



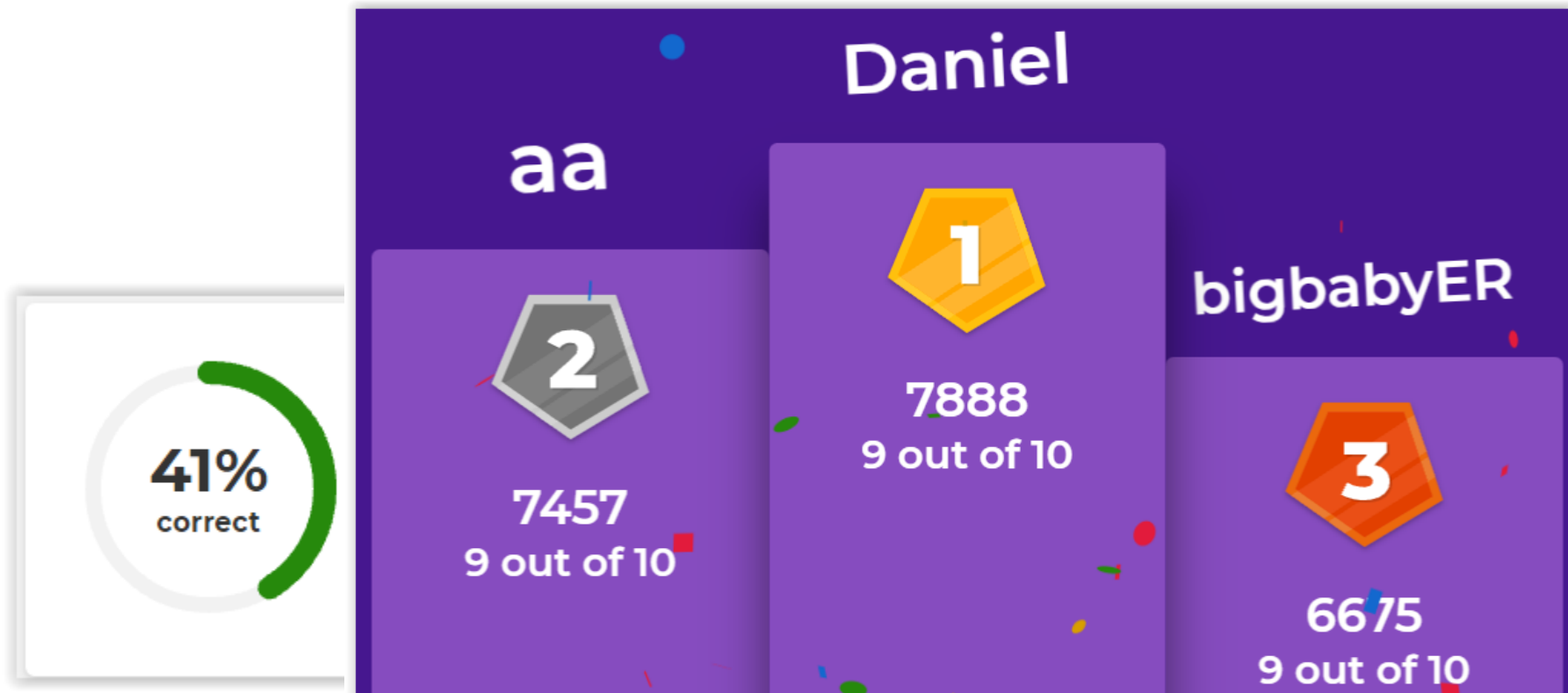
# Økt 10 (av 12)

DB1102 Databaser

Per Lauvås / [per.lauvas@kristiania.no](mailto:per.lauvas@kristiania.no)

Yuan Lin / [yuan.lin@kristiania.no](mailto:yuan.lin@kristiania.no)

# Gårsdagens Kahoot - pallen



# Gårdagens Kahoot

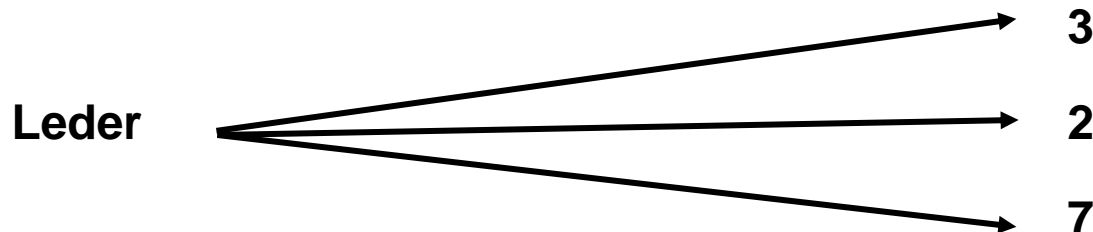
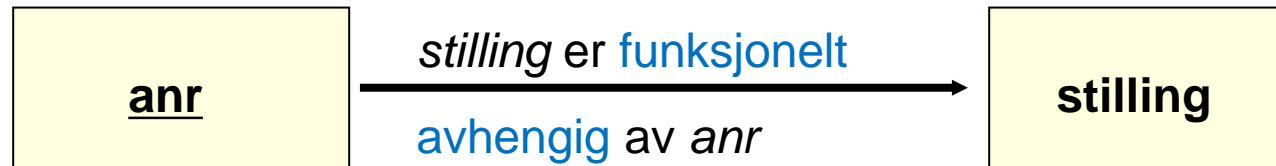
4 -Quiz Hvis A og B er attributter i en relasjon R, så er B funksjonelt avhengig av A hvis

1 of 4 < >



<input checked="" type="checkbox"/>	hver verdi av A er assosiert med flere verdier av B.	✗ <div></div>	11
<input checked="" type="checkbox"/>	hver verdi av A er assosiert med en og bare en verdi av B.	✓ <div></div>	21
<input checked="" type="checkbox"/>	hver verdi av A er assosiert med en og bare en verdi av R.	✗ <div></div>	5
<input checked="" type="checkbox"/>	hver verdi av B er assosiert med en og bare en verdi av A.	✗ <div></div>	30
<input type="checkbox"/>	No answer	✗ <div></div>	7

# Funksjonell avhengighet



# Lærdom fra gårsdagens Zoom

- Hvis dere leverer arbeidskravet i gruppe: Hver enkelt medlem i gruppen må første aktivere Peergrade (ved å gå inn på PeerGrade i Canvas).
- Jeg har oppdatert «Alt du trenger å vite om arbeidskravet...» i Canvas.

# Dagens temaer

Dagens pensum: [Læreboka kapittel 8](#)

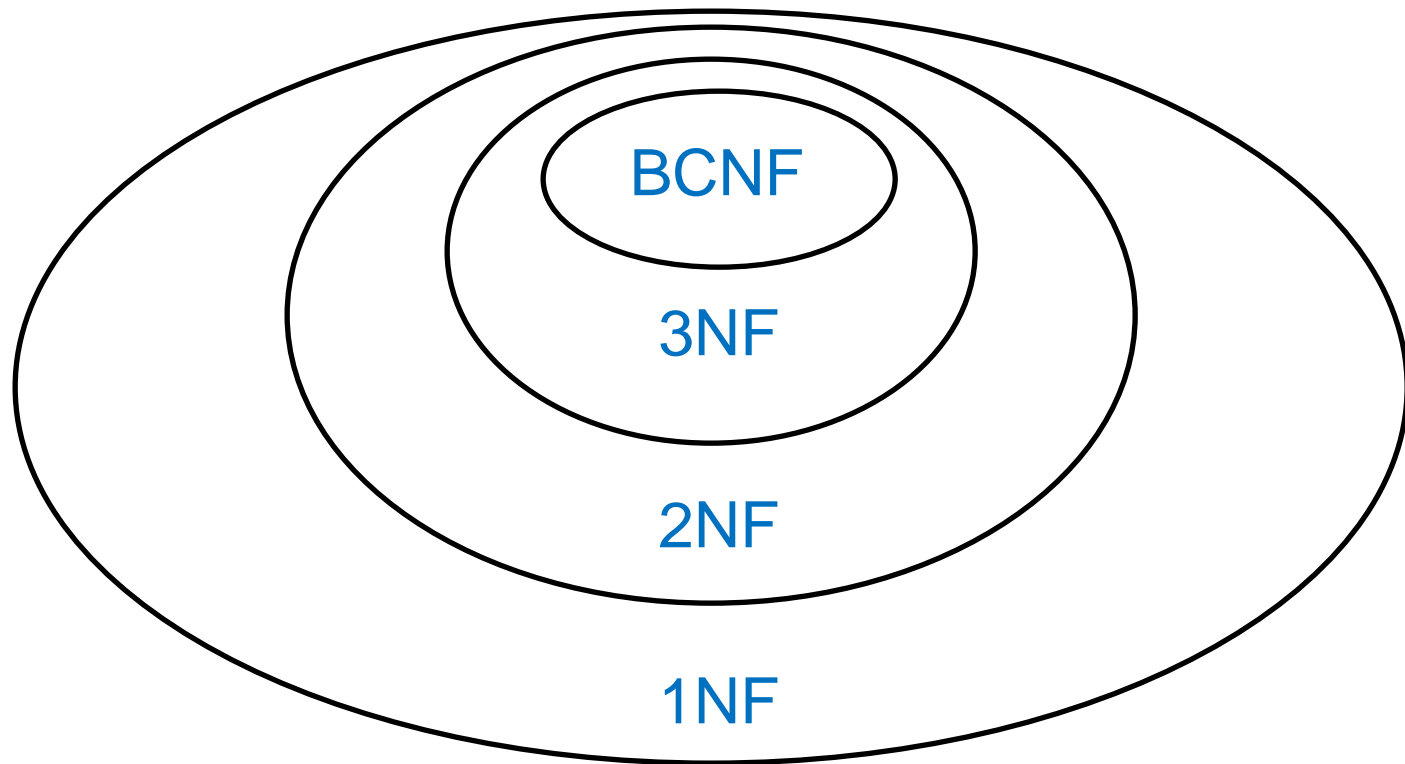
- Dagens tema, mer [Normalisering](#):
  - Boyce-Codd normalform (BCNF)
  - Denormalisering

# **MER NORMALISING:**

## **Boyce-Codd normalform (BCNF)**

# Normalformer opp til BCNF

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





# 1NF til 3NF, repetisjon

- 1NF:
  - Cellene inneholder bare *ett* dataelement. (Er atomære.)
- 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  $B \rightarrow C \Rightarrow A \rightarrow C$ .)

# Unnormalisert form (UNF), repetisjon

- Vi registrerer leieforhold som omfatter leietaker (lnr, lnavn), prospekt (pnr, padresse, fradato, tildato, leie) og eier (enr, enavn).
  - En tabell er unormalisert dersom den inneholder celler med mer enn ett dataelement.

<i>lnr</i>	<i>lnavn</i>	<i>pnr</i>	<i>padresse</i>		<i>fradato</i>	<i>tildato</i>	<i>leie</i>	<i>enr</i>	<i>enavn</i>
5	Hansen	6	Aveien	1	01-JUL-94	01-SEP-96	3500	4	Olsen
?	?	1	Bveien	8	01-SEP-96	01-JAN-98	4400	12	Larsen
9	Persen	6	Aveien	1	01-SEP-96	01-SEP-97	3500	4	Olsen
?	?	2	Dveien	6	01-SEP-97	01-SEP-98	5500	6	Alfsen
?	?	5	Eveien	3	01-SEP-98	null	5000	6	Alfsen

# Første normalform (1NF)

- En tabell er på **1NF** dersom:
  - Cellene inneholder bare ett dataelement.
  - (NB: "Ett dataelement" kan godt være flere ord, f.eks. i en varchar celle.)
- Vi kan fjerne cellene med flere dataelementer ved å lage én rad per element:
  - Da inneholder hvert felt kun en verdi.
  - Primærnøkkelen blir her sammensatt av (lnr, pnr).

## Leietaker\_Leiekontrakt\_Prospekt\_Eier

lnr	lnavn	pnr	padresse	fradato	tildato	leie	enr	enavn
---	-----	---	-----	-----	-----	----	---	-----
5	Hansen	6	Aveien 1	01-JUL-94	01-SEP-96	3500	4	Olsen
5	Hansen	1	Bveien 8	01-SEP-96	01-JAN-98	4400	12	Larsen
9	Persen	6	Aveien 1	01-SEP-96	01-SEP-97	3500	4	Olsen
9	Persen	2	Dveien 6	01-SEP-97	01-SEP-98	5500	6	Alfsen
9	Persen	5	Eveien 3	01-SEP-98	null	5000	6	Alfsen

# Gårdagens Kahoot

8 -Quiz En tabell er på 1NF dersom

4 of 4 < >



Cellene inneholder én og bare én verdi.



24



Alle rader er unike.



9



Alle rader har en primærnøkkel.



13



Alle celler inngår i bare én kolonne.



18



No answer



10

# Andre normalform (2NF)

- En tabell er på 2NF dersom:
  - Den er på 1NF.
  - Og ingen delmengde ("subset") av primærnøkkelen er determinant for en eller flere andre kolonner (ingen "delvis avhengighet" til PK).
  - Merk: Om en tabell er på 1NF og PK består av én kolonne, er tabellen også på 2NF!

Ok: avhengighet til hele PK.

lnr	lnavn	pnr	padresse	fradato	tildato	leie	enr	enavn
5	Hansen	6	Aveien 1	01-JUL-94	01-SEP-96	3500	4	Olsen
5	Hansen	1	Bveien 8	01-SEP-96	01-JAN-98	4400	12	Larsen
9	Persen	6	Aveien 1	01-SEP-96	01-SEP-97	3500	4	Olsen
9	Persen	2	Dveien 6	01-SEP-97	01-SEP-98	5500	6	Alfsen
9	Persen	5	Eveien 3	01-SEP-98	null	5000	6	Alfsen

Ikke ok: delvis avhengighet til PK.

# Gårdagens Kahoot

9 -Quiz En tabell er på 2NF dersom den er på 1NF og

3 of 4 < >



vi ikke har noen delvis avhengighet til PK



22



vi kun har en delvis avhengighet til PK



17



vi har en sammensatt primærnøkkel



18



vi ikke har noen primærnøkkel



5



No answer



12

## Andre normalform (2NF) – forts.

- Primærnøkkelen er sammensatt.
  - Vi skiller ut kolonnene som er avhengige av deler av primærnøkkelen og plasserer disse i nye tabeller:

Leietaker

lnr	lnavn
---	-----
5	Hansen
9	Persen

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

Prospekt\_Eier

pnr	padresse	leie	enr	enavn
---	-----	-----	---	-----
6	Aveien 1	3500	4	Olsen
1	Bveien 8	4400	12	Larsen
2	Dveien 6	5500	6	Alfsen
5	Eveien 3	5000	6	Alfsen

# Tredje normalform (3NF)

- En tabell er på 3NF dersom:
  - Den er på 2NF.
  - Og ingen ikke-primærnøkkel-attributt har en transitiv avhengighet til en primærnøkkel. (Ingen  $A \rightarrow B$  og  $B \rightarrow C \Rightarrow A \rightarrow C$ .)
- Alternativt kan vi si det slik: Alle ikke-primærnøkkel-attributter er kun funksjonelt avhengige av primærnøkkel.

pnr	adresse	leie	enr	enavn
6	Aveien	1 3500	4	Olsen
1	Bveien	8 4400	12	Larsen
2	Dveien	6 5500	6	Alfsen
5	Eveien	3 5000	6	Alfsen

Merk: Tabellene  
**Leietaker** og  
**Leiekontrakt** er  
allerede på 3NF.



## Tredje normalform (3NF) – forts.

- Vi fjerner transitivt avhengige attributter fra tabellen og plasserer de i en ny tabell.
- **Determinanten blir nøkkelen.**

Prospekt

pnr	adresse	leie	enr
---	-----	----	---
6	Aveien 1	3500	4
1	Bveien 8	4400	12
2	Dveien 6	5500	6
5	Eveien 3	5000	6

Eier

enr	enavn
---	-----
4	Olsen
12	Larsen
6	Alfsen

# 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 supernøkler.
  - Kan også sies som: Alle minimale determinanter er kandidatnøkler.
- Sagt mer folkelig:
  - BCNF = det kan ikke eksistere determinanter som kan ha samme innhold på flere rader i tabellen.

# 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.

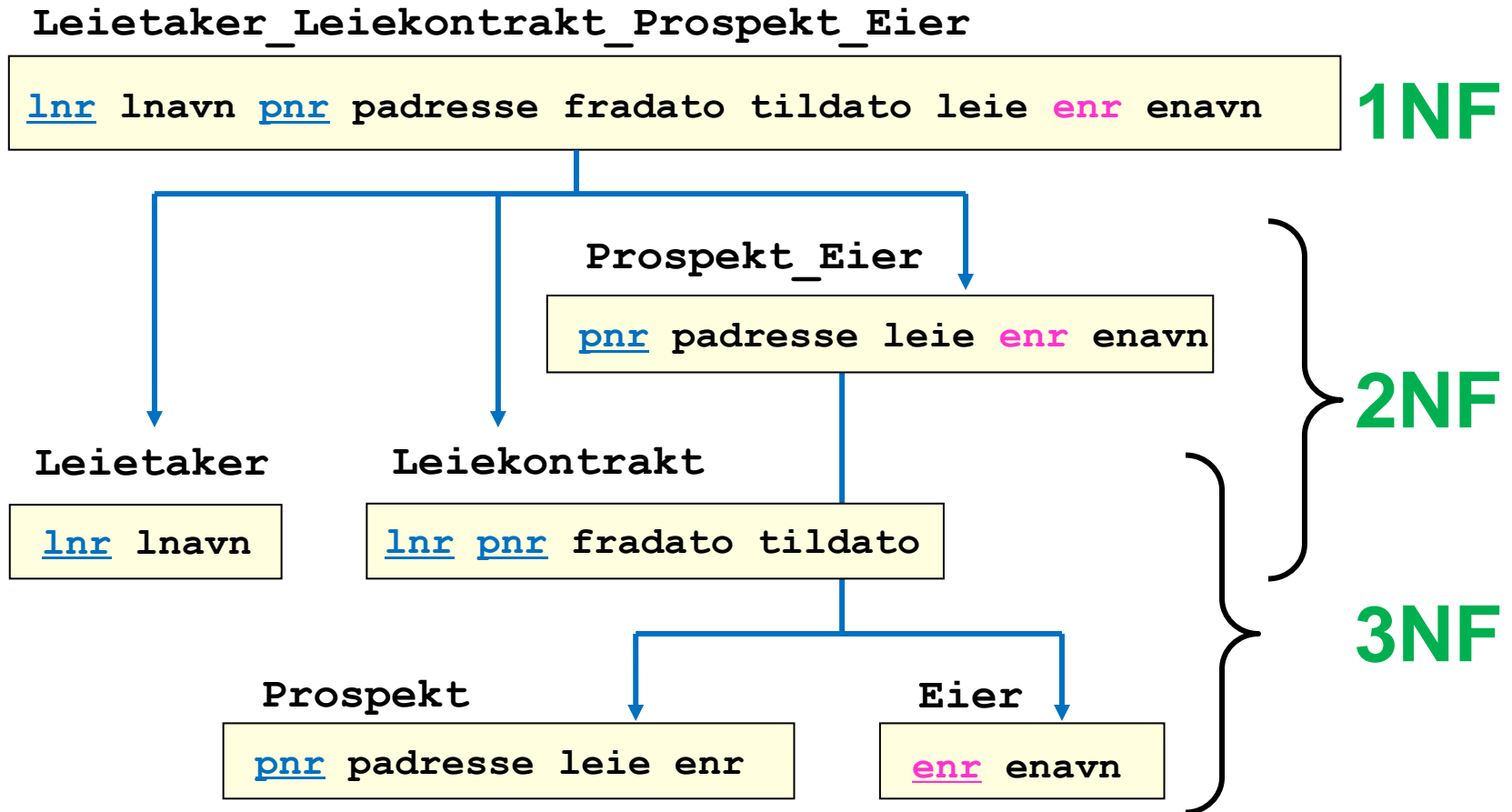
- Primærnøkkel

- 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

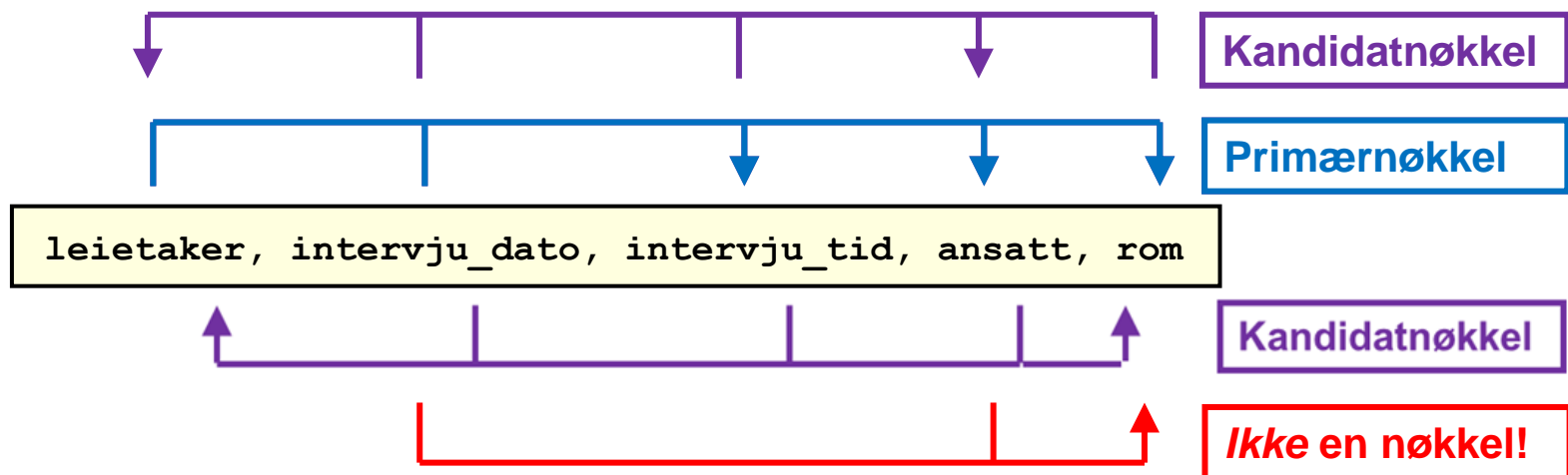
# En skikkelig nøtt!

- La oss gjøre et tappert forsøk...
- Hvilke 3 kandidatnøkler har vi her?

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

# leietaker\_intervju – forts.

- Oppgave: Hvilke 3 kandidatnøkler har vi her?



- Men finner du en determinant til?
  - Vi har PK, 2 kandidatnøkler, 1 ekstra avhengighet.

## leietaker\_intervju – forts.

- PK og de 2 kandidatnøkklene er ok i BCNF.
- Den 4. determinanten (intervju\_dato, ansatt → rom) er ikke godkjent i BCNF!  
Den er nemlig ikke en kandidatnøkkel...
- Vi må splitte tabellen i to for å oppnå BCNF.

# leietaker\_intervju – forts.

- Splitter tabellen i to nye:

Leietaker\_Intervju

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



Intervju

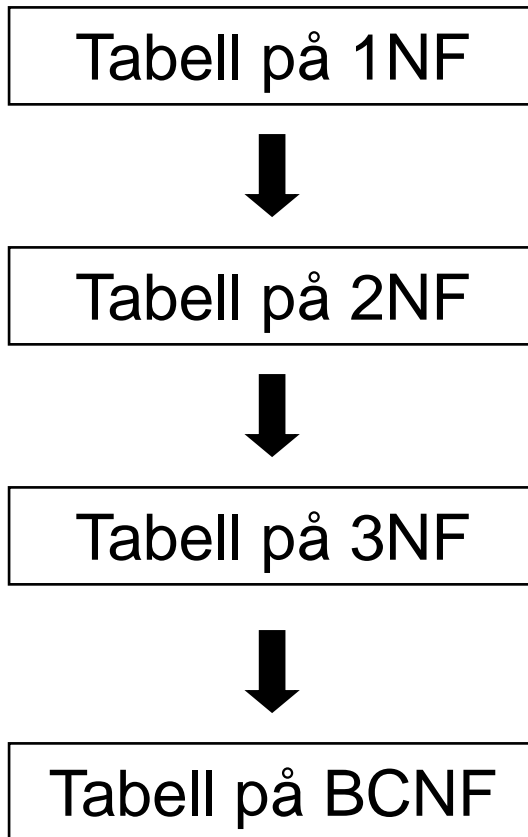
lnr	ivju_dato	ivju_tid	ansatt
76	13-MAY-11	10:30	5
56	13-MAY-11	12:00	5
74	13-MAY-11	12:00	7
56	01-JUL-11	10:30	5



Ansatt\_Rom

ivju_dato	ansatt	rom
13-MAY-11	5	101
13-MAY-11	7	102
01-JUL-11	5	102

# 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!

# Neste uke

- Diverse småtemaer som vi har igjen.
- Pensum til neste gang: Kap. 9.3, 10.2.1, 10.3.2, 11.2.2