**Consumer Complaints - Case Study**

**Group C**

**SUBMITTED BY: Umair Khalid**

**Group Members:**

**Aiman Mukhtar**

**Muhammad Waleed Yasin**

**Osama Yusuf Hassan**

**Umair Khalid**

**Ghazi Haider**

# **Introduction:**

We have selected the Consumer Complaints dataset which presents consumers' financial complaints. These complaints have been collected by the Consumer Financial Protection Bureau (CFPB).

The reason behind selecting this dataset is that we intend to determine how the analysis was performed on complaints, how it has impacted the business to grow and the pertaining factors that have been measured in case of complaints.

# **Stakeholders:**

1. Financial Institutions
2. (CFPB), US Agency

# **Background:**

CFPB is the Consumer Financial Protection Bureau, a U.S. government agency dedicated to making sure you are treated fairly by banks, lenders and other financial institutions.

The CFPB publishes complaints sent to companies for response in the Consumer Complaint Database after the company responds, confirming a commercial relationship with the consumer, or after 15 days, whichever comes first.

Complaints can give us insights into problems people are experiencing in the marketplace and help us regulate consumer financial products and services under existing federal consumer financial laws, enforce those laws judiciously, and educate and empower consumers to make informed financial decisions.

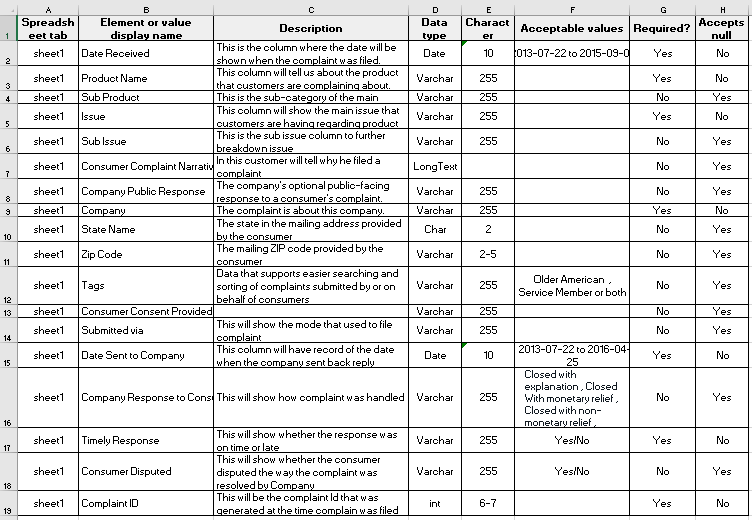
# **Tools:**

* MS Excel (For data dictionary)
* Lucid Chart (For ER-Diagram)
* Google Slides
* Google Sheets
* SQL (Structured Query Language)
* MySQL Workbench

# 

# **Data Dictionary:**

The consumer Complaints dataset has 18 attributes and Compliant ID is a Primary Key.



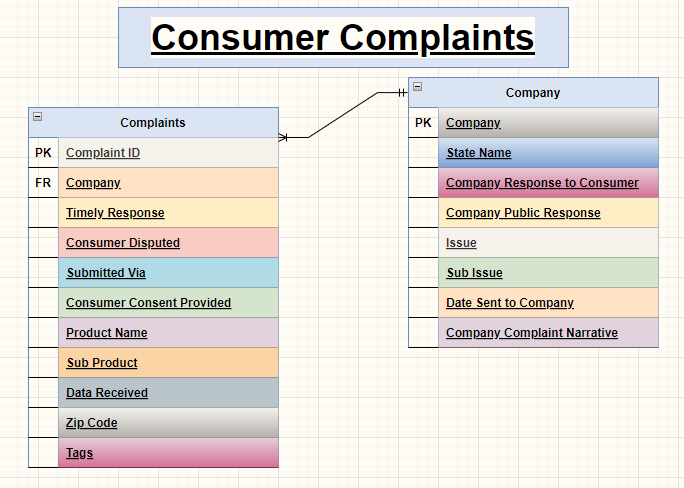
# 

# 

# 

# **Entity Relationship Diagram (ERD):**

We have divided our table into two to make an ERD (Entity Relationship Diagram). Otherwise, we have worked on a single table comprising 18 fields.



# 

# **Working:**

First of all, we have converted the **date fields** in the MS Excel file of our Consumer Complaint data set in **yyyy-mm-dd** format and removed spaces with underscores the in fields heading. This is because **MySQL Workbench** does not accept spaces in the attribute name. Then we saved it in xlsx format so that all changes remain intact. Since MySQL imports the data set in a CSV file format only, we again converted the xlsx file in CSV format.

We imported data in MySQL Workbench with all fields in text and **Complaint\_ID** in **bigint**. It took around 2 hours to upload the **Consumer Complaint** data set because it has 65500 rows. Once it was imported, we proceeded towards addressing the business problems.

# **Business Problems:**

1. **The customers lodge their complaints through various modes of submission. The research team decides to identify the most frequent mode of submission into ensure that the respective team is reasonably equippedThe given query is to identify the mode of submission adopted most frequently by customers.**

QUERY :

SELECT Submitted\_via,COUNT(\*) AS value\_occurrence

FROM project01.consumerdata01

GROUP BY Submitted\_via

ORDER BY value\_occurrence DESC

LIMIT 1;

RESULT :

****

1. **The research team decides to figure out the product whthat has received the most mplaints in a specific month of a specific year.**

QUERY :

SELECT Product\_Name, count(Product\_Name) AS value\_occurrence

FROM project01.consumerdata01

WHERE monthname(Date\_Received) = 'July'

AND year(Date\_Received) = '2013'

GROUP BY Product\_Name

ORDER BY value\_occurrence DESC

LIMIT 1;

RESULT :



1. **Once the research team has performed a month by month analysis, it proceeds on to perform an year by year analysis on each product and numbers of complaints associated with it.**

**The given query returns the list of products along with the number of complaints each product receives year by year.**

QUERY :

SELECT Product\_Name , count(Product\_Name) AS no\_of\_complaints

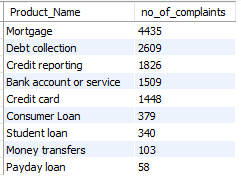
FROM project01.consumerdata01

WHERE year(Date\_Received) = '2013'

GROUP BY Product\_Name

ORDER BY no\_of\_complaints DESC;

RESULT :



1. **Once the research team has performed a month by month analysis, it proceeds on to ascertain the key issues and their number of occurrences in each year.**

**Given query returns the key issues ordered by their number of occurrences (from highest to the lowest) year by year.**

QUERY :

SELECT Issue, COUNT(Issue) AS value\_occurrence

FROM project01.consumerdata01

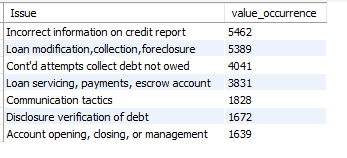
WHERE year(Date\_Received) = '2014'

GROUP BY Issue

HAVING value\_occurrence > 1000

ORDER BY value\_occurrence DESC;

RESULT :



1. **The research team decides to assess the complaints intensity for every product in each state.**

**The given query returns the number of complaints received by every state for each product separately.**

QUERY :

SELECT State\_Name, Product\_Name ,

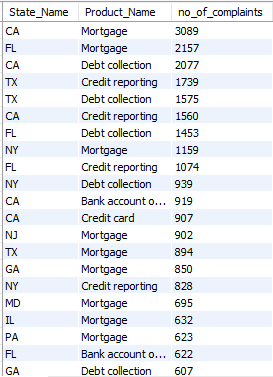
count(State\_Name) AS no\_of\_complaints

FROM project01.consumerdata01

GROUP BY State\_Name, Product\_Name

ORDER BY no\_of\_complaints DESC;

RESULT :



1. **The research team observes that it took unusually long for some complaints to be addressed. After an in-depth investigation it was discovered that it took several days for those complaints to be forwarded to their respective financial companies.**

**The given query returns the ID number of the top 5 complaints that took the longest to be forwarded to their financial companies once received. Also return the number of days before they were forwarded.**

QUERY :

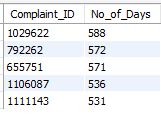
SELECT Complaint\_ID , datediff(Date\_Sent\_to\_Company , Date\_Received) AS No\_of\_Days

FROM project01.consumerdata01

ORDER BY No\_of\_Days DESC

LIMIT 5;

RESULT :



1. **The research team attempts to study the performance of each department.**

**Given query returns the consumer timely response rate.**

QUERY :

SELECT sum(CASE WHEN Timely\_Response = 'Yes' THEN 1 ELSE 0 END)

/count(\*) AS Timely\_response\_rate

FROM project01.consumerdata01;

RESULT :



1. **Now the research team has to evaluate the performance of the customer care team. It decides it will base its evaluation on customer complaints dispute rate. If the dispute rate is high it will have to work on improving the customer care team performance.**

**The given query returns that consumer complaints dispute rate.**

QUERY :

SELECT sum(CASE WHEN Consumer\_Disputed = 'Yes' THEN 1 ELSE 0 END)

/count(Consumer\_Disputed) AS Dispute\_Rate

FROM project01.consumerdata01;

RESULT :



1. **The team computes the consumer complaint resolution rate to assess the performance of customer care team**

**The given query returns the consumer complaint resolution rate.**

QUERY :

SELECT sum(CASE WHEN Consumer\_Disputed = 'No' THEN 1 ELSE 0 END)

/count(Consumer\_Disputed) AS Resolution\_Rate

FROM project01.consumerdata01;

RESULT :



1. **The research team wants to publish a list of the most responsive companies. The company with the most numbers of timely responses by their customer team will be part of the list.**

**The given query returns the names of the top 5 most responsive companies. Also, return the number of responses each of them has made.**

QUERY :

SELECT DISTINCT Company , COUNT(Timely\_Response) AS value\_occurrence

FROM project01.consumerdata01

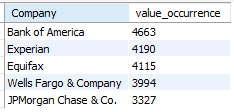
WHERE Timely\_Response = 'Yes'

GROUP BY Company

ORDER BY value\_occurrence DESC

LIMIT 5;

RESULT :



1. **Also, the research team decides to offer its expert advice to companies which have received most complaints for their products. This will help the team in identifying which companies need the most help.**

**The given query returns the top 5 companies with the most number of complaints. Also, return the number of complaints each of them has received.**

QUERY :

SELECT Company , COUNT(\*) AS no\_of\_complaints

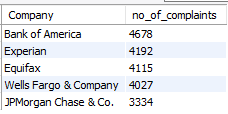
FROM project01.consumerdata01

GROUP BY Company

ORDER BY no\_of\_complaints DESC

LIMIT 5;

RESULT :



1. **The research team plans on identifying the product which has the most number of complaints associated with them. This will help them investigate any shortcomings related to those products so that they could mitigate the prevailing issues.**

**The given query returns the top 5 products with the most number of complaints. Also, return the number of complaints each of them has received.**

QUERY :

SELECT Product\_Name , Count(Product\_Name) AS No\_of\_Complaints

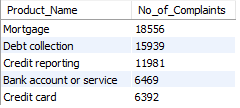
FROM project01.consumerdata01

GROUP BY Product\_Name

ORDER BY No\_of\_complaints DESC

LIMIT 5;

RESULT :



1. **Once the team has addressed the issues faced by products with the most number of complaints, it sets on to identify the product with the most disputes.**

**The given query returns the top 5 products with most disputes along with the number of occurrences.**

QUERY :

SELECT Product\_Name ,

sum(case when Consumer\_Disputed = 'Yes' then 1 else 0 end) AS No\_of\_disputes

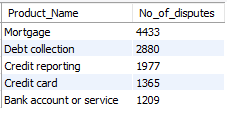
FROM project01.consumerdata01

GROUP BY Product\_Name

ORDER BY No\_of\_disputes DESC

LIMIT 5 ;

RESULT :



# 

# **Conclusion:**

* Complaints were filed by six different modes, namely via Web, Phone, Referral, Fax, Email & Postal mail.
* We have found that most complaints have been submitted via the Web.
* It has been found that maximum complaints are regarding Mortgage during the month of July of 2013.
* The State of California has received the most complaints.
* The State of Florida has received the most complaints regarding Mortgage and least complaints regarding bank accounts orsssssssse**.**
* It may take years to forward a complaint. We have extracted 5 complaint IDs that took more than a year to forward a complaint.
* Dispute Rate is around 20% which could be improved later. whereas, the complaint Resolution rate is 78%.
* Bank of America has received the maximum complaints and they respond timely.
* By Product: From 2013-2015, mortgages received the most complaints.
* By Company: From 2013 to 2015, Bank of America received the most complaints.
* By States: Most complaints were filed by California during the period 2013-2015.
* Florida was in second place ina number of complaints
* By Product: Mortgage had the highest number of disputes throughout 2013-2015.
* By Company: The number of disputes at Bank of America during the period 2013-2015 was the highest.
* Wells Fargo & Company was in second place.
* JPMorgan Chase & Co. was in third place.

# **Act:**

Based on the analysis of the CFPB complaints dataset from 2013 to 2015:

* Several companies should focus on California, Florida, and Texas since they contribute a large part of the complaints they receive.