



UNIVERSITETET I BERGEN

KANDIDAT

135

PRØVE

# INF115 0 Databaser og modellering

Emnekode	INF115
Vurderingsform	Skriftlig eksamen
Starttid	14.05.2024 09:00
Sluttid	14.05.2024 12:00
Sensurfrist	--
PDF opprettet	31.05.2024 13:16

**Seksjon 1**

Oppgave	Tittel	Oppgavetype
<b>i</b>	Om eksamen	Informasjon eller ressurser

**Multiple choice / short questions**

Oppgave	Tittel	Oppgavetype
1	SQL: reserved words	Flervalg
2	SQL: valid statement	Flervalg (flere svar)
3	SQL: Aggregation	Flervalg
4	Shiny: Front-end	Flervalg
5	Shiny: Component	Flervalg
6	Indeks	Flervalg
7	SQL: Primary/foreign	Flervalg (flere svar)
8	Normalization: Invalid form	Flervalg (flere svar)
9	Concurency: Dining philosophers	Flervalg
10	The problem with relational models	Flervalg (flere svar)

**SQL queries**

Oppgave	Tittel	Oppgavetype
11	SQL queries	Tekstfelt

**ER-diagram & Normalization**

Oppgave	Tittel	Oppgavetype
12	Normalization / ER diagrams	Langsvar

**Oblig score**

Oppgave	Tittel	Oppgavetype
13	Poeng fra oblig 1 og 2	Muntlig

**1 SQL: reserved words**

Hvilket er ikke et "reserved word" for SQL:

**Velg ett alternativ:**

- ☐ DISTINCT
- ☐ SELECT
- ☐ WHERE
- ☒ BECAUSE

**Knytte håndtegninger til denne oppgaven?**  
Bruk følgende kode:

**9 4 1 3 6 2 3**

## 2 SQL: valid statement

Hvilke av disse er "valid SQL statements" ?

**Velg ett eller flere alternativer**

- ☒ SELECT Polarbears FROM EndangeredSpecies
- ☐ FIND EXTINCT Polarbears IN Svalbard
- ☒ SELECT AnsattNr, Etternavn, Lønn FROM Ansatt
- ☐ GET AnsattNr FROM Polarbears

Knytte håndtegninger til denne oppgaven?  
Bruk følgende kode:

**2 2 6 8 2 5 4**

## 3 SQL: Aggregation

Hvilken av disse er ikke en "reserved SQL aggregation word"

**Velg ett alternativ:**

- ☒ UPPER
- ☐ AVG
- ☐ COUNT
- ☐ MIN

Knytte håndtegninger til denne oppgaven?  
Bruk følgende kode:

**3 9 5 0 1 1 3**

## 4 Shiny: Front-end

Shiny er en såkalt front-end application, i motsetning til databasen som er en del av back-end. Hvilken av disse er ikke en del av front-end

**Velg ett alternativ:**

- ☐ En action button for å aktivere en event
- ☒ Et SQL join call på 2 tabeller, tidligere spurt om fra brukeren til databasen.
- ☐ En DataTable component som viser data fra et SQL join call på 2 tabeller tidligere spurt om fra brukeren.
- ☐ En DataTable component som viser data fra en database

**Knytte håndtegninger til denne oppgaven?**  
Bruk følgende kode:

**9 3 8 8 0 4 8**

## 5 Shiny: Component

Shiny components er delt i:

Input components (input text, select one box, checkbox etc)

Output components (DataTable, interactive map, interactive graph etc)

Hvilke components-typer kan hente sin informasjon fra en database?

**Velg ett alternativ:**

- ☐ Ingen, en database definerer aldri innholdet til en komponent.
- ☒ Output components
- ☐ Input & output components
- ☐ Input components

**Knytte håndtegninger til denne oppgaven?**  
Bruk følgende kode:

**2 7 2 2 1 0 1**

## 6 Indeks

Hvis en indeksert tabell over alle gener i mennesket inneholder 3 kolonner:  
Gen-navn, kromosom, kromosom-posisjon,  
og hvis mennesket har 24 kromosomer, og filen er sortert og indeksert på kromosom,  
hvor mange rader må en indeks ha for å beskrive alle "indeks hopp"  
mellom kromosomer i tabellen ?  
(indeksen inneholder informasjon også om kromosom 1)

**Velg ett alternativ:**

- ☐  $(24 \cdot 23) / 2 = 276$
- ☒ 24
- ☐ 48
- ☐  $24 - 1 = 23$

Knytte håndtegninger til denne  
oppgaven?  
Bruk følgende kode:

**7 5 0 2 0 7 0**

## 7 SQL: Primary/foreign

Velg riktig utsagn angående "value restriction" for primary og foreign keys.

**Velg ett eller flere alternativer**

- ☐ Primary keys can contain duplicates
- ☒ Foreign keys can contain Null values.
- ☒ Primary keys must not contain duplicates
- ☐ Foreign keys must not contain Null values.

Knytte håndtegninger til denne  
oppgaven?  
Bruk følgende kode:

**1 7 6 0 8 7 1**

## 8 Normalization: Invalid form

Hvilke av disse er ikke en normaliseringsform?

**Velg ett eller flere alternativer**

- ☒ -1NF
- ☐ 1NF
- ☐ BCNF
- ☐ 0NF/UNF
- ☐ 4NF

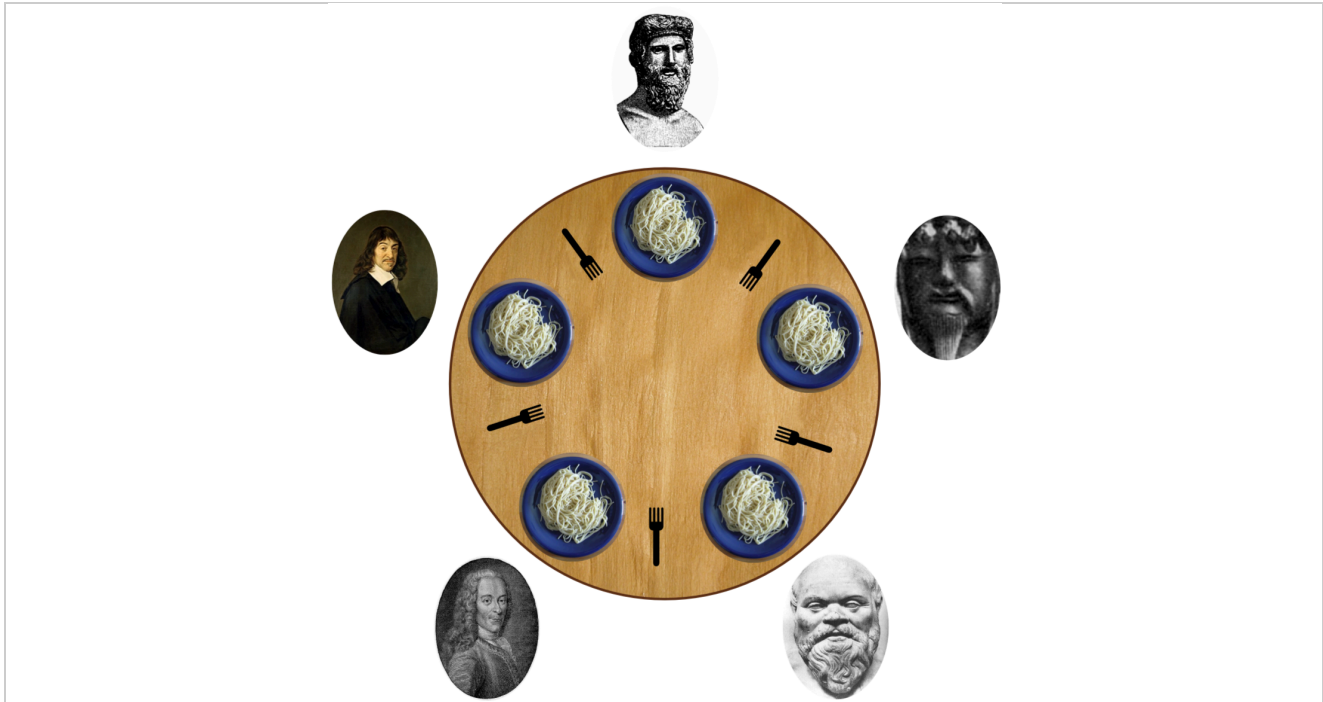
**Knytte håndtegninger til denne oppgaven?**

Bruk følgende kode:

**8 6 7 9 6 0 1**



## 9 Concurrency: Dining philosophers



Dining philosophers problem, er et kjent dilemma i concurrency (når flere prosesser kjører sammen og bruker samme data).

I dette problemet, veksler filosofer mellom å tenke og å spise. Hver filosof har én gaffel, men kan kun spise med to (gaffelen til høyre og venstre for seg).

Hva er max antall filosofer som kan spise samtidig i dette bildet?

**Velg ett alternativ:**

☐ 4

☒ 2

☐ 1

☐ 3

**Knytte håndtegninger til denne oppgaven?**

Bruk følgende kode:

**4 6 6 7 6 3 1**

## 10 The problem with relational models

Relasjonsmodeller som SQL kan ha et stort problem når det gjelder "real world application". Dette er grunnen til at alternativer som NoSQL også finnes for å dempe krav om spesifiserte relasjoner. Hvilket/hvilke problem er det?

**Velg ett eller flere alternativer**

- ☒ Det er et problem at ikke alle tabeller har en direkte relasjon (altså uten å gå igjennom andre tabeller).
- ☐ Det er et problem at ikke alle tabeller kan ha foreign keys.
- ☐ Det er et problem at SQLite ikke kan ha mer enn 100 rader per tabell.
- ☒ Det er et problem at data må være ryddet opp/standardisert slik at relasjoner kan bygges igjennom tabeller.

**Knytte håndtegninger til denne oppgaven?**  
Bruk følgende kode:

**9 2 3 9 8 2 3**

## 11 SQL queries

Se på en filmdatabase med følgende tabeller. Primary keys er merket med understreking, foreign keys med stjerne\*:

Movies (Id, Title, Language, Year, Runtime, Budget, Revenue),

People (Id, Name),

Actors (Movie\_id\*, Actor\_id\*, Character) with

FOREIGN KEY( Movie\_id ) REFERENCES Movies( Id ) and

FOREIGN KEY( Actor\_id ) REFERENCES People( Id ).

(a)

Skriv en SQL-spørring som viser **movie titles, budget og revenue** for filmer med et budsjett over 50 millioner dollar.

**Skriv ditt svar her**

```
SELECT
  title as movie titles,
  budget,
  revenue
FROM
  movies
WHERE
  budget > 50000000;
```

(b)

Skriv en SQL-spørring som viser **de 5 topp språkene sortert etter gjennomsnittlig inntekt** (average revenue) for alle filmer på det språket. Utskriften skal vise for eksempel:

en, 1264125.54

fr, 26354.34

...

**Skriv ditt svar her**

```
SELECT
  language, avg_revenue
FROM(
  SELECT
    language, AVG(revenue) OVER(PARTITION BY language) as avg_revenue
  FROM
    movies
  GROUP BY
    language
)
ORDER BY
  avg_revenue DESC
```

```
LIMIT 5;  
DESC(max-min);
```

(c)

Skriv en SQL-spørring som viser **hvor mange** filmer som har en karakter som heter nøyaktig "Fred" (dvs. ikke "Fred Astaire"). Bruk kun en av tabellene.

**Skriv ditt svar her**

```
SELECT  
  COUNT(*)  
FROM(  
  SELECT  
    DISTINCT movie_id  
  FROM  
    actors  
  WHERE  
    Character='Fred'  
);
```

(d)

Skriv en SQL-spørring som viser **alle filmtitler** på filmer som har en karakter som heter nøyaktig "Fred".

**Skriv ditt svar her**

```
SELECT  
  m.title  
FROM  
  movies m  
  JOIN actors a ON m.id=a.movie_id  
WHERE  
  a.character='Fred';
```

(e)

Skriv en SQL-spørring som viser filmen og skuespilleren der skuespillere har "Fred" som en *del* av navnet **og** har spilt en karakter som heter *nøyaktig* "Fred". Utskriften skal vise **movie, actor name, character name**, for eksempel:

Thursday the 12th, Frederick Oswald, Fred  
Saving Private Fred, Freddie Tungsten, Fred  
...

**Skriv ditt svar her**

```
SELECT  
  m.title, p.name, a.character
```

```
FROM
  movies m
  JOIN actors a ON m.id=a.movie_id
  JOIN people p ON a.actor_id=p.id
WHERE
  p.name LIKE '%Fred%' AND a.character='Fred';
```

**Knytte håndtegninger til denne oppgaven?**  
Bruk følgende kode:

**4 4 0 6 9 1 1**

## 12 Normalization / ER diagrams

Et forskningsprosjekt ønsker å undersøke hvilke kjente personer som møtte hverandre på en serie middagsselskaper i London på 1800-tallet, basert på lesekilder som omtaler festene, som brev og dagbøker.

En av forskerne har samlet følgende informasjon om festdeltakerne, alt i én stor regnearkfil. De har lagt merke til at de trenger å skrive den samme informasjonen mange ganger i regnearket og har bedt om hjelp til å organisere dataene sine bedre.

På en gitt dato er det bare én fest som skjer. Dette eksemplet viser noen personer fra 2 forskjellige fester som skjedde på 1830-06-01 and 1830-07-15

Party data							
date	guest	birth	death	related to	source book	source page	source quote
1830-06-01	Charles Hallam	1809	1882		Abc (1953)	12	Also Charles and his wife attended
1830-06-01	Emma Hallam	1808	1896	wife of Charles Hallam	Abc (1953)	12	Also Charles and his wife attended
1830-06-01	George Porter	1810	1857		Klm (1977)	37	On 1 June, George Porter attended an evening get-together with Jane Pollock. Charles Hallam was there, too
1830-06-01	Jane Pollock				Klm (1977)	37	On 1 June, George Porter attended an evening get-together with Jane Pollock. Charles Hallam was there, too
1830-06-01	Charles Hallam	1809	1882		Klm (1977)	37	On 1 June, George Porter attended an evening get-together with Jane Pollock. Charles Hallam was there, too
1830-07-15	Charles Hallam	1809	1882		Abc (1953)	23	15 July - Charles attending dinner party

For denne øvelsen kan du anta følgende:

- På en gitt dato er det bare én fest som skjer.
- Enhver fest har minst én kjent gjest.
- Ingen gjester har samme navn i hele datasettet.
- Vi kjenner ikke alle fødsels- eller dødsår for gjestene.
- For noen gjester kjenner vi et forhold mellom dem.
- Kolonnen *related\_to* kan ha flere forbindelser ("onkel til XYZ, gift med ABC")
- En gitt bok kan gi mer enn ett sitat på en gitt side.
- Noen ganger er det mer enn én kilde som nevner den samme gjesten på samme fest (Charles på 1830-06-01 er nevnt i to forskjellige bøker).

## DEL A BØR BESVARES HER DIGITALT

## DEL B BØR TEGNES PÅ SCANTRON PAPIR

### DEL A - BESVARES HER DIGITALT

Foreslå et databaseskjema som er normalisert så mye som mulig og forklar dine valg kort. Du står fritt til å legge til attributter hvis de er nyttige, men du kan ikke utelate informasjon som er i regnearket. Bruk notasjonen fra emnet:

Tabellnavn (primærnøkkel, fremmednøkkel\*, andre\_attributter...)

Du trenger ikke å merke null / non-null.

**Hint:** Bare attributtene *birth* og *death* kan inneholde nullverdier i de endelige normaliserte tabellene. Alle andre attributter er alltid definert.

**Hint:** Du trenger minst 4 tabeller.

### DEL B - TEGNES PÅ SCANTRON PAPIR

Tegn et ER-diagram som gjenspeiler den foreslåtte løsningen fra del A. Du trenger ikke extended-ER-funksjoner. Tegn alle enhetene, alle attributter og alle relasjoner (med mengdetall) som løsningen din trenger.

### Skriv ditt svar til del A her

```
Guest(id,first_name,last_name,birth,death)
```

```
Related(guest1_id*,guest2_id*,relation) with  
FOREIGN KEY( guest1_id) REFERENCES Guest(id) and
```

FOREIGN KEY( guest2\_id) REFERENCES Guest(id).

Party(party\_id,date)

Source(source\_id,name,year)

Source\_quote(quote\_id,source\_id\*,page,quote) with  
FOREIGN KEY(source\_id) REFERENCES Source(source\_id).

Party\_data(party\_id\*,guest\_id\*,quote\_id\*) with  
FOREIGN KEY(party\_id) REFERENCES party(party\_id) and  
FOREIGN KEY(guest\_id) REFERENCES guest(id) and  
FOREIGN KEY(quote\_id) REFERENCES source\_quote(quote\_id)

Jeg deler regnearket opp i 6 tabeller og gir 5 av tabellene unike ider for å gjøre det lettest mulig å føre ny data og som gjør at databasen har mindre duplikat av data, altså at det oppstår mindre redundansy. Jeg lagde spesifikt Party grunnet redundancy siden hvis det er en stor database så sparer man mye plass med å kun skrive party\_id sammenlignet med date, selv om date er en candidate key, samme med Guest, der er det bedre å ha en id selv om ingen har samme navn og first\_name og last\_name kunne blitt brukt som primary keys alene. Når alt av dataen er avhengig av hverandre så fører det også til at det ikke oppstår aggregate,update og delete feil grunnet dataen er utdatert eller ført med slurvfeil. Det blir også ingen partial dependency, alle verdiene er atomic, og det er ingen trival-transitive dependency.

mine assumptions (for cardinality på B):

- det kan være en source som ikke blir brukt.
- gjester må ikke ha vært på middagsfest
- en fest må ikke ha loggført data

**for å gjøre relationshipet Relation mer tydelig mtp cardinality lagde jeg en forstørret modell til venstre på arket mellom ()**

Ord: 236

**Knytte håndtegninger til denne oppgaven?**  
Bruk følgende kode:

**7 9 3 5 0 5 0**



Håndtegning 1 av 1

Fyll inn oppgavekode og emneinformasjon på alle skissearkene

Fill out question code and test information on every sheet

Oppgavekode  
Question codeDato  
DateEmnekode  
Subject codeKandidatnummer  
Candidate numberOppgavenummer  
Question numberSidetall  
Page number

7935050

5/14/2024 INF115

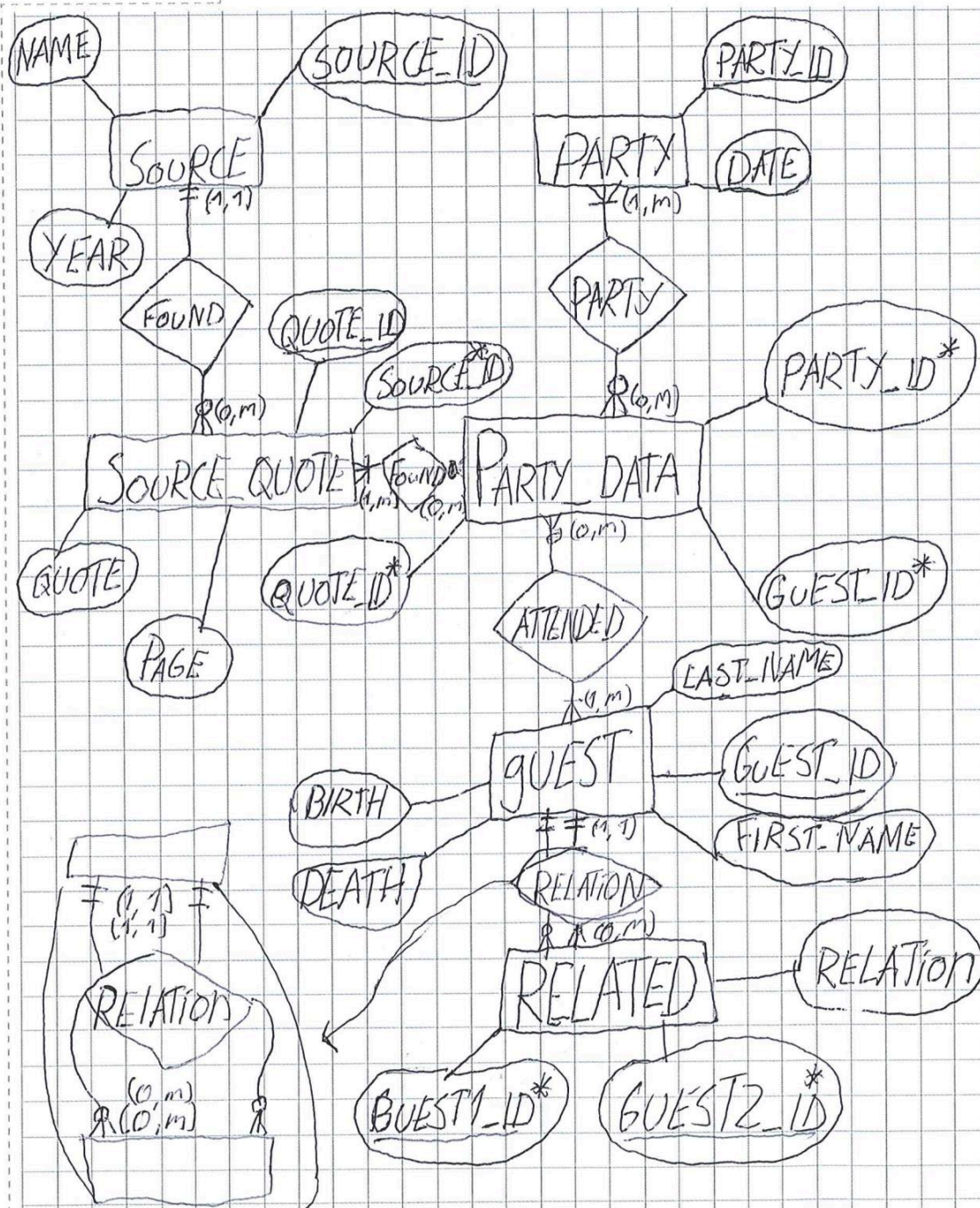
135

12

1/1 av/of

0	0	0	0	0	0
1	1	1	1	1	1
2	2	2	2	2	2
3	3	3	3	3	3
4	4	4	4	4	4
5	5	5	5	5	5
6	6	6	6	6	6
7	7	7	7	7	7
8	8	8	8	8	8
9	9	9	9	9	9

Tegneområde Drawing area



## 13 Poeng fra oblig 1 og 2

Du trenger ikke gjøre noe her.

Vi bruker denne delen for å registrere poengsummen fra obligene.

**Knytte håndtegninger til denne oppgaven?**  
Bruk følgende kode:

**3 3 6 9 9 6 3**