

Extrack Transform Load Using Talend8

Data Engineer IDX Partners

Presented by Sihar Pangaribuan



Sihar Pangaribuan

About Me

I am a career switcher from a Radio Frequency engineer to the field of Data Science. To deepen my knowledge in the data realm, I attended a 3-month Data Science bootcamp at Hacktiv8. I possess skills and knowledge in various areas including database management systems, SQL queries, exploratory data analysis, Python programming, data visualization using tools such as Tableau, Looker, and Data Studio, as well as machine learning and deep learning modeling. Additionally, I have experience with ETL tools like Pentaho, Talend, and Airflow. I am open to opportunities in the data world, encompassing roles in Data Science, Data Engineering, and Data Analysis



Experience

PT. Sinergi Aitikom 2019-2023

Radio Frequency Engineer
Job desk: Perform analysis and reporting of the results of checking telecommunication signals.

Technology / Tools: Microsoft Exel, Micosoft Power Point, Microsoft Power Bi, Nemo Analize, Tems, Mapinfo, Google Earth

Case Study

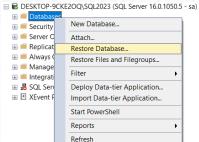


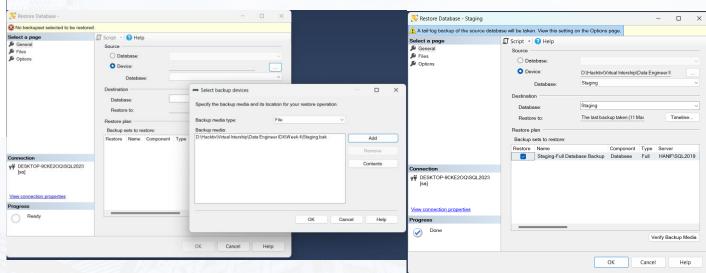
Case Study Background

One of the clients of ID/X Partners, operating in the e-commerce sector, has a requirement to create a Data Warehouse sourced from several tables in the source database. This Data Warehouse will consist of one Fact table and several Dimension tables.

- 1.Perform the Import/Restore of the Staging Database.
- 2.Create a Database named DWH_Project and generate Fact and Dimension Tables from the tables within the Staging Database.
- 3.Develop an ETL Job within the Talend application to transfer data from Staging to the Data Warehouse. Specifically for the DimCustomer Table, perform data transformation by converting data in the FirstName and LastName columns to all capital letters, then merge these two columns into a single column named CustomerName.
- 4. Generate a Stored Procedure (SP) to display a summary of sales orders based on delivery status.







```
-- crreate database
create database DWH_Project;
```

```
-- create table DimCustomer
Create table DimCustomer(
        CustomerID int not null Primary Key,
        CustomerName varchar(100) not null,
        Age int not null,
        Gender varchar(50) not null,
        City varchar(50) not null,
        NoHp varchar(50) not null
);
```

```
-- create table DimProduct

CREATE TABLE DimProduct (

ProductID int not null PRIMARY KEY,

ProductName varchar(255) not null,

ProductCategory varchar(255) not null,

ProductUnitPrice int null

);
```



```
-- crate table DimStatusOrder

CREATE TABLE DimStatusOrder(

StatusID int not null PRIMARY KEY,

StatusOrder varchar(50) not null,

StatusOrderDesc varchar(50) not null

);
```

```
-- crate table FactSalesOrder

CREATE TABLE FactSalesOrder(
OrderID int not null PRIMARY KEY,

CustomerID int not null,

ProductID int not null,

Quantity int not null,

Amount int not null,

StatusID int not null,

OrderDate date not null,

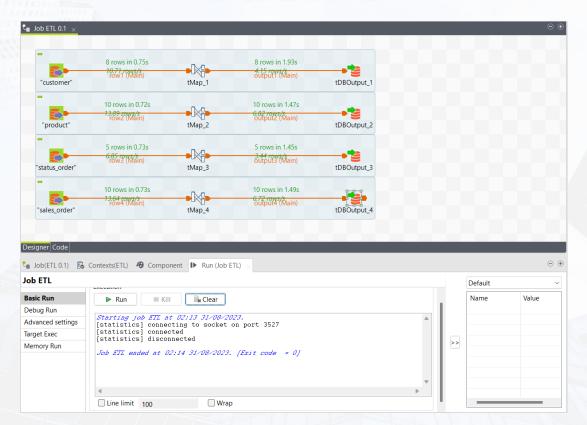
CONSTRAINT fk_customer FOREIGN KEY (CustomerID) REFERENCES DimCustomer (CustomerID),

CONSTRAINT fk_product FOREIGN KEY (ProductID) REFERENCES DimProduct (ProductID),

CONSTRAINT fk_statusorder FOREIGN KEY (StatusID) REFERENCES DimStatusOrder (StatusID)

);
```







```
CREATE PROCEDURE summary status order
    @StatusId INT
BEGIN
    SELECT
        fso.OrderId,
        dso.StatusId,
        c.CustomerName,
        p.ProductName,
        fso.Quantity,
        dso.StatusOrder
    FROM FactSalesOrder AS fso
    join DimCustomer AS c ON fso.CustomerId = c.CustomerId
    join DimProduct AS p ON fso.ProductId = p.ProductId
    join DimStatusOrder AS dso ON fso.StatusId = dso.StatusId
   WHERE dso.StatusId = @StatusId;
END;
EXEC summary status order @StatusID = 1;
```



Video Presentation/ Tutorial

https://youtu.be/8hc89P55rjg?si=vMxhP6Qh_DoK3va3

Github

https://github.com/siharp/Data_engineer_IDX_Virtual_Intership

Thank You





