Project Report

Census API Data - Annual Business Survey 2019

M08: Pandas and Visualizations

Dev10 - Group 4

Sargis Abrahamyan, Justin Bartell, Ben Hines, Lindsey Oh

17 January 2022

Introduction

This report features visualizations built from the United States Census Bureau's 2017-2019 Annual Business Survey (ABS) to help show the greatest predictors of business success. Business success is defined in this report purely by average annual pay per employee, and it does not consider other factors such as: employee size of business, years in business, revenue, etc. The predictors investigated in this report include owner demographics, geographic location (state), and industry.

Feedback Utilized

Because we heavily modified our questions and created completely new visualizations from the Creating Visualizations While Exploring Data Exercise, some of the specific feedback given to us could not be incorporated. However, some of the general feedback we received and incorporated were:

Yihua Liu: "make sure graphs have titles (try to put something more descriptive than just the column name), labeled axes, legends, etc."

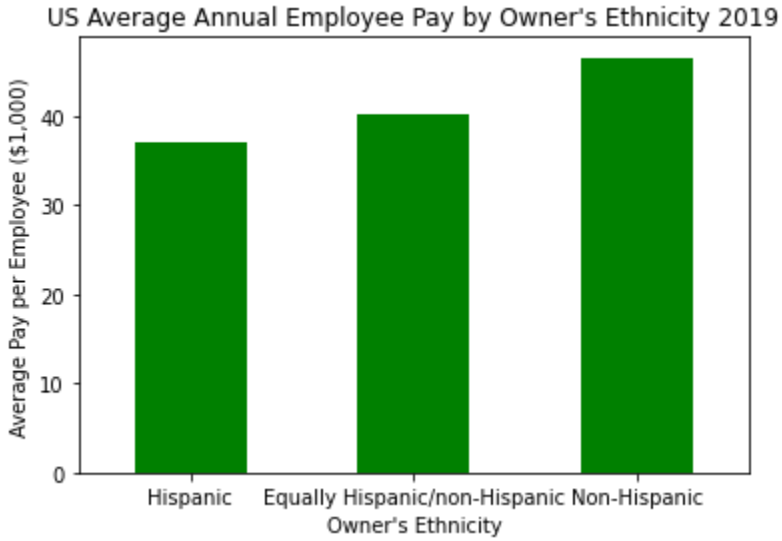
Thomas Seeber: "Consider adding a regression line to scatter plots if the data looks linear. The Legends don't make sense with some of the bar charts and try to rename features to be easy to understand labels. Try to experiment with non-default colours."

Owner Demographics vs. Employee Pay

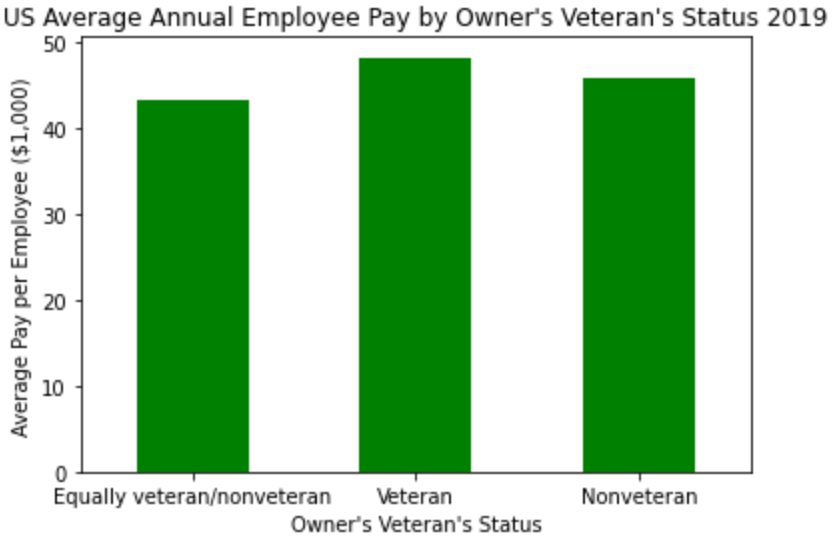
How does an owner's demographic information predict the success of their company, as defined by annual pay per employee?

The owner demographics information available in the ABS Characteristics of Business Owners dataset are sex, ethnicity, race, and veteran status. Other information related to the type of ownership of business included whether a business was the number of owners, whether the business was family owned, and whether spouses were involved in the business. This section focuses only on 2019 data.

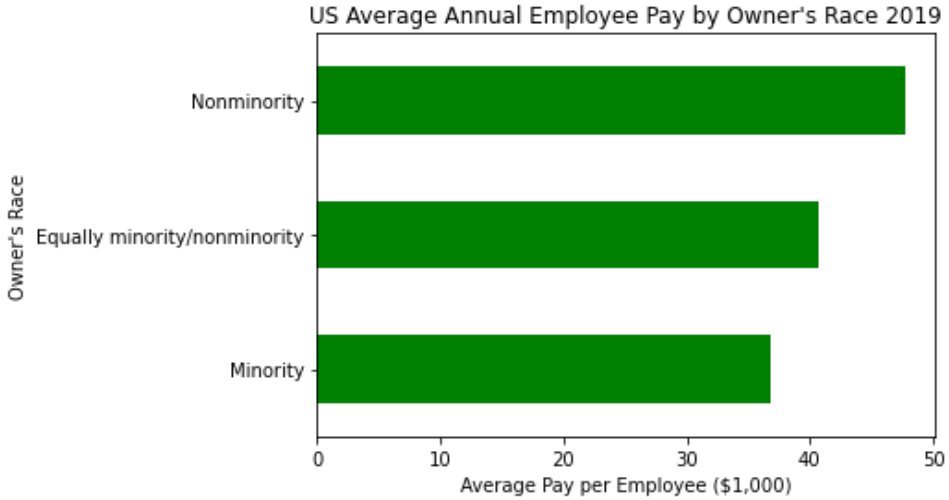
There appears to be differences in average annual employee pay based on the owner's sex. While employees were paid on average $49,352 from male-owned businesses, they were only paid on average $37-39,000 from female-owned and equally male/female-owned businesses.



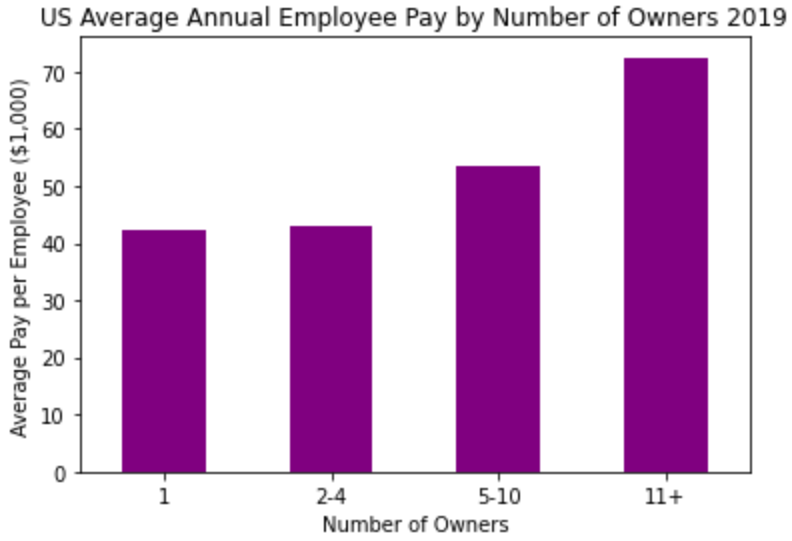
Differences in average annual employee pay also appear to occur based on the owner's ethnicity, with higher pay from non-Hispanic businesses. Non-Hispanic businesses pay their employees, on average, $46,570 while Hispanic businesses pay nearly $10,000 less: $36,975. Businesses that are owned equally between Hispanic and non-Hispanic owners paid their employees $40,286.



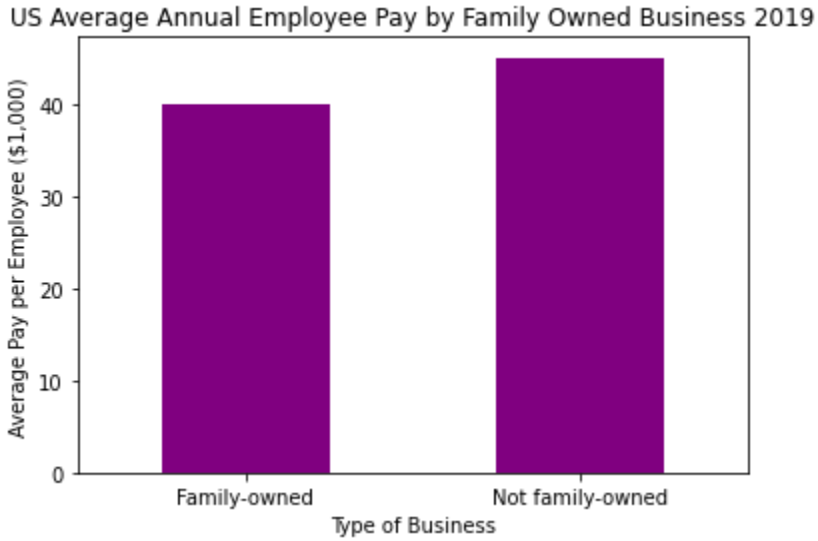
Differences in average annual employee pay also appear to occur based on the owner's veteran status. Businesses owned by veterans paid their employees the most, on average ($48,306), while businesses owned equally between veterans and nonveterans paid the least ($43,396). Businesses owned primarily by nonveterans paid their employees $46,005 on average.



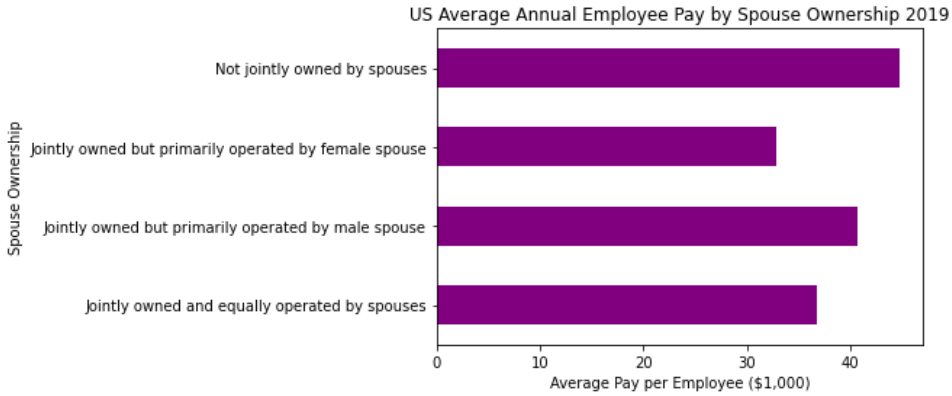
Whether the owners were primarily owned by minorities also seem to lead to differences in average annual employee pay. Businesses owned by nonminority owners paid their employees, on average, over $10,000 annually. Nonminority owners paid their employees $47,775 while minority owners paid their employees $36,847. Businesses equally owned by nonminority and minority owners paid their employees $40,742.



The number of owners a business was owned by is also related to the average annual employee pay, with a greater number of owners correlating with higher pay. If a business was owned by just a single owner, it paid their employees on average $42,186. A business owned by 11 or more owners paid their employees, on average, $72,475.



Whether the business was family-owned also seems to lead to disparities in employee pay. Non-family-owned businesses tended to pay their employees over $5,000 more annually. Family-owned businesses paid their employees $39,892 while non-family-owned businesses paid their employees $45,015.



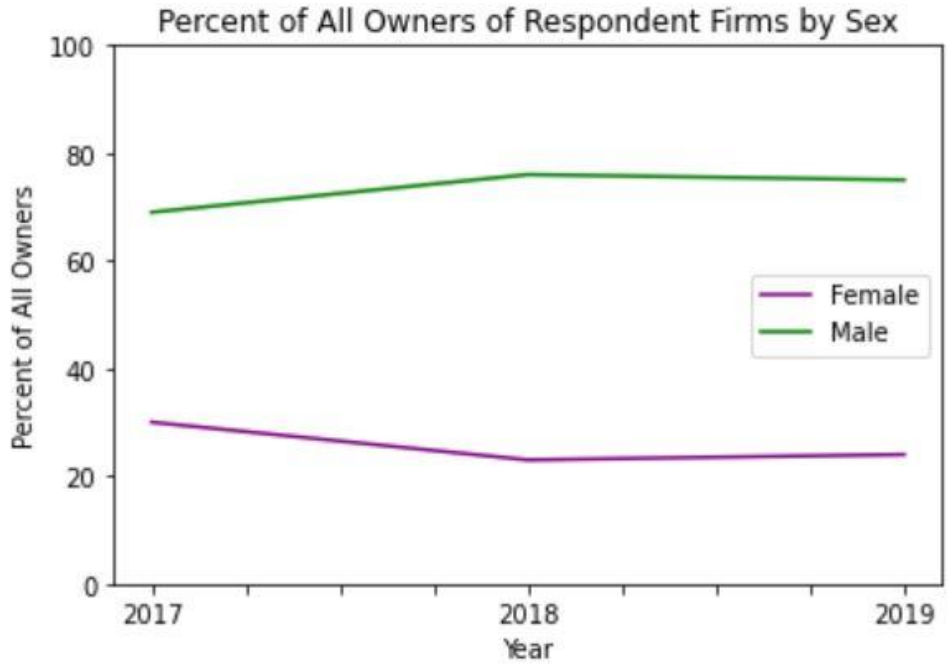
Finally, whether the business was operated with spouses—and whether a male or female spouse was the main operator—seemed to lead to disparities in employee pay. Businesses not jointly owned by spouses had the greatest pay ($44,755). Businesses jointly owned by male spouses had the next highest ($40,727), and businesses jointly owned by female spouses had the lowest ($32,921).

Overall, owner demographics lead to disparities in average annual employee pay—some as great as 20% discrepancy. The statistical power of this is uncertain, since both the total employee pay and number of employees in each of these fields have a standard error that ranged from 0.7-12.6%. Distributions of a ratio of two random variables lead to an undefined standard error, making statistical significance of this data difficult to determine.

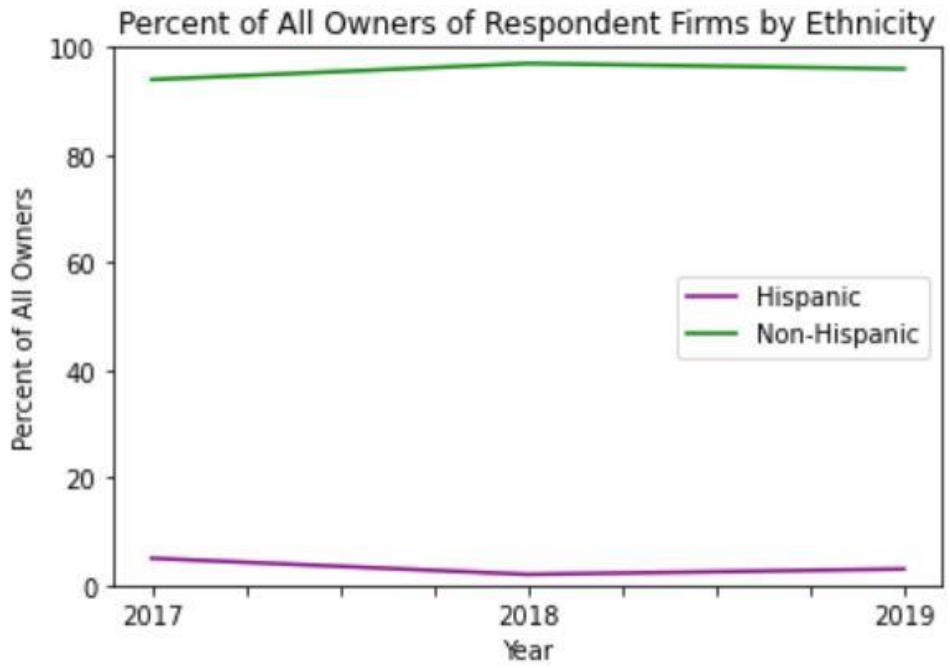
Owner Demographics Over Time

How have the demographics of business owners changed between 2017-2019?

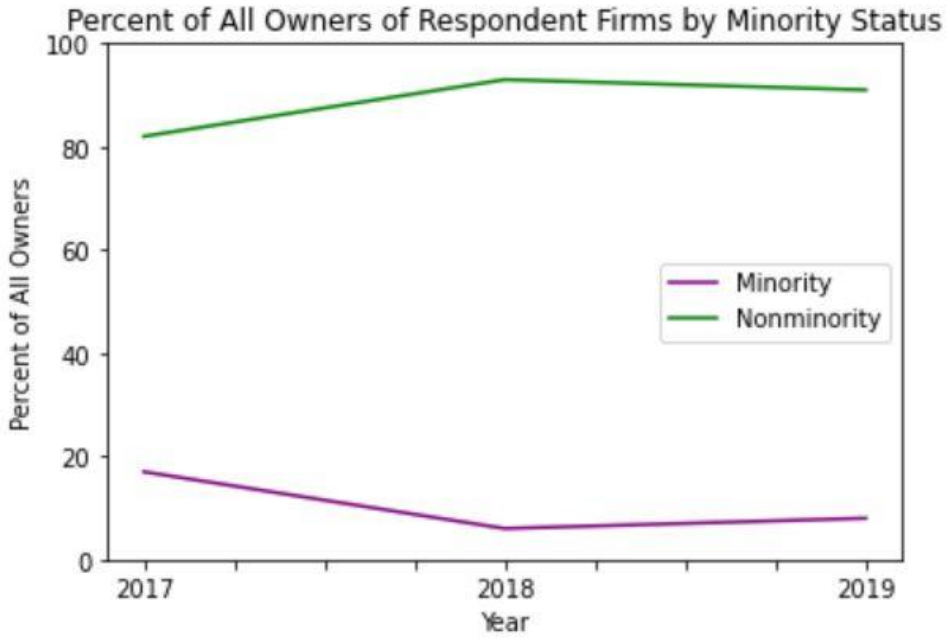
The demographics information available in the ABS Characteristics of Business Owners dataset are sex, ethnicity, race, and veteran status. Because the dataset spans only three years, there is limited insight that can be gleaned from a regression line. Therefore, the regression line was not included in the visualizations. The number of owners of respondent firms dropped drastically after 2017. The respective number of owners of respondent firms are: 3,264,885 in 2017; 146,045 in 2018; 133,653 in 2019. To compare across years, percentages were calculated and graphed.



In 2017, the approximate percentages for female and male business owners were 30% and 70%, respectively. In 2018, the approximate percentages for female and male business owners were 23% and 77%, respectively. In 2019, the approximate percentages for female and male business owners were 24% and 76%, respectively. Over the three years, the percent of all respondent female owners decreased slightly while the percent of all respondent male owners increased slightly. If the assumption is that the balance ought to be 50% female, 50% male, then the percent of respondent female business owners would have to increase by about 26% from its value in 2019.

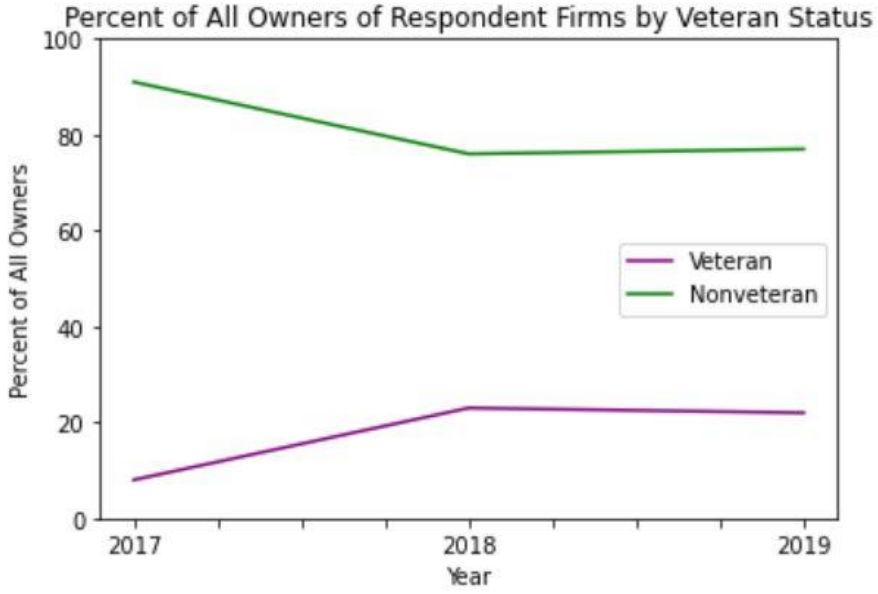


In 2017, the approximate percentages for Hispanic and Non-Hispanic business owners were 5% and 95%, respectively. In 2018, the approximate percentages for Hispanic and Non-Hispanic business owners were 2% and 98%, respectively. In 2019, the approximate percentages for Hispanic and Non-Hispanic business owners were 3% and 97%, respectively. Over the three years, the percent of all respondent Hispanic owners decreased slightly while the percent of all respondent Non-Hispanic owners increased slightly. If the assumption is that the balance ought to be 50% Hispanic, 50% Non-Hispanic, then the percent of respondent Hispanic business owners would have to increase by about 46% from its value in 2019.



When this survey was conducted, business owners were instructed to select all from the following list that applied: American Indian and Alaska Native, Asian, Black or African American, Native Hawaiian and Other Pacific Islander, White. As a result, the corresponding data did not lend itself well to percentages. The race categories Minority and Nonminority, however, were structured such that the components added up to the total. Therefore, this data was chosen for this visualization.

In 2017, the approximate percentages for minority and nonminority business owners were 17% and 82%, respectively. In 2018, the approximate percentages for minority and nonminority business owners were 6% and 94%, respectively. In 2019, the approximate percentages for minority and nonminority business owners were 8% and 91%, respectively. Over the three years, the percent of all respondent minority owners decreased noticeably while the percent of all respondent nonminority owners increased noticeably. If the assumption is that the balance ought to be 50% minority, 50% nonminority, then the percent of respondent minority business owners would have to increase by about 41% from its value in 2019.



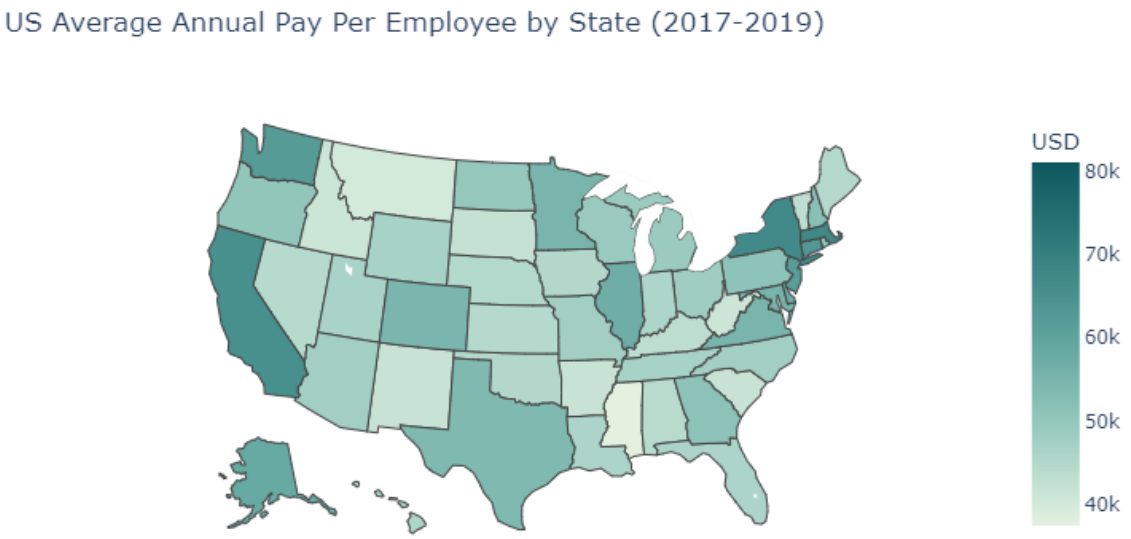
In 2017, the approximate percentages for veteran and nonveteran business owners were 8% and 92%, respectively. In 2018, the approximate percentages for veteran and nonveteran business owners were 23% and 77%, respectively. In 2019, the approximate percentages for veteran and nonveteran business owners were 22% and 78%, respectively. Over the three years, the percent of all respondent veteran owners increased noticeably while the percent of all respondent nonveteran owners decreased noticeably. If the assumption is that the balance ought to be 50% veteran, 50% nonveteran, then the percent of respondent veteran business owners would have to increase by about 28% from its value in 2019.

In conclusion, the demographics of business owners have not changed significantly between 2017-2019. More can be done to promote inclusivity of marginalized groups in the population of business owners at large.

State vs. Avg Annual Pay Per Employee

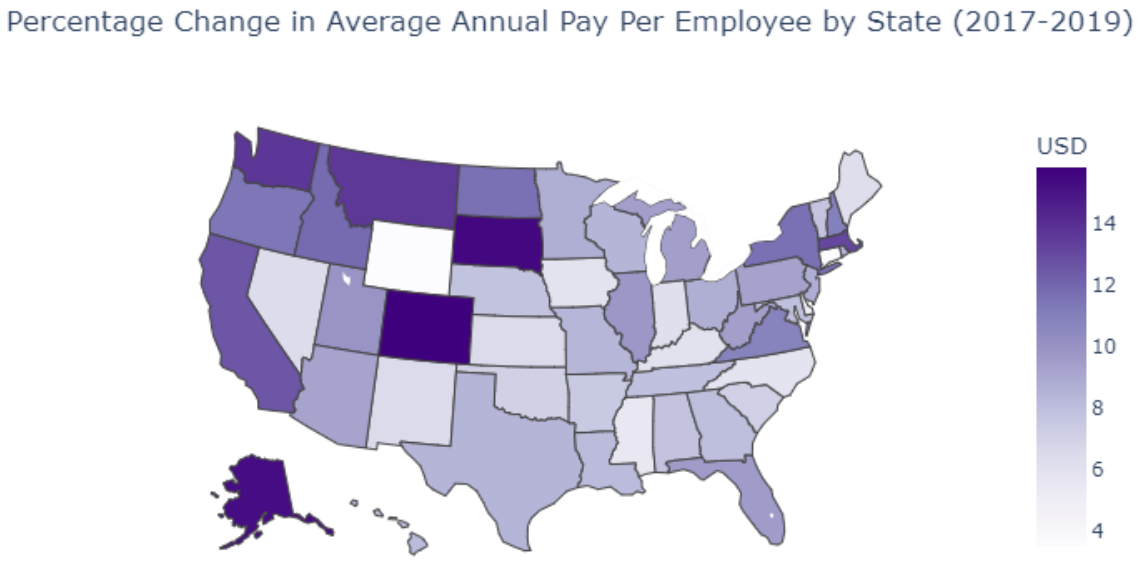
How does a business's state predict the success of their company, as defined by annual pay per employee?

To calculate average annual pay per employee by state, total annual pay in a state was divided by total number of employees in the state.



Between 2017 and 2019 with national average being $50,211, highest paid employees were in District of Columbia with average annual pay of $81,171, followed by Massachusetts $67,514, New York $67,032, California $65,603, and Connecticut $63,602. Lowest average paid employees were reported to be in Mississippi, Montana, and West Virginia with respective annual salaries $37,426, $39,987, $40,999.

To see in which states employee salaries are increasing, the percentage change of annual pay per employee for each state was calculated between 2019 and 2017.

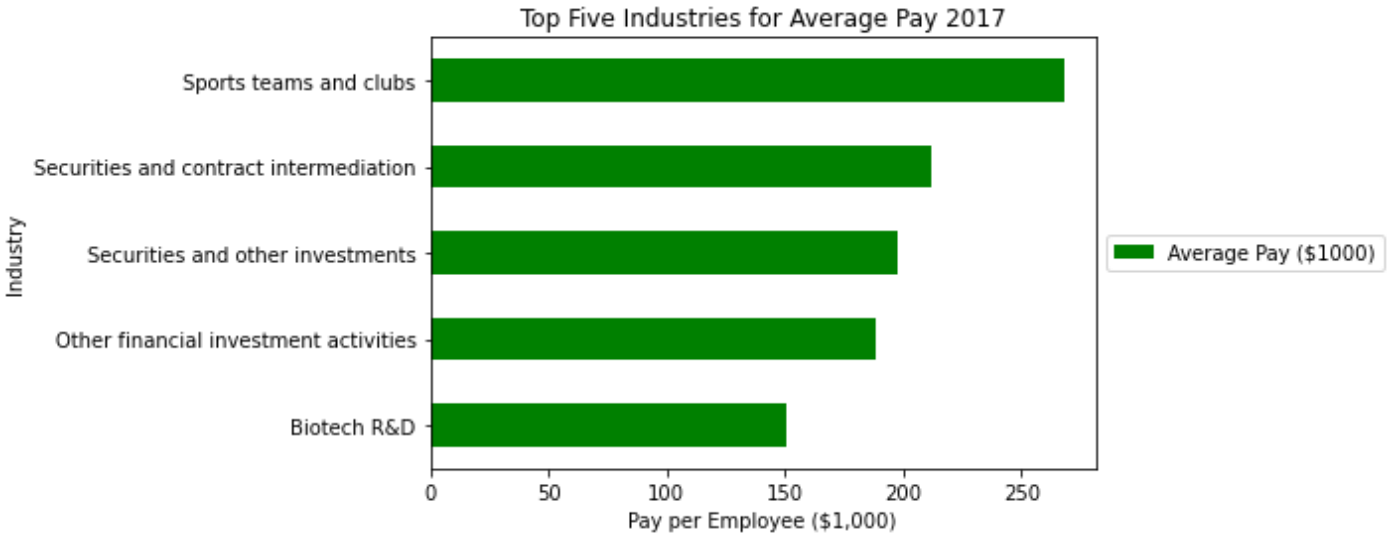


National average increase was 9.09 %. Highest increases in average annual pay were in Colorado, District of Columbia, South Dakota, Alaska with respective increases of 15.86%, 15.68%, 15.59% and 15.38 %. Lowes changes in average employee pay were in Wyoming 3.5%, Connecticut 3.55%, Delaware 4.81% and Mississippi 5.47%.

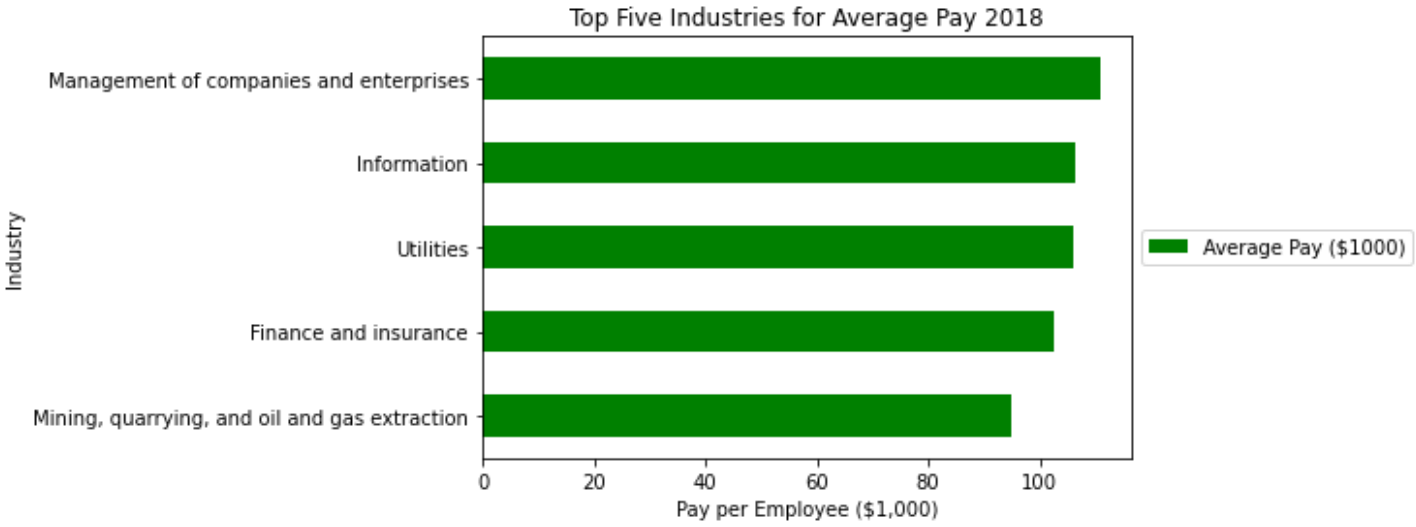
Industry vs. Employee Pay

How does a business's industry predict the pay of their employees between 2017 - 2019?

Top Five Industries for Average Pay 2017 - 2019

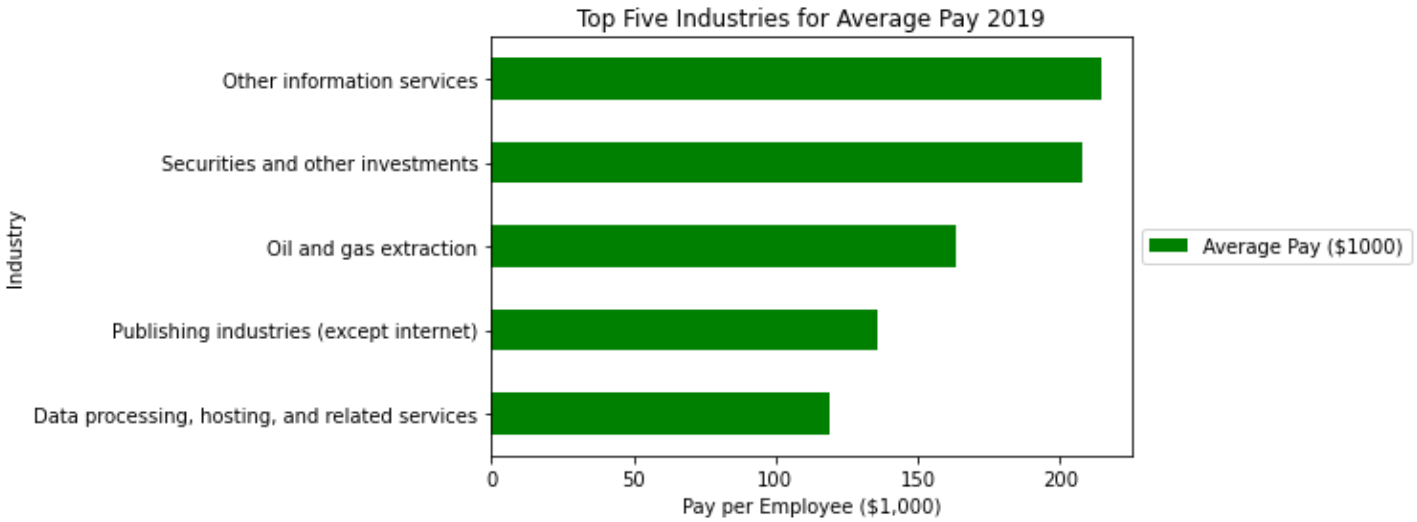


Seen here, the sports industry tops employee pay with an average in excess of $250,000 per year. What follows are sub-industries within the financial world hovering around an average of $200,000 per year. Research and development in biotech is fifth with an average of $150,000 per year pay for employees.



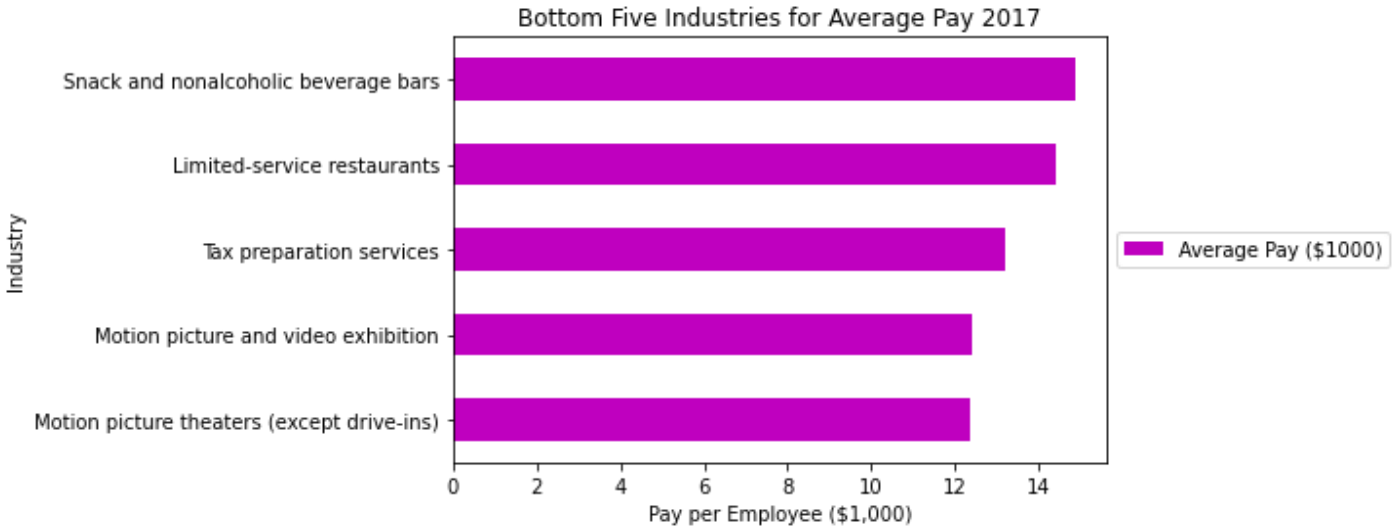
In 2018, company management tops average employee pay, at just above $100,000 per year. Something of note is that the financial sector, while still in the top five, displays a much lower average pay at approximately $100,000 per year. A new entry compared to 2017, mining and

resource extraction, is seen to show up at number five with an average pay of just below $100,000 per year.

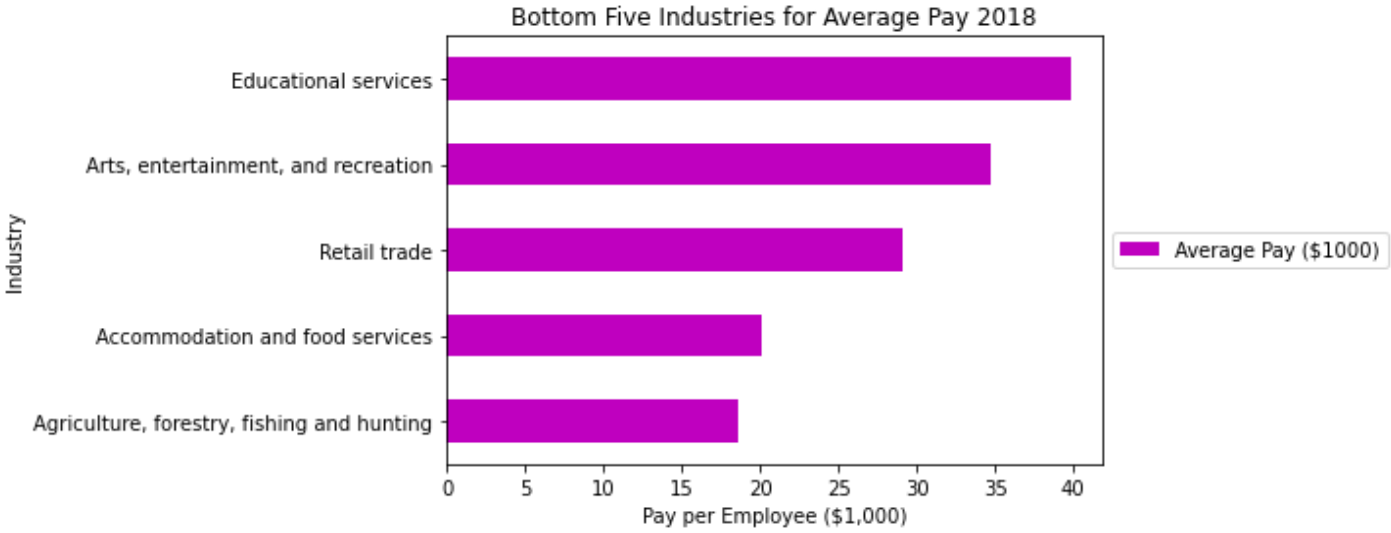


Top industries for pay in 2019 sees information services top in average pay, while the financial world is represented yet again by ‘Securities and other investments. Notable here is the return of resource extraction in the entry for ‘Oil and gas extraction’, with an average pay of above $150,000.

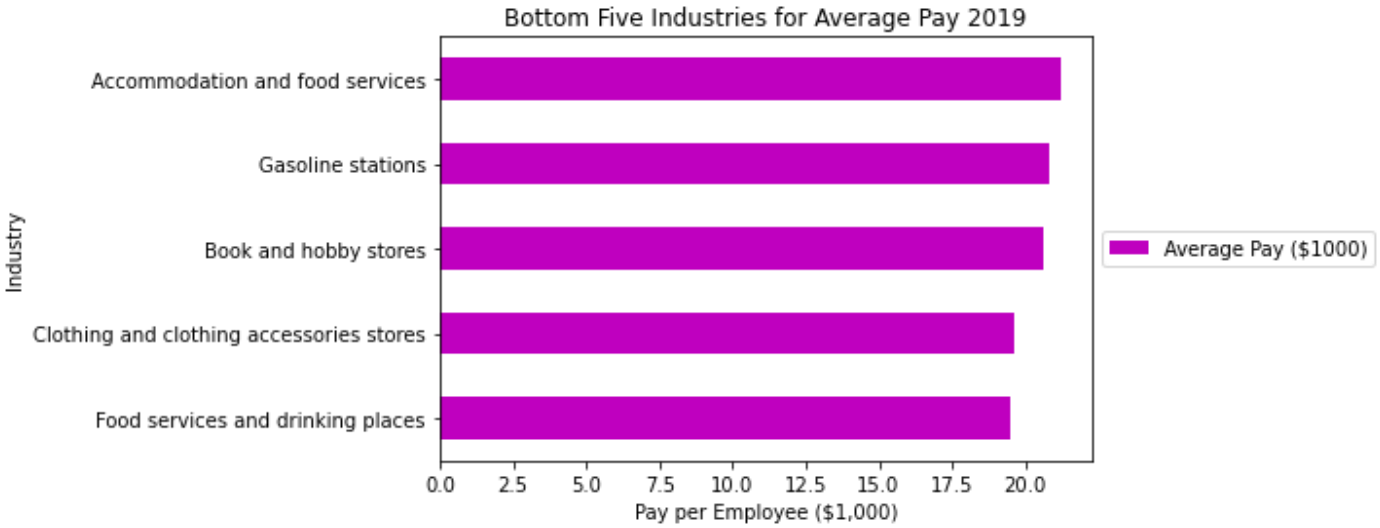
Bottom Five Industries for Average Pay 2017 - 2019



For the bottom five industries for pay in 2017, we see food service at approximately $14,000 per year, with movie theater workers at the bottom, making $12,000 per year.



In 2018 the education industry shows to pay an average of $40,000 per year, with the service industry represented again with ‘Accommodation and food services’. General retail, credited as ‘Retail trade’ is seen paying an average between $30,000 and $35,000 per year.



In 2019 the service industries show again amongst the lowest paying industries, along with other service and retail industries, such as ’Gasoline stations’ and clothing retail.

Overall, a general picture is painted of financial industries, along with management, and resource extraction, as being the most consistently high paying. This is contrasted with low paying industries populated most often with companies from the retail and service sectors. While this general pattern repeats across all three years, there is variability in how the industries are explicitly described and categorized. For example, 'Accommodation and food services’ in the

2018 chart seems to have been relabeled as ‘’Food services and drinking places’ in 2019. What is likely happening is the census participants are varying their responses as to how their given company is categorized across years, and therefore the specific label varies, but the nature of the industry which the response captures remains consistent. However, it must be taken into account that the data does not capture the work status of the represented employees. It can be assumed that the financial sector employs full-time workers more consistently, while service and retail sectors might see a high incidences of part-time workers, and therefore pay within these sectors may be artificially deflated when directly compared. Ultimately, pay is nearly an order of magnitude lower in the service sectors than it is in finance, management, and other top paying industries across all three years.

Conclusion

Employee pay appears to be affected by owner demographics, geographic location by state, and industry of the business. In general: male, non-minority, veteran, multiple-owner, non-family businesses tended to pay their employees the best. There are discrepancies among the states of upwards of over $40,000 per year. While average pay per industry tended to change year-to-year, the highest paid industries tended to be paid around $150,00-$200,000 while the least paid industries tended to pay around $20,000 per year.

Causation was not explored throughout this analysis, nor could a predictive model be created from the aggregated data given from the Annual Business Survey’s APIs. Statistical significance was not calculated based on the difficulty of calculating standard error on the ratio of two random, non-independent random variables. Further, because this data was aggregated, it was not possible to test for outliers.

References

U.S. Government (2021, October 14). *Annual Business Survey (ABS) APIs*. United States Census Bureau. Retrieved January 14, 2022, from https://www.census.gov/data/developers/data-sets/abs.html