KING OF A DECADE

ABSTRACT

In this paper, we describe the reasons of CS: GO's popularity.

We found some reasons to explain it and we will tell the reasons with visualitions of them. These reasons are, CS: GO is a free to play game, have low system requirements, it offers people many choices on their preferences and streaming success on streaming platforms. Then we decide to study on streaming platform's data. Our final goal is predicting CS: GO popularity for next months from today using Forecast Prophet regression.

Game Popularity Concept

Popularity is the extent to which a person, idea, place, item, or other concept is liked or status gained by other people. It can be caused by liking, mutual liking, interpersonal attraction, and similar factors. It can be caused by social status, dominance, superiority and similar factors. For games, its measured by asking two questions: How many people are playing the game and how many people watching.

Keywords

Game; Stream; CS:GO; Apex Legends; Rainbow Six Siege; Player; Viewer; Streaming Platform;

1. INTRODUCTION

A game's popularity is directly connected with its profit. If a company wants to make money from a game, it has to maintain the games popularity. From our research, game popularity is connected with people preferences and streaming platform success. We decided to work on data from the most popular game service Steam and most popular streaming platform Twitch.

2. Adding Libraries in Jupyter Notebook

As a first step, we decided the libraries that we will need for our project. Te libraries we use are: numpy, pandas, matplotlib.pyplot, os, seaborn, plotly.graph_objects. And we use subliblaries which are: make_subplots, RandomForestRegressor, model_selection, train_tset_split, GridSearchCV, cross_val_score, mean_squared_error, r2_score, mean_absolute_error.

SAMPLE:

```
Import numpy as np
Import matplot.pyplot as plt
From sklearn import model_section
From sklearn.metrics import r2_score
```

3. DATASETS

After the libraries are imported we started to find Datasets for CS:GO and its popular competitions player numbers and twitch viewer numbers for turkey and global. After that, we started to do reshape data to serve our subject best. We cleared the unnecessary data from them.

Finally we decided to use 2 dataset. First Valve_Player_Data.csv, Second Twitch_game_data.csv. They are both easy to use and accessible.

3.1 Valve_Player_Data.csv

We use this data set for Steam user data. This Dataset provides us average player of games, games profits, percentage gained and peak player numbers. We use this Dataset for regression.

	Month_Year	Avg_players	Gain	Percent_Gain	Peak_Players
0	September 2021	512350.920	268.96	+0.05%	942519
1	August 2021	512081.960	6014.60	+1.19%	802544
2	July 2021	506067.360	-43279.72	-7.88%	763523
3	June 2021	549347.080	-110541.81	-16.75%	929940
4	May 2021	659888.890	-63457.63	-8.77%	1087197

3.2 Twitch_game_data.csv

We use this Dataset for Twitch viewer data. This Dataset provide us some information about Twitch users. These are: Game names, Timeline, hours watched, hours streamed, peak viewer numbers, peak channels, Streamers, average viewer numbers, average channels viewers and average viewer ratio.

Rank	Game	Month	Year	Hours_watched	Hours_Streamed	Peak_viewers	Peak_channels
1	League of Legends	1	2016	94377226	1362044 hours	530270	2903
2	Counter-Strike: Global Offensive	1	2016	47832863	830105 hours	372654	2197
3	Dota 2	1	2016	45185893	433397 hours	315083	1100
4	Hearthstone	1	2016	39936159	235903 hours	131357	517
5	Call of Duty: Black Ops III	1	2016	16153057	1151578 hours	71639	3620

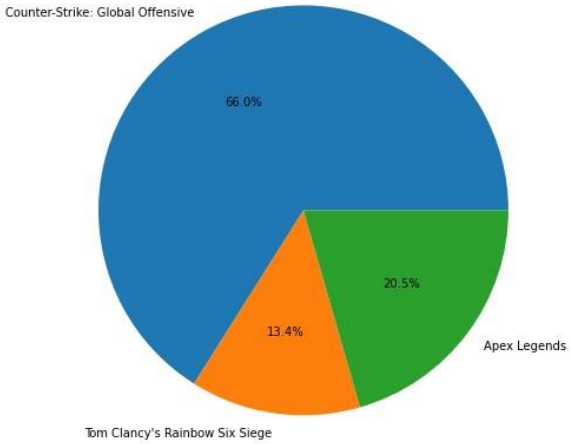
4. Data Visualize

When we are done with reshaping our data to their most suitable version for our project, we move into the our next step; visualizing the data. We

use some library functions for this. Our main goal when using visualization data was show that CS:GO's popularity over other games. We used player numbers for this. We used graphics to show it's popularity in a better and cleaner way.

4.1 Pie Chart

The pie chart is a pictorial representation of data that makes it possible to visualize the relationships between the parts and the whole of a variable. Learn how and when to use it.



4.2 Bar Chart

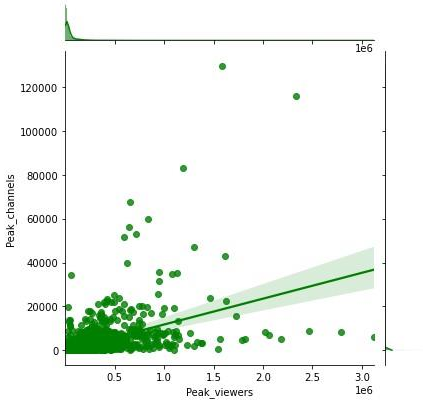
A bar chart displays categorical data with rectangular bars whose length or height corresponds to the value of each data point.

Bar charts can be visualized using vertical or horizontal bars. Bar charts are best used to compare a single category of data or several. When comparing more than one category of data, the bars can be grouped together to created a grouped bar chart.

Bar charts use volume to demonstrate differences between each bar. Because of this, bar charts should always start at zero. When bar charts do not start at zero, it risks users misjudging the difference between data values

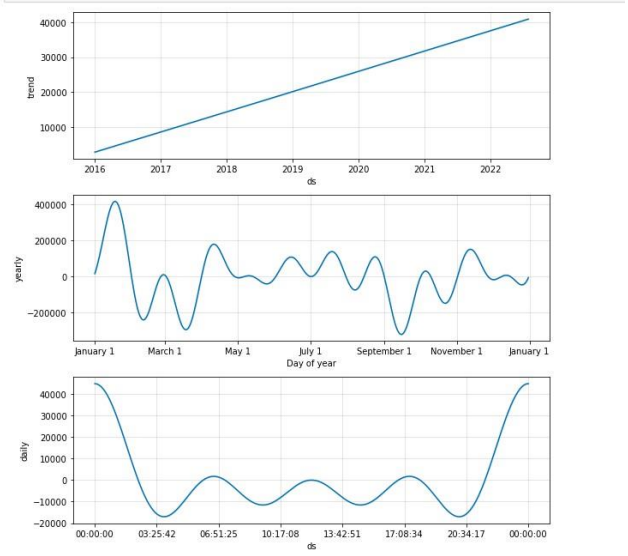
4.3 Scatterplot Graph

A scatter plot is a data visualization that displays the values of two different variables as points. The data for each point is represented by its horizontal and vertical position on the visualization. Additional variables can be encoded by labels, markers, color, transparency, size (bubbles), and creating 'small multiples' of scatter plots.



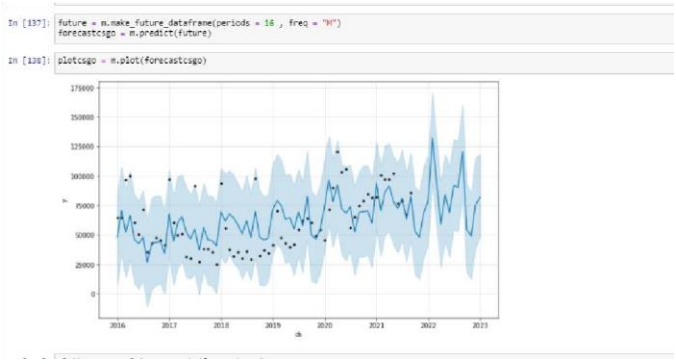
4.4 Line Chart

A line chart or line graph displays the evolution of one or several numeric variables. Data points are connected by straight line segments. It is similar to a scatter plot except that the measurement points are ordered (typically by their x-axis value) and joined with straight line segments. A line chart is often used to visualize a trend in data over intervals of time – a time series – thus the line is often drawn chronologically.



5. Regression with Prophet by Forecast

Prophet is a procedure for forecasting time series data based on an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects. It works best with time series that have strong seasonal effects and several seasons of historical data. Prophet is robust to missing data and shifts in the trend, and typically handles outliers well.



5.1 R^2

Compares your models predictions to the mean of the targets. Values can range from negative infinity (a very poor model) to 1. For example if all your model does is predict the mean of the targets, it's R^2 value would be 0. And if your model perfectly predicts a range of numbers it's R^2 value would be 1

```
In [24]: y_test_mean=np.full(len(y_test),y_test.mean())

In [25]: r2_score(y_test,y_test_mean)
Out[25]: 0.0

In [26]: r2_score(y_test,y_test)
Out[26]: 1.0
```

6. REFERENCES

1. Makarov, I., Savostyanov, D., Litvyakov, B., & Ignatov, D. I. (2017, July). Predicting winning team and probabilistic ratings in "Dota 2" and "Counter-Strike: Global Offensive" video games. In International Conference on Analysis of Images, Social Networks and Texts (pp. 183-196). Springer, Cham.
2. Lin, D., Bezemer, C. P., & Hassan, A. E. (2017). Studying the urgent updates of popular games on the steam platform. Empirical Software Engineering, 22(4), 2095-2126.
3. J., Cuadrado, F., Tyson, G., & Uhlig, S. (2015, December). Behind the game: Exploring the twitch streaming platform. In 2015 International Workshop on Network and Systems Support for Games (NetGames) (pp. 1-6). IEEE.
4. Huang, M. L., Huang, T. H., & Zhang, J. (2009, July). TreemapBar: Visualizing additional dimensions of data in bar chart. In 2009 13th International Conference Information Visualisation (pp. 98-103). IEEE.
5. Furtado, P. (2021). Epidemiology SIR with Regression, Arima, and Prophet in Forecasting Covid-19. In Engineering Proceedings (Vol. 5, No. 1, p. 52). Multidisciplinary Digital Publishing Institute.
6. Xie, C., Wen, H., Yang, W., Cai, J., Zhang, P., Wu, R., ... & Huang, S. (2021). Trend analysis and forecast of daily reported incidence of hand, foot and mouth disease in Hubei, China by Prophet model. Scientific reports, 11(1), 1-8.