CS 434 Assignment 4

Members:

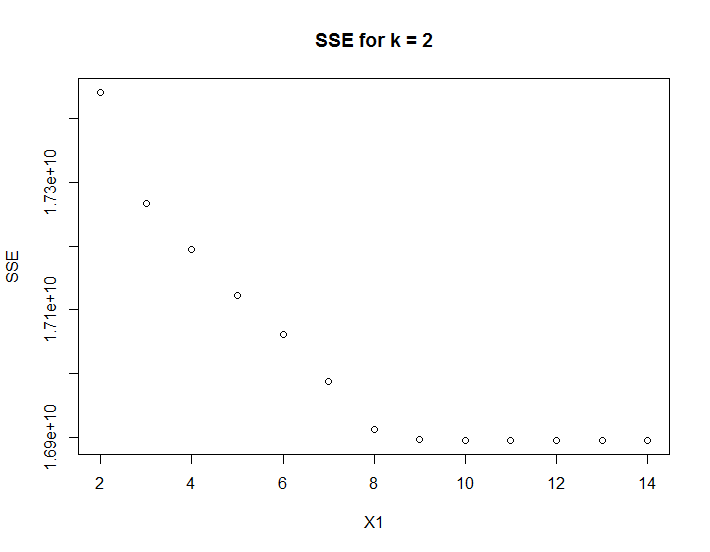
Alessandro Lim

Bryan Liaw

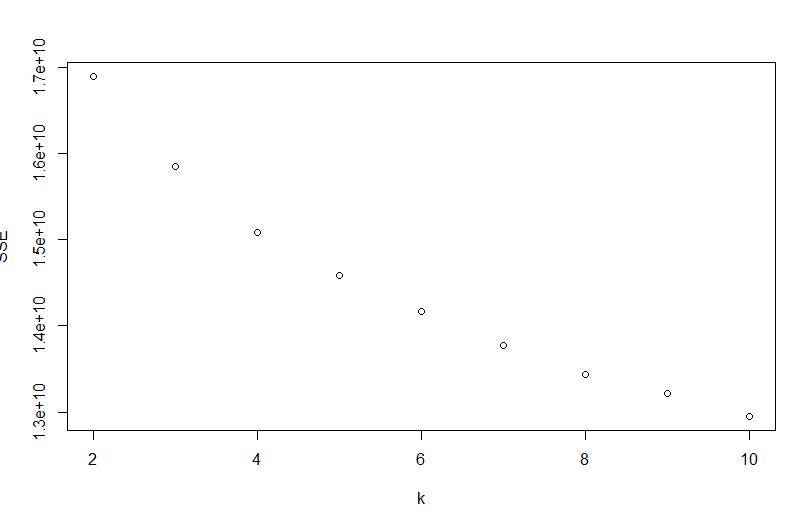
Songjian Luan

**Kmeans**

1. The SSE of a typical kmeans with k=2 (note, we only plotted the SSE 14 times)



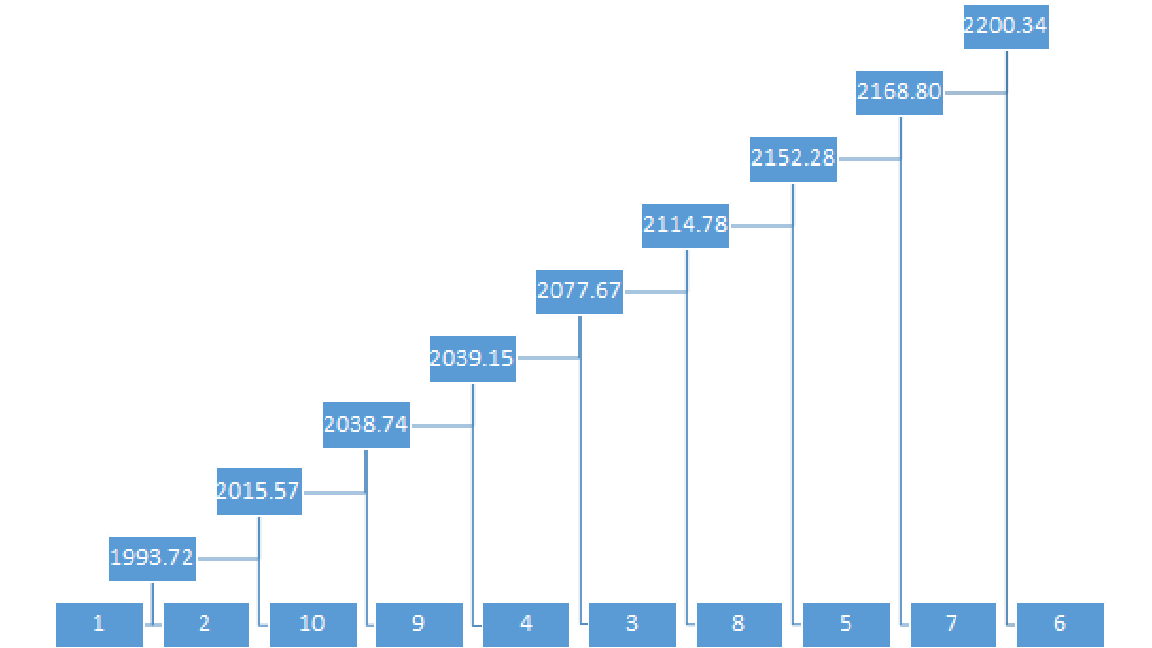
1. SSE of k from 2 to 10:



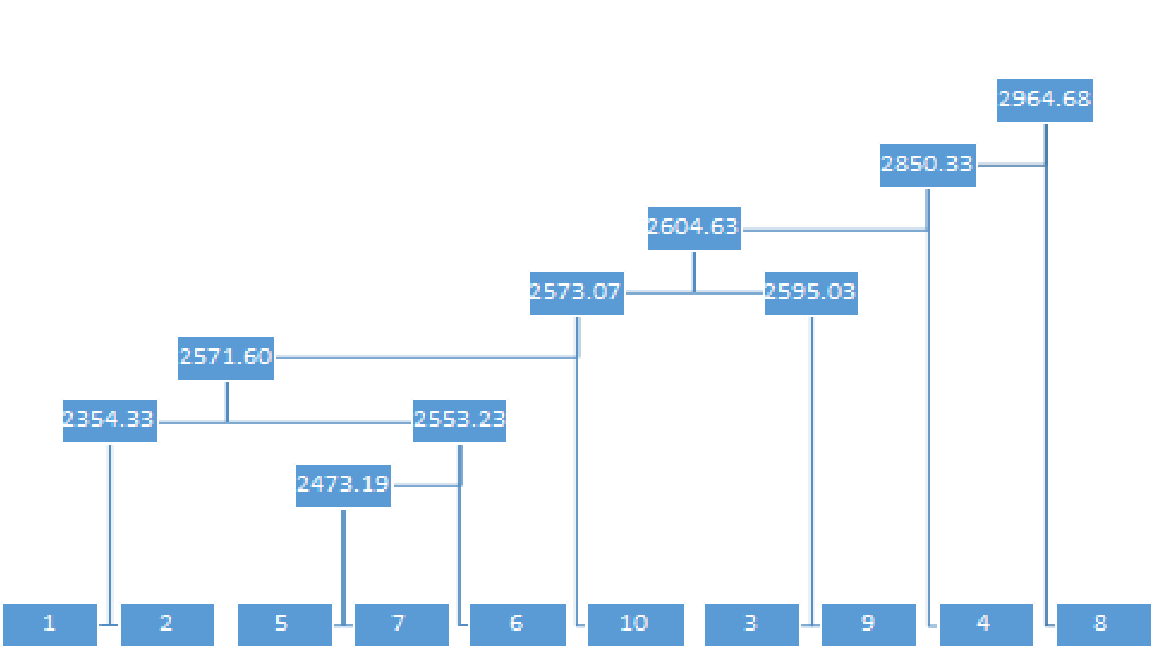
The best K to use seems to be K = 8, since that since to be the “knee” as the SSE plot des not change that much anymore.

**HAC**

1. Single-linked HAC Dendogram:



The point where we separate the cluster is between the height 2039 and 2077, since the height difference is the most extreme in this case.

1. Complete linked HAC:

The best point to separate the cluster is between the cluster with the height 2604 and 2850, since the height difference is the most extreme in this case.