



Økt 8 (av 12)

DB1100 Databaser

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

Dagens temaer

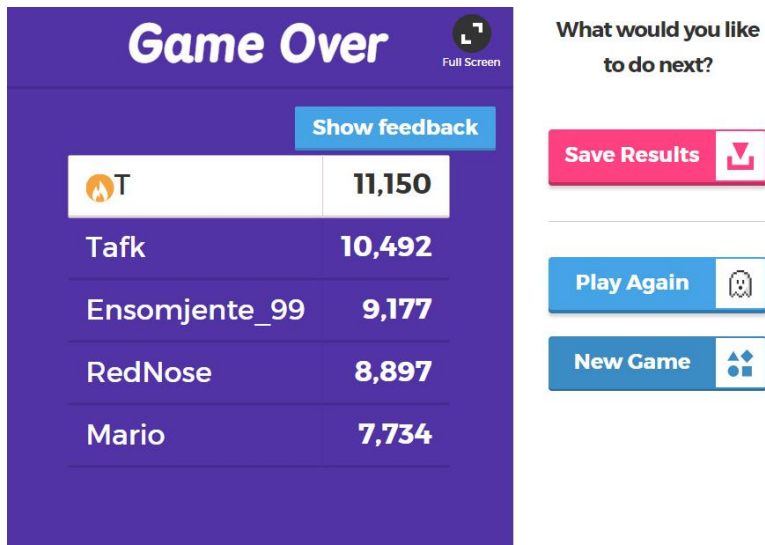
Dagens tema: [Modellering ER, del 2 \(av 2\)](#).

– Dagens pensum: (samme som forrige uke) [Læreboka, kapittel 7 & 8.1](#)

- Fra forrige økt: Resultater Kahoot
- Status LearnER
- Dagens innhold: Mer datamodellering med ER

Resultat, forrige Kahoot

- Forrige gang ble det ikke så fornuftige tall fra Kahoot, da den viste drøye 500 "deltakere" på det meste.
 - Vi har hatt problemer med nettet i auditoriet, men med all sannsynlighet var det én til noen få (sagt på DB-språk: "1:M" O_o) av dere som saboterte Kahoot-seansen sist, ved å lage 3-4 serier spambot'er á ca 100 bot'er per serie. :'-)
- Jeg kan vise topp-lista fra sist, den ble ikke ødelagt av botene som deltok:



The screenshot shows the 'Game Over' screen of a Kahoot! quiz. On the left, a purple sidebar contains the text 'Game Over' and a 'Full Screen' icon. Below this is a 'Show feedback' button. The main area displays a leaderboard with the following entries:

Rank	Player Name	Score
1	T	11,150
2	Tafk	10,492
3	Ensomjente_99	9,177
4	RedNose	8,897
5	Mario	7,734

On the right, there is a section titled 'What would you like to do next?' with three buttons: 'Save Results' (pink), 'Play Again' (blue), and 'New Game' (blue).

- La meg innrømme med en gang: Kahoot er **ikke** laget for å være hacking-sikkert!
 - Tvert imot, det er enkelt å spam'e Kahoot-bots, det krever overhodet ikke mad skillz.
 - (Ergo gir det heller ikke særlig bragging rights?)
 - Jeg håper de som hadde behov for å teste dette, ble ferdige sist gang, så vi kan ha vanlig Kahoot igjen i dag, og se om dere fortsatt slår fjorårets kull?! :-)

Status LearnER

- LearnER problemet vi opplevde sist, vedr. [innstillinger](#), [ble fikset](#) torsdag kveld i forrige uke! :-)
 - Så vidt jeg vet fungerer LearnER godt nå (som i "ingen kjente *alvorlige* bugs").
 - Mwn meld i fra om feil dere måtte finne (melding eller mail), så tar jeg de videre.
- NB, LearnER er litt [sær på ÉN ting](#):
 - Om det skal være [minimum 0 eller 1 på koplingene](#).
 - Om oppgaven ikke spesifiserer dette godt, og dere får feil på koplinger, prøv den andre varianten. (Prøv med 0 om dere har satt 1, og vice versa).
- Har du andre [utfordringer med LearnER?](#)
 - I så fall, i øvingstimene: [Spør meg](#), spør [veilederne](#) eller spør [hverandre](#)! :-)
- NB: En av dere etterspurte sist screenshots av LearnER løsningsforslag.
 - Så vidt jeg vet fungerer LearnER godt nå, derfor har jeg ikke laget dette.
 - Snakk med meg eller en veileder i øvingstimene om dette fortsatt er et problem.

Konseptuell (begrepsmessig) og logisk modell

- **Konseptuell (begrepsmessig) modell** beskrives i kapittel 7 i læreboka.
 - Dette er den enkleste modellen.
 - Den angir ikke fremmednøkler.
 - Når vi benytter enkel kråkefot eller UML notasjon angir den heller ikke attributter.
 - I LearnER begynner vi eventuelt med konseptuell modell før vi går til den neste:
- **Logisk modell** beskrives i kapittel 8.1 i læreboka.
 - Dette er en mer utfyllende modell.
 - Den har alltid med nøkler (PK og FK).
 - Den har alltid med attributter.
 - I LearnER er logisk modell alltid med.

ER diagram med Kråkefot og UML

- Vi kan vise modeller av databasen som [ER diagram](#).
 - Dette kan gjøres på litt forskjellige måter. Her vist med noen LearnER eksempler:

- **Konseptuell** modell

- Enkel Kråkefot

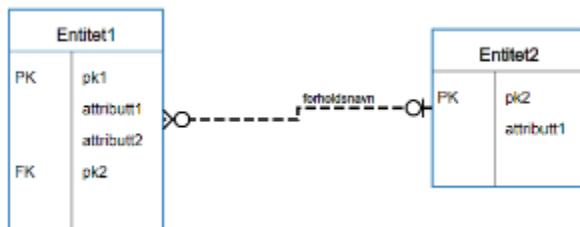


- UML

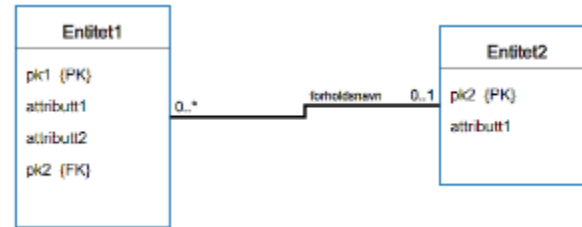


- **Logisk** modell:

- Kråkefot



- UML



Identifiserende og ikke-identifiserende forhold

- Når vi benytter **kråkefot** notasjon (*ikke* enkel kråkefot eller UML) kan vi i den **logiske modellen** tydeliggjøre svake entiteter og identifiserende relasjoner.
- **Svake entiteter:**
 - Eksistensen til forekomster av den svake entiteten avhenger av eksistensen til forekomsten av entiteten den er knyttet til.
 - Identifikatoren til en svak entitet er helt eller delvis arvet fra andre entiteter.
- **Identifiserende forhold:**
 - Relasjoner mot svake entiteter, som skal føre til arv av identifikator, kalles **identifiserende**, og tegnes som **heltrukne linjer**.
 - Alle andre forhold kalles **ikke-identifiserende** og tegnes som **stiplede linjer**.
- Eksempel fra læreboka: (*kap. 7.4*)
 - Saga Kino og salen Saga 2.

Eksempel case fra forrige økt: prosjektstyring

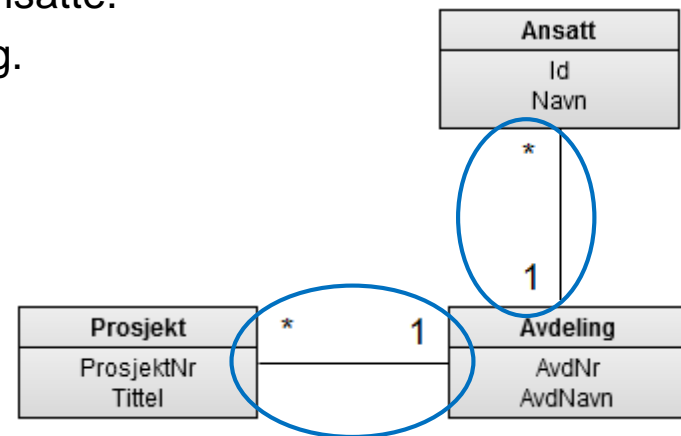
- Et firma ønsker å få oversikt over sine prosjekter. De har leid oss inn for å lage en databaseløsning som ordner dette.
- De ønsker spesifikt å få oversikt over følgende:
 - Hvilken **avdeling** (**nummer**, **navn**) eier hvert prosjekt?
 - Hvilke **prosjekter** (**nummer**, **tittel**) involverer hvilke ansatte?
 - Hvor mye **tid** benytter hver **ansatt** (**id**, **navn**) per prosjekt?
(NB: Denne siste er litt vanskelig å plassere på rett sted.)
- Modellerings spørsmål:
 - Hvilke entiteter (kommende tabeller) må vi ha?
Og hvilke attributter skal plasseres i entitetene?
 - Hva er relasjonene (koplingene) mellom entitetene?
(Hvordan hører de sammen: En til en? En til mange? Mange til mange?)

Entiteter og attributter (repetisjon fra sist)

- Ut fra spesifikasjonen kommer vi fram til et behov for følgende entiteter:
 - Avdeling
 - Prosjekt
 - Ansatt
- Videre trenger vi følgende attributter:
 - Avdeling: AvdNr, AvdNavn
 - Prosjekt: ProsjektNr, Tittel
 - Ansatt: Id, Navn
- Og så må **tid** plasseres inn et eller annet sted.

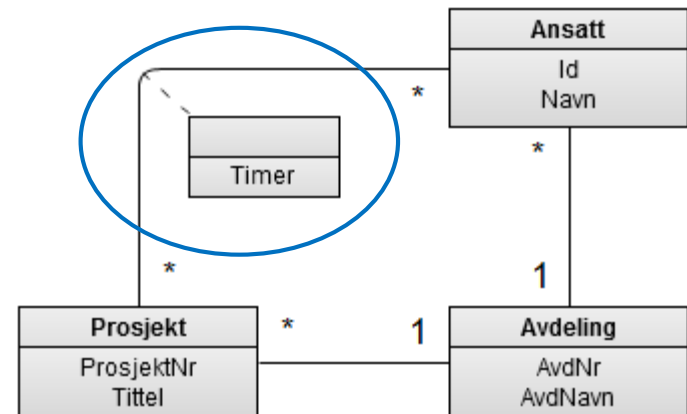
Relasjoner

- Relasjonene er ikke oppgitt i spesifikasjonen, men det er logisk(?) å anta relasjoner som oppgitt under.
 - (Vi bør tidlig i prosjektet få de bekreftet av kunden!)
- Relasjonen avdeling & ansatt:
 - En avdeling kan ha mange (symbol: *) ansatte.
 - En ansatt tilhører én (symbol: 1) avdeling.
- Relasjonen prosjekt & avdeling:
 - Et prosjekt tilhører én (symbol: 1) avdeling.
 - En avdeling kan ha mange (symbol: *) prosjekter.



Relasjoner – forts.

- Relasjonen prosjekt & ansatt:
 - Et prosjekt kan bemannes av mange (*) ansatte.
 - En ansatt kan jobbe parallelt på mange (*) prosjekter.
- Modellen begynner å falle på plass! :-)
- Men vi har én attributt igjen:
 - Vi trenger å vite ”tid per ansatt per prosjekt”.
 - Men putte denne hvor?
 - På relasjonen prosjekt & ansatt!



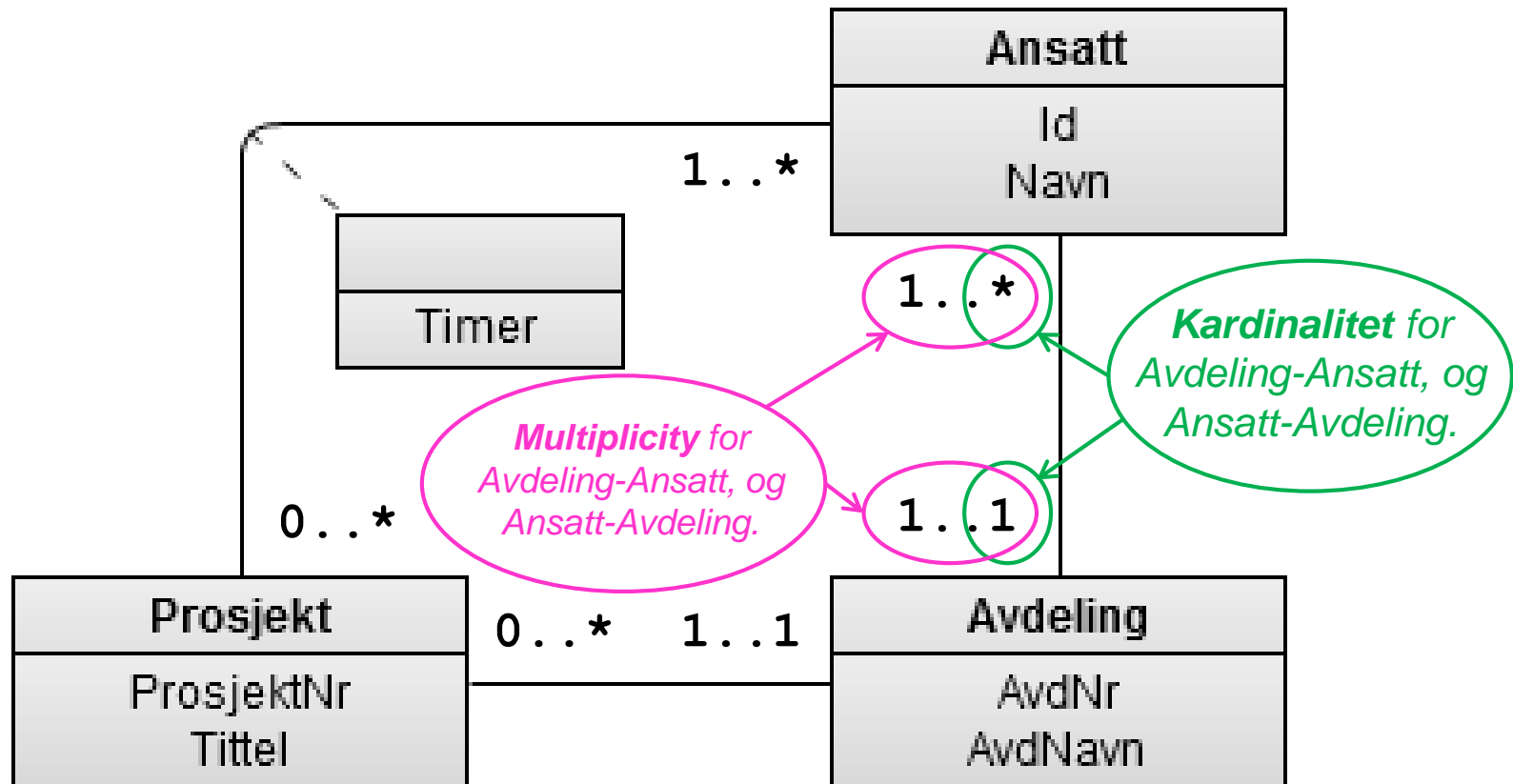
Koblingsentiteter

- Vi la inn et attributt på relasjonen mellom Ansatt og Prosjekt.
 - Relasjonen er i dette tilfellet et mange-til-mange-forhold: En ansatt kan delta i mange prosjekter. Et prosjekt kan ha mange prosjektdeltakere (ansatte).
- Når vi har et **mange-til-mange-forhold** introduserer vi en **koblingsentitet**, og **attributtene legges i denne entiteten**.
- Vi kaller overgangen fra mange-til-mange-forhold til koblingsentitet for *entitisering*.
- Kapittel 7.5 i boka forklarer dette godt! :-)

Relasjoner – nye begreper (pluss litt repetisjon)

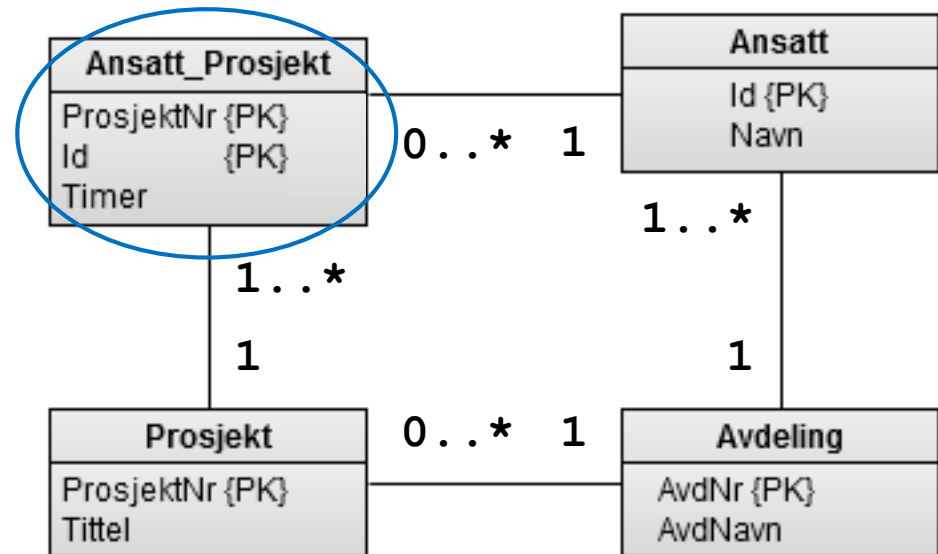
- I forrige økt snakket vi om **kardinalitet ifbm. ER-modellering**:
 - Kardinaliteten angir **det største antall koplinger en entitet kan ha**.
 - Eksempel: En ansatt kan maks tilhøre 1 avdeling. Kardinaliteten til ansatt i "ansatt-avdeling relasjonen" er derfor 1. (For avdeling i samme relasjon er den "**".)
- **Deltakelse**: Der kardinalitet angir det største antall koplinger en entitet kan ha, angir deltakelse **det minste antall koplinger en entitet kan ha**.
 - Eksempel: En ansatt kan minimum tilhøre 1 avdeling. Deltakelsen til ansatt i "ansatt-avdeling relasjonen" er derfor 1. (For avdeling i samme relasjon er den også 1.)
 - Deltakelse angis før kardinalitet. I UML: også med ".." i mellom de.
- **"Multiplicity"**: Deltakelse og kardinalitet sett sammen utgjør "multiplicity".
 - Eksempel: En ansatt kan ha fra 1 til 1 avdeling. (Og en avdeling kan ha fra 1 til mange ansatte.)
 - Multiplicity skrives i **UML**: "0..1", "1..1" (evt bare "1"), "0..*", "1..*".

EAR-skjema, ord/uttrykk



ER-skjema til database

- ER til database, punkt 1/6:
 - Markér PK for alle entiteter.
 - Fjern mange-til-mange relasjoner ved å lage nyekoplingsentiteter.
 - Koplingsentitetene skal inneholde PK fra de to entitetene de kopler sammen, samt eventuelle relasjonsattributter.



ER skjema til database – forts.

- ER til database, punkt 2/6:
 - Hver entitet blir en tabell med samme navn.
 - Attributtene til en entitet blir kolonner i tabellen.
 - I UML er det vanlig å angi primary key ved å sette {PK} bak navnet.
 - (Underline er en annen vanlig primary key notasjon.)

ER-skjema til database – forts.

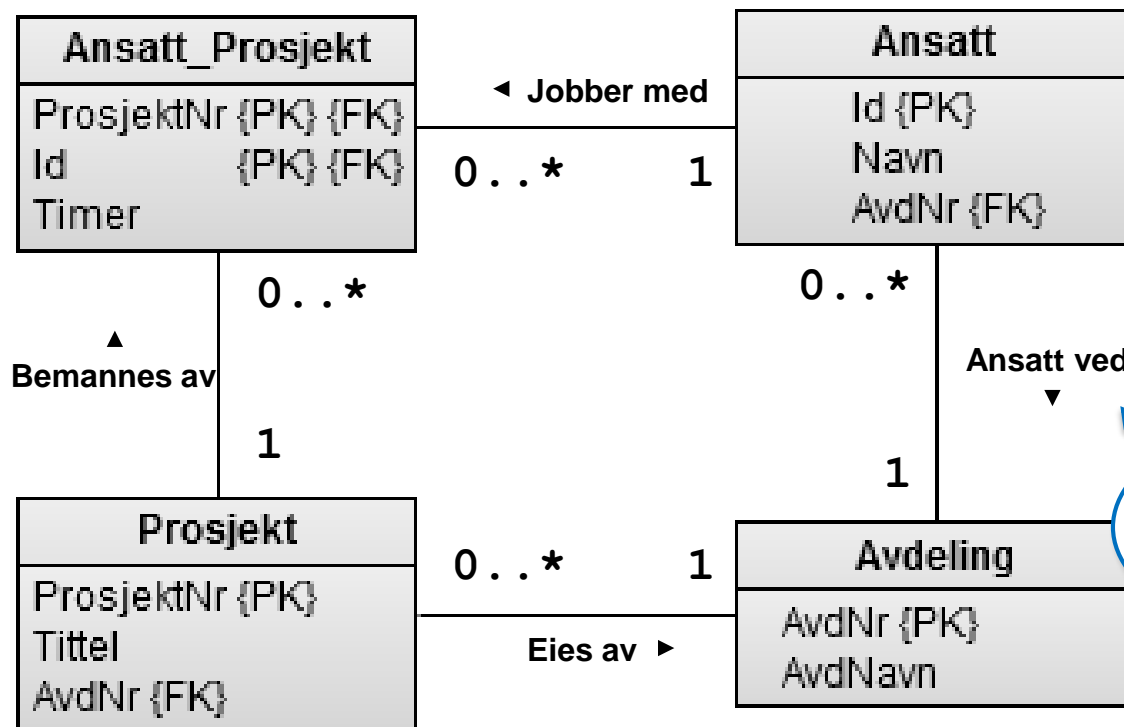
- ER til database, punkt 3/6:
 - Én til mange relasjoner realiseres ved at PK på én-siden kopieres til tabellen på mange-siden, og blir fremmednøkkel (FK) der.
 - UML notasjon, foreign key: {FK}.
 - (Annen vanlig notasjon: stiplet linje.)

ER-skjema til database – forts.

- Resten (punkt 4 til 6) ikke relevant for oss foreløpig, men nevner de allikevel:
- ER til database, punkt 4/6:
 - Én til én relasjoner kan realiseres på flere måter.
 - Sjekk kapittel 8.1.5 i boka.
- ER til database, punkt 5/6 & 6/6:
 - Normalisering og de-normalisering av databasen.
 - (Kommer på de neste to forelesningene.)

Ferdig databasemodell

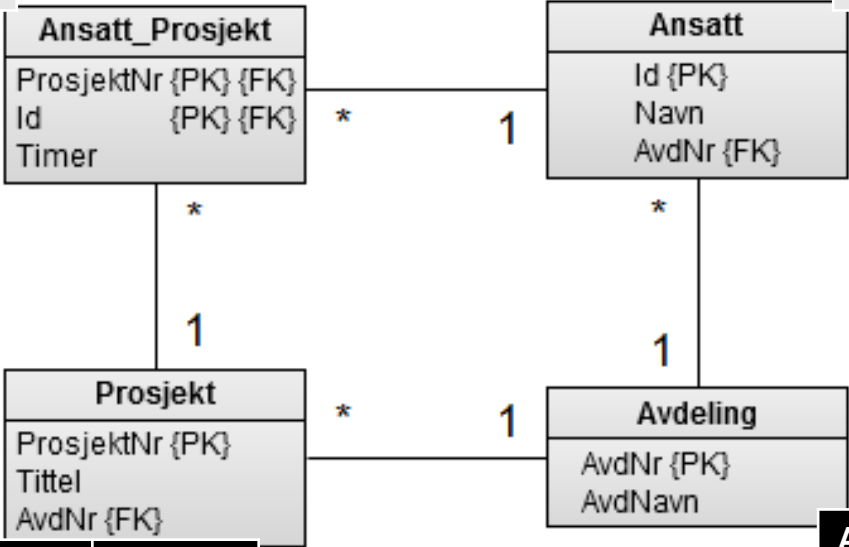
- Etter å ha gjort punkt 1 til 6 (i praksis 1 til 3) ser vår ferdige databasemodell ut som følger:



Navn på
relasjoner

ProsjektNr	Id	Timer
4	11	5
2	13	6
3	14	12
4	13	17

Id	Navn	AvdNr
11	Jon	111
12	Ida	113
13	Ole	111
14	Eli	114



ProsjektNr	Tittel	AvdNr
1	Grindgut	111
2	Apollo	113
3	Ulv	113
4	Follobanen	111

AvdNr	AvdNavn
111	Oslo
112	Bergen
113	Mandal
114	Bodø

Nøkler – alltid greit å repetere disse

- Disse introduserte vi sist:
 - **Supernøkkel**. En eller flere kolonner som danner en unik identifikator for en rad i en tabell. NB: Kan godt være flere kolonner enn det minste, unike utvalget.
 - **Kandidatnøkkel**. En supernøkkel som ikke kan reduseres til færre kolonner om den fortsatt skal være en unik identifikator.
- Og hvorfor skal vi lære disse?!? (Godt spørsmål!)
 - Vi trenger de når vi skal normalisere databasen (kommende forelesninger).
- Disse bør dere kunne godt nå:
 - **Primærnøkkel** (PK). Den kandidatnøkkel vi velger å bruke som vår unike identifikator for en rad i en tabell.
 - **Fremmednøkkel** (FK). Kolonne(r) som viser til primærnøkkel i en annen (evt. samme) tabell.

Koplingsentiteten

- Eksempel på hvordan innhold i koplingsentiteten Ansatt_Prosjekt kan se ut:

ProsjektNr (fra Prosjekt)	Id (fra Ansatt)	Timer
1001	1	12
1002	1	44
1002	2	20
1002	3	125
1003	2	25
1004	2	5
1004	3	10
1005	1	10
1006	1	20
1006	2	125

Surrogatnøkkel?

- På forrige slide danner attributtene ProsjektNr og Id sammen primær-nøkkelen.
- I stedet for å benytte disse to attributtene sammen, kunne vi lagt til et nytt attributt, en **surrogatnøkkel**:
 - Surrogatnøkkel forklares i pensumboka slik: **Et løpenummer uten «betydning» som kun brukes for entydig identifikasjon.** Surrogatnøkler blir som regel generert automatisk av systemet.
- Interessant!
 - Hva lønner seg å benytte?
 - En surrogat PK eller den «naturlige», sammensatte primærnøkkelen?
 - Finnes det et enkelt svar?
- Ikke noe fasitsvar, men les mer her:
 - <https://stackoverflow.com/questions/1029485/use-composite-keys-or-always-use-surrogate-keys>
 - <https://stackoverflow.com/questions/23850396/composite-vs-surrogate-keys-for-referential-integrity-in-6nf>

Ny oppgave om nøkler

- Hvilke supernøkler har vi her?
- Hvilke kandidatnøkler har vi her?

Code	Name	Continent	Region	SurfaceArea	IndepYear	Population	LifeExpectancy	GNP	GNPold	LocalName	GovernmentForm	HeadOfState	Capital	Code2
ABW	Aruba	North America	Caribbean	193.00	NULL	103000	78.4	828.00	793.00	Aruba	Nonmetropolitan T...	Beatrix	129	AW
AFG	Afghani...	Asia	Souther...	652090.00	1919	22720000	45.9	5976.00	NULL	Afganistan/A...	Islamic Emirate	Mohammad Omar	1	AF
AGO	Angola	Africa	Central ...	1246700.00	1975	12878000	38.3	6648.00	7984.00	Angola	Republic	JosÃ© Eduard...	56	AO
AIA	Anguilla	North America	Caribbean	96.00	NULL	8000	76.1	63.20	NULL	Anguilla	Dependent Territor...	Elisabeth II	62	AI
ALB	Albania	Europe	Souther...	28748.00	1912	3401200	71.6	3205.00	2500.00	ShqipÃ«ria	Republic	Rexhep Mejdani	34	AL
AND	Andorra	Europe	Souther...	468.00	1278	78000	83.5	1630.00	NULL	Andorra	Parliamentary Copri...		55	AD
ANT	Netherl...	North America	Caribbean	800.00	NULL	217000	74.7	1941.00	NULL	Nederlandse...	Nonmetropolitan T...	Beatrix	33	AN
ARE	United ...	Asia	Middle East	83600.00	1971	2441000	74.1	37966...	36846.00	Al-Imarat al-...	Emirate Federation	Zayid bin Sulta...	65	AE
ARG	Argentina	South Amer...	South A...	2780400.00	1816	37032000	75.1	34023...	323310.00	Argentina	Federal Republic	Fernando de la...	69	AR
ARM	Armenia	Asia	Middle East	29800.00	1991	3520000	66.4	1813.00	1627.00	Hajastan	Republic	Robert KotÃsar...	126	AM
ASM	America...	Oceania	Polynesia	199.00	NULL	68000	75.1	334.00	NULL	Amerika Sam...	US Territory	George W. Bush	54	AS
ATA	Antarctica	Antarctica	Antarctica	13120000.00	NULL	0	NULL	0.00	NULL	Ã-	Co-administrated		NULL	AQ
ATF	French ...	Antarctica	Antarctica	7780.00	NULL	0	NULL	0.00	NULL	Terres austr...	Nonmetropolitan T...	Jacques Chirac	NULL	TF
ATG	Antigua...	North America	Caribbean	442.00	1981	68000	70.5	612.00	584.00	Antigua and ...	Constitutional Mona...	Elisabeth II	63	AG
AUS	Australia	Oceania	Australia ...	7741220.00	1901	18886000	79.8	35118...	392911.00	Australia	Constitutional Mona...	Elisabeth II	135	AU
AUT	Austria	Europe	Western...	83859.00	1918	8091800	77.7	21186...	206025.00	Ã-sterreich	Federal Republic	Thomas Klestil	1523	AT
AZE	Azerbaijan	Asia	Middle East	86600.00	1991	7734000	62.9	4127.00	4100.00	AzÃrbaycan	Federal Republic	HeydÃr Ãliyev	144	AZ
BND	Burundi	Africa	Central ...	27834.00	1962	5595000	45.2	992.00	992.00	Burundi/Uburu	Republic	Pierre Nkurunziza	553	BT

Videre arbeid i dag (og kort om neste gang)

- Nå:
 - Lese kap. 7 & 8.1 i pensumboka. (Om du ikke gjorde det til sist.)
 - Se opp til 6 studentvideoer om modellering (linket nederst på Canvas). Finner du noen favoritter?!
 - Begynne på øvingsoppgavene.
- Etterpå:
 - Øving (flipped) 12:15 - 15:00, sjekk TimeEdit for rom.
 - Teori (flipped) 15:15 - 16:00, auditoriet.
- Neste gang: (torsdag om 1 uke)
 - Normalisering, del 1.
 - *Pensum til neste gang: Kapittel 8.2.1 - 8.2.4 (bare 8 sider!)*