

Homework 3 (EDA part 2)

EDA - **Default of Credit Card Clients** (100 points)

Due Date: Monday Feb 8 at 11:59 pm

Instruction:

- This is a group-work assignment!
- You are expected to submit the **.ipynb** file and the exported **.html**.
- Only one member in each group needs to submit the assignment. It will be automatically submitted for the rest of group members.
- This is a long assignment, start early!

Question 1 I need you to do a complete EDA on credit Card Default data set. **Please note that just generating tables and graphs is not enough!** You should make some comments for each table and graph. Once you are done with cleaning and restructuring the variables, you can save it as a new csv file and use it for your future classification exercises. We will revisit this data set later in the course.

Hint 1: make sure you explore other people's codes for this Kaggle competition. There are MANY good EDAs written for this data set on Kaggle!

Hint 2: There is not necessarily a right of wrong answer for this exercise but I expect you to deliver something like what we did in class for the wage data set!

Where can you find the data set?

Please find the **.csv** file on the Kaggle website! scroll down to the **Data Explorer** section and you should find a csv file named as **UCI-Credit-Card.csv** 2.73MB

Have fun with Python and GOOD LUCK!