

0



arnaldo.carlos.lucas@gmail.com (https://academy.semantix.com.br/courses/108/users/5901)

7 de fev de 2022

Boa noite.

No exercício 8 **Qual nome e quantidade das 5 crianças mais nascidas em 2016 do sexo masculino e feminino**, o professor fala que se quiser coletar as 5 crianças mais nascidas de cada sexo totalizando 5 de cada, só precisa aumentar o limit 5 para 10, minha dúvida é: Isso não aconteceu por coincidência dos dados terem exatamente o top 10 das crianças com 5 de cada sexo? Se por exemplo fosse as 8 primeiras crianças mais nascidas sendo meninas (F), e usasse essa mesma query, retornaria 8 F e 2 M, faz sentido?

Obrigado!

Editado por arnaldo.carlos.lucas@gmail.com (https://academy.semantix.com.br/courses/108/users/5901) em 7 Fev em 23:54

← Responder (3 curtidas)

0



<u>lucas.pam19@gmail.com (https://academy.semantix.com.br/courses/108/users/5936)</u>

8 de fev de 2022

Bom dia Arnaldo,

Tive exatamente a mesma dúvida que você, e para solucionar eu utilizei a clausura **UNION ALL** do hive e funcionou, embora a query tenha ficado um pouco mais pesada

para retornar os valores:

(select nome,sexo,max(frequencia) as c1 from nascimento where ano =2016 and sexo = 'F' group by nome,sexo order by c1 desc limit 5) union all (select nome,sexo,max(frequencia) as c1 from nascimento where ano =2016 and sexo = 'M' group by nome,sexo order by c1 desc limit 5);

← Responder / (7 curtidas)

0



arnaldo.carlos.lucas@gmail.com (https://academy.semantix.com.br/courses/108/users/5901)

8 de fev de 2022

Boa Lucas! Eu tinha pensado em fazer algo do tipo, mas quando olhei a correção que fiquei com isso na cabeça hahahaha. Boa saída mano.

← Responder / (2 curtidas)

0



Rodrigo Augusto Rebouças (https://academy.semantix.com.br/courses/108/users/5)

9 de fev de 2022

Obrigado pela observação Arnaldo, realmente não tinha reparado nisso.

A resposta do Lucas teria mais sentido, obrigado por compartilhar.

← Responder ← (3 curtidas)

0



vasc.edson@gmail.com (https://academy.semantix.com.br/courses/108/users/4484)

9 de fev de 2022

Eu fiz com subselect:

select * from (select nome, sexo, sum(frequencia) as qtd from edson.nascimento

where ano = 2016 group by nome, sexo order by qtd desc limit 10) fo order by sexo, qtd desc;

0



souza.chagas.felipe@gmail.com (https://academy.semantix.com.br/courses/108/users/5584)

7 de fev de 2022

•

Olá Professor,

Quando tento executar a query:

SELECT nome, COUNT(nome) AS qtde FROM nascimento WHERE ano=2017 GROUP BY nome;

É retornada a seguinte mensagem:

Unknown HS2 problem when communicating with Thrift server.

Error: org.apache.thrift.transport.TTransportException: java.net.SocketException: Broken pipe (Write failed) (state=08S01,code=0)

Depois disso o contêiner do Hive cai e preciso usar o docker-compose start para que o contêiner suba novamente.

Quando executei as queries mais simples dos exercícios anteriores e do primeiro exercício dessa lista, funcionou normalmente, o problema ocorre quando executo essas queries com função de agregação e que retornam mais dados.

No momento estou utilizando Arch Linux em um equipamento que possui apenas 4Gb de Ram e um processador Core 2Duo.

É possível realizar configuração para que o cluster não caia ou a configuração da máquina não possui o mínimo para o cluster?

Desde já agradeço!

0



delanosilva_182@hotmail.com (https://academy.semantix.com.br/courses/108/users/5706)

8 de fev de 2022

.

Boa tarde Souza!

Passei por isso tbm e agora não desligo mas meu pc. Fiquei quase um dia com esse problema de comunicação.

Eu tava achando que era erro meu, mas não foi.

0



souza.chagas.felipe@gmail.com (https://academy.semantix.com.br/courses/108/users/5584)

8 de fev de 2022

.

Boa noite Delano!

Cara, consegui resolver fazendo o seguinte (Utilizo o sistema operacional Arch Linux):

- Arquivo docker-compose.yml:
 - Aumentei a memória para o hive-server.
- No terminal rodei os seguintes comandos:
 - docker-compose down
 - docker-compose pull (Não sei se foi necessário)
 - docker-compose up -d

Feito isso, não tive mais problemas do hive-server caindo na tentativa de rodar alguma consulta.

Espero ter ajudado

Editado por souza.chagas.felipe@gmail.com (https://academy.semantix.com.br/courses/108/users/5584) em 8 Fev em 21:05

0



Rodrigo Augusto Rebouças (https://academy.semantix.com.br/courses/108/users/5)

9 de fev de 2022

Obrigado por compartilhar a resolução, caso queiram alterar os recursos do serviço, pode ser editado o docker-compose da seguinte maneira:

deploy:

resources:

limits:

cpus: 0.50

memory: 512M

reservations:

cpus: 0.25

memory: 128M

No caso do hive, se alterar o limite da memória para **1GB**, será o suficiente.



0

rafaelsanchesz@hotmail.com (https://academy.semantix.com.br/courses/108/users/1743)

8 de fev de 2022

Boa noite,

Tive um erro quando fui subir o docker: "Ports are not available: listen tcp 0.0.0.0/50070: bind: An attempt was made to access a socket in a way forbidden by its access permissions."

se caso tiverem passando pelo mesmo problema, consegui resolver ele, segue os passos a baixo:

Entrar no prompt como administrador

aplicar o comando - net stop winnat

aplicar o comando - docker start 'container_name' ou você pode ir até o 'ubunto' e dar um start no docker e verificar se subiu todos, se sim então voltar para o cmd

aplicar o comando - net start winnat

Bom foi o que solucionou meu problema, perdi um tempinho até conseguir e espero ajudar.



0

Rodrigo Augusto Rebouças (https://academy.semantix.com.br/courses/108/users/5)

9 de fev de 2022

Obrigado por postar a sua resolução e na aula de problemas de instalação com windows também explico esse motivo.

0



renanfsvieira@gmail.com (https://academy.semantix.com.br/courses/108/users/5866)

9 de fev de 2022

Boa noite galera,

Estou um pouco perdido e confuso com tanto comando em linha de comando :s :s

Alguém teria um mapa mental ou anotações de todas operações utilizadas nas aulas com as ferramentas de big data que possa compartilhar? ficarei grato.

0



elnataoliveira@gmail.com (https://academy.semantix.com.br/courses/108/users/5697)

11 de fev de 2022

Estou organizando aos poucos, talvez possa te ajudar.

https://github.com/elnataoliveira/data-engineer-semantix (https://github.com/elnataoliveira/data-engineer-semantix)

Obs.: Tem alguns diretórios a mais, mas os de docker, hadoop e hive tentei organizar de acordo com o curso.

Editado por elnataoliveira@gmail.com (https://academy.semantix.com.br/courses/108/users/5697) em 11 Fev em 9:54

0



renanfsvieira@gmail.com (https://academy.semantix.com.br/courses/108/users/5866)

11 de fev de 2022

•

Muito obg Elnato, irei acompanhar.

Responder (1 curtida)

0



Rodrigo Augusto Rebouças (https://academy.semantix.com.br/courses/108/users/5)

12 de fev de 2022

.

Parabéns pelo git, até o final do curso estará com muito conteúdo, rs

← Responder

21/02/22, 17:59 Assunto: Dúvidas - Aula 6

0



msena2210@gmail.com (https://academy.semantix.com.br/courses/108/users/5754)

Quarta-feira

Professor,

Percebi que existe um registro null na tabela pop, ao tentar deletar com o comando abaixo, existe uma diretiva de segurança que não me permitiu Como poderia excluir esse registro?

0: jdbc:hive2://localhost:10000>

delete from pop_parquet where pop_parquet.zip_code = NULL;

Error: Error while compiling statement: FAILED: SemanticException [Error 10294]: Attempt to do update or delete using transaction manager that does not support these operations. (state=42000,code=10294)

← Responder (2 curtidas)

0



Rodrigo Augusto Rebouças (https://academy.semantix.com.br/courses/108/users/5)

Sexta-feira

•

Você não pode deletar ou atualizar, pois é uma tabela externa, apenas usa para consultas.

← Responder / }