

Tidy Tuesday - Week 2

Grace Hu

15/03/2021

Task at hand

For this week, I will be looking at the Bechdel Test data taken from [here](#). Upon reading the [FiveThirtyEight](#) article that inspired this week's data, I wanted to try and recreate two of the plots included in the article: "The Bechdel Test Over Time" and "Median Budget For Films Since 1990".

Challenge 01

Recreating "The Bechdel Test Over Time".

Data Wrangling

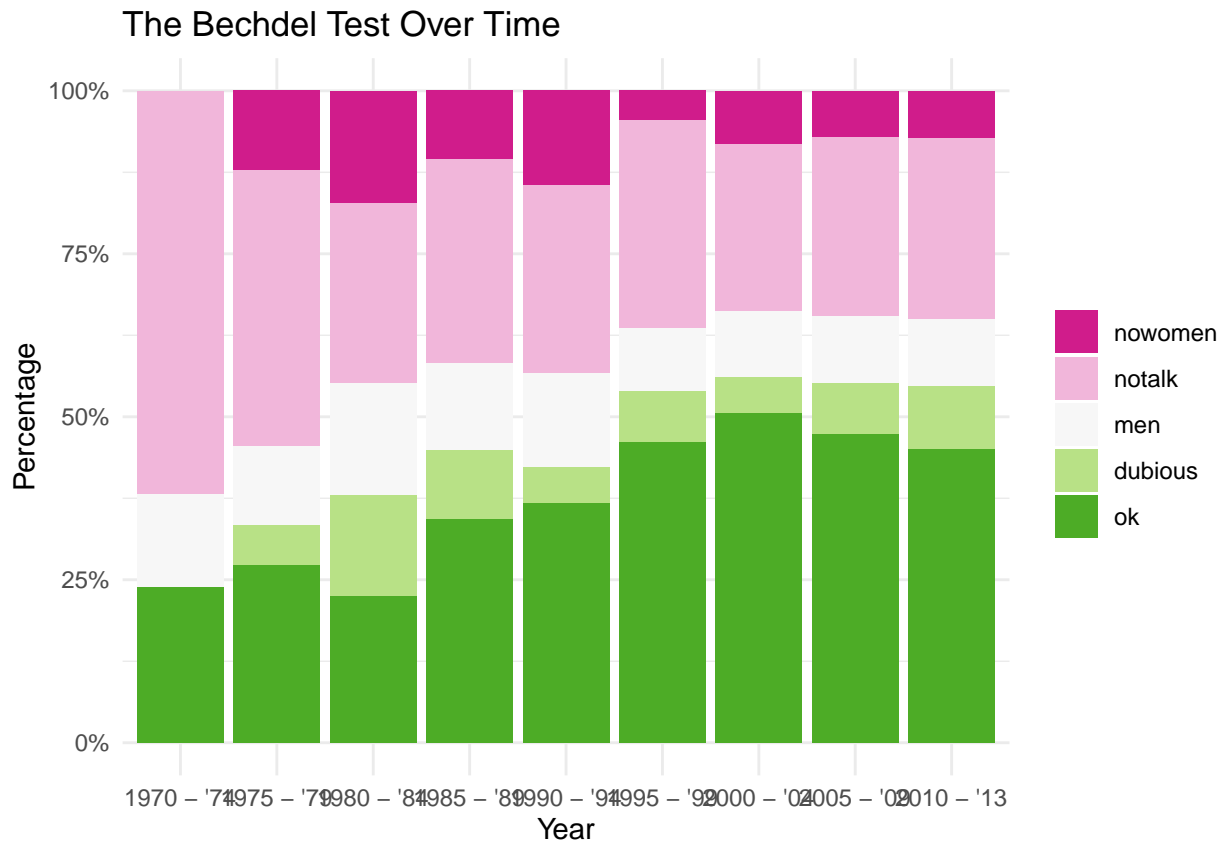
```
movies <- raw_movies %>%
  mutate(clean_test = fct_relevel(
    as_factor(clean_test), c("nowomen", "notalk", "men", "dubious", "ok")
  ))

movies_acc <- movies %>%
  mutate(year_acc = as_factor(case_when(
    year >= 1970 & year < 1975 ~ "1970 - '74",
    year >= 1975 & year < 1980 ~ "1975 - '79",
    year >= 1980 & year < 1985 ~ "1980 - '84",
    year >= 1985 & year < 1990 ~ "1985 - '89",
    year >= 1990 & year < 1995 ~ "1990 - '94",
    year >= 1995 & year < 2000 ~ "1995 - '99",
    year >= 2000 & year < 2005 ~ "2000 - '04",
    year >= 2005 & year < 2010 ~ "2005 - '09",
    year >= 2010 & year < 2015 ~ "2010 - '13",
  ))) %>%
  mutate(year_acc = fct_relevel(year_acc, levels(year_acc)[9:1]))
```

Data Visualization

```
ggplot(data = movies_acc, aes(x = year_acc, fill = clean_test)) +
  geom_bar(position = "fill") +
  scale_fill_brewer(palette="PiYG") +
  scale_y_continuous(labels=scales::percent_format()) +
```

```
labs(
  title = "The Bechdel Test Over Time",
  x = "Year",
  y = "Percentage"
) +
theme_minimal() +
theme(legend.title=element_blank())
```



Challenge 02

Recreating “Median Budget For Films Since 1990”.

Data Wrangling

```
budget <- movies %>%
  filter(year >= 1990) %>%
  mutate(clean_test = as_factor(case_when(
    as.character(clean_test) == "dubious" ~ "ok",
    TRUE ~ as.character(clean_test)
  ))) %>%
  mutate(clean_test = fct_rev(fct_relevel(clean_test, "notalk", after=2))) %>%
  group_by(clean_test) %>%
  summarise(median_budget = median(budget) / 1000000)
```

Data Visualization

```
ggplot(data = budget, aes(x = clean_test, y = median_budget)) +  
  geom_bar(stat = "identity", position="dodge", fill="#0190d4") +  
  geom_text(aes(label=median_budget), hjust=-0.3) +  
  labs(  
    title = "The Bechdel Test Over Time",  
    subtitle="2013 dollars",  
    x = element_blank(),  
    y = element_blank()  
  ) +  
  scale_x_discrete(labels=c(  
    "Fewer than two women",  
    "Women don't talk to each other",  
    "Women only talk about men",  
    "Passes Bechdel Test")) +  
  theme_minimal() +  
  theme(  
    legend.title = element_blank(),  
    text = element_text(size=18)  
  ) +  
  coord_flip()
```

