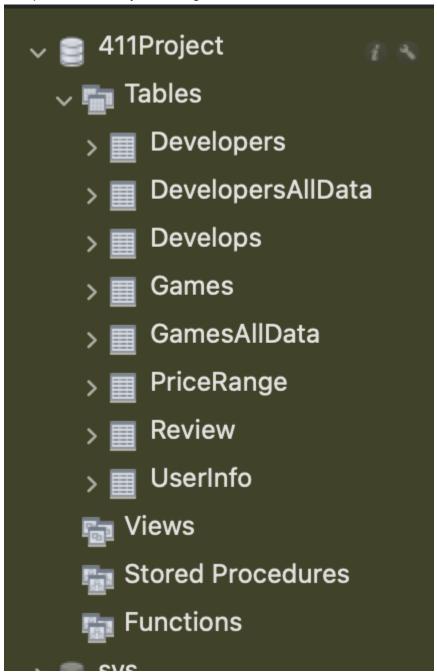
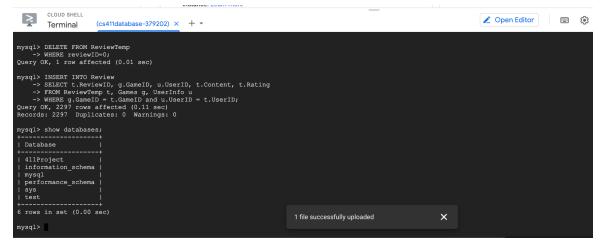
1. Implement at least 4 main tables

We implemented 6 tables: Developers, Develops, Games, PriceRange, Review, and UserInfo. DevelopersAllData and GamesAllData are helper tables that we created when uploading the csv file to the database, and we will not be using these two tables in our final product, so they can be ignored.



Here is a picture of the connection to GCP:



2. In the Database Design markdown or pdf, provide the Data Definition Language (DDL) commands you all used to create each of these tables in the database.

```
CREATE TABLE Games
  GameID INT primary key.
  ResponseName varchar(256),
  ReleaseDate varchar(100),
  RecommendationCount INT.
  AchievementCount int,
  ControllerSupport
                    varchar(10),
  IsFree varchar(10),
  FreeVerAvail varchar(10),
  PurchaseAvail varchar(10),
  SubscriptionAvail varchar(10),
  PlatformWindows varchar(10),
  PlatformLinux varchar(10),
  PlatformMac varchar(10),
  CategorySinglePlayer varchar(10),
  CategoryMultiplayer varchar(10),
  CategoryCoop varchar(10),
  CategoryMMO varchar(10),
  CategoryInAppPurchase varchar(10),
  CategoryIncludeSrcSDK varchar(10),
  CategoryIncludeLevelEditor varchar(10),
  CategoryVRSupport varchar(10),
  GenrelsNonGame varchar(10),
  GenrelsIndie varchar(10),
  GenrelsAction varchar(10),
  GenrelsAdventure varchar(10),
  GenrelsCasual varchar(10),
  GenrelsStrategy varchar(10),
```

```
GenrelsRPG varchar(10),
  GenrelsSimulation varchar(10),
  GenrelsEarlyAccess varchar(10),
  GenrelsFreeToPlay varchar(10),
  GenrelsSports varchar(10),
  GenrelsRacing varchar(10),
  GenrelsMassivelyMultiplayer varchar(10),
  PriceFinal DOUBLE,
  SupportedLanguages varchar(512)
);
CREATE TABLE Developers
  DeveloperID INT not null Primary Key,
  Developer varchar(256),
  Country varchar(100),
  Notes varchar(300)
);
CREATE TABLE Develops (
  GameID int NOT NULL,
  DeveloperID int NOT NULL,
  PRIMARY KEY (GameID),
  FOREIGN KEY (DeveloperID) REFERENCES Developers (DeveloperID),
  FOREIGN KEY (GameID) REFERENCES Games (GameID)
);
Create Table UserInfo(
      UserID INT not null Primary Key,
      Email VARCHAR(100),
      Username VARCHAR(20),
      Password VARCHAR(40),
      DeviceBackground VARCHAR(10)
);
Create Table Review(
      ReviewID int not null primary key,
      GameID int.
      UserID int,
      Content VARCHAR(1000),
      Rating INT,
      Foreign key (GameID) references Games(GameID),
      Foreign key (UserID) references UserInfo(UserID)
```

```
);
Create Table PriceRange(
Grade int not null primary key,
LowerPrice INT,
UpperPrice INT
);
```

3. Insert data into these tables. You should insert at least 1000 rows each in three of the tables. Try to use real data, but if you cannot find a good dataset for a particular table, you may use auto-generated data.

We have inserted data to all the tables that we have implemented. Among these data, UserInfo and UserReviews are auto generated.

```
mysql> select count(UserID) from UserInfo;
+-----+
| count(UserID) |
+-----+
| 2999 |
+-----+
1 row in set (0.00 sec)
```

```
mysql> select count(ReviewID) from Review;
+-----+
| count(ReviewID) |
+-----+
| 2297 |
+-----+
1 row in set (0.01 sec)
```

Games table is populated with selected columns of our original dataset.

```
mysql> select count(gameID) from Games
    ->;
+-----+
| count(gameID) |
+-----+
| 13356 |
+-----+
1 row in set (0.01 sec)
```

For Price Range, we divide price ranges into different grades, every 5 dollar is a different grade, up to \$2500, which is a little more than the cost of the most expensive game on steam.

```
mysql> select count(grade) from PriceRange;
+-----+
| count(grade) |
+-----+
| 499 |
+-----+
1 row in set (0.00 sec)
```

For Developers, we could not find any datasets or tools that find game developers for us. Therefore, we pulled a list of game developers (companies) from wikipedia including 700+ different companies. We then randomly assigned each game with a random game developer using the following python code:

```
[19] df = pd.read_csv("test1.csv")
[21] df['developerID'] = [random.randint(0, 722) for i in range(0, len(df))]
```

We will be clarifying this in our final product to make sure no confusion arises.

We also implemented one relation: develops relation. In this relation, each game has a developer. Here is a count of the number of games in the develops relation (should be equal to the number of games):

```
mysql> select count(GameID) from Develops;
+-----+
| count(GameID) |
+-----+
| 13344 |
+-----+
1 row in set (0.01 sec)
```

- 4. As a group, develop two advanced SQL queries related to the project that are different from one another. The two advance queries are expected to be part of your final application. The queries should each involve at least two of the following SQL concepts: Join of multiple relations, Set operations, Aggregation via GROUP BY, Subqueries
- 5. Execute your advanced SQL queries and provide a screenshot of the top 15 rows of each query result (you can use the LIMIT clause to select the top 15 rows). If your output is less than 15 rows, say that in your output.

## Query 1

SELECT g.ResponseName AS GameName, AVG(r.Rating) AS AverageRating, g.ReleaseDate AS ReleaseDate, d.Developer AS DeveloperName FROM Games g JOIN Develops dv ON g.GameID = dv.GameID JOIN Developers d ON dv.DeveloperID = d.DeveloperID JOIN Review r ON g.GameID = r.GameID WHERE r.Rating > 4.0 GROUP BY g.ResponseName, g.ReleaseDate, d.Developer

This query used Join of multiple relations and Aggregation via GROUP BY.

ORDER BY AVG(r.Rating) DESC:

This query gets the list of top-rated games along with their average rating, release date, and the names of the developers who developed them, for all games that have a rating greater than 4.0. At the end of the query, we added LIMIT 15 to only show the top 15 results.

## Query 2

SELECT ResponseName, PlatformWindows, PlatformLinux, PlatformMac, PriceFinal FROM Games

WHERE PriceFinal >= (SELECT LowerPrice FROM PriceRange WHERE Grade = 3)

AND PriceFinal <= (SELECT UpperPrice FROM PriceRange WHERE Grade = 3)

AND PlatformWindows = 'True'

UNION

SELECT ResponseName, PlatformWindows, PlatformLinux, PlatformMac, PriceFinal FROM Games

WHERE PriceFinal >= (SELECT LowerPrice FROM PriceRange WHERE Grade = 3)

AND PriceFinal <= (SELECT UpperPrice FROM PriceRange WHERE Grade = 3)

AND PlatformLinux = 'True'

UNION

SELECT ResponseName, PlatformWindows, PlatformLinux, PlatformMac, PriceFinal FROM Games

WHERE PriceFinal >= (SELECT LowerPrice FROM PriceRange WHERE Grade = 3)

AND PriceFinal <= (SELECT UpperPrice FROM PriceRange WHERE Grade = 3)

AND PlatformMac = 'True';

This query uses subqueries set operations.

This query find the game name, platform (Windows, Linux, or Mac), and price, for all games with a price within the specified range (in this case, grade 3) for any of the three platforms. At the end of the query, we added LIMIT 15 to only show the top 15 results.

```
mysql> SELECT ResponseName, PlatformWindows, PlatformLinux, PlatformMac, PriceFinal
      -> WHERE PriceFinal >= (SELECT LowerPrice FROM PriceRange WHERE Grade :
     -> AND PriceFinal <= (SELECT UpperPrice FROM PriceRange WHERE Grade = 3)
-> AND PlatformWindows = 'True'
      -> SELECT ResponseName, PlatformWindows, PlatformLinux, PlatformMac, PriceFinal
     -> FROM Games
-> wHERE PriceFinal >= (SELECT LowerPrice FROM PriceRange WHERE Grade = 3)
-> AND PriceFinal <= (SELECT UpperPrice FROM PriceRange WHERE Grade = 3)
-> AND PlatformLinux = 'True'
      -> SELECT ResponseName, PlatformWindows, PlatformLinux, PlatformMac, PriceFinal
      -> WHERE PriceFinal >= (SELECT LowerPrice FROM PriceRange WHERE Grade
     -> AND PriceFinal <= (SELECT UpperPrice FROM PriceRange WHERE Grade = 3)
-> AND PlatformMac = 'True
     -> LIMIT 15;
  ResponseName
                                                                                        | PlatformWindows | PlatformLinux | PlatformMac | PriceFinal |
  Counter-Strike: Global Offensive
                                                                                                                                                                    14.99
                                                                                          True
  Dangerous Waters
Space Empires V
                                                                                                                                         False
False
                                                                                                                                                                    14.99
14.99
                                                                                                                    False
                                                                                          True
                                                                                                                   False
  Quake III Arena
                                                                                          True
                                                                                                                   False
                                                                                                                                         False
False
                                                                                                                                                                    14.99
14.99
  Quake IV
                                                                                          True
                                                                                                                   False
  QUAKE III: Team Arena
Warhammer(r) 40000: Dawn of War(r) - Game of the Year Edition
Warhammer(r) 40000: Dawn of War(r) - Dark Crusade
Condemned: Criminal Origins
                                                                                                                    False
                                                                                                                                         False
                                                                                                                                                                    14.99
                                                                                                                                         False
                                                                                                                                                                    12.99
                                                                                          True
                                                                                                                   False
                                                                                                                   False
                                                                                                                                         False
False
                                                                                                                                                                    12.99
                                                                                                                   False
                                                                                          True
                                                                                                                                                                    14.95
  Medieval II: Total War(tm) Kingdoms
Heroes of Annihilated Empires
                                                                                                                    True
                                                                                                                                         True
False
                                                                                                                                                                    11.99
                                                                                                                                                                    14.99
                                                                                          True
                                                                                                                   False
  Lost Planet(tm): Extreme Condition
Jade Empire(tm): Special Edition
                                                                                                                   False
                                                                                                                                         False
                                                                                          True
                                                                                                                                         False
                                                                                                                   False
                                                                                                                                                                    14.99
  Just Cause 2
Master Levels for Doom II
                                                                                                                   False
False
                                                                                                                                          False
                                                                                                                                                                    14.99
15 rows in set (0.00 sec)
```

6. Indexing: As a team, for each advanced query:

Use the EXPLAIN ANALYZE command to measure your advanced query performance before adding indexes.

Explore adding different indices to different attributes on the advanced query. For each indexing design you try, use the EXPLAIN ANALYZE command to measure the query performance after adding the indices.

Report on the index design you all select and explain why you chose it, referencing the analysis you performed in (b).

Note that if you did not find any difference in your results, report that as well. Explain why you think this change in indexing did not bring a better effect to your query.

## Query 1

### Without Index:

## Index 1:

CREATE INDEX RDIndex ON Games (ReleaseDate);

## **ALTER TABLE Games**

## DROP INDEX RDIndex;

This line of sql command is to drop this index so it does not affect our later testing.

#### Index 2:

CREATE INDEX RTIndex ON Review (Rating);

```
mysql: XVIAIN NAINIZE

-> SELECT g.ResponseName AS GameName, ANG(r.Rating) AS AverageRating, g.ReleaseDate AS ReleaseDate, d.DeveloperID JOIN Review r ON g.GameID = r.GameID

-> MIRESE r.Rating > 4.0

-> GROUP BY g.ResponseName, g.ReleaseDate, d.Developer

-> ORICE BY AVG(r.Rating) DESC;

| EXFLAIN |

| -> Sort: AverageDating DESC (actual time=11.560.11.731 row=1303 loop=1)

-> Nails sense on ecemporary (actual time=0.002.0.160 cow=1309 loop=1)

-> Appropriate using responsely table (actual time=0.574.10.827 row=1303 loop=1)

-> Nested loop inner join (cost=1622.5 row=1381) (actual time=0.002.0.9.85 row=1381 loop=1)

-> Nested loop inner join (cost=1622.5 row=1381) (actual time=0.003.0.985 row=1381 loop=1)

-> Filter: ((r.Rating > 4) and (r.GameID is not null)) (cost=0.22.20 row=1381 (actual time=0.003.0.0985 row=1381) opp=1)

-> Single-row index lookup on dw using FRIMARY (GameID=GameID) (cost=0.25 row=1) (actual time=0.001.0.001 row=1 loop=181)

-> Single-row index lookup on dw using FRIMARY (BeveloperID=dv.DeveloperID) (cost=0.25 row=1) (actual time=0.001.0.001 row=1 loop=181)

-> Single-row index lookup on dw using FRIMARY (BeveloperID=dv.DeveloperID) (cost=0.25 row=1) (actual time=0.001.0.001 row=1 loop=181)

-> Single-row index lookup on dw using FRIMARY (BeveloperID=dv.DeveloperID) (cost=0.25 row=1) (actual time=0.001.0.001 row=1 loop=181)

-> Single-row index lookup on dw using FRIMARY (BeveloperID=dv.DeveloperID) (cost=0.25 row=1) (actual time=0.001.0.001 row=1 loop=181)
```

**ALTER TABLE Review** 

DROP INDEX RTIndex;

Index 3:

CREATE INDEX Devindex ON Develops (GameID, DeveloperID);

```
PAYED ANTER TRANSE SERVICE

SINGE PRODUCTS TO DEPOSITE SERVICE

SINGE PRODUCTS TO DEPOSITE SERVICES

ORREY OR, O rows affected (0.02 sec)

Records: O Deplicates: O Marnings: O

SUMPLIANCE THERE NEWL PROVIDES (GAMEID, DeveloperID);

ORREY OR, O rows affected (0.13 sec)

Records: O Deplicates: O Warnings: O

SAMPLE SEPIAMIN MENTER

SELECT G. Records: O Deplicates: O Warnings: O

SAMPLE SEPIAMIN MENTER

SELECT G. Records: AS Gameinse, AVG(r. Rating) AS AverageBating, g.ReleaseDate AS ReleaseDate, d.Developer AS DeveloperRome

SELECT G. Records: AS Gameinse, AVG(r. Rating) AS AverageBating, g.ReleaseDate AS ReleaseDate, d.Developer ID JOIN Review r ON q.GameID = r.GameID

MIREM F.Mating) 125C;

SOURCE BY AVERGAMENT AS AVERAGE AS AVERAGED ONLY Developer of ON dv.DeveloperID = d.DeveloperID JOIN Review r ON q.GameID = r.GameID

**ADDITIONAL PROPERTY OF A SECTION OF A SE
```

# ALTER TABLE Develops DROP INDEX DevIndex;

It seems that Index 2 is the fastest. The total actual time for index 2 is 10.574 - 11.731, which is the fastest among the four. Additionally, index 2's query plan estimates that there are 1381 rows that satisfy the join conditions between the Review, Games, and Develops tables, which is the most among all other plans. The cost for index 2's query plan is higher is explainable because there are more matching rows between tables in index 2's query plan. Therefore, we will be using index plan 2.

Query 2 Without Index:

```
Septimental Management (State Constructions, Flatfordinos, Flatfordinos, Flatfordinos, Flatfordinos, Flatfordinos, Flatfordinos (State - 3)

- Mills Finishinal > (ELLIC Lowerines Finish Flatfordinos) (Mills of the Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finish Flatfordinos) (Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finish Flatfordinos) (Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finish Flatfordinos) (Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finish Flatfordinos) (Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finish Flatfordinos) (Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishina) (Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishina) (Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 3)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 4)

- Mills Finishinal > (ELLIC Lowerines Finishinal - 4)

- Mills Finishinal - (ELLIC Lowerines Finishinal - 4)

- Mills Finishinal - (ELLIC Low
```

## Index Plan 1:

## CREATE INDEX PFIdx ON Games (PriceFinal);

```
Secretary of the form of force (pitterfunks)

Secretary of the form of force (pitterfunks)

Secretary of the form of force (pitterfunk)

Secretary of the form of force (pitterfunk)

Secretary of the force (pitterfunk)

SECRIT Managements, Pitterfunks, Pitterfunks, Pitterfunks

SECRIT Managements, Pitterfunks (collection for Pitterfunks)

SECRIT Managements, Pitterfunks (collection for Pitterfunks)

SECRIT Managements, Pitterfunks, Pitterfunks, Pitterfunks

SECRIT Managements, Pitterfunks, Pitterfu
```

ALTER TABLE Games DROP INDEX PFIdx;

Index Plan 2: CREATE INDEX WIdx ON Games (PlatformWindows); CREATE INDEX LIdx ON Games (PlatformLinux); CREATE INDEX MIdx ON Games (PlatformMac);

```
ql> CREATE INDEX Widx ON Games (PlatformWindows);
ry OK, 0 rows affected (0.15 sec)
ords: 0 Duplicates: 0 Warnings: 0
   mysql> CREATE INDEX LIdx ON Games (PlatformLinux);
Query OK, 0 rows affected (0.18 sec)
Records: 0 Duplicates: 0 Warnings: 0
   mysql> CREATE INDEX MIdx ON Games (PlatformMac);
Query OK, 0 rows affected (0.13 sec)
Records: 0 Duplicates: 0 Warnings: 0
   mysql> EXPLAIN ANALYZE

-> SELECT ResponseName, PlatformWindows, PlatformLinux, PlatformMac, PriceFinal
-> FROM Games

-> WHERE PriceFinal >= (SELECT LowerPrice FROM PriceRange WHERE Grade = 3)
-> AND PriceFinal -= (SELECT UpperPrice FROM PriceRange WHERE Grade = 3)
-> AND PlatformWindows = 'True'
-> INTOM
                                                                               sponseName, PlatformWindows, PlatformLinux, PlatformMac, PriceFinal
                                                 M Games
RE PriceFinal >= (SELECT LowerPrice FROM PriceRange WHERE Grade = 3)
PriceFinal <= (SELECT UpperPrice FROM PriceRange WHERE Grade = 3)
PlatformLinux = 'True'
                                      ELECT ResponseName, PlatformWindows, PlatformLinux, PlatformMac, PriceFinal
                                 FROM Games
WHERE PriceFinal >= (SELECT LowerPrice FROM PriceRange WHERE Grade = 3)
AND PriceFinal <= (SELECT UpperPrice FROM PriceRange WHERE Grade = 3)
AND PlatformMac = 'True';
      EXPLAIN
                       Table scan on cunion temporary> (cost=0.01..22.34 rows=1587) (actual time=0.003..0.160 rows=1254 loops=1)

>> Union materialize with deduplication (cost=967.67..989.99 rows=1587) (actual time=41.675..41.902 rows=1254 loops=1)

-> Filter: ((Games.PriceFinal >= (select #2)) and (Games.PriceFinal <= (select #3))) (cost=290.82 rows=741) (actual time=0.375..26.107 rows=1279 loops=1)

-> Index lookup on Games using Widx (FlatfornWindows=True*) (cost=290.82 rows=6688) (actual time=0.267..24.750 rows=13354 loops=1)

-> Select #2 (subquery in condition; run only once)

-> Rows fetched before execution (cost=0.00..0.00 rows=1) (actual time=0.000..0.000 rows=1 loops=1)

-> Filter: ((Games.PriceFinal >= (select #5)) and (Games.PriceFinal <= (select #6))) (cost=250.71 rows=340) (actual time=0.146..5.382 rows=371 loops=1)

-> Index lookup on Games using Lidx (FlatfornLinux=True*) (cost=250.71 rows=3057) (actual time=0.126..5.120 rows=3057 loops=1)

-> Select #5 (subquery in condition; run only once)

-> Rows fetched before execution (cost=0.00..0.00 rows=1) (actual time=0.000..0.000 rows=1 loops=1)

-> Filter: ((Games.PriceFinal >= (select #5)) and (Games.PriceFinal <= (select #9)) (cost=250.72 rows=3057 loops=1)

-> Filter: ((Games.PriceFinal >= (select #9)) and (Games.PriceFinal <= (select #9)) (cost=250.72 rows=3057 loops=1)

-> Rows fetched before execution (cost=0.00..0.00 rows=1) (actual time=0.000..0.000 rows=1 loops=1)

-> Select #6 (subquery in condition; run only once)

-> Rows fetched before execution (cost=0.00..0.00 rows=1) (actual time=0.000..0.000 rows=1 loops=1)

-> Select #6 (subquery in condition; run only once)

-> Rows fetched before execution (cost=0.00..0.00 rows=1) (actual time=0.000..0.000 rows=1 loops=1)

-> Select #6 (subquery in condition; run only once)

-> Rows fetched before execution (cost=0.00..0.00 rows=1) (actual time=0.000..0.000 rows=1 loops=1)

-> Select #6 (subquery in condition; run only once)

-> Rows fetched before execution (cost=0.00..0.00 rows=1) (actual time=0.000..0.000 rows=1 lo
1 row in set (0.04 sec)
```

ALTER TABLE Games DROP INDEX WIdx; ALTER TABLE Games DROP INDEX LIdx; ALTER TABLE Games DROP INDEX MIdx;

### Index Plan 3:

CREATE INDEX Priceldx ON Games (PriceFinal, PlatformWindows, PlatformLinux, PlatformMac);

```
Section 10 Delicates | Section 10 Delicates |
```

## ALTER TABLE Games DROP INDEX Priceldx;

It seems that Index plan 3 is the fastest. The total actual time for index 3 is 8.129-8.356, which is the fastest among the four. Additionally, index 3's query plan estimates that there are 1279 rows that satisfy the conditions for each query. The cost for index 3's query plan is higher is explainable because there are more matching rows between tables in index 3's query plan. Therefore, we will be using index plan 3.