ABS 2019 Project Report

Kai, Keith-Jordan, Addison, and Caitlin

10/11/21

Table of Contents

[Introduction](#)
- [Questions](#)
- [Feedback](#)

[Where are High Revenue Firms Located and Do They Affect the Number of Firms per Owner?](#)

[How do female-owned businesses compare to male-owned businesses in terms of employee pay and number of employees within a given state?](#)

[Which States are Utilizing Advanced Technologies?](#)

[How does the number of business owners vary by race and sex in the United States?](#)

[Conclusion](#)
- [Where are High Revenue Firms Located and Do They Affect the Number of Firms per Owner?](#)
- [How do female-owned businesses compare to male-owned businesses in terms of employee pay and number of employees?](#)
- [Which States are Utilizing Advanced Technologies?](#)
- [How does the number of business owners vary by race and sex in the United States?](#)

[References](#)

Introduction

The US Census Bureau conducts the Annual Business Survey (ABS) to obtain demographic information on firms and business owners, as well other business characteristics such as innovation and technology (US Census Bureau, 2021). To better understand businesses at the state and national levels, we use Pandas to explore datasets including “Characteristics of Businesses”, “Characteristics of Business Owners”, and “Technology Characteristics of Businesses” (US Census Bureau, 2019).

Questions

- Where are High Revenue Firms Located and Do They Affect the Number of Firms per Owner?
- How do female-owned businesses compare to male-owned businesses in terms of employee pay and number of employees within a given state?
- Which States are Utilizing Advanced Technologies?
- How does the number of business owners vary by race and sex in the United States?

Feedback

- slide 3: fix the bin size in the histogram (it doesn't appear to be exactly 1%; if that was intentional, please indicate the bin size explicitly)

- slide 6: since there isn't a big difference between these two states, it might be more interesting to pick two additional states that exhibit the biggest contrast

- slide 7: include the slope and r-squared of the regression line shown

- slide 15: replace the code numbers on the x-axis with the meaning

To address this feedback line by line, the histogram bin size was updated in the Jupyter Notebook (Technology_ETL.ipynb). The histogram was not included in this report due to space limitations and because similar information is conveyed in the included bar chart and maps.

The feedback regarding the two pie charts for the slideshow is addressed here in this report and reflected in the Jupyter Notebook ('Business Owner Sex ETL.ipynb'). Instead of using Wisconsin and Minnesota data as used in the slideshow, Delaware, Idaho, and Nevada were chosen as the three states had similar ratios in terms of the breakdown of the sex of the owner and the average payroll per employee.

The scatterplot from slide 7 is not included in the report. The plot was used during the exploration of data phase of the project, the use of average payroll per employee (pay/emp) was chosen as a way to help normalize the data so states with large populations and the resulting high payrolls did not affect the visuals and analysis.

The feedback regarding replacing the codes has been addressed and the changes have been reflected both in the report and in the Jupyter Notebook ('Business Owners US- ETL.ipynb').

Where are High Revenue Firms Located and Do They Affect the Number of Firms per Owner?

The data in the Company Summary table provides information on the number of firms in each state grouped by years in business and revenue size for the year 2018. The data in the Characteristics of Business Owners was used to find the number of business owners by state. By filtering the data from the Company Summary to businesses with revenue of $1,000,000 or more, it can be plotted on a map where high revenue businesses are located (**Fig. 1a**).

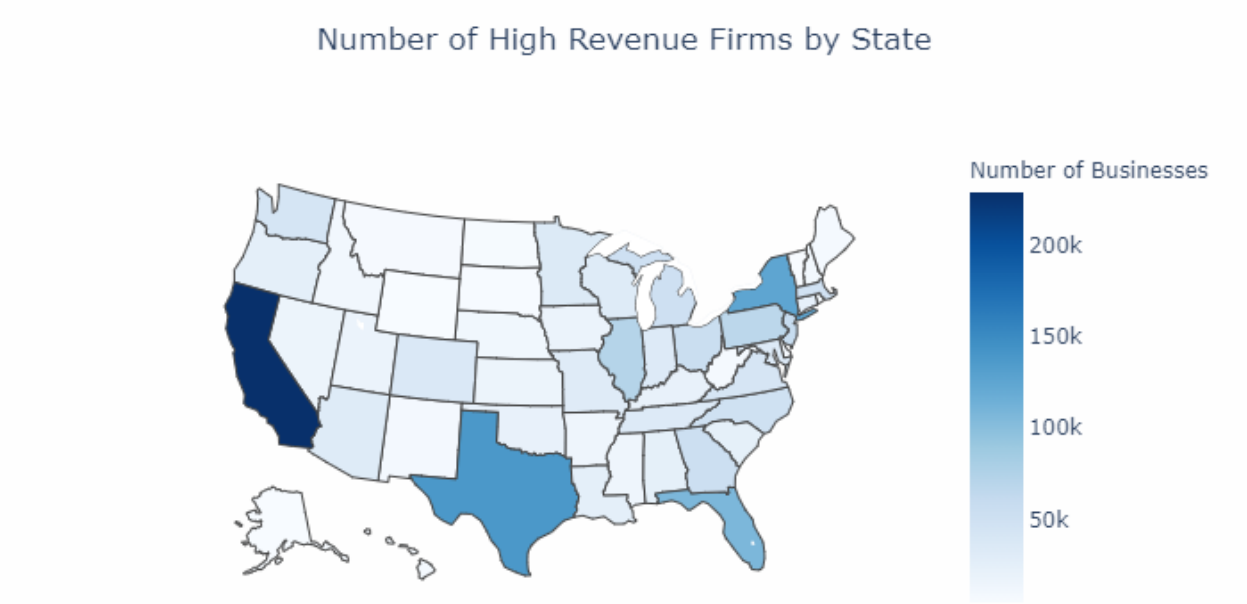

Figure 1a: *Number of High Revenue Firms per State*

As to be expected, the states with the highest populations have the highest number of high revenue businesses due to them having more businesses in general. In order to get a better sense of the concentration of high revenue businesses, they can be normalized as a percentage of total respondent businesses per state (**Fig. 1b**).

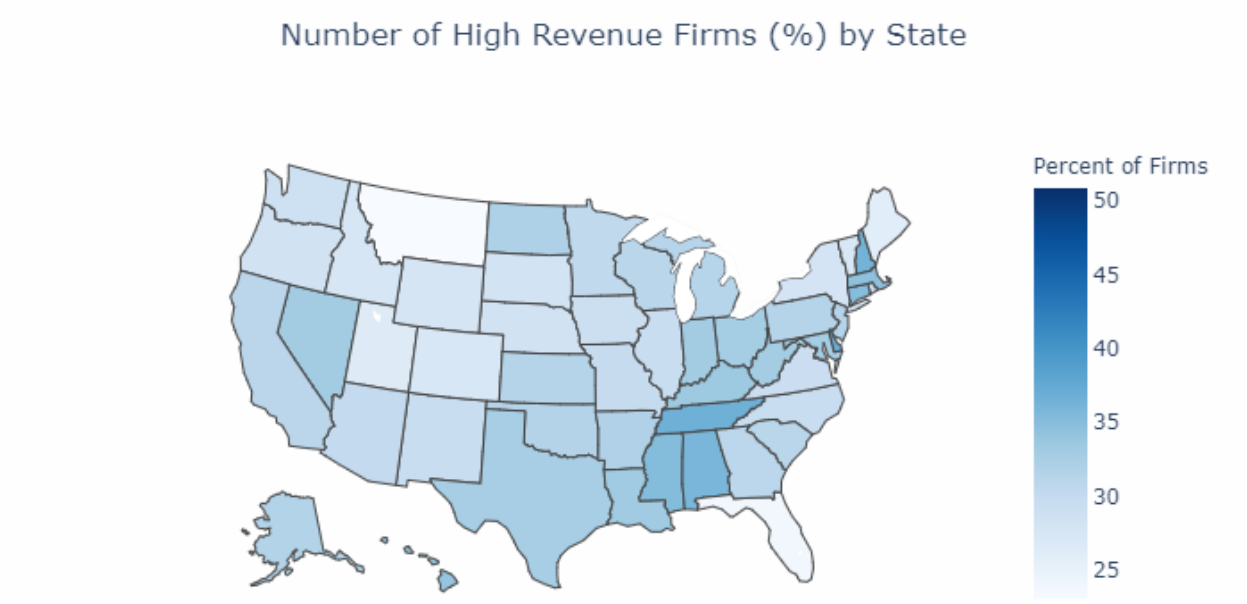

Figure 1b: *High Revenue Firms per State as Percentage of Total Firms*

This shows a higher percentage of firms with high revenue being in the eastern US, with the highest being in DC at just over 50% of businesses having $1,000,000 in revenue in 2018. From here the number of businesses of owners can be found by merging the company summary table with the characteristics of business owners table and comparing their columns (**Fig. 1c**).

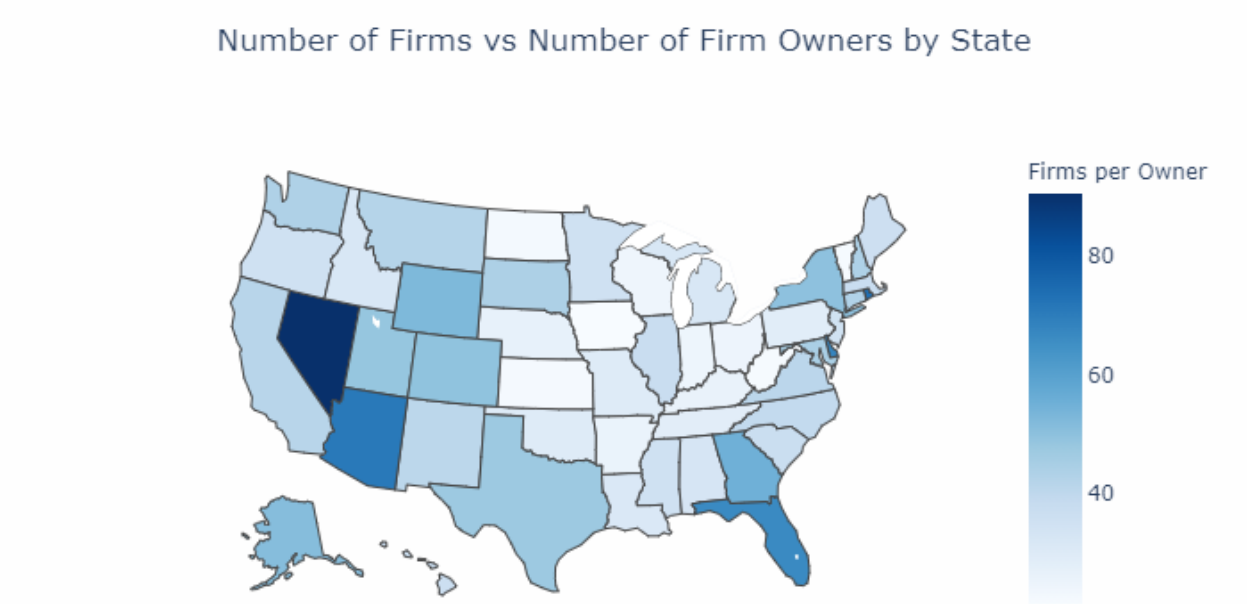

Figure 1c: *Number of Firms per Owner per State*

This shows that many of the states with high numbers of firms per owner don't rank highly on either the quantity of high revenue businesses or the concentration of high revenue businesses maps. For example, Nevada has a much higher rate of firms to owners than other states, but it doesn't rank highly on either high revenue business map. To look further at how the location of high revenue businesses affects or doesn't affect the number of firms per owner, the number of firms per owner can be plotted against the quantity and percentage of high revenue firms for each state (**Fig. 1d**).

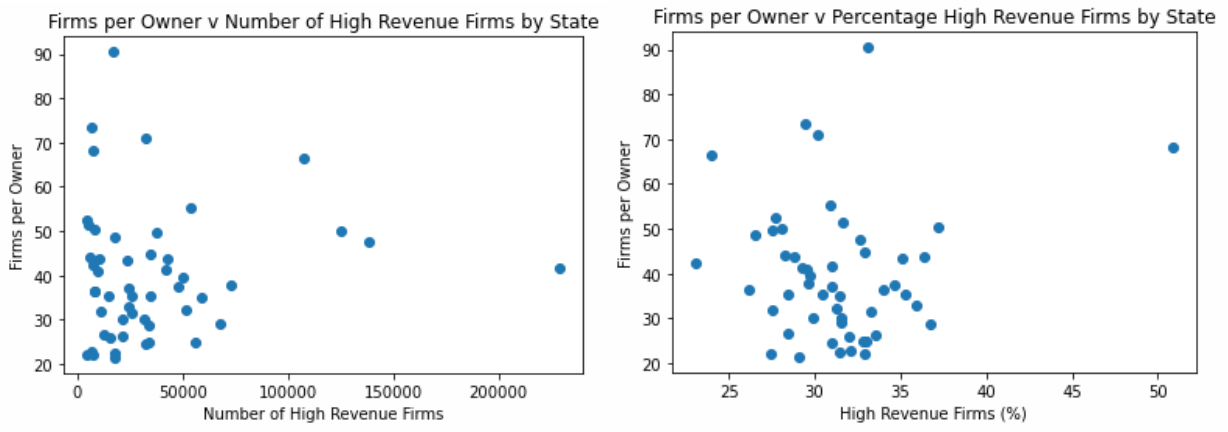

Figure 1d: Left: *Firms per Owner vs Quantity of High Revenue Businesses per State*

Right: *Firms per Owner vs Percentage of High Revenue Businesses per State*

These confirm the previous conclusion from the maps that neither high quantity nor high concentration of high revenue businesses increase the number of businesses per owner at the state level. This means that high revenue businesses either don't cause their owners to open more businesses than smaller business owners, or that the existence of larger businesses allows for more owners with few businesses to offset the high revenue owners' expansion into other businesses.

How do female-owned businesses compare to male-owned businesses in terms of employee pay and number of employees within a given state?

The data pulled from the Company Summary table gave the sex code of the business owner, the sex label of the business owner, the number of employees in the state who work for a business with that particular type of business owner, and the total annual payroll of a business with that particular type of business owner. The data here already grouped all industries together so there is no breakdown by industry and as the question is asking about male versus female business owners, those two categories have been retained. Additionally, two new calculations were done to the dataframe in order to help normalize the data to help with visuals. An average annual pay per employee column was added which took the annual payroll column and divided it by the number of employees column. And a ratio of the female average annual pay per employee to male average annual pay per employee, the ratio is explained further in the section.

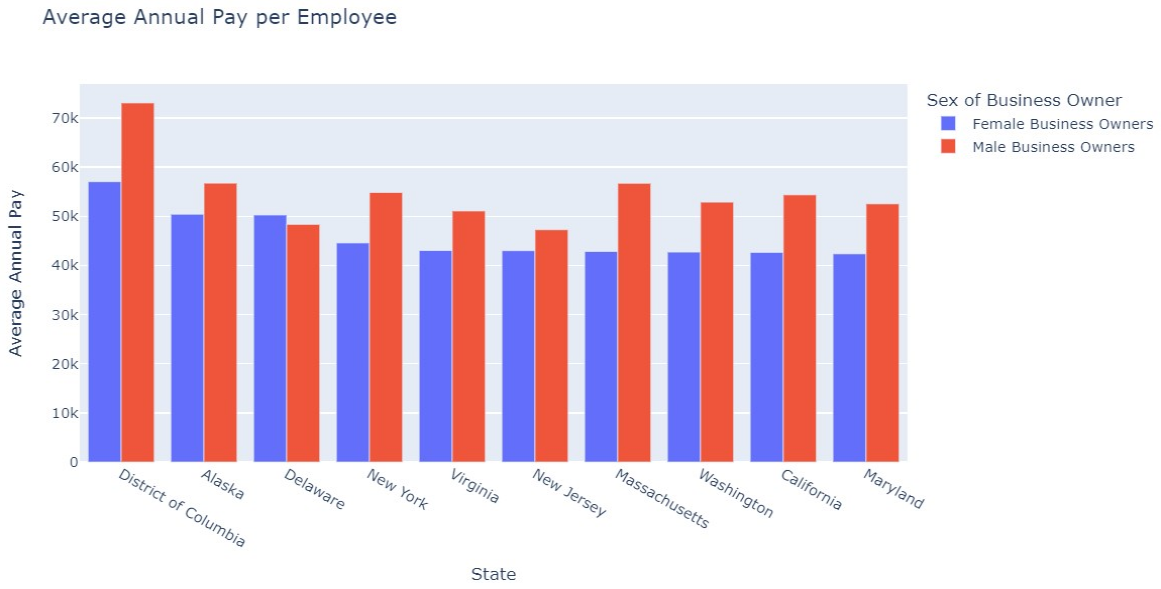

Figure 2a: *Average pay sorted by female-owned*

To answer the first part of the question, regarding employee pay. When sorting by average annual pay per employee for a female-owned business, the District of Columbia pays employees the most, with Alaska and Delaware among the top 3 (**Fig. 2a**). But what is evident in this graph is that the pay for

employees of male-owned businesses is higher in almost every state except Delaware which brings up a new question. Are there any other states where female business owners pay their employees more than male business owners?

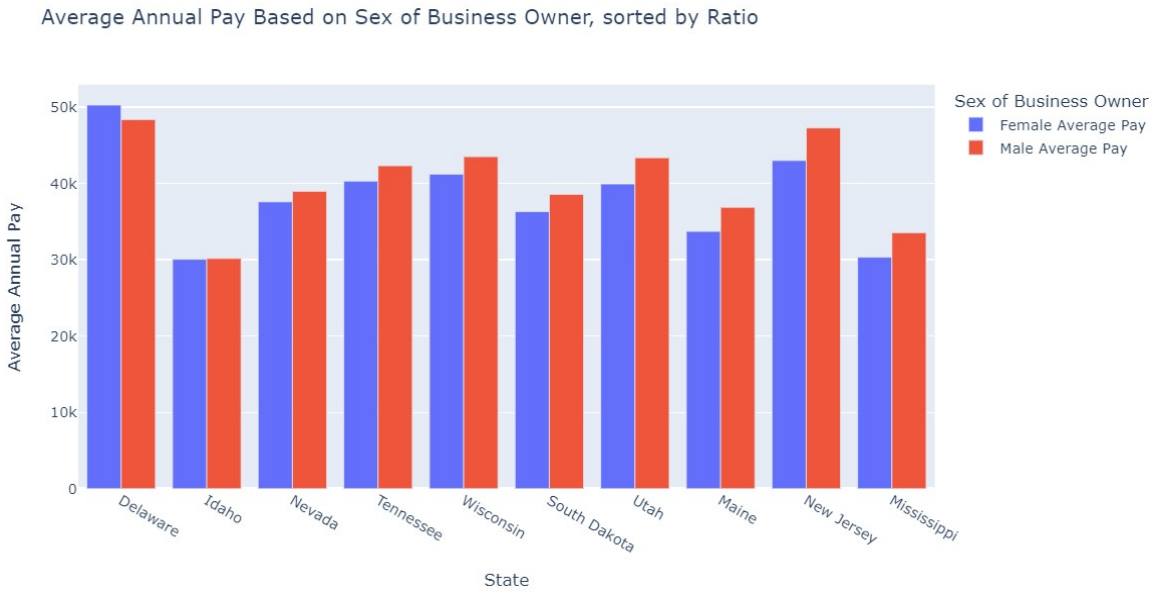

Figure 2b: *Average pay sorted by ratio of average pay*

Seen in the plot to the right (**Fig. 2b**) the differences in average annual employee pay between male and female business owners is much smaller than shown in Figure 2a. To address the question raised above, if any other states have female business owners paying their employees more than male business owners, a new approach was taken. The new approach calculated the ratio between the average pay per employee of female-owned businesses and the average pay per employee of male-owned businesses. If the ratio is greater than 1, a given state's female business owner's average annual pay per employee is larger than that of a male business owner. Seen in Figure 2b, Delaware is the only state in which the ratio is greater than 1, but the top 10 states also have differing levels of pay per employee. So a higher average annual pay per employee does not mean the ratio is close to 1 as states with lower annual pay per employee also had a high ratio between the two types of business owners.

After looking at the average pay per employee based on the sex of the business owner, next look at how many employees work for a business owner who is male or female.

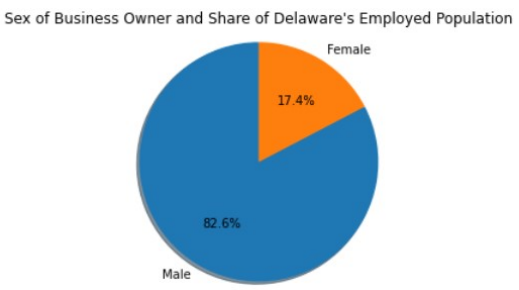

Figure 3c: *Delaware*

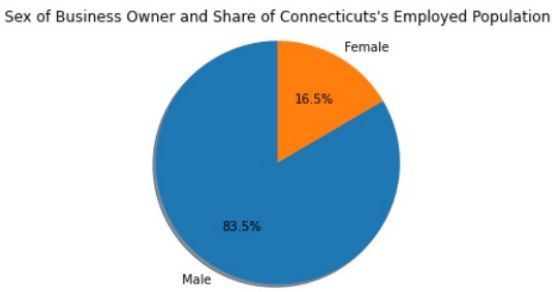

Figure 3d: *Connecticut*

These two pie charts are almost identical. But, as mentioned above, Delaware was the state with the highest ratio and in fact, was the only state where female-owned businesses had a higher average annual pay per employee than male-owned businesses. In Connecticut's case, it was last in the ratio ranking with a value of 0.6 (of states in which data was fully available to find the ratio). So the share of the employed population who work for female-owned businesses is not related to the average annual pay per employee ratio when compared to the pay of employees of male-owned businesses.

Which States are Utilizing Advanced Technologies?

The dataset "Technology Characteristics of Businesses" provides the number of firms that utilize different technologies such as artificial intelligence, cloud computing, specialized equipment, etc. The amount of firms per state that utilize AI or Robotics can be plotted in a bar graph to show which states utilize AI and robotics the most (**Fig. 3a**) .

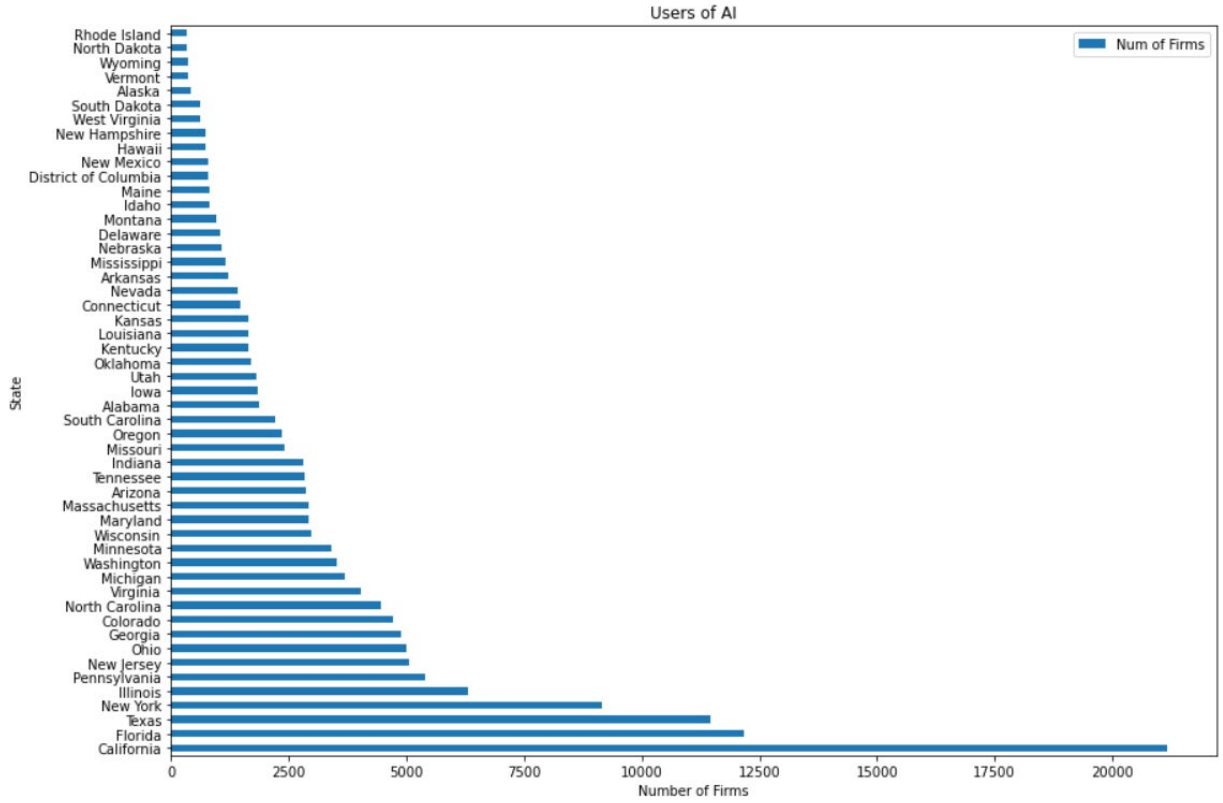

Figure 3a: *Number of Firms Using AI per State*

An issue with this representation, however, is that the number of firms within a state strongly correlates with the population. In both AI use and Robotics use, California, Florida, Texas, New York, and Illinois lead. With a higher population, a state should have a higher number of firms, and therefore have higher representation in plots of raw count per state.

In a sense, the question of which states utilize advanced technologies the most has been answered. However, additional analysis by normalizing values against the number of total respondent firms per state would provide additional insight towards how prevalent advanced technologies are in each

individual state. As part of the data transformation process, DataFrames were merged together in order to produce these insights.

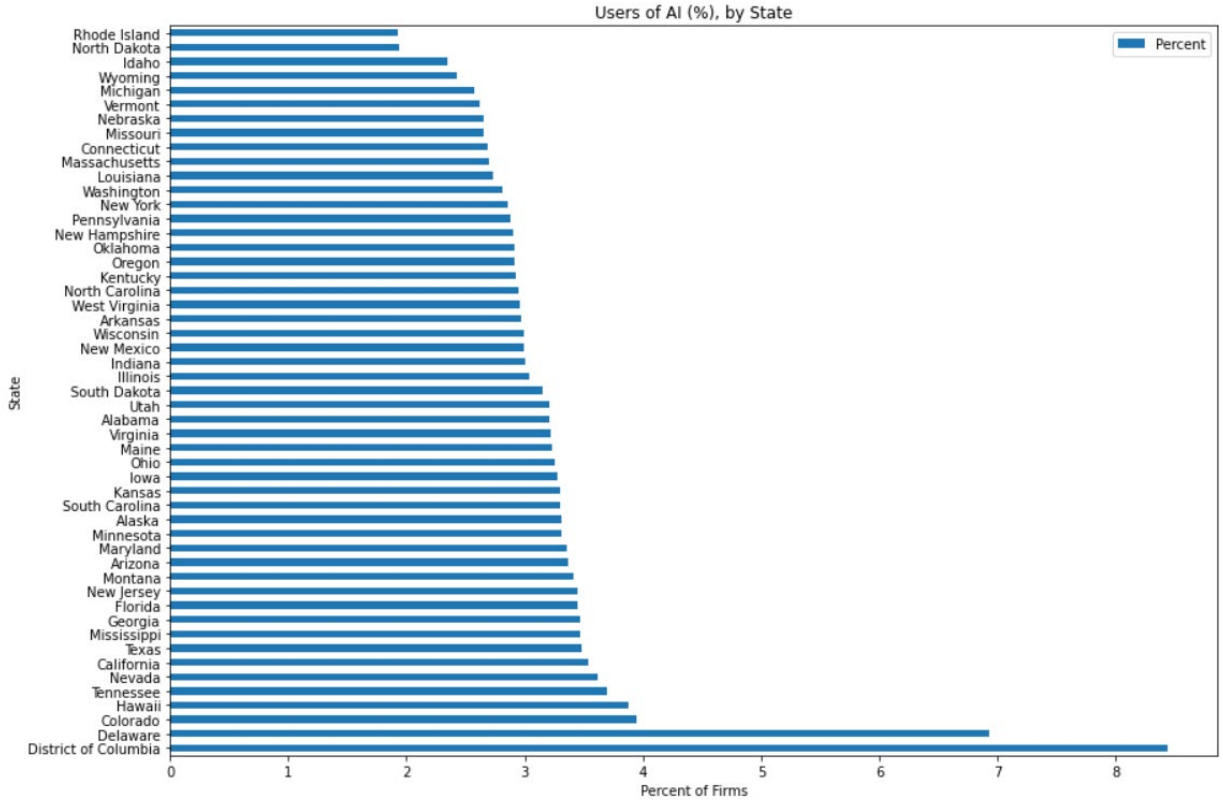

Figure 3b: *Percent of Firms Using AI per State*

Calculating the percentage of firms that utilize AI reveals that firms in Delaware and the District of Columbia utilize AI at a rate much higher than other states (**Fig. 3b and Fig. 3d**). For D.C. in particular, this may point to law firms, election campaign managers, or lobbyists making use of AI in the decision-making process. Additional research on the industries of firms operating in DC may reveal more information as to how the use of AI differs in DC from other states.

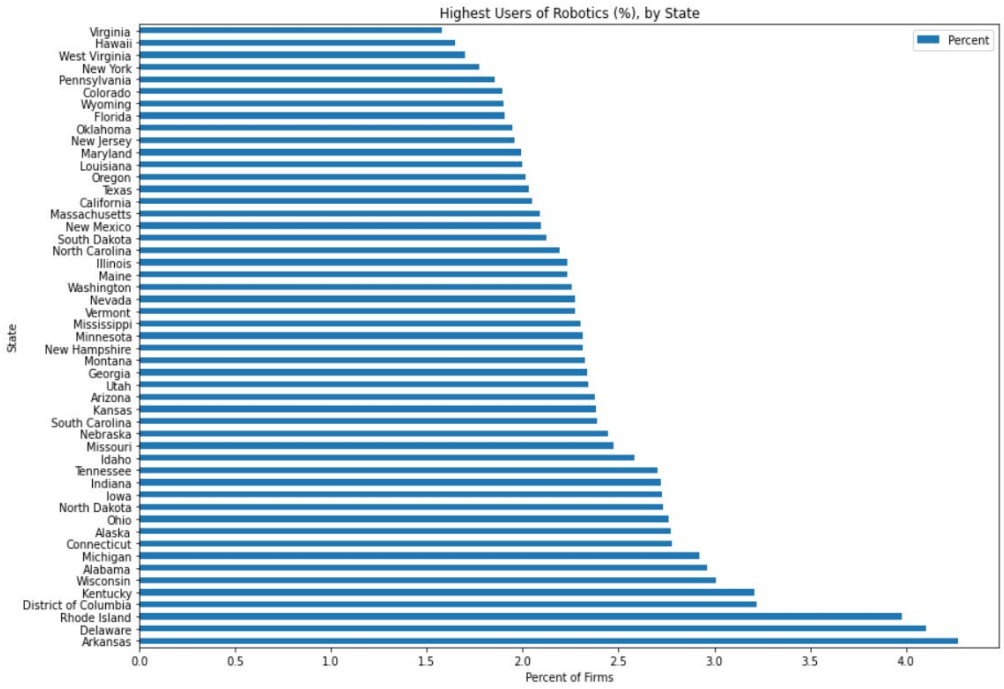

Figure 3c: *Percent of Firms Using Robotics per State*

After looking at state's use of AI, we were also interested in how states use complementary advanced technology such as robotics (**Fig. 3c and 3d).** In contrast to our initial hypothesis that AI use would correlate well with robotics use, top robotics users such as Arkansas, Rhode Island, and Kentucky lag behind in adopting AI use. This could be explained by these robotics being more heavily utilized in manufacturing firms, whereas AI may be utilized more by software or product design firms. Additional research towards these individual states' ecomines would be necessary to support these hypotheses.

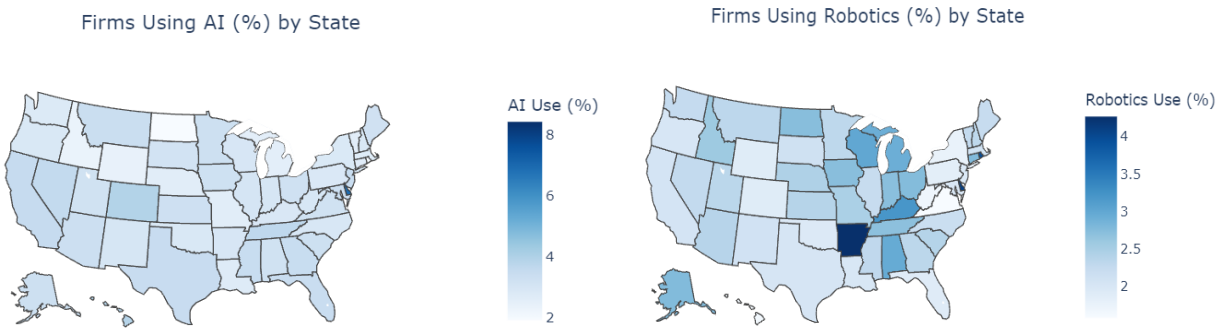

Figure 3d: Left*: Percent of Firms Using AI. R****ight****: Percent of Firms Using Robotics*

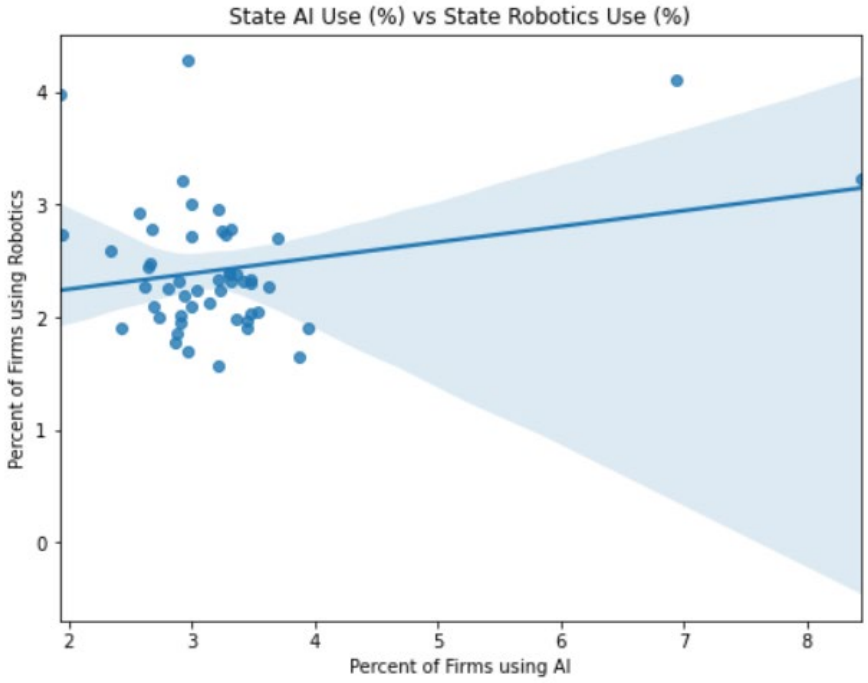

Figure 3e: *Regression Plot of Probiotics Use (%) vs AI Use (%) with regression line and confidence interval*

The lack of correlation between AI use and Robotics use in US states can be clearly shown in a regression plot prepared with the Seaborn package (**Fig 3e**). When plotting these values against each other, the regression line shows no confidence in any correlation. This finding is consistent with previous findings, and provides additional support that despite AI and Robotics both being advanced technologies, they're utilized in very different industries.

How does the number of business owners vary by race and sex in the United States?

The dataset “Characteristics of Business Owners” provides the number of business owners that are present in the United States based on their race, gender and veteran status. In order to answer the question the data had to be filtered so that the veteran status would not be taken into account. After cleaning and filtering the data we were able to plot graphs to see the breakdown of the amount of business owners by their demographics.

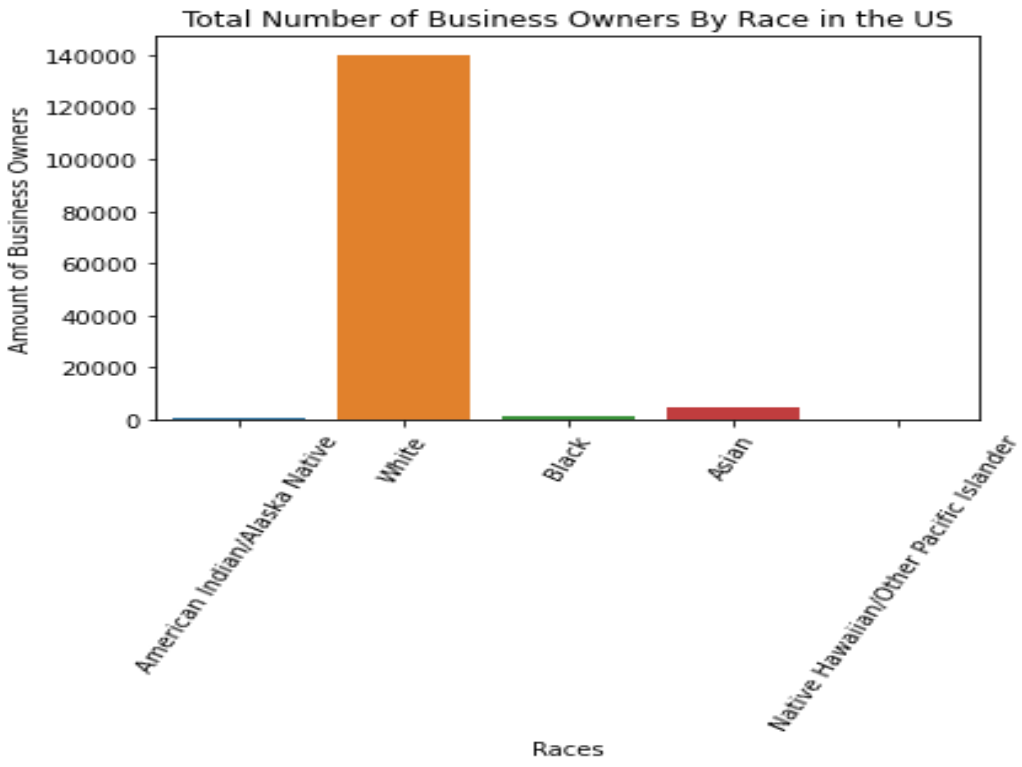

***Figure 4a**: Bar Chart of Total Number of Business Owners By Race in the US*

As seen in figure 4a, there is a huge gap between the amount of White business owners and the number of business owners belonging to other races in the United States. This is expected as the population of white people in the United States is much higher than that of other races. An unexpected

result however, was that Asain Americans owned more businesses than Black Americans when Black Americans had a larger population in the U.S than Asians. Also, it can be noted that the amount of Native Hawaiian business owners was negligible when compared to the number of owners recorded for other races.

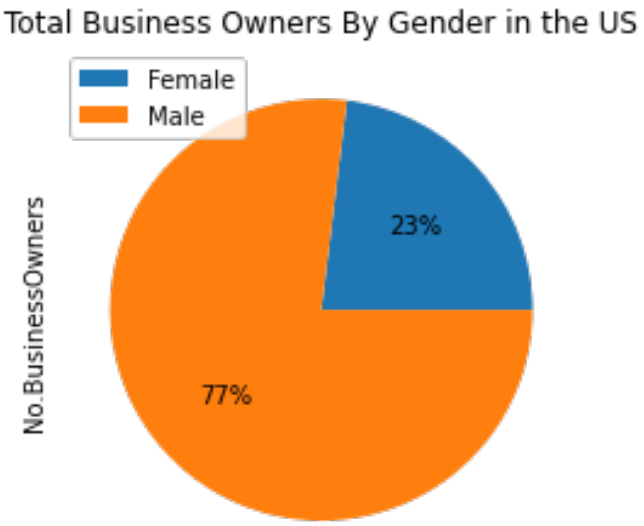

***Figure 4b**: Pie Chart Showing the Difference Between Male and Female Business Owners in the U.S.*

Based on the data, it is clear that in the U.S there are much more male business owners than female business owners. This is not a surprise as American culture is rooted in patriarchy. It was not until recently that a push for equity and equality for women has been seen in American society. Also, studies have shown that men are more likely to try to start businesses because they are overconfident. As a result, this result was somewhat expected.

Conclusion

Where are High Revenue Firms Located and Do They Affect the Number of Firms per Owner?

High revenue firms are located in the highest quantities in the states with the highest populations and in the highest proportion in the east coast. However these seem to not play a big role in the number of firms by the number of owners per state, with the highest consolidation of businesses being in Nevada, a place with neither a high population or a high percentage of high revenue businesses. Additional research would need to be done to see if high revenue businesses play any role at all in consolidation.

How do female-owned businesses compare to male-owned businesses in terms of employee pay and number of employees?

The District of Columbia's female-owned businesses pay their employees the most when compared to other states' female-owned businesses. But when compared within the state to the male-owned businesses, Delaware's female-owned businesses outperform the male-owned businesses in terms of average annual pay per employee. When it comes to the number of employees in a state who work for a female-owned business, there is little variation from state-to-state and even in extreme examples of comparing different ends of the average annual employee pay ratio spectrum, the share of employees who work for female-owned businesses is very similar. For further research, including the different industries would be important going forward. It can allow for understanding if female-owned businesses are more common in certain industries, possibly lower-paying industries which affects the comparison in pay with male-owned businesses.

Which States are Utilizing Advanced Technologies?

DC and Delaware lead the US as having the highest percent of firms that utilize AI technology. Arkansas, Delaware, and Rhode Island lead the US as having the highest percent of firms that utilize

robotics technology. AI use does not correlate well with robotics use, suggesting that the two technologies are used by separate industries. Additional research into individual state's firms may provide supporting evidence.

How does the number of business owners vary by race and sex in the United States?

White Americans constitute the majority of business owners in the United States. Also, despite having a lower population than African Americans, Asians have the second highest population of business owners. It can also be noted that despite there being slightly more females than males in the U.S, males have a much higher population of business owners than females.

References

U.S. Census Bureau. (2019*). Annual Business Survey, 2019.* [Dataset]. https://api.census.gov/data/2018/abscs.html

U.S. Census Bureau. (2021*). About the Annual Business Survey (ABS)* . https://www.census.gov/programs-surveys/abs/about.html#:~:text=The%20Annual%20Business%20Survey%20(ABS)%20provides%20information%20on%20selected%20economic,%2C%20race%2C%20and%20veteran%20status