Principles of Analytic Graphics

* Show comparisons
* Show causality, mechanism, explanation, systematic structure
* Show multivariate data
* Integration of evidence
* Describe and document the evidence with appropriate labels, scales, sources, etc.
* Content is king

plotting systems

* Base: “artist’s palette” model
* Lattice: Entire plot specified by one function
* ggplot2: Mixes elements of base and lattice

Important base graphics Parameters

* pch: the plotting symbol
* lty: the line type
* lwd: the line width
* col: the plotting color
* xlab: character string for x-axis label
* ylab: character string for y-axis label
* las: the orientation of the axis labels on the plot
* bg: the background color
* mar: the margin size
* oma: the outer margin size
* mfrow: number of plots per row, column (plots are filled row-wise)
* mfcol: number of plots per row, column (plots are filled column-wise)

Important Base Plotting functions

* plot: make a scatterplot or other type of plot depending on the class of the object being plotted
* lines: add lines to plot, given a vector x values and a corresponding vector of y values
* points: add points to a plog
* text: add text labels to a plot using specified x, y coordinates
* title: add annotations to x, y axis labels, title, subtitle, outer margin
* mtext: add arbitrary text to the margins of the plot
* axis: add axis ticks/labels

Graphic Devices

* dev.cur(): active graphic device
* dev.set(int): sets active graphic device
* dev.copy() copys a plot
* dev.off() closes graphic device

Lattice Functions

* xyplot: scatterplots
* bwplot: boxplot
* histogram: histrograms
* stripplot: boxplot but with actual points
* dotplot: plot dots on “biolin strings”
* splom: scatterplot matrix: like pairs in base ploting system
* levelplot, contourplot: for plotting “image” data

Basic Components of a ggplot2 plot

* A data frame
* aesthetic mappings: how data are mapped to color, size
* geoms: geometric object like points, lines and shapes
* facets: for conditional plots
* stats: statistical transformations like binning, quantiles, smoothing
* scales: what scale an aesthetic map uses
* coordinate system

qplot = workhorse of ggplot2

PCA = Principal Component Analysis

SVD = Singular Value Decomposition

How do you deal with missing data in PCA and SVD?

* Impute the data