Let's go to the movies!
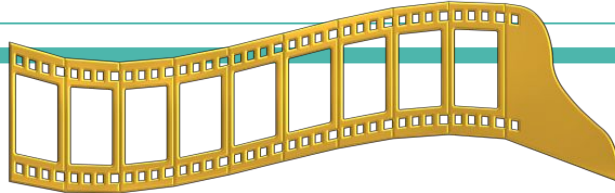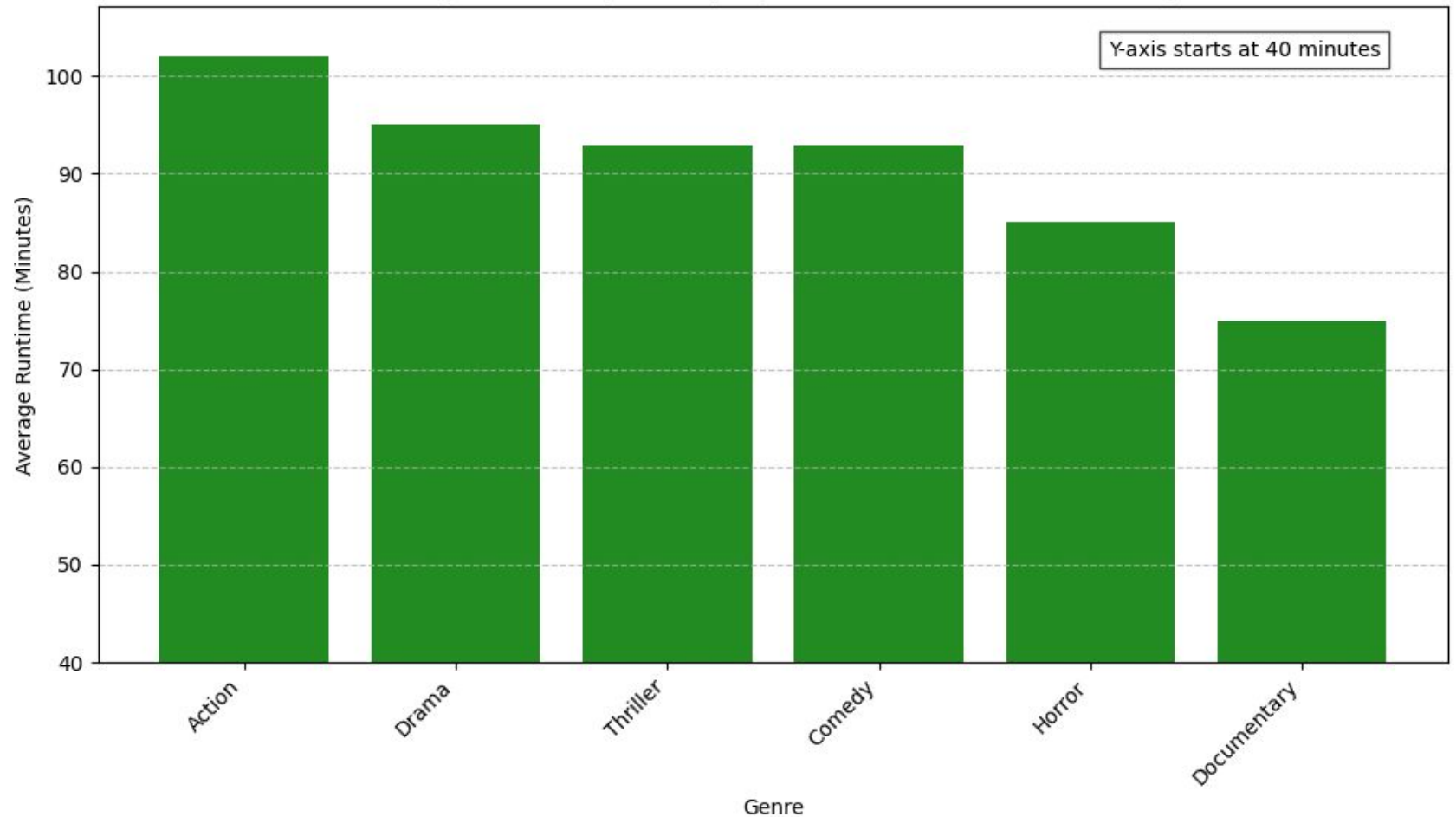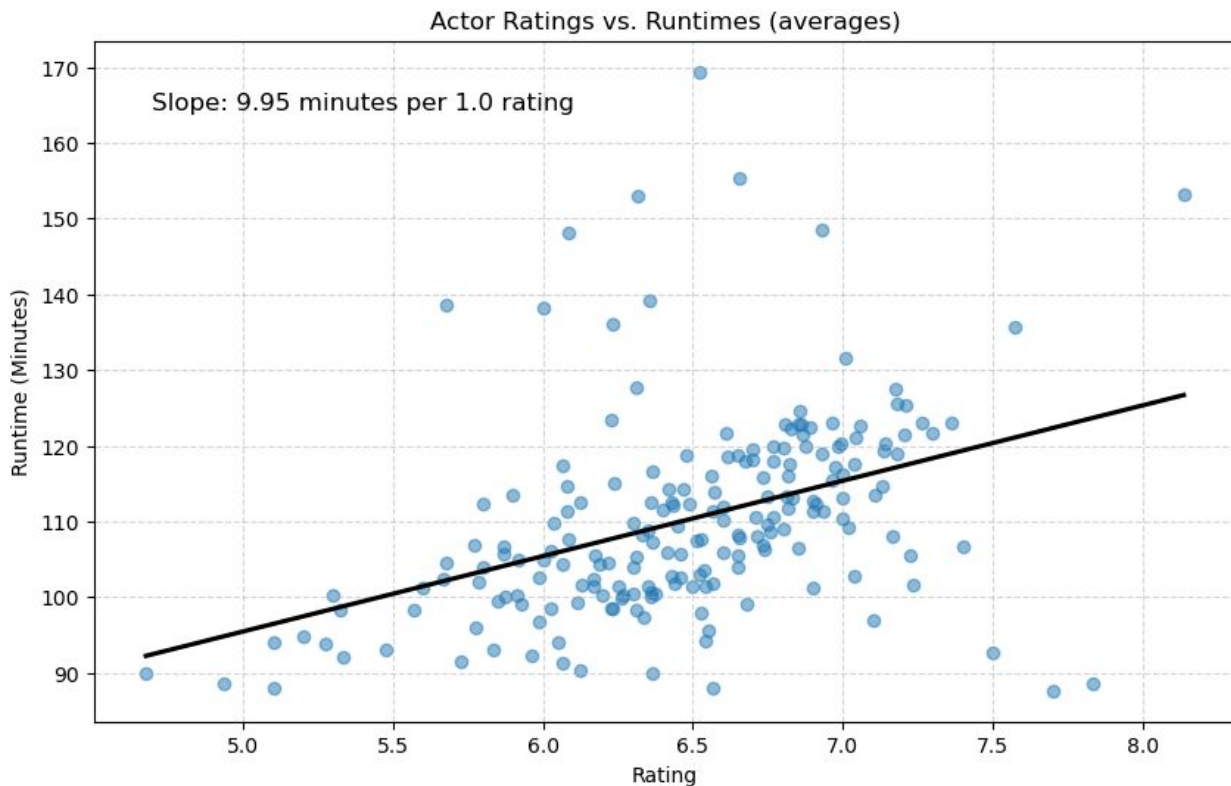
Paul Dennis

Initial Exploration: IMDB Data 2010-2021

- The data contains records from 2010 to 2021. We're focusing on 2010-2019. While this will make our conclusions more relevant for modern movies, it limits our ability to find trends over time.
- Many movies have multiple genres and multiple directors. This makes categorization based on those categories complicated.
- Movie ratings - We're lacking data on box office performance, or critical acclaim. IMDB's ratings are user-generated. We need to keep this in mind when using this rating.
- Among single-genre titles, Documentary is by far the most common, with Drama second.
- In such titles, Documentaries have the shortest runtime (as expected) with Action being the longest (about 30% longer than Docs on average)

Average Runtime by Genre (Single Genre Movies, >= 1000 Films)

Y-axis starts at 40 minutes

Audience Engagement, Narrative Films

Actor Ratings vs. Runtimes (averages)

Slope: 9.95 minutes per 1.0 rating

There is a noticeable correlation between an actor's average rating and their average runtime.

It's reasonable to conclude that more highly rated actors, due to their star power or acting talent, are a factor in keeping audiences engaged for longer films.

IMBD data from 2010-2019 inclusive. Actors with at least 3 movies, whose movies all have at least 3000 votes, no documentaries

Audience Engagement - Further Analysis

- Within current data:
  - Director/Writer - examine this trend for directors and/or writers
  - Genre Breakdown - is this trend true for all genres?

- With additional data
  - Other measures of audience engagement - critical acclaim such as RT ratings, CinemaScore data, and box office performance
  - Other decades - examine these trends outside of the 2010s

"Funniest" Actors of the Decade (Comedy Ratings)

Actor	Average Rating
Rakshit Shetty	8.3
Tomokazu Sugita	7.9
Brad Pitt	7.7
Humayun Saeed	7.7
Ryan Gosling	7.7
Taron Egerton	7.3
Benjamin Bratt	7.2
Javed Sheikh	7.2
Zachary Levi	7.2
Tom Hiddleston	7.1

- Representing actors from diverse locations emphasizes the global nature of comedy
- There is no gender diversity - this suggests a substantial bias in these ratings.
- Online ratings are particularly subject to gender biases, but this is also possibly an indication of bias in the industry.

(Actors with min 3 movies that include comedy, min 1000 votes)

Credits

- Data from IMDB, provided with exercise
- See Jupyter file for more complete documentation/explanation
- https://github.com/pauldennis2/dsc-sql-summative-lab

Thank you! Please throw away your popcorn on the way out.