

## PROJECT REPORT

This report will explore various questions relating to census data from the 2019 Annual Business Survey (ABS). It includes the ETL process used to collect and parse the data from the census API, as well as visualizations showing what we discovered.

The Company Summary(CS) dataset includes data from 2018 and we used it at the granularity of the U.S. as a whole and at the state level. It has many categorical columns to separate out aggregations for different business types. These include demographic data on the business owner(s), buckets of employment sizes, buckets of business revenues, buckets of years in business, and different industries. We cut out the numeric code columns for these categories, keeping just the human-readable labels. The data value columns are totals for different categories of business that include the number of employees, the number of businesses, annual payroll, and annual sales or revenue.

The Characteristics of Business Owners(CBO) dataset is similar to the Company Summary dataset, but with more of a focus on the business owners. The main differences are that it includes more demographic and categorical columns about the owners and that it only has the total number of owners for its data values.

Bureau, US Census. "Annual Business Survey (ABS) Apis." *Census.gov*, US Census Bureau, 28 Oct. 2022, <https://www.census.gov/data/developers/data-sets/abs.2019.html>.

### ETL Process

See Jupyter Notebook

### Findings

In this section, we will explore the questions we had, showing visualizations, and either answer those questions, or indicate if more research is needed.

#### Questions:

What are the factors that affect pay in the United States?

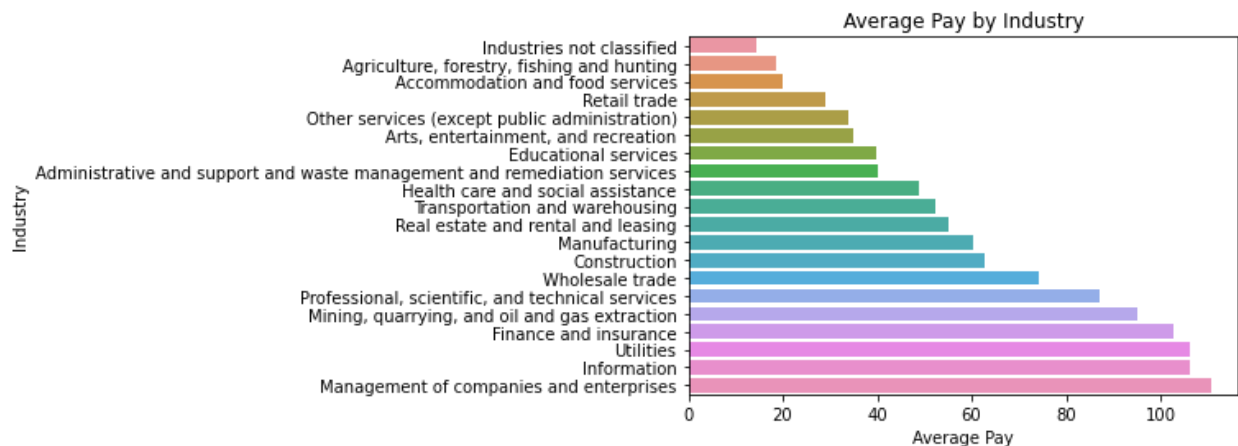
What are the characteristics of business owners based on race?

Is there a relationship between location and how successful a business is?

Is there a relationship between the race of the owner and how successful a business is?

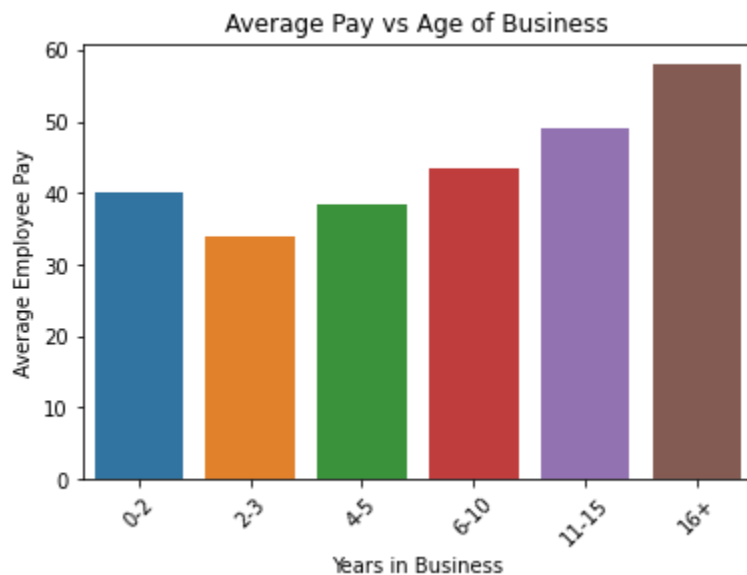
Is there a relationship between the gender of the owner and how successful a business is?

We investigated a few different business characteristics that could affect average pay. First, we looked at differences between industries to see if there was any significant influence.

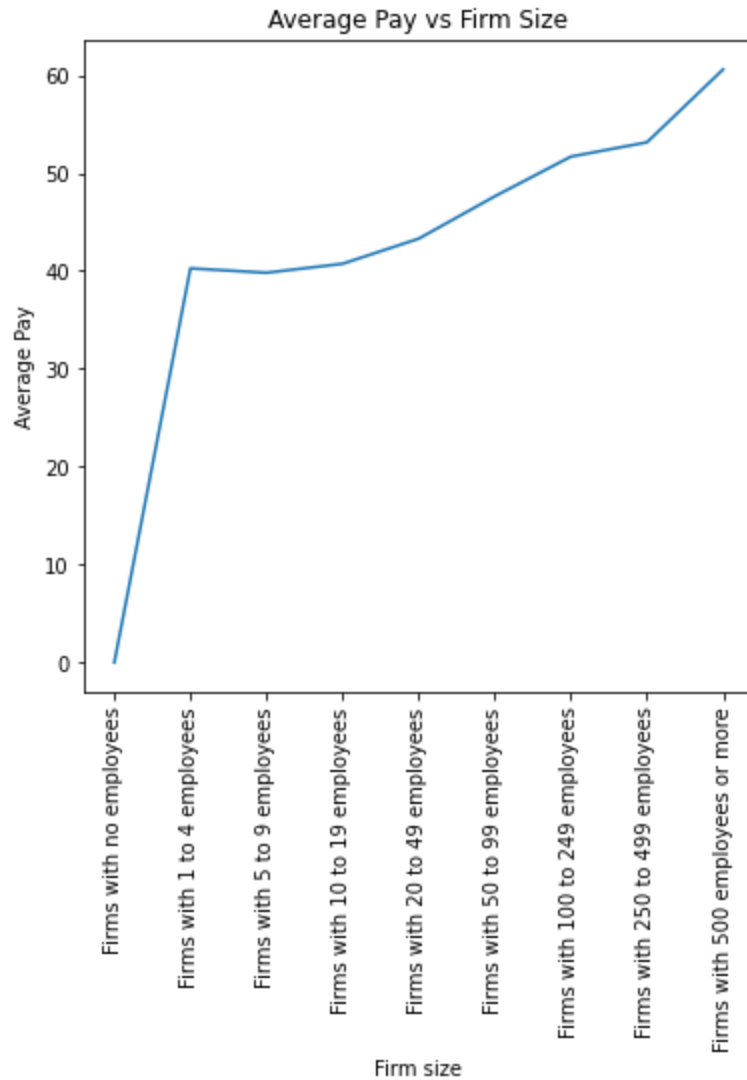


The industries with the highest average pay (annual payroll/number of employees) are, in ascending order, Finance and insurance, Utilities, Information, and Management of companies and enterprises. The industries with the lowest average pay are those businesses that are not classified, Agriculture, forestry, fishing and hunting, Accommodation and food services, and Retail trade. None of the highest or lowest paying industries are especially surprising. The utilities came as a bit of a surprise, and I would like to know some examples of businesses were not classified. From this data, we can conclude that average pay varies significantly from industry to industry.

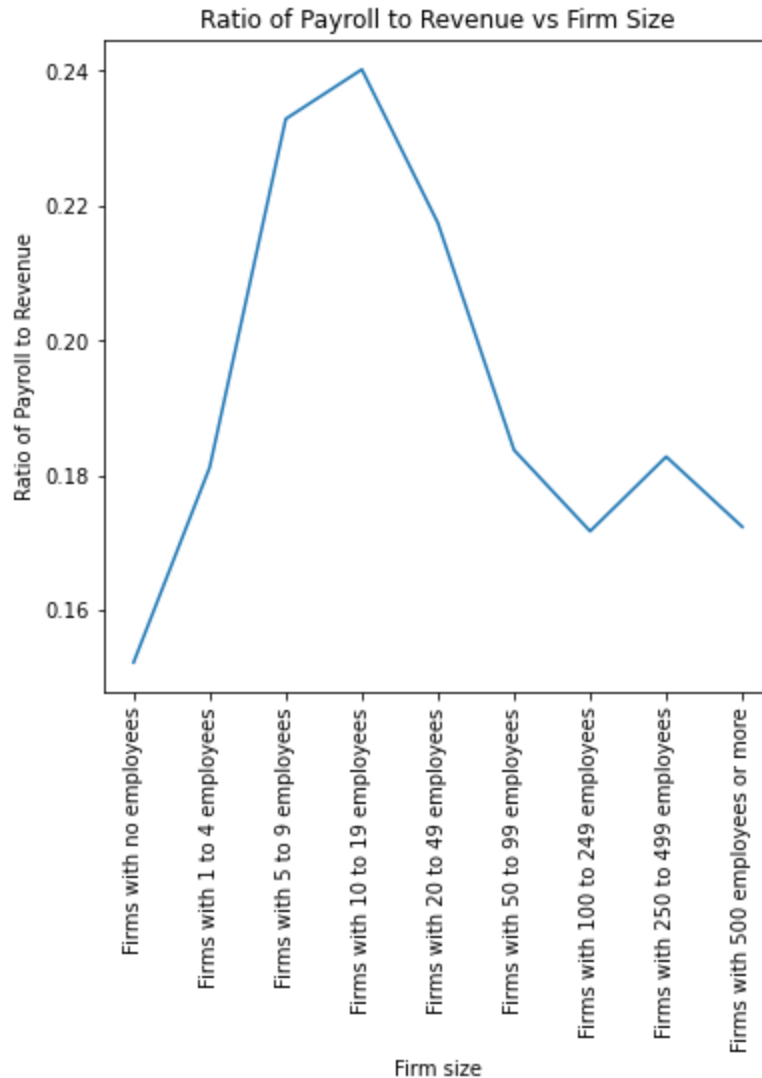
Then, we looked at ages and sizes of business firms.



Except for the most immature of businesses, average pay steadily increases with the age of the business. This does not necessarily mean that the median pay increases, because of the exponential growth of upper management salaries, which could be outliers that increase the mean.



As firms get larger, average pay gets higher. However, similar to the previous graph, this could be skewed by outlier upper management salaries, which are often several orders of magnitude larger than median employee salaries.

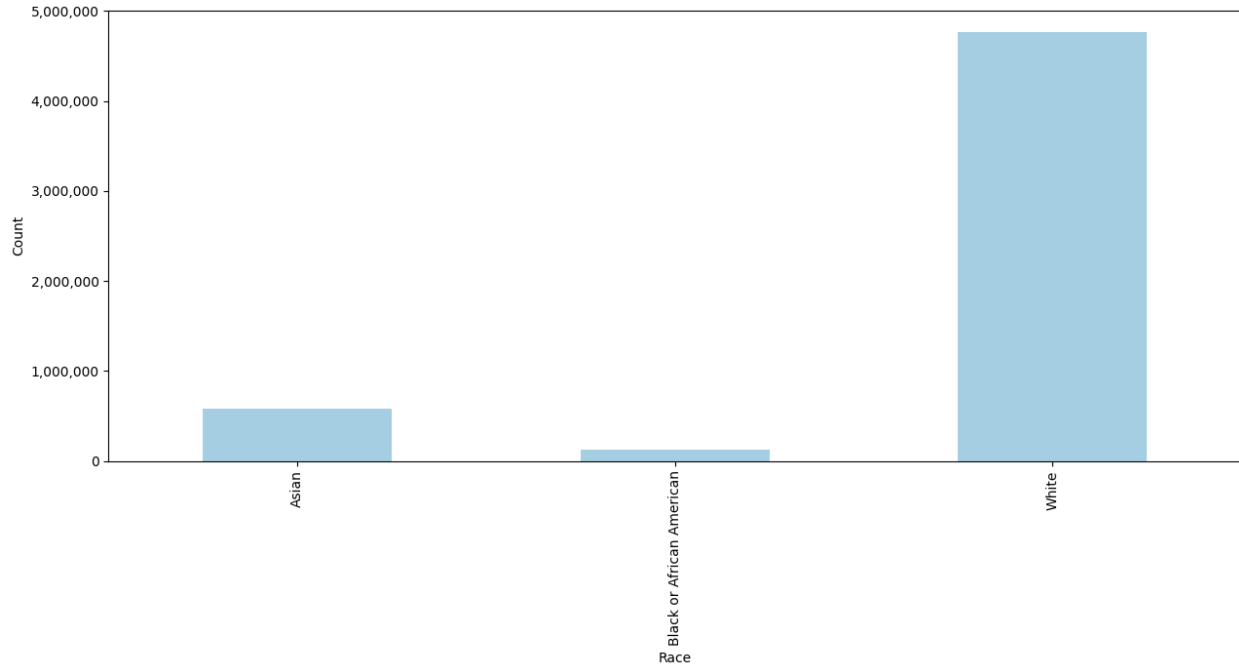


The ratio of payroll to revenue shows us what proportion of a company's total revenue is spent on salaries. I assume with finer granularity on the right side of this graph, there would be a smoother trend down, then up, then down again as firms become huge, and likely spend a lesser proportion on their payroll.

The oddity in this graph is that firms with no employees still have a payroll, so there is something not quite right there. The only way that this works is if the business owner is drawing a salary themselves. However, from what I could find about tax law, the owner counts as an employee if they draw a salary, and if they take an owner's draw instead, that does not end up on payroll. So, it is confusing how that data is collected.

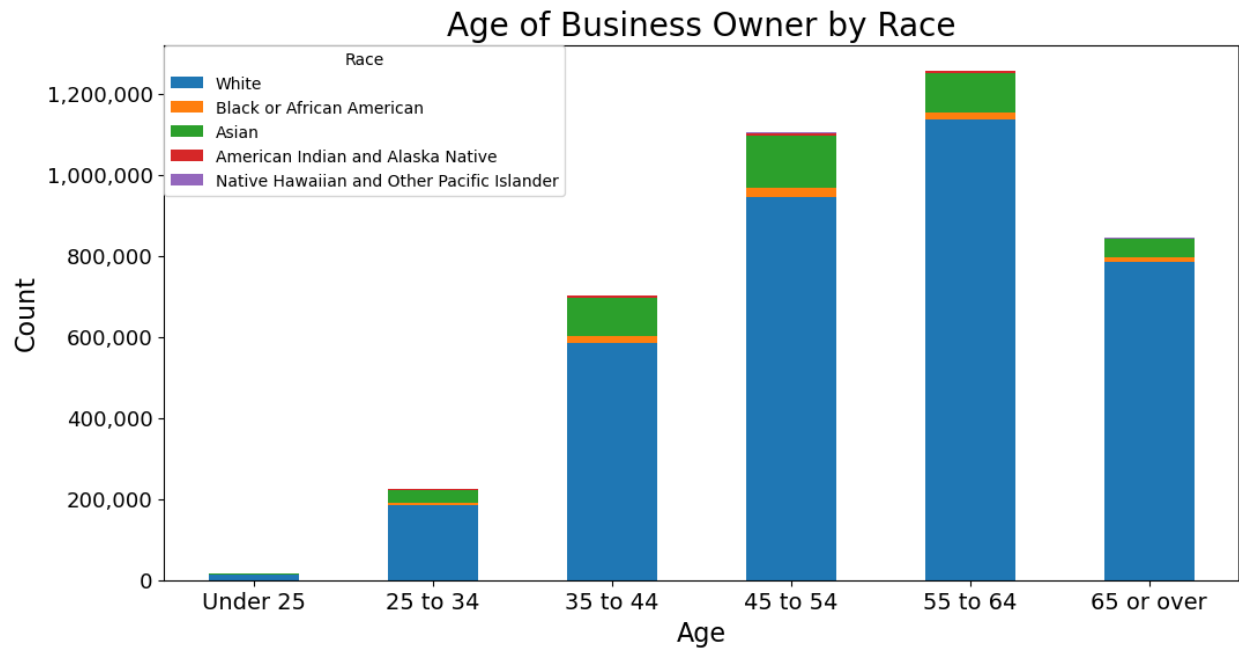
In conclusion, annual pay in the United States is significantly impacted by the industry of the business. For the other business characteristics of age and size, the graphs appear to show a direct correlation, but this correlation would be proven better if we had access to median pay data, since the highest earners in a business may be outliers that shift the mean higher. Further information on how the firms with no employees are using their payroll would be useful as well. In later research, it would be useful to look at how owner demographic data impacts average pay.

**What are the characteristics of business owners based on race?**  
**Amount of Business Owners by Race Group**

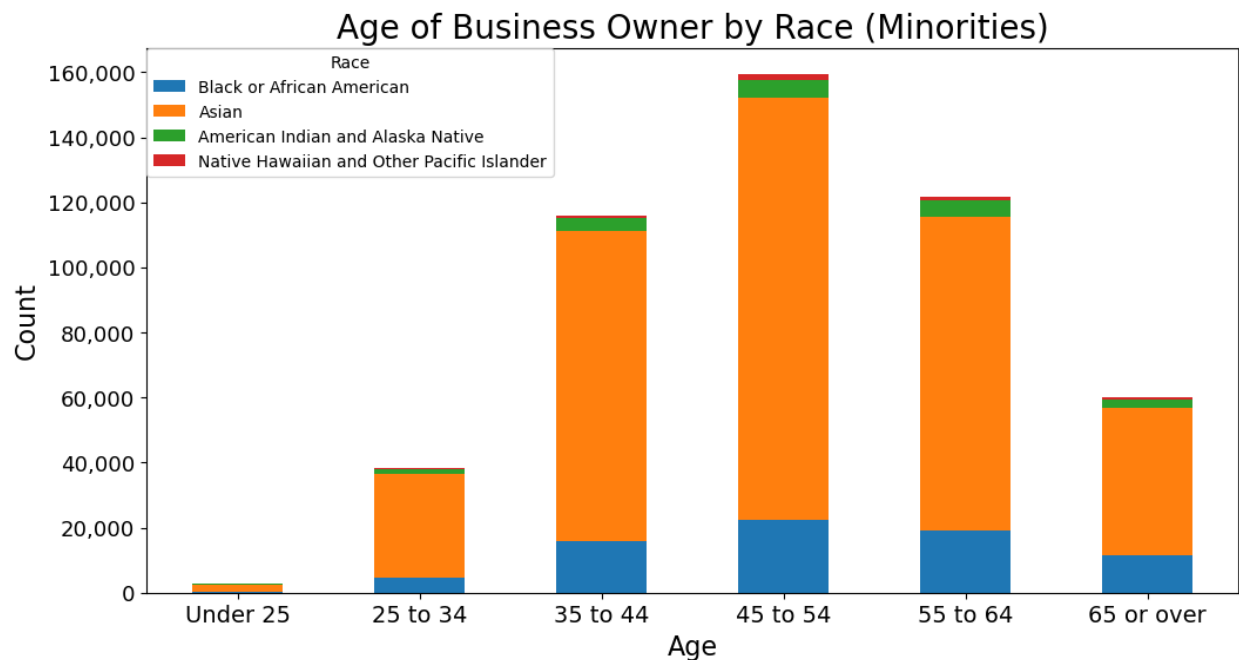


The visual above is displaying the distribution of races among business owners. There is a greater number of White business owners than any other race.

## Age of Business Owner

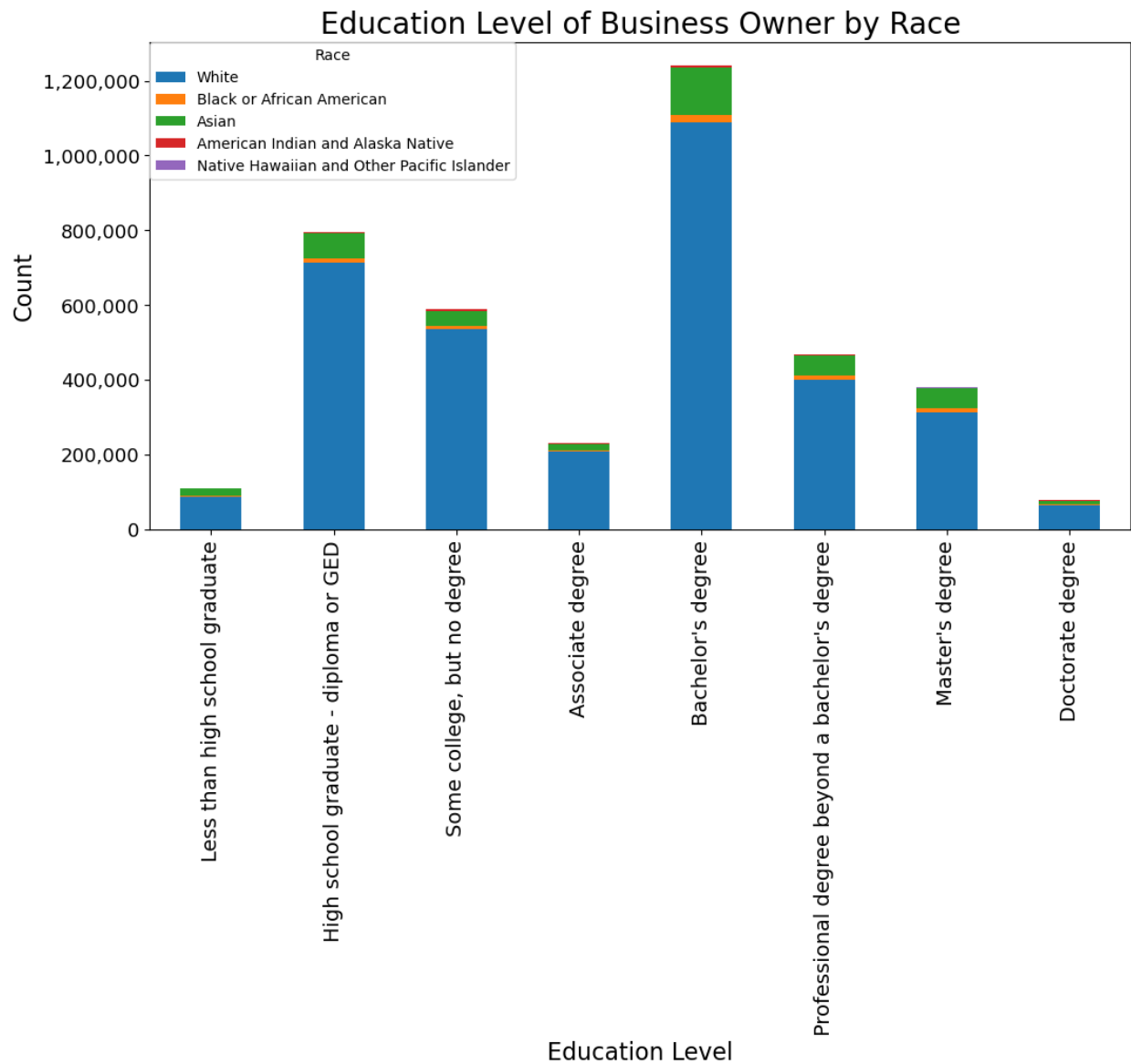


The age of most owners was over the age of 45, with most business owners being 55-64 years old.

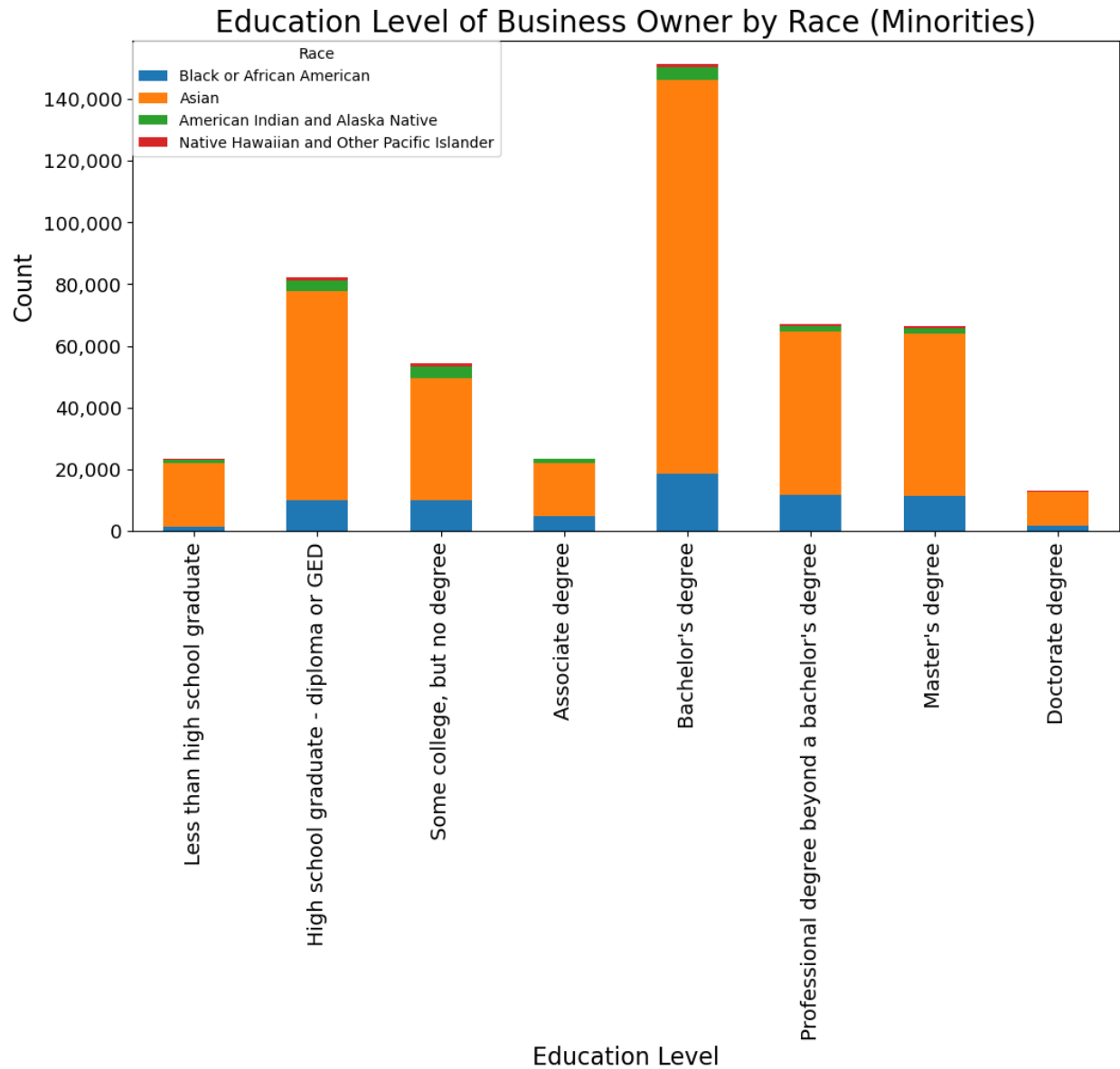


But minority business owners tended to be younger, with most falling between 35-64, and the highest proportion being 45-54 years old.

## Education Level of Business Owner



The vast majority of business owners have a bachelor's degree. More business owners have less education than a bachelor's than have more.

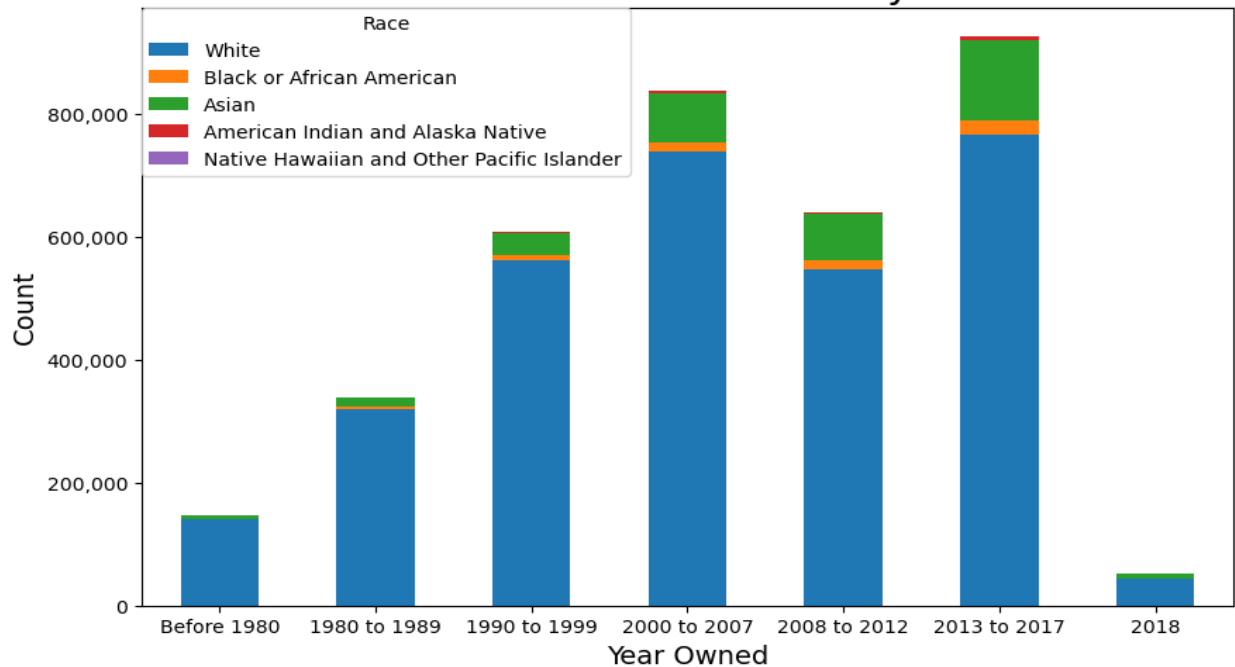


With minorities, it seems the opposite. While most business owners have a bachelor's, more business owners have education beyond a bachelor's than have education less than a bachelor's. Minority business owners are generally more educated.



## Year the Business Owner started Owning the Business

### Year Owned the Business by Race



A large portion of business owners received control of a business between 2000-2007 and 2013-2017. There is a dip in 2008-2012, likely due to the 2008 recession. Not many people would want to (or even could) start a business during economic turmoil.

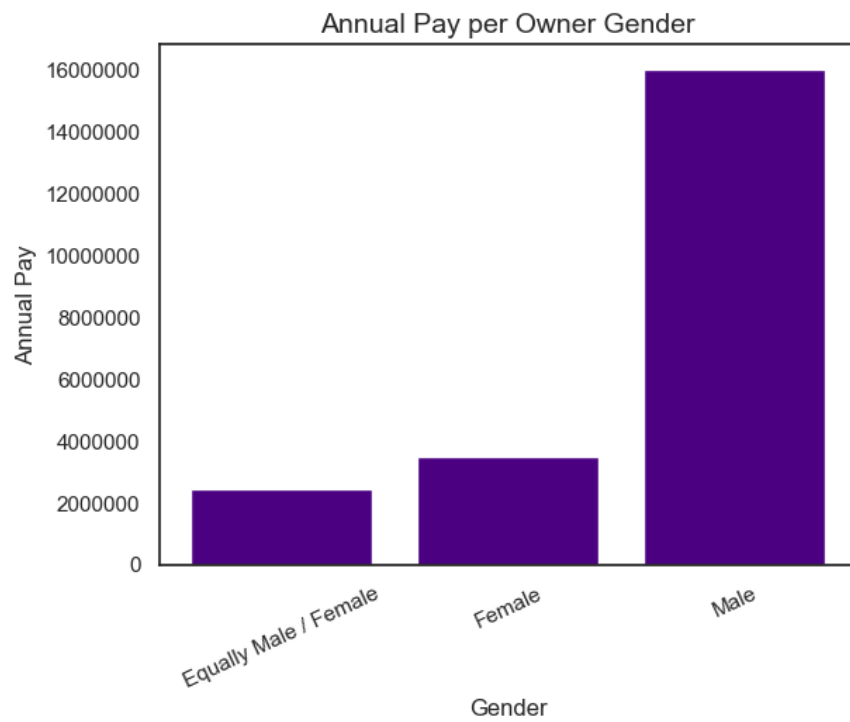
### Year Owned the Business by Race (Minorities)



There's a similar trend in current minority business owners with most minorities opening their businesses in 2000-2007 and 2013-2017, with a slight dip from 2008-2012, but not as steep as the dip from all businesses. Perhaps minority business owners from out of the country saw an opportunity to start a new business when there was less competition.

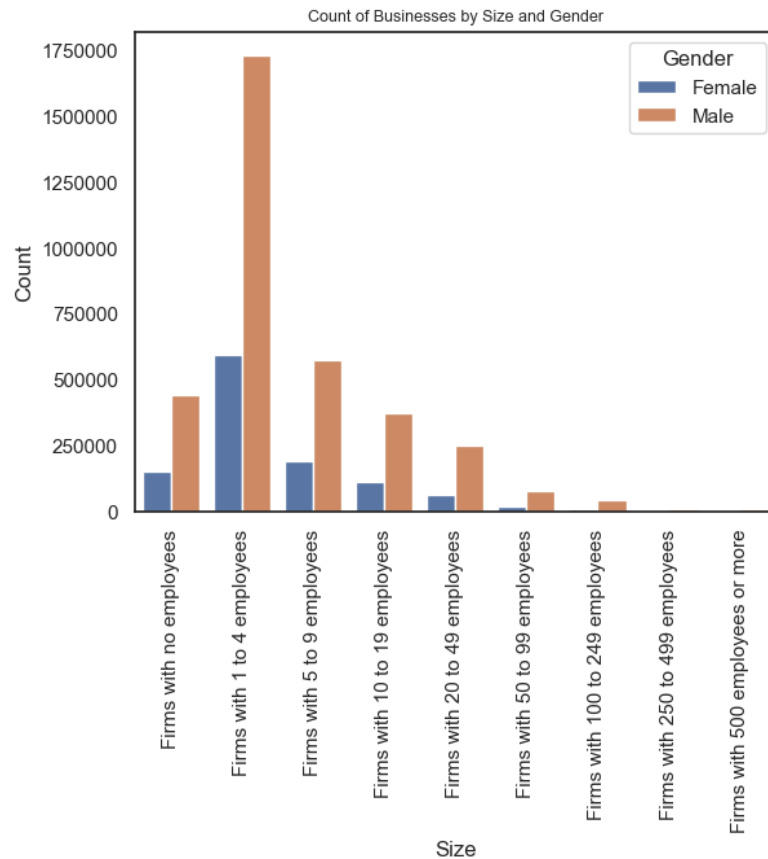
**What are the characteristics of business owners based on gender?**

**Annual Pay per Owner Gender**



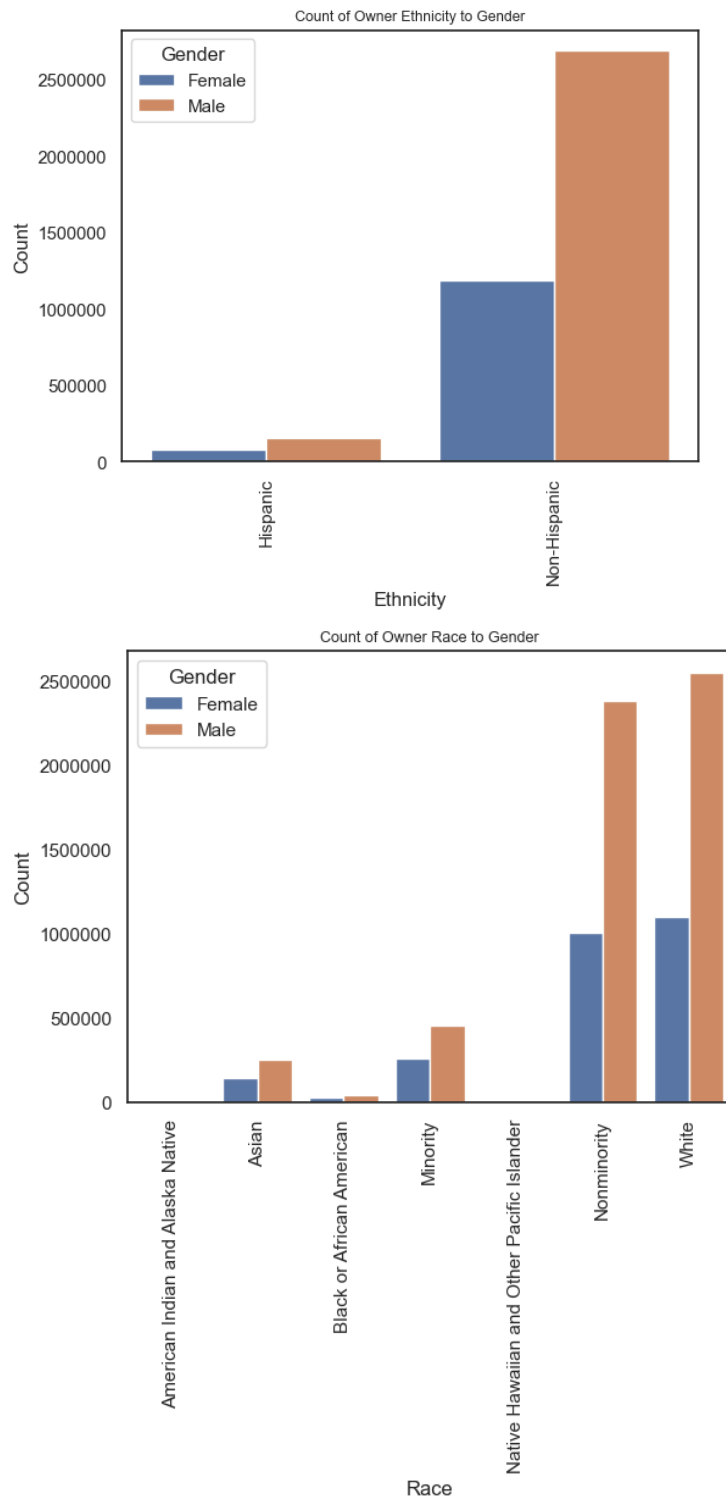
This visualization shows the average annual pay for employees of businesses that have male owners, female owners or equal male / female owners. Note that the Average Pay is the average amount the average business owner of that sex pays out total, not per employee.

## Count of Businesses according to Size and Gender



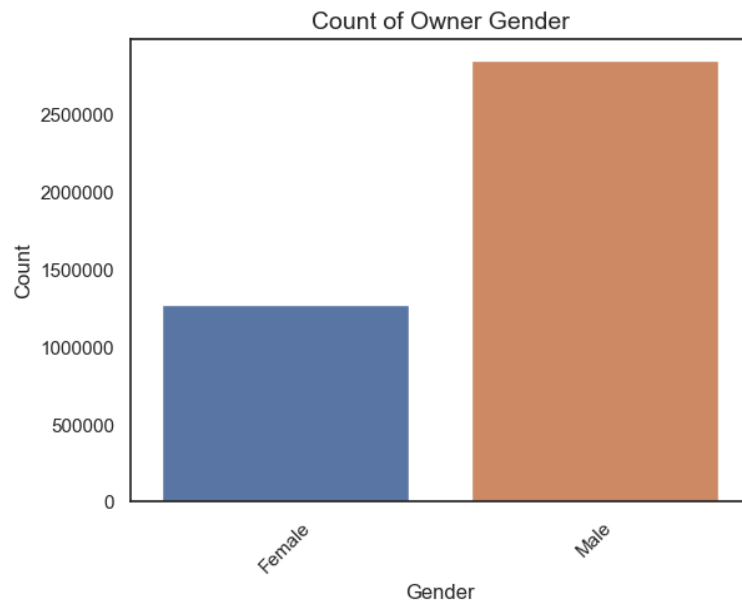
This visualization shows the count of employees in businesses by the gender of the owner. Blue represents a business where the owners are equally male / female. The orange bars represent businesses run by females and the green bars represent businesses run by males. There is a slight right skewed bell curve in the data. Firms with 1 to 4 employees have the greatest number of employees and then gradually fall off as the number of employees increase.

## Count of Race and Ethnicity of Business Owners by Gender



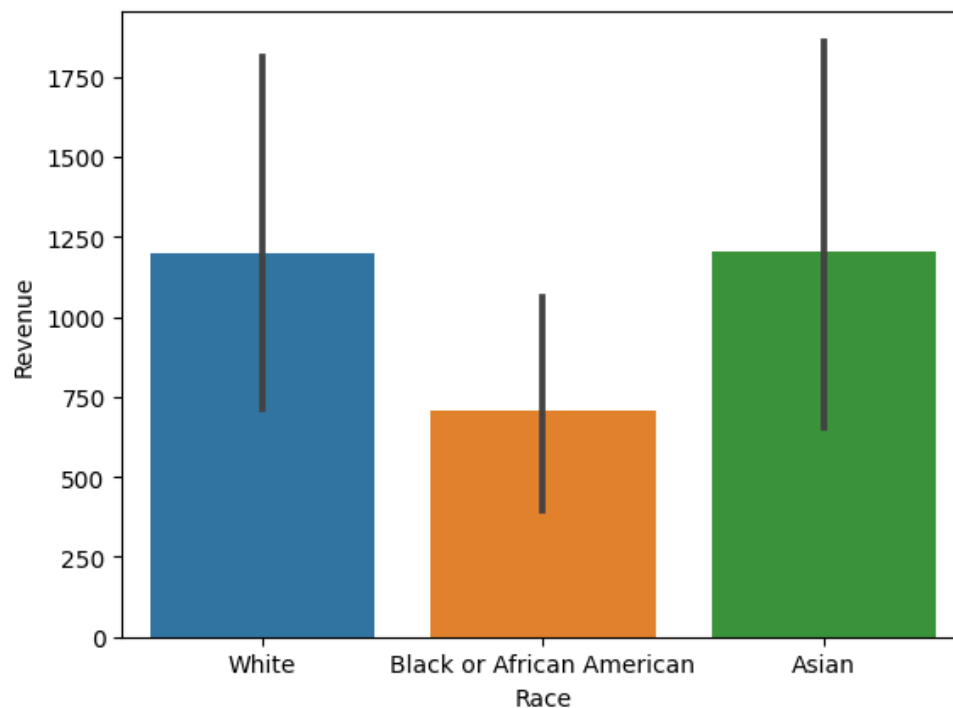
This visualization shows the count of business owners by race and grouped by gender. As shown in the graph, the largest demographic of owners is white males. The largest demographic of female owners is also white.

## Count of Business Owner Gender



This visualization shows the count of female and male owners as gathered by the survey.

## How does owner race affect revenue?

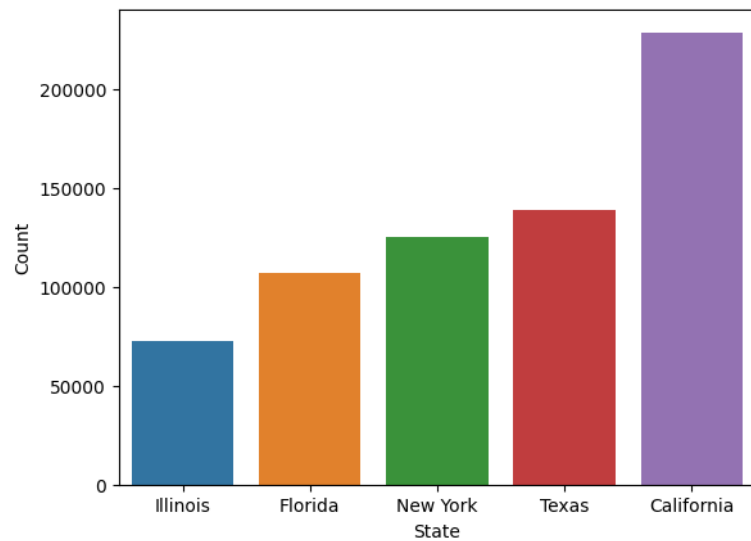
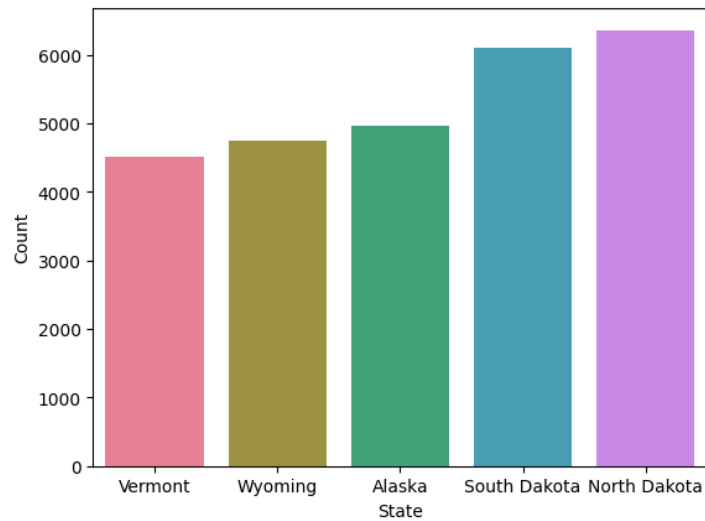


This visual is displaying the average revenue per business owner, measured by sales (RCPPEMP) in thousands of dollars, accrued by each major race group.

Most business owners are white, but they don't necessarily make more revenue based on being the majority. The average Asian business owner is bringing in similar amounts and the average

African American business owner is bringing in a bit over half the amount as an average White business owner.

**What states have the most and least high earning firms?**



The visual above is displaying the top 5 states with the highest value sales and the bottom 5 states with the least amount of high value sales.