manipalglobal
Academy of Data Science

OPERATIONS ON DATA FRAMES

Introduction

This assignment would help to understand on how to use built-in datasets in R and manipulate them using loops and functions. A dataset named “airquality” is built into the R-studio. It consists of data about daily air quality measurements in New York from May to September 1973. The data set is a data frame with 154 observations on six variables. The problem statement is to find the maximum wind speed for each month on the data set “airquality”, and return a new data frame with the computed values. Since the data is collected only from month 5 to month 9, every month will just have one day for maximum wind speed to be calculated.

Objective

Demonstrate the ability to manipulate the data frames.

Self-Assessment Questions

Steps involved:

a) Create a function “windmax” which takes a data frame as an argument

b) Create an empty vector “a” to populate it in the following steps

c) Using the subset function compute the maximum wind speed for each month (month 5 to 9) and assign it to a variable calc_speed

d) Populate the empty vector “a” with calc_speed using the rbind function

e) Call the function “windmax” by passing the airquality data frame as an argument