Pandalytics Project Report

Members:

Shane Sarabdial, Edward Oh, Noelle Kiesz, Tristen Brewer

Table of Contents

[INTRODUCTION 2](#_Toc133170000)

[Data Source 2](#_Toc133170001)

[Annual Business Survey (ABS) 2](#_Toc133170002)

[1. Company Summary (CS) 2](#_Toc133170003)

[2. Characteristics of Businesses (CB) 2](#_Toc133170004)

[3. Characteristics of Business Owners (CBO) 2](#_Toc133170005)

[4. Module Business Characteristics (MCB) 2](#_Toc133170006)

[Analysis 2](#_Toc133170007)

[Company Summary Analysis 2](#_Toc133170008)

[Characteristics of Businesses Analysis 4](#_Toc133170009)

[Characteristics of Business Owners Analysis 7](#_Toc133170010)

[Module Business Characteristics 8](#_Toc133170011)

# INTRODUCTION

This project was design 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 a 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.

### 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](https://www2.census.gov/programs-surveys/abs/technical-documentation/api/API2020-company-summary.pdf).

### Characteristics of Businesses (CB)

The CB dataset provides information on 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](https://www2.census.gov/programs-surveys/abs/technical-documentation/api/ABS_API_CB.pdf).

### Characteristics of Business Owners (CBO)

The CBO provides information on citizenship status, age group, reason for owning a business, education and many more categories. You can read the technical information about the dataset [here](https://www2.census.gov/programs-surveys/abs/technical-documentation/api/ABS_API_CBO.pdf).

### 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](https://www2.census.gov/programs-surveys/abs/technical-documentation/api/ABS_API_MCB-10-18-2022.pdf).

# 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 businesses 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 went down slightly as well from 828,967 to 812,034.

Chart, pie chart

Description automatically generated

Chart, bar chart

Description automatically generated

In terms of payroll, female-owned business had a slight increase in total payroll while male-owned business has dropped.

Chart, waterfall chart

Description automatically generated

By state, Hawaii and Wyoming had the greatest 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

Looking at the data from 2020 on the number of employees working from home in each sector, not a lot can be seen in the first heatmap. It was used, however, to see if there were any interesting sectors that could be looked at closer.

Text

Description automatically generated

The only blue area was an area of interest that the ‘Furniture and furnishings stores’ has the highest number of employees working from home two to 4 days out of the 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.

Chart

Description automatically generated

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 this was the case.

To delve into the four sectors more, gender was split out to look at the differences between the amount of time males vs. females spend working from home. A filter was also used to only be looking 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 both being 26-50% of females working from home five days a week in the ‘Ambulatory health care services’ and in ‘Merchant wholesalers’. Interestingly, where Furniture and home furnishing stores had originally had the highest number of employees working from home, does not have employees consistently working from home five days out of the week.

Chart, bar chart

Description automatically generated

The last graph looks at the number of male and female employees responding across 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 sector.

Chart

Description automatically generated

## 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.

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

Chart, treemap chart

Description automatically generated

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

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

In the next graph we wanted to look for similar trends but with gender instead.

Chart, bar chart

Description automatically generated

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

## 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 male, female, and equally male and female owned businesses. In the graph we see there are a lot of male owned business which received much more assistance during the pandemic.

Chart, bar chart

Description automatically generatedThere were less than half of the amount of female owned business that received assistance in every money group except in the $5k-less than $10k, but even in that group female owned businesses had approximately 25k less business receiving assistance.

Chart, bar chart

Description automatically generatedWe also look at the various races represented in the data. We again see white owners received a lot 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 being Asian owned businesses for every group.

Chart

Description automatically generated with low confidence

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

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

Much 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. There were more white owned businesses requesting help than all other race group combined over every money grouping from less than $5k all the way to a million dollars or more.

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

Chart, bar chart

Description automatically generated

Chart

Description automatically generated