



# **PREDICTING VOTING SCORE USING 2018 IMDB**

**HUILIN TANG**

**UCLA STATISTICS 418 – TOOLS IN DATA SCIENCE  
SPRING 2019**



# ATTAINING DATA – WEB SCRAPER

All

Watch Now  
For Free

Movies, TV  
& Showtimes


Celebs, Events  
& Photos

News & Community

## Released between 2018-01-01 and 2019-01-01 (Sorted by Number of Votes Descending)

1-50 of 291,602 titles. | [Next »](#) View Mode: [Compact](#) | [Details](#)

Sort by: [Popularity](#) | [A-Z](#) | [User Rating](#) | **[Number of Votes](#)** | [US Box Office](#) | [Runtime](#) | [Year](#) | [Release Date](#) | [Date of Your Rating](#) | [Your Rating](#)



1. **Avengers: Infinity War** (2018)


PG-13 | 149 min | Action, Adventure, Sci-Fi

★ **8.5** ☆ [Rate this](#) **68** Metascore

The Avengers and their allies must be willing to sacrifice all in an attempt to defeat the powerful Thanos before his blitz of devastation and ruin puts an end to the universe.

Directors: [Anthony Russo](#), [Joe Russo](#) | Stars: [Robert Downey Jr.](#), [Chris Hemsworth](#), [Mark Ruffalo](#), [Chris Evans](#)

Votes: **666,993** Gross: \$678.82M



2. **Black Panther** (2018)

PG-13 | 134 min | Action, Adventure, Sci-Fi

★ **7.3** ☆ [Rate this](#) **88** Metascore

T'Challa, heir to the hidden but advanced kingdom of Wakanda, must step forward to lead his people into a new future and must confront a challenger from his country's past

IMDb  
score

```
<div class="list-item-content">
  <h3 class="list-item-header">
    <span class="list-item-index unbold text-primary">1.
    </span>
    <a href="/title/tt4154756/?ref=adv_li_tt">Avengers:
    Infinity War</a>
    <span class="list-item-year text-muted unbold">(2018)
    </span>
  </h3>
  <p class="text-muted">...</p>
  <div class="ratings-bar">
    <div class="inline-block ratings-imdb-rating" name="ir"
      data-value="8.5">...</div>
    <div class="inline-block ratings-user-rating">...</div>
    <div class="inline-block ratings-metascore">...</div>
  </div>
  <p class="text-muted">...</p>
  <p class="text-muted">...</p>
</div>
```

html.scriptsOn.gr\_imdb\_com body#styleguide-v2.fixed script

list-item mode-advanced

1 of 50

Cancel

Styles Event Listeners DOM Breakpoints Properties Accessibility

Filter

:hov .cls +

Console What's New x

Highlights from the Chrome 74 update

Highlight all nodes affected by CSS property

Hover over a CSS property like padding or margin in the Styles pane to highlight all nodes affected by that declaration.

Lighthouse v4 in the Audits panel

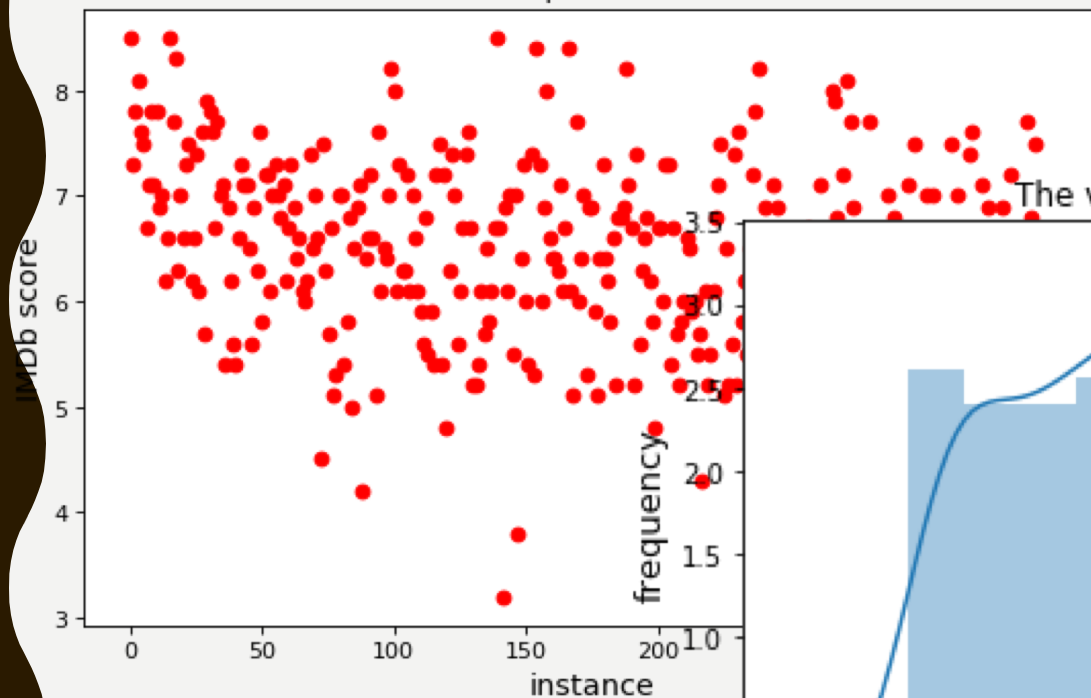
Featuring a new "tap targets" audit for checking that mobile links and buttons are properly sized, and a new UI for PWA reports.

WebSocket binary message viewer

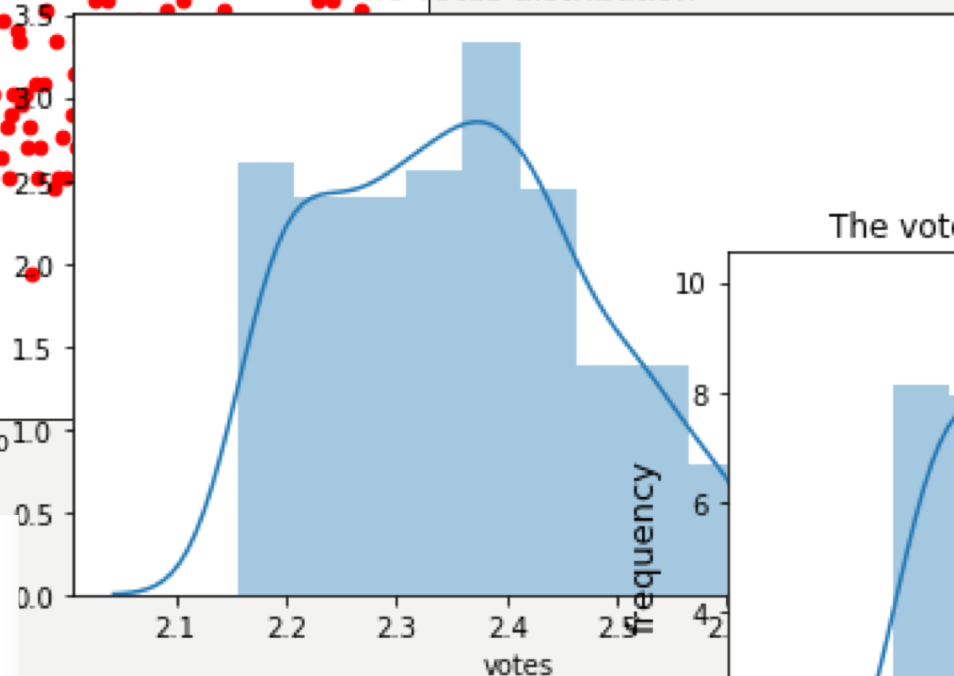


# DATA EXPLORATION ANALYSIS

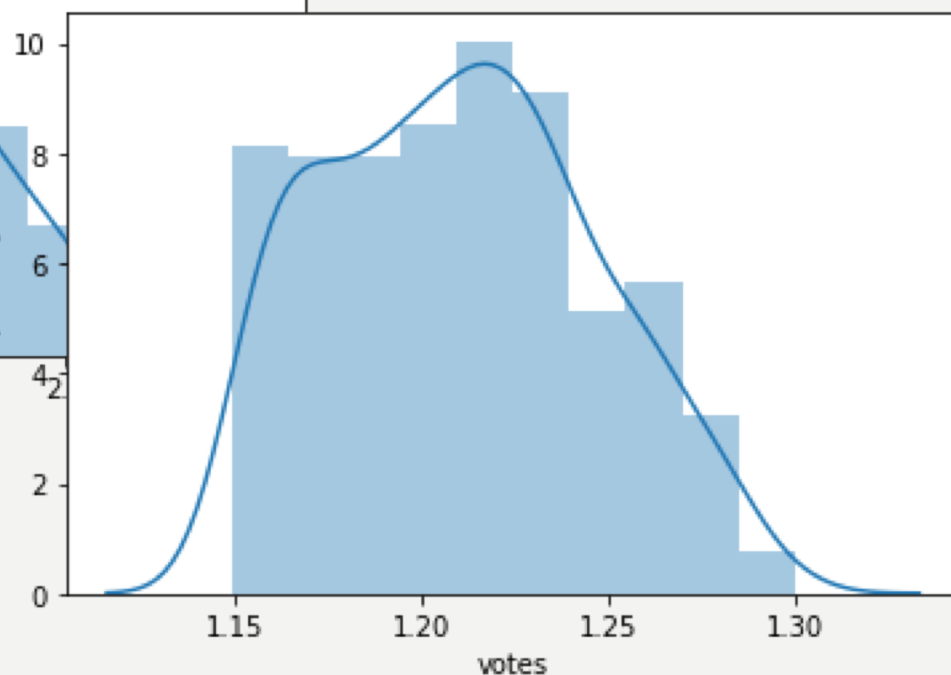
The scatterplot of IMDb score



The votes distribution

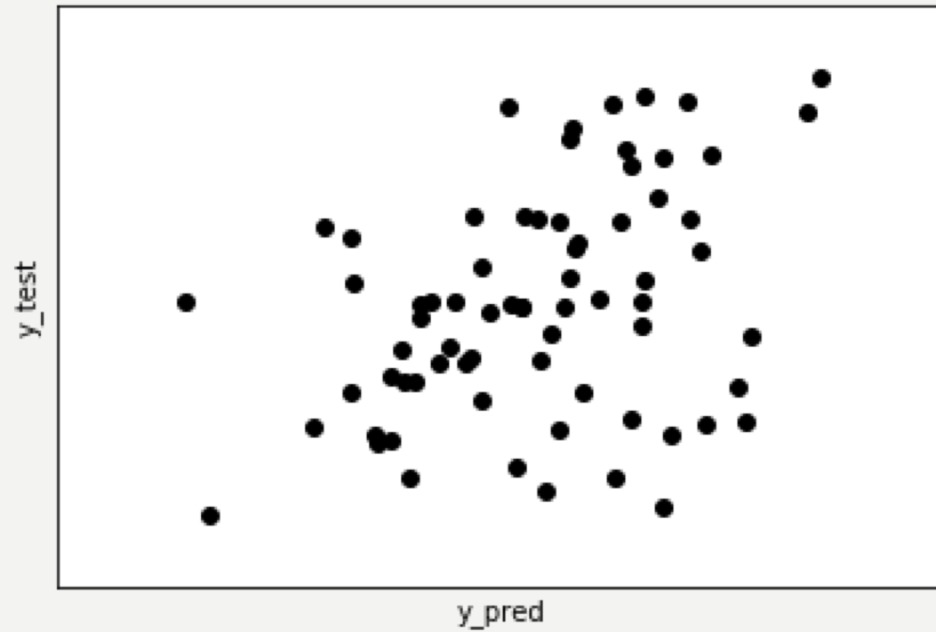


The votes distribution after transformation



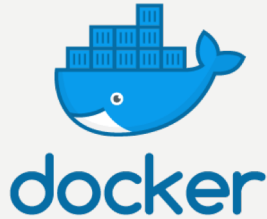
	imdb	metascore	movie	votes	year
0	8.5	68	Avengers: Infinity War	667105	(2018)
1	7.3	88	Black Panther	514171	(2018)
2	7.8	66	Deadpool 2	389509	(2018)
3	8.1	49	Bohemian Rhapsody	342427	(2018)
4	7.6	82	A Quiet Place	303571	(2018)

# MODEL ANALYSIS



Mean squared error: 0.0010889836135416554

# DEPLOYMENT



- Docker & Flask & AWS
- Input: Imdb score, Metascore
- Output: Predicted Votes
- `curl -H "Content-Type: application/json" -X POST -d '{"imdb":"6.7", "metascore":"53"}' "http://localhost:5000/votes"`
- Predicted result

```
[s-164-67-229-202:docker MsBrilliant$ curl http://localhost:5000/]  
[The server is up - good job!!!s-164-67-229-202:docker MsBrilliant$ curl -H "Cont  
  POST -d '{"imdb":"6.7", "metascore":"53"}' "http://localhost:5000/votes"  
{  
  "predicted votes": 52869.39191900185  
}
```