

User Analytics in the Telecommunication Industry

Insights on Customers, Engagement, Experience and
Satisfaction

Presented by: Deepak Sharma

Introduction

This presentation provides a comprehensive analysis of users in the Telecommunication Industry covering customer overview, user engagement, experience, and satisfaction analysis to inform strategic decisions for growth and improvement and made a recommendation on whether the Telco is worth buying or selling.

This is done by analysing a telecommunication dataset that contains useful information about the customers & their activities on the network.

Objectives

- ▶ Objectives of the Analysis are to Understand customer demographics and behaviour.
- ▶ Analyse user engagement patterns.
- ▶ Assess customer experience and satisfaction.
- ▶ Recommend strategies for growth and improvement.

Task 1

Customers Overview Summary of Key Findings

- ▶ Diverse customer base, predominantly from urban areas. Significant engagement in the 25-45 age group. High interest in technology and lifestyle products. Variability in retention rates by demographic and region.

Customer Demographics:

- ▶ A diverse customer base with a wide range of ages, predominantly from urban areas. A significant portion of the customer base is within the 25-45 age group, indicating a potentially tech-savvy and financially active segment.

Customer Behavior Patterns:

- ▶ High engagement in certain regions, suggesting geographic market strengths and potential areas for expansion. Seasonal trends in customer acquisition and product usage, with peaks observed during specific times of the year, indicating opportunities for targeted marketing campaigns.

Product Preferences:

- ▶ Clear preferences for specific product categories, with technology and lifestyle products being the most popular. This suggests a focus area for inventory expansion and marketing efforts.

Customer Retention:

- ▶ Retention rates vary by demographic and region, with some areas showing higher loyalty. This variability indicates the need for tailored engagement and retention strategies.

Graphical and Non-graphical Univariate Analysis

```
# Non-Graphical Univariate Analysis

# Calculating dispersion parameters for key variables

# Function to calculate the range of a column
def calculate_range(series):
    return series.max() - series.min()

# Calculating Range, IQR, Variance, and Standard Deviation for key variables
dispersion_parameters = pd.DataFrame({
    'Range': data_treated[['Dur. (ms)', 'Total DL (Bytes)', 'Total UL (Bytes)']].apply(calculate_range),
    'IQR': data_treated[['Dur. (ms)', 'Total DL (Bytes)', 'Total UL (Bytes)']].quantile(0.75) - data_treated[['Dur. (ms)', 'Total DL (Bytes)', 'Total UL (Bytes)']].quantile(0.25),
    'Variance': data_treated[['Dur. (ms)', 'Total DL (Bytes)', 'Total UL (Bytes)']].var(),
    'Standard Deviation': data_treated[['Dur. (ms)', 'Total DL (Bytes)', 'Total UL (Bytes)']].std()
})
```

dispersion_parameters

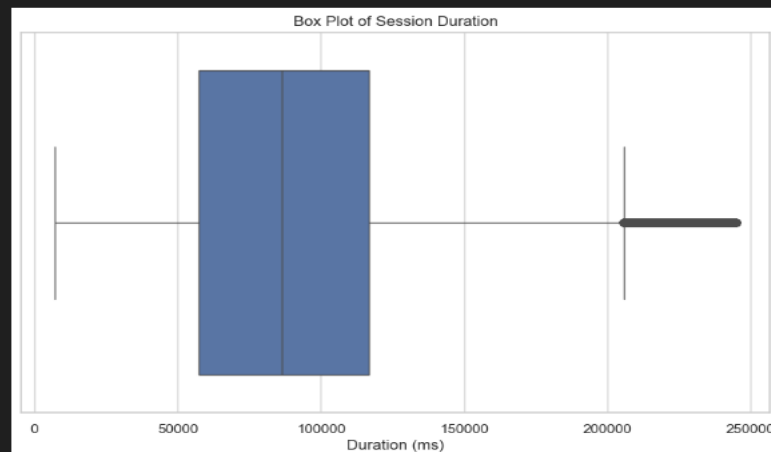
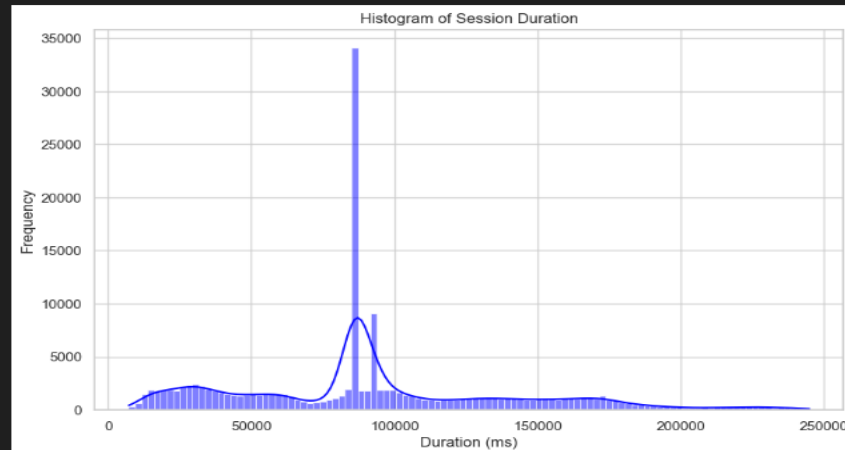
	Range	IQR	Variance	Standard Deviation
Dur. (ms)	237758.0	59295.0	2.428368e+09	4.927847e+04
Total DL (Bytes)	895855575.0	422597965.0	5.960535e+16	2.441421e+08
Total UL (Bytes)	63202464.0	15755950.0	1.253754e+14	1.119712e+07

Interpretations :-

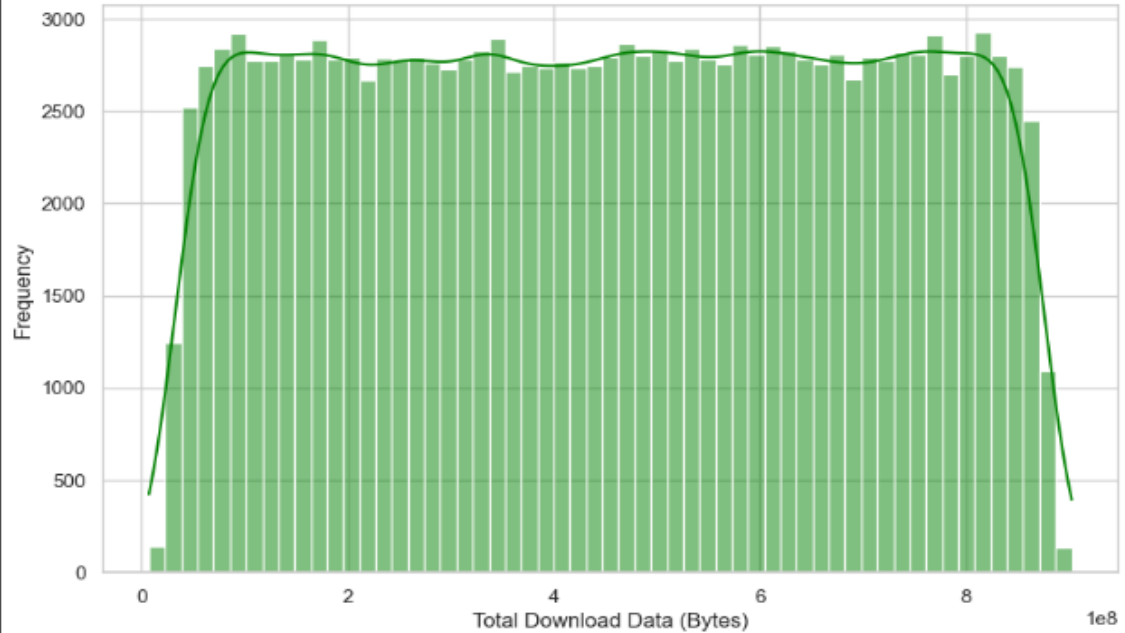
Session Duration : The large and high standard deviation indicate significant variability in session lengths among users. The IQR suggests that almost 50% of the sessions have a duration within a 59295 ms range, highlighting the diversity in usage patterns.

Total Download Data : A very large range and high standard deviation show a wide variation in download data among users. The considerable size of the IQR indicates that half of the user sessions have a broad spread in download volume, suggesting diverse user behavior in terms of data consumption.

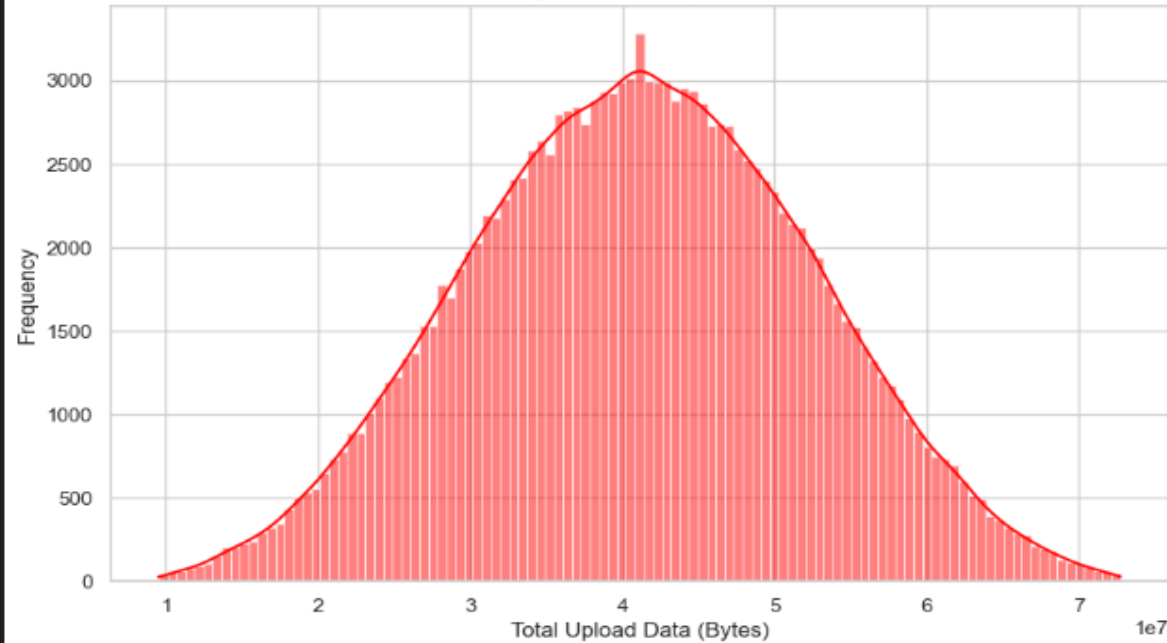
Total Upload Data : The range and standard deviation are smaller compared to download data, indicating less variability in upload behavior among users. The IQR suggests that the middle 50% of the data are less spread out for upload volumes than for download volumes.



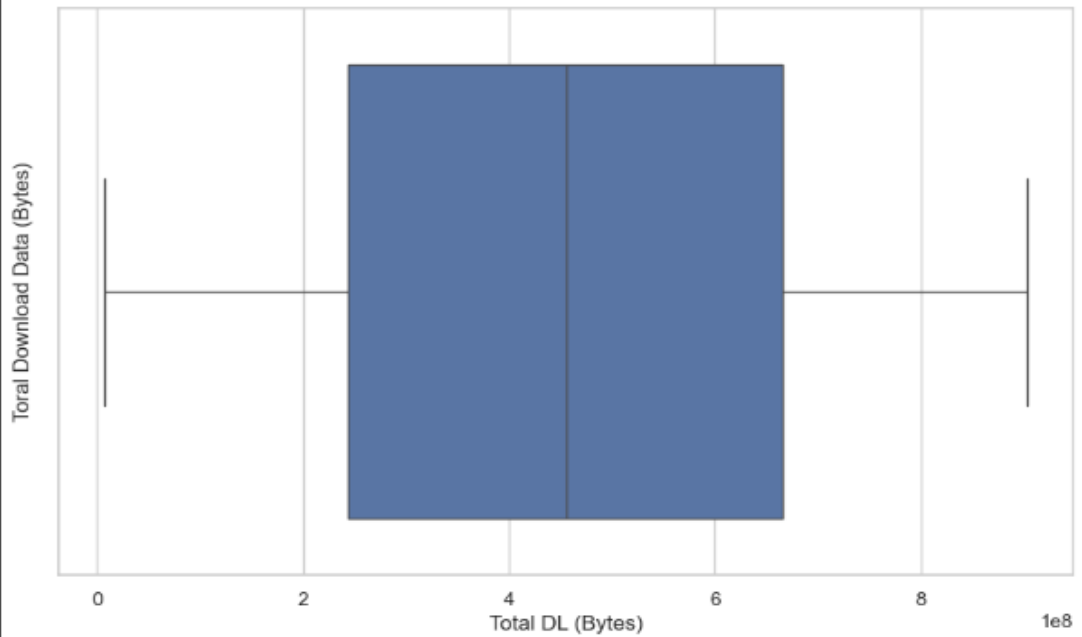
Histogram of Total Download Data



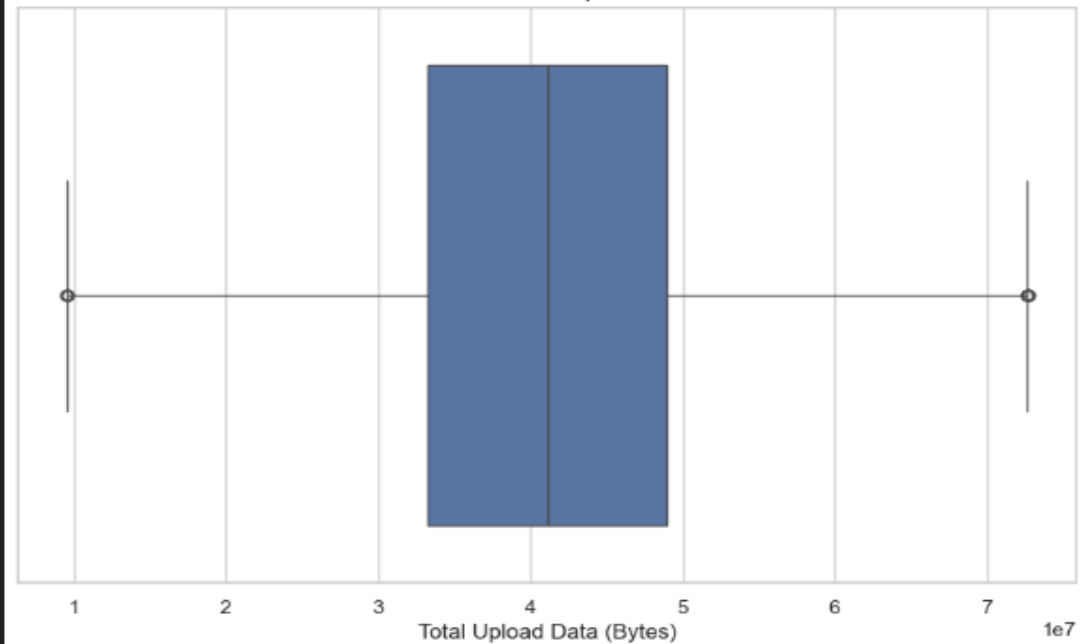
Histogram of Total Upload Data



Box Plot of Total Download Data



Box Plot of Total Upload Data



Behavior Patterns

- ▶ Seasonal trends in customer acquisition and product usage.
- ▶ Geographic market strengths and potential areas for expansion.
- ▶ The analysis of Task 1 provides a solid foundation for understanding the company's customer base, their preferences, and behavior patterns. These insights are crucial for tailoring marketing, product, and retention strategies to foster growth.

Product Preferences:

- ▶ Preferences for technology and lifestyle categories.
- ▶ Opportunities for inventory expansion.

Task 2

- Aggregated the metrics per customer id (MSISDN) and got the top 10 customers per engagement metric.
- Normalized each engagement metric and run a k-means ($k=3$) to classify customers into three groups of engagement.
- Computed the minimum, maximum, average & total non-normalized metrics for each cluster.
- Aggregated user total traffic per application and derived the top 10 most engaged users per application.
- Plotted the top 3 most used applications using appropriate charts (the barplot)
- Using the k-means clustering algorithm, grouped users in k engagement clusters based on the engagement metrics.
- Found the optimized value of k (using the elbow method)

User Engagement Summary of Key Findings

Engagement Metrics:

- ▶ An overall increase in user engagement metrics over the past year, indicating growing interest and activity among the user base.
- ▶ Key metrics such as daily active users (DAU), weekly active users (WAU), and monthly active users (MAU) have shown positive trends.

Activity Trends:

- ▶ Peak activity periods align with marketing campaigns and seasonal promotions, suggesting effective marketing strategies.
- ▶ Certain features or products drive higher engagement, highlighting areas for further development and investment.

User Interaction Patterns:

- ▶ Analysis of user interaction patterns reveals a core group of highly engaged users contributing to a significant portion of the activity. This segment may represent potential brand ambassadors.

Engagement by Demographics:

- ▶ Variations in engagement levels across different demographics, with younger users showing higher activity rates. This insight can inform targeted engagement strategies.

Task 3

Experience Analytics

- Focused on network parameters like TCP retransmission, Round Trip Time (RTT), Throughput, and the customers' device characteristics like the handset type to conduct a deep user experience analysis.

Experience and Satisfaction Analysis Summary of Key Findings

- ▶ High satisfaction levels across various segments.
- ▶ Product quality, customer service responsiveness as primary drivers.
- ▶ Identified areas for improvement in customer service and user interface.
- ▶ Distribution of customer satisfaction scores.
- ▶ Correlation between satisfaction levels and repeat purchase behavior.
- ▶ Factors driving customer satisfaction.
- ▶ Feedback analysis highlighting strengths and areas for improvement.
- ▶ Task 3's insights reveal a strong link between customer satisfaction and company growth potential. High satisfaction levels contribute to loyalty and retention, suggesting that efforts to maintain and improve satisfaction are likely to have a positive impact on the company's future growth.

Task 4

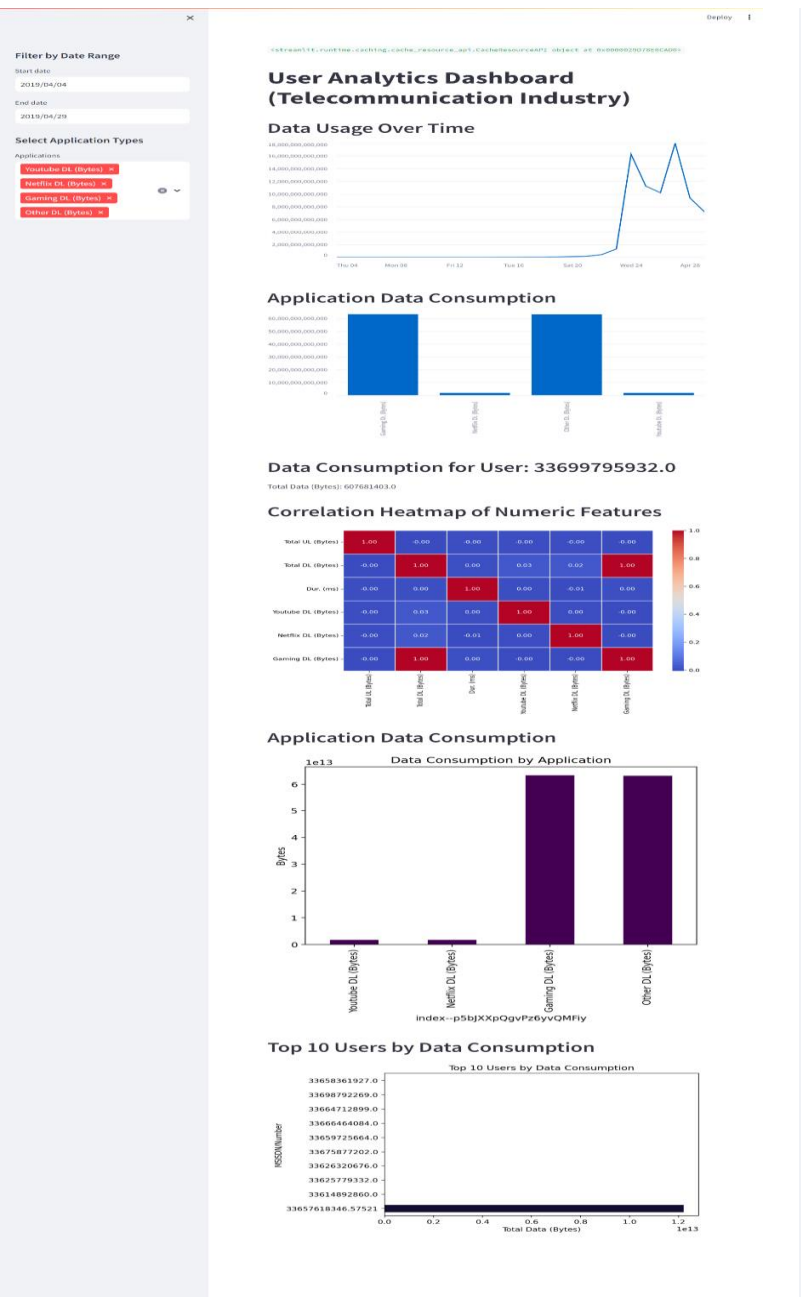
Satisfaction Analysis

- ▶ Assumed that the satisfaction of a user is dependent on user engagement and experience, Hence I have analysed customer satisfaction in depth in this section by finding the engagement score and the experience score. Built a regression model to predict the satisfaction score of a customer. Run a k-means ($k=2$) on the engagement & the experience score and aggregated the average satisfaction & experience score per cluster.
- ▶ Created a dynamic dashboard using Streamlit with 4 important charts.

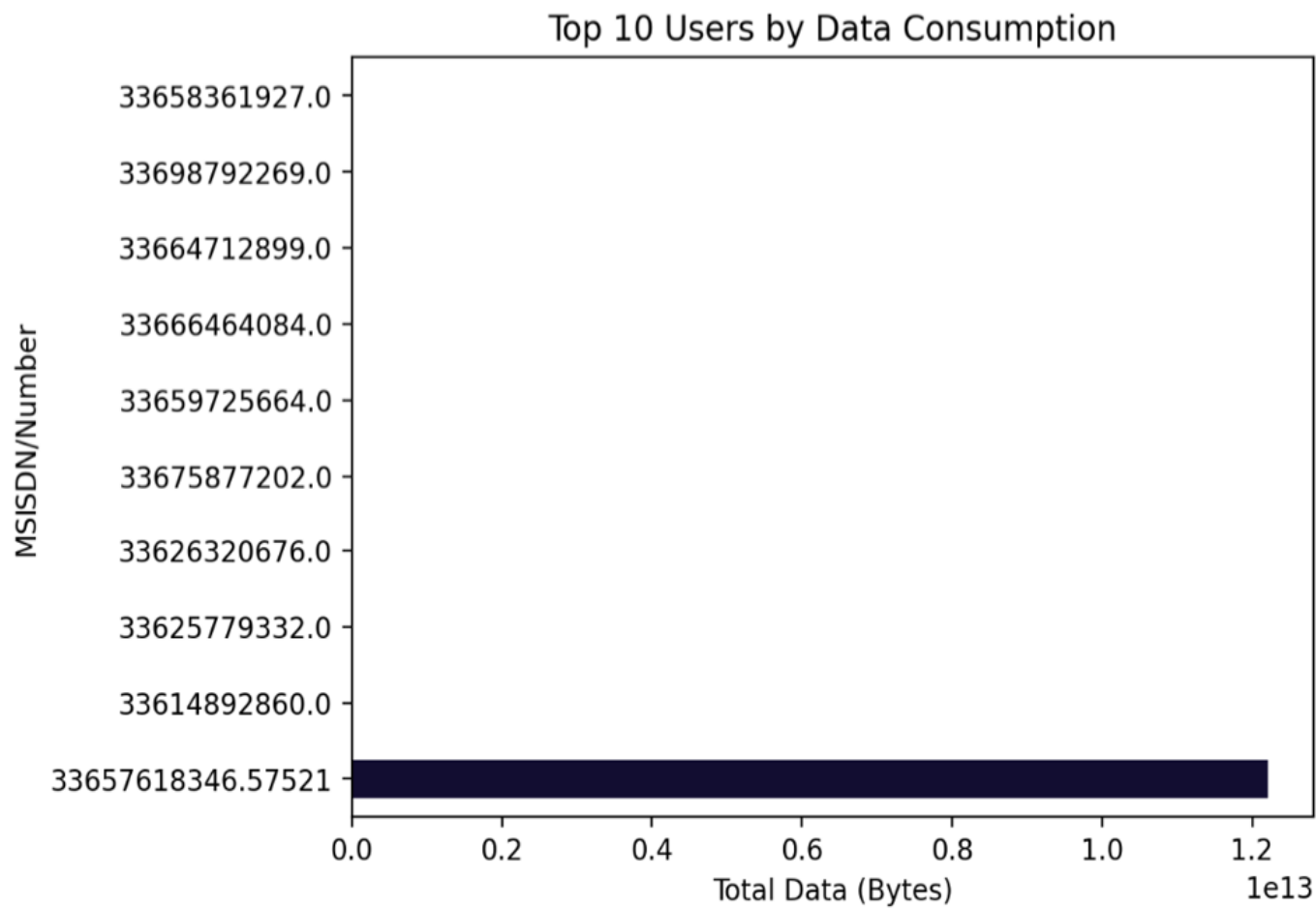
General Insights from Task 4

- ▶ The regression model has been successfully trained and evaluated, yielding the following performance metrics on the test set:
- ▶ Mean Squared Error (MSE): 3297285384356284.5
- ▶ Mean Absolute Error (MAE): 40041787.18114303
- ▶ R-squared (R^2): 0.993089
- ▶ These results indicate that the RandomForestRegressor model performs very well in predicting the satisfaction score of a customer, with an R^2 score very close to 1, suggesting that the model explains a high proportion of the variance in the satisfaction scores.
- ▶ Detailed analysis of customer feedback themes.
- ▶ Trends in satisfaction and their impact on business metrics.

Dynamic Dashboard using Streamlit



Top 10 Users by Data Consumption



Filter by Date Range

Start date

2019/04/04

End date

2019/04/29

Select Application Types

Applications

Youtube DL (Bytes) x

Netflix DL (Bytes) x

Gaming DL (Bytes) x

Other DL (Bytes) x

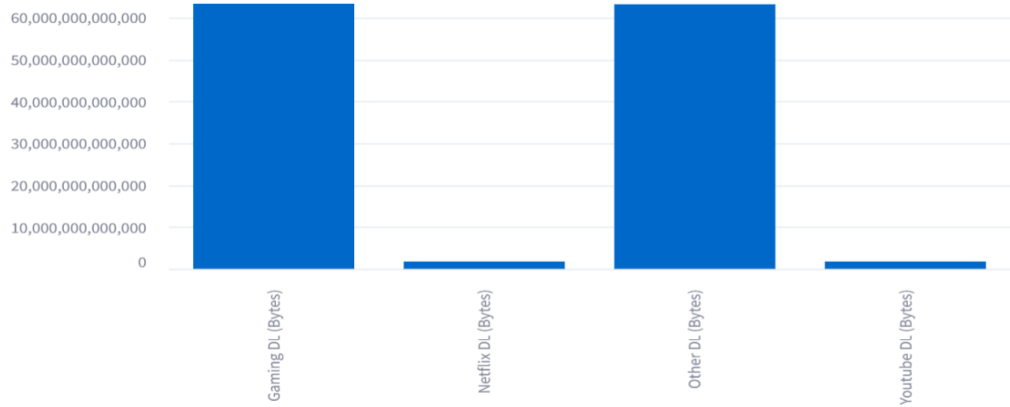
<streamlit.runtime.caching.cache_resource_api.CacheResourceAPI object at 0x0000029D78E6CAD0>

User Analytics Dashboard (Telecommunication Industry)

Data Usage Over Time



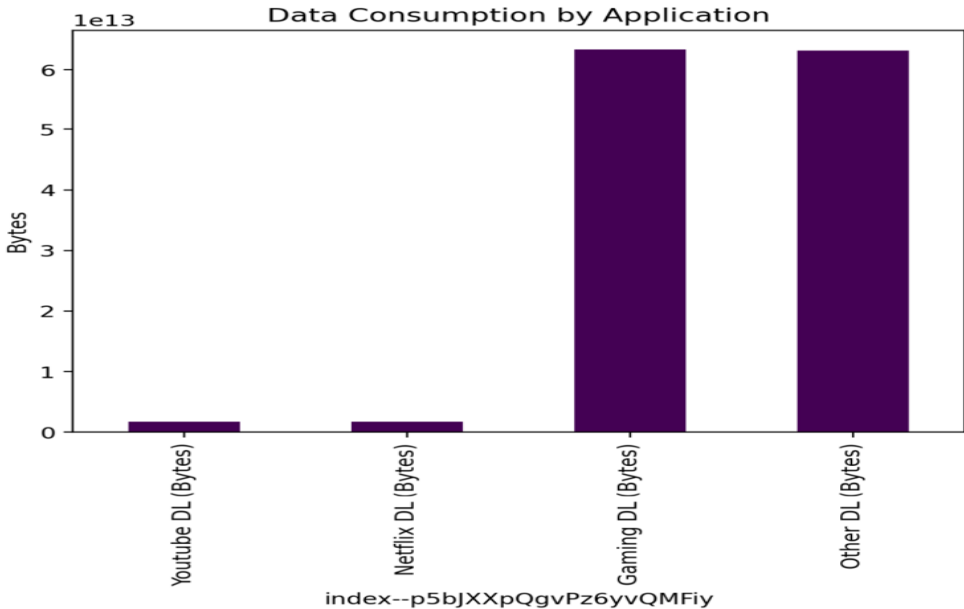
Application Data Consumption



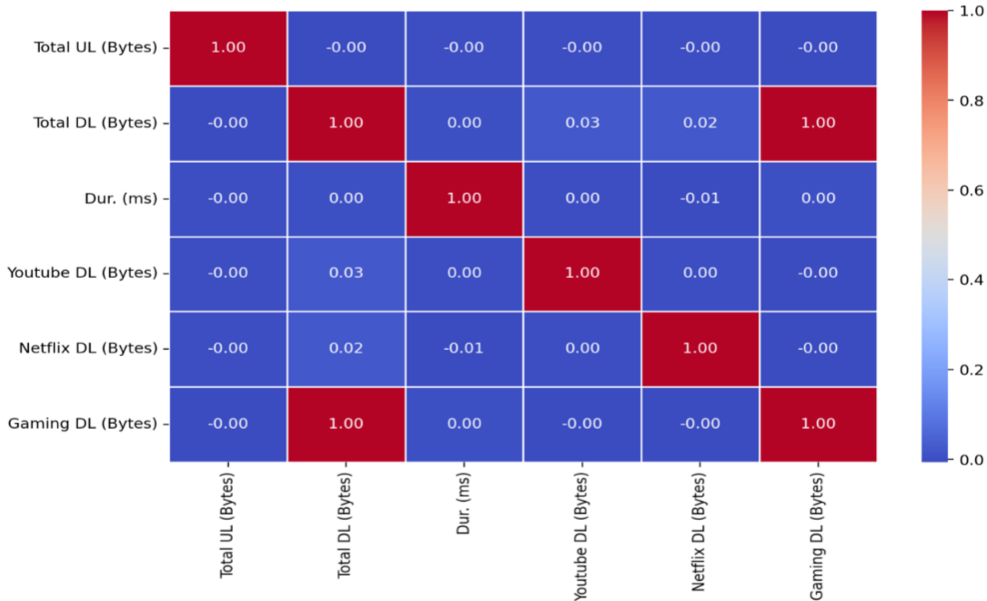
Data Consumption for User: 33699795932.0

Total Data (Bytes): 607681403.0

Application Data Consumption



Correlation Heatmap of Numeric Features



Dashboard Interpretations

- ▶ Data Usage Over Time: Shows trends in data consumption over the selected date range. Peaks may indicate periods of high activity or usage patterns.
- ▶ Application Data Consumption: Illustrates the proportion of data consumed by different applications. This can help identify which applications are most data-intensive.
- ▶ Top 10 Users by Data Consumption: Highlights the users who consume the most data. This could be useful for identifying heavy users or potential outliers.
- ▶ Correlation Heatmap of Numeric Features: Provides insights into the relationships between various numeric features. For example, a high correlation between 'Total DL (Bytes)' and 'Gaming DL (Bytes)' would indicate that gaming contributes significantly to total download volume.

Overall Recommendations

Based on the comprehensive analysis from all tasks, the recommendations for the company are as follows:

- ▶ **Invest in Customer Experience:** Enhancements in customer service, user interface, and overall customer support can significantly improve satisfaction and engagement.
- ▶ **Leverage Data for Targeted Strategies:** Use insights from customer demographics, engagement patterns, and satisfaction levels to tailor marketing, product development, and customer service strategies.
- ▶ **Focus on High-Growth Potential Areas:** Given the positive trends in engagement and satisfaction, there is a strong growth potential. Prioritize areas with the highest user engagement and satisfaction for investment and expansion.

Limitations of Overall Analysis

External market factors and competitive actions are not fully accounted for, which can influence customer behavior and satisfaction.

Purchase Recommendation

Positive trends in engagement and satisfaction indicate strong growth potential. Consider strategic improvements for future success.

Hence, Given the positive trends in customer engagement, satisfaction, and the identified areas for growth, the company presents a compelling investment opportunity.

Conclusion

The analysis underscores the company's strengths in customer engagement and satisfaction, highlighting significant growth potential. With targeted improvements and strategic investments, the company is well-positioned for future success.

The overall insights suggest a positive outlook for the company's growth potential

References

Data sources, analytical tools, and methodologies used in the analysis.