Virtual class Hackathon

*A Research Article*

*on*

**DATA ANALYSIS ON COVID-19**

**USING PYTHON**

*Submitted* By

**Varun B G (19BCS111)**

**Venigalla Manideep (19BCS112)**

*Under the guidance of*

Dr.Uma Sheshadri

Professor

IIIT Dharwad



INDIAN INSTITUTE OF INFORMATION TECHNOLOGY

DHARWAD

Introduction

World Health Organisation (WHO) has declared COVID-19 as a Public Health Emergency of International Concern (PHEIC) on 30 January 2020. The COVID-19 infectious disease is fast spreading across the countries, impacting the health of large numbers of people, and thus requires immediate actions to prevent the disease at the community level across the GLOBE.

As a Student community , it is a concern for the community as many of our students

travel across the states and countries for their project work, and internship work in future.

The result of this RESEARCH cum Data TRENDS study of this Class Virtual hackathon is

therefore crucial. The final outcome is aimed towards understanding the infectious disease

data pattern in the GLOBAL population such that the impact of COVID-19 has minimum

effect on our students.

Objective

To understand the distribution of data in COVID-19 Cases and perform various data processing operations using DATA STRUCTURES concept for the benefit of student community for their further studies and internships abroad.

Also this project helps to know the below mentioned points.

* Analyze the data of COVID-19 using suitable graph.
* Comment on how safe a country is based on the above data.
* Compare countries on various parameters on the growth of COVID-19.
* See the country which has the highest recovery rate.
* Get to know the latest report of COVID in a country

Problems Selected

1. Identity the countries as HIGH RISK TRAVEL destination countries for Internship or Project work for next two years. Draw a trend graph of death growing rate for that country
2. Find the average death time and average recovery time of COVID-19 in various country

Data Types used in Project

* List
* Dictionary
* Strings
* Sets

We have used Pandas for dealing with the dataframe and matplotlib for drawing the graph.

Data Files and References

* Time Series of confirmed, recovered and death cases.

Centre for systems science and engineering at Johns Hopkins University

<https://github.com/CSSEGISandData/COVID-19/blob/master/csse_covid_19_data/csse_covid_19_time_series/>

Approach towards this project

* We used pandas to read .csv format of data and converted it into dataframe.
* We used the commands in pandas to deal with selecting the dataframes.
* We then used lists, dictionary and sets to store the data that we wanted to deal with.
* We used the sorting techniques for dictionary and lists and wrote the output to csv file
* We used the matplotlib package to deal with visual representation of data

Observation

* It can be seen that USA and countries in Europe have been affected too much with this disease. With almost seven lakhs confirmed cases and fifty thousand deaths , USA can be said as most unsafe country .
* The growth of this virus is increasing very rapidly as time is progressing and is very dangerous if no solution is found out.

Results

* After analyzing the data it can be concluded that the countries that are not secure for students for their further studies or internships are:

|  |
| --- |
| * + United States |
| * + Spain |
| * + Italy |
| * + Germany |
| * + France |
| * + United Kingdom |
| * + China |
| * + Iran |
| * + Turkey |
| * + Belgium |
| * + Brazil |