Exploration of Album Ratings Over The Last Five Years

Catherine Wolk & Sarah Smith

Our Dataset:

Web-scraped Pitchfork for the following information:

- Album Name
- Artist
- Rating Score
- Date Album Released
- Reviewer
- Reviewer Position

Have a dataset of 6,943 ratings

EDA:


pandas
$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$
BeautifulSoup
MySQL
amazon
web services
SciPy

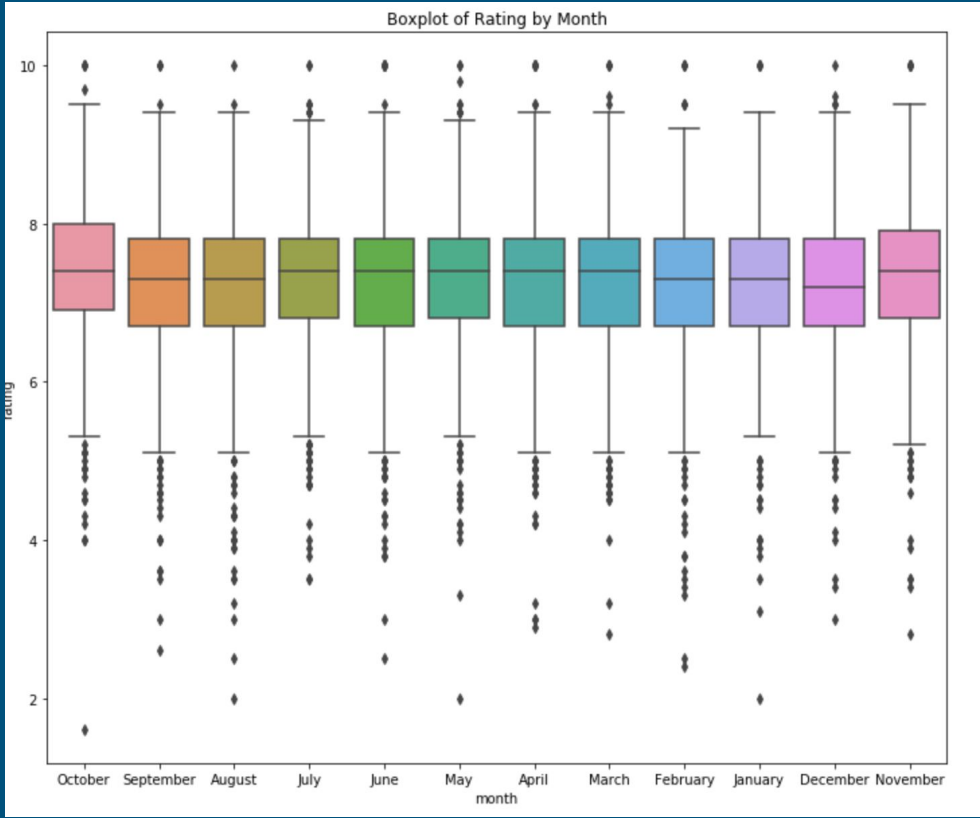
Question 1:
Is there a difference between average overall rating depending on the month it is released?

H_o: There is no significant difference in average rating between months

H_a: There is a significant difference in average rating between months

Test Used:
One way ANOVA
Alpha = 0.01 (due to large sample size)

An overview of rating grouped by month



	rating							
	count	mean	std	min	25%	50%	75%	max
month								
April	653.0	7.212864	1.019557	2.9	6.8	7.30	7.8	10.0
August	604.0	7.129139	1.023300	2.0	6.7	7.30	7.8	10.0
December	289.0	7.192734	1.070168	3.0	6.7	7.30	7.8	10.0
February	610.0	7.138525	1.028178	2.4	6.7	7.30	7.8	10.0
January	550.0	7.166545	1.025650	2.0	6.7	7.30	7.8	10.0
July	601.0	7.174709	0.970426	2.0	6.8	7.30	7.7	10.0
June	650.0	7.224615	1.039872	2.5	6.7	7.40	7.8	10.0
March	673.0	7.215305	0.951684	2.8	6.7	7.40	7.8	10.0
May	655.0	7.299542	0.942579	2.0	6.8	7.40	7.8	10.0
November	484.0	7.273760	1.067996	2.0	6.8	7.40	7.9	10.0
October	548.0	7.318613	0.993866	1.6	6.9	7.40	7.9	10.0
September	626.0	7.226198	1.000568	2.6	6.7	7.35	7.8	10.0

One Way Anova Results:

	df	sum_sq	mean_sq	F	PR(>F)
month	11.0	22.806598	2.073327	2.04303	0.021099
Residual	6931.0	7033.784153	1.014830	NaN	NaN

F-value = 2.04, p-value = 0.02

Here p-value > alpha, therefore we fail to reject the null.

Power of 1 therefore very low chance of a type II error.

Tukey's HSD Results:

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
April	August	-0.0837	0.9	-0.2696	0.1022	False
April	December	-0.0201	0.9	-0.2528	0.2126	False
April	February	-0.0743	0.9	-0.2598	0.1111	False
April	January	-0.0463	0.9	-0.2369	0.1443	False
April	July	-0.0382	0.9	-0.2243	0.148	False
April	June	0.0118	0.9	-0.1707	0.1942	False
April	March	0.0024	0.9	-0.1785	0.1833	False
April	May	0.0867	0.9	-0.0954	0.2688	False
April	November	0.0609	0.9	-0.1366	0.2584	False
April	October	0.1057	0.7862	-0.085	0.2965	False
April	September	0.0133	0.9	-0.1709	0.1976	False
August	December	0.0636	0.9	-0.172	0.2992	False
August	February	0.0094	0.9	-0.1797	0.1984	False
August	January	0.0374	0.9	-0.1567	0.2315	False
August	July	0.0456	0.9	-0.1442	0.2353	False
August	June	0.0955	0.871	-0.0907	0.2816	False
August	March	0.0862	0.9	-0.0984	0.2708	False
August	May	0.1704	0.1093	-0.0154	0.3562	False
August	November	0.1446	0.4425	-0.0563	0.3455	False
August	October	0.1895	0.0641	-0.0048	0.3838	False
August	September	0.0971	0.8633	-0.0908	0.2849	False
December	February	-0.0542	0.9	-0.2894	0.181	False
December	January	-0.0262	0.9	-0.2655	0.2131	False
December	July	-0.018	0.9	-0.2538	0.2177	False
December	June	0.0319	0.9	-0.201	0.2647	False
December	March	0.0226	0.9	-0.209	0.2542	False
December	May	0.1068	0.9	-0.1258	0.3394	False
December	November	0.081	0.9	-0.1638	0.3259	False
December	October	0.1259	0.8447	-0.1135	0.3653	False
December	September	0.0335	0.9	-0.2008	0.2677	False
February	January	0.028	0.9	-0.1656	0.2217	False
February	July	0.0362	0.9	-0.1531	0.2255	False
February	June	0.0861	0.9	-0.0996	0.2717	False
February	March	0.0768	0.9	-0.1073	0.2609	False
February	May	0.161	0.1635	-0.0243	0.3463	False
February	November	0.1352	0.5392	-0.0652	0.3357	False
February	October	0.1801	0.0981	-0.0137	0.3739	False
February	September	0.0877	0.9	-0.0997	0.275	False
January	July	0.0082	0.9	-0.1862	0.2025	False
January	June	0.0581	0.9	-0.1327	0.2489	False
January	March	0.0488	0.9	-0.1405	0.2381	False
January	May	0.133	0.4903	-0.0575	0.3235	False
January	November	0.1072	0.8517	-0.098	0.3125	False
January	October	0.1521	0.3399	-0.0467	0.3508	False
January	September	0.0597	0.9	-0.1328	0.2521	False
July	June	0.0499	0.9	-0.1365	0.2363	False
July	March	0.0406	0.9	-0.1442	0.2254	False
July	May	0.1248	0.5464	-0.0612	0.3109	False
July	November	0.0991	0.9	-0.1021	0.3002	False
July	October	0.1439	0.3966	-0.0506	0.3384	False
July	September	0.0515	0.9	-0.1366	0.2396	False
June	March	-0.0093	0.9	-0.1904	0.1718	False
June	May	0.0749	0.9	-0.1074	0.2573	False
June	November	0.0491	0.9	-0.1486	0.2469	False
June	October	0.094	0.9	-0.097	0.285	False
June	September	0.0016	0.9	-0.1828	0.186	False
March	May	0.0842	0.9	-0.0965	0.265	False
March	November	0.0585	0.9	-0.1378	0.2547	False
March	October	0.1033	0.8049	-0.0862	0.2928	False
March	September	0.0109	0.9	-0.172	0.1938	False
May	November	-0.0258	0.9	-0.2232	0.1716	False
May	October	0.0191	0.9	-0.1716	0.2097	False
May	September	-0.0733	0.9	-0.2574	0.1107	False
November	October	0.0449	0.9	-0.1606	0.2503	False
November	September	-0.0476	0.9	-0.2469	0.1518	False
October	September	-0.0924	0.9	-0.2851	0.1002	False

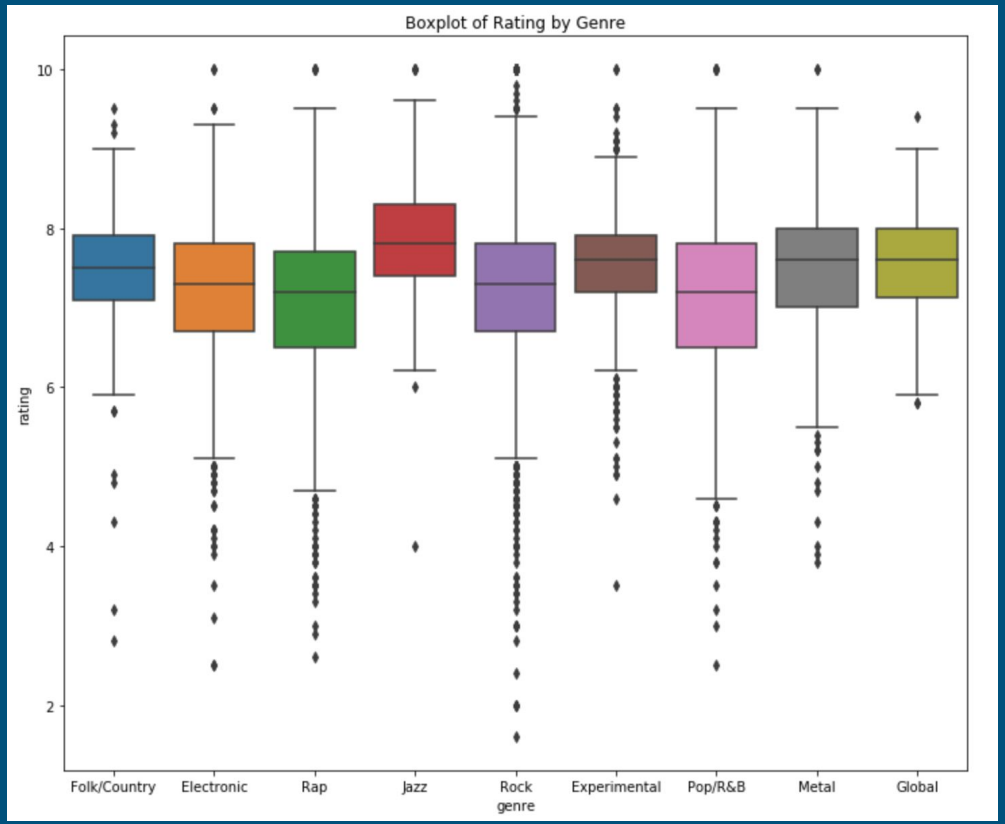
Question 2:
Is there a difference between average overall rating in genres?

H_o: There is no significant difference in average rating between genres

H_a: There is a significant difference in average rating between genres

Test Used:
One way ANOVA
Alpha = 0.01 (due to large sample size)

An overview of rating grouped by genre



	rating							
	count	mean	std	min	25%	50%	75%	max
genre								
Electronic	1164.0	7.169072	0.900801	2.5	6.700	7.3	7.8	10.0
Experimental	646.0	7.495666	0.729349	3.5	7.200	7.6	7.9	10.0
Folk/Country	262.0	7.425954	0.798858	2.8	7.100	7.5	7.9	9.5
Global	66.0	7.524242	0.787413	5.8	7.125	7.6	8.0	9.4
Jazz	152.0	7.848684	0.835946	4.0	7.400	7.8	8.3	10.0
Metal	277.0	7.407942	0.920032	3.8	7.000	7.6	8.0	10.0
Pop/R&B	684.0	7.110526	1.121169	2.5	6.500	7.2	7.8	10.0
Rap	973.0	7.004625	1.057595	2.6	6.500	7.2	7.7	10.0
Rock	2186.0	7.201784	1.081740	1.6	6.700	7.3	7.8	10.0

One Way Anova Results:

	df	sum_sq	mean_sq	F	PR(>F)
genre	8.0	193.209168	24.151146	24.401382	2.566511e-37
Residual	6401.0	6335.357890	0.989745	NaN	NaN

We have a p-value which is significantly low and our f distribution is larger.

This suggests results are statistically significant.

However, we then tested for effect size.

The effect size was found to be significantly low therefore we fail to reject the null.

Power = 1 so low chance of Type II error

Tukey's HSD Results:

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
Electronic	Experimental	0.3266	0.001	0.1751	0.478	True
Electronic	Folk/Country	0.2569	0.0051	0.0458	0.468	True
Electronic	Global	0.3552	0.1094	-0.0354	0.7458	False
Electronic	Jazz	0.6796	0.001	0.4134	0.9458	True
Electronic	Metal	0.2389	0.01	0.0325	0.4452	True
Electronic	Pop/R&B	-0.0585	0.9	-0.2073	0.0902	False
Electronic	Rap	-0.1644	0.0045	-0.2985	-0.0304	True
Electronic	Rock	0.0327	0.9	-0.0793	0.1447	False
Experimental	Folk/Country	-0.0697	0.9	-0.2958	0.1564	False
Experimental	Global	0.0286	0.9	-0.3703	0.4275	False
Experimental	Jazz	0.353	0.0027	0.0747	0.6313	True
Experimental	Metal	-0.0877	0.9	-0.3094	0.134	False
Experimental	Pop/R&B	-0.3851	0.001	-0.5545	-0.2158	True
Experimental	Rap	-0.491	0.001	-0.6477	-0.3344	True
Experimental	Rock	-0.2939	0.001	-0.4321	-0.1556	True
Folk/Country	Global	0.0983	0.9	-0.3269	0.5234	False
Folk/Country	Jazz	0.4227	0.001	0.108	0.7375	True
Folk/Country	Metal	-0.018	0.9	-0.284	0.248	False
Folk/Country	Pop/R&B	-0.3154	0.001	-0.5397	-0.0912	True
Folk/Country	Rap	-0.4213	0.001	-0.6362	-0.2065	True
Folk/Country	Rock	-0.2242	0.0167	-0.426	-0.0224	True
Global	Jazz	0.3244	0.4001	-0.1306	0.7795	False
Global	Metal	-0.1163	0.9	-0.5391	0.3065	False
Global	Pop/R&B	-0.4137	0.0343	-0.8116	-0.0158	True
Global	Rap	-0.5196	0.0013	-0.9123	-0.127	True
Global	Rock	-0.3225	0.1888	-0.7081	0.0632	False
Jazz	Metal	-0.4407	0.001	-0.7523	-0.1292	True
Jazz	Pop/R&B	-0.7382	0.001	-1.015	-0.4614	True
Jazz	Rap	-0.8441	0.001	-1.1133	-0.5748	True
Jazz	Rock	-0.6469	0.001	-0.9058	-0.388	True
Metal	Pop/R&B	-0.2974	0.001	-0.5173	-0.0776	True
Metal	Rap	-0.4033	0.001	-0.6135	-0.1931	True
Metal	Rock	-0.2062	0.0319	-0.403	-0.0093	True
Pop/R&B	Rap	-0.1059	0.4532	-0.2599	0.0481	False
Pop/R&B	Rock	0.0913	0.4792	-0.044	0.2265	False
Rap	Rock	0.1972	0.001	0.0782	0.3161	True

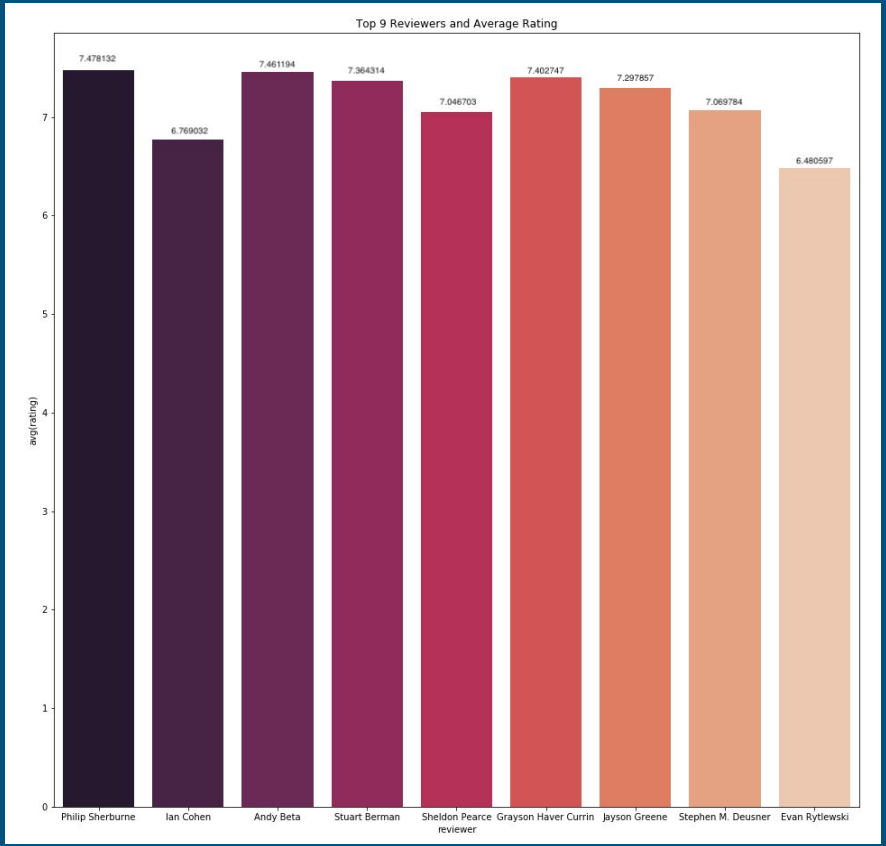
Question 3:
Is there a difference between average overall rating between writers?

H_o: There is no significant difference in average rating between writers

H_a: There is a significant difference in average rating between writers

Test Used:
One way ANOVA
Alpha = 0.01 (due to large sample size)

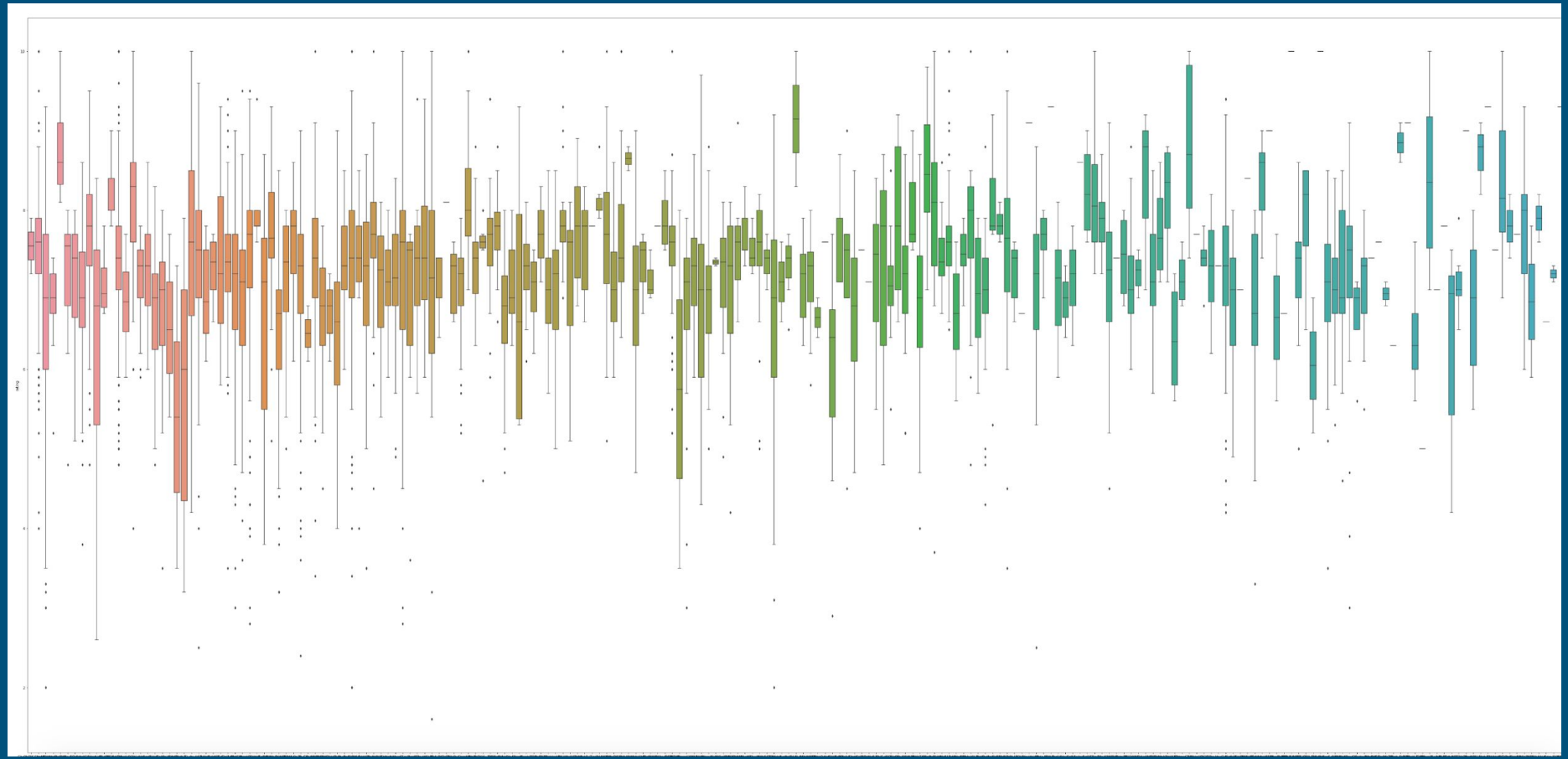
An overview of rating grouped by reviewer



An overview of rating grouped by top 9 reviewers determined by review count

	reviewer	review_count	avg(rating)
0	Philip Sherburne	343	7.478134
1	Ian Cohen	310	6.769032
2	Andy Beta	268	7.461194
3	Stuart Berman	255	7.364314
4	Sheldon Pearce	182	7.046703
5	Grayson Haver Currin	182	7.402747
6	Jayson Greene	140	7.297857
7	Stephen M. Deusner	139	7.069784
8	Evan Rytlewski	134	6.480597

An overview of rating grouped by reviewer



One Way Anova Results:

	df	sum_sq	mean_sq	F	PR(>F)
C(reviewer)	353.0	1303.004794	3.691232	4.227194	3.072283e-121
Residual	6589.0	5753.585956	0.873211	NaN	NaN

We have a p-value which is significantly low again.

This suggests results are statistically significant.

However, we then tested for effect size.

The effect size was found to be significantly low therefore we fail to reject the null.

Power = 1 so low chance of Type II error

Summary:

So, we conclude that:

- It doesn’t matter what month you release your album.
- Some genres have a small difference in average rating than others, but when looked at overall there is no significant difference.
- While there are minor differences, reviewers have the same statistical review average.