



Processo Seletivo

Data Science

Projeto Data Analysis - SQL

- Esta é uma provinha de SQL baseada no modelo de consultas do [Google Big Query](#).
- Utilize o schema da tabela abaixo para desenvolver os SELECTs.
- Queremos ver se você consegue elaborar a estrutura do SELECT (não se preocupe com erros de formatação ou de sintaxe **específicos** do Big Query, mas sim, a sintaxe e estrutura de SQL).
- **Não é** necessário entrar no Google BigQuery e executar as consultas, apenas faça a query que achar correta.

Sugestões de estudo:

- [Estudo - SQL](#)

Importante: Deixar indicação na prova de qual SQL foi utilizado para responder às questões. Ex: Standard SQL no BigQuery.



Considere a [seguinte tabela](#) sobre corridas de táxi em Nova Iorque:

Nome do campo	Tipo	Descrição
vendor_id	STRING	A code indicating the TPEP provider that provided the record. 1= Creative Mobile Technologies, LLC; 2= VeriFone Inc
pickup_datetime	DATETIME	The date and time when the meter was engaged.
dropoff_datetime	DATETIME	The date and time when the meter was disengaged.
passenger_count	INTEGER	The number of passengers in the vehicle. This is a driver-entered value
trip_distance	NUMERIC	The elapsed trip distance in miles reported by the taximeter.
rate_code	STRING	The final rate code in effect at the end of the trip. 1= Standard rate 2=JFK 3=Newark 4=Nassau or Westchester 5=Negotiated fare 6=Group ride
store_and_fwd_flag	STRING	This flag indicates whether the trip record was held in vehicle memory before sending to the vendor, aka "store and forward," because the vehicle did not have a connection to the server. Y= store and forward trip N= not a store and forward trip
payment_type	STRING	A numeric code signifying how the passenger paid for the trip. 1= Credit card 2= Cash 3= No charge 4= Dispute 5= Unknown 6= Voided trip
fare_amount	NUMERIC	The time-and-distance fare calculated by the meter
extra	NUMERIC	Miscellaneous extras and surcharges. Currently, this only includes the \$0.50 and \$1 rush hour and overnight charges.
mta_tax	NUMERIC	\$0.50 MTA tax that is automatically triggered based on the metered rate in use
tip_amount	NUMERIC	Tip amount – This field is automatically populated for credit card tips. Cash tips are not included
tolls_amount	NUMERIC	Total amount of all tolls paid in trip.
imp_surcharge	NUMERIC	\$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015.
total_amount	NUMERIC	The total amount charged to passengers. Does not include cash tips
pickup_location_id	STRING	TLC Taxi Zone in which the taximeter was engaged
dropoff_location_id	STRING	TLC Taxi Zone in which the taximeter was disengaged



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?

