Group Type: Individual
Name: Alex Lindberg
Email: alexlindberg149@gmail.com
Country: United States
College: Ohio State University
Specialization: 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).

Business Understanding:
The key variables that we need to predict are whether or not a client will buy the companies product. Understanding this will lead the company to be much more efficient in where they designate their customer service resources.

Project Lifecycle:
Understand the data, explore the data (data visualization), prepare the data (involves wrangling and feature engineering), test different models, select a model, report performance of the model, deploy the model using flask, convert ML metrics to business metrics (explain in simple terms), place results into presentation.

Data Intake Report: NEXT PAGE

Github Repo: https://github.com/AlexLindberg9

Data Intake Report

Name: Data Science Project Report
Report Date: 08/19/2022
Internship Batch: LISUM11:30
Version: 1.0
Data Intake by: Alex Lindberg
Data Storage Location : https://github.com/AlexLindberg9

Total number of Observations: 45211
Total number of files: 4
Total number of features: 17
Base format of the file: csv
Size of the data: 5.8 MB, 8 KB, 586 KB

Proposed Approach:
If two rows have identical client features (age, job, marital status, education, housing, default and loan) then this will identify a duplicate client and one of the rows will need to be removed. I have the following assumptions: 'Job' contains all possible jobs that these clients could have, everyone has some sort of contact type, 'poutcome' contains all possible outcomes of the previous campaign and no other outcomes are possible.