Visualization 2

Applied Multivariate Statistics – Spring 2012
Goals

- Bubbleplots
- Parallel Coordinate Plots
- Glyphplots (stars)
- Teaser: rggobi, googleVis (not for exam)
Bubbleplot

- Standard scatter plot, but the plotting symbols vary according to third variable
Parallel Coordinate Plots

- Easily overcrowded
Glyphplots for continuous data

- Each data sample is represented by a symbol (=glyph) with some aspects
- Depending on data values, aspects are more or less pronounced
- Very good, if you have few samples (<50) and not too many variables (<10)
Glyphplots: Stars

• Which cities are special?
• Which cities are like New Orleans?
• Seattle and Miami are quite far apart; how do they compare?
If data is not continuous…

- Categorical: See next lecture
- Mixed: Very hard
  - parallel coordinate plots might work
  - use colors or plot symbols
Export graphics

- “Export” button in Rstudio
- Functions jpeg(), pdf(), etc. in R
R commands to know

- Symbols
- Parcoord
- Stars
Teaser: ggobi

- Nice tool for brushing (highlight a sample) and identification of a sample
- In theory: Can be linked to R via package rggobi
- In practice: Hard to set up properly
Teaser: R package “googleVis”

- Many useful visualization functions
- Output can be easily embedded in webpage

Example: MotionChart
Show development of data over time
Next week

- Visualizing categorical data and making inference
- Detecting outliers in many dimensions