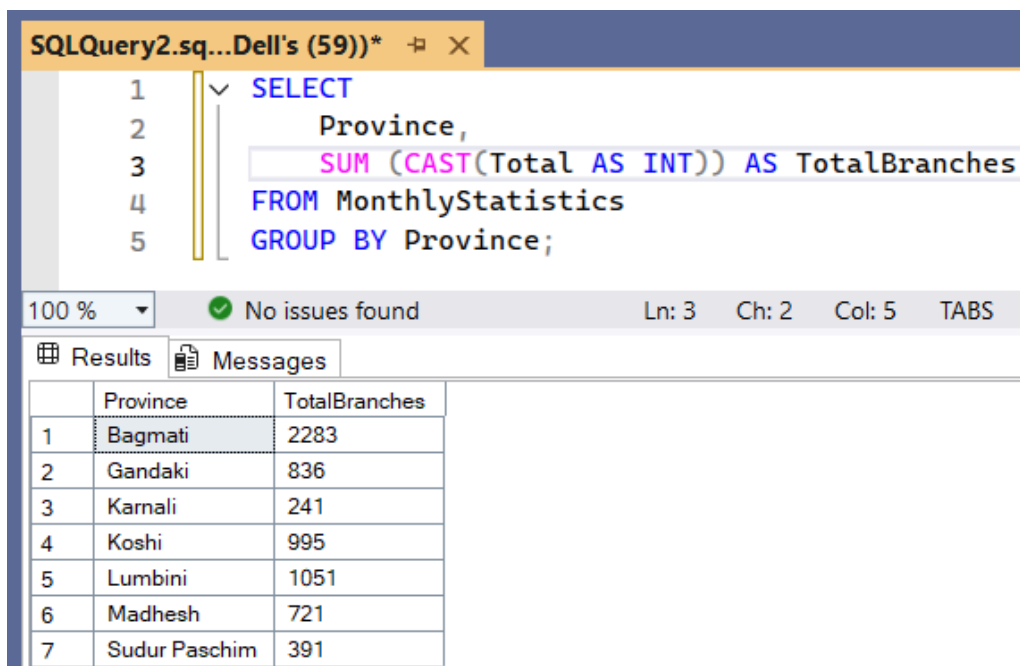# Lab Work 6: Data Aggregation and Reporting with SQL Stored Procedures

Objective:

Learn how to aggregate, analyze, and report branch and population data by writing SQL queries and automating them within a stored procedure. Understand key SQL concepts like grouping, aggregation, filtering, unions, and ranking.

Steps:

1. Total Branches by Province: Write a query to calculate the total number of branches in each province by summing the "Total" column grouped by province.

2. Average Population per Branch by Province: Write a query to find the average population per branch for each province using the "Pop._Per_Branch" column.

3. Identify Districts with Inconsistent Branch Counts: Find districts where the sum of Class_A, Class_B, and Class_C does not match the value in the "Total" column, indicating data inconsistency.

4. Provinces with More than 10 Branches: Write a query to list provinces where the total number of branches exceeds 100, using SQL's HAVING clause.

5. List Provinces and Districts with Type Labels: Create a combined list of all provinces and districts, tagging each row with a type label — either "Province" or "District".

6. Districts with High Branch Counts in Specific Classes: Identify districts where Class_A branches exceed 10 or Class_B branches exceed 15, showcasing filtering with OR conditions and UNION.

```
31   ∨ SELECT District, Province, Class_A, Class_B
32     FROM Monthlystatistics
33     WHERE Class_A > 10
34     UNION
35     SELECT District, Province, Class_A, Class_B
36     FROM Monthlystatistics
37     WHERE Class_B > 15;
```

91 %   ▼   ✔ No issues found

⊞ Results 📖 Messages

	District	Province	Class_A	Class_B
1	Achham	Sudur Paschim	18	3
2	Arghakhanchi	Lumbini	21	7
3	Baglung	Gandaki	50	9
4	Baitadi	Sudur Paschim	22	2
5	Bajhang	Sudur Paschim	21	1
6	Bajura	Sudur Paschim	13	1
7	Banke	Lumbini	82	29
8	Bara	Madhesh	81	13
9	Bardiya	Lumbini	45	26
10	Bhaktapur	Bagmati	125	19

7. District with Highest Population in Each Province: Write a query to find the district(s) with the highest population in every province.

```
39   ∨ SELECT m.*
40     FROM Monthlystatistics m
41     JOIN (
42         SELECT Province, MAX(Population) AS MaxPopulation
43         FROM Monthlystatistics
44         GROUP BY Province
45     ) p ON m.Province = p.Province AND m.Population = p.MaxPopulation;
```

91 %   ▼   ✔ No issues found   ◄   ►   Ln: 45   Ch: 67

⊞ Results 📖 Messages

	Province	District	Class_A	Class_B	Class_C	Total	Population	Pop_Per_Branch	month_year
1	Sudur Paschim	Kailali	132	30	5	167	904666	5417.161676646707	Jestha 2082
2	Madhesh	Dhanusa	89	8	13	110	867747	7888.609090909091	Jestha 2082
3	Lumbini	Kapilbastu	86	18	6	110	682961	6208.736363636363	Jestha 2082
4	Koshi	Jhapa	139	54	12	205	998054	4868.556097560976	Jestha 2082
5	Karnali	Mugu	9	0	0	9	64549	7172.111111111111	Jestha 2082
6	Gandaki	Kaski	205	77	19	301	600051	1993.5249169435217	Jestha 2082
7	Bagmati	Chitawan	165	53	18	236	719859	3050.25	Jestha 2082

8. Top 3 Districts with the Highest Number of Branches Overall: Retrieve the top three districts ranked by total branches across all provinces.

```
47   SELECT TOP 3 Province, District, Total
48     FROM Monthlystatistics
49     ORDER BY CAST(Total AS INT) DESC;
```

91 %   ⊘ No issues found

Results **Messages**

	Province	District	Total
1	Bagmati	Kathmandu	1135
2	Lumbini	Rupandehi	326
3	Gandaki	Kaski	301

9. Number of Districts and Average Population per Province: For each province, calculate the number of districts it contains and the average population across those districts.


```
51   SELECT
52     Province,
53       COUNT (District) AS District_Count, AVG(CAST(Population AS BIGINT)) AS Avg_Population
54     FROM Monthlystatistics
55     GROUP BY Province;
```

91 %   ⊘ No issues found   Ln: 54   Ch: 23   TABS   C

Results **Messages**

	Province	District_Count	Avg_Population
1	Bagmati	13	470528
2	Gandaki	11	224220
3	Karnali	10	168841
4	Koshi	14	354386
5	Lumbini	12	426839
6	Madhesh	8	764325
7	Sudur Paschim	9	299420

10. Provinces with Average Branches per District Greater Than 20: Identify provinces where the average number of branches per district exceeds 20.


```
57   SELECT
58     Province,
59       AVG(CAST(Total AS INT)) AS Avg_Branches
60     FROM Monthlystatistics
61     GROUP BY Province
62     HAVING AVG(CAST(Total AS INT)) > 20;
```

91 %   ⊘ No issues found

Results **Messages**

	Province	Avg_Branches
1	Bagmati	175
2	Gandaki	76
3	Karnali	24
4	Koshi	71
5	Lumbini	87
6	Madhesh	90
7	Sudur Paschim	43

11. Combine all above queries into a single stored procedure that inserts their results into corresponding summary tables for automated reporting.

```sql
CREATE PROCEDURE sp_GenerateMonthlyReports
AS
BEGIN
        --Provinces with More than 100 Branches
        INSERT INTO Prov_Branches_Summary
        SELECT Province, SUM(CAST(Total AS INT)) AS Total_Branches
        FROM Monthlystatistics
        GROUP BY Province
        HAVING SUM(CAST (Total AS INT)) > 100;
        --List Provinces and Districts with Type Labels
        INSERT INTO Prov_Dist_Type
        SELECT DISTINCT Province AS Name, Province AS Type
        FROM Monthlystatistics
        UNION
        SELECT District AS Name, 'District' AS Type
        FROM Monthlystatistics;
        --Districts where Class A 10 OR Class_B > 15
        INSERT INTO High_Class_Branches
        SELECT District, Province, Class_A, Class_B
        FROM Monthlystatistics
        WHERE CAST (Class_A AS INT) > 10
        UNION
        SELECT District, Province, Class_A, Class_B
        FROM Monthlystatistics
        WHERE CAST(Class_B AS INT) > 15;
        --District with Highest Population in Each Province
        INSERT INTO MaxPop_District
        SELECT m.*
        FROM Monthlystatistics m
        JOIN (
            SELECT Province, MAX(CAST(Population AS BIGINT)) AS MaxPop
            FROM Monthlystatistics m
            GROUP BY Province
        )p ON m. Province = p. Province
        AND CAST(m.Population AS BIGINT) = p.MaxPop;
        --Top 3 Districts by Branches
        INSERT INTO Top3_Branches
        SELECT TOP 3 District, Province, CAST(Total AS INT) AS Total
        FROM Monthlystatistics
        ORDER BY CAST (Total AS INT) DESC;
        --Number of Districts & Average Population
        INSERT INTO Prov_Dist_Avg
        SELECT Province, COUNT (District) AS District_Count,
            AVG(CAST (Population AS BIGINT)) AS Avg_Population
        FROM Monthlystatistics
        GROUP BY Province;
        --Provinces with Avg Branches 20
        INSERT INTO High_Avg_Branches
        SELECT Province, AVG(CAST(Total AS INT)) AS Avg_Branches
        FROM Monthlystatistics
        GROUP BY Province
        HAVING AVG(CAST(Total AS INT)) > 20
END;
```

Messages

Commands completed successfully.

Completion time: 2025-08-28T20:35:46.7083585+05:45

Conclusion

The key takeaways from this lab include the following important aspects:

- Practice creating and managing multiple summary tables.
- Understand how to refresh reports programmatically by truncating and inserting data.
- Learn about SQL Server Agent or other scheduling tools to automate stored procedure execution.
- Reflect on how data automation aids timely and accurate business decision-making.