Consumer Complaint Analysis

Michael Min
Cory Morales
Scott Young
Michael Gilroy

Project Description

We're looking at a consumer complaints dataset for patterns in complaints over time. This could include companies' business practices that cause complaints, or common types of complaints over time.

Prior Work

- [Deloitte analysis of CFPB dataset](#)
  - Analysis of legislature's impact on complaints
  - Graphs and histograms plotting mortgage complaints
  - Statistics on how the cases closed (in favor of consumer or not)
- [USPIRG article on mortgage complaints in CFPB](#)
  - Citing relevant facts in relation to mortgage complaints
  - CFPB actions in regards to incoming complaints

Dataset

- Consumer complaints submitted to the Consumer Financial Protection Bureau
  - https://www.consumerfinance.gov/data-research/consumer-complaints/

Proposed Work

- Data Cleaning: Removing the inconsistencies in names and filling in incomplete cells

- Data Preprocessing: Assigning new values to the data's nominal attributes for the sake of data collection

- Data integration is not planned to be a factor in this project because of how we are currently only using one dataset

List of Tools

- Python
- Numpy/matplotlib/pandas
- WEKA

Evaluation

- Histograms of types of complaints
- Complaints over time by company
- Histograms of company's total complaints
- Using support/confidence to make statistical statements about the dataset