**ROAD ACCIDENT DATA ANALYSIS REPORT BASED ON DATA SET OBTAINED FROM KAGGLE**

**SUMMARY:** This report is based on an analysis of a Dataset downloaded from Kaggle. The Dataset is about a collection of road accident data for a period of two, i.e.; 2021 to 2022 in England. The purpose of the analysis is to derive some insights based on certain requirements as outlined below. This report is a comparison of results earlier obtained from a dashboard created in Power BI earlier, which can be referred to. This analysis is performed in SQL.

**REQIUREMENTS:**

**TOOLS USED:**  Microsoft Power BI and SQL(Microsoft SQL Server).

**ASSUMPTIONS:** It is assumed that the audience understands how Microsoft SQL Server Management Studio environment is setup.

**REPORT**

**METHODOLGY:** Queries will be fired in SQL for each specific requirement and the result pasted under the query as a way of validating the output with Power BI.

**SQL QUERIES/RESULTS**

1. Primary KPIs without filters for other conditions
2. Current Year Total Casualties (CY Total Casualties) – 2022

SELECT SUM(number\_of\_casualties) AS CY\_Casualties

FROM road\_accident

WHERE YEAR(accident\_date) = '2022'

Output:



1. Current Year Total Accidents - 2022

SELECT DISTINCT(COUNT(accident\_index) )AS CY\_Accidents

FROM road\_accident

WHERE YEAR(accident\_date) = '2022'

Output:



1. Current Year Fatal Accidents

SELECT

SUM(number\_of\_casualties)

FROM road\_accident

WHERE YEAR(accident\_date) = '2022' AND accident\_severity = 'Fatal'

Output:



1. Current Year Serious Casualties

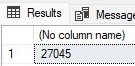
SELECT

SUM(number\_of\_casualties)

FROM road\_accident

WHERE YEAR(accident\_date) = '2022' AND accident\_severity = 'Serious'

Output:



1. Current Year Slight Casualties

SELECT

SUM(number\_of\_casualties)

FROM road\_accident

WHERE YEAR(accident\_date) = '2022' AND accident\_severity = 'Slight'

Output:



1. Current Year Casualties by Vehicle Type

SELECT

CASE

WHEN vehicle\_type IN('Agricultural vehicle') THEN 'Agricultural'

WHEN vehicle\_type IN ('Car','Taxi/Private hire car') THEN 'Cars'

WHEN vehicle\_type IN ('Motorcycle 125cc and under','Motorcycle 50cc and under', 'Motorcycle over 125cc and up to 500cc','Motorcyle over 500cc','Pedal cycle','Ridden horse') THEN 'Bikes'

WHEN vehicle\_type IN ('Bus or coach (17 or more pass seats)','Minibus (8 - 16 passenger seats)') THEN 'Buses'

WHEN vehicle\_type IN ('Goods 7.5 tonnes mgw and over','Goods over 3.5t. and under 7.5t','Van / Goods 3.5 tonnes mgw or under') THEN 'Trucks'

ELSE 'Others'

END As vehicle\_group,

SUM(number\_of\_casualties) as CY\_Casualties

FROM dbo.road\_accident

WHERE YEAR(accident\_date) = '2022'

GROUP BY

CASE

WHEN vehicle\_type IN('Agricultural vehicle') THEN 'Agricultural'

WHEN vehicle\_type IN ('Car','Taxi/Private hire car') THEN 'Cars'

WHEN vehicle\_type IN ('Motorcycle 125cc and under','Motorcycle 50cc and under', 'Motorcycle over 125cc and up to 500cc','Motorcyle over 500cc','Pedal cycle','Ridden horse') THEN 'Bikes'

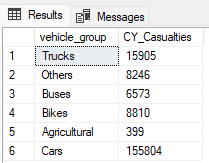
WHEN vehicle\_type IN ('Bus or coach (17 or more pass seats)','Minibus (8 - 16 passenger seats)') THEN 'Buses'

WHEN vehicle\_type IN ('Goods 7.5 tonnes mgw and over','Goods over 3.5t. and under 7.5t','Van / Goods 3.5 tonnes mgw or under') THEN 'Trucks'

ELSE 'Others'

END

Output:



1. Casualties by Month by Year

SELECT DATENAME(MONTH, accident\_date) AS Month\_Name,

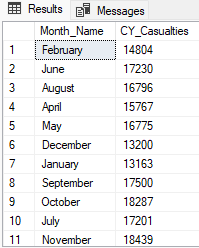
SUM(number\_of\_casualties) AS CY\_Casualties

FROM road\_accident

WHERE YEAR(accident\_date) = '2022'

GROUP BY DATENAME(MONTH, accident\_date)

Output:



SELECT DATENAME(MONTH, accident\_date) AS Month\_Name,

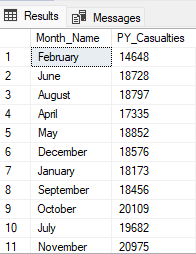
SUM(number\_of\_casualties) AS PY\_Casualties

FROM road\_accident

WHERE YEAR(accident\_date) = '2021'

GROUP BY DATENAME(MONTH, accident\_date)

Output:



viii) Casualties by Road Type by Year

SELECT road\_type,

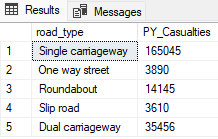
SUM(number\_of\_casualties) as PY\_Casualties

FROM dbo.road\_accident

WHERE YEAR(accident\_date) = '2021'

GROUP BY road\_type

Output:



SELECT road\_type,

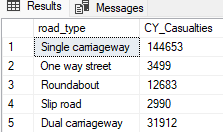
SUM(number\_of\_casualties) as CY\_Casualties

FROM dbo.road\_accident

WHERE YEAR(accident\_date) = '2022'

GROUP BY road\_type

Output:



ix) Casualties by Urban/Rural

SELECT

CASE

WHEN urban\_or\_rural\_area IN ( 'Urban') THEN 'Urban'

ELSE 'Rural'

END AS Area\_Type,

SUM(number\_of\_casualties) as PY\_Casualties

FROM dbo.road\_accident

WHERE YEAR(accident\_date) = '2021'

GROUP BY

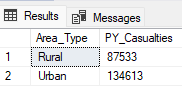
CASE

WHEN urban\_or\_rural\_area IN ( 'Urban') THEN 'Urban'

ELSE 'Rural'

END

Output:



x) Percentage Casualties by Urban/Rural

SELECT

urban\_or\_rural\_area, CAST(SUM(number\_of\_casualties) AS DEC(10,2))\*100/

(SELECT CAST(SUM(number\_of\_casualties) AS DEC(10,2)) FROM road\_accident WHERE YEar(accident\_date) = '2022')

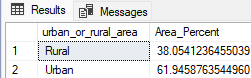
AS Area\_Percent

FROM road\_accident

WHERE YEAR(accident\_date) = '2022'

GROUP BY urban\_or\_rural\_area

Output:



xi) Casualties by Lighting conditions

SELECT

CASE

WHEN light\_conditions IN ('Darkness - lighting unknown','Darkness - lights lit','

Darkness - lights unlit','Darkness - no lighting') THEN 'Night'

ELSE 'Day'

END AS lighting\_group,

CAST(CAST(SUM(number\_of\_casualties) AS DEC(10,2)) \* 100/

(SELECT CAST(SUM(number\_of\_casualties) AS DEC (10,2)) FROM road\_accident WHERE YEAR(accident\_date) = '2022') AS DEC(10,2))

AS Percent\_lighting

FROM dbo.road\_accident

WHERE YEAR(accident\_date) = '2022'

GROUP BY

CASE

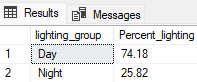
WHEN light\_conditions IN ('Darkness - lighting unknown','Darkness - lights lit','

Darkness - lights unlit','Darkness - no lighting') THEN 'Night'

ELSE 'Day'

END

Output:



xii) Casualty count by top ten Cities/Towns

SELECT

TOP 10 local\_authority, SUM(number\_of\_casualties) AS Total\_Casualties

FROM

dbo.road\_accident

GROUP BY

local\_authority

ORDER BY Total\_Casualties DESC

Output:

