

# **Facteurs influençants la prise des transports en commun**

RACHIDI Mustapha & SAUNIER Florent & SAADALLAH Malek

**Janvier 2023**

## Introduction

Ce projet se base sur des données récoltées en 2010 dans la région Grenobloise. L'étude a pour but de déterminer les facteurs influençant la prise des transports en commun . Pour cela nous nous sommes pris comme limites : le réseau Mtag qui comprend les bus qualifiés de "ville" (Nous n'avons pas pris en compte les bus régionaux comme par exemple le bus Grenoble - Chamrousse) et le réseau du trammway dont les lignes depuis 2010 ont été augmentées.

## Articles de la littérature

### Familiarisation avec la base de données

La base de données contient 30 702 lignes et 116 colonnes ce qui correspond à nos variables , on peut la qualifier de base de données "moyenne" mais qui saura nous occuper. Concernant le nombre de valeurs manquantes, toutes variables confondues nous avons 971 658 valeurs manquantes soit 27.3% de notre base de données. De plus, 0% des lignes ont toutes leurs valeurs et c'est 21% des colonnes qui n'ont pas de valeurs manquantes. Il peut être intéressant de voir où sont les valeurs manquantes.

*Visualisation valeurs manquantes*

```
data_2<-DB_projet_full[,c(22:30)]
vis_miss(
  data_2,
  cluster = FALSE,
  sort_miss = FALSE,
  show_perc = TRUE,
  show_perc_col = TRUE,
  large_data_size = 9e+05,
  warn_large_data = TRUE
)
```

