

Sample Final

April 14, 2022 9:10 AM



Final Sample
converted

Fundamentals of Data Engineering (COE848)

Sample Final Exam

This is a CLOSED BOOK exam. Textbooks, notes, laptops, calculators, personal digital assistants, cell phones, and Internet access are NOT allowed.

Please read each question carefully, and write your answers legibly in the space provided. You may do the questions in any order you wish, but please USE YOUR TIME WISELY.

When you are finished, please hand in your exam paper and make sure you are signed out.
Good luck!

Student Name: _____
Student ID: _____
Score: ____ %

Question	Maximum Mark	Received
1	16	
2	5	
3	15	
4	14	
5	15	
6	15	
Total	80	

1 of 9

Question 1 (15 Marks): Explain the following in 2-3 sentences:

make sure to be comfortable
with uses of nosql concepts

a) Transitive Dependency (nosql)

$A \rightarrow B \rightarrow C$ then $A \rightarrow C$

B is dependent on A & C dependent on B so can say C dependent on A

b) Consistency in NoSQL databases

c) Graph Databases (nosql)

→ can answer questions over structural graphs
↳ these questions not easy to answer in relational DB (need to join etc.)

d) Key-Value Databases (nosql)

↳ key-value compared to relational databases :

→ much faster

→ can't have complex queries

→ very simple of basic structure keys & values

2 of 9

Question 4 (14 Marks): Assume you are in charge of managing the program committee for an important conference. The following database stores information about papers submitted to the conference (Paper table), reviewers on the program committee (Reviewer table), and the assignment of reviewers to papers (Reviews table). Each reviewer on the program committee will have to review a set of papers. Each paper will be reviewed by some subset of reviewers.

Paper (pid, title)
Reviewer (rid, name)
Reviews (rid, pid)

a) Write a SQL query that finds all papers with fewer than three reviewers assigned to them. The output of the query should be a list of paper titles. The result should also include papers without any reviewers assigned to them.

→ need LEFT join to include papers with no reviews, must put all attributes in select into group statement?

select title from paper left join reviews as R where P.pid = R.pid

group by P.pid, title having count

Answer:
select pid, title from paper LEFT outer join reviews as R
ON P.pid = R.pid GROUP BY rid, title HAVING COUNT(*) < 3

picks paper & reviews
where pid's match, but
also has rid's where
it doesn't have a rid

5 of 9

b) Write a SQL query that finds the reviewers with the most papers assigned to them. There can be more than one such reviewer. The output of the query should be a list of reviewer names. A reviewer should be listed if no other reviewer has strictly more papers to review.

finding # of papers assigned to reviewers

→ need to join reviews and papers & all rows assigned and max of all rows assigned means inner join
(select R.name, R.rid from reviewer as R
join reviews as R
on R.rid = R.rid
group by R.name, R.rid
having
(count(*) > (select count(*) from reviews group by rid)))

↑
needed to
count so we
can find max
of all rows
assigned

Question 5 (Mark 15): Refer to the following schema design and answer the queries using SQL queries:

Country (CName, Continent, GDP, Population)
Politician (PName, Gender, Age, Office, Country, Region)
Head_of_Government (Country, PNName)
Alliance_Membership (Country, Alliance)

a) List the countries in the database where all of the senators (office=senator) are women.

→ have to think countries where ALL senators are women
find countries where it has senators with mix of genders
then exclude it from countries who only has men

select h.Country from Head_of_Government as H, Politician as P
where h.PName = p.PName and gender = "male" and office = "senator"
..... extra

rid	titles	count
rid 1		10
rid 2		5
:		2
rid 3		3
:		;

reviewer	joining	reviews
1	sarah	R1
1	sarah	R2
1	Sarah	R3
2	John	R4

1	Sarah	3
2	John	1

then you group by

↑
Comparing this
to assuming
else in
subquery

if you just select
countries you'll
get countries
that have
senators
in them

6 of 9

- b) List the countries in the database that are in NATO (Alliance=Nato) but do not have a female head of government.

7 of 9

Question 6 (15 Marks): Consider the following DTD Document.

```
<!DOCTYPE Hogwarts [  
<!ELEMENT Hogwarts (Rooms, Teachers, Courses) >  
<!ELEMENT Rooms (Room*) >  
<!ELEMENT Room EMPTY >  
<!ATTLIST Room  
    name ID #REQUIRED  
    nrSeats CDATA #IMPLIED >  
  
<!ELEMENT Teachers (Teacher*) >  
<!ELEMENT Teacher (Title*) >  
<!ELEMENT Title (#PCDATA) >  
<!ATTLIST Teacher  
    name ID #REQUIRED  
    room IDREF #REQUIRED >  
  
<!ELEMENT Courses (Course*) >  
<!ELEMENT Course (Class*) >  
<!ATTLIST Course  
    name ID #REQUIRED  
    teacher IDREF #REQUIRED  
    nrStudents CDATA #IMPLIED >  
<!ELEMENT Class EMPTY >  
<!ATTLIST Class  
    day CDATA #REQUIRED  
    hour CDATA #REQUIRED >  
]>
```

- a) Write an XPath expression to find all courses that have at least 20 students

8 of 9

b) Write an XPath expression to list all professors at the school

c) Find all rooms that are used on Mondays