Sports Analytics

Aaron Nielsen, Department of Statistics, Colorado State University

2022-06-02

Contents

About

This book serves as the course textbook for the following courses at Colorado State University:

- STAT 351 (Sports Statistics and Analytics 1)
- STAT 451 (Sports Statistics and Analytics 2

CSU students contributed to the creation of this book. Many thanks to the following student collaborators:

- Levi Kipp
- Ellie Martinez
- Isaac Moorman

Current Tasks

Updated: "2022-06-02"

Team Tasks and Tips

1. Find datasets from various sports to use as examples for EDA and later chapters
2. Show how to get basic sumamry statistics from these datasets using dplyr, tidy
3. Describe and calculate useful team and individual (descriptive statistics). Example: Baseball: calculate AVG, OBP, OPS, WOBA
4. (High quality) Visualizations using ggplot
5. Look for relevant "sports" R packages
6. Include examples from CSU and Colorado sports teams when possible
7. Sports to be included: Baseball/Softball, Football, Basketball, Soccer, Hockey, Volleyball
8. Sports to be potentially included: Lacrosse, Cricket, Handball,

---

Aaron:

Sports:

Chapters: Currently working to add content to chapters 1-4

Ellie:

Sports: Soccer, Volleyball

Chapters: EDA, Probability

Levi:

Sports: Basketball, Hockey

Chapters: EDA, Probability

Isaac:

Sports: Baseball, Football, Tennis

Chapters: EDA, Scraping

Chapter 1

Exploratory Data Analysis

1.1 Getting Started With R

1.1.1 Installing R

For this class, you will be using R Studio to complete statistical analyses on your computer.

To begin using R Studio, you will need to install "R" first and then install "R Studio" on your computer.

Step 1: Download R
(a) Visit https://www.r-project.org/
(b) Click **CRAN** under **Download** (c) Select any of the mirrors
(d) Click the appropriate link for your type of system (Mac, Windows, Linux)
(e) Download R on this next page.
(For Windows, this will say **install R for the first time**. For Mac, this will be under **Latest release** and will be something like **R-4.1.0.pkg** – the numbers may differ depending on the most recent version)
(f) Install R on your computer

Step 2: Download R Studio
(a) Visit https://www.rstudio.com/products/rstudio/download/#download
(b) Click to download
(c) Install R Studio on your computer

Step 3: Verify R Studio is working
(a) Open R Studio

(b) Let's enter a small dataset and calculate the average to make sure everything is working correctly.
(c) In the console, type in the following dataset of Sammy Sosa's season home run totals from 1998–2002:

```
sosa.HR <- c(66,63,50,64,49)
```

(d) In the console, calculate the average season home run total for Sammy Sosa between 1998–2002:

```
mean(sosa.HR)
```

```
## [1] 58.4
```

(e) Did you find Slammin' Sammy's average home run total from 1998–2002 was 58.4? If so, you should be set up correctly!

1.1.2 Some R Basics

For the following examples, let's consider Peyton Manning's career with the Denver Broncos. In his four seasons with the Broncos, Manning's passing yard totals were: 4659, 5477, 4727, 2249. Let's enter this data into R. To enter a vector of data, use the **c()** function.

```
peyton <- c(4659, 5477, 4727, 2249)
```

To look at the data you just put in the variable *peyton*, type *peyton* into the console and press enter.

```
peyton
```

```
## [1] 4659 5477 4727 2249
```

Some basic function for calculating summary statistics include **summary**, **mean()**, **median()**, **var()**, and **sd()**.

```
summary(peyton)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    2249    4056    4693    4278    4914    5477
```

```
mean(peyton)
```

```
## [1] 4278
```

```
sd(peyton)
```

```
## [1] 1402.522
```

R allows you to install additional packages (collections of functions) that aren't offered in the base version of R. To install a package, use **install.packages()** and to load a package, use **library()**.

One package that we will use frequently is **tidyverse**. This package includes several other packages and functions such as **ggplot** (plotting function), **dplyr** (data manipulation package), and **stringr** (string manipulation package).

```
install.packages("tidyverse")
library("tidyverse")
```

You will also need to know how to load datasets from files. For this class, we will typically provide data files is .csv format.

Here is how to load a file:

```
# load readr package and load example dataset
library(readr)
NFL_2021_Team_Passing <- read_csv("data/NFL_2021_Team_Passing.csv")

# we can look at the header (first few entries) using "head()"
head(NFL_2021_Team_Passing)
```

```
## # A tibble: 6 x 25
##      Rk Tm              G   Cmp   Att `Cmp%`   Yds    TD `TD%`   Int `Int%`   Lng
##   <dbl> <chr>       <dbl> <dbl> <dbl>  <dbl> <dbl> <dbl> <dbl> <dbl>  <dbl> <dbl>
## 1     1 Tampa Bay~     17   492   731   67.3  5229    43   5.9    12    1.6    62
## 2     2 Los Angel~     17   443   674   65.7  4800    38   5.6    15    2.2    72
## 3     3 Dallas Co~     17   444   647   68.6  4800    40   6.2    11    1.7    73
## 4     4 Kansas Ci~     17   448   675   66.4  4791    37   5.5    13    1.9    75
## 5     5 Los Angel~     17   406   607   66.9  4642    41   6.8    18    3       79
## 6     6 Las Vegas~     17   429   628   68.3  4567    23   3.7    14    2.2    61
## # ... with 13 more variables: `Y/A` <dbl>, `AY/A` <dbl>, `Y/C` <dbl>,
## #   `Y/G` <dbl>, Rate <dbl>, Sk <dbl>, SKYds <dbl>, `Sk%` <dbl>, `NY/A` <dbl>,
## #   `ANY/A` <dbl>, `4QC` <dbl>, GWD <dbl>, EXP <dbl>
```

1.2 Descriptive Statistics

1.2.1 Definitions

Definition 1.1. A ***population*** is a well-defined complete collection of objects.

Definition 1.2. A ***sample*** is a subset of the population.

Example 1.1. Suppose we are interested in studying Peyton's Manning's season passing yards totals. How could you define the population and what is one possible sample?

Definition 1.3. ***Quantitative data*** is numeric data or numbers. It can be broken into two further categories: discrete and continuous data.

Definition 1.4. ***Discrete data*** is quantitative data with a finite or countably infinite number of values.

Definition 1.5. ***Continuous data*** is quantitative data with an uncountably infinite number of values or data taken from an interval.

Example 1.2. What are possible discrete and continuous data associated with Peyton Manning?

Definition 1.6. ***Qualitative data*** refers to names, categories, or descriptions. It can also be broken down into two further categories, nominal data and ordinal data.

Definition 1.7. ***Nominal data*** is qualitative data with no natural ordering.

Definition 1.8. ***Ordinal data*** is qualitative data with a natural ordering.

Example 1.3. What are possible nominal and ordinal data associated with Peyton Manning?

1.2.2 Descriptive Statistics

While we will learn about some descriptive statistics that are unique to specific sports, there are some descriptive statistics that are frequently used in many applications.

1.2.2.1 Descriptive Statistics for Quantitative Data

There are different descriptive statistics depending on the type of data you are analyzing. We will begin by looking at descriptive statistics for quantitative data.

To begin, let $x_1, x_2, \ldots, x_n$ represent a numerical dataset with a sample of size n, where x_i is the i^{th} value in the dataset.

Definition 1.9. The ***sum*** of the data values is given by: $\sum_{i=1}^{n} x_i = x_1 + x_2 + \ldots + x_n$

Definition 1.10. The ***sample mean*** (or sample average), $\bar{x}$, of the numerical dataset is given by $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

Definition 1.11. The ***population mean*** (or population average), μ, is the mean value for the entire population.

The mean can be thought of as a measure of center or more generally, a measure of location.

Example 1.4. Recall that Peyton Manning's season passing yards total while with the Broncos were: 4659, 5477, 4727, 2249. Calculate the sample mean of these values.

```r
# Calculate the sample of Peyton Manning's passing yards season totals with Colts
peyton.broncos <- c(4659, 5477, 4727, 2249)
mean(peyton.broncos)
```

```
## [1] 4278
```

In sports statistics, we often have to choose between using a descriptive statistic that summarizes a quantity versus a descriptive statistic that summarizes a rate. For instance, in basketball, we can compare two players based on how many points they score in a game (total quantity) or we can compare two players based on how many points per minute played (rate statistic). Many applications in sports analytics focus more on rate statistics rather than quantity statistics. Why?

We can measure the spread or variability of a dataset using *variance* and *standard devatiation*.

Definition 1.12. The ***sample variance***, s^2, of the numerical dataset is a measure of spread and is given by $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

Definition 1.13. The ***sample standard deviation***, s, of the numerical dataset is a measure of spread and is given by $s = \sqrt{s^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$

Definition 1.14. The ***population variance***, σ^2, is the variance for an entire population.

Definition 1.15. The ***population standard deviation***, σ, is the standard deviation for an entire population.

We often prefer to work with standard deviations as a measure of spread as opposed to variance because standard deviations are given in our original units.

```
# Calculate the variance and standard deviation of Peyton Manning's passing yards seas
var(peyton.broncos) # units: yards^2
```

```
## [1] 1967068
```

```
sd(peyton.broncos) # units: yards
```

```
## [1] 1402.522
```

Definition 1.16. The **sample median**, $\tilde{x}$, of a numerical dataset is the middle value when the data are ordered from smallest to largest. In other words, let $x_1, x_2, \ldots, x_n$ be the (unordered) dataset and let $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ be the same dataset but ordered from smallest to largest. If n is odd, then $\tilde{x} = x_{(n+1)/2}$ and if n is even, then $\tilde{x} = \frac{1}{2} \cdot \left[x_{(\frac{n}{2})} + x_{(\frac{n+1}{2})}\right]$.

Example 1.5. Calculate the sample median of Peyton Manning's season passing yards total while with the Colts (3739, 4135, 4413, 4131, 4200, 4267, 4557, 3747, 4397, 4040, 4002, 4500, 4700).

Like sample mean, sample median is a measure of center. It gives you an idea of where the "middle" of your dataset is.

We can calculate sample mean and sample median in R as follows:

```
# Calculate the median of Peyton Manning's passing yards season totals with Broncos an
peyton.colts <- c(3739, 4135, 4413, 4131, 4200, 4267, 4557, 3747, 4397, 4040, 4002, 45
median(peyton.broncos)
```

```
## [1] 4693
```

```
median(peyton.colts)
```

```
## [1] 4200
```

Definition 1.17. A ***percentile*** is a measure of relative standing. The p^{th} percentile is the number where at least p% of the data values are less than or equal to this number.

Definition 1.18. A ***quantile*** is a measure of relative standing and are the cut points for breaking a distribution of values into equal sized bins.

Definition 1.19. A ***quartile*** is a measure of relative standing and are the cut points for breaking a distribution of values into four equal parts.

```
# Calculate the 10th and 90th percentile of Peyton Manning's passing yards season totals with Co
quantile(peyton.colts,0.10)
```

```
##  10%
## 3798
```

```
quantile(peyton.colts,0.90)
```

```
##    90%
## 4545.6
```

```
quantile(peyton.colts,c(0.1,0.9))
```

```
##    10%    90%
## 3798.0 4545.6
```

Special percentiles:
1. 25th percentile = 1st quartile = Q_1
2. 50th percentile = 2nd quartile = $Q_2 = \tilde{x}$
3. 75th percentile = 3rd quartile = Q_3

Definition 1.20. ***Range*** is a measure of spread, measures the full width of a dataset, and is given by: $Range = Max - Min$.

Definition 1.21. ***Interquartile range*** is a measure of spread, measures the width of the middle 50% of a dataset, and is given by: $IQR = Q_3 - Q_1$.

Definition 1.22. A ***five number summary*** describes the center, spread, and edges of a dataset and is given by: $(Min, Q_1, Q_2, Q_3, max)$.

```
summary(peyton.colts)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    3739    4040    4200    4218    4413    4700
```

```
quantile(peyton.colts,c(0,0.25,0.5,0.75,1))
```

```
##   0%  25%  50%  75% 100%
## 3739 4040 4200 4413 4700
```

1.2.2.2 Descriptive Statistics for Qualitative Data

In sports statistics, we also encounter qualitative (categorical) data which is names or labels which has its own descriptive statistics.

To begin, let $x_1, x_2, \ldots, x_n$ represent a categorical dataset with a sample of size n, where x_i is the i^{th} value in the dataset.

Definition 1.23. The ***proportion*** of sampled data that fall into a category is given by: $p = \frac{\text{\# in category}}{\text{\# total}}$

' 'Proportion" and "Probability" are often used interchangeably. Both have a minimum value of 0 and a maximum value of 1.

Definition 1.24. The ***percentage*** of sampled data that fall into a category is given by: $P\% = 100 \cdot p = 100 \cdot \frac{\text{\# in category}}{\text{\# total}}$

Percentages in this context can have a minimum value of 0% and a maximum value of 100%.

Example 1.6. In 2014, Peyton Manning started as quarterback for the Denver Broncos. The result of the Broncos' 16-game season was:
Win, Win, Loss, Win, Win, Win, Win, Loss, Win, Loss, Win, Win, Win, Win, Loss, Win

Calculate the proportion and percentage of Broncos' winning games in 2014.

```
broncos2014 <- c("Win", "Win", "Loss", "Win", "Win", "Win", "Win", "Loss", "Win", "Los
broncos.prop <- sum(broncos2014 == "Win")/length(broncos2014); broncos.prop
```

```
## [1] 0.75
```

```
broncos.perc <- 100*broncos.prop; broncos.perc
```

```
## [1] 75
```

We can also build a frequency table that summarizes the categories and their occurrences using **table()** in R. Note that **table()** works for quantitative and qualitative data.

```
table(broncos2014)
```

```
## broncos2014
## Loss  Win
##    4   12
```

1.3 Visualizations

Conveying information visually is also an important part in providing a description of a dataset.

R provides some basic plotting functions such as **plot**, **hist**, and **barplot**. These plotting functions are simple and not always very clean looking.

In this class, we will use analogous plotting functions in **ggplot2** that are much improved plotting functions.

If you have already installed the **tidyverse** package, it should have also installed the **ggplot2** package.

```
# You have likely already installed the tidyverse package but if not, use the following command u
# install.packages("tidyverse")

# Load the tidyverse package (which includes ggplot2)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.6     v dplyr   1.0.9
## v tibble  3.1.7     v stringr 1.4.0
## v tidyr   1.2.0     v forcats 0.5.1
## v purrr   0.3.4

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

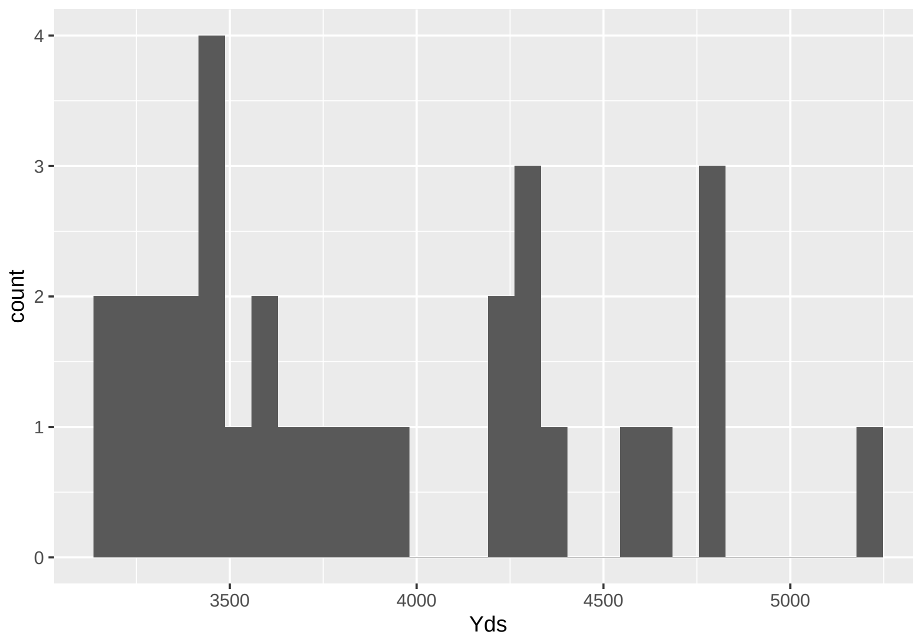
Let's load the file "NFL_2021_Team_Passing.csv" which contains NFL Team Passing Statistics, 2021

```
library(readr)
NFL_2021_Team_Passing <- read_csv("data/NFL_2021_Team_Passing.csv")
```

Histograms are one of the most common and basic ways to visualize a dataset's distribution of values. To make a histogram, you will use **ggplot** and **geom_histogram**.

Example 1.7. Create a histogram of the NFL Team Passing Yards in 2021.

```
NFL_2021_Team_Passing %>% ggplot(aes(x=Yds)) + geom_histogram()
```

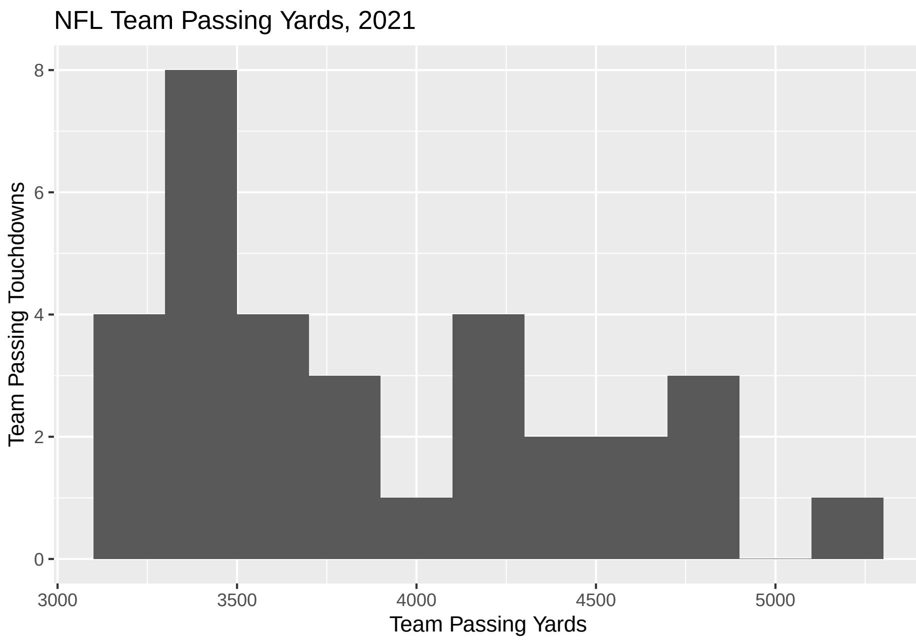


Notice how %>% is used to **pipe** the dataset into ggplot. This is using the pipe function from the **dplyr** package.

By default, **geom_histogram** uses 30 bins but this is customizable. Let's make the bins have a width of 200.

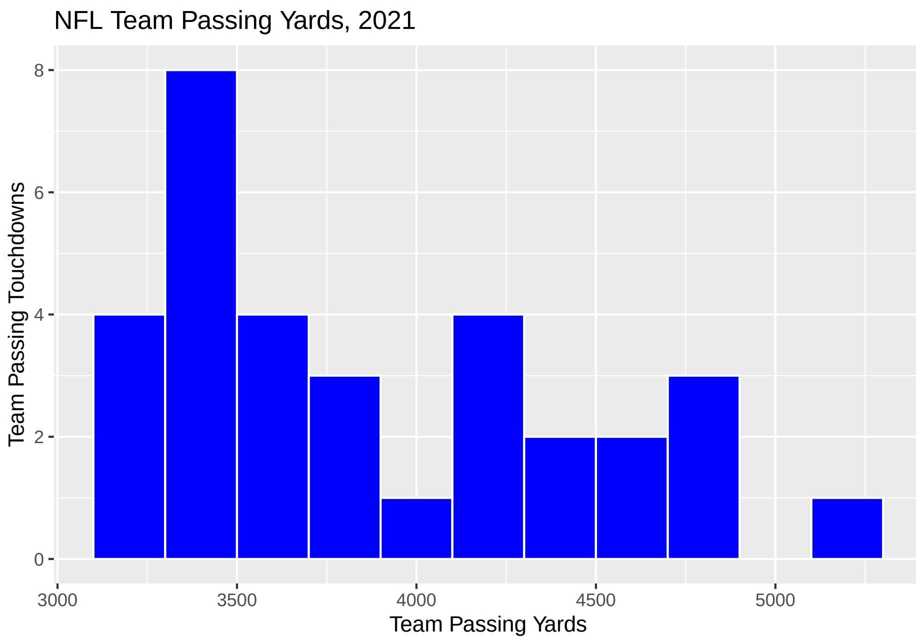
All good visualizations have good labels. Let's improve the axis labels and give the figure a title.

```
NFL_2021_Team_Passing %>% ggplot(aes(x=Yds)) +
  geom_histogram(binwidth = 200) +
  labs(x="Team Passing Yards",y="Team Passing Touchdowns",title="NFL Team Passing Yards
```



We also have numerous options to change the appearance of plots when using **ggplot**. Let's change the bins color to *blue* and change the bin borders to *white*.

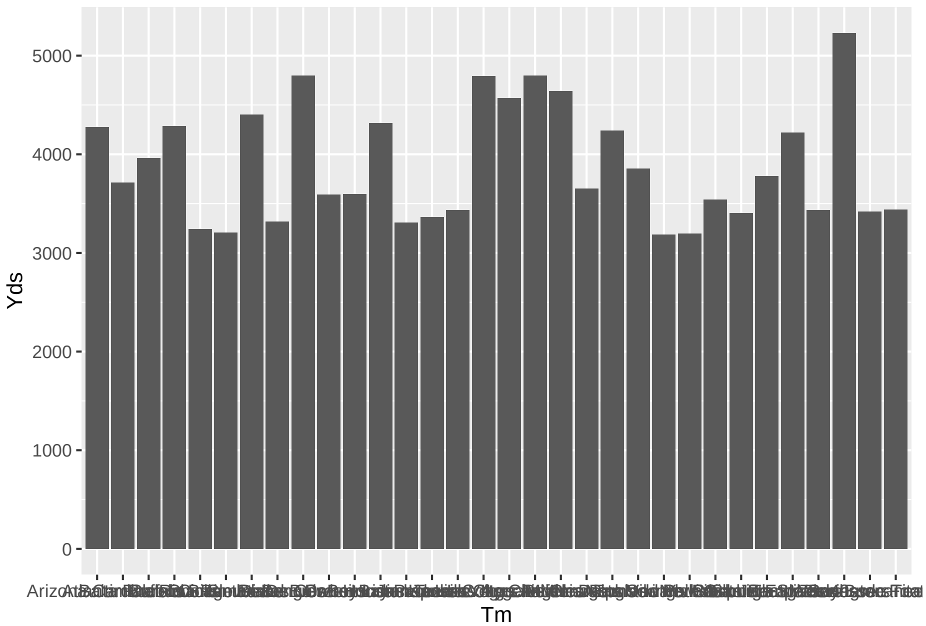
```
NFL_2021_Team_Passing %>% ggplot(aes(x=Yds)) +
  geom_histogram(color = "white", fill = "blue", binwidth = 200) +
  labs(x="Team Passing Yards",y="Team Passing Touchdowns",title="NFL Team Passing Yards, 2021")
```



We can also create bar plots using ggplot using the **geom_bar** function.

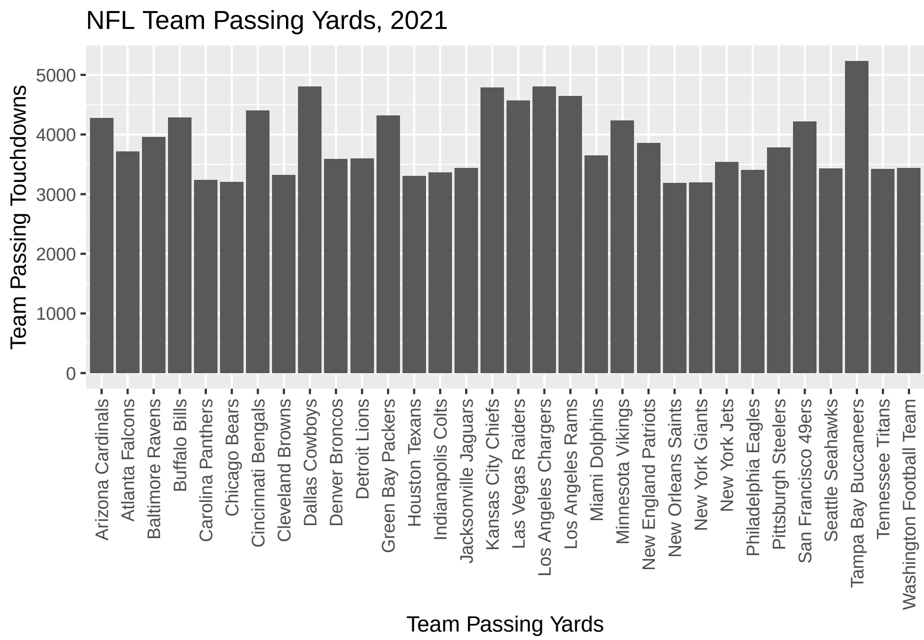
Example 1.8. Create a bar plot with teams on the horizontal axis and passing touchdowns on the vertical axis.

```
NFL_2021_Team_Passing %>% ggplot(aes(x=Tm,y=Yds)) +
  geom_bar(stat="identity")
```



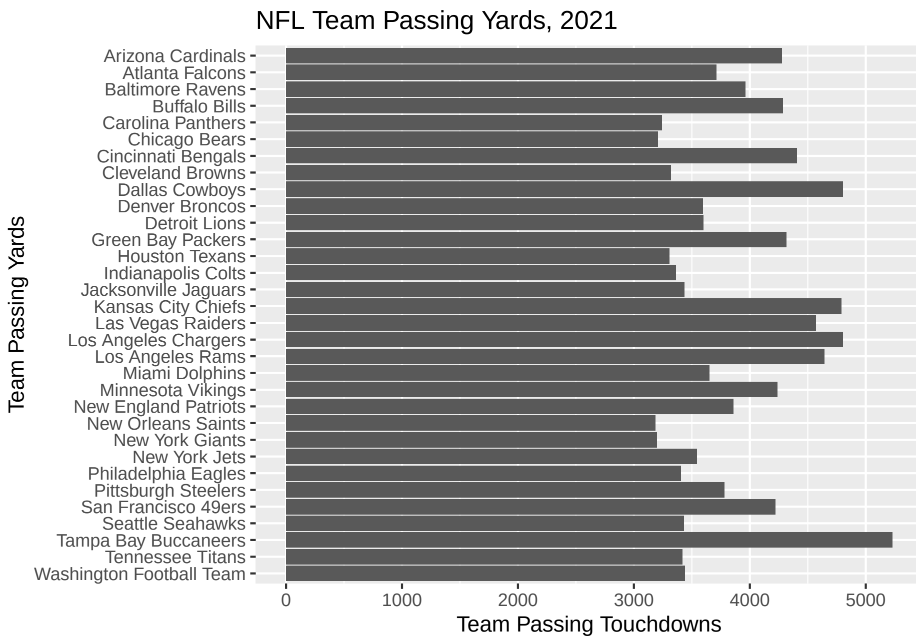
The team labels are a complete mess. Let's fix this and make some adjustments to the axis labels and figure title.

```
NFL_2021_Team_Passing %>% ggplot(aes(x=Tm,y=Yds)) +
  geom_bar(stat="identity") +
  labs(x="Team Passing Yards",y="Team Passing Touchdowns",title="NFL Team Passing Yards, 2021") +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```



We can flip this graph if we like as well. Note that when we flip the graph, our labels get in reverse ordering, so this can be fixed using **fct_rev()** which is part of the **forcats** package.

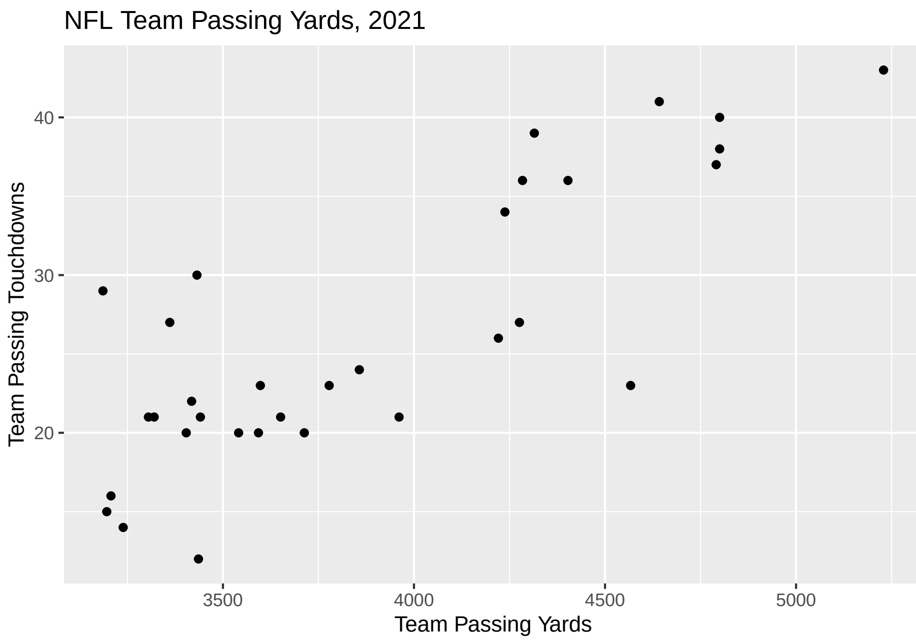
```
NFL_2021_Team_Passing %>%
  ggplot(aes(x=fct_rev(Tm),y=Yds)) +
  geom_bar(stat="identity") +
  labs(x="Team Passing Yards",y="Team Passing Touchdowns",title="NFL Team Passing Yards
  coord_flip()
```



Another common and useful visualization is a scatterplot which shows the relationship between two numeric variable. In ggplot, you use **geom_point()**.

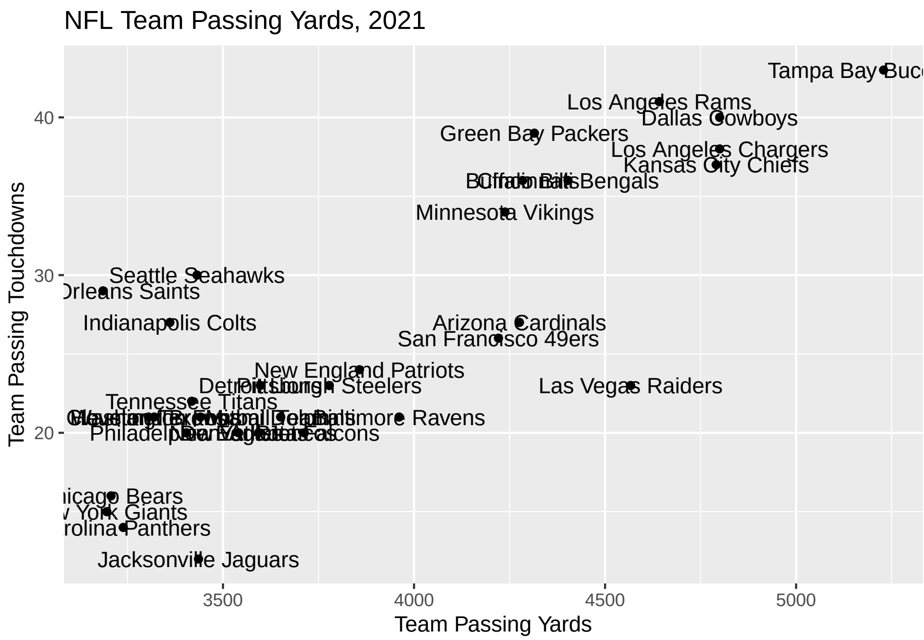
Example 1.9. Create a scatterplot of Team Passing Yards and Team Passing Touchdowns from the NFL 2021 dataset.

```
NFL_2021_Team_Passing %>%
  ggplot(aes(x=Yds,y=TD,label=Tm)) +
  geom_point() +
  labs(x="Team Passing Yards",y="Team Passing Touchdowns",title="NFL Team Passing Yards, 2021")
```



We may want to include team labels on this plot, however, it can get messy very quickly with a lot of points.

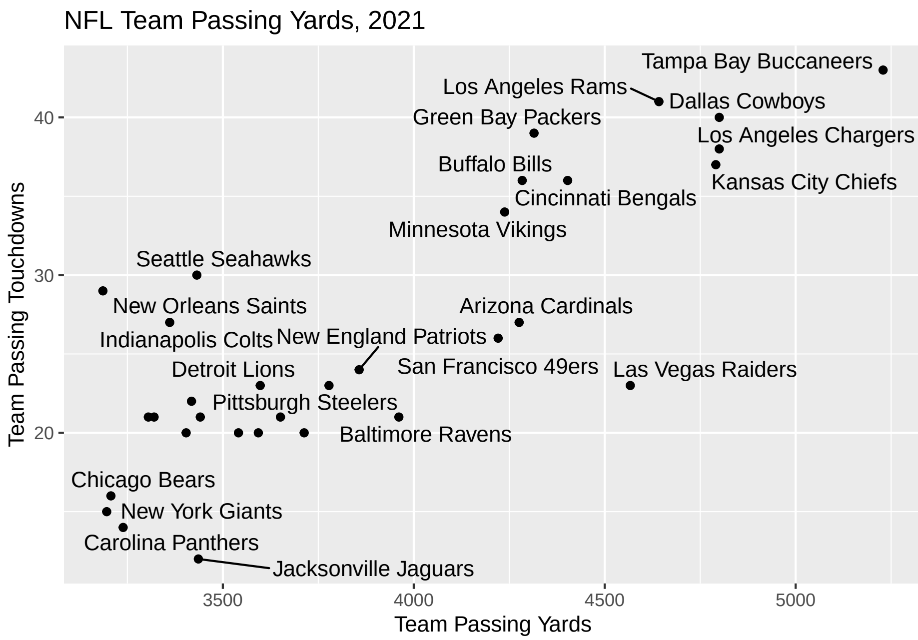
```
NFL_2021_Team_Passing %>%
  ggplot(aes(x=Yds,y=TD,label=Tm)) +
  geom_point() +
  labs(x="Team Passing Yards",y="Team Passing Touchdowns",title="NFL Team Passing Yards
  geom_text()
```



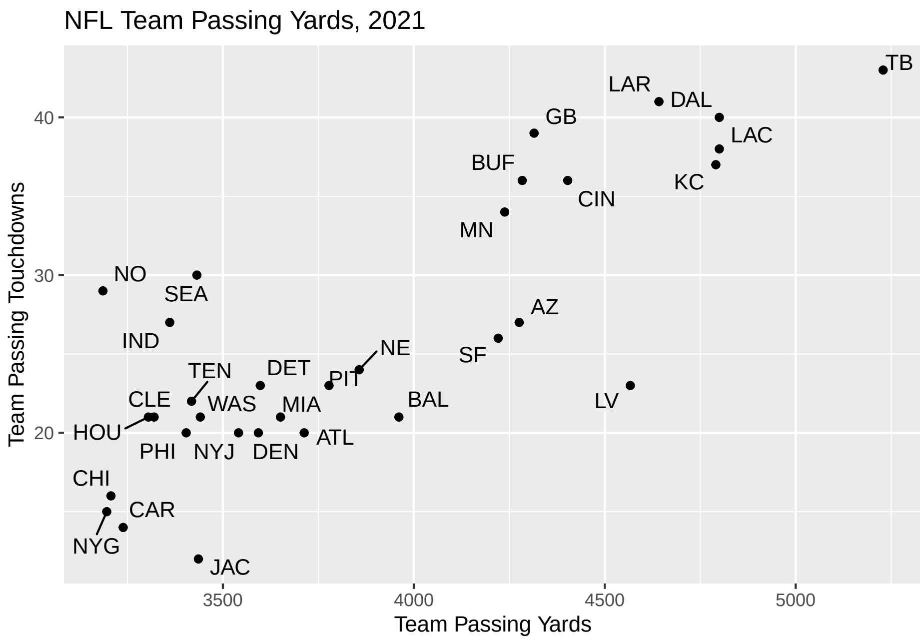
Many sports leagues have around 30 teams, so a clean scatterplot with labels can be tricky to make. Here are some options below.

```
# install ggrepel package
library(ggrepel)
NFL_2021_Team_Passing %>%
  ggplot(aes(x=Yds,y=TD,label=Tm)) +
  geom_point() +
  labs(x="Team Passing Yards",y="Team Passing Touchdowns",title="NFL Team Passing Yards, 2021") +
  geom_text_repel()
```

```
## Warning: ggrepel: 9 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```




```
NFL_2021_Team_Passing$Abbr <- c("TB","LAC","DAL","KC","LAR","LV","CIN","GB","BUF","AZ"
                                "BAL","NE","PIT","ATL","MIA","DET","DEN","NYJ","WAS","
                                "TEN","PHI","IND","CLE","HOU","CAR","CHI","NYG","NO")
NFL_2021_Team_Passing %>%
  ggplot(aes(x=Yds,y=TD,label=Abbr)) +
  geom_point() +
  labs(x="Team Passing Yards",y="Team Passing Touchdowns",title="NFL Team Passing Yard
  geom_text_repel(box.padding = 0.3)
```



1.4 Baseball

1.5 Football

1.6 Basketball

1.7 Soccer

1.8 Volleyball

1.9 Hockey

For this example, we'll use a set of NHL data from moneypuck.com. First, let's load the data into R and open the data frame.

```
nhl_2022_data <- read_csv("https://moneypuck.com/moneypuck/playerData/seasonSummary/2021/regular/

head(nhl_2022_data)
```

```
## # A tibble: 6 x 107
##   team...1 season name  team...4 position  situation games_played
##   <chr>     <dbl> <chr> <chr>    <chr>     <chr>            <dbl>
```

name	situation	games_played	xGoalsPercentage	corsiPercentage
WPG	other	82	0.49	0.50
WPG	all	82	0.49	0.50
WPG	5on5	82	0.49	0.49
WPG	4on5	82	0.16	0.14
WPG	5on4	82	0.86	0.86
CBJ	other	82	0.52	0.49
CBJ	all	82	0.45	0.48
CBJ	5on5	82	0.45	0.48

```
## 1 WPG        2021 WPG   WPG      Team Level other              82
## 2 WPG        2021 WPG   WPG      Team Level all                82
## 3 WPG        2021 WPG   WPG      Team Level 5on5               82
## 4 WPG        2021 WPG   WPG      Team Level 4on5               82
## 5 WPG        2021 WPG   WPG      Team Level 5on4               82
## 6 CBJ        2021 CBJ   CBJ      Team Level other              82
## # ... with 100 more variables: xGoalsPercentage <dbl>, corsiPercentage <dbl>,
## #   fenwickPercentage <dbl>, iceTime <dbl>, xOnGoalFor <dbl>, xGoalsFor <dbl>,
## #   xReboundsFor <dbl>, xFreezeFor <dbl>, xPlayStoppedFor <dbl>,
## #   xPlayContinuedInZoneFor <dbl>, xPlayContinuedOutsideZoneFor <dbl>,
## #   flurryAdjustedxGoalsFor <dbl>, scoreVenueAdjustedxGoalsFor <dbl>,
## #   flurryScoreVenueAdjustedxGoalsFor <dbl>, shotsOnGoalFor <dbl>,
## #   missedShotsFor <dbl>, blockedShotAttemptsFor <dbl>, ...
```

We can create nice looking tables using the "kableExtra'' package. Let's look at the first eight rows and a samll selection of columns of the data frame and format the table output using a kable table.

```
library("kableExtra")
```

```
##
## Attaching package: 'kableExtra'

## The following object is masked from 'package:dplyr':
##
##     group_rows
```

```
nhl_2022_data[1:8, c(3,6:9)] %>% kbl() %>% kable_styling()
```

This dataset includes a *lot* of covariates. It also splits these data by different game situations: even-strength (5 on 5), power play (5 on 4), etc. Let's subset the data to include all game situations.

Use the `nrow` command to check the number of columns in the new data frame. Check: Is it the same as the number of teams in the league for the 2021-2022 season?

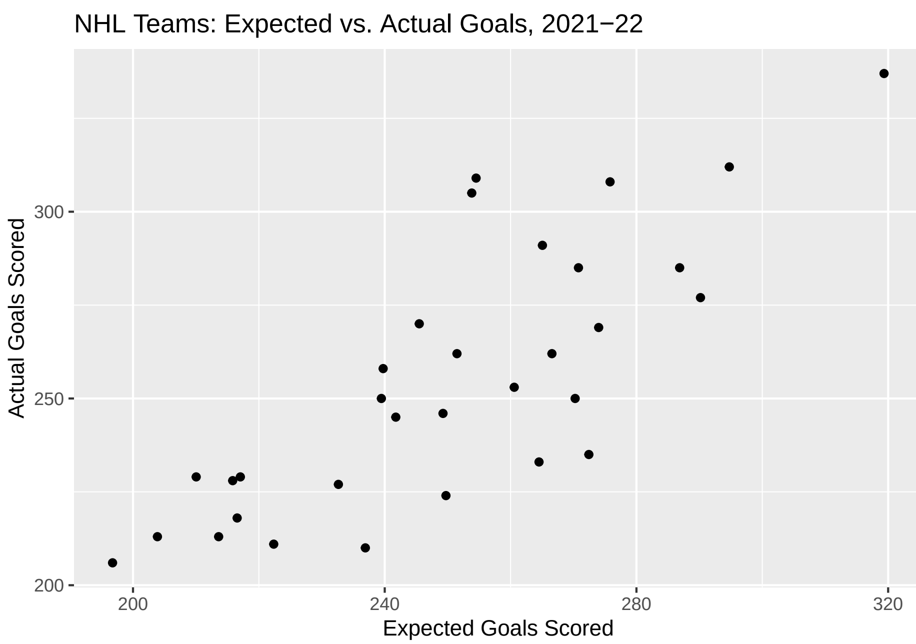
```
nhl_data_all <- filter(nhl_2022_data, situation == "all")

nrow(nhl_data_all)
```

```
## [1] 32
```

The dataset includes an Expected Goals statistic for each team in the `xGoalsFor` column. Let's plot this quantity against the team's actual number of goals scored; this is given by the `goalsFor` column.

(Remember to always have a good title and axis labels!)

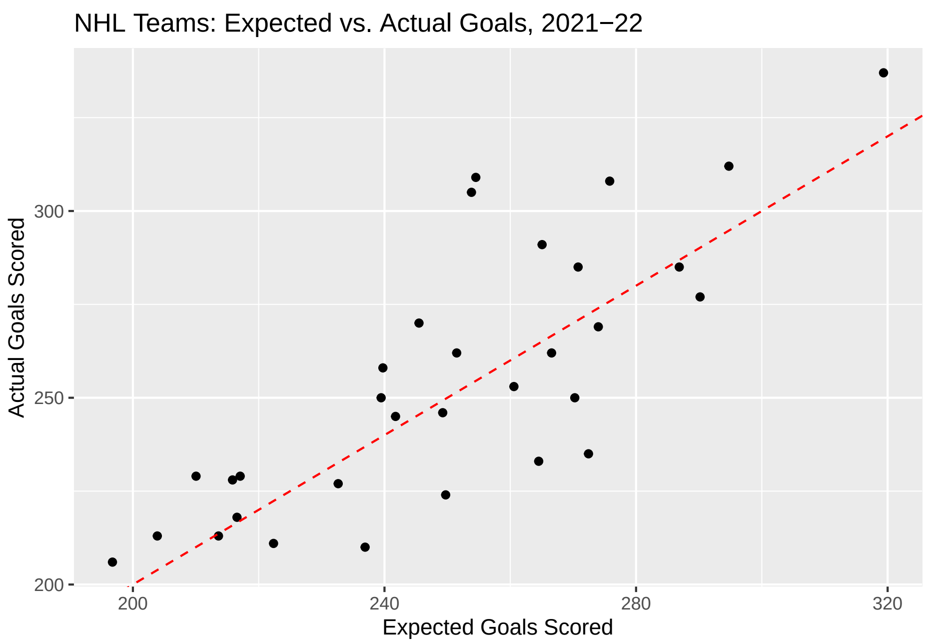
```
ggplot(data=nhl_data_all, aes(x=xGoalsFor, y=goalsFor)) + labs(x="Expected Goals Scored", y="Actu
```



As expected, there is a general positive correlation between expected and actual goals ($r \approx 0.8$). However, there is some variability - for example, the Kings only scored 7 more actual goals than the Ducks, despite having 56.6 more expected goals.

Let's add a line to the graph using the `geom_abline` function corresponding to the line $y = x$, the line on which data points would fall if expected goals were equal to actual goals. We can also customize the line's color and type.

```
ggplot(data=nhl_data_all, aes(x=xGoalsFor, y=goalsFor)) + labs(x="Expected Goals Scored
```



Note: A slope of 0 and an intercept of 1 are actually the default parameters for the function.

Q: What does it mean for a team's data point to fall below this line? Above it?

A: If the data point is below the line, it means the expected goals were greater than the actual goals; if the data point is above the line, it means the actual goals were greater than the expected goals.

Q: Do you think that a team's expected goals would be more likely to be closer to its actual goals for a ten-game stretch, an entire season, or five consecutive seasons? Why?

A: We would expect that as sample size increases, the result would become closer to expectation. So, actual goals would be most likely closer to expected goals over a span of five seasons.

Chapter 2

Probability

Chapter Preview

Simply put, probability is the study of randomness. In this chapter, we will define probability, learn rules of probability, and apply these rules to sports data.

2.1 Definitions

Definition 2.1. An ***experiment*** is any activity or process whose outcome is subject to uncertainty.

Definition 2.2. The ***sample space*** of an experiment, denoted by Ω or $\mathcal{S}$, is the set of all possible outcomes of that experiment.

Definition 2.3. An ***event*** is any collection (subset) of outcomes contained in the sample space, Ω.

Example 2.1.

Example 2.2.

2.2 Set Theory

For the following examples, suppose that we are interested in the batting outcomes of a plate appearance in softball.

Let A be the event that the batter gets walked, let B be the event that the batter gets a hit, let C be the event that the batter strikes out, and let D be the event that the batter makes it to first base at the end of their at bat.

We will define a handful of set operations to help us when we begin calculating the probability of different events occurring.

Definition 2.4. The ***compliment*** of an event A, denoted by A^c or A', is the set of all outcomes in Ω that are not contained in A.

Example 2.3. Draw a Venn diagram illustrating A^c and describe the event.

Definition 2.5. The ***union*** of two events A and B, denoted by $A \cup B$ and read "A or B", is the event consisting of all outcomes that are either in A or B or in both.

Example 2.4. Draw a Venn diagram illustrating $A \cup D$ and describe the event.

Definition 2.6. The ***intersection*** of two events A and B, denoted by $A \cap B$ and read "A and B", is the event consisting of all outcomes that are in both A and B.

Example 2.5. Draw a Venn diagram illustrating $A \cap D$ and describe the event.

Definition 2.7. The ***difference*** of two events A and B, denoted by $A\,/\,B$ and read "difference of A and B", is the event consisting of all outcomes that are in A but not in B.

Example 2.6. Draw a Venn diagram illustrating $D\,/\,A$ and describe the event.

Definition 2.8. Two events A and B are said to be ***disjoint*** (or ***mutually exclusive***) if $A \cap B = \emptyset$

Example 2.7. Are the events A and B disjoint? How about A and D?

2.3 Probability Axioms and Properties

There are some basic assumptions of "axioms" which are the foundation of the theory of probability. Andrey Kolmogorov first described these axioms in 1933.

2.3.1 Axioms of Probability

1. $P(A) \geq 0$, for any event A
2. $P(\Omega) = 1$
3. If $A_1, A_2, A_3, ...$ is a collection of disjoint events, then: $P(\cup_{i=1}^{\infty} A_i) = P(A_1 \cup A_2 \cup ...) = \sum_{i=1}^{\infty} P(A_i)$

Note that all probabilities are between 0 and 1, that is, for any event A, $0 \leq P(A) \leq 1$.

We can convert to percentages by multiplying probabilities by 100, however, this is a set that is only done after all calculations have been completed.

2.3.2 Properties of Probability

- $P(\emptyset) = 0$

- $P(A^c) = 1 - P(A)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A \cup B \cup C) =$
  $P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$
- $P([A \cup B]^c) = P(A^c \cap B^c)$
- $P([A \cap B]^c) = P(A^c \cup B^c)$

2.4 Laws of Probability

Definition 2.9. Let A and B be two events such that $P(B) > 0$. Then the ***conditional probability*** of A given B, written $P(A|B)$, is given by: $P(A|B) = \frac{P(A \cap B)}{P(B)}$

Example 2.8. In 2001, Barry Bonds broke the single season home run record with 73 home runs. In this season, he had 664 plate appearances, 156 hits, 177 walks and 9 hit by pitches. Given that Bonds reached base (via hit, walk, or HBP), what was the probability that he got a hit?

Theorem 2.1 (Multiplication Rule)**.** *For any two events A and B, $P(A \cap B) = P(B|A) \cdot P(A)$.*

Definition 2.10. Events $A_1, A_2, ..., A_n$ are said to form a ***partition*** of a sample space Ω if both:
(i) $A_i \cap A_j = \emptyset$ $(i \neq j)$
(ii) $\cup_{i=1}^{n} A_i = \Omega$

Theorem 2.2 (Law of Total Probability)**.** *Suppose events $A_1, A_2, ..., A_n$ form a partition of Ω, then: $P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + ... P(B|A_n)P(A_n)$*

Theorem 2.3 (Bayes Theorem: simple version)**.** *Suppose events B and C form a partition of Ω, then: $P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|C)P(C)}$*

Theorem 2.4 (Bayes Theorem)**.** *Suppose events $B_1, B_2, \ldots, B_n$ form a partition of Ω, then:* $P(B_k|A) = \frac{P(B_k \cap A)}{P(A)} = \frac{P(A|B_k)P(B_k)}{P(A|B_1)P(B_1)+P(A|B_2)P(B_2)+\ldots+P(A|B_n)P(B_n)}$

2.5 Combinatorics

Combinatorics is the mathematical study of counting, particularly with respect to permutations and combinations.

Definition 2.11. The ***factorial function** ($n!$)* is defined for all positive integers by: $n! = n \cdot (n-1) \cdot \ldots 2 \cdot 1$

Note that $0! \equiv 1$ and $1! \equiv 1$.

Definition 2.12. An ordered subset is called a ***permutation***. The number of permutations of size k that can be formed from the n elements in a set is given by: $P_{n,k} = \frac{n!}{(n-k)!}$

Definition 2.13. An unordered subset is called a ***combination***. The number of combinations of size k that can be formed from the n elements in a set is given by: $C_{n,k} = \binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$

Theorem 2.5 (Product Rule for Ordered Pairs)**.** *If the first element of an ordered pair can be selected in n_1 ways and for each of the these n_1 ways the second element of the pair can be selected in n_2 ways, then the number of pairs is $n_1 \cdot n_2$.*

Theorem 2.6 (Generalized Product Rule)**.** *Suppose a set consists of k elements (k-tuples) and that there are n_1 possible choices for the first element, n_2 possible choices for the second element, ... , and n_k possible choices for the k^{th} element, then there are $n_1 \cdot n_2 \cdot \ldots \cdot n_k$ possible k-tuples.*

2.6 Random Variables

2.7 Some examples

Over the course of a season, a hockey player scored a goal 30% of the time during a home game, and $P(player\ scores \mid away\ game) = .18$. Assume all games are either home or away.

Q: What is the probability the player scored a goal in any game if there were an equal number of home and away games?

A: $P(score) = P(score|home)P(home) + P(score|away)P(away) = .3(.5) + .18(.5) = .24$

Q: What is the probability the player scored a goal in any game if there were twice as many home games as away games?

A: $P(score) = P(score|home)P(home) + P(score|away)P(away) = .3(\frac{2}{3}) + .18(\frac{1}{3}) = .26$

Q: What is the probability the player scored a goal in any game if the ratio of home games to away games is 2:3?

A: $P(score) = P(score|home)P(home) + P(score|away)P(away) = .3(\frac{2}{5}) + .18(\frac{3}{5}) = .228$

2.7.1 Sets and Conditional Probability

100 sports fans in Colorado were polled and it was found that 64 had attended either a Denver Nuggets or Colorado Avalanche game at Ball Arena (formerly Pepsi Center). 34 people had seen only a Nuggets game, while 17 had seen both a Nuggets and an Avalanche game.

Q: How many people saw an Avalanche game but not a Nuggets game?

A: 64 - 34 - 17 = 13

Q: What is the probability that a randomly selected person in the poll had been to a Nuggets game?

A: (34 + 17) / 100 = .51

Q: What is the probability that a randomly selected person that had been to a game at Ball Arena had been to a Nuggets game?

A: (34 + 17) / 64 = .797

Q: What is the probability that a randomly selected person had been to a Nuggets game given they had been to an Avalanche game?

A: 17 / (17 + 13) = .567

2.7.2 Binomial Probability

Two baseball teams are playing a 4-game series. The home team has a .65 probability of winning each game, and the away team a .35 probability. Assume each game is independent.

I used baseball in this example because it's the sport that most often has 4-game series, but it could easily be replaced by another sport.

Find the following probabilities.

(a) The road team wins exactly 1 game.

$\binom{4}{1}\ .65^3\ .35^1 = \binom{4}{3}\ .65^3\ .35^1 \approx .384$

```
dbinom(1, 4, .35)
```

```
## [1] 0.384475
```

```
dbinom(3, 4, .65)
```

```
## [1] 0.384475
```

(b) The home team wins exactly 2 games.

$\binom{4}{2}$ $.65^2$ $.35^2 \approx .311$

```
dbinom(2, 4, .65)
```

```
## [1] 0.3105375
```

```
dbinom(2, 4, .35)
```

```
## [1] 0.3105375
```

(c) The road team wins at least 2 games.

$\binom{4}{2}$ $.65^2$ $.35^2 + \binom{4}{3}$ $.65^1$ $.35^3 + .35^4 = 1 - [.65^4 + \binom{4}{1}$ $.65^3$ $.35^1] \approx .437$

```
pbinom(1.9, 4, .35, lower.tail=F)
```

```
## [1] 0.4370187
```

```
pbinom(2, 4, .65, lower.tail=T)
```

```
## [1] 0.4370187
```

(d) The series ends in a sweep.

$.65^4 + .35^4 \approx .194$

```
dbinom(4, 4, .65) + dbinom(4, 4, .35)
```

```
## [1] 0.1935125
```

```
.65^4 + .35^4
```

```
## [1] 0.1935125
```

2.7.3 Binomial Coefficient Symmetry

Playoff series for a certain sports league are played as a best-of-seven series, with one team hosting four games and the opposing team hosing three. An executive for the league wishes to know the number of ways the home and away games can be assigned. (One such combination is A-A-B-B-A-B-A, the format used by the NBA and NHL for their best-of-seven series.) What is the total number of combinations?

Answer: Since there are a fixed number of games (seven) and a fixed number of games that must be given to the lower-seeded team (four), there are $\binom{7}{4} = \frac{7!}{4!\,\cdot\,(7-4)!} = 35$ ways to create a home-away pattern for the seven-game series.

However, instead of thinking about the number of ways to assign the games to the team that gets four home games, what if we thought about the number of ways to assign games to the team that gets three home games?

That would be $\binom{7}{3}$. We can use the `choose` command in R to find this quantity.

```
choose(7,3)
```

```
## [1] 35
```

It turns out that this binomial coefficient is also equal to 35.

Theorem: $\binom{n}{k} = \binom{n}{n-k}$

$\binom{n}{k} = \frac{n!}{k!\,\cdot\,(n-k)!}$

$\binom{n}{n-k} = \frac{n!}{(n-k)!\,\cdot\,(n-(n-k))!} = \frac{n!}{(n-k)!\,\cdot\,k!} = \binom{n}{k}$

2.7.4 Binomials and Multinomials

Suppose we are curious about probabilities regarding the results of a soccer team's next five games.

Wait!!! A soccer game has three possible outcomes (win, lose, draw)! We can't use the binomial distribution, since it limits us to two possible outcomes!

It depends. If we are interested in the probability that a soccer team wins 2 of their next 5 games, we can use the binomial distribution. We can create the following partition of the sample space of outcomes: (Win) and (Win^C), where the second set includes both losing and drawing.

Then, the formula would be represented as:

$\binom{5}{2}\ P(Win)^2\ P(Win^C)^{(5-2)}$

If we are interested in the probability of the team winning two of the next five games, drawing two, and losing one, we cannot use the binomial theorem. That involves three outcomes, and would be represented as a multinomial.

2.7.5 Geometric (First Success) RVs

Caution: Some references parameterize the Geometric distribution based on the number of failures before the first success, rather than the trial on which the first success occurs. This changes the PMF, mean, and variance, so be careful.

```
set.seed(2022)
geometric <- rgeom(100, 1/3)
head(geometric, 20)
```

```
##  [1]  2  5  1  3 12  7  1  4  2  2  1  1  1  2  0  0  0  4  3  0
```

Some of the values were 0, which could not happen if R was considering the number of the trial on which the first success occurred. You can add 1 to the values given by `R` to arrive at the First Success distribution.

```
first_success <- geometric + 1
head(first_success, 20)
```

```
##  [1]  3  6  2  4 13  8  2  5  3  3  2  2  2  3  1  1  1  5  4  1
```

```
mean(first_success)
```

```
## [1] 3.03
```

The mean of this sample of variables is 3.03, which is close to the expected mean of $\frac{1}{p} = 3$.

2.7.6 Geometric Distribution - Hockey

Suppose the number of shots needed by a hockey team in order to score their first goal, X, is modeled by a Geometric($\frac{1}{10}$) random variable.

Q: What is the probability that it takes more than 3 shots to score the first goal?

A: $P(X > 3) = P(X = 4) + P(X = 5) + P(X = 6) + ...$

This is an infinite series, so let's use the Law of Total Probability.

$P(X > 3) = 1 - P(X \leq 3) = 1 - [P(X = 1) + P(X = 2) + P(X = 3)] = 1 - [(\frac{1}{10}) + (\frac{9}{10})^1(\frac{1}{10}) + (\frac{9}{10})^2(\frac{1}{10})] = .729$

Chapter 3

Monte Carlo Simulation

Chapter 4

Statistical Inference

4.1 One Sample and Two Sample t-tests and confidence intervals

Chapter 5

Correlation

Chapter 6

Linear Regression

Chapter 7

Data Scraping

Chapter 8

Principal Component Analysis

Chapter 9

Clustering

Chapter 10

Classification

Chapter 11

Decision Trees

11.1 Random Forests

11.2 Gradient Boosting

Chapter 12

Non-parametric Statistics

Chapter 13

Baseball

Chapter 14

Football

Chapter 15

Basketball

Chapter 16

Soccer

Chapter 17

Hockey

Chapter 18

Volleyball

18.1 Resources

Women's Volleyball D1 Statistics

Chapter 19

Other Sports

Chapter 20

Ellie's Stuff

20.1 To Be Added

20.1.1 Topic 1

```
library(worldfootballR)
```

```
# Get "Squad Standard Stats" Data
fb_big5_advanced_season_stats(season_end_year = 2021, stat_type = "standard", team_or_player = "t
```

```
##     Season_End_Year           Squad           Comp Team_or_Opponent Num_Players
## 1              2021          Alavés        La Liga             team          30
## 2              2021          Alavés        La Liga         opponent          30
## 3              2021          Angers        Ligue 1             team          31
## 4              2021          Angers        Ligue 1         opponent          31
## 5              2021         Arminia     Bundesliga             team          26
## 6              2021         Arminia     Bundesliga         opponent          26
## 7              2021         Arsenal Premier League             team          29
## 8              2021         Arsenal Premier League         opponent          29
## 9              2021     Aston Villa Premier League             team          24
## 10             2021     Aston Villa Premier League         opponent          24
## 11             2021        Atalanta        Serie A             team          30
## 12             2021        Atalanta        Serie A         opponent          30
## 13             2021   Athletic Club        La Liga             team          27
## 14             2021   Athletic Club        La Liga         opponent          27
## 15             2021 Atlético Madrid        La Liga             team          25
## 16             2021 Atlético Madrid        La Liga         opponent          25
## 17             2021        Augsburg     Bundesliga             team          25
```

```
## 18           2021        Augsburg     Bundesliga         opponent          25
## 19           2021       Barcelona        La Liga             team          25
## 20           2021       Barcelona        La Liga         opponent          25
## 21           2021   Bayern Munich     Bundesliga             team          29
## 22           2021   Bayern Munich     Bundesliga         opponent          29
## 23           2021       Benevento        Serie A             team          29
## 24           2021       Benevento        Serie A         opponent          29
## 25           2021           Betis        La Liga             team          25
## 26           2021           Betis        La Liga         opponent          25
## 27           2021         Bologna        Serie A             team          37
## 28           2021         Bologna        Serie A         opponent          37
## 29           2021        Bordeaux        Ligue 1             team          28
## 30           2021        Bordeaux        Ligue 1         opponent          28
## 31           2021           Brest        Ligue 1             team          26
## 32           2021           Brest        Ligue 1         opponent          26
## 33           2021        Brighton Premier League             team          27
## 34           2021        Brighton Premier League         opponent          27
## 35           2021         Burnley Premier League             team          25
## 36           2021         Burnley Premier League         opponent          25
## 37           2021           Cádiz        La Liga             team          34
## 38           2021           Cádiz        La Liga         opponent          34
## 39           2021        Cagliari        Serie A             team          32
## 40           2021        Cagliari        Serie A         opponent          32
## 41           2021      Celta Vigo        La Liga             team          30
## 42           2021      Celta Vigo        La Liga         opponent          30
## 43           2021         Chelsea Premier League             team          27
## 44           2021         Chelsea Premier League         opponent          27
## 45           2021         Crotone        Serie A             team          31
## 46           2021         Crotone        Serie A         opponent          31
## 47           2021  Crystal Palace Premier League             team          24
## 48           2021  Crystal Palace Premier League         opponent          24
## 49           2021           Dijon        Ligue 1             team          33
## 50           2021           Dijon        Ligue 1         opponent          33
## 51           2021        Dortmund     Bundesliga             team          26
## 52           2021        Dortmund     Bundesliga         opponent          26
## 53           2021           Eibar        La Liga             team          30
## 54           2021           Eibar        La Liga         opponent          30
## 55           2021  Eint Frankfurt     Bundesliga             team          25
## 56           2021  Eint Frankfurt     Bundesliga         opponent          25
## 57           2021           Elche        La Liga             team          30
## 58           2021           Elche        La Liga         opponent          30
## 59           2021         Everton Premier League             team          29
## 60           2021         Everton Premier League         opponent          29
## 61           2021      Fiorentina        Serie A             team          30
## 62           2021      Fiorentina        Serie A         opponent          30
## 63           2021        Freiburg     Bundesliga             team          24
```

```
## 64            2021         Freiburg     Bundesliga         opponent          24
## 65            2021           Fulham Premier League             team          28
## 66            2021           Fulham Premier League         opponent          28
## 67            2021            Genoa        Serie A             team          34
## 68            2021            Genoa        Serie A         opponent          34
## 69            2021           Getafe       La Liga             team          31
## 70            2021           Getafe       La Liga         opponent          31
## 71            2021          Granada       La Liga             team          34
## 72            2021          Granada       La Liga         opponent          34
## 73            2021    Hellas Verona        Serie A             team          32
## 74            2021    Hellas Verona        Serie A         opponent          32
## 75            2021       Hertha BSC     Bundesliga             team          32
## 76            2021       Hertha BSC     Bundesliga         opponent          32
## 77            2021       Hoffenheim     Bundesliga             team          30
## 78            2021       Hoffenheim     Bundesliga         opponent          30
## 79            2021           Huesca       La Liga             team          27
## 80            2021           Huesca       La Liga         opponent          27
## 81            2021            Inter        Serie A             team          25
## 82            2021            Inter        Serie A         opponent          25
## 83            2021         Juventus        Serie A             team          30
## 84            2021         Juventus        Serie A         opponent          30
## 85            2021             Köln     Bundesliga             team          28
## 86            2021             Köln     Bundesliga         opponent          28
## 87            2021            Lazio        Serie A             team          30
## 88            2021            Lazio        Serie A         opponent          30
## 89            2021     Leeds United Premier League             team          23
## 90            2021     Leeds United Premier League         opponent          23
## 91            2021   Leicester City Premier League             team          27
## 92            2021   Leicester City Premier League         opponent          27
## 93            2021             Lens        Ligue 1             team          27
## 94            2021             Lens        Ligue 1         opponent          27
## 95            2021          Levante       La Liga             team          29
## 96            2021          Levante       La Liga         opponent          29
## 97            2021       Leverkusen     Bundesliga             team          30
## 98            2021       Leverkusen     Bundesliga         opponent          30
## 99            2021            Lille        Ligue 1             team          21
## 100           2021            Lille        Ligue 1         opponent          21
## 101           2021        Liverpool Premier League             team          28
## 102           2021        Liverpool Premier League         opponent          28
## 103           2021          Lorient        Ligue 1             team          27
## 104           2021          Lorient        Ligue 1         opponent          27
## 105           2021             Lyon        Ligue 1             team          29
## 106           2021             Lyon        Ligue 1         opponent          29
## 107           2021       M'Gladbach     Bundesliga             team          23
## 108           2021       M'Gladbach     Bundesliga         opponent          23
## 109           2021         Mainz 05     Bundesliga             team          31
```

```
## 110            2021        Mainz 05     Bundesliga         opponent          31
## 111            2021 Manchester City Premier League             team          24
## 112            2021 Manchester City Premier League         opponent          24
## 113            2021  Manchester Utd Premier League             team          29
## 114            2021  Manchester Utd Premier League         opponent          29
## 115            2021       Marseille        Ligue 1             team          31
## 116            2021       Marseille        Ligue 1         opponent          31
## 117            2021            Metz        Ligue 1             team          27
## 118            2021            Metz        Ligue 1         opponent          27
## 119            2021           Milan        Serie A             team          29
## 120            2021           Milan        Serie A         opponent          29
## 121            2021          Monaco        Ligue 1             team          30
## 122            2021          Monaco        Ligue 1         opponent          30
## 123            2021     Montpellier        Ligue 1             team          25
## 124            2021     Montpellier        Ligue 1         opponent          25
## 125            2021          Nantes        Ligue 1             team          25
## 126            2021          Nantes        Ligue 1         opponent          25
## 127            2021          Napoli        Serie A             team          25
## 128            2021          Napoli        Serie A         opponent          25
## 129            2021   Newcastle Utd Premier League             team          27
## 130            2021   Newcastle Utd Premier League         opponent          27
## 131            2021            Nice        Ligue 1             team          31
## 132            2021            Nice        Ligue 1         opponent          31
## 133            2021           Nîmes        Ligue 1             team          33
## 134            2021           Nîmes        Ligue 1         opponent          33
## 135            2021         Osasuna        La Liga             team          28
## 136            2021         Osasuna        La Liga         opponent          28
## 137            2021       Paris S-G        Ligue 1             team          33
## 138            2021       Paris S-G        Ligue 1         opponent          33
## 139            2021           Parma        Serie A             team          42
## 140            2021           Parma        Serie A         opponent          42
## 141            2021      RB Leipzig     Bundesliga             team          25
## 142            2021      RB Leipzig     Bundesliga         opponent          25
## 143            2021     Real Madrid        La Liga             team          30
## 144            2021     Real Madrid        La Liga         opponent          30
## 145            2021   Real Sociedad        La Liga             team          30
## 146            2021   Real Sociedad        La Liga         opponent          30
## 147            2021           Reims        Ligue 1             team          29
## 148            2021           Reims        Ligue 1         opponent          29
## 149            2021          Rennes        Ligue 1             team          32
## 150            2021          Rennes        Ligue 1         opponent          32
## 151            2021            Roma        Serie A             team          31
## 152            2021            Roma        Serie A         opponent          31
## 153            2021   Saint-Étienne        Ligue 1             team          40
## 154            2021   Saint-Étienne        Ligue 1         opponent          40
## 155            2021       Sampdoria        Serie A             team          27
```

```
## 156             2021        Sampdoria        Serie A         opponent           27
## 157             2021         Sassuolo        Serie A             team           28
## 158             2021         Sassuolo        Serie A         opponent           28
## 159             2021       Schalke 04     Bundesliga             team           42
## 160             2021       Schalke 04     Bundesliga         opponent           42
## 161             2021          Sevilla        La Liga             team           26
## 162             2021          Sevilla        La Liga         opponent           26
## 163             2021   Sheffield Utd Premier League             team           27
## 164             2021   Sheffield Utd Premier League         opponent           27
## 165             2021     Southampton Premier League             team           29
## 166             2021     Southampton Premier League         opponent           29
## 167             2021           Spezia        Serie A             team           35
## 168             2021           Spezia        Serie A         opponent           35
## 169             2021       Strasbourg        Ligue 1             team           26
## 170             2021       Strasbourg        Ligue 1         opponent           26
## 171             2021        Stuttgart     Bundesliga             team           29
## 172             2021        Stuttgart     Bundesliga         opponent           29
## 173             2021           Torino        Serie A             team           29
## 174             2021           Torino        Serie A         opponent           29
## 175             2021        Tottenham Premier League             team           24
## 176             2021        Tottenham Premier League         opponent           24
## 177             2021          Udinese        Serie A             team           33
## 178             2021          Udinese        Serie A         opponent           33
## 179             2021    Union Berlin     Bundesliga             team           27
## 180             2021    Union Berlin     Bundesliga         opponent           27
## 181             2021         Valencia        La Liga             team           30
## 182             2021         Valencia        La Liga         opponent           30
## 183             2021       Valladolid        La Liga             team           32
## 184             2021       Valladolid        La Liga         opponent           32
## 185             2021       Villarreal        La Liga             team           29
## 186             2021       Villarreal        La Liga         opponent           29
## 187             2021   Werder Bremen     Bundesliga             team           27
## 188             2021   Werder Bremen     Bundesliga         opponent           27
## 189             2021        West Brom Premier League             team           30
## 190             2021        West Brom Premier League         opponent           30
## 191             2021         West Ham Premier League             team           24
## 192             2021         West Ham Premier League         opponent           24
## 193             2021        Wolfsburg     Bundesliga             team           26
## 194             2021        Wolfsburg     Bundesliga         opponent           26
## 195             2021           Wolves Premier League             team           27
## 196             2021           Wolves Premier League         opponent           27
##      Age Poss MP_Playing Starts_Playing Min_Playing Mins_Per_90_Playing Gls Ast
## 1   28.7 42.7         38            418        3420                  38  35  21
## 2   27.5 57.3         38            418        3420                  38  57  44
## 3   27.9 46.2         38            418        3420                  38  39  23
## 4   26.0 53.8         38            418        3420                  38  57  41
```

```
## 5    26.0 42.2            34            374          3060                   34   23   16
## 6    25.9 57.8            34            374          3060                   34   51   36
## 7    25.9 53.8            38            418          3420                   38   53   38
## 8    26.7 46.2            38            418          3420                   38   35   25
## 9    25.2 48.1            38            418          3420                   38   52   38
## 10   26.7 51.9            38            418          3420                   38   45   30
## 11   27.3 54.9            38            418          3420                   38   89   65
## 12   27.0 45.1            38            418          3420                   38   46   36
## 13   26.6 49.8            38            418          3420                   38   43   31
## 14   27.6 50.2            38            418          3420                   38   41   29
## 15   27.3 52.4            38            418          3420                   38   65   54
## 16   27.5 47.6            38            418          3420                   38   22   15
## 17   27.0 41.3            34            374          3060                   34   35   23
## 18   26.0 58.7            34            374          3060                   34   51   37
## 19   26.4 65.8            38            418          3420                   38   80   51
## 20   27.4 34.2            38            418          3420                   38   36   20
## 21   27.3 61.1            34            374          3060                   34   98   75
## 22   25.9 38.9            34            374          3060                   34   44   41
## 23   28.7 42.3            38            418          3420                   38   40   26
## 24   27.2 57.7            38            418          3420                   38   69   52
## 25   27.9 53.9            38            418          3420                   38   50   35
## 26   27.4 46.1            38            418          3420                   38   47   31
## 27   27.2 50.9            38            418          3420                   38   49   38
## 28   27.3 49.1            38            418          3420                   38   63   40
## 29   27.6 49.7            38            418          3420                   38   41   32
## 30   25.8 50.3            38            418          3420                   38   55   37
## 31   24.9 49.3            38            418          3420                   38   47   31
## 32   26.0 50.7            38            418          3420                   38   64   43
## 33   25.8 51.3            38            418          3420                   38   39   24
## 34   26.4 48.7            38            418          3420                   38   44   30
## 35   28.3 41.7            38            418          3420                   38   32   20
## 36   26.6 58.3            38            418          3420                   38   54   47
## 37   28.8 34.3            38            418          3420                   38   33   16
## 38   27.5 65.7            38            418          3420                   38   57   41
## 39   26.5 44.6            38            418          3420                   38   43   30
## 40   27.3 55.4            38            418          3420                   38   59   45
## 41   26.6 53.7            38            418          3420                   38   55   39
## 42   27.4 46.3            38            418          3420                   38   56   37
## 43   26.0 61.4            38            418          3420                   38   56   38
## 44   26.3 38.6            38            418          3420                   38   35   26
## 45   27.8 46.4            38            418          3420                   38   44   27
## 46   27.1 53.6            38            418          3420                   38   90   68
## 47   29.1 40.1            38            418          3420                   38   39   29
## 48   26.4 59.9            38            418          3420                   38   66   47
## 49   25.7 46.2            38            418          3420                   38   25   15
## 50   26.0 53.8            38            418          3420                   38   69   44
```

```
## 51  25.6 59.8          34            374         3060                   34  74  57
## 52  26.0 40.2          34            374         3060                   34  45  33
## 53  28.2 49.6          38            418         3420                   38  29  20
## 54  27.5 50.4          38            418         3420                   38  52  40
## 55  27.2 54.0          34            374         3060                   34  63  52
## 56  25.8 46.0          34            374         3060                   34  51  40
## 57  28.1 47.6          38            418         3420                   38  33  25
## 58  27.4 52.4          38            418         3420                   38  54  43
## 59  26.3 46.5          38            418         3420                   38  45  32
## 60  26.4 53.5          38            418         3420                   38  47  33
## 61  26.6 46.2          38            418         3420                   38  44  31
## 62  27.4 53.8          38            418         3420                   38  57  39
## 63  26.2 47.0          34            374         3060                   34  51  34
## 64  26.0 53.0          34            374         3060                   34  50  35
## 65  25.2 49.9          38            418         3420                   38  26  18
## 66  26.6 50.1          38            418         3420                   38  52  34
## 67  28.3 44.7          38            418         3420                   38  46  32
## 68  27.2 55.3          38            418         3420                   38  56  40
## 69  26.7 42.8          38            418         3420                   38  27  14
## 70  27.4 57.2          38            418         3420                   38  40  26
## 71  27.4 41.0          38            418         3420                   38  46  30
## 72  27.6 59.0          38            418         3420                   38  64  41
## 73  26.2 49.8          38            418         3420                   38  38  30
## 74  27.0 50.2          38            418         3420                   38  46  31
## 75  25.4 49.4          34            374         3060                   34  40  29
## 76  26.1 50.6          34            374         3060                   34  51  35
## 77  25.7 51.1          34            374         3060                   34  51  38
## 78  26.0 48.9          34            374         3060                   34  52  38
## 79  28.0 48.3          38            418         3420                   38  32  23
## 80  27.3 51.7          38            418         3420                   38  51  37
## 81  27.7 52.4          38            418         3420                   38  84  62
## 82  26.9 47.6          38            418         3420                   38  34  19
## 83  27.7 57.1          38            418         3420                   38  76  59
## 84  27.0 42.9          38            418         3420                   38  36  21
## 85  24.9 46.3          36            374         3060                   34  33  23
## 86  26.2 53.7          36            374         3060                   34  60  43
## 87  29.4 53.1          38            418         3420                   38  59  40
## 88  26.9 46.9          38            418         3420                   38  53  33
## 89  26.1 57.6          38            418         3420                   38  60  45
## 90  26.5 42.4          38            418         3420                   38  52  36
## 91  26.5 54.6          38            418         3420                   38  64  45
## 92  26.5 45.4          38            418         3420                   38  48  29
## 93  26.4 51.4          38            418         3420                   38  53  35
## 94  26.0 48.6          38            418         3420                   38  51  25
## 95  28.0 52.4          38            418         3420                   38  45  35
## 96  27.4 47.6          38            418         3420                   38  55  42
```

```
## 97  25.4 59.9       34        374       3060          34  52  41
## 98  25.9 40.1       34        374       3060          34  37  27
## 99  25.4 53.9       38        418       3420          38  62  39
## 100 26.2 46.1       38        418       3420          38  21  17
## 101 26.8 62.4       38        418       3420          38  65  43
## 102 26.6 37.6       38        418       3420          38  42  29
## 103 26.8 44.6       38        418       3420          38  48  26
## 104 25.8 55.4       38        418       3420          38  67  44
## 105 25.8 54.9       38        418       3420          38  76  46
## 106 26.3 45.1       38        418       3420          38  41  24
## 107 26.9 51.9       34        374       3060          34  63  44
## 108 25.6 48.1       34        374       3060          34  56  44
## 109 25.8 40.5       34        374       3060          34  39  23
## 110 25.9 59.5       34        374       3060          34  55  35
## 111 26.1 63.9       38        418       3420          38  82  55
## 112 26.4 36.1       38        418       3420          38  31  17
## 113 25.6 55.8       38        418       3420          38  70  51
## 114 26.6 44.2       38        418       3420          38  42  28
## 115 27.2 53.9       38        418       3420          38  53  45
## 116 25.8 46.1       38        418       3420          38  46  36
## 117 25.7 45.3       38        418       3420          38  42  28
## 118 26.3 54.7       38        418       3420          38  46  25
## 119 24.4 51.9       38        418       3420          38  72  47
## 120 27.4 48.1       38        418       3420          38  40  26
## 121 24.8 56.0       38        418       3420          38  76  52
## 122 26.2 44.0       38        418       3420          38  41  25
## 123 28.0 45.3       38        418       3420          38  60  42
## 124 25.9 54.7       38        418       3420          38  60  39
## 125 25.8 43.7       40        418       3420          38  47  28
## 126 26.0 56.3       40        418       3420          38  54  29
## 127 27.1 55.2       38        418       3420          38  83  56
## 128 27.1 44.8       38        418       3420          38  40  26
## 129 27.1 38.2       38        418       3420          38  44  26
## 130 26.6 61.8       38        418       3420          38  59  41
## 131 23.5 54.5       38        418       3420          38  49  26
## 132 26.1 45.5       38        418       3420          38  53  28
## 133 25.8 44.4       38        418       3420          38  40  26
## 134 25.9 55.6       38        418       3420          38  71  52
## 135 27.9 43.1       38        418       3420          38  36  25
## 136 27.4 56.9       38        418       3420          38  45  31
## 137 26.7 63.5       38        418       3420          38  85  52
## 138 25.9 36.5       38        418       3420          38  27  17
## 139 28.0 47.7       38        418       3420          38  39  26
## 140 27.1 52.3       38        418       3420          38  80  62
## 141 25.0 60.0       34        374       3060          34  59  44
## 142 26.1 40.0       34        374       3060          34  32  27
```

```
## 143 27.9 59.7           38           418         3420                 38  64  53
## 144 27.5 40.3           38           418         3420                 38  27  17
## 145 25.0 55.1           38           418         3420                 38  58  39
## 146 27.4 44.9           38           418         3420                 38  37  24
## 147 24.6 44.8           38           418         3420                 38  39  23
## 148 26.0 55.2           38           418         3420                 38  49  34
## 149 25.5 59.5           38           418         3420                 38  50  39
## 150 26.1 40.5           38           418         3420                 38  40  30
## 151 26.7 52.5           38           418         3420                 38  67  51
## 152 27.3 47.5           38           418         3420                 38  52  36
## 153 25.4 48.3           38           418         3420                 38  40  28
## 154 25.8 51.7           38           418         3420                 38  52  34
## 155 27.9 45.1           38           418         3420                 38  52  37
## 156 27.1 54.9           38           418         3420                 38  53  35
## 157 26.5 60.8           38           418         3420                 38  61  36
## 158 27.2 39.2           38           418         3420                 38  55  46
## 159 25.9 44.6           34           374         3060                 34  24  20
## 160 26.0 55.4           34           374         3060                 34  81  65
## 161 28.1 61.2           38           418         3420                 38  52  36
## 162 27.3 38.8           38           418         3420                 38  31  20
## 163 26.7 41.5           38           418         3420                 38  19  13
## 164 26.7 58.5           38           418         3420                 38  60  45
## 165 26.6 52.2           38           418         3420                 38  47  33
## 166 26.5 47.8           38           418         3420                 38  67  44
## 167 25.2 52.2           38           418         3420                 38  52  37
## 168 27.3 47.8           38           418         3420                 38  72  50
## 169 26.8 45.6           38           418         3420                 38  49  26
## 170 26.1 54.4           38           418         3420                 38  57  38
## 171 24.1 51.5           34           374         3060                 34  54  42
## 172 26.1 48.5           34           374         3060                 34  54  38
## 173 26.9 47.2           38           418         3420                 38  49  28
## 174 27.1 52.8           38           418         3420                 38  68  48
## 175 27.2 51.7           38           418         3420                 38  66  50
## 176 26.5 48.3           38           418         3420                 38  42  24
## 177 27.2 46.1           38           418         3420                 38  40  30
## 178 27.1 53.9           38           418         3420                 38  58  35
## 179 27.3 44.9           34           374         3060                 34  50  39
## 180 26.0 55.1           34           374         3060                 34  41  31
## 181 25.2 46.9           38           418         3420                 38  48  35
## 182 27.6 53.1           38           418         3420                 38  51  29
## 183 28.1 45.1           38           418         3420                 38  34  21
## 184 27.4 54.9           38           418         3420                 38  57  44
## 185 28.0 55.6           38           418         3420                 38  57  33
## 186 27.5 44.4           38           418         3420                 38  42  25
## 187 26.3 43.8           34           374         3060                 34  35  25
## 188 26.0 56.2           34           374         3060                 34  54  41
```

```
## 189 26.4 37.6         38            418        3420                  38  33  20
## 190 26.6 62.4         38            418        3420                  38  73  52
## 191 27.8 42.9         38            418        3420                  38  60  46
## 192 26.6 57.1         38            418        3420                  38  43  34
## 193 25.6 51.4         34            374        3060                  34  57  48
## 194 26.1 48.6         34            374        3060                  34  36  27
## 195 26.3 49.3         38            418        3420                  38  34  21
## 196 26.6 50.7         38            418        3420                  38  49  34
##     G_minus_PK PK PKatt CrdY CrdR Gls_Per Ast_Per G+A_Per G_minus_PK_Per
## 1           30  5     8   96    8    0.92    0.55    1.47           0.79
## 2           54  3     3   80    3    1.50    1.16    2.66           1.42
## 3           34  5     6   65    2    1.03    0.61    1.63           0.89
## 4           50  7     8   73    3    1.50    1.08    2.58           1.32
## 5           22  1     2   52    1    0.68    0.47    1.15           0.65
## 6           46  5     6   63    2    1.50    1.06    2.56           1.35
## 7           47  6     6   49    5    1.39    1.00    2.39           1.24
## 8           33  2     3   74    2    0.92    0.66    1.58           0.87
## 9           47  5     6   71    4    1.37    1.00    2.37           1.24
## 10          40  5     6   73    7    1.18    0.79    1.97           1.05
## 11          83  6     7   67    3    2.34    1.71    4.05           2.18
## 12          41  5     8   87    4    1.21    0.95    2.16           1.08
## 13          39  4     5   82    3    1.13    0.82    1.95           1.03
## 14          36  5     6   80    5    1.08    0.76    1.84           0.95
## 15          60  5     7  100    0    1.71    1.42    3.13           1.58
## 16          21  1     4   59    3    0.58    0.39    0.97           0.55
## 17          33  2     5   70    4    1.03    0.68    1.71           0.97
## 18          45  6     9   54    1    1.50    1.09    2.59           1.32
## 19          76  4     8   72    2    2.11    1.34    3.45           2.00
## 20          31  5     5   76    6    0.95    0.53    1.47           0.82
## 21          89  9    10   44    3    2.88    2.21    5.09           2.62
## 22          43  1     3   50    0    1.29    1.21    2.50           1.26
## 23          35  5     7   95    5    1.05    0.68    1.74           0.92
## 24          65  4     5   76    6    1.82    1.37    3.18           1.71
## 25          42  8    11   93    8    1.32    0.92    2.24           1.11
## 26          39  8    10   83    2    1.24    0.82    2.05           1.03
## 27          46  3     4   82    4    1.29    1.00    2.29           1.21
## 28          56  7     9   84    1    1.66    1.05    2.71           1.47
## 29          37  4     4   78    4    1.08    0.84    1.92           0.97
## 30          48  7     8   66    3    1.45    0.97    2.42           1.26
## 31          44  3     4   62    4    1.24    0.82    2.05           1.16
## 32          55  9    11   78    5    1.68    1.13    2.82           1.45
## 33          33  6     9   49    6    1.03    0.63    1.66           0.87
## 34          38  6     7   51    2    1.16    0.79    1.95           1.00
## 35          29  3     3   48    0    0.84    0.53    1.37           0.76
## 36          51  3     3   32    2    1.42    1.24    2.66           1.34
## 37          29  4     4   81    3    0.87    0.42    1.29           0.76
```

```
## 38            48  9   10   87   4   1.50   1.08   2.58          1.26
## 39            39  4    5   77   3   1.13   0.79   1.92          1.03
## 40            54  5    5   76   5   1.55   1.18   2.74          1.42
## 41            48  7    7  109   5   1.45   1.03   2.47          1.26
## 42            49  7    8  102   6   1.47   0.97   2.45          1.29
## 43            48  8   10   51   3   1.47   1.00   2.47          1.26
## 44            32  3    4   56   3   0.92   0.68   1.61          0.84
## 45            35  9    9   89   4   1.16   0.71   1.87          0.92
## 46            85  5    5   90   5   2.37   1.79   4.16          2.24
## 47            36  3    4   56   2   1.03   0.76   1.79          0.95
## 48            63  3    4   44   4   1.74   1.24   2.97          1.66
## 49            22  3    5   76   5   0.66   0.39   1.05          0.58
## 50            62  7   10   54   1   1.82   1.16   2.97          1.63
## 51            71  3    7   44   1   2.18   1.68   3.85          2.09
## 52            39  6    6   62   1   1.32   0.97   2.29          1.15
## 53            25  4    9   71   3   0.76   0.53   1.29          0.66
## 54            49  3    5   66   2   1.37   1.05   2.42          1.29
## 55            55  8    8   82   1   1.85   1.53   3.38          1.62
## 56            47  4    6   64   0   1.50   1.18   2.68          1.38
## 57            31  2    3   99   3   0.87   0.66   1.53          0.82
## 58            48  6   10   79   2   1.42   1.13   2.55          1.26
## 59            41  4    5   59   2   1.18   0.84   2.03          1.08
## 60            44  3    4   66   1   1.24   0.87   2.11          1.16
## 61            38  6    6   86   5   1.16   0.82   1.97          1.00
## 62            49  8   11   81   2   1.50   1.03   2.53          1.29
## 63            45  6    6   62   0   1.50   1.00   2.50          1.32
## 64            46  4    6   54   0   1.47   1.03   2.50          1.35
## 65            23  3    6   67   3   0.68   0.47   1.16          0.61
## 66            45  7    7   40   1   1.37   0.89   2.26          1.18
## 67            42  4    4   87   2   1.21   0.84   2.05          1.11
## 68            49  7    7   69   1   1.47   1.05   2.53          1.29
## 69            24  3    5  120   7   0.71   0.37   1.08          0.63
## 70            33  7    8  106   6   1.05   0.68   1.74          0.87
## 71            43  3    5   97   6   1.21   0.79   2.00          1.13
## 72            58  6   12  104   6   1.68   1.08   2.76          1.53
## 73            35  3    3   91   1   1.00   0.79   1.79          0.92
## 74            40  6    7   81   2   1.21   0.82   2.03          1.05
## 75            36  4    6   67   3   1.18   0.85   2.03          1.06
## 76            44  7    8   68   3   1.50   1.03   2.53          1.29
## 77            46  5    6   68   4   1.50   1.12   2.62          1.35
## 78            46  6    9   47   2   1.53   1.12   2.65          1.35
## 79            29  3    4   69   2   0.84   0.61   1.45          0.76
## 80            46  5    5   64   2   1.34   0.97   2.32          1.21
## 81            76  8    9   61   2   2.21   1.63   3.84          2.00
## 82            29  5    7   84   3   0.89   0.50   1.39          0.76
## 83            68  8   10   80   6   2.00   1.55   3.55          1.79
```

```
## 84       30  6   9  85  2  0.95  0.55  1.50     0.79
## 85       29  4   6  63  1  0.97  0.68  1.65     0.85
## 86       52  8   9  67  0  1.76  1.26  3.03     1.53
## 87       53  6  10 107  5  1.55  1.05  2.61     1.39
## 88       43 10  10  87  3  1.39  0.87  2.26     1.13
## 89       56  4   4  61  1  1.58  1.18  2.76     1.47
## 90       44  8   9  49  2  1.37  0.95  2.32     1.16
## 91       54 10  12  61  0  1.68  1.18  2.87     1.42
## 92       44  4   4  61  1  1.26  0.76  2.03     1.16
## 93       43 10  12  94  7  1.39  0.92  2.32     1.13
## 94       40 11  11  76  7  1.34  0.66  2.00     1.05
## 95       42  3   6  68  1  1.18  0.92  2.11     1.11
## 96       52  3   4  87  7  1.45  1.11  2.55     1.37
## 97       49  3   3  58  0  1.53  1.21  2.74     1.44
## 98       36  1   2  50  3  1.09  0.79  1.88     1.06
## 99       57  5   5  68  2  1.63  1.03  2.66     1.50
## 100      20  1   2  76  7  0.55  0.45  1.00     0.53
## 101      59  6   6  40  0  1.71  1.13  2.84     1.55
## 102      38  4   8  56  2  1.11  0.76  1.87     1.00
## 103      39  9  10  69  3  1.26  0.68  1.95     1.03
## 104      57 10  11  77  6  1.76  1.16  2.92     1.50
## 105      65 11  11  66 10  2.00  1.21  3.21     1.71
## 106      36  5   7  77  8  1.08  0.63  1.71     0.95
## 107      53 10  11  61  2  1.85  1.29  3.15     1.56
## 108      50  6   6  75  3  1.65  1.29  2.94     1.47
## 109      34  5   5  63  1  1.15  0.68  1.82     1.00
## 110      49  6   7  73  4  1.62  1.03  2.65     1.44
## 111      77  5   9  46  2  2.16  1.45  3.61     2.03
## 112      23  8  10  61  2  0.82  0.45  1.26     0.61
## 113      60 10  11  64  1  1.84  1.34  3.18     1.58
## 114      38  4   4  64  3  1.11  0.74  1.84     1.00
## 115      49  4   7  98  9  1.39  1.18  2.58     1.29
## 116      43  3   3  81  6  1.21  0.95  2.16     1.13
## 117      35  7  11  84  4  1.11  0.74  1.84     0.92
## 118      34 12  13  76  6  1.21  0.66  1.87     0.89
## 119      57 15  20  83  4  1.89  1.24  3.13     1.50
## 120      37  3   5  83  4  1.05  0.68  1.74     0.97
## 121      64 12  15  78  7  2.00  1.37  3.37     1.68
## 122      36  5   5  89 12  1.08  0.66  1.74     0.95
## 123      58  2   5  65  7  1.58  1.11  2.68     1.53
## 124      48 12  12  79  4  1.58  1.03  2.61     1.26
## 125      39  8   9  70  4  1.24  0.74  1.97     1.03
## 126      44 10  12  88  3  1.42  0.76  2.18     1.16
## 127      76  7   8  76  3  2.18  1.47  3.66     2.00
## 128      35  5   5 107  6  1.05  0.68  1.74     0.92
## 129      38  6   7  65  3  1.16  0.68  1.84     1.00
```

```
## 130           55  4    5   60    3    1.55    1.08    2.63            1.45
## 131           42  7    7   73    4    1.29    0.68    1.97            1.11
## 132           46  7    7   66    1    1.39    0.74    2.13            1.21
## 133           33  7    8   60    5    1.05    0.68    1.74            0.87
## 134           65  6   10   87    9    1.87    1.37    3.24            1.71
## 135           31  5    6   80    5    0.95    0.66    1.61            0.82
## 136           40  5    8   65    0    1.18    0.82    2.00            1.05
## 137           73 12   13   79    7    2.24    1.37    3.61            1.92
## 138           26  1    2   59    9    0.71    0.45    1.16            0.68
## 139           34  5    5   91    1    1.03    0.68    1.71            0.89
## 140           77  3    3   71    1    2.11    1.63    3.74            2.03
## 141           51  8   10   57    0    1.74    1.29    3.03            1.50
## 142           29  3    3   64    2    0.94    0.79    1.74            0.85
## 143           61  3    3   59    2    1.68    1.39    3.08            1.61
## 144           20  7    8   87    3    0.71    0.45    1.16            0.53
## 145           49  9   12   81    1    1.53    1.03    2.55            1.29
## 146           34  3    5   91    6    0.97    0.63    1.61            0.89
## 147           34  5    5   77    8    1.03    0.61    1.63            0.89
## 148           44  5    8   58    3    1.29    0.89    2.18            1.16
## 149           47  3    3   79    5    1.32    1.03    2.34            1.24
## 150           37  3    4   70    6    1.05    0.79    1.84            0.97
## 151           60  7    8   91    3    1.76    1.34    3.11            1.58
## 152           47  5    6   80    6    1.37    0.95    2.32            1.24
## 153           34  6    9   76    2    1.05    0.74    1.79            0.89
## 154           46  6    9   77    1    1.37    0.89    2.26            1.21
## 155           47  5    6   83    3    1.37    0.97    2.34            1.24
## 156           46  7    9   77    1    1.39    0.92    2.32            1.21
## 157           48 13   14   79    4    1.61    0.95    2.55            1.26
## 158           51  4    5   82    4    1.45    1.21    2.66            1.34
## 159           24  0    3   74    2    0.71    0.59    1.29            0.71
## 160           74  7    8   59    2    2.38    1.91    4.29            2.18
## 161           46  6    7   79    2    1.37    0.95    2.32            1.21
## 162           27  4    6  109    1    0.82    0.53    1.34            0.71
## 163           16  3    4   73    3    0.50    0.34    0.84            0.42
## 164           57  3    4   39    1    1.58    1.18    2.76            1.50
## 165           42  5    6   52    3    1.24    0.87    2.11            1.11
## 166           59  8    9   58    3    1.76    1.16    2.92            1.55
## 167           48  4    5  100    5    1.37    0.97    2.34            1.26
## 168           63  9   10   83    3    1.89    1.32    3.21            1.66
## 169           39 10   12   58    3    1.29    0.68    1.97            1.03
## 170           51  6    8   68    2    1.50    1.00    2.50            1.34
## 171           48  6    8   65    2    1.59    1.24    2.82            1.41
## 172           49  5    7   69    3    1.59    1.12    2.71            1.44
## 173           44  5    6   78    4    1.29    0.74    2.03            1.16
## 174           59  9   10  101    6    1.79    1.26    3.05            1.55
## 175           61  5    5   57    2    1.74    1.32    3.05            1.61
```

```
## 176          32 10    10   77    1    1.11     0.63     1.74            0.84
## 177          36  4     4   65    2    1.05     0.79     1.84            0.95
## 178          44 14    14   84    4    1.53     0.92     2.45            1.16
## 179          45  5     7   57    2    1.47     1.15     2.62            1.32
## 180          37  4     5   64    2    1.21     0.91     2.12            1.09
## 181          40  8    10   82    5    1.26     0.92     2.18            1.05
## 182          46  5     8   99    5    1.34     0.76     2.11            1.21
## 183          27  7     8   93    4    0.89     0.55     1.45            0.71
## 184          51  6     7   96    2    1.50     1.16     2.66            1.34
## 185          45 12    13   67    5    1.50     0.87     2.37            1.18
## 186          35  7     9   78    4    1.11     0.66     1.76            0.92
## 187          30  5     5   67    3    1.03     0.74     1.76            0.88
## 188          50  4     6   68    3    1.59     1.21     2.79            1.47
## 189          29  4     4   51    4    0.87     0.53     1.39            0.76
## 190          66  7    11   50    1    1.92     1.37     3.29            1.74
## 191          58  2     4   50    3    1.58     1.21     2.79            1.53
## 192          40  3     5   52    1    1.13     0.89     2.03            1.05
## 193          55  2     3   58    3    1.68     1.41     3.09            1.62
## 194          33  3     5   61    2    1.06     0.79     1.85            0.97
## 195          30  4     4   55    1    0.89     0.55     1.45            0.79
## 196          42  7     8   62    6    1.29     0.89     2.18            1.11
##     G+A_minus_PK_Per xG_Expected npxG_Expected xA_Expected npxG+xA_Expected
## 1               1.34        42.4          36.5        25.5             62.0
## 2               2.58        50.8          48.5        37.0             85.5
## 3               1.50        41.7          37.2        27.4             64.6
## 4               2.39        50.9          45.6        35.0             80.6
## 5               1.12        30.7          29.2        22.1             51.3
## 6               2.41        55.7          51.2        40.8             92.1
## 7               2.24        53.5          49.0        36.6             85.5
## 8               1.53        44.3          42.1        29.9             71.9
## 9               2.24        52.9          48.5        37.1             85.6
## 10              1.84        52.9          48.4        32.7             81.1
## 11              3.89        77.2          72.0        55.9            128.0
## 12              2.03        42.3          36.2        25.6             61.8
## 13              1.84        44.4          40.7        31.1             71.8
## 14              1.71        36.3          31.9        24.6             56.5
## 15              3.00        52.4          46.8        37.8             84.6
## 16              0.95        32.7          29.7        23.2             52.9
## 17              1.65        36.1          32.3        24.5             56.9
## 18              2.41        55.6          49.0        37.2             86.2
## 19              3.34        78.9          72.6        57.4            130.0
## 20              1.34        39.6          35.8        24.2             60.0
## 21              4.82        75.8          68.2        53.3            121.5
## 22              2.47        41.0          38.7        31.7             70.4
## 23              1.61        41.6          36.2        27.3             63.5
## 24              3.08        66.1          62.5        44.0            106.5
```

```
## 25                  2.03        46.0          37.8        29.9              67.7
## 26                  1.84        45.5          38.0        29.6              67.7
## 27                  2.21        53.4          50.5        37.4              87.9
## 28                  2.53        64.7          58.4        42.7             101.0
## 29                  1.82        41.9          38.9        30.2              69.1
## 30                  2.24        56.7          50.8        38.9              89.7
## 31                  1.97        47.5          44.4        33.7              78.2
## 32                  2.58        55.6          47.3        34.9              82.2
## 33                  1.50        51.6          44.8        33.0              77.8
## 34                  1.79        37.7          32.4        23.9              56.3
## 35                  1.29        39.9          37.6        27.1              64.8
## 36                  2.58        57.6          55.4        43.6              99.0
## 37                  1.18        34.0          31.0        19.0              50.0
## 38                  2.34        55.0          47.5        38.1              85.6
## 39                  1.82        45.5          41.8        30.6              72.4
## 40                  2.61        63.1          59.4        48.2             107.6
## 41                  2.29        47.7          42.4        33.2              75.6
## 42                  2.26        49.5          43.4        32.5              76.0
## 43                  2.26        64.0          56.4        42.4              98.8
## 44                  1.53        32.8          29.8        22.0              51.8
## 45                  1.63        38.5          31.8        24.5              56.3
## 46                  4.03        71.6          67.9        51.6             119.5
## 47                  1.71        32.4          29.5        20.9              50.4
## 48                  2.89        57.5          54.4        39.8              94.2
## 49                  0.97        31.5          27.7        20.3              48.0
## 50                  2.79        75.3          67.7        48.8             116.6
## 51                  3.76        68.6          62.9        51.0             113.9
## 52                  2.12        40.4          35.8        25.7              61.5
## 53                  1.18        41.5          34.7        26.2              60.8
## 54                  2.34        47.0          43.2        33.8              77.0
## 55                  3.15        56.1          50.0        43.1              93.1
## 56                  2.56        48.4          43.6        33.3              76.9
## 57                  1.47        30.5          28.2        21.9              50.1
## 58                  2.39        58.4          51.2        43.2              94.4
## 59                  1.92        47.1          43.4        32.9              76.3
## 60                  2.03        52.0          48.9        33.5              82.5
## 61                  1.82        45.2          40.7        30.5              71.2
## 62                  2.32        55.1          46.6        38.9              85.5
## 63                  2.32        46.2          41.7        29.2              70.9
## 64                  2.38        53.7          49.2        37.3              86.5
## 65                  1.08        41.3          36.8        27.8              64.6
## 66                  2.08        52.9          47.7        36.2              83.9
## 67                  1.95        37.8          34.8        25.1              59.9
## 68                  2.34        54.5          49.4        38.3              87.7
## 69                  1.00        32.7          28.9        20.1              48.9
## 70                  1.55        40.3          34.1        25.0              59.1
```

```
## 71              1.92        39.5          35.9        23.8             59.6
## 72              2.61        53.7          44.8        34.6             79.4
## 73              1.71        43.5          41.3        31.1             72.4
## 74              1.87        50.3          45.2        33.8             79.0
## 75              1.91        42.4          37.9        29.8             67.7
## 76              2.32        47.0          40.9        32.1             73.0
## 77              2.47        51.8          47.3        35.9             83.2
## 78              2.47        53.3          46.7        37.5             84.2
## 79              1.37        35.3          32.5        23.8             56.2
## 80              2.18        44.3          40.7        29.5             70.2
## 81              3.63        74.8          68.0        53.4            121.4
## 82              1.26        38.6          33.0        24.7             57.6
## 83              3.34        74.2          66.7        51.4            118.1
## 84              1.34        38.2          31.5        23.1             54.6
## 85              1.53        39.1          34.7        24.6             59.3
## 86              2.79        55.3          48.5        37.9             86.4
## 87              2.45        59.8          52.4        40.2             92.7
## 88              2.00        47.8          40.3        28.4             68.7
## 89              2.66        57.5          54.5        42.5             97.0
## 90              2.11        62.9          56.0        41.6             97.6
## 91              2.61        56.0          46.9        33.1             80.0
## 92              1.92        47.7          44.7        34.2             78.9
## 93              2.05        56.0          46.9        34.7             81.6
## 94              1.71        47.6          39.2        27.3             66.5
## 95              2.03        41.0          36.4        28.5             64.9
## 96              2.47        53.9          51.0        38.7             89.8
## 97              2.65        47.2          45.0        33.8             78.8
## 98              1.85        40.4          38.8        27.7             66.5
## 99              2.53        50.0          46.2        36.1             82.4
## 100             0.97        26.7          25.2        19.4             44.6
## 101             2.68        72.6          68.1        50.0            118.1
## 102             1.76        45.3          39.3        29.7             69.1
## 103             1.71        48.8          41.6        31.2             72.7
## 104             2.66        53.1          44.9        33.2             78.1
## 105             2.92        84.6          76.4        55.0            131.5
## 106             1.58        43.2          37.9        25.1             63.1
## 107             2.85        55.2          47.0        34.6             81.7
## 108             2.76        44.7          40.1        30.8             70.9
## 109             1.68        47.1          43.4        34.7             78.1
## 110             2.47        50.9          45.8        36.0             81.7
## 111             3.47        73.3          66.6        51.0            117.5
## 112             1.05        31.4          23.7        15.6             39.2
## 113             2.92        60.2          51.9        40.5             92.4
## 114             1.74        42.2          39.3        28.2             67.5
## 115             2.47        47.5          42.2        34.5             76.7
## 116             2.08        46.0          43.7        32.6             76.3
```

```
## 117                1.66        41.3         33.2        25.6          58.8
## 118                1.55        49.3         39.6        30.2          69.7
## 119                2.74        71.2         56.6        43.4         100.0
## 120                1.66        45.7         42.0        31.5          73.6
## 121                3.05        66.9         55.4        44.7         100.1
## 122                1.61        33.1         29.4        20.6          50.0
## 123                2.63        47.9         44.2        30.7          74.9
## 124                2.29        65.7         56.6        44.8         101.4
## 125                1.76        48.9         42.1        26.7          68.8
## 126                1.92        55.8         46.9        33.6          80.4
## 127                3.47        67.7         61.5        45.5         107.0
## 128                1.61        41.6         37.8        26.4          64.2
## 129                1.68        41.0         35.5        26.1          61.6
## 130                2.53        54.0         50.2        37.9          88.1
## 131                1.79        46.9         41.7        26.3          68.1
## 132                1.95        52.1         46.8        34.3          81.2
## 133                1.55        41.4         35.8        27.3          63.1
## 134                3.08        64.9         57.1        44.2         101.3
## 135                1.47        37.6         33.2        23.4          56.7
## 136                1.87        51.3         45.4        34.8          80.2
## 137                3.29        82.4         73.1        54.6         127.7
## 138                1.13        37.3         35.8        27.1          62.8
## 139                1.58        40.0         36.2        28.6          64.9
## 140                3.66        61.4         59.4        45.3         104.7
## 141                2.79        65.3         58.0        44.7         102.7
## 142                1.65        29.1         26.8        22.1          48.9
## 143                3.00        61.5         59.2        46.0         105.3
## 144                0.97        36.5         30.4        23.8          54.2
## 145                2.32        60.4         51.6        39.2          90.7
## 146                1.53        36.6         32.8        24.4          57.3
## 147                1.50        33.3         29.7        20.9          50.5
## 148                2.05        61.1         55.4        40.3          95.6
## 149                2.26        50.0         47.7        37.5          85.2
## 150                1.76        40.0         37.0        26.2          63.2
## 151                2.92        68.9         62.9        49.2         112.2
## 152                2.18        51.4         46.7        34.0          80.7
## 153                1.63        48.2         41.4        30.8          72.2
## 154                2.11        48.5         41.7        30.5          72.2
## 155                2.21        46.2         41.7        30.0          71.7
## 156                2.13        56.9         50.3        35.7          86.0
## 157                2.21        59.2         48.8        35.5          84.4
## 158                2.55        54.9         51.1        38.8          89.9
## 159                1.29        27.0         24.5        18.3          42.9
## 160                4.09        68.0         62.1        48.9         111.0
## 161                2.16        50.9         45.6        34.8          80.4
## 162                1.24        35.0         30.6        20.5          51.1
```

```
## 163             0.76        31.4          28.4        22.6           51.0
## 164             2.68        62.4          59.5        43.4          102.8
## 165             1.97        42.4          37.9        27.4           65.3
## 166             2.71        54.2          47.4        36.3           83.7
## 167             2.24        46.7          42.9        28.9           71.8
## 168             2.97        69.3          61.8        48.9          110.7
## 169             1.71        52.6          43.2        30.0           73.2
## 170             2.34        46.6          40.6        31.2           71.8
## 171             2.65        53.2          47.1        38.1           85.1
## 172             2.56        53.3          48.0        35.4           83.4
## 173             1.89        50.2          45.5        32.5           78.1
## 174             2.82        59.0          51.5        37.3           88.8
## 175             2.92        54.5          50.7        35.5           86.2
## 176             1.47        49.5          41.9        33.0           75.0
## 177             1.74        42.9          40.1        27.4           67.4
## 178             2.08        52.1          41.6        31.4           73.0
## 179             2.47        45.8          40.4        31.1           71.5
## 180             2.00        41.3          37.3        28.9           66.2
## 181             1.97        43.9          36.4        27.2           63.6
## 182             1.97        53.9          47.7        32.5           80.2
## 183             1.26        37.1          31.2        22.1           53.3
## 184             2.50        51.8          46.5        35.7           82.1
## 185             2.05        58.0          48.3        40.5           88.9
## 186             1.58        43.5          36.7        25.4           62.0
## 187             1.62        33.8          30.0        21.5           51.5
## 188             2.68        51.0          46.3        35.2           81.5
## 189             1.29        33.8          30.7        21.2           52.0
## 190             3.11        67.7          59.4        44.1          103.5
## 191             2.74        53.9          50.7        36.8           87.5
## 192             1.95        48.3          44.5        35.3           79.8
## 193             3.03        49.5          47.4        37.6           85.0
## 194             1.76        42.0          38.1        29.5           67.6
## 195             1.34        39.9          36.8        26.5           63.3
## 196             2.00        45.8          39.8        30.2           69.9
##     xG_Per xA_Per xG+xA_Per npxG_Per npxG+xA_Per
## 1     1.12   0.67      1.79     0.96        1.63
## 2     1.34   0.97      2.31     1.28        2.25
## 3     1.10   0.72      1.82     0.98        1.70
## 4     1.34   0.92      2.26     1.20        2.12
## 5     0.90   0.65      1.55     0.86        1.51
## 6     1.64   1.20      2.84     1.51        2.71
## 7     1.41   0.96      2.37     1.29        2.25
## 8     1.17   0.79      1.95     1.11        1.89
## 9     1.39   0.98      2.37     1.28        2.25
## 10    1.39   0.86      2.25     1.27        2.13
## 11    2.03   1.47      3.50     1.90        3.37
```

```
## 12    1.11   0.67      1.79     0.95         1.63
## 13    1.17   0.82      1.99     1.07         1.89
## 14    0.96   0.65      1.60     0.84         1.49
## 15    1.38   0.99      2.37     1.23         2.23
## 16    0.86   0.61      1.47     0.78         1.39
## 17    1.06   0.72      1.78     0.95         1.67
## 18    1.63   1.09      2.73     1.44         2.54
## 19    2.08   1.51      3.59     1.91         3.42
## 20    1.04   0.64      1.68     0.94         1.58
## 21    2.23   1.57      3.80     2.01         3.57
## 22    1.21   0.93      2.14     1.14         2.07
## 23    1.10   0.72      1.81     0.95         1.67
## 24    1.74   1.16      2.90     1.64         2.80
## 25    1.21   0.79      2.00     1.00         1.78
## 26    1.20   0.78      1.98     1.00         1.78
## 27    1.41   0.98      2.39     1.33         2.31
## 28    1.70   1.12      2.82     1.54         2.66
## 29    1.10   0.80      1.90     1.02         1.82
## 30    1.49   1.02      2.52     1.34         2.36
## 31    1.25   0.89      2.14     1.17         2.06
## 32    1.46   0.92      2.38     1.24         2.16
## 33    1.36   0.87      2.23     1.18         2.05
## 34    0.99   0.63      1.62     0.85         1.48
## 35    1.05   0.71      1.76     0.99         1.70
## 36    1.52   1.15      2.66     1.46         2.60
## 37    0.89   0.50      1.39     0.82         1.31
## 38    1.45   1.00      2.45     1.25         2.25
## 39    1.20   0.81      2.00     1.10         1.91
## 40    1.66   1.27      2.93     1.56         2.83
## 41    1.25   0.87      2.13     1.11         1.99
## 42    1.30   0.86      2.16     1.14         2.00
## 43    1.68   1.12      2.80     1.48         2.60
## 44    0.86   0.58      1.44     0.78         1.36
## 45    1.01   0.64      1.66     0.84         1.48
## 46    1.89   1.36      3.24     1.79         3.14
## 47    0.85   0.55      1.40     0.78         1.33
## 48    1.51   1.05      2.56     1.43         2.48
## 49    0.83   0.53      1.36     0.73         1.26
## 50    1.98   1.29      3.27     1.78         3.07
## 51    2.02   1.50      3.52     1.85         3.35
## 52    1.19   0.76      1.94     1.05         1.81
## 53    1.09   0.69      1.78     0.91         1.60
## 54    1.24   0.89      2.13     1.14         2.03
## 55    1.65   1.27      2.92     1.47         2.74
## 56    1.42   0.98      2.40     1.28         2.26
## 57    0.80   0.58      1.38     0.74         1.32
```

```
## 58    1.54   1.14       2.67      1.35          2.48
## 59    1.24   0.87       2.11      1.14          2.01
## 60    1.37   0.88       2.25      1.29          2.17
## 61    1.19   0.80       1.99      1.07          1.87
## 62    1.45   1.02       2.47      1.23          2.25
## 63    1.36   0.86       2.22      1.23          2.09
## 64    1.58   1.10       2.68      1.45          2.54
## 65    1.09   0.73       1.82      0.97          1.70
## 66    1.39   0.95       2.34      1.26          2.21
## 67    0.99   0.66       1.66      0.92          1.58
## 68    1.43   1.01       2.44      1.30          2.31
## 69    0.86   0.53       1.39      0.76          1.29
## 70    1.06   0.66       1.72      0.90          1.56
## 71    1.04   0.63       1.67      0.94          1.57
## 72    1.41   0.91       2.32      1.18          2.09
## 73    1.15   0.82       1.97      1.09          1.91
## 74    1.32   0.89       2.21      1.19          2.08
## 75    1.25   0.88       2.13      1.11          1.99
## 76    1.38   0.94       2.33      1.20          2.15
## 77    1.52   1.06       2.58      1.39          2.45
## 78    1.57   1.10       2.67      1.37          2.48
## 79    0.93   0.63       1.55      0.85          1.48
## 80    1.17   0.78       1.94      1.07          1.85
## 81    1.97   1.40       3.37      1.79          3.19
## 82    1.02   0.65       1.66      0.87          1.52
## 83    1.95   1.35       3.30      1.76          3.11
## 84    1.01   0.61       1.61      0.83          1.44
## 85    1.15   0.72       1.88      1.02          1.74
## 86    1.63   1.12       2.74      1.43          2.54
## 87    1.57   1.06       2.63      1.38          2.44
## 88    1.26   0.75       2.00      1.06          1.81
## 89    1.51   1.12       2.63      1.43          2.55
## 90    1.66   1.10       2.75      1.47          2.57
## 91    1.47   0.87       2.35      1.23          2.11
## 92    1.25   0.90       2.16      1.18          2.08
## 93    1.47   0.91       2.39      1.23          2.15
## 94    1.25   0.72       1.97      1.03          1.75
## 95    1.08   0.75       1.83      0.96          1.71
## 96    1.42   1.02       2.44      1.34          2.36
## 97    1.39   0.99       2.38      1.32          2.32
## 98    1.19   0.81       2.00      1.14          1.96
## 99    1.32   0.95       2.27      1.22          2.17
## 100   0.70   0.51       1.21      0.66          1.17
## 101   1.91   1.32       3.23      1.79          3.11
## 102   1.19   0.78       1.97      1.04          1.82
## 103   1.29   0.82       2.11      1.09          1.91
```

```
## 104   1.40   0.87       2.27     1.18          2.06
## 105   2.23   1.45       3.67     2.01          3.46
## 106   1.14   0.66       1.80     1.00          1.66
## 107   1.62   1.02       2.64     1.38          2.40
## 108   1.31   0.91       2.22     1.18          2.09
## 109   1.39   1.02       2.41     1.28          2.30
## 110   1.50   1.06       2.55     1.35          2.40
## 111   1.93   1.34       3.27     1.75          3.09
## 112   0.83   0.41       1.24     0.62          1.03
## 113   1.58   1.07       2.65     1.37          2.43
## 114   1.11   0.74       1.85     1.03          1.78
## 115   1.25   0.91       2.16     1.11          2.02
## 116   1.21   0.86       2.07     1.15          2.01
## 117   1.09   0.67       1.76     0.87          1.55
## 118   1.30   0.79       2.09     1.04          1.83
## 119   1.87   1.14       3.02     1.49          2.63
## 120   1.20   0.83       2.03     1.11          1.94
## 121   1.76   1.18       2.94     1.46          2.63
## 122   0.87   0.54       1.41     0.77          1.32
## 123   1.26   0.81       2.07     1.16          1.97
## 124   1.73   1.18       2.91     1.49          2.67
## 125   1.29   0.70       1.99     1.11          1.81
## 126   1.47   0.88       2.35     1.23          2.12
## 127   1.78   1.20       2.98     1.62          2.82
## 128   1.09   0.69       1.79     1.00          1.69
## 129   1.08   0.69       1.77     0.93          1.62
## 130   1.42   1.00       2.42     1.32          2.32
## 131   1.24   0.69       1.93     1.10          1.79
## 132   1.37   0.90       2.27     1.23          2.14
## 133   1.09   0.72       1.81     0.94          1.66
## 134   1.71   1.16       2.87     1.50          2.67
## 135   0.99   0.62       1.61     0.87          1.49
## 136   1.35   0.92       2.27     1.19          2.11
## 137   2.17   1.44       3.61     1.92          3.36
## 138   0.98   0.71       1.69     0.94          1.65
## 139   1.05   0.75       1.81     0.95          1.71
## 140   1.62   1.19       2.81     1.56          2.76
## 141   1.92   1.32       3.24     1.70          3.02
## 142   0.86   0.65       1.51     0.79          1.44
## 143   1.62   1.21       2.83     1.56          2.77
## 144   0.96   0.63       1.59     0.80          1.43
## 145   1.59   1.03       2.62     1.36          2.39
## 146   0.96   0.64       1.61     0.86          1.51
## 147   0.88   0.55       1.43     0.78          1.33
## 148   1.61   1.06       2.67     1.46          2.52
## 149   1.31   0.99       2.30     1.25          2.24
```

```
## 150   1.05   0.69       1.74     0.97          1.66
## 151   1.81   1.30       3.11     1.66          2.95
## 152   1.35   0.89       2.25     1.23          2.12
## 153   1.27   0.81       2.08     1.09          1.90
## 154   1.28   0.80       2.08     1.10          1.90
## 155   1.21   0.79       2.00     1.10          1.89
## 156   1.50   0.94       2.44     1.32          2.26
## 157   1.56   0.94       2.49     1.29          2.22
## 158   1.44   1.02       2.47     1.34          2.37
## 159   0.79   0.54       1.33     0.72          1.26
## 160   2.00   1.44       3.44     1.83          3.26
## 161   1.34   0.92       2.26     1.20          2.12
## 162   0.92   0.54       1.46     0.80          1.34
## 163   0.83   0.60       1.42     0.75          1.34
## 164   1.64   1.14       2.78     1.56          2.71
## 165   1.11   0.72       1.84     1.00          1.72
## 166   1.43   0.96       2.38     1.25          2.20
## 167   1.23   0.76       1.99     1.13          1.89
## 168   1.82   1.29       3.11     1.63          2.91
## 169   1.38   0.79       2.17     1.14          1.93
## 170   1.23   0.82       2.05     1.07          1.89
## 171   1.56   1.12       2.68     1.38          2.50
## 172   1.57   1.04       2.61     1.41          2.45
## 173   1.32   0.86       2.18     1.20          2.05
## 174   1.55   0.98       2.53     1.35          2.34
## 175   1.43   0.93       2.37     1.33          2.27
## 176   1.30   0.87       2.17     1.10          1.97
## 177   1.13   0.72       1.85     1.05          1.77
## 178   1.37   0.83       2.20     1.10          1.92
## 179   1.35   0.91       2.26     1.19          2.10
## 180   1.22   0.85       2.06     1.10          1.95
## 181   1.15   0.72       1.87     0.96          1.67
## 182   1.42   0.86       2.27     1.25          2.11
## 183   0.98   0.58       1.56     0.82          1.40
## 184   1.36   0.94       2.30     1.22          2.16
## 185   1.53   1.07       2.59     1.27          2.34
## 186   1.14   0.67       1.81     0.97          1.63
## 187   0.99   0.63       1.63     0.88          1.52
## 188   1.50   1.04       2.54     1.36          2.40
## 189   0.89   0.56       1.45     0.81          1.37
## 190   1.78   1.16       2.94     1.56          2.72
## 191   1.42   0.97       2.39     1.33          2.30
## 192   1.27   0.93       2.20     1.17          2.10
## 193   1.45   1.11       2.56     1.39          2.50
## 194   1.24   0.87       2.10     1.12          1.99
## 195   1.05   0.70       1.75     0.97          1.67
```

```
## 196   1.21   0.79      2.00     1.05        1.84
##                                                                           Url
## 1                    https://fbref.com/en/squads/8d6fd021/2020-2021/Alaves-Stats
## 2                    https://fbref.com/en/squads/8d6fd021/2020-2021/Alaves-Stats
## 3                    https://fbref.com/en/squads/69236f98/2020-2021/Angers-Stats
## 4                    https://fbref.com/en/squads/69236f98/2020-2021/Angers-Stats
## 5                  https://fbref.com/en/squads/247c4b67/2020-2021/Arminia-Stats
## 6                  https://fbref.com/en/squads/247c4b67/2020-2021/Arminia-Stats
## 7                  https://fbref.com/en/squads/18bb7c10/2020-2021/Arsenal-Stats
## 8                  https://fbref.com/en/squads/18bb7c10/2020-2021/Arsenal-Stats
## 9              https://fbref.com/en/squads/8602292d/2020-2021/Aston-Villa-Stats
## 10             https://fbref.com/en/squads/8602292d/2020-2021/Aston-Villa-Stats
## 11               https://fbref.com/en/squads/922493f3/2020-2021/Atalanta-Stats
## 12               https://fbref.com/en/squads/922493f3/2020-2021/Atalanta-Stats
## 13           https://fbref.com/en/squads/2b390eca/2020-2021/Athletic-Club-Stats
## 14           https://fbref.com/en/squads/2b390eca/2020-2021/Athletic-Club-Stats
## 15         https://fbref.com/en/squads/db3b9613/2020-2021/Atletico-Madrid-Stats
## 16         https://fbref.com/en/squads/db3b9613/2020-2021/Atletico-Madrid-Stats
## 17               https://fbref.com/en/squads/0cdc4311/2020-2021/Augsburg-Stats
## 18               https://fbref.com/en/squads/0cdc4311/2020-2021/Augsburg-Stats
## 19              https://fbref.com/en/squads/206d90db/2020-2021/Barcelona-Stats
## 20              https://fbref.com/en/squads/206d90db/2020-2021/Barcelona-Stats
## 21          https://fbref.com/en/squads/054efa67/2020-2021/Bayern-Munich-Stats
## 22          https://fbref.com/en/squads/054efa67/2020-2021/Bayern-Munich-Stats
## 23              https://fbref.com/en/squads/4fcb34fd/2020-2021/Benevento-Stats
## 24              https://fbref.com/en/squads/4fcb34fd/2020-2021/Benevento-Stats
## 25             https://fbref.com/en/squads/fc536746/2020-2021/Real-Betis-Stats
## 26             https://fbref.com/en/squads/fc536746/2020-2021/Real-Betis-Stats
## 27                https://fbref.com/en/squads/1d8099f8/2020-2021/Bologna-Stats
## 28                https://fbref.com/en/squads/1d8099f8/2020-2021/Bologna-Stats
## 29               https://fbref.com/en/squads/123f3efe/2020-2021/Bordeaux-Stats
## 30               https://fbref.com/en/squads/123f3efe/2020-2021/Bordeaux-Stats
## 31                  https://fbref.com/en/squads/fb08dbb3/2020-2021/Brest-Stats
## 32                  https://fbref.com/en/squads/fb08dbb3/2020-2021/Brest-Stats
## 33  https://fbref.com/en/squads/d07537b9/2020-2021/Brighton-and-Hove-Albion-Stats
## 34  https://fbref.com/en/squads/d07537b9/2020-2021/Brighton-and-Hove-Albion-Stats
## 35                https://fbref.com/en/squads/943e8050/2020-2021/Burnley-Stats
## 36                https://fbref.com/en/squads/943e8050/2020-2021/Burnley-Stats
## 37                  https://fbref.com/en/squads/ee7c297c/2020-2021/Cadiz-Stats
## 38                  https://fbref.com/en/squads/ee7c297c/2020-2021/Cadiz-Stats
## 39              https://fbref.com/en/squads/c4260e09/2020-2021/Cagliari-Stats
## 40              https://fbref.com/en/squads/c4260e09/2020-2021/Cagliari-Stats
## 41             https://fbref.com/en/squads/f25da7fb/2020-2021/Celta-Vigo-Stats
## 42             https://fbref.com/en/squads/f25da7fb/2020-2021/Celta-Vigo-Stats
## 43                https://fbref.com/en/squads/cff3d9bb/2020-2021/Chelsea-Stats
## 44                https://fbref.com/en/squads/cff3d9bb/2020-2021/Chelsea-Stats
```

```
## 45              https://fbref.com/en/squads/3074d7b1/2020-2021/Crotone-Stats
## 46              https://fbref.com/en/squads/3074d7b1/2020-2021/Crotone-Stats
## 47       https://fbref.com/en/squads/47c64c55/2020-2021/Crystal-Palace-Stats
## 48       https://fbref.com/en/squads/47c64c55/2020-2021/Crystal-Palace-Stats
## 49                https://fbref.com/en/squads/8dfb7350/2020-2021/Dijon-Stats
## 50                https://fbref.com/en/squads/8dfb7350/2020-2021/Dijon-Stats
## 51             https://fbref.com/en/squads/add600ae/2020-2021/Dortmund-Stats
## 52             https://fbref.com/en/squads/add600ae/2020-2021/Dortmund-Stats
## 53                https://fbref.com/en/squads/bea5c710/2020-2021/Eibar-Stats
## 54                https://fbref.com/en/squads/bea5c710/2020-2021/Eibar-Stats
## 55    https://fbref.com/en/squads/f0ac8ee6/2020-2021/Eintracht-Frankfurt-Stats
## 56    https://fbref.com/en/squads/f0ac8ee6/2020-2021/Eintracht-Frankfurt-Stats
## 57                https://fbref.com/en/squads/6c8b07df/2020-2021/Elche-Stats
## 58                https://fbref.com/en/squads/6c8b07df/2020-2021/Elche-Stats
## 59              https://fbref.com/en/squads/d3fd31cc/2020-2021/Everton-Stats
## 60              https://fbref.com/en/squads/d3fd31cc/2020-2021/Everton-Stats
## 61           https://fbref.com/en/squads/421387cf/2020-2021/Fiorentina-Stats
## 62           https://fbref.com/en/squads/421387cf/2020-2021/Fiorentina-Stats
## 63             https://fbref.com/en/squads/a486e511/2020-2021/Freiburg-Stats
## 64             https://fbref.com/en/squads/a486e511/2020-2021/Freiburg-Stats
## 65               https://fbref.com/en/squads/fd962109/2020-2021/Fulham-Stats
## 66               https://fbref.com/en/squads/fd962109/2020-2021/Fulham-Stats
## 67                https://fbref.com/en/squads/658bf2de/2020-2021/Genoa-Stats
## 68                https://fbref.com/en/squads/658bf2de/2020-2021/Genoa-Stats
## 69               https://fbref.com/en/squads/7848bd64/2020-2021/Getafe-Stats
## 70               https://fbref.com/en/squads/7848bd64/2020-2021/Getafe-Stats
## 71              https://fbref.com/en/squads/a0435291/2020-2021/Granada-Stats
## 72              https://fbref.com/en/squads/a0435291/2020-2021/Granada-Stats
## 73        https://fbref.com/en/squads/0e72edf2/2020-2021/Hellas-Verona-Stats
## 74        https://fbref.com/en/squads/0e72edf2/2020-2021/Hellas-Verona-Stats
## 75           https://fbref.com/en/squads/2818f8bc/2020-2021/Hertha-BSC-Stats
## 76           https://fbref.com/en/squads/2818f8bc/2020-2021/Hertha-BSC-Stats
## 77           https://fbref.com/en/squads/033ea6b8/2020-2021/Hoffenheim-Stats
## 78           https://fbref.com/en/squads/033ea6b8/2020-2021/Hoffenheim-Stats
## 79               https://fbref.com/en/squads/c6c493e6/2020-2021/Huesca-Stats
## 80               https://fbref.com/en/squads/c6c493e6/2020-2021/Huesca-Stats
## 81       https://fbref.com/en/squads/d609edc0/2020-2021/Internazionale-Stats
## 82       https://fbref.com/en/squads/d609edc0/2020-2021/Internazionale-Stats
## 83             https://fbref.com/en/squads/e0652b02/2020-2021/Juventus-Stats
## 84             https://fbref.com/en/squads/e0652b02/2020-2021/Juventus-Stats
## 85                 https://fbref.com/en/squads/bc357bf7/2020-2021/Koln-Stats
## 86                 https://fbref.com/en/squads/bc357bf7/2020-2021/Koln-Stats
## 87                https://fbref.com/en/squads/7213da33/2020-2021/Lazio-Stats
## 88                https://fbref.com/en/squads/7213da33/2020-2021/Lazio-Stats
## 89         https://fbref.com/en/squads/5bfb9659/2020-2021/Leeds-United-Stats
## 90         https://fbref.com/en/squads/5bfb9659/2020-2021/Leeds-United-Stats
```

```
## 91           https://fbref.com/en/squads/a2d435b3/2020-2021/Leicester-City-Stats
## 92           https://fbref.com/en/squads/a2d435b3/2020-2021/Leicester-City-Stats
## 93                     https://fbref.com/en/squads/fd4e0f7d/2020-2021/Lens-Stats
## 94                     https://fbref.com/en/squads/fd4e0f7d/2020-2021/Lens-Stats
## 95                  https://fbref.com/en/squads/9800b6a1/2020-2021/Levante-Stats
## 96                  https://fbref.com/en/squads/9800b6a1/2020-2021/Levante-Stats
## 97         https://fbref.com/en/squads/c7a9f859/2020-2021/Bayer-Leverkusen-Stats
## 98         https://fbref.com/en/squads/c7a9f859/2020-2021/Bayer-Leverkusen-Stats
## 99                    https://fbref.com/en/squads/cb188c0c/2020-2021/Lille-Stats
## 100                   https://fbref.com/en/squads/cb188c0c/2020-2021/Lille-Stats
## 101               https://fbref.com/en/squads/822bd0ba/2020-2021/Liverpool-Stats
## 102               https://fbref.com/en/squads/822bd0ba/2020-2021/Liverpool-Stats
## 103                 https://fbref.com/en/squads/d2c87802/2020-2021/Lorient-Stats
## 104                 https://fbref.com/en/squads/d2c87802/2020-2021/Lorient-Stats
## 105                    https://fbref.com/en/squads/d53c0b06/2020-2021/Lyon-Stats
## 106                    https://fbref.com/en/squads/d53c0b06/2020-2021/Lyon-Stats
## 107         https://fbref.com/en/squads/32f3ee20/2020-2021/Monchengladbach-Stats
## 108         https://fbref.com/en/squads/32f3ee20/2020-2021/Monchengladbach-Stats
## 109                https://fbref.com/en/squads/a224b06a/2020-2021/Mainz-05-Stats
## 110                https://fbref.com/en/squads/a224b06a/2020-2021/Mainz-05-Stats
## 111         https://fbref.com/en/squads/b8fd03ef/2020-2021/Manchester-City-Stats
## 112         https://fbref.com/en/squads/b8fd03ef/2020-2021/Manchester-City-Stats
## 113       https://fbref.com/en/squads/19538871/2020-2021/Manchester-United-Stats
## 114       https://fbref.com/en/squads/19538871/2020-2021/Manchester-United-Stats
## 115               https://fbref.com/en/squads/5725cc7b/2020-2021/Marseille-Stats
## 116               https://fbref.com/en/squads/5725cc7b/2020-2021/Marseille-Stats
## 117                    https://fbref.com/en/squads/f83960ae/2020-2021/Metz-Stats
## 118                    https://fbref.com/en/squads/f83960ae/2020-2021/Metz-Stats
## 119                   https://fbref.com/en/squads/dc56fe14/2020-2021/Milan-Stats
## 120                   https://fbref.com/en/squads/dc56fe14/2020-2021/Milan-Stats
## 121                  https://fbref.com/en/squads/fd6114db/2020-2021/Monaco-Stats
## 122                  https://fbref.com/en/squads/fd6114db/2020-2021/Monaco-Stats
## 123             https://fbref.com/en/squads/281b0e73/2020-2021/Montpellier-Stats
## 124             https://fbref.com/en/squads/281b0e73/2020-2021/Montpellier-Stats
## 125                  https://fbref.com/en/squads/d7a486cd/2020-2021/Nantes-Stats
## 126                  https://fbref.com/en/squads/d7a486cd/2020-2021/Nantes-Stats
## 127                  https://fbref.com/en/squads/d48ad4ff/2020-2021/Napoli-Stats
## 128                  https://fbref.com/en/squads/d48ad4ff/2020-2021/Napoli-Stats
## 129        https://fbref.com/en/squads/b2b47a98/2020-2021/Newcastle-United-Stats
## 130        https://fbref.com/en/squads/b2b47a98/2020-2021/Newcastle-United-Stats
## 131                    https://fbref.com/en/squads/132ebc33/2020-2021/Nice-Stats
## 132                    https://fbref.com/en/squads/132ebc33/2020-2021/Nice-Stats
## 133                   https://fbref.com/en/squads/1cbf5f9e/2020-2021/Nimes-Stats
## 134                   https://fbref.com/en/squads/1cbf5f9e/2020-2021/Nimes-Stats
## 135                 https://fbref.com/en/squads/03c57e2b/2020-2021/Osasuna-Stats
## 136                 https://fbref.com/en/squads/03c57e2b/2020-2021/Osasuna-Stats
```

```
## 137    https://fbref.com/en/squads/e2d8892c/2020-2021/Paris-Saint-Germain-Stats
## 138    https://fbref.com/en/squads/e2d8892c/2020-2021/Paris-Saint-Germain-Stats
## 139                  https://fbref.com/en/squads/eab4234c/2020-2021/Parma-Stats
## 140                  https://fbref.com/en/squads/eab4234c/2020-2021/Parma-Stats
## 141             https://fbref.com/en/squads/acbb6a5b/2020-2021/RB-Leipzig-Stats
## 142             https://fbref.com/en/squads/acbb6a5b/2020-2021/RB-Leipzig-Stats
## 143            https://fbref.com/en/squads/53a2f082/2020-2021/Real-Madrid-Stats
## 144            https://fbref.com/en/squads/53a2f082/2020-2021/Real-Madrid-Stats
## 145          https://fbref.com/en/squads/e31d1cd9/2020-2021/Real-Sociedad-Stats
## 146          https://fbref.com/en/squads/e31d1cd9/2020-2021/Real-Sociedad-Stats
## 147                  https://fbref.com/en/squads/7fdd64e0/2020-2021/Reims-Stats
## 148                  https://fbref.com/en/squads/7fdd64e0/2020-2021/Reims-Stats
## 149                 https://fbref.com/en/squads/b3072e00/2020-2021/Rennes-Stats
## 150                 https://fbref.com/en/squads/b3072e00/2020-2021/Rennes-Stats
## 151                   https://fbref.com/en/squads/cf74a709/2020-2021/Roma-Stats
## 152                   https://fbref.com/en/squads/cf74a709/2020-2021/Roma-Stats
## 153          https://fbref.com/en/squads/d298ef2c/2020-2021/Saint-Etienne-Stats
## 154          https://fbref.com/en/squads/d298ef2c/2020-2021/Saint-Etienne-Stats
## 155              https://fbref.com/en/squads/8ff9e3b3/2020-2021/Sampdoria-Stats
## 156              https://fbref.com/en/squads/8ff9e3b3/2020-2021/Sampdoria-Stats
## 157               https://fbref.com/en/squads/e2befd26/2020-2021/Sassuolo-Stats
## 158               https://fbref.com/en/squads/e2befd26/2020-2021/Sassuolo-Stats
## 159             https://fbref.com/en/squads/c539e393/2020-2021/Schalke-04-Stats
## 160             https://fbref.com/en/squads/c539e393/2020-2021/Schalke-04-Stats
## 161                https://fbref.com/en/squads/ad2be733/2020-2021/Sevilla-Stats
## 162                https://fbref.com/en/squads/ad2be733/2020-2021/Sevilla-Stats
## 163       https://fbref.com/en/squads/1df6b87e/2020-2021/Sheffield-United-Stats
## 164       https://fbref.com/en/squads/1df6b87e/2020-2021/Sheffield-United-Stats
## 165            https://fbref.com/en/squads/33c895d4/2020-2021/Southampton-Stats
## 166            https://fbref.com/en/squads/33c895d4/2020-2021/Southampton-Stats
## 167                 https://fbref.com/en/squads/68449f6d/2020-2021/Spezia-Stats
## 168                 https://fbref.com/en/squads/68449f6d/2020-2021/Spezia-Stats
## 169             https://fbref.com/en/squads/c0d3eab4/2020-2021/Strasbourg-Stats
## 170             https://fbref.com/en/squads/c0d3eab4/2020-2021/Strasbourg-Stats
## 171              https://fbref.com/en/squads/598bc722/2020-2021/Stuttgart-Stats
## 172              https://fbref.com/en/squads/598bc722/2020-2021/Stuttgart-Stats
## 173                 https://fbref.com/en/squads/105360fe/2020-2021/Torino-Stats
## 174                 https://fbref.com/en/squads/105360fe/2020-2021/Torino-Stats
## 175      https://fbref.com/en/squads/361ca564/2020-2021/Tottenham-Hotspur-Stats
## 176      https://fbref.com/en/squads/361ca564/2020-2021/Tottenham-Hotspur-Stats
## 177                https://fbref.com/en/squads/04eea015/2020-2021/Udinese-Stats
## 178                https://fbref.com/en/squads/04eea015/2020-2021/Udinese-Stats
## 179           https://fbref.com/en/squads/7a41008f/2020-2021/Union-Berlin-Stats
## 180           https://fbref.com/en/squads/7a41008f/2020-2021/Union-Berlin-Stats
## 181               https://fbref.com/en/squads/dcc91a7b/2020-2021/Valencia-Stats
## 182               https://fbref.com/en/squads/dcc91a7b/2020-2021/Valencia-Stats
```

```
## 183               https://fbref.com/en/squads/17859612/2020-2021/Valladolid-Stats
## 184               https://fbref.com/en/squads/17859612/2020-2021/Valladolid-Stats
## 185               https://fbref.com/en/squads/2a8183b3/2020-2021/Villarreal-Stats
## 186               https://fbref.com/en/squads/2a8183b3/2020-2021/Villarreal-Stats
## 187            https://fbref.com/en/squads/62add3bf/2020-2021/Werder-Bremen-Stats
## 188            https://fbref.com/en/squads/62add3bf/2020-2021/Werder-Bremen-Stats
## 189     https://fbref.com/en/squads/60c6b05f/2020-2021/West-Bromwich-Albion-Stats
## 190     https://fbref.com/en/squads/60c6b05f/2020-2021/West-Bromwich-Albion-Stats
## 191          https://fbref.com/en/squads/7c21e445/2020-2021/West-Ham-United-Stats
## 192          https://fbref.com/en/squads/7c21e445/2020-2021/West-Ham-United-Stats
## 193                https://fbref.com/en/squads/4eaa11d7/2020-2021/Wolfsburg-Stats
## 194                https://fbref.com/en/squads/4eaa11d7/2020-2021/Wolfsburg-Stats
## 195  https://fbref.com/en/squads/8cec06e1/2020-2021/Wolverhampton-Wanderers-Stats
## 196  https://fbref.com/en/squads/8cec06e1/2020-2021/Wolverhampton-Wanderers-Stats
```

Chapter 21

Levi's stuff

21.1 Added

21.1.1 Hockey - ggplot

For this example, we'll use a set of NHL data from moneypuck.com. First, let's load the data into R and open the data frame.

```
nhl_2022_data <- read.csv("https://moneypuck.com/moneypuck/playerData/seasonSummary/2021/regular/

#view(nhl_2022_data)
```

This dataset includes a *lot* of covariates. It also splits these data by different game situations: even-strength (5 on 5), power play (5 on 4), etc. Let's subset the data to include all game situations.

Use the `nrow` command to check the number of columns in the new data frame. Check: Is it the same as the number of teams in the league for the 2021-2022 season?

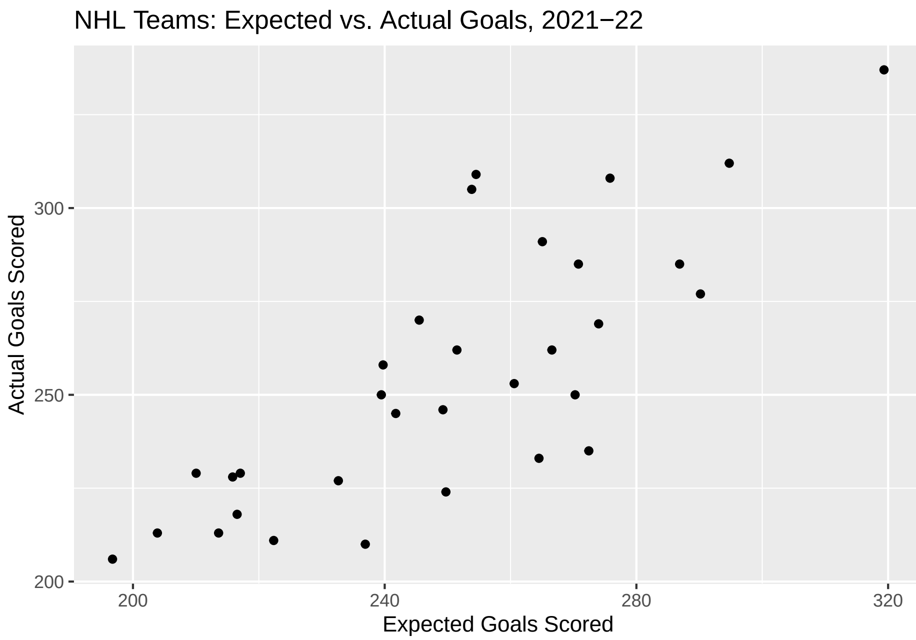
```
nhl_data_all <- filter(nhl_2022_data, situation == "all")
#view(nhl_data_all)

nrow(nhl_data_all)
```

```
## [1] 32
```

The dataset includes an Expected Goals statistic for each team in the `xGoalsFor` column. Let's plot this quantity against the team's actual number of goals scored; this is given by the `goalsFor` column.

(Remember to always have a good title and axis labels!)

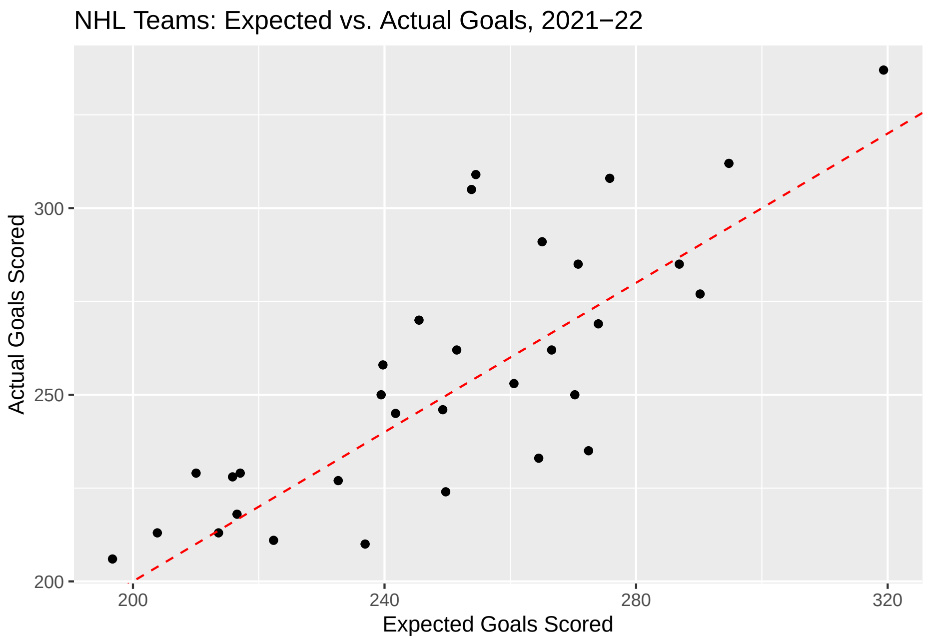
```
ggplot(data=nhl_data_all, aes(x=xGoalsFor, y=goalsFor)) + labs(x="Expected Goals Scored
```



As expected, there is a general positive correlation between expected and actual goals ($r \approx 0.8$). However, there is some variability - for example, the Kings only scored 7 more actual goals than the Ducks, despite having 56.6 more expected goals.

Let's add a line to the graph using the `geom_abline` function corresponding to the line $y = x$, the line on which data points would fall if expected goals were equal to actual goals. We can also customize the line's color and type.

```
ggplot(data=nhl_data_all, aes(x=xGoalsFor, y=goalsFor)) + labs(x="Expected Goals Scored
```



Note: A slope of 0 and an intercept of 1 are actually the default parameters for the function.

Q: What does it mean for a team's data point to fall below this line? Above it?

A: If the data point is below the line, it means the expected goals were greater than the actual goals; if the data point is above the line, it means the actual goals were greater than the expected goals.

Q: Do you think that a team's expected goals would be more likely to be closer to its actual goals for a ten-game stretch, an entire season, or five consecutive seasons? Why?

A: We would expect that as sample size increases, the result would become closer to expectation. So, actual goals would be most likely closer to expected goals over a span of five seasons.

21.2 To Be Added

21.2.1 Law of Total Probability - Hockey

Over the course of a season, a hockey player scored a goal 30% of the time during a home game, and $P(player\ scores \mid away\ game) = .18$. Assume all games are either home or away.

Q: What is the probability the player scored a goal in any game if there were an equal number of home and away games?

A: $P(score) = P(score|home)P(home) + P(score|away)P(away) = .3(.5) + .18(.5) = .24$

Q: What is the probability the player scored a goal in any game if there were twice as many home games as away games?

A: $P(score) = P(score|home)P(home) + P(score|away)P(away) = .3(\frac{2}{3}) + .18(\frac{1}{3}) = .26$

Q: What is the probability the player scored a goal in any game if the ratio of home games to away games is 2:3?

A: $P(score) = P(score|home)P(home) + P(score|away)P(away) = .3(\frac{2}{5}) + .18(\frac{3}{5}) = .228$

21.2.2 Sets and Conditional Probability

100 sports fans in Colorado were polled and it was found that 64 had attended either a Denver Nuggets or Colorado Avalanche game at Ball Arena (formerly Pepsi Center). 34 people had seen only a Nuggets game, while 17 had seen both a Nuggets and an Avalanche game.

Q: How many people saw an Avalanche game but not a Nuggets game?

A: 64 - 34 - 17 = 13

Q: What is the probability that a randomly selected person in the poll had been to a Nuggets game?

A: (34 + 17) / 100 = .51

Q: What is the probability that a randomly selected person that had been to a game at Ball Arena had been to a Nuggets game?

A: (34 + 17) / 64 = .797

Q: What is the probability that a randomly selected person had been to a Nuggets game given they had been to an Avalanche game?

A: 17 / (17 + 13) = .567

21.2.3 Binomial Probability

Two baseball teams are playing a 4-game series. The home team has a .65 probability of winning each game, and the away team a .35 probability. Assume each game is independent.

I used baseball in this example because it's the sport that most often has 4-game series, but it could easily be replaced by another sport.

Find the following probabilities.

(a) The road team wins exactly 1 game.

$\binom{4}{1}.65^3\ .35^1 = \binom{4}{3}.65^3\ .35^1 \approx .384$

```
dbinom(1, 4, .35)
```

```
## [1] 0.384475
```

```
dbinom(3, 4, .65)
```

```
## [1] 0.384475
```

(b) The home team wins exactly 2 games.

$\binom{4}{2}.65^2\ .35^2 \approx .311$

```
dbinom(2, 4, .65)
```

```
## [1] 0.3105375
```

```
dbinom(2, 4, .35)
```

```
## [1] 0.3105375
```

(c) The road team wins at least 2 games.

$\binom{4}{2}.65^2\ .35^2 + \binom{4}{3}.65^1\ .35^3 + .35^4 = 1 - [.65^4 + \binom{4}{1}.65^3\ .35^1] \approx .437$

```
pbinom(1.9, 4, .35, lower.tail=F)
```

```
## [1] 0.4370187
```

```
pbinom(2, 4, .65, lower.tail=T)
```

```
## [1] 0.4370187
```

(d) The series ends in a sweep.

$.65^4 + .35^4 \approx .194$

```
dbinom(4, 4, .65) + dbinom(4, 4, .35)
```

```
## [1] 0.1935125
```

```
.65^4 + .35^4
```

```
## [1] 0.1935125
```

21.2.4 Binomial Coefficient Symmetry

Playoff series for a certain sports league are played as a best-of-seven series, with one team hosting four games and the opposing team hosing three. An executive for the league wishes to know the number of ways the home and away games can be assigned. (One such combination is A-A-B-B-A-B-A, the format used by the NBA and NHL for their best-of-seven series.) What is the total number of combinations?

Answer: Since there are a fixed number of games (seven) and a fixed number of games that must be given to the lower-seeded team (four), there are $\binom{7}{4} = \frac{7!}{4!\,\cdot\,(7-4)!} = 35$ ways to create a home-away pattern for the seven-game series.

However, instead of thinking about the number of ways to assign the games to the team that gets four home games, what if we thought about the number of ways to assign games to the team that gets three home games?

That would be $\binom{7}{3}$. We can use the `choose` command in R to find this quantity.

```
choose(7,3)
```

```
## [1] 35
```

It turns out that this binomial coefficient is also equal to 35.

Theorem: $\binom{n}{k} = \binom{n}{n-k}$

$\binom{n}{k} = \frac{n!}{k!\,\cdot\,(n-k)!}$

$\binom{n}{n-k} = \frac{n!}{(n-k)!\,\cdot\,(n-(n-k))!} = \frac{n!}{(n-k)!\,\cdot\,k!} = \binom{n}{k}$

21.2.5 Binomials and Multinomials

Suppose we are curious about probabilities regarding the results of a soccer team's next five games.

Wait!!! A soccer game has three possible outcomes (win, lose, draw)! We can't use the binomial distribution, since it limits us to two possible outcomes!

It depends. If we are interested in the probability that a soccer team wins 2 of their next 5 games, we can use the binomial distribution. We can create the following partition of the sample space of outcomes: (Win) and (Win^C), where the second set includes both losing and drawing.

Then, the formula would be represented as:

$\binom{5}{2}\ P(Win)^2\ P(Win^C)^{(5-2)}$

If we are interested in the probability of the team winning two of the next five games, drawing two, and losing one, we cannot use the binomial theorem. That involves three outcomes, and would be represented as a multinomial.

21.2.6 Geometric (First Success) RVs

Caution: Some references parameterize the Geometric distribution based on the number of failures before the first success, rather than the trial on which the first success occurs. This changes the PMF, mean, and variance, so be careful.

```
set.seed(2022)
geometric <- rgeom(100, 1/3)
head(geometric, 20)
```

```
##  [1]  2  5  1  3 12  7  1  4  2  2  1  1  1  2  0  0  0  4  3  0
```

Some of the values were 0, which could not happen if R was considering the number of the trial on which the first success occurred. You can add 1 to the values given by `R` to arrive at the First Success distribution.

```
first_success <- geometric + 1
head(first_success, 20)
```

```
##  [1]  3  6  2  4 13  8  2  5  3  3  2  2  2  3  1  1  1  5  4  1
```

```
mean(first_success)
```

```
## [1] 3.03
```

The mean of this sample of variables is 3.03, which is close to the expected mean of $\frac{1}{p} = 3$.

21.2.7 Geometric Distribution - Hockey

Suppose the number of shots needed by a hockey team in order to score their first goal, X, is modeled by a Geometric($\frac{1}{10}$) random variable.

Q: What is the probability that it takes more than 3 shots to score the first goal?

A: $P(X > 3) = P(X = 4) + P(X = 5) + P(X = 6) + ...$

This is an infinite series, so let's use the Law of Total Probability.

$P(X > 3) = 1 - P(X \leq 3) = 1 - [P(X = 1) + P(X = 2) + P(X = 3)] = 1 - [(\frac{1}{10}) + (\frac{9}{10})^1(\frac{1}{10}) + (\frac{9}{10})^2(\frac{1}{10})] = .729$

Chapter 22

Isaac's stuff

22.1 To Be Added

22.1.1 Topic 1

Chapter 23

Aaron's stuff

23.1 Notes for Chapter 2 (Probability)

Axioms of Probability:

1. $P(A) \geq 0$
2. $P(\Omega) = 1$
3. If $A_1, A_2, \ldots, A_n$ are disjoint events, then $P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$

Theorem 23.1 (Bayes theorem)**.** *Let A and B be events in Ω such that $P(B) > 0$. Then we have the following:*

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

23.2 Suggested Readings

23.2.1 Moneyball

Moneyball, Chapter 2, How to Find a Ballplayer (Lewis, 2004)

Near the end of the chapter (page 40), Michael Lewis give a list of players the Oakland Athletics hoped to draft. How did these players turn out? Find the WAR for each of the players in their pre-free agency years and compare it against the Rockies draft picks in the same rounds from the same draft.

23.2.2 Future Value

Future Value, Chapter 7, How to Scout (Longenhagen and McDaniel, 2020)

If a player receives a running grade of 40, approximately what proportion of MLB players have a lower have a lower running grade?

For a given tool, about 95% of all player grades fall between what two bounds? (Consider the middle 95% of the distribution of grades.)

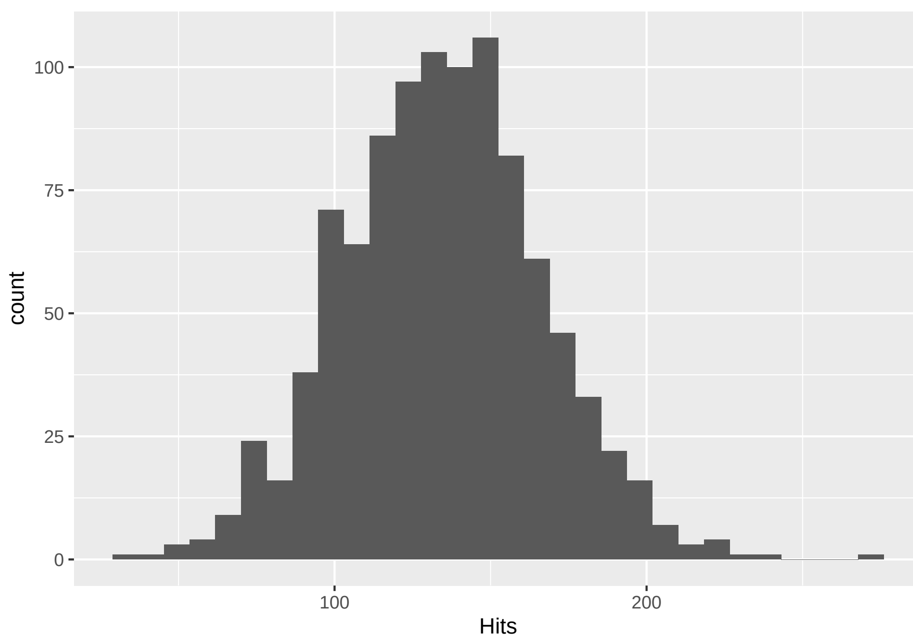
23.3 Notes for Chapter 4 (Simulation)

23.3.1 Baseball Simulation Example

```
library(tidyverse)
```

This is a baseball example for chapter 4.

```
set.seed(2022)
n.sims <- 1000
hits <- rep(0,n.sims)
avg <- 0.300
atbats.mean <- 450
atbats.sd <- 100
sim.atbats <- round(rnorm(n.sims,atbats.mean,atbats.sd))

for(i in 1:n.sims){
  sim.hits <- rbinom(1,sim.atbats[i],avg)
  hits[i] = sim.hits
}
hits.df <- data.frame(Hits=hits)
hits.df %>% ggplot(aes(x=Hits)) + geom_histogram()
```



Reference: Blocks

23.4 Equations

Here is an equation.

$$f\left(k\right) = \binom{n}{k} p^k \left(1 - p\right)^{n-k} \tag{23.1}$$

You may refer to using `\@ref(eq:binom)`, like see Equation (23.1).

23.5 Theorems and proofs

Labeled theorems can be referenced in text using `\@ref(thm:tri)`, for example, check out this smart theorem 23.2.

Theorem 23.2. *For a right triangle, if c denotes the* length *of the hypotenuse and a and b denote the lengths of the* ***other*** *two sides, we have*

$$a^2 + b^2 = c^2$$

Read more here https://bookdown.org/yihui/bookdown/markdown-extensions-by-bookdown.html.

23.6 Callout blocks

The R Markdown Cookbook provides more help on how to use custom blocks to design your own callouts: https://bookdown.org/yihui/rmarkdown-cookbook/custom-blocks.html

Reference: Footnotes and citations

23.7 Footnotes

Footnotes are put inside the square brackets after a caret `^[]`. Like this one [1].

23.8 Citations

Reference items in your bibliography file(s) using `@key`.

For example, we are using the **bookdown** package (Xie, 2016) (check out the last code chunk in index.Rmd to see how this citation key was added) in this sample book, which was built on top of R Markdown and **knitr** (Xie, 2015) (this citation was added manually in an external file book.bib). Note that the `.bib` files need to be listed in the index.Rmd with the YAML `bibliography` key.

The RStudio Visual Markdown Editor can also make it easier to insert citations: https://rstudio.github.io/visual-markdown-editing/#/citations

[1] This is a footnote.

Chapter 24

References

Bibliography

Lewis, M. (2004). *Moneyball: The art of winning an unfair game*. WW Norton & Company.

Longenhagen, E. and McDaniel, K. (2020). *Future Value: The battle for baseball's soul and how teams will find the next superstar*. Triumph Books.

Xie, Y. (2015). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.

Xie, Y. (2016). *bookdown: Authoring Books and Technical Documents with R Markdown*. R package version 0.3.9.