

```

# Load required libraries
library(readxl)
library(stats)
library(data.table)
library(dplyr)
library(ggplot2)
library(survey)
library(stargazer)
library(ggplot2)

# Read the Excel file
V_CG <- read_excel("V_CG.xlsx")
V_TG <- read_excel("V_TG.xlsx")
W_CG <- read_excel("W_CG.xlsx")
W_TG <- read_excel("W_TG.xlsx")

V_CG$Horodateur <- NULL
V_TG$Horodateur <- NULL
W_CG$Horodateur <- NULL
W_TG$Horodateur <- NULL

V_TG$`Wat is uw favoriete seizoen?` <- NULL
V_TG$`Wat is uw favoriete keuken?` <- NULL
V_TG$`Als er maar één van deze twee gerechten bestond, welke zou u
dan kiezen?` <- NULL
W_TG$`Quel est votre saison préférée ?` <- NULL
W_TG$`Quelle est votre type de cuisine préféré ?` <- NULL
W_TG$`Si seulement un de ces deux plats devait exister, lequel
choisiriez-vous ?` <- NULL

# Column numbers to be renamed
cols_to_rename <- 12:51
# New names for simplicity
new_col_names <- c("No_1", "15_1", "25_1", "35_1", "50_1",
                  "No_2", "15_2", "25_2", "35_2", "50_2",
                  "No_3", "15_3", "25_3", "35_3", "50_3",
                  "No_4", "15_4", "25_4", "35_4", "50_4",
                  "No_5", "15_5", "25_5", "35_5", "50_5",
                  "No_6", "15_6", "25_6", "35_6", "50_6",
                  "No_7", "15_7", "25_7", "35_7", "50_7",
                  "No_8", "15_8", "25_8", "35_8", "50_8")

# Rename the columns
colnames(V_CG)[cols_to_rename] <- new_col_names
colnames(W_CG)[cols_to_rename] <- new_col_names
colnames(V_TG)[cols_to_rename] <- new_col_names
colnames(W_TG)[cols_to_rename] <- new_col_names

#Merging all distances responses regarding the allowances in a
single column
V_CG <- V_CG %>%

```

```

mutate(NoAllowance = coalesce(V_CG$No_1, V_CG$No_2, V_CG$No_3,
V_CG$No_4, V_CG$No_5, V_CG$No_6, V_CG$No_7, V_CG$No_8))
V_CG <- V_CG %>%
  mutate(Allowance15 = coalesce(V_CG$"15_1", V_CG$"15_2",
V_CG$"15_3", V_CG$"15_4", V_CG$"15_5", V_CG$"15_6", V_CG$"15_7",
V_CG$"15_8"))
V_CG <- V_CG %>%
  mutate(Allowance25 = coalesce(V_CG$"25_1", V_CG$"25_2",
V_CG$"25_3", V_CG$"25_4", V_CG$"25_5", V_CG$"25_6", V_CG$"25_7",
V_CG$"25_8"))
V_CG <- V_CG %>%
  mutate(Allowance35 = coalesce(V_CG$"35_1", V_CG$"35_2",
V_CG$"35_3", V_CG$"35_4", V_CG$"35_5", V_CG$"35_6", V_CG$"35_7",
V_CG$"35_8"))
V_CG <- V_CG %>%
  mutate(Allowance50 = coalesce(V_CG$"50_1", V_CG$"50_2",
V_CG$"50_3", V_CG$"50_4", V_CG$"50_5", V_CG$"50_6", V_CG$"50_7",
V_CG$"50_8"))

V_TG <- V_TG %>%
  mutate(NoAllowance = coalesce(V_TG$No_1, V_TG$No_2, V_TG$No_3,
V_TG$No_4, V_TG$No_5, V_TG$No_6, V_TG$No_7, V_TG$No_8))
V_TG <- V_TG %>%
  mutate(Allowance15 = coalesce(V_TG$"15_1", V_TG$"15_2",
V_TG$"15_3", V_TG$"15_4", V_TG$"15_5", V_TG$"15_6", V_TG$"15_7",
V_TG$"15_8"))
V_TG <- V_TG %>%
  mutate(Allowance25 = coalesce(V_TG$"25_1", V_TG$"25_2",
V_TG$"25_3", V_TG$"25_4", V_TG$"25_5", V_TG$"25_6", V_TG$"25_7",
V_TG$"25_8"))
V_TG <- V_TG %>%
  mutate(Allowance35 = coalesce(V_TG$"35_1", V_TG$"35_2",
V_TG$"35_3", V_TG$"35_4", V_TG$"35_5", V_TG$"35_6", V_TG$"35_7",
V_TG$"35_8"))
V_TG <- V_TG %>%
  mutate(Allowance50 = coalesce(V_TG$"50_1", V_TG$"50_2",
V_TG$"50_3", V_TG$"50_4", V_TG$"50_5", V_TG$"50_6", V_TG$"50_7",
V_TG$"50_8"))

W_CG <- W_CG %>%
  mutate(NoAllowance = coalesce(W_CG$No_1, W_CG$No_2, W_CG$No_3,
W_CG$No_4, W_CG$No_5, W_CG$No_6, W_CG$No_7, W_CG$No_8))
W_CG <- W_CG %>%
  mutate(Allowance15 = coalesce(W_CG$"15_1", W_CG$"15_2",
W_CG$"15_3", W_CG$"15_4", W_CG$"15_5", W_CG$"15_6", W_CG$"15_7",
W_CG$"15_8"))
W_CG <- W_CG %>%
  mutate(Allowance25 = coalesce(W_CG$"25_1", W_CG$"25_2",
W_CG$"25_3", W_CG$"25_4", W_CG$"25_5", W_CG$"25_6", W_CG$"25_7",
W_CG$"25_8"))
W_CG <- W_CG %>%
  mutate(Allowance35 = coalesce(W_CG$"35_1", W_CG$"35_2",
W_CG$"35_3", W_CG$"35_4", W_CG$"35_5", W_CG$"35_6", W_CG$"35_7",
W_CG$"35_8"))

```

```

W_CG <- W_CG %>%
  mutate(Allowance50 = coalesce(W_CG$"50_1", W_CG$"50_2",
W_CG$"50_3", W_CG$"50_4", W_CG$"50_5", W_CG$"50_6", W_CG$"50_7",
W_CG$"50_8"))

W_TG <- W_TG %>%
  mutate(NoAllowance = coalesce(W_TG$No_1, W_TG$No_2, W_TG$No_3,
W_TG$No_4, W_TG$No_5, W_TG$No_6, W_TG$No_7, W_TG$No_8))
W_TG <- W_TG %>%
  mutate(Allowance15 = coalesce(W_TG$"15_1", W_TG$"15_2",
W_TG$"15_3", W_TG$"15_4", W_TG$"15_5", W_TG$"15_6", W_TG$"15_7",
W_TG$"15_8"))
W_TG <- W_TG %>%
  mutate(Allowance25 = coalesce(W_TG$"25_1", W_TG$"25_2",
W_TG$"25_3", W_TG$"25_4", W_TG$"25_5", W_TG$"25_6", W_TG$"25_7",
W_TG$"25_8"))
W_TG <- W_TG %>%
  mutate(Allowance35 = coalesce(W_TG$"35_1", W_TG$"35_2",
W_TG$"35_3", W_TG$"35_4", W_TG$"35_5", W_TG$"35_6", W_TG$"35_7",
W_TG$"35_8"))
W_TG <- W_TG %>%
  mutate(Allowance50 = coalesce(W_TG$"50_1", W_TG$"50_2",
W_TG$"50_3", W_TG$"50_4", W_TG$"50_5", W_TG$"50_6", W_TG$"50_7",
W_TG$"50_8"))

```

#Deleting the now useless columns

```

cols_to_delete <- 12:51
V_CG <- V_CG[, -cols_to_delete, with = FALSE]
V_TG <- V_TG[, -cols_to_delete, with = FALSE]
W_CG <- W_CG[, -cols_to_delete, with = FALSE]
W_TG <- W_TG[, -cols_to_delete, with = FALSE]

```

```

setnames(V_CG, "Bent u een:", "Sex")
setnames(V_TG, "Bent u een:", "Sex")
setnames(W_CG, "Vous êtes un/une :", "Sex")
setnames(W_TG, "Vous êtes un/une :", "Sex")

```

```

setnames(V_CG, "Hoe oud bent u?", "Age")
setnames(V_TG, "Hoe oud bent u?", "Age")
setnames(W_CG, "Quel âge avez-vous ?", "Age")
setnames(W_TG, "Quel âge avez-vous ?", "Age")

```

```

setnames(V_CG, "Wat is uw postcode?", "ZIP")
setnames(V_TG, "Wat is uw postcode?", "ZIP")
setnames(W_CG, "Quel est votre code postal ?", "ZIP")
setnames(W_TG, "Quel est votre code postal ?", "ZIP")

```

```

setnames(V_CG, "Bent u een student?", "Student")
setnames(V_TG, "Bent u een student?", "Student")
setnames(W_CG, "Êtes-vous étudiant ?", "Student")
setnames(W_TG, "Êtes-vous étudiant ?", "Student")

```

```

setnames(V_CG, "Op een schaal van 0 tot 10, hoe Vlaams voelt u

```

```
zich?", "W/F feeling")
setnames(V_TG, "Op een schaal van 0 tot 10, hoe Vlaams voelt u
zich?", "W/F feeling")
setnames(W_CG, "Sur une échelle de 0 à 10, à quel point vous sentez-
vous wallon ?", "W/F feeling")
setnames(W_TG, "Sur une échelle de 0 à 10, à quel point vous sentez-
vous wallon ?", "W/F feeling")
```

```
setnames(V_CG, "Wat is uw hoofdvervoerwijze om op uw werk te komen?
(het vervoermiddel dat u het meest gebruikt gedurende het jaar en
waarmee u de grootste afstand aflegt)", "Transport")
setnames(V_TG, "Wat is uw hoofdvervoerwijze om op uw werk te komen?
(het vervoermiddel dat u het meest gebruikt gedurende het jaar en
waarmee u de grootste afstand aflegt)", "Transport")
setnames(W_CG, "Quel est votre mode de transport dominant pour vous
rendre à votre travail ? (celui que vous prenez le plus souvent sur
l'année et qui vous fait parcourir la plus grande distance)",
"Transport")
setnames(W_TG, "Quel est votre mode de transport dominant pour vous
rendre à votre travail ? (celui que vous prenez le plus souvent sur
l'année et qui vous fait parcourir la plus grande distance)",
"Transport")
```

```
setnames(V_CG, "Hebt u een bedrijfswagen?", "Company car")
setnames(V_TG, "Hebt u een bedrijfswagen?", "Company car")
setnames(W_CG, "Disposez-vous d'une voiture de société ?", "Company
car")
setnames(W_TG, "Disposez-vous d'une voiture de société ?", "Company
car")
```

```
setnames(V_CG, "Hebt u een tankkaart dankzij uw werk?", "Fuel card")
setnames(V_TG, "Hebt u een tankkaart dankzij uw werk?", "Fuel card")
setnames(W_CG, "Disposez-vous d'une carte carburant par votre
travail ?", "Fuel card")
setnames(W_TG, "Disposez-vous d'une carte carburant par votre
travail ?", "Fuel card")
```

```
setnames(V_CG, "Als u naar uw werk fietst, krijgt u dan een
fietsvergoeding?", "Bike_allowance")
setnames(V_TG, "Als u naar uw werk fietst, krijgt u dan een
fietsvergoeding?", "Bike_allowance")
setnames(W_CG, "Si vous vous rendez au travail à vélo, bénéficiez-
vous d'une indemnité vélo ?", "Bike_allowance")
setnames(W_TG, "Si vous vous rendez au travail à vélo, bénéficiez-
vous d'une indemnité vélo ?", "Bike_allowance")
```

```
colnames(V_CG)[10] <- "Weekly commutes"
colnames(V_TG)[10] <- "Weekly commutes"
colnames(W_CG)[10] <- "Weekly commutes"
colnames(W_TG)[10] <- "Weekly commutes"
```

```
setnames(V_CG, "Hoe ver van uw huis werkt u?", "Distance")
setnames(V_TG, "Hoe ver van uw huis werkt u?", "Distance")
setnames(W_CG, "À quelle distance de votre domicile travaillez-
```

```

vous ?", "Distance")
setnames(W_TG, "À quelle distance de votre domicile travaillez-
vous ?", "Distance")

# Add a column of 'Yes' for the treatment group and of 'No' for the
control group
V_CG$Treatment <- "No"
V_TG$Treatment <- "Yes"
W_CG$Treatment <- "No"
W_TG$Treatment <- "Yes"

#Merge the control and treatment group for Walloons and Flemings
VL_db <- rbind(V_CG, V_TG)
WA_db <- rbind(W_CG, W_TG)

##Translation
#VL
VL_db$Sex <- gsub("Man", "M", VL_db$Sex)
VL_db$Sex <- gsub("Vrouw", "F", VL_db$Sex)
VL_db$Sex <- gsub("Andere", "O", VL_db$Sex)
VL_db$Sex <- gsub("Dat zeg ik liever niet", "NP", VL_db$Sex)

VL_db$Age <- gsub(" jaar", "", VL_db$Age)
VL_db$Age <- gsub(" en ouder", "+", VL_db$Age)

VL_db$Student <- gsub("Ja", "Yes", VL_db$Student)
VL_db$Student <- gsub("Neen", "No", VL_db$Student)

VL_db$Transport <- gsub("Wagen", "Car", VL_db$Transport)
VL_db$Transport <- gsub("Camionette", "Car", VL_db$Transport)
VL_db$Transport <- gsub("carpoolen", "Car", VL_db$Transport)
VL_db$Transport <- gsub("vrachtwagen", "Car", VL_db$Transport)
VL_db$Transport <- gsub("Vrachtwagen", "Car", VL_db$Transport)
VL_db$Transport <- gsub("Wagen maar carpool", "Car",
VL_db$Transport)
VL_db$Transport <- gsub("Carpool", "Car", VL_db$Transport)
VL_db$Transport <- gsub("[]", "", VL_db$Transport)
VL_db$Transport <- gsub("Openbaar vervoer trein, bus, metro, tram",
"Public transport", VL_db$Transport)
VL_db$Transport <- gsub("Fiets, elektrische fietsFiets, elektrische
fiets", "Fiets, elektrische fiets", VL_db$Transport)
VL_db$Transport <- gsub("Fiets, elektrische fiets", "Bike",
VL_db$Transport)
VL_db$Transport <- gsub("Motor, scooter,...", "Moto",
VL_db$Transport)
VL_db$Transport <- gsub("Te voet", "Walking", VL_db$Transport)
VL_db$Transport <- gsub("Speedpedelec", "Bike", VL_db$Transport)

VL_db$`Company car` <- gsub("Ja", "Yes", VL_db$`Company car`)
VL_db$`Company car` <- gsub("Neen", "No", VL_db$`Company car`)

VL_db$`Fuel card` <- gsub("Ja", "Yes", VL_db$`Fuel card`)
VL_db$`Fuel card` <- gsub("Neen", "No", VL_db$`Fuel card`)

```

```

VL_db$`Bike_allowance` <- gsub("Ja", "Yes", VL_db$`Bike_allowance`)
VL_db$`Bike_allowance` <- gsub("Neen", "No", VL_db$`Bike_allowance`)
VL_db$`Bike_allowance` <- gsub("Ik fiets naar mijn werk niet",
"Other", VL_db$`Bike_allowance`)
VL_db$`Bike_allowance` <- gsub("Ik fiets niet naar mijn werk",
"Other", VL_db$`Bike_allowance`)

```

```

VL_db$`Weekly commutes` <- gsub("[()]", "", VL_db$`Weekly commutes`)
VL_db$`Weekly commutes` <- gsub(" traditioneel voltijds", "",
VL_db$`Weekly commutes`)
VL_db$`Weekly commutes` <- gsub("Meer dan 6", "6+", VL_db$`Weekly
commutes`)

```

```

VL_db$`Distance` <- gsub("meer dan 60 km", "60+", VL_db$`Distance`)
VL_db$`Distance` <- gsub(" km", "", VL_db$`Distance`)
VL_db$`Distance` <- gsub(" 50-60", "51-60", VL_db$`Distance`)

```

```

VL_db$NoAllowance <- gsub("Wagen", "Car", VL_db$NoAllowance)
VL_db$NoAllowance <- gsub("Camionette", "Car", VL_db$NoAllowance)
VL_db$NoAllowance <- gsub("carpoolen", "Car", VL_db$NoAllowance)
VL_db$NoAllowance <- gsub("vrachtwagen", "Car", VL_db$NoAllowance)
VL_db$NoAllowance <- gsub("Vrachtwagen", "Car", VL_db$NoAllowance)
VL_db$NoAllowance <- gsub("Wagen maar carpool", "Car",
VL_db$NoAllowance)
VL_db$NoAllowance <- gsub("Carpool", "Car", VL_db$NoAllowance)
VL_db$NoAllowance <- gsub("[()]", "", VL_db$NoAllowance)
VL_db$NoAllowance <- gsub("Openbaar vervoer trein, bus, metro,
tram", "Public transport", VL_db$NoAllowance)
VL_db$NoAllowance <- gsub("Fiets, elektrische fietsFiets,
elektrische fiets", "Fiets, elektrische fiets", VL_db$NoAllowance)
VL_db$NoAllowance <- gsub("Fiets, elektrische fiets", "Bike",
VL_db$NoAllowance)
VL_db$NoAllowance <- gsub("Step", "Bike", VL_db$NoAllowance)
VL_db$NoAllowance <- gsub("Motor, scooter,...", "Moto",
VL_db$NoAllowance)
VL_db$NoAllowance <- gsub("Te voet", "Walking", VL_db$NoAllowance)
VL_db$NoAllowance <- gsub("Speedpedelec", "Bike", VL_db$NoAllowance)

```

```

VL_db$Allowance15 <- gsub("Wagen", "Car", VL_db$Allowance15)
VL_db$Allowance15 <- gsub("Camionette", "Car", VL_db$Allowance15)
VL_db$Allowance15 <- gsub("carpoolen", "Car", VL_db$Allowance15)
VL_db$Allowance15 <- gsub("vrachtwagen", "Car", VL_db$Allowance15)
VL_db$Allowance15 <- gsub("Vrachtwagen", "Car", VL_db$Allowance15)
VL_db$Allowance15 <- gsub("Wagen maar carpool", "Car",
VL_db$Allowance15)
VL_db$Allowance15 <- gsub("Carpool", "Car", VL_db$Allowance15)
VL_db$Allowance15 <- gsub("[()]", "", VL_db$Allowance15)
VL_db$Allowance15 <- gsub("Openbaar vervoer trein, bus, metro,
tram", "Public transport", VL_db$Allowance15)
VL_db$Allowance15 <- gsub("Fiets, elektrische fietsFiets,
elektrische fiets", "Fiets, elektrische fiets", VL_db$Allowance15)
VL_db$Allowance15 <- gsub("Fiets, elektrische fiets", "Bike",
VL_db$Allowance15)

```

```

VL_db$Allowance15 <- gsub("Step", "Bike", VL_db$Allowance15)
VL_db$Allowance15 <- gsub("Motor, scooter,...", "Moto",
VL_db$Allowance15)
VL_db$Allowance15 <- gsub("Te voet", "Walking", VL_db$Allowance15)
VL_db$Allowance15 <- gsub("Speedpedelec", "Bike", VL_db$Allowance15)

```

```

VL_db$Allowance25 <- gsub("Wagen", "Car", VL_db$Allowance25)
VL_db$Allowance25 <- gsub("Camionette", "Car", VL_db$Allowance25)
VL_db$Allowance25 <- gsub("carpoolen", "Car", VL_db$Allowance25)
VL_db$Allowance25 <- gsub("vrachtwagen", "Car", VL_db$Allowance25)
VL_db$Allowance25 <- gsub("Vrachtwagen", "Car", VL_db$Allowance25)
VL_db$Allowance25 <- gsub("Wagen maar carpool", "Car",
VL_db$Allowance25)
VL_db$Allowance25 <- gsub("Carpool", "Car", VL_db$Allowance25)
VL_db$Allowance25 <- gsub("[()]", "", VL_db$Allowance25)
VL_db$Allowance25 <- gsub("Openbaar vervoer trein, bus, metro,
tram", "Public transport", VL_db$Allowance25)
VL_db$Allowance25 <- gsub("Fiets, elektrische fietsFiets,
elektrische fiets", "Fiets, elektrische fiets", VL_db$Allowance25)
VL_db$Allowance25 <- gsub("Fiets, elektrische fiets", "Bike",
VL_db$Allowance25)
VL_db$Allowance25 <- gsub("Step", "Bike", VL_db$Allowance25)
VL_db$Allowance25 <- gsub("Motor, scooter,...", "Moto",
VL_db$Allowance25)
VL_db$Allowance25 <- gsub("Te voet", "Walking", VL_db$Allowance25)
VL_db$Allowance25 <- gsub("Speedpedelec", "Bike", VL_db$Allowance25)

```

```

VL_db$Allowance35 <- gsub("Wagen", "Car", VL_db$Allowance35)
VL_db$Allowance35 <- gsub("Camionette", "Car", VL_db$Allowance35)
VL_db$Allowance35 <- gsub("carpoolen", "Car", VL_db$Allowance35)
VL_db$Allowance35 <- gsub("vrachtwagen", "Car", VL_db$Allowance35)
VL_db$Allowance35 <- gsub("Vrachtwagen", "Car", VL_db$Allowance35)
VL_db$Allowance35 <- gsub("Wagen maar carpool", "Car",
VL_db$Allowance35)
VL_db$Allowance35 <- gsub("Carpool", "Car", VL_db$Allowance35)
VL_db$Allowance35 <- gsub("[()]", "", VL_db$Allowance35)
VL_db$Allowance35 <- gsub("Openbaar vervoer trein, bus, metro,
tram", "Public transport", VL_db$Allowance35)
VL_db$Allowance35 <- gsub("Fiets, elektrische fietsFiets,
elektrische fiets", "Fiets, elektrische fiets", VL_db$Allowance35)
VL_db$Allowance35 <- gsub("Fiets, elektrische fiets", "Bike",
VL_db$Allowance35)
VL_db$Allowance35 <- gsub("Step", "Bike", VL_db$Allowance35)
VL_db$Allowance35 <- gsub("Motor, scooter,...", "Moto",
VL_db$Allowance35)
VL_db$Allowance35 <- gsub("Te voet", "Walking", VL_db$Allowance35)
VL_db$Allowance35 <- gsub("Speedpedelec", "Bike", VL_db$Allowance35)

```

```

VL_db$Allowance50 <- gsub("Wagen", "Car", VL_db$Allowance50)
VL_db$Allowance50 <- gsub("Camionette", "Car", VL_db$Allowance50)
VL_db$Allowance50 <- gsub("carpoolen", "Car", VL_db$Allowance50)
VL_db$Allowance50 <- gsub("vrachtwagen", "Car", VL_db$Allowance50)
VL_db$Allowance50 <- gsub("Vrachtwagen", "Car", VL_db$Allowance50)

```

```

VL_db$Allowance50 <- gsub("Wagen maar carpool", "Car",
VL_db$Allowance50)
VL_db$Allowance50 <- gsub("Carpool", "Car", VL_db$Allowance50)
VL_db$Allowance50 <- gsub("[()]", "", VL_db$Allowance50)
VL_db$Allowance50 <- gsub("Openbaar vervoer trein, bus, metro,
tram", "Public transport", VL_db$Allowance50)
VL_db$Allowance50 <- gsub("Fiets, elektrische fietsFiets,
elektrische fiets", "Fiets, elektrische fiets", VL_db$Allowance50)
VL_db$Allowance50 <- gsub("Fiets, elektrische fiets", "Bike",
VL_db$Allowance50)
VL_db$Allowance50 <- gsub("Step", "Bike", VL_db$Allowance50)
VL_db$Allowance50 <- gsub("Motor, scooter,...", "Moto",
VL_db$Allowance50)
VL_db$Allowance50 <- gsub("Te voet", "Walking", VL_db$Allowance50)
VL_db$Allowance50 <- gsub("Speedpedelec", "Bike", VL_db$Allowance50)

```

```
#WA
```

```

WA_db$Sex <- gsub("Homme", "M", WA_db$Sex)
WA_db$Sex <- gsub("Femme", "F", WA_db$Sex)
WA_db$Sex <- gsub("Autre", "O", WA_db$Sex)
WA_db$Sex <- gsub("Je ne souhaite pas le préciser", "NP", WA_db$Sex)

```

```

WA_db$Age <- gsub(" ans","", WA_db$Age)
WA_db$Age <- gsub(" et plus","+", WA_db$Age)

```

```

WA_db$Student <- gsub("Oui", "Yes", WA_db$Student)
WA_db$Student <- gsub("Non", "No", WA_db$Student)

```

```

WA_db$Transport <- gsub("Voiture", "Car", WA_db$Transport)
WA_db$Transport <- gsub("Voiture électrique", "Car",
WA_db$Transport)
WA_db$Transport <- gsub("Camion de société", "Car", WA_db$Transport)
WA_db$Transport <- gsub("Camion", "Car", WA_db$Transport)
WA_db$Transport <- gsub("[()]", "", WA_db$Transport)
WA_db$Transport <- gsub("Transport en commun train, bus, métro,
tram", "Public transport", WA_db$Transport)
WA_db$Transport <- gsub("Vélo, vélo électrique", "Bike",
WA_db$Transport)
WA_db$Transport <- gsub("Moto, scooter,...", "Moto",
WA_db$Transport)
WA_db$Transport <- gsub("Marche à pied", "Walking", WA_db$Transport)

```

```

WA_db$`Company car` <- gsub("Oui", "Yes", WA_db$`Company car`)
WA_db$`Company car` <- gsub("Non", "No", WA_db$`Company car`)

```

```

WA_db$`Fuel card` <- gsub("Oui", "Yes", WA_db$`Fuel card`)
WA_db$`Fuel card` <- gsub("Non", "No", WA_db$`Fuel card`)

```

```

WA_db$`Bike_allowance` <- gsub("Oui", "Yes", WA_db$`Bike_allowance`)
WA_db$`Bike_allowance` <- gsub("Non", "No", WA_db$`Bike_allowance`)
WA_db$`Bike_allowance` <- gsub("Je ne rends pas au travail à vélo",
"Other", WA_db$`Bike_allowance`)

```

```

WA_db$`Weekly commutes` <- gsub("[()]", "", WA_db$`Weekly commutes`)

```

```

WA_db$`Weekly commutes` <- gsub(".0", "", WA_db$`Weekly commutes`)
WA_db$`Weekly commutes` <- gsub(" temps plein classique", "",
WA_db$`Weekly commutes`)
WA_db$`Weekly commutes` <- gsub("Plus de 6", "6+", WA_db$`Weekly
commutes`)

```

```

WA_db$`Distance` <- gsub("[+]", "", WA_db$`Distance`)
WA_db$`Distance` <- gsub(" de 60", "60+", WA_db$`Distance`)
WA_db$`Distance` <- gsub(" km", "", WA_db$`Distance`)
WA_db$`Distance` <- gsub(" 50-60", "51-60", WA_db$`Distance`)

```

```

WA_db$NoAllowance <- gsub("Voiture", "Car", WA_db$NoAllowance)
WA_db$NoAllowance <- gsub("Voiture électrique", "Car",
WA_db$NoAllowance)
WA_db$NoAllowance <- gsub("Camion de société", "Car",
WA_db$NoAllowance)
WA_db$NoAllowance <- gsub("Camion", "Car", WA_db$NoAllowance)
WA_db$NoAllowance <- gsub("[()]", "", WA_db$NoAllowance)
WA_db$NoAllowance <- gsub("Transport en commun train, bus, métro,
tram", "Public transport", WA_db$NoAllowance)
WA_db$NoAllowance <- gsub("Vélo, vélo électrique", "Bike",
WA_db$NoAllowance)
WA_db$NoAllowance <- gsub("Moto, scooter,...", "Moto",
WA_db$NoAllowance)
WA_db$NoAllowance <- gsub("Marche à pied", "Walking",
WA_db$NoAllowance)

```

```

WA_db$Allowance15 <- gsub("Voiture", "Car", WA_db$Allowance15)
WA_db$Allowance15 <- gsub("Voiture électrique", "Car",
WA_db$Allowance15)
WA_db$Allowance15 <- gsub("Camion de société", "Car",
WA_db$Allowance15)
WA_db$Allowance15 <- gsub("Camion", "Car", WA_db$Allowance15)
WA_db$Allowance15 <- gsub("[()]", "", WA_db$Allowance15)
WA_db$Allowance15 <- gsub("Transport en commun train, bus, métro,
tram", "Public transport", WA_db$Allowance15)
WA_db$Allowance15 <- gsub("Vélo, vélo électrique", "Bike",
WA_db$Allowance15)
WA_db$Allowance15 <- gsub("Trottinette électrique", "Bike",
WA_db$Allowance15)
WA_db$Allowance15 <- gsub("Moto, scooter,...", "Moto",
WA_db$Allowance15)
WA_db$Allowance15 <- gsub("Marche à pied", "Walking",
WA_db$Allowance15)

```

```

WA_db$Allowance25 <- gsub("Voiture", "Car", WA_db$Allowance25)
WA_db$Allowance25 <- gsub("Voiture électrique", "Car",
WA_db$Allowance25)
WA_db$Allowance25 <- gsub("Camion de société", "Car",
WA_db$Allowance25)
WA_db$Allowance25 <- gsub("Camion", "Car", WA_db$Allowance25)
WA_db$Allowance25 <- gsub("[()]", "", WA_db$Allowance25)
WA_db$Allowance25 <- gsub("Transport en commun train, bus, métro,
tram", "Public transport", WA_db$Allowance25)

```

```

WA_db$Allowance25 <- gsub("Vélo, vélo électrique", "Bike",
WA_db$Allowance25)
WA_db$Allowance25 <- gsub("Moto, scooter,...", "Moto",
WA_db$Allowance25)
WA_db$Allowance25 <- gsub("Marche à pied", "Walking",
WA_db$Allowance25)

WA_db$Allowance35 <- gsub("Voiture", "Car", WA_db$Allowance35)
WA_db$Allowance35 <- gsub("Voiture électrique", "Car",
WA_db$Allowance35)
WA_db$Allowance35 <- gsub("Camion de société", "Car",
WA_db$Allowance35)
WA_db$Allowance35 <- gsub("Camion", "Car", WA_db$Allowance35)
WA_db$Allowance35 <- gsub("[()]", "", WA_db$Allowance35)
WA_db$Allowance35 <- gsub("Transport en commun train, bus, métro,
tram", "Public transport", WA_db$Allowance35)
WA_db$Allowance35 <- gsub("Vélo, vélo électrique", "Bike",
WA_db$Allowance35)
WA_db$Allowance35 <- gsub("Moto, scooter,...", "Moto",
WA_db$Allowance35)
WA_db$Allowance35 <- gsub("Marche à pied", "Walking",
WA_db$Allowance35)

WA_db$Allowance50 <- gsub("Voiture", "Car", WA_db$Allowance50)
WA_db$Allowance50 <- gsub("Voiture électrique", "Car",
WA_db$Allowance50)
WA_db$Allowance50 <- gsub("Camion de société", "Car",
WA_db$Allowance50)
WA_db$Allowance50 <- gsub("Camion", "Car", WA_db$Allowance50)
WA_db$Allowance50 <- gsub("[()]", "", WA_db$Allowance50)
WA_db$Allowance50 <- gsub("Transport en commun train, bus, métro,
tram", "Public transport", WA_db$Allowance50)
WA_db$Allowance50 <- gsub("Vélo, vélo électrique", "Bike",
WA_db$Allowance50)
WA_db$Allowance50 <- gsub("Moto, scooter,...", "Moto",
WA_db$Allowance50)
WA_db$Allowance50 <- gsub("Marche à pied", "Walking",
WA_db$Allowance50)

```

```
##Filter
```

```
#Nonexisting postcode + Brussels municipalities
```

```
ZIP_up <- 9992
ZIP_down <- 1213
```

```
VL_db <- subset(VL_db, !(ZIP <= ZIP_down | ZIP >= ZIP_up))
WA_db <- subset(WA_db, !(ZIP <= ZIP_down | ZIP >= ZIP_up))
```

```
#Dealing with NA values
```

```
sum(is.na(VL_db))
sum(is.na(WA_db))
```

```
which(is.na(VL_db))
```

```

which(is.na(WA_db))

#Majority of NA values are in the Bike_allowance question, replace
for all
#respondents that do not go by bike by "Other" as it was the
expected answer already
for (i in 1:nrow(VL_db)) {
  # Check if variables have allowed values
  if (VL_db$Transport[i] != "Bike") {
    VL_db$`Bike_allowance`[i] <- "Other"
  }
}

for (i in 1:nrow(WA_db)) {
  # Check if variables have allowed values
  if (WA_db$Transport[i] != "Bike") {
    WA_db$`Bike_allowance`[i] <- "Other"
  }
}

#No more NA values for Walloons,
which(is.na(VL_db))
VL_db$Allowance15[1939] <- "Car"
#This respondent choice were all "Car", so we can confidently fill
the NA value

#Inconsistent mobility choice
#VL

# Vector to store row indices to be deleted
rows_to_delete <- c()

# Loop through rows of data frame
for (i in 1:nrow(VL_db)) {
  # Enforce value constraints for other 4 variables
  if (VL_db$Allowance15[i] != VL_db$NoAllowance[i] &
VL_db$Allowance15[i] != "Bike" |
    VL_db$Allowance25[i] != VL_db$NoAllowance[i] &
VL_db$Allowance25[i] != "Bike" |
    VL_db$Allowance35[i] != VL_db$NoAllowance[i] &
VL_db$Allowance35[i] != "Bike" |
    VL_db$Allowance50[i] != VL_db$NoAllowance[i] &
VL_db$Allowance50[i] != "Bike") {
    rows_to_delete <- c(rows_to_delete, i)
  }
}
VL_db <- VL_db[-rows_to_delete,]
rows_to_delete <- c()

for (i in 1:nrow(VL_db)) {
  # Enforce value constraints for other 4 variables
  if (VL_db$Transport[i] != VL_db$NoAllowance[i] &
    VL_db$Transport[i] != VL_db$Allowance15[i] &

```

```

        VL_db$Transport[i] != VL_db$Allowance25[i] &
        VL_db$Transport[i] != VL_db$Allowance25[i] &
        VL_db$Transport[i] != VL_db$Allowance50[i]) {
    rows_to_delete <- c(rows_to_delete, i)
  }
}
VL_db <- VL_db[-rows_to_delete,]
rows_to_delete <- c()

#WA
# Vector to store row indices to be deleted
rows_to_delete <- c()

# Loop through rows of data frame
for (i in 1:nrow(WA_db)) {
  # Enforce value constraints for other 4 variables
  if (WA_db$Allowance15[i] != WA_db$NoAllowance[i] &
      WA_db$Allowance15[i] != "Bike" |
      WA_db$Allowance25[i] != WA_db$NoAllowance[i] &
      WA_db$Allowance25[i] != "Bike" |
      WA_db$Allowance35[i] != WA_db$NoAllowance[i] &
      WA_db$Allowance35[i] != "Bike" |
      WA_db$Allowance50[i] != WA_db$NoAllowance[i] &
      WA_db$Allowance50[i] != "Bike") {
    rows_to_delete <- c(rows_to_delete, i)
  }
}
WA_db <- WA_db[-rows_to_delete,]
rows_to_delete <- c()

for (i in 1:nrow(WA_db)) {
  # Enforce value constraints for other 4 variables
  if (WA_db$Transport[i] != WA_db$NoAllowance[i] &
      WA_db$Transport[i] != WA_db$Allowance15[i] &
      WA_db$Transport[i] != WA_db$Allowance25[i] &
      WA_db$Transport[i] != WA_db$Allowance25[i] &
      WA_db$Transport[i] != WA_db$Allowance50[i]) {
    rows_to_delete <- c(rows_to_delete, i)
  }
}
WA_db <- WA_db[-rows_to_delete,]
rows_to_delete <- c()

#Delete non-usable data
#VL
rows_to_keep_VL <- c()
allowed_vals <- c("Car", "Bike", "Public transport", "Moto",
"Walking")
# Loop through rows of data frame
for (i in 1:nrow(VL_db)) {
  # Check if variables have allowed values

```

```

    if (VL_db$NoAllowance[i] %in% allowed_vals &
        VL_db$Allowance15[i] %in% allowed_vals &
        VL_db$Allowance25[i] %in% allowed_vals &
        VL_db$Allowance35[i] %in% allowed_vals &
        VL_db$Allowance50[i] %in% allowed_vals) {
        rows_to_keep_VL <- c(rows_to_keep_VL, i)
    }
}
VL_db <- VL_db[rows_to_keep_VL,]

#WA
rows_to_keep_WA <- c()
allowed_vals <- c("Car", "Bike", "Public transport", "Moto",
"Walking")
# Loop through rows of data frame
for (i in 1:nrow(WA_db)) {
    # Check if variables have allowed values
    if (WA_db$NoAllowance[i] %in% allowed_vals &
        WA_db$Allowance15[i] %in% allowed_vals &
        WA_db$Allowance25[i] %in% allowed_vals &
        WA_db$Allowance35[i] %in% allowed_vals &
        WA_db$Allowance50[i] %in% allowed_vals) {
        rows_to_keep_WA <- c(rows_to_keep_WA, i)
    }
}
WA_db <- WA_db[rows_to_keep_WA,]

#Check for 50-50 split even after data cleaning
sum(VL_db$Treatment == "Yes")
sum(WA_db$Treatment == "Yes")

##Creation of a variable that is 1 if respondents changed in
function of the allowance
## and 0 if respondents did not change

#VL
vals_NoAllowance <- unique(VL_db$NoAllowance)
for (val1 in vals_NoAllowance) {
    # Loop through rows of data frame
    for (i in 1:nrow(VL_db)) {
        # Check if NoAllowance has current value
        if (VL_db$NoAllowance[i] == val1) {
            # Enforce value constraints for other 4 variables
            if (VL_db$Allowance15[i] == val1 &
                VL_db$Allowance25[i] == val1 &
                VL_db$Allowance35[i] == val1 &
                VL_db$Allowance50[i] == val1) {
                VL_db$Change[i] <- 0
            }
            else VL_db$Change[i] <- 1
        }
    }
}
}

```

```

#WA
vals_NoAllowance <- unique(WA_db$NoAllowance)
for (val1 in vals_NoAllowance) {
  # Loop through rows of data frame
  for (i in 1:nrow(WA_db)) {
    # Check if NoAllowance has current value
    if (WA_db$NoAllowance[i] == val1) {
      # Enforce value constraints for other 4 variables
      if (WA_db$Allowance15[i] == val1 &
          WA_db$Allowance25[i] == val1 &
          WA_db$Allowance35[i] == val1 &
          WA_db$Allowance50[i] == val1) {
        WA_db$Change[i] <- 0
      }
      else WA_db$Change[i] <- 1
    }
  }
}

#Creation of a variable that is 1 if a CAR user switched to the
bike, 0 otherwise
vals_NoAllowance <- unique(VL_db$NoAllowance)
# Loop through rows of data frame
for (i in 1:nrow(VL_db)) {
  # Check if NoAllowance has current value
  if (VL_db$Transport[i] == "Car") {
    # Enforce value constraints for other 4 variables
    if (VL_db$Allowance15[i] != "Car" |
        VL_db$Allowance25[i] != "Car" |
        VL_db$Allowance35[i] != "Car" |
        VL_db$Allowance50[i] != "Car") {
      VL_db$ChangeCar[i] <- 1
    }
    else VL_db$ChangeCar[i] <- 0
  }
  else VL_db$ChangeCar[i] <- 0
}

# Loop through rows of data frame
for (i in 1:nrow(WA_db)) {
  # Check if NoAllowance has current value
  if (WA_db$Transport[i] == "Car") {
    # Enforce value constraints for other 4 variables
    if (WA_db$Allowance15[i] != "Car" |
        WA_db$Allowance25[i] != "Car" |
        WA_db$Allowance35[i] != "Car" |
        WA_db$Allowance50[i] != "Car") {
      WA_db$ChangeCar[i] <- 1
    }
    else WA_db$ChangeCar[i] <- 0
  }
  else WA_db$ChangeCar[i] <- 0
}

```

```

#Same but only between NoAll and All15
#VL
vals_NoAllowance <- unique(VL_db$NoAllowance)
for (val1 in vals_NoAllowance) {
  # Loop through rows of data frame
  for (i in 1:nrow(VL_db)) {
    # Check if NoAllowance has current value
    if (VL_db$NoAllowance[i] == val1) {
      # Enforce value constraints for other 4 variables
      if (VL_db$Allowance15[i] == val1) {
        VL_db$ChangeNo15[i] <- 0
      }
      else VL_db$ChangeNo15[i] <- 1
    }
  }
}

vals_Allowance15 <- unique(VL_db$Allowance15)
for (val1 in vals_Allowance15) {
  # Loop through rows of data frame
  for (i in 1:nrow(VL_db)) {
    # Check if NoAllowance has current value
    if (VL_db$Allowance15[i] == val1) {
      # Enforce value constraints for other 4 variables
      if (VL_db$Allowance25[i] == val1) {
        VL_db$Change1525[i] <- 0
      }
      else VL_db$Change1525[i] <- 1
    }
  }
}

#WA
vals_NoAllowance <- unique(WA_db$NoAllowance)
for (val1 in vals_NoAllowance) {
  # Loop through rows of data frame
  for (i in 1:nrow(WA_db)) {
    # Check if NoAllowance has current value
    if (WA_db$NoAllowance[i] == val1) {
      # Enforce value constraints for other 4 variables
      if (WA_db$Allowance15[i] == val1) {
        WA_db$ChangeNo15[i] <- 0
      }
      else WA_db$ChangeNo15[i] <- 1
    }
  }
}

vals_Allowance15 <- unique(WA_db$Allowance15)
for (val1 in vals_Allowance15) {
  # Loop through rows of data frame
  for (i in 1:nrow(WA_db)) {

```

```

# Check if NoAllowance has current value
if (WA_db$Allowance15[i] == val1) {
  # Enforce value constraints for other 4 variables
  if (WA_db$Allowance25[i] == val1) {
    WA_db$Change1525[i] <- 0
  }
  else WA_db$Change1525[i] <- 1
}
}
}

```

```

#Same but only between NoAll and All25
#VL
vals_NoAllowance <- unique(VL_db$NoAllowance)
for (val1 in vals_NoAllowance) {
  # Loop through rows of data frame
  for (i in 1:nrow(VL_db)) {
    # Check if NoAllowance has current value
    if (VL_db$NoAllowance[i] == val1) {
      # Enforce value constraints for other 4 variables
      if (VL_db$Allowance15[i] == val1 &
          VL_db$Allowance25[i] == val1 ) {
        VL_db$ChangeNo25[i] <- 0
      }
      else VL_db$ChangeNo25[i] <- 1
    }
  }
}
}

```

```

#WA
vals_NoAllowance <- unique(WA_db$NoAllowance)
for (val1 in vals_NoAllowance) {
  # Loop through rows of data frame
  for (i in 1:nrow(WA_db)) {
    # Check if NoAllowance has current value
    if (WA_db$NoAllowance[i] == val1) {
      # Enforce value constraints for other 4 variables
      if (WA_db$Allowance15[i] == val1 &
          WA_db$Allowance25[i] == val1) {
        WA_db$ChangeNo25[i] <- 0
      }
      else WA_db$ChangeNo25[i] <- 1
    }
  }
}
}

```

```

#Same but only between All25 and All35
#VL
vals_Allowance25 <- unique(VL_db$Allowance25)
for (val1 in vals_Allowance25) {
  # Loop through rows of data frame
  for (i in 1:nrow(VL_db)) {
    # Check if NoAllowance has current value
    if (VL_db$Allowance25[i] == val1) {

```

```

    # Enforce value constraints for other 4 variables
    if (VL_db$Allowance35[i] == val1) {
      VL_db$Change2535[i] <- 0
    }
    else VL_db$Change2535[i] <- 1
  }
}

```

```

#WA
vals_Allowance25 <- unique(WA_db$NoAllowance)
for (val1 in vals_Allowance25) {
  # Loop through rows of data frame
  for (i in 1:nrow(WA_db)) {
    # Check if NoAllowance has current value
    if (WA_db$Allowance25[i] == val1) {
      # Enforce value constraints for other 4 variables
      if (WA_db$Allowance35[i] == val1) {
        WA_db$Change2535[i] <- 0
      }
      else WA_db$Change2535[i] <- 1
    }
  }
}

```

##People that currently take the bike

```

#VL
for (i in 1:nrow(VL_db)) {
  if (VL_db$Transport[i] != "Bike") {
    VL_db$BikeDefault[i] <- 0
  }
  else VL_db$BikeDefault[i] <- 1
}

```

```

for (i in 1:nrow(VL_db)) {
  if (VL_db$NoAllowance[i] != "Bike") {
    VL_db$BikeNo[i] <- 0
  }
  else VL_db$BikeNo[i] <- 1
}

```

```

for (i in 1:nrow(VL_db)) {
  if (VL_db$Allowance15[i] != "Bike") {
    VL_db$Bike15[i] <- 0
  }
  else VL_db$Bike15[i] <- 1
}

```

```

for (i in 1:nrow(VL_db)) {
  if (VL_db$Allowance25[i] != "Bike") {
    VL_db$Bike25[i] <- 0
  }
  else VL_db$Bike25[i] <- 1
}

```

```

}

for (i in 1:nrow(VL_db)) {
  if (VL_db$Allowance35[i] != "Bike") {
    VL_db$Bike35[i] <- 0
  }
  else VL_db$Bike35[i] <- 1
}

for (i in 1:nrow(VL_db)) {
  if (VL_db$Allowance50[i] != "Bike") {
    VL_db$Bike50[i] <- 0
  }
  else VL_db$Bike50[i] <- 1
}

#WA
for (i in 1:nrow(WA_db)) {
  if (WA_db$Transport[i] != "Bike") {
    WA_db$BikeDefault[i] <- 0
  }
  else WA_db$BikeDefault[i] <- 1
}

for (i in 1:nrow(WA_db)) {
  if (WA_db$NoAllowance[i] != "Bike") {
    WA_db$BikeNo[i] <- 0
  }
  else WA_db$BikeNo[i] <- 1
}

for (i in 1:nrow(WA_db)) {
  if (WA_db$Allowance15[i] != "Bike") {
    WA_db$Bike15[i] <- 0
  }
  else WA_db$Bike15[i] <- 1
}

for (i in 1:nrow(WA_db)) {
  if (WA_db$Allowance25[i] != "Bike") {
    WA_db$Bike25[i] <- 0
  }
  else WA_db$Bike25[i] <- 1
}

for (i in 1:nrow(WA_db)) {
  if (WA_db$Allowance35[i] != "Bike") {
    WA_db$Bike35[i] <- 0
  }
  else WA_db$Bike35[i] <- 1
}

for (i in 1:nrow(WA_db)) {
  if (WA_db$Allowance50[i] != "Bike") {

```

```

    WA_db$Bike50[i] <- 0
  }
  else WA_db$Bike50[i] <- 1
}

```

```

#% of people that would take the bike with a 0.50€/km allowance
(sum(VL_db$Change)+sum(VL_db$BikeDefault))/nrow(VL_db)
(sum(WA_db$Change)+sum(VL_db$BikeDefault))/nrow(WA_db)

```

```

#% of people that would take the bike with a 0.25€/km allowance
(sum(VL_db$ChangeNo25)+sum(VL_db$BikeDefault))/nrow(VL_db)
(sum(WA_db$ChangeNo25)+sum(WA_db$BikeDefault))/nrow(WA_db)

```

```

#Internal validity

```

```

#VL

```

```

sum(VL_db$Sex == "M" & VL_db$Treatment == "Yes")
sum(VL_db$Sex == "M" & VL_db$Treatment == "No")
sum(VL_db$Sex == "F" & VL_db$Treatment == "Yes")
sum(VL_db$Sex == "F" & VL_db$Treatment == "No")
sum(VL_db$Sex == "0" & VL_db$Treatment == "Yes")
sum(VL_db$Sex == "0" & VL_db$Treatment == "No")
sum(VL_db$Age == "18-25" & VL_db$Treatment == "Yes")
sum(VL_db$Age == "18-25" & VL_db$Treatment == "No")
sum(VL_db$Age == "26-35" & VL_db$Treatment == "Yes")
sum(VL_db$Age == "26-35" & VL_db$Treatment == "No")
sum(VL_db$Age == "36-45" & VL_db$Treatment == "Yes")
sum(VL_db$Age == "36-45" & VL_db$Treatment == "No")
sum(VL_db$Age == "46-55" & VL_db$Treatment == "Yes")
sum(VL_db$Age == "46-55" & VL_db$Treatment == "No")
sum(VL_db$Age == "56-65" & VL_db$Treatment == "Yes")
sum(VL_db$Age == "56-65" & VL_db$Treatment == "No")
sum(VL_db$Age == "66+" & VL_db$Treatment == "Yes")
sum(VL_db$Age == "66+" & VL_db$Treatment == "No")
sum(VL_db$`Company car` == "Yes" & VL_db$Treatment == "Yes")
sum(VL_db$`Company car` == "Yes" & VL_db$Treatment == "No")
sum(VL_db$`Company car` == "No" & VL_db$Treatment == "Yes")
sum(VL_db$`Company car` == "No" & VL_db$Treatment == "No")

```

```

#t-tests

```

```

control_group <- VL_db[VL_db$Treatment == "No", ]
treatment_group <- VL_db[VL_db$Treatment == "Yes", ]

```

```

mean(control_group$`W/F feeling`, na.rm = TRUE)
mean(treatment_group$`W/F feeling`, na.rm = TRUE)

```

```

id_c <- as.numeric(control_group$`W/F feeling`)
id_t <- as.numeric(treatment_group$`W/F feeling`)
t.test(id_c, id_t, paired = FALSE)

```

```

gender_c <- as.numeric(factor(control_group$Sex))
gender_t <- as.numeric(factor(treatment_group$Sex))

```

```

t.test(gender_c, gender_t, paired = FALSE)

age_c <- as.numeric(factor(control_group$Age))
age_t <- as.numeric(factor(treatment_group$Age))
t.test(age_c, age_t, paired = FALSE)

cc_c <- as.numeric(factor(control_group`Company car`))
cc_t <- as.numeric(factor(treatment_group`Company car`))
t.test(cc_c, cc_t, paired = FALSE)

#WA
sum(WA_db$Sex == "M" & WA_db$Treatment == "Yes")
sum(WA_db$Sex == "M" & WA_db$Treatment == "No")
sum(WA_db$Sex == "F" & WA_db$Treatment == "Yes")
sum(WA_db$Sex == "F" & WA_db$Treatment == "No")
sum(WA_db$Sex == "O" & WA_db$Treatment == "Yes")
sum(WA_db$Sex == "O" & WA_db$Treatment == "No")
sum(WA_db$Sex == "NP" & WA_db$Treatment == "Yes")
sum(WA_db$Sex == "NP" & WA_db$Treatment == "No")
sum(WA_db$Age == "18-25" & WA_db$Treatment == "Yes")
sum(WA_db$Age == "18-25" & WA_db$Treatment == "No")
sum(WA_db$Age == "26-35" & WA_db$Treatment == "Yes")
sum(WA_db$Age == "26-35" & WA_db$Treatment == "No")
sum(WA_db$Age == "36-45" & WA_db$Treatment == "Yes")
sum(WA_db$Age == "36-45" & WA_db$Treatment == "No")
sum(WA_db$Age == "46-55" & WA_db$Treatment == "Yes")
sum(WA_db$Age == "46-55" & WA_db$Treatment == "No")
sum(WA_db$Age == "56-65" & WA_db$Treatment == "Yes")
sum(WA_db$Age == "56-65" & WA_db$Treatment == "No")
sum(WA_db$Age == "66+" & WA_db$Treatment == "Yes")
sum(WA_db$Age == "66+" & WA_db$Treatment == "No")
sum(WA_db`Company car` == "Yes" & WA_db$Treatment == "Yes")
sum(WA_db`Company car` == "Yes" & WA_db$Treatment == "No")
sum(WA_db`Company car` == "No" & WA_db$Treatment == "Yes")
sum(WA_db`Company car` == "No" & WA_db$Treatment == "No")

sum(WA_db$Treatment == "Yes")
sum(WA_db$Treatment == "No")

sum(VL_db$Treatment == "Yes")
sum(VL_db$Treatment == "No")

#t-tests
control_group <- WA_db[WA_db$Treatment == "No", ]
treatment_group <- WA_db[WA_db$Treatment == "Yes", ]

mean(control_group`W/F feeling`, na.rm = TRUE)
mean(treatment_group`W/F feeling`, na.rm = TRUE)

id_c <- as.numeric(control_group`W/F feeling`)
id_t <- as.numeric(treatment_group`W/F feeling`)
t.test(id_c, id_t, paired = FALSE)

gender_c <- as.numeric(factor(control_group$Sex))

```

```

gender_t <- as.numeric(factor(treatment_group$Sex))
t.test(gender_c, gender_t, paired = FALSE)

age_c <- as.numeric(factor(control_group$Age))
age_t <- as.numeric(factor(treatment_group$Age))
t.test(age_c, age_t, paired = FALSE)

cc_c <- as.numeric(factor(control_group$`Company car`))
cc_t <- as.numeric(factor(treatment_group$`Company car`))
t.test(cc_c, cc_t, paired = FALSE)

sum(control_group$Change == 1)/sum(nrow(control_group))
sum(treatment_group$Change == 1)/sum(nrow(treatment_group))

sum(control_group$ChangeNo15 == 1)/sum(nrow(control_group))
sum(treatment_group$ChangeNo15 == 1)/sum(nrow(treatment_group))

sum(control_group$ChangeNo25 == 1)/sum(nrow(control_group))
sum(treatment_group$ChangeNo25 == 1)/sum(nrow(treatment_group))

sum(control_group$Change1525 == 1)/sum(nrow(control_group))
sum(treatment_group$Change1525 == 1)/sum(nrow(treatment_group))

sum(control_group$Change2535 == 1)/sum(nrow(control_group))
sum(treatment_group$Change2535 == 1)/sum(nrow(treatment_group))

##Descriptive statistics
#VL
#ZIP code
ZIP_up1 <- 4000
ZIP_up2 <- 8000
ZIP_down1 <- 1500
VLinWA <- subset(VL_db, ZIP_down1 > ZIP | ZIP >= ZIP_up1 & ZIP <
ZIP_up2)
mean(VLinWA$`W/F feeling`)

sum(VL_db$ZIP == "2000" | VL_db$ZIP == "2018" | VL_db$ZIP == "2020" |
VL_db$ZIP == "2030" | VL_db$ZIP == "2040" | VL_db$ZIP == "2060" |
VL_db$ZIP == "2060" | VL_db$ZIP == "2100" | VL_db$ZIP == "2130"
| VL_db$ZIP == "2140" | VL_db$ZIP == "2170" | VL_db$ZIP == "2180" |
VL_db$ZIP == "2600" | VL_db$ZIP == "2610" | VL_db$ZIP == "2660" |
VL_db$ZIP == "3000" | VL_db$ZIP == "3001" | VL_db$ZIP == "3010"
| VL_db$ZIP == "3012" | VL_db$ZIP == "3018" | VL_db$ZIP == "3050"
| VL_db$ZIP == "8000" | VL_db$ZIP == "8200" | VL_db$ZIP == "8310" |
VL_db$ZIP == "8380" | VL_db$ZIP == "9000"
| VL_db$ZIP == "9030" | VL_db$ZIP == "9031" | VL_db$ZIP == "9032" |
VL_db$ZIP == "9040" | VL_db$ZIP == "9041"
| VL_db$ZIP == "9042")
#individual ZIP codes count
length(unique(VL_db$ZIP))

#Gender
MT <- sum(VL_db$Sex == "M" & VL_db$Treatment == "Yes")

```

```

MC <- sum(VL_db$Sex == "M" & VL_db$Treatment == "No")
FT <- sum(VL_db$Sex == "F" & VL_db$Treatment == "Yes")
FC <- sum(VL_db$Sex == "F" & VL_db$Treatment == "No")
MT/(MT+FT)
FT/(MT+FT)
MC/(MC+FC)
FC/(MC+FC)

#Global population Transport mode used
sum(VL_db$Transport == "Car")/nrow(VL_db)
sum(VL_db$Transport == "Bike")/nrow(VL_db)
sum(VL_db$Transport == "Public transport")/nrow(VL_db)

#Age groups details
car1825 = 0
bike1825 = 0
train1825 = 0
car2635 = 0
bike2635 = 0
train2635 = 0
car3645 = 0
bike3645 = 0
train3645 = 0
car4655 = 0
bike4655 = 0
train4655 = 0
car5665 = 0
bike5665 = 0
train5665 = 0
car66 = 0
bike66 = 0
train66 = 0

for (i in 1:nrow(VL_db)) {
  if (VL_db$Age[i] == "18-25" & VL_db$Transport[i] == "Car") {
    car1825 <- car1825 + 1
  }
  else if (VL_db$Age[i] == "18-25" & VL_db$Transport[i] == "Bike")
  {
    bike1825 <- bike1825 + 1
  }
  else if (VL_db$Age[i] == "18-25" & VL_db$Transport[i] == "Public
transport"){
    train1825 <- train1825 + 1
  }
  else if (VL_db$Age[i] == "26-35" & VL_db$Transport[i] == "Car"){
    car2635 <- car2635 + 1
  }
  else if (VL_db$Age[i] == "26-35" & VL_db$Transport[i] == "Bike")
  {
    bike2635 <- bike2635 + 1
  }
  else if (VL_db$Age[i] == "26-35" & VL_db$Transport[i] == "Public
transport"){

```

```

    train2635 <- train2635 + 1
  }
  else if (VL_db$Age[i] == "36-45" & VL_db$Transport[i] == "Car"){
    car3645 <- car3645 + 1
  }
  else if (VL_db$Age[i] == "36-45" & VL_db$Transport[i] == "Bike")
  {
    bike3645 <- bike3645 + 1
  }
  else if (VL_db$Age[i] == "36-45" & VL_db$Transport[i] == "Public
transport"){
    train3645 <- train3645 + 1
  }
  else if (VL_db$Age[i] == "46-55" & VL_db$Transport[i] == "Car"){
    car4655 <- car4655 + 1
  }
  else if (VL_db$Age[i] == "46-55" & VL_db$Transport[i] == "Bike")
  {
    bike4655 <- bike4655 + 1
  }
  else if (VL_db$Age[i] == "46-55" & VL_db$Transport[i] == "Public
transport"){
    train4655 <- train4655 + 1
  }
  else if (VL_db$Age[i] == "56-65" & VL_db$Transport[i] == "Car"){
    car5665 <- car5665 + 1
  }
  else if (VL_db$Age[i] == "56-65" & VL_db$Transport[i] == "Bike")
  {
    bike5665 <- bike5665 + 1
  }
  else if (VL_db$Age[i] == "56-65" & VL_db$Transport[i] == "Public
transport"){
    train5665 <- train5665 + 1
  }
  else if (VL_db$Age[i] == "66+" & VL_db$Transport[i] == "Car"){
    car66 <- car66 + 1
  }
  else if (VL_db$Age[i] == "66+" & VL_db$Transport[i] == "Bike"){
    bike66 <- bike66 + 1
  }
  else if (VL_db$Age[i] == "66+" & VL_db$Transport[i] == "Public
transport"){
    train66 <- train66 + 1
  }
}

```

```

car1825/sum(VL_db$Age == "18-25")
bike1825/sum(VL_db$Age == "18-25")
train1825/sum(VL_db$Age == "18-25")

```

```

car2635/sum(VL_db$Age == "26-35")
bike2635/sum(VL_db$Age == "26-35")
train2635/sum(VL_db$Age == "26-35")

```

```
car3645/sum(VL_db$Age == "36-45")
bike3645/sum(VL_db$Age == "36-45")
train3645/sum(VL_db$Age == "36-45")
```

```
car4655/sum(VL_db$Age == "46-55")
bike4655/sum(VL_db$Age == "46-55")
train4655/sum(VL_db$Age == "46-55")
```

```
car5665/sum(VL_db$Age == "56-65")
bike5665/sum(VL_db$Age == "56-65")
train5665/sum(VL_db$Age == "56-65")
```

```
car66/sum(VL_db$Age == "66+")
bike66/sum(VL_db$Age == "66+")
train66/sum(VL_db$Age == "66+")
```

```
sum(VL_db$Age == "18-25")/nrow(VL_db)
sum(VL_db$Age == "26-35")/nrow(VL_db)
sum(VL_db$Age == "36-45")/nrow(VL_db)
sum(VL_db$Age == "46-55")/nrow(VL_db)
sum(VL_db$Age == "56-65")/nrow(VL_db)
sum(VL_db$Age == "66+")/nrow(VL_db)
```

```
#For external validity (without 66+ group)
sum(VL_db$Age == "18-25")/sum(VL_db$Age == "18-25" | VL_db$Age ==
"26-35" | VL_db$Age == "36-45"
                                | VL_db$Age == "46-55" | VL_db$Age ==
"56-65")
sum(VL_db$Age == "26-35")/sum(VL_db$Age == "18-25" | VL_db$Age ==
"26-35" | VL_db$Age == "36-45"
                                | VL_db$Age == "46-55" | VL_db$Age ==
"56-65")
sum(VL_db$Age == "36-45")/sum(VL_db$Age == "18-25" | VL_db$Age ==
"26-35" | VL_db$Age == "36-45"
                                | VL_db$Age == "46-55" | VL_db$Age ==
"56-65")
sum(VL_db$Age == "46-55")/sum(VL_db$Age == "18-25" | VL_db$Age ==
"26-35" | VL_db$Age == "36-45"
                                | VL_db$Age == "46-55" | VL_db$Age ==
"56-65")
sum(VL_db$Age == "56-65")/sum(VL_db$Age == "18-25" | VL_db$Age ==
"26-35" | VL_db$Age == "36-45"
                                | VL_db$Age == "46-55" | VL_db$Age ==
"56-65")
```

```
#Distance details
car05 = 0
bike05 = 0
train05 = 0
car0610 = 0
bike0610 = 0
train0610 = 0
```

```

car1120 = 0
bike1120 = 0
train1120 = 0
car2130 = 0
bike2130 = 0
train2130 = 0
car3140 = 0
bike3140 = 0
train3140 = 0
car4150 = 0
bike4150 = 0
train4150 = 0
car5160 = 0
bike5160 = 0
train5160 = 0
car61100 = 0
bike61100 = 0
train61100 = 0

for (i in 1:nrow(VL_db)) {
  if (VL_db$Distance[i] == "0-5" & VL_db$Transport[i] == "Car") {
    car05 <- car05 + 1
  }
  else if (VL_db$Distance[i] == "0-5" & VL_db$Transport[i] ==
"Bike"){
    bike05 <- bike05 + 1
  }
  else if (VL_db$Distance[i] == "0-5" & VL_db$Transport[i] ==
"Public transport"){
    train05 <- train05 + 1
  }
  else if (VL_db$Distance[i] == "6-10" & VL_db$Transport[i] ==
"Car"){
    car0610 <- car0610 + 1
  }
  else if (VL_db$Distance[i] == "6-10" & VL_db$Transport[i] ==
"Bike"){
    bike0610 <- bike0610 + 1
  }
  else if (VL_db$Distance[i] == "6-10" & VL_db$Transport[i] ==
"Public transport"){
    train0610 <- train0610 + 1
  }
  else if (VL_db$Distance[i] == "11-20" & VL_db$Transport[i] ==
"Car"){
    car1120 <- car1120 + 1
  }
  else if (VL_db$Distance[i] == "11-20" & VL_db$Transport[i] ==
"Bike"){
    bike1120 <- bike1120 + 1
  }
  else if (VL_db$Distance[i] == "11-20" & VL_db$Transport[i] ==
"Public transport"){
    train1120 <- train1120 + 1
  }
}

```

```

    }
    else if (VL_db$Distance[i] == "21-30" & VL_db$Transport[i] ==
"Car"){
        car2130 <- car2130 + 1
    }
    else if (VL_db$Distance[i] == "21-30" & VL_db$Transport[i] ==
"Bike"){
        bike2130 <- bike2130 + 1
    }
    else if (VL_db$Distance[i] == "21-30" & VL_db$Transport[i] ==
"Public transport"){
        train2130 <- train2130 + 1
    }
    else if (VL_db$Distance[i] == "31-40" & VL_db$Transport[i] ==
"Car"){
        car3140 <- car3140 + 1
    }
    else if (VL_db$Distance[i] == "31-40" & VL_db$Transport[i] ==
"Bike"){
        bike3140 <- bike3140 + 1
    }
    else if (VL_db$Distance[i] == "31-40" & VL_db$Transport[i] ==
"Public transport"){
        train3140 <- train3140 + 1
    }
    else if (VL_db$Distance[i] == "41-50" & VL_db$Transport[i] ==
"Car"){
        car4150 <- car4150 + 1
    }
    else if (VL_db$Distance[i] == "41-50" & VL_db$Transport[i] ==
"Bike"){
        bike4150 <- bike4150 + 1
    }
    else if (VL_db$Distance[i] == "41-50" & VL_db$Transport[i] ==
"Public transport"){
        train4150 <- train4150 + 1
    }
    else if (VL_db$Distance[i] == "51-60" & VL_db$Transport[i] ==
"Car"){
        car5160 <- car5160 + 1
    }
    else if (VL_db$Distance[i] == "51-60" & VL_db$Transport[i] ==
"Bike"){
        bike5160 <- bike5160 + 1
    }
    else if (VL_db$Distance[i] == "51-60" & VL_db$Transport[i] ==
"Public transport"){
        train5160 <- train5160 + 1
    }
    else if (VL_db$Distance[i] == "60+" & VL_db$Transport[i] == "Car")
{
        car61100 <- car61100 + 1
    }
    else if (VL_db$Distance[i] == "60+" & VL_db$Transport[i] ==

```

```

"Bike"){
  bike61100 <- bike61100 + 1
}
else if (VL_db$Distance[i] == "60+" & VL_db$Transport[i] ==
"Public transport"){
  train61100 <- train61100 + 1
}
}

car05/sum(VL_db$Distance == "0-5")
bike05/sum(VL_db$Distance == "0-5")
train05/sum(VL_db$Distance == "0-5")

car0610/sum(VL_db$Distance == "6-10")
bike0610/sum(VL_db$Distance == "6-10")
train0610/sum(VL_db$Distance == "6-10")

car1120/sum(VL_db$Distance == "11-20")
bike1120/sum(VL_db$Distance == "11-20")
train1120/sum(VL_db$Distance == "11-20")

car2130/sum(VL_db$Distance == "21-30")
bike2130/sum(VL_db$Distance == "21-30")
train2130/sum(VL_db$Distance == "21-30")

car3140/sum(VL_db$Distance == "31-40")
bike3140/sum(VL_db$Distance == "31-40")
train3140/sum(VL_db$Distance == "31-40")

car4150/sum(VL_db$Distance == "41-50")
bike4150/sum(VL_db$Distance == "41-50")
train4150/sum(VL_db$Distance == "41-50")

car5160/sum(VL_db$Distance == "51-60")
bike5160/sum(VL_db$Distance == "51-60")
train5160/sum(VL_db$Distance == "51-60")

car61100/sum(VL_db$Distance == "60+")
bike61100/sum(VL_db$Distance == "60+")
train61100/sum(VL_db$Distance == "60+")

#W/F Feeling
mean(VL_db$`W/F feeling`, na.rm = TRUE)

#Company car ownership
sum(VL_db$`Company car` == "Yes")/nrow(VL_db)
sum(VL_db$`Fuel card` == "Yes")/nrow(VL_db)
sum(VL_db$`Company car` == "Yes" & VL_db$`Fuel card` == "Yes")
sum(VL_db$`Company car` == "Yes" & VL_db$`Fuel card` == "No")
sum(VL_db$`Company car` == "No" & VL_db$`Fuel card` == "Yes")
sum(VL_db$`Company car` == "No" & VL_db$`Fuel card` == "No")

#Distance

```

```

sum(VL_db$Distance == "0-5")/nrow(VL_db)
sum(VL_db$Distance == "6-10")/nrow(VL_db)
sum(VL_db$Distance == "11-20")/nrow(VL_db)
sum(VL_db$Distance == "21-30")/nrow(VL_db)
sum(VL_db$Distance == "31-40")/nrow(VL_db)
sum(VL_db$Distance == "41-50")/nrow(VL_db)
(sum(VL_db$Distance == "51-60")+sum(VL_db$Distance == "50-60"))/
nrow(VL_db)
sum(VL_db$Distance == "60+")/nrow(VL_db)

unique(VL_db$Distance)
unique(WA_db$Distance)

which(is.na(VL_db$Distance))

#Number of commutes/week (15 NA values)
sum(VL_db$`Weekly commutes` == "1.0" | VL_db$`Weekly commutes` ==
"1", na.rm = TRUE)/(nrow(VL_db)-15)
sum(VL_db$`Weekly commutes` == "2.0" | VL_db$`Weekly commutes` ==
"2", na.rm = TRUE)/(nrow(VL_db)-15)
sum(VL_db$`Weekly commutes` == "3.0", VL_db$`Weekly commutes` ==
"3", na.rm = TRUE)/(nrow(VL_db)-15)
sum(VL_db$`Weekly commutes` == "4.0" | VL_db$`Weekly commutes` ==
"4", na.rm = TRUE)/(nrow(VL_db)-15)
sum(VL_db$`Weekly commutes` == "5", na.rm = TRUE)/(nrow(VL_db)-15)
sum(VL_db$`Weekly commutes` == "6.0" | VL_db$`Weekly commutes` ==
"6", na.rm = TRUE)/(nrow(VL_db)-15)
sum(VL_db$`Weekly commutes` == "6+", na.rm = TRUE)/(nrow(VL_db)-15)

sum((VL_db$`Weekly commutes` == "1.0" | VL_db$`Weekly commutes` ==
"1") & VL_db$Transport == "Car", na.rm = TRUE)/sum(VL_db$`Weekly
commutes` == "1.0" | VL_db$`Weekly commutes` == "1", na.rm = TRUE)
sum((VL_db$`Weekly commutes` == "2.0" | VL_db$`Weekly commutes` ==
"2") & VL_db$Transport == "Car", na.rm = TRUE)/sum(VL_db$`Weekly
commutes` == "2.0" | VL_db$`Weekly commutes` == "2", na.rm = TRUE)
sum((VL_db$`Weekly commutes` == "3.0" | VL_db$`Weekly commutes` ==
"3") & VL_db$Transport == "Car", na.rm = TRUE)/sum(VL_db$`Weekly
commutes` == "3.0" | VL_db$`Weekly commutes` == "3", na.rm = TRUE)
sum((VL_db$`Weekly commutes` == "4.0" | VL_db$`Weekly commutes` ==
"4") & VL_db$Transport == "Car", na.rm = TRUE)/sum(VL_db$`Weekly
commutes` == "4.0" | VL_db$`Weekly commutes` == "4", na.rm = TRUE)
sum(VL_db$`Weekly commutes` == "5" & VL_db$Transport == "Car",
na.rm = TRUE)/sum(VL_db$`Weekly commutes` == "5", na.rm = TRUE)
sum((VL_db$`Weekly commutes` == "6.0" | VL_db$`Weekly commutes` ==
"6") & VL_db$Transport == "Car", na.rm = TRUE)/sum(VL_db$`Weekly
commutes` == "6.0" | VL_db$`Weekly commutes` == "6", na.rm = TRUE)
sum((VL_db$`Weekly commutes` == "6.0+" | VL_db$`Weekly commutes` ==
"6+") & VL_db$Transport == "Car", na.rm = TRUE)/sum(VL_db$`Weekly
commutes` == "6.0+" | VL_db$`Weekly commutes` == "6+", na.rm = TRUE)

sum((VL_db$`Weekly commutes` == "1.0" | VL_db$`Weekly commutes` ==
"1") & VL_db$Transport == "Bike", na.rm = TRUE)/sum(VL_db$`Weekly
commutes` == "1.0" | VL_db$`Weekly commutes` == "1", na.rm = TRUE)

```

```

sum((VL_db$`Weekly commutes` == "2.0" | VL_db$`Weekly commutes` ==
"2") & VL_db$Transport == "Bike" , na.rm = TRUE)/sum(VL_db$`Weekly
commutes` == "2.0" | VL_db$`Weekly commutes` == "2", na.rm = TRUE)
sum((VL_db$`Weekly commutes` == "3.0" | VL_db$`Weekly commutes` ==
"3") & VL_db$Transport == "Bike" , na.rm = TRUE)/sum(VL_db$`Weekly
commutes` == "3.0" | VL_db$`Weekly commutes` == "3", na.rm = TRUE)
sum((VL_db$`Weekly commutes` == "4.0" | VL_db$`Weekly commutes` ==
"4") & VL_db$Transport == "Bike" , na.rm = TRUE)/sum(VL_db$`Weekly
commutes` == "4.0" | VL_db$`Weekly commutes` == "4", na.rm = TRUE)
sum(VL_db$`Weekly commutes` == "5" & VL_db$Transport == "Bike" ,
na.rm = TRUE)/sum(VL_db$`Weekly commutes` == "5", na.rm = TRUE)
sum((VL_db$`Weekly commutes` == "6.0" | VL_db$`Weekly commutes` ==
"6") & VL_db$Transport == "Bike" , na.rm = TRUE)/sum(VL_db$`Weekly
commutes` == "6.0" | VL_db$`Weekly commutes` == "6", na.rm = TRUE)
sum((VL_db$`Weekly commutes` == "6.0+" | VL_db$`Weekly commutes` ==
"6+") & VL_db$Transport == "Bike" , na.rm = TRUE)/sum(VL_db$`Weekly
commutes` == "6.0+" | VL_db$`Weekly commutes` == "6+", na.rm = TRUE)

```

```

sum((VL_db$`Weekly commutes` == "1.0" | VL_db$`Weekly commutes` ==
"1") & VL_db$Transport == "Public transport" , na.rm = TRUE)/
sum(VL_db$`Weekly commutes` == "1.0" | VL_db$`Weekly commutes` ==
"1", na.rm = TRUE)
sum((VL_db$`Weekly commutes` == "2.0" | VL_db$`Weekly commutes` ==
"2") & VL_db$Transport == "Public transport" , na.rm = TRUE)/
sum(VL_db$`Weekly commutes` == "2.0" | VL_db$`Weekly commutes` ==
"2", na.rm = TRUE)
sum((VL_db$`Weekly commutes` == "3.0" | VL_db$`Weekly commutes` ==
"3") & VL_db$Transport == "Public transport" , na.rm = TRUE)/
sum(VL_db$`Weekly commutes` == "3.0" | VL_db$`Weekly commutes` ==
"3", na.rm = TRUE)
sum((VL_db$`Weekly commutes` == "4.0" | VL_db$`Weekly commutes` ==
"4") & VL_db$Transport == "Public transport" , na.rm = TRUE)/
sum(VL_db$`Weekly commutes` == "4.0" | VL_db$`Weekly commutes` ==
"4", na.rm = TRUE)
sum((VL_db$`Weekly commutes` == "5") & VL_db$Transport == "Public
transport" , na.rm = TRUE)/sum(VL_db$`Weekly commutes` == "5", na.rm
= TRUE)
sum((VL_db$`Weekly commutes` == "6.0" | VL_db$`Weekly commutes` ==
"6") & VL_db$Transport == "Public transport" , na.rm = TRUE)/
sum(VL_db$`Weekly commutes` == "6.0" | VL_db$`Weekly commutes` ==
"6", na.rm = TRUE)
sum((VL_db$`Weekly commutes` == "6.0+" | VL_db$`Weekly commutes` ==
"6+") & VL_db$Transport == "Public transport" , na.rm = TRUE)/
sum(VL_db$`Weekly commutes` == "6.0+" | VL_db$`Weekly commutes` ==
"6+", na.rm = TRUE)

```

```

unique(VL_db$`Weekly commutes`)
which(is.na(VL_db$`Weekly commutes`))

```

#Bike allowance

```

sum(VL_db$Bike_allowance == "No", na.rm = T)
sum(VL_db$Bike_allowance == "Yes", na.rm = T)
sum(VL_db$Bike_allowance == "Other")
520/(520+58)

```

```

#Bike_allowance effect
sum(VL_db$Transport == "Bike")/nrow(VL_db)
sum(VL_db$NoAllowance == "Bike")/nrow(VL_db)
sum(VL_db$Allowance15 == "Bike")/nrow(VL_db)
sum(VL_db$Allowance25 == "Bike")/nrow(VL_db)
sum(VL_db$Allowance35 == "Bike")/nrow(VL_db)
sum(VL_db$Allowance50 == "Bike")/nrow(VL_db)

sum(VL_db$Transport == "Car")/nrow(VL_db)
sum(VL_db$NoAllowance == "Car")/nrow(VL_db)
sum(VL_db$Allowance15 == "Car")/nrow(VL_db)
sum(VL_db$Allowance25 == "Car")/nrow(VL_db)
sum(VL_db$Allowance35 == "Car")/nrow(VL_db)
sum(VL_db$Allowance50 == "Car")/nrow(VL_db)

sum(VL_db$Transport == "Public transport")/nrow(VL_db)
sum(VL_db$NoAllowance == "Public transport")/nrow(VL_db)
sum(VL_db$Allowance15 == "Public transport")/nrow(VL_db)
sum(VL_db$Allowance25 == "Public transport")/nrow(VL_db)
sum(VL_db$Allowance35 == "Public transport")/nrow(VL_db)
sum(VL_db$Allowance50 == "Public transport")/nrow(VL_db)

#WA
#ZIP code
WAINVL <- subset(WA_db, ZIP_down1 <= ZIP & ZIP < ZIP_up1 | ZIP >=
ZIP_up2)
mean(WAINVL$`W/F feeling`)
sort(unique(WAINVL$ZIP))

#Living in 4 major WA cities
sum(WA_db$ZIP == "4000" | WA_db$ZIP == "4020" | WA_db$ZIP == "4030" |
WA_db$ZIP == "4031" | WA_db$ZIP == "4032" | WA_db$ZIP == "4053" |
WA_db$ZIP == "4430" | WA_db$ZIP == "5000" | WA_db$ZIP == "5001"
| WA_db$ZIP == "5002" | WA_db$ZIP == "5003" | WA_db$ZIP == "5004" |
WA_db$ZIP == "5020" | WA_db$ZIP == "5021" | WA_db$ZIP == "5022" |
WA_db$ZIP == "5024" | WA_db$ZIP == "5100" | WA_db$ZIP == "5101"
| WA_db$ZIP == "6000" | WA_db$ZIP == "6001" | WA_db$ZIP == "6010"
| WA_db$ZIP == "6020" | WA_db$ZIP == "6030" | WA_db$ZIP == "6031" |
WA_db$ZIP == "6032" | WA_db$ZIP == "6040"
| WA_db$ZIP == "6041" | WA_db$ZIP == "6042" | WA_db$ZIP == "6043" |
WA_db$ZIP == "6044" | WA_db$ZIP == "6060"
| WA_db$ZIP == "6061" | WA_db$ZIP == "6111" | WA_db$ZIP == "7000" |
WA_db$ZIP == "7010" | WA_db$ZIP == "7011" | WA_db$ZIP == "7012" |
WA_db$ZIP == "7020"
| WA_db$ZIP == "7021" | WA_db$ZIP == "7024" | WA_db$ZIP == "7030" |
WA_db$ZIP == "7031" | WA_db$ZIP == "7032" | WA_db$ZIP == "7033" |
WA_db$ZIP == "7034")

#individual ZIP codes count
length(unique(WA_db$ZIP))

#Gender

```

```

MT <- sum(WA_db$Sex == "M" & WA_db$Treatment == "Yes")
MC <- sum(WA_db$Sex == "M" & WA_db$Treatment == "No")
FT <- sum(WA_db$Sex == "F" & WA_db$Treatment == "Yes")
FC <- sum(WA_db$Sex == "F" & WA_db$Treatment == "No")
MT/(MT+FT)
FT/(MT+FT)
MC/(MC+FC)
FC/(MC+FC)

#Global population Transport mode used
sum(WA_db$Transport == "Car")/nrow(WA_db)
sum(WA_db$Transport == "Bike")/nrow(WA_db)
sum(WA_db$Transport == "Public transport")/nrow(WA_db)

#Age groups details
car1825 = 0
bike1825 = 0
train1825 = 0
car2635 = 0
bike2635 = 0
train2635 = 0
car3645 = 0
bike3645 = 0
train3645 = 0
car4655 = 0
bike4655 = 0
train4655 = 0
car5665 = 0
bike5665 = 0
train5665 = 0
car66 = 0
bike66 = 0
train66 = 0

for (i in 1:nrow(WA_db)) {
  if (WA_db$Age[i] == "18-25" & WA_db$Transport[i] == "Car") {
    car1825 <- car1825 + 1
  }
  else if (WA_db$Age[i] == "18-25" & WA_db$Transport[i] == "Bike"){
    bike1825 <- bike1825 + 1
  }
  else if (WA_db$Age[i] == "18-25" & WA_db$Transport[i] == "Public
transport"){
    train1825 <- train1825 + 1
  }
  else if (WA_db$Age[i] == "26-35" & WA_db$Transport[i] == "Car"){
    car2635 <- car2635 + 1
  }
  else if (WA_db$Age[i] == "26-35" & WA_db$Transport[i] == "Bike"){
    bike2635 <- bike2635 + 1
  }
  else if (WA_db$Age[i] == "26-35" & WA_db$Transport[i] == "Public
transport"){
    train2635 <- train2635 + 1
  }
}

```

```

    }
    else if (WA_db$Age[i] == "36-45" & WA_db$Transport[i] == "Car"){
      car3645 <- car3645 + 1
    }
    else if (WA_db$Age[i] == "36-45" & WA_db$Transport[i] == "Bike"){
      bike3645 <- bike3645 + 1
    }
    else if (WA_db$Age[i] == "36-45" & WA_db$Transport[i] == "Public
transport"){
      train3645 <- train3645 + 1
    }
    else if (WA_db$Age[i] == "46-55" & WA_db$Transport[i] == "Car"){
      car4655 <- car4655 + 1
    }
    else if (WA_db$Age[i] == "46-55" & WA_db$Transport[i] == "Bike"){
      bike4655 <- bike4655 + 1
    }
    else if (WA_db$Age[i] == "46-55" & WA_db$Transport[i] == "Public
transport"){
      train4655 <- train4655 + 1
    }
    else if (WA_db$Age[i] == "56-65" & WA_db$Transport[i] == "Car"){
      car5665 <- car5665 + 1
    }
    else if (WA_db$Age[i] == "56-65" & WA_db$Transport[i] == "Bike"){
      bike5665 <- bike5665 + 1
    }
    else if (WA_db$Age[i] == "56-65" & WA_db$Transport[i] == "Public
transport"){
      train5665 <- train5665 + 1
    }
    else if (WA_db$Age[i] == "66+" & WA_db$Transport[i] == "Car"){
      car66 <- car66 + 1
    }
    else if (WA_db$Age[i] == "66+" & WA_db$Transport[i] == "Bike"){
      bike66 <- bike66 + 1
    }
    else if (WA_db$Age[i] == "66+" & WA_db$Transport[i] == "Public
transport"){
      train66 <- train66 + 1
    }
  }
}

```

```

car1825/sum(WA_db$Age == "18-25")
bike1825/sum(WA_db$Age == "18-25")
train1825/sum(WA_db$Age == "18-25")

```

```

car2635/sum(WA_db$Age == "26-35")
bike2635/sum(WA_db$Age == "26-35")
train2635/sum(WA_db$Age == "26-35")

```

```

car3645/sum(WA_db$Age == "36-45")
bike3645/sum(WA_db$Age == "36-45")
train3645/sum(WA_db$Age == "36-45")

```

```

car4655/sum(WA_db$Age == "46-55")
bike4655/sum(WA_db$Age == "46-55")
train4655/sum(WA_db$Age == "46-55")

car5665/sum(WA_db$Age == "56-65")
bike5665/sum(WA_db$Age == "56-65")
train5665/sum(WA_db$Age == "56-65")

car66/sum(WA_db$Age == "66+")
bike66/sum(WA_db$Age == "66+")
train66/sum(WA_db$Age == "66+")

sum(WA_db$Age == "18-25")/nrow(WA_db)
sum(WA_db$Age == "26-35")/nrow(WA_db)
sum(WA_db$Age == "36-45")/nrow(WA_db)
sum(WA_db$Age == "46-55")/nrow(WA_db)
sum(WA_db$Age == "56-65")/nrow(WA_db)
sum(WA_db$Age == "66+")/nrow(WA_db)

#For external validity (without 66+ group)
sum(WA_db$Age == "18-25")/sum(WA_db$Age == "18-25" | WA_db$Age ==
"26-35" | WA_db$Age == "36-45"
                                | WA_db$Age == "46-55" | WA_db$Age ==
"56-65")
sum(WA_db$Age == "26-35")/sum(WA_db$Age == "18-25" | WA_db$Age ==
"26-35" | WA_db$Age == "36-45"
                                | WA_db$Age == "46-55" | WA_db$Age ==
"56-65")
sum(WA_db$Age == "36-45")/sum(WA_db$Age == "18-25" | WA_db$Age ==
"26-35" | WA_db$Age == "36-45"
                                | WA_db$Age == "46-55" | WA_db$Age ==
"56-65")
sum(WA_db$Age == "46-55")/sum(WA_db$Age == "18-25" | WA_db$Age ==
"26-35" | WA_db$Age == "36-45"
                                | WA_db$Age == "46-55" | WA_db$Age ==
"56-65")
sum(WA_db$Age == "56-65")/sum(WA_db$Age == "18-25" | WA_db$Age ==
"26-35" | WA_db$Age == "36-45"
                                | WA_db$Age == "46-55" | WA_db$Age ==
"56-65")

#Distance details
car05 = 0
bike05 = 0
train05 = 0
car0610 = 0
bike0610 = 0
train0610 = 0
car1120 = 0
bike1120 = 0
train1120 = 0
car2130 = 0
bike2130 = 0

```

```

train2130 = 0
car3140 = 0
bike3140 = 0
train3140 = 0
car4150 = 0
bike4150 = 0
train4150 = 0
car5160 = 0
bike5160 = 0
train5160 = 0
car61100 = 0
bike61100 = 0
train61100 = 0

for (i in 1:nrow(WA_db)) {
  if (WA_db$Distance[i] == "0-5" & WA_db$Transport[i] == "Car") {
    car05 <- car05 + 1
  }
  else if (WA_db$Distance[i] == "0-5" & WA_db$Transport[i] ==
"Bike"){
    bike05 <- bike05 + 1
  }
  else if (WA_db$Distance[i] == "0-5" & WA_db$Transport[i] ==
"Public transport"){
    train05 <- train05 + 1
  }
  else if (WA_db$Distance[i] == "6-10" & WA_db$Transport[i] ==
"Car"){
    car0610 <- car0610 + 1
  }
  else if (WA_db$Distance[i] == "6-10" & WA_db$Transport[i] ==
"Bike"){
    bike0610 <- bike0610 + 1
  }
  else if (WA_db$Distance[i] == "6-10" & WA_db$Transport[i] ==
"Public transport"){
    train0610 <- train0610 + 1
  }
  else if (WA_db$Distance[i] == "11-20" & WA_db$Transport[i] ==
"Car"){
    car1120 <- car1120 + 1
  }
  else if (WA_db$Distance[i] == "11-20" & WA_db$Transport[i] ==
"Bike"){
    bike1120 <- bike1120 + 1
  }
  else if (WA_db$Distance[i] == "11-20" & WA_db$Transport[i] ==
"Public transport"){
    train1120 <- train1120 + 1
  }
  else if (WA_db$Distance[i] == "21-30" & WA_db$Transport[i] ==
"Car"){
    car2130 <- car2130 + 1
  }
}

```

```

    else if (WA_db$Distance[i] == "21-30" & WA_db$Transport[i] ==
"Bike"){
        bike2130 <- bike2130 + 1
    }
    else if (WA_db$Distance[i] == "21-30" & WA_db$Transport[i] ==
"Public transport"){
        train2130 <- train2130 + 1
    }
    else if (WA_db$Distance[i] == "31-40" & WA_db$Transport[i] ==
"Car"){
        car3140 <- car3140 + 1
    }
    else if (WA_db$Distance[i] == "31-40" & WA_db$Transport[i] ==
"Bike"){
        bike3140 <- bike3140 + 1
    }
    else if (WA_db$Distance[i] == "31-40" & WA_db$Transport[i] ==
"Public transport"){
        train3140 <- train3140 + 1
    }
    else if (WA_db$Distance[i] == "41-50" & WA_db$Transport[i] ==
"Car"){
        car4150 <- car4150 + 1
    }
    else if (WA_db$Distance[i] == "41-50" & WA_db$Transport[i] ==
"Bike"){
        bike4150 <- bike4150 + 1
    }
    else if (WA_db$Distance[i] == "41-50" & WA_db$Transport[i] ==
"Public transport"){
        train4150 <- train4150 + 1
    }
    else if (WA_db$Distance[i] == "51-60" & WA_db$Transport[i] ==
"Car"){
        car5160 <- car5160 + 1
    }
    else if (WA_db$Distance[i] == "51-60" & WA_db$Transport[i] ==
"Bike"){
        bike5160 <- bike5160 + 1
    }
    else if (WA_db$Distance[i] == "51-60" & WA_db$Transport[i] ==
"Public transport"){
        train5160 <- train5160 + 1
    }
    else if (WA_db$Distance[i] == "60+" & WA_db$Transport[i] == "Car")
{
        car61100 <- car61100 + 1
    }
    else if (WA_db$Distance[i] == "60+" & WA_db$Transport[i] ==
"Bike"){
        bike61100 <- bike61100 + 1
    }
    else if (WA_db$Distance[i] == "60+" & WA_db$Transport[i] ==
"Public transport"){

```

```

    train61100 <- train61100 + 1
  }
}

car05/sum(WA_db$Distance == "0-5")
bike05/sum(WA_db$Distance == "0-5")
train05/sum(WA_db$Distance == "0-5")

car0610/sum(WA_db$Distance == "6-10")
bike0610/sum(WA_db$Distance == "6-10")
train0610/sum(WA_db$Distance == "6-10")

car1120/sum(WA_db$Distance == "11-20")
bike1120/sum(WA_db$Distance == "11-20")
train1120/sum(WA_db$Distance == "11-20")

car2130/sum(WA_db$Distance == "21-30")
bike2130/sum(WA_db$Distance == "21-30")
train2130/sum(WA_db$Distance == "21-30")

car3140/sum(WA_db$Distance == "31-40")
bike3140/sum(WA_db$Distance == "31-40")
train3140/sum(WA_db$Distance == "31-40")

car4150/sum(WA_db$Distance == "41-50")
bike4150/sum(WA_db$Distance == "41-50")
train4150/sum(WA_db$Distance == "41-50")

car5160/sum(WA_db$Distance == "51-60")
bike5160/sum(WA_db$Distance == "51-60")
train5160/sum(WA_db$Distance == "51-60")

car61100/sum(WA_db$Distance == "60+")
bike61100/sum(WA_db$Distance == "60+")
train61100/sum(WA_db$Distance == "60+")

#Number of commutes/week (7 "0" values)
sum(WA_db$`Weekly commutes` == "1", na.rm = TRUE)/(nrow(WA_db)-7)
sum(WA_db$`Weekly commutes` == "2", na.rm = TRUE)/(nrow(WA_db)-7)
sum(WA_db$`Weekly commutes` == "3", na.rm = TRUE)/(nrow(WA_db)-7)
sum(WA_db$`Weekly commutes` == "4", na.rm = TRUE)/(nrow(WA_db)-7)
sum(WA_db$`Weekly commutes` == "5", na.rm = TRUE)/(nrow(WA_db)-7)
sum(WA_db$`Weekly commutes` == "6", na.rm = TRUE)/(nrow(WA_db)-7)
sum(WA_db$`Weekly commutes` == "6+", na.rm = TRUE)/(nrow(WA_db)-7)

which(WA_db$`Weekly commutes` == 0)

sum(WA_db$`Weekly commutes` == "1" & WA_db$Transport == "Car",
na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "1", na.rm = TRUE)
sum(WA_db$`Weekly commutes` == "2" & WA_db$Transport == "Car",
na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "2", na.rm = TRUE)
sum(WA_db$`Weekly commutes` == "3" & WA_db$Transport == "Car",
na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "3", na.rm = TRUE)

```

```

sum(WA_db$`Weekly commutes` == "4" & WA_db$Transport == "Car" ,
na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "4", na.rm = TRUE)
sum(WA_db$`Weekly commutes` == "5" & WA_db$Transport == "Car" ,
na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "5", na.rm = TRUE)
sum(WA_db$`Weekly commutes` == "6" & WA_db$Transport == "Car" ,
na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "6", na.rm = TRUE)
sum(WA_db$`Weekly commutes` == "6+" & WA_db$Transport == "Car" ,
na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "6+", na.rm = TRUE)

```

```

sum(WA_db$`Weekly commutes` == "1" & WA_db$Transport == "Bike" ,
na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "1", na.rm = TRUE)
sum(WA_db$`Weekly commutes` == "2" & WA_db$Transport == "Bike" ,
na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "2", na.rm = TRUE)
sum(WA_db$`Weekly commutes` == "3" & WA_db$Transport == "Bike" ,
na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "3", na.rm = TRUE)
sum(WA_db$`Weekly commutes` == "4" & WA_db$Transport == "Bike" ,
na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "4", na.rm = TRUE)
sum(WA_db$`Weekly commutes` == "5" & WA_db$Transport == "Bike" ,
na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "5", na.rm = TRUE)
sum(WA_db$`Weekly commutes` == "6" & WA_db$Transport == "Bike" ,
na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "6", na.rm = TRUE)
sum(WA_db$`Weekly commutes` == "6+" & WA_db$Transport == "Bike" ,
na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "6+", na.rm = TRUE)

```

```

sum(WA_db$`Weekly commutes` == "1" & WA_db$Transport == "Public
transport" , na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "1", na.rm
= TRUE)
sum(WA_db$`Weekly commutes` == "2" & WA_db$Transport == "Public
transport" , na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "2", na.rm
= TRUE)
sum(WA_db$`Weekly commutes` == "3" & WA_db$Transport == "Public
transport" , na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "3", na.rm
= TRUE)
sum(WA_db$`Weekly commutes` == "4" & WA_db$Transport == "Public
transport" , na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "4", na.rm
= TRUE)
sum(WA_db$`Weekly commutes` == "5" & WA_db$Transport == "Public
transport" , na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "5", na.rm
= TRUE)
sum(WA_db$`Weekly commutes` == "6" & WA_db$Transport == "Public
transport" , na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "6", na.rm
= TRUE)
sum(WA_db$`Weekly commutes` == "6+" & WA_db$Transport == "Public
transport" , na.rm = TRUE)/sum(WA_db$`Weekly commutes` == "6+",
na.rm = TRUE)

```

```

unique(WA_db$`Weekly commutes`)
#Bike allowance
sum(WA_db$Bike_allowance == "Yes")
sum(WA_db$Bike_allowance == "No")
sum(WA_db$Bike_allowance == "Other")
49/(49+14)

```

```

#W/F feeling

```

```
mean(WA_db$`W/F feeling`)
```

```
#Company car ownership
```

```
sum(WA_db$`Company car` == "Yes")/nrow(WA_db)
sum(WA_db$`Fuel card` == "Yes")/nrow(WA_db)
sum(WA_db$`Company car` == "Yes" & WA_db$`Fuel card` == "Yes")
sum(WA_db$`Company car` == "Yes" & WA_db$`Fuel card` == "No")
sum(WA_db$`Company car` == "No" & WA_db$`Fuel card` == "Yes")
sum(WA_db$`Company car` == "No" & WA_db$`Fuel card` == "No")
```

```
#Distance
```

```
sum(WA_db$Distance == "0-5")/nrow(WA_db)
sum(WA_db$Distance == "6-10")/nrow(WA_db)
sum(WA_db$Distance == "11-20")/nrow(WA_db)
sum(WA_db$Distance == "21-30")/nrow(WA_db)
sum(WA_db$Distance == "31-40")/nrow(WA_db)
sum(WA_db$Distance == "41-50")/nrow(WA_db)
(sum(WA_db$Distance == "51-60")+sum(WA_db$Distance == "50-60"))/
nrow(WA_db)
sum(WA_db$Distance == "60+")/nrow(WA_db)
```

```
#Bike_allowance effect
```

```
sum(WA_db$Transport == "Bike")/nrow(WA_db)
sum(WA_db$NoAllowance == "Bike")/nrow(WA_db)
sum(WA_db$Allowance15 == "Bike")/nrow(WA_db)
sum(WA_db$Allowance25 == "Bike")/nrow(WA_db)
sum(WA_db$Allowance35 == "Bike")/nrow(WA_db)
sum(WA_db$Allowance50 == "Bike")/nrow(WA_db)
```

```
sum(WA_db$Transport == "Car")/nrow(WA_db)
sum(WA_db$NoAllowance == "Car")/nrow(WA_db)
sum(WA_db$Allowance15 == "Car")/nrow(WA_db)
sum(WA_db$Allowance25 == "Car")/nrow(WA_db)
sum(WA_db$Allowance35 == "Car")/nrow(WA_db)
sum(WA_db$Allowance50 == "Car")/nrow(WA_db)
```

```
sum(WA_db$Transport == "Public transport")/nrow(WA_db)
sum(WA_db$NoAllowance == "Public transport")/nrow(WA_db)
sum(WA_db$Allowance15 == "Public transport")/nrow(WA_db)
sum(WA_db$Allowance25 == "Public transport")/nrow(WA_db)
sum(WA_db$Allowance35 == "Public transport")/nrow(WA_db)
sum(WA_db$Allowance50 == "Public transport")/nrow(WA_db)
```

```
sum(WA_db$Transport == "Walking")/nrow(WA_db)
sum(WA_db$NoAllowance == "Walking")/nrow(WA_db)
sum(WA_db$Allowance15 == "Walking")/nrow(WA_db)
sum(WA_db$Allowance25 == "Walking")/nrow(WA_db)
sum(WA_db$Allowance35 == "Walking")/nrow(WA_db)
sum(WA_db$Allowance50 == "Walking")/nrow(WA_db)
```

```
##IDENTITY TREATMENT
```

```

# Binary dependent variable
y_VLTot <- VL_db$Change
y_VLNo15 <- VL_db$ChangeNo15
y_VLNo25 <- VL_db$ChangeNo25
y_VL1525 <- VL_db$Change1525
y_VL2535 <- VL_db$Change2535
y_VLCar <- VL_db$ChangeCar
x1_VL <- VL_db$Treatment # Binary explanatory variable
x2_VL <- VL_db$`Company car`
x3_VL <- VL_db$Age
x4_VL <- VL_db$Sex

# Convert column 3 to a factor
x3_VL <- as.factor(x3_VL)

# Perform logistic regression
model_SVLTot <- glm(y_VLTot ~ x1_VL, data = VL_db, family =
binomial)
model_SVLNo15 <- glm(y_VLNo15 ~ x1_VL, data = VL_db, family =
binomial)
model_SVLNo25 <- glm(y_VLNo25 ~ x1_VL, data = VL_db, family =
binomial)
model_SVL1525 <- glm(y_VL1525 ~ x1_VL, data = VL_db, family =
binomial)
model_SVL2535 <- glm(y_VL2535 ~ x1_VL, data = VL_db, family =
binomial)

model_CVLTot <- glm(y_VLTot ~ x1_VL + x2_VL + x3_VL + x4_VL, data =
VL_db, family = binomial)
model_CVLNo15 <- glm(y_VLNo15 ~ x1_VL + x2_VL + x3_VL + x4_VL, data
= VL_db, family = binomial)
model_CVLNo25 <- glm(y_VLNo25 ~ x1_VL + x2_VL + x3_VL + x4_VL, data
= VL_db, family = binomial)
model_CVL1525 <- glm(y_VL1525 ~ x1_VL + x2_VL + x3_VL + x4_VL, data
= VL_db, family = binomial)
model_CVL2535 <- glm(y_VL2535 ~ x1_VL + x2_VL + x3_VL + x4_VL, data
= VL_db, family = binomial)

stargazer(model_SVLTot, model_SVLNo15, model_SVLNo25, model_SVL1525,
model_SVL2535, title = "Flanders' treatment effect", align = TRUE,
dep.var.labels=c("0 to 0.5", "0 to 0.15", "0 to 0.25",
"0.15 to 0.25", "0.25 to 0.35"), covariate.labels=c("Treatment"),
no.space = TRUE)

stargazer(model_CVLTot, model_CVLNo15, model_CVLNo25, model_CVL1525,
model_CVL2535, title = "Flanders' treatment effect with control
variables", align = TRUE,
dep.var.labels=c("0 to 0.5", "0 to 0.15", "0 to 0.25",
"0.15 to 0.25", "0.25 to 0.35"),
covariate.labels=c("Treatment","Company Car", "26-35", "36-45",

```

```

"46-55","56-65", "66+", "Men", "Not specified", "Other"),
  no.space = TRUE)

compcar <- glm(y_VLNo15 ~ x1_VL + x2_VL, data = VL_db, family =
binomial)
summary(compcar)
agegroup <- glm(y_VLNo15 ~ x1_VL + x3_VL, data = VL_db, family =
binomial)
gender <- glm(y_VLNo15 ~ x1_VL + x4_VL, data = VL_db, family =
binomial)

stargazer(model_SVLNo15, compcar, title = "Flanders' treatment
effect with and without controlling for car ownership",
  align = TRUE, covariate.labels = c("Treatment", "Company
car"), no.space = TRUE)

#WA
# Binary dependent variable
y_WATot <- WA_db$Change
y_WANo15 <- WA_db$ChangeNo15
y_WANo25 <- WA_db$ChangeNo25
y_WA1525 <- WA_db$Change1525
y_WA2535 <- WA_db$Change2535
y_WACar <- WA_db$ChangeCar
x1_WA <- WA_db$Treatment # Binary explanatory variable
x2_WA <- WA_db$`Company car`
x3_WA <- WA_db$Age
x4_WA <- WA_db$Sex

# Convert column 3 to a factor
x3_WA <- as.factor(x3_WA)

# Perform logistic regression
model_SWATot <- glm(y_WATot ~ x1_WA, data = WA_db, family =
binomial)
model_SWANo15 <- glm(y_WANo15 ~ x1_WA, data = WA_db, family =
binomial)
model_SWANo25 <- glm(y_WANo25 ~ x1_WA, data = WA_db, family =
binomial)
model_SWA1525 <- glm(y_WA1525 ~ x1_WA, data = WA_db, family =
binomial)
model_SWA2535 <- glm(y_WA2535 ~ x1_WA, data = WA_db, family =
binomial)

model_CWATot <- glm(y_WATot ~ x1_WA + x2_WA + x3_WA + x4_WA, data =
WA_db, family = binomial)
model_CWANo15 <- glm(y_WANo15 ~ x1_WA + x2_WA + x3_WA + x4_WA, data
= WA_db, family = binomial)
model_CWANo25 <- glm(y_WANo25 ~ x1_WA + x2_WA + x3_WA + x4_WA, data
= WA_db, family = binomial)
model_CWA1525 <- glm(y_WA1525 ~ x1_WA + x2_WA + x3_WA + x4_WA, data
= WA_db, family = binomial)
model_CWA2535 <- glm(y_WA2535 ~ x1_WA + x2_WA + x3_WA + x4_WA, data

```

```

= WA_db, family = binomial)

stargazer(model_SWATot, model_SWANo15, model_SWANo25, model_SWA1525,
model_SWA2535, title = "Wallonia's treatment effect", align = TRUE,
          dep.var.labels=c("0 to 0.5", "0 to 0.15", "0 to 0.25",
"0.15 to 0.25", "0.25 to 0.35"), covariate.labels=c("Treatment"),
          no.space = TRUE)

stargazer(model_CWATot, model_CWANo15, model_CWANo25, model_CWA1525,
model_CWA2535, title = "Wallonia's treatment effect with control
variables", align = TRUE,
          dep.var.labels=c("0 to 0.5", "0 to 0.15", "0 to 0.25",
"0.15 to 0.25", "0.25 to 0.35"),
covariate.labels=c("Treatment","Company Car", "26-35", "36-45",
"46-55","56-65", "66+", "Men", "Not specified"),
          no.space = TRUE)

#Weighted average for women/men repartition
#VL
sum(VL_db$Sex == "F")/nrow(VL_db)
weightsVL <- ifelse(x6 == "F", 0.7474, 0.5052)

# Perform the logistic regressions

wVLTot <- glm(y_VLTot ~ x1_VL, data = VL_db, family = binomial,
weights = weightsVL)
wVLNo15 <- glm(y_VLNo15 ~ x1_VL, data = VL_db, family = binomial,
weights = weightsVL)
wVLNo25 <- glm(y_VLNo25 ~ x1_VL, data = VL_db, family = binomial,
weights = weightsVL)
wVL1525 <- glm(y_VL1525 ~ x1_VL, data = VL_db, family = binomial,
weights = weightsVL)
wVL2535 <- glm(y_VL2535 ~ x1_VL, data = VL_db, family = binomial,
weights = weightsVL)

stargazer(wVLTot, wVLNo15,wVLNo25,wVL1525,wVL2535, title =
"Flanders' weighted average results", align = TRUE,
          dep.var.labels=c("0 to 0.5", "0 to 0.15", "0 to 0.25",
"0.15 to 0.25", "0.25 to 0.35"), covariate.labels=c("Treatment"),
          no.space = TRUE)

#WA
sum(WA_db$Sex == "F")/nrow(WA_db)
weightsWA <- ifelse(x6_WA == "F", 0.7929, 0.5109)

wWATot <- glm(y_WATot ~ x1_WA, data = WA_db, family = binomial,
weights = weightsWA)
wWANo15 <- glm(y_WANo15 ~ x1_WA, data = WA_db, family = binomial,
weights = weightsWA)
wWANo25 <- glm(y_WANo25 ~ x1_WA, data = WA_db, family = binomial,
weights = weightsWA)
wWA1525 <- glm(y_WA1525 ~ x1_WA, data = WA_db, family = binomial,

```

```

weights = weightsWA)
wWA2535 <- glm(y_WA2535 ~ x1_WA, data = WA_db, family = binomial,
weights = weightsWA)

stargazer(wWATot, wWANO15, wWANO25, wWA1525, wWA2535, title =
"Wallonia's weighted average results", align = TRUE,
          dep.var.labels=c("0 to 0.5", "0 to 0.15", "0 to 0.25",
"0.15 to 0.25", "0.25 to 0.35"), covariate.labels=c("Treatment"),
          no.space = TRUE)

#Graph of bike allowances

#VL
# Create the data points
x <- c(0, 0.15, 0.25, 0.35, 0.5)
y <- c(0.1778, 0.2544, 0.3259, 0.4164, 0.5263)

# Create the dataframe with the data points
df <- data.frame(x = x, y = y)

# Create the plot
ggplot(df, aes(x = x, y = y)) +
  geom_point(size = 1) +
  geom_hline(yintercept = 0.2128, linetype = "dashed", color =
"red") +
  geom_vline(xintercept = 0.25, linetype = "dashed", color = "blue")
+
  scale_y_continuous(limits = c(0, 1), expand = c(0, 0), labels =
scales::percent) +
  scale_x_continuous(limits = c(0 - 0.05, 0.5 + 0.05), expand = c(0,
0), labels = scales::dollar_format(prefix = "", suffix = "/km")) +
  labs(x = "Price per kilometer (€/km)", y = "Percentage of bike
commuters") +
  theme_minimal()

# Create the dataframe with the data points
df <- data.frame(x = x, y = y)

# Perform linear regression
lm_model <- lm(y ~ x, data = df)

# Print the regression summary
summary(lm_model)

# Create the data points
x <- c(0, 0.15, 0.25, 0.35, 0.5)
y <- c(0.0316, 0.0732, 0.1097, 0.1705, 0.2604)

# Create the dataframe with the data points
df <- data.frame(x = x, y = y)

# Create the plot
ggplot(df, aes(x = x, y = y)) +
  geom_point() +
  # Add a line connecting the

```

```

points
  geom_hline(yintercept = 0.0391, linetype = "dashed", color = "red")
+ # Add a horizontal line at 21.28%
  geom_vline(xintercept = 0.25, linetype = "dashed", color = "blue")
+
  scale_y_continuous(limits = c(0, 1), expand = c(0,0), labels =
scales::percent) + # Set the y-axis limits and format as percentage
  scale_x_continuous(limits = c(0 - 0.05, 0.5 + 0.05), expand =
c(0,0), labels = scales::dollar_format(prefix = "", suffix = "/km"))
+ # Set the x-axis limits and format as currency
  labs(x = "Price per kilometer (€/km)", y = "Percentage of bike
commuters") +
  theme_minimal() # Use a minimalistic theme

# Create the dataframe with the data points

# Perform linear regression
lm_model <- lm(y ~ x, data = df)

data <- data.frame(
  Labels = c("18-25", "26-35", "36-45", "46-55", "56-65", "66+"),
  Shares = c(7.22, 23.41, 29.14, 27.77, 12.08, 0.37)
)

# Create the pie chart
ggplot(data, aes(x = "", y = Shares, fill = Labels)) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar("y", start = 0) +
  labs(title = "Shares by Age Group",
       caption = "18-25, 26-35, 36-45, 46-55, 56-65, 66+") +
  theme_minimal() +
  theme(axis.text.x = element_blank())

```