

# Project (SQL): Analyzing Unicorn Companies

**# IMPORTING & CLEANING DATA**

**# CASE STUDIES**

*by Jenn Lai*

## Introduction

Did you know that the average return from investing in stocks is 10% per year! But who wants to be average?!

You have been asked to support an investment firm by analyzing trends in high-growth companies. They are interested in understanding which industries are producing the highest valuations and the rate at which new high-value companies are emerging. Providing them with this information gives them a competitive insight as to industry trends and how they should structure their portfolio looking forward.

You have been given access to their unicorns database, which contains the following tables:

dates

Column	Description
company_id	A unique ID for the company.
date_joined	The date that the company became a unicorn.
year_founded	The year that the company was founded.

funding

Column	Description
company_id	A unique ID for the company.
valuation	Company value in US dollars.
funding	The amount of funding raised in US dollars.
select_investors	A list of key investors in the company.

## industries

Column	Description
company_id	A unique ID for the company.
industry	The industry that the company operates in.

## companies

Column	Description
company_id	A unique ID for the company.
company	The name of the company.
city	The city where the company is headquartered.
country	The country where the company is headquartered.
continent	The continent where the company is headquartered.

# 1. Review the dataset

SELECT \* FROM dates

[29]

<div><div>×</div><div>Collapse</div></div>				Rows per page <div>10</div> 1074 rows <div>↓</div>	
	company_id ▾	date_joined ▾		year_founded ▾	
0	189	2017-06-24T00:00:00.000Z		1919	
1	848	2021-06-01T00:00:00.000Z		2019	
2	556	2022-02-15T00:00:00.000Z		2011	
3	999	2021-11-17T00:00:00.000Z		2020	
4	396	2021-10-21T00:00:00.000Z		2021	
5	931	2021-09-24T00:00:00.000Z		2020	
6	364	2019-02-01T00:00:00.000Z		2015	
7	732	2022-01-27T00:00:00.000Z		2017	
8	906	2021-06-10T00:00:00.000Z		2017	
9	72	2015-01-23T00:00:00.000Z		2006	
<div>⏪ &lt; Page 1 of 108 &gt; ⏩</div>				Rows per page <div>10</div> 1074 rows <div>↓</div>	
<div><div>+</div> Add Middleware</div> <div><div>+</div> Add Code</div> <div><div>+</div> Add SQL</div>					

SELECT * FROM funding [30]				
<div> <span>×</span> Collapse         Rows per page <span>10</span> 1074 rows <span>↓</span> </div>				
	company_id ▾	valuation ▾	funding ▾	select_investors ▾
0	189	4000000000	0	EQT Partners
1	848	1000000000	100000000	Dragonfly Captial, Qiming Venture Partners, DST Global
2	556	2000000000	100000000	Blackstone, Bessemer Venture Partners
3	999	1000000000	100000000	Goldman Sachs Asset Management, 3L
4	396	2000000000	100000000	Insight Partners, Softbank Group, Connect Ventures
5	931	1000000000	100000000	Tencent Holdings, Hillhouse Capital Management
6	364	2000000000	101000000	K2VC, Lightspeed China Partners, Sky9 Capital
7	732	1000000000	102000000	Initialized Capital, General Catalyst, Kraken Ventures
8	906	1000000000	105000000	Delian Capital, China International Capital Corporation, Sequoia Capital China
9	72	8000000000	105000000	Accel Partners, Sequoia Capital
<div> <span>&lt;&lt;</span> <span>&lt;</span> Page 1 of 108 <span>&gt;</span> <span>&gt;&gt;</span> Rows per page <span>10</span> 1074 rows <span>↓</span> </div>				

SELECT * FROM industries [31]		
<div> <span>×</span> Collapse         Rows per page <span>10</span> 1074 rows <span>↓</span> </div>		
	company_id ▾	industry ▾
0	189	Health
1	848	Fintech
2	556	Internet software & services
3	999	Internet software & services
4	396	Fintech
5	931	Consumer & retail
6	364	E-commerce & direct-to-consumer
7	732	Fintech
8	906	Hardware
9	72	Hardware
<div> <span>&lt;&lt;</span> <span>&lt;</span> Page 1 of 108 <span>&gt;</span> <span>&gt;&gt;</span> Rows per page <span>10</span> 1074 rows <span>↓</span> </div>		

SELECT * FROM companies [32]					
<div> <span>×</span> Collapse         Rows per page <span>10</span> 1074 rows <span>↓</span> </div>					
	company_id ▾	company ▾	city ▾	country ▾	continent ▾
0	189	Otto Bock HealthCare	Duderstadt	Germany	Europe
1	848	Matrixport		Singapore	Asia
2	556	Cloudinary	Santa Clara	United States	North America
3	999	PLACE	Bellingham	United States	North America
4	396	candy.com	New York	United States	North America
5	931	HAYDON	Shanghai	China	Asia
6	364	eDaili	Shanghai	China	Asia
7	732	CoinTracker	San Francisco	United States	North America
8	906	EcoFlow	Shenzhen	China	Asia
9	72	DJI Innovations	Shenzhen	China	Asia
<div> <span>&lt;&lt;</span> <span>&lt;</span> Page 1 of 108 <span>&gt;</span> <span>&gt;&gt;</span> Rows per page <span>10</span> 1074 rows <span>↓</span> </div>					

## 2. Find out companies becoming unicorn between 2019~2021

```
SELECT
    company_id AS unicorn_company,
    valuation,
    industry,
    date_part('year', date_joined) AS year
FROM
    funding
LEFT JOIN dates USING(company_id)
LEFT JOIN industries USING(company_id)
WHERE
    date_part('year', date_joined) >= 2019
    AND date_part('year', date_joined) <= 2021
    AND valuation >= 1000000000;
```

### Result :

	unicorn_company ▾	valuation ▾	industry ▾	year ▾
0	848	1000000000	Fintech	2021
1	999	1000000000	Internet software & services	2021
2	396	2000000000	Fintech	2021
3	931	1000000000	Consumer & retail	2021
4	364	2000000000	E-commerce & direct-to-consumer	2019
5	906	1000000000	Hardware	2021
6	983	1000000000	Fintech	2019
7	844	1000000000	Fintech	2021
8	894	1000000000	Fintech	2021
9	813	1000000000	Fintech	2021

<< < Page 1 of 74 > >>

Rows per page 10 ▾ 732 rows ▾

### 3. Find out the Top3 industries having the most unicorns

```
WITH unicorn AS (  
    SELECT  
        company_id AS unicorn_company,  
        industry,  
        date_part('year', date_joined) AS year  
    FROM  
        funding  
    LEFT JOIN dates USING(company_id)  
    LEFT JOIN industries USING(company_id)  
    WHERE  
        date_part('year', date_joined) >= 2019  
        AND date_part('year', date_joined) <= 2021  
        AND valuation >= 1000000000  
)  
SELECT  
    industry,  
    count(unicorn_company)  
FROM  
    unicorn  
GROUP BY  
    industry  
ORDER BY  
    count DESC  
LIMIT  
    3;
```

#### Result :

	industry ▾	count ▾
0	Fintech	173
1	Internet software & services	152
2	E-commerce & direct-to-consumer	75

3 rows ▾

The top3 industries having most companies becoming unicorns between 2019~2021 are “Fintech”, “Internet software & services”, and “E-commerce & direct-to-consumer”.

## 4. Trends for the top-performing industries

```
WITH unicorn AS (
    SELECT
        company_id AS unicorn_company,
        valuation,
        industry,
        date_part('year', date_joined) AS year
    FROM
        funding
        LEFT JOIN dates USING(company_id)
        LEFT JOIN industries USING(company_id)
    WHERE
        date_part('year', date_joined) >= 2019
        AND date_part('year', date_joined) <= 2021
        AND valuation >= 1000000000
)
SELECT
    industry,
    year,
    count(unicorn_company) AS num_unicorns,
    ROUND(AVG(valuation) / 1000000000, 2) AS average_valuation_billions
FROM
    unicorn
WHERE
    industry IN (
        'Fintech',
        'Internet software & services',
        'E-commerce & direct-to-consumer'
    )
GROUP BY
    industry,
    year
ORDER BY
    industry,
    year DESC;
```

### Result :

	industry ▾	year ▾	num_unicorns ▾	average_valuation_billions ▾
0	E-commerce & direct-to-consumer	2021	47	2.47
1	E-commerce & direct-to-consumer	2020	16	4
2	E-commerce & direct-to-consumer	2019	12	2.58
3	Fintech	2021	138	2.75
4	Fintech	2020	15	4.33
5	Fintech	2019	20	6.8
6	Internet software & services	2021	119	2.15
7	Internet software & services	2020	20	4.35
8	Internet software & services	2019	13	4.23