

DATABASE DESIGN

I. DDL Commands :

1. Main_table – has bulk csv data

```
CREATE TABLE main_data(  
Year INT,  
Term VARCHAR(10),  
YearTerm VARCHAR(10),  
Subject VARCHAR(10),  
Number INT,  
CourseTitle VARCHAR(200),  
SchedType VARCHAR(10),  
gA INT,  
gAplus INT,  
gAmin INT,  
gB INT,  
gBplus INT,  
gBmin INT,  
gC INT,  
gCplus INT,  
gCmin INT,  
gD INT,  
gDplus INT,  
gDmin INT,  
gF INT,  
gW INT,  
Instructor VARCHAR(100)  
);
```

2. Department – has list of all departments

```
CREATE TABLE Department(  
  DeptId VARCHAR(10) PRIMARY KEY,  
  DeptName VARCHAR(30)  
);
```

3. Courses – list of all courses

```
CREATE TABLE Courses(  
  CourseId VARCHAR(20) PRIMARY KEY,  
  Number INT ,  
  Title VARCHAR(50),  
  DeptId VARCHAR(10),  
  FOREIGN KEY (DeptId) REFERENCES Department (DeptId)  
);
```

//DATA DOESN'T HAVE CRN SO MODIFIED

4. Instructors – list of instructors

```
CREATE TABLE Instructors(  
  Eid VARCHAR(20) PRIMARY KEY,  
  Name VARCHAR(70)  
);
```

// DON'T HAVE DEPT_ID FOR INSTRUCTORS YET AND NO EMAIL ID TOO

5. Course Offerings - list of course offerings

```
CREATE TABLE Course_Offerings(  
  Eid VARCHAR(20) REFERENCES Instructors(Eid),  
  CourseId VARCHAR(20) REFERENCES Courses(CRN),  
  Semester VARCHAR(20) ,
```

A1 INT PRIMARY KEY,
A2 INT,
A3 INT,
B1 INT,
B2 INT,
B3 INT,
C1 INT,
C2 INT,
C3 INT,
D1 INT,
D2 INT,
D3 INT,
F INT,
W INT,
PRIMARY KEY (Eid, CourseId, Semester,)
);

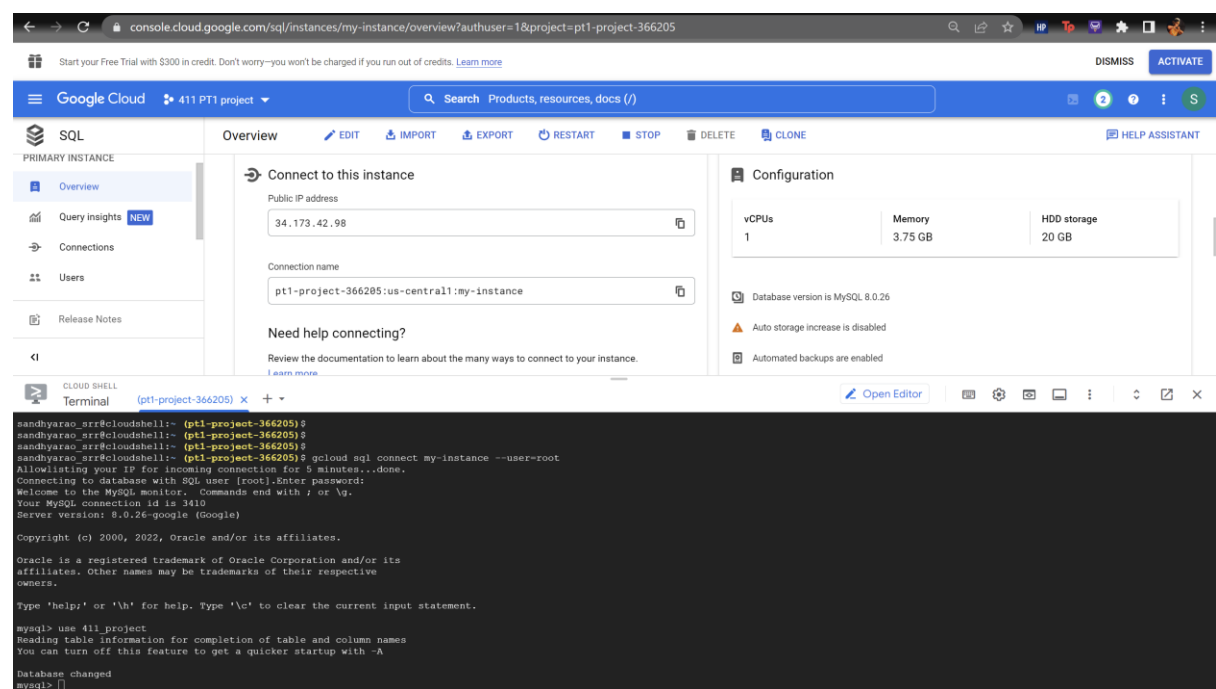
II. Count of Tables: (for tables with >1000 rows)

```
mysql> select count(*) from Courses;
+-----+
| count(*) |
+-----+
|      6005 |
+-----+
1 row in set (0.00 sec)
```

```
mysql> select count(*) from Instructors;
+-----+
| count(*) |
+-----+
|      8678 |
+-----+
1 row in set (0.00 sec)
```

```
mysql> select count(*) from main_data;
+-----+
| count(*) |
+-----+
|     61557 |
+-----+
1 row in set (0.02 sec)
```

III. GCP Connection:



Iv. ADVANCED SQL QUERIES

```
mysql> select Instructor,Eid from main_data m JOIN Instructors i ON (m.Instructor = i.Name) where Subject="CS" UNION (select Instructor,Eid from main_data m JOIN Instructors i ON (m.Instructor = i.Name) where Subject="IE") LIMIT 15;
```

```
+-----+-----+
| Instructor | Eid |
+-----+-----+
|           | e1  |
| Caccamo, Marco | e1008 |
| Caesar, Matthew C | e1014 |
| Campbell, R H | e1040 |
```

Alawini, Abdussalam A e106
Challen, Geoffrey W e1159
Chan, Timothy M e1173
Chang, Kevin C e1182
Chapman, William L e1192
Charalambides, Minas e1193
Chekuri, Chandra S e1210
Chen, Yi-Shuan e1245
Chowdhary, Girish e1342
Clifton, Rhiannon L e1424
Cook, Lucas T e1481

+-----+-----+

15 rows in set (0.04 sec)

2.