Bank Marketing Campaign Data Analysis

Executive Summary

Introduction

This project analyzes a bank marketing campaign dataset, focusing on understanding client behaviors and evaluating the effectiveness of direct marketing campaigns by a Portuguese bank. The aim is to provide actionable insights that can enhance client targeting, improve subscription rates, and optimize marketing strategies.

Key Findings

- **Demographic indicators**—such as age, job type, marital status, and education—have noticeable patterns influencing campaign responses.
- **Contact metrics**—number and duration of contacts plus previous outcomes—are key predictors of campaign success.
- **Economic context**, including consumer price index and Euribor rates, is relevant to client decision-making.
- **Preprocessing**, like handling missing data and standardizing categories, is essential for reliable analysis.

Recommendations

- Focus future campaigns on demographic segments with higher subscription rates.
- Time campaigns around months with above-average effectiveness.
- Use economic indicators to refine predictions and campaign timing.

Technical Report

1. Introduction

This section details the journey from data cleaning to exploratory data analysis and key findings supporting predictive analytics and marketing improvement.

2. Data Overview

Key Fields:

Column	Description
client_id	Unique client identifier
age	Client's age
job	Occupation (e.g., housemaid, services, admin)
marital	Marital status (e.g., married)
education	Education (e.g., basic.4y, high_school)
credit_default	Whether client has credit in default
mortgage	Client mortgage indicator
number_contacts	Number of contacts in current campaign
contact_duration	Duration of last contact (seconds)
previous_campaign_contacts	Number of previous campaign contacts
previous_outcome	Previous campaign result
campaign_outcome	Current campaign result (yes/no)
month, day, year	Timing of last contact
cons_price_idx	Consumer price index
euribor_three_months	3-month Euribor rate

3. Methodology

Data Cleaning

- Transformed date columns into consistent datetime formats.
- Standardized categorical columns (job, education).
- Imputed or removed missing values for analysis integrity.

Sample cleaning code:

```
import pandas as pd

# df = pd.read_csv('bank_campaign.csv')
# df['last_contact_date'] = pd.to_datetime(df[['year','month','day']])
# df.fillna(method='ffill', inplace=True)
```

Exploratory Data Analysis (EDA)

- Distribution plots for age, occupation, marital status, and education.
- Analysis of contact frequency, duration, and campaign success.
- Correlation heatmaps to examine the relationship between economic indicators and subscription likelihood.

4. Exploratory Data Analysis

Age Distribution

- Majority are aged 30–60.

Job and Marital Status

- Most clients work in services or admin jobs.
- Married status is most common.

Campaign Outcomes

- Longer contact duration and prior positive contacts are linked to higher success.
- Certain months (notably May) show increased campaign effectiveness.
- Clients without credit/default or mortgage are more likely to subscribe.

5. Key Results

- Higher subscription rates for clients with prior positive engagements.
- Seasonality and economic environment matter for campaign success.
- Data cleaning and proper feature engineering boost prediction accuracy.

6. Conclusions

- Deploy personalized, well-timed campaigns for likely subscribers.
- Continuously update campaign strategy using new client and economic information.
- Consider machine learning for even better targeting.