Data Analysis using Excel

Checkpoint 1: Understanding the Business Context and Define Metrics

Introduction

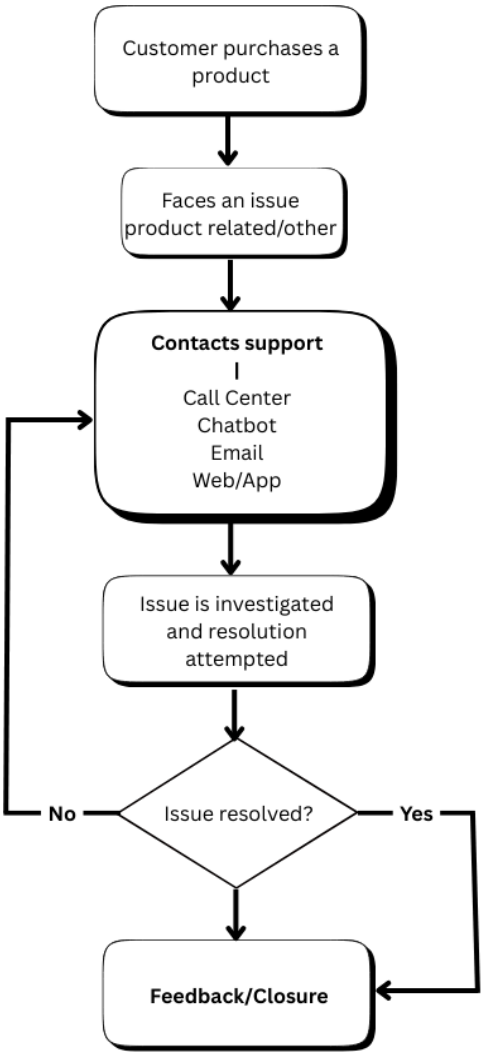
Flipkart is one of India’s largest and most trusted e-commerce platforms, serving millions of customers across the country with a vast array of products ranging from electronics and fashion to home essentials and lifestyle goods.

To maintain its competitive edge and build lasting customer loyalty, Flipkart places a strong emphasis on delivering exceptional customer service.

Flipkart is experiencing a decline in customer retention and wants to identify if customer service operations are impacting satisfaction and loyalty.

This project analyzes customer support data to define key metrics, establish their relationships, and formulate testable hypotheses for future analysis, with the goal of uncovering actionable insights to enhance customer retention.

User Journey Map

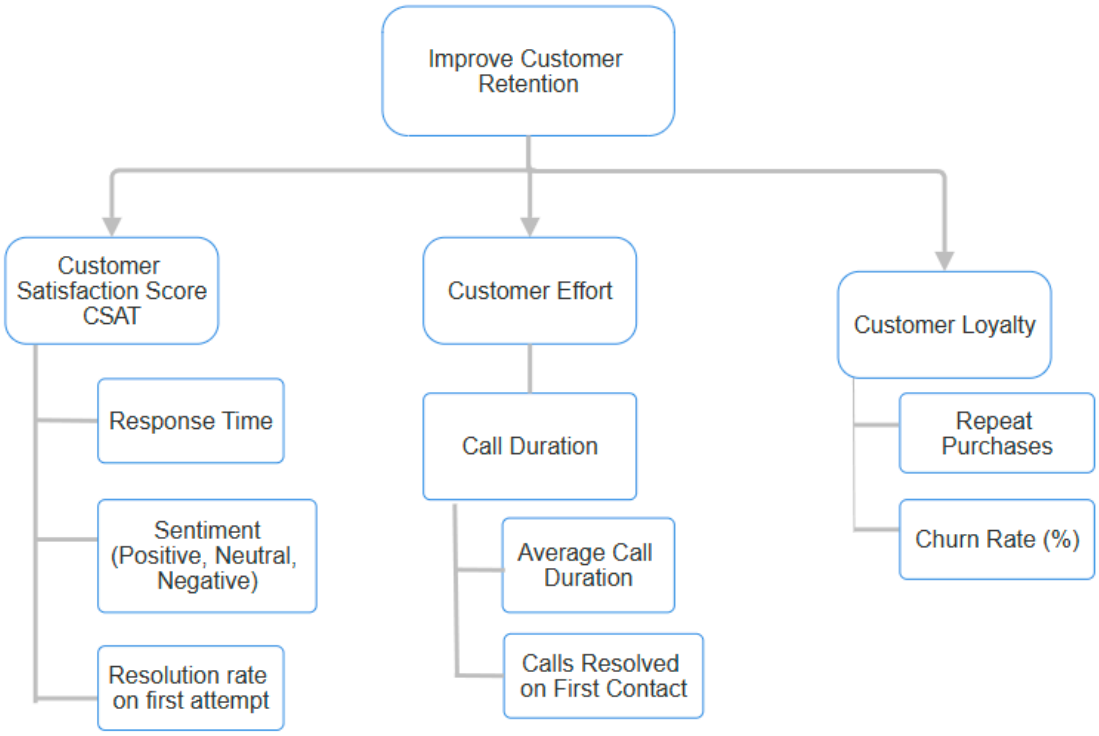


This flowchart illustrates the typical journey a Flipkart customer follows when seeking support after purchasing a product:

- Customer purchases a product: The journey begins after a successful order on Flipkart.
- Faces an issue (product-related/other): When the customer encounters any concern, whether related to the product itself or the shopping experience, the need for support arises.
- Contact support (Call Center, Chatbot, Email, Web/App): The customer chooses a preferred support channel—be it calling the support center, chatting via the app or website, sending an email, or using the web-based form.
- Issue is investigated and resolution attempted: The support team or automated assistant reviews the problem and attempts resolution, providing updates as required.
- Issue resolved?
    - If yes, the customer proceeds to the Feedback/Closure phase, where satisfaction is recorded via CSAT surveys or other feedback tools, and the case is closed.
    - If no, the case loops back to Contacts support for further action or escalation, ensuring persistent effort toward resolution.

Ensuring fast response times and resolving issues on the first contact helps increase Customer Satisfaction (CSAT) scores.
Higher CSAT scores contribute to improved customer loyalty and retention, making efficient support crucial to the overall user experience.

Metric Tree



Overall Structure

- The business goal is to improve customer retention.
- This is influenced by three main pillars: Customer Satisfaction (CSAT), Customer Effort, and Customer Loyalty.
- Each pillar is measured through a set of actionable, quantifiable sub-metrics.

Metric Explanations & Equations

- Customer Satisfaction Score (CSAT):
  - *Response Time (RT):*

$$\text{Average Response Time} = \frac{\sum RT \text{ for all tickets}}{\text{Total tickets}}$$

- *Sentiment*: Percentage or count of feedback categorized as Positive, Neutral, or Negative.
- *Resolution Rate on First Attempt (FCR):*

$$FCR\ (\%) = \frac{Issues\ resolved\ on\ first\ contact}{Total\ issues} \times 100$$

- Customer Effort:
  - *Call Duration (CD):*

$$\text{Average Call Duration} = \frac{\Sigma\ Call\ Duration}{Total\ Calls}$$

  - *Calls Resolved on First Contact:* Count or percentage, same as FCR shown above.

- Customer Loyalty:
  - Repeat Purchases:

$$\text{Repeat Purchase Rate (\%)} = \frac{Customers\ making > 2\ purchases}{Total\ Customers} \times 100$$

  - *Churn Rate (%):*

$$\text{Churn Rate} = 100\% - \text{Retention Rate}$$

If data on customers at the start and end of a period is available,

$$\text{Retention Rate} = \frac{Customers\ at\ end\ of\ period - New\ Customers}{Customers\ at\ start\ of\ period} \times 100$$

Impact of CSAT on Customer Loss

To estimate the impact of Customer Satisfaction Score (CSAT) on the number of customers lost, we use the formula:

$$\text{Estimated Customers Lost} = \left(1 - \frac{CSAT}{100}\right) \times \#\text{ customers in support queue}$$

- When CSAT is high, the estimated number of customers lost is low.
- When CSAT is low, the estimated number of customers lost is high, indicating dissatisfaction drives customer churn.

Key Metrics

- **Customer Satisfaction Score (CSAT):** Measures how satisfied customers are with the support experience, usually on a scale of 1-10.
- **Response Time:** Average time taken by support to respond to a customer query or call.
- **Sentiment (Positive, Neutral, Negative):** Qualitative measure of customer feedback tone expressing satisfaction or dissatisfaction.
- **Resolution Rate on First Attempt:** Percentage of issues resolved successfully in the first customer interaction without follow-ups.
- **Customer Effort:** Represents the effort customers expend to get their issues resolved, impacting satisfaction.
- **Call Duration:** Average length of calls between customers and support agents.
- **Customer Loyalty:** Measures repeat business and long-term engagement of customers with the company.
- **Churn Rate (%):** Percentage of customers lost over a specific period, calculated as 100% minus retention rate.

Listing Hypothesis

Hypothesis 1: Certain call centers have lower CSAT scores compared to others.

Hypothesis 2: Calls with response times exceeding the SLA (Above SLA) have lower CSAT scores.

Hypothesis 3: Longer call durations are associated with lower CSAT scores.

Hypothesis 4: Negative or Very Negative customer sentiment correlates with lower CSAT scores.

Hypothesis 5: Specific issue reasons (Billing, Payments, Service Outage) have differing average call durations.

Hypothesis 6: Customers contacting through different channels experience different CSAT levels.

Hypothesis 7 : Regional differences in call duration and customer satisfaction

Analysis Summary for Final Hypotheses

Hypothesis 1: Certain call centers have lower CSAT scores compared to others.

Key Findings:
Average CSAT scores show Chennai has the highest satisfaction, while Kolkata has the lowest. ***"While Chennai shows the highest average CSAT, Delhi handles a much larger call volume, making its performance more operationally significant overall."*** However, the differences are marginal.

Conclusion:
While Kolkata performs slightly worse, the small gap indicates no major call center stands out as an urgent concern. Potential location-specific improvement opportunities exist but require further detailed investigation.

Hypothesis 2: Calls with response times exceeding the SLA (Above SLA) have lower CSAT scores.

Key Findings:
Average CSAT scores are similar across response time categories (Within SLA, Below SLA, Above SLA), with no consistent trend of lower CSAT for delayed responses.

Conclusion:
Response time adherence to SLA does not appear to significantly affect customer satisfaction in the data sample. SLA generally impacts CSAT negatively, but in this dataset Above SLA surprisingly shows higher CSAT. This anomaly suggests resolution quality may matter more than speed, and requires further analysis.

Hypothesis 3: Longer call durations are associated with lower CSAT scores.

Key Findings:
Correlation between call duration and CSAT is near zero, and average satisfaction does not vary meaningfully with call length.

Conclusion:
Duration alone does not predict customer satisfaction, suggesting call quality or issue resolution effectiveness holds more importance than length.

Hypothesis 4: Negative or Very Negative customer sentiment correlates with lower CSAT scores.

Key Findings:
Strong, clear relationship between sentiment and CSAT—very negative sentiment corresponds to a CSAT around 2.5, while very positive sentiment corresponds to a CSAT near 9.5.

Conclusion:

1 in 3 calls results in negative sentiment—improvement here can drive major CSAT gains.
Sentiment is a powerful predictor of satisfaction and a key area for driving customer experience improvements to reduce churn risk.

Hypothesis 5: Specific issue reasons (Billing, Payments, Service Outage) have differing average call durations.

Key Findings:
Average call durations and CSAT scores vary marginally across issue types, all clustering near 25 minutes for duration and around 5.5 for CSAT.

Conclusion:
Issue type does not significantly differentiate customer experience or call length in this dataset. Focus may shift to other factors for performance improvement.

Hypothesis 6: Customers contacting through different channels experience different CSAT levels.

Key Findings:
CSAT variation by contact channel is minor, with Web and Call Center slightly above Chatbot and Email, but differences are small.

Conclusion:
Channel choice has limited impact on customer satisfaction at the aggregate level, indicating consistent service quality across channels.

Hypothesis 7: Regional differences in call duration and customer satisfaction

Key Findings:
Average call durations and CSAT scores by region differ only slightly, with the North East region showing marginally higher duration and higher CSAT. No pronounced adverse impact found.

Conclusion:
Regional differences in call duration and satisfaction exist but are not stark. Language or infrastructure challenges hypothesized for Tier 2 cities require finer data granularity or qualitative insights for confirmation.