California's Unlevel Playing Field

Emma Durler, Kylie Scharf, Nicole Yee, Curtis LeQuang

Table of contents

```r
library(tidyverse)
library(writexl)
#loading in data
tuesdata <- tidytuesdayR::tt_load('2022-03-29')
sports_dat <- tuesdata$sports
```

```r
#subsetting and moving data
library(tidyverse)
sports_dat <- sports_dat |>
  filter(sports == "Baseball" | sports == "Softball") |>
  mutate(
    rev_men = ifelse(is.na(rev_men), 0, rev_men),
    rev_women = ifelse(is.na(rev_women), 0, rev_women),
    exp_men = ifelse(is.na(exp_men), 0, exp_men),
    exp_women = ifelse(is.na(exp_women), 0, exp_women),
  ) |>
  filter(state_cd == "CA")

sports_dat <- sports_dat |>
  rename(Sport = sports)
```

```r
plot <- sports_dat |>
  group_by(Sport, year) |>
  filter(total_exp_menwomen != 0) |>
  summarise(mean_exp = mean(total_exp_menwomen), stan = sd(total_exp_menwomen)/sqrt(n())) |>
  ggplot(aes(y = mean_exp, x = year, color = Sport, shape=Sport, fill=Sport)) +
   geom_ribbon(aes(ymin = mean_exp - stan, ymax = mean_exp + stan), alpha=.5, color=NA) +
  geom_point(size=4) +
```

```
  geom_line() +
  labs(x = "Year", subtitle = "Average Yearly Expense", y = "", title = "Fig 1. California Co
theme(plot.caption = element_text(hjust = 0)) +
scale_y_continuous(labels = scales::dollar)

plot
```
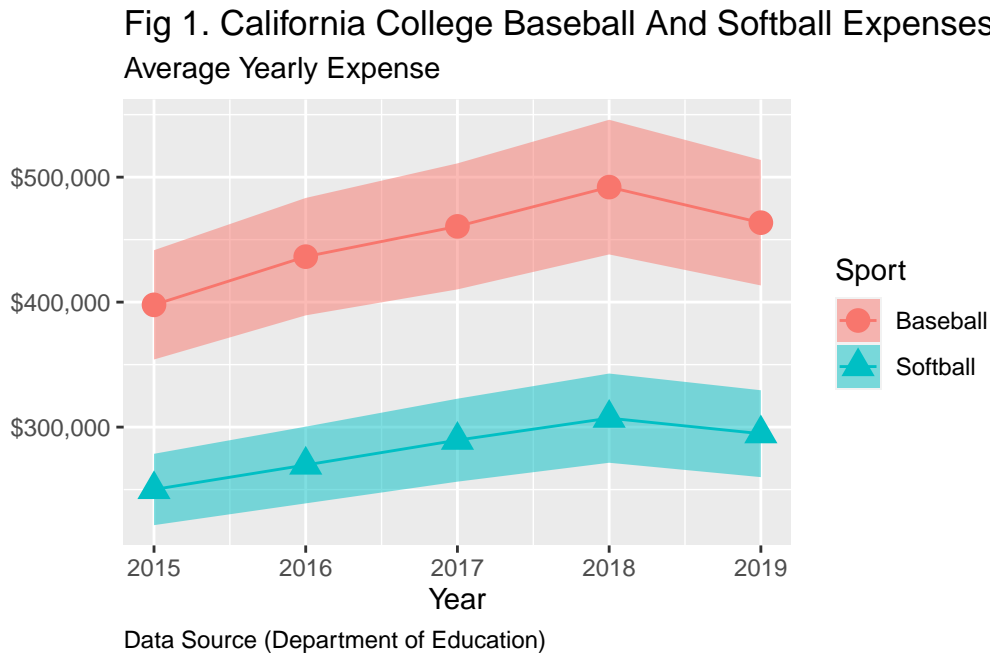

Figure 1: Above entails an analysis of the differences in average expenses from 2015 to 2019
for collegiate baseball and softball in California. The Department of Education
collects data yearly on collegiate sports budgets in accordance to The Equity in
Althletics Disclosure Act. The graph above consists of such data subsetted to contain
soley softball and baseball teams from California in the years 2015 to 2019. The
expense for a particular year averaged over all California softball teams is plotted
as a blue triangle, wheras the expense averaged over all California baseball teams
in a particular year is plotted as a red circle. The highlighted region represents
adding or subtracting one standard error of the average from the plotted averages.
There appears to be more variability in the average expenses for baseball than for
softball in any particular year in California. The average expenses for baseball are
consistantly signifigantly larger than the expenses for women, although the expenses
for both sports tend to increase at a similiar rate over time.

```
sports_dat |>
  group_by(Sport, year) |>
  filter(exp_men != 0 | exp_women !=0) |>
  summarise(mean_exp_men = mean(exp_men), mean_exp_women = mean(exp_women))
```

```
# A tibble: 10 x 4
# Groups:   Sport [2]
   Sport      year mean_exp_men mean_exp_women
   <chr>     <dbl>        <dbl>          <dbl>
 1 Baseball  2015      397818.              0
 2 Baseball  2016      436380.              0
 3 Baseball  2017      460609.              0
 4 Baseball  2018      492033.              0
 5 Baseball  2019      463565.              0
 6 Softball  2015           0         250097.
 7 Softball  2016           0         269675.
 8 Softball  2017           0         289536.
 9 Softball  2018           0         307159
10 Softball  2019           0         294723.
```