

**LAPORAN UAS**  
**DATA WAREHOUSE**

**Kelompok 4**



Disusun Oleh:

<b>Arimbi Putri Hapsari</b>	<b>2341760016</b>
<b>Karina Ika Indasa</b>	<b>2341760042</b>
<b>Khuzaima Filla Januardha</b>	<b>2341760078</b>
<b>Paudra Akbar Buana</b>	<b>2341760063</b>
<b>Reza Angelina Febriyanti</b>	<b>2341760015</b>

**Jurusan Teknologi Informasi**  
**[D4 Teknik Informatika/D4 Sistem Informasi Bisnis]**  
**Politeknik Negeri Malang**  
**2024**

## 1. Cari Studi Kasus Pengelolaan Data

- Link Dataset:

<https://www.kaggle.com/datasets/andrexibiza/grocery-sales-dataset>

Dalam Dataset diatas telah memenuhi kriteria:

- Data Karyawan = **employees**
- Data Penjualan = **sales**
- Data Inventaris = **products**
- Data Pelanggan = **customers**

## 2. Pilih Studi Kasus yang Memenuhi Kriteria

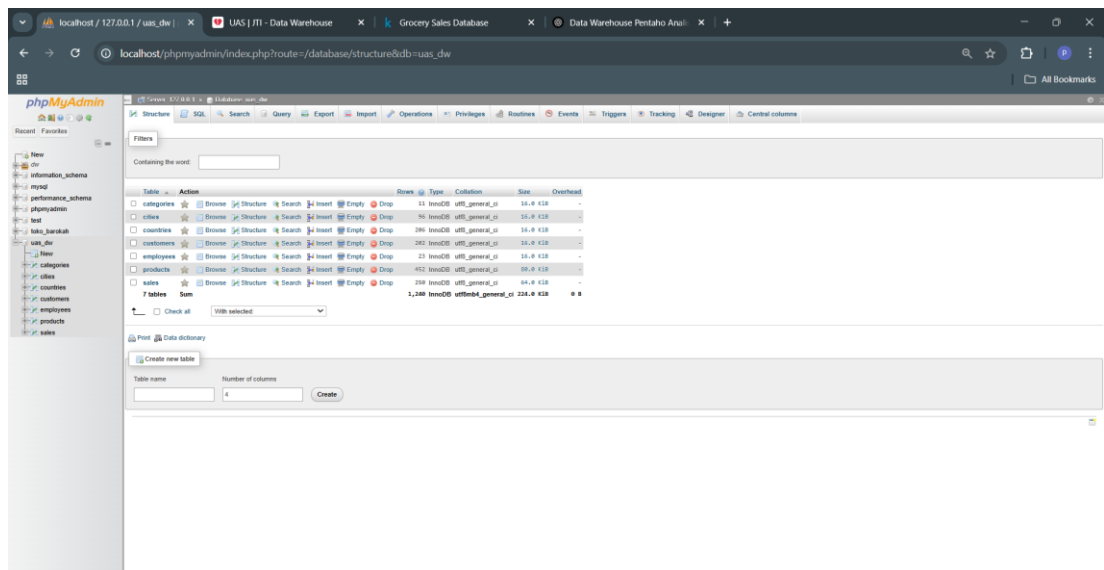
- Studi Kasus:

Melakukan pengambilan data dari tiga tabel utama dalam data warehouse, yaitu *Fact\_Sales*, *Dim\_Date*, *Dim\_Employees*, Dan *Dim\_Customers* untuk menyusun laporan penjualan karyawan. Laporan ini bertujuan menampilkan total penjualan selama tahun 2017 yang dilakukan oleh karyawan dengan EmployeeID = 22.

Hasil ini digunakan sebagai dasar evaluasi kinerja individual dalam mendukung pengambilan keputusan manajerial.

## 3. Ambil Data Dummy dari Studi Kasus Terpilih

- Database:

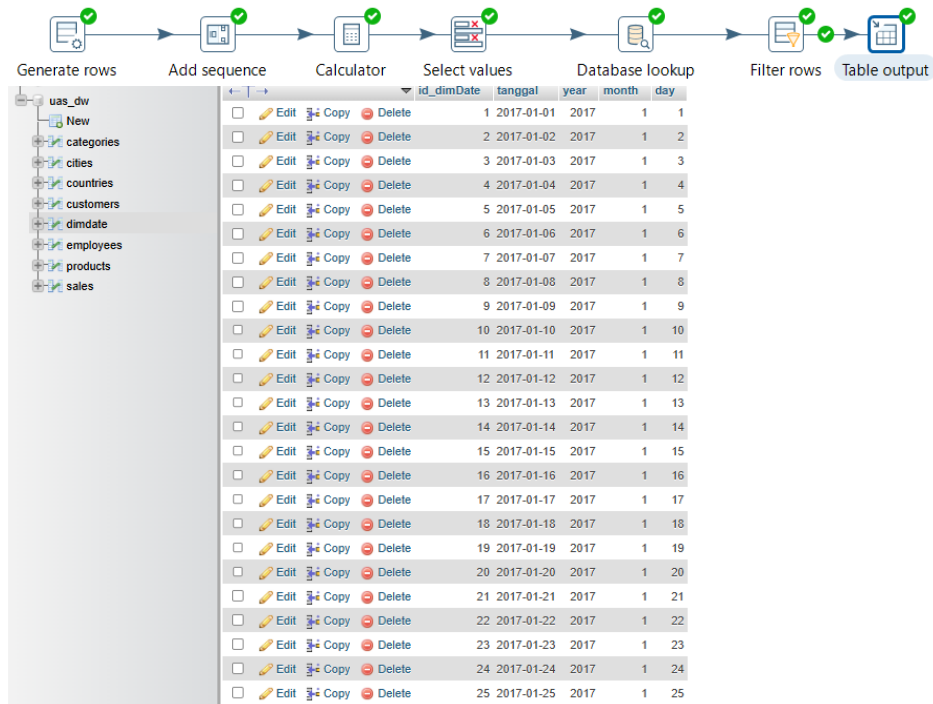


Impor data dummy yang bertipe (csv) dan buat database setelah itu masukkan tiap file csv menjadi 1 tabel. Dan pada dataset saya berisi 7 file/table yaitu:

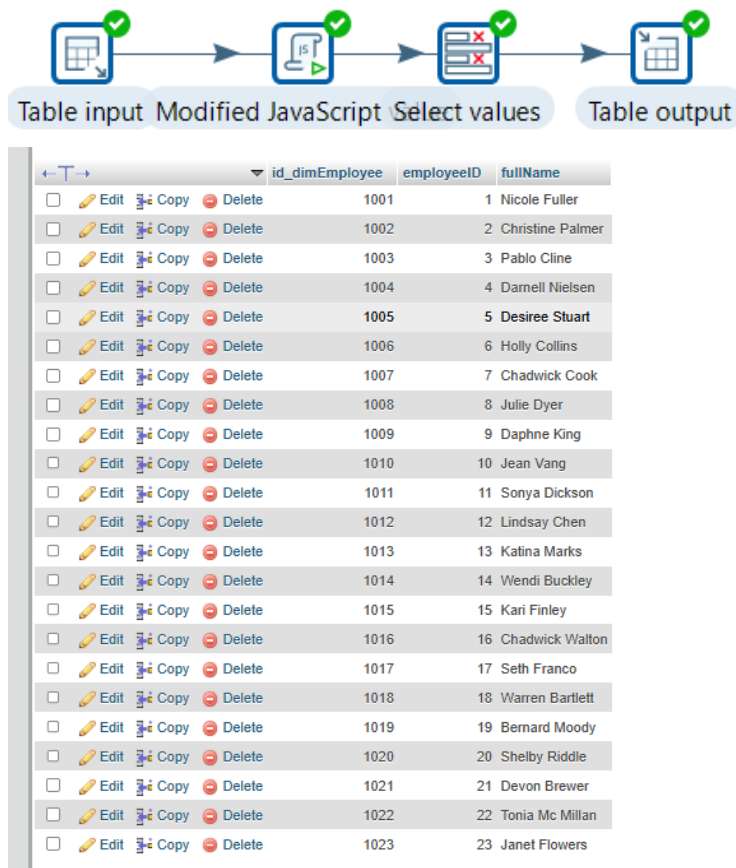
- categories.csv
- cities.csv
- countries.csv
- customers.csv
- employees.csv
- products.csv
- sales.csv

#### 4. Rancang Skema Bintang (Star Schema) untuk Data Warehouse

##### - Dimensi Waktu



##### - Dimensi Pegawai



- Dimensi Customer

Diagram flow: Table input → Modified JavaScript → Select values → Table output

	id_dimCust	customerID	fullName
<input type="checkbox"/>	1001	0	FirstName LastName
<input type="checkbox"/>	1002	1	Stefanie Frye
<input type="checkbox"/>	1003	2	Sandy Kirby
<input type="checkbox"/>	1004	3	Lee Zhang
<input type="checkbox"/>	1005	4	Regina Avery
<input type="checkbox"/>	1006	5	Daniel Mccann
<input type="checkbox"/>	1007	6	Dennis Zuniga
<input type="checkbox"/>	1008	7	Myra Jefferson
<input type="checkbox"/>	1009	8	Teddy Parsons
<input type="checkbox"/>	1010	9	Annie Stein
<input type="checkbox"/>	1011	10	Herman Pope
<input type="checkbox"/>	1012	11	Jenifer Winters
<input type="checkbox"/>	1013	12	Stefanie Gordon
<input type="checkbox"/>	1014	13	Chastity Gillespie
<input type="checkbox"/>	1015	14	Evelyn Sweeney
<input type="checkbox"/>	1016	15	Jeannie Gardner
<input type="checkbox"/>	1017	16	Rickey Ho
<input type="checkbox"/>	1018	17	Bobbie Simon
<input type="checkbox"/>	1019	18	Lesley Orr
<input type="checkbox"/>	1020	19	Shawna Herring
<input type="checkbox"/>	1021	20	Theresa Berg
<input type="checkbox"/>	1022	21	Tasha Heath
<input type="checkbox"/>	1023	22	Karla Sutton
<input type="checkbox"/>	1024	23	Otis Miller
<input type="checkbox"/>	1025	24	Alisa Ruiz

- Fakta QTY

Diagram flow: Table input → Database lookup → Database lookup 2 → Database lookup 3 → Select values → Table output

id_dimDate	id_dimEmployee	id_dimCust	totalPrice
22	1006	1116	9.60
244	1016	1019	2.80
156	1013	1150	9.40
49	1008	1088	5.40
23	1010	1189	0.10
220	1013	1077	8.80
30	1014	1019	5.30
239	1003	1064	5.30
105	1016	1059	7.00
57	1022	1102	5.40
222	1005	1131	7.60
186	1017	1146	8.30
13	1022	1133	5.60
7	1007	1027	6.60
244	1019	1135	3.40
201	1016	1193	2.10
21	1008	1153	9.60
250	1017	1186	5.00
189	1023	1065	4.70
195	1018	1171	5.80
159	1016	1054	7.40
206	1018	1158	0.60
53	1001	1022	0.80
147	1023	1038	9.90
74	1007	1120	4.30

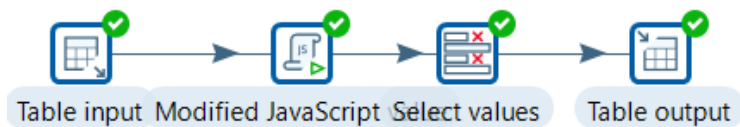
## 5. Rancang Proyek ETL Berdasarkan Skema Bintang

- **Extract:** Mengambil data dari sumber data dummy (file CSV, database, API, dsb).
- **Transform:** Membersihkan data, mengubah format data, konversi tipe data, agregasi, dan transformasi lainnya sesuai kebutuhan.
- **Load:** Memasukkan data yang sudah bersih dan siap ke dalam tabel fakta dan dimensi di data warehouse.

### - Dimensi Waktu



### - Dimensi Pegawai



### - Dimensi Customer



### - Fakta QTY



## 6. Jelaskan Langkah-Langkah Proses ETL

- **Extract:** Membaca data dummy dari sumber, misalnya membaca file CSV berisi data karyawan atau data lain.
  - Membuat table baru Bernama uas\_dw, setelah itu Impor data dummy yang telah di download pada Kaggle juga bertipe (csv) setelah itu masukkan tiap file csv menjadi 1 tabel
- **Transform:** Melakukan pembersihan data seperti memperbaiki format tanggal, normalisasi nama atau kategori, menghapus data duplikat, menyesuaikan struktur data sesuai skema star schema.
  - Memperbaiki tipe data yang berbeda pada tiap table sebelum membuat table dimensi, dan disamakan nantinya Ketika ingin membuat table dimensi
- **Load:** Memasukkan data ke tabel fakta dan dimensi dengan memperhatikan integritas data, kunci primer, dan relasi antar tabel.
  - Mengambil id pada tiap table nantinya pada table dimensi yang ingin dibuat, lalu beberapa value yang ingin diambil sesuai dengan studi kasus yang ada diatas.

## 7. Studi Kasus

Query:

```
1 SELECT e.fullName,d.tanggal, f.totalPrice
2 FROM dimemployees e
3 JOIN factqty f
4 ON e.id_dimEmployee = f.id_dimEmployee
5 JOIN dimdate d ON f.id_dimDate = d.id_dimDate
6 WHERE d.year = 2017
7 AND e.employeeID = 22
8 GROUP BY d.tanggal,e.employeeID
9 ORDER BY `d`.`tanggal` ASC
```

Output:

fullName	tanggal	totalPrice
Tonia Mc Millan	2017-01-01	5.40
Tonia Mc Millan	2017-01-02	5.40
Tonia Mc Millan	2017-01-03	5.40
Tonia Mc Millan	2017-01-04	5.40
Tonia Mc Millan	2017-01-05	5.40
Tonia Mc Millan	2017-01-06	5.40
Tonia Mc Millan	2017-01-07	5.40
Tonia Mc Millan	2017-01-08	5.40
Tonia Mc Millan	2017-01-09	5.40
Tonia Mc Millan	2017-01-10	5.40
Tonia Mc Millan	2017-01-11	5.40
Tonia Mc Millan	2017-01-12	5.40
Tonia Mc Millan	2017-01-13	5.40
Tonia Mc Millan	2017-01-14	5.40
Tonia Mc Millan	2017-01-15	5.40
Tonia Mc Millan	2017-01-16	5.40
Tonia Mc Millan	2017-01-17	5.40
Tonia Mc Millan	2017-01-18	5.40
Tonia Mc Millan	2017-01-19	5.40
Tonia Mc Millan	2017-01-20	5.40
Tonia Mc Millan	2017-01-21	5.40
Tonia Mc Millan	2017-01-22	5.40
Tonia Mc Millan	2017-01-23	5.40
Tonia Mc Millan	2017-01-24	5.40
Tonia Mc Millan	2017-01-25	5.40
Tonia Mc Millan	2017-01-26	5.40
Tonia Mc Millan	2017-01-27	5.40