Nagalaxmi Eepuri

Week 4: Reporting, Documentation, and Portfolio Building

Key Findings and Insights from the dataset, such as

Top busiest ports and states for crossings are San Ysidro -California, El Paso -Texas, Laredo-Texas, Hidalgo -Texas, Calexico-California, Buffalo Nagar falls-New York, Brownsville-Texas, Otay Mesa-California.

Measure Type Patterns: Understanding which crossing types are dominant helps align border resources with actual traffic, supporting better service delivery. From the dataset, there are higher in personal vehicle passengers, personal vehicle and trucks that cross US-Canada and US- Mexico borders.

Peak Crossing Months: Identifying high-traffic months for each border aids in seasonal workforce planning and infrastructure utilization. By analysing the dataset, I found the peak crossing months are March, May and December.

Yearly Trends: Analyzing trends over years supports strategic planning by revealing shifts in cross-border mobility patterns, which may relate to economic or policy changes.

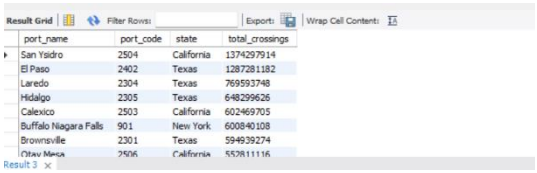
1. The busiest ports and states for crossings.

This query will give you a combined view of the busiest state-port combinations by total crossings, sorted in descending order.

QUERY:

```
SELECT port_name,port_code, state, SUM(value) AS total_crossings

FROM border_crossing_entry_data

GROUP BY port_name, port_code, state

ORDER BY total_crossings DESC

LIMIT 8;
```

OUTPUT: Here are the top 8 busiest ports and states for crossings.



port_name	port_code	state	total_crossings
San Ysidro	2504	California	1374297914
El Paso	2402	Texas	1287281182
Laredo	2304	Texas	769593748
Hidalgo	2305	Texas	648299626
Calexico	2503	California	602469705
Buffalo Niagara Falls	901	New York	600840108
Brownsville	2301	Texas	594939274
Otay Mesa	2506	California	553811116

Description:

1. SUM(value) calculates the total number of crossings for each port and each state.
2. The GROUP BY clause groups the data by port and state.
3. ORDER BY total_crossings DESC orders the results to show the busiest ports and states at the top.
4. LIMIT 8 restricts the results to the top 8 busiest ports and states.

---

2. Patterns and trends by border (U.S.-Mexico vs U.S.-Canada).

To analyze patterns and trends by border (U.S.-Mexico vs. U.S.-Canada) using the border_crossing_entry_data table, we can approach this by looking at the overall trends in crossings over

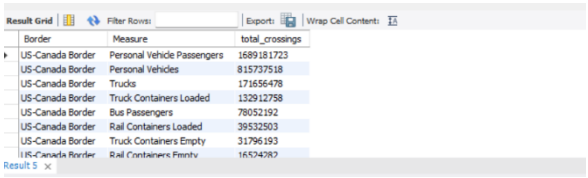
time for each border. This will help identify differences or similarities in traffic patterns between the two borders.

I. Analyze Patterns by Measure Type for Each Border

QUERY:

```
SELECT Border, Measure, SUM(Value) AS total_crossings

FROM border_crossing_entry_data

GROUP BY Border, Measure

ORDER BY Border, total_crossings DESC;
```

OUTPUT:



Border	Measure	total_crossings
US-Canada Border	Personal Vehicle Passengers	1689181723
US-Canada Border	Personal Vehicles	815737518
US-Canada Border	Trucks	171656478
US-Canada Border	Truck Containers Loaded	132912758
US-Canada Border	Bus Passengers	78052192
US-Canada Border	Rail Containers Loaded	39532503
US-Canada Border	Truck Containers Empty	31796193
US-Canada Border	Rail Containers Empty	16524282

Description:

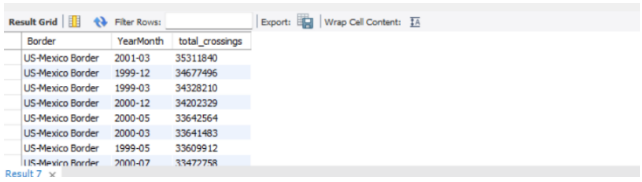
1. This query aggregates the total crossings by measure type for each border.
2. It helps to identify which types of crossings are most common at each border, indicating patterns such as whether trucks or pedestrians dominate traffic.

II. Peak Crossing Months by Border

QUERY:

```
SELECT Border, DATE_FORMAT(Date, '%Y-%m') AS YearMonth, SUM(Value) AS total_crossings

FROM border_crossing_entry_data

GROUP BY Border, YearMonth

ORDER BY  total_crossings DESC

LIMIT 350;
```

OUTPUT:



Border	YearMonth	total_crossings
US-Mexico Border	2001-03	35311840
US-Mexico Border	1999-12	34677496
US-Mexico Border	1999-03	34328210
US-Mexico Border	2000-12	34202329
US-Mexico Border	2000-05	33642564
US-Mexico Border	2000-03	33641483
US-Mexico Border	1999-05	33609912
US-Mexico Border	2000-07	33472758

Description:

1. This query finds the months with the highest number of crossings for each border.
2. Limiting to the top 12 results will show the peak crossing months, which can highlight seasonal peaks in traffic.

III. Yearly Trends for Each Border

Nagalaxmi Eepuri

QUERY:

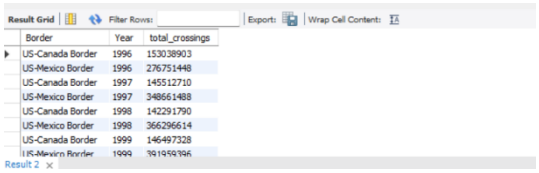
SELECT Border, YEAR(Date) **AS** Year, **SUM**(value) **AS** total_crossings

FROM border_crossing_entry_data

GROUP BY Border, Year

ORDER BY Year, Border;

OUTPUT:



Border	Year	total_crossings
US-Canada Border	1996	153038903
US-Mexico Border	1996	276751448
US-Canada Border	1997	145512710
US-Mexico Border	1997	348661488
US-Canada Border	1998	142291790
US-Mexico Border	1998	366296614
US-Canada Border	1999	146497328
US-Mexico Border	1999	381959396

Description:

1. This query groups the data by border and year to analyze trends in crossing volumes annually.
2. It will help identify if there are any long-term increases or decreases in traffic at each border.

---

3. How the number of crossings varies by measure type, year, and month

I. The number of crossings varies by measure type and year

QUERY:

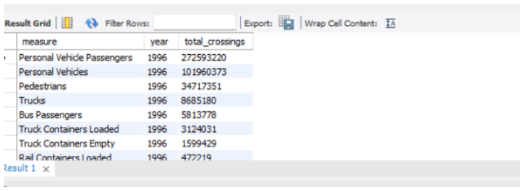
SELECT measure, YEAR(date) **AS** year, **SUM**(value) **AS** total_crossings

FROM border_crossing_entry_data

GROUP BY measure, YEAR(date)

ORDER BY year, total_crossings **DESC;**

OUTPUT:



measure	year	total_crossings
Personal Vehicle Passengers	1996	272593220
Personal Vehicles	1996	101960373
Pedestrians	1996	34717351
Trucks	1996	8685180
Bus Passengers	1996	5813778
Truck Containers Loaded	1996	3124031
Truck Containers Empty	1996	1599429
Rail Containers Loaded	1996	472219

II. The number of crossings varies by measure type, year and month.

QUERY:

SELECT measure, DATE_FORMAT(date, '%Y-%m') **AS** YearMonth, **SUM**(value) **AS** total_crossings

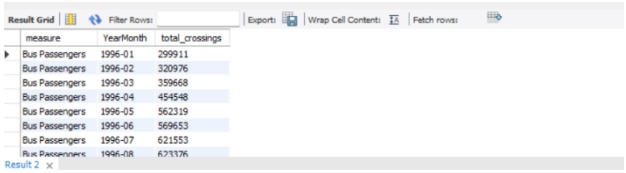
FROM border_crossing_entry_data

GROUP BY measure, YearMonth

ORDER BY measure, YearMonth;

Nagalaxmi Eepuri

OUTPUT:



measure	YearMonth	total_crossings
Bus Passengers	1996-01	299911
Bus Passengers	1996-02	320976
Bus Passengers	1996-03	359668
Bus Passengers	1996-04	454548
Bus Passengers	1996-05	562319
Bus Passengers	1996-06	569653
Bus Passengers	1996-07	621553

Description:

1. DATE_FORMAT(date, '%Y-%m') combines the year and month into a single column, making it easier to read the trends.
2. The ORDER BY measure, year_month orders the data by measure type and then chronologically, highlighting how crossing numbers change over time.

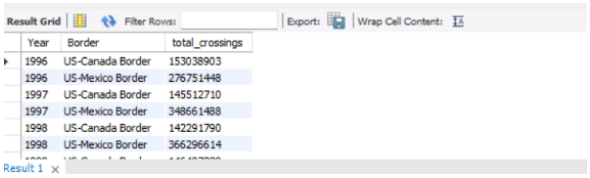
This approach will give you a detailed view of how the number of crossings varies by measure type, year, and month, making it easier to identify trends, patterns, and seasonal variations for each crossing type.

Total crossings per year and border.

QUERY:

```
SELECT YEAR(Date) AS Year, Border, SUM(Value) AS total_crossings
FROM border_crossing_entry_data
GROUP BY Border, YEAR(Date)
 ORDER BY Year, Border;
```

OUTPUT:



Year	Border	total_crossings
1996	US-Canada Border	153038903
1996	US-Mexico Border	276751448
1997	US-Canada Border	145512710
1997	US-Mexico Border	348661488
1998	US-Canada Border	142291790
1998	US-Mexico Border	366296614

Description:

- YEAR(date) extracts the year from the date column.
- border represents the type of border (e.g., U.S.-Canada or U.S.-Mexico).
- SUM(value) calculates the total number of crossings for each year and border.
- GROUP BY YEAR(date), border groups the data by year and border.
- ORDER BY year, border sorts the result by year and border.

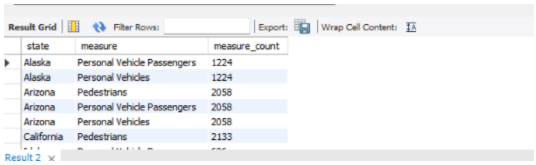
Most common measure types by state.

QUERY:

```
SELECT bc.state, bc.measure, COUNT(*) AS measure_count
```

```
FROM border_crossing_entry_data bc
GROUP BY bc.state, bc.measure
HAVING COUNT(*) = (
    SELECT MAX(measure_count)
    FROM (
      SELECT state, measure, COUNT(*) AS measure_count
      FROM border_crossing_entry_data
      GROUP BY state, measure
    ) AS subquery
    WHERE subquery.state = bc.state
  )
ORDER BY bc.state;
```

OUTPUT:



state	measure	measure_count
Alaska	Personal Vehicle Passengers	1224
Alaska	Personal Vehicles	1224
Arizona	Pedestrians	2058
Arizona	Personal Vehicle Passengers	2058
Arizona	Personal Vehicles	2058
California	Pedestrians	2133

Description:

- The main query counts the occurrences of each measure type grouped by state.
- The HAVING clause compares the count of each measure type to the maximum count found for that specific state.
- The subquery generates the counts for each measure and state and finds the maximum count for each state.

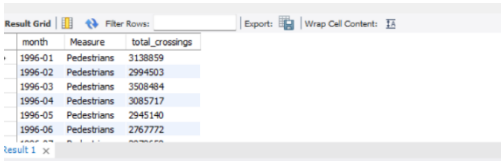
Monthly trends in truck and pedestrian crossings.

QUERY:

```
SELECT DATE_FORMAT(Date, '%Y-%m') AS month, Measure, SUM(Value) AS total_crossings
FROM border_crossing_entry_data
WHERE Measure IN ('Truck', 'Pedestrians')
GROUP BY DATE_FORMAT(Date, '%Y-%m'), Measure
ORDER BY month, Measure;
```

OUTPUT:

Nagalaxmi Eepuri



month	Measure	total_crossings
1996-01	Pedestrians	3138859
1996-02	Pedestrians	2994503
1996-03	Pedestrians	3508484
1996-04	Pedestrians	3085717
1996-05	Pedestrians	2945140
1996-06	Pedestrians	2767772

Description:

 DATE_FORMAT(date, '%Y-%m') extracts the year and month from the date column in YYYY-MM format to group data by month.

 measure represents the type of crossing, such as truck or pedestrian.

 SUM(value) calculates the total number of crossings for each measure type.

 The WHERE clause filters the results to include only truck and pedestrian crossings.

 GROUP BY DATE_FORMAT(date, '%Y-%m'), measure groups the data by month and measure type.

 ORDER BY month, measure sorts the results by month and measure type.

Key Findings and Insights:

- The busiest ports and states for crossings are
- San Ysidro California
- El Paso Texas
- Laredo Texas
- Hdalgo Texas
- Calexico California
- Buffalo nagar falls New York
- Brownscille Texas
- Otay Mesa California