le tableau brut (tibble) directement. Toutefois, ce type de tableau n'est pas mis en forme pour une publication. Il faut donc l'éviter pour le projet.

3 Analyse statistique

Après avoir exploré vos données, vous devez les analyser de façon appropriée. Pour imprimer le tableau d'un modèle linéaire, on utilise la fonction `tab_model()` de l'extension `sjPlot`. C'est une fonction très utile et déjà discuté en classe. Voici l'exemple pour le fichier `villes2.csv` : on exécute un modèle `note ~ ville`, qui suit naturellement de la Figure 1.

Tableau 2: Résultats du modèle statistique

Predictors	Estimates	note CI	p
(Intercept)	67.01	64.04 – 69.98	<0.001
ville [Montréal]	2.57	-1.63 – 6.78	0.228
ville [Québec]	7.16	2.96 – 11.37	0.001
Observations	150		
R ² / R ² adjusted	0.073 / 0.061		

Il faut maintenant interpréter les coefficients $\hat{\beta}$, naturellement. De plus, vous pouvez ajouter une version plus visuelle des résultats : une figure représentant le modèle. En effet, la figure pourrait *remplacer* le tableau, puisqu'elle affichera à peu près les mêmes informations. On utilise la fonction `plot_model()` de l'extension `sjPlot`.

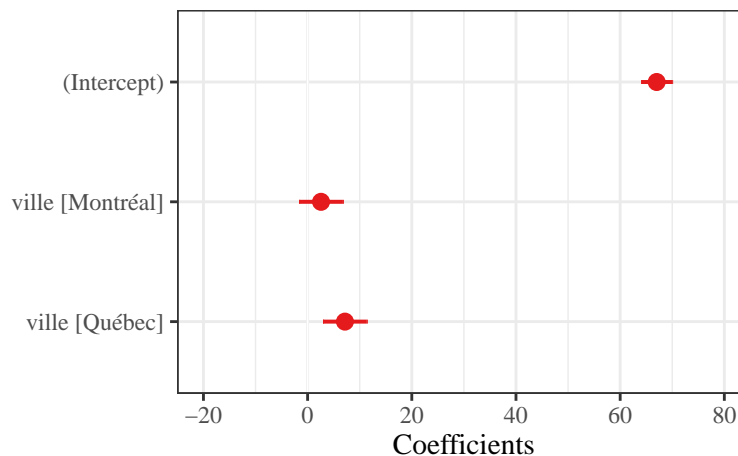


Figure 2: Les coefficients du modèle et leurs IC à 95 %

Veillez noter que la figure ne montre pas le coefficient de détermination R^2 . Vous pouvez donc compléter la figure en ajoutant cette valeur dans votre texte manuellement, par exemple (si vous décidez de ne pas utiliser le tableau). Une alternative serait d'utiliser le code en ligne suivant : `summary(mod)$r.squared |> round(2)`, ce qui renverra 0.07.

Références

BARNIER Julien, 2023, *Introduction à R et au tidyverse*.

Appendices

Code d'exploration des données

```
# Les extensions utilisées :  
library(knitr) # Pour les tableaux  
library(sjPlot) # Pour les modèles statistiques  
library(tidyverse)  
  
# Le fichier des données :  
d <- read_csv("../donnees/villes2.csv")  
  
# Le tableau/résumé des données :  
d |>  
  summarize(M = mean(note),  
            ET = sd(note),  
            .by = ville) |>  
  mutate(across(where(is_double), ~round(.x, 2))) |>  
  kable()
```

Code des figures

```
# Figure des données :  
ggplot(data = d, aes(x = ville, y = note)) +  
  geom_boxplot() +  
  labs(y = "Note",  
       x = "Ville") +  
  theme_bw(base_family = "Times")  
  
# Figure du modèle :  
plot_model(mod, show.intercept = TRUE) +  
  labs(title = NULL,  
       y = "Coefficients",  
       x = NULL) +  
  theme_bw(base_family = "Times")
```

Code de l'analyse

```
# Code du modèle :  
mod <- lm(note ~ ville, data = d)  
tab_model(mod)
```