

Nome: Guilherme Guimarães Vieira Lourenço da Silva

Parte 1 – SQL/Consultas

* Além do resultado, informe a query/consulta utilizada para alcançá-lo.

1) Selecione os Nome e CPF de todos usuários que possuem um CPF que inicie pelos dígitos 55, ordene pelo nome em ordem alfabética.	
Resultado	name cpf Clare 556014 Kay 558803 Mia 558758
SUA QUERY	select name, cpf from users where cpf like '55%' order by name;

2) Faça uma lista que contenha: Id, Nome, CPF e e-mail de todos usuários que vivem no estado do Rio de Janeiro;	
Resultado	id name cpf email 31 Quon 958558 dignissim.Maecenas.ornare@eleifend.org 35 Eleanor NULL 36 Stephanie 908679 NULL 61 Bianca 572735 NULL 70 Sawyer 649277 NULL 75 Elvis 999369 NULL 76 Colby 908849 NULL 78 William 691457 NULL 110 Jackson 753217 NULL 111 Warren 999893 NULL 154 Madison 824486 NULL 156 Joelle 581331 NULL
SUA QUERY	select users.id, name, cpf, email from users left join profile on users.id = profile.user_id where State = 'Rio de Janeiro';

3) Selecione Id, Nome e E-mail e Estado dos usuários que possuem um e-mail cadastrado	
Resultado	id name email state 59 Sandra scelerisque.neque.Nullam@aliquet.ca São Paulo 55 Josiah Vivamus@antedictumcursus.ca Rio Grande do Sul 94 Neve iaculis.quis.pede@feugiat.edu Bahia 11 Iona et@facilisis.org Rio Grande do Sul

	95 Macaulay molestie.tellus@turpisegestas.edu Santa Catarina 11 Iona Cras.eu@facilisis.com Rio Grande do Sul 65 Mia dignissim@ipsumnonarcu.co.uk Ceará 57 Octavius laoreet.lectus.quis@acrisusMorbi.net Paraná 68 Wyoming ac@tristiqueac.edu São Paulo 64 Rinah a.feugiat@Donectempuslorem.com Pará 66 Kirk lorem.ut@Aliquamtincidunt.ca Paraná 68 Wyoming Phasellus@metusurnaconvallis.com São Paulo 59 Sandra Mauris@eratnonummy.net São Paulo 50 Barry vestibulum@nuncsit.co.uk São Paulo 64 Rinah Nulla@sit.org Pará 53 Clare auctor.non.feugiat@maurishoncus.net Minas Gerais 88 Leigh ornare.sagittis@adipiscing.org Paraná 95 Macaulay quam.quis@metusurna.ca Santa Catarina 31 Quon dignissim.Maecenas.ornare@eleifend.org Rio de Janeiro
SUA QUERY	select users.id, name, email, state from users, profile where users.id = profile.user_id and email != "";

4) Selecione os cinco estados com maior quantidade de usuário/cadastros;	
Resultado	São Paulo Minas Gerais Rio de Janeiro Ceará Paraná
SUA QUERY	select state, count(*) from users group by state order by count(*) desc limit 5;

5) Você encontrou alguma anomalia no conjunto de dados fornecidos? Se sim, quais? Como resolveria?	
Sua resposta	<p>Sim. Existem muitos dados faltantes, tanto na tabela users quanto no campo user_id (na tabela profile) sendo o último o mais problemático, visto que ele é responsável pelo relacionamento entre as duas tabelas. Outras anomalias também ocorrem na tabela users como todos os usuários fazendo com a data de nascimento igual (o que é possível dependendo do contexto) e datas de criação em datas futuros, o que claramente não é possível.</p> <p>Para os problemas de dados faltantes não há muito o que se possa fazer dado a natureza dos dados que estão faltando. Dados como CPF e ZipCode não podem ser inferidos computacionalmente. Dependendo do percentual de linhas com dados faltantes no dataset é viável apenas excluí-las. Caso elas representem uma grande parte do dataset é possível, em alguns casos, buscá-las em fontes externas</p>

	<p>como sites de órgãos públicos e etc.</p> <p>Já para o caso das datas de nascimento e criação a solução é mais simples. Se todos nasceram exatamente no mesmo dia esse campo não acrescenta muito significado ao conjunto de dados então pode ser removido. Quanto às datas de criação inconsistentes, suas linhas podem ser removidas já que representam uma pequena parte do conjunto.</p> <p>*As Soluções foram pensadas não levando em conta controle sobre a obtenção dos dados. Assumindo que os dados foram recebidos com o propósito serem analisados.</p>
<ul style="list-style-type: none"> Esta pergunta deve ser respondida com texto. Não é obrigatório adicionar nenhum tipo de conteúdo de suporte. 	

Parte 2

* Aqui não é necessário disponibilizar os códigos SQL. **

** Seja objetivo, respostas grandes e complexas não resolverão os problemas

*** 3 a 5 linhas são suficientes para descrever. Se achar cabível, utilize lista / *bullet* points.

Situação (1) Problema de acesso à plataforma

Um usuário com problema de cadastro um usuário entra em contato através de um dos nossos canais de comunicação, informa que está com um problema em seu cadastro e que tenta recuperar sua senha, mas seu código de confirmação não chega em seu e-mail.

(a) Quais informações são necessárias para iniciar uma investigação na base de dados para entender o problema da usuária?

Sua resposta: Nome, CPF e ZipCode

(b) Descreva uma solução possível:

Sua resposta:

Nem todos os emails estão ligados a um usuário, para descobrir se essa é a situação da usuária é necessário descobrir o Id dela. Para descobrir isso bastaria procurar pelo CPF da cliente, entretanto o CPF dela pode não estar cadastrado, então, pode se usar o nome e ZipCode visto que o nome pode se repetir e ZipCode serviria para descobrir quem é a usuária.

Situação (2) Usuário que gostaria de alterar e-mail de acesso

A usuária Sandra Castelo entrou em contato solicitando a alteração do seu e-mail pois perdeu acesso ao e-mail antigo.

A usuária apenas informou que seu e-mail antigo, e gostaria de alterar para sandra.castelo@meu-email.com

(a) Com a situação descrita acima, você acredita ser possível resolver o problema?

Sua resposta: Não. Assumindo que os dados pessoais já foram verificados, só há uma Sandra no conjunto de dados, entretanto, ela possui dois emails, sendo necessário saber qual email ela deseja alterar.

* Não ficou claro se o email antigo foi passado, então assumi que não foi.

(b) Caso não, quais dados você considera necessário para realizar a tarefa?

Sua resposta: email que deve ser alterado