

LAPORAN PRAKTIKUM DATA WAREHOUSE
JOBSHEET 3 : DATABASE ANALYTICAL C-D



Oleh :

Danica Nasywa Putrinier (2341760122)

Kelas SIB 2B / 05

PROGRAM STUDI D-IV SISTEM INFORMASI BISNIS JURUSAN
TEKNOLOGI INFORMASI
POLITEKNIK NEGERI MALANG

Jl. Soekarno Hatta No.9, Jatimulyo, Kec. Lowokwaru, Kota Malang, Jawa Timur

65141

C. Fakta Pembayaran

1. Buatlah sebuah tabel pada database OLAP yang telah dibuat (database dw_legendVehicle).
 - nama tabel: FactOmset
 - id_dimEmployees (foreign key ke dimEmployees)
 - id_dimDate (foreign key ke dimDate)
 - amount (jumlah pembayaran dengan tipe decimal(10,2))

```
✓ MySQL returned an empty result set (i.e. zero rows). (Query took 0.0441 seconds.)

CREATE TABLE FactOmset (id_dimEmployees INT, id_dimDate INT, amount DECIMAL(10,2),
FOREIGN KEY (id_dimEmployees) REFERENCES dimEmployees(id_dimEmployees), FOREIGN KEY
(id_dimDate) REFERENCES dimDate(id_dimDate));

[ Edit inline ] [ Edit ] [ Create PHP code ]
```

2. Buat Transformation baru pada PDI Spoon. Gunakan objevt-object yang dihubunugkan sebagai berikut:
 - Table Input: Dignakan untuk mengambil data transaksi dari OLTP
 - Select values: digunakan untuk memilih field yang akan digunakan untuk OLAP
 - Database lookup (1) : digunakan untuk mencocokkan data pada tabel dimEmployee untuk mengambil id_dimEmployee.
 - Database lookup (2) : digunakan untuk mencocokkan data pada tabel dimDate untuk mengambil id_dimDate.
 - Databse lookup (3): digunakan untuk mencocokkan data pada tabel factomset untuk melihat data yang sama atau tidak.
 - Filter rows: digunakan untuk memiih data yang sudah ada pada tabel factomset tidak dimasukkan lagi.
 - Table output: digunakan untuk memasukkan data pada tabel factOmset.



3. Konfigurasi pada Table input adalah untuk mendapatkan data dari OLTP sesuai dengan hasil dari query berikut. query tersebut akan menampilkan seluruh isi field dari employee hingga transaksi yang didapat dilihat dari payments yang didapat.

Table input

Step name: Table input

Connection: conn_oltp_resources

SQL:

```
select *
from employees e
left join employees r on e.reportsTo=r.employeeNumber
inner join customers c on e.employeeNumber=c.salesRepEmployeeNumber
inner join payments p on c.customerNumber=p.customerNumber;
```

Line 1 Column 0

Store column info in step meta data ☐

Enable lazy conversion ☐

Replace variables in script? ☐

Insert data from step

Execute for each row? ☐

Limit size: 0

Help OK Preview Cancel

4. Konfigurasi pada Select Values adalah untuk menghapus semua field kecuali employeeNumber, lastname dari employee, firstname dari employee, jobTitle

Select values

Step name: Select values

Select & Alter Remove Meta-data

Fields:

#	Fieldname	Rename to
1	employeeNumber	streamEmployeeNumber
2	lastName	streamLastName
3	firstName	streamFirstName
4	extension	
5	email	
6	officeCode	
7	reportsTo	
8	jobTitle	streamJobTitle
9	employeeNumber_1	
10	lastName_1	streamLastNameBoss
11	firstName_1	streamFirstNameBoss
12	extension_1	
13	email_1	
14	officeCode_1	
15	reportsTo_1	
16	jobTitle_1	
17	customerNumber	
18	customerName	
19	contactLastName	
20	contactFirstName	
21	phone	
22	addressLine1	
23	addressLine2	
24	city	
25	state	
26	postalCode	

Select & Alter Remove Meta-data

Fields to remove:

#	Fieldname
1	extension
2	email
3	officeCode
4	reportsTo
5	employeeNumber_1
6	extension_1
7	email_1
8	officeCode_1
9	reportsTo_1
10	jobTitle_1
11	customerNumber
12	customerName
13	contactLastName
14	contactFirstName
15	phone
16	addressLine1
17	addressLine2
18	city
19	state
20	postalCode
21	country
22	salesRepEmployeeNumber
23	creditLimit
24	customerNumber_1
25	checkNumber

dari employee, lastname dari manager, firstname dari manager, payment date dan amount.

- Konfigurasi pada tabel database lookup adalah dengan melakukan komparasi field stream (output dari proses sebelumnya) dengan field isi data pada tabel dimEmployees. Jika data tersebut cocok maka akan diambil id_dimEmployees

Step name: Database lookup

Connection: conn_dw_destination

Lookup schema: dw_legendvehicle

Lookup table: dimemployees

Enable cache? ☐

Cache size in rows (0=cache): 0

Load all data from table ☐

The key(s) to look up the value(s):

#	Table field	Comparator	Field1	Field2
1	employeeNumber	=	streamEmployeeNumber	
2	firstName	=	streamFirstName	
3	lastName	=	streamLastName	
4	jobTitle	=	streamJobTitle	
5	boss_firstName	=	streamFirstNameBoss	
6	boss_lastName	=	streamLastNameBoss	

Values to return from the lookup table:

#	Field	New name	Default	Type
1	id_dimEmployees			None

nya.

Step name: Database lookup.2

Connection: conn_dw_destination

Lookup schema: dw_legendvehicle

Lookup table: dimdate

Enable cache? ☐

Cache size in rows (0=cache): 0

Load all data from table ☐

The key(s) to look up the value(s):

#	Table field	Comparator	Field1	Field2
1	date	=	streamDate	

Values to return from the lookup table:

#	Field	New name	Default	Type
1	id_dimDate			None

- Konfigurasi pada tabel database lookup yang kedua adalah dengan melakukan komparasi field stream (output dari proses sebelumnya untuk field date) dengan field isi data pada tabel dimDate. Jika data tersebut cocok maka akan diambil id_dimDate nya.
- Konfigurasi pada tabel database lookup yang ketiga adalah dengan melakukan

komparasi id_dimDate dan id_dimEmployees yang diambil dari proses lookup sebelumnya dengan field isi data pada tabel factOmset. Jika data tersebut cocok

Step name: Database lookup 3

Connection: conn_dw_destination

Lookup schema: dw_legendvehicle

Lookup table: factomset

Enable cache? ☐

Cache size in rows (0=cache): 0

Load all data from table ☐

The key(s) to look up the value(s):

#	Table field	Comparator	Field1	Field2
1	id_DimDate	=	id_dimDate	
2	id_DimEmployees	=	id_dimEmployees	
3	amount	=	streamAmount	

Values to return from the lookup table:

#	Field	New name	Default	Type
1	id_DimDate			None
2	id_DimEmployees			None
3	amount			None

data tidak akan dimasukkan dalam tabel factOmset.

8. Pada bagian filter rows untuk melihat data belum ada pada tabel factOmset dengan melihat bahwa output dari lookup sebelumnya bernilai NULL.

Step name: Filter rows

Send 'true' data to step: Table output

Send 'false' data to step:

The condition:

☐

id_DimDate_1 IS NULL

AND

id_DimEmployees_1 IS NULL

AND

amount IS NULL

9. Konfigurasi terakhir pada table output adalah dengan melakukan mapping data output dari proses sebelumnya ke dalam field pada tabel factOmset.

Table output

Step name: **Table output**

Connection: **conn_dw_destination**

Target schema: **dw_legendvehicle**

Target table: **factomset**

Commit size: **1000**

Truncate table: ☐

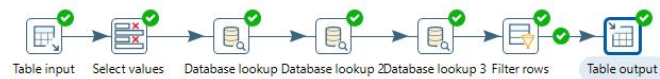
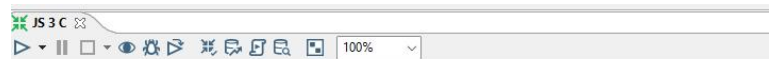
Ignore insert errors: ☐

Specify database fields: ☒

Main options | **Database fields**

Fields to insert:

#	Table field	Stream field
1	id_DimDate	id_dimDate
2	id_DimEm...	id_dimEmple...
3	amount	streamAmou...



Execution Results

Logging | Execution History | Step Metrics | Performance Graph | Metrics | Preview data

2025/04/22 20:47:34 - Table input.0 - Finished reading query, closing connection

2025/04/22 20:47:34 - Table input.0 - Finished processing (I=273, O=0, R=0, W=273, U=0, E=0)

2025/04/22 20:47:34 - Select values.0 - Finished processing (I=0, O=0, R=273, W=273, U=0, E=0)

2025/04/22 20:47:34 - Database lookup.0 - Finished processing (I=273, O=0, R=273, W=273, U=0, E=0)

2025/04/22 20:47:35 - Database lookup 2.0 - Finished processing (I=273, O=0, R=273, W=273, U=0, E=0)

2025/04/22 20:47:35 - Database lookup 3.0 - Finished processing (I=0, O=0, R=273, W=273, U=0, E=0)

2025/04/22 20:47:35 - Filter rows.0 - Finished processing (I=0, O=0, R=273, W=273, U=0, E=0)

2025/04/22 20:47:35 - Table output.0 - Finished processing (I=0, O=273, R=273, W=273, U=0, E=0)

2025/04/22 20:47:35 - Spoon - The transformation has finished!!

10. Jalankan proses transformation tersebut.

Hasil Transformation

Server: localhost:3306 Database: dw_legendvehicle Table: factomset

Browse Structure SQL Search Insert Export

Current selection does not contain a unique column. Grid edit, checkbox, Edit, Copy and De

Showing rows 0 - 24 (273 total, Query took 0.0003 seconds.)

`SELECT * FROM 'factomset'`

☐ Profiling [Edit inline] [Edit] [Explain SQL] [Create PHP code] [Refresh]

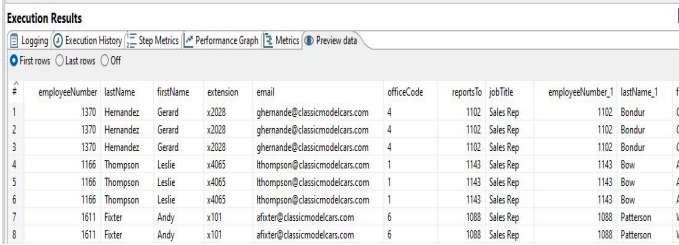
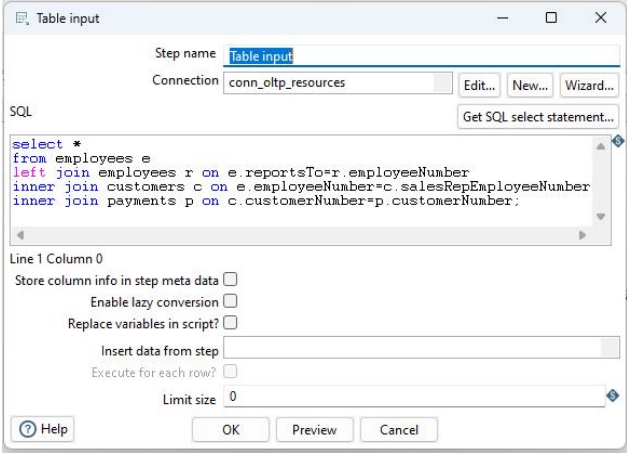
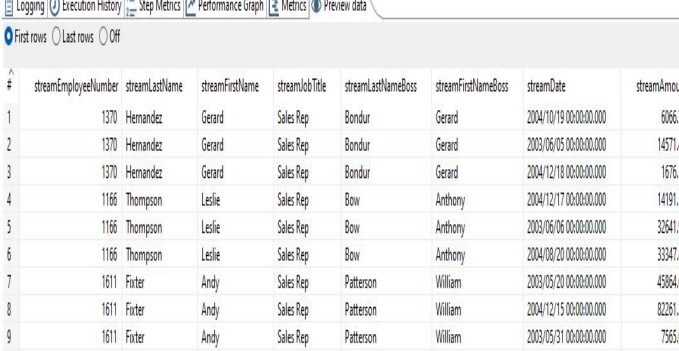
1 > >> ☐ Show all Number of rows: 25 Filter rows: Se

Extra options

id_dimEmployees	id_dimDate	amount
14	658	6066.78
14	156	14571.44
14	718	1676.14
8	717	14191.12
8	157	32641.98
8	598	33347.88
18	140	45864.03
18	715	82261.22
18	151	7565.08
18	435	44894.74
14	684	19501.82
14	586	47924.19
14	784	49523.67
17	47	50218.95
17	301	1491.38

Tugas 3

1. Buka preview tab pada execution result area di setiap proses object.
amati input dan output data yang ada. bandingkan di setiap prosesnya.
jelaskan perbedaan disetiap prosesny

Proses Objek	SS data input	SS data output	Keterangan
Table Input	data transaksi dari OLTP yaitu (employee, manager, customer, dan payments	 data yang sama seperti query employees	mengambil data transaksi dari database OLTP
Select values	 semua field hasil query	 hanya field yang dipilih (employeeNumber, lastname, firstname, jobTitle, manager, paymentDate, amount)	menghapus field yang tidak dibutuhkan

Database lookup (1)

Select values

Step name: **Select values**

Select & Alter Remove Meta-data

Fields:

#	Fieldname	Rename to
1	employeeNumber	streamEmployeeNumber
2	lastName	streamLastName
3	firstName	streamFirstName
4	extension	
5	email	
6	officeCode	
7	reportsTo	
8	jobTitle	streamJobTitle

data dari select values

Logging Execution History Step Metrics Performance Graph Metrics Preview data

First rows Last rows Off

#	streamEmployeeNumber	streamLastName	streamFirstName	streamJobTitle	streamLastNameBoss	streamFirstNameBoss	streamDate	streamAmount
1	1370	Hernandez	Gerard	Sales Rep	Bondur	Gerard	2004/10/19 00:00:00.000	6066.78
2	1370	Hernandez	Gerard	Sales Rep	Bondur	Gerard	2003/06/05 00:00:00.000	14571.44
3	1370	Hernandez	Gerard	Sales Rep	Bondur	Gerard	2004/12/18 00:00:00.000	1676.14
4	1166	Thompson	Leslie	Sales Rep	Bow	Anthony	2004/12/17 00:00:00.000	14191.12

data dengan tambahan id_dimEmployees dari dimEmployee

Mencari id_dimemployees berdasarkan employeeNumber

Database lookup (2)

Database lookup

Step name: **Database lookup**

Connection: conn_dw_destination Edit... New... Va

Lookup schema: dw_legendvehicle Br

Lookup table: dimemployees Br

Enable cache? ☐

Cache size in rows (0=cache): 0

Load all data from table ☐

The key(s) to look up the value(s):

#	Table field	Comparator	Field1	Field2
1	employeeNumber	=	streamEmployeeNumber	
2	firstName	=	streamFirstName	
3	lastName	=	streamLastName	
4	jobTitle	=	streamJobTitle	
5	boss_firstName	=	streamFirstNameBoss	
6	boss_lastName	=	streamLastNameBoss	

Values to return from the lookup table:

#	Field	New name	Default	Type
1	id_dimEmployees			None

data dari database lookup 1

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

First rows Last rows Off

#	streamEmployeeNumber	streamLastName	streamFirstName	streamJobTitle	streamLastNameBoss
1	1370	Hernandez	Gerard	Sales Rep	Bondur
2	1370	Hernandez	Gerard	Sales Rep	Bondur
3	1370	Hernandez	Gerard	Sales Rep	Bondur
4	1166	Thompson	Leslie	Sales Rep	Bow
5	1166	Thompson	Leslie	Sales Rep	Bow

data dengan tambahan id_dimdatedari dimDate

mencari id_dimDate berdasarkan payment

Database lookup (3)

Database lookup

Step name: Database lookup 2

Connection: conn_dw_destination

Lookup schema: dw_legendvehicle

Lookup table: dimdate

Enable cache? ☐

Cache size in rows (0=cache): 0

Load all data from table: ☐

The key(s) to look up the value(s):

#	Table field	Comparator	Field1	Field2
1	date	=	streamDate	

Values to return from the lookup table:

#	Field	New name	Default	Type
1	id_dimDate			None

data dari database lookup 2

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

First rows Last rows Off

#	streamEmployeeNumber	streamLastName	streamFirstName	streamJobTitle	streamLastNameBoss
1	1370	Hernandez	Gerard	Sales Rep	Bondur
2	1370	Hernandez	Gerard	Sales Rep	Bondur
3	1370	Hernandez	Gerard	Sales Rep	Bondur
4	1166	Thompson	Leslie	Sales Rep	Bow
5	1166	Thompson	Leslie	Sales Rep	Bow
6	1166	Thompson	Leslie	Sales Rep	Bow
7	1611	Fixter	Andy	Sales Rep	Patterson

data dengan tambahan id factomset jika sudah ada

mengecek
apakah
kombinasi
id_dimemploye
es dari
id_dimDate
sudah ada di
factomset

Filter rows

Database lookup

Step name

Database lookup 3

Connection

conn_dw_destination

Lookup schema

dw_legendvehicle

Lookup table

factomset

Enable cache?

☐

Cache size in rows (0=cache all rows)

0

Load all data from table

☐

The key(s) to look up the value(s):

#	Table field	Comparator	Field1	Field2
1	id_DimDate	=	id_dimDate	
2	id_DimEmployees	=	id_dimEmployees	
3	amount	=	streamAmount	

Values to return from the lookup table :

#	Field	New name	Default	Type
1	id_DimDate			None
2	id_DimEmployees			None
3	amount			None

data dari database lookup 3

Database lookup 3					
First rows					
#	streamEmployeeNumber	streamLastName	streamFirstName	streamJobTitle	streamLastNameBoss
1	1370	Hernandez	Gerard	Sales Rep	Bondur
2	1370	Hernandez	Gerard	Sales Rep	Bondur
3	1370	Hernandez	Gerard	Sales Rep	Bondur
4	1166	Thompson	Leslie	Sales Rep	Bow
5	1166	Thompson	Leslie	Sales Rep	Bow
6	1166	Thompson	Leslie	Sales Rep	Bow
7	1611	Fixter	Andy	Sales Rep	Patterson
8	1611	Fixter	Andy	Sales Rep	Patterson

data yang belum ada dalam factomset

menyaring data yang sudah ada agar tidak dimasukkan lagi

Table output

Filter rows

Step name

Filter rows

Send 'true' data to step:

Table output

Send 'false' data to step:

The condition:

id_DimDate_1 IS NULL

AND

id_DimEmployees_1 IS NULL

AND

amount IS NULL

data hasil filter rows

First rows

Last rows

Off

#	streamEmployeeNumber	streamLastName	streamFirstName	streamJobTitle	streamLastNameBoss
1	1370	Hernandez	Gerard	Sales Rep	Bondur
2	1370	Hernandez	Gerard	Sales Rep	Bondur
3	1370	Hernandez	Gerard	Sales Rep	Bondur
4	1166	Thompson	Leslie	Sales Rep	Bow
5	1166	Thompson	Leslie	Sales Rep	Bow
6	1166	Thompson	Leslie	Sales Rep	Bow
7	1611	Fixter	Andy	Sales Rep	Patterson
8	1611	Fixter	Andy	Sales Rep	Patterson
9	1611	Fixter	Andy	Sales Rep	Patterson
10	1611	Fixter	Andy	Sales Rep	Patterson

data masuk ke tabel factomset

memasukkan data baru ke dalam tabel OLAP

2. Jika proses itu di ulangi (di run kembali) apakah data akan redudant?

Tidak. Karena terdapat proses pengecekan (Database Lookup ke-3 dan Filter Rows) yang memastikan hanya data baru yang disimpan. Jadi, transformasi bersifat **idempotent** (tidak menyebabkan duplikasi walau dijalankan berulang).

D. Jobs

1. Buka Jobs pada **File - New - Jobs**

2. Gunakan 5 objects dan hububngkan sesuai urutan sebagai berikut:

- **Start**: Awal job
- **Transformation 1**: Menghasilkan data dimDate
- **Transformation 2**: Menghasilkan data dimEmployees
- **Transformation 3**: Menghasilkan factOmset
- **Success**: Tanda job selesai

Job menjalankan ketiga transformasi secara berurutan, memastikan bahwa seluruh proses ETL dijalankan otomatis dan efisien.




3. Konfigurasi dilakukan pada setiap transformation untuk mengambil file tranformation yang telah dibuat pada bagian A, B dan C.

File Transformation A


The screenshot shows the 'Transformation' configuration window. The 'Entry Name' field is set to 'Transformation'. The 'Transformation' field shows the path 'E:\KULIAH\Kuliah Semester 4\Data Warehouse\Transformation' with a 'Browse...' button. Below this are tabs for 'Options', 'Logging', 'Arguments', and 'Parameters'. The 'Options' tab is active, showing 'Run configuration:' set to 'Pentaho local'. Under the 'Execution' section, the following options are listed:

- ☐ Execute every input row
- ☐ Clear results rows before execution
- ☐ Clear results files before execution
- ☒ Wait for remote transformation to complete
- ☐ Follow local abort to remote transformation
- ☐ Suppress result data from remote transformation

File Transformation B

 Transformation

Entry Name:


Transformation:
 

Options

Logging

Arguments

Parameters

Run configuration:
 

Execution

☐ Execute every input row

☐ Clear results rows before execution

☐ Clear results files before execution

☒ Wait for remote transformation to complete

☐ Follow local abort to remote transformation

☐ Suppress result data from remote transformation

File Transformation C

Transformation

Entry Name:
Transformation 3

Transformation:
E:/KULIAH/Kuliah Semester 4/Data Warehouse/JS 3 C.ktr Browse...

Options | Logging | Arguments | Parameters

Run configuration:
Pentaho local

Execution

- ☐ Execute every input row
- ☐ Clear results rows before execution
- ☐ Clear results files before execution
- ☒ Wait for remote transformation to complete
- ☐ Follow local abort to remote transformation
- ☐ Suppress result data from remote transformation

4. Jalankan jobs tersebut ketika semua transformation telah dihubungkan dengan file tranformation yang telah dibuat sebelumnya.



Execution Results

Logging History Job metrics Metrics



2025/04/22 22:21:03 - Spoon - Starting job...
2025/04/22 22:21:03 - JS 3 D - Start of job execution
2025/04/22 22:21:03 - JS 3 D - Starting entry [Transformation]
2025/04/22 22:21:04 - Transformation - Using run configuration [Pentaho local]
2025/04/22 22:21:04 - Transformation 1Js3 - Dispatching started for transformation [Transformation 1Js3]
2025/04/22 22:21:04 - Table output.0 - Connected to database [conn_dw_destination] (commit=1000)
2025/04/22 22:21:04 - Generate rows.0 - Finished processing (I=0, O=0, R=0, W=1825, U=0, E=0)
2025/04/22 22:21:04 - Add sequence.0 - Finished processing (I=0, O=0, R=1825, W=1825, U=0, E=0)
2025/04/22 22:21:04 - Calculator.0 - Finished processing (I=0, O=0, R=1825, W=1825, U=0, E=0)
2025/04/22 22:21:04 - Select task.0 - Finished processing (I=0, O=0, R=1825, W=1825, U=0, E=0)

Tugas 4

1. Buka desain database dari dw_legendvehicle pada DBMS, bandingkan design tersebut dengan desain db OLTP legendVehicle pada jobsheet 2. analisislah dan ceritakan perbedaannya.

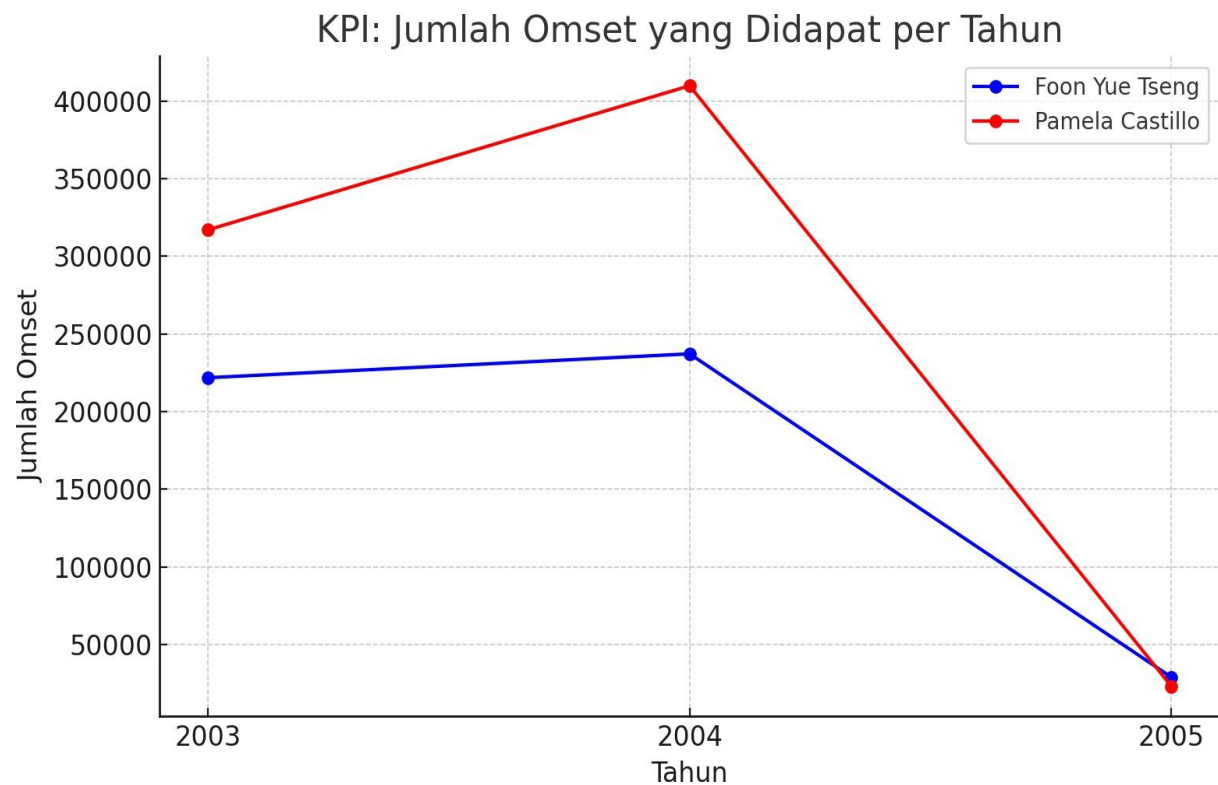
No.	Aspek	OLTP (Operational Database)	OLAP (Data Warehouse)
1	Tujuan	Untuk transaksi harian	Untuk analisis data historis jangka panjang
2	Struktur Data	Normalisasi tinggi, banyak relasi	Denormalisasi, fokus ke kemudahan analisis
3	Contoh Tabel	customers, orders, vehicles	dimDate, dimEmployees, factOmset
4	Performansi	Optimal untuk operasi CRUD cepat	Optimal untuk query analitik yang kompleks

2. Buatlah report pertahun untuk KPI "**Jumlah omset yang didapat**" pada **Foon Yue Tseng** dan **Pamela Castillo**. Serta gambarkan grafiknya (grafik garis).

Showing rows 0 - 1 (2 total, Query took 0.0037 seconds.)

```
SELECT emp.firstName AS Nama, SUM(CASE WHEN d.year = 2003 THEN f.amount ELSE 0
END) AS '2003', SUM(CASE WHEN d.year = 2004 THEN f.amount ELSE 0 END) AS '2004',
SUM(CASE WHEN d.year = 2005 THEN f.amount ELSE 0 END) AS '2005' FROM factOmsset f
JOIN dimEmployees emp ON f.id_dimEmployees = emp.id_dimEmployees JOIN dimdate d ON
f.id_dimDate = d.id_dimDate WHERE (emp.firstName = 'Foon Yue' AND emp.lastName =
'Tseng') OR (emp.firstName = 'Pamela' AND emp.lastName = 'Castillo') GROUP BY
```

Nama	2003	2004	2005
Foon Yue	221887.03	237255.26	29070.38
Pamela	317104.78	409910.07	23187.02



3. Jelaskan perbedaan query saat mendapatkan data pada nomor 2 dengan query pada saat Jobsheet 2!

Jobsheet 2 (OLTP)	Tugas ini (OLAP)
Query langsung ke tabel transaksi seperti orders, payments	Query ke factOmset, dimEmployees, dimDate
Banyak join antar tabel kompleks	Query lebih sederhana dan fokus pada analisis (agregasi)
Tidak fokus pada agregasi historis	Fokus ke waktu dan total per entitas

4. Simpulkan dengan bahasa sendiri, apa perbedaan OLTP dan OLAP?

- **OLTP (Online Transaction Processing):** Digunakan untuk mengelola aktivitas harian. Fokus pada efisiensi penyimpanan dan performa transaksi (insert/update/delete).
- **OLAP (Online Analytical Processing):** Digunakan untuk menganalisis data dalam jangka panjang. Dirancang agar performa analisis dan laporan cepat melalui struktur dim