Basic Statistics and R Programming
Assignment 1* - Summary Document

Aalok Sharma Kafle

June, 2022

Problem 1

The objective is to plot the empirical CDF for Maximum Daily Ozone using Beard Plotting Position Formula. The daily ozone value is maximum value aggregated from hourly data of each day.

For Beard Plotting position, the probability of $P(X \geq x)$ is given by:

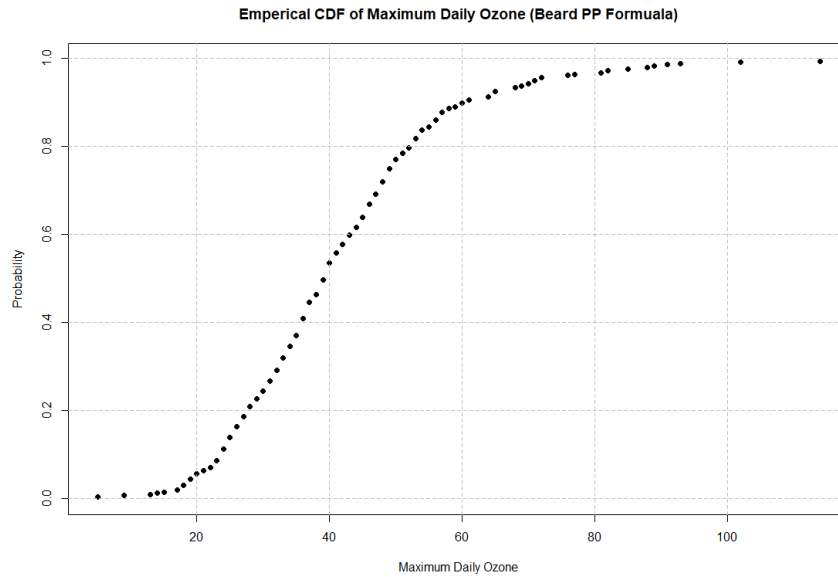$$P(X \geq x) = (i - 0.31)/(N + 0.38)$$

Figure 1: Empirical CDF for Maximum Daily Ozone using Beard Plotting Position Formula

---

Problem 2

The objective is to plot the histogram of mean daily temperature using Freedman-Diaconis Binning Formula. The daily mean temperature value is average value aggregated from hourly data of each day.
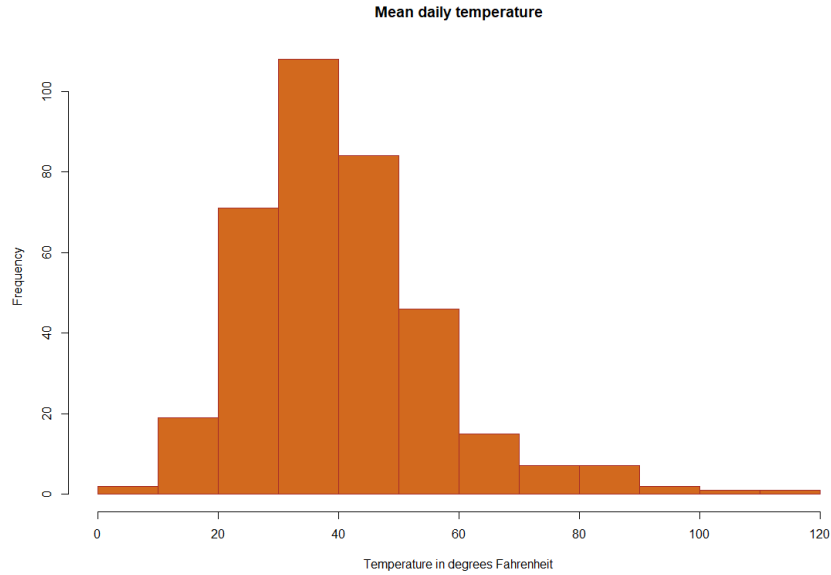

Mean daily temperature

Figure 2: Mean Daily Temperature's Histogram using Freedman-Diaconis Binning

Problem 3

The objective is to compute the 1st ,5th, 10th, 25th, 50th, 75th, 90th, 95th, 99th percentile of daily maximum values of the oxides of nitrogen. The daily maximum value is aggregated into daily format from hourly data of oxides of nitrogen.

1%	5%	10%	25%	50%	75%	90%	95%	99%
3.90	7.70	9.90	15.20	22.25	36.50	62.10	87.60	151.50

Table 1: Percentile Table for Daily Maximum Values of the Oxides of Nitrogen

Problem 4

The objective is to create a correlation plot between Daily max Ozone, Daily min Dew Point Temperature,Maximum daily solar radiation, Maximum daily wind speed and Daily maximum concentration of oxides of nitrogen.
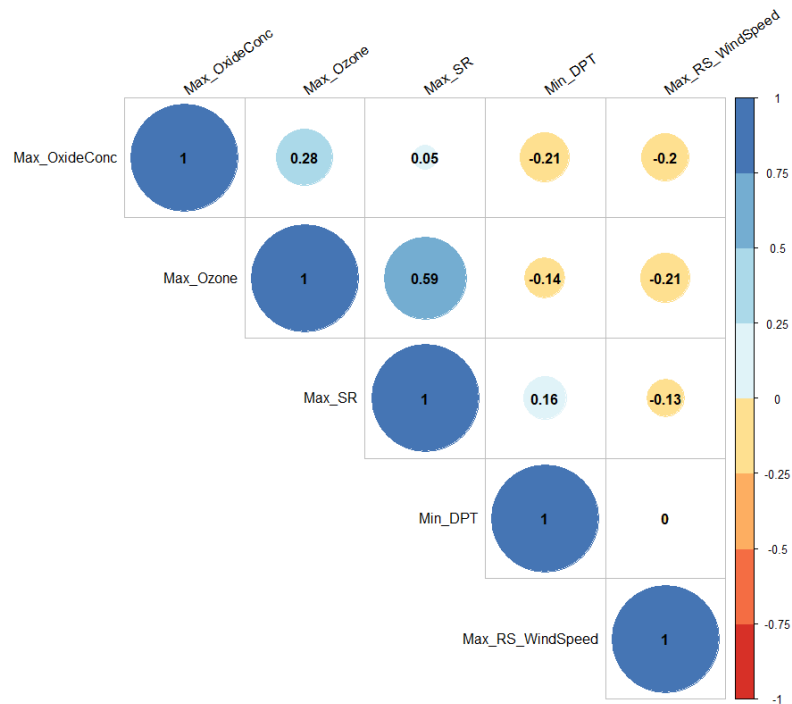

Figure 3: Correlation plot using Spearman Rank Correlation Coefficient

Figure label Description
Max_OxideConc - Daily Maximum Concentration of Oxides of Nitrogen
Max_Ozone - Daily maximum Ozone Concentration
Max_SR - Maximum Daily Solar Radiation
Min_DPT - Daily Minimum Dew Point Temperature
Max_RS_WindSpeed - Maximum Daily Resultant Wind Speed

Problem 5

The objective is to use 90 degree Quadrants of Wind Direction and compute the probability of Ozone being less than or equal to 10 ppm or higher by creating a contingency table showing joint, marginal probabilities. Second, we will also compute the conditional probability for Ozone concentration being greater than 10ppb AND Resultant Wind Direction with 0 to 90 degrees.

$$P(Ozone>10 \mid Wind\ Direction = NE(0 - 90]).$$

Here, Ozone (POC 3) is measured in parts per billion and Resultant Wind Direction (POC 1) is measured in degrees compass with North as 0 degrees and South as 180 degrees.

First, the count of all combined occurrences of Ozone level and Wind Direction Quadrant is arranged in the following table.

Ozone / WD	NE	SE	SW	NW
<10	303	539	342	383
>= 10	1611	3191	942	1084

Table 2: Count of all combined occurrences of Ozone level and Wind Direction.

Contingency table showing all the joint and marginal probabilities is listed on following table.

Ozone / WD	<10 ppb	>= 10ppb	Total
0-90 Quad	0.03609291	0.19189994	**0.22799285**
90-180 Quad	0.06420488	0.38010721	**0.44431209**
180-270 Quad	0.04073853	0.11220965	**0.15294818**
270-360 Quad	0.04562239	0.12912448	**0.17474687**
Total	0.1866587	0.8133413	1

Table 3: Contingency table showing all joint and marginal probabilities.