Description

Dataset of 550 000 observations about the black Friday in a retail store, it contains different kinds of variables either numerical or categorical. It contains missing values.

User\_ID

Product\_ID - Id Product

Gender – Boolean (M/F)

Age - Age of customer

Occupation - Occupation of each customer

City\_Category

Stay\_In\_Current\_City\_Years

Marital\_Status

Product\_Category\_1

Product\_Category\_2

Product\_Category\_3

Purchase-Purchase amount in dollars

**Approach**

## #Purchase Amount Prediction - Linear Regression.

# #ML model (Random Forest) because most features are discrete, Random Forest is expected to fit the data well.