

Team Name: Gold Team

CS5200 DBMS Project - Game Ranker

Project Team: Eric Studley; Jake Dec; Tom Diaz

Team Name: Gold Team

CS5200 DBMS Project - Game Ranker

PM3: Business Insights

Videos

Intro & Q1 - <https://youtu.be/HtyMeYrZBwQ>

Q5 - <https://www.youtube.com/watch?v=BheHurM90so>

Q9 - <https://youtu.be/uUPZwaT8lFY>

## Table of Contents

<b>Value Proposition</b>	<b>4</b>
<b>Questions</b>	<b>5</b>
Q1 - What games have the highest average playtime per user that have played the game?	5
Rationale	5
Query	5
Results	6
Q2 - What games have a large number of purchases with little/no playtime?	6
Rationale	6
Query	6
Results	7
Q3 - What are the most prolific genre types?	8
Rationale	8
Query	8
Results	9
Q4 - What games are rated most highly by critics?	9
Rationale	9
Query	9
Results	10
Q5 - What are the top three games of a given genre from both critic and user scores?	10
Rationale	10
Query	11
Results	12
Q6 - What games are available on the most platforms?	12
Rationale	12
Query	12
Results	13
Q7 - Which publisher published the most owned game?	13
Rationale	13
Query	13
Results	14

Q8 - What are the highest rated genres?	14
Rationale	14
Query	14
Results	15
Q9 - What's the average playtime for the most popular genres?	16
Rationale	16
Query	16
Results	17
Q10 - What games are rated most highly by steam users?	18
Rationale	18
Query	18
Results	19
Changes to DML	20
Genre	20

## Value Proposition

- Game Ranker is a database of video games with **recommendations, reviews, average time played, and availability per console.**
- Game Ranker offers **recommendations of new games for players who want:**
  - Games of a specific genre
  - Games of over/under a specific time played
  - Games of particular quality or replay value.
- Game Ranker also allows developers to investigate information about the gaming industry. (e.g if the length of a particular game, genre, platform availability has an effect on game sales.)

## Questions

**Q1 - What games have the highest average playtime per user that have played the game?**

### Rationale

This is a query that fits in with recommending games to users. One metric that is not used in many recommendations is the actual playtime for a given game. While this is only publicly available on some platforms, such as our steam game ownership, it can provide some valuable insight on games that are a great value.

### Query

```
-- What games have the longest average playtime?
-- Filtered to games that are owned by at least 1% of our sampled
user base
-- Limited to 20 entries
-- Sorted by playtime

SELECT
    games.GameName,
    AVG(filtered.PlayTime) AS AveragePlayTime,
    COUNT(filtered.PlayTime) AS Users,
    COUNT(filtered.PlayTime) / MAX(filtered.UserIdFk) * 100 AS
PercentageOwnership
FROM
    (SELECT * FROM userhasgame
    WHERE PlayTime > 0) AS filtered
    INNER JOIN
    games ON filtered.GameIdFk = games.GameId
GROUP BY gameIdFk
HAVING COUNT(filtered.playtime) > MAX(filtered.UserIdFk) *.01
ORDER BY AveragePlayTime DESC
LIMIT 20;
```

## Results

GameName	AveragePlayTime	Users	PercentageOwnership
Arma 3	259.8666666348775	18	1.8163
Counter-Strike	256.71800001099706	50	5.1282
Counter-Strike Global Offensive	243.67881355017929	118	11.8355
Counter-Strike Source	224.81249995449824	56	5.7085
Dota 2	197.81397059238424	408	40.8000
DC Universe Online	172.89230769872665	13	1.4790
Call of Duty Modern Warfare 2 - Multiplayer	163.08571427492868	21	2.1472
Total War SHOGUN 2	161.81176469606513	17	1.7653
Sid Meier's Civilization V	153.331707322016	41	4.2797
DayZ	135.86153848612537	26	2.6236
Football Manager 2011	134.66666666666666	3	1.0345
Total War ROME II - Emperor Edition	132.24285712696258	21	2.1084
Mass Effect	121.6400000691414	10	1.1481
Warframe	108.71785715115922	28	3.1854
The Sims(TM) 3	108	1	1.7241
Rocket League	102.75294121223338	17	1.9518
The Elder Scrolls V Skyrim	101.92857147114617	56	5.6509
Path of Exile	90.02857146731445	14	1.4113
Rust	85.90769230631682	13	1.3252
Mass Effect 2	85.79999998637608	7	1.0687

## Q2 - What games have a large number of purchases with little/no playtime?

### Rationale

A game that was purchased for little to no time may indicate a game that either didn't meet the users expectations. In general we may not want to recommend games like this to users, as it presents a bad entertainment value per dollar.

### Query

```
-- What games were purchased but have little to no playtime.
-- Less than 2 hours of playtime
-- Limited to 20 entries
-- Sorted by number of users who own but haven't played the game
```

```

SELECT
    games.GameName,
    AVG(filtered.PlayTime) AS AveragePlayTime,
    COUNT(filtered.PlayTime) AS Users
FROM
    (SELECT
        *
    FROM
        userhasgame
    WHERE
        PlayTime <= 2) AS filtered
    INNER JOIN
        games ON filtered.GameIdFk = games.GameId
GROUP BY GameIdFk
ORDER BY Users DESC
LIMIT 20;

```

## Results

Result Grid       Filter Rows: <input type="text"/>   Export:    Wrap Cell Content:    Fetch n			
	GameName	AveragePlayTime	Users
▶	Dota 2	0.9757575781056376	132
	Half-Life 2 Lost Coast	0.08295454592867331	88
	Unturned	0.31111111372341343	81
	Half-Life 2 Deathmatch	0.13823529530097456	68
	Team Fortress 2	0.8631578983183492	57
	Counter-Strike Condition Zero Deleted Scenes	0.10754717071101351	53
	Warframe	0.16538461526999107	52
	Counter-Strike Condition Zero	0.08636363561857831	44
	Deathmatch Classic	0.00975609788807427	41
	Heroes & Generals	0.4609756120821325	41
	Ricochet	0.019512195412705585	41
	War Thunder	0.15250000283122062	40
	Day of Defeat	0.010256410409242678	39
	Day of Defeat Source	0.27428571517978395	35
	Counter-Strike	0.23428571500948497	35
	Counter-Strike Source	0.16764706241733887	34
	Portal	0.5588235270012828	34
	Left 4 Dead 2	0.4000000013939796	31
	Robocraft	0.17333333492279052	30
	Half-Life Opposing Force	0.06666666666666667	30

### Q3 - What are the most prolific genre types?

#### Rationale

If someone is thinking about developing, or publishing, a game, looking at how many similar games on the market may be an indication that a certain genre of game is popular, or perhaps over-crowded.

#### Query

```
-- What are the most prolific genre types?
```

```
SELECT
    genres.Genre, COUNT(genres.Genre) AS NumberGames
FROM
    gameisgenre
    INNER JOIN
    genres ON genres.GenreId = gameisgenre.GenreIdFk
GROUP BY genres.Genre
ORDER BY NumberGames DESC
LIMIT 20;
```



## Results



The screenshot shows a database application window with a 'Result Grid' tab selected. The grid displays a table with two columns: 'Genre' and 'NumberGames'. The data is sorted by 'NumberGames' in descending order. The first row is highlighted with a mouse cursor.

Genre	NumberGames
Misc	9476
Action	7667
Adventure	5293
Sports	5244
Shooter	4586
Role-Playing	4551
Platform	3445
Strategy	3266
Puzzle	3162
Racing	3030
Simulation	2737
Fighting	2085
Action	1718
General	1484
Shooter	1022
Action Adv...	697
Sports	686
Sim	649
2D	634
Sci-Fi	626

## Q4 - What games are rated most highly by critics?

### Rationale

When trying to recommend games to a user, ones that are most praised by critics is one metric that could be looked at.

### Query

```
-- Most highly rated games of all time by critics
-- Limit 20;
```

```
SELECT
    GameName, AVG(criticreviews.Score) AS AverageScore
FROM
    reviews
```

```

INNER JOIN
criticreviews ON reviews.ReviewId = criticreviews.ReviewIdFk
INNER JOIN
games ON games.GameId = reviews.GameIdFk
GROUP BY GameName
ORDER BY AverageScore DESC
LIMIT 20;

```

## Results

GameName	AverageScore
Grand Theft Auto V	96.17460317460318
Portal 2	94.96
The Last of Us Remastered	94.08823529411765
BioShock Infinite	93.625
The Legend of Zelda: Ocarina of Time 3D	93.53846153846153
Persona 4 Golden	93.53571428571429
Super Mario 3D World	92.73170731707317
Kentucky Route Zero - Act III	92.5
Super Smash Bros. for Wii U	92.36363636363636
The Witcher 3: Wild Hunt	92
Bloodborne	92
Mark of the Ninja	91.8
The Walking Dead: Episode 5 - No Time Left	91.7
Journey	91.7
Fire Emblem: Awakening	91.59375
Divinity: Original Sin II - Definitive Edition	91.5
Batman: Arkham City	91.5
Mario Kart 8 DLC Pack 2	91.5
NieR: Automata - Become as Gods Edition	91.46153846153847
Bayonetta 2	91.31481481481481

**Q5 - What are the top three games of a given genre from both critic and user scores?**

## Rationale

This query allows a user to search for the three most popular games reviewed by critics and users. It averages the average critic score with the average user score after normalizing to 0-100. These games will surely be enjoyable for fans of the queried genre. "Role-Playing" in the query can be replaced with any other genre in the Genre table.

## Query

```
-- What game from a genre has highest average critic + user score?
-- We want the average of both the critic or user score, even if a
game doesn't have both types of scores.
-- Left join from the review table to get critic reviews, user
reviews, and game genre.
-- Group by GameName
-- Limited to 3 entries
-- Sorted by combined score

SELECT
    GameName, genres.Genre,
    (AVG(IFNULL
        (criticreviews.Score, IFNULL(userreviews.Score * 10, 0)))
    +
    (AVG(IFNULL
        (userreviews.Score * 10, IFNULL(criticreviews.Score, 0)))))
    / 2 AS ComboScore
FROM
    reviews
    LEFT JOIN criticreviews
        ON reviews.ReviewId = criticreviews.ReviewIdFk
    LEFT JOIN games
        ON games.GameId = reviews.GameIdFk
    LEFT JOIN userreviews
        ON reviews.ReviewId = userreviews.ReviewIdFk
    LEFT JOIN gameisgenre
        ON games.GameId = gameisgenre.GameIdFk
    LEFT JOIN genres
        ON genres.GenreId = gameisgenre.GenreIdFk
GROUP BY GameName
HAVING genres.Genre = "Role-Playing"
ORDER BY ComboScore DESC
LIMIT 10;
```

## Results

	GameName	Genre	ComboScore
►	Skies of Arcadia	Role-Playing	99.23076923076923
	Paper Mario	Role-Playing	96.20253164556962
	Persona 4 Golden	Role-Playing	93.88368580060423
	Baldur's Gate II: Shadows of Amn	Role-Playing	93.36941813261164
	Persona 5	Role-Playing	91.935303514377
	Undertale	Role-Playing	91.52173913043478
	Divinity: Original Sin II - Definitive Edition	Role-Playing	91.41379310344827
	NieR: Automata - Become as Gods Edition	Role-Playing	90.13333333333334
	The Witcher 3: Wild Hunt - Hearts of Stone	Role-Playing	89.83870967741936
	The Witcher 3: Wild Hunt - Blood and Wine	Role-Playing	89.16455696202532

## Q6 - What games are available on the most platforms?

### Rationale

This query finds which games are available on a large amount of platforms. Users may want a game they can play on their phone, on their computer, and on their gaming console at any point in time.

### Query

```
-- What game from the games table has the most entries in the
gameonplatform table?
```

```
-- Group by GameName
```

```
-- Sorted by platform count
```

```
-- Limited to 10 entries
```

```
SELECT
    GameName, COUNT(gameonplatform.PlatformIdFk) as PlatformCount
FROM
    games
    INNER JOIN
    gameonplatform ON gameonplatform.GameIdFk = games.GameId
GROUP BY GameName
ORDER BY PlatformCount DESC
LIMIT 10;
```

## Results

	GameName	PlatformCount
►	Plants vs. Zombies	20
	Monopoly	15
	Double Dragon	14
	Samurai Shodown	13
	Space Invaders	12
	Elite	12
	Jeopardy!	12
	Lemmings	12
	Pier Solar and the Great Architects	12
	Wheel of Fortune	12

## Q7 - Which publisher published the most owned game?

### Rationale

This query finds which publisher published the game owned by the most individual users. This would allow a user to find who created the most popular game and check other games created by that publisher as they're obviously wildly popular.

### Query

```
-- What publisher owns the game with the highest users owning that game?
```

```
-- Group by GameIdFk
```

```
-- Sorted by count of games
```

```
-- Limited to the top entry
```

```
SELECT PublisherName, COUNT(*) AS GameCount
FROM userhasgame
    INNER JOIN
        games ON games.GameId = userhasgame.GameIdFk
    INNER JOIN
        publishers ON publishers.PublisherId = games.PublisherIdFk
GROUP BY GameIdFk
ORDER BY GameCount DESC
LIMIT 1;
```

## Results

	PublisherName	GameCount
▶	Valve Corporation	408

## Q8 - What are the highest rated genres?

### Rationale

Users may not know the most popular games of the year, let alone the most popular genres. Filtering by the most popular genre across critics is another way in which the user may tract down a new favorite.

This query will showcase the top-ten genres, grouped by average critic score.

### Query

```
#What are the highest rated genres?
#Joined lookup tables for genres onto games, and from critic reviews
onto reviews
#Reviews and games joined to cross-reference critic reviews and game
genres
#Top 10 genres listed; grouped by their average critic score
#Descending order, with the top ranked genre displayed first.
USE GameRater;
SELECT Genre, Avg(Score) 'Average Critic Score' from games
LEFT JOIN GameIsGenre on GameIsGenre.GameIdFK = games.GameId
LEFT JOIN Genres on Genres.GenreId = GameIsGenre.GenreIdFk
LEFT JOIN Reviews on games.GameId = Reviews.GameIdFk
LEFT JOIN CriticReviews on Reviews.ReviewId =
CriticReviews.ReviewIdFk
GROUP BY Genre
ORDER BY 2 Desc
LIMIT 10;
```

## Results

Genre	Average Critic Score
Application	88.2666666666667
City Building	85.55
Basketball	84.66350067842605
Console-style RPG	84.57916666666667
Music Maker	84.38181818181818
Action RPG	84.17016381373091
Massively Multiplayer Online	84.00878293601004
Rail	84.00735294117646
Scrolling	83.958751393534
Motocross	83.9

## Q9 - What's the average playtime for the most popular genres?

### Rationale

Some genres may be generously ranked, or reviewed, however offer little replay value. Consequently, it can be in a user's best interest instead to search for the average playtime across the most popular genres. This instead can allow the user to find the genre that offers the most bang for their buck.

This query showcases the top twenty genres across users, grouped by users' average playtime.

### Query

```
#What is the average playtime for the most popular genres?
#Join genre lookup tables onto games, and whether a user owns a game
onto the game table
#Group entries by the genre, and calculate their average score
#Show the top 20 genres by average playtime
Use GameRater;
SELECT Genre, Avg(Playtime)as 'Average Playtime' from games
LEFT JOIN GameIsGenre on GameIsGenre.GameIdFK = games.GameId
LEFT JOIN Genres on Genres.GenreId = GameIsGenre.GenreIdFk
INNER JOIN UserHasGame on Games.GameId = UserHasGame.GameIdFk
GROUP BY Genre
ORDER BY 2 DESC
LIMIT 20;
```



## Results

Genre	Avg(Playtime)
Football	457.8
Traditional	314.8857142789023
Management	281.1318840533495
MOBA	188.09093023642552
Soccer	177.70784315875932
Team	168.36851854236036
Sports	149.3496093993308
Real-Time	143.93108808345744
Strategy	141.63654240026662
4X	104.67343749722932
Fantasy	93.92223416031943
Breeding/Constructing	89.79999987284343
Futuristic	83.47857145894142
MMO	77.11587301742226
Sim	73.40000004238553
Turn-Based	68.7543689296807
Sandbox	59.31683673998531
Tactical	58.81707747266169
Japanese-Style	57.673469416036895
Western-Style	55.1284090697088

## Q10 - What games are rated most highly by steam users?

### Rationale

Sometimes the user review is more valuable than a critic review. Moreover, it's possible that the user only utilizes Steam for gaming, rather than gaming consoles. By looking at games rated most highly only by steam users, this removes confounding data of those that game on other platforms.

This query displays the top twenty games, grouped by their average user rating, where the steam id must be populated. A Steam Id is unique to steam users. In the event the database is populated with a null steam value, this entry would not be included in the result set.

### Query

```
#What Games Are Rated Most Highly By Steam Users?
#join user reviews onto the user, then cross-reference whether that
particular owns a game
#group the game names by average user ratings
#only show the top 20 results and where the steamid is not null
#if in the data set the steam id was null, this would indicate the
user was not a steam user.
```

```
Use GameRater;
SELECT GameName, Avg(Score) as 'Average User Rating' from games
INNER JOIN UserHasGame on games.GameId = UserHasGame.GameIdFk
INNER JOIN Users on UserHasGame.UserIdFK = Users.UserId
INNER JOIN Reviews on games.GameId = Reviews.GameIdFk
INNER JOIN UserReviews on Users.UserID = UserReviews.UserIdFk
WHERE SteamID IS NOT NULL
GROUP BY GameName
ORDER BY 2 DESC
LIMIT 20;
```

## Results

GameName	Average User Rating
Bastion	10
Undertale	10
Half-Life 2	10
The Binding of Isaac	10
The Vanishing of Ethan Carter	10
Astebreed	10
Super Meat Boy	10
Braid	10
LIMBO	10
Terraria	10
VVVVVV	10
HAWKEN	10
Endless Legend	10
DARK SOULS II	10
Xeodrifter	10
Ikaruga	10
Football Manager 2013	9.4
Dota 2	6.666666666666667

## Changes to DML

### Genre

In the data set of video games from metacritic, genres names were entered as a flat string with multiple genres concatenated by semicolons. A [python script was created](#) to generate a new csv with those genres unrolled to allow us to insert the individual genres into our database. This gave us much more insight into which genres were available and which games had been considered a certain type. No changes were made to our team's design or UML or DDL; a new LOAD DATA with the generated csv was added in the DML file.