BANK INSTITUTION TERM DEPOSIT SUBSCRIPTION ANALYSIS: EDA
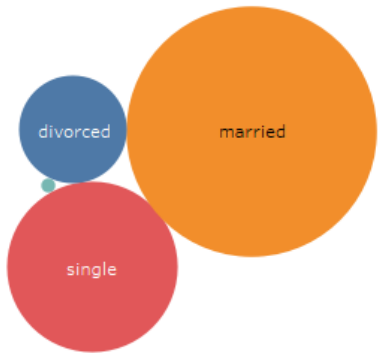
File created on: 8/25/2020 5:06:14 PM

BANK INSTITUTION TERM DEPOSIT SUBSCRIPTION ANALYSIS
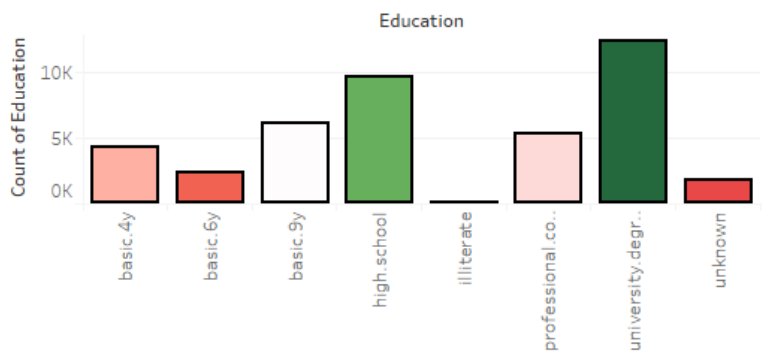
The four plots below indicates the value counts of each category under the following categorical variables: *MaritalStatus, Education, JobType* and *Term Deposit Subscription (target variable).*
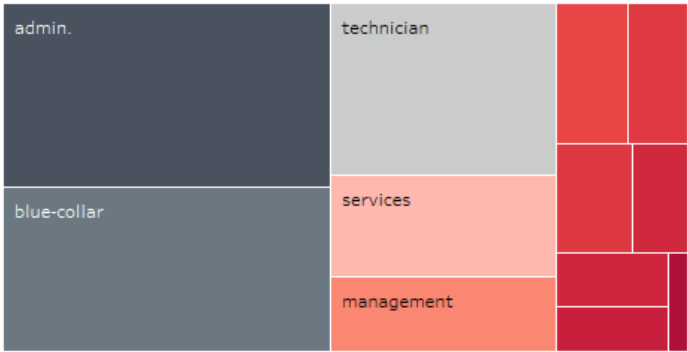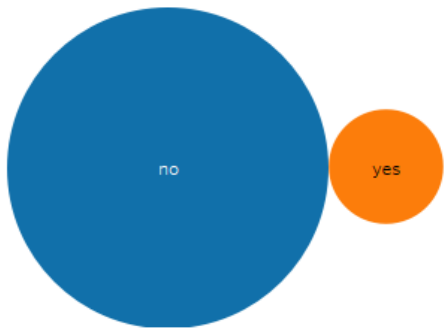The latter has a high degree of class imbalance.

Marital Status

Education Types Popularity

Job Types

Term Deposit Subscription (Target Variable) Counts

BANK INSTITUTION TERM DEPOSIT SUBSCRIPTION ANALYSIS

These 3 plots are counts plots of the variables: ***Housing, CreditDefaults*** and ***LoanStatus.***
The *unknown* category has the least count in the ***Housing*** and ***LoanStatus*** features while the *yes* category has the least count in the ***CreditDefaults*** feature.
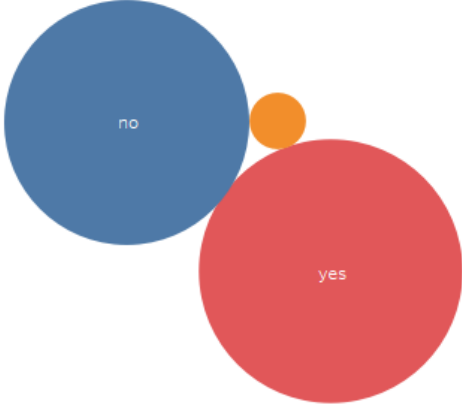
Housing
- no
- unknown
- yes

Loan
- no
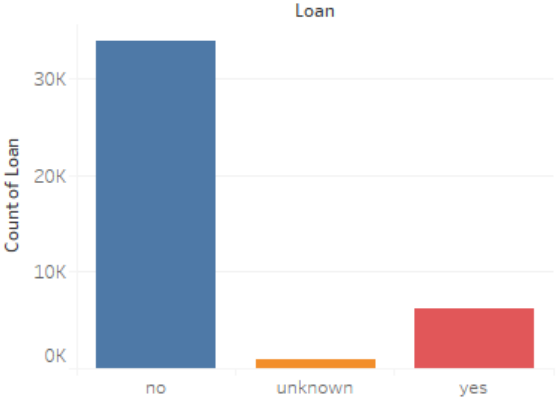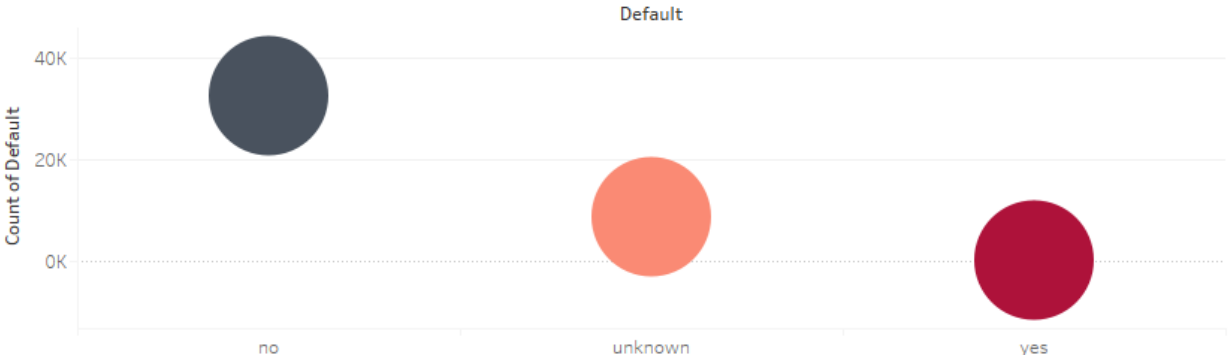- unknown
- yes

Count of Default
3 — 32,588

Housing Loans
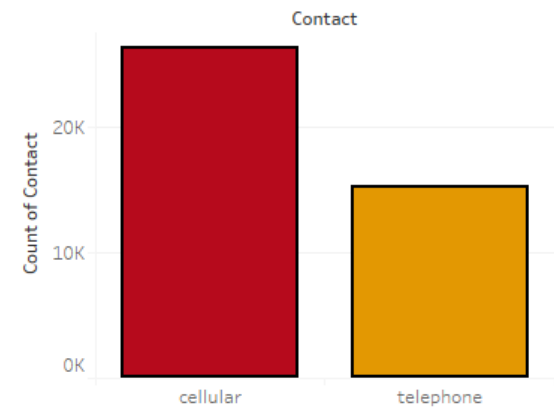

Loan Status

Credit Defaults Value Counts

BANK INSTITUTION TERM DEPOSIT SUBSCRIPTION ANALYSIS

In these 4 plots, 3 of them are count plots of the features: **Contact, DayOfWeek** and **Month.** The remaining one plot is a histogram showing distribution of the last contact with customer duration.

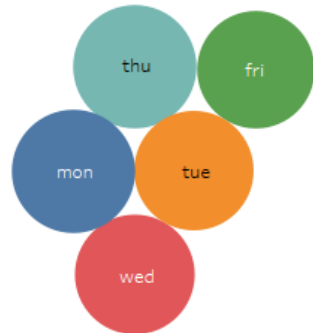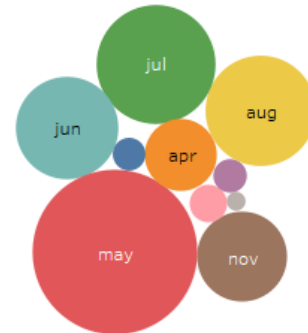## Communication Type Value Counts

Last Contact Duration Distribution
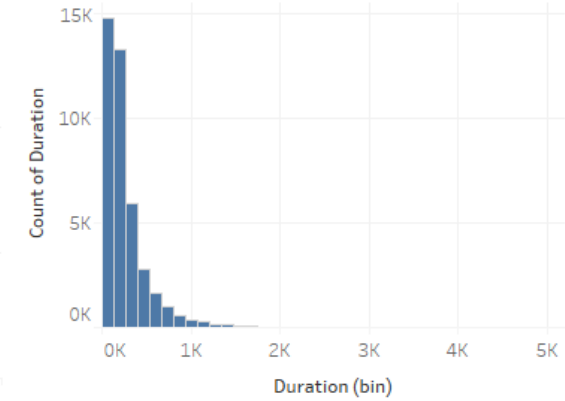
DayOfWeek Value Counts
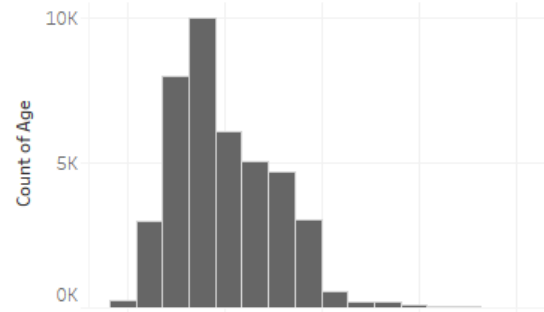

Last Month of Contact Value Counts
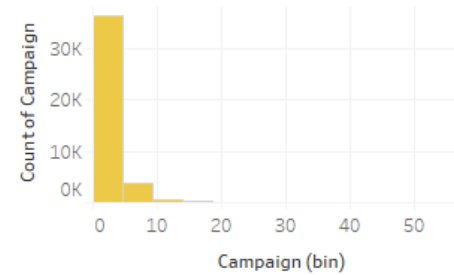
BANK INSTITUTION TERM DEPOSIT SUBSCRIPTION ANALYSIS

These 6 histograms show the distribution of the following numerical features: *Age, ConsumerConfidenceIndex, ConsumerPriceIndex, Euribor 3 MonthRate, Campaigns* and *LastContactDuration.*
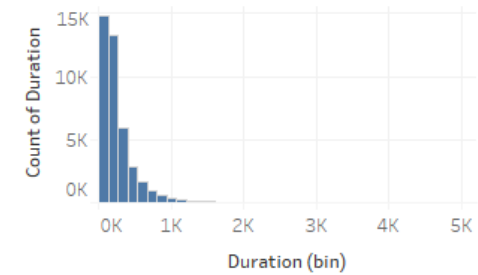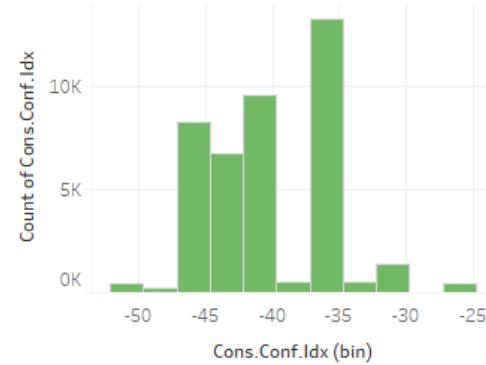

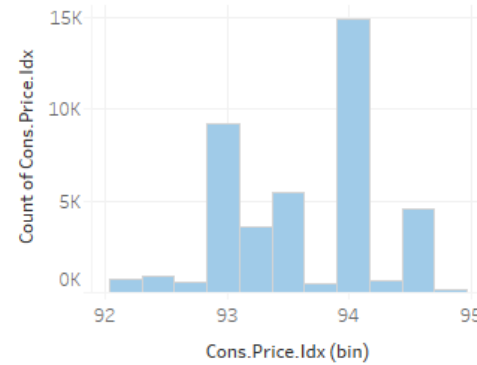Distribution of customers' ages.

Number of Campaigns Distribution

Last Contact Duration Distribution
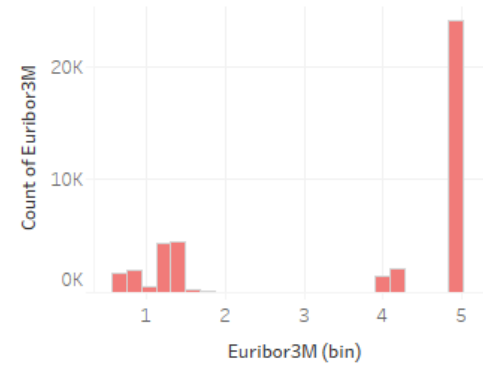
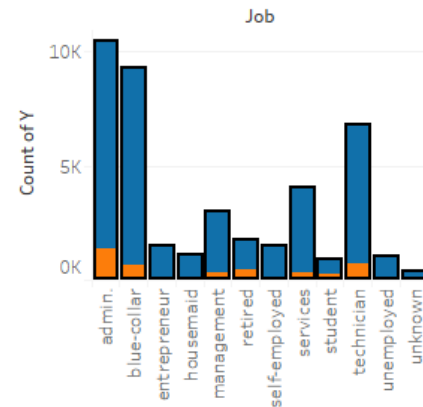Consumer Confidence Index

Consumer Price Index

Euribor 3 month rate

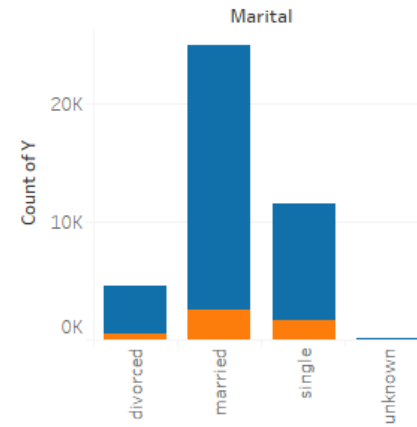# BANK INSTITUTION TERM DEPOSIT SUBSCRIPTION ANALYSIS

These 6 plots show the relationship between the categorical variables and the target variable.
Due to the class imbalance observed in the first workbook, the **no** class in the target variable is bound to have a greater percentage.
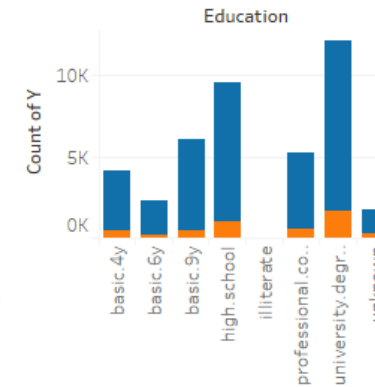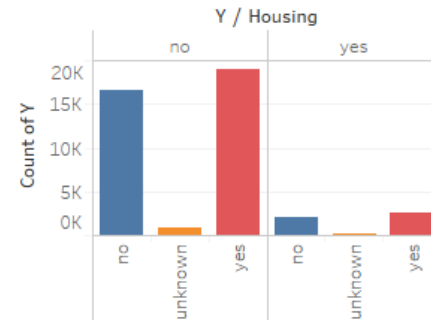
Y
- no
- yes

Housing
- no
- unknown
- yes

Loan
- no
- unknown
- yes

Default
- no
- unknown
- yes

Subscriptions Under Different Job Types.

Job

Subscriptions Under Different Marital Statuses

Marital

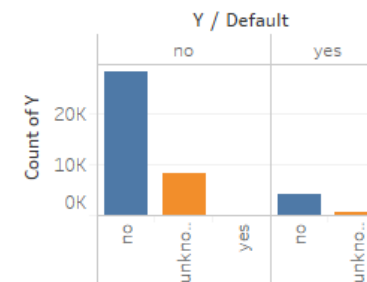### Subscriptions Under Different Education Categories

Education

Subscriptions Under Different Housing Loan States

Y / Housing

Subscriptions Under Different Loan Statuses

Y / Loan
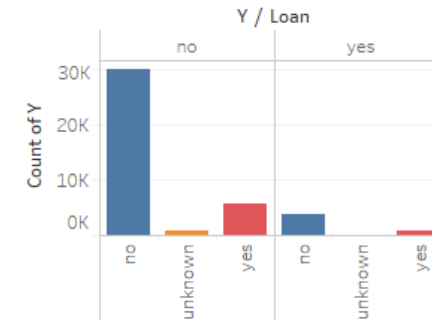
Subscriptions Under Different Credit in Default Classes

Y / Default

BANK INSTITUTION TERM DEPOSIT SUBSCRIPTION ANALYSIS

This grouped barplot shows the relationship between the **DayOfWeek** and **Month** in relation to the count of the target variable. May appears to be the most active month while December being the least active.

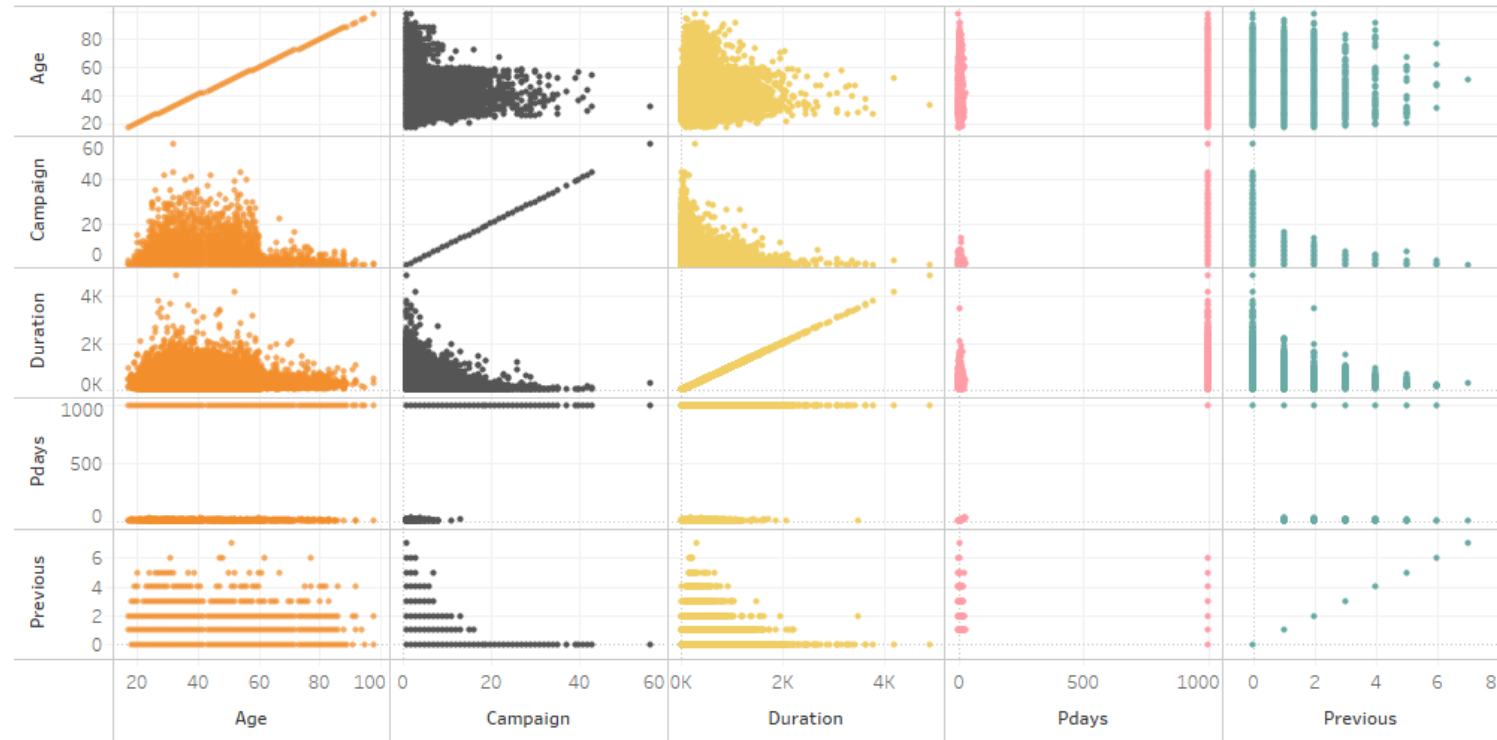### DayOfWeek/Month in relation to the Subscription.

BANK INSTITUTION TERM DEPOSIT SUBSCRIPTION ANALYSIS

The plot below shows the relationships between different numerical features.
None of the pairs shows any linear relationship.

Pairplot

BANK INSTITUTION TERM DEPOSIT SUBSCRIPTION ANALYSIS

The plot below shows the relationships between different numerical features.
None of the pairs shows any linear relationship. The Euribor feature has quite a number of null values.

Pairplot