SMDE FIRST ASSIGNMENT (40% OF THE FINAL MARK)

FIRST QUESTION: ANOVA (25% OF THE FIRST ASSIGNMENT).

Download the laptop prices data set from the web site of kaggle (Laptop Price Prediction Cleaned Dataset) and read the data set in R.

The data set consists of following variables: Company, Type Name, Ram, Weight, Price, Touch Screen indicator, IPS indicator, Ppi, Cpu_brand, HDD, SSD, Gpu_brand, Os (Operation System).

a) Read the laptop price data set and create a sub dataset including only laptop brands "Dell", "Acer" and "Hp". Summarize the variable "company" and checked the overall distribution of Price and Weight for this subset. (10p)

b) The objective is to analyze the relationship between the brand of the computer and its price and its weight. First test the assumptions of the statistical method by using corresponding test and plots. Write your conclusions about the assumptions. If there is any violated assumption, interpret what should be done? Are both variables suitable for the corresponding analysis? If not, explain why? (30p)

c) After fulfilling the assumptions, apply the related statistical method and interpret your findings. Does the brand of the computer have significant effect on its price and its weight? (just consider the variable(s) that fulfill the assumptions tested in section (b)) (35p)

d) Analyze the effect of brand and touch screen characteristics together on the price. Analyze whether the interaction of two term is significant. Interpret your findings. (Do not forget to confirm the assumptions!) (25 p)

NOTE: Do not forget to do multiple comparisons! Apply post hoc tests to see where the differences source from. Apply three different post hoc tests and compare their findings.