

Mini Math

Week 2

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 boxplot 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:

```
1 ## Load Packages
2 library(ggplot2)
3 library(tidyverse)
4
5 lego <- read.csv('lego.csv')
6
7 ## Question 1
8 hist(lego$Price)
9
10 ## Question 2
11 glimpse(lego)
12 lego %>% filter(rating>3) %>% ggplot(aes(x=Price))+geom_histogram()
13
14 ## Question 3
15 barplot(table(lego$theme))
16
17 ## Question 4
18 easy <- lego %>% filter(difficulty=='Easy')
19 hard <- lego %>% filter(difficulty=='Challenging')
20 boxplot(easy$Price, hard$Price)
21
22 ## Question 5
23 lego %>% arrange(desc(Price)) %>% select(theme) %>% head(100) %>% table() %>%
  barplot()
```