DATA ANALYSIS PORTFOLIO

PREPARED BY

Professional background

As a final year student of public health, I have always been interested in the application of data analysis in decision-making. My academic journey has been a fulfilling one, with a cumulative grade point average of 4.35 out of 5.00. This has given me a solid foundation in public health, which I have complemented with practical experience through an internship at the Asaba Specialist Hospital.

During my internship, I was privileged to work with the hospital's public health department, where I was actively involved in data analysis. This was a great opportunity for me to put my analytical skills to work as I helped to analyze research data using SPSS. My experience with SPSS sparked an interest in data analysis and its potential in making informed decisions.

Furthermore, I have taken several courses in data analysis at the entry-level, which have helped me gain valuable experience and skills. I am proficient in the use of various data analysis tools, including Excel, R, Power BI, and Tableau. My ability to use these tools effectively has given me the confidence to handle complex data sets and derive meaningful insights from them.

My fascination with the potential of data analysis in making informed decisions has continued to grow. I am convinced that data analysis is a critical component of public health, as it enables us to make informed decisions that can improve the health outcomes of populations. I believe that with my skills and experience, I am well-positioned to contribute to the field of public health through data analysis.

In conclusion, my academic background, practical experience, and proficiency in data analysis tools have equipped me with the necessary skills to make meaningful contributions to the field of public health. I am excited about the possibilities that data analysis offers in improving health outcomes, and I am committed to using my skills and knowledge to make a positive impact in this field.

Portfolio outline

Udemy Project

Description

- As the data analyst for Udemy, I was instructed by my manager to identify the performance level of the courses on Udemy, measure the revenue of the courses, and also identify the high-performance courses that suggest courses with larger number of students/ interests, so that opportunities for revenue increment can be identified.

Methodology

- I studied the course subscription trends for over four courses obtained from Udemy from a period of 2011-2017. The data was modified, duplicates were removed, and blank sheets expunged and inconsistent data modified using Google sheets. The clean data was analyzed using Google sheets and Tableau was used to create the necessary visuals. This analysis is aimed at identifying possible opportunities to increase revenue and track the performance of the following subjects; Web Development, Graphic Design, Business Finance and Musical instruments.

Results

Information obtained from the dataset includes;

- Web Development is the most popular subject, it has the highest number of subscribers, Seventeen (17) out the top 20 courses are from Web Development, and has also provided the highest revenue.
- There are opportunities in revenue increment in Web Development, seeing as it is responsible for most of the revenue generation.
- Reviews from other subscribers may help increase sales of Web development courses thereby further increasing the revenue

- Web Development also has the largest amount of free courses, some can be converted to paid courses and can be easily done by picking the free courses from the top 20 table, and this would also help to further increase revenue.

The Problem

- The business problem is to identify Performance level of the courses on Udemy, measure the revenue of the courses, and also indicate the high-performance courses as a means to identify opportunities for revenue increment. The analysis is expected within three (3) weeks.

- Data gotten from Web Development, Graphic Design, Business Finance, and Musical Instruments will provide information. The data is provided in descriptive, tabular and pictorial to enable its understanding.

- This analysis seeks to understand the total number of subscribers, subject with the highest revenue, and the subject with the highest positive reviews to understand the performance of various subjects.

DATA DESIGN

I studied the course subscription trends for over four courses obtained from Udemy from a period of 2011-2017. The data was modified, duplicates were removed, and blank sheets expunged and inconsistent data modified using Google sheets. The clean data was analyzed using Google sheets and Tableau was used to create the necessary visuals. This analysis is aimed at identifying possible opportunities to increase revenue and track the performance of the following subjects; Web Development, Graphic Design, Business Finance and Musical instruments.

The results show that Web Development is the most popular subject, it has the highest number of subscribers, Seventeen (17) out the top 20 courses are from Web Development, and has also provided the highest revenue.

It also shows that Web Development does has the highest number of ratings.

FINDINGS.

FINDING 1

TABLE 1; Total Subscriptions by Subject

As reported by the data provided on different course titles and subjects, Web Development occupies the largest percentage of subscribers (67.9%) which in total 7,981,935.

subject	SUM of num_subscribers
Business Finance	1868711
Graphic Design	1063148
Musical Instruments	846689
Web Development	7981935
Grand Total	**11760483**

FINDING 2

This analysis shows that Web Development has the highest average reviews

## Figure 1; Average review by Subject;
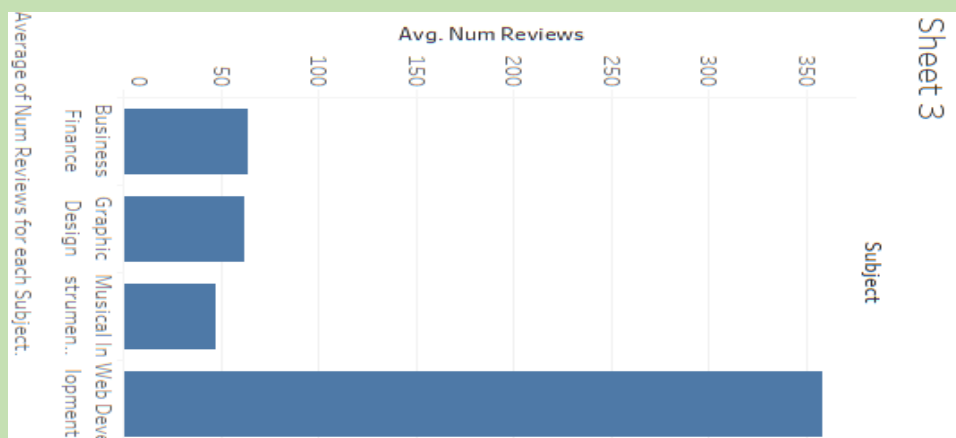
FINDINGS 3

This analysis indicates Web development as the subject with highest average price.

Table 2; Subject by Average price

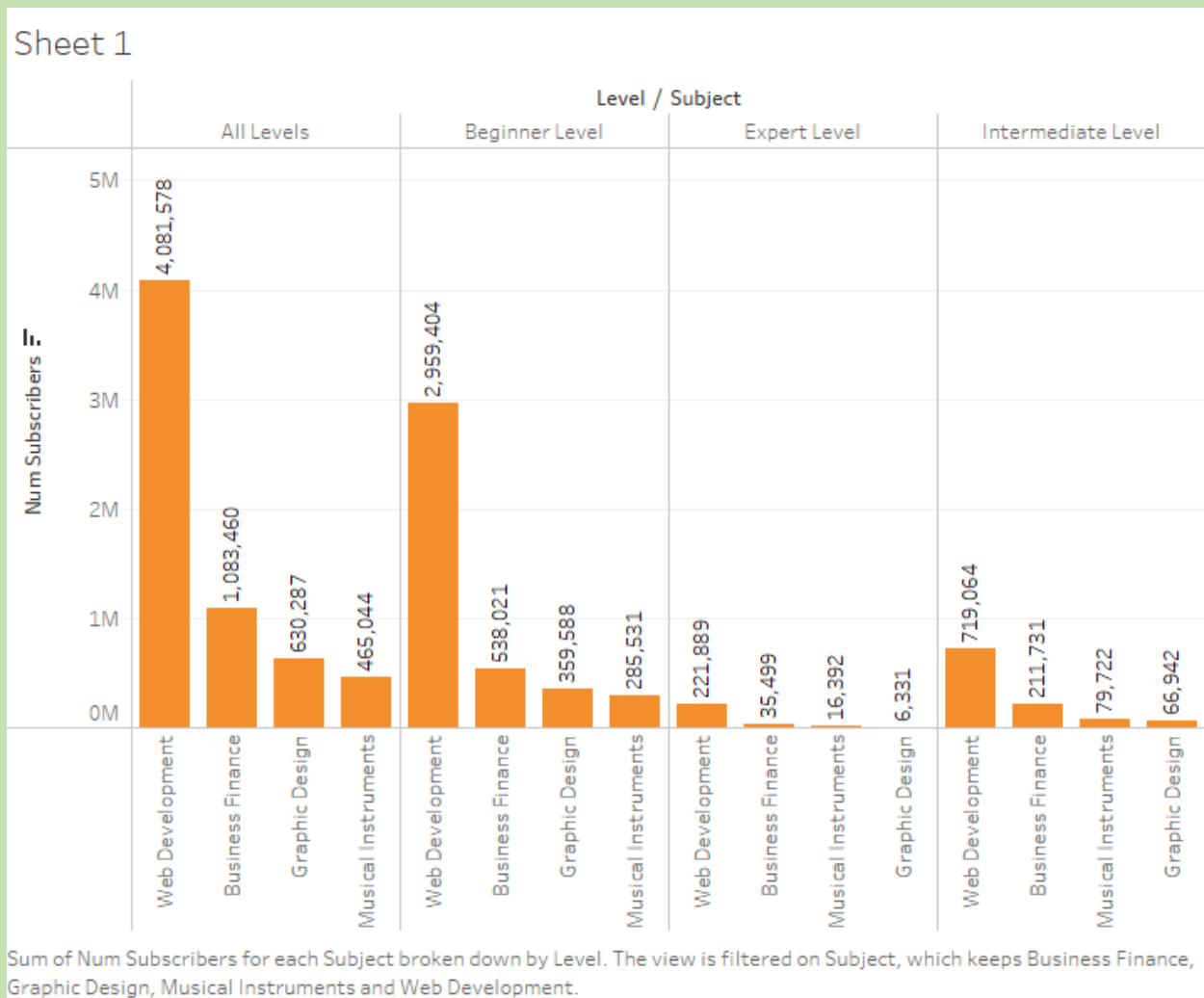| subject	AVERAGE of price
Business Finance	68.7
Graphic Design	57.9
Musical Instruments	49.6
Web Development	77.0
Grand Total	**66.1**

Figure 2; Average price by subject

Average of Price for each Subject. Color shows details about Payment Type.
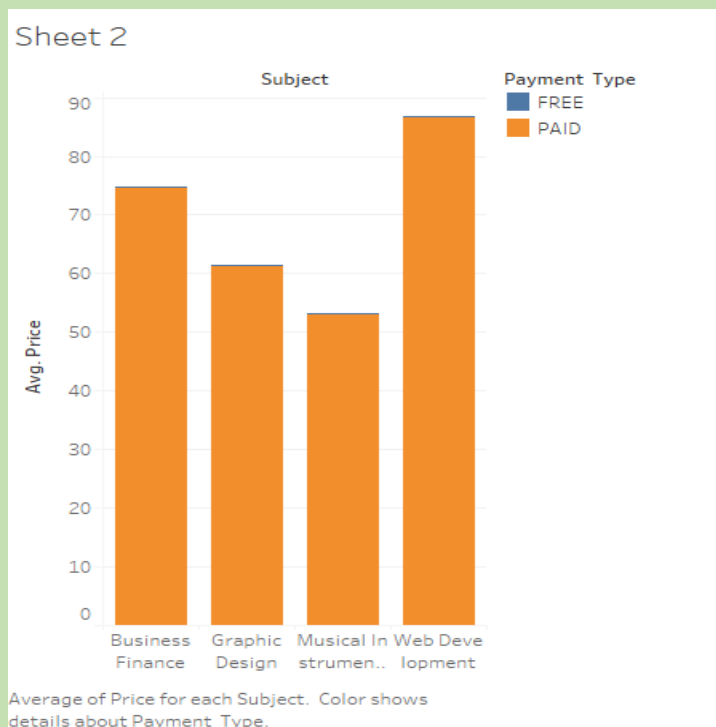
Findings 4

This shows the distribution of subscribers across all levels and all subject. It also shows that Web development has the highest amount of subscribers across each level.

Figure 3; Levels by Subject

Sum of Num Subscribers for each Subject broken down by Level. The view is filtered on Subject, which keeps Business Finance, Graphic Design, Musical Instruments and Web Development.

Findings 5

This shows that Web development has the highest number of both free and paid classes.

Table 3; Subject by Payment type

SUM of num_subscribers	payment _type		
subject	FREE	PAID	Grand Total
Business Finance	603228	1265483	1868711
Graphic Design	284821	778327	1063148
Musical Instruments	304735	541954	846689
Web Development	2393767	5588168	7981935
Grand Total	**3586551**	**8173932**	**11760483**

Figure 5; Payment type by Subject

This shows that Web development has the highest number of both free and paid classes at 20.35% and 47.52% respectively.
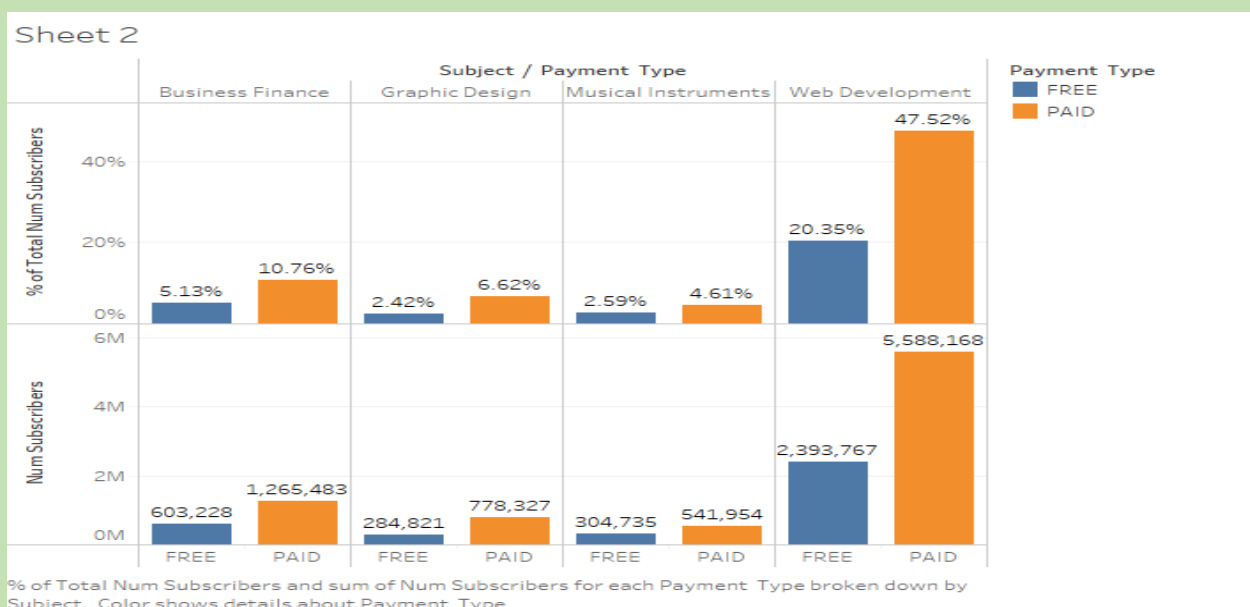
MY ANALYSIS

My primary perception of the data is that is that Web Development would record the highest number of subscribers amongst the other courses because there is a high rise in the skill market for the course.

My conclusion reached on the basis of evidence that this dataset has provided includes;

- A possible reduction in revenue generation due to the high amount of free courses offered on the Udemy website and the low percentage of reviews, which is not an encouraging factor in the boosting of sales.
- Although Web Development is the highest revenue generating course on the Udemy website, all other courses suffer from the same issues, high amount of free courses and low reviews.

The whys analysis would help determine the root course of the problem and provide possible insights on how to solve the problem.

Analysis

Whys analysis

Problem: Low revenue generation

Why?

Possible reason: low sales of courses

Why?

Possible reason: Large percentage of free courses

Why?

Possible reason; low reviews, awareness and marketing strategies.

CONCLUSION

- The data shows that Web Development is the most popular subject and is responsible for the highest revenue generation between the years.

- There are various opportunities to increase revenue generation on all the courses. If the amount of free courses are reduced across all four subject areas, especially Web Development because it is the most popular subject and has the highest percentage of free courses, this would help increase revenue.

- Reviews from subscribers and new marketing and advertising strategies can help boost sales of courses, especially Web Development.

- There is also a need to determine the total profit in relation to the number of people enrolled for different courses.

Capstone project

Description

❖ This project is based on video game sales, their ranks, and publishers, sales in different regions and globally. I did a review of the data, where I identified some problems and some possible impacts on Production Company and stakeholders. The problems identified led to the formation of my hypothesis, which includes; drastic reduction of sales, reduction in revenue and reduction in profit.

Methodology

❖ The data was modified, duplicates were removed, and blank sheets expunged and inconsistent data modified using Google sheets. The clean data was analyzed using Google sheets and Tableau was used to create the necessary visuals.

❖ This study aims to track performance of video game sales, identify possible revenue increase, and identify any other hidden problems.

Results

Information provided from this analysis includes

❖ After global sales peaked in 2008, it was identified that global sales reduced every year after 2008 and began to drastically reduce from 2016 down to 2020.

❖ The action genre generated the highest percentage of global sales.

❖ The publishing of video games through Nintendo drives up the sales of video games as it occupies 89.15% of the top 20 video game sales.

❖ Video games with Ps2 platform had the highest global sales.

Data design

I studied the sales of video games from 1980-2020. The data was modified, duplicates were removed, and blank sheets expunged and inconsistent data modified using Google sheets. The clean data was analyzed using Google sheets and Tableau was used to create the necessary visuals.

Tableau was used for visualizations due to its capacity to absorb large data sets and give more in-depth information on the visuals created such as percentages and numbers required to understand the visuals clearly.

This analysis is aimed at identifying possible opportunities to increase revenue and track the sales performance of the video games, identify genres, regions, publishers and regions with highest sales globally.
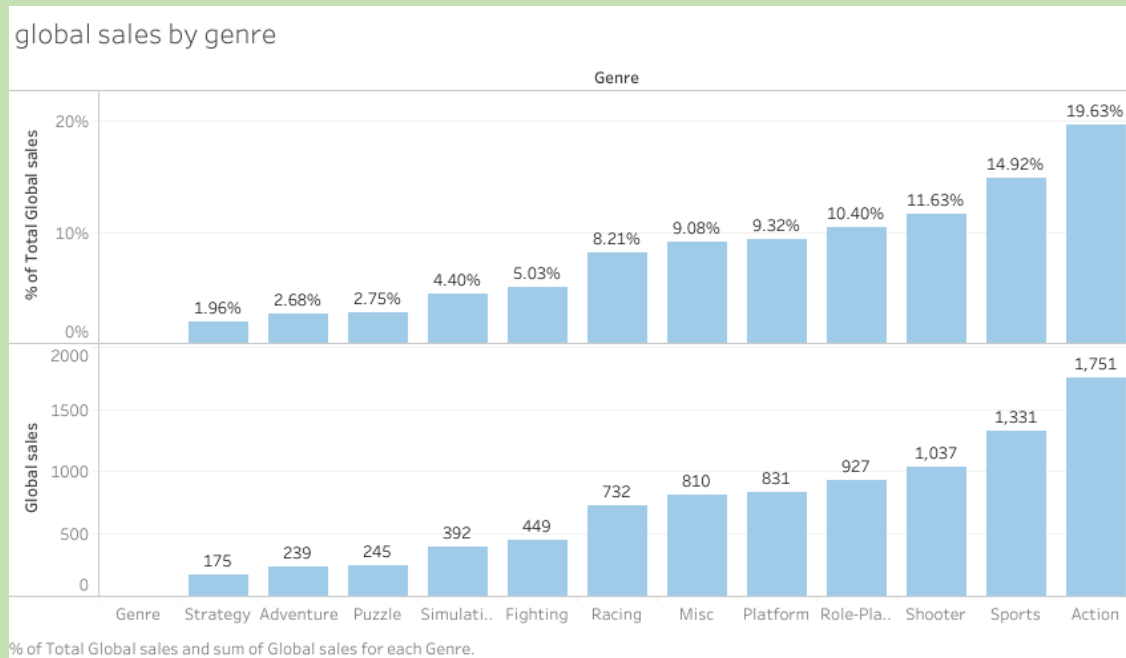
FINDINGS

FINDINGS 1

TABLE 1; Total sales by genre

This shows that action recorded the highest sales across North America, Europe, Japan, other sales and global sales respectively.

Genre	SUM of NA_sales	SUM of EU_sales	SUM of JP_sales	SUM of other_sales	SUM of Global_sales
Action	877.83	525.00	159.95	187.38	1751.18
Adventure	105.8	64.13	52.07	16.81	239.04
Fighting	223.59	101.32	87.35	36.68	448.91
Misc	410.24	215.98	107.76	75.32	809.96
Platform	447.05	201.63	130.77	51.59	831.37
Puzzle	123.78	50.78	57.31	12.55	244.95
Racing	359.42	238.39	56.69	77.27	732.04
Role-Playing	327.28	188.06	352.31	59.61	927.37
Shooter	582.6	313.27	38.28	102.69	1037.37
Simulation	183.31	113.38	63.7	31.52	392.2
Sports	683.35	376.85	135.37	134.97	1330.93
Strategy	68.7	45.34	49.46	11.36	175.12
Grand Total	**$4392.95**	**$2434.13**	**$1291.02**	**$797.75**	**$8920.44**

Figure 1; Global sales by Genre

global sales by genre

Genre

20%	19.63%

% of Total Global sales

1.96% 2.68% 2.75% 4.40% 5.03% 8.21% 9.08% 9.32% 10.40% 11.63% 14.92%

Global sales

175 239 245 392 449 732 810 831 927 1,037 1,331 1,751

Genre Strategy Adventure Puzzle Simulati.. Fighting Racing Misc Platform Role-Pla.. Shooter Sports Action

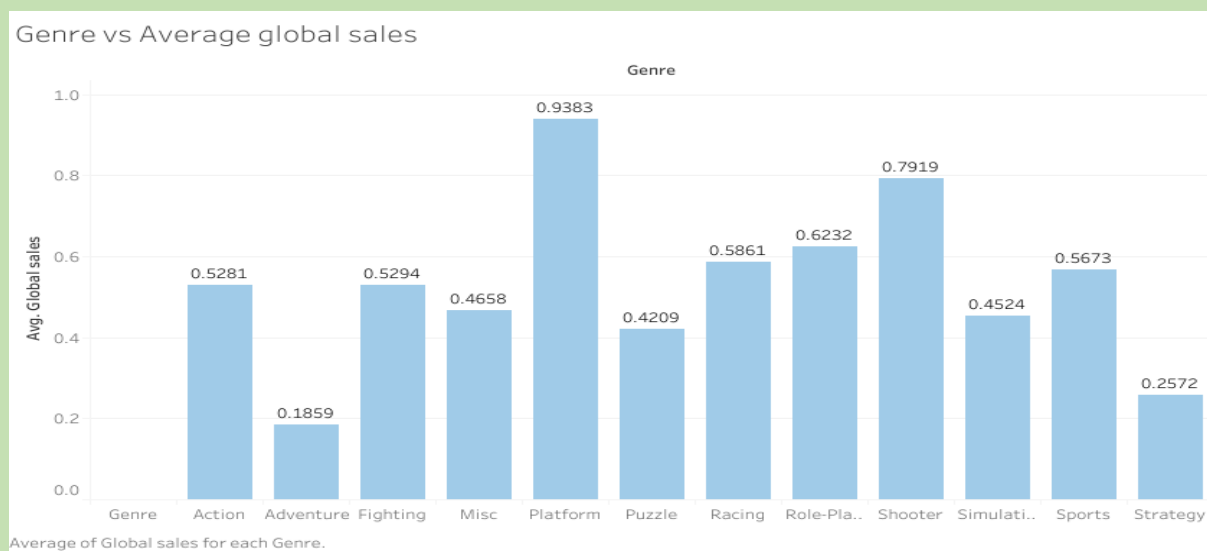% of Total Global sales and sum of Global sales for each Genre.

Findings 2;

Figure 2; Average global sales by Genre

This shows that platform recorded the highest average global sales at 0.9383 million.

Genre vs Average global sales

Genre

Avg. Global sales

0.5281 0.1859 0.5294 0.4658 0.9383 0.4209 0.5861 0.6232 0.7919 0.4524 0.5673 0.2572

Genre Action Adventure Fighting Misc Platform Puzzle Racing Role-Pla.. Shooter Simulati.. Sports Strategy

Average of Global sales for each Genre.
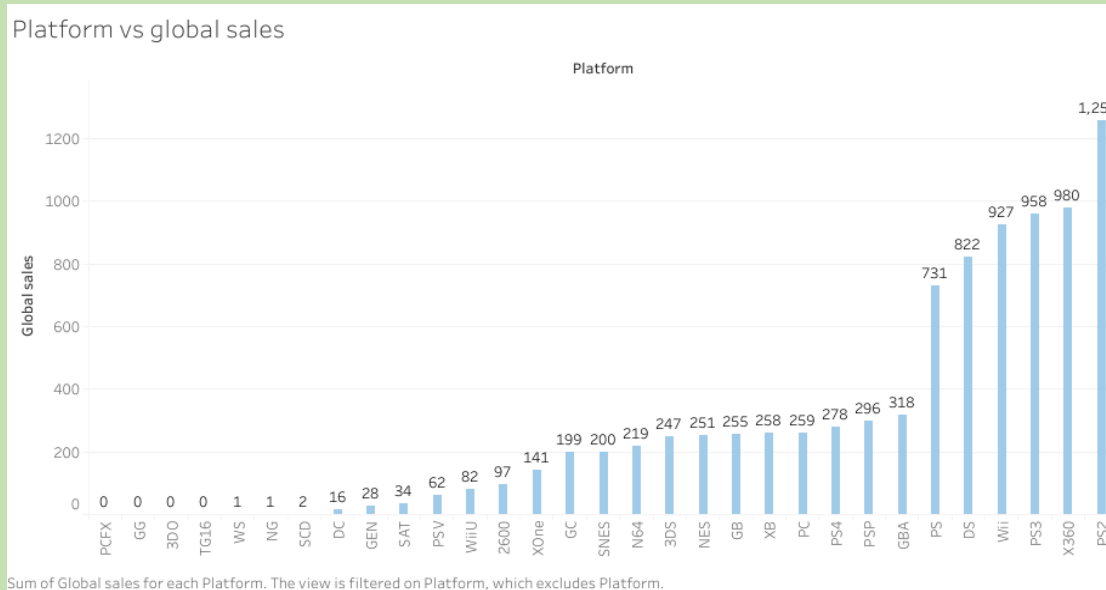
Findings 3

Figure3; Yearly Global sales

In 2008 the highest global sales was recorded at 678.9 million and thereafter there was constant declining in global sales.
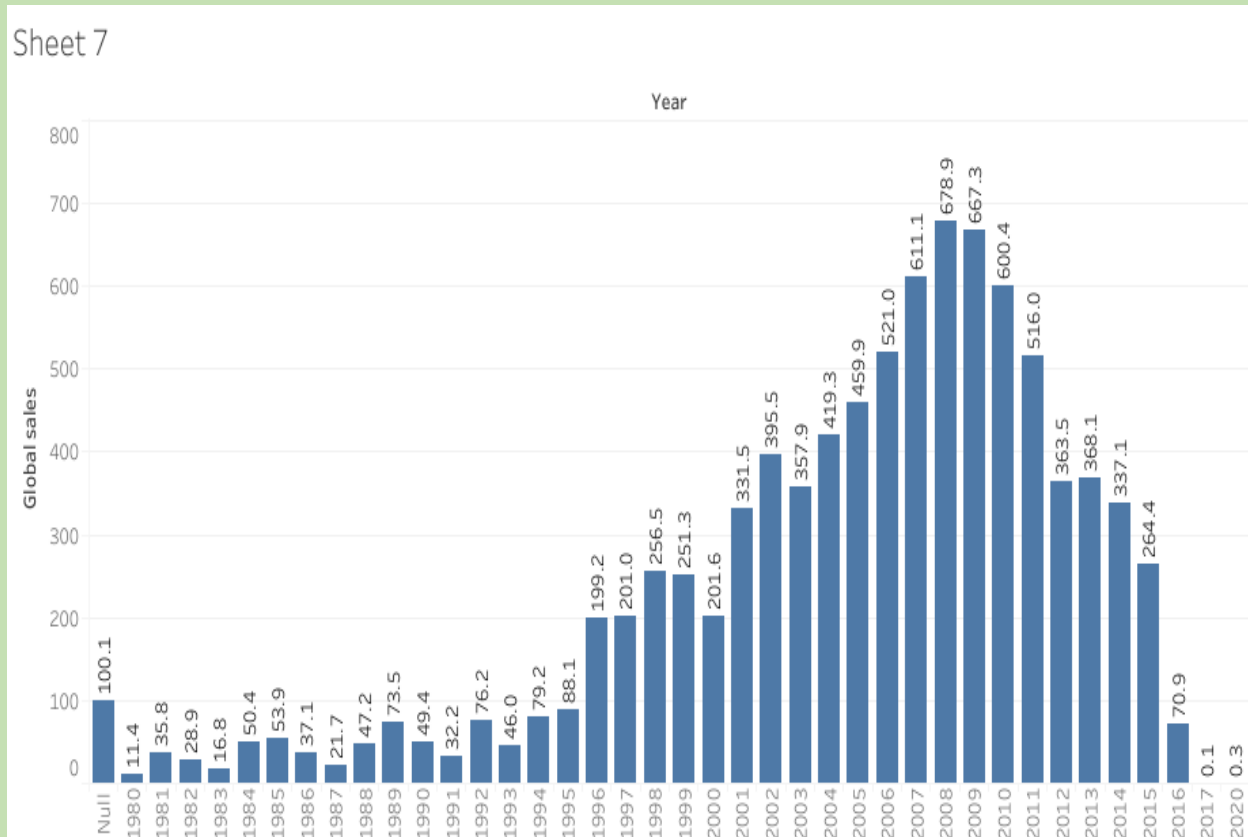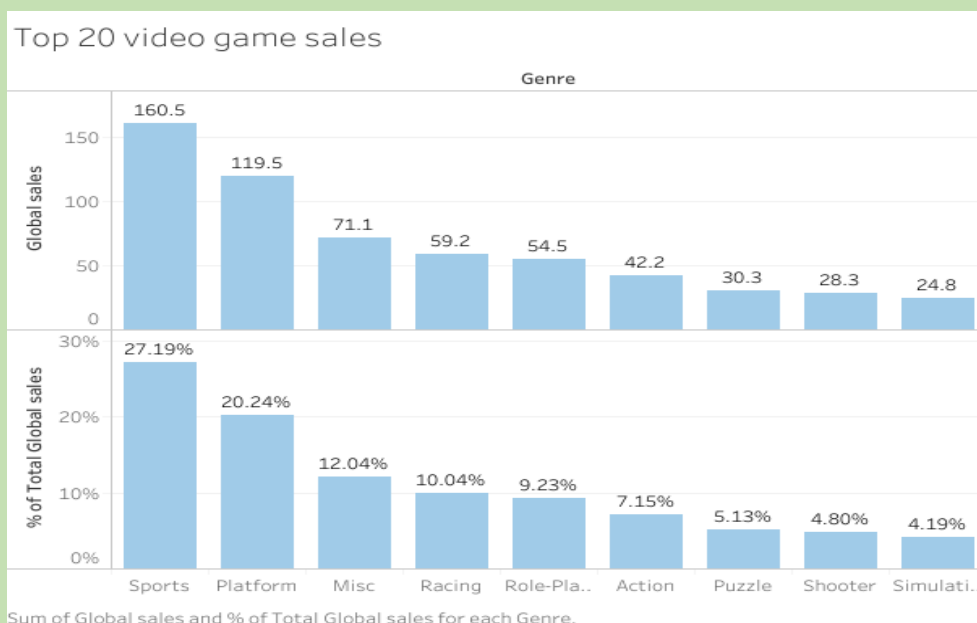
Findings 4

Figure 4; global sales by platform

This data indicates that the platform Ps2 recorded the highest global sales.

Platform vs global sales

Sum of Global sales for each Platform. The view is filtered on Platform, which excludes Platform.

Findings 5

Figure 5; Top 20 video game sales.

The data shows that the genre sports recorded the highest global sates at 27.19% amongst the top 20 video game sales.

Top 20 video game sales

Sum of Global sales and % of Total Global sales for each Genre.

Findings 6

Figure 6; Top 20 video game global sales by publisher.

This data shows that Nintendo has recorded the highest sales globally amongst the Top 20 video game sales as it is responsible for 89.15% of sales.

Sheet 6

Publisher

Sum of Global sales and % of Total Global sales for each Publisher.
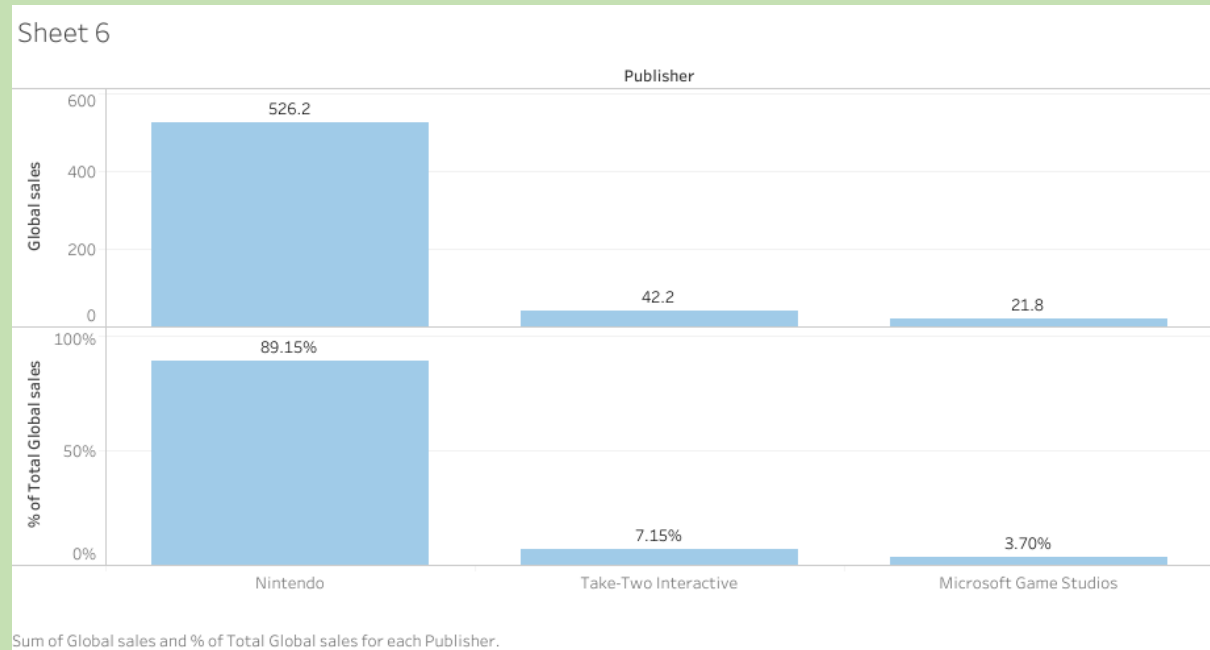
MY ANALYSIS

Based on the data analyzed there is an evident reduction in sales which could be due to many factors including; changes in the video game market of different regions, inadequate or improper marketing and advertisement strategies, high cost of video games, low reviews from users.

My deduction includes; change in video game market in different regions mainly because a genre that performs well in North America is not guaranteed to perform the same way in other regions. It also includes, improper marketing strategies, low reviews from users and cost of the video games.

The whys analysis would help determine the root course of the problem and provide possible insights on how to solve the problem.

Whys analysis

Problem; reduction in sales.

Why?

Possible reason; inadequate awareness of product.

Why?

Possible reason; improper marketing and advertising strategies.

Why?

Possible reason; difference in video game market in various region

Why?

Possible root cause; inadequate study / analysis of the dynamics of video game market in various regions.

CONCLUSION

❖ There was a continuous decline of sales after the year 2008.

❖ There are various opportunities for increase of revenue.

❖ The video game market of various regions should be adequately studied.

❖ Marketing and advertisement strategies should be developed based on the analysis of the various video game markets of various regions so as to boost sales globally.

❖ Publishing of games through Nintendo should be increased because it boosts Sales. According to this data 89.19%of the top 20 video games sales globally was contributed by Nintendo.

❖ Further analysis showing the different genres and their performances in various regions should be done. This would help in identifying what genres perform well in what regions, which can also lead to revenue increase.