ESPORTS TOURNAMENT MACHINE LEARNING MODEL

Blair Martin

THE GOAL

- The goal is to build a machine learning model that can predict an increase in twitch viewers based on tournament data.
- This would increase Twitch's revenue due to the twitch ad revenue increase.
- Additional if we could gain further information on which locations and games are most effective.

THIS IS WHAT WE'RE WORKING WITH

The target: Twitch viewers

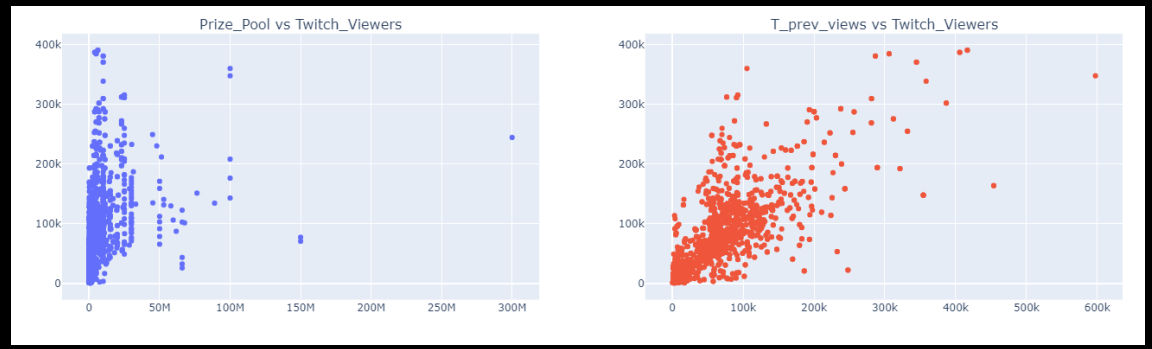
<u>The variables</u>

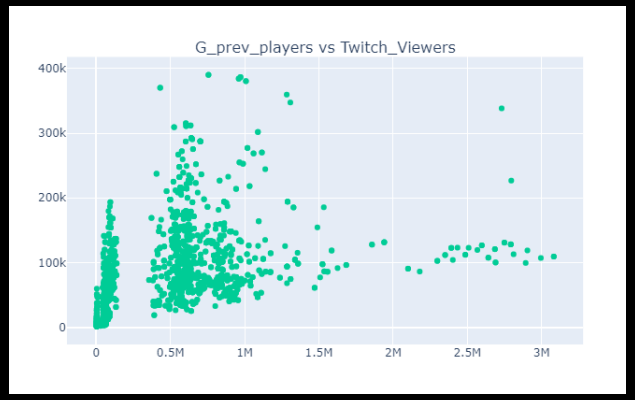
- Previous days twitch viewers
- Previous days game players
- Games(6)
- Countries(69)
- Players on day of the tournament
- Prize pool
- Number of days tournament was run < 5

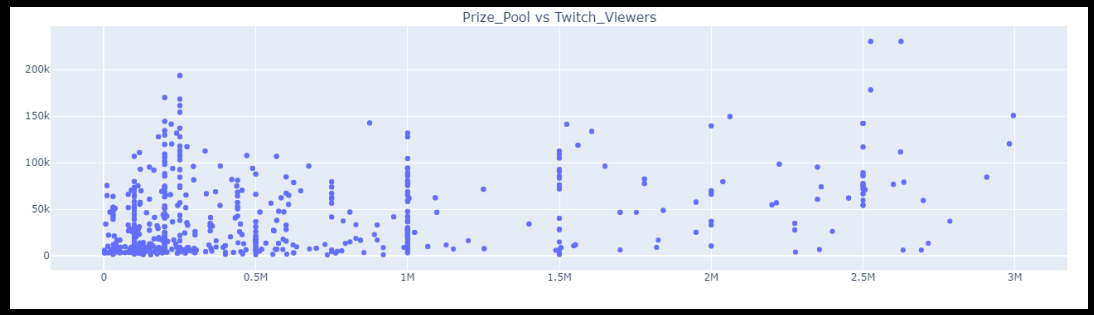
<u>Ideally</u>

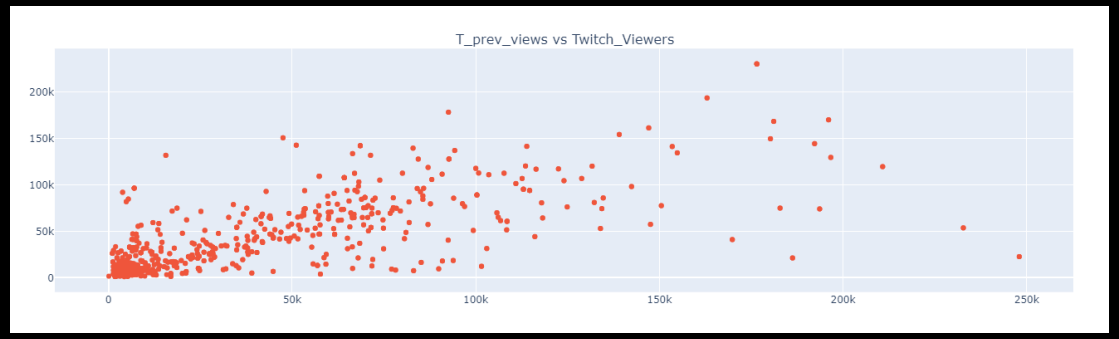
- Prize pool
- Countries
- The game

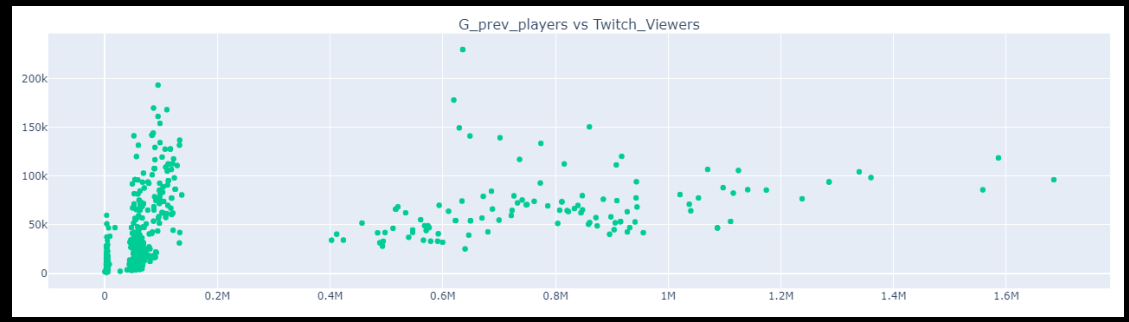
LINEAR ASSUMPTIONS

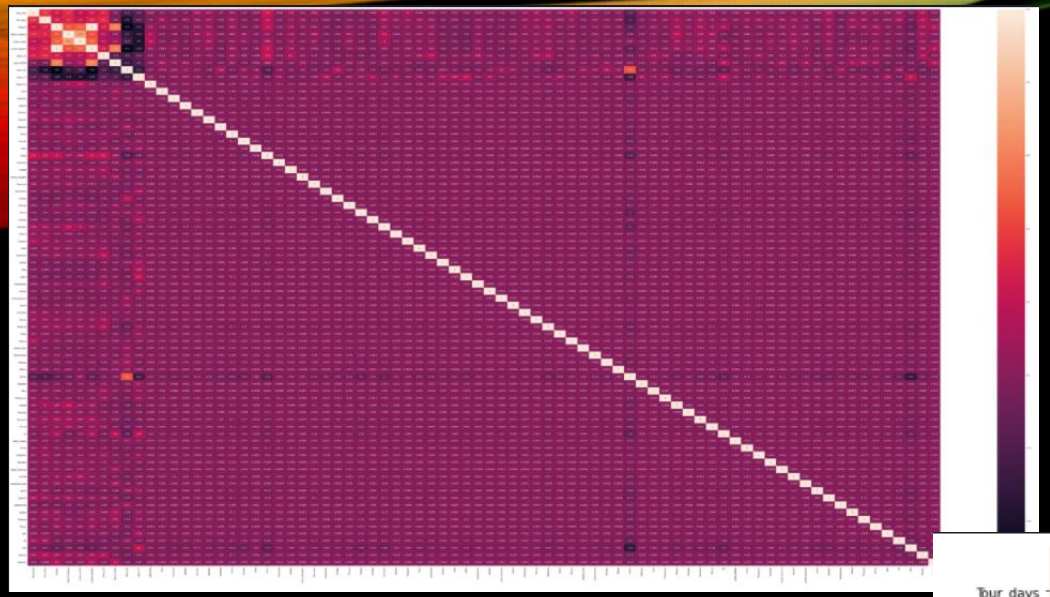










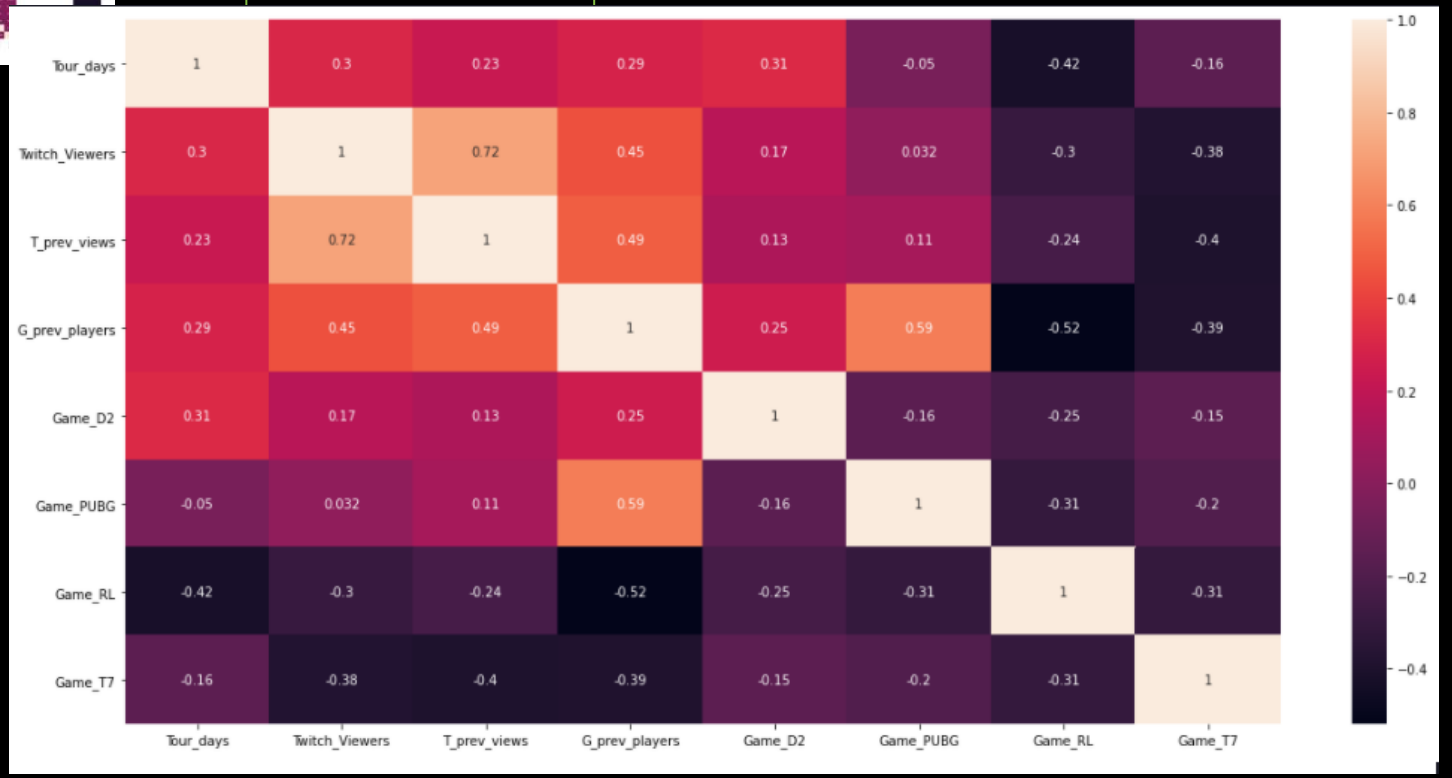


Full heatmap of all variables

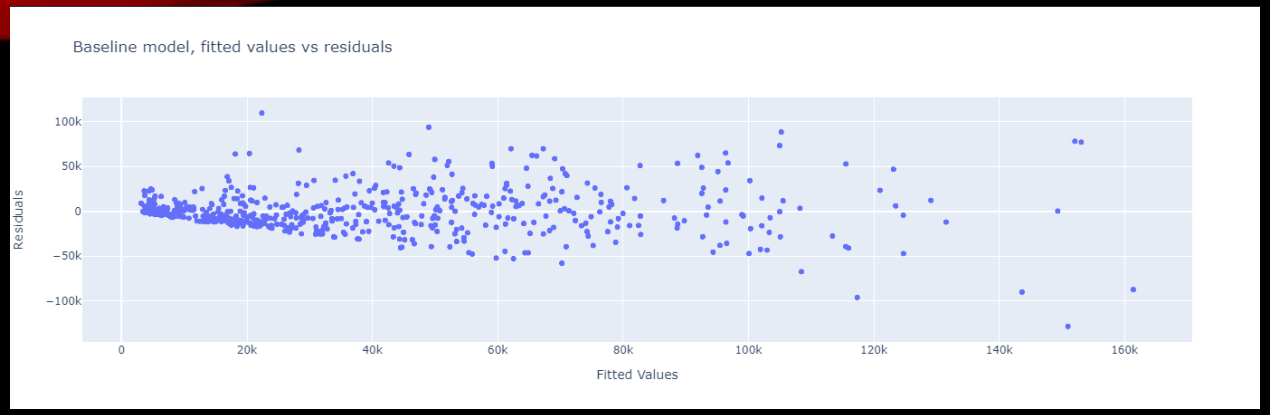
Final heatmap after dropping columns.

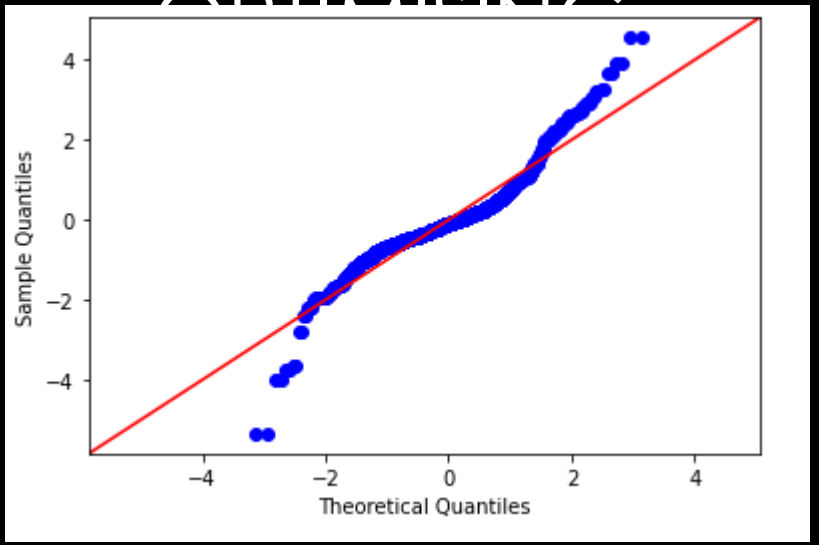
Players = Players before tournament

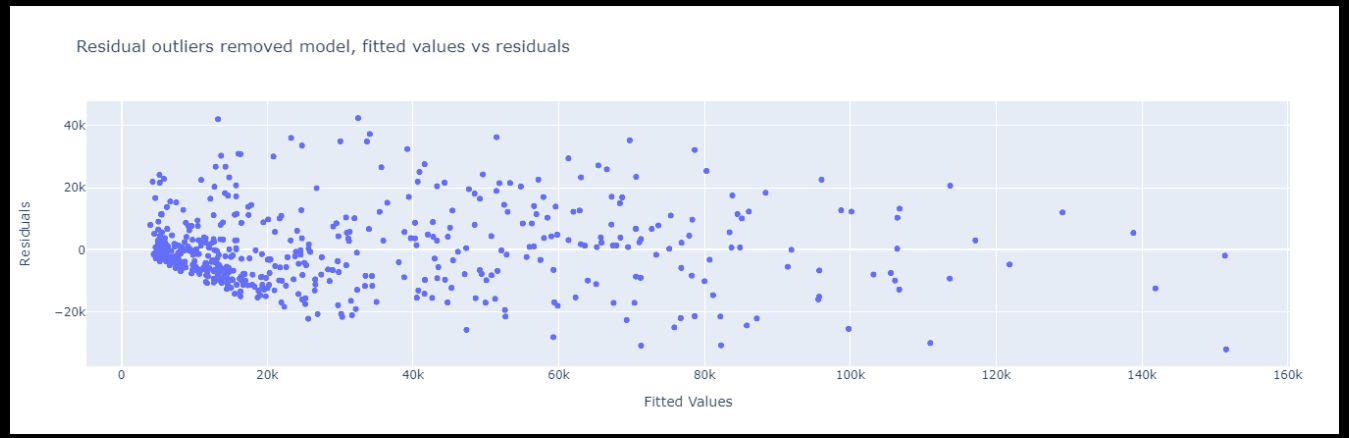
Removed Players

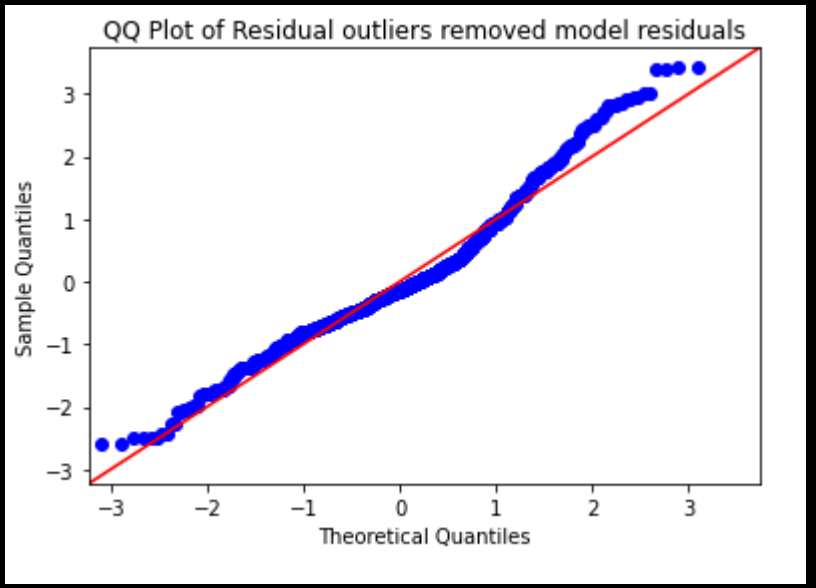


OPTIMISING









Final R^2 = 84.9%

COEFFICIENTS

The data

- Prize Pool: 0.0082:1
- Previous twitch viewers: 0.6759:1
- Game previous players: 0.0013:1
- Tekken: -3415 + 6285
- Belgium: 18320 + 6285
- France: 9005 + 6285
- Holland: 81120 + 6285
- Norway: 22050 + 6285

The results

- Prize pool has no significant impact
- Tournaments do boost twitch viewers
- Country matters!
- Tekken sucks!

OPTIMISATION RESULTS

OLS Regression Results

Dep. Variable:	Twitch_Viewers	R-squared:	0.634
Model:	OLS	Adj. R-squared:	0.631
Method:	Least Squares	F-statistic:	189.8
Date:	Tue, 01 Jun 2021	Prob (F-statistic):	1.05e-253
Time:	18:51:06	Log-Likelihood:	-13978.
No. Observations:	1215	AIC:	2.798e+04
Df Residuals:	1203	BIC:	2.804e+04
Df Model:	11		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	1.053e+04	1419.678	7.414	0.000	7740.680	1.33e+04
Prize_Pool	0.0102	0.001	8.338	0.000	0.008	0.013
T_prev_views	0.5590	0.020	28.534	0.000	0.521	0.597
G_prev_players	0.0111	0.003	4.283	0.000	0.006	0.016
Game_D2	9987.2203	4576.747	2.182	0.029	1007.927	1.9e+04
Game_T7	-8499.9770	1823.065	-4.662	0.000	-1.21e+04	-4923.237
Belgium	2.065e+04	6098.027	3.386	0.001	8684.586	3.26e+04
Europe	1.334e+04	5926.816	2.252	0.025	1716.348	2.5e+04
France	9331.2008	3933.562	2.372	0.018	1613.796	1.7e+04
Netherlands	5.534e+04	1.4e+04	3.963	0.000	2.79e+04	8.27e+04
Norway	1.552e+04	5759.991	2.694	0.007	4215.948	2.68e+04
Thailand	1.944e+04	8612.031	2.257	0.024	2543.549	3.63e+04

Omnibus:	147.439	Durbin-Watson:	0.708
Prob(Omnibus):	0.000	Jarque-Bera (JB):	800.005
Skew:	0.415	Prob(JB):	1.91e-174
Kurtosis:	6.888	Cond. No.	1.80e+07

OLS Regression Results

Dep. Variable:	Twitch_Viewers	R-squared:	0.849
Model:	OLS	Adj. R-squared:	0.848
Method:	Least Squares	F-statistic:	718.6
Date:	Tue, 01 Jun 2021	Prob (F-statistic):	0.00
Time:	18:51:06	Log-Likelihood:	-11199.
No. Observations:	1033	AIC:	2.242e+04
Df Residuals:	1024	BIC:	2.246e+04
Df Model:	8		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	6285.2572	770.467	8.158	0.000	4773.382	7797.132
Prize_Pool	0.0082	0.001	12.671	0.000	0.007	0.009
T_prev_views	0.6759	0.014	48.430	0.000	0.649	0.703
G_prev_players	0.0113	0.002	7.518	0.000	0.008	0.014
Game_T7	-3515.4165	965.570	-3.641	0.000	-5410.138	-1620.695
Belgium	1.832e+04	3621.480	5.058	0.000	1.12e+04	2.54e+04
France	9005.6665	2200.377	4.093	0.000	4687.903	1.33e+04
Netherlands	8.112e+04	8808.627	9.210	0.000	6.38e+04	9.84e+04
Norway	2.205e+04	3957.973	5.570	0.000	1.43e+04	2.98e+04

Omnibus:	82.884	Durbin-Watson:	0.912
Prob(Omnibus):	0.000	Jarque-Bera (JB):	108.383
Skew:	0.672	Prob(JB):	2.92e-24
Kurtosis:	3.844	Cond. No.	1.88e+07

ACCURACY R^2

- Train set: 84.5%
- Test Set: 76.%
- Validation set: 80%

- K-fold cross validation: 57%