

Tidy Tuesday

Erik King-Vargas

12/6/23

Table of contents

Introduction	3
1 Week of 11/27/2023	4
1.1 Data Cleaning	4
1.2 Ridgeline Plot	5
2 Week of 12/4/2023	7

Introduction

This Bookdown repository serves as a place to store my TidyTuesday submissions and showcase the code and deeper analysis on each dataset. As I continue to analyze the weekly datasets I'll upload the code here and link to the new page in my social media posts. Thanks for reading!

1 Week of 11/27/2023

Loading up some libraries that will make this analysis easier, namely `tidyverse` and `lubridate`. Also, we'll load the data from the `tidytuesdayR` package.

```
library(tidyverse)
library(viridis)
library(hrbrthemes)
library(ggribes)
library(lubridate)

tuesdata <- tidytuesdayR::tt_load('2023-11-28')

drwho_episodes <- tuesdata$drwho_episodes
drwho_directors <- tuesdata$drwho_directors
drwho_writers <- tuesdata$drwho_writers
```

1.1 Data Cleaning

First we join the data together in one flat table, then create a year variable for each episode based on when it was aired. We also turn the year variable into a factor so that it will be plotted correctly later on.

```
# join all data together

drwho_episodes <- left_join(drwho_episodes, drwho_directors,
                           by = "story_number")

drwho_episodes <- left_join(drwho_episodes, drwho_writers,
                           by = "story_number")

# get a year aired variable for each episode
drwho_episodes <- drwho_episodes %>%
  mutate(year_aired = floor_date(first_aired, "year"))
```

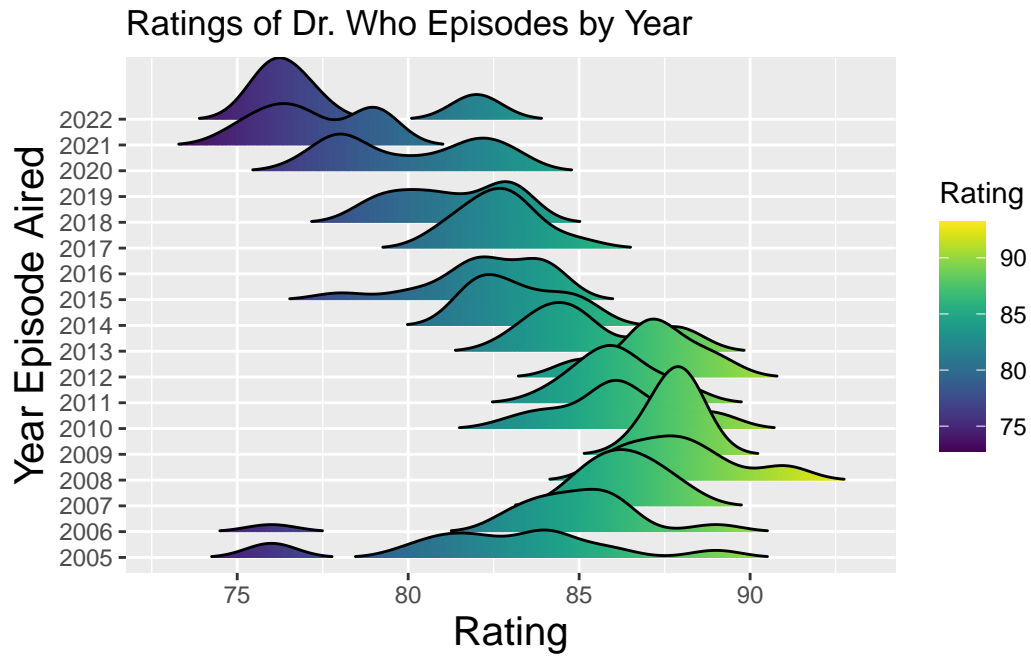
```
# strip out the month and day
drwho_episodes$year_aired <- year(drwho_episodes$year_aired)

drwho_episodes$year_aired <- as.factor(drwho_episodes$year_aired)
```

1.2 Ridgeline Plot

Finally, we visualize the Dr. Who episode ratings by year the episodes were aired. An apparent decline in ratings can be seen as the show has progressed through 2022.

```
drwho_episodes %>%
  group_by(year_aired) %>%
  ggplot(aes(x = rating, y = year_aired, group = year_aired,
            fill = ..x..)) +
  geom_density_ridges_gradient(scale = 3, rel_min_height = 0.01) +
  scale_fill_viridis(name = "Rating", option = "D") +
  labs(title = 'Ratings of Dr. Who Episodes by Year',
       x = "Rating",
       y = "Year Episode Aired") +
  theme(
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8),
    axis.title.x = element_text(size = 15),
    axis.title.y = element_text(size = 15)
  )
```



2 Week of 12/4/2023