**COUNTRIES WITH A HUGE LAND AREA HAVE THE HIGHEST RATE DEATH RATE DUE AIR POLLUTION**

SUBMITTED TO

DR JONG-KYOU KIM

BY

ELIJAH KOJO DANSO APPAU

This report summarizes all the primary statistical modeling and analysis results associated with proving whether there is a correlation between air pollution and land to forest ratio.

To collect reliable sample data for this study I resorted to the internet and the links can be found under the reference title

Introduction

Countries with a huge land area have the highest rate death rate due air pollution

Air pollution is one of the major factors of death in the world today, from the pollutant from factories to the minute ones even from burning of charcoal and the rate at which this is rising is very sad. Air pollution is the contamination of the air due to the presence of substances in the atmosphere that are harmful to the health of humans and other living beings or cause damage to the climate or to materials. Most of these effects takes place in these low-income countries. Photosynthesis is the process by which plants use sunlight, water, and carbon dioxide to create oxygen and energy in the form of sugar. Therefore, logically if the is a balance between the intake and release of carbon dioxide the rate of pollution should be cancelled out.

*Hypothesis*

This leads me to my hypothesis; I believe the larger the forest area ratio is to the lower the air pollution therefore in the long run the lower the death rate. Trees with sunlight and sugar during photosynthesis on a regular day would utilize carbon dioxide to make their food and respire oxygen which we then use for our daily respiration; a country with a lot of this process going on is sure to have less pollution and therefore less death rates.

Analysis And Implication

Two datasets were used in the proof of this hypothesis, one had countries and their forest to land ratio from 1960 to 2020, while the other had countries with their corresponding air pollution, with the pandas merge, I was able to merge the data on country to have a compile data. The problem with this was that for the air pollution we had multiple columns of countries and their corresponding air pollution due to the fact that the air pollution was recorded every quarter or every month for most of the countries, I then found the mean using the mean function in pandas and also realised I had empty columns because during the early 1970 resources were not sufficient to collect the forest to land ratio of every single country so a dropna() function was used to remove these empty column

The next step was to find the country largest air pollution and that fell on China while the country with the least pollution was Suriname. Later on, a correlation chart was drawn using seaborns function and it proved to me negative. It turned out that China and Qatar were part of outliers, while the rest of the countries followed the trend which supported my hypothesis. I then tried to predict the forest ratio of Suriname which did not go well as the result was not realistic. The country with the lowest air pollution had more forest ratio compared to the second most air polluted country and this trend could be seen through the graph except some outliers.

Forest Area and Air pollution Scatter Plot

Chart, scatter chart

Description automatically generated

Forest Area And Air Pollution Scatter Plot with Line of Best Fit

Chart, scatter chart

Description automatically generated

Forest Ratio To Land (Suriname) Change Over The Years

Chart, line chart

Description automatically generated

Percentage Change of Suriname Forest Area Graph

Chart, line chart

Description automatically generated

Global Air Pollution Prediction Graph

Chart

Description automatically generated

Correlation Graph between Forest Area and Air Pollution

Graphical user interface, application

Description automatically generated

Conclusion

‘Some data indicate that increased neighbourhood green space is associated with lower surface and air temperatures as well as possible reduced exposures to air pollution at the household level.’ ( Zupancic, Westmacott & Bulthuis,2015 ).

Based on my analysis and research from the David Suzuki Foundation the hypothesis was true

The greater the land size ratio the bigger the air pollution. Apart From Qatar and China which were outliers, from the data and analysis done by pandas

We have a negative correlation meaning that the smaller the forest area the greater the pollution. The world today has an increased rate of global warming there’s nothing we can do to stop it, but we owe it to ourselves to try and slow it down and minimize the effect.

References

# Pavankalyan(2021)Air Polllution

# <https://www.kaggle.com/datasets/pavan9065/air-pollution?resource=download>

# (Zupancic, Westmacott & Bulthuis,2015)The Impact Of Heat On Green Space And Air Pollution In Urban Communities

<https://davidsuzuki.org/wp-content/uploads/2017/09/impact-green-space-heat-air-pollution-urban-communities.pdf>

The World Bank,Food and Agriculture Organization(2020)Forest area(% of land area)

<https://data.worldbank.org/indicator/AG.LND.FRST.ZS?end=2020&start=1961>