**INTERNSHIP PROJECT REPORT**

**ON**

**“VIDEO GAMES ANALYSIS”**

**SUBMITTED BY:**

**PRAYAS SAMAL**

**EXECUTIVE SUMMARY**

The "Video Game Sales Analysis" project report examines the variables influencing video games' financial performance. To find patterns, trends, and statistically significant relationships between different game qualities and their commercial performance, the analysis employs a quantitative research methodology to look at a sizable dataset of video game sales. The paper explores the market dynamics for video games and offers insights that industry stakeholders may find useful.

The study makes use of Kaggle data, which includes a range of video games from various platforms and time periods. Total sales, platform, genre, publisher, developer, year of release, and reviewer ratings are among the important characteristics that are examined. Descriptive statistics, grouping and aggregation, exploratory data analysis (EDA), statistical inference (including t-tests, ANOVA, and chi-squared tests), and correlation analysis are all part of the analytical approach. Python is used for the data analysis, along with libraries like scipy.stats, matplotlib, seaborn, and pandas.

Several significant insights are revealed by the analysis's findings:

• 64,016 video game titles with 14 features—including qualitative and numerical data—make up the dataset.

• The distribution of total sales is skewed, and a small number of best-selling titles dominate the market.

• Japan and other regions have smaller market shares than North America and the PAL region, which contribute the most to overall sales.

• Platforms (such as the PS2 and X360) and major publishers (like Activision and Electronic Arts) have shown impressive overall sales performance.

• Among the most popular game genres are action and adventure games.

• There is a weak positive association between critic scores and overall sales, and critic reviews are often positive.

Missing data, possible data discrepancies, the inability to prove causation, a limited collection of features, the scope of the study, and the potential for bias in statistical tests are just a few of the limitations noted in the report**.**

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| **SERIAL NO** | **PARTICULARS** | **PAGE NO** |
| 1. | EXECUTIVE SUMMARY | 2 |
| 2. | OBJECTIVES | 4 |
| 3. | DATA ANALYSIS | 5-17 |
| 4. | CONCLUSION | 18 |

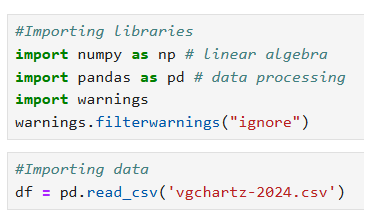
**OBJECTIVES**

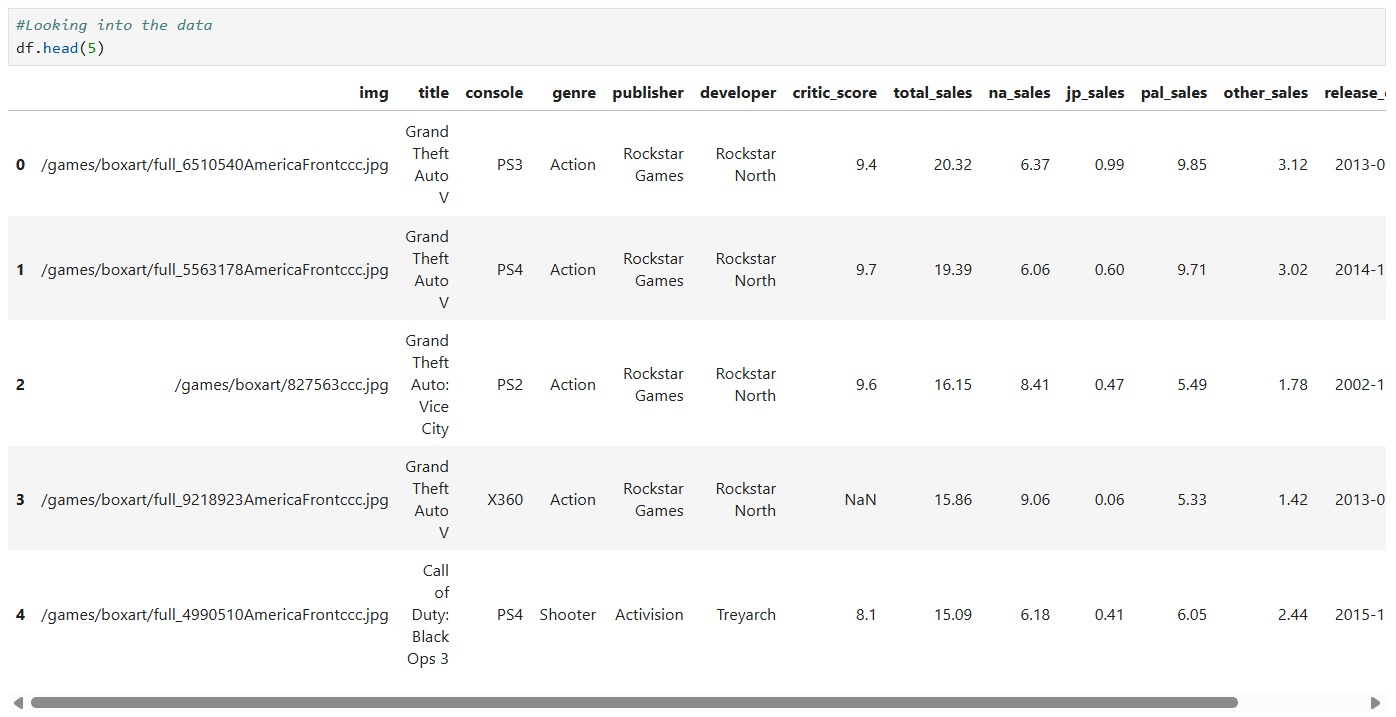
"Video Game Sales Analysis" is defined for the purposes of this project report as the methodical process of looking at and analysing quantitative data pertaining to the commercial performance of video games. This procedure entails:

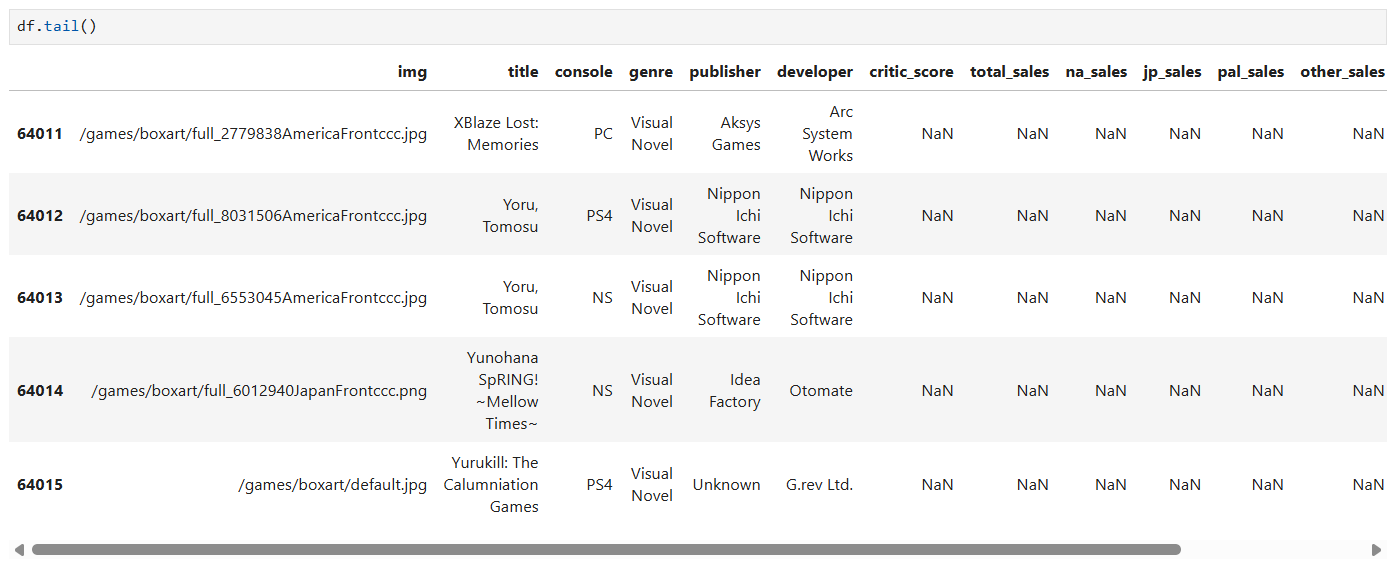
* Data Acquisition and Preparation: Compiling unprocessed data on video game sales numbers, possibly including variables like genre, publisher, developer, release date, critic ratings, platform (console, PC, mobile), revenue generated (both globally and in particular regions like North America, Japan, and Europe), and units sold. Data organization and cleansing for analysis are also part of this step.
* Descriptive statistics and exploratory data analysis (EDA): determining summary metrics for sales data and other pertinent numerical variables, such as mean, median, and standard deviation. Using data visualizations (such as bar charts, histograms, and scatter plots) to comprehend the distribution of important variables and spot patterns, trends, and possible links.
* Comparative analysis is the study of sales success in many categories. For example, it compares the overall sales of various game consoles, genres, publishers, or time periods (e.g., year of release). This could entail figuring out and displaying variations in average or total sales.
* Examining possible relationships or correlations between various factors and video game sales is known as relationship analysis. For instance, employing statistical tests such as the chi-squared test to investigate the relationship between console and genre or correlation coefficients and scatter plots to investigate the relationship between critic scores and overall sales
* Statistical Inference: By using statistical tests (such as ANOVA and t-tests) to determine the significance of observed associations or differences in the data, conclusions can be made about the larger population of video games outside of the particular dataset.

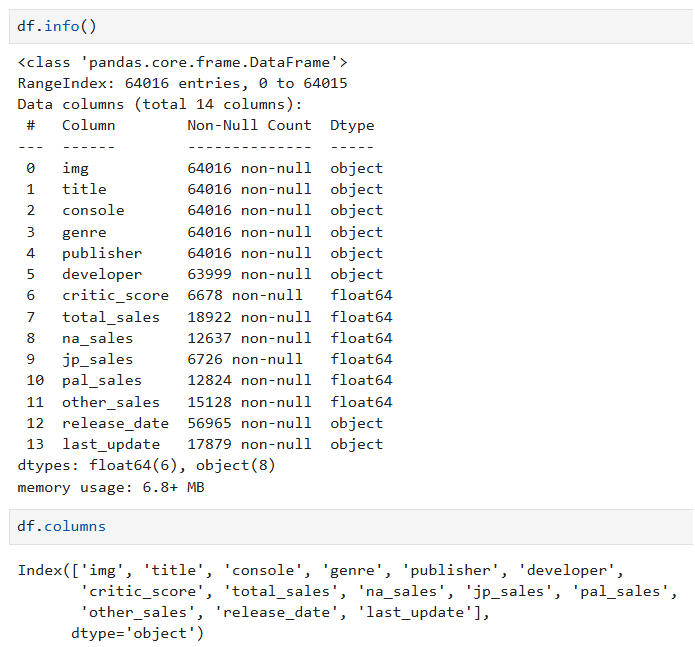
The ultimate goal of this "Video Game Sales Analysis," as defined within this report, is to extract meaningful insights from the sales data to understand the factors influencing commercial success in the video game industry, identify trends, and potentially inform future strategies or predictions. The specific techniques and variables analysed will be detailed within the methodology section of this report.

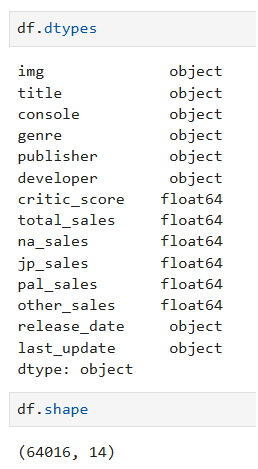
**DATA ANALYSIS**

****

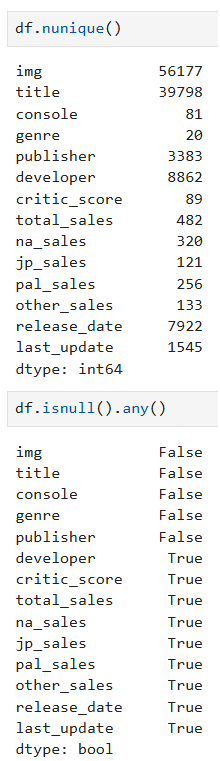
****

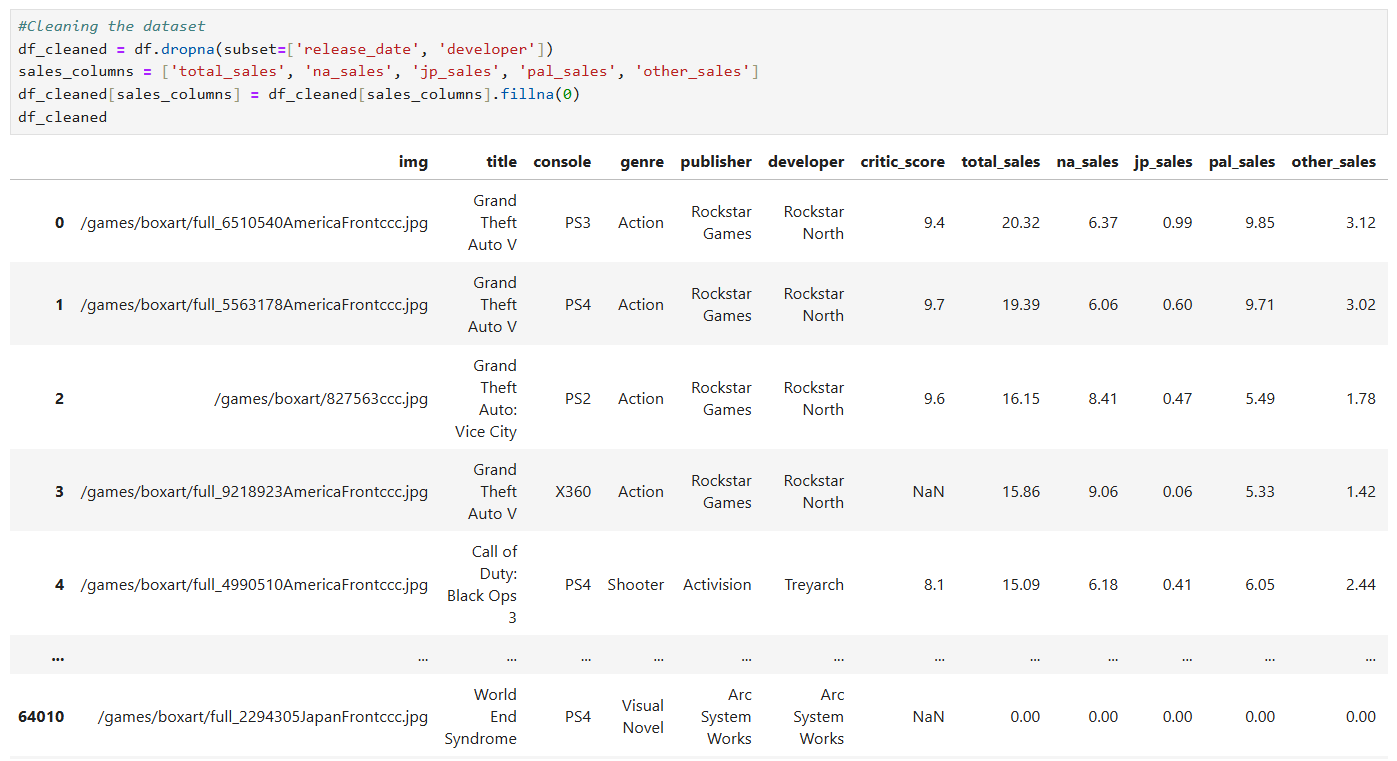
****

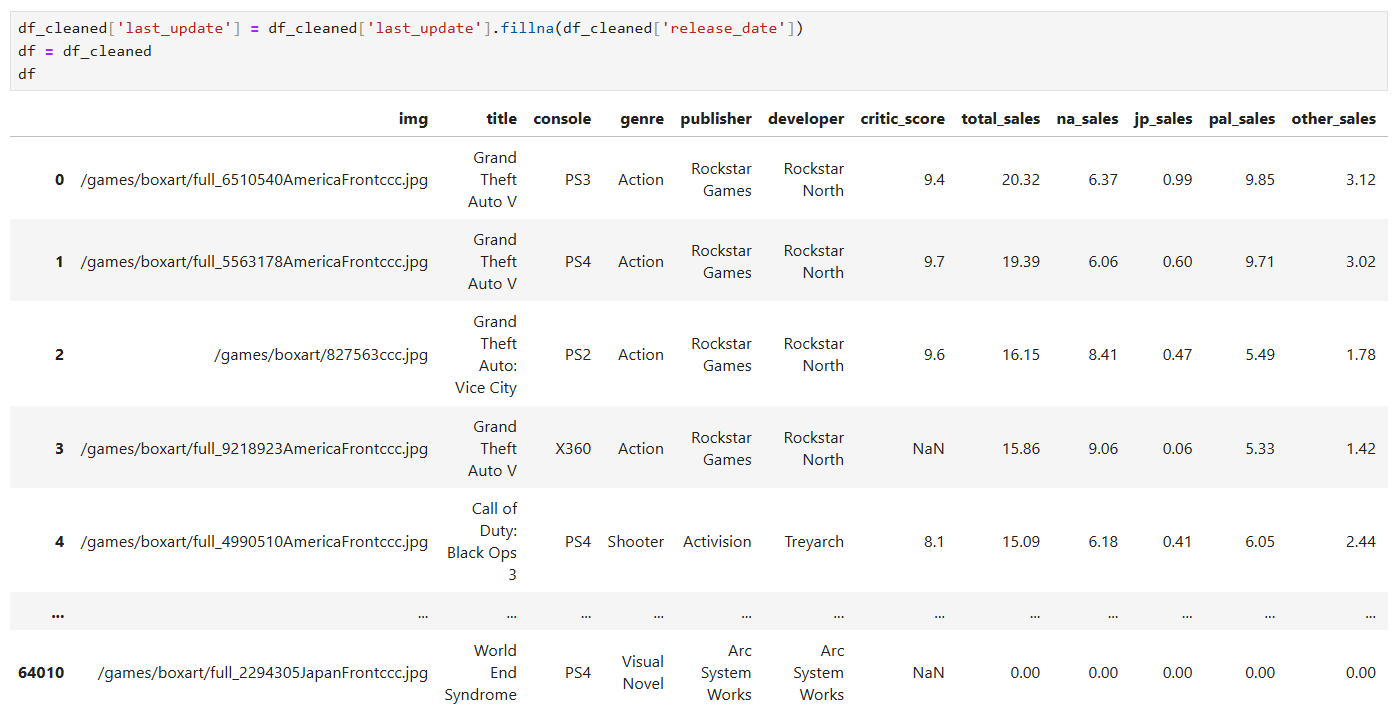
****

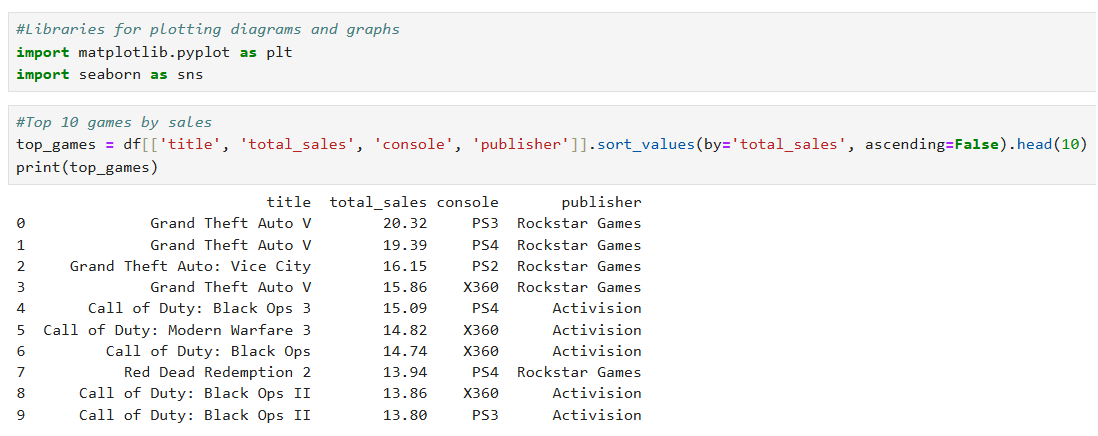
****

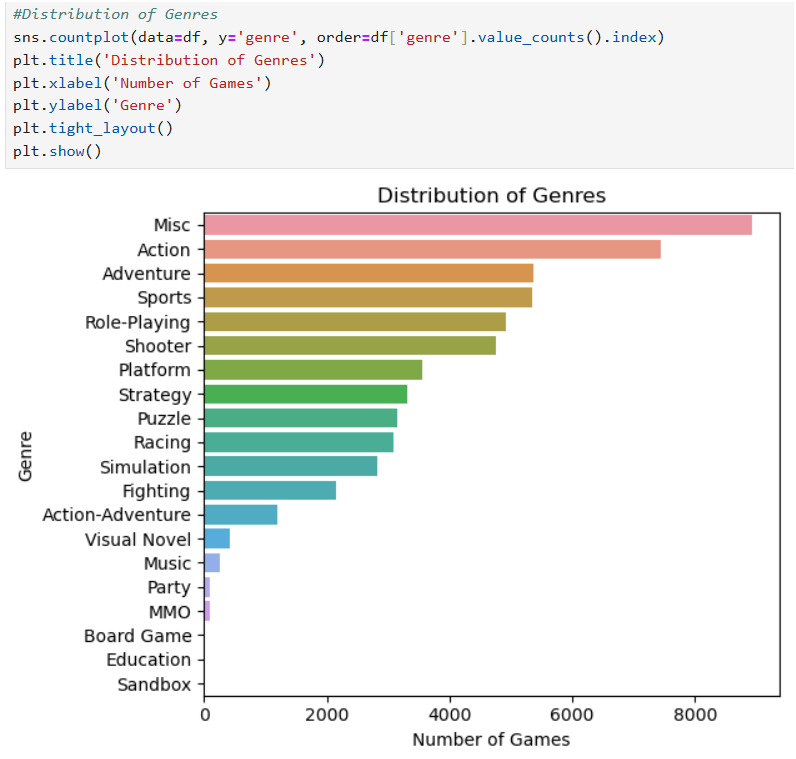
****

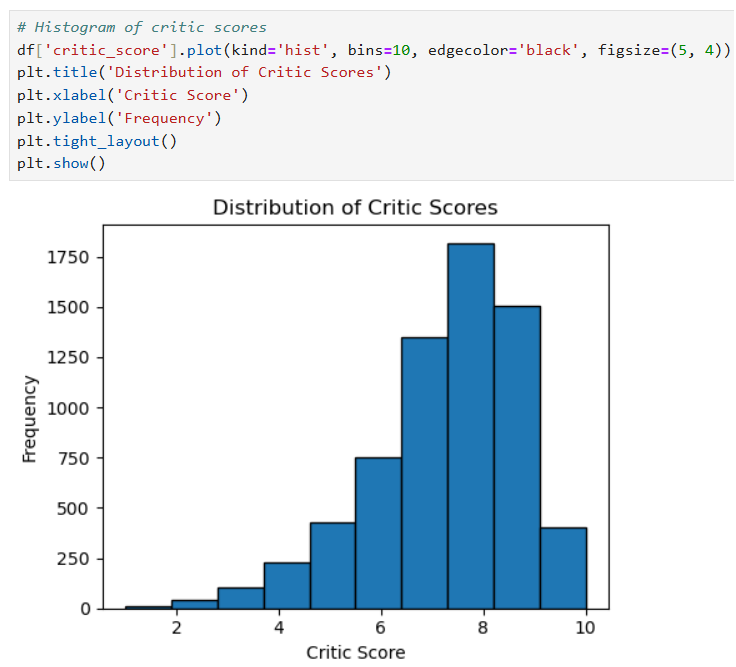
****

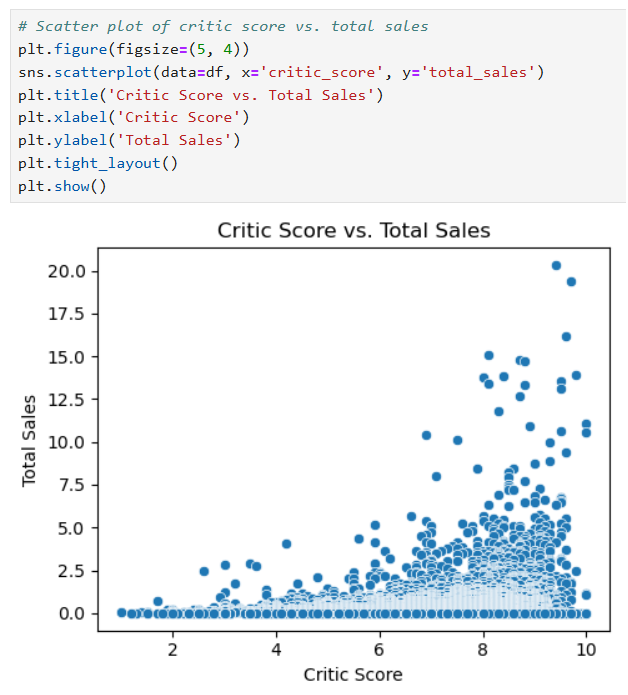
****

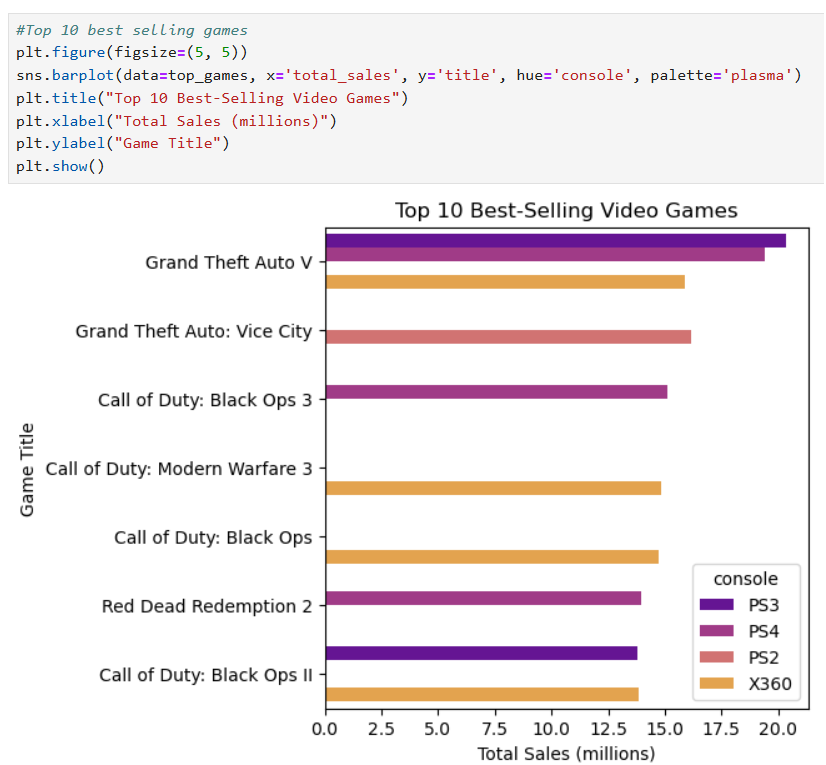
****

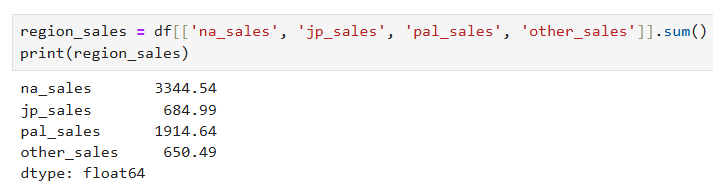
****

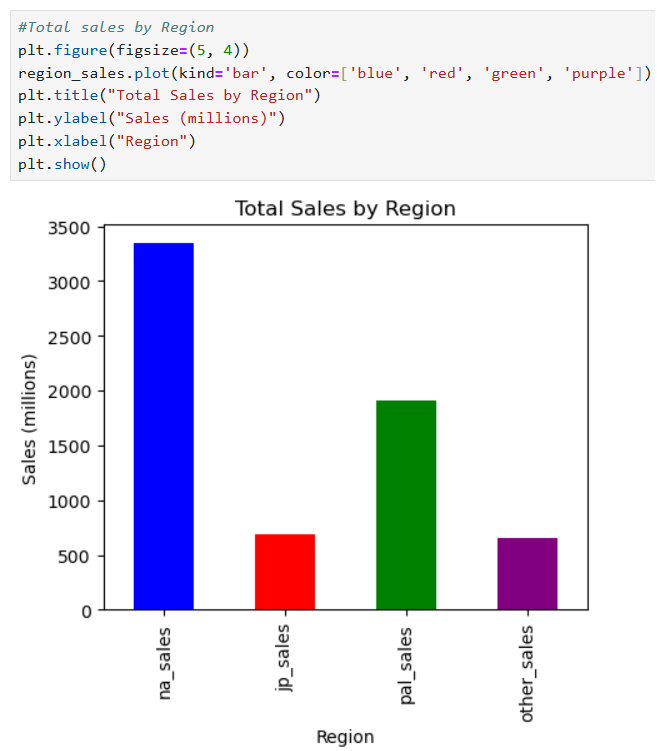
****

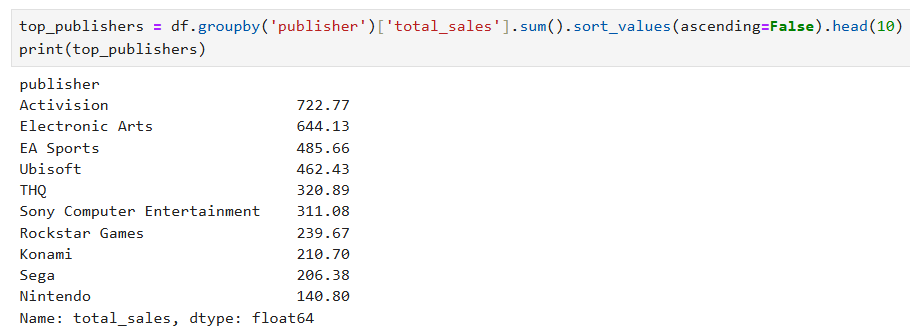
****

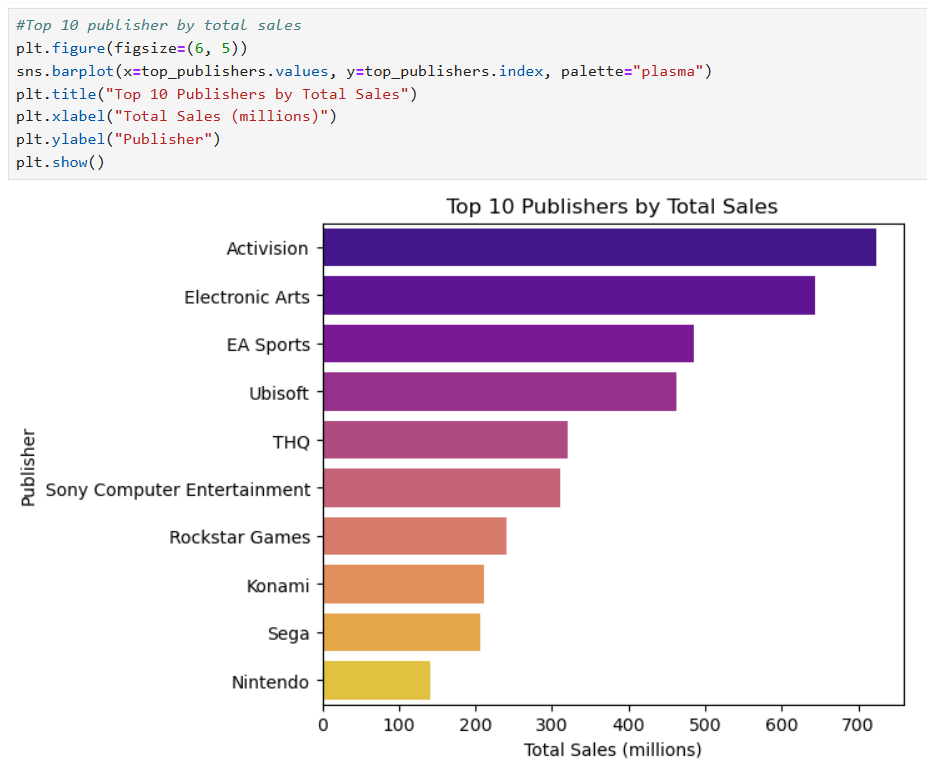
****

****

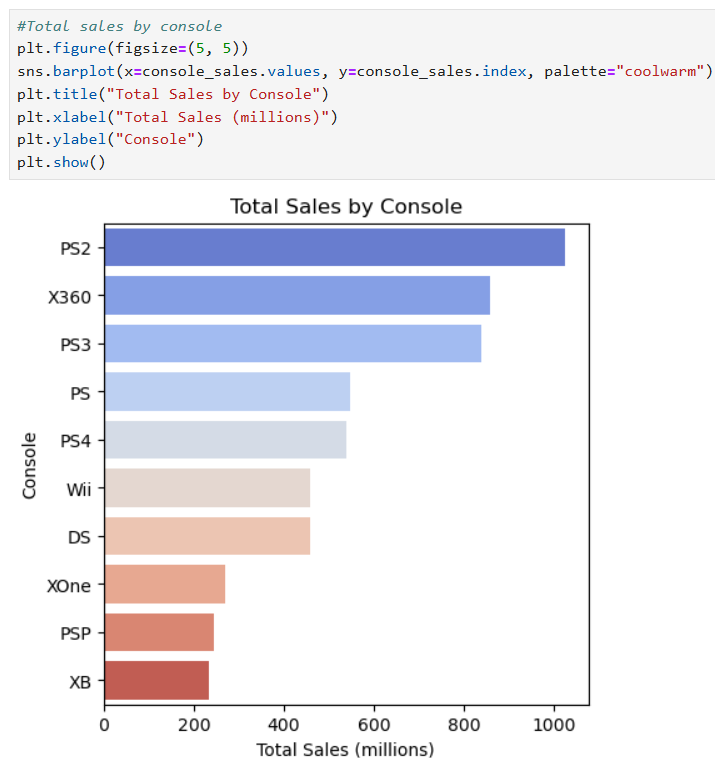
****

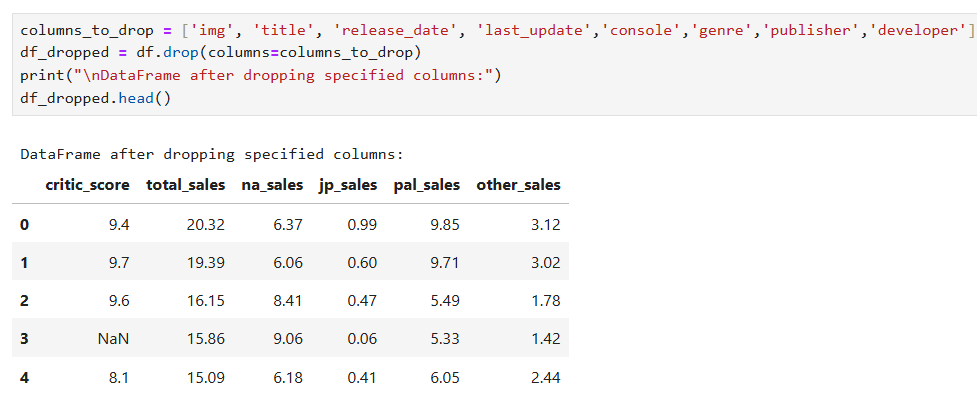
****

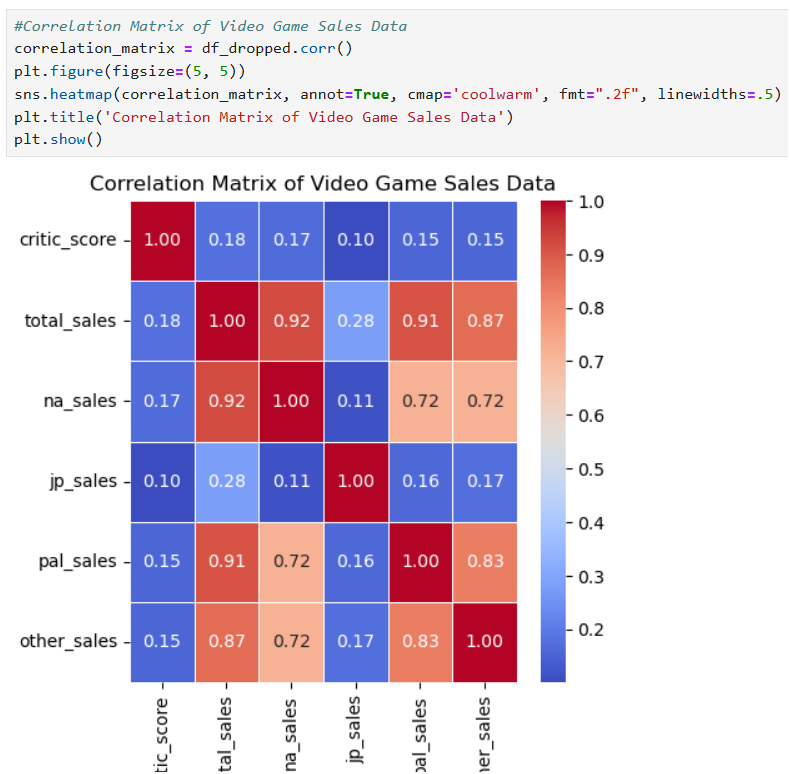
****

****

****

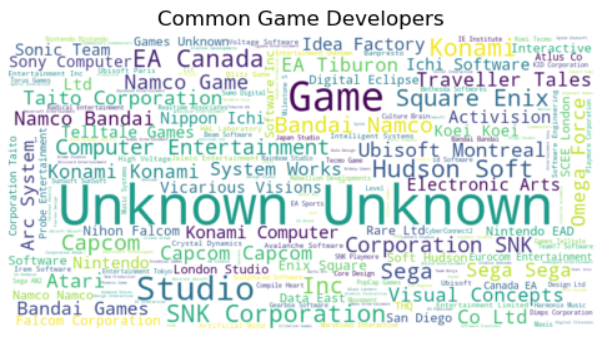
****

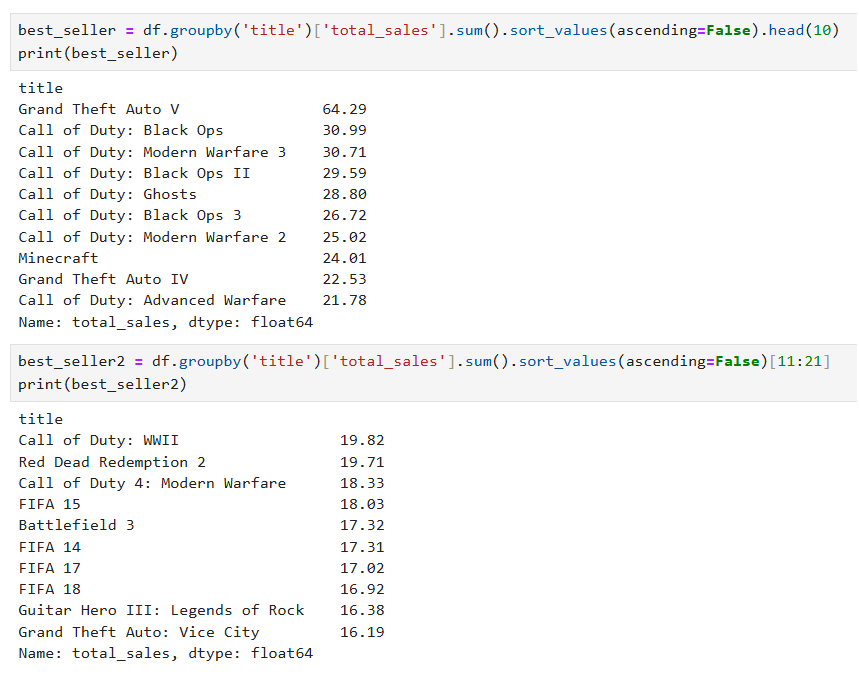
****

****

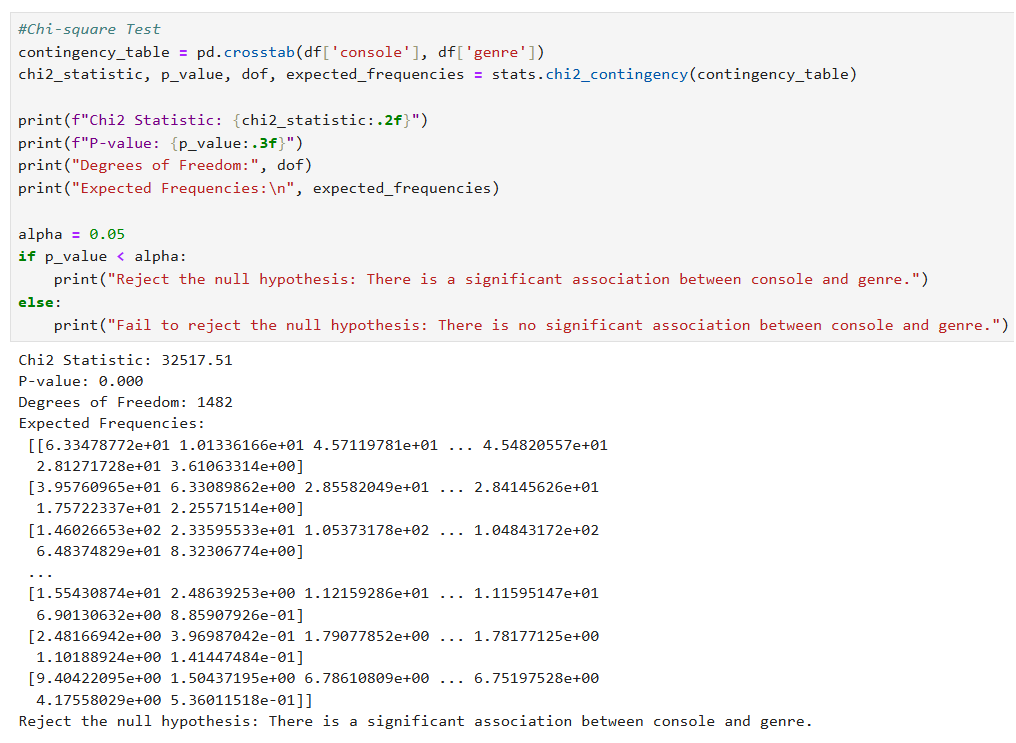
****

****

****

****

****

****

**CONCLUSION**

Several important insights into the elements affecting the industry's commercial performance have been uncovered by this examination of the video game sales dataset.

* Data Characteristics: The dataset includes 64,016 video game titles with 14 features. It includes both numerical (such as sales numbers and critic ratings) and categorical (such as genre, publisher, and console) data. Careful data cleaning and preparation are required because a sizable amount of the data, especially sales numbers and critic scores, have missing values.
* Sales Distribution: A few number of best-selling titles dominate the market, and the distribution of overall sales is severely skewed. One game that performs exceptionally well on a variety of platforms is Grand Theft Auto V.
* Regional Sales: While Japan and other regions have lesser market shares, North America and the PAL region make significant contributions to overall sales.
* Performance of publishers and systems: A number of publishers, including Activision and Electronic Arts, as well as systems such as the PS2 and X360, have shown impressive overall sales results.
* Analysis of Genre: Among the most popular gaming genres are action and adventure games.
* Critic ratings: The dataset generally shows a trend toward good evaluations, with critic ratings concentrated in the higher range. Total sales and critic scores have a weakly positive association.

Statistical Relationships:

* The genre of the game and the gaming system are statistically significantly correlated.
* The average reviewer evaluations for action and sports games differ statistically significantly.
* The difference between PS3 and PS4 system sales is statistically significant.