**DE Lab Programs for A.Y: 2022-23**

1. Write a high level program for the following:

Suppose that the data for analysis includes the attribute age. The age values for the data tuples are (in increasing order)

13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 36, 40, 45, 52, 70.

1. What is the mean of the data? What is median?
2. What is the mode of the data? Comment on the data’s modality (bimodal, trimodal, etc..)
3. What is mid-range of the data?
4. Can you find the first quartile(Q1) and the third quartile (Q3) of the data?
5. Give the five number summary of the data.
6. Write a C program to calculate the *correlation coefficient. Use the following data to check your code.*

Suppose a hospital tested the age and body fat data for 18 randomly selected adults with the following result:



Are these two variables positively or negatively correlated?

1. Suppose that the data for analysis includes the attribute *age*. The *age* values for the data tuples are (in increasing order) 13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70. Write a C program to implement *smoothing by bin means* to smooth the data, using a bin depth of 3.
2. Write a C program to implement:

(a) min-max normalization

(b) z-score normalization

(c) Normalization by decimal scaling.

1. Write a program to find frequent item sets using apriori algorithm for the following transactional database. Let min sup = 60% and min conf = 80%



1. Write high level language programs to implement classification techniques.
2. Write high level language programs to implement clustering techniques.