# MTH312 - Project6 - Group 15

**Author — Ankita Goyal**
Roll No. 231080016

**Author — Chakravartula Vinay**
Roll No. 231080029

**Author — Ashwani Kumar**
Roll No. 231080025

**Author — Sameer Verma**
Roll No. 220949

## Introduction

This project focuses on analyzing consumer complaints related to Unfair, Deceptive, or Abusive Acts or Practices (UDAAP) from the Consumer Financial Protection Bureau (CFPB) database. The objective is to identify trends, patterns, and underlying themes in UDAAP complaints across various financial institutions, and assess their relationship with economic indicators and complaint resolution outcomes.

### Data Collection and Filtering Approach

To ensure that the complaints analyzed are genuinely related to Unfair, Deceptive, or Abusive Acts or Practices (UDAAP), we adopted a targeted data collection and filtering strategy using the Consumer Financial Protection Bureau (CFPB) complaint database. The steps are outlined below:

1. **Narrative-Based Filtering:** We began by filtering the complaints to include only those that contain consumer narratives, which provide detailed descriptions of consumer experiences. This is critical, as structured fields like *issue* or *product* often lack the specificity needed to capture UDAAP-related elements.

2. **Keyword-Based UDAAP Identification:** We compiled a list of approximately 20 UDAAP-specific keywords to identify potential violations. These include terms such as *"do not owe debt," "unauthorized," "mislead," "threaten," "frequent calls," "lie," "without notice,"* and others. These keywords were applied to the consumer narratives to extract complaints that are more likely to reflect UDAAP concerns.

3. **Refined Dataset Creation:** Only complaints that contained one or more of the above keywords in the narrative were retained. This resulted in a dataset enriched with contextually relevant UDAAP cases.

4. **Temporal Filtering:** We restricted the data to complaints submitted from **January 2016** onward. This cutoff was selected based on our observation that a significant and consistent volume of narrative-based complaints began appearing only after this point in time.

## Trend in UDAAP Complaints Over Time

We began by visualizing the monthly trend in UDAAP-related complaints over the entire dataset and then zoomed in to examine the trend for the most recent two years.
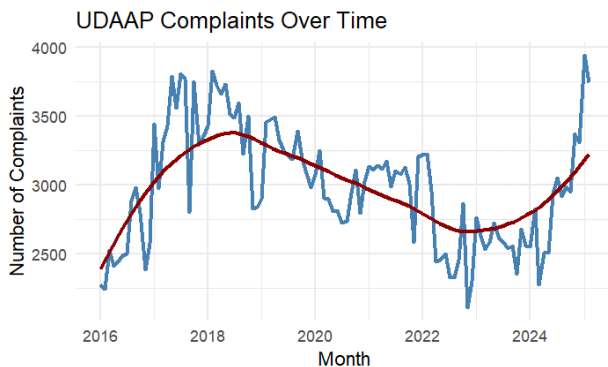


Monthly UDAAP Complaints (2016–Present)
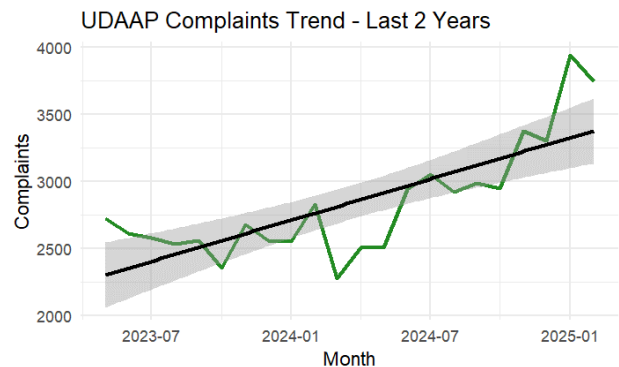


Monthly UDAAP Complaints (Last Two Years)

We can observe an upward trend in the graph for the last two years. To determine whether this increase is statistically significant, we created a simple linear regression model with the number of complaints as the response variable and time (in months) as the predictor.

The model showed that the number of complaints increases by about **1 complaint per month** on average. The associated **p-value of 0.0487** indicates that this upward trend is statistically significant at the 5% level.

In conclusion, although the monthly increase in complaints is gradual, it reflects a steady and meaningful rise in complaints related to unfair or deceptive financial practices.

## Banks with steepest growth in complaints

To identify which banks have experienced the fastest growth in complaint volumes over time, we followed these steps:

1. We aggregated the UDAAP-related complaints monthly for each company.
2. To avoid noise from companies with too few data points, we filtered for companies that appeared in at least 24 different months (i.e., at least 24 times).
3. We fit a linear regression model to quantify growth in complaints. A positive slope means complaints are increasing, and a higher slope indicates faster growth.
4. We then compared the slopes across companies to identify which ones are seeing the steepest growth in complaint volume.

The top 5 companies that show the steepest growth in UDAAP-related complaints over time are TransUnion Intermediate Holdings Inc., Equifax Inc., Experian Information Solutions Inc., Navy Federal Credit Union, and Resurgent Capital Services L.P.

TransUnion, Equifax, and Experian are the three major credit bureaus in the U.S., responsible for maintaining credit reports for hundreds of millions of consumers. So, potential reasons for the steep growth in complaints can be :

- **Increased Consumer Base:** There are more consumers, so naturally, there are more complaints.
- **Data Breaches:** Major data breaches, such as the Equifax breach in 2017, which exposed the personal information of over 147 million Americans, led to a significant loss of trust among consumers and increased public scrutiny.
- **Debt Collection Practices:** As a debt collector, Resurgent is often involved in handling charged-off debts, which can lead to confusion or disputes about the validity of the debt.

## Time Series Modeling and Forecasting of UDAAP Complaints

As we saw in the overall complaints graph above, there is no clear increasing or decreasing trend; instead, the graph exhibits several ups and downs. We selected the SARIMA (Seasonal Autoregressive Integrated Moving Average) model, which is well-suited for such time series data as it captures both non-stationary trends and possible seasonal fluctuations.

To validate our choice, we examined the ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function) plots.
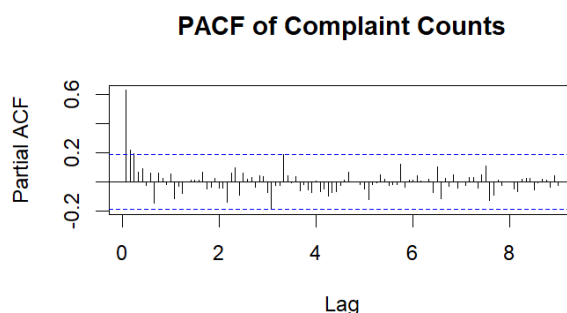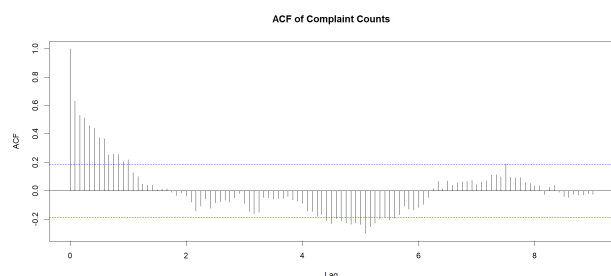


The ACF plot shows an autocorrelation structure that is clearly non-stationary. The up-and-down wave pattern suggests a possible cyclical pattern or seasonality. The PACF plot shows a significant spike only at lag 1, with all other lags falling within the confidence bounds. This implies that only the previous month's complaints have a direct impact on the current month, and further lags do not contribute additional explanatory power.

The SARIMA model fits the data well, with a relatively low forecasting error. Mean Absolute Percentage Error (MAPE), a metric that measures the average absolute percentage difference between predicted and actual values, below 10% is typically considered a strong result in time series forecasting. Our MAPE is **7.58%**, indicating a good model fit.
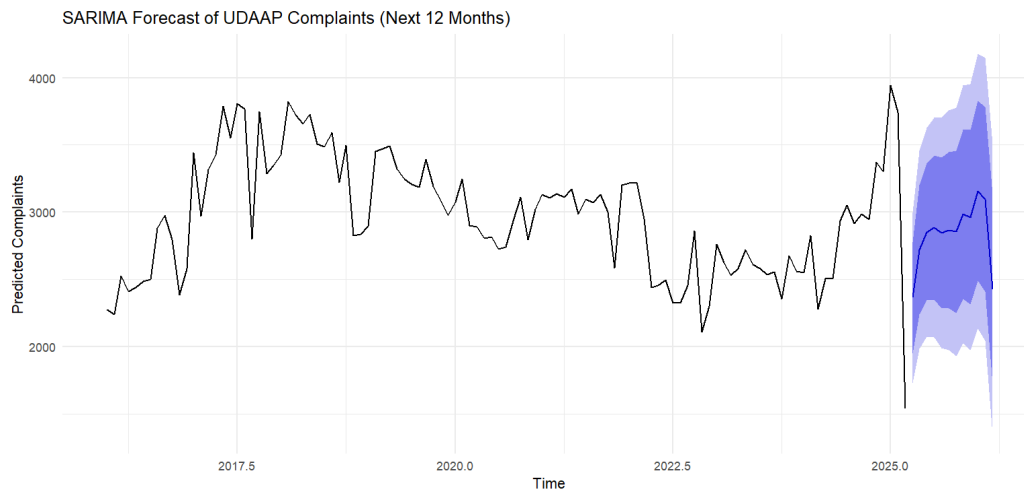


The forecast for next 12 months shows complaint levels stabilizing between **2600–3400**, with possible peaks up to ∼4000 in the higher confidence range. The shaded blue region in the plot represents this confidence range—dark blue indicates the 80% confidence intervals, while lighter blue represents the 95% intervals. The wide intervals toward the end reflect greater uncertainty in long-term predictions.

Based on this projection, UDAAP complaints are likely to remain elevated, possibly driven by continued consumer issues, financial stress, or regulatory reporting improvements.

## Macroeconomic Indicators and UDAAP Complaints

We ran a linear regression to test whether monthly UDAAP complaints are correlated with the monthly unemployment rate (Feb 2023 – Mar 2025). The p-value of **0.00139** indicates that complaints are significantly correlated with unemployment rates. As unemployment increases, UDAAP-related complaints tend to rise. One possible explanation is that during periods of higher unemployment, individuals may experience reduced income and become more financially vulnerable. This can lead to increased reliance on credit products such as loans and credit cards, potentially resulting in more disputes, misunderstandings, or adverse financial events that trigger complaints.

Similarly, we checked for correlations using **Annual unemployment rate (2016–2024)** and **Monthly inflation rate (Jan 2016 – Mar 2025)** and we found no evidence of a relationship between the *annual unemployment rate* and complaint volume.

However, there is a significant *negative* relationship between the **inflation rate** and complaints. It appears that when inflation rises, fewer complaints are filed. A potential explanation is that higher inflation often leads to increased interest rates, which can reduce credit card usage, loans, mortgage applications. As a result, fewer financial interactions might lead to fewer UDAAP-related complaints being submitted.

## Themes in UDAAP Complaint Narratives

After processing consumer complaint narratives and filtering for UDAAP-related keywords, we identified key themes across major credit reporting companies. The analysis highlights that **debt collection** and **credit reporting** dominate consumer concerns. Consumers most frequently report issues such as incorrect information, debt not owed, and ineffective investigations, improper use of report.

For the institutions with the highest number of reported complaints, **Equifax** received the most complaints related to investigation issues, suggesting a perceived lack of responsiveness or transparency in how disputes are handled. **TransUnion** and **Experian** both face consistent reports of inaccurate data, which may indicate systemic issues with data validation and quality control. The issue of **unauthorized use of credit reports** appears across all major bureaus, pointing toward potential UDAAP risks involving consumer privacy and consent.

## Delays in Resolution Across Major Banks

We analyzed UDAAP-related complaints for signs of delays in resolution across four major U.S. banks: **JP Morgan Chase**, **Discover**, **American Express**, and **Capital One**.

Among these institutions, only **JP Morgan Chase** exhibited a notable pattern of delayed resolution. A deeper look into the complaint narratives revealed that the underlying theme often involved **misleading information**. Consumers felt they were given inaccurate or confusing explanations regarding fees, terms, or account actions.

## Evaluation of Relationship Between Negative Sentiment and Resolution Times

To identify negative sentiment in complaints, we analyzed consumer narratives using the Bing Sentiment Lexicon. This lexicon classifies words based on their emotional tone as either negative or non-negative. If a complaint narrative contains words like *threat*, *complain*, *disappoint*, *fraud*, *scam*, *error*, *mistake*, and other similar negative terms, we classify the complaint as "negative." Complaints without these words are considered "non-negative."

To evaluate the relationship between sentiment and resolution times for specific banks, we first visualized the proportion of timely and delayed responses across complaints grouped by sentiment category.



Proportion of Timely Responses by Complaint Sentiment (Balanced Data)
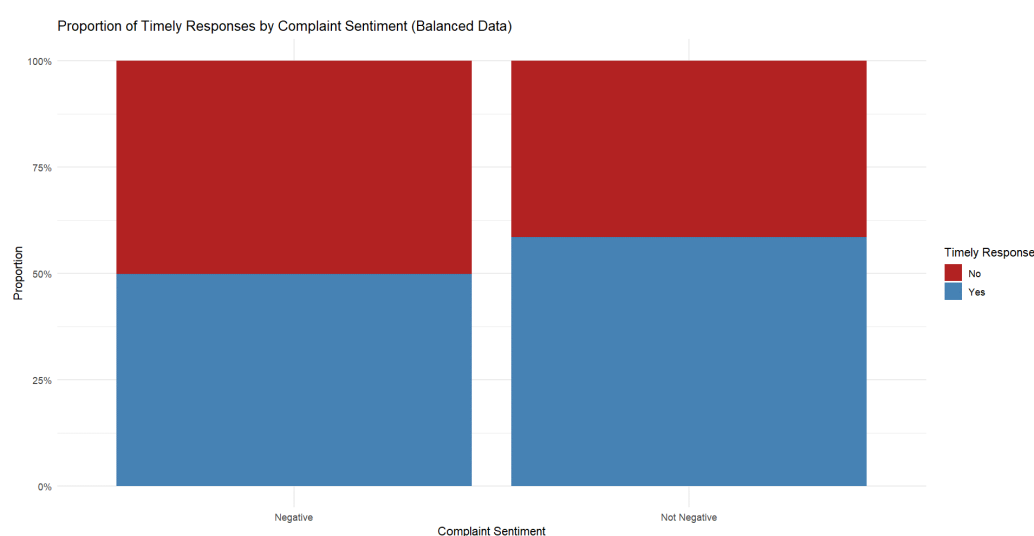
We see that "Not Negative" complaints received timely responses more often than "Negative" complaints, indicating a possible correlation between negative sentiment and delayed responses. To assess the significance of this observation, we conducted a **Chi-Square Test of Independence**. the results ($\chi^2 = 11.156$, p = 0.0008378) suggest a significant relationship between sentiment and timely resolution. Specifically, complaints with negative sentiment are less likely to receive timely responses compared to those with neutral or less emotionally charged language.

This finding implies that the tone of a complaint may influence how quickly companies respond, which could have implications for fairness and consumer trust.

## References

[1] Consumer Financial Protection Bureau. Consumer Complaint Database. https://www.consumerfinance.gov/data-research/consumer-complaints/search/?dateRange=All

[2] Statista. Monthly Unemployment Rate in the U.S. (Seasonally Adjusted). https://www.statista.com/statistics/273909/seasonally-adjusted-monthly-unemployment-rate-in-the-us/

[3] Trading Economics. United States Inflation Rate (CPI). https://tradingeconomics.com/united-states/inflation-cpi