DDL Commands

```
CREATE TABLE Users(
      NetID VARCHAR(15),
      Username VARCHAR(200),
      Password VARCHAR(200),
      IsGuest BOOLEAN.
      PRIMARY KEY(Username));
CREATE TABLE Instructors(
      NetID VARCHAR(15),
      Name VARCHAR(100),
      DeptAbv VARCHAR(200),
      PRIMARY KEY (NetID),
      FOREIGN KEY (DeptAbv) REFERENCES Departments(DeptAbv));
CREATE TABLE Departments(
      DeptAby VARCHAR(10),
      DeptName VARCHAR(200),
      Location VARCHAR(200),
      URL VARCHAR(200),
      PhoneNumber VARCHAR(20),
      PRIMARY KEY (DeptAbv));
CREATE TABLE Courses(
      CRN VARCHAR(10),
      CourseName VARCHAR(200),
      CourseNumber VARCHAR(10),
      Description VARCHAR(500),
      DeptAbv VARCHAR(200),
      PRIMARY KEY(CRN),
      FOREIGN KEY(DeptAbv) REFERENCES Departments(DeptAbv));
CREATE TABLE Reviews(
      ReviewID VARCHAR(200),
      Rating INT,
      Comment VARCHAR(500),
```

IsRecommended BOOLEAN, RequiresTextbook BOOLEAN, Upvotes INT, Downvotes INT, Username VARCHAR(200), CRN VARCHAR(10), InstructorNetID VARCHAR(15),

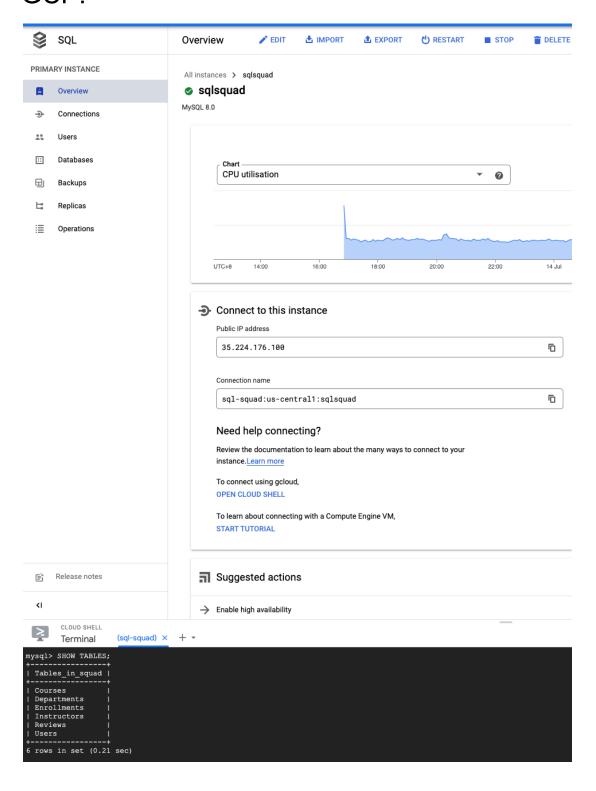
PRIMARY KEY(ReviewID), FOREIGN KEY(Username) REFERENCES Users(Username), FOREIGN KEY(InstructorNetID) REFERENCES Instructors(NetID), FOREIGN KEY(CRN) REFERENCES Courses(CRN));

CREATE TABLE Enrollments(

Username VARCHAR(200), CRN VARCHAR(10), InstructorNetID VARCHAR(15), YearTerm VARCHAR(10), GPA REAL,

PRIMARY KEY(Username, CRN, YearTerm),
FOREIGN KEY(Username) REFERENCES Users(Username),
FOREIGN KEY(CRN) REFERENCES Courses(CRN),
FOREIGN KEY(InstructorNetID) REFERENCES Instructors(NetID));

Proof of Implementation / Terminal Connection on GCP:



Proof of at least 1000 entries in 3 tables:

```
mysql> SELECT COUNT(*) FROM Courses;
+-----+
| COUNT(*) |
+-----+
| 12212 |
+-----+
1 row in set (0.24 sec)

mysql> SELECT COUNT(*) FROM Reviews;
+-----+
| COUNT(*) |
+-----+
| 1000 |
+-----+
1 row in set (0.22 sec)

mysql> SELECT COUNT(*) FROM Instructors;
+------+
| COUNT(*) |
+-------+
| Tow in set (0.21 sec)
```

2 Different Advanced SQL Queries:

Query 1: Search for Course: Average Rating, and relevant Course Details

purpose: this allows users to see the general rating and relevant details for courses, specific to the CRN, that they may want to take.

SELECT AVG(r.Rating), c.Description, c.CourseName, c.CourseNumber, c.DeptAbv FROM Reviews r LEFT JOIN Courses c ON r.CRN = c.CRN LEFT JOIN Enrollments e ON c.CRN = e.CRN GROUP BY c.CRN LIMIT 15;

```
EMERIC AND ADDRESS OF COM = C.OM = C.OM LET JOHN Extraces to C.OM = C.OM LET JOHN Extraces to C.OM = C.OM =
```

Query 2: Instructor with the Highest Rating grouped by DeptAbv and CourseNumber

Purpose: This query allows users to filter for instructors that have the highest ratings for the courses that they might potentially want to take, in order to make a better choice when picking instructors. This is specific to the Department Abbreviation, for example 'CS', and the Course Number, so it finds the professor or professors with the highest average rating for 'CS100' for example.

SELECT DISTINCT i.Name, s2.DeptAbv, s2.CourseNumber, s2.maxRating
FROM (SELECT MAX(s1.avgRating) as maxRating, s1.DeptAbv, s1.CourseNumber
FROM (SELECT AVG(r.Rating) as avgRating, c.DeptAbv, c.CourseNumber, r.InstructorNetID
FROM Reviews r LEFT JOIN Courses c ON r.CRN = c.CRN GROUP BY c.DeptAbv,
c.CourseNumber, r.InstructorNetID) AS s1 GROUP BY s1.DeptAbv, s1.CourseNumber) AS s2
LEFT JOIN (SELECT AVG(r2.Rating) as avgRating, c2.DeptAbv, c2.CourseNumber,
r2.InstructorNetID FROM Reviews r2 LEFT JOIN Courses c2 ON r2.CRN = c2.CRN GROUP BY
c2.DeptAbv, c2.CourseNumber, r2.InstructorNetID) AS s3 ON s2.maxRating = s3.avgRating
AND s2.DeptAbv = s3.DeptAbv AND s2.CourseNumber = s3.CourseNumber LEFT JOIN
Instructors i ON s3.InstructorNetID = i.NetID
LIMIT 15;

Indexing Analysis

Query 1

Query Details:

```
SELECT AVG(r.Rating), c.Description, c.CourseName, c.CourseNumber,
c.DeptAbv
FROM Reviews r LEFT JOIN Courses c ON r.CRN = c.CRN LEFT JOIN Enrollments e
ON c.CRN = e.CRN
GROUP BY c.CRN
LIMIT 15;
```

EXPLAIN ANALYZE results:

| EXPLAIN

- | -> Limit: 15 row(s) (actual time=8.779..8.783 rows=15 loops=1)
 - -> Table scan on <temporary> (actual time=0.002..0.005 rows=15 loops=1)
 - -> Aggregate using temporary table (actual time=8.777..8.781 rows=15 loops=1)
- -> Nested loop left join (cost=801.75 rows=1000) (actual time=0.071..5.774 rows=1177 loops=1)
- -> Nested loop left join (cost=451.75 rows=1000) (actual time=0.058..2.536 rows=1000 loops=1)
- -> Table scan on r (cost=101.75 rows=1000) (actual time=0.040..0.355 rows=1000 loops=1)
- -> Single-row index lookup on c using PRIMARY (CRN=r.CRN) (cost=0.25 rows=1) (actual time=0.002..0.002 rows=1 loops=1000)
- -> Index lookup on e using CRN (CRN=c.CRN) (cost=0.25 rows=1) (actual time=0.002..0.003 rows=1 loops=1000)

```
1 row in set (0.21 sec)
3 Different Indexing Designs:
Indexing on c.DeptAbv results:
I EXPLAIN
| -> Limit: 15 row(s) (actual time=9.361..9.366 rows=15 loops=1)
  -> Table scan on <temporary> (actual time=0.002..0.005 rows=15 loops=1)
    -> Aggregate using temporary table (actual time=9.360..9.364 rows=15 loops=1)
       -> Nested loop left join (cost=825.37 rows=1185) (actual time=0.071..6.242 rows=1177
loops=1)
         -> Nested loop left join (cost=451.75 rows=1000) (actual time=0.061..2.926
rows=1000 loops=1)
            -> Table scan on r (cost=101.75 rows=1000) (actual time=0.044..0.375 rows=1000
loops=1)
            -> Single-row index lookup on c using PRIMARY (CRN=r.CRN) (cost=0.25 rows=1)
(actual time=0.002..0.002 rows=1 loops=1000)
         -> Index lookup on e using CRN (CRN=c.CRN) (cost=0.26 rows=1) (actual
time=0.002..0.003 rows=1 loops=1000)
1 row in set (0.22 sec) (this is poorer in performance compared to the raw one, 0.22 > 0.21)
Indexing on c.CourseNumber:
| EXPLAIN
| -> Limit: 15 row(s) (actual time=9.357..9.362 rows=15 loops=1)
  -> Table scan on <temporary> (actual time=0.002..0.005 rows=15 loops=1)
    -> Aggregate using temporary table (actual time=9.357..9.360 rows=15 loops=1)
       -> Nested loop left join (cost=825.37 rows=1185) (actual time=0.105..6.218 rows=1177
loops=1)
         -> Nested loop left join (cost=451.75 rows=1000) (actual time=0.075..3.035
rows=1000 loops=1)
            -> Table scan on r (cost=101.75 rows=1000) (actual time=0.053..0.419 rows=1000
loops=1)
            -> Single-row index lookup on c using PRIMARY (CRN=r.CRN) (cost=0.25 rows=1)
(actual time=0.002..0.002 rows=1 loops=1000)
         -> Index lookup on e using CRN (CRN=c.CRN) (cost=0.26 rows=1) (actual
time=0.002..0.003 rows=1 loops=1000)
1 row in set (0.22 sec) (this is poorer in performance compared to the raw one, 0.22 > 0.21)
```

Indexing on r.Rating:

| EXPLAIN

```
| -> Limit: 15 row(s) (actual time=8.617..8.621 rows=15 loops=1)
-> Table scan on <temporary> (actual time=0.002..0.004 rows=15 loops=1)
-> Aggregate using temporary table (actual time=8.616..8.619 rows=15 loops=1)
-> Nested loop left join (cost=825.37 rows=1185) (actual time=0.060..5.598 rows=1177 loops=1)
-> Nested loop left join (cost=451.75 rows=1000) (actual time=0.051..2.767 rows=1000 loops=1)
-> Table scan on r (cost=101.75 rows=1000) (actual time=0.034..0.370 rows=1000 loops=1)
-> Single-row index lookup on c using PRIMARY (CRN=r.CRN) (cost=0.25 rows=1) (actual time=0.002..0.002 rows=1 loops=1000)
-> Index lookup on e using CRN (CRN=c.CRN) (cost=0.26 rows=1) (actual time=0.002..0.003 rows=1 loops=1000)
| 1 row in set (0.22 sec)
```

Conclusion

It is unlikely that indexing will help here since we are finding the average rating, hence the need to go through each and every row. So indexes do not speed up this query at all.

Query 2

Query Details:

SELECT DISTINCT i.Name, s2.DeptAbv, s2.CourseNumber, s2.maxRating
FROM (SELECT MAX(s1.avgRating) as maxRating, s1.DeptAbv, s1.CourseNumber
FROM (SELECT AVG(r.Rating) as avgRating, c.DeptAbv, c.CourseNumber, r.InstructorNetID
FROM Reviews r LEFT JOIN Courses c ON r.CRN = c.CRN GROUP BY c.DeptAbv,
c.CourseNumber, r.InstructorNetID) AS s1 GROUP BY s1.DeptAbv, s1.CourseNumber) AS s2
LEFT JOIN (SELECT AVG(r2.Rating) as avgRating, c2.DeptAbv, c2.CourseNumber,
r2.InstructorNetID FROM Reviews r2 LEFT JOIN Courses c2 ON r2.CRN = c2.CRN GROUP BY
c2.DeptAbv, c2.CourseNumber, r2.InstructorNetID) AS s3 ON s2.maxRating = s3.avgRating
AND s2.DeptAbv = s3.DeptAbv AND s2.CourseNumber = s3.CourseNumber LEFT JOIN
Instructors i ON s3.InstructorNetID = i.NetID
LIMIT 15:

EXPLAIN ANALYZE Results:

No Indexes:

```
I EXPLAIN
| -> Limit: 15 row(s) (cost=5346.01..5346.01 rows=0) (actual time=11.744..11.747 rows=15
loops=1)
  -> Table scan on <temporary> (cost=2.50..2.50 rows=0) (actual time=0.000..0.002 rows=15
loops=1)
    -> Temporary table with deduplication (cost=5348.51..5348.51 rows=0) (actual
time=11.742..11.744 rows=15 loops=1)
       -> Limit table size: 15 unique row(s)
         -> Nested loop left join (cost=5346.01 rows=0) (actual time=11.645..11.709 rows=15
loops=1)
            -> Nested loop left join (cost=2674.26 rows=0) (actual time=11.631..11.657
rows=15 loops=1)
              -> Table scan on s2 (cost=2.50..2.50 rows=0) (actual time=0.001..0.003
rows=13 loops=1)
                -> Materialize (cost=2.50..2.50 rows=0) (actual time=6.001..6.005 rows=383
loops=1)
                   -> Table scan on <temporary> (actual time=0.000..0.026 rows=383
loops=1)
                     -> Aggregate using temporary table (actual time=5.851..5.899 rows=383
loops=1)
                       -> Table scan on s1 (cost=2.50..2.50 rows=0) (actual
time=0.001..0.043 rows=575 loops=1)
                          -> Materialize (cost=2.50..2.50 rows=0) (actual time=5.190..5.275
rows=575 loops=1)
                            -> Table scan on <temporary> (actual time=0.001..0.052
rows=575 loops=1)
                               -> Aggregate using temporary table (actual time=4.804..4.890
rows=575 loops=1)
                                 -> Nested loop left join (cost=451.75 rows=1000) (actual
time=0.079..3.261 rows=1000 loops=1)
                                   -> Table scan on r (cost=101.75 rows=1000) (actual
time=0.057..0.426 rows=1000 loops=1)
                                    -> Single-row index lookup on c using PRIMARY
(CRN=r.CRN) (cost=0.25 rows=1) (actual time=0.003..0.003 rows=1 loops=1000)
              -> Index lookup on s3 using <auto key1> (avgRating=s2.maxRating,
DeptAbv=s2.DeptAbv, CourseNumber=s2.CourseNumber) (actual time=0.001..0.002 rows=1
loops=13)
                -> Materialize (cost=0.00..0.00 rows=0) (actual time=5.644..5.649 rows=575
loops=1)
                  -> Table scan on <temporary> (actual time=0.001..0.086 rows=575
loops=1)
                     -> Aggregate using temporary table (actual time=4.627..4.763 rows=575
loops=1)
```

```
-> Nested loop left join (cost=451.75 rows=1000) (actual time=0.068..2.726 rows=1000 loops=1)
-> Table scan on r2 (cost=101.75 rows=1000) (actual time=0.058..0.431 rows=1000 loops=1)
-> Single-row index lookup on c2 using PRIMARY (CRN=r2.CRN) (cost=0.25 rows=1) (actual time=0.002..0.002 rows=1 loops=1000)
-> Single-row index lookup on i using PRIMARY (NetID=s3.InstructorNetID) (cost=0.25 rows=1) (actual time=0.003..0.003 rows=1 loops=15)
| 1 row in set (0.22 sec)

3 Different Indexing Designs:
Indexing on Reviews(Rating):
| EXPLAIN
| -> Limit: 15 row(s) (cost=5346.01..5346.01 rows=0) (actual time=10.007..10.010 rows=15 loops=1)
```

-> Table scan on <temporary> (cost=2.50..2.50 rows=0) (actual time=0.000..0.002 rows=15

-> Nested loop left join (cost=5346.01 rows=0) (actual time=9.921..9.975 rows=15

-> Table scan on s2 (cost=2.50..2.50 rows=0) (actual time=0.000..0.002

-> Nested loop left join (cost=2674.26 rows=0) (actual time=9.910..9.937 rows=15

-> Materialize (cost=2.50..2.50 rows=0) (actual time=5.338..5.341 rows=383

-> Aggregate using temporary table (actual time=5.189..5.236 rows=383

-> Materialize (cost=2.50..2.50 rows=0) (actual time=4.563..4.641

-> Aggregate using temporary table (actual time=4.158..4.241

-> Table scan on <temporary> (actual time=0.001..0.051

-> Table scan on <temporary> (actual time=0.000..0.026 rows=383

-> Table scan on s1 (cost=2.50..2.50 rows=0) (actual

-> Temporary table with deduplication (cost=5348.51..5348.51 rows=0) (actual

loops=1)

loops=1)

loops=1)

loops=1)

loops=1)

loops=1)

rows=13 loops=1)

rows=575 loops=1)

rows=575 loops=1)

rows=575 loops=1)

time=10.006..10.008 rows=15 loops=1)

time=0.001..0.046 rows=575 loops=1)

-> Limit table size: 15 unique row(s)

```
-> Nested loop left join (cost=451.75 rows=1000) (actual
time=0.076..2.654 rows=1000 loops=1)
                                   -> Table scan on r (cost=101.75 rows=1000) (actual
time=0.054..0.404 rows=1000 loops=1)
                                   -> Single-row index lookup on c using PRIMARY
(CRN=r.CRN) (cost=0.25 rows=1) (actual time=0.002..0.002 rows=1 loops=1000)
              -> Index lookup on s3 using <auto key1> (avgRating=s2.maxRating,
DeptAbv=s2.DeptAbv, CourseNumber=s2.CourseNumber) (actual time=0.002..0.002 rows=1
loops=13)
                -> Materialize (cost=0.00..0.00 rows=0) (actual time=4.587..4.592 rows=575
loops=1)
                  -> Table scan on <temporary> (actual time=0.001..0.045 rows=575
loops=1)
                     -> Aggregate using temporary table (actual time=3.747..3.824 rows=575
loops=1)
                       -> Nested loop left join (cost=451.75 rows=1000) (actual
time=0.045..2.279 rows=1000 loops=1)
                          -> Table scan on r2 (cost=101.75 rows=1000) (actual
time=0.034..0.373 rows=1000 loops=1)
                          -> Single-row index lookup on c2 using PRIMARY (CRN=r2.CRN)
(cost=0.25 rows=1) (actual time=0.002..0.002 rows=1 loops=1000)
           -> Single-row index lookup on i using PRIMARY (NetID=s3.InstructorNetID)
(cost=0.25 rows=1) (actual time=0.002..0.002 rows=1 loops=15)
1 row in set (0.22 sec)
```

Indexing on Instructors(NetID)

Conclusion: