Student course analysis – data analytics project

use project;

select \* from `cloud counseleage data analytics project`;

-- 1) how many unique students are include in the dataset

select count(distinct `email id`) from `cloud counseleage data analytics project`;

-- 2) what is average GPA of students

select round(avg(cgpa),2) as avg\_gpa from `cloud counseleage data analytics project`;

-- 3) distribution of students among different graduation years

select `Year of Graduation`,count(`ï»¿first name`) as total\_students from `cloud counseleage data analytics project`

group by `Year of Graduation`

order by total\_students asc;

-- 4) distribution of students experience with python programing language.

select `Year of Graduation`,round(avg(`Experience with python (Months)`),2) as month\_experience from `cloud counseleage data analytics project`

group by `Year of Graduation`;

-- 5) average family income of students

select round(avg(`family income in lakh`),0) as avg\_family\_income\_lakhs from `cloud counseleage data analytics project`;

-- 6) how does gpa vary among different colleges?

select `College Unknownme`,round(avg(cgpa),2) as avg\_cgpa from `cloud counseleage data analytics project`

group by `College Unknownme`

order by avg\_cgpa desc limit 5 ;

-- 7) any outliers in number of course completed

-- yes it is present remove outlier for better accuracy and getting proper insights

-- 8) average gpa from each city

select `City`,round(avg(cgpa),2) as avg\_cgpa from `cloud counseleage data analytics project`

group by `City`

order by avg\_cgpa desc limit 5 ;

-- 9) can we identify any relationship between family income and gpa?

select \* from `cloud counseleage data analytics project`;

select `family income in lakh`,round(avg(cgpa),2) as avg\_cgpa from `cloud counseleage data analytics project`

group by `family income in lakh`;

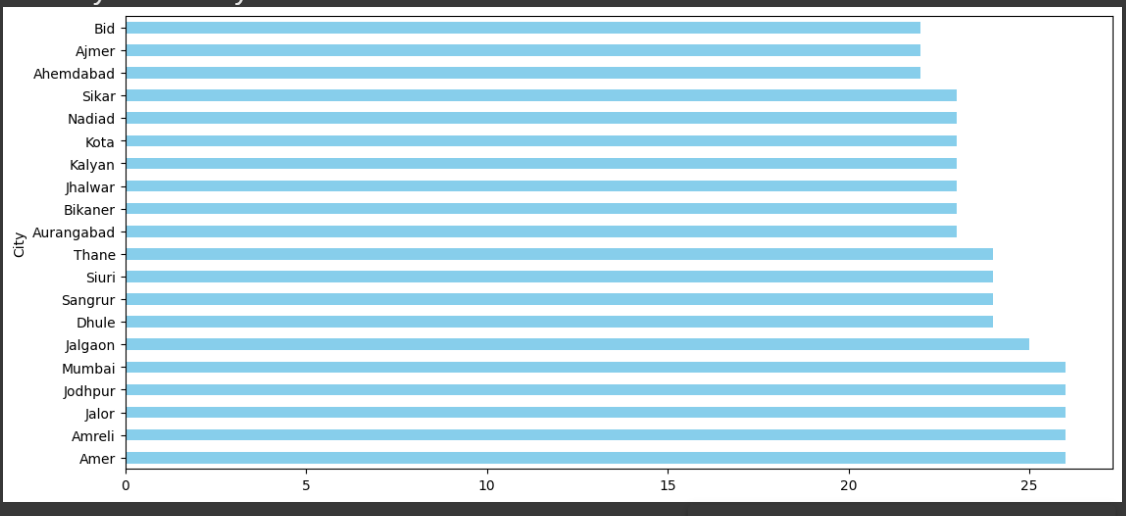
-- no any relationship here we can se family income of 5 lakh have more avg\_cgpa than family income of 7 lakh

-- 10) how many students from various cities (done in python pandas )

-- done using python library like matplotlib and seaborn we just see top 20 city from students is coming

-- to see run sql scripts below

-- select City,count(Quantity) as total from `cloud counseleage data analytics project` group by City order by total desc;



-- 11) how does expected salary vary based on factrs like gpa ,family income ,experience with python

SELECT

`Expected salary (Lakh)`,

`family income in lakh` AS family\_income,

cgpa AS cgpa,

`Experience with python (Months)` AS python\_exp

FROM

`cloud counseleage data analytics project`

order by `Expected salary (Lakh)`;

-- it is not fully varies for whole data , little data varies for expected salary having cgpa and python\_exp more , family\_income doesn't have relationship with expected salary

-- 12)whcih eventt tend to attract more students from specific field of study

select \* from `cloud counseleage data analytics project`;

select `how do you know about this event ?`,count(quantity) as total\_students from `cloud counseleage data analytics project`

group by `how do you know about this event ?`

order by total\_students desc

limit 10;

-- 13 ) do students in learship position during the collge year tend to have higheer gpa or better expected salary?

select max(CGPA),avg(`Expected salary (Lakh)`) from `cloud counseleage data analytics project`

where `Leadership- skills` = "yes"

order by avg(`Expected salary (Lakh)`)desc;

-- 14) how many students graduating by end of year 2024?

select count(quantity) as total\_students from `cloud counseleage data analytics project`

where `Year of Graduation` <= 2024;

-- 15)which promotoion channel bring in more students participating for the event

select `how do you know about this event ?`,count(quantity) as total\_students\_participate from `cloud counseleage data analytics project`

group by `how do you know about this event ?`

order by total\_students\_participate desc

limit 10;

-- 16)find total number students who attented the events related to data science?

select count(quantity) as total\_students from `cloud counseleage data analytics project`

where events like "Data%" or "Artificial%" or "Hello%";

-- 17)those who have high cgpa & more experience in language those who had expectiotion for salary?

select max(CGPA),max(`Experience with python (Months)`),avg(`Expected salary (Lakh)`) from `cloud counseleage data analytics project`

order by avg(`Expected salary (Lakh)`)desc;

-- yes it diifer

-- 18)how many students know about event from their collagous which of these top 3 colleges ?

with cte as (

select `College Unknownme`,count(quantity) as total\_students,dense\_rank() over(order by count(quantity) desc) as rnk

from `cloud counseleage data analytics project`

where `how do you know about this event ?`='Others'

group by `College Unknownme`

order by total\_students desc

)

select \* from cte

where rnk<=5