BDAT 1010 BUSINESS INTELLIGENCE

GROUP 13

ASSIGNMENT 2 - TALEND

Q1: Create an ETL job to read the data of employee, which is in the following format-

Employee.csv

The output data should be stored in MSSQL database table.

Process:

A screenshot of a computer

Description automatically generated

Output:

A computer screen with a white screen

Description automatically generated

Q2: Create an ETL job to read the data of “Covid19 data.csv” and store it into the MSSQL database table.

Process:

A computer screen shot of a computer

Description automatically generated

Output:

A screenshot of a computer

Description automatically generated

Q3: Create an ETL job to read the data from the following JSON file and stored it into the CSV file

Store.json

Process:

A screenshot of a computer

Description automatically generated

Output:

A screenshot of a computer

Description automatically generated

Q4: Create an ETL job to read the data from the following XML file and store it into excel file-

Book.xml

Process:

A computer screen shot of a computer

Description automatically generated

Output:

A screenshot of a computer

Description automatically generated

Q5: Create an ETL job to read the data from following MSSQL table to Excel file.

DimCustomer

Process:

A screenshot of a computer

Description automatically generated

Output: A screenshot of a computer

Description automatically generated

Q6: Create an ETL job to load the data from three different files as shown below and join these three different files based on the key columns (DeptId, AddrID). The output will store it into the CSV file call Employee.csv.

EmpAddr.txt

EmpDept.csv

Employee.xlsx

Mapping:

A screenshot of a computer

Description automatically generated

Process:  
A screenshot of a computer

Description automatically generated

Output:

A screenshot of a computer

Description automatically generated

Q7: Create an ETL job to load the data from “Product.xlsx” file having data for Product and Order, join them based on the key columns (ProductID). The output will store it into the MySQL database table call Sales. [Use join to join this table]

Process:

A computer screen shot of a computer

Description automatically generated

Mapping:

A screenshot of a computer

Description automatically generated

Output:

A computer screen shot of a computer screen

Description automatically generated

Q8: Create an ETL job to load from “ActiveEmployee.xlsx” file, load the data into the MSSQL database table based on the following conditions-

 Split the data into three different files based on the conditions-

o If the salary <= 250000 then call it LowRange

o If the salary > 250000 and salary <= 500000 then call it MediumRange

o If the salary > 500000 then call it HighRange

 Combine the first name and last name of the employee and call it EmployeeName

 Calculate the age of the individual employee and call it Age column (int data type)

 Evaluate if Gender="M” then “Male” and if Gender= “F” then “Female”, call it Gender.

 Evaluate if Country Code="CAN” then “Canada”, if Country Code=”CHN” then “China” and so on, call this as a Country.

 Evaluate if Length of zipcode=4 then add two “00” as the beginning of the data, if Length of zipcode=5 then add a “0” as the beginning of the data and so on, make this column’s each data as 6. Ex- (1890 then 001890)

Process: A computer screen shot of a computer screen

Description automatically generated

Mapping:

A screenshot of a computer

Description automatically generated

Q10: Create an ETL job Achieve all the files (in .zip format) from a directory.

Process:

A screenshot of a computer

Description automatically generated

Output:

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Q11: Create an ETL job to declare the number starting from 101 to 105, capture the value of each iteration. Use a flat file to capture the output. [tForLoop]

Process:

A computer screen shot of a computer

Description automatically generated

Value:

A screenshot of a computer

Description automatically generated

Output:

A screenshot of a computer

Description automatically generated

Q12: Create an ETL job to read the data of “Financial Sample.xlsx” sort the data by (UnitsSold and SalePrice) columns and store it into a MSSQL Database table.

Process: A computer screen shot of a computer screen

Description automatically generated

Sorting:

A screenshot of a computer

Description automatically generated

Output:

A screenshot of a computer

Description automatically generated

|  |
| --- |
| **Name of students who has participated in the assignment.** |
| **Student name: Akshaykumar Vijay Thakare** |
| **Student name: Jay Sureshbhai Mangukiya** |
| **Student name: Vivek Jilesingh Maan** |
| **Student name:** |
| **Student name:** |

|  |
| --- |
| **Name of students who has not participated in the assignment.** |
| **Student name:** |
| **Student name:** |
| **Student name:** |
| **Student name:** |
| **Student name:** |