

**Goal: For the given marketing dataset, predict whether the customer will subscribe or not for a term deposit (yes/no).**

**What is a term deposit?**

A term deposit is like a fixed deposit where consumer invests their money for a fixed time for fixed interest rate.

**Information about the data:**

- Data has 17 columns and 45211 rows.
- Calculated counts of each entry from the respective feature.
- Calculated count of unique entries from the features.
- Data has not missing values.

**Visualization of Categorical Features:**

- Plotted distribution of categorical features.
- Plotted distribution of categorical features concerning output label.

**Visualization of Numerical Features:**

- Plotted distribution of continuous features.
- Plotted boxplot to get an idea about the outliers.

**Correlation Matrix:**

- Plotted heatmap

**Visualization of Labelled Feature:**

- Data is imbalanced.

**Data Pre-processing:**

- Removed outliers & unnecessary columns.
- Data Encoded.
- Did Undersampling.

**Splitting of Data:**

Data is split with a test size of 75%.

**Building of Classifier:**

Various classifiers are developed with their feature importance.