[Click here](https://data.sba.gov/dataset/ppp-foia) to visit the webpage for downloading the data related to this project. This project would judge your creativity and data handling & processing skills.

**Introduction:**

The Paycheck Protection Program is a loan designed to provide a directincentive for small businesses to keep their workers on the payroll duringthe time of COVID-19.

**Objective:**

Process and analyze the public data provided by SBA.gov for each US state toextract meaningful insights from features in the dataset.

**Processing:**

Download multiple datasets (state-wise), from the link provided above and merge them to form one large data-set.

**Analysis:**

You are free to perform any analysis on the data, come up with interesting and meaningful insights through analysis. The analysis can be performed on any state data, and will give different results as the data changes. Compare the data for various states, trends for loan amount above and below $ 150,000, predict behaviours and the list goes on. Share with us, whatever you can come up with after going through this data.

Some sample analysis points are stated below:

• LOAN AMOUNT:

o State Average vs National Average.

o Average loan for a particular city.

o Loan Amount grouped by other KPIs like Business Type, Race,Gender etc.

o Find a linear/non-linear relation between number of jobs reportedand loan amount.

o Amount of loan given by each lender etc.

• CITY:

o City wise distribution of businesses granted loans.

o Highest loan lender in each city.

o Compare cities on the basis of loan amounts etc.

* BUSINESS TYPE:

o Number of jobs reported by different business types.

o Average Loan Amount per business type.

o Compare loan amounts of Male Owned businesses VS FemaleOwned businesses etc.

• DATE APPROVED:

o Find months in which high amount of loans were sanctioned.

o Find how many jobs were reported in the past months and predicthow many could be reported in the upcoming months.

o Find total number of loans approved in each month and use it forTime Series Forecasting to predict how many loans would beapproved in the upcoming months etc.

You can come up with similar, creative analysis points, and show how well you can handle a large data-set.