

### **Insight 1: High demand for legal services between 2Pm&10pm; Shortage of attorneys/volunteers in Texas and Florida**

To promote the efficiency of resources, we want to enhance the efficiency of lawyer resource allocation by analyzing the volume and timing of inquiries received at different times of the day. By visualizing the barplot, we discover that the peak hour when clients are likely to seek legal advice is between 2 Pm and 10 Pm. We conclude that we should assign more lawyers during the afternoon and nighttime to meet the high demand for services. This approach not only ensures that clients receive timely responses, but also maximizes the efficiency of lawyer resources.

Secondly, in future volunteer recruitment, we should adjust attorney numbers in each state based on the case-to-attorney ratio for fair distribution. This would ensure that each state has an appropriate number of legal professionals to handle the caseload and provide adequate representation to all citizens, regardless of their location. From our analysis, we find that attorneys in South Dakota and Arizona have a significantly higher case-to-attorney ratio than other states (1:34.54 and 1: 32.91 respectively), suggesting that more attorneys or volunteers are needed to balance the caseload. This approach promotes fair distribution of legal resources which directly helps improve the quality of legal services in those states.

### **Insight 2: Issues regarding family and children, housing and homelessness are increasing in popularity.**

We conducted an analysis of client inquiries to identify the most frequent and trending law category. By examining the types of legal issues that customers are seeking help with, we can address the areas where pro bono attorneys are needed the most. We discovered that in recent years, the most frequent topics among customer inquiries in the US were family and children law cases, followed by issues related to housing and homelessness. Observing the trend plot, we also identified a notable positive trend in these two categories, indicating that we can anticipate a continuous increase in these two categories in the future. As we pinpoint our analysis to be state-specific, we find a similar, consistent pattern in the two most representative states, Texas and Mississippi. Texas has the highest number of client inquiries whereas Mississippi has the lowest average income. Given this pattern, we can reasonably extrapolate that our finding is likely to hold true for every state in the US. Based on the results, ABA should consider hiring more lawyers with specific expertise in rising legal categories such as family and children law or housing and homelessness law. To conclude, this approach helps ABA to effectively manage its human resources, ensuring that high-quality legal services can be provided to its clients at all times.

### **Insight 3:**

Lastly, we gathered data on the proportion of each question category and discovered that although every law category shows an increasing trend, individual rights and income maintenance law are low in proportion of eligible questions for each category (4.47% and 1.44% respectively). Furthermore, individual right and income maintenance law were the two highest law categories in proportion of unanswered questions among all categories. This implies that additional attorneys may not be required in these categories, but training programs that concentrate on developing skills in these areas of expertise would likely improve the performance of volunteers in providing legal services that are currently receiving less attention than needed.

We noticed that around 18% of eligible questions have category 'other' that might be confusing for the attorney and prevent them from understanding the issues of their clients better. In light of this issue, we estimated a classification rule that aims to predict the category of questions based on clients' characteristics, including age, gender, etc. Specifically, we applied Linear Discriminant Analysis and Quadratic Discriminant Analysis and based on variable selection, our estimation rule has age and number of people in the household as predictors with a leave-one-out cross validation error rate of 46.5%. We can also infer that both age and number of people in the household could be correlated with some underlying characteristics of individuals that are important in deciding their questions. However, it should be noticed that our data is highly unbalanced on the distribution of question categories, which undermines the accuracy of our classification rule and leaves room for future improvement. In addition, we do encounter violations of normality assumptions for both models due to the discrete nature of our predictors. Nevertheless, this preliminary classification rule has revealed the great potential in the development of such prediction algorithms.

