NextWatch
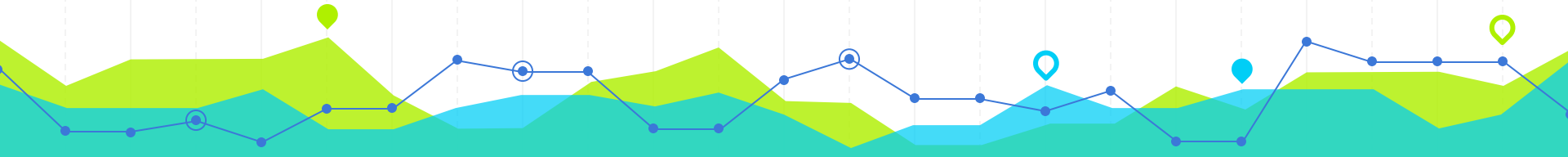
OCRUG Hackathon
Team 6

Bryan Sam, Ellie Park, Hafsa Aasi, Jay Mantuhac, Jing Liao

SUMMARY

◉ We have watched most of the best movies on our streaming services.

◉ Not much movie-making took place since the pandemic began.

◉ Many cinemas faced bankruptcies during the pandemic and are eager to generate revenue again.

◉ **Our key question: What genres of movies should we recommend to customers as they start coming back to the theater?**
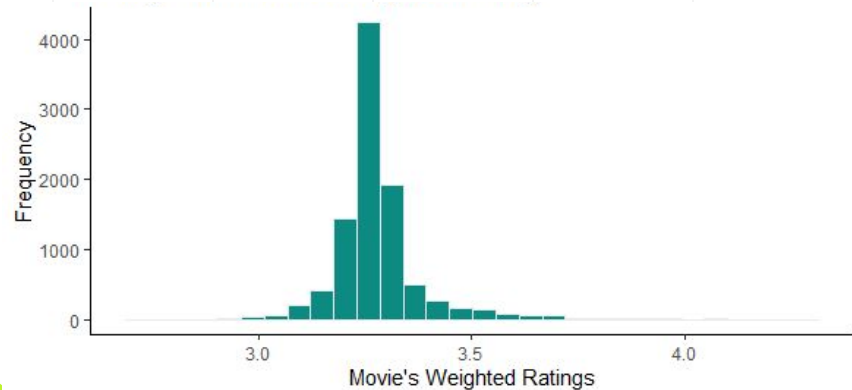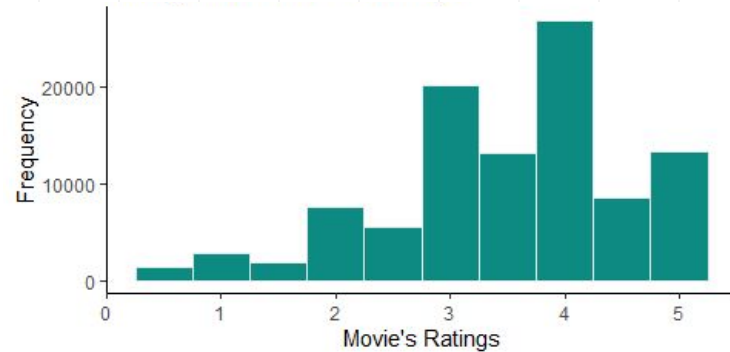
HISTOGRAMS OF UNWEIGHTED AND WEIGHTED MOVIE RATINGS

Weighted Rating (WR) = $(\frac{v}{v+m} \cdot R) + (\frac{m}{v+m} \cdot C)$

where,

- v is the number of votes for the movie
- m is the minimum votes required to be listed in the chart
- R is the average rating of the movie
- C is the mean vote across the whole report

Inspired by IMDB's most guessed algorithm

LINEAR REGRESSION MODEL
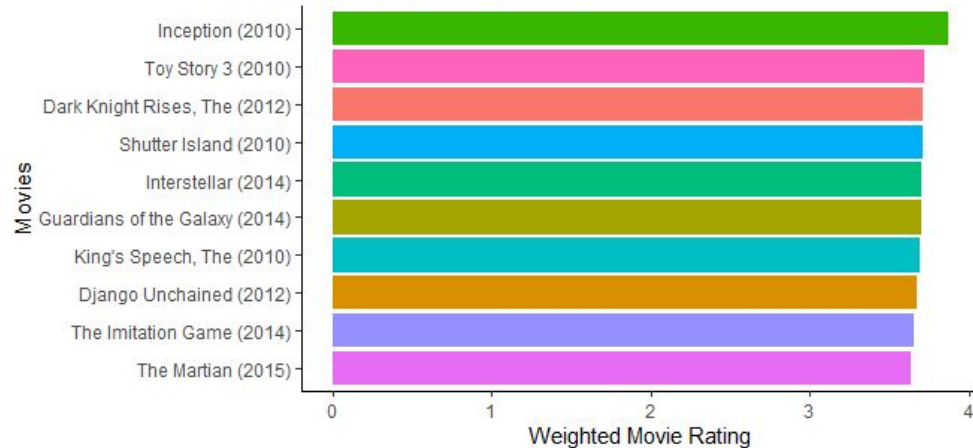
```
Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)        3.337394   0.014871 224.429  < 2e-16 ***
genres_Action      0.017868   0.013984   1.278 0.202307
genres_Adventure   0.023503   0.014353   1.638 0.102542
genres_Animation  -0.014483   0.020198  -0.717 0.473908
genres_Children   -0.000589   0.023858  -0.025 0.980321
genres_Comedy      0.014643   0.013222   1.107 0.268971
genres_Crime       0.020882   0.016949   1.232 0.218873
genres_Documentary -0.003509  0.027116  -0.129 0.897110
genres_Drama       0.033908   0.013494   2.513 0.012495 *
genres_Fantasy     0.019378   0.020782   0.932 0.351829
`genres_Film-Noir` 0.004084   0.067756   0.060 0.951977
genres_Horror     -0.020762   0.026902  -0.772 0.440863
genres_IMAX        0.063135   0.018360   3.439 0.000666 ***
genres_Musical     0.063069   0.050528   1.248 0.212914
genres_Mystery     0.006767   0.022273   0.304 0.761478
genres_Romance    -0.021297   0.019413  -1.097 0.273482
`genres_Sci-Fi`    0.032380   0.015398   2.103 0.036300 *
genres_Thriller    0.018535   0.014584   1.271 0.204747
genres_War         0.009128   0.031148   0.293 0.769696
genres_Western     0.059009   0.039654   1.488 0.137755
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.09336 on 305 degrees of freedom
Multiple R-squared:  0.1502,    Adjusted R-squared:  0.09724
F-statistic: 2.837 on 19 and 305 DF,  p-value: 9.361e-05
```


Top 10 Movies After 2010

INSIGHTS AND RECOMMENDATION

◉ According to our model, IMAX is an important feature in highly rated films.

◉ Can higher rating potential be associated with an increase in revenue?

◉ Movie theaters can do more research on whether customers prefer IMAX movies or IMAX tends to select films based on their potential for popularity