



# Analysis of CFPB's Consumer Complaints Dataset

1.

# OUR DATA SET

# The Dataset

## DESCRIPTION

Our data set contains the list of consumer complaints received by the Consumer Finance Protection Bureau (CFPB) about financial product services offered by banks and other financial institutions across the United States.

The data set contains complaints received by the CFPB from December 2011 till January 2017.

## SOURCE AND LICENSING

The data set was retrieved from the data.gov website.

There is no license associated with the data set and its intended for public access and use.



**719,794 Complaints**

Total Rows



**18 Parameters**

Total Columns



**5 Years**

Time Period

2.

# DATA CLEANING

- ▷ Overall, our data set was quite clean. Majority of the columns had no missing values.
- ▷ For the 'state' column, some of the values were missing. We used the 'Zip Code' column to map the missing states.
- ▷ For the variables 'Consumer Consent Provided' & 'Tags', majority of the values were missing. We removed this columns for the purpose of our analysis.
- ▷ Data type of some of the variables were changed. Ex: 'Medium of Complaint' was of type integer, it was changed to character type.
- ▷ Parsed 'Date' of type string to extract year, month and day as integers.

3.

# PAST RESEARCH

- ▷ 2016 Analysis of Regulatory Violations and Consumer Complaints (Harrow & Moon, 2017)
- ▷ Calling on CFPB for Help: Telling Stories and Consumer Protection (Foohey, 2016)
- ▷ CFPB's consumer complaint database: Analysis reveals valuable insights (Deloitte, 2013)
- ▷ Analysis and Study of CFPB Consumer Data Related to Mortgage Servicing Activities (Daganhardt, n.d.)



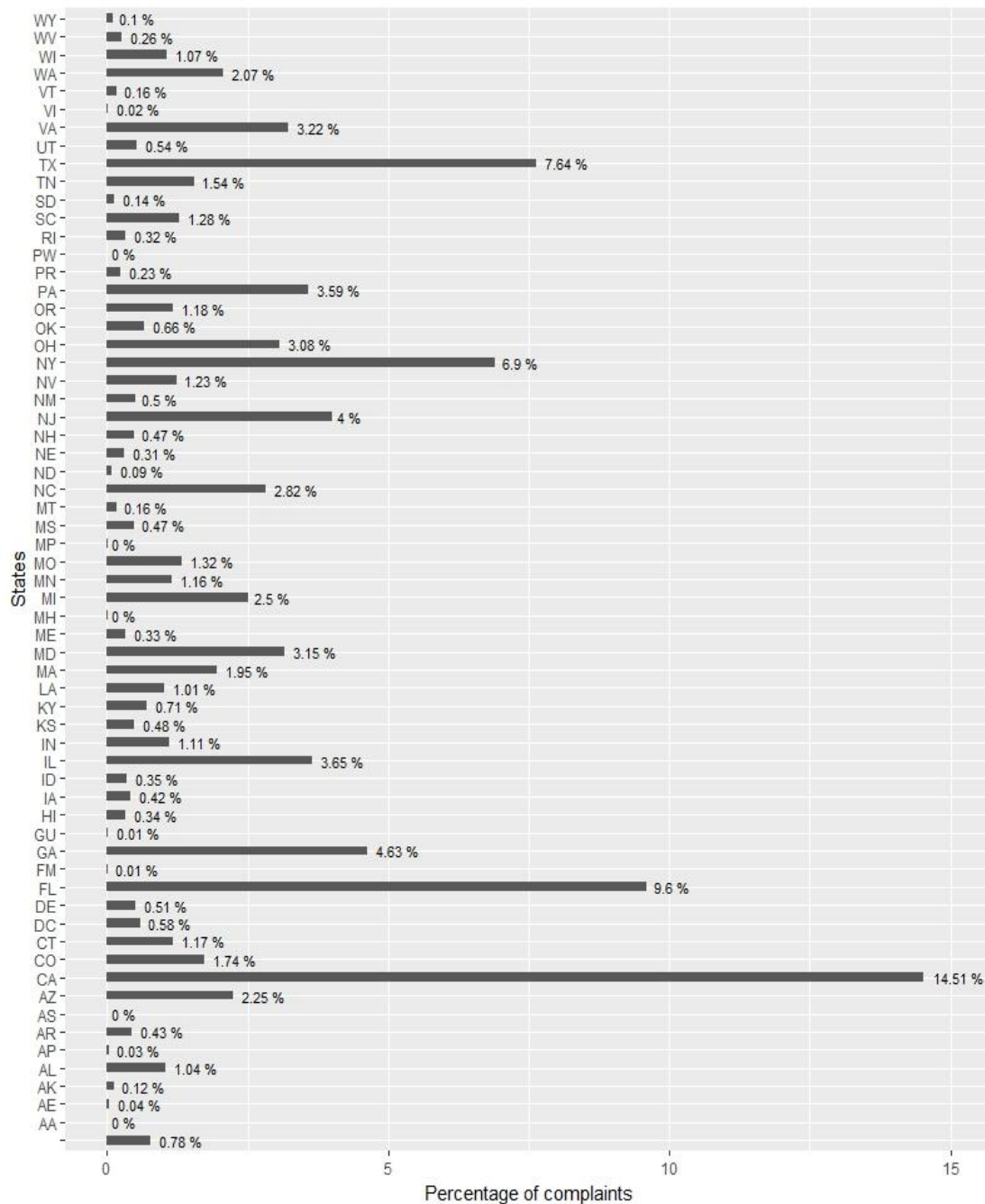
4.

# RESEARCH QUESTIONS

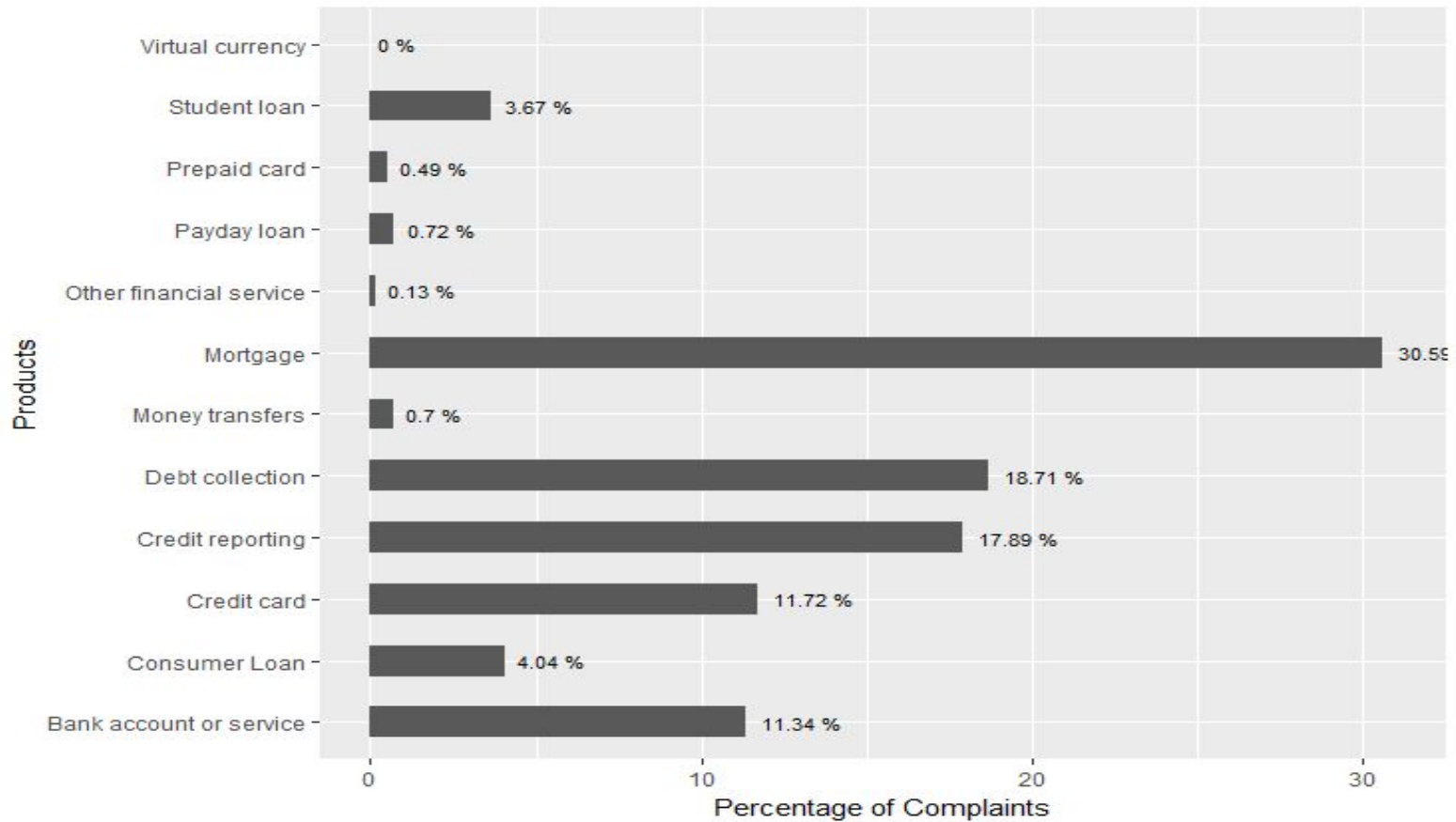
- ▷ Given a consumer complaint, can we predict the response provided by the associated financial institution?
- ▷ Predicting whether the response given by the financial institution would be disputed by a consumer.
- ▷ Predicting the medium via which a complaint is received.
- ▷ Given a consumer complaint, can we predict the location where it originated from?
- ▷ Predicting the emotion (sadness, anger, disgust, fear, etc.) of the consumer from their narrative.

5.

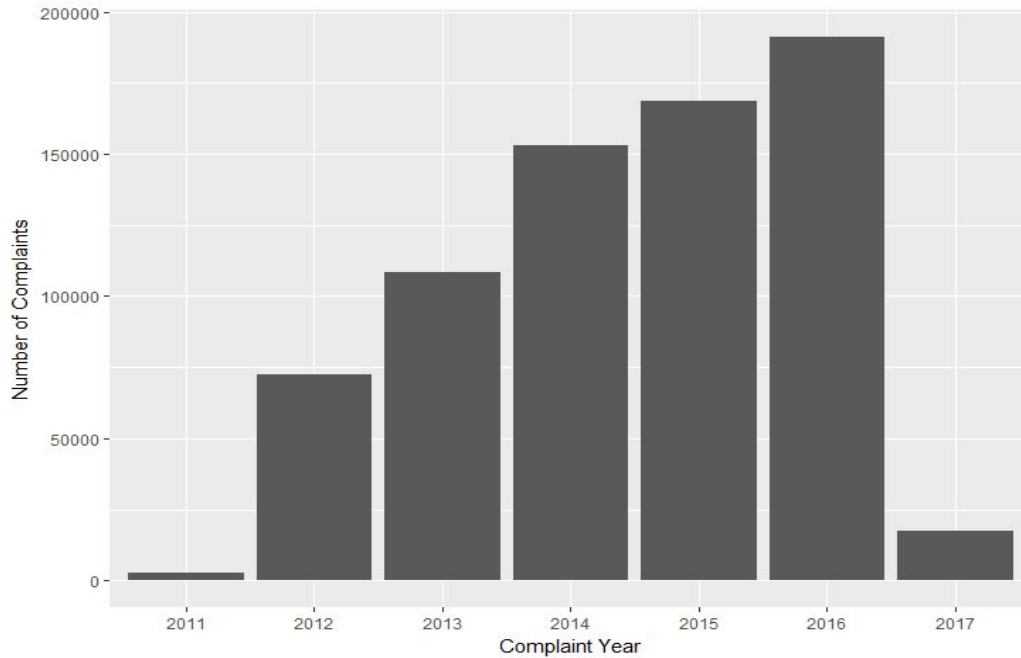
# DESCRIPTIVE STATISTICS



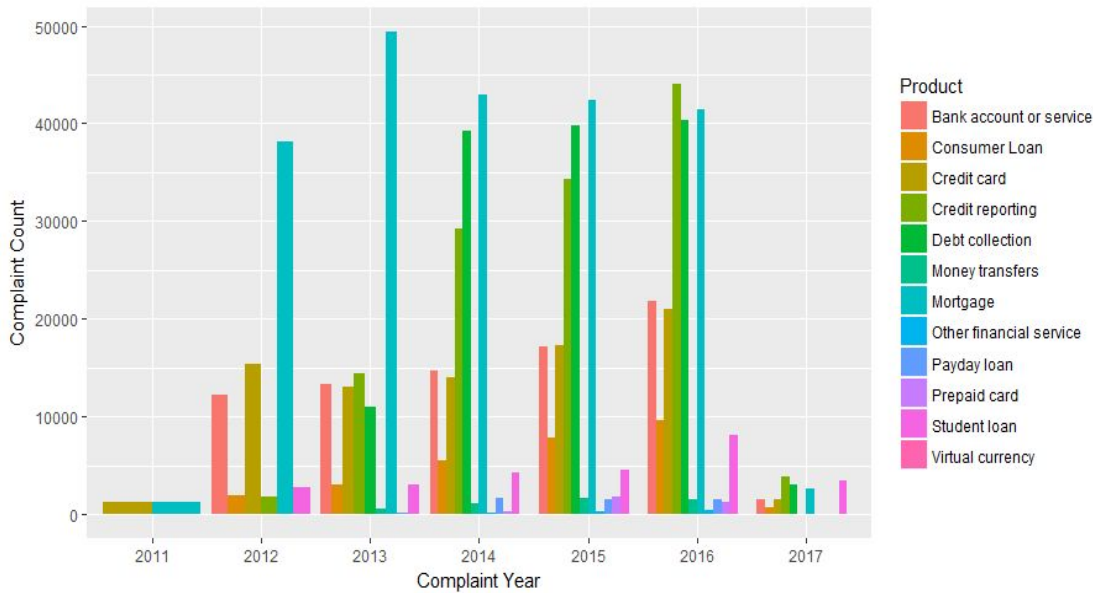
State wise  
distribution of  
consumer  
complaints



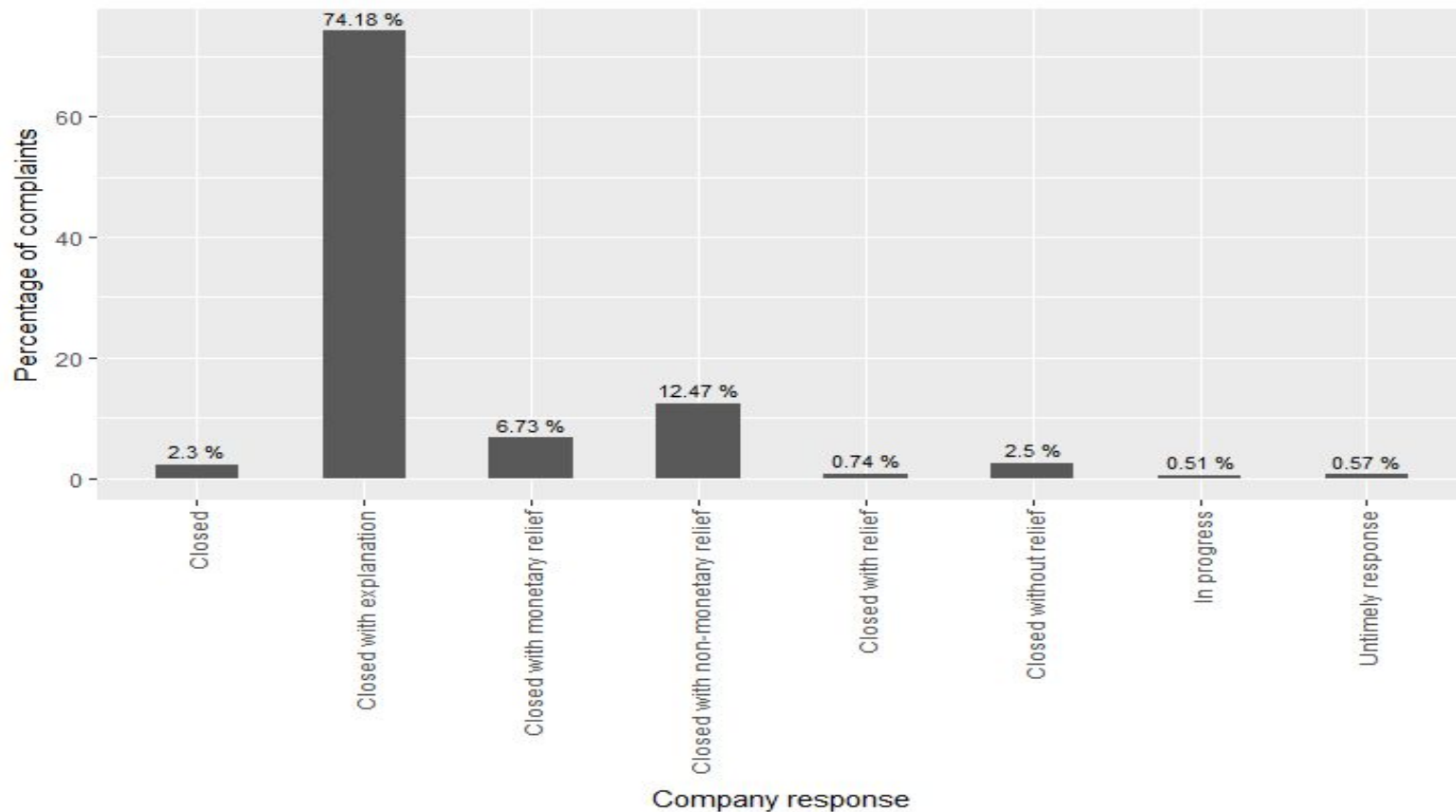
Product wise distribution of complaints



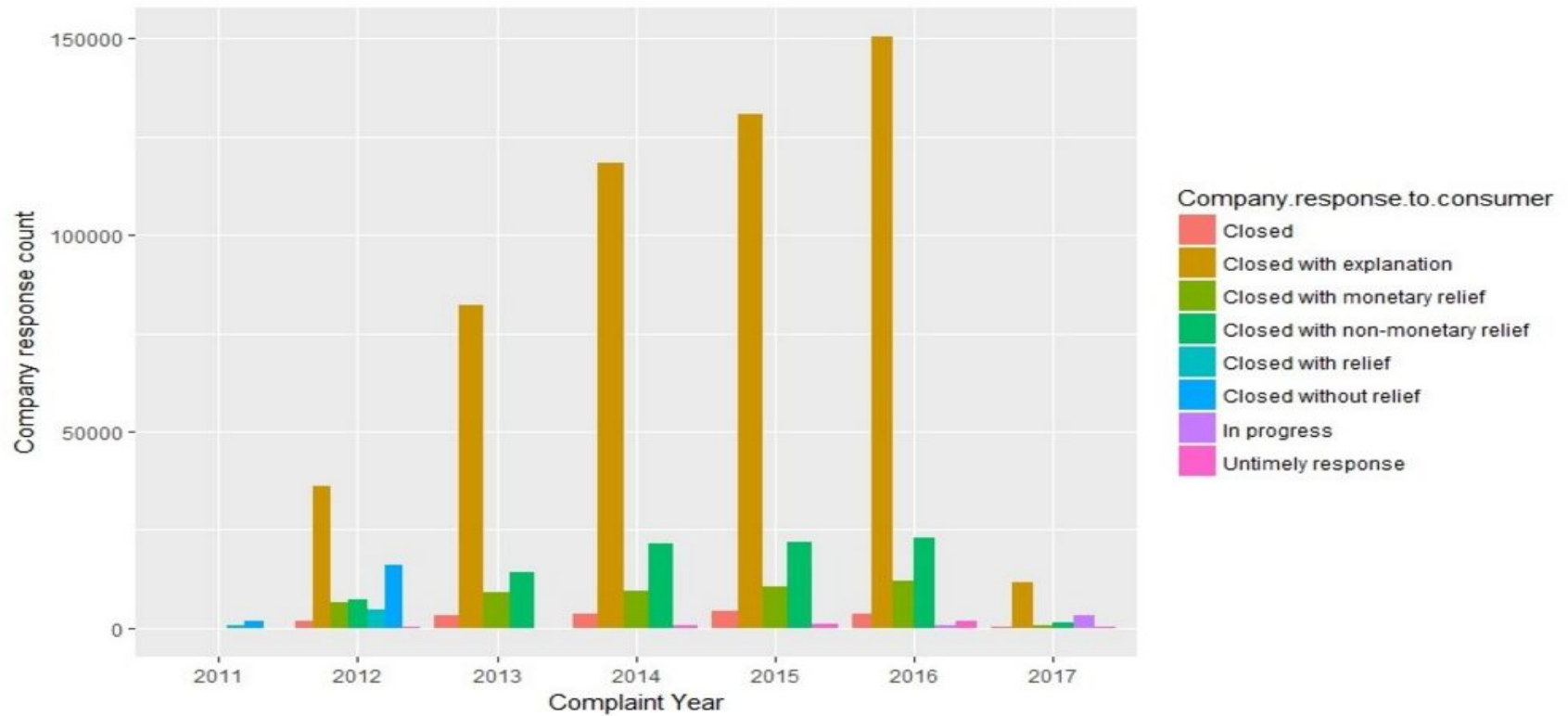
Number of complaints received year on year



Number of complaints received year on year, split by type of product



Distribution of type of response given by a firm for a given complaint



*Figure 7 Year wise complaint-response distribution*

# Year wise distribution of complaint response



Thanks!

**Any questions?**

# REFERENCES

- ▷ Daganhardt, D. *Analysis and Study of CFPB Consumer Complaint Data Related to Mortgage Servicing Activities*. . Retrieved from [http://dsnews.com/wp-content/uploads/sites/25/2015/04/WP\\_DNA\\_CFPBComplaintTrackingWhitePaper\\_FullReport\\_FINAL.pdf](http://dsnews.com/wp-content/uploads/sites/25/2015/04/WP_DNA_CFPBComplaintTrackingWhitePaper_FullReport_FINAL.pdf)
- ▷ Harrow, R., & Moon, C. (2017). 2016 analysis of regulatory violations & consumer complaints. Retrieved February 25, 2017, from <https://www.valuepenguin.com/analysis-regulatory-penalties-consumer-complaints>
- ▷ [3] Foohey, Pamela, Calling on the CFPB for Help: Telling Stories and Consumer Protection (December 9, 2016). 80 Law & Contemporary Problems (Forthcoming); Indiana Legal Studies Research Paper No. 356. Available at SSRN: <https://ssrn.com/abstract=2883242>
- ▷ [4] Retrieved February 25, 2017, from [https://www2.deloitte.com/content/dam/Deloitte/se/Documents/financial-services/CFPBConsumerComplaintDatabase091913US\\_FSI\\_.pdf](https://www2.deloitte.com/content/dam/Deloitte/se/Documents/financial-services/CFPBConsumerComplaintDatabase091913US_FSI_.pdf)