

# IN2090 – Databaser og datamodellering

## 02 – Relasjonsmodellen

Leif Harald Karlsen  
leifhka@ifi.uio.no

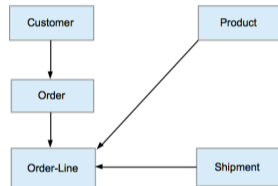


Universitetet i Oslo

# Litt historie

---

- ◆ Et av de første databasesystemet ble utformet av Bachman og Williams i 1964
- ◆ Ble solgt under navnet IDS (*Integrated Data Store*)
- ◆ IDS var en *nettverksdatabase* og var designet for bruk fra et programmeringsspråk
- ◆ I 1968 kom IBM med IMS (*Information Management System*), en *hierarkisk database* som forenkling av nettverksdatabaser
- ◆ IMS var nær enerådende for administrativ databehandling
- ◆ Fortsatt en del store firma som har legacy-systemer som bruker IMS

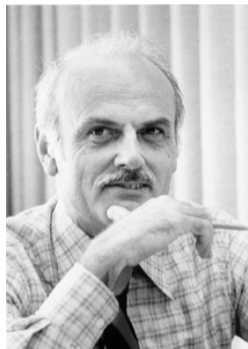


Eksempel  
nettverksdatabaseskjema

## Litt mer historie

---

- ◆ I 1970 presenterte E. F. Codd sin *relasjonsmodell*
- ◆ Dette var den teoretiske beskrivelsen av den da nye typen databaser kalt *relasjonsdatabaser*
- ◆ Relasjonsdatabaser er enkle å beskrive og bruke, men vanskelig å implementere
- ◆ Først i 1977 klarte Oracle å lage et databasesystem som fortjener betegnelsen *relasjonell*
- ◆ Relasjonsdatabaser er den klart mest brukte typen database



Edgar Frank "Ted" Codd

# Hva er relasjonsmodellen?

---

CustomerID (int)	Name (text)	Birthdate (date)	NrProducts (int)
0	Anna Consuma	1978-10-09	19
1	Peter Young	2009-03-01	1
2	Carla Smith	1986-06-14	8
3	Sam Penny	1961-01-09	14
4	John Mill	1989-11-16	8
5	Yvonne Potter	1971-04-12	6

- ◆ Forrige uke: Relasjonsdatabaser inneholder relasjoner/tabeller som består av kolonner og rader
- ◆ Relasjonsmodellen: formell/matematisk beskrivelse av denne måten å representere data
- ◆ Presis beskrivelse av:
  - ◆ hva en relasjon er (hva de består av, når er de like, osv.)
  - ◆ hvordan man beskriver data med relasjoner (struktur, identitet, referanser)
  - ◆ hvordan man kan hente ut og manipulere data representert med relasjoner
- ◆ Altså: komplett beskrivelse av hvordan bruke relasjoner for å beskrive data

# Relasjoner og relasjonsdatabaser

CustomerID (int)	Name (text)	Birthdate (date)	NrProducts (int)
0	Anna Consuma	1978-10-09	19
1	Peter Young	2009-03-01	1
2	Carla Smith	1986-06-14	8
3	Sam Penny	1961-01-09	14
4	John Mill	1989-11-16	8
5	Yvonne Potter	1971-04-12	6

- ◆ *Relasjon*: en signatur og en mengde tupler
- ◆ *Mengde*: Samling av unike elementer (uordnet)
  - ◆ F.eks.  $\{1, 2, 3\} = \{3, 1, 2\}$
- ◆ *Tuppel*: Liste med verdier (ordnet)
  - ◆ F.eks.  $(1, 2, 'Ole') \neq ('Ole', 1, 2)$
- ◆ *Relasjonsdatabase*: mengde med relasjoner med unikt navn

(Customers(CustomerID, Name, Birthdate, NrProducts),  
{(0, Anna Consuma, 1978-10-09, 19),  
(1, Peter Young, 2009-03-01, 1),  
(2, Carla Smith, 1986-06-14, 8),  
(3, Sam Penny, 1961-01-09, 14),  
(4, John Mill, 1989-11-16, 8),  
(5, Yvonne Potter, 1971-04-12, 6)})

# Signatur og skjema

---

Customers

CustomerID (int)	Name (text)	Birthdate (date)	NrProducts (int)
0	Anna Consuma	1978-10-09	19
1	Peter Young	2009-03-01	1
2	Carla Smith	1986-06-14	8
3	Sam Penny	1961-01-09	14
4	John Mill	1989-11-16	8
5	Yvonne Potter	1971-04-12	6

- ◆ En *relasjonssignatur* består av et navn og en mengde attributter/kolonner
- ◆ Skrives ofte

Customers(CustomerID (int), Name (text), Birthdate (date), NrProducts (int))

eller

Customers(CustomerID, Name, Birthdate, NrProducts)

- ◆ Med mer matematisk notasjon (hvor et par av A og B skriver A: B)

Customers: {CustomerID: int, Name: text, Birthdate: date, NrProducts: int}

- ◆ Inneholder også nøkler (mer om dette i neste video)
- ◆ En signatur gir deg det du trenger å vite for å bruke relasjonen
- ◆ *Relasjonsskjema*: mengde med relasjonssignaturer

# Attributt

---

Customers

CustomerID (int)	Name (text)	Birthdate (date)	NrProducts (int)
0	Anna Consuma	1978-10-09	19
1	Peter Young	2009-03-01	1
2	Carla Smith	1986-06-14	8
3	Sam Penny	1961-01-09	14
4	John Mill	1989-11-16	8
5	Yvonne Potter	1971-04-12	6

- ◆ Intuitivt: En kolonne
- ◆ Består av et navn og en type (av og til kalt domene)
- ◆ To attributter kan ikke ha samme navn (i samme relasjon)

Customers

CustomerID (int)	Name (text)	Birthdate (date)	NrProducts (int)
0	Anna Consuma	1978-10-09	19
1	Peter Young	2009-03-01	1
2	Carla Smith	1986-06-14	8
3	Sam Penny	1961-01-09	14
4	John Mill	1989-11-16	8
5	Yvonne Potter	1971-04-12	6

- ◆ Intuitivt: en rad
- ◆ Formelt: mengde med par av attributtnavn og verdier, f.eks.:  
`{CustomerID: '2', Name: 'Carla Smith', Birthdate: '1986-06-14', NrProducts: '8'}`
- ◆ Alså ikke helt likt det vi vanligvis kaller tuppel, f.eks.:  
`('2', 'Carla Smith', '1986-06-14', '8')`
- ◆ Vi kan bruke attributt-navn fremfor posisjon i tuplene! (F.eks. `t.Name`)
- ◆ Gitt en rekkefølge på attributtene kan vi droppe dem i tuplene
- ◆ Tuppel kalles av og til *instans*



Customers

CustomerID (int)	Name (text)	Birthdate (date)	NrProducts (int)
0	Anna Consuma	1978-10-09	19
1	Peter Young	2009-03-01	1
2	Carla Smith	1986-06-14	8
3	Sam Penny	1961-01-09	14
4	John Mill	1989-11-16	8
5	Yvonne Potter	1971-04-12	6

- ◆ Kun atomære verdier (med samme type som attributten/kolonnen)
- ◆ Atomær er et litt vagt begrep og avhenger av bruk
  - ◆ Kan brukes direkte, uten å måtte "tolkes" eller "pakkes ut" først
- ◆ **NULL**: Markerer manglende eller ukjent verdi

# Full matematisk representasjon

---

Customers

CustomerID (int)	Name (text)	Birthdate (date)	NrProducts (int)
0	Anna Consuma	1978-10-09	19
1	Peter Young	2009-03-01	1
2	Carla Smith	1986-06-14	8
3	Sam Penny	1961-01-09	14
4	John Mill	1989-11-16	8
5	Yvonne Potter	1971-04-12	6

---

```
( Customers: { CustomerID: int, Name: text, Birthdate: date, NrProducts: int },  
  { { CustomerId: '0', Name: 'Anna Consuma', Birthdate: '1978-10-09', NrProducts: '19' },  
    { CustomerId: '1', Name: 'Peter Young', Birthdate: '2009-03-01', NrProducts: '1' },  
    { CustomerId: '2', Name: 'Carla Smith', Birthdate: '1986-06-14', NrProducts: '8' },  
    { CustomerId: '3', Name: 'Sam Penny', Birthdate: '1961-01-09', NrProducts: '14' },  
    { CustomerId: '4', Name: 'John Mill', Birthdate: '1989-11-16', NrProducts: '8' },  
    { CustomerId: '5', Name: 'Yvonne Potter', Birthdate: '1971-04-12', NrProducts: '6' } } )
```

# Eksempler: Ugyldig relasjon

---

Person

Epost (text)	(text)	Nummer (text)	By	Gate (text)	Nummer (int)	Postnummer (int)
ola@online.no	Ola Norman	01028912345	Oslo	Gateveien	2	ukjent
ali@hmail.net	Ali Nilsen	05109354321	Bergen	Nedre gate	første	5432
kari@epost.no	Kari Mo	vet ikke	Oslo	Veigata	7	2341

# Eksempler: Like og ulike relasjoner

Student

Brukernavn (text)	Navn (text)	Program (text)
olano	Ola Norman	PROSA
alini	Ali Nilsen	DESIGN
karimo	Kari Mo	PROSA

?  
=

Student

Brukernavn (text)	Program (text)	Navn (text)
alini	DESIGN	Ali Nilsen
olano	PROSA	Ola Norman
karimo	PROSA	Kari Mo

?  
=

```
(  
(Student, {Brukernavn: text, Navn: text, Program: text}),  
{  
  {Brukernavn: 'olano', Navn: 'Ola Norman', Program: 'PROSA'},  
  {Brukernavn: 'alini', Navn: 'Ali Nilsen', Program: 'DESIGN'},  
  {Brukernavn: 'karimo', Navn: 'Kari Mo', Program: 'PROSA'}  
}  
)
```

```
(  
(Student, {Brukernavn: text, Program: text, Navn: text}),  
{  
  {Brukernavn: 'alini', Program: 'DESIGN', Navn: 'Ali Nilsen'},  
  {Brukernavn: 'olano', Program: 'PROSA', Navn: 'Ola Norman'},  
  {Brukernavn: 'karimo', Program: 'PROSA', Navn: 'Kari Mo'}  
}  
)
```

Takk for nå!

---

Neste video handler om nøkler.