**Reading Data**

**From Excel**

library(readxl)

xl\_data <- read\_excel("./data/gpa.xlsx")

head(xl\_data)

**haven Package**

library(haven)

**From SAS**

sas\_data <- read\_sas("./data//money.sas7bdat")

head(sas\_data)

**From SPSS**

spss\_data <- read\_sav("./data//airline\_passengers.sav")

head(spss\_data)

**From STATA**

stata\_data <- read\_dta("./data//stata\_sampledata\_crime.dta")

head(stata\_data)

**From JSON**

library(jsonlite)

url <- "http://fantasy.premierleague.com/web/api/elements/1"

json\_data <- fromJSON(url)

head(json\_data)

**tidyr**

Functions

* gather(): make *wide* data *long*
  + used *key-value* pair
* spread(): make *long* data *wide*
  + using *key* and *value*
* separate(): splits single column into multiple columns
* unite(): combines multiple columns into single column

library(readr)

jj.df <- read\_csv("./data/stockprice.csv")

jj.df

This data is considered wide since the time variable (represented as quarters) is structured such that each quarter represents a variable.

To represent time as a variable, we reshape the data.

**gather**

library(tidyr)

library(magrittr)

jj\_long <- jj.df %>%

gather(Quarter, Price, Qtr.1:Qtr.4)

jj\_long

**separate**

jj\_long <- jj\_long %>%

separate(Quarter, c("Time\_Interval","Interval\_ID"))

**unite**

jj\_long\_united <- jj\_long %>%

unite(Qtr, Time\_Interval, Interval\_ID, sep=".")

**spread**

jj\_wide <- jj\_long\_united %>%

spread(Qtr, Price)