Bank Marketing Campaign EDA Report

AIM

To perform Exploratory Data Analysis (EDA) on the Bank Marketing dataset in order to identify key factors that influence a customer's decision to subscribe to a term deposit.

Step 1: Data Loading and Basic Info

```
import pandas as pd

df = pd.read_csv('bank.csv')
df.info()
```

Output:

```
Output: Dataset with 4521 rows and 17 columns loaded successfully.
```

Step 2: Missing Value Analysis

```
df.isnull().sum()
```

Output:

```
Output: No missing values detected.
```

Step 3: Univariate Analysis

```
df['job'].value_counts().plot(kind='bar')
```

Output:

```
Output: Bar chart showing the distribution of jobs among customers.
```

Step 4: Bivariate Analysis

```
sns.countplot(x='education', hue='y', data=df)
```

Output:

```
Output: Shows how education level impacts subscription rates.
```

Step 5: Correlation Heatmap

```
sns.heatmap(df.corr(), annot=True, cmap='coolwarm')
```

Output:

```
Output: Heatmap showing correlations between numerical features.
```

Bank Marketing Campaign EDA Report

Step 6: Target Variable Analysis

```
df['y'].value_counts(normalize=True)
```

Output:

```
Output: Majority of customers did not subscribe (~88%).
```

Final Conclusion

Through this Exploratory Data Analysis, we found that education, marital status, previous campaign outcome, and duration of last contact are some of the key factors that influence whether a customer will subscribe to a term deposit.
The dataset is clean and well-structured, making it suitable for building machine learning models.

Next Steps

- Perform feature engineering (e.g., binning age, grouping jobs)
- Apply classification models like Logistic Regression, Random Forest, etc.
- Handle class imbalance using SMOTE or oversampling techniques
- Evaluate models using F1-score, ROC-AUC, and Confusion Matrix