

Write-Up

For the database design, I decided to make a genre table because a genre was associated with many records in the CSV file, and created repetitive information. The table has a one-to-many relationship with a game table because each record from the CSV file showed that a single genre was used to describe many games. I added a genre ID primary key to help connect the genre table to a game table.

I decided to make a game table because a game was associated with many records in the CSV file and created repetitive information. The table has a many-to-one relationship with the genre table because the data showed that many games had the same genre. The table also has a one-to-many relationship with a release table because a game can have many releases. I added a game ID primary key because it would be needed to help connect the game table to a release table. I also needed to add a foreign key that referenced the genre ID from the genre table to connect a game to its genre.

I decided to make a publisher and platform table both with a one-to-many relationship with a release table because a platform and publisher can have many releases. I added a publisher and platform ID primary key because it would be needed to help connect the two tables to a release table.

I decided to make a release table with a many-to-one relationship with the game, publisher, and platform table because many games have multiple platforms and multiple publishers and the information caused a many-to-many relationship, so a linking table was needed. Three foreign keys were needed to connect the game, publisher, and platform IDs to the release table and a release ID primary key was created to uniquely identify all this information. I added the sales information to this table because the release table should have all unique combinations of game, publisher, and platform IDs therefore the sales should be the same for a unique release ID.

I decided not to include the rank and global sale fields in the design because that information relied on the other types of sales, and changing the value for any type of sales would cause these values to change.

The issue I ran into when migrating the data from the CSV file to the tables was that the CSV file had bad data. Two records had the same information, but different ranking, and two records had the same game, platform, and publisher IDs but different sales values. I decided to not handle this bad data, but it could be fixed if need be.