**Synopsis- project 1**

I took data on co2 emission, particulate matter, and respiratory death rate in this study. And then I generated a table showing the top 1 and bottom 1 countries with the highest co2 emission, particulate matter, renewable, non-renewable shares, and death rate. And as per the table, it shows Argentina is in the top 1 and Switzerland is at the bottom list when compared with 40 different countries.

Objective

My objective is to know how particulate matter and co2 emission are related to air quality, which of them contribute more to the bad air, and what relationship they have with each other. As we know the increase in both if not controlled is not good for air quality, so to see whether the increase in co2 leads to an increase in particulate or not.

Methods used for the presentation of data.

I generated a hypothesis, bar chart, Box plot, scatter diagram, R-value, p-value, T-value, and line chart for clear visualization and analysis of data. The results, methods used for analysis, and interpretation of results are below.

The hypothesis under the study

1. null hypothesis – there is no statistical significance between co2 emission and particulate matter.

2. alternative hypothesis – There is a significant relationship exists between co2 emission and particulate matter.

In this statistical analysis, the null hypothesis states no significant statistical relationship exists between CO2 emissions and particulate matter. It indicates that the two variables are independent of each other. On the other hand, the alternative hypothesis suggests that there is a meaningful correlation between the two variables and that CO2 emissions are associated with changes in particulate matter.

T-value

The calculated t-statistic of -10.6208 and the corresponding P value of 1.7471 both reject the null hypothesis, providing evidence of a significant relationship between CO2 emission and particulate matter. The T- statistic indicates the difference between the observed correlation and what would be expected if there were no relationship between the variables. in this project, the t-statistic had the largest negative value that represents stronger evidence against the null hypothesis. and the **t-statistic shows there is a strong relationship between CO2 emissions and particulate matter.**

p-value

The p-value represents the probability of observing a correlation as strong as the one observed if there were no relationships between the variables. **a very low P-value such as 1.7471 in the dataset indicates that it is highly unlikely to obtain such a strong correlation alone and supports an alternative hypothesis.**

R-value

The R-value of 0.5298 shows the direction and strength of the correlations between the two variables. **The value of R indicates that there is a strong relationship between CO2 emission and particulate matter.**

**Since all the values are against the null hypothesis, it shows that there is a significant relationship between co2 emission and particulate matter.**

**Bar chart analysis**

As per the bar plot, among all countries, India emits more CO2, and it is followed by Japan and Germany. From this, we can say that. These countries are more responsible for bad air than others.

**Line chart**

We can see the line chart is a bit unclear. As it shows, the relationship between the two variables is not entirely linear. At some points it shows inversely related and in some it shows as directly related to the study. So, I don’t feel it is a good method for this study.

**Scatter plot**

The p-value of 1.7471 suggests a significant relationship between CO2 emissions and particulate matter. The regression value of 0.52 indicates a moderate to strong negative correlation between these variables. The coefficient of determination of 0.27 indicates that 27% of the variation in particulate matter can be explained by the variation in CO2 emissions.