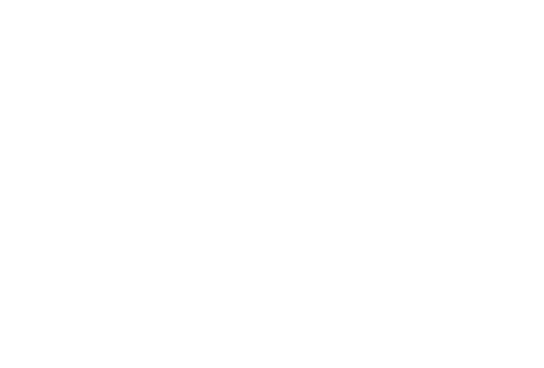


Milgard School of Business, MSBA

White Be Boundless tagline graphic­­­­­DOCUMENT APPROVAL SHEET

TBANLT 510

MODULE 5-ETL EXERCISE

gold boundless bar

**DOCUMENT IDENTIFICATION:**

MSBA AY 21-22/TBANLT510/ASSIGNMENTS/MODULE5 ETL EXCERCISE

**DOCUMENT AUTHOR:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **AUTHORED BY/ ID** | **DEPT./ ROLE** | **COHORT** | **DATE** | **CHECKED BY** | **DATE** |
| Teja Alluru  (1150077) | MSBA/ Student | 2021-2022 | 28-Nov-2021 | Shephali Jain  (2110257) | 28-Nov-2021 |

**DOCUMENT APPROVAL:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **APPROVED BY/ ID** | **DEPT.** | **SCHOOL** | **SIGNATURE** | **DATE** |
| Teja Alluru (1150077) | MSBA/ Student | University of Washington, Tacoma, Milgard School of Business |  | 28-Nov-2021 |

**GRADED BY:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **GRADED BY** | **Role/DEPT.** | **SCHOOL** | **SIGNATURE** | **DATE** |
| Micheal Turek | Faculty/  MSBA | University of Washington, Tacoma, Milgard School of Business |  | 28-Nov-2021 |

*AUTHORED BY: Authored by mean of a person/entity responsible for the creation of the document.*

*APPROVED BY: Approval means a person/entity responsible for the approval of the document content.*

*AUTHORISED BY: Graded by means a person/entity having the authority to grade the assignment.*

*COPYRIGHT: This document was released for the viewing of the MSBA department only.*

INDEX OF CONTENTS

gold boundless bar

[DOCUMENT APPROVAL SHEET 1](#_Toc88998280)

[INDEX OF CONTENTS 3](#_Toc88998281)

[1 ASSIGNMENT TASKS 4](#_Toc88998282)

[2 DATA WAREHOUSE SCHEMA 4](#_Toc88998283)

[3 COMPLETE CREATE DATABASE AND TABLES IN SSMS AND SQL SERVER INTEGRATION SERVICES TUTORIAL 5](#_Toc88998284)

[4 PROVIDE NAME OF YOUR DATABASE 6](#_Toc88998285)

[5 SCREENSHOTS OF DIMENSIONS AND FACT TABLE 7](#_Toc88998286)

[6 PROVIDE TOTAL NUMBER OF ROW COUNT FOR EACH TABLE LOADED ARE AS FOLLOWS 10](#_Toc88998287)

[7 PROVIDE PARTIAL SCREEN SHOTS OF YOUR CREATED AND POPULATED TABLES (dim\_time, dim\_channel, fact\_interactions) 13](#_Toc88998288)

INDEX OF FIGURES

gold boundless bar

[Figure 2‑1: Data Warehouse Schema – Hubspot Data 4](#_Toc88998297)

[Figure 3‑1: Remote Connection – CBA Virtual Machine 5](#_Toc88998298)

[Figure 4‑1: SSMS – Database Created (ETL Exercise) 6](#_Toc88998299)

[Figure 5‑1: Dimension Table – dim\_time 7](#_Toc88998300)

[Figure 5‑2: Dimension Table – dim\_channel 8](#_Toc88998301)

[Figure 5‑3: Fact Table – fact\_interactions 9](#_Toc88998302)

[Figure 6‑1: Dimension Table – Row Count – dim\_channel(3) 10](#_Toc88998303)

[Figure 6‑2: Dimension Table – Row Count – dim\_time(185) 11](#_Toc88998304)

[Figure 6‑3: Fact Table – Row Count – fact\_interaction (938) 12](#_Toc88998305)

[Figure 7‑1: Dimension Table Screenshot – dim\_channel 13](#_Toc88998306)

[Figure 7‑2: Dimension Table Screenshot – dim\_time 14](#_Toc88998307)

[Figure 7‑3: Fact Table Screenshot – fact\_interactions 15](#_Toc88998308)

# ASSIGNMENT TASKS

The following tasks were the point of focus for the ETL assignment

1. Create Database and Tables in SSMS: Step by step tutorial to create your database on SSMS
2. SQL Server Integration Services Tutorial: step by step tutorial on how to use SSIS to perform ETL

# DATA WAREHOUSE SCHEMA

The sample data warehouse scheme we are going to create a database for is shown in Figure 2‑1 below.

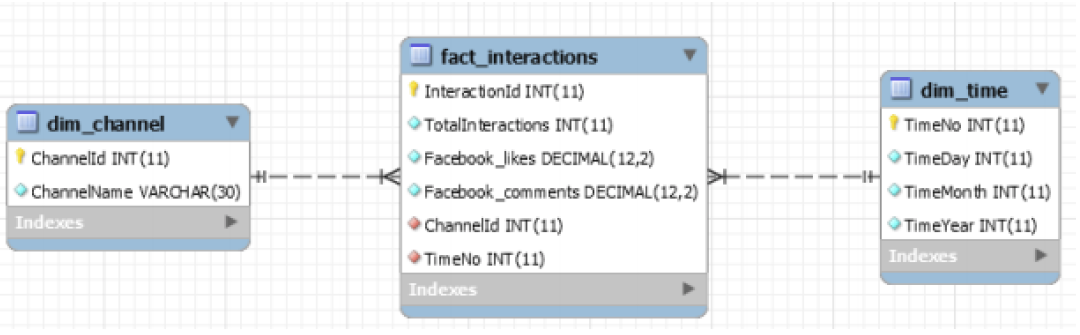


Figure 2‑1: Data Warehouse Schema – Hubspot Data

# COMPLETE CREATE DATABASE AND TABLES IN SSMS AND SQL SERVER INTEGRATION SERVICES TUTORIAL

The database has been created using SSIS, SSMS, and visual studio 2019. CBA virtual machines are accessed and the assignment is completely executed in the virtual system through a remote connection.

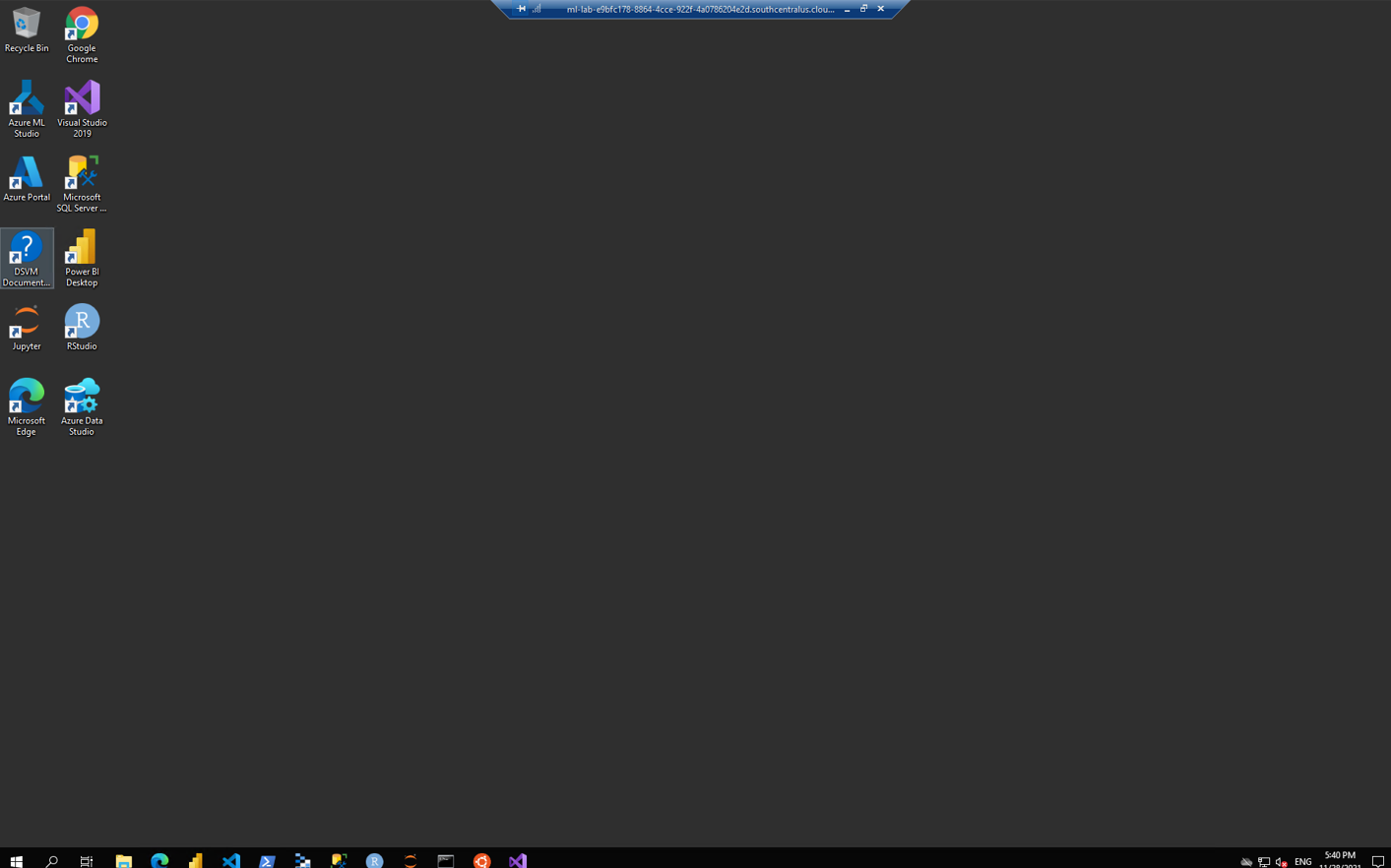


Figure 3‑1: Remote Connection – CBA Virtual Machine

# PROVIDE THE NAME OF YOUR DATABASE

The database by the name ETL exercise is created in SSMS using the SQL script. The screenshot of the created database is presented in Figure 4‑1 below.

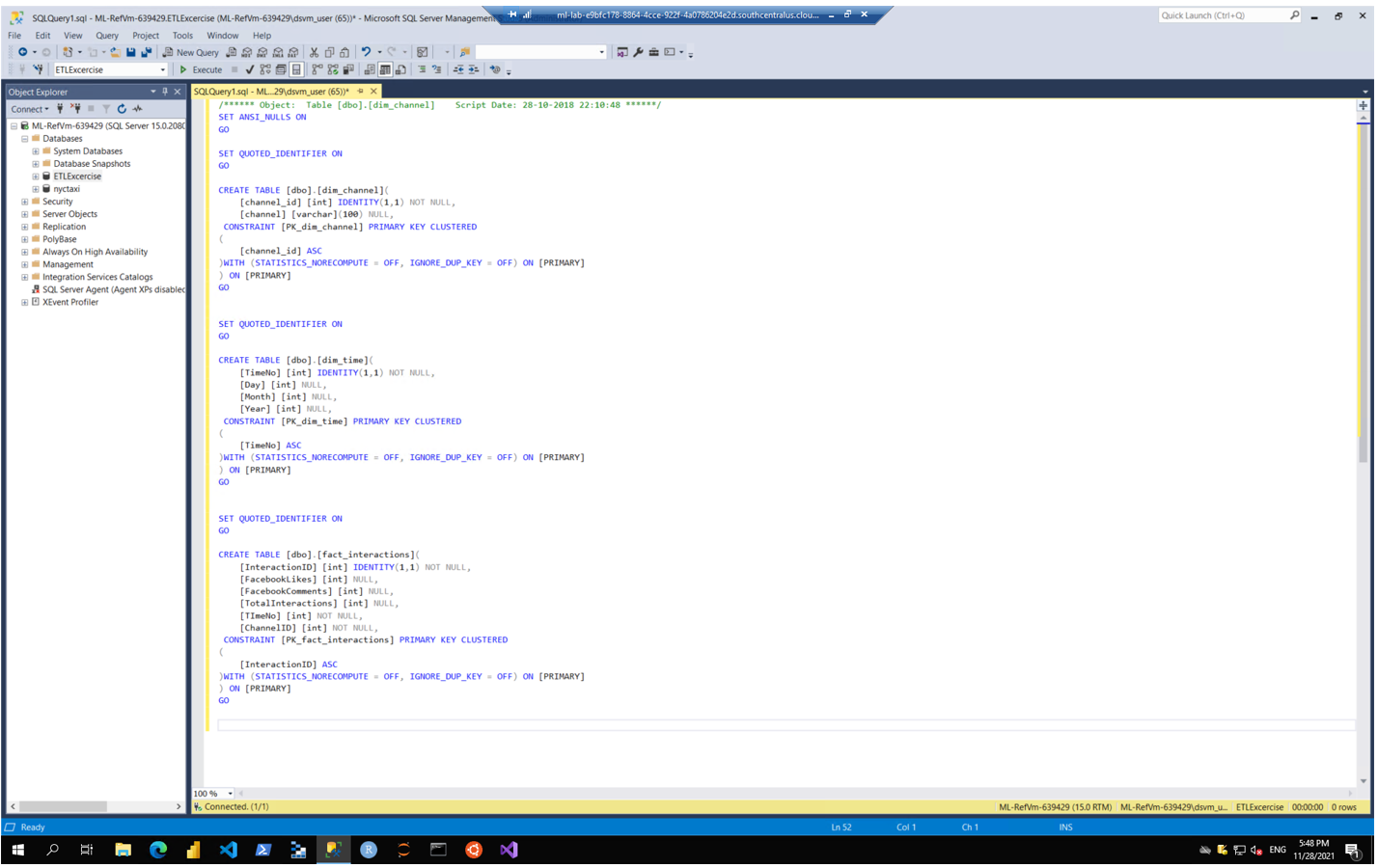


Figure 4‑1: SSMS – Database Created (ETL Exercise)

# SCREENSHOTS OF DIMENSIONS AND FACT TABLE

The screenshots of the dimensions and fact table creation are presented in Figure 5‑1 to Figure 5‑3 below.

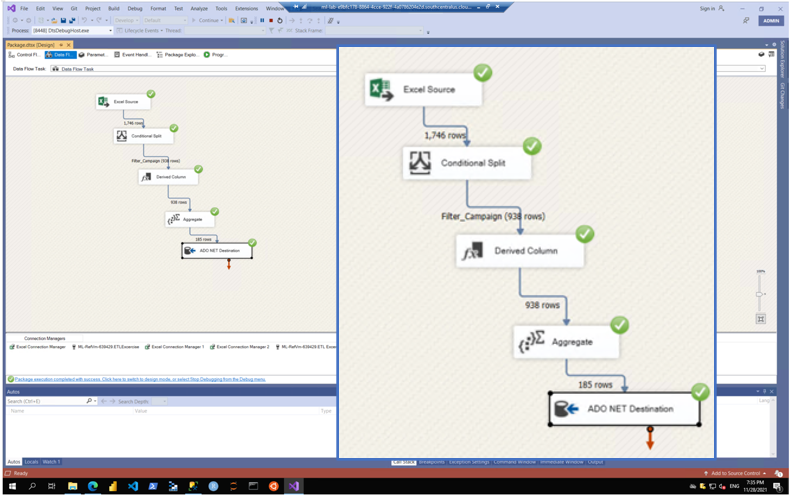


Figure 5‑1: Dimension Table – dim\_time

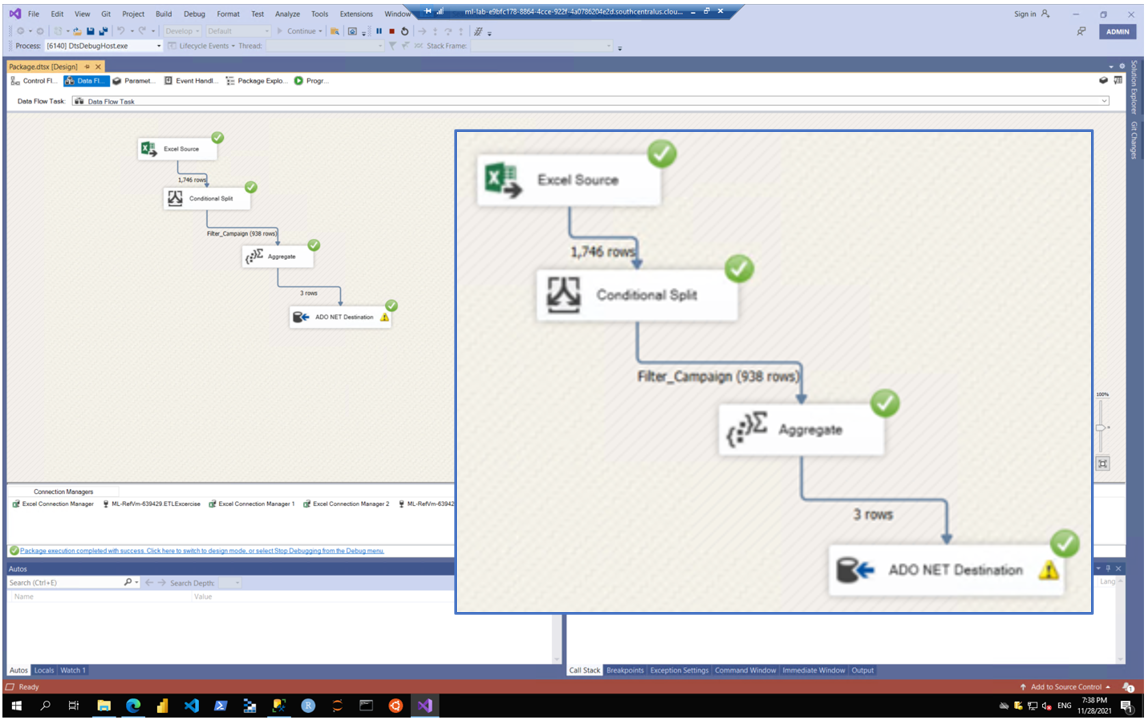


Figure 5‑2: Dimension Table – dim\_channel

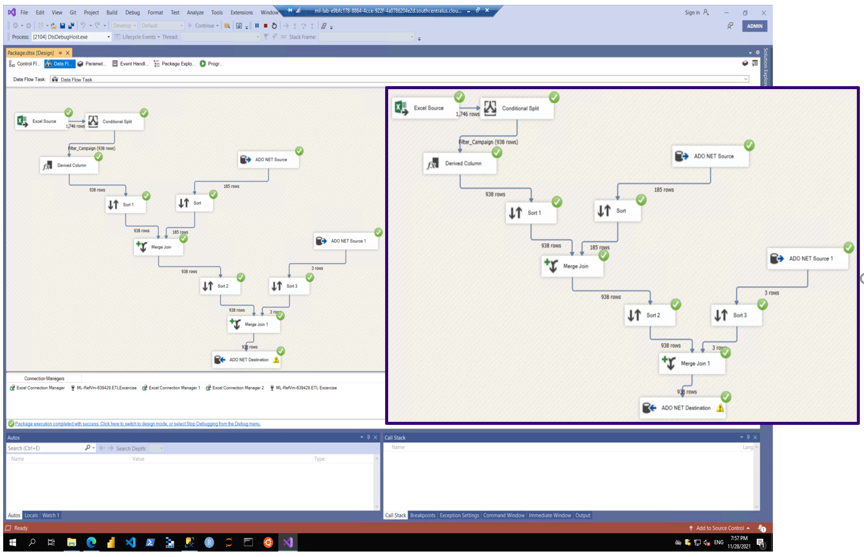


Figure 5‑3: Fact Table – fact\_interactions

# PROVIDE TOTAL NUMBER OF ROW COUNTS FOR EACH TABLE LOADED ARE AS FOLLOWS

The total number of rows for dim\_channel is 3, dim\_time is 185 and facts\_interactions is 938. The screenshots of the same are presented in Figure 6‑1 to Figure 6‑3 below.

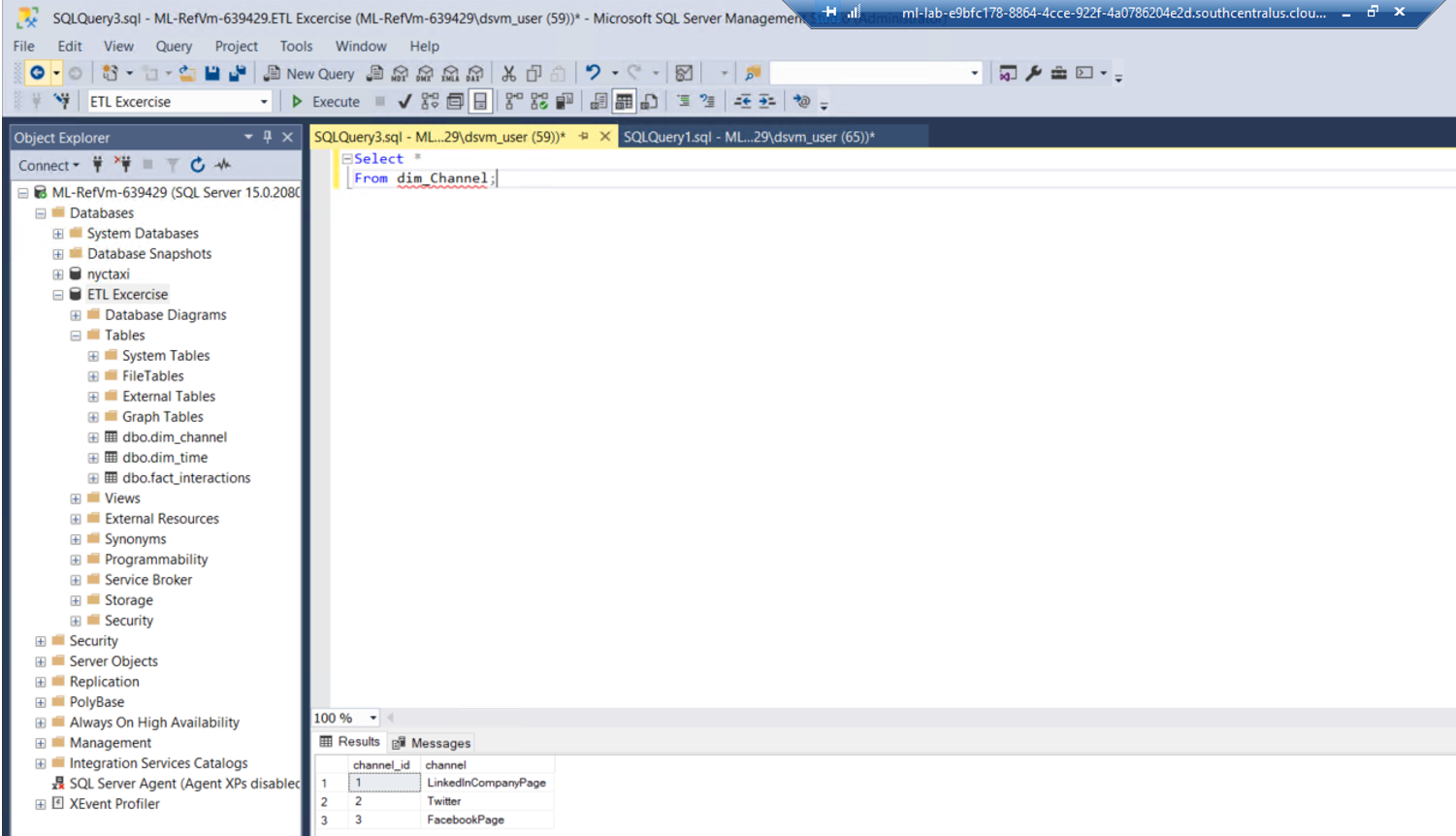


Figure 6‑1: Dimension Table – Row Count – dim\_channel(3)

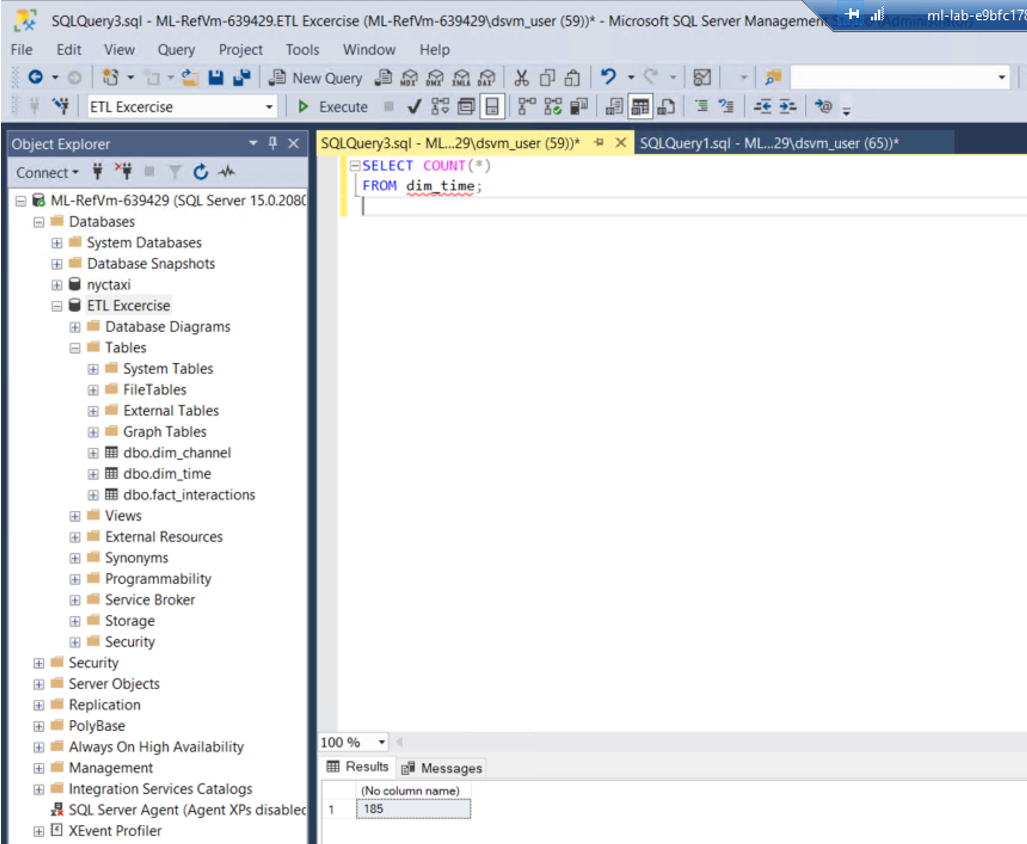


Figure 6‑2: Dimension Table – Row Count – dim\_time(185)

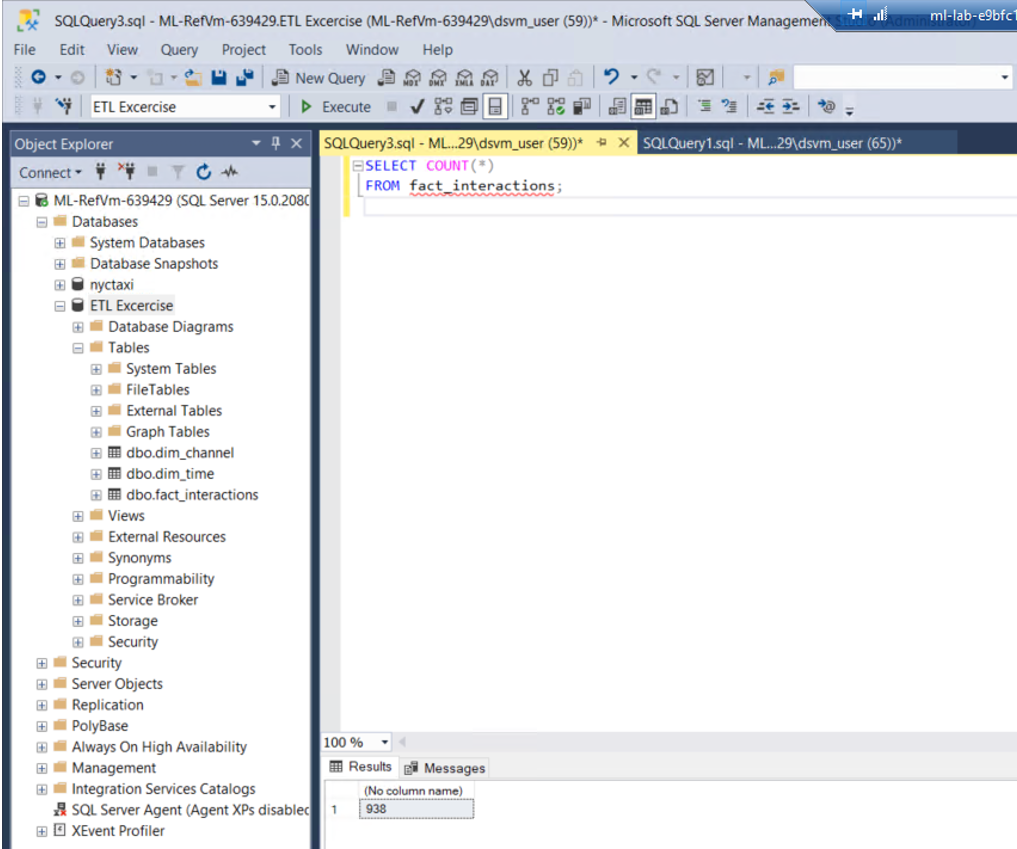


Figure 6‑3: Fact Table – Row Count – fact\_interaction (938)

# PROVIDE PARTIAL SCREENSHOTS OF YOUR CREATED AND POPULATED TABLES (dim\_time, dim\_channel, fact\_interactions)

The screenshots for dim\_channel, dim\_time, and facts\_interactions are presented in Figure 7‑1 to Figure 7‑3 below.

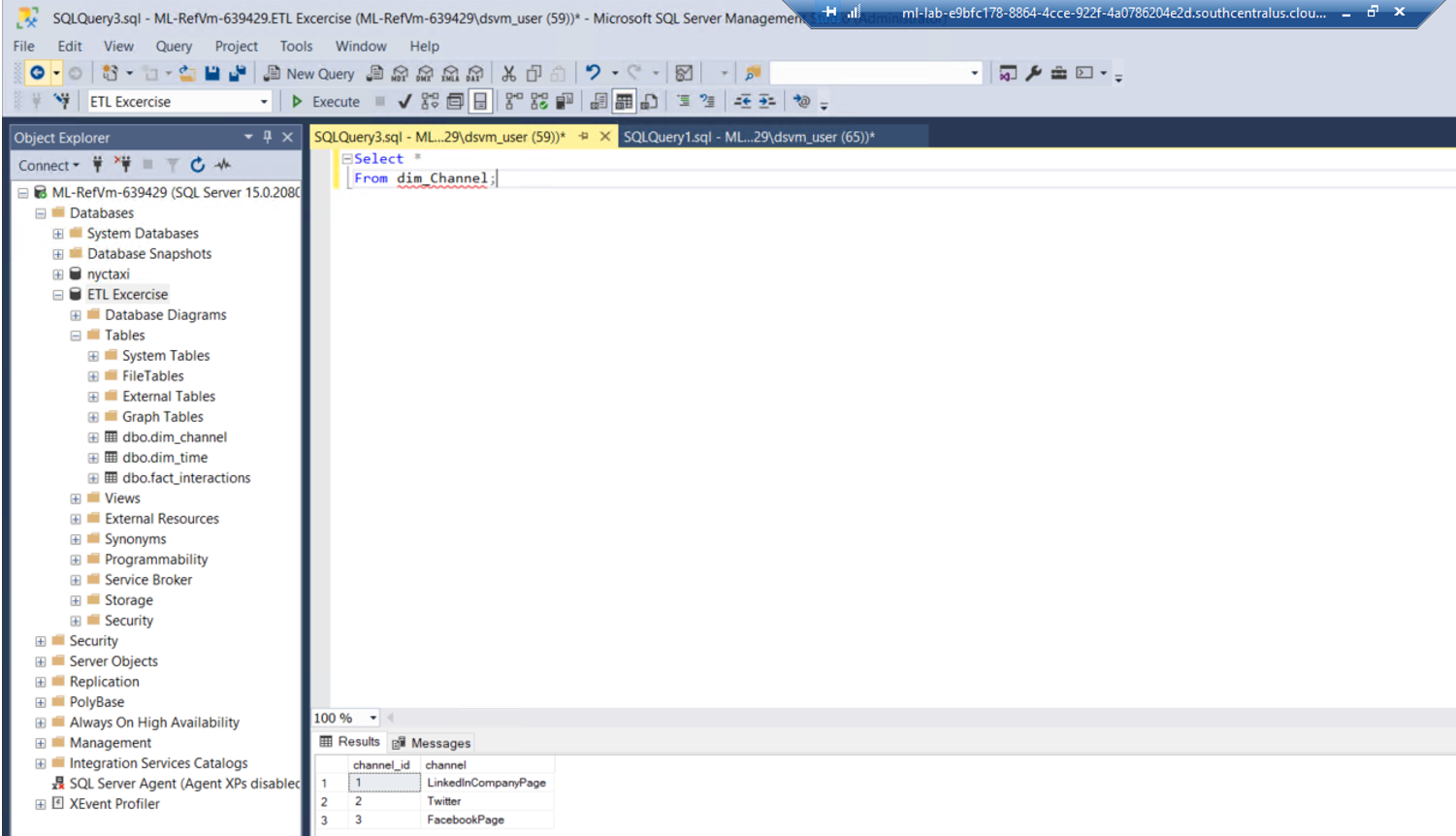


Figure 7‑1: Dimension Table Screenshot – dim\_channel

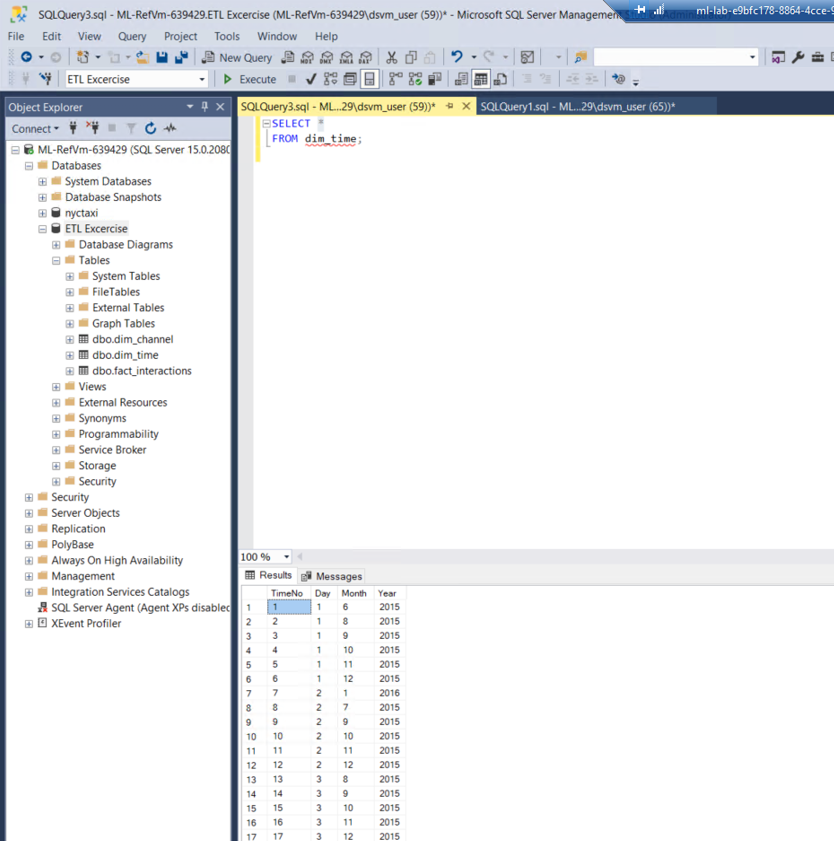


Figure 7‑2: Dimension Table Screenshot – dim\_time

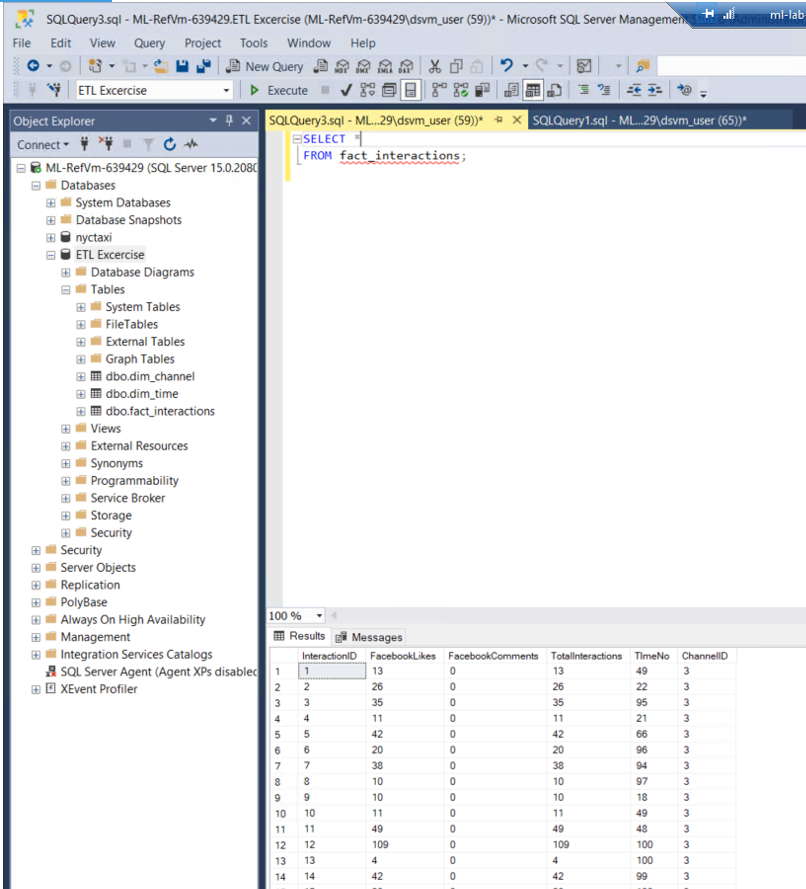


Figure 7‑3: Fact Table Screenshot – fact\_interactions