# **BIG DATA MANAGEMENT**

# POST GRADUATE DIPLOMA IN DATA ENGINEERING

**ASSIGNMENT - 6** 

# **SUBMITTED BY:**

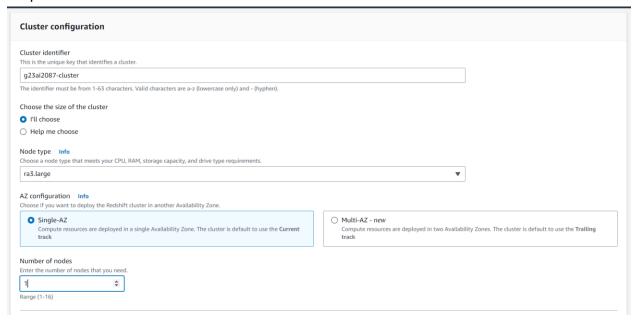
NIRAJ BHAGCHANDANI [G23AI2087]



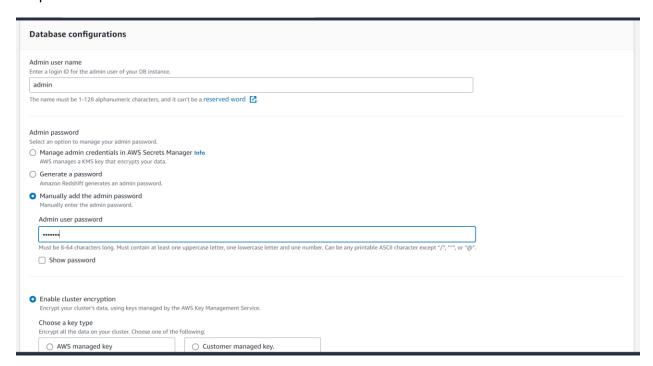
SUBMISSION DATE: 15th December, 2024

DEPARTMENT OF AIDE INDIAN INSTITUTE OF TECHNOLOGY, JODHPUR

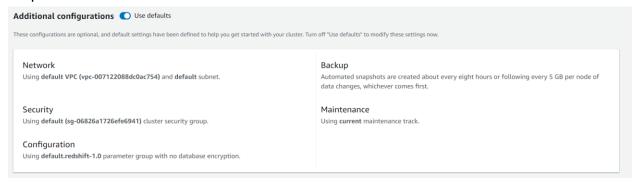
# Step-1:



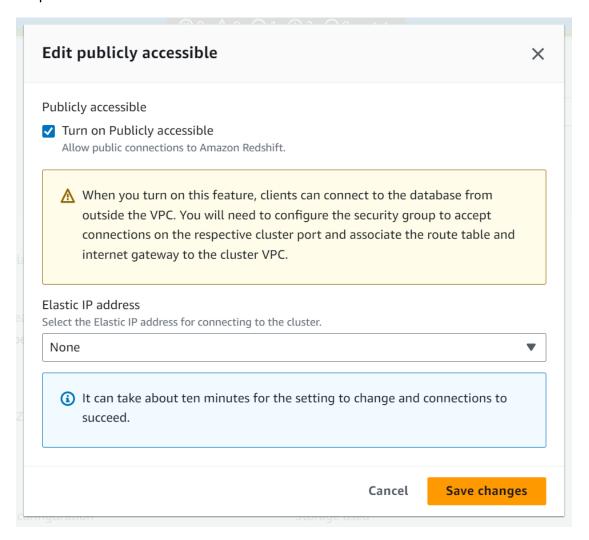
# Step-2:



# Step -3:

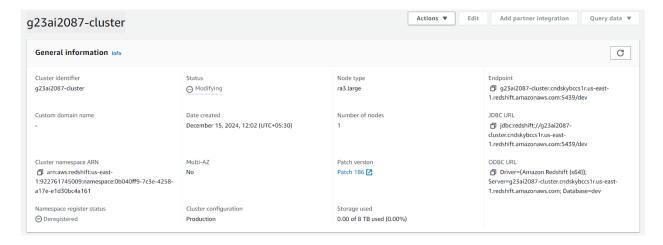


# Step-4:

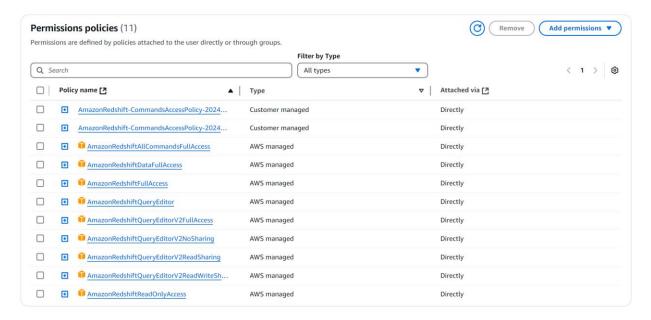


Step - 5:

# Big Data Management | Assignment - 6 | Trimister - 3 | IIT Jodhpur



# Step-6:

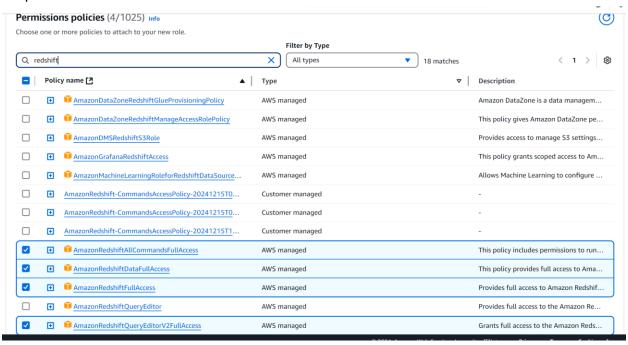


# Step - 7:



# Big Data Management | Assignment - 6 | Trimister - 3 | IIT Jodhpur

# Step - 8:



# Step - 9:



#### Step - 10



Now do the following tasks with the help of the starter code provided below

1. Write the method connect() to make a connection to the database. [5]

#### Code:

```
1. public Connection connect() {
           try {
3.
               Class.forName("com.amazon.redshift.jdbc42.Driver");
4.
               Properties properties = new Properties();
5.
               properties.setProperty("user", masterUsername);
               properties.setProperty("password", password);
6.
7.
8.
               this.con = DriverManager.getConnection(redshiftUrl, properties);
9.
10.
               System.out.println("Connection to Redshift established
   successfully.");
11.
           } catch (ClassNotFoundException e) {
12.
               System.out.println("Error: Redshift JDBC driver not found.");
13.
               e.printStackTrace();
           } catch (SQLException e) {
14.
               System.out.println("Failed to connect to the database.");
15.
16.
               e.printStackTrace();
17.
18.
           return con;
19.
```

#### Output:

```
C:\Windows\system32\cmd.exe - java -cp Drivers\*; AmazonRedshift

D:\iitj\tri-3\BDM\iitj-bdm-assignment\Assignment - 6>java -cp Drivers\*; AmazonRedshift

===== Amazon Redshift Menu =====

1. Connect to the database

2. Drop all tables

3. Create schema and tables

4. Insert TPC-H Data

5. Execute Query 1

6. Execute Query 2

7. Execute Query 3

78. Close the connection

9. Exit

|Enter your choice: 1
|Connection to Redshift established successfully.
```

2. Method close() to close the connection to the database. [5]

#### Code:

```
1. public void close() {
2.
           System.out.println("Closing database connection.");
3.
           try {
                if (con != null && !con.isClosed()) {
4.
5.
                    con.close();
6.
7.
           } catch (SQLException e) {
8.
                System.out.println("Error closing the connection.");
9.
10.
```

## Output:

```
ENC:\Windows\system32\cmd.exe - java --add-opens java.base/java.net=ALL-UNNAMED -cp "Drivers/*;." AmazonRedshift

===== Amazon Redshift Menu =====

1. Connect to the database

2. Drop all tables

3. Create schema and tables

4. Insert TPC-H Data

45. Execute Query 1

6. Execute Query 2

7. Execute Query 3

8. Close the connection

9. Exit
Enter your choice: 8

Closing database connection.
```

3. Method drop() to drop all the tables from the database. Note: The database schema name will be dev. [5]

```
7.
               while (rs.next()) {
8.
                   String tableName = rs.getString("tablename");
9.
                   String dropTableQuery = "DROP TABLE IF EXISTS dev." + tableName;
10.
                    stmt.executeUpdate(dropTableQuery);
                   System.out.println("Dropped table: " + tableName);
11.
               }
12.
13.
           } catch (SQLException e) {
               System.out.println("Error dropping tables: " + e.getMessage());
14.
15.
           }
16.
17. Output:
18. Method create() to create the database dev and the tables. [5]
19. Code:
20.public void create() throws SQLException {
           System.out.println("Creating the 'dev' schema and tables...");
21.
22.
           String createSchemaQuery = "CREATE SCHEMA IF NOT EXISTS dev";
23.
           try (Statement stmt = con.createStatement()) {
24.
                stmt.executeUpdate(createSchemaQuery);
25.
               System.out.println("Schema 'dev' created.");
26.
           } catch (SQLException e) {
27.
               System.out.println("Error creating schema: " + e.getMessage());
28.
29.
           File ddlFolder = new File("ddl");
           File[] ddlFiles = ddlFolder.listFiles((dir, name) ->
   name.endsWith(".sql"));
           if (ddlFiles != null) {
31.
32.
               for (File ddlFile : ddlFiles) {
33.
                   try {
34.
                        String ddlQuery = new
   String(Files.readAllBytes(ddlFile.toPath()));
                        if (ddlFile.getName().equals("tpch create.sql")) {
35.
                            // Ensure the 'tpch_create.sql' contains CREATE TABLE
36.
   queries
                            if (ddlQuery.toUpperCase().contains("CREATE TABLE")) {
37.
38.
                                try (Statement stmt = con.createStatement()) {
39.
                                    stmt.executeUpdate(ddlQuery);
                                    System.out.println("Created table(s) from " +
40.
   ddlFile.getName());
41.
                                }
42.
                            } else {
                                System.out.println("No CREATE TABLE queries found in
43.
   tpch_create.sql.");
44.
                            }
45.
                        } else {
46.
                            System.out.println("Skipping non-CREATE TABLE SQL file: "
   + ddlFile.getName());
47.
48.
                   } catch (IOException e) {
                        System.out.println("Error reading DDL file: " +
49.
   ddlFile.getName() + " - " + e.getMessage());
50.
                    } catch (SQLException e) {
                        System.out.println("Error executing DDL query for file " +
                        " - " + e.getMessage());
   ddlFile.getName() +
52.
```

```
53.     }
54.     } else {
55.         System.out.println("No DDL files found in 'ddl' folder.");
56.     }
57.     }
Output:
```

4. Write the method insert() to add the standard TPC-H data into the database. The DDL files are in the ddl folder. Hint: Files are designed so can read entire file as a string and execute it as one statement. May need to divide up into batches for large files. [10]

```
1. public void insert() {
           File dataFolder = new File("ddl");
           File[] dataFiles = dataFolder.listFiles((dir, name) ->
3.
   name.endsWith(".sql") && !name.equals("tpch_create.sql"));
4.
5.
           if (dataFiles != null) {
6.
               // Use a CountDownLatch to wait for all insertions to complete
7.
               CountDownLatch latch = new CountDownLatch(dataFiles.length);
8.
9.
               for (File dataFile : dataFiles) {
                   executorService.submit(() -> {
10.
11.
                       processFile(dataFile);
                       latch.countDown(); // Decrease latch count when each file is
12.
   processed
13.
                   });
14.
               }
15.
               try {
16.
                   // Wait until all insert tasks are completed
17.
                   latch.await();
18.
                   System.out.println("All data insertions completed.");
19.
               } catch (InterruptedException e) {
                   System.out.println("Error waiting for insertions to complete: " +
20.
   e.getMessage());
21.
22.
           } else {
23.
               System.out.println("No data files found in the folder.");
           }
24.
25.
       private void processFile(File dataFile) {
26.
27.
           try {
28.
                // Read the SQL query from the file
29.
               String sqlQuery = new String(Files.readAllBytes(dataFile.toPath()));
30.
31.
               // Log the start of the process
32.
               System.out.println("Processing file: " + dataFile.getName());
33.
34.
               // Open the Statement for executing the insert query
35.
               try (Statement stmt = con.createStatement()) {
```

```
36.
                    // Split the SQL query into individual statements (assuming
   multiple INSERT statements in the file)
37.
                    String[] insertStatements = sqlQuery.split(";");
38.
                    int totalStatements = insertStatements.length; // Total number of
   statements in the file
39.
40.
                    int batchCount = 0;
41.
                    int recordCount = 0;
42.
43.
                    // Process each INSERT statement
                    for (String statement : insertStatements) {
44.
45.
                        if (statement.trim().isEmpty()) continue;
46.
47.
                        // Add the statement to the batch
48.
                        stmt.addBatch(statement.trim());
49.
                        batchCount++;
50.
                        // Execute the batch every 500 records
51.
                        if (batchCount % 500 == 0) {
52.
53.
                            stmt.executeBatch(); // Execute the batch
54.
                            recordCount += 500;
55.
56.
                            // Calculate and display progress
57.
                            double percentageCompleted = (recordCount / (double)
   totalStatements) * 100;
                            int remainingRecords = totalStatements - recordCount;
58.
                            System.out.printf("File: %s | Inserted: %d | Remaining:
   %d | Progress: %.2f%%%n",
60.
                                    dataFile.getName(), recordCount,
   remainingRecords, percentageCompleted);
61.
                        }
62.
63.
                    if (batchCount > 0) {
64.
                        stmt.executeBatch();
                        recordCount += batchCount;
65.
                        double percentageCompleted = (recordCount / (double)
   totalStatements) * 100;
67.
                        int remainingRecords = totalStatements - recordCount;
                        System.out.printf("File: %s | Inserted: %d | Remaining: %d |
68.
   Progress: %.2f%%n",
                                dataFile.getName(), recordCount, remainingRecords,
69.
   percentageCompleted);
70.
                    }
71.
                    System.out.println("Insertion completed for file: " +
   dataFile.getName());
73.
74.
               } catch (SQLException e) {
                    System.out.println("Error executing statement: " +
75.
   e.getMessage());
76.
                    e.printStackTrace();
77.
                }
78.
           } catch (IOException e) {
```

```
C:\Windows\system32\cmd.exe - java --add-opens java.base/java.net=ALL-UNNAMED -cp "Drivers/*,." AmazonRedshift
 ==== Amazon Redshift Menu =====
1. Connect to the database
2. Drop all tables
 3. Create schema and tables
4. Insert TPC-H Data
5. Execute Query 1
6. Execute Query 2
7. Execute Query 3
8. Close the connection
9. Exit
Enter your choice: 4
Processing file: nation.sql
Processing file: customer.sql
Processing file: orders.sql
Processing file: lineitem.sql
File: customer.sql | Inserted: 500 | Remaining: 1000 | Progress: 33.33% File: nation.sql | Inserted: 25 | Remaining: 0 | Progress: 100.00% Insertion completed for file: nation.sql
File: orders.sql | Inserted: 500 | Remaining: 14500 | Progress: 3.33%
Processing file: part.sql
File: customer.sql | Inserted: 1000 | Remaining: 500 | Progress: 66.67%
```

5. Write the method query1() that returns the most recent top 10 orders with the total sale and the date of the order for customers in America. [5]

```
1. public ResultSet query1() throws SQLException {
2.
           System.out.println("Executing query #1.");
3.
           String query = " SELECT o.O_ORDERKEY AS order_key, " +
                        " SUM(1.L_EXTENDEDPRICE) AS total_sale, " +
4.
                        " o.O_ORDERDATE AS order_date " +
5.
                        " FROM ORDERS o " +
6.
                        " JOIN LINEITEM 1 ON o.O_ORDERKEY = 1.L_ORDERKEY " +
7.
                        " JOIN CUSTOMER c ON o.O_CUSTKEY = c.C_CUSTKEY " +
8.
                        " JOIN NATION n ON c.C_NATIONKEY = n.N_NATIONKEY " +
9.
                        " WHERE n.N NAME = 'UNITED STATES' " +
10.
                        " GROUP BY o.O_ORDERKEY, o.O_ORDERDATE " +
11.
                        " ORDER BY o.O_ORDERDATE DESC " +
12.
                        " LIMIT 10 ";
13.
14.
           if (con == null) {
               throw new SQLException("Connection is null. Please connect to the
   database first.");
16.
           }
17.
18.
           Statement stmt = con.createStatement();
```

```
19. return stmt.executeQuery(query);
20. }
```

```
🖭 C:\Windows\system32\cmd.exe - java --add-opens java.base/java.net=ALL-UNNAMED -cp "Drivers/*;." AmazonRedshift
   == Amazon Redshift Menu =====
1. Connect to the database
2. Drop all tables
. Create schema and tables
. Insert TPC-H Data
. Execute Query 1
 . Execute Query
 . Execute Query 3
  Close the connection
. Exit
Enter your choice: 5
Executing query #1.
 Order Key | Total Sale | Order Date
 28615
                 202064.00
                                 2018-12-25
 4484
                 312330.00
                                 2018-12-24
 23938
                 181490.00
                                 2018-12-06
 49028
                 21803.00
                                 2018-12-06
  50114
                 208291.00
                                 2018-11-29
 30531
                 72428.00
                                 2018-11-27
 30816
                 41644.00
                                 2018-11-26
                 65230.00
 6247
                                 2018-11-22
 18342
                 260993.00
                                 2018-11-19
 5831
                 139126.00
                                 2018-11-17
```

6. Write the method query2() that returns the customer key and the total price a customer spent in descending order, for all urgent orders that are not failed for all customers who are outside Europe and belong to the largest market segment. The largest market segment is the market segment with the most customers. [10]

```
1. public ResultSet query2() throws SQLException {
2.
           System.out.println("Executing query #2.");
           String segmentQuery = "SELECT C_MKTSEGMENT " +
3.
4.
                                  "FROM CUSTOMER " +
                                  "GROUP BY C MKTSEGMENT " +
5.
                                  "ORDER BY COUNT(C_CUSTKEY) DESC " +
6.
7.
                                  "LIMIT 1";
8.
9.
           Statement stmt = con.createStatement();
10.
           ResultSet segmentResult = stmt.executeQuery(segmentQuery);
11.
           String largestSegment = null;
12.
           if (segmentResult.next()) {
               largestSegment = segmentResult.getString("C_MKTSEGMENT");
13.
14.
           }
15.
           if (largestSegment != null) {
                String query = "WITH NonEuropeanCustomers AS ( " +
16.
```

```
17.
                                    SELECT C.C CUSTKEY " +
18.
                                    FROM CUSTOMER C " +
                               ...
19.
                                    JOIN NATION N ON C.C NATIONKEY = N.N NATIONKEY "
20.
                                    JOIN REGION R ON N.N_REGIONKEY = R.R_REGIONKEY "
                                    WHERE R.R_NAME != 'EUROPE' " +
21.
                               "), " +
22.
                               "FilteredCustomers AS ( " +
23.
24.
                                    SELECT C.C CUSTKEY " +
25.
                                    FROM CUSTOMER C " +
26.
                                    WHERE C.C_MKTSEGMENT = ? " +
27.
                                    AND C.C CUSTKEY IN (SELECT C CUSTKEY FROM
   NonEuropeanCustomers) " +
                               "), " +
28.
                               "UrgentOrders AS ( " +
29.
30.
                                    SELECT 0.0_CUSTKEY AS CustomerKey,
   SUM(L.L EXTENDEDPRICE) AS TotalSpent " +
                                    FROM ORDERS O " +
31.
32.
                                    JOIN LINEITEM L ON O.O_ORDERKEY = L.L_ORDERKEY "
                               11
                                    WHERE 0.0 ORDERPRIORITY = '1-URGENT' " +
33.
                                      AND O.O_ORDERSTATUS != 'F' " +
34.
                               ...
                                      AND O.O CUSTKEY IN (SELECT C CUSTKEY FROM
35.
   FilteredCustomers) " +
                                    GROUP BY 0.0 CUSTKEY " +
36.
                               ") " +
37.
                               "SELECT U.CustomerKey, U.TotalSpent " +
38.
                               "FROM UrgentOrders U " +
39.
                               "ORDER BY U.TotalSpent DESC";
40.
41.
               PreparedStatement preparedStatement = con.prepareStatement(query);
42.
43.
                preparedStatement.setString(1, largestSegment); // Set the largest
  market segment dynamically
44.
               return preparedStatement.executeQuery();
45.
           System.out.println("No largest market segment found.");
46.
47.
           return null;
48.
```

```
===== Amazon Redshift Menu =====
1. Connect to the database
2. Drop all tables
3. Create schema and tables
4. Insert TPC-H Data
5. Execute Query 1
6. Execute Query 2
7. Execute Query 3
8. Close the connection
9. Exit
Enter your choice: 6
Executing query #2.
Query #2 Results:
```

Customer Key	Total Spent
962	934276.00
1052	833829.00
103	772835.00
1279	740936.00
1061	734090.00
664	658041.00
1331	654336.00
1415	622849.00
334	620177.00
1316	611696.00
1334	601828.00
1144	601810.00
1345	574253.00
340	570314.00
1013	555097.00
1027	545509.00
694	541472.00
1253 818	538708.00 529872.00
1124	526399.00
229	522166.00
380	513735.00
835	511995.00
575	504550.00
1214	503437.00
1268	489049.00
188	475074.00
767	457100.00
995	456724.00
932	453079.00
134	438663.00
1486	437668.00
1075	425282.00
512	422890.00
1 329	421274.00 413687.00
662	404588.00
844	403041.00
649	402030.00
674	400317.00
508	395928.00
1223	386810.00
814	382438.00
1046	374573.00
803	368071.00
592	366214.00
938	365671.00
185	363520.00
580	359714.00
1414 709	355840.00 355246.00
968	355246.00 353066.00
553	344943.00
298	334711.00
1163	334638.00
1100	327751.00
1009	327475.00
1115	319143.00

```
805
568
                 316412.00
                 312407.00
1433
                 308212.00
1237
                 308206.00
1202
                 308009.00
1091
                 306579.00
1400
                 305537.00
77
                 304452.00
1295
                 299133.00
860
                 297132.00
1430
                 296181.00
1235
                 292461.00
1453
                 292215.00
220
                 290881.00
610
                 284676.00
152
                 283341.00
476
                 279524.00
722
                 275900.00
986
                 271477.00
73
                 259207.00
428
                 257247.00
116
                 248579.00
1318
                 244310.00
223
                 244030.00
1208
                 243368.00
211
                 242005.00
1405
                 238524.00
1475
                 237839.00
944
                 237457.00
40
                 235070.00
475
                 234605.00
790
                 233068.00
280
                 232699.00
1201
                 230542.00
392
                 224661.00
523
                 224453.00
1183
                 220834.00
1460
                 217412.00
98
                 215161.00
1312
                 212232.00
                 210170.00
653
400
                 204820.00
1040
                 198800.00
1114
                 194242.00
1006
                 193509.00
224
                 191248.00
278
                 191091.00
1085
                 190796.00
1180
                 189196.00
1412
                 187654.00
865
                 183851.00
766
                 182136.00
763
                 175047.00
170
                 171616.00
1196
                 169535.00
1447
                 167667.00
1184
                 166621.00
64
                 164606.00
350
                 162400.00
200
                 159039.00
784
                 146507.00
```

```
113
                146103.00
826
                144242.00
670
                143332.00
515
                142196.00
448
                141324.00
328
                139529.00
802
                138073.00
1396
                130749.00
811
                128303.00
347
                117633.00
1358
                117503.00
905
                116995.00
548
                115519.00
1330
                112946.00
1357
                109970.00
1370
                108208.00
1261
                106892.00
562
                105994.00
904
                101642.00
1385
                101477.00
296
                100987.00
542
                99021.00
1082
                96900.00
728
                96476.00
793
                93905.00
647
                91988.00
535
                91529.00
859
                90163.00
518
                89866.00
602
                89745.00
13
                86102.00
419
                81138.00
121
                77998.00
1277
                76457.00
109
                74685.00
557
                73613.00
205
                71835.00
32
                64622.00
1468
                58353.00
623
                53127.00
785
                48653.00
890
                43337.00
1033
                41460.00
386
                33790.00
221
                32520.00
1292
                28556.00
478
                10756.00
```

7. Write the method query3() that returns a count of all the line items that were ordered within the six years starting on April 1st, 1997 group by order priority. Make sure to sort by order priority in ascending order. [10]

```
String query = "SELECT o.O ORDERPRIORITY, COUNT(1.L LINENUMBER) AS
   lineitem count " +
4.
                           "FROM orders o " +
5.
                           "JOIN lineitem 1 ON o.O ORDERKEY = 1.L ORDERKEY " +
                           "WHERE o.O_ORDERDATE >= '1997-04-01' " +
6.
                           "AND o.O_ORDERDATE < DATEADD(year, 6, '1997-04-01') " +
7.
                           "GROUP BY o.O_ORDERPRIORITY " +
8.
9.
                           "ORDER BY o.O_ORDERPRIORITY ASC";
10.
           Statement stmt = con.createStatement();
11.
           ResultSet rs = stmt.executeQuery(query);
12.
13.
           return rs;
14.
```

```
🖭 C:\Windows\system32\cmd.exe - java --add-opens java.base/java.net=ALL-UNNAMED -cp "Drivers/*;," AmazonRedshift
 ==== Amazon Redshift Menu =====
1. Connect to the database
2. Drop all tables
3. Create schema and tables
4. Insert TPC-H Data
 5. Execute Query 1
5. Execute Query 2
7. Execute Query 3
8. Close the connection
9. Exit
Enter vour choice: 7
Executing query #3.
 Order Priority | Line Item Count
  1-URGENT
                      1387
  2-HIGH
                      1303
  3-MEDIUM
                      1287
  4-NOT SPECIFIED
                      1530
  5-LOW
                      1268
```

#### Full Code:

```
    import java.math.BigDecimal;

import java.nio.file.Files;
import java.sql.Connection;
import java.sql.DriverManager;
5. import java.sql.PreparedStatement;
import java.sql.ResultSet;
import java.sql.ResultSetMetaData;
import java.sql.SQLException;
import java.sql.Statement;
10.import java.util.Arrays;
11.import java.util.List;
12.import java.util.Properties;
13.import java.util.Scanner;
14.import java.io.File;
15.import java.io.IOException;
16.import java.util.concurrent.CountDownLatch;
```

```
17.import java.util.concurrent.ExecutorService;
18. import java.util.concurrent.Executors;
19.
20./**
21. * Performs SQL DDL and SELECT queries on a MySQL database hosted on AWS RDS.
23. * java --add-opens java.base/java.net=ALL-UNNAMED -cp "Drivers/*;."
   AmazonRedshift
24. */
25.public class AmazonRedshift {
       /**
26.
27.
        * Connection to database
        */
28.
29.
       static final String redshiftUrl = "jdbc:redshift://g23ai2087-
   cluster.cndskybccs1r.us-east-1.redshift.amazonaws.com:5439/dev";
       static final String masterUsername = "admin"; // Replace with your Redshift
30.
   admin username
       static final String password = "IITj1234"; // Replace with your Redshift
   password
32.
       private static ExecutorService executorService;
       /**
33.
34.
        * Main method is only used for convenience. Use JUnit test file to verify
  your
        * answer.
35.
36.
37.
        * @param args
        *
                      none expected
        * @throws SQLException
39.
40.
                               if a database error occurs
41.
42.
       public static void main(String[] args) {
43.
           Scanner scanner = new Scanner(System.in);
44.
           AmazonRedshift q = new AmazonRedshift();
45.
           // Initialize the executor service for parallel execution
46.
           executorService = Executors.newFixedThreadPool(4);
47.
           ResultSet rs;
48.
           while (true) {
49.
               // Display menu
               System.out.println("\n==== Amazon Redshift Menu =====");
50.
51.
               System.out.println("1. Connect to the database");
52.
               System.out.println("2. Drop all tables");
53.
               System.out.println("3. Create schema and tables");
               System.out.println("4. Insert TPC-H Data");
54.
55.
               System.out.println("5. Execute Query 1");
56.
               System.out.println("6. Execute Query 2");
               System.out.println("7. Execute Query 3");
57.
               System.out.println("8. Close the connection");
58.
59.
               System.out.println("9. Exit");
60.
               System.out.print("Enter your choice: ");
61.
               int choice = scanner.nextInt();
62.
               try {
63.
64.
                   switch (choice) {
65.
                       case 1:
```

```
66.
                         q.connect();
67.
                         break;
68.
                     case 2:
69.
                         q.drop();
70.
                         break;
71.
                     case 3:
72.
                         q.create();
73.
                         break;
74.
                     case 4:
75.
                         q.insert();
76.
                         break;
77.
78.
                     case 5:
79.
                         if (q.con == null) {
                             System.out.println("Please connect to the database
80.
   first.");
81.
                         } else {
82.
                             try {
83.
                                 rs = q.query1();
84.
85.
                                 // Display header for the table
                                 System.out.println("===============
86.
   =======");
                                 System.out.println(" | Order Key | Total
87.
   Sale
         Order Date
                        |");
88.
                                 =======");
89.
                                 // Process and display each row in the ResultSet
90.
91.
                                 while (rs.next()) {
                                    System.out.printf("| %-12d | %-12.2f | %-13s
92.
   | n",
                                                     rs.getInt("order_key"),
93.
      // Column alias from the query
94.
                                                     rs.getDouble("total_sale"),
     // Column alias from the query
95.
                                                     rs.getDate("order_date"));
      // Column alias from the query
96.
97.
98.
                                 =======");
99.
                             } catch (SQLException e) {
                                    System.out.println("Error processing result
100.
   set: " + e.getMessage());
101.
                                }
102.
103.
                            break;
104.
                         case 6:
105.
                            try {
106.
                                rs = q.query2();
107.
                                if (rs != null) {
                                    System.out.println("Query #2 Results:");
108.
```

```
109.
                                =======");
                                System.out.printf("%-15s %-20s%n", "Customer
110.
  Key", "Total Spent");
                                System.out.println("==============
111.
  =======");
                                while (rs.next()) {
112.
                                    System.out.printf("%-15d %-20.2f%n",
113.
114.
                                          rs.getInt("CustomerKey"),
115.
                                          rs.getDouble("TotalSpent"));
116.
                                System.out.println("==============
117.
  ======="";
118.
                             } else {
                                System.out.println("No results found for Query
119.
  #2.");
120.
121.
                          } catch (SQLException e) {
                             System.out.println("Error executing Query #2: " +
122.
  e.getMessage());
123.
                          }
124.
                          break;
125.
                      case 7:
126.
                      if (q.con == null) {
127.
                          System.out.println("Please connect to the database
  first.");
128.
                      } else {
129.
                          rs = q.query3();
                          ======");
                          System.out.println(" | Order Priority
131.
                                                           | Line Item
           |");
  Count
                          132.
  ======");
                         while (rs.next()) {
133.
                             String orderPriority =
   rs.getString("O_ORDERPRIORITY");
                             int lineItemCount = rs.getInt("lineitem_count");
135.
                             System.out.printf("| %-16s | %-18d |\n",
  orderPriority, lineItemCount);
137.
                          138.
  ======");
139.
                      }
140.
                          break;
141.
                      case 8:
                          q.close();
142.
143.
                          break;
144.
                      case 9:
145.
                          System.out.println("Exiting program...");
                          scanner.close();
146.
                          System.exit(0); // Exit the program
147.
                          break;
148.
149.
                      default:
```

```
150.
                                System.out.println("Invalid choice. Please try
   again.");
151.
                               break;
152.
                       }
                   } catch (SQLException e) {
153.
                       System.out.println("Error: " + e.getMessage());
154.
155.
                   } catch (Exception e) {
                       System.out.println("An unexpected error occurred: " +
156.
   e.getMessage());
157.
158.
               }
159.
           }
160.
           /**
161.
            * Makes a connection to the database and returns connection to caller.
162.
163.
           * @return
164.
                      connection
165.
            * @throws SQLException
166.
167.
                                    if an error occurs
            */
168.
169.
           // Redshift connection details
170.
171.
172.
173.
           private Connection con;
174.
175.
           public Connection connect() {
176.
               try {
177.
                   Class.forName("com.amazon.redshift.jdbc42.Driver");
178.
                   Properties properties = new Properties();
                   properties.setProperty("user", masterUsername);
179.
                   properties.setProperty("password", password);
180.
181.
                   this.con = DriverManager.getConnection(redshiftUrl, properties);
182.
183.
                   System.out.println("Connection to Redshift established
184.
   successfully.");
185.
               } catch (ClassNotFoundException e) {
                   System.out.println("Error: Redshift JDBC driver not found.");
186.
187.
                   e.printStackTrace();
188.
               } catch (SQLException e) {
                   System.out.println("Failed to connect to the database.");
189.
190.
                   e.printStackTrace();
191.
192.
               return con;
           }
193.
194.
            * Closes connection to database.
195.
196.
197.
           public void close() {
198.
               System.out.println("Closing database connection.");
199.
200.
                   if (con != null && !con.isClosed()) {
```

```
201.
                       con.close();
202.
                   }
203.
               } catch (SQLException e) {
204.
                   System.out.println("Error closing the connection.");
205.
              }
206.
          }
207.
208.
          public void drop() {
209.
               System.out.println("Dropping all tables in the 'dev' schema...");
               String dropQuery = "SELECT tablename FROM pg_tables WHERE schemaname =
210.
   'dev'":
211.
212.
              try (Statement stmt = con.createStatement()) {
213.
                   ResultSet rs = stmt.executeQuery(dropQuery);
214.
                   while (rs.next()) {
                       String tableName = rs.getString("tablename");
215.
216.
                       String dropTableQuery = "DROP TABLE IF EXISTS dev." +
   tableName;
217.
                       stmt.executeUpdate(dropTableQuery);
                       System.out.println("Dropped table: " + tableName);
218.
219.
                   }
220.
              } catch (SQLException e) {
221.
                   System.out.println("Error dropping tables: " + e.getMessage());
              }
222.
223.
          }
224.
225.
          public void create() throws SQLException {
226.
              System.out.println("Creating the 'dev' schema and tables...");
               String createSchemaQuery = "CREATE SCHEMA IF NOT EXISTS dev";
227.
228.
              try (Statement stmt = con.createStatement()) {
229.
                   stmt.executeUpdate(createSchemaQuery);
230.
                   System.out.println("Schema 'dev' created.");
231.
               } catch (SQLException e) {
232.
                   System.out.println("Error creating schema: " + e.getMessage());
233.
234.
              File ddlFolder = new File("ddl");
235.
              File[] ddlFiles = ddlFolder.listFiles((dir, name) ->
   name.endsWith(".sql"));
              if (ddlFiles != null) {
236.
237.
                   for (File ddlFile : ddlFiles) {
238.
                       try {
239.
                           String ddlQuery = new
   String(Files.readAllBytes(ddlFile.toPath()));
240.
                           if (ddlFile.getName().equals("tpch_create.sql")) {
241.
                               // Ensure the 'tpch create.sql' contains CREATE TABLE
   queries
                               if (ddlQuery.toUpperCase().contains("CREATE TABLE")) {
242.
                                   try (Statement stmt = con.createStatement()) {
243.
244.
                                       stmt.executeUpdate(ddlQuery);
                                       System.out.println("Created table(s) from " +
   ddlFile.getName());
246.
                                   }
247.
                               } else {
```

```
248.
                                   System.out.println("No CREATE TABLE queries found
   in tpch create.sql.");
249.
                               }
                           } else {
250.
251.
                               System.out.println("Skipping non-CREATE TABLE SQL
   file: " + ddlFile.getName());
252.
253.
                       } catch (IOException e) {
                           System.out.println("Error reading DDL file: " +
   ddlFile.getName() + " - " + e.getMessage());
                       } catch (SQLException e) {
255.
                           System.out.println("Error executing DDL query for file " +
256.
   ddlFile.getName() + " - " + e.getMessage());
257.
                       }
258.
                   }
259.
               } else {
260.
                   System.out.println("No DDL files found in 'ddl' folder.");
               }
261.
262.
263.
           public void insert() {
264.
               File dataFolder = new File("ddl");
265.
               File[] dataFiles = dataFolder.listFiles((dir, name) ->
   name.endsWith(".sql") && !name.equals("tpch_create.sql"));
266.
267.
              if (dataFiles != null) {
268.
                   // Use a CountDownLatch to wait for all insertions to complete
                   CountDownLatch latch = new CountDownLatch(dataFiles.length);
269.
270.
271.
                   for (File dataFile : dataFiles) {
272.
                       executorService.submit(() -> {
273.
                           processFile(dataFile);
                           latch.countDown(); // Decrease latch count when each file
274.
  is processed
275.
                       });
276.
                   }
277.
                   try {
                       // Wait until all insert tasks are completed
278.
279.
                       latch.await();
                       System.out.println("All data insertions completed.");
280.
281.
                   } catch (InterruptedException e) {
                       System.out.println("Error waiting for insertions to complete:
     + e.getMessage());
283.
                   }
284.
              } else {
285.
                   System.out.println("No data files found in the folder.");
286.
287.
           private void processFile(File dataFile) {
288.
289.
290.
                   // Read the SQL query from the file
291.
                   String sqlQuery = new
   String(Files.readAllBytes(dataFile.toPath()));
292.
293.
                   // Log the start of the process
```

```
294.
                   System.out.println("Processing file: " + dataFile.getName());
295.
296.
                   // Open the Statement for executing the insert query
                   try (Statement stmt = con.createStatement()) {
297.
298.
                       // Split the SQL query into individual statements (assuming
   multiple INSERT statements in the file)
                       String[] insertStatements = sqlQuery.split(";");
299.
300.
                       int totalStatements = insertStatements.length; // Total number
   of statements in the file
301.
302.
                       int batchCount = 0;
303.
                       int recordCount = 0;
304.
305.
                       // Process each INSERT statement
                       for (String statement : insertStatements) {
306.
307.
                           if (statement.trim().isEmpty()) continue;
308.
309.
                           // Add the statement to the batch
310.
                           stmt.addBatch(statement.trim());
311.
                           batchCount++;
312.
                           // Execute the batch every 500 records
313.
                           if (batchCount % 500 == 0) {
314.
315.
                               stmt.executeBatch(); // Execute the batch
316.
                               recordCount += 500;
317.
318.
                               // Calculate and display progress
319.
                               double percentageCompleted = (recordCount / (double)
   totalStatements) * 100;
320.
                               int remainingRecords = totalStatements - recordCount;
                               System.out.printf("File: %s | Inserted: %d |
321.
   Remaining: %d | Progress: %.2f%%%n",
322.
                                        dataFile.getName(), recordCount,
   remainingRecords, percentageCompleted);
323.
324.
                       if (batchCount > 0) {
325.
                           stmt.executeBatch();
326.
327.
                           recordCount += batchCount;
                           double percentageCompleted = (recordCount / (double)
328.
   totalStatements) * 100;
329.
                           int remainingRecords = totalStatements - recordCount;
330.
                           System.out.printf("File: %s | Inserted: %d | Remaining: %d
   | Progress: %.2f%%n",
                                   dataFile.getName(), recordCount, remainingRecords,
   percentageCompleted);
332.
333.
334.
                       System.out.println("Insertion completed for file: " +
   dataFile.getName());
335.
336.
                   } catch (SQLException e) {
337.
                       System.out.println("Error executing statement: " +
e.getMessage());
```

```
338.
                       e.printStackTrace();
339.
                   }
340.
341.
              } catch (IOException e) {
                   System.out.println("Error reading file: " + dataFile.getName() + "
   - " + e.getMessage());
                   e.printStackTrace(); // Optional: log the stack trace for more
343.
  details
344.
              }
          }
345.
346.
347.
          /**
348.
349.
           * Query returns the most recent top 10 orders with the total sale and the
  date
           * of the
350.
351.
           * order in `America`.
352.
           * @return
353.
354.
                     ResultSet
           * @throws SQLException
355.
356.
                                   if an error occurs
357.
358.
          public ResultSet query1() throws SQLException {
359.
              System.out.println("Executing query #1.");
              String query = " SELECT o.O_ORDERKEY AS order_key, " +
360.
                           " SUM(1.L EXTENDEDPRICE) AS total sale, " +
361.
                           " o.O_ORDERDATE AS order_date " +
362.
                           " FROM ORDERS o " +
363.
                           " JOIN LINEITEM 1 ON o.O_ORDERKEY = 1.L_ORDERKEY " +
364.
                           " JOIN CUSTOMER c ON o.O_CUSTKEY = c.C_CUSTKEY " +
365.
                           " JOIN NATION n ON c.C NATIONKEY = n.N NATIONKEY " +
366.
                           " WHERE n.N_NAME = 'UNITED STATES' " +
367.
                           " GROUP BY o.O_ORDERKEY, o.O_ORDERDATE " +
368.
                           " ORDER BY o.O_ORDERDATE DESC " +
369.
370.
                           " LIMIT 10 ";
371.
              if (con == null) {
                   throw new SQLException("Connection is null. Please connect to the
372.
   database first.");
373.
              }
374.
375.
              Statement stmt = con.createStatement();
376.
              return stmt.executeQuery(query);
377.
          }
378.
379.
380.
          /**
381.
382.
           * Query returns the customer key and the total price a customer spent in
383.
           * descending
           * order, for all urgent orders that are not failed for all customers who
384.
  are
385.
           * outside Europe belonging
386.
           * to the highest market segment.
```

```
387.
388.
             @return
389.
                      ResultSet
           * @throws SQLException
390.
                                   if an error occurs
391.
           */
392.
393.
          public ResultSet query2() throws SQLException {
394.
               System.out.println("Executing query #2.");
              String segmentQuery = "SELECT C_MKTSEGMENT " +
395.
                                     "FROM CUSTOMER " +
396.
                                     "GROUP BY C MKTSEGMENT " +
397.
398.
                                     "ORDER BY COUNT(C_CUSTKEY) DESC " +
399.
                                     "LIMIT 1";
400.
401.
              Statement stmt = con.createStatement();
402.
              ResultSet segmentResult = stmt.executeQuery(segmentQuery);
403.
              String largestSegment = null;
404.
              if (segmentResult.next()) {
                   largestSegment = segmentResult.getString("C_MKTSEGMENT");
405.
406.
407.
              if (largestSegment != null) {
                   String query = "WITH NonEuropeanCustomers AS ( " +
408.
409.
                                       SELECT C.C_CUSTKEY " +
                                  ...
410.
                                       FROM CUSTOMER C " +
411.
                                       JOIN NATION N ON C.C NATIONKEY = N.N NATIONKEY
                                       JOIN REGION R ON N.N REGIONKEY = R.R REGIONKEY
                                       WHERE R.R NAME != 'EUROPE' " +
413.
                                  "), " +
414.
                                  "FilteredCustomers AS ( " +
415.
                                       SELECT C.C CUSTKEY " +
416.
                                  ...
                                       FROM CUSTOMER C " +
417.
418.
                                       WHERE C.C MKTSEGMENT = ? " +
419.
                                       AND C.C_CUSTKEY IN (SELECT C_CUSTKEY FROM
   NonEuropeanCustomers) " +
420.
                                  "), " +
                                  "UrgentOrders AS ( " +
421.
422.
                                       SELECT 0.0_CUSTKEY AS CustomerKey,
   SUM(L.L EXTENDEDPRICE) AS TotalSpent " +
423.
                                       FROM ORDERS O " +
424.
                                       JOIN LINEITEM L ON O.O ORDERKEY = L.L ORDERKEY
                                       WHERE 0.0 ORDERPRIORITY = '1-URGENT' " +
425.
                                         AND O.O_ORDERSTATUS != 'F' " +
426.
427.
                                         AND O.O_CUSTKEY IN (SELECT C_CUSTKEY FROM
   FilteredCustomers) " +
                                       GROUP BY 0.0 CUSTKEY " +
428.
                                  ") " +
429.
430.
                                  "SELECT U.CustomerKey, U.TotalSpent " +
431.
                                  "FROM UrgentOrders U " +
                                  "ORDER BY U.TotalSpent DESC";
432.
433.
434.
                   PreparedStatement preparedStatement = con.prepareStatement(query);
```

```
435.
                   preparedStatement.setString(1, largestSegment); // Set the largest
   market segment dynamically
436.
                   return preparedStatement.executeQuery();
437.
438.
              System.out.println("No largest market segment found.");
439.
               return null;
440.
          }
441.
          /**
442.
           * Query returns all the lineitems that was ordered within the six years
443.
  from
           * January 4th,
444.
445.
           * 1997 and the orderpriority in ascending order.
446.
           * @return
447.
448.
                      ResultSet
           * @throws SQLException
449.
450.
                                   if an error occurs
451.
452.
          public ResultSet query3() throws SQLException {
453.
              System.out.println("Executing query #3.");
454.
              String query = "SELECT o.O_ORDERPRIORITY, COUNT(1.L_LINENUMBER) AS
  lineitem_count " +
                              "FROM orders o " +
455.
                              "JOIN lineitem 1 ON o.O_ORDERKEY = 1.L_ORDERKEY " +
456.
457.
                              "WHERE o.O ORDERDATE >= '1997-04-01' " +
458.
                              "AND o.O ORDERDATE < DATEADD(year, 6, '1997-04-01') " +
459.
                              "GROUP BY o.O_ORDERPRIORITY " +
                              "ORDER BY o.O_ORDERPRIORITY ASC";
460.
461.
              Statement stmt = con.createStatement();
462.
              ResultSet rs = stmt.executeQuery(query);
463.
464.
              return rs;
465.
          }
466.
467.
468.
           * Do not change anything below here.
469.
           */
470.
471.
          /**
472.
           * Converts a ResultSet to a string with a given number of rows displayed.
473.
           * Total rows are determined but only the first few are put into a string.
474.
           * @param rst
475.
476.
                             ResultSet
           * @param maxrows
477.
478.
                             maximum number of rows to display
           * @return
479.
480.
                      String form of results
481.
           * @throws SQLException
                                   if a database error occurs
482.
483.
484.
          public static String resultSetToString(ResultSet rst, int maxrows) throws
SQLException {
```

```
485.
               StringBuffer buf = new StringBuffer(5000);
486.
               int rowCount = 0;
487.
               ResultSetMetaData meta = rst.getMetaData();
488.
               buf.append("Total columns: " + meta.getColumnCount());
               buf.append('\n');
489.
               if (meta.getColumnCount() > 0)
490.
491.
                   buf.append(meta.getColumnName(1));
492.
               for (int j = 2; j <= meta.getColumnCount(); j++)</pre>
                    buf.append(", " + meta.getColumnName(j));
493.
               buf.append('\n');
494.
495.
               while (rst.next()) {
496.
                   if (rowCount < maxrows) {</pre>
497.
                        for (int j = 0; j < meta.getColumnCount(); j++) {</pre>
498.
                            Object obj = rst.getObject(j + 1);
499.
                            buf.append(obj);
                            if (j != meta.getColumnCount() - 1)
500.
                                buf.append(", ");
501.
502.
                        buf.append('\n');
503.
504.
                   }
505.
                   rowCount++;
506.
               buf.append("Total results: " + rowCount);
507.
508.
               return buf.toString();
509.
           }
510.
           /**
511.
512.
            * Converts ResultSetMetaData into a string.
513.
514.
              @param meta
515.
                           ResultSetMetaData
            * @return
516.
517.
                      string form of metadata
518.
            * @throws SQLException
                                     if a database error occurs
519.
520.
            */
           public static String resultSetMetaDataToString(ResultSetMetaData meta)
   throws SQLException {
               StringBuffer buf = new StringBuffer(5000);
522.
               buf.append(meta.getColumnName(1) + " (" + meta.getColumnLabel(1) + ",
523.
                        meta.getColumnType(1) + "-" + meta.getColumnTypeName(1) + ", "
524.
525.
                        meta.getColumnDisplaySize(1) + ", " + meta.getPrecision(1) +
                        meta.getScale(1) + ")");
526.
               for (int j = 2; j <= meta.getColumnCount(); j++) {
   buf.append(", " + meta.getColumnName(j) + " (" +</pre>
527.
   meta.getColumnLabel(j) + ",
                            meta.getColumnType(j) + "-" + meta.getColumnTypeName(j) +
529.
                            meta.getColumnDisplaySize(j) + ", " + meta.getPrecision(j)
530.
531.
                            meta.getScale(j) + ")");
```

```
532.     }
533.     return buf.toString();
534.     }
535. }
```