# Homework 1
# Preprocessing

Due Date: Sept 16th
Assigned Date: Sept 9th

Please follow the instructions below

1. Import the required libraries
2. Load the dataset
3. Take care of the missing data ('nan' value): fill them with average value
4. Encode the categorical variable:
    ① Encode the independent variable
    ② Encode the dependent variable
5. Split train set and test set with parameters: test_ size = 0.2, random_ state = 0
6. Apply feature scaling

**Write your name and USC-ID at the top of your code.**
**Upload your code in .ipynb format with the running result in it.**
**Upload your code's screenshot (make it in .pdf format) or export it to .pdf**
**You are recommended to use Google Colab for your homework**

New Coders **READ ME!!**:
If you are not familiar with Python, pip, and conda(mamba):
    **1.** I strongly recommend using Google Colab
    **2.** If you want to run Python on your own device, do your own research / ask your friend

If you get an error:
    **1.** read the error, try to understand it
    **2.** If you do not understand it, it is fine, copy paste the error into Google, click on the first link, generally there will be an answer/explanation to it
    **3.** If it turns into another error, congrats, you are moving forward, repeat 1 & 2 until it is fully resolved!
    **4.** If you cannot fix it for 1.5 hours, seek help to other students