Video Game Sales Analysis — End-to-End Data Analytics Project Report

1. Project Title

End-to-End Video Game Sales Analysis using SQL, Python, and Power BI

2. Project Overview

The global gaming industry generates billions of dollars annually, with sales spread across multiple regions, platforms, and genres. Understanding patterns in video game sales is crucial for publishers, developers, and business stakeholders to make strategic decisions.

This project performs a complete end-to-end data analytics workflow, starting from raw data extraction to data cleaning, analysis, and interactive dashboard visualization.

The analysis was conducted using:

• SQL for data extraction and querying
• Python for data cleaning and exploratory analysis
• Power BI for visualization and dashboard creation

3. Project Objectives

The primary goals of this project were:

• To analyze global video game sales performance
• To identify top genres, platforms, and publishers
• To understand regional sales trends over time
• To evaluate the relationship between ratings and sales
• To build an interactive dashboard for business insights

4. Tools & Technologies Used

Database & Querying

• MySQL / SQL

Programming & Analysis

• Python
• Pandas
• NumPy
• Matplotlib
• Seaborn

Visualization

• Power BI Desktop
• DAX Calculations

5. Dataset Description

The dataset contains historical video game sales information.

Key Attributes:

- Game Title
- Platform
- Primary Genre
- Publisher
- Release Year
- North America Sales
- Europe Sales
- Japan Sales
- Other Region Sales
- Global Sales
- Game Ratings

6. Project Workflow

This project followed a complete analytics pipeline:

Data Collection → Data Cleaning → Data Analysis → Visualization → Insights

7. SQL Phase — Data Extraction & Preparation

In the first stage, SQL was used to explore and prepare the dataset.

Tasks Performed:

• Imported dataset into SQL database
• Retrieved important fields using SELECT queries
• Filtered data based on conditions
• Aggregated sales data using GROUP BY
• Calculated total regional and global sales

Sample Analysis Done:

- Total sales by genre
- Top platforms by global sales
- Publisher-wise revenue comparison

SQL helped structure the data efficiently before moving to deeper analysis.

8. Python Phase — Data Cleaning & Exploratory Analysis

Python was used for preprocessing and performing exploratory data analysis (EDA).

Steps Performed:

1. Loaded dataset using Pandas
2. Handled missing and null values
3. Removed duplicate records
4. Converted data types properly
5. Created new calculated fields
6. Generated summary statistics

Exploratory Data Analysis Performed:

• Distribution of global sales
• Sales comparison across regions
• Trend analysis over years
• Correlation between ratings and sales
• Identification of top-selling categories

Python visualizations helped uncover hidden patterns before building dashboards.

---

9. Power BI Phase — Dashboard Development

After cleaning and analysis, Power BI was used to build an interactive dashboard.

9.1 Key Performance Indicators (KPIs)

The dashboard includes:

• Total Global Sales
• Total North America Sales
• Total Europe Sales
• Total Japan Sales
• Average Global Sales
• Average Game Rating

These KPIs provide a quick snapshot of overall performance.

9.2 Dashboard Visualizations

1. Top Game Genres by Global Sales

Shows most profitable game categories.
Insight: Adventure and Shooter genres dominate.

---

2. Top Platforms by Global Sales

Displays best-performing gaming consoles.
Insight: Xbox 360, PS3, and Wii lead in sales.

---

3. Global Sales Trend Over Years

Shows sales growth over time.
Insight: Rapid growth occurred during the mid-2000s gaming boom.

---

4. Top Publishers by Global Sales

Identifies market-leading companies.
Insight: Nintendo leads global sales.

---

5. Regional Sales Trend by Year

Compares NA, EU, and JP performance.
Insight: North America consistently shows highest revenue.

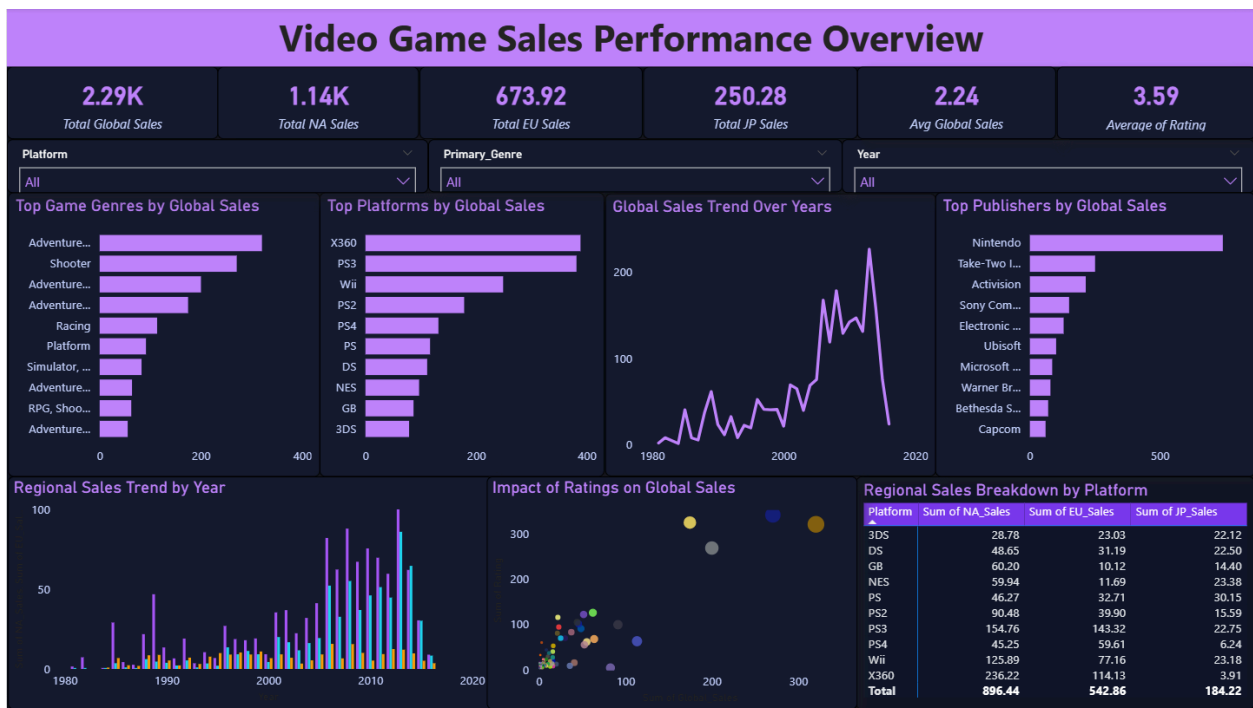### 6. Impact of Ratings on Global Sales

Scatter plot analyzing correlation.
Insight: Higher-rated games generally achieve higher sales.

7. Regional Sales Breakdown by Platform

Matrix visualization showing regional performance per platform.

10. Power BI Dashboard Preview

Insight: "The dashboard provides an interactive view of global video game sales performance across genres, platforms, regions, and publishers."

11. Key Insights from the Project

- Adventure and Shooter genres generate highest revenue
- North America is the largest gaming market
- Xbox 360 and PS3 are top-selling platforms
- Nintendo dominates publishing industry
- Higher ratings positively impact global sales

12. Business Impact of Analysis

This analysis helps stakeholders:

• Identify profitable game genres
• Understand regional demand
• Optimize marketing strategies
• Improve platform investment decisions
• Predict future gaming trends

---

13. Challenges Faced

• Handling missing sales values
• Cleaning inconsistent genre labels
• Managing large categorical data
• Designing a balanced dashboard layout

---

14. Conclusion

This end-to-end project demonstrates the complete data analytics workflow, starting from raw data querying in SQL to deep analysis in Python and interactive visualization in Power BI.

The project successfully transforms raw gaming data into actionable insights that can support business decision-making in the gaming industry.

---

15. Future Scope

• Build predictive sales forecasting model
• Integrate real-time gaming data
• Perform customer behavior analysis
• Add machine learning recommendation system

---

16. Author

Muskan Bhardwaj
Aspiring Data Analyst
Skilled in SQL | Python | Power BI

END OF REPORT