



PROJECT ON-
Explorin Academy Trends in Startups

Database:
TrendsInStartups Explorin

Hand On Platform:
<https://dumbmatter.com/csv-sql-live/>

NAME-DIYA RANA SUBMITTED TO- PRASHANT JINDAL
MBA 1ST SEMESTER
ROLL NO. 242014063

Explorin Academy Trends in Startups

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

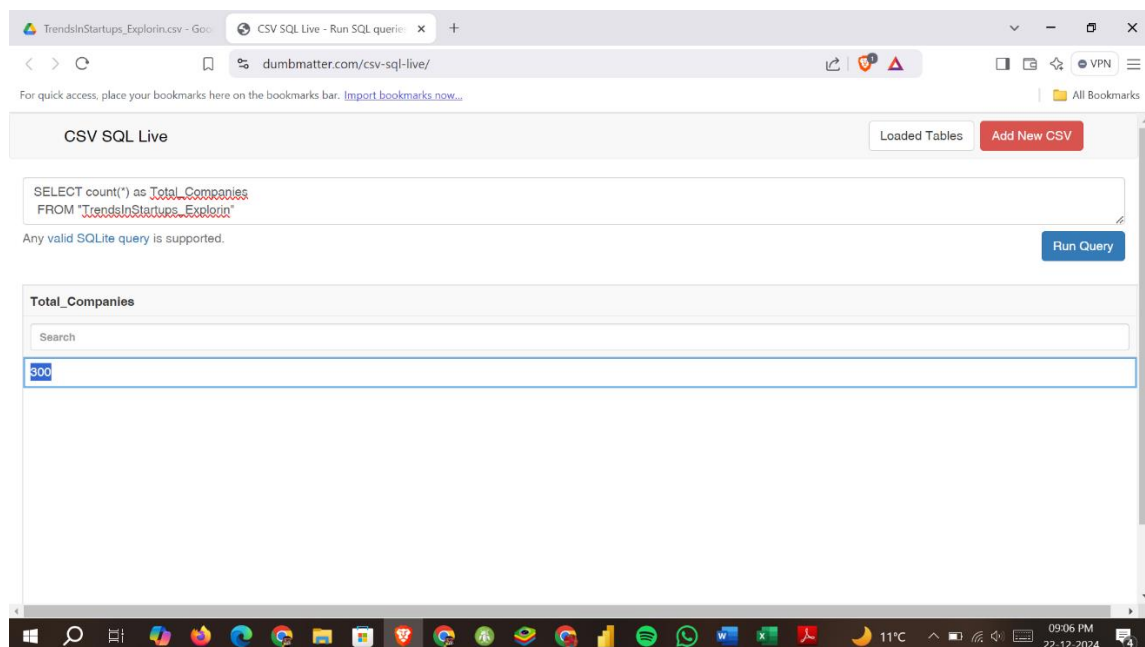
Task:

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

```
SELECT count(*) as TOTAL_COMPANIES
```

```
FROM "TrendsInStartups_Explorin"
```

OUTPUT = 300

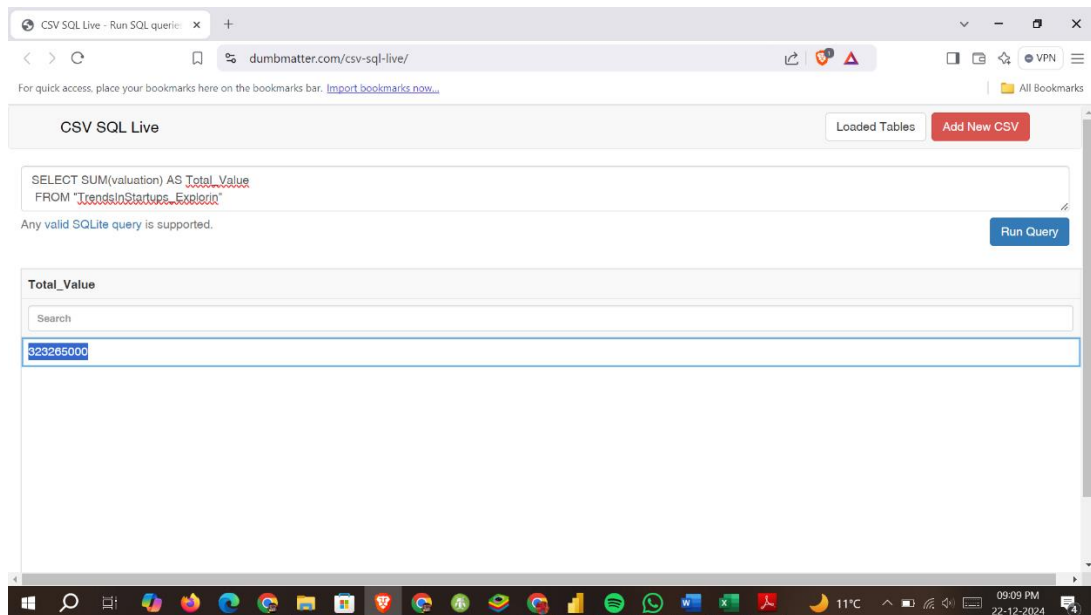


2 .Determine the total value of all companies in the dataset.

```
SELECT SUM(VALUATION)
```

```
FROM "TrendsInStartups_Explorin"
```

OUTPUT = 323265000

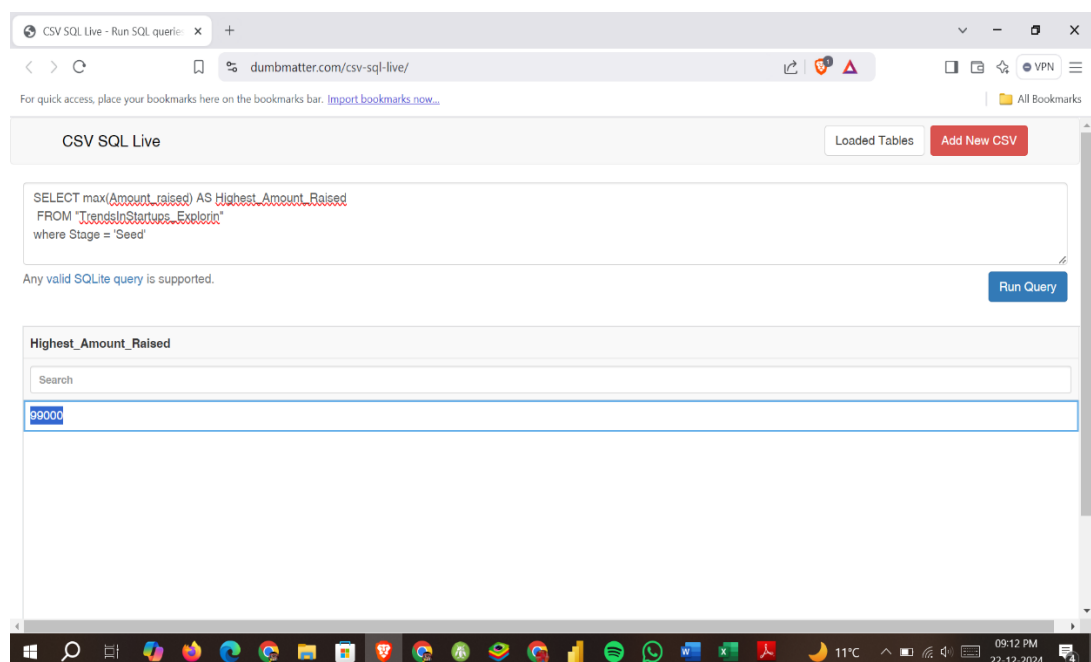


3 .Find the highest amount raised by a startup at the 'Seed' stage.

```
SELECT MAX(Amount_raised) AS highest_amount_raised
FROM "TrendsInStartups_Explorin"
```

WHERE Stage = 'Seed'

OUTPUT = 99000



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 = 2002

The screenshot shows the CSV SQL Live web application. The query entered is: `SELECT min(founded_year) as Oldest_Company_Founded FROM "TrendsInStartups_Explorin"`. The result is displayed in a table with the column `Oldest_Company_Founded` and a single row containing the value `2002`.

Oldest_Company_Founded
2002

5 .Calculate the average valuation within each startup category.

SELECT avg(valuation) ,sector

FROM "TrendsInStartups_Explorin"

group by sector

The screenshot shows the CSV SQL Live web application. The query entered is: `SELECT sector, avg(valuation) FROM "TrendsInStartups_Explorin" group by sector`. The result is displayed in a table with two columns: `avg(valuation)` and `sector`.

avg(valuation)	sector
954177.2151898735	E-commerce
834605.2631578947	Finance
1376193.3333333333	Healthcare
1187062.5	Technology

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

SELECT location as NAME, count(location) as num_startups

FROM "TrendsInStartups_Explorin"

Group by location

The screenshot shows the CSV SQL Live web application. The query entered is: `SELECT location as NAME, count(*) as num_startups FROM "TrendsInStartups_Explorin" GROUP BY location`. The results table shows the following data:

NAME	num_startups
California	78
Massachusetts	72
New York	74
Texas	73
Washington	3

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

SELECT DISTINCT location, avg(size) as avg_size

FROM "TrendsInStartups_Explorin"

group by location

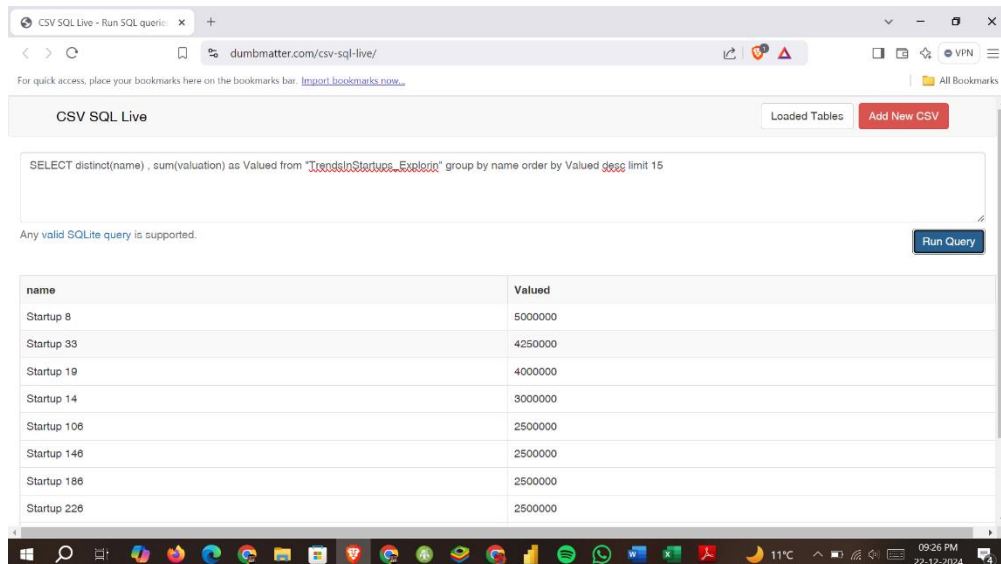
having AVG(size) > 500

output = null

The screenshot shows the CSV SQL Live web application. The query entered is: `SELECT DISTINCT location, avg(size) as avg_size FROM "TrendsInStartups_Explorin" group by location having AVG(size) > 500`. The results section shows "No rows returned."

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

SELECT distinct(name) , sum(valuation) as Valued from "TrendsInStartups_Explorin" group by name order by Valued desc limit 15

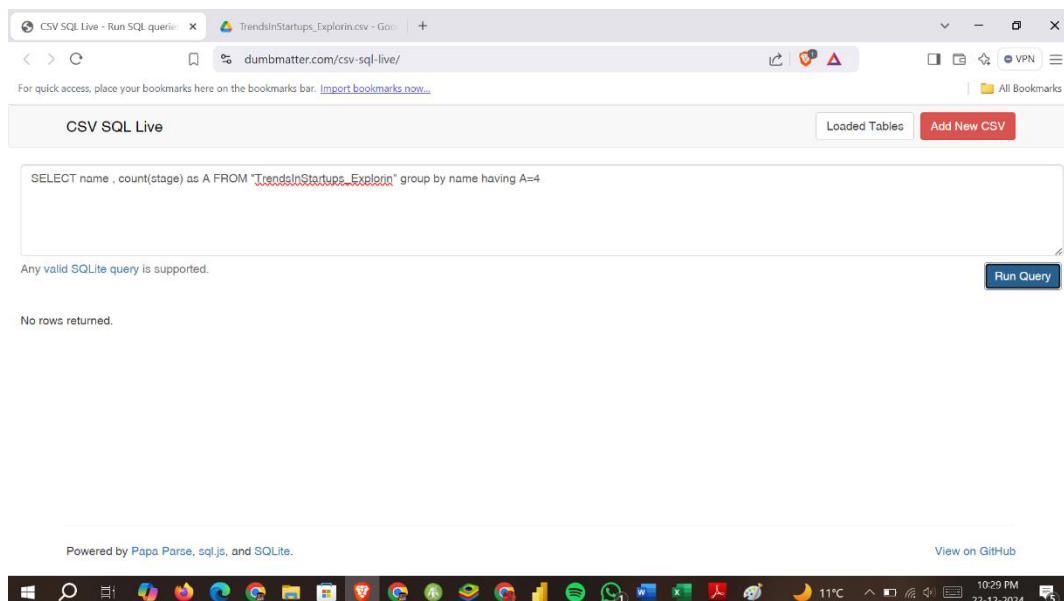


The screenshot shows the CSV SQL Live interface. The query entered is: `SELECT distinct(name) , sum(valuation) as Valued from "TrendsInStartups_Explorin" group by name order by Valued desc limit 15`. The results are displayed in a table with two columns: 'name' and 'Valued'.

name	Valued
Startup 8	5000000
Startup 33	4250000
Startup 19	4000000
Startup 14	3000000
Startup 106	2500000
Startup 146	2500000
Startup 186	2500000
Startup 226	2500000

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



The screenshot shows the CSV SQL Live interface. The query entered is: `SELECT name , count(stage) as A FROM "TrendsInStartups_Explorin" group by name having A=4`. The results section displays "No rows returned."

name	A
------	---

Powered by Papa Parse, sql.js, and SQLite. [View on GitHub](#)

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.csv 3" 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

name	growth_percentage
<input type="text" value="Search"/>	<input type="text" value="Search"/>
Startup 1	-64.28571428571429
Startup 2	-76.66666666666667
Startup 4	-88
Startup 6	-43.75
Startup 7	-92
Startup 9	-50