Executive Post Graduate Certification in Data Analytics- Project

Case Study: Bank Marketing Campaign Analysis

Dataset Overview

The dataset contains information about a marketing campaign carried out by a Portuguese banking institution. The marketing campaigns were based on phone calls. The objective is to identify patterns and insights that can help improve future marketing strategies.

Objectives

The goal of this case study is to perform a comprehensive analysis of the dataset using Python, SQL, and Power BI, with a focus on statistical analysis, machine learning, and advanced database queries and visualizations.

Dataset Link -
https://drive.google.com/file/d/1RXd7kM4dqfblxBUQBxHORyeSD9qTABd9/view?usp=drive_link

Python Assignment

1. Data Cleaning and Preprocessing:

- Handle missing values: Identify and appropriately handle missing values in the dataset.
- Standardize columns: Ensure all categorical data is standardized (e.g., consistent capitalization, removing leading/trailing spaces).
- Perform encoding for y column(categorical)

2. Exploratory Data Analysis (EDA):

- Demographic Analysis: Analyze the age distribution and identify trends based on job, marital status, and education.
- Loan Analysis: Investigate the relationship between housing loan and personal loan status.
- Campaign Analysis: Examine the distribution of campaign contacts and their success rate.

3. Statistical Analysis:

- Hypothesis Testing: Conduct hypothesis testing to determine if there is a significant difference in balance between clients with and without a housing loan.

- Correlation Analysis: Analyze the correlation between campaign-related variables (e.g., duration, number of contacts) and the success rate of term deposit subscriptions.

4. Machine Learning:

- Predictive Modeling: Build a classification model to predict whether a client will subscribe to a term deposit based on the provided features. Evaluate the model using appropriate metrics (e.g., accuracy, precision, recall).

SQL Assignment

1. Complex Queries:

- Top Jobs by Balance:Write a query to find the top 5 job types with the highest average balance.
- Campaign Success Rate: Use joins to calculate the success rate of the campaign for different education levels.

2. Views:

- Client Summary View: Create a view that summarizes key information for each client, including average balance, number of loans, and contact details.
- Campaign Performance View: Create a view that shows the performance of the campaign by month, including the number of contacts and success rate.

3. Indexes:

- Index Creation: Create indexes on columns that are frequently queried (e.g., age, job, marital) to optimize query performance.

4. Stored Procedures:

- Update Balance Procedure: Create a stored procedure to update the balance of clients based on their transaction history.

Power BI Assignment

1. Data Transformation:

- Import and Transform: Import the dataset and perform necessary transformations, including handling missing values and standardizing data formats.

2. Visualizations:

- Dashboard Creation: Create a dashboard to visualize demographic information (e.g., age distribution, job types) and campaign performance.
- Loan Analysis: Develop visualizations to show trends in housing and personal loan status.

3. Advanced Analysis:

- DAX Functions: Use DAX functions to calculate custom metrics, such as average balance by marital status or education level.
- Interactive Visualizations: Create interactive visualizations that allow users to filter data by campaign success, contact month, and client demographics.

4. Reports and Insights:

- Detailed Report: Generate a detailed report summarizing key insights from the dataset, such as the most effective contact methods and client demographics most likely to subscribe to a term deposit.

5. Power BI Services:

- Row-Level Security (RLS): Implement RLS to ensure that different users can only access data relevant to their role.
- Power BI Service Integration: Publish the report to Power BI Service and configure scheduled refreshes and data alerts.