# Introduction

## Data

The data is about unicorn startup companies that are privately held companies with a valuation of $1 billion or more. The word unicorn was used to express the rarity of such companies as they were once considered rare and mythical. However, over the past decade, there has been a significant increase in the number of unicorn startups, especially in the technology and e-commerce sectors. These companies are often seen as disruptors and game-changers in their respective industries and attract significant attention from investors and the media.

## Analysis

In this portfolio project, I analyzed the mentioned dataset containing information about unicorn companies, including their valuations, industry, country, city, and select investors. I posed insightful questions about the data and used SQL queries to derive meaningful insights. I explored trends in valuations over time, geographical distribution of high-valuation companies, investor patterns, and industry correlations. I leveraged SQL functions like AVG, SPLIT, and UNNEST to perform calculations and transform data for analysis. The project aimed to showcase data analysis skills, from formulating questions to writing SQL queries for complex analyses.

--  cleaning/correcting

  UPDATE `our-bruin-395513.unicorn\_companies.unicorn`

  SET Country = 'united Kingdom' WHERE Country = 'London';

--descriptive analysis

-- companies with highest valuation and their industry

  SELECT [Company] as company\_name, [Valuation\_\_\_B\_] as value , [Industry] as sector

  FROM `our-bruin-395513.unicorn\_companies.unicorn`

  order by Valuation\_\_\_B\_ DESC;

--total number of companies in each country and their total valuation

  SELECT [Country], COUNT(Company) AS Company\_count, SUM(CAST( Valuation\_\_\_B\_ AS int)) as total\_valuation

  from  `our-bruin-395513.unicorn\_companies.unicorn`

  group by Country

  ORDER BY company\_count DESC;

--number of companies joining each year

  SELECT EXTRACT(YEAR FROM (Date\_joined)) as year, count (\*) AS no\_of\_comp

  from `our-bruin-395513.unicorn\_companies.unicorn`

  GROUP BY year

  ORDER BY year;

--Top 5 most common industries in data set

  SELECT [industry], COUNT([Industry]) as num\_of\_industry

  FROM `our-bruin-395513.unicorn\_companies.unicorn`

  GROUP BY Industry

  ORDER BY num\_of\_industry DESC

  LIMIT 5;

--temporal analysis

  SELECT EXTRACT (YEAR from(Date\_joined)) AS year, AVG(Valuation\_\_\_B\_) AS avg\_VALUE

  FROM `our-bruin-395513.unicorn\_companies.unicorn`

  GROUP BY date\_joined

  ORDER BY Date\_Joined

--companies average valuation over time / correlation between joing year and valuation

  SELECT EXTRACT (YEAR from(Date\_joined)) AS years, ROUND(AVG(Valuation\_\_\_B\_),3) AS avg\_VALUE

  FROM `our-bruin-395513.unicorn\_companies.unicorn`

  GROUP BY years

  ORDER BY years

--geographical Analysis

--top 5 countries with highest number of companies

  SELECT Country , count(company) as num\_of\_comp

  FROM `our-bruin-395513.unicorn\_companies.unicorn`

  group by country

  order by num\_of\_comp DESC

  Limit 5

  SELECT Country , AVG(Valuation\_\_\_B\_) as avgvalue\_of\_comp

  FROM `our-bruin-395513.unicorn\_companies.unicorn`

  group by country

  order by num\_of\_comp DESC

  Limit 5

--industry analysis

--highest valuation

  SELECT INDUSTRY, AVG(Valuation\_\_\_B\_) as avg\_value

  from `our-bruin-395513.unicorn\_companies.unicorn`

  GROUP BY Industry

  order by avg\_value DESC

  LIMIT 1;

-- relation between industry and valuation

  SELECT INDUSTRY, AVG(Valuation\_\_\_B\_) as avg\_value

  from `our-bruin-395513.unicorn\_companies.unicorn`

  GROUP BY Industry

  order by avg\_value DESC

--industries attracting more companies

  SELECT Industry, COUNT(Company) as num\_of\_comp

  FROM `our-bruin-395513.unicorn\_companies.unicorn`

  GROUP BY Industry

  ORDER BY NUM\_OF\_COMP DESC;

--company perfomance analysis

SELECT [Country], SUM(CAST( Valuation\_\_\_B\_ AS int)) as total\_valuation

from  `our-bruin-395513.unicorn\_companies.unicorn`

group by Country;

-- TOP 20 Companies achieving highest valuation in least amount of years.

  SELECT Company,Industry , Valuation\_\_\_B\_, DATE\_DIFF(CURRENT\_DATE(), Date\_Joined, year ) AS years\_since\_joining

  FROM `our-bruin-395513.unicorn\_companies.unicorn`

  ORDER BY Valuation\_\_\_B\_ DESC,  years\_since\_joining ASC

  LIMIT 20;

--Corelations

  --corelation in country and sector with most companies

  SELECT Country, Industry, COUNT(COMPANY) AS num\_of\_companies,

  FROM `our-bruin-395513.unicorn\_companies.unicorn`

  GROUP BY Country, Industry

  ORDER BY num\_of\_companies DESC, Country;

  --corelation in country and sector with most companies along with thier avg\_valuation

  SELECT Country, Industry, COUNT(COMPANY) AS num\_of\_companies, ROUND(AVG(Valuation\_\_\_B\_),3) AS AVG\_VALUE\_OF\_SECTOR

  FROM `our-bruin-395513.unicorn\_companies.unicorn`

  GROUP BY Country, Industry

  ORDER BY num\_of\_companies DESC, Country;

  --Corelation in industry/sector and valuation , finding which industry has highest valuation

  SELECT Industry, ROUND(AVG(Valuation\_\_\_B\_),3) AS AVG\_VALUE\_OF\_industry

  FROM `our-bruin-395513.unicorn\_companies.unicorn`

  GROUP BY Industry

  ORDER BY AVG\_VALUE\_OF\_industry  DESC;