

# STAT 530/J530

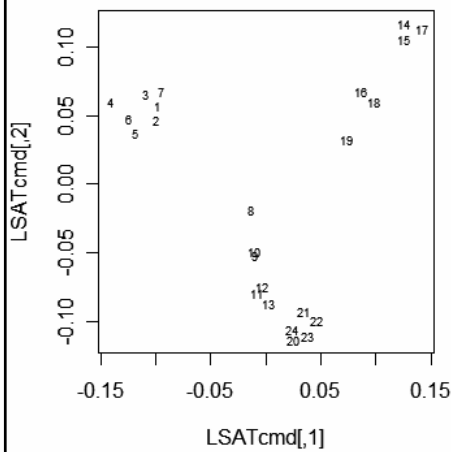
## October 25th, 2005

Instructor: Brian Habing  
Department of Statistics  
LeConte 203  
Telephone: 803-777-3578  
E-mail: habing@stat.sc.edu

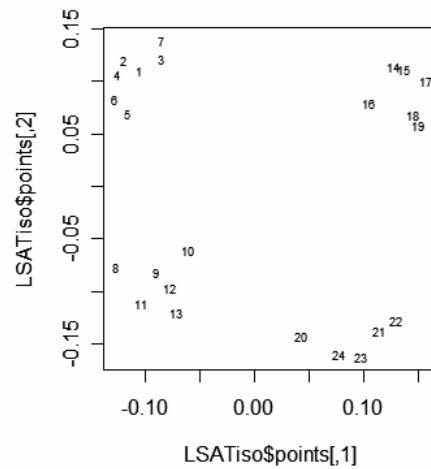


## Four Paragraph Test

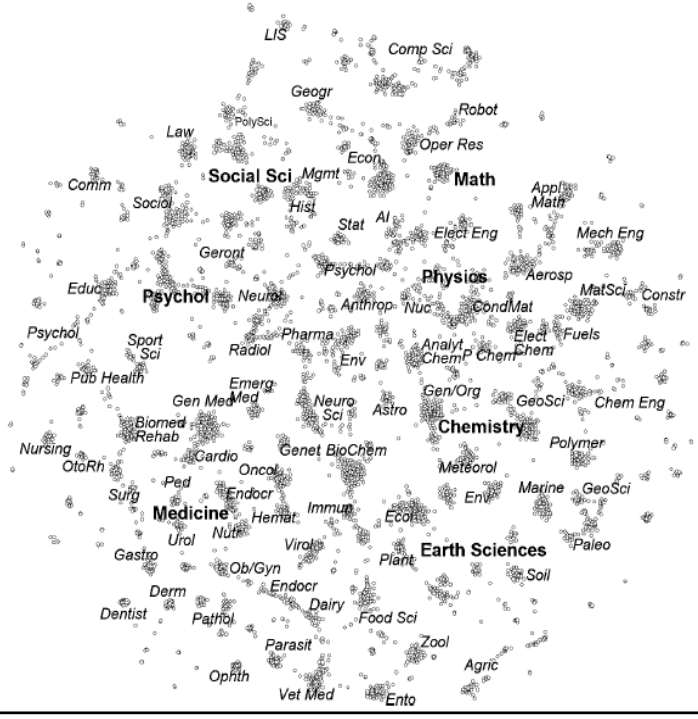
**Classical**



**Isometric**



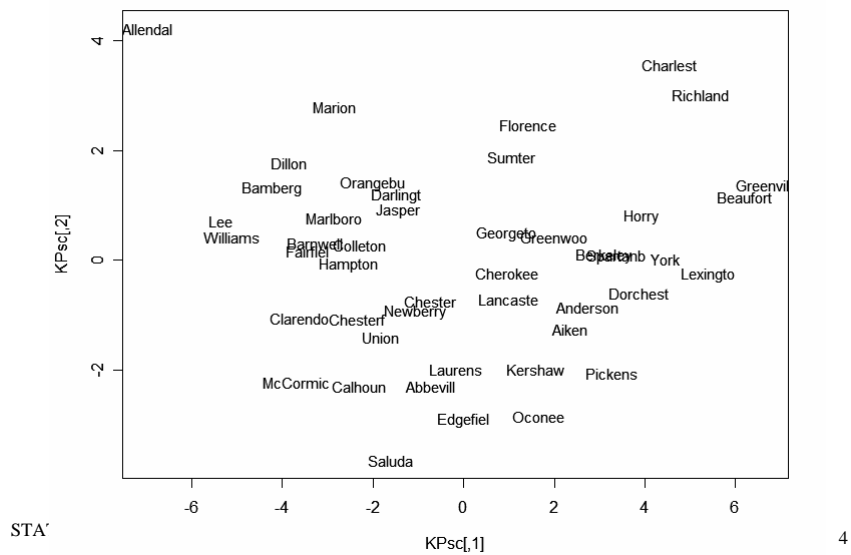
Boyack, K.W.,  
 Klavans, R., &  
 Borner, K.  
 (2005).  
 Mapping the  
 backbone of  
 science.  
*Scientometrics*,  
 64, 351-374.



STAT 530/J530 B.Hat

## South Carolina Counties

Karl Pearson Distance



## Cluster Analysis

Like MDS, cluster analysis uses distances to help visualize the data. The goal of cluster analysis is to group the items into related groups or clusters, instead of attempting to form an entire map.



## Agglomerative Hierarchical Clustering

- (0) Start with each item as a separate cluster.
- (1) Identify the pair of clusters with the smallest distance between them and merge them. This reduces the number of clusters by 1.
- (2) Repeat step 1 until all of the items are in a single cluster.



## Distance Between Clusters?

- Complete Linkage, Farthest Neighbor, Compact
- Single Linkage, Nearest Neighbor, Connected
- Average (Mean of the Distances)
- Centroid (Distance of the Means)
- Ward's Method is related to the sum of squares in ANOVA

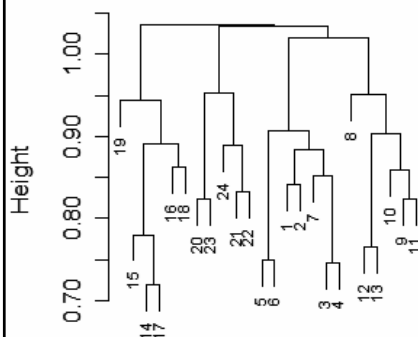
STAT 530/J530 B.Habing Univ. of S.C.



7

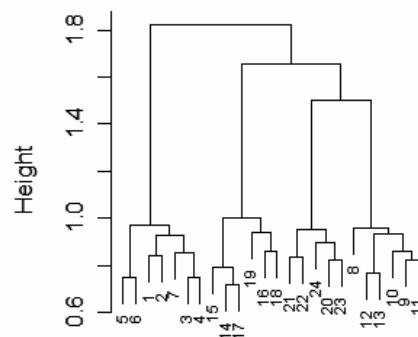
## Four Paragraph Test

Cluster Dendrogram



as.dist(-1 \* cc + 1)  
hclust(\*, "complete")  
STAT 530/J530 B.Habing Univ. of S.C.

Cluster Dendrogram



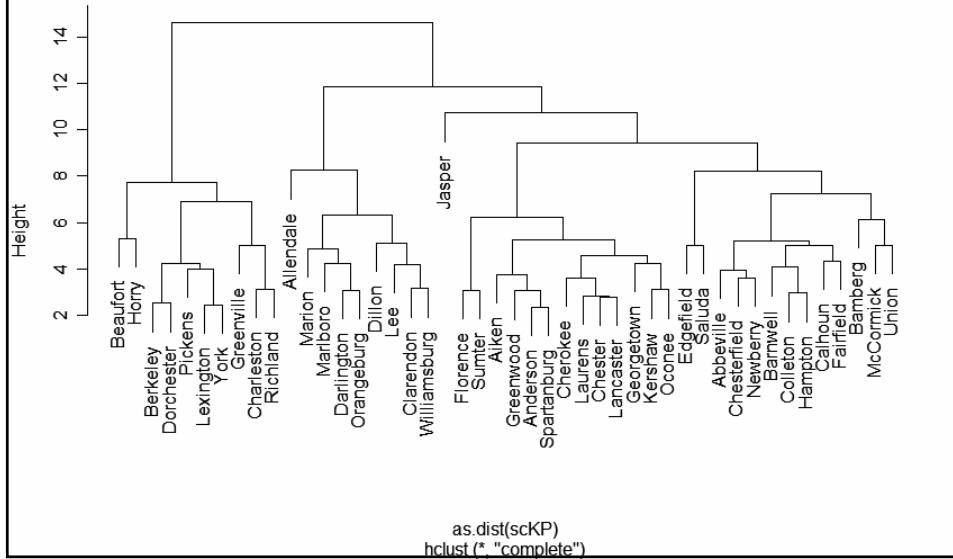
as.dist(-1 \* cc + 1)  
hclust(\*, "ward")



8

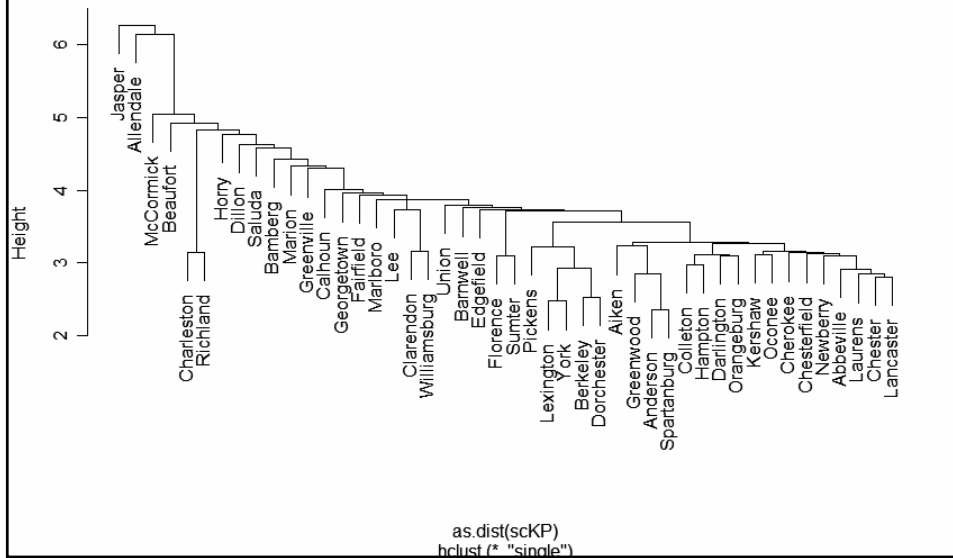
# South Carolina Counties

Cluster Dendrogram



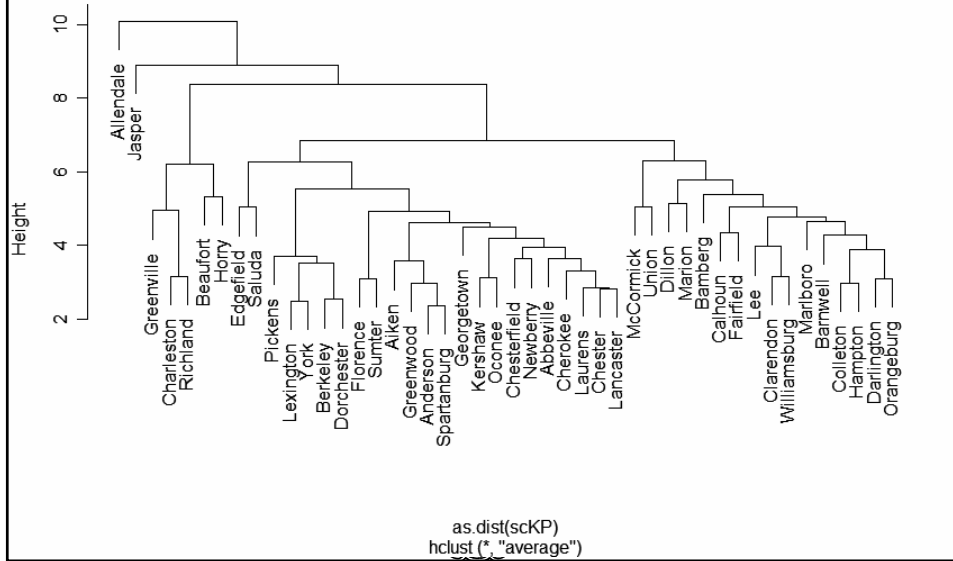
# South Carolina Counties

Cluster Dendrogram



# South Carolina Counties

Cluster Dendrogram



# South Carolina Counties

Cluster Dendrogram

