SMDE FIRST ASSIGNMENT (20% OF THE FINAL MARK, INDIVIDUAL)

FIRST QUESTION: VISUALISATION, CHI SQUARE AND T-TEST (15% OF THE FIRST ASSIGNMENT).

Download the data set decathlon from the web site of Kaggle given in the link below and read it in R.
https://www.kaggle.com/drisskaouthar/decathlon#decathlon.csv

a) Analyze the distribution of "X100m" according to the type of competition by using boxplot. Write your conclusion.

b) Create a new categorical variable with two categories from the variable "X100m" by using 11 seconds as the cut-off point. Make a cross table from the new categorical variable and the "Competition". Are these two variables independent? Write your conclusion by checking marginal probabilities and test the independency of two variables by using Chi-Square test.

c) Visualize the distribution of quantitative variables by using proper graph. Which of these variables follow a Normal distribution?

d) Generate three Normally distributed random variables of length 50. Two of them should have the same mean, different standard deviations while the third one has a different mean but the same standard deviation with the first distribution. Use t test to compare mean differences between three variables.

e) Test if there is a difference between two type of competitions according to the variables "X100m" and "X400m" by using t test.