Twitch Gaming

Analysis of StreamTime, Following, and being a Twitch Partner as Factors of Growth in Viewers for Streamers

AUTHOR
Brianna Mateo,Gracia,Bareti,Nina Hernandez

PUBLISHED
December 20, 2022

Abstract

The purpose of our analysis is to examine the relationship that streaming time, followers, and being a Twitch partner have on the effect of your viewer growth as a Twitch streamer. Twitch Streamers are effectively full-time gamers. The best Twitch streamers make a living through paid subscriptions, donations, and influencer marketing. Using data from Kaggle, this Twitch Data set contains the Top 1000 Streamers from the past year on the streaming platform. We fit our multiple linear regression models with viewers gained as the response. We found evidence against our primary hypothesis that having a long stream time increases your growth in viewers, as we observed little to no relationship between stream time and viewers gained. Additionally, the following count does not appear

to change the relationship between stream time and viewers gained for all Twitch streamers, partnered or not. Overall, the strength of the correlation between the independent and dependent variables suggests that further analysis must be conducted to determine the extent of growth of our dependent variables.

Introduction

The future of gaming is streaming. Over the past decade, there has been a rise in live stream platforms available for viewers and streamers, such as YouTube Live and Twitch. Streaming has become a full-time position for many, notably Twitch, a leading streaming source, making up 1.8% of total U.S. Internet traffic and ranking fourth in live streaming activity.[1]

The streaming platform is aimed towards social entertainment, allowing individuals to live stream video content to an online audience,reaching 45 billion minutes watched per month in 2018 alone.[2] Prior studies on Twitch live streaming have focused on the behavioral attributes that lead to viewer growth.[3] These studies have found that people like to engage with streamers that emulate their interests, whether it be in a particular game or sport, and that viewers enjoy the social interaction gained from live platforms. [4] In this analysis, we are considering Twitch streamers' attributes associated with streaming growth. Specifically, we are considering the role of stream time and Twitch partnerships on the number of views gained. Being a Twitch Partner is decided by the Twitch platform itself.[5] As a Twitch Partner, you have access to special features that can help you grow your channel, though being a partner is decided by the Twitch platform itself. Our main question is whether streaming time alone predicts a streamer's growth in Twitch viewers or whether there are other important factors.

To support our main analysis question, we also explore:

- *Does streaming time and having a big following influence your growth in viewers?*
- *Does being a twitch partner and having a long stream time affect your stream time?*
- *Being a partner, stream time , and a big following impact your growth in viewers?*

Our primary hypothesis is that folks with a longer stream time have more viewers and the twitch partnership does not impact the number of viewers gained. Our secondary hypothesis is that a large follower base influences the number of viewers gained.

Methods

Data

This study uses data from Kaggle, a Google subsidiary and platform providing various data sets by data scientists and machine learning practitioners. The 'Top Streamers on Twitch' data set provides data from the top 1,000 Twitch streamers from 2020 and the most relevant information regarding these streamers specifically. The data was collected by a Twitch fan under the name Aayush Mishra.

Variables

This analysis is focused on what best predicts growth in viewers for Twitch streamers, thus ***'views_gained'*** will serve as our response variable. We immediately chose ***'stream_time'***(in minutes) as an explanatory variable hypothesizing that growth in your viewers may be associated with the number of minutes you are streaming. Continually, we were curious about the ***'followers'*** (numeric) variable and explored if the size of your following platform may or may not affect your growth in viewers. Additionally, we included a binary variable ***'partnered'*** indicating whether Twitch Streamers being partnered correlates with their growth in viewers this past year, recorded as 'TRUE' or 'FALSE'. We used Twitch Partnerships as a covariate variable, to explore whether being a Twitch Partner gains you more viewership in conjunction with your stream time, or are other variables offer significant predictions of the growth in viewership.

Model Analysis: *Bivariate Model*

Does streaming time alone predict your growth in Twitch viewers, or are there other important factors?

To find the multiple regression model that best explains the number of views gained, we conducted several model comparisons using tools such as Nested F-Tests, ANOVA tables, and tests for multicollinearity. Overall, four models were chosen to answer each of our questions.

The bivariate model that best predicts views gained using only streaming time will be our simplest model to compare to our other models to find what best affects growth in viewers in conjunction with stream time.

Results: *Bivariate Model*

A simple linear regression model was generated to determine whether a relationship exists between the amount of time a twitch user streams and the number of viewers they gain. The coefficients suggest that for a user who has streamed for zero minutes, the expected amount of views gained would be about 9,405,000, and for every additional minute a user streams, the number of views gained would be approximately 18.78 on average. The intercept and slope coefficients have p-values less than 0.05, indicating both are statistically significant. However, the adjusted R-squared value is extremely weak (0.00315). The low r-squared value suggests the amount of variance accounted for in the data is not extensively explained by a user's streaming time. Nonetheless, the data supports our initial hypothesis of a relationship between stream time and views gained by a twitch user, despite it being a weak correlation. We used our bivariate model to compare to our later models.

Model Analysis: *Covarite Model 1*

How does having a big following affect growth after accounting for your stream time?

An interaction model with predictors, **'streamtime'** and **'followers'** was created to answer whether an increase in followers and streaming time affects views gained. The interaction was necessary for predicting viewership growth based on the results of comparing other models.

Results: *Covarite Model 1*

Since the linear model was not a good fit for the data, an interaction model was run to determine the relationship between a twitch user's streaming time and another predictor: the number of followers a user has. The model produced has four coefficients: intercept, stream_time, followers, and stream_time: followers – the interaction term. The "stream_time" coefficient suggests for each additional minute spent streaming, the number of views gained for a streamer with zero followers is expected to be about 2.201. In comparison, the followers coefficient

represents a 3.67 increase in the number of views gained for every additional follower a user has if they have streamed for zero minutes. The interaction term (4.55e-05) can be interpreted as the change in slope on the “stream_time” coefficient with each increase in followers and vice versa. The model has an R-squared value of 0.097, and the interaction and followers coefficients are statistically significant. However, the “stream_time” coefficient has a high p-value (0.835), indicating that stream time may not be a necessary predictor for this model.

Model Analysis: *Covarite Model 2*

How does being partnered affect your growth after accounting for your stream time?

For our second cpvariate model, between **‘streamtime’** and **‘partnered’** we found that the increase in predictive accuracy in an interaction model is not statistically significant enough to add to the models, thus resulting in an additive model. An additive model was run to answer our second subquestion about whether the number of views gained can be explained by streaming time and the difference between being partnered with Twitch.

Results: *Covarite Model 2*

The model’s adjusted R-squared value of 0.003 and the first coefficient stream_timee, (18.71) represents the number of views gained for each minute increase in the streaming total, regardless of whether a user is partnered with the twitch streaming site. The partneredTRUE coefficient represents the difference in the number of views gained for a user with a partnership with twitch compared to a user without a partnership. The model suggests that a twitch user with a partnership loses about -5.916e+06 views more than twitch users who do not have a partnership. PartneredTRUE is the only coefficient within the model that is not statistically significant (0.27).

Model Analysis: *Covarite Model 3*

How does having a big following and being a Twitch partner affect your viewer growth after accounting for your stream time?

Lastly, we chose our other covariate model using ***‘streamtime’*** ,

'followers,' and ***'partnered'*** as response variables. We found that the interactive effect of ***'streamtime'*** and ***'followers'***, combined with ***'partnered',*** best answers its respective sub-question for our research, based on the results of the nested F-Tests, and model comparisons using ANOVA.

Results:*Covarite Model 3*

To determine whether stream time, followers acquired, and partnership status all influenced the number of views gained, a mixed interaction-additive model was created where stream time and followers interacted. The adjusted R-squared value was 0.099, and the partneredTRUE and stream_time: followers coefficient were statistically significant. A nested F-test was constructed to test if the slope depends on partnership status using the interaction model with stream time and followers as predictor variables and a full model with partnered, stream time, and followers as predictor variables. The nested F-test returns a p-value of 0.081, suggesting the slope of stream time and followers is not dependent on partnership and does change the model's effectiveness. Therefore, the complexity of the mixed model was unnecessary, supporting our second hypothesis that partnership does not impact the number of viewers gained for a twitch user.

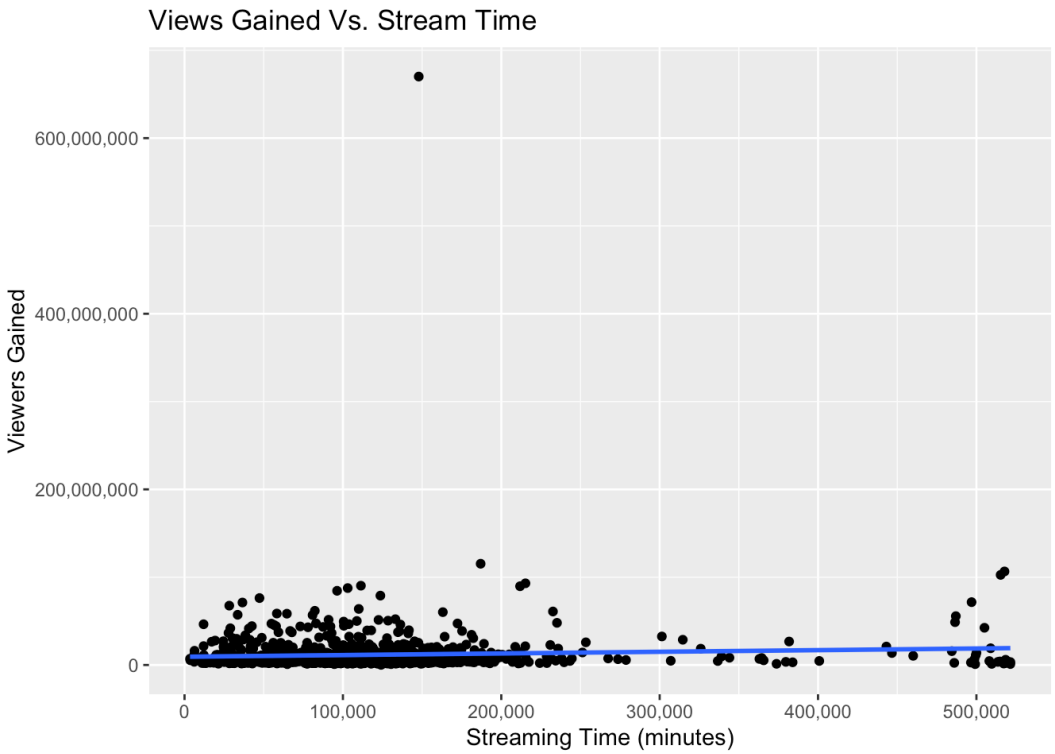


Figure 1: Plot of Views Gained Vs. Stream Time

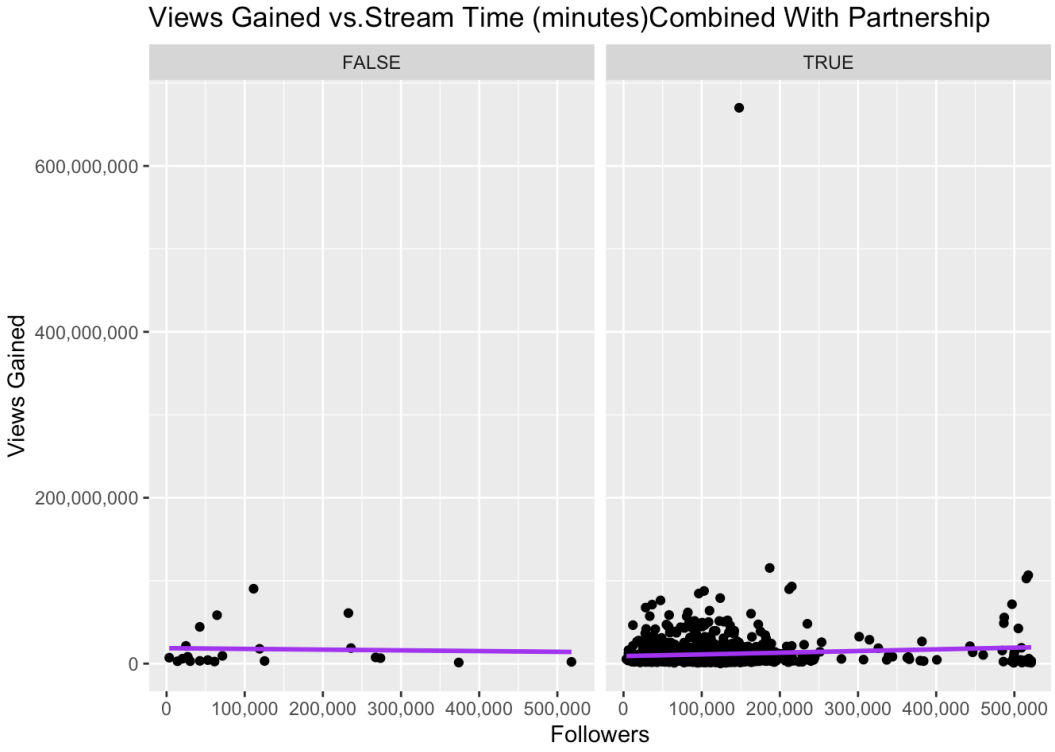


Figure 2: Plot of Views Gained Vs. Stream Time & Partnership

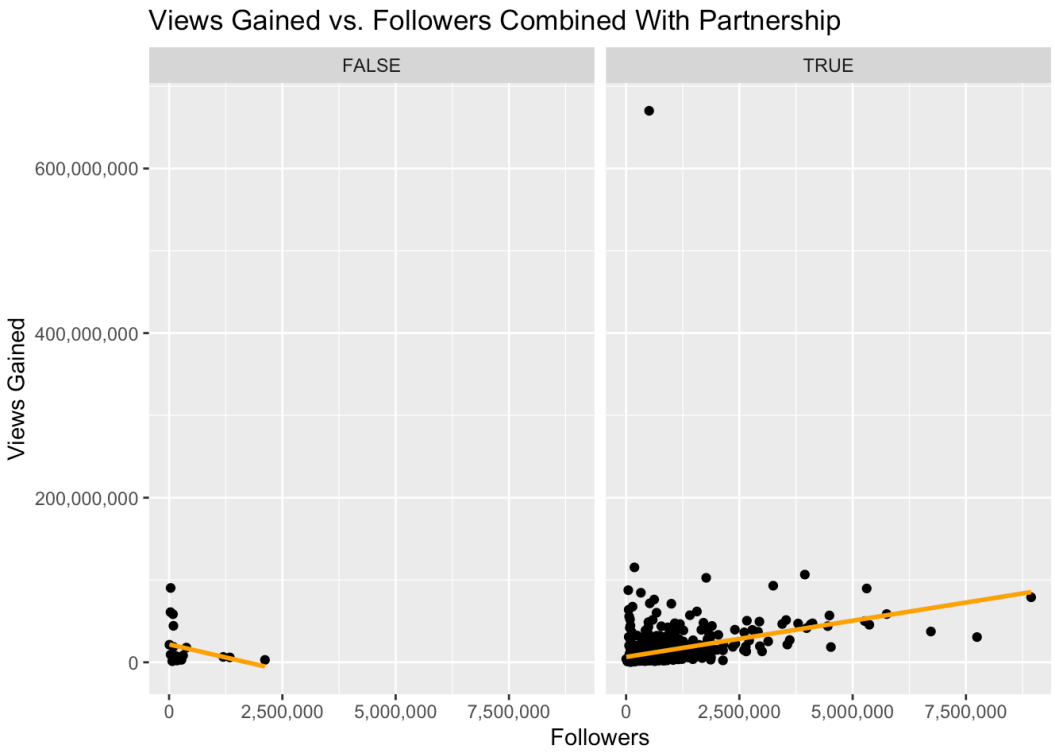


Figure 3: Plot of Views Gained Vs.Followers & Partnership

Discussion

The simple linear model used to investigate the extent of streaming time on the number of views gained showed a weak correlation between the two variables. The p-value of the streaming time coefficient is evidence that supports a relationship between a twitch user's streaming time and the number of views gained exists. However, the R-squared value strongly indicates that streaming time alone does not predict the number of viewers gained. Therefore, we cannot use this model to predict the growth in viewers. Instead, there appear to be other variables that may have a more significant effect on twitch users' viewership status.

The interaction model used to investigate whether the number of followers and the number of minutes a twitch user streams affects the total number of views gained showed that the number of followers significantly affects views gained. In this model, the interaction and followers' coefficients were statistically significant; therefore, we can conclude that there is sufficient evidence of a relationship between followers and views gained. The interaction term between followers and streaming implies that as the amount of streaming time increases for a user, the views gained will differ

based on the number of followers the user starts with. The negative value for the effect of the interaction term implies the higher the streaming time is, the fewer followers a user has on average. Similarly, the higher the follower base, the less time a streamer streams for. In this context, the interaction makes sense since smaller streamers would spend more time streaming to gain more followers and views. However, the stream_time coefficient not being statistically significant, which contradicts the analysis of our first model. This strongly indicated this coefficient was not a necessary predictor for this model. Despite this, the increase in the R-squared value from the simple linear model to the interaction indicates the number of followers a user has is significant in explaining the number of views gained and more necessary as a predictor than streaming time. Therefore, evidence supports our second hypothesis that followers' basis holds a more significant influence on the number of viewers gained (compared to streaming time). Nonetheless, the models allude to a positive relationship between stream time and followers, with a minimal increase between the two variables.

In our first hypothesis, we believed there would be no difference in views gained between users partnered with twitch compared to those not. The data supported our hypothesis and suggested no difference in the number of views gained by a twitch user who is and is not partnered with the twitch streaming site (determined by our third model). The negative coefficient for partneredTRUE that was produced may be due to the number of ads and promotions users must do during their streams, which may result in them losing disinterested viewers.

Overall, there was enough evidence to determine the statistical significance of followers, streaming time, and partnership status regarding views gained. The models indicate that streaming time alone cannot predict the number of views gained by a twitch follower and other predictors such as followers. Additionally, our models support our hypothesis about seeing no difference in views gained by a user who is and isn't partnered with Twitch. However, the strength of the correlation between the independent and dependent variables suggests that further analysis must be conducted to determine the extent of growth of our dependent variable.

Limitations

The limitations of our analysis include the data lacking variability as it only held information regarding the top 1,000 streamers from only a year-long period. For a platform that has been functioning for 5+ years, this data set would've been more beneficial if it was collecting data over a greater period of time rather than just a year, given that our data set also lacks an unequal variance distribution.[6] Additionally, the data doesn't provide information on the demographic of the top 1,000 streamers being examined or the particular content the streamers provide to users. This limitation doesn't allow us to explore if any trends could be occurring midst a particular type of content and viewer growth.

Despite our models' inconclusive results, if there were more context regarding the content these streamers provide, it would provide a better understanding of how to interpret Twitch viewer growth best—creating more definitive results within our models.

References

Mu Hu, Mingli Zhang, Yu Wang(2017).Why do audiences choose to keep watching on live video streaming platforms? An explanation of dual identification framework.Computers in Human Behavior. Volume 75. Pages 594-606,ISSN 0747-5632, https://doi.org/10.1016/j.chb.2017.06.006

VanKempJ, M. P. (2019, July). The drivers of live-streaming engagement on Twitch - Tilburg University. Retrieved December 21, 2022, from http://arno.uvt.nl/show.cgi?fid=150183

Harpstead, Erik & Rios, Juan & Seering, Joseph & Hammer, Jessica. (2019). Toward a Twitch Research Toolkit: A Systematic Review of Approaches to Research on Game Streaming. 111-119. 10.1145/3311350.3347149.

Zorah Hilvert-Bruce, James T. Neill, Max Sjöblom, Juho Hamari(2018). Social motivations of live-streaming viewer engagement on Twitch, Computers in Human Behavior,Volume 84.Pages 58-67, ISSN 0747-5632 https://doi.org/10.1016/j.chb.2018.02.013.

Twitch revenue and usage statistics (2022). Business of Apps. (2022, September 6). Retrieved December 20, 2022, from

https://www.businessofapps.com/data/twitch-statistics/

Appendix

Final Models

Does streaming time alone predict your growth in viewers or are their other important factors?

a. Does streaming time and having a big following influence your growth in viewers?

b.Does being a twitch partner and having a long stream time affect your growth in viewers?

c. Being a partner, stream time , and a big following impact your growth in viewers?

F-Tests, Model Comparisons, Anova Tables

Does streaming time alone predict your growth in viewers or are their other important factors?

```
Call:
lm(formula = views_gained ~ stream_time, data = twitchdata)

Residuals:
      Min        1Q    Median        3Q       Max 
-18101637  -7540955  -4973026    495257 657955435 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 9.405e+06  1.361e+06   6.912 8.53e-12 ***
stream_time 1.878e+01  9.215e+00   2.038   0.0418 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 24870000 on 998 degrees of freedom
Multiple R-squared:  0.004144,  Adjusted R-squared:  0.003146 
F-statistic: 4.152 on 1 and 998 DF,  p-value: 0.04184
```

a. Does streaming time and having a big following influence your growth in

viewers?

```
Call:
lm(formula = views_gained ~ stream_time + followers, data =
twitchdata)

Residuals:
      Min        1Q    Median        3Q       Max
-41907608  -5577263  -3055065    567756 658287718

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.466e+06  1.451e+06   2.388  0.01713 *
stream_time 2.636e+01  8.877e+00   2.970  0.00305 **
followers   8.815e+00  9.421e-01   9.357  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 23850000 on 997 degrees of freedom
Multiple R-squared:  0.08453,   Adjusted R-squared:  0.08269
F-statistic: 46.03 on 2 and 997 DF,  p-value: < 2.2e-16
```

b.Does being a twitch partner and having a long stream time affect your growth in viewers?

```
Analysis of Variance Table

Model 1: views_gained ~ stream_time + partnered
Model 2: views_gained ~ stream_time * partnered
  Res.Df        RSS Df  Sum of Sq      F Pr(>F)
1    997 6.1635e+17
2    996 6.1604e+17  1 3.0941e+14 0.5002 0.4796
```

c. Being a partner, stream time , and a big following impact your growth in viewers?

```
Call:
lm(formula = views_gained ~ (stream_time + followers) *
partnered,
    data = twitchdata)

Residuals:
```

```
      Min        1Q    Median        3Q       Max
-43428913  -5361044  -2835588    677272 658440949

Coefficients:
                            Estimate Std. Error t value
Pr(>|t|)
(Intercept)                2.589e+07  8.597e+06   3.011
0.00267 **
stream_time               -2.728e+01  4.030e+01  -0.677
0.49857
followers                 -1.494e+01  1.040e+01  -1.437
0.15111
partneredTRUE             -2.292e+07  8.723e+06  -2.628
0.00873 **
stream_time:partneredTRUE  5.511e+01  4.131e+01   1.334
0.18246
followers:partneredTRUE    2.401e+01  1.044e+01   2.300
0.02168 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 23790000 on 994 degrees of freedom
Multiple R-squared:  0.09199,   Adjusted R-squared:  0.08742
F-statistic: 20.14 on 5 and 994 DF,  p-value: < 2.2e-16


Call:
lm(formula = views_gained ~ (stream_time * followers) +
partnered,
    data = twitchdata)

Residuals:
      Min        1Q    Median        3Q       Max
-35697598  -5429655  -3201308    620076 658247889

Coefficients:
                        Estimate Std. Error t value Pr(>|t|)
(Intercept)            1.520e+07  5.254e+06   2.893   0.0039
**
stream_time            1.747e+00  1.055e+01   0.166   0.8685
followers              3.652e+00  1.551e+00   2.354   0.0188 *
partneredTRUE         -8.911e+06  5.105e+06  -1.745   0.0812 .
stream_time:followers  4.632e-05  1.097e-05   4.222 2.65e-05
***
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 23640000 on 995 degrees of freedom
Multiple R-squared:  0.1028,    Adjusted R-squared:  0.09924
F-statistic: 28.51 on 4 and 995 DF,  p-value: < 2.2e-16

Analysis of Variance Table

Model 1: views_gained ~ (stream_time + followers) + partnered
Model 2: views_gained ~ (stream_time * followers) * partnered
  Res.Df        RSS Df  Sum of Sq      F    Pr(>F)
1    996 5.6590e+17
2    992 5.5271e+17  4 1.3191e+16 5.9189 0.0001043 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Analysis of Variance Table

Model 1: views_gained ~ (stream_time + followers) * partnered
Model 2: views_gained ~ (stream_time * followers) + partnered
  Res.Df        RSS Df  Sum of Sq F Pr(>F)
1    994 5.6267e+17
2    995 5.5595e+17 -1 6.7263e+15

Analysis of Variance Table

Model 1: views_gained ~ (stream_time * followers) * partnered
Model 2: views_gained ~ (stream_time * followers) + partnered
  Res.Df        RSS Df   Sum of Sq      F Pr(>F)
1    992 5.5271e+17
2    995 5.5595e+17 -3 -3.2339e+15 1.9347 0.1223


Call:
lm(formula = views_gained ~ (stream_time * followers) *
partnered,
    data = twitchdata)

Coefficients:
                       (Intercept)
stream_time
                         2.344e+07
5.750e+01
                         followers
partneredTRUE
```

```
                          -1.722e+00
-1.747e+07
              stream_time:followers
stream_time:partneredTRUE
                          -6.425e-04
-5.392e+01
            followers:partneredTRUE
stream_time:followers:partneredTRUE
                           5.759e+00
6.869e-04


Call:
lm(formula = views_gained ~ (stream_time * followers) +
partnered,
    data = twitchdata)

Coefficients:
          (Intercept)            stream_time
followers
            1.520e+07              1.747e+00
3.652e+00
        partneredTRUE  stream_time:followers
           -8.911e+06              4.632e-05


Call:
lm(formula = views_gained ~ (stream_time * followers) *
partnered,
    data = twitchdata)

Residuals:
      Min        1Q    Median        3Q       Max
-36410463  -5343834  -3111370    662570 658246695

Coefficients:
                                    Estimate Std. Error t
value Pr(>|t|)
(Intercept)                         2.344e+07  8.723e+06
2.687  0.00734 **
stream_time                         5.750e+01  7.482e+01
0.769  0.44231
followers                          -1.722e+00  1.427e+01
-0.121  0.90398
partneredTRUE                      -1.747e+07  8.876e+06
```

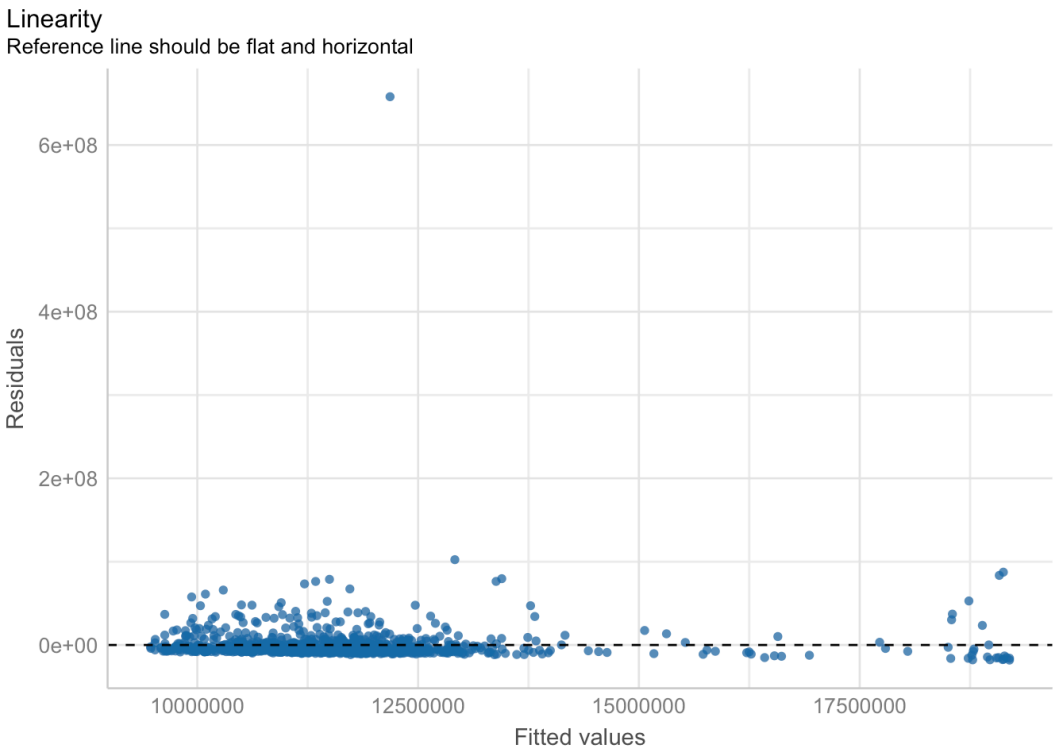
```
-1.968  0.04939 *
stream_time:followers               -6.425e-04  4.793e-04
-1.341  0.18033
stream_time:partneredTRUE           -5.392e+01  7.560e+01
-0.713  0.47590
followers:partneredTRUE              5.759e+00  1.435e+01
0.401  0.68834
stream_time:followers:partneredTRUE  6.869e-04  4.794e-04
1.433  0.15224
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 23600000 on 992 degrees of freedom
Multiple R-squared:  0.1081,    Adjusted R-squared:  0.1018
F-statistic: 17.17 on 7 and 992 DF,  p-value: < 2.2e-16

Analysis of Variance Table

Model 1: views_gained ~ (stream_time * followers) * partnered
Model 2: views_gained ~ (stream_time * followers) + partnered
  Res.Df        RSS Df   Sum of Sq      F Pr(>F)
1    992 5.5271e+17
2    995 5.5595e+17 -3 -3.2339e+15 1.9347 0.1223

Analysis of Variance Table

Model 1: views_gained ~ (stream_time * followers) + partnered
Model 2: views_gained ~ (stream_time * followers) * partnered
  Res.Df        RSS Df  Sum of Sq      F Pr(>F)
1    995 5.5595e+17
2    992 5.5271e+17  3 3.2339e+15 1.9347 0.1223

          stream_time             followers
stream_time:followers
             1.449065              2.784521
3.055234

          stream_time             followers
partnered
             1.449946              2.784606
1.003600
stream_time:followers
             3.060156

stream_time   partnered
```

```
1.00005     1.00005
```

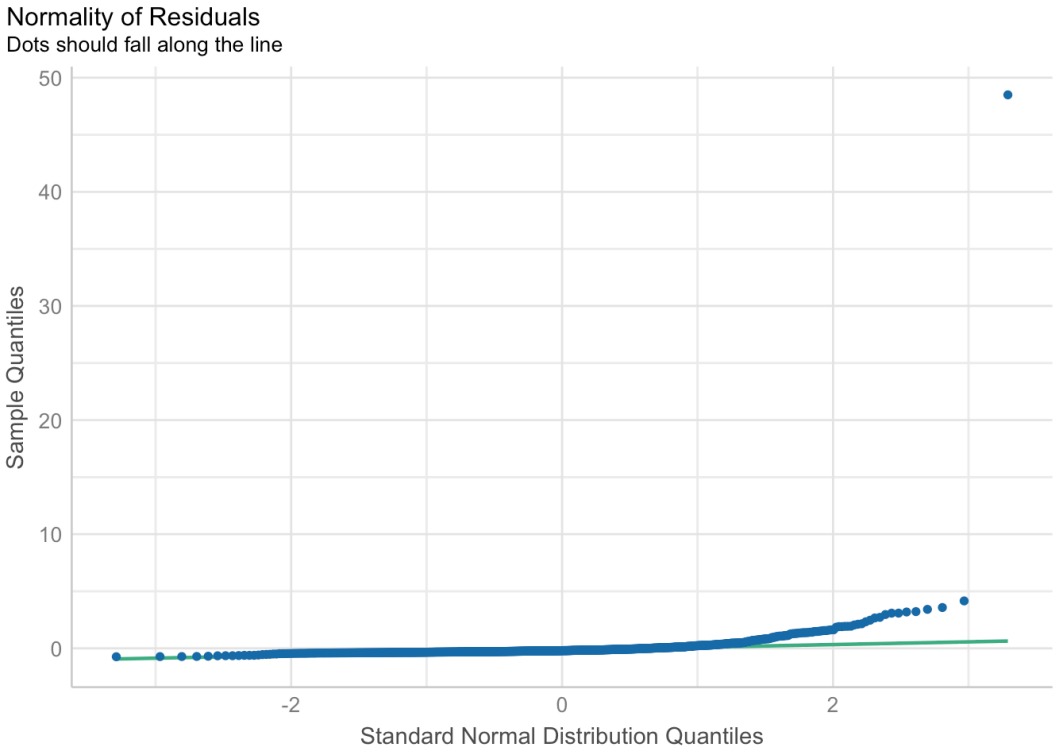
Checking Conditions:

Model 1: viewsgained_stream

$NCV

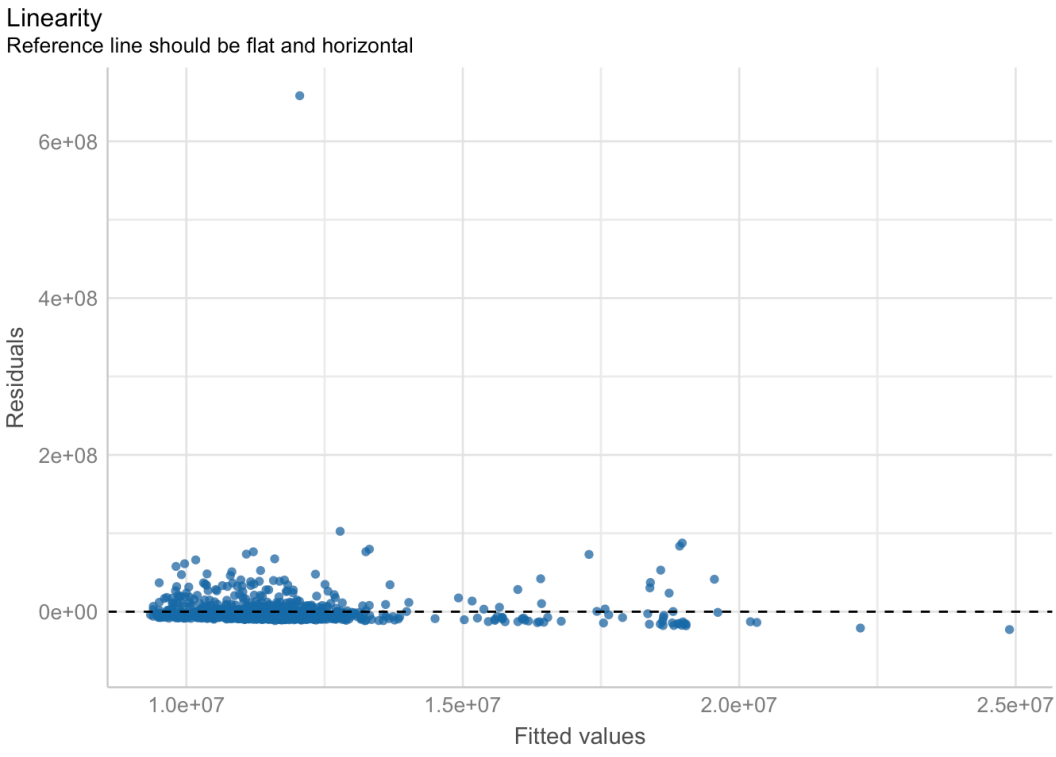


$QQ

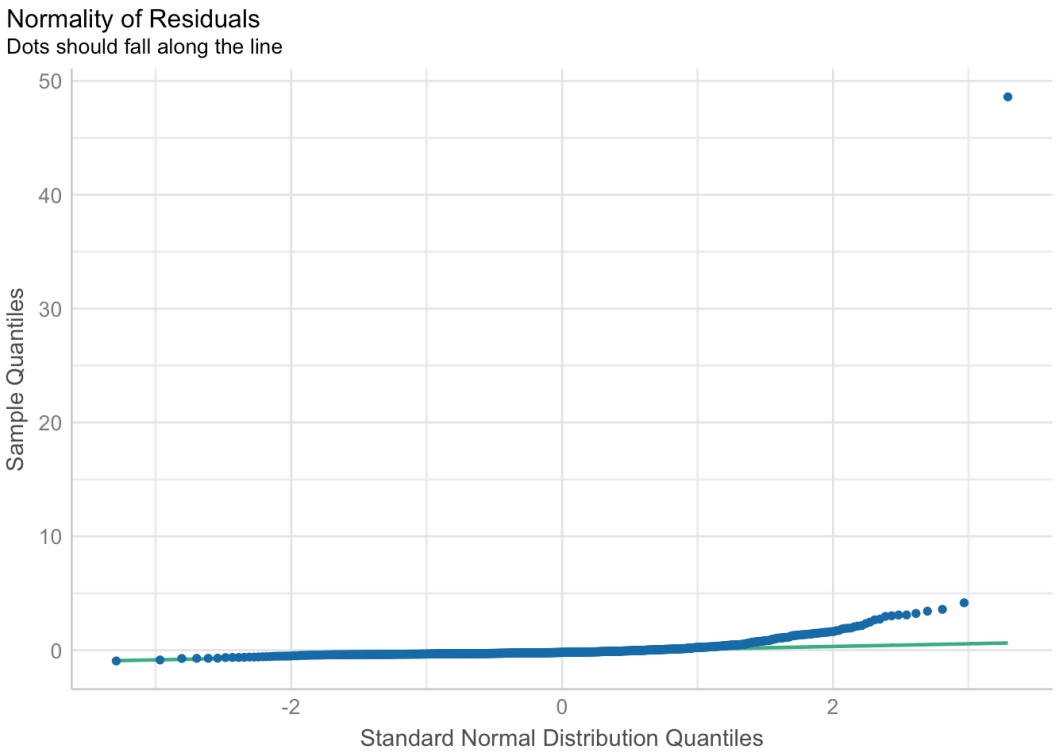


Model 2: stream_partnered_reduced

$NCV

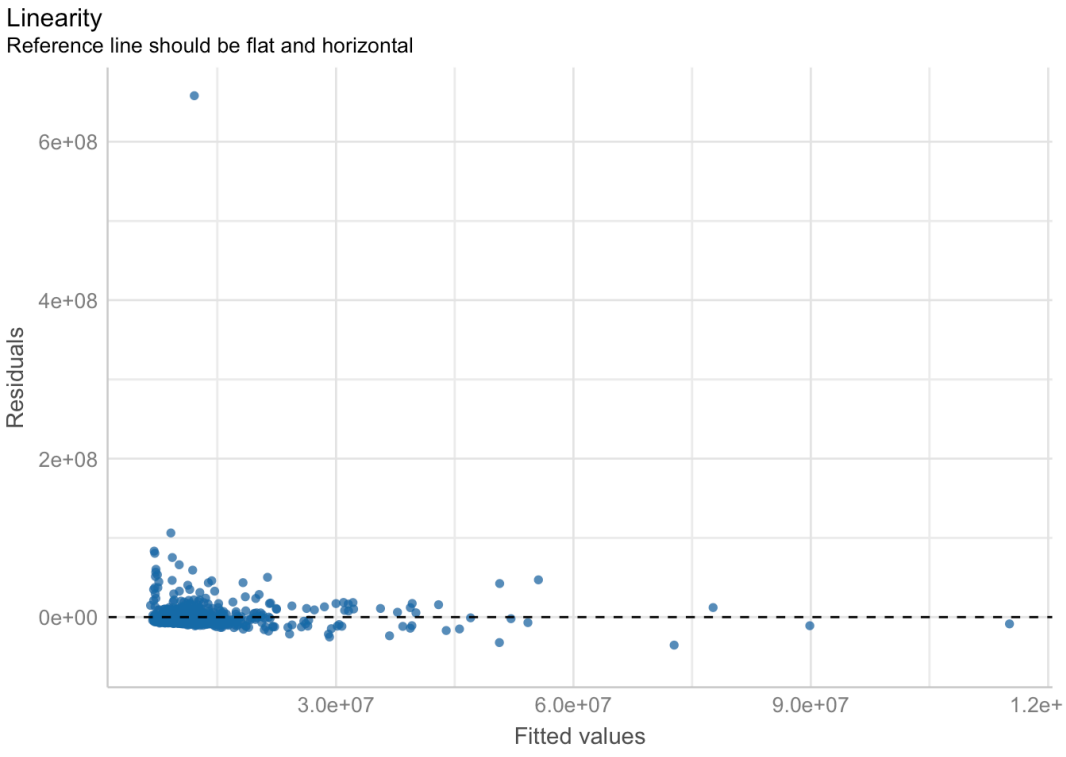


$QQ

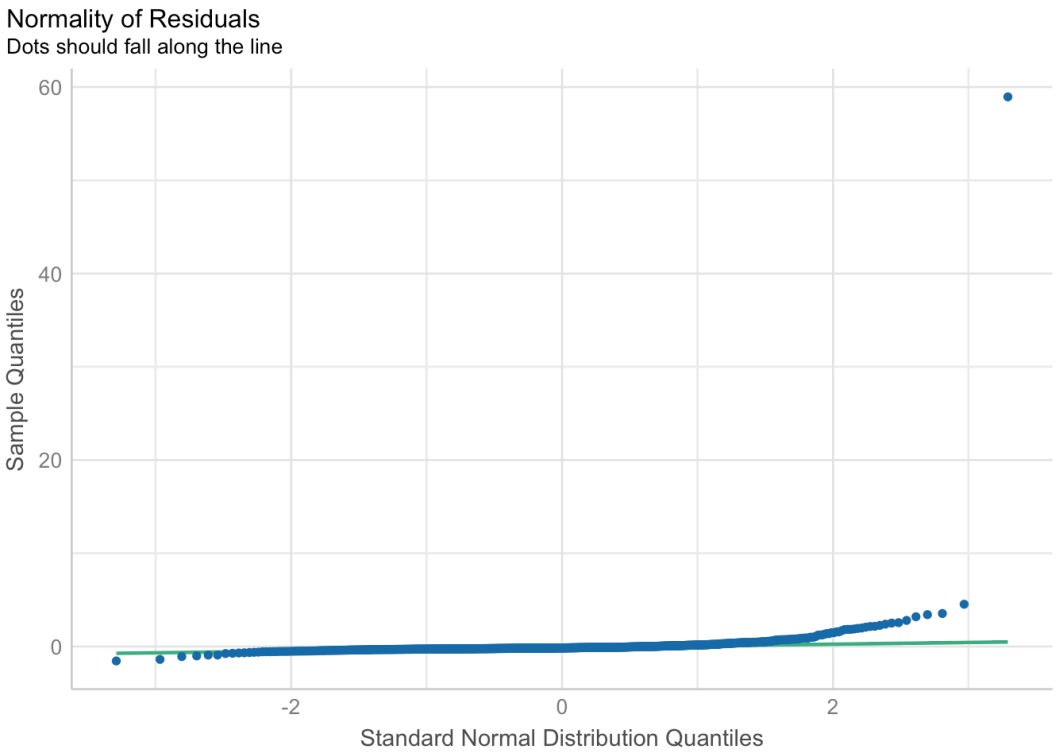


Model 3: viewsgained_stream_followers_full

$NCV

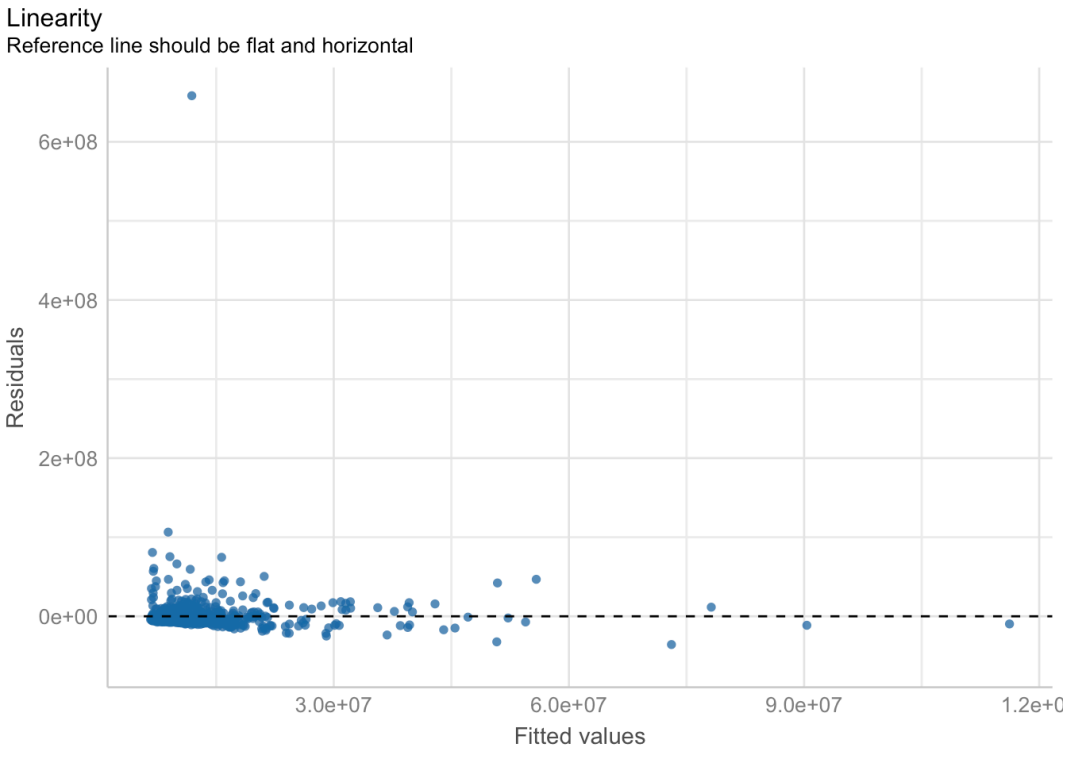


$QQ

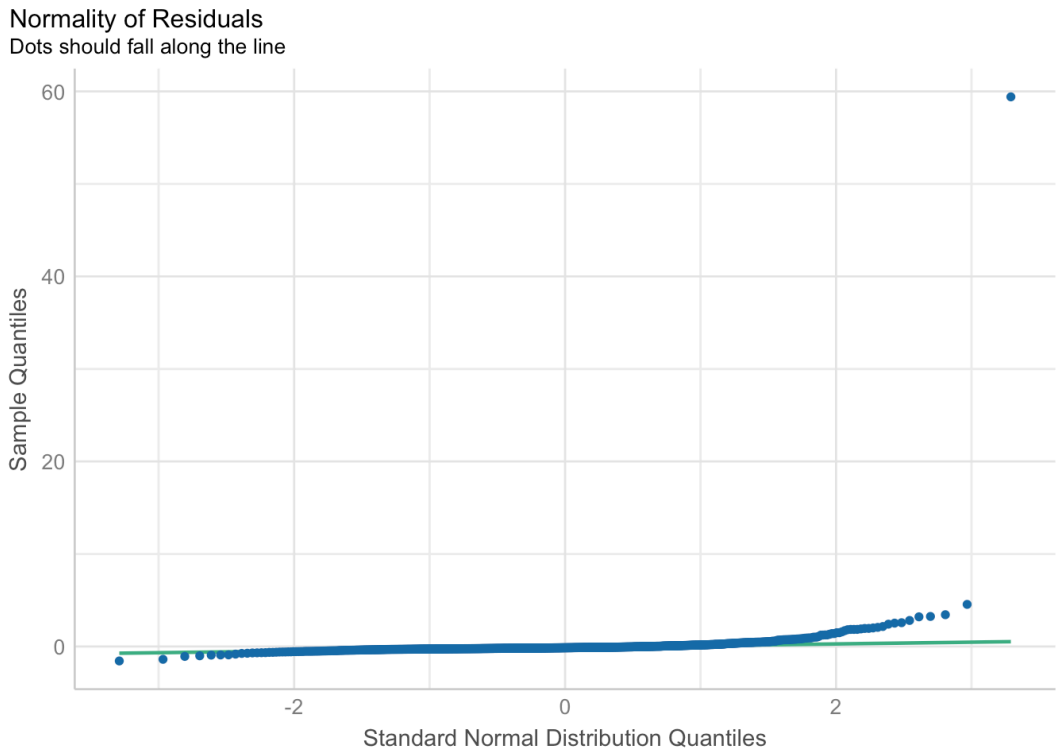


Model 4: stream_followers_partnered_mixed2

$NCV



$QQ



Footnotes

1. Twitch Gaming:Explanation↩
2. Engagement↩
3. Twitch Review↩
4. Social Motivation↩
5. Twitch Gaming:Explanation↩
6. Twitch Statistics↩