AIR QUALITY ANALYSIS

STAT40830 - Advanced Data Programming With R

Yashita Paras Jain - 23201169

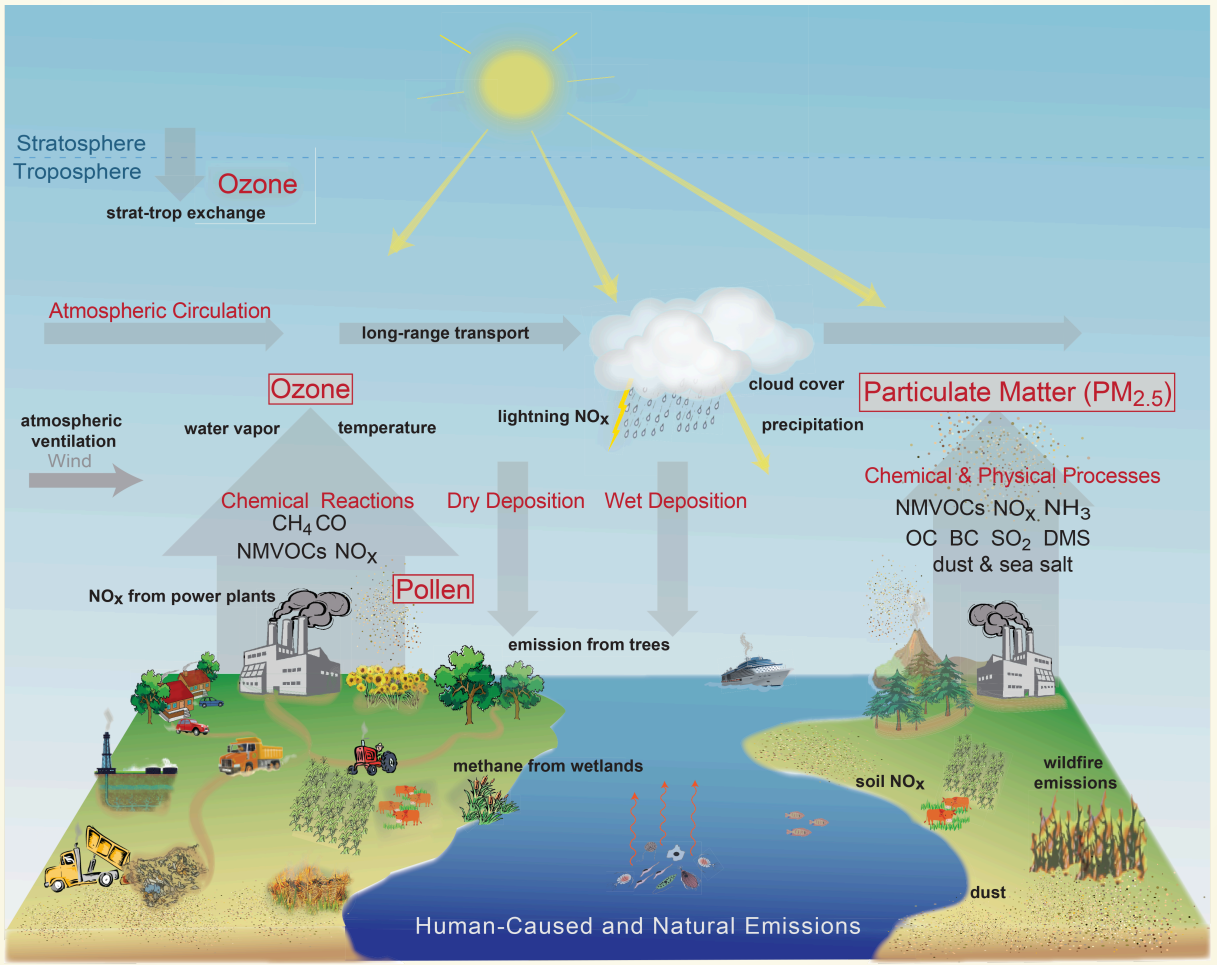
2024-06-25

TABLE OF CONTENTS

AIR QUALITY OVERVIEW

Analysis of the **airquality** dataset using Quarto with RStudio is proposed. The dataset records daily air quality measurements in New York City from May to September 1973. It includes variables such as ozone concentration, wind speed, and temperature.

Exploration of data, summary statistics and plot visualisation is done to understand the relationship between ozone levels and wind speed.



National Climate Assessment

DATASET INTRODUCTION

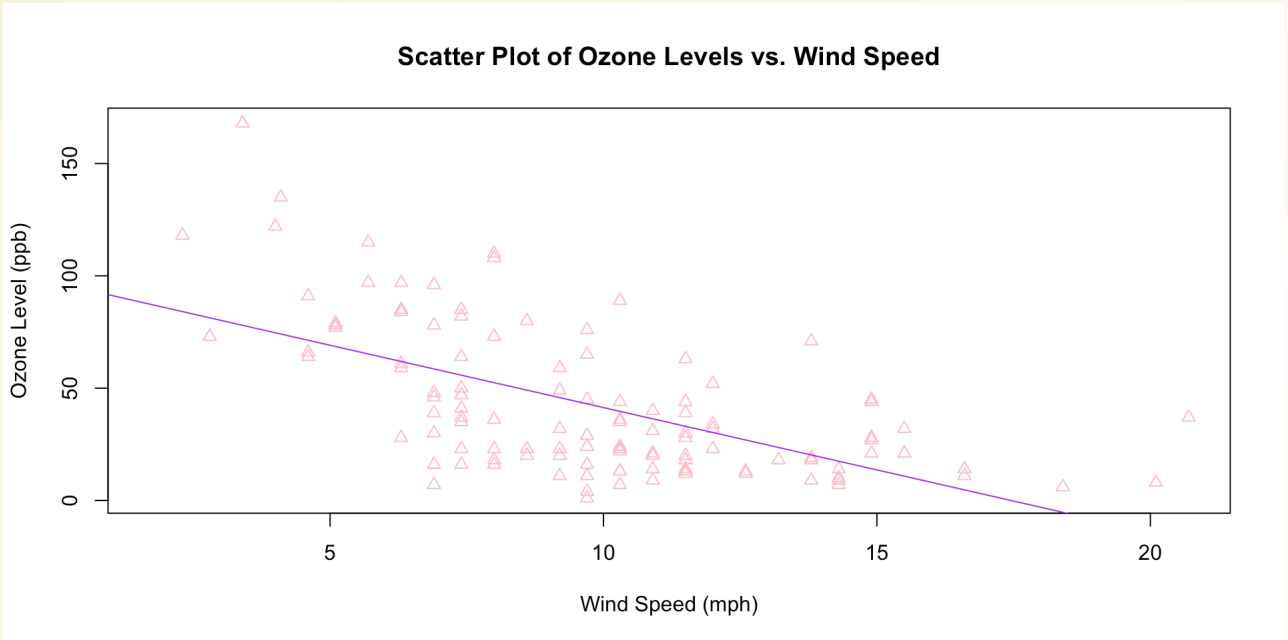
```
  Ozone Solar.R Wind Temp Month Day
1    41     190  7.4   67     5   1
2    36     118  8.0   72     5   2
3    12     149 12.6   74     5   3
4    18     313 11.5   62     5   4
5    NA      NA 14.3   56     5   5
6    28      NA 14.9   66     5   6
```

Has 153 observations on the following 6 variables:

- Ozone: Ozone concentration (ppb)
- Solar.R: Solar radiation (lang)
- Wind: Wind speed (mph)
- Temp: Temperature (degrees F)
- Month: Month (1 = Jan, 2 = Feb, ..., 12 = Dec)
- Day: Day of the month

The 4 variables - Ozone, solar radiation, wind and temperature are of important concern.

SCATTER PLOT



INTERPRETATION

The above scatter plot displays the relationship between wind speed and ozone levels.

- There seems to be a negative correlation between wind speed and ozone levels. Hence, as wind speed increases, ozone levels tend to decrease.
- This inverse relationship might be due to the dispersal effect of wind, reducing the concentration of ozone in the air.
- The red line represents the linear regression fit, reinforcing the negative trend between wind speed and ozone levels.
- Data points are scattered, indicating variability in the relationship but still supporting the overall negative trend.

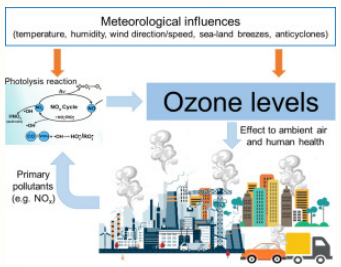
The analysis provides insights into the impact of wind speed on air quality, particularly ozone concentration, in New York City.

STATISTICAL SUMMARY

OZONE LEVELS

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
   1.00   18.00   31.50   42.13   63.25  168.00      37
```

- Ozone levels range from 1 ppb to 168 ppb.
- The mean ozone level is approximately 42.13 ppb, and the median is 31.50 ppb.
- The interquartile range (IQR) is 45 ppb (63.25 - 18.25), indicating a wide variability in ozone levels.



Science Direct Vol 18,102809

WIND SPEED

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.700   7.400   9.700   9.958  11.500  20.700
```

- Wind Speeds range from 1.7 mph to 20.7 mph.
- The mean wind speed is 9.96 mph, and the median is 9.70 mph.
- The IQR is 3.35 mph (11.50 - 8.15), showing moderate variability in wind speeds.

This presentation demonstrates how exploratory data analysis can be used to uncover important patterns and relationships in environmental datasets.