CAPSTONE PROJECT

Fandango Movie rating discrepancy analysis using python

Presented By:
Lokendra Sowmiyan S
3rd year – Electrical and Electronic Engineering
SSM INSTITUTE OF ENGINEERING AND TECHNOLOGY

edunet foundation

OUTLINE

- Problem Statement
- Proposed System/Solution
- System Development Approach
- Algorithm & Deployment
- Result
- Conclusion
- Future Scope
- References

Problem Statement

- “Explore and analyze the potential rating discrepancies in Fandango movie ratings compared to other movie rating platforms. Utilize Python to gather, clean, and analyze data, aiming to uncover any biases or inconsistencies in Fandango's rating system compared to objective movie rating sources like IMDb or Rotten Tomatoes. Identify patterns, outliers, and potential factors contributing to any observed differences in ratings."

Proposed Solution

Data Collection:

- Obtain movie ratings data from Fandango and another reliable source (e.g., IMDb).

Data Cleaning:

- Clean the data to ensure accuracy and consistency.

Data Analysis:

- Calculate summary statistics (mean, median, standard deviation, etc.) for both Fandango and the other source.
- Visualize the distribution of ratings from both sources using histograms or boxplots.
- Perform hypothesis testing to determine if there's a significant difference between the ratings.

System Approach

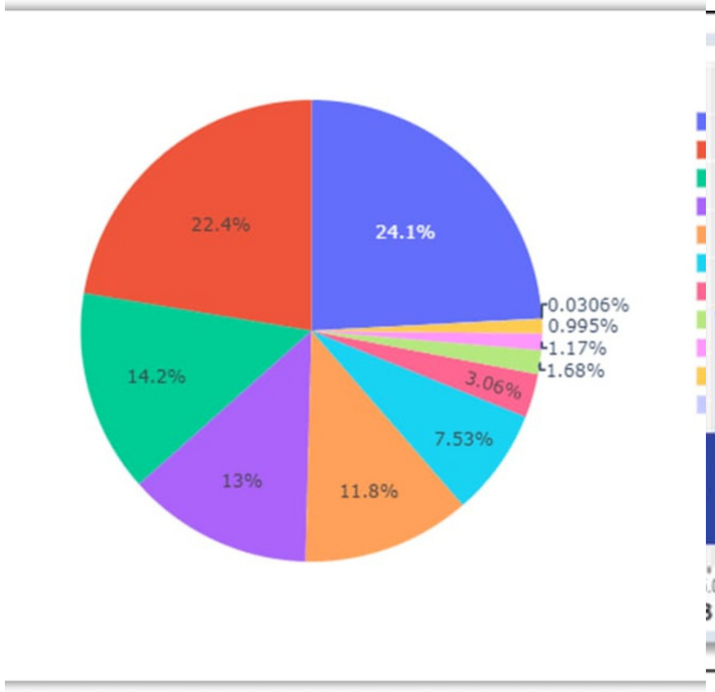
- **Problem Definition**:
- Clearly define the objective of the analysis, such as understanding the extent of rating inflation on Fandango compared to other platforms.
- **Scope Definition:**
- Determine the scope of the analysis, including which movies, time period, and comparison platforms will be included.
- **Data Collection:**
- Gather Fandango ratings data using web scraping or an API.
- Collect ratings data from alternative sources like IMDb or Rotten Tomatoes.
- Ensure data integrity and completeness.
- **Data Preprocessing:**
- Clean the data by handling missing values, inconsistencies, and outliers.
- Normalize ratings to a common scale if necessary.
- Explore the data to understand its distribution and characteristics.

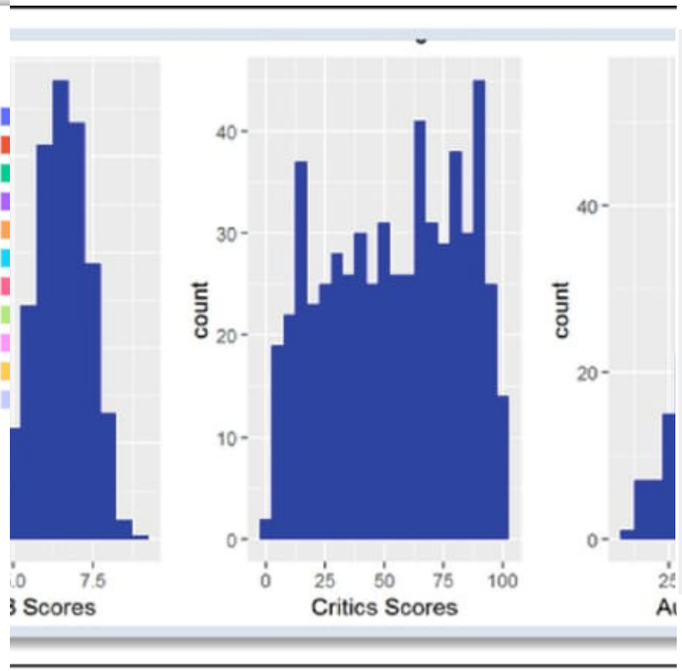
Algorithm & Deployment

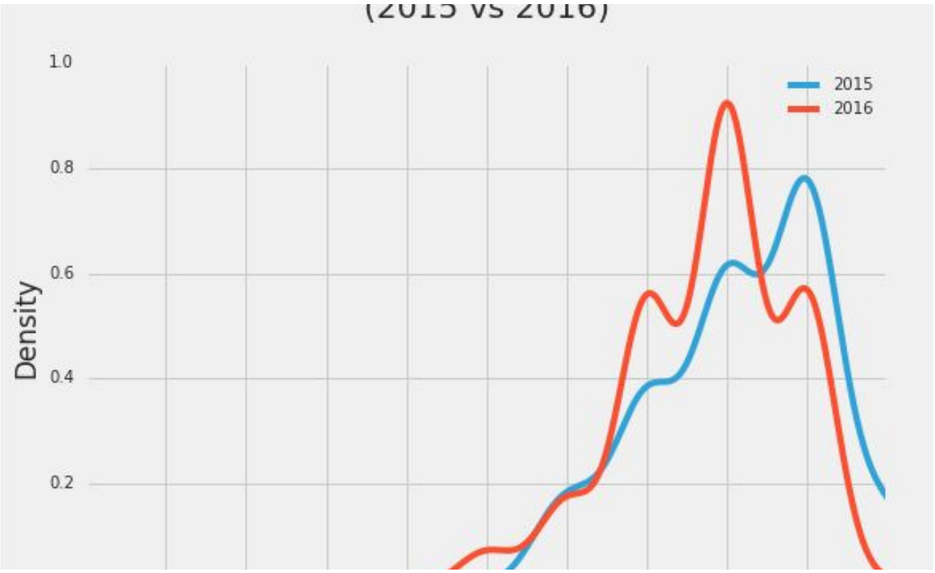
- **Algorithm Development:**

1. **Data Collection:** Utilize web scraping or APIs to gather Fandango movie ratings data and ratings from alternative sources such as IMDb or Rotten Tomatoes.
2. **Data Preprocessing**: Clean the collected data, handle missing values, and normalize ratings if needed.
3. **Analysis**: Calculate summary statistics, visualize rating distributions, and conduct hypothesis testing to identify discrepancies between Fandango ratings and ratings from other sources.
4. **Insights Generation:** Interpret the analysis results to understand the reasons behind rating differences and provide actionable insights.

Result


24.1%
22.4%
14.2%
13%
11.8%
7.53%
3.06%
1.68%
1.17%
0.995%
0.0306%


count
Critics Scores
count


(2015 vs 2016)
Density
2015
2016

Conclusion

- Our analysis showed that there's indeed a slight difference between Fandango's ratings for popular movies in 2015 and Fandango's ratings for popular movies in 2016. We also determined that, on average, popular movies released in 2016 were rated lower on Fandango than popular movies released in 2015.

Future scope

“Analyzing Fandango movie rating discrepancies using Python has a promising future scope. You could build predictive models to detect inconsistencies, create data visualization tools for insights, or develop automated monitoring systems. Exploring sentiment analysis or machine learning algorithms to understand user reviews could also enhance analysis accuracy. Additionally, integrating with other movie databases for comparative analysis would be beneficial.”

References

- https://www.kaggle.com/datasets
- https://pandas.pydata.org/pandas-docs/stable/user guide/index.html
- https://seaborn.pydata.org/
- https://matplotlib.org/stable/contents.html
- https://chat.openal.com

THANK YOU

edunet
foundation