

LAPORAN HASIL JOBSHEET 3 BAGIAN C DAN D



DISUSUN OLEH:

DIAJENG SEKAR ARUM

2341760070

PROGRAM STUDI D-IV SISTEM INFORMASI BISNIS

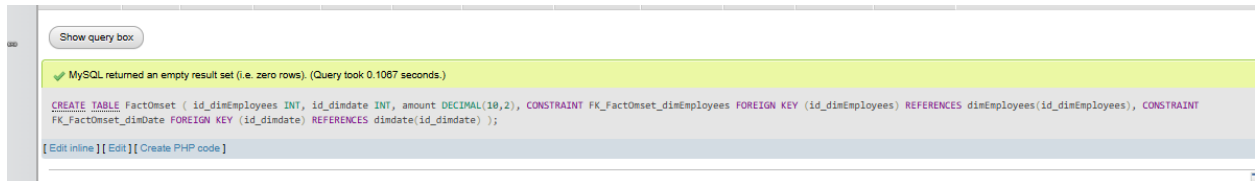
JURUSAN TEKNOLOGI INFORMASI

POLITEKNIK NEGERI MALANG

2024

FAKTA PEMBAYARAN

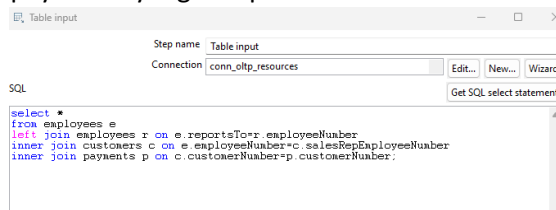
1. Buatlah sebuah tabel pada database OLAP yang telah dibuat (database dw_legendVehicle).
nama tabel: FactOmset
Field: id_dimEmployees int FK tabel dimEmployee
Field: id_dimDate int FK tabel dimDate
Field: amount decimal(10,2)



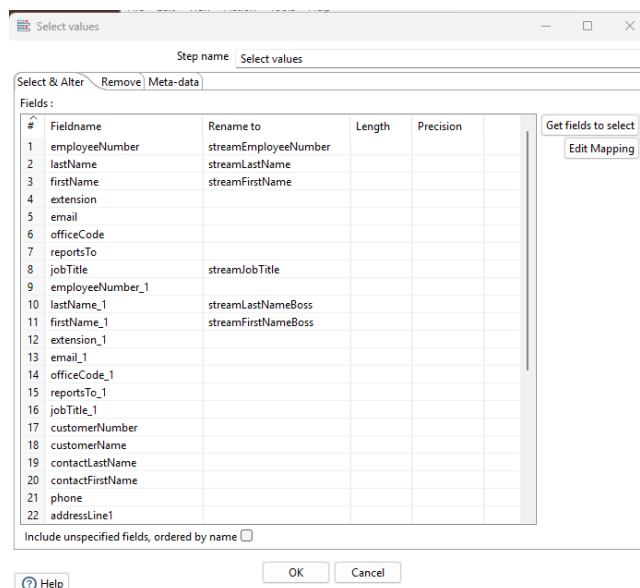
2. Buat Transformation baru pada PDI Spoon. Gunakan object-object yang dihubungkan sebagai berikut:



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.



4. Konfigurasi pada Select Values adalah untuk menghapus semua field kecuali employeeNumber, lastname dari employee, firstname dari employee, jobTitle dari employee, lastname dari manager, firstname dari manager, payment date dan amount.



Select values

Step nameSelect values

Select & AlterRemoveMeta-data

Fields:

| # | Fieldname | Rename to | Length | Precision |
|----|------------------------|--------------|--------|-----------|
| 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 | | | |
| 27 | country | | | |
| 28 | salesRepEmployeeNumber | | | |
| 29 | creditLimit | | | |
| 30 | customerNumber_1 | | | |
| 31 | checkNumber | | | |
| 32 | paymentDate | streamDate | | |
| 33 | amount | streamAmount | | |

Get fields to select

Edit Mapping

Include unspecified fields, ordered by name

Help

OK

Cancel

Select values

Step nameSelect values

Select & AlterRemoveMeta-data

Fields to remove:

| # | Fieldname |
|----|------------------------|
| 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 |

Get fields to remove

Help

OK

Cancel

- 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 nya.

Database lookup

Step nameDatabase lookup

Connection

conn_dw_destination

Edit...New...Wizard...

Lookup schema

dw_legendvehicle

Browse...

Lookup table

dimemployees

Browse...

Enable cache?

Cache size in rows (0=cache)

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 | | | |

Do not pass the row if the lookup fails

Fail on multiple results?

Order by

Help

OK

Cancel

Get Fields

Get lookup fields

6. 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.

The screenshot shows the 'Database lookup' configuration window for 'Database lookup 2'. The 'Connection' is 'conn_dw_destination', 'Lookup schema' is 'dw_legendvehicle', and 'Lookup table' is 'dimdate'. The 'Enable cache?' checkbox is unchecked. The 'Cache size in rows (0=cache)' is set to 0. The 'Load all data from table' checkbox is unchecked. The 'Key(s) to look up the value(s)' table has one row: 'date' compared with 'streamDate'. The 'Values to return from the lookup table' table has one row: 'id_dimDate'. The 'Do not pass the row if the lookup fails' checkbox is unchecked, and 'Fail on multiple results?' is also unchecked. The 'Order by' field is empty.

| # | Table field | Comparator | Field1 | Field2 |
|---|-------------|------------|------------|--------|
| 1 | date | = | streamDate | |

| # | Field | New name | Default | Type |
|---|------------|----------|---------|------|
| 1 | id_dimDate | | | |

7. 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 data tidak akan dimasukkan dalam tabel factOmset.

The screenshot shows the 'Database lookup' configuration window for 'Database lookup 3'. The 'Connection' is 'conn_dw_destination', 'Lookup schema' is 'dw_legendvehicle', and 'Lookup table' is 'factomset'. The 'Enable cache?' checkbox is unchecked. The 'Cache size in rows (0=cache)' is set to 0. The 'Load all data from table' checkbox is unchecked. The 'Key(s) to look up the value(s)' table has three rows: 'id_dimdate' compared with 'id_dimDate', 'id_dimEmployees' compared with 'id_dimEmployees', and 'amount' compared with 'streamAmount'. The 'Values to return from the lookup table' table has three rows: 'id_dimEmployees', 'id_dimdate', and 'amount'. The 'Do not pass the row if the lookup fails' checkbox is unchecked, and 'Fail on multiple results?' is also unchecked. The 'Order by' field is empty.

| # | Table field | Comparator | Field1 | Field2 |
|---|-----------------|------------|-----------------|--------|
| 1 | id_dimdate | = | id_dimDate | |
| 2 | id_dimEmployees | = | id_dimEmployees | |
| 3 | amount | = | streamAmount | |

| # | Field | New name | Default | Type |
|---|-----------------|----------|---------|------|
| 1 | id_dimEmployees | | | |
| 2 | id_dimdate | | | |
| 3 | amount | | | |

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

The screenshot shows the 'Filter rows' configuration window. The 'Step name' is 'Filter rows'. The 'Send 'true' data to step:' and 'Send 'false' data to step:' fields are empty. The 'The condition:' section shows a logical expression: 'id_dimdate_1 IS NULL AND (id_dimEmployees_1 IS NULL AND amount IS NULL)'. The 'To edit a subcondition, simply click on it' text is present, along with a '+' button to add more conditions.

```

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 kedalam field pada tabel factOmset.

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_dimEmployees | id_dimEmployees |
| 3 | amount | streamAmount |

Get fields
Enter field mapping

Help OK Cancel SQL

10. Jalankan proses transformation tersebut.

Table input → Select values → Database lookup → Database lookup 2 → Database lookup 3 → Filter rows → Table output

Execution Results

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

First rows | Last rows | Off

| # | streamEmployeeNumber | streamLastName | streamFirstName | streamJobTitle | streamLastNameBoss | streamFirstNameBoss | streamDate |
|---|----------------------|----------------|-----------------|----------------|--------------------|---------------------|-------------------------|
| 1 | 1370 | Hernandez | Gerard | Sales Rep | Bondur | Gerard | 2004/10/19 00:00:00.000 |
| 2 | 1370 | Hernandez | Gerard | Sales Rep | Bondur | Gerard | 2003/06/05 00:00:00.000 |
| 3 | 1370 | Hernandez | Gerard | Sales Rep | Bondur | Gerard | 2004/12/18 00:00:00.000 |
| 4 | 1166 | Thompson | Leslie | Sales Rep | Bow | Anthony | 2004/12/17 00:00:00.000 |
| 5 | 1166 | Thompson | Leslie | Sales Rep | Bow | Anthony | 2003/06/06 00:00:00.000 |
| 6 | 1166 | Thompson | Leslie | Sales Rep | Bow | Anthony | 2004/08/20 00:00:00.000 |

Hasil pada phpMyAdmin

Showing rows 0 - 24 (25 total. Query took 0.0007 seconds)

SELECT * FROM factomset

Profiling | Edit inline | Edit | Explain SQL | Create PHP code | Refresh

1 | > >> | Show all | Number of rows: 25 | Filter rows: Search this table | Sort by key: Name

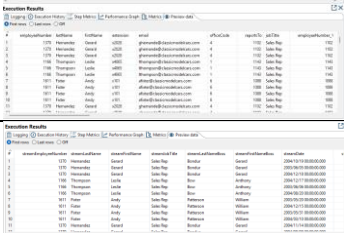
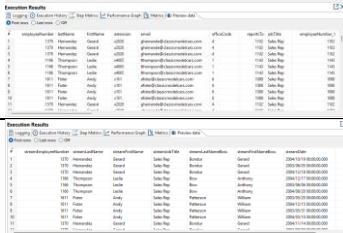
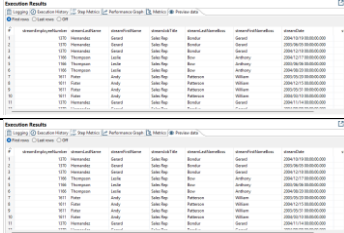
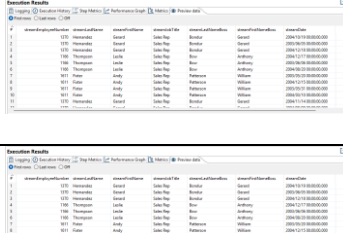
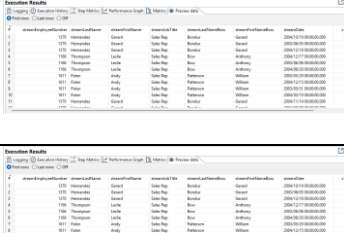
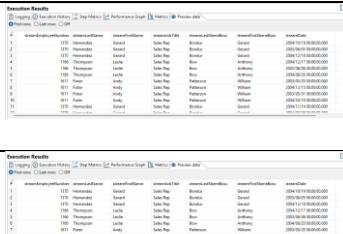
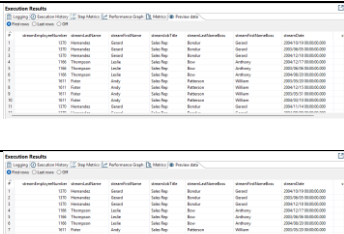
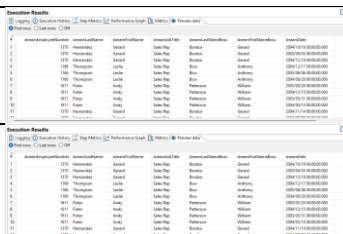
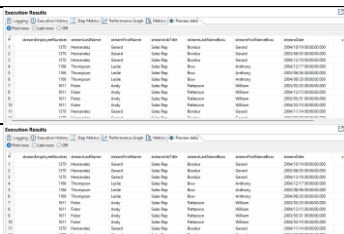
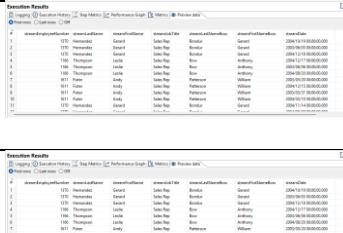
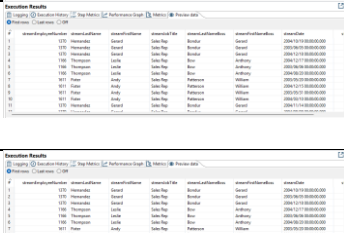
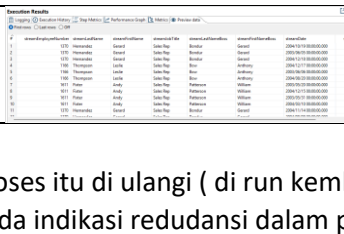
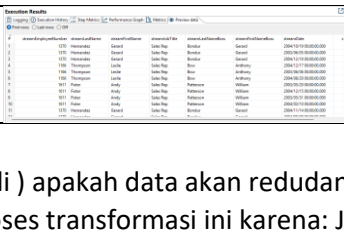
Extra options

| id_dimEmployees | id_dimdate | amount |
|-----------------|------------|-----------|
| 14 | 050 | 6080.78 |
| 14 | 150 | 14871.44 |
| 14 | 710 | 1678.14 |
| 8 | 717 | 14191.52 |
| 8 | 157 | 32641.96 |
| 9 | 550 | 33347.95 |
| 10 | 140 | 45894.03 |
| 10 | 715 | 82291.22 |
| 10 | 151 | 7893.08 |
| 10 | 435 | 44694.74 |
| 14 | 604 | 19501.82 |
| 14 | 580 | 47624.16 |
| 14 | 734 | 46623.67 |
| 17 | 47 | 80219.96 |
| 17 | 301 | 1491.38 |
| 17 | 674 | 17678.32 |
| 17 | 680 | 34838.14 |
| 7 | 104 | 101942.40 |

Console

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 prosesnya.

| Proses Objek | SS Data Input | SS Data Output | Keterangan |
|---------------------|---|--|---|
| Table Input | — |  | Mengambil seluruh data dari database sumber. |
| Select values |  |  | Menyaring kolom yang hanya diperlukan untuk proses selanjutnya. |
| Database lookup (1) |  |  | Menambahkan data dari tabel lain berdasarkan key tertentu. |
| Database lookup (2) |  |  | Lookup kedua, memperkaya data dengan informasi tambahan. |
| Database lookup (3) |  |  | Lookup ketiga untuk melengkapi data dari tabel lain. |
| Filter rows |  |  | Menyaring data berdasarkan kondisi tertentu (misal status aktif). |
| Table Output |  |  | Menyimpan data akhir ke dalam tabel target di database. |

2. Jika proses itu di ulangi (di run kembali) apakah data akan redundant? tidak ada indikasi redundansi dalam proses transformasi ini karena: Jumlah input sama dengan output di setiap langkah transformasi:
Table input.0 membaca 273 record.
Database lookup memproses 273 record tanpa error.
Filter rows.0 memproses 273 record tetapi tidak menulis ulang (W=0).

JOBS

1. Buka Jobs pada File - New - Jobs
2. Gunakan 5 objects dan hububngkan sesuai urutan sebagai berikut:



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

- Bagian a

The screenshot shows the 'Transformation' configuration dialog box. The 'Entry Name' is 'Transformation'. The 'Transformation' file path is 'D:\kuliah\sem4\data warehouse\pentaho\yb3-1.ktr'. The 'Options' tab is selected, showing 'Run configuration' set to 'Pentaho local'. Under 'Execution', the following options are visible:

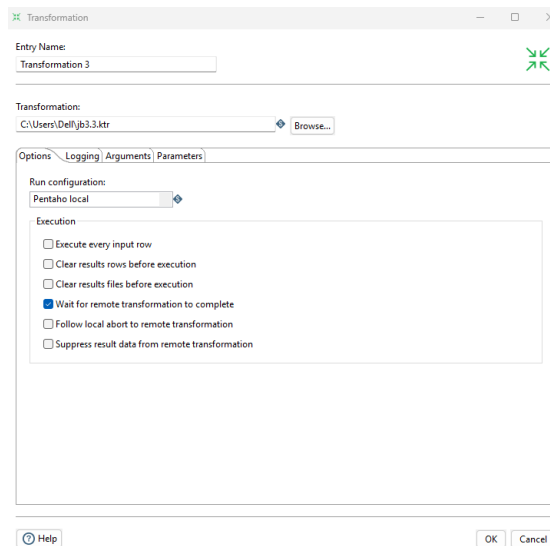
- ☐ Execute every input row
- ☐ Clear results rows before execution
- ☐ Clear results files before execution
- ☒ Wait for remote transformation to complete
- ☐ Follow local abort to remote transformation
- ☐ Suppress result data from remote transformation

- Bagian b

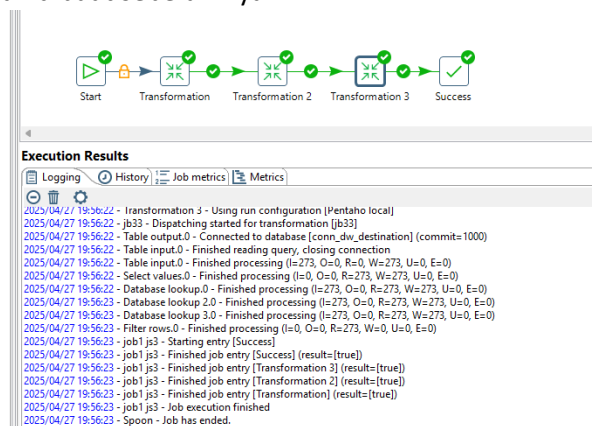
The screenshot shows the 'Transformation' configuration dialog box for 'Transformation 2'. The 'Entry Name' is 'Transformation 2'. The 'Transformation' file path is 'D:\kuliah\sem4\data warehouse\pentaho\yb 3.2.ktr'. The 'Options' tab is selected, showing 'Run configuration' set to 'Pentaho local'. Under 'Execution', the following options are visible:

- ☐ 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

- Bagian c



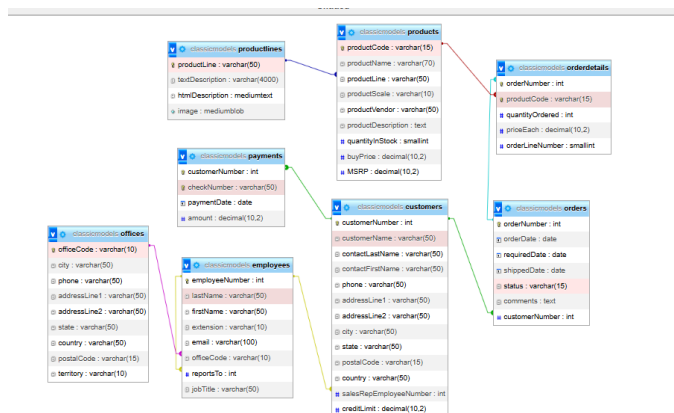
4. Jalankan jobs tersebut ketika semua transformation telah dihubungkan dengan file tranformation yang telah dibuat sebelumnya.



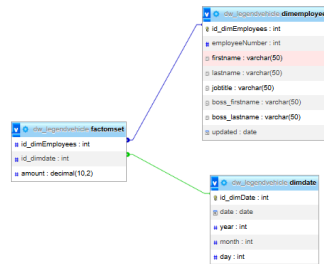
TUGAS 4

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

- Oltp



- Olap



| OLAP (dw_legendvehicle) | OLTP (legendVehicle) |
|--|---|
| Digunakan untuk analisis data dan pembuatan laporan bisnis | Digunakan untuk memproses transaksi harian |
| Struktur tabel lebih sederhana dan ringkas (hasil denormalisasi) | Struktur tabel kompleks dan banyak (tingkat normalisasi tinggi) |
| Berisi data yang sudah dirangkum dan bersifat historis | Berisi data transaksi yang detail dan terkini |
| Memanfaatkan basis data multidimensional | Menggunakan basis data relasional |

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

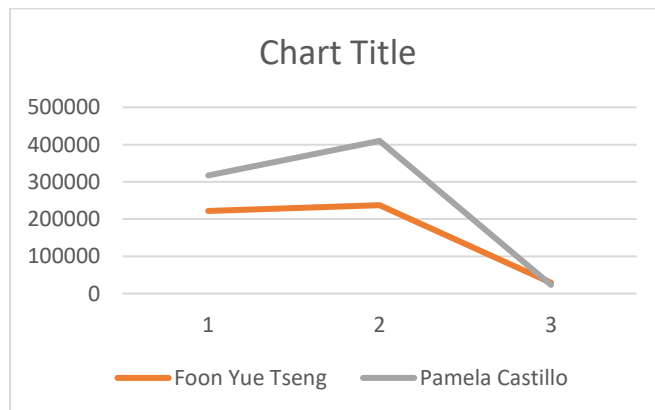
```

1 SELECT
2   CONCAT(e.firstname, ' ', e.lastname) AS name,
3   d.year AS Tahun,
4   SUM(t.amount) AS Total_Omset
5 FROM factomset t
6 JOIN dmemployees e ON t.id_dmEmployees = e.id_dmEmployees
7 JOIN dmdate d ON t.id_dmDate = d.id_dmDate -- Menggunakan id_dmDate sebagai foreign key
8 WHERE CONCAT(e.firstname, ' ', e.lastname) IN ('Foon Yue Tseng', 'Pamela Castillo')
9 GROUP BY name, d.year
10 ORDER BY name, Tahun
11
  
```

Hasil:

| name | 1 | Tahun | 2 | Total_Omset |
|-----------------|---|-------|---|-------------|
| Foon Yue Tseng | | 2003 | | 221887.03 |
| Foon Yue Tseng | | 2004 | | 237255.26 |
| Foon Yue Tseng | | 2005 | | 29070.38 |
| Pamela Castillo | | 2003 | | 317104.78 |
| Pamela Castillo | | 2004 | | 409910.07 |
| Pamela Castillo | | 2005 | | 23187.02 |

Grafik:



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

- Query nomor 2

```
1 SELECT
2   CONCAT(e.firstname, ' ', e.lastname) AS name,
3   d.year AS Tahun,
4   SUM(t.amount) AS Total_Omset
5 FROM factOmset t
6 JOIN dimEmployees e ON t.id_dimEmployees = e.id_dimEmployees
7 JOIN dimDate d ON t.id_dimDate = d.id_dimDate -- Menggunakan id_dimDate sebagai foreign key
8 WHERE CONCAT(e.firstname, ' ', e.lastname) IN ('Foon Yue Tseng', 'Pamela Castillo')
9 GROUP BY name, d.year
10 ORDER BY name, Tahun;
11
```

- Query js 2

```
1 SELECT
2   e.employeeNumber,
3   CONCAT(e.firstName, ' ', e.lastName) AS employee_name,
4   YEAR(o.orderDate) AS tahun,
5   SUM(od.quantityOrdered * od.priceEach) AS total_omset
6 FROM employees e
7 JOIN customers c ON e.employeeNumber = c.salesRepEmployeeNumber
8 JOIN orders o ON c.customerNumber = o.customerNumber
9 JOIN orderdetails od ON o.orderNumber = od.orderNumber
10 WHERE e.firstName IN ('Foon Yue', 'Pamela') AND e.lastName IN ('Tseng', 'Castillo')
11 GROUP BY e.employeeNumber, tahun
12 ORDER BY tahun, employee_name;
13
```

- Perbedaan

| QUERY NO 2 (OLAP) | QUERY JOBSHEET 2 (OLTP) |
|--|--|
| Memanfaatkan OLAP (dw_legendvehicle) menggunakan tabel factOmset, dimEmployees, dan dimDate | Menggunakan OLTP (legendVehicle) dengan tabel orders, orderdetails, dan payments |
| Melakukan join dengan dimEmployees untuk mengambil nama karyawan serta dimDate untuk mengambil tahun transaksi | Melakukan join antara orders, orderdetails, dan customers untuk memperoleh rincian transaksi |
| Menyajikan laporan tahunan berupa total omzet per karyawan dalam bentuk tabel | Menampilkan daftar transaksi lengkap dengan jumlah item, harga satuan, dan total pembayaran |

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

- OLTP** (Online Transaction Processing) adalah sistem yang digunakan untuk mengelola transaksi harian secara cepat dan akurat. Data di OLTP bersifat detail, real-time, dan tersimpan dalam tabel-tabel yang sangat terstruktur (ternormalisasi) untuk menghindari duplikasi dan menjaga konsistensi data.
- OLAP** (Online Analytical Processing), sebaliknya, digunakan untuk menganalisis data dalam jumlah besar, biasanya untuk keperluan laporan bisnis atau pengambilan keputusan. Data di OLAP sudah diringkas, bersifat historis, dan tabel-tabelnya dibuat lebih sederhana (denormalisasi) untuk mempercepat proses analisis.