



# Formação Engenheiro de Dados

Analítico

# Criar Database no Hive

```
beeline
!connect jdbc:hive2://
create database locadora;
use locadora;
```

# Ingestão para SQOOP

```
sqoop import --connect
jdbc:postgresql://127.0.0.1/locadora --
username postgres --password 123456 --
table clientes --hive-import --create-
hive-table --hive-table
LOCADORA.CLIENTES;
```

# Quais veículos foram locados em determinados períodos?

```
CREATE VIEW locperiodo
```

```
AS
```

```
select veic.modelo as Modelo,  
MONTH(loc.datalocacao) as Mes,  
YEAR(loc.datalocacao) as Ano, COUNT(*)  
as Total from locacao loc  
  
join veiculos veic on (loc.idveiculo =  
veic.idveiculo) GROUP BY veic.modelo ,  
MONTH(loc.datalocacao) ,  
YEAR(loc.datalocacao) order by Total  
desc;
```

```
Select * from locperiodo;
```

locperiodo.modelo	locperiodo.mes	locperiodo.ano	locperiodo.total
BMW 750i xDrive	6	2019	10
BMW 750i xDrive	5	2019	10
Volvo XC90 T8 Hybrid	4	2019	8
Volvo XC90 T8 Hybrid	6	2019	7
BMW 750i xDrive	3	2019	5
Infiniti Q50S 3.7 Sedan	4	2019	5
Infiniti Q50S 3.7 Sedan	3	2019	5
Mercedes-Benz S65 AMG Coupe	5	2019	5
Volvo XC90 T8 Hybrid	3	2019	4
Volvo XC90 T8 Hybrid	2	2019	4
Volvo XC90 T8 Hybrid	5	2019	4
Infiniti Q50S 3.7 Sedan	5	2019	4
Mercedes-Benz S65 AMG Coupe	3	2019	4
Tesla Model S P90D	2	2019	3
Tesla Model S P90D	5	2019	3
Mercedes-Benz S65 AMG Coupe	4	2019	2
Tesla Model S P90D	6	2019	2
Tesla Model S P90D	4	2019	2
Tesla Model S P90D	3	2019	2
Mercedes-Benz S65 AMG Coupe	6	2019	2
Infiniti Q50S 3.7 Sedan	6	2019	2
Infiniti Q50S 3.7 Sedan	2	2019	2
BMW 750i xDrive	4	2019	2
BMW 750i xDrive	2	2019	2
Mercedes-Benz S65 AMG Coupe	2	2019	1

# Quais despachantes locaram quais veículos?

```

CREATE VIEW locdespachante
AS
select desp.nome nome, veic.modelo as Modelo,
MONTH(loc.datalocacao) as Mes, YEAR(loc.datalocacao)
as Ano, COUNT(*) as Total from locacao loc
join veiculos veic on (loc.idveiculo = veic.idveiculo)
join despachantes desp on (loc.iddespachante =
desp.iddespachante)
GROUP BY desp.nome , veic.modelo,
MONTH(loc.datalocacao) , YEAR(loc.datalocacao) order
by Total desc;

```

Select \* from locdespachante;

locdespachante.nome	locdespachante.modelo	locdespachante.mes	locdespachante.ano	locdespachante.total
Carminda Pestana	BMW 750i xDrive	6	2019	5
Matilde Rebouças	Volvo XC90 T8 Hybrid	4	2019	3
Graça Ornelas	Infiniti Q50S 3.7 Sedan	3	2019	2
Viviana Sequeira	Tesla Model S P90D	2	2019	2
Carminda Pestana	Infiniti Q50S 3.7 Sedan	4	2019	2
Graça Ornelas	BMW 750i xDrive	6	2019	2
Emídio Dornelles	Tesla Model S P90D	5	2019	2
Roque Vásquez	BMW 750i xDrive	5	2019	2
Emídio Dornelles	BMW 750i xDrive	3	2019	2
Graça Ornelas	BMW 750i xDrive	4	2019	2
Graça Ornelas	BMW 750i xDrive	3	2019	2
Felisabela Dornelles	Volvo XC90 T8 Hybrid	6	2019	2
Uriel Queiroz	BMW 750i xDrive	5	2019	2
Viviana Sequeira	Volvo XC90 T8 Hybrid	6	2019	2
Graça Ornelas	Volvo XC90 T8 Hybrid	2	2019	2
Graça Ornelas	Mercedes-Benz S65 AMG Coupe	5	2019	2
Noémia Orrica	Infiniti Q50S 3.7 Sedan	5	2019	2
Carminda Pestana	Volvo XC90 T8 Hybrid	4	2019	2
Graça Ornelas	Volvo XC90 T8 Hybrid	3	2019	1
Viviana Sequeira	Volvo XC90 T8 Hybrid	4	2019	1
Viviana Sequeira	Volvo XC90 T8 Hybrid	3	2019	1
Viviana Sequeira	Volvo XC90 T8 Hybrid	2	2019	1
Viviana Sequeira	Tesla Model S P90D	5	2019	1
Viviana Sequeira	Mercedes-Benz S65 AMG Coupe	2	2019	1
Viviana Sequeira	Infiniti Q50S 3.7 Sedan	3	2019	1
Viviana Sequeira	BMW 750i xDrive	5	2019	1
Viviana Sequeira	BMW 750i xDrive	2	2019	1
Uriel Queiroz	Volvo XC90 T8 Hybrid	5	2019	1
Uriel Queiroz	Tesla Model S P90D	6	2019	1
Uriel Queiroz	Tesla Model S P90D	2	2019	1

# Qual o faturamento por período?

```
CREATE VIEW locfaturamento
AS
select MONTH(loc.datalocacao) as Mes,
YEAR(loc.datalocacao) as Ano , SUM(loc.total) as
Total from locacao loc
GROUP BY MONTH(loc.datalocacao) ,
YEAR(loc.datalocacao) order by mes, ano;
```

```
Select * from locfaturamento;
```

locfaturamento.mes	locfaturamento.ano	locfaturamento.total
2	2019	23852.0
3	2019	39239.0
4	2019	36407.0
5	2019	50498.0
6	2019	44530.0

# Quais clientes locaram quais veículos?

```
CREATE VIEW locclientes  
AS  
  
select cli.nome, MONTH(loc.datalocacao) as Mes,  
YEAR(loc.datalocacao) as Ano , count(*) as  
Total from locacao loc  
  
join clientes cli on (cli.idcliente =  
loc.idcliente)  
  
GROUP BY cli.nome,MONTH(loc.datalocacao) ,  
YEAR(loc.datalocacao) order by cli.nome;
```

```
Select * from locclientes;
```

locclientes.nome	locclientes.mes	locclientes.ano	locclientes.total
Alvito Espargosa	6	2019	1
Alvito Espargosa	3	2019	3
Blasco Canto	6	2019	1
Carminda Veríssimo	2	2019	1
Carminda Veríssimo	3	2019	1
Carminda Veríssimo	5	2019	1
Cleusa Lamego	4	2019	2
Cleusa Lamego	5	2019	1
Clóvis Carrasco	5	2019	1
Clóvis Carrasco	6	2019	3
Cátia Fróis	4	2019	2
Cátia Fróis	5	2019	1
Catia Frois	6	2019	2
Emílio Faro	4	2019	2
Emílio Faro	3	2019	1
Emílio Faro	5	2019	2
Emílio Faro	6	2019	1
Hipólito Granja	2	2019	1
Hipólito Granja	5	2019	2
Iara Cardoso	2	2019	1
Iara Cardoso	3	2019	2
Iara Cardoso	4	2019	2
Iara Cardoso	6	2019	1
Isaura Farias	2	2019	1
Iuri Alancastre	2	2019	1
Iuri Alancastre	3	2019	3