**EXPLORING ACADEMY**

**TRENDS IN STARTUP**

**Project – 1**

***Problem:***

To analyze a dataset of startup companies using SQL queries, exploring various metrics to understand trends in the startup ecosystem.

***Database:***  TrendsInStartups\_Explorin.csv

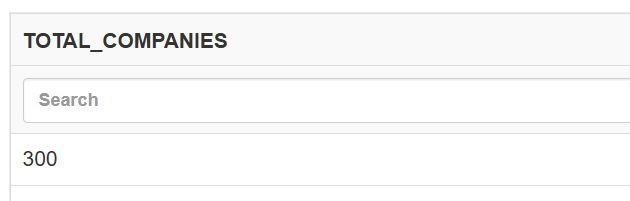
***TASK WITH SOLUTIONS GIVEN BELOW:***

1.Calculate the total number of companies in the dataset.

SELECT count (\*) as TOTAL\_COMPANIES

FROM "TrendsInStartups\_Explorin"

OUTPUT

****

2.Determine the total value of all companies in the dataset.

SELECT SUM(VALUATION)

FROM "TrendsInStartups\_Explorin"

OUTPUT



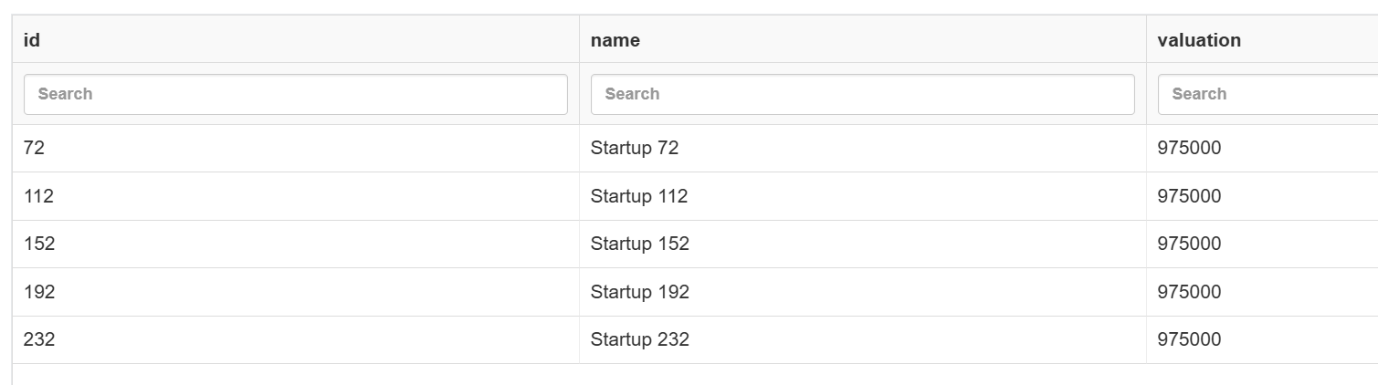
3.Find the highest amount raised by a startup at the 'Seed' stage.

SELECT ID, NAME, VALUATION

FROM "TrendsInStartups\_Explorin"

ORDER BY valuation DESC LIMIT 5

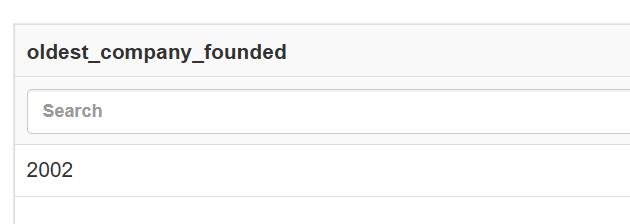
OUTPUT



4. Identify the year when the oldest company on the list was founded.

SELECT MIN(founded\_year) as oldest\_company\_founded

FROM "TrendsInStartups\_Explorin"

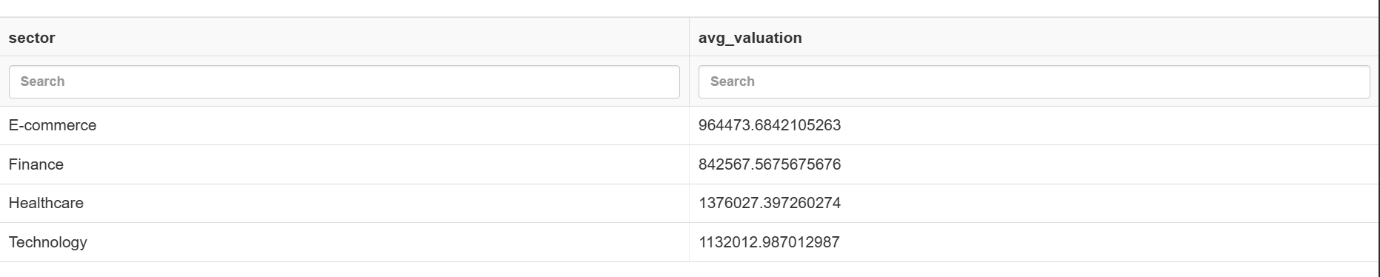
OUTPUT

5. Calculate the average valuation within each startup category.

SELECT sector, avg(valuation) as avg\_valuation

FROM "TrendsInStartups\_Explorin"

group by sector

OUTPUT 

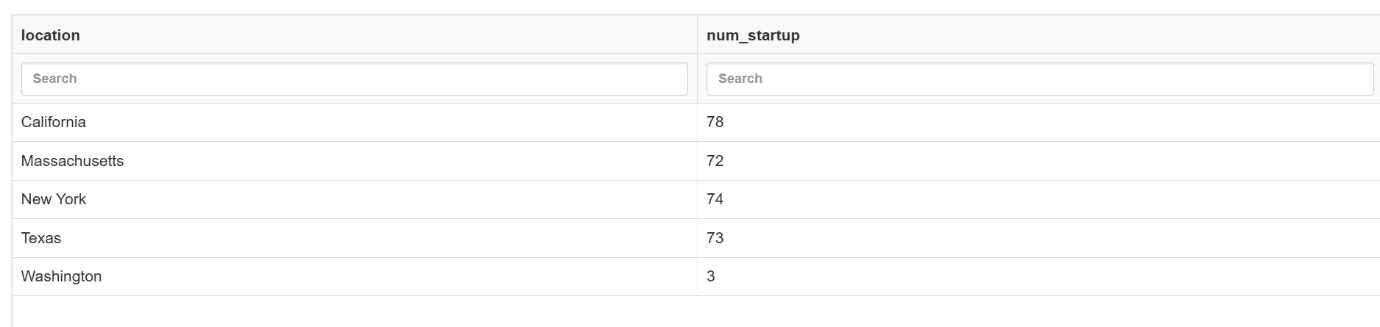
6. Determine the top locations with the highest number of startups.

SELECT LOCATION, count(location) as num\_startup

FROM "TrendsInStartups\_Explorin"

GROUP by location

OUTPUT



7. Calculate the average size of startups in each location where the average size exceeds 500.

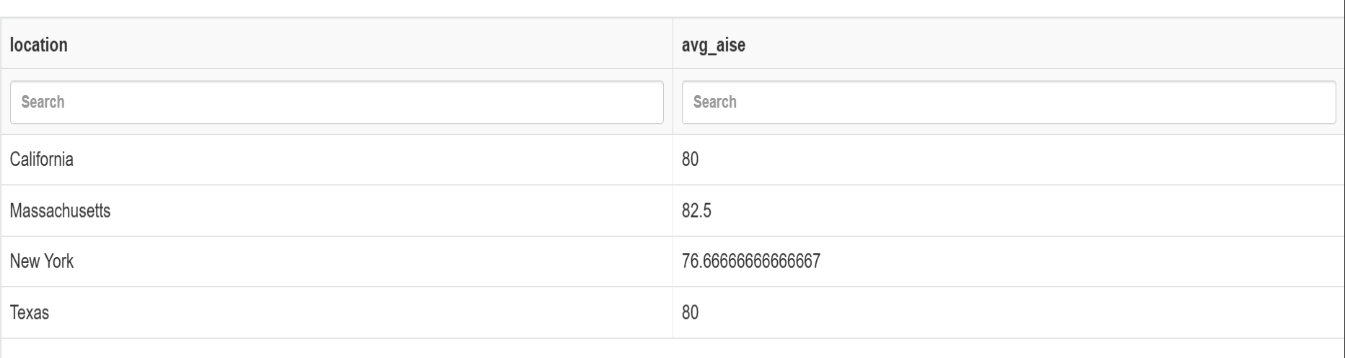
SELECT DISTINCT location, avg(size) as avg\_aise

FROM "TrendsInStartups\_Explorin"

where size > 500

group by location

OUTPUT



8. Find the top 5% of startups with the highest valuations.

SELECT count(id)

FROM "TrendsInStartups\_Explorin"

ORDER by valuation desc limit 15

OUTPUT



9. Identify startups that have raised funding in every stage (Seed, Series A, Series B, etc.).

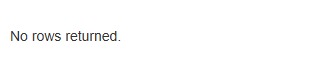
SELECT name , count(stage) as A

FROM "TrendsInStartups\_Explorin"

group by name

having A=4

OUTPUT



10.Calculate the percentage growth in valuation from Seed stage to Series A for each startup.

SELECT s1.name, ((max(s1.valuation) - min(s2.valuation)) \* 100.0 / min(s2.valuation)) as growth\_percentage

from "TrendsInStartups\_Explorin" as s1

join "TrendsInStartups\_Explorin.csv 3" as s2 on s1.name=s2.name

where s1.stage = 'Seed' and s2.stage = 'Series A' group by s1.name

OUTPUT

