

TP2_Tritz : Etude de l'évolution des taux de vol de biens du ménage dans des maisons depuis 1999 (option 2)

Apolline Tritz

2024-02-05

```
library(readr)
library(haven)
library(tidyverse)
```

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
v dplyr      1.1.3      v purrr      1.0.2
v forcats    1.0.0      v stringr    1.5.0
v ggplot2    3.4.3      v tibble     3.2.1
v lubridate  1.9.3      v tidyr      1.3.0
-- Conflicts ----- tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()     masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become
```

Question 1 : Question de recherche

Selon Amelia Thatcher, travaillant pour la Gendarmerie Royale du Canada, les infractions contre les biens sont fréquentes dans les villes canadiennes (Thatcher 2015). Au Canada, les introductions par effraction et les vols de véhicules sont des types de crimes les plus représentés dans les données officielles (Thatcher 2015). Il y aurait, selon l'étude effectuée en 2015, une introduction par effraction toutes les trois minutes (Thatcher 2015). Mon analyse porte sur les vols de biens du ménage et non sur les introductions par effraction, mais cela revient presque à la même chose, car la majorité du temps lorsqu'il y a une entrée par effraction dans un logement, elle est suivie d'une tentative ou d'un vol de bien. Il s'agit seulement d'un problème d'appellation pour les statistiques criminelles. Étant donné cette statistique assez

impressionnante, on pourrait se demander si cela a toujours été le cas. C’est pourquoi, ma première analyse a pour but d’être seulement descriptive et non explicative. Cette première partie va répondre à la question de l’évolution des tentatives et des vols de biens du ménage, au sein des maisons depuis 1999.

Les vols de biens du ménage peuvent aller d’un vol de vélo à un vol de biens plus coûteux. Les vols de biens plus coûteux sont généralement bien connus des services de police, car il faut faire une déclaration du vol à la police pour commencer les démarches avec les assurances. La majorité des personnes, pour un vol de faible valeur, préfère ne pas perdre du temps dans les démarches administratives et remplacer le bien volé. C’est pourquoi, afin d’avoir une évolution la plus ressemblante possible à la réalité, j’ai choisi les Enquêtes Sociales Générales (ESG), sondages de victimisation. Les avantages des Enquêtes Sociales Générales (ESG) sont qu’on possède des informations qui ne sont pas forcément signalés à la police. De plus, elles sont réalisées à intervalle régulières, ce qui nous permet de faire des analyses temporelles. Un des inconvénients qu’il faut prendre en compte est que le répondant va peut-être communiquer une information sur infraction qu’il a subi, alors qu’il a subi cette infraction à l’extérieur de la période étudiée. C’est pourquoi, pour la base de données de 1999, j’ai sélectionné une variable nommée « V1A », afin de ne garder que les observations qui correspondaient à une infraction commise dans l’année précédant la collecte de données. La première base de données, celle de 1999, comporte seulement les réponses pour les dix provinces ; tandis que les deux autres, celles de 2009 et 2019, comportent les réponses des dix provinces et des territoires.

Les variables qui m’ont intéressée pour cette étude sont THFTHP (vol de biens du ménage); A_THFTHP (tentative de vol de biens du ménage); DWELINCC, DWELINC_C et WHR_100 (lieux où se sont déroulés l’infraction) et V1A (l’infraction s’est déroulée dans les 12 mois précédant la collecte (seulement pour l’ESG de 1999)).

Question 2 : Données et méthodes

Dans un premier temps, j’ai importé mes 3 bases de données trouvées sur Odesi, dans R. Ces bases de données sont les résultats aux Enquêtes de Sociales Généralisées (ESG) sur la victimisation des années 1999, 2009 et 2019. J’avais seulement pour l’année 1999, un fichier .csv, donc j’ai décidé de tous les télécharger en fichier .sav. J’ai importé ces bases de données grâce à la fonction “read_sav” et j’ai enregistré chacune des bases de données dans son objet “ESG_date” correspondant. Une fois cela effectué, j’ai dû convertir ces fichiers .sav en fichiers .csv, grâce à la fonction “write.csv” et je les ai enregistrées dans un nouvel objet, “ESG_date”.

J’ai pu commencer mon travail de nettoyage de ma base de données. Pour ce faire, pour chaque base de données, j’ai sélectionné les variables qui me seront utiles pour la suite de l’analyse. Pour chacune des bases de données, je les ai enregistrées dans un nouvel objet “ESG_date_1”. Dans l’objet “ESG_1999_1”, j’ai sélectionné les variables : THFTHP, DWELINCC, V1A, A_THFTHP. Dans l’objet “ESG_2009_1”, j’ai sélectionné les variables

: THFTHP, DWELINC_C et A_THFTHP. Dans l'objet ESG_2019_1, j'ai sélectionné les variables : THFTHP, A_THFTHP et WHR_100.

Pour les 3 bases de données différentes, dans un nouvel objet ESG_date_2, j'ai créé une nouvelle variable "Vol_ménage_maison" où chaque observation qui était égale à 1 (1= il y a eu un vol de biens du ménage) dans la variable THFTHP prenait la valeur 1. Si l'observation n'était pas égale à 1, elle prenait la valeur de 0. J'ai fait la même manipulation pour les 3 bases de données différentes, avec la création de la variable "Tentative_vol_ménage_maison" où chaque observation qui était égale à 1 (1= il y a eu une tentative de vol de biens du ménage) dans la variable A_THFTHP prenait la valeur 1. Si l'observation n'était pas égale à 1, elle prenait la valeur de 0.

Dans les trois bases de données ESG_date_3, j'ai filtré les valeurs des variables correspondants aux lieux de l'infraction, pour ne conserver seulement les infractions ayant eu lieu dans une maison, qu'elle soit individuelle, attachée ou familiale pour la base de données correspondant à l'année 1999. J'ai aussi filtré pour la base de données de 1999, les observations égales à 1 et 5 de la variable V1A, pour seulement ne retenir que les observations qui ont été commises dans l'année précédant la récolte de données. J'ai par la suite enlevé dans les 3 bases de données, les observations correspondants aux données non disponibles (NA). Ensuite, j'ai créé une nouvelle variable VOL, où j'ai additionné les résultats des deux nouvelles variables créées précédemment ("Vol_ménage_maison" et "Tentative_vol_ménage_maison"). J'ai rendu ces deux variables numériques, par conséquent la nouvelle variable "Vol" est numérique et contient seulement les données où l'infraction s'est déroulée dans une maison. Dans les nouveaux objets ESG_date_moyenne, je fais les moyennes de la variable VOL. Par la suite, je crée un data frame où je rentre manuellement les moyennes trouvées pour chaque année, soit 1999, 2009, 2019, afin de rassembler dans un même objet mes trois moyennes provenant des 3 bases de données différentes. Cela me permettra donc de représenter visuellement mes résultats.

Question 3

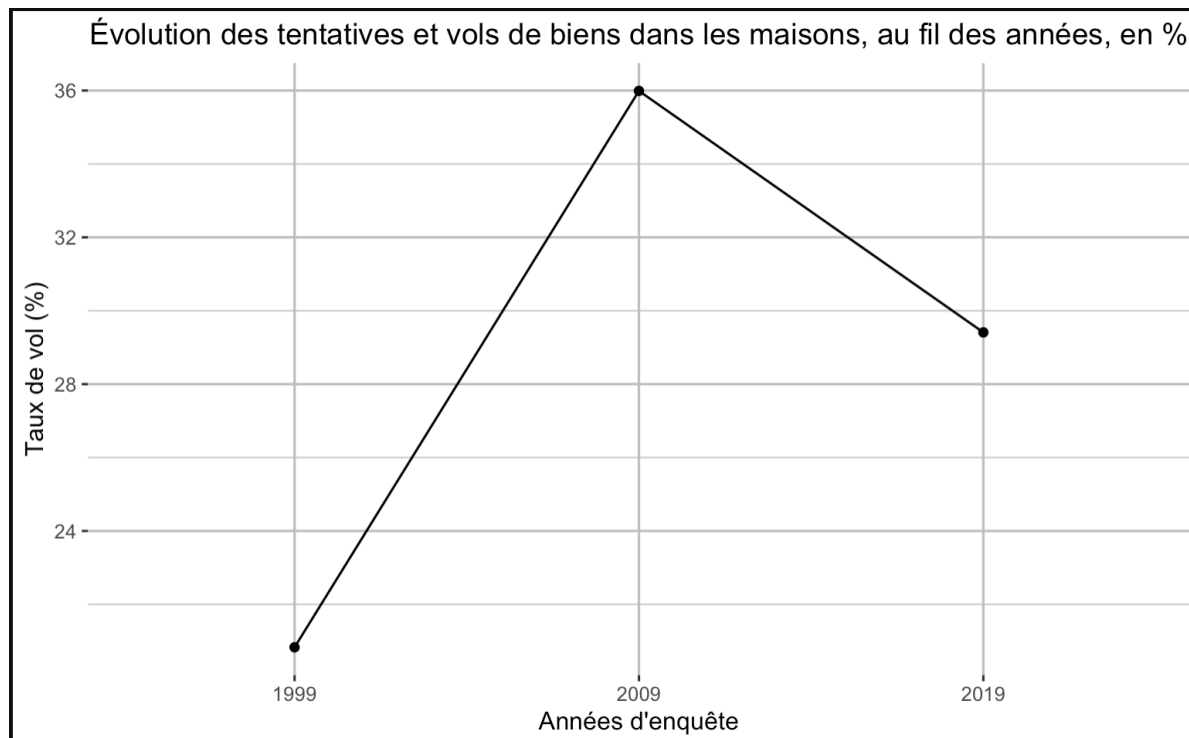


Figure 1: Graphique présentant l'évolution des tentatives et des vols des biens du ménage dans une maison, de 1999 à 2019

La base de données de 1999, une fois nettoyée comporte seulement 48 données, comparative-ment aux deux autres qui comptent 3331 données (2009) et 5340 (2019). À la lumière de cette observation, il est légitime de remettre en question la fiabilité du taux de tentatives/vols de biens du ménage pour l'année 1999. La faiblesse du nombre d'observations retenues peut s'expliquer par l'évolution ultérieure des technologies facilitant la collecte de données, qui n'étaient peut-être pas aussi avancées à l'époque.

Toutefois, grâce à cette étude descriptive, nous pouvons constater qu'une augmentation des tentatives/vols des biens dans une maison, a eu lieu entre 1999 (20,83%) et 2009 (35,99%) même si cette dernière reste à relativiser. A partir de 2009, on remarque une baisse de 6,6% dans les tentatives/vols de biens du ménage, ce qui amène à un taux de 29,41% en 2019.

Pour la continuité de cette étude, il revêt une importance cruciale de démêler les facteurs qui ont contribué à cette diminution. Comprendre ces éléments permettra d'identifier des leviers susceptibles d'être exploités pour maintenir cette tendance positive. L'exploration des divers

facteurs significatifs, quant à la diminution des vols de biens du ménage, sera essentielle pour élaborer des stratégies efficaces en vue de renforcer la sécurité des biens du ménage.

Bibliographie

Thatcher, Amelia. « Crimes contre les biens | Gendarmerie royale du Canada », 15 octobre 2015. <https://www.rcmp-grc.gc.ca/fr/gazette/crimes-contre-les-biens>.

Annexe code

```
ESG_1999 <- read_sav("/Users/apolline/Desktop/FAS1001/fas_1001_Tritz/_tp/_tp2/ESG_1999.sav")
ESG_2009 <- read_sav("/Users/apolline/Desktop/FAS1001/fas_1001_Tritz/_tp/_tp2/ESG_2009.sav")
ESG_2019 <- read_sav("/Users/apolline/Desktop/FAS1001/fas_1001_Tritz/_tp/_tp2/ESG_2019")
```

```
write.csv(ESG_1999, "/Users/apolline/Desktop/FAS1001/fas_1001_Tritz/_tp/_tp2/ESG_1999.csv", )
write.csv(ESG_2009, "/Users/apolline/Desktop/FAS1001/fas_1001_Tritz/_tp/_tp2/ESG_2009.csv", )
write.csv(ESG_2019, "/Users/apolline/Desktop/FAS1001/fas_1001_Tritz/_tp/_tp2/ESG_2019.csv", )
```

```
ESG_1999_ <- read_csv("/Users/apolline/Desktop/FAS1001/fas_1001_Tritz/_tp/_tp2/ESG_1999.csv")
```

```
Rows: 10087 Columns: 217
```

```
-- Column specification -----
```

```
Delimiter: ","
```

```
chr (1): VBSCRNO
```

```
dbl (216): RECID, V1, V1A, V2, DWELINCC, V5, V6, V7, V8, V8A_C01, V8A_C02, V...
```

```
i Use `spec()` to retrieve the full column specification for this data.
```

```
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
ESG_2009_ <- read_csv("/Users/apolline/Desktop/FAS1001/fas_1001_Tritz/_tp/_tp2/ESG_2009.csv")
```

```
Rows: 7096 Columns: 224
```

```
-- Column specification -----
```

```
Delimiter: ","
```

```
chr (1): SCREEN_INC_TYPE
```

```
dbl (223): RECID, CIR_D010, WHR_Q100, WHR_Q110, DWELINC_C, WHR_Q120, WHR_Q13...
```

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.

```
ESG_2019_ <- read_csv("/Users/apolline/Desktop/FAS1001/fas_1001_Tritz/_tp/_tp2/ESG_2019.csv")
```

Rows: 5522 Columns: 567

```
-- Column specification -----
Delimiter: ","
chr   (1): SCR_N_INC
dbl   (64): PUMFID, CIR_D010, WHR_100, PCA_20, PCA_40, CIR_090, CIR_100, CIR_...
lg1   (502): WGHT_VIC, ADJWTVIC, WVCBS001, WVCBS002, WVCBS003, WVCBS004, WVCBS...
```

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.

```
ESG_1999_1 <- ESG_1999_ |> select(THFTHP, DWELINCC, V1A, A_THFTHP)
ESG_2009_1 <- ESG_2009_ |> select(THFTHP, DWELINCC, A_THFTHP)
ESG_2019_1 <- ESG_2019_ |> select(THFTHP, A_THFTHP, WHR_100)
```

```
ESG_1999_2 <- ESG_1999_1 |> mutate(Vol_ménage_maison = ifelse(THFTHP == 1, "1", "0"))
ESG_2009_2 <- ESG_2009_1 |> mutate(Vol_ménage_maison = ifelse(THFTHP == 1, "1", "0"))
ESG_2019_2 <- ESG_2019_1 |> mutate(Vol_ménage_maison = ifelse(THFTHP == 1, "1", "0"))
```

```
ESG_1999_3 <- ESG_1999_2 |> mutate(Tentative_vol_ménage_maison = ifelse(A_THFTHP == 1, "1", "0"))
ESG_2009_3 <- ESG_2009_2 |> mutate(Tentative_vol_ménage_maison = ifelse(A_THFTHP == 1, "1", "0"))
ESG_2019_3 <- ESG_2019_2 |> mutate(Tentative_vol_ménage_maison = ifelse(A_THFTHP == 1, "1", "0"))
```

```
ESG_1999_3 |> filter(DWELINCC == 1 | DWELINCC == 2 | DWELINCC == 3)
```

A tibble: 821 x 6

	THFTHP	DWELINCC	V1A	A_THFTHP	Vol_ménage_maison	Tentative_vol_ménage_maison
	<dbl>	<dbl>	<dbl>	<dbl>	<chr>	<chr>
1	2	1	NA	2	0	0
2	2	1	NA	2	0	0
3	1	1	NA	2	1	0
4	2	1	NA	2	0	0
5	2	1	NA	2	0	0
6	2	1	1	2	0	0
7	2	1	NA	2	0	0
8	NA	1	NA	NA	<NA>	<NA>

```

  9      1      1    NA      2 1      0
10      2      2    NA      2 0      0
# i 811 more rows

```

```
ESG_2009_3 |> filter(DWELINC_C == 1)
```

```

# A tibble: 2,213 x 5
  THFTHP DWELINC_C A_THFTHP Vol_ménage_maison Tentative_vol_ménage_maison
  <dbl>   <dbl>   <dbl> <chr>           <chr>
1      2      1      2 0      0
2      1      1      2 1      0
3      1      1      2 1      0
4      2      1      2 0      0
5      2      1      2 0      0
6      1      1      2 1      0
7      1      1      2 1      0
8      2      1      1 0      1
9      2      1      2 0      0
10     2      1      2 0      0
# i 2,203 more rows

```

```
ESG_2019_3 |> filter(WHR_100 == 1)
```

```

# A tibble: 3,314 x 5
  THFTHP A_THFTHP WHR_100 Vol_ménage_maison Tentative_vol_ménage_maison
  <dbl>   <dbl>   <dbl> <chr>           <chr>
1      2      2      1 0      0
2      2      2      1 0      0
3      2      2      1 0      0
4      2      2      1 0      0
5      2      2      1 0      0
6      2      2      1 0      0
7      1      2      1 1      0
8      1      2      1 1      0
9      1      2      1 1      0
10     1      2      1 1      0
# i 3,304 more rows

```

```
ESG_1999_3 |> filter(V1A == 1 | V1A == 5)
```

```
# A tibble: 654 x 6
  THFTHP DWELINCC V1A A_THFTHP Vol_ménage_maison Tentative_vol_ménage_maison
    <dbl>    <dbl> <dbl>    <dbl> <chr>                <chr>
1     NA      NA     5      NA <NA>                <NA>
2     NA      NA     5      NA <NA>                <NA>
3     NA      NA     5      NA <NA>                <NA>
4     NA      NA     5      NA <NA>                <NA>
5     NA      NA     1      NA <NA>                <NA>
6      2      1     1      2 0                    0
7     NA      NA     5      NA <NA>                <NA>
8     NA      NA     1      NA <NA>                <NA>
9      2      NA     1      2 0                    0
10    NA      NA     5      NA <NA>                <NA>
# i 644 more rows
```

```
ESG_1999_3_sansNA <- na.omit(ESG_1999_3)
ESG_2009_3_sansNA <- na.omit(ESG_2009_3)
ESG_2019_3_sansNA <- na.omit(ESG_2019_3)
```

```
ESG_1999_moyenne <- ESG_1999_3_sansNA |> mutate(VOL = as.numeric(Vol_ménage_maison) + as.numeric(Tentative_vol_ménage_maison))
ESG_2009_moyenne <- ESG_2009_3_sansNA |> mutate(VOL = as.numeric(Vol_ménage_maison) + as.numeric(Tentative_vol_ménage_maison))
ESG_2019_moyenne <- ESG_2019_3_sansNA |> mutate(Vol_ménage_maison = as.numeric(Vol_ménage_maison) + as.numeric(Tentative_vol_ménage_maison))
```

```
ESG_1999_moyenne |> summarise(mean_vol = mean(VOL, na.rm = TRUE))
```

```
# A tibble: 1 x 1
  mean_vol
    <dbl>
1    0.208
```

```
ESG_2009_moyenne |> summarise(mean_vol = mean(VOL, na.rm = TRUE))
```

```
# A tibble: 1 x 1
  mean_vol
    <dbl>
1    0.360
```



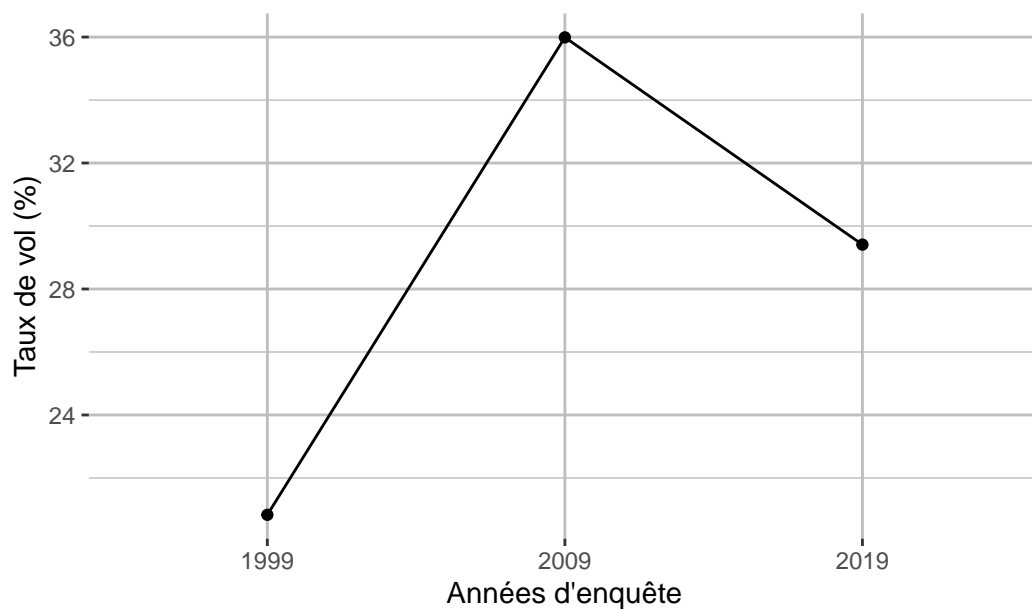
```
ESG_2019_moyenne |> summarise(mean_vol = mean(Vol_ménage_maison, na.rm = TRUE))
```

```
# A tibble: 1 x 1
  mean_vol
  <dbl>
1    0.294
```

```
Vol_Final <- data.frame(
  Année = c(1999, 2009, 2019),
  Taux_vol = c(20.83, 35.99, 29.41),
  Groupe = "Total")
```

```
ggplot(data = Vol_Final,
       mapping = aes(x= as.factor(Année) , y= Taux_vol, group = Groupe)) + geom_point() + geom_line() +
scale_x_discrete(labels = c("1999", "2009", "2019")) +
labs(x = "Années d'enquête", y = "Taux de vol (%)") +
ggtitle("Évolution des tentatives et vols de biens dans les maisons, au fil des années, en France") +
theme(plot.title = element_text(hjust = 0.5), panel.background = element_rect(fill = "white"))
```

Évolution des tentatives et vols de biens dans les maisons, au fil des années, en France



```
nrow(ESG_1999_3_sansNA)
```

```
[1] 48
```

```
nrow(ESG_2009_3_sansNA)
```

```
[1] 3331
```

```
nrow(ESG_2019_3_sansNA)
```

```
[1] 5340
```