

# Rapport de projet

## Data Science

Luca Debeir, Simon Gayet, Alexandre Girbal, David Meier

Lien de l'application shiny : [Shiny Data Science - Sujet 7](#)

Lien du github : [Github - Data Science](#)

Lien des slides de la soutenance : [Soutenance - Data Science](#)

# Table des matières

<b>Table des Figures</b>	<b>2</b>
<b>Introduction</b>	<b>3</b>
<b>Utilisation par rapport à la fréquence</b>	<b>4</b>
Problématique	4
Données	4
Analyse	5
Hypothèse n°1:	5
Hypothèse n°2:	7
Hypothèse n°3:	8
Hypothèses n°4 et 5:	8
Conclusion du premier axe	13
<b>Utilisation par rapport aux usages</b>	<b>14</b>
Problématique	14
Données	14
Analyse	15
Hypothèse n°1:	15
Hypothèse n°2:	17
Hypothèse n°3:	18
Données	18
Analyse	18
Conclusion du second axe	20
<b>Conclusion</b>	<b>21</b>
<b>Annexes</b>	<b>22</b>
Annexe 1:	22
Annexe 2:	31
Annexe 3:	38
Annexe 4 : Server Shiny App	46
Annexe 5 : UI Shiny App	52

## Table des Figures

- Figure 1 : Diagramme sur R, répartition du nombre d'autos
- Figure 2 : Nombre de km parcourus en moto en fonction du km parcourus en voiture
- Figure 3 : Tableau de contingence de la fréquence en moto et en auto
- Figure 4 : Tableau de contingence de la fréquence en fonction du véhicule
- Figure 5 : Part de l'utilisation des véhicules par les utilisateurs
- Figure 6 : ACM par rapport à la fréquence, à la situation familiale et à l'assurance
- Figure 7 : Moyenne de km parcourus en moto en fonction de la situation familiale
- Figure 8 : Moyenne de km parcourus en auto en fonction de la situation familiale
- Figure 9 : Moyenne de km parcourus en moto et en auto en fonction de la situation maritale
- Figure 10 : Moyenne de km parcourus en auto en fonction de la situation maritale
- Figure 11 : Moyenne de km parcourus en moto en fonction de la situation maritale
- Figure 12 : Répartition des usages pour les motos et les automobiles
- Figure 13 : Répartition des usages pour les automobiles
- Figure 14 : Répartition des usages pour les motos
- Figure 15 : AFC sur les variables Véhicule et Usage
- Figure 16 : Tableau de contingence entre les variables fréquence automobile et usage
- Figure 17 : Tableau de contingence entre les variables fréquence moto et usage

## Introduction

Dans le cadre du projet data science, nous avons analysé les données fournies par la Mutuelle des Motards, une société proposant assurances et mutuelles aux conducteurs de deux-roues en France. Le sujet qui nous a été proposé est le suivant:

### **Sujet 7 : Comparer l'utilisation de la voiture et celle du 2-roues. Est-ce que les gros rouleurs en voiture sont aussi les gros rouleurs en 2-roues ?**

Pour répondre à cette problématique nous disposons d'un questionnaire sur les motos et un autre sur les autos. Nous avons croisé les données pour ne garder que les utilisateurs ayant répondu à ces deux questionnaires pour pouvoir comparer l'utilisation de ces deux véhicules. Parmi ces individus 36.5 % possèdent une auto, c'est pourquoi il semblait intéressant de comparer l'utilisation de la voiture à celle des deux roues sur cet échantillon selon les 2 axes d'études suivants:

- *L'utilisation par rapport à la fréquence*
- *L'utilisation par rapport aux usages*

Afin de traiter nos axes, nous avons dû définir les gros rouleurs comme étant des individus parcourant plus de kilomètres qu'un kilométrage annuel arbitraire. Nous avons considéré que ce kilométrage serait la moyenne française des km parcourus tous véhicules confondus.

Ainsi nous savons avant de commencer un traitement des données approfondi que 18,45% des personnes ayant répondu aux questionnaires et ayant une voiture et une moto sont considérées comme des gros rouleurs en moto contre 33.11% qui sont considérées comme des gros rouleurs en voiture.

# 1. Utilisation par rapport à la fréquence

## Problématique

Les individus considérés comme des gros rouleurs seront les personnes roulant plus de **17 000 km/an**, ce qui est la moyenne française en 2015 d'après l'Argus.fr. Nous avons choisi une source externe pour établir ce que l'on considère comme un gros rouleur car l'échantillon des données n'est pas représentatif de la population étant donnée que le questionnaire est rempli sur la base du volontariat parmi les personnes inscrites à la mutuelle des motards. D'autre part les personnes roulant moins de 17 000 km/an seront à l'inverse considéré comme petits rouleurs.

Nous souhaitons alors comparer l'utilisation de la voiture et de la moto par rapport à cette séparation faite sur l'échantillon étudié.

Nous avons formulé cinq hypothèses pour traiter cet axe:

- H1: *"Il y a une corrélation négative entre la fréquence d'utilisation de la moto et celle de la voiture."*
- H2: *"Les répondants ayant une voiture utilisent plus leur voiture que leur moto."*
- H3: *"Utilisation de la moto et de la voiture en corrélation le sexe."*
- H4: *"Les parents célibataires sont plus aptes à conduire une moto."*
- H5: *"Plus il y a d'enfant à charge plus c'est un gros rouleur en voiture."*

Nous commencerons par sélectionner les données à analyser afin de répondre à nos hypothèses.

## Données

Les données proviennent de réponse à des questionnaires envoyé par la mutuelle des motards à leurs abonnés dont un axé sur les motos et un autre sur les voitures. Nous croisons les données pour ne garder que les utilisateurs ayant répondu aux deux questionnaires. Nous traitons dans cette partie les km parcourus (c'est-à-dire la fréquence d'utilisation de chaque véhicule) en fonction de différents paramètres.

Le questionnaire des motos permet de renseigner jusqu'à 6 cylindrées et pour chacune d'entre elles de renseigner le nombre de km parcourus par an. Nous avons donc effectué une somme de ces kilométrages pour chaque individu afin d'obtenir une donnée "nombre total de km parcourus à moto". D'autre part le questionnaire des autos permet de renseigner le nombre de voitures mais demande le kilométrage seulement pour la première voiture la plus utilisée. Nous avons donc utilisé cette donnée en sachant que cela fausse probablement les résultats pour certains individus.

De plus grâce aux questionnaires nous connaissons certaines informations personnelles de chaque individu ( le nombre d'enfants, la situation familiale et maritale ...), qui sont utilisées pour répondre aux hypothèses émises précédemment.

Les données brutes sont disponibles à l'adresse suivante:

<https://imag.umontpellier.fr/~toulemonde/filesIG4/ProjetDataScience/>

## Analyse

### Hypothèse n°1:

Afin de comparer les fréquences d'utilisation de ces véhicules, nous sommes partis de l'hypothèse qu'il y a une corrélation négative entre la fréquence d'utilisation de la moto et celle de la voiture.

Nous avons comme données, les km/an de chaque 2-roues possédé par les individus; nous les avons donc sommés pour obtenir le nombre de kilomètres parcouru en un an par une personne avec toutes ses motos confondus. Cependant, nous possédons le nombre de km/an seulement de la voiture principale de chaque individu. Or comme on peut le voir ci-dessous les répondant possèdent dans près de **55%** des cas plus d'une seule voiture.

C'est pour cette raison que nos prochaines analyses concernant le nombre de kilomètres parcouru par an en voiture seront certainement à modérer.

Répartition du nombre d'autos

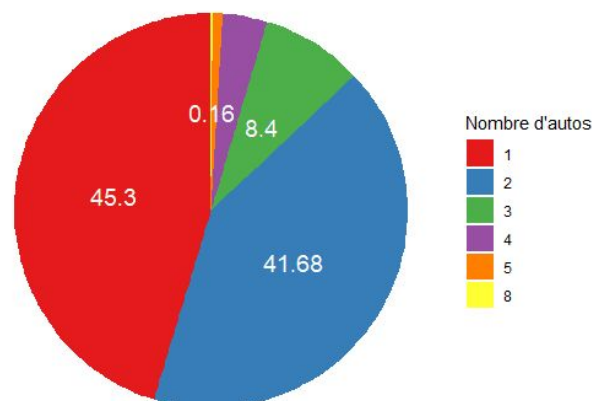


Figure 1: Diagramme R, répartition du nombre d'autos

Nous avons donc réalisé un graphe comparant la fréquence d'utilisation de la voiture et de la moto avec en abscisse le nombre de kilomètres/an avec la voiture principale et en ordonnée le nombre de kilomètres/an toutes motos confondu.

Nous avons vu qu'il y avait certaines données incohérentes, c'est pourquoi nous avons réduit l'intervalle des km/an de 0 à 100 000.

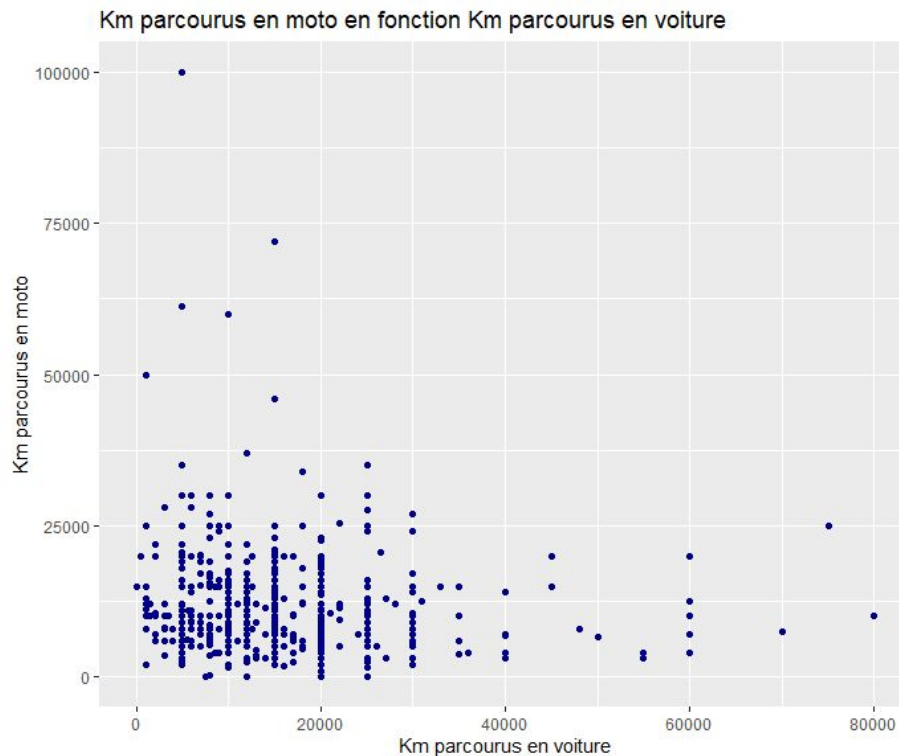


Figure 2 : Nombre de km parcourus en moto en fonction du km parcourus en voiture

On remarque la tendance suivante : si on roule beaucoup avec un véhicule on roule moins avec l'autre. On ne peut cependant pas déterminer qu'il y a une corrélation négative entre la fréquence d'utilisation de la moto et celle de la voiture.

Nous avons retraité les données correspondant aux kilomètres par an dans le but d'obtenir les variables qualitatives **rouleur moto** et **rouleur auto** définies précédemment.

Nous avons réalisé un test du chi-2 sur ces variables.

	GrosRouleurAuto	PetitRouleurAuto
GrosRouleurMoto	19	44
PetitRouleurMoto	177	367

Figure 3 : Tableau de contingence de la fréquence en moto et en auto

```
> chisq.test(tabContingenceRouleur)
```

```
Pearson's Chi-squared test with Yates' continuity correction
```

```
data: tabContingenceRouleur
X-squared = 0.057523, df = 1, p-value = 0.8105
```

Le test ne nous permet pas de rejeter l'hypothèse d'indépendance car : la valeur de X-squared est inférieure à 3,84.

### Hypothèse n°2:

On se demande donc quel est le moyen de transport le plus utilisé parmi notre panel réduit. Afin de déterminer s'il existe un lien entre la fréquence d'utilisation et le véhicule, nous avons ci-dessous le tableau de contingence qui lie les données entre voiture et moto, à partir du nombre de kilomètres parcourus sur une année avec le véhicule principal.

	Auto	Moto
Gros rouleur	196	63
Petit rouleur	411	544

Figure 4 : Tableau de contingence de la fréquence en fonction du véhicule

Le test du khi-2 sur ce tableau nous donne le résultat suivant :

```
> chisq.test(tabContingenceRouleur2)
```

```
Pearson's Chi-squared test with Yates' continuity correction
```

```
data: tabContingenceRouleur2
```

```
X-squared = 85.519, df = 1, p-value < 2.2e-16
```

Le test nous permet de rejeter l'hypothèse d'indépendance car : la valeur de X-squared > 3,84. Donc on a mis en évidence le lien entre la fréquence d'utilisation et les véhicules, mais nous ne pouvons pas en conclure plus avec cette analyse.



Nous pouvons voir sur le graphe ci-dessous qu'effectivement la voiture semble plus utilisée que la moto et ce malgré le fait que l'on connaisse la fréquence d'utilisation seulement de la voiture principale.

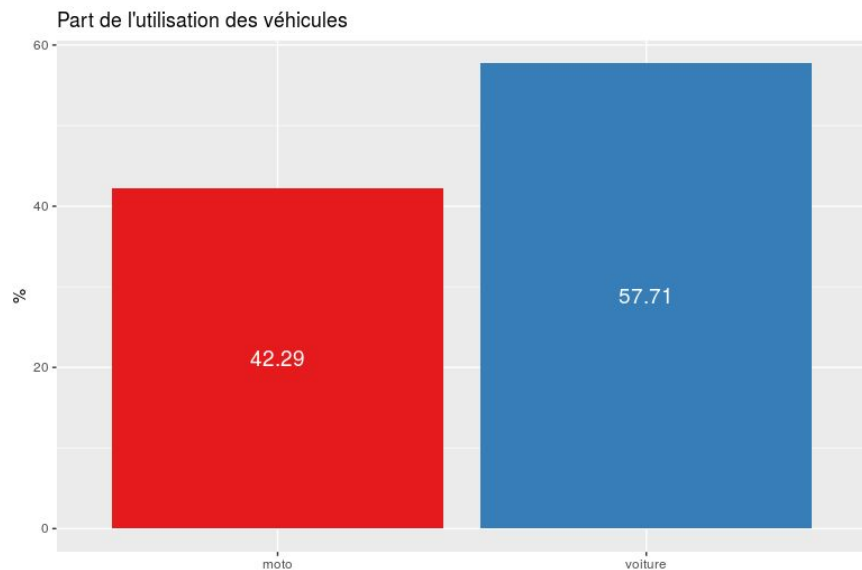


Figure 5 : Part de l'utilisation des véhicules par les utilisateurs

### Hypothèse n°3:

Nous avons par la suite cherché à vérifier l'hypothèse suivante:

*“L'utilisation de la moto et de la voiture est en corrélation avec le sexe.”*

Mais malheureusement il reste seulement une trentaine de femme dans notre jeu de données sur 607 individus. Ces chiffres nous semblant peu significatif nous avons décidé de ne pas traiter cette hypothèse.

### Hypothèses n°4 et 5:

Dans cette partie nous allons analyser les données de manière à répondre à ces hypothèses:

*“Les parents célibataires sont plus aptes à conduire une moto qu'une voiture.”*

*“Plus on a d'enfants à charge plus on est un gros rouleur en voiture.”*

Nous avons traité les données pour faire ressortir les deux variables qualitatives suivantes : **situation familiale, avec ou sans enfant**.

Nous nous sommes également questionné sur le lien entre le fait qu'un répondant soit assuré à la mutuelle des motards ou non.

Dans le but de voir les liens entre ces variables et nos variables précédemment définies : **rouleur moto** et **rouleur voiture** nous avons pensé à réaliser une ACM car c'est

la méthode adaptée aux tableaux de données où  $p$  variables qualitatives décrivent un ensemble d'individus.

Pour bien comprendre le résultat de l'ACM ci-dessous voici l'explication des variables:

**Assure\_mutuelle\_1** : Les individus assurés à la mutuelle des motards.

**Assure\_mutuelle\_0** : Les individus qui ne sont pas assurés à la mutuelle des motards.

**en\_couple\_1** : les individus qui sont en couples.

**en\_couple\_0** : les individus célibataires.

**avec\_enfant\_1** : les individus avec des enfants.

**avec\_enfant\_0** : les individus sans enfants.

**gros\_rouleur\_auto\_1** : Les individus qui roule plus de 17 000 Km/an avec leur voiture principale.

**gros\_rouleur\_auto\_0** : Les individus qui roule moins de 17 000 Km/an avec leur voiture principale.

**gros\_rouleur\_moto\_1** : Les individus qui roule plus de 17 000 Km/an avec toutes leurs motos confondues.

**gros\_rouleur\_moto\_0** : Les individus qui roule moins de 17 000 Km/an avec toutes leurs motos confondues.

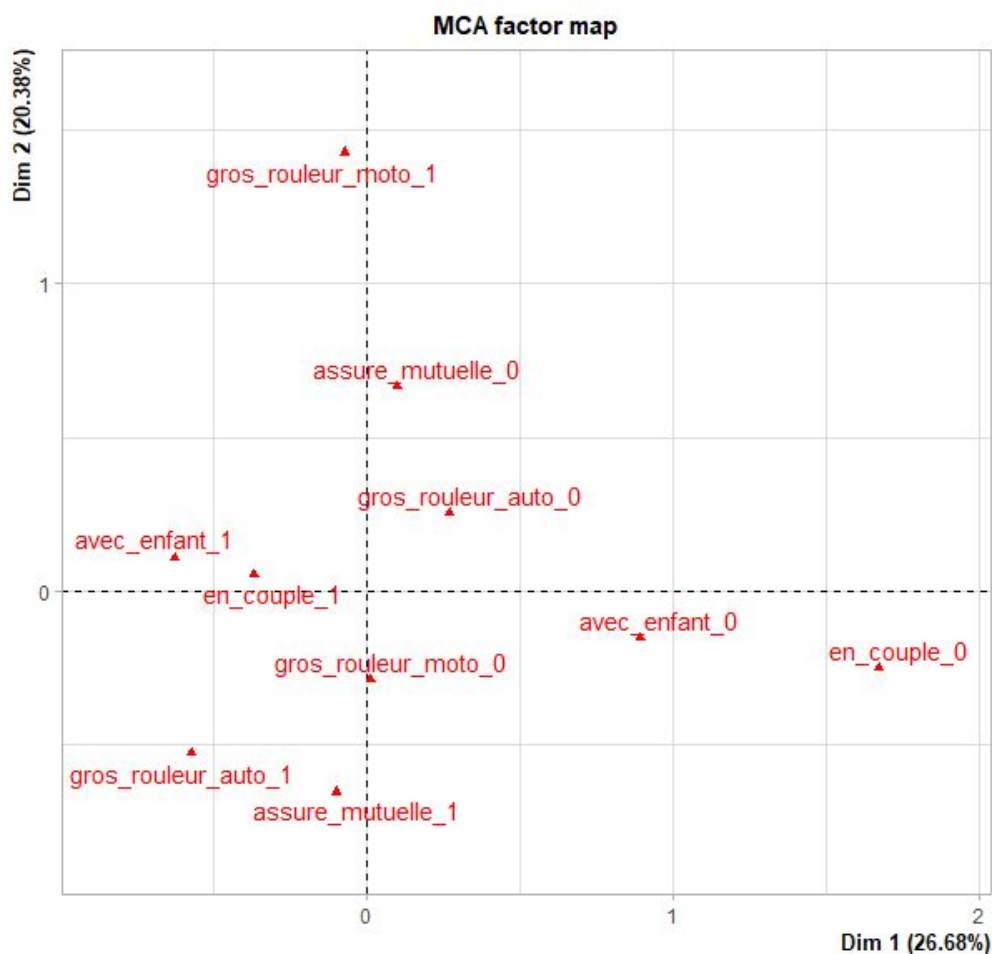


Figure 6 : ACM par rapport à la fréquence, à la situation familiale, à la situation maritale et à l'assurance

Nous voyons sur cette ACM que les gros rouleurs en autos semblent être assurés à la mutuelle des motards. Mais sans clusters évidents, nous n'avons pas su tirer d'informations claires et répondant à nos autres hypothèses.

On a donc réalisé des diagrammes que nous expliquerons par la suite pour vérifier nos deux hypothèses.

Tout d'abord nous avons cherché à savoir si une famille roulait plus suivant le nombre d'enfants à charge. Comme on peut le voir sur la figure 7 ci-dessous, cela semble être le cas pour les motos.

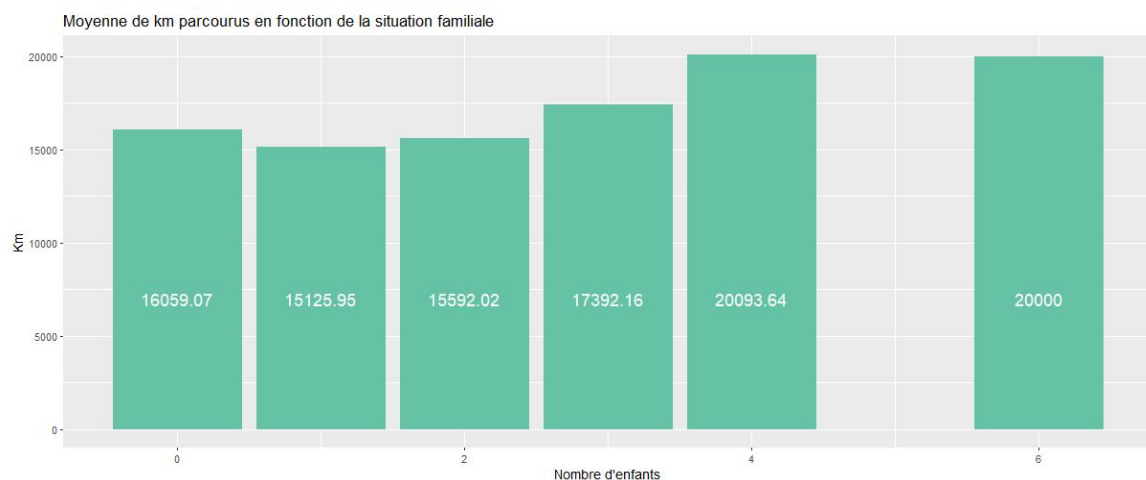


Figure 7 : Moyenne de km parcourus en moto en fonction de la situation familiale

Par contre contrairement à notre hypothèse on remarque que les familles n'utilisent pas plus la voiture lorsqu'ils ont plus d'enfants à charge. Cependant même si nous utilisons la moyenne des kilomètres parcourus pour analyser ces données, il faut cependant préciser que notre population est trop petite pour confirmer ou non nos hypothèses. En effet il y a une seule famille à 6 enfants sur 607 individus considérés.



Figure 8 : Moyenne de km parcourus en auto en fonction de la situation familiale

On rappelle la seconde hypothèse que nous cherchons à vérifier dans cette partie :

*“Les parents célibataires sont plus aptes à conduire une moto qu’une voiture.”*

Pour y répondre nous avons réalisé trois diagrammes montrant la situation maritale respectivement le nombre de kilomètres parcourus en auto et à moto puis ceux parcourus en auto et enfin ceux parcourus à moto.

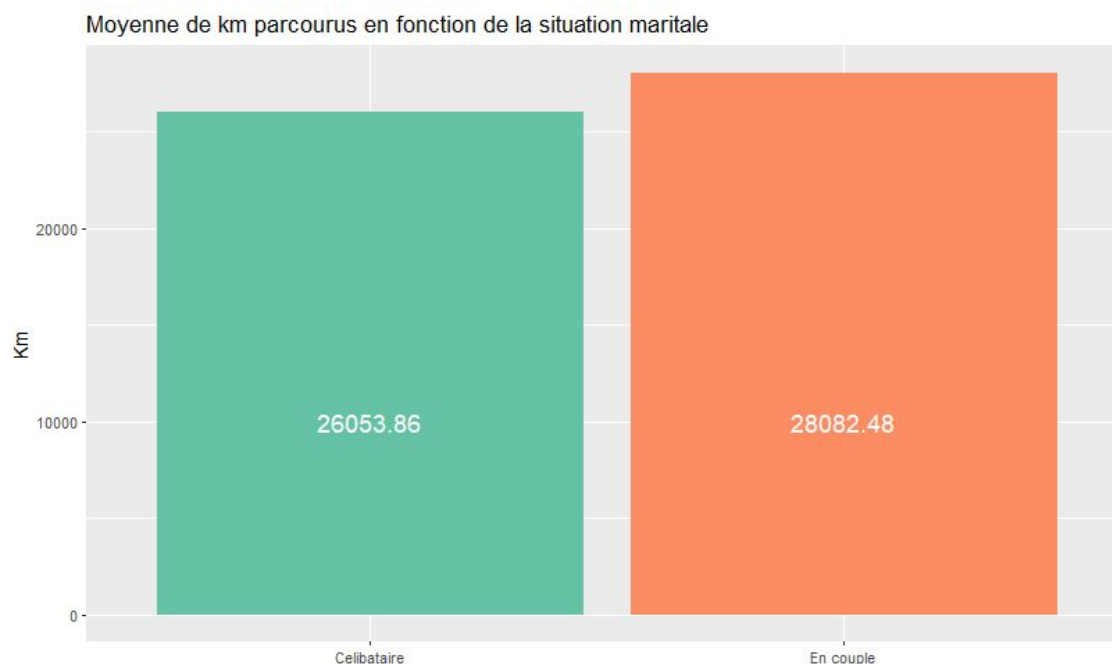


Figure 9 : Moyenne de km parcourus en moto et en auto en fonction de la situation maritale

Comme nous le voyons sur ce graphe, les individus en couple semblent rouler en moyenne plus que les individus célibataires.

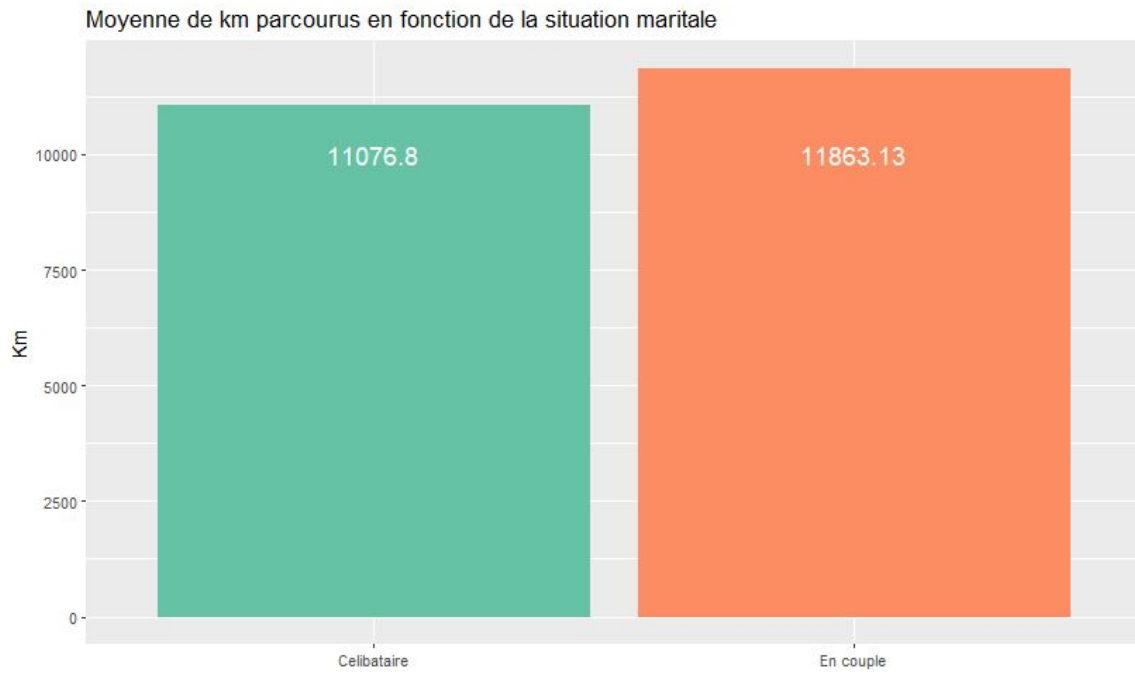


Figure 10 : Moyenne de km parcourus en auto en fonction de la situation maritale

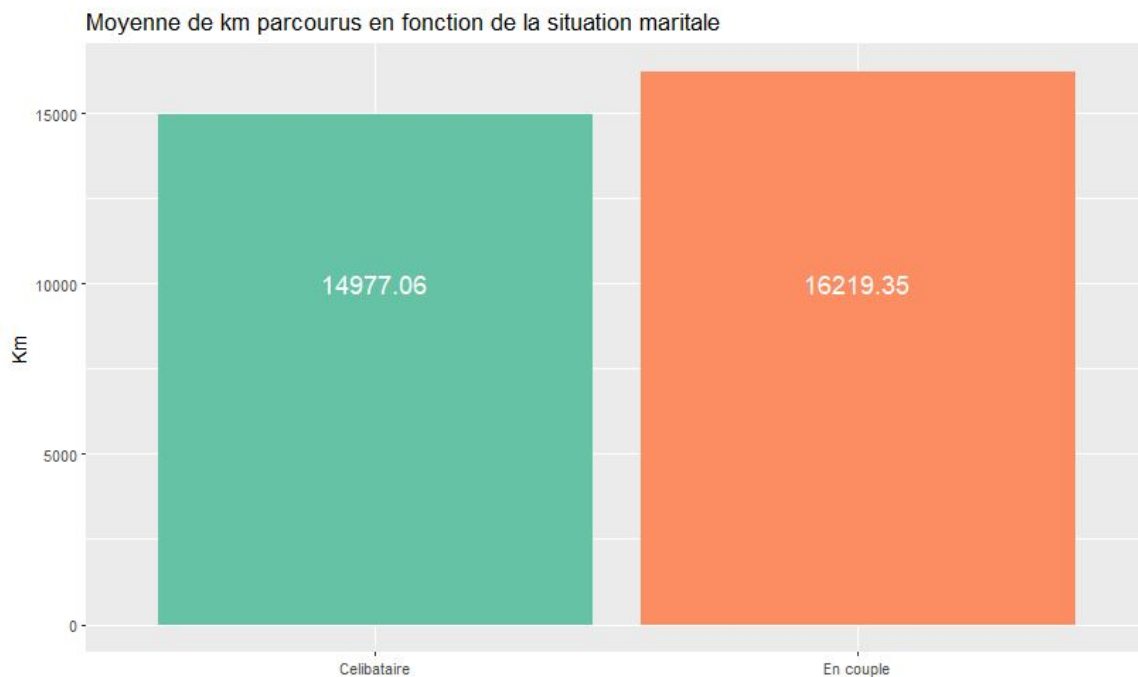


Figure 11 : Moyenne de km parcourus en moto en fonction de la situation maritale

On constate à l'aide des figures 10 et 11 que les individus roulent environ le même nombre de kilomètres par an en auto contrairement au fait que la moto semble être plus utilisée par les couples ce qui contredit notre hypothèse à savoir : *“Les parents célibataires sont plus aptes à conduire une moto qu’une voiture.”*

## Conclusion du premier axe

Dans cette partie nous avons cherché à comparer l'utilisation de l'auto et des 2-roues suivant la fréquence d'utilisation .

L'analyse faite pour répondre à notre première hypothèse ne nous permet pas de conclure que les gros rouleurs en voitures roulent moins à moto ni que les gros rouleurs à motos roulent moins en voitures.

Cependant il nous a permis de relever une certaine tendance. En effet, les très gros utilisateurs d'un type de véhicule avaient tendance à moins utiliser l'autre.

L'analyse de nos données nous a permis de nous rendre compte que les répondants roulent davantage en auto qu'en 2-roues ce qui nous a permis de valider notre deuxième hypothèse.

Ensuite par manque de données nous n'avons pu savoir s'il y avait une corrélation entre le sexe d'un individu et le nombre de kilomètres parcourus en auto et moto.

Pour finir, nous avons constaté que contrairement à ce que l'on pensait, les individus en couple roulaient plus à moto que ceux célibataires. De plus, bien que les familles semblent utiliser plus la moto lorsque leur nombre d'enfants à charge augmente, elle n'utilise pas plus l'auto pour autant.

## 2. Utilisation par rapport aux usages

### Problématique

Nous rappelons que le but de notre étude est de déterminer les facteurs qui influent sur la comparaison entre voitures et 2-roues au travers de la problématique suivante :

*“ Quels sont les différents facteurs qui amènent à la comparaison entre voiture et 2-roues? “*

Le facteur qui est étudié dans cette partie c'est l'usage des véhicules. Nous l'étudierons au travers de catégorie d'individus pour lesquelles nous avons formulé trois hypothèses :

- h1 *“Les personnes qui ont des voitures et des motos ne les utilisent pas pour les mêmes usages.”*
- h2. *“Les gros rouleurs sont ceux qui travaillent avec leurs voitures / motos.”*
- h3. *“Etre un gros rouleur est indépendant de l'usage du véhicule.”*

Notre démarche sera d'abord de comprendre les usages pour l'ensemble des données sans catégoriser les individus. Puis nous analyserons plus finement pour comprendre la répartition des usages en fonction du type de véhicule et de la catégorie d'utilisateurs pour déterminer s'il y a corrélation. Cette analyse nous permettra de conclure sur les différentes hypothèses que nous avons formulées.

### Données

Comme pour la partie précédente les données proviennent de réponse à des questionnaires envoyé par la mutuelle des motards à leurs abonnés. Un questionnaire correspond aux motos et un autre aux voitures. Nous avons croisé les données pour ne garder que les utilisateurs ayant répondu aux deux. La donnée importante de cette partie c'est l'usage qui est fait des véhicules. Le questionnaire des motos permet de renseigner jusqu'à 6 cylindrées et pour chacune d'entre elles de sélectionner un usage parmi : loisir, trajet domicile-travail, trajets professionnels. Le questionnaire des voitures permet de renseigner un véhicule et pour ce dernier de sélectionner un ou plusieurs usages parmi : sport ou autres loisirs, trajets domicile travail, trajets professionnels, les courses, sorties, vacances ou week-end. La manière de renseigner les usages et leurs types sont différents en fonction du type de véhicule, cela empêche de réaliser des comparaisons pertinentes. Pour palier à cela nous avons regroupé les usages : sport ou autres loisirs, sorties, vacances ou week-end ; en une catégorie loisir. Nous avons choisi de ne pas prendre en compte les course et magasin qui correspondait à une catégorie non proposer aux motos, cette catégorie reste pertinente et elle pourra être utilisée lors de la conclusion de l'analyse. Nous avons ainsi trois catégories identiques pour les usages des voitures et des motos. Nos individus sont catégorisés de la même manière que dans la partie précédente (gros rouleur, petit rouleur, situation familiale).

## Analyse

### Hypothèse n°1:

Dans cette partie nous allons analyser les données de manière à répondre à cette hypothèse.

« Les personnes qui ont des voitures et des motos ne les utilisent pas pour les mêmes usages. »

Pour commencer on s'intéresse à l'usage que font les individus de leur véhicule de manière globale.

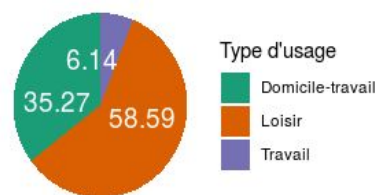
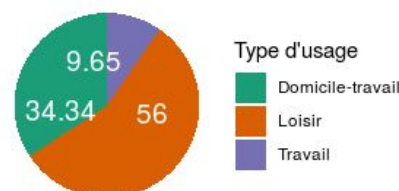


Figure 12 : Répartition des usages pour les motos et les automobiles

On constate que l'usage majoritaire correspond au loisir, l'utilisation domicile-travail est aussi beaucoup représentée. Par contre peu de personnes utilisent leur véhicule dans le cadre du travail.

On s'intéresse maintenant spécifiquement aux automobiles, nous avons remarqué dans le questionnaire que l'utilisateur pouvait sélectionner plusieurs usages différents, nous avons donc fait un graphe de l'usage global de l'automobile.

Répartition des usages des autos

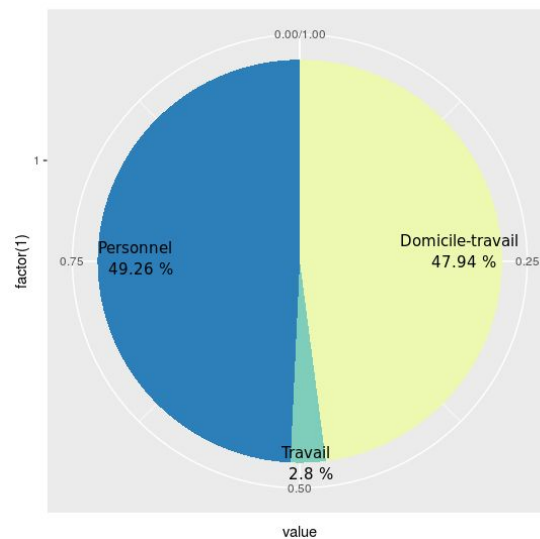


ramme R, répartition des usages des autos

Figure 13 : Répartition des usages pour les automobiles

Pour les motos, nous avons remarqué que l'utilisateur pouvait mettre plusieurs motos différentes, mais que pour une moto, l'utilisateur avait un seul choix par véhicule. Nous avons donc fait un graphe sur l'usage principal de la moto. La catégorie « Personnel » correspond à la catégorie « Loisir ».





*Figure 14 : Répartition des usages pour les motos*

En comparant les 2 graphes on remarque déjà une petite différence entre les usages qui sont faits des voitures et des motos. Les motos sont moins utilisées dans le cadre du travail ce qui paraît logique, il est très rare de se déplacer à moto dans le cadre du travail. La moto est aussi favorisée pour les trajets domicile-travail une fois de plus cela peut s'expliquer de manière logique, on est en général tout seul pour aller au travail tandis que les loisirs peuvent être pratiqués à plusieurs.

Nous nous appuyons maintenant sur une Analyse Factorielle des Correspondances (AFC) pour confirmer ce que l'on observe en comparant grossièrement les données.

La première étape consiste à émettre l'hypothèse  $h_0$  : « Le type d'utilisation qui est faite par les individus des véhicules est indépendante du type de véhicule ». Le test du khi2 permet de rejeter cette hypothèse et de considérer que le type d'utilisation dépend du véhicule. Il est donc pertinent de réaliser une AFC.

La réalisation de l'AFC nous donne ce diagramme :

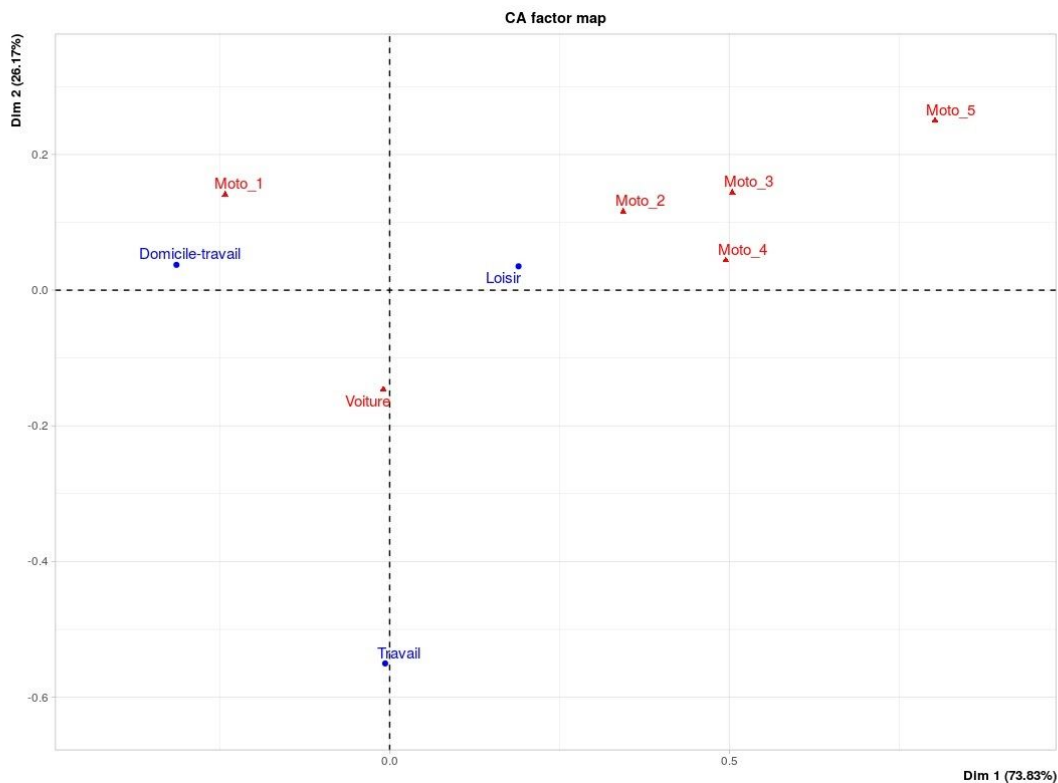


Figure 15 : AFC sur les variables Véhicule et Usage

Tout d'abord on remarque la représentation des axes est de 100%, nos données sont donc totalement représentées par ces axes. Trois groupes de véhicules se distinguent, la voiture, la moto principale puis les motos secondaires. L'axe 1 représente principalement les modalités Domicile-travail et loisir, la modalité travail est représentée par l'axe 2. On constate que la voiture est bien privilégiée pour le travail, on remarque que la moto principale est plus destinée au trajet domicile travail tandis que les motos secondaires correspondent plus aux loisirs.

On peut conclure que l'hypothèse est validée, l'usage des véhicules dépend en effet du type de véhicule. Cet usage dépend aussi de l'importance du véhicule. Il est important de nuancer le résultat en rappelant que pour la voiture les usagers avaient la possibilité de sélectionner plusieurs usages ce qui explique son positionnement central. Cela signifie que cette modalité est plus nuancée que les autres.

## Hypothèse n°2:

Dans cette partie nous allons analyser les données de manière à répondre à cette hypothèse.  
Les gros rouleurs sont ceux qui travaillent avec leurs voitures / motos.

Après avoir récolté les données nécessaires, nous n'avons pas pu conclure sur cette hypothèse par manque de données. En effet, nous avons seulement 16 personnes sur 607 qui utilisent leur voiture pour les trajets professionnels, et 17 personnes sur 607 qui utilisent leur moto.

Nous n'avons donc pas pu conclure sur cette hypothèse.

## Hypothèse n°3:

Dans cette partie nous allons analyser les données de manière à répondre à cette hypothèse :  
*Être un gros rouleur est indépendant de l'usage du véhicule.*

### Données

Les données importantes de cette partie sont : la fréquence d'utilisation et les usages des véhicules.

### Analyse

Nous souhaitons comparer l'utilisation de la voiture et de la moto par rapport à la fréquence. Nous définissons la fréquence d'utilisation en fonction du nombre de kilomètres parcouru par an avec un véhicule. Et l'usage du véhicule en fonction du premier choix pour l'automobile, et le choix du premier véhicule pour la moto. Nous avons donc deux variables avec plusieurs modalités : la variable "*Véhicule*" avec "*Gros rouleur*" et "*Petit rouleur*" (définies précédemment : gros rouleur si plus de 17000 km sur un an avec le véhicule), et la variable "*Usage*" avec :

- pour l'AFC avec les automobiles :
  - "Travail" pour les trajets professionnels
  - "Domicile-travail" pour les trajets domicile-travail
  - "Personnel" pour les trajets personnels (sport, week-end, vacance..)
  - "Course" pour les trajets liés aux courses
- pour l'AFC avec les motos :
  - "Travail" pour les trajets professionnels
  - "Domicile-travail" pour les trajets domicile-travail
  - "Personnel" pour les trajets personnels (sport, week-end, vacance..)

Dans le but de voir les liens entre ces variables, nous avons pensé à réaliser 2 AFC car c'est la méthode adaptée aux tableaux de données où 2 variables qualitatives décrivent un ensemble d'individus (une AFC pour chaque type de véhicule : Automobile et Moto). Pourquoi 2 analyses différentes selon le type de véhicule? Tout simplement parce que nous n'avons pas les mêmes données pour ceux-ci. Pour l'automobile, l'utilisateur peut sélectionner plusieurs usages différents. Nous avons sélectionné le premier dans la base de données, mais ceux-ci peut porter à confusion. Le résultat de cet AFC ne sera donc pas concluante à 100%. Contrairement à l'AFC sur la moto. En effet, pour celle-ci, l'utilisateur n'a qu'un choix possible, donc nous avons l'usage principale de la moto.

	Personnel	DomicileTravail	Travail	Course
<b>Gros rouleur Auto</b>	110	61	11	14
<b>Petit rouleur Auto</b>	210	90	5	106

Figure 16 : Tableau de contingence entre les variables fréquence automobile et usage

Ci-dessus, nous avons le tableau de contingence pour l'AFC sur les automobiles. Celui-ci confirme aussi notre réponse à notre deuxième hypothèse (très peu de données pour bien analyser). Une fois ce tableau de fait, nous avons effectué un test du khi-deux afin de tester ces deux hypothèses :

H0 : le fait d'être un gros ou un petit rouleur est indépend de l'usage de l'automobile

H1 : les deux variables sont dépendantes

```
> chisq.test(tabContingenceAuto)
```

Pearson's Chi-squared test

```
data: tabContingenceAuto
```

```
X-squared = 38.248, df = 3, p-value = 2.504e-08
```

Ici X-squared > 7,82 donc on rejette l'hypothèse d'indépendance. On a donc mis en évidence un lien entre les 2 variables.

	Personnel	DomicileTravail	Travail
<b>Gros rouleur Moto</b>	18	39	6
<b>Petit rouleur Moto</b>	281	252	11

Figure 17 : Tableau de contingence entre les variables fréquence moto et usage

Ci-dessus, nous avons le tableau de contingence pour l'AFC sur les automobiles. Celui-ci confirme aussi notre réponse à notre deuxième hypothèse (très peu de données pour bien analyser). Une fois ce tableau de fait, nous avons effectué un test du khi-deux afin de tester ces deux hypothèses :

H0 : le fait d'être un gros ou un petit rouleur est indépend de l'usage de la moto

H1 : les deux variables sont dépendantes

```
> chisq.test(tabContingenceMoto)
```

```
Pearson's Chi-squared test
```

```
data: tabContingenceMoto  
X-squared = 20.312, df = 2, p-value = 3.884e-05
```

Ici X-squared > 5,99 donc on rejette l'hypothèse d'indépendance. On a donc mis en évidence un lien entre les 2 variables.

⇒ Suite à nos deux tests du khi-deux, et donc à nos 2 AFC, nous avons pu conclure qu'il y a un lien entre la fréquence d'utilisation d'un véhicule et l'usage de ce véhicule.

## Conclusion du second axe

Dans cette partie nous avons cherché à savoir si l'usage qui est fait des véhicules est un facteur de comparaison entre voitures et 2-roues. L'analyse liée à notre première hypothèse nous permet de conclure qu'il existe un lien entre le type de véhicule et l'usage qu'il en est fait. Ce lien peut permettre d'expliquer pourquoi certains utilisateurs vont privilégier un certain type de véhicule. Ce lien peut être caractérisé pas des contraintes professionnelles ou le besoin de se déplacer à plusieurs. L'analyse de la troisième hypothèse renseigne que le type d'usage qui est fait des véhicules est aussi lié au kilomètres que l'on parcourt avec. Il est assez évident qu'une personne utilisant son véhicule dans le cadre du travail réalisera plus de km qu'une personne qui ne l'utilise que dans un cadre privé et cela peu importe le type de véhicule.

## Conclusion

Au cours de l'analyse des données fournis par la Mutuelle des Motards nous avons comparé l'utilisation de la voiture et celle du 2-roues. Pour cela nous avons choisi deux axes d'études, la fréquence d'utilisation d'un véhicule et l'usage qu'on en fait. Notre but était de comprendre si le type de véhicule (2-roues et voiture) avait un impact sur nos axes. L'analyse des usages montre que ces derniers diffère en fonction du type de véhicule. On peut déduire de tout cela que lorsque les individus choisissent d'utiliser un certain type de véhicule ils prennent en compte l'usage qu'ils en auront. Le volume d'utilisation peut jouer un rôle mais nous n'avons pas réussi à le démontrer. Tout comme il ne semble pas y avoir de corrélation entre la manière d'utiliser les motos et la manière d'utiliser les voitures.

# Annexes

## Annexe 1:

```
#----- INITIALISATION DATA -----#
install.packages("readxl")
library("readxl")

bdd_2_roue <- read_excel("Data/bdd_2_roue.xls")
View(bdd_2_roue)

bdd_auto <- read_excel("Data/bdd_auto-modif.xlsx")
View(bdd_auto)

#renommage des 2 premieres colonnes de la base 2-roue
colnames(bdd_2_roue)[colnames(bdd_2_roue)=="...1"] <- "Identifiant"
colnames(bdd_2_roue)[colnames(bdd_2_roue)=="...2"] <- "Date"

#creation de la base de donnees conjointes (avec les personnes utilisant auto et
2-roue)
bdd <- merge(bdd_auto, bdd_2_roue, by="Identifiant")

#colnames(bdd_moto_2019)[colnames(bdd_moto_2019)=="IDENTIFIANT"] <- "Identifiant"
#creation de la 2eme base en prenant en compte les donnees 2019
#bdd_2019 <- merge(bdd, bdd_moto_2019, by="Identifiant")
#View(bdd_2019)

#####
#----- Installation des packages et chargement des library-----#
install.packages("ggplot2")
install.packages("dplyr")
library(ggplot2)
library(dplyr)
```

```
#####
#----- Usage table -----#

# début de la création d'une table avec tout les usages, ça correspondrait a la
table utile

# pour répondre a cet axe, il faut récupérer tout les usage et ainsi que le nombre
de km associé

usage <- bdd %>%
  select(Identifiant,      nbKmVoiture="Q17      [1]",      usageVoiture1="Q15      [1]",
usageVoiture2="Q15 [2]",
        usageVoiture3="Q15 [3]", usageVoiture4="Q15 [4]", usageVoiture5="Q15 [5]",
        usageVoiture6="Q15 [6]", usagev1="Usage 1", kmv1="Nb km 1",
usagev2="...54",
        kmv2="...56", usagev3="...65", kmv3="...67", usagev4="...76", kmv4="...78",
        usagev5="...88", kmv5="...90")

#-----Divison des usages-----#
usages <- bdd %>%
  select(Identifiant, usage_auto="Q15 [1]",usage_moto1="Usage 1")
usages[is.na(usages)] <- 0

#fonctions pour transformer les donnees numériques en texte
to_readable_auto <- function(code){
  if (code == 1)
    return ("Personnel")
  if (code == 2)
    return ("Domicile-travail")
  if (code == 3)
    return ("Travail")
  if (code == 4)
    return ("Courses")
  if (code == 5)
    return ("Personnel")
  if (code == 6)
    return ("Personnel")
}
to_readable_auto_vector <- Vectorize(to_readable_auto, vectorize.args = "code")

to_readable_moto <- function(code){
  if (code == 0)
```



```

    return ("Non renseigné")
  if (code == 1)
    return ("Personnel")
  if (code == 2)
    return ("Domicile-travail")
  if (code == 3)
    return ("Travail")
}

to_readable_moto_vector <- Vectorize(to_readable_moto, vectorize.args = "code")

#tableau avec id + usage principal de l'auto et de la moto1
usages <- usages %>%
  mutate(Usage_auto = to_readable_auto_vector(usage_auto)) %>%
  mutate(Usage_moto = to_readable_moto_vector(usage_moto1))

nbUsager <- nrow(bdd)

usageMotoPersonnelGlobal <- sum(usages['Usage_moto'] == 'Personnel')
usageMotoDomicileTravailGlobal <- sum(usages['Usage_moto'] == 'Domicile-travail')
usageMotoTravailGlobal <- sum(usages['Usage_moto'] == 'Travail')

#Camembert usage principal moto

usagePrincipalMoto <- data.frame(
  group = c("Domicile-travail", "Travail", "Personnel"),
  value = c(round((usageMotoDomicileTravailGlobal/nbUsager),4),
    round((usageMotoTravailGlobal/nbUsager),4),
    round((usageMotoPersonnelGlobal/nbUsager),4))
)

library(extrafont)

loadfonts(device="win")

usagePrincipalMoto$group <- factor(usagePrincipalMoto$group, levels =
rev(usagePrincipalMoto$group))

ggplot(data = usagePrincipalMoto, mapping = aes(x = factor(1), y = value, fill =
group)) +

```

```

geom_bar(width=1, stat = "identity") +
coord_polar(theta = "y") +
scale_fill_brewer(type = "seq", direction = -1, palette= "YlGnBu", guide = F) +
geom_text(aes(x = c(1.3, 1.5, 1.3),
                y = value/2 + c(0, cumsum(value)[-length(value)]),
                label=paste(group, "\n", value*100, "%")), family = "Consolas")

usageAutoPersonnelGlobal <- sum(usages['Usage_auto'] == 'Personnel')
usageAutoDomicileTravailGlobal <- sum(usages['Usage_auto'] == 'Domicile-travail')
usageAutoTravailGlobal <- sum(usages['Usage_auto'] == 'Travail')
usageAutoCourseGlobal <- sum(usages['Usage_auto'] == 'Courses')

#Camembert usage principal auto
usagePrincipalAuto <- data.frame(
  group = c("Personnel", "Domicile-travail", "Travail", "Courses"),
  value = c(round((usageAutoPersonnelGlobal/nbUsager), 4),
             round((usageAutoDomicileTravailGlobal/nbUsager), 4),
             round((usageAutoTravailGlobal/nbUsager), 4),
             round((usageAutoCourseGlobal/nbUsager), 4))
)

library(extrafont)

loadfonts(device="win")

usagePrincipalAuto$group <- factor(usagePrincipalAuto$group, levels =
rev(usagePrincipalAuto$group))

ggplot(data = usagePrincipalAuto, mapping = aes(x = factor(1), y = value, fill =
group)) +
  geom_bar(width=1, stat = "identity") +
  coord_polar(theta = "y") +
  scale_fill_brewer(type = "seq", direction = -1, palette= "YlGnBu", guide = F) +
  geom_text(aes(x = c(1.3, 1.5, 1.3, 1.3),
                y = value/2 + c(0, cumsum(value)[-length(value)]),
                label=paste(group, "\n", value*100, "%")), family = "Consolas")

#Camembert usage global auto
usageGlobalAutoPerso <-
round(((usageVoiture[1,][2]+usageVoiture[5,][2]+usageVoiture[6,][2])/usageVoiture[7,][2]), 4)

```

```
usageGlobalAutoDT <- round((usageVoiture[2,][2]/usageVoiture[7,][2]),4)
usageGlobalAutoT <- round((usageVoiture[3,][2]/usageVoiture[7,][2]),4)
usageGlobalAutoC <- round((usageVoiture[4,][2]/usageVoiture[7,][2]),4)

usageGlobalAuto <- data.frame(
  group = c("Personnel", "Domicile-travail", "Travail", "Coursess"),
  value = c(usageGlobalAutoPerso[1,],
            usageGlobalAutoDT[1,],
            usageGlobalAutoT[1,],
            usageGlobalAutoC[1,])
)

library(extrafont)

loadfonts(device="win")

usageGlobalAuto$group <- factor(usageGlobalAuto$group, levels =
rev(usageGlobalAuto$group))

ggplot(data = usageGlobalAuto, mapping = aes(x = factor(1), y = value, fill =
group)) +
  geom_bar(width=1, stat = "identity") +
  coord_polar(theta = "y") +
  scale_fill_brewer(type = "seq", direction = -1, palette= "YlGnBu", guide = F) +
  geom_text(aes(x = c(1.3, 1.5, 1.3, 1.3),
                y = value/2 + c(0, cumsum(value)[-length(value)]),
                label=paste(group,"\n",value*100, "%")), family = "Consolas")

##### Données pour diviser les individus en 2 groupes (gros rouleur VS petit
rouleur) #####

grosRouleurAuto <- 0
grosRouleurMoto <- 0
petitRouleurAuto <- 0
petitRouleurMoto <- 0

repartitionRouleur <- function(code) {
```

```
if (code > 17000)
  return (1)
else
  return (0)
}

for (i in (bdd$`Q17 [1]`))
  if (repartitionRouleur(strtoi(i)) == 1) {
    grosRouleurAuto <- grosRouleurAuto + 1
  } else {
    petitRouleurAuto <- petitRouleurAuto + 1
  }

for (i in (bdd$`Nb km 1`))
  if (repartitionRouleur(strtoi(i)) == 1) {
    grosRouleurMoto <- grosRouleurMoto + 1
  } else {
    petitRouleurMoto <- petitRouleurMoto + 1
  }

### Tableau de contingence ###

usageAutoPersonnelGrosRouleur <- 0
usageAutoTravailGrosRouleur <- 0
usageAutoDTGrosRouleur <- 0
usageAutoCourseGrosRouleur <- 0

usageAutoPersonnelPetitRouleur <- 0
usageAutoTravailPetitRouleur <- 0
usageAutoDTPetitRouleur <- 0
usageAutoCoursePetitRouleur <- 0

usageMotoPersonnelGrosRouleur <- 0
usageMotoTravailGrosRouleur <- 0
usageMotoDTGrosRouleur <- 0

usageMotoPersonnelPetitRouleur <- 0
usageMotoTravailPetitRouleur <- 0
usageMotoDTPetitRouleur <- 0
```

```
#répartition pour les voitures avec le premier choix coché

#pas super mais pas le choix

for (i in 1:nrow(bdd))
  if (bdd$`Q15 [1]`[i] == 1 || bdd$`Q15 [1]`[i] == 5 || bdd$`Q15 [1]`[i] == 6) {
    if (repartitionRouleur(strtoi(bdd$`Q17 [1]`[i])) == 1) {
      usageAutoPersonnelGrosRouleur <- usageAutoPersonnelGrosRouleur + 1
    } else {
      usageAutoPersonnelPetitRouleur <- usageAutoPersonnelPetitRouleur + 1
    }
  } else if (bdd$`Q15 [1]`[i] == 2) {
    if (repartitionRouleur(strtoi(bdd$`Q17 [1]`[i])) == 1) {
      usageAutoDTGrosRouleur <- usageAutoDTGrosRouleur + 1
    } else {
      usageAutoDTPetitRouleur <- usageAutoDTPetitRouleur + 1
    }
  } else if (bdd$`Q15 [1]`[i] == 3) {
    if (repartitionRouleur(strtoi(bdd$`Q17 [1]`[i])) == 1) {
      usageAutoTravailGrosRouleur <- usageAutoTravailGrosRouleur + 1
    } else {
      usageAutoTravailPetitRouleur <- usageAutoTravailPetitRouleur + 1
    }
  } else {
    if (repartitionRouleur(strtoi(bdd$`Q17 [1]`[i])) == 1) {
      usageAutoCourseGrosRouleur <- usageAutoCourseGrosRouleur + 1
    } else {
      usageAutoCoursePetitRouleur <- usageAutoCoursePetitRouleur + 1
    }
  }
}

for (i in 1:nrow(bdd))
  if (bdd$`Usage 1`[i] == 1) {
    if (repartitionRouleur(strtoi(bdd$`Nb km 1`[i])) == 1) {
      usageMotoPersonnelGrosRouleur <- usageMotoPersonnelGrosRouleur + 1
    } else {
      usageMotoPersonnelPetitRouleur <- usageMotoPersonnelPetitRouleur + 1
    }
  } else if (bdd$`Usage 1`[i] == 2) {
    if (repartitionRouleur(strtoi(bdd$`Nb km 1`[i])) == 1) {
      usageMotoDTGrosRouleur <- usageMotoDTGrosRouleur + 1
    } else {
      usageMotoDTPetitRouleur <- usageMotoDTPetitRouleur + 1
    }
  }
}
```

```

    }
  } else {
    if (repartitionRouleur(strtoi(bdd$`Nb km 1`[i])) == 1) {
      usageMotoTravailGrosRouleur <- usageMotoTravailGrosRouleur + 1
    } else {
      usageMotoTravailPetitRouleur <- usageMotoTravailPetitRouleur + 1
    }
  }
}

## Création du tableau de contingence Auto ##

tabContingenceAuto <- data.frame(
  Personnel=c(usageAutoPersonnelGrosRouleur,
              usageAutoPersonnelPetitRouleur),
  DomicileTravail=c(usageAutoDTGrosRouleur,
                     usageAutoDTPetitRouleur),
  Travail=c(usageAutoTravailGrosRouleur,
             usageAutoTravailPetitRouleur),
  Course=c(usageAutoCourseGrosRouleur,
            usageAutoCoursePetitRouleur),
  row.names=c("Gros rouleur Auto", "Petit rouleur Auto")
)

resCA<-CA(tabContingenceAuto)

tabContingenceMoto <- data.frame(
  Personnel=c(usageMotoPersonnelGrosRouleur,
              usageMotoPersonnelPetitRouleur),
  DomicileTravail=c(usageMotoDTGrosRouleur,
                     usageMotoDTPetitRouleur),
  Travail=c(usageMotoTravailGrosRouleur,
             usageMotoTravailPetitRouleur),
  row.names=c("Gros rouleur Moto", "Petit rouleur Moto")
)

resCA<-CA(tabContingenceMoto)

##### AFC rouleur auto/moto

GrosRouleurAutoetMoto <- 0
GrosRouleurAutoetPetitRouleurMoto <- 0

```

```
GrosRouleurMotoetPetitRouleurAuto <- 0
PetitRouleurAutoetMoto <- 0

for (i in 1:nrow(bdd))
  if (repartitionRouleur(strtoi(bdd$`Nb km 1`[i])) == 1) {
    if (repartitionRouleur(strtoi(bdd$`Q17 [1]`[i])) == 1) {
      GrosRouleurAutoetMoto <- GrosRouleurAutoetMoto + 1
    } else {
      GrosRouleurMotoetPetitRouleurAuto <- GrosRouleurMotoetPetitRouleurAuto + 1
    }
  } else if (repartitionRouleur(strtoi(bdd$`Nb km 1`[i])) == 0) {
    if (repartitionRouleur(strtoi(bdd$`Q17 [1]`[i])) == 1) {
      GrosRouleurAutoetPetitRouleurMoto <- GrosRouleurAutoetPetitRouleurMoto + 1
    } else {
      PetitRouleurAutoetMoto <- PetitRouleurAutoetMoto + 1
    }
  }
}

tabContingenceRouleur <- data.frame(
  GrosRouleurAuto=c(GrosRouleurAutoetMoto,GrosRouleurAutoetPetitRouleurMoto),
  PetitRouleurAuto=c(GrosRouleurMotoetPetitRouleurAuto,PetitRouleurAutoetMoto),
  row.names=c("GrosRouleurMoto", "PetitRouleurMoto")
)

tabContingenceRouleur2 <- data.frame(
  Auto=c(grosRouleurAuto,petitRouleurAuto),
  Moto=c(grosRouleurMoto,petitRouleurMoto),
  row.names=c("Gros rouleur", "Petit rouleur")
)

resAFCRouleur2 <- CA(tabContingenceRouleur2)
```

## Annexe 2:

```
#####
#----- Diagramme du nombre d'autos -----#
intro <- bdd %>%
  select(Identifiant, Q1="Q1 [1]") %>%
  arrange(Q1) %>%
  group_by(Q1) %>%
  summarise("Nombre_de_personnes"= n())

View(intro)

intro2 <- intro %>%
  arrange(desc(Q1)) %>%
  mutate(prop = round(Nombre_de_personnes / sum(intro$Nombre_de_personnes) *100,2))
%>%
  mutate(ypos = cumsum(prop)- 0.5*prop )
#View(intro2)

#On ne met rien sur l'axe x pour pouvoir en faire un cercle
diag_intro <- ggplot(intro2, aes(x = "", y = prop, fill = factor(Q1))) +
  geom_bar(width = 1,stat = "identity") +
  theme_void() +
  labs(fill="Nombre d'autos",
        x=NULL,
        y=NULL,
        title="Rpartition du nombre d'autos",
        caption="Figure 1: Diagramme R, rpartition du nombre d'autos") +
  geom_text(aes(y = ypos, label = prop), color = "white", size=5,check_overlap =
TRUE) +
  scale_fill_brewer(palette="Set1") +
  coord_polar("y", start=0)
#diag_intro

#####
```



```
#----- Diagramme des gros rouleurs en moto/auto-----#

gros_rouleurs <- bdd %>%
  select(Identifiant, moto1="Nb km 1", moto2="...56", moto3="...67", moto4="...78",
kmAuto="Q17 [1]", Assure_Mutuelle = "Assur Mutuelle", Nombre_personnes_foyer =
"Nombre personnes au foyer", Situation_familiale = "Situation familiale")

#class(gros_rouleurs$moto1)
#On doit changer le type pour pouvoir sommer
gros_rouleurs$moto1 <- as.numeric(as.character(gros_rouleurs$moto1))
gros_rouleurs$moto2 <- as.numeric(as.character(gros_rouleurs$moto2))
gros_rouleurs$moto3 <- as.numeric(as.character(gros_rouleurs$moto3))
gros_rouleurs$moto4 <- as.numeric(as.character(gros_rouleurs$moto4))
#class(gros_rouleurs$moto1)

#On transforme les valeurs nulles en 0 pour pouvoir sommer

gros_rouleurs[is.na(gros_rouleurs)] <- 0
gr_sum <- gros_rouleurs %>%
  mutate(kmMoto = rowSums(.[2:5])) %>%
  arrange(desc(kmMoto))
View(gr_sum)

gr_sum_total <- gr_sum %>%
  mutate(kmTotaux = kmMoto + kmAuto)

moyenne_km_an_echantillon <- mean(gr_sum_total$kmTotaux) #27718.2 km

moyenne_km_an <- 17000

#0=petit rouleur, 1= gros rouleur
isGrosRouleur_moto <- function(km) {
  if (km <= moyenne_km_an)
    return (0)
  if (km > moyenne_km_an)
    return (1)
}

isGrosRouleur_moto_vector <- Vectorize(isGrosRouleur_moto, vectorize.args = "km")

isGrosRouleur_auto <- function(km) {
  if (km <= moyenne_km_an)
    return (0)
```

```
if (km > moyenne_km_an)
  return (1)
}

isGrosRouleur_auto_vector <- Vectorize(isGrosRouleur_auto, vectorize.args = "km")

isAssure <- function(num) {
  if (num == "1")
    return (1)
  if (num == "2")
    return (0)
}

isAssure_vector <- Vectorize(isAssure, vectorize.args = "num")

situation_familiale <- function(num) {
  if (num == "1")
    return ("Seul(e) sans enfant")
  if (num == "2")
    return ("Seul(e) avec enfant(s)")
  if (num == "3")
    return ("En couple sans enfant")
  if (num == "4")
    return ("En couple avec enfant(s)")
  if (num == "5")
    return ("Autre")
}

situation_familiale_vector <- Vectorize(situation_familiale, vectorize.args =
"num")

is_en_couple <- function(situation) {
  if (situation == "Seul(e) sans enfant")
    return ("Celibataire")
  if (situation == "Seul(e) avec enfant(s)")
    return ("Celibataire")
  if (situation == "En couple sans enfant")
    return ("En couple")
  if (situation == "En couple avec enfant(s)")
    return ("En couple")
  if (situation == "Autre")
    return ("Celibataire")
}

is_en_couple_vector <- Vectorize(is_en_couple, vectorize.args = "situation")
```

```

hasEnfants <- function(situation){
  if (situation == "Seul(e) sans enfant")
    return (0)
  if (situation == "Seul(e) avec enfant(s)")
    return (1)
  if (situation == "En couple sans enfant")
    return (0)
  if (situation == "En couple avec enfant(s)")
    return (1)
  if (situation == "Autre")
    return (0)
}

hasEnfants_vector <- Vectorize(hasEnfants, vectorize.args = "situation")

type_rouleur <- gr_sum_total %>%
  mutate(gros_rouleur_moto = isGrosRouleur_moto_vector(kmMoto)) %>%
  mutate(gros_rouleur_auto = isGrosRouleur_auto_vector(kmAuto)) %>%
  mutate(assure_mutuelle = isAssure_vector(Assure_Mutuelle)) %>%
  mutate(situation_familiale = situation_familiale_vector(Situation_familiale)) %>%
  mutate(en_couple = is_en_couple_vector(situation_familiale)) %>%
  mutate(avec_enfant = hasEnfants_vector(situation_familiale)) %>%

select(Identifiant,gros_rouleur_moto,gros_rouleur_auto,assure_mutuelle,en_couple,av
ec_enfant,kmTotaux,kmAuto,kmMoto)

View(type_rouleur)

situation_couple <- type_rouleur %>%
  group_by(en_couple) %>%
  summarise(moy_km_moto = round(mean(kmAuto, na.rm=TRUE),2),
            moy_km_auto = round(mean(kmMoto),2),
            moy_km_total = round(mean(kmTotaux),2)
  )
View(situation_couple)

kmEtudie <- situation_couple$moy_km_auto
kmEtudie <- situation_couple$moy_km_moto

ggplot(data=situation_couple, aes_string(x="en_couple", y=kmEtudie, fill =
"en_couple")) +

```

```

geom_bar(stat="identity") +
scale_fill_brewer(palette="Set1") +
labs(fill="situation",
      x="Situation",
      y="Km",
      title="Moyenne de km parcourus en fonction de la situation maritale") +
geom_text(aes_string(y = kmEtudie / 2, label = kmEtudie ), color = "white",
size=5) +
theme(legend.position="none")

#AFC#####
install.packages(c("FactoMineR", "factoextra"))
install.packages("questionr")
library("FactoMineR")
library("factoextra")
library("questionr")
type_rouleur2 <- type_rouleur
str(type_rouleur2)

for(i in 2:6){
  type_rouleur2[,i]<-as.factor(type_rouleur2[,i])
}

res.mca <- MCA(type_rouleur2[,2:6])
plot(res.mca,invisible=c("ind"))

tab_situation_familiale_moto <-
table(type_rouleur$situation_familiale,type_rouleur$gros_rouleur_moto,type_rouleur$
gros_rouleur_auto)
chisq <- chisq.test(tab_situation_familiale_moto)
chisq
tab_situation_familiale_moto <- rprop(tab_situation_familiale_moto)
tab_situation_familiale_moto
res.ca <- MCA(type_rouleur2[,2:6])

#####

sup_moto <- gr_sum %>%
  group_by(kmMoto) %>%
  filter(kmMoto >= moyenne_km_an)
inf_moto <- gr_sum %>%
  group_by(kmMoto) %>%

```

```

filter(kmMoto < moyenne_km_an)

sup_auto <- gr_sum %>%
  group_by(auto) %>%
  filter(auto >= moyenne_km_an)
inf_auto <- gr_sum %>%
  group_by(auto) %>%
  filter(auto < moyenne_km_an)

View(sup_auto)
View(inf_auto)

#TODO: afficher la part de gros rouleurs en moto
#les gros rouleurs en voiture,
#les gros rouleurs en moto qui sont des gros rouleurs en voiture
#les gros rouleurs en voiture qui sont des gros rouleurs en moto
# partir de ces diagrammes et calculs on dduit le reste
#str(sup_auto)

#resum_gros_rouleurs <- data.frame("rows_bdd" = dim(bdd)[1],
#                                "rows_sup_auto" = dim(sup_auto)[1],
#                                "rows_sup_moto" = dim(sup_moto)[1],
#                                "rows_inf_auto" = dim(inf_auto)[1],
#                                "rows_inf_moto" = dim(inf_moto)[1],
#                                "modalite de separation" = "Moyenne fr: 17000
km")

effectifs <- data.frame(stat=c("Total étudié", "Gros rouleurs en moto", "Petits
rouleurs en moto", "Gros rouleurs en auto", "Petits rouleurs en auto"),

effectif=c(dim(bdd)[1], dim(sup_moto)[1], dim(inf_moto)[1], dim(sup_auto)[1], dim(inf_a
uto)[1]))
View(effectifs)

pourcentage_gros_rouleurs_fr <- effectifs[-5,][-3,][-1,] %>%
  mutate(pourcentage = round(effectif / dim(bdd)[1] * 100, 2))
pourcentage_petits_rouleurs_fr <- effectifs[-4,][-2,][-1,] %>%
  mutate(pourcentage = round(effectif / dim(bdd)[1] * 100, 2))

View(pourcentage_gros_rouleurs_fr)

```

```

View(pourcentage_petits_rouleurs_fr)

diag_pourcentage_GR_fr <-
  ggplot(data=pourcentage_gros_rouleurs_fr, aes(x=stat, y=pourcentage, fill
=factor(stat))) +
  geom_bar(stat="identity") +
  scale_fill_brewer(palette="Set1") +
  labs(fill="Vehicule",
        x=NULL,
        y="%",
        title="Répartition du nombre de gros rouleurs en fonction du véhicule",
        caption="Figure 2: Répartition du nombre de gros rouleurs en fonction du
véhicule") +
  geom_text(aes(y = pourcentage / 2, label = pourcentage ), color = "white",
size=5,check_overlap = TRUE)+
  theme(legend.position="none")
diag_pourcentage_GR_fr

diag_pourcentage_PR_fr <-
  ggplot(data=pourcentage_petits_rouleurs_fr, aes(x=stat, y=pourcentage, fill
=factor(stat))) +
  geom_bar(stat="identity") +
  scale_fill_brewer(palette="Set1") +
  labs(fill="Vehicule",
        x=NULL,
        y="%",
        title="Répartition du nombre de petits rouleurs en fonction du véhicule",
        caption="Figure 2: Répartition du nombre de petits rouleurs en fonction du
véhicule") +
  geom_text(aes(y = pourcentage / 2, label = pourcentage ), color = "white",
size=5,check_overlap = TRUE) +
  theme(legend.position="none")
diag_pourcentage_PR_fr

```

## Annexe 3:

```
#####
#----- Usage des voitures -----#

usage1 <- bdd %>%
  select(usage="Q15 [1]")
usage2 <- bdd %>%
  select(usage="Q15 [2]")
usage3 <- bdd %>%
  select(usage="Q15 [3]")
usage4 <- bdd %>%
  select(usage="Q15 [4]")
usage5 <- bdd %>%
  select(usage="Q15 [5]")
usage6 <- bdd %>%
  select(usage="Q15 [6]")

usagecor <- cbind(usage1,usage2,usage3,usage4,usage5,usage6)
# on ne veux pas que les 3 usages personnels s'additionne cela donnerais
l'impression que
#que cet usage est plus représenté que les autre,usage3,usage4,usage5,usage6
to_filter_Usage <- function(usage){
  dim <- dim(usage)[1]
  for (i in 1:dim){
    for(j in 1:6){
      if(j != 1 && (usage[i,1] %in% c(1,5,6)) && (usage[i,j] %in% c(1,5,6))){
        usage[i,j] <- NA
      }
      if(j != 2 && (usage[i,2] %in% c(1,5,6)) && (usage[i,j] %in% c(1,5,6))){
        usage[i,j] <- NA
      }
      if(j != 3 && (usage[i,3] %in% c(1,5,6)) && (usage[i,j] %in% c(1,5,6))){
        usage[i,j] <- NA
      }
      if(j != 4 && (usage[i,4] %in% c(1,5,6)) && (usage[i,j] %in% c(1,5,6))){
```

```

        usage[i,j] <- NA
    }
    if(j != 5 && (usage[i,5] %in% c(1,5,6)) && (usage[i,j] %in% c(1,5,6))){
        usage[i,j] <- NA
    }
}
}
return (usage)
}

usagecor <- to_filter_Usage(usagecor)

#attribution des valeur corrigé au different usage
usage1[,1] <- usagecor[,1]
usage2[,1] <- usagecor[,2]
usage3[,1] <- usagecor[,3]
usage4[,1] <- usagecor[,4]
usage5[,1] <- usagecor[,5]
usage6[,1] <- usagecor[,6]

#correspond au nombre d'apparition de chaque type d'usage
usageVoiture <- rbind(usage1,usage2) %>% rbind(usage3) %>% rbind(usage4) %>%
rbind(usage5) %>% rbind(usage6) %>%
    arrange(usage) %>%
    group_by(usage) %>%
    summarise("Voiture"= n())
View(usageVoiture)

#regrouper les type d'utilisation des voiture pour qu'ils correspondent au moto
usageVoiture[1,2] <- usageVoiture[1,2] + usageVoiture[5,2] + usageVoiture[6,2]

#----- Usage des motos -----#

usage1 <- bdd %>%
    select(usage = "Usage 1")
usage2 <- bdd %>%
    select(usage = "...55")
usage3 <- bdd %>%
    select(usage="...66")
usage4 <- bdd %>%
    select(usage="...77")
usage5 <- bdd %>%

```



```

select(usage="...89")

#correspond au nombre d'apparition de chaque type d'usage
# c'est l'usage total il n'est plus utile
usageMoto <- rbind(usage1,usage2) %>% rbind(usage3) %>% rbind(usage4) %>%
rbind(usage5) %>%
  arrange(usage) %>%
  group_by(usage) %>%
  summarise("Nombre_de_personnes"= n())
View(usageMoto)

#-- AFC Usage par véhicule --#

#Usage par moto
usageMoto1 <- usage1 %>%
  arrange(usage) %>%
  group_by(usage) %>%
  summarise("Moto_1"= n())
usageMoto1[4,2] <- 0

usageMoto2 <- usage2 %>%
  arrange(usage) %>%
  group_by(usage) %>%
  summarise("Moto_2"= n())
usageMoto2[4,2] <- 0

usageMoto3 <- usage3 %>%
  arrange(usage) %>%
  group_by(usage) %>%
  summarise("Moto_3"= n())
usageMoto3[4,2] <- 0

usageMoto4 <- usage4 %>%
  arrange(usage) %>%
  group_by(usage) %>%
  summarise("Moto_4"= n())
usageMoto4[4,2] <- 0

usageMoto5 <- usage5 %>%
  filter() %>%

```

```

  arrange(usage) %>%
  group_by(usage) %>%
  summarise("Moto_5"= n())
usageMoto5[4,2] <- 0
#probleme avec les valeur manquante
usageMoto5[2:4,2] <- 0

AFC_usage <- usageVoiture[1:4,2]
AFC_usage[,2] <- usageMoto1[,2]
AFC_usage[,3] <- usageMoto2[,2]
AFC_usage[,4] <- usageMoto3[,2]
AFC_usage[,5] <- usageMoto4[,2]
AFC_usage[,6] <- usageMoto5[,2]

#renome les ligne
#rownames(AFC_usage) <- c("Personnel","Domicile-travail", "Travail", "Course")

#on enlève les course, on pourra essayer en les rajoutant
AFC_usage <- AFC_usage[1:3,]
rownames(AFC_usage) <- c("Loisir","Domicile-travail", "Travail")

# 1. convertir les données en tant que table
dt <- as.table(as.matrix (AFC_usage))

View(dt)

# 2. Graphique
balloonplot(t (dt), main = "AFC_usage", xlab = "", ylab = "",
            label = FALSE, show.margins = FALSE)

# test du khi²
chisq <- chisq.test (AFC_usage)

#mise en place de l'AFC
res_AFC <- CA (AFC_usage, graph = TRUE)

#analyse des valeur propre
eig.val <- get_eigenvalue (res_AFC)

# Contributions des colonnes à la dimension 1
fviz_contrib(res_AFC, choice = "col", axes = 1, top = 10)

```

```

# Contributions des colonnes à la dimension 2
fviz_contrib(res_AFC, choice = "col", axes = 2, top = 10)

# Contributions des lignes à la dimension 1
fviz_contrib(res_AFC, choice = "row", axes = 1, top = 10)
# Contributions des lignes à la dimension 2
fviz_contrib(res_AFC, choice = "row", axes = 2, top = 10)

# repel = TRUE pour éviter le chevauchement de texte
fviz_ca_biplot (res_AFC)
ggplot(res_AFC)

#----- Usage table -----#

# début de la création d'une table avec tout les usages, ça correspondrait a la
table utile pour répondre a cet axe, il faut récupérer tout les usage et ainsi que
le nombre de km associé

usage <- bdd %>%
  select(Identifiant,      nbKmVoiture="Q17      [1]",      usageVoiture1="Q15      [1]",
usageVoiture2="Q15 [2]",
        usageVoiture3="Q15 [3]", usageVoiture4="Q15 [4]", usageVoiture5="Q15 [5]",
        usageVoiture6="Q15 [6]", usageM1="Usage 1", kmM1="Nb km 1",
usageM2="...55",
        kmM2="...56", usageM3="...66", kmM3="...67", usageM4="...77", kmM4="...78",
        usageM5="...89", kmM5="...90")

#fonctions pour transformer les donnees en qq chose de lisible
to_readable_auto <- function(code){
  if (is.na(code))
    return ("")
  if (code == 1)
    return ("Personnel")

```

```
if (code == 2)
  return ("Domicile-travail")
if (code == 3)
  return ("Travail")
if (code == 4)
  return ("Course")
if (code == 5)
  return ("Personnel")
if (code == 6)
  return ("Personnel")
}

to_readable_auto_vector <- Vectorize(to_readable_auto, vectorize.args = "code")

to_readable_moto <- function(code){
  if (is.na(code))
    return ("")
  if (code == 1)
    return ("Personnel")
  if (code == 2)
    return ("Domicile-travail")
  if (code == 3)
    return ("Travail")
}

to_readable_moto_vector <- Vectorize(to_readable_moto, vectorize.args = "code")

usage_readable <- usage %>%
  mutate(usageVoiture1 = to_readable_auto_vector(usageVoiture1)) %>%
  mutate(usageVoiture2 = to_readable_auto_vector(usageVoiture2)) %>%
  mutate(usageVoiture3 = to_readable_auto_vector(usageVoiture3)) %>%
  mutate(usageVoiture4 = to_readable_auto_vector(usageVoiture4)) %>%
  mutate(usageVoiture5 = to_readable_auto_vector(usageVoiture5)) %>%
  mutate(usageVoiture6 = to_readable_auto_vector(usageVoiture6)) %>%
  mutate(usageM1 = to_readable_moto_vector(usageM1)) %>%
  mutate(usageM2 = to_readable_moto_vector(usageM2)) %>%
  mutate(usageM3 = to_readable_moto_vector(usageM3)) %>%
  mutate(usageM4 = to_readable_moto_vector(usageM4)) %>%
  mutate(usageM5 = to_readable_moto_vector(usageM5))

View(usage)
View(usage_readable)
```

```

#-- données usage plot introduction --#

usageVoiturePLOT <- usageVoiture[1:3,2]
usageVoiturePLOT <- mutate(usageVoiturePLOT ,prop = round(Voiture /
sum(usageVoiturePLOT$Voiture) *100,2))%>%
  mutate(ypos = cumsum(prop)- 0.5*prop )

dt <- data.frame(
  type = c("Loisir","Domicile-travail", "Travail"),
  value = usageVoiturePLOT[,1],
  prop = usageVoiturePLOT[,2],
  ypos = usageVoiturePLOT[,3]
)

#On ne met rien sur l'axe x pour pouvoir en faire un cercle
diag_intro <- ggplot(dt, aes(x = "", y = prop, fill = factor(type))) +
  geom_bar(width = 1,stat = "identity") +
  theme_void() +
  labs(fill="Type d'usage",
        x=NULL,
        y=NULL,
        title="Répartition des usages des autos",
        caption="Figure 1: Diagramme R, répartition des usages des autos") +
  geom_text(aes(y = ypos, label = prop), color = "white", size=5,check_overlap =
TRUE) +
  scale_fill_brewer(palette="Dark2") +
  coord_polar("y", start=0)
diag_intro

usageMotoPLOT <- usageMoto[1:3,2]
usageMotoPLOT <- mutate(usageMotoPLOT ,prop = round(Nombre_de_personnes /
sum(usageMotoPLOT$Nombre_de_personnes) *100,2))%>%
  mutate(ypos = cumsum(prop)- 0.5*prop )

dtm <- data.frame(
  type = c("Loisir","Domicile-travail", "Travail"),
  value = usageMotoPLOT[,1],
  prop = usageMotoPLOT[,2],
  ypos = usageMotoPLOT[,3]
)

```

```

#On ne met rien sur l'axe x pour pouvoir en faire un cercle
diag_intro_moto <- ggplot(dtm, aes(x = "", y = prop, fill = factor(type))) +
  geom_bar(width = 1, stat = "identity") +
  theme_void() +
  labs(fill="Type d'usage",
        x=NULL,
        y=NULL,
        title="Répartition des usages des motos",
        caption="Figure 2: Diagramme R, répartition des usages des motos") +
  geom_text(aes(y = ypos, label = prop), color = "white", size=5, check_overlap =
TRUE) +
  scale_fill_brewer(palette="Dark2") +
  coord_polar("y", start=0)
diag_intro_moto

usagePLOT <- usageMoto[1:3,2] + usageVoiture[1:3,2]
usagePLOT <- mutate(usagePLOT, prop = round(Nombre_de_personnes /
sum(usagePLOT$Nombre_de_personnes) *100,2))%>%
  mutate(ypos = cumsum(prop) - 0.5*prop)
usagePLOT

dtall <- data.frame(
  type = c("Loisir", "Domicile-travail", "Travail"),
  value = usagePLOT[,1],
  prop = usagePLOT[,2],
  ypos = usagePLOT[,3]
)

#On ne met rien sur l'axe x pour pouvoir en faire un cercle
diag_intro_moto <- ggplot(dtall, aes(x = "", y = prop, fill = factor(type))) +
  geom_bar(width = 1, stat = "identity") +
  theme_void() +
  labs(fill="Type d'usage",
        x=NULL,
        y=NULL,
        title="Répartition des usages pour les motos et les autos",
        caption="Figure 2: Diagramme R, répartition des usages pour les motos et les
autos") +
  geom_text(aes(y = ypos, label = prop), color = "white", size=5, check_overlap =
TRUE) +
  scale_fill_brewer(palette="Dark2") +
  coord_polar("y", start=0)

```

diag\_intro\_moto

## Annexe 4 : Server Shiny App

```
#
# This is the server logic of a Shiny web application. You can run the
# application by clicking 'Run App' above.
#
# Find out more about building applications with Shiny here:
#
#   http://shiny.rstudio.com/
#

library(shiny)
library("readxl")
library(ggplot2)
library(dplyr)

bdd_auto <- read_excel("C:/Users/alexa/Documents/R/dataScience/Data/bdd_auto-modif.xlsx")
bdd_2_roue <- read_excel("C:/Users/alexa/Documents/R/dataScience/Data/bdd_2_roue.xls")
colnames(bdd_2_roue)[colnames(bdd_2_roue)=="...1"] <- "Identifiant"
colnames(bdd_2_roue)[colnames(bdd_2_roue)=="...2"] <- "Date"
bdd <- merge(bdd_auto, bdd_2_roue, by="Identifiant")

gros_rouleurs <- bdd %>%
  select(Identifiant, moto1="Nb km 1", moto2="...56", moto3="...67",
moto4="...78", kmAuto="Q17 [1]", Nombre_personnes_foyer = "Nombre personnes au
foyer", Situation_familiale = "Situation familiale", Nombre_enfants = "Nombre
enfants")

gros_rouleurs$moto1 <- as.numeric(as.character(gros_rouleurs$moto1))
gros_rouleurs$moto2 <- as.numeric(as.character(gros_rouleurs$moto2))
gros_rouleurs$moto3 <- as.numeric(as.character(gros_rouleurs$moto3))
gros_rouleurs$moto4 <- as.numeric(as.character(gros_rouleurs$moto4))
gros_rouleurs$Nombre_enfants <-
as.numeric(as.character(gros_rouleurs$Nombre_enfants))
```

```
gros_rouleurs[is.na(gros_rouleurs)] <- 0
gr_sum <- gros_rouleurs %>%
  mutate(kmMoto = rowSums(.[2:5])) %>%
  arrange(desc(kmMoto))

gr_sum_total <- gr_sum %>%
  mutate(kmTotaux = kmMoto + kmAuto)

isGrosRouleur_moto <- function(km){
  if (km <= moyenne_km_an)
    return (0)
  if (km > moyenne_km_an)
    return (1)
}
isGrosRouleur_moto_vector <- Vectorize(isGrosRouleur_moto, vectorize.args = "km")

isGrosRouleur_auto <- function(km){
  if (km <= moyenne_km_an)
    return (0)
  if (km > moyenne_km_an)
    return (1)
}
isGrosRouleur_auto_vector <- Vectorize(isGrosRouleur_auto, vectorize.args = "km")

isAssure <- function(num){
  if (num == "1")
    return (1)
  if (num == "2")
    return (0)
}
isAssure_vector <- Vectorize(isAssure, vectorize.args = "num")

situation_familiale <- function(num){
  if (num == "1")
    return ("Seul(e) sans enfant")
  if (num == "2")
    return ("Seul(e) avec enfant(s)")
  if (num == "3")
    return ("En couple sans enfant")
  if (num == "4")
```



```

        return ("En couple avec enfant(s)")
    if (num == "5")
        return ("Autre")
}

situation_familiale_vector <- Vectorize(situation_familiale, vectorize.args =
"num")

is_en_couple <- function(situation){
    if (situation == "Seul(e) sans enfant")
        return ("Celibataire")
    if (situation == "Seul(e) avec enfant(s)")
        return ("Celibataire")
    if (situation == "En couple sans enfant")
        return ("En couple")
    if (situation == "En couple avec enfant(s)")
        return ("En couple")
    if (situation == "Autre")
        return ("Celibataire")
}

is_en_couple_vector <- Vectorize(is_en_couple, vectorize.args = "situation")

hasEnfants <- function(situation){
    if (situation == "Seul(e) sans enfant")
        return (0)
    if (situation == "Seul(e) avec enfant(s)")
        return (1)
    if (situation == "En couple sans enfant")
        return (0)
    if (situation == "En couple avec enfant(s)")
        return (1)
    if (situation == "Autre")
        return (0)
}

hasEnfants_vector <- Vectorize(hasEnfants, vectorize.args = "situation")

type_rouleur <- gr_sum_total %>%
    mutate(gros_rouleur_moto = isGrosRouleur_moto_vector(kmMoto)) %>%
    mutate(gros_rouleur_auto = isGrosRouleur_auto_vector(kmAuto)) %>%
    mutate(situation_familiale = situation_familiale_vector(Situation_familiale))%>%

```

```

mutate(en_couple = is_en_couple_vector(situation_familiale)) %>%
mutate(avec_enfant = hasEnfants_vector(situation_familiale)) %>%

select(Identifiant,gros_rouleur_moto,gros_rouleur_auto,en_couple,Nombre_enfants,kmT
otaux,kmAuto,kmMoto)

situation_couple <- type_rouleur %>%
  group_by(en_couple) %>%
  summarise(moy_km_moto = round(mean(kmAuto, na.rm=TRUE),2),
            moy_km_auto = round(mean(kmMoto),2),
            moy_km_total = round(mean(kmTotaux),2)
  )

situation_famille <- type_rouleur %>%
  group_by(Nombre_enfants) %>%
  summarise(moy_km_moto = round(mean(kmAuto, na.rm=TRUE),2),
            moy_km_auto = round(mean(kmMoto),2),
            moy_km_total = round(mean(kmTotaux),2)
  ) %>%
  filter(Nombre_enfants < 10)
View(situation_famille)

#####
# Define server logic required to draw a histogram
shinyServer(function(input, output) {

  output$plot1 <- renderPlot({

    #INPUT
    moyenne_km_an <- input$separation

    sup_moto <- gr_sum %>%
      group_by(kmMoto) %>%
      filter(kmMoto >= moyenne_km_an)
    inf_moto <- gr_sum %>%
      group_by(kmMoto) %>%
      filter(kmMoto < moyenne_km_an)

    sup_auto <- gr_sum %>%

```

```

    group_by(kmAuto) %>%
    filter(kmAuto >= moyenne_km_an)
inf_auto <- gr_sum %>%
    group_by(kmAuto) %>%
    filter(kmAuto < moyenne_km_an)

    effectifs <- data.frame(stat=c("Total etudie", "Gros rouleurs en
moto", "Petits rouleurs en moto", "Gros rouleurs en auto", "Petits rouleurs en auto"),
effectif=c(dim(bdd) [1], dim(sup_moto) [1], dim(inf_moto) [1], dim(sup_auto) [1], dim(inf_a
uto) [1]))

    pourcentage_gros_rouleurs_fr <- effectifs[-5,][-3,][-1,] %>%
    mutate(pourcentage = round(effectif / dim(bdd) [1] * 100,2))

    ggplot(data=pourcentage_gros_rouleurs_fr, aes(x=stat, y=pourcentage, fill
=factor(stat))) +
    geom_bar(stat="identity") +
    scale_fill_brewer(palette="Set1") +
    labs(fill="Vehicule",
         x=NULL,
         y="%",
         title="Repartition du nombre de gros rouleurs en fonction du
vehicule") +
    geom_text(aes(y = pourcentage / 2, label = pourcentage ), color =
"white", size=5, check_overlap = TRUE)+
    theme(legend.position="none")
  })

  output$plot2 <- renderPlot({

    kmEtudie <- input$km

    ggplot(data=situation_couple, aes_string(x="en_couple", y=kmEtudie, fill =
"en_couple")) +
    geom_bar(stat="identity") +
    scale_fill_brewer(palette="Set2") +
    labs(fill="situation",
         x="Situation",
         y="Km",
         title="Moyenne de km parcourus en fonction de la situation
maritale") +

```

```
      geom_text(aes_string(y = 10000 , label = kmEtudie ), color = "white",
size=5) +
      theme(legend.position="none")
    })

    output$plot3 <- renderPlot({

      kmEtudie <- input$kmFamille

      ggplot(data=situation_famille, aes_string(x="Nombre_enfants", y=kmEtudie,
fill = factor("Nombre_enfants")))) +
        geom_bar(stat="identity") +
        scale_fill_brewer(palette="Set2") +
        labs(fill="situation",
              x="Situation",
              y="Km",
              title="Moyenne de km parcourus en fonction de la situation
familiale") +
        geom_text(aes_string(y = 7000 , label = kmEtudie ), color = "white",
size=5) +
        theme(legend.position="none")
    })
  })
```

## Annexe 5: UI Shiny App

```
#
# This is the user-interface definition of a Shiny web application. You can
# run the application by clicking 'Run App' above.
#
# Find out more about building applications with Shiny here:
#
#   http://shiny.rstudio.com/
#

library(shiny)

# Define UI for application that draws a histogram
shinyUI(
  fluidPage(

    # Application title
    titlePanel("Comparaison de l'utilisation de la voiture et celle du
2-roues"),

    mainPanel(
      tabsetPanel(
        tabPanel("Frequence",
          wellPanel(
            radioButtons("separation", "Kilometrage de separation
des types de rouleurs:",
              c("Moyenne de l'echantillon" = "27718.2",
                "Moyenne francaise" = "17000")
            )
          ),
        tabsetPanel(
          tabPanel("Repartition des gros rouleurs",

            plotOutput("plot1")
          ),
          tabPanel("Situation Maritale",
            wellPanel(
```

```
radioButtons("km", "Km a comparer",
             c("Auto" = "moy_km_auto",
               "Moto" = "moy_km_moto",
               "Total" = "moy_km_total")
            ),
plotOutput("plot2")
),
tabPanel("Situation Familiale",
        wellPanel(
            radioButtons("kmFamille", "Km a comparer",
                         c("Auto" = "moy_km_auto",
                           "Moto" = "moy_km_moto",
                           "Total" = "moy_km_total")
                        ),
            plotOutput("plot3")
        )
    ),
tabPanel("Usages",
        plotOutput("")
    )
)
)
)
```