WEEK 7 DELIVERABLES

Group Name: JSSN

Name	Email	Country	University /Company	Specialization
Mha Luqman Salameh	mahasa179@gmail.com	Jordan	AABU	Data Science
Almudena Zhou Ramírez	almu180@gmail.com	Spain	UNED	Data Science
Shoug Alotaibi	Shouga.1417@gmail.com	Saudi Arabia	KFUPM	Data Science
Peter Okwukogu	peter.okwukogu@gmail.com	Nigeria	Colab Kaduna	Data Science

PROBLEM DESCRIPTION

ABC bank wants to sell term deposits to their customers, but they want to identify particular customers with a higher propensity to buy.

There's a need to identify these customers so they can focus their marketing campaigns efficiently and effectively.

BUSINESS UNDERSTANDING

A term deposit is a fixed-term investment that includes the deposit of money into an account at a financial institution. Term deposit investments usually carry short-term maturities ranging from one month to a few years and will have varying levels of required minimum deposits.

Term deposits have significant advantages over normal deposits because financial institutions can use them for safe investments. With term deposits, banks can invest in businesses that have higher rates of return and profit returns.

Additionally, the bank will like to save costs through efficient and effective marketing campaigns

PROJECT LIFE CYCLE ALONG WITH THE DEADLINE

1. **Identifying the problem**: ABC wants to sell its term deposits to customers. However, the bank would like to know which customers have a higher probability of buying the term deposits.
   Why is this important? The bank campaign for selling term deposits needs to be efficient and effective. This will reduce costs and save time because they will be selling to customers with a higher propensity to buy the term deposits.
2. **Data Investigation and Cleaning**: The data has been provided, so we are going to document the data quality, confirm if the data has problems e.g. missing values, unusual values, outliers etc., and confirm how we would solve the data problems.
3. **Perform Exploratory Data Analysis:** Use this process to accept or reject hypotheses or answer questions developed from the business understanding and problem identification.
4. **Propose Modeling Techniques**: After EDA, we should have conceptualized our problem, answered crucial questions, then we propose a model type to build for prediction.
5. **Model Selection and Model Building**

DATA INTAKE REPORT

Group Name: JSSN
Report Date:
Internship Batch:
Data Intake By JSSN
Data Storage Location:

Tabular data details: **bank-additional-full.csv**

Total number of observations	41188
Total number of files	1
Total number of features	21
Base format of the file	.csv
Size of the data	5.83MB

PROPOSED APPROACH FOR DATA QUALITY ANALYSIS

1. Confirm if the number of examples (observations) and inputs(features) are consistent with the data description document. Check if column names, numbers of rows and columns are consistent with the data description document
2. Check if there are duplicate values and if these duplicate values affect the data consistency.
3. Check for missing data and how the missing data affects data integrity
4. Check for the uniqueness of data values across columns to confirm for unusual values. For instance, is someone having an unusual age value like 200?

GITHUB REPO LINK

iPablo26/DataScientistsGroupProject (github.com)