CIS Final Project

1. Business analytics projects have a practical goal that analysts try to reach or questions they try to answer. Accordingly, state one or two of the most important questions you are trying to answer in your business context with the final project.

- What correlations or patterns exist between factors such as genre, platform, release year, global sales, and rating concerning sales performance, and how can these insights inform future game development and marketing strategies?

- What are the most common behaviors or patterns observed in the dataset, and how do they vary across platforms?

- What are the key trends in top video game sales over the past decade, and how have they evolved?

- Can we identify any correlations or patterns between specific game genres, platforms, and sales performance? How do factors like release timing, marketing strategies, and critical reception influence sales?

- What insights can we draw from demographic data about top-selling games? Are there notable differences in sales patterns between regions, age groups, or other demographic segments? How might these insights inform future marketing and development strategies in the gaming industry?

1. Data is an essential element of analytics. Finding a good dataset can be challenging and requires some research. Therefore, do some initial exploration and find two or three sources of data that you plan to use; list at least two sources even if you have settled with one dataset.

- <https://www.kaggle.com/datasets/gregorut/videogamesales>

- <https://www.kaggle.com/datasets/ashaheedq/video-games-sales-2019>

1. To start planning the possible data analysis, create a table(s) that describes the datasets you plan to use. This table(s) should list all the columns you plan to use and provide a brief description of each column.

Dataset 1 columns: (['Rank', 'Name', 'Platform', 'Year', 'Genre', 'Publisher', 'NA\_Sales',

'EU\_Sales', 'JP\_Sales', 'Other\_Sales', 'Global\_Sales'],

|  |  |
| --- | --- |
| Dataset 1 | <https://www.kaggle.com/datasets/gregorut/videogamesales> |
| Rank | Ranking of overall sales |
| Name | The name of the game |
| Platform | Platform of the game release (PC, PS4, Wii, Xbox, etc.) |
| Year | Year of the game release |
| Genre | Genre of the game |
| Publisher | Name of the publisher of the game |
| NA\_Sales | Sales in North America (in millions) |
| EU\_Sales | Sales in Europe (in millions) |
| JP\_Sales | Sales in Japan (in millions) |
| Other\_Sales | Sales in the rest of the world (in millions) |
| Global\_Sales | Total worldwide sales (in millions) |

Dataset 2 Columns: ['Rank', 'Name', 'Genre', 'ESRB\_Rating', 'Platform', 'Publisher', 'Developer', 'VGChartz\_Score', 'Critic\_Score', 'User\_Score', 'Total\_Shipped', 'Global\_Sales', 'NA\_Sales', 'PAL\_Sales', 'JP\_Sales', 'Other\_Sales', 'Year', 'Last\_Update', 'url', 'status', 'Vgchartzscore', 'img\_url']

|  |  |
| --- | --- |
| Dataset 2 | <https://www.kaggle.com/datasets/ashaheedq/video-games-sales-2019> |
| Rank | Ranking of overall sales |
| Name | Name of the game |
| Genre | Genre of the game |
| ESRB Rating | ‘Entertainment Software Ratings Board’ Rating |
| Platform | Platform of the game (PC, PS4, Xbox, etc.) |
| Publisher | Publisher of the game |
| Developer | Developer of the game |
| VGChartz Score | Video Game Chartz.com score <https://www.vgchartz.com/gamedb/> |
| Critic Score | Critic score of the game 1-10 |
| User Score | Users’ score of the game 1-10 |
| Total Shipped | Total shipped copies of the game |
| Global Sales | Total worldwide sales (in millions) |
| NA Sales | Sales in North America (in millions) |
| PAL Sales | Sales in Europe (in millions) |
| JP Sales | Sales in Japan (in millions) |
| Other Sales | Sales in the rest of the world (in millions) |
| Year | Year of the game release |
| Last Update | Last day of updates on the game |

1. After finding the questions and datasets you plan to use, list three steps that you plan to take toward the final project.
   1. I plan to cross-reference the data and find correlations and accuracy between both datasets and the data provided in the dataset to see which dataset will give better results.
   2. Continue research to find more insight into why the games were so popular in the year they were released and see if it had to do anything with platform releases or what the gaming environment was like at the time.
   3. Create visualizations to show trends in the demographics, year, and sales performance.
2. Prepare a short presentation (5-8 minutes) to pitch your project idea to the class.

**Deliverable**

To submit your work, create a text document and submit your answers in an organized manner. I recommend that you export your text document to a PDF. Additionally, submit the slides you have prepared for your short presentation.

-

Dataset ideas:

-call of duty dataset

-top game sales dataset

-las vegas casino revenue from 19

Dataset links:

- <https://github.com/leomaurodesenv/game-datasets/tree/master>

- <https://www.kaggle.com/datasets/faisaljanjua0555/most-played-games-of-all-time>

- <https://www.kaggle.com/datasets/gregorut/videogamesales>

-<https://github.com/vinzlercodes/Gaming-Indutry-Analysis/blob/master/data/Video_Games_Sales_as_at_22_Dec_2016%20(improved).csv>

- <https://www.kaggle.com/datasets/ashaheedq/video-games-sales-2019>

- <https://github.com/Activision/cwl-data/blob/master/data/data-2019-08-18-champs.csv>

-https://gaming.library.unlv.edu/reports/NV\_1984\_present.pdf