California's Unlevel Playing Field

Emma Durler, Kylie Scharf, Nicole Yee, Curtis LeQuang

Table of contents

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

```
#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") |>
  rename(Sport = sports) |>
  group_by(Sport, year) |>
  pivot_wider(id_cols = year:sector_name, names_from = Sport, values_from = exp_men:exp_women) |>
  mutate(diff = exp_men_Baseball - exp_women_Softball)
```

```
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 C
theme(plot.caption = element_text(hjust = 0)) +
scale_y_continuous(labels = scales::dollar)
```

```
Error in `group_by()`:
! Must group by variables found in `.data`.
x Column `Sport` is not found.
```

```
plot
```

```
function (x, y, ...)
UseMethod("plot")
<bytecode: 0x1458c1468>
<environment: namespace:base>
```

```
library(rstatix)
library(kableExtra)
table <- sports_dat |>
  group_by(Sport, year) |>
  filter(total_exp_menwomen != 0) |>
  summarise(mean_exp = mean(total_exp_menwomen)) |>
  pivot_wider(id_cols = year, names_from = Sport, values_from = mean_exp) |>
  rename(Year=year, Expenses_Men=Baseball, Expenses_Women=Softball)
```

```
Error in `group_by()`:
! Must group by variables found in `.data`.
x Column `Sport` is not found.
```

```
kable(table,
      caption="CA Collegiate Baseball and Softball Expenses",
      col.names = c("Year",
                    "Expenses Baseball ($)",
                    "Expenses Softball ($)")) |>
  kable_classic() |>
  kable_styling(full_width = FALSE, position = "center") |>
  footnote(general="Source: Department of Education")
```

```
Error in as.data.frame.default(x): cannot coerce class '"function"' to a data.frame
```

```r
t.test(sports_dat$exp_men_Baseball, sports_dat$exp_women_Softball, paired=TRUE)
```

```
    Paired t-test

data:  sports_dat$exp_men_Baseball and sports_dat$exp_women_Softball
t = 14.236, df = 699, p-value < 2.2e-16
alternative hypothesis: true mean difference is not equal to 0
95 percent confidence interval:
 130301.7 171991.8
sample estimates:
mean difference
       151146.8
```