

# Présentation R

Arthur Lemoine, François Somville, Alexandre Antippas

Janvier 2019

# Plan

1. Introduction
2. Exploration du Dataset
3. Statistiques descriptives
4. Analyse des composantes principales (ACP)
5. Clustering
6. Bonus

## 1. Introduction

Le dataset contient toutes les factures émises par un magasin en ligne basé en Angleterre. Le magasin vend des cadeaux uniques pour toutes les occasions.

Le dataset contient: nb de lignes

```
##      InvoiceDate
## 1 1/12/10 08:26
## 2 1/12/10 08:26
## 3 1/12/10 08:26
## 4 1/12/10 08:26
## 5 1/12/10 08:26
```

```
##      InvoiceDate
## 541905 9/12/11 12:50
## 541906 9/12/11 12:50
## 541907 9/12/11 12:50
## 541908 9/12/11 12:50
## 541909 9/12/11 12:50
```

# Exploration des données dans le Dataset

Nombre d'enregistrements

```
## [1] 541909
```

Exploration des variables

```
## [1] 8
```

Ajouter la signification des variables

Nombre de pays différents

```
## [1] 145572      -5
```