# IMPACT OF CO2 EMISSION FROM DIFFERENT SOURCES

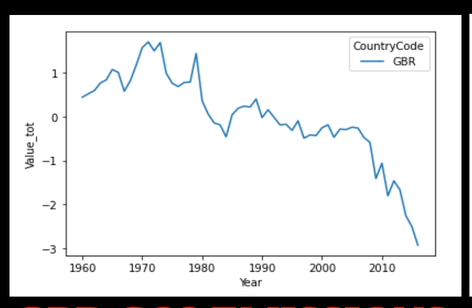
## Background

Understanding where carbon emission is the most helps fighting excessive emission by concentrating efforts on one of the three states : Solid, Liquid and Gaseous. The aim of this report is to display the distribution of various states of co2 emission from the total amount.

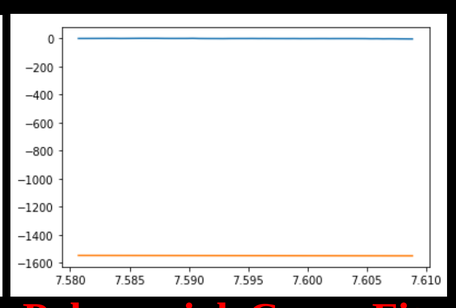
## Data

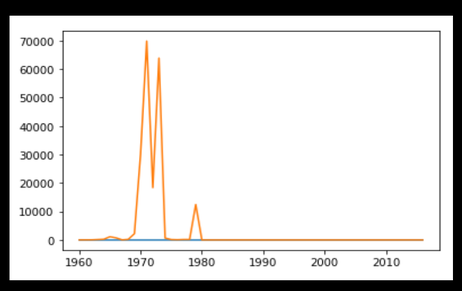
The data is taken from the world data bank dataset with the CO2 emission indicators

## ANALYSIS



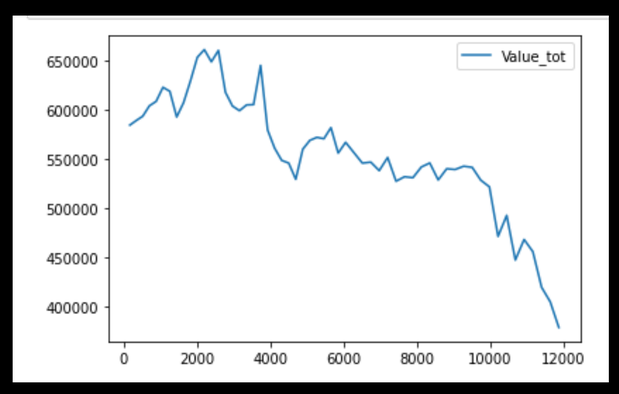

GBR CO2 EMISSIONS



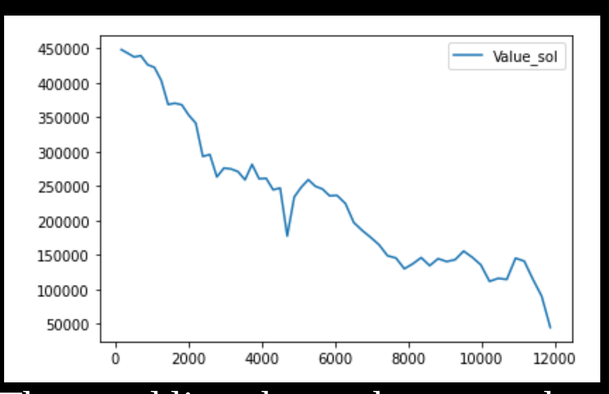


## TOAL CO2 EMISSIONS

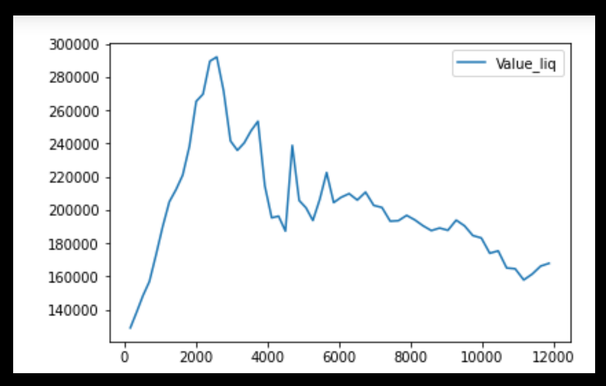



The trend line shows that over the years CO2 emission is on a constant decline meaning the waste are being handled efficiently.

## CO2 FROM SOLIDS

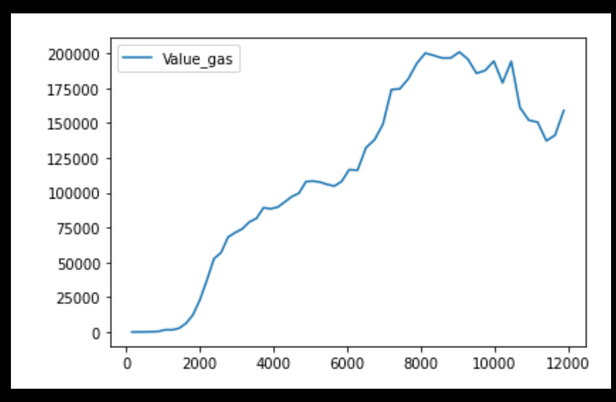



The trend line shows that over the years CO2 emission due to solids is on a constant decline meaning the solid waste are being handled efficiently.



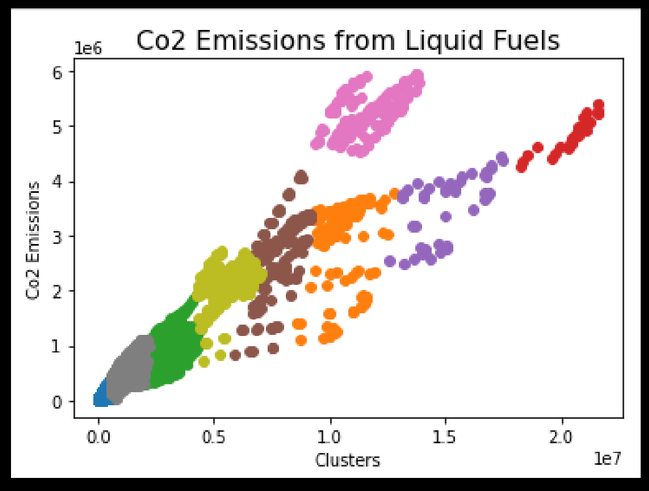

## CO2 FROM LIQUIDS

The trend line shows that over the years CO2 emission due to liquids is stationary meaning the liquid waste which emit CO2emission are not increasing overtime.

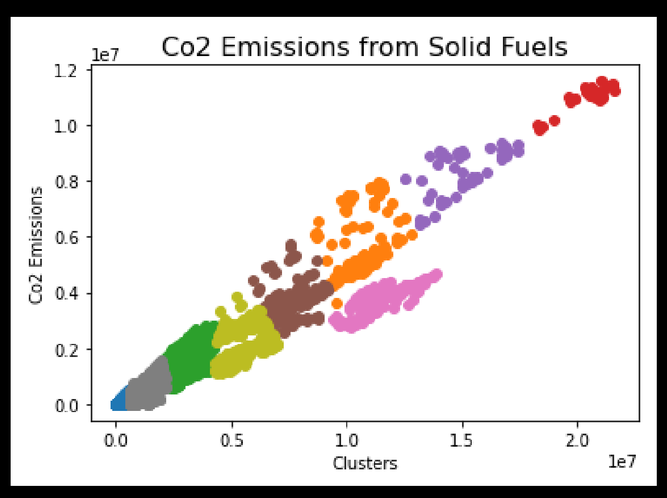



## CO2 FROM GASES

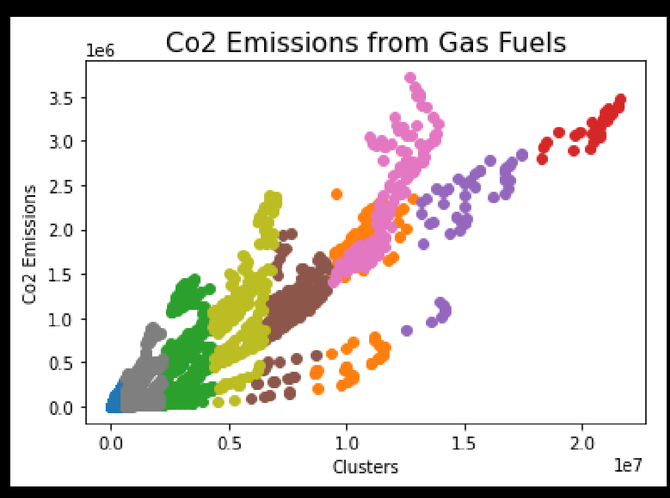
The trend line shows that over the years CO2 emission due to gases is on a constant decline meaning the gaseous waste are constantly on the rise and will eventually become a problem in the near future.



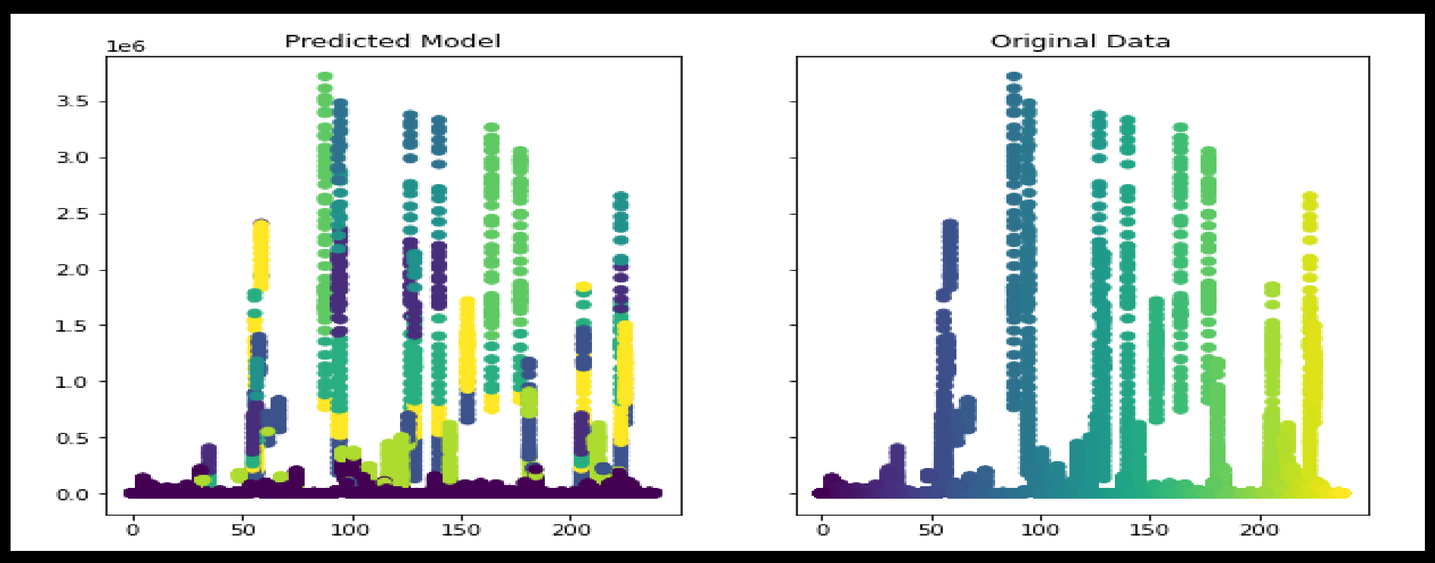

The clustering of countries based on the CO2 emission by liquid waste is shown here.



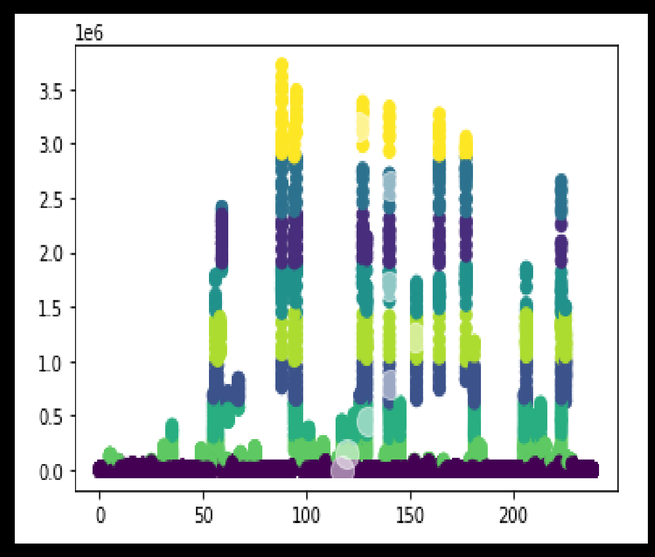

The clustering of countries based on the CO2 emission by solid waste is shown here.




The clustering of countries based on the CO2 emission by gaseous waste is shown above.




Comparison of Predictive Data with Original Data




## Conclusion:

Although the colouring between the two plots is different, We can see that predictive model did a fairly good job of predicting the clusters within the dataset. We can also see that the model was not perfect - look at the data points along a cluster's edge, We can see that it occasionally misclassified an observation from the dataset.

Hence Clustering of Data is a powerful technique to analyze and do the predictions.

## References:


Mahdavi, M., Neutatz, F., Visengeriyeva, L. and Abedjan, Z., 2019. Towards automated data cleaning workflows. *Machine Learning*, *15*, p.16.

Ramírez-Gallego, S., Krawczyk, B., García, S., Woźniak, M. and Herrera, F., 2017. A survey on data preprocessing for data stream mining: Current status and future directions. *Neurocomputing*, *239*, pp.39-57.