# Manage Packages

See # Manage R packages in Useful R code.r

# Import Data

dataset <- readr::read\_csv(file.choose())

dataset <- readxl::read\_excel(file.choose())

dataset <- haven::read\_sav(file.choose())

# Data Cleaning

View(dataset)

glimpse(dataset)

str(dataset)

# See column classes  
dataset %>%  
 dplyr::summarise\_all(class) %>%  
 tidyr::gather(variable, class)

# change to factor

dataset <- dataset %>% mutate(column = as.factor(column))

# order factor levels

dataset <- dataset %>% mutate(column = fct\_relevel(column,”one”,”two”,”etc”))

dataset <- dataset %>%

dplyr::mutate(ordinal2 = dplyr::recode\_factor(ordinal,

"first" = "1st",

"second" = "2nd",

"third" = "3rd",

"fourth" = "4th",

.ordered = TRUE))

# create new variable

dataset <- dataset %>% mutate (nominal2 = nominal)

# reverse code variable

Dataset$ordinal2 <- recode(dataset$ordinal2, first=4, second=3, third=2, fourth=1)

# add label to variable

dataset <- dataset %>% labelled::set\_variable\_labels(ordinal2 = "Variable Label for ordinal2")

# add labels to values in variable

dataset <- dataset %>% labelled::set\_value\_labels(ordinal2 = c(low=1, "medium low"=2, "medium high"=3, high=4))

# Data Exploration

summarytools::freq(dataset[c("nominal","ordinal")])

summarytools::descr(dataset[c("interval","ratio")])

view(dfSummary(dataset, plain.ascii = FALSE, style = "grid",

graph.magnif = 0.75, valid.col = FALSE, tmp.img.dir = "/tmp"))

ggplot(data=dataset, aes(x=nominal)) +

geom\_bar(stat="count")

ggplot(data=dataset, aes(x=ordinal)) +

geom\_bar(aes(y = (..count..)/sum(..count..)))

ggplot(data=dataset, aes(x= "", y = interval)) +

geom\_boxplot()

ggplot(data=dataset, aes(x=ratio)) +

geom\_histogram(binwidth = .2)

ggplot(data=dataset, aes(x=interval, y=ratio)) +

geom\_point()

qplot(dataset$nominal)

qplot(dataset$ordinal)

qplot(dataset$interval)

qplot(dataset$ratio)

qplot(dataset$nominal,dataset$ratio)

# Data Modeling

# Data Export

readr::write\_csv(dataset, path = “folder/fileName.csv”)

readr::write\_rds(dataset, path = "week2data.rds")

haven::write\_sav(dataset, path = "week2data.sav")

openxlsx::write.xlsx(dataset, file = "week2data.xlsx")

# Notes

# Format for noting packages and functions

# Can be pasted directly into code.

# This format avoids function masking

# when different packages use the same function name.

packageName::functionName(arguments)

# Useful Shortcuts

Run line of code – Ctrl-Enter

# Resources

https://rstudio.com

Useful sheets - rstudio.com/resources/cheatsheets/

Analysis process cheat sheet - <https://www.business-science.io/r-cheatsheet.html>

Data wrangling: <https://suzan.rbind.io/2018/01/dplyr-tutorial-1/>