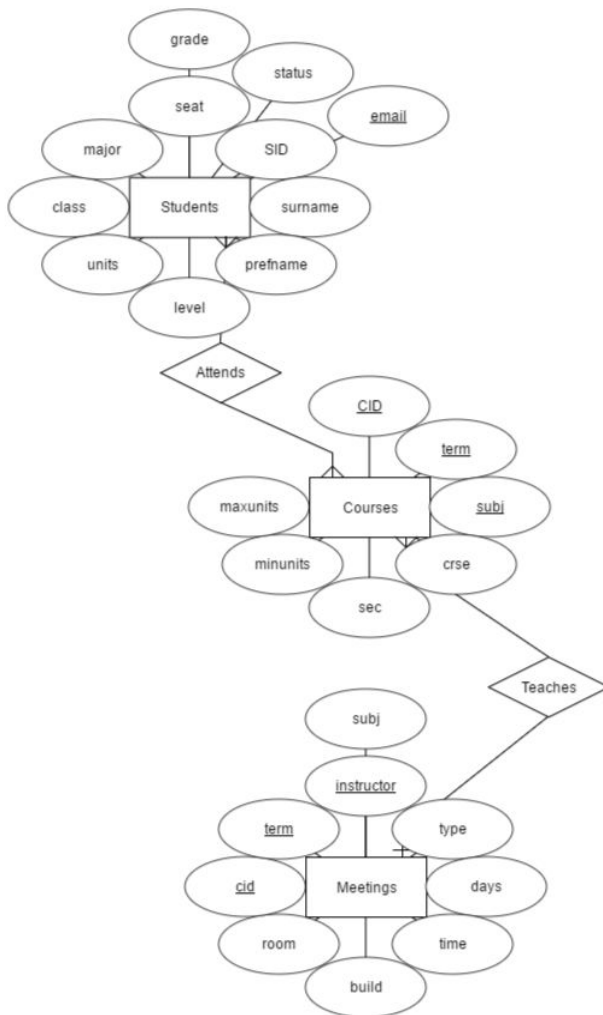Michael Casias 912408211
Jarred Stein 912355526
ECS 165A
Professor Nitta
12/3/2016

<p align="center">Homework 4 Report</p>

ER Diagram of the Tables

Explaining the tables

     Our database schema had 4 tables: Courses, Students, Meetings, and Attends. The first three tables came directly from the csv files, while Attends is used to map the relation of Students table to the Courses table.

The elements of the Courses table are: CID int, TERM int, SUBJ text, CRSE int, SEC int, MINUNITS float, MAXUNITS float. We parsed the UNITS column of the csvs into min and max to more easily deal with variable unit courses. The PRIMARY KEY is CID, TERM, SUBJ. This assumes that CIDs can be reused across terms, but not in the same one.

The elements of the Meetings table are: INSTRUCTOR text, TYPE text, DAYS text, TIME text, BUILD text, ROOM text, CID int, TERM int, and SUBJ text. The PRIMARY KEY is INSTRUCTOR, CID, TERM, SUB. This key allows for an instructor to teach multiple classes per quarter. This table also references the FOREIGN KEY CID, TERM, and SUBJ from the courses table. This is so we can keep track of what course each meeting is for.

The elements of the Students table are: SEAT int, SID bigint, SURNAME text, PREFNAME text, LEVEL text, UNITS text, CLASS text, MAJOR text, GRADE text, STATUS text, EMAIL text. The PRIMARY KEY is SID, assuming each student get's their own unique SID that will never be reused. We also assume that each student's EMAIL is going to be unique.

The elements of the Attends table are: SID int, CID int, TERM int, SUBJ int, TERM int, SUBJ text, GRADE float, and MAJOR text. The PRIMARY KEY is SID, CID, TERM, SUBJ. This of course assumes that a student can only take a class once per term, but they can take as many classes at once as they wish. The two FOREIGN KEYS are SID from the Students table and CID, TERM, SUBJ from the Courses table. This table gives a record of every course that every student took, the grade they got in the class, and what their major was when they took it.

Updating the tables

There are several approaches to updating the tables in the postgreSQL server; for updating tuples, one would only need to run a command in the main.cpp file with the send_query() command. Our tables contain the keywords 'ON UPDATE CASCADE' for the foreign keys to maintain referential integrity between the tables on an update. Additionally, to avoid any possible errors regarding duplicate information in the csv files, we included the keywords 'ON CONFLICT DO NOTHING'.

Functional/Multivalued Dependencies

STUDENTS:

Besides the trivial FD/ MVD (key determines the rest of the tuple), SID and email values in the student table is guaranteed to be unique.

COURSES:

The CID and term tuples would determine what class, section, and number of units are for that class during that term. Subject and CRSE would also determine the number of units for that class.

MEETINGS:

The instructor, type, and term columns would multivalue determine the rest of the information about that particular course (day and time and location the class or discussion meets as well as what subject the class is). If you remove type from the dependency, It is a MVD instead of a FD because an instructor can teach lectures and discussions, for instance.

ATTENDS:

Two types of dependencies we'd expect to hold for this relation is the following: SID,CID,term -> subj,grade,major and SID,CID ->> term, subj,grade,major.

Assumptions for Question 3 Queries

3a.

None.

3b.

None.

3c.

Assumption: We are only considering A+ to F letter grades when considering easiest/hardest professors, since those grades did not lend themselves to determining a GPA.

3d.

Assumption: Same as before, we disregard any grades that are not in the range A+ to F.

3e.

Assumption: we ignore classes with NULL times for day/room/building since we cannot confirm that there was conflict in the times.

3f.

Just as in c and d, we disregard grades that are not between A+ and F. We also do not account for the number of students in that major, which is why there are

polarizing results since there are few students in the major, so if there's only one student in it who receives an A, the average grade for that major is an A.

3g.
        Assumption: When calculating percent of students for top 5 majors, we used the number of students in the major, not the number of students who transferred.

Query Outputs

3a.
RESULTS FOR PROBLEM 3(a)
units: 1 percent of students: 0.918049
units: 2 percent of students: 0.018361
units: 3 percent of students: 0.455964
units: 4 percent of students: 41.6488
units: 5 percent of students: 1.68309
units: 6 percent of students: 0.159128
units: 7 percent of students: 0.257054
units: 8 percent of students: 11.8643
units: 9 percent of students: 1.26079
units: 10 percent of students: 0.211151
units: 11 percent of students: 0.220332
units: 12 percent of students: 4.18936
units: 13 percent of students: 0.820124
units: 14 percent of students: 0.220332
units: 15 percent of students: 0.208091
units: 16 percent of students: 2.52157
units: 17 percent of students: 0.407002
units: 18 percent of students: 0.137707
units: 19 percent of students: 0.18973
units: 20 percent of students: 1.45358

3b.
RESULTS FOR PROBLEM 3(b)
units: 1 average GPA: 3.40702
units: 2 average GPA: 3.56039
units: 3 average GPA: 3.25445
units: 4 average GPA: 2.69025
units: 5 average GPA: 2.8419
units: 6 average GPA: 3.12663
units: 7 average GPA: 2.83382

units: 8 average GPA: 2.75149
units: 9 average GPA: 2.723
units: 10 average GPA: 2.94275
units: 11 average GPA: 2.85248
units: 12 average GPA: 2.85367
units: 13 average GPA: 2.80787
units: 14 average GPA: 2.94179
units: 15 average GPA: 3.00915
units: 16 average GPA: 3.16603
units: 17 average GPA: 3.35388
units: 18 average GPA: 3.44481
units: 19 average GPA: 3.37161
units: 20 average GPA: 3.60227

3c.
RESULTS FOR PROBLEM 3(c)
Professors with the highest average grades
instructor: Russo, Angel J. average grade: 3.95
instructor: O'donnell, Madison G. average grade: 3.95
Professors with the lowest average grades
instructor: Turner, Emily A. average grade: 1.7

3d.
RESULTS FOR PROBLEM 3(d)
Easiest professors in the ABC 100 level courses
instructor: Dodson, Nicole M. average grade: 3.4628
Hardest professors in the ABC 100 level courses
instructor: Adams, Emily G. average grade: 1.78182

3e.
RESULTS FOR PROBLEM 3(e)
Courses with schedule conflicts
term: 200906 courses: ABC 104, ABC 107
term: 199303 courses: ABC 105, ABC 105
term: 199406 courses: ABC 105, ABC 107
term: 199506 courses: ABC 105, ABC 107
term: 199601 courses: ABC 108, ABC 108
term: 200906 courses: ABC 201, ABC 221
term: 200606 courses: DEF 201, DEF 258
term: 200606 courses: DEF 201, DEF 258
term: 200606 courses: DEF 201, DEF 258
term: 200606 courses: DEF 201, DEF 258
term: 200201 courses: DEF 252, DEF 252

term: 199506 courses: DEF 258, DEF 258

3f.
RESULTS FOR PROBLEM 3(f)
Majors with the best average GPA for ABC courses:
Major: O275 average GPA: 4
Major: O151 average GPA: 4
Major: O193 average GPA: 4
Major: O207 average GPA: 4
Major: O167 average GPA: 4
Major: O139 average GPA: 4
Major: O255 average GPA: 4
Major: O100 average GPA: 4
Major: O113 average GPA: 4
Major: O169 average GPA: 4
Major: O179 average GPA: 4
Major: O176 average GPA: 4
Majors with the worst average GPA for ABC courses:
Major: O281 average GPA: 0
Major: O152 average GPA: 0
Major: O263 average GPA: 0
Major: O279 average GPA: 0

3g.
RESULTS FOR PROBLEM 3(g)
Percent of students who transferred in to ABC major: 25.4602
Top 5 majors that transferred in to ABC major:
Major: DEF2 number of students: 216
Major: OT16 number of students: 197
Major: DEF1 number of students: 116
Major: OT26 number of students: 99
Major: OT35 number of students: 92
Percent of DEF2 major: 13.3251
Percent of OT16 major: 15.2124
Percent of DEF1 major: 4.42917
Percent of OT26 major: 8.28452
Percent of OT35 major: 8.34089