# Mengde

VIKTIG EKSAMENSRELATERT!

En mengde er en samling av dataverdier som kan være enkeltverdier eller sammensatte verdier. Ingen av verdiene kan forekomme mer enn en gang i mengden.

**Kartesiske produktet**: mengden av ordnede par.

**Relasjon**: delmengden av det kartesiske produktet som er sann.

**Attributtets domene**: mengden av verdier et attributt kan anta.

Tabellformen er en måte å presentere relasjonen på.

*I dette eksempelet eksisterer det en database med ansattnummer og navn.*

**Kartesiske produktet**: Ansnr X Navn = {<1, Åse>, <2, Åse>, <3, Åse>, <1, Anne>, <2, Anne>, <3, Anne>

|  |  |
| --- | --- |
| **AnsNr** | **Navn** |
| 2 | Åse |
| 3 | Anne |

**Relasjon**: {<2, Åse>, <3, Anne>}

*Dette er to ulike måter å presentere sannheten på. Delmengde av det kartesiske produkt*

Egenskaper ved slike tabeller som ovenfor:

1. Duplikater eksisterer ikke.
2. Tuplene (radene) er uordnet.
3. Attributtene (kolonnene) er uordnet.
4. Verdiene er atomiske (relasjonen kalles *normalisert*). Hver «boks» kun en verdi.

Relasjonens *grad* = antall attributter (kolonner).

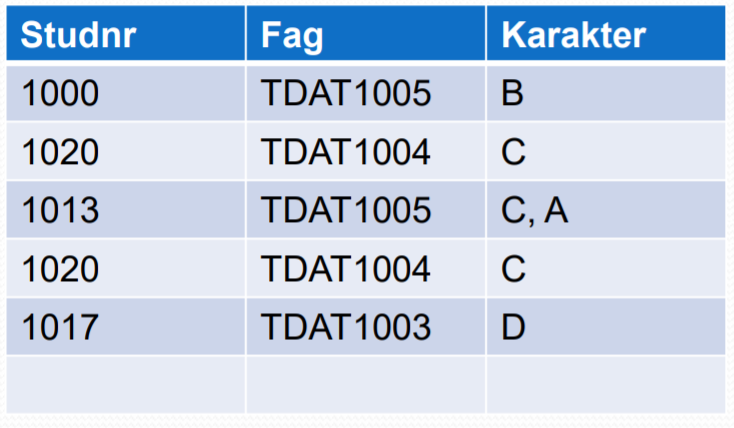
Relasjonens *kardinalitet* = antall tuppler (relasjonens tilstand, varierer over tid).

En *relasjonsdatabase* er en samling navngitte, normaliserte relasjoner (kun atomiske verdier).

**Denne tabellen oppfyller disse kravene:** B og C.

A oppfylles ikke siden rad 2 og 4 er like.

B oppfylles ikke siden karakter kolonnen i rad 3 har to verdier, altså den er ikke atomisk.



# Entitetsintegritet

Hva skal til for å identifisere en entitet i tabellen?

*Entitet:* et objekt uten operasjoner i objektorientert tankegang.

**På samme måte som entiteter (objekter) i den virkelige verden er identifiserbare, skal også den enkelte tuppel i en relasjon være identifiserbar. Ingen rader skal være like.**

**Nøkler**

*Supernøkler*: en mengde attributter som gjør en tuppel entydig.

* F.eks. id, fødslNr, ansattNr eller både ansattNr og navn eller flere kan kombineres.

*Kandidatnøkkel*: minimal supernøkkel.

* Id, fødslNr, ansattNr

*Primærnøkkel*: en kandidatnøkkel som vi velger til å identifisere en tuppel. Hele eller deler av denne kan ikke være NULL (altså at verdien mangler, det er ikke det samme som en tekststreng som består av blanke tegn, og heller ikke tallverdien 0).

*Alternative nøkler*: kandidatnøkler som ikke velges til primærnøkkel.

*Surrogatnøkkel*: godt alternativ til en informasjonsbærende primærnøkkel. Gjerne løpenummer generert av databasesystemet.

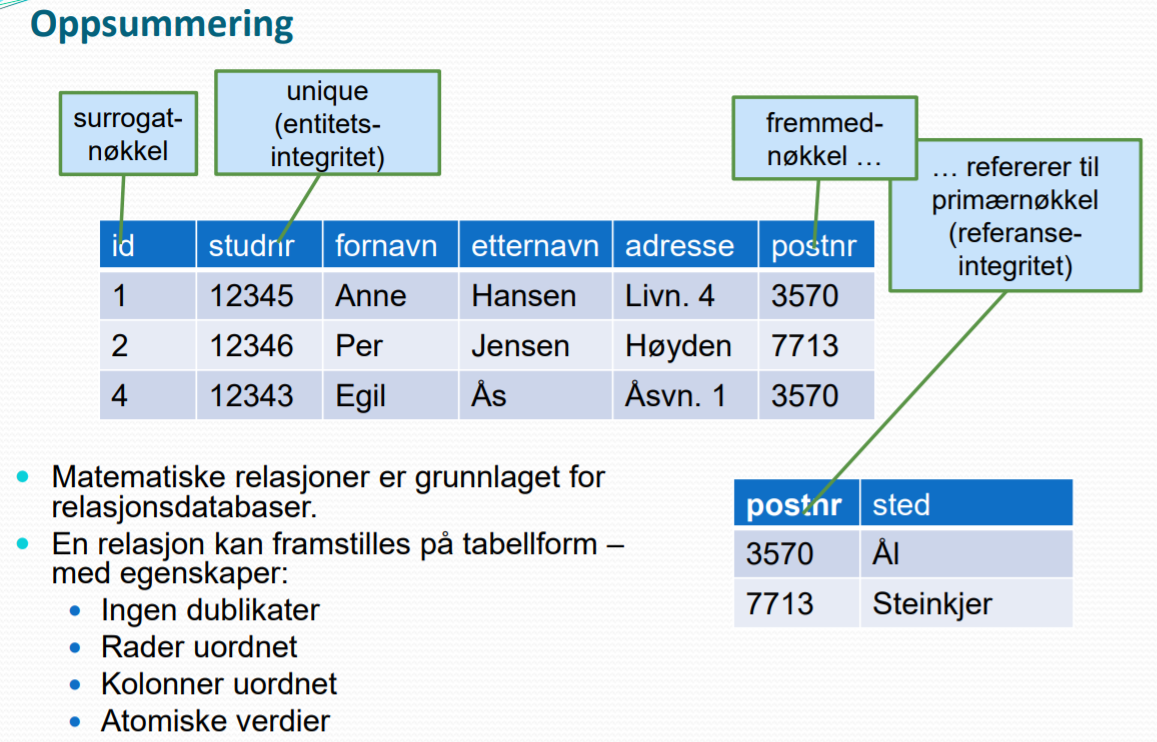
*Fremmednøkkel*: brukes til å koble sammen tabeller.

* Nøkkelen i R2 og primærnøkkelen i R1 må være definert på samme domene.
* Markeres med stjerne i tabeller.

Valg av *primærnøkkel*:

* Må være unik
* Unngå vanligvis tekster (samt navn og adresse).
* Tall fungerer godt som primærnøkler (fødslNr, studNr, automatisk generert løpenummer).

Entity Relationship Diagram (ER diagram) ligner på klassediagram, men er noe annet.



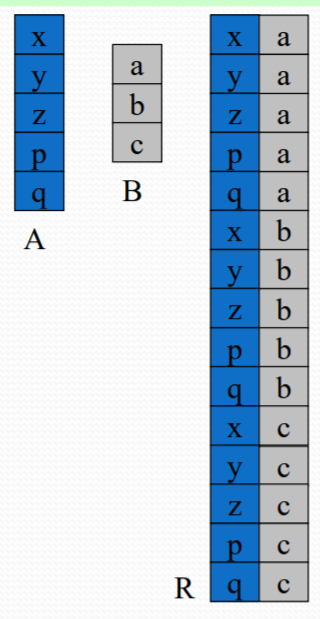
# Relasjonsalgebra

*Ett sett med operasjoner som utføres på relasjoner. Resultatet er alltid en ny relasjon.*

**Algebraoperasjonene**: seleksjon, projeksjon, produkt og forening.

**Seleksjon/restriksjon**: Begrenser antall rader. Filtre ut spesifikke tuppler.

**Projeksjon/reduksjon**: Begrenser antall kolonner.



Det kartesiske produktet / kryssproduktet er alle mulige kombinasjoner av tuppler fra to relasjoner A og B.

**Oppsummering**

