

# Respondendo 5 perguntas de negócio com SQL

Vinicius Olivedo

## Contexto

Ao longo dos últimos meses, tenho prestado serviços como *freelancer* no regime de **home-office**. Ao longo da jornada, foram completos centenas de projetos para clientes ao redor do mundo. No entanto, para aumentar as receitas e entender melhor o negócio, os dados de cada projeto passaram a ser devidamente registrados. Isso permitiu aplicar análise de dados voltada à tomada de decisões, bem como para a obtenção de insights.

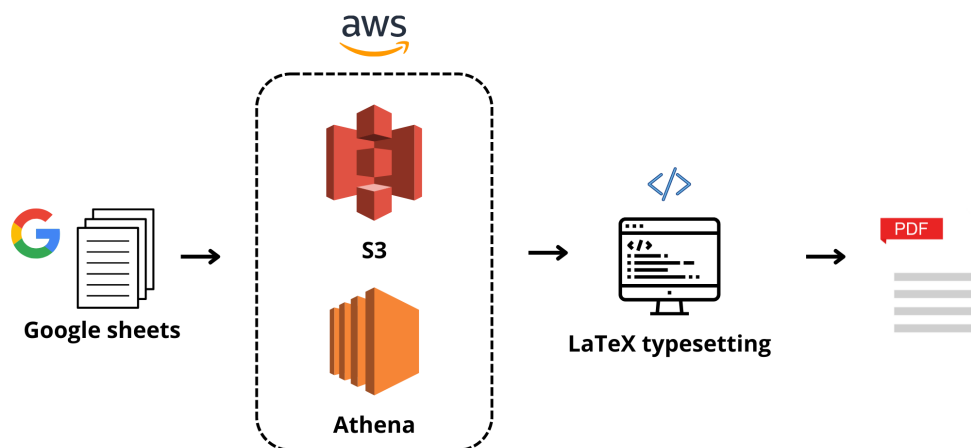
Neste documento, mostro um pouco destas análises realizadas na nuvem da *Amazon Web Service* (AWS) e por meio de consultas SQL. O dataset compreendeu 144 projetos registrados.

## Definições

**Descrição do problema:** busca-se saber a resposta para 5 perguntas acerca dos negócios *freelancer* ao longo de 2020-2022 (atual):

1. Qual foi a receita líquida por ano?
2. Qual foi o número de projetos completos em cada ano?
3. Qual é a base de clientes por país?
4. Qual é a média de valor-hora (em U\$D) por fonte?
5. Qual é a fonte com mais projetos em cada ano?

## Pipeline



## 🔖 Queries

### </> Receita líquida por ano

```
1 SELECT
2     year,
3     SUM(amount_earned) as total_earned
4 FROM
5     projects
6 GROUP BY
7     year;
```

#	year	total_earned
1	2021	5262.530
2	2022	6320.900
3	2020	2686.790

### </> Projetos completos por ano

```
1 SELECT
2     year,
3     COUNT(project_id) as
4     projects_completed
5 FROM
6     projects
7 GROUP BY year;
```

#	year	projects_completed
1	2020	26
2	2021	71
3	2022	47

### </> Clientes por país:

```
1 SELECT
2     country,
3     COUNT(country) as customers
4 FROM projects
5 GROUP BY country
6 ORDER by customers DESC;
```

#	country	customers
1	USA	60
2	Brazil	20
3	UK	16
4	Netherlands	12
5	Germany	8
6	Saudi Arabia	6
7	India	3
8	Spain	3
9	Denmark	2
10	Greece	2
11	Canada	2
12	Finland	1
13	Philippines	1
14	Colombia	1
15	Ireland	1
16	France	1
17	Egypt	1
18	Korea	1
19	Thailand	1
20	Italy	1
21	Belgium	1

#### </> **Valor-hora médio por fonte:**

```
1 SELECT
2     source ,
3     AVG(hourly_rate) AS mean_hr
4 FROM projects
5 GROUP BY source, year
6 HAVING year = '2022';
```

#	source	mean_hr
1	Direct	40.29
2	Workan	34.27
3	Fiverr	23.98
4	Upwork	16.88

**</> Fonte com mais projetos completos em cada ano:**

```
1 SELECT
2     year,
3     COUNT(project_id) AS projects,
4     source
5 FROM
6     projects
7 GROUP BY
8     source, year
9 ORDER BY
10    1, 3 DESC;
```

#	year	projects	source
1	2020	6	Study.com
2	2020	17	Fiverr
3	2020	3	Direct
4	2021	17	Study.com
5	2021	48	Fiverr
6	2021	6	Direct
7	2022	1	Workana
8	2022	1	Upwork
9	2022	33	Fiverr
10	2022	12	Direct