1. i) Write a function called kelvin_to_celsius() that takes a temperature in Kelvin and returns that temperature in Celsius (**Hint:** To convert from Kelvin to Celsius you subtract 273.15)
(ii) Write suitable R code to compute the mean, median ,mode of the following values c(90, 50, 70, 80, 70, 60, 20, 30, 80, 90, 20)
(iii) Write R code to find 2nd highest and 3rd Lowest value of above problem.

2. Explore the airquality dataset. It contains daily air quality measurements from New York during a period of five months:
• Ozone: mean ozone concentration (ppb),
• Solar.R: solar radiation (Langley),
• Wind: average wind speed (mph),
• Temp: maximum daily temperature in degrees Fahrenheit,
• Month: numeric month (May=5, June=6, and so on),
• Day: numeric day of the month (1-31).

i. Compute the mean temperature(don't use build in function)

ii.Extract the first five rows from airquality.

iii.Extract all columns from airquality except Temp and Wind

iv.Which was the coldest day during the period?

v.How many days was the wind speed greater than 17 mph?

3. (i)Get the Summary Statistics of air quality dataset

(ii)Melt airquality data set and display as a long – format data?

(iii)Melt airquality data and specify month and day to be "ID variables"?

(iv)Cast the molten airquality data set with
respect to month and date features

(v) Use cast function appropriately and compute the average of Ozone, Solar.R , Wind and temperature per month?

4.(i) Find any missing values(na) in features and drop the missing values if its less than 10% else replace that with mean of that feature.

(ii) Apply a linear regression algorithm
using Least Squares Method on “Ozone” and “Solar.R”

(iii)Plot Scatter plot between Ozone and
Solar and add regression line created by above model