Fandango movie ratings

Emmanuel Messori

15/09/2021

Point of departure

In October 2015, data journalist Dave Hickey analysed data from film reviews coming from the website . In his analysis he states :

> Of the 437 films with at least one review, 98 percent had a 3-star rating or higher and 75 percent had a 4-star rating or higher.
>
> It seemed nearly impossible for a movie to fail by Fandango's standards.
>
> When I focused on movies that had 308 or more user reviews,9 none of the 209 films had below a 3-star rating. Seventy-eight percent had a rating of 4 stars or higher.

link to his article

In comparing Fandango's scores to the most popular competitors' website he found out that most of the times they were higher.

Eventually he illustrated a huge discrepancy between the rating displayed in the page html code and the rating actually displayed, caused by rounding up to the nearest half star (e.g. a rating of 4.1 would be rounded to 4.5)

Fandango's officials replied that the biased rounding off was caused by a bug in their system rather than being intentional, and they promised to fix the bug as soon as possible. Presumably, this has already happened.

Nowadays the rating displayed are coming from the site (which by the way was purchased by Fandango)

In this project, we'll analyze more recent movie ratings data(2016-2017) to determine whether there has been any change in Fandango's rating system after Hickey's analysis.

Retrieving the data

```
library(tidyverse)
old_data <- read_csv("https://raw.githubusercontent.com/mircealex/Movie_ratings_2016_17/master/fandango_
new_data <- read_csv("https://raw.githubusercontent.com/mircealex/Movie_ratings_2016_17/master/movie_rat
```

We're going to focus on the columns of the two datasets that offer information about Fandango's ratings.

```
old_data %>%
  select(FILM, Fandango_Stars, Fandango_Ratingvalue, Fandango_votes, Fandango_Difference) -> old_data_fi
new_data %>%
  select(movie, year, fandango) -> new_data_filtered
```

Purposive Sampling

Our population of interest was made in theory by all the movies present in the database in the years 2016-2017. But we are not working with random samples here, since ‘`fandango_score_comparison.csv` (Hickey’s data):

> contains every film that has a Rotten Tomatoes rating, a RT User rating, a Metacritic score, a Metacritic User score, and IMDb score, and at least 30 fan reviews on Fandango

and `movie_ratings_16_17.csv`

> contains movie ratings data for 214 of the most popular movies (with a significant number of votes) released in 2016 and 2017.

New analysis goal

The initial goal of our analysis would be to compare to samples of movies of the Fandango website, one which comprises films released just before Hickey’s analysis and another more recent one, to see whether there has been a change in the rating system.

Since the samples are not representative of all the movies in the database regardless of the releasing year, we will consider the samples to be representative at least of the most popular movies. Our objective is now to check whether there’s any difference in Fandango’s ratings for popular movies in 2015 and Fandango’s ratings for popular movies in 2016-2017.

The term “popular” is vague and we need to define it with precision before continuing. Let’s see if we can use Hickey’s benchmark of 30 fan ratings and consider a movie as “popular” only if it has 30 fan ratings or more on Fandango’s website.

Checking the samples

The `old_data` dataset which contains Hickey’s selection, as stated in the description contains movies which have at least 30 reviews (information we can check on the data):

```
summary(old_data$Fandango_votes)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    35.0   222.2  1446.0  3848.8  4439.5 34846.0
```

There are at least 35 reviews.

To see if the new sample is still representative for the new analysis question, we should check whether these movies have at least 30 reviews (like for Hickey’s data). But since Fandango has since abandoned the fan rating system, we could compare for example the number of ratings present on the Rotten Tomatoes site. But, again, since the user base is different we cannot really say with confidence whether these review numbers are comparable to the Fandango fan ratings. We’ll just content ourselves for now to make the assumption that the sample is representative.

We’ll filter the data to focus just on the films released in 2015 and 2016.

```
old_data_filtered %>%
  separate(FILM, c("title", "year"), sep =  "\\(") %>%
  mutate(year = parse_number(year)) %>%
  filter(year==2015) -> movies2015
head(movies2015)
```

title	year	Fandango_Stars	Fandango_Ratingvalue	Fandango_votes	Fandango_Difference
Avengers: Age of Ultron	2015	5.0	4.5	14846	0.5
Cinderella	2015	5.0	4.5	12640	0.5
Ant-Man	2015	5.0	4.5	12055	0.5
Do You Believe?	2015	5.0	4.5	1793	0.5
Hot Tub Time Machine 2	2015	3.5	3.0	1021	0.5
The Water Diviner	2015	4.5	4.0	397	0.5

```
new_data_filtered %>%
  filter(year == 2016) -> movies2016
head(movies2016)
```

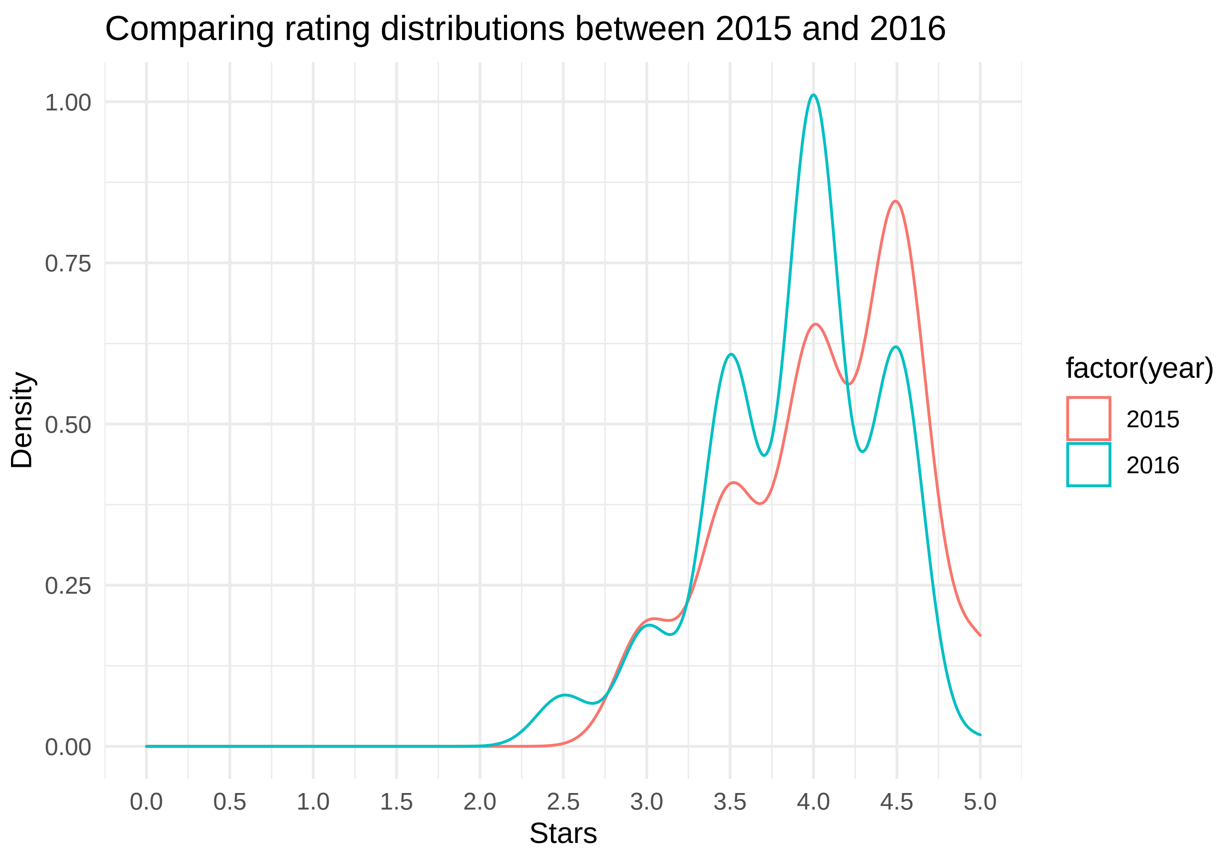
movie	year	fandango
10 Cloverfield Lane	2016	3.5
13 Hours	2016	4.5
A Cure for Wellness	2016	3.0
A Hologram for the King	2016	3.0
A Monster Calls	2016	4.0
A Street Cat Named Bob	2016	4.5

Distribution comparison

Let's now compare the two rating distributions with density kernel plots. For the 2015 movies, we'll use the `Fandango_Stars` variable, containing the official rating:

```
movies2015 %>%
  select (1:3) %>%
  rename(movie = title, fandango = Fandango_Stars) %>%
  bind_rows(movies2016) -> all_movies

  ggplot(all_movies, aes(fandango, color = factor(year))) +
  geom_density() +
  theme_minimal() +
  labs(title="Comparing rating distributions between 2015 and 2016",
       x="Stars", y="Density") +
  scale_x_continuous(breaks = seq(0,5,0.5), limits = c(0,5))
```



Though both distributions are left skewed, the 2016 distribution shows an evident shift toward the left. It seems that the 2016 movies are lower rated. The min value is 2.5, and peak density is at 4 (the rating from 2015 peak at 4.5).

Frequency Tables

```
count(movies2015, Fandango_Stars) %>%
  mutate(prop = round(n / sum(n) * 100, 2))
```

Fandango_Stars	n	prop
3.0	11	8.53
3.5	23	17.83
4.0	37	28.68
4.5	49	37.98
5.0	9	6.98

```
count(movies2016, fandango) %>%
  mutate(prop = round(n / sum(n) * 100, 2))
```

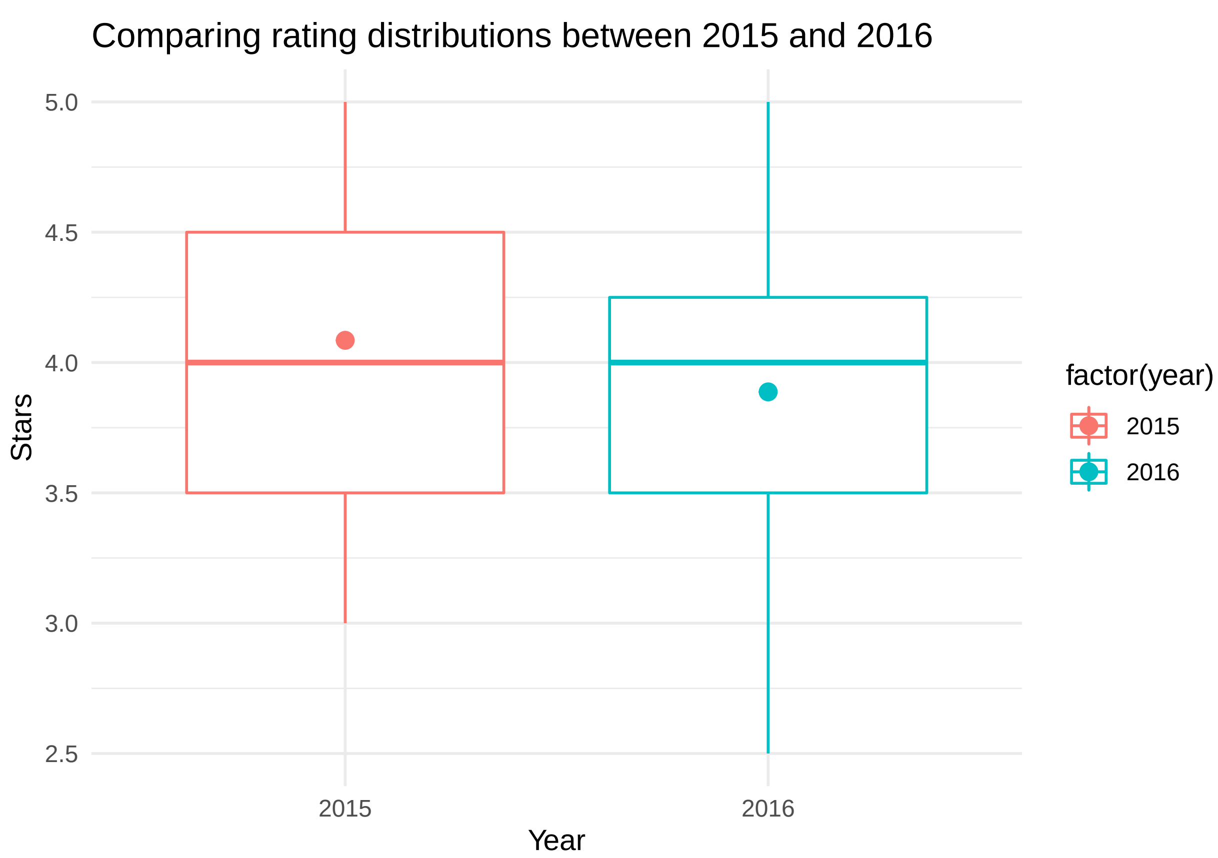
fandango	n	prop
2.5	6	3.14
3.0	14	7.33
3.5	46	24.08
4.0	77	40.31
4.5	47	24.61
5.0	1	0.52

The difference between the two distributions is less evident through the frequency tables: though 2015 rating have higher percentages for the highest ratings(4.5 and 5), 2016 movies have a higher proportion of entries in the mid range (4.0).

Studying the tendency

The 2016 mean rating is ~ 0.2 stars lower than in 2015.

```
ggplot(all_movies, aes(factor(year), fandango, color = factor(year))) +
geom_boxplot() +
theme_minimal() +
labs(title="Comparing rating distributions between 2015 and 2016",
     x="Year", y="Stars") +
scale_y_continuous(breaks = seq(0,5,0.5)) +
stat_summary(fun=mean)
```



The two distributions have the same median but the 2016 has (as already mentioned) a lower mean and also lower 75th quartile:

```
print(summary(movies2015$Fandango_Stars))
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##   3.000   3.500   4.000   4.085   4.500   5.000
```

```
print(summary(movies2016$fandango))
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##   2.500   3.500   4.000   3.887   4.250   5.000
```

Finally we can use a t.test to confirm that this difference in means is statistically significant:

```
t.test(fandango ~ year, data=all_movies)
```

```
## 
##  Welch Two Sample t-test
## 
## data:  fandango by year
## t = 3.2955, df = 264.6, p-value = 0.001117
## alternative hypothesis: true difference in means between group 2015 and group 2016 is not equal to 0
## 95 percent confidence interval:
##  0.07963503 0.31603850
## sample estimates:
## mean in group 2015 mean in group 2016 
##           4.085271           3.887435
```

End notes

With a very low p-value there is little chance that this difference is random. So indeed there has been a slight decrease in the mean ratings between 2015 and 2016. Maybe Fandango has really fixed the rating system after all.