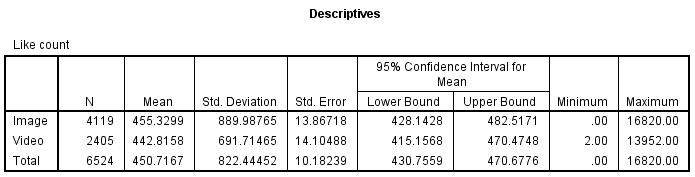
**Impact of ‘type of post’ on ‘likes’:**

We will use ANOVA because independent variable ‘type of post’ is categorical while dependent variable ‘like count’ is quantitative.

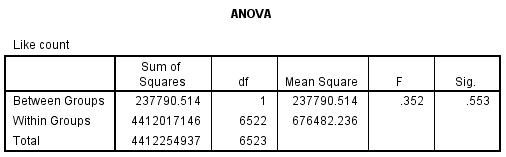


Average number of likes of image posts was 455.3299 while average number of likes of video posts was 442.8158 i.e., image posts garnered more likes as compared to video posts. This is also shown in the means plot below.

Variation in number of likes of image posts (standard deviation = 889.99) was higher than variation in number of likes of video posts (standard deviation = 691.71).

Chart, line chart

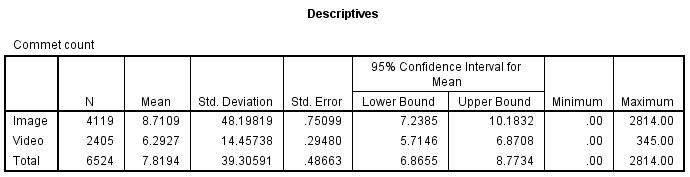
Description automatically generated



But the difference between average likes of image posts and average likes of video posts was statistically insignificant as the ANOVA p-value was 0.553 i.e., higher than alpha of 5%.

**Impact of ‘type of post’ on ‘comment count’:**

We will use ANOVA because independent variable ‘type of post’ is categorical while dependent variable ‘comment count’ is quantitative.

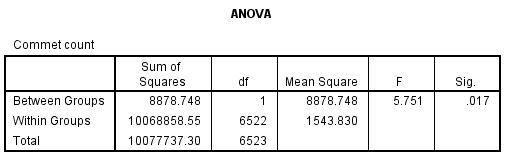


Average number of comments on image posts was 8.71 while average number of comments on video posts was 6.29 i.e., image posts garnered more comments as compared to video posts. This is also shown in the means plot below.

Variation in number of comments of image posts (standard deviation = 48.19) was higher than variation in number of comments of video posts (standard deviation = 14.45).

Chart, line chart

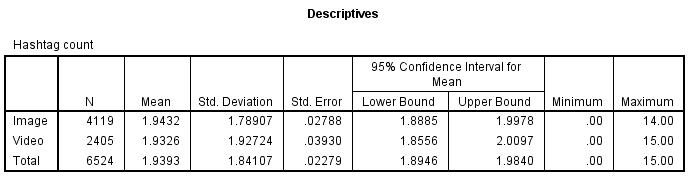
Description automatically generated



And the difference between average comments on image posts and average comments on video posts was statistically significant as the ANOVA p-value was 0.017 i.e., less than alpha of 5%.

**Impact of ‘type of post’ on ‘hashtag count’:**

We will use ANOVA because independent variable ‘type of post’ is categorical while dependent variable ‘hashtag count’ is quantitative.

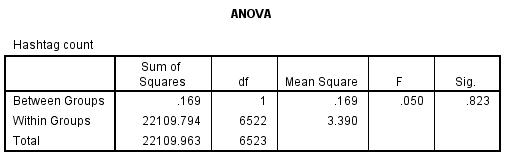


Average number of hashtags of image posts was 1.94 while average number of hashtags of video posts was 1.93 i.e., image posts garnered more hashtags as compared to video posts. This is also shown in the means plot below.

Variation in number of hashtags of image posts (standard deviation = 1.78) was lower than variation in number of hashtags of video posts (standard deviation = 1.92).

Chart, line chart

Description automatically generated

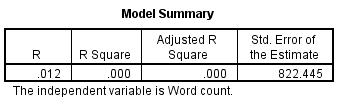


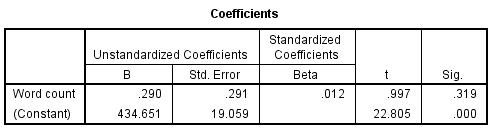
But the difference between average hashtags of image posts and average hashtags of video posts was statistically insignificant as the ANOVA p-value was 0.823 i.e., higher than alpha of 5%.

**Impact of ‘word count of post’ on ‘likes’:**

We will use regression (linear, quadratic, and cubic) because independent variable ‘word count of post’ as well as dependent variable ‘like count’ are both quantitative.

Linear regression:

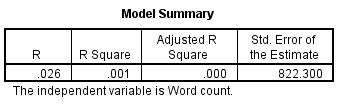


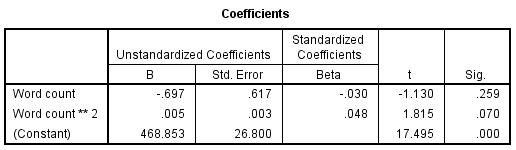


The R-squared of the linear model is 0.000 which implies that 0% of the variation in the dependent variable ‘like count’ is explained by the independent variable ‘word count of post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of linear model is 0.290 which implies that 1 word rise in word count of post leads to a 0.290 rise in number of likes garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘word count of post’ on ‘like count’ is statistically insignificant at the 5% significance level as the p-value of 0.319 is higher than alpha of 0.05.

Quadratic regression:

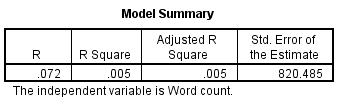


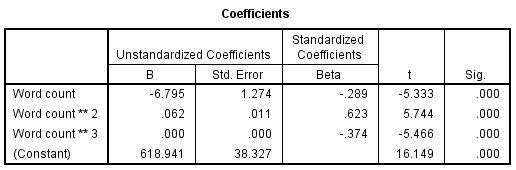


The R-squared of the quadratic model is 0.001 which implies that 0.1% of the variation in the dependent variable ‘like count’ is explained by the independent variables ‘word count of post’ and ‘square of word count of post’. Still this is an extremely low value, and it indicates very poor goodness of fit.

The coefficient of ‘word count of post’ and ‘square of word count of post’ is -0.697 and 0.005, respectively. This implies that x words rise in word count of post leads to a -0.697x + 0.005x^2 change in number of likes garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘word count of post’ on ‘like count’ is statistically insignificant at the 5% significance level as the p-value of ‘word count of post’ and ‘square of word count of post’ is 0.259 and 0.070, respectively i.e., higher than alpha of 0.05.

Cubic regression:





The R-squared of the quadratic model is 0.005 which implies that 0.5% of the variation in the dependent variable ‘like count’ is explained by the independent variables ‘word count of post’, ‘square of word count of post’, and ‘cube of word count of post’. This value is higher than that for linear and quadratic model, and hence it indicates best goodness of fit in comparison with other models. Thus, we reject linear and quadratic models and choose cubic model i.e., we conclude that relationship between ‘word count of post’ and ‘like count’ is cubic.

The coefficient of ‘word count of post’, ‘square of word count of post’, and ‘square of word count of post’ is -6.795, 0.062 and 0.000, respectively. This implies that x words rise in word count of post leads to a -6.795x + 0.062x^2 + 0.000x^3 change in number of likes garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘word count of post’ on ‘like count’ is statistically significant at the 5% significance level as the p-value of ‘word count of post’, ‘square of word count of post’ and ‘cube of word count of post’ is 0.000, 0.000, and 0.000, respectively i.e., lower than alpha of 0.05.

Curve estimation:

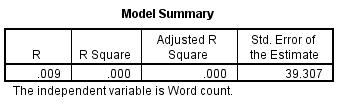
Chart, scatter chart

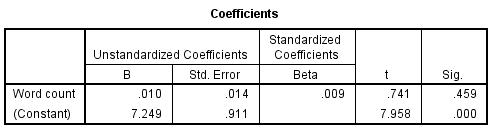
Description automatically generated

**Impact of ‘word count of post’ on ‘comment count’:**

We will use regression (linear, quadratic, and cubic) because independent variable ‘word count of post’ as well as dependent variable ‘comment count’ are both quantitative.

Linear regression:

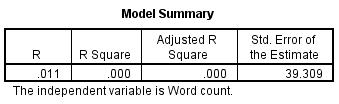


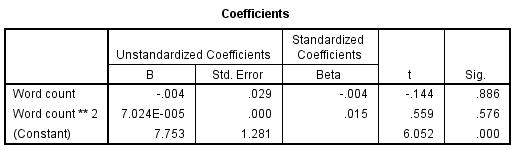


The R-squared of the linear model is 0.000 which implies that 0% of the variation in the dependent variable ‘comment count’ is explained by the independent variable ‘word count of post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of linear model is 0.010 which implies that 1 word rise in word count of post leads to a 0.010 rise in number of comments garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘word count of post’ on ‘comment count’ is statistically insignificant at the 5% significance level as the p-value of 0.3459 is higher than alpha of 0.05.

Quadratic:

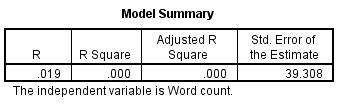


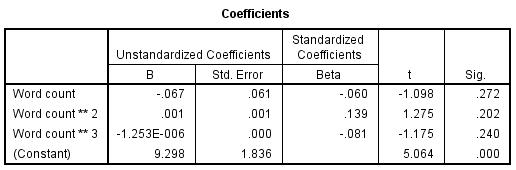


The R-squared of the quadratic model is 0.000 which implies that 0% of the variation in the dependent variable ‘comment count’ is explained by the independent variables ‘word count of post’ and ‘square of word count of post’. Still this is an extremely low value, and it indicates very poor goodness of fit.

The coefficient of ‘word count of post’ and ‘square of word count of post’ is -0.004 and 7.024E-005, respectively. This implies that x words rise in word count of post leads to a -0.004x + 7.024E-005x^2 change in number of comments garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘word count of post’ on ‘comment count’ is statistically insignificant at the 5% significance level as the p-value of ‘word count of post’ and ‘square of word count of post’ is 0.886 and 0.576, respectively i.e., higher than alpha of 0.05.

Cube:





The R-squared of the quadratic model is 0.000 which implies that 0% of the variation in the dependent variable ‘comment count’ is explained by the independent variables ‘word count of post’, ‘square of word count of post’, and ‘cube of word count of post’. This value is not any better than that for linear and quadratic model, and hence it indicates that none of the models have adequate goodness of fit. Thus, we reject all models (linear, quadratic as well as cubic model) i.e., we conclude that there is no relationship between ‘word count of post’ and ‘comment count’.

The coefficient of ‘word count of post’, ‘square of word count of post’, and ‘square of word count of post’ is -0.067, 0.001 and -1.253E-006, respectively. This implies that x words rise in word count of post leads to a -0.067x + 0.001x^2 – 1.253E-006x^3 change in number of comments garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘word count of post’ on ‘comment count’ is statistically insignificant at the 5% significance level as the p-value of ‘word count of post’, ‘square of word count of post’ and ‘cube of word count of post’ is 0.272, 0.202, and 0.240, respectively i.e., higher than alpha of 0.05.

Curve estimation:

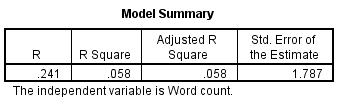
Graphical user interface

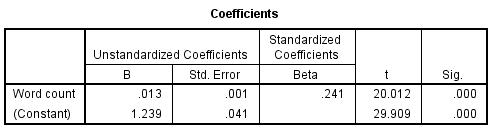
Description automatically generated with medium confidence

**Impact of ‘word count of post’ on ‘hashtag count’:**

We will use regression (linear, quadratic, and cubic) because independent variable ‘word count of post’ as well as dependent variable ‘hashtag count’ are both quantitative.

Linear regression:

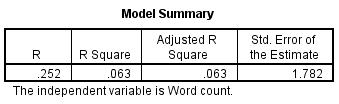


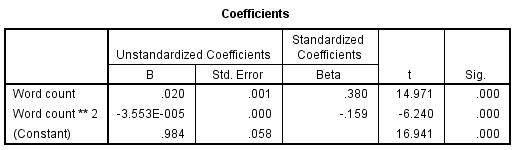


The R-squared of the linear model is 0.058 which implies that 5.8% of the variation in the dependent variable ‘hashtag count’ is explained by the independent variable ‘word count of post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of linear model is 0.013 which implies that 1 word rise in word count of post leads to a 0.013 rise in number of hashtags garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘word count of post’ on ‘hashtag count’ is statistically significant at the 5% significance level as the p-value of 0.000 is lower than alpha of 0.05.

Quadratic:

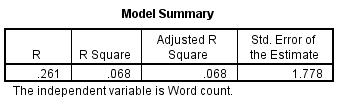


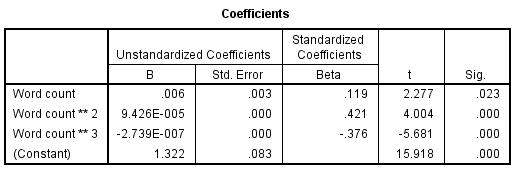


The R-squared of the quadratic model is 0.063 which implies that 6.3% of the variation in the dependent variable ‘hashtag count’ is explained by the independent variables ‘word count of post’ and ‘square of word count of post’. Still this is an extremely low value, and it indicates very poor goodness of fit.

The coefficient of ‘word count of post’ and ‘square of word count of post’ is 0.020 and -3.553E-005, respectively. This implies that x words rise in word count of post leads to a 0.020x – 3.553E-005x^2 change in number of hashtags garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘word count of post’ on ‘hashtag count’ is statistically significant at the 5% significance level as the p-value of ‘word count of post’ and ‘square of word count of post’ is 0.000 and 0.000, respectively i.e., lower than alpha of 0.05.

Cube:

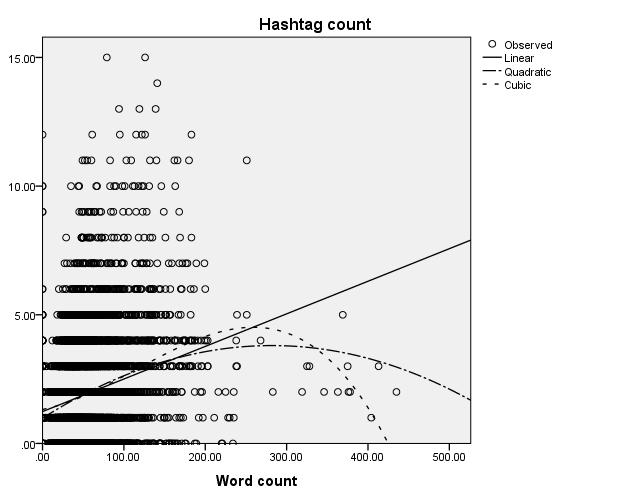




The R-squared of the quadratic model is 0.068 which implies that 6.8% of the variation in the dependent variable ‘hashtag count’ is explained by the independent variables ‘word count of post’, ‘square of word count of post’, and ‘cube of word count of post’. This value is higher than that for linear and quadratic model, and hence it indicates best goodness of fit in comparison with other models. Thus, we reject linear and quadratic models and choose cubic model i.e., we conclude that relationship between ‘word count of post’ and ‘hashtag count’ is cubic.

The coefficient of ‘word count of post’, ‘square of word count of post’, and ‘square of word count of post’ is 0.006, 9.426E-005 and -2.739E-007, respectively. This implies that x words rise in word count of post leads to a 0.006x + 9.426E-005x^2 – 2.739E-007x^3 change in number of hashtags garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘word count of post’ on ‘hashtag count’ is statistically significant at the 5% significance level as the p-value of ‘word count of post’, ‘square of word count of post’ and ‘cube of word count of post’ is 0.000, 0.000, and 0.000, respectively i.e., lower than alpha of 0.05.

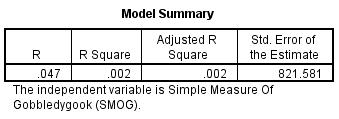
Curve estimation:

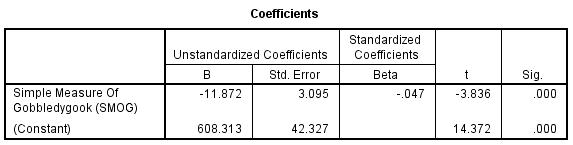


**Impact of ‘Simple Measure Of Gobbledygook (SMOG)’ on ‘likes’:**

We will use regression (linear, logarithmic, inverse, quadratic, and cubic) because independent variable ‘SMOG’ as well as dependent variable ‘like count’ are both quantitative.

Linear regression:

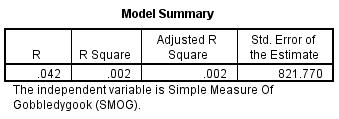


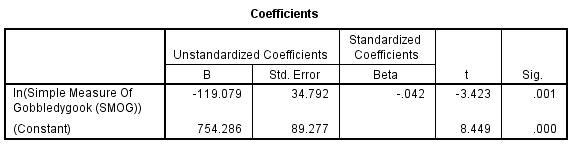


The R-squared of the linear model is 0.002 which implies that 0.2% of the variation in the dependent variable ‘like count’ is explained by the independent variable ‘SMOG of post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of linear model is -11.872 which implies that 1 unit rise in SMOG of post leads to a 11.872 unit fall in number of likes garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘SMOG of post’ on ‘like count’ is statistically significant at the 5% significance level as the p-value of 0.000 is lower than alpha of 0.05.

Logarithmic regression:

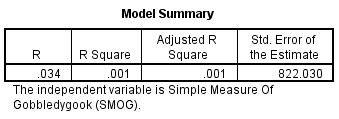


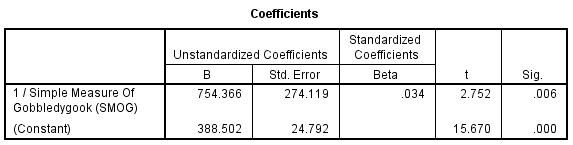


The R-squared of the logarithmic model is 0.002 which implies that 0.2% of the variation in the dependent variable ‘like count’ is explained by the independent variable ‘natural log of SMOG of post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of logarithmic model is -119.079 which implies that 1 percent rise in SMOG of post leads to a 119.079/100 = 1.119079% unit fall in number of likes garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘natural log of SMOG of post’ on ‘like count’ is statistically significant at the 5% significance level as the p-value of 0.001 is lower than alpha of 0.05.

Inverse regression:

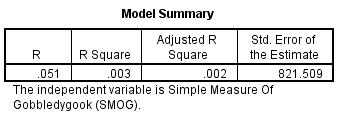


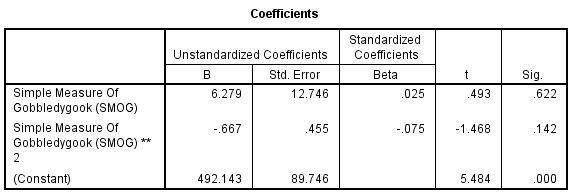


The R-squared of the inverse model is 0.001 which implies that 0.1% of the variation in the dependent variable ‘like count’ is explained by the independent variable ‘inverse of SMOG of post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of inverse model is 754.366 which implies that reducing SMOG of post by a factor of 1 lead to a 754.366 unit rise in number of likes garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘inverse of SMOG of post’ on ‘like count’ is statistically significant at the 5% significance level as the p-value of 0.006 is lower than alpha of 0.05.

Quadratic regression:

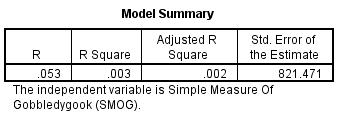


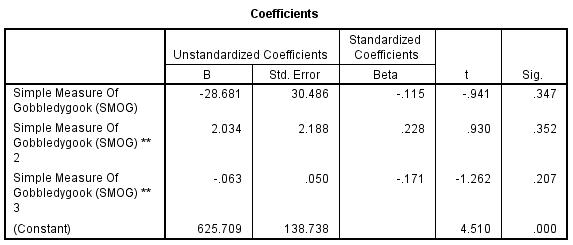


The R-squared of the quadratic model is 0.003 which implies that 0.3% of the variation in the dependent variable ‘like count’ is explained by the independent variables ‘SMOG of post’ and ‘square of SMOG of post’. Still this is an extremely low value, and it indicates very poor goodness of fit.

The coefficient of ‘SMOG of post’ and ‘square of SMOG of post’ is 6.279 and -0.667, respectively. This implies that x units rise in SMOG of post leads to a 6.279x - 0.667x^2 change in number of likes garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘SMOG of post’ on ‘like count’ is statistically insignificant at the 5% significance level as the p-value of ‘SMOG of post’ and ‘square of SMOG of post’ is 0.622 and 0.142, respectively i.e., higher than alpha of 0.05.

Cubic regression:





The R-squared of the quadratic model is 0.002 which implies that 0.2% of the variation in the dependent variable ‘like count’ is explained by the independent variables ‘SMOG of post’, ‘square of SMOG of post’, and ‘cube of SMOG of post’. This value is not any better than that for other models, and hence it indicates that none of the models have adequate goodness of fit. Thus, we reject all models (linear, quadratic as well as cubic mode i.e., we conclude that there is no relationship between ‘SMOG’ and ‘like count’. But still there is relationship between ‘inverse of SMOG’ and ‘like count’.

The coefficient of ‘SMOG of post’, ‘square of SMOG of post’, and ‘square of SMOG of post’ is -28.68, 2.034 and -0.063, respectively. This implies that x units rise in SMOG of post leads to a -28.68x + 2.034x^2 - 0.063x^3 change in number of likes garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘SMOG of post’ on ‘like count’ is statistically significant at the 5% significance level as the p-value of ‘SMOG of post’, ‘square of SMOG of post’ and ‘cube of SMOG of post’ is 0.347, 0.352, and 0.207, respectively i.e., higher than alpha of 0.05.

Curve estimation:

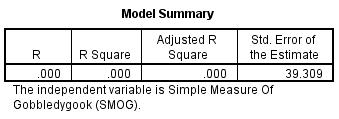
Chart, scatter chart

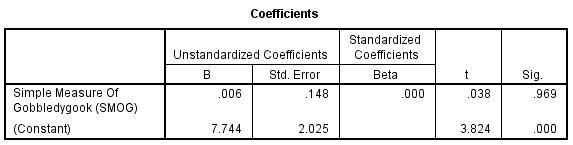
Description automatically generated

**Impact of ‘Simple Measure Of Gobbledygook (SMOG)’ on ‘comments’:**

We will use regression (linear, logarithmic, inverse, quadratic, and cubic) because independent variable ‘SMOG’ as well as dependent variable ‘comment count’ are both quantitative.

Linear regression:

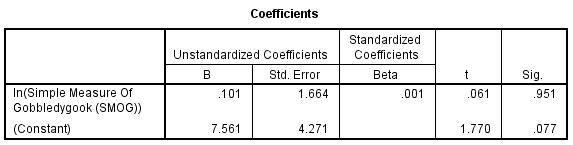
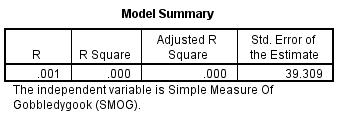




The R-squared of the linear model is 0.000 which implies that 0% of the variation in the dependent variable ‘comment count’ is explained by the independent variable ‘SMOG of post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of linear model is 0.006 which implies that 1 unit rise in SMOG of post leads to a 11.872 unit fall in number of comments garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘SMOG of post’ on ‘comment count’ is statistically significant at the 5% significance level as the p-value of 0.969 is higher than alpha of 0.05.

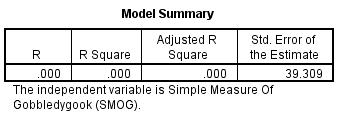
Logarithmic regression:

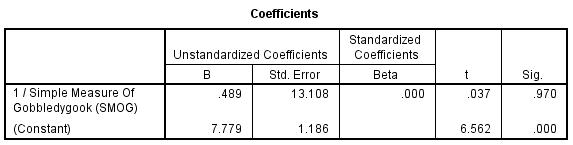


The R-squared of the logarithmic model is 0.000 which implies that 0% of the variation in the dependent variable ‘comment count’ is explained by the independent variable ‘natural log of SMOG of post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of logarithmic model is 0.101 which implies that 1 percent rise in SMOG of post leads to a 0.101/100 = 0.00101% unit fall in number of comments garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘natural log of SMOG of post’ on ‘comment count’ is statistically insignificant at the 5% significance level as the p-value of 0.951 is higher than alpha of 0.05.

Inverse regression:

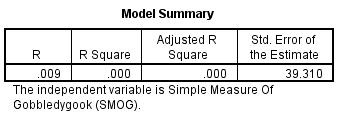


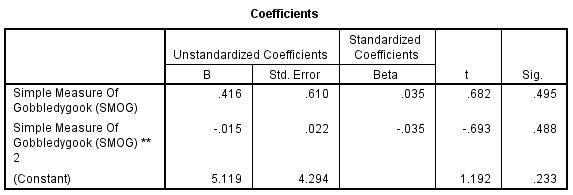


The R-squared of the inverse model is 0.000 which implies that 0% of the variation in the dependent variable ‘comment count’ is explained by the independent variable ‘inverse of SMOG of post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of inverse model is 0.489 which implies that reducing SMOG of post by a factor of 1 lead to a 0.489 unit rise in number of comments garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘inverse of SMOG of post’ on ‘comment count’ is statistically insignificant at the 5% significance level as the p-value of 0.970 is higher than alpha of 0.05.

Quadratic regression:

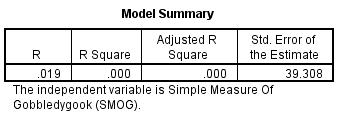


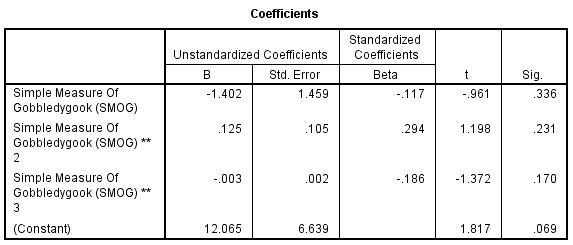


The R-squared of the quadratic model is 0.000 which implies that 0% of the variation in the dependent variable ‘comment count’ is explained by the independent variables ‘SMOG of post’ and ‘square of SMOG of post’. Still this is an extremely low value, and it indicates very poor goodness of fit.

The coefficient of ‘SMOG of post’ and ‘square of SMOG of post’ is 0.416 and -0.015, respectively. This implies that x units rise in SMOG of post leads to a 0.416x - 0.015x^2 change in number of comments garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘SMOG of post’ on ‘comment count’ is statistically insignificant at the 5% significance level as the p-value of ‘SMOG of post’ and ‘square of SMOG of post’ is 0.495 and 0.488, respectively i.e., higher than alpha of 0.05.

Cubic regression:





The R-squared of the quadratic model is 0.000 which implies that 0% of the variation in the dependent variable ‘comment count’ is explained by the independent variables ‘SMOG of post’, ‘square of SMOG of post’, and ‘cube of SMOG of post’. This value is not any better than that for other models, and hence it indicates that none of the models have adequate goodness of fit. Thus, we reject all models (linear, quadratic as well as cubic mode i.e., we conclude that there is no relationship between ‘SMOG’ and ‘comment count’. But still there is relationship between ‘inverse of SMOG’ and ‘comment count’.

The coefficient of ‘SMOG of post’, ‘square of SMOG of post’, and ‘square of SMOG of post’ is -1.402, 0.125 and -0.003, respectively. This implies that x units rise in SMOG of post leads to a -1.402x + 0.125x^2 - 0.003x^3 change in number of comments garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘SMOG of post’ on ‘comment count’ is statistically significant at the 5% significance level as the p-value of ‘SMOG of post’, ‘square of SMOG of post’ and ‘cube of SMOG of post’ is 0.336, 0.231, and 0.170, respectively i.e., higher than alpha of 0.05.

Curve estimation:

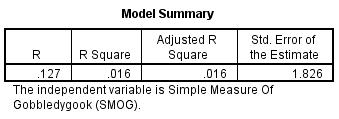
Graphical user interface

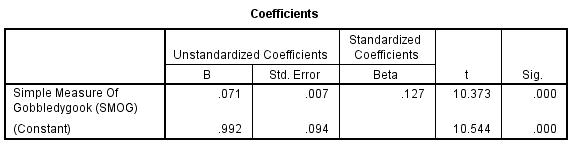
Description automatically generated with medium confidence

**Impact of ‘Simple Measure Of Gobbledygook (SMOG)’ on ‘hashtag count’:**

We will use regression (linear, logarithmic, inverse, quadratic, and cubic) because independent variable ‘SMOG’ as well as dependent variable ‘hashtag count’ are both quantitative.

Linear regression:

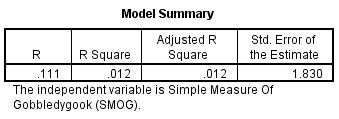


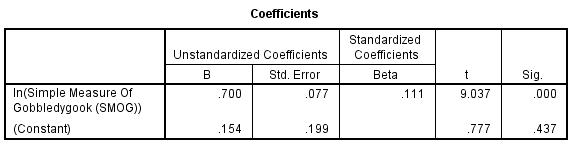


The R-squared of the linear model is 0.016 which implies that 1.6% of the variation in the dependent variable ‘hashtag count’ is explained by the independent variable ‘SMOG of post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of linear model is 0.071 which implies that 1 unit rise in SMOG of post leads to a 0.071 unit fall in number of hashtags garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘SMOG of post’ on ‘hashtag count’ is statistically significant at the 5% significance level as the p-value of 0.000 is lower than alpha of 0.05.

Logarithmic regression:

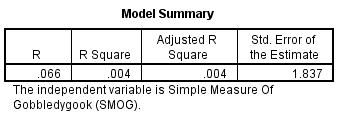


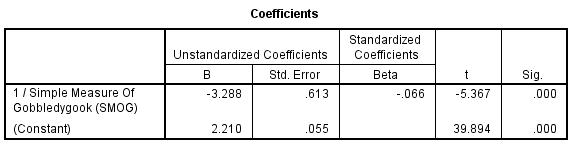


The R-squared of the logarithmic model is 0.012 which implies that 1.2% of the variation in the dependent variable ‘hashtag count’ is explained by the independent variable ‘natural log of SMOG of post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of logarithmic model is 0.700 which implies that 1 percent rise in SMOG of post leads to a 0.700/100 = 0.007% unit rise in number of hashtags garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘natural log of SMOG of post’ on ‘hashtag count’ is statistically significant at the 5% significance level as the p-value of 0.000 is less than alpha of 0.05.

Inverse regression:

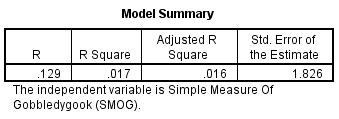


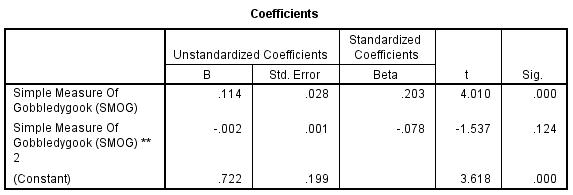


The R-squared of the inverse model is 0.004 which implies that 0.4% of the variation in the dependent variable ‘hashtag count’ is explained by the independent variable ‘inverse of SMOG of post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of inverse model is -3.288 which implies that reducing SMOG of post by a factor of 1 lead to a 3.288 unit fall in number of hashtags garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘inverse of SMOG of post’ on ‘hashtag count’ is statistically significant at the 5% significance level as the p-value of 0.000 is lower than alpha of 0.05.

Quadratic regression:

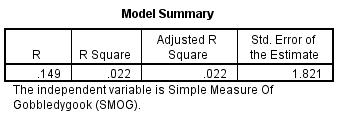


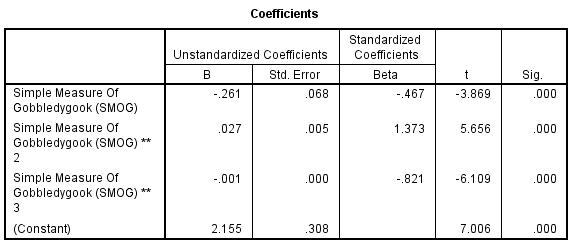


The R-squared of the quadratic model is 0.017 which implies that 1.7% of the variation in the dependent variable ‘hashtag count’ is explained by the independent variables ‘SMOG of post’ and ‘square of SMOG of post’. Still this is an extremely low value, and it indicates very poor goodness of fit.

The coefficient of ‘SMOG of post’ and ‘square of SMOG of post’ is 0.114 and -0.002, respectively. This implies that x units rise in SMOG of post leads to a 0.114x - 0.002x^2 change in number of hashtags garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘SMOG of post’ on ‘hashtag count’ is statistically insignificant at the 5% significance level as the p-value of ‘square of SMOG of post’ is 0.124 i.e., higher than alpha of 0.05.

Cubic regression:

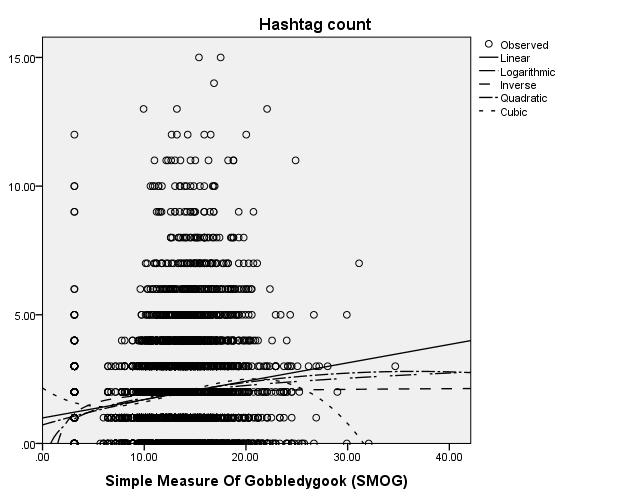




The R-squared of the quadratic model is 0.022 which implies that 2.2% of the variation in the dependent variable ‘hashtag count’ is explained by the independent variables ‘SMOG of post’, ‘square of SMOG of post’, and ‘cube of SMOG of post’. This value is better than that for other models, and hence it indicates that cubic model has adequate goodness of fit. Thus, we reject all other models and conclude that there is cubic relationship between ‘SMOG’ and ‘hashtag count’.

The coefficient of ‘SMOG of post’, ‘square of SMOG of post’, and ‘square of SMOG of post’ is -0.261, 0.027 and -0.001, respectively. This implies that x units rise in SMOG of post leads to a -0.261x + 0.027x^2 - 0.001x^3 change in number of hashtags garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘SMOG of post’ on ‘hashtag count’ is statistically significant at the 5% significance level as the p-value of ‘SMOG of post’, ‘square of SMOG of post’ and ‘cube of SMOG of post’ is 0.000, 0.000, and 0.000, respectively i.e., lower than alpha of 0.05.

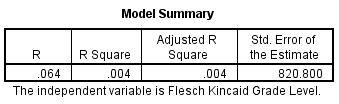
Curve estimation:

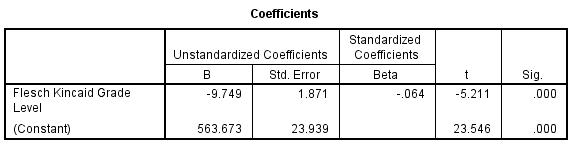


**Impact of ‘Flesch Kincaid’ on ‘likes’:**

We will use regression (linear, logarithmic, inverse, quadratic, and cubic) because independent variable ‘FLESCH KINCAID’ as well as dependent variable ‘like count’ are both quantitative.

Linear regression:

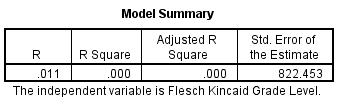


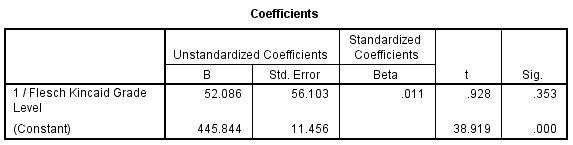


The R-squared of the linear model is 0.004 which implies that 0.4% of the variation in the dependent variable ‘like count’ is explained by the independent variable ‘FLESCH KINCAID of post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of linear model is -9.749 which implies that 1 unit rise in FLESCH KINCAID of post leads to a 9.749 unit fall in number of likes garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘FLESCH KINCAID of post’ on ‘like count’ is statistically significant at the 5% significance level as the p-value of 0.000 is lower than alpha of 0.05.

Inverse regression:

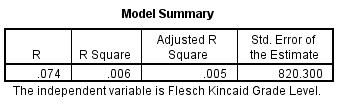


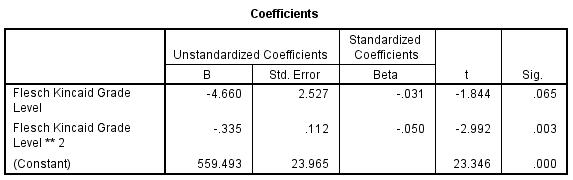


The R-squared of the inverse model is 0.000 which implies that 0% of the variation in the dependent variable ‘like count’ is explained by the independent variable ‘inverse of FLESCH KINCAID of post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of inverse model is 52.086 which implies that reducing FLESCH KINCAID of post by a factor of 1 lead to a 52.086 unit rise in number of likes garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘inverse of FLESCH KINCAID of post’ on ‘like count’ is statistically insignificant at the 5% significance level as the p-value of 0.353 is higher than alpha of 0.05.

Quadratic regression:

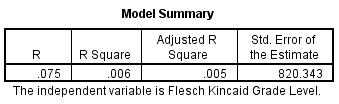


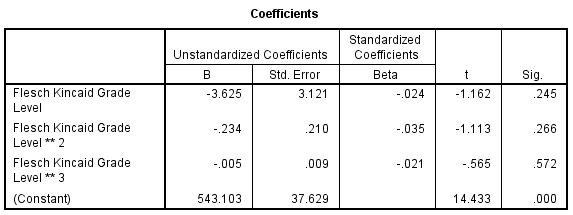


The R-squared of the quadratic model is 0.006 which implies that 0.6% of the variation in the dependent variable ‘like count’ is explained by the independent variables ‘FLESCH KINCAID of post’ and ‘square of FLESCH KINCAID of post’. Still this is an extremely low value, and it indicates very poor goodness of fit.

The coefficient of ‘FLESCH KINCAID of post’ and ‘square of FLESCH KINCAID of post’ is -4.660 and -0.335, respectively. This implies that x units rise in FLESCH KINCAID of post leads to a -4.660x - 0.335x^2 change in number of likes garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘FLESCH KINCAID of post’ on ‘like count’ is statistically insignificant at the 5% significance level as the p-value of ‘FLESCH KINCAID of post’ is 0.065 i.e., higher than alpha of 0.05.

Cubic regression:





The R-squared of the quadratic model is 0.006 which implies that 0.6% of the variation in the dependent variable ‘like count’ is explained by the independent variables ‘FLESCH KINCAID of post’, ‘square of FLESCH KINCAID of post’, and ‘cube of FLESCH KINCAID of post’. This value is not any better than that for other models, and hence it indicates that none of the models have adequate goodness of fit. Thus, we reject all models (linear, quadratic as well as cubic mode) i.e., we conclude that there is no relationship between ‘FLESCH KINCAID’ and ‘like count’. But still there is relationship between ‘inverse of FLESCH KINCAID’ and ‘like count’.

The coefficient of ‘FLESCH KINCAID of post’, ‘square of FLESCH KINCAID of post’, and ‘square of FLESCH KINCAID of post’ is -3.625, -0.234 and -0.005, respectively. This implies that x units rise in FLESCH KINCAID of post leads to a -3.625x – 0.234x^2 - 0.005x^3 change in number of likes garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘FLESCH KINCAID of post’ on ‘like count’ is statistically significant at the 5% significance level as the p-value of ‘FLESCH KINCAID of post’, ‘square of FLESCH KINCAID of post’ and ‘cube of FLESCH KINCAID of post’ is 0.245, 0.266, and 0.572, respectively i.e., higher than alpha of 0.05.

Curve estimation:

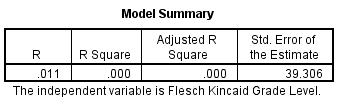
Chart, scatter chart

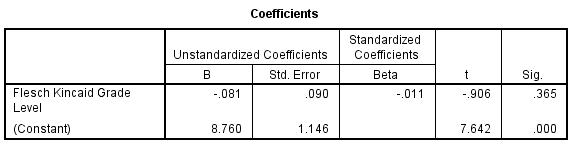
Description automatically generated

**Impact of ‘Flesch Kincaid’ on ‘comments’:**

We will use regression (linear, logarithmic, inverse, quadratic, and cubic) because independent variable ‘FLESCH KINCAID’ as well as dependent variable ‘comment count’ are both quantitative.

Linear regression:

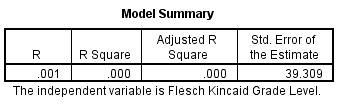


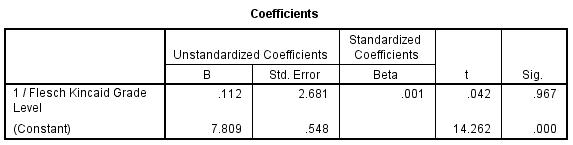


The R-squared of the linear model is 0.000 which implies that 0% of the variation in the dependent variable ‘comment count’ is explained by the independent variable ‘FLESCH KINCAID of post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of linear model is -0.081 which implies that 1 unit rise in FLESCH KINCAID of post leads to a 0.081 unit fall in number of comments garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘FLESCH KINCAID of post’ on ‘comment count’ is statistically insignificant at the 5% significance level as the p-value of 0.365 is higher than alpha of 0.05.

Inverse regression:

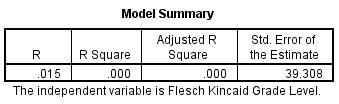


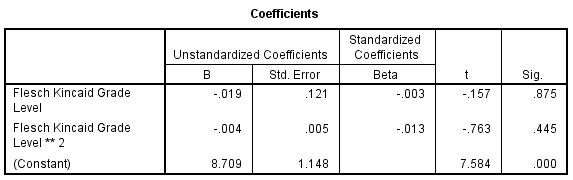


The R-squared of the inverse model is 0.000 which implies that 0% of the variation in the dependent variable ‘comment count’ is explained by the independent variable ‘inverse of FLESCH KINCAID of post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of inverse model is 0.112 which implies that reducing FLESCH KINCAID of post by a factor of 1 lead to a 0.112 unit rise in number of comments garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘inverse of FLESCH KINCAID of post’ on ‘comment count’ is statistically insignificant at the 5% significance level as the p-value of 0.967 is higher than alpha of 0.05.

Quadratic regression:

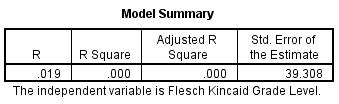


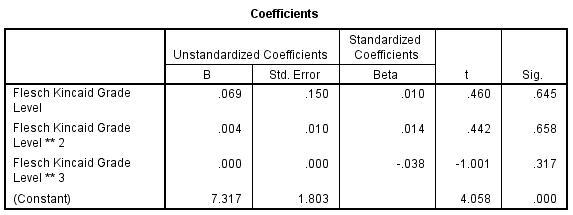


The R-squared of the quadratic model is 0.000 which implies that 0% of the variation in the dependent variable ‘comment count’ is explained by the independent variables ‘FLESCH KINCAID of post’ and ‘square of FLESCH KINCAID of post’. Still this is an extremely low value, and it indicates very poor goodness of fit.

The coefficient of ‘FLESCH KINCAID of post’ and ‘square of FLESCH KINCAID of post’ is -0.019 and -0.004, respectively. This implies that x units rise in FLESCH KINCAID of post leads to a -0.019x - 0.004x^2 change in number of comments garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘FLESCH KINCAID of post’ on ‘comment count’ is statistically insignificant at the 5% significance level as the p-value of ‘FLESCH KINCAID of post’ is 0.875 i.e., higher than alpha of 0.05.

Cubic regression:

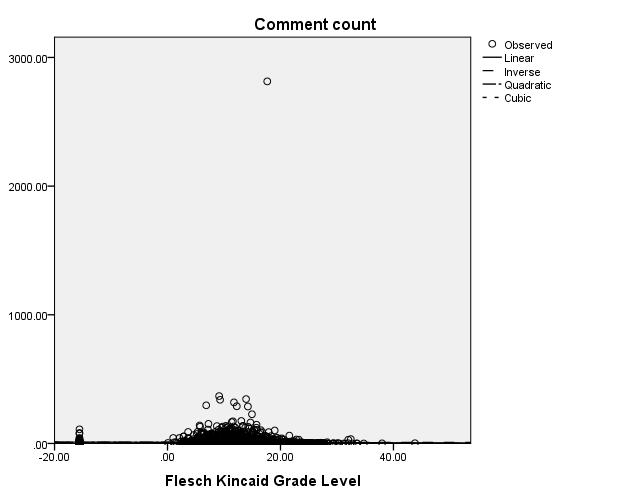




The R-squared of the quadratic model is 0.000 which implies that 0% of the variation in the dependent variable ‘comment count’ is explained by the independent variables ‘FLESCH KINCAID of post’, ‘square of FLESCH KINCAID of post’, and ‘cube of FLESCH KINCAID of post’. This value is not any better than that for other models, and hence it indicates that none of the models have adequate goodness of fit. Thus, we reject all models (linear, quadratic as well as cubic mode) i.e., we conclude that there is no relationship between ‘FLESCH KINCAID’ and ‘comment count’. But still there is relationship between ‘inverse of FLESCH KINCAID’ and ‘comment count’.

The coefficient of ‘FLESCH KINCAID of post’, ‘square of FLESCH KINCAID of post’, and ‘square of FLESCH KINCAID of post’ is 0.069, 0.004 and 0.000, respectively. This implies that x units rise in FLESCH KINCAID of post leads to a 0.069x + 0.004x^2 + 0.000x^3 change in number of comments garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘FLESCH KINCAID of post’ on ‘comment count’ is statistically significant at the 5% significance level as the p-value of ‘FLESCH KINCAID of post’, ‘square of FLESCH KINCAID of post’ and ‘cube of FLESCH KINCAID of post’ is 0.645, 0.658, and 0.317, respectively i.e., higher than alpha of 0.05.

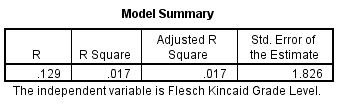
Curve estimation:

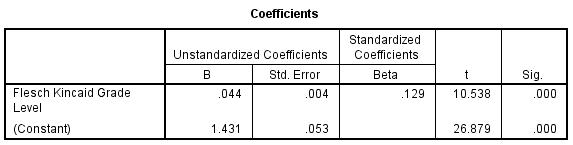


**Impact of ‘Flesch Kincaid’ on ‘hashtags’:**

We will use regression (linear, logarithmic, inverse, quadratic, and cubic) because independent variable ‘FLESCH KINCAID’ as well as dependent variable ‘hashtag count’ are both quantitative.

Linear regression:

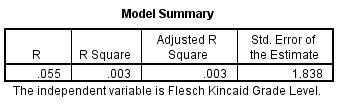


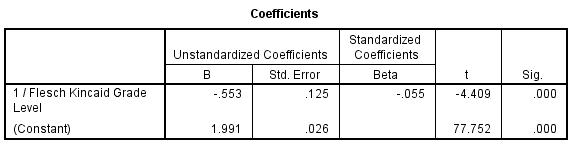


The R-squared of the linear model is 0.017 which implies that 1.7% of the variation in the dependent variable ‘hashtag count’ is explained by the independent variable ‘FLESCH KINCAID of post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of linear model is 0.044 which implies that 1 unit rise in FLESCH KINCAID of post leads to a 0.044 unit fall in number of hashtags garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘FLESCH KINCAID of post’ on ‘hashtag count’ is statistically significant at the 5% significance level as the p-value of 0.000 is lower than alpha of 0.05.

Inverse regression:

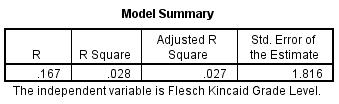


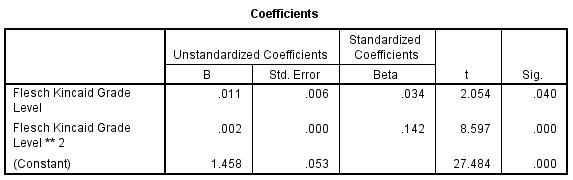


The R-squared of the inverse model is 0.003 which implies that 0.3% of the variation in the dependent variable ‘hashtag count’ is explained by the independent variable ‘inverse of FLESCH KINCAID of post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of inverse model is -0.553 which implies that reducing FLESCH KINCAID of post by a factor of 1 lead to a 0.553 unit fall in number of hashtags garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘inverse of FLESCH KINCAID of post’ on ‘hashtag count’ is statistically significant at the 5% significance level as the p-value of 0.000 is lower than alpha of 0.05.

Quadratic regression:

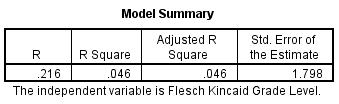


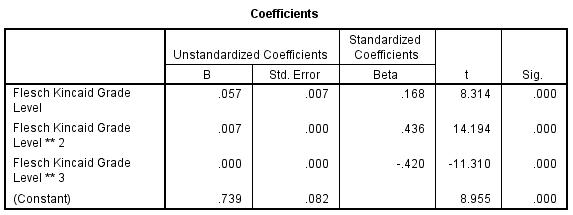


The R-squared of the quadratic model is 0.028 which implies that 2.8% of the variation in the dependent variable ‘hashtag count’ is explained by the independent variables ‘FLESCH KINCAID of post’ and ‘square of FLESCH KINCAID of post’. Still this is an extremely low value, and it indicates very poor goodness of fit.

The coefficient of ‘FLESCH KINCAID of post’ and ‘square of FLESCH KINCAID of post’ is 0.011 and 0.002, respectively. This implies that x units rise in FLESCH KINCAID of post leads to a 0.011x + 0.002x^2 change in number of hashtags garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘FLESCH KINCAID of post’ on ‘hashtag count’ is statistically significant at the 5% significance level as the p-value of ‘FLESCH KINCAID of post’ is 0.040 i.e., less than alpha of 0.05.

Cubic regression:





The R-squared of the quadratic model is 0.046 which implies that 4.6% of the variation in the dependent variable ‘hashtag count’ is explained by the independent variables ‘FLESCH KINCAID of post’, ‘square of FLESCH KINCAID of post’, and ‘cube of FLESCH KINCAID of post’. This value is better than all models and hence we conclude that relationship between ‘FLESCH KINCAID’ and ‘hashtag count’ is cubic.

The coefficient of ‘FLESCH KINCAID of post’, ‘square of FLESCH KINCAID of post’, and ‘square of FLESCH KINCAID of post’ is 0.057, 0.007 and 0.000, respectively. This implies that x units rise in FLESCH KINCAID of post leads to a 0.057x + 0.007x^2 + 0.000x^3 change in number of hashtags garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘FLESCH KINCAID of post’ on ‘hashtag count’ is statistically significant at the 5% significance level as the p-value of ‘FLESCH KINCAID of post’, ‘square of FLESCH KINCAID of post’ and ‘cube of FLESCH KINCAID of post’ is 0.000, 0.000, and 0.000, respectively i.e., lower than alpha of 0.05.

Curve estimation:

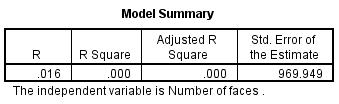
Diagram

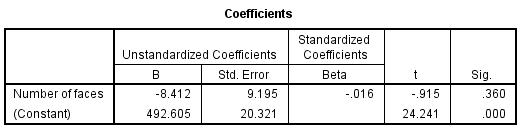
Description automatically generated

**Impact of ‘number of faces in post’ on ‘like count’:**

We will use regression (linear, quadratic, and cubic) because independent variable ‘number of faces in post’ as well as dependent variable ‘like count’ are both quantitative.

Linear regression:

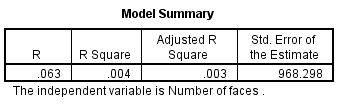


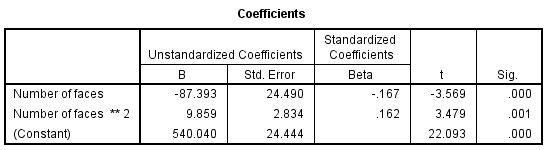


The R-squared of the linear model is 0.000 which implies that 0% of the variation in the dependent variable ‘like count’ is explained by the independent variable ‘number of faces in post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of linear model is -8.412 which implies that 1 unit rise in number of faces in post leads to a 8.412 unit fall in number of hashtags garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘number of faces in post’ on ‘like count’ is statistically insignificant at the 5% significance level as the p-value of 0.350 is higher than alpha of 0.05.

Quadratic:

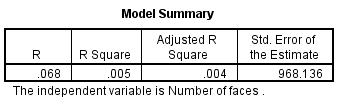


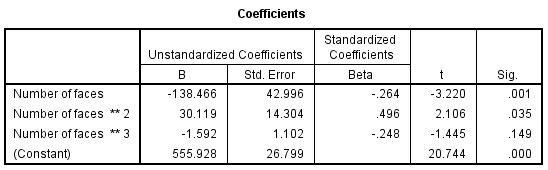


The R-squared of the quadratic model is 0.004 which implies that 0.4% of the variation in the dependent variable ‘like count’ is explained by the independent variables ‘number of faces in post’ and ‘number of faces in post’. Still this is an extremely low value, and it indicates very poor goodness of fit.

The coefficient of ‘number of faces in post’ and ‘square of number of faces in post’ is -87.393 and 9.859, respectively. This implies that x words rise in word count of post leads to a -87.393x + 9.859x^2 change in number of likes garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘number of faces in post’ on ‘like count’ is statistically significant at the 5% significance level as the p-value of ‘number of faces in post’ and ‘square of number of faces in post’ is 0.000 and 0.001, respectively i.e., lower than alpha of 0.05.

Cubic regression:





The R-squared of the quadratic model is 0.005 which implies that 0.5% of the variation in the dependent variable ‘like count’ is explained by the independent variables ‘number of faces in post’, ‘square of number of faces in post’, and ‘cube of number of faces in post’. This value is higher than that for linear and quadratic model, and hence it indicates best goodness of fit in comparison with other models. Thus, we reject linear and quadratic models and choose cubic model i.e., we conclude that relationship between ‘number of faces in post’ and ‘like count’ is cubic.

The coefficient of ‘number of faces in post’, ‘square of number of faces in post’, and ‘square of number of faces in post’ is -138.466, 30.119 and -1.592, respectively. This implies that x words rise in number of faces in post leads to a -138.466x + 30.119x^2 – 1.592x^3 change in number of likes garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘number of faces in post’ on ‘like count’ is statistically significant at the 5% significance level as the p-value of ‘number of faces in post’ and ‘square of number of faces in post’ is 0.001 and 0.035, respectively i.e., lower than alpha of 0.05.

Curve estimation:

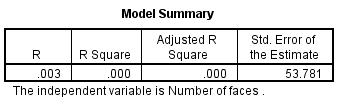
Chart

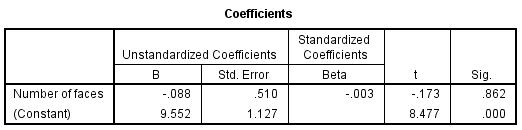
Description automatically generated

**Impact of ‘number of faces in post’ on ‘comment count’:**

We will use regression (linear, quadratic, and cubic) because independent variable ‘number of faces in post’ as well as dependent variable ‘comment count’ are both quantitative.

Linear regression:

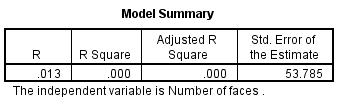


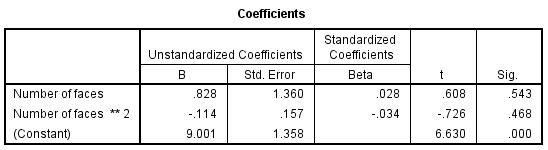


The R-squared of the linear model is 0.000 which implies that 0% of the variation in the dependent variable ‘comment count’ is explained by the independent variable ‘number of faces in post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of linear model is -0.088 which implies that 1 unit rise in number of faces in post leads to a 0.088 unit fall in number of hashtags garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘number of faces in post’ on ‘comment count’ is statistically insignificant at the 5% significance level as the p-value of 0.862 is higher than alpha of 0.05.

Quadratic:

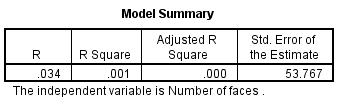


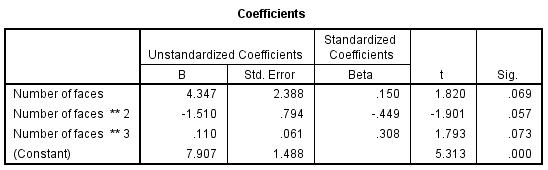


The R-squared of the quadratic model is 0.000 which implies that 0% of the variation in the dependent variable ‘comment count’ is explained by the independent variables ‘number of faces in post’ and ‘number of faces in post’. Still this is an extremely low value, and it indicates very poor goodness of fit.

The coefficient of ‘number of faces in post’ and ‘square of number of faces in post’ is 0.828 and -0.114, respectively. This implies that x words rise in word count of post leads to a 0.828x - 0.114x^2 change in number of comments garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘number of faces in post’ on ‘comment count’ is statistically insignificant at the 5% significance level as the p-value of ‘number of faces in post’ and ‘square of number of faces in post’ is 0.543 and 0.468, respectively i.e., higher than alpha of 0.05.

Cubic regression:

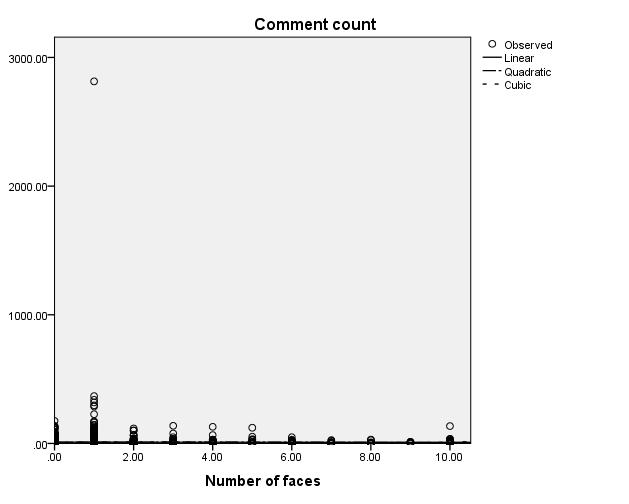




The R-squared of the cubic model is 0.001 which implies that 0.1% of the variation in the dependent variable ‘comment count’ is explained by the independent variables ‘number of faces in post’, ‘square of number of faces in post’, and ‘cube of number of faces in post’. This value is higher than that for linear and quadratic model, and hence it indicates best goodness of fit in comparison with other models. Thus, we reject linear and quadratic models and choose cubic model i.e., we conclude that relationship between ‘number of faces in post’ and ‘comment count’ is cubic.

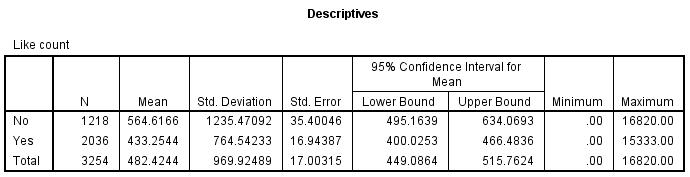
The coefficient of ‘number of faces in post’, ‘square of number of faces in post’, and ‘square of number of faces in post’ is 4.347, -1.510 and 0.110, respectively. This implies that x words rise in number of faces in post leads to a 4.347x – 1.510x^2 + 0.110x^3 change in number of comments garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘number of faces in post’ on ‘comment count’ is statistically insignificant at the 5% significance level as the p-value of ‘number of faces in post’, ‘square of number of faces in post’ and ‘cube of number of faces in post’ are 0.069, 0.057, and 0.073 respectively i.e., higher than alpha of 0.05.

Curve estimation:



**Impact of ‘existence of face in the post’ on ‘like count’:**

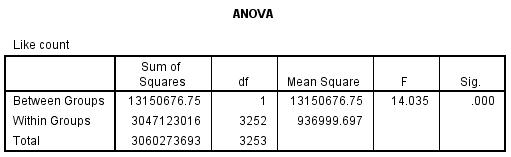
We will use ANOVA because independent variable ‘existence of face in the post’ is categorical while dependent variable ‘like count’ is quantitative.



Average number of likes of posts in which face exists was 433.25 while average number of likes of posts in which face doesn’t exist was 564.62 i.e., posts in which face exists garnered less likes as compared to posts in which face doesn’t exist. This is also shown in the means plot below.

Chart

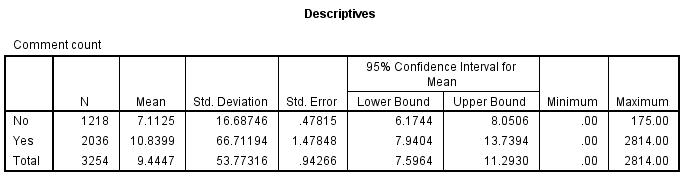
Description automatically generated



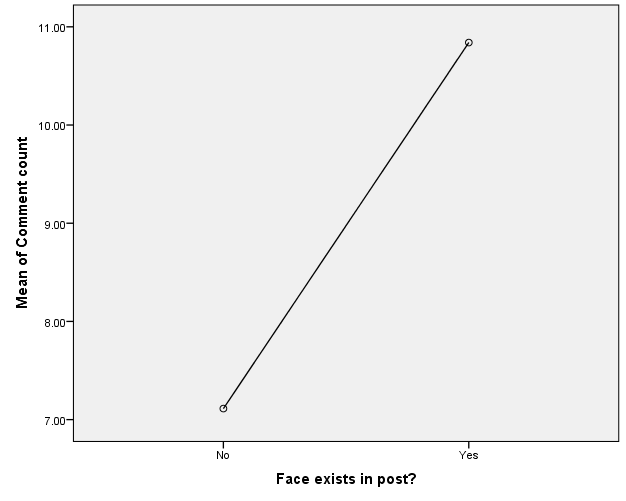
And the difference between average likes of posts in which face exists and average likes of posts in which face doesn’t exist was statistically significant as the ANOVA p-value was 0.000 i.e., lower than alpha of 5%.

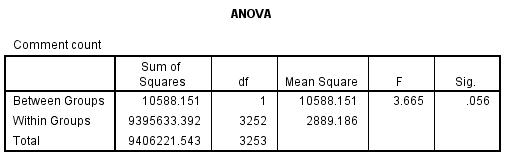
**Impact of ‘existence of face in the post’ on ‘comment count’:**

We will use ANOVA because independent variable ‘existence of face in the post’ is categorical while dependent variable ‘comment count’ is quantitative.



Average number of comments of posts in which face exists was 10.84 while average number of comments of posts in which face doesn’t exist was 7.11 i.e., posts in which face exists garnered more comments as compared to posts in which face doesn’t exist. This is also shown in the means plot below.

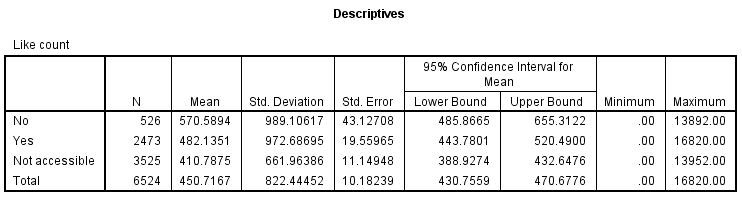




But the difference between average comments of posts in which face exists and average comments of posts in which face doesn’t exist was statistically insignificant as the ANOVA p-value was 0.056 i.e., higher than alpha of 5%.

**Impact of ‘existence of text in the post’ on ‘like count’:**

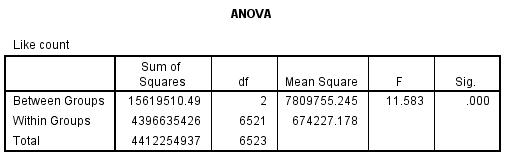
We will use ANOVA because independent variable ‘existence of text in the post’ is categorical while dependent variable ‘like count’ is quantitative.



Average number of likes of posts in which text exists was 482.1351 while average number of likes of posts in which text doesn’t exist was 570.5894 i.e., posts in which text exists garnered less likes as compared to posts in which text doesn’t exist. This is also shown in the means plot below.

Chart, line chart

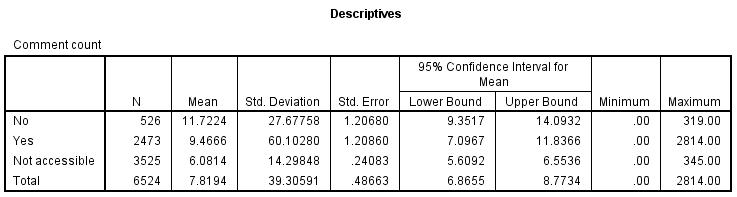
Description automatically generated



And the difference between average likes of posts in which text exists and average likes of posts in which text doesn’t exist was statistically significant as the ANOVA p-value was 0.000 i.e., lower than alpha of 5%.

**Impact of ‘existence of text in the post’ on ‘comment count’:**

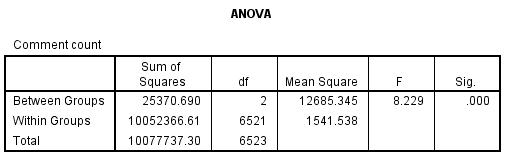
We will use ANOVA because independent variable ‘existence of text in the post’ is categorical while dependent variable ‘comment count’ is quantitative.



Average number of comments of posts in which text exists was 9.46 while average number of comments of posts in which text doesn’t exist was 11.72 i.e., posts in which text exists garnered less comments as compared to posts in which text doesn’t exist. This is also shown in the means plot below.

Chart

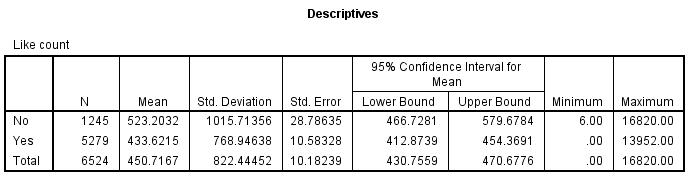
Description automatically generated



And the difference between average comments of posts in which text exists and average comments of posts in which text doesn’t exist was statistically significant as the ANOVA p-value was 0.000 i.e., lower than alpha of 5%.

**Impact of ‘agreement between readability readings of the post’ on ‘like count’:**

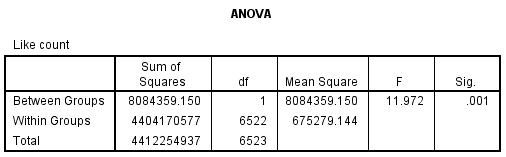
We will use ANOVA because independent variable ‘agreement between readability readings of the post’ is categorical while dependent variable ‘like count’ is quantitative.



Average number of likes of posts in which there is agreement between readability ratings was 433.6215 while average number of comments of posts in which there is no agreement between readability ratings was 523.2032 i.e., posts in which there is agreement between readability ratings garnered less comments as compared to posts in which there is no agreement between readability ratings. This is also shown in the means plot below.

Chart, line chart

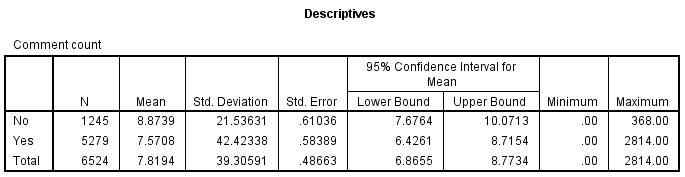
Description automatically generated



And the difference between average likes of posts in which there is agreement between readability readings and average likes of posts in which there is no agreement between readability readings was statistically significant as the ANOVA p-value was 0.001 i.e., lower than alpha of 5%.

**Impact of ‘agreement between readability readings of the post’ on ‘comment count’:**

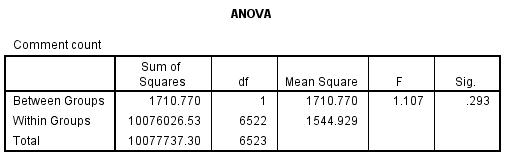
We will use ANOVA because independent variable ‘agreement between readability readings of the post’ is categorical while dependent variable ‘comment count’ is quantitative.



Average number of comments of posts in which there is agreement between readability ratings was 7.5708 while average number of comments of posts in which there is no agreement between readability ratings was 8.8739 i.e., posts in which there is agreement between readability ratings garnered less comments as compared to posts in which there is no agreement between readability ratings. This is also shown in the means plot below.

Chart

Description automatically generated

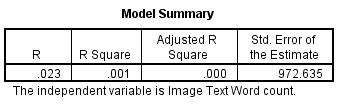


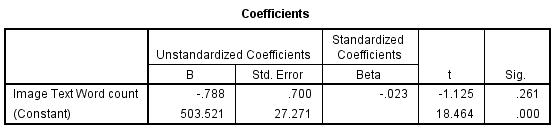
And the difference between average comments of posts in which there is agreement between readability readings and average comments of posts in which there is no agreement between readability readings was statistically insignificant as the ANOVA p-value was 0.293 i.e., higher than alpha of 5%.

**Impact of ‘image text wordcount’ on ‘like count’:**

We will use regression (linear, quadratic, and cubic) because independent variable ‘image text wordcount’ as well as dependent variable ‘like count’ are both quantitative.

Linear regression:

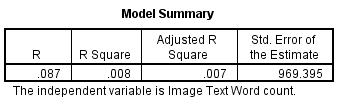


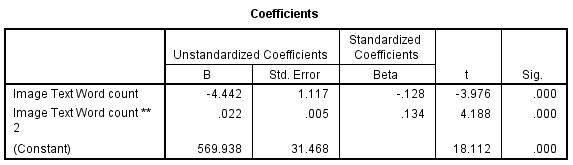


The R-squared of the linear model is 0.001 which implies that 0.1% of the variation in the dependent variable ‘like count’ is explained by the independent variable ‘wordcount of text in the image in the post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of linear model is -0.788 which implies that 1 word rise wordcount of text in the image in the post leads to a 0.788 unit fall in number of likes garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘wordcount of text in the image in the post’ on ‘like count’ is statistically insignificant at the 5% significance level as the p-value of 0.261 is higher than alpha of 0.05.

Quadratic regression:

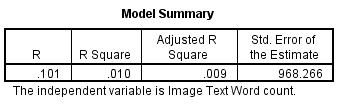


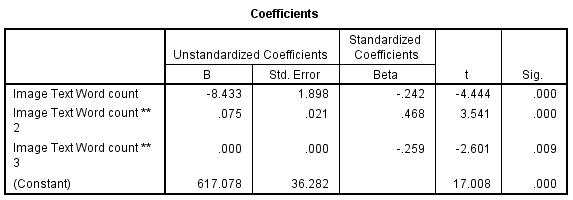


The R-squared of the linear model is 0.008 which implies that 0.8% of the variation in the dependent variable ‘like count’ is explained by the independent variables ‘wordcount of text in the image in the post’ and ‘square of wordcount of text in the image in the post’. This is still an extremely low value, and it indicates very poor goodness of fit.

The slope coefficients of quadratic model are -4.442 and 0.022 which implies that x word rise wordcount of text in the image in the post leads to a -4.442x + 0.022\*x^2 units change in number of likes garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘wordcount of text in the image in the post’ and ‘square of wordcount of text in the image in the post’ on ‘like count’ is statistically significant at the 5% significance level as the p-value of 0.000 is lower than alpha of 0.05.

Cubic regression:

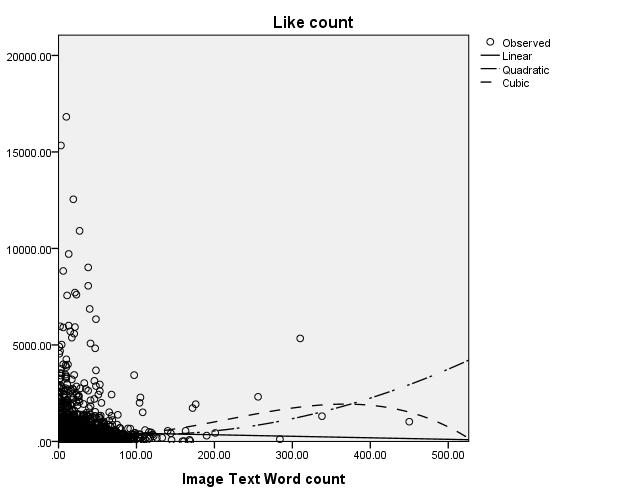




The R-squared of the cubic model is 0.010 which implies that 1% of the variation in the dependent variable ‘like count’ is explained by the independent variables ‘wordcount of text in image in the post’, ‘square of wordcount of text in image in the post’, and ‘cube of wordcount of text in image in the post’. This value is higher than that for linear and quadratic model, and hence it indicates best goodness of fit in comparison with other models. Thus, we reject linear and quadratic models and choose cubic model i.e., we conclude that relationship between ‘wordcount of text in image in the post’ and ‘like count’ is cubic.

The coefficient of ‘wordcount of text in image in the post’, ‘square of wordcount of text in image in the post’, and ‘cube of wordcount of text in image in the post’ is -8.433, 0.075 and 0.000, respectively. This implies that x words rise in wordcount of text in image in the post leads to a -8.433x + 0.075x^2 + 0.000x^3 change in number of likes garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘wordcount of text in image in the post’ on ‘like count’ is statistically significant at the 5% significance level as the p-value of ‘wordcount of text in image in the post’, ‘square of wordcount of text in image in the post’ and ‘cube of wordcount of text in image in the post’ are 0.000, 0.000, and 0.009 respectively i.e., lower than alpha of 0.05.

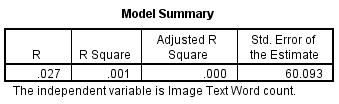
Curve estimation:

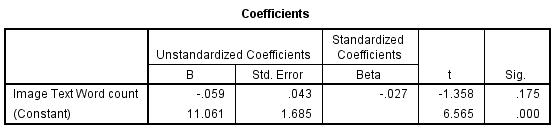


**Impact of ‘image text wordcount’ on ‘comment count’:**

We will use regression (linear, quadratic, and cubic) because independent variable ‘image text wordcount’ as well as dependent variable ‘comment count’ are both quantitative.

Linear regression:

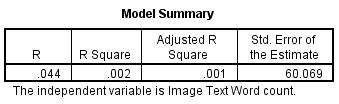


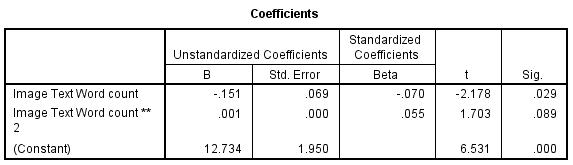


The R-squared of the linear model is 0.001 which implies that 0.1% of the variation in the dependent variable ‘comment count’ is explained by the independent variable ‘wordcount of text in the image in the post’. This is an extremely low value, and it indicates very poor goodness of fit.

The slope coefficient of linear model is -0.059 which implies that 1 word rise wordcount of text in the image in the post leads to a 0.059 unit fall in number of comments garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘wordcount of text in the image in the post’ on ‘comment count’ is statistically insignificant at the 5% significance level as the p-value of 0.175 is higher than alpha of 0.05.

Quadratic regression:

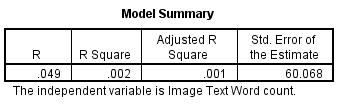


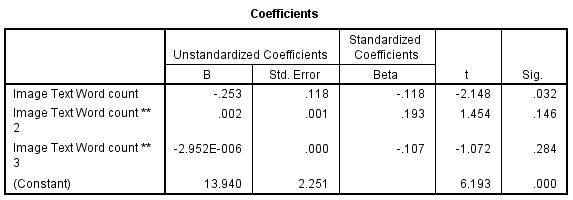


The R-squared of the linear model is 0.002 which implies that 0.2% of the variation in the dependent variable ‘comment count’ is explained by the independent variables ‘wordcount of text in the image in the post’ and ‘square of wordcount of text in the image in the post’. This is still an extremely low value, and it indicates very poor goodness of fit.

The slope coefficients of quadratic model are -0.151 and 0.001 which implies that x word rise wordcount of text in the image in the post leads to a -0.151x + 0.001\*x^2 units change in number of comments garnered by the post, all other factors remaining constant as in ceteris peribus. But the impact of ‘wordcount of text in the image in the post’ and ‘square of wordcount of text in the image in the post’ on ‘comment count’ is statistically insignificant at the 5% significance level as the p-value of 0.029 and 0.089 is higher than alpha of 0.05.

Cubic regression:





The R-squared of the cubic model is 0.002 which implies that 0.2% of the variation in the dependent variable ‘comment count’ is explained by the independent variables ‘wordcount of text in image in the post’, ‘square of wordcount of text in image in the post’, and ‘cube of wordcount of text in image in the post’. This value is higher than that for linear and quadratic model, and hence it indicates best goodness of fit in comparison with other models. Thus, we reject linear and quadratic models and choose cubic model i.e., we conclude that relationship between ‘wordcount of text in image in the post’ and ‘comment count’ is cubic.

The coefficient of ‘wordcount of text in image in the post’, ‘square of wordcount of text in image in the post’, and ‘cube of wordcount of text in image in the post’ is -0.253, 0.02 and 0.000, respectively. This implies that x words rise in wordcount of text in image in the post leads to a -0.253x + 0.02x^2 + 0.000x^3 change in number of comments garnered by the post, all other factors remaining constant as in ceteris peribus. And the impact of ‘wordcount of text in image in the post’ on ‘comment count’ is statistically significant at the 5% significance level as the p-value of ‘wordcount of text in image in the post’, ‘square of wordcount of text in image in the post’ and ‘cube of wordcount of text in image in the post’ are 0.032, 0.146, and 0.284 respectively i.e., lower than alpha of 0.05.

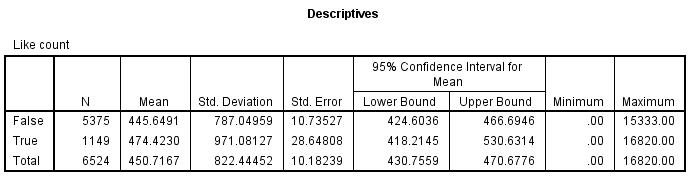
Curve estimation:

Graphical user interface

Description automatically generated with medium confidence

**Impact of ‘brand name existence in post’ on ‘like count’:**

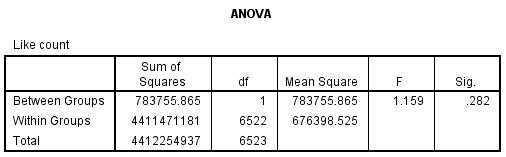
We will use ANOVA because independent variable ‘brand name existence in post’ is categorical while dependent variable ‘like count’ is quantitative.



Average number of likes of posts in which brand name exists was 474.4230 while average number of likes of posts in which brand name does not exist was 445.6491 i.e., posts in which brand name exists garnered more likes as compared to posts in which brand names doesn’t exist. This is also shown in the means plot below.

Chart, line chart

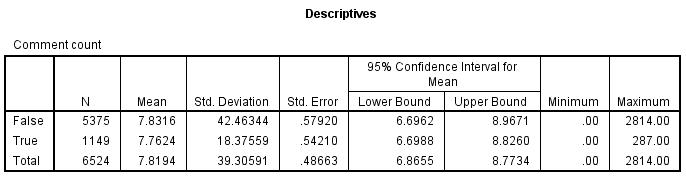
Description automatically generated



And the difference between average likes of posts in which there was brand name and average likes of posts in which there was no brand name was statistically insignificant as the ANOVA p-value was 0.282 i.e., higher than alpha of 5%.

**Impact of ‘brand name existence in post’ on ‘comment count’:**

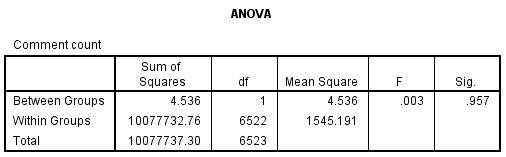
We will use ANOVA because independent variable ‘brand name existence in post’ is categorical while dependent variable ‘comment count’ is quantitative.



Average number of comments of posts in which brand name exists was 7.76 while average number of comments of posts in which brand name does not exist was 7.83 i.e., posts in which brand name exists garnered more comments as compared to posts in which brand names doesn’t exist. This is also shown in the means plot below.

Chart, line chart

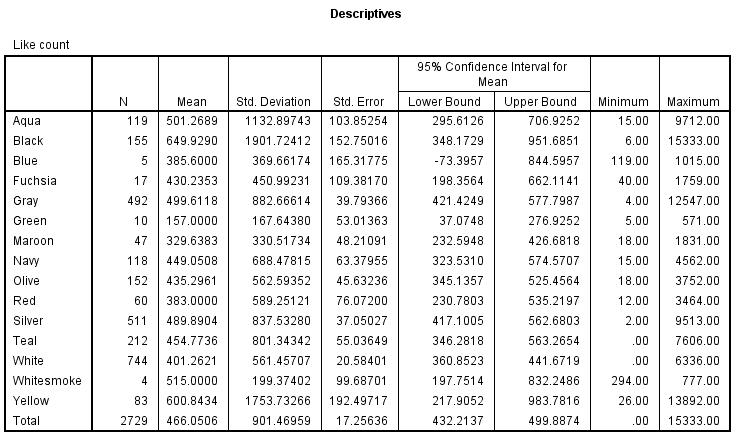
Description automatically generated



And the difference between average comments of posts in which there was brand name and average comments of posts in which there was no brand name was statistically insignificant as the ANOVA p-value was 0.957 i.e., higher than alpha of 5%.

**Impact of ‘color of post’ on ‘like count’:**

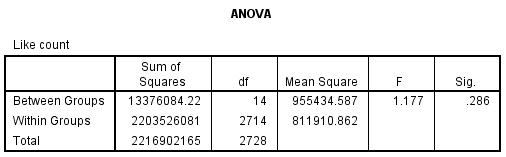
We will use ANOVA because independent variable ‘color of post’ is categorical while dependent variable ‘like count’ is quantitative.



Average number of likes was highest for posts in which dominant color was ‘black’ (mean = 649.92) while average number of likes was lowest for posts in which dominant color was ‘green’ (mean = 157.00).

Chart, line chart

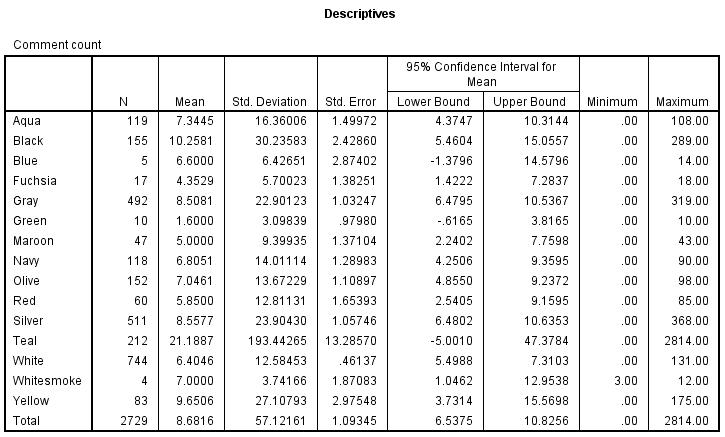
Description automatically generated



But the difference between average likes of posts in terms of different colors was statistically insignificant as the ANOVA p-value was 0.286 i.e., higher than alpha of 5%.

**Impact of ‘color of post’ on ‘comment count’:**

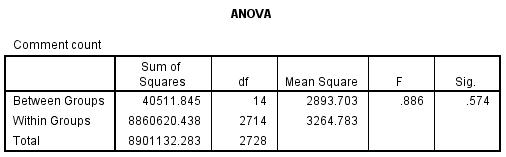
We will use ANOVA because independent variable ‘color of post’ is categorical while dependent variable ‘comment count’ is quantitative.



Average number of comments was highest for posts in which dominant color was ‘teal’ (mean = 21.1887) while average number of comments was lowest for posts in which dominant color was ‘green’ (mean = 1.600).

Chart, line chart

Description automatically generated



But the difference between average comments of posts in terms of different colors was statistically insignificant as the ANOVA p-value was 0.574 i.e., higher than alpha of 5%.