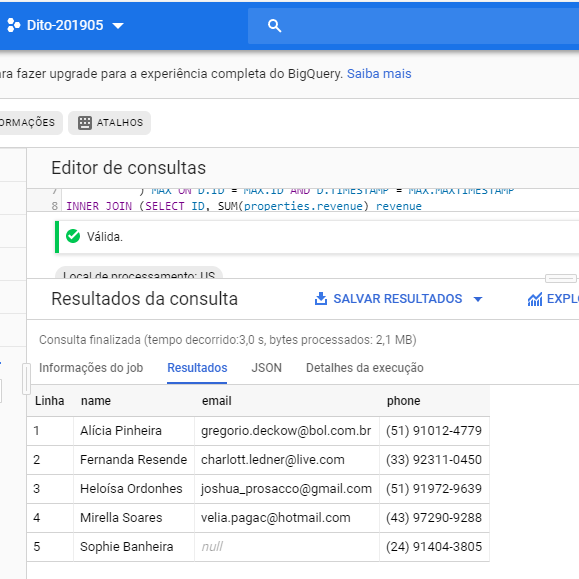
**Desafio Dito**

**01 - SQL**

1. Qual o nome, e-mail e telefone das 5 pessoas que mais geraram receita?
   1. Resultado:



* 1. Lógica:
  2. Selecionar o último registro válido de cada cliente
  3. Agrupamento dos clientes que mais geraram receita. Utilizei a query disponibilizada
  4. Junção a partir do ID da query de receita para buscar na identify os dados dos cliente.
  5. Query:

SELECT traits.name, traits.email, traits.phone

FROM `dito-data-scientist-challenge.tracking.dito` D

INNER JOIN (SELECT MAX(TIMESTAMP) MAXTIMESTAMP, ID

FROM `dito-data-scientist-challenge.tracking.dito`

WHERE type = 'identify' AND traits.phone IS NOT NULL

GROUP BY ID

) MAX ON D.ID = MAX.ID AND D.TIMESTAMP = MAX.MAXTIMESTAMP

INNER JOIN (SELECT ID, SUM(properties.revenue) revenue

FROM `dito-data-scientist-challenge.tracking.dito`

WHERE type = 'track'

GROUP BY id

ORDER BY revenue DESC

LIMIT 5

)MAX\_RECEITA ON D.ID = MAX\_RECEITA.ID

WHERE type = 'identify'

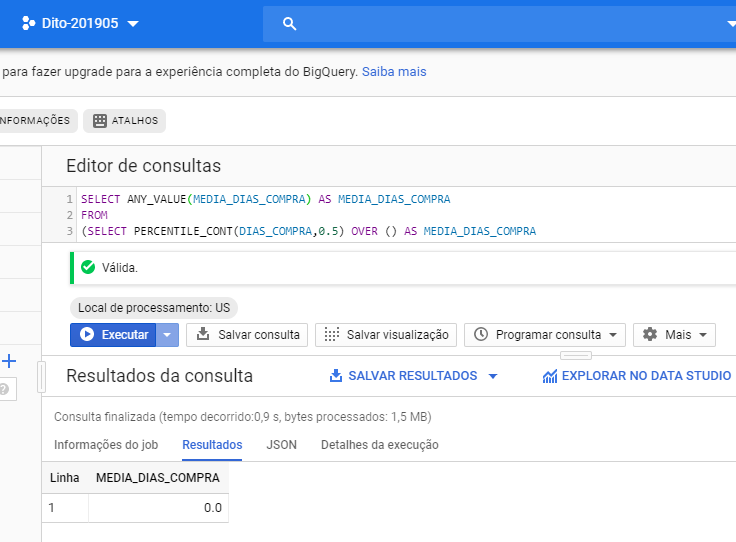
order by 1

1. De quantos em quantos dias, em média, as pessoas compram? Use a mediana como

média.

1. Resultado:

O resultado retornou com a mediana em zero, pois a maioria das compras são realizadas no mesmo dia, diferindo em horas. Ou seja, em média os clientes estão realizando compras no mesmo dia.



1. Lógica:
2. Selecionar por ID as datas de compra;
3. Criar uma coluna a partir da TIMESTAMP para possibilitar a subtração entre as datas;
4. Realizar a subtração da data anterior pela data atual para saber de quantos em quantos dias é realizada a compra;
5. Verificar qual é o percentil50, pois é onde se encontra a mediana;
6. Query:

SELECT

ANY\_VALUE(MEDIA\_DIAS\_COMPRA) AS MEDIA\_DIAS\_COMPRA

FROM (

SELECT

PERCENTILE\_CONT(DIAS\_COMPRA,

0.5) OVER () AS MEDIA\_DIAS\_COMPRA

FROM (

SELECT

ID,

DATE\_DIFF(DATE(NEW\_TIMESTAMP2),DATE(TIMESTAMP), DAY) AS DIAS\_COMPRA

FROM (

SELECT

TIMESTAMP,

ID,

LEAD(TIMESTAMP) OVER (PARTITION BY ID ORDER BY ID ASC) AS NEW\_TIMESTAMP2

FROM (

SELECT

TIMESTAMP,

ID

FROM

`dito-data-scientist-challenge.tracking.dito`

WHERE

TYPE = 'track'

ORDER BY

ID,

TIMESTAMP ))))