Exercício 1

A )

select traits.name, traits.email, traits.phone

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

( select id i, max(timestamp) k from `dito-data-scientist-challenge.tracking.dito` where traits.name is not null and traits.email is not null and traits.phone is not null

group by id ) T on i = id and k = timestamp

where id in (

select X from ( SELECT id X,

SUM(properties.revenue) revenue

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

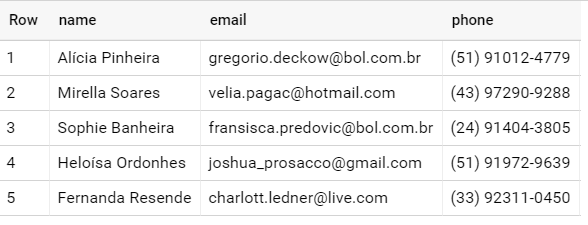
WHERE type = 'track'

GROUP BY id

ORDER BY revenue DESC

LIMIT 5 ) T1 )

Explicação: primeiro, eu selecionei os 5 Ids com maiores rentas, depois eu selecionei a maior timestamp em que os campos de nome, e-mail e telefone não estivesse como nulos. Logo em seguida, eu uni( join ) o timestamp com os Ids para que não houvessem registros repetidos com o mesmo id



B)

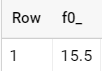
select ( 1 + max - min ) / 2 from (

select extract ( dayofyear from ( select min(timestamp)

from `dito-data-scientist-challenge.tracking.dito`) at time zone 'America/Los\_Angeles') Min,

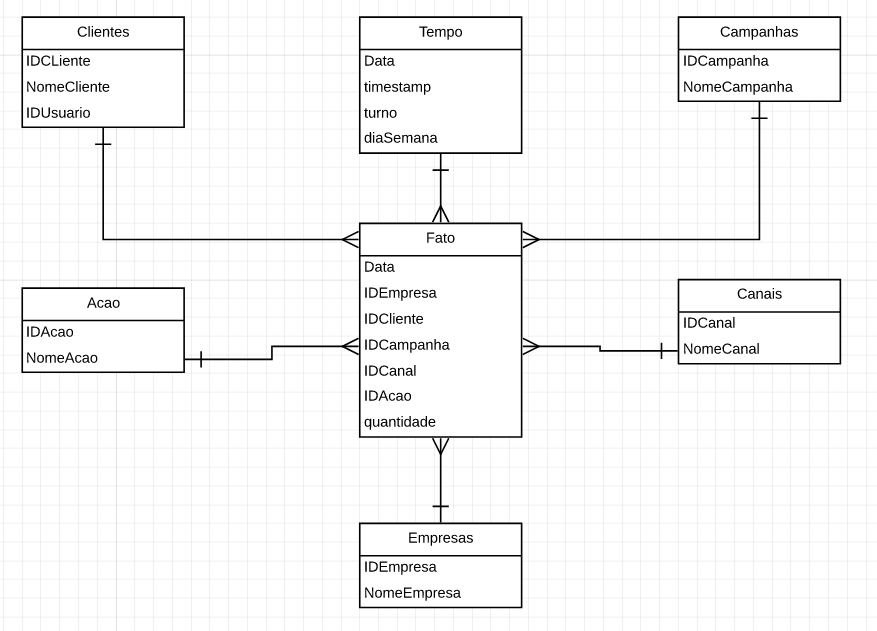
extract ( dayofyear from ( select max(timestamp)

from `dito-data-scientist-challenge.tracking.dito`) at time zone 'America/Los\_Angeles') Max ) T



Explicação: Eu peguei a menor e maior data e transformei em dias. Logo em seguida, eu subtrair da soma o mínimo para descontar a quantidade de dias que não foi utilizada. Dividindo por 2, encontra-se a mediana

Exercício 2



1. Pode-se verificar qual campanha obteve maior alcance diante do número de pessoas que visualizaram a campanha em algum dos canais
2. Pode-se verificar qual canal cada cliente prefere utilizar para saber em qual canal a campanha terá preferência para aquele tipo de cliente
3. Pode-se verificar qual o turno( manhã, tarde, noite ) e dia da semana a campanha obteve mais efeito

Exercício 3

01/2005 08/2005 03/2006 01/2007 08/2007 12/2010

08/2012 12/2013 01/2014 09/2014 11/2014 02/2015 01/2018