



Video Media and Attention Span - An Analysis

Luke Gulson



Research Questions

- Is there evidence of people's shortened attention span with the media we consume?
- Do people more often choose short form content over long form content?



Datasets

- First Set: five datasets - over 20,000 shows and movies from Netflix, HBO Max, Disney+, Prime Video, and Hulu.
- Second Set: one dataset - 1500 of trending YouTube videos March 2007 - June 2016



ETL

```
netflix_df = pd.read_csv(r'netflix_titles.csv') #import Netflix movies and shows data to a DF
hbo_max_df = pd.read_csv(r'hbo_max_titles.csv') #import HBO Max movies and shows data to a DF
disney_df = pd.read_csv(r'disney_titles.csv') #import Disney+ movies and shows data to a DF
amazonp_df = pd.read_csv(r'amazonp_titles.csv') #import Amazon Prime movies and shows data to a DF
hulu_df = pd.read_csv(r'hulu_titles.csv') #import Hulu movies and shows data to a DF
youtube_df = pd.read_csv(r'Trending videos on youtube dataset.csv') #import trending YouTube video data to a DF
```

#Creates one cohesive DataFrame by merging each streaming service's DataFrame together.

```
combined_df = netflix_df.merge(hbo_max_df, on = ['id', 'title', 'type', 'description', 'release_year', 'age_certification', 'runtime', 'genre'])
combined_df = combined_df.merge(disney_df, on = ['id', 'title', 'type', 'description', 'release_year', 'age_certification', 'runtime', 'genre'])
combined_df = combined_df.merge(amazonp_df, on = ['id', 'title', 'type', 'description', 'release_year', 'age_certification', 'runtime', 'genre'])
combined_df = combined_df.merge(hulu_df, on = ['id', 'title', 'type', 'description', 'release_year', 'age_certification', 'runtime', 'genre'])
```

#Removes all instances of shows that appear on multiple platforms such that they only appear once and with no duplicates.

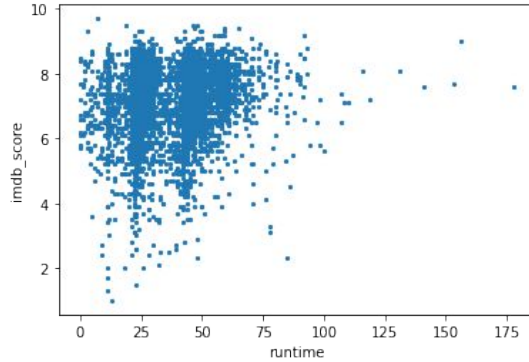
```
combined_df.drop_duplicates()
```

#Makes a DataFrame which only contains observations that are shows, and not movies.

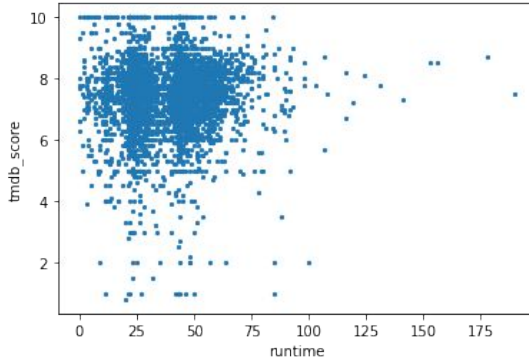
```
all_shows_df = combined_df[combined_df.type == 'SHOW']
```

Show Runtime vs Scores

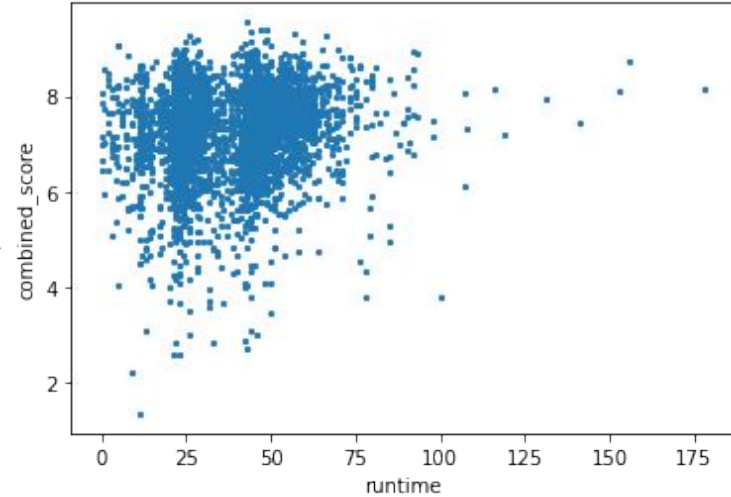
IMDB



tMDB



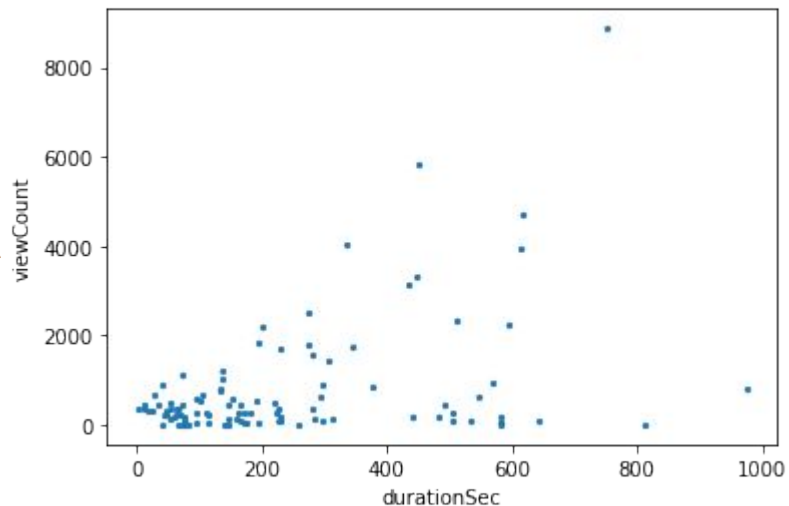
Combined Average



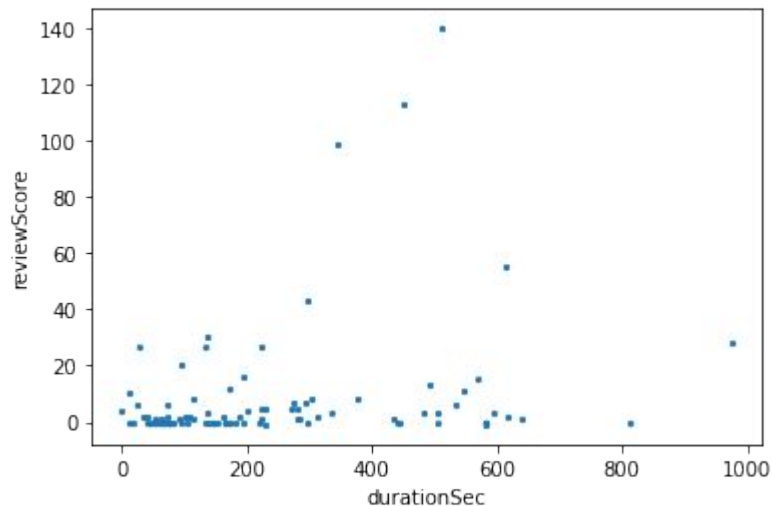
Correlation??

Youtube Video Length vs Reception

Number of Views



Like/Dislike Score



Hypothesis



<null>

- The longer the duration of a piece of visual media, the worse it will be received.
- Null Hypothesis
- Outside variables: Age, Release Date, Content Type, Graphic Content/Age Restrictions, Production Country.

Conclusions

- Won't always get your desired result
- The length of video media doesn't change how we consume it.
- Future Analysis: Other variables, changes over time, other media sources.
- People are Unpredictable





Questions?

Thank you very much!

