Mini Project #2(a)

Goal: practice the basic visualization tools used in visual analytics

- use data from mini project #1 (or other)
- client-server system: python for processing (server), D3 for VIS (client)
- non-CS students can use plotly Dash for VIS & python or R to compute

Task 1: basic dimension reduction and data visualization with PCA

- use PCA to compute the Eigenvectors of the data and visualize the Eigenvalues as a scree plot
- add an interaction element into the scree plot that allows the user to mark and select the intrinsic dimensionality index (d_i)
- plot the data into a PCA-based biplot

Task 2: visualize the data with a scatterplot matrix

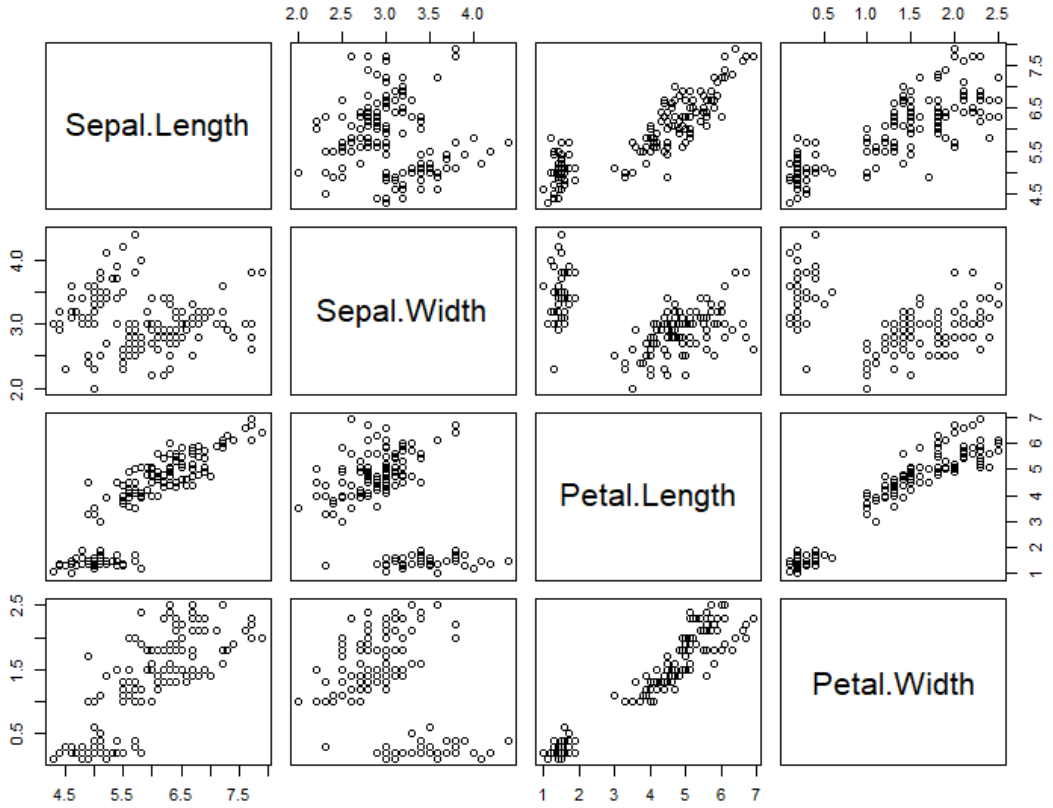
- use the PCA components $\leq d_i$ to obtain the 4 attributes with the highest squared sum of PCA loadings and list them in a table on the webpage
- use these four attributes and construct a scatterplot matrix

Task 3: use k-means to find clusters and color the points by cluster ID

- use the elbow method to find the best k (visualize the function on k)

What's a Scatterplot Matrix?

All possible bivariate scatterplots arranged into a matrix



Scoring and Due Dates

Each (task) bullet item carries 10 points

- an extra 10 pts for overall elegant implementation and function

Don't forget to

- label the axes and tick marks where appropriate
- show color legends where appropriate
- provide a meaningful header on each plot

Due date

- due March 7, end of day
- no group project, single only

Deliverables

Submit on Brightspace

- voice-narrated video file to show all features of your software in action
- in the video discuss any interesting observations you were able to make in the data
- 2-3 page report
  - describe interesting observations (beyond the video)
  - mention anything noteworthy about implementation (beyond the video)
- zip file with complete source code as well as the data
- submit video as an extra file

Grading

Grading

- TA will pick students at random for thorough code review sessions
- you better know your code !!!
- so, please do not just copy code beyond the D3 templates
- or even worse, videotape someone else's program