

## Prova técnica consultas SQL

Data 17/06/2022

**Klayton Crul Correa**

E-mail: [klayton12341@live.com](mailto:klayton12341@live.com)

Linkedin: <https://www.linkedin.com/in/klayton-crul>

Github: <https://github.com/klaytoncrul/data-science-projects>

Telefone: (45)998023698

### **Objetivo:**

Estágio em Ciência de dados

### **Dados utilizados:**

tlc\_yellow\_trips\_2018

### **Fonte dos dados:**

[https://docs.google.com/spreadsheets/d/18V9jsGMn1KSb7vsEN4Fh0CQNIKweE5ctFrmGDEeWG\\_k/edit#gid=460814509](https://docs.google.com/spreadsheets/d/18V9jsGMn1KSb7vsEN4Fh0CQNIKweE5ctFrmGDEeWG_k/edit#gid=460814509)

### **Exercício proposto:**

Utilizando a documentação de SQL do Google BigQuery e baseando no schema da tabela acima e nos dados presentes na aba "TABELA" da planilha fornecida, crie as consultas SQL correspondentes às seguintes questões:

Perguntas:

- 1) Qual foi a receita de cada tipo de pagamento no dia 15 de Março de 2018?
- 2) Considere que corridas de táxi válidas tenham de 1 a 5 passageiros. Qual a quantidade de corridas feitas com cada número de passageiros?
- 3) Considerando apenas as corridas que houveram pedágios (tolls), qual a média do valor pago em pedágios por corrida?
- 4) Qual a hora que mais começaram corridas?

## 1) Qual foi a receita de cada tipo de pagamento no dia 15 de Março de 2018?

**total\_amount\_receita** foi de **543,81**

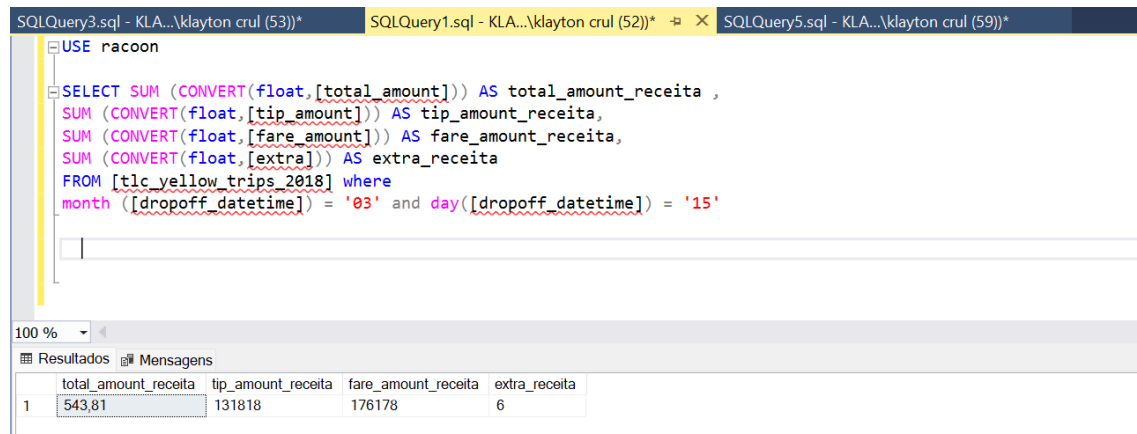
**tip\_amount\_receita** foi de **131818**

**fare\_amount\_receita** foi de **176818**

**extra\_receita** foi de **6**

USE racoon

```
SELECT SUM (CONVERT(float,[total_amount])) AS total_amount_receita ,  
SUM (CONVERT(float,[tip_amount])) AS tip_amount_receita,  
SUM (CONVERT(float,[fare_amount])) AS fare_amount_receita,  
SUM (CONVERT(float,[extra])) AS extra_receita  
FROM [tlc_yellow_trips_2018] where  
month ([dropoff_datetime]) = '03' and day([dropoff_datetime]) = '15'
```



The screenshot shows a SQL Server Enterprise Manager window with three tabs: 'SQLQuery3.sql - KLA...\\klayton crul (53))\*', 'SQLQuery1.sql - KLA...\\klayton crul (52))\*', and 'SQLQuery5.sql - KLA...\\klayton crul (59))\*'. The active tab is 'SQLQuery5.sql'. The query editor shows the same SQL query as above. Below the editor, the 'Resultados' (Results) pane is visible, showing a table with the following data:

	total_amount_receita	tip_amount_receita	fare_amount_receita	extra_receita
1	543,81	131818	176178	6

## 2) Considere que corridas de táxi válidas tenham de 1 a 5 passageiros. Qual a quantidade de corridas feitas com cada número de passageiros?

Quantidade de corridas por passageiro:

1 passageiro = **1590**

2 passageiros = **795**

3 passageiros = **530**

4 passageiros = **397,5**

5 passageiros = **318**

### Código:

USE racoon

```
SELECT SUM ([passenger_count])/1 AS passenger_count_1 ,  
SUM ([passenger_count])/2 AS passenger_count_2,  
SUM ([passenger_count])/3 AS passenger_count_3,
```

```
SUM ([passenger_count])/4 AS passenger_count_4,
SUM ([passenger_count])/5 AS passenger_count_5
FROM [tlc_yellow_trips_2018]
```

The screenshot shows a SQL Server Enterprise Manager interface. The top pane displays a query in 'SQLQuery2.sql' with the following code:

```
USE racoon
SELECT SUM ([passenger_count])/1 AS passenger_count_1 ,
SUM ([passenger_count])/2 AS passenger_count_2,
SUM ([passenger_count])/3 AS passenger_count_3,
SUM ([passenger_count])/4 AS passenger_count_4,
SUM ([passenger_count])/5 AS passenger_count_5
FROM [tlc_yellow_trips_2018]
```

The bottom pane shows the 'Resultados' (Results) tab with a single row of data:

	passenger_count_1	passenger_count_2	passenger_count_3	passenger_count_4	passenger_count_5
1	1590	795	530	397,5	318

**3) Considerando apenas as corridas que houveram pedágios (tolls), qual a média do valor pago em pedágios por corrida?**

A média de valor pago em pedágio por corridas é de **15,0748106333343**

**Código:**

```
use racoon
SELECT avg([tolls_amount])/SUM (CONVERT(float,[total_amount])) as
tolls_amount_average
FROM [tlc_yellow_trips_2018] where [tolls_amount] > '0'
```

The screenshot shows a SQL Server Enterprise Manager interface. The top pane displays a query in 'SQLQuery4.sql' with the following code:

```
use racoon
SELECT avg([tolls_amount])/SUM (CONVERT(float,[total_amount])) as tolls_amount_average
FROM [tlc_yellow_trips_2018] where [tolls_amount] > '0'
```

The bottom pane shows the 'Resultados' (Results) tab with a single row of data:

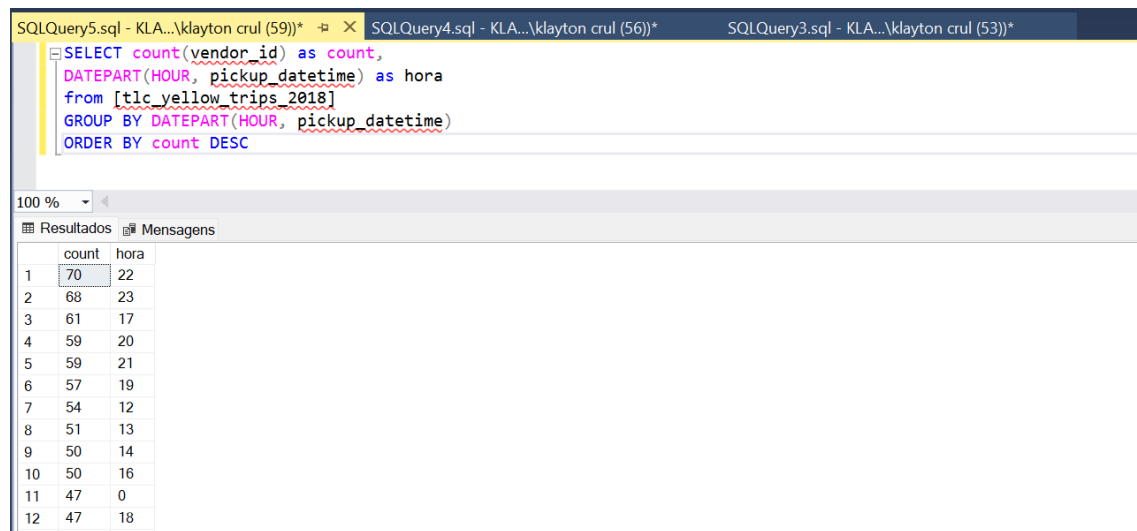
	tolls_amount_average
1	15,0748106333343

**Qual a hora que mais começaram corridas?**

De acordo com os dados o horário em que acontecem mais corridas é as **22** horas com **70** corridas.

### Código:

```
SELECT count(vendor_id) as count,  
DATEPART(HOUR, pickup_datetime) as hora  
from [tlc_yellow_trips_2018]  
GROUP BY DATEPART(HOUR, pickup_datetime)  
ORDER BY count DESC
```



The screenshot shows a SQL Server Enterprise Manager interface. At the top, there are three tabs: 'SQLQuery5.sql - KLA...\klayton crul (59))\*', 'SQLQuery4.sql - KLA...\klayton crul (56))\*', and 'SQLQuery3.sql - KLA...\klayton crul (53))\*'. The active tab is 'SQLQuery5.sql'. Below the tabs, the SQL query is displayed in a text editor. The query is:   
`SELECT count(vendor_id) as count,  
DATEPART(HOUR, pickup_datetime) as hora  
from [tlc_yellow_trips_2018]  
GROUP BY DATEPART(HOUR, pickup_datetime)  
ORDER BY count DESC`  
Below the query editor, there is a '100 %' zoom level indicator and a 'Resultados' (Results) tab. The 'Resultados' tab is active, showing a table with two columns: 'count' and 'hora'. The table contains 12 rows of data, with the first row having a count of 70 and a hora of 22. The results are ordered by count in descending order.

	count	hora
1	70	22
2	68	23
3	61	17
4	59	20
5	59	21
6	57	19
7	54	12
8	51	13
9	50	14
10	50	16
11	47	0
12	47	18