R BYOD CA1 QUESTIONS

CASE1: Consider a scenario where CALL CENTER data needs to be stored, having following attributes:

[Caller Phone Number, Caller Name, Caller Location and Call Duration]

TASK1: Create the appropriate data structure to save the required data. Call center got five calls where

1st Caller Name is ANIKET, talking from Delhi, his phone is 9041 and had called for 9 mins.

2nd Caller is from Chandigarh calling from 9888 number for 15 mins and telling her name as MANISHA

3rd Caller is also from DELHI and was talking for continuous 30 mins, his registered mobile number and name is 9646 and RAMAN respectively.

4th Caller mentioned himself as RAJNIKANT and explained his problem in 20secs from 7777 number from AP

5th Caller drops his call just after calling his number saved as 8899.

TASK2: After the calls and storing the data, later we feel to add attribute GENDER and PROBLEM PRIORTY in dataset, so add GENDER and PROBLEM PRIORTY [CRITICAL, NORMAL, MODERATE] in the dataset with appropriate data type and levels.

TASK3: After 10 mins 3rd Caller again calls and gets his name rectified as DAMAN.

TASK4: Team leader found problem in accessing data so he had given a requirement of saving Every Caller data with number as CALLER1, CALLER2, CALLER3, CALLER4 and CALLER5 without making it as an attribute.

CASE2: Consider you are analyzing MEDICAL COLLEGE DATA and you have to analyze between the patient personal data, his/her previous medical history and parent's medical history. You found that data is not tidy and needs to be stored in a correct manner again.

TASK: Store the data in data structure in such a way that when you enter PATIENT ID then three further tables PATIENT BIODATA, MEDICAL HISTORY, PARENTS MEDICAL HISTORY for that patient will be accessed directly. Store the data for 2 patients having PATIENT ID as [P101 and P105]

Attributes for the Patients in Their tables:

[PATIENT BIODATA: Patient Name, Patient ID, Patient Gender, Patient Age, Patient Job, Patient Job hours]

[PATIENT MEDICAL HISTORY: Patient ID, Blood Group, Blood Sugar Level, Blood Pressure, Heart Beat, ESI, RBC Count, WBC Count]

[PARENTS'S MEDICAL HISTORY: Patient ID, Father Name, Blood Group, Avg. Blood Sugar, Avg. Blood Pressure]

(Store data in these tables for P101 and P105 according to your choice.)

R BYOD CA1 QUESTIONS

CASE3: Analyst have been given the dataset of Titanic to analyze the survival rate of passengers but before working on it, analyst have felt that data is separated among 3 files.

TASK1: Analyst want to work on single file having all 12 Variables and 1309 Observations.

TASK2: Data before analyzing needs to be cleaned so clean the data as required.

(HINT: Data structure, Data Type, Data Level, Variable Name, NA values, Repetition and Length of data needs to be considered)

TASK3: Show the brief summary of data.

CASE4: We have mtcars dataset which needs to be further improvised and some new columns are required to add in dataset by giving dataset a new name as NEWMTCARS.

TASK: A new factor is to be stored named as MILEAGE in which data will be either LOW, MEDIUM, GOOD. The cars whose mpg below 20 will be saved as LOW, for cars mpg below 25 and above 20 will comes under MEDIUM and for those having mpg above 25 are GOOD.

CASE5: PVR Cinemas wants to save their boxoffice collection where matrix columns will be defined by movie names and matrix rows will be defined by movie sequels.
[Dhoom1, Dhoom 2, and Dhoom 3 – Box office collection: 20cr, 40cr, 80cr respectively]
[Golmaal 1, Golmaal 2 and Golmaal 3 – Box office collection: 50cr, 65cr, 90cr respectively]
[Housefull 1, Housefull 2, Housefull 3 –Box office collection: 30cr, 65cr,45cr respectively]

TASK1: Create a matrix for this data and analyze.
TASK2: The average collection for movie.
TASK3: Percentage increase/decrease in collection after each sequel.