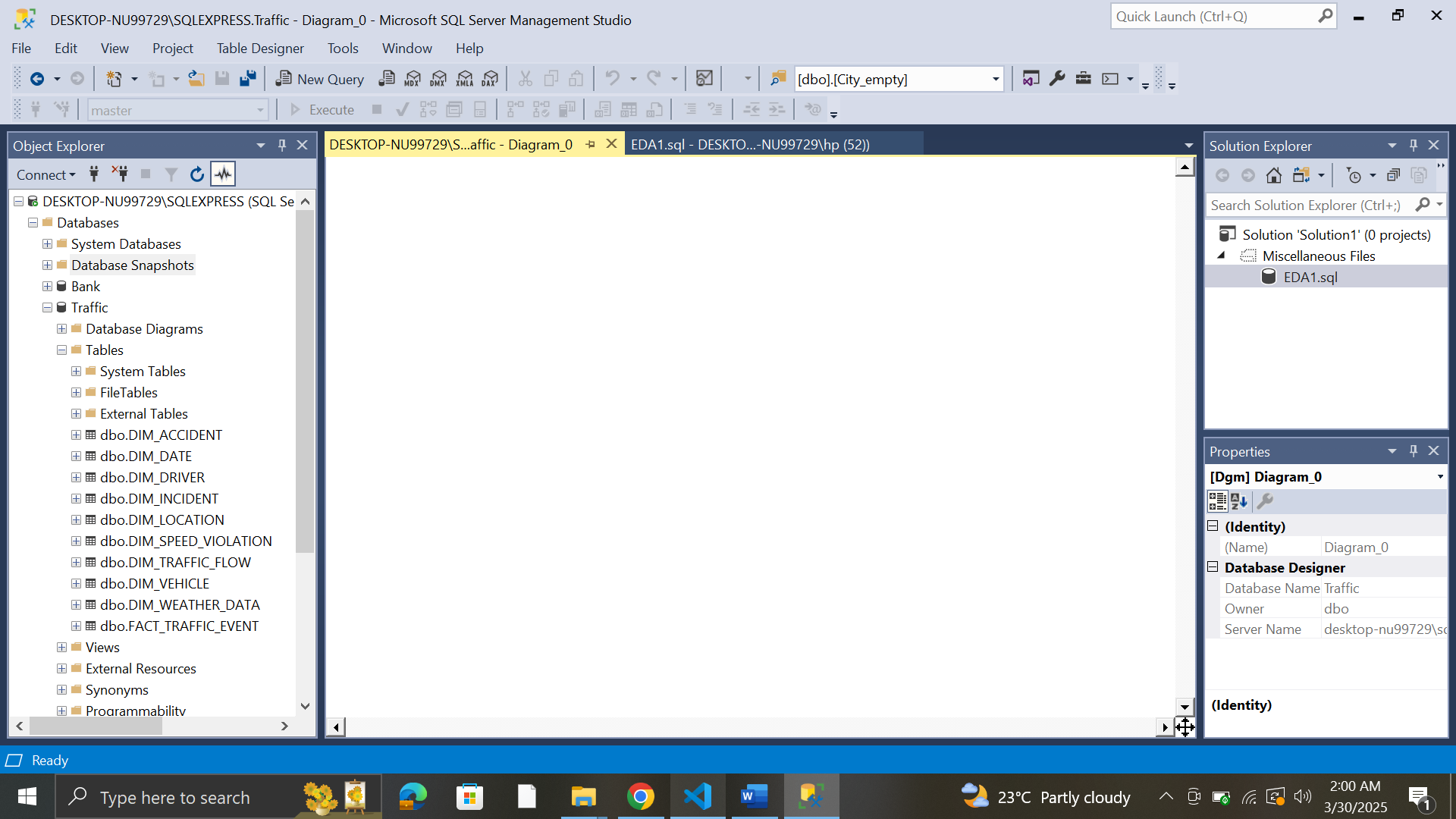
# ETL IMPLEMENTATION USING PYTHON AND SQL SERVER MANAGEMENT STUDIO

A computer screen with a white box

AI-generated content may be incorrect.



AFTER ETL IMPLEMENTATION

BEFORE ETL IMPLEMENTATION

A screenshot of a computer

AI-generated content may be incorrect.STAR SCHEMA IMAGE (SSMS)

# LIST OF SCREEN SHOTS FOR DIMENSIONS AND FACT TABLE

1. DIM VEHICLE

A screenshot of a table

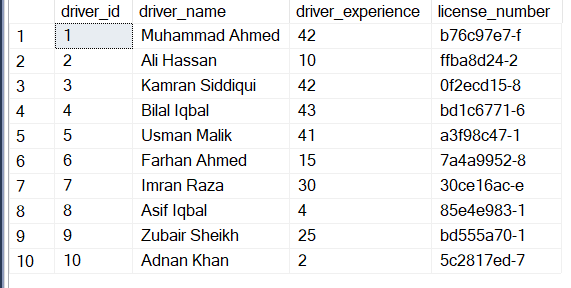
AI-generated content may be incorrect.

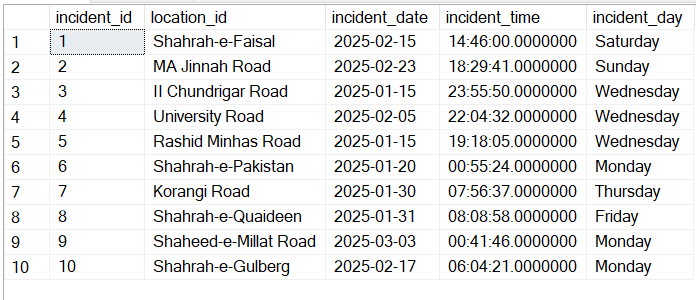
1. DIM SPEEDVIOLATION

A table with numbers and text

AI-generated content may be incorrect.

1. DIM DRIVER



1. DIM INCIDENT
2. A screenshot of a computer

   AI-generated content may be incorrect.DIM LOCATION
3. DIM ACCIDENT

A computer screen with a white screen

AI-generated content may be incorrect.

1. A computer screen with a white screen

   AI-generated content may be incorrect.DIM DATE
2. DIM WEATHER DATA

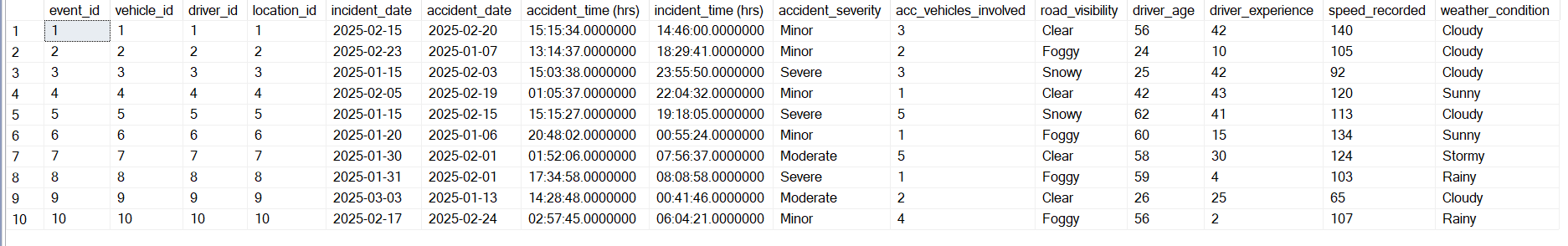
A screenshot of a data

AI-generated content may be incorrect.

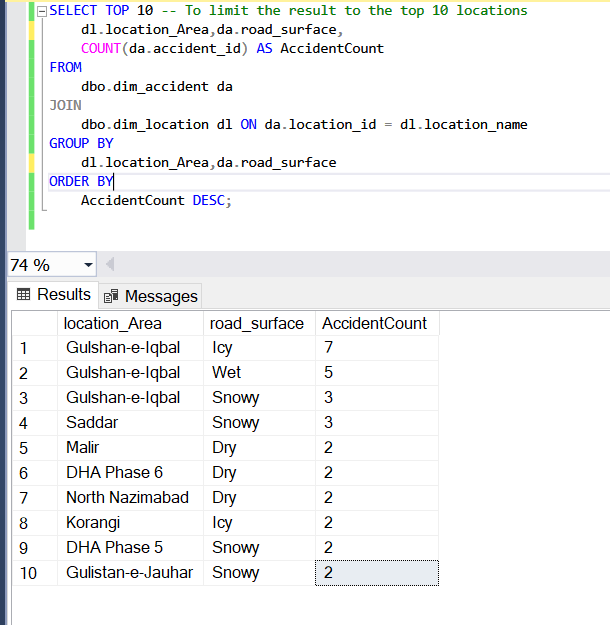
1. DIM TRAFFIC FLOW

A screenshot of a data

AI-generated content may be incorrect.

1.  FACT TRAFFIC EVENT

## A computer screen with text boxes AI-generated content may be incorrect.SQL QUERY 1. Identify the top 10 location areas with the highest traffic flow between February and March 2025, and analyze the correlation with road conditions (e.g., wet, icy, snowy) and number of accidents.



## A screenshot of a computer AI-generated content may be incorrect.SQL QUERY 2. What are the average speed violations recorded for each vehicle type (Car, Bus, Truck, Motorcycle) during peak hours (7-9 AM and 4-6 PM) in February 2025.

## SQL QUERY 3. What is the incident frequency at locations (e.g., Shahrah-e-Faisal, University Road) from January 2025 onwards.

A screenshot of a computer

AI-generated content may be incorrect.

## SQL QUERY 4. Calculate the percentage of inactive cameras at locations with the highest traffic flow (Top 5) and assess its impact in February 2025.

WITH TopLocations AS (

-- 1. Find Top 5 Location Areas by Traffic Flow in February 2025

SELECT TOP 10

dl.location\_Area,

SUM(tf.vehicle\_count) AS TotalTraffic

FROM

dbo.dim\_traffic\_flow tf

JOIN

dbo.dim\_location dl ON tf.location\_id = dl.location\_id

WHERE

tf.traffic\_flow\_date >= '2025-02-01' AND tf.traffic\_flow\_date < '2025-03-01'

GROUP BY

dl.location\_Area

ORDER BY

TotalTraffic DESC

),

CameraStatus AS (

-- 2. Get Camera Status for each Location Area

SELECT

tl.location\_Area,

CASE

WHEN dl.camera\_status = 'Inactive' THEN 1

ELSE 0

END AS IsInactive

FROM

TopLocations tl

JOIN

dbo.dim\_location dl ON tl.location\_Area = dl.location\_Area

)

-- 3. Calculate Percentage of Inactive Cameras by Location Area

SELECT

location\_Area,

(SUM(IsInactive \* 1.0) / COUNT(\*)) \* 100 AS PercentageInactive

FROM

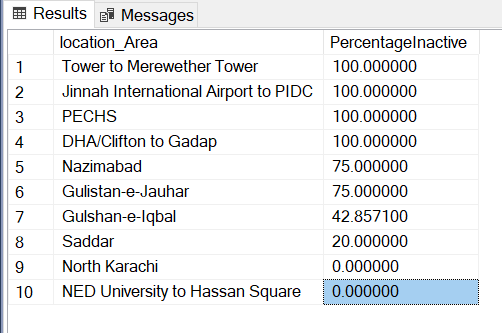
CameraStatus

GROUP BY

location\_Area

ORDER BY

PercentageInactive DESC;



## SQL QUERY 5. Identify drivers aged under 25 or younger with the highest number of speed violations, and analyze their associated vehicle types (e.g., Motorcycle, Car) to prioritize targeted safety intervention.

