Project Title: NYC PPP Loan Distribution Data (Loans > 150K)

Group Members:

Anne Ramer Ciara Spencer Chris Barlow

Project Description/Outline:

The topic of our analysis centers on PPP (Paycheck Protection Program) Loan distribution data provided by the SBA. As of July 2020, access to the data was made available to the public after a lawsuit was filed against the SBA by 5 major news media organizations demanding access to records, enforcing the federal freedom-of-information law. The SBA retains all records of PPP Loans as disbursed of the \$669 billion-dollar program included as part of the \$2 trillion Cares Act. The program enabled banks, upon approval of application, to collect interest on these loans given to businesses requiring the majority of it to be spent on payroll. The relatively short application to be completed for approval required the applicant to provide information on the number of employees it would be retaining payroll for, the average monthly payroll in order to determine loan amount, the business name, address, and type, as well as it's EIN which provides information of its industry affiliate.

Included within the official data provided by the SBA are additional columns for Gender, Race, and Veteran status. These were optional declarations of the application and therefore sufficient data could not be provided for thorough analysis on these categorizations of loan recipients. The additional columns included provide information on Loan Range, Lender, Approval Date, and Congressional District of each loan.

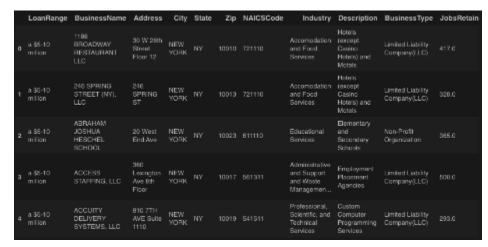
Provided within the dataset specific to NYC, as acquired from data.beta.nyc is information on NYC business loan recipients over 150K as of June 2020, coinciding with a significant decrease in sought loans. After a quick depletion of funds nearly two weeks into the programs initial opening to small businesses on April 3rd, a replenishment of \$310 billion was made on April 27th, nearly half of which was still available mid-June.

Based on research of other analysis and depictions of statistics on PPP Loan Data state-to-state and nationally, for our project we decided on the focus of our analysis to be on the economics of the loans in reference to industry, jobs retained, and five tier loan ranges for recipients of loans specific to NYC.

Because these loans are going to industry and corporations, our focus is not on whether this money went to what we're concerned as efforts we determine to be important, but rather whether the higher range loans amount went to industry enabling retainment of most jobs. According to other analyses, 1 percent of PPP Loan borrowers got over a quarter of the total loan money. Sources also reference more loan recipients in industry supportive of remote work rather than front-line industry that were required to shutdown and/or limit business. Our research will look to see if the data aligns with this specific to NYC data.

-Link to dataset and screenshot of metadata

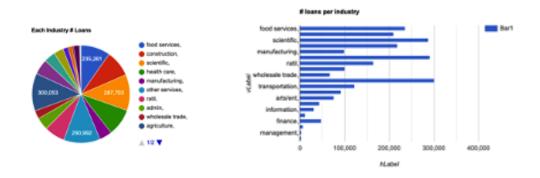
https://data.beta.nyc/dataset/nyc-ppp-recipients-over-150k



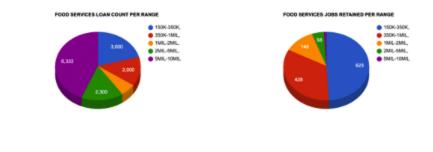
Our project will illustrate our NYC PPP loan data analysis as determined by functions applied to variables of Industry, Loan Range, and Jobs Retained.

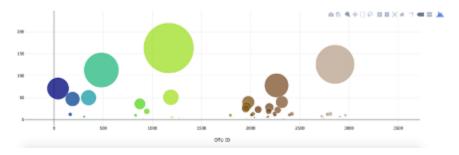


Relative visualizations accordant with our analysis in final dashboard display for *total industry* statistics:



Relative visualizations accordant with our analysis in final dashboard display for *selected industry* statistics:





Jobs Retained per selected Industry based on Loan Amount: x-axis of loan amount, y-axis of jobs retained, marker size of jobs retained (each loan) //hover data to include specific values

Charts made using https://www.rapidtables.com/tools/bar-graph.html

FILTER TABLE OPTION:



-Link to primary GitHub repository

https://github.com/cspence001/nyc_ppp_loan