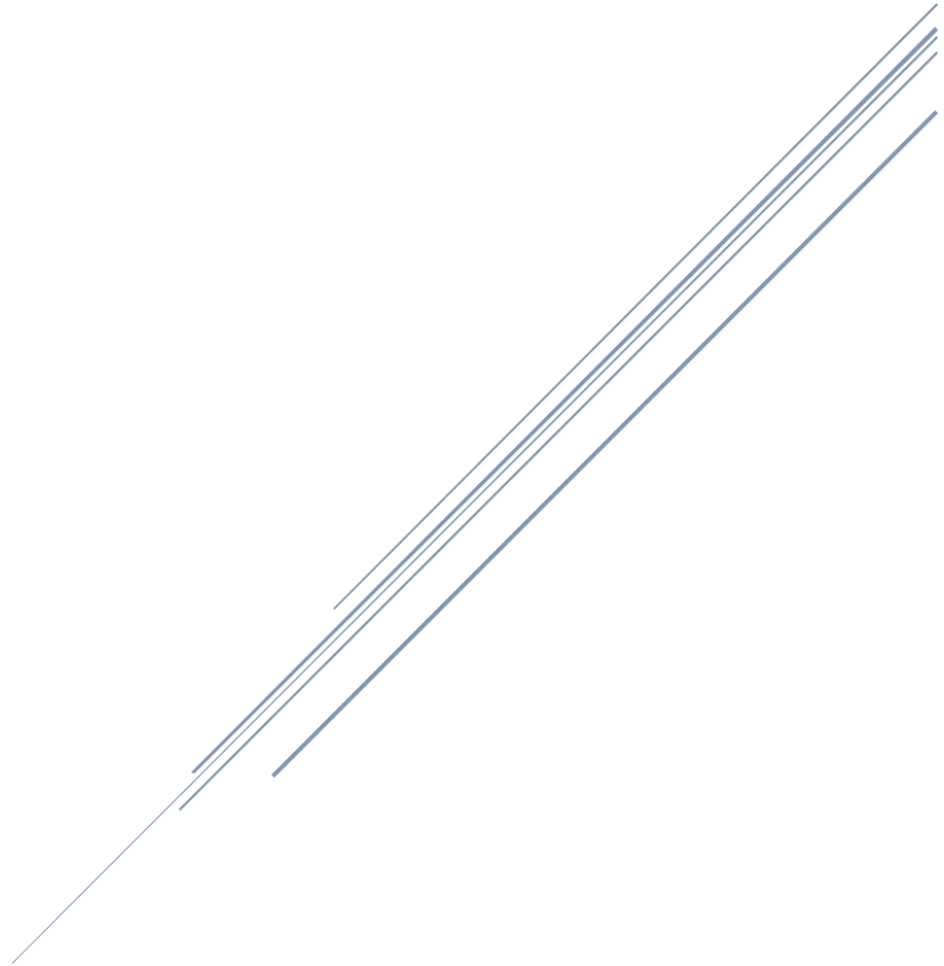# ANNUAL BUSINESS SURVEY APIS

Azure Eller, Cole Ballard, Seungmin Kim, Widnie Dorilas

PANDAS & VISUALIZATIONS ASSESSMENT

7/15/22

Identifying Sources

The sources used in this report are the Annual Business Survey (ABS) APIs. The APIs used are the state-level API access of 'Company Summary' for the years 2019 and 2020. API Keys were requested to gain access to the datasets. Using Jupyter notebooks, the data was imported, cleaned, transformed, merged and exported as a CSV. Using the cleaned and transformed data questions were formed and visualizations were used to attempt to answer them.

Initial Questions

1. How has the number of minority women owned businesses changed between 2019 and 2020?

2. How has the number of black-owned businesses changed in 2019 and 2020 for each gender?

3. What is the difference between male and female owned businesses?

4. What are the top 5 ethnic groups who own businesses in the northeast?

5. What are the top 5 ethnic groups who own businesses in the west?

6. What are the top 5 ethnic groups who own businesses in the south?

7. What are the top 5 ethnic groups who own businesses in the midwest?

8. Which states lost the most black-owned businesses between 2019 and 2020?

9. Which states gained the most black-owned businesses between 2019 and 2020?

10. Which 5 states have the highest percentage of women-owned businesses?

11. Which 5 states have the lowest percentage of women-owned businesses?

12. How has the ethnic makeup of business ownership changed from 2019 to 2020?

Answer Finding Process

Finding the answers to the questions had a matter to do with how the ETL was done. The data was imported into two pandas dataframes. Once the dataframes were explored further, the columns that were found to be unnecessary to the research were dropped. There was extra data checking to make sure there weren't null values in the dataframes and they looked as expected. Once that was confirmed, the dataframes were merged to create one so that it would be more efficient to make visualizations to answer the questions.

That one dataframe was then used as the starter base to make the visualizations that would answer the questions. To find the answers that were specifically based on gender, race, and ethnicity, there was a document provided by the U.S. Census that explained the technical information for the data that was being worked with, essentially a data dictionary. Once the data dictionary was found and put into use, the table was then able to be filtered using that information to make sure the visualizations would accurately reflect the data that was being used.

Answering Questions

After doing the visualizations, the questions were answered. Some of the answers were close to what they were assumed to be, and other answers were not what was expected and revealed new information. The first two questions ask about the difference between minority women and black-owned businesses from 2019 to 2020. The visualizations showed a clear increase, particularly in states like Florida, California, and New York.
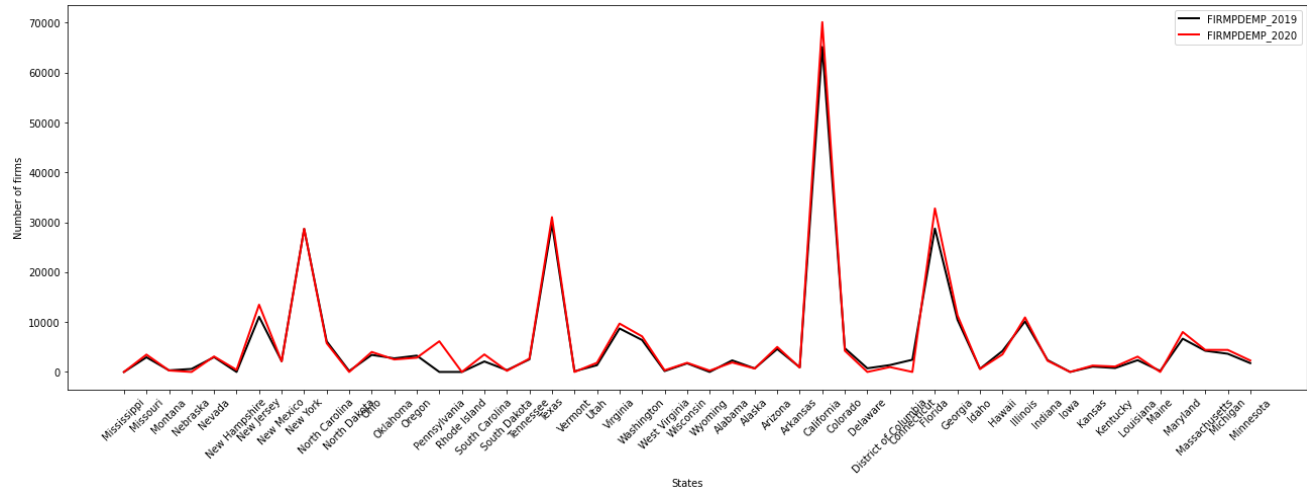
The difference between male and female-owned businesses was clear, especially when looking at the line graph for states such as California, New Mexico, and Tennessee where male-owned businesses are majority of the ownership, but there are states like Kentucky where the number of female-owned businesses surpasses male-owned businesses.

Searching for the top 5 race groups that owned businesses for each region in the United States, bar graphs are needed. In the Northeast region of the United States, the top group was white-owned businesses by a huge margin. From 2019 to 2020 there was not a notable change for white-owned businesses averaging well over 800,000 for both years. The next three groups, however, are much less when compared to white-owned businesses. The second highest group is African-American-owned businesses which had around 500,000 for both 2019 and 2020.

The top 5 business owner race groups for the West and Midwest regions of the US were very similarly distributed to the Northeast. White owners sit at the highest with over 1 million businesses while the next three race groups are Black, Native, and Hawaiian sit with just a fraction of that. The main thing to note is the disparity between the White-owned businesses and other groups.

The southern region had a slight increase in the proportion of Black-owned businesses to White-owned businesses, however, the difference between them is still there. The Native and Hawaiin race groups have the least number of businesses owned for all regions by a significant margin. Considering the proportional disparity between the white-owned businesses and other race groups, it may be more insightful to visualize the top race groups excluding White-owned businesses since they are significant outliers in each region.

The answers to the top 5 states that had the highest percentage of women-owned businesses and the top 5 states that had the lowest percentage of women-owned businesses were not what they were predicted to be. Hawaii is in the lead with the highest percentage of women-owned businesses with almost 25% of businesses. Considering Hawaii does not have as high of a population as some of the other top 5 states including Virginia and Florida, it was not predicted that it would lead. South Dakota is at the top for states with the lowest percentage of women-owned businesses at 13%.

Figuring out which states lost the most and gained the most black-owned businesses was interesting when looking at the graphs considering a lot of the states that lost black-owned businesses don't have as large of a black population as states that gained businesses. The bar graphs showed just about exactly how many businesses were lost with Washington state losing about 1,200 businesses and California gaining almost 25,000 new businesses.

States That Gained the Most Black Owned Businesses Between 2019 & 2020

States That Lost the Most Black Owned Businesses Between 2019 & 2020
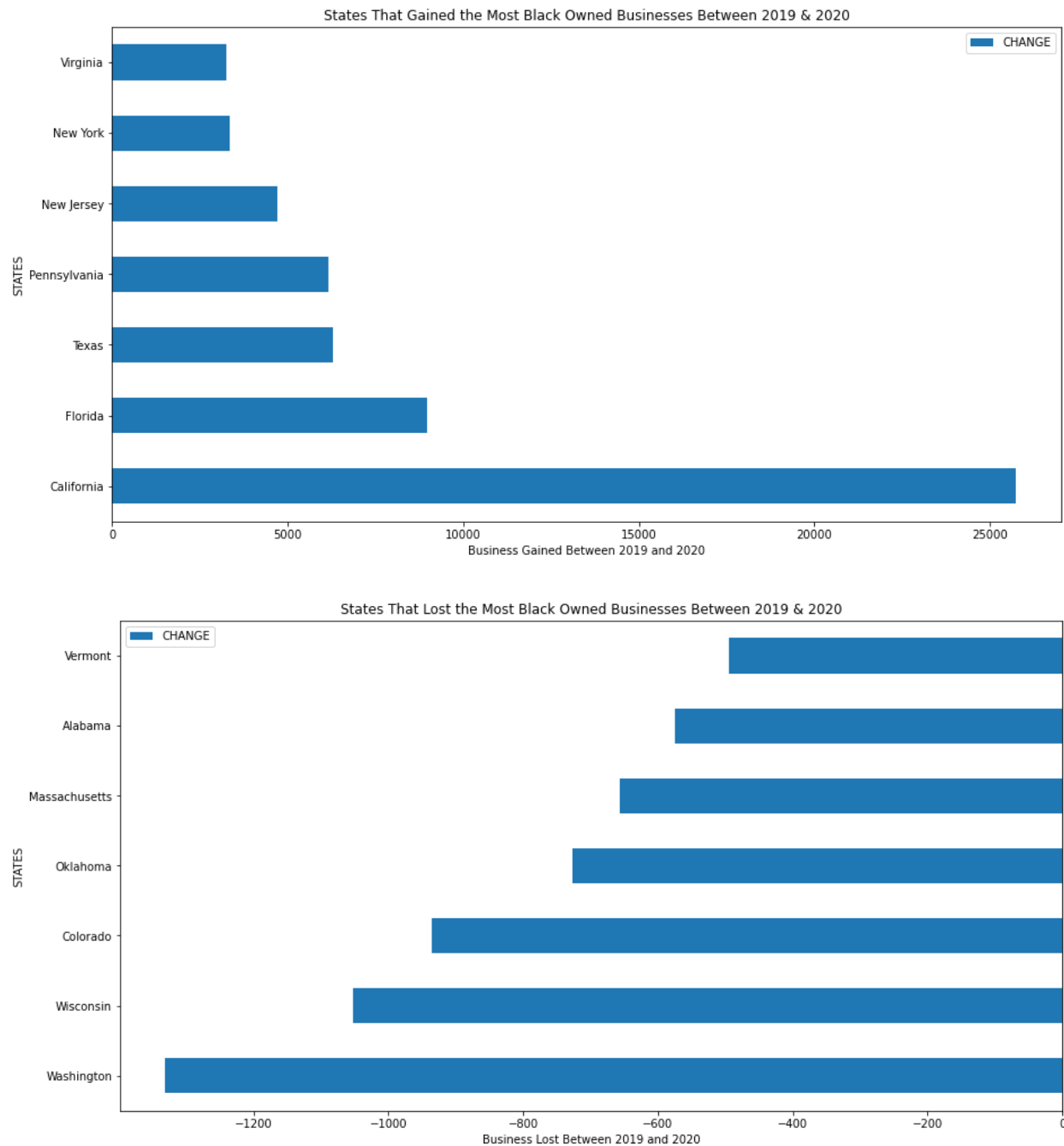
Figuring out how the ethnic makeup of business ownership changed between 2019 and 2020, a bar graph was used to see the percentage difference between the years. It showed that Native Hawaiian and Native Pacific Islander owned businesses decreased by 22% while American Indian and Alaska Native owned businesses increased by almost 10%. This is compared to white owned businesses which showed almost no change in increasing or decreasing between the years.

Conclusion

In closing, there needs to be more research done into the insights of the datasets we used. For example, gender is fluid so it might be more helpful to consider that when having businesses categorized by gender. Additionally, it would also be helpful to have the state population broken down by race and gender to see how the number of businesses is proportional to the population.

The answers to the questions asked were certainly interesting and broadened the knowledge of those topics. However, the answers to those questions also led to more questions in how we can dive deeper into this topic and find out exactly how the survey is being conducted and who is answering it.

Sources

Technical Information for ABS APIs

Annual Business Survey API for 2019

Annual Business Survey API for 2020