**Python Mini Project**

The World before and during Covid-19

**ABSTRACT**

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.”

* Industries:

1. Work pattern of employees
2. Profits and Losses faced by the different types of Industries

* Climate:

1. Factors contributing to glacial temperatures
2. Carbon foot prints

* Population: (area wise)

1. Birth rate
2. Death rate
3. Divorce rate

**Chapter 1: INTRODUCTION**

* 1. **: Back ground/Motivation**

We could not find a project which contained the effects of Covid-19 on multiple variables (climate, population, industries, etc.) integrated into a single project. Hence, we decided to choose this problem statement to visualize and interpret the difference that this pandemic has brought into the world.

## 1.2: Problem Definition

Our project is helpful in visualizing several differences brought about at industrial, climatic and public level due to the pandemic by comparing historical data (before 2020) to that of the years 2020-21.

We also plan to implement machine learning module to interpret post pandemic stock prices and performance of different industries based on the current data

## 1.3: Scope / Assumptions

### Scope:

* Main aim of our project is to visualize Covid-19 related data through different data sources.
* We are currently using Jupyter Notebook as it is efficient to run code and check for errors.
* The libraries used to visualize data in our project are Matplotlib, Plotly and Folium.
* To store and clean the data we make use of the Pandas library.
* To perform numerical calculations on the data we are making use of NumPy library in our project.
* Matplotlib is a very common and powerful tool used to visualize data effectively and it also has a great community in case we face any difficulties.
* Plotly is also another library used specifically for graphs and gives the user great customization options.
* Folium is used to obtain a visual representation of a world map thus enabling us to gather the geographical insights about the data.
* Pandas is a very powerful tool which is commonly used to handle structured data which is stored in many different file formats.
* NumPy library provides vectorization of mathematical operations which significantly optimizes the computation speed.

**Chapter 2: LITERATURE SURVEY**

* 1. **: Related Work**

1. Coronatracker

* Visualizes heat maps according to frequency of cases across the world.
* Tracks the frequency of COVID cases across the globe.
* Helps us to recognize the COVID hot-spots.

<https://www.coronatracker.com/analytics>

1. Worldometer

* Detailed line graphs and bar charts which are plotted against accurate data gives us an overview of the following:

1. Death rate due to COVID
2. Total no. Of cases
3. Recovery rate
4. Monthly new cases
5. Monthly death rate
6. An insight into the daily corona cases

<https://www.worldometers.info/coronavirus/country/india/>

* 1. **: Existing System**

|  |  |  |  |
| --- | --- | --- | --- |
| **Sector** | **Existing System** | **Outcome of the Existing System.** | **Difference between the existing system and our system.** |
| 1. Public/  Industrial | “Covid-19 PANDEMIC INIDA”- M.Sc. (Data Science) – SEM II Department of Computer Science.  FERGUSSON COLLEGE (AUTONOMOUS), | Visualization of the affected population numbers and India’s GDP during the pandemic. The system only shows the effects of Covid-19 on India. | The existing system visualizations are static from limited data sources; our system makes use of several data sources for dynamic visualizations on world map. |
| 2.Industrial | Analysing the Impact of Coronavirus on the Stock Market using Python, Google Sheets and Google Finance- adilmoujahid.com | Data gathering and visualization of S&P 500 companies (USA) and how they were affected during the pandemic. | Our system visualizes not just the S&P 500 companies but also the NIFTY-50 companies’ data and the trend of NIFTY and SENSEX over the years. |
| 3.Public/  Industrial | Unemployment, total (% of total labor force) (modeled ILO estimate) -worldbank.org | Visualizes Unemployment rates of different countries and the World as a whole over the years 1991-2020. | Our system visualizes a comparison between the unemployment rates of India and the World on a single line graph over the years 1991-2020. |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |