# Visualization of Video Game Sales Data Using Dash by Plotly

Mohit Pokharel, Sampurna K.C

Student, Department of Computer Science & Engineering

Kathmandu University, Dhulikhel, Nepal

pokhrelmohit@gmail.com  , sampurnakc19@gmail.com

**Abstract: This paper presents the visualization of the video game sales dataset. We have applied different visualization methods to obtain valuable insights into our data. This dataset contains a list of video games with sales greater than 100,000 copies. The data set encompasses a variety of information including the game's ranking of overall sales, name, platform of the game's release, year of the game's release, genre, publisher, and sales details. Using the powerful tool of data visualization and an interactable dashboard, this study aims to find the hidden relations and rules within the dataset.**
**We aim to find the connection between the game sales and the year of release and genre.**

## 1. INTRODUCTION

The Video Game Sales database contains a total of 16600 entries and has a list of video games with sales greater than 100,000 copies. It has 6 categorical data and 5 numerical data. The games are arranged in ascending order of the most sales with the number in the millions.

## 2. SCATTER PLOT

In the dashboard created the first option is the plot of global sales of the game vs the release year. As the sample was too numerous in the count, we applied two filters that selected the games based on their genre and the release platform. This limited the number of data that were =needed to be shown on the diagram itself.
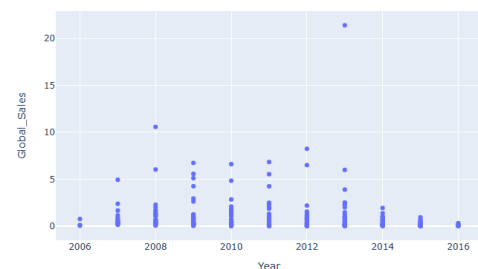


*Fig-2.1: Scatter plot showing the Action genre on PS3*

The plot in fig-1 shows us the period that action games were popular in the gaming community. We can see that action games were popular from the initial period of PS3 and were relevant till the end time of the PS3. The plot ended in 2016 because the PS3 was then discontinued.
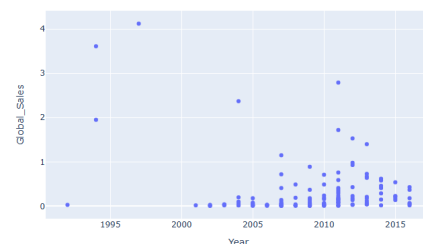


*Fig-2.2: Scatter plot showing the Shooter genre on PC*

The plot above shows that shooter games on PC have been present since the very early period of gaming. They were low in numbers but were popular. After 2000 PCs started to become stronger and the number of people playing games also increased, increasing the number of sales.

## 3. BAR GRAPH

Similar to the filter that we have applied in the scatter plot, we have applied a filter in the creation of the bar graph too as the data samples were too numerous and the space to represent them was limited.
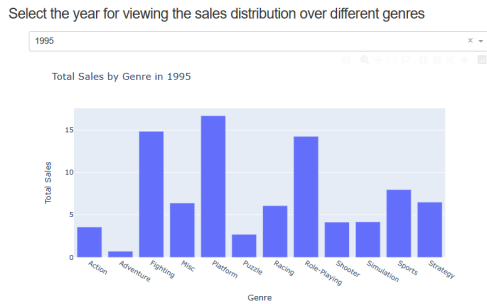


*Fig-3.1: Bar graph showing the total sales vs genre in 1995*

This shows us what the gaming trend was in the year 1995. In that era, most sold games belonged to the platform, fighting, and role-playing genre with the Adventure and Puzzle genres getting relatively lower sales.
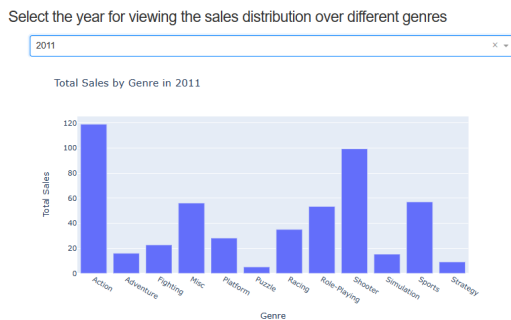


*Fig-3.2: Bar graph showing the total sales vs genre in 2011*

As the times changed, so did the people's preferences. In the year 2011, Action games and the Shooter games were the most popular with Puzzles, Strategy, and simulation getting lower sales.

We can see a huge increase in the sales of the action genre which can be credited to the release of games like GTA: Vice City and GTA: San Andreas.

## 4. PIE CHART

The pie chart in the dashboard gives us a comparative representation of the sales based on the genre of the games for the selected year.
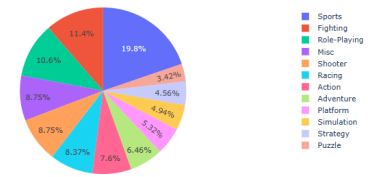


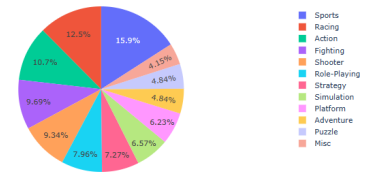*Fig- 4.1: Pie Chart of the Games released in 1996.*



*Fig- 4.1: Pie Chart of the Games released in 1996.*

Even when time had passed for just a year we can see the changes in the relative sales of the games belonging to a certain genre. Puzzles became more popular, sports became a little less popular, and so on.

## 5. CLASSIFIERS



*Fig 5.2 The classifier's accuracy.*

The classifier in our dashboard allows us to select a category and one or more classes for comparison. In this case, "Genre" has been selected as the category, and "Role-Playing", "Racing", and "Puzzle" have been chosen as classes. The results display the accuracy of two classifiers: the Decision Tree Classifier and the Random Forest Classifier.

## 6.  METHODOLOGY

The charts were created using Plotly which updated the DCC components of the dash. At first, the CSV file was loaded into the Python file. No transformations were needed on the file itself. Then the charts created using Plotly were displayed on a dashboard using HTML and deployed online on videogamessales.onrender.com. We can set the filters and get the relevant data on the fields we want.

## 7.  CONCLUSIONS

The Datases provided a timeline of the game genre that was the most popular. We can see the game type that was most popular in a specific period and how the trends changed over time. We can see the clusters of games that represent the trends.

## 8.  REFERENCES

[1] Dash Documentation & User Guide.
https://dash.plotly.com/

[2] Plotly Python Graphing Library
https://plotly.com/python/

[3] Video Game Sales
https://www.kaggle.com/datasets/gregorut/videogamesales/data