Video Streaming Platform: Which Has Better Shows?

Alison Yao (yy2564)

Television streaming has been disruptive to traditional television in recent years. Due to COVID-19, video streaming platforms experienced a boost in subscribers when people were quarantining at home. As a TV show addict, I have been watching shows from Netflix for years and also other newer platforms in recent years. Therefore, after seeing the TV viewership data in class, it is of interest for me to understand more about the relationship between television streaming platforms and genres and the ratings.

I collected a sample of 179 observations of various types of TV shows on various streaming platforms. The response is the Average Audience Score (**Rating(%)**), which is not technically a score but the percentage of users who rated a show 3.5 stars or higher. The two categorical predictors are **Genre** (Action & Adventure, Animation, Comedy, Drama, Horror, Kids & Family, Musical & Performing Arts, Mystery & Suspense, Science Fiction & Fantasy) and **Platform** (Amazon Prime Video, Apple TV+, Disney+, HBO Max, Hulu, Netflix). The first few observations are listed below:

Title	Platform	Genre	Rating(%)
Lore	Amazon	Mystery & Suspense	43
The Pale Horse	Amazon	Mystery & Suspense	46
Monsterland	Hulu	Horror	46
Muppets Now	Disney+	Animation	49
Genera+ion	HBO Max	Comedy	52

All data are collected from Rotten Tomatoes. They provide lists of the best TV shows and movies on different streaming platforms. For example, the list of Apple TV+ shows is here: https://editorial.rottentomatoes.com/guide/apple-tv-plus-shows-and-movies-ranked/. The list consists of both movies and TV shows, so I picked out the TV shows and documented the title, platform, genre and average audience score. Please note that the list is sorted by Average Tomatometer (critic score) over 60%, but I prefer audience feedback, so I am using Average Audience Score as the response.

Let's first look at the counts of TV shows in each combination of categories. Netflix, Hulu and Amazon Prime Video have more than ten years of history while HBO Max, Apple TV+ and Disney+ were launched only a couple of months ago. Therefore, there are plenty of combinations of categories with 0 count. Also, some genres have fewer TV shows in general.

Musical & Performing Arts is a Disney+ special, while Kid's Family only has 3 counts in total. There are also fewer Horror and Mystery & Suspense shows. For example, I don't think parents would appreciate it if Disney+ have Horror shows for their children.

Rows: Genre Columns: Platform

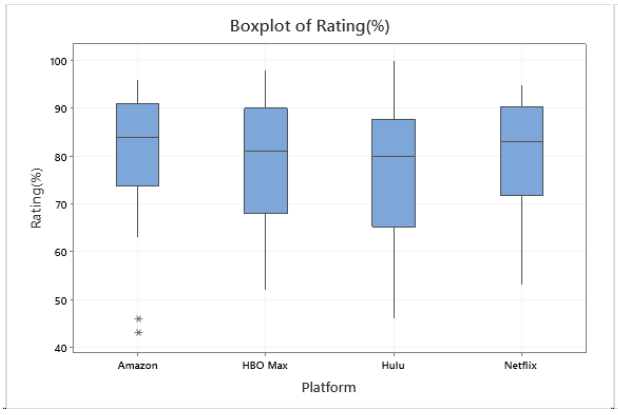
	Amazon	Apple TV+	Disney+	HBO Max	Hulu	Netflix	All
Action & Adventure	4	0	2	1	2	8	17
Animation	2	1	1	3	2	2	11
Comedy	13	2	2	4	12	7	40
Drama	20	7	0	4	17	21	69
Horror	1	0	0	0	4	5	10
Kids & Family	0	0	2	1	0	0	3
Musical & Performing Arts	0	0	2	0	0	0	2
Mystery & Suspense	3	2	0	2	0	4	11
Science Fiction & Fantasy	3	0	2	1	3	7	16
All	46	12	11	16	40	54	179

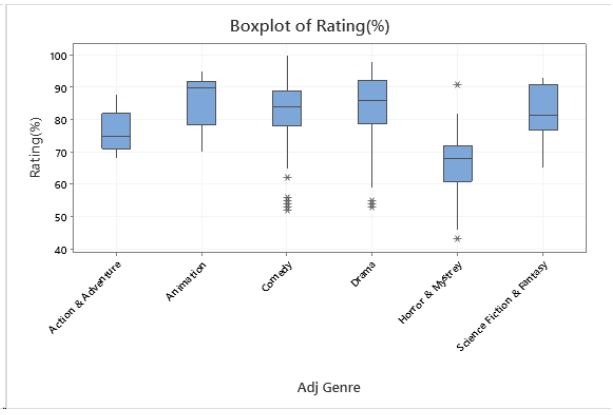
Therefore, I will take out the Apple TV+ and Disney+ from the columns and Kids & Family and Musical & Performing Arts from the rows. I will merge Horror and Mystery & Suspense into one category, Horror & Mystery, because they are quite similar in terms of giving people heart attacks. We are still left with 155 observations, more than 86% of the original data. The adjusted counts are listed below:

Rows: Adj Genre Columns: Platform

	Amazon	HBO Max	Hulu	Netflix	All
Action & Adventure	4	1	2	8	15
Animation	2	3	2	2	9
Comedy	13	4	12	7	36
Drama	20	4	17	21	62
Horror & Mystrey	4	2	4	9	19
Science Fiction & Fantasy	3	1	3	7	14
All	46	15	40	54	155

Here, HBO Max Action & Adventure and Science Fiction & Fantasy only have 1 datapoint, so we will need to pay special attention to them later in the analysis. Let's look at the boxplots:





The average ratings across different platforms are similar, with Amazon and Netflix being slightly higher. There seems to be some indication of nonconstant variance: Amazon has less variability than Hulu. The average ratings across genres are very different.

Horror & Mystery is significantly lower than every other genre. Action & Adventure is second to last. Animation has the highest average rating. Nonconstant variance is more noticeable here. Action & Adventure has much less variability than, for example, Drama.

Now, let's attempt to fit a two-way ANOVA model for 155 Action & Adventure, Animation, Comedy, Drama, Horror & Mystery and Science Fiction & Fantasy TV shows on Amazon Prime, HBO Max, Hulu and Netflix.

General Linear Model: Rating(%) versus Platform, Adj Genre

Method

Factor coding (-1, 0, +1)

Factor Information

Factor	Type	Levels	Values
Platform	Fixed	4	Amazon, HBO Max, Hulu, Netflix
Adj Genre	Fixed	6	Action & Adventure, Animation, Comedy, Drama, Horror & Mystrey, Science Fiction & Fantasy

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Platform	3	531.0	177.0	1.44	0.234
Adj Genre	5	4382.3	876.5	7.13	0.000
Platform*Adj Genre	15	2004.3	133.6	1.09	0.374
Error	131	16101.7	122.9		
Total	154	23776.8			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
11.0867	32.28%	20.39%	*

Coefficients

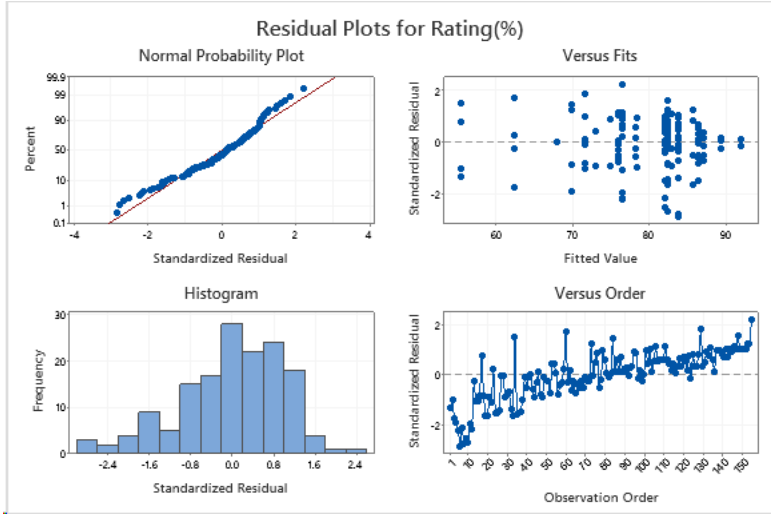
Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	77.87	1.27	61.19	0.000	
Platform					
Amazon	1.75	2.03	0.86	0.390	3.33
HBO Max	-1.19	2.70	-0.44	0.661	3.52
Hulu	-3.28	2.14	-1.53	0.128	3.44
Adj Genre					
Action & Adventure	-2.81	3.35	-0.84	0.404	2.65
Animation	7.71	3.32	2.33	0.022	2.05
Comedy	0.98	2.11	0.46	0.644	1.70
Drama	5.65	1.92	2.94	0.004	1.84
Horror & Mystrey	-13.45	2.70	-4.98	0.000	1.95
Platform*Adj Genre					
Amazon Action & Adventure	1.44	4.89	0.29	0.769	4.29
Amazon Animation	4.67	5.83	0.80	0.425	3.84
Amazon Comedy	5.86	3.18	1.85	0.067	2.41
Amazon Drama	-2.77	2.87	-0.96	0.337	3.42
Amazon Horror & Mystrey	-10.67	4.48	-2.38	0.019	3.75
HBO Max Action & Adventure	-2.87	7.61	-0.38	0.706	8.00
HBO Max Animation	4.94	5.51	0.90	0.372	3.13
HBO Max Comedy	-7.66	4.52	-1.70	0.092	3.14
HBO Max Drama	3.42	4.43	0.77	0.442	5.15
HBO Max Horror & Mystrey	4.77	5.79	0.82	0.411	5.17
Hulu Action & Adventure	3.21	5.89	0.55	0.586	5.63
Hulu Animation	-3.81	5.87	-0.65	0.518	3.89
Hulu Comedy	0.93	3.29	0.28	0.778	2.50
Hulu Drama	1.82	3.01	0.60	0.547	3.54
Hulu Horror & Mystrey	1.36	4.53	0.30	0.765	3.83

Regression Equation

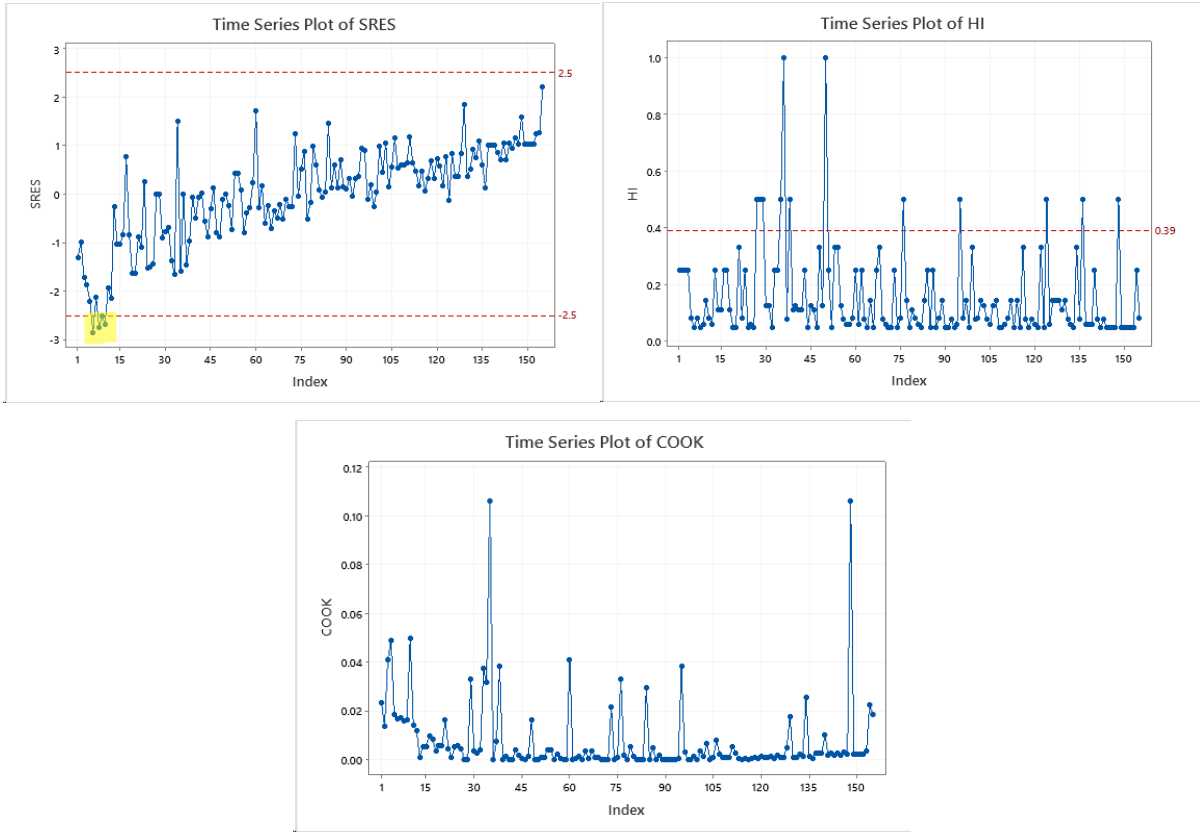
Rating(%) = 77.87 + 1.75 Platform_Amazon - 1.19 Platform_HBO Max - 3.28 Platform_Hulu
+ 2.71 Platform_Netflix - 2.81 Adj Genre_Action & Adventure
+ 7.71 Adj Genre_Animation + 0.98 Adj Genre_Comedy + 5.65 Adj Genre_Drama
- 13.45 Adj Genre_Horror & Mystrey + 1.92 Adj Genre_Science Fiction & Fantasy
+ 1.44 Platform*Adj Genre_Amazon Action & Adventure
+ 4.67 Platform*Adj Genre_Amazon Animation + 5.86 Platform*Adj Genre_Amazon Comedy - 2.77 Platform*Adj Genre_Amazon Drama - 10.67 Platform*Adj Genre_Amazon Horror & Mystrey + 1.46 Platform*Adj Genre_Amazon Science Fiction & Fantasy
- 2.87 Platform*Adj Genre_HBO Max Action & Adventure
+ 4.94 Platform*Adj Genre_HBO Max Animation - 7.66 Platform*Adj Genre_HBO Max Comedy + 3.42 Platform*Adj Genre_HBO Max Drama + 4.77 Platform*Adj Genre_HBO Max Horror & Mystrey - 2.60 Platform*Adj Genre_HBO Max Science Fiction & Fantasy
+ 3.21 Platform*Adj Genre_Hulu Action & Adventure - 3.81 Platform*Adj Genre_Hulu Animation + 0.93 Platform*Adj Genre_Hulu Comedy + 1.82 Platform*Adj Genre_Hulu Drama + 1.36 Platform*Adj Genre_Hulu Horror & Mystrey
- 3.51 Platform*Adj Genre_Hulu Science Fiction & Fantasy
- 1.78 Platform*Adj Genre_Netflix Action & Adventure
- 5.80 Platform*Adj Genre_Netflix Animation + 0.87 Platform*Adj Genre_Netflix Comedy - 2.47 Platform*Adj Genre_Netflix Drama + 4.54 Platform*Adj Genre_Netflix Horror & Mystrey + 4.64 Platform*Adj Genre_Netflix Science Fiction & Fantasy

Means

Term	Fitted Mean	SE Mean
Platform		
Amazon	79.62	2.23
HBO Max	76.68	3.37
Hulu	74.59	2.43
Netflix	80.58	1.91
Adj Genre		
Action & Adventure	75.06	3.80
Animation	85.58	3.75
Comedy	78.85	2.06
Drama	83.52	1.77
Horror & Mystrey	64.42	2.92
Science Fiction & Fantasy	79.79	3.73
Platform*Adj Genre		
Amazon Action & Adventure	78.25	5.54
Amazon Animation	92.00	7.84
Amazon Comedy	86.46	3.07
Amazon Drama	82.50	2.48
Amazon Horror & Mystrey	55.50	5.54
Amazon Science Fiction & Fantasy	83.00	6.40
HBO Max Action & Adventure	71.0	11.1
HBO Max Animation	89.33	6.40
HBO Max Comedy	70.00	5.54
HBO Max Drama	85.75	5.54
HBO Max Horror & Mystrey	68.00	7.84
HBO Max Science Fiction & Fantasy	76.0	11.1
Hulu Action & Adventure	75.00	7.84
Hulu Animation	78.50	7.84
Hulu Comedy	76.50	3.20
Hulu Drama	82.06	2.69
Hulu Horror & Mystrey	62.50	5.54
Hulu Science Fiction & Fantasy	73.00	6.40
Netflix Action & Adventure	76.00	3.92
Netflix Animation	82.50	7.84
Netflix Comedy	82.43	4.19
Netflix Drama	83.76	2.42
Netflix Horror & Mystrey	71.67	3.70
Netflix Science Fiction & Fantasy	87.14	4.19



There is nonconstant variance and non-normality. The normal plot seems to have a longer left tail than a right tail. For the sake of alleviating nonconstant variance, I tried taking log on the response, but the assumption fitting does not improve. Therefore, I will then check the unusual observations to see if the model will fit the assumptions better. For the genre&platform combinations with count 1, I manually fill in standard residual and Cook's distance as 0 (leverage values are 1).



There are four outliers: *Designated Survivor, The Irregulars, A Teacher* and *Emily In Paris*. They are not the tv shows with the lowest ratings, but they do have unusually low ratings in their own group (Netflix Drama, Hulu Drama, Netflix Comedy). Reasons why they are outliers are hard to explain. In the HI plot, the reference line is 2.5 * (23 + 1) / 155 = 0.39 (p = 3 + 5 + 15 = 23). There are plenty of leverage points, two of which show up very unusual in the Cook plot, but they are all underrepresented platform+genre combination with only 1 or 2 counts. Therefore, here we only take out the four outliers and do another two-way ANOVA model fitting.

General Linear Model: Rating(%) versus Platform, Adj Genre

Method

Factor coding (-1, 0, +1)

Factor Information

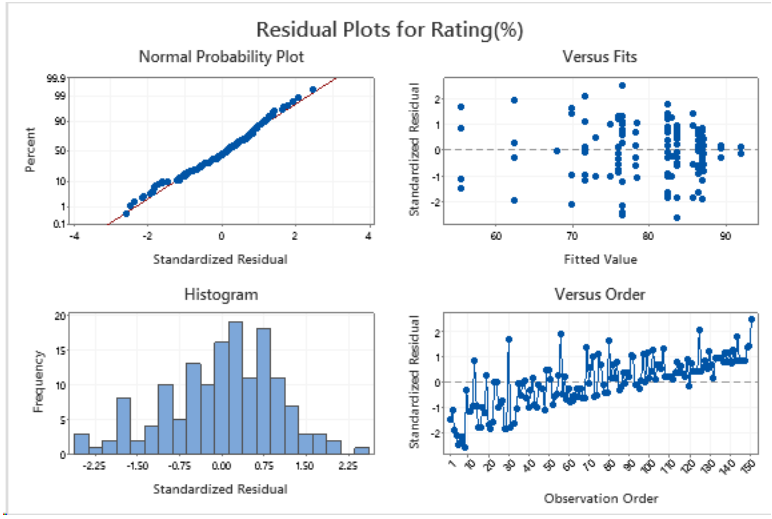
Factor	Type	Levels	Values
Platform	Fixed	4	Amazon, HBO Max, Hulu, Netflix
Adj Genre	Fixed	6	Action & Adventure, Animation, Comedy, Drama, Horror & Mystrey, Science Fiction & Fantasy

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Platform	3	692.9	230.97	2.36	0.074
Adj Genre	5	4829.1	965.82	9.87	0.000
Platform*Adj Genre	15	2010.2	134.01	1.37	0.172
Error	127	12421.2	97.80		
Total	150	21093.7			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
9.88963	41.11%	30.45%	*



The normal plot looks better but nonconstant variance persists (the middle seems to have a larger variance than the left and especially the right). The interaction effect is not statistically significant, with a p-value of 0.172. Therefore, I will fit another two-way ANOVA model with only the main effects.

General Linear Model: Rating(%) versus Platform, Adj Genre

Method

Factor coding (-1, 0, +1)

Factor Information

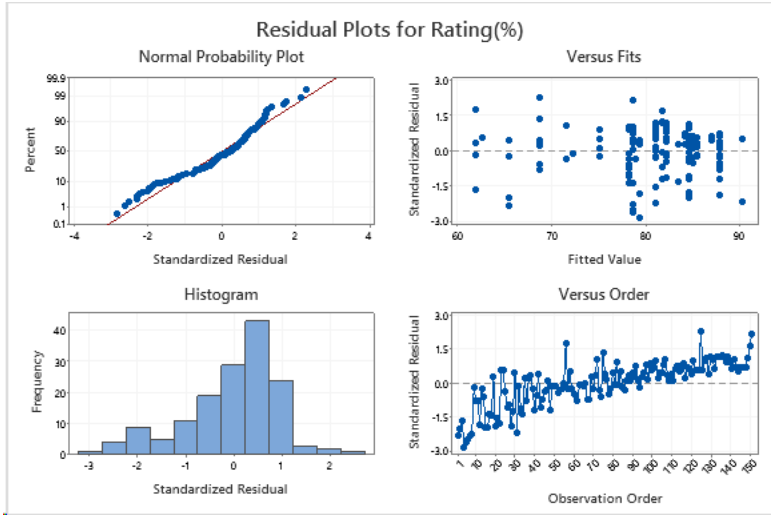
Factor	Type	Levels	Values
Platform	Fixed	4	Amazon, HBO Max, Hulu, Netflix
Adj Genre	Fixed	6	Action & Adventure, Animation, Comedy, Drama, Horror & Mystrey, Science Fiction & Fantasy

Analysis of Variance

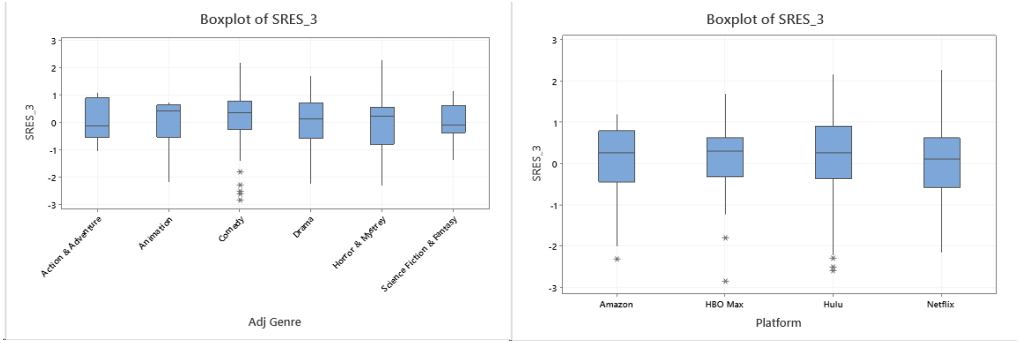
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Platform	3	1068	356.06	3.50	0.017
Adj Genre	5	6087	1217.31	11.98	0.000
Error	142	14431	101.63		
Lack-of-Fit	15	2010	134.01	1.37	0.172
Pure Error	127	12421	97.80		
Total	150	21094			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
10.0811	31.58%	27.73%	22.76%



As we can see, both the Platform effect and the Adj Genre effect are statistically significant, but the assumptions are less well-satisfied. The problem of nonconstant variance seems to have worsened and the boxplots suggest the same, even though the Levene's tests are not significant (I am not putting the plots here but the p-values are about 0.2, 0.3, much bigger than 0.1).



I would like to try taking log to solve it. Interestingly, the 4 outliers in models that are not logged are exactly the same as those in the logged models. Therefore, I will try using Logged Rating (base 10) as the new response to fit the data without the 4 outliers and add the interaction effect back in to check its significance.

General Linear Model: Logged Rating versus Platform, Adj Genre

Method

Factor coding (-1, 0, +1)

Factor Information

Factor	Type	Levels	Values
Platform	Fixed	4	Amazon, HBO Max, Hulu, Netflix
Adj Genre	Fixed	6	Action & Adventure, Animation, Comedy, Drama, Horror & Mystrey, Science Fiction & Fantasy

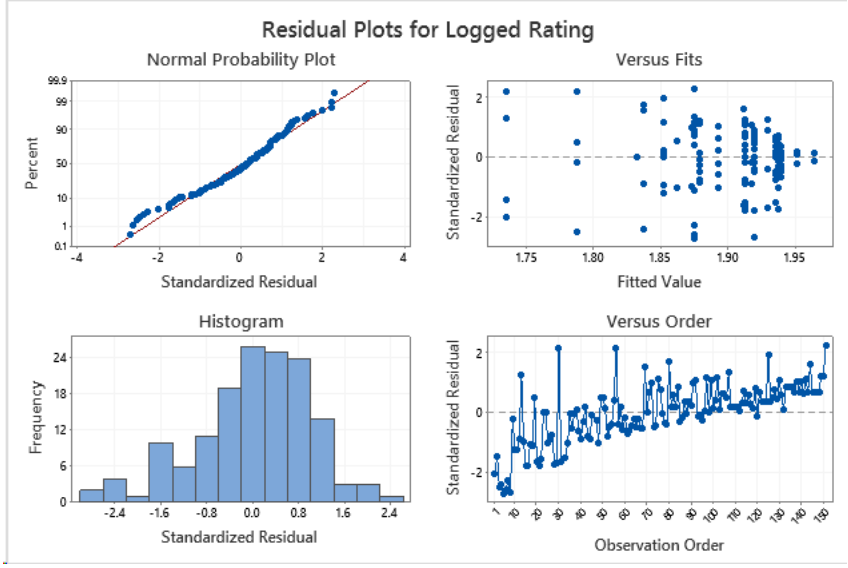
Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Platform	3	0.02350	0.007832	2.32	0.078
Adj Genre	5	0.17408	0.034817	10.33	0.000
Platform*Adj Genre	15	0.08033	0.005355	1.59	0.086
Error	127	0.42815	0.003371		
Total	150	0.74692			

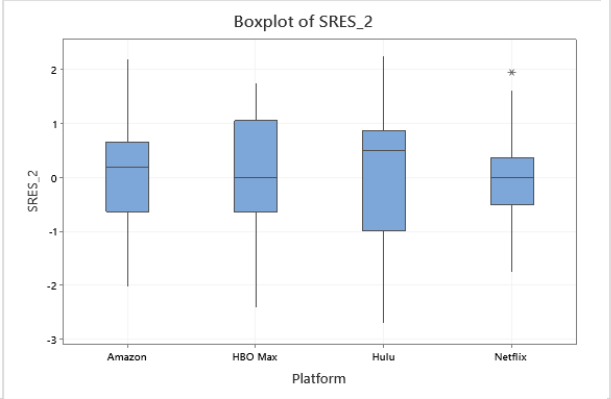
Model Summary

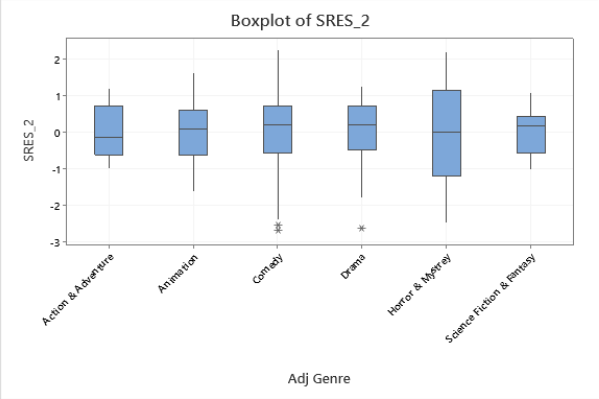
S	R-sq	R-sq(adj)	R-sq(pred)
0.0580627	42.68%	32.30%	*

This time, there is marginal evidence to reject the null hypothesis that interaction effect doesn't exist.



The assumption seems to fit a little bit better. The normal plot is a little bit straighter in the middle, while nonconstant variance still exists. The box plots also suggest the same.





Let's use the Levene's test to check again.

General Linear Model: absres_2 versus Platform, Adj Genre

Method

Factor coding (-1, 0, +1)

Factor Information

Factor	Type	Levels	Values
Platform	Fixed	4	Amazon, HBO Max, Hulu, Netflix
Adj Genre	Fixed	6	Action & Adventure, Animation, Comedy, Drama, Horror & Mystrey, Science Fiction & Fantasy

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Platform	3	2.661	0.8869	3.12	0.028
Adj Genre	5	2.944	0.5889	2.07	0.073
Platform*Adj Genre	15	17.497	1.1665	4.11	0.000
Error	127	36.073	0.2840		
Total	150	61.310			

The F-test of the nonconstant variance problem of the interaction effect is highly statistically significant, showing that there is indeed nonconstant variance and it is related to both Platform and Adj Genre. Therefore, we can go on to perform Weighted Least Squares to solve this problem. We need the weights first:

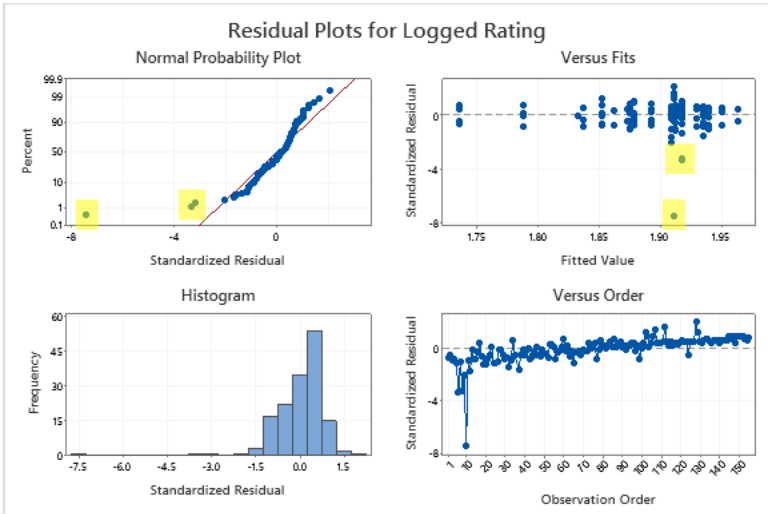
95% Bonferroni Confidence Intervals for Standard Deviations

Platform	Adj Genre	N	StDev	CI
Amazon	Action & Adventure	4	0.52734	(0.015810, 73.144)
Amazon	Animation	2	0.16261	(*, *)
Amazon	Comedy	13	0.66029	(0.276992, 2.054)
Amazon	Drama	20	1.03092	(0.681418, 1.839)
Amazon	Horror & Mystrey	4	2.04714	(0.124537, 139.932)
Amazon	Science Fiction & Fantasy	3	0.99849	(*, *)
HBO Max	Action & Adventure	1	*	(*, *)
HBO Max	Animation	3	0.21458	(*, *)
HBO Max	Comedy	4	1.99360	(0.086504, 191.055)
HBO Max	Drama	4	1.32310	(0.041080, 177.204)
HBO Max	Horror & Mystrey	2	0.00000	(*, *)
HBO Max	Science Fiction & Fantasy	1	*	(*, *)
Hulu	Action & Adventure	2	1.40030	(*, *)
Hulu	Animation	2	1.43362	(*, *)
Hulu	Comedy	12	1.69999	(0.905893, 4.272)
Hulu	Drama	16	1.00189	(0.366905, 3.377)
Hulu	Horror & Mystrey	4	1.93568	(0.034759, 448.249)
Hulu	Science Fiction & Fantasy	3	0.89605	(*, *)
Netflix	Action & Adventure	8	0.78767	(0.246096, 4.065)
Netflix	Animation	2	2.28418	(*, *)
Netflix	Comedy	6	0.28650	(0.072082, 2.307)
Netflix	Drama	19	0.69038	(0.366229, 1.549)
Netflix	Horror & Mystrey	9	1.06319	(0.406342, 4.199)
Netflix	Science Fiction & Fantasy	7	0.43319	(0.087490, 3.790)

Individual confidence level = 99.7619%

The weights are calculated as 1 / (StDev ** 2). For those marked * or have a standard deviation that equals 0, I manually set their weights to be 1.

Then, let's fit a WLS using all observations, including the 4 outliers. Maybe they will no longer be outliers. I will continue to use Logged Rating as the response because the slight right tail exists throughout all model fitting. But unfortunately, 3 of the 4 previous outliers, *Designated Survivor, The Irregulars* and *Emily In Paris*, are still apparently outlying.



Taking them out results in more outliers popping up, so I ended up excluding 7 outliers in total and fitting the following model.

General Linear Model: Logged Rating versus Platform, Adj Genre

Method

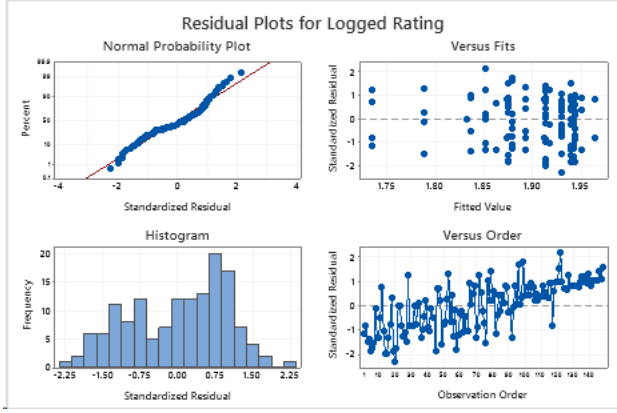
Factor coding	(-1, 0, +1)
Weights	wt

Factor Information

Factor	Type	Levels	Values
Platform	Fixed	4	Amazon, HBO Max, Hulu, Netflix
Adj Genre	Fixed	6	Action & Adventure, Animation, Comedy, Drama, Horror & Mystrey, Science Fiction & Fantasy

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Platform	3	0.01062	0.003540	1.42	0.239
Adj Genre	5	0.09059	0.018118	7.28	0.000
Platform*Adj Genre	15	0.05520	0.003680	1.48	0.123
Error	124	0.30847	0.002488		
Total	147	0.66832			



The assumptions fit much better. However, the interaction effect is insignificant, so I started over from the beginning with all outliers and fit WLS models without interaction effect. Now, I have excluded 6 outliers: *Designated Survivor*, *The Irregulars*, *Emily In Paris*, *The Teacher*, *Little Fires Everywhere* and *Ratched*. The current model is as follows:

General Linear Model: Logged Rating versus Platform, Adj Genre

Method

Factor coding	(-1, 0, +1)
Weights	wt

Factor Information

Factor	Type	Levels	Values
Platform	Fixed	4	Amazon, HBO Max, Hulu, Netflix
Adj Genre	Fixed	6	Action & Adventure, Animation, Comedy, Drama, Horror & Mystrey, Science Fiction & Fantasy

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Platform	3	0.03754	0.012515	4.64	0.004
Adj Genre	5	0.26544	0.053087	19.69	0.000
Error	140	0.37751	0.002697		
Lack-of-Fit	15	0.05144	0.003429	1.31	0.203
Pure Error	125	0.32607	0.002609		
Total	148	0.68453			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.0519278	44.85%	41.70%	37.72%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	1.90307	0.00432	440.31	0.000	
Platform					
Amazon	0.00579	0.00483	1.20	0.232	2.68
HBO Max	-0.00777	0.00693	-1.12	0.264	3.90
Hulu	-0.01653	0.00858	-1.93	0.056	3.96
Adj Genre					
Action & Adventure	-0.02859	0.00869	-3.29	0.001	2.07
Animation	0.05473	0.00694	7.88	0.000	2.93
Comedy	0.01830	0.00558	3.28	0.001	1.65
Drama	0.02183	0.00630	3.47	0.001	1.72
Horror & Mystrey	-0.0795	0.0128	-6.20	0.000	3.31

Regression Equation

Logged Rating = 1.90307 + 0.00579 Platform_Amazon - 0.00777 Platform_HBO Max - 0.01653 Platform_Hulu + 0.01851 Platform_Netflix - 0.02859 Adj Genre_Action & Adventure + 0.05473 Adj Genre_Animation + 0.01830 Adj Genre_Comedy + 0.02183 Adj Genre_Drama - 0.0795 Adj Genre_Horror & Mystrey + 0.01321 Adj Genre_Science Fiction & Fantasy

Means

Term	Fitted Mean	SE Mean
Platform		
Amazon	1.90886	0.00580
HBO Max	1.89530	0.00861
Hulu	1.8865	0.0108
Netflix	1.92158	0.00489
Adj Genre		
Action & Adventure	1.8745	0.0100
Animation	1.95780	0.00606
Comedy	1.92138	0.00617
Drama	1.92490	0.00650
Horror & Mystrey	1.8236	0.0153
Science Fiction & Fantasy	1.91628	0.00880

The typical Rating of Amazon TV shows is 10**1.90886 = 81.1%; HBO Max is 78.6%; Hulu is 77.0%; Netflix is 83.4%. The typical Rating for Action & Adventure shows is 74.9%; Animation is as high as 90.75%; Comedy is 83.9%; Drama is 84.12%; Horror & Mystery is only 66.65; Science Fiction & Fantasy is 82.4%. The rating differences between the genres can be quite large.

The two main effects are both highly statistically significant. The pure error lack of fit test, on the other hand, has a p-value of 0.203, indicating that the interaction effect is indeed not statistically significant. Since there is no interaction effect, we do not need to look at the interaction plot. But the main effects are important, so let's look at some comparisons:

Tukey Pairwise Comparisons: Platform

Grouping Information Using the Tukey Method and 95% Confidence

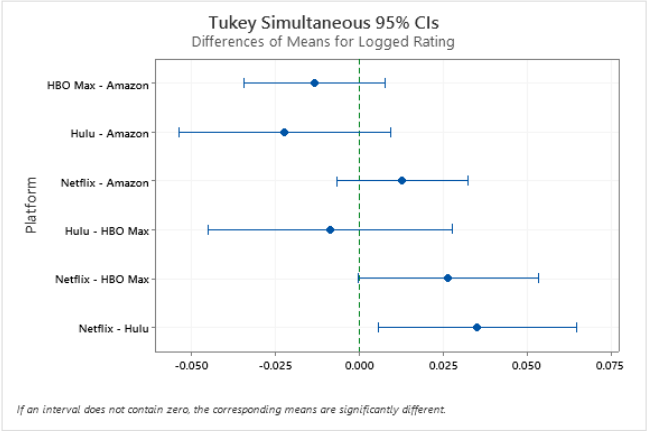
Platform	N	Mean	Grouping	
Netflix	50	1.92158	A	
Amazon	46	1.90886	A	B
HBO Max	15	1.89530	A	B
Hulu	38	1.88654		B

Means that do not share a letter are significantly different.

Tukey Simultaneous Tests for Differences of Means

Difference of Platform Levels	Difference of Means	SE of Difference	Simultaneous 95% CI	T-Value	Adjusted P-Value
HBO Max - Amazon	-0.01356	0.00807	(-0.03455, 0.00743)	-1.68	0.337
Hulu - Amazon	-0.0223	0.0121	(-0.0538, 0.0091)	-1.85	0.256
Netflix - Amazon	0.01272	0.00747	(-0.00671, 0.03215)	1.70	0.326
Hulu - HBO Max	-0.0088	0.0139	(-0.0450, 0.0275)	-0.63	0.923
Netflix - HBO Max	0.0263	0.0103	(-0.0006, 0.0531)	2.55	0.057
Netflix - Hulu	0.0350	0.0113	(0.0057, 0.0644)	3.10	0.012

Individual confidence level = 98.97%



We can see that given Adj Genre, Netflix and Hulu are very different from each other in terms of Logged Rating (logged Rotten Tomato Average Audience Score), while Netflix, Amazon and HBO Max are similar and Amazon, HBO Max and Hulu are similar. Adj Genre, on the other hand, shows much more distinctive differences.

Tukey Pairwise Comparisons: Adj Genre

Grouping Information Using the Tukey Method and 95% Confidence

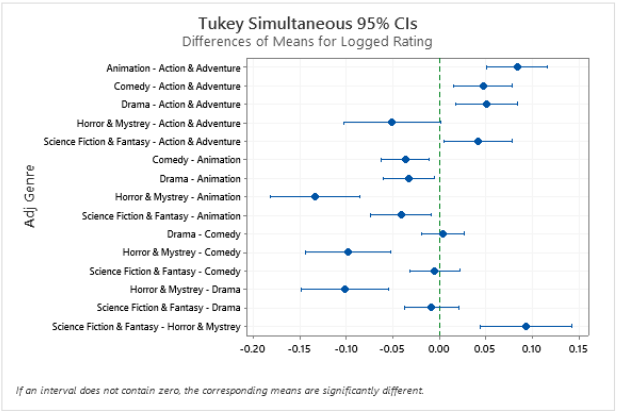
Adj Genre	N	Mean	Grouping		
Animation	9	1.95780	A		
Drama	57	1.92490		B	
Comedy	35	1.92138		B	
Science Fiction & Fantasy	14	1.91628		B	
Action & Adventure	15	1.87449			C
Horror & Mystrey	19	1.82359			C

Means that do not share a letter are significantly different.

Tukey Simultaneous Tests for Differences of Means

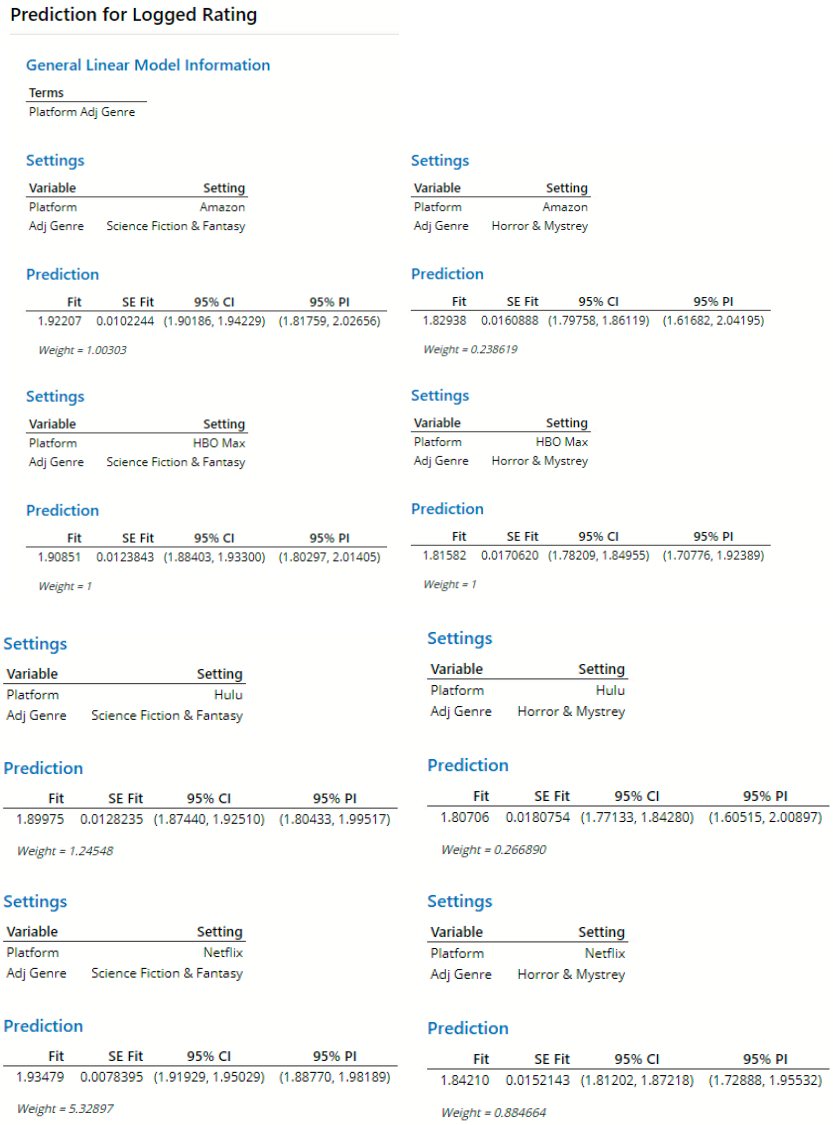
Difference of Adj Genre Levels	Difference of Means	SE of Difference	Simultaneous 95% CI	T-Value	Adjusted P-Value
Animation - Action & Adventure	0.0833	0.0115	(0.0501, 0.1165)	7.26	0.000
Comedy - Action & Adventure	0.0469	0.0110	(0.0152, 0.0786)	4.28	0.000
Drama - Action & Adventure	0.0504	0.0116	(0.0170, 0.0839)	4.36	0.000
Horror & Mystrey - Action & Adventure	-0.0509	0.0180	(-0.1029, 0.0011)	-2.83	0.058
Science Fiction & Fantasy - Action & Adventure	0.0418	0.0127	(0.0050, 0.0786)	3.28	0.016
Comedy - Animation	-0.03642	0.00891	(-0.06220, -0.01064)	-4.09	0.001
Drama - Animation	-0.03290	0.00950	(-0.06038, -0.00542)	-3.46	0.009
Horror & Mystrey - Animation	-0.1342	0.0166	(-0.1822, -0.0862)	-8.08	0.000
Science Fiction & Fantasy - Animation	-0.0415	0.0114	(-0.0744, -0.0087)	-3.66	0.005
Drama - Comedy	0.00352	0.00806	(-0.01979, 0.02683)	0.44	0.998
Horror & Mystrey - Comedy	-0.0978	0.0159	(-0.1437, -0.0519)	-6.16	0.000
Science Fiction & Fantasy - Comedy	-0.00510	0.00932	(-0.03206, 0.02187)	-0.55	0.994
Horror & Mystrey - Drama	-0.1013	0.0163	(-0.1484, -0.0543)	-6.23	0.000
Science Fiction & Fantasy - Drama	-0.0086	0.0101	(-0.0377, 0.0205)	-0.86	0.956
Science Fiction & Fantasy - Horror & Mystrey	0.0927	0.0170	(0.0437, 0.1417)	5.47	0.000

Individual confidence level = 99.56%



Given Platform, Animation has the highest Logged Rating. Drama, Comedy and Science Fiction & Fantasy next and then Action & Adventure and Horror have the lowest.

Then, let's look at some predictions. I present the results for Science Fiction & Fantasy and Horror & Mystery, because they happen to be very different from the comparisons above and are my favorite genres.



Prediction for Logged Rating

General Linear Model Information

Terms
Platform Adj Genre

Settings

Variable	Setting
Platform	Amazon
Adj Genre	Science Fiction & Fantasy

Prediction

Fit	SE Fit	95% CI	95% PI
1.92207	0.0102244	(1.90186, 1.94229)	(1.81759, 2.02656)

Weight = 1.00303

Settings

Variable	Setting
Platform	Amazon
Adj Genre	Horror & Mystrey

Prediction

Fit	SE Fit	95% CI	95% PI
1.82938	0.0160888	(1.79758, 1.86119)	(1.61682, 2.04195)

Weight = 0.238619

Settings

Variable	Setting
Platform	HBO Max
Adj Genre	Science Fiction & Fantasy

Prediction

Fit	SE Fit	95% CI	95% PI
1.90851	0.0123843	(1.88403, 1.93300)	(1.80297, 2.01405)

Weight = 1

Settings

Variable	Setting
Platform	HBO Max
Adj Genre	Horror & Mystrey

Prediction

Fit	SE Fit	95% CI	95% PI
1.81582	0.0170620	(1.78209, 1.84955)	(1.70776, 1.92389)

Weight = 1

Settings

Variable	Setting
Platform	Hulu
Adj Genre	Science Fiction & Fantasy

Prediction

Fit	SE Fit	95% CI	95% PI
1.89975	0.0128235	(1.87440, 1.92510)	(1.80433, 1.99517)

Weight = 1.24548

Settings

Variable	Setting
Platform	Hulu
Adj Genre	Horror & Mystrey

Prediction

Fit	SE Fit	95% CI	95% PI
1.80706	0.0180754	(1.77133, 1.84280)	(1.60515, 2.00897)

Weight = 0.266890

Settings

Variable	Setting
Platform	Netflix
Adj Genre	Science Fiction & Fantasy

Prediction

Fit	SE Fit	95% CI	95% PI
1.93479	0.0078395	(1.91929, 1.95029)	(1.88770, 1.98189)

Weight = 5.32897

Settings

Variable	Setting
Platform	Netflix
Adj Genre	Horror & Mystrey

Prediction

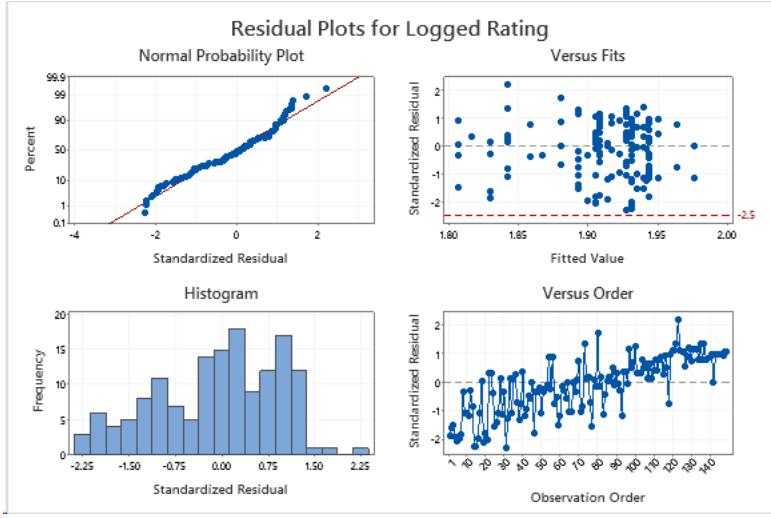
Fit	SE Fit	95% CI	95% PI
1.84210	0.0152143	(1.81202, 1.87218)	(1.72888, 1.95532)

Weight = 0.884664

We can see that Science Fiction & Fantasy has higher predicted logged ratings than Horror & Mystery, which is consistent with the results from the comparisons above. The prediction within each Adj Genre is almost the same (Science Fiction & Fantasy is about 10**1.9 = 79.43(%); Horror & Mystery is about 10**1.82 = 66.07(%)). The intervals are quite similar as well, but the Prediction Interval of Netflix seems a bit narrower in both genres. Netflix's Prediction Interval of Science Fiction & Fantasy and Horror & Mystery are

(10**1.8877, 10**1.98189) = (77.2%, 95.94%) and (10**1.729, 10**1.955) = (53.58%, 90.15%).

The assumptions fit fairly okay, although far from perfect. Nonconstant variance problem is mitigated to a degree but there is still a violation of normality.



Interestingly, the Levene's test and the boxplots are a little bit contradictory. Both the Levene's tests with and without interaction effect show high p-values, meaning that there is no longer a problem of nonconstant variance.

General Linear Model: absres versus Platform, Adj Genre

Method

Factor coding (-1, 0, +1)

Factor Information

Factor	Type	Levels	Values
Platform	Fixed	4	Amazon, HBO Max, Hulu, Netflix
Adj Genre	Fixed	6	Action & Adventure, Animation, Comedy, Drama, Horror & Mystrey, Science Fiction & Fantasy

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Platform	3	0.5834	0.19446	0.52	0.668
Adj Genre	5	0.3791	0.07583	0.20	0.960
Platform*Adj Genre	15	1.2188	0.08125	0.22	0.999
Error	125	46.5714	0.37257		
Total	148	48.6002			

General Linear Model: absres versus Platform, Adj Genre

Method

Factor coding (-1, 0, +1)

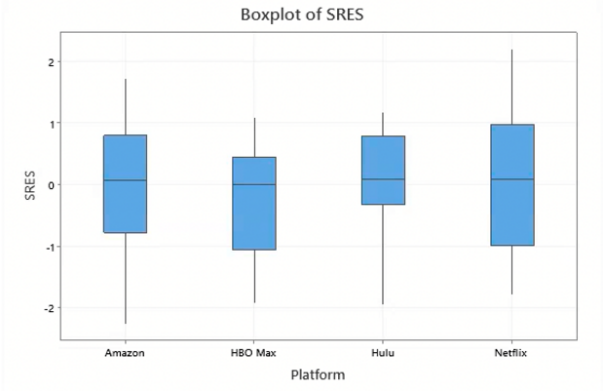
Factor Information

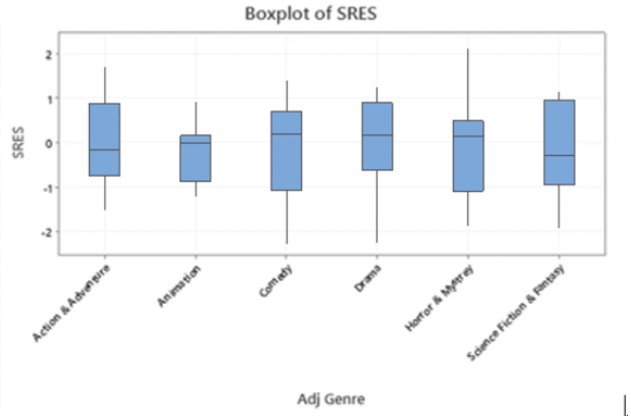
Factor	Type	Levels	Values
Platform	Fixed	4	Amazon, HBO Max, Hulu, Netflix
Adj Genre	Fixed	6	Action & Adventure, Animation, Comedy, Drama, Horror & Mystrey, Science Fiction & Fantasy

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Platform	3	0.5824	0.19414	0.57	0.637
Adj Genre	5	0.2135	0.04270	0.13	0.987
Error	140	47.7901	0.34136		
Lack-of-Fit	15	1.2188	0.08125	0.22	0.999
Pure Error	125	46.5714	0.37257		
Total	148	48.6002			

However, the boxplots show otherwise. There still seems to have a little bit of nonconstant variance, especially among the different genres.





But we are going to stick with this model for now. Comparing AICc also shows that this model is the best compared to the ones having interaction effect and only having one main effect.

General Linear Model: Logged Rating versus Platform, Adj Genre

Method

Factor coding	(-1, 0, +1)
Weights	wt

Factor Information

Factor	Type	Levels	Values
Platform	Fixed	4	Amazon, HBO Max, Hulu, Netflix
Adj Genre	Fixed	6	Action & Adventure, Animation, Comedy, Drama, Horror & Mystrey, Science Fiction & Fantasy

Analysis of Variance

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Platform	3	0.04158	6.07%	0.03754	0.012515	4.64	0.004
Adj Genre	5	0.26544	38.78%	0.26544	0.053087	19.69	0.000
Error	140	0.37751	55.15%	0.37751	0.002697		
Lack-of-Fit	15	0.05144	7.51%	0.05144	0.003429	1.31	0.203
Pure Error	125	0.32607	47.63%	0.32607	0.002609		
Total	148	0.68453	100.00%				

Model Summary

S	R-sq	R-sq(adj)	PRESS	R-sq(pred)	AICc	BIC
0.0519278	44.85%	41.70%	0.426303	37.72%	-483.27	-454.82

General Linear Model: Logged Rating versus Platform, Adj Genre

Method

Factor coding	(-1, 0, +1)
Weights	wt

Factor Information

Factor	Type	Levels	Values
Platform	Fixed	4	Amazon, HBO Max, Hulu, Netflix
Adj Genre	Fixed	6	Action & Adventure, Animation, Comedy, Drama, Horror & Mystrey, Science Fiction & Fantasy

Analysis of Variance

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Platform	3	0.04158	6.07%	0.01048	0.003492	1.34	0.265
Adj Genre	5	0.26544	38.78%	0.09076	0.018152	6.96	0.000
Platform*Adj Genre	15	0.05144	7.51%	0.05144	0.003429	1.31	0.203
Error	125	0.32607	47.63%	0.32607	0.002609		
Total	148	0.68453	100.00%				

Model Summary

S	R-sq	R-sq(adj)	PRESS	R-sq(pred)	AICc	BIC
0.0510742	52.37%	43.60%	*	*	-466.12	-401.59

General Linear Model: Logged Rating versus Platform

Method

Factor coding	(-1, 0, +1)
Weights	wt

Factor Information

Factor	Type	Levels	Values
Platform	Fixed	4	Amazon, HBO Max, Hulu, Netflix

Analysis of Variance

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Platform	3	0.04158	6.07%	0.04158	0.013861	3.13	0.028
Error	145	0.64295	93.93%	0.64295	0.004434		
Lack-of-Fit	20	0.31688	46.29%	0.31688	0.015844	6.07	0.000
Pure Error	125	0.32607	47.63%	0.32607	0.002609		
Total	148	0.68453	100.00%				

Model Summary

S	R-sq	R-sq(adj)	PRESS	R-sq(pred)	AICc	BIC
0.0665892	6.07%	4.13%	0.704132	0.00%	-415.11	-400.51

General Linear Model: Logged Rating versus Adj Genre

Method

Factor coding	(-1, 0, +1)
Weights	wt

Factor Information

Factor	Type	Levels	Values
Adj Genre	Fixed	6	Action & Adventure, Animation, Comedy, Drama, Horror & Mystrey, Science Fiction & Fantasy

Analysis of Variance

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Adj Genre	5	0.26948	39.37%	0.26948	0.053895	18.57	0.000
Error	143	0.41505	60.63%	0.41505	0.002902		
Lack-of-Fit	18	0.08898	13.00%	0.08898	0.004943	1.90	0.022
Pure Error	125	0.32607	47.63%	0.32607	0.002609		
Total	148	0.68453	100.00%				

Model Summary

S	R-sq	R-sq(adj)	PRESS	R-sq(pred)	AICc	BIC
0.0538746	39.37%	37.25%	0.451527	34.04%	-475.94	-455.71

The model without interaction effect has an AICc score of -483, which is much smaller than the model with interaction effect (-466), the model with only Platform effect (-415) and the model with only Adj Genre effect (-475).

In conclusion, the final model of choice is the two-way ANOVA model with Weighted Least Squares for 149 Action & Adventure, Animation, Comedy, Drama, Horror & Mystery and Science Fiction & Fantasy TV shows on Amazon Prime, HBO Max, Hulu and Netflix with both main effects and no interaction effect. Among the four streaming television platforms I examined, only Netflix and Hulu seem to have TV shows with very different ratings. It is possible that they have very different decision processes regarding what shows to produce and what shows to stream on their platforms. Other than that, Netflix, Amazon Prime and HBO Max can be studied as a group given the lack of difference during comparison. Same with Amazon Prime, HBO Max and Hulu. Among all the genres, all platforms have the highest ratings in Animation and lowest rating in Horror & Mystery and Action & Adventure. What's interesting to me is that these three genres all have smaller sizes of samples (9 Animation, 14 Horror & Mystery, 19 Action & Adventure) but they have quite polarized ratings. Other genres with more TV shows being selected by these platforms have more mediocre ratings on average. Perhaps showing fewer mainstream genres is a somewhat risky move and Animation is having a higher reward than Horror & Mystery and Action & Adventure. These are just conjectures, but this model has potentially pointed us to a clearer way to study the similarities and differences between different platforms and genres. However, we also need to understand that this model is

nowhere perfect. We only have 2 predictors. That is most probably not enough to model the complicated market of television streaming.