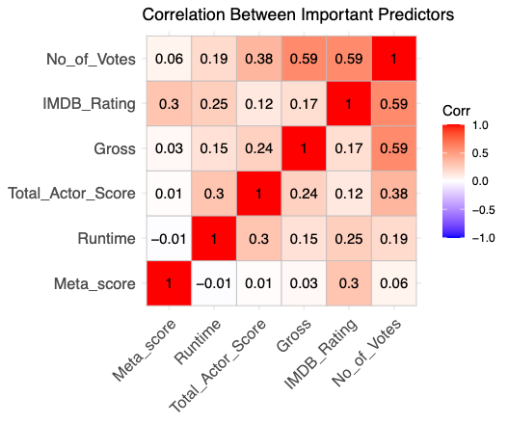


0.2.2 Models

Based on the correlation plot, we will be evaluating three models. (add any transformations – log?) that have been made. Do we further explain the actor score?

0.2.2.1 Model One For the first model we are examining only the primary beta without any covariates, the Gross variable:

$$Metascore = \beta_0 + \beta_1(Gross)$$

```
model_one <- lm(Meta_score ~ Gross, data=movies)
model_one
```

```
##
## Call:
## lm(formula = Meta_score ~ Gross, data = movies)
##
## Coefficients:
## (Intercept)        Gross
##   7.771e+01   -3.372e-09
```

From this model we can see that for a score of 77.71, the film's Gross will not have a perceived impact with a change of -0.000000003372 for every point increase.

0.2.2.2 Model Two For the second model we are examining the primary beta (Gross) with the covariate of IMDB_Rating:

$$Metascore = \beta_0 + \beta_1(Gross) + \beta_2(IMDBRating)$$

```
model_two <- lm(Meta_score ~ Gross + IMDB_Rating, data=movies)
model_two
```

```
##
## Call:
## lm(formula = Meta_score ~ Gross + IMDB_Rating, data = movies)
##
## Coefficients:
## (Intercept)        Gross  IMDB_Rating
##  -2.189e+01   -7.479e-09    1.259e+01
```

Explain the results - HERE

0.2.2.3 Model three

For the second model we are examining the primary beta (Gross) with the covariates of IMDB_Rating and Total_Actor_Rating:

$$Metascore = \beta_0 + \beta_1(Gross) + \beta_2(IMDBRating) + \beta_3(TotalActorScore)$$

```
model_three <- lm(Meta_score ~ Gross + IMDB_Rating + Total_Actor_Score, data = movies)
model_three
```

```
##
## Call:
## lm(formula = Meta_score ~ Gross + IMDB_Rating + Total_Actor_Score,
##     data = movies)
##
## Coefficients:
##       (Intercept)              Gross        IMDB_Rating  Total_Actor_Score
##        -2.402e+01         -5.531e-09           1.302e+01          -6.624e-02
```

Explain the results - HERE

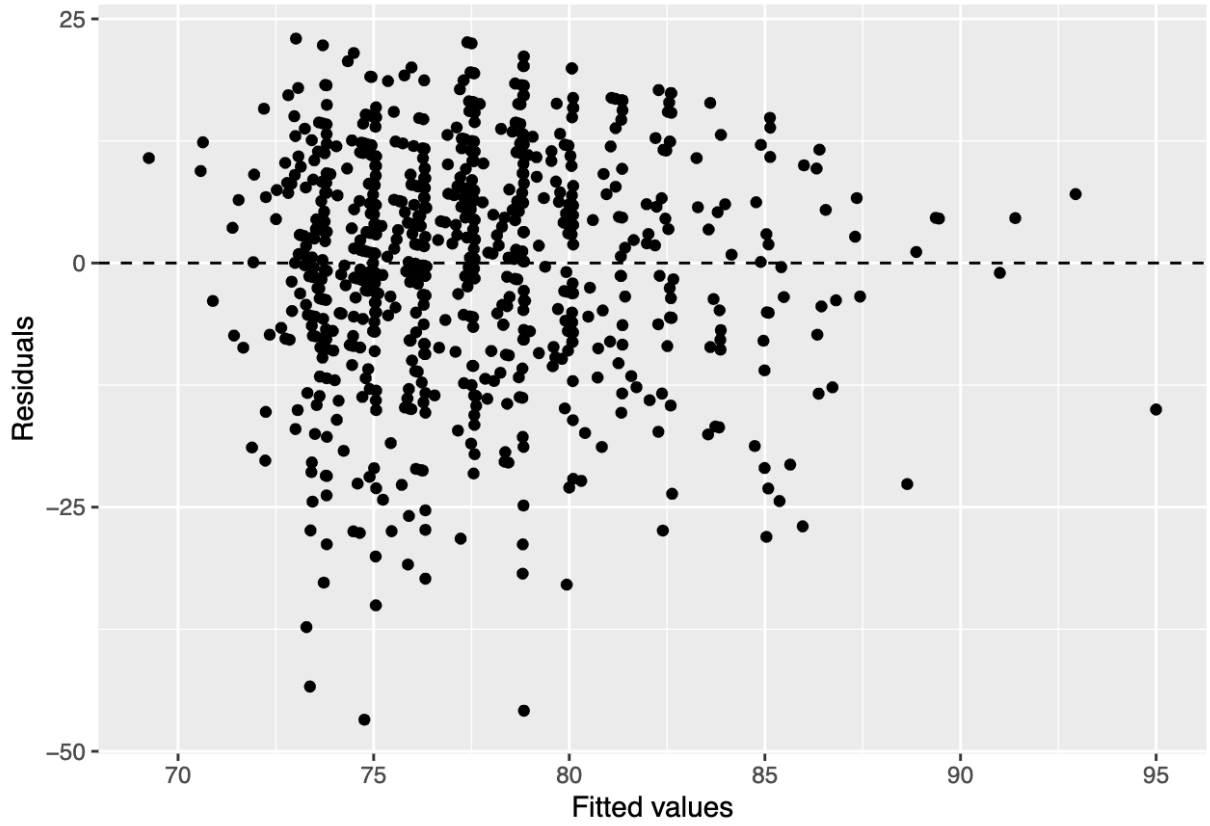
```
##           1           2           3           4           5           6
##  2.38282439 22.74237436  8.09085029 12.48047546 18.30195648 17.56138831
##           7           8           9          10          11          12
## 16.65120071 16.61400663 -2.72614968 -11.58787667 15.35130258  5.40089751
##          13          14          15          16          17          18
## 12.30782394 10.44237159 -4.13449840 12.44519113  5.26676510  5.66492933
##          20          22          23          24          25          26
## 18.46721583 -3.07872357  1.31275866 18.32116349 14.01745173 -16.25143796
##          27          28          29          30          32          34
## -18.51851882 -12.37511219  7.72813234 13.37556616 20.28816143 -17.58157050
##          35          36          37          38          39          40
## 10.33140166 -20.66829390 -11.53372145  7.73366728  7.39709211 -10.07978440
##          41          42          43          44          45          46
## -15.69008582 -0.63403577 -13.64698575 11.71292331 -2.02199361  2.32768696
##          48          49          50          51          52          53
##  9.99745131  2.30519878 19.39516133 22.29070904 18.28780460 21.28731881
##          54          57          59          60          61          62
## -2.70714449  1.30417278  9.92876685  3.18177373 -7.42371218  3.99447122
##          63          64          65          67          68          69
##  3.83625033  1.79842399 -10.69071625 18.04195764 11.32531202 -0.71036018
##          70          71          73          74          75          76
##  2.37339397 -1.70473610  8.12407437 -11.56431489 16.56872841 11.55331282
##          79          82          85          89          90          91
## 19.28818450 22.41122707  0.82418567 -0.71042862 17.31119095  2.31037694
##          94          95          96          97          98          99
## -8.30627051 11.40325538 -8.60070617 -22.61047630 -9.70048667  6.72595288
##         100         101         102         103         104         105
## -7.24593449 -0.70959656 17.93401021 -9.45781511  1.29680401 -1.55642315
##         107         108         109         110         111         112
##  6.57442300 10.46251238 -12.55898155 -18.67034408  8.32599172 16.38255835
##         113         114         115         116         117         118
##  5.82544129 -0.69181280  6.47931206 -3.66216437 22.43840562 16.34997519
##         119         120         121         124         125         127
## 20.33201875 22.29804485 21.31699277 17.30654254 22.29260104 20.29142260
##         132         133         139         144         145         146
## -8.42584703 10.47107991 -4.71225860  2.30880669 -6.66669272 -14.28107346
```

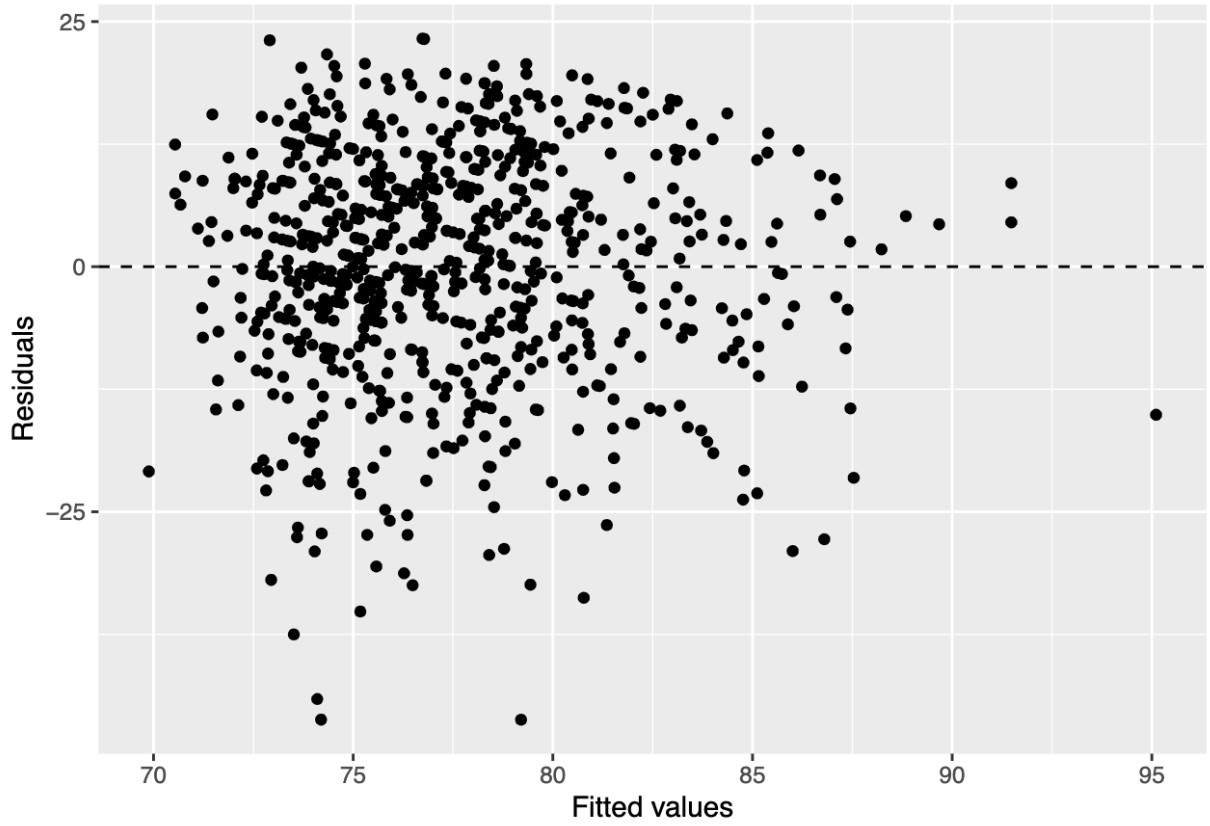
```
##          809          810          811          812          814          816
##  -4.69903790  -1.57142196   9.40801412 -30.52351649   5.32104023   3.36093091
##          817          818          819          820          821          822
##  -7.70211995   1.30786096 -11.45200019 -13.62074990  -4.55031213  -1.45813258
##          824          825          826          827          828          829
##   4.32342067 -15.23613266 -13.43482245  -8.69199418 -11.69561423   5.81482717
##          831          832          833          834          835          836
##   6.29246365  -1.56442665 -15.63775477 -23.68862360  -1.68391028  -2.67689596
##          837          839          840          841          842          843
##  12.32299839  -0.53065247 -10.44664347 -16.55353924   8.47204521   0.56848727
##          845          846          849          850          851          852
##   9.44574283  -6.70625548  -4.30977992   9.38774271   5.37155663   2.31104771
##          853          854          856          857          859          862
##  16.33996244  12.40831252   8.43740441  -4.55998974  -5.61714684   9.45950458
##          863          866          867          874          876          879
##  12.32570797  -3.69620572  12.28743187  13.28906197  18.54490987  -6.62500060
##          880          881          882          884          885          887
##   5.76778758   8.28916394  14.30716424 -17.62911371  -7.66849501   5.29010512
##          888          889          890          891          892          893
##  -9.39825711   8.65085372  -5.65090139   3.53110334   4.33945263   4.12608776
##          894          895          896          897          898          899
##   4.44536951   4.65740139  14.29093916  10.37783699  -0.58559167   1.99049071
##          901          902          903          904          905          906
##  -4.69890299  -9.57447841   2.35200770  11.01322971   2.86470267 -12.70934647
##          907          908          909          911          912          913
##  -4.86461556  -5.59467452 -11.55064463   9.54220341  -7.70200300  -4.45784787
##          915          916          917          918          919          920
## -20.00788026 -23.84962492   1.31902748 -41.47686133   4.34496722 -11.58246500
##          921          922          923          924          925          926
##   3.28956868  -1.41426952  -5.64429151 -11.67800749  -7.62967847  -9.57830483
##          927          928          929          930          931          932
## -25.71272560   1.30549074 -25.00253124  -5.63488433 -21.35021208 -15.63131090
##          933          934          935          936          937          938
## -31.52390703 -10.70235331  -4.71039491 -32.71089558  -1.69194738   3.30660788
##          939          940          941          942          943          944
##  -7.65781276 -21.71259315 -10.53847368  -9.66870341 -47.51737081  -4.56078247
##          946          947          948          949          950          951
##  -1.53616920  10.33319000 -12.64184936  -3.38726172 -25.53405568 -24.53941401
##          952          953          954          955          956          957
##   2.73261724  -3.54188526 -13.66192737  -0.68823885   0.41000022  -6.30600231
##          958          960          961          962          963          964
## -36.67673092 -11.66426070 -28.54717576 -25.69994303   6.43287077 -19.37549372
##          965          966          967          968          969          970
## -15.70924615  -8.47158469  -0.12654698  10.29219299 -21.57481487   0.31420745
##          971          972          973          974          975          976
##  -9.53426343  12.28878251 -11.70669570 -13.74913038 -17.48794125  -1.39973541
##          977          978          979          980          981          982
##  10.66338989  -1.44717217  -5.54572482  -9.49286092   4.29450414  -9.31050347
##          983          984          985          986          988          989
##  -0.67071140 -12.63690721  -3.45425080  -1.56774557 -18.59472237  12.73267021
##          990          991          992          993          995          998
##   9.29048797  -0.71039657 -27.70809765 -12.23443503  18.33372175   7.39010318

##             1             2             3             4             5
```

```
## -14.998659739   7.057915501  -3.433097415  -1.004603424   4.599475342
##             6             7             8             9            10
##   6.651817896   4.633194395   4.550704991 -12.726725249 -22.637872289
##            11            12            13            14            15
##   5.445049187  -4.444958659   1.130811426   1.906184567 -13.373203675
##            16            17            18            20            22
##   2.694627407  -3.483277317  -3.818034942  10.002635613 -10.990344093
##            23            24            25            26            27
##  -7.339921115   9.678719207   6.222954486 -24.373391851 -26.965726564
##            28            29            30            32            34
## -20.647677847  -0.418701065   6.017184378  11.605526892 -23.628591270
##            35            36            37            38            39
##   2.960587121 -28.038737697 -18.740281517   0.852735978   0.106276091
##            40            41            42            43            44
## -16.733535218 -23.087068035  -7.962759657 -20.991480250   6.024540155
##            45            46            48            49            50
##  -8.605366204  -5.047651384   3.437758937  -5.097525890  12.101993971
##            51            52            53            54            57
##  14.870338618  10.863897120  13.862819712  -8.865739342  -4.840639776
##            59            60            61            62            63
##   4.544591427   0.541333450 -10.801519477  -1.309688731  -1.660592666
##            64            65            67            68            69
##  -2.526673955 -16.829304627  12.795627132   5.206243057  -6.872871122
##            70            71            73            74            75
##  -3.687120300  -7.860397980   2.977746464 -17.548970362  10.746094440
##            76            79            82            85            89
##   5.711905579  13.123901237  16.396786328  -3.428188215  -5.613861318
##            90            91            94            95            96
##  12.434086785  -2.567718538 -12.717515284   6.638268234 -13.370517732
##            97            98            99           100           101
## -27.392186029 -14.591811953   2.353950084 -11.583701416  -5.612015969
##           102           103           104           105           107
##  13.815381788 -14.053612468  -3.597820719  -6.272306404   2.017885552
##           108           109           110           111           112
##   5.769689015 -17.277980432 -22.307151041   3.466912081  11.592366148
##           113           114           115           116           117
##   1.574596516  -5.572574974   1.806947514  -8.506820377  17.716224803
##           118           119           120           121           124
##  11.520102865  15.480278902  17.404931218  16.446954121  12.423777487
##           125           127           132           133           139
##  17.392857894  15.390244335 -11.723551726   7.047851767  -8.358758285
##           144           145           146           147           148
##  -1.312039467 -10.257701791 -17.402471400   8.831444944  -5.485576269
##           149           150           151           152           153
## -13.351511146  11.940972284  16.921618220  13.743849061 -18.832509609
##           156           158           159           160           162
##  -9.812856151   0.681361377  -1.324606190  -8.082911908 -15.330690296
##           163           165           166           168           169
##   9.123410351  -4.855500689  -8.042456565   4.396685050 -14.885254439
##           171           172           173           174           179
##   4.648429196  -8.739042455  14.671090983   7.815042357   9.649353808
##           180           184           187           188           190
##   4.730653047  16.827127998  15.640589013  -6.359745014  16.640883240
##           191           192           196           200           201
```

```
##  16.640237149  16.677659132  14.927441369   3.914330752  -6.074556186
##           203           204           205           207           210
##  -1.408423954   6.009092234  -3.077440577   7.900560501   8.341314901
##           211           213           214           216           217
##   0.153999479  -8.598038275  16.565185121  16.323154239  -5.899144757
##           218           219           220           223           224
##   1.778994905  13.236280288   4.905735777  -9.644460773  11.051475407
##           226           227           228           231           232
##  -6.993120237   7.748782996 -12.089472739  -3.473457322  -6.963411044
##           233           235           236           237           238
##  11.454865970   6.650871496  -0.924697396 -16.092368102   9.942853392
##           240           242           243           244           245
##   1.899428312 -10.576429158  12.747526892  -3.869572103   2.939586593
##           246           248           249           250           251
##   1.067513065   3.899847237 -13.905638210 -22.098739442  10.838780754
##           252           253           254           255           256
##  -7.098069854  -8.969809443   4.919769666   3.022734173   5.083387676
##           260           261           262           263           264
##  -3.059279856  19.929564057  -3.096188667  -9.090854224   4.086367443
##           267           268           269           270           271
##  -0.383770393  -4.709636617  12.935348171  -2.084358991   5.903030335
##           272           273           275           276           278
## -22.997599957 -32.934320100  19.936501683   4.145138884  -2.950766402
##           279           280           284           285           289
##   6.265623272  -9.223912663  12.116204303  -2.869333855  12.020610139
##           292           294           301           305           306
##  15.899740748  -5.070388722  10.457979905   7.235175119  10.971117943
##           308           310           312           313           314
##  -8.070987817  16.902681985  15.900843046   5.931929668  -7.028882057
##           315           317           328           329           330
##  18.385159968   6.931929668  16.288526009  -9.454572735   2.866866118
##           331           333           334           335           337
##   1.710719689  10.169754403 -11.851168180 -13.799885460   3.166040783
##           338           339           340           341           342
##  -7.821878242  -5.160016103  -0.349795144   2.846927671  11.349700948
##           343           344           345           346           347
## -28.222920060  -1.468147031 -11.708854718  -7.840767020  10.196505703
##           348           349           350           353           355
## -15.572323390 -11.126226717   8.159281913 -28.811461250   6.215370512
##           356           357           358           359           360
##   0.958363860 -45.841455500  -5.180208506  11.191732601   1.362639316
##           361           362           363           364           365
## -13.618254838 -14.412491955   7.859667195  -6.839732061  -4.287317142
##           366           368           369           370           371
##  13.203282066  18.702424871   2.410758703   4.653633929   3.162238728
##           372           373           374           375           376
##   6.176289370  -4.825912675  -9.380455922 -24.828148553  -3.840246953
##           377           378           379           381           383
## -13.557423313 -20.346000970  13.113718064 -18.830070513   9.169552711
##           384           385           386           388           389
##  -1.673574756 -17.810224718   7.204874793   3.170807439   1.200330629
##           390           392           393           394           395
##   6.331685216   5.341527987  12.258827834  -4.414173184  -2.837657490
##           396           397           398           399           400
```


25
0
−25
−50
Residuals
70
75
80
85
90
95
Fitted values




```
#we are going to need to a transform. Thoughts on what to use?
model_one$coefficients
```

```
##   (Intercept)         Gross
##  7.771275e+01 -3.372100e-09
```

```
model_two$coefficients
```

```
##   (Intercept)         Gross   IMDB_Rating
## -2.189141e+01 -7.478678e-09  1.259162e+01
```

```
model_three$coefficients
```

```
##       (Intercept)             Gross       IMDB_Rating Total_Actor_Score
##     -2.402407e+01     -5.530582e-09      1.301950e+01     -6.623691e-02
```

Assumptions:

NOTES: Add LM models, coeffs and risd

0.3 Results - (Sean)

```
stargazer(model_one, model_two, model_three, type="latex", title="Movie OLS Model Results")
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Mon, Apr 11, 2022 - 01:17:16 PM

You should display all of your model specifications in a regression table, using a package like to format your output. It should be easy for the reader to find the coefficients that represent key effects near the

Table 2: Movie OLS Model Results

	Dependent variable:		
	Meta_score		
	(1)	(2)	(3)
Gross	−0.000	−0.000*	−0.000
	(0.000)	(0.000)	(0.000)
IMDB_Rating		12.592***	13.019***
		(1.525)	(1.541)
Total_Actor_Score			−0.066*
			(0.036)
Constant	77.713***	−21.891*	−24.024**
	(0.547)	(12.079)	(12.116)
Observations	750	750	750
R^2	0.001	0.084	0.089
Adjusted R^2	−0.0004	0.082	0.085
Residual Std. Error	12.496 (df = 748)	11.971 (df = 747)	11.952 (df = 746)
F Statistic	0.700 (df = 1; 748)	34.447*** (df = 2; 747)	24.144*** (df = 3; 746)
Note:			*p<0.1; **p<0.05; ***p<0.01

top of the regression table, and scan horizontally to see how they change from specification to specification. Make sure that you display the most appropriate standard errors in your table.

In your text, comment on both statistical significance and practical significance. *You may want to include statistical tests besides the standard t-tests for regression coefficients. Here, it is important that you make clear to your audience the practical significance of any model results. How should the product change as a result of what you have discovered? Are there limits to how much change you are proposing? What are the most important results that you have discovered, and what are the least important?*

0.4 Model Limitations - Kate & Mick

Statistical limitations of Model: *As a team, evaluate all of the large sample model assumptions. However, you do not necessarily want to discuss every assumption in your report. Instead, highlight any assumption that might pose significant problems for your analysis. For any violations that you identify, describe the statistical consequences. If you are able to identify any strategies to mitigate the consequences, explain these strategies. Note that you may need to change your model specifications in response to violations of the large sample model.*

A statistical concern of our model is the independence of the data. The movies in our data occur over time. Movie that follow preceding movies are very likely influenced by their predecessors. In addition the production of high-grossing and popular movies tend to come from the movie industry in Hollywood. This insular production phenomenon can result in competitive and creative interactions. These potential violations of large sample assumption of independence must be recognized as potential limitations in our model's validity.

Structural limitations of Model: *What are the most important* omitted variables *that you were not able to measure and include in your analysis? For each variable you name, you should* reason about the direction of bias *caused by omitting this variable and whether the omission of this variable calls into question the core results you are reporting. What data could you collect that would resolve any omitted variables bias?*

The relationship between variables explaining human behavior is complex and interrelated. The overall outcome is to generate a predictor of human behavior, a predictor of what movies will generate the most views on the new streaming platform. The models include the variables of the film's total gross, IMDB rating, and total actor score. Many other factors may contribute to the number of views on a streaming platform. Some of these factors were omitted by choice, while others were not available at the analysis time.

Intentionally Omitted Variables: The movie title and the year the movie was released were both intentionally omitted from the models. This data had been collected and was available at the time of analysis but was omitted as it was determined it had little impact on the viewability of the movie on the new streaming platform.

Further Examination Two factors that could impact Meta scores were runtime and genre of movie. Runtime ranged from 45 minutes to 321 minutes with an expected value of 123 minutes. The amount of time a movie takes to run may influence its attractiveness to viewers. Further examination of this variable is suggested.

The genre of the movie may also influence its attractiveness to viewers. Popular genres of the times could guide more viewers to stream a movie. Many movies are classified into multiple genres, such as "Action, Adventure" or "Drama, Action, Biography, Crime". The double or triple classification can make selecting a genre that will draw viewers to the streaming platform challenging and unpredictable. While further examination of the relationship between genre and the number of viewers streaming the movie is suggested, the current models omitted genre to maximize the models' reliable predictive power.

Unintentionally Omitted Variables: Factors not available at the time of analysis include, but are not limited to;

Total Movie Budgets Number and Types of Awards and Honors the Movie received Number and Locations of Countries where the movie was released Genre Fan Base Size Pre-existing Fan Base Size (from books or other media types) Number of languages the move was translated into Adjusted Gross (movies over five years old) The number of times the film has been released

Omitted variables, both intentionally and unintentionally omitted variables, can potentially impact and influence the bias of the models. These factors may change the model's overall fit to reality and sway the model's predictive power. Three unintentionally omitted factors that were not available at the time of analysis that could impact the fit of the models are total movie budget, number and types of awards and honors the movie received, and number and locations of countries where the film was released.

Total Movie Budget A movie may have a large budget for many reasons, high profile actors or directors, large casts, intricate sets and costumes, complicated stunts, large advertising campaigns, or complications with production. Some of these factors may positively influence the outcome of viewership, but others may negatively affect the streaming of a movie. Therefore, the influence of budget on the current coefficients is unknown and needs further exploration. The 1000 movies in the IMDB dataset would need to be coordinated with other datasets to fill the budget amounts for every 1000 movies on the list. Compiling the data may take multiple datasets since the range of film spans numerous decades, different movie studios, and a wide range of genres.

Number and Types of Awards and Honors the Movie Received From Oscars to BAFTA, a movie can gain many awards and honors. Awards and honors may increase the public's desire to watch the film. Awards can bring visibility to a movie that did not have a large advertising budget, showcase a new rising actor or celebrate the life of a well-known one, and get a movie to a country or audience that has never been exposed to the film. Due to all of these interactions, the number and types of awards may positively affect the streaming viewership. This positive relationship may alter the coefficients and result in a better fit model to reality. The number and type of awards may have a positive bias and pull the model away from zero.

Number and Locations of Countries where the Movie was Released Movies are released all over the world to a variety of audiences. Often, movies are even dubbed into a wide range of languages to reach a greater audience better. Films with the budget and the audience base to be released in more countries are more likely to have a larger audience. Therefore, the effect of the number of countries the movie is released in, and viewership of a movie on the new streaming platform is predicted as a positive relationship. Further

exploration should be completed to fully realize the impact of the number of countries the film is released in and the rating. If the relationship is positive, the bias will draw the model coefficients away from the zero of reality. Multiple data sources will need to be mined to generate information about each of the 1000 movies in the IMDB dataset.

0.5 Conclusion

Make sure that you end your report with a discussion that distills key insights from your estimates and addresses your research question.