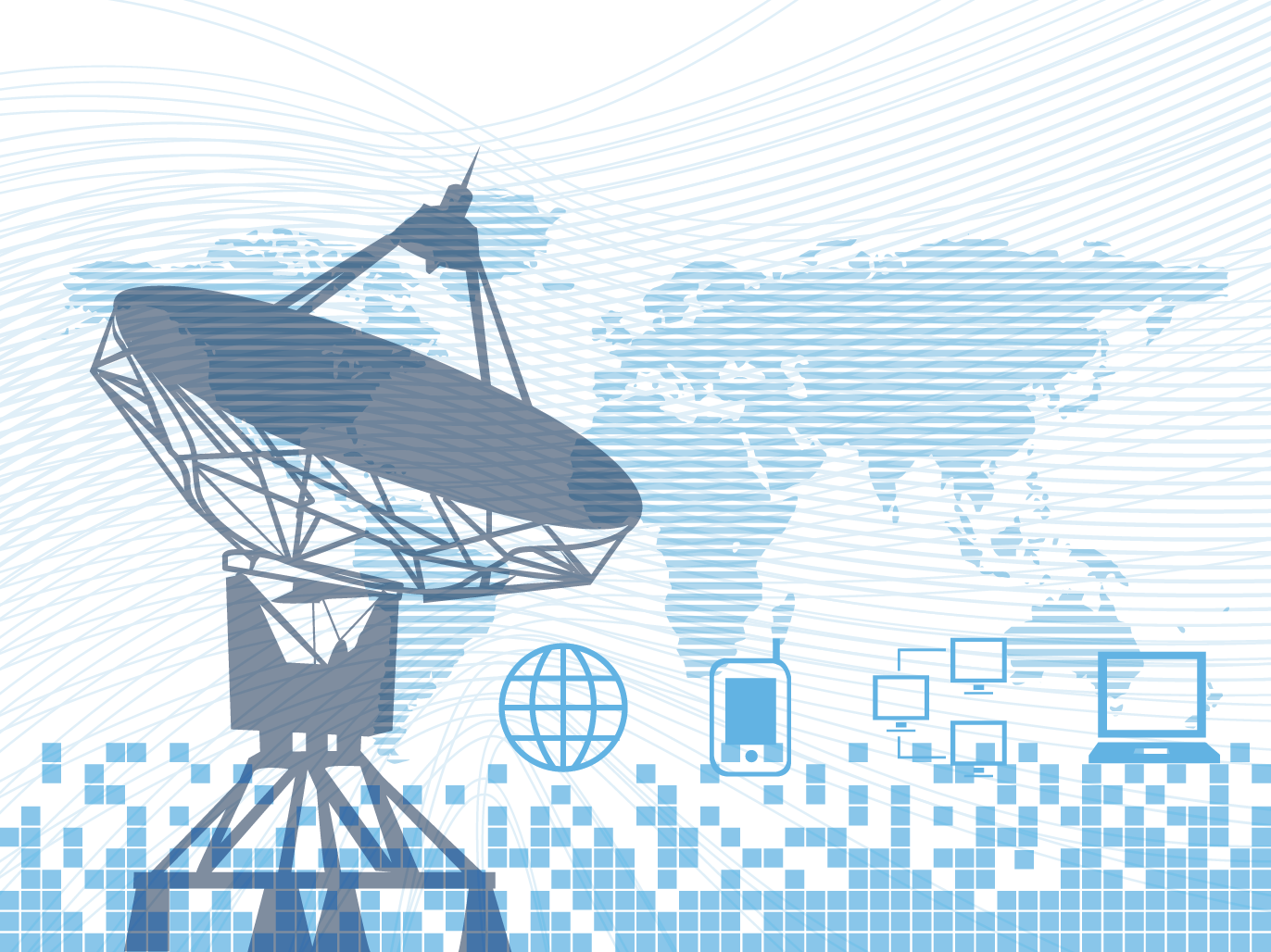
**Exploratory Data Analysis on Telecom Users Data Set**



**Introduction**

The Telecom Users data set contains 23 columns and 6050 rows of information. I imported varies libraries such as pandas, Numpy, Seaborn and Matplotlib to assist my analysis and loaded the excel file into Jupyter notebook.

**Objectives**

1. To glean any patterns or relationships from the data
2. To make recommendations to stakeholders for future areas of growth and potential savings.

**DATA CLEANING**

In summary I carried out the following:

* I identified missing values were present in two columns ‘Provider’ and ‘Age’ I decided to remove these from my dataset as they were not relevant.
* I identified the types of the columns and decided to change the ‘Total Charges’ column from an object to a float to assist with my analysis to follow.
* I removed the ‘Unnamed: 0’ column as this was not required.

**MISSING DATA**

In summary I carried out the following:

* I identified null values and removed the two columns ‘Provider’ and ‘Age’ as their fields could not be determined.

**DATA STORIES AND VISUALISATIONS**

In summary I carried out the following:

**Telecom Users - Non Senior Citizens**

**Male and Female subscribers with a Phone Service vs Monthly Charges with a histogram using Matplotlib and Seaborn**

A picture containing diagram, screenshot, plot, text

Description automatically generated

Based on the side-by-side comparison of the histograms, it appears that there are some differences and similarities in the distributions of monthly charges for male and female subscribers with a phone service.

For male subscribers, the highest count of around 650 is for monthly charges of approximately £20, while their second highest count of around 280 is for monthly charges of approximately £85.

For female subscribers, the highest count of around 600 is for monthly charges of approximately £20, while their second highest count of around 250 is for monthly charges of approximately £75.

These figures suggest that both male and female subscribers have similar distributions of monthly charges, with the majority having monthly charges of around £20 for the Phone Service.

**Male and Female subscribers with a Internet Service vs Monthly Charges with a histogram using Matplotlib and Seaborn**

A picture containing diagram, screenshot, plot, pixel

Description automatically generated

Based on the side-by-side comparison of the histograms, it appears that there are some differences and similarities in the distributions of monthly charges for male and female subscribers with a internet service.

For male subscribers, the highest count of around 200 is for monthly charges of approximately £85, while their second highest count of around 175 is for monthly charges of approximately £90.

For female subscribers, the highest count of around 190 is for monthly charges of approximately £85, while their second highest count of around 180 is for monthly charges of approximately £80.

These figures suggest that both male and female subscribers have similar distributions of monthly charges, with the majority having monthly charges of around £85 for the Internet Service.

**Male and Female subscribers with Device Protection vs Monthly Charges with a histogram using Matplotlib and Seaborn**

A picture containing diagram, screenshot, plot, pixel

Description automatically generated

Based on the side-by-side comparison of the histograms, it appears that there are some differences and similarities in the distributions of monthly charges for male and female subscribers with Device Protection.

For male subscribers, the highest count of around 130 is for monthly charges of approximately £85, while their second highest count of around 100 is for monthly charges of approximately £110.

For female subscribers, the highest count of around 110 is for monthly charges of approximately £100, while their second highest count of around 90 is for monthly charges of approximately £85.

These figures suggest that male subscribers have a slightly higher count when monthly charges are lower, while female subscribers have a higher count when monthly charges are higher for Device Protection. This could indicate that female subscribers are more willing to pay higher monthly charges for Device Protection than male subscribers.

.

**Male and Female subscribers with Streaming TV Service vs Monthly Charges with a histogram using Matplotlib and Seaborn**

A picture containing diagram, screenshot, plot, pixel

Description automatically generated

Based on the side-by-side comparison of the histograms, it appears that there are some differences and similarities in the distributions of monthly charges for male and female subscribers with Streaming TV Service.

For male subscribers, the highest count of around 120 is for monthly charges of approximately £100, while their second highest count of around 110 is for monthly charges of approximately £110.

For female subscribers, the highest count of around 105 is for monthly charges of approximately £100, while their second highest count of around 100 is for monthly charges of approximately £95.

This could indicate that Streaming TV Service is a popular service among both male and female subscribers, and that both groups are willing to pay similar monthly charges for this service. The upward trend in Monthly Charges could suggest that this is a growing area of interest for non senior citizens.

**Male and Female subscribers with Paperless Billing vs Monthly Charges with a histogram using Matplotlib and Seaborn**

A picture containing screenshot, diagram, plot

Description automatically generated

Based on the side-by-side comparison of the histograms, it appears that there are some differences and similarities in the distributions of monthly charges for male and female subscribers with Paperless Billing.

For male subscribers, the highest count of around 200 is for monthly charges of approximately £20, while their second highest count of around 175 is for monthly charges of approximately £75.

For female subscribers, the highest count of around 180 is for monthly charges of approximately £20, while their second highest count of around 175 is for monthly charges of approximately £75.

These figures suggest that both male and female subscribers have similar distributions of monthly charges, with the majority having monthly charges of around £20 for the Paperless Billing..

**Comparison of Male and Female subscribers count with a Phone Service using a Barplot in Seaborn**

A picture containing text, screenshot, colorfulness

Description automatically generated

Based on the side-by-side comparison of the Barplots, it appears that there are higher male subscribers with a phone service around 2300 to 2200 female subscribers.

**Comparison of Male and Female subscribers count with a Internet Service using a Barplot in Seaborn**

A picture containing screenshot, text, colorfulness

Description automatically generated

Based on the side-by-side comparison of the Barplots, it appears that there are higher male subscribers with a internet service around 1950 to 1850 female subscribers.

.

**Comparison of Male and Female subscribers count with Device Protection using a Barplot in Seaborn**

A picture containing text, screenshot, colorfulness, diagram

Description automatically generated

Based on the side-by-side comparison of the Barplots, it appears that there are slightly higher male subscribers with device protection around 850 to 800 female subscribers.

.

**Comparison of Male and Female subscribers count with Streaming TV using a Barplot in Seaborn**

A picture containing screenshot, text, colorfulness

Description automatically generated

Based on the side-by-side comparison of the Barplots, it appears that there are higher male subscribers with Streaming TV service around 950 to 900 female subscribers.

.

**Comparison of Male and Female subscribers count with Paperless Billing using a Barplot in Seaborn**

**A picture containing screenshot, text, colorfulness, software

Description automatically generated**

Based on the side-by-side comparison of the Barplots, it appears that there are slightly higher male subscribers with paperless billing around 1450 to 1400 female subscribers.

**Telecom Users - Senior Citizens**

**Total Count of subscribers that are Senior Citizens vs not Senior Citizens with a Barplot using Matplotlib**

A picture containing text, screenshot, diagram, rectangle

Description automatically generated

**Total Count of subscribers that are Senior Citizens vs not Senior Citizens with a Barplot using Matplotlib**

A blue and orange pie chart

Description automatically generated with medium confidence

**Male and Female Senior Citizen subscribers with a Phone Service vs Monthly Charges with a histogram using Matplotlib and Seaborn**

A picture containing diagram, screenshot, plot, line

Description automatically generated

Based on the side-by-side comparison of the histograms, it appears that there are some differences and similarities in the distributions of monthly charges for male and female senior citizen subscribers with a phone service.

For male subscribers, the highest count of around 65 is for monthly charges of approximately £85, while their second highest count of around 60 is for monthly charges of approximately £90.

For female subscribers, the highest count of around 70 is for monthly charges of approximately £85, while their second highest count of around 60 is for monthly charges of approximately £80.

These figures suggest that both male and female subscribers have similar distributions of monthly charges, with the majority having monthly charges of around £85 for the Phone Service. This could indicate that Phone Service is a popular service among both male and female senior subscribers, and that both groups are willing to pay similar monthly charges for this service.

**Male and Female Senior Citizen subscribers with a Internet Service vs Monthly Charges with a histogram using Matplotlib and Seaborn**

A picture containing diagram, screenshot, plot, colorfulness

Description automatically generated

Based on the side-by-side comparison of the histograms, it appears that there are some differences and similarities in the distributions of monthly charges for male and female senior citizen subscribers with a internet service.

For male subscribers, the highest count of around 65 is for monthly charges of approximately £100, while their second highest count of around 65 is for monthly charges of approximately £95.

For female subscribers, the highest count of around 70 is for monthly charges of approximately £85, while their second highest count of around 60 is for monthly charges of approximately £110.

These figures suggest that female subscribers have a higher count of subscribers when monthly charges are lower, while male subscribers have a higher count when monthly charges are higher for an internet service. This could indicate that female subscribers are more sensitive to price changes for an internet service than male subscribers.

**Male and Female Senior Citizen subscribers with Device Protection vs Monthly Charges with a histogram using Matplotlib and Seaborn**

A picture containing diagram, plot, screenshot, line

Description automatically generated

Based on the side-by-side comparison of the histograms, it appears that there are some differences and similarities in the distributions of monthly charges for male and female senior citizen subscribers with Device Protection.

For male subscribers, the highest count of around 45 is for monthly charges of approximately £105, while their second highest count of around 40 is for monthly charges of approximately £100.

For female subscribers, the highest count of around 50 is for monthly charges of approximately £105, while their second highest count of around 40 is for monthly charges of approximately £95.

These figures suggest that male subscribers have a slightly higher count when monthly charges are lower, while female subscribers have a higher count when monthly charges are higher for Device Protection. This could indicate that female subscribers are more willing to pay higher monthly charges for Device Protection than male subscribers.

**Male and Female Senior Citizen subscribers with Streaming TV Service vs Monthly Charges with a histogram using Matplotlib and Seaborn**

A picture containing diagram, plot, screenshot

Description automatically generated

Based on the side-by-side comparison of the histograms, it appears that there are some differences and similarities in the distributions of monthly charges for male and female senior citizen subscribers with Streaming TV Service.

For male subscribers, the highest count of around 55 is for monthly charges of approximately £105, while their second highest count of around 40 is for monthly charges of approximately £100.

For female subscribers, the highest count of around 50 is for monthly charges of approximately £100, while their second highest count of around 40 is for monthly charges of approximately £105.

Looking at the overall distribution, the figures are quite low and this could indicate that Streaming TV Service isn't a popular service among both male and female senior subscribers. It is only when monthly charges are £100 and above that the count of subscribers increases for both male and female. With the right marketing and investment, this could would be a growth area for senior citizens as new Streaming TV customers.

**Male and Female Senior Citizen subscribers with Paperless Billing vs Monthly Charges with a histogram using Matplotlib and Seaborn**

A picture containing diagram, screenshot, plot, text

Description automatically generated

Based on the side-by-side comparison of the histograms, it appears that there are some differences and similarities in the distributions of monthly charges for male and female senior citizen subscribers with Paperless Billing.

For male subscribers, the highest count of around 65 is for monthly charges of approximately £95, while their second highest count of around 55 is for monthly charges of approximately £100.

For female subscribers, the highest count of around 65 is for monthly charges of approximately £75, while their second highest count of around 55 is for monthly charges of approximately £95.

These figures suggest that both male and female subscribers have similar distributions of monthly charges, but male seniors have a higher count with a higher Paperless Billing than female subscribers.

**Telecom Users – Total Charges**

**Total Charges across the whole dataset with a pie chart using Matplotlib**

**A blue circle with white text

Description automatically generated with medium confidence**

**Distribution of Total Charges across the whole dataset with a Dist plot using Matplotlib**

**A picture containing text, diagram, plot, screenshot

Description automatically generated**

**Heatmap of Total Charges across the whole dataset by gender using Seaborn**

**A picture containing text, screenshot, font

Description automatically generated**

**Pie Chart percentage of Total Charges vs Online Backup using Matplotlib**

**A picture containing text, screenshot, diagram, font

Description automatically generated**

**Pie Chart percentage of Total Charges vs Phone Service using Matplotlib**

A picture containing screenshot, diagram, circle, font

Description automatically generated

**Pie Chart percentage of Total Charges vs Streaming Movies using Matplotlib**

A picture containing text, screenshot, font, diagram

Description automatically generated

**Pie Chart percentage of Total Charges vs Streaming TV using Matplotlib**

A picture containing text, screenshot, diagram, font

Description automatically generated

**Pie Chart percentage of Total Charges vs Multiple Lines using Matplotlib**

A picture containing text, screenshot, diagram, font

Description automatically generated

**Pie Chart percentage of Total Charges vs Churn using Matplotlib**

A red and blue pie chart

Description automatically generated with medium confidence

**Pie Chart percentage of Total Charges vs Senior Citizen or Non Senior Citizen Subscriber using Matplotlib**

A red and blue pie chart

Description automatically generated with medium confidence

**Total Charges vs Device Protection vs Contract**

**A picture containing text, screenshot, font, number

Description automatically generated**

**Bar Chart percentage of Total Charges vs Contract using Matplotlib**

A picture containing text, screenshot, diagram, plot

Description automatically generated

**Pair Plot of Device Protection vs tenure vs Total Charges vs Contract using Seaborn**

A screenshot of a graph

Description automatically generated with low confidence

My findings were as follows:

My conclusions are as follows:

**THIS REPORT WAS WRITTEN BY: LOUISE MARIE RANDALL**

