GROUP NAME: DG TEAM

NAME	EMAIL	COUNTRY	COLLEGE/COMPANY	SPECIALIZATION
Sean Chen	seancchen98@gmail.com	United States	Data Glaciers	Data Analyst
Allen Ayodeji	allenzeereeh@gmail.com	Nigeria	Data Glaciers	Data Science

PROBLEM DESCRIPTION : ABC Bank wants to sell it's term deposit product to customers and before launching the product they want to develop a model which help 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 UNDERSTANDING : The Bank wants to use a ML model to shortlist customers whose chances of buying the product is more so that their marketing channel (telemarketing, SMS/email marketing etc) can focus only on those customers whose chances of buying the product is more. This will save resources and their time (which is directly involved in the cost (resource billing)).

WHAT TYPE OF DATA DO YOU HAVE FOR ANALYSIS: There are four datasets:
1) Bank-additional-full.csv with all examples (41188) and 20 inputs, ordered by date (from May 2008 to November 2010), very close to the data analyzed in [Moro et al., 2014]
2) Bank-additional.csv with 10% of the examples (4119), randomly selected from 1, and 20 inputs.
3) Bank-full.csv with all examples and 17 inputs, ordered by date (older version of this dataset with less inputs).
4) Bank.csv with 10% of the examples and 17 inputs, randomly selected from 3 (older version of this dataset with less inputs).

WHAT ARE THE PROBLEMS IN THE DATA: There are some values that are unknown in the dataset. Some of these include the columns for “Default”, “Housing”, “Education”, and “Loan”.

WHAT APPROACHES ARE YOU TRYING TO APPLY ON YOUR DATA SET TO OVERCOME THE PROBLEMS: We used one hot encoding to convert the “unknown” elements to numeric form.

REPO LINK: https://github.com/ALLEN-AYODEJI/BANK-MARKETING-CAMPAIGN-