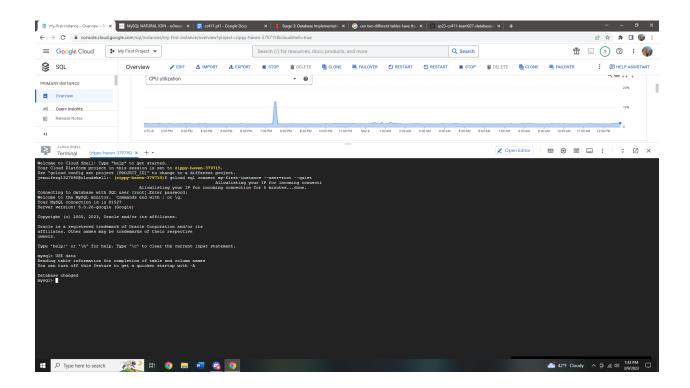
Divya Machineni Michael Ma Jennifer Gaw Kaan Yigit

CONNECTION:



DDL COMMANDS:

```
CREATE TABLE GameData (
     GameID int,
     GameName varchar (255),
     Description varchar (1000),
     Rating int,
     Price REAL,
     Reviews varchar (255),
     PCreq varchar (1000),
     Developer int,
     RequiredAge int,
     Language varchar (255),
     PRIMARY KEY (GameID)
);
CREATE TABLE GenreData (
     GameID int,
     Singleplayer varchar(10),
     Multiplayer varchar (10),
     Coop varchar (10),
     MMO varchar(10),
     InAppPurchase varchar(10),
     VRSupport varchar (10),
     NonGame varchar (10),
     Indie varchar(10),
     Action varchar (10),
     Adventure varchar (10),
     Casual varchar(10),
     Strategy varchar (10),
     RPG varchar (10),
     Simulation varchar (10),
     EarlyAccess varchar(10),
     FreeToPlay varchar(10),
     Sports varchar(10),
     Racing varchar (10),
     MassivelyMultiplayer varchar(10),
     PRIMARY KEY (GameID)
     FOREIGN KEY (GameID) REFERENCES Game(GameID)
```

```
CREATE TABLE Region (
    RegionName varchar(255),
    PRIMARY KEY (RegionName)
);

CREATE TABLE PriceRange (
    Range int,
    PriceRangeMin int,
    PriceRangeMax int,
    PRIMARY KEY (Range)
);
```

TABLE COUNTS:

```
mysql> SELECT COUNT(*) FROM GameData
-> ;
+-----+
| COUNT (*) |
| 13358 |
1 row in set (0.01 sec)
mysql> SELECT COUNT(*) FROM GenreData;
| COUNT (*) |
| 13358 |
1 row in set (0.01 sec)
mysql> SELECT COUNT(*) FROM Region;
| COUNT (*) |
| 1000 |
+-----
1 row in set (0.00 sec)
mysql> SELECT COUNT(*) FROM PriceRange;
| COUNT (*) |
     1000 |
1 row in set (0.00 sec)
mysql>
```

```
SELECT COUNT(*) FROM GameData;

SELECT COUNT(*) FROM GenreData;

SELECT COUNT(*) FROM Region;

SELECT COUNT(*) FROM PriceRange;
```

ADVANCED QUERY 1:

```
SELECT DISTINCT Price, COUNT(GameName)

FROM GameData NATURAL JOIN GenreData

WHERE Price < 10 AND (Multiplayer = 'TRUE' OR Action = 'TRUE')

GROUP BY Price

UNION

SELECT DISTINCT Price, COUNT(GameName)

FROM GameData NATURAL JOIN GenreData

WHERE Price > 30 AND SinglePlayer = 'TRUE'

GROUP BY Price

ORDER BY Price DESC

LIMIT 15;
```

```
### STREET DISTINCT Frice, COUNT (GameName) FROM GameData MATURAL JOIN GenreData MIREE Frice < 10 AMD (Multiplayer = 'TROE') GROUP BY Frice ONION SELECT DISTINCT Frice, COUNT (GameName) FROM GameData MATURAL JOIN GenreData Miree Frice < 10 AMD (Multiplayer = 'TROE') GROUP BY Frice ONION SELECT DISTINCT Frice, COUNT (GameName) FROM GameData MATURAL JOIN GenreData Ma
```

++		+					
Price	COUNT (GameName)	1					
++		+					
449.99	1	1					
234.99	1	1					
149.99	1	I					
99.99	7	I					
79.99	4	Ι					
64.99	1	I					
59.99	44	Ι					
54.99	4	I					
49.99	53	I					
47.99	1	I					
44.99	12	I					
42.49	1	T					
40.49	1	I					
39.99	114	I					
35.99	4	T					
++							
15 rows in	set (0.06 sec)						
<u> </u>							

ADVANCED QUERY 2:

SELECT Rating, RequiredAge, COUNT(GameName)

```
FROM GameData

WHERE GameName IN (SELECT GameName

FROM GameData NATURAL JOIN GenreData

WHERE Multiplayer = 'TRUE' AND Action = 'TRUE')

GROUP BY Rating, RequiredAge

ORDER BY Rating DESC, RequiredAge DESC

LIMIT 15;
```

+		+		+-	+
Ra	ating	Reg	uiredAge	Т	COUNT (GameName)
+		+		+-	+
1	96	1	17	1	1
1	96	1	0	Т	1
I	94	1	0	Т	1
I	93	1	17	Т	2
1	92	1	17	Т	2
1	92	1	0	Т	1
I	91	I .	18	1	1
I	91	1	0	Т	6
1	90	1	17	Т	2
I	90	1	0	Т	1
I	89	1	17	Т	1
I .	89	1	0	Т	4
I	88	I _	17	I	3
I	88	I	0	I	6
I	87	I	18	I	1
+		+		+-	+
15	rows i	n set	(0.04 s∈	c)	

INDEXING:

Advanced Query 1:

Original:

```
|-> Limit: 15 row(s) (cost=2.50 rows=0) (actual time=0.030..0.031 rows=15 loops=1)
-> Sort: Price DESC, limit input to 15 row(s) per chunk (cost=2.50 rows=0) (actual time=0.029..0.030 rows=15 loops=1)
-> Sort: Price DESC, limit input to 15 row(s) per chunk (cost=2.50 rows=0) (actual time=0.029..0.030 rows=15 loops=1)
-> Table scan on (union temporary) (cost=2.50 rows=0) (actual time=0.001.0.005 rows=99 loops=1)
-> Table scan on (temporary) (actual time=0.01.0.005 rows=00 loops=1)
-> Aggregate using temporary table (actual time=31.509..31.517 rows=80 loops=1)
-> Aggregate using temporary table (actual time=31.509..31.517 rows=80 loops=1)
-> Table scan on (sembata.Gameble GenreData.Gamell) (cost=46587.25 rows=31103) (actual time=13.217..29.979 rows=5241 loops=1)
-> Table scan (acust=0.76 rows=366) (actual time=0.025..11.661 rows=10305 loops=1)
-> Table scan on GenreData (cost=0.76 rows=1030) (actual time=0.025..10.466 rows=13358 loops=1)
-> Table scan on GenreData (cost=0.76 rows=13360) (actual time=0.027..8.258 rows=13358 loops=1)
-> Table scan on GenreData (cost=0.76 rows=1366) (actual time=0.027..8.258 rows=13358 loops=1)
-> Aggregate using temporary table (actual time=24.337..24.339 rows=19 loops=1)
-> Table scan on GenreData.GameDl (cost=0.97 rows=1366) (actual time=0.150..10.729 rows=396 loops=1)
-> Table scan on GenreData (cost=0.97 rows=1366) (actual time=0.029..10.362 rows=11687 loops=1)
-> Table scan on GenreData (cost=0.97 rows=13360) (actual time=0.029..10.362 rows=11687 loops=1)
-> Table scan on GenreData (cost=0.97 rows=13360) (actual time=0.00.10..10.29 rows=3358 loops=1)
-> Stah
-> Filter: (GenreData.Singleplayer = 'TRUE') (cost=1376.25 rows=13360) (actual time=0.029..10.362 rows=11687 loops=1)
-> Table scan on GenreData (cost=0.97 rows=13360) (actual time=0.00.18..7.991 rows=13358 loops=1)
-> Table scan on GenreData (cost=0.97 rows=13360) (actual time=0.00.18..7.991 rows=13358 loops=1)
-> Table scan on GenreData (cost=0.97 rows=13360) (actual time=0.018..7.991 rows=13358 loops=1)
```

Index 1:

```
mysql> CREATE INDEX Singleplayer_idx on GenreData(Singleplayer);
Query OK, 0 rows affected (0.15 sec)
Records: 0 Duplicates: 0 Warnings: 0
```

Results:

```
| -> Limit: 15 row(s) (cost=2.50 rows=0) (actual time=0.029..0.031 rows=15 loops=1)
-> Sort: Price DESC, limit input to 15 row(s) per chunk (cost=2.50 rows=0) (actual time=0.028..0.029 rows=15 loops=1)
-> Table scan on comino temporaryy (cost=2.50 rows=0) (actual time=0.0028..0.029 rows=99 loops=1)
-> Table scan on comino temporaryy (cost=2.50 rows=0) (actual time=0.028..0.029 rows=99 loops=1)
-> Table scan on comporaryy (actual time=0.001..0.005 rows=00 loops=1)
-> Table scan on comporaryy (actual time=0.001..0.005 rows=00 loops=1)
-> Table scan on (Cost=0.50 rows=00 loops=1)
-> Table scan on (Cost=0.50 rows=00 loops=1)
-> Table scan on GameData.GameID (cost=0.76 rows=100) (actual time=0.028..12.238 rows=10305 loops=1)
-> Table scan on GameData (cost=0.76 rows=1030) (actual time=0.028..12.238 rows=10305 loops=1)
-> Table scan on GameData (cost=0.76 rows=1030) (actual time=0.026..11.035 rows=13358 loops=1)
-> Table scan on GameData (cost=0.76 rows=1030) (actual time=0.026..13.035 rows=1358 loops=1)
-> Table scan on GameData (cost=0.76 rows=1030) (actual time=0.025..8.443 rows=13358 loops=1)
-> Table scan on GameData (cost=0.77.096 rows=19 loops=1)
-> Table scan on GameData (cost=0.77.096 rows=19 loops=1)
-> Table scan on GameData (cost=0.34 rows=1030) (actual time=0.208..10.925 rows=396 loops=1)
-> Table scan on GameData (cost=0.34 rows=1030) (actual time=0.208..10.925 rows=396 loops=1)
-> Table scan on GameData (cost=0.34 rows=1030) (actual time=0.208..10.925 rows=396 loops=1)
-> Table scan on GameData (cost=0.34 rows=1030) (actual time=0.208..10.925 rows=396 loops=1)
-> Table scan on GameData (cost=0.34 rows=1030) (actual time=0.208..10.925 rows=396 loops=1)
-> Table scan on GameData (cost=0.34 rows=1030) (actual time=0.208..10.925 rows=396 loops=1)
-> Table scan on GameData (cost=0.34 rows=1030) (actual time=0.208..10.925 rows=396 loops=1)
-> Table scan on GameData (cost=0.34 rows=306) (actual time=0.208..10.925 rows=396 loops=1)
-> Table scan on GameData (cost=0.34 rows=306) (actual time=0.208..10.925 rows=
```

Explanation:

When we only index GenreData. Singleplayer, we can see that the cost decreases from 1376.25 to 788.75. This is because the data is stored in Main memory so it can be accessed a lot faster than without indexing Singleplayer.

Index 2:

```
mysql> CREATE INDEX Price ON GameData(price);
```

Results:

Explanation:

When we only index Price, we can see that it affects both parts of the union query. For the query before the Union:

We can see that the cost relating to GameData.Price decreases from 0.76 to 0.4. This could be because the indexing of the Price would help access the data faster since it is stored in the main memory.

For the guery after the Union:

We can see that the cost relating to GameData.Price increased from 0.97 to 178.46. However, the cost relating to GenreData.SinglePlayer decreased from 1376.25 to 8.08. These two variables are linked via this SQL statement: WHERE Price > 30 AND SinglePlayer = `TRUE'. Therefore, the changes in the costs could be due to the Price condition being executed first after indexing it, so this condition would filter out more rows than without the indexing. Thus, for the SinglePlayer condition, there are less rows for it to filter out and that's why the cost lowers for SinglePlayer instead.

Index 3:

```
mysql> CREATE INDEX Singleplayer_idx on GenreData(Singleplayer);
ERROR 1061 (42000): Duplicate key name 'Singleplayer_idx'
mysql> CREATE INDEX Price_idx on GameData(Price);
Query OK, 0 rows affected (0.12 sec)
Records: 0 Duplicates: 0 Warnings: 0
```

Results:

```
|-> Limit: 15 row(s) (cost=2.50 rows=0) (actual time=0.039.0.041 rows=15 loops=1)
|-> Sort: Price DESC, limit input to 15 row(s) per chunk (cost=2.50 rows=0) (actual time=0.038.0.039 rows=15 loops=1)
|-> Table scan on (union temporaryy (cost=2.50 rows=0) (actual time=0.000.0.005 rows=99 loops=1)
|-> Union materialize with deduplication (cost=2.50.2.50 rows=0) (actual time=44.702.44.705 rows=99 loops=1)
|-> Table scan on comporaryy (cost=2.50 rows=0) (actual time=0.003.0.005 rows=0) loops=1)
|-> Aggregate using temporary table (actual time=0.01.0.005 rows=0 loops=1)
|-> Aggregate using temporary table (actual time=0.01.0.005 rows=00 loops=1)
|-> Table scan on comporary (cost=0.63 rows=10.005)
|-> Table scan on commontate (cost=0.63 rows=10.005)
|-> Table scan on Commontate (cost=0.63 rows=10.005)
|-> Table scan on Commontate (cost=0.136.25 rows=10.006)
|-> Table scan on Comporary (actual time=0.026.0.1.0.290 rows=13360) (actual time=0.026.0.1.0.005 rows=13360)
|-> Table scan on Comporary (actual time=0.10.0.002 rows=1306)
|-> Table scan on Comporary (actual time=0.10.0.002 rows=1306)
|-> Table scan on Commontate (cost=1376.25 rows=1306) (actual time=0.020.9.0.306 rows=13358 loops=1)
|-> Aggregate using temporary table (actual time=0.11.0.002 rows=1300)
|-> Table scan on Commontate (cost=1376.25 rows=160) (cost=244.702 rows=1326) (actual time=1.250..12.403 rows=288 loops=1)
|-> Table scan on Commontate (actual time=1.010 (cost=244.702 rows=1326) (actual time=0.022..1.0.101 rows=11687 loops=1)
|-> Table scan on Commontate (actual time=0.012..7.700 rows=1380 loops=1)
|-> Table scan on Commontate (actual time=0.012..7.700 rows=1380 loops=1)
|-> Table scan on Commontate (actual time=0.012..7.700 rows=1380 loops=1)
|-> Table scan on Commontate (actual time=0.012..7.700 rows=1380 loops=1)
|-> Table scan on Commontate (actual time=0.012..7.700 rows=1380 loops=1)
|-> Table scan on Commontate (actual time=0.012..7.700 rows=1380 loops=1)
|-> Table scan on Commontate (actual time=0.012..7.700 rows=1380 loops=1)
|-> Table sc
```

Explanation:

When we index both Singleplayer and Price, we can see that the effects from index 2 still holds (explained above in the index 2 section). However, the key difference is that the Singleplayer cost decreased even more than the cost seen in index 2. For reference, with no indexing, the cost was 1376.25. With the conditions listed in index 2, the cost was 8.08. Yet with the conditions listed in index 3, the cost was 4.76. This is because compared to index 2, we have also indexed SinglePlayer so it will be a lot faster to access the data.

Advanced Query 2:

Original:

```
mysql> EXPLAIN ANALYZE SELECT Rating, RequiredAge, COUNT(GameName)
   -> FROM GameData
   -> WHERE GameName IN (SELECT GameName
   -> FROM GameData NATURAL JOIN GenreData
   -> WHERE Multiplayer = 'TRUE' AND Action = 'TRUE')
   -> GROUP BY Rating, RequiredAge
   -> ORDER BY Rating DESC, RequiredAge DESC
   -> LIMIT 15;
```

Index 1:

```
mysql> CREATE INDEX Multiplayer_idx on GenreData(Multiplayer);
Query OK, 0 rows affected, 1 warning (0.13 sec)
Records: 0 Duplicates: 0 Warnings: 1
```

```
| -> Limit: 15 row(s) (actual time=42.397..42.399 rows=15 loops=1)
| -> Sort: GameData.Rating DESC, GameData.Rating DESC, limit input to 15 row(s) per chunk (actual time=42.396..42.397 rows=15 loops=1)
| -> Sort: GameData.Rating DESC, GameData.Rating DESC, limit input to 15 row(s) per chunk (actual time=42.396..42.397 rows=15 loops=1)
| -> Angregate using temporary table (actual time=42.334..42.348 rows=130 loops=1)
| -> Nested loop inner join (cost=7365195731.5 rows=7365107024) (actual time=20.585..41.427 rows=2109 loops=1)
| -> Filter: (GameData.GameName is not null) (cost=1818.73 rows=110.30 (actual time=0.55.11.57) rows=1338 loops=1)
| -> Filter: (GameData.GameName is not null) (cost=1818.73 rows=1388 loops=1)
| -> Single=row index lookup on cubuquery2 using saute distinct keys (actual time=0.000..0.000 rows=0 loops=13358)
| -> Materialize with deduplication (cost=1818.448.62..78148.62 rows=1618873) (actual time=2.089..28.277 rows=2091 loops=1)
| -> Filter: (GameData.GameName is not null) (cost=161856.130 rows=1618873) (actual time=7.655..19.135 rows=2130 loops=1)
| -> Table scan on GameData (cost=0.36 rows=1030) (cost=161867.3) rows=1038 loops=1)
| -> Table scan on GameData (cost=0.36 rows=1030) (cost=267.52 rows=1468) (actual time=0.027..6.435 rows=2106 loops=1)
| -> Filter: (GenreData.Nation' = "TRUE") (cost=267.52 rows=3481) (actual time=0.024..5.800 rows=3481 loops=1)
| -> Index lookup on GenreData using Multi (Multiplayer="TRUE") (cost=267.52 rows=3481) (actual time=0.024..5.800 rows=3481 loops=1)
```

Explanation:

When we solely index GenreData.Multiplayer, we can see that the actualtime decreases from 0.027.. to 0.024. This is because the data is stored in Main memory so it can be accessed a lot faster than without indexing Multiplayer.

Index 2:

```
mysql>
mysql> Create Index Rating on GameData(rating)
   ->;
Query OK, 0 rows affected (0.10 sec)
Records: 0 Duplicates: 0 Warnings: 0
```

Explanation:

Indexing by GameData.Rating has improved the performance of the query by reducing the amount of data that needs to be processed, limiting the number of rows that need to be scanned, and improving the efficiency of the joins. Index lookup time on action has decreased from 0.022 to 0.020.

Index 3:

```
| -> Limit: 15 row(s) (actual time=49.361..49.363 rows=15 loops=1)
-> Sort: GameData. Rating DESC, GameData. RequiredAge DESC, limit input to 15 row(s) per chunk (actual time=49.360..49.361 rows=15 loops=1)
-> Nagregate using temporary (actual time=0.001..0.008 rows=130 loops=1)
-> Nested loop inner join (cost=1785619579.15 rows=17856170924) (actual time=27.177..48.430 rows=2109 loops=1)
-> Filter: (GameData. GameName is not null) (cost=1383.75 rows=11030) (actual time=0.028..10.594 rows=13358 loops=1)
-> Single-row index lookup on <subquery2> using dauto_distinct key> (GameName=GameData. GameName) (actual time=0.000..0.000 rows=0 loops=13358)
-> Materialize with deduplication (cost=1781448.62..1781448.62 rows=1618873) (actual time=9.423..25.577 rows=2130 loops=1)
-> Filter: (GameData.GameName is not null) (cost=1619561.30 rows=1618873) (actual time=9.423..25.577 rows=2130 loops=1)
-> Table scan on GameData (cost=0.36 rows=1030) (actual time=0.017..14.212 rows=13558 loops=1)
-> Table scan on GameData (cost=0.36 rows=1030) (actual time=0.017..14.212 rows=13558 loops=1)
-> Hash
-> Filter: (GenreData.'Action' = 'TRUE') (cost=267.52 rows=1468) (actual time=0.037..8.746 rows=2106 loops=1)
-> Index lookup on GenreData using Multi (Multiplayer="TRUE') (cost=267.52 rows=3481) (actual time=0.034..8.068 rows=3481 loops=1)
```

Explanation:

Indexing by GenreData.Action, has enabled us to reduce the actual time of the table scan on GameData from 0.024 to 0.017. However it does have some downsides, unlike the previous queries it does not decrease the lookup time on Multiplayer, yet increase it.