We have considered:

exploratory univariate analysis

exploratory regression analysis

exploratory anova

exploratory contingency table analysis

exploratory time series analysis

There are others:

exploratory multivariate analysis

datum - a vector $(y_1, \ldots, y_p)$

exploratory process data analysis

datum - curve, image, spike train, tesselation, video, shape, trees, ...

spatial eda

Exploratory multivariate analysis

graphical displays

    Chernoff plots

    residual plots

    star plots

multivariate least squares

$$\underset{n \times q}{Y} \sim \underset{n \times p}{X} \underset{p \times q}{B}$$

normal equations

$$\underset{p \times n}{X^T} \underset{n \times p}{X} \underset{p \times q}{\hat{B}} = \underset{p \times n}{X^T} \underset{n \times q}{Y}$$

residuals

$$\underset{n \times q}{E} = Y - X \hat{B}$$

$$\underset{1 \times q}{E_i} = Y_i - X_i \hat{B} \qquad i = 1, \ldots, n$$

univariate derived value

given $\underset{q \times q}{A} \geqslant 0$

$$E_i A E_i^T \qquad \text{scalar}$$

e.g. $\underset{\sim}{A} = \underset{\sim}{I}$

$\underset{\sim}{A} = \underset{\sim}{S}_E^{-1}$

Employ scalar eda methods
  residual plots
  quantile plots

  . . .

Other dimension reduction
procedures:
  latents of $\underset{\sim}{S}_Y$

projection pursuit

PRIM (grand tour)

interactive/dynamic

MANOVA

robust variants

Tukey and Wilk

"The basic general intent of data analysis is simply stated: to seek through a body of data for interesting relationships and information and to exhibit the results in such a way as to make them recognizable to the data analyzer and recordable for posterity."

11/17/04

Wal-Mart

Hurricanes Charley and Francis

"Wal-Mart has 460 terabytes of data"

".. internet has less than half as much"

"Wal-Mart hoards its information
obsessively"

".. takes pains to keep the information
secret"

"Wal-Mart considers data to be a top priority"

"tiny transmitters that would track every item that it sells"

"data enhancement - customer's name and address, telephone number, email"

"uses its mountain of data to push for greater efficiency"

"We are pretty near real time."

"Scan-based trading - manufacturers own each product until it is sold"

"shedding $50 billion of inventory"

"personal check... We don't mine that data"

"which customer bought what"

"You can do some association there"

"products sell as part of a larger basket"

" Shoppers might buy..., but not all these products need to be priced at rock-bottom "

" sex-discrimination ... Wal-Mart has the ability to use its human-resources database...to determine whether women were fairly promoted and paid. "