**NAME:** thotalo muzammil

**REGISTER NUMBER:** 713921104055

**NAAN MUDHALVAN ID:** au713921104055

**EMAIL ID:** tmuzammil55@gmail.com

**TITLE: COVID-19 VACCINES ANALYSIS**

# Introduction:

COVID-19 vaccines have been highly effective in preventing severe illness, hospitalization, and death from the virus. However, there is still much to learn about how to optimize vaccine deployment strategies to maximize their impact. This paper will conduct an in-depth analysis of COVID-19 vaccine data, including efficacy, distribution, and adverse effects, to provide insights that can aid policymakers and health organizations in making informed decisions about vaccine deployment.

**Phase 4:**

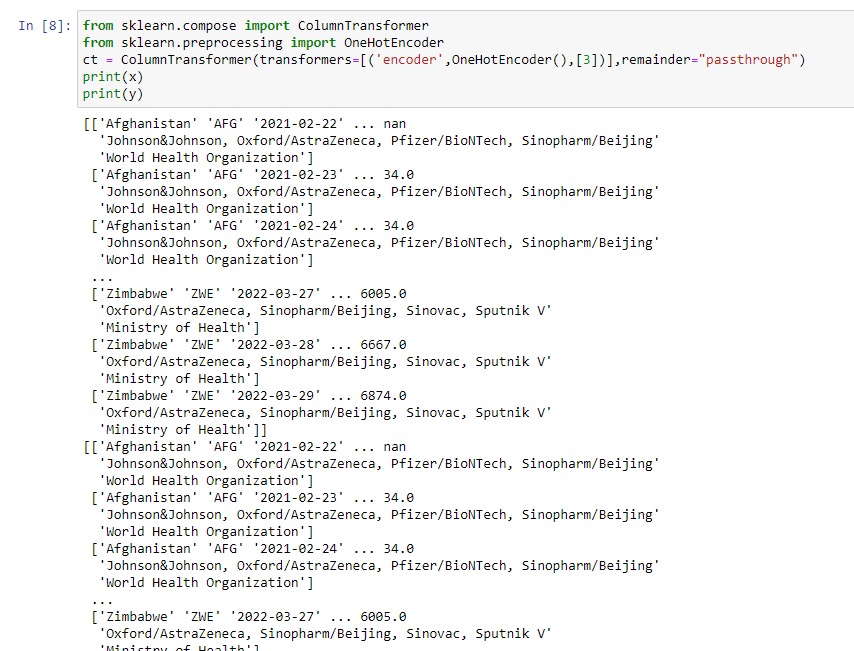
In this part will continue building the project by performing the following things

* Encoding categorical data
* Splitting the data set into test set and train set
* Feature scaling and Feature engineering
* Model training and Evaluation
* Different analysis and visualizing data.

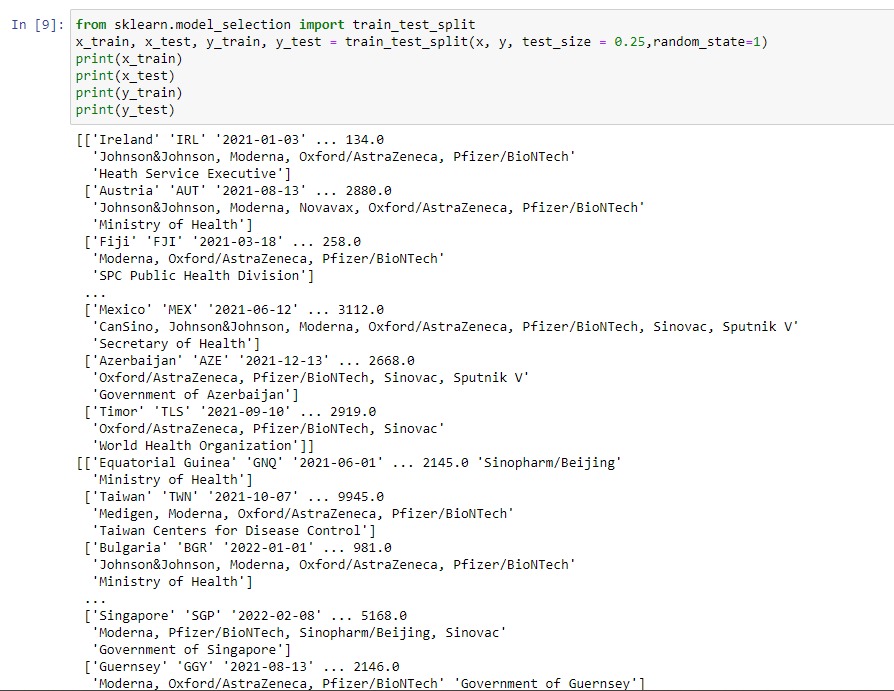
**Dataset:** <https://www.kaggle.com/datasets/gpreda/covid-world-vaccination-progress>

**1.Encoding the categorical data:**

Encoding categorical data is a process of converting categorical data into integer format so that the data with converted categorical values can be provided to the models to give and improve the predictions. Encoding categorical data is one of such tasks which is considered crucial. As we know, most of the data in real life come with categorical string values and most of the machine learning models work with integer values only and some with other different values which can be understandable for the model. All models basically perform mathematical operations which can be performed using different tools and techniques. But the harsh truth is that mathematics is totally dependent on numbers. So in short we can say most of the models require numbers as the data, not strings or not anything else and these numbers can be float or integer.



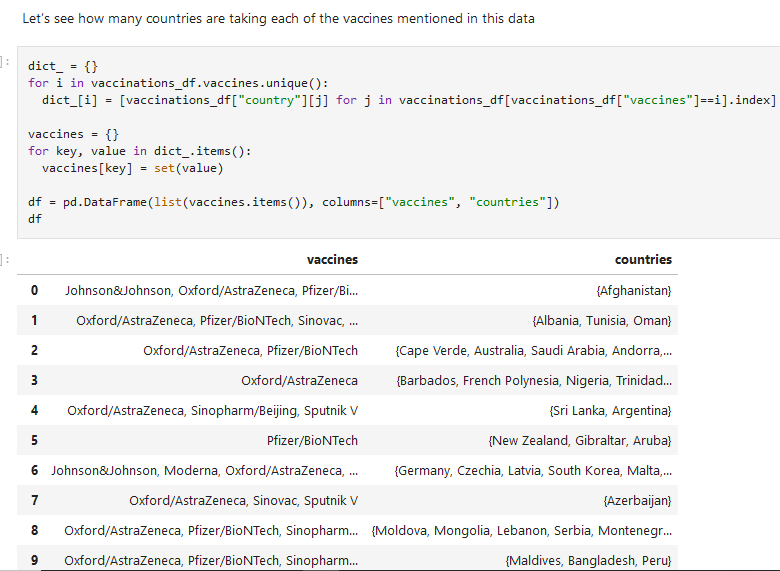
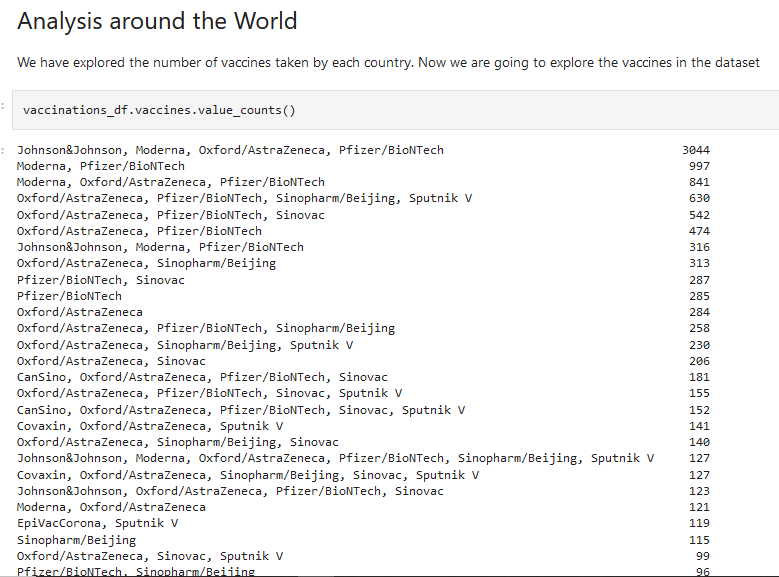
**2. Splitting the data set into test set and train set:**



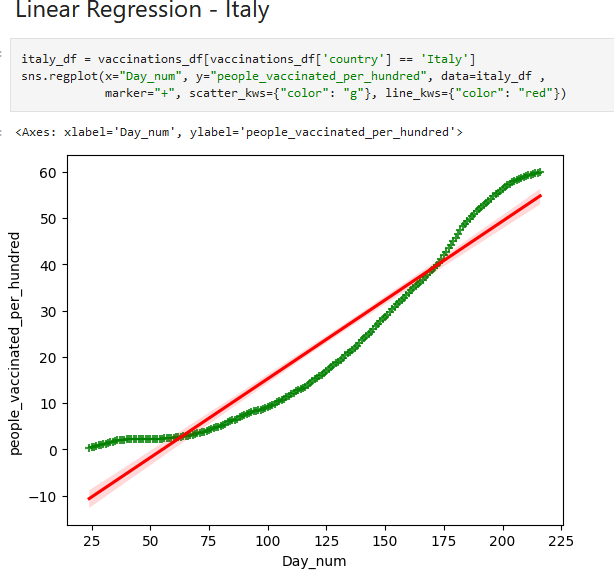
**3. Feature scaling:**



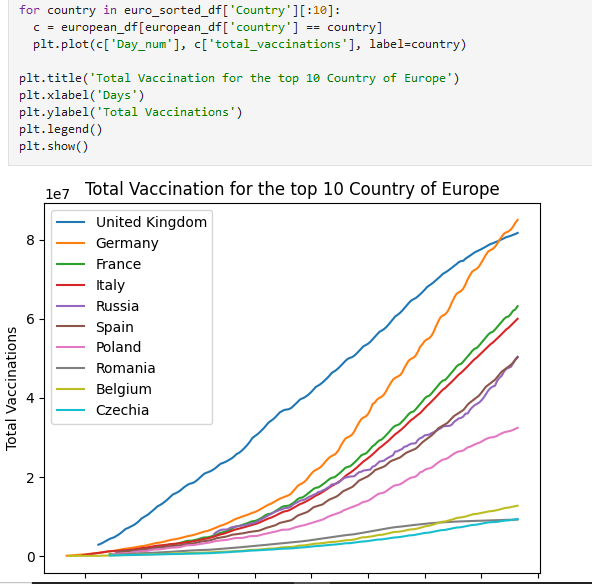
**4.modeltraining and evaluation:**



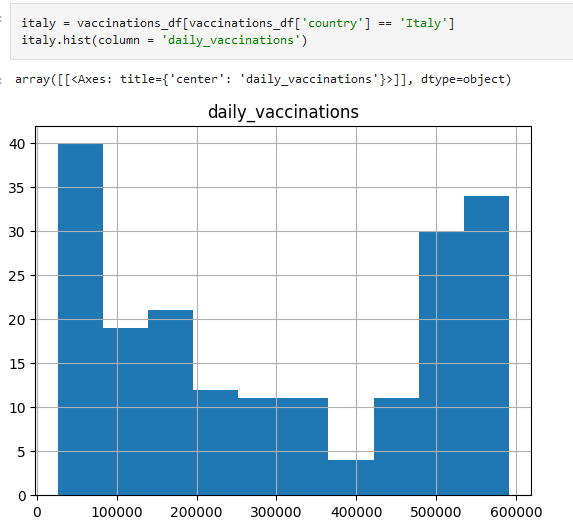
**5.different analysis:**



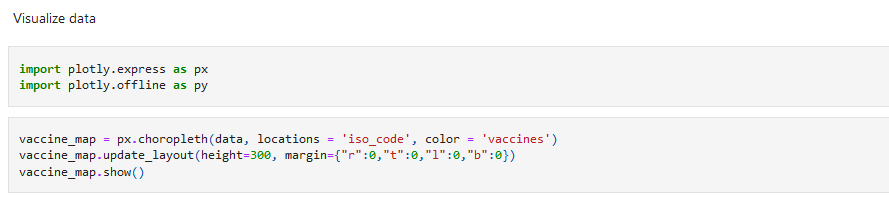
**Total vaccinations in Europe country:**

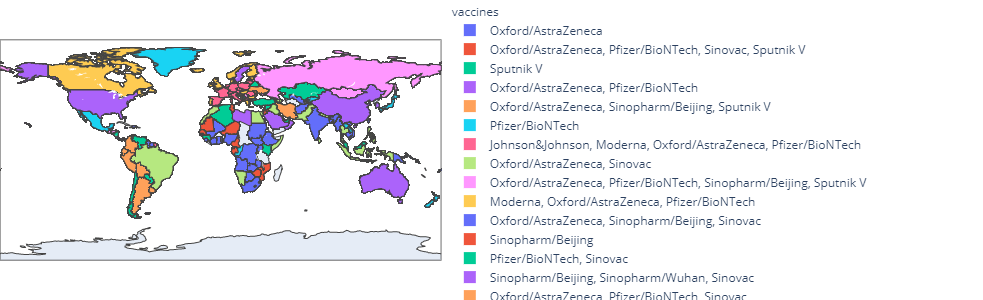
to

**Daily vaccinations:**



**6.visualizing the data:**





**Conclusion:**

All the following things Encoding categorical data, Splitting the data set into test set and train set Feature scaling, Feature engineering, Model training, Evaluation, Different analysis and visualizing data are analysed.