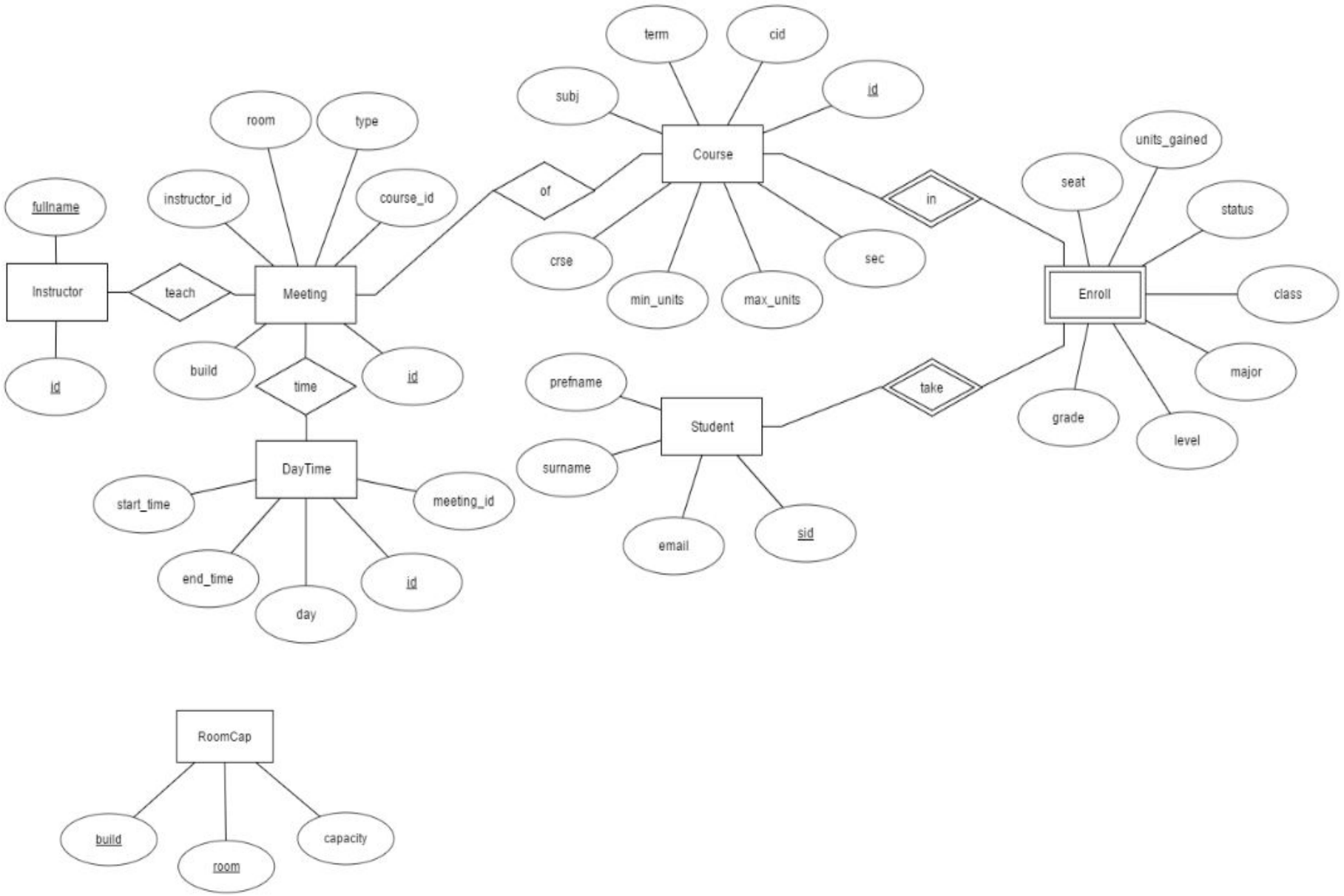
term
cid
subj
id
Course
room
type
instructor_id
course_id
of
in
seat
units_gained
status
fullname
Instructor
teach
Meeting
crse
min_units
max_units
sec
Enroll
class
build
time
id
major
prefname
take
grade
level
DayTime
Student
surname
start_time
meeting_id
email
sid
end_time
id
day
RoomCap
build
room
capacity

Database Schema Writeup

The 6 entities Instructor, Meeting, Daytime, Student, Course, and Enroll represent the csv data. For all of the tables, there are NOT NULL constraints for fields that either have foreign key constraints, unique constraints, or primary key constraints. Some tables like Meeting do not have a NOT NULL constraint on its Instructor foreign key because it might be variable.

The Instructor entity contains unique full names of instructors and a primary key of the ID. It is reasonable to think that the entity might get additional fields in the future. As of now the full_name unique constraint is satisfied, but if there are instructors with conflicting names, then the constraint should be removed. The ID serves as the primary key for this reason.

The Meeting entity has the ID as the primary key. It has course_id and instructor_id as foreign key constraints. It also has a one to many DayTime relationship. We separated the time from Meeting so that specific days could be easily queried. It is also possible that a Meeting contains no reverse relations to any DayTime instances--this would be a variable Meeting. The Meeting is connected many to one related to Instructor and Course.

The Course entity has the ID as the primary key. It has a unique constraint for term and cid together. We split the units into min_units and max_units to represent a range and update that range more easily. There is a one Course to many Enroll relationship.

The Enroll entity has sid, cid, and term as the primary key. It is a weak entity that gains its key from Course and Student. The entity itself serves to relate a Student to a Course and add state information about the student. The sid has a foreign key constraint with Student and the cid and term have a foreign key constraint with Course. There is a unique constraint around sid, cid, and term so that it can be used as a primary key.

The Student entity has sid as the primary key. The email has a unique constraint. Most of the information about a student is broken out into Enroll. If you wanted to determine the current major or level of a student you would have to check the most current Enroll row with a relation to the Student. For checking their total units, you would aggregate the units_gained with passing grades.

The separate entity RoomCap is built for the convenience of recording the maximum capacity of each room--required by query 5b. The build and room are the primary key and have a unique together constraint as a result. On entry of the csv data, capacity is null. When the data is all entered, the calc_roomcap function calculates the estimated room capacity.

Functional Dependencies:

sid->email, surname, prefname
email->sid, surname, prefname
term, cid -> subj, crse, min_units, max_units, sec, Course.id
Course.id->term, cid, subj, crse, min_units, max_units, sec
term, cid, sid -> seat, units_gained, status, class, major, level, grade
term, cid, seat-> sid, units_gained, status, class, major, level, grade
Meeting.id->course_id, type, room, Instructor_id, build, meeting_id, DayTime.id, day,
start_time, end_time, Instructor.id, fullname

Instructor.id->fullname
term, day, start_time, end_time, full_name->build, room
term, day, start_time, end_time, build, room->full_name RoomCap.build,
RoomCap.room->capacity

A)

Units taken	Percent of students
1	0.031924
2	0.013351
3	0.018182
4	0.471793
5	0.070662
6	0.011683
7	0.016778
8	0.101082
9	0.03254
10	0.014838
11	0.017068
12	0.10848
13	0.04517
14	0.014188
15	0.007628
16	0.010505
17	0.005775
18	0.001234
19	0.000726

20	0.000418

B)

Units taken	Average GPA
1	0.214469
2	0.697108
3	2.289995
4	2.130529
5	2.019768
6	1.969245
7	2.081111
8	2.170058
9	2.073166
10	2.113291
11	2.14114
12	2.093817
13	2.060128
14	2.102853
15	2.159226
16	2.225336
17	2.24871
18	2.174019
19	2.214141
20	2.121216

C)
Hardest Overall Instructor: Turner, Emily A. 1.7
Easiest Overall Instructor: O'donnell, Madison G. 3.95

D)
By Pass Rate
Course: 114
Hardest Instructor: Perry, Katherine V. 0.75
Easiest Instructor: Williams, Victoria M. 1.0

Course: 113
Hardest Instructor: Diaz, Michelle H. 0.87
Easiest Instructor: Herring, Nathan L. 1.0

Course: 112
Hardest Instructor: 0.71
Easiest Instructor: Morris, Evan E. 1.0

By GPA
Course: 108
Hardest Instructor: Green, Isabella M. 2.0
Easiest Instructor: Olson, Jennifer D. 3.17

Course: 109
Hardest Instructor: Fisher, Caleb K. 3.04
Easiest Instructor: Parsons, Mia E. 3.33

Course: 111
Hardest Instructor: Morris, Evan E. 3.96
Easiest Instructor: Morris, Evan E. 3.96

Course: 110
Hardest Instructor: Cobb, Sophie A. 3.54
Easiest Instructor: Cobb, Sophie A. 3.54

Course: 102
Hardest Instructor: Parsons, Mia E. 3.18
Easiest Instructor: Parsons, Mia E. 3.18

Course: 103
Hardest Instructor: Moore, Isabella M. 2.8
Easiest Instructor: Donaldson, Matthew C. 3.5

Course: 101
Hardest Instructor: Diaz, Riley I. 2.21
Easiest Instructor: Logan, Jackson J. 3.21

Course: 106
Hardest Instructor: Whitehead, William A. 1.98
Easiest Instructor: Dodson, Nicole M. 3.46

Course: 107
Hardest Instructor: Bates, Logan Q. 2.45
Easiest Instructor: Edwards, Maya N. 3.14

Course: 104
Hardest Instructor: Miller, Emma J. 1.99
Easiest Instructor: Murphy, Melanie S. 3.27

Course: 105
Hardest Instructor: Adams, Emily G. 1.82
Easiest Instructor: Williams, Victoria M. 3.37

E)
The rows that look like duplicates actually have different DayTimes.

Subject1	Number1	Subject2	Number2	Term
'ABC'	'104'	'ABC'	'107'	'200906'
'ABC'	'104'	'ABC'	'108'	'200106'
'ABC'	'104'	'ABC'	'108'	'200106'
'ABC'	'104'	'ABC'	'221'	'200906’
'ABC'	'104'	'ABC'	'108'	'200106'
'ABC'	'105'	'ABC'	'107'	'199406'
'ABC'	'105'	'ABC'	'107'	'199406'
'ABC'	'105'	'DEF'	'201'	'200106'
'ABC'	'105'	'ABC'	'107'	'199506'
'ABC'	'105'	'DEF'	'250'	'200606'

'ABC'	'105'	'ABC'	'107'	'199406'
'ABC'	'108'	'DEF'	'201'	'200906'
'DEF'	'201'	'DEF'	'258'	'200606'

F)
Best Performing Major:
O207 4.0
Worst Performing Major:
O153 0.5

G)
Percentage of students who transfer into an ABC major: 0.0454350752662

Top 5 majors with highest percentage of students who transfer into an ABC major

Previous major	Percent
O171	1
O276	1
O169	0.5
OT97	0.5
O230	0.4

5a)
ABC222:
>95%: ABC 221 0.969
>90%: ABC 108 0.916
>80%: ABC 104 0.82

ABC210:
>95%: ABC 209 0.958
>85%: ABC 108 0.867
>75%: ABC 104 0.783
>75%: ABC 221 0.778
ABC203:
>90% ABC 202 0.926
>80%: ABC 108 0.824
>75%: ABC 221 0.785