Game of Quartiles: A Data-Driven Exploration of Ratings and Genres

Uncovering Patterns in Video Game Ratings Across Genres and Time

By: Eden Nuhamovici
Community Tech - Data Sciences, Toronto,
Canada
Email: enuhamovici@gmail.com

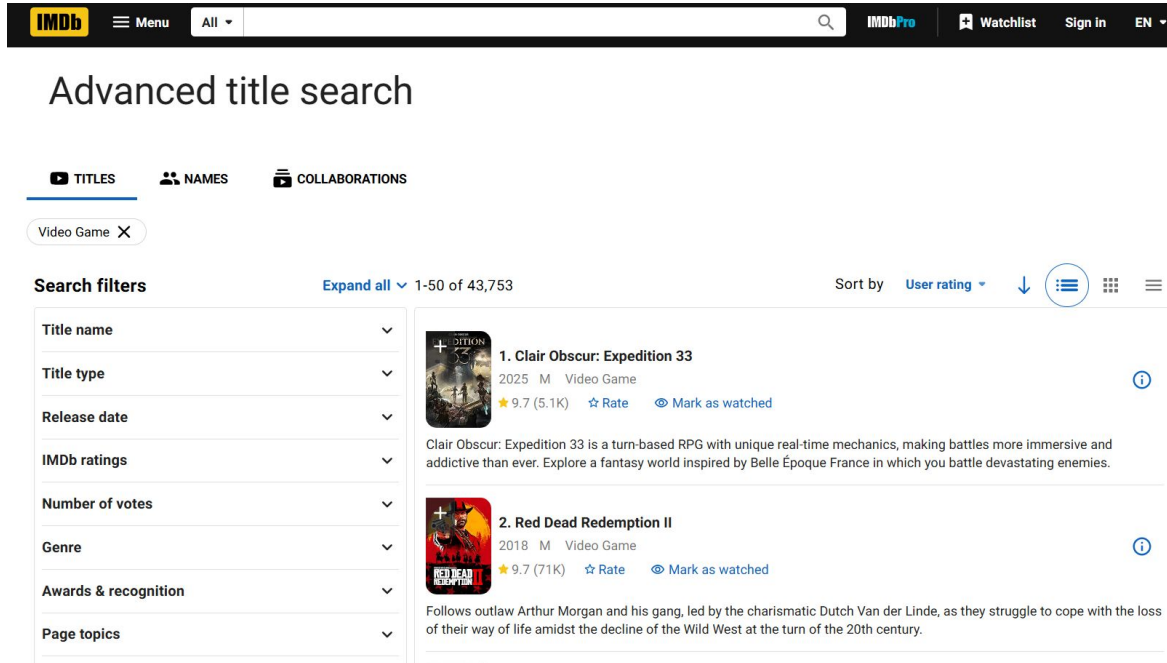Chris Avellone, lead creative designer at *Fallout: New Vegas* developer Obsidian Entertainment, has claimed that the company did not receive a bonus payment for its work on *New Vegas* because it did not garner a high enough Metacritic score. The game was released in 2010 for Windows PC, PS3 and Xbox 360, and within a month of launch had shipped 5 million units and made $300 million in sales. However, Obsidian was told that it would only receive a bonus payment from publisher Bethesda if the game received an 85 or more score from aggregate reviews website Metacritic, says Avellone. The game actually scored 84 on PC and Xbox 360, and 82 on PS3. Avellone tweeted, "*Fallout: New Vegas* was a straight payment, no royalties, only a bonus if we got an 85+ on Metacritic, which we didn't." This isn't the first time this sort of practice has been heard of, with reports in the past from numerous publishers who have based royalties for video game releases on how well a game does on Metacritic. Gamasutra has contacted Avellone to clarify, and to ask whether he believes receiving this bonus would have allowed the company to dodge the latest round of layoffs this week.

Introduction

- Ratings are important for the success of a video game as they can influence popularity and purchases.

- **Platforms like IMDb** offer ratings and data that help players decide if a game is worth buying.

Data Source

- The IMDB Video Games dataset comes from Kaggle.

- Consists of 20,803 games spanning the years 1952 to 2027 (for planned release).

- **Genres of games tracked:**
  - **Action**
  - **Adventure**
  - **Comedy**
  - **Crime**
  - **Family**
  - **Fantasy**
  - **Mystery**
  - **Sci-Fi**
  - **Thriller**

  (Not mutually exclusive in classification)

IMDB Video Games

Around 20K titles for data and game enthusiasts

Data Card   Code (7)   Discussion (1)   Suggestions (0)

About Dataset

Inspiration
This data collection is part of a research project on game based learning.

Context
The purpose of this dataset is to gain insights into the trends of game genre popularity.
I am particularly interested in the following questions:

1. Which game genres are the most popular?
2. What are the trends in popularity for each game genre?

Usability ⓘ
9.71

License
CC0: Public Domain

Expected update frequency
Never

Tags

Arts and Entertainment

Education

Analysis Questions

Is there a relationship between genre of a video game and the rating it received?

Is there a relationship between age rating of a game and the rating it received?

What is the trend with game ratings over time? Are they getting better or worse over time? What about on the genre level?

Data Cleaning & Preparation

- I removed entries with missing ratings- the analytical dataset consists of ~11,600 records.
- Made sure that my ratings were numeric and consistently formatted.
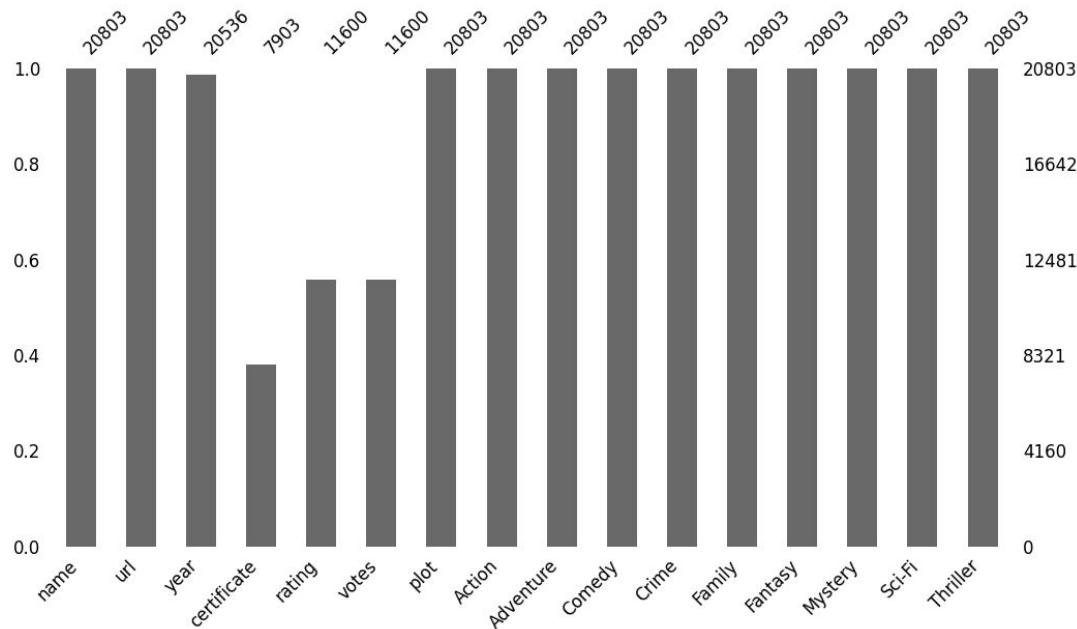
Proposed Approach

- The proposed method of choice was using quartiles to break the data apart to minimize any outlier or skew in data.

- Used bar charts to get more insightful looks in the data such as deciphering what years contains the most records of data.
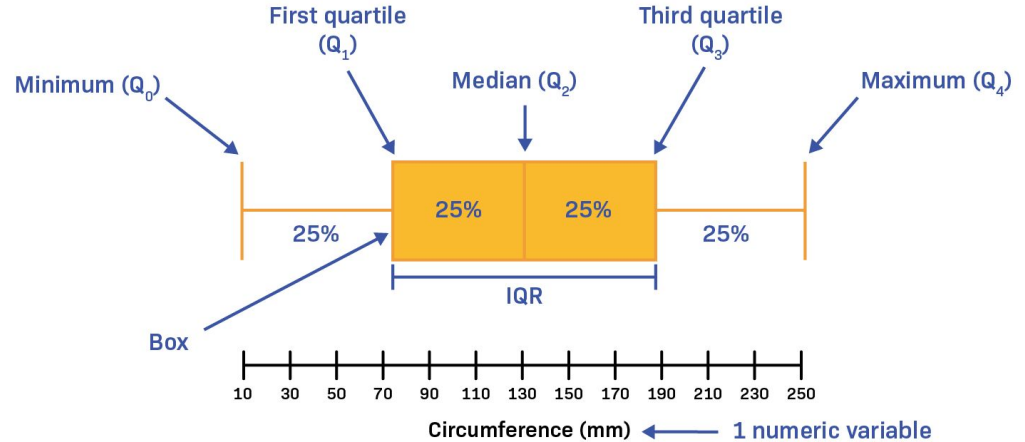
More methodology

Analytical Approach

● Used histograms to visualize
  distribution of ratings across age
  categories.

● I Grouped data by genre to help
   spot the differences in rating
   patterns.
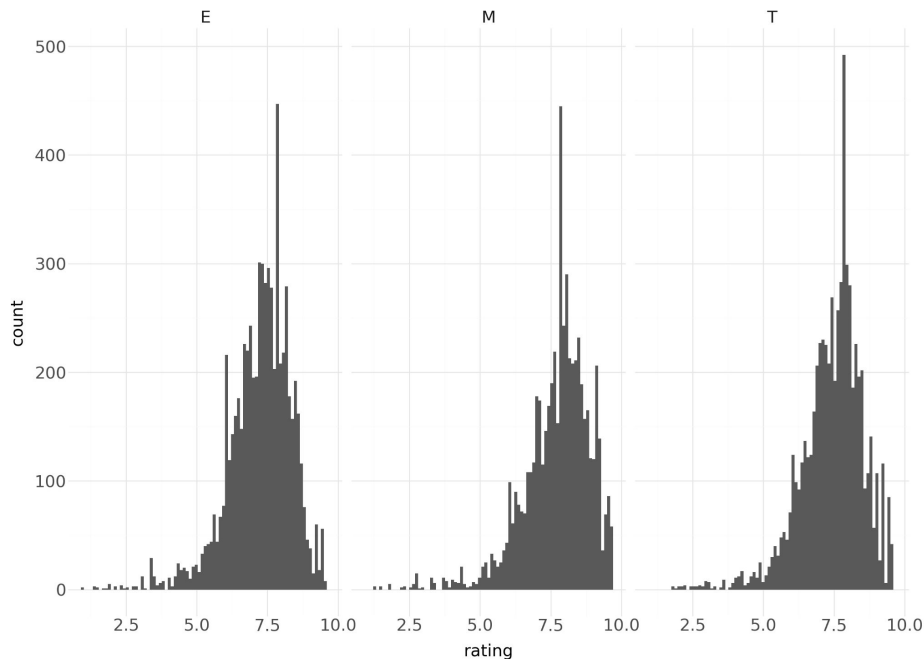
Results

- Histograms are to split the data in three ways:
  - Excluding games classified as action and/or adventure exclusively (1)

  - The rating distribution of games that are action AND adventure (including other genres) (2)

  - The rating distribution for ONLY action/adventure games excluding genres (3)

(1)

(2)

(3)

Results

Distribution of IMDb Video Game Ratings by Genre
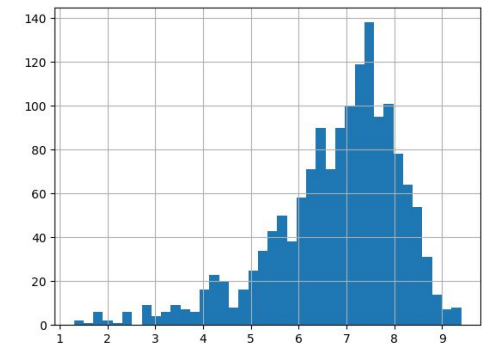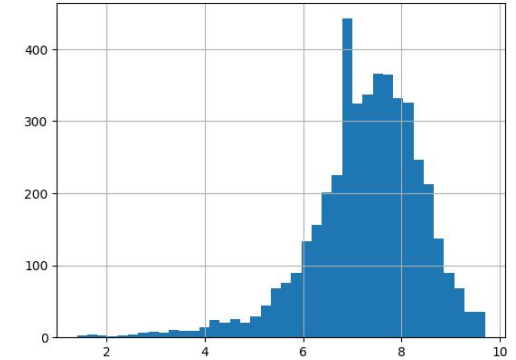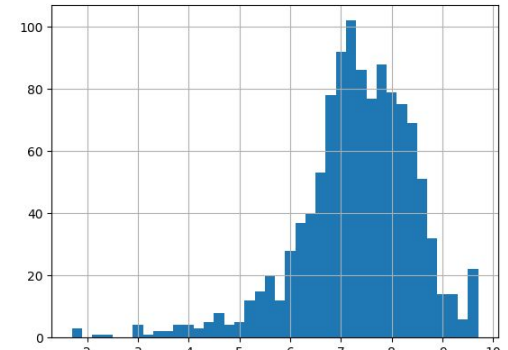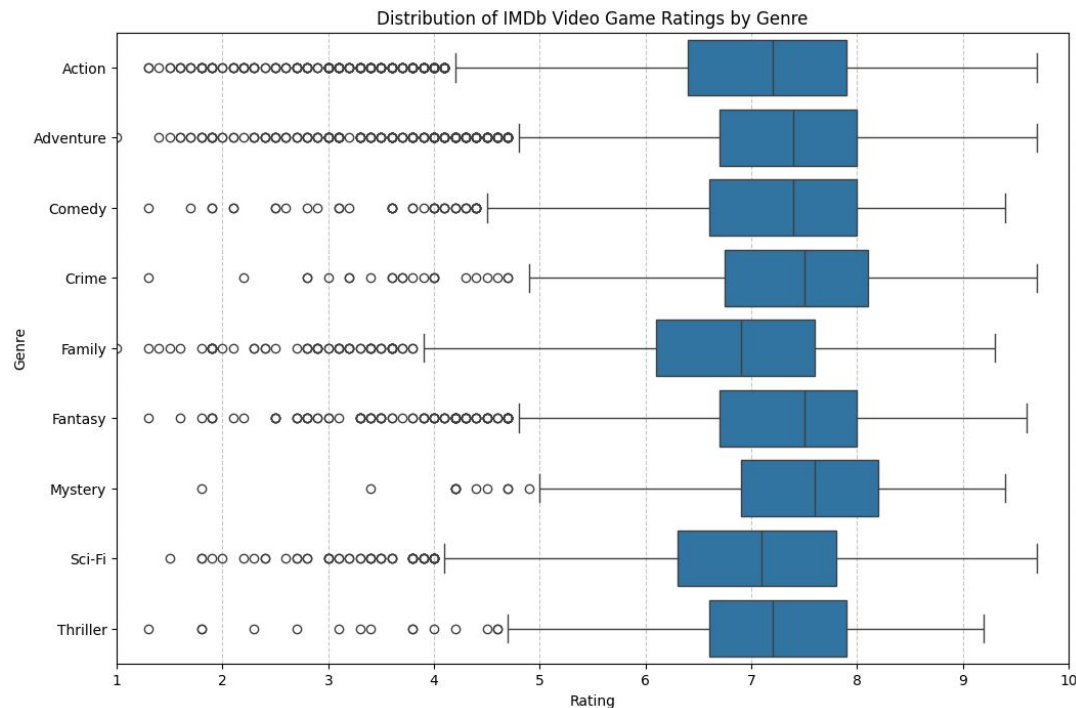
genre	min	Q1	Median	Q3	max
Action	1.3	6.4	7.2	7.9	9.7
Adventure	1	6.7	7.4	8	9.7
Comedy	1.3	6.6	7.4	8	9.4
Crime	1.3	6.75	7.5	8.1	9.7
Family	1	6.1	6.9	7.6	9.3
Fantasy	1.3	6.7	7.5	8	9.6
Mystery	1.8	6.9	7.6	8.2	9.4
Sci-Fi	1.5	6.3	7.1	7.8	9.7
Thriller	1.3	6.6	7.2	7.9	9.2

Even more results

- Genre can have a significant influence on ratings in games with action and adventure consistently performing better.

- Action and adventure continued to stay the most dominant throughout the years.

- Age certificate does not have a significant influence on how a game is rated or its popularity.

Genre Trends Over Years (Q4 count)
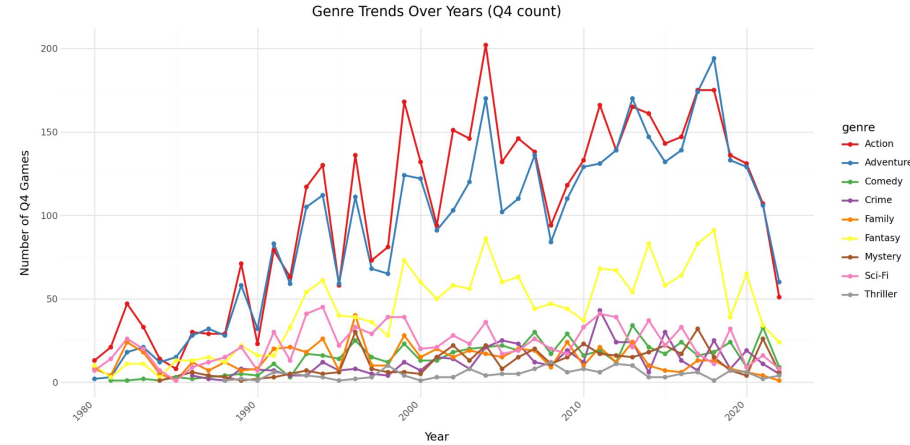
Discussion

- There is a significant connection between the genre of a game to the rating of a game.

- There is a slight connection between age and rating although it's almost negligible.

- Action and adventure continue to dominate through the years with ratings seemingly decreasing overtime

Coefficient	Estimate	p-value
Intercept	6.2449	0.0
Year[2000s]	0.9898	0.0
Year[2010s]	0.9201	0.0
Year[2020s+]	0.8799	0.0
Year[1990s]	0.7458	0.0
Genre[Adventure]	0.4916	0.0
Year[2000s] × Genre[Adventure]	-0.4087	0.0
Genre[Fantasy]	0.433	0.0
Year[2010s] × Genre[Adventure]	-0.3635	0.0
Genre[Mystery]	1.0028	0.0001
Year[2020s+] × Genre[Adventure]	-0.3404	0.0005
Genre[Comedy]	0.7037	0.0005
Year[1990s] × Genre[Adventure]	-0.2848	0.0011
Year[2010s] × Genre[Comedy]	-0.6908	0.0011
Year[2000s] × Genre[Comedy]	-0.6827	0.0014
Genre[Family]	-0.2653	0.0019
Year[2020s+] × Genre[Fantasy]	-0.3747	0.0021
Year[2020s+] × Genre[Mystery]	-0.8262	0.0051
Year[1990s] × Genre[Comedy]	-0.5996	0.0062
Year[2000s] × Genre[Mystery]	-0.7398	0.0067
Year[1990s] × Genre[Fantasy]	-0.2811	0.0093
Year[2010s] × Genre[Mystery]	-0.6972	0.0095
Year[2010s] × Genre[Fantasy]	-0.2541	0.0119
Year[2000s] × Genre[Fantasy]	-0.2528	0.0128
Year[2020s+] × Genre[Comedy]	-0.5733	0.0135
Year[1990s] × Genre[Mystery]	-0.5367	0.06
Genre[Crime]	0.3906	0.0691
Year[1990s] × Genre[Crime]	-0.3831	0.1159
Genre[Sci-Fi]	-0.1227	0.1192
Year[1990s] × Genre[Sci-Fi]	0.1423	0.1469
Year[2000s] × Genre[Crime]	-0.2542	0.2675
Year[2000s] × Genre[Family]	-0.1108	0.28
Year[2010s] × Genre[Sci-Fi]	0.0953	0.3187
Year[2010s] × Genre[Thriller]	-0.2909	0.4299
Year[2000s] × Genre[Sci-Fi]	0.0613	0.5285
Year[1990s] × Genre[Thriller]	0.2488	0.5399
Year[2020s+] × Genre[Crime]	-0.1645	0.5406
Year[2020s+] × Genre[Thriller]	-0.2424	0.5411
Year[2020s+] × Genre[Family]	-0.0938	0.5922
Genre[Thriller]	0.1551	0.663
Year[2000s] × Genre[Thriller]	-0.1608	0.6674
Year[2010s] × Genre[Family]	-0.0383	0.7205
Year[2010s] × Genre[Crime]	-0.068	0.7664
Year[2020s+] × Genre[Sci-Fi]	0.0269	0.8412
Year[1990s] × Genre[Family]	0.0182	0.8677

References

(1)   Talay, M. A. (2022, September 2). *IMDB video games*. Kaggle.
       https://www.kaggle.com/datasets/muhammadadiltalay/imdb-video-games

(2)   Rose, M. (2023, November 19). *Report: Obsidian missed fallout: New Vegas bonus by one Metacritic point*. Report:
       Obsidian missed Fallout: New Vegas bonus by one Metacritic point.
       https://www.gamedeveloper.com/business/report-obsidian-missed-i-fallout-new-vegas-i-bonus-by-one-metacritic-point
(3)

Thank You!

The Community TECH • IMDb