# MTH 312: Project 6
# Predicting UDAAP: Analyzing Consumer Complaints for Financial Fairness

**Arnab Das**
(231080020)

**Gaurav Tomar**
(231080039)

**Soumen Konai**
(231080091)

**Jayant Vikash Jha**
(220479)

## Abstract

A credit card is a card from a bank that lets people buy things now and pay later. It is easy to use but sometimes banks treat customers unfairly. In this project, we study customer complaints against big US banks about unfair, deceptive, or abusive acts (UDAAP). We use real data from the Consumer Financial Protection Bureau (CFPB). First, we will see how the number of complaints has changed over time. We will find which banks have the fastest growth in complaints. We will use time-series analysis to show these trends. We will also build machine learning models to predict future risks. We will check if complaints rise when inflation or unemployment is high. By reading the complaint texts, we will find the most common problems for each bank. We will also see if bad complaints take longer to solve. Finally, we will study if government actions reduce complaints. Our project will help find hidden risks for customers and banks.

## Introduction

Credit cards are a common tool people use to buy things without paying right away. Banks give credit cards to customers with the promise that they will pay back the money later, often with some extra fees or interest. While credit cards make life easier, they can also cause problems if banks act unfairly. Sometimes banks hide important information, charge extra fees without warning, or make mistakes that harm customers. These are called UDAAP issues — Unfair, Deceptive, or Abusive Acts or Practices.

The Consumer Financial Protection Bureau (CFPB) collects complaints from people who face such problems with banks. By studying these complaints, we can better understand how banks treat their customers and how often unfair practices happen.

In this project, we will look at complaint data from the CFPB website. We will find out which banks get the most complaints and how complaint numbers change over time. We will use machine learning to predict future risks and study if outside factors like inflation affect complaints. We will also check if government actions, like fines or warnings, help to reduce bad behavior. Our goal is to spot patterns, find major complaint themes, and see how these problems affect customers. This study will give a full picture of credit card risks from a customer's point of view.

## Dataset Description

The dataset we are working with contains consumer complaints mainly related to credit cards and credit reporting. Each row represents a single complaint filed with regulatory authorities in the

United States. It includes details like the type of product and issue, a description written by the consumer (if available), and information about the company and location involved. Important fields are the dates when the complaint was received and sent to the company, the channel through which the complaint was submitted, the company's response, and whether the response was timely. Each complaint has a unique ID. Overall, the dataset gives a broad view of customer experiences with financial services and helps spot trends over time and across different locations.

## Data Preprocessing

We first imported the data using `read.csv("complaints.csv")`. The date columns, `Date.received` and `Date.sent.to.company`, were properly converted to Date format. Some columns were renamed to simpler names for easier coding, such as changing `Consumer.complaint.narrative` to `complaint_text`. We filtered out any rows where the complaint text was missing. Then, we cleaned the text by making everything lowercase and removing extra spaces. Categorical fields were changed into factors to make them easier to work with. Finally, we checked the structure and summary of the dataset to make sure everything looked correct before starting the analysis.

## TASK-1: UDAAP Complaint Trend Analysis

To investigate UDAAP-related complaints, we identified records containing "UDAAP" in either the `Issue`, `Sub.issue`, or `complaint_text` fields. Monthly counts were generated by extracting the month from each complaint's received date, then visualized in a time series graph with three-month intervals.

The data reveals three distinct periods:

- A stable baseline (2014-2022) with 5-20 monthly complaints
- A growth phase beginning late 2022
- A dramatic surge in 2024 exceeding 100 complaints/month

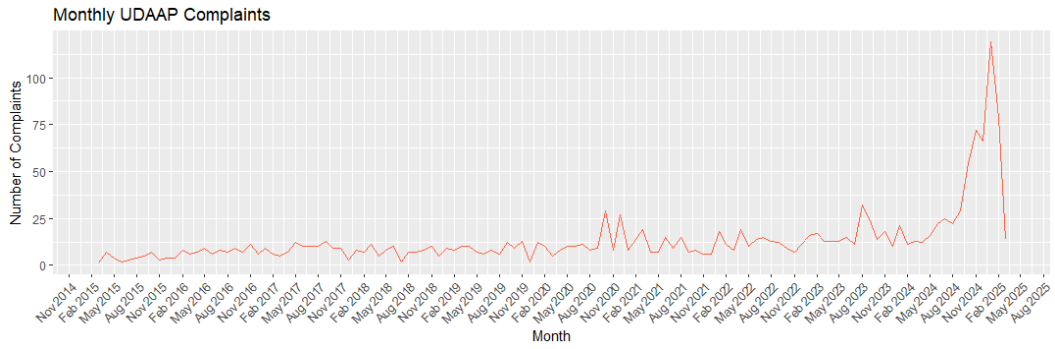


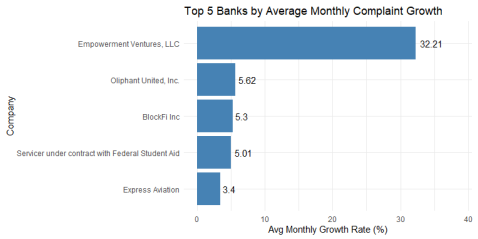

Figure 1: Monthly UDAAP Complaints

While early 2025 shows some moderation from peak levels, complaint volumes remain substantially higher than pre-2022 levels. This pattern suggests multiple possible drivers including increased consumer awareness, regulatory changes, or rising incidence of violations - particularly during the critical 2024 surge period that merits focused examination.

## TASK-2: Identifying Banks with Rising Complaint Trends

Our analysis of complaint trends revealed that Empowerment Ventures, LLC experienced the steepest increase in customer grievances among all financial institutions examined. The bank's complaint

volume showed a dramatic pattern - after peaking at over 600 complaints in a single month during 2015, numbers dropped sharply and stabilized at significantly lower levels post-2016.




| company | avg_growth |
|---|---|
| *<chr>* | *<dbl>* |
| Empowerment Ventures, LLC | 32.2 |
| Oliphant United, Inc. | 5.62 |
| BlockFi Inc | 5.30 |
| Servicer under contract with Federal Student Aid | 5.01 |
| Express Aviation | 3.40 |

Figure 2: Complaint Trends

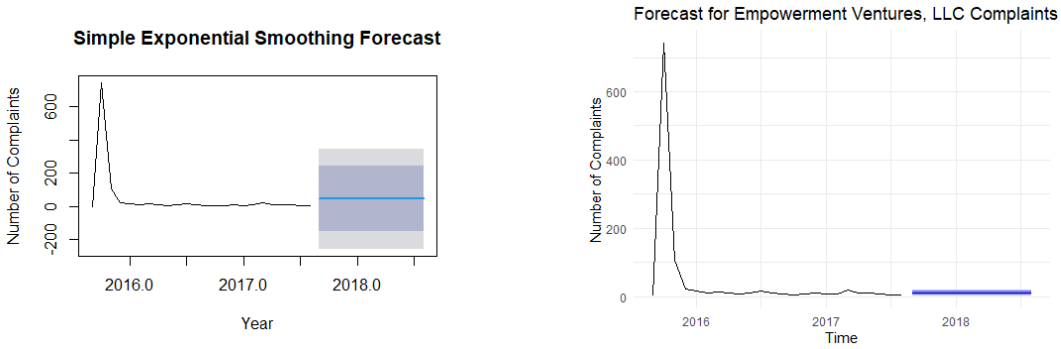



Figure 3: Forecasted Trends

| | Point Forecast | Lo 80 | Hi 80 | Lo 95 | Hi 95 |
|---|---|---|---|---|---|
| Sep 2017 | 9.600679 | 4.378979 | 14.82238 | 1.614779 | 17.58658 |
| Oct 2017 | 9.600679 | 4.378979 | 14.82238 | 1.614779 | 17.58658 |
| Nov 2017 | 9.600679 | 4.378979 | 14.82238 | 1.614779 | 17.58658 |
| Dec 2017 | 9.600679 | 4.378979 | 14.82238 | 1.614779 | 17.58658 |
| Jan 2018 | 9.600679 | 4.378979 | 14.82238 | 1.614779 | 17.58658 |
| Feb 2018 | 9.600679 | 4.378979 | 14.82238 | 1.614779 | 17.58658 |
| Mar 2018 | 9.600679 | 4.378979 | 14.82238 | 1.614779 | 17.58658 |
| Apr 2018 | 9.600679 | 4.378979 | 14.82238 | 1.614779 | 17.58658 |
| May 2018 | 9.600679 | 4.378979 | 14.82238 | 1.614779 | 17.58658 |
| Jun 2018 | 9.600679 | 4.378979 | 14.82238 | 1.614779 | 17.58658 |
| Jul 2018 | 9.600679 | 4.378979 | 14.82238 | 1.614778 | 17.58658 |
| Aug 2018 | 9.600679 | 4.378979 | 14.82238 | 1.614778 | 17.58658 |

Figure 4: Forecasted Points

This trend suggests the institution likely implemented effective corrective measures around 2015-2016, potentially through improved customer service, policy changes, or regulatory compliance efforts. Our forecasting indicates complaint levels should remain stable at these reduced rates, demonstrating how targeted interventions can successfully address customer satisfaction issues. The case of Empowerment Ventures serves as a notable example of complaint trend reversal following operational improvements.

## TASK-3: UDAAP Complaint Analysis

Our examination of unfair banking practice complaints revealed several important patterns:

The data shows a clear upward trend in complaints, with increasing seasonal variations in recent years. We found strong connections to economic factors - complaints rise sharply with unemployment (92% correlation) while decreasing moderately with inflation (55% correlation).

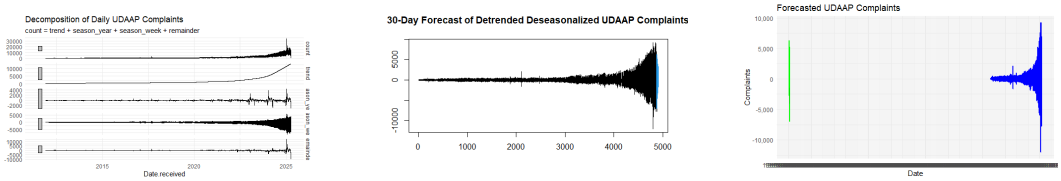



Figure 5: Forecasted Plots

Our forecasting models identified these patterns but faced challenges:

- Predictions become less certain further into the future
- Even after accounting for trends and seasons, complaint volumes remain volatile
- The best model still had significant forecasting errors

| | Unem_ploymentrate | CPIinflationrate | No. of Complaints |
|---|---|---|---|
| Unem_ploymentrate | 1.000 | -0.807 | 0.922 |
| CPIinflationrate | -0.807 | 1.000 | -0.554 |
| No. of Complaints | 0.922 | -0.554 | 1.000 |

Figure 6

```
Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
    Min      1Q  Median      3Q     Max
-7499.1   -72.6    -1.8    70.9  9810.0

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.567376   5.549835   0.463    0.644
z.lag.1     -2.169069   0.083012 -26.130  < 2e-16 ***
z.diff.lag1  0.931042   0.078230  11.901  < 2e-16 ***
z.diff.lag2  0.573216   0.067204   8.529  < 2e-16 ***
z.diff.lag3  0.309558   0.055587   5.569  2.7e-08 ***
z.diff.lag4 -0.009443   0.044257  -0.213    0.831
z.diff.lag5 -0.331921   0.033360  -9.950  < 2e-16 ***
z.diff.lag6 -0.580799   0.022718 -25.566  < 2e-16 ***
z.diff.lag7  0.013199   0.014901   0.886    0.376
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 387.1 on 4860 degrees of freedom
Multiple R-squared:  0.9263,    Adjusted R-squared:  0.9262
F-statistic:  7638 on 8 and 4860 DF,  p-value: < 2.2e-16


Value of test-statistic is: -26.1296 341.4446

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.43 -2.86 -2.57
phi1  6.43  4.59  3.78
```

Figure 7: Forecasting Model

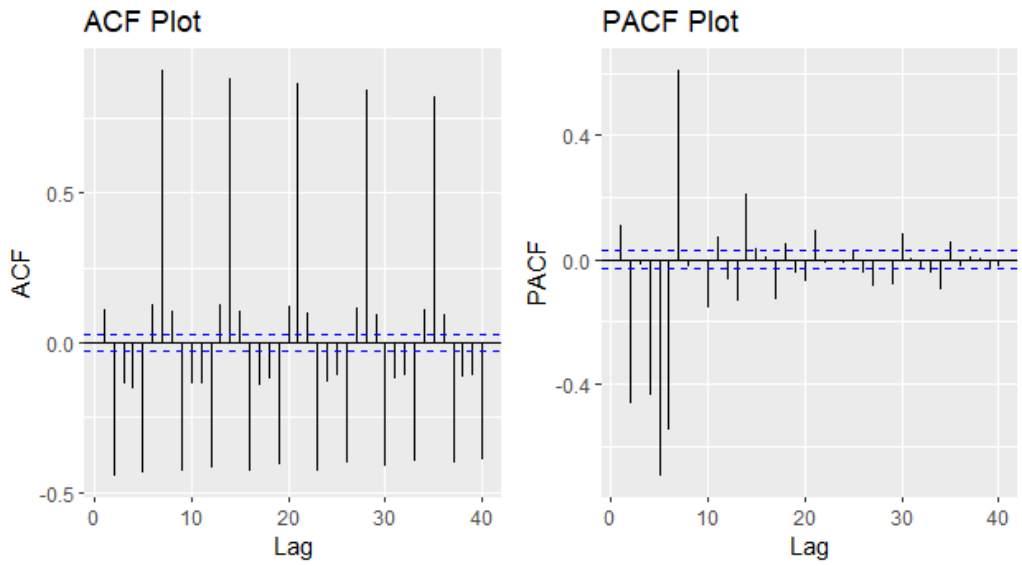



Figure 8: ACF & PACF Plots

While economic indicators help explain complaint patterns, the inherent variability makes precise predictions difficult. Future improvements might incorporate additional economic factors to enhance forecast accuracy.

## TASK-4: UDAAP Complaint Theme Analysis

We identified key themes in UDAAP-related banking complaints through text analysis of consumer reports. Our approach involved filtering complaints using terms like ”unfair” and ”deceptive,” then processing the text by removing irrelevant words, placeholders, and numbers.

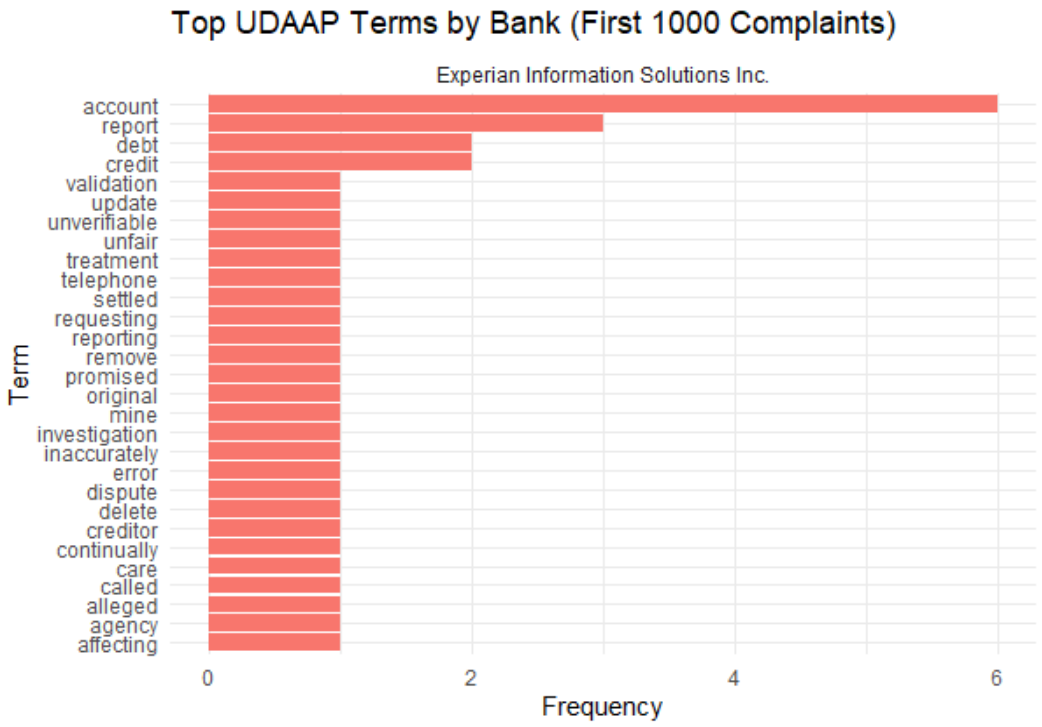



Figure 9: UDAAP Terms by Bank

The analysis revealed each bank's top complaint themes through word frequency analysis. We visualized these patterns in comparative bar charts showing the most common issues per institution. The results highlight:

- Specific problem areas for individual banks (like excessive fees or poor servicing)
- Common concerns shared across multiple institutions
- Variations in complaint severity between different banks

These findings help pinpoint where consumer protection efforts should focus, providing both banks and regulators with clear targets for improvement. The visual comparison makes it easy to see which issues affect which institutions and identify systemic problems in the banking sector.

## TASK-5: Comparative Analysis of Complaint Resolution Delays

This task focuses on analyzing the resolution delays in consumer complaints related to four leading U.S. banks—**JPMorgan Chase**, **Discover Bank**, **American Express**, and **Capital One**. The objective was to quantify how long these institutions typically take to resolve issues and identify recurring themes in complaints that experienced delayed resolution.

### Resolution Time Distribution

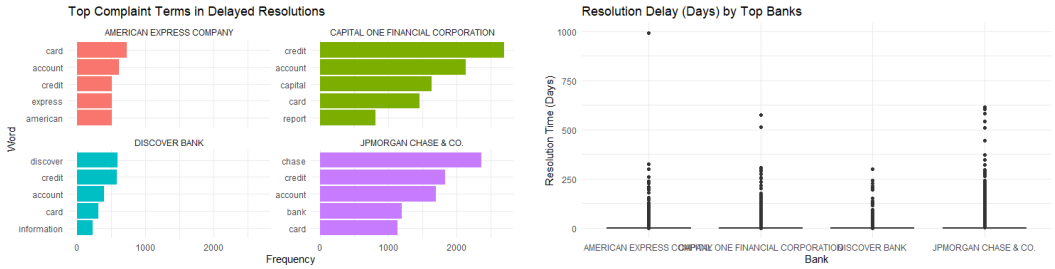



Figure 10: Resolution Time Distribution

To understand the variation in how long it takes for complaints to be resolved, we visualized the distribution of resolution times (in days) using a boxplot. The plot, titled *"Resolution Delay (Days) by Top Banks"*, highlighted long-tailed distributions across all four banks. Most complaints were addressed promptly; however, a subset of complaints exhibited extended delays, in some cases exceeding 500 days. Notably, **American Express** recorded a case with an extraordinary delay nearing 1000 days. While the median resolution time appeared relatively consistent across institutions, **JPMorgan Chase** and **Capital One** displayed broader variability, indicating inconsistency in processing times.

### Textual Analysis of Delayed Complaints

To explore the nature of complaints with prolonged resolution times, we isolated those that took more than 10 days to close. From these, the textual data in the complaint narratives was tokenized and cleaned to remove irrelevant terms. We then extracted the top five most frequently occurring terms for each bank. This helped surface dominant themes associated with delayed complaint handling.

| Bank | Frequent Complaint Themes (Top Words) |
|---|---|
| **JPMorgan Chase & Co.** | *chase, credit, account, bank, card* — highlighting general banking concerns and credit issues. |
| **Discover Bank** | *discover, credit, account, card, information* — suggesting difficulties around credit access and misinformation. |
| **American Express Company** | *card, account, credit, express, american* — primarily focused on disputes regarding cards and account maintenance. |
| **Capital One Financial Corp.** | *credit, account, capital, card, report* — indicating frequent credit reporting errors and account-related issues. |

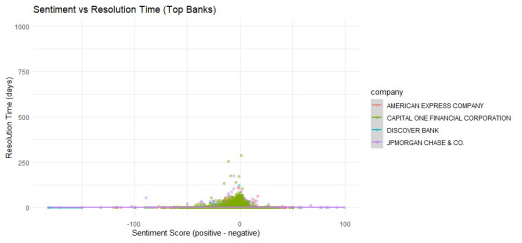
### Conclusion

The resolution of consumer complaints across these top banks shows a considerable range, with severe outliers in all cases. American Express recorded the most extreme delay. The thematic analysis reveals that complaints centered around **credit**, **accounts**, and **cards** consistently dominate the delayed cases across all institutions. These recurring issues suggest that disputes involving financial products such as credit cards and account management are inherently more complex and may require better procedural frameworks, faster verification, and improved customer service practices to avoid prolonged consumer dissatisfaction.

## Task 6: Impact of Negative Sentiment on Resolution Time

We looked at how the sentiment of a complaint (whether it's more positive or negative) affects how long it takes for banks to solve the issue. In the scatter plot, most of the points are packed around the middle, showing that most complaints are neutral and get resolved pretty quickly. Even for very

positive or very negative complaints, the time taken to fix the problems doesn't seem to change much just by looking at the plot.




| company | n | correlation |
|---|---|---|
| <chr> | <int> | <dbl> |
| AMERICAN EXPRESS COMPANY | 43537 | -0.0233 |
| CAPITAL ONE FINANCIAL CORPORATION | 125544 | 0.00133 |
| DISCOVER BANK | 36374 | 0.0287 |
| JPMORGAN CHASE & CO. | 138512 | -0.0000415 |

Figure 11: Correlation Values

We also checked the correlation between sentiment and resolution time for each bank. The numbers we got are very close to zero, meaning there's almost no link between how someone feels in their complaint and how long the bank takes to sort it out. For example, American Express has a small negative correlation (-0.0233), and others like Capital One, Discover Bank, and JPMorgan Chase also have very tiny values. So, overall, sentiment doesn't really seem to affect how fast complaints are resolved.

## Additional Tasks 1 & 2

### Introduction

In this section, we examine the relationship between regulatory enforcement actions and UDAAP (Unfair, Deceptive, or Abusive Acts or Practices) complaint volumes, focusing on Discover Bank. Two significant enforcement events were identified—one in August 2015 and another in October 2020. Our analysis explores whether UDAAP complaints exhibit a discernible pattern around these dates, specifically investigating whether complaints escalate prior to enforcement actions and subsequently recede as the bank adjusts its practices. For each case, we take the count of UDAAP complaints three months before the enforcement action and three months after that.

### First Enforcement Action - August 2015

The first instance of enforcement action taken against Discover Bank was in August 2015.

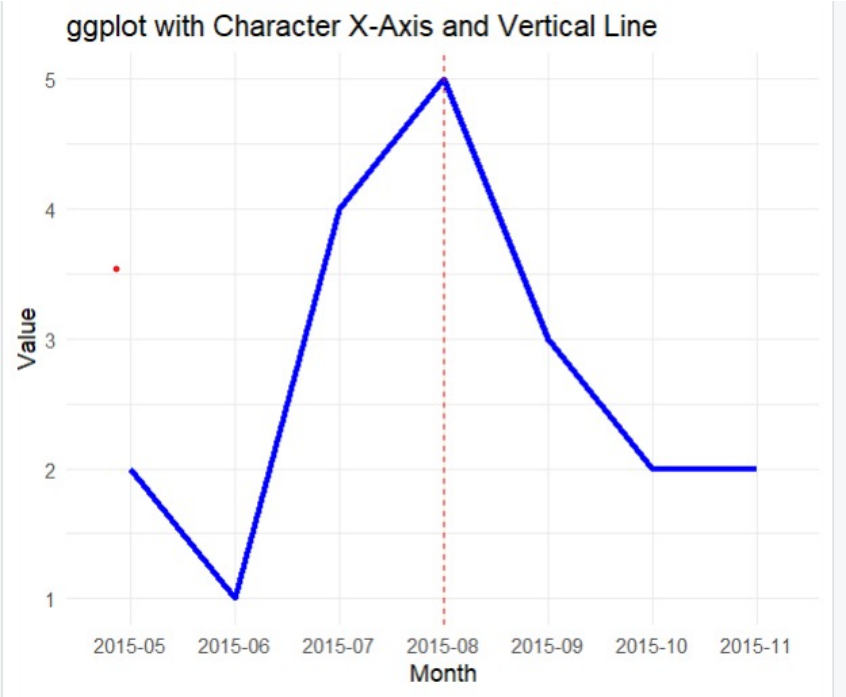



Figure 12: Month vs UDAAP complaints from May-2015 to November-2015

**Second Enforcement Action - October 2020**

The second instance of enforcement action taken against Discover Bank was in October 2020.

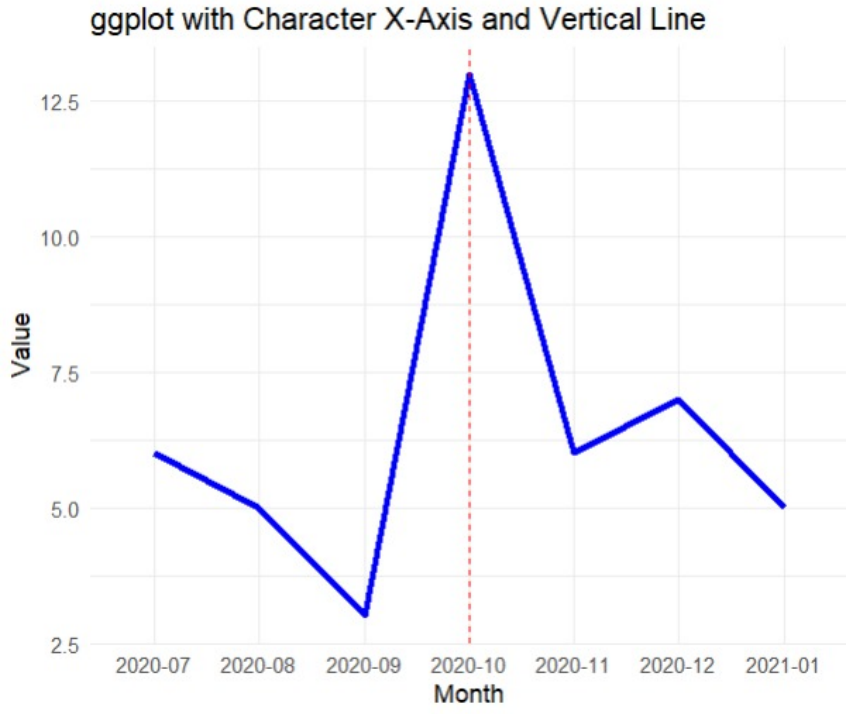



Figure 13: Month vs UDAAP complaints from July-2020 to January-2021

**Conclusion**

In both cases, we can see a slight increase in the number of UDAAP complaints just before the enforcement action was taken. Also, it can be seen that the number of complaints show a steady decrease after Enforcement Actions taken by CFPB.

## References


- **Consumer Financial Protection Bureau (CFPB)**, Consumer Complaint Database.
  `https://www.consumerfinance.gov/data-research/consumer-complaints/`
  Official open-access database of consumer complaints submitted to the CFPB, used as the primary dataset for this analysis.
- **Office of the Comptroller of the Currency (OCC)**, Unfair, Deceptive, or Abusive Acts or Practices (UDAAP) — Compliance Guidelines.
  `https://www.occ.treas.gov`
  Regulatory interpretation and enforcement criteria for identifying UDAAP violations.
- Hu, M., & Liu, B. (2004). *Mining and Summarizing Customer Reviews*.
  In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.
  Foundational work on opinion mining and sentiment classification from customer complaints and reviews.
- Feldman, R. (2013). *Techniques and Applications for Sentiment Analysis*.
  Communications of the ACM, 56(4), 82–89.
  Surveys various natural language processing techniques applied to sentiment and opinion analysis in real-world data.
- Silge, J., & Robinson, D. (2017). *Text Mining with R: A Tidy Approach*.
  O'Reilly Media.
  Practical guide for text analysis in R using tidy principles, especially useful for tokenization, sentiment scoring, and visualization.
- Poursabzi-Sangdeh, F., Goldstein, D. G., Hofman, J. M., Wortman Vaughan, J., & Wallach, H. (2021).
- CFPB enforcement actions database