

# **LAPORAN UJIAN AKHIR SEMESTER**

## **DATA WAREHOUSE**



Disusun Oleh:

Maharani Wirawan	2341760111
Muhammad Fairuz Daffa A.	2341760079
Renald Agustinus	2341760090
Sufyan Dwi Bagaskara	2341760110

**Jurusan Teknologi Informasi**

**D4 Sistem Informasi Bisnis**

**Politeknik Negeri Malang**

**2025**

**Dataset:** Employee/HR Dataset (All in One)

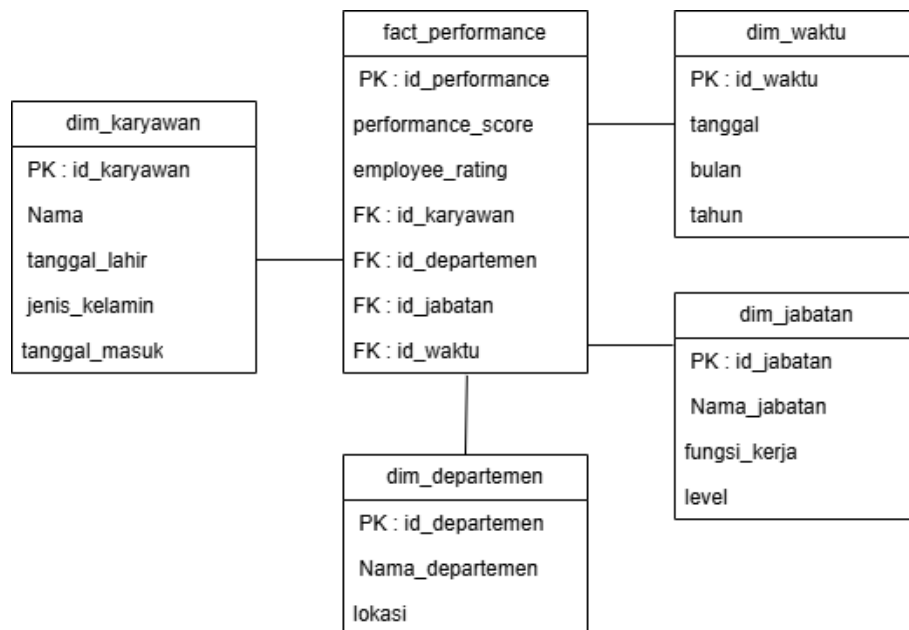
<https://www.kaggle.com/datasets/ravindrasinghrana/employeeedataset>

## Study Case:

Analisis kinerja (performance) karyawan

## 1. Skema Star Schema

Gambar berikut menggambarkan desain data warehouse bertipe *Star Schema*:



### Tabel Dimensi:

- 1) **dim\_karyawan**  
Menyimpan data dasar karyawan seperti nama, tanggal lahir, jenis kelamin, dan tanggal masuk kerja.
- 2) **dim\_departemen**  
Menyimpan informasi departemen dan lokasi.
- 3) **dim\_jabatan**  
Menyimpan nama jabatan, fungsi kerja, dan level jabatan.
- 4) **dim\_waktu**  
Menyimpan data waktu dalam bentuk tanggal, bulan, dan tahun.

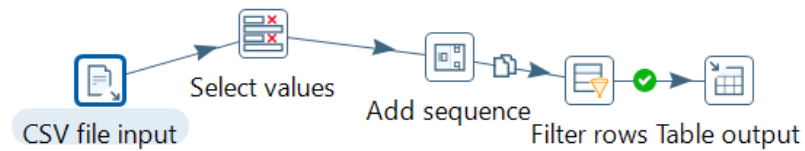
### Tabel Fakta:

- **fact\_performance**  
Menyimpan skor kinerja (**performance\_score**) dan rating pegawai (**employee\_rating**) yang berelasi dengan dimensi di atas.
-

## 2. Proses ETL (Extract – Transform – Load)

### A. Transformasi 1 – dim\_karyawan

- **Input:** File `employee_data.csv`
- **Langkah-langkah:**

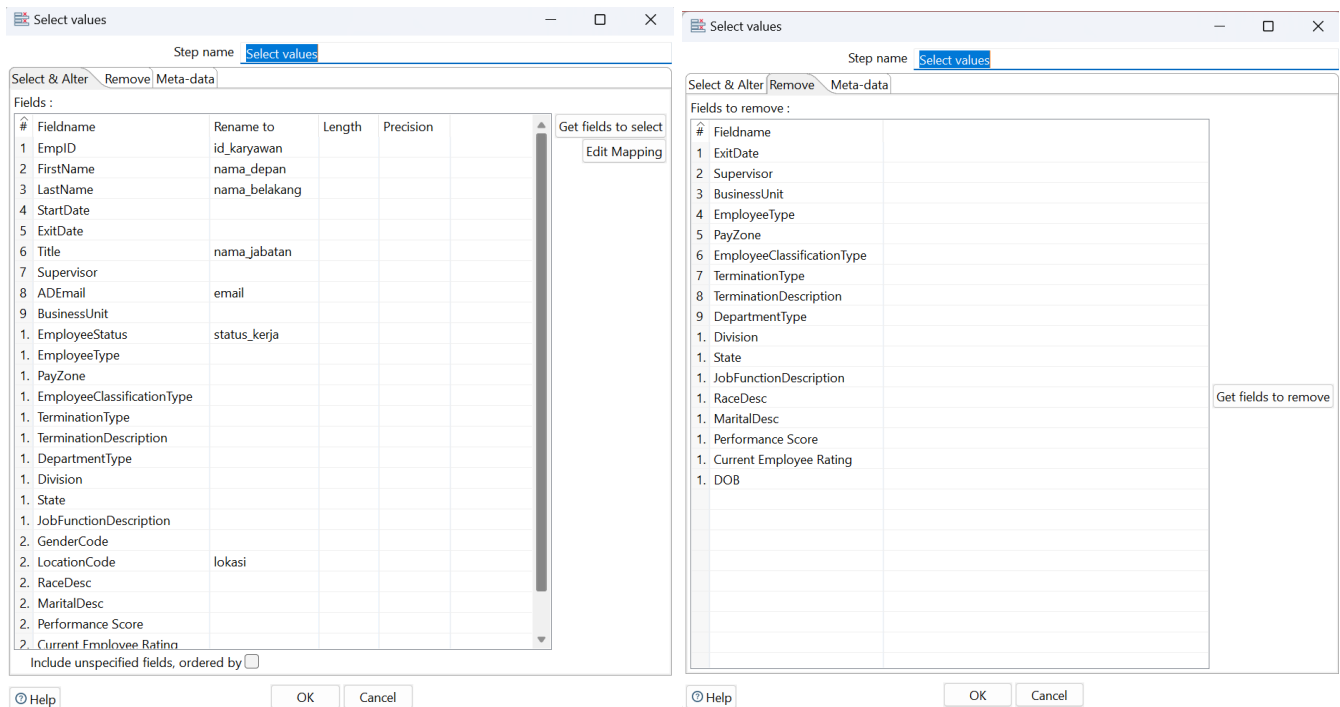


1) Membaca file CSV menggunakan input file.

The screenshot shows the 'CSV file input' configuration window. The 'Step name' is 'CSV file input'. The 'Filename' is '\$(InternalEntry.CurrentDirectory)/employee\_data.csv'. The 'Delimiter' is ',' and the 'Enclosure' is '"'. The 'NIO buffer size' is 50000. The 'Lazy conversion?' checkbox is checked. The 'Header row present?' checkbox is checked. The 'Add filename to result' checkbox is unchecked. The 'The row number field name (optional)' field is empty. The 'Running in parallel?' checkbox is unchecked. The 'New line possible in fields?' checkbox is unchecked. The 'Format' is 'mixed'. The 'File encoding' is 'UTF-8'. Below the configuration fields is a table with columns: #, Name, Type, Format, Length, Precision, Currency, Decimal, Group, Trim type.

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Trim type
1	EmpID	Integer	#	15	0	USD	-	,	none
2	FirstName	String		11		USD	-	,	none
3	LastName	String		9		USD	-	,	none
4	StartDate	Date	yyyy-MM-dd			USD	-	,	none
5	EndDate	String		9		USD	-	,	none
6	Title	String		23		USD	-	,	none
7	Supervisor	String		18		USD	-	,	none
8	ADEmail	String		34		USD	-	,	none
9	BusinessUnit	String		4		USD	-	,	none
1.	EmployeeStatus	String		12		USD	-	,	none
1.	EmployeeType	String		9		USD	-	,	none
1.	PayZone	String		6		USD	-	,	none
1.	EmployeeClassificationType	String		9		USD	-	,	none
1.	TerminationType	String		11		USD	-	,	none
1.	TerminationDescription	String		49		USD	-	,	none

2) Melakukan pembersihan dan seleksi field yang diperlukan.



3) Database Lookup untuk mencegah duplikasi (pengecekan ID>Nama).

The image shows the 'Add sequence' dialog box. It contains the following fields and options:

- Step name: Add sequence
- Name of value: id\_karyawan
- Use a database to generate the sequence:
  - Use DB to get sequence? ☐
  - Connection: dim\_employee
  - Schema name:
  - Sequence name: SEQ\_
- Use a transformation counter to generate the sequence:
  - Use counter to calculate sequence? ☒
  - Counter name (optional):
  - Start at value: 0
  - Increment by: 1
  - Maximum value: 99999999

4) Filter Rows untuk menyaring data baru.

**Filter rows**

Step name:

Send 'true' data to step:

Send 'false' data to step:

The condition:

(String)

##### 5) Output ke tabel dim\_karyawan.

**Table output**

Step name:

Connection:

Target schema:

Target table:

Commit size:

Truncate table: ☒

Ignore insert errors: ☐

Specify database fields: ☒

Main options | Database fields

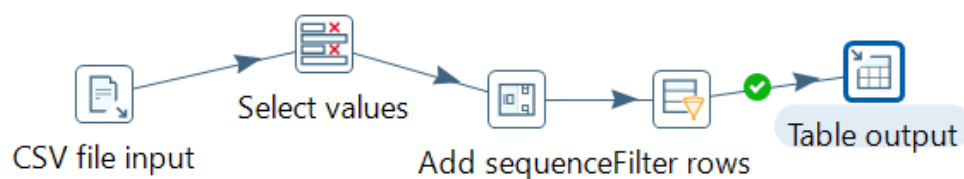
Fields to insert:

#	Table field	Stream field
1	id_karyaw...	id_karyawan
2	first_name	nama_depan
3	last_name	nama_belak...
4	StartDate	StartDate
5	email	email
6	status_kerja	status_kerja
7	jenis_kela...	GenderCode

Server: localhost:3306 » Database: dimemployee » Table: dim\_karyawan

	first_name	last_name	jenis_kelamin	status_kerja	StartDate	email	id_karyawan
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Susan	Exantus	Female	Active	2019-08-29	susan.exantus@bilearner.com	1001
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Sandra	Martin	Female	Active	2022-12-12	sandra.martin@bilearner.com	1002
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Keyla	Del Bosque	Female	Active	2023-03-08	keyla.del bosque@bilearner.com	1003
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Andrew	Szabo	Male	Active	2020-05-29	andrew.szabo@bilearner.com	1004
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Luke	Patronick	Male	Active	2022-09-16	luke.patronick@bilearner.com	1005
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Colby	Andreola	Female	Active	2021-08-15	colby.andreola@bilearner.com	1006
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Judith	Carabbio	Female	Active	2019-08-28	judith.carabbio@bilearner.com	1008
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Adell	Saada	Female	Active	2019-06-13	adell.saada@bilearner.com	1009
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Jasmin	Shah	Male	Active	2020-05-13	jasmin.shah@bilearner.com	1013
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Kole	Quinn	Male	Active	2022-12-31	kole.quinn@bilearner.com	1014
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Ansley	Jackson	Female	Active	2018-12-03	ansley.jackson@bilearner.com	1015
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Jayda	Reese	Female	Active	2020-02-21	jayda.reese@bilearner.com	1016
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Julien	Whitehead	Female	Active	2020-09-07	julien.whitehead@bilearner.com	1017
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Alan	Haynes	Female	Active	2020-06-19	alan.haynes@bilearner.com	1018
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Frederick	Howe	Male	Active	2022-09-19	frederick.howe@bilearner.com	1022
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Nickolas	Davila	Male	Active	2018-11-09	nickolas.davila@bilearner.com	1023
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Kasey	Boyer	Female	Active	2018-12-30	kasey.boyer@bilearner.com	1024
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Giovanni	Jenkins	Female	Active	2023-05-11	giovanni.jenkins@bilearner.com	1025
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Alexis	Moss	Female	Active	2022-11-03	alexis.moss@bilearner.com	1026
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Joanna	Murphy	Male	Active	2020-06-14	joanna.murphy@bilearner.com	1027
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Brittany	Hurley	Female	Active	2020-04-28	brittany.hurley@bilearner.com	1029
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Niko	Hamilton	Male	Active	2019-10-29	niko.hamilton@bilearner.com	1030
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Ruben	Martinez	Male	Active	2021-04-28	ruben.martinez@bilearner.com	1031
<input type="checkbox"/> Edit <input type="checkbox"/> Copy <input type="checkbox"/> Delete	Kiley	Haynes	Male	Active	2022-04-19	kiley.haynes@bilearner.com	1032

## B. Transformasi 2 – dim\_departemen



- **Input:** Query OLTP
- **Langkah-langkah:**

1) Mengambil data departemen (id, tipe, lokasi).

CSV file input

Step name: CSV file input

Filename: \$(Internal.Entry.Current.Directory)/employee\_data.csv Browse...

Delimiter: , Insert TAB

Enclosure: "

NIO buffer size: 50000

Lazy conversion? ☒

Header row present? ☒

Add filename to result ☐

The row number field name (optional):

Running in parallel? ☐

New line possible in fields? ☐

Format: mixed

File encoding:

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Trim type
1	EmpID	Integer	#	15	0	USD	.	,	none
2	FirstName	String		11		USD	.	,	none
3	LastName	String		9		USD	.	,	none
4	StartDate	Date	yyyy-MM-dd			USD	.	,	none
5	ExitDate	String		9		USD	.	,	none
6	Title	String		23		USD	.	,	none
7	Supervisor	String		18		USD	.	,	none
8	ADEmail	String		34		USD	.	,	none
9	BusinessUnit	String		4		USD	.	,	none
1.	EmployeeStatus	String		12		USD	.	,	none
1.	EmployeeType	String		9		USD	.	,	none
1.	PayZone	String		6		USD	.	,	none
1.	EmployeeClassificationType	String		9		USD	.	,	none
1.	TerminationType	String		11		USD	.	,	none
1.	TerminationDescription	String		49		USD	.	,	none

Help OK Get Fields Preview Cancel

2) Select values untuk menyaring field.

Select values

Step name: Select values

Select & Alter Remove Meta-data

Fields :

#	Fieldname	Rename to	Length	Precision
1	DepartmentType			
2	LocationCode			

Get fields to select Edit Mapping

3) Database Lookup + Filter Rows.

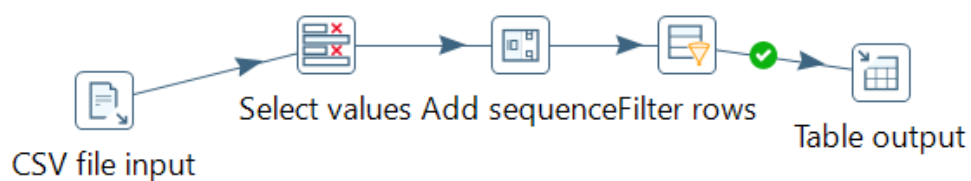




Server: localhost:3306 » Database: dimemployee » Table: dim\_departemen

	id_departemen	tipe_departemen	lokasi
<input type="checkbox"/> Edit Copy Delete	1	Production	34904
<input type="checkbox"/> Edit Copy Delete	2	Production	6593
<input type="checkbox"/> Edit Copy Delete	3	Sales	2330
<input type="checkbox"/> Edit Copy Delete	4	Sales	58782
<input type="checkbox"/> Edit Copy Delete	5	Sales	33174
<input type="checkbox"/> Edit Copy Delete	6	Sales	6050
<input type="checkbox"/> Edit Copy Delete	7	Sales	90007
<input type="checkbox"/> Edit Copy Delete	8	Sales	97756
<input type="checkbox"/> Edit Copy Delete	9	Sales	78789
<input type="checkbox"/> Edit Copy Delete	10	Sales	78207
<input type="checkbox"/> Edit Copy Delete	11	Sales	46204
<input type="checkbox"/> Edit Copy Delete	12	Sales	30428
<input type="checkbox"/> Edit Copy Delete	13	Sales	80820
<input type="checkbox"/> Edit Copy Delete	14	Sales	40220
<input type="checkbox"/> Edit Copy Delete	15	Sales	89139
<input type="checkbox"/> Edit Copy Delete	16	IT/IS	2810
<input type="checkbox"/> Edit Copy Delete	17	IT/IS	2621
<input type="checkbox"/> Edit Copy Delete	18	Sales	44553
<input type="checkbox"/> Edit Copy Delete	19	Sales	5360
<input type="checkbox"/> Edit Copy Delete	20	Sales	16325
<input type="checkbox"/> Edit Copy Delete	21	Sales	43481
<input type="checkbox"/> Edit Copy Delete	22	Sales	50705
<input type="checkbox"/> Edit Copy Delete	23	Sales	5168
<input type="checkbox"/> Edit Copy Delete	24	Sales	11765

### C. Transformasi 3 – dim\_jabatan



- **Langkah:**
  - 1) Proses serupa dengan departemen.

CSV file input

Step name: CSV file input

Filename: \$(Internal.Entry.Current.Directory)/employee\_data.csv Browse...

Delimiter: , Insert TAB

Enclosure: "

NIO buffer size: 50000

Lazy conversion? ☒

Header row present? ☒

Add filename to result ☐

The row number field name (optional):

Running in parallel? ☐

New line possible in fields? ☐

Format: mixed

File encoding:

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Trim type
1	EmpID	Integer	#	15	0	USD	.		none
2	FirstName	String		11		USD	.		none
3	LastName	String		9		USD	.		none
4	StartDate	Date	yyyy-MM-dd			USD	.		none
5	EndDate	String		9		USD	.		none
6	Title	String		23		USD	.		none
7	Supervisor	String		18		USD	.		none
8	ADEmail	String		34		USD	.		none
9	BusinessUnit	String		4		USD	.		none
1.	EmployeeStatus	String		12		USD	.		none
1.	EmployeeType	String		9		USD	.		none
1.	PayZone	String		6		USD	.		none
1.	EmployeeClassificationType	String		9		USD	.		none
1.	TerminationType	String		11		USD	.		none
1.	TerminationDescription	String		49		USD	.		none

Help OK Get Fields Preview Cancel

Select values

Step name: Select values

Select & Alter Remove Meta-data

Fields:

#	Fieldname	Rename to	Length	Precision
1	Title			
2	BusinessUnit			
3	EmployeeType			

Get fields to select Edit Mapping

Add sequence

Step name: Add sequence

Name of value: id\_jabatan

Use a database to generate the sequence

Use DB to get sequence? ☐

Connection: Dim\_jabatan Edit... New... Wizard...

Schema name:  Schemas...

Sequence name: SEQ\_ Sequences...

Use a transformation counter to generate the sequence

Use counter to calculate sequence? ☒

Counter name (optional):

Start at value: 1

Increment by: 1

Maximum value: 999999999

Help OK Cancel

**Filter rows**

Step name:

Send 'true' data to step:

Send 'false' data to step:

The condition:

+

id\_jabatan IS NOT NULL

AND

Title IS NOT NULL

AND

BusinessUnit IS NOT NULL

AND

EmployeeType IS NOT NULL

2) Menyimpan data jabatan ke tabel dim\_jabatan.

**Table output**

Step name:

Connection:

Target schema:

Target table:

Commit size:

Truncate table: ☐

Ignore insert errors: ☐

Specify database fields: ☒

**Main options** | **Database fields**

Partition data over tables: ☐

Partitioning field:

Partition data per month: ☒

Partition data per day: ☐

Use batch update for inserts: ☒

Is the name of the table defined in a field? ☐

Field that contains name of table:

Store the tablename field: ☒

Return auto-generated key: ☐

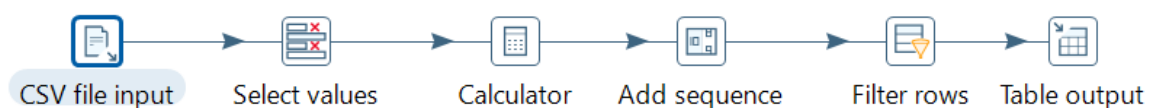
Name of auto-generated key field:

Server: localhost:3306 » Database: dimemployee » Table: dim\_jabatan

	id_jabatan	nama_jabatan	fungsi_kerja	level
<input type="checkbox"/> Edit Copy Delete	1	Production Technician I	CCDR	Contract
<input type="checkbox"/> Edit Copy Delete	2	Production Technician I	EW	Contract
<input type="checkbox"/> Edit Copy Delete	3	Area Sales Manager	PL	Full-Time
<input type="checkbox"/> Edit Copy Delete	4	Area Sales Manager	CCDR	Contract
<input type="checkbox"/> Edit Copy Delete	5	Area Sales Manager	TNS	Contract
<input type="checkbox"/> Edit Copy Delete	6	Area Sales Manager	BPC	Contract
<input type="checkbox"/> Edit Copy Delete	7	Area Sales Manager	WBL	Full-Time
<input type="checkbox"/> Edit Copy Delete	8	Area Sales Manager	CCDR	Contract
<input type="checkbox"/> Edit Copy Delete	9	Area Sales Manager	NEL	Contract
<input type="checkbox"/> Edit Copy Delete	10	Area Sales Manager	BPC	Part-Time
<input type="checkbox"/> Edit Copy Delete	11	Area Sales Manager	SVG	Contract
<input type="checkbox"/> Edit Copy Delete	12	Area Sales Manager	MSC	Full-Time
<input type="checkbox"/> Edit Copy Delete	13	Area Sales Manager	EW	Full-Time
<input type="checkbox"/> Edit Copy Delete	14	Area Sales Manager	CCDR	Contract
<input type="checkbox"/> Edit Copy Delete	15	Area Sales Manager	BPC	Part-Time
<input type="checkbox"/> Edit Copy Delete	16	Area Sales Manager	PYZ	Full-Time
<input type="checkbox"/> Edit Copy Delete	17	Area Sales Manager	WBL	Full-Time
<input type="checkbox"/> Edit Copy Delete	18	Area Sales Manager	NEL	Contract
<input type="checkbox"/> Edit Copy Delete	19	Area Sales Manager	PL	Contract
<input type="checkbox"/> Edit Copy Delete	20	Area Sales Manager	BPC	Part-Time
<input type="checkbox"/> Edit Copy Delete	21	Area Sales Manager	CCDR	Part-Time
<input type="checkbox"/> Edit Copy Delete	22	Area Sales Manager	SVG	Full-Time
<input type="checkbox"/> Edit Copy Delete	23	Area Sales Manager	EW	Contract
<input type="checkbox"/> Edit Copy Delete	24	Area Sales Manager	MSC	Contract

Console

#### D. Transformasi 4 – dim\_waktu



### 1) Generate Rows untuk 5 tahun (1825 hari).

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Trim type
1	EmpID	Integer	#	15	0	USD	.	,	none
2	FirstName	String		11		USD	.	,	none
3	LastName	String		9		USD	.	,	none
4	StartDate	Date	yyyy-MM-dd			USD	.	,	none
5	ExitDate	String		9		USD	.	,	none
6	Title	String		23		USD	.	,	none
7	Supervisor	String		18		USD	.	,	none
8	ADEmail	String		34		USD	.	,	none
9	BusinessUnit	String		4		USD	.	,	none
1.	EmployeeStatus	String		12		USD	.	,	none
1.	EmployeeType	String		9		USD	.	,	none
1.	PayZone	String		6		USD	.	,	none
1.	EmployeeClassificationType	String		9		USD	.	,	none
1.	TerminationType	String		11		USD	.	,	none
1.	TerminationDescription	String		49		USD	.	,	none

### 2) Tambahkan sequence hari, kalkulasi tanggal, ekstraksi tahun/bulan/hari.

#	Fieldname	Rename to	Length	Precision
1	StartDate			

#	New field	Calculation	Field A	Field B	Field C	Value type	Length	Precision	Remove	Conversion mask	Decimal symbol	Grouping symbol	Curren
1	tanggal	Day of month of date A	StartDate			Integer			N				
2	bulan	Month of date A	StartDate			Integer			N				
3	tahun	Year of date A	StartDate			Integer			N				

### 3) Lookup untuk mencegah duplikat.

**Add sequence**

Step name: Add sequence

Name of value: id\_waktu

Use a database to generate the sequence

Use DB to get sequence? ☐

Connection: dim\_waktu Edit... New... Wizard...

Schema name:  Schemas...

Sequence name: SEQ\_ Sequences...

Use a transformation counter to generate the sequence

Use counter to calculate sequence? ☒

Counter name (optional):

Start at value: 1

Increment by: 1

Maximum value: 999999999

Help OK Cancel

**Filter rows**

Step name: Filter rows

Send 'true' data to step:

Send 'false' data to step:

The condition:

☐ +

id\_waktu IS NOT NULL

AND

tanggal IS NOT NULL

AND

bulan IS NOT NULL

AND

tahun IS NOT NULL

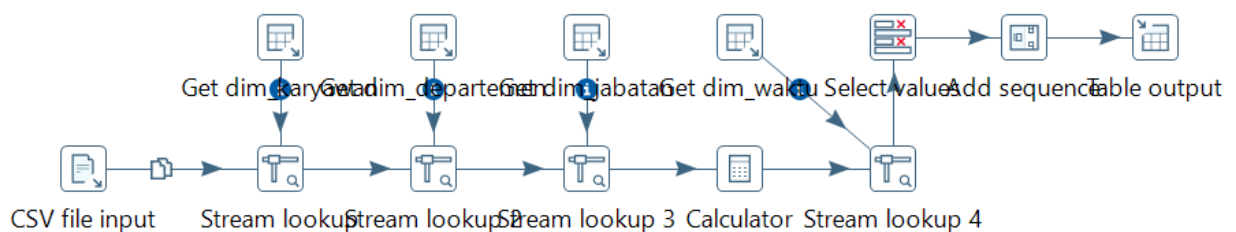
Help OK Cancel

4) Output ke dim\_waktu.

[illegible]

Server: localhost:3306 » Database: dimemployee » Table: dim_waktu											
Browse		Structure		SQL		Search		Insert		Export	
<div>← T →</div>				id_waktu	tanggal	bulan	tahun				
<input type="checkbox"/>	Edit	Copy	Delete	1	20	9	2019				
<input type="checkbox"/>	Edit	Copy	Delete	2	11	2	2023				
<input type="checkbox"/>	Edit	Copy	Delete	3	10	12	2018				
<input type="checkbox"/>	Edit	Copy	Delete	4	21	6	2021				
<input type="checkbox"/>	Edit	Copy	Delete	5	29	6	2019				
<input type="checkbox"/>	Edit	Copy	Delete	6	17	1	2020				
<input type="checkbox"/>	Edit	Copy	Delete	7	6	4	2022				
<input type="checkbox"/>	Edit	Copy	Delete	8	6	11	2020				
<input type="checkbox"/>	Edit	Copy	Delete	9	18	8	2018				
<input type="checkbox"/>	Edit	Copy	Delete	10	21	1	2022				
<input type="checkbox"/>	Edit	Copy	Delete	11	4	8	2023				
<input type="checkbox"/>	Edit	Copy	Delete	12	10	8	2018				
<input type="checkbox"/>	Edit	Copy	Delete	13	25	5	2022				
<input type="checkbox"/>	Edit	Copy	Delete	14	5	12	2019				
<input type="checkbox"/>	Edit	Copy	Delete	15	28	4	2019				
<input type="checkbox"/>	Edit	Copy	Delete	16	9	7	2019				
<input type="checkbox"/>	Edit	Copy	Delete	17	5	4	2021				
<input type="checkbox"/>	Edit	Copy	Delete	18	28	11	2021				
<input type="checkbox"/>	Edit	Copy	Delete	19	16	1	2021				
<input type="checkbox"/>	Edit	Copy	Delete	20	24	8	2021				
<input type="checkbox"/>	Edit	Copy	Delete	21	26	5	2020				
<input type="checkbox"/>	Edit	Copy	Delete	22	1	10	2019				
<input type="checkbox"/>	Edit	Copy	Delete	23	10	5	2023				
<input type="checkbox"/>	Edit	Copy	Delete	24	1	9	2020				

## E. Transformasi Fact Table – fact\_performance



- **Input:** Gabungan data karyawan, jabatan, departemen, waktu, dan nilai performance.





2) Lookup ke masing-masing dimensi untuk ambil ID relasi.

Stream lookup

Step name: Stream lookup

Lookup step: Get dim\_karyawan

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

#	Field	LookupField
1	EmpID	id_karyawan

Specify the fields to retrieve :

#	Field	New name	Default	Type
1	id_karyawan	FK_id_karyawan		None

Preserve memory (costs CPU) ☒

Key and value are exactly one integer field ☐

Use sorted list (i.s.o. hashtable) ☐

[Help](#) [OK](#) [Cancel](#) [Get Fields](#) [Get lookup fields](#)

Table input

Step name: Get dim\_karyawan

Connection: fact\_performance [Edit...](#) [New...](#) [Wizard...](#)

SQL [Get SQL select statement...](#)

```
SELECT id_karyawan FROM dim_karyawan;
```

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](#)

Stream lookup

Step name

Lookup step

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

#	Field	LookupField
1	DepartmentType	tipe_departemen

Specify the fields to retrieve :

#	Field	New name	Default	Type
1	id_departemen	FK_id_departemen		None

Preserve memory (costs CPU) ☒

Key and value are exactly one integer field ☐

Use sorted list (i.s.o. hashtable) ☐

Table input

Step name

Connection

SQL

```
SELECT id_departemen, tipe_departemen FROM dim_departemen;
```

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

Stream lookup

Step name

Lookup step

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

#	Field	LookupField
1	Title	nama_jabatan

Specify the fields to retrieve :

#	Field	New name	Default	Type
1	id_jabatan	FK_id_jabatan		None

☒ Preserve memory (costs CPU)  
☐ Key and value are exactly one integer field  
☐ Use sorted list (i.s.o. hashtable)

Table input

Step name

Connection

SQL

```
SELECT id_jabatan, nama_jabatan FROM dim_jabatan;
```

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

Calculator

Step name

calculator

☒ Throw an error on non existing files

Fields:

#	New field	Calculation	Field A	Field B	Field C	Value type	Length	Precision	Remove	Conversion mask	Decimal symbol	Grouping syn
1	day_from_start_date	Day of month of date A	StartDate			Integer			N			
2	month_from_start_date	Month of date A	StartDate			Integer			N			
3	year_from_start_date	Year of date A	StartDate			Integer			N			

Help

OK Cancel

Stream lookup

Step name

Stream lookup 4

Lookup step

Get dim\_waktu

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

#	Field	LookupField
1	day_from_start_date	tanggal
2	month_from_start_date	bulan
3	year_from_start_date	tahun

Specify the fields to retrieve :

#	Field	New name	Default	Type
1	id_waktu	FK_id_waktu		None

Preserve memory (costs CPU) ☒

Key and value are exactly one integer field ☐

Use sorted list (i.s.o. hashtable) ☐

Help

OK

Cancel

Get Fields

Get lookup fields

Table input

Step name: Get dim\_waktu

Connection: fact\_performance Edit... New... Wizard...

SQL Get SQL select statement...

```
SELECT id_waktu, tanggal, bulan, tahun FROM dim_waktu;
```

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

3) Lookup ke tabel fact\_performance untuk hindari duplikasi.

Add sequence

Step name: Add sequence

Name of value: id\_performance

Use a database to generate the sequence

Use DB to get sequence? ☐

Connection: fact\_performance Edit... New... Wizard...

Schema name  Schemas...

Sequence name SEQ\_ Sequences...

Use a transformation counter to generate the sequence

Use counter to calculate sequence? ☒

Counter name (optional)

Start at value 1

Increment by 1

Maximum value 999999999

Help OK Cancel

4) Filter Rows dan Table Output ke tabel fakta.

Table output

Step name: Table output

Connection: fact\_performance [Edit...] [New...] [Wizard...]

Target schema: dimemployee [Browse...]

Target table: fact\_performance [Browse...]

Commit size: 1000

Truncate table: ☐

Ignore insert errors: ☐

Specify database fields: ☒

Main options | Database fields

Fields to insert:

#	Table field	Stream field
1	performa...	Performanc...
2	employee...	Current Em...
3	id_departe...	FK_id_depar...
4	id_jabatan	FK_id_jabatan
5	id_karyaw...	FK_id_karya...
6	id_waktu	FK_id_waktu
7	id_perfor...	id_performa...

[Get fields]

[Enter field mapping]

[?] Help [OK] [Cancel] [SQL]

### 3. Eksekusi JOB ETL

File `etl_fact_performance.ktr` dijalankan melalui Jobs yang berisi:

- Start
- Transformasi dim\_karyawan
- Transformasi dim\_departemen
- Transformasi dim\_jabatan
- Transformasi dim\_waktu
- Transformasi fact\_performance
- Success

**Output:** Data lengkap dimasukkan ke dalam masing-masing tabel di data warehouse tanpa duplikasi.

#### 4. Perbandingan OLTP vs OLAP

Aspek	OLTP (Operasional)	OLAP (Analitik)
Tujuan	Transaksi harian	Analisis data historis
Desain	Relational Model	Star Schema
Tabel	Banyak, berelasi kompleks	Lebih sedikit, relasi sederhana
Redundansi	Minim	Wajar demi kecepatan query
Jenis Query	INSERT, UPDATE, DELETE, SELECT	SELECT (agregasi: COUNT, SUM, AVG)
Contoh	Data transaksi pegawai	Analisis performa pegawai per tahun

## 5. Kesimpulan

- Proyek ini menggunakan konsep **Star Schema** untuk menganalisis performa karyawan.
- Transformasi dilakukan secara bertahap agar data bersih, konsisten, dan tidak duplikat.
- Tools ETL seperti Pentaho sangat membantu dalam mengotomatisasi proses pembuatan data warehouse.
- Dengan adanya tabel fakta dan dimensi, analisis data seperti performa tahunan dan komparasi antar jabatan dapat dilakukan dengan mudah.