Analytics with *R*: Assignment-1

1. How to get a description of the version of R and its attached packages used in the current session?
2. Create object called “abc” by assigning the number 3.
3. Create different vectors (“a” is numeric, “b” is character and “c” is logical).
4. List all the objects in the current session.
5. Create vector “x” with the values {4, 4, 5, 6, 7, 2, 9}.
   a. Calculate the number of observation (n), mean, sum, max, min, variance of the vector ‘x”.
   b. Also print the 3rd element, odd positions elements and elements from 2 to 6 positions.
6. Create 6x4 matrix(6 rows and 4 columns) using 1 to 24 numbers.
7. Create data frame with the below vectors
   a. StoreID - (111, 208, 113, 408)
   b. Tenure - (25, 34, 28, 52)
   c. StoreType - ("Type1", "Type2", "Type1", "Type1")
   d. status - ("Poor", "Improved", "Excellent", "Poor")
8. Print the data in different programs.
   a. only storeId, tenure
   b. only storetype and status
   c. only tenure
9. Create different factors using the below vectors and label the values and order the levels for factors “outcome” and “status”.
   a. ethnicity - White", "African amrican", "White", "Asian"
   b. status - Poor", "Improved", "Excellent", "Poor"
   c. outcome - c(1, 3, 2, 4, 3, 1, 1) – labels- “Poor", "Average", "Good", "Excellent"

10. Create list called “mylist” with the title "My First List" and with the below objects called “ages” with below h, j and k. Also print the different combinations of objects.
    a. h –numeric vector with the values 25, 26, 18, 39
    b. j – matrix with 5 rows and 2 columns with the values 1 to 10
    c. k – character vector with the values "one", "two", "three"
11. Read the stores.csv data set and find the summary statistics for all the columns.
12. Use with() function, calculate the summary of operatingcost in the stores.csv data set. What is the difference?
13. Apply below functions on stores data frame and observe the outputs.
    a. class(stores)
    b. names(stores)
    c. length(stores)
    d. dim(stores)
    e. str(stores)
    f. head(stores)
    g. tail(stores)
    h. fix(stores)
14. Create new variable by OperatingCost + AcqCostPercust using different methods
    a. Simple calculation and assignment operators
    b. Using transform function
15. Create new variable store class as follows.
    a. If total sales<120 then “Low Perform store”
    b. If total sales>=120 and total sales<240 then “Average Perform store”
    c. If total sales>240 then “High Perform store”
16. Rename the variable “AcqCostPercust” with “AcqCost”

17. How to find missing values in the stores dataset and recode missing values with 0? How to delete the missing values (if there are any) from the data set?
18. Sort the stores data as follows
    a. Create new dataset (newstores) with sorted data by "Storetype"
    b. Create new dataset (newstores) with sorted data by location(ascending) and totalsales in descending order.
19. Create vectors with date values with date formats.
    a. "2014-06-22", "2014-02-13"
    b. "01/05/1965", "08/16/1975"
20. Create subsets of data from stores data as following conditions.
    a. Only Columns 5,7, 8,9
    b. Excluding columns 5, 7, 8,9
    c. Selecting first 10 observations
    d. Observations with Storetype= Apparel and totalsales>100
    e. Columns (storecode, storename, location and totalsales) with totalsales between 100 and 300.
    f. Include all columns from Storecode to Basketsize with storetype=Electronics and totalsale>100