SalesMind AI Ideal Client Profiles

STATS 141XP: Final Report II

Stella Koh, Fathima Shaikh, Theresa Wellington, Emmy Su,
Nam Vien, Site Li, Ashton Jin

2025-03-16

1 Introduction

In the evolving landscape of Business-to-Business (B2B) sales, accurately identifying and reaching Ideal Client Profiles (ICP) remains a persistent challenge. This study aims to enhance lead generation through data-driven methodologies. Utilizing a dataset of outreach efforts within the medical industry, we implemented data cleaning, feature engineering, and logistic regression modeling, incorporating Synthetic Minority Oversampling Technique (SMOTE) to address class imbalance. Our analysis identified key factors influencing successful client engagement, including the significance of call timing and sector-based targeting. Furthermore, we researched and evaluated third-party APIs to optimize mobile number accuracy and reduce outdated data, providing actionable insights for improving AI-driven prospecting tools. The findings contribute to the advancement of AI-enhanced sales strategies by refining data validation techniques and optimizing lead-scoring models, ultimately increasing connection rates and operational efficiency in B2B outreach.

2 Background

SalesMind AI offers conversational AI to enhance sales efficiency of businesses through scalability and customer satisfaction. The company has human-like AI sales agents with specific training to boost sales call results. It helps businesses mitigate hold time, have better and more accurate connections, and easily facilitates appointment booking.

One of the biggest advantages of AI in sales is the sheer scalability. Many calls can happen simultaneously, and humans are only needed for the few, complex cases. Other companies and industries are also taking part in AI, such as Alta. They can handle prospecting, research, outreach, and scheduling using existing sales tools. SalesMind AI stands out because of their advanced AI voice technology. Their AI sales agents are adaptable, and can give personalized connections, handle objections, and instantly set appointments. However, they are grappling with finding the right customers.

They face the same problems in B2B sales as many others: finding their Ideal Client Profiles. An Ideal Client Profile defines the perfect customer for an organization and their services. Through extensive data analysis, it is possible to discover the most relevant and high-converting prospects across different industries.

Another goal of SalesMind AI is in the work of API's. The ideal API would be inexpensive, have a high mobile number accuracy, and a robust amount of data for each individual or company. Validation techniques and increasing statistical power would be necessary to further improve the API. This report will provide criteria for the best API's which will be cross checked with existing API's, with the goal of finding the right fit for SalesMind AI.

This report will consist of data cleaning, predictive modeling, and visualizations, and will be interpreted and visualized for anyone to understand.

3 Methods

3.1 Data Cleaning

SalesMind AI provided a dataset of 3,074 observations and 71 features of their recent cold calls conducted from January 12, 2025 to February 12, 2025. These cold calls were targeted specifically towards the businesses in the medical industry. A large majority of calls consisted of errors in the outcome of the call, which is primarily used as the predictor variable. Such observations with missing crucial information were removed, resulting in 629 complete observations remaining of interest. About 12 missing values in the state-level information were manually corrected.

3.2 Feature Engineering

To include more informative features in the dataset, we underwent an extensive feature engineering process. Given the business name and website url, we categorized all the medical businesses into 14 specialized segments, ranging from Physical Therapy & Rehabilitation to Medical Facilities & Surgery Centers through ChatGPT prompt engineering. For the purposes of model predicting, we further grouped those 14 specialized segments into 3 broader sectors: generalized medical services, specialized medical services, and healthcare consulting (Appendix 6.1).

As the dataset included state-level information, we attempted to include relevant geographic attributes. From the U.S. Census Bureau (2023), we extracted the total population and median income by state as additional predictors. To account for the variation of medical progress across states, we also extracted healthcare spending per capita and annual healthcare growth rate from the 2020 National Healthcare Expenditure Survey.

Many other steps were taken to restructure the original dataset for modeling purposes. Two dummy variables were created to identify if the person of contact was a "CEO" or a "COO." We also extracted the specific month, day, and hour of call.

3.3 Modelling Approaches

Due to the highly imbalanced nature of our dataset (see 5.1), we turned to logistic regression utilizing the SMOTE technique. Through this technique, we synthetically generated new samples for the minority class ("Appointment Booked") to ensure the robustness of our model results.

To cross-check results, we also utilized tree-based modelling–specifically, XGBoost modelling, which deals well with imbalance data using weights. We hypertuned parameters through a grid search with a 5-fold cross validation, and assessed results using relevant diagnostic graphs and tables.

Both models above were used to predict 3 different variables of interest–if an appointment was booked, if an appointment or follow-up was booked, or the number was incorrect–for a total of 6 models.

3.4 API Research Approach

Our API research was driven by the goal of improving lead accuracy and enhancing data validation through a solution with the highest accuracy and lowest cost. We evaluated multiple APIs to determine the most effective method for verifying phone numbers, categorizing line types, and ensuring call success rates. We considered the benefits of aggregating APIs for improved accuracy or using a single API to minimize redundancy and keep costs manageable.

Given the challenges of outdated contact information, low call connection rates, and uncertainty about third-party data sources, our main focus was to ensure phone numbers are valid, active, and correctly categorized before outreach. Additionally, we considered APIs that offer data enrichment to personalize sales calls. To compare potential strategies, we evaluated APIs based on criteria such as features, cost, accuracy, and scalability. Our research led us to form a step-by-step workflow aimed at lowering API costs while maintaining data validation, enhancing accuracy, and introducing cost-effective lead enrichment.

4 Results

4.1 Understanding the Outcome

The outcome distribution of the calls exhibits a notable imbalance, with the majority categorized under "Voicemail" and "Not Interested." Only 4 of 629 calls resulted in an "Appointment Booked" outcome (Figure 1). This suggests considerable challenges in both reaching potential leads and effectively engaging them. A high volume of calls going to voicemail may indicate issues such as suboptimal call timing, incorrect or outdated contact information, or low prospect responsiveness. Given this pattern, a deeper analysis of call timing and response trends could help identify optimal windows for successful outreach, thereby increasing the likelihood of securing appointments. Additionally, leveraging advanced APIs for phone number validation and call scheduling may enhance accuracy in dialing and improve engagement rates.

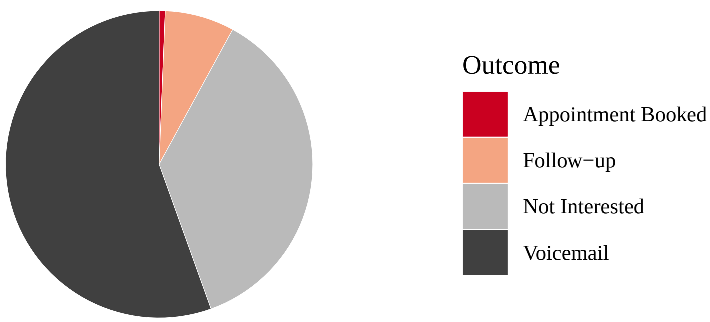

Figure 1: Call Outcome Frequency

Another outcome of interest was how frequent calls were made using the wrong number. Approximately 19.71% of calls or 124 out of 629 calls reached the wrong number (Figure 2). This is not a small percentage and suggests there is room for call accuracy improvements. Note that we were only interested in complete calls in this analysis, so the actual percentage of incorrect phone numbers pulled may be–and is likely to be–much higher after accounting for incomplete calls.

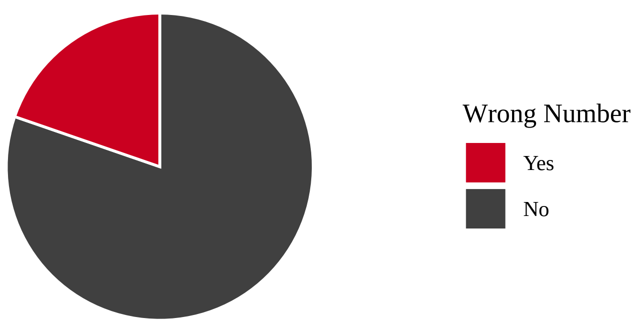

Figure 2: Wrong Number Frequency

4.2 Specialized Segment

The data shows a huge disproportion in the distribution of calls across various specialized medical business segments (Figure 3). The majority of outreach efforts have been concentrated on General Medical Services, while niche segments such as Medical Research and Pharmaceuticals receive significantly fewer calls.

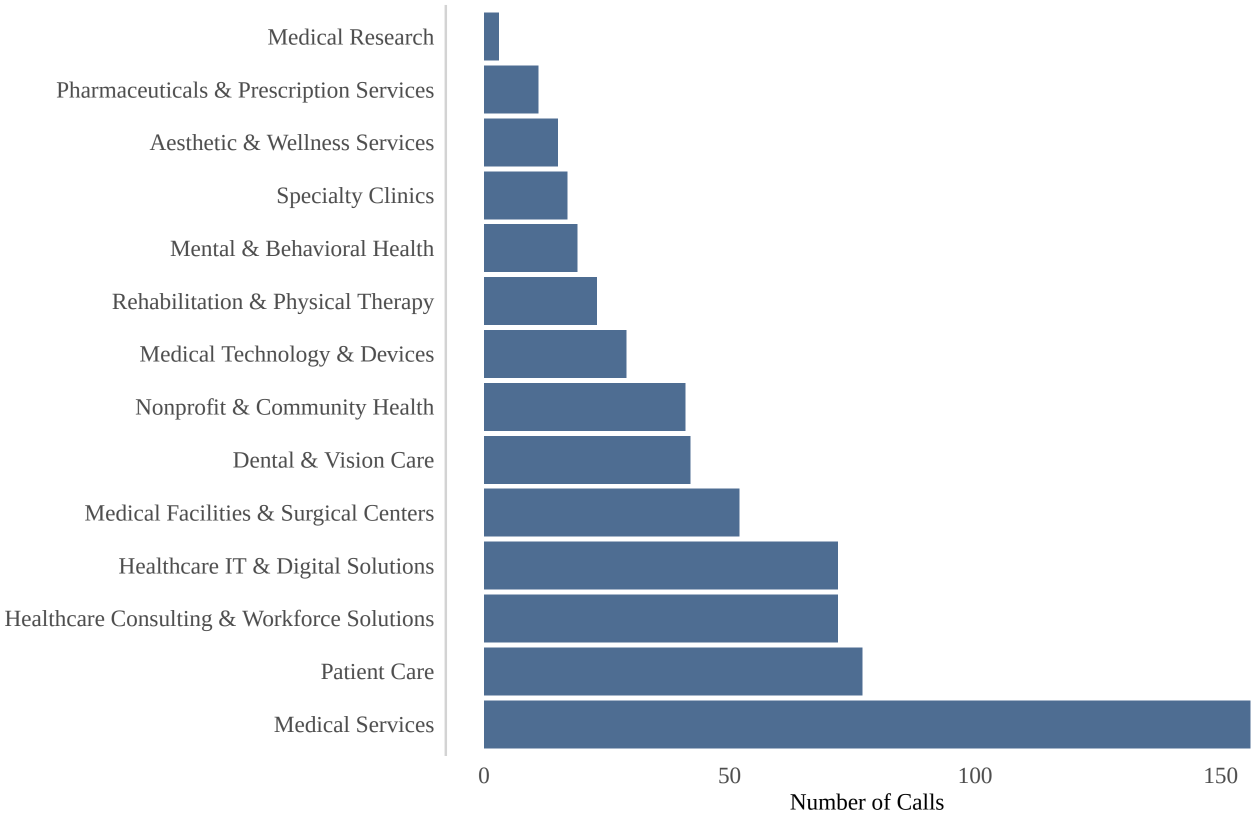

Figure 3: Number of Calls by Specialized Segment

Most of the booked appointments originate from Patient Care, Medical Facilities & Surgical Centers, and Healthcare IT & Digital Solutions, indicating that these sectors exhibit higher responsiveness and conversion potential (Figure 4). This trend suggests that certain fields are more receptive to outreach efforts, and a strategic reallocation of resources could optimize results. Shifting outreach efforts toward specialized fields, particularly those already demonstrating higher appointment conversion rates, could enhance engagement and lead quality. These niche segments may face less competition and demonstrate greater receptiveness to tailored solutions, which may lead to higher response rates.

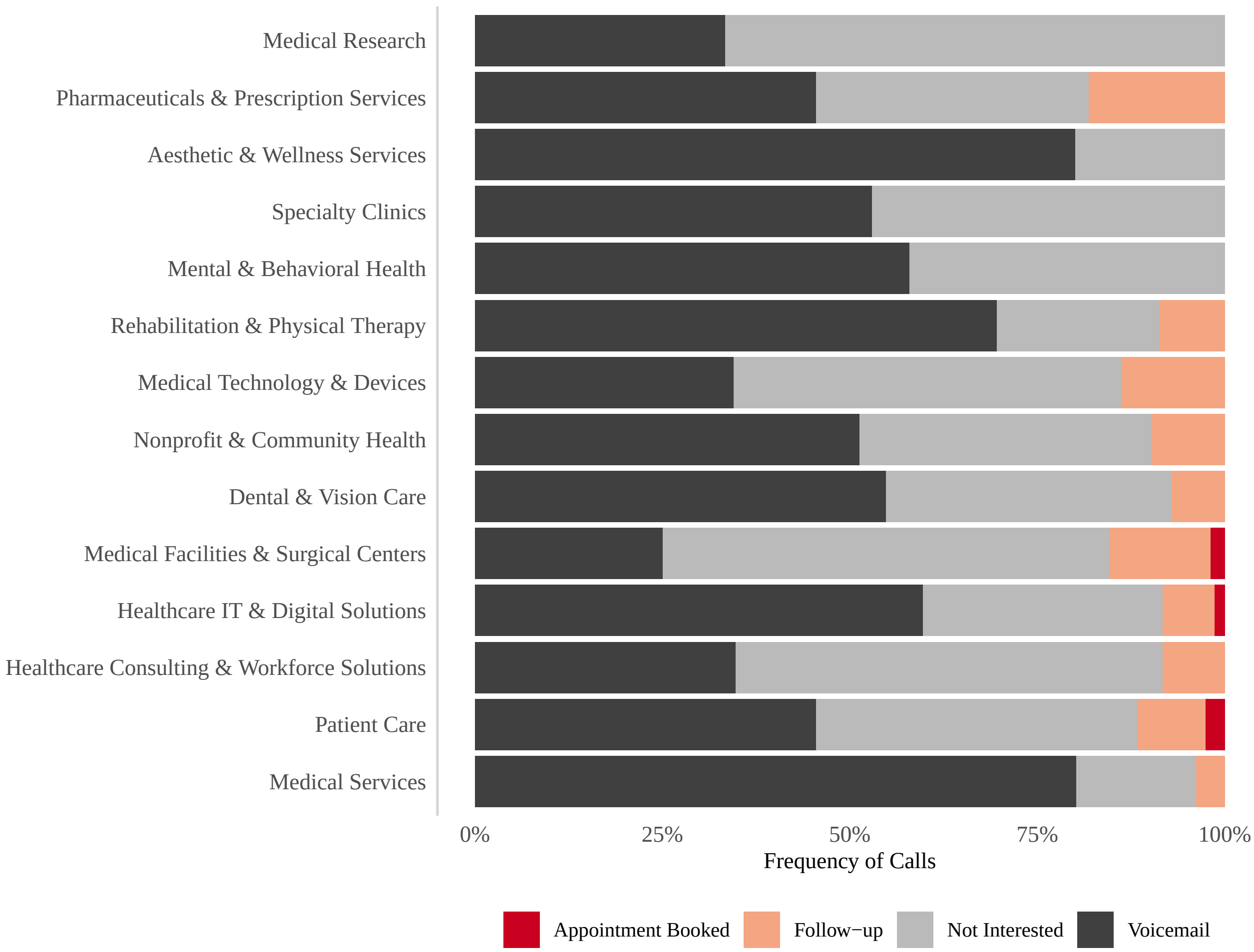

Figure 4: Call Outcome by Specialized Segment

4.3 Grouped Segment

To ensure a more balanced comparison industry engagement, specialized segments were grouped into three main categories: General Medical Services, Specialized Medical Services, and Healthcare Consulting. We still note that General Medical Services and Specialized Medical Services have higher number of calls than Healthcare Consulting (Figure 5, left), but overall, the differences are much less extreme than before.

Now, while the number of calls per segment is the highest in Medical Services, they also experience the highest levels of disengagement with the lowest proportion of interested respondents (Figure 5, right). On top of that, calls towards General Medical Services experience the lowest percentage of actual people picking up the phone, with over 60% of calls going straight towards voicemail instead. Just looking at the proportion of appointments booked (in red), General Medical Services has a lower rate compared to Specialized Medical Services and Healthcare Consulting, which appear to be more comparable to each other. Ultimately, these visuals suggest that General Medical Services is not the most ideal business segment to appeal towards.

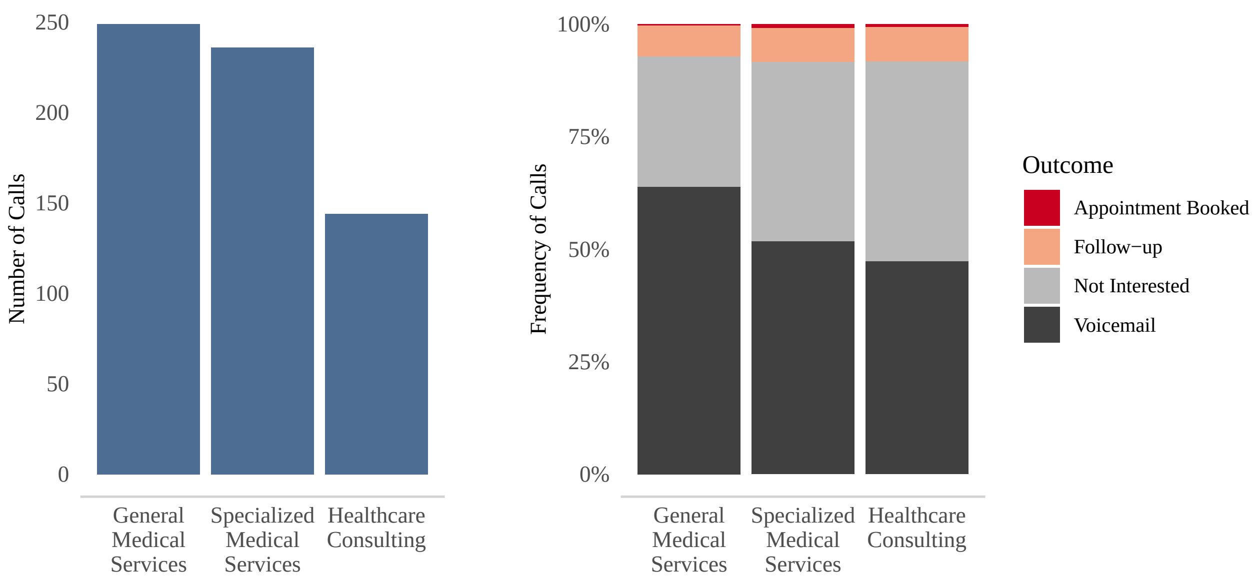

Figure 5: Number of Calls (Left) and Outcome (Right) by Grouped Segment

4.4 Location

The current outreach strategy is heavily concentrated in the Eastern Time Zone, consisting of over half of all calls alone (Figure 6, left). Pacific follows a far second, suggest more calls end up targeting the more populated, coastal areas of the U.S. Central and Mountain are similar in frequency, both under 100 calls.

While most calls occur in the Eastern Time Zone, there is mixed evidence that this region has better conversation rates than the other (Figure 6, right). The Eastern Zone does have the highest proportion of interested respondents but a lower proportion of appointments booked, but both these are likely attributed to the sheer number of cold calls made. We see that the Mountain Time Zone appears to have a slightly higher rate of appointments booked compared to other regions, but given its low total, it is likely not significant. A more even distribution of calls and optimizing call timing based on regional business hours could potentially help improve engagement, but we would likely require more data from non-Eastern time zones to confirm our idea.

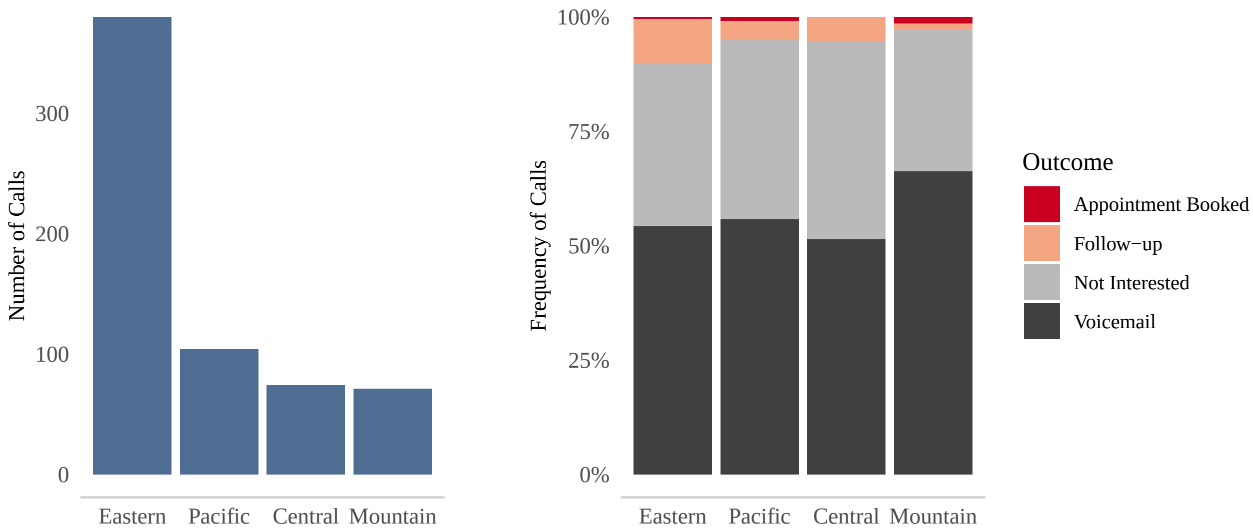

Figure 6: Number of Calls (Left) and Outcome (Right) by Time Zone

Increasing the level of granularity to state-level dynamics, we take a look at how the number of calls and interested calls vary geographically. We also compare these results with economic factors at play–specifically, the median income and healthcare spending per capita in each state.

First, note that the four states with appointments booked were **Utah**, **California**, **Georgia**, and **Maryland**. States with higher engagement rates are primarily concentrated in the Northeast, parts of California, and the Midwest (Figure 7). The highest call volumes are observed in California, Texas, Florida, and the Northeast, which aligns with large populations and major healthcare hubs. However, higher call volume does not necessarily equate to higher interest rates. When comparing total calls to engaged calls, California received the most calls, yet New York and Georgia showed higher engagement levels. Refer to Appendix 6.2 for full-size maps.

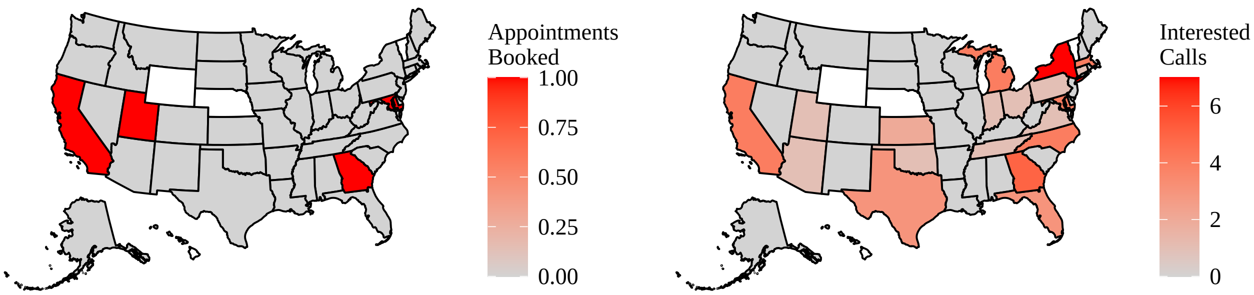

Figure 7: Calls With an 'Appointment Booked' (Left) and 'Interested' (Right) Outcome

In Figure 8, it is difficult to see a clear correlation between income and engagement. We do see higher median incomes around the coastal areas where there are more appointments booked and interested respondents, perhaps suggesting that higher-income states may still offer better conversion opportunities. Healthcare spending per capita is generally concentrated around the Northeast, which is not quite where appointments were booked, indicating that targeting areas with higher healthcare spending per capita may be counterproductive.

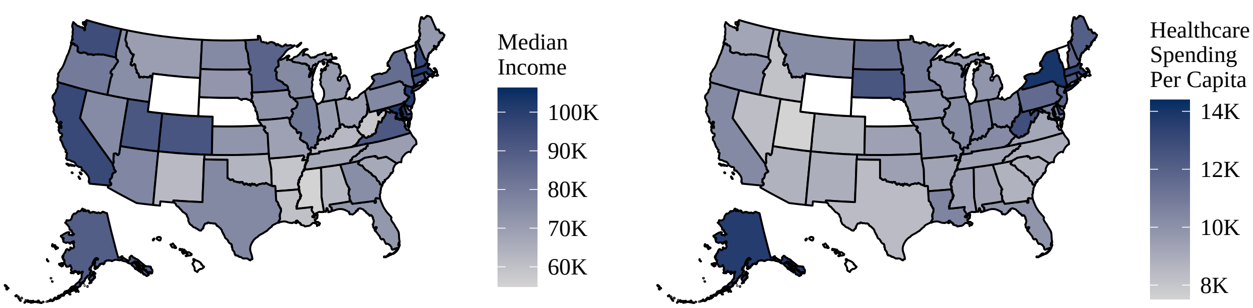

Figure 8: Median Income (Left) and Healthcare Spending Per Capita (Right) by State

4.5 Call Timing and Duration

The majority of calls occur between 12 PM and 6 PM of recipient's local time, with peak engagement between 4 PM and 6 PM (Figure 9). During this time window, higher portions of follow-ups and booked appointments occur compared to other times. The outcome suggests calls made between 4 PM to 6 PM are more likely to be successful.

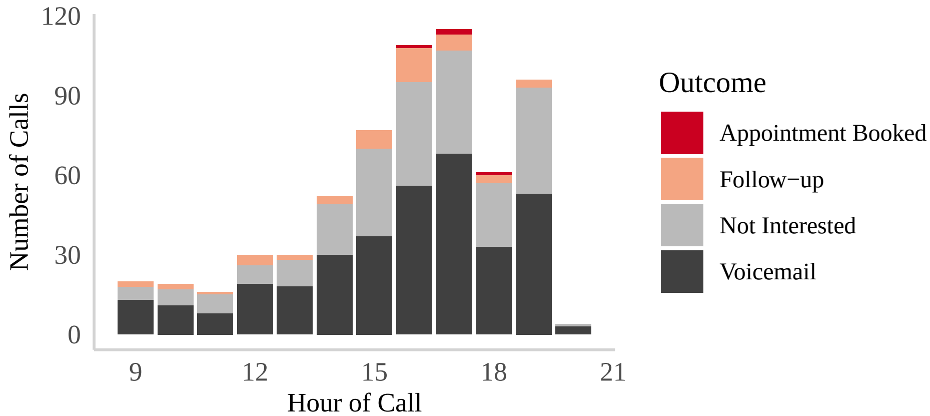

Figure 9: Outcome by Time of Call

The boxplot of call duration by outcome demonstrates a relationship between call length and call outcome (Figure 10). Instead of longer calls leading to appointments, it is more likely that the outcome itself determines the call duration—people who are more engaged and willing to book an appointment naturally spend more time on the call, while those who are not interested tend to end the conversation quickly. Thus, while this visualization does give insight into how call duration is correlated with different outcomes, we cannot use it as a valid explanatory variable to predict outcome.

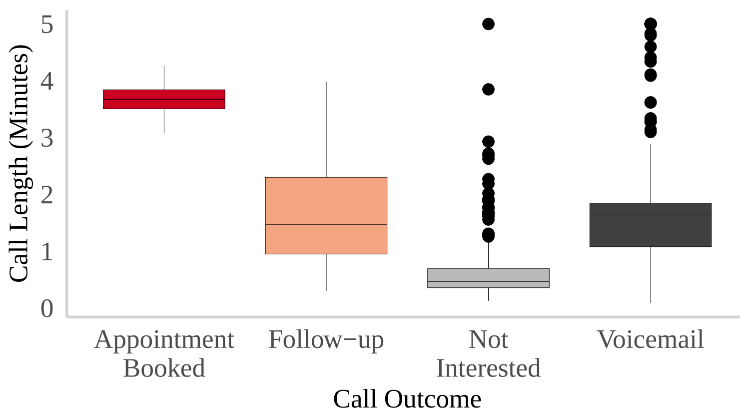

Figure 10: Outcome By Call Duration

4.6 'Appointment Booked' Outcome Models

4.6.1 'Appointment Booked' Logistic Regression

Table 1: Appointments Booked Outcome - SMOTE

term	Estimate	Std.Error	z value	p value	Significance	Odds
(Intercept)	-24.671	1199.594	-0.021	0.984		
Hour	0.982	0.107	9.186	0.000	***	2.669
Generalized	-2.084	0.350	-5.956	0.000	***	0.124
Consulting	2.364	0.318	7.437	0.000	***	10.636
Is.CEO	22.965	1199.590	0.019	0.985		
Is.COO	21.177	1199.589	0.018	0.986		
Wrong.number	-21.217	658.168	-0.032	0.974		
Timezone.Offset	2.203	0.221	9.991	0.000	***	9.053
Median.Income	0.181	0.016	11.480	0.000	***	1.199
Total.Population	0.009	0.014	0.650	0.516		
Healthcare.Spending	-1.707	0.169	-10.115	0.000	***	0.181
Healthcare.Growth	-0.036	0.776	-0.046	0.963		

The first model of primary interest was a logistic regression on Appointments Booked (1 if booked, 0 otherwise) using the SMOTE technique. Several variables are significant. We find that for every hour later the call is made, the odds of a successful appointment being booked increases by 2.669 which equates to a 166.9% increase in chance. Between the grouped segments, we find that specialized medical services have significantly higher odds than generalized medical services but lower odds than healthcare consulting businesses. This confirms our observations above that the highest proportion of appointments occur in healthcare consulting, specialized medical services, then generalized medical services.

Regarding geographic attributes, we do find that the Time Zone Offset is also a statistically significant predictor. A positive coefficient of 2.203 means that for every hour closer to GMT (every hour eastwards), we expect the odds of an appointment being booked increases by a factor of approximately 9. We do have to keep in mind that most calls did occur in the Eastern Time Zone, so this may be a result of the disproportionality of call distribution and sample size.

For every thousand increase in median incomes, we also see a ~1.2 increase in odds of appointments booked. This suggests that businesses operating within states of a comparatively wealthier population have higher chances of booking appointments. Interestingly, we see the opposite occur for healthcare spending per capita. For every thousand dollar increase in healthcare spending per capita, the log-likelihood of an appointment being booked decreases by 1.707. This is equivalent to a decrease in odds of 81.9%.

All other variables are not significant at a p-value threshold of 0.001. Though CEO's and COO's have large coefficients indicating higher log likelihood, they both have high p-values indicating non-significance from 0. However, we cannot interpret this as calling CEOs and COOs make no difference from calling other employees because our dataset largely consists of only C-suites, making comparison invalid. Being a wrong number also has a large negative coefficient, indicating a huge decrease in odds, but again, it is not statistically different from 0. Total population of the state and healthcare growth similarly have negligible influences on the odds of appointment booked, with small coefficients and large p-values.

4.6.2 'Appointment Booked' XGBoost Model

The XGBoost Model is able to fit the data extremely well, with only 3 false positives out of the 629 observations (Table 2; Appendix 6.3.1).

Table 2: 'Appointment Booked' XGBoost Model Confusion Matrix

	Actual 'No'	Actual 'Yes'
Predicted 'No'	621	0
Predicted 'Yes'	3	5

While we cannot directly measure the specific effect of each variable on the outcome due to the nonlinearity of the tree model approach, we are able to take a glimpse into which variables provide the most insight into classifying appointments (Figure 11; Appendix 6.3.2). Specifically, we see that the hour of the call was by far the most important feature, providing the largest improvement in prediction accuracy and affects a large proportion of splits. The next few features useful for prediction were all economic state-level data: median income, healthcare growth percentage, total population, and healthcare spending. These results suggest that regions with varying levels of healthcare investment are likely to react differently to AI calling services.

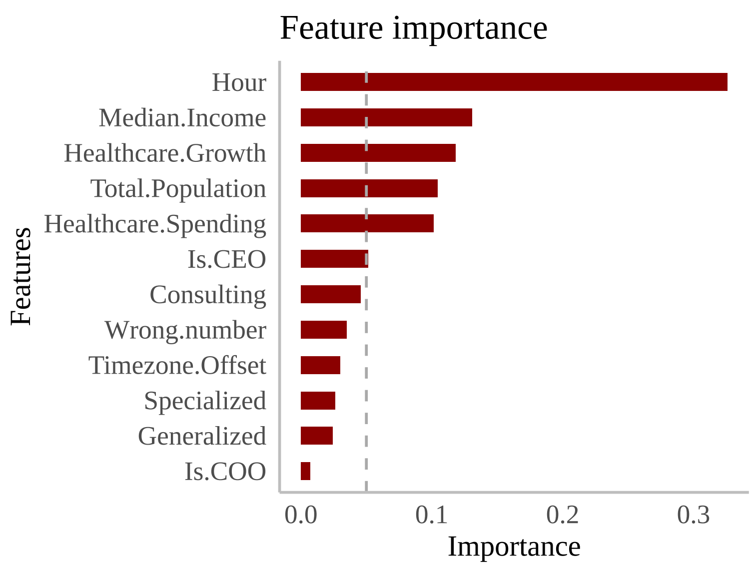

Figure 11: Feature Importance of 'Appointment Booked' XGBoost Model

The rest of the variables appear to be much less useful in model prediction, either falling at or below the 0.05 threshold. We find that calling the CEO is much less predictive than the broader economic and timing factors discussed above. Healthcare consulting appears to contribute much more to the accuracy of the model compared to the two other segments of specialized and generalized medical services. Calling the wrong number has a minimal influence on the outcome as well. All these results agree with the logistic regression results above.

4.7 'Is Interested' Outcome Modelling

4.7.1 'Is Interested' Logistic Regression

We were also interested in seeing how the logistic regression estimates would change when we include follow-ups as a successful outcome. We find a few notable changes. First, the hour the call is made is no longer significantly non-zero. As we saw earlier in Figure 9, follow-ups are fairly distributed across all hours, unlike appointments which were clustered in a singular area. Economic factors of healthcare spending and median income, which were useful in predicting appointments, are now no longer significant as well. However, we do find that total population is now positively significant, indicating businesses from states with larger populations are more likely to express interest.

Table 3: Interested Outcome - SMOTE

term	Estimate	Std.Error	z value	p value	Significance	Odds
(Intercept)	2.886	1.360	2.121	0.034	*	17.914
Hour	0.045	0.033	1.359	0.174		
Generalized	-0.312	0.159	-1.959	0.050		
Consulting	0.224	0.182	1.226	0.220		
Is.CEO	-0.449	0.273	-1.644	0.100		
Is.COO	-0.267	0.297	-0.902	0.367		
Wrong.number	1.450	0.163	8.888	0.000	***	4.261
Timezone.Offset	0.632	0.105	6.017	0.000	***	1.881
Median.Income	0.009	0.007	1.406	0.160		
Total.Population	0.023	0.007	3.082	0.002	**	1.023
Healthcare.Spending	-0.041	0.060	-0.696	0.487		
Healthcare.Growth	-0.154	0.313	-0.494	0.622		

A new variable of strong significance is Wrong Number. This seems very counterintuitive, as this indicates that reaching an unintended recipient leads to a higher likelihood of a follow-up. Sifting through call summaries made this phenomenon much more clear. Many follow-ups occurred when the call reached a receptionist or assistant, who would offer to pass the message along. Only a few were truly an accurate expression of interest in the AI calling services. Thus, we we have given more emphasis on the results of our prior "Appointments Booked" model than our "Is Interested" model in our later analysis.

4.7.2 'Is Interested' XGBoost Model

While the XGBoost Model is highly accurate in predicting if a call would result in interest, it is less accurate than predicting appointments. We see a higher false positive rate, with 40 misclassified negative observations (Table 4). Still, the high ROC curve (see Appendix 6.4.1) indicates a good fit.

Table 4: 'Is Interested' XGBoost Model Confusion Matrix

	Actual 'No'	Actual 'Yes'
Predicted 'No'	539	0
Predicted 'Yes'	40	50

The feature importance plot, however, is vastly different from the results observed in the logistic regression model. Rather, results are more similar to the logistic regression model and XGBoost model for appointments booked. The similar timing and economic variables prove to be the most useful in contributing to model accuracy.

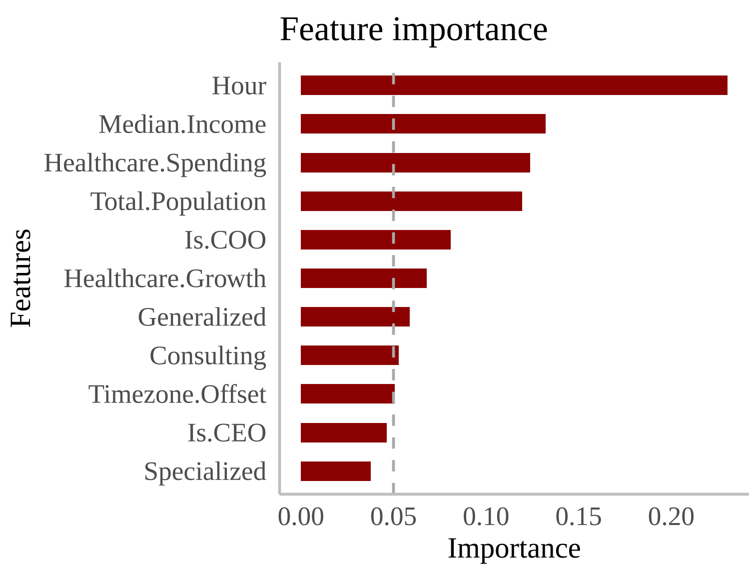

Figure 12: Feature Importance of 'Is Interested' XGBoost Model

4.8 'Wrong Number' Models

4.8.1 'Wrong Number' Logistic Regression

In an attempt to analyze mobile number accuracy, we ran a logistic regression model on calls where the intended recipient was not reached to pinpoint potential identifying characteristics. Note that this logistic regression model still used the SMOTE technique, as less than 20% of the completed calls made reached the wrong number.

Table 5: Wrong Number Model - SMOTE

term	Estimate	Std.Error	z value	p value	Significance	Odds
(Intercept)	-2.508	1.292	-1.941	0.052		
Hour	0.021	0.034	0.612	0.541		
Generalized	-0.336	0.157	-2.143	0.032	*	0.715
Consulting	-0.077	0.177	-0.437	0.662		
Is.CEO	0.049	0.276	0.177	0.859		
Is.COO	-0.114	0.302	-0.378	0.706		
Timezone.Offset	-0.006	0.090	-0.066	0.948		
Median.Income	0.001	0.006	0.167	0.868		
Total.Population	0.003	0.006	0.425	0.671		
Healthcare.Spending	-0.103	0.063	-1.647	0.099		
Healthcare.Growth	0.673	0.316	2.133	0.033	*	1.961

Only two variables are significant, but with fairly high p-values not less than our more conservative 0.01 threshold. Still, we can look to these as potential methods of improvement. Generalized medical services appear to decrease the odds of reaching the wrong number by 28.5% compared to calling specialized medical services, meaning that we tend to get higher number accuracy rates in that industry segment. It is likely that these businesses in the generalized medical services, such as hospitals, surgery centers, clinics, etc. have more standardized and public contact lists, enabling higher number accuracy. However, this is not likely to increase interested respondents. From our earlier models, we have determined that Generalized Medical Services are the least likely to book appointments out of all three business groupings. Thus, though calling businesses within Generalized Medical Services may increase number accuracy, we are likely not to see a proportional increase in appointments booked.

Meanwhile, for every percentage increase in healthcare growth rate, we find that the odds of calling the wrong number increases by 96.1%. In areas with quickly growing healthcare growth, there may be rapid organizational restructuring and expansion, leading to outdated contact lists and more wrong numbers. Thus, it appears that calling states with less rapid healthcare growth may increase call accuracy.

Results suggest that none of the other factors are statistically significant in predicting wrong numbers. This means that timing of the call, location of the call, and other economic factors will not help improve phone number accuracy. Unfortunately, this gives us little insight into improving phone mobile accuracy besides searching for the best API.

4.8.2 'Wrong Number' XGBoost Model

The XGBoost Model similarly shows a much poorer accuracy rate of predicting wrong numbers compared to before. We have 3 false positives and 75 false negatives as well as a lower ROC curve (see Appendix 6.5.1).

Table 6: 'Wrong Number' XGBoost Model Confusion Matrix

	Actual 'No'	Actual 'Yes'
Predicted 'No'	502	75
Predicted 'Yes'	3	49

In terms of feature importance, we find that the same variables as the previous models are still the most useful. The timing of the call, state-level information, and economic progress provide the most insight on predicting wrong numbers. The position of the intended recipient and segment of business are much less useful in prediction. Ultimately, this does not provide us much insight into the prediction of wrong numbers in correlation to the logistic regression above.

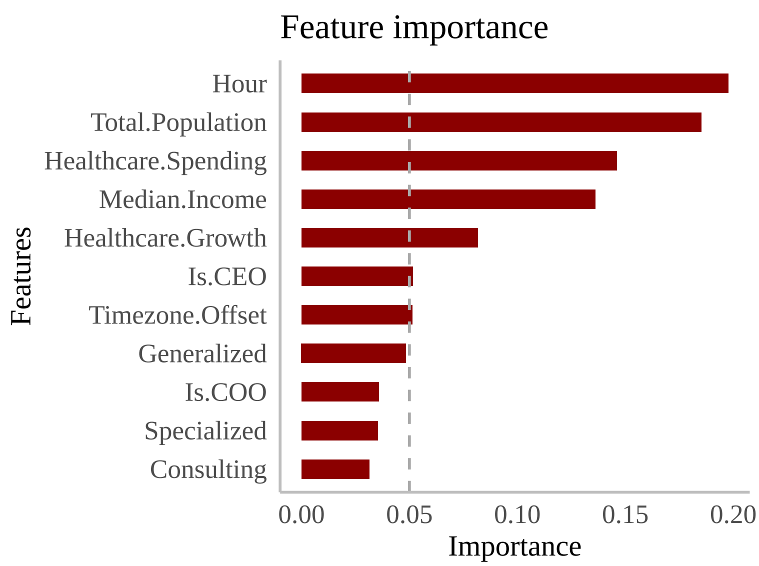

Figure 13: Feature Importance of 'Wrong Number' XGBoost Model

4.9 API Results

Based on our API research and evaluation, we recommend a workflow that prioritizes high accuracy and low API cost, while providing comprehensive data to improve outreach. This workflow includes four stages: lead sourcing, an initial validation step, comprehensive phone number validation, and finally real-time lead enrichment.

For the first stage of lead sourcing, we recommend using Apollo.io or Airscale.io, both of which offer highly accurate B2B contact data at a reasonable price. The paid plans for both APIs start at $49 per month, and Apollo offers a free tier for more limited usage. Our data analysis can help refine ICPs for optimal search criteria.

Next, we propose an initial validation step if the lead sourcing is done through a different, lower-cost API, or if additional accuracy is warranted. This initial validation step will filter out invalid or incorrectly formatted numbers, which will reduce unnecessary API calls and lower cost. We recommend using in-house scripts or basic validation tools to remove obviously invalid numbers before sending them to Twilio Lookup for the deeper phone number validation stage. Another option for this initial validation step would be to use NumVerify, a cost-effective API which would filter out invalid numbers and also provide line type information. Then, only the valid mobile numbers could be sent to Twilio Lookup to more accurately verify line type and determine line status. The first approach avoids any extra API costs but requires computational resources, while the second option introduces an additional API expense but improves accuracy. Both methods should theoretically result in a lower cost, as the number of requests to Twilio Lookup should decrease.

For deeper validation, we recommend to continue utilizing Twilio Lookup, which is regarded as the industry standard for phone number validation. This API provides Line Type service to confirm line type (landline, mobile, or VoIP) as well as a Line Status service to determine whether the number is active and reachable. Using this API will ensure that outreach efforts only focus on real, reachable prospects which will minimize unsuccessful calls. Specifically, the Line Type and Line Status services included in Twilio Lookup should prove to be highly effective for SalesMind AI's purposes.

Lastly, If additional data enrichment is wanted with recent company updates, we recommend using an API such as NewsData.io to scrape news sites for updates on high-value leads. Another option is to utilize an open-source LinkedIn scraper which would effectively function the same way. This stage will find noteworthy updates on leads, such as recent changes in leadership, major announcements, or funding rounds. This real-time data enrichment stage will provide more recent information to add to the more historical data found in lead sourcing APIs, which will allow for a personalized sales pitch to increase conversion rates.

Following this API workflow can reduce unnecessary costs, maintain data accuracy, validate phone numbers, enrich high-value leads, and ultimately improve call success rates.

5 Conclusions and Limitations

The results reveal factors that influence the success of AI sales calls, and there are strategies and drawbacks that stem from them. For example, grouping industries showed that the highest proportion of appointments booked were from Patient Care, Medical Facilities & Surgical Centers, and Healthcare IT & Digital Solutions. In contrast, generalized medical services were the least engaged, which suggests that moving towards specialized medical fields and consulting services may be better. Also, the data showed that reaching CEOs and COOs did not significantly impact appointment booking. It is difficult to make a conclusion from this finding because most people called were either CEOs or COOs, but it may help to contact people in other positions to provide a valid comparison test.

Additionally, targeting wealthier and more populated regions may enhance conversion rates, as the Median Income was a statistically significant and positive predictor for appointments booked whereas Total Population was significant for interested callers. Calls made further east in the United States, particularly in the Eastern and Central Time Zones, were more likely to result in an appointments. States with lower healthcare spending per capita also showed higher engagement, possibly revealing that underdeveloped healthcare markets are more receptive to AI-driven outreach.

The time of the call is a significant predictor of appointment success, as calling during the later business hours can lead to a higher rate of booking an appointment. Therefore, adjusting calling times to prioritize late afternoon outreach could improve engagement.

Our main constraint lied in our inability to verify the accuracy and precision of name and company data due to paywalls for most APIs. While the proposed three-step strategy for phone number validation and lead enrichment aimed to minimize unnecessary API requests, direct verification of API results would significantly improve data reliability. The high cost of API calls remains a challenge, but this research indicates that changing call timing and targeting specific demographics can also improve outreach efficiency.

Another major limitation was the use of state-level aggregated data, which is likely too broad for precise targeting of medical entities. Geographic and economic data at the zip code or census tract level would give a more detailed understanding of external factors influencing call outcomes. We could attempt to map mobile area codes, which are provided in the dataset, to zip codes, but it is not a one-to-one relationship and would result in a fairly messy combination. Additionally, mapping outreach efforts based on user's mobile area code versus business headquarters could enable us to gauge how remote work may influence call interest.

Finally, the analysis of follow-ups and wrong-number outcomes did not give useful insights. Many follow-ups were from unintended recipients, such as receptionists or assistants passing along messages, rather than from people who were interested. Similarly, wrong-number modeling did not provide meaningful indicators of number accuracy that could improve respondent engagement rates, suggesting that improving the contact data quality would require API validation rather than predictive modeling.

Overall, this research highlights the importance of strategic call timing, specific industry outreach, and prioritizing geography. While API validation is costly, it is a necessary component of data accuracy. These findings offer practical adjustments that can enhance engagement and optimize resource allocation for AI-driven sales.

6 Appendix

6.1 Broader Segments Breakdown

The following table shows the classification of specialized medical segments into broader groupings used for modelling.

Broader Segment	Specialized Segment
General Medical Services	Medical Facilities & Surgical Centers
	Medical Services
	Nonprofit & Community Health
Specialized Medical Services	Aesthetic & Wellness Services
	Dental & Vision Care
	Medical Research
	Medical Technology & Devices
	Mental & Behavioral Health
	Patient Care
	Pharmaceuticals & Prescription Services
	Rehabilitation & Physical Therapy
	Specialty Clinics
Healthcare Consulting	Healthcare Consulting & Workforce Solutions
	Healthcare IT & Digital Solutions

6.2 Full-Size State Level Maps

6.2.1 Total Calls By State

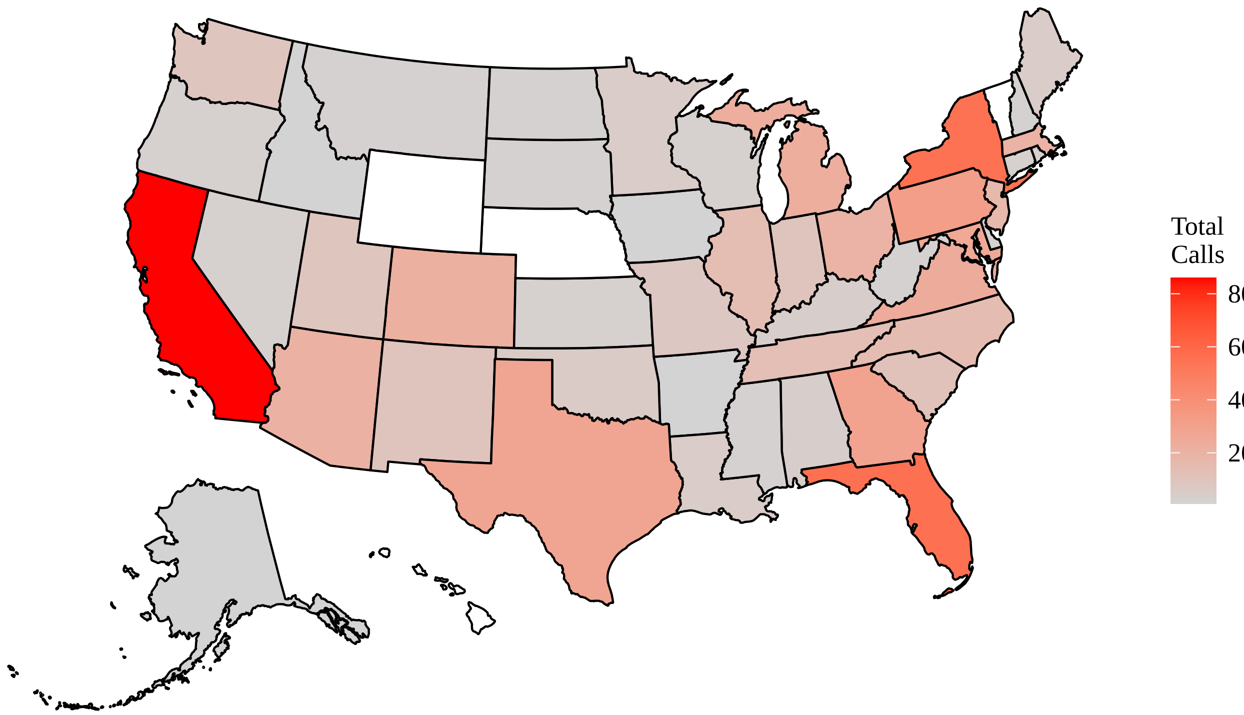

Figure 14: Total Calls by State

6.2.2 Interested Calls By State

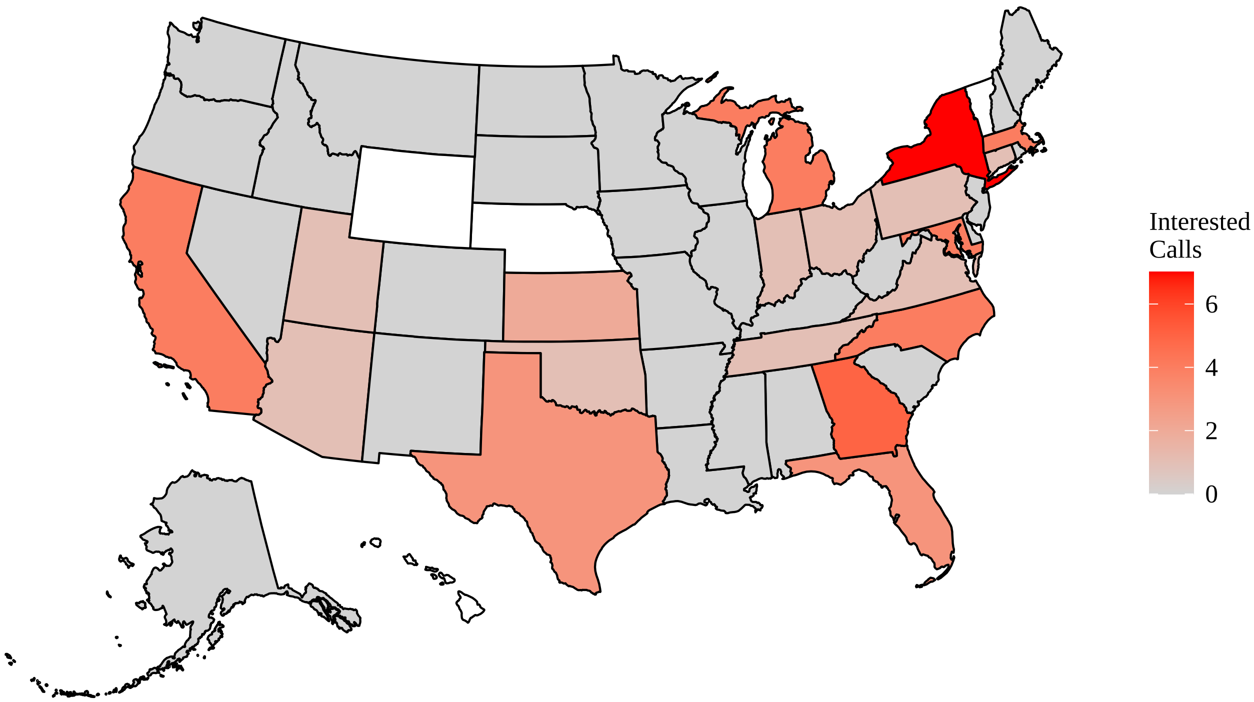

Figure 15: Interested Calls by State

6.2.3 Appointments Booked By State

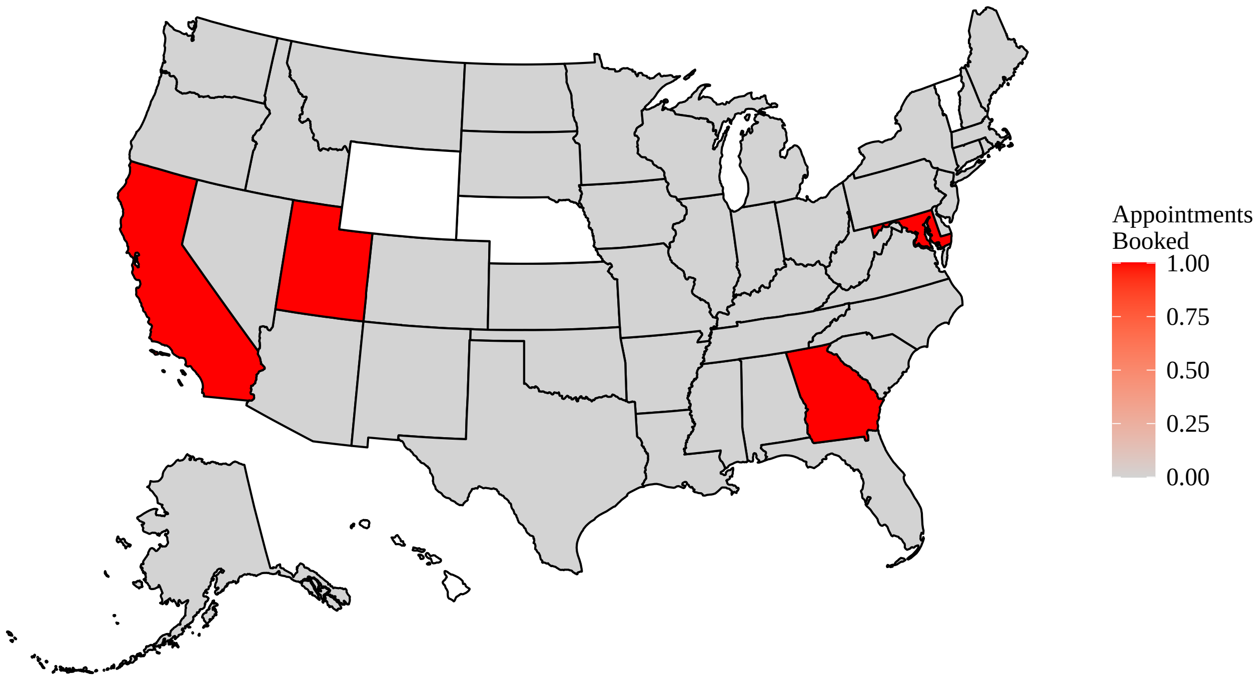

Figure 16: Appointments Booked By State

6.2.4 Median Income By State

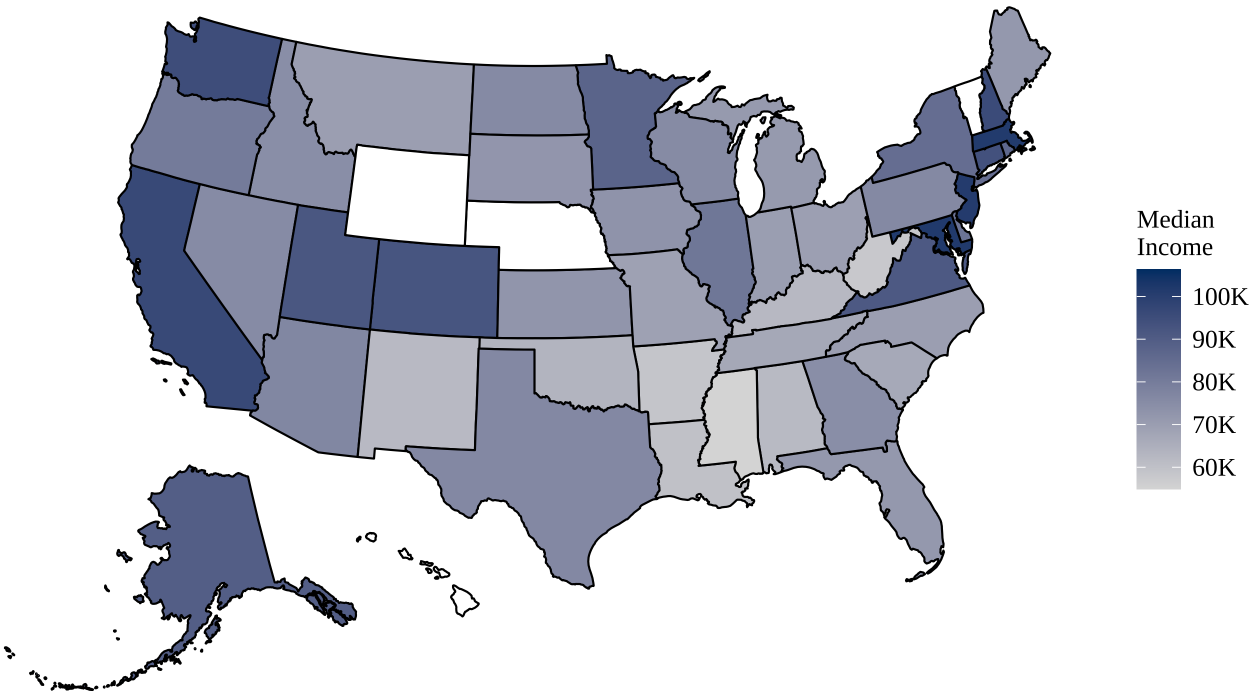

Figure 17: Median Income By State

6.2.5 Total Population By State

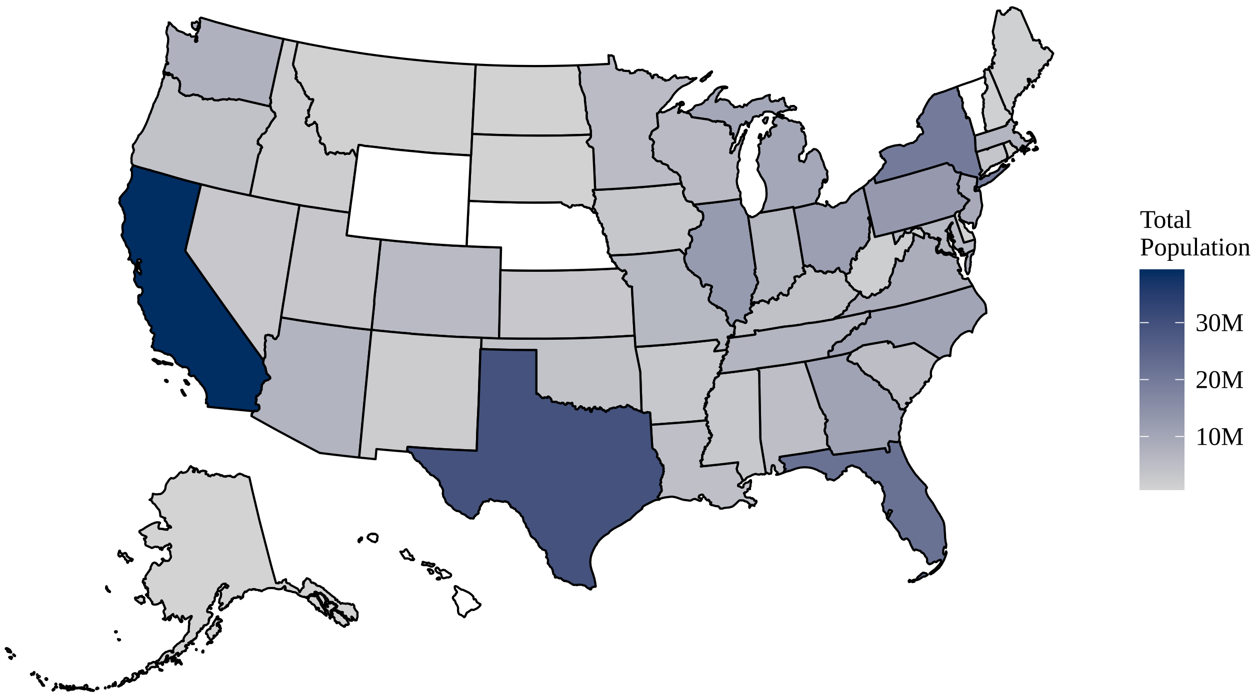

Figure 18: Total Population By State

6.2.6 Healthcare Spending Per Capita By State

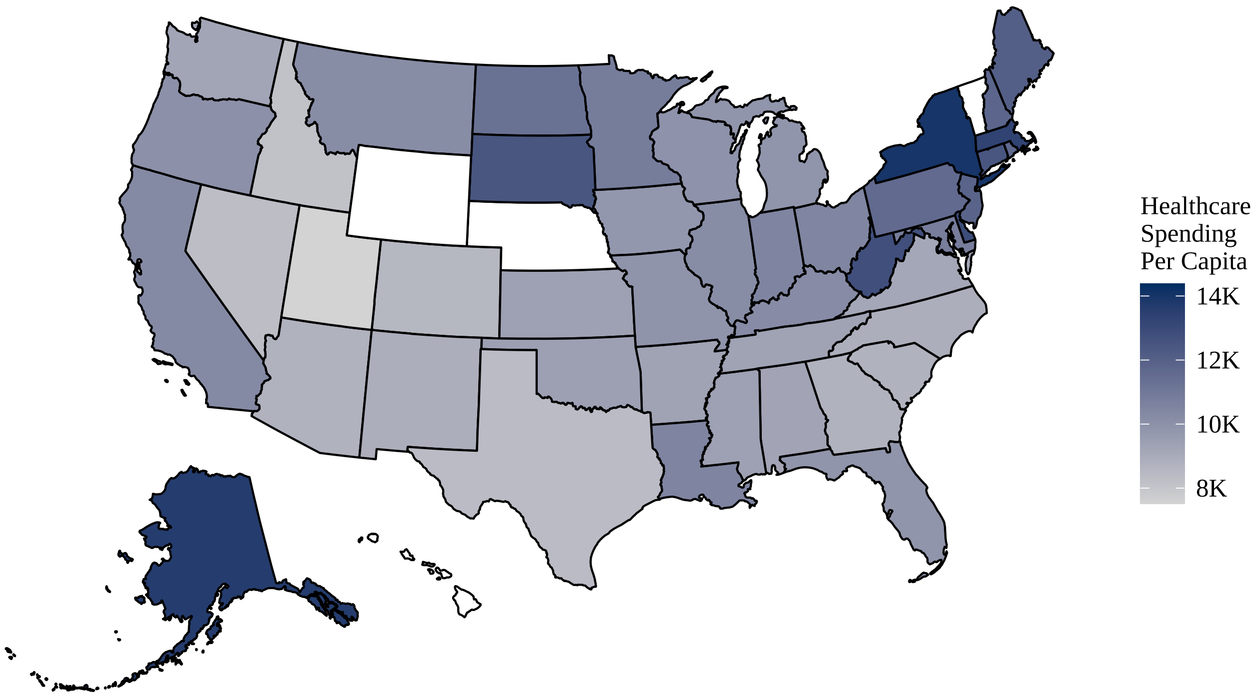

Figure 19: Healthcare Spending Per Capita By State

6.2.7 Healthcare Growth By State

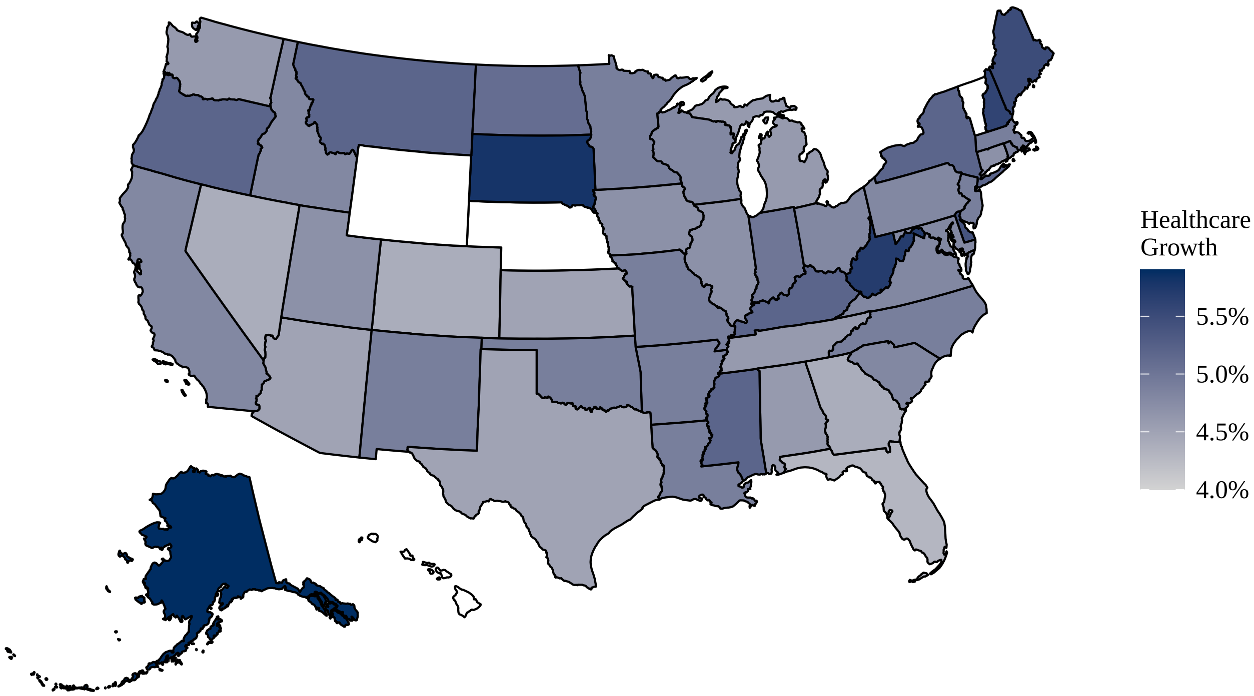

Figure 20: Healthcare Growth By State

6.3 'Appointment Booked' Diagnostic Results

6.3.1 'Appointment Booked' XGBoost Model ROC Curve

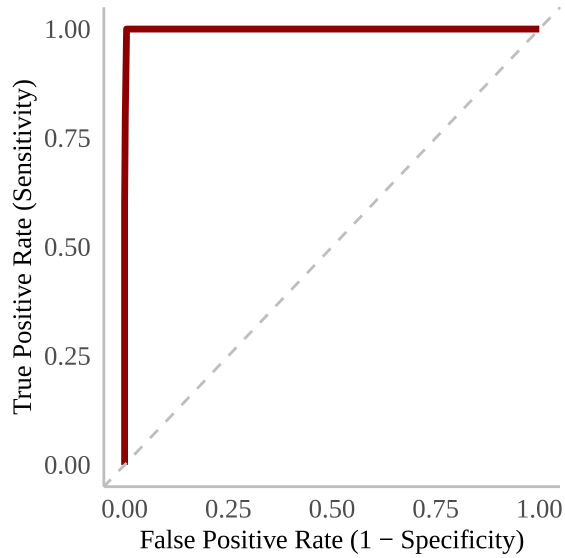

Figure 21: 'Appointment Booked' XGBoost Model ROC Curve

6.3.2 'Appointment Booked' XGBoost Model Feature Importance Matrix

Table 8: 'Appointment Booked' Model Feature Importance Table

Feature	Gain	Cover	Frequency	Importance
Hour	0.326	0.224	0.160	0.326
Median.Income	0.131	0.118	0.147	0.131
Healthcare.Growth	0.118	0.115	0.078	0.118
Total.Population	0.104	0.103	0.127	0.104
Healthcare.Spending	0.101	0.148	0.115	0.101
Is.CEO	0.052	0.053	0.036	0.052
Consulting	0.046	0.055	0.096	0.046
Wrong.number	0.035	0.063	0.021	0.035
Timezone.Offset	0.030	0.039	0.086	0.030
Specialized	0.026	0.022	0.044	0.026
Generalized	0.024	0.043	0.044	0.024
Is.COO	0.007	0.017	0.045	0.007

6.4 ‘Is Interested’ XGBoost Model Diagnostic Results

6.4.1 ‘Is Interested’ XGBoost Model ROC Curve

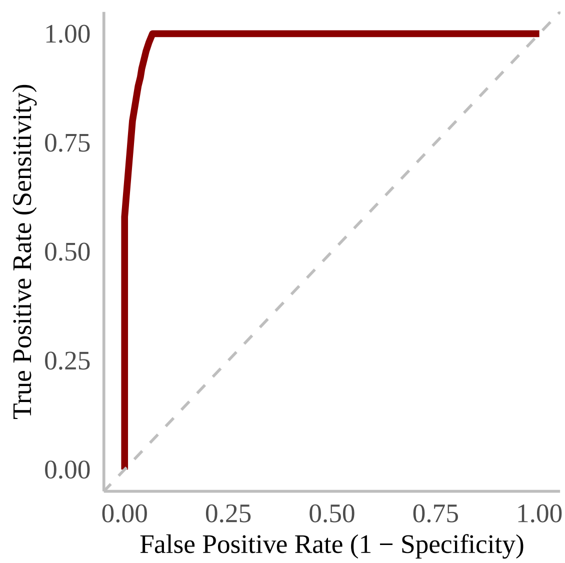

Figure 22: ‘Is Interested’ XGBoost Model ROC Curve

6.4.2 ‘Is Interested’ XGBoost Model Feature Importance Matrix

Table 9: ‘Is Interested’ Model Feature Importance Table

Feature	Gain	Cover	Frequency	Importance
Hour	0.230	0.204	0.235	0.230
Median.Income	0.132	0.139	0.132	0.132
Healthcare.Spending	0.124	0.136	0.107	0.124
Total.Population	0.119	0.156	0.118	0.119
Is.COO	0.081	0.050	0.059	0.081
Healthcare.Growth	0.068	0.096	0.079	0.068
Generalized	0.059	0.045	0.065	0.059
Consulting	0.053	0.039	0.057	0.053
Timezone.Offset	0.051	0.059	0.042	0.051
Is.CEO	0.046	0.038	0.058	0.046
Specialized	0.038	0.037	0.048	0.038

6.5 'Wrong Number' XGBoost Model Diagnostic Results

6.5.1 'Wrong Number' XGBoost Model ROC Curve

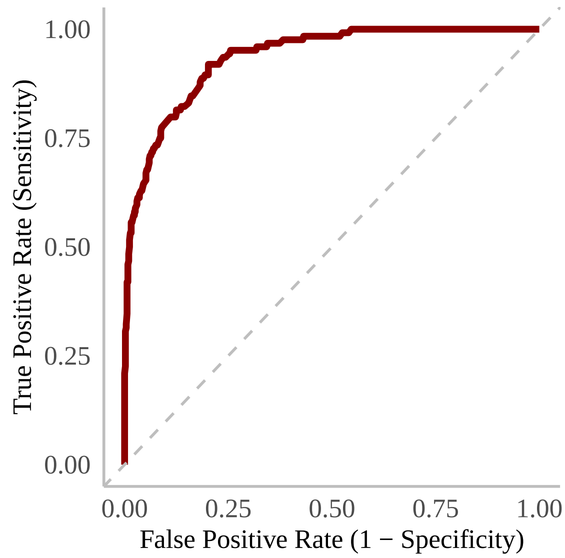

Figure 23: 'Wrong Number' XGBoost Model ROC Curve

6.5.2 'Wrong Number' XGBoost Model Feature Importance Matrix

Table 10: 'Wrong Number' Model Feature Importance Table

Feature	Gain	Cover	Frequency	Importance
Hour	0.198	0.195	0.203	0.198
Total.Population	0.185	0.206	0.164	0.185
Healthcare.Spending	0.146	0.141	0.129	0.146
Median.Income	0.136	0.138	0.132	0.136
Healthcare.Growth	0.082	0.104	0.084	0.082
Is.CEO	0.051	0.041	0.055	0.051
Timezone.Offset	0.051	0.053	0.049	0.051
Generalized	0.048	0.039	0.054	0.048
Is.COO	0.036	0.029	0.037	0.036
Specialized	0.035	0.030	0.050	0.035
Consulting	0.031	0.024	0.042	0.031

References

AltaHQ. 2025. "Alta: Private Market Investment Platform." https://www.altahq.com/.

Centers for Medicare and Medicaid Services. 2020. "State Health Expenditure Data." https://www.cms.gov/data-research/statistics-trends-and-reports/national-health-expenditure-data/state-residence.

Factors.AI. 2025. "AI-Powered Sales Solutions." https://www.factors.ai/solutions/sales.

HubSpot. 2025. "HubSpot AI: Artificial Intelligence for Marketing & Sales." https://www.hubspot.com/products/artificial-intelligence.

NumVerify. 2025. "NumVerify: Free Phone Number Validation & Lookup API." https://numverify.com/.

People Data Labs. 2025. "Sales & Marketing Data Solutions." https://www.peopledatalabs.com/use-cases/sales-and-marketing.

SalesMind.ai. 2025. "SalesMind AI Partners with UCLA Professor Dave Zes and Statistics Student Teams to Solve Lead Generation Data Challenges." *Send2Press Newswire*. https://www.send2press.com/wire/salesmind-ai-partners-with-ucla-professor-dave-zes-and-statistics-student-teams-to-solve-lead-generation-data-challenges/.

Twilio. 2025. "Twilio Lookup: User Authentication & Identity." https://www.twilio.com/en-us/user-authentication-identity/lookup.

U.S. Census Bureau. 2023. "U.S. Census Bureau Data Tables." https://data.census.gov/all.