

# Class 1 R Notebook

```
## Warning: package 'readr' was built under R version 4.2.3
```

## Read in Dataset

```
# loads a package needed to use the read_csv() function
# install package before first using it for the first time

lego <- read_csv("lego.csv")

## Rows: 1304 Columns: 15
## -- Column specification -----
## Delimiter: ","
## chr (7): Set_Name, Theme, Ages, Packaging, Weight, Availability, Size
## dbl (8): Item_Number, Pieces, Price, Amazon_Price, Year, Pages, Minifigures,...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.

# Alternative way to load lego (remove # to use)
# You can load a csv file directly from internet (i.e. Github)
# lego <- read_csv("https://raw.githubusercontent.com/JA-McLean/STOR455/master/data/lego.csv")
```

## Preview Dataset

```
head(lego) #Vertically (Top 6 Rows by Default)
```

```
## # A tibble: 6 x 15
##   Item_Number Set_Name      Theme Pieces Price Amazon_Price Year Ages Pages
##   <dbl> <chr>          <chr>  <dbl> <dbl>      <dbl> <dbl> <chr> <dbl>
## 1    41916 Extra Dots - Se~ DOTS    109  3.99        3.44  2020 Ages~    NA
## 2    41908 Extra Dots - Se~ DOTS    109  3.99        3.99  2020 Ages~    NA
## 3    11006 Creative Blue B~ Clas~    52  4.99        4.93  2020 Ages~    37
## 4    11007 Creative Green ~ Clas~    60  4.99        4.93  2020 Ages~    37
## 5    41901 Funky Animals B~ DOTS    33  4.99        4.99  2020 Ages~    NA
## 6    41902 Sparkly Unicorn~ DOTS    33  4.99        4.99  2020 Ages~    NA
## # i 6 more variables: Minifigures <dbl>, Packaging <chr>, Weight <chr>,
## #   Unique_Pieces <dbl>, Availability <chr>, Size <chr>
```

```
str(lego) #Horizontally
```

```
## spc_tbl_ [1,304 x 15] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
## $ Item_Number : num [1:1304] 41916 41908 11006 11007 41901 ...
## $ Set_Name : chr [1:1304] "Extra Dots - Series 2" "Extra Dots - Series 1" "Creative Blue Bricks
## $ Theme : chr [1:1304] "DOTS" "DOTS" "Classic" "Classic" ...
## $ Pieces : num [1:1304] 109 109 52 60 33 33 33 33 33 33 ...
## $ Price : num [1:1304] 3.99 3.99 4.99 4.99 4.99 4.99 4.99 4.99 4.99 4.99 ...
```

```
## $ Amazon_Price : num [1:1304] 3.44 3.99 4.93 4.93 4.99 4.99 4.99 4.99 4.99 4.99 ...
## $ Year          : num [1:1304] 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 ...
## $ Ages          : chr [1:1304] "Ages_6+" "Ages_6+" "Ages_4+" "Ages_4+" ...
## $ Pages         : num [1:1304] NA NA 37 37 NA NA NA NA NA NA ...
## $ Minifigures   : num [1:1304] NA NA NA NA NA NA NA NA NA NA ...
## $ Packaging      : chr [1:1304] "Foil pack" "Foil pack" "Box" "Box" ...
## $ Weight        : chr [1:1304] NA NA NA NA ...
## $ Unique_Pieces : num [1:1304] 6 6 28 36 10 9 9 12 10 9 ...
## $ Availability  : chr [1:1304] "Retail" "Retail" "Retail" "Retail" ...
## $ Size          : chr [1:1304] "Small" "Small" "Small" "Small" ...
## - attr(*, "spec")=
## .. cols(
## ..   Item_Number = col_double(),
## ..   Set_Name = col_character(),
## ..   Theme = col_character(),
## ..   Pieces = col_double(),
## ..   Price = col_double(),
## ..   Amazon_Price = col_double(),
## ..   Year = col_double(),
## ..   Ages = col_character(),
## ..   Pages = col_double(),
## ..   Minifigures = col_double(),
## ..   Packaging = col_character(),
## ..   Weight = col_character(),
## ..   Unique_Pieces = col_double(),
## ..   Availability = col_character(),
## ..   Size = col_character()
## .. )
## - attr(*, "problems")=<externalptr>
```

## Summary of Amazon Price

### Numerically

```
#dataframe$variable
```

```
mean(lego$Amazon_Price, na.rm=T)
```

```
## [1] 57.8232
```

```
median(lego$Amazon_Price, na.rm=T)
```

```
## [1] 37.325
```

```
sd(lego$Amazon_Price, na.rm=T)
```

```
## [1] 66.26777
```

### Visually

Example of how to add a summary to R Code. You can write paragraphs outside the code chunks like this. For example, I may want to tell my audience that the visual below shows a histogram of the Amazon price of lego sets, and we can clearly see that the distribution is extremely right-skewed.

```
hist(lego$Amazon_Price)
```

**Histogram of lego\$Amazon\_Price**

