Perform the following written operations:

1. Read the file in Zip format and get it into R.

data <- read.table("airquality.zip", nrows=10, header=T, quote="\"", sep=",")

2. Create Univariate for all the columns.

describe(airquality)

vars n mean sd median trimmed mad min max range

Ozone 1 116 42.13 32.99 31.5 37.80 25.95 1.0 168.0 167

Solar.R 2 146 185.93 90.06 205.0 190.34 98.59 7.0 334.0 327

Wind 3 153 9.96 3.52 9.7 9.87 3.41 1.7 20.7 19

Temp 4 153 77.88 9.47 79.0 78.28 8.90 56.0 97.0 41

Month 5 153 6.99 1.42 7.0 6.99 1.48 5.0 9.0 4

Day 6 153 15.80 8.86 16.0 15.80 11.86 1.0 31.0 30

skew kurtosis se

Ozone 1.21 1.11 3.06

Solar.R -0.42 -1.00 7.45

Wind 0.34 0.03 0.28

Temp -0.37 -0.46 0.77

Month 0.00 -1.32 0.11

Day 0.00 -1.22 0.72

1. Check for missing values in all columns.

colSums(is.na(AirQualityUCI))

install.packages("mice")

library(mice)

1. Impute the missing values using appropriate methods.

5. Create bi-variate analysis for all relationships.

6. Test relevant hypothesis for valid relations.

7. Create cross tabulations with derived variables.

8. Check for trends and patterns in time series.

9. Find out the most polluted time of the day and the name of the chemical compound.