

- <http://datafest.stat.ucla.edu/past-datafests/2022-presentation-slides/>
- Actual Pro bono website:
 - <https://abafreelegalanswers.org/>
- Main Topic
 - The ABA would like to anticipate the sorts of legal questions that arise so that they can prepare volunteers to address those questions, better know how and when to recruit lawyers with specific expertise, and know how to advise state partners on general trends they're seeing.
 - Goals
 - Do sentiment analysis with the question info
 - PCA
 - Time of day
 - Categories
 - Question: How can ABA best allocate their resources to help clients?
 - *How many hours are each attorney working?
 - What are the top 5 states asking questions? TX, IN, FL, SC, IL
 - How many are actually getting answered?
 - What are the top states of unanswered questions?
 - ID, PA, KS, GA, US
 - What are the top states of answered questions?
 - VA, TN, SD, FL, NY
 - What are the top categories/sub within each state: answered and unanswered?
 - What time of day are most questions being asked/answered?
 - evening
 - Ratio of attorneys working on specific days of the week/time?
 - What day of the week are most questions being asked?
 - Tuesday
 - What day of the week are most questions being closed?
 - Friday

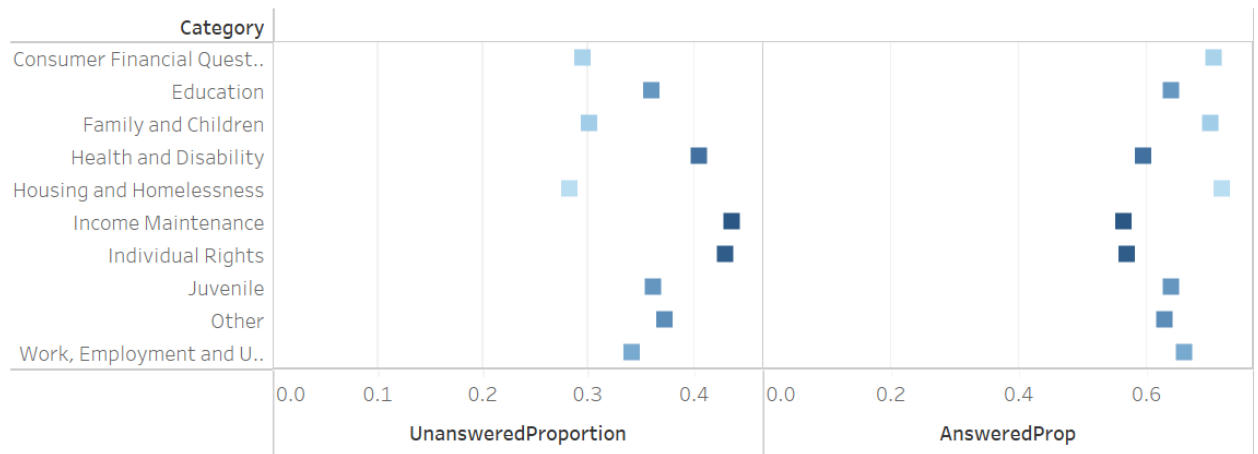
- Demographics vs categories of questions being asked
 - Demographics vs categories of questions being answered
 - Geographic Location of client and attorney zipcodes
 - Chicago, IL: ABA Headquarters
 - Optional: PowerBI/Tableau Dashboard of Interactive heatmap and a few graphs of the above answered questions
 - Text analysis if we have time
- Datasets to Consider
 - Not Using (at this time)
 - attorneytimeentries.csv
 - categories.csv
 - statesites.csv
 - subcategories.csv
 - Looking into
 - attorneys.csv
 - clients.csv
 - Text Analysis
 - questionposts.csv
 - questions.csv
- 3/31
 - Individual Data Exploration
- 04/01
 - Review Individual Data Exploration
 - Discuss plan for Sentiment Analysis
 - Build model for text analysis
 - Create visualizations
 - Create presentation
 - Create 1 page document
- Sites to look at
 - Text Mining + Sentiment Analysis

- <https://www.red-gate.com/simple-talk/databases/sql-server/bi-sql-server/text-mining-and-sentiment-analysis-with-r/>

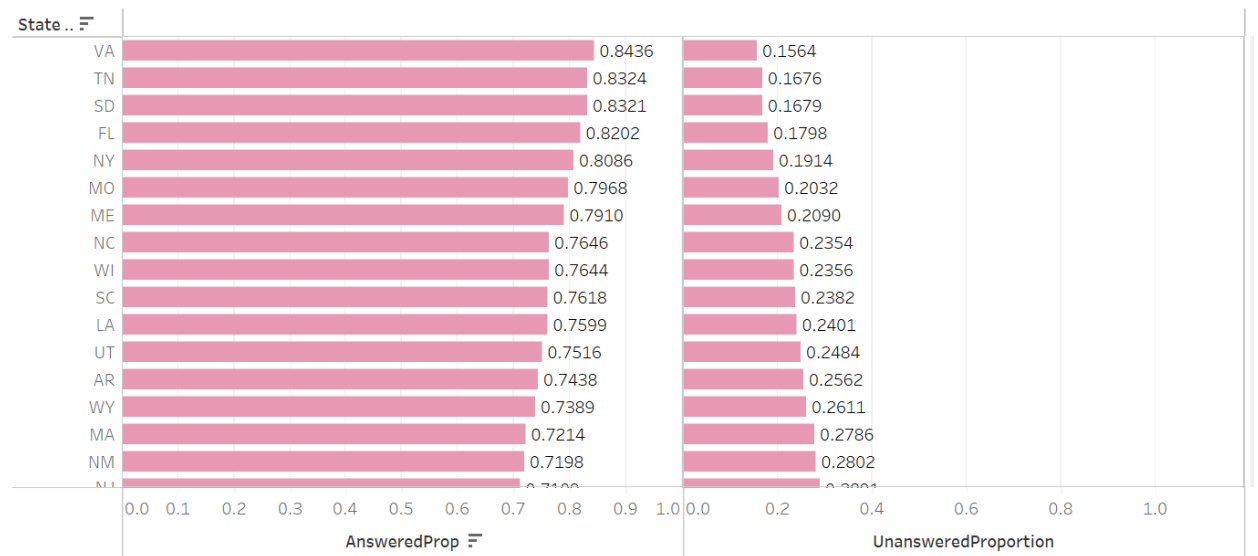
- Tableau

Imprisoned values with Answered vs Unanswered

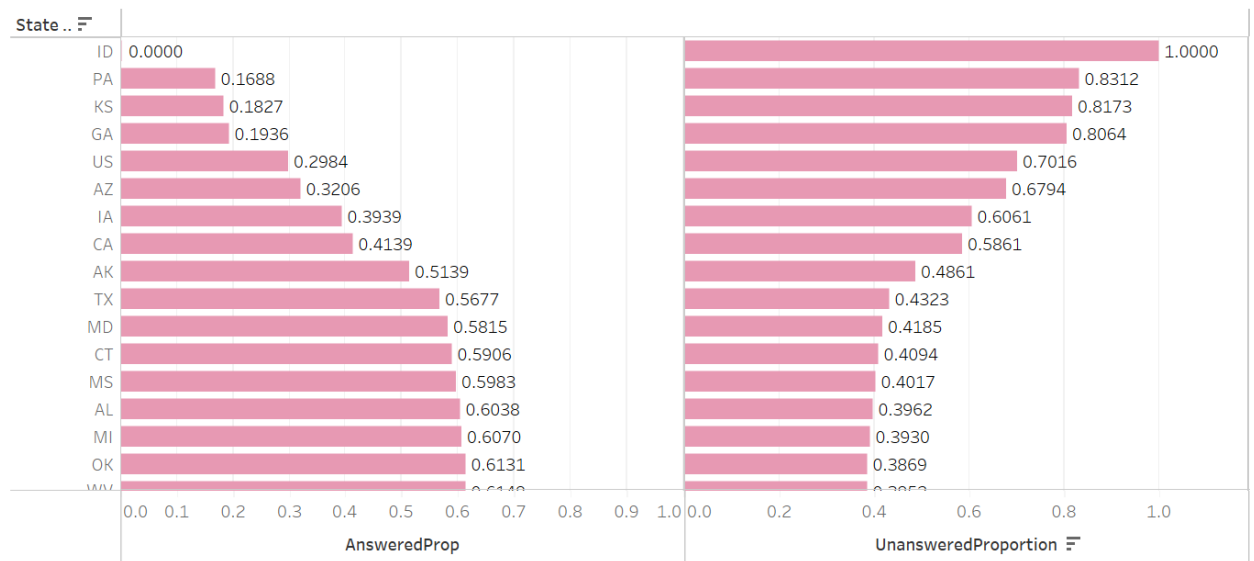
AnsweredI..	Imprisoned	
	No	Yes
0	61,331	10
1	125,923	20



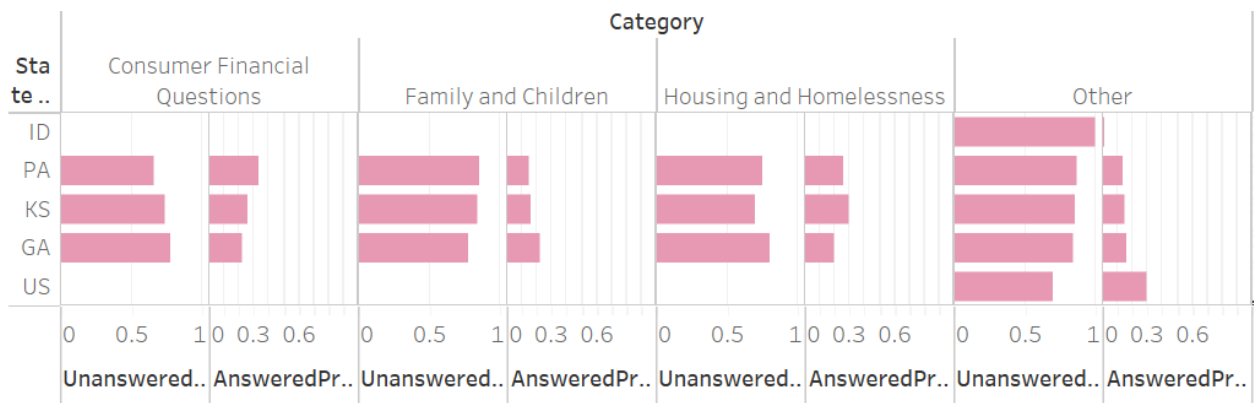
Percentage of Answered and Unanswered Questions with Most Answered at Top



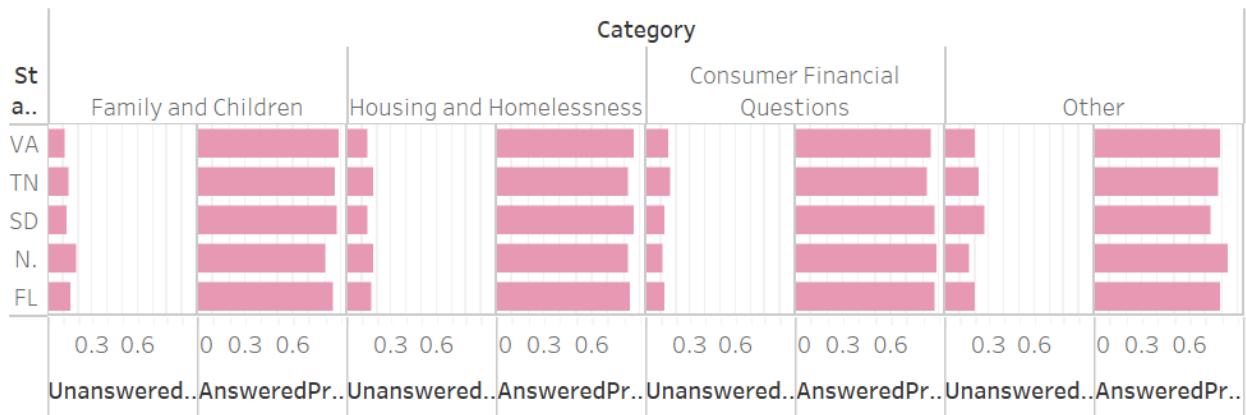
Percentage of Answered and Unanswered Questions with Least Answered at Top



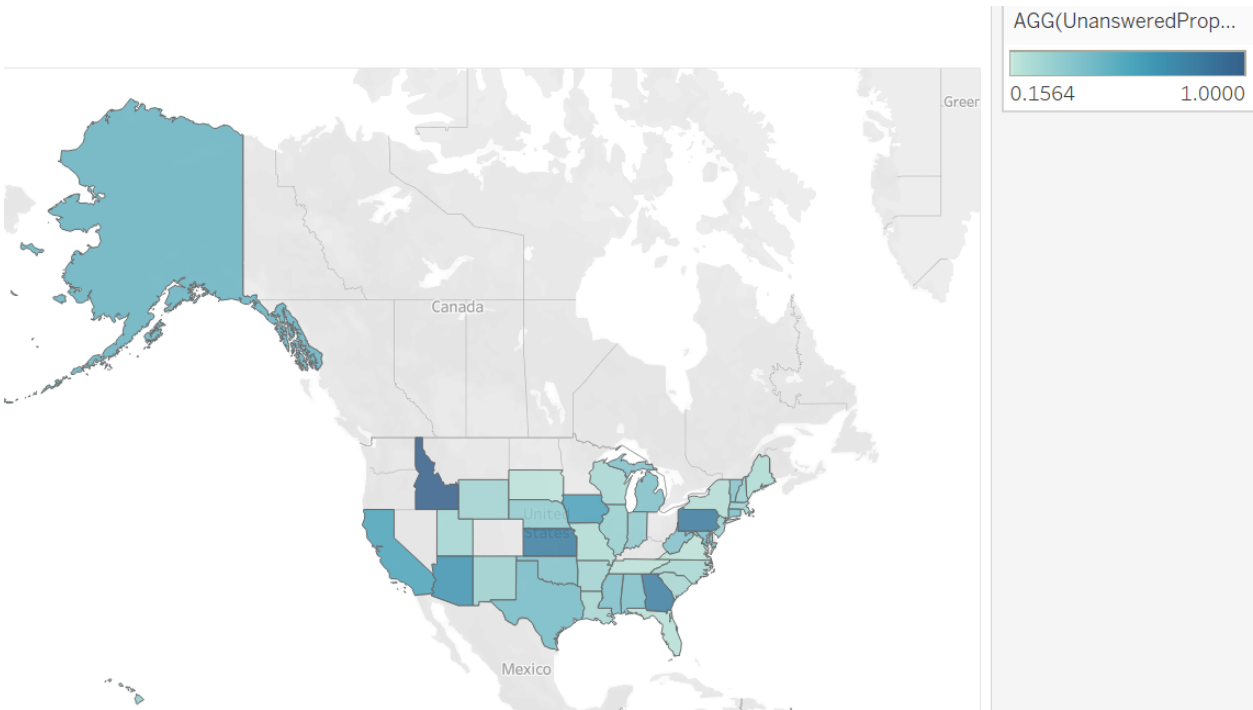
Percentage of answered questions for least answered questions within most popular categories



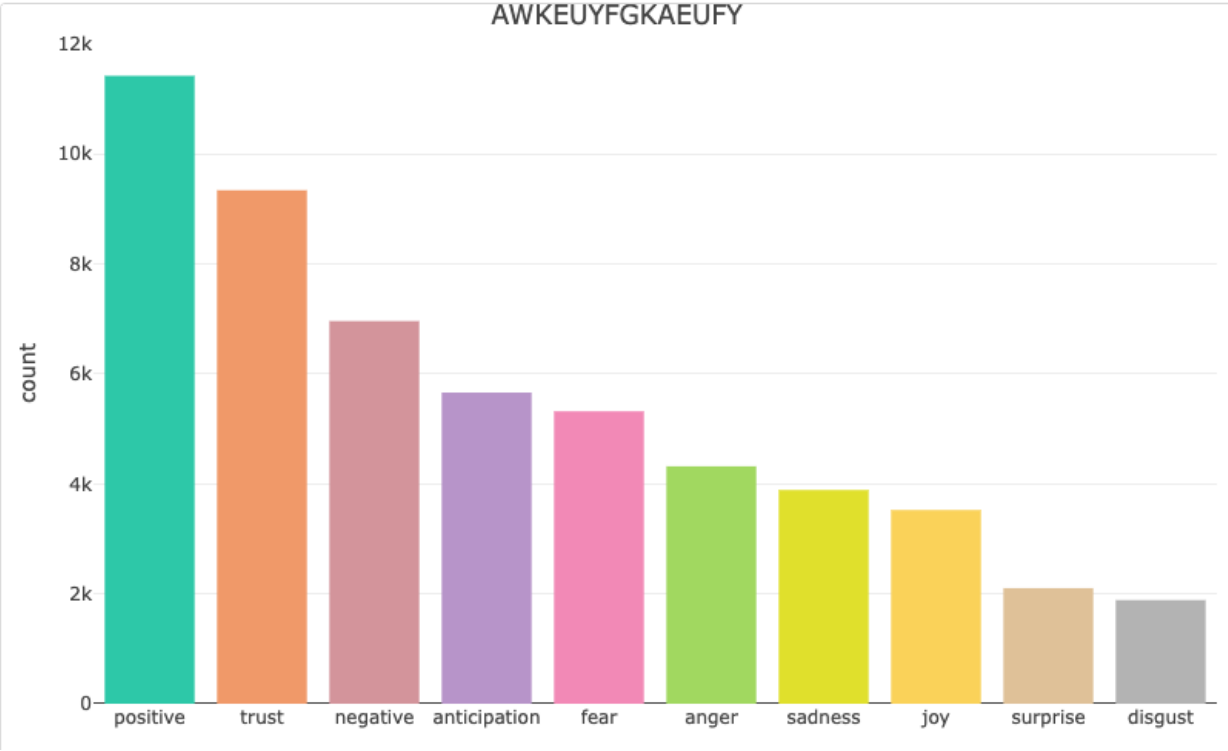
Percentage of answered questions for most answered questions within most popular categories



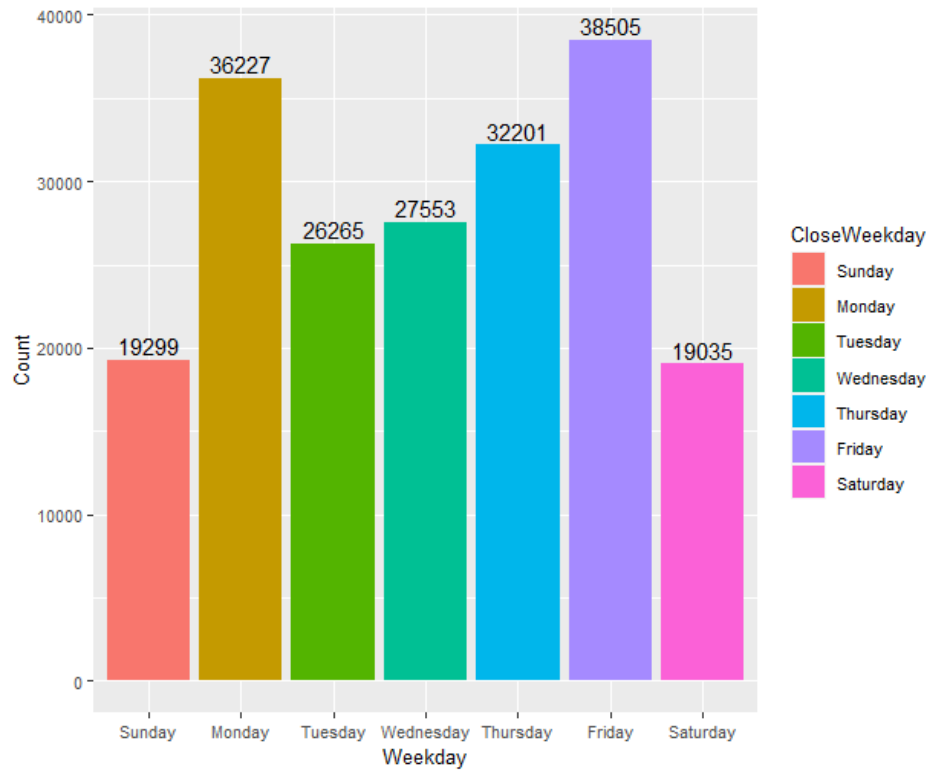
Map showing Unanswered proportion by state



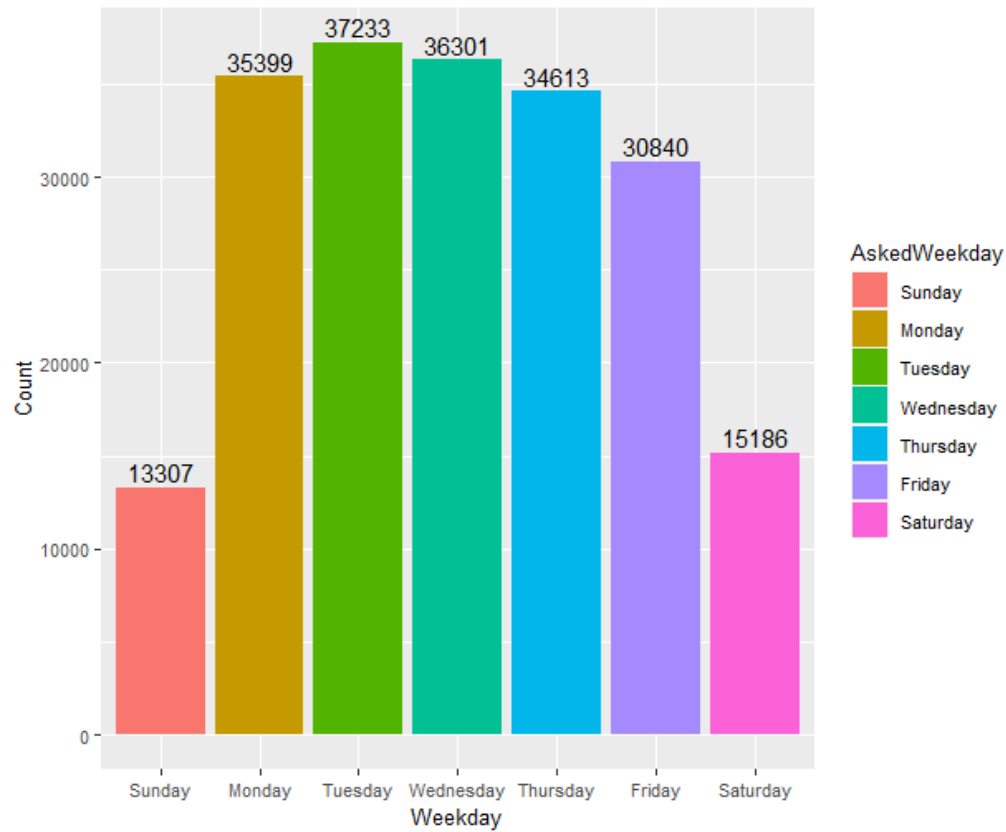
- ARKANSAS



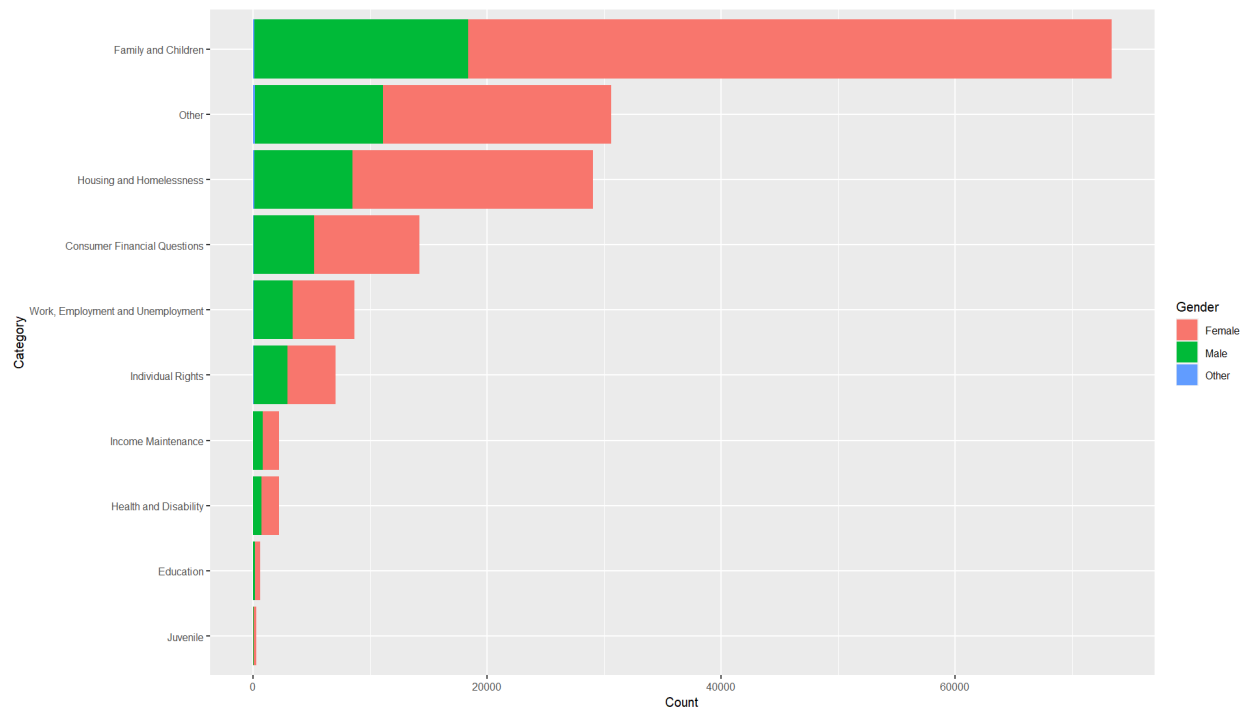
Question Closed by Day of the Week



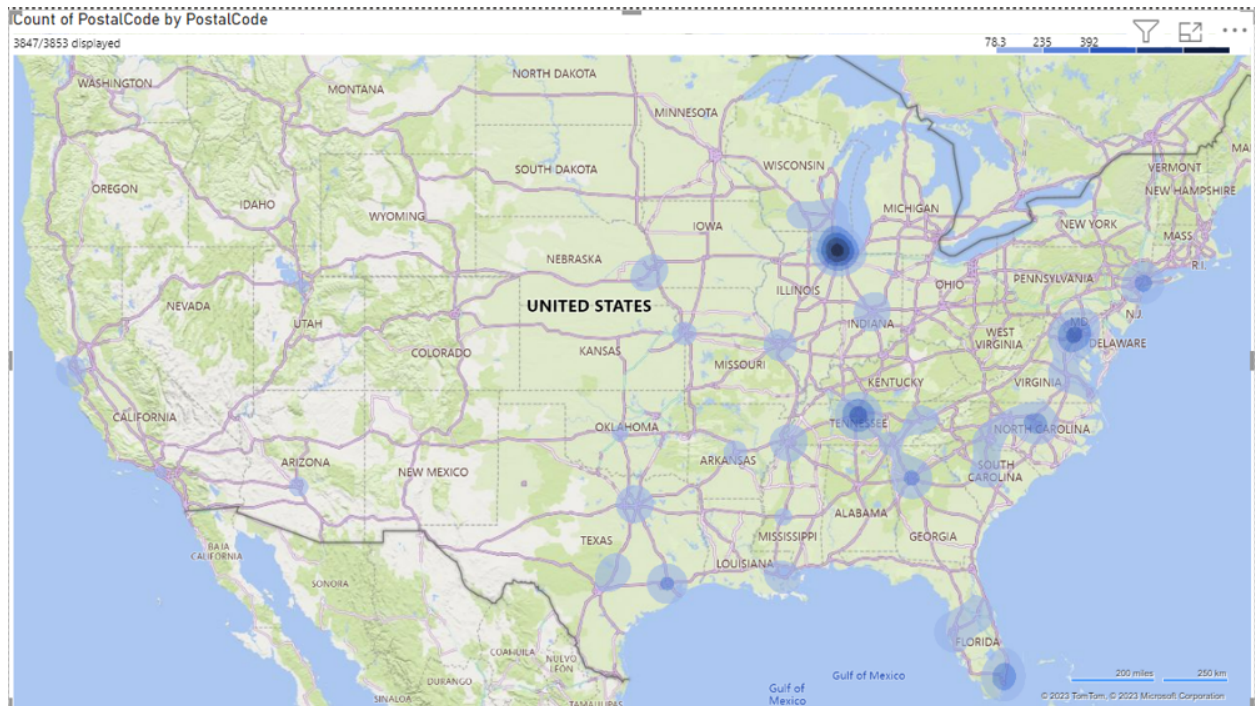
Question Asked by Day of the Week



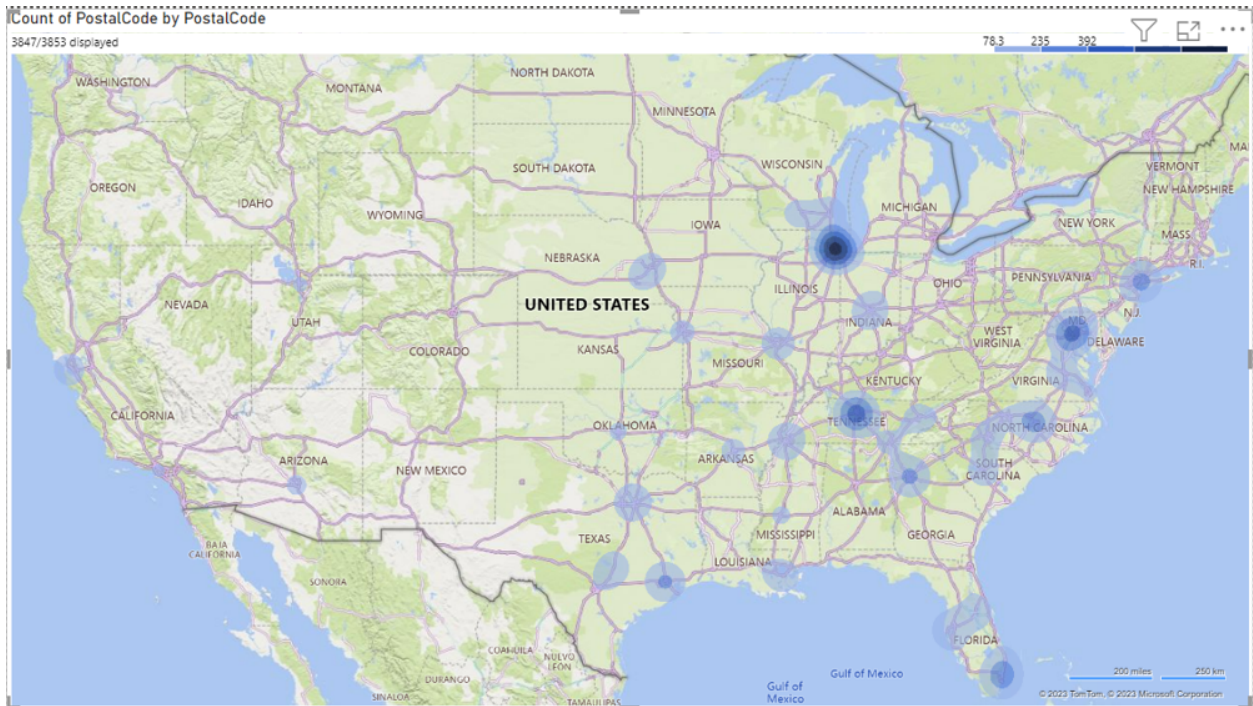
Category Breakdown by Gender



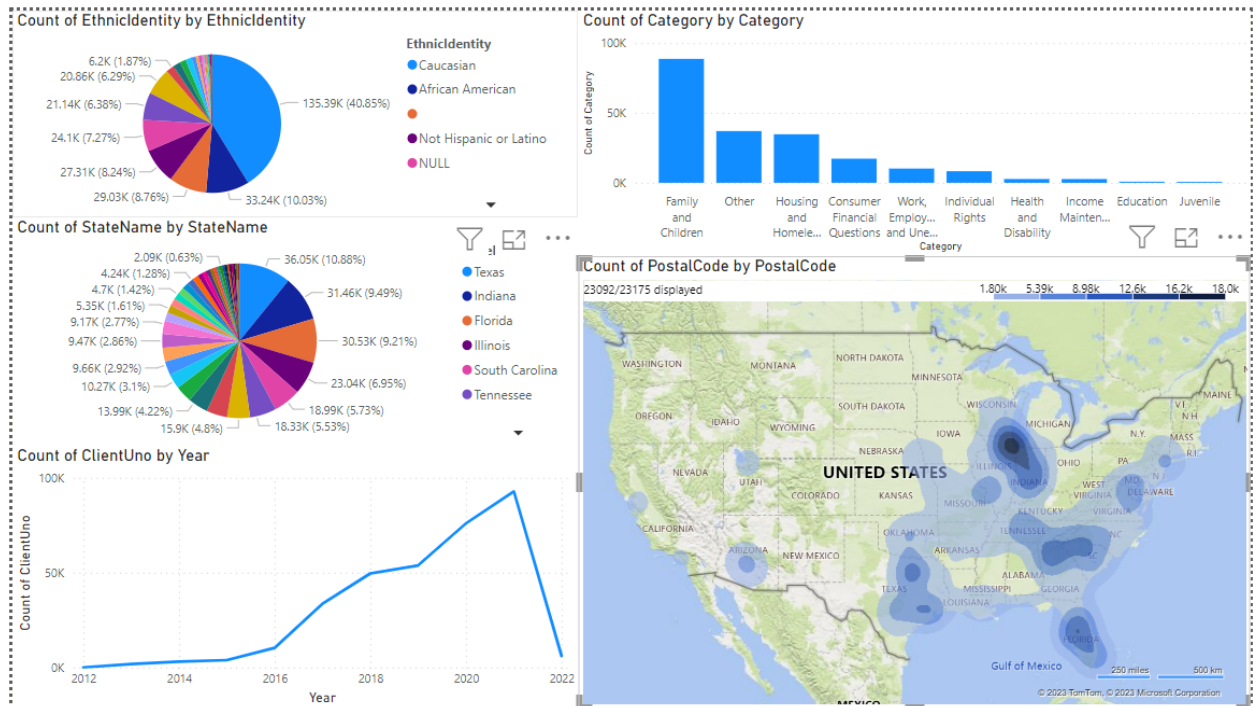
Attorney Zipcode Heatmap



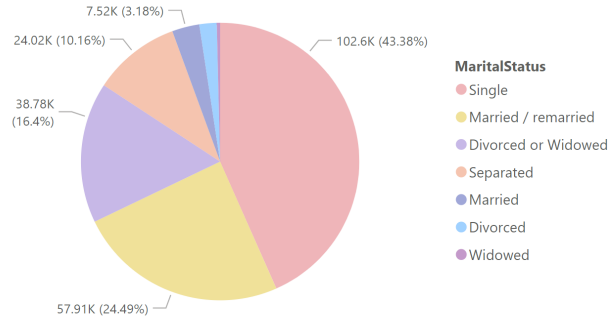
Client Zipcode Heatmap



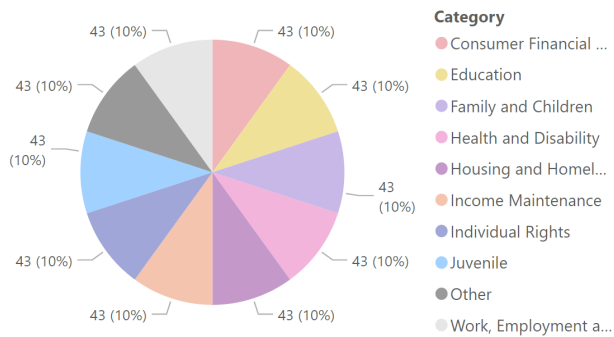
PowerBI Dashboard



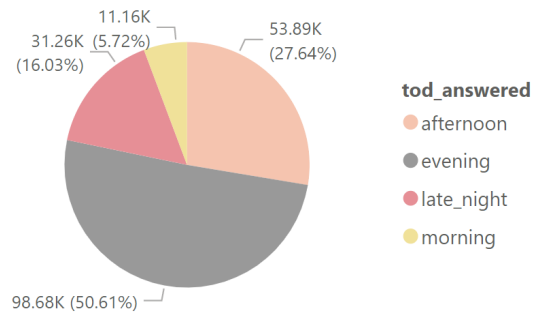
Total # Questions Asked by MaritalStatus



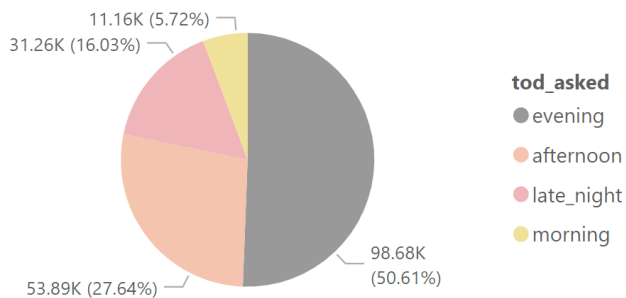
Total # Questions Asked by Category



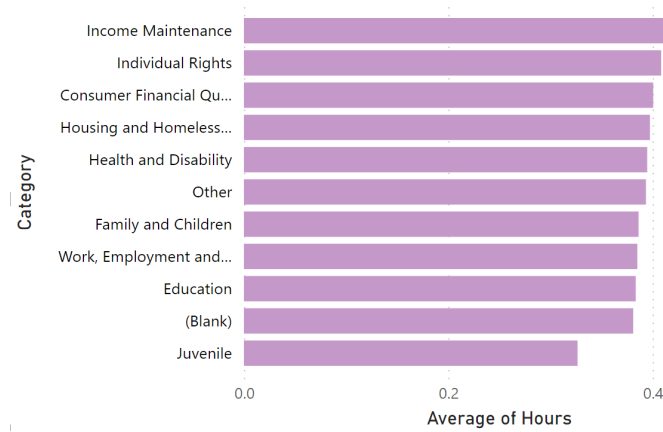
Count of Id by Time of Day Answered



Count of # Questions Asked by Time of Day



Average of Hours Attorneys Spent by Category



Sum of Hours Attorneys Spent by Category

