The use case will be working on customer Payment Delays model.

▪You are firstly expected to find out the reasons for customer payment delays by doing an Exploratory Data Analysis and suggest counter measures.

▪Secondly you are expected to build a machine learning model and use it on new data to identify top 300 customers that are expected to delay in payments.

The Challenge

The overall goal is to determine if a customer is about to register a payment delay or not.

The following steps need to be accomplished:

1) Do an Explorative Data Analysis (EDA) and document interesting patterns, data issues and ideas.

Document your findings in an Jupyter Notebook file which can easily presented.

2) Find a model to predict if a customer will register a payment delay or not. Consider selecting your features and data pre-processing depending on your models used.

3) You are given a list (test.csv) of 2,000 customers to asses if they are likely to register payment delays in the following period. Since your action time is limited, you need to extract 300 customers from the list who will then be addressed directly by the company.

Please submit the indices [1,2000] of the first 300 customers from the test.csv list which are likely to register payment delays.

You will present your findings in a 10-15min talk addressing:

- The results of the EDA

- Models used, model selected and performance characteristics