Investigating the Correlation of Economic and Demographic Characteristics for Businesses

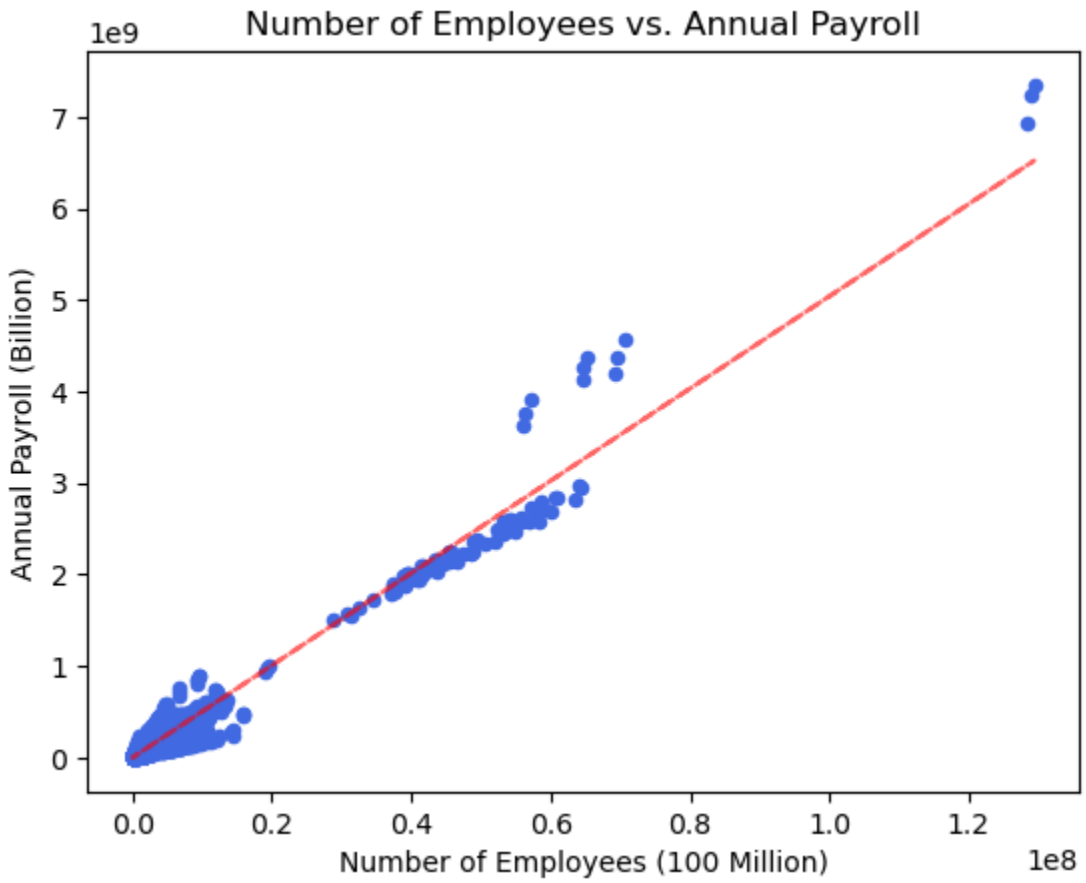
Group Members:

Miko Le
Islam Orabi
Sam Moe

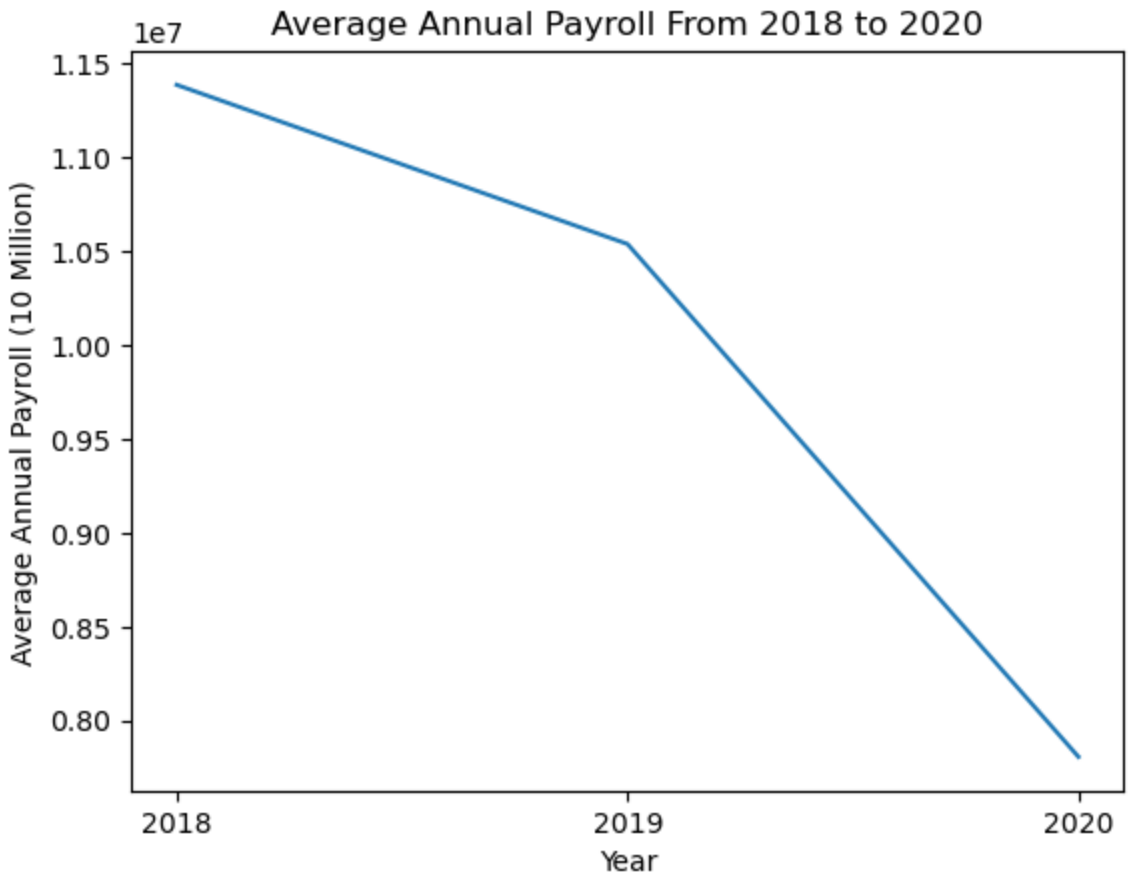
Introduction

Factors such as gender and race/ethnicity have meaningful correlations with economic characteristics in businesses in America. From 2018 to 2020, trends in demographic breakdowns were observed and the economic characteristics varied from business to business. The Annual Business Survey conducted by the US Census Bureau looks at the demographic characteristics for businesses and business owners by gender, race/ethnicity and veteran status. To investigate the impact of demographics on economic characteristics for businesses, the API URL paths for 2018, 2019 and 2020 datasets were retrieved from the Annual Business Survey (ABS) APIs webpage on the US Census Bureau website. Using the API URL paths, GET requests were performed and transformed into JSON objects. From the JSON objects, DataFrames were created for each year's datasets. The three DataFrames were concatenated into one DataFrame and transformed into a final DataFrame that was used for data analysis and visualizations. Using Matplotlib, Seaborn and Plotly, charts were created to observe trends in the data.

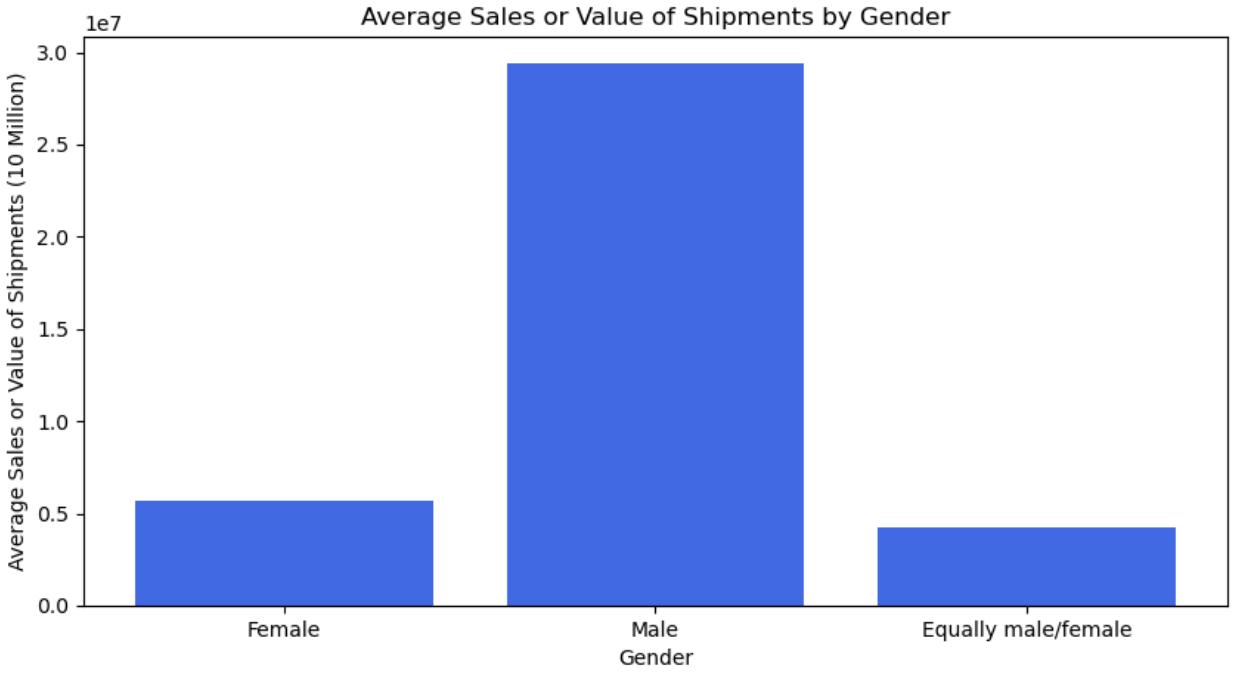
Analysis



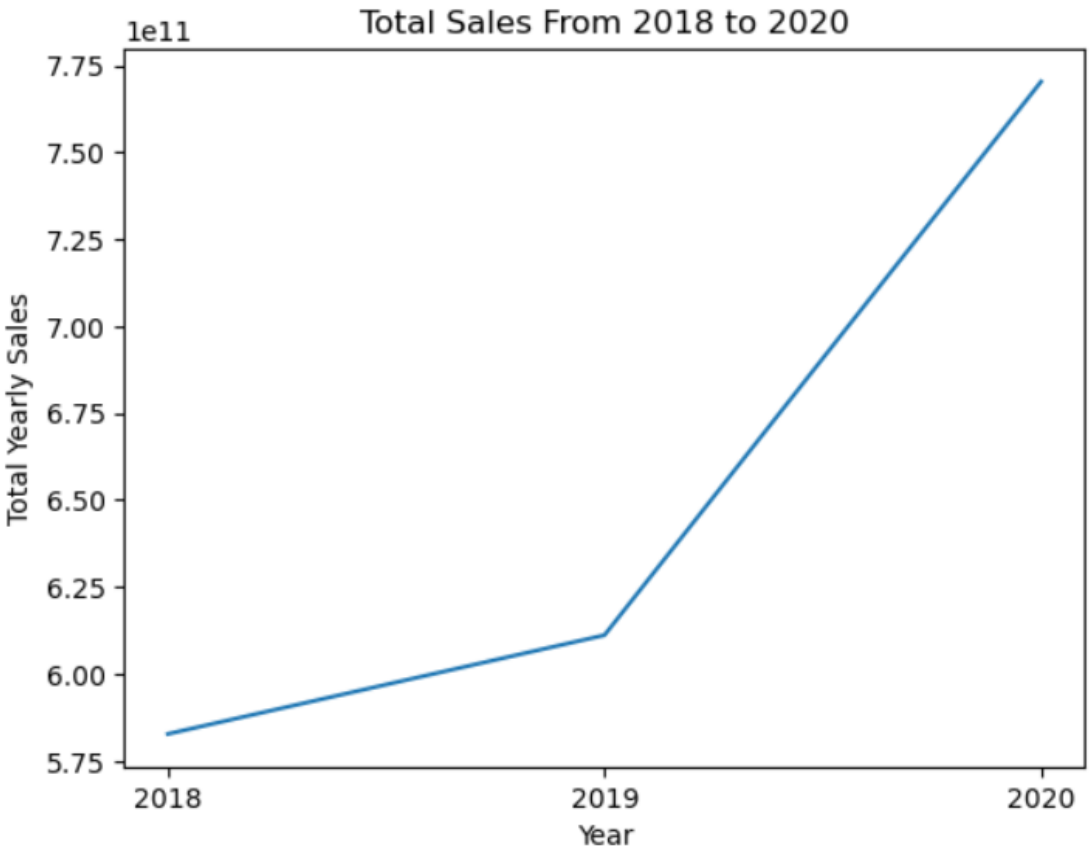
Visualization 1: The scatter chart represents the correlation between the number of employees and the annual payroll in businesses. A trendline was added to show the direction and best fit of the two variables. There is a positive correlation between the number of employees and the annual payroll. As the number of employees increases, the annual payroll increases.



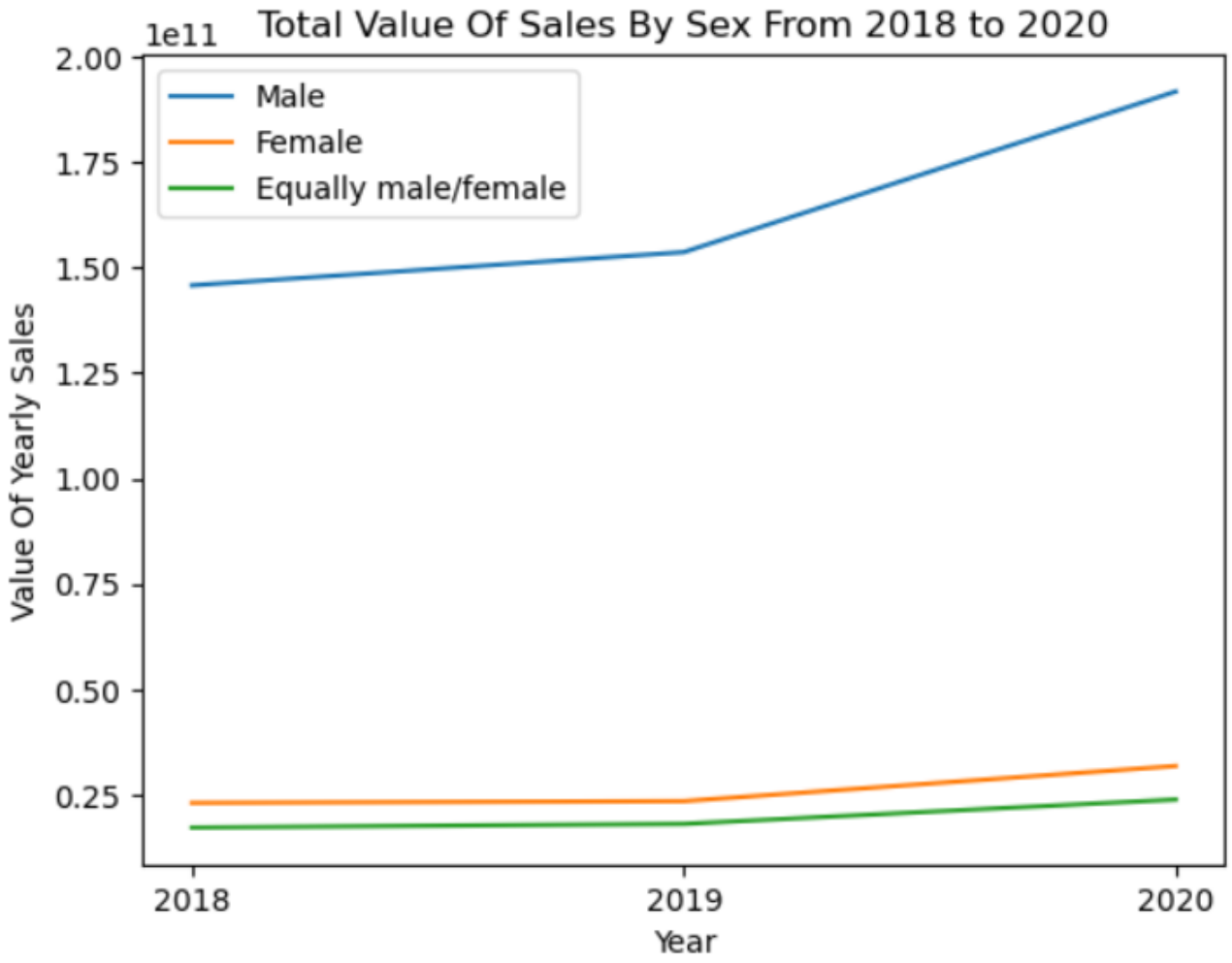
Visualization 2: The line chart represents the average annual payroll from 2018 to 2020. Over the years, the average annual payroll decreased. From 2018 to 2019, there was a decline in the average annual payroll. After 2019, the average annual payroll rapidly dropped below 10 million dollars.



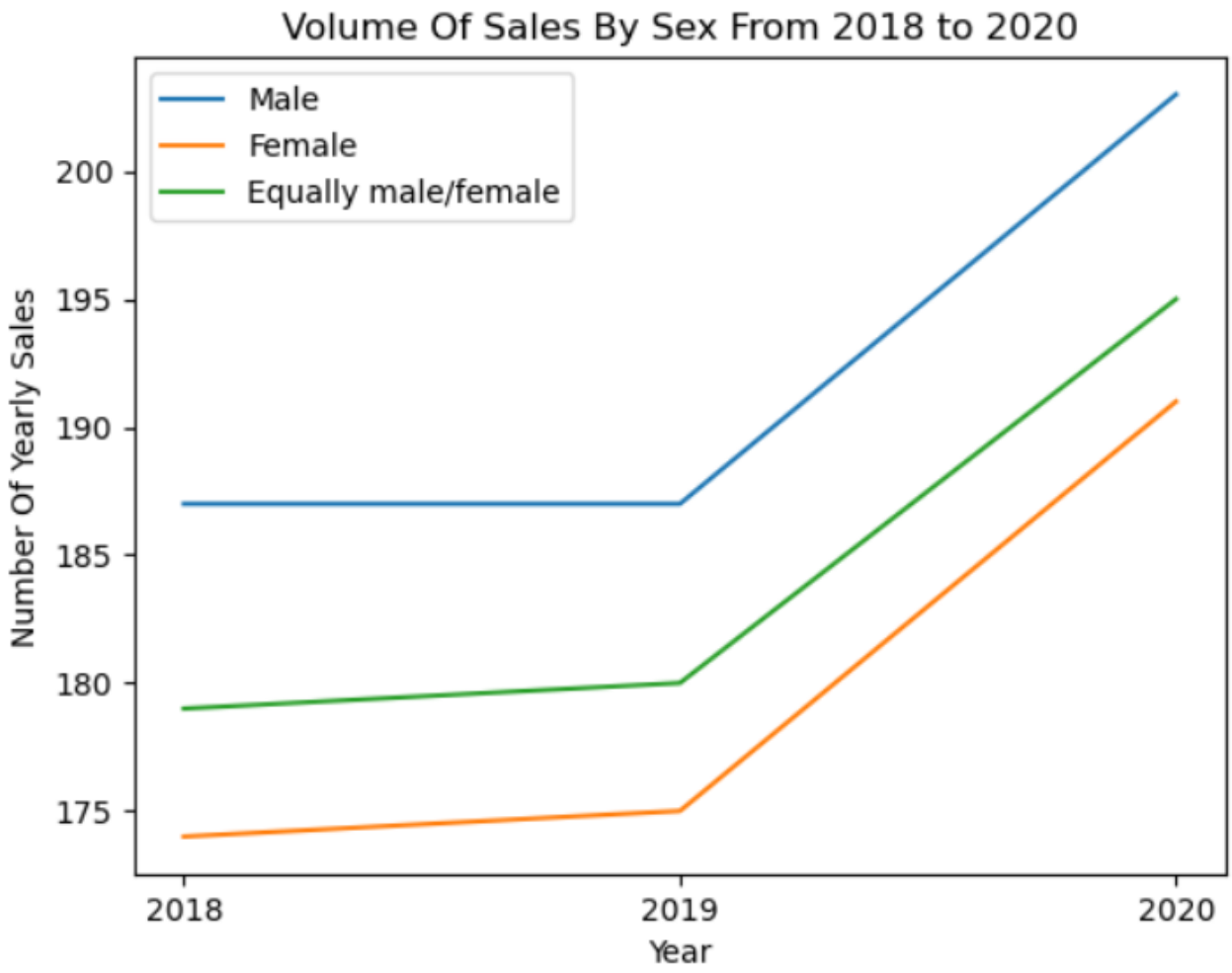
Visualization 3: The bar chart represents the average sales or value of shipments by gender. There is a gender discrepancy between males and females such that male businesses have larger sales or value of shipments than female businesses. Businesses of equal males and females have the lowest sales or value of shipments than male and female businesses.



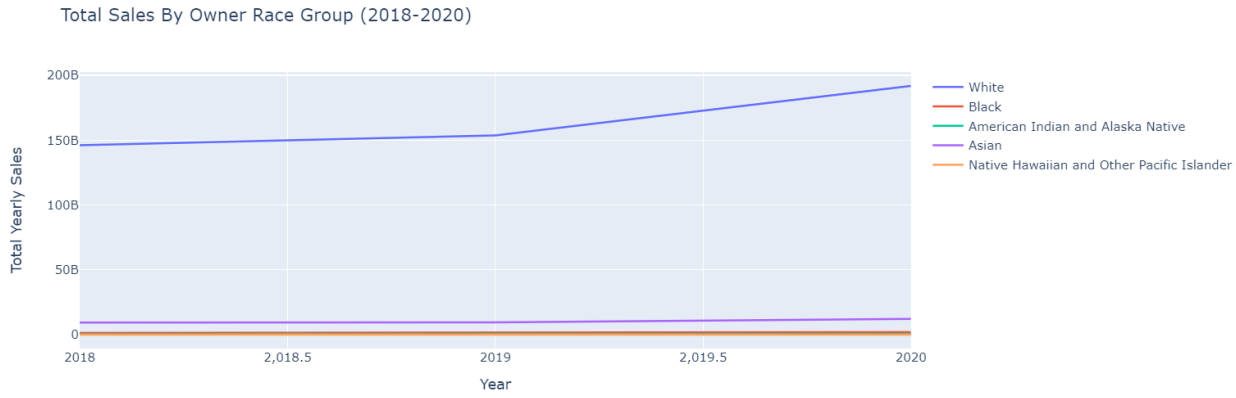
Visualization 4: The line chart shows the total value of all sales for the years 2018 through 2020. There is an increase in sales from 2018 to 2019, followed by a drastic acceleration in sales from 2019 to 2020. This is contrary to the decrease in average payroll within the same timeframe, meaning the average amount of payroll per sale has fallen drastically in recent years. (Matplotlib, 2023)



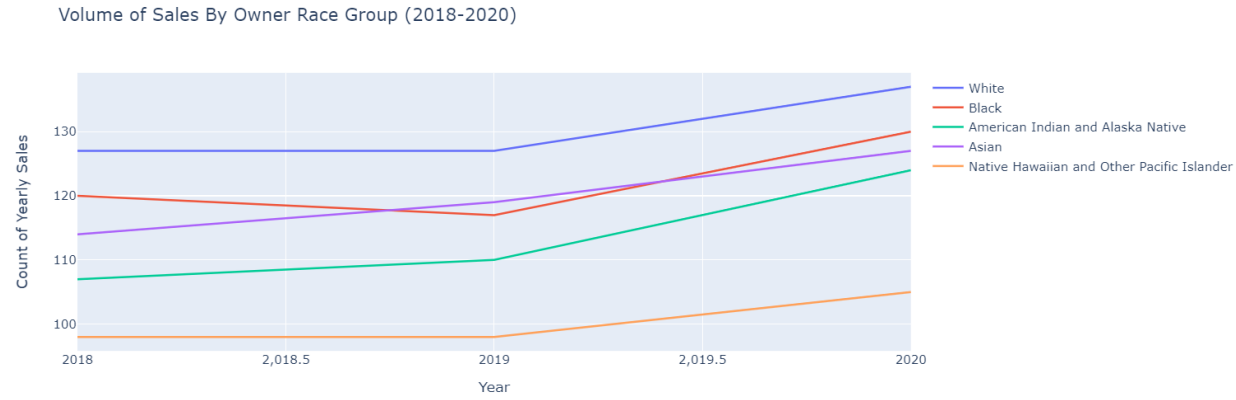
Visualization 5.1: This line graph shows the total value of sales broken down by the sex of majority owners of businesses. We can see the total value of sales was highest in companies with a majority male ownership, followed by companies with female ownership, and companies with mixed gender ownership had the lowest sales for each year. We also note that while all groups experienced an increase in total sales from 2019 to 2020, the increase was much higher among companies with male ownership. (Matplotlib, 2023)



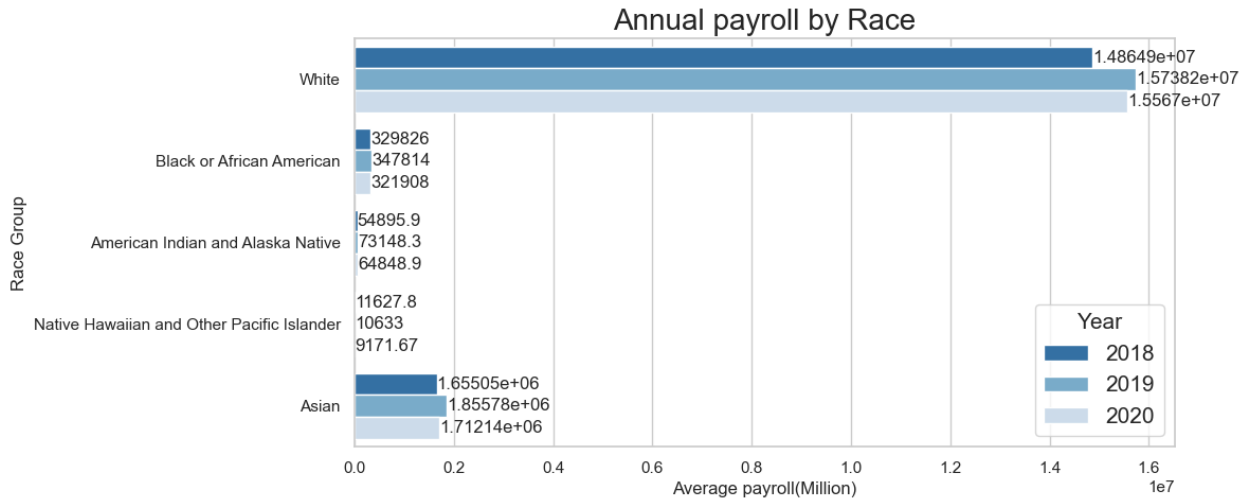
Visualization 5.2: While there was a large discrepancy between male, female, and equally owned business when it comes to total value of sales, this discrepancy is less pronounced when looking at the number of sales per year. Male owned companies still held the largest share of sales each year, but this difference was much smaller, indicating that male owned businesses must have a higher average value per sale. (Matplotlib, 2023)



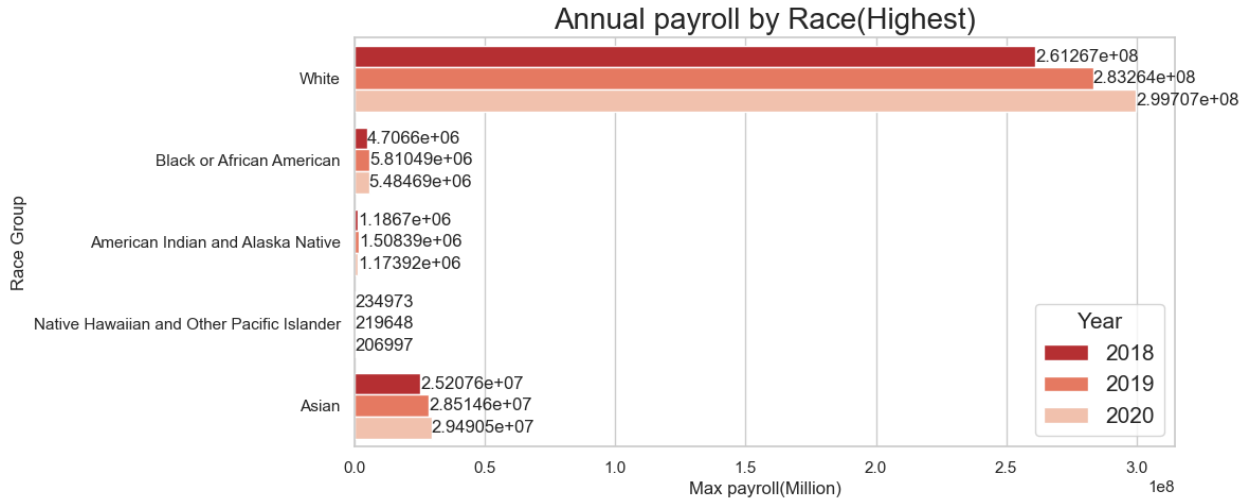
Visualization 6.1: This line graph compares total yearly sales of businesses broken down by race group of owners. Here we can see that there is a large discrepancy between white owned businesses and businesses primarily owned by other races. (Plotly, 2023)



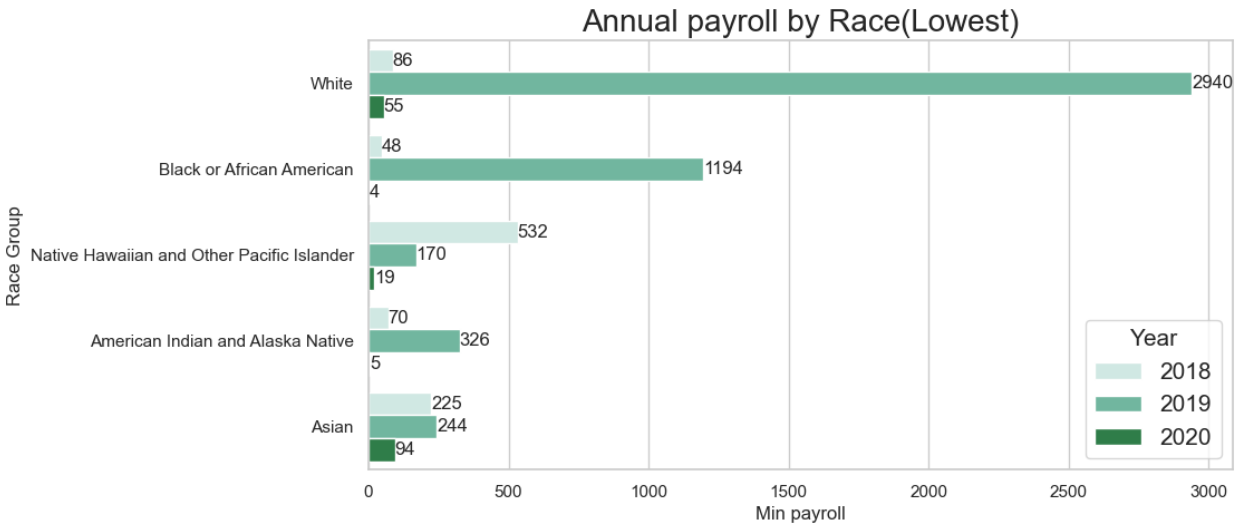
Visualization 7.1: To gain a better understanding of the relationship between race and payroll, we created a chart to show the average payroll for each race. By looking at the chart, we can see that White is leading by a landslide, following that we have Asian, African American, American Indian, and lastly is Native hawaiian/Pacific Islander.



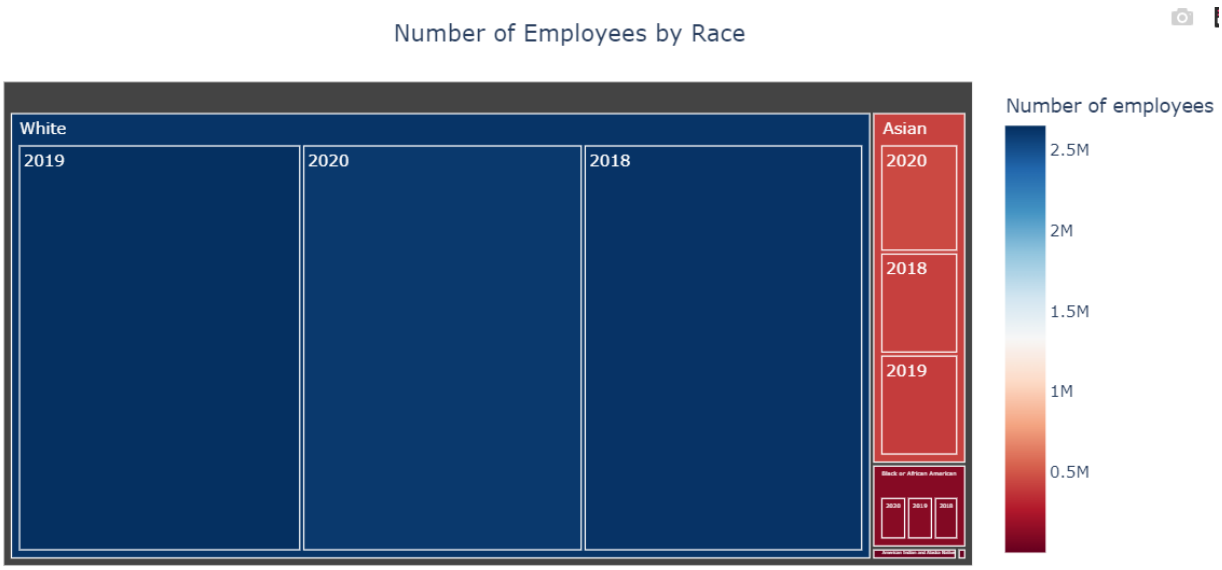
Visualization 7.2: We wanted to see how big of a gap there is in payroll between different races by comparing the highest earnings for each race. This yielded similar results to average payroll, in first place is White with the highest payroll in 2020(near 3Million); and Native Hawaiian/other Pacific Islander in last with around 200k.



Visualization 7.3: The bar chart below displays the lowest income for each race for all three years. Compared to the highest payroll for each category, there is a drastically different payroll. One notable observation would be payroll during the year 2020, when the Covid-19 pandemic occurred. With the pandemic, many lost their jobs and became financially unstable which could explain the small payroll amounts.



Visualization 8: The following treemap exhibits information regarding the diversity of employees for the year 2018, 2019, and 2020. As can be seen below, White employees are the vast majority then Asian, then African American, and there a small amount of American Indian & Native Hawaiian employees. These numbers seem to be static through the years, not too much changes in terms of numbers of employees; this can be applied to all races.



Conclusion

Through our analysis we were able to identify trends and answer several questions. We found that discrepancies do exist in payroll, sales and numbers of employees among U.S businesses when it comes to sex and racial background of their owners. White owned and male owned businesses account for the vast majority in sale value, yet they do not make up the overwhelming majority of businesses, nor do they make up the majority of sales in terms of volume. This may be an issue that needs further examination as to how and why these discrepancies exist. Additionally, our analysis found trends that seem to be consistent with the covid-19 pandemic that began in 2019, as we saw a large decrease in average payroll during this period.