Internship Program : Soulvibe.Tech

PROJECT TITLE:

“Analyzing Socioeconomic and Demographic Influences on Income"

BATCH NAME : SVT/DAINT/2025/06/B09.

BY

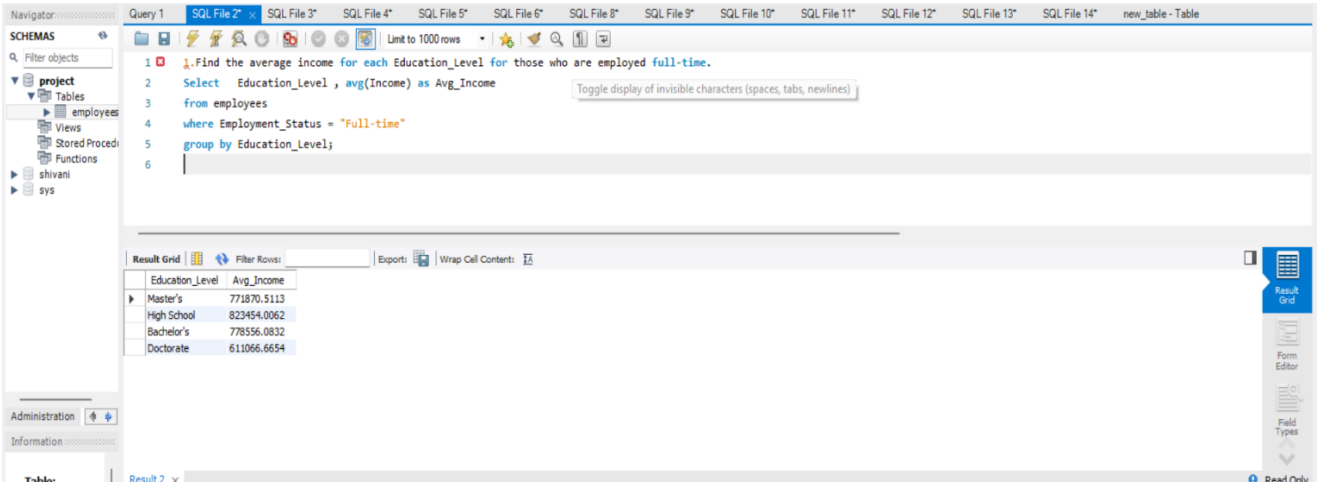
K.SHIVANI

Introduction

Overview of the main objective

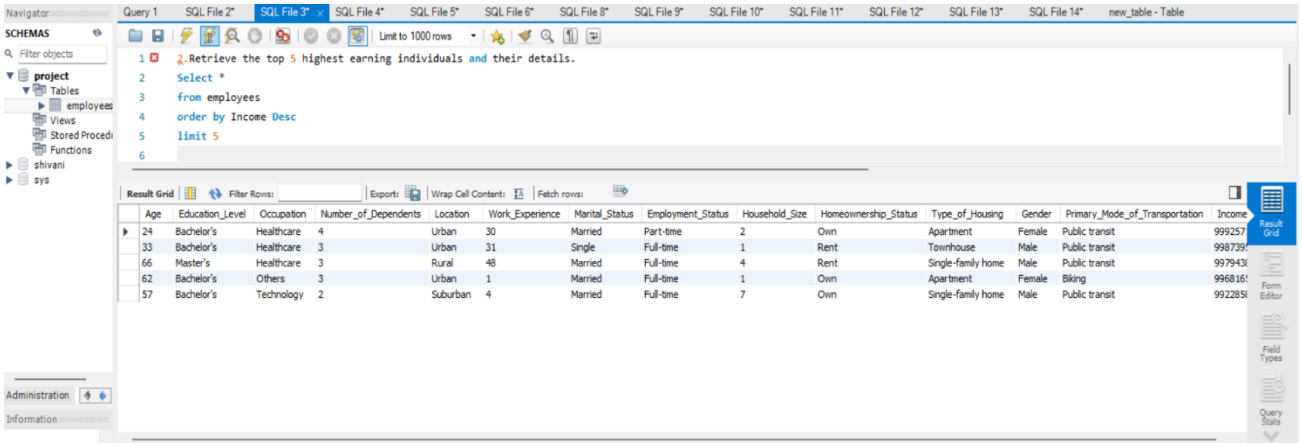
I conducted a thorough analysis of socioeconomic and demographic data for this assignment using SQL. I gained important insights into the variables affecting income levels and living circumstances by using queries to filter, group, and aggregate data on income, household characteristics, employment, and education. A greater comprehension of the connections between individual demographics and economic results was made possible by this analytical method. The results offer a strong basis for strategic planning and well-informed decision-making.

1. Find the average income for each Education_Level for those who are employed full-time?



In order to examine how education affects salaries in the full-time workforce, this query determines the average income for each level of education among those who are employed full-time.

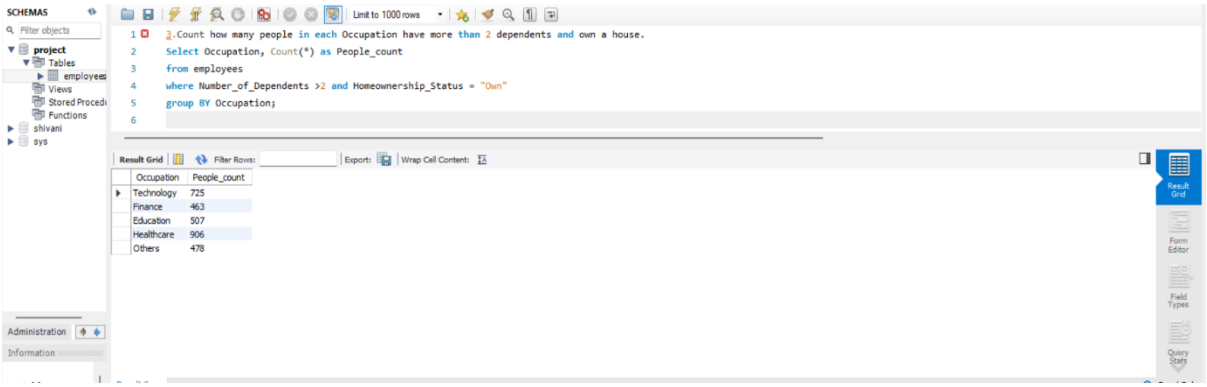
2. Retrieve the top 5 highest earning individuals and their details?



Age	Education_Level	Occupation	Number_of_Dependents	Location	Work_Experience	Marital_Status	Employment_Status	Household_Size	Homeownership_Status	Type_of_Housing	Gender	Primary_Mode_of_Transportation	Income
24	Bachelor's	Healthcare	4	Urban	30	Married	Part-time	2	Own	Apartment	Female	Public transit	999257
33	Bachelor's	Healthcare	3	Urban	31	Single	Full-time	1	Rent	Townhouse	Male	Public transit	998739
66	Master's	Healthcare	3	Rural	48	Married	Full-time	4	Rent	Single-family home	Male	Public transit	997943
62	Bachelor's	Others	3	Urban	1	Married	Full-time	1	Own	Apartment	Female	Biking	996816
57	Bachelor's	Technology	2	Suburban	4	Married	Full-time	7	Own	Single-family home	Male	Public transit	992285

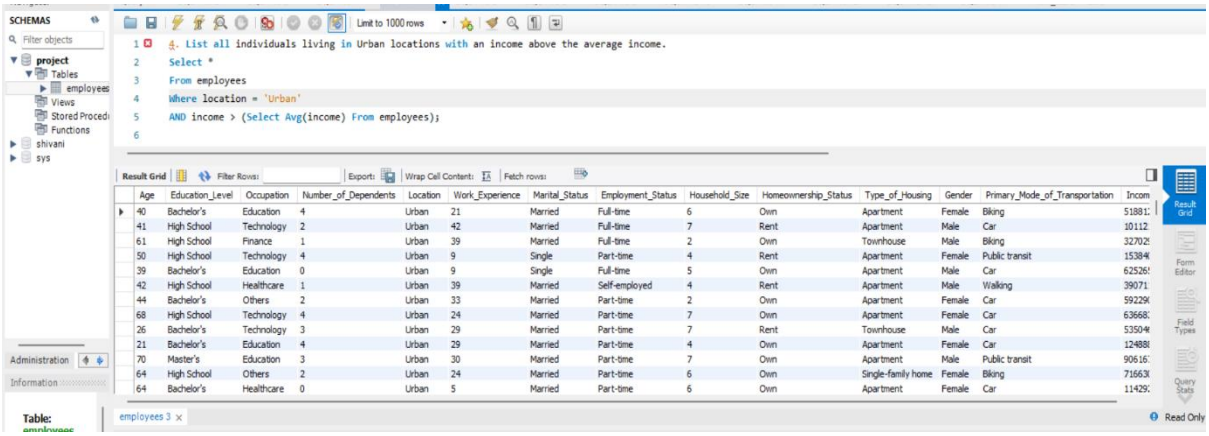
This query limits the output to the top five entries and sorts the dataset in descending order of income to return the top five highest earning persons along with all of their details. For more research or comparison, it assists in identifying high-income profiles.

3. Count how many people in each Occupation have more than 2 dependents and own a house?



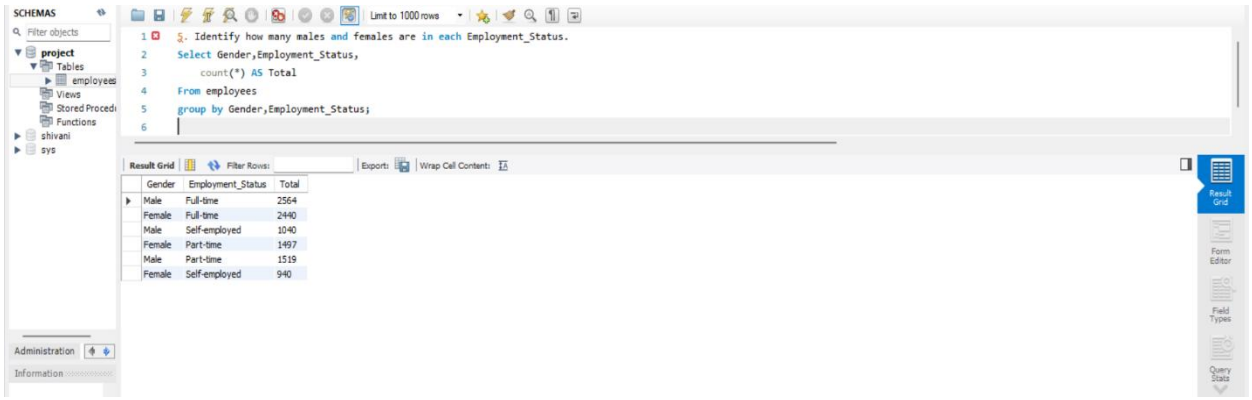
The purpose of this query is to examine how homeownership, family size, and occupation are related. Those who own their homes and have more than two dependents are the target audience. A better insight of which occupational groupings are more likely to support larger families while keeping homeownership is provided by the query, which groups the data by occupation and counts the number of people who match both criteria. Assessing housing patterns across various work sectors and finding financially solid professions can both benefit from this knowledge.

4. List all individuals living in Urban locations with an income above the average income?



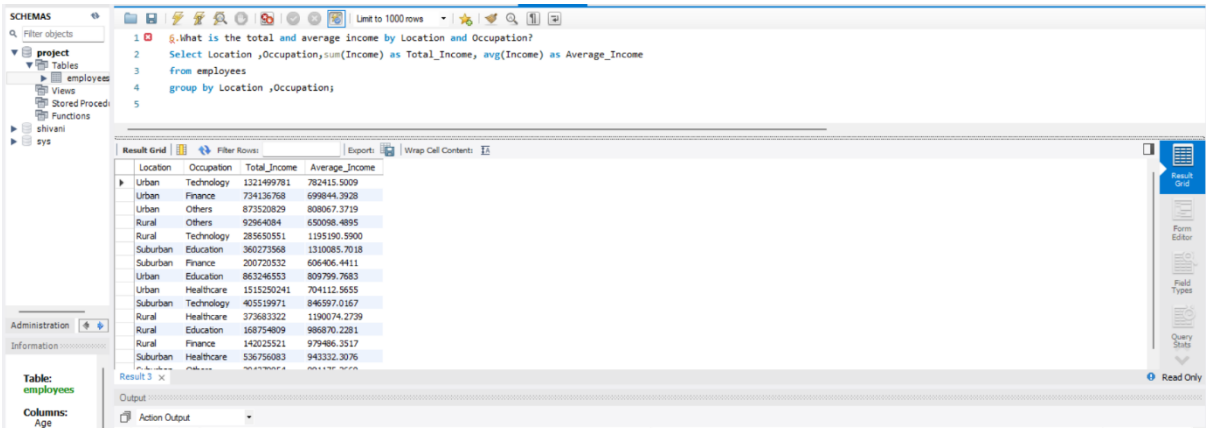
This search finds people with higher-than-average incomes who live in cities. By highlighting high incomes in urban areas, it offers valuable information for economic research, resource planning, and targeted marketing.

5. Identify how many males and females are in each Employment_Status?



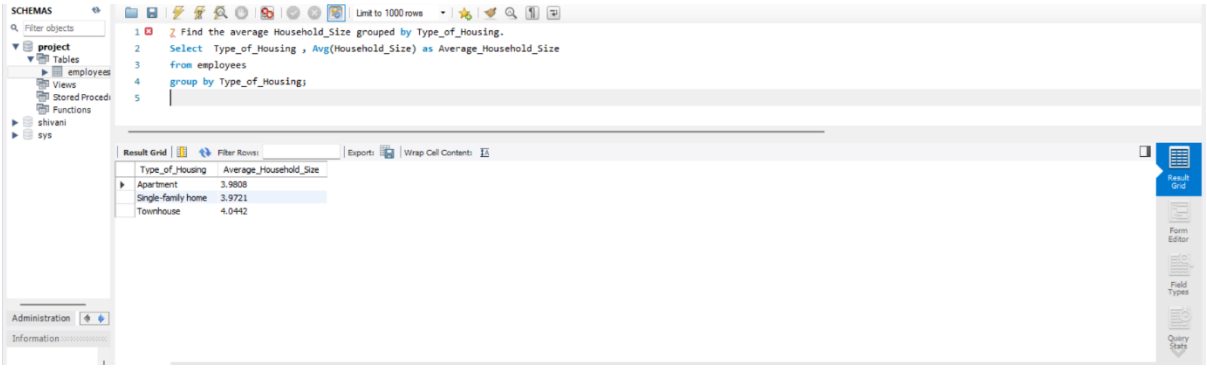
This query counts the number of men and women in each Employment_Status category. Understanding the gender distribution across various employment types and identifying potential differences in gender are made easier by classifying the data by gender and employment status.

6. What is the total and average income by Location and Occupation?



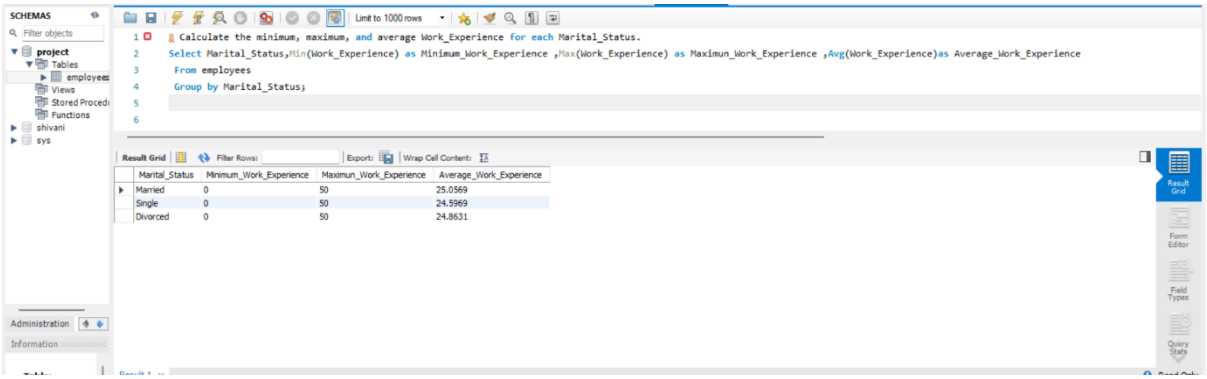
This query determines the average and total income by occupation and location. It provides important insights into how incomes differ by profession and geography by combining income data from various employment roles and geographic locations. Understanding economic differences and adjusting company or policy initiatives accordingly are made easier with the use of that information.

7. Find the average Household_Size grouped by Type_of_Housing?



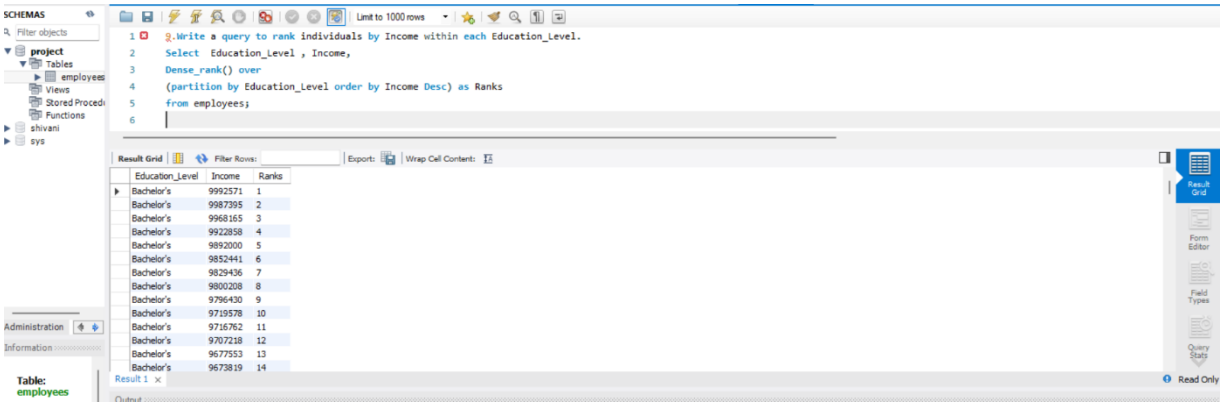
The average number of household members for each form of reflecting is calculated by this query. It provides information on family structures and living space efficiency by displaying the differences in household size across different housing categories.

8. Calculate the minimum, maximum, and average Work_Experience for each Marital_Status?



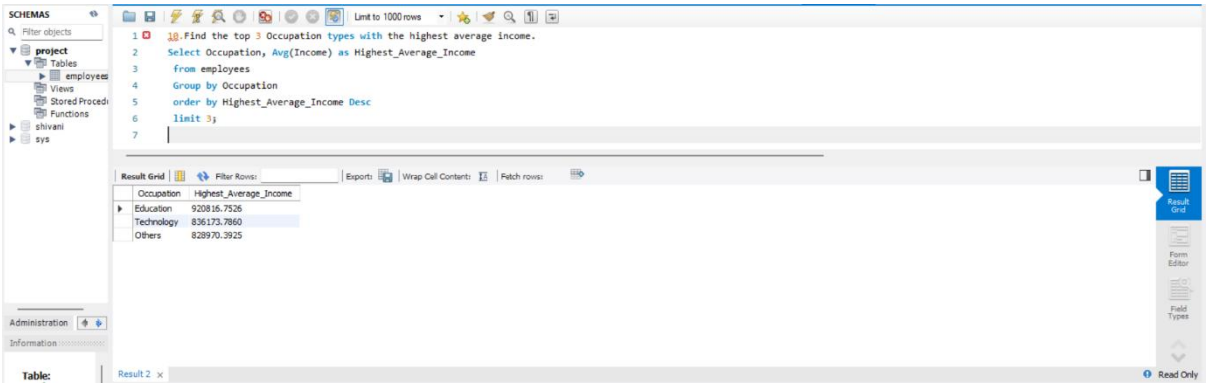
The minimum, maximum, and average work experience for each individual sorted by Marital_Status is determined by this query. It provides awareness on how job experience differs amongst various marital groups, which can help with research on demographic trends and career growth.

9. Write a query to rank individuals by Income within each Education_Level.



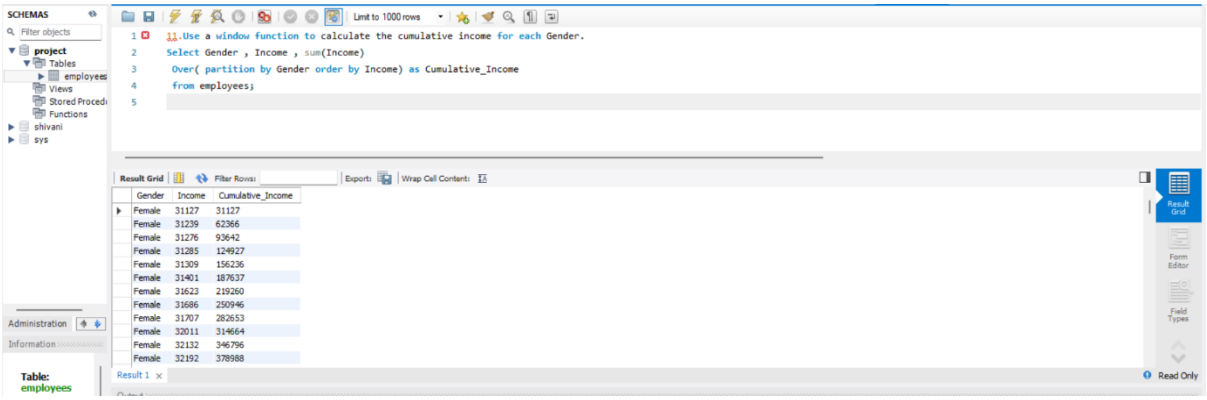
People in each educational level are ranked by their income in this regard. By ranking each individual based on how much they earn as compared to those with the same educational background, it becomes easier to identify high earners and to understand how income is distributed among various educational groupings.

10. Find the top 3 Occupation types with the highest average income?



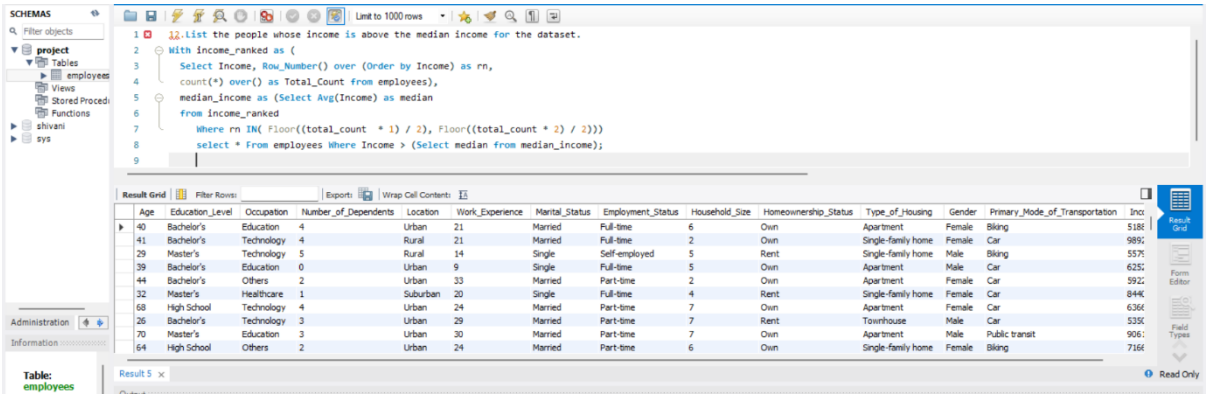
By determining the mean salary for each occupation and ranking them appropriately, this query determines the top three occupations with the greatest average income. This research offers important insights into wage discrepancies across professions by highlighting the most expensive job categories within the dataset.

11. Use a window function to calculate the cumulative income for each Gender?



This query determines each gender's cumulative income using a window function. It offers insight into the distribution and aggregation patterns of income between males and females by adding up the incomes in each gender group in a running total order.

12. List the people whose income is above the median income for the dataset?



The list of people whose income is higher than the dataset's median income is gathered by this query. It identifies individuals in the upper half of the income range and highlights the higher earners by comparing each person's income to the median value. This provides insights into income distribution.

CONCLUSION

- In order to find trends in income, employment, education, and living conditions, I Investigated important socioeconomic and demographic factors using a SQL-based analysis of the employee's table. I was able to learn more about lifestyle characteristics, household dynamics, and geographic patterns by composing questions.

- Important Lessons Learnt: SQL made structured and effective data exploration possible.

- I gained a deeper knowledge of data linkages through querying.

- It will assist us in understanding employee's behaviour.


THANKYOU