Unlocking Insights for Flipkart


Introduction

- The problem statement tells us that **"The goal is to identify specific issues within customer service operations that may be affecting customer satisfaction and retention rates**". So, we will focus only on the customer service aspect of customer retention.

- Customer retention can be influenced by other factors, but we are not focusing on those in this analysis.

- Our task is to improve customer retention by analysing customer service operations at Flipkart. To do this, let's examine the steps a customer takes when interacting with customer support. Here is the user journey for a customer trying to buy a product on Flipkart.

User Journey Map
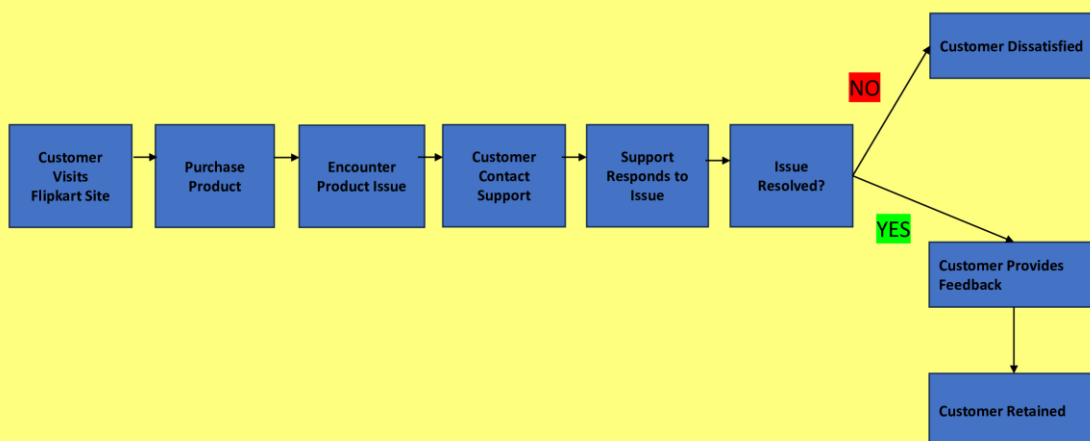
So, in above user journey the customer is doing:

- Customer Visits Flipkart:
  - The journey begins when the customer browses the Flipkart platform, exploring products.
- Customer Purchases Product:
  - The customer selects and buys a product, initiating the transaction.
- Customer Encounters Product Issue:
  - After receiving the product, the customer experiences an issue or has questions about the product.
- Customer Contacts Support:
  - The customer reaches out to Flipkart's support team for help with their issue.
- Support Responds to Issue:
  - Flipkart support responds to the customer, aiming to resolve the issue.
- Issue Resolved?
  - Yes: If the issue is resolved, the customer is satisfied and proceeds to provide feedback.
  - No: If the issue is not resolved, the customer may be dissatisfied, which could lead to potential churn.
- Customer Provides Feedback:
  - If the issue is resolved, the customer provides feedback, which Flipkart can use to improve service quality.
- Customer Retained:
  - A positive resolution and feedback increase the likelihood of customer retention.

Metric Tree:
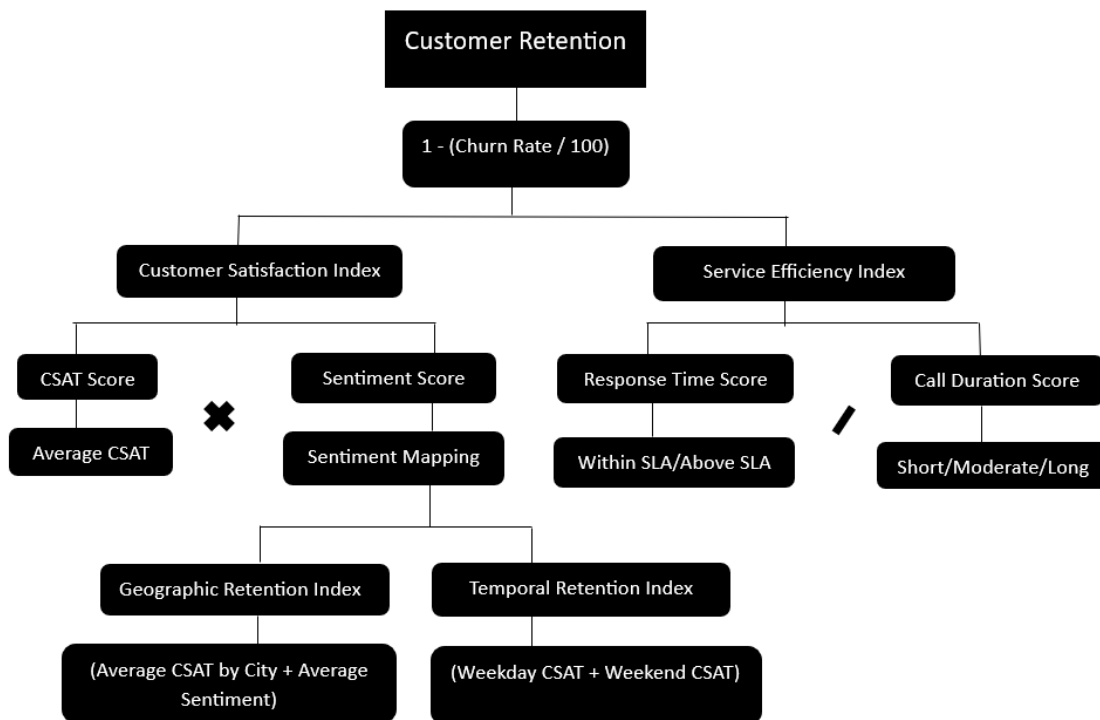
Details on Metric Tree

1. **Customer Retention** = 1 - (Churn Rate / 100)

Retention is inversely related to churn rate. A lower churn rate indicates higher retention, and this relationship provides a direct measure of how well the company retains its customers.

2. **Customer Satisfaction Index:**

This index combines CSAT Score and Sentiment Score to show a weighted satisfaction metric. The CSAT Score (e.g., on a 1-10 scale) gives a baseline satisfaction measure, while the Sentiment Score (e.g., Very Positive = 5, Very Negative = 1) adjusts this based on how customers feel about their experience. Multiplying them gives an overall measure of customer experience, where high satisfaction and positive sentiment lead to higher retention.
Formula: Customer Experience Index = CSAT Score * Sentiment Score

3. **Service Efficiency Index:**

Response Time Compliance / Avg Call Duration

This metric combines Response Time Compliance and Average Call Duration to measure service effectiveness. Compliance with SLA (1 if within SLA, 0 if not) shows if response times meet standards, while Call Duration (Short = 1, Long = 0) indicates how quickly issues are resolved. Dividing SLA compliance by call duration helps highlight areas where faster service improves retention. A higher score indicates efficient service.

4. **Geographic Retention Index:**

Geographic Retention Index = Avg CSAT by City + Avg Sentiment by City

Different cities may have unique challenges affecting satisfaction. Calculating Average CSAT and Average Sentiment by city highlights geographic patterns in customer retention. Adding these averages gives an index that identifies areas needing improvement. For example, low scores in certain cities might prompt city-specific initiatives to boost retention.

5.**Temporal Retention Index:**

Temporal Retention Index = Average Weekday CSAT + Average Weekend CSAT
This index helps measure retention trends based on the day of service. Differences between weekday and weekend CSAT scores may indicate that customer satisfaction fluctuates based on time, possibly due to staffing or service quality on weekends. Adding weekday and weekend CSAT scores gives a single index to highlight time-based retention trends

Defining Key Metric :

Finally, the key metrics for this analysis are:

- CSAT Score: Measures customer satisfaction based on their support experience.
- Sentiment: Captures the emotional tone of customer feedback, indicating their feelings toward the service.
- Percentage of Responses Within SLA: Shows how often support responses meet the agreed time limits, reflecting promptness.
- Call Duration: Tracks the length of support calls, indicating how quickly customer issues are addressed.

Data Cleaning :

- **Handling Missing Data :**

  i) Given that over half of the entries in the CSAT_Score field are missing, cleaning could theoretically be applied. However, it was decided not to fill these values, as they represent direct customer feedback. The absence of a score may indicate that the customer opted not to provide feedback following problem resolution.

    (a) An alternative approach involves imputing missing CSAT_Score values by calculating an average CSAT_Score based on the associated Sentiment, producing an approximate CSAT score. This imputed score has been included in the cleaned dataset for reference but remains unanalyzed, as it does not reflect the actual customer scores.

  ii) Missing values in the city and state fields are likely due to system-related errors.

- **Removing Duplicates :**

  i) The ID column uniquely identifies each customer, ensuring that all entries are distinct despite possible similarities in customer names. No duplicate entries exist within this column.

- **Standardizing the Data :**

  i) The Name field has been standardized by removing trailing spaces and ensuring uniform formatting. Symbols such as "@" have been replaced with spaces to maintain consistency.

  ii) The timestamp column has been reformatted to standardize dates in the dd-mm-yyyy (eg: 29-10-2023) format.

- **Data Processing :**

  i) An additional column, Call Duration Category, was introduced to categorize into segments (<15 minutes (short), <30 (moderate), and >30 minutes (long)) to support targeted insights and streamlined categorization.

ii) The day of the week has been extracted from the call_timestamp field to enhance temporal analysis. The Week Number has been extracted as well in a separate column.

iii) The *Call Resolution column was added to track whether customer issues were resolved during the initial interaction. This metric helps assess support efficiency and its impact on customer satisfaction.
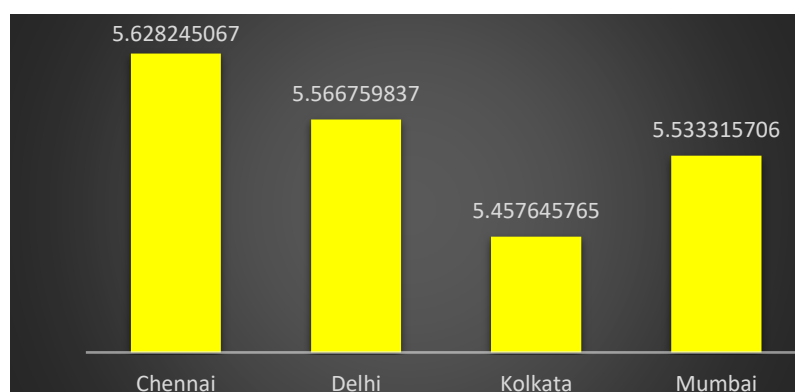
Hypothesis:

So we are done with the metric tree and also know key metrics to focus on. Let's make a list of hypothesis that we can check and analyse while doing EDA

1.Call Center Performance

- **Hypothesis 1: Certain call centers demonstrate consistently lower satisfaction scores, suggesting performance disparities across locations.**

  Analysis: Upon examining the data, **Kolkata** demonstrated the lowest average CSAT score across all call centers, indicating potential issues in meeting customer expectations in this location.This disparity in CSAT scores between call centers could indicate differences in process efficiency, resource availability, or regional challenges affecting satisfaction.

  Conclusion: The hypothesis is validated as certain call centers indeed demonstrated consistently lower or higher satisfaction scores.

- **Hypothesis 2: First-Call Resolution Improves Overall Satisfaction Scores**

Analysis: The data shows that FCR rates and CSAT scores are not directly proportional. For instance, "**Payments**" has the lowest FCR rate (24.75%) but the highest CSAT score (5.63), while "**Service Outage**" has the highest FCR rate (27.73%) with a slightly lower CSAT (5.5). This inverse relationship suggests that higher FCR rates alone don't guarantee higher customer satisfaction.

Hypothesis 3		
Row Labels	FCR By Issue Type	Average of csat_score
Billing Question	26.21%	5.532004981
Payments	24.75%	5.634469697
Service Outage	27.73%	5.508125

Conclusion: The assumption that First Call Resolution improves overall satisfaction scores is contradicted by this data. Despite a lower FCR rate, "Payments" has a higher CSAT score, suggesting that factors beyond FCR influence customer satisfaction.

2. Regional Influence on Customer Experience

- **Hypothesis 1: Customers from particular cities or states show lower satisfaction levels, suggesting regional service challenges.**

    Analysis: *Tripura* exhibited the lowest average CSAT score, suggesting regional challenges in customer service or potential dissatisfaction unique to this area.

Hypothesis 1	
Row Labels	**Average of csat_score**
Andaman and Nicobar Islands	5.519230769
Andhra Pradesh	5.616071429
Arunachal Pradesh	5.963636364
Assam	5.774390244
Bihar	5.593548387
Chandigarh	5.2
Chhattisgarh	5.571428571
Dadra and Nagar Haveli	5.403846154
Daman and Diu	5.291666667
Delhi	5.683673469
Goa	5.806451613
Gujarat	5.683840749
Haryana	5.494983278
Himachal Pradesh	5.616438356
Jammu and Kashmir	5.543478261
Jharkhand	5.691542289
Karnataka	5.395061728
Kerala	5.569230769
Lakshadweep	5.425
Madhya Pradesh	5.24925816
Maharashtra	5.513780919
Manipur	5.358490566
Meghalaya	5.578947368
Mizoram	5.6
Nagaland	5.777777778
Odisha	5.76127321
Puducherry	5.918367347
Punjab	5.315457413
Rajasthan	5.50310559
Sikkim	5.924528302
Tamil Nadu	5.53647587
Telangana	5.709677419
Tripura	4.9
Uttar Pradesh	5.524480712
Uttarakhand	5.886792453
West Bengal	5.471861472

Conclusion: The analysis supports the hypothesis, identifying *Tripura* as a region with the lowest satisfaction levels, which could indicate underlying service challenges. Further investigation into the specific service needs and support infrastructure in Tripura is recommended to address regional satisfaction effectively.

3. Impact of Date and Time on Satisfaction

- **Hypothesis 1: Lower satisfaction scores on weekends may indicate staffing or availability issues affecting service quality.**

  Analysis: Unlike weekends, where CSAT scores remained stable, **Thursday** alone showed a marked decline, hinting that specific challenges on this day may affect service quality.

| Hypothesis 1	
Row Labels	**Average of csat_score**
Sunday	5.57493188
Monday	5.545079147
Tuesday	5.540594059
Wednesday	5.614780903
Thursday	5.425905598
Friday	5.562371134
Saturday	5.556825824

Conclusion: Thursday stand out as a day with lower CSAT scores, suggesting operational factors unique to that day may be affecting service quality

- **Hypothesis 2:How quickly an issue is resolved; whether within or beyond the expected SLA time,might affect customer satisfaction, with faster resolutions likely leading to happier customers.**

  Analysis: The data reveals that a majority of issues (62.49%) were resolved within the SLA, demonstrating the call center's effectiveness in meeting expected response times.
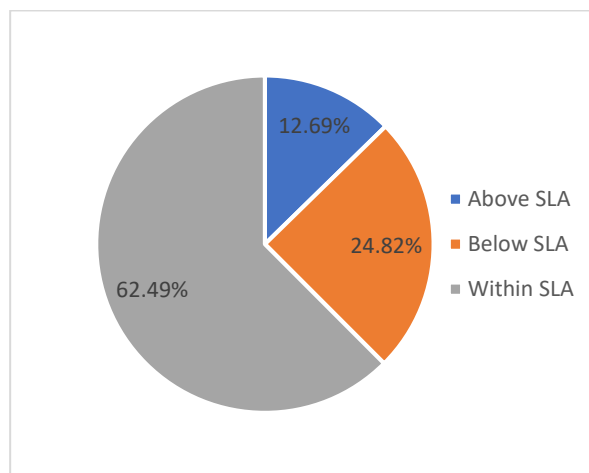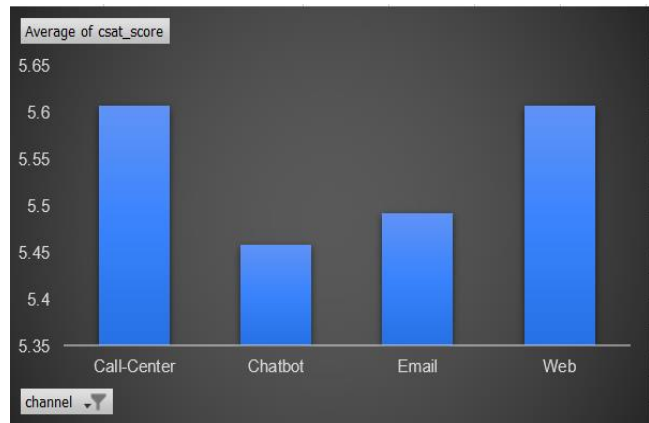

  Conclusion: To enhance customer satisfaction further, attention could be focused on the cases exceeding SLA to understand and address the reasons for delays, as this group might impact overall satisfaction.

4. Support Channel Effectiveness

- **Hypothesis 1: Different support channels (e.g., phone, chatbot) exhibit varying effectiveness, with certain channels achieving higher satisfaction scores.**

  Analysis: Interestingly, chatbot interactions recorded lower response times but did not consistently translate into higher CSAT scores, suggesting that quick responses do not always equate to greater satisfaction for complex issues.
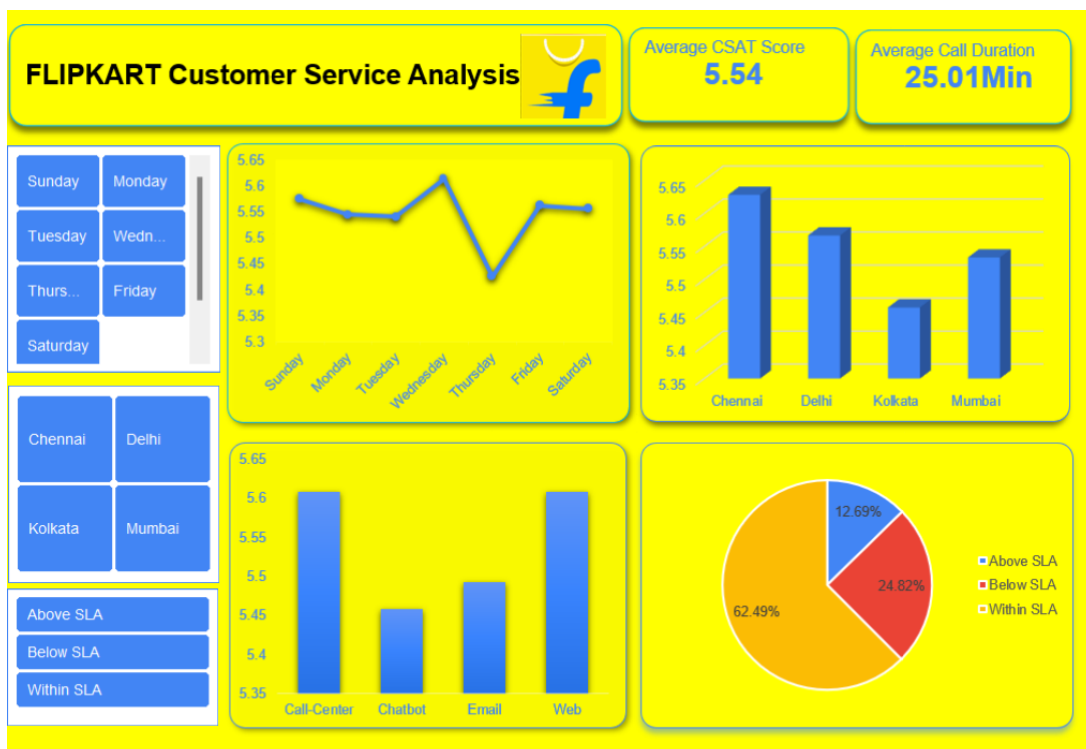
Conclusion: The call center and web appears to be the most effective support channels in terms of customer satisfaction. Encouraging the use of this channel or improving the performance of others could lead to higher overall CSAT scores.

Dashboard

Developed an interactive Dashboard to present key insights

Recommendations:

- o **Improve Chatbot Effectiveness:** The data indicates lower customer satisfaction for interactions through the chatbot compared to the call center. Enhancing the chatbot's ability to handle complex queries and providing an easy escalation path to human support can improve the customer experience.
- o **Prioritize Response Time for High-Sentiment Issues**: Since quicker response times are associated with higher satisfaction, ensure that issues with potentially negative sentiment or high importance are resolved within SLA to enhance customer retention.