Patrice Shumate, Benjamin Zheng, Jimmy Nguyen, Saatvik Gangavarapu, Sumaya Shameem

Introduction

M.Sequel is creating a database that will provide information about vintage games released in the late 1900's-2000's. The data included in the database will allow users to access information such as the publisher, genre, and the number of sales each game generated in each region. The database will be based on game compatibility with three consoles developed by Nintendo: the Game Boy, the Super Nintendo Entertainment System, and the Nintendo 64. We have 20 games for each console for a total of 60 games. This will enable users to begin a comparative analysis of various entities used in the table. Most of the data in our database comes from a sales data analysis of over 16,000 games from Kaggle. The data for the developer table and the release date for each game will come from Wikipedia.

Database Description

Logical Design
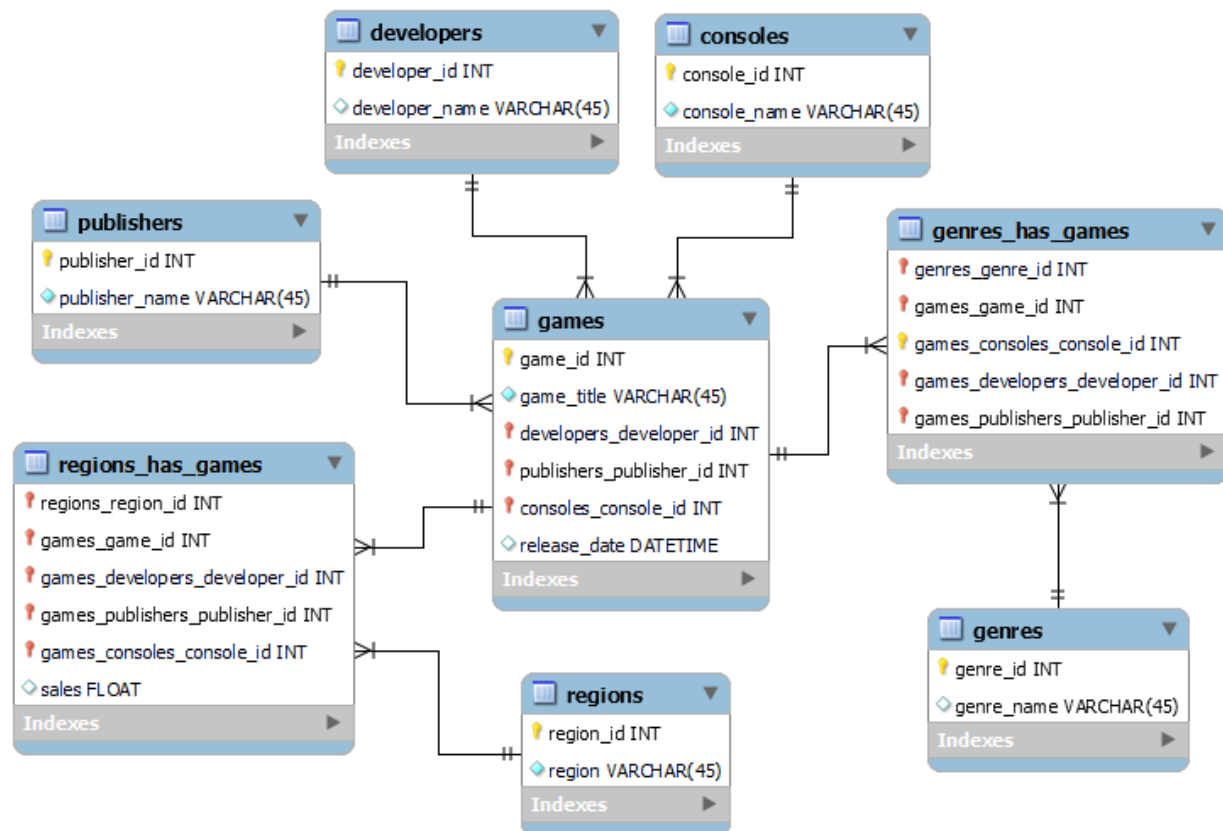
The above ERD is a visual representation of our entries video games database. This can provide information on how the tables are connected. The logical design also provides information on each table's primary and foriegn keys. This design has helped with the creation of the queries by providing insight on how to join tables to achieve the desired result set.

Physical Database

Our database consists of 8 tables focusing on video games compatible with Game Boy, Super Nintendo Entertainment System, and Nintendo 64. One of our tables is the game table itself which has the name of the game along with the publisher, developer, and console that matches with it. Our next table is the console table consisting of the three previously mentioned consoles. We then have the genre table consisting of 30 unique genres. The next two tables are the publisher and developer tables, both having unique rows. We then have the region tables consisting of North America, Japan, and the remaining regions. We then have our two linking tables. Genres_has_games links each game to whichever genres they are categorized in. Regions_has_games links each game to each region and provides the number of sales the game generated for each region.

Sample Data

Games Table

game_id	game_name	publishers_publisher_id	developers_developer_id	consoles_console_id	release_date
1	Pokemon Red/Blue	1	1	1	02-27-1996 0:00:00
2	Tetris	1	1	1	06-14-1989 0:00:00
3	Pokemon Gold/Silver	1	1	1	11-21-1999 0:00:00
4	Super Mario Land	1	1	1	04-21-1989 0:00:00
5	Dr. Mario	1	1	1	06-27-1990 0:00:00
6	Kirby's Dream Land	1	7	1	04-27-1992 0:00:00
7	F-1 Race	1	7	1	11-02-1984 0:00:00
8	Donkey Kong Land	1	2	1	06-26-1995 0:00:00
9	The Legend of Zelda: Link's Awakening	1	1	1	06-06-1993 0:00:00
10	Dragon Warrior Monsters	19	28	1	09-26-1998 0:00:00
11	Yu-Gi-Oh! Duel Monsters	17	26	1	12-16-1998 0:00:00
12	Baseball	1	1	1	12-07-1983 0:00:00
13	Disney's DuckTales	2	6	1	09-14-1989 0:00:00
14	Game de Hakken!! Tamagotchi 2	6	29	1	10-17-1997 0:00:00
15	SolarStriker	1	30	1	01-26-1990 0:00:00

There are a total of 60 games in the games table reflective of 20 video games for each of the three consoles. The sample data included in the database is based on real information and therefore we did our best to maintain that integrity throughout our implementation of the tables. With this in mind, we formatted the data in a way that would be easily comprehensible to those interested in the database.

Regions_Have_Games Table (many-to-many)

regions_region_id	games_game_id	games_developers_developer_id	games_publishers_publisher_id	games_consoles_console_id	sales
1	1	1	1	1	11.27
2	1	1	1	1	10.22
3	1	1	1	1	9.89
1	2	1	1	1	23.2
2	2	1	1	1	4.22
3	2	1	1	1	2.84
1	3	1	1	1	9
2	3	1	1	1	7.2
3	3	1	1	1	6.89
1	4	1	1	1	10.83
2	4	1	1	1	4.18
3	4	1	1	1	3.13
1	5	1	1	1	2.18
2	5	1	1	1	2

This is the linking table that links regions to games. Each game will populate three rows for each of the regions along with their respective publisher and developer for a grand total of 180 rows. The last row shows the number of sales (in millions) that each game generated for each region (NA, JP, and PAL).

Views/Queries

Query #	Req. A (4 with JOIN)	Req. B (3 with filtering)	Req. C (2 with aggregation)	Req. D (1 with linking table)	Req. E (1 with subquery)
1 - games and genre	X			X	
2 - total_sales	X		X	X	
3 - num_games_from_PAL	X	X	X	X	
4 - subquery	X	X		X	X
5	X	X	X		

Changes From Original Design

We had many changes from our initial project proposal and progress report. Firstly, as mentioned in the progress report, we now want to show how many sales a game has generated within three regions (NA, JP, and the remaining other regions). We also now have a set in stone amount of games we want for each console, being 20 per console. We were also initially doing both consoles created by both Xbox and Nintendo, however the first Xbox console was created in 2001 which doesn't fit our vintage requirement. Therefore, we decided to focus solely on Nintendo consoles. In terms of the ERD, we also made some changes. We were originally doing a many-to-many relationship between consoles and games but we realized that vintage games usually launch on only one console at the time of its release so now it is a one-to-many relationship. Our sample data had one genre for each game, but we changed it to a many-to-many relationship where a game can have multiple genres now and a genre can be for many games. Originally, we didn't know how we wanted the sales table to work but decided on a many-to-many relationship where a game can be released in multiple regions and a region can have multiple games.

Database Ethics Considerations

After assessing our database's inclusiveness, diversity, and equity, we came across a few biases that are present in our vintage video games database and also present in almost all media outlets. Accessibility to video games is a very popular concern to many video game users and publishers. Not everyone is able to access every vintage video game, nor do they have the means to, which leads to a mildly biased database. Our vintage video games database can be said to be biased towards consumers who speak certain languages specific to most video games. However, this is the case with almost all video game databases. Not all video games are catered to all video game consumers. Some vintage video games have been released only in specific countries, therefore the only languages available are the ones specific to those countries.

An accessibility concern we have regarding our database is that certain types of vintage video games are banned in many countries. For example, there are some vintage video games that are banned in Venezuela, Brazil, China, Australia, and a couple of other countries. Another good question we can ask and consider in regards to accessibility, is whether people with disabilities are able to access and play these video games as anyone else is able to. Video game

publishers have been working on inclusivity and accessibility for disabled people, such as the visually impaired. They have created and are constantly creating audio games that use synthetic speech so that blind people are able to play their games just as easily as someone who isn't visually impaired. Our database does include some games catered to people with disabilities.

As for video game inclusivity and diversity, we also have to consider the demographics of the video game creators. According to TechCrunch, content greatly depends on the producers' experiences. The International Game Developers Association conducted a survey in 2019 and came to the conclusion that 81% of video game creators worldwide are white, which shows that there is a discrepancy in diversity. Not only are most video game creators white, but the character representation in video games also reflects this. Only 10.7% of characters are black and as little as 2.7% of characters are Latinx.

When creating a database, it is possible for there to be a great amount of ethical concerns, however this does also depend on the type of database being created. Several categories have been examined prior to following through with this database idea. Our vintage video games database will provide information about vintage games released in the late 1900's- 2000's. Firstly, there is no issue regarding privacy because individual data is not being analyzed. We are analyzing data about the games, not about the users. It can be common for some to use copyright material when creating their databases, however this was not the case for us. Our database has drawn upon several sources that are publicly available. These include sales statistics and general data about the video game that is available. By retrieving information found publicly and double checking what we were using, this helped us in making sure we were adhering to all copyright laws. As a result of this, we do not foresee any legal or ethical issues occurring in the future.

Lessons Learned

There were several lessons learned through the course of this project. One issue was figuring out what direction we wanted to take for this project. One important lesson we learned was that things won't always go according to plan, especially during the brainstorming phase of the design process. We did have a general idea which we built on after much collaboration and advice. The idea was always to create a database based on vintage video games. However, we had a tough time figuring out exactly which video games and consoles to focus on.

After doing some research on database ethics, we learned that we would not have any ethical problems within our database because we were using publicly stored information. In addition to that, another key lesson came from us beginning to create the actual database. After we agreed that we wanted our database to include vintage video games, our next main issue was finding the sample data that had what we desired. For our sample data, we decided to try and use Wikipedia pages that had a list of games that were made for a console. However, we learned that that dataset would not suit our needs in regards to how we envisioned our database to look like. The Wikipedia dataset only had the publisher, developer, and the release date for different regions. After much thought and research, fortunately we came across a website called Kaggle. Kaggle is a community filled with data scientists, researchers, and students; Kaggle enabled us to find our dataset.

Another lesson we learned arose when we began creating our tables. We realized we would have to utilize the normalization process. We made sure to carry out each form in the normalization process by figuring out the relationships between tables to prevent any data redundancies and to remove any anomalies.

Potential Future Work

Potentially, there could be plans to further expand the database in order to strengthen SQL skills. This can be done through adding more data, creating more many-to-many relationships in the database, and adding more information in a few tables. One of the most prominent problems within our database resides in the database that we use in order to create it. If there are plans for future works, we would most definitely start by adding most (if not all) of the data from each console. It is also critical that our database can provide more relationships as new games provide for a number of consoles in today's age. To do this, we need to create new relationships, through the consoles and publishers table to allow for such new constructs. Additionally, there should be more information about companies who develop and publish these sorts of games. Our future work would provide more information for the developers and publishers table with such information like address of company, location of headquarters, number of staff/departments, and more. A dataset for vintage games can also be an inspiration for a future dataset for more modern games.

Citations

Smith, Gregory. "Video Game Sales." *Kaggle*, 26 Oct. 2016, Retrieved May 5, 2022, from https://www.kaggle.com/datasets/gregorut/videogamesales.

"List of Super Nintendo Entertainment System games." *Wikipedia*, Wikimedia Foundation, 1 May 2022, Retrieved May 5, 2022, from https://en.wikipedia.org/wiki/List_of_Super_Nintendo_Entertainment_System_games

"List of GameCube games." *Wikipedia*, Wikimedia Foundation, 22 Apr. 2022, Retrieved May 5, 2022, from https://en.wikipedia.org/wiki/List_of_GameCube_games

"List of Nintendo 64 games." *Wikipedia*, Wikimedia Foundation, 30 Apr. 2022, Retrieved May 5, 2022, from https://en.wikipedia.org/wiki/List_of_Nintendo_64_games