Pandalytics Project Report

Members:

Shane Sarabdial, Edward Oh, Noelle Kiesz, Tristen Brewer

Table of Contents

Introduction

This project was designed to allow us to demonstrate our abilities with APIs, python programming, cleaning and transforming data with Pandas, using GitHub, and creating visualizations with Matplotlib, Seaborn, and Plotly. We are tasked with the following :

1. Perform an ETL from the US Census Bureau.
2. Create visualization with the plotting libraries.

Data Source

Annual Business Survey (ABS)

The ABS provides information on selected economic and demographic characteristics for businesses and their owners by sex, ethnicity, race, and veteran status. There are four datasets in the ABS. We choose to work with data from the years 2020 and 2019. The data source can be found [here](#).

1. ### Company Summary (CS)
   The CS dataset adds years in business, firm size, state data and payroll information, and urban and rural classification to the dataset. You can read the technical information about the dataset [here](#).

2. ### Characteristics of Businesses (CB)
   The CB dataset provides information on the number of owners, type of business, work-from-home status, types of customers, and many more categories. You can read the technical information about the dataset [here](#).

3. ### Characteristics of Business Owners (CBO)
   The CBO provides information on citizenship status, age group, the reason for owning a business, education, and many more categories. You can read the technical information about the dataset [here](#).

4. ### Module Business Characteristics (MCB)
   The MCB adds technology use, management practices, and financing characteristics of the business. You can read all about the MCB dataset [here](#).

Initial Questions

After taking a broad look at the datasets, we have a few initial questions we want to explore more.

1. Has the number of male-owned vs. female-owned businesses changed from 2019 to 2020?
2. Has the payroll changed between males and females from 2019 to 2020?

3. What industries support working from home?
4. What are the top reasons for owning a business by race and gender?
5. During the pandemic, who received the most assistance by gender and race?

Analysis

Company Summary Analysis

Compared to 2019, there was a slight percentage increase in female-owned businesses (21.9% vs. 21.7). The percentage of male-owned companies did not change and remained the same at 64%.

Female-owned businesses increased from 1,217,339 to 1,223,701, while male-owned businesses decreased from 3,584,220 to 3,547,204. Equally, male and female-owned businesses also decreased slightly from 828,967 to 812,034.
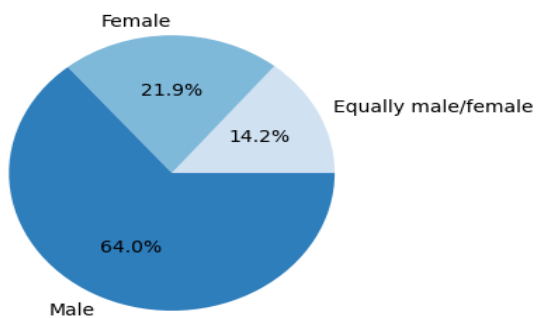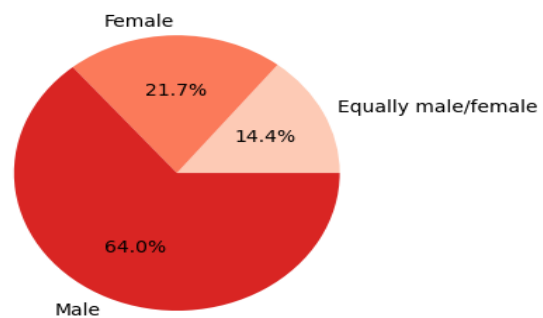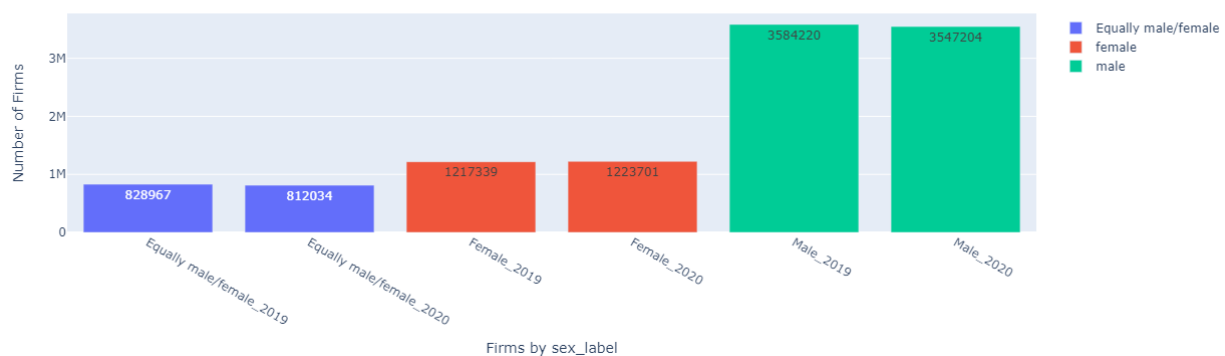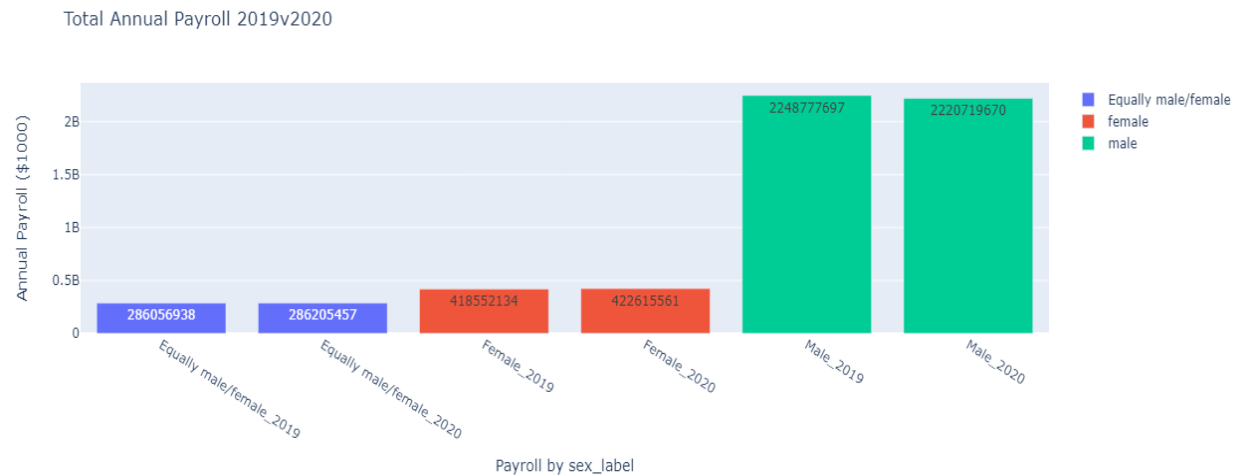
In terms of payroll, the female-owned business had a slight increase in total payroll while the male-owned business dropped.

Total Annual Payroll 2019v2020

By state, Hawaii and Wyoming had the most significant percentage drop in total annual payroll (11.48% and 11.25%, respectively), while Delaware and South Dakota had the highest percentage increase (9.74% and 8.59%, respectively).

Characteristics of Businesses Analysis

Examining the data from 2020 on the number of employees working from home in each sector, not much can be seen in the first heatmap. However, it was used to see if any interesting sectors could be looked at more closely.

Number of Workers Working from home by Sector
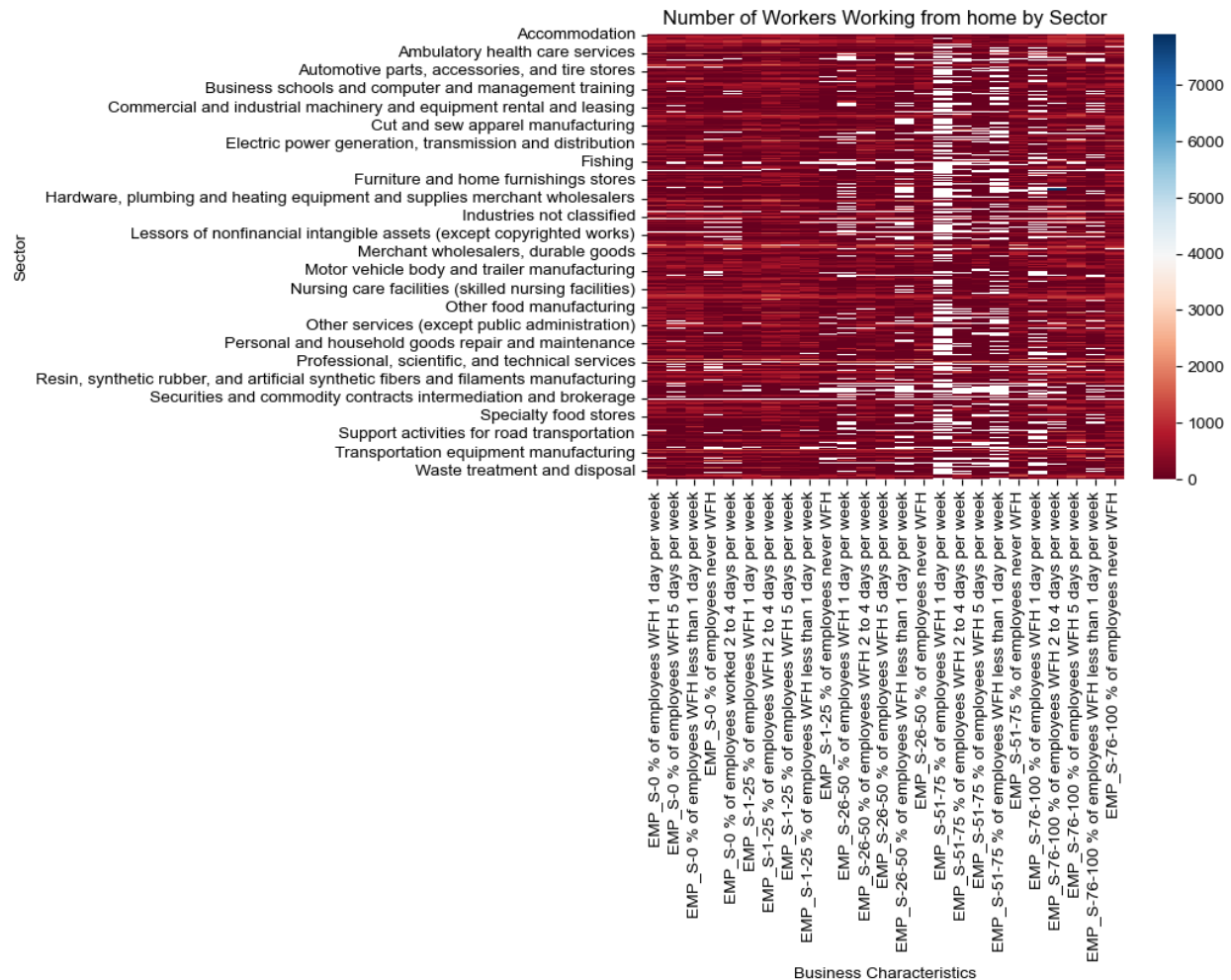
The only blue area was an area of interest in which the 'Furniture and furnishings stores' had the highest number of employees working from home two to 4 days a week. The other sectors of interest were: 'Ambulatory health care services,' 'Merchant wholesalers, durable goods,' and 'Professional, scientific, and technical services.' Those sectors were used for the second heatmap to zoom into the data to see if there were more interesting findings.

This heatmap shows that the most common answer to the percentage of employees and time spent working at home was '51-75% of employees WFH 1 day a week' across all of the four sectors that were looked at.

To delve into the four sectors more, gender was split out to examine the differences between the amount of time males vs. females spend working from home. A filter was also used to look

at the answers for working from home five days a week. Overall, looking at the graph below, males typically spent more time working from home when compared to females. The two exceptions are 26-50% of females working from home five days a week in the Ambulatory health care services and Merchant wholesalers. Interestingly, where Furniture and home furnishing stores originally had the highest number of employees working from home do not have employees consistently working from home five days a week.

Number of Workers Working from home by Sector
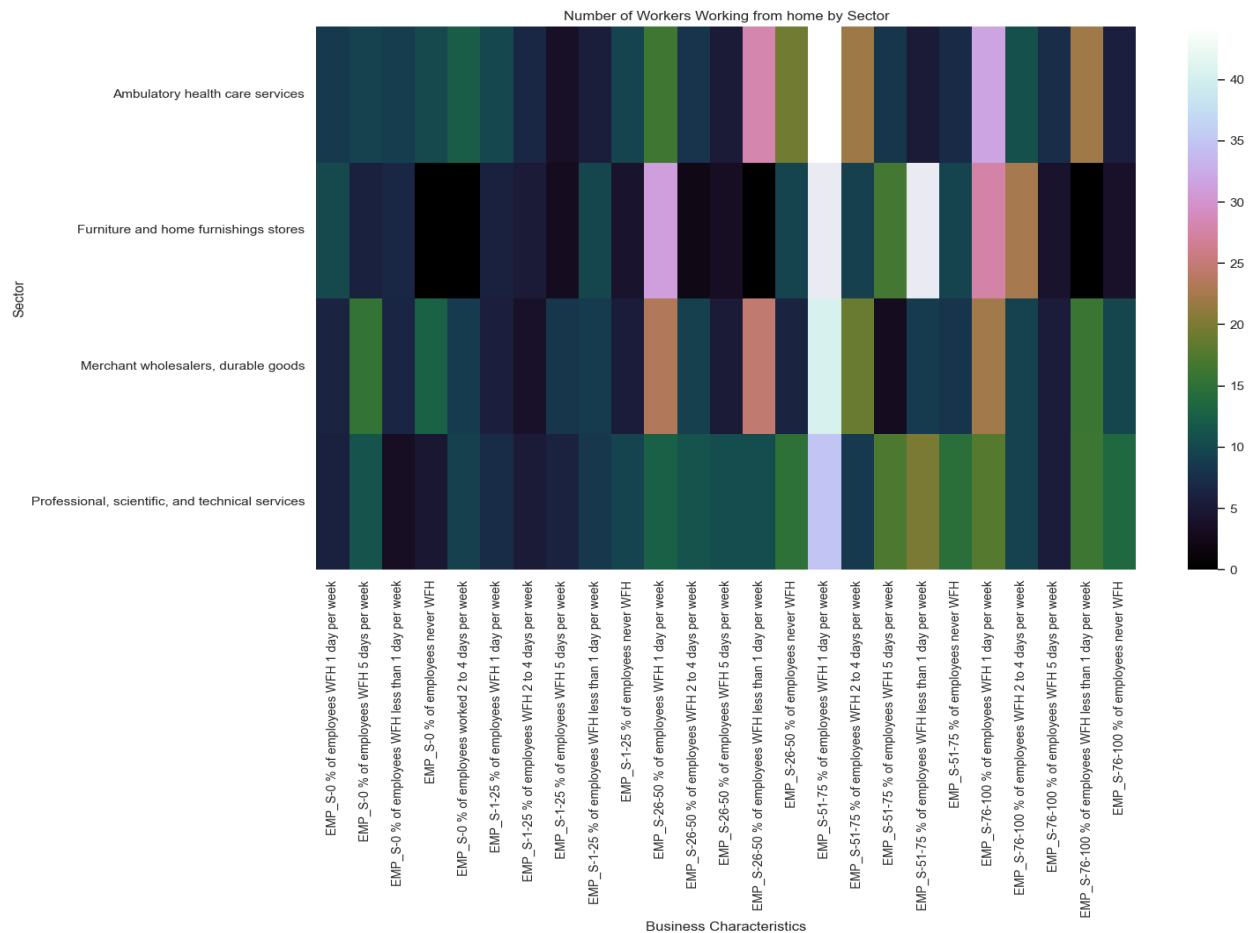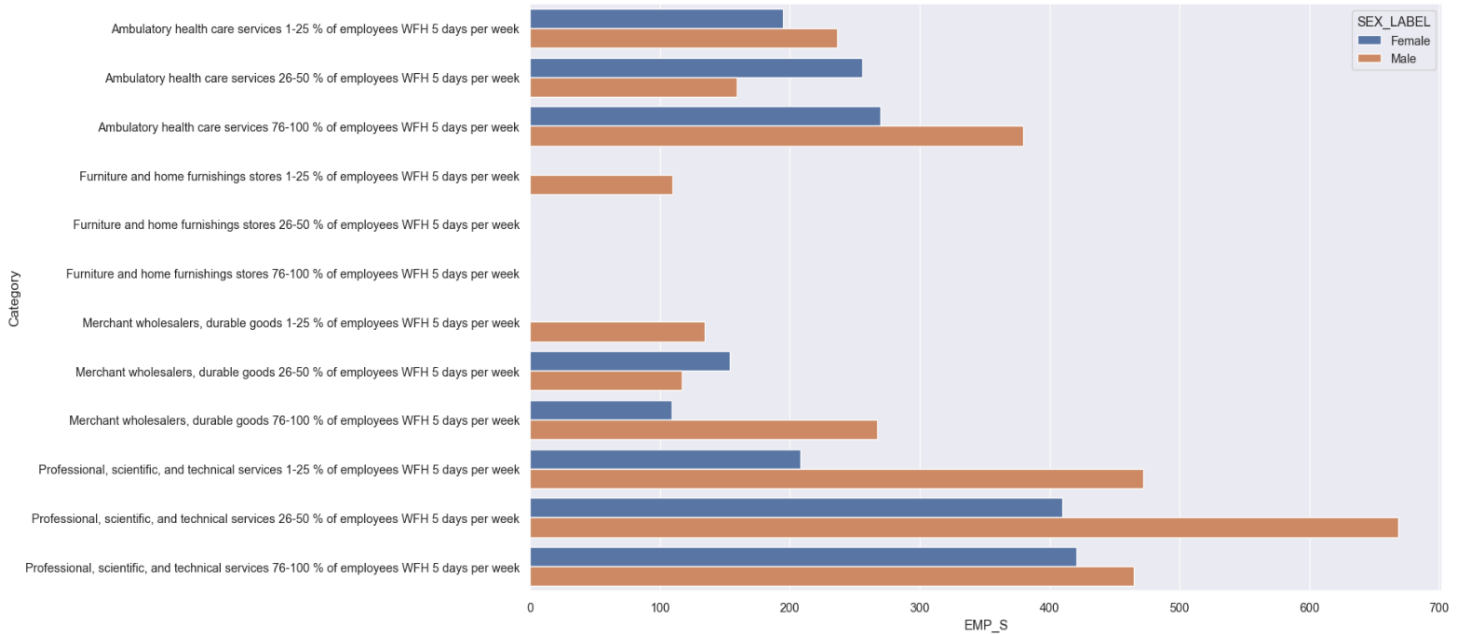
The last graph examines the number of male and female employees responding to all questions. Overall, the males were higher in all categories. This could be due to the sectors that were chosen. They could have a higher proportion of males who work in the industry.
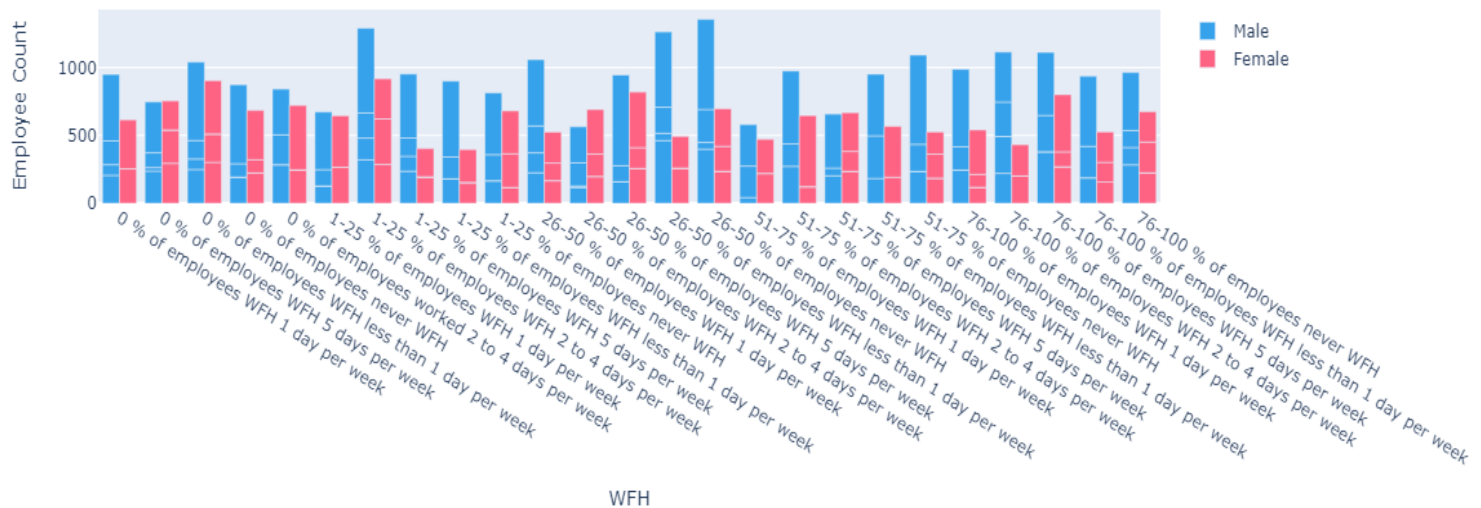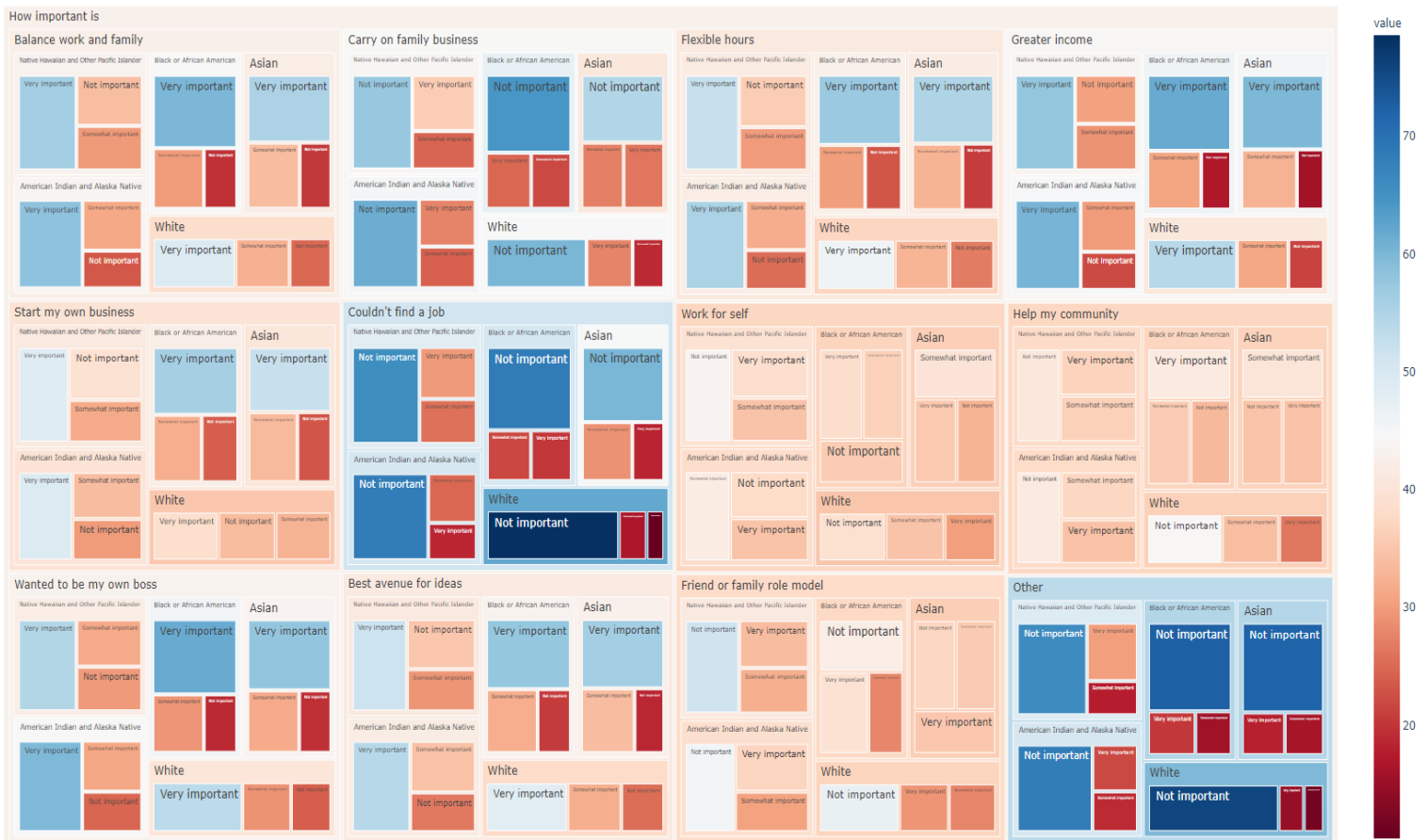
Characteristics of Business Owners Analysis

The CBO survey asked owners, "How important to Owner 1 (2, 3, or 4) are each of the following reasons for owning this business? Select one for each row. " [Not important, Somewhat important, or Very important]. We wanted to know how each race responded.


Most Important Reason to Own A Business

There was a lot of data, and we found that a tree map was the best way to get a broad overview of how people responded.

In the tree map above, we see a few trends.

1. There is an equal distribution of answers in the "Work for self" and "Help my community" questions.
2. Most races responded "not important" to the "Couldn't find a job," with whites overwhelmingly agreeing.
3. Generally, most people responded, "Very important" to "Wanted to be my own boss," "balance of work and family," and "Greater Income."

Most Important to Own a Business by Gender
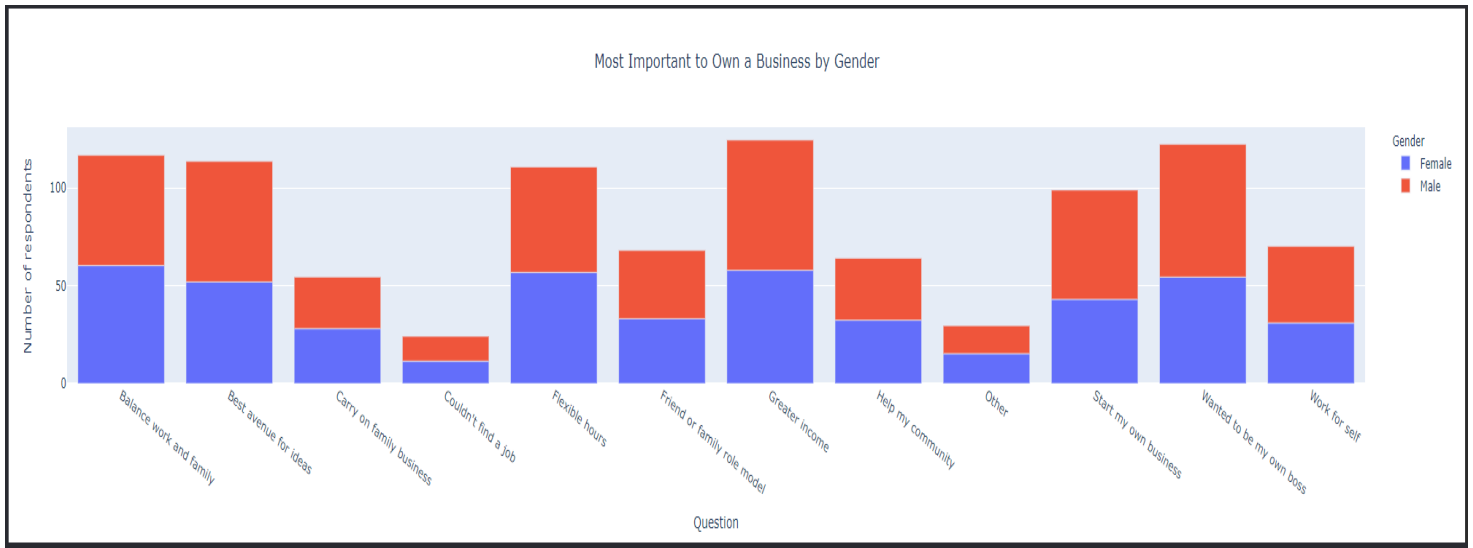
In the following graph, we wanted to look for similar trends but with gender instead.

In this bar graph, we filtered the data to look and gender and where respondents chose "very important" to the 12 questions asked. We see that males and females find a balance of work and family, helping their community, and flexible hours almost equally important. Males will generally find it more important to be their own boss, start their own business and have a greater income than females.

In our last chart, we zeroed in on owners who wanted to be their own bosses and find out what industry they were likely to work in. The graph below shows the top 10 industries of those with this characteristic.

Wanted to be my own boss: Very important
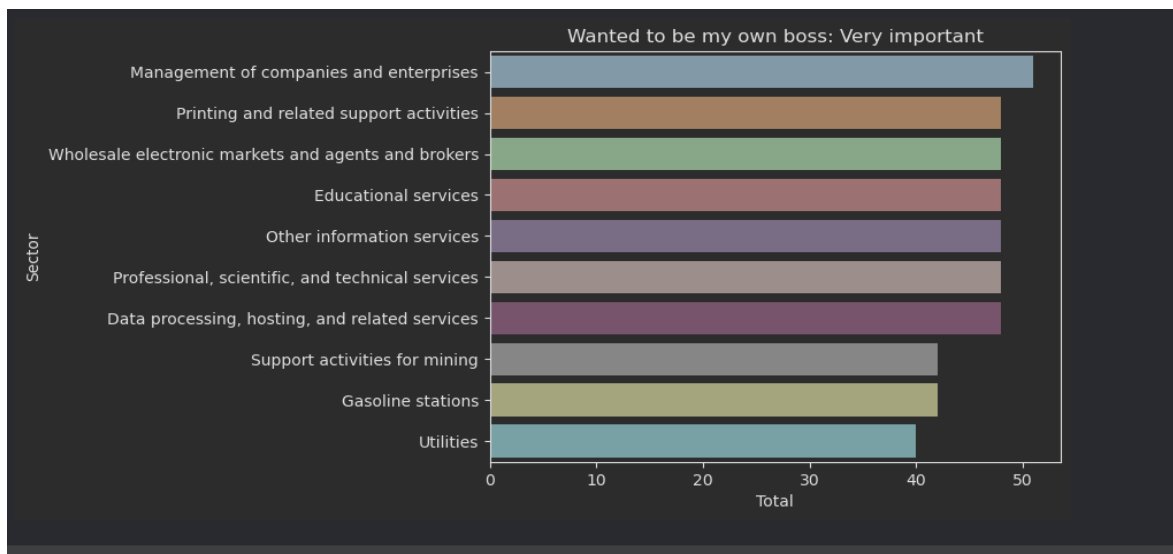
As expected, people who find it very important to be their own boss work in management and brokering. Interestingly, information services, data processing, education, and printing services are so high on the list.

Module Business Characteristics

In the project, we wanted to see how government assistance was given out to businesses in the US and how it differed for companies with owners of different races and sex.

For the first plot, we look at the male, female, and equally male and female-owned businesses. In the graph, we see a lot of male-owned businesses that received much more assistance during the pandemic.

Less than half of the female-owned businesses received assistance in every money group except in the $5k-less than $10k. Still, even in that group, female-owned businesses had approximately 25k fewer businesses receiving aid.
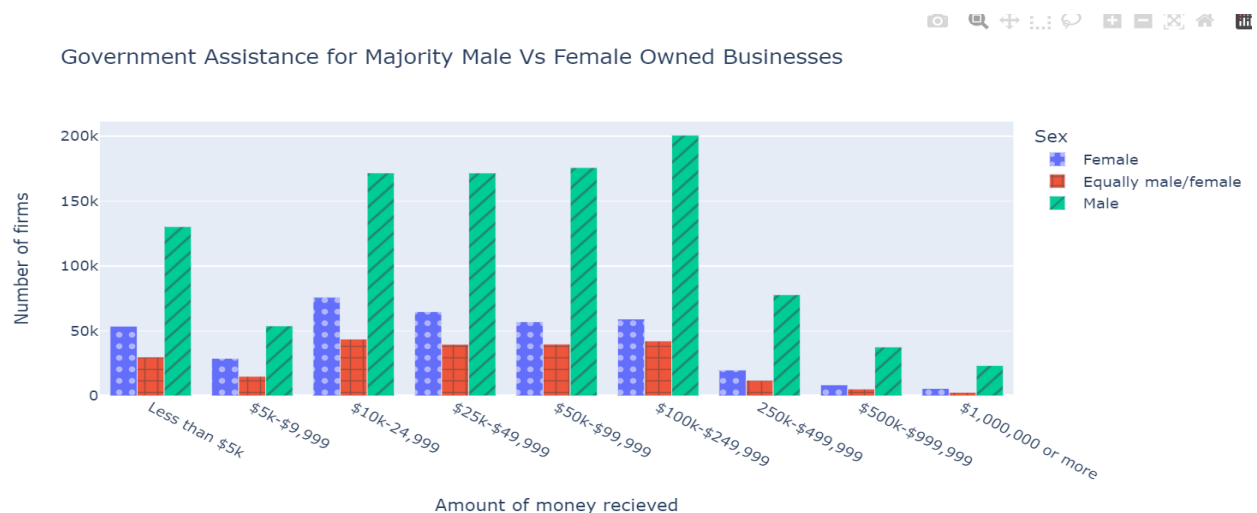
We also look at the various races represented in the data. We again see white owners received much more assistance than any other race. More than double the amount for every money group. From less than 5k. to a million or more. The second most are Asian owned businesses for every group.

Government Assistance by Race Group

Government Assistance for White Owned Business

To give an example. For less than $5k of assistance received, there were approximately 180k white-owned businesses, while about 26k businesses were Asian owned. There were even fewer of all the other race groups.

Also, we looked at the number of firms that requested a certain amount of government help based on their race groups.

Like before, White-owned businesses requested government assistance much more than any other race, second being Asian-owned businesses.

However, it was still not close. More white-owned businesses were requesting help than all other race groups combined over every money grouping from less than $5k to a million dollars or more.

For example, in the grouping of $25k to less than $50k, approximately 270k white-owned businesses requested this much government assistance. For all the other race groups combined, less than 50k businesses were requesting this much help.

Government Assistance Requested by race group

Government Assistance Requested for white business onwers
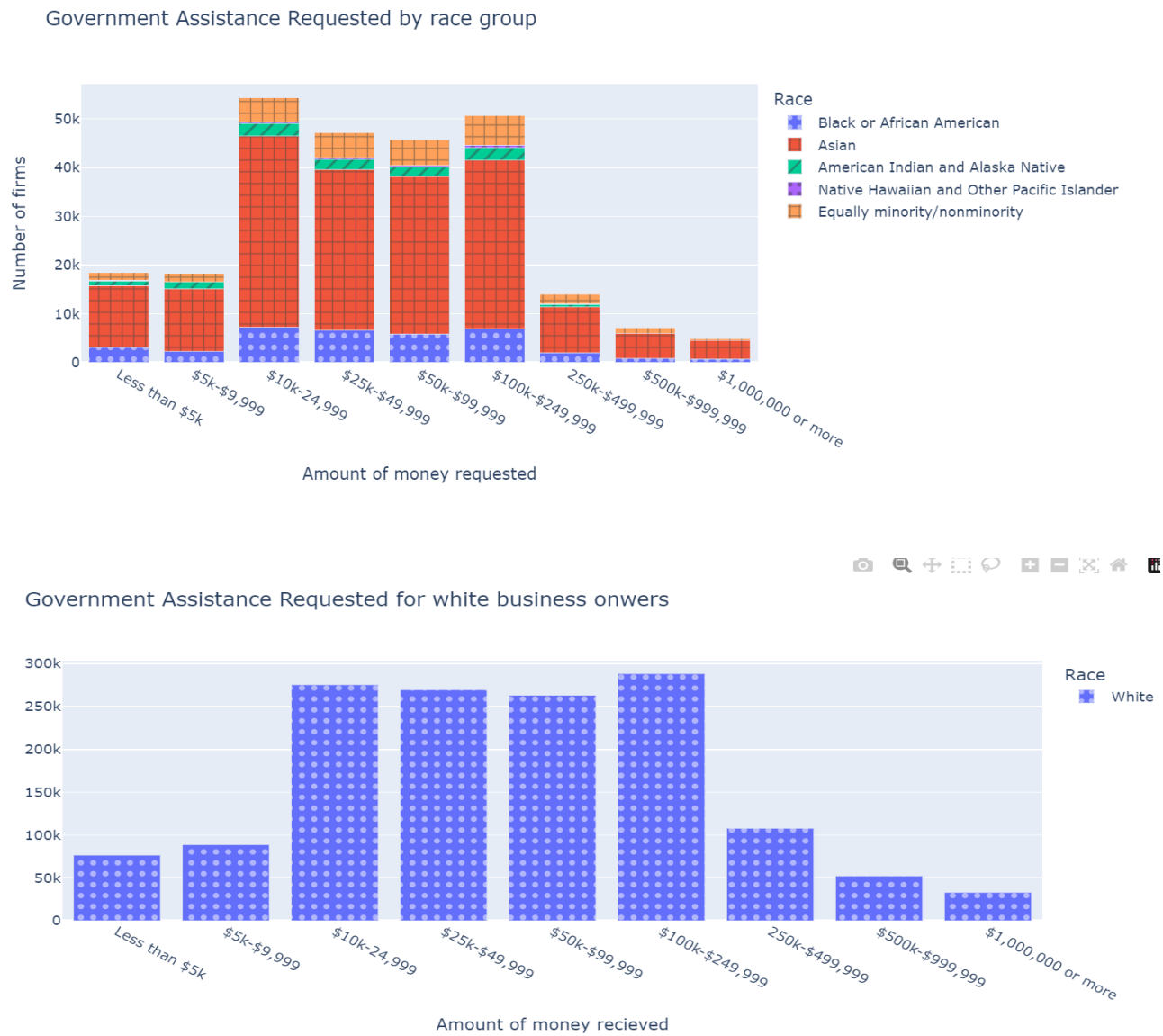
Conclusions

Our initial questions were the following

1. Has the number of male-owned vs. female-owned businesses changed from 2019 to 2020?
2. Has the payroll changed between males and females from 2019 to 2020?
3. What industries support working from home?
4. What are the top reasons for owning a business by race and gender?
5. During the pandemic, who received the most assistance by gender and race?

After conducting our analysis, we found a marginal change between the number of female-owned and male-owned businesses from 2019 to 2020 and their payroll. Males worked from home more than females. Interestingly, the furnishing industry, supported work from home more than any other industry during the pandemic. The top reason for owning a business for both gender and race is more significant income, being your boss, and flexible hours. During the pandemic, more white-owned companies received aid, but this was because most respondents were white.

The census data was massive, and finding trends and even more challenging. One thing we could do differently is to look at a longer time horizon. The year-to-year data changes marginally, making finding movements difficult. Looking at a 10 or even 20-year time horizon, we will see more exciting results.