**data structures:**

* array: contain data of the same type
  + vector: 1D
  + matrix: 2D
  + array: 3+ dimensions
* data frame:
  + row\*column data format
  + each column is a vector
* list: an ordered collection of (possibly different) types

packages:

* install.packages (“package\_name”)
* library (“package\_name”)
* remove.packages (“package\_name”)

working directory:

* getwd() #get working directory
* setwd(“~/desktop”) # set working directory

reading csv files:

x <- read.csv(“InvestB.csv”, header=TRUE)

**descriptive statistics:**

mean (x)

sd (x) #standard deviation

hist (x) #draw a basic histogram

var (x) #variance

median (x)

quantile (x, probs) #probs [0,1]

range (x)

sum (x)

min (x)

max (x)

colMeans (InvestB, na.rm=TRUE) #ignore empty cell (remove NA values)

boxplot (InvestB)

**bivariate data:**

cor(Toothbrush)

plot (Toothbrush)

plot(Toothbrush$Price, Toothbrush$Function) #dataframe$column: identify column in a data frame

attach (Toothbrush) #call column in a data object by name without having to specify the object name (column name is unique)

plot (Price, Function)

fitted = lm(Function~Price) #linear model

Plot (Price, Function)

abline (fitted)

**estimation/hypothesis testing:**

boxplot (WorkerA, WorkerB)

t.test(WorkerA) #default confidence level 95%

t.test (WorkerA, conf.level = 0.55) #specify confidence level

t.test (WorkerA, WorkerB)