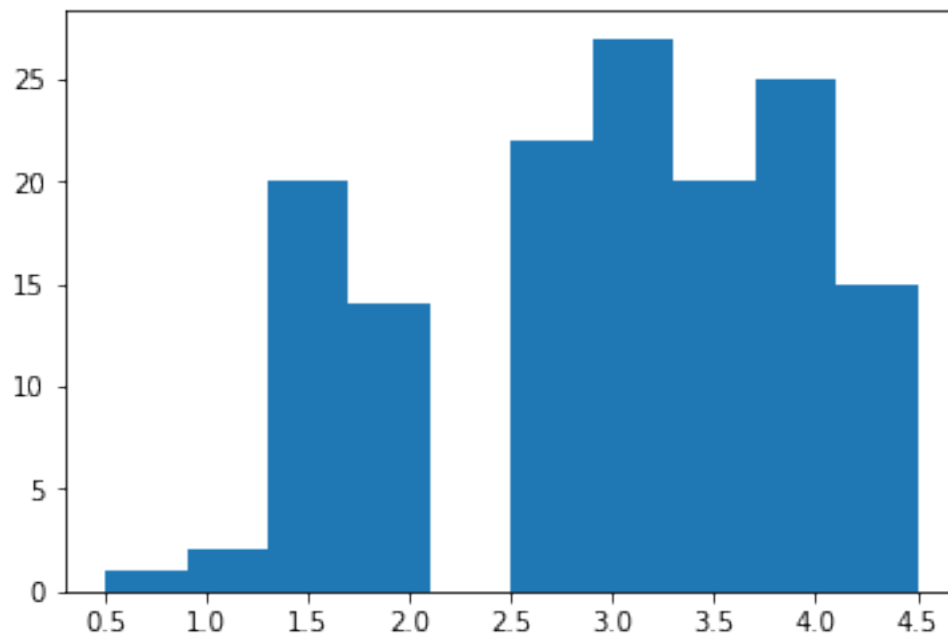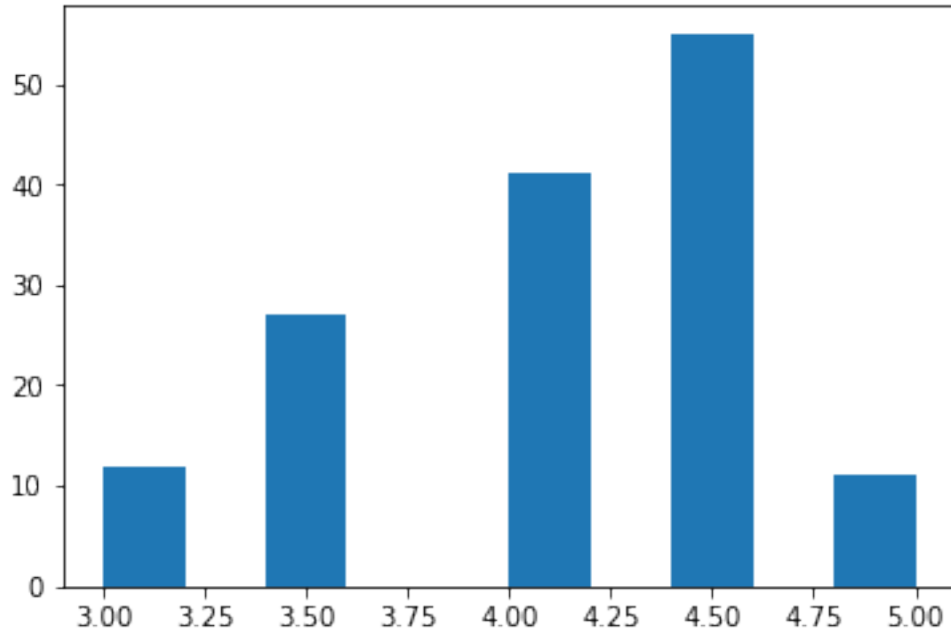# Reviews_Exploration

October 16, 2017

```
In [1]: import pandas as pd
        import matplotlib.pyplot as plt
        import numpy as np
        from scipy.stats import pearsonr, linregress
        file = "fandango_score_comparison.csv"
        movies = pd.read_csv(file)

In [2]: %matplotlib inline
        plt.hist(movies["Metacritic_norm_round"]);
```



```
In [3]: plt.hist(movies["Fandango_Stars"]);
```

# 1 Fandango vs. Metacritic Reviews

Reviewss for Fandango is 3.0 and above with reviews being centered around 4.0 and 4.50 stars. On the other hand, reviews for Metacritic spans from 0.0 stars to 4.5 stars with reviews being centered around 3.0 and 3.5.

```python
In [4]: FS_mean = movies["Fandango_Stars"].mean()
        MC_mean = movies["Metacritic_norm_round"].mean()
        FS_med = movies["Fandango_Stars"].median()
        MC_med = movies["Metacritic_norm_round"].median()
        FS_std = np.std(movies["Fandango_Stars"])
        MC_std = np.std(movies["Metacritic_norm_round"])

        print("Mean of Fandango is:", FS_mean)
        print("Mean of Metacritic is:", MC_mean)
        print("Median of Fandango is:", FS_med)
        print("Median of Metacritic is:", MC_med)
        print("Standard Deviation of Fandango is:", FS_std)
        print("Standard Deviation of Metacritic is:", MC_std)

Mean of Fandango is: 4.089041095890411
Mean of Metacritic is: 2.9726027397260273
Median of Fandango is: 4.0
Median of Metacritic is: 3.0
Standard Deviation of Fandango is: 0.5385321612699531
```

```
Standard Deviation of Metacritic is: 0.9875610297038636
```

## 2 Fandango and Metacritic's Methodology

Metacritic's methodologies were very transparent and they wait until they have four critic's ratings before the aggregate the ratings to get a final score. Fandango, on the other hand, isn't transparent about their aggregation at all. Furthermore, Fandango said the error was due to a software bug when asked about their ratings.

## 3 Fandango and Metacritic's Number Differences

The low median of Metacritic can be attributed to a number of very low movie ratings decrease the median. The high mean of Fandango can be attributed to a number of very high movie ratings increase the mean.

Fandango's low standard deviation is due to their ratings being clustered between 3 and 5 and not as spread out as Metacritic's (whose numbers range from 0.0 to 5.0

In general, Fandango reviews seem to be higher than Metacritic. There does seeem to be some studio influence for Fandango ratings. Although, one can only speculate as their aggregation technique is unknown.

```
In [5]: plt.scatter(movies["Fandango_Stars"], movies["Metacritic_norm_round"]);
```
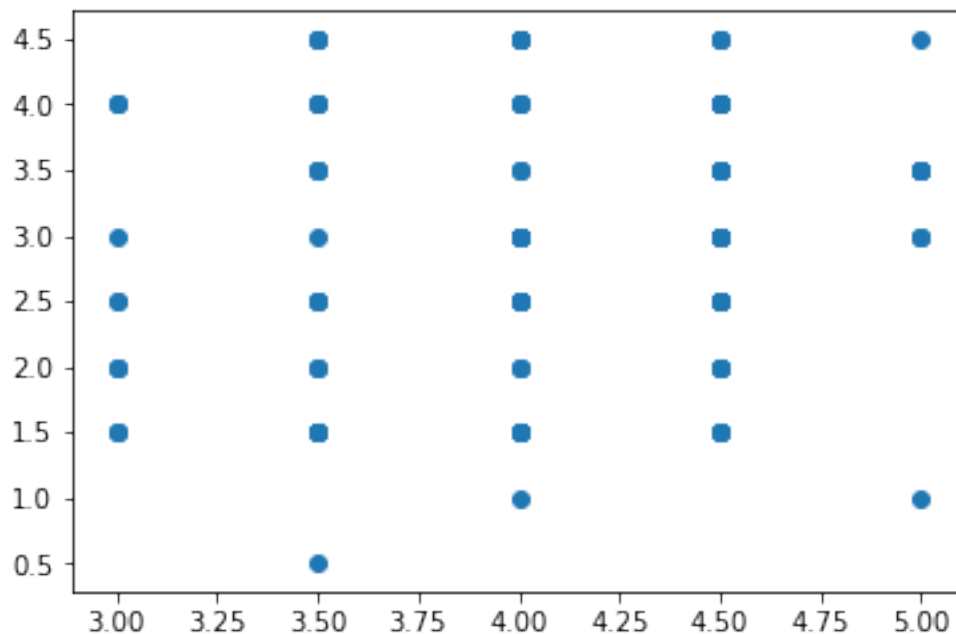


```
In [6]: movies["fm_diff"] = np.abs(movies["Metacritic_norm_round"]
                                  - movies["Fandango_Stars"])
```

3

```
In [7]: movies.sort_values(by="fm_diff", ascending=False).head(5)

Out[7]:                          FILM  RottenTomatoes  RottenTomatoes_User  Metacrit
        3        Do You Believe? (2015)              18                   84
        85             Little Boy (2015)              20                   81
        47                  Annie (2014)              27                   61
        19                 Pixels (2015)              17                   54
        134    The Longest Ride (2015)              31                   73


             Metacritic_User  IMDB  Fandango_Stars  Fandango_Ratingvalue  RT_norm
        3                4.7   5.4             5.0                   4.5     0.90
        85               5.9   7.4             4.5                   4.3     1.00
        47               4.8   5.2             4.5                   4.2     1.35
        19               5.3   5.6             4.5                   4.1     0.85
        134              4.8   7.2             4.5                   4.5     1.55


             RT_user_norm    ...     RT_norm_round  RT_user_norm_round  \
        3            4.20    ...               1.0                 4.0
        85           4.05    ...               1.0                 4.0
        47           3.05    ...               1.5                 3.0
        19           2.70    ...               1.0                 2.5
        134          3.65    ...               1.5                 3.5


             Metacritic_norm_round  Metacritic_user_norm_round  IMDB_norm_round  \
        3                      1.0                         2.5              2.5
        85                     1.5                         3.0              3.5
        47                     1.5                         2.5              2.5
        19                     1.5                         2.5              3.0
        134                    1.5                         2.5              3.5


             Metacritic_user_vote_count  IMDB_user_vote_count  Fandango_votes  \
        3                            31                  3136            1793
        85                           38                  5927             811
        47                          108                 19222            6835
        19                          246                 19521            3886
        134                          49                 25214            2603


             Fandango_Difference  fm_diff
        3                    0.5      4.0
        85                   0.2      3.0
        47                   0.3      3.0
        19                   0.4      3.0
        134                  0.0      3.0

        [5 rows x 23 columns]

In [8]: r_value, p_value = pearsonr(movies["Fandango_Stars"], movies["Metacritic_n

        r_value
```

## 4  Fandango and Metacritic Correlation

The low correlation between Fandango and Metacritic signifies Fandango's ratings aren't just inflated, they're fundamentally different. It also appears that Fandango inflates scores overall, and inflates scores depending on the movie.

```
In [9]: slope, intercept, r_value, p_value, stderr_slope = linregress(
            movies["Metacritic_norm_round"], movies["Fandango_Stars"])
        pred_1 = 1 * slope + intercept
        pred_5 = 5 * slope + intercept
        plt.scatter(movies["Metacritic_norm_round"], movies["Fandango_Stars"])
        plt.plot([1,5],[pred_1,pred_5])
        plt.xlim(1,5)
        plt.show()
```