

Membuat Data Warehouse dan Stored Procedure untuk salah satu client ID/X Partners

ID/X Partners - Data Engineer

Presented by

ALEX PUJI FIRMANSYAH

Alex Puji Firmansyah

About You

Mahasiswa aktif semester 4 yang sangat tertarik dengan dunia data. Melalui proyek-proyek seperti membangun data warehouse dengan BigQuery dan melakukan ETL dengan Pandas, saya telah mengasah kemampuan saya dalam SQL, Python, dan berbagai tools data engineering lainnya.

Saat ini, saya sedang aktif mempelajari cloud computing untuk memperluas pengetahuan saya tentang infrastruktur data yang scalable. Saya sangat tertarik untuk bergabung dengan tim yang inovatif dan berkontribusi dalam proyek-proyek data yang menarik.

***notes: just a sample page**



Lamongan, Jawa Timur



alexpujifirmansyah080803@gmail.com



www.linkedin.com/in/alex-puji-firmansyah

Courses and Certification

Data Engineer Challenge With Python | <https://academy.dqlab.id/certificate/pdf/DQLABPRJC3URHTAG> | Okt, 2024.

Data Engineer Challenge with SQL | <https://academy.dqlab.id/certificate/pdf/DQLABSQLTSOJETHT> | Okt, 2024.

Project Simple ETL with Pandas | <https://academy.dqlab.id/certificate/pdf/DQLABDEPROAFPMIE> | Okt, 2024.

Data Engineer Learnig Path at Google Cloud Skills | <https://www.credly.com/users/alex-puji-firmansyah> | Sep , 2024.

Cloud Practitioner Essentials (Belajar Dasar AWS Cloud) | <https://www.dicoding.com/certificates/L4PQ1L604X01> | Agu, 2024.

Pentaho Data Integration Fundamentals | <https://academy.dqlab.id/certificate/pdf/DQLABPDI10DCOUNQ> | Mei, 2024.

Merancang Database dan Membangun Karir Data Engineer | <https://coursenet.s3.ap-southeast-3.amazonaws.com/Course-Net/CompletionCertificateALEX%20PUJI%20FIRMANSYAH-109398.pdf> | Mei, 2024.

***notes: just a sample page**

About Company

ID/X Partners adalah perusahaan konsultan yang didirikan pada tahun 2002 oleh mantan banker dan konsultan manajemen yang memiliki pengalaman luas dalam siklus kredit dan manajemen proses, pengembangan skor, dan manajemen kinerja. Perusahaan ini memiliki pengalaman yang luas dalam melayani korporasi di Asia dan Australia, serta berbagai industri seperti jasa keuangan, telekomunikasi, manufaktur, dan ritel.

The logo for id/x partners, consisting of the text "id/x" in white on a light blue background, followed by "partners" in white on a dark blue background.

id/x partners

ID/X Partners menawarkan jasa konsultasi yang spesialis dalam menggunakan solusi analitik data dan pengambilan keputusan (DAD) yang dikombinasikan dengan manajemen risiko dan pemasaran terintegrasi untuk membantu klien meningkatkan profitabilitas portofolio dan proses bisnis.

Project Portfolio

Deskripsi Proyek

Salah satu client dari perusahaan ID/X Partners yang bergerak di industri perbankan, memiliki kebutuhan untuk membuat sebuah Data Warehouse dari beberapa sumber data yang berbeda yang tersimpan di dalam sistem mereka. Beberapa sumber data tersebut antara lain seperti file Excel, CSV, dan database SQL Server.

Latar Belakang

Permasalahan yang mereka hadapi saat ini adalah mereka kesulitan untuk mengekstrak data dari berbagai sumber (excel, csv, database) secara bersamaan sehingga pelaporan dan analisis data mereka selalu mengalami keterlambatan.

Problem statement

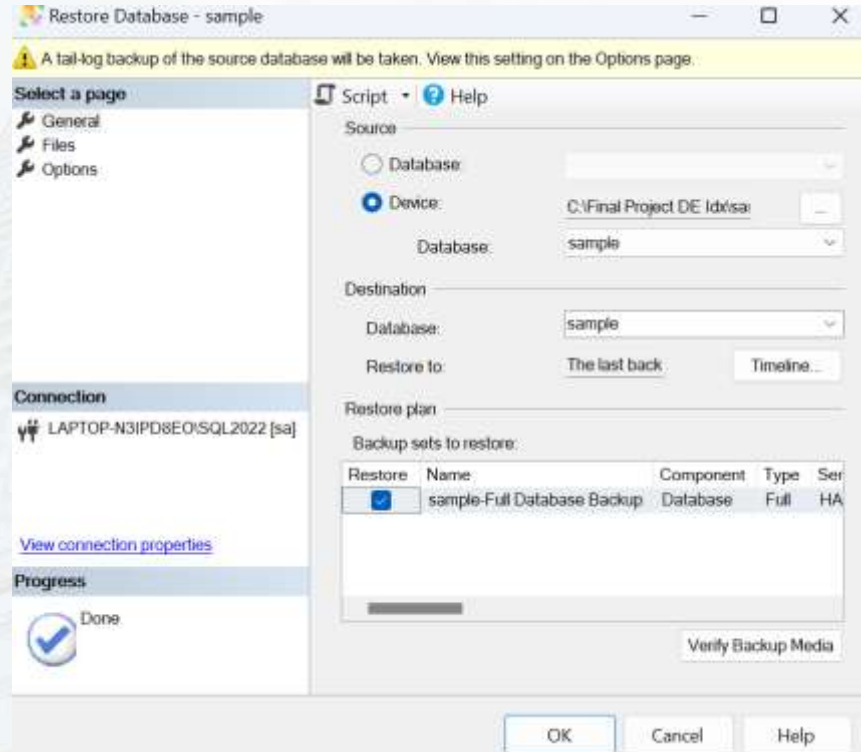
Perusahaan ini membutuhkan sebuah solusi yang dapat mengintegrasikan data dari berbagai sumber, membersihkan dan mentransformasi data, membangun data warehouse yang efisien, dan menyediakan alat pelaporan yang user-friendly. Dengan demikian, perusahaan dapat mengatasi masalah keterlambatan dan ketidakakuratan data, meningkatkan efisiensi proses pengambilan keputusan, dan mengurangi biaya operasional.

Project explanation video [here!](#)

1. Data Warehouse Creation

Restore Database:

Klik kanan pada database -> pilih
Restore Database -> pada general
bagian Source klik device -> lalu klik
titik 3(...) -> pada Backup file pilih file ->
klik add -> pilih file sample.bak -> klik ok
-> lalu klik ok lagi.



1. Data Warehouse Creation

Create Data Warehouse DWH → CREATE DATABASE DWH

Create Data Tabel FactTransaction

```
46 -- Tabel FactTransaction
47 CREATE TABLE FactTransaction (
48     Transaction_Id INT PRIMARY KEY NOT NULL, -- Primary Key
49     Account_Id INT NOT NULL, -- Foreign Key ke DimAccount
50     Transaction_Date DATETIME2 NOT NULL,
51     Amount DECIMAL NOT NULL,
52     Transaction_Type VARCHAR(50) NOT NULL,
53     Branch_Id INT NOT NULL, -- Foreign Key ke DimBranch
54
55     FOREIGN KEY (Account_Id) REFERENCES DimAccount(Account_Id),
56     FOREIGN KEY (Branch_Id) REFERENCES DimBranch(Branch_Id)
57 );
58
59
```

Create Tabel Dimension

```
15 -- Tabel DimCustomer
16 CREATE TABLE DimCustomer (
17     Customer_Id INT PRIMARY KEY NOT NULL, -- Primary Key
18     Customer_Name VARCHAR(50) NOT NULL,
19     Address VARCHAR(50) NOT NULL,
20     City_Name VARCHAR(50) NOT NULL,
21     State_Name VARCHAR(50) NOT NULL,
22     Age int NOT NULL,
23     Gender VARCHAR(50) NOT NULL,
24     Email VARCHAR(50) NOT NULL
25 );
26
27 -- Tabel DimAccount
28 CREATE TABLE DimAccount (
29     Account_Id INT PRIMARY KEY NOT NULL, -- Primary Key
30     Customer_Id INT NOT NULL, -- Foreign Key ke DimCustomer
31     Account_Type VARCHAR(50) NOT NULL,
32     Balance INT NOT NULL,
33     Date_Opened DATETIME NOT NULL,
34     Status VARCHAR(50) NOT NULL,
35     FOREIGN KEY (Customer_id) REFERENCES DimCustomer(Customer_Id)
36 );
37
38 -- Tabel DimBranch
39 CREATE TABLE DimBranch (
40     Branch_Id INT PRIMARY KEY NOT NULL, -- Primary Key
41     Branch_Name VARCHAR(50) NOT NULL,
42     Branch_Location VARCHAR(50) NOT NULL
43 );
44
```

1. Data Warehouse Creation

Result →

Results Messages

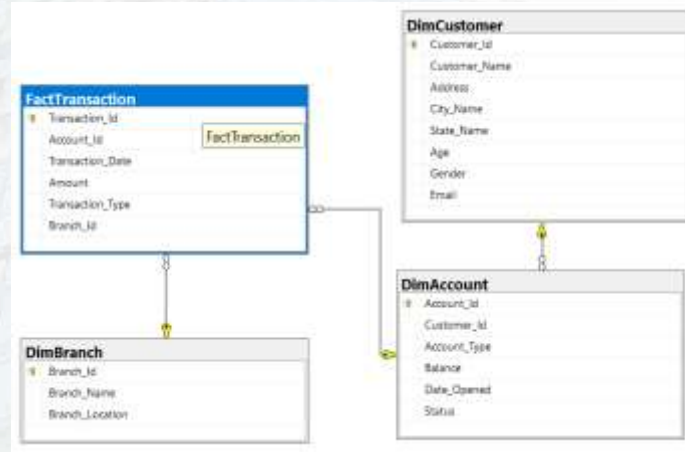
	Customer_Id	Customer_Name	Address	City_Name	State_Name	Age	Gender	Email
1	1	SHELLY JUWITA	JL BOULEVARD NO. 31	KELAPA GADING	JAKARTA UTARA	25	FEMALE	shelly@gmail.com
2	2	BOBI RINALDO	JL MANGGA NO. 1	TANJUNG PRIOK	JAKARTA UTARA	31	MALE	Bob@gmail.com

	Account_Id	Customer_Id	Account_Type	Balance	Date_Opened	Status
1	1	1	saving	1500000	2020-05-01 09:00:00.000	active
2	2	2	saving	500000	2020-06-01 10:00:00.000	active

	Branch_Id	Branch_Name	Branch_Location
1	1	KC Jakarta	Jl. Gatot Subroto No 13
2	2	KC Bogor	Jl. Padjajaran No 43

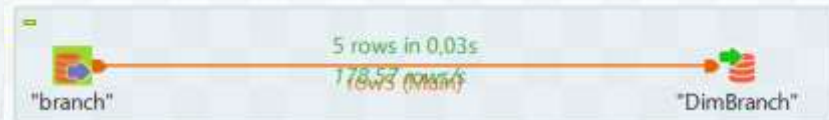
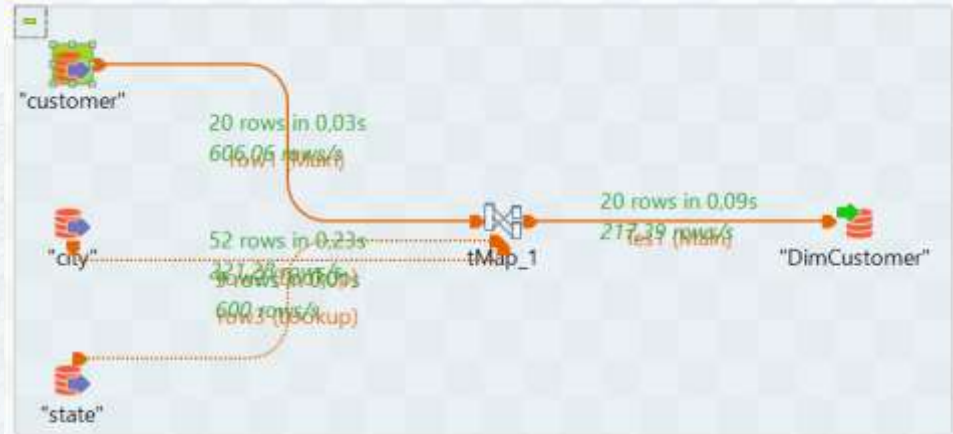
	Transaction_Id	Account_Id	Transaction_Date	Amount	Transaction_Type	Branch_Id
1	1	1	2024-01-17 09:10:00.0000000	100000	Deposit	1
2	2	2	2024-01-17 10:10:00.0000000	1000000	Deposit	1

Result Diagram Tabel →



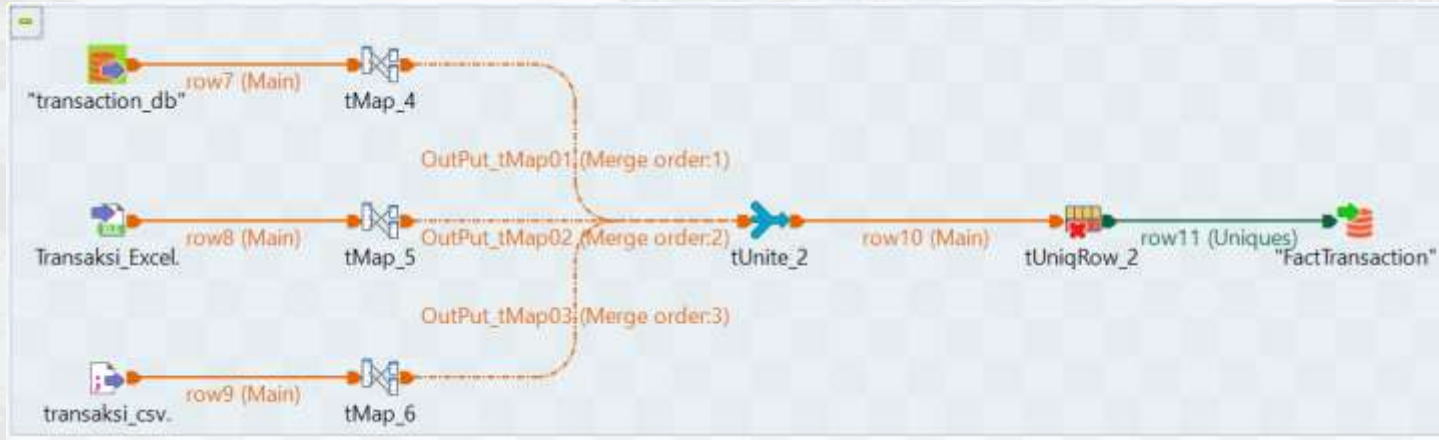
2. Create ETL Job for Dimension Table

Pada proses tMap sebelum ke output DimCustomer dilakukan perubahan data dari kolom tersebut diubah menjadi huruf kapital kecuali untuk kolom CustomerID, Age dan Email. Adapun tabel DimAccount dan DimBranch langsung ke output tanpa ada perubahan.



3. Create ETL Job for Fact Table

Membuat job ETL untuk menggabungkan data transaksi (transaksi_excel, transaksi_csv, transaksi_db) menjadi satu di tabel FactTransaction. Lalu pada tUniq dilakukan pembersihan data agar tidak ada row yang duplikat di dalam tabel FactTransaction.



4. Create Stored Procedure DailyTransaction

Membuat dua Stored Procedure **DailyTransaction** untuk menghitung banyaknya transaksi beserta total nominalnya setiap harinya, agar mereka mendapatkan ringkasan data dengan cepat.

Query SQL yang diberikan bertujuan untuk **menghitung dan menampilkan ringkasan transaksi harian** dalam rentang waktu tertentu.

	Date	TotalTransactions	TotalAmount
1	2024-01-17	2	1100000
2	2024-01-18	4	11250000
3	2024-01-19	3	5400000
4	2024-01-20	4	4000000
5	2024-01-21	2	2000000
6	2024-01-22	10	5180000

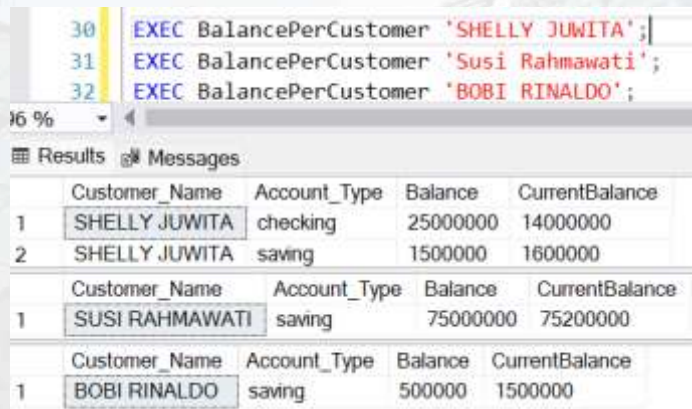
```
1 CREATE PROCEDURE DailyTransaction
2     @start_date DATE, -- Parameter input untuk menentukan rentang waktu
3     @end_date DATE -- Parameter output untuk menentukan rentang waktu
4 AS
5 BEGIN
6     SELECT
7         CAST(Transaction_Date AS DATE) AS Date, --Mengubah tipe data
8         COUNT(*) AS TotalTransactions, -- Menghitung jumlah transaksi
9         SUM(Amount) AS TotalAmount -- Menjumlahkan total nominal
10    FROM
11        FactTransaction
12    WHERE
13        Transaction_Date BETWEEN @start_date AND @end_date
14    GROUP BY
15        CAST(Transaction_Date AS DATE) --- Mengelompokkan data berdasarkan tanggal
16    ORDER BY
17        Date; -- Mengurutkan hasil berdasarkan tanggal secara ascending
18 END
19
20 EXEC DailyTransaction '2024-01-01', '2024-01-31'
```

4. Create Stored Procedure BalancePerCustomer

Membuat dua Stored Procedure **BalancePerCustomer** untuk mengetahui sisa balance per customer, agar mereka mendapatkan ringkasan data dengan cepat.

Tujuan Akhir:

Menggunakan logika ini, semua transaksi, baik berupa penambahan (Deposit) maupun pengurangan (Withdrawal), dihitung untuk menentukan perubahan bersih pada saldo akun.



	Customer_Name	Account_Type	Balance	CurrentBalance
1	SHELLEY JUWITA	checking	25000000	14000000
2	SHELLEY JUWITA	saving	1500000	1600000
	Customer_Name	Account_Type	Balance	CurrentBalance
1	SUSI RAHMAWATI	saving	75000000	75200000
	Customer_Name	Account_Type	Balance	CurrentBalance
1	BOBI RINALDO	saving	500000	1500000

```
1 CREATE PROCEDURE BalancePerCustomer
2     @name NVARCHAR(100) -- parameter @name yang digunakan untuk
3 AS
4 BEGIN
5     SELECT
6         c.Customer_Name,
7         a.Account_Type,
8         a.Balance,
9         a.Balance + SUM(CASE WHEN t.transaction_type = 'Deposit'
10 THEN t.Amount ELSE -t.Amount END) AS CurrentBalance
11     -- CASE WHEN digunakan untuk menentukan apakah transaksi
12
13 FROM
14     DimAccount a -- INNER JOIN antara tabel DimAccount dan D:
15     INNER JOIN DimCustomer c ON a.customer_id = c.customer_id
16     LEFT JOIN FactTransaction t ON a.account_id = t.account_id --
17
18 WHERE -- digunakan untuk memfilter data berdasarkan nama cu:
19     c.Customer_Name = @name AND
20     a.status = 'active'
21
22 GROUP BY
23     c.Customer_Name, a.Account_Type, a.Balance;
24 END
25
26 EXEC BalancePerCustomer 'SUSI RAHMAWATI';
```


Thank You



id/x partners



Berikut beberapa link dokumen:

- Link Github Hasil Pengerjaan :
https://github.com/alexpujifirmansyah/Final_Task_ID_X-Partners_Data-Engineer__ALEX_PUJI_FIRMANSYAH
- Link video Youtube :
<https://www.youtube.com/watch?v=gjwpr0888dM&feature=youtu.be>