Tidy Tuesday

Erik King-Vargas

12/6/23

Table of contents

Introduction

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.

```r
library(tidyverse)
library(viridis)
library(hrbrthemes)
library(ggridges)
library(lubridate)

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

```
Downloading file 1 of 3: `drwho_episodes.csv`
Downloading file 2 of 3: `drwho_directors.csv`
Downloading file 3 of 3: `drwho_writers.csv`
```

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

1.1 Data Cleaning

```r
# 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
```

```r
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)
```

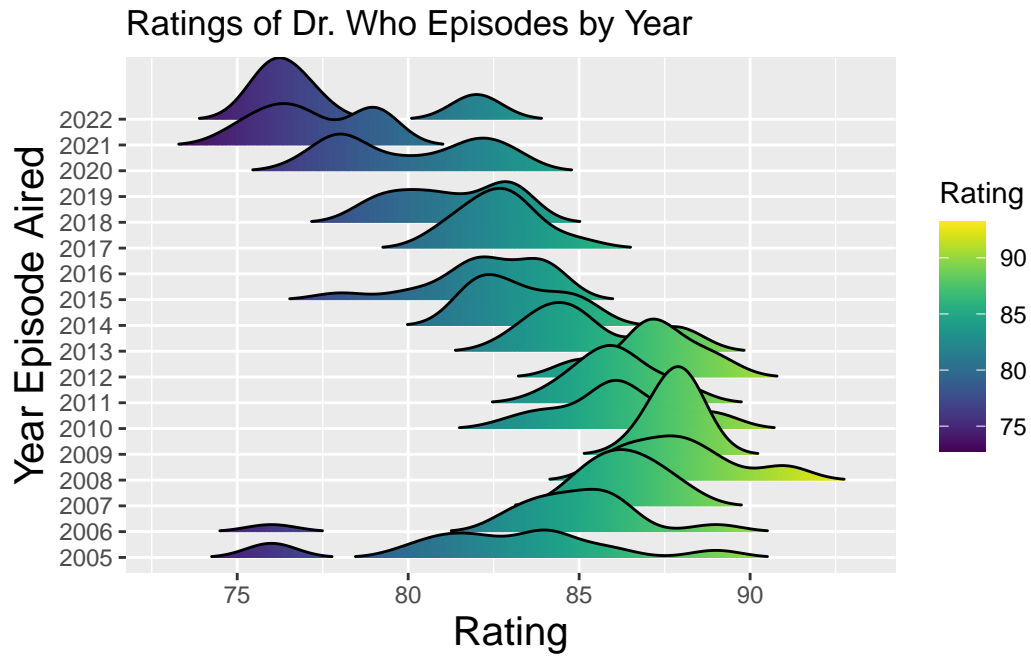## 1.2 Ridgeline Plot

```r
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_ipsum()+

  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)
  )
```

Ratings of Dr. Who Episodes by Year

2 Week of 12/4/2023