For the Python bridge course team project, we plan to use a data set from the US Census Bureau which contains statistics for all U.S. firms by Industry, Veteran Status, and Ethnicity for the U.S., States, Metro Areas, Counties, and Places based on the 2012 Survey of Business Owners.   The universe for the 2012 Survey of Business Owners (SBO) includes all U.S. firms operating during 2012 with receipts of $1,000 or more which are classified in the North American Industry Classification System (NAICS).

We plan to find answers for the questions below:

1. Which sectors are the biggest employers?

For this analysis, we will find the sum of paid employees by sector. We will visualize the results in the form of a pie chart. If there are too many categories in the pie chart, we will include the top 4 employers in the chart and group the remaining industry sectors into one category.

1. Do Veterans own more firms (wholly or partially) in some sectors than others? - Frank, you can have this one if you want.

We will get a sum of veterans owned firms by industry sector. We will also get a sum all privately held firms by sector (this will not include publicly held firms). We will then get a ratio by dividing the two sums and compare this ratio for all sectors.

1. Which sectors (NAICS Code) have the highest earnings per paid employee?

We will calculate the sum of earnings for each sector for firms with paid employees, as well as sum of paid employees for each sector. We will then find the ratio of the two to get revenue in dollars per employee.

1. Which sectors have the highest percentage of privately held firms? Which sectors have the highest percentage of publicly held firms?

For this analysis, we will find the sum of ‘All firms classifiable by gender, ethnicity, race, and veteran status’ which are privately held. This sum would be grouped by industry sector. We will then find sum of ‘Publicly held and other firms not classifiable by gender, ethnicity, race, and veteran status’ and group them by industry sector.

We will then visualize these in the form of a column chart with the sum of publicly traded and privately owned firms on Y axis and industry sector on the x axis.