XP Booster
March 9

This week you will be working **SOLO**. You have exactly 15 minutes to finish each of the following questions. For this challenge you will be using the Lego dataset.

1. Visualize **and describe** the distribution for lego prices using a histogram. Draw the histogram and include your description at the bottom of this page.
2. Visualize the distribution of prices for lego sets with a rating of greater than 3. There is no need to draw this distribution of describe, you only need to write the code that will produce it.
3. Create a barplot showing the number of sets belonging to each lego theme.
4. Create a pairwise boxpllot comparing the prices of 'Easy' lego sets and 'Hard' lego sets. Describe how the two distributions differ.
5. Create a barplot for the themes that the top 100 priced lego sets belong. Describe the plot below.

Solution:

```
## Load Packages
library(ggplot2)
library(tidyverse)

lego <- read.csv('lego.csv')

## Question 1
hist(lego$Price)

## Question 2
glimpse(lego)
lego %>% filter(rating>3) %>% ggplot(aes(x=Price))+geom_histogram()

## Question 3
barplot(table(lego$theme))

## Question 4
easy <- lego %>% filter(difficulty=='Easy')
hard <- lego %>% filter(difficulty=='Challenging')
boxplot(easy$Price,hard$Price)

## Question 5
lego %>% arrange(desc(Price)) %>% select(theme) %>% head(100) %>% table() %>%
    barplot()
```