COVID 19 : Current and future prediction analysis

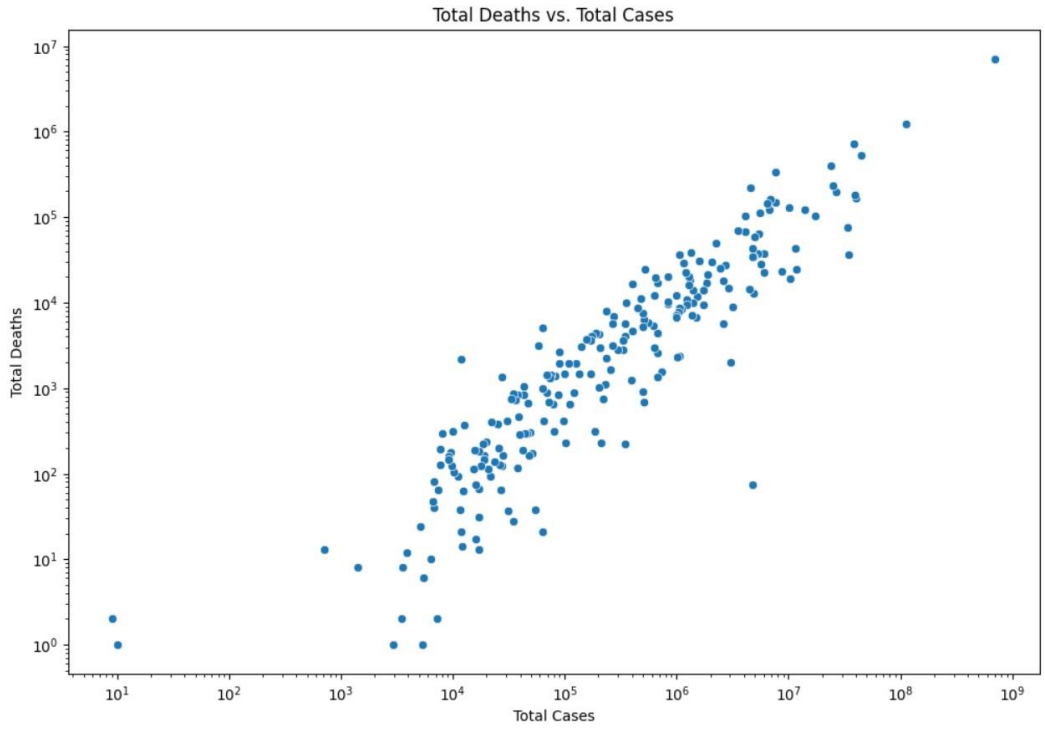
Introduction:

The global health crisis caused by the COVID-19 pandemic has altered not just health systems but also economy and social institutions. We examine the COVID-19 data in this study in order to identify patterns, make forecasts, and provide context for the current state of affairs..

Data Collection and Preprocessing:

We carefully collected information from reliable sources, such as international organisations and federal health agencies. Compiling data on overall cases, fatalities, recoveries, and other pertinent indicators was required for this.

Preprocessing was essential to guaranteeing our analysis's correctness. When needed, we scaled features, fixed data types, and dealt with missing information. The basis for a solid and trustworthy analysis is laid by these procedures.



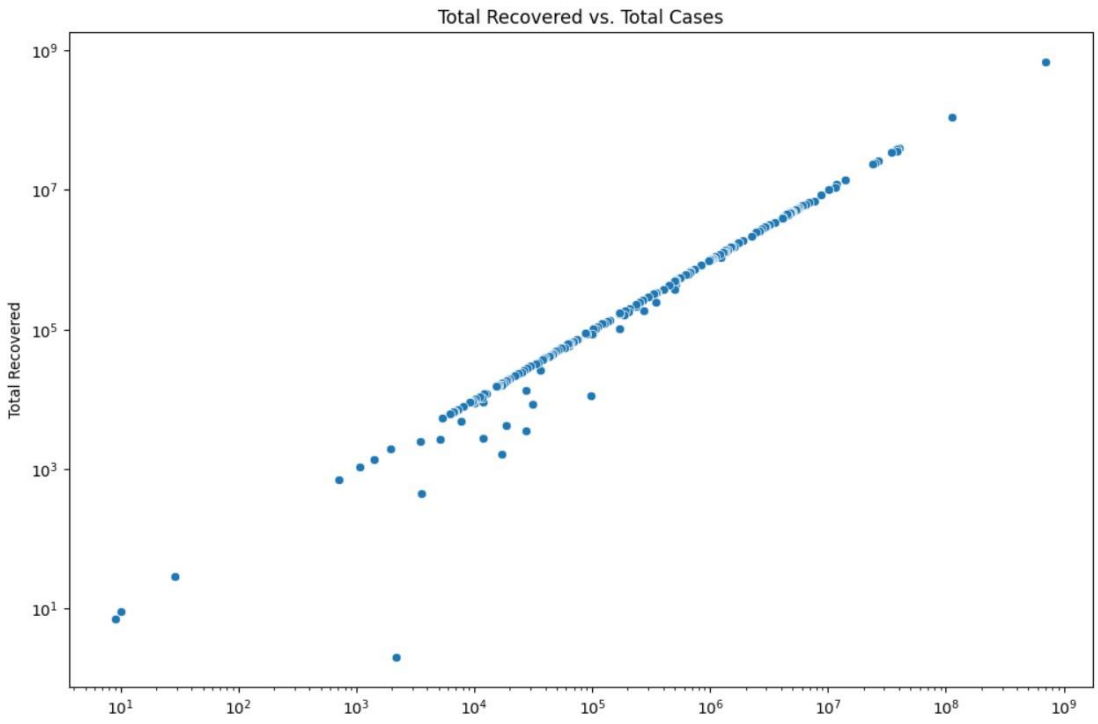
Exploratory Data Analysis:

Our preliminary investigation shed important light on the transmission and consequences of Covid-19. Key data including the total number of cases, fatalities, and recoveries were highlighted by descriptive statistics, which provided an overview of the pandemic's scope.

Finding correlations between variables and comprehending trends across time were made possible in large part by visualisations. While scatter plots showed relationships between several pandemic features, line plots showed the development of cases.

	Active Cases_text	Country_text	Last Update	New Cases_text	\
0	697743209	World	4/24/2024 4:08	0.0	
1	110600595	USA	4/24/2024 3:08	NaN	
2	44501823	India	4/24/2024 4:08	NaN	
3	39970918	France	4/23/2024 23:07	NaN	
4	38645968	Germany	4/24/2024 4:08	NaN	

	New Deaths_text	Total Cases_text	Total Deaths_text	Total Recovered_text
0	0.0	704753890	7010681	675,619,811
1	NaN	111820082	1219487	109,814,428
2	NaN	45035393	533570	NaN
3	NaN	40138560	167642	39,970,918
4	NaN	38828995	183027	38,240,600



Feature Engineering:

We created new elements like the case fatality rate, recovery rate, and active cases percentage in order to obtain a deeper understanding of the data. These measures provide complex viewpoints on the extent and adaptability of the impacted areas.

We also used log transformation to improve the performance of our predictive models on skewed variables. This conversion made sure that the bases of our analysis were solid and sound statistically.

Model Development:

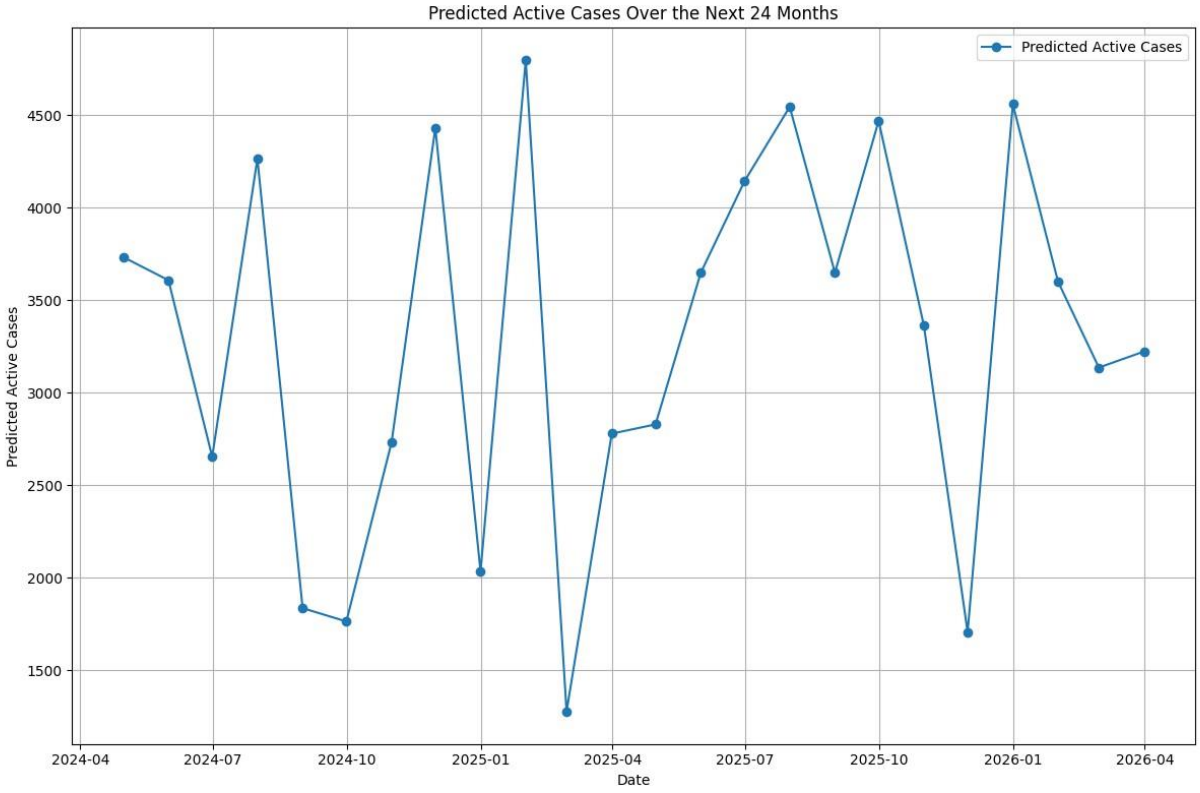
Modern methods were employed to build our prediction models, with a particular emphasis on Random Forest Regression. Our algorithms were able to predict future trends with accuracy by utilising data like the recovery rate, case fatality rate, and percentage of current cases.

R-squared scores and mean squared error, two validation metrics, showed how reliable our models were. These measurements showed how well our method captured the intricacies of the COVID-19 epidemic.

Future Predictions:

Looking ahead, we generated forecasts for the next 24 months based on our trained models. By simulating future scenarios, we gained valuable insights into potential trajectories of the pandemic.

Visual representations of our predictions provided clarity and context, allowing stakeholders to anticipate and prepare for future challenges. These insights are crucial for informed decision-making and resource allocation in the fight against COVID-19.



Conclusion:

Our research clarifies the complex dynamics of the COVID-19 pandemic and provides useful information to the public, governments, and medical professionals. We may navigate these difficult times with more efficacy and resilience by utilising data-driven tactics. We must cooperate and remain vigilant as we try to resolve this global catastrophe as a team. By continuous evaluation and

adjustment, we can lessen the effects of COVID-19 and create a more robust and healthy future for everybody.