



Data Glacier

Your Deep Learning Partner

Exploratory Data Analysis

<Marketing Strategy>

<2023.1.20>

Agenda

Executive Summary

Problem Statement

Approach

EDA

EDA Summary

Model Suggestion

Executive Summary

- Clean and preprocess the dataset for analysis
- Analyze features and their associations with customer decision
- Summarize the feature importance and suggest ML models

Problem Statement

- This project aims to fix the problem that there are too many potential clients associate with a Portuguese banking institution. The bank marketing campaigns are primarily based on phone calls. However, there are too many potential clients, and it is impossible to call all of them. Hence, some machine learning techniques are needed to classify the customers and predict if they are going to subscribe the term deposit or not. By implementing ML models, we are able to limit the number of 'potential clients' to an acceptable number that make sure the marketing strategies can be carried on.

Approach

- Data Loaded & Cleaning & Preprocessing

Perform basic cleaning & Preprocessing

```
[5]: my_df.isnull().any()
```

```
Out[5]: age                False
        job                False
        marital            False
        education           False
        default            False
        housing            False
        loan               False
        contact            False
        month              False
        day_of_week        False
        duration           False
        campaign           False
        pdays              False
        previous           False
        poutcome           False
        emp.var.rate       False
        cons.price.idx     False
        cons.conf.idx      False
        euribor3m          False
        nr.employed        False
        y                  False
        dtype: bool
```

```
In [7]: my_df.duplicated().any()
```

```
Out[7]: True
```

```
In [11]: my_df = my_df.drop_duplicates()
         my_df.shape
```

```
Out[11]: (41176, 21)
```

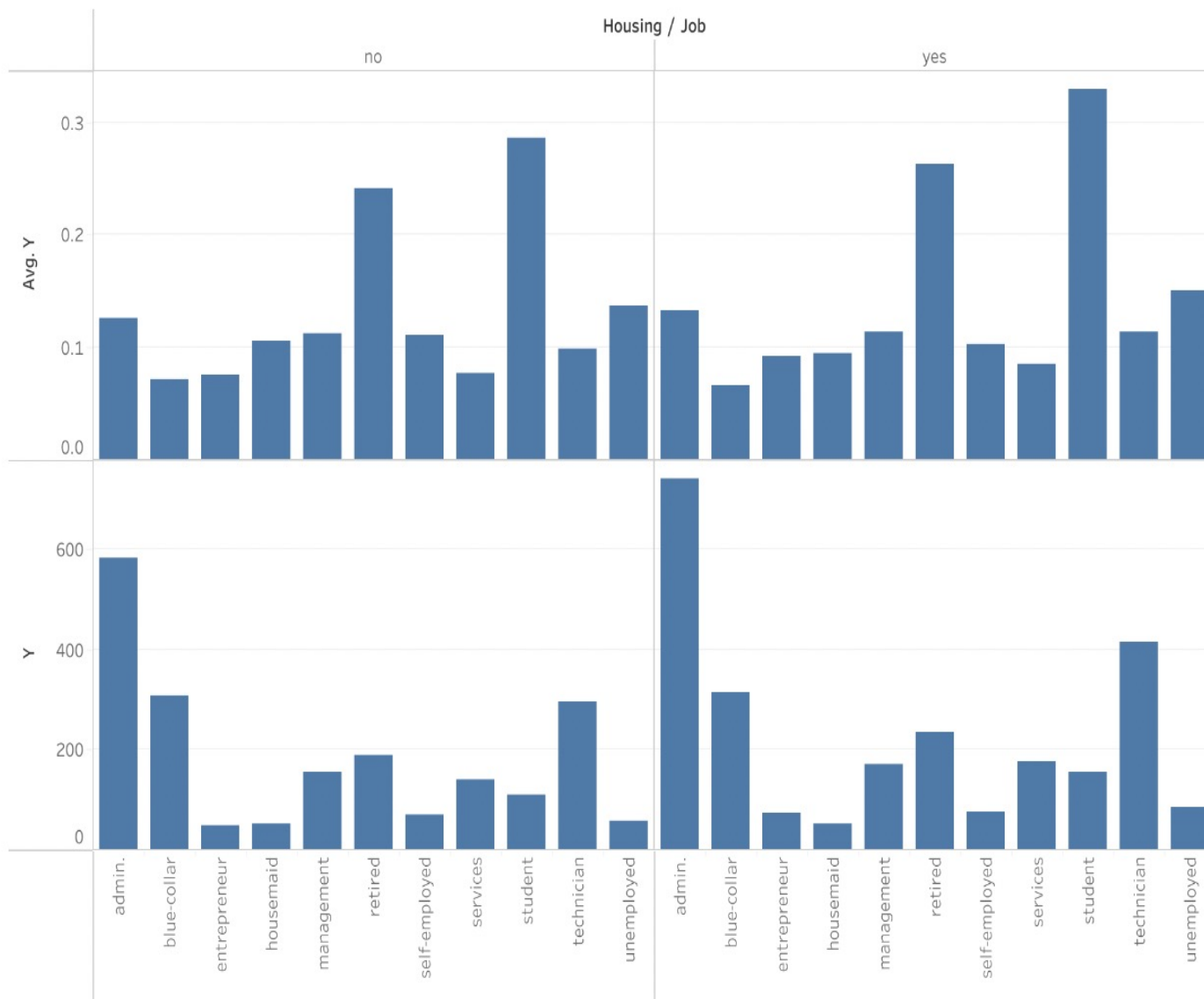
Approach

- EDA analysis & Visualization using Tableau
- Machine Learning Algorithms

EDA

Housing & Job Types

Housing & Job Types vs Subscription Decision

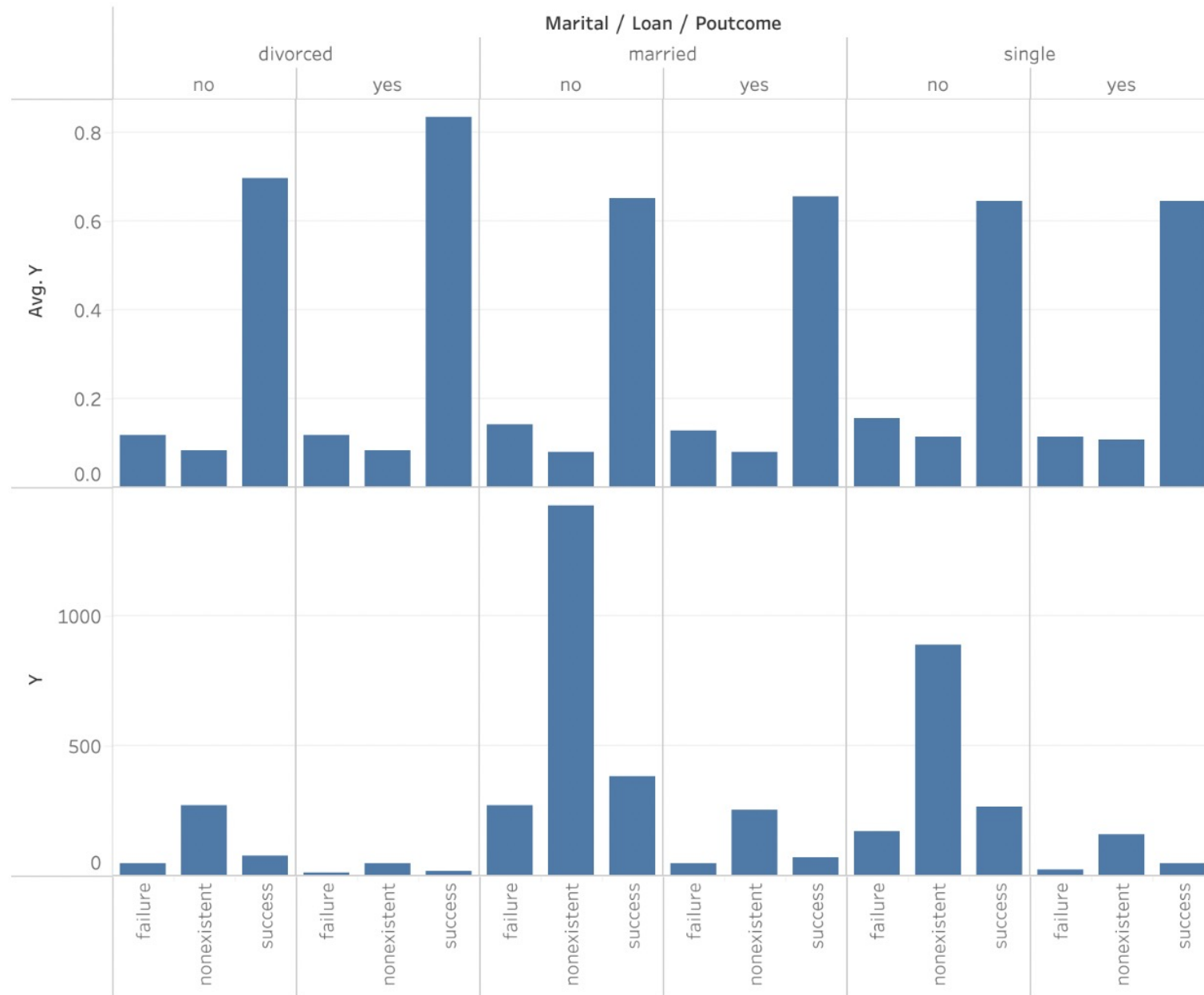


Average of Y and sum of Y for each Job broken down by Housing. The view is filtered on Housing and Job. The Housing filter keeps no and yes. The Job filter excludes unknown.

EDA

- Marital & Loan & Previous Marketing Campaign

Marital/Loan/Previous Marketing Campaign Outcome

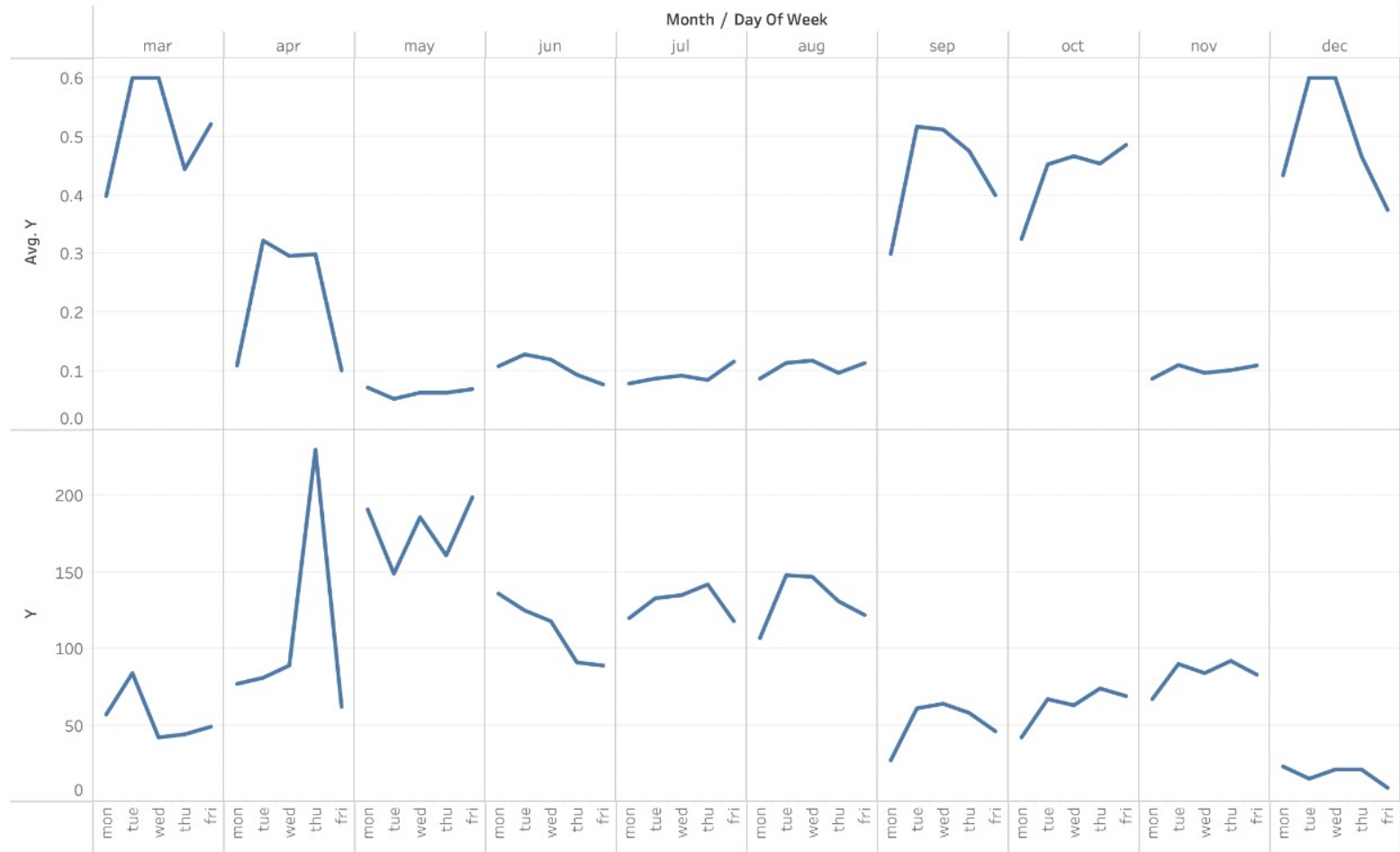


Average of Y and sum of Y for each Poutcome broken down by Marital and Loan. The data is filtered on Housing and Job. The Housing filter keeps no and yes. The Job filter excludes unknown. The view is filtered on Marital, which keeps divorced, married and single.

EDA

- Seasonality

Seasonality

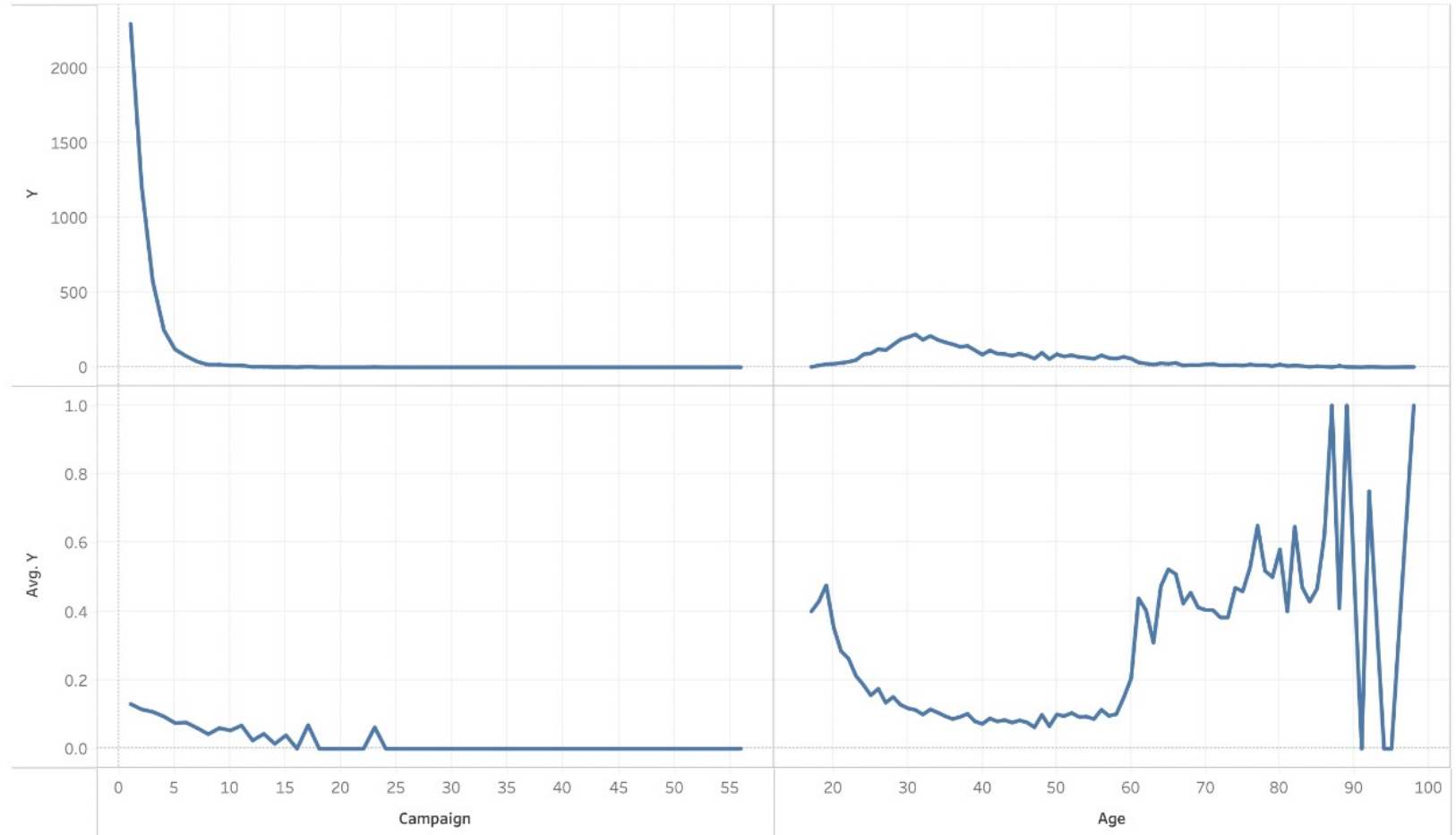


The trends of average of Y and sum of Y for Day Of Week broken down by Month.

EDA

- Number of Contact

Number of contact in this Campaign & Age

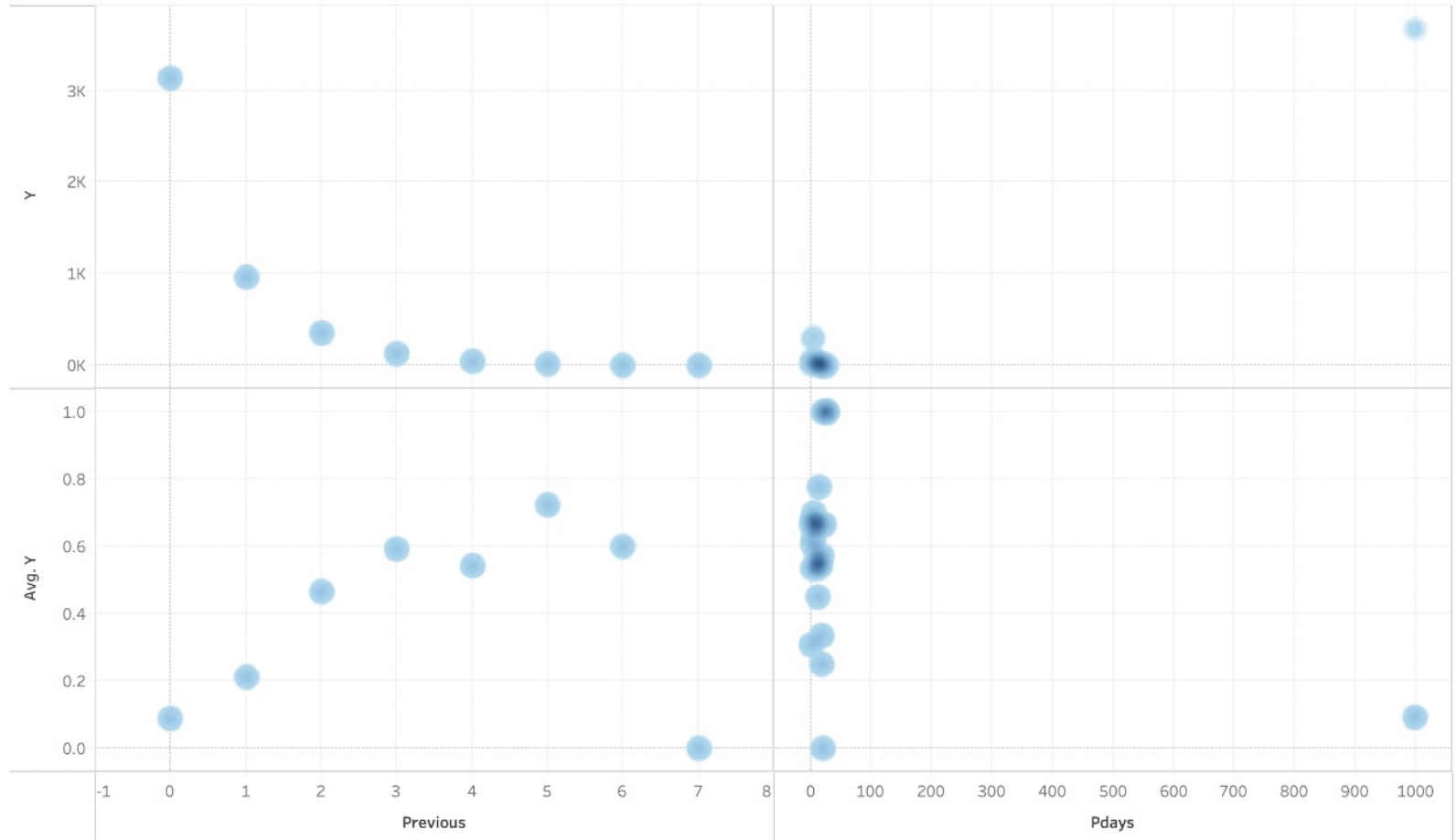


The trends of sum of Y and average of Y for Campaign and Age.

EDA

- Number of days since last contact & previous contact

Number of previous campaign and number of days since clients were contacted the last time



The plots of sum of Y and average of Y for Previous and Pdays.

EDA Summary

- As above, we analyzed the features and their potential influence to the outcome subscription decision. We found that Number of contact in the current campaign, number of previous contact, job type, education have relatively stronger influence on the outcome decision, while some other feature could help us categorize customers and help us make strategies with respect to different group of customer.¶

ML Model Suggestion

- **Linear Model (LDA, Logistic Regression)**
- **Bayesian Model (Naïve Bayes, Gaussian NB)**
- **SVM (Linear SVM, SVM with different kernels)**
- **KNN**
- **Tree Based models (XGBoost, Random Forest, Extra Trees)**
- **Neural Networks (Multi-Layer Perceptron)**

Thank You