

# Development phase-1

## Data warehousing using IBM Cloud Db2 warehouse

### SQL script for creating tables in Db2 Warehouse :

The given SQL command using a Db2 SQL client, such as the Db2 Command Line Processor, Db2 Developer Center, or other SQL client tools. Make sure you have the necessary privileges to create tables in your Db2 Warehouse database.

```
student@admin1-HP-Pro-3330-MT: ~  
MariaDB [(none)]> use new;  
Database changed  
MariaDB [new]> create table customers(  
-> customer_id INT PRIMARY KEY,  
-> customer_name VARCHAR(100),  
-> email VARCHAR(100));  
Query OK, 0 rows affected (0.008 sec)  
  
MariaDB [new]> create table orders(  
-> order_id INT PRIMARY KEY,  
-> customer_id INT,  
-> order_date DATE,  
-> total_amount DECIMAL(10,2));  
Query OK, 0 rows affected (0.008 sec)  
  
MariaDB [new]> desc customers;  
+-----+-----+-----+-----+-----+-----+  
| Field          | Type          | Null | Key | Default | Extra |  
+-----+-----+-----+-----+-----+-----+  
| customer_id    | int(11)       | NO   | PRI | NULL    |       |  
| customer_name  | varchar(100)  | YES  |     | NULL    |       |  
| email          | varchar(100)  | YES  |     | NULL    |       |  
+-----+-----+-----+-----+-----+-----+  
3 rows in set (0.001 sec)  
  
MariaDB [new]> desc orders;  
+-----+-----+-----+-----+-----+-----+  
| Field          | Type          | Null | Key | Default | Extra |  
+-----+-----+-----+-----+-----+-----+  
| order_id       | int(11)       | NO   | PRI | NULL    |       |  
| customer_id    | int(11)       | YES  |     | NULL    |       |  
| order_date     | date          | YES  |     | NULL    |       |  
| total_amount   | decimal(10,2) | YES  |     | NULL    |       |  
+-----+-----+-----+-----+-----+-----+  
4 rows in set (0.001 sec)  
  
MariaDB [new]>
```

```
student@admin1-HP-Pro-3330-MT: ~
student@admin1-HP-Pro-3330-MT:~$ mysql -u root -p
Enter password:
Welcome to the MariaDB monitor.  Commands end with ; or \g.
Your MariaDB connection id is 41
Server version: 10.6.12-MariaDB-0ubuntu0.22.04.1 Ubuntu 22.04

Copyright (c) 2000, 2018, Oracle, MariaDB Corporation Ab and others.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

MariaDB [(none)]> use new;
Database changed
MariaDB [new]> create table customers(
-> customer_id INT PRIMARY KEY,
-> customer_name VARCHAR(100),
-> email VARCHAR(100));
Query OK, 0 rows affected (0.008 sec)

MariaDB [new]> create table orders(
-> order_id INT PRIMARY KEY,
-> customer_id INT,
-> order_date DATE,
-> total_amount DECIMAL(10,2));
Query OK, 0 rows affected (0.008 sec)

MariaDB [new]>
```

## Identifying data sources(csv file) :

FileHomeInsertPage LayoutFormulasDataReviewViewKutools™Kutools PlusHelp

CutCopyFormat PainterClipboard

Calibri11A A B I U Bold Italic Underline Font Alignment

Wrap Text Merge & Center

GeneralConditional FormattingFormat as TableCell StylesInsertDeleteFormatCells

AutoSumFillClear

Sort & Find & Filter > Select > Editing

Add-InsAdd-ins

POSSIBLE DATA LOSS Some features might be lost if you save this workbook in the comma-delimited (.csv) format. To preserve these features, save it in an Excel file format. Don't show againSave As...

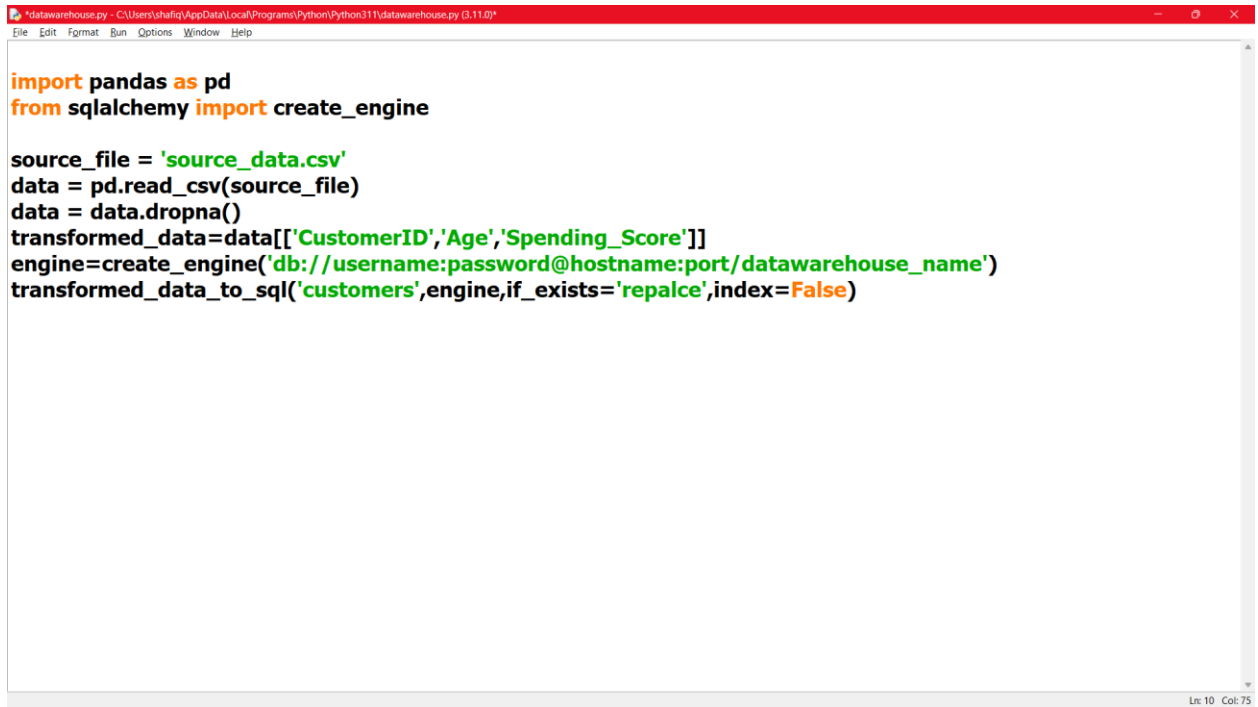
A1

CustomerID

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Customer	Genre	Age	Annual_In		Spending_Score												
2	1	Male	19	15	39													
3	2	Male	21	15	81													
4	3	Female	20	16	6													
5	4	Female	23	16	77													
6	5	Female	31	17	40													
7	6	Female	22	17	76													
8	7	Female	35	18	6													
9	8	Female	23	18	94													
10	9	Male	64	19	3													
11	10	Female	30	19	72													
12	11	Male	67	19	14													
13	12	Female	35	19	99													
14	13	Female	58	20	15													
15	14	Female	24	20	77													
16	15	Male	37	20	13													
17	16	Male	22	20	79													
18	17	Female	35	21	35													
19	18	Male	20	21	66													
20	19	Male	52	23	29													
21	20	Female	35	23	98													
22	21	Male	35	24	35													
23	22	Male	25	24	73													

Accessibility Unavailable

## Process to integrate them into the data Warehouse :



```
*datawarehouse.py - C:\Users\shafiq\AppData\Local\Programs\Python\Python311\datawarehouse.py (3-11.0)*
File Edit Format Run Options Window Help

import pandas as pd
from sqlalchemy import create_engine

source_file = 'source_data.csv'
data = pd.read_csv(source_file)
data = data.dropna()
transformed_data=data[['CustomerID','Age','Spending_Score']]
engine=create_engine('db://username:password@hostname:port/datawarehouse_name')
transformed_data_to_sql('customers',engine,if_exists='replace',index=False)
```

Ln: 10 Col: 75