Integrated Demographic and Behavioural Analysis Project

1. Age Distribution: What is the distribution of ages among the dataset? Are there any age groups that stand out?

Query –

select

case

when age >=18 and age < 35 then 'Young'

when age >=35 and age <=50 then 'Middle Age'

else 'Old'

end as age\_group,

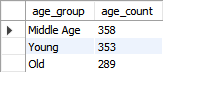
count(\*) as age\_count

from data

group by age\_group

order by age\_count desc;

output-



Ans - After run SQL query I found “Middle Age” Age\_group is the Highest count of individuals, followed by “Young” Age\_group. The “Old” Age\_group is slightly less count of individuals compared to other two age\_groups. Therefore, while there is a relatively balanced distribution among the age groups, the middle-aged individuals appear to be slightly more prevalent in the dataset.

1. Gender Representation: How does gender distribution look like in the dataset? Is there any significant gender imbalance?

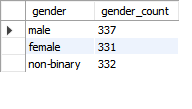
Query –

select gender, count(\*) as gender\_count

from data

group by gender;

output -



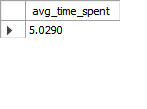
Ans - After the query , I got Male – 337 , Female – 331 and Non-binary(LGBTQ) - 332 , this data doesn’t have a significant gender imbalance. The Number of Males, Females and Non-binary(LGBTQ) are relatively close . Therefore, the dataset appears to have a reasonably balanced representation across gender categories.

1. Time Spent Analysis: What is the average time spent on the platform? Are there any patterns based on age, gender, or location?

Query –

select avg(time\_spent) as avg\_time\_spent

from data;



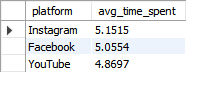
Ans – after run query for average time spent is 5.02 hours.

Query 2 -

select platform, avg(time\_spent) as avg\_time\_spent

from data

group by platform;



Ans- after run the query for average time spent by platforms , in this Instagram is the highest average time spent which is 5.15 hours followed closely Facebook average time spent is 5.05 hours, and Youtube average time spent is 4.86 hours. If we look to the results

We thought that average time spent according to platforms is not have too much difference, but when we talk about social media every minute make a huge difference.

Query 3 –

select avg(time\_spent) as average\_time\_spent,

avg(case when age >=18 and age <35 then time\_spent end) as avg\_time\_spent\_young,

avg(case when age >= 35 and age <50 then time\_spent end) as avg\_time\_spent\_middle\_aged,

avg(case when age >= 50 then time\_spent end) as avg\_time\_spent\_old,

avg(case when gender = 'Male' then time\_spent end) as avg\_time\_spent\_male,

avg(case when gender = 'Female' then time\_spent end) as avg\_time\_spent\_female,

avg(case when gender = 'Non-Binary' then time\_spent end) as avg\_time\_spent\_non\_binary,

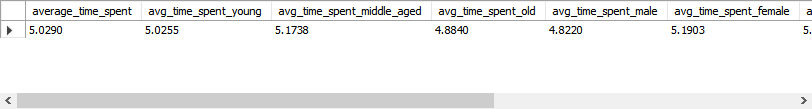
avg(case when location = 'United Kingdom' then time\_spent end) as avg\_time\_spent\_UK,

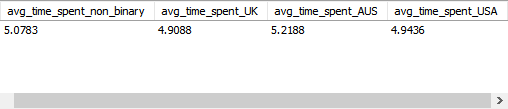
avg(case when location = 'Australia' then time\_spent end) as avg\_time\_spent\_AUS,

avg(case when location = 'United States' then time\_spent end) as avg\_time\_spent\_USA

from data;

output –





I wrote a different query for the same question

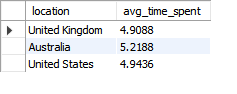
Query –

select location , avg(time\_spent) as avg\_time\_spent

from data

group by location ;

output-



Ans- Users from Australia have the highest average time spent, at 5.21 hours.

Users from the United States have the second-highest average time spent, at 4.94 hours.

Users from the UK have the lowest average time spent, at 4.90 hours.

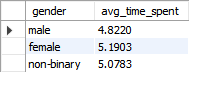
Query –

select gender ,avg(time\_spent) as avg\_time\_spent

from data

group by gender;

output –



Ans - Females have the highest average time spent, at 5.19 hours.

Non-binary (LGBTQ) have an average time spent of 5.07 hours.

Males have the lowest average time spent, at 4.82 hours

Query –

select

case

when age >=18 and age < 35 then 'Young'

when age >=35 and age <=50 then 'Middle Age'

else 'Old'

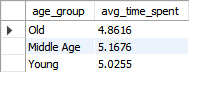
end as age\_group,

avg(time\_spent) as avg\_time\_spent

from data

group by age\_group;

output –



Ans – The average time spent by different age groups is as follows:

“Middle-aged” age group have the highest average time spent, at 5.16 hours.

“Young“ age group have an average time spent of 5.02 hours.

“Old” age group have the lowest average time spent, at 4.86 hours.

1. Platform Preferences: Which platform is most popular among users? Is there a difference in platform preference based on demographics or interests?

Query 1 -

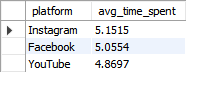
select platform ,avg(time\_spent) as avg\_time\_spent

from data

group by platform

order by avg\_time\_spent desc;

output –



Ans - The most popular platform among users, based on the average time spent, is Instagram, with an average time spent of 5.1515 hours. Facebook follows with an average time spent of 5.0554 hours, and YouTube comes next with an average time spent of 4.8697 hours. Therefore, Instagram appears to be the most popular platform among the users in terms of time spent.

Query 2 –

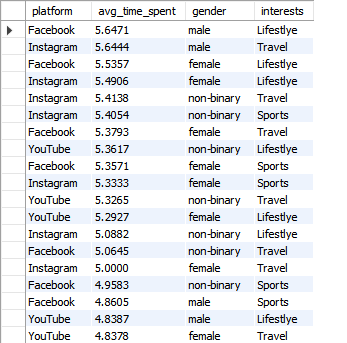
select platform, avg(time\_spent) as avg\_time\_spent ,gender, interests

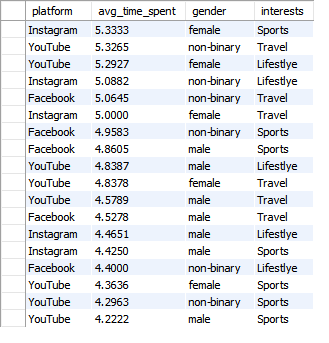
from data

group by platform, gender, interests

order by avg\_time\_spent;

output –





Answer - According to Demographics : -

Male: Facebook has the highest average time spent (5.6471 hours), followed closely by Instagram (5.6444 hours). YouTube follows with a slightly lower average time spent (4.8387 hours).

Female: Instagram is the most popular platform with the highest average time spent (5.4906 hours), followed by Facebook (5.5357 hours) and YouTube (5.2927 hours).

Non-binary: Instagram is the preferred platform with the highest average time spent (5.4138 hours), followed by Facebook (5.0645 hours) and YouTube (4.2963 hours).

According to Interests : -

Lifestyle: Instagram is the most popular platform among individuals interested in Lifestyle, with the highest average time spent (5.4082 hours), followed by Facebook (5.4418 hours) and YouTube (4.8677 hours).

Travel: Instagram is the preferred platform among individuals interested in Travel, with the highest average time spent (5.3154 hours), followed by Facebook (5.1818 hours) and YouTube (4.9490 hours).

Sports: Instagram is also popular among individuals interested in Sports, with the highest average time spent (4.7358 hours), followed by Facebook (4.7594 hours) and YouTube (4.3926 hours).

1. Location Insights: How does user distribution vary across different locations?

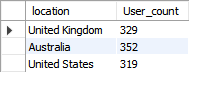
Query 1 –

select location, count(\*) as User\_count

from data

group by location;

output –



ans - This distribution reveals variations in user presence across locations, with Australia boasting the largest user base, followed by the United Kingdom and the United States.

1. Interest Analysis: What are the most common interests among users? Are there any trends or patterns in interests based on demographics or location?

Query 1 –

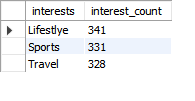
select interests , count(\*) as interest\_count

from data

group by interests

order by interest\_count desc;

output –



Ans - These results indicate that Lifestyle is the most common interest among users, followed closely by Sports and Travel.

Query 2 –

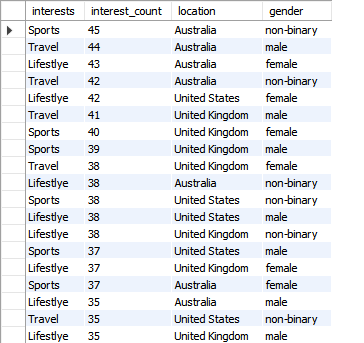
select interests , count(\*) as interest\_count,location, gender

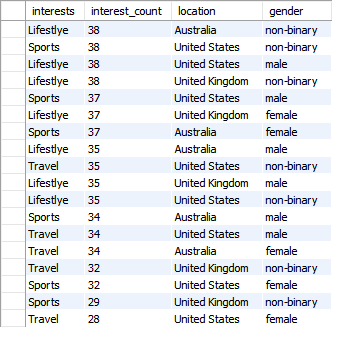
from data

group by interests,location, gender

order by interest\_count desc;

output –





Ans - Location-based Trends:

In Australia, the most common interest is Sports, followed closely by Lifestyle and Travel.

In the United States, Lifestyle is the predominant interest among users, followed by Sports and Travel.

In the United Kingdom, Travel seems to be the most common interest, followed by Lifestyle and Sports.

Demographic-based Trends:

Among non-binary individuals, Lifestyle appears to be the most common interest across all locations.

Among males, Travel is relatively more popular, especially in Australia and the United Kingdom.

Among females, Lifestyle is a prevalent interest, particularly in Australia and the United States.

1. Profession Analysis: What are the predominant professions among users? Is there any correlation between profession and income level?

Query –

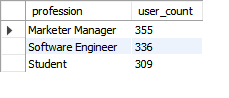
select profession, count(\*) as user\_count

from data

group by profession

order by user\_count desc;

output -



Ans- These results indicate that Marketer Manager is the most common profession among users, followed closely by Software Engineer and Student.

Query 2 –

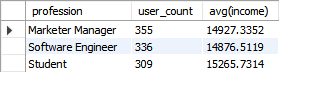
select profession, count(\*) as user\_count, avg(income)

from data

group by profession

order by user\_count desc;

output –



Ans - While the income levels are relatively close, it seems that Students have a slightly higher average income compared to Marketer Managers and Software Engineers. However, without additional statistical analysis, such as calculating correlation coefficients or conducting regression analysis, it's challenging to determine the strength and direction of the correlation between profession and income level definitively.

8. Income Distribution: Are there any significant differences based on demographics or location?

Query –

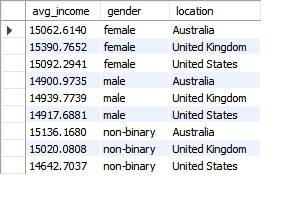
select avg(income) as avg\_income, gender, location

from data

group by gender, location

order by gender, location;

output –



Ans - Gender-based Differences:

Among females, the average income level is highest in the United Kingdom ($15,390.7652), followed by Australia ($15,062.6140) and the United States ($15,092.2941).

Among males, the average income level is highest in the United Kingdom ($14,939.7739), followed by Australia ($14,900.9735) and the United States ($14,917.6881).

Among non-binary individuals, the average income level is highest in Australia ($15,136.1680), followed by the United Kingdom ($15,020.0808) and the United States ($14,642.7037).

Location-based Differences:

Across all genders, the United Kingdom has the highest average income level, followed by Australia and the United States.

9 . Debt Status: What proportion of users are in debt? Is there any correlation between debt status and other variables like income or homeownership?

Query –

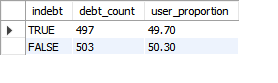
select indebt,count(\*) as debt\_count,

cast(count(\*) \* 100 / (select count(\*) from data)as decimal(10, 2)) as user\_proportion

from data

group by indebt;

output –



Ans- Therefore, approximately 49.70% of users are in debt, while the remaining 50.30% are not in debt.

Query 2 –

select indebt, count(\*) as debt\_count,

avg(income) as avg\_income,

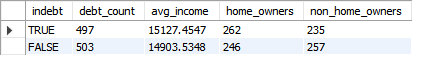
sum(case when isHomeOwner = 'True' then 1 else 0 end) as home\_owners,

sum(case when isHomeOwner = 'False' then 1 else 0 end) as non\_home\_owners

from data

group by indebt;

Output –



Ans – Income:

The average income for users in debt (TRUE) is $15,127.4547, while for users not in debt (FALSE), it is $14,903.5348. This suggests that users in debt tend to have a slightly higher average income compared to those not in debt.

Homeownership:

Among users in debt (TRUE), there are 262 homeowners and 235 non-homeowners.

Among users not in debt (FALSE), there are 246 homeowners and 257 non-homeowners.

This indicates that there might not be a strong correlation between debt status and homeownership, as the counts are relatively balanced between homeowners and non-homeowners in both debt and non-debt categories.

10 . Homeownership Analysis: What percentage of users are homeowners? Is there any correlation between homeownership and other variables like income or car ownership?

Query 1-

select count(\*) ,cast(count(\*) \* 100.0 / (SELECT count(\*) FROM data)as decimal(10,2)) isHomeOwner

from data

group by isHomeOwner

having isHomeOwner = 'True';

output –



Ans – Therefore, 50.80% of users are homeowners.

Query 2 –

select isHomeOwner,

avg(income) as avg\_income,

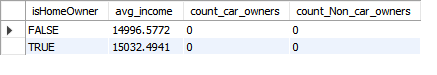
sum(case when Owns\_Car = 'Yes' then 1 else 0 end)as car\_owners,

sum(case when Owns\_Car = 'No' then 1 else 0 end)as Non\_car\_owners

from data

group by isHomeOwner;

output –



Ans –

Income:

The average income for users who are homeowners (TRUE) is $15,032.4941, while for users who are not homeowners (FALSE), it is $14,996.5772. This suggests that homeowners tend to have a slightly higher average income compared to non-homeowners.

Car Ownership:

All users who are homeowners (TRUE) and non-homeowners (FALSE) have a reported percentage of 0% for both car ownership and non-car ownership. This suggests that there might not be a correlation between homeownership and car ownership based on the provided data.

Overall, there seems to be a weak correlation between homeownership and income, with homeowners having a slightly higher average income. However, there doesn't appear to be a correlation between homeownership and car ownership based on the provided data. Further analysis with more detailed data or additional variables might provide a clearer picture of these relationships.

11 . Car Ownership: How many users own a car? Is there any correlation between car ownership and other variables like income or location?

Query 1 –

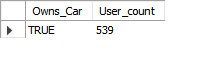
select Owns\_Car ,count(\*) as User\_count

from data

group by Owns\_Car

Having Owns\_Car = 'True';

Output –



Ans - Therefore, there are 539 users who own a car.

Query 2 –

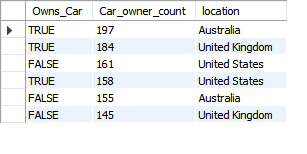
select Owns\_Car ,count(\*) as Car\_owner\_count, location

from data

group by Owns\_Car, location

order by Car\_owner\_count desc;

output –



Ans - Location:

Australia: 197 users

United Kingdom: 184 users

United States: 158 users

Among users who do not own a car (FALSE), the counts for the same locations are:

United States: 161 users

Australia: 155 users

United Kingdom: 145 users

This indicates that there might be some correlation between car ownership and location, as the distribution of car ownership varies across different countries.

Query –

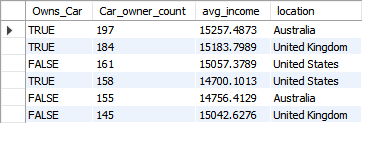
select Owns\_Car ,count(\*) as Car\_owner\_count, avg(income)as avg\_income ,location

from data

group by Owns\_Car , location

order by Car\_owner\_count desc;

output –



Ans –

Income:

Among users who own a car (TRUE), the average income levels for each location are as follows:

Australia: $15,257.4873

United Kingdom: $15,183.7989

United States: $14,700.1013

Among users who do not own a car (FALSE), the average income levels for each location are as follows:

United States: $15,057.3789

Australia: $14,756.4129

United Kingdom: $15,042.6276

This suggests that there might be some correlation between car ownership and income, as users who own a car tend to have slightly higher average incomes compared to those who do not own a car, especially in Australia and the United Kingdom.