

## Assignment to Introduction to Data Science

Deadline: Tuesday, 9 May 2023

Question 1: K-Means / 10 Marks

A fitness club would like to follow the evolution of its members at given time intervals. To do so, it would like to create two clusters (K=2) of its members using the K-means algorithm.

Members	Height (Cm)	Weight (Kg)
Pierre	185	72
Jetheme	170	56
Paul	168	60
José	179	68
Alain	182	72
Claire	188	77
Ketsia	180	71
Jean	180	70
Bosco	183	84
Meira	180	88
Luc	180	67
Anne	177	76

**Required:** Create the two clusters taking Jetheme and Bosco as the initial centers.

## Question 2: Apriori / 10 Marks

Given the following dataset:

Transaction ID	Items
T1	Sugar, Bread, Butter
T2	Sugar, Chips
T3	Bread, Sugar, Cake
T4	Cake, Sugar, Bread, Yogurt
T5	Butter, Bread

**Required:** Use the Apriori Algorithm to discover strong association rules among the items given a minimum support of 40% and a confidence level of 70%