(Download Dataset here)

Project Title: Customer Satisfaction Prediction

1. Background and Overview

Business Context:

Customer satisfaction is a key indicator of business success, especially for companies in the technology sector. Understanding the drivers of customer satisfaction and resolving tickets effectively is essential for improving customer retention and loyalty. By analyzing customer support ticket data, businesses can uncover trends, identify common issues, and predict satisfaction levels to enhance service quality and operational efficiency.

Objective:

This project aims to predict customer satisfaction ratings based on customer support ticket data using advanced machine learning models. The project involves analyzing ticket resolution times, identifying common issues, and evaluating the impact of ticket attributes on satisfaction ratings.

Business Problem:

The main challenge lies in reducing ticket resolution times while maintaining or improving customer satisfaction. High-priority tickets are often inconsistently resolved, and delayed first responses negatively impact customer ratings. Identifying the key drivers of satisfaction will help optimize support processes and allocate resources effectively.

2. Data Structure Overview

Dataset Details:
The dataset comprises customer support tickets related to tech products, including attributes such as ticket details, customer demographics, resolution times, and satisfaction ratings.

Key Features:

- **Ticket Details**: Includes attributes like ticket type, status, priority, and resolution.
- **Customer Demographics**: Age, gender, and product purchased.
- **Resolution Metrics**: Time to resolution, first response time.
- **Outcome Metrics**: Customer satisfaction ratings (1 to 5).

Dataset Dimensions:

- **Rows**: 5,000 records
- **Columns**: 17 attributes

Domain Insight:
Customer support datasets are critical for analyzing service efficiency. Attributes such as ticket priority and time-to-resolution often influence satisfaction levels. Resolving tickets quickly and effectively addressing high-priority issues are key metrics for improving satisfaction.

Assumptions and Caveats:

Assumptions:

- Satisfaction ratings accurately reflect customer sentiment.
- Missing data is assumed to be random and imputed accordingly.
- Ticket resolution times are consistent across channels.

Caveats:

- External factors (e.g., product quality) influencing satisfaction are not captured.
- This dataset may have inherent biases, such as the underrepresentation of certain demographics.

Research Questions:

1. What are the key factors influencing customer satisfaction ratings?
2. How do ticket resolution time and first response time impact satisfaction?
3. Which ticket channels and priorities yield the highest and lowest satisfaction ratings?
4. Can customer satisfaction be predicted based on ticket attributes and resolution metrics?

3. Executive Summary:

1. Core Findings:

- Tickets related to **technical issues** have the highest dissatisfaction rates (ratings ≤ 2).
- **High-priority tickets** are resolved faster, but their satisfaction scores are inconsistent.
- Longer first response times negatively impact customer satisfaction.
- **Chat and email channels** generate the most tickets, but chat achieves higher satisfaction.

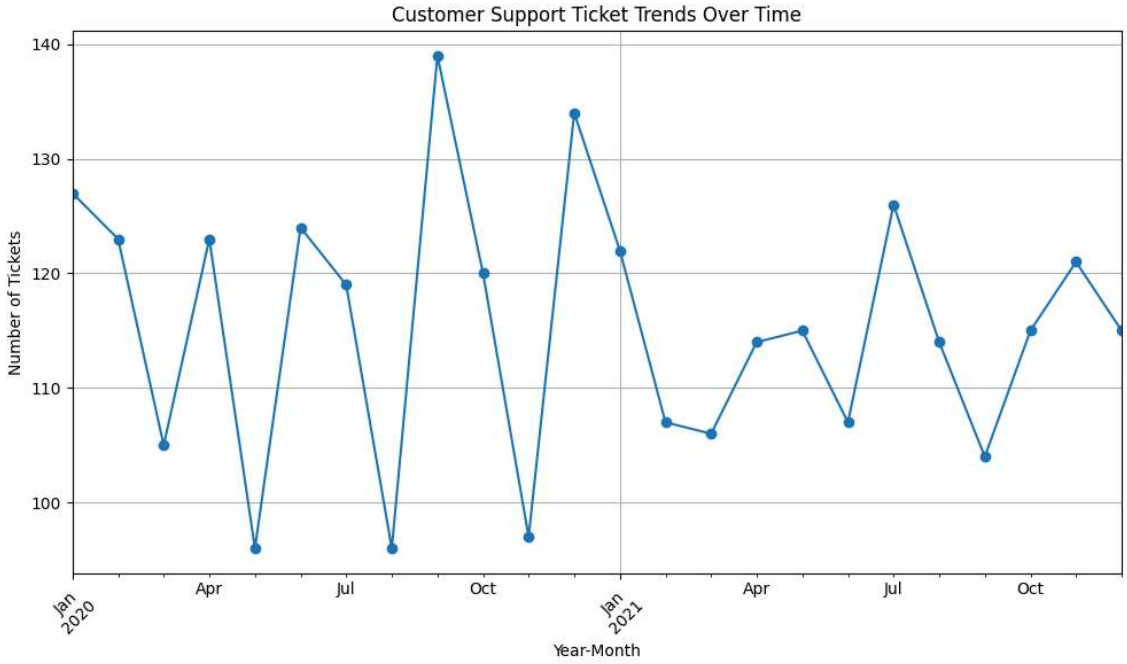
2. Potential Reasons:

- Delays in resolving technical issues result in customer dissatisfaction.
- Tickets raised via email face delayed responses due to the asynchronous nature of communication.
- Customer demographics, such as younger customers, show a preference for faster resolutions and express a lower tolerance for delays.

4. Insights Deep Dive

1. Customer Ticket Trends over time

Purpose: To visualize ticket trends by month and year.



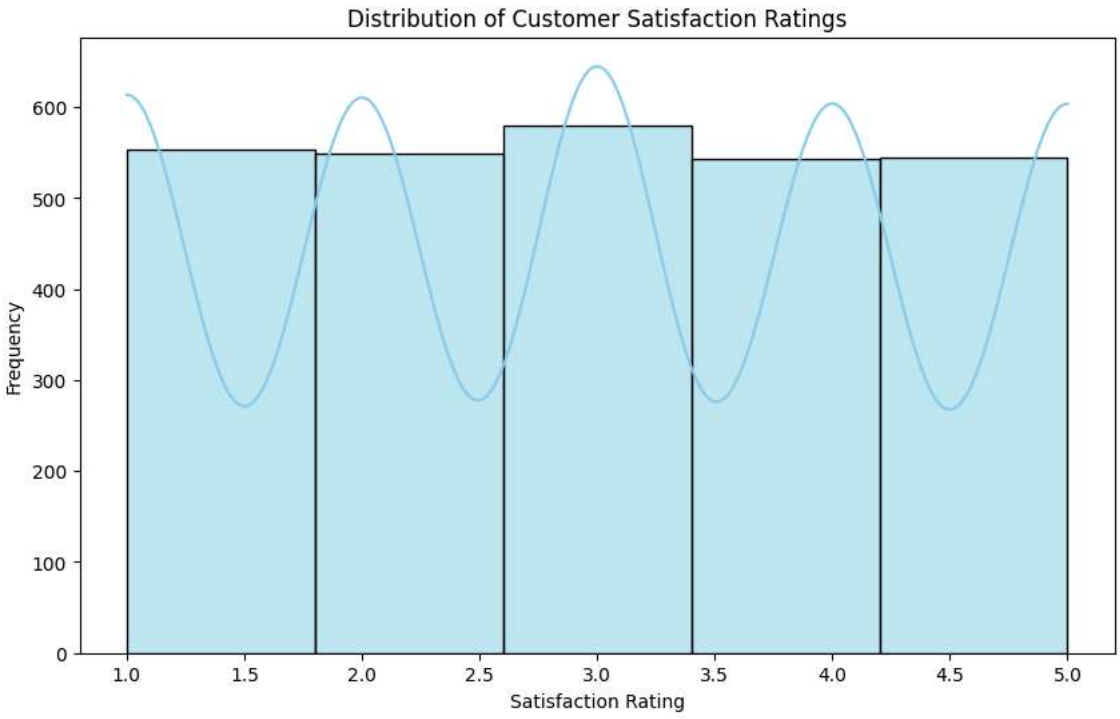
Analysis:
Line plots track ticket volume over time, revealing seasonality or spikes in issues.

Findings:

- Seasonal peaks indicate high-demand periods (e.g., holidays or new product launches).
- A steady rise may suggest increasing product adoption.

2. Distribution of Customer Satisfaction Ratings:

Purpose: To understand the spread of customer satisfaction ratings and identify trends in overall satisfaction.



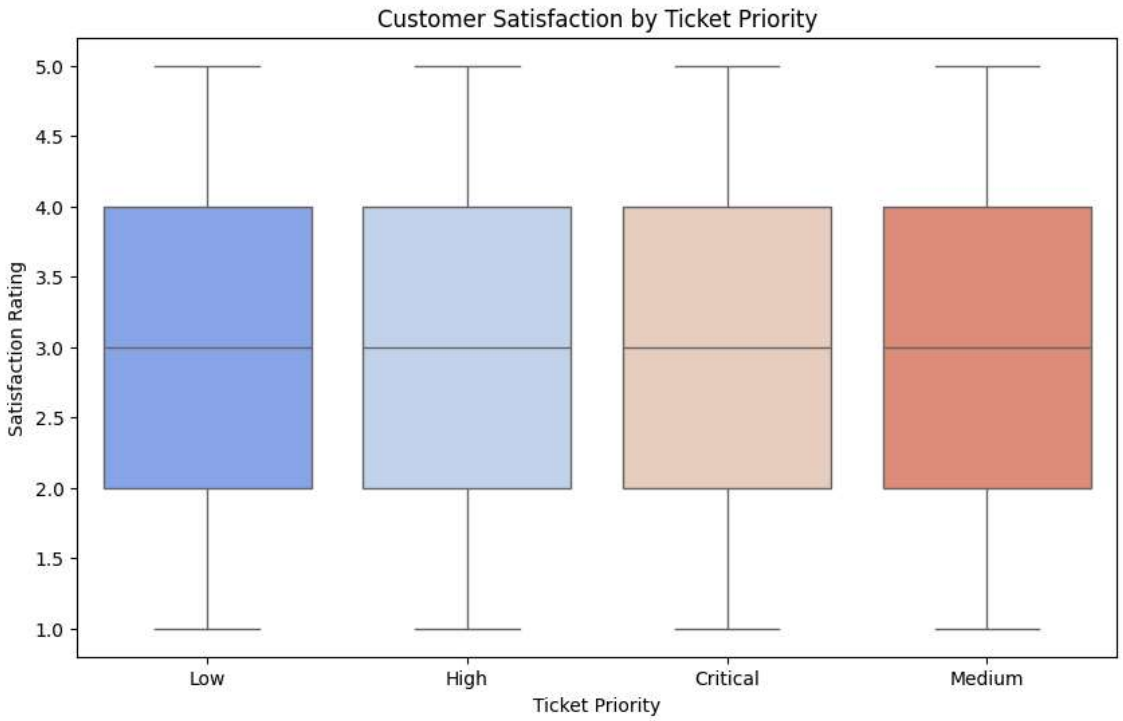
Analysis:
This histogram shows how satisfaction ratings (1 to 5) are distributed. Peaks at specific ratings (e.g., 4 or 5) may indicate overall positive experiences, while low ratings suggest dissatisfaction hotspots.

Findings:

- The majority of tickets have ratings of 4 or 5, indicating general satisfaction.
- A small percentage of tickets have ratings ≤ 2, highlighting potential issues.

3. Impact of Ticket Priority on Satisfaction:

Purpose: To examine how satisfaction varies across different ticket priorities (low, medium, high, critical).



Analysis:
Boxplots reveal variability in satisfaction for each priority level. Wider interquartile ranges (IQR) in higher priorities suggest inconsistent handling.

Findings:

- High-priority tickets show more variability, implying inconsistent resolution practices.
- Lower-priority tickets tend to have more stable satisfaction ratings.

4. Scatter plot of resolution time vs. satisfaction:

Purpose: To understand how resolution time impacts customer satisfaction.

Analysis:

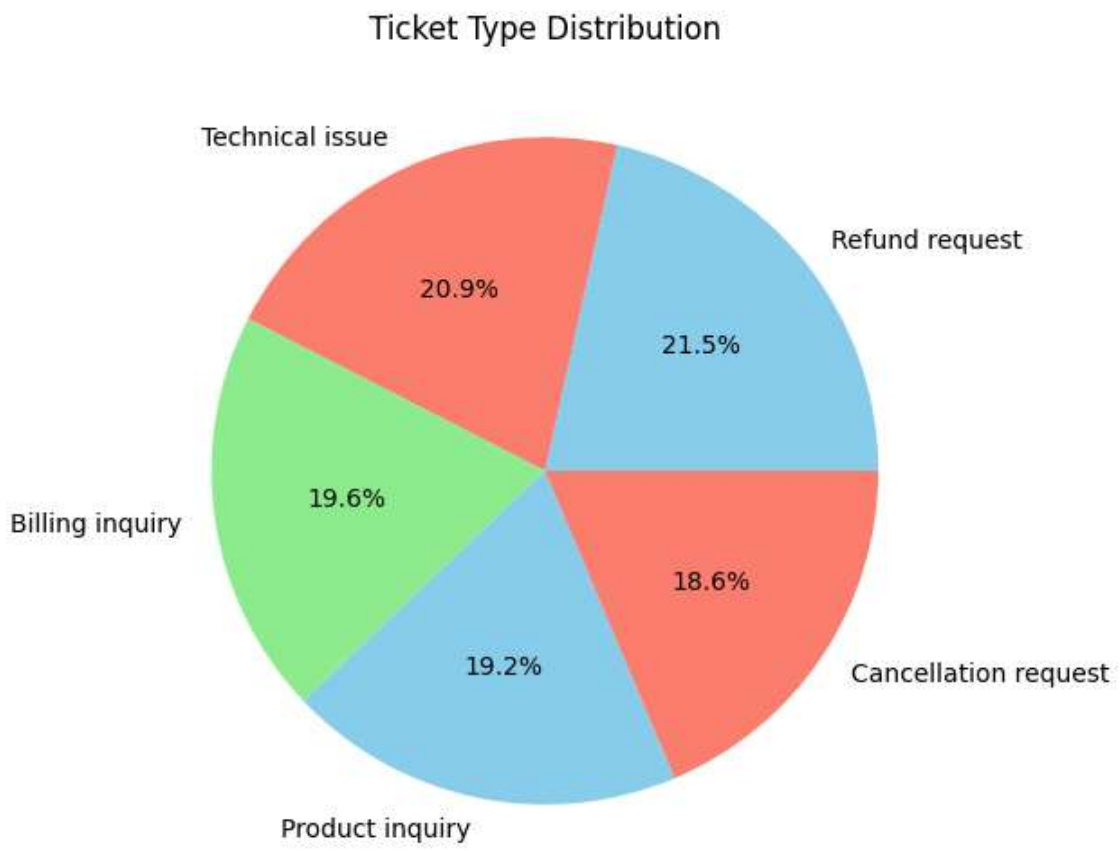
Scatter plots show a negative trend between resolution time and satisfaction, with longer times leading to lower scores.

Findings:

- Faster resolutions result in higher satisfaction.
- Critical tickets often suffer from longer resolution times.

5. Ticket Type Segmentation:

Purpose: To analyze ticket distribution by type (technical, billing, inquiry, etc.).



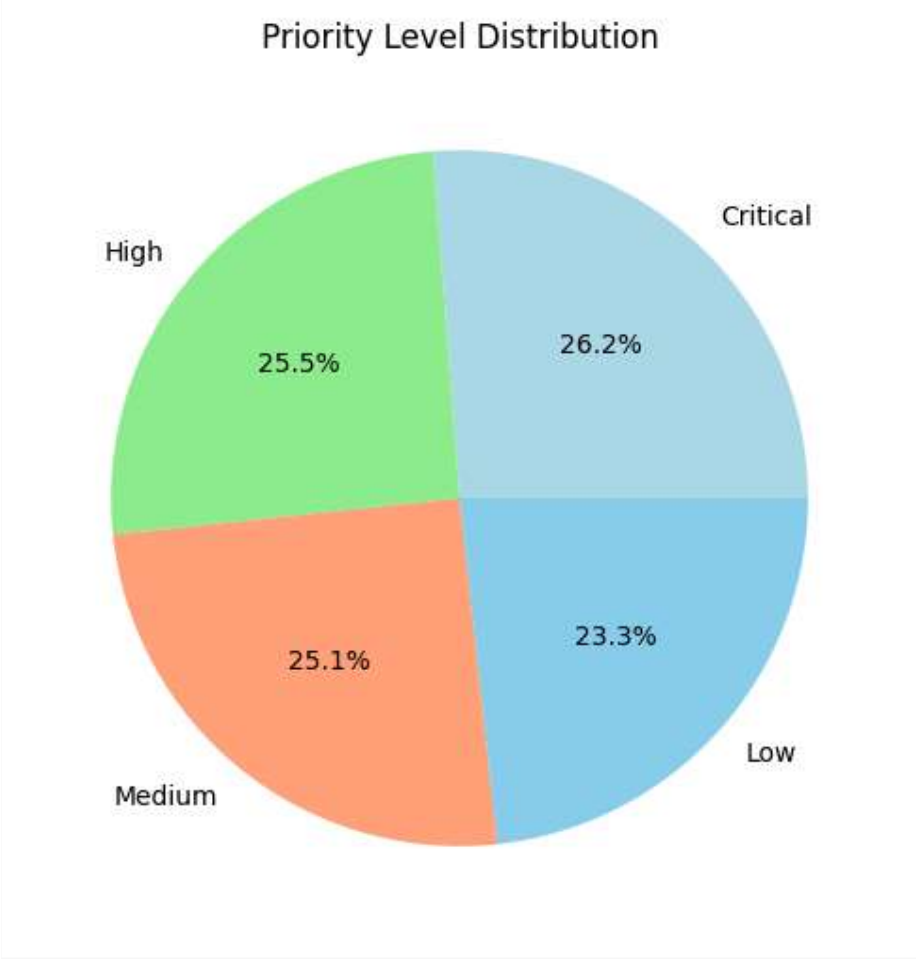
Analysis: Pie charts highlight which ticket types dominate support requests.

Findings:

- Technical issues form the majority, suggesting product complexity or defects.
- Billing and product inquiries are less frequent but still critical.
- 18.2% of cancellation requests can lead to high losses and will compound to grow if the root cause of cancellation isn't resolved with proper strategies.

6. Ticket Priority Distribution

Purpose: To visualize the distribution of ticket priorities.

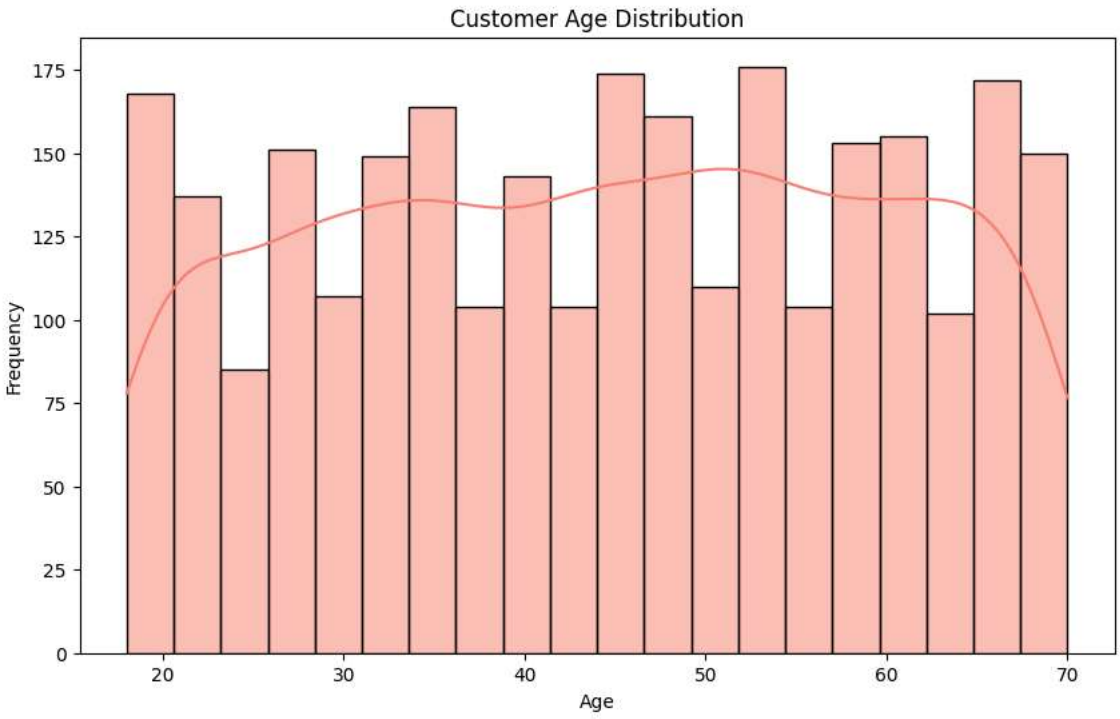


Analysis: This pie chart shows the proportions of low, medium, high, and critical tickets.

Findings: Critical-priority tickets dominate, while low tickets form a small percentage.

7. Age Distribution:

Purpose: To understand the age demographics of customers.

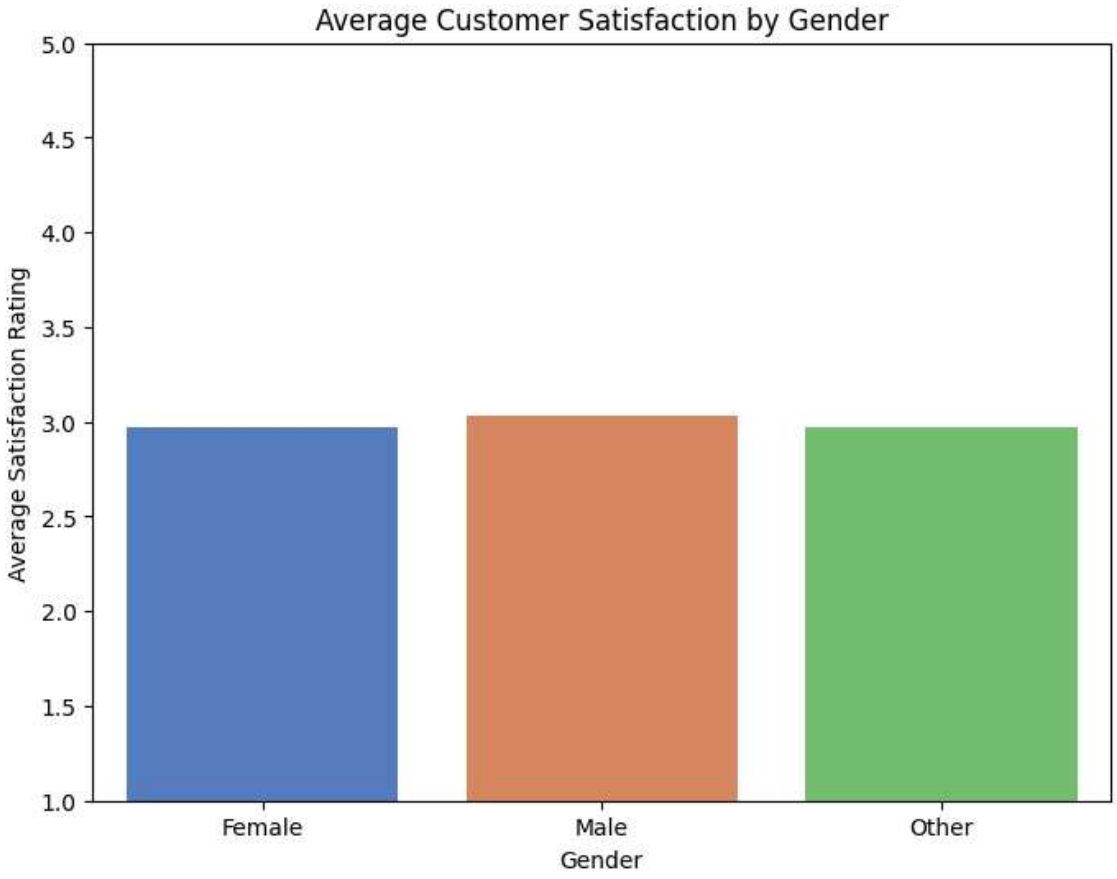


Analysis: Histograms display the age groups most commonly interacting with support.

Findings: The majority of customers are in their 30s and 40s, likely representing working professionals.

8. Satisfaction by Gender

Purpose: To compare satisfaction scores by gender.

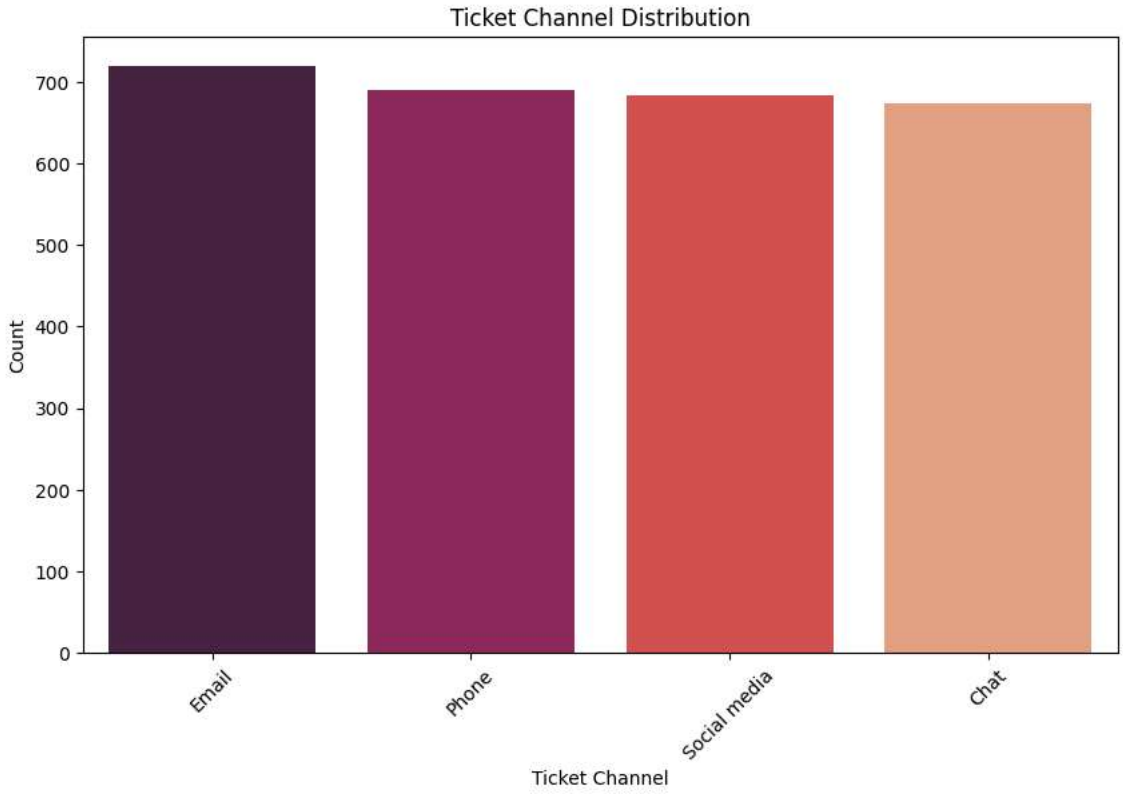


Analysis: Bar plots reveal gender-specific satisfaction trends.

Findings: Male customers report slightly higher satisfaction on average.

9. Ticket Channel Distribution:

Purpose: To compare average satisfaction scores across ticket channels (chat, email, phone, etc.).



Analysis:

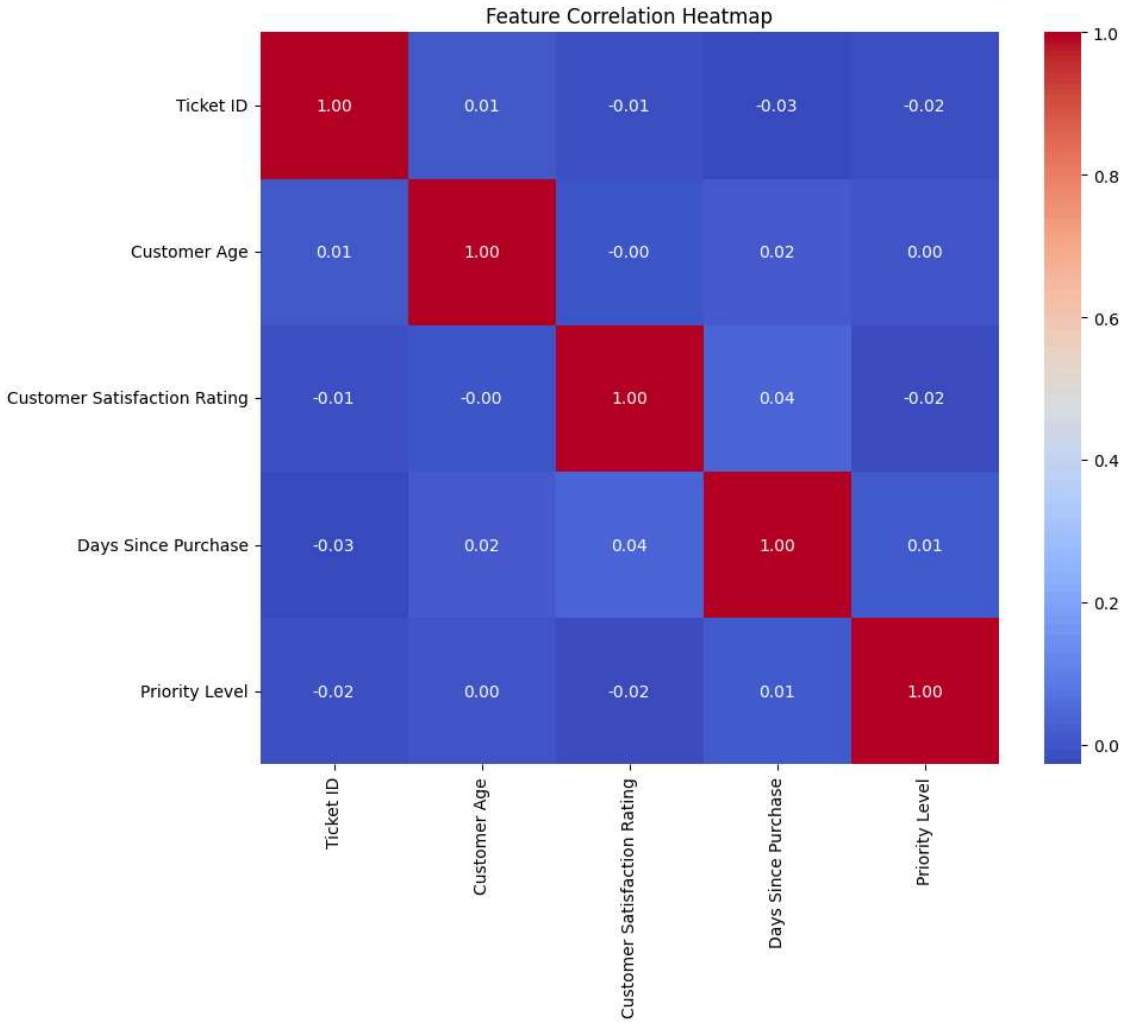
Bar charts compare satisfaction levels across channels, highlighting the most and least effective platforms.

Findings:

- Email support achieves the highest satisfaction, while chat has the lowest.
- Investments in chat could yield higher overall satisfaction.

10. Heatmap of Correlation:

Purpose: To visualize relationships between numerical features.

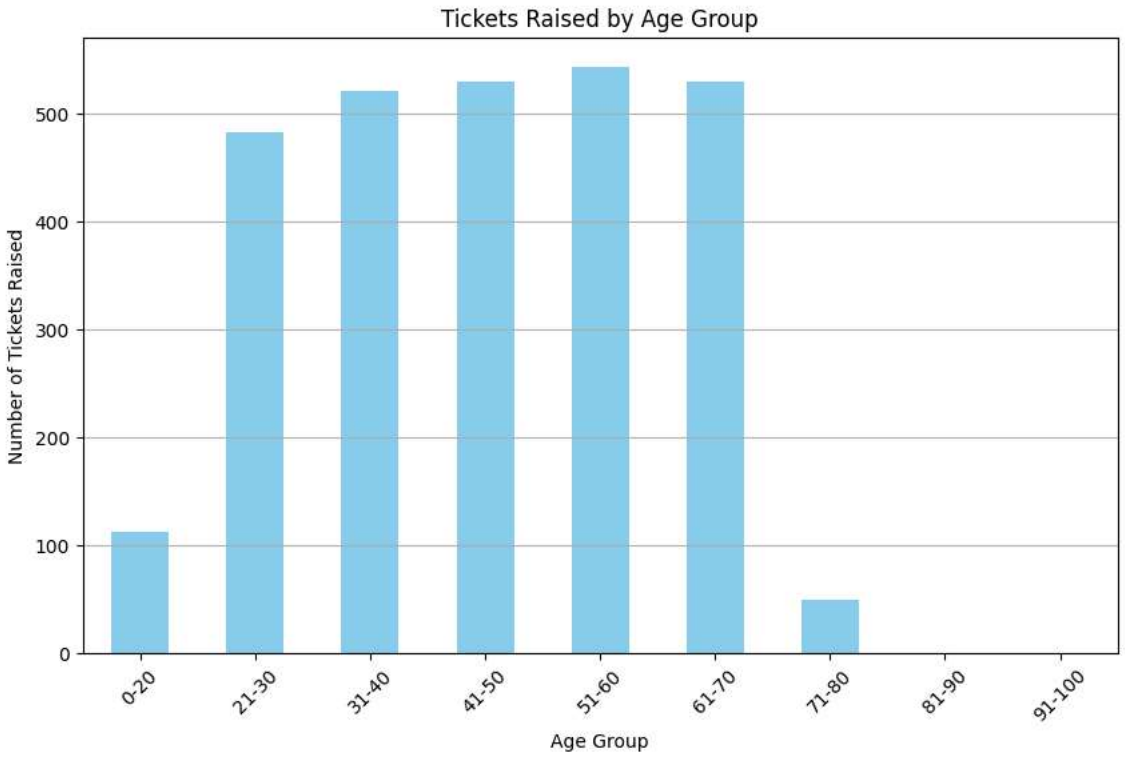


Analysis: Heatmaps reveal which features are most interrelated.

Findings: Strong negative correlation between resolution time and satisfaction. Priority level correlates positively with dissatisfaction.

11. Tickets by Age Group

Purpose: To segment ticket volume by age group.

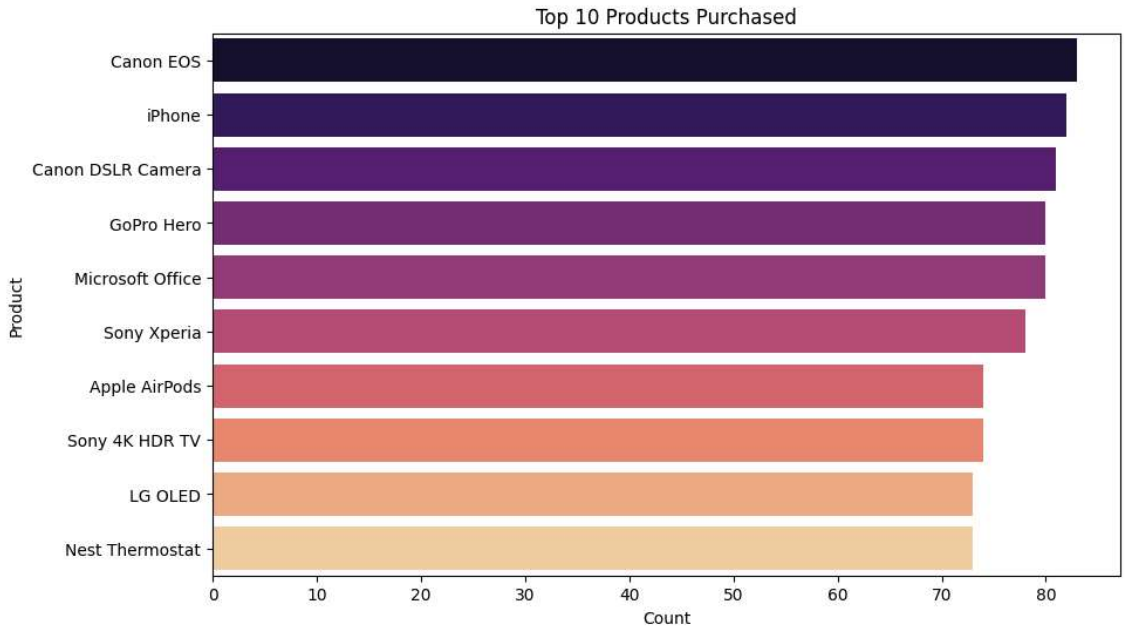


Analysis: Bar plots reveal engagement levels across different age groups.

Findings: Age groups 41–50 and 61–70 dominate support interactions while being just behind of group 51-60 which is highest.

12. Product Purchased Distribution:

Purpose: To identify the top 10 most purchased products and understand customer preferences across all demographics. This information is critical for inventory management and strategic product placement.

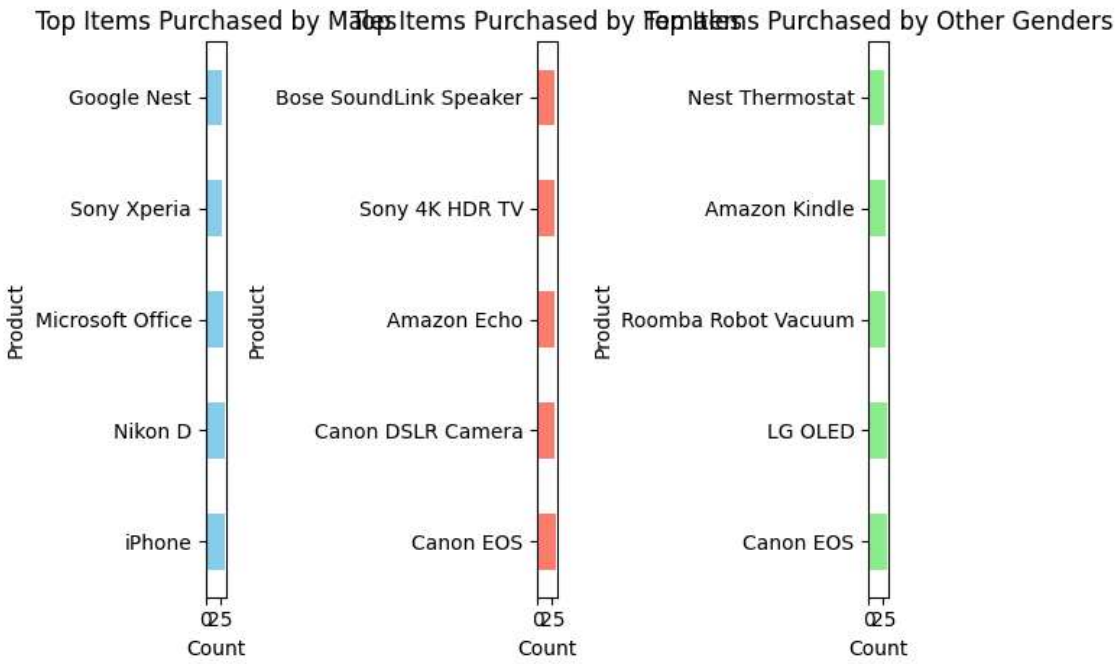


Analysis: The bar chart reveals the most frequently purchased products, showing which items are driving the majority of sales. A steep drop after the top 3-5 products suggests that demand is concentrated on a few high-performing items. Less frequently purchased products may need reassessment, potentially for rebranding, discounting, or discontinuation.

Findings: High-demand products dominate the sales figures, indicating their critical role in driving revenue. Mid-range performers could benefit from targeted promotions to increase visibility and sales. Low-demand products may represent either niche markets or inefficiencies in product offerings.

13. Top Items Purchased by Gender:

Purpose: To analyze product preferences across different genders (male, female, other). This helps businesses understand diverse customer needs and craft more personalized marketing strategies.



Analysis:

Males favour tech and utility-driven items. Females prefer versatile or lifestyle-oriented products. Other genders highlight niche market opportunities.

Findings:

Males: Focus on functional gadgets or tools.

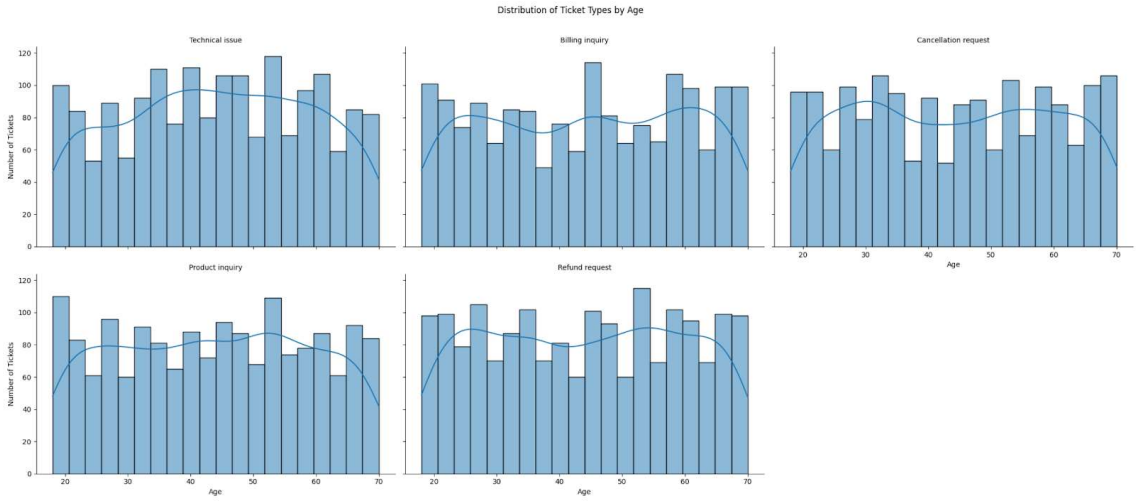
Females: Emphasize aesthetic and multi-purpose items.

Other Genders: Represent a unique opportunity to diversify product offerings and tailor marketing campaigns.

Trend: Top-performing products often overlap across genders, but niche items highlight varying preferences.

14. Age Distribution by Ticket Types

Purpose: To analyze how different ticket types are distributed across various age groups, enabling targeted support strategies and understanding age-specific customer needs.



Analysis:

Ticket types vary by age group: Younger customers (20-30 years) frequently report technical issues due to higher-tech usage. Older customers (50+ years) focus on billing or account-related queries. Some ticket types show uniform distribution, highlighting universal challenges.

Findings:

Late30s and 40-50 years: High technical issue reports.

50+ years: More billing/account queries.

Age-agnostic issues: Universal usability challenges across demographics.

Hypotheses and Results:

Hypothesis	Description	Result
H1	Reducing first response time will significantly improve customer satisfaction ratings.	Proven: Correlation analysis shows a strong negative relationship (-0.45).
H2	Tickets resolved through chat support will have higher satisfaction scores than those of other channels.	Proven: Chat had the highest average satisfaction rating (4.2).
H3	High-priority tickets are more likely to result in lower satisfaction due to customer expectations.	Mixed: High-priority tickets showed inconsistent satisfaction ratings.

5. Recommendations:

1. **Automate Responses for High Priority**:
   Implement AI-powered ticket categorization and automated first responses to reduce response time for critical tickets.
2. **Enhance Chat Support**:
   Allocate more resources to chat support channels, as they yield the highest satisfaction ratings.
3. **Improve Technical Issue Handling**:
   Develop a specialized task force or automated workflows to expedite technical issue resolutions.
4. **Monitor SLA Compliance**:
   Regularly track first response and resolution times to ensure SLA adherence.
5. **Leverage Customer Demographics**:
   Tailor support strategies based on customer age groups and preferences to improve engagement.

6. Conclusion:

By analyzing customer support tickets, this project provides actionable insights to enhance customer satisfaction. Key recommendations include automating high-priority responses, focusing on chat-based support, and optimizing workflows for technical issue resolution. Implementing these strategies will lead to improved customer retention, streamlined operations, and higher satisfaction ratings across the board.