

Introduction to *Data* and *Database*

Basic SQL

Learning Progress Review Week 3

By:
MARVEL TEAM

Fikrie | Natalia | Satria

1. *Introduction to Data and Database*

Apa yang dimaksud dengan *Data* dan *Database*?

Apa itu Data?

Data merupakan sebuah hasil dari pengukuran atau pengamatan suatu variabel yang bentuknya dapat berupa kata-kata maupun angka. Pada dasarnya, data dapat digunakan sebagai alat bantu untuk membuat keputusan.

Tipe – tipe Data dibedakan menjadi 3, yaitu:

1. *Structured Data* : Jenis data yang dapat di proses, disimpan dan diambil dalam format tetap. Jenis data ini disimpan dalam bentuk tabel, baris dan kolom yang normalnya disimpan dalam excel atau *spreadsheet*. Contoh : data penjualan, data karyawan dalam *database* perusahaan.
2. *Semi Structured Data* : jenis data yang dimasukkan ke dalam sebuah tabel, tetapi skemanya tidak sama dengan tabel biasa yang hanya terdiri dari baris dan kolom. Contoh : file xml, json.
3. *Unstructured Data* : jenis data dengan bentuk yang tidak dikenal, harus disimpan dengan format khusus karena tidak memiliki struktur yang spesifik seperti *data structured*. Contoh : data teks, video, suara.

Dimana kita menyimpan Data?

Database atau basis data adalah kumpulan informasi yang disimpan di dalam komputer *secure systematic* sehingga dapat diperiksa menggunakan suatu program komputer untuk memperoleh informasi dari basis data tersebut.

Tipe – tipe Database berdasarkan kegunaannya, antara lain :

OLTP (Online Transaction Processing) : adalah sistem yang berorientasi proses yang memproses transaksi secara langsung melalui komputer yang terhubung dalam jaringan.

Contoh : PostgreSQL, MySQL, MariaDB, MongoDB

OLAP (Online Analytical Processing) : adalah suatu proses yang digunakan untuk melakukan permintaan terhadap data dalam bentuk yang kompleks dan menganalisa data yang bervolume besar. Maka dari itu OLAP seringkali disebut analisis data multidimensi.

Contoh : Data Warehouse, RedShift (AWS), BigQuery (GCP), Teradata

Unveiling Big Data

Big Data adalah kumpulan proses yang terdiri dari volume data dalam jumlah besar yang terstruktur maupun tidak terstruktur dan digunakan untuk membantu kegiatan bisnis. Istilah *Big Data* muncul setelah penemuan Hadoop Technology.



Hadoop adalah sebuah piranti lunak garapan Apache Software Foundation yang sifatnya *open source* yang digunakan untuk menangani big data yang jumlahnya sangat besar.

Big Data tercipta ketika penyimpanan dan pengolahan data lebih murah sehingga perusahaan mau menyimpan lebih banyak data.

Unveiling Big Data

Tipe – tipe server terdiri dari :

On - Premise

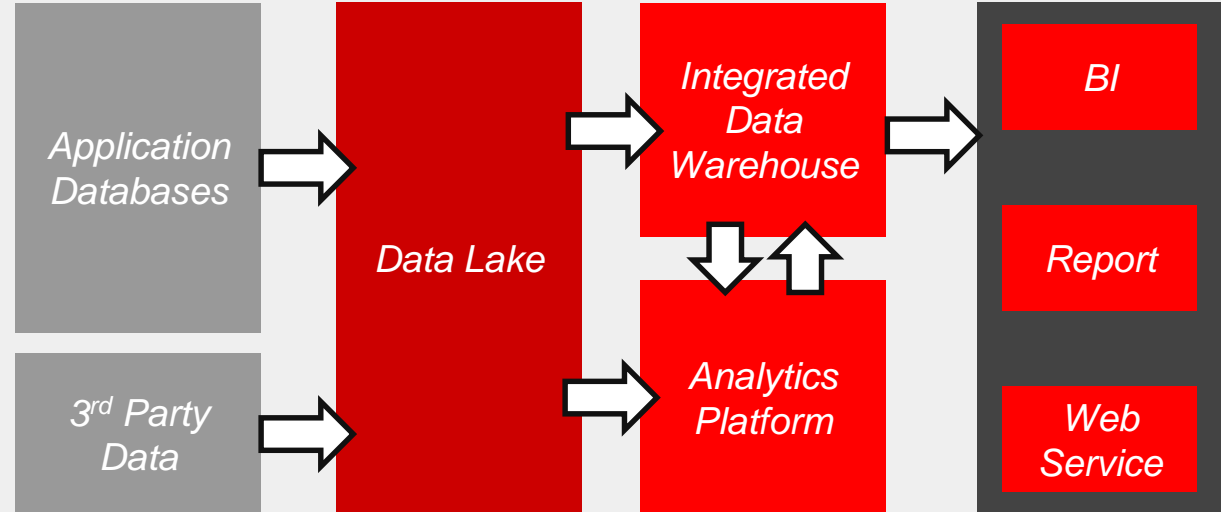
- Dapat mengontrol seluruh system.
- Data dapat disimpan dan dikelola langsung oleh perusahaan.
- Sistem administrasi dimiliki sendiri.
- Biaya investasi tinggi tetapi dapat digunakan dalam jangka waktu panjang.
- Perusahaan yang biasa menggunakan adalah perusahaan *mature* atau pemerintahan.

Cloud

- Dimiliki oleh pihak ketiga seperti Google, Amazon, IBM, dll).
- Sudah memiliki layanan bawaan.
- Sebagian keamanan dan infrastruktur dikontrol oleh pihak ketiga.
- Biaya investasi terjangkau.
- Perusahaan yang biasa menggunakan adalah perusahaan *start-up* atau *fast-growing*.

End to end Data Architecture

Arsitektur Data yang Ideal



2. Basic SQL

Let's Code!

Apa itu SQL?

SQL (**Structured Query Language**) adalah bahasa khusus yang digunakan untuk menarik data yang terstruktur dari *database*.

SQL adalah bahasa di atas bahasa. Dikarenakan ketika menarik data, di belakang layar SQL menjadi bahasa pemrograman lain, dia akan diterjemahkan menjadi Java, C++ dll, ketika data sudah terambil maka akan dikembalikan ke dalam bentuk SQL.

SQL adalah bahasa yang paling sederhana dan paling relevan untuk digunakan saat ini. SQL juga berbeda dengan bahasa lainnya, dikarenakan SQL ini tidak sensitif terhadap huruf kapital, spasi dan enter sehingga mudah untuk dipahami.



Apa yang membuat *Database* lebih baik dari *Spreadsheet*?

Limitasi Spreadsheet :

- a. Memiliki batasan kolom, baris dan sel;
- b. Ketika data banyak, sistem akan berjalan lambat;
- c. Kesulitan ketika akan menggabungkan data.

Keunggulan Database :

- a. Lebih mudah untuk menggabungkan data;
- b. Ketika data banyak, sistem akan tetap berjalan normal;
- c. Bisa diaplikasikan dalam semua software.



Mengapa menggunakan PostgreSQL?

Terdapat beberapa penyimpanan di dalam SQL, seperti My SQL, Maria DB, PostgreSQL dll. Paling familiar adalah My SQL dikarenakan lebih dulu muncul.

Namun, **PostgreSQL** diyakini memiliki keunggulan dari jenis penyimpanan SQL lainnya. Terdapat fitur-fitur tertentu di dalam **PostgreSQL** yang tidak dimiliki oleh jenis penyimpanan SQL yang lain, seperti :

- a. Sistem keamanan yang berbeda;
- b. Dapat menyimpan berbagai jenis tipe data di dalam satu file;
- c. Memiliki performa yang jauh lebih baik;
- d. Jaminan terhadap *recovery* data dalam keadaan yang tidak diduga.



Let's Code!

SQL mempunyai struktur bahasa *query* yang mirip dengan Bahasa Inggris.

SELECT DATA

SELECT "COLUMN_NAME"
FROM "TABLE_NAME"

```
select  
  *  
from datasource.employees as e;
```

* berarti memilih seluruh kolom.

```
select  
  employee_id  
  , first_name  
  , job_id  
from datasource.employees as e;
```

Memilih beberapa kolom.

```
select  
  distinct  
  job_id  
from datasource.employees as e;
```

Apabila kolom terdapat nilai duplikat, kita dapat menampilkan nilai uniknya.

Let's Code!

CREATE TABLE

Mari lihat contoh!

```
create table batch_4.profile_natalia (  
  id int  
  , nama varchar  
  , birthdate date  
  , ktp bigint  
  , badge varchar  
)
```

Jangan lupa tanda kurung () ketika membuat table!

Perhatikan tipe data yang digunakan!

varchar : huruf

date : tanggal dengan format yyyy-mm-dd

numeric

int

bigint

float

} Penulisan angka

INSERT TABLE

Mari lihat contoh!

```
insert into batch_4.profile_natalia values  
(1, 'Udin Amir', '1990-01-01', 1234567891011, 'WNI'),  
(2, 'Andi Tanuwijaya', '1991-02-02', 1234567891012, 'WNI'),  
(3, 'Michael Angelo', '1992-03-03', 1234567891013, 'WNA'),  
(4, 'Mira Putri', '1993-04-04', 1234567891014, 'WNI'),  
(5, 'Angelica McMillan', '1994-05-05', 1234567891015, 'WNA');
```

Masukkan data berdasarkan urutan kolom-kolom yang dibuat dan jenis data di dalam kolom tersebut.

UPDATE TABLE

Mari lihat contoh!

```
update batch_4.profile_natalia  
set birthdate = '1998-10-09'  
where id = 3
```

Meng-update atau mengubah data yang sudah ada dalam kolom.

where digunakan Ketika ingin memilih spesifik data.

MENAMBAH KOLOM BARU

ALTER TABLE ...
ADD ...

```
alter table batch_4.profile_natalia add alamat varchar;
```

Jangan lupa memasukkan jenis datanya apakah varchar/int/numeric, dll.

MENGHAPUS KOLOM

ALTER TABLE ...
DROP COLUMN

```
alter table batch_4.profile_natalia drop column alamat;
```

Let's Code!

MENGHAPUS BARIS

DELETE FROM ...
WHERE

```
delete  
from batch_4.profile_natalia  
where ID = 5;
```

Tidak direkomendasikan menghapus baris atau tabel, kecuali tabel milik sendiri. Hati – hati dalam menghapus ketika di dunia kerja.

MEMBERSIHKAN TABEL

TRUNCATE TABLE ...

```
truncate table batch_4.profile_natalia;
```

MENGHAPUS TABEL

DROP TABLE ...

Special Thanks to :



Slide template by SlideCarnival