Project 5
Counting and Probability

Project Instructions

In this project, you will use R to solve counting and probability problems. To gain the most benefit from this project, avoid calculating numeric values and entering them into R. Instead, use R to do all necessary calculations.

Setting up Your Project

When working in RStudio, be careful with cloud drives. RStudio does **not** always play well with cloud drives' longer file path names. You will be best served by using a local drive whenever possible until you are comfortable troubleshooting technical issues.

1. Create a new project called "Lastname-Project5".

2. Create an R Script file within your project called "Lastname-Project5.R".

3. Include your name, the date, and the class in a comment as the first line of the script.

4. Include any script initialization code and library loading.

Testing Your Solution

You can evaluate your project using the project5_tests.R test file. This will only work if you store results in the exact variable names specified in bold and in parentheses at the end of the problem.

Problems

Using the ball-dataset.

1. Download the data set **ball-dataset.csv** and read it into your script. Each ball in the dataset is represented by a color (red, blue, green, or yellow) and a label (A, B, C, D, or E).

2. Create a frequency table as a data.frame or tibble that contains counts for each color of ball (**freq_color**).

```
# A tibble: 4 × 2
  color   counts
  <chr>    <int>
1 blue       307
2 green      192
3 red        408
4 yellow      93
```
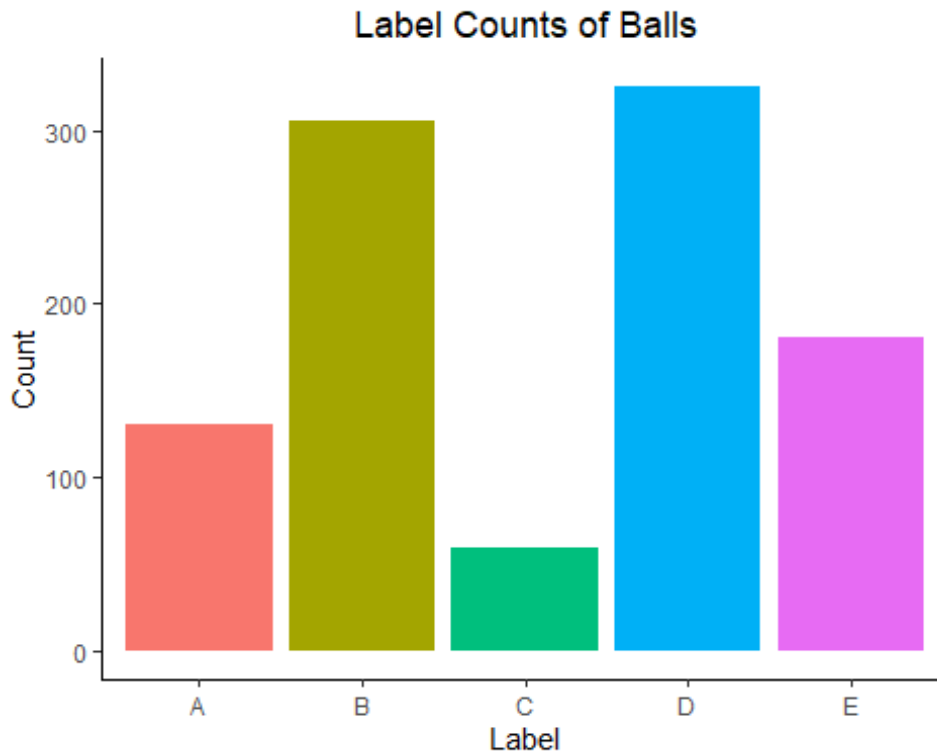
3. Create a frequency table as a data.frame or tibble that contains counts for each label of ball (**freq_label**).

```
# A tibble: 5 × 2
  label counts
  <chr>  <int>
1 A        130
2 B        305
3 C         59
4 D        325
5 E        181
```

4. Create a bar chart of the ball data set representing the counts of the different colors.
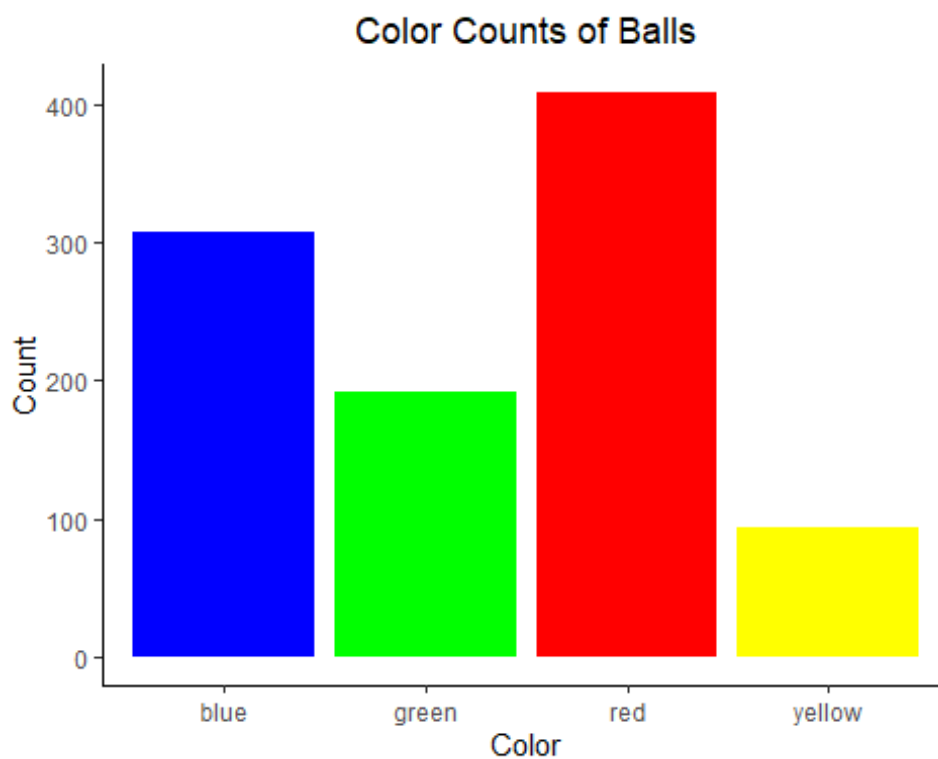
5. Create a bar chart of the ball data set representing the counts of the different labels.

6. What is the probability of drawing a green ball (**prob6_result**)?

7. What is the probability of drawing a blue or a red ball (**prob7_result**)?

8. What is the probability of drawing a ball with a label of A or C (**prob8_result**)?

9. What is the probability of drawing a yellow ball with a D (**prob9_result**)?

10. What is the probability of drawing a yellow ball or a ball with a D (**prob10_result**)?

11. What is the probability of drawing a blue ball followed by a red ball without replacement (**prob11_result**)?

12. What is the probability of drawing four green balls in a row without replacement (**prob12_result**)?

13. What is the probability of drawing a red ball followed by a ball with a B without replacement (**prob13_result**)?

14. [Challenge] Write a function called **my_factorial** that computes the factorial of a given number.

    – my_factorial (0) = 0
    – my_factorial(3) = 6
   – my_factorial(5) = 120

Creating a coin flipping data frame

For the following problems, consider an **UNFAIR COIN** that has a probability **.6** of landing on heads.

15. Manually create a data.frame or tibble that contains all possible outcomes of flipping the coin four times (**coin_outcomes**).

```
# A tibble: 16 × 4
   first second third fourth
   <chr> <chr>  <chr> <chr>
 1 H     H      H     H
 2 H     H      H     T
 3 H     H      T     H
 4 H     H      T     T
 5 H     T      H     H
 6 H     T      H     T
 7 H     T      T     H
 8 H     T      T     T
 9 T     H      H     H
10 T     H      H     T
11 T     H      T     H
12 T     H      T     T
13 T     T      H     H
14 T     T      H     T
15 T     T      T     H
16 T     T      T     T
```

16. Compute the probability of each row outcome and store it as a column in the data.frame or tibble. You can do this manually or programmatically (**coin_outcomes**).

17. There are 5 possible outcomes in our coin dataset if we count the number of heads in each row. For example, the row "H H H H" has 4 heads and the row "H T H T" has 2 heads. Compute the probability of each of the 5 possible outcomes (**num_heads_prob**).

18. What is the probability of an outcome of three heads (**prob18_result**)?

19. What is the probability of an outcome of two heads or four heads (**prob19_result**)?

20. What is the probability of an outcome of less than or equal to three heads (**prob20_result**)?

21. Create a bar chart where the *x*-axis is the outcome, and the *y*-axis is the probability.

Probability Distibution of Heads for 4 flips
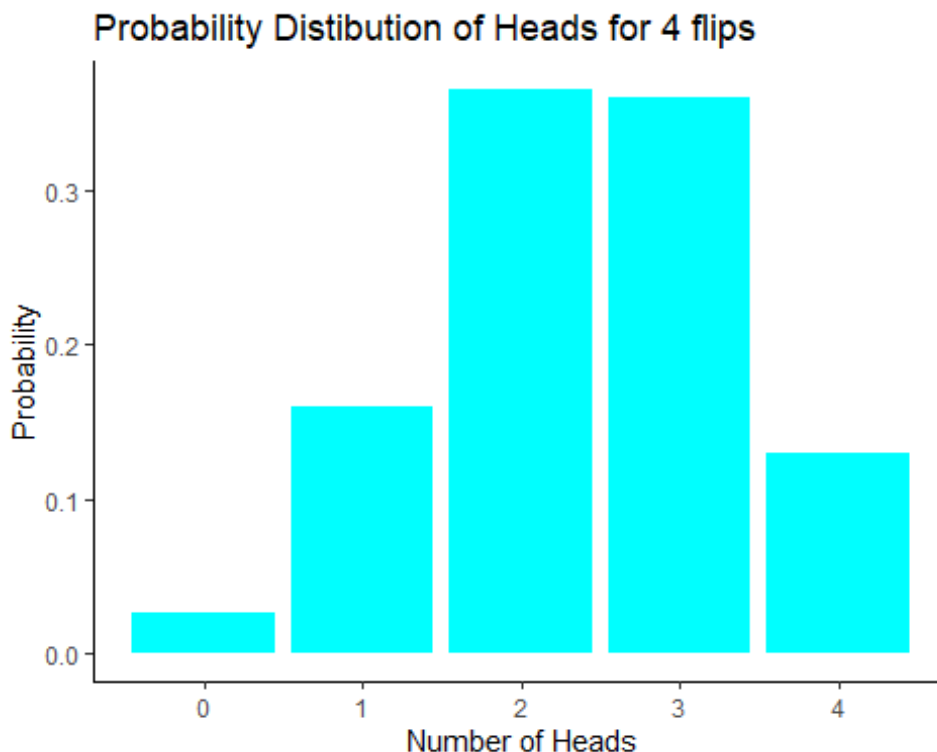
Soccer Games (Optional Challenge Problems)

The following problems consider a soccer team with a 75% chance of winning a game at home and a 50% chance of winning away games. Consider that the team is about to play 10 games: five at home and five away.

22. What is the probability that they will win exactly 10 games (**prob22_result**)?

23. What is the probability that they will win more than one game (**prob23_result**)?

24. How many different ways could you pick five games at random and have three home games and two away games (**prob24_result**)?

Submitting to Canvas

When you are satisfied with your solution, do the following:

1. Submit one (1) file in Canvas. Your R script should be named **Lastname_Project5.R**.

Congratulations on completing Project 5!