Week 1, R

Intro to Statistical Learning Ch2.3.

R Commands List

- `c()` (concatenate) for vectors.
- `→, ←, =` - Assignment operators
- `?funcname` - Meta/Info abt function
- `x + y` for vectors works like `np.array()`. Adds corresponding elements form x and y. Doesn't concatenate x and y.
- `length(x)` equivalent to `len(x)`.
- `ls()` - List all data and functions defined in the script. `rm(<inputs>)` - Delete specified inputs. Delete all by passing `ls()` inside `rm()`.

  Similar to `ls` and `rm` in CMD.
- `matrix(data, nrow, ncol)` - DS to create matrices. Pass in vector to be converted to matrix in data. Automatically separated based on `nrow, ncol`. Here, the matrix population is done column by column; i.e. first $a_{11}, \ldots, a_{n1}, a_{12}, \ldots, a_{n2}, \ldots$.
- `matrix(data, nrow, ncol, byrow=TRUE)` populates it row by row.
- General functions
  - Single number input - `sqrt()`
  - Vector input - `mean(), var(), sd()`
- `rnorm(n, mean, sd)` - Generates vector of n random numbers chosen from Normal dist.
- `cor(x, y)` - Corr(x, y)
- `set.seed()` fixes the seed and we'll get same result for `rnorm(n)` after setting that.
- `plot(x, y, xlab, ylab, main)` - Vectors x and y. `xlab, ylab` labels for axes. `main` is for overall.

- Saving output of an R plot. `pdf(filename)` preceding the `plot(x, y, col=<colour>)` followed by `dev.off()`. `dev.off()` to tell R we're done plotting and save.
- `seq(a, b, length)` - Sequence of numbers. Like `linspace` in Numpy.
- `contour()` - Contour plot. Inputs
  - x values. List must be increasing.
  - y values. "
  - z value. Matrix of elements
- `persp(x, y, z, theta, phi)` - 3d plot (perspective). `theta` and `phi` control viewing angle for 3d plot.

  Essentially you can view the contours for normal at point defined by params.
- Also `image()`.

- `read.table(filename, header, strings)` - Loading data from .data file. Loads from current dir.
  - `header` is Bool for whether or not first line has var names. `strings` is to identify missing data.
- `fix()` - View data in a spreadsheet-like window.
- `read.csv()` - Read csv file. Same params.
- `dim(data)` - Outputs dimensions of data. `names(data)` - Var names. `summary(data)` - Literally summary.
- `plot(database$column, database$column)` - Must mention the name of the data variable. R can't directly find columns.
- `attach(Auto)` - Make variables of data available directly by name; i.e. you don't need to do `Data$Column.`
- `as.factor()` - Convert quantitative data to qualitative. **Plotting this var will automatically produce box plots.**
- `hist()` - Histogram.

- `pair()` - Scatterplot for every subset of variables. Can also produce only for a subset by entering it within the function.

Some Cmds

- `A = matrix(1: 16, 4, 4)`
  - `1:16` - range
  - `4, 4` - nrows, ncols
- `A[2, 3]` - Indexing
-

Things I Found

- `y = 10 + x` gives a vector y with all elements of x incremented by 10.
- `outer()` -