**Exploratory Data Analysis**

**What is EDA?**

Exploratory Data Analysis (EDA) refers to the method of studying and exploring record sets to apprehend their predominant traits, discover patterns, locate outliers, and identify relationships between variables.

**Approach followed for EDA:**

For performing exploratory data analysis on given COVID-19 country based segregated report in jupyter notebook we first imported proper python libraries and used their different functions to get inference.

* **Pandas:** 1. To read our CSV file.

2. To get shape of dataset.

3. To get quick summary of dataset. (by .describe( ))

4. To get columns with their datatype. (by .info( ))

5. To check number of unique values. (by .nunique( ))

6. To check for null value. (by .isnull( ))

* **Matplotlib:** To show all the plots.
* **Seaborn:** To draw different plots to get relation between different attributes and get better inference in visualised format.
* **Sklearn:** For encoding our data. (to change string part that is not understandable by machine is converted into number format)

**Work done on given dataset:**

* By using above mentioned functions of pandas we get shape, number of rows and columns of our given dataset.
* We get number of unique values in different columns of this dataset.
* Then by checking for null values in the dataset and we get to know that there is no null value in the given dataset.
* By scatter pair plot we get relation between different values.
* By label encoding we give encoding number to our numbers.
* By hot label encoding we represented categorical variables as binary vectors.