Team Member Details:

Name	Email	Country	College/Company	Specialization
Jane Condon	jane.condon@stonybrook.edu	United States	Stony Brook University	Data Science
Jeffery Su	Jefferysu556@gmail.com	Canada	University of Toronto	Data Science
Seyed Mojtaba Hejazi	Mojtaba.hejazi991368@gmail.com	Italy		Data Science

Problem Description:

ABC Bank wants to sell its term deposit product to customers and before launching the product they want to develop a model which helps them in understanding whether a particular customer will buy their product or not (based on customer's past interaction with bank or other Financial Institution).

Data Cleansing and Transformation Techniques

First, we checked the data for missing values and found that there were zero missing values in both the full dataset and the smaller dataset, so we did not need to remove any missing or NA values. We also checked for outliers using data visualization. After looking at both scatterplots and boxplots of the numerical variables, we noticed that most of the numerical variables (‘age’, 'balance', 'duration', 'campaign', 'pdays', and 'previous') had many outliers. Rather than removing the outliers, we used various transformation techniques on the data to deal with both the skewness and outliers in the data. Some of the transformations used were a log transformation, square-root transformation, and box-cox transformation.

Github Repo URL: https://github.com/janecondon/Data-Glacier-Internship-Group-Project.git