

BASIS DATA

NORMALISASI

NORMALISASI

Normalisasi :

Teknik/pendekatan yang digunakan dalam membangun disain lojik database relasional melalui organisasi himpunan data dengan tingkat <u>ketergantungan fungsional dan keterkaita</u>n yang tinggi sedemikian sehingga menghasilkan struktur tabel yang normal.

Tujuan :

- Minimalisasi redundansi (pengulangan data)
- Memudahkan identifikasi entitas
- Mencegah terjadinya anomali

Beberapa bentuk normal (normal forms, NF) :

- ✓ 1NF, 2NF, 3NF, BCNF based on keys and functional dependencies
- ✓ 4NF, 5NF
 based on keys and multi-valued dependencies)

FIRST NORMAL FORM (1NF)

- Suatu relasi disebut memenuhi bentuk normal pertama (1NF) jika dan hanya jika setiap atribut dari relasi tersebut hanya memiliki nilai tunggal dan tidak ada pengulangan grup atribut dalam baris.
- Bentuk 1NF tidak boleh mengandung grup atribut yang berulang.
- Tujuan membentuk 1NF:
 - ✓ semantik tabel menjadi lebih eksplisit (say anything once).
 - ✓ semua operator aljabar relasional dapat diaplikasikan pada tabel.

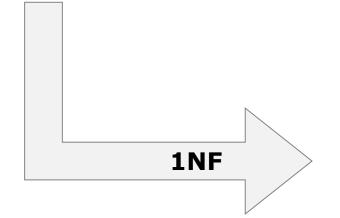
FIRST NORMAL FORM (1NF)

Tabel: Sales

IDSales	NamaSales	Telepon
ADN006	Yeni, SE	3517261, 3520165
ADN007	Memey	4744621,08122861427
ADN008	Tina	08566241521
ADN009	Ir. Yanto	7265122, 7123910
ADN010	Made	6723192

non-atomic

Unnormalized Not 1NF



IDSales	NamaSales	Telepon
ADN006	Yeni, SE	3517261
ADN006	Yeni, SE	3520165
ADN007	Memey	4744621
ADN007	Memey	08122861427
ADN008	Tina	08566241521
ADN009	Ir. Yanto	7265122
ADN009	Ir. Yanto	7123910
ADN010	Made	6723192

FIRST NORMAL FORM (1NF)

Unnormalized Not 1NF

Tabel: Buku repeated

ISBN	Thn_Terbit	ID_Pengarang	Nama_Pengarang	ID_Pengarang	Nama_Pengarang
12-1202-19222	1992	K0121	Aris M	K1021	Kosim P
11-1090-29101	2001	K1021	Kosim P		
11-1090-29102	2001	K2091	K Odelia	K0121	Aris M
12-1201-90871	2002	K2092	Renaldi	K2091	K Odelia
13-2089-12910	2001	K2019	Samsuri J		

1NF

ISBN	Thn_Terbit	ID_Pengarang	Nama_Pengarang
12-1202-19222	1992	K0121	Aris M
12-1202-19222	1992	K1021	Kosim P
11-1090-29101	2001	K1021	Kosim P
11-1090-29102	2001	K2091	K Odelia
11-1090-29102	2001	K0121	Aris M
12-1201-90871	2002	K2092	Renaldi
12-1201-90871	2002	K2091	K Odelia
13-2089-12910	2001	K2019	Samsuri J

- Suatu relasi disebut memenuhi bentuk normal kedua (2NF) jika dan hanya jika :
 - 1. memenuhi 1NF
 - 2. setiap atribut yang bukan kunci utama tergantung secara fungsional terhadap semua atribut kunci dan bukan hanya sebagian atribut kunci (fully functionally dependent).
- Untuk normalisasi ke bentuk 2NF, maka tabel 1NF didekomposisi menjadi beberapa tabel yang masing-masing memenuhi 2NF.
- Bila terdapat ketergantungan parsial maka: eliminate.
- Tujuan membentuk 2NF:
 - √ semantik tabel 2NF menjadi lebih eksplisit (fully FD)
 - √ mengurangi update anomali yang masih mungkin terjadi pada 1NF

Contoh:

```
Diketahui tabel R=(A,B,C,D,E) ; A,B kunci utama (primary key) dengan FD : A,B \rightarrow C,D,E maka tabel R memenuhi 2NF sebab : A,B \rightarrow C,D,E berarti : A,B \rightarrow C, A,B \rightarrow D dan A,B \rightarrow E
```

Jadi semua atribut bukan kunci utama tergantung penuh pada (A,B).

Contoh:

Bagaimana bila R = (A,B,C,D,E) tetapi dengan

 $FD: (A,B) \rightarrow (C,D)$ dan $B \rightarrow E$. Apakah memenuhi 2NF?

Jelas bahwa R bukan 2NF karena ada atribut E yang bergantung hanya pada atribut B saja dan bukan terhadap (A,B).

Dari FD : $(A,B) \rightarrow (C,D)$ juga mencerminkan bahwa hanya C dan D saja yang bergantung secara fungsional terhadap (A,B), tidak untuk E.

Jadi bukan 2NF.

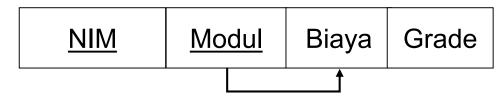
Untuk mengubah menjadi 2NF, lakukan dekomposisi menjadi :

R1 = (A,B,C,D) dan R2 = (B,E). Tampak R1 dan R2 memenuhi 2NF.

Contoh:

Diketahui Workshop = (NIM, Modul, Biaya, Grade)

Peserta Workshop



Tabel biaya peserta workshop

NIM	<u>Modul</u>	Biaya	Grade
P11.2004.0129	VB.Net	250000	Α
P11.2004.0130	Prolog	100000	Α
P11.2004.0129	Prolog	100000	В
P11.2004.0201	Delphi 6	150000	Α
P11.2004.0250	VB.Net	250000	В

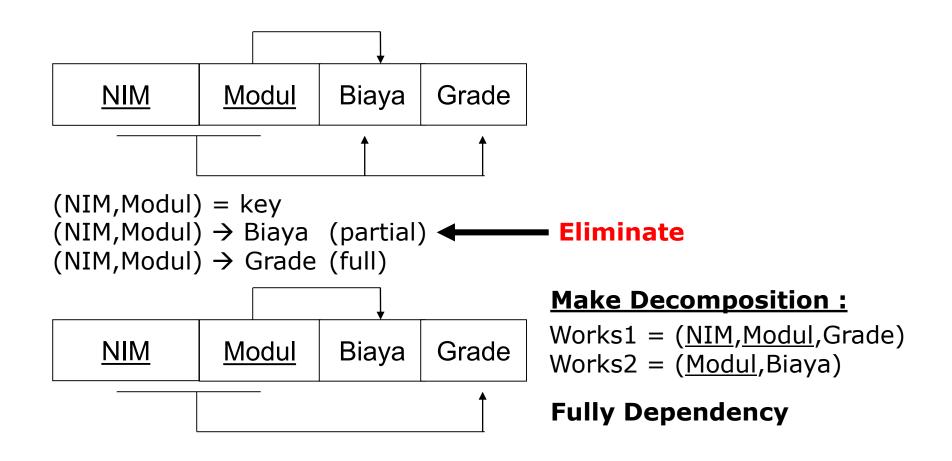
Key: NIM+Modul

FD : Modul → Biaya

(Biaya ditentukan oleh Modul yang

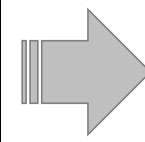
diambil mahasiswa)

- 1NF
- Not 2NF
 Sebab dalam tabel ini, biaya tidak bergantung penuh pada atribut kunci (NIM, Modul)



Workshop

NIM	Modul	Biaya	Grade
P11.2004.0129	VB.Net	250000	Α
P11.2004.0130	Prolog	100000	Α
P11.2004.0129	Prolog	100000	В
P11.2004.0201	Delphi 6	150000	Α
P11.2004.0250	VB.Net	250000	В



<u>NIM</u>	<u>Modul</u>	Grade
P11.2004.0129	VB.Net	Α
P11.2004.0130	Prolog	Α
P11.2004.0129	Prolog	В
P11.2004.0201	Delphi 6	Α
P11.2004.0250	VB.Net	В

Works1

	1
<u>Modul</u>	Biaya
VB.Net	250000
Prolog	100000
Delphi 6	150000

Works2

- Suatu relasi disebut memenuhi bentuk normal ketiga (3NF) jika dan hanya jika :
 - 1. memenuhi 2NF
 - 2. setiap atribut yang bukan kunci tidak tergantung secara fungsional terhadap atribut bukan kunci yang lain dalam relasi tsb (tidak terdapat ketergantungan transitif pada atribut bukan kunci).

Atau

- Suatu relasi disebut memenuhi bentuk normal ketiga (3NF) jika dan hanya jika setiap FD nontrivial : X → A, di mana X dan A atribut (atau kompositnya), memenuhi salah satu kondisi :
 - 1. X adalah superkey
 - 2. A merupakan anggota candidate key (A disebut prime attribute)

- Jika suatu relasi sudah memenuhi 2NF tapi tidak memenuhi 3 NF, maka untuk normalisasi ke bentuk 3NF, tabel 2NF didekomposisi menjadi beberapa tabel hingga masing-masing memenuhi 3NF.
- Tujuan membentuk 3NF:
 - semantik tabel 3NF menjadi lebih eksplisit (fully FD hanya pada primary key).
 - ✓ menghindari update anomali yang masih mungkin terjadi pada 2NF.

Catatan:

Jika suatu relasi memenuhi 2NF dan hanya memiliki tepat satu atribut yang bukan kunci utama maka relasi tsb memenuhi 3NF

Contoh:

Diketahui tabel R=(A,B,C,D,E); A,B kunci utama (primary key) dengan FD: A,B → C,D,E dan C → D,E maka R bukan 3NF sebab: Atribut D dan E (bukan kunci utama) bergantung secara fungsional pada C (yang juga bukan kunci utama).

Melalui FD :

 \checkmark Diketahui A,B \rightarrow C,D,E.

Karena sifat refleksif maka A,B \rightarrow A,B. Sehingga A,B \rightarrow A,B,C,D,E

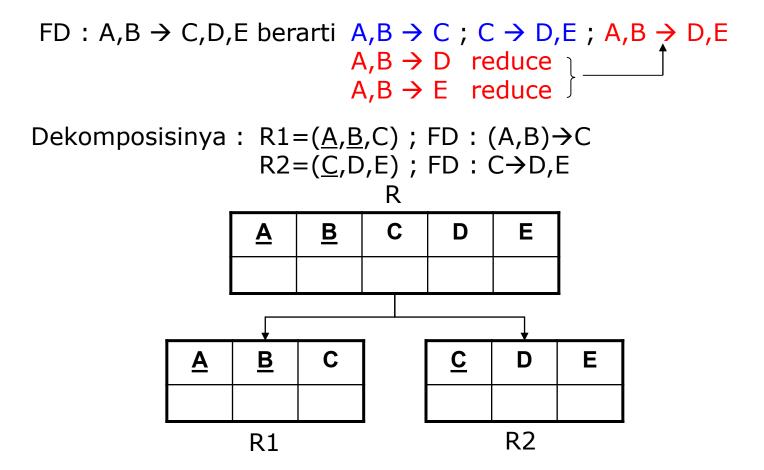
(A,B): Superkey.

 \checkmark Diketahui C \rightarrow D,E.

Karena sifat refleksif maka $C \rightarrow C$. Sehingga $C \rightarrow C$, D,E.

Karena $C \rightarrow A,B,C,D,E$ maka C bukan superkey.

- ✓ Tidak memenuhi definisi 3NF. Jadi R bukan 3NF.
- Agar R memenuhi 3NF maka didekomposisi menjadi : R1=(A,B,C) dan R2=(C,D,E) sehingga R1 dan R2 memenuhi 3NF.



Misal diketahui struktur informasi dari suatu dokumen supplier :

S	Status	City	F	Q.
			Р	Qty
S1	20	LONDON	P1	300
			P2	200
			Р3	400
			P4	200
			P5	100
			P6	100
S2	10	PARIS	P1	300
			P2	400
S 3	10	PARIS	P2	200
S 4	20	LONDON	P2	200
			P4	399
			P5	400

Akan dibentuk suatu tabel dengan skema $TPS=(\underline{S}, Status, City, \underline{P}, Qty)$ dengan (S,P) = primary key dan berlaku FD:

 $S \rightarrow Status$

 $S \rightarrow City$

City → Status

Lakukan normalisasi dari 1NF hingga 3NF.

TPS

<u>S</u>	Status	City	<u>P</u>	Qty
S 1	20	LONDON	P1	300
S1	20	LONDON	P2	200
S1	20	LONDON	P3	400
S1	20	LONDON	P4	200
S1	20	LONDON	P5	100
S1	20	LONDON	P6	100
S2	10	PARIS	P1	300
S2	10	PARIS	P2	400
S3	10	PARIS	P2	200
S4	20	LONDON	P2	200
S4	20	LONDON	P4	399
S4	20	LONDON	P5	400

- 1NF
- Not 2NF

Problem:

- Redundansi → inconsistency (low speed process)
- Anomaly : S→(Status, City) tapi kita tidak bisa insert data (S5,30,JAKARTA) tanpa diikuti data P (khususnya) dan Q. Menghapus 1 baris data akan jg merusak keutuhan informasi.
- Solusi : Dekomposisi menjadi : TPS1 dan TPS2

TPS1

<u>S</u>	Status	City
S1	20	LONDON
S2	10	PARIS
S3	10	PARIS
S4	20	LONDON

- 1NF
- 2NF
- Not 3NF (trans.)S → CityCity → Status
- Sekarang kita dapat menambah data (S5,30,JAKARTA) secara aman
- Tapi masih ada anomaly: Karena City → Status maka kita tidak bisa entry data City baru sebelum Status punya nilai. Penghapusan 1 baris sebagian data City juga bisa merusak keutuhan informasi S.
- Selain itu, masih ada redundansi pada Status dan City

TPS2

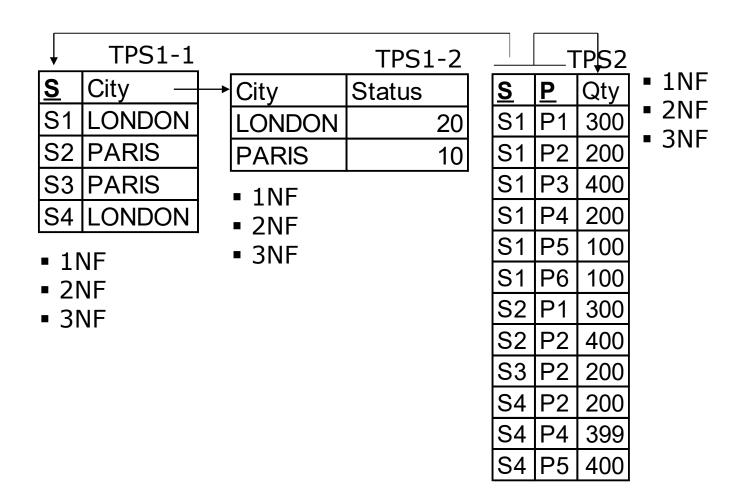
<u>S</u>	<u>P</u>	Qty
S1	P1	300
S1	P2	200
S1	P3	400
S1	P4	200
S1	P5	100
S1	P6	100
S2	P1	300
S2	P2	400
S3	P2	200
S4	P2	200
S4	P4	399
S4	P5	400

- 1NF
- 2NF
- 3NF

Redundansi partial

- → tidak potensial
- → lebih baik dari redundan sebelumnya

Tidak mungkin menghilangkan semua redundan tapi buat yang minimal



1. Diberikan skema relasi R = (A,B,C,D,E,F,G,H,I,J,K) dengan ketergantungan fungsional:

$$A \rightarrow B,C,D;C \rightarrow D;E \rightarrow F;A,E \rightarrow G,H,I,J,K;I \rightarrow J,K$$

Apakah R memenuhi 3NF? Jika tidak, rancanglah skema relasi R sedemikian sehingga memenuhi bentuk 3NF. Bila Saudara melakukan dekomposisi tabel, lengkapi dengan uji dekomposisi dan uji lossless.

- 2. Diketahui R = $(\underline{A},\underline{B},C,D,E,F,G,H)$ dimana (A,B) : primary key
 - Ketergantungan fungsional yang berlaku (FD):

$$A \rightarrow C,F ; B \rightarrow G,H ; A,B \rightarrow D,E dan D \rightarrow E$$

- a. Jika diketahui bahwa R memenuhi 1NF, apakah R memenuhi 2NF? 3NF?
- b. Jika tidak, rancanglah skema relasi R sedemikian sehingga memenuhi bentuk 2NF dan 3NF.
- c. Lengkapi dengan uji dekomposisi dan uji lossless.

- Suatu relasi disebut memenuhi BCNF jika dan hanya jika setiap determinan yang ada pada relasi tersebut adalah candidate key.
- Definisi yang lain :
 Suatu relasi disebut memenuhi BCNF jika untuk setiap FD nontrivial :
 X → A atribut X adalah superkey.
- Untuk normalisasi ke bentuk BCNF, maka tabel 3NF didekomposisi menjadi beberapa tabel yang masing-masing memenuhi BCNF.
- Tujuan membentuk BCNF :
 - semantik multiple candidate key menjadi lebih eksplisit (FD hanya pada candidate key).
 - menghindari update anomali yang masih mungkin terjadi pada 3NF.

Dari definisi 3NF dan BCNF, maka apabila suatu relasi memenuhi BCNF pasti memenuhi 3NF, tetapi belum tentu sebaliknya.

Contoh:

```
Diketahui tabel R=(A,B,C)
dengan FD : A \rightarrow B dan B \rightarrow C maka R bukan BCNF, sebab :
A superkey ?
  A \rightarrow B (diketahui)
  A \rightarrow B dan B \rightarrow C maka A \rightarrow C (transitif)
  A \rightarrow A (refleksif)
   Sehingga A \rightarrow (A,B,C) atau A \rightarrow R. Jadi A superkey.
B superkey ?
  B→C (diketahui)
  B \rightarrow B (refleksif)
   Tapi B\rightarrowA. Sehingga B\rightarrowA,B,C atau B bukan superkey.
Agar R memenuhi BCNF maka didekomposisi menjadi :
R1=(A,B); FD: A \rightarrow B dan
 R2=(B,C); FD: B \rightarrow C.
sehingga R1 dan R2 masing-masing memenuhi BCNF. Sebab A dan B dua-duanya sekarang
 menjadi superkey.
```

Contoh:

Diketahui tabel R=(A,B,C) dengan FD : $AB \rightarrow C$ dan $C \rightarrow B$. Apakah :

- 3NF?
- BCNF?

Jawab:

R memenuhi 3NF karena :

 $AB \rightarrow C$; maka $AB \rightarrow ABC$, atau $A \rightarrow R$. Jadi AB superkey dari R.

 $C \rightarrow B$; maka $AC \rightarrow AB$, atau $AC \rightarrow ABC$ dan $AC \rightarrow R$. Jadi AC juga superkey (sekaligus juga candidate key) dari R

Karena AB superkey dan C subset candidate key maka R memenuhi 3NF

R bukan BCNF karena :
AB avasaskas tatani C kultura ava

AB superkey tetapi C bukan superkey.

Students

sid	name	age
53666	Jones	18
53668	Smith	18
53669	Melissa	17
53670	Hilden	19

Students=(sid, name, age)

FD : sid \rightarrow name, age

• BCNF, sebab sid superkey

Pinjam

idpinjam	sid	bid	date
P-01	53666	B002	10/11/2005
P-02	53668	B001	10/11/2005
P-03	53668	B004	11/12/2005
P-04	53670	B002	14/11/2005

Books

bid	title	year
B001	MySQL	2002
B002	Algorithm	2003
B003	Visual Foxpro 6.0	2003
B004	Visual basic 6.0	2005

Books=(bid, title, year)

FD : bid \rightarrow title, year

• BCNF, sebab bid superkey

Pinjam=(idpinjam, sid, bid, date)

FD : idpinjam \rightarrow bid, date

 Bukan BCNF, sebab idpinjam bukan superkey idpinjam → sid

idpinjam	sid	bid	date
P-01	53666	B002	10/11/2005
P-02	53668	B001	10/11/2005
P-03	53668	B004	11/12/2005
P-04	53670	B002	14/11/2005

Didekomposisi menjadi:

Pinjam1

idpinjam	sid
P-01	53666
P-02	53668
P-03	53668
P-04	53670

FD trivial

→ BCNF

Pinjam2

idpinjam	bid	date
P-01	B002	10/11/2005
P-02	B001	10/11/2005
P-03	B004	11/12/2005
P-04	B002	14/11/2005

idpinjam → bid, date idpinjam superkey → BCNF

PERBANDINGAN ANTARA BCNF DAN 3NF

- Itu selalu untuk mengurai suatu relasi ke dalam relasi di 3NF dan
 - dekomposisinya adalah lossless
 - dependensi dipertahankan
- Itu selalu untuk mengurai suatu relasi ke dalam relasi di BCNF dan
 - dekomposisinya adalah lossless
 - tidak mungkin mempertahankan ketergantungan.

PERBANDINGAN ANTARA BCNF DAN 3NF

Contoh kasus redundansi pada 3NF

Jadwal = (Nim, Modul, Dosen)

 $FD = \{Dosen \rightarrow Modul\}$

Relasi ini memenuhi 3NF, karena tidak ada ketergantungan transitif.

Tetapi tidak memenuhi BCNF karena dari Dosen → Modul maka Dosen

bukan candidate key.

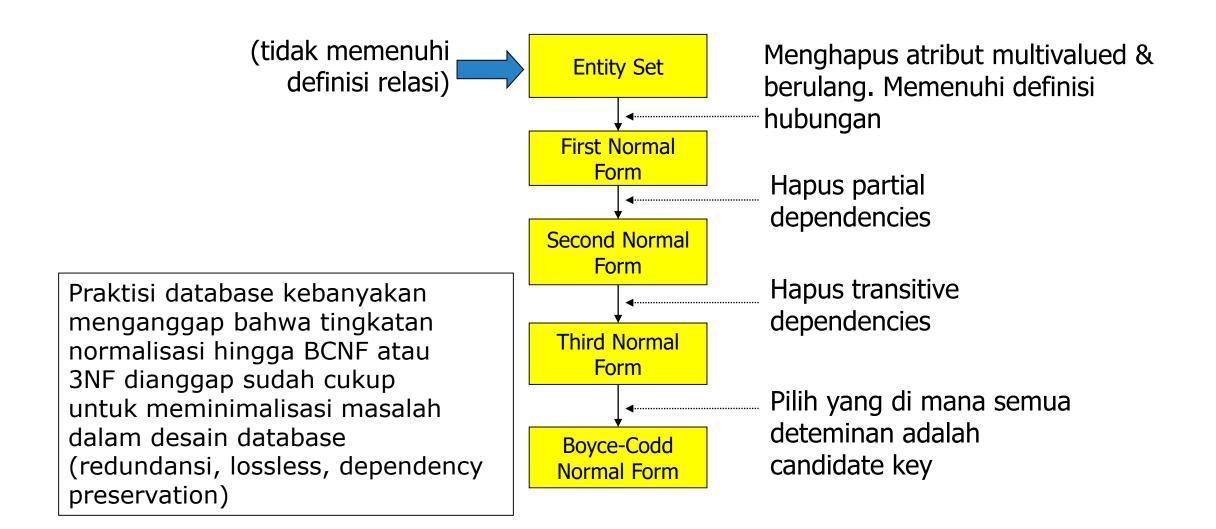
Alternatif yang dilakukan adalah dekomposisi tabel menjadi :

NIM	Modul	Dosen		NIM	Dosen	Dosen	Modul
P11.2004.0129	VB.Net	Ajib		P11.2004.0129	Ajib	Ajib	VB.Net
P11.2004.0130	Prolog	Aris		P11.2004.0130	Aris	Aris	Prolog
P11.2004.0129	VB Net	Ajib		P11.2004.0129	Ajib	Jono	Prolog
P11.2004.0201	VB Net	Budi	7	P11.2004.0201	Budi	Budi	VB.Net
P11.2004.0250	Prolog	Jono		P11.2004.0250	Jono		
P11.2004.0260	VB.Net	Budi		P11.2004.0260	Budi		
	NOT BC	NF			ВС	NF	

TUJUAN DESAIN

- Tujuan untuk desain basis data relasional adalah:
 - **BCNF.**
 - Lossless join.
 - Dependency preservation.
- Jika tidak dapat mencapai ini, menerima salah satunya
 - Kurangnya Dependency preservation
 - Redundansi karena penggunaan 3NF

TAHAP DESAIN



TANTANGAN PERANCANGAN BASIS DATA

Desainer harus membuat kompromi desain yang dipicu oleh tujuan yang saling bertentangan:

- standar desain (desain keanggunan atau kesetiaan),
 untuk mengembangkan desain "baik": Tanpa Cadangan, Tidak Berlebihan, Pelestarian Ketergantungan
- kecepatan pemrosesan (kinerja), dan

persyaratan pengguna tepat waktu

- kecepatan pemrosesan tinggi adalah prioritas "teratas" (efisiensi)
- meminimalkan jumlah dan kompleksitas hubungan atau tabel
- persyaratan informasi
 kemampuan untuk mengirimkan semua permintaan dan pelaporan yang ditentukan

Contoh bentuk kompromi yang populer:

→ Denormalisasi (pelanggaran normalisasi)

DENORMALISASI

Design Standards Vs (proceessing speed, information requirements)

 Normalisasi hanyalah merupakan teknik pendekatan yang digunakan untuk mendapatkan desain database (lojik) yang baik dan bukan sebagai "aturan baku" DBMS yang harus digunakan.

Bersifat "Normatif", memungkinkan untuk dilanggar dengan alasan : Kecepatan Proses (Efisiensi) dan Pelayanan Informasi Tepat Waktu

- Bentuk-bentuk Denormalisasi :
 - Membuat Atribut Turunan pada Tabel, mis: Cost=Qty*Price
 - Atribut yang Berlebihan, mis: NIM mahasiswa sudah mencerminkan program studi mahasiswa, tetapi dalam tabel mahasiswa dibuat atribut Program Studi.
 - Summary Table (mis: summary table for report)
 - Membiarkan relasi transitif dalam satu tabel untuk kemudahan proses
- Konsekuensi Denormalisasi :
 Redundancy, Not Atomic, Worst Space dll

LATIHAN

Berdasarkan formulir tersebut,

- Rancanglah tabel penyimpanan datanya
- Lakukan normalisasi hingga 3NF atau
 BCNF



NOTA PELANGGARAN SOPIR

Nomor	:	 	 	 	 	 	

Alamat				
Kota		Propinsi		Kodepos
No. SIM		Sex	Tanggal Lahir	TB BB
KENDARAAN No. Kendaraan	Warna 	Tahun	Tipe	No. Lambung
Supervisor Area				
_okasi				
Catatan Pelanggar	'an			
Findakan Peringa	atan.	Tidak diijinkan	mengemudi kendal	′aa∩ selama 3 hari.
Tindakan Peringa	atan.		mengemudi kendal	raan selama 3 hari.
Kembal	atan. Ii ke pangkalan.		- Semar	raan selama 3 hari. ang,// as Pemantau

LATIHAN

- 1. Berikan alasan anda mengapa suata basis data harus berada pada kondisi normal!
- 2. Kondisi apa saja yang harus dipenuhi, bahwa suatu basis data di katakan memenuhi normalisasi!
- 3. Diberikan suatu data sebagai berikut di bawah ini. Dari data yang diberikan belum memenuhi kondisi normal, maka tentukanlah proses normalisasi untuk menghasilkan data yang memenuhi kondisi normal!

No	Class	Time	Time Day Teacher Start		Start	Room	Remark
1	B.1I	17.00-18.30	Tuesday; Thursday	Ms. Avi; Ms. Oki	19/08/2004	A202	Run
2	B.5I	15.30-17.00	Tuesday; Friday	Ms. Beta; Ms. Susi	20/08/2004	A302	Run
3	B.1J	17.00-18.30	Monday; Thursday	Ms. Galuh; Ms. Avi	23/08/2004	A301	Run
4	B.2J	17.00-18.30	Tuesday; Thursday	Mr. Aris; Ms. Beta	24/08/2004	A102	Run
5	B.3J	15.30-17.00	Tuesday; Thursday	Mr. Aris; Ms. Oki	05/08/2004	A103	Run
6	B.2F	15.30-17.00	Monday; Thursday	Ms. Galuh; Mr. Hery	19/08/2004	A203	Run
7	B.1I	18.30-20.00	Wednesday; Friday	Ms. Ria; Ms. Galuh	04/08/2004	A203	Pending