



Økt 10 (av 12)

DB1100 Databaser

(Tomas Sandnes / tomas.sandnes@kristiania.no)

Dagens temaer

Dagens pensum: [Læreboka kapittel 8, fra 8.2.5 og utover](#)

- Fra forrige økt:
 - Resultater Kahoot
 - Repetisjon Normalisering
- Dagens tema, mer [Normalisering](#):
 - Boyce-Codd normalform (BCNF)
 - Denormalisering

Resultat, forrige Kahoot

- Nok en gang slo dere fjorårets studenter, stillingen er nå **6-1 til dere**.
- Dere: **68,42 % riktig svar!**
 - I fjor: 64,74 % riktig.
 - Klar rundeseier til dere! :-)

Game Over	
Show feedback	
90k&NoWifi	12,361
gulpo	12,253
Ensomjente_99	12,158
Per_Uten_H	11,988
KebabRoy	11,858

DB1100 - økt 9 - Normalisering	
Played on	25 Oct 2018
Hosted by	Tomas_Sandnes
Played with	102 players
Played	10 of 10 questions

Overall Performance	
Total correct answers (%)	68,42%
Total incorrect answers (%)	31,58%
Average score (points)	5961,20 points

- NB: **Ensomjente_99** har ligget på **topplista 3 uker på rad** nå!
 - Meget imponerende!
- MEN: Jeg savner 1..* kobling mot **MedstudentSomInteressererSeg!** ;-)

Lærdom fra Kahoot

Spm. 4: "Hvis A og B er attributter i en relasjon R, så er B funksjonelt avhengig av A hvis"

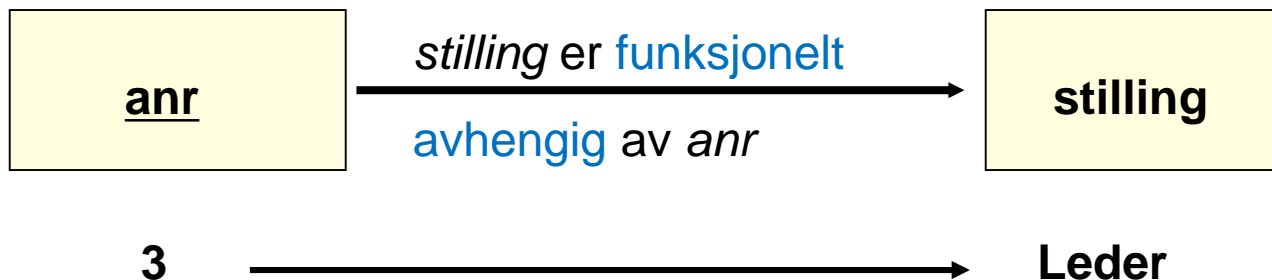
- A) "hver verdi av A er assosiert med flere verdier av B" (26 svar)
- B) "det for hver verdi av A kan identifiseres kun én verdi av B" (43 svar)
- C) "hver verdi av A er assosiert med en og bare en verdi av R" (13 svar)
- D) "det for hver verdi av B finnes eksakt én verdi av A" (18 svar)

Kanskje lettest å forklare med et eksempel:

- La *A* være *anr*, *B* være *stilling*.

Riktig svar:

- Hvis *anr* og *stilling* er attributter i en relasjon *R*, så er *stilling* funksjonelt avhengig av *anr* hvis det for hver verdi av *anr* kan identifiseres kun én verdi av *stilling*



Lærdom fra Kahoot #2

Spm. 7: "B er fullstendig funksjonelt avhengig av A hvis B er funksjonelt avhengig av"

A) "A" (26 svar)

B) "A, men ingen delmengde ("subset") av A" (43 svar)

C) "B" (3 svar)

D) "C" (23 svar)

- *"... er fullstendig funksjonelt avhengig av ..." er et annet uttrykk for:
"... har ingen delvis avhengighet til ..."*
- *Så dette er regelen for 2. normalform (2NF), skrevet med A og B. Altså:*
 - *Ingen delmengde ("subset") av A er determinant for B.
(Dvs. B har ingen delvis avhengighet til A.)*

Er NF, determinant, m.m. fortsatt litt hårete?

- Det kan godt være at deler av databaseemnet fortsatt har vanskelige temaer?
- I så fall – forslag:
 - Ta en titt på en (eller fler) av [studentvideoene fra i fjor](#)! Du finner de nederst i lista med eksterne ressurser for DB1100 på Canvas.
Eksempel: [Olav Sundfør \(fjorårsstudent, veileder i DB1100 i år\) sin videoliste](#)
 - Les, en gang til, dagens avsnitt i læreboka.
 - Google etter mer info (artikler og videoer).
 - Spør veiledere og foreleser i øvingstimene! :-)
- Og: Gjør øvingsoppgavene!
 - (Husk: Løsningsforslag tilgjengeliggjøres når det er én øvingstime igjen.)

MER NORMALISING:

Boyce-Codd normalform (BCNF)

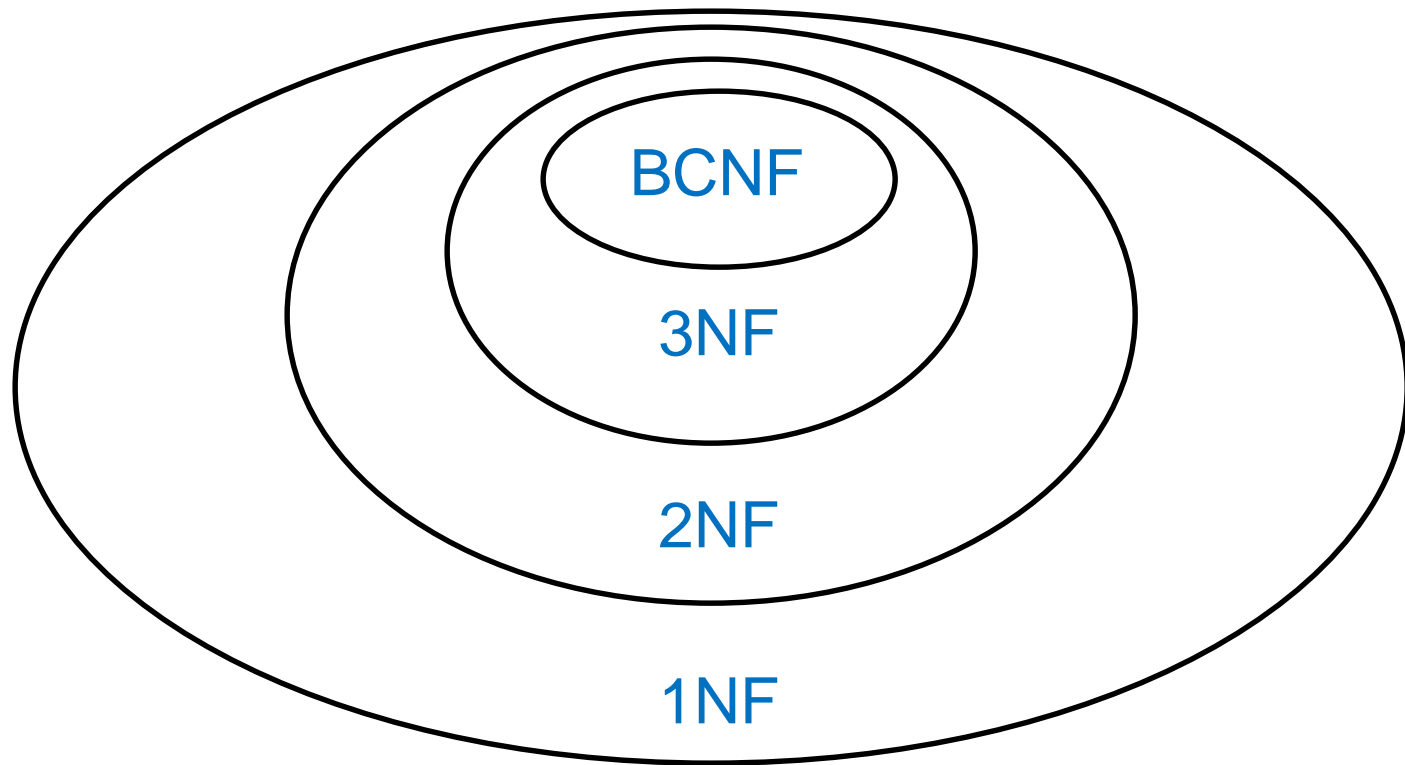
Denormalisering

1NF til 3NF, repetisjon

- 1NF:
 - Cellene inneholder bare *ett* dataelement.
- 2NF, dersom den er på 1NF og:
 - Ingen avhengighet til *deler av primærnøkkelen*. (Ingen delmengde av PK er determinant for kolonner som ikke er del av PK.)
- 3NF, dersom den er på 2NF og:
 - Ingen kolonne utenfor primærnøkkelen har en *transitiv* (indirekte) avhengighet til *primærnøkkelen*. ($A \rightarrow B$ og $B \rightarrow C \Rightarrow A \rightarrow C$.)

Normalformer opp til BCNF

- Enhver tabell på BCNF er på 3NF, osv.:



Boyce-Codd normalform

- Boyce-Codd normalform (BCNF) er en videreutvikling av 3NF. (Den befinner seg mellom 3NF og 4NF.)
- En tabell er på BCNF dersom:
 - Alle determinanter er kandidatnøkler.

Determinant, repetisjon

- Determinant:
 - En kolonne (eller fler) i en tabell som er bestemmende for minst en annen kolonne i tabellen.
- Generelt:
 - Ved $A \rightarrow B$ (B funksjonelt avhengig av A) er A determinant for B.
(Men B er ikke determinant for A.)

Noen nøkler, repetisjon

- Supernøkkel

- En eller flere kolonner som danner en unik identifikator for hver rad i en tabell.
NB: Kan godt være flere kolonner enn det minste, unike utvalget.
("Alle kolonner i en tabell til sammen" danner derfor alltid en supernøkkel.)

- Kandidatnøkkel

- En supernøkkel som ikke kan reduseres til færre kolonner om den fortsatt skal være en unik identifikator.

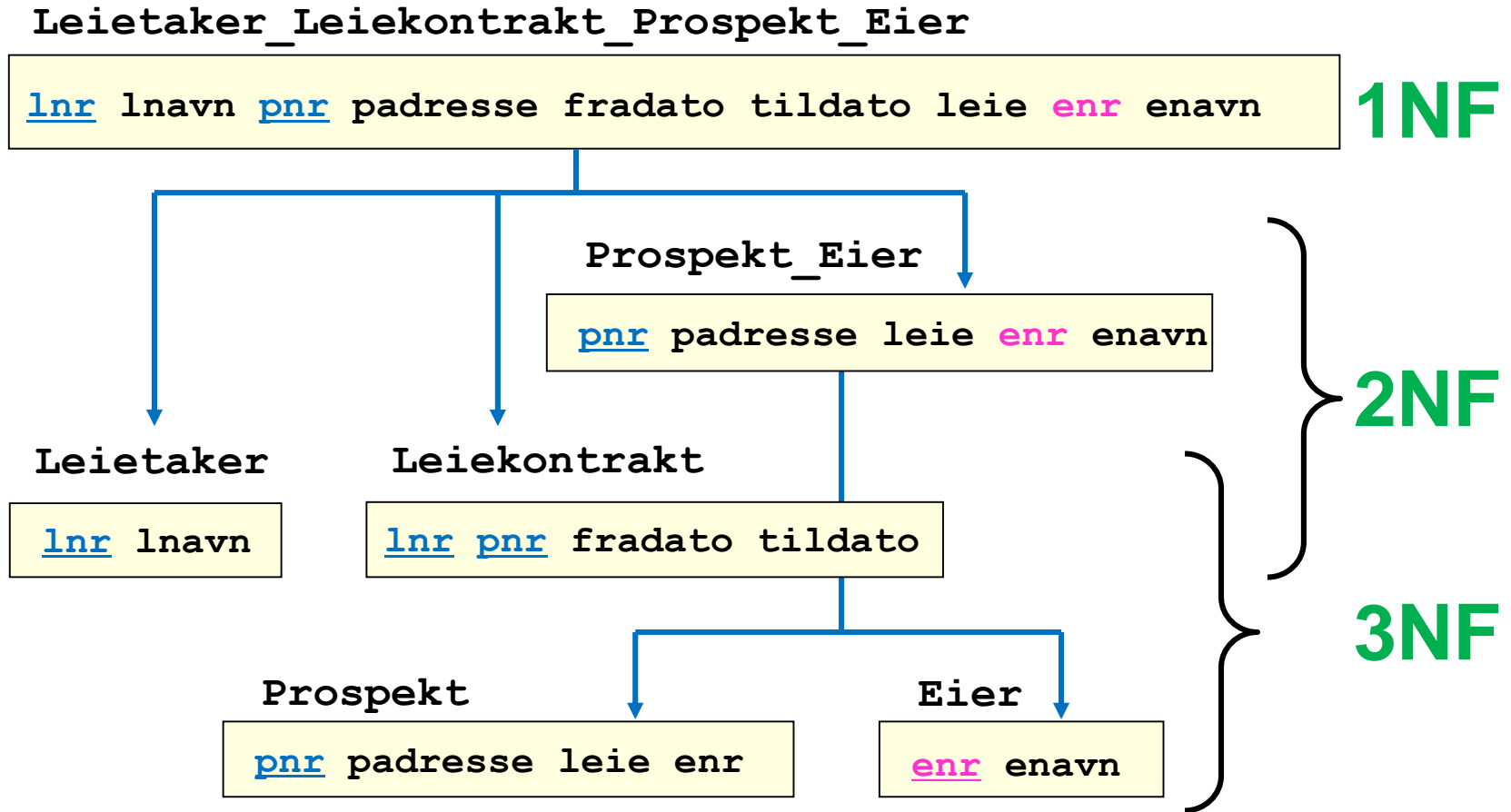
- Primærnøkkel

- Den kandidatnøkkel som velges som identifikator for hver rad i en tabell.

BCNF – forts.

- Som regel oppfylles BCNF automatisk når 3NF oppfylles!
- La oss se tilbake på forrige ukes case, som vi normaliserte til 3NF, og se om denne også oppfyller BCNF.

Oppsummering: fra 1NF - 3NF



Eksempel: fra 3NF til BCNF

- BCNF krav:
 - Alle determinanter er også kandidatnøkler.
- Tabellene **Leietaker**, **Prospekt** og **Eier** er på BCNF da de bare har én determinant hver: (deres PK)
 - Leietaker: lnr, lnavn « **Ok**
 - Prospekt: pnr, padresse, leie, enr « **Ok**
 - Eier: enr, enavn « **Ok**

Eksempel: fra 3NF til BCNF

- Hva med Leiekontrakt?

Leiekontrakt

lnr	pnr	fradato	tildato
---	---	-----	-----
5	6	01-JUL-94	01-SEP-96
5	1	01-SEP-96	01-JAN-98
9	6	01-SEP-96	01-SEP-97
9	2	01-SEP-97	01-SEP-98
9	5	01-SEP-98	null

- Før vi finner determinantene, må vi kjenne til forretningsreglene:
 - En klient kan bare leie et gitt prospekt én gang.
 - En klient kan bare leie ett prospekt av gangen.
 - Oppgave: Finn determinantene.

Eksempel: fra 3NF til BCNF

lnr	pnr	fradato	tildato
---	---	-----	-----
5	6	01-JUL-94	01-SEP-96
5	1	01-SEP-96	01-JAN-98
9	6	01-SEP-96	01-SEP-97
9	2	01-SEP-97	01-SEP-98
9	5	01-SEP-98	null

- **Determinantene:**

- lnr, pnr → fradato, tildato « **Ok**
- lnr, fradato (eller tildato) → pnr, tildato (eller fradato) « **Ok**
- pnr, fradato (eller tildato) → lnr, tildato (eller fradato) « **Ok**

- **Konklusjon:**

- Alle determinanter er også kandidatnøkler!
- Ergo er forrige ukes eksempel ikke bare på 3NF, men også på BCNF.

Eks. BCNF: leietaker_intervju

- Vi bygger videre på forrige ukes eks. (Utleie av leiligheter.)
- Skal modellere intervju prosessen. Forutsetninger:
 - En leietaker blir kun intervjuet én gang på en dag ifbm. leie av leilighet. (Men kan intervjues flere ganger over flere dager.)
 - Hver ansatt holder alle sine intervjuer for en gitt dag på ett eneste rom. (Et rom kan fortsatt deles av flere ansatte.)

Intervju_Ansatt_Rom

lnr	ivju_dato	ivju_tid	ansatt	rom
---	-----	-----	-----	---
76	13-MAY-11	10:30	5	101
56	13-MAY-11	12:00	5	101
74	13-MAY-11	12:00	7	102
56	01-JUL-11	10:30	5	102

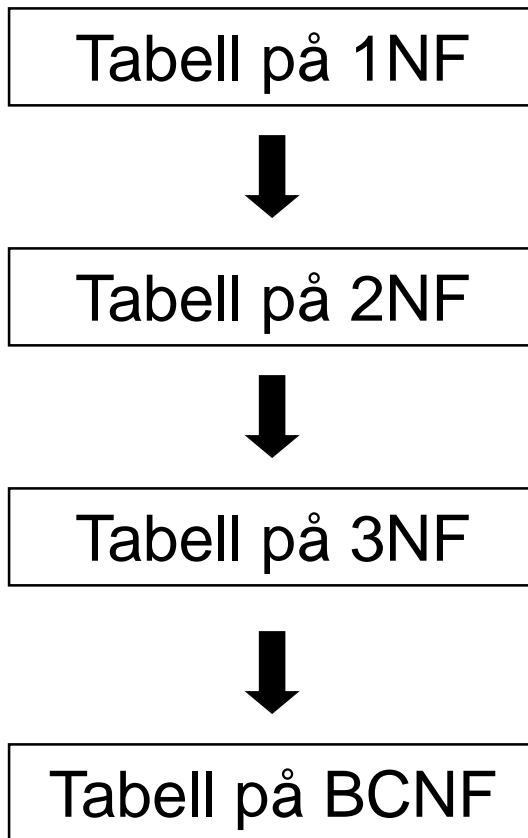
Oppgave som løses i samlingen

- Denne oppgaven går vi gjennom sammen i samlingen.
 - Tenk gjennom den på forhånd! :-)
- Hvilke 3 kandidatnøkler har vi her?

```
leietaker, intervju_dato, intervju_tid, ansatt, rom
```

- Finner dere en determinant til?

Fra 1NF til BCNF



Fjern delvis
avhengighet.

Fjern transitiv
avhengighet.

Fjern resterende
redundans.

Denormalisering

- Med **denormalisering** menes formelt:
 - Endring av tabeller i en database slik at normaliseringsgraden reduseres med en eller fler i forhold til den opprinnelige. (Eks: BCNF til 2NF.)
- Utgangspunktet er gjerne at JOIN tar tid:
 - Vi kan gjøre en SELECT raskere hvis vi slipper å joine.

Denormalisering – forts.

- **Fordeler** med denormalisering:
 - Du unngår mye kobling av tabeller (færre joins).
 - Hastigheten på databasen kan derfor øke betraktelig ved oppslag i store mengder data.
- **Ulemper:**
 - Implementeringen blir vanskeligere.
 - Dobbeltlagring (redundans).
 - Tregere lagring/oppdatering.
 - Fleksibiliteten minker.

Denormalisering – forts.

- Så når er det aktuelt å denormalisere?
 - Når databasen benyttes til mye SELECTS og lite updates.
 - Når systemet ikke kan oppfylle ytelseskravene sine med en normalisert database.
- NB: Denormalisering er unntaket fra det vanlige!
 - Både i dette emnet og i arbeidslivet, tenk 3NF/BCNF med mindre dere får spesifikk beskjed om noe annet!

Videre arbeid i dag (og kort om neste gang)

- Nå:
 - Lese kap. 8 fra 8.2.5 og utover.
 - Se studentvideoer fra opp til 6 studenter (linket nederst på Canvas).
Finner du noen favoritter?!
 - Begynne på øvingsoppgavene.
- Etterpå:
 - Øving (flipped, 2 timer) 12:15 - 14:00, sjekk TimeEdit for rom.
Teori (flipped, 2 timer) 14:15 - 16:00, auditoriet.
- Neste gang: torsdag om 5 uker!
 - Diverse småtemaer vi har igjen.
 - Pensum til neste gang: Kap. 9.3, 10.3.2, 11.1.5, 11.4, 13.3, 15.1, 15.6