8th Semester Project (UDP)

Sales Insights

Problem Statement

With more than 2.5 billion gamers from all over the world, more gamers are switching towards mobile gaming compared to traditional video game consoles which are slowly moving out of phase.

I'll create a predictive model on video gaming market revenue based on the data and also identify what are the most important features that affects the revenue of the video games.

Additionally, through analysis, I would like to identify the console among PS, XBOX, Nintendo, PlayStation is the console that will possibly generate the most sales and effectively maximizing profits and minimizing the cost for further game developments.

Executive Summary

With the data, I did some 'tidying up' on the dataset and to ensure that model is able to predict video games revenue to its best accuracy.

After which, I went on to analysis the importance of each feature which will affect the selling price, with the help of visualization such as Heatmap, Histograms and Scatterplot.

The ultimate outcome will be, with the prediction for the sales volume, I will use the data to convince stakeholders of the company to set aside an appropriate amount of budget for future game development, focus on building games based on the most popular video gaming platforms, effectively boost up and maximize the profit of the company while minimizing the cost.

EDA & Key Takeaways

North America region plays the most important part in determining the overall sales of the video gaming industry, which contributed to 40-50% of the market share. Followed by PAL regions, primarily Europe, Australia, Brazil, and regions of Africa and Asia.

There was a shift of trend of popular genre from Sports from late 1990s and slowly evolving to Action games being popular now, we can see that there is no one genre that could well sustain through the years.

There seems to be an increase of sales volume during the festive seasons October, November and declining from December compared to other months.

PlayStation had the greatest number of sales and dominated at the top because they had licensing on exclusive distribution on certain popular games and they have been around for a long time before Xbox was launched. In addition, Nintendo and PlayStation being a Japanese product, many Japanese game developer will only launch games based on PlayStation platform, resulting in Xbox being unpopular in Japanese market.

There seem to be a decline in the total sales over the past 10 years. One reason could be the rise of mobile gaming is taking away sales volume from console gaming and people would prefer to play-on-the-go for the convenience.

Modeling

Baseline

- The mean of total sales is used as a baseline score of around 400,000 for total sales.
- Our objective is to create a model that has the lowest RMSE score.

Preprocessing

- Standard Scaler was used to scale our model.

Evaluation

- We want to create a production model that can provide useful feature importance information.
- Linear Regression, Lasso, Ridge Regression and Random Forest were considered.
- Random Forest had the best score among all models.

Summary

The important features that affect overall total sales are critic score, release month of November followed by platform to be PlayStation.

Through this model, we are able to predict the total sales with a RMSE score of 660,000, what this means is that **our model's prediction is on average off by $660,000.** Given the better performance of the model, it is clear that we achieved the objective of our project successfully.

The ultimate outcome will be, with the details of the total sales prediction for the sales volume, I will use the data to convince stakeholders of the company to set aside an amount of budget for any new game development, focus on building games based on PlayStation and also allocating more marketing strategies and ensuring that the critics give a good score for our games.

Recommendation

Additional outside research found that video game downloadable contents and microtransactions to drive $4 billion in sales alone. Therefore, much consideration should be made to increase content with new updates and additional features in the game. This will greatly increase the sales revenue of the game developed as well.

With the genre of Action and Sports games being popular, perhaps we can brainstorm and focus on developing games that are based on those genres which will be of most interest to most gamers now. One of the ideas could be to create a game whereby it has to deal with stock market trading, business creation be it legal or illegal sort of business and enable players get to buy luxurious ranging from cars to even purchasing a soccer team which could be a linkup business opportunity with other sports games.

Additional consideration for future game development is that we can create a game which players get to play on their main tv game console and at the same time have an mobile game app which is linked. So, in this way, we are able to capture the upcoming trend of mobile gaming and ride the wave on the potential revenues that possible to be explode in the years to come.

Data Dictionary

- Name – the title of games, object
- Genre – genre of the game, object
- Platform – platform of the game released (i.e., PS, PS4, etc.), object
- Developer – developer of the game, object
- Vgchartz_Score – aggregate score compiled by Vgchartz staff, float
- Critic_Score – aggregate score compiled by independent games site, float
- User_Score – aggregate score compiled by gamers, float
- Total_Shipped – total shipments from manufacturers, float
- NA_Sales – sales in North America (in millions), float
- PAL_Sales – sales is Asia, Europe, Argentina, Brazil and Paraguay (in million), float
- JP_Sales – sales in Japan (in million), float
- Other_Sales - Sales in the rest of the world, float
- Release Date - Date of the game's release, datetime
- Last Update - Time the game was last updated, datetime