**Census Annual Business Survey 2019**

**Project Report**

**Group 3: Luis, Will, Phil, Sharif**

**Introduction**

Businesses have several characteristics when it comes to identifying changes in demographic, ownership, and sales components. By identifying those component changes within businesses’ development process government entities, industry leaders, analysts and other interested parties may establish their working programs and agendas. Obtaining those type of data can be quite expensive and at times restricted to individual parties. Besides, those data need to be updated on frequent basis to depict current state of changes in industry, which makes it significantly more challenging for individual party. Therefore there’s a strong need in set of dataset that could satisfy all those requirements.

**Dataset Introduction**

The Annual Business Survey (ABS) provides information on selected economic and demographic characteristics for businesses and business owners by sex, ethnicity, race, and veteran status. Further, the survey measures research and development (for microbusinesses), new business topics such as innovation and technology, as well as other business characteristics. Statistics from the ABS may be used by government program officials, industry organization leaders, economic and social analysts, business entrepreneurs, and domestic and foreign researchers in academia, business, and government. ABS is updated annually to reflect ongoing changes in the industry. Below we listed ABS dataset sources that we used to conduct our work:

ANNUAL BUSINESS SURVEY (ABS) APIs. Accessed on 4/22/2022

- <https://www.census.gov/data/developers/data-sets/abs.2019.html>

COMPANY SUMMARY.

- Examples: [*api.census.gov/data/2018/abscs/examples.html*](https://api.census.gov/data/2018/abscs/examples.html)

- Variables: [*api.census.gov/data/2018/abscs/variables.html*](https://api.census.gov/data/2018/abscs/variables.html)

CHARACHTERISTICS OF BUSINESSES

- Examples: [*api.census.gov/data/2018/abscb/examples.html*](https://api.census.gov/data/2018/abscb/examples.html)

*-* Variables: [*api.census.gov/data/2018/abscb/variables.html*](https://api.census.gov/data/2018/abscb/variables.html)

CHARACHTERISTICS OF BUSINESS OWNERS

**-** Examples: [*api.census.gov/data/2018/abscbo/examples.html*](https://api.census.gov/data/2018/abscbo/examples.html)

*-* Variables: [*api.census.gov/data/2018/abscbo/variables.html*](https://api.census.gov/data/2018/abscbo/variables.html)

### TECHNOLOGY CHARACTERISTICS OF BUSINESSES

**-** Examples: [*api.census.gov/data/2018/abstcb/examples.html*](https://api.census.gov/data/2018/abstcb/examples.html)

*-* Variables: [*api.census.gov/data/2018/abstcb/variables.html*](https://api.census.gov/data/2018/abstcb/variables.html)

In this research, exploration, analysis, and findings our group concentrated on separate areas of the ABS dataset by identifying specific questions that would be answered by relevant work and findings of each member. Below is list of questions we each decided to investigate:

**Luis:**

* + Question 1: Comparing firms with least amount of years in business to firms with most amount of years in business, do the number of employees grow as years in business grows? (first 5 states).
  + Question 2: Is there a relationship between the number of owners for all firms and the number of employees?
  + Question 3: What is the total number of minority employees in all firms for the first 10 states?

**Will:**

What impact does having Female ownership have on payroll, sales, and employees?

**Phil:**

Please list all of your chosen questions in numeric order

**Sharif:**

With my exploration of the ABS datasets, I intend to answer the following questions:

1. Does higher number of employees always lead to higher volume of sales?

2. Does higher number of owners dictate higher number of employees?

3. Ethnicity categories don’t quite have same share in total number of employees. I’ll check 5 states data with highest number of employees.

**Discoveries**

During our work, several interesting findings were made that either proved or rejected validity of previously chosen hypothesis. We, also, separated those findings based on each member’s research:

**Luis:**

**Chart

Description automatically generated** For the first question that I was trying to answer, the results were what I expected and therefore were not really astonishing to me. For the first question I decided I wanted to focus on the first five states by alphabetical order because if I tried to show the data for each state, the visualization would not fit, and it would be too crowded. I also decided to focus on firms with the most years in business and least years in business because looking at every one of the categories in the “Years In Business” column would make my visualization too crowded and difficult to read. Also, looking at the firms with least experience and most experience would pave the way for interesting results. From the results it is safe to conclude that the firms that have been in business between 11-15 years have a greater number of employees than those firms that have been around for less than two years. This makes sense because the longer a firm is around, the greater the likelihood of expanding the amount of employees.

For my second visualization, I wanted to focus on whether there is a relationship between the number of owners of firms and the number of employees for each state. However, from the results we can see that there as the number of employees by state descends, the number of owners for the firms in the states does not descend. This means there is not a relationship between the number of owners in a firm and the number of employees in those firms. For my visualization below the dark green represents the number of employees and the light green represents the number of owners.

Chart, text

Description automatically generated with medium confidence

For my third visualization, I wanted to figure out how many employees in the first 10 states are minorities for the firms in those states. I decided to only focus on the first 10 states because the visualization would be too long to be able to understand if I did this for all states. The visualization gives you an idea of what states have the most number of employees for their firms and which states have the least. However, this visualization is strongest when it is used along with the table that I created so that you can see the exact numbers for each state. It is of no surprise that out of the first 10 alphabetical states, California has the most number of employees for its firms. It number is significantly greater than the other 9 states.

Chart, line chart

Description automatically generated with medium confidence

**Will:**

A boxplot of payroll by owner sex, to see initial discrepancies:

Chart, box and whisker chart

Description automatically generated

Clearly, these are all right skewed, with the amount of skew being higher for Male owned companies.

Two bar graphs of median Payroll by owner sex:

Chart, bar chart

Description automatically generated

Chart, bar chart

Description automatically generated

We see the difference is starker for the us data as opposed to the states data. (note different scales)

Sales vs Employees Scatter plot (zoomed in):

Chart

Description automatically generated

Chart, scatter chart

Description automatically generated

Of note, a number of male owned companies have rather higher sales per employee than the general trend. Why this occurs would be an interesting question for further study.

Also, the high-end outliers seem to also be above the lower-end trend, this is also worth looking into with a larger dataset.

Payroll vs Sales scatter plots (full and zoomed in):

Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generated

We see that a number of male owned businesses have a lower ratio of Payroll to Sales than the general trend. These also seem to be the same Sales numbers as those of the higher performing companies from the previous section. This is interesting, and the cause thereof also worthy of further study.

Also, the high-end outliers seem to also be below the lower-end trend, this is also worth looking into with a larger dataset.

Payroll vs Employees scatter plot:

Chart

Description automatically generated

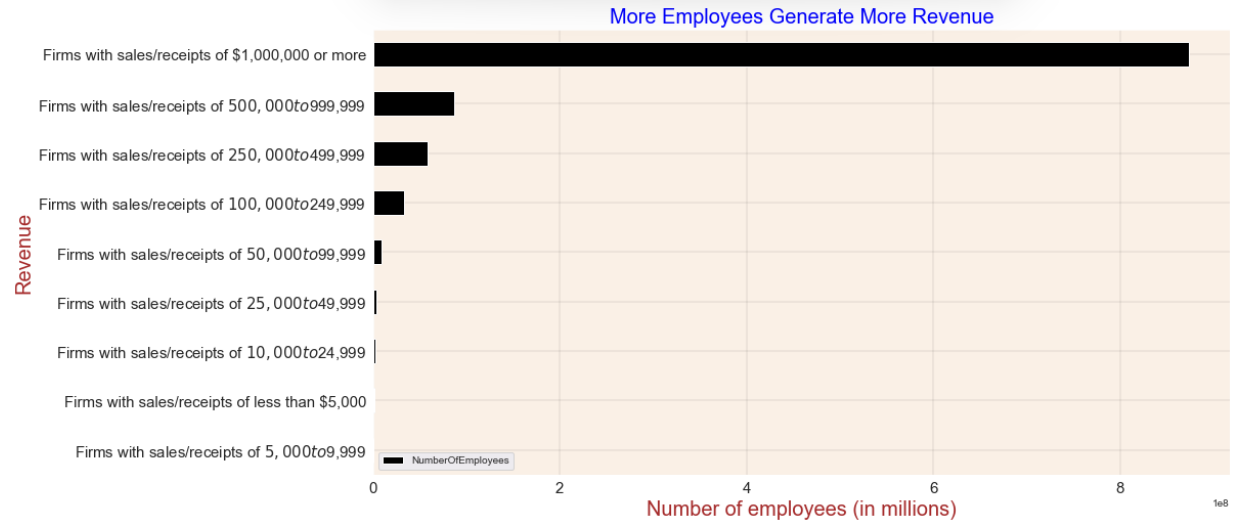
Overall some companies are a bit higher than the trend and some are lower but not in any seemingly significant way (the outlier trend seems to match the rest of the data) implying the payroll to employee ratio is fairly standard, which is somewhat curious, and worthy of further study.

**Phil:**

Please list your findings and visuals numerically based on order of questions mentioned above. Don’t forget to comment your findings below each of your visuals.

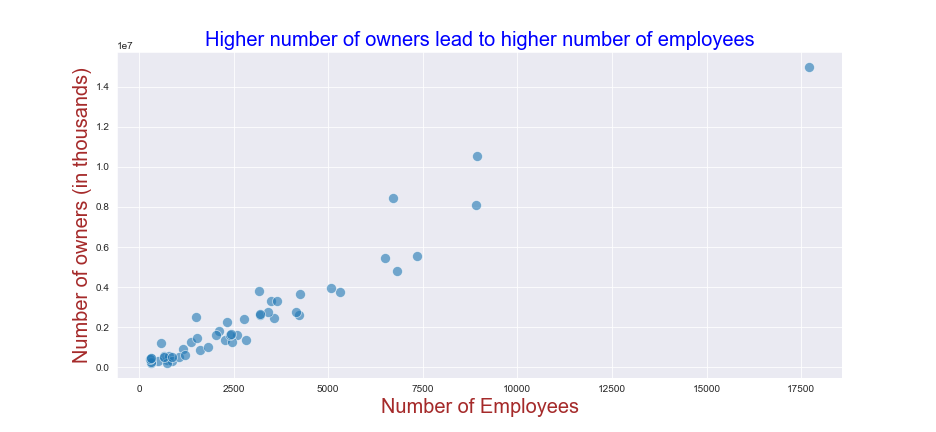
**Sharif:**

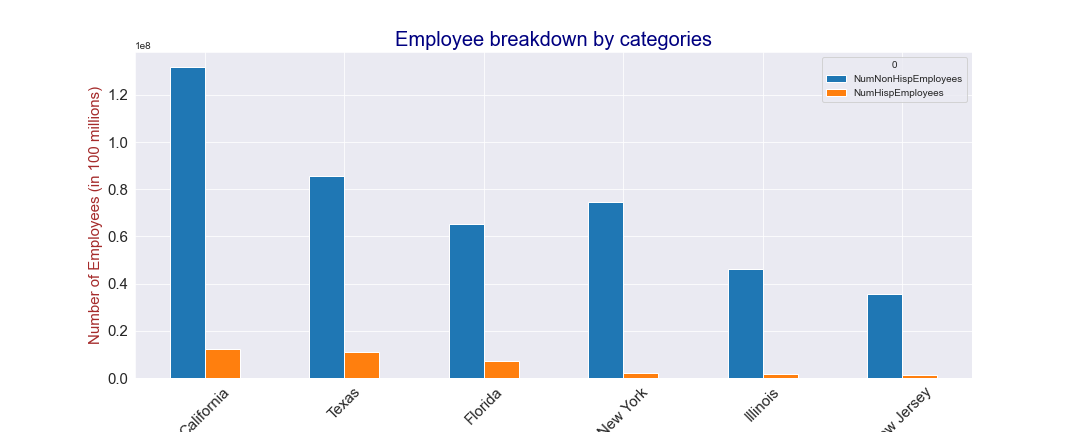
1. Based on my analysis I discovered that higher number of employees do dictate higher volume of sales:



As seen from visualization above, while higher number of employees lead to higher volume of sales, that impact is not same for all group of employee sizes. Companies with sales volume up until a million account for only 400-500 thousand of employees, while companies making over a million may count for up to 9-10 million of all employees.

2. It seems that when more owners involved in a business the more employees they tend to hire:



From plot above we can evidence positive correlation between number of owners and number of employees. It takes more owners to oversea the business and its increasing number of employees. However, some of those owners may represent a pull of owners that just bought ownership shares without actual management stake in company.3. When checking data for 5 largest states based on number of employees, I found the following to be true:

I confirmed that share of employee categories within total number of employees are not the same. As seen above, employees of Hispanic origin only represent 1/7th -1/8th of total employees in 5 largest states based on number of employees. This may indicate inequal access to jobs to different categories of employees, however this may also just be because the small share of those minority categories within total population of the states too.

**Conclusion**

ABS dataset serves as a powerful tool to monitor changes in private sector for different consumers of that data and use it to make necessary changes in their course of work. Conclusions we made may bring some light to an interested party’s work or find solutions to a potential problem.

**Luis**:

The first question I was trying to answer was: Comparing firms with least amount of years in business to firms with most amount of years in business, do the number of employees grow as years in business grows? (first 5 states). For this question I saw that yes, as years in business grows, so do the number of employees within that business. My second dealt with finding whether there is a relationship between the number of owners for all firms and the number of employees for all states. The answer to that was no. As the number of employees is decreasing, you can see that the number of owners goes up and down, so there was no relationship there. For my last question I was trying to figure out what the total number of minority employees in all firms for the first 10 states was. Although my visualization came in handy to give you an idea of the answer to my question, my visualization is strengthened by the table I created for it.

**Will**:  
 Overall, most large companies are owned by men, as well as the companies which most diverge from the overall trend between employment numbers and sales, which raises very interesting questions going forward about if that discrepancy is caused by some other factor, possibly industry, or just a holdover from times before women were a major factor in the corporate workforce (the old boys club nature of business). One interesting trend between the state and us data was that while the graphs have fairly similar shape though drastically different magnitudes, the differences are more pronounced in the us data as opposed to the state data. This could be a result of the state dataset being larger, as well as caused by other not looked into reasons, but it does raise questions going forward, most importantly: “In which states is this difference least pronounced?” Finally, while male owned companies more often outperform the general projection on sales by employment, they slightly underperform when looking at payroll vs sales. While these companies are right in line with the payroll/employees trend, having significantly higher sales then others with similar employee numbers but having a lower payroll/sales ratio implies the company is devoting a lower percentage of revenue to its workers than the general trend, and these seeming to only be male owned businesses is intriguing.

In the end, most of the more lucrative companies in the survey are owned by men, which isn’t really surprising, and overall, the sex of the owner appears to have little bearing on the employee compensation, but in a select range of company sizes and sales there are some deviations which would be worth further study.

**Phil**:  
 Add your conclusion based on your work

**Sharif**:

Generally, my findings allow me to conclude that my hypothesis were confirmed by the analysis of data that Census ABS provided. Higher number of employees can generate higher volume of sales. There is a pattern where businesses with higher number of owners also have higher number of employees. And, lastly, Ethnicity categories are not equally represented within total number of employees in businesses.  
 Census data is quite challenging to work with and takes time and effort to explore and find necessary piece of material. However, it was very useful experience and hope makes future data exploration an easier task.