

Thais Helena Dias

*Desafio Técnico*

1. Metodologia de trabalho

Inicialmente as tabelas foram visualizadas no Python, para limpeza e tratamento dos dados. Nessa etapa os dados foram explorados a fim de detectar valores nulos e duplicatas e corrigir algumas discrepâncias.

Em seguida com as tabelas ‘limpas’ foi feita a conexão com o banco de dados MySQL para que algumas consultas iniciais fossem realizadas com o intuito de se extrair alguns insights do negocio e por fim o banco de dados foi exportado para o PowerBI para que algumas visualizações fossem feitas a partir de um dashboard.

Com as conexões dos programas citados acima, a atualização dos dados pode ser feita de maneira automatizada, então qualquer dado que precise ser corrigido no Python por exemplo, é atualizado no MySQL ou PowerBI.

1. Dashboard e insights
2. Consultas e modificação na estrutura de dados (MySQL)

O primeiro passo no MySQL foi checar o relacionamento entre as tabelas e os tipos de dados utilizados na construção do banco de dados.

Algumas estruturas de tabela foram alteradas, por exemplo: colunas do tipo ‘data’ que estavam como ‘text’ foram modificadas, entre outros, que podem ser verificados a seguir no script completo, assim como algumas consultas iniciais foram feitas para responder questões relevantes a partir do dataset.

/\* alterando tipo de dados das tabelas\*/

/\*tabela clients\*/

alter table clients

change user\_id user\_id varchar(150) Primary key not null;

alter table clients

change sign\_up\_date sign\_up\_date datetime;

/\*tabela professionals \*/

alter table professionals

change professional\_user\_id professional\_user\_id varchar(150) primary key not null;

alter table professionals

change sign\_up\_date sign\_up\_date datetime;

alter table professionals

change professional\_type professional\_type varchar(150);

alter table professionals

change gender gender enum('f','m','nb');

/\*tabela events\*/

alter table events

change session\_id session\_id varchar(150) primary key not null;

alter table events

change session\_date session\_date datetime;

alter table events

change client\_user\_id client\_user\_id varchar(150);

alter table events

change professional\_user\_id professional\_user\_id varchar(150);

/\*tabela ratings\*/

alter table ratings

change client\_user\_id client\_user\_id varchar(150);

alter table ratings

change professional\_user\_id professional\_user\_id varchar(150);

/\* consultas SQL\*/

/\* total de usuarios cadastrados\*/

select count(distinct user\_id) as Usuarios

from clients;

/\*quantas vezes cada usuario usou o a sessao de terapia\*/

select distinct(count(c.user\_id)) as usuarios, count(e. session\_id) as sessoes

from clients c

join events e

on c.user\_id = e.client\_user\_id;

/\*quantidade de vezes que cada profissional fez atendimento\*/

select professional\_user\_id, count(professional\_user\_id) as qtde\_atendimentos

from events

group by professional\_user\_id

order by count(professional\_user\_id) desc;

/\*quantos profissionais diferentes estao cadastrados\*/

select count(distinct(professional\_user\_id))

from professionals;

/\* data do primeiro e ultimo cadastro de clientes\*/

select max(sign\_up\_date) as ultimo\_cadastro, min(sign\_up\_date) as primeiro\_cadastro

from clients;

/\* numero de avaliaçoes\*/

select count(rating\_id), value

from ratings;

/\*numero de avaliaçoes por nota\*/

SELECT DISTINCT

(value) AS valor, COUNT(rating\_id) AS nota\_avaliaçao

FROM

ratings

GROUP BY value;

/\*quantas avaliaçoes estao visiveis e tem nota menor que 5\*/

select count(rating\_id)

from ratings

where value <> '5' and is\_visible like '%1';

/\* especialidade mais cara e mais barata\*/ /\* preço medio dos profisisonais\*/

select avg(price) as media\_de\_preço, professional\_type as especialidade

from professionals

group by professional\_type

order by avg(price) Desc;

/\*quantos profissionais na categoria gold, diamond, silver\*/

select count(\*)

from professionals

where is\_gold like'%1';

select count(\*)

from professionals

where is\_diamond like'%1';

select count(\*)

from professionals

where is\_silver like'%1';

/\*distinçao de genero dos profissionais\*/

select count(gender)

from professionals

where gender like'f';

select count(gender)

from professionals

where gender like'm';

select count(gender)

from professionals

where gender like'nb';

/\* data do ultimo cadastro de profissional\*/

select max(sign\_up\_date)

from professionals;

/\*quantos profissionais aceitam meia hora de consulta\*/

select count(professional\_user\_id)

from professionals

where is\_accepting\_half\_hour like '%1';

select count(professional\_user\_id)

from professionals

where is\_accepting\_half\_hour not like '%1';

/\*principais razoes de contato\*/

select count(contact\_reasons)

from events;

/\* ansiedade\*/

select count(contact\_reasons)

from events

where contact\_reasons like '%anxi%';

/\*auto conhecimento\*/

select count(contact\_reasons)

from events

where contact\_reasons like '%self%';

/\*relacionamentos\*/

select count(contact\_reasons)

from events

where contact\_reasons like '%relationship%';

/\*conflitos familiares\*/

select count(contact\_reasons)

from events

where contact\_reasons like '%family%';

/\*relaionados a carreira\*/

select count(contact\_reasons)

from events

where contact\_reasons like '%career%';

/\*relacionados a mudanças de cidade, estado etc...\*/

select count(contact\_reasons)

from events

where contact\_reasons like '%change%';

/\* quais profissionais tiveram avaliaçao menor que 4/ qual o genero e o tipo de especialidade desses profissionais\*/

select count(rating\_id), gender, professional\_type

from professionals p

join ratings r

on p.professional\_user\_id = r.professional\_user\_id

where r.value <4

group by gender;

/\* diferença de preço com profissionais de avaliacao 5 e os outros\*/

select count(r.rating\_id), p.price, p.professional\_type, p.professional\_user\_id

from professionals p

join ratings r

on p.professional\_user\_id = r.professional\_user\_id

where r.value >4

group by professional\_user\_id

order by price desc;

/\* tipo de categorias(gold, silver, diamond) e avaliaçoes\*/

select count(p.is\_gold), value

from professionals p

join ratings r

on p.professional\_user\_id = r.professional\_user\_id

where p.is\_gold like'%1%' and value ='5';

select count(p.is\_silver), value

from professionals p

join ratings r

on p.professional\_user\_id = r.professional\_user\_id

where p.is\_silver like'%1%' and value ='5';

select count(p.is\_diamond), value

from professionals p

join ratings r

on p.professional\_user\_id = r.professional\_user\_id

where p.is\_diamond like'%1%' and value ='5';