Problem Description

Objective: Build machine learning models to predict whether income exceeds $50K/yr based on census data.

Requirements:

* Please use python jupyter notebook to finish this exercise
* Perform feature exploration, feature engineering and modelling steps and explain each step using basic comments
* For feature exploration performing data visualization would be a bonus.
* Strongly recommend trying out various algorithms such as logistic regression, decision trees, random forest, xgboost etc.
* You can use python libraries such as pandas, numpy, sklearn, seaborn, matplotlib, plotly etc.

Submission:

* For each model built, please specify the evaluation metric (ROC, accuracy) for the testing dataset
* Please upload a python notebook which includes all the analysis code and visualizations to your github account, and share us the link of the github repository

Hint: To install the jupyter notebook using Anaconda on your local machine, you can follow the tutorial below:

<http://jupyter.readthedocs.io/en/latest/install.html>