Will there be an increase in the number of people infected / dead from the corona virus in a month?

Authors: Or Shemesh & Snir Shaharabani

Datasets:

- The “covid_19_deaths” database includes the number of deaths in each country from 22/1/20 to 22/7/20. Each country has a latitude and longitude that indicate its geographical location.

- The “covid_19_confirmed” database includes the parameters of the previous database only this time it is about the number of infected instead of the number of deaths.

Time Series Algorithms:

To predict the effect of the corona virus, we learned about three different time series algorithms.
We first convert the datasets to date and number of cases (death or confirmed).
The user will decide on which country to make a prediction, because each country has a different rate of infection or mortality.

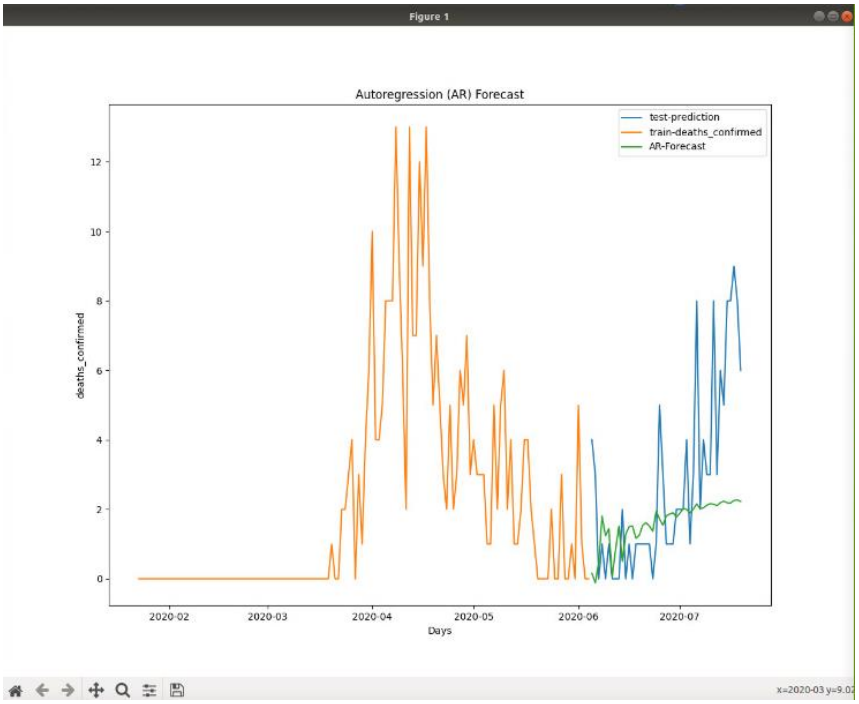
In the following pictures, we demonstrated the prediction of the number of dead and sick in Israel.

- **Autoregression (AR)**

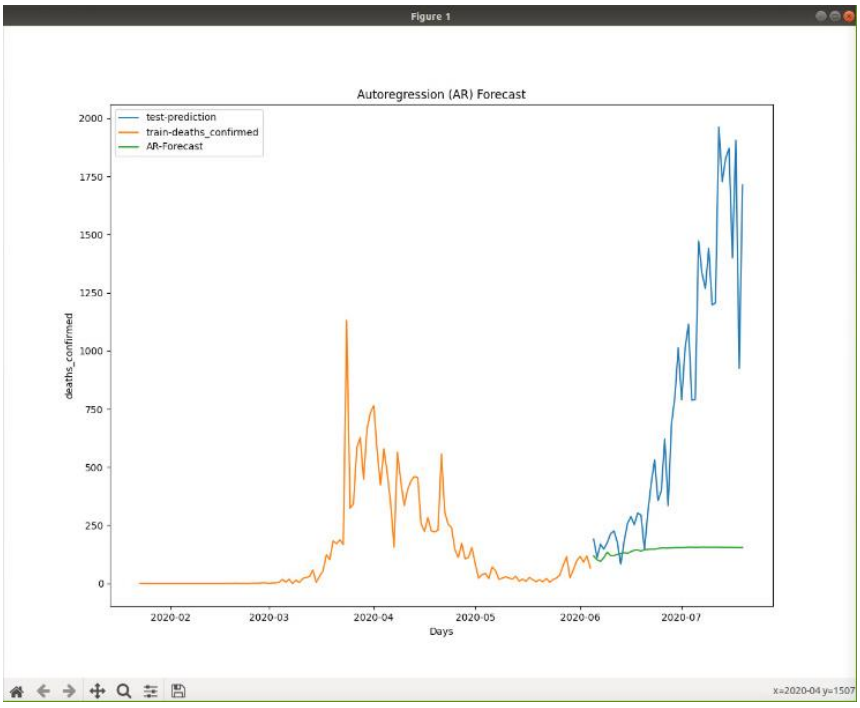
  Definition: Autoregression is a time series model that uses observations from previous time steps as input to a regression equation to predict the value at the next time step.

  Images:

  Deaths in Israel:



Confirmed cases in Israel:

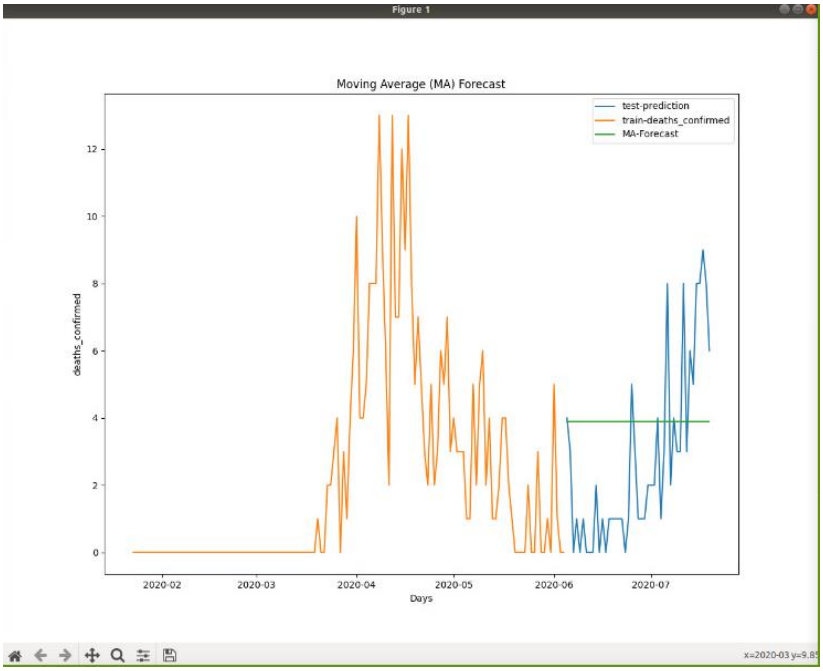


- **Moving Average (MA)**

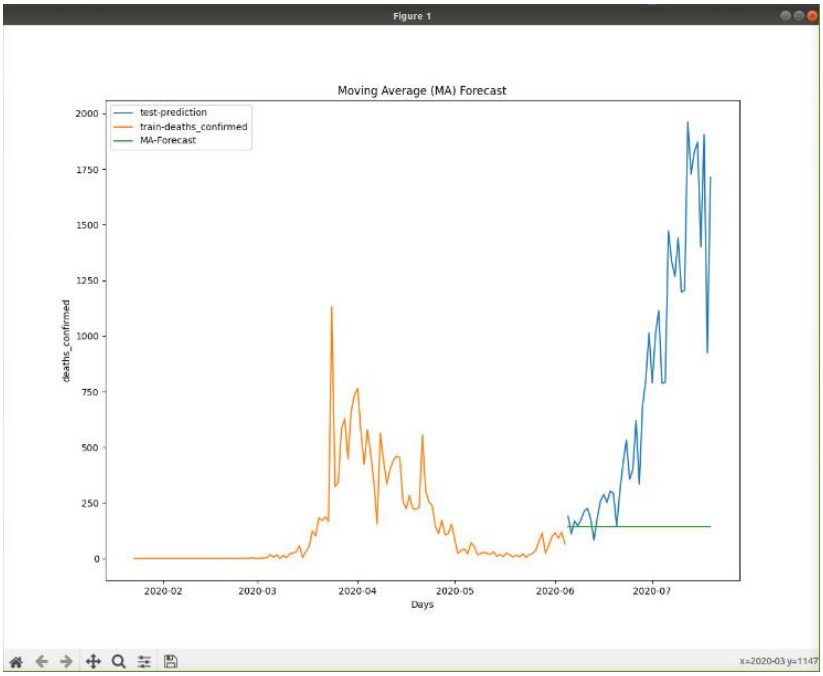
  Definition: Moving Average is a simple technical analysis tool that smooths out price data by creating a constantly updated average price. The average is taken over a specific period of time, like 10 days, 20 minutes, 30 weeks or any time period the trader chooses.

  Images:

  Deaths in Israel:



  Confirmed cases in Israel:



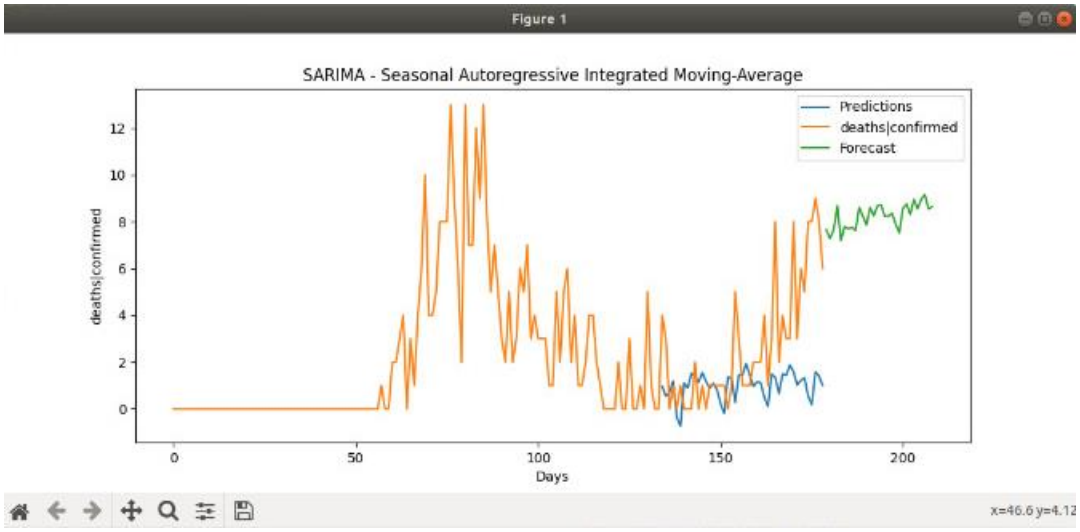
- **Seasonal Autoregressive Integrated Moving Average (SARIMA)**

  Definition: SARIMA models are a general time series model, and is used to analyze and forecast data which have an additional seasonal component. We derive values for p, d, and q in order to make the time series stationary.
  A stationary series has a constant mean and variance.

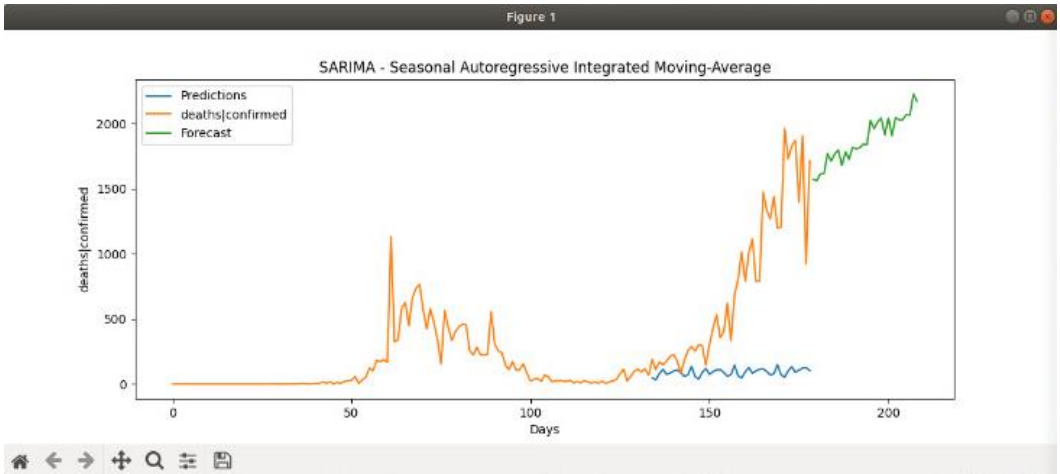
$$\text{ARIMA} \quad \underbrace{(p,d,q)}_{\substack{\uparrow \\ \text{Non-seasonal part} \\ \text{of the model}}} \quad \underbrace{(P,D,Q)_m}_{\substack{\uparrow \\ \text{Seasonal part} \\ \text{of the model}}}$$

Images:

Deaths in Israel:



Confirmed cases in Israel:



Conclusion :

As can be seen from the various models, especially from the latest model (SARIMA), for the next 30 days in Israel, the number of people infected and the number of dead from the corona virus will continue to rise.

Links :

- https://www.geeksforgeeks.org/python-arima-model-for-time-series-forecasting
- https://github.com/advaitsave/Introduction-to-Time-Series-forecasting-Python/blob/master/Time%20Series%20in%20Python.ipynb
- https://www.rdocumentation.org/packages/forecast/versions/8.12/topics/auto.arima
- https://www.statsmodels.org/dev/examples/notebooks/generated/statespace_sarimax_stata.html