Injury Analysis

Carmen Canedo

11 October 2020

Loading libraries

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------- tidyverse 1.3.0.9000 --

## v ggplot2 3.3.2     v purrr   0.3.4
## v tibble  3.0.3     v dplyr   1.0.2
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0

## Warning: package 'ggplot2' was built under R version 4.0.2

## Warning: package 'tibble' was built under R version 4.0.2

## Warning: package 'tidyr' was built under R version 4.0.2

## Warning: package 'dplyr' was built under R version 4.0.2

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

```
library(likert)
```

```
## Warning: package 'likert' was built under R version 4.0.2

## Loading required package: xtable

## Warning: package 'xtable' was built under R version 4.0.2

##
## Attaching package: 'likert'

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

Reading in data

The data used are as follows:

- Youth emergency visits among US children for sports-related brain injuries from 2010-2016
- US annual average emergency department visits for all sports injuries 2010-2016
- US emergency department visits ages 5 to 24 from 2010 to 2016 by gender and activity
- Percent of concussions in US youth that play sports as of 2017
- Survey on the long-term effects of concussions in sports
- Average annual number of US college female sports injuries during the fall season from 2009 to 2014
- Average annual number of US college male sports injuries during the fall season from 2009 to 2014
- Average annual number of US college for all seasons from 2009 to 2014 by gender

```
youth_ed_tbi_visits <- read_csv("data/injury-data/ed-visits-among-us-children-for-sports-related-brain-
```

```
## Parsed with column specification:
## cols(
##   Sport = col_character(),
##   Count = col_number(),
##   `Contact Sport` = col_logical()
## )
```

```
general_ed_visits <- read_csv("data/injury-data/us-annual-average-emergency-department-visits-for-sports
```

```
## Parsed with column specification:
## cols(
##   age = col_character(),
##   count = col_number()
## )
```

```
sport_ed_visits <- read_csv("data/injury-data/us-emergency-department-visits-ages-5-24-2010-2016-by-gend
```

```
## Parsed with column specification:
## cols(
##   Activity = col_character(),
##   Male = col_double(),
##   Female = col_double()
## )
```

```
num_concussions_2017 <- read_csv("data/injury-data/number-of-concussions-suffered-by-us-youth-that-play
```

```
## Parsed with column specification:
## cols(
##   concussions = col_character(),
##   female = col_double(),
##   male = col_double()
## )
```

```
survey_tbi_effects <- read_csv("data/injury-data/survey-on-concussions-suffered-in-sports-and-long-term
```

```
## Parsed with column specification:
## cols(
##   Statement = col_character(),
##   `Strongly agree` = col_double(),
##   `Somewhat agree` = col_double(),
##   `Somewhat disagree` = col_double(),
##   `Strongly disagree` = col_double(),
##   `Not at all sure` = col_double()
## )
```

```
female_college_sports <- read_csv("data/injury-data/annual-number-of-us-college-female-sports-injuries-
```

```
## Parsed with column specification:
## cols(
##   Sport = col_character(),
##   Competition = col_number(),
##   Practice = col_number(),
##   Overall = col_number()
## )
```

```
male_college_sports <- read_csv("data/injury-data/average-number-of-male-sports-injuries-during-us-colle
```

```
## Parsed with column specification:
## cols(
##   Sport = col_character(),
##   Competition = col_number(),
##   Practice = col_number(),
##   Overall = col_number()
## )
```

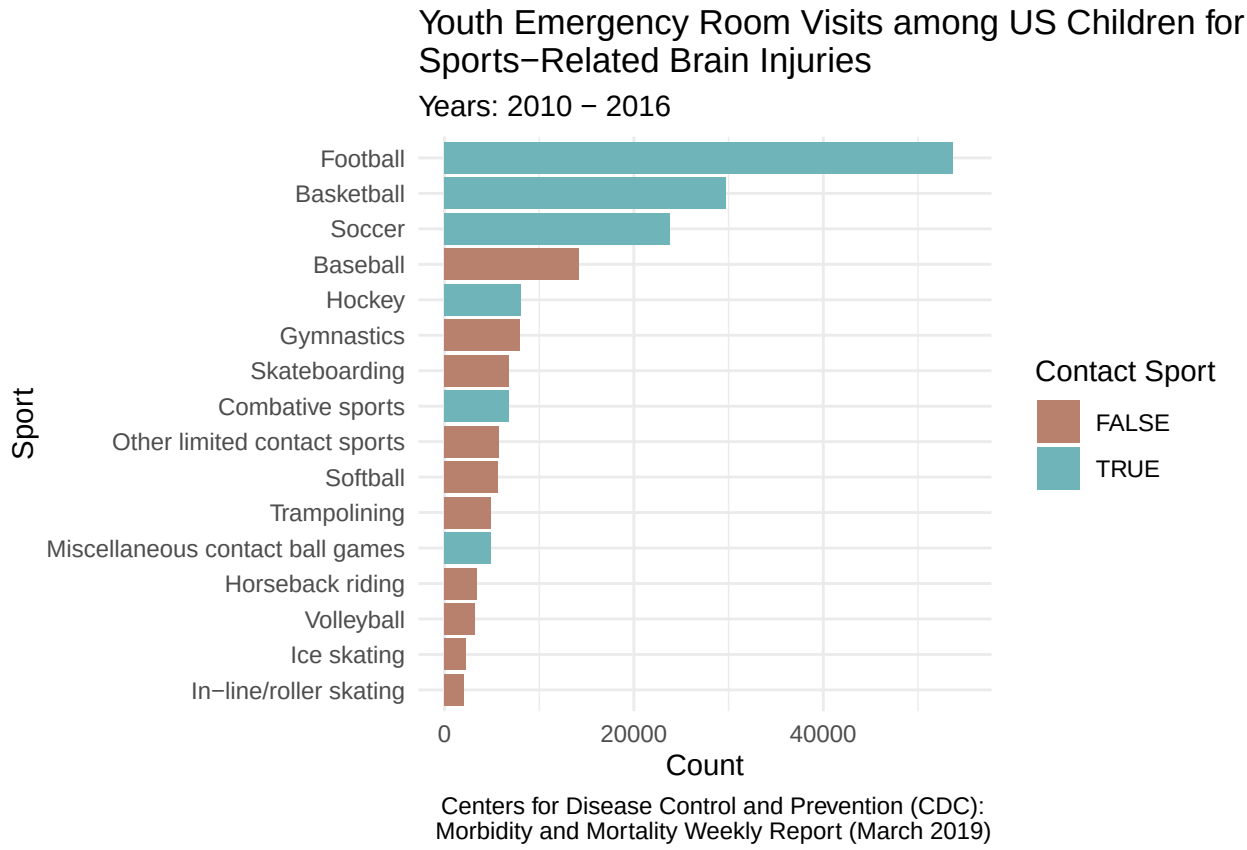
```
avg_college_sports <- read_csv("data/injury-data/annual-number-of-college-sports-injuries-in-the-us-by-g
```

```
## Parsed with column specification:
## cols(
##   Category = col_character(),
##   Competition = col_number(),
##   Practice = col_number(),
##   Overall = col_number()
## )
```

Bar Graph: Youth emergency visits for brain injuries

```
# Getting rid of total counts
youth_ed_tbi_visits <- youth_ed_tbi_visits %>%
  filter(Sport != "Total contact count") %>%
  filter(Sport != "Total limited contact count")
```

```r
# Plotting bar graph
youth_ed_tbi_visits %>%
  ggplot(aes(x = reorder(Sport, Count))) +
  geom_col(aes(y = Count,
               fill = `Contact Sport`)) +
  ylim(0, 55000) +
  scale_fill_manual(values = c("TRUE" = "#6eb4b8", "FALSE" = "#b8816e")) +
  coord_flip() +
  theme_minimal() +
  labs(title = "Youth Emergency Room Visits among US Children for \nSports-Related Brain Injuries",
       subtitle = "Years: 2010 - 2016",
       caption = "Centers for Disease Control and Prevention (CDC): \n Morbidity and Mortality Weekly Re
       x = "Sport",
       y = "Count")
```



Histogram

Decided not to do

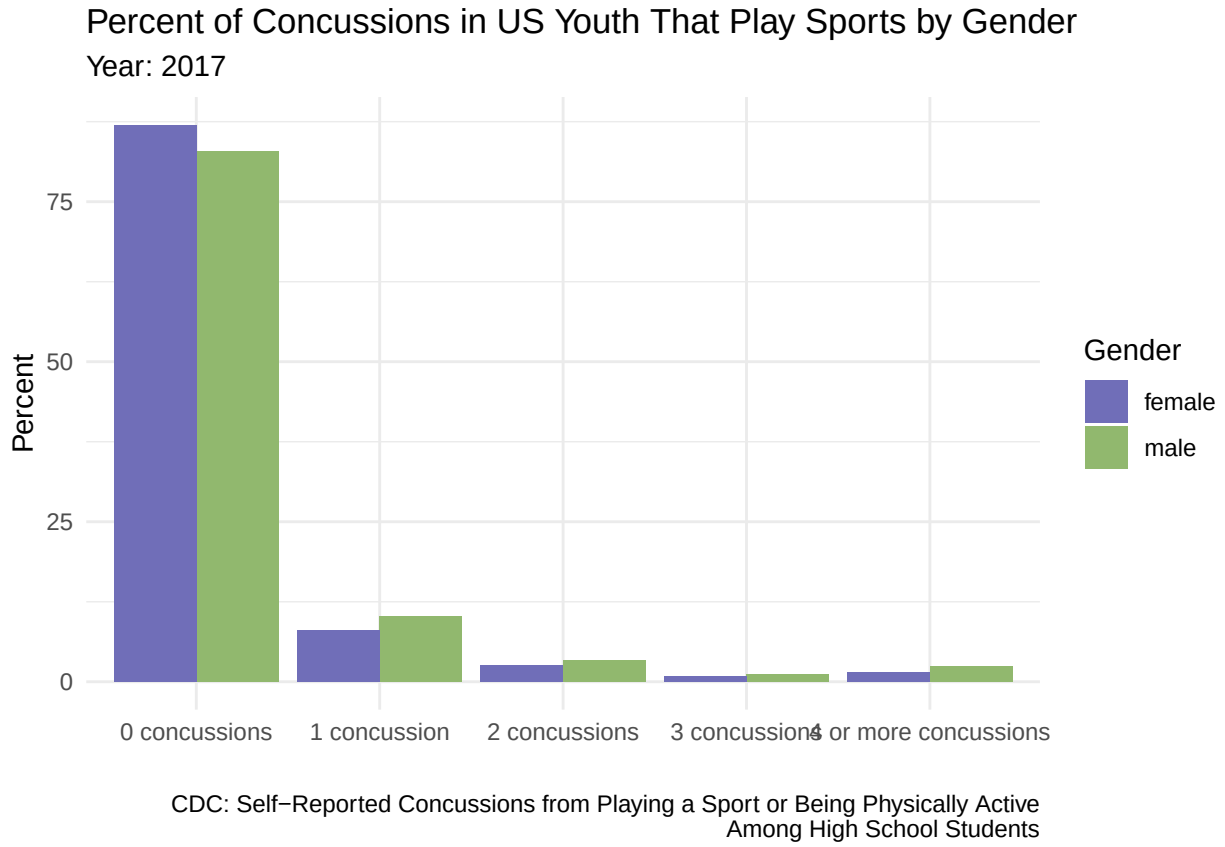
Count

Decided not to do

Bar Graph: Number of concussions by gender

```
num_concussions_2017 <- num_concussions_2017 %>%
  pivot_longer(cols = c(male, female)) %>%
  rename(gender = name)


num_concussions_2017 %>%
   ggplot(aes(x = concussions, y = value, fill = gender)) +
   geom_col(position = "dodge") +
  scale_fill_manual(values = c("female" = "#706eb8",
                               "male" = "#91b86e"),
                    name = "Gender") +
  labs(title = "Percent of Concussions in US Youth That Play Sports by Gender",
       subtitle = "Year: 2017",
       caption = "CDC: Self-Reported Concussions from Playing a Sport or Being Physically Active
Among High School Students",
       x = "",
       y = "Percent") +
  theme_minimal()
```



Stacked Bar Graph

```
text_statements <- c("Players who suffer a head injury should be required to take a minimum set amount
                     "Helmets should be changed to better protect against concussions.",
                     "The risks of playing football are widely known and players participate at their ou
                     "There should be a standardized test used to determine if and when injured players
                     "Aggressive tackles which are more prone to leading to head injuries should be res

survey_levels <- c(
  "Strongly agree",
  "Somewhat agree",
  "Somewhat disagree",
  "Strongly disagree",
  "Not at all sure"
)
```

```
# Getting rid of Statement column to convert numbers to percents
survey_interim <- survey_tbi_effects %>%
  select(-Statement)

# Calculating percentages
data_percentage <- survey_interim %>%
  apply(2,
        function(x){
          perc <- x / sum(x)
          round(perc, digits = 2)
          })

# Converting to tibble and adding final column so it equals 1
diff_interim <- data_percentage %>%
  as_tibble() %>%
  select_if(is.numeric) %>%
  mutate(sum = rowSums(.[1:5])) %>%
  mutate(diff = 1 - sum)

# Saving into data_percentage
data_percentage <- diff_interim %>%
  select(-sum)


# Adding the statement column back
data_percentage <- cbind(survey_tbi_effects$Statement, data_percentage) %>%
  as_tibble()  %>%
  slice(1:5)


# Renaming column
data_percentage <- data_percentage %>%
  rename(Statement = `survey_tbi_effects$Statement`)

# Pivot long
long_percentage <- data_percentage %>%
  pivot_longer(cols = contains("r")) %>%
```

```
  rename(response = name)
```

```
set_colors <- c("#91b86e", "#956eb8", "#b86e90", "#6eb895", "#b8816e", "#b6b86e")
xlabels <- c("1", "2", "3", "4", "5")

# Making graph
stacked_survey <- long_percentage %>%
  ggplot(aes(x = Statement, y = value, fill = response)) +
  geom_col(position = "fill") +
  scale_fill_manual(values = set_colors, name = "Response") +
  scale_x_discrete(labels = xlabels) +
  labs(title = "Survey Responses on The Long-Term Effects of Concussions in Sports",
       subtitle = "2014",
       caption = "Source: Harris Interactive | The Harris Poll #92",
       y = "") +
  theme_light()

stacked_survey
```