

Navigating the Consumer Complaint Database

Unveiling Insights through Data Visualization

Data Visualization for Business Intelligence

Tandon School of Engineering

05/12/2023

Group 7

Alvina Bohora

Aditi Rani Uppari

Siddharth Bhardwaj

Sriram Suresh Sharma



Contents

- 1. Description of Scenario**
 - 1.1. Project Definition
 - 1.2. Objectives
 - 1.3 Meeting Record
- 2. Data sources**
- 3. Data Cleaning**
- 4. Dashboard**
 - 4.1. Description of Dashboard
 - 4.2. Overview
 - 4.3 Visualization
- 5. Tools and Technologies**
- 6. Target Audience**
- 7. Project Takeaways**
- 8. Citations**

Description of Scenario

1.1 Description

The Customer Complaint Dataset is a comprehensive collection of customer complaints received by the US Consumer Financial Protection Bureau (CFPB) since 2011. This database contains approximately 1.5 million individual complaints after cleaning the dataset, about a variety of financial goods and services. The dataset provides researchers with a one-of-a-kind opportunity to dig into the realm of consumer complaints and study the common challenges encountered by consumers in their contacts with financial institutions. Each complaint entry contains useful information such as the date of filing, the product or service category, the concern expressed, the firm involved, and the state of the customer.

Researchers can get insights into numerous elements of customer complaints by examining this dataset. They can study the number and distribution of complaints across various goods and services, find common concerns or patterns, investigate resolution timetables, and even assess consumer emotion indicated in complaint narratives.

Accessing and using the customer Complaint Dataset can help the research community better understand customer experiences, shape regulatory choices, and drive business best practices. Researchers may get the dataset and associated documentation on Kaggle by clicking on the following link:

<https://www.kaggle.com/datasets/utkarshx27/consumer-complaint>

1.2 Objectives

The purpose of this research is to highlight the enormous usefulness of the Consumer Complaint Dataset as a tool for businesses to obtain insights, increase customer happiness, and drive corporate success. The study intends to contribute to the field of consumer-centric business practices by giving practical advice based on data analysis. The study intends to accomplish the following goals by evaluating this dataset:

- **Identify Common Pain Points:-** Attempts to find the most frequent pain points indicated by customers in their complaints through extensive analysis. Understanding these frequent concerns allows businesses to handle them more proactively, resulting in better goods and services.
- **Benchmarking Against Competitors:-** The purpose is to examine and assess the performance of various organizations in managing consumer complaints. Companies can obtain a competitive advantage and find areas for development by examining complaint trends among rivals.
- **Enhance Customer Service: -** The goal of the study report is to give practical ideas for bettering customer service. Companies can discover areas where customer service can be improved by reviewing complaint data, leading to increased satisfaction and loyalty.

- **Insights into Consumer Behavior:-** The Consumer Complaint Database allows businesses to acquire useful insights into consumer behavior. Companies may learn the typical pain points and disappointments mentioned by consumers by examining the dataset. This insight enables them to modify their products and services to better fit the demands and preferences of their customers.
- **Regulatory and Policy Insights:-** The findings from the Consumer Complaint Database can be used to guide regulatory and policy choices. Policymakers can design effective legislation to safeguard consumer interests by recognizing frequent consumer concerns. These data may also be used by businesses to proactively align their activities with regulatory requirements and industry standards.

1.3 Meeting Record:

Meeting no	Meeting Date	Meeting location	Names			
			Alvina	Siddharth	Aditi	Sriram
1	04-20-2023	In class	P	P	P	P
2	04-27-2023	In class	P	P		P
3	05-02-2023	Library	P		P	
4	05-03-2023	Zoom	P	P		
5	05-04-2023	In class	P	P		P
6	05-08-2023	Library	P	P		P
7	05-09-2023	Zoom	P	P		P
8	05-13-2023	Library	P	P	P	P
9	05-14-2023	Library	P	P	P	P

Data Source

Brief Description

The Kaggle-accessible Consumer Complaint Database is a helpful resource for businesses, analysts, and academics for various reasons:

- **Financial sectors:** The database includes complaints made against Financial Products and Services.
- **Well-Organized Dataset (Uncleaned Data):** The dataset is well-structured and includes a variety of relevant information such as complaint ID, date received, product, issue, customer complaint narrative, corporate public response, company, state, ZIP code, and more. This extensive dataset contains a wealth of information for in-depth research and exploration.
- **Data-Driven Decision Making:** Organizations may uncover common issues and patterns by examining the information, allowing them to make data-driven decisions to increase customer satisfaction. Understanding the most common consumer pain points and concerns enables firms to solve these issues and improve their goods, services, and overall customer experience.

- **Customer Relationship Development:** Leveraging the Consumer Complaint Database allows businesses to foster stronger relationships with their customers. By actively listening to customer complaints and addressing their concerns, companies can demonstrate their commitment to customer satisfaction, building trust and loyalty.
- **Opportunities for Research and Analysis:** The database is a useful resource for researchers and analysts, providing chances to conduct studies and develop insights. Exploring correlations, patterns, and interactions between variables can result in significant findings that contribute to academic research and industrial expertise.

In conclusion, the Kaggle customer Complaint Database is a helpful resource that provides a well-organized and comprehensive dataset of customer complaints from many businesses. Its availability helps enterprises, analysts, and researchers to gather insights, enhance goods and services, and strengthen customer connections by gaining a better grasp of consumer requirements and preferences.

Elements

The elements of the Consumer Complaint Database include:

- Complaint ID
- Date Received
- Product
- Sub-product
- Issue
- Sub-issue
- Consumer Complaint Narrative
- Company Public Response
- Company
- State
- ZIP Code
- Tags
- Consumer Consent Provided
- Submitted via
- Date Sent to Company
- Company Response to Consumer
- Timely Response
- Consumer Disputed

This consists of a variety of customer complaints, corporate answers, and associated information. They provide useful information on the kind of complaints, product categories, geographic distribution, and consumer feedback. These columns may be analyzed to uncover patterns, trends, and opportunities for improvement in goods, services, and customer experiences.

Data Cleaning

Data Cleaning Process:

1. **Loading the Dataset:** The dataset was loaded into a pandas DataFrame using the `pd.read_csv()` function.
2. **Exploring the Dataset:**
 - The shape of the DataFrame was examined to understand the size and dimensions of the dataset.
 - The first 50 rows of the DataFrame were displayed to gain initial insights into the data structure.
3. **Handling Missing Values:**
 - Missing values in categorical columns were filled with the text 'Missing' using the `fillna()` function.
 - Any remaining rows with missing values were dropped using the `dropna()` function.
4. **Data Type Conversion:**
 - The 'Date received' and 'Date sent to company' columns were converted to datetime format using the `pd.to_datetime()` function.
 - This conversion enables easier manipulation and analysis of date-related information.
5. **Creating a New Column:**
 - A new column named 'new_col' was created based on a condition that checked if the 'Date received' was the same as 'Date sent to company'.
 - The resulting column indicated whether the dates matched or not, providing potential insights into the complaint handling process.

6. Dropping Unnecessary Columns:

- The 'Date sent to company' & 'new_col' columns were dropped from the DataFrame as they were deemed unnecessary for further analysis.

7. Cleaning ZIP Code:

- The 'ZIP code' column was converted to a string data type using the `astype()` function.
- Any decimal points present in the ZIP codes were removed using the `str.replace()` method.
- Leading zeros were added to ensure all ZIP codes had a consistent 5-digit format using the `str.zfill()` method.

8. Data Visualization Libraries:

- The `matplotlib`, `pyplot` and `seaborn` libraries were imported.
- These libraries are typically used for data visualization purposes, allowing for the creation of various charts, plots, and graphs.

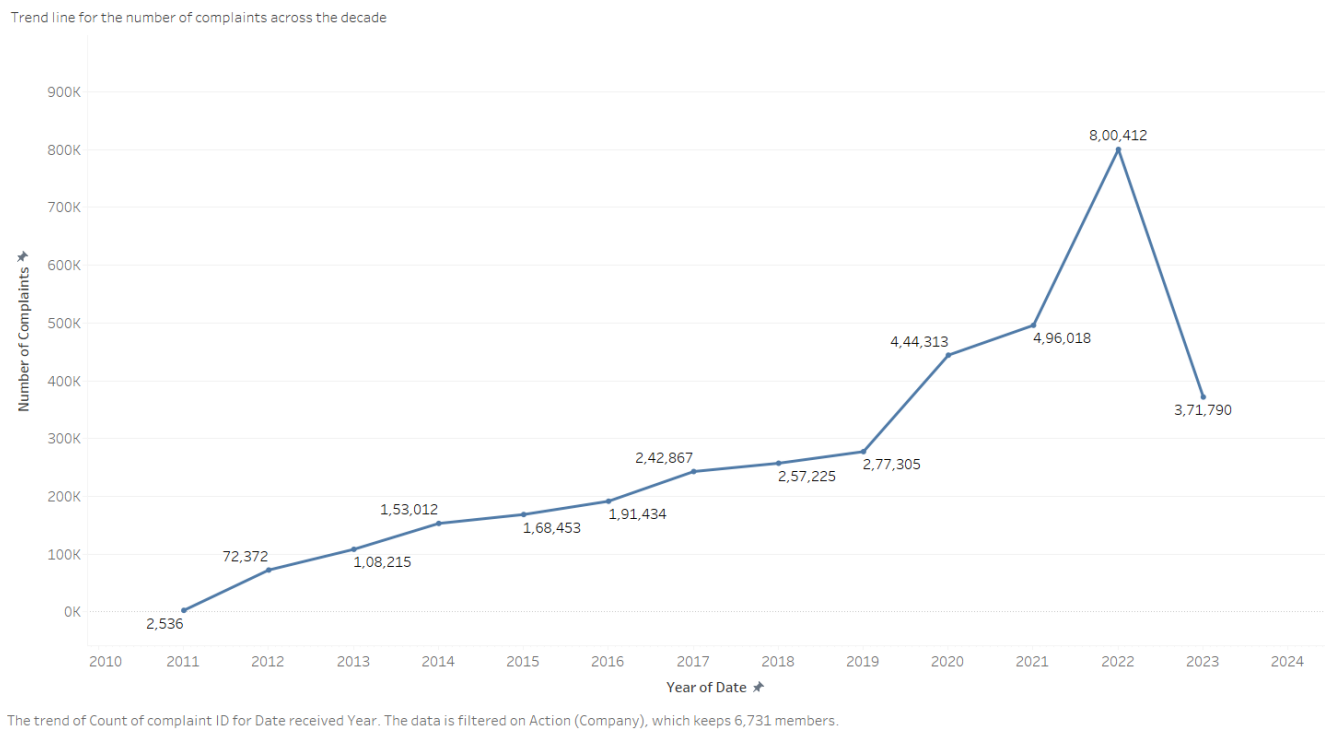
The data cleaning procedure for the Consumer Complaint Database included numerous critical processes, such as managing missing values, switching data types, establishing new columns, and deleting extraneous information. By following this procedure, the dataset has been prepped for future analysis, guaranteeing that it is correct and ready for the extraction of useful insights. The clean dataset provides a strong platform for researchers, analysts, and companies to look into consumer complaints, uncover patterns, and make data-driven decisions to improve customer happiness and effectively resolve issues.

Dashboard

4.1 Description

Dashboard-1

Upward Trend in the Number of Complaints Over the Last Decade



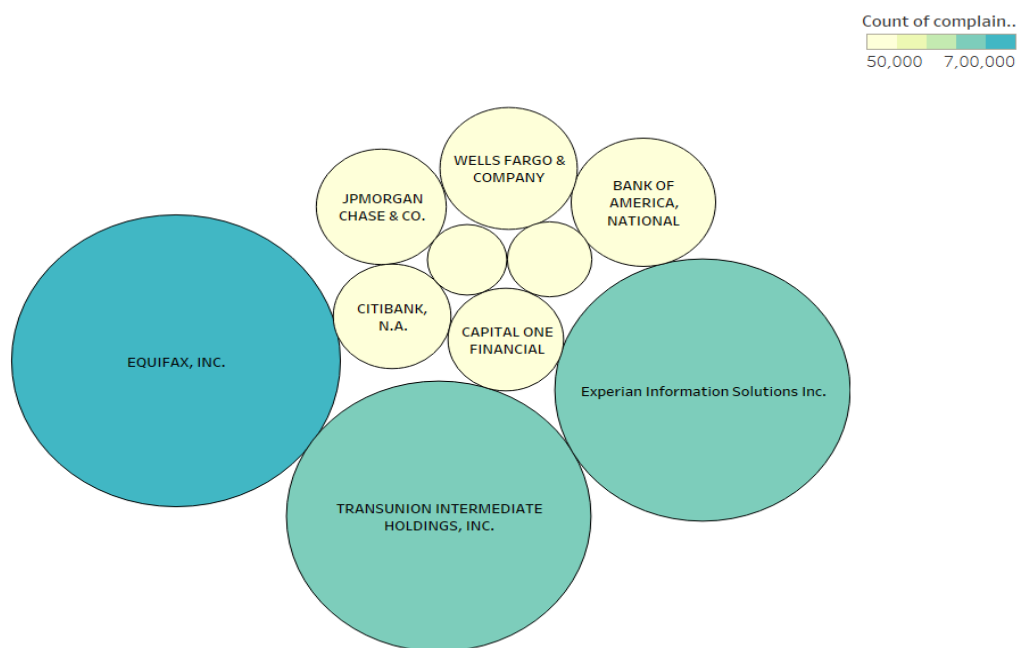
Complaints Across the Decade **trendline visualization** depicts a decade's worth of consumer complaints from 2011 to 2023. This graphic offers useful insights into the shifting patterns and trends of complaint volume. The trendline shows that the number of complaints has fluctuated over time, with a significant increase from 2011 to 2022, peaking at 800,412 complaints. However, there was a considerable decline in complaints in 2023, with a total of 371,790 registered.

The research encourages additional investigation into the underlying variables causing these variations, such as changes in consumer behavior, changes in corporate policies, or external market dynamics. Understanding these characteristics can lead to higher customer satisfaction and more effective treatment of emerging problems.

Incorporating this trendline graphic into your research endeavor provides a brief yet comprehensive review of the complaint trends discovered over the decade. It improves comprehension of long-term patterns and changes in customer feedback, allowing for a thorough examination of the consumer complaint environment.

The Top 10 Companies with the Most Complaints

Top 10 Companies with most number of complaints



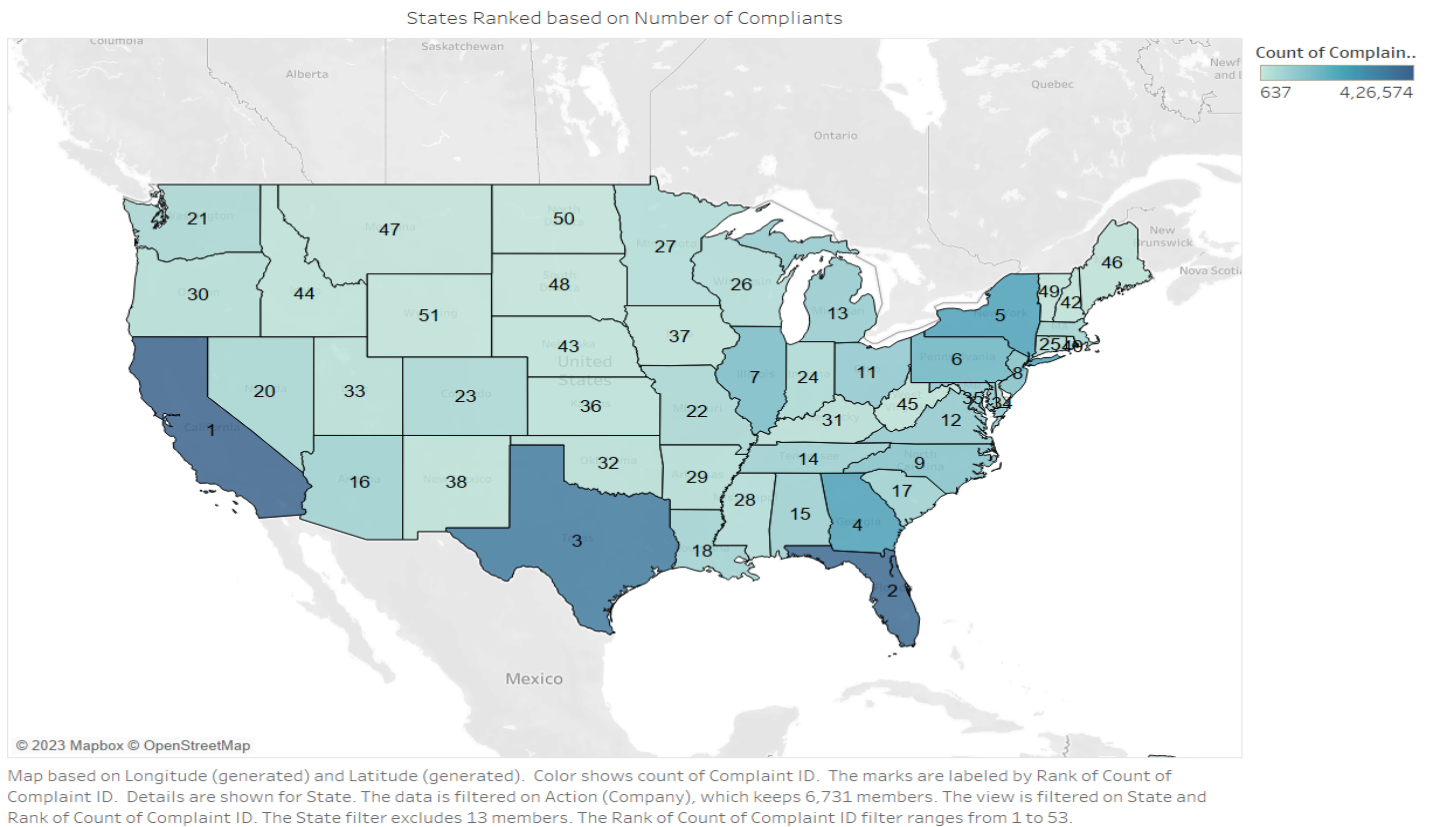
Company. Color shows Count of complaint ID. Size shows % of Total Count of Complaint ID. The marks are labeled by Company. The view is filtered on Company, which keeps 10 of 6,731 members.

The **bubble pack visualization** depicts the complaint volumes as well as the proportional proportions of the top ten corporations. This allows for the quick identification of organizations with the highest complaint counts and provides significant insights for decision-making and improvement initiatives to efficiently handle consumer issues.

- Equifax, Inc.: Equifax, Inc. has the most complaints among the top ten companies, accounting for 28.05% of the total number of complaints with 651,105. This indicates a high level of dissatisfaction with Equifax's services.
- TransUnion Intermediate Holdings, Inc.: TransUnion ranks second with 558,665 complaints, accounting for 24.07% of all complaints. The large number of complaints indicates a significant level of customer dissatisfaction with TransUnion's offerings.
- Experian Information Solutions Inc. ranked third with 525,458 complaints, accounting for 22.64% of all complaints. This shows a significant volume of consumer complaints against Experian's services.
- Bank of America, National Association: While Bank of America receives less complaints than the top three firms, it still receives a large amount of them. With 125,668 complaints, it accounts for 5.41% of the total.

- Wells Fargo & Company has 113,591 complaints, accounting for 4.89% of all complaints. Despite being lower than the top three corporations, this nevertheless implies a significant amount of customer unhappiness.
- JPMorgan Chase & Co.: JPMorgan Chase has 101,800 complaints, accounting for 4.39% of all complaints. This indicates a moderate level of consumer dissatisfaction with the company's services.
- Citibank, N.A.: Citibank has 83,488 complaints, accounting for 3.60% of all complaints. While it receives fewer complaints than some other companies, it still receives a large amount of them.
- Capital One Financial Corporation: With 80,545 complaints, Capital One accounts for 3.47% of all complaints. This implies that consumers are moderately dissatisfied with Capital One's services.
- Synchrony Financial has 43,042 complaints, accounting for 1.85% of the total number. Synchrony Financial's complaint volume is low in comparison to the other top 10 corporations.
- Navient Solutions, LLC: Navient Solutions has 38,067 complaints, accounting for 1.64% of all complaints. It has fewer complaints than the other top ten companies.

Unveiling Consumer Complaints-Based State Rankings



We discovered surprising insights though **mapping** about the number of complaints submitted across different states after delving into the dataset. The rankings focus light on the states that have seen a high level of consumer complaints. Let us look more closely at the findings:

- California (CA) takes the top spot, with an astounding total of 426,574 complaints. This suggests a significant uproar from consumers within the state, making it a top priority.
- Florida (FL) comes in second place, not far behind California, with 404,922 complaints. Residents of the state have expressed their

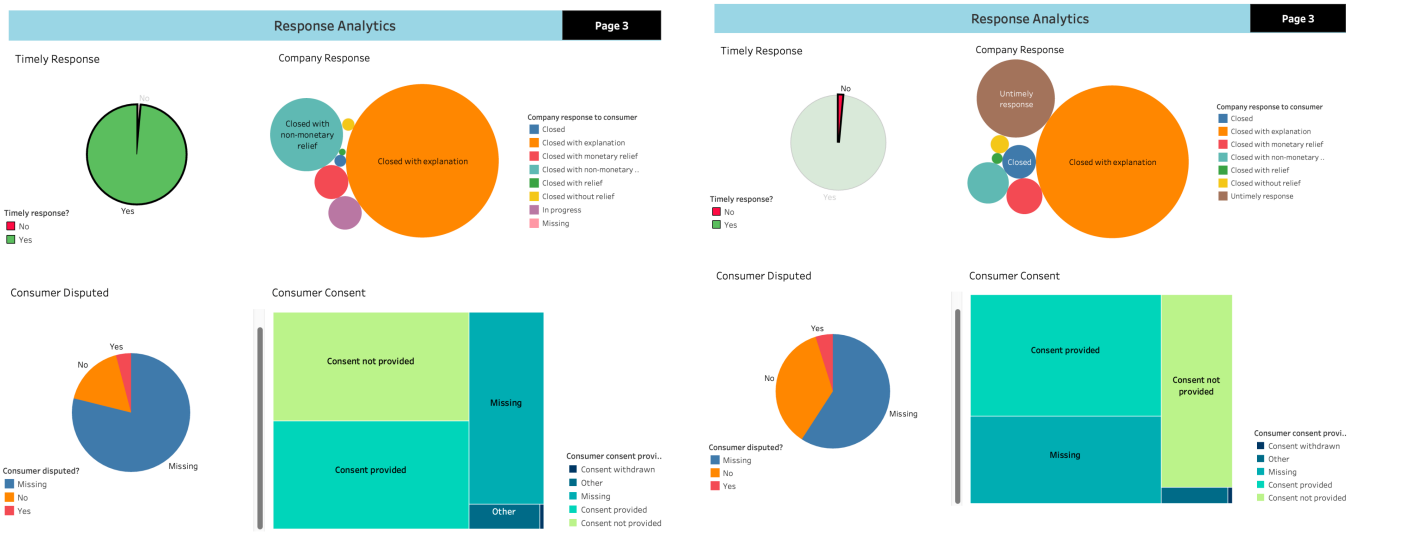
unhappiness, indicating the need for focused emphasis on addressing their concerns.

- Georgia (GA) takes the third rank, with a significant number of 234,105 complaints. Consumers in Georgia are clearly vocal about their complaints and seek redress.
- New York (NY) occupies the fourth position in the standings, with 231,777 complaints. Residents of the state have made their voices heard, highlighting the significance of addressing consumer issues.
- Texas (TX) closes out the top five, with a staggering total of 355,958 complaints. The Lone Star State is home to a sizable number of disgruntled customers who are looking for effective solutions to their problems.

These rankings offer an enthralling peek into the consumer complaint environment, highlighting states that are dealing with a large number of complaints. Understanding these regional tendencies allows firms to effectively allocate resources, prioritize remedies, and improve customer satisfaction. Businesses can take proactive efforts to reduce concerns, improve their offers, and develop a harmonious relationship with their valued clients by delving deeper into the underlying causes of these complaints.

Dashboard-2

Consumer Dispute



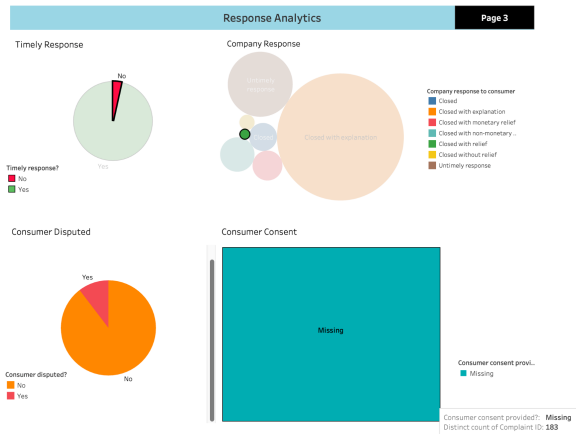
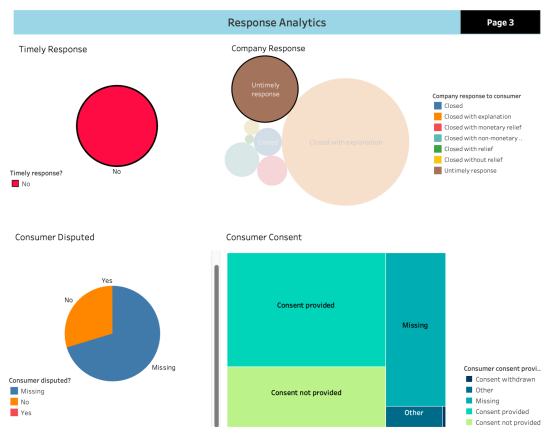
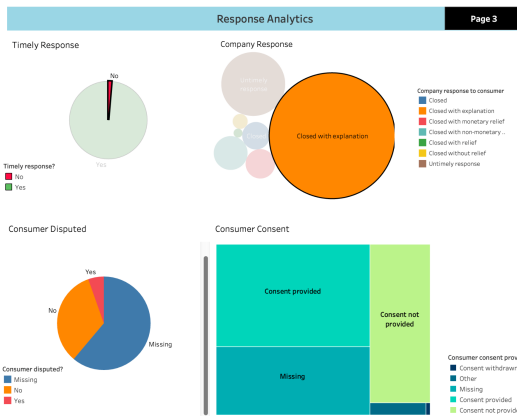
Based on the visualization of the **pie chart** and the data received from the tree map, the following conclusions can be drawn:

- Missing Consumer Disputed Status:** A missing consumer disputed status is present in the majority of complaints (78.57%). This suggests that for a major percentage of the dataset, information on whether the consumer challenged the complaint or not is not available.

- **No Consumer Challenge:** In approximately 17.29% of cases, the consumer did not challenge the complaint. This shows that a sizable number of customers accepted the company's resolution or did not see the need to take the issue further.
- **Consumer Dispute:** A tiny percentage of complaints (4.14%) show that the complainant disputed the complaint. This means that these customers were displeased with the company's resolution and chose to contest the complaint or seek additional action.
- The **large percentage of missing customer disputed** status emphasizes the difficulty in measuring the influence of consumer disputes on the complaint resolution process. In disputed cases, lacking data may impede the capacity to draw conclusive findings or judge the effectiveness of the resolution.
- **Consumer Perception:** The prevalence of consumer disputes, albeit in a reduced proportion, demonstrates that there are times when consumers feel compelled to contest the company's resolve. This emphasizes the significance of proper complaint management and the need for businesses to adequately address consumer complaints.

These findings provide an overview of the dataset's consumer disputed status and emphasize the need for additional analysis and exploration to better analyze the influence of disputes on complaint resolution and customer satisfaction.

Timely Response



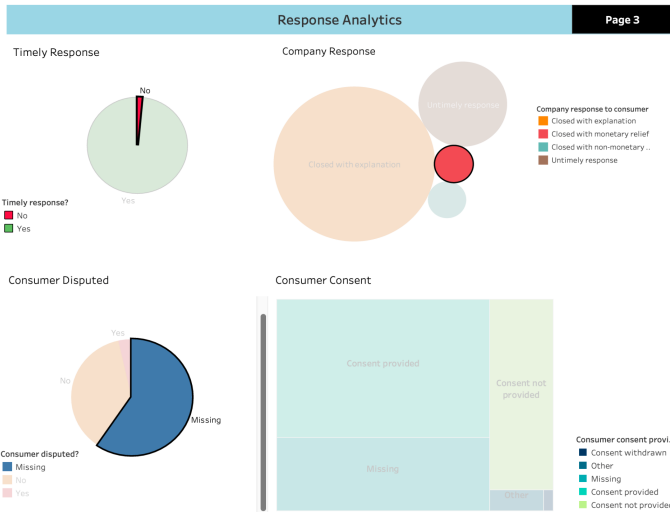
The following conclusions can be drawn from the pie chart representation and the data provided by clicking on it:

- Timely Response:** The majority of complaints (98.54%) received a "Yes" for a timely response, suggesting that the companies addressed the consumer complaints within the timeframe intended.

- **Delayed Response:** A tiny percentage of complaints (1.46%) scored a "No" for a timely response, suggesting that organizations did not respond promptly or within the specified timeframe in some cases.
- **Customer Satisfaction:** The high percentage of timely responses implies that businesses are proactive in responding to client complaints, which improves overall customer satisfaction.
- **Service Efficiency:** According to the data, the majority of consumer complaints were handled in a timely manner, suggesting that businesses emphasize efficient complaint handling processes and aim to fix issues as soon as possible.
- **Room for Improvement:** While the number of delayed replies is low, there is always room for improvement in responding to consumer complaints quickly. Companies might concentrate on lowering the number of complaints that result in response delays.

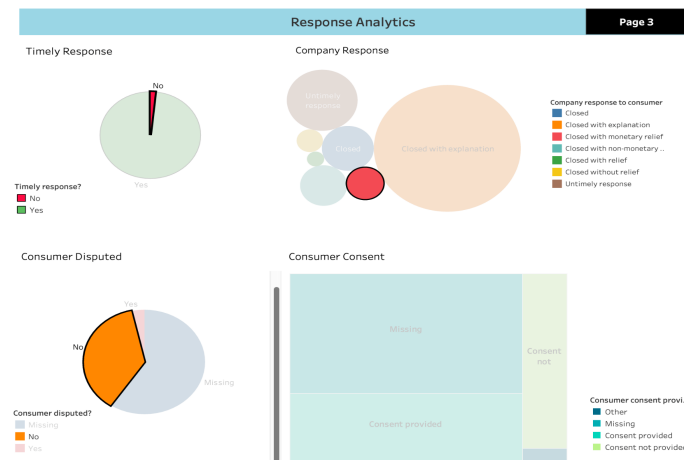
These findings shed light on the companies' reaction to consumer complaints and emphasize the necessity of prompt resolution in preserving customer happiness.

Company Response



Certainly! The packed bubbles representation of the "Company response to consumer" data, including the number of complaints for each category, is explained below

- "Closed with the explanation" (2,675,819 complaints):** This category has the most complaints, showing that businesses have provided explanations or resolutions for a large number of consumer concerns. It implies that businesses have made an attempt to answer consumer concerns and provide extensive explanations for their activities.



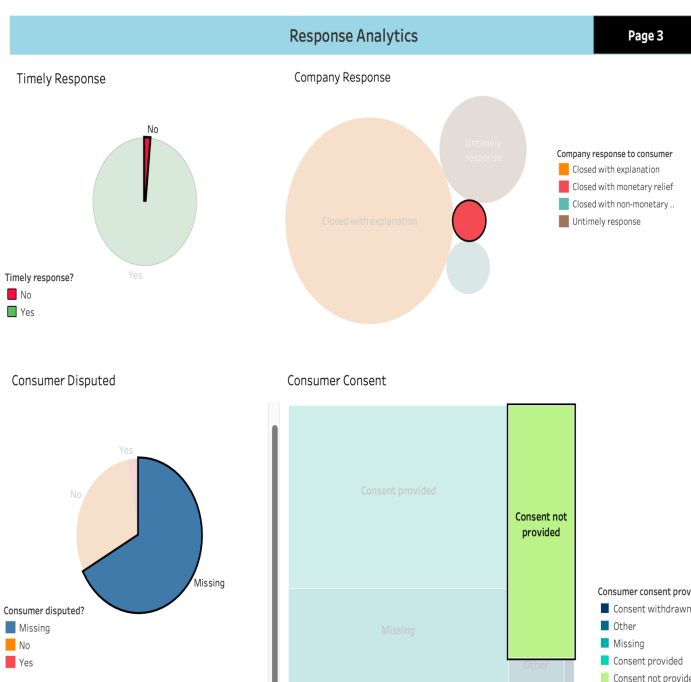
- "Closed with non-monetary relief" (602,986 complaints):** This category has a large number of complaints, showing that companies have handled a large number of consumer difficulties with non-monetary relief. This could involve activities like providing

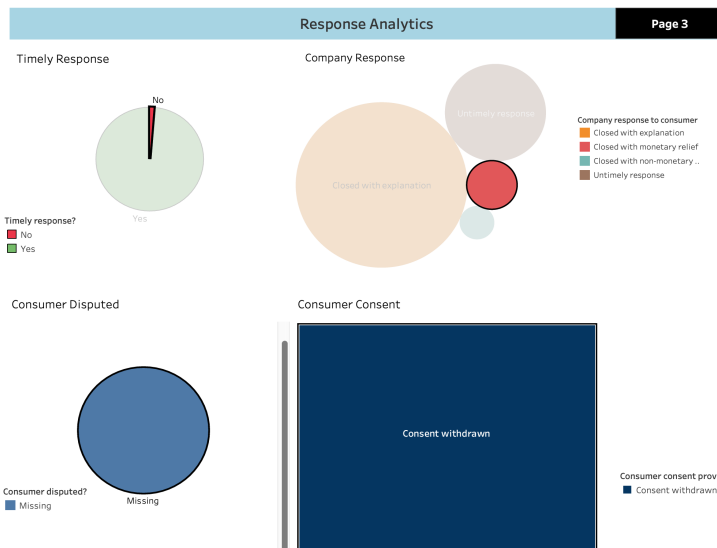
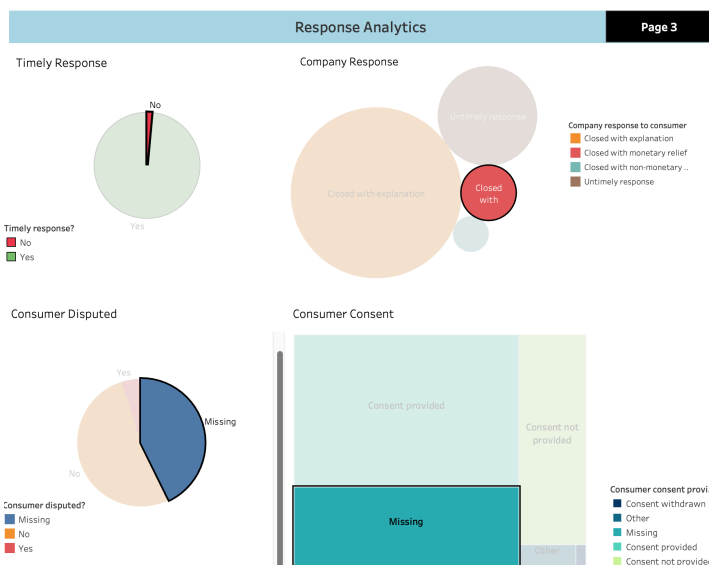
replacements, repairs, or other non-monetary remedies to consumer complaints.

- **"Closed with monetary relief" (131,403 complaints):** This category includes a significant number of complaints for which corporations paid monetary compensation or financial relief as a resolution. It implies that businesses have accepted responsibility and paid cash compensation to address consumer complaints.
- **"Closed without relief" (17,868 complaints):** This category contains a smaller number of complaints in which corporations closed the cases without offering any kind of relief or resolution to the consumers. It suggests that the companies did not appropriately address or resolve certain consumer complaints.
- **"In progress" (125,555 complaints):** This category shows many ongoing complaints that the companies are still addressing or investigating. These concerns have not yet been resolved, showing that the companies are working hard to fix them.
- **"Untimely response" (9,402 complaints):** This category indicates the number of complaints when organizations responded later than expected. It shows that some businesses have encountered delays in responding to consumer concerns.
- **"Missing" (4 complaints):** This category contains a relatively small number of complaints in which response information is missing or unavailable. These scenarios may necessitate additional inquiry or clarification to determine the causes of the missing data.

By considering the number of complaints in each category, the visualization provides a quantitative understanding of how companies have responded to consumer complaints. It offers insights into the volume of complaints addressed with explanations, non-monetary and monetary relief, unresolved cases, ongoing investigations, delayed responses, and missing response information. These insights can be valuable for assessing the effectiveness of company complaint resolution processes and identifying areas for improvement in addressing consumer grievances.

Consumer Consent





The following information is shown by the tree map visualization built using the "Consumer consent provided?" dataset:

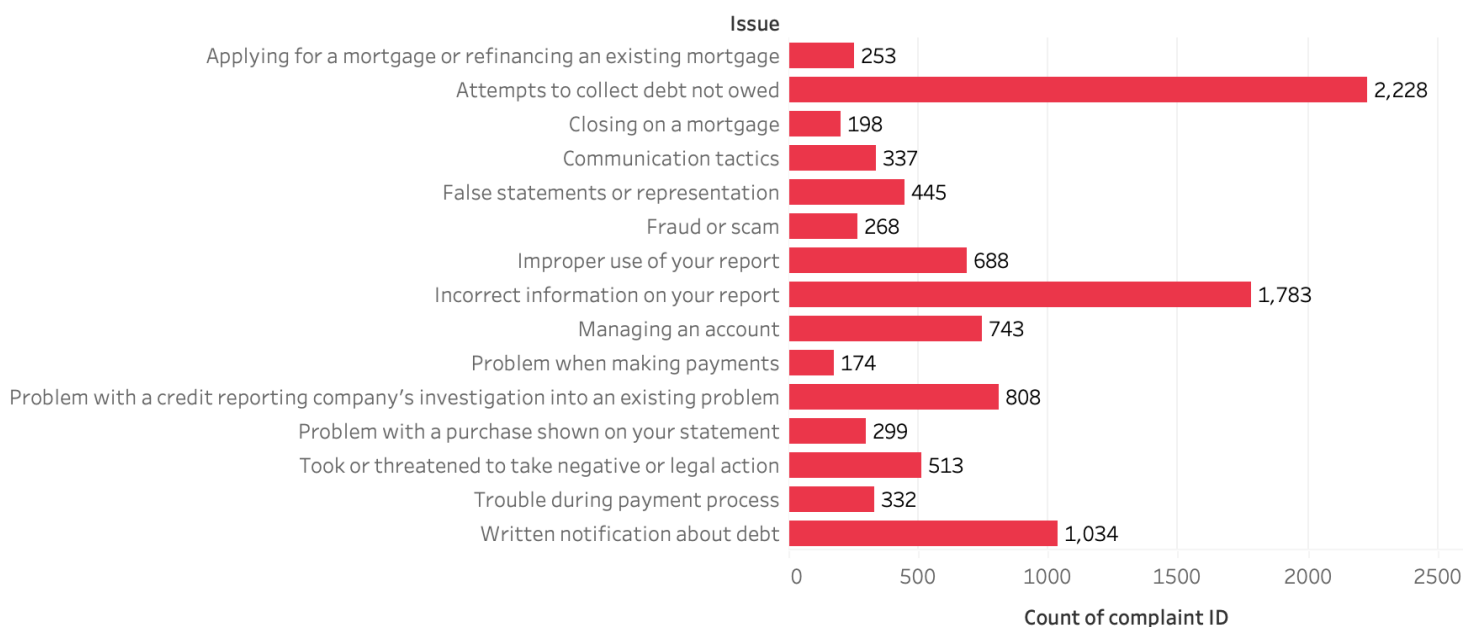
- **"Consent not provided" (1,296,407 separate complaints):** This group includes the most distinct complaints, showing that a considerable number of consumers did not consent to the use or sharing of their complaints.
- **"Consent provided" (1,292,959 separate complaints):** This group indicates a comparable amount of distinct complaints where consumers have explicitly consented for their complaints to be utilized or shared. It suggests that a sizable proportion of consumers are eager to disclose information about their complaints.
- **"Missing" (883,365 distinct complaints):** This category represents a substantial number of distinct complaints in which consent information is missing or unavailable. To understand the reasons for the missing data, further study or explanation may be required.

- **"Other" (106,836 distinct complaints):** This category covers a subgroup of distinct complaints where consumer permission is classified differently or comes under a different category. More information would be required to assess the particular nature of these concerns.
- **"Consent withdrew" (6,385 separate complaints):** This category indicates a relatively small number of distinct complaints in which consumers provided consent but then retracted it. It shows that a small number of consumers changed their minds about using or sharing their complaint information.

We have built this dashboard as every component of this dashboard is interlinked with each other and you will see how each and every component will contribute to our analysis. It provides a clear understanding of the distribution and proportions of customer consent statuses by presenting the data as a tree map. The scale and color coding of the tree map provide a visual representation of the different categories, making it easier to identify the largest and smallest segments. This visualization aids in understanding the prevalence of consent provided, consent not granted, missing consent information, and incidents where consent was withdrawn, allowing enterprises to make informed decisions about customer data handling and privacy concerns.

Dashboard-3

Top 10 client Issues that companies have not responded to for the last 3 years



We constructed a **horizontal bar chart** to depict the prevalence of various concerns and their respective timely response status. The graphic depicts the percentage of total complaint IDs for each concern. The following are the important findings:

- "Attempts to collect debt not owed" accounts for 17.10% of total complaint IDs. This shows a high volume of complaints about debt-collecting activities.

- With 13.68% of all complaint IDs, "incorrect information on your report" comes in second. This indicates a significant number of complaints about credit report mistakes.
- The percentages for "Written notification about debt" and "Problem with a credit reporting company's investigation into an existing problem" are 7.93% and 6.20%, respectively. These difficulties also account for a significant share of the complaints.
- Other issues with percentages ranging from 5.28% to 5.70%, showing a high number of complaints, include "Managing an account," "Improper use of your report," and "False statements or representation."

It is worth noting that all these issues had a "No" response for timely response, indicating that the complaints were not addressed on time. This visualization provides a clear overview of the distribution of complaint IDs for each issue and identifies areas where fast response needs to be improved.

Top 10 Products with the greatest number of complaints that are being ignored

Consumer Complaint Distribution by Product

The **tree map representation** provides a clear overview of consumer complaint distribution by product. The proportion of total complaints for each product category is represented by the size of each rectangle.

According to the data, the top three complaint categories were "Debt collection" (35.8% of all complaints), "Credit reporting, credit repair services, or

Top 10 Products with most number of complaints which are being ignored



other personal consumer reports" (25.0%), and "Credit card or prepaid card" (10.2%).

High-Impact Product Categories: Examining the number of complaint IDs inside each product category revealed more information about the most influential locations. The "Debt collection" category has the most complaint IDs (4,665), indicating a high volume of consumer complaints.

This was followed by "Credit reporting, credit repair services, or other personal consumer reports" (3,263) and "Mortgage" (931 complaint IDs).

These findings imply that these specific product categories demand additional attention as well as potential improvements in consumer experiences.

These findings shed light on the distribution and impact of consumer complaints across various product categories.

Further examination and analysis of the specific issues within each category can provide useful insights for addressing consumer complaints and improving customer happiness.

4. 2 Overview of Dashboard

Dashboard 1 Overview:

This dashboard represents two data visualizations that offer insightful information about consumer complaints over a decade are discussed in the article. The first graphic displays the trendline of complaints from 2011 to 2023, emphasizing the peaks and valleys as well as possible causes of the changes.

The second visualization gives a rapid overview of the proportionate magnitude of each organization's complaints by presenting the top ten organizations with the most complaints in a bubble pack format. These visualizations aid academics and decision-makers in spotting long-term trends and shifts in customer input, enabling more efficient problem-solving and greater client happiness.

Dashboard 2 Overview :

Dashboard 2 offers insights on three elements consumer dispute, timely response, and consumer consent

The bulk of complaints (78.57%) lack information on whether the customer challenged the complaint or not, according to a review of the data from the pie chart and tree map of consumer complaints. A tiny number (4.14%) of cases show that the consumer challenged the complaint, whereas the majority (around 17.29% of cases) show that the consumer did not contest the allegation. The lack of consumer dispute status highlights how challenging it is to gauge how conflicts affect the handling of complaints. But the existence of consumer disputes highlights the significance of good complaint management and the requirement for firms to effectively resolve consumer complaints. To fully understand how conflicts affect the handling of complaints and consumer satisfaction, more research is required.

The statistics and pie chart representation offer insights into how the businesses handled customer complaints. The majority of complaints (98.54%) got a response right away, indicating that companies are proactive in handling customer issues. This significant increase in timely responses suggests enhanced client satisfaction and operational effectiveness. Although there aren't many delayed responses, there is always room for improvement when it comes to immediately addressing customer issues. The results highlight the necessity for businesses to provide prompt resolution of customer complaints top priority in order to increase customer satisfaction.

The "Consumer consent provided?" dataset's tree map visualization provides a clear picture of the distribution and ratios of customer consent statuses. The majority of the groups, "Consent not provided," show that many customers did

not agree to the usage or dissemination of their complaints. Similar numbers of separate complaints are seen in the "Consent provided" group, indicating that many customers are open to disclosing their grievances. A large number of distinct complaints in the "Missing" group lack consent information, underscoring the need for additional research or justification. While "Consent withdrew" has a small number of distinct complaints indicating customers changed their minds about using or sharing their complaint information, the "Other" group includes a subgroup of distinct complaints where consumer permission is classified differently or falls under a different category. In general, the tree map representation helps with comprehension of the prevalence of given, withheld, and missing consent information, as well as instances where consent was revoked, enabling businesses to make informed decisions about handling consumer data and privacy issues.

Dashboard 3 Overview:

The data analysis presented in this dashboard provides insights into the prevalence and impact of consumer complaints across various product categories. A horizontal bar chart was used to depict the percentage of total complaint IDs for different concerns and their respective timely response status. The findings suggest that issues related to debt collection and credit reporting are the most prevalent, with a high volume of consumer complaints. Moreover, all issues highlighted in the chart had a "No" response for timely response, indicating the need for improvements in addressing consumer complaints. The treemap representation provides a clear overview of consumer complaint distribution by product, with debt collection, credit reporting, and a credit card or prepaid card being the top three complaint categories. Further examination of specific issues within each category can

provide useful insights for addressing consumer complaints and improving customer satisfaction.

4.3 Visualization

Throughout our team's project, we used a variety of visualization approaches to effectively evaluate and show the data. These visualizations were critical in identifying patterns, trends, and insights in the consumer complaints dataset. Here are the visualizations that we used:

The trendline analysis was used to discover long-term trends and patterns in the data. We were able to visualize the overall direction and size of changes over time by fitting trendlines to the data points. This study assisted us in understanding how consumer complaints have changed over time and whether there are any discernible trends or movements.

Bubble Pack: We utilized bubble pack charts to investigate hierarchical links within the data. Using this visualization technique, we were able to display several variables with varying sizes of bubbles. Each bubble's size indicated the magnitude or value of the variable, allowing us to compare and study several variables at the same time. This visualization was especially helpful in discovering correlations and similarities between various complaint categories.

Maps: Understanding the geographical distribution of consumer complaints was aided by geospatial visualization using maps. We were able to discover hotspots or places with a high number of complaints by mapping complaint data onto regions or countries. This visualization tool revealed geographic

patterns and differences in consumer complaints. It also aided us in determining the effect of location on the sorts and frequency of complaints.

Horizontal bar charts were used to compare the frequency or magnitude of different complaint types. We visualized the relative magnitude and frequencies of the complaints by portraying each category as a horizontal bar. This method made it simple to identify the most common or major complaint categories. It also made it easier to compare different categories and identify areas that need more attention or improvement.

We were able to show the data in a visually pleasing and instructive manner by utilizing various visualization tools as a team. Each visualization served a distinct purpose, adding to our overall comprehension of the consumer complaints dataset. The usage of visualizations promoted good communication and cooperation within our team, allowing us to draw relevant insights and make data-driven decisions.

Tools and Technologies

The following tools and technologies are being used based on the current code and your statement of utilizing Python and Tableau:

a) Python

Python is utilized as the programming language for data cleaning, preprocessing, and analysis, as indicated by the code excerpt. Python is a popular data analysis language with many modules and packages for processing data and performing various analytical tasks.

b) Pandas

The code imports the Pandas module, a sophisticated Python data manipulation and analysis tool. Pandas offers data structures and functions for efficient data processing, including reading and writing data, cleaning data, filtering, grouping, and aggregating.

c) Numpy

The code includes the NumPy library, which is an essential Python module for scientific computing. NumPy provides efficient arrays and numerical operations for large-scale mathematical and statistical computations.

d) Matplotlib and Seaborn

The code includes the Matplotlib and Seaborn libraries, which are used in Python for data visualization. Matplotlib provides a versatile and extensive collection of plotting tools, but Seaborn adds statistical visualization features that make it easier to generate useful and visually appealing graphs.

e) Tableau

While the code excerpt contains no Tableau-specific actions, you stated utilizing Tableau as a data visualization tool. Tableau is a robust data visualization software that makes it possible to create interactive and aesthetically pleasing representations and dashboards. It has an easy-to-use interface and a variety of data connections for importing and visualizing data.

f) Excel

Excel will be utilized to store and manage the information because it is a convenient tool for organizing and analyzing tabular data. It supports the simple import and export of CSV files, making it ideal for data storage and manipulation activities. Sorting, filtering, and formulae in Excel will help with data cleansing, transformation, and basic exploratory analysis.

g) PowerPoint

PowerPoint will be utilized to generate dynamic and visually appealing presentations that will highlight the project's findings and insights. It has an easy-to-use interface for producing slides with text, photos, charts, and other visual elements. Slide transitions and animations, for example, enable the production of interesting and impactful presentations that effectively communicate the evaluated data and essential points.

As a result, the present code use Python (with libraries such as Pandas, NumPy, Matplotlib, Seaborn, and Stats models) for data manipulation, analysis, and modeling. Tableau is used for data visualization and building interactive dashboards. Together, these tools enable efficient data processing, statistical analysis, class imbalance handling, and the construction of visually appealing visualizations to draw insights and effectively present findings.

Target Audience

The audience for your research will differ depending on the setting and goal of your investigation. However, given that you are working on the Consumer Complaint Database and hoping to obtain insights to enhance financial goods and services, the following audience groups may find your work relevant:

- **Financial Companies and Institutions:**

Banks, credit card firms, mortgage lenders, insurance companies, and other financial organizations are included. They may profit from your study by finding frequent consumer pain points and difficulties, which will allow them to enhance their goods, services, and customer happiness. Your insights may be useful in making decisions about

- **Regulatory and Policy-making Bodies:**

Government authorities, consumer protection groups, and financial regulatory bodies can use your study to better understand consumer issues and tendencies. This data can help them develop laws, regulations, and guidelines to safeguard consumer interests and promote fair practices in the business.

- **Market Researchers and Analysts:**

Market research, consumer behavior analysis, and competitive analysis professionals may find your work useful. They may use the insights gained from your research to better understand market trends, customer preferences, and the competitive environment across various financial products and services

- **Academics and Researchers:**

Our study and analysis can help academic studies and research in finance, consumer behavior, data analysis, and data visualization. Using the Consumer Complaint Database, researchers may expand on your work, confirm findings, and investigate new research issues.

- **General Public**

While not the intended target, the results of your investigation may be of interest to the wider public. Increased openness and understanding of customer complaints in the financial industry can enable people to make more educated judgments when purchasing financial products or services. It might also be used as a forum for customers to express their complaints and push for better industry practices.

To successfully reach and engage your target audience, it is critical to analyze your study's unique aims and objectives and customize your communication and distribution techniques appropriately.

Project Takeaway

- **Collaborative Effort:** Our team effectively collaborated to achieve the project's objectives. Each team member provided their unique talents and expertise, allowing us to efficiently tackle various elements of the project. The team's cohesiveness was critical in generating a thorough end product.
- **Effective Data Cleaning and Analysis:** We ensured the dataset's accuracy and quality through meticulous data cleaning and preparation. To get important insights from the data, our team used a variety of technical talents, including data visualization, regression analysis, and correlation analysis. This comprehensive analysis created a solid foundation for the project's outcomes.
- **Project Management and Planning:** Our team's painstakingly developed project plan served as a roadmap for the project's effective execution. We assigned tasks, established deadlines, and tracked progress to guarantee a seamless workflow and on-time completion. This methodical approach helped us to collaborate effectively and create a well-structured final report.
- **Impactful Tableau representations:** Using Tableau, we created informative representations that improved the comprehension and interpretation of the data. The horizontal bar charts and other visual components represented the findings clearly, making it easier for stakeholders to understand the information given.

- **Continuous Learning and Improvement:** We had a growth attitude throughout the project and explored opportunities to broaden our knowledge and abilities. We experimented with several methodologies, such as predictive regression and exploratory data analysis, and recognized Tableau's worth as a sophisticated data visualization tool. This agility and eagerness to learn aided our professional development and the project's overall success.
- **Effective Communication and Coordination:** Effective communication and coordination among team members were critical for overcoming hurdles and staying on track with the project. Even though we were faced with unanticipated situations, we kept open lines of communication, resolved concerns quickly, and supported one another to guarantee project milestones were accomplished.

In conclusion, the collaborative attitude of our team, excellent data analysis, precise planning, impactful visualizations, continual learning, and efficient communication were critical components in the project's success. We generated a high-quality final deliverable that met the project objectives and gave significant insights to Data Scientists as a result of our collaborative efforts.

Citation

Dataset: Utkarshx27. (May 01, 2023). Consumer Complaint Database. Kaggle.

Retrieved from: <https://www.kaggle.com/datasets/utkarshx27/consumer-complaint>