BYOD CA1

All questions are compulsory
Upload your .R script file in a zipped folder on UMS
Time allowed: 2 Hours

ROLL NO: **REG NO:**

Q.1 Consider you are working as a DATA ANALYST for Cancer hospital, and you have been given the BREAST CANCER DATASET for predicting that weather patient is suffered or not.

For reference take the dataset: BREAST CANCER [https://www.tinyurl.com/examdatasets]

TASK: While working on DATASET, you came to know that it has outliers in some attributes. So before proceeding towards creating MODEL, try to handle outliers and write the R script for it.

Q.2 Predict the school of painters from Painters dataset (Mass package).

For reference take the dataset: Painters [Use Package MASS in R Studio]

TASK: Use an appropriate algorithm to predict SCHOOL of painters if attribute values are

Composition = 12 **Drawing = 13** **Color = 9** **Expression = 8**

Q.3 You have been given the dataset about some lung's patients having following attributes:

LUNGCAPDATA (LungCap, Age, Height, Smoke, Gender, Caesarean)

- LungCap: It's the lung capacity of the person
- Age: It's how old is the person
- Height: It's how tall is the person
- Smoke: If the person smokes or doesn't smoke
- Gender: If are male or female
- Caesarean: If they're born by caesarean

For reference take the dataset: LUNGCAPDATA [https://www.tinyurl.com/examdatasets]

TASK: You have to create prediction model for following:

i. Handle NAs
ii. Handle Outliers
iii. All attributes need to be in same range of values.
iv. While entering Patient LungCAP capacity, Age, Height, Smoke and Gender model will predict weather person born was Caesarean or not.