



# Scraping Television Shows from Metacritic

## Web Scraping Project

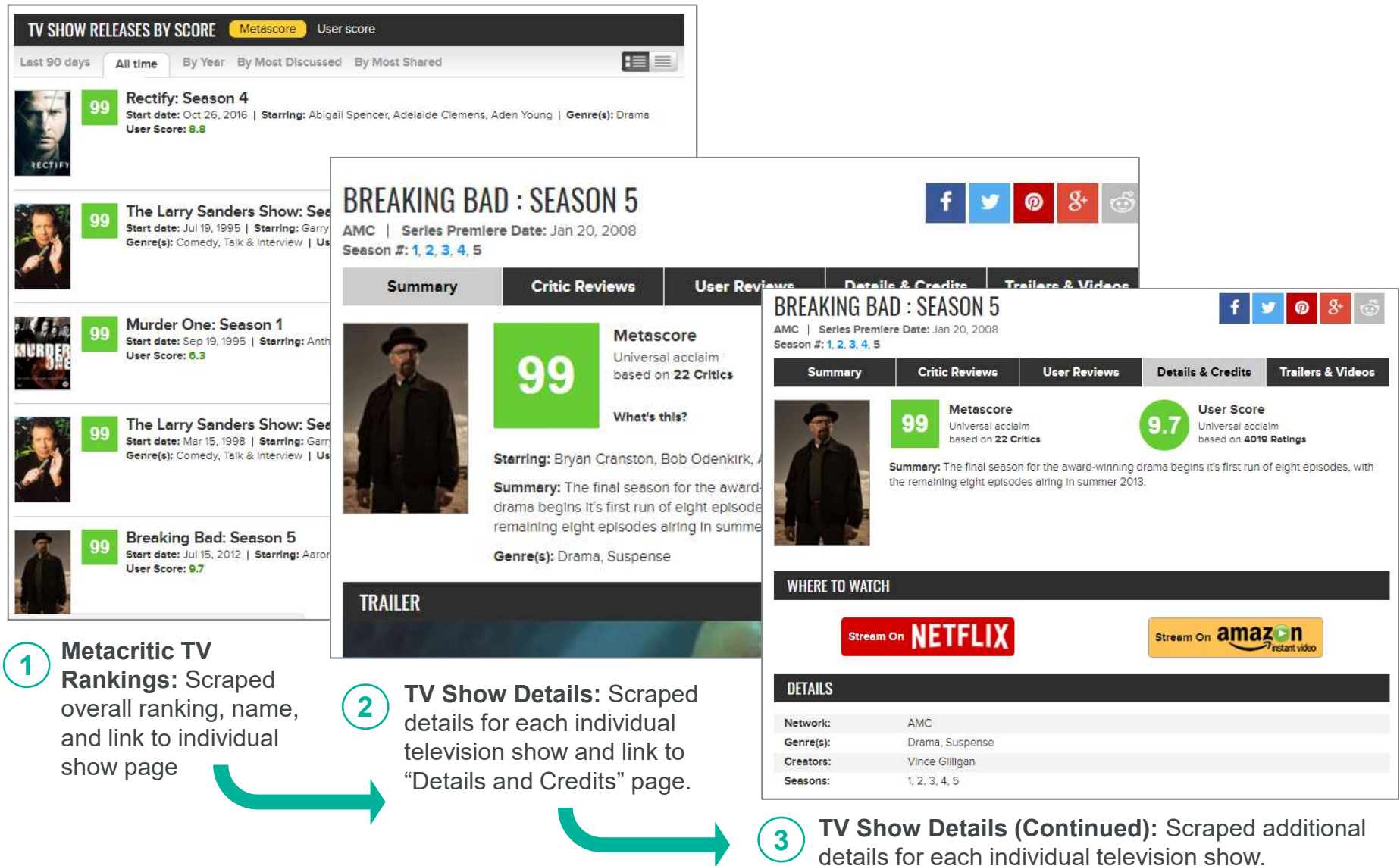
Julia Goldstein

AUGUST 2, 2017

# Motivation & Background

- This period has been dubbed “the new golden age of television” and the era of “peak TV”.
- The Financial Times recently wrote, “We used to complain that there was never anything good on TV. Now we complain about the opposite. ”
- I scraped television show data from Metacritic to answer the following:
  - ① Are there really more “good” shows on now than ever before?
  - ② Are subscription/online networks displacing traditional channels?
  - ③ Should we trust the critics – do critics and users usually agree?
  - ④ What should I actually watch next?

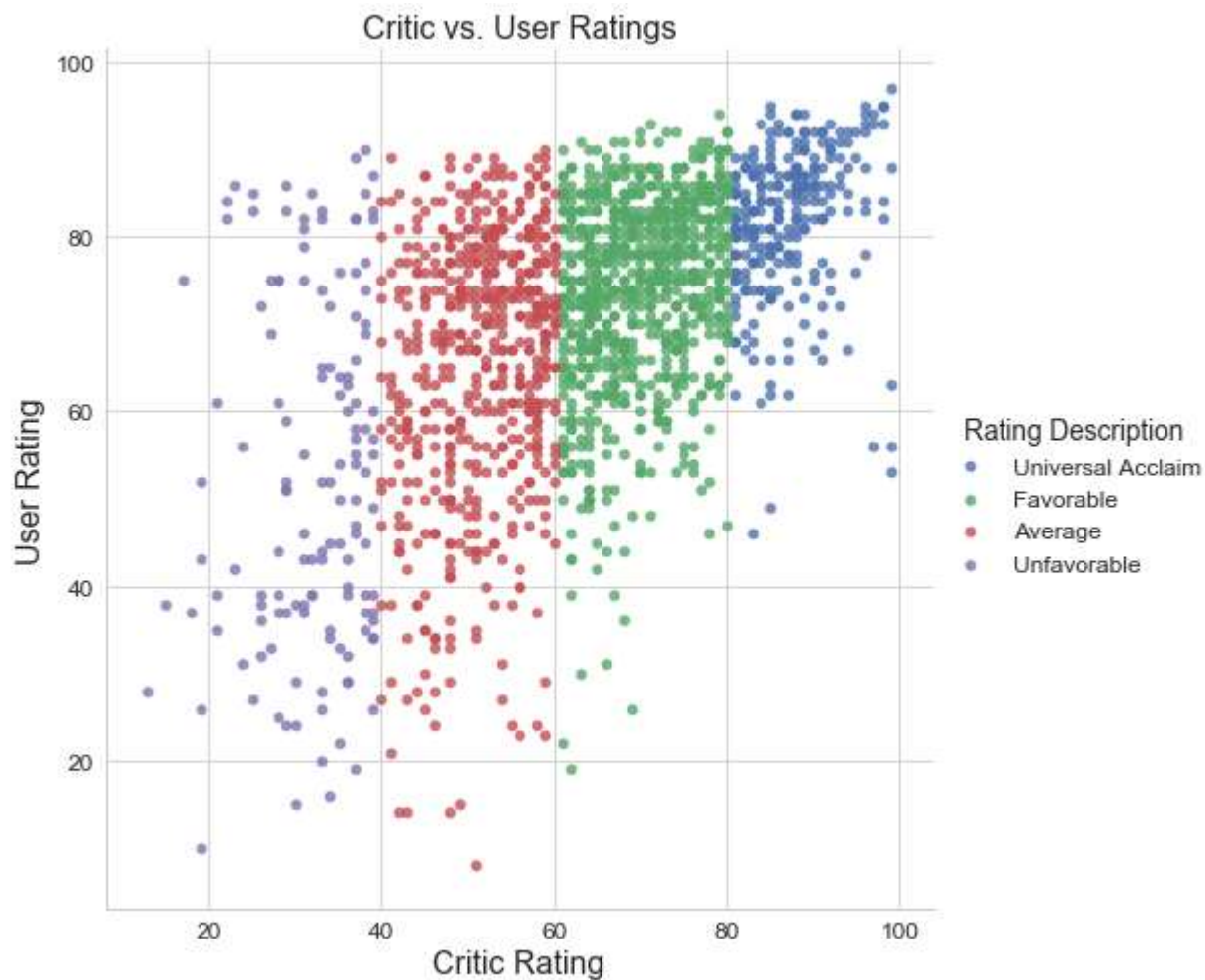
# Web Scraping Process



# Data Description

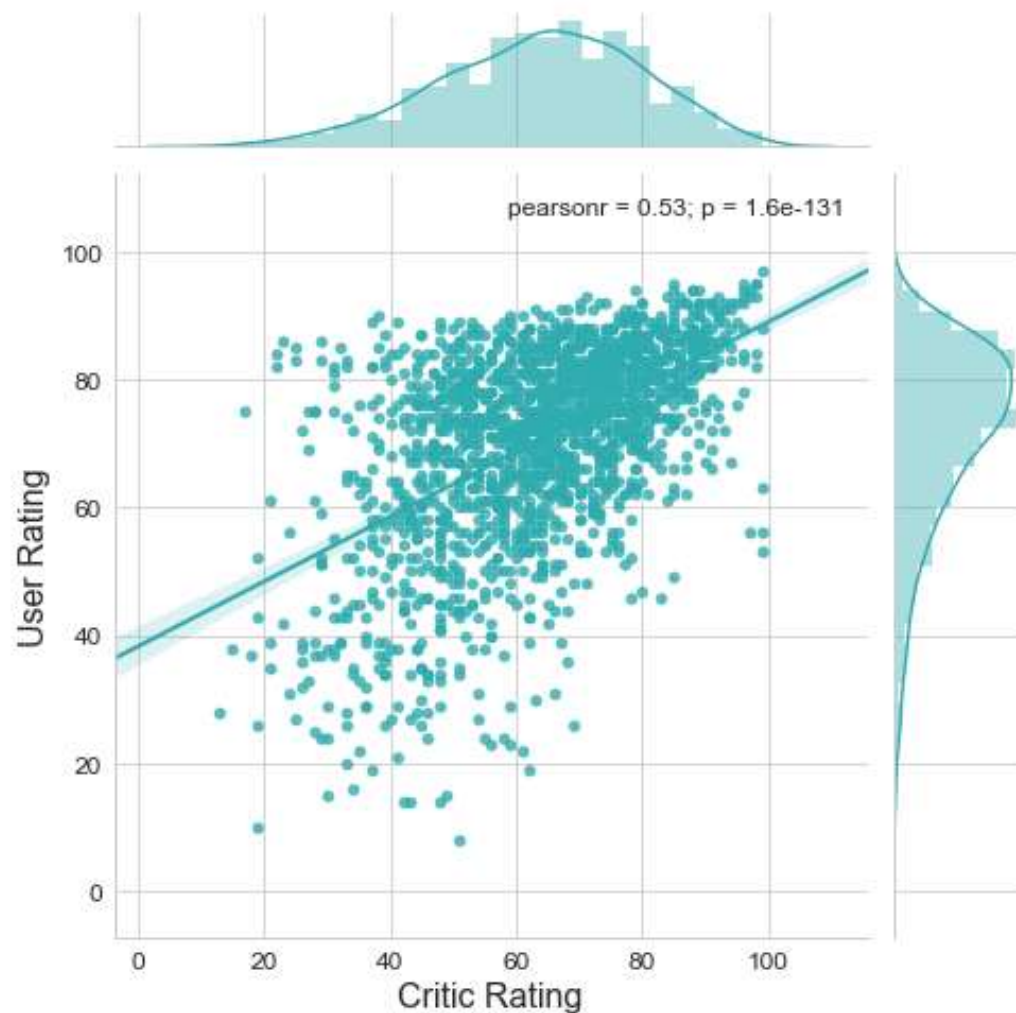
- **Scraping Metacritic yielded the following variables:**
  1. **Show Name**
  2. **Overall Metacritic Rank**
  3. **Critic Rating**
  4. **Number of Critic Reviews**
  5. **Critic Rating Description** (from “Universal Acclaim” to “Unfavorable”)
  6. **User Rating**
  7. **Number of User Reviews**
  8. **Number of Seasons**
  9. **Network**
  10. **Series Premiere Date**
  11. **Season Premiere Date**
  12. **Length** (in minutes)
  13. **Genre(s)**
- **The TV section of Metacritic was launched in 2005, so although the dataset contains shows from prior years, it is likely missing many shows from that time period.**

# Do Critics and Users Agree?



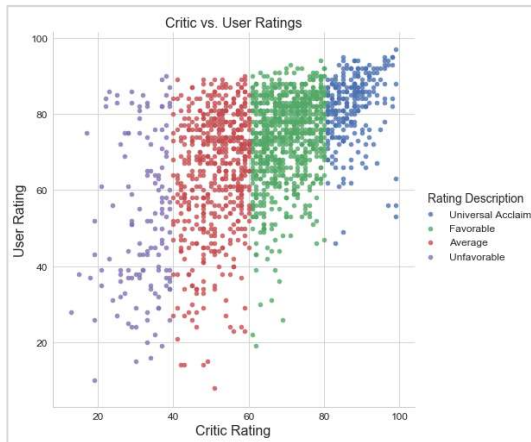
The relationship looks fairly linear, especially as critic ratings improve.

# Do Critics and Users Agree?



There is a significant moderate, positive relationship between critic and user ratings.

# Do Critics and Users Agree?



Was there really more consensus around higher rated shows?

```
# statistical tests
from scipy import stats
unfavorable = metacritic.loc[metacritic['RDescrip'] == "Unfavorable"]
stats.ttest_ind(unfavorable["CriticR"], unfavorable["UserR"], nan_policy="omit")
# unlikely to have same mean
```

```
Ttest_indResult(statistic=-12.200532429830098, pvalue=6.7662115735591363e-28)
```

```
average = metacritic.loc[metacritic['RDescrip'] == "Average"]
stats.ttest_ind(average["CriticR"], average["UserR"], nan_policy="omit")
# unlikely to have same mean
```

```
Ttest_indResult(statistic=-20.178664117289706, pvalue=6.9568416589098526e-78)
```

```
favorable = metacritic.loc[metacritic['RDescrip'] == "Favorable"]
stats.ttest_ind(favorable["CriticR"], favorable["UserR"], nan_policy="omit")
# unlikely to have same mean
```

```
Ttest_indResult(statistic=-9.4490049787126775, pvalue=1.071042474222479e-20)
```

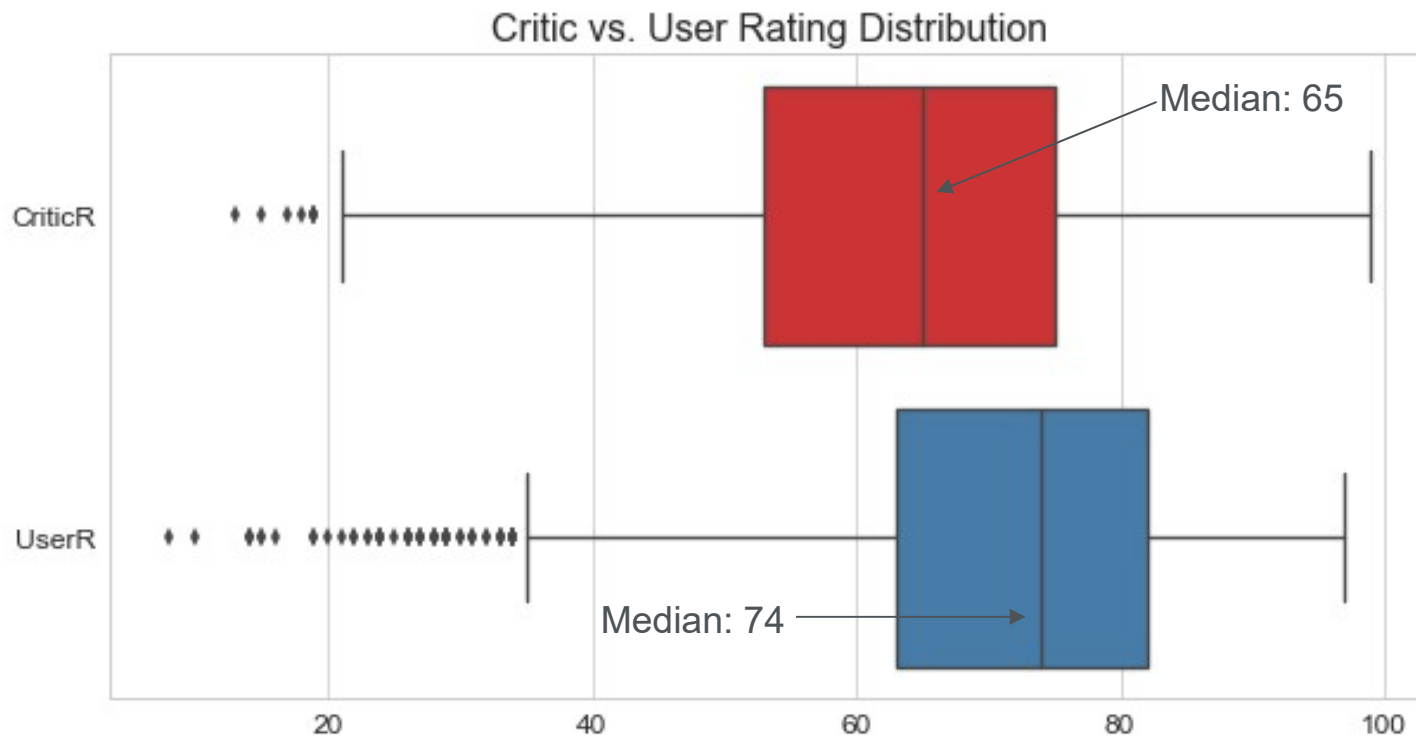
```
acclaim = metacritic.loc[metacritic['RDescrip'] == "Universal Acclaim"]
stats.ttest_ind(acclaim["CriticR"], acclaim["UserR"], nan_policy="omit")
# unlikely to have same mean
```

```
Ttest_indResult(statistic=8.592223892683263, pvalue=9.4680693122106597e-17)
```

The population means of critic and user ratings are different for all test groups ("Unfavorable", "Average", "Favorable" and "Universal Acclaim").



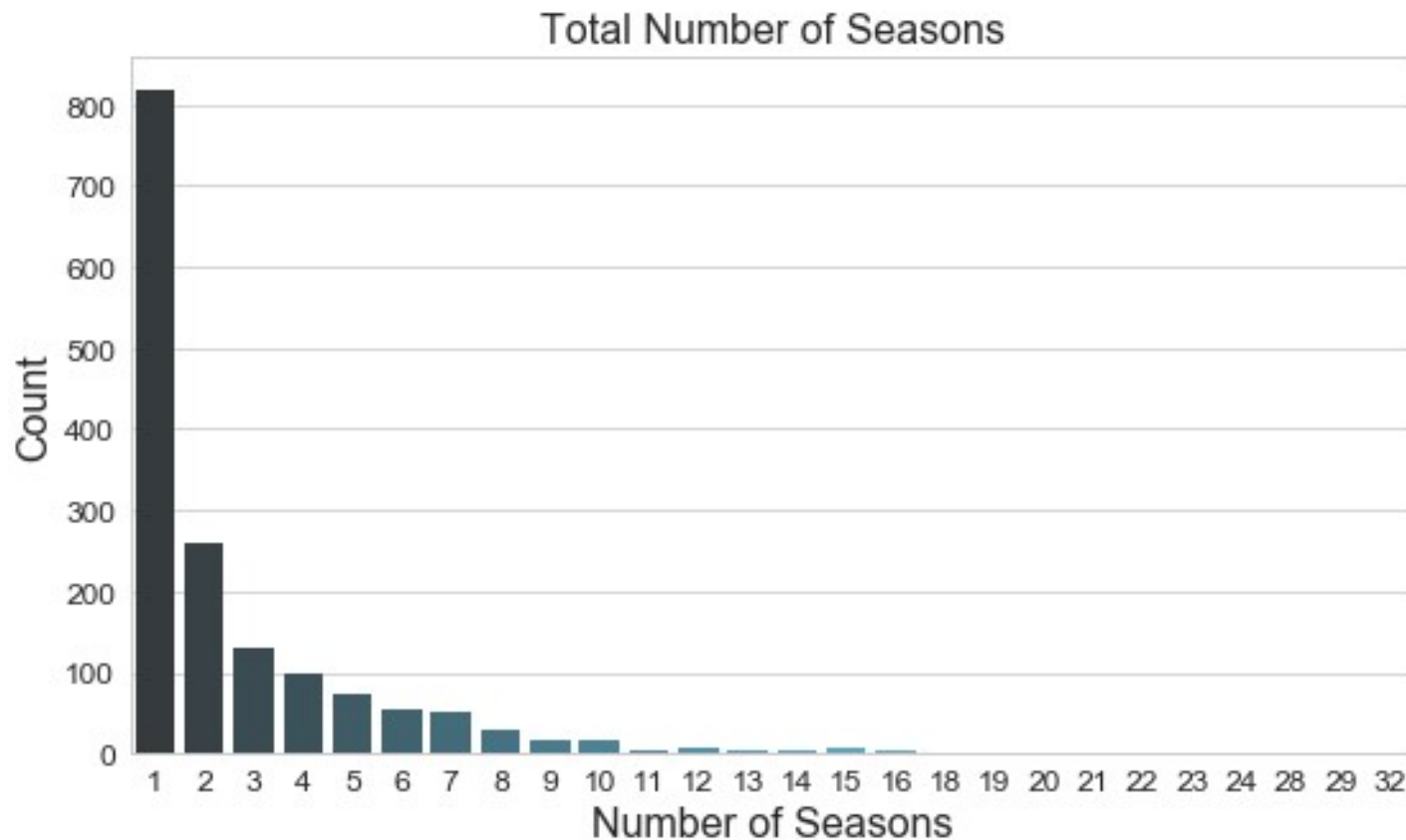
# Do Critics and Users Agree?



Critics are “harsher” than users – their median review is lower.

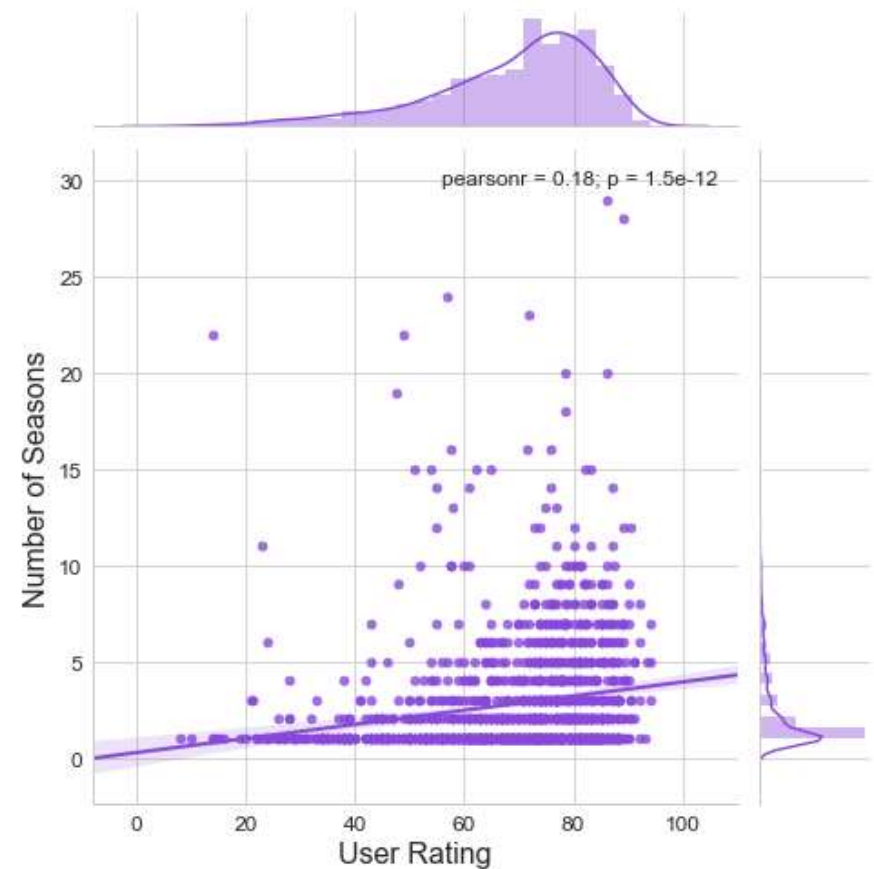
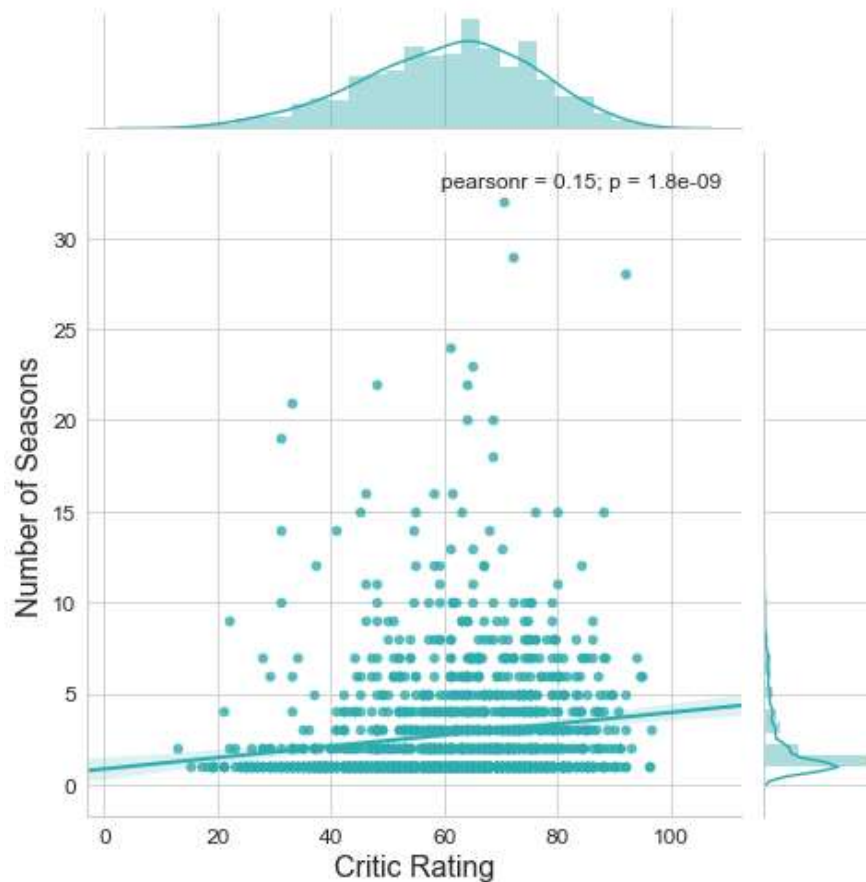


# Do Ratings Matter for a Show's Success?



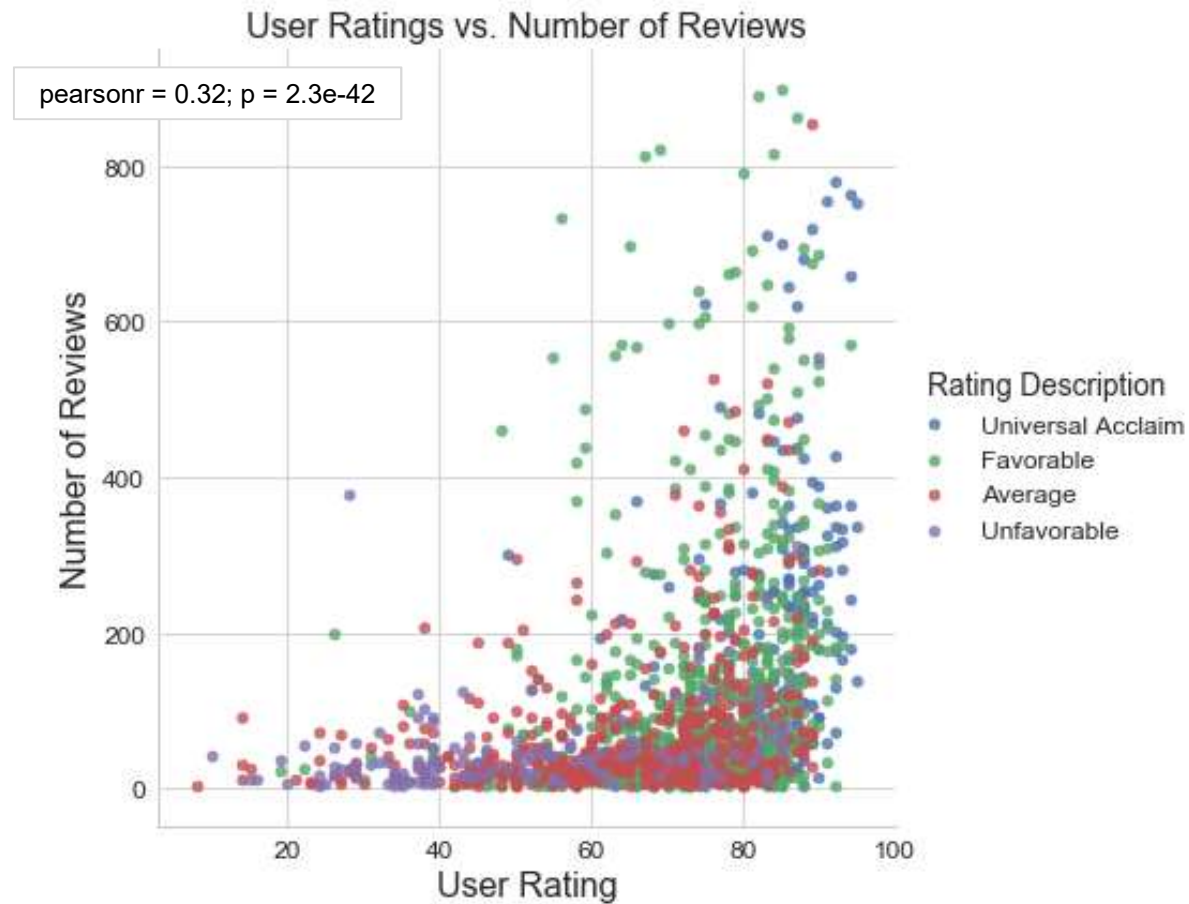
The majority of shows last only one season – and the number decreases from there.

# Do Ratings Matter for a Show's Success?



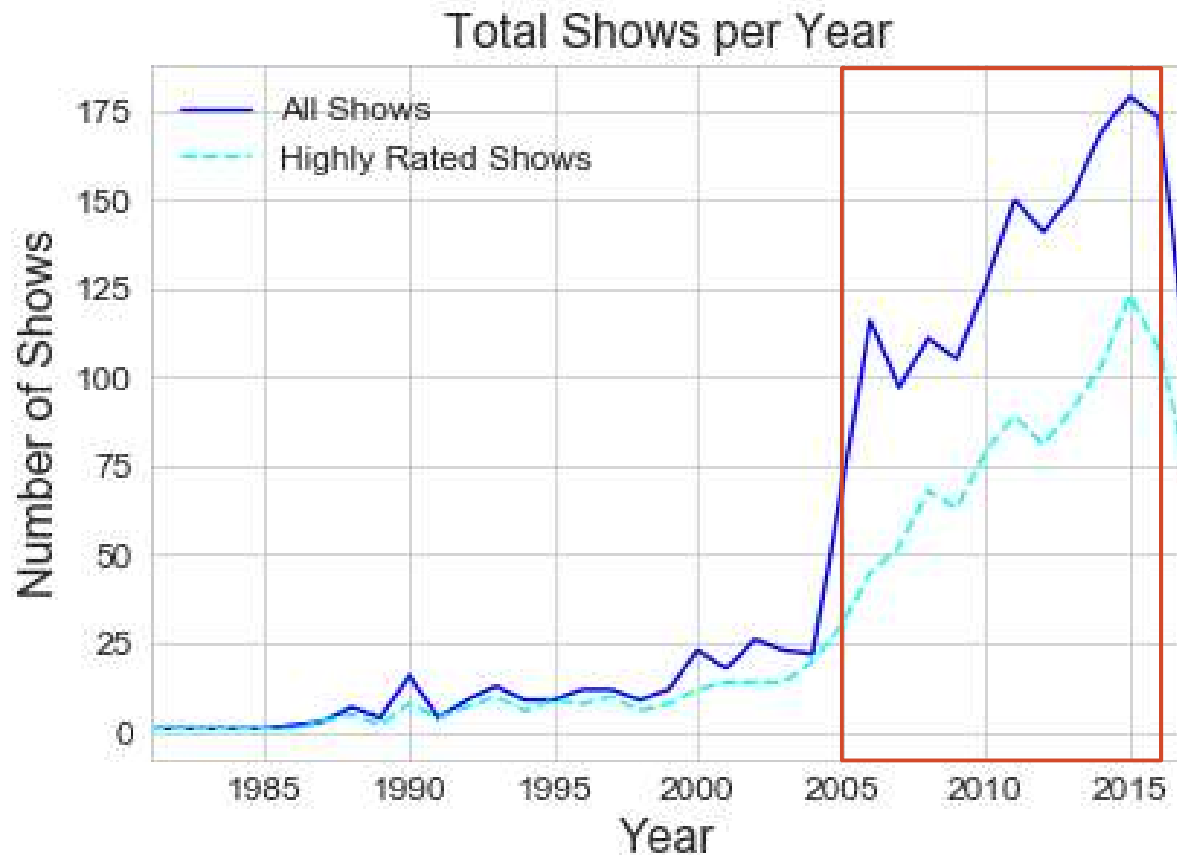
There is a weak (but significant) linear relationship between critic and user ratings and number of seasons.

# Are Users More Likely to Review a Good Show?



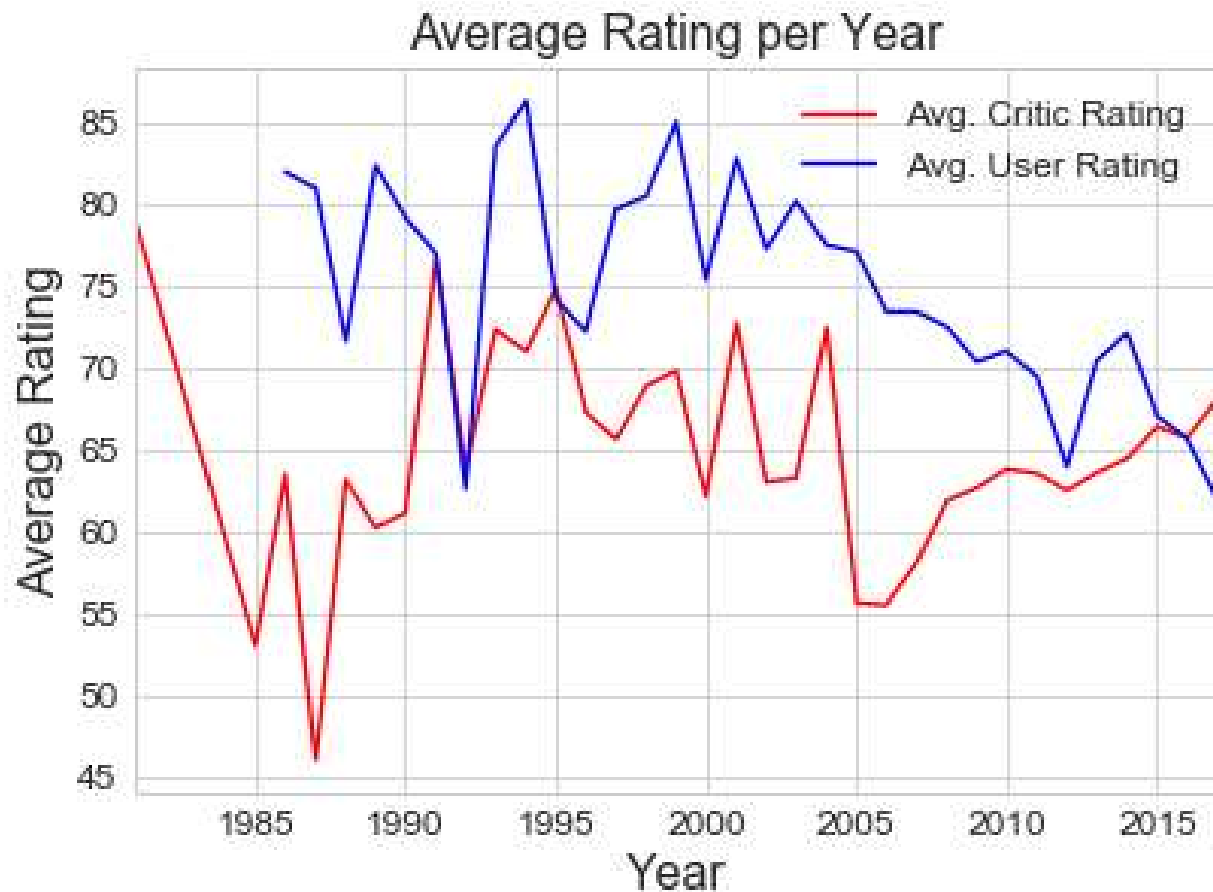
There is a weak (but significant) linear relationship between user ratings and number of reviews.

# Is This the Era of Peak TV?



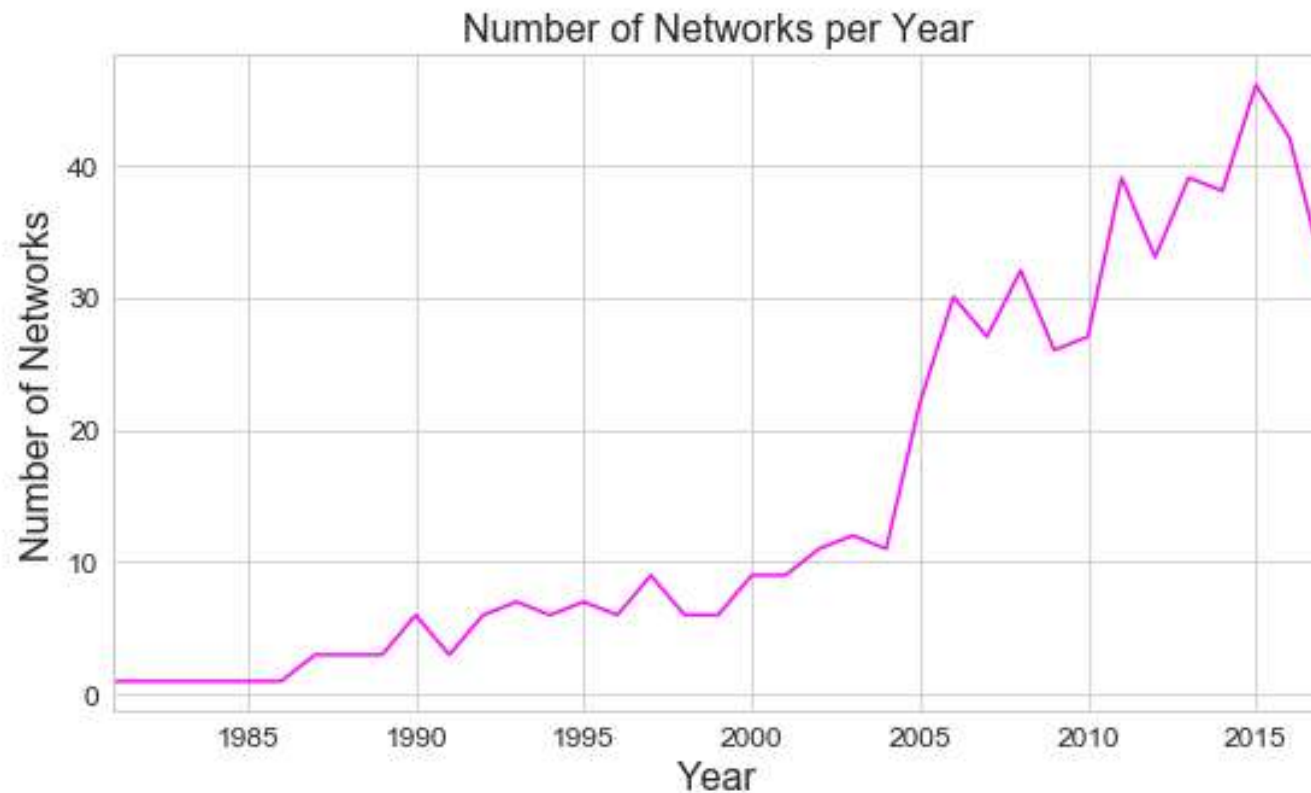
There has been an increase in the total number of shows on the air, as well as an increase in “good” shows (shows with critic ratings > 60).

# Is This the Era of Peak TV?



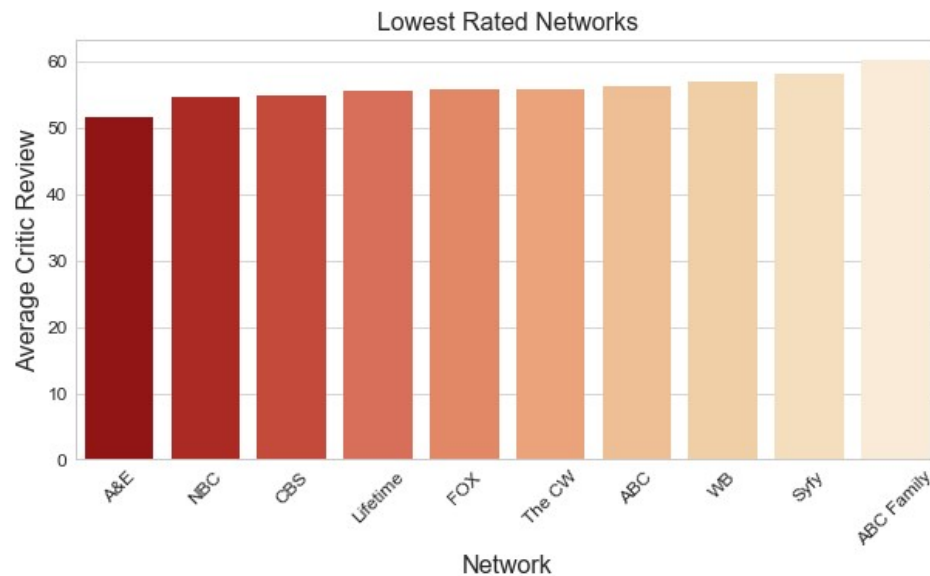
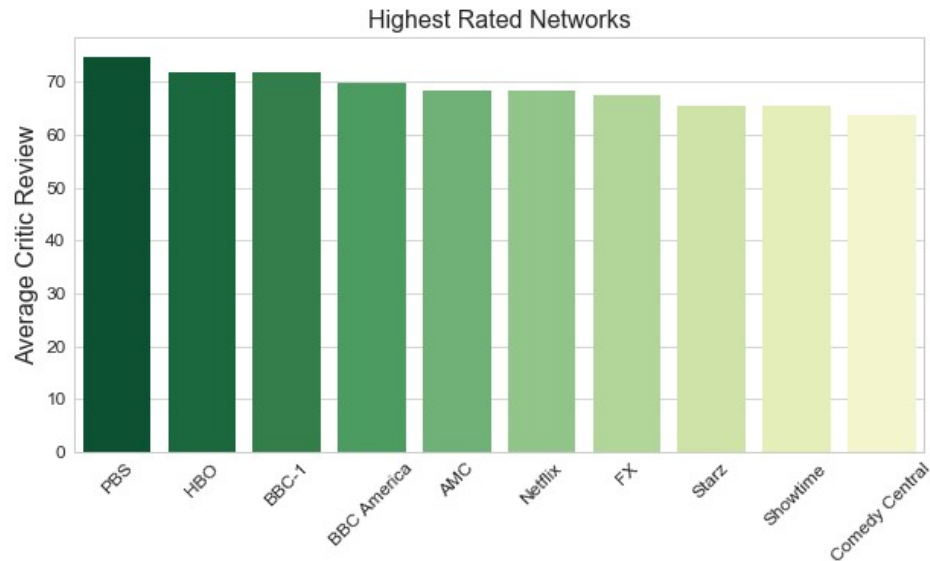
Shows are not getting “better” – average critic and user ratings have not increased.

# Are There More Networks Producing Shows?



There has been an overall increase in the number of networks producing at least one TV show.

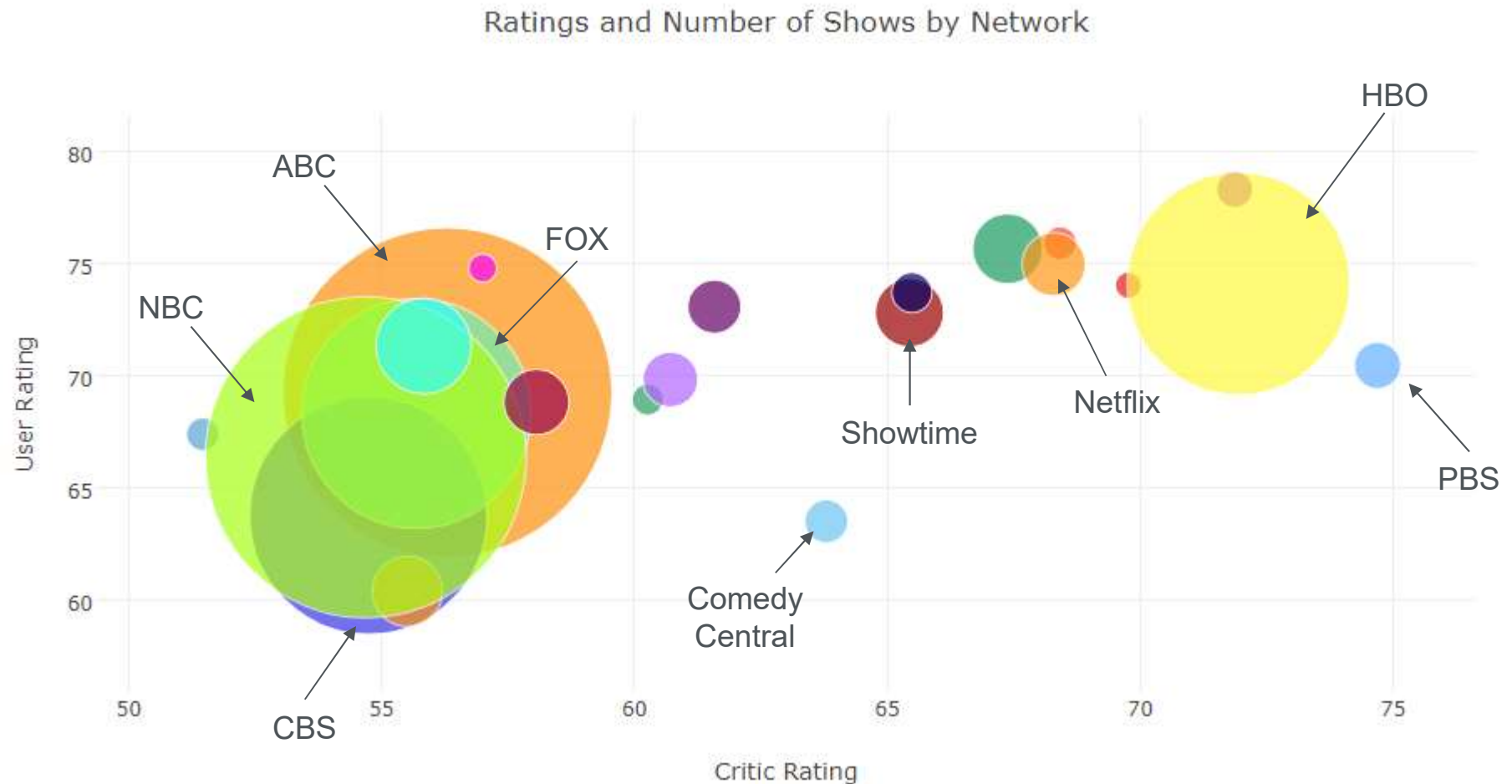
# What are the Highest & Lowest Rated Networks?



Of all networks with more than 15 existing shows, PBS and HBO shows have the highest average critic ratings. A&E and NBC have the worst ratings.

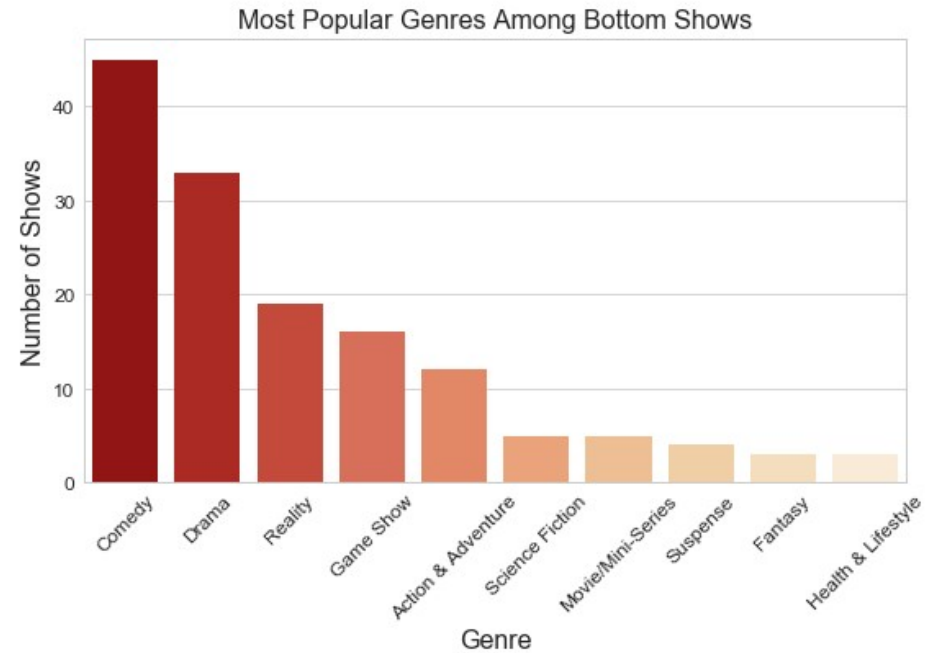
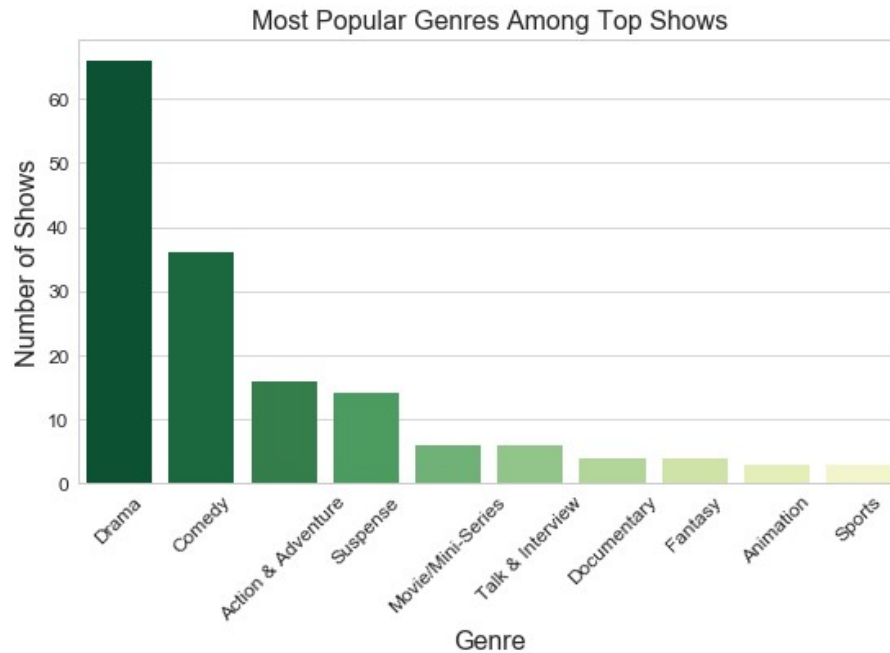


# What are the Highest & Lowest Rated Networks?



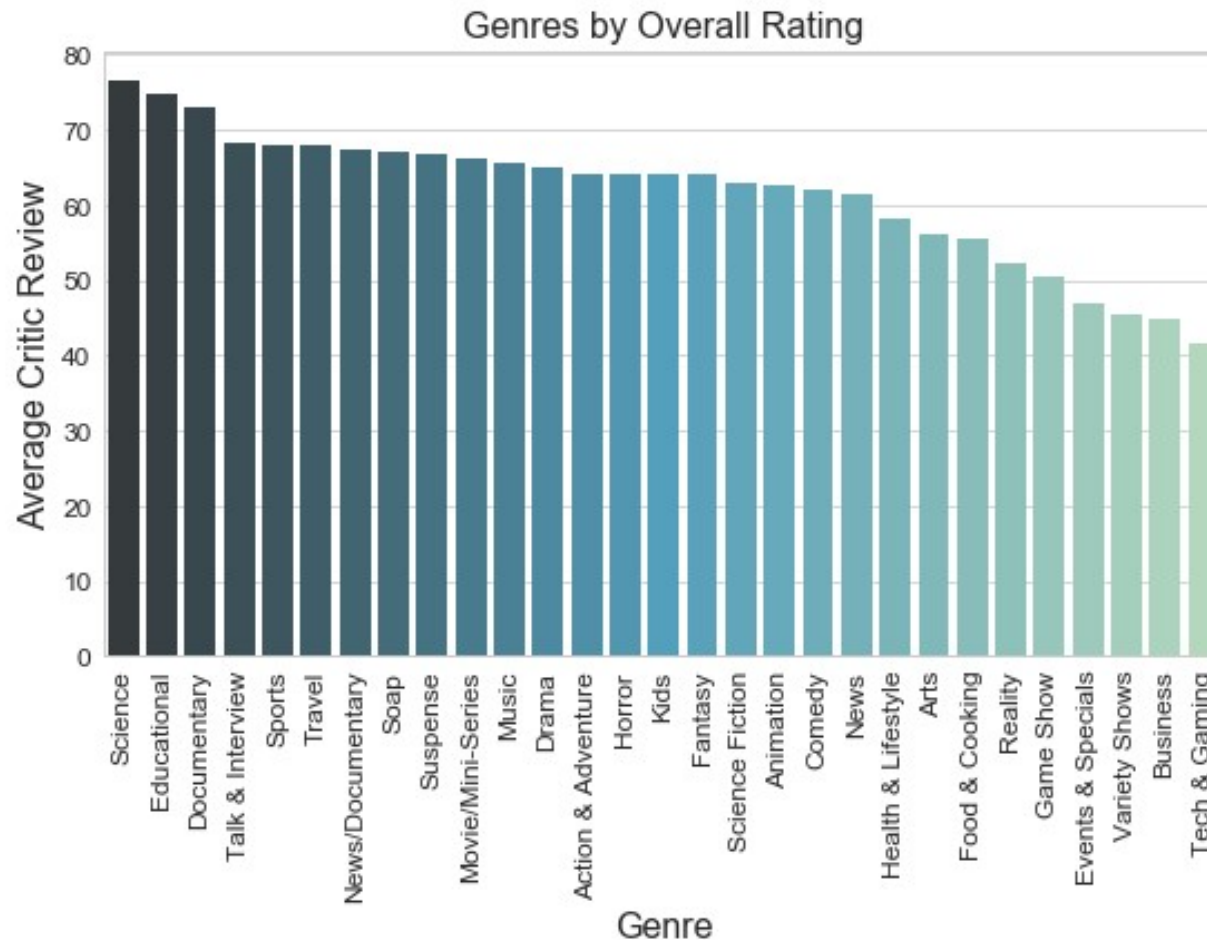
Traditional basic-cable channels are lagging behind in terms of creating “peak” TV.

# How do the Best and Worst Shows Compare?



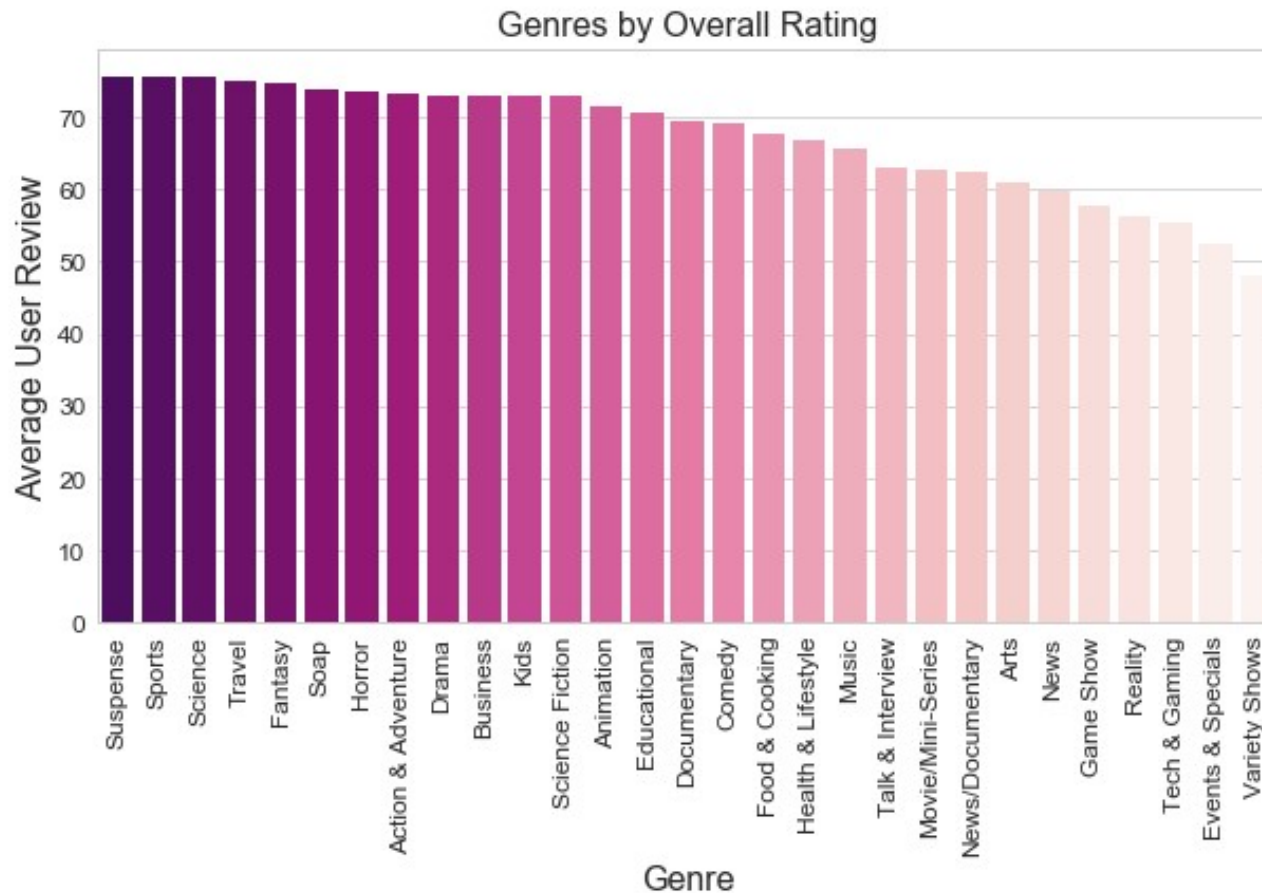
Among the 100 highest rated shows, Drama and Comedy are the most common genres. Among the 100 lowest, Reality shows are very common.

# What Genres to Critics Like and Dislike?



Critics rate science and educational programming highly. The lowest rated genres are variety, business, and gaming shows.

# What Genres to Critics Like and Dislike?



Users prefer sports and suspenseful shows. The lowest rated genres are variety, reality, and gaming shows.

# What Should I Watch Next?

## Best Dramas – All Time

Overall Rank	Show	Average Rating
1.	The Sopranos	93.4
2.	The Wire	92.7
3.	Breaking Bad	91.2
4.	Fargo	90.2
5.	Rectify	90.0
6.	Deadwood	89.0
7.	Sherlock	88.5
8.	Justified	87.6
9.	The Americans	87.4
10.	Game of Thrones	87.0

## Best Comedies – All Time

Overall Rank	Show	Average Rating
1.	The Office (UK)	90.7
2.	Fargo	90.2
3.	Silicon Valley	86.6
4.	Louie	86.6
5.	Arrested Development	86.2
6.	Curb Your Enthusiasm	86.1
7.	Better Call Saul	85.8
8.	Catastrophe	84.3
9.	BoJack Horseman	83.2
10.	Veep	82.6

\* Only includes shows with 3 or more seasons

# Conclusions

- Overall Findings:
  - ① Critic and user ratings are moderately correlated, but critics tend to be pickier.
  - ② TV may not be getting better, but there are more shows to watch (and more good shows).
  - ③ Subscription/digital networks are in a creative race to the top, leaving traditional cable channels behind.
  - ④ Critics love educational programming and hate reality television, and users agree (for the most part).