Real Time Traffic Project

Under Super Vision: DR. Amira Yousef

Ahmed Salama Soliman Hasan

Mohamed Nasr

Osama Ahmed Reda Hegazy

Sara Hisham Ahmed Mohamed

Sherif Gamal Kamed AbdElwahed

Yousef Ahmed Mohamed Ibrahiem

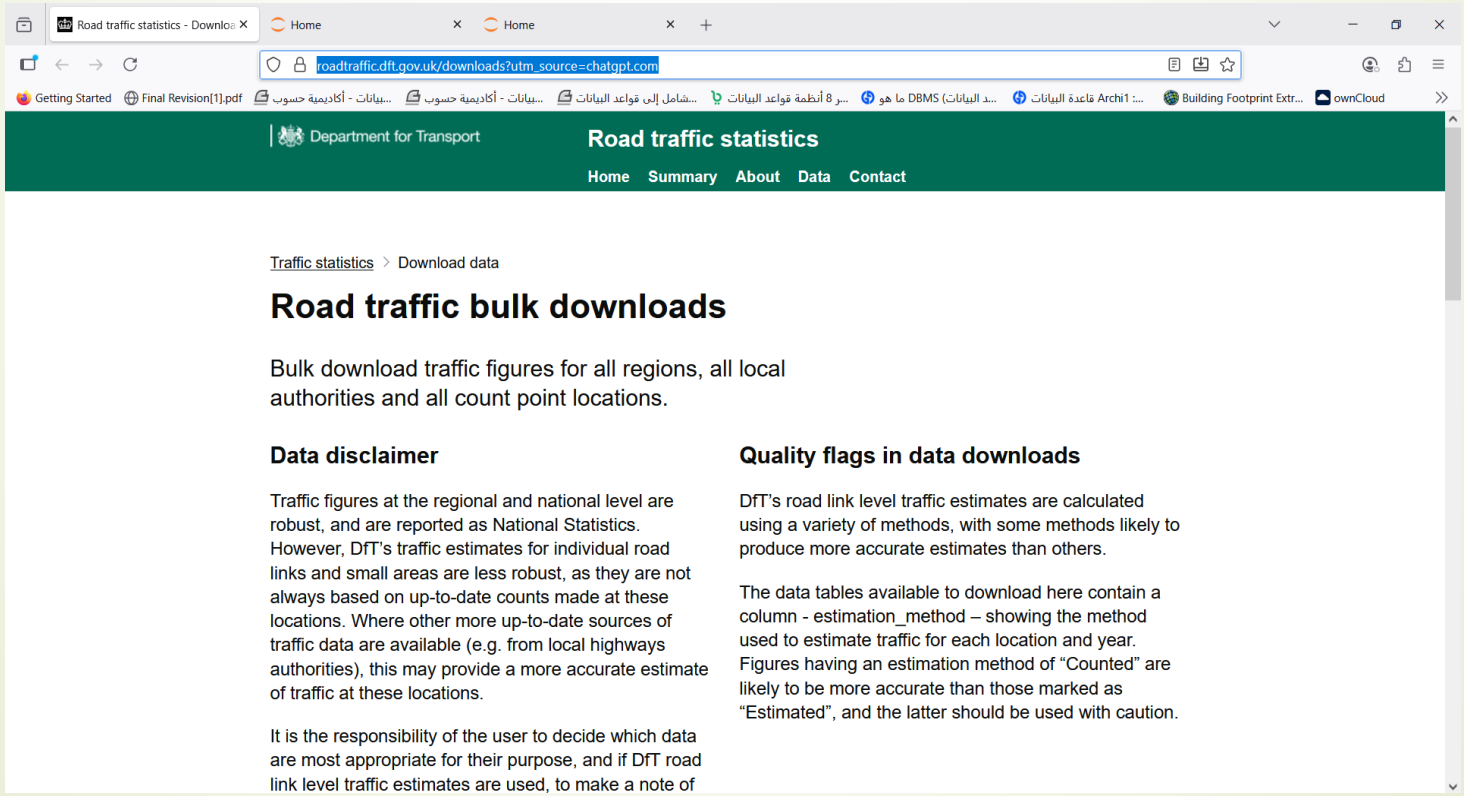
Zakaria Yehia Ahmed

The beneficiaries of the "Real Time Traffic Project" include:

1. **Commuters and Drivers** – who can access up-to-date traffic information, avoid congestion, and reduce travel time.
2. **Public Transportation Authorities** – who can optimize routes, schedules, and fleet management using real-time data.
3. **Emergency Services** – such as police, ambulance, and fire departments, enabling faster response times and route planning.
4. **Urban Planners and Decision Makers** – who can use traffic data to improve infrastructure planning and traffic flow management.
5. **Logistics and Delivery Companies** – who can enhance operational efficiency and reduce fuel consumption through route optimization.
6. **Environmental Agencies** – benefiting from reduced vehicle emissions due to improved traffic flow.
7. **Citizens and Local Communities** – who experience safer roads, less congestion, and better air quality.

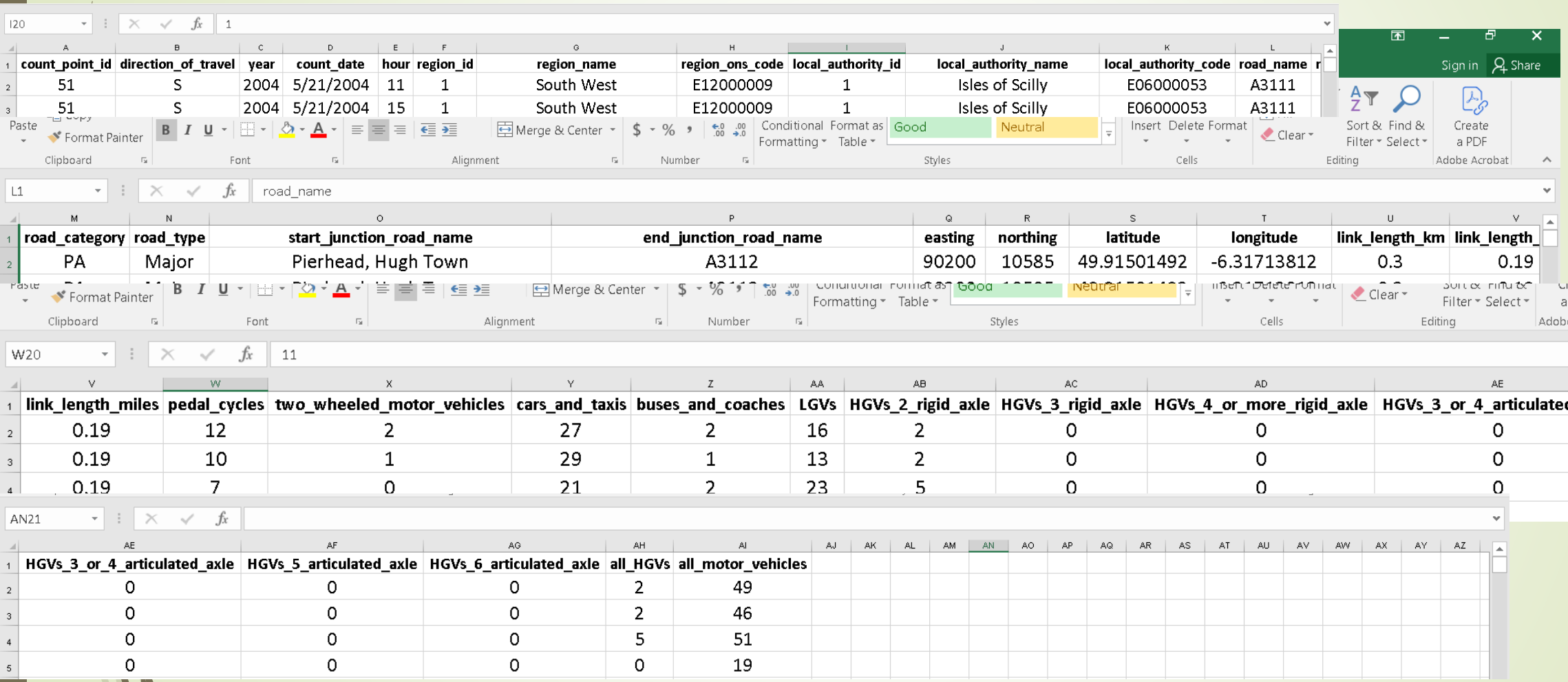
Data Source

https://roadtraffic.dft.gov.uk/downloads?utm_source=chatgpt.com



Data Before

(worked on 46476 out of 10,485,76records)



count_point_id	direction_of_travel	year	count_date	hour	region_id	region_name	region_ons_code	local_authority_id	local_authority_name	local_authority_code	road_name
51	S	2004	5/21/2004	11	1	South West	E12000009	1	Isles of Scilly	E06000053	A3111
51	S	2004	5/21/2004	15	1	South West	E12000009	1	Isles of Scilly	E06000053	A3111

road_category	road_type	start_junction_road_name	end_junction_road_name	easting	northing	latitude	longitude	link_length_km	link_length_
PA	Major	Pierhead, Hugh Town	A3112	90200	10585	49.91501492	-6.31713812	0.3	0.19

link_length_miles	pedal_cycles	two_wheeled_motor_vehicles	cars_and_taxis	buses_and_coaches	LGVs	HGVs_2_rigid_axle	HGVs_3_rigid_axle	HGVs_4_or_more_rigid_axle	HGVs_3_or_4_articulated_axle	HG
0.19	12	2	27	2	16	2	0	0	0	
0.19	10	1	29	1	13	2	0	0	0	
0.19	7	0	21	2	23	5	0	0	0	

HGVs_3_or_4_articulated_axle	HGVs_5_articulated_axle	HGVs_6_articulated_axle	all_HGVs	all_motor_vehicles
0	0	0	2	49
0	0	0	2	46
0	0	0	5	51
0	0	0	0	19

شرح معاني الحقول:(Columns)

اسم الحقل	المعنى
count_point_id	رقم تعريف نقطة العدّ (مكان تم فيه عدّ المركبات).
direction_of_travel	اتجاه السفر (زي شمال، جنوب، شرق، غرب).
year	سنة تسجيل البيانات.
count_date	التاريخ الفعلي للعدّ (يوم العدّ).
hour	الساعة (الوقت اللي تم فيه العدّ – غالبًا من 0 إلى 23).
region_id	رقم تعريف المنطقة (رقم داخلي يستخدم لتحديد المنطقة).
region_name	اسم المنطقة) زي "East Midlands" أو."London"
region_ons_code	كود المنطقة حسب مكتب الإحصاء الوطني) ONS للاستخدام الإحصائي.(
local_authority_id	رقم تعريف السلطة المحلية (بلدية أو مجلس محلي).
local_authority_name	اسم السلطة المحلية) زي."Leeds City Council"
local_authority_code	كود السلطة المحلية (رمز مختصر).
road_name	اسم الطريق) زي A1، M25).
road_category	فئة الطريق (زي طريق سريع، طريق رئيسي، محلي، الخ).
road_type	نوع الطريق (قد يكون مزيد من التفصيل عن الفئة – مفرد، مزدوج، الخ).
start_junction_road_name	اسم الطريق عند بداية المفرق أو التقاطع.
end_junction_road_name	اسم الطريق عند نهاية المفرق أو التقاطع.
easting	الإحداثي الشرقي (نظام الإحداثيات البريطانية).
northing	الإحداثي الشمالي (نظام الإحداثيات البريطانية).
latitude	دائرة العرض (إحداثيات جغرافية).
longitude	خط الطول (إحداثيات جغرافية).
link_length_km	طول المقطع الطرقي بالكيلومترات.
link_length_miles	طول المقطع الطرقي بالأميال.
pedal_cycles	عدد الدراجات الهوائية.
two_wheeled_motor_vehicles	عدد الدراجات النارية (سكوتر، موتوسيكل).
cars_and_taxis	عدد السيارات العادية والتاكسي.
buses_and_coaches	عدد الحافلات.
LGVs	سيارات النقل الخفيف – (Light Goods Vehicles) زي الفانات الصغيرة.
HGVs_2_rigid_axle	شاحنات ثقيلة بعدد 2 محور صلب.
HGVs_3_rigid_axle	شاحنات ثقيلة بعدد 3 محاور صلبة.
HGVs_4_or_more_rigid_axle	شاحنات ثقيلة بعدد 4 أو أكثر من المحاور الصلبة.
HGVs_3_or_4_articulated_axle	شاحنات ثقيلة مفصلية بعدد 3 أو 4 محاور.
HGVs_5_articulated_axle	شاحنات ثقيلة مفصلية بعدد 5 محاور.
HGVs_6_articulated_axle	شاحنات ثقيلة مفصلية بعدد 6 محاور أو أكثر.
all_HGVs	إجمالي الشاحنات الثقيلة.(HGVs)
all_motor_vehicles	إجمالي جميع المركبات الآلية (من موتوسيكلات لحد الشاحنات).

Software Used

- Jupyter , Python
- SQL Server
- Jithub

Data Preparation and Geospatial Processing Workflow

- Reading and cleaning data
- Transforming time columns
- Checking for outliers
- Exporting data to an SQL database
- Creating a GeoDataFrame and exporting data to GeoJSON and Shapefile
- Data quality report

1. Initial Setup and Data Loading
"import pandas as pd - import numpy as np - import geopandas as gpd)

Load the CSV data

----

2. Data Exploration and Cleaning

----

3. Data Validation and Quality Checks

----

4. Data Transformation and Feature Engineering

==========

5. Save to Parquet (optional)

----

7. Aggregate Data for Reporting

----

8. Save GeoJSON (Optional for Spatial Analysis)

----

9. Data Quality Report

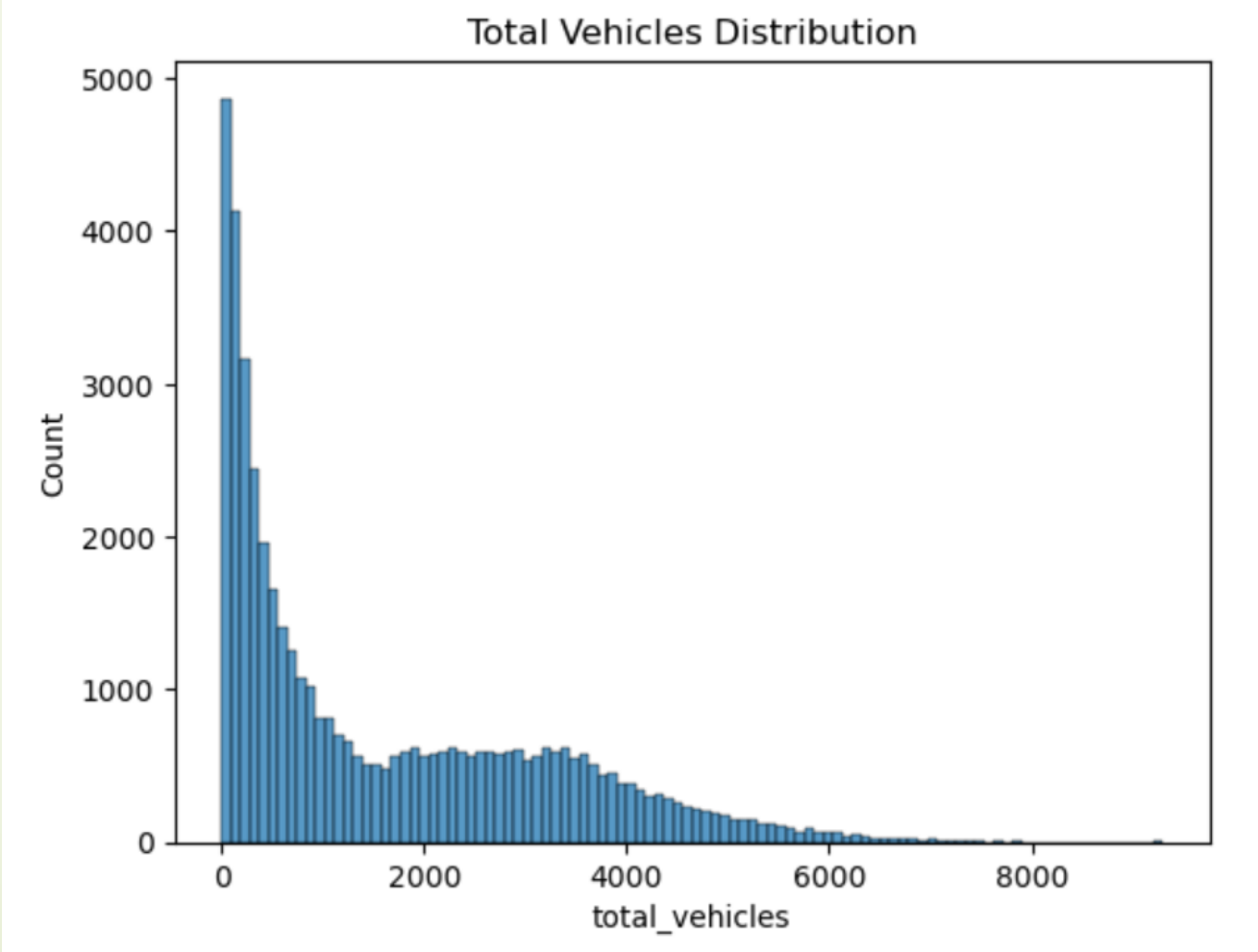
GIS Map

قيم غير منطقية(Logical Errors)

العمود	التحقق
hour	هل القيم بين 0 و 23 فقط؟
total_vehicles	هل في قيم سالبة أو أكبر من مليون؟
link_length_km	هل يوجد قيم سالبة؟
vehicle_density	هل القيمة منطقية؟ مثلًا مش 100,000 مركبة/كم
bicycle_percentage وأhgv_percentage	هل أكثر من 1 أو أقل من 0؟ (نسب غير منطقية)

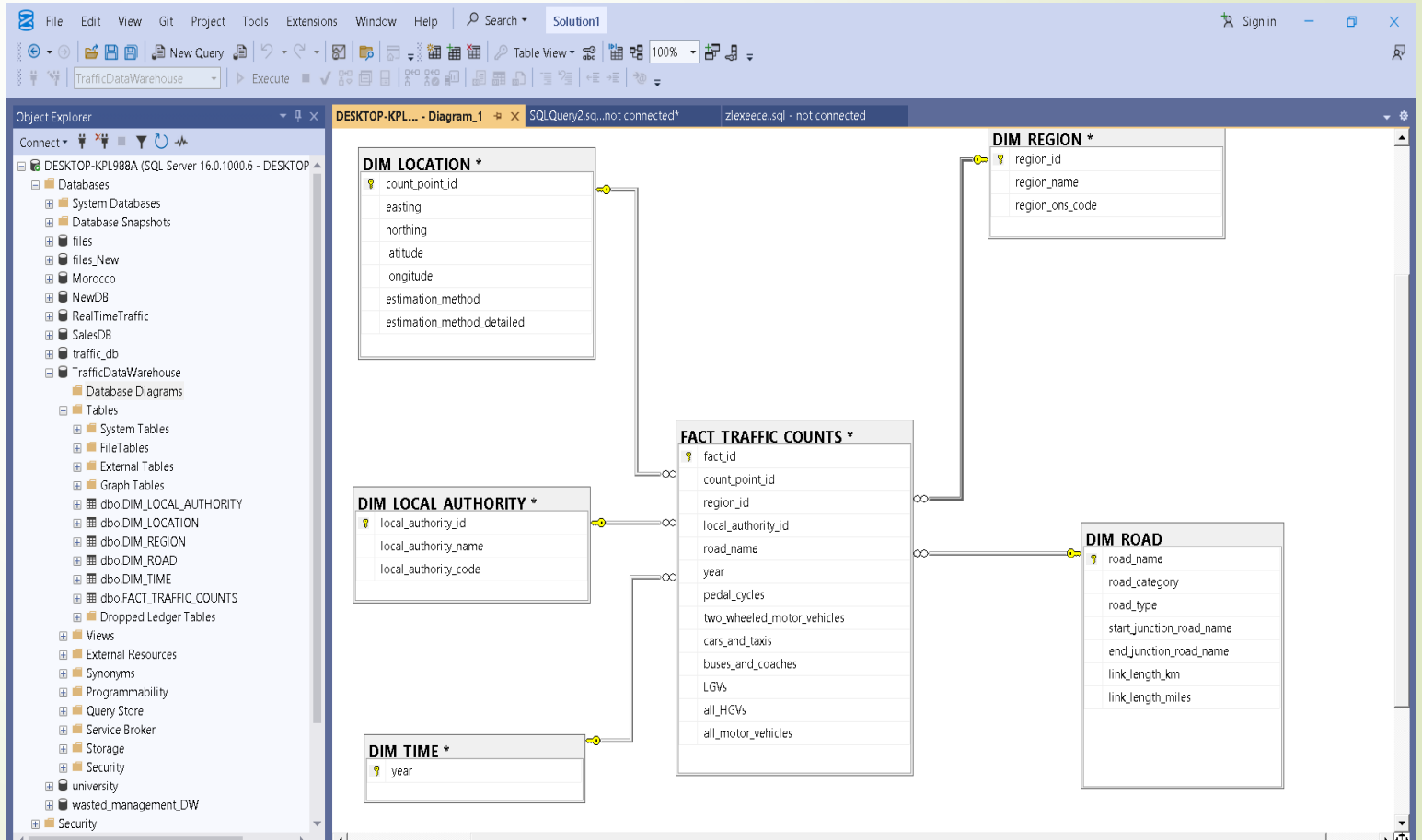
- Illogical Dates
- Have future dates?
- Some records have missing dates (NaT)?
- Geographic points outside the UK?
- If the data is from the UK, such as DFT
- Make sure the latitude and longitude are within the UK Approximately

Data Distribution If the data has a strong skew or many outliers, it can affect statistical analysis or modeling



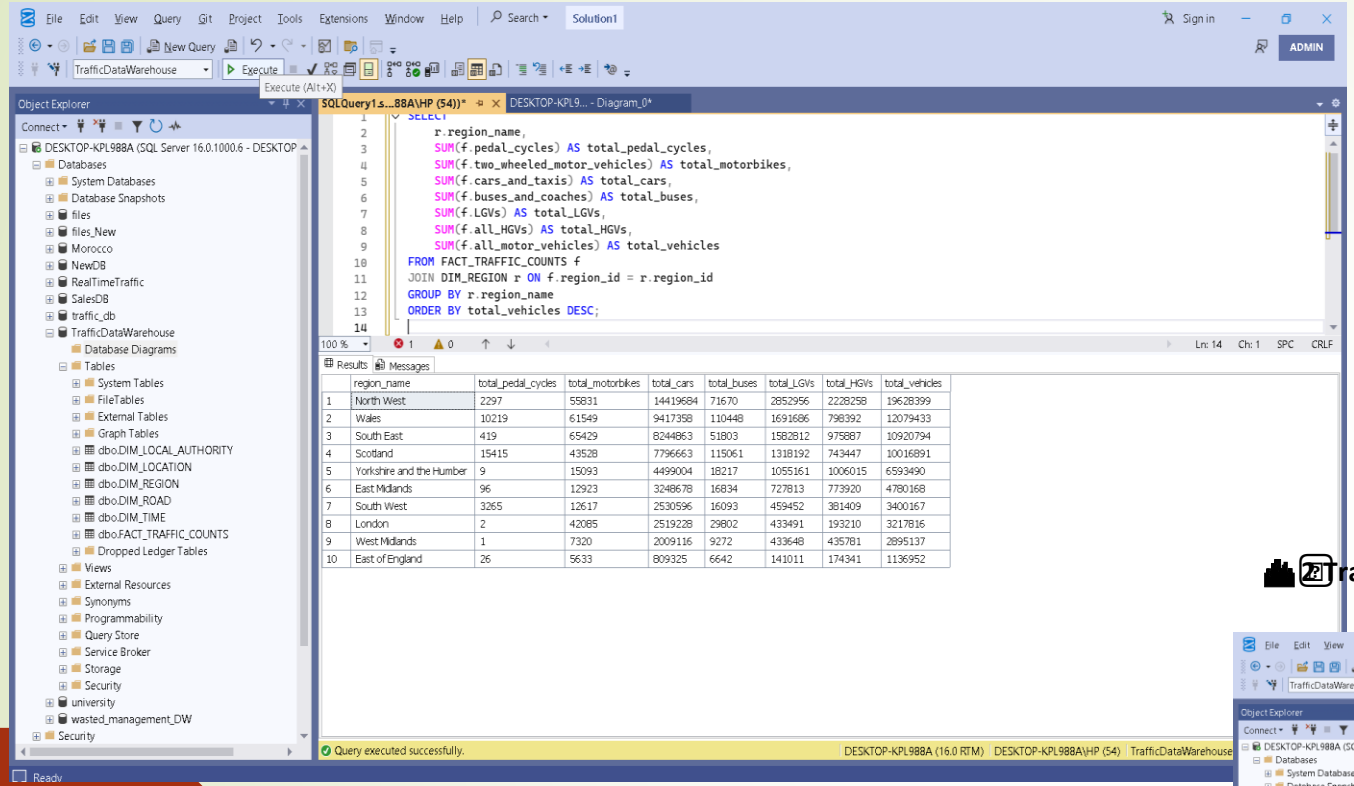
Data Warehouse

Data After Restructure (Star Schema)



Data Analysis:

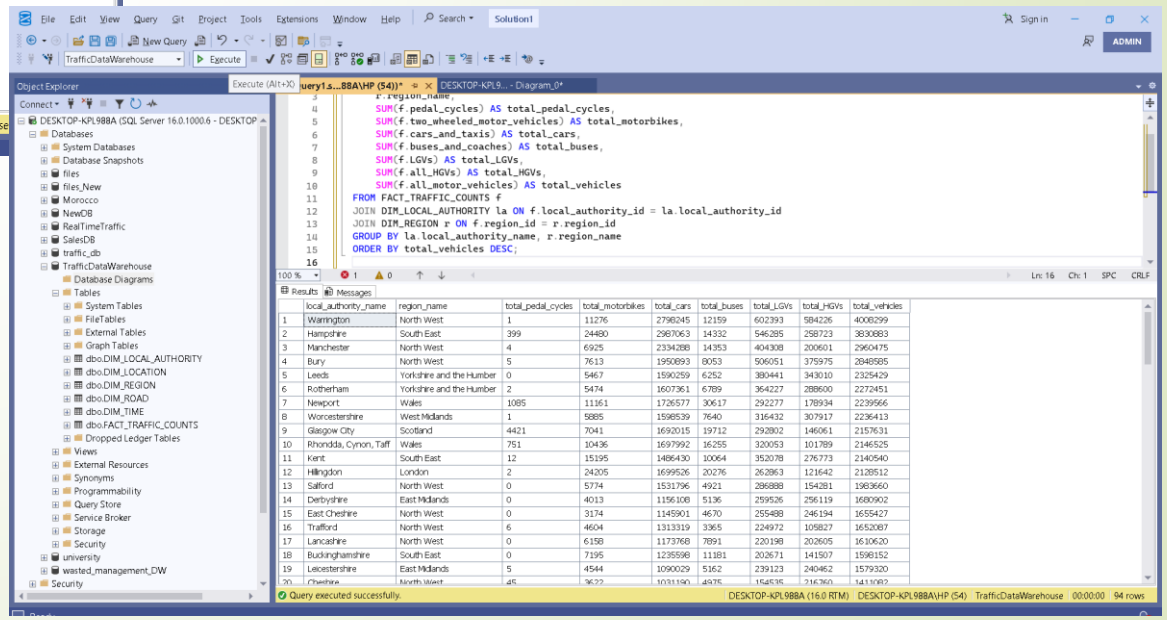
1️⃣ Traffic by Region



```
SELECT
    r.region_name,
    SUM(f.pedal_cycles) AS total_pedal_cycles,
    SUM(f.two_wheeled_motor_vehicles) AS total_motorbikes,
    SUM(f.cars_and_taxis) AS total_cars,
    SUM(f.buses_and_coaches) AS total_buses,
    SUM(f.LGVs) AS total_LGVs,
    SUM(f.all_HGVs) AS total_HGVs,
    SUM(f.all_motor_vehicles) AS total_vehicles
FROM FACT_TRAFFIC_COUNTS f
JOIN DIM_REGION r ON f.region_id = r.region_id
GROUP BY r.region_name
ORDER BY total_vehicles DESC;
```

	region_name	total_pedal_cycles	total_motorbikes	total_cars	total_buses	total_LGVs	total_HGVs	total_vehicles
1	North West	2297	55831	14419684	71670	2852956	2228258	19628399
2	Wales	10219	61549	9417358	110448	1691686	798392	12079433
3	South East	419	65429	8244863	51803	1582812	975887	10920794
4	Scotland	15415	43528	7796663	115061	1318192	743447	10016891
5	Yorkshire and the Humber	9	15093	4499004	18217	1055161	1006015	6593490
6	East Midlands	96	12923	3248678	16834	727813	773920	4780168
7	South West	3265	12617	2530596	16093	459452	381409	3400167
8	London	2	42085	2519228	29802	433491	193210	3217816
9	West Midlands	1	7320	2009116	9272	433648	435781	2895137
10	East of England	26	5633	809325	6642	141011	174341	1136952

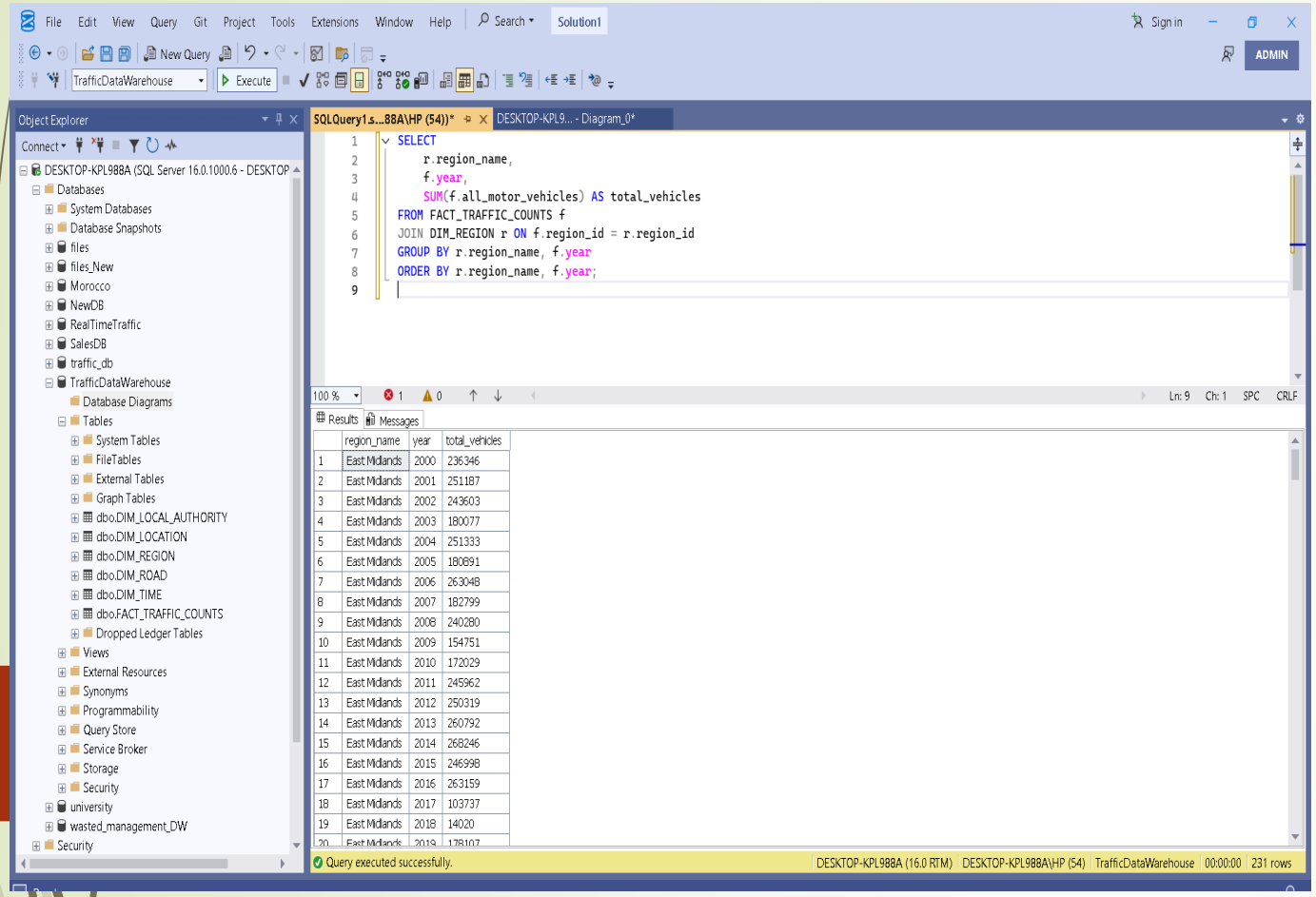
2️⃣ Traffic by Local Authority



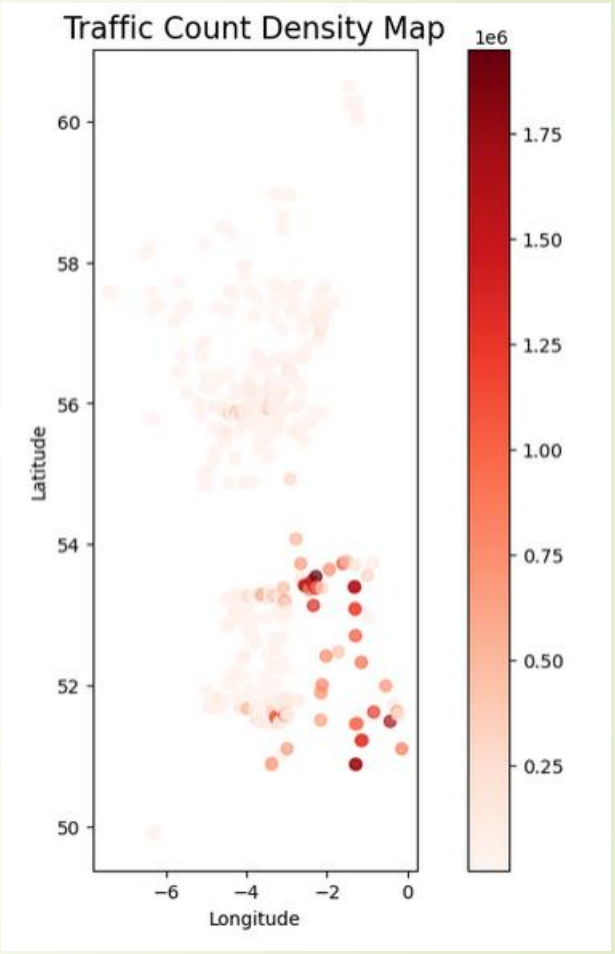
```
SELECT
    la.local_authority_name,
    r.region_name,
    SUM(f.pedal_cycles) AS total_pedal_cycles,
    SUM(f.two_wheeled_motor_vehicles) AS total_motorbikes,
    SUM(f.cars_and_taxis) AS total_cars,
    SUM(f.buses_and_coaches) AS total_buses,
    SUM(f.LGVs) AS total_LGVs,
    SUM(f.all_HGVs) AS total_HGVs,
    SUM(f.all_motor_vehicles) AS total_vehicles
FROM FACT_TRAFFIC_COUNTS f
JOIN DIM_LOCAL_AUTHORITY la ON f.local_authority_id = la.local_authority_id
JOIN DIM_REGION r ON f.region_id = r.region_id
GROUP BY la.local_authority_name, r.region_name
ORDER BY total_vehicles DESC;
```

	local_authority_name	region_name	total_pedal_cycles	total_motorbikes	total_cars	total_buses	total_LGVs	total_HGVs	total_vehicles
1	Warrington	North West	1	11276	2798245	12159	602393	584226	4008299
2	Hampshire	South East	399	24480	2987063	14332	546285	258723	3830883
3	Manchester	North West	4	6925	2334288	14353	404308	200601	2960475
4	Bury	North West	5	7613	1950893	8053	506051	375975	2848585
5	Leeds	Yorkshire and the Humber	0	5467	1590259	6252	380441	343010	2325429
6	Rotherham	Yorkshire and the Humber	2	5474	1607361	6789	364227	288600	2272451
7	Newport	Wales	1085	11161	1726577	30617	292277	178934	2239566
8	Worcestershire	West Midlands	1	5885	1598539	7640	316432	307917	2236413
9	Glasgow City	Scotland	4421	7041	1692015	19712	292802	146061	2157631
10	Rhondda, Cynon, Taff	Wales	751	10436	1697992	16255	320053	101789	2146525
11	Kent	South East	12	15195	1486430	10064	352078	276773	2140540
12	Hillingdon	London	2	24205	1699526	20276	262863	121642	2128512
13	Salford	North West	0	5774	1531796	4921	286888	154281	1983660
14	Derbyshire	East Midlands	0	4013	1156108	5136	259526	256119	1680902
15	East Cheshire	North West	0	3174	1145901	4670	255488	246194	1655427
16	Trafford	North West	6	4604	1313319	3365	224972	105827	1652087
17	Lancashire	North West	0	6158	1173768	7891	220198	202605	1610620
18	Buckinghamshire	South East	0	7195	1235598	11181	202671	141507	1598152
19	Leicestershire	East Midlands	5	4544	1090029	5162	239123	240462	1579320

📈 Traffic Trend by Region (Over Time)



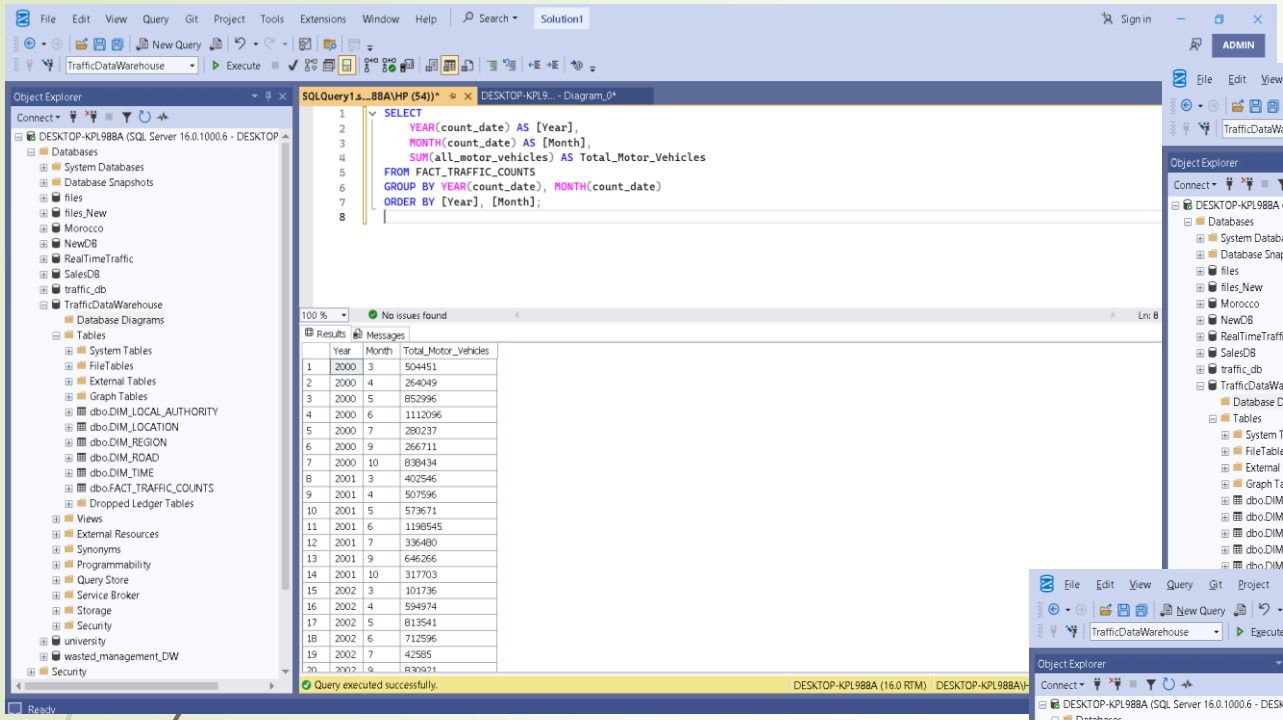
Plot:



Data Analysis – Date time:🎯 Objective of the Analysis Traffic data analysis (Traffic Counts) by time, to understand temporal patterns such as:

- Distribution of the number of vehicles by year / month / day / hour
- Identifying Peak Hours
- Calculating the total and average number of vehicles at different times
- Understanding temporal trends (daily, weekly, yearly)

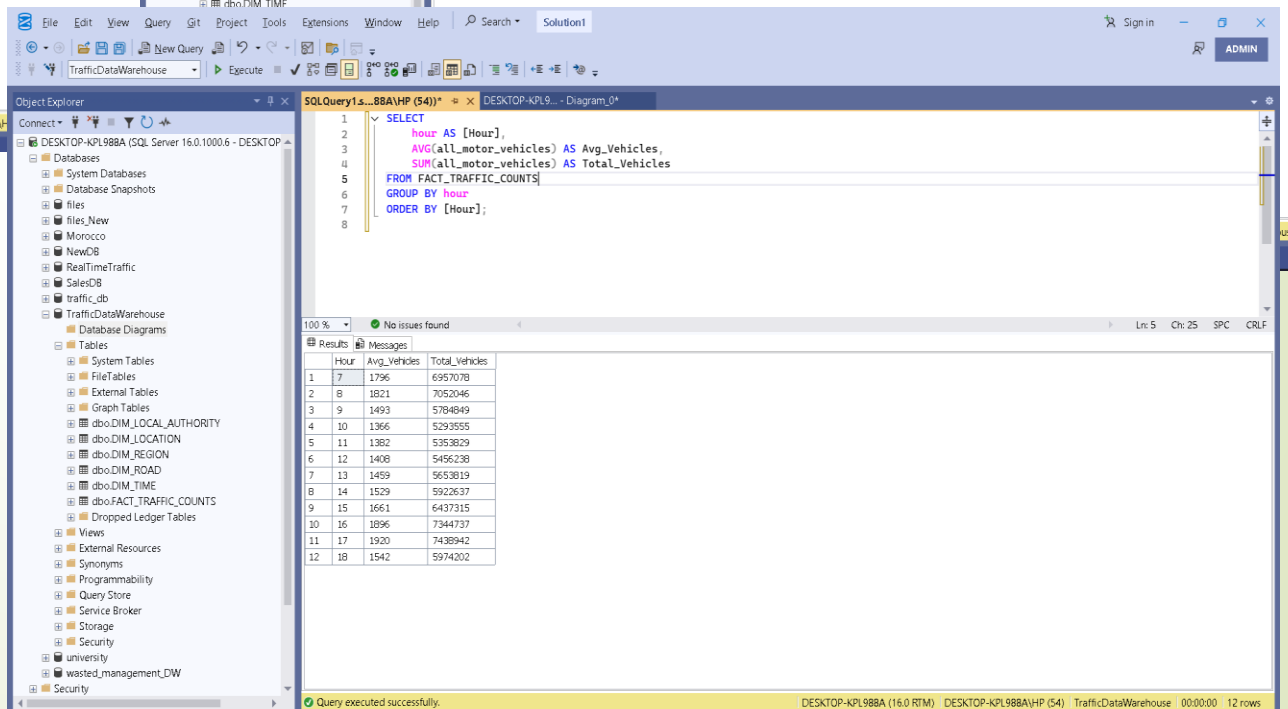
2. Analysis of total vehicles by year and month



```
SELECT
    YEAR(count_date) AS [Year],
    MONTH(count_date) AS [Month],
    SUM(all_motor_vehicles) AS Total_Motor_Vehicles
FROM FACT_TRAFFIC_COUNTS
GROUP BY YEAR(count_date), MONTH(count_date)
ORDER BY [Year], [Month];
```

	Year	Month	Total_Motor_Vehicles
1	2000	3	504451
2	2000	4	264049
3	2000	5	852996
4	2000	6	1112096
5	2000	7	280237
6	2000	9	266711
7	2000	10	838434
8	2001	3	402546
9	2001	4	507596
10	2001	5	573671
11	2001	6	1198545
12	2001	7	336480
13	2001	9	646266
14	2001	10	317703
15	2002	3	101736
16	2002	4	594974
17	2002	5	813541
18	2002	6	712596
19	2002	7	42585

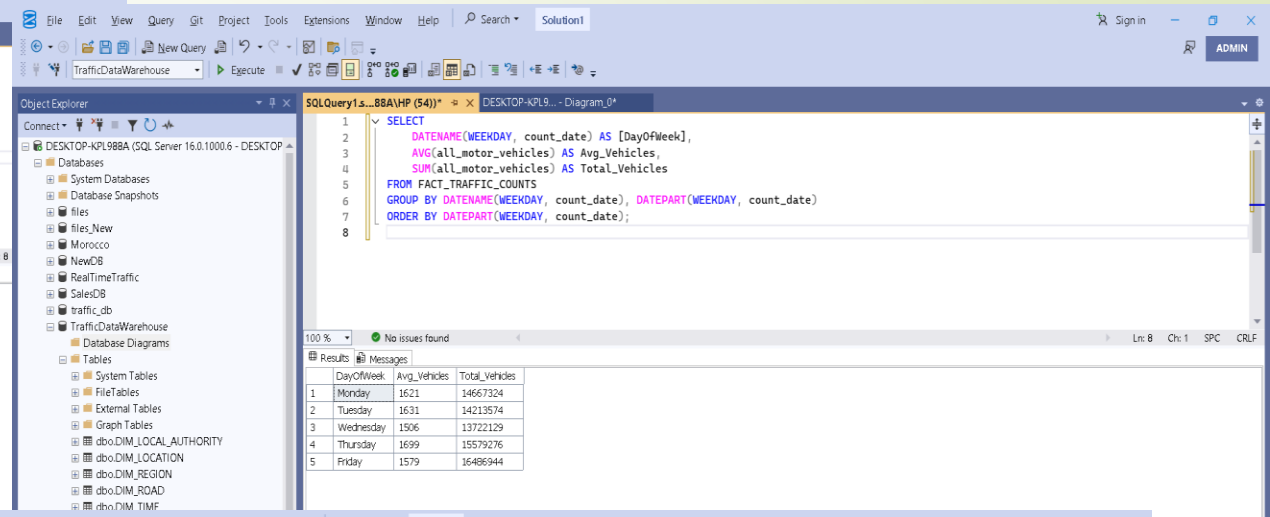
3. Analysis of the number of vehicles by hour - Peak Hour Analysis.



```
SELECT
    hour AS [Hour],
    AVG(all_motor_vehicles) AS Avg_Vehicles,
    SUM(all_motor_vehicles) AS Total_Vehicles
FROM FACT_TRAFFIC_COUNTS
GROUP BY hour
ORDER BY [Hour];
```

	Hour	Avg_Vehicles	Total_Vehicles
1	7	1796	6957078
2	8	1821	7052046
3	9	1493	5784849
4	10	1366	5293555
5	11	1382	5353829
6	12	1408	5456238
7	13	1459	5653819
8	14	1529	5922637
9	15	1661	6437315
10	16	1896	7344737
11	17	1920	7438942
12	18	1542	5974202

4. Analysis by day of the week

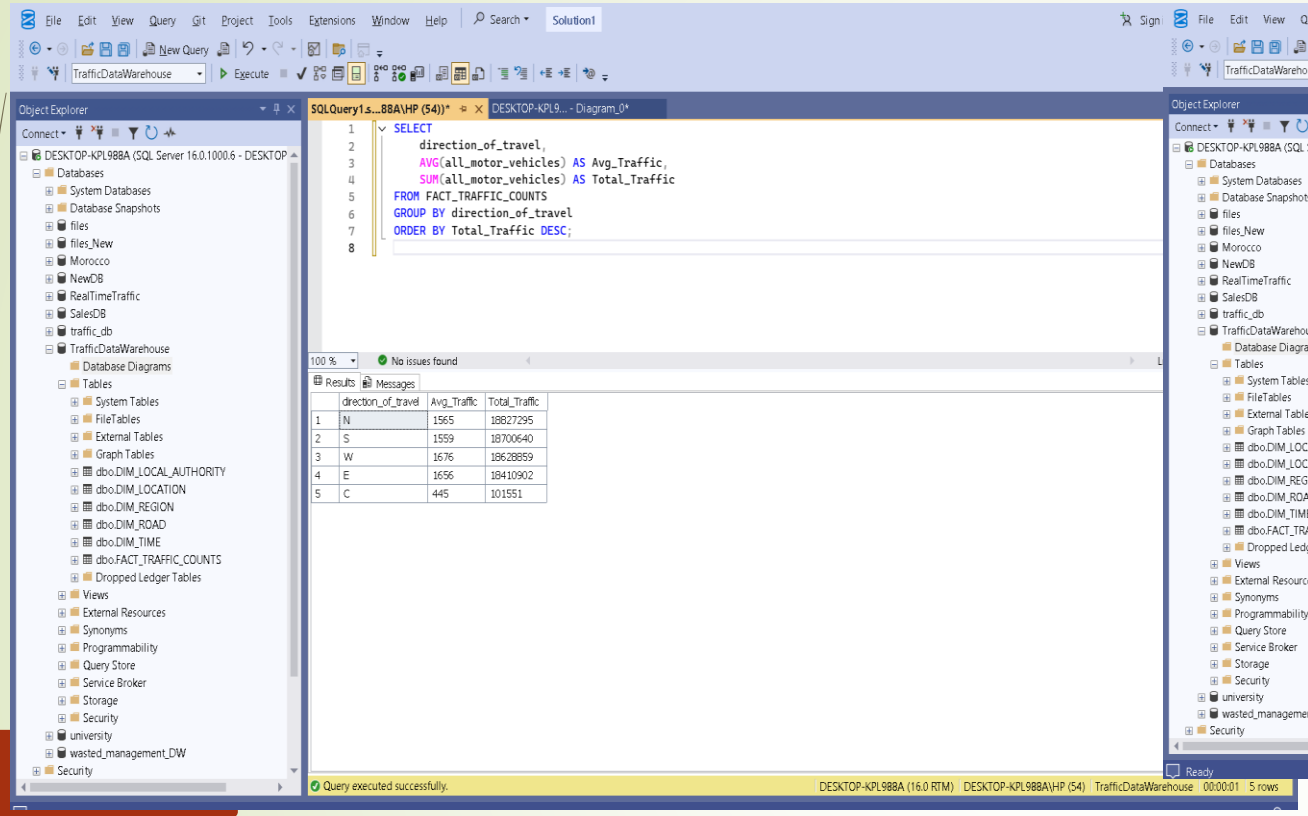


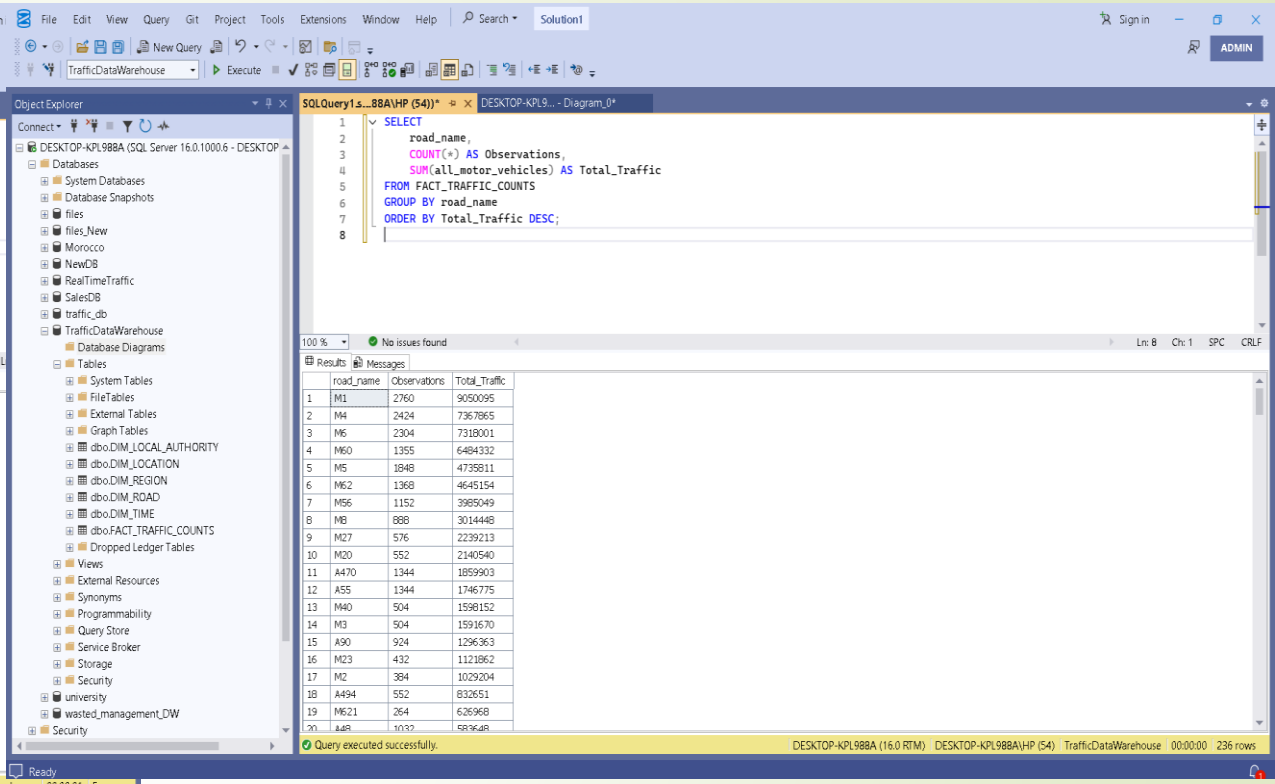
```
SELECT
    DATENAME(WEEKDAY, count_date) AS [DayOfWeek],
    AVG(all_motor_vehicles) AS Avg_Vehicles,
    SUM(all_motor_vehicles) AS Total_Vehicles
FROM FACT_TRAFFIC_COUNTS
GROUP BY DATENAME(WEEKDAY, count_date), DATEPART(WEEKDAY, count_date)
ORDER BY DATEPART(WEEKDAY, count_date);
```

	DayOfWeek	Avg_Vehicles	Total_Vehicles
1	Monday	1621	14667324
2	Tuesday	1631	14213574
3	Wednesday	1506	13722129
4	Thursday	1699	15579276
5	Friday	1579	16486944

5. Direction Analysis (Direction of Travel)

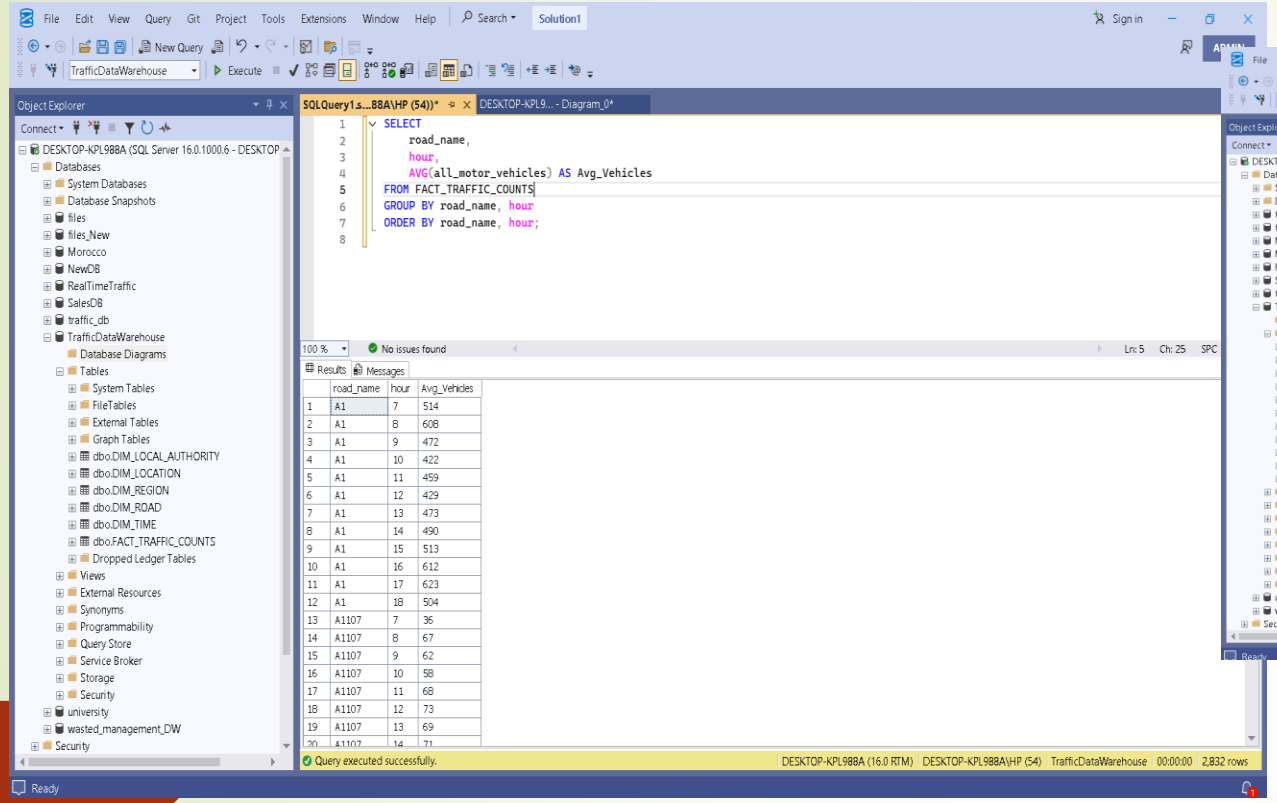
6. Road Analysis Road Name

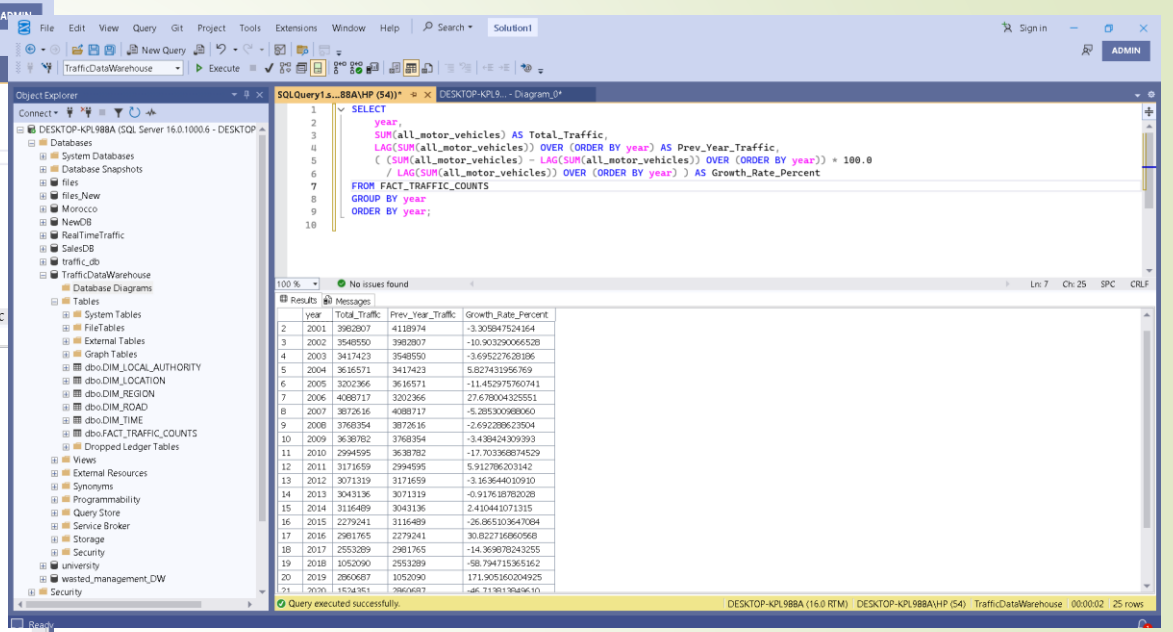




7. Combining temporal and spatial analysisFor example — traffic per road by the hour

8. Analysis of the Annual Growth Rate






Thank
you! :)