

Capstone project
Predictive Modelling for COVID-19 in Public Health

written By Dashuwar Kromnan

3mtt fellow with fellow id: [FE/23/71765037](#)

EXECUTIVE SUMMARY

In this project, I have analyzed the COVID-19 pandemic using data on cases, deaths, and recoveries. The aim of this analysis is to understand how the pandemic spread over time in different countries and regions. Furthermore, I applied machine learning techniques to predict future outcomes, such as the increase in cases or deaths, based on historical data.

aims and objective

1. To understand how COVID-19 spread over time in different countries and regions,
2. Analyze the cumulative number of confirmed cases, deaths, and recoveries.
3. Visualize data to identify patterns in the distribution of cases, recoveries, and deaths.
4. Apply machine learning models to predict outcomes such as the increase in cases or deaths in the future.

Data Used

I will use data with the following features:

Lat (Latitude): To determine geographical locations north or south.

Long (Longitude): To determine geographical locations east or west.

Date: The date of the daily report.

Confirmed: The cumulative number of confirmed cases up to this day.

Deaths: The cumulative number of deaths up to this day.

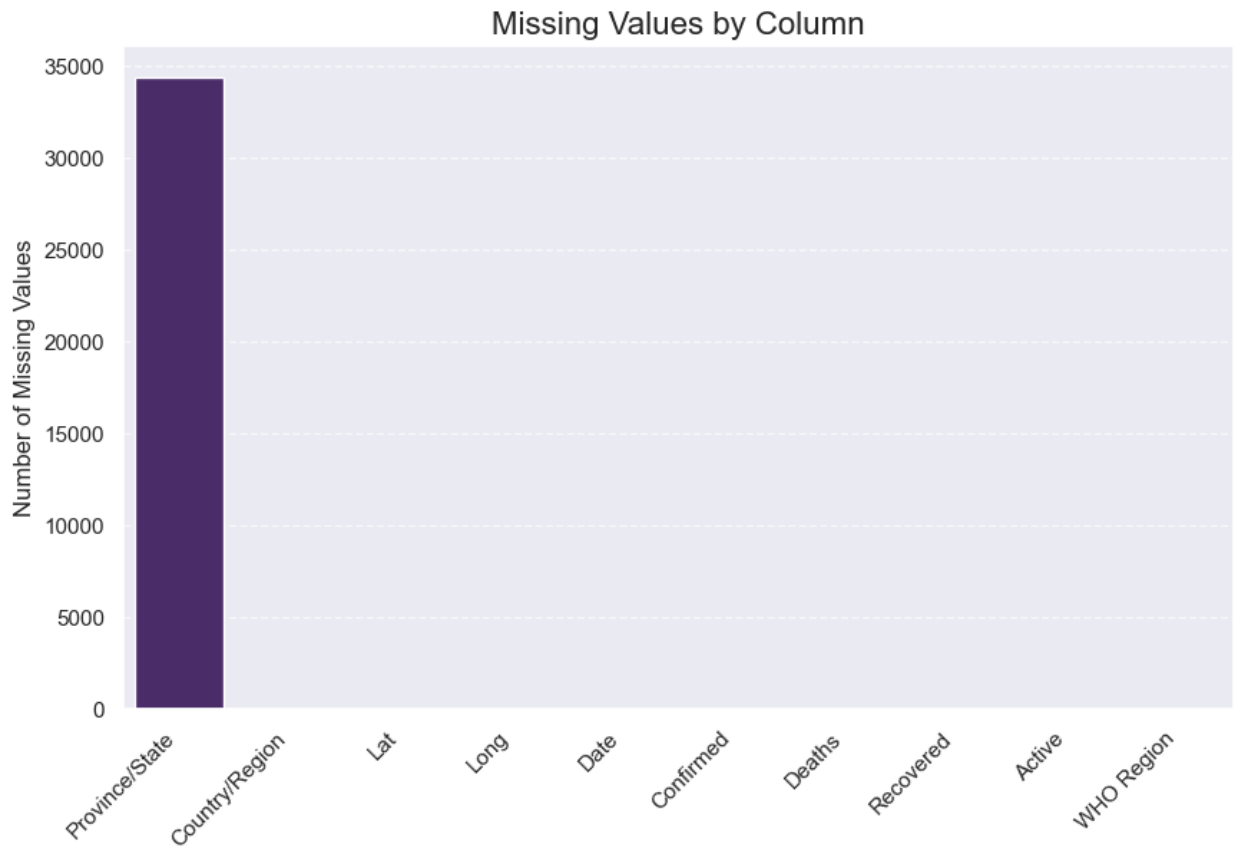
Recovered: The cumulative number of recovered cases up to this day.

Steps

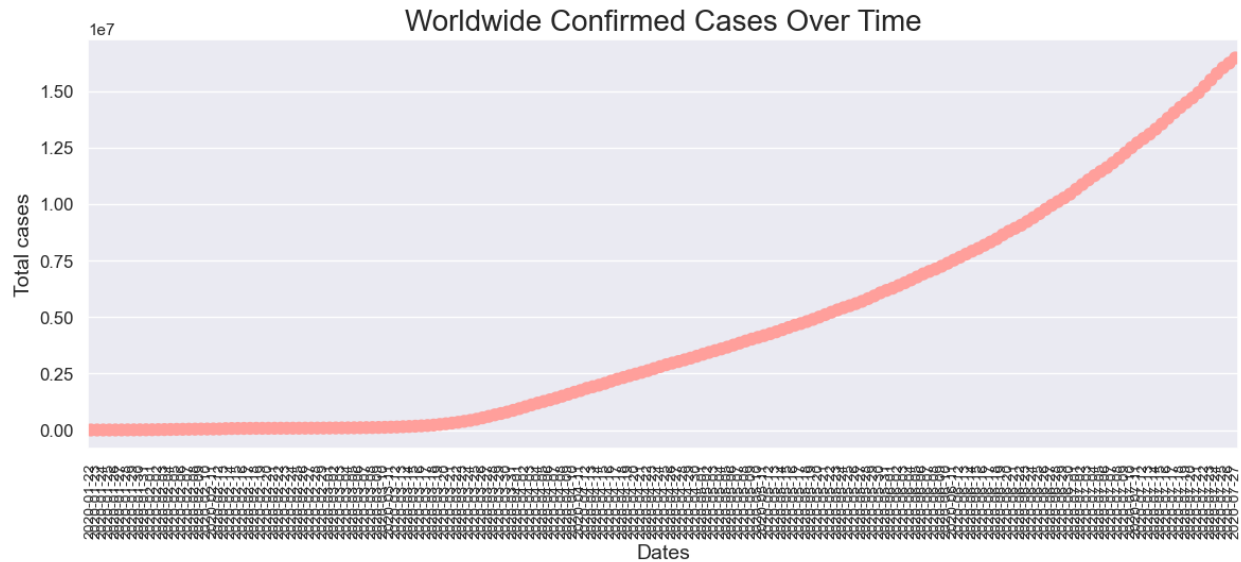
1. Imported the necessary libraries and data.
2. Clean and prepare the data for analysis.
3. Explore and visualize general patterns/ exploratory data analysis
4. Train machine learning models to predict and analyze complex patterns.

EXPLORATORY DATA ANALYSIS (EDA)

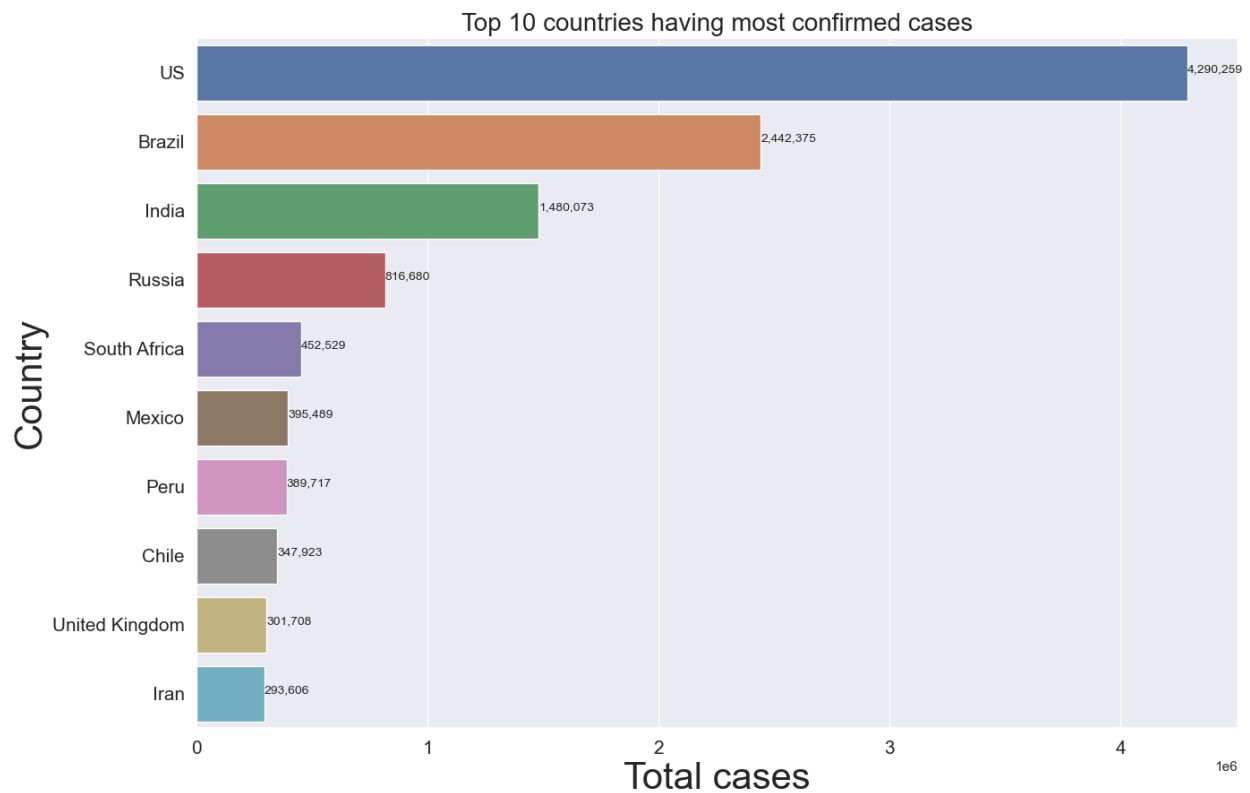
Preliminary investigations show's that, there were no missing values in terms of numerical information as shown in the chat below



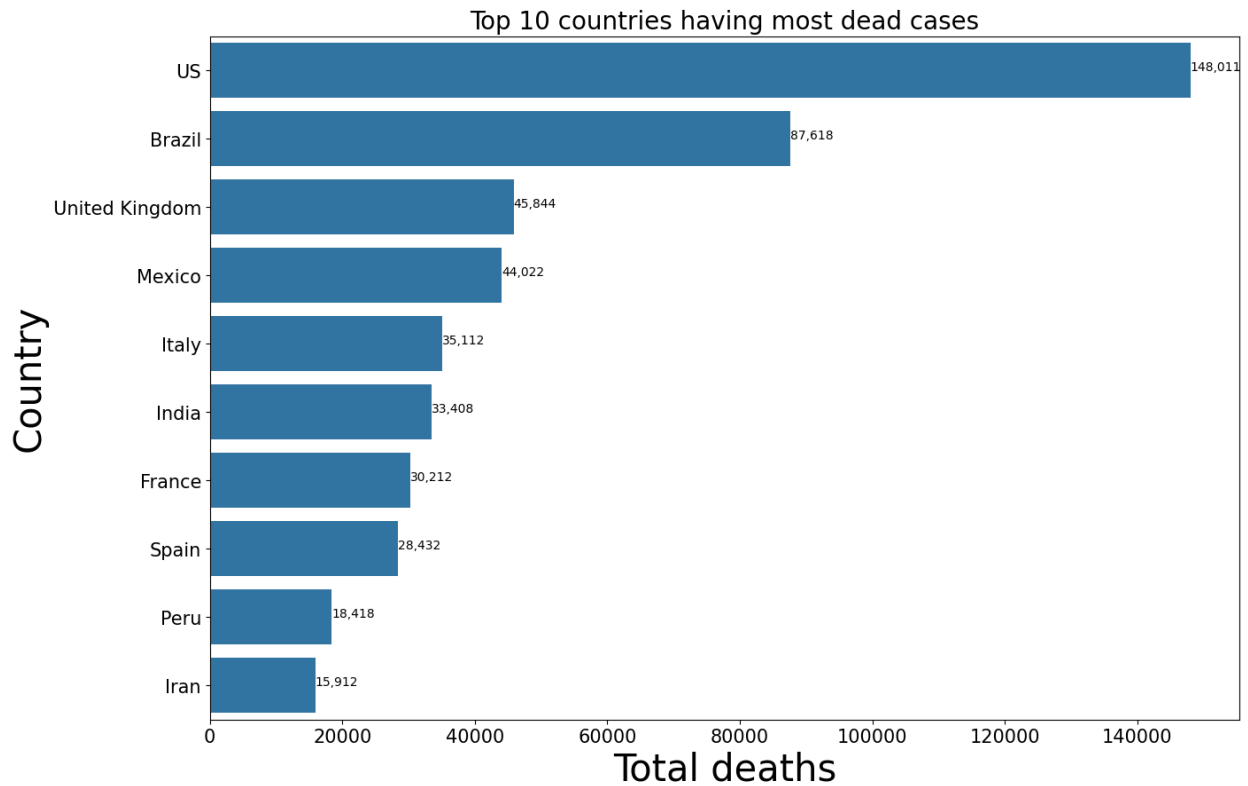
I also noticed that number of cases keeps increasing yearly across various countries of the world. This was also proven by the chat presented below.



further investigations revealed the top 10 countries with the highest number of confirmed cases with America topping the chat.

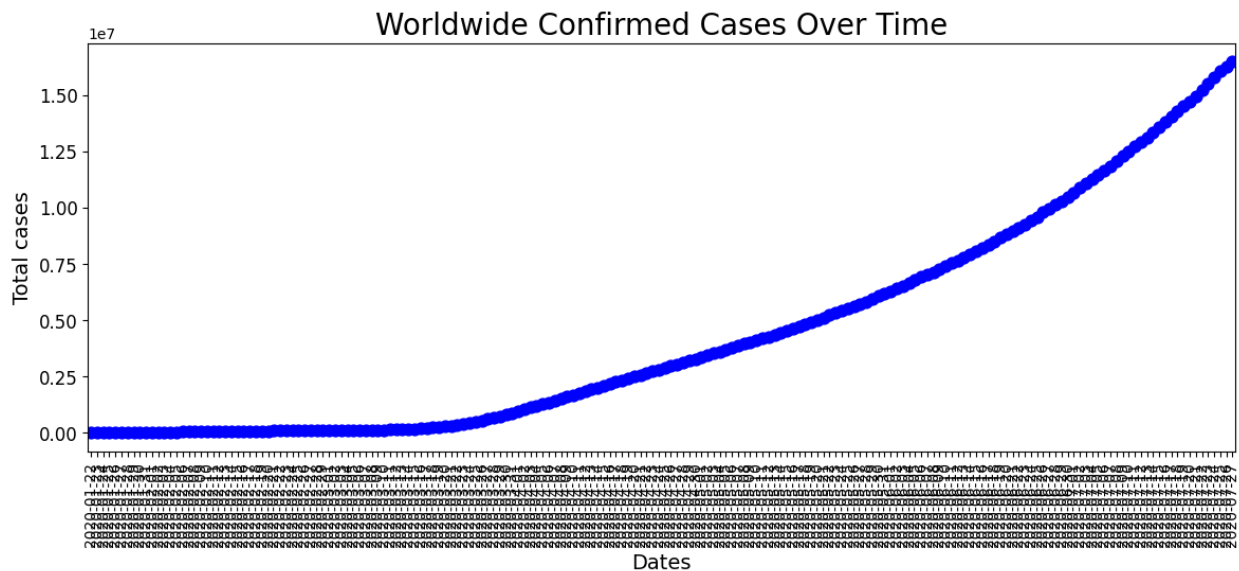


The country with the highest recorded cases is the United States of America reason being that it is the country with the highest confirmed cases

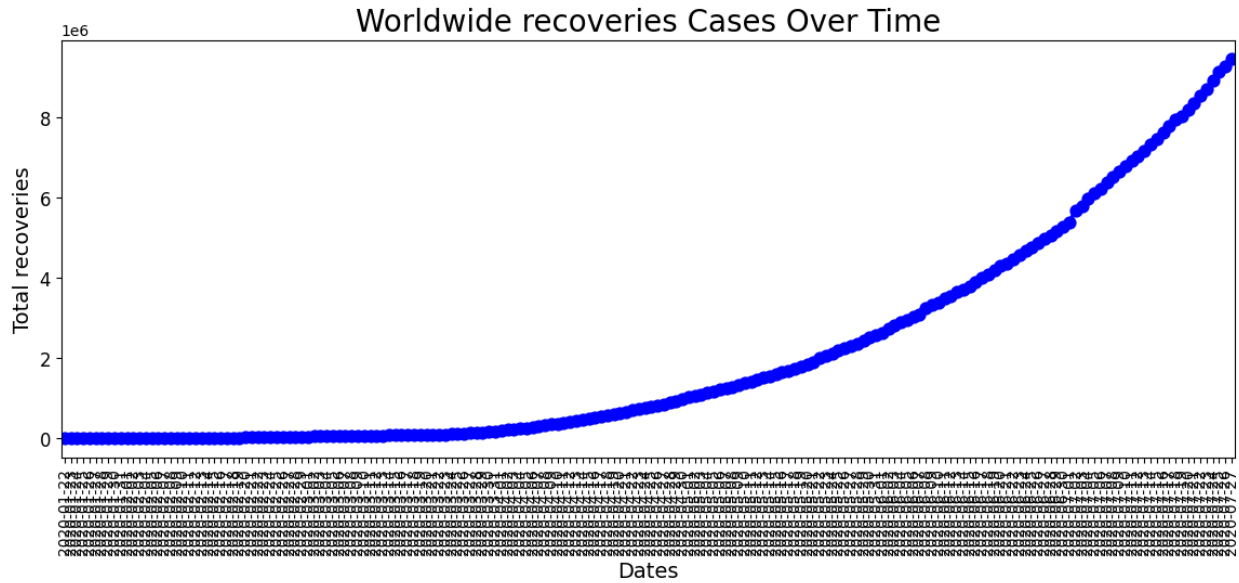


Trend analysis

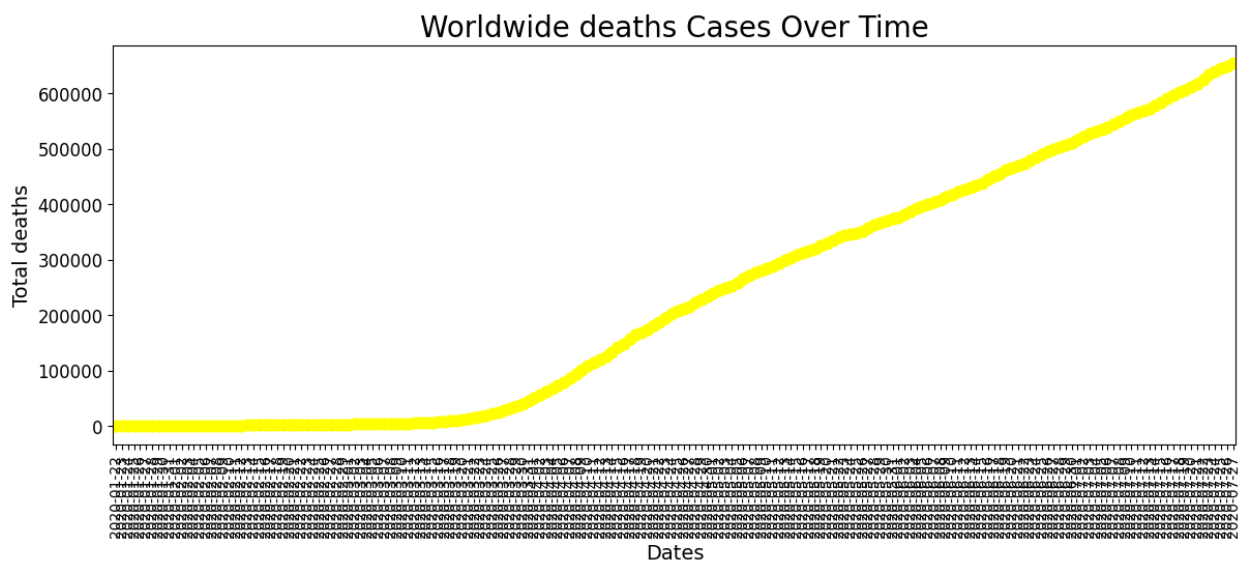
As the number of confirmed cases were increasing,



More people were also recovering, probably as a result of awareness and increase in personal hygiene and advent of vaccine .



With all these, number of deaths keeps increasing but not at the same rate with confirmed and recovery cases.



Model development

Models were built using several methods and predictions were also made

From the model developed, one can easily predict/ forecast if persons tested will be positive or not based on the symptoms presented.