# Final Assignment

June 8, 2024

Your manager was impressed by the work you did last week on the PPP data now wants you to build on it. The following tasks have been given to you.

## Task 1

Your first task is to redo the analysis for all the PPP data, but now focusing on relevant variables. More specifically you are now expected to

1. Load into an indexed Pandas data frame the variables*: Loan Amount, Borrower's State, Date Approved and NAICS Code* from all available PPP files (keep only two-digit NAICS).
2. Add in the data frame a *LoanProgram* variable that takes the value "PPP" for all rows.

Next you need to include data from a different program called Economic Injury Disaster Loan.

## Task 2

1. Download the EIDL data from SBA (link). The EIDL is another US government program designed to help business recover from COVID impact.
2. Load into an indexed Pandas data frame the variables: ACTIONDATE, PRIMPLACEOFPERFORMANCECD, FACEVALUEOFDIRECTLOANORLOANGUARANTEE for all the EIDL loans.
3. Create a new variable in the EIDL data frame that takes the value "EIDL" for all rows.
4. Retrieve the State from the PRIMPLACEOFPERFORMANCECD variable.
5. Merge the EIDL and PPP by State. At the end you end up with something that looks like:

| Index | Date | Amount | State | LoanProgram | NAICS |
|---|---|---|---|---|---|
| 1 | 1/1/2020 | 10000 | NC | PPP | 32 |
| 2 | 3/3/2020 | 240000 | CA | EIDL | |

## Task 3

1. Compute the total Number of Loans and Total Amount per State and NAICS code.
2. The attached file contains data from the US Census's Annual Business Survey. Load the data into a Pandas data frame and use the variables you think are appropriate to normalize the Number of Loans and Total Loan Amount.
3. Plot two heat maps of the US by State, that show the normalized variables.

## Task 4

1. Prepare a short report (max 2 pages) of your work commenting on the results of the two maps. The report should contain a brief description of the two loan programs, a description of your work and a commentary on which states and industries were helped the most/least by the programs.
2. The following plot is taken from Opportunity Insight's website https://tracktherecovery.org/ and shows the consumer spending differences between North Carolina and Michigan. In your report you should comment on whether the difference in improvement rate between the two states could be explained by the PPP and EIDL programs.