Data Analysis is the process of bringing order and structure to collected data. It turns data into information teams can use. Data visualization is the process of putting data into a chart, graph, or other visual format that helps inform analysis and interpretation. Analysis and Visualization of datasets has always been a helpful for various reasons whether it’s for improvement of customer experience or business plans, etc. These all aspects require the analysis of the data. In 2020, the world has seen a paradigm shift across many industries, businesses, climate and to human life itself due to the COVID pandemic. The Government and many private organizations need to know the damage caused by the pandemic for reasons ranging from public welfare to business strategies. These calculations are very important for the growth and robustness of the National economy. To calculate and analyze the effects, we need data regarding the damage. Data is available as clusters in the many nooks and crannies of the internet. This data is then collected as a whole and then merged into a data-set. Even when data is amassed into data sets, it is still an enormous task to sort and make meaning out of it. This data can be simplified and visualized using various Python libraries like matplotlib, NumPy, pandas, etc. In this project the main goal is to implement the Python tools to simplify, analyse, visualize and predict different aspects under the banner “Impact of COVID - 19 on industries, climate and population.”

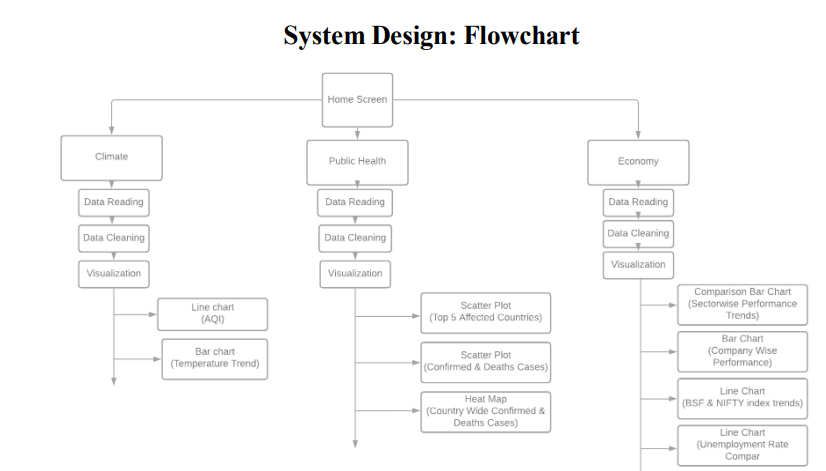
To analysis the impact of the pandemic across the globe in fields such as public health, economic effects, climate changes.- • Visualize the data by using various visualization tools available with python • Cleaning the data to improve the data quality and overall productivity • Use machine learning to predict stock market • The trained LSTM model will help us visualize how the stock market is affected due to the pandemic and based on past results will also be able to predict where the market is heading

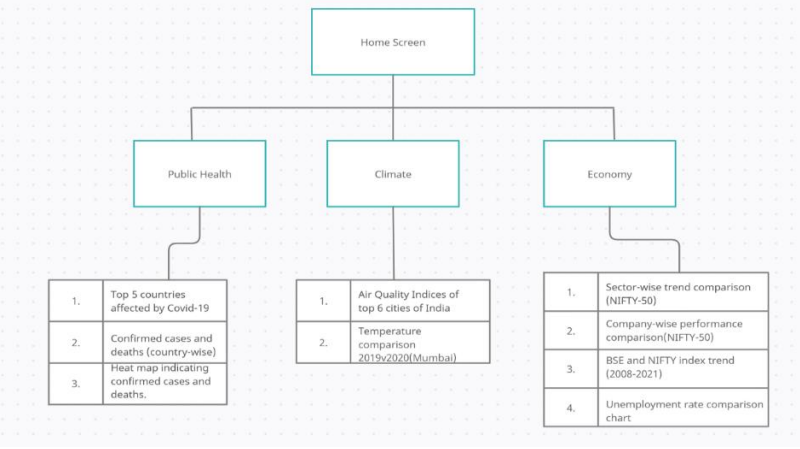
Requirement Analysis:

• Visualization Modules: Stream-lit, Matplotlib, Plotly, Folium.

• Data Reading & Storage Modules: Pandas, Pandas Data Reader, MS Excel.

• Computation Modules: NumPy, Date-Time, SK-learn, TensorFlow.





**System Requirements:**

**Hardware requirements:**

Processor: Pentium(R)Dual Core CPU RAM: 2 GB

**Software requirements:**

Operating system: Windows 7/8/10 Environment:Streamlit and Jupyter Notebook Python Version: 3.7+ The following libraries and modules are required for project implementation: o NumPy o Streamlit o Pandas o Matplotlib o Plotly o Sci-kit learn o TensorFlow o Folium o Pandas data-reader o Datetime

**Solution Methodology**

**Climate Screen:**

Data used is static and is read into a Pandas data-frame. • For the first graph, the data is segregated into 6 data frames of 6 different cities. Then the AQI indices are given as parameters to the plot function which are selected from a drop-down list in the Streamlit app. • For the second graph, the data is cleaned using dropna and fillna functions, segregated into different time intervals(2019 and 2020) and visualized in the form a comparison bar chart.

