Video Game Sales Data Analysis and Visualization Project

The Video Game Sales dataset was obtained via Kaggle, and it provides insights into video game sales from 1980 through 2020. The dataset contains information about games that sold more than 100,00 copies. There are more than 16,500 different videogames listed in the data. The data contains 16,598 records and 11 different columns. Columns contain information such as the games Rank, Name, Platform, Release Year, Genre, Publisher, North America Sales, Europe Sales, Japan Sales, Sales in other regions, and Global Sales. The data was downloaded is a CSV format, but was converted into a regular workbook to be uploaded into Microsoft SQL Server Management Studio. My goal for this project was to generate insights using SQL and visualize those results using Tableau Public.

Before beginning my analysis, I first ran this query to ensure that no issues would occur when working with the data.

Text

Description automatically generated

The data loads correctly, so now I want to check if there are any missing values present.

Text

Description automatically generated

The results return about 271 NULL values in the Year column. There are also 56 missing values in the publisher column. I run the following queries to remove the rows with the missing data.

Text

Description automatically generated

Now that the data no longer contains missing values, we can begin the analysis.

1. **How many games are in each Genre?**

First, we run this query.

Text

Description automatically generated

Here, I want to see the Genre column and count of all the genres renamed as Amount. I group them by the Genre and order them by the amount descending so the highest amounts appear on top.

Table

Description automatically generated with medium confidence

* Action games had the most releases at 3,251 between 1980 and 2020.
* Puzzle Games had the least releases at 570 between 1980 and 2020.

Chart, bar chart

Description automatically generated

1. **Which 5 Years released the most games?**

Text

Description automatically generated

For this question, I will be working with the Year column and Count of all the Video Game names renamed as amount. They will be grouped by the year and ordered by the amount to have the highest appear on top.

Table

Description automatically generated

* 2009 released the most games at 1,431.
* 2011 released the least games 1,136.

Chart, bar chart, histogram

Description automatically generated

1. **What were the global sales per year?**

**Text

Description automatically generated**

For this query, I am working with the Year column and Global Sales summed up, rounded to 2 decimal places, and renamed as Global Sales. They were grouped by the Year and ordered by the Global Sales.

Table

Description automatically generated

Table

Description automatically generated

* 2008 had the highest global sales at $678.9 million.
* 2017 had the lowest global sales at $0.05 million.

Chart, line chart

Description automatically generated

1. **What were the total global sales for each genre?**

Text

Description automatically generated

Here, I am looking at the Genre column and Global Sales summed up and rounded to 2 decimal places. The new column was renamed as Global Sales, then it was grouped by the genre and ordered by global sales in a descending order.

Table

Description automatically generated

* Action has the highest global sales at $1,722 Million.
* Strategy games have the lowest global sales at $173 Million.

Chart, bar chart

Description automatically generated

1. **What was the most sold and produced Genre in what year?**

A screenshot of a computer

Description automatically generated with medium confidence

In this query, I am looking at the Year, Genre, and a count of the total number of genres renamed as Amount. I also run a Sub query to get the summed up global sales where the Year is 2009 and the Genre is Action. Then, it is grouped by the Year and Genre and ordered by the Amount descending.



We can see that Action genre games was the most produced and sold in 2009.

1. **Which Game Platform performed the best in Global Sales**

**Text

Description automatically generated**

Here, I want to see the Platform and Global Sales column summed up and rounded to 2 decimal places. They were grouped by the Platform and ordered by Sales descending.

Table

Description automatically generatedTable

Description automatically generated

* We can see that the PlayStation 2 performed the best with over 1 million dollars in sales.
* The PCFX performed the worst having the least number of sales overall.

Chart

Description automatically generated

1. **What are the 10 best performing Video Games?**

Text

Description automatically generated

Here, I choose to view only 10 results using the Name, Year, Genre, and Global Sales Columns. The results are then ordered by the Global Sales descending.

Graphical user interface, application

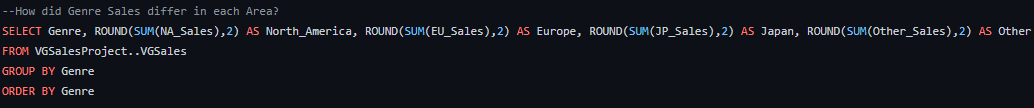
Description automatically generated

* Wii Sports is the best performing Video game of all time having over $82.71 million in global sales.
* The tenth best performing video game would be Duck Hunt have over $28.31 million in Global Sales.

Chart

Description automatically generated

1. **How did Genre Sales differ in each area?**



I wanted to see how genre sales were affected based on the region it was sold in. Here, I selected the Genre column, and each Regional column was summed up and rounded to 2 decimal places. Then they were grouped and ordered by the Genre.



* Action, Adventure, Fighting, Misc, Platform, Puzzle, Racing, Shooter, Simulation, Sports, and Strategy games performed the best in North America.
* Role Playing Games performed the best in Japan.

Table

Description automatically generated

1. **What were the Top 10 Game Publishers?**

Text

Description automatically generated

Here, I’m looking at the game publisher based on the amount of games they released over the years. I select the Publisher column, and a Count of how many times each name appears renamed in Published Games. They were grouped by the Publisher and order by the count descending.

Graphical user interface, text

Description automatically generated with medium confidence

* Electronic Arts published the most games at 1,339.
* Take-Two Interactive published 412 games making it the 10th publisher with the most games.

**Chart, bar chart

Description automatically generated with medium confidence**

1. **Game Publisher Global Sales**

Text

Description automatically generated

Now, I want to see how much in Global Sales each Publisher was able to make. I select the Publisher column and sum Global Sales rounding it to 2 decimal places. Then it is grouped by the Publisher and order by Sales descending.

Graphical user interface, application

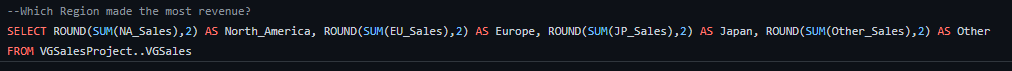
Description automatically generated

* Despite not having the most amount of published games, Nintendo is number one is it comes to global sales. Electronic Arts comes behind in a close second.

Chart, bubble chart

Description automatically generated

1. **Which Region made the most revenue?**



I want to see which region made the most revenue. I select the Region columns summing them up and rounding to 2 decimal places.

Graphical user interface, text, application

Description automatically generated

* North America made the most revenue overall.
* The “Other” Region made the least amount of revenue.

Chart, pie chart

Description automatically generated