

Task 2 – Working with Database : Load Data

Usecase Skenario

Sebuah Perusahaan Retail ingin mendapatkan Infrastruktur Business Intelligence menggunakan Data Warehouse, namun data-data yang dimiliki masih tersebar dari berbagai sumber, salah satunya berada di file lokal dari sistem yang sudah berjalan.

Terdapat data transaksi penjualan selama bulan Januari 2019 – Juni 2019:



transactionApr2019.csv 0 downloads, 554.12 KB



transactionFeb2019.csv 0 downloads, 553.47 KB



transactionJan2019.csv 0 downloads, 555.44 KB



transactionJun2019.csv 0 downloads, 555.40 KB



transactionMar2019.csv 0 downloads, 553.08 KB



transactionMay2019.csv 0 downloads, 554.25 KB

```

1 "?SalesID","Sales Area","Product Name","Country","State","Quantity","Order Date","Price Paid"
2 "1000752","Region North 1","Red Bull Zero","United Kingdom","Wales","439","2019-01-01","912"
3 "1000305","Region West","Monster Energy Absolutely Zero","Netherlands","South Holland","1179","2019-01-01","1,71"
4 "1000980","Region North 2","Monster Citron","Norway","Akershus","346","2019-01-01","1,76"
5 "1000510","Region Central","Monster Ultra White","Austria","Styria","54","2019-01-01","1,99"
6 "1000897","Region Central","Monster Citron","Austria","Styria","88","2019-01-01","1,35"
7 "1000368","Region West","Big Pump","Belgium","Antwerp","735","2019-01-01","1,09"
8 "1000313","Region North 2","Monster Energy Absolutely Zero","Norway","Buskerud","81","2019-01-01","2,02"
9 "1000837","Region West","Monster Ultra Red","Netherlands","Zeeland","2244","2019-01-01","872"
10 "1000098","Region West","Red Bull","Belgium","West Flanders","455","2019-01-01","0,95"
11 "1000206","Region North 2","Red Bull","Sweden","Östergötland","301","2019-01-01","1,25"
12 "1001007","Region South-West","Monster Citron","Spain","Extremadura","335","2019-01-01","1,09"
13 "1001008","Region South-West","Monster Citron","Spain","Galicia","1172","2019-01-01","1,09"
14 "1000993","Region South-West","Monster Citron","Portugal","Braga","438","2019-01-01","1,08"
15 "1000544","Region South","Monster Ultra White","Italy","Calabria","11","2019-01-01","1,69"
16 "1000227","Region Central","Monster Energy Absolutely Zero","Austria","Upper Austria","806","2019-01-01","1,55"
17 "1000470","Region South-West","Big Pump","Spain","Castile and León","3285","2019-01-01","0,93"
18 "1000797","Region Central","Monster Ultra Red","Germany","North Rhine-Westphalia","6895","2019-01-01","0,99"
19 "1000728","Region South-West","Red Bull Zero","Spain","Madrid","2927","2019-01-01","0,8"

```

Tugas Anda sebagai seorang Data Engineer saat ini adalah:

1. Melakukan pembacaan data dari file lokal.
2. Menggabungkan data (Union) dari 6 file lokal menjadi satu tabel yang disimpan di file lokal (AllTransactions.csv)
3. Loading Data (AllTransactions.csv) yang sudah digabungkan ke dalam satu tabel di database **msib_task**
4. **Nama Tabel:** AllTransactions_(nama_kelas)_(nama_lengkap), contoh:
AllTransactions_DE_Jeni_Sudirman
5. Membuat dokumentasi dari langkah-langkah yang sudah dilakukan.

Unduh File yang dibutuhkan untuk tugas ini di link berikut:

<https://www.mediafire.com/folder/nvapvasis4z2t/Task2 - 20230918>

Database Connection:

Database: msib_task

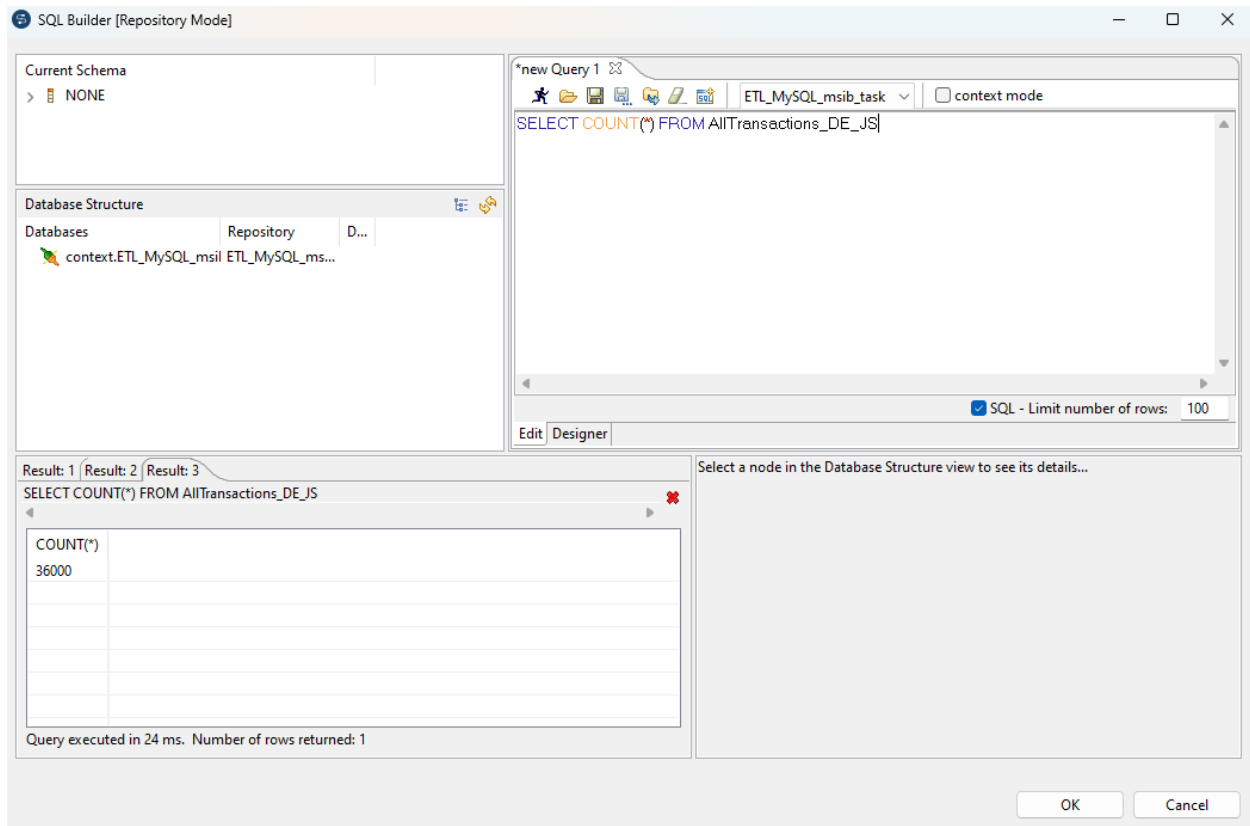
Server: 34.128.94.208

username: etl_msib

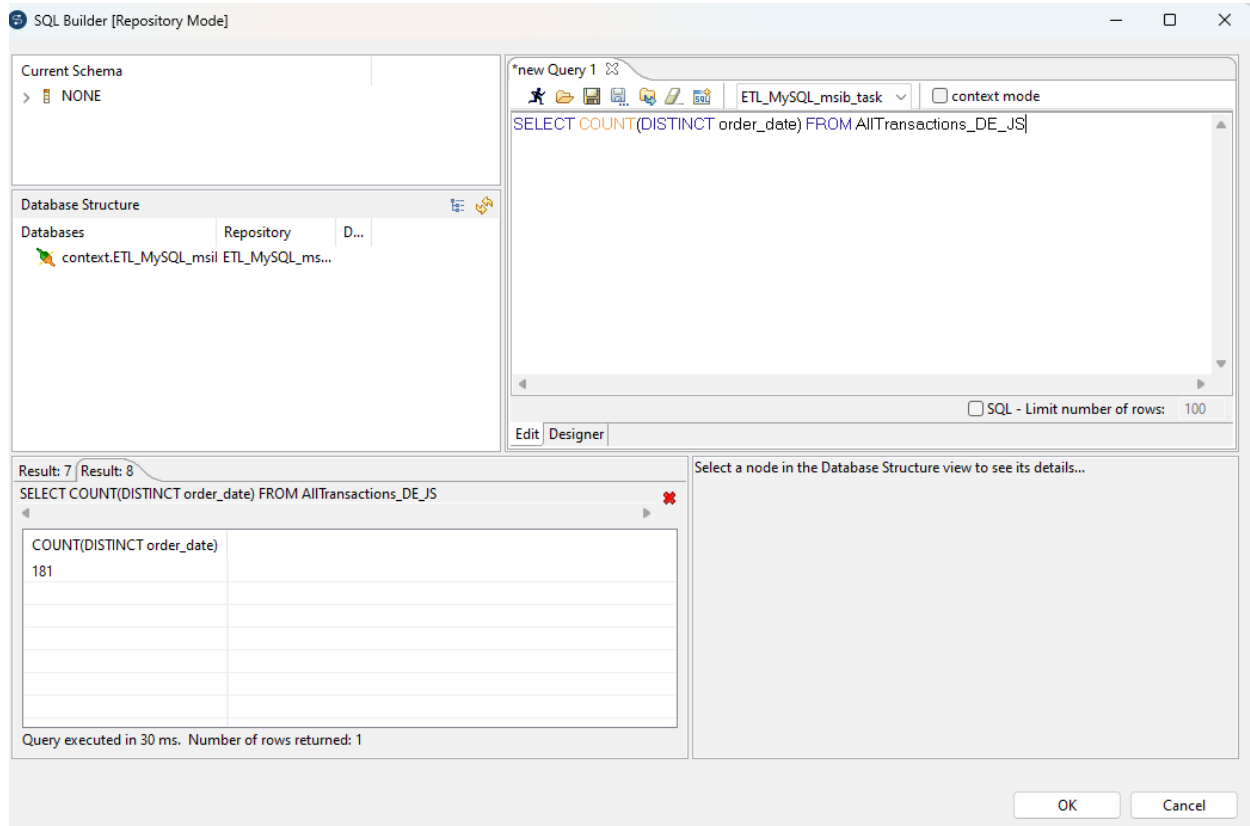
password: lihat di deskripsi grup kelas ini

Pertanyaan:

1. Berapakah total baris data yang anda miliki? Tunjukkan hasilnya dalam query di talend editor seperti gambar di bawah ini:



2. Ada berapa jumlah tanggal yang unik? Tunjukkan hasilnya dalam query di talend editor seperti gambar di bawah ini:



Sertakan screenshot untuk menjawab pertanyaan di atas ke dalam file dokumentasi yang sudah anda buat.

Submission:

1. File Project yang sudah di export dalam bentuk .zip. Nama Project:
Task2ETL_(nama_kelas)_(nama_mahasiswa).zip
2. File pdf dokumentasi dengan nama Task2ETL_(nama_kelas)_(nama_mahasiswa).pdf

FAQ

Q: Kapan deadline tugas ini?

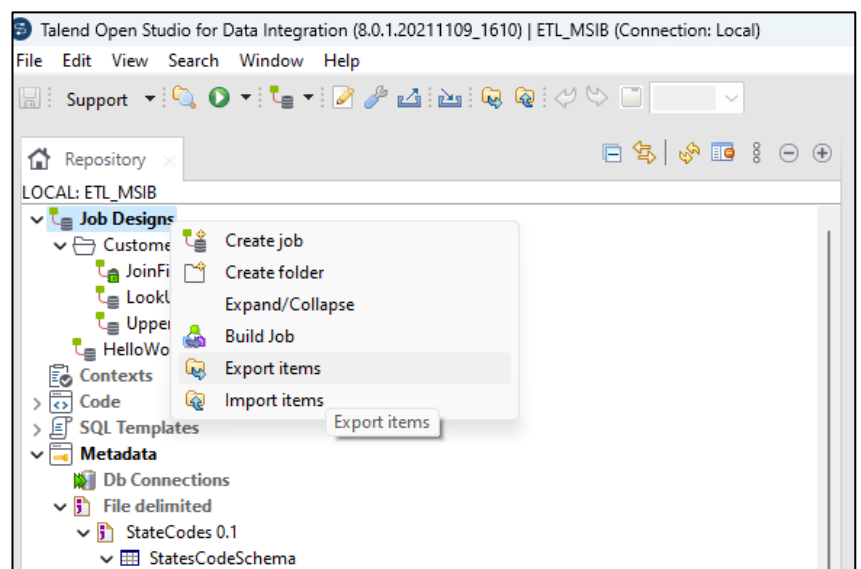
A: Kamis, 21 September 2023 Jam 23.59.

Q: Apakah kedua file submission tersebut perlu dikumpulkan lagi menjadi satu file .zip?

A: Tidak, total ada 2 file submission yang perlu di upload ke kognitify.

How to: Export Job

1. Klik kanan di **JobDesign** > pilih **Export Items**



2. Di Jendela Export items, pilih **Select archive File**:, lalu Browse lokasi penyimpanan file .zip, beri nama sesuai format. Pastikan semua item yang diminta sudah di ceklis semua. Lalu klik Finish.

