## Introduction

Efficient resource management is a critical success factor for any organization. By better understanding the organization’s staff, time, and other supporting resources, an organization can improve performance while improving operational efficiency at the same time. The Small Business Development Centers (SBDCs), a nationwide business consulting network funded by the Small Business Administration to support small business across the country. There is at least one SBDC center in almost all states, each is tasked with serving the businesses within that state. Furthermore, those state-level centers are split into multiple regional centers to better serve the local businesses in those areas. This structure ensures that all small businesses have access to the consulting services needed to develop as a business. However, with the limited resources available, in addition to those resources being funded by taxpayer’s money, it is crucial that those centers allocate their staff, time, and efforts accordingly to maximize the impact they have on local businesses.

In collaboration with the Nebraska SBDC, this project aims to analyze the Nebraska SBDC centers’ resources, focusing on staffing numbers, consulting hours, and performance metrics by looking at client milestone successes. I aim to find insights into resource allocation and client success by developing an interactive app designed to serve as a tool for supporting data-driven decision making. This paper outlines the analysis process, the development of the tool, key findings, and opportunities for further research.

## SBDC Human Capital

### Data Pre-Processing

The initial SBDC human capital dataset was gathered through a survey of center directors from 8 states and a total of 61 regional centers. The survey aimed to capture both center-specific data, such as the number of staff, number of clients, total counseling hours, and availability of HR support. Additionally, the survey captures regional data, including the population of the region served and whether the center is in an urban or rural area. However, the dataset provided was in a wide format, where each row represented a state, and then multiple columns contained the survey answers for the individual centers within that state. For instance, we have a row dedicated to all Nebraska centers, then we have 10 columns corresponding to each survey questions for the first Nebraska center. Following that, we have another 10 columns corresponding to the survey questions of the second Nebraska center, so on and so forth. The first step in the data pre-processing process involved transforming this wide-form dataset into a long form dataset, where each row represented a single center and a column for each of the survey answers for that specific center, in addition to two identifier columns, including the state name and the center number.

Another critical step in the pre-processing phase of this project was standardizing the dataset. Since the survey was designed as an open-ended survey for all questions, that led to inconsistencies in the responses across the 61 centers given we had 61 different center directors answering this survey. To do good analysis and generate quality results, I ensured uniformity across all the variables. For instance, while some centers reported part-time staff as 0.5 and full-time staff as 1.0, others did not distinguish between employment types. Even though a small sample of centers used this method of differentiating between full-time and part-time employees, I used this method across the dataset following the guidance of the Nebraska center director. Another example of inconsistent survey responses is in the urban/rural center question where we received different answers many variations, such as “urban,” “Urban,” “mostly urban,” etc. I manually standardized these answers into binary categories, encoding urban centers as 1 and rural centers as 0. For variables like HR support, responses such as “yes,” “Yes,” “no,” and “No” were similarly standardized to binary values, where 1 represented “Yes” and 0 represented “No.”

Table . Human Capital Standardized Variables

|  |  |
| --- | --- |
| **Variable Name** | **Description** |
| State | The state where the SBDC center is located. |
| Center Number | A center identifier for each regional center within a state. |
| Total Staff | Number of employees in a center. 1 for a full-time staff and 0.5 for a part-time staff. |
| Total Counseling Hours | Total preparation + contact time recorded for all the consultants within the center. |
| HR Support | Indicate if a center has HR support. 1 if it does and 0 if it does not. |
| Population | The population of the region the center serves. |
| Has employees reporting | Indicates if a center has employees beyond the center director. 1 if it does and 0 if it does not. |
| Program-Related Activity Time | Number of hours a consultant spent on activities other than client consulting, such as consultant training and client outreach. |
| Is Urban | Indicates if the center is in an urban or rural area. 1 if it is in an urban area and 0 if it is in a rural area. |

Table 1. shows all the variables in a standardized form. This pre-processing phase is a critical part of this project, and any misstep here will negatively impact the analysis and the results. Given that I am not a domain expert, I discussed what I did in this step with the Nebraska SBDC state director to ensure that my interpretation and processing of the data aligned with the intended meaning of the survey responses.

### Counseling Hours Estimates

With the dataset pre-processed and standardized, I moved forward with my analysis and development of an interactive map. The first objective was to predict the expected counseling hours for each center using a machine-learning based approach. To do so, I explored different machine-leaning and non-machine-learning approaches, including a simple Non-Negative Least Squares (NNLS) regression model, an Extreme Gradient Boosting (XGBoost) regressor, a Random Forest Regressor, and a sensitivity analysis.

To compare between those methods, I used the Root Mean Squared Error (RMSE) as the primary comparison metric, which captures the magnitude of prediction errors.

Table . Counseling Hours RMSE Comparison

|  |  |
| --- | --- |
| **Method** | **RMSE** |
| NNLS | 585.17 |
| Random Forrest Regressor | 199.9 |
| XGBoost | 828.59 |
| Sensitivity Analysis | 276.65 |

The result of the comparison is shown in Table 2. where we can clearly see that the Random Forrest Regressor significantly outperformed the other methods with the lowest RMSE of 199.0. This shows that the random forrest regressor does the best job at capturing linear and non-linear relationships and interactions between variables in our dataset. Therefore, I moved forward with this method to predict expected counseling hours for each center. The results were then used to calculate additional insights, such as the expected hours per consultant and the expected hours per consultant.

One of the advantages of using a tree-based model, such as the Random Forest Regressor, is its ability to provide variable importance data, which offers valuable insights into the factors driving the model's predictions.

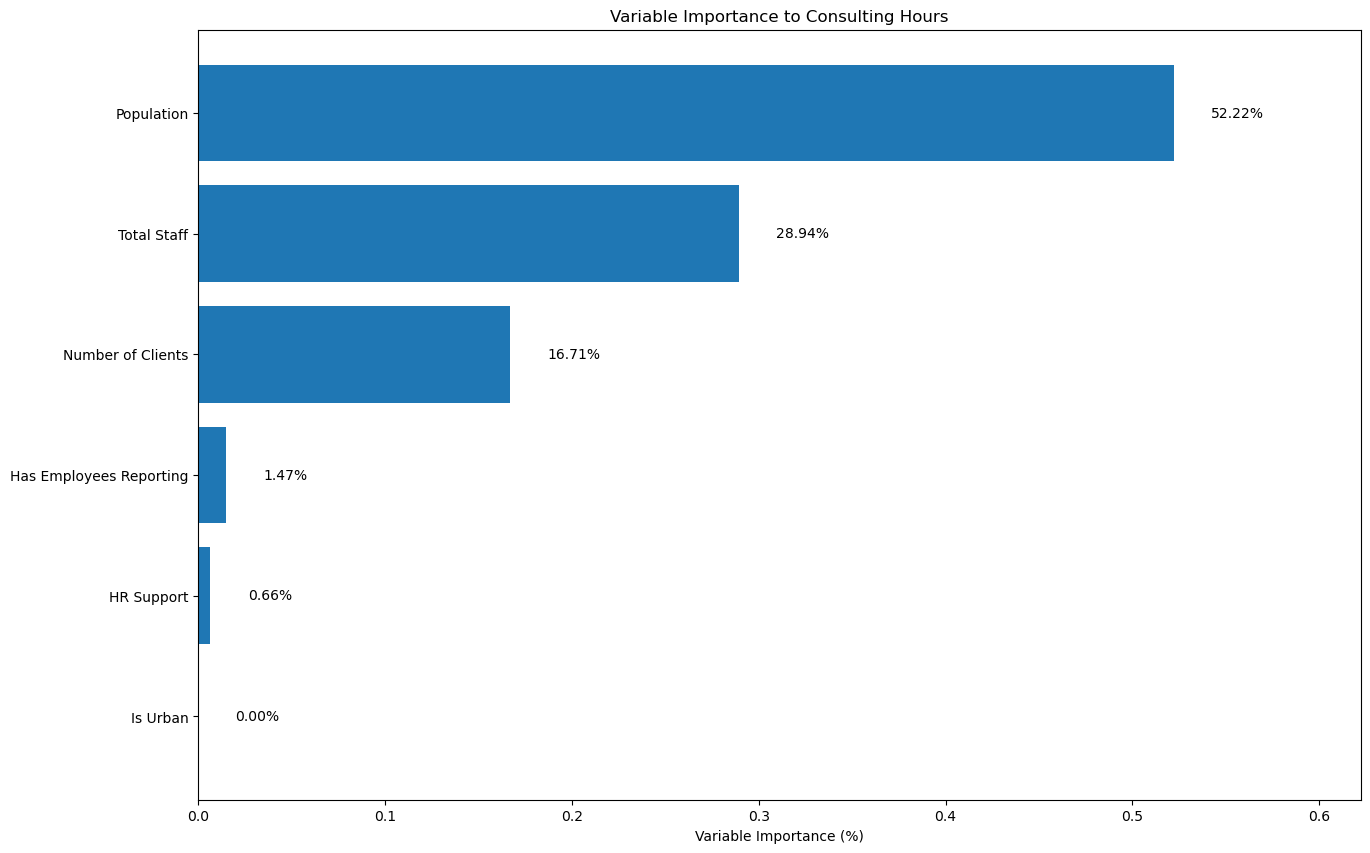


Figure . Random Forrest Regressor Variable Importance

Figure 2. shows all the variables used in our predictions and how important each is to the final predictions. As shown, the majority of the output can be explained by the population variable as it has 52.22% of the total importance. This suggest that the population of the region each center serves players a huge role in estimating the number of hours that center is expected council, this is likely because of the expected demand of a higher population drives the service hours of a center. Total Staff, contributing 28.94%, is the second most influential variable, reflecting the direct relationship between staffing levels and the ability to dedicate more time to consulting with clients. The third most influential variable is the number of clients with an importance of 16.71%, which is the first surprise as my initial hypothesis was that the number of staff would be the most important variable. However, it seems that it is significantly less influential than the population.

On the other hand, variables such as “Has Employees Reporting,” “HR Support,” and “Is Urban” have low to no influence on our estimates. Therefore, this reveals that we can make quality estimates using the population size, number of employees, and number of clients. I believe that these findings can be useful for the SBDC, as they highlight actionable strategies for increasing counseling hours. For instance, centers could consider increasing staffing levels, expanding client outreach to serve more businesses, or adjusting center regional boundaries to account for population changes and better align resources with demand.

Those results were incorporated into the interactive app for the Nebraska SBDC to conduct and present the same analysis in a dynamic and user-friendly way. The cleaned and standardized dataset, which contains important variables, such as the Total Staff, Number of Clients, Population, Total Counseling Hours, HR Support, and Is Urban is fed into the program. Based on each center's distinct features, the app uses this information to estimate the anticipated counseling hours using the Random Forest Regressor model.

Once the user inputs the data, the app returns a range of output as a result of the analysis. First, it displays the forecast for each center's anticipated counseling hours as a table. Apart from these forecasts, the app computes two crucial performance indicators: Hours Per Consultant, which is the average number of counseling hours divided by the number of employees to evaluate workload distribution, and Hours Per Client, which divides the counseling hours by the total number of clients to determine the average amount of attention each client receives. The user then has the option to download this data for further analysis.

In addition to the estimates, the app visualizes those results by comparing the centers within state and states with one another. Users can choose a state from a drop-down menu, and the app will create a bar plot at the center level that shows the projected hours per consultant for each center in that state. Comparing centers and identifying ones that could need more resources or modifications is made simple by this tool. A high-level overview of resource consumption patterns is also provided by the app, which offers state-level reports that show the average hours per consultant and hours per client across all states.

A graph of blue rectangular bars

Description automatically generated

Figure . SBDC Human Capital App Output (Estimated Hours Per Consultant for Nebraska Centers)

Figure 2. shows a sample output if the user selects “Nebraska” from the drop-down menu. This shows the distribution of counseling hours throughout the several regional centers in Nebraska. The height of each bar, which represents a particular center, shows the average number of counseling hours per consultant at that center. Centers like Kearney and Wayne, for example, have far greater hours per consultant (590.88 and 533.18 hours, respectively), suggesting that consultants there are handling a heavier burden than consultants in other places.

On the other hand, centers like Scotts Bluff (167.76 hours) and North Platte (156.16 hours) have fewer average hours per consultant, indicating that their workloads may be lighter. Decision-makers can more easily identify centers that could require increased staffing support or resource reallocation by using this kind of visualization, which helps highlight potential discrepancies in resource consumption within the state.

A blue and purple bars

Description automatically generated

Figure . App Output for The State-Level Comparisons

Figure 3. shows the state-level comparisons from the app’s output. In the “Average Hours Per Consultant by State” plot (left), the graph shows how Nebraska is in the middle of the pack when compared to other states. Nebraska reports 349.01 hours per consultant. This suggests that Nebraska's consultants may have a smaller burden than those in North Dakota (874.64 hours) and Wyoming (575.33 hours). In contrast, Nebraska has the greatest average number of hours per client (5.82), out of all the states, in the "Average Hours Per Client by State" plot (right). This suggests that although Nebraska consultants work fewer hours overall, they spend a lot more time with each client. This balance points to a significant emphasis on offering clients comprehensive support, which may help them succeed in the state.

### Ideal Number of Staff

Following estimating the counseling time per center, the Nebraska SBDC wanted to gage their staffing levels to see if they are over-staffed, under-staffed, or evenly staffed. This section of the analysis uses a sensitivity analysis approach to determine the optimal number of employees for each center. The objective is to determine how staffing levels should change to accommodate different needs for the number of clients, counseling hours, and geographical population.

In this analysis, we start by selecting our input variables that will estimate the ideal number of staff per center. Those variables are “Total Counseling Time,” “Population,” “Number of Clients,” “Is Urban,” “HR Support,” and “Has Employees Reporting.” Once we have selected our input variables, to deal with skewed data and better the models stability, in the analysis we start by applying a log transformation for variables such as Population and Total Counseling Time. To preprocess the data and forecast the total number of employees required for each center, a machine learning pipeline is established. Sensitivity analysis is used to account for little changes in inputs in order to guarantee reliable outcomes.

Key predictors (counseling hours, population, and client numbers) are gradually modified by percentages ranging from -10% to +10% in the sensitivity analysis. The pipeline forecasts the number of employees needed in the simulated environment for every modification. The Ideal Staff for each facility is then determined by averaging these forecasts. This approach ensures accurate and reasonable staffing estimates by taking into consideration uncertainties and slight variations in workload.

These results are also incorporated in the app for the Nebraska SBDC to utilize. The sensitivity analysis for staffing is carried out by the app using the same pre-processed dataset that was used for the counseling hours estimates. A straightforward dataframe including the ideal staff values for each center as well as the present staff levels is used to illustrate the findings. This makes it simple for users to spot any discrepancies between the ideal and existing workforce numbers. The app also offers the ability to download the data as an Excel file for additional analysis.

## SBDC Client Milestone Analysis

Besides analyzing the SBDC centers' human capital and resource allocation, the Nebraska SBDC requested me to calculate the number of sessions and hours a consultant must work with a client until the client is successful. A client's success is determined by achieving specific milestones. We have access to two different kinds of milestones in this project: capital milestones, which indicate that customers have secured financial funding, and company established milestones, which indicate that clients have successfully launched a business. In this part I will analyze the datasets that have been supplied to evaluate the link between counseling sessions and milestone achievement.

### Client/Milestone Datasets

To address the questions regarding how much a consultant has to spend with a client before they are successful, three datasets were provided by the Nebraska SBDC:

1. Counseling Hours: This dataset includes the client counseling sessions for all seven Nebraska SBDC centers over the previous ten years. It comprises all of the time consultants have spent with clients, broken down into other categories such session date, contact hours, and preparation.
2. Capital Milestone: This contains all the client capital-related milestones over the past 10 years. It contains details on the milestone dates, the clients' identities, and the amount of capital funding received.
3. Business Established: A milestone record showing when clients launched a business in the past ten years. Client identities and milestone dates are included.

### Capital Milestone Analysis

In this part of the analysis, we use both the Consulting Hours Dataset and the Capital Milestones Dataset, previously mentioned, and merges them on the Client ID to associate client session records with their capital funding milestones. To concentrate on the period preceding funding, we refine the data to encompass only consulting sessions that took place on or before the Capital Reporting Date. This guarantees that the analysis encompasses the aggregate number of sessions and counseling hours dedicated to clients before obtaining funding.

Once the datasets are merged, we aggregate the outcomes by Client ID, Funding Type, and Date. This compilation summarizes the total sessions and hours spent with clients monthly, classified by the type of funding received. The users (Nebraska SBDC) can utilize this analysis using another part of the app interface. After uploading the datasets, the application produces a summary table displaying Client ID, Funding Type, Month-Year, Total Sessions, and Total Hours. Users can interact with the analysis by applying filters for date range and funding type, with the ability to select multiple funding types or to include all funding types by default.

Following the application of the filters, the app computes high-level key metrics, such as the number of milestones within the specified range, the aggregate average of sessions, and the total client counseling hours average before the capital funding milestone. The app produces two visual representations: a line graph showing average sessions over time and another line graph illustrating average hours over time. These visualizations make it easier for users to identify the trends in customer contact preceding capital milestones. Users also have the ability to download the filtered dataset and the plots for additional reporting or analysis.

Table . Capital Funding Milestone Analysis Results

|  |  |
| --- | --- |
| Number of clients that received capital funding | 6,860 |
| Average number of sessions | 2.83 |
| Average counseling hours | 4.98 |

Table 3 describes the counseling efforts necessary for assisting clients in attaining capital funding milestones. The data indicates that 6,860 clients successfully obtained financial funding. Clients attended an average of 2.83 sessions, accumulating a total of 4.98 counseling hours prior to obtaining funding. These findings underscore the effectiveness of the SBDC's consulting efforts, illustrating that a relatively small investment of time per client may produce substantial results. The data in Table 3 can act as a standard for centers to enhance their resource planning and allocation. By comprehending these averages, SBDC centers can align consultant efforts with client requirements, improve their consulting methodologies, and sustain balance between resource allocation and client achievement.

### Business Established Milestone Analysis

The business established milestone analysis parallels the methodology used for capital funding milestones, although it highlights understanding the time commitment necessary for clients to effectively launch a business. This research employs the Consulting Hours and Business Established datasets, which are merged using the Client ID. This ensures that only consulting sessions prior to the client's business establishment date are incorporated in the study.

The analysis begins by aggregating the overall number of sessions and total counseling hours dedicated to each client before their business establishment milestone. The values are later aggregated and averaged weekly to analyze trends over time. Similar to the capital funding analysis, the application allows users to filter the data using a date range filter, allowing targeted examination of specific time intervals. When the application of filters and the generation of results, the app presents three primary insights: the total count of milestones, denoting the number of clients who successfully launched a business within the specified timeframe; the average number of sessions, reflecting the typical number of consulting sessions required; and the average counseling hours, signifying the mean duration spent with each client before achieving the business establishment milestone.

To enhance the interpretability of the results, the application employs two-line plots for data visualization. The initial plot depicts the average number of sessions over time, while the plot that follows represents the average counseling hours over time. These visualizations facilitate the identification of trends and consulting patterns linked to successful commercial enterprises. The program allows users to download the dataset as and export the created plots for further research or sharing.

Table . Business Established Milestone Analysis Results

|  |  |
| --- | --- |
| Number of clients with a business established milestone | 1,006 |
| Average number of sessions | 3.11 |
| Average counseling hours | 2.23 |

The findings from the milestone analysis of the business, presented in Table 4, provide important insights into the consulting activities necessary for customers to effectively develop a business. A total of 1,006 milestones were documented during the specified time frame, indicating the number of clients that attained this milestone. Clients, on average, engaged in 3.11 sessions prior to effectively launching their firm. The sessions collectively totaled an average of 2.23 counseling hours for each individual.

In comparison to the capital milestone results in Table 3, the business established milestone requires a somewhat greater average number of sessions (3.11 vs. 2.83) but much less total counseling hours (2.23 vs. 4.98). The difference indicates that initiating a firm typically requires more frequent yet brief consultations with consultants, while ng financial backing usually requires longer, more extensive counseling sessions.

These findings underline the complex nature of the counseling services offered by SBDC. The difference in time commitment illustrates the unique problems and procedures linked to achieving each category of milestone. For business entities, regular follow-ups and direction might be enough, whereas funding requires more extensive and time-intensive assistance to get financial resources. This differentiation offers essential context for resource distribution and strategic planning, enabling centers to enhance their initiatives in assisting clients to attain certain milestones efficiently.

## Conclusion

This research effectively analyzed the resource allocation of Nebraska SBDC centers and its correlation with client success by evaluating staffing levels, consulting hours, and significant milestones. The project utilized an extensive dataset and created an interactive tool, providing vital insights about the average consulting duration necessary for customers to reach critical milestones, including securing capital backing or launching a business. The results indicated that capital fundraising milestones need fewer sessions with extended counseling hours, while startup milestones comprised shorter, more frequent sessions. These findings can assist SBDC centers in optimizing their resources, ensuring that consultants' time and efforts are matched with the distinct needs of clients and the objectives they seek to accomplish.

The interactive application created for this project functions as a useful tool for facilitating data-driven decision-making at the Nebraska SBDC. The application allows users to assess consulting hours, staff productivity, and customer milestones through diverse filters, hence augmenting center directors' capacity for more efficient resource allocation. This study highlights the significance of effective resource management in attaining optimal impact with constrained resources. Subsequently, these insights and methods can be applied to additional SBDC networks nationwide, facilitating future research into resource management and performance enhancement among comparable organizations.