Analysis Report

ANALYSIS OF TOP 200 YOUTUBERS

USING ONE-WAY ANOVA AND LINEAR REGRESSION

Quyen Thuc Nguyen

30 November 2022

Table of contents

List of figures

List of tables

I.   Introduction

Youtube is the second-most popular online video sharing and social media platform worldwide with the number of monthly active users being 2.5 billion. With user-generated contents, Youtube has undoubtedly created an innovative alternate entertainment industry and the most famous channels have the power to influence an individual's intention to the mind-set of a whole community on a global scale. The popularity of a Youtube channel can be determined by the number of subscribers, likes, views, comments, or audience engagement rate, which are significantly important characteristics to study further.

Several authors have also conducted research on Youtube's popularity and its influencing factors of different characteristics. In research from Fan and Lian, the results analysed through linear regression and artificial neural network show users' remarkable tendency of watching films and listening to music and concentration in India and the U.S. (2022). Another study from Yang et al. reveals that shorter videos attract more views and social endorsement cues, and engagement rate are positively related (2022), meanwhile Velho et al. stated that view forecasts could be influenced by the number of likes, the efficiency of video production, video age and formats (2020) and Lopezosa et al. believed viewership and the number of subscribers were highly and positively correlated (2019). On a relevant theme, as per Bärtl's research on '*Youtube channels, uploads, and views*' (2018), even though older channels have a significant likelihood of attracting a sizable audience, younger channels always have a slim chance to succeed rapidly, depending on choosing category wisely.

In most of the journals with a related research topic on Youtube channels' popularity, very few papers investigate the number of total likes per channel. This report will conduct a statistical analysis on the Top 200 Youtubers to unravel whether different main video categories have different averages of boost index as well as examine influencing factors of a channel's total likes.

II.   Methodologies
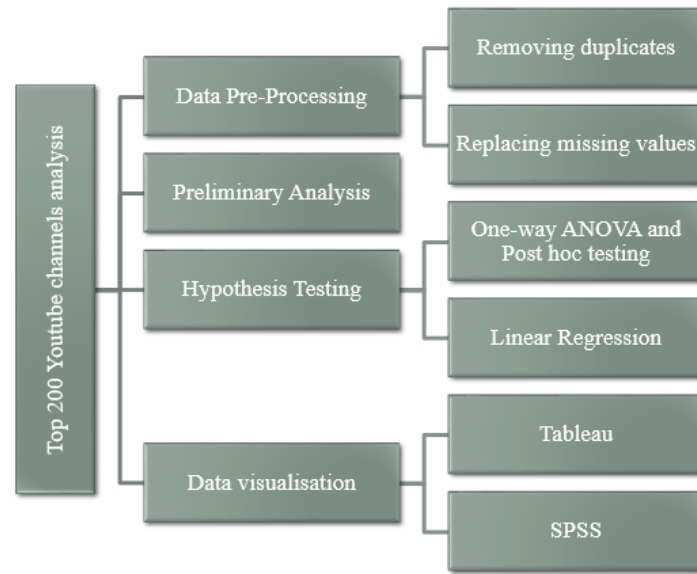
Figure 1 Analysis process of the Top 200 Youtubers dataset

1. Data pre-processing

The original Top 200 Youtubers dataset records each Youtuber as one observation. 875 observations indicating duplicates which are removed to remain 200 unique observations.

After duplicates are eliminated, missing values (Appendix A) are replaced. Channels with missing values in 'Category' are classified based on the category of other channels in the same main video categories and main topics. Missing values in 'Country', and other continuous variables cannot be interpreted and remain null as no other variables can be used as references in the current dataset.

2. Descriptive statistics analysis

This section conducts an overview to summarise outstanding features of the dataset, as well as univariate descriptive statistics to study the measures of central tendency, dispersion and distribution and identify outliers of some main variables.

3. Explorative analysis

In this section, hypothesis testing is conducted using One-way ANOVA and Linear regression with a significance level of 0.05.

- Hypothesis I: Youtube main video categories have different averages of normalised boost index.

SPSS ANOVA with Tukey post hoc test is used to examine this hypothesis. Required assumptions to perform ANOVA test:

- A categorical variable and a continuous variable are assessed, in which the continuous variable is normalised.
- More than 3 different groups across the same measure with independent observations are compared.
- The Tukey post hoc test is conducted to examine which pairs of main video categories have different average of total likes.

As SPSS post hoc test is not performed when at least one group has fewer than two cases, main video categories with only one observation are combined into other related main video categories (See Appendix B for details).

- Hypothesis II: There is relationship between the number of total likes and the numbers of total views, followers, and comments average.

Pearson's correlation coefficients and linear regression analysis are used to test this hypothesis for the following reasons:

- The linear correlations between independently observed and continuous variables are examined.
- The purpose is to determine how much of the variation in the total likes can be explained by the total views, followers, and comments average.

III.   Preliminary analysis

1. Overview of the dataset

The dataset consists of 200 observations with 8 nominal and 14 continuous variables. Details of each variable are as follows:

Table 1 Description of variables

Variables	Levels of measurement	Descriptions
Country	Nominal	Name of the countries
Channel Name	Nominal	Name of the channels
Category	Nominal	Category of the channels
Main Video Category	Nominal	Main video categories
username	Nominal	Username of the channels
followers	Scale	Number of followers
Main topic	Nominal	Mainly discussed topics
More topics	Nominal	Other discussed topics
Likes	Scale	Total likes
Boost Index	Scale	Boost index value
Engagement Rate	Scale	Rate of engagement with the users
Engagement Rate 60days	Scale	Rate of engagement with the users for 60 days
Views	Scale	Total views
Views Avg.	Scale	Average number of views
Avg. 1 Day	Scale	Average number of views for one day
Avg. 3 Day	Scale	Average number of views for 3 days
Avg. 7 Day	Scale	Average number of views for 7 days
Avg. 14 Day	Scale	Average number of views for 14 days
Avg. 30 day	Scale	Average number of views for 30 days
Avg. 60 day	Scale	Average number of views for 60 days
Comments Avg	Scale	Average number of comments
Youtube Link	Nominal	Link of the channel

Overall, most Youtubers within top 200 are in the U.S. and India. The majority of channels produce contents in 'Gaming & Apps', 'Music', 'Entertainment' and 'Lifestyles' across different categories and subgenres (Fig. 2). Moreover, statistics show that audiences in the U.S. mostly watch 'Entertainment' and 'Education' video, meanwhile in India, 'Music' is preferred over 'Entertainment', being the most viewed main video category in India.
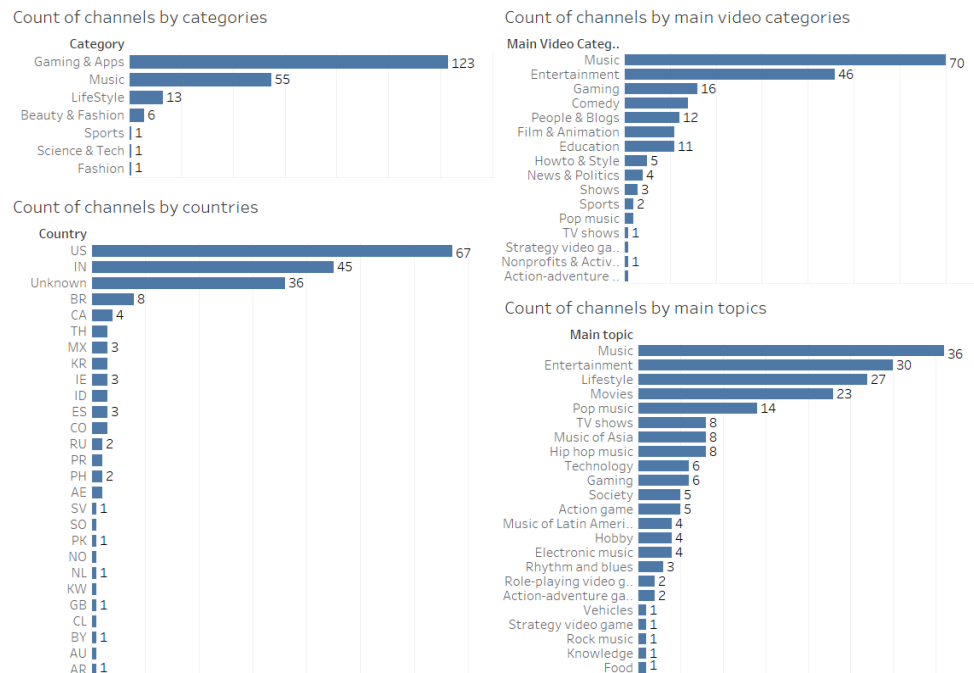
Figure 2 Count of channels by countries, categories, main video categories and main topics

Figure 3 Number of Views by Countries and Main video categories

From Fig. 4, some channels rank in top 10 by total likes, total views, and followers. For instance, T-Series is leading in both total views and subscribers, and MrBeast, PewDiePie, SET India, and WWE are also the familiar names in these three categories. On the contrary, there are no channels in the top 10 by engagement rate classified in other 3 categories. This might indicate that a channel with a high number of likes possibly has high numbers of total views and followers as well, from which Hypothesis II is developed and testing results are discussed in Explorative analysis section.

Figure 4 Top 10 channels by total likes, views, followers, and engagement rate

2. Descriptive statistics analysis

Table 2 Descriptive statistics of Followers, Likes, Boost Index, Engagement rate, Engagement rate in 60days, Views and Views average

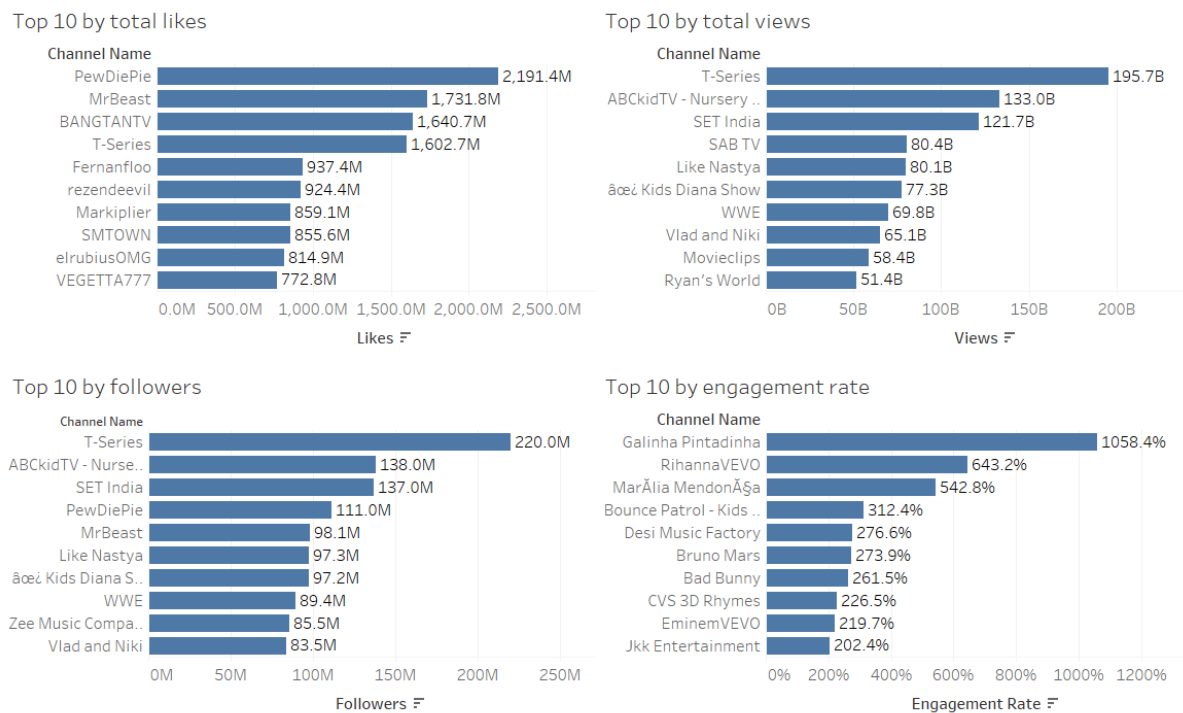|		Followers	Likes	Boost Index	Engagement Rate (%)	Engagement Rate 60days (%)	Views	Views Average
Descriptive Statistics								
N	Valid	200	198	200	199	199	199	199
	Missing	0	2	0	1	1	1	1
Mean		39,477,500	183,132,950.564	64.975	46.065	7.700	20,810,409,566.271	16,639,304.402
Median		32,550,000	72,701,577.840	70.000	12.026	2.563	16,380,874,723.000	2,911,652.646
Mode		27,800,000	790,616[a]	75[a]	0[a]	0[a]	994,418,044[a]	681[a]
Std. Deviation		22,622,785	309,438,973.290	16.487	107.203	17.762	21,251,818,422.316	40,821,295.266
Variance		5.1179E+14	9.57525E+16	272	11,492	315	4.5164E+20	1.66638E+15
Skewness		4.013	3.543	-1.504	5.984	5.521	4.581	6.179
Std. Error of Skewness		0.172	0.173	0.172	0.172	0.172	0.172	0.172
Kurtosis		23.318	15.524	2.188	46.436	36.666	28.870	52.502
Std. Error of Kurtosis		0.342	0.344	0.342	0.343	0.343	0.343	0.343
Range		196,000,000	2,190,614,926	87.000	1,058.382	151.902	194,666,326,372	423,922,818
Minimum		24,000,000	790,616	1.000	0.026	0.003	994,418,044	681
Maximum		220,000,000	2,191,405,542	88.000	1,058.408	151.904	195,660,744,416	423,923,499
Percentiles	25	27,400,000	25,056,792	60.000	2.748	0.617	11,507,336,712	681,340
	50	32,550,000	72,701,578	70.000	12.026	2.563	16,380,874,723	2,911,653
	75	42,225,000	190,414,331	76.750	46.615	6.535	23,356,797,129	14,483,462

a. Multiple modes exist. The smallest value is shown

Table 3 Descriptive statistics of Average views in 1 day, 3 days, 7 days, 14 days, 30 days, and 60 days and Comments average.

		Average views in 1 Day	Average views in 3 Day	Average views in 7 Day	Average views in 14 Day	Average views in 30 day	Average views in 60 day	Comments Average
Descriptive Statistics								
N	Valid	89	125	144	154	163	172	185
	Missing	111	75	56	46	37	28	15
Mean		217,398.50	560,598.52	1,166,708.70	2,061,996.49	2,566,732.19	3,156,209.41	14,772.42
Median		44,195.00	122,221.78	239,976.39	360,166.75	568,718.40	861,749.71	1,771.24
Mode		8,062	703[a]	1,514[a]	2,549[a]	2,509[a]	2,553[a]	0[a]
Std. Deviation		492,052.182	1,063,116.802	2,937,921.150	6,526,750.418	7,006,757.997	7,476,551.364	30,454.564
Variance		2.42115E+11	1.13022E+12	8.63138E+12	4.25985E+13	4.90947E+13	5.58988E+13	927480490
Skewness		4.361	3.349	7.061	7.883	6.489	4.806	3.242
Std. Error of Skewness		0.255	0.217	0.202	0.195	0.190	0.185	0.179
Kurtosis		23.240	12.981	64.861	74.326	52.799	26.682	11.905
Std. Error of Kurtosis		0.506	0.430	0.401	0.389	0.378	0.368	0.355
Range		3,472,182	6,595,298	29,939,507	68,774,767	68,774,807	53,832,675	199,523
Minimum		456	703	1,514	2,549	2,509	2,553	0
Maximum		3,472,638	6,596,001	29,941,021	68,777,316	68,777,316	53,835,228	199,523
Percentiles	25	14,299.00	31,670.86	71,308.18	100,137.42	161,278.94	210,586.01	86.68
	50	44,195.00	122,221.78	239,976.39	360,166.75	568,718.40	861,749.71	1,771.24
	75	148,595.92	646,941.75	913,448.13	1,328,607.18	2,008,041.57	2,691,784.28	13,537.53

a. Multiple modes exist. The smallest value is shown

Overall, all continuous variables within the dataset are not normally distributed as the means, medians, and modes are not equal for each variable (See Table 2 & 3) and they all have a considerable numbers of outliers.

Except for Boost index with a left-skewed distribution (negative skewness of -1.504, Table 3), all other continuous variables have right-skewed distribution (positive skewness). This implies that within the Top 200, the outliers are the most popular Youtubers as only a few of them have exponentially greater values in channels statistical data. Additionally, all considered variables have kurtosis values of greater than 1, meaning their distributions are extremely peaked.

Besides, their high standard deviations and variances reveals that the data points spread out substantially far away from the means.

The bellows are the histograms and boxplots of variables considered for hypothesis testing.

- Boost Index
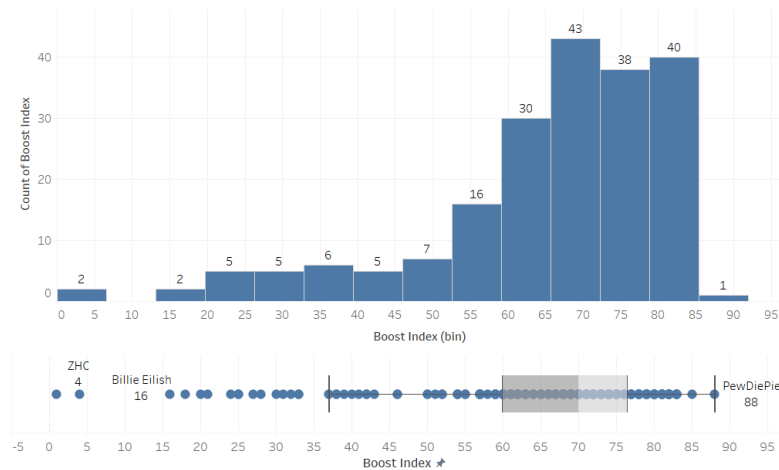
Figure 5 Histograms and boxplots of Boost index

As discussed above, the histogram for 'Boost index' negatively skewed. Most channels in the top 200 have the boost index values range from 55 to 85. Some of the outliers are PewDiePie, ZHC and Billie Eilish channels with the boost index values of 88, 16, and 4 respectively.

- Followers
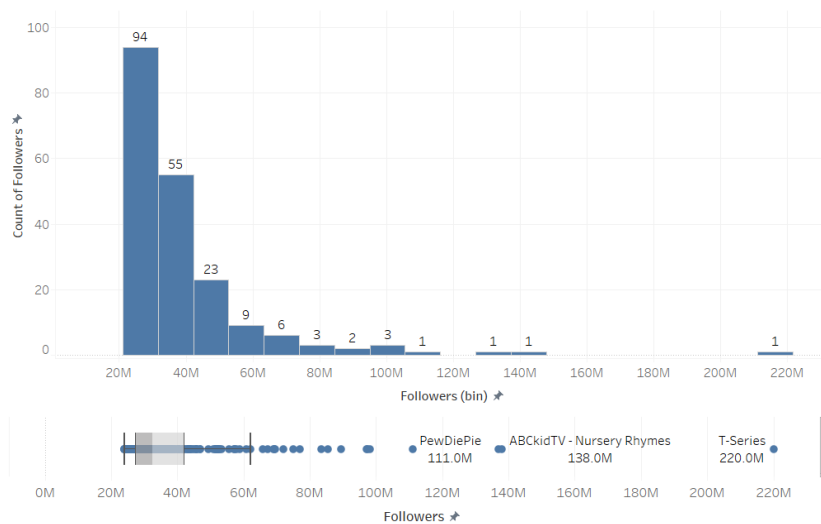
Figure 6 Histogram and boxplot of Followers

The majority of channels within top 200 have the numbers of followers vary from 20 milluon to 50 million. Only a few most popular channels such as T-Series, ABCkid TV, and PewDiePie have significantly more followers (Fig. 6).

- Likes
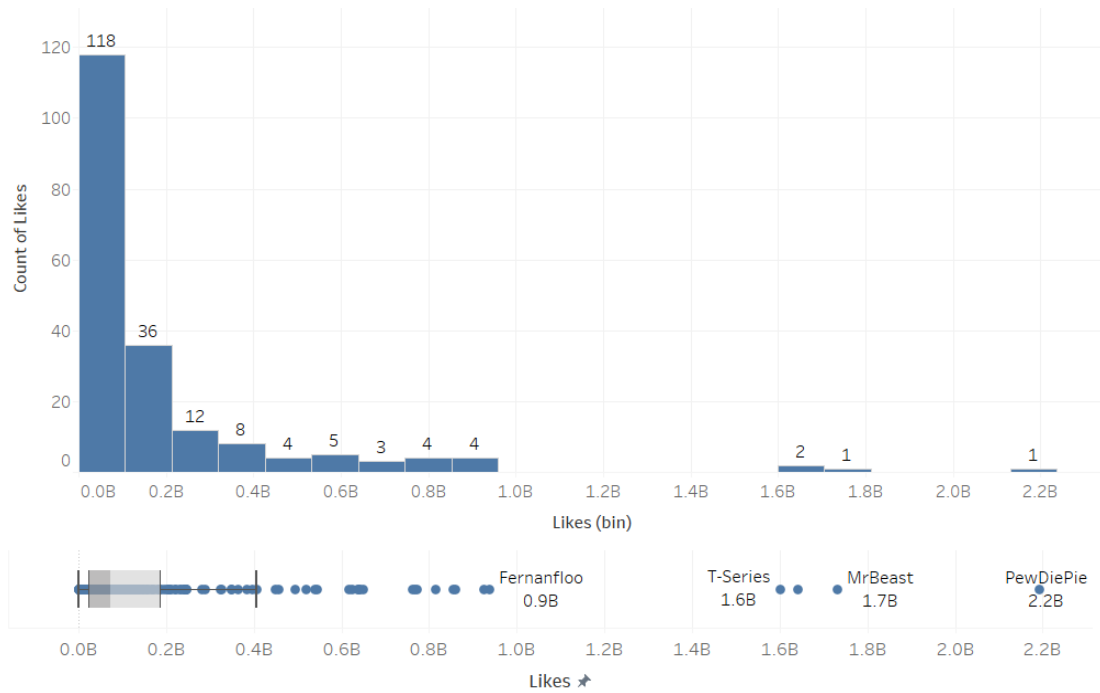
Figure 7 Histogram and boxplot of total likes

Most channels in the top 200 have total likes of less than 200 million, meanwhile some outlying channels have the exponentially greater total likes, especially top 10 channels (Fig. 4).

- Views
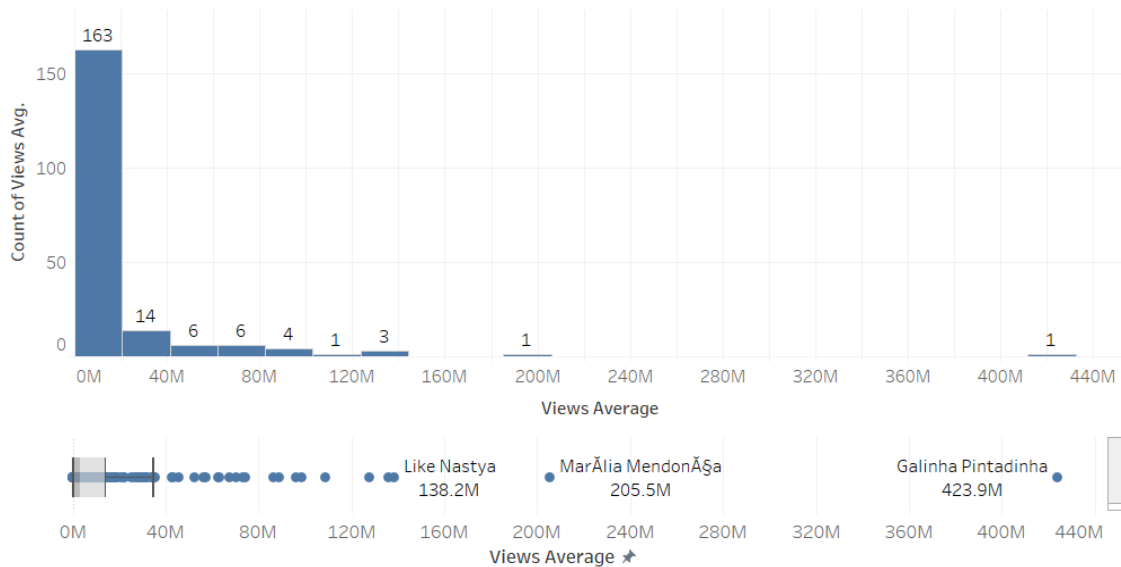
Figure 8 Histogram and boxplots of Total views

177 out of 200 channels in the top 200 have the viewership of less than 40 million counts. Likewise, the outliers fall within the top 10 channels by view as shown in Fig. 4.

9
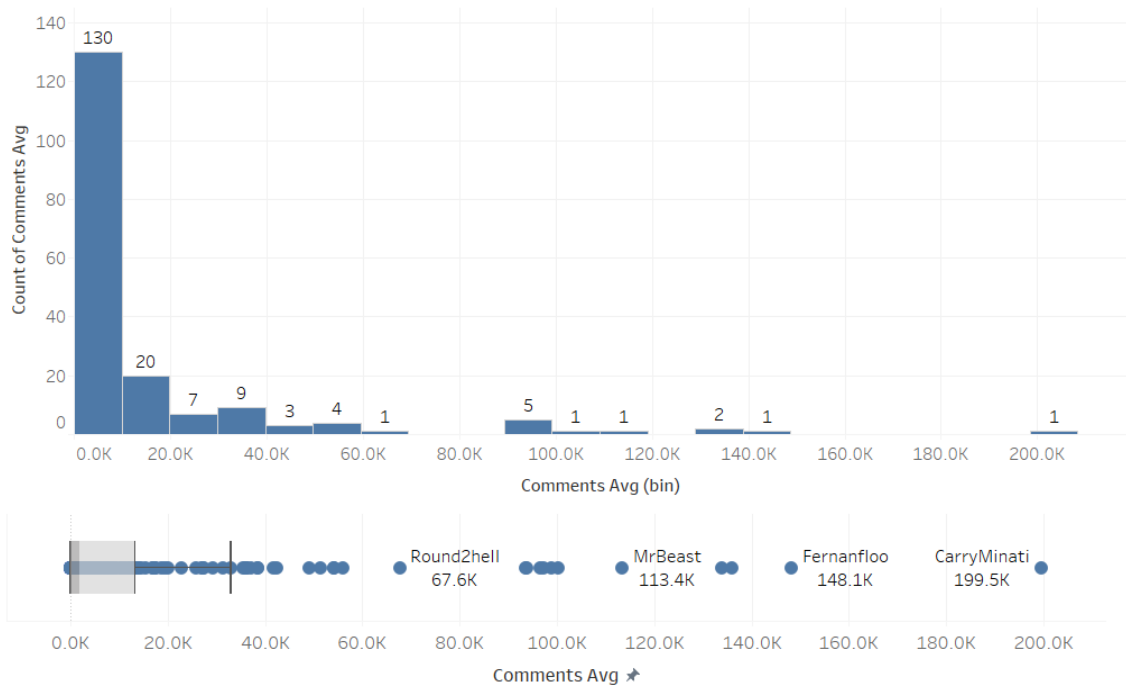
- Comments average

Figure 9 Histogram and boxplot of Comments average

75% channels of the top 200 have the numbers of comments average of less than 20,000. Surprisingly, except for MrBeast, channels that have the highest comments average are not the outliers in other categories.

IV. Explorative analysis

1. Hypothesis I

The developed hypotheses are:

H_0: Youtube main video categories have equal averages of normalised boost index.

H_1: Youtube main video categories have different averages of normalised boost index.

Significant level $\alpha = 0.05$.

The One-way ANOVA test results from SPSS are as follows:

Table 4 Descriptives table of boost index by main video categories

Descriptives

nomalisedBI

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Music	70	-.1077	1.02913	.12300	-.3530	.1377	-2.97	1.09
Education	11	-.0371	.70277	.21189	-.5092	.4350	-1.58	.79
Shows	4	.5019	.78517	.39259	-.7475	1.7513	-.67	1.03
Gaming	18	.2846	1.03246	.24335	-.2289	.7980	-2.73	1.40
Entertainment	46	.1360	.89039	.13128	-.1284	.4004	-3.88	1.09
People & Blogs	12	-.3270	1.19148	.34395	-1.0840	.4300	-2.42	1.03
Sports	2	.9720	.08578	.06065	.2013	1.7426	.91	1.03
Howto & Styles	5	-.4473	1.86579	.83441	-2.7640	1.8694	-3.70	.97
Film & Animation	11	-.2852	.88190	.26590	-.8777	.3073	-2.12	.73
News & Politics	5	.3048	.61854	.27662	-.4632	1.0728	-.36	1.21
Pop music	2	-2.2123	.38599	.27294	-5.6803	1.2557	-2.49	-1.94
Comedy	14	.3438	.63328	.16925	-.0219	.7094	-1.15	1.09
Total	200	.0000	1.00000	.07071	-.1394	.1394	-3.88	1.40
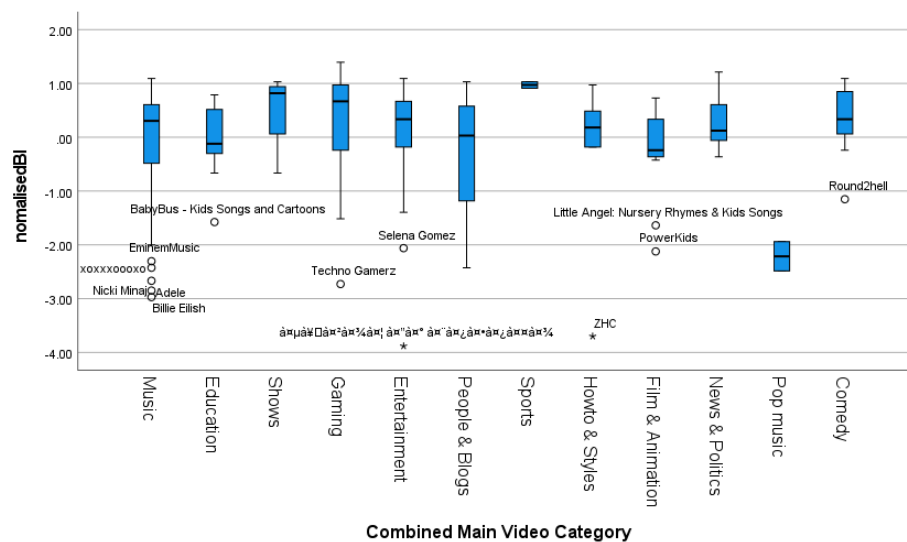
Figure 10 Boxplots of likes by main video category

As seen from Table 4 and Fig. 11, 'Sports' have the highest mean of normalised boost index value, followed by that of 'Shows', 'News & Politics' and 'Comedy'.

11

Tests of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
nomalisedBI	Based on Mean	1.479	11	188	.142
	Based on Median	.731	11	188	.708
	Based on Median and with adjusted df	.731	11	137.746	.707
	Based on trimmed mean	1.237	11	188	.265

Tests of homogeneity of variances are conducted to examine whether the distributions of boost index values across main video categories are comparable, in other words, homogeneous. The p-value of 0.142, which is greater than the significance level of 0.05, implies observations between main categories have similar variances, and thus the assumption of homogeneity of variance is met and ANOVA can be performed.

ANOVA

nomalisedBI

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	21.118	11	1.920	2.029	.028
Within Groups	177.882	188	.946		
Total	199.000	199			

Results of one-way ANOVE (Table 4.) indicates a statistically significant difference between main video categories with $F(11,188) = 2.029$ and p-value equals 0.028. Since the p-value is less than the significance level of 0.05, the null hypothesis is rejected. It can be concluded that the averages of normalised boost index values across main video categories are different.

Multiple Comparisons Summary: Pairs with significant means difference						
Dependent Variable: Nomalised Boost index Tukey HSD						
(I) Combined Main Video Category		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Gaming	Pop music	2.49686	0.72502	0.033	0.0972	4.8965
Entertainment	Pop music	2.34831	0.70261	0.045	0.0228	4.6738
Pop music	Gaming	-2.49686	0.72502	0.033	-4.8965	-0.0972
Pop music	Entertainment	-2.34831	0.70261	0.045	-4.6738	-0.0228
Pop music	Comedy	-2.55607	0.73531	0.030	-4.9898	-0.1224
Comedy	Pop music	2.55607	0.73531	0.030	0.1224	4.9898

From the summary of the Tukey post hoc test results (See Table 7), 'Pop music' has the significantly means different of boost index in pair comparisons with other 3 groups of

'Gaming', 'Entertainment', and 'Comedy'. There is no statistically significant means difference between other variables (See Appendix C for completed results).

2. Hypothesis II

The developed hypotheses are:

H_0: There is no relationship between the number of total likes and the numbers of total views, followers and comments average.

H_2: There is relationship between the number of total likes and the numbers of total views, followers and comments average.

Significant level $\alpha = 0.05$.

Pearson's correlation coefficients (r-value) between each 2 variables of the number of total likes, total views, followers and comments average are calculated as follows:

Table 8 Correlations between Likes, Views, Followers, and Comments average

Correlations

		Followers	Likes	Views
Likes	Pearson Correlation	.462		
	Sig. (2-tailed)	<.001		
	N	198		
Views	Pearson Correlation	.829	.187	
	Sig. (2-tailed)	<.001	.008	
	N	199	198	
Comments average	Pearson Correlation	.053	.420	-.194
	Sig. (2-tailed)	.477	<.001	.008
	N	185	185	185

From table 6, the r-values are significant and positive between 'Likes' and 'Followers' (0.462), 'Likes' and 'Views' (0.187), and 'Likes' and 'Comments average' (0.420). The p-values of those pairs' correlation coefficients are less than 0.01, implying strong evidences for positive linear relationships. Hence, the null hypothesis is rejected.

Linear regression models with 'Likes' as dependent variables ($\hat{y}$) and independent variables being 'Followers' (x_1), 'Views' (x_2), 'Comment average' (x_3). Model 1 has all considered factors, however 'Followers' and 'Views' have moderate multicollinearity as their VIFs are between 1 and 5 (see Table 9), other models are developed to compare for the fittest. Summary of developed linear regression models are as follows:

13

Table 9 Models summary

Models Summary

Model	Explanatory variables	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	Comments average, Followers, Views	64.3%	41.4%	40.4%	245,550,705.6
2	Comments average, Followers	61.3%	37.6%	36.9%	252,516,975.4
3	Followers, Views	58.0%	33.6%	33.0%	253,368,835.1
4	Comments average, Views	50.5%	25.5%	24.7%	275,994,642.6
5	Followers	46.2%	21.3%	20.9%	275,202,473.1
6	Comments average	42.0%	17.6%	17.2%	289,379,664.4
7	Views	18.7%	3.5%	3.0%	304,750,618.8

Table 10 Equations of developed linear regression models

Models	Equations
1	$Likes\ (\hat{y}) = 10.656 * Followers(x_1) - 0.005 * Views(x_2) + 3,208.06 * Comments(x_3) - 162,951,062.741$
2	$Likes\ (\hat{y}) = 6.258 * Followers(x_1) + 4,140.618 * Comments(x_3) - 116,741,272.574$
3	$Likes\ (\hat{y}) = 13.399 * Followers(x_1) - 0.009 * Views(x_2) - 156,815,261.036$
4	$Likes\ (\hat{y}) = 0.004 * Views(x_2) + 4,963.983 * Comments(x_3) + 31,893,554.372$
5	$Likes\ (\hat{y}) = 6.286 * Followers(x_1)$
6	$Likes\ (\hat{y}) = 4,386.502 * Comments(x_3) + 126,946,800.725$
7	$Likes\ (\hat{y}) = 0.003 * Views(x_2) + 126,648,705.582$

Table 11 Coefficients of Model 1

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	95.0% Confidence Interval for B Lower Bound	95.0% Confidence Interval for B Upper Bound	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	-162951062.74	39,393,237.474		-4.137	<.001	-240,680,106.677	-85,222,018.806		
	Followers	10.656	1.523	.762	6.995	<.001	7.650	13.662	.273	3.666
	Views	-.006	.002	-.376	-3.387	<.001	-.009	-.002	.263	3.798
	Comments average	3,208.060	655.815	.307	4.892	<.001	1,914.035	4,502.086	.821	1.217

a. Dependent Variable: Likes

From the ANOVA result tables generated by SPSS, all developed models have significant p-values of less than 0.001. Coefficients of independent variables in each equation illustrated in Table 7 are also significant as their p-values are less than 0.05 (See Appendix F for detailed results of each model).

The most significantly appropriate model so far is the Model 1 with the highest R^2 and adjusted R^2. These statistics reveal that approximately 40% of the variation in the number of total likes can be explained with the numbers of comments average, followers, and views.

A 3D scatterplot for linear regression model 1 are drafted below with Likes on the Y-axis, Views and Followers on the X and Z axis, respectively. Colour scale of Comments average visualises that the lower the values, the darker the colour and deeper the data points and the higher the values, the brighter the colour and closer the data points. From Fig. 12, heteroskedasticity is clearly demonstrated as data points are not equally scattered. In other words, the variance of residuals is not constant across observations.
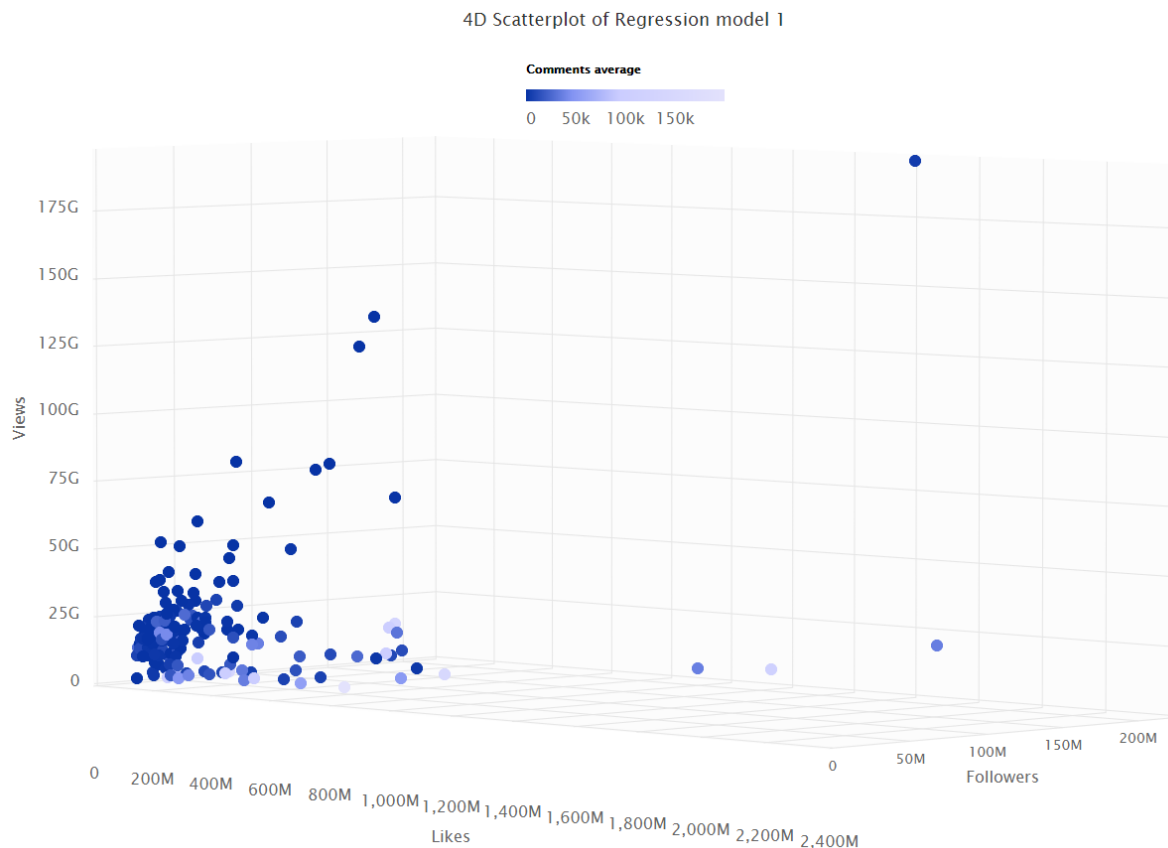
Figure 11 Scatterplot of linear regression model 1

V. Limitations

This study consists of some limitations. All variables have remarkable numbers of outliers, of which presence could exaggerate error rates and significant distortions of parameter and statistics estimates when conducting either parametric or nonparametric hypothesis testing. As a consequence, some parameter estimates are more susceptible to outliers than others, specifically the mean and least squares estimations. Moreover, as shown in Fig. 12, outliers result in heteroskedasticity in the developed linear regression model, which could affect prediction results from this model. However, the aim of the analysis is not forecasting, thus heteroskedasticity is only observed and not adjusted.

VI. Conclusion

The top 200 Youtubers dataset reveals the major audience preferences are videos about gaming, music and entertainment, as well as the spatial spread of subscribers mostly focuses in the U.S. and India. The study of Fan and Lian also supports the same premise. Additionally, the most popular Youtube channels are T-Series, PewDiePie, and MrBeast as they are all in the top 10 by total likes, views, and followers.

Furthermore, the explorative analysis results show that Youtube main video categories have different averages of normalised boost index, specifically between 'Pop music' and 'Gaming', 'Entertainment', and 'Comedy'. Last but not least, 40% of changes in a channel's total likes can be explained with its numbers of followers, views, and comments average.

APPENDICES

Appendix A – Summary of missing values

Variables	Number of valid values	Number of missing values	Faction of missing values
Channel name	200	0	0%
Username	200	0	0%
Country	162	38	23%
Category	164	36	22%
Main Video Category	199	1	1%
Main topic	199	1	1%
More topics	200	0	0%
Followers	200	0	0%
Likes	198	2	1%
Boost Index	200	0	0%
Engagement Rate	199	1	1%
Engagement Rate 60days	199	1	1%
Views	199	1	1%
Views Average	199	1	1%
Average views in 1 Day	89	111	125%
Average views in 3 Day	125	75	60%
Average views in 7 Day	144	56	39%
Average views in 14 Day	154	46	30%
Average views in 30 day	163	37	23%
Average views in 60 day	172	28	16%
Comments Average	185	15	8%
Youtube Link	200	0	0%

APPENDICES

Appendix B – Combination in main video categories for One-way ANOVA and Tukey post hoc test

Before combining		After combining	
Main Video Category	**Number of cases**	**Main Video Category**	**Number of cases**
Shows	3	Shows	4
TV shows	1		
Gaming	16	Gaming	18
Action-adventure game	1		
Strategy video game	1		
News & Politics	4	News & Politics	5
Nonprofits & Activism	1		
Entertainment	46	Entertainment	46
People & Blogs	12	People & Blogs	12
Sports	2	Sports	2
Howto & Styles	5	Howto & Styles	5
Film & Animation	11	Film & Animation	11
Pop music	2	Pop music	2
Comedy	14	Comedy	14
Music	70	Music	70
Education	11	Education	11

APPENDICES

Appendix C – Tukey post hoc test: Multiple comparisons table

<div align="center">

Multiple Comparisons

</div>

Dependent Variable: Nomalised Boost index

Tukey HSD

(I) Combined Main Video Category	(J) Combined Main Video Category	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Music	Education	-.07058	.31549	1.000	-1.1148	.9736
	Shows	-.60956	.50006	.987	-2.2646	1.0455
	Gaming	-.39222	.25707	.932	-1.2430	.4586
	Entertainment	-.24366	.18462	.976	-.8547	.3674
	People & Blogs	.21936	.30392	1.000	-.7865	1.2252
	Sports	-1.07961	.69757	.925	-3.3884	1.2292
	Howto & Styles	.33965	.45028	1.000	-1.1507	1.8300
	Film & Animation	.17755	.31549	1.000	-.8666	1.2217
	News & Politics	-.41244	.45028	.999	-1.9028	1.0779
	Pop music	2.10464	.69757	.112	-.2042	4.4134
	Comedy	-.45143	.28478	.912	-1.3940	.4911
Education	Music	.07058	.31549	1.000	-.9736	1.1148
	Shows	-.53898	.56795	.998	-2.4187	1.3408
	Gaming	-.32164	.37227	.999	-1.5538	.9105
	Entertainment	-.17309	.32647	1.000	-1.2536	.9075
	People & Blogs	.28994	.40604	1.000	-1.0539	1.6338
	Sports	-1.00904	.74773	.971	-3.4839	1.4658
	Howto & Styles	.41023	.52465	1.000	-1.3262	2.1467
	Film & Animation	.24812	.41477	1.000	-1.1247	1.6209
	News & Politics	-.34186	.52465	1.000	-2.0783	1.3946
	Pop music	2.17522	.74773	.146	-.2996	4.6500
	Comedy	-.38085	.39192	.998	-1.6780	.9163

APPENDICES

Shows	Music	.60956	.50006	.987	-1.0455	2.2646
	Education	.53898	.56795	.998	-1.3408	2.4187
	Gaming	.21734	.53769	1.000	-1.5623	1.9970
	Entertainment	.36589	.50706	1.000	-1.3124	2.0442
	People & Blogs	.82892	.56160	.945	-1.0298	2.6877
	Sports	-.47006	.84240	1.000	-3.2582	2.3181
	Howto & Styles	.94921	.65252	.951	-1.2105	3.1089
	Film & Animation	.78710	.56795	.965	-1.0927	2.6669
	News & Politics	.19712	.65252	1.000	-1.9626	2.3568
	Pop music	2.71420	.84240	.064	-.0739	5.5023
	Comedy	.15813	.55148	1.000	-1.6671	1.9834
Gaming	Music	.39222	.25707	.932	-.4586	1.2430
	Education	.32164	.37227	.999	-.9105	1.5538
	Shows	-.21734	.53769	1.000	-1.9970	1.5623
	Entertainment	.14855	.27043	1.000	-.7465	1.0436
	People & Blogs	.61158	.36251	.872	-.5882	1.8114
	Sports	-.68739	.72502	.998	-3.0870	1.7122
	Howto & Styles	.73187	.49173	.942	-.8956	2.3594
	Film & Animation	.56977	.37227	.930	-.6623	1.8019
	News & Politics	-.02022	.49173	1.000	-1.6477	1.6073
	Pop music	2.49686[*]	.72502	.033	.0972	4.8965
	Comedy	-.05921	.34663	1.000	-1.2065	1.0880
Entertainment	Music	.24366	.18462	.976	-.3674	.8547
	Education	.17309	.32647	1.000	-.9075	1.2536
	Shows	-.36589	.50706	1.000	-2.0442	1.3124
	Gaming	-.14855	.27043	1.000	-1.0436	.7465
	People & Blogs	.46302	.31531	.947	-.5806	1.5066
	Sports	-.83595	.70261	.989	-3.1614	1.4895
	Howto & Styles	.58332	.45805	.981	-.9327	2.0993
	Film & Animation	.42121	.32647	.979	-.6593	1.5018
	News & Politics	-.16877	.45805	1.000	-1.6848	1.3472

APPENDICES

	Pop music	2.34831[*]	.70261	.045	.0228	4.6738
	Comedy	-.20776	.29691	1.000	-1.1905	.7749
People & Blogs	Music	-.21936	.30392	1.000	-1.2252	.7865
	Education	-.28994	.40604	1.000	-1.6338	1.0539
	Shows	-.82892	.56160	.945	-2.6877	1.0298
	Gaming	-.61158	.36251	.872	-1.8114	.5882
	Entertainment	-.46302	.31531	.947	-1.5066	.5806
	Sports	-1.29897	.74293	.843	-3.7579	1.1599
	Howto & Styles	.12029	.51777	1.000	-1.5934	1.8340
	Film & Animation	-.04181	.40604	1.000	-1.3857	1.3021
	News & Politics	-.63180	.51777	.987	-2.3455	1.0819
	Pop music	1.88528	.74293	.324	-.5736	4.3442
	Comedy	-.67079	.38267	.841	-1.9373	.5957
Sports	Music	1.07961	.69757	.925	-1.2292	3.3884
	Education	1.00904	.74773	.971	-1.4658	3.4839
	Shows	.47006	.84240	1.000	-2.3181	3.2582
	Gaming	.68739	.72502	.998	-1.7122	3.0870
	Entertainment	.83595	.70261	.989	-1.4895	3.1614
	People & Blogs	1.29897	.74293	.843	-1.1599	3.7579
	Howto & Styles	1.41927	.81383	.845	-1.2743	4.1129
	Film & Animation	1.25716	.74773	.875	-1.2177	3.7320
	News & Politics	.66718	.81383	1.000	-2.0264	3.3608
	Pop music	3.18425	.97272	.056	-.0352	6.4037
	Comedy	.62819	.73531	.999	-1.8055	3.0619
Howto & Styles	Music	-.33965	.45028	1.000	-1.8300	1.1507
	Education	-.41023	.52465	1.000	-2.1467	1.3262
	Shows	-.94921	.65252	.951	-3.1089	1.2105
	Gaming	-.73187	.49173	.942	-2.3594	.8956
	Entertainment	-.58332	.45805	.981	-2.0993	.9327
	People & Blogs	-.12029	.51777	1.000	-1.8340	1.5934
	Sports	-1.41927	.81383	.845	-4.1129	1.2743

APPENDICES

	Film & Animation	-.16211	.52465	1.000	-1.8986	1.5743
	News & Politics	-.75209	.61520	.987	-2.7883	1.2841
	Pop music	1.76499	.81383	.574	-.9286	4.4586
	Comedy	-.79108	.50677	.921	-2.4684	.8862
Film & Animation	Music	-.17755	.31549	1.000	-1.2217	.8666
	Education	-.24812	.41477	1.000	-1.6209	1.1247
	Shows	-.78710	.56795	.965	-2.6669	1.0927
	Gaming	-.56977	.37227	.930	-1.8019	.6623
	Entertainment	-.42121	.32647	.979	-1.5018	.6593
	People & Blogs	.04181	.40604	1.000	-1.3021	1.3857
	Sports	-1.25716	.74773	.875	-3.7320	1.2177
	Howto & Styles	.16211	.52465	1.000	-1.5743	1.8986
	News & Politics	-.58998	.52465	.993	-2.3264	1.1465
	Pop music	1.92709	.74773	.300	-.5477	4.4019
	Comedy	-.62897	.39192	.905	-1.9261	.6682
News & Politics	Music	.41244	.45028	.999	-1.0779	1.9028
	Education	.34186	.52465	1.000	-1.3946	2.0783
	Shows	-.19712	.65252	1.000	-2.3568	1.9626
	Gaming	.02022	.49173	1.000	-1.6073	1.6477
	Entertainment	.16877	.45805	1.000	-1.3472	1.6848
	People & Blogs	.63180	.51777	.987	-1.0819	2.3455
	Sports	-.66718	.81383	1.000	-3.3608	2.0264
	Howto & Styles	.75209	.61520	.987	-1.2841	2.7883
	Film & Animation	.58998	.52465	.993	-1.1465	2.3264
	Pop music	2.51708	.81383	.092	-.1765	5.2107
	Comedy	-.03899	.50677	1.000	-1.7163	1.6383
Pop music	Music	-2.10464	.69757	.112	-4.4134	.2042
	Education	-2.17522	.74773	.146	-4.6500	.2996
	Shows	-2.71420	.84240	.064	-5.5023	.0739
	Gaming	-2.49686*	.72502	.033	-4.8965	-.0972
	Entertainment	-2.34831*	.70261	.045	-4.6738	-.0228

APPENDICES

	People & Blogs	-1.88528	.74293	.324	-4.3442	.5736
	Sports	-3.18425	.97272	.056	-6.4037	.0352
	Howto & Styles	-1.76499	.81383	.574	-4.4586	.9286
	Film & Animation	-1.92709	.74773	.300	-4.4019	.5477
	News & Politics	-2.51708	.81383	.092	-5.2107	.1765
	Comedy	-2.55607[*]	.73531	.030	-4.9898	-.1224
Comedy	Music	.45143	.28478	.912	-.4911	1.3940
	Education	.38085	.39192	.998	-.9163	1.6780
	Shows	-.15813	.55148	1.000	-1.9834	1.6671
	Gaming	.05921	.34663	1.000	-1.0880	1.2065
	Entertainment	.20776	.29691	1.000	-.7749	1.1905
	People & Blogs	.67079	.38267	.841	-.5957	1.9373
	Sports	-.62819	.73531	.999	-3.0619	1.8055
	Howto & Styles	.79108	.50677	.921	-.8862	2.4684
	Film & Animation	.62897	.39192	.905	-.6682	1.9261
	News & Politics	.03899	.50677	1.000	-1.6383	1.7163
	Pop music	2.55607[*]	.73531	.030	.1224	4.9898

*. The mean difference is significant at the 0.05 level.