Project Report: 2019 ABS

Group 5: Albert, Meghan, Sarvani, Stephen

Introduction

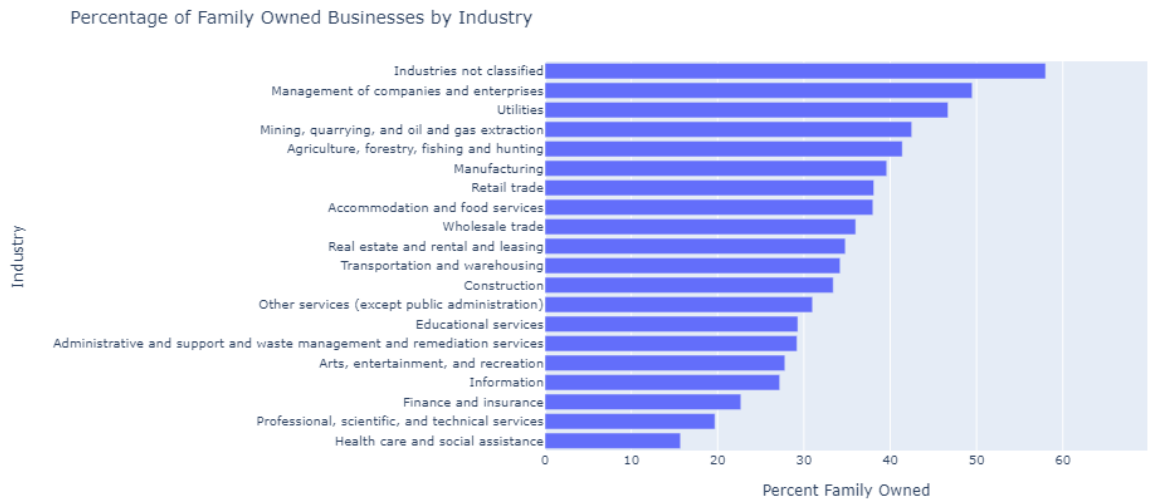
Starting in 2007, the Annual Business Survey (ABS) has been conducted jointly by the U.S. Census Bureau and the National Center for Science and Engineering Statistics. The 2019 survey samples approximately 300,000 randomly selected non-farm employer businesses filing the 941, 944, or 1120 tax forms within the United States. This project focuses on three of the four available data topics: company summary, business characteristics, and owner characteristics.

This paper is divided into four sections, each section is respectively guided by the following questions:

1. What factors impact how likely a business is family owned? Are there certain characteristics family owned businesses share?
2. What is the distribution of education level and fields of study among business owners? How does this distribution change when veteran status is considered?
3. Can anything interesting be gleaned from the veteran data with respect to the level of education (a second look), sex, or age of the veteran business owner respondents?
4. How does industry, location, and race affect a company's annual payroll?

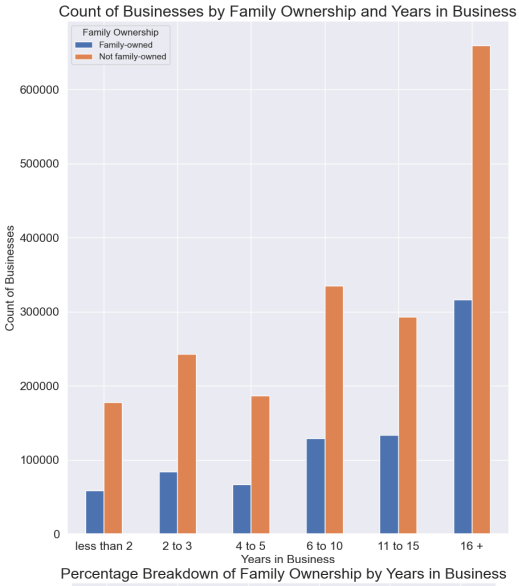
Family Ownership

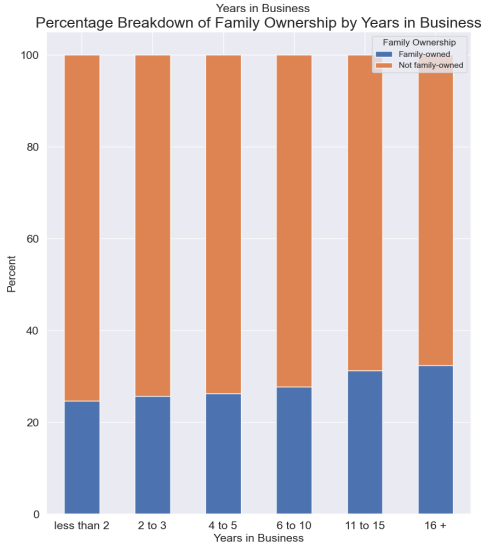
Our group was interested in examining how the prevalence of family ownership in businesses varies by industry, age of the business, sex of the majority owner, and geography. To answer these questions, two queries with different parameters were called on the characteristics of businesses API. Details of these queries can be found in the ETL report. It is important to note that the following charts exclude observations where family ownership was not reported or not applicable. All observations either reported as family owned, or not family owned.

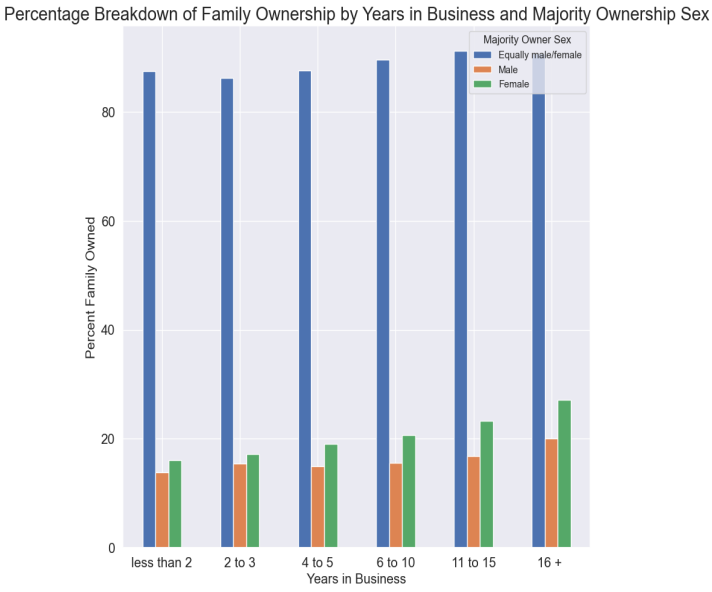


The figure above gives us insight into how family ownership varies by industry. The plot shows that most industries have somewhere between 30% and 40% of businesses that are family owned. The industry with the lowest percentage of family owned businesses is health care and social assistance. The industry with the highest percentage (disregarding unclassified industries) is company and enterprise management.

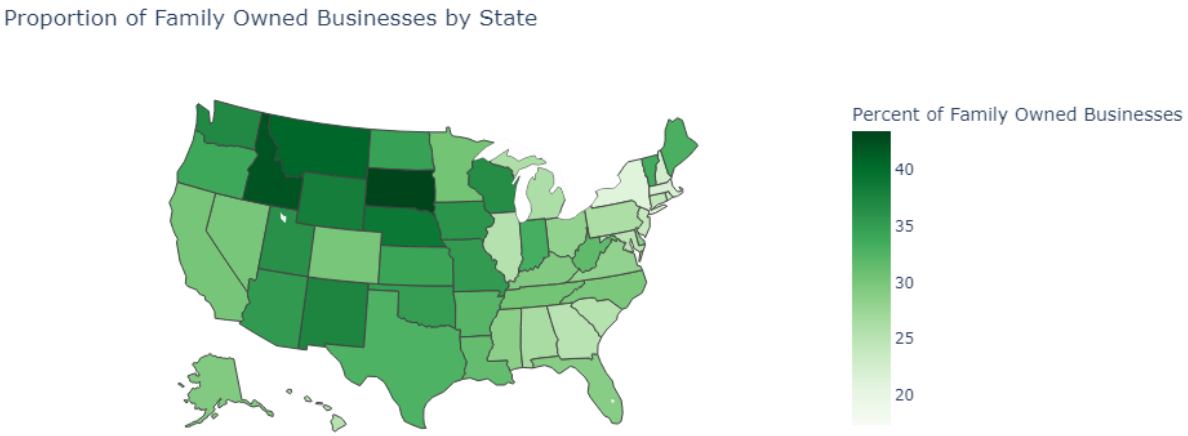
Next, we examine how family ownership prevalence changes with how many years a firm has been in business. The plot on the right shows us a few things: the United States has a higher count of older businesses compared to newer businesses, regardless of family ownership. For every 'years in business' breakout, there is a higher percentage of businesses that are not family owned. Lastly, the prevalence of family owned firms seems to increase with years in business. The change is subtle but can be seen more clearly in the 100% stacked bar chart located on the right. We also look at how this distribution changes when sex of the business's majority owner is considered. The most distinct variation in the plot below is that equally male/female owned businesses have a significantly higher proportion of family owned businesses compared to female and male-owned. This makes sense as one can imagine many family businesses being jointly owned by spouses. It is also interesting to note that for all 'years in business' breakouts, male owned businesses have the lowest proportion of family owned businesses.





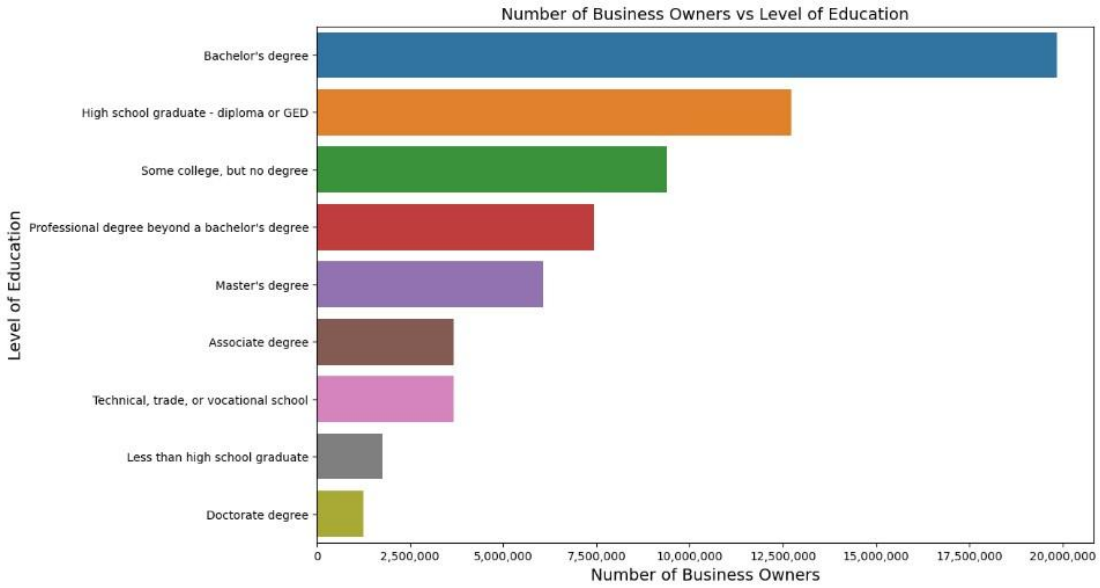


Lastly, we examine the geographic patterns of family ownership. In the map below, states with a darker shade have a higher proportion of family owned businesses, with a maximum percentage of about 43% in South Dakota. The state with the lowest proportion is New York at around 21%. There also seems to be a cluster of high family ownership states in the Midwest whereas Northeastern and Southeastern states tend to have lower family ownership.

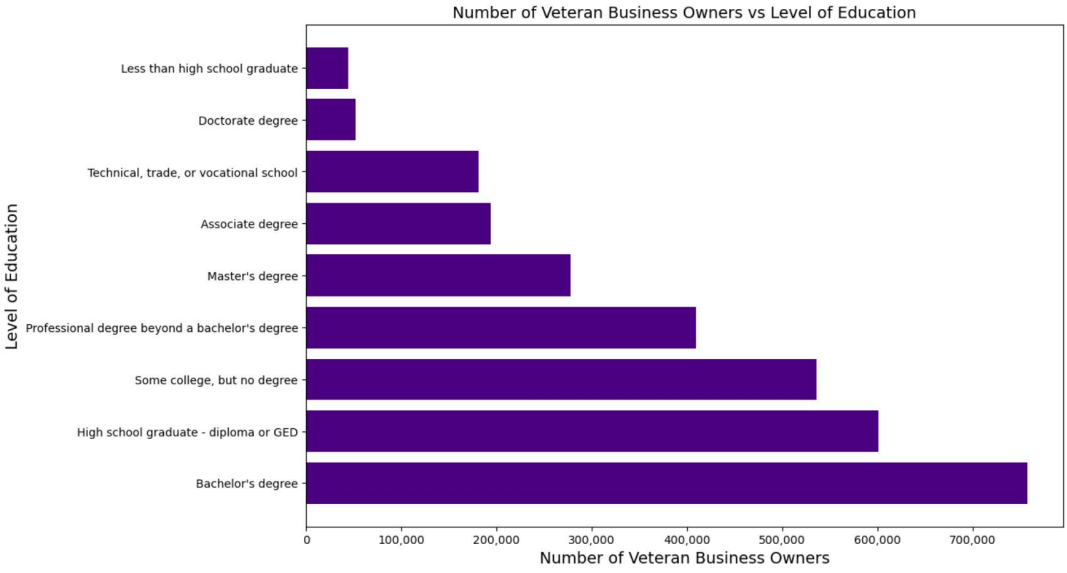


Education

Our group sought to investigate the level of education and educational majors of both overall and veteran business owners. To do this, we utilized the "Business Owners Characteristics" API and the ETL steps are shown in the ETL report. In order to provide a more thorough analysis, we chose to exclude "Total Reporting" and "Item Not Reported" from the filtering process. Based on our hypothesis, we anticipate that the most common degrees held by business owners will be a bachelor's degree and a master's degree.

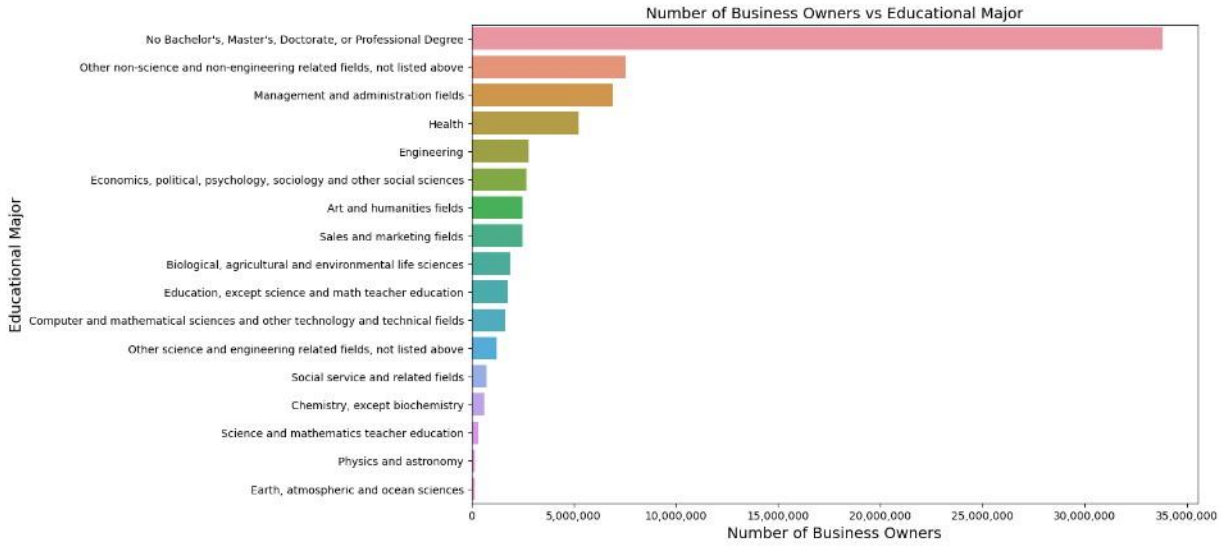


Based on the data presented in the above visualization, it is clear that the majority of business owners possess a bachelor's degree, while fewer have attained higher levels of education. Additionally, it is noteworthy that there are approximately 13 million business owners who have only completed high school or obtained a GED, and that there are more business owners without any diploma compared to those with a doctorate degree. This shows that our hypothesis is partially correct in stating that a Bachelor's degree does have the most business owners. However, there are more business owners who have obtained a high school diploma or a GED than business owners who have a master's degree.

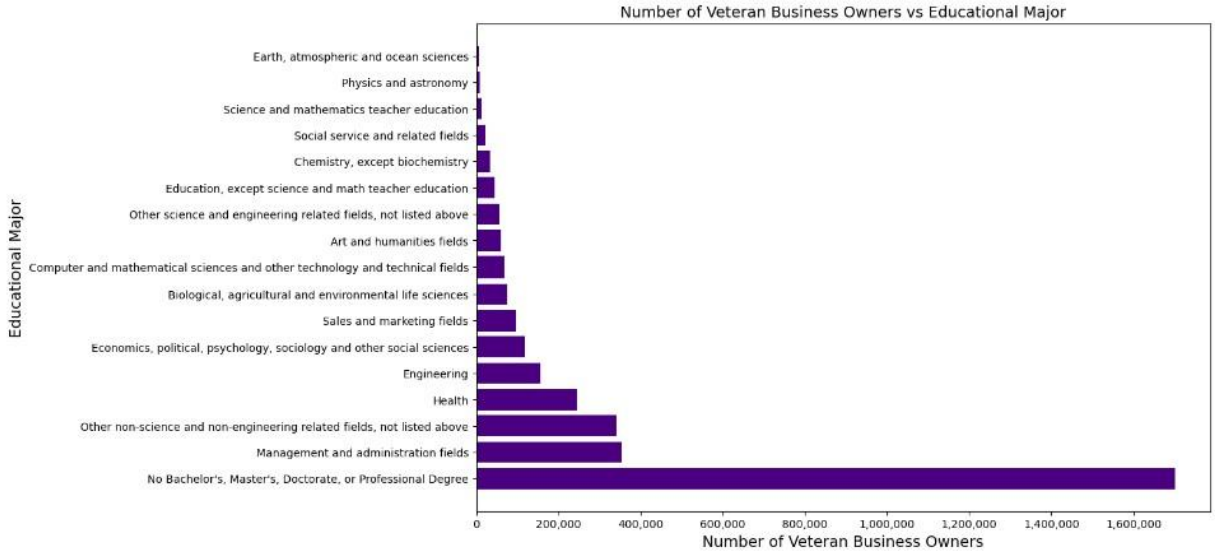


.

This visualization examines the level of education among veteran business owners and illustrates that approximately 750,000 of them possess a bachelor's degree. The number of individuals with a doctorate degree is higher than the number of those who have not graduated from high school. The distribution of educational achievement among veteran business owners appears to be similar to that of all business owners. It is worth noting that the number of veterans with a doctorate degree is higher than the number who have not completed high school. This could be because military service often requires at least a bachelor's degree for certain positions, such as commissioned officer.



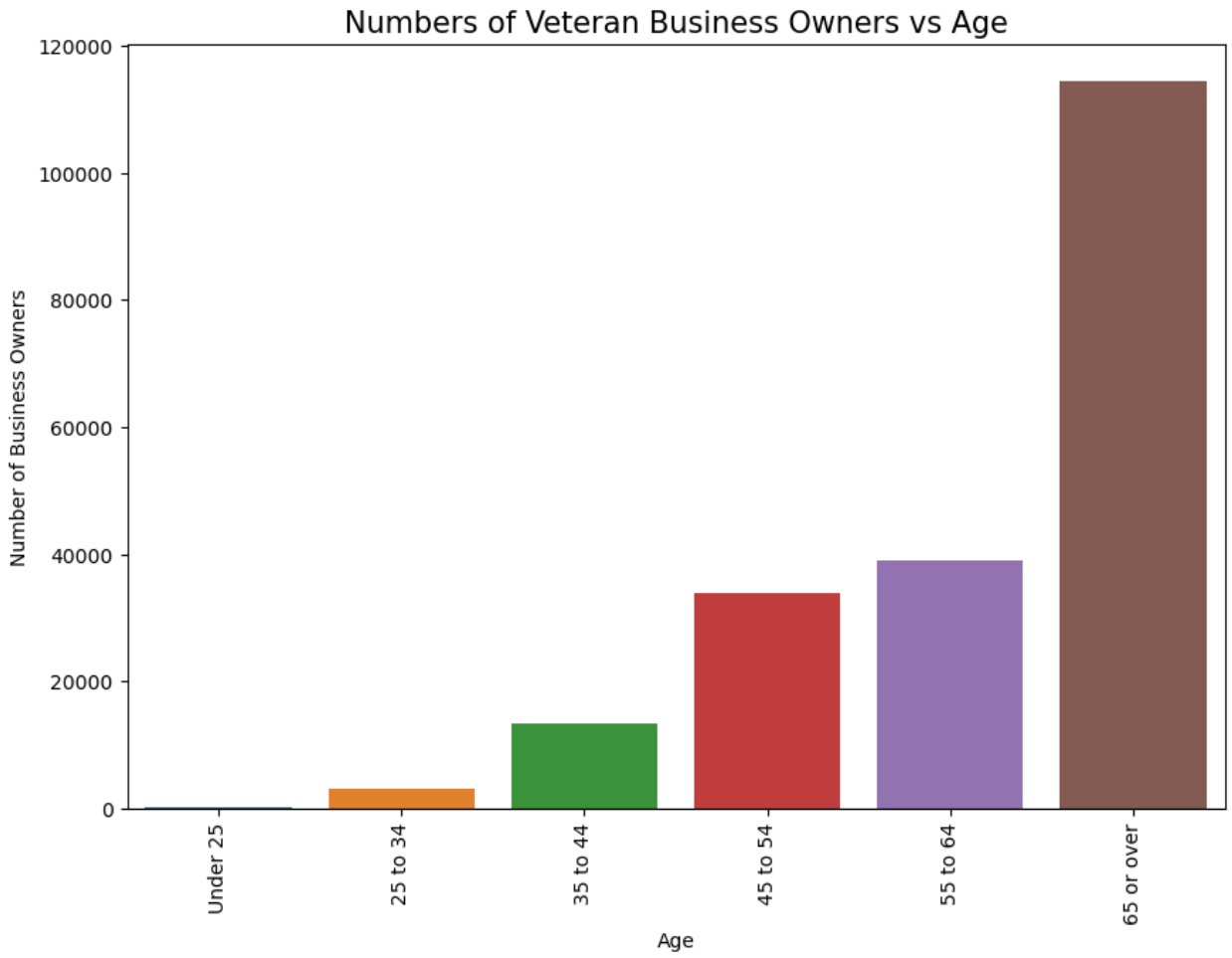
The following visualization presents data on the field of study among business owners. Excluding “No Bachelor's, Master's, Doctorate, or Professional Degree", the largest group, comprising approximately 9 million business owners, studied fields other than science or engineering. The second largest group, consisting of approximately 8 million business owners, studied management and administration. The smallest group studied Earth, atmospheric, and ocean sciences. It would be more useful to separate the category of "No Bachelor's, Master's, Doctorate, or Professional Degree" individually in order to better understand the data. This bar row includes approximately 34 million business owners.



The following visualization displays data on the field of study among veteran business owners. The two largest categories are management and administration and fields other than science or engineering. There appear to be more veteran business owners who studied management and administration than those who studied non-science and non-engineering

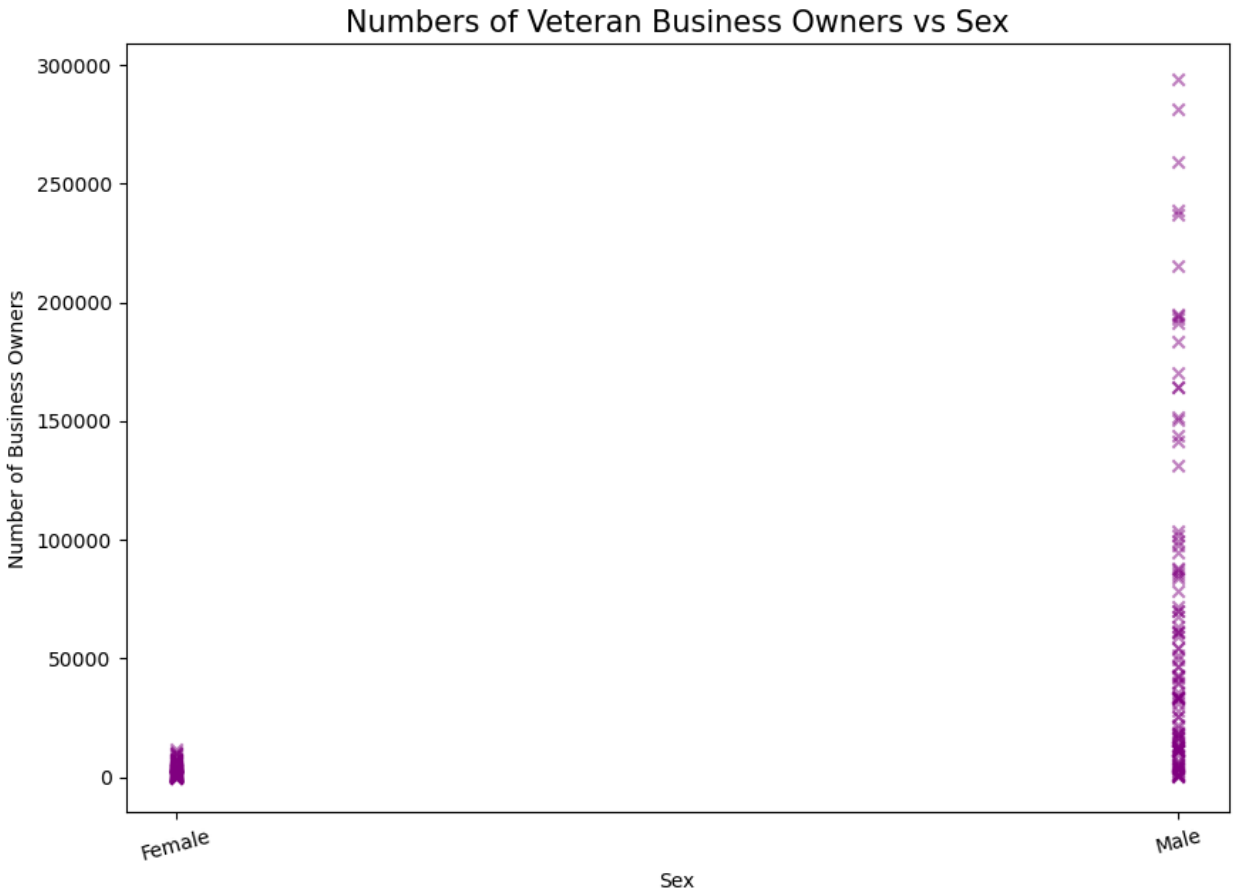
fields. The smallest group studied Earth, atmospheric, and ocean sciences, which is consistent with the previous visualization. As previously mentioned, it would be more insightful to further break down the category of "No Bachelor's, Master's, Doctorate, or Professional Degree" into individual categories such as bachelor's, master's, doctorate, and professional degree to facilitate a more thorough analysis of the data.

Deeper Dive Into Veteran Data

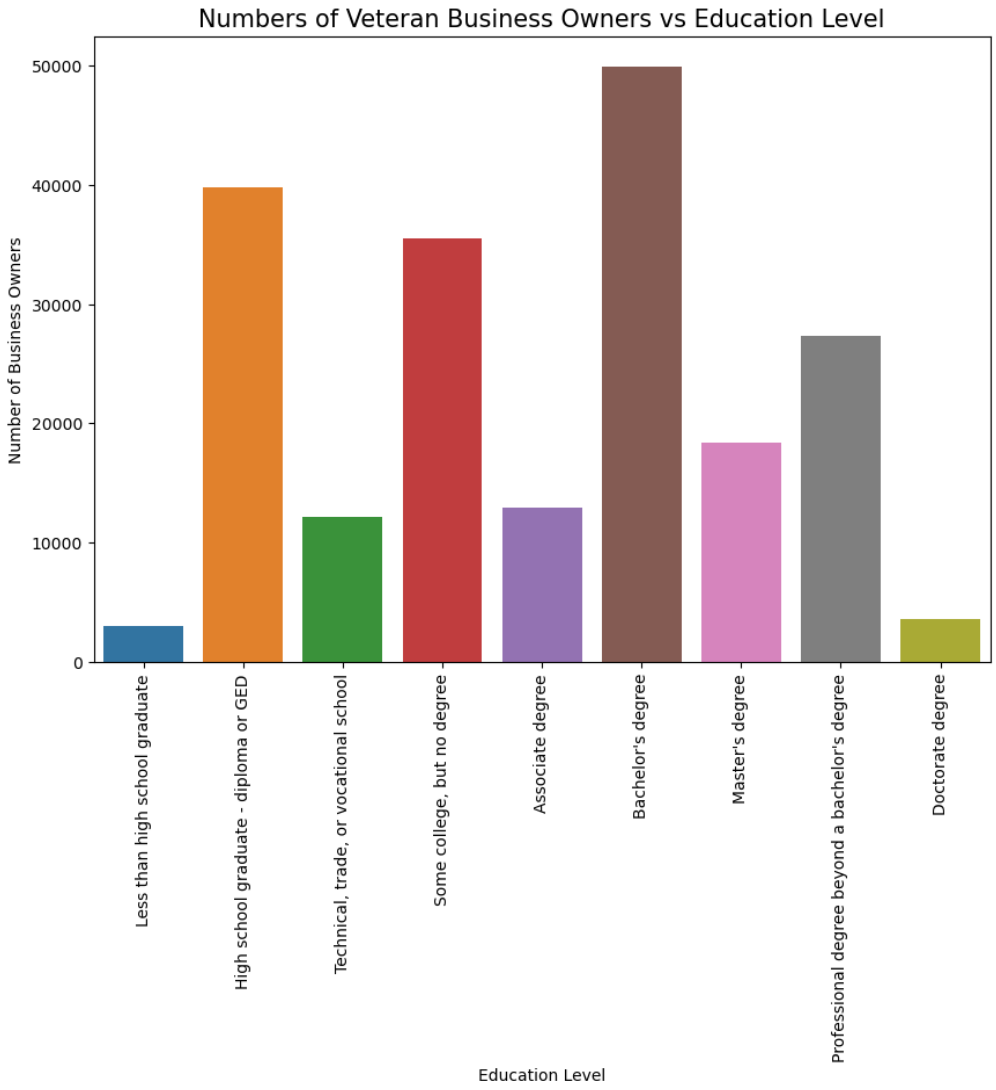


There seems to be a strong correlation between the number of veteran business owners (VBO) and their respective age with an upward trend. The final bucket could potentially be misleading as it includes a potential range of the previous four buckets combined, but the preceding buckets (25 to 34, 35 to 44, 45 to 54, 55 to 64) still maintain the aforementioned strong correlation.

The age restriction and initial contract length corroborates the lack of VBO under 25, and the slow increase in the amount of VBO can easily be explained by service members leaving varying contract lengths with leadership experience–a skill set that lends itself well to business ownership.



While it was no surprise that the majority of VBO responded with a male identification, the sheer overshadowing amount was almost baffling. This could be a result of poorly framed survey questions that didn't allow the option outside of the predated binary, or these data could strongly convey the lack of female veterans who choose to become business owners.



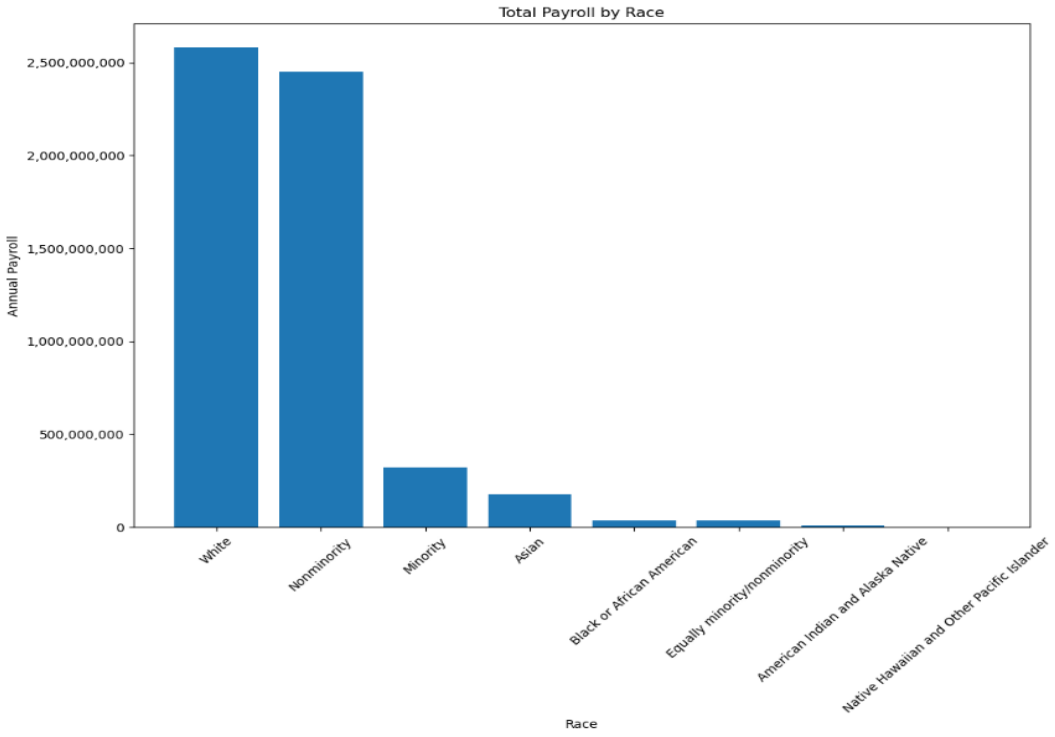
A secondary look was taken into the relationship between education level and the number of related VBO using a modified pull request from the characteristics of business owners repository, and the resulting analysis supports the previous analysis.

Veterans who discharge with a status of anything other than less-than-honorable receive financial help in the form of the Post 9/11 GI Bill (and similar programs) that can help offset the rising costs of tuition and other education related expenses. Additionally, the majority of veterans leave their service with at least some level of college in the form of transfer credits from technical training received in the military required to perform the roles and responsibilities of their respective MOS.

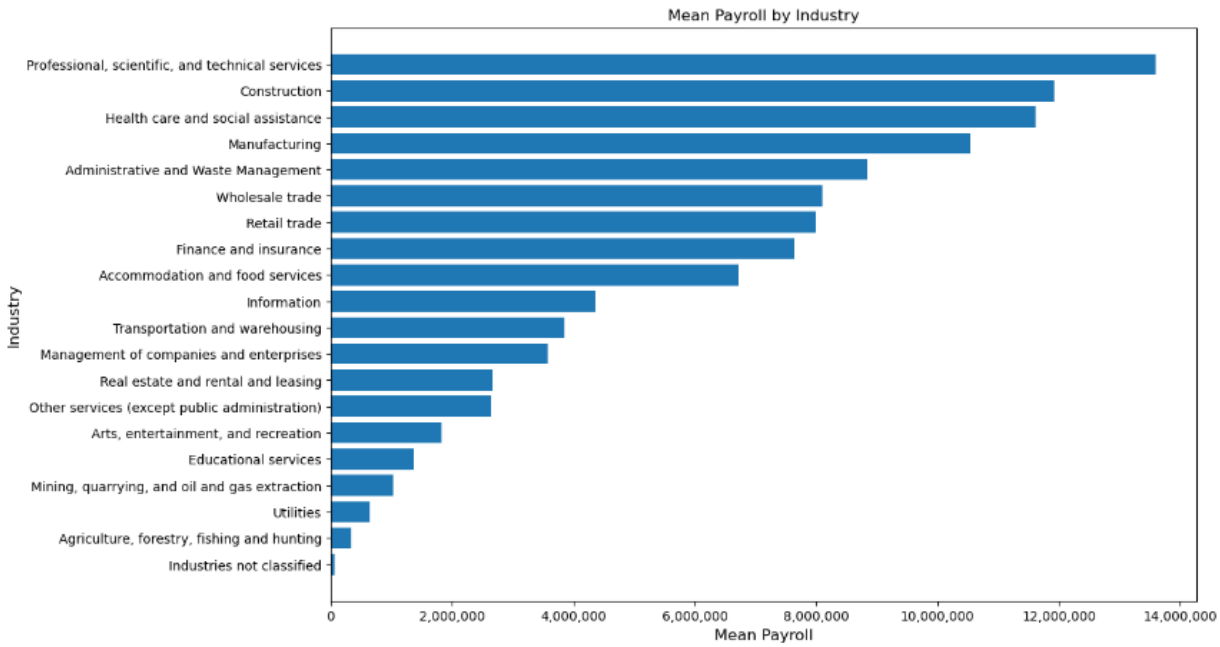
If a veteran joined the military with a year or two of college courses and tacked their technical training onto their transcript, they could potentially only need another year or two post-separation to acquire a four year degree. There are also opportunities to pursue degrees prior to separation, leaving veterans' benefits post-separation available to help the veteran chase higher than bachelor level degrees.

Annual Payroll Overview

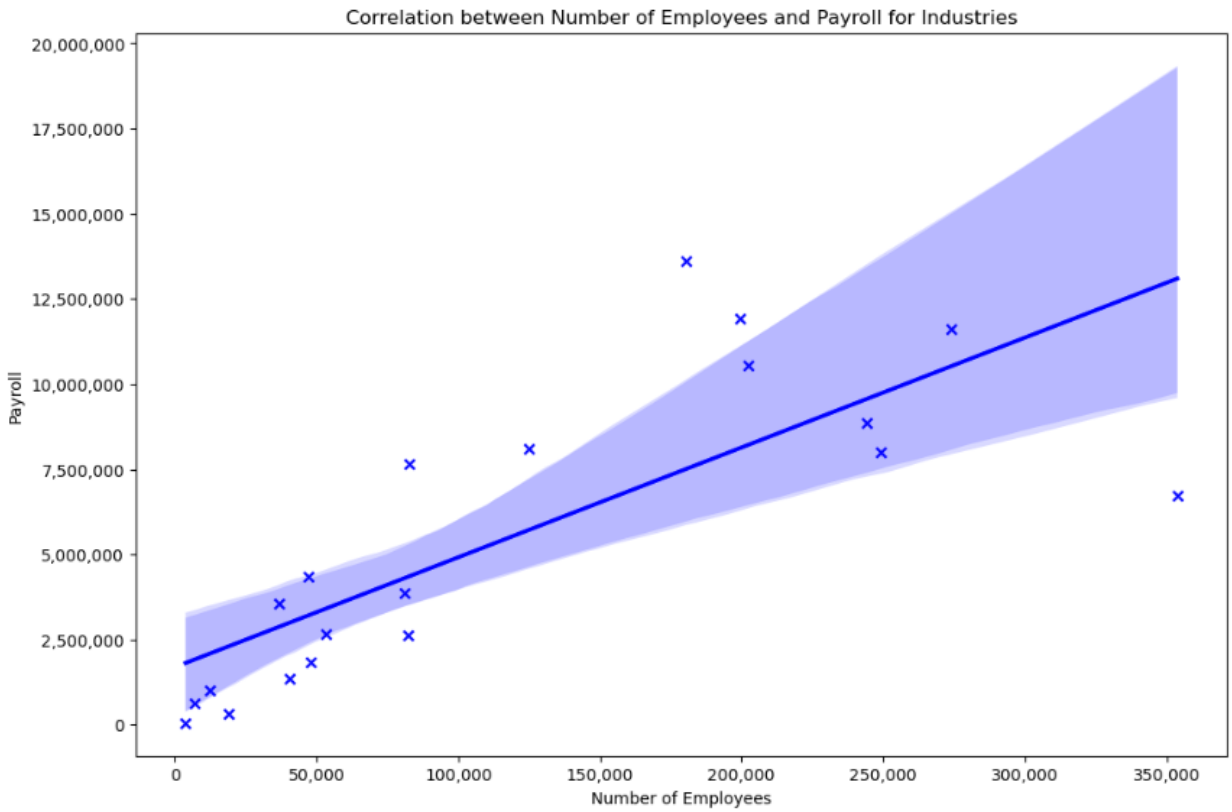
A broad question worth asking is how can the annual payroll of companies be summarized by different factors, such as industry, owner race and state. We hypothesize that the highest grossing industries, non-minorities and highest grossing states will have the highest payrolls on average. Some data was omitted from these visualizations to give a more meaningful representation of the sample.



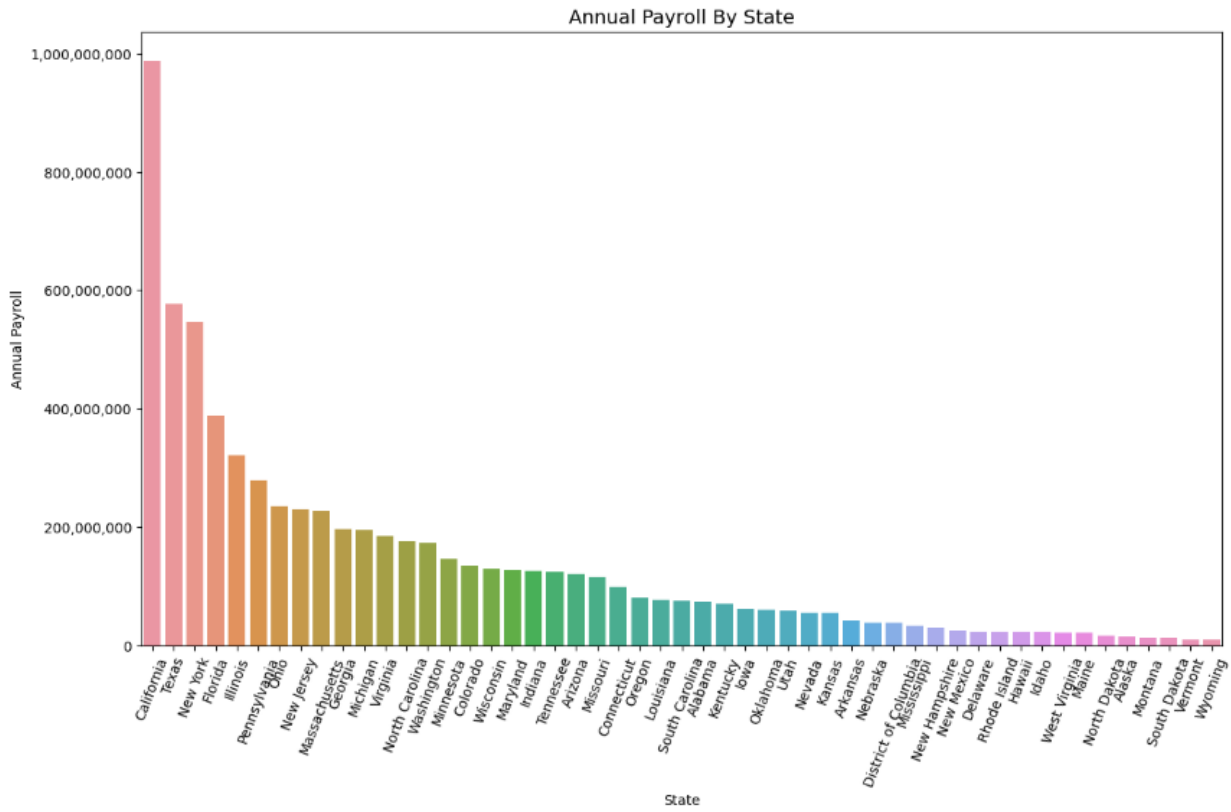
Upon viewing the total payroll by race, you will notice a drastically smaller national payroll than expected. One main drawback of the US census data is that much of the payroll data for race is labeled ‘Classified’ and ‘Unclassified’. Those values were omitted, lowering the total payroll numbers. Despite this manipulation, the general breakdown confirms the hypothesis that a majority of payroll would be controlled by white, non-minority business owners.



Another interesting topic to break down is the mean payroll by industry. There are a number of ways to interpret this graphic, and lead to further questions. One question would be to break down each industry by average firm size. Naturally, larger firms will have larger payrolls. Another explanation could simply be that certain industries do have higher payrolls on average. This would be expected for professional, scientific, and technical services, but not necessarily for construction. To further analyze this question, look at the scatter plot below:



This plot was created using the regplot function. It shows the expected payroll for companies of certain sizes. Looking at this visualization, it is clear that company size and payroll are not following the explicit trend of company size leading to increased payroll. About half of the data lies outside of a 95% confidence interval, which indicates that the number of employees alone is not a great predictor of payroll. This lends itself to the idea that most industries have other factors, on average, affecting payroll.



Annual payroll by state certainly does follow the trend that more employees leads to higher annual payroll. Unlike the previous datasets that project the averages across race and industry, this graphic projects totals. It is unsurprising that states with higher populations have higher payrolls. However it is not merely a population chart, as companies in some states in the northeast and on the west coast have higher annual payrolls relative to population than midwestern states. This graphic illustrates that location does have an effect on annual payroll.

Data Sources

U.S Census Bureau. (2019). Annual Business Survey (ABS) API. Retrieved Dec 22, 2022 from https://www.census.gov/data/developers/data-sets/abs.2019.html