Improving Credit Card Services

**Abstract**:

The Consumer Complaint Database is a federal managed data collection of complaints about consumer financial products and services. In total, there are about 2 million records/entries in the database, ranging from December 1, 2011 to current date. More than 8% of the complaints are relevant to general purpose credit card. In addition, the Consumer Financial Protection Bureau (CFPB) received approximately 154,500 credit or consumer reporting complaints in 2019, a 23% increase from 2018. Complaints with credit card accounts 8% of the total, amount at 29,900, a 4% increase from 2018.

This project will focus on credit card complaints in the database. Through analyzing records of credit card relevant complaints, it is aimed at finding patterns of these complaints and concluding with suggestions to relevant stakeholders on improving services and minimizing complaints on credit card services.

1. **Problem Statement and Business Case**

The Consumer Complaint Database of Customer Financial Protection Bureau (CFPB) was introduced in July 2012, underscoring CFPB’s intent to fulfill two core objectives: enforcing federal consumer protection laws more vigorously and analyzing consumers, financial services providers and market activities[[1]](#footnote-1). The Database is a rich resource for data analysts and data scientists searching for historical as well as emerging trends about consumer complaints with relevant to financial services products, including reasons for those complaints and actions financial institutions are taking to resolve them.

The Database generally updates daily and is open for public access.

CFPB complaint data is organized by products. Financial institutions would benefit from investigating on the dataset. Doing so not only about back-looking complaint analysis, but also identify areas that institutions could develop advanced analytics aimed at identifying and rectifying potential compliance issues before they escalate to a higher level of regulatory and reputational risk.

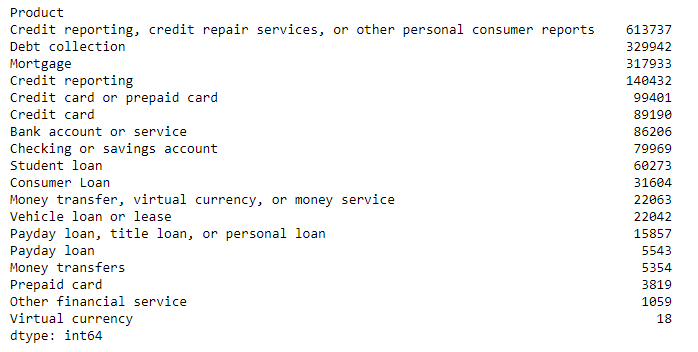
This project focuses on the sector of credit card. More than 8% of the complaints in the database are relevant to general purpose credit card issued by financial institutions. In addition, the Consumer Financial Protection Bureau (CFPB) received approximately 154,500 credit or consumer reporting complaints in 2019, a 23% increase from 2018. Complaints with credit card accounts 8% of the total, amount at 29,900, a 4% increase from 2018.

The research aim is to reviewing and analyzing records of credit card relevant complaints, and providing suggestions for financial institutions on how could they improve their services, minimize complaints, and improve consumer experiences. Through the research, the following questions are expected to be answered:

* What are the patterns of these complaints?
* Is there any improvement in some areas?
* Is there any emerging trend on complaints?
* How did the financial institutions respond to the complaints?
* What are the areas the financial institutions should focus on?
* Recommendations on how financial institutions could improve?
* ….

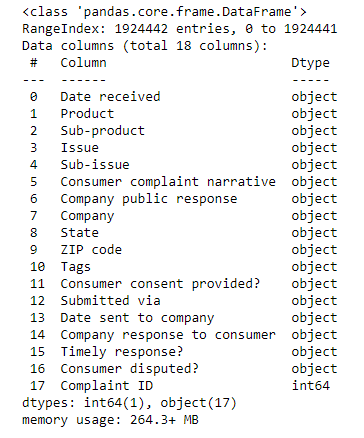
1. **Data Science Workflow**

The database is organized by product. In total, there are 18 different products in the database. However, the grouping into different products is not very clear. For example, at least two products mentioned credit card: ‘Credit card’, and ‘Credit card or prepaid card’. In the latter category, there are prepaid card, government issued card, store credit card, and general purpose credit card. As our focus is on credit card issued by financial institutions, the very first step is to clean up and data and prepare the data ready for further analysis.



Then exploratory data analysis will be deployed to derive patterns and trends from the dataset. This section focuses on not only the complaints itself but also how financial institutions resolve the complaints.

Lastly, machine learning models will be built to further analyze the complaints and advanced analytic skills will be applied.



1. **Data Collection**

The Consumer Complaint Database is a collection of complaints about consumer financial products and services that CFPB sent to companies for response. Complaints are published after the company responds, confirming a commercial relationship with the consumer, or after 15 days, whichever comes first. Complaints referred to other regulators, such as complaints about depository institutions with less than $10 billion in assets, are not published in the Consumer Complaint Database. The database generally updates daily and is open for public access.

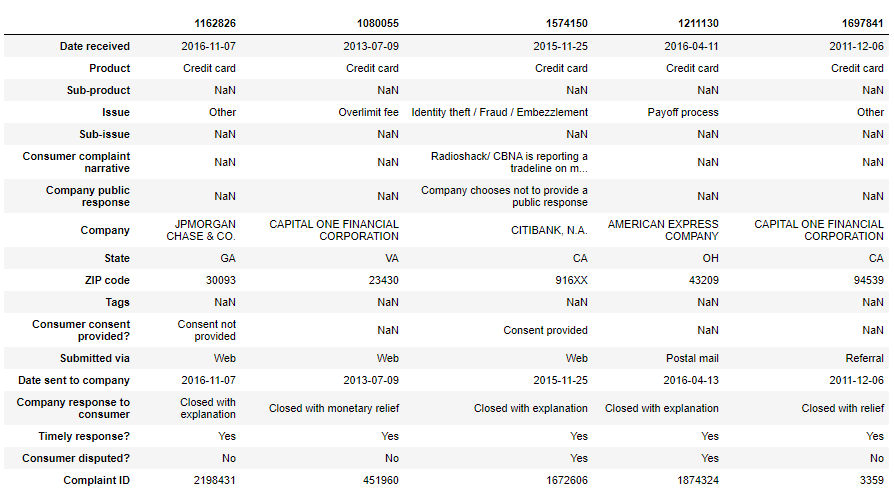
Currently there are about 2 million records/entries in the database, ranging from December 1, 2011 to current date. The data could be retrieved online from: <https://catalog.data.gov/dataset/consumer-complaint-database>.

1. **Data Processing, Preparation, & Feature Engineering**

Cleaning the dataset requires identifying outliers and handling missing data. In this process we will also standardize the timestamp since it is a time series dataset. Other techniques include proper indexing. Outliers will be neglected and missing data will be handled separately.

1. **Machine Learning: Model Selection**

Classification models and regression models will be adopted. I would propose to use different models to compare their performance and select the best one based on performance indicators.



1. **Model Persistence: Deployment, Training, & Data Pipelines**

1. Consumer Financial Protection Bureau. “About us/Core functions.” Retrieved from: <http://www.consumerfinance.gov/the-bureau/>. [↑](#footnote-ref-1)