Table of Contents

Defining the problem

Business questions to be answered using data

1. How many customers are there in total?
2. What is the distribution of gender amongst the data?
3. How many active and inactive customers are there?
4. What are the reasons for leaving? Which of these have the highest churn?
5. What are the contract terms that exist?
6. How many customers have multiple lines, internet subscription and phone service?

Gathering the data

Dataset Source:

https://www.kaggle.com/code/farazrahman/telco-customer-churn-logisticregression/input

Purpose Of Project

The purpose of this analysis is to use charts and visualization in Tableau to answer the business questions stated under the business problem definition section.

```
RangeIndex: 7043 entries, 0 to 7042
Data columns (total 33 columns):
 #   Column             Non-Null Count  Dtype  
---  ------             --------------  -----  
 0   CustomerID         7043 non-null   object 
 1   Count              7043 non-null   int64  
 2   Country            7043 non-null   object 
 3   State              7043 non-null   object 
 4   City               7043 non-null   object 
 5   Zip Code           7043 non-null   int64  
 6   Lat Long           7043 non-null   object 
 7   Latitude           7043 non-null   float64
 8   Longitude          7043 non-null   float64
 9   Gender             7043 non-null   object 
 10  Senior Citizen     7043 non-null   object 
 11  Partner            7043 non-null   object 
 12  Dependents         7043 non-null   object 
 13  Tenure Months      7043 non-null   int64  
 14  Phone Service      7043 non-null   object 
 15  Multiple Lines     7043 non-null   object 
 16  Internet Service   7043 non-null   object 
 17  Online Security    7043 non-null   object 
 18  Online Backup      7043 non-null   object 
 19  Device Protection  7043 non-null   object 
 20  Tech Support       7043 non-null   object 
 21  Streaming TV       7043 non-null   object 
 22  Streaming Movies   7043 non-null   object 
 23  Contract           7043 non-null   object 
 24  Paperless Billing  7043 non-null   object 
 25  Payment Method     7043 non-null   object 
 26  Monthly Charges    7043 non-null   float64
 27  Total Charges      7043 non-null   object 
 28  Churn Label        7043 non-null   object 
 29  Churn Value        7043 non-null   int64  
 30  Churn Score        7043 non-null   int64  
 31  CLTV               7043 non-null   int64  
 32  Churn Reason       1869 non-null   object 
dtypes: float64(3), int64(6), object(24)
memory usage: 1.8+ MB
```

Data Cleansing

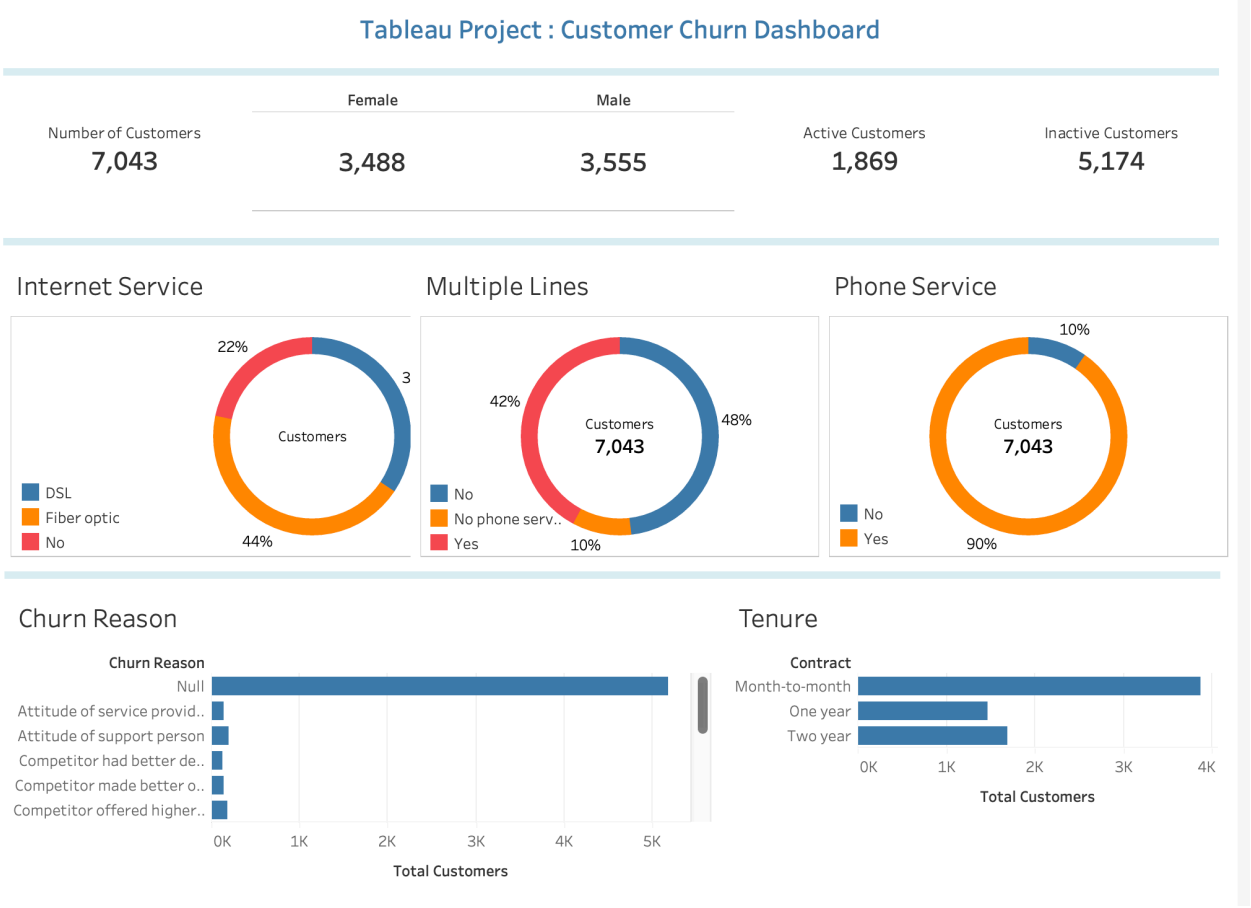
During the data cleansing stage, the data was reviewed to identify any missing values as seen in the below:

```
missing_values = customer_data.isnull().sum()
print('The missing values in each column:')
print(missing_values)
```

```
The missing values in each column:
CustomerID             0
Count                  0
Country                0
State                  0
City                   0
Zip Code               0
Lat Long               0
Latitude               0
Longitude              0
Gender                 0
Senior Citizen         0
Partner                0
Dependents             0
Tenure Months          0
Phone Service          0
Multiple Lines         0
Internet Service       0
Online Security        0
Online Backup          0
Device Protection      0
Tech Support           0
Streaming TV           0
Streaming Movies       0
Contract               0
Paperless Billing      0
Payment Method         0
Monthly Charges        0
Total Charges          0
Churn Label            0
Churn Value            0
Churn Score            0
CLTV                   0
Churn Reason        5174
dtype: int64
```

The above shows that about 70% of the churn reason values is missing from the data. The rest of the columns do not have missing information

Data Analysis & Findings



Conclusion

There are about 7043 customers in total with about 26% who are active. The rest account for inactive customers. Almost 50% of the customers are males and 49% are female.

For internet service subscription, 22% of the customers do not have an internet subscription. Only 10% have multiple lines and 90% have a phone service with a telco. About 73% of the churn reason data is unavailable, however, most customers complained that the attitude of support person was the reason they left.

Limitations and Future Work

Customer churn data is crucial for businesses because it provides insights into why customers stop using a product or service, allowing companies to identify critical pain points in the customer journey, develop targeted retention strategies, and ultimately improve customer satisfaction and loyalty, thereby minimizing revenue loss by preventing customers from leaving.

There is a lot more analysis and value from the data above that could be explored further. However, the future work will be extending this to uncover more business insights and making predictions on which customers are likely to churn in a period using some features from the data.

Link to workbook.

https://github.com/techieabbiee/CustomerChurnAnalysis