Homework 1

Air Quality Dataset Analysis

Gary Walsh - 19210998

Introduction

The `airquality` dataset from the `datasets` package in R contains daily air quality measurements in New York, from May to September 1973. The dataset includes measurements of Ozone, Solar Radiation, Wind, and Temperature.

In this document, we will plot the relationship between Ozone levels and Wind speed. The data will be color-coded based on the month to provide additional insights.

Approach

1. **Dataset:** We use the `airquality` dataset from the `datasets` package.
2. **Plot Type:** Scatter plot with color-coding for months and a smoother line to show the trend.
3. **Descriptive Statistics:** Mean Ozone levels and mean Wind speed.

Descriptive Statistics

Table 1 below presents some descriptive statistics for the airquality dataset.

Table 1: Descriptive Statistics for Ozone and Wind

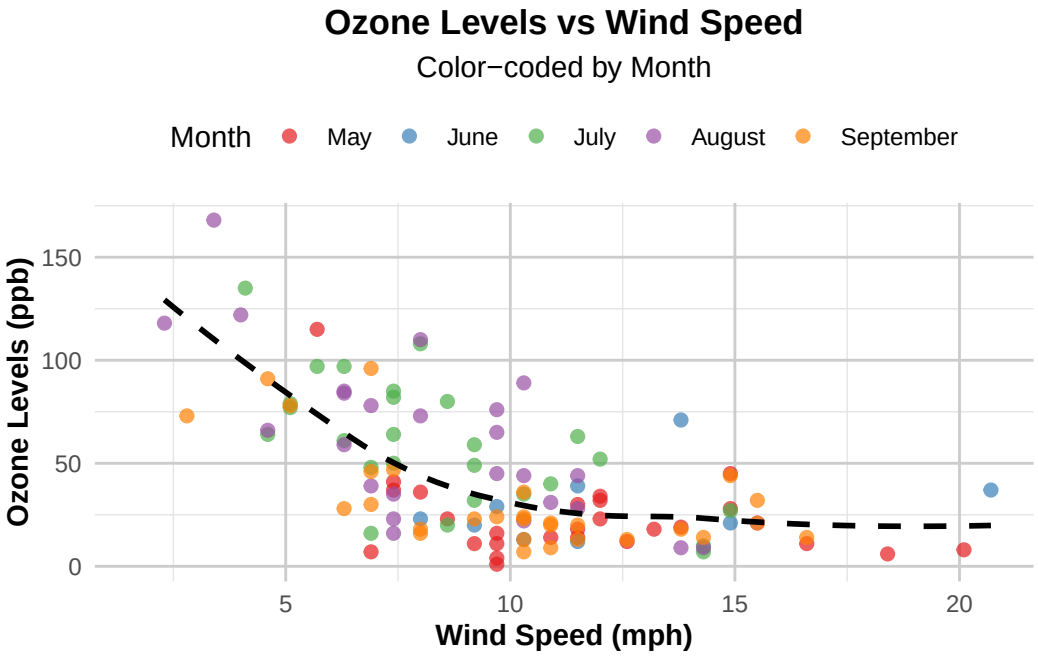
Statistic	Ozone	Wind
Mean	42.12931	9.957516
Median	31.50000	9.700000
Std Dev	32.98788	3.523001
Min	1.00000	1.700000
Max	168.00000	20.700000

Summary

Ozone levels exhibit high variability with extreme values, suggesting significant fluctuations in air quality. Wind speeds are more stable, with measurements closely clustering around the mean.

Plot 1

The scatter plot in fig 1 below shows the relationship between Ozone levels and Wind speed. The points are color-coded based on the month, and a smoother line is added to illustrate the trend.



Interpretation of the Plot

1. Inverse Relationship: Ozone levels generally decrease as Wind speed increases, suggesting higher wind speeds help disperse Ozone.

2. Seasonal Variation:

- May and June: Lower Ozone levels across various wind speeds.
- July and August: Higher Ozone levels at lower wind speeds, indicating severe Ozone pollution in mid-summer.
- September: Mixed data but overall lower Ozone levels as fall approaches.

2. Data Dispersion:

- High variability in Ozone levels at lower wind speeds, especially in July and August.
- Reduced variability and lower Ozone levels at higher wind speeds.

Monthly Summary Statistics for Ozone Levels

Table 2 below presents the monthly summary statistics for Ozone levels.

Table 2: Monthly Summary Statistics for Ozone Levels

Month	Mean	Median	Std_Dev	Min	Max
May	23.61538	18	22.22445	1	115
June	29.44444	23	18.20790	12	71
July	59.11538	60	31.63584	7	135
August	59.96154	52	39.68121	9	168
September	31.44828	23	24.14182	7	96

Additional Simple Plot: Ozone Levels Distribution

The histogram in fig 3 below shows the distribution of Ozone levels.