

# Comprehensive Analysis of Cumulus Financial Service Complaints



IE 6600: Project 2

## Group 10

Krishna Teja Samudrala

Susilkessav Seshadri Bhuvaneswari

[samudrala.k@northeastern.edu](mailto:samudrala.k@northeastern.edu)

[seshadribhuvaneswa.s@northeastern.edu](mailto:seshadribhuvaneswa.s@northeastern.edu)

Percentage of Effort Contributed by Student 1: 50%

Percentage of Effort Contributed by Student 2: 50%

Submission Date: 10 December 2023

## Introduction

Understanding client complaints becomes critical for the finance industry in an era where consumer satisfaction and financial transparency are key. Through our examination of these complaints, we hope to find recurring themes, pinpoint common problems, and offer regulatory agencies and financial institutions useful information.

Here's a breakdown of its structure and the type of information it contains:

- Complaint ID: A unique identifier for each complaint.
- Date Submitted: The date when the complaint was submitted.
- Product: The type of financial product related to the complaint (e.g., Credit card, Debt collection, Mortgage).
- Sub-product: A more specific category within the product (e.g., General-purpose credit card, Checking account).
- Issue: The primary issue or concern raised in the complaint.
- Sub-issue: Further details about the issue.
- Company public response: Any public response provided by the company.
- Company: The name of the company the complaint is against.
- State: The state of the complainant.
- ZIP code: The ZIP code of the complainant.
- Tags: Additional tags or labels for the complaint.
- Consumer consent provided?: Indicates if the consumer provided consent for sharing their complaint.
- Submitted via: The medium through which the complaint was submitted (e.g., Web, Referral).
- Date Received: The date the company received the complaint.
- Company response to consumer: How the company responded to the complaint.
- Timely response?: Indicates if the response was timely.
- Consumer disputed?: Indicates if the consumer disputed the response.

The public dataset web URL of Financial Consumer Complaints - <https://data.world/markbradboune/rwfd-real-world-fake-data/workspace/file?filename=Financial+Consumer+Complaints.csv>

The data includes submission and received dates, allowing for time-based analysis and trend identification over specific periods. A multidimensional analysis that sheds light on several facets of customer satisfaction and the efficacy of institutional response mechanisms is made possible by the diversity of data.

## Problem Statement

The financial sector, which is essential to daily life and savings, is frequently criticized for how it handles customer complaints. In this industry, complaints can vary from minor inconveniences to serious issues that compromise consumer confidence and legal compliance. The problem at hand is multifaceted.

First, there is the challenge of effectively categorizing and understanding the nature of these complaints. They range from issues related to mortgage services, credit cards, and bank accounts to more intricate problems involving debt collection, credit reporting, and loan servicing.

The responsiveness of financial institutions to these complaints is critical. A timely and satisfactory response not only resolves individual issues but also enhances overall consumer trust in the financial system. On the other hand, delayed or inadequate responses can exacerbate dissatisfaction, leading to increased regulatory scrutiny and potential reputational damage.

Furthermore, the geographical distribution of these complaints can highlight regional disparities in financial services, possibly influenced by economic, demographic, or regulatory variances. Understanding these geographical patterns is essential for both national and regional financial entities and regulators.

# Methodology

Cleaning tools that we have used in this project are:

## 1. Google Co-lab:

In Google Co-lab, we have made cleaning and data formatting and removing null values, and dropping the duplicate values.

## 2. Tableau Prep:

With the use of Tableau prep, we came up with different visualizations and mainly used for analyzing the data in multiple ways.

Code Snippet that we have used in Google Collab to Clean the data:

```
import pandas as pd
import io
data = pd.read_csv(io.BytesIO(uploaded['Cleared_Financial_Consumer_Complaints.csv']))
```

|       | Complaint ID | Product                     | Sub-product                                | Issue                               | Sub-issue                             | Company public response                           | Company           | State | ZIP code | Tags         | Consumer consent provided? | Submitted via | Date Received | Company response to consumer | Timely response? | Consumer disputed? | Date Submitted |
|-------|--------------|-----------------------------|--|-------------------------------------|---------------------------------------|---|-------------------|-------|----------|--------------|----------------------------|---------------|---------------|------------------------------|------------------|--------------------|----------------|
| 0     | 2824926      | Credit card or prepaid card | General-purpose credit card or charge card | Closing your account                | Company closed your account           | Not Provided                                      | Cumulus Financial | FL    | 33179    | Not Provided | Consent not provided       | Web           | 2018-02-24    | Closed with explanation      | Yes              | Not Provided       | 2018-02-24     |
| 1     | 1515991      | Debit collection            | Credit card                                | Improper contact or sharing of info | Contacted employer after asked not to | Not Provided                                      | Cumulus Financial | CA    | 93561    | Not Provided | Consent not provided       | Web           | 2015-06-12    | Closed with explanation      | Yes              | No                 | 2015-06-12     |
| 2     | 1145261      | Bank account or service     | Checking account                           | Deposits and withdrawals            | ---                                   | Not Provided                                      | Cumulus Financial | CA    | 91792    | Not Provided | Not Provided               | Referral      | 2014-12-08    | Closed with explanation      | Yes              | No                 | 2014-12-09     |
| 3     | 2573783      | Mortgage                    | Conventional home mortgage                 | Trouble during payment process      | ---                                   | Company has responded to the consumer and the ... | Cumulus Financial | CA    | 92629    | Not Provided | Consent not provided       | Web           | 2017-07-13    | Closed with explanation      | Yes              | Not Provided       | 2017-07-13     |
| 4     | 2573113      | Mortgage                    | Conventional home mortgage                 | Struggling to pay mortgage          | ---                                   | Not Provided                                      | Cumulus Financial | NY    | 10018    | Not Provided | Consent not provided       | Web           | 2017-07-13    | Closed with explanation      | Yes              | Not Provided       | 2017-07-13     |
| ---   | ---          | ---                         | ---  | ---                                 | ---                                   | ---   | ---               | ---   | ---      | ---          | ---                        | ---           | ---           | ---                          | ---              | ---                | ---            |
| 75508 | 3080971      | Debit collection            | Credit card debt                           | Communication tactics               | Frequent or repeated calls            | Company has responded to the consumer and the ... | Cumulus Financial | GA    | 31701    | Not Provided | Consent provided           | Web           | 2018-11-22    | Closed with monetary relief  | Yes              | Not Provided       | 2018-11-22     |
| 75509 | 3111624      | Checking or savings account | Checking account                           | Managing an account                 | Deposits and withdrawals              | Not Provided                                      | Cumulus Financial | CA    | 90046    | Not Provided | Not Provided               | Phone         | 2016-12-28    | Closed with monetary relief  | Yes              | Not Provided       | 2016-12-28     |
| 75510 | 3141804      | Checking or savings account | Checking account                           | Managing an account                 | Deposits and withdrawals              | Company has responded to the consumer and the ... | Cumulus Financial | MD    | 20874    | Not Provided | Not Provided               | Referral      | 2019-02-04    | Closed with explanation      | Yes              | Not Provided       | 2019-02-01     |
| 75511 | 3081601      | Debit collection            | I do not know                              | Attempts to collect debt not listed | Debt was report or identity theft     | Company has responded to the consumer and the ... | Cumulus Financial | GA    | 302XX    | Not Provided | Consent provided           | Web           | 2018-11-23    | Closed with monetary relief  | Yes              | Not Provided       | 2018-11-23     |

Upon loading the dataset, examine its overall structure, taking note of the columns, data types, and initial rows. This was crucial to understand the nature of the data and identify potential areas requiring cleaning.

```
# Calculate the percentage of missing data in each column
missing_percentage = (data.isnull().sum() / len(data)) * 100

missing_percentage
```

|                              |           |
|------------------------------|-----------|
| Complaint ID                 | 0.000000  |
| Date Submitted               | 0.000000  |
| Product                      | 0.000000  |
| Sub-product                  | 0.000000  |
| Issue                        | 0.000000  |
| Sub-issue                    | 0.994531  |
| Company public response      | 63.800935 |
| Company                      | 0.000000  |
| State                        | 3.831123  |
| ZIP code                     | 5.783110  |
| Tags                         | 88.143763 |
| Consumer consent provided?   | 69.892739 |
| Submitted via                | 0.000000  |
| Date Received                | 0.000000  |
| Company response to consumer | 0.000000  |
| Timely response?             | 0.000000  |
| Consumer disputed?           | 48.927999 |
| dtype: float64               |           |

Our initial analysis revealed several columns with missing data, including 'Sub-issue', 'Company public response', 'State', 'ZIP code', 'Tags', 'Consumer consent provided?', and 'Consumer disputed?'

## Code Snippet of Python for Checking the missing data percentage :

```
# Filling missing values with appropriate placeholders
columns_with_placeholders = {
    'Sub-issue': 'Not Provided',
    'Company public response': 'Not Provided',
    'State': 'Unknown',
    'ZIP code': 'Unknown',
    'Consumer consent provided?': 'Not Provided',
    'Consumer disputed?': 'Not Provided'
}

for column, placeholder in columns_with_placeholders.items():
    data[column].fillna(placeholder, inplace=True)

# Considering dropping the 'Tags' column due to high percentage of missing data
data.drop('Tags', axis=1, inplace=True)

# Verify the changes
cleaned_data = data
print(f'Sum of missing data: {cleaned_data.isnull().sum()}\nData after cleaned: {cleaned_data.head()}')
```

|                              |   |
|------------------------------|---|
| ZIP code                     | 0 |
| Consumer consent provided?   | 0 |
| Submitted via                | 0 |
| Date Received                | 0 |
| Company response to consumer | 0 |
| Timely response?             | 0 |
| Consumer disputed?           | 0 |
| Date Submitted               | 0 |

dtype: int64

For columns with a lower percentage of missing data ('Sub-issue', 'State', 'ZIP code'), we filled missing values with a placeholder "Not Provided" or "Unknown".

In columns with a significant portion of missing data ('Company public response', 'Consumer consent provided?', 'Consumer disputed?'), we also used "Not Provided" to maintain data consistency without skewing the analysis.

```
[ ] # Standardizing date formats
# Converting 'Date Submitted' and 'Date Received' to datetime format
cleaned_data['Date Submitted'] = pd.to_datetime(cleaned_data['Date Submitted'], errors='coerce')
cleaned_data['Date Received'] = pd.to_datetime(cleaned_data['Date Received'], errors='coerce')

# Dropping the original 'Date Submitted' column after standardization
cleaned_data.drop('Date Submitted', axis=1, inplace=True)

# Check the changes and the types of these columns now
cleaned_data.dtypes, cleaned_data.head()
```

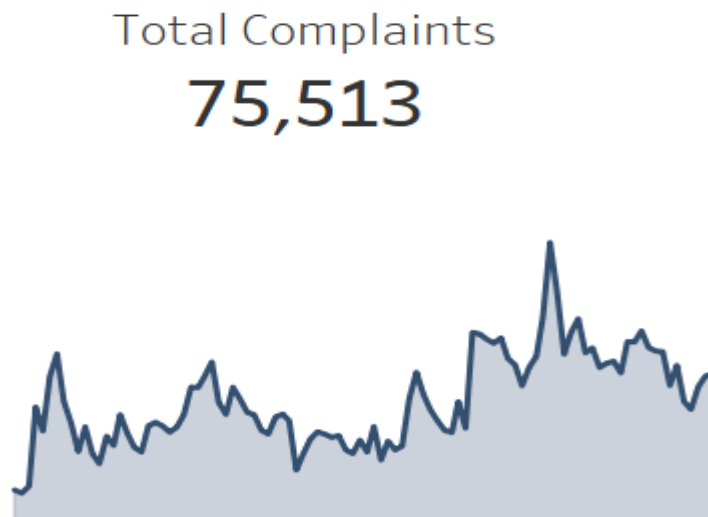
|                              |                |
|------------------------------|----------------|
| (Complaint ID                | int64          |
| Product                      | object         |
| Sub-product                  | object         |
| Issue                        | object         |
| Sub-issue                    | object         |
| Company public response      | object         |
| Company                      | object         |
| State                        | object         |
| ZIP code                     | object         |
| Consumer consent provided?   | object         |
| Submitted via                | object         |
| Date Received                | datetime64[ns] |
| Company response to consumer | object         |
| Timely response?             | object         |
| Consumer disputed?           | object         |

After that, we converted the data, classifying and encoding different text fields in order to make quantitative analysis easier. Products and issues, for example, were grouped to facilitate better aggregation and comparison. Upon completion of the cleaning, the cleaned and dataset will be transitioned into Tableau for advanced analysis and visualization. This transition marks a significant step in our journey to uncover deeper insights and patterns within the data. To explore various dimensions of the data, including trends over time, geographic patterns, and correlations between different complaint types and financial products. This phase is critical in transforming our initial findings into actionable insights that can guide decision-making and strategy for financial institutions and regulatory bodies.

In summary, the rigorous data cleaning and exploratory analysis set the stage for a more in-depth and insightful exploration in Tableau, paving the way for impactful outcomes from this project.

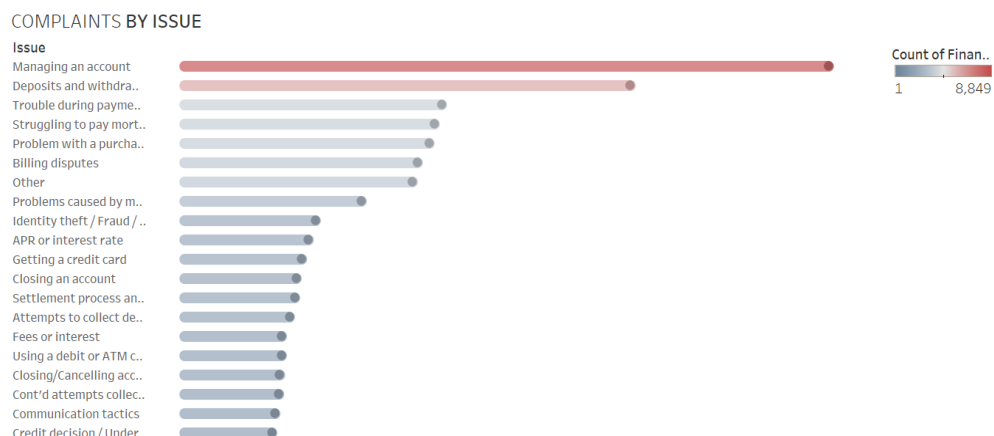
## Design Process

The dashboard was meticulously designed to offer a user-friendly interface that categorises complex data into digestible visual formats while offering the viewer a chance to do a “what-if” analysis. We employed a variety of graphs, each chosen for its ability to best represent the corresponding data:



Time Series Graph for Total Complaints:

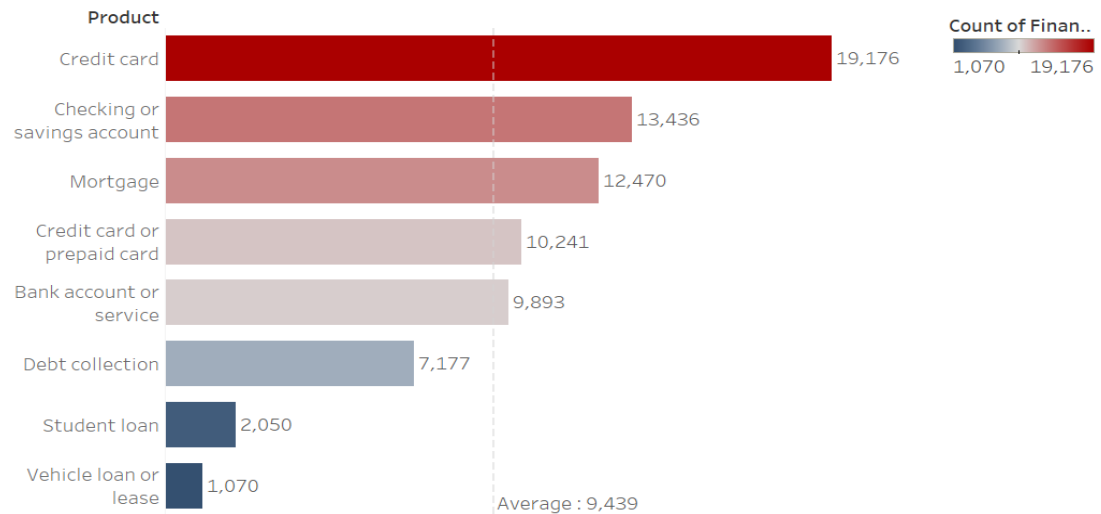
- To depict the volume of complaints over time, we used a line graph . Its continuous nature allows for quick identification of trends and patterns.



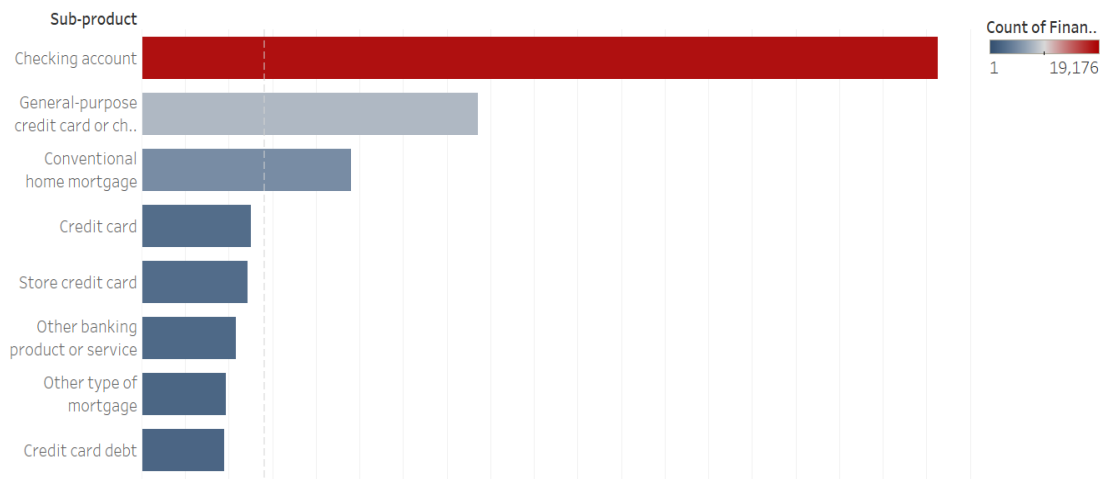
Horizontal Bar Charts for Complaints by Media:

- These were selected to compare the volume of complaints across different submission methods, usual assessment of the most prevalent channels customers use.

## COMPLAINTS BY PRODUCT



## CUSTOMER BY SUB-PRODUCT



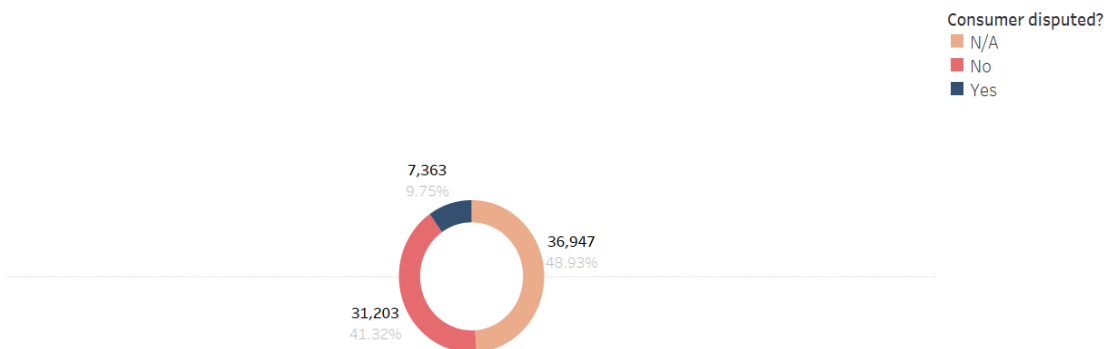
Bar Charts for Complaints with Average line by Product and Sub-Product with Navigation :

- This approach allows us to not only see the total number of complaints per product but also the distribution across sub-categories and compare with the average giving an estimate on the impact of that field compared to others, giving insight into specific areas of concern and we have drilled down by creating navigation .

## COMPLAINTS BY TAG



## CUSTOMER DISPUTED

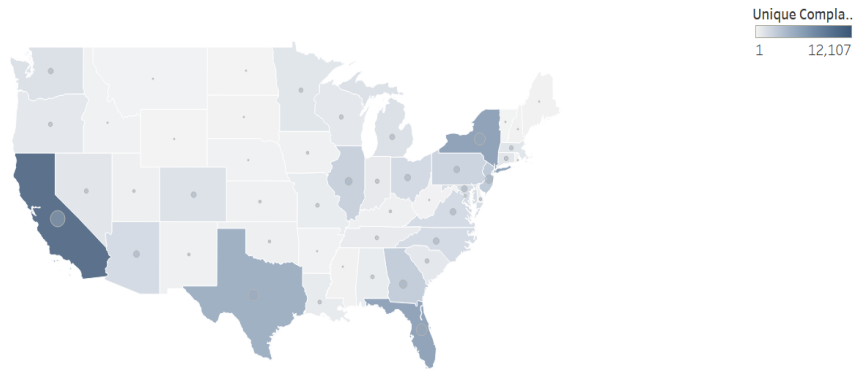


Donut Chart for Complaints by Tag and Dispute with Navigation :

- The donut chart provides a clear visual distinction between disputed and undisputed complaints, highlighting customer satisfaction or contention with the resolution process. In addition to this another chart to describe the tag (i.e Older American, Service Member, General) of the complaints was also used to profile and understand their different problems .



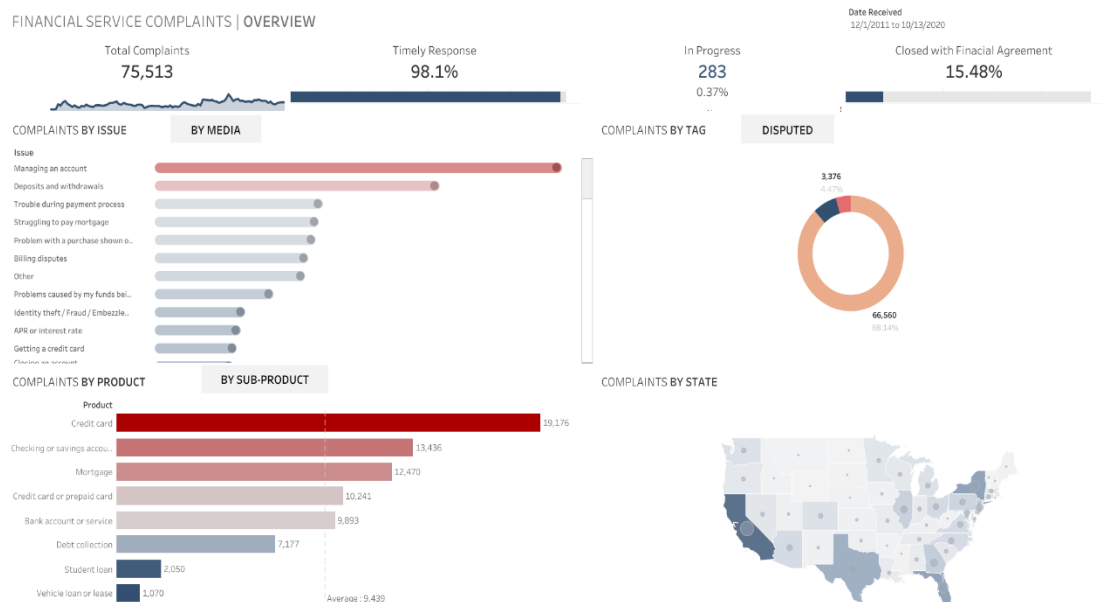
## COMPLAINTS BY STATE



Choropleth Map for Complaints by State:

- A geographical heat map effectively showcases the density and distribution of complaints across the United States, pinpointing areas with higher complaint volumes.

Please find below one sample dashboard:



Link to Dashboard : [Financial Service Complaints | Tableau Public](#)

## Key Insights

- **Complaint Medium:** The dominance of web submissions underscores the shift to digital platforms, signalling the need for robust online customer service portals.
- **Products and Services:** High complaint volumes for credit cards and bank accounts necessitate a focused strategy to address these key pain points.
- **Response Effectiveness:** While the timeliness of response is commendable, the rate of financial agreements points to potential improvements in resolving customers' grievances that lead to such agreements.
- **Dispute Resolution:** The low dispute rate suggests either effective initial resolutions or a lack of follow-through from consumers, an area that may benefit from deeper qualitative analysis.
- **Geographical Trends:** The state-wise distribution of complaints can guide localised customer service improvements and regulatory compliance.

## Conclusion

This comprehensive analysis of financial consumer complaints provides essential insights into the industry's customer service landscape. By identifying key issues, assessing response efficiency, and recognizing geographical trends, this report aims to contribute to the enhancement of consumer satisfaction in the financial sector. The findings serve as a valuable resource for financial institutions to refine their products and services, and for regulatory bodies to strengthen consumer protection mechanisms. Through visual storytelling, our dashboard takes the reader on a data journey from the macro-observations to the granular details of financial service complaints. Each graph was selected not just for aesthetic appeal but for its ability to convey complex data in an intuitive manner. The insights drawn from this visualisation highlight the critical areas for service enhancement and customer engagement in the institution.