**Lab 8**

Download the Allstate Claims Severity training data file from kaggle.

1. Do hierarchical clustering of the samples using the numeric features, and retrieve the two largest clusters. Then for each of these two clusters, find out how many samples with loss > the median loss; and how many with loss < the median loss.
2. Do kmeans clustering of the samples using the numeric features, by setting k=2. Then for each of these two clusters, find out how many samples with loss > the median loss; and how many with loss < the median loss.
3. Do PCA of the data using the numeric features and use the first two PCs to represent a sample and re-do Question 1 and Question 2. Also generate barplot of the variance of the first five PCs.

Hint: use prcomp function in R