Tidy Tuesday - Week 3

Grace Hu

23/03/2021

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

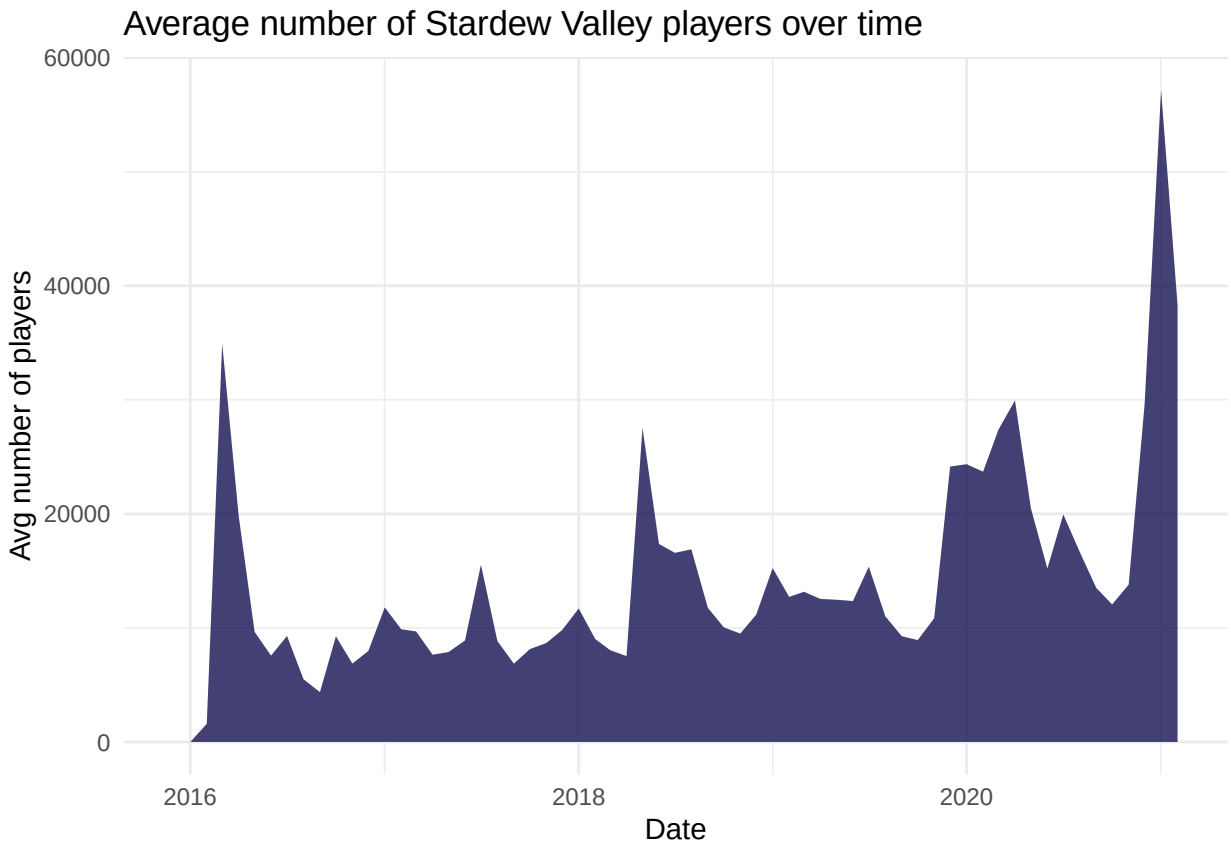
Stardew Valley Stats (one of my favourite video games!)

Data Wrangling

```
stardew_data <- games_cleaned %>%
  filter(gamename == "Stardew Valley")
```

Data Visualization

```
ggplot(data = stardew_data, aes(x = date, y = avg)) +
  geom_area(stat = "identity", fill = "#151152", alpha=0.8) +
  labs(
    title = "Average number of Stardew Valley players over time",
    x = "Date",
    y = "Avg number of players"
  ) +
  theme_minimal()
```



Challenge 02

Top 5 games in 2020 (released before 2020)

Data Wrangling

```
game_release_date <- games_cleaned %>%
  filter(is.na(gain)) %>%
  arrange(date) %>%
  distinct(gamename, .keep_all = TRUE) %>%
  mutate(release_date = date) %>%
  select(gamename, release_date)

top_game_names <- games_cleaned %>%
  left_join(game_release_date, by="gamename") %>%
  filter(year == 2020 & release_date <= as.Date("2020-01-01")) %>%
  group_by(gamename) %>%
  mutate(avg_2020 = mean(avg)) %>%
  arrange(-avg) %>%
  distinct(gamename) %>%
  head(n=5)

top_games <- games_cleaned %>%
  filter(year == 2020 & gamename %in% top_game_names$gamename)
```

Data Visualization

```
ggplot(data = top_games, aes(x = date, fill = gamename)) +
  geom_area(aes(y = avg), alpha=0.6) +
  labs(
    title = "Top 5 Games in 2020",
    subtitle="By highest average number of players over the year",
    x = "Date",
    y = "Avg number of players"
  ) +
  theme_minimal() +
  scale_fill_brewer(palette="Dark2")
```