**Recoded Dataset Explaination:**

**Gender**: Male=0, Female=1

**Family history of obesity**: No=0, Yes=1

**Frequency of Consumption of Vegetables(FAVC)**: No=0, Yes=1

**Consumption of Food between meals(CAEC)**: No=0, Sometimes=1, Frequently=2, Always=3

**Consumption of Alcohol(CALC)**: No=0, Sometimes = 1, Frequently=2, Always=3

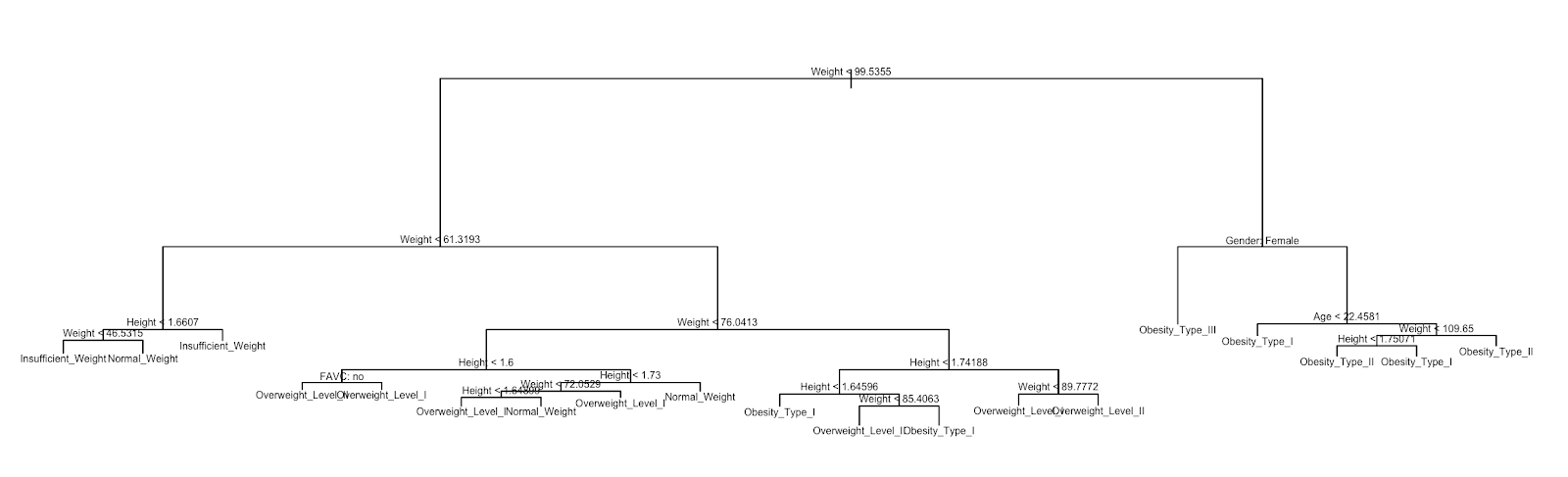
**Smoke**: No=0, Yes=1

**Calories consumption monitoring(SCC)**: No=0, Yes=1

**Transportation used(MTrans):** Walking=0, Public transportation=1, bike=2, motorbike=3, automobile=4

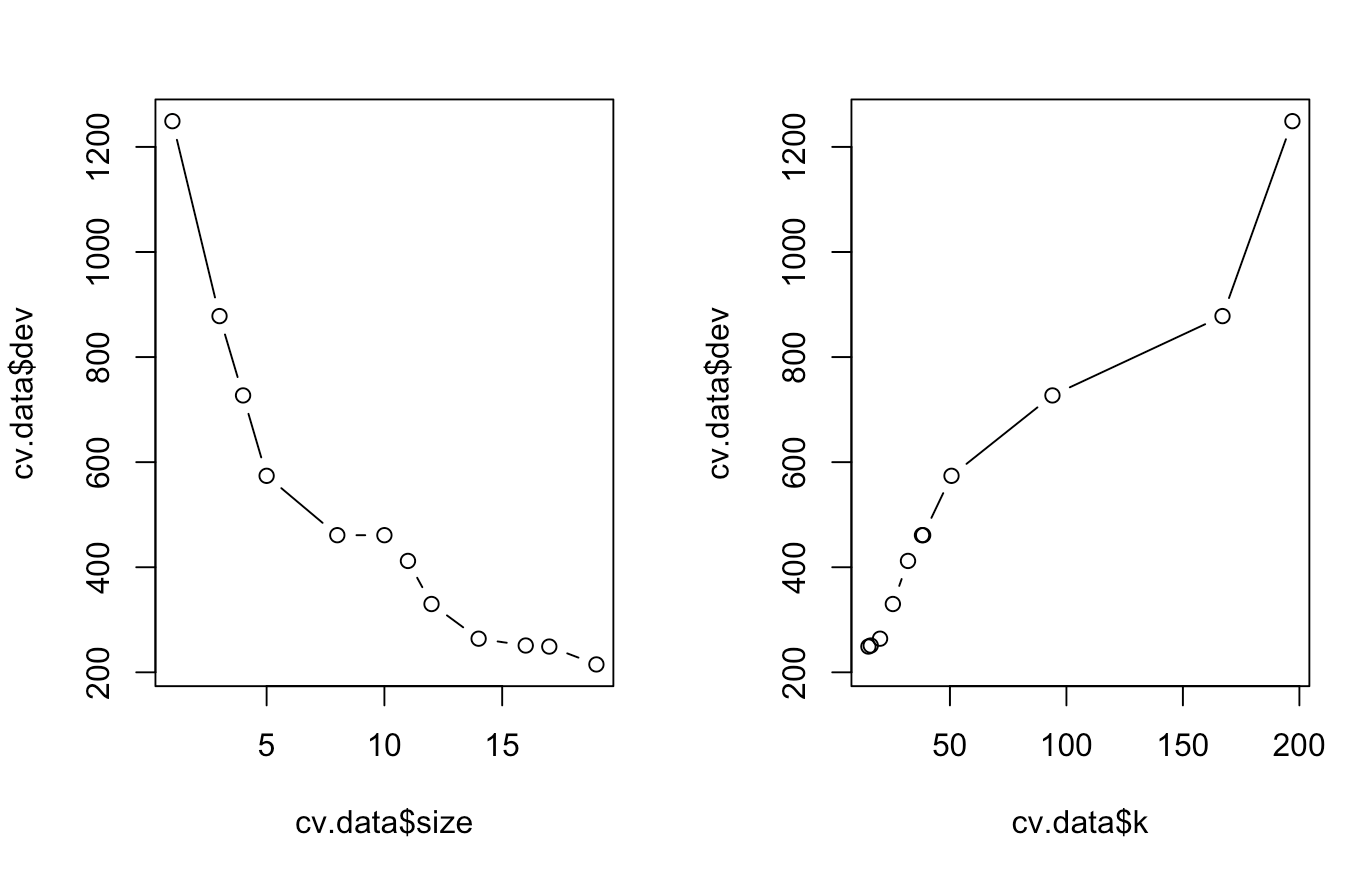
**BMI Level(Nobesity):** Underweight=0, normal weight =1, overweight level I=2, overweight level II=3, obesity type I=4, obesity type II = 5, obesity type III = 6

**Classification Tree Diagram:**

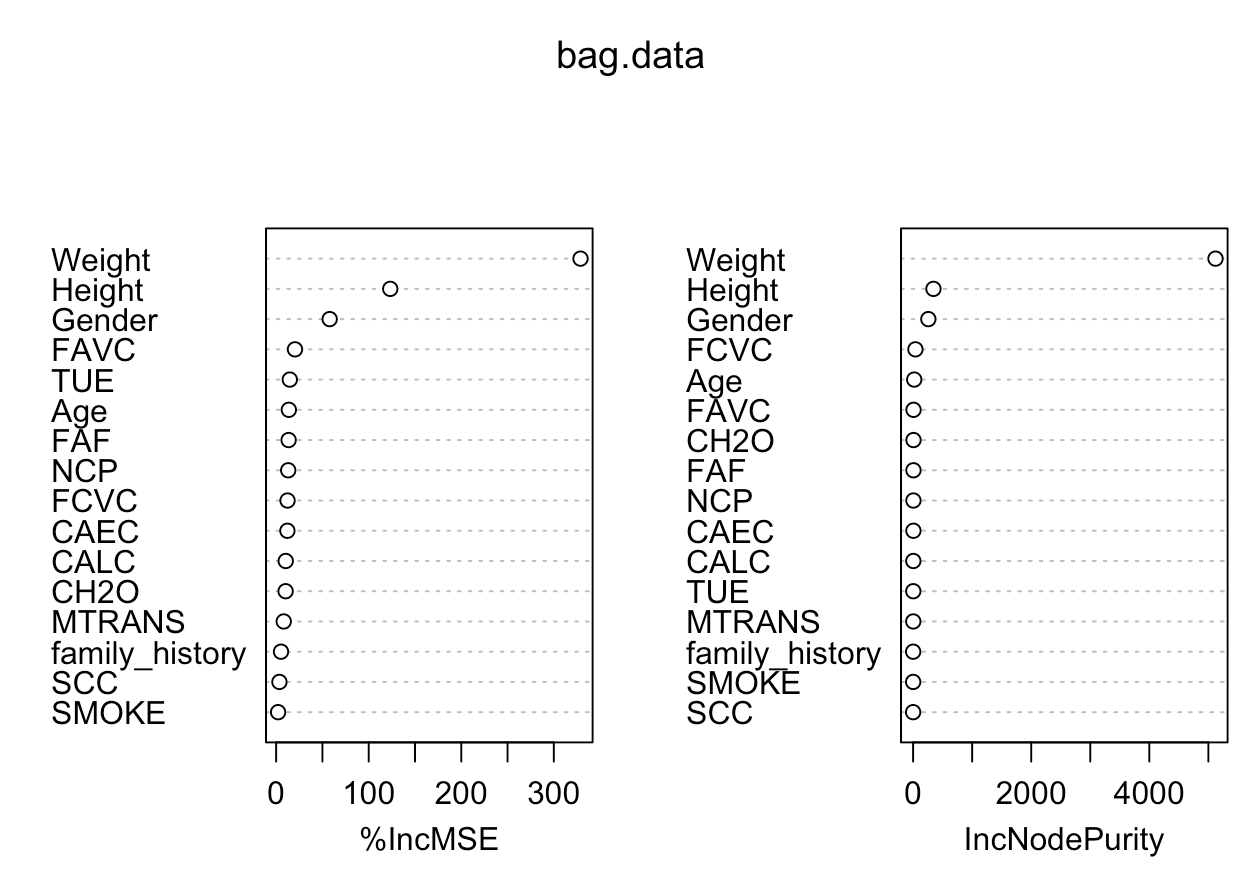


**Classification Tree Node Selection:**

# of terminal nodes= 19

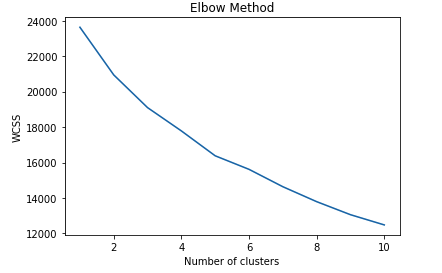


**Random Forest Importance Plot:**

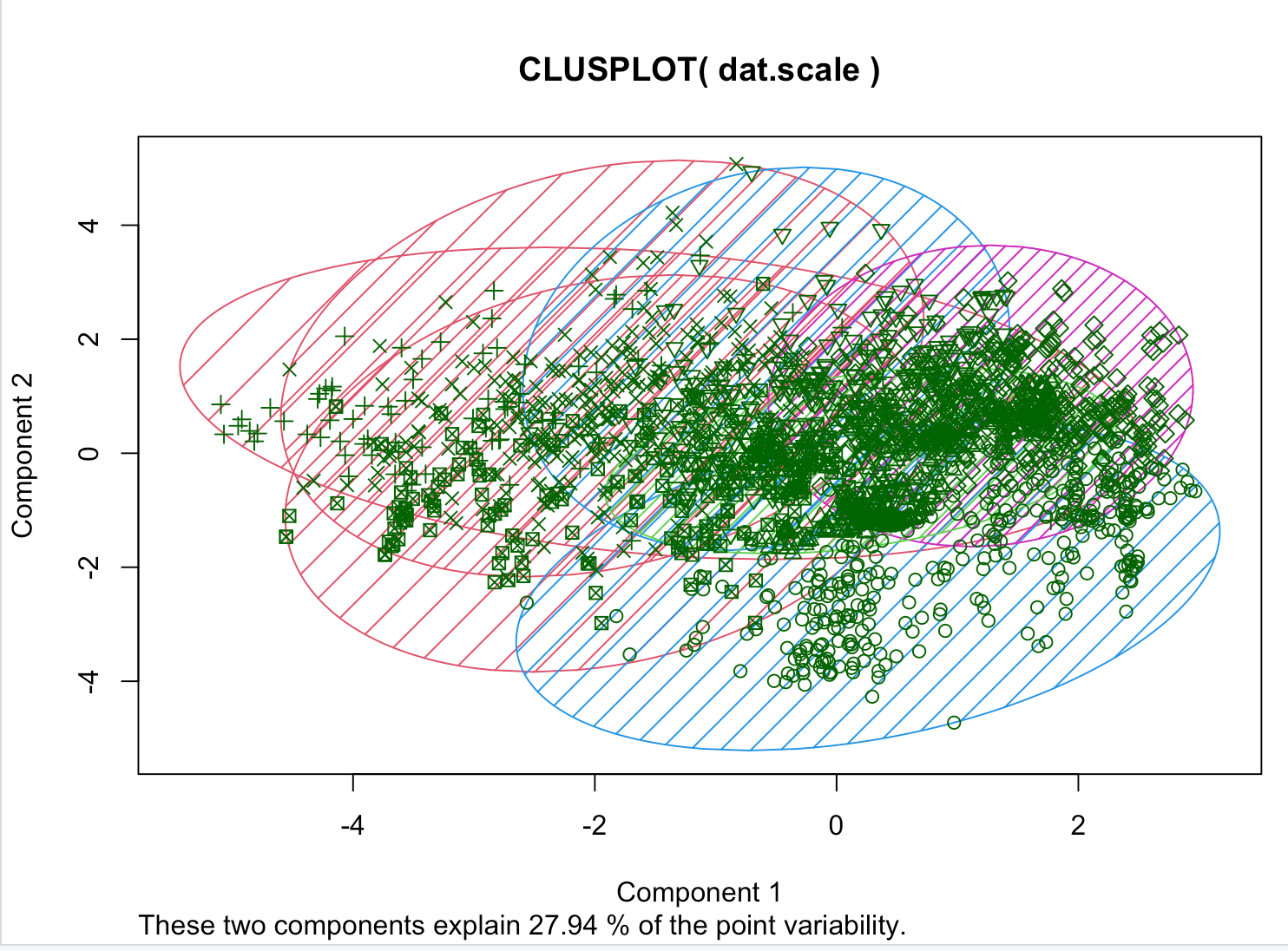


**Elbow Point for Clustering:**

Since no obvious elbow point, we decided to use k=7, which is the number of the obesity levels.



**K-Means Clustering Diagram**



**Frequency and Normality Check:**

