The effect of the covid-19 vaccine on the number of deaths case study Saudi Arabia

Introduction:

The COVID-19 pandemic has brought the whole world to a halt and many people dead. Anti-vaccine groups are cropping up all over the world, even though the advantages of vaccination have been shown numerous times. The purpose of this data project is to look at the effect of coronavirus vaccines on coronavirus mortality. Several parties will benefit from this project: state governments, the health sector, and therefore all members of society. As well as vaccine production companies in the event that a close relationship appears between the number of vaccinated people and the decrease in the number of those who died due to Corona virus.

Dataset:

The dataset is from Kaggle "**covid vaccination vs death ratio**" /kaggle/input/covid-vaccination-vs-death/covid_vaccination_vs_death_ratio.csv . It's size is 2.19 MB. The dataset is containing **10 features** ["country", "iso_code", "date", "total_vaccinations", "people_vaccinated", "people_fully_vaccinated", "New_deaths", "population", "ratio"] and **20333 records**.

Features	Description
country	country name
iso_code	iso code for each country
date	date that this data belong
total_vaccinations	number of all doses of COVID vaccine usage in that country
people_vaccinated	number of people who got at least one shot of COVID vaccine
people*fully*vaccinated	number of people who got full vaccine shots
New_deaths	number of daily new deaths
population	2021 country population
ratio	% of vaccinations in that country at that date = people_vaccinated/population * 100

Tools:

-Jupiter environment, Numpy, seaborn, sklearn and pandas, libraries will imported for modeling and evaluation, matplotlib and Seaborn for visualization.

Preprocessing:

Data cleaning by make sure there are no null or empty spaces values and delete unnamed columns.

Firstly, the database has been plotted to visualize point of dataset and see relationship between features. Figure1shows total vaccinations increases over time. As well as figure 2 shows new deaths decreases over time. Now we need to examination this by machine learning algorethims.

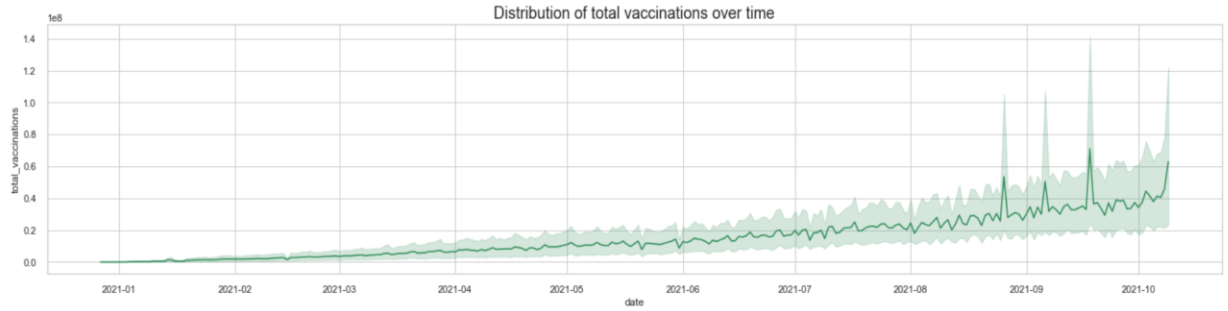

Figure 1

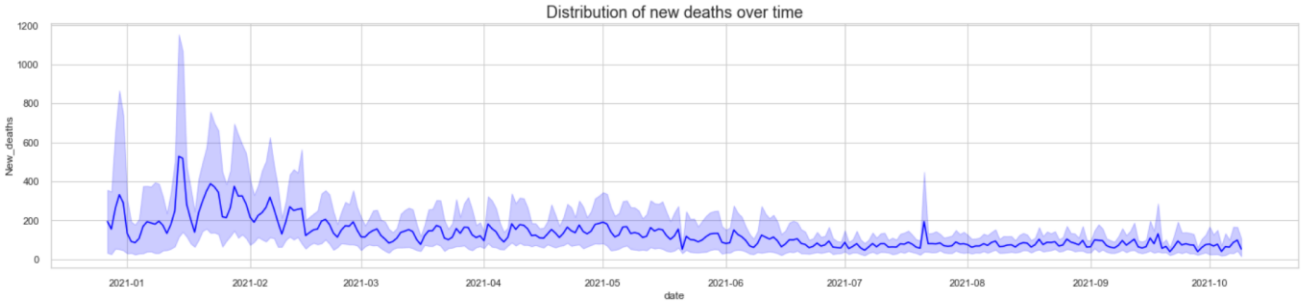

Figure 2

Algorithms:

In this project two algorithms have been used. Firstly, Liner regression with ratio and new_deaths on Saudi Arabia and Bahrain data. Then, Polynomial Regression has been used only on Saudi Arabia data. This data has been selected in order to find out the relationships between the number of new deaths in Saudi and the ratio of vaccinations. It has been found that the increasing number of vaccinations make a smaller number of new deaths.

Result:

	Linear Regression	Polynomial Regression
MAE	1.28	1.16
MSE	2.45	2
R2	0.85	0.88

Result of Bahrain was not satisfactory due to data noise and imbalanced.