

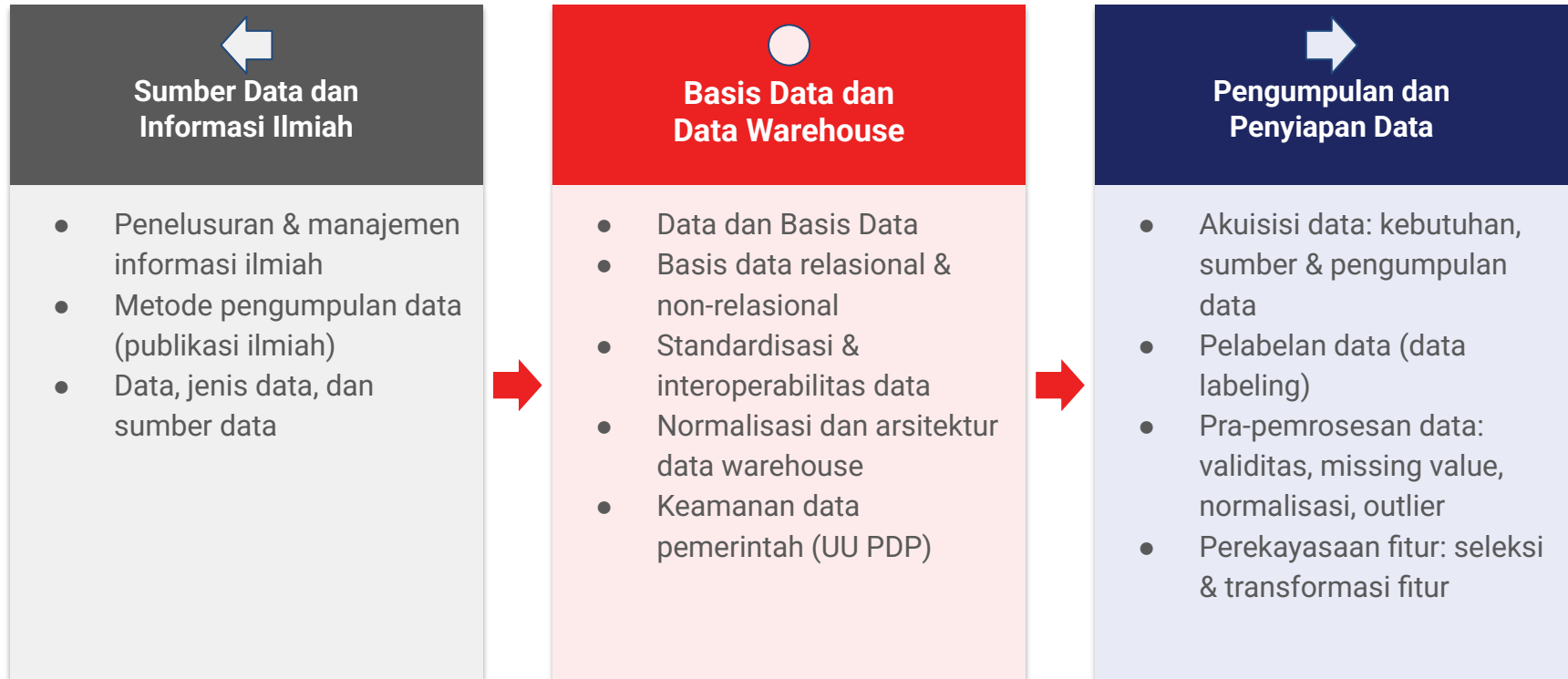
Database dan Data Warehouse

Pelatihan Penguatan Teknis Substantif Jabatan Fungsional Analis Data Ilmiah (JF ADI)
Pegawai Lembaga Kebijakan Pengadaan Barang/Jasa Pemerintah (LKPP)

Andria Arisal

Online, 24 Juni 2026

Posisi Materi dalam Program Pelatihan



Struktur basis data yang baik (sesi ini) menentukan bagaimana data dikumpulkan dan disiapkan (sesi selanjutnya).

Tujuan dan Indikator Hasil Belajar

Hasil Belajar: Peserta mampu menjelaskan konsep dasar basis data, standardisasi data, keamanan data, dan data warehouse secara proporsional.

Setelah mengikuti sesi ini, peserta mampu:

1. Menjelaskan arsitektur basis data relasional dan non-relasional.
2. Menguraikan prinsip standardisasi data dan interoperabilitas.
3. Menerapkan prinsip normalisasi untuk merancang struktur tabel yang baik.
4. Menjelaskan arsitektur data warehouse dan perannya dalam analitik pemerintahan.
5. Menjelaskan aspek keamanan data dan mitigasi kebocoran informasi.



Studi kasus sepanjang sesi ini menggunakan (asumsi) data pengadaan barang/jasa pemerintah (LKPP)

Apa itu Data?

Bagian 1

Pengantar Sistem Basis Data

Konsep dasar, RDBMS vs NoSQL, bahasa SQL (DDL/DML/DCL), serta keamanan akses dasar.

Apa itu Sistem Basis Data?

Apa itu Sistem Basis Data?

Definisi Inti

Basis data:

- kumpulan data terstruktur yang saling berelasi, disimpan dan dikelola secara sistematis.

DBMS (Database Management System):

- perangkat lunak untuk membuat, mengelola, dan mengamankan basis data.
- Contoh DBMS: PostgreSQL, MySQL, Oracle, SQL Server, MongoDB.

Mengapa Penting bagi LKPP

- Mengelola data penyedia, paket pengadaan, tender, dan kontrak secara konsisten.
- Menjadi sumber kebenaran tunggal (single source of truth) untuk audit dan pelaporan.
- Mendukung sistem SPSE, e-katalog, dan integrasi dengan instansi lain.
- ...

Basis Data Relasional (RDBMS)

Karakteristik

- Data disimpan dalam tabel (baris & kolom) yang saling berelasi melalui key.
- Menjamin sifat ACID (Atomicity, Consistency, Isolation, Durability)
- Skema (struktur tabel) bersifat tetap (fixed schema) dan harus didefinisikan di awal.
- Diakses menggunakan bahasa SQL (Structured Query Language).

Contoh Penggunaan di LKPP

- Tabel penyedia, paket_pengadaan, tender, kontrak yang saling terhubung melalui foreign key.
- Transaksi pencatatan kontrak harus konsisten — nilai kontrak tidak boleh tercatat ganda.
- Cocok untuk data transaksional pengadaan yang memerlukan integritas tinggi.

ACID (Atomicity, Consistency, Isolation, Durability)

Basis Data Non-Relasional (NoSQL)

Document Store

Menyimpan data sebagai dokumen JSON/BSON — fleksibel untuk data semi-terstruktur, mis. log aktivitas portal SPSE.

contoh: MongoDB, CouchDB

Key-Value Store

Penyimpanan cepat berbasis kunci-nilai — cocok untuk cache sesi pengguna aplikasi pengadaan.

contoh: Redis, DynamoDB

Column-Family Store

Optimal untuk data dalam volume besar dan kecepatan tulis tinggi, mis. log transaksi e-purchasing.

contoh: Cassandra, HBase.

Graph Database

Memodelkan relasi kompleks — mis. jejaring afiliasi antar-penyedia untuk deteksi indikasi persekongkolan tender.

contoh: Neo4j

BASE (Basically Available, Soft state, Eventually consistent)

Perbandingan DBMS

Fitur	RDBMS	NoSQL
Model Data	Terstruktur, <i>Schema on Write</i>	Fleksibel, <i>Schema on Read</i>
Skalabilitas	Vertikal	Horizontal
Konsistensi	ACID	BASE
Bahasa	SQL	Bervariasi, API
Fleksibilitas	Rendah	Tinggi
Kompleksitas	Cocok untuk data terstruktur dan relasi kompleks	Cocok untuk data tidak terstruktur/semi-terstruktur

Memilih Jenis Basis Data yang Tepat

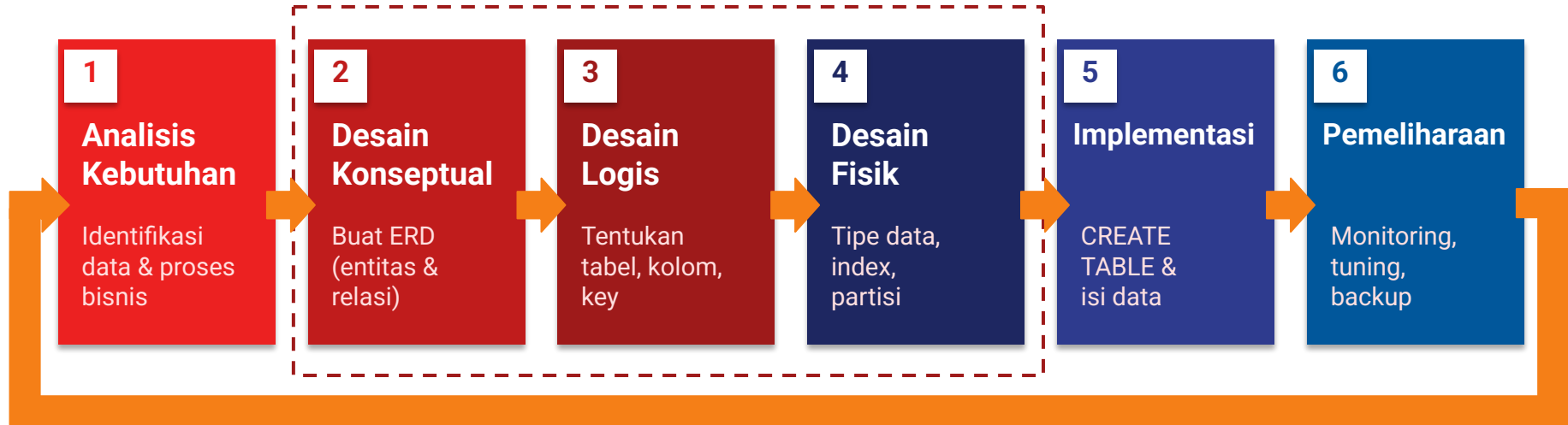
Kebutuhan	Jenis Basis Data	Alasan
Transaksi kontrak & pembayaran	Relasional (RDBMS)	Memerlukan integritas data & konsistensi tinggi (ACID)
Log akses portal SPSE	Document / Key-Value	Volume besar, struktur fleksibel, kecepatan tulis tinggi
Deteksi afiliasi penyedia	Graph Database	Fokus pada relasi antar-entitas, bukan tabel datar
Laporan & analitik historis	Data Warehouse (relasional)	Memerlukan agregasi dan query analitik kompleks

Bagian 2

Struktur dan Desain Basis Data

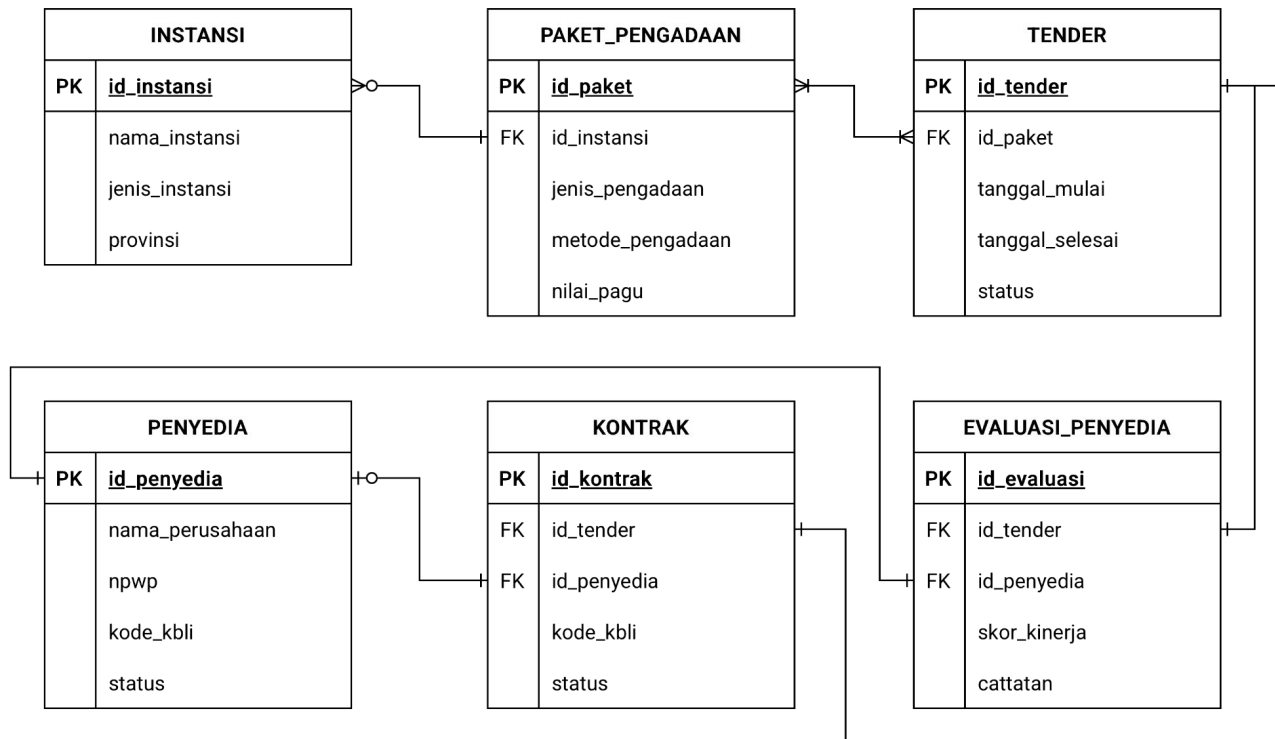
Siklus desain basis data, pemodelan ERD, serta tabel master vs transaksi dan jenis-jenis kunci.

Siklus Desain Basis Data



⚠ Setiap perubahan kebutuhan bisnis pengadaan (misal aturan e-purchasing baru) akan memutar siklus ini kembali ke tahap analisis.

Contoh Skema Basis Data: Pengadaan Barang/Jasa Pemerintah



asumsi

Tabel Master vs Transaksi, dan Jenis Kunci

Tabel Master vs Transaksi

- Tabel Master:
data relatif statis dan menjadi rujukan, mis. instansi, penyedia, ref_kbli.
- Tabel Transaksi:
mencatat kejadian/peristiwa yang terus bertambah, mis. tender, kontrak, evaluasi_penyedia.
- Tabel transaksi umumnya berisi foreign key yang merujuk ke tabel master.

Jenis-Jenis Kunci (Key)

- Primary Key: pengidentifikasi unik suatu baris (id_kontrak).
- Foreign Key: rujukan ke primary key tabel lain (id_penyedia pada tabel kontrak).
- Candidate Key: kolom yang berpotensi jadi primary key (npwp juga unik).
- Composite Key: kombinasi beberapa kolom sebagai key gabungan.
- Surrogate Key: key buatan sistem (auto-increment) tanpa makna bisnis.

Standardisasi dan Interoperabilitas Data

Prinsip standardisasi data pemerintah, data induk dan kode referensi, constraint, serta interoperabilitas antar-sistem.

Standardisasi Data: Definisi dan Regulasi

Definisi (DAMA International, 2015)

- Standardisasi data: proses menyepakati format, struktur, dan makna data agar konsisten lintas sistem dan organisasi.
- Mencakup: format penulisan, satuan ukur, kode referensi, dan metadata.

Landasan Regulasi

- Peraturan BPS No. 4 Tahun 2020: Petunjuk Teknis Standar Data Statistik.
- Perpres No. 39 Tahun 2019: Satu Data Indonesia — data pemerintah harus memenuhi 4 prinsip:
 - satu standar data,
 - metadata baku,
 - interoperable,
 - mengacu kode referensi/data induk.
- Loshin (Master Data Management, 2010): kualitas data dimulai dari standardisasi definisi.

Akibat Tanpa Standar: Data Penyedia di 3 Sistem

Sistem	Identitas Penyedia	Kode Wilayah	Status
SPSE	NPWP: 01.234.567.8-409.000	Kab. Bandung	Aktif (teks)
e-Katalog	NPWP: 012345678409000	Bandung	1 (kode angka)
SIKaP	NPWP: 01.234.567.8409.000	BANDUNG, JABAR	true (boolean)

Data Induk dan Kode Referensi

Data Induk (Master Data)

- Data inti yang dijadikan acuan bersama oleh banyak sistem, mis. data instansi, data wilayah, data penyedia terverifikasi.
- Dikelola terpusat agar tidak terjadi duplikasi/inkonsistensi antar-sistem pengadaan.

Kode Referensi

- KBLI (Klasifikasi Baku Lapangan Usaha Indonesia): mengklasifikasikan bidang usaha penyedia.
- Kode wilayah administrasi (Kemendagri): provinsi, kabupaten/kota.
- Kode jenis pengadaan & metode pengadaan (Barang/Jasa Konstruksi/Jasa Konsultansi/Jasa Lainnya; Tender/Seleksi/Pengadaan Langsung/E-purchasing).

Interoperabilitas Data: Empat Lapisan

Teknis	Protokol API, format data (JSON/XML), jaringan, autentikasi.	Sistem dapat terhubung dan berkomunikasi
Semantik	Definisi data bersama, kode referensi, ontologi (mis. KBLI).	Data bermakna sama di seluruh instansi
Organisasi	Peran, proses, dan tanggung jawab kepemilikan data yang disepakati.	Semua pihak tahu siapa pemilik dan pengelola data
Hukum / Keamanan	Hak akses, dasar hukum pemrosesan, perlindungan data (UU 27/2022).	Pertukaran data sah, dapat diaudit, dan aman



Permenkominfo No. 1/2023 mewajibkan interoperabilitas teknis dan semantik untuk layanan data SPBE.

Bagian 4

Komunikasi dengan Data

Metode komunikasi dengan data dalam basis data

Constraint SQL untuk Menjaga Standar Data

Constraint memaksa data mengikuti aturan/standar pada level basis data:

NOT NULL — kolom wajib diisi, mis. nama_perusahaan tidak boleh kosong.

UNIQUE — nilai tidak boleh berulang, mis. npwp harus unik per penyedia.

PRIMARY KEY — identitas unik tiap baris, mis. id_kontrak.

FOREIGN KEY — menjaga referensi valid antar-tabel, mis. id_penyedia pada kontrak harus ada di tabel penyedia.

CHECK — nilai harus sesuai aturan, mis. CHECK (nilai_kontrak > 0).

DEFAULT / ENUM — nilai baku, mis. status DEFAULT 'Proses' atau
ENUM ('Tender', 'Seleksi', 'Pengadaan Langsung', 'E-purchasing').

SQL: DDL, DML, dan DCL

DDL (Data Definition Language) – mendefinisikan struktur tabel:

```
CREATE TABLE penyedia (id_penyedia VARCHAR(10) PRIMARY KEY, nama_perusahaan VARCHAR(100) NOT NULL, npwp VARCHAR(20) UNIQUE, kode_kbli VARCHAR(6));
```

DML (Data Manipulation Language) – mengelola isi data:

```
INSERT INTO penyedia VALUES ('P01', 'PT Mitra Karya', '01.234.xxx', '4690');  
UPDATE penyedia SET status='Aktif' WHERE id_penyedia='P01';
```

DQL (Data Query Language) – mengambil isi data:

```
SELECT p.id_penyedia, p.nama_perusahaan, k.tahun FROM penyedia p JOIN kontrak k USING (id_penyedia) WHERE p.status='Aktif' AND YEAR(k.tanggal_kontrak)=2026;
```

DCL (Data Control Language) – mengatur hak akses:

```
GRANT SELECT ON kontrak TO auditor_lkpp; REVOKE DELETE ON kontrak FROM staff_input;
```

 **best practice: gunakan DCL untuk membatasi siapa yang dapat menghapus data kontrak – bukan hanya siapa yang dapat membacanya.**

Latihan 1: Merancang Tabel Penyedia

Skenario

Tim pengadaan di instansi Anda ingin menyimpan data penyedia yang mendaftar di sistem informasi pengadaan. Setiap penyedia memiliki NPWP unik, nama perusahaan, kode KBLI, dan status (Aktif/Non-aktif/Diblokir).

Tugas Anda:

1. Tentukan kolom apa saja yang dibutuhkan untuk tabel penyedia.
2. Tentukan primary key dan constraint (NOT NULL, UNIQUE) yang sesuai.
3. Tulis pernyataan CREATE TABLE untuk tabel tersebut.
4. Diskusikan: data apa yang tidak boleh diakses oleh seluruh staf (lihat keamanan data)?

Latihan 2: Temukan Masalah Kualitas Data

Tinjau contoh data penyedia berikut, dari tiga sistem berbeda di lingkungan pengadaan. Masalah apa yang Anda temukan?

NPWP	Nama Perusahaan	Kode KBLI	Provinsi	Status
01.234.xxx	PT Mitra Karya	4690	Jawa Barat	Aktif
01.234.xxx	PT. Mitra Karya	46900	Jabar	aktif
02.567.xxx	CV Sumber Jaya	41011	DKI Jakarta	A
02567000	Sumber Jaya CV	41.011	Jakarta	1

Refleksi:

1. Masalah kualitas data apa yang Anda lihat? (konsistensi, format, duplikasi)
2. Elemen mana yang perlu distandardisasi?
3. Field mana yang termasuk data pribadi menurut UU No. 27/2022 (PDP)?
4. Bagaimana Anda akan memperbaiki masalah ini?

Database Optimization

Normalisasi: Mengatasi anomali dan menerapkan bentuk normal (1NF–3NF).

Indexing: Mempercepat pencarian

Mengapa Normalisasi? Anomali pada Tabel Tak Ternormalisasi

Tabel kontrak_penyedia (belum dinormalisasi):

id_kontrak	id_penyedia	nama_perusahaan	npwp	nilai_kontrak
K001	P01	PT Mitra Karya	01.234.xxx	450,000,000
K002	P01	PT Mitra Karya	01.234.xxx	120,000,000
K003	P02	CV Sumber Jaya	02.567.xxx	85,000,000

Anomali Insert

Tidak bisa menambah penyedia baru tanpa ada kontraknya.

Anomali Update

Jika nama PT Mitra Karya berubah, harus diubah di banyak baris (K001 & K002).

Anomali Delete

Jika kontrak K003 dihapus, seluruh data CV Sumber Jaya ikut hilang.

Mengapa Normalisasi? Anomali pada Tabel Tak Ternormalisasi

Tabel kontrak_penyedia (belum dinormalisasi):

id_kontrak	id_penyedia	nama_perusahaan	npwp	nilai_kontrak
K001	P01	PT Mitra Karya	01.234.xxx	450,000,000
K002	P01	PT Mitra Karya	01.234.xxx	120,000,000
K003	P02	CV Sumber Jaya	02.567.xxx	85,000,000

Anomali Insert

Tidak bisa menambah penyedia baru tanpa ada kontraknya.

Anomali Update

Jika nama PT Mitra Karya berubah, harus diubah di banyak baris (K001 & K002).

Anomali Delete

Jika kontrak K003 dihapus, seluruh data CV Sumber Jaya ikut hilang.

Proses Normalisasi: Dari 1NF ke 3NF (bisa sampai 4NF, 5NF, ...)



contoh:

Hasil 3NF untuk

kontrak_penyedia: memisahkan tabel penyedia dan tabel kontrak

penyedia(id_penyedia, nama_perusahaan, npwp)

kontrak (id_kontrak, id_penyedia, nilai_kontrak)

> menghilangkan ketiga anomali sekaligus.

Latihan 3: Normalisasi Tabel Paket Pengadaan

Skenario

Tabel `paket_pengadaan_lengkap` menyimpan:

`id_paket`, `nama_paket`, `id_instansi`, `nama_instansi`, `alamat_instansi`, `id_penyedia_pemenang`,
`nama_penyedia_pemenang`
dalam satu tabel besar.

Tugas Anda:

1. Identifikasi dependensi fungsional pada tabel ini.
2. Tentukan anomali apa yang berpotensi muncul (Insert/Update/Delete).
3. Pecah tabel ini menjadi struktur 3NF (minimal 3 tabel terpisah).
4. Gambarkan foreign key yang menghubungkan tabel-tabel hasil normalisasi.

Optimasi Basis Data: Indexing & Kinerja Query

Apa itu Index?

Struktur data tambahan yang memungkinkan database menemukan baris dengan cepat tanpa membaca seluruh tabel — seperti daftar indeks buku yang menunjuk langsung ke halaman.

Kapan Membuat Index?

- Kolom sering difilter di WHERE atau JOIN (id_penyedia, id_instansi)
- Kolom Foreign Key yang merujuk tabel lain
- Kolom ORDER BY / GROUP BY pada query laporan besar
- Kolom high-cardinality: no_kontrak, NPWP, kode_paket

 **Trade-off:** Index mempercepat READ tetapi memperlambat INSERT/UPDATE/DELETE dan membutuhkan storage tambahan.

Jenis Index

B-Tree

Default; range, sort, LIKE "abc%"

Hash

Equality (=) cepat; tidak bisa range

Composite

Multi-kolom; urutan kolom menentukan efektivitas

Partial

Subset baris: WHERE status='aktif'

Contoh SQL — Basis Data LKPP

(asumsi)

```
-- Kolom FK: cari tender per penyedia
CREATE INDEX idx_tender_penyedia
ON TENDER(id_penyedia);

-- Composite: laporan per instansi & tanggal
CREATE INDEX idx_kontrak_inst_tgl
ON KONTRAK(id_instansi, tanggal_kontrak);

-- Unique index: pastikan NPWP tidak ganda
ALTER TABLE PENYEDIA
ADD CONSTRAINT uq_npwp UNIQUE (npwp);
```

Latihan 4: Identifikasi Kolom Index pada SPSE

Skenario

Query berikut berjalan lambat di portal SPSE LKPP karena tabel TENDER memiliki 5 juta baris:

```
SELECT * FROM TENDER WHERE id_instansi = 101 AND status_tender = 'selesai'  
ORDER BY tanggal_penutupan DESC;
```

Tugas Anda:

1. Kolom mana yang paling perlu diindex? Jelaskan alasannya.
2. Apakah index composite atau index terpisah lebih baik di sini? Mengapa?
3. Jika kolom status_tender hanya memiliki 3 nilai (aktif/selesai/batal), apakah index tetap berguna? Jelaskan konsep selectivity.
4. Tuliskan perintah CREATE INDEX yang Anda rekomendasikan.

Bagian 6

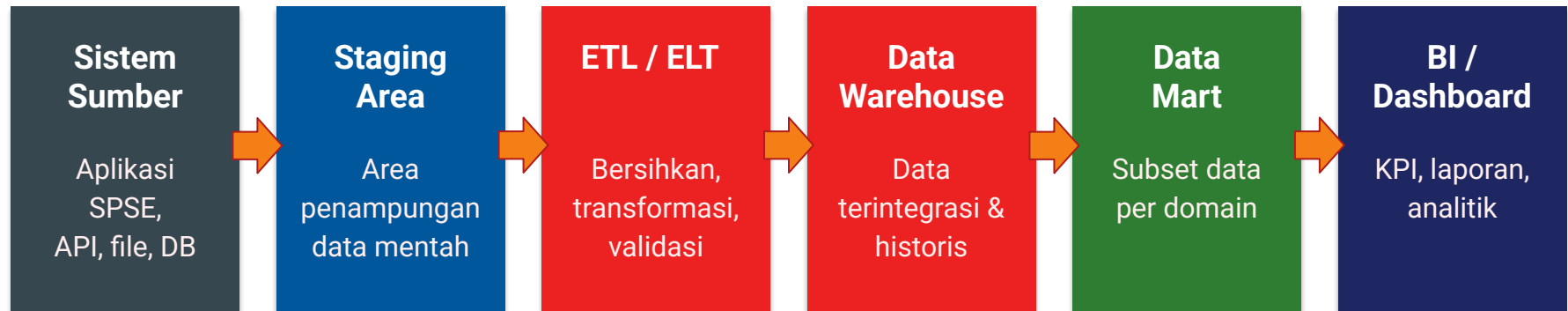
Data Warehouse

Definisi, arsitektur, pemodelan dimensional (star schema), dan operasi OLAP untuk analitik pengadaan.

OLTP vs Data Warehouse (OLAP)

Aspek	OLTP (Basis Data Operasional)	Data Warehouse (OLAP)
Tujuan	Mencatat transaksi harian	Analisis & pengambilan keputusan
Struktur data	Ternormalisasi (3NF)	Denormalisasi (star/snowflake)
Contoh	Input kontrak baru di SPSE	Tren nilai pengadaan 5 tahun terakhir
Volume per query	Sedikit baris, sering	Banyak baris, agregasi, jarang
Sumber definisi	—	Inmon (1993): koleksi data subject-oriented, terintegrasi, time-variant, non-volatile

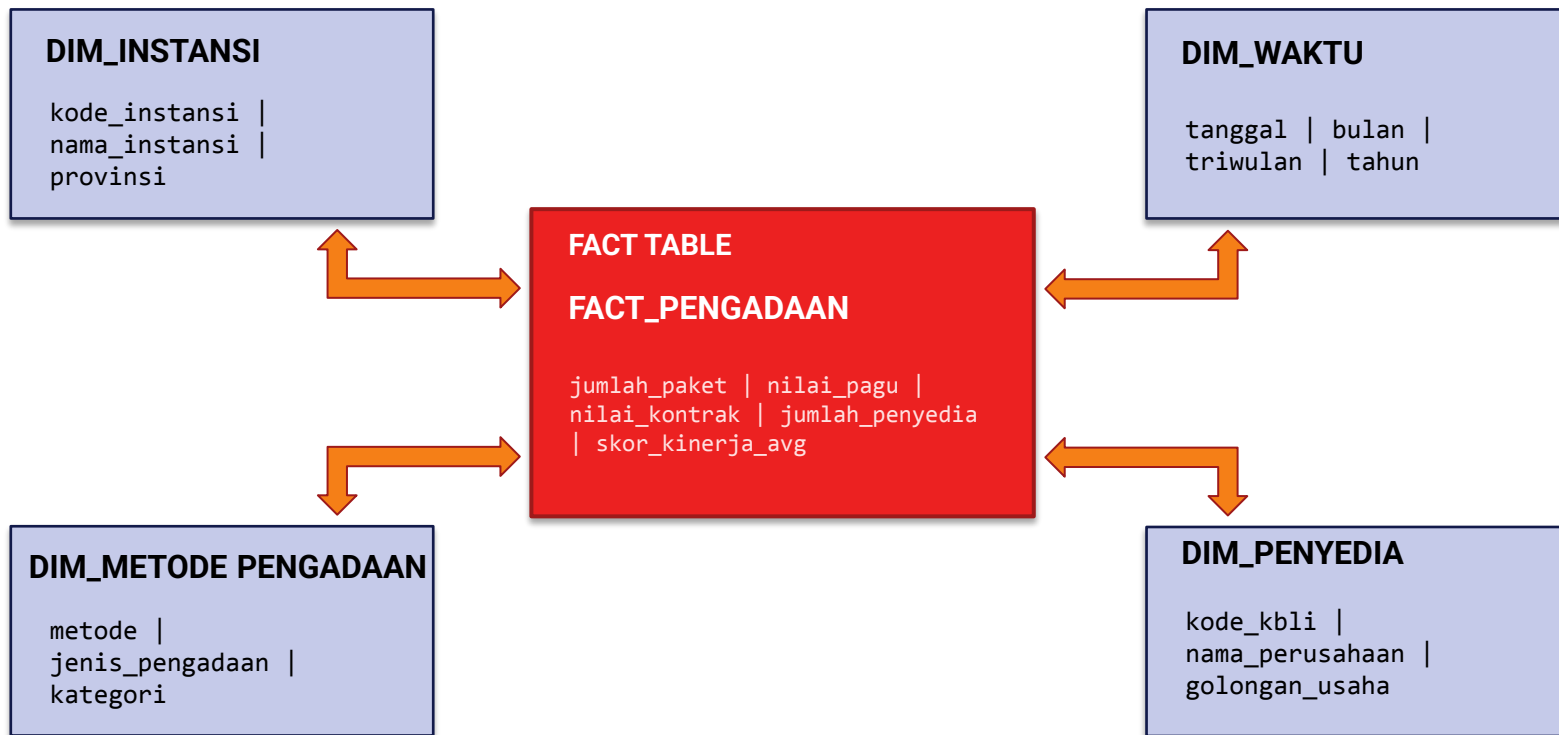
Arsitektur Data Warehouse



Contoh dashboard: realisasi pengadaan, kinerja penyedia, capaian TKDN, monitoring pagu vs. realisasi anggaran

 **Garbage in, Garbage out**

Star Schema: Model Data Warehouse (untuk Data Pengadaan)



* Foreign key pada Fact Table menghubungkan ke setiap Dimension Table.

Operasi OLAP: Menjelajahi Data Pengadaan Multidimensi



Hierarki: Nasional -> Provinsi -> Instansi/Satker.

Roll-up meringkas nilai pengadaan ke level lebih tinggi; drill-down membuka rincian per satker.

Slice & Dice

Memotong data cube: contoh "pengadaan Jasa Konstruksi di Jawa Barat, triwulan 2".

Pivot (Rotate)

Menukar sumbu analisis: dari Waktu x Instansi menjadi Penyedia x Metode Pengadaan.

Pemanfaatan Data Warehouse untuk LKPP

Monitoring Realisasi Pengadaan

Membandingkan nilai pagu vs realisasi kontrak per instansi dan triwulan secara nasional.

Evaluasi Kinerja Penyedia

Mengagregasi skor kinerja penyedia lintas kontrak untuk mendukung keputusan blacklist/whitelist.

Analisis Trend Metode Pengadaan

Melihat pergeseran tren Tender vs E-purchasing dari waktu ke waktu.

Deteksi Anomali & Potensi Persekongkolan

Mengidentifikasi pola tidak wajar pada pemenang tender lintas paket dan wilayah.

Latihan 5: Merancang Star Schema

Skenario

Unit Monitoring & Evaluasi ingin membangun dashboard untuk memantau total nilai kontrak per provinsi, per jenis pengadaan, dan per triwulan.

Tugas Anda:

1. Tentukan fact table dan ukuran (measure) yang relevan.
2. Tentukan dimension table yang diperlukan beserta atributnya.
3. Gambarkan skema star schema sederhana untuk kasus ini.
4. Sebutkan satu contoh operasi OLAP (roll-up/drill-down/slice/pivot) yang akan digunakan pengguna dashboard.

Bagian 6

Keamanan Data Pemerintah

Risiko keamanan, klasifikasi data, serta kontrol mitigasi teknis dan tata kelola.

Keamanan Data Pemerintah: Risiko & Klasifikasi

Vektor Risiko Utama

	Akses tidak sah ke data penyedia/kontrak
	Kebocoran data pribadi (NPWP, KTP pengurus)
	Serangan SQL injection pada portal pengadaan
	Ransomware / malware
	Konfigurasi cloud/server yang salah
	Backup tidak aman
	Berbagi data tanpa kendali antar-instansi
	Ancaman dari dalam (insider threat)

Tingkat Klasifikasi Data

Publik	Data terbuka – tanpa batasan (mis. RUP)
Internal	Penggunaan internal instansi – berbagi terbatas
Terbatas	Data operasional sensitif (evaluasi penyedia)
Rahasia	Berdampak besar jika bocor (HPS, dok. penawaran)
Data Pribadi	Diatur UU 27/2022 PDP (NPWP, NIK pengurus)
Data Pribadi Spesifik	Keuangan, biometrik, data sensitif lain

Mitigasi Keamanan: Kontrol Teknis & Tata Kelola

Kontrol Teknis

- Role-Based Access Control (RBAC) per peran (panitia, auditor, penyedia)
- Prinsip least privilege
- Multi-Factor Authentication (MFA)
- Enkripsi data saat disimpan & dikirim
- Audit log lengkap atas akses dan perubahan data
- Masking data sensitif (NPWP, NIK)
- Backup terenkripsi & prosedur DR
- Pemantauan & peringatan anomali akses

Kontrol Tata Kelola & Kebijakan

- Penetapan kepemilikan data (data owner) per unit
- Perjanjian berbagi data (PKS/DSA) antar-instansi
- SOP permintaan dan persetujuan akses data
- Kebijakan klasifikasi dan retensi data
- Pelatihan kesadaran keamanan secara berkala
- Rencana tanggap insiden (IRP)
- Audit keamanan berkala
- Kepatuhan terhadap UU 27/2022 PDP

 **Tidak semua data yang boleh dibagikan harus dibagikan. Tidak semua perlu dalam bentuk mentah. Tidak semua pengguna yang berwenang perlu akses unduh.**

Rangkuman

1

Basis data
mendukung
transaksi harian

2

Normalisasi
menghindari
anomali data

3

Standardisasi
membuat data
dapat
diintegrasikan

4

Gudang data
mengaktifkan
analitik & OLAP

5

Keamanan data
harus dirancang
sejak awal

"Data pemerintah yang baik bukan hanya besar dan cepat – tetapi akurat, terstandar, aman, interoperabel, dan akuntabel."

Refleksi:

1. Apa satu kesenjangan standardisasi pada data di instansi Anda saat ini?
2. Data apa di instansi Anda yang paling diuntungkan oleh data warehouse?
3. Sebutkan satu kontrol keamanan yang mungkin belum dimiliki sistem Anda.

Pengumpulan dan Penyiapan Data

Yang Anda Bawa ke Sesi Selanjutnya

- Desain skema basis data: menentukan field apa yang harus dikumpulkan
- Tipe data & constraint: menentukan format dan nilai yang valid
- Aturan standardisasi: memastikan data yang dikumpulkan konsisten
- Klasifikasi keamanan: menentukan siapa yang boleh mengumpulkan apa

Sesi Selanjutnya Akan Membahas

- Akuisisi Data: kebutuhan, sumber & pengumpulan data
- Pelabelan Data (data labeling)
- Prapemrosesan Data: validitas, missing value, normalisasi, outlier
- Perekayasa Fitur: seleksi & transformasi fitur

Pertanyaan lanjutan:

Bagaimana Anda akan merancang skema basis data untuk upaya pengumpulan data berikutnya di instansi Anda?