Features Affecting Anime Popularity and Scoring

Dexter Luu

Introduction

Animation has emerged as a popular form of media among both young ones and adults and the industry is still improving in the era of technological advances. As of 2015, the global animation industry was estimated to be worth $244 billion with some major markets being the United States, Canada, Japan, China, France, Britain, Korea and Germany(Research and Markets). One of the most popular forms of animation, anime, comes from Japan. Japan's anime market was worth more than $18 billion in 2016, which was 9.9% higher than the previous year(The Asahi Shimbun). With such a large, and increasing, demand for this form of Japanese media, it is imperative to know what customers want to see when choosing which show or movie to indulge in.

Generally, good TV shows and movies share qualities that draw customers in. An online study in June 2016 showed that the genre of the content and the cast members were the two most important factors when deciding what to watch(Statista). This study's aim is to see if those important factors hold true for anime while also looking at other factors available to see what features scored best and were most popular.

Data and Methodology

The data was scraped off of myanimelist.net using Python and the data was analyzed using R. The data consists of the title, rank, score, genre, studio, voice actors, and media type for the 950 most popular anime as of September 2018. A head view of the data can be seen below in table 1. The rank of the anime is the popularity of it based on users who said they watched it on the website. The score is the mean score from user ratings. In anime, there are three main types of media: TV, movies, and OVAs. TV and movies are just like any other TV show and movie while OVAs are original video animations. OVAs can vary in length and

Table 1. *The head of the data scraped from myanimelist.net. The voice actors column was truncated to four entries for formatting but the data had up to 10 entries per anime.*

X	Title	Rank	Score	studio	media.type
0	Death Note	1	8.66	Madhouse	TV
1	Shingeki no Kyojin	2	8.48	Wit Studio	TV
2	Sword Art Online	3	7.60	A-1 Pictures	TV
3	Fullmetal Alchemist: Brotherhood	4	9.24	Bones	TV
4	One Punch Man	5	8.71	Madhouse	TV
5	Tokyo Ghoul	6	7.99	Studio Pierrot	TV

X	genres
0	c("Mystery", "Police", "Psychological", "Supernatural", "Thriller", "Shounen")
1	c("Action", "Military", "Mystery", "Super Power", "Drama", "Fantasy", "Shounen")
2	c("Action", "Adventure", "Fantasy", "Game", "Romance")
3	c("Action", "Adventure", "Comedy", "Drama", "Fantasy", "Magic", "Military", "Shounen")
4	c("Action", "Sci-Fi", "Comedy", "Parody", "Super Power", "Supernatural", "Seinen")
5	c("Action", "Mystery", "Horror", "Psychological", "Supernatural", "Drama", "Seinen")

X	voice.actors
0	c("Yamaguchi, Kappei", "Miyano, Mamoru", "Nakamura, Shidou", "Hidaka, Noriko")
1	c("Ishikawa, Yui", "Kaji, Yuki", "Inoue, Marina", "Kamiya, Hiroshi")
2	c("Matsuoka, Yoshitsugu", "Tomatsu, Haruka", "Taketatsu, Ayana", "Itou, Kanae")
3	c("Park, Romi", "Kugimiya, Rie", "Miki, Shinichiro", "Fujiwara, Keiji")
4	c("Furukawa, Makoto", "Ishikawa, Kaito", "Nakamura, Yuuichi", "Yuuki, Aoi")
5	c("Kobori, Yurie", "Amamiya, Sora", "Kugimiya, Rie", "Miyano, Mamoru")

are about a show or movie but do not have to follow the story. Another media type is a special and is like an OVA in that it does not need to follow the story. As we can see from the genres section, animes, like other media, can fall under multiple genres. Lastly, the voice actors column contains up to ten of the voice actors for the most popular characters in the anime. The table displays only 4 voice actors for formatting.

The data was aggregated by feature for each column and various summary statistics were found for each feature to see if certain characteristics appeared to perform better than others. Only features with more than twenty occurrences were analyzed. This twenty count threshold was much more easily satisfied by features in the genre and voice actor columns because these variables could take on multiple values while the studio and media type variables could only be one value.

Results and Discussion

The first variable that was looked at was the studio. The average rank and score for each studio were found and the minimum, maximum, median, and mean values were tabulated for the mean values. These values can be seen in Table 2 below. In the table, it is shown that Kyoto Animation was the lowest ranked and the highest scored on average. The rankings are based on popularity so it is natural to expect that higher scored shows would have lower rank numbers because better shows should

Table 2. *Summary statistics for the studio variable. The values are the min, max, and median values of the average rank or score for a studio. The mean column is the mean of the means.*

	Average.rank	Rank.stds	Rank.feature	Average.score	Score.stds	Score.feature
Min	320.118	261.003	Kyoto Animation	7.496	0.471	Gonzo
Max	592.478	217.162	Xebec	7.989	0.583	Kyoto Animation
Median	475.038	271.592	Brain's Base	7.768	0.567	Shaft
Mean	468.695	70.246	NA	7.760	0.137	NA

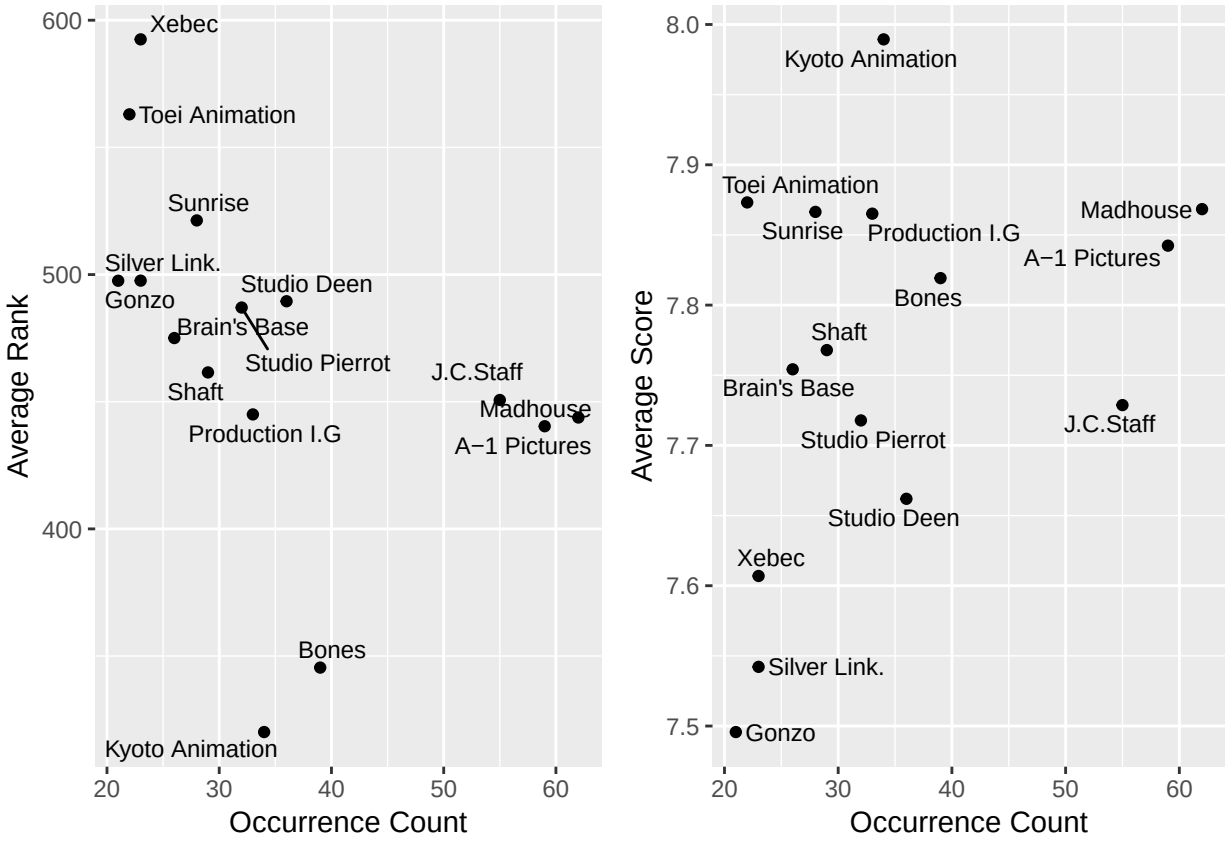

Figure 1. *Average rank vs. occurrence count(left) and average score vs. occurrence count(right).*

draw in more views. This general trend can be seen in Figure 1 for the rest of the studios.

A similar analysis was done for the media type. However, unlike the studios, there were only four types of media that had more than twenty instances in the dataset. Because of this, it was unnecessary to make a

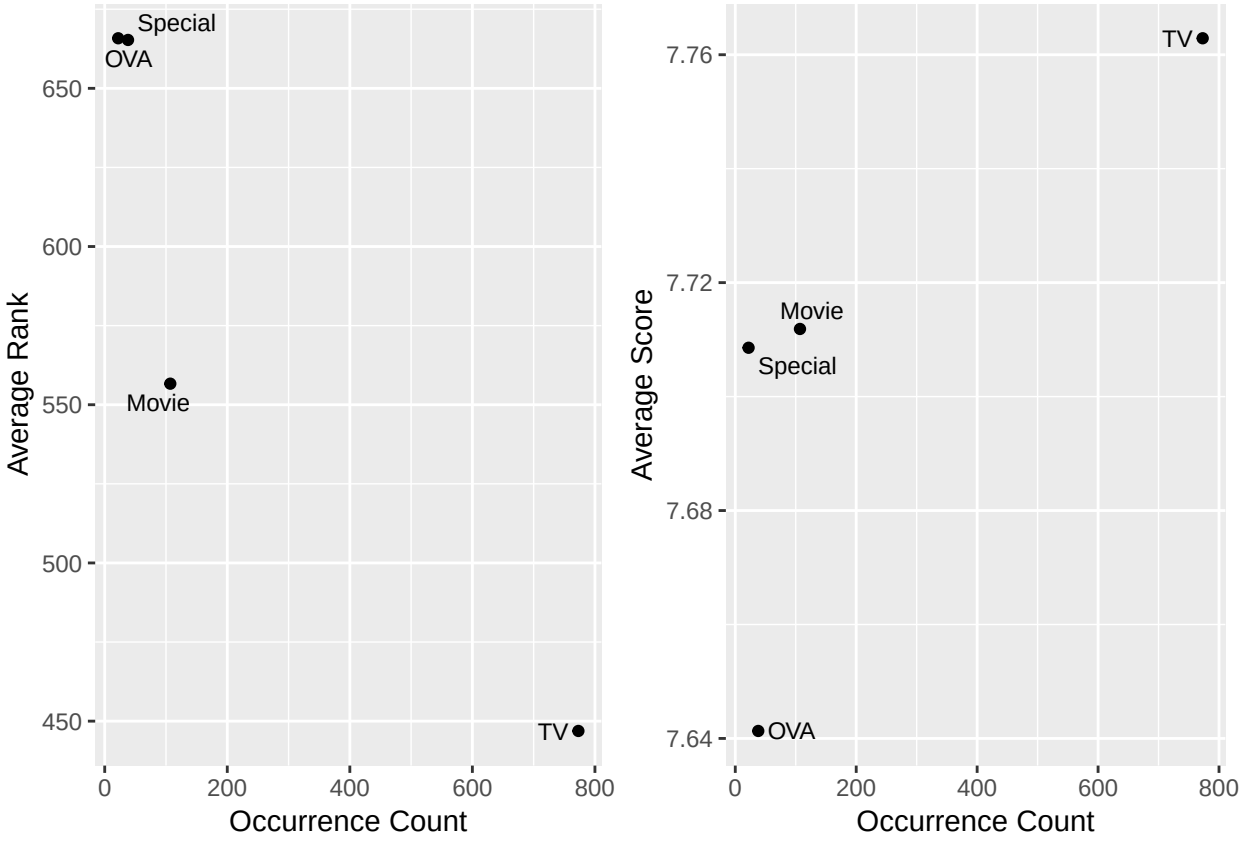

Figure 2. *Average rank vs occurrence count and average score vs occurrence count for media type.*

summary table as the values can be seen in Figure 2. The ratings for the four different media types were all very similar with the difference between the max score and min score being only about 0.12. It seems that viewers didn't particularly think that one form of media had better quality than another.

Although the score did not seem to change between media type, it seems that viewers tended to watch TV shows most often with movies coming in second. TV shows and movies generally have better storylines then OVAs and specials. OVAs and specials could be based off of an existing anime but do not have to follow the story. Because of this, the storyline in these could be seen as unimportant and appear less desirable to watch. The difference in rankings for TV shows and movies can be attributed to the longer stories of a series. Movies are usually a one and done when it comes to watching while a series is watched weekly and can possibly span multiple years depending on the show. Because of the longer story, the show can stay relevant for longer leading to more popularity and more reviews on myanimelist.net. The popularity of TV shows could also be seen in the amount of shows that are made. In this sample of anime, the TV show count was almost eight times higher than movies which was the second highest. If TV shows were not so popular, production companies probably would not make so many.

The next variable that was analyzed was the genre of an anime. With genres being voted the biggest factor in determining what to watch, it was expected that genres would have a wide range of average ranks. Like what was done for the studio variable, summary statistics were found for the different genres and the minimum, maximum, median, and mean values for the average rank and scores of each genre are shown in Table 3. As we can see from Table 3, the difference between the minimum and maximum average ranks was only approximately 138 ranks while the standard deviation was 249 for both of the values. A similar scenario can be said for the minimum and maximum scores. The difference in scores was about 0.41, which is not terribly low, but the standard deviation was 0.94 and 0.51 for the minimum and maximum, respectively.

Table 3. *Summary statistics for the genre variable. The values are the min, max, and median values of the average rank or score for a genre. The mean column is the mean of the means.*

	Average.rank	Rank.stds	Rank.feature	Average.score	Score.stds	Score.feature
Min	390.531	249.677	Vampire	7.463	0.942	Sports
Max	528.000	249.931	Martial Arts	7.874	0.511	Magic
Median	457.519	268.290	Romance	7.755	0.577	Shoujo
Mean	466.657	34.795	NA	7.754	0.079	NA

The small difference between ranks can be a result of an oversaturation of the market with popular genres. If there are so many shows that are classified as being a drama, then it is a given that there will be several small name animes with the drama classification that don't ever become popular. Additionally, an oversaturation of the market can result in the production of several bad animes. These bad animes would then be rated poorly and bring down the rating of the genre overall. A spread of the average rank and scores can be seen in Figure 3 below.

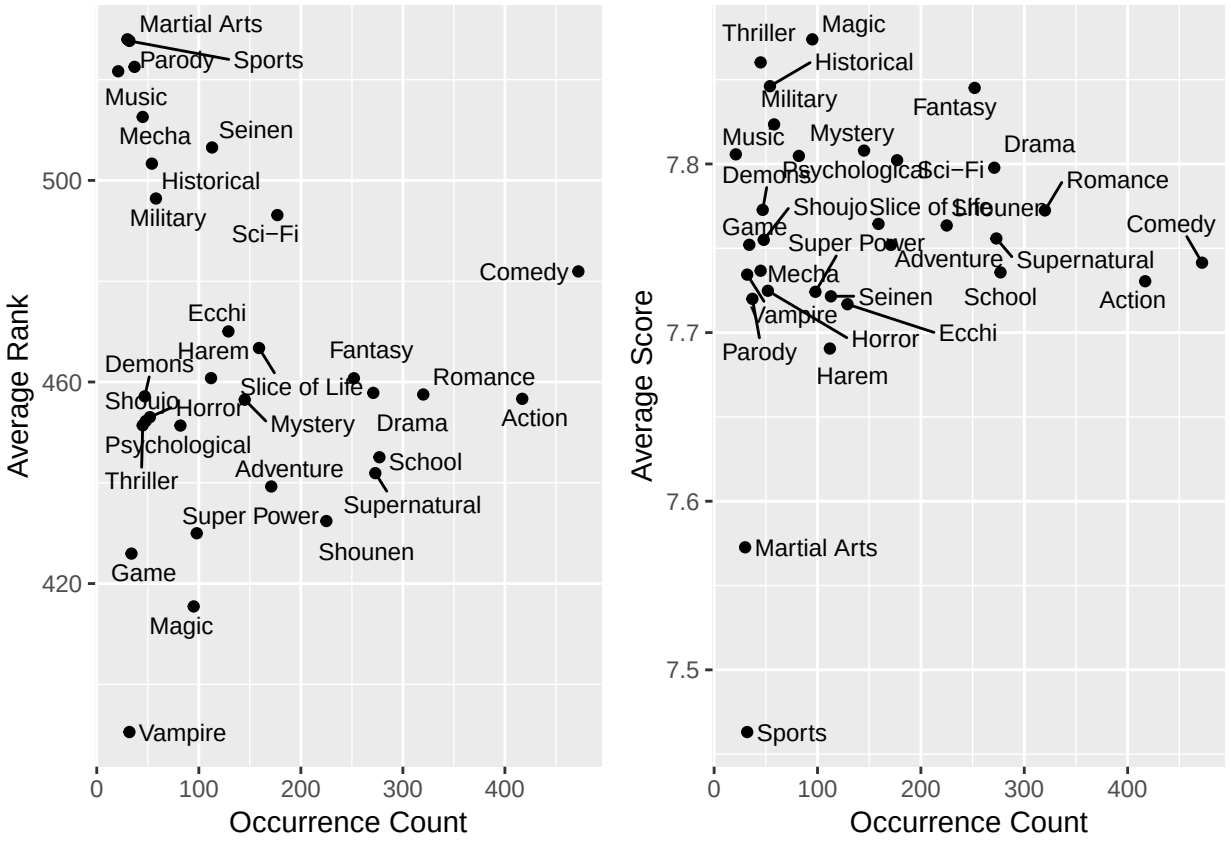

Figure 3. *Average rank vs occurrence count and average score vs occurrence count for genre.*

As shown in the rank plot, the top three genre counts are comedy, action, and romance. All three of these genres hover near the mean value of 467 and don't perform better which was expected. Again, there are some points that do stand out, but with such high standard deviations it cannot be definitively said that one genre performs better than another.

The last variable explored was voice actors. Like genres, the cast members, which voice actors would fall under, was voted as a very important factor in deciding what to watch according to the 2016 online survey. Table 4 shows that the range of ranks and scores is much larger for different voice actors than it was for the other variables. The difference between the maximum and minimum ranks is about 324 and the difference in scores is about 0.70. Also, the standard deviations for these two measures are about equal or less than the ranges which was unlike the other variables where generally, the standard deviations were larger than the distance from the minimum and maximum values. A plot of all average ranks and average scores can be seen

in Figure 4.

Table 4. *Summary statistics for the voice actor variable. The values are the min, max, and median values of the average rank or score for a voice actor. The mean column is the mean of the means.*

	Average.rank	Rank.stds	Rank.feature	Average.score	Score.stds	Score.feature
Min	300.138	273.611	Ishikawa, Kaito	7.301	0.759	Inoue, Kazuhiko
Max	624.650	245.889	Morikubo, Shoutarou	8.031	0.561	Yukana
Median	458.143	302.873	Uchiyama, Yumi	7.772	0.586	Fukuen, Misato
Mean	461.711	57.905	NA	7.761	0.116	NA

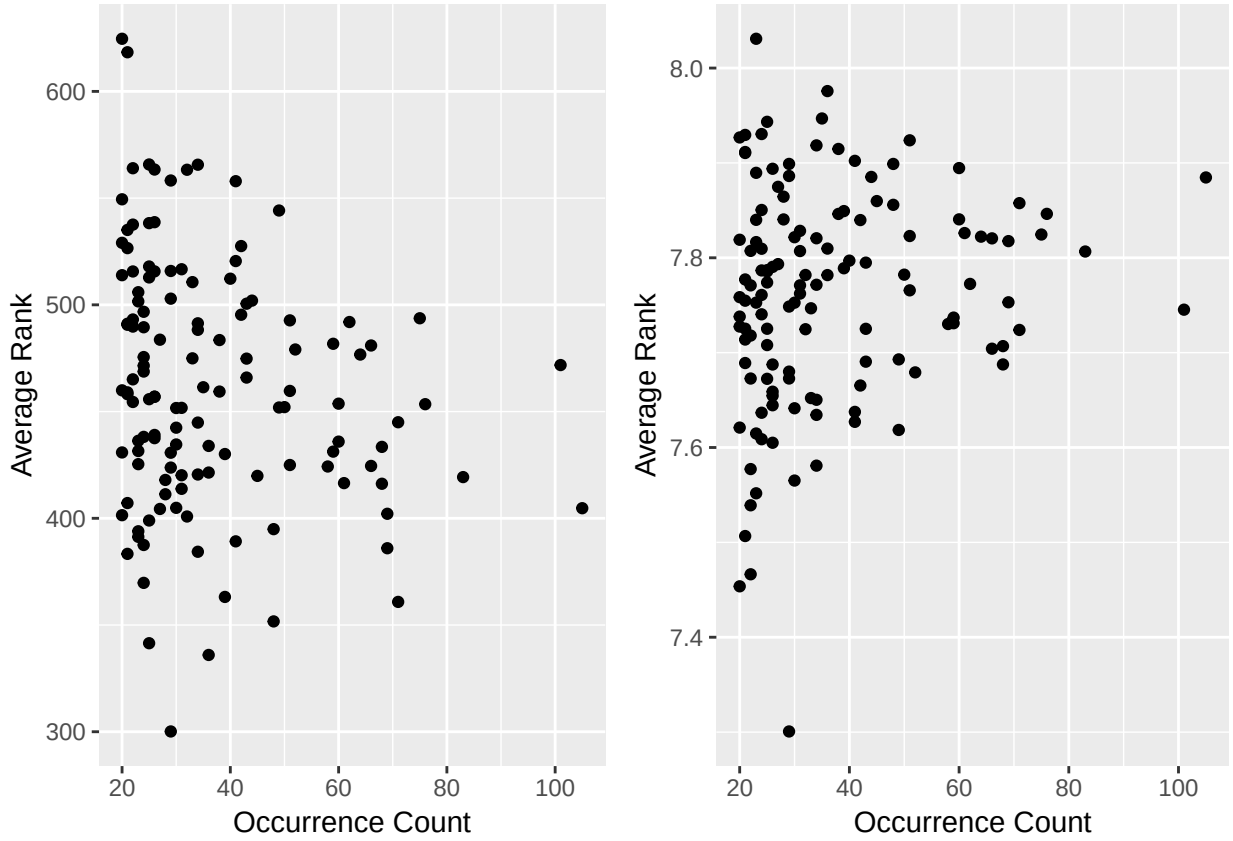

Figure 4. *Average rank vs occurrence count and average score vs occurrence count for voice actor.*

Again, we can see that the range of the average ranks is large when we compare that to the range of the other variables. A similar trend can be seen with the scores. This larger range could mean that the voice actors have a larger affect on which anime is popular. To see the exact names of all the features and their average score and rank, refer to Table A1 in Appendix A. Lastly, it is good to reiterate that when taking into account the large standard deviations of all variables, it would make sense to make an initial judgement that one cannot say the features and their ratings are statistically different from each other.

Conclusion

Data on the popularity and score of an anime was explored to see if the studio that produced it, the type of media it is, the genre it is classified as, or the voice actors in it affect the success of the anime. Each variable had features with large standard deviations so it cannot be said that any feature was statistically significant in affecting the success in this data sample. If one wanted to make a clain regardless, the ratio of the range of the data to the standard deviation of those two points making the range was largest for the voice actors variable. Therefore, voice actors appeared to be the most significant factor. In the future, an

analysis on a larger and more random dataset should be done. This data consisted of the top 950 shows, but with over 15000 shows, a better sample can be obtained.

Bibliography

Research and Markets. "Global Animation Industry Report 2017: The Most Authoritative Global Animation Industry Analysis." GlobeNewswire News Room, "GlobeNewswire", 16 Jan. 2017.

"Japan's Anime Market in 2016 Worth More than 2 Trillion Yen." Translated by The Asahi Shimbun, The Asahi Shimbun, 10 Dec. 2017.

Statista. "Factors Influencing the Decision What to Watch in the U.S. 2016 | Statistic." Statista.

Appendix A: Tables

Table A1. *The average ranks and scores as well as their standard deviations are organized by descending count number for the number of times that voice actor appeared in the dataset. The minimum count number was set to 20.*

Voice.actor	Average.Rank	Rank.std	Average.Score	Score.std
Hanazawa, Kana	404.728	278.958	7.885	0.554
Sakurai, Takahiro	471.790	298.261	7.745	0.540
Sawashiro, Miyuki	419.299	291.641	7.807	0.547
Fukuyama, Jun	453.461	276.923	7.846	0.643
Kugimiya, Rie	493.712	290.655	7.825	0.574
Kaji, Yuki	360.831	295.229	7.858	0.579
Kamiya, Hiroshi	445.000	277.841	7.724	0.656
Miyano, Mamoru	385.970	285.998	7.818	0.761
Ono, Daisuke	402.118	269.525	7.753	0.668
Horie, Yui	433.463	276.002	7.707	0.595
Suwabe, Junichi	416.182	280.140	7.688	0.659
Hayami, Saori	424.530	264.739	7.820	0.544
Kitamura, Eri	480.954	263.940	7.704	0.601
Kayano, Ai	476.710	286.091	7.822	0.541
Noto, Mamiko	491.951	263.883	7.772	0.559
Nakamura, Yuuichi	416.475	293.117	7.826	0.510
Hikasa, Yoko	435.915	256.094	7.841	0.562
Seki, Tomokazu	453.695	277.725	7.895	0.613
Ishida, Akira	481.763	260.895	7.737	0.602
Okamoto, Nobuhiko	431.237	268.068	7.731	0.496
Namikawa, Daisuke	424.263	277.592	7.730	0.556
Sugita, Tomokazu	479.080	267.277	7.679	0.557
Itou, Shizuka	459.667	272.766	7.766	0.656
Sakamoto, Maaya	424.961	267.230	7.823	0.589
Tamura, Yukari	492.729	295.433	7.924	0.579
Toyosaki, Aki	452.080	251.294	7.782	0.508
Fujiwara, Keiji	451.911	275.542	7.693	0.584
Saito, Chiwa	544.191	264.765	7.619	0.548
Kawasumi, Ayako	394.915	269.725	7.899	0.476
Yuuki, Aoi	351.688	286.282	7.856	0.696
Tomatsu, Haruka	419.864	266.575	7.860	0.581
Morikawa, Toshiyuki	501.977	297.238	7.885	0.509
Inoue, Marina	465.930	296.032	7.725	0.594
Katou, Emiri	474.805	295.162	7.690	0.525
Koyasu, Takehito	500.535	286.391	7.795	0.530
Koshimizu, Ami	495.357	296.783	7.665	0.757
Miki, Shinichiro	527.537	285.219	7.840	0.579
Ohara, Sayaka	520.512	294.122	7.638	0.724
Satou, Rina	557.927	261.421	7.627	0.672
Taketatsu, Ayana	389.195	260.093	7.902	0.534
Suzumura, Kenichi	512.250	296.340	7.797	0.575
Hosoya, Yoshimasa	430.103	265.536	7.849	0.591

(continued)

Voice.actor	Average.Rank	Rank.std	Average.Score	Score.std
Matsuoka, Yoshitsugu	363.158	263.332	7.789	0.494
Kobayashi, Yuu	459.368	272.584	7.915	0.647
Nakahara, Mai	483.500	305.958	7.846	0.560
Kimura, Ryouhei	433.914	287.330	7.810	0.629
Uchida, Maaya	421.429	294.832	7.782	0.478
Uchiyama, Kouki	335.971	262.765	7.976	0.554
Sakura, Ayane	461.400	305.913	7.947	0.590
Fukuen, Misato	444.844	264.888	7.772	0.586
Kakihara, Tetsuya	565.647	275.044	7.634	0.660
Kanemoto, Hisako	488.294	253.058	7.581	0.590
Nazuka, Kaori	420.515	275.650	7.918	0.544
Park, Romi	384.324	285.327	7.821	0.615
Shimono, Hiro	491.353	241.606	7.650	0.597
Asumi, Kana	474.935	237.058	7.747	0.614
Mizuki, Nana	510.636	283.702	7.652	0.694
Iguchi, Yuka	400.844	252.196	7.782	0.550
Yoshino, Hiroyuki	563.281	288.352	7.725	0.647
Ise, Mariya	413.806	258.060	7.762	0.472
Koyama, Rikiya	420.161	276.969	7.828	0.471
Nabatame, Hitomi	516.677	260.726	7.807	0.561
Nakata, Jouji	451.742	288.961	7.771	0.480
Hanae, Natsuki	442.448	291.607	7.641	0.491
Hino, Satoshi	451.667	270.026	7.753	0.644
Irino, Miyu	434.586	304.010	7.565	0.993
Komatsu, Mikako	404.900	236.702	7.822	0.566
Inoue, Kazuhiko	502.897	275.621	7.301	0.759
Inoue, Kikuko	558.250	263.809	7.749	0.377
Ishikawa, Kaito	300.138	273.611	7.886	0.606
Itou, Kanae	430.724	255.682	7.673	0.421
Nakai, Kazuya	423.759	274.357	7.899	0.518
Okiayu, Ryotaro	515.862	303.069	7.680	0.618
Touyama, Nao	417.963	269.237	7.864	0.460
Yukino, Satsuki	411.261	240.117	7.840	0.616
Orikasa, Fumiko	404.370	324.771	7.875	0.843
Tachiki, Fumihiko	483.667	264.866	7.793	0.622
Hatano, Wataru	563.385	279.093	7.605	0.761
Hayashibara, Megumi	457.000	247.822	7.688	0.467
Hidaka, Rina	439.038	280.141	7.659	0.678
Ishizuka, Unshou	538.731	269.934	7.655	0.762
Konishi, Katsuyuki	456.923	291.118	7.894	0.550
Mizuhashi, Kaori	515.682	281.707	7.645	0.574
Sugiyama, Noriaki	437.538	287.506	7.790	0.510
Endou, Aya	538.280	305.962	7.672	0.536
Hoshi, Soichiro	565.760	280.156	7.774	0.648
Osaka, Ryota	455.792	296.181	7.943	0.419
Suzuki, Tatsuhisa	341.478	240.063	7.787	0.639
Takagaki, Ayahi	398.960	236.720	7.786	0.532

(continued)

Voice.actor	Average.Rank	Rank.std	Average.Score	Score.std
Taniyama, Kishou	512.840	317.754	7.725	0.711
Yusa, Kouji	517.960	264.330	7.708	0.554
Akasaki, Chinatsu	471.500	305.606	7.850	0.662
Amamiya, Sora	369.750	300.237	7.787	0.505
Hiyama, Nobuyuki	489.500	298.248	7.761	0.516
Kaida, Yuuko	496.750	270.700	7.637	0.600
Maeno, Tomoaki	475.542	297.317	7.740	0.446
Sakaguchi, Daisuke	387.500	276.805	7.930	0.455
Yahagi, Sayuri	468.750	238.514	7.609	0.558
Yamadera, Kouichi	438.143	287.141	7.810	0.540
Chihara, Minori	501.609	289.614	7.552	0.733
Hirata, Hiroaki	431.522	318.174	7.817	0.551
Sato, Satomi	436.304	325.186	7.890	0.483
Taneda, Risa	393.913	280.162	7.840	0.385
Toyoguchi, Megumi	425.364	263.823	7.753	0.510
Tsuda, Kenjirou	505.913	232.431	7.615	0.615
Yukana	391.304	280.997	8.031	0.561
Hirakawa, Daisuke	454.500	221.827	7.466	1.085
Hirano, Aya	465.091	328.391	7.539	1.069
Minagawa, Junko	489.818	315.253	7.771	0.766
Takeuchi, Junko	537.591	300.496	7.577	0.695
Tanaka, Rie	564.000	238.565	7.718	0.746
Ueda, Kana	515.636	296.774	7.807	0.519
Watanabe, Akeno	493.136	301.923	7.673	0.647
Fujimura, Ayumi	618.333	239.555	7.689	0.480
Minase, Inori	491.143	278.003	7.714	0.477
Ogura, Yui	526.571	277.401	7.507	0.479
Ono, Kensho	383.333	299.952	7.725	0.620
Ootsuka, Akio	459.000	287.644	7.777	0.490
Toyonaga, Toshiyuki	407.143	255.880	7.755	0.502
Uchiyama, Yumi	458.143	302.873	7.930	0.521
Ueda, Yuji	535.095	288.741	7.910	0.476
Uesaka, Sumire	490.667	273.835	7.912	0.570
Asanuma, Shintaro	513.895	246.843	7.454	0.588
Hirohashi, Ryou	549.450	293.060	7.758	0.604
Morikubo, Shoutarou	624.650	245.889	7.621	0.599
Nojima, Kenji	459.950	305.611	7.819	0.696
Ono, Yuuki	401.450	278.395	7.728	0.510
Shimazaki, Nobunaga	529.053	336.288	7.927	0.759
Tanaka, Atsuko	430.850	234.824	7.738	0.480

Table A2. *The average ranks and scores as well as their standard deviations are organized by descending count number for the number of times that genre appeared in the dataset. The minimum count number was set to 20.*

Genre	Average.Rank	Rank.std	Average.Score	Score.std
Comedy	481.945	276.320	7.741	0.603
Action	456.657	279.677	7.730	0.631
Romance	457.519	268.290	7.772	0.572
School	445.069	272.674	7.736	0.568
Supernatural	441.912	270.695	7.756	0.594
Drama	457.841	281.866	7.798	0.613
Fantasy	460.738	282.704	7.845	0.562
Shounen	432.400	280.084	7.764	0.618
Sci-Fi	493.124	284.783	7.802	0.615
Adventure	439.263	286.269	7.752	0.609
Slice of Life	466.736	267.434	7.764	0.587
Mystery	456.483	283.583	7.808	0.594
Ecchi	470.039	248.377	7.717	0.588
Seinen	506.540	266.460	7.722	0.616
Harem	460.768	246.725	7.691	0.548
Super Power	429.980	292.133	7.724	0.627
Magic	415.463	268.586	7.874	0.511
Psychological	451.354	304.306	7.805	0.562
Military	496.414	314.971	7.823	0.650
Historical	503.333	270.752	7.846	0.568
Horror	453.019	296.565	7.725	0.608
Shoujo	452.250	246.845	7.755	0.577
Demons	457.149	275.202	7.773	0.622
Mecha	512.600	302.883	7.737	0.595
Thriller	451.400	301.860	7.860	0.648
Parody	522.541	299.494	7.720	0.693
Game	425.941	276.658	7.752	0.522
Sports	527.688	262.060	7.463	0.942
Vampire	390.531	249.677	7.734	0.629
Martial Arts	528.000	249.931	7.573	0.490
Music	521.667	258.242	7.806	0.617

Table A3. *The average ranks and scores as well as their standard deviations are organized by descending count number for the number of times that studio appeared in the dataset. The minimum count number was set to 20.*

Studio	Average.Rank	Rank.std	Average.Score	Score.std
Madhouse	443.790	271.589	7.868	0.604
A-1 Pictures	440.356	293.946	7.842	0.548
J.C.Staff	450.673	270.456	7.729	0.594
Bones	345.462	280.302	7.819	0.650
Studio Deen	489.500	267.408	7.662	0.611
Kyoto Animation	320.118	261.003	7.989	0.583
Production I.G	445.000	295.804	7.865	0.578
Studio Pierrot	487.031	292.228	7.718	0.542
Shaft	461.517	255.473	7.768	0.567
Sunrise	521.286	324.885	7.866	0.634
Brain's Base	475.038	271.592	7.754	0.674
Silver Link.	497.609	262.042	7.542	0.581
Xebec	592.478	217.162	7.607	0.372
Toei Animation	563.000	282.989	7.873	0.702
Gonzo	497.571	259.649	7.496	0.471