

# Data Engineering Project : Building a Data Warehouse

**Project-Based Internship ID/X  
Partners - Data Engineer**  
Presented by  
**OKKY DARMAWAN KOSTIDJAN**

## Data Engineer Project



**Okky  
Darmawan  
Kostidjan**

### About You

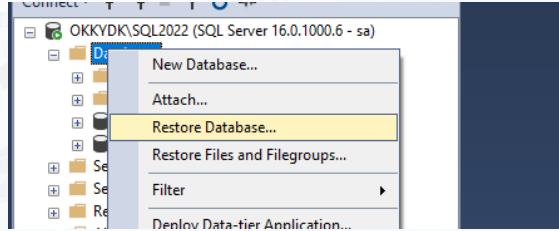
**"I am a qualified analyst with 5 years of experience in the field of analysis, specifically as an Inventory Analyst. My are of expertise are in maintaining stock availability, analyzing causes of product damage, and generating sales and analytical reports. "**

- Import / Restore Database
- Create a Database
- Create ETL Job
- Create Store Procedure

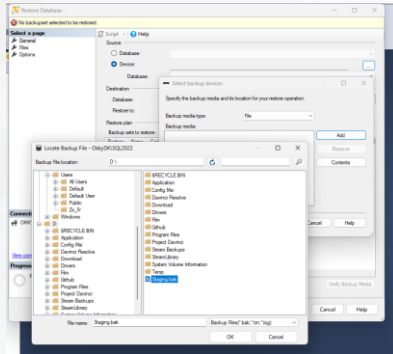
# **Import / Restore Database**

**Import / Restore Database in SSMS**

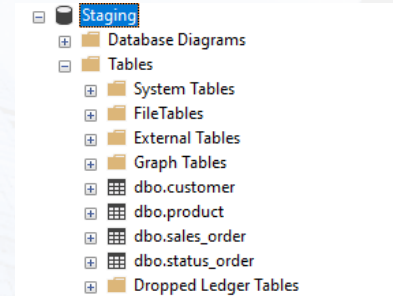
# Import / Restore Database



Klik kanan pada file Database lalu pilih restore Database



Pilih Device -> Add -> Pilih file .bak yang akan di restore / import. Kemudian pilih OK.



Ketika selesai maka akan muncul Database "Staging" seperti pada gambar di atas beserta tabel data di dalamnya

# **Create a Database**

**Create a New Database in SSMS**

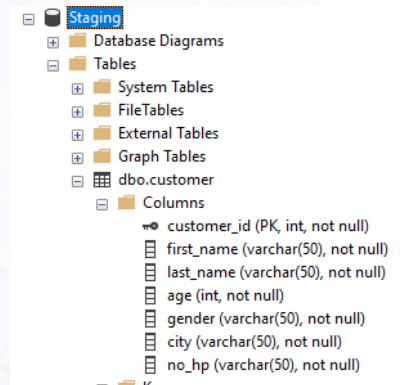


# Create a Database

```
SQLQuery1.sql - OKK...2.staging (sa (54)) * > SQLQuery1.sql - OKK...2.staging (sa (57))
USE [staging]
GO

/***** Object: Table [dbo].[customer]    Script Date: 31/12/2023 11:01:02 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO

CREATE TABLE [dbo].[customer](
    [customer_id] [int] NOT NULL,
    [first_name] [varchar](50) NOT NULL,
    [last_name] [varchar](50) NOT NULL,
    [age] [int] NOT NULL,
    [gender] [varchar](50) NOT NULL,
    [city] [varchar](50) NOT NULL,
    [no_hp] [varchar](50) NOT NULL,
    CONSTRAINT [PK_customer] PRIMARY KEY CLUSTERED
    (
        [customer_id] ASC
    ) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]
) ON [PRIMARY]
GO
```



```
USE [staging]
GO

/***** Object: Table [dbo].[sales_order]    Script Date: 31/12/2023 11:03:53 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO

CREATE TABLE [dbo].[sales_order](
    [order_id] [int] NOT NULL,
    [customer_id] [int] NOT NULL,
    [product_id] [int] NOT NULL,
    [quantity] [int] NOT NULL,
    [amount] [int] NOT NULL,
    [status_id] [int] NOT NULL,
    [order_date] [date] NOT NULL,
    CONSTRAINT [PK_sales_order] PRIMARY KEY CLUSTERED
    (
        [order_id] ASC
    ) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]
) ON [PRIMARY]
GO

ALTER TABLE [dbo].[sales_order] WITH CHECK ADD CONSTRAINT [FK_CustomerOrder] FOREIGN KEY([customer_id])
REFERENCES [dbo].[customer] ([customer_id])
GO

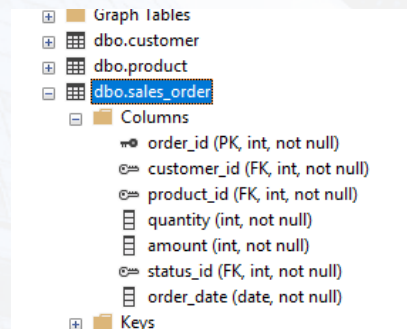
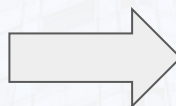
ALTER TABLE [dbo].[sales_order] CHECK CONSTRAINT [FK_CustomerOrder]
GO

ALTER TABLE [dbo].[sales_order] WITH CHECK ADD CONSTRAINT [FK_ProductOrder] FOREIGN KEY([product_id])
REFERENCES [dbo].[product] ([product_id])
GO

ALTER TABLE [dbo].[sales_order] CHECK CONSTRAINT [FK_ProductOrder]
GO

ALTER TABLE [dbo].[sales_order] WITH CHECK ADD CONSTRAINT [FK_StatusOrder] FOREIGN KEY([status_id])
REFERENCES [dbo].[status_order] ([status_id])
GO

ALTER TABLE [dbo].[sales_order] CHECK CONSTRAINT [FK_StatusOrder]
GO
```

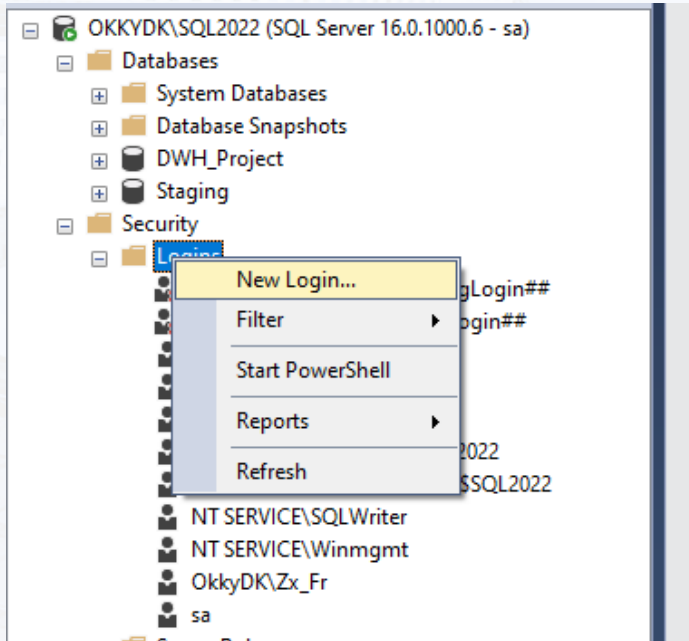


[Query lengkap bisa klik pada taunan ini.](#)

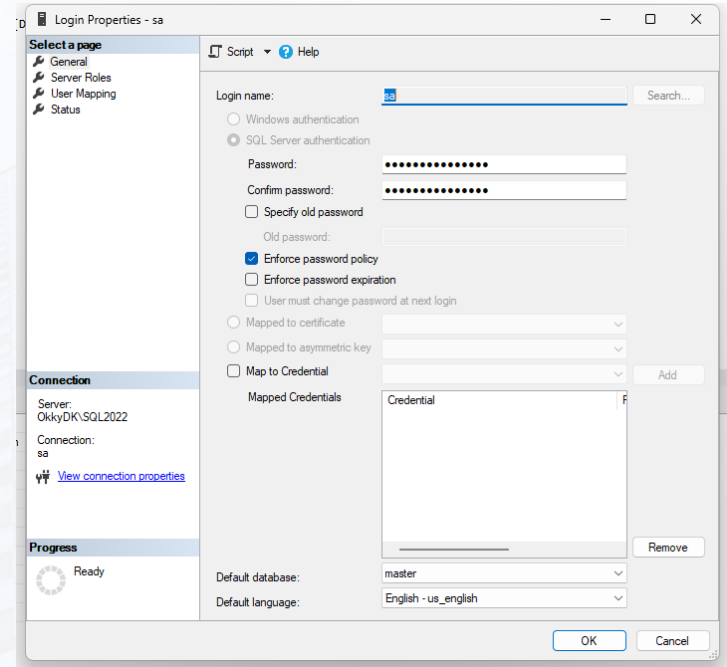
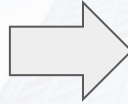
# Create ETL Job

Create a ETL JOB in Talend

# Create Account for SQL Server Authentication



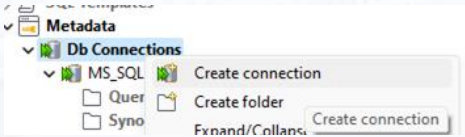
Klik kanan Db Connections pilih create connection



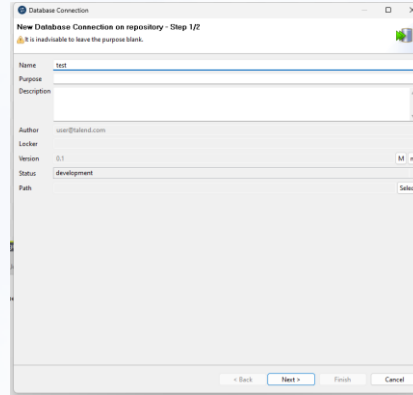
Pilih SQL Server authentication isi login name, password dan pastikan default database sesuai dengan database yang akan digunakan pada aplikasi talend selanjutnya



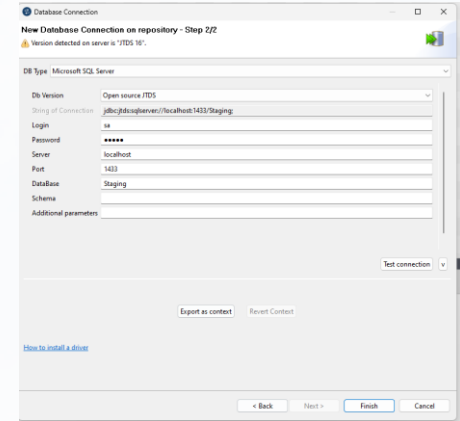
# Create a Db Connections



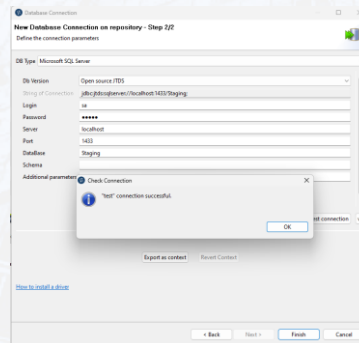
Klik kanan Db Connections pilih create connection



Beri nama lalu klik Next



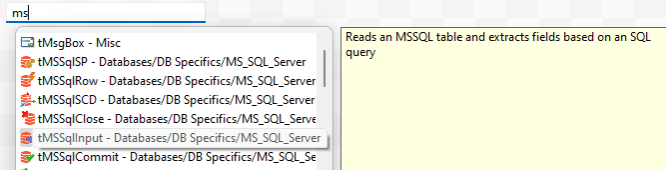
Isi cell kosong pada halaman berikut sesuai user yang telah dibuat pada SSMS lalu klik test connection



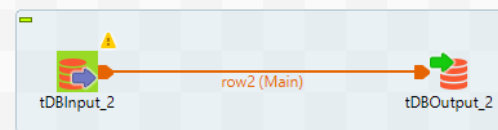
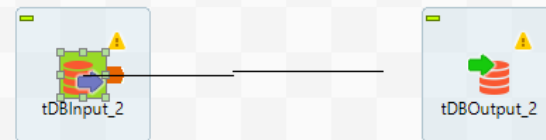
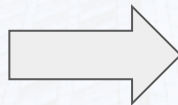
Apabila data yang di input sudah benar maka ketika klik test connection akan muncul notif seperti berikut. Lalu ok dan Finish



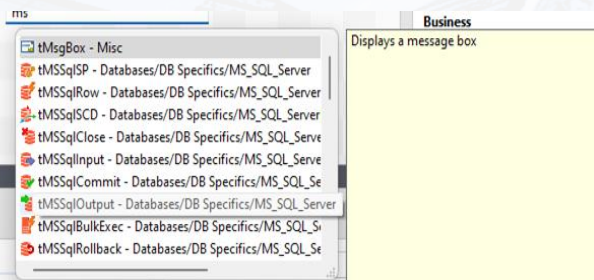
# Create ETL Job



Pilih tMSSqlInput untuk memilih database asal

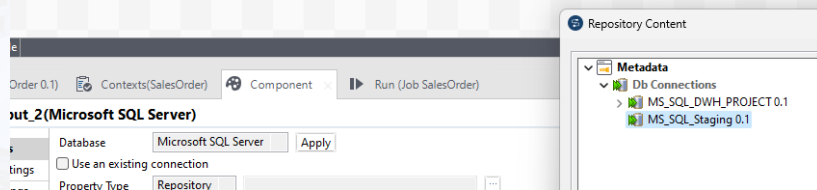
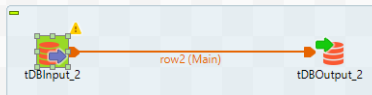


Klik tDBInput dan tarik garis ke tDBOutput untuk membuat relasi antara dua database tersebut

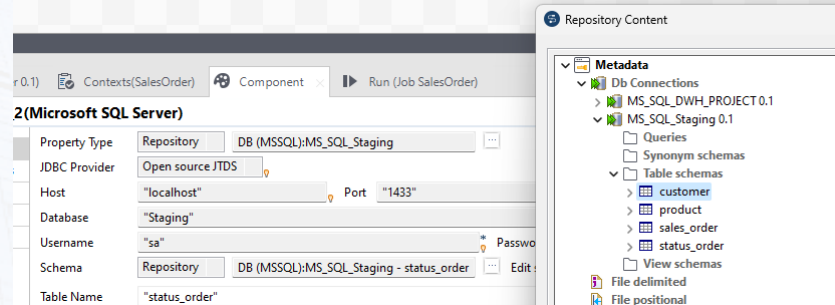
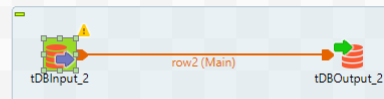


Pilih tMSSqlOutput untuk memilih database output

# Create ETL Job

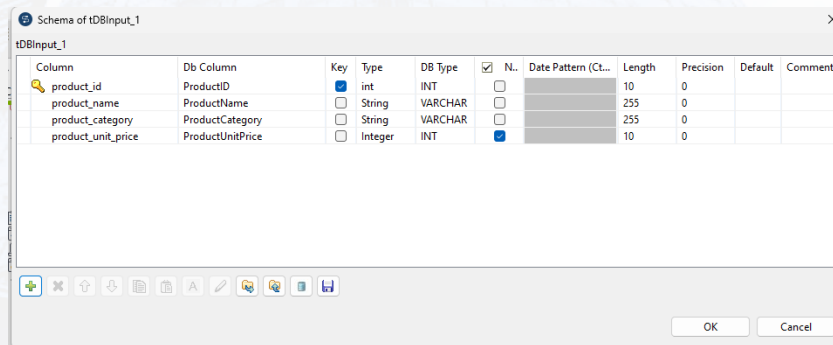
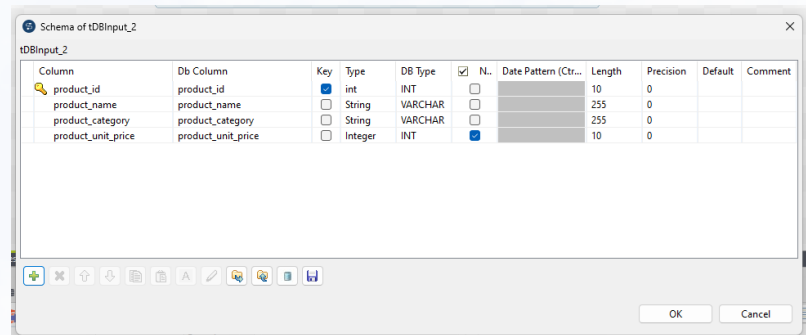
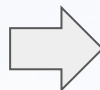
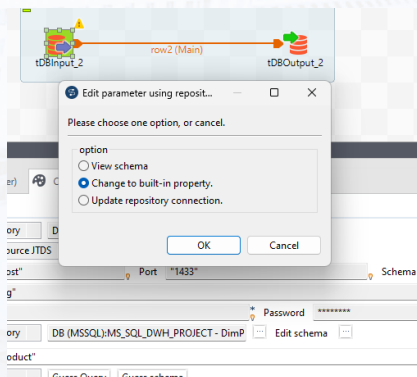


Klik tDB input, pilih menu Component dan ubah Prototype Type menjadi Repository dan pilih Db connection yang akan dijadikan database asal (dalam case ini database Staging).



Ubah Schema menjadi Repository dan pilih Tabel connection yang akan dijadikan tabel asal.

# Create ETL Job



Samakan nama kolom di bawah  
Db Column sesuai dengan nama  
kolom database tujuan

Klik icon titik tiga di sebelah kanan Edit Schema  
pilih Change to built-in property.

# Create ETL Job

(Microsoft SQL Server)

Database: Microsoft SQL Server [Apply]

☐ Use an existing connection

Property Type: Repository DB (MSSQL):MS\_SQL\_Staging [...]

JDBC Provider: Open source JTDS [v]

Host: "localhost" [v] Port: "1433" [v] Schema: "dbo" [v]

Database: "Staging" [v]

Username: "sa" [v] The variable attached to this parameter is: \_\_DBNAME\_\_ [v]

Schema: Built-In [v] Edit schema [...]

Table Name: "product" [...]

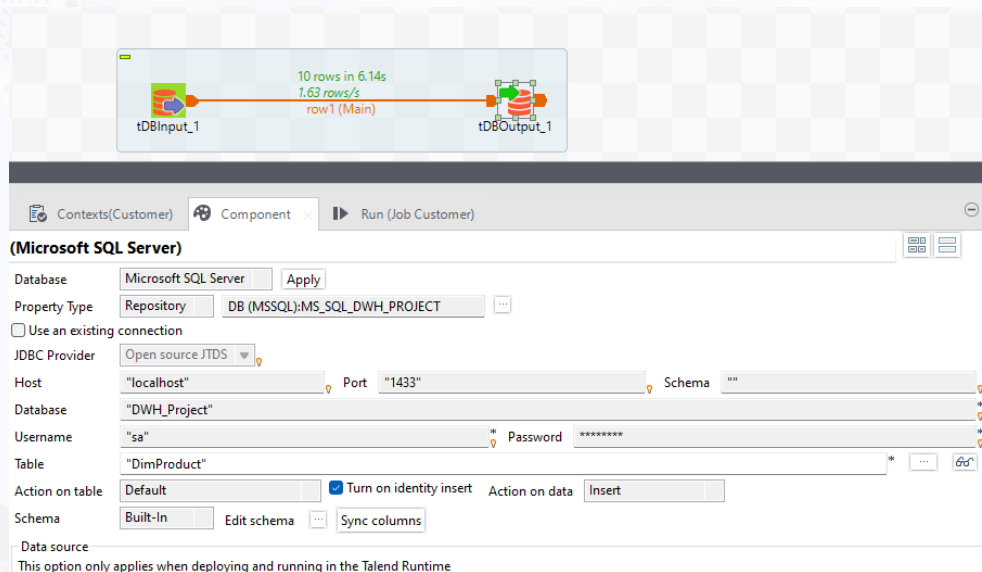
Query Type: Built-In [v] Guess Query [v] Guess schema [v]

Query: "SELECT \* FROM product" [v]

Lalu isi Query dengan data dari database yang akan disalin ke database tujuan



# Create ETL Job



10 rows in 6.14s  
1.63 rows/s  
row1 (Main)

tDBInput\_1 tDBOutput\_1

Contexts(Customer) Component Run (Job Customer)

**(Microsoft SQL Server)**

Database: Microsoft SQL Server Apply

Property Type: Repository DB (MSSQL):MS\_SQL\_DWH\_PROJECT

☐ Use an existing connection

JDBC Provider: Open source JTDS

Host: localhost Port: 1433 Schema:

Database: DWH\_Project

Username: sa Password: \*\*\*\*\*

Table: DimProduct

Action on table: Default ☒ Turn on identity insert Action on data: Insert

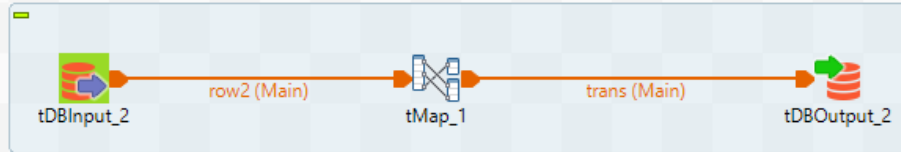
Schema: Built-In Edit schema Sync columns

Data source

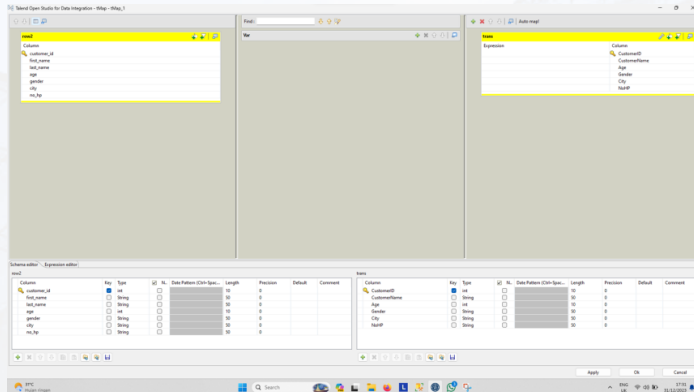
This option only applies when deploying and running in the Talend Runtime

Lakukan hal yang sama dengan tDBOutput untuk database tujuan namun untuk menu Schema pastikan nama tabel sama dengan klik menu Sync Column lalu cek dengan klik menu titik tiga yang berada di sebelah kiri menu sync columns

# Create ETL Job

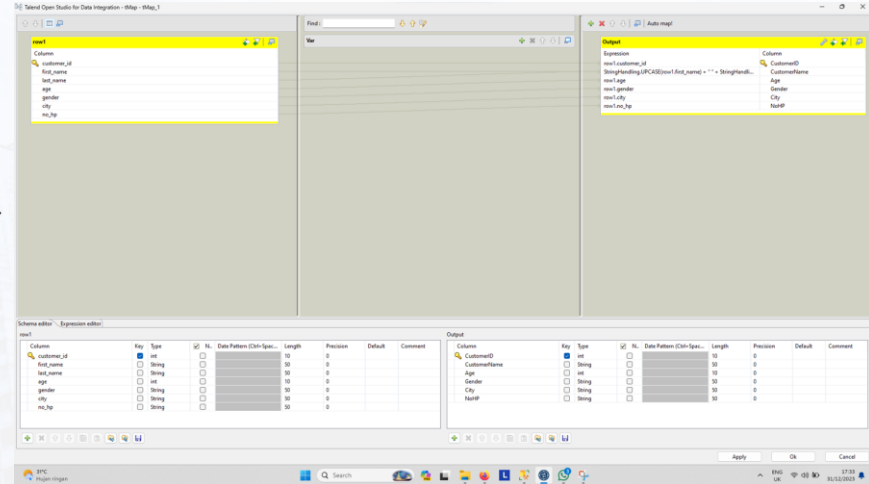
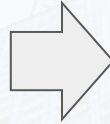
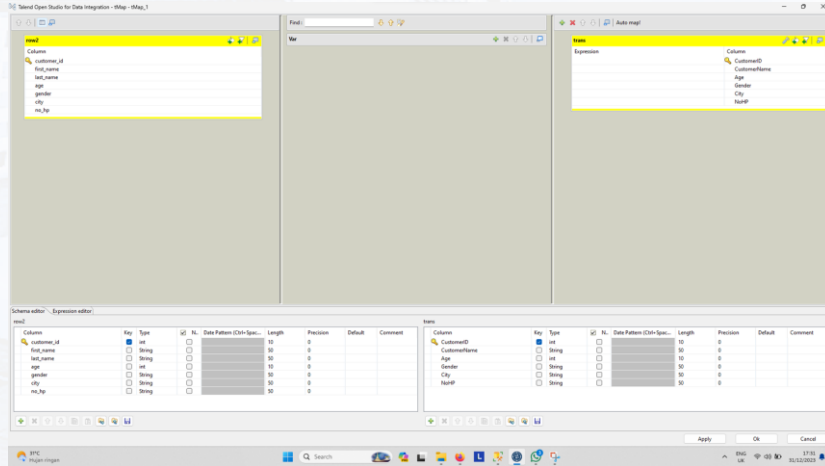


Untuk melakukan edit atau transformasi tambahkan tMap diantara tDBInput dan tDBOutput kemudian tarik garis dan beri nama fungsi tersebut.



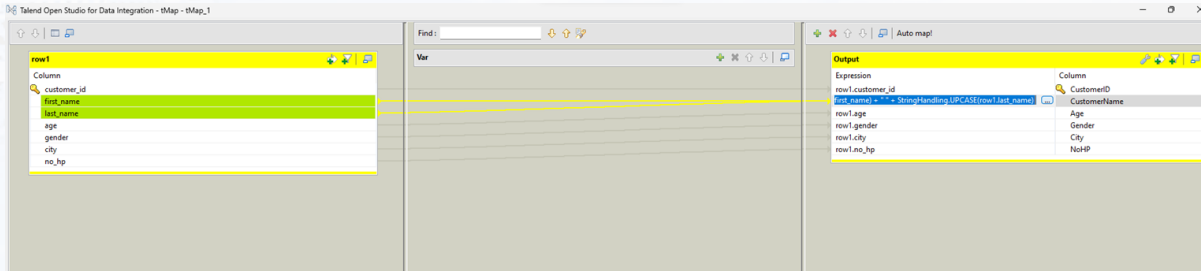
Klik tMap hingga muncul menu seperti gambar di samping

# Create ETL Job

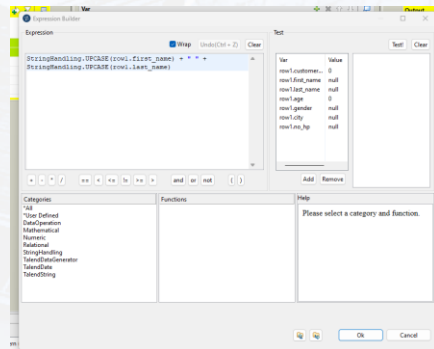


Edit nama dan jumlah kolom agar sesuai dengan nama kolom pada database tujuan pada menu kanan bawah.  
Tarik garis juga pada setiap kolom yang sesuai dari database asal dengan database tujuan

# Create ETL Job

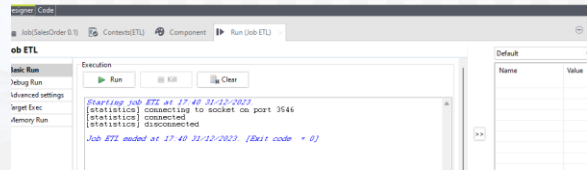


Klik baris di bawah Output untuk melakukan transformasi lalu klik ikon titik tiga yang muncul pada bagian tengah baris dan masukkan query perintah transformasi yang diinginkan.

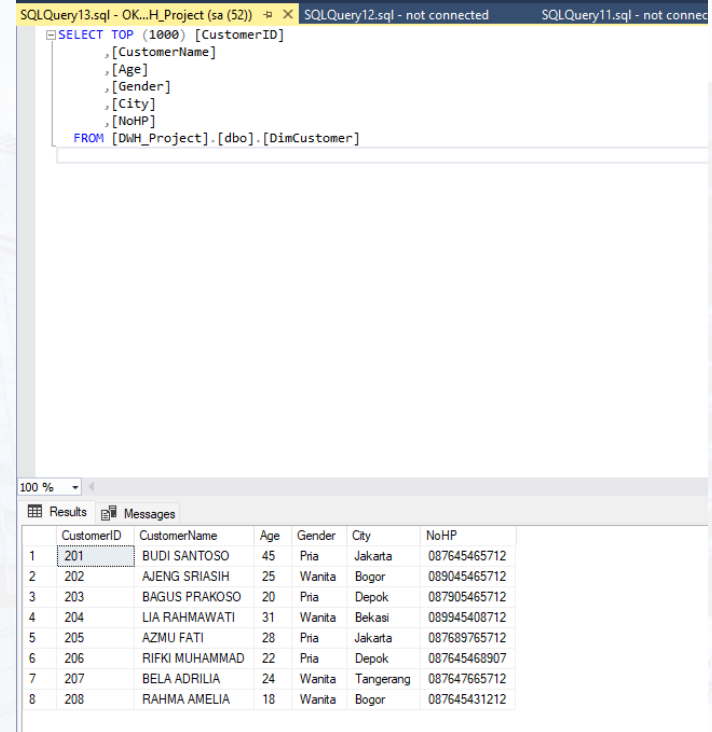


Pada gambar sebelah kiri adalah contoh transformasi untuk menggabungkan kolom first\_name dan last\_name menjadi satu kolom CustomerName dengan semua huruf diubah menjadi huruf kapital, kemudian klik OK apabila sudah selesai.

# Create ETL Job



Apabila sudah selesai semua, pilih menu Run dan klik Run, tunggu beberapa saat. Jika sudah selesai dan tidak muncul tulisan berwarna merah maka proses berjalan dengan lancar dan tidak ada masalah, silahkan cek pada SSMS isi tabel tujuan.



SQLQuery13.sql - OK...H\_Project (sa (52))

SQLQuery12.sql - not connected

SQLQuery11.sql - not connected

```
SELECT TOP (1000) [CustomerID]
, [CustomerName]
, [Age]
, [Gender]
, [City]
, [NoHP]
FROM [DimH_Project].[dbo].[DimCustomer]
```

100 %

Results

	CustomerID	CustomerName	Age	Gender	City	NoHP
1	201	BUDI SANTOSO	45	Pria	Jakarta	087645465712
2	202	AJENG SRIASIH	25	Wanita	Bogor	089045465712
3	203	BAGUS PRAKOSO	20	Pria	Depok	087905465712
4	204	LIA RAHMAWATI	31	Wanita	Bekasi	089945408712
5	205	AZMU FATI	28	Pria	Jakarta	087689765712
6	206	RIFKI MUHAMMAD	22	Pria	Depok	087645468907
7	207	BELA ADRILIA	24	Wanita	Tangerang	087647665712
8	208	RAHMA AMELIA	18	Wanita	Bogor	087645431212



# Create Store Procedure

Create a store procedure in SSMS

# Create Store Procedure

SQLQuery14.sql - OK...H\_Project (sa (64)) \* SQLQuery11.sql - not connected SQLQuery10.sql - not connected

```
CREATE or ALTER PROCEDURE summary_order_status
    @StatusID INT
AS
BEGIN
    SELECT
        fo.OrderID,
        c.CustomerName,
        p.ProductName,
        fo.Quantity,
        ds.StatusOrder
    FROM
        FactSalesOrder fo
        INNER JOIN DimCustomer c ON fo.CustomerID = c.CustomerID
        INNER JOIN DimProduct p ON fo.ProductID = p.ProductID
        INNER JOIN DimStatusOrder ds ON fo.StatusID = ds.StatusID
    WHERE
        fo.StatusID = @StatusID;
END;

EXEC summary_order_status @StatusID = 2;
```

100 %

Results Messages

	OrderID	CustomerName	ProductName	Quantity	StatusOrder
1	1301	LIA RAHMAWATI	Converse Cap Original	2	Awaiting Shipment
2	1304	AJENG SRIASIH	T-Shirt Polo Nevada	2	Awaiting Shipment
3	1307	RAHMA AMELIA	Pull & Bear T-Shirt	1	Awaiting Shipment

# Thank You



id/x partners