**MINI PROJECT SYNOPSIS ON**

**COVID-19 DATA ANALYSIS**

**UNDER THE GUIDENCE OF**

**Mr. Amir Khan**

**Technical Trainer**

**T&D dept.**

**GLA University, Mathura.**

Under taken by

**NIMISHA PACHAURI 181500430**

**PARAKH TIWARI 181500450**

**Objective:** To analyse the data of covid19 contain and spread across the country and the world among the population.

**Introduction:** The project aims at deeply analysing the contain and spread of the world pandemic Corona virus, also known as Covid19 in Indian subcontinent and other parts of the world, which has made a foot fall in December 2019 from the city of Wuhan, China. The project uses various technologies for analysing the spread of the virus through charts, graphs, tables, maps and various representations methods.

Today, as we all know how much the virus is spreading and making its roots across the country and the world, the data that we will represent in this project, will give us a thorough study of the numbers upto which the virus is spreading among the locals, how steeply the number of patients are rising, the number of recoveries, active patients and the number of deceased. It also gives a thorough study of the data of specified places where the cases are more, places where the cases are less, states which are able to control the disease, the population effected, groups of ages of the people who are effected, who among males and females are effected, who are at higher risk of infection, who has lower risk of infection and many other deep studies related to the pandemic.

The project also gives a time series analysis of the pandemic which helps us to predict the future projection of the disease, which will help us to plan the future actions that should be taken for controlling and stopping its contain.

This could help us to understand the extent of corona virus spread in the country and the world and, in turn, it could help us to study what effects the measures have done, taken in the past, and now what shall be done in future to control the spread.

**Category of project:** Descriptive Analysis

**Working Methodology:** The project uses python as a working methodology. Various libraries of python is been used in this project to make this project working more effective, efficient and apprehensive. The project works on Jupyter Notebook along with other technologies where each tech used in this project has a different purpose. Some techs give us data analysed in graphical form, whereas some techs represent data tabular form and all techs used in this project gives a better cumulative result and the purpose of the project is been fulfilled.

**Details Of Technology Used:** The main platform on which the project works if Jupyter Notebook. The other techs include

are matplotlib, plotly, plotly.express , plotly.graph\_objects , cufflinks, plotly.offline , folium.

1. *Jupyter Notebook:* The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and explanatory text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, machine learning and much more.
2. Matplotlib: It is a plotting library for the Python programming language and its numerical mathematics extension Numpy. It provides an object-oriented API for embedding plots into applications using general-purpose GUI toolkits like Tkniter, wxPython, Qt, or GTK+

1. Plotly: Built on top of the Plotly JavaScript library (plotly. js), plotly enables Python users to create beautiful interactive web-based visualizations that can be displayed in Jupyter notebooks, saved to standalone HTML files, or served as part of pure Python-built web applications using Dash.
2. *Plotly.express:* Plotly Express is the easy-to-use, high-level interface to Plotly, which operates on a variety of types of data and produces easy-to-style figures. Plotly Express provides functions to visualize a variety of types of data. Most functions such as px. bar or px.
3. *Plotly.graph\_objects:* To use plotly in object-oriented way, we use this.
4. *Cufflinks:* It is another library that connects the Pandas data frame with Plotly enabling users to create visualizations directly from Pandas. The library binds the power of Plotly with the flexibility of Pandas for easy plotting.
5. *Plotly.offline:* To use plotly offline, we use this.
6. *Folium:* folium makes it easy to visualize data that's been manipulated in Python on an interactive leaflet map. It enables both the binding of data to a map for choropleth visualizations as well as passing rich vector/raster/HTML visualizations as markers on the map.

***What is Pandas?***

It is a high-level data manipulation tool developed by Wes McKinney. It is built on the Numpy package and its key data structure is called the DataFrame. DataFrames allow you to store and manipulate tabular data in rows of observations and columns of variables. There are several ways to create a DataFrame.

***What is NumPy?***

NumPy is a python library used for working with arrays. It also has functions for working in domain of linear algebra, fourier transform, and matrices. NumPy was created in 2005 by Travis Oliphant. It is an open source project and you can use it freely.

***What is DataFrame?***

**Pandas DataFrame** is two-dimensional size-mutable, potentially heterogeneous tabular data structure with

labeled axes (rows and columns). A Data frame is a two-dimensional data structure, i.e., data is aligned in a tabular fashion in rows and columns. Pandas DataFrame consists of three principal components, the **data**, **rows**, and **columns.**

**Scope Of The Project: The project will reduce the searching and analysing of covid19 data for the users and analysts who has to search at different places for such data, maintain the record of the data and has to generate reports of the contain and spread of the data. The project will also help the government, doctors, police, media, journalists and other concerned authorities who require the data for analysing purpose and report making and to place the future actions that should be taken for controlling the same, knowing the past mistakes.**

**Contribution Of The Project: The project contributes for methods to be taken for controlling and stopping pandemic. The data can give huge success in efficient fight against the disease, can save lives, could reduce deaths and could give the authorities effective ideas for measures to be taken worldwide.**

**References: The source of the data used in the project is:**

* **google.com**
* **mohfw.gov.in (The official site of Ministry Of Health And Family Welfare of India)**
* **covid19india.org**
* **worldometers.info/coronavirus**
* **Some other sites as well for collecting data.**
* ***(Data upto 31st July, 2020)***

**Conclusion: All it concludes that with the study and hard work of our team, this is something we could contribute for the fight against the ongoing pandemic through our Data Analysis on the same and the useful information that can be extracted which could help in analysing the situation, taking control of the measures taken in the past, predicting future situations carving different measures to fight against it and bringing life back to normal, saving billions of life.**

**THANK YOU!**