Fall 2019 SML 201 Midterm 1 Cheat Sheet

Ch2.1

- Functions/techniques that help you to get familiar with the dataset: `dim()`, `names()`, `class()`, `head()`, `tail()`, `str()`, `summary()`
- Statistic functions: `mean(..., na.rm = FALSE)`, `sum(..., na.rm = FALSE)`, min(..., na.rm = FALSE), max(..., na.rm = FALSE), median(..., na.rm = FALSE), range(..., na.rm = FALSE);

Ch2.2

- Data extraction: `$` by name; `[,]` by indices; with conditions or logical arguments
- Combining conditions: `&`, `|`, `!`
- Data manipulation: `na.omit()`, `sort(..., decreasing = FALSE)`, `unique()`, `rank()`, `%in%`, `order()`
- Statistic functions: `sd(..., na.rm = ...)`, `var(..., na.rm = ...)`, `quantile(..., na.rm =..., p=...)`, `IQR(..., na.rm =...)`
- Graphical summary functions: `hist(..., main = ..., xlab = ..., ylab = ..., freq = ..., breaks = ..., main = ..., ...)`

Ch2.3

- Create a function: `function(input.variable1 = , input.variable2 = ,...){... return(...)}`
- Statistic functions: `tapply(X = , INDEX = , FUN = )`, `cor(x = , y = )`
- Data manipulation: `merge(x = , y = , by = , by.x = , by.y = , all = , all.x = , all.y = )`
- Graphical summary functions: `boxplot(y ~ grp, xlab = , ylab = , main = , names = , ...)`, plot(x = , y = , xlab = , ylab = , main = ,...)

Ch2.4

- ggpairs(..., aes(colour = , alpha =))
- ggplot(...) + geom_function(mapping = aes(x= , y= , colour =) + facet_grid(...~...)
- geom_boxplot()
- geom_histogram()
- geom_line()
- geom_point()
- geom_smooth()

Ch4.2

- To fit linear models: `lm(formula, data,...)`
- To make matrix plots: `ggpairs(data, upper = list(continuous = wrap(...)), lower = list(continuous = wrap(...))) + ...`

Additional functions from Precepts

- seq(to=, from =, by = , length=)
- rep(x = , times = , length.out =)
- >=, <=, >, <, ==
- par(mfrow =),