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)
```