# Call Centre Data Analysis Summary

1. My Name and Contact Information

- Name: Akash Rawat

- Contact Information: [akashrawat851.ar@gmail.com](mailto:akashrawat851.ar@gmail.com)

[Mca22.akash.rawat@sunstone.edu.in](mailto:Mca22.akash.rawat@sunstone.edu.in)

## Task 1: Data Exploration

Dataset Overview

- **Initial Rows Displayed:**

- Displayed the first 10 rows of the dataset.

- Identified key columns including `\_id`, `user`, `chatStatus`, `guru`, `guruName`, `consultationType`, `website`, `refundStatus`, and more.

**Dataset Shape:**

- Rows: 28,027

- Columns: 45

**Missing Data:**

- Identified columns with missing data such as `chatStatus`, `chatSeconds`, `timeDuration`, `callChannel`, `amount`, etc.

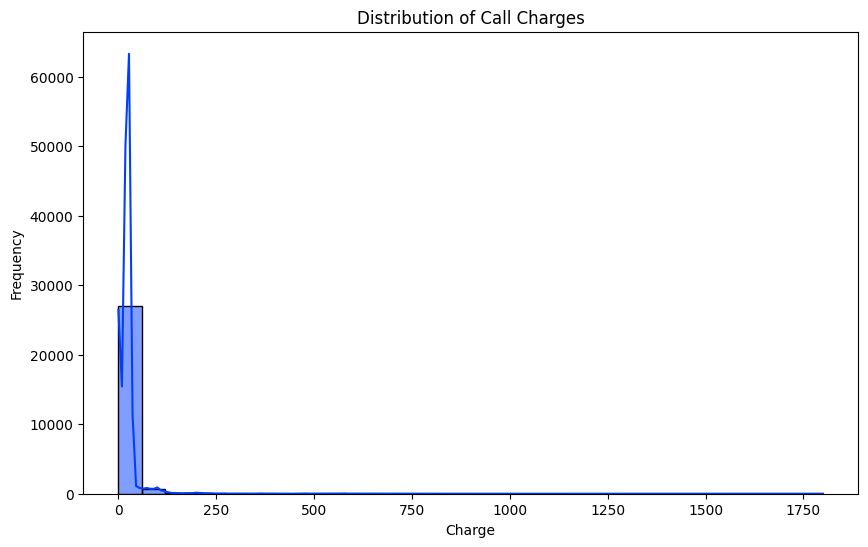
- Filled missing values for key columns with their mean values.

**Summary Statistics for Key Numeric Columns:**

- Calculated summary statistics for `chatSeconds`, `timeDuration`, `amount`, `astrologerOnCallDuration`, `userOnCallDuration`.

**Visualization:**

- Created a histogram to visualize the distribution of call charges.



## Task 2: Call Centre Performance Metrics

**Average Talk Time for Different Call Activities:**

- Calculated the average talk time for `Call`, `Chat`, `Complementary`, and `public\_live\_Call`.

**Most Common Source of Calls:**

- Identified `gurucool` as the most common source of calls.

**Total Earnings and Spending:**

- Total Earnings (Users): 3,329,182.085 seconds

- Total Spending (Users): 665,441.986 units

- Total Earnings (Masters): 3,609,649.765 seconds

- Total Spending (Masters): 99,146.571 units

**Correlation Between Talk Time and Charge:**

- Correlation Coefficient: 0.4696

## Task 3: Call Handling Analysis

**Average Connect Delay:**

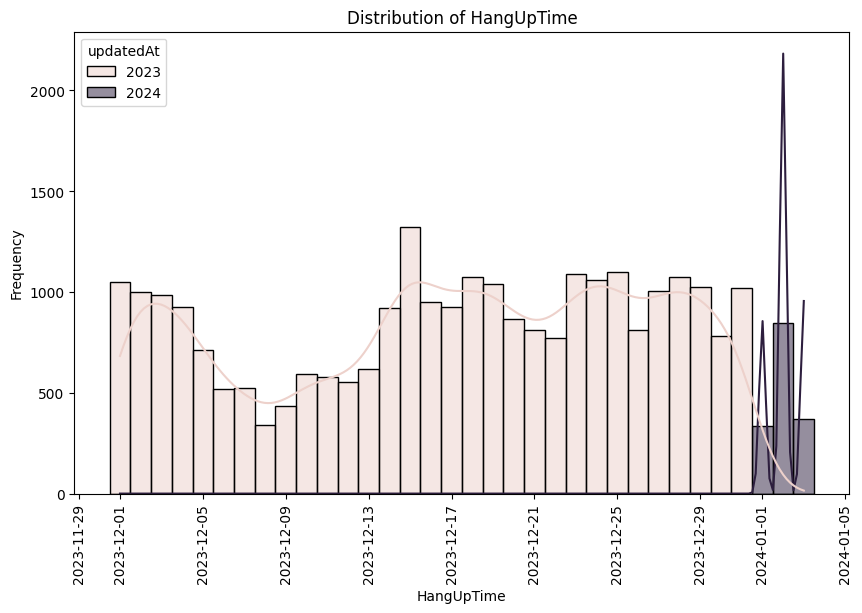
- Calculated the average connect delay as 3,404.697 seconds.

**Most Common Reason for Call Disconnection:**

- The most common disconnection reasons were `completed`, `no-answer`, `busy`, `failed`, and `incomplete`.

**Hang Up Time Patterns:**

- Created a histogram to visualize the distribution of hang-up times.

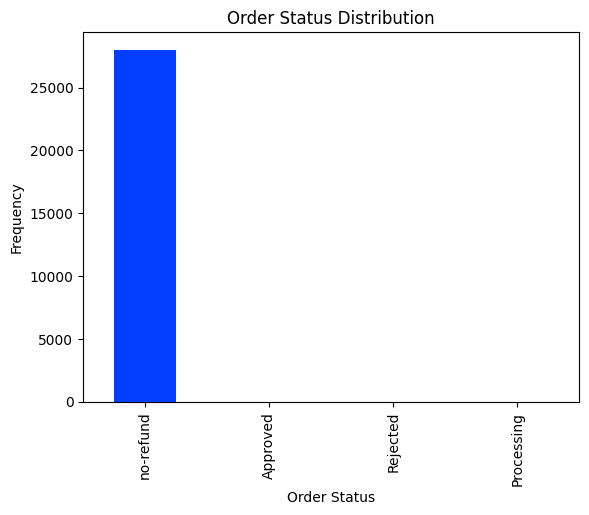


## Task 4: Order and Refund Analysis

**Order Status Distribution for Calls:**

- Most common refund status: `no-refund`

- Visualization: Created a bar chart showing the distribution of order status.



**Total Refund Amount:**

- Total Refund Amount: 665,441.986 units

## Task 5: Insights and Recommendations

**Insights:**

- The correlation between talk time and charge suggests that longer calls lead to higher charges.

- Users' spending is significantly higher than their earnings, indicating a need to revise pricing strategies.

**Recommendations:**

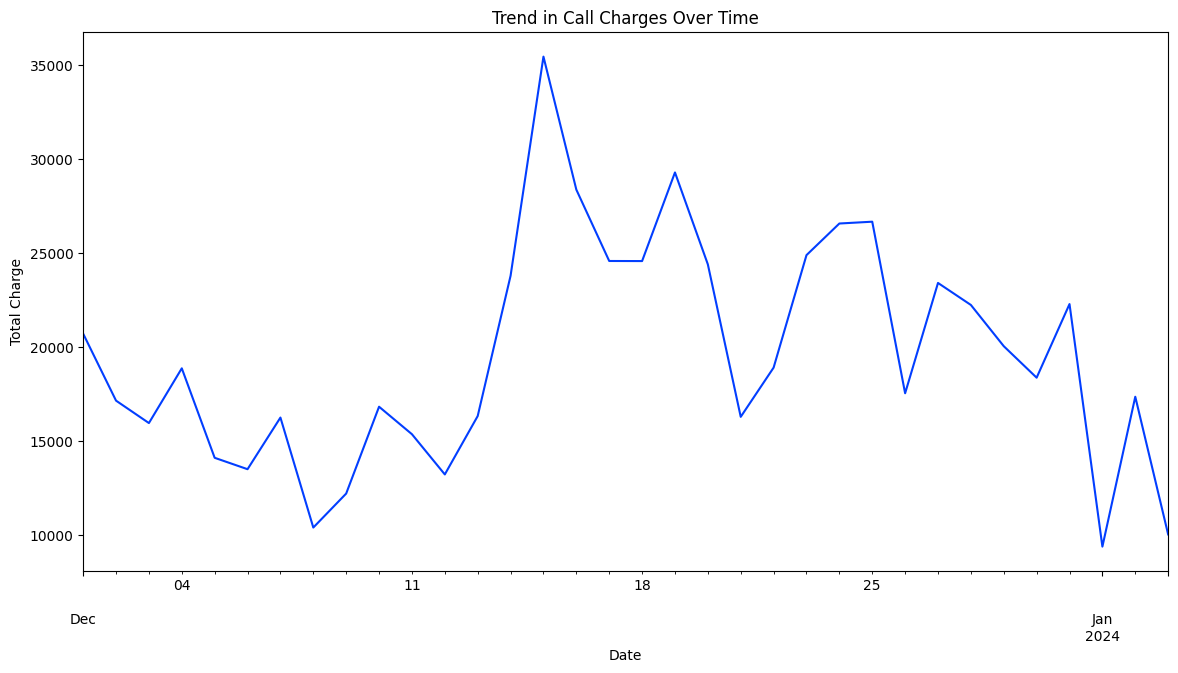
- Implement strategies to reduce average connect delay, such as improving call routing or agent training.

- Monitor common disconnect reasons and address any technical or process issues contributing to call drops.

## Task 6: Data Visualization

**Trend in Call Charges Over Time:**

- Created a line chart to show the trend in call charges over time.

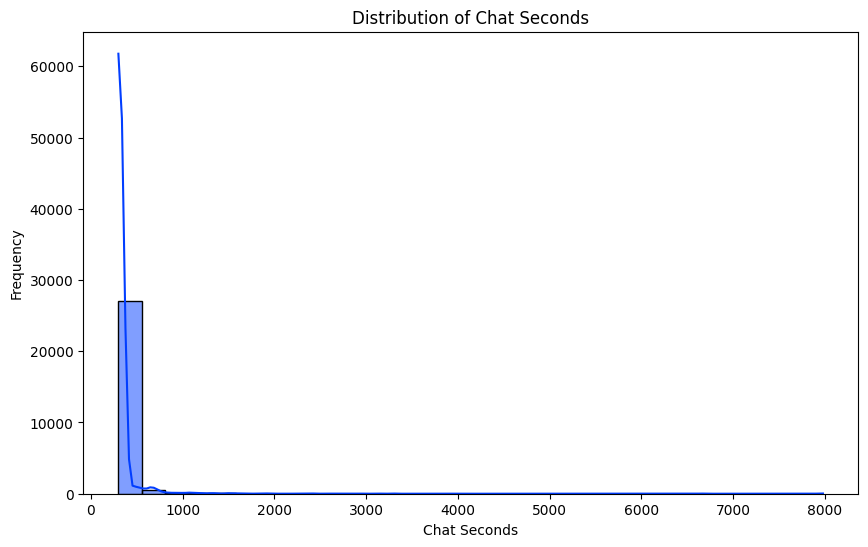


**Relationship Between Talk Time and User Spend:**

- Created a scatter plot to analyze the relationship between talk time and user spend.

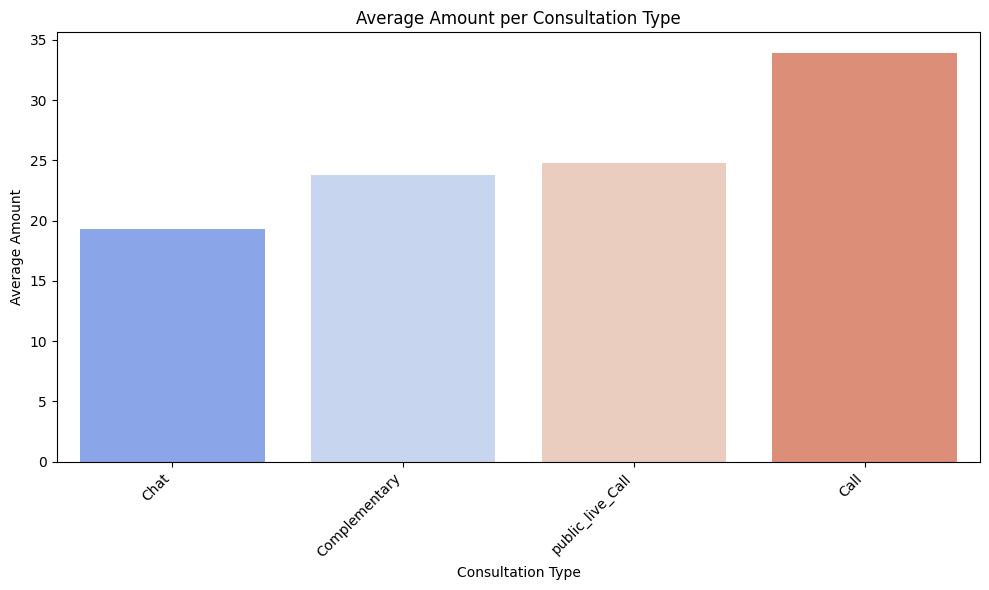
**Distribution of Chat Seconds:**

- Created a histogram to visualize the distribution of chat seconds.



**Average Amount per Consultation Type:**

- Created a bar plot to show the average amount per consultation type.

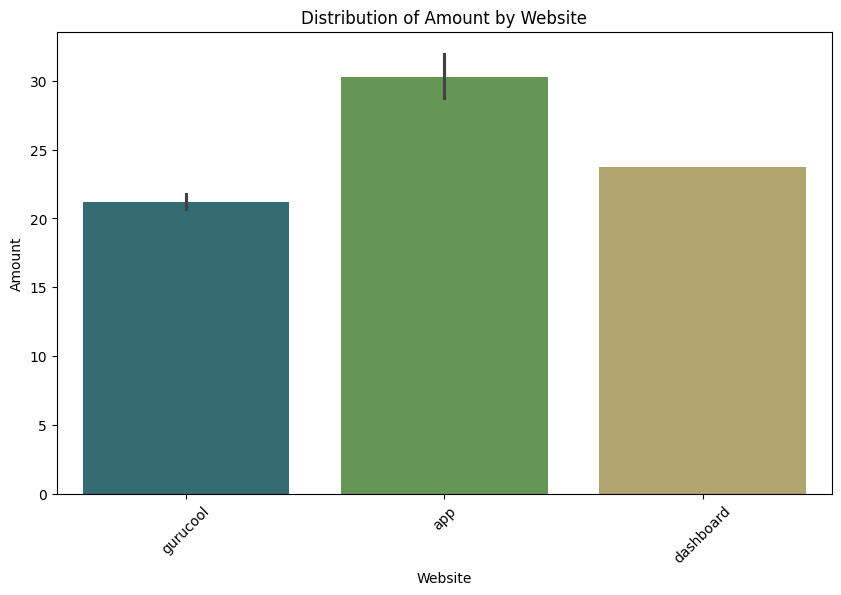


**Distribution of Time Duration:**

- Created a histogram to visualize the distribution of time duration.

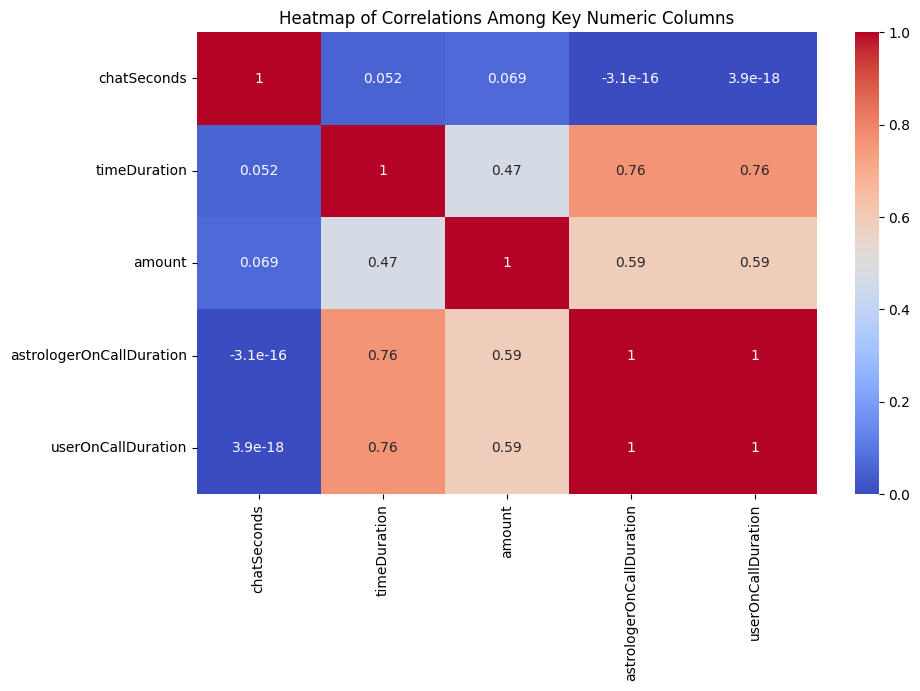
**Distribution of Amount by Website:**

- Created a bar plot to show the distribution of amount by website.



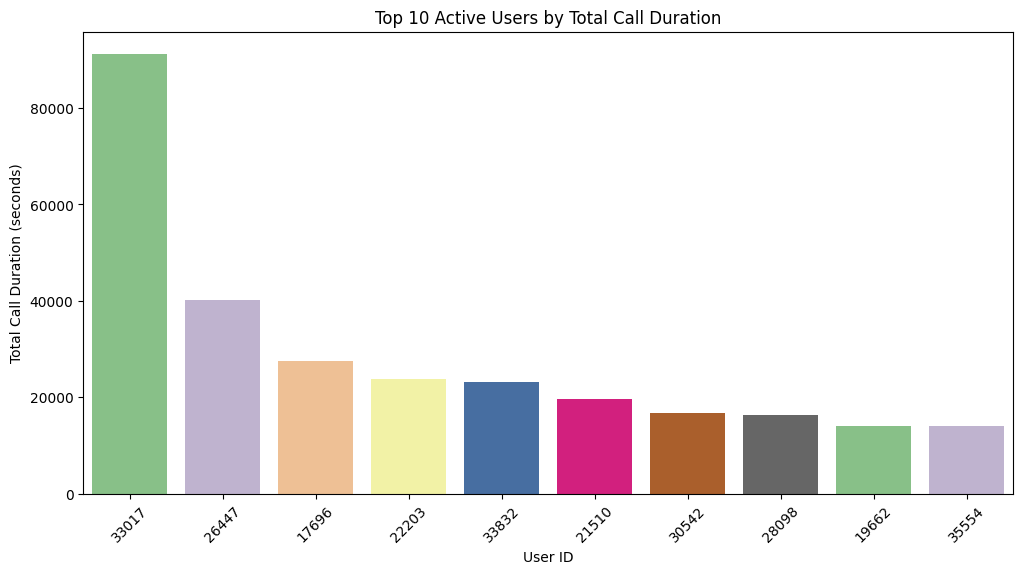
**Heatmap of Correlations Among Key Numeric Columns:**

- Created a heatmap to show correlations among key numeric columns.



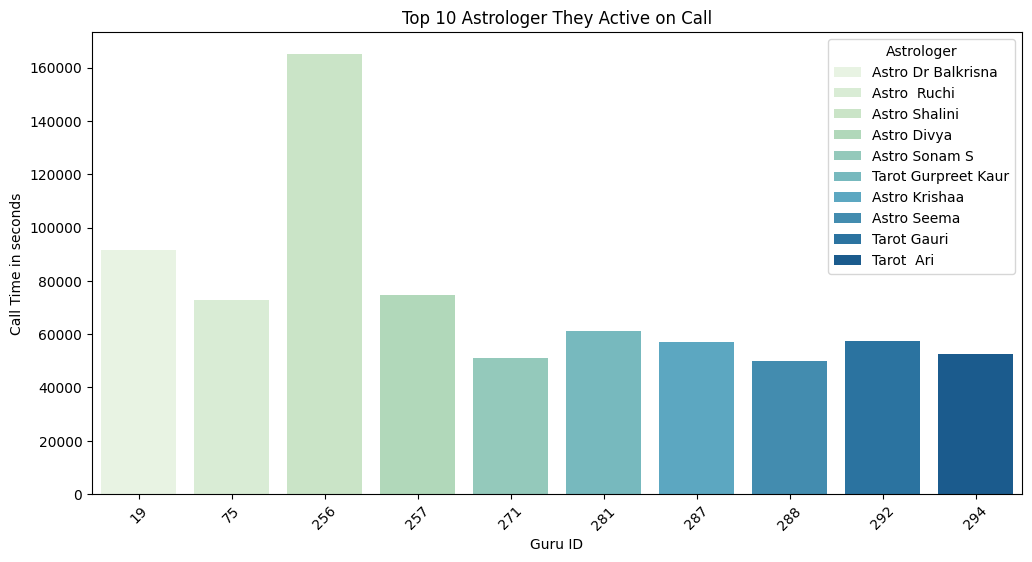
**Top 10 Active Users:**

- Identified and visualized the top 10 active users by total call duration.



**Top 10 Users Actively Connected with Astrologers:**

- Identified the top 10 users who were most actively connected with astrologers.



**Insights and Visualizations:**

1. User Engagement:

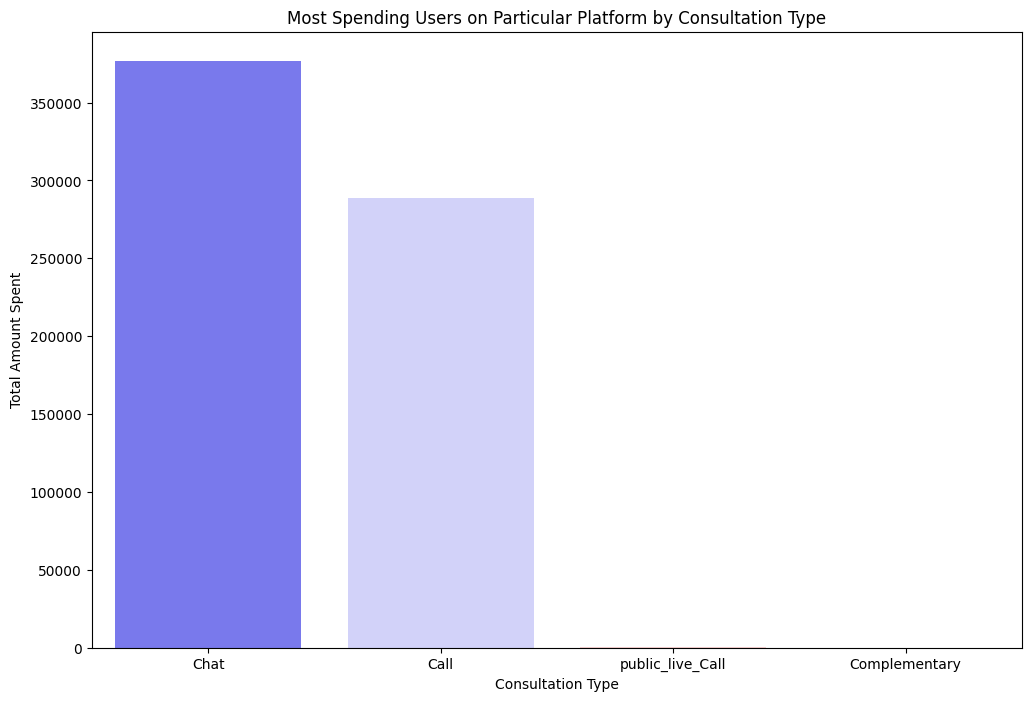
- Significant variation in call durations among top users.

- User 33017 has a notably higher engagement level.

**2. Usage Patterns:**

- Skewed distribution with a small number of users contributing a large portion of total call time.

**3. Most Spending User On Particular Platform:**

****