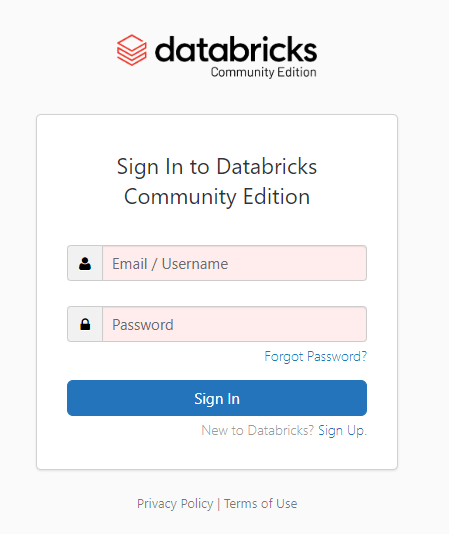
Olá, ----------!

Tudo bem?

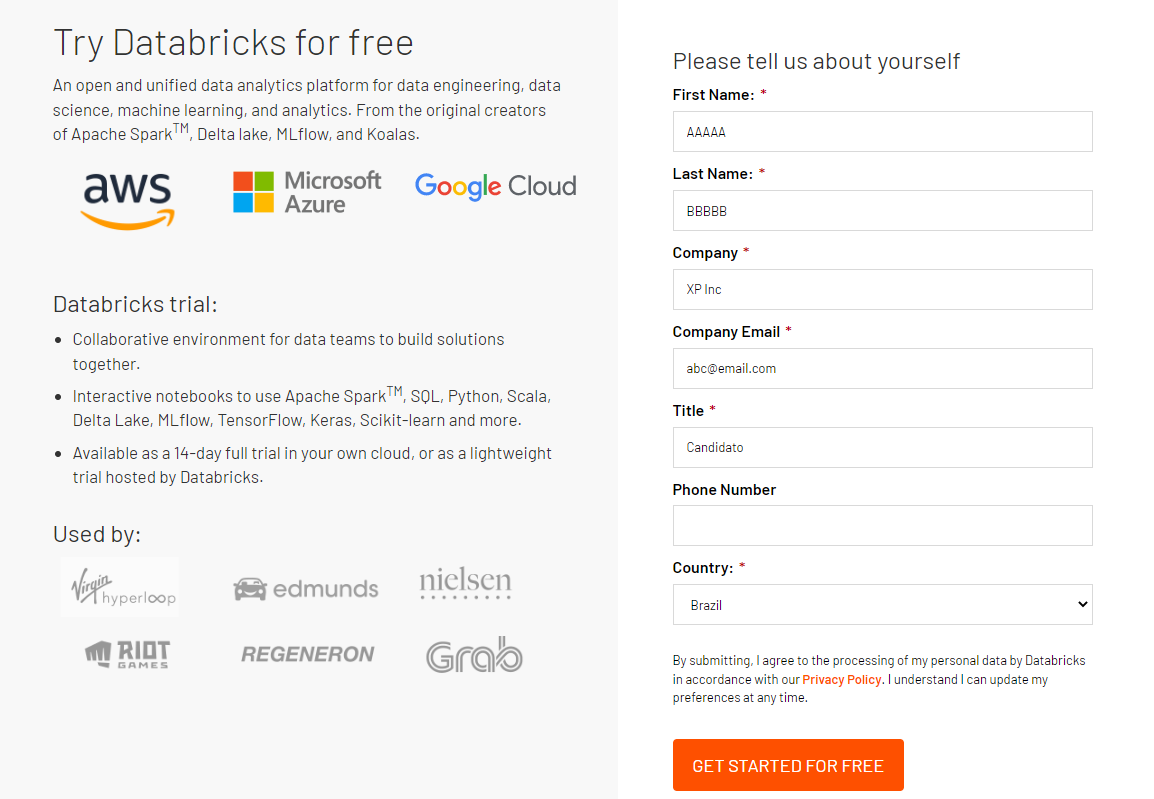
Muito obrigado por participar dessa etapa, queremos que você conheça um pouco mais do nosso dia a dia aqui na XP, para isso vamos simular um case que têm relação com o nosso negócio. Além disso, vamos apresentar uma de nossas principais ferramentas de trabalho, o Databricks e vai ser muuito legal compartilhar com você um pouco de como usamos essa ferramenta incrível de consulta, transformação e visualização de dados.

Bem... já sei o que você está pensando caso nunca tenha usado o Databricks, mas não se preocupe, nosso objetivo é te inserir no universo XP, então vamos dar o passo a passo de como usar a versão gratuita. Esse teste foi desenvolvido para ser realizado utilizando o Databricks, mas **se você tem familiaridade com outras ferramentas** que permitam abrir os arquivos disponibilizados em .dbc (extensão de notebook) e .csv, **você pode ficar à vontade para utilizá-las**. Caso tenha optado por seguir com o Databricks, vamos juntos pelo passo a passo!

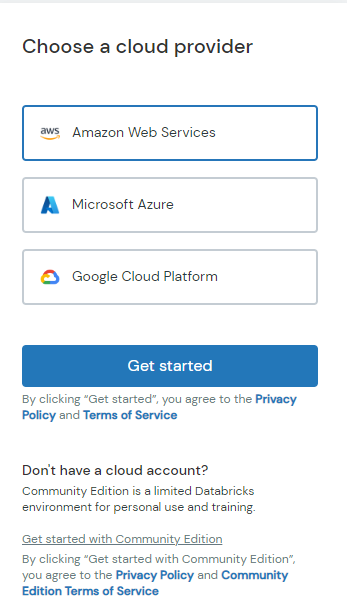
1. Acesse o site [Databricks - Sign In](https://community.cloud.databricks.com/login.html)
2. Na página inicial você irá ver uma tela de Login e para fazer seu cadastro, basta clicar em **Sign Up.**



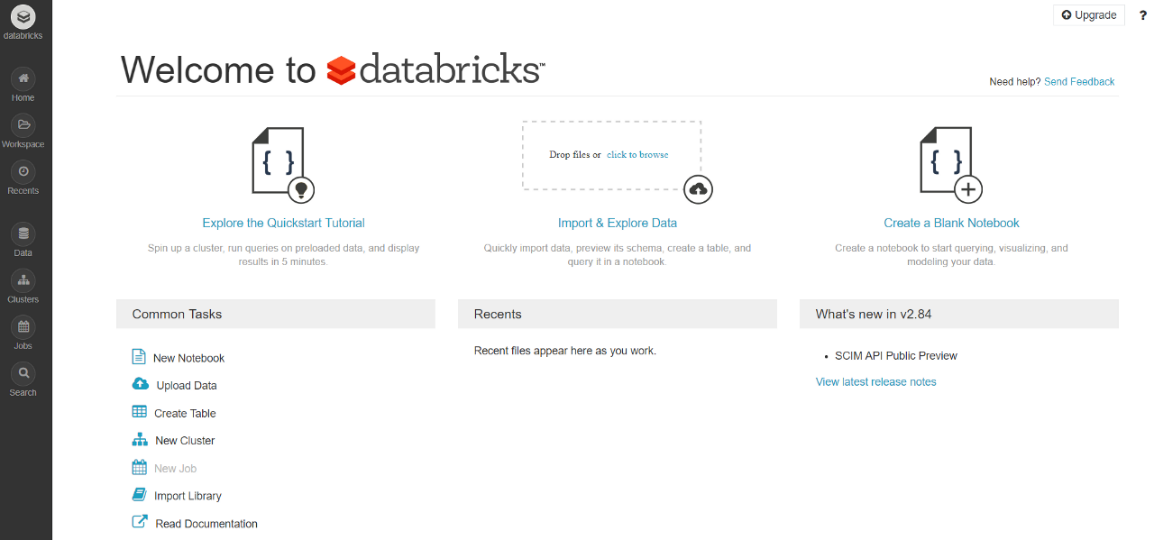
1. Em seguida, você será levado para uma página de cadastro e deve preencher as informações que são pedidas (Nome, Sobrenome etc.). No campo de empresa você pode colocar “XP Inc” e no campo de e-mail da empresa, insira seu e-mail pessoal. Após preencher as informações, é só clicar em **Get started**.



1. Na página para escolher o cloud provider, você irá selecionar **Get started with Community Edition**.

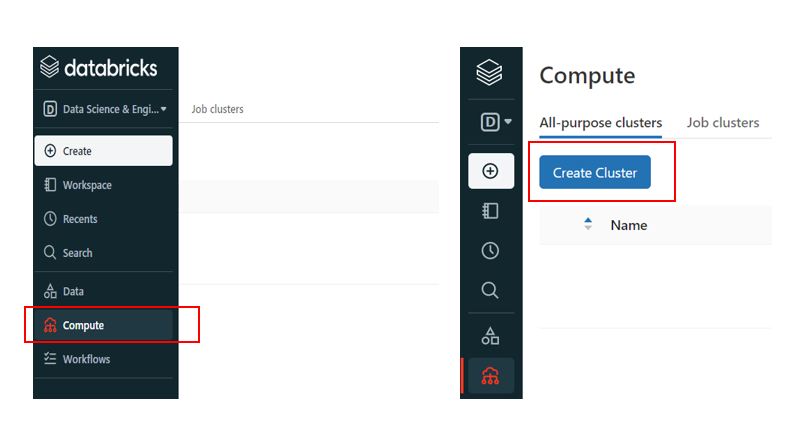


1. Você receberá uma mensagem no e-mail para confirmar o seu e-mail cadastrado no Databricks Community. Após confirmação e (não esqueça de olhar a caixa de Spam), você será direcionado para a página do Databricks Communitiy na qual você irá em para criar uma senha, que deve conter pelo menos uma letra maiúscula e um símbolo.
2. Agora você está na página inicial do databricks e está tudo pronto para começar!

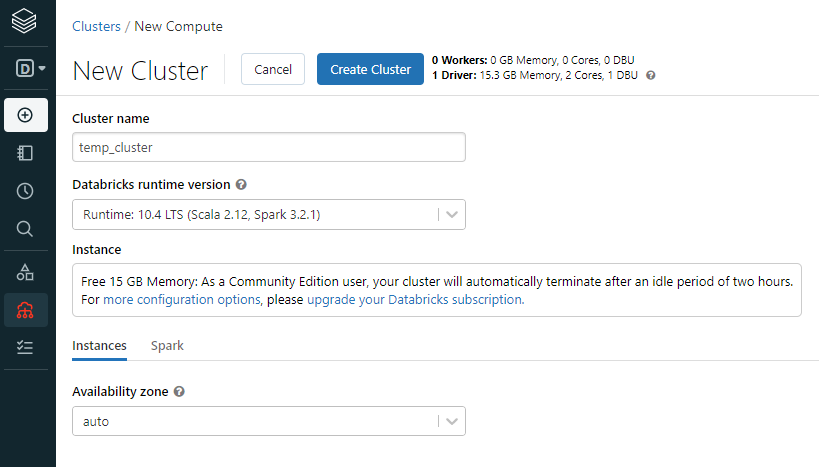


Para conseguir realizar o teste técnico, você receber um arquivo em CSV que deverá ser inserido manualmente no Databricks Community, para que você o utilize como tabela de consulta.

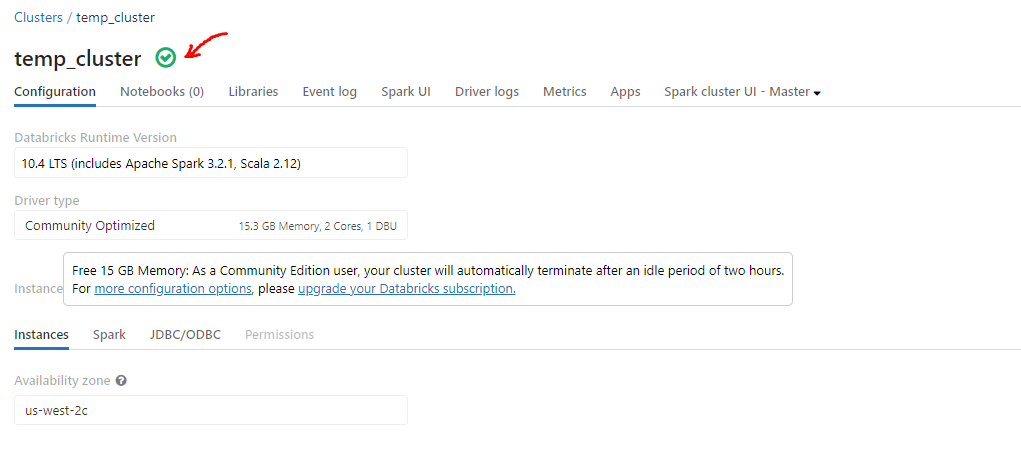
1. Na caixa de ferramentas do lado esquerdo, selecione **Compute** e depois selecione **Create Cluster**.



1. Ao criar o cluster, você precisará dar um nome e clicar na opção **Create Cluster**, mantendo as configurações padrão.

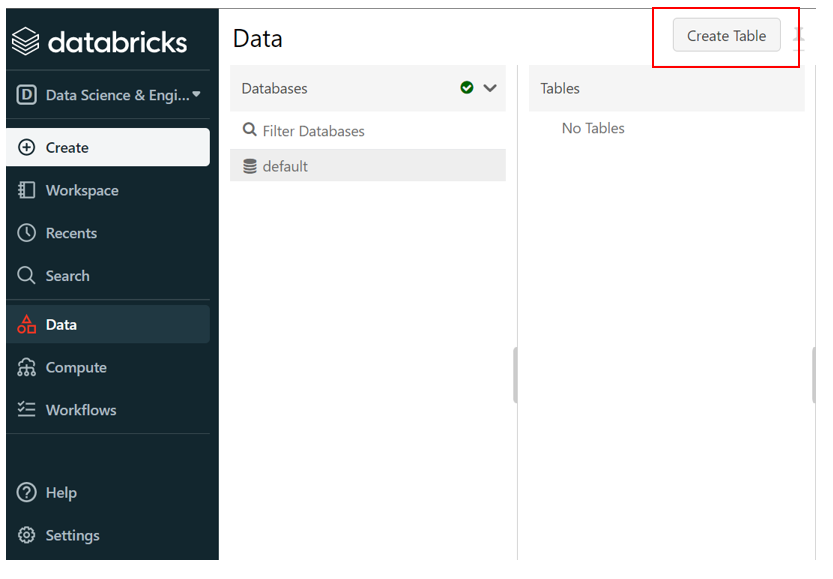


1. Após isso o cluster levará alguns minutos até ser criado de fato, aguarde até que o símbolo da imagem apareça (nesse exemplo colocamos o nome do cluster de “temp\_cluster” evidenciando a possibilidade de cria-lo com qualquer nome)

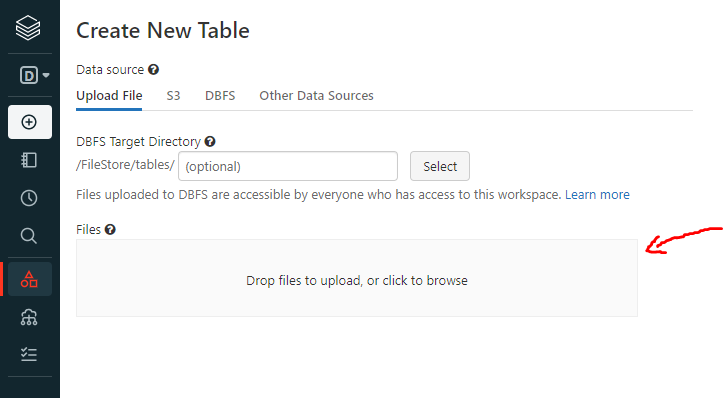


Por ser um conta gratuita o cluster pode acabar desligando sozinho depois de 2 horas, caso isso ocorra você deverá refazer os processos a partir da etapa 8

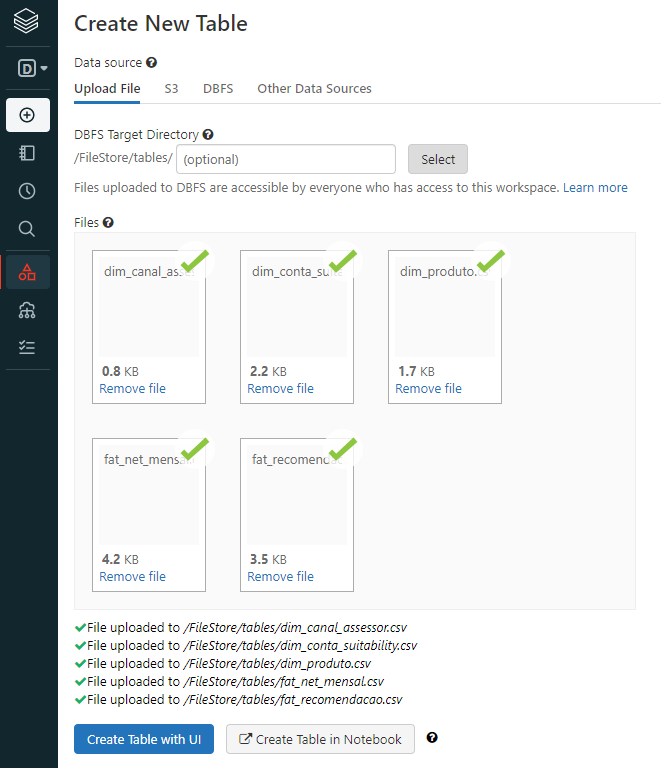
1. Depois de carregar seu cluster recém-criado, você deverá fazer o upload das tabelas que foram disponibilizadas no formato CSV. Na caixa de ferramentas, selecione **Data,** em seguida **Create Table**



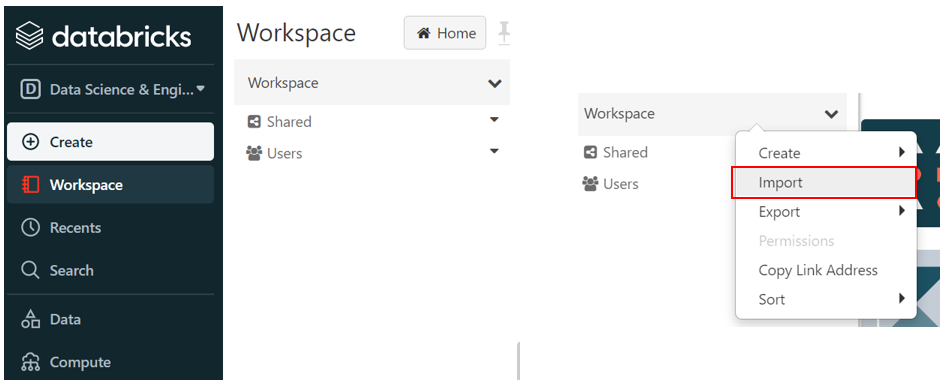
1. O próximo passo é carregar os arquivos CSV no repositório do Databricks, para isso basta arrastá-los até a caixa cinza exibida na imagem ou simplesmente clicar nessa caixa e em seguida selecionar os arquivos em sua máquina



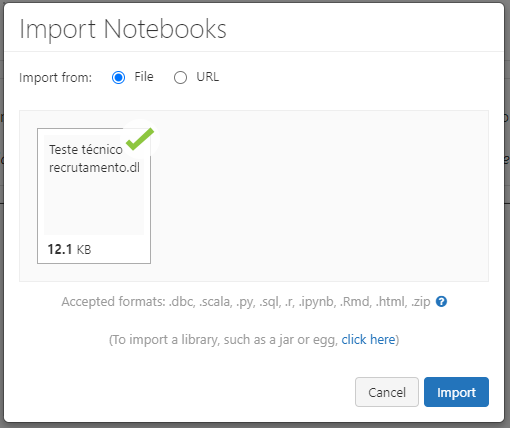
1. Após selecionar todos os arquivos CSV eles serão carregados e ficarão com um símbolo de “check” verde em cima deles, após isso essa etapa está finalizada, não precisa clicar mais em nada nessa tela, chegou a hora de criar nosso notebook e criar nossas tabelas de fato



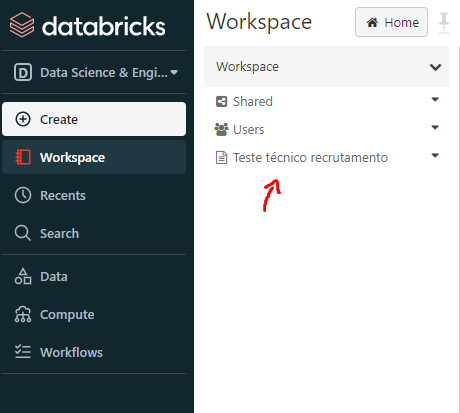
1. Por fim, após carregar os arquivos CSV para o Databricks, basta carregar o notebook disponibilizado e seguir as instruções contidas nele. (O notebook é o arquivo com formato DBC). Na barra de ferramentas à esquerda selecione **Workspace** e clique com botão direito do mouse para abrir uma caixa de opções, conforme a imagem abaixo, selecione **Import**.



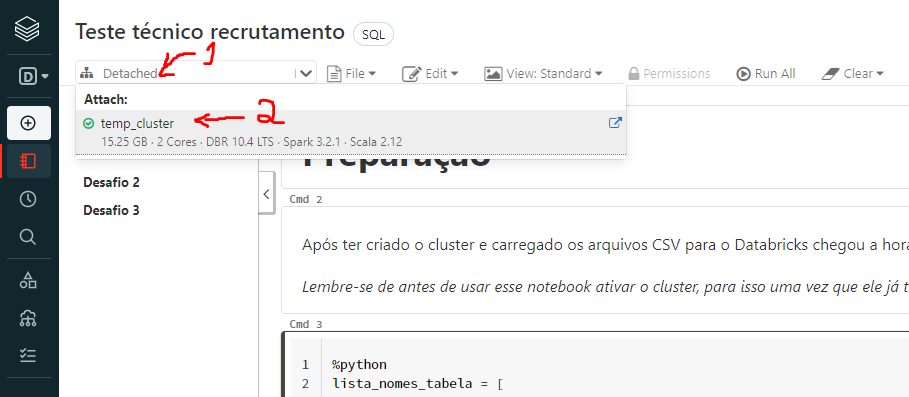
1. Nesse passo, carregue o arquivo DBC enviado e confirme a importação.



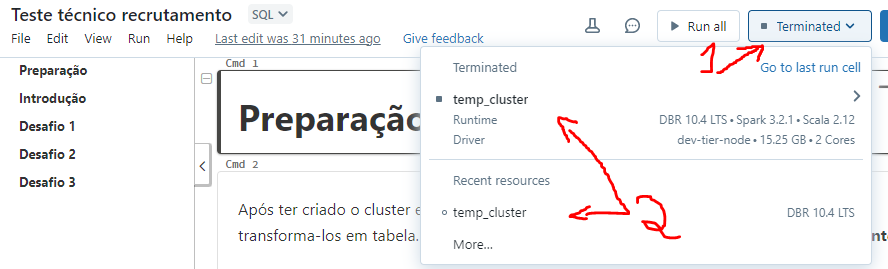
1. Com isso você poderá acessar o notebook que preparamos para você clicando no nome dele:



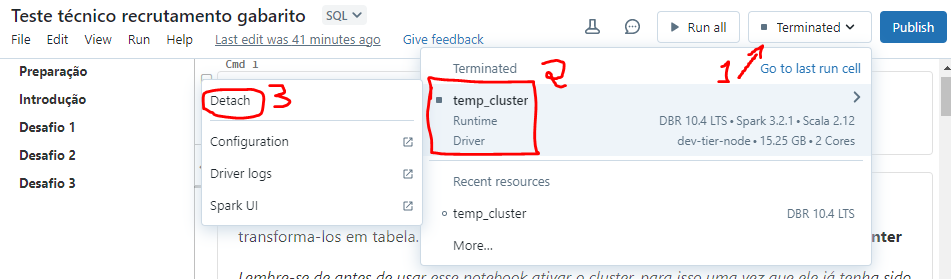
1. Agora dentro do notebook você precisará se conectar no cluster que criou mais cedo para isso basta clicar no texto escrito “Detached” e selecionar o cluster que você criou. Detached significa que o cluster não está conectado, se no seu caso o cluster já estiver conectado não precisa se preocupar com essa etapa:



Dependendo da versão de seu Databricks, o layout pode estar um pouco diferente, ao invés de aparecer “Detached” pode estar escrito “Terminated” ou “Connect” e o botão estará no outro canto da tela, mas o processo é o mesmo, basta clicar nele e selecionar o cluster que você criou, como exibido na imagem a seguir:

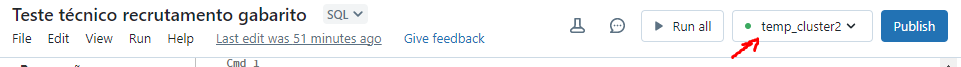


Se ainda assim ficar aparecendo “Terminated” ao invés de “Connected” ou o nome do cluster, clique no cluster e selecione “Detach” como mostra a imagem a seguir:

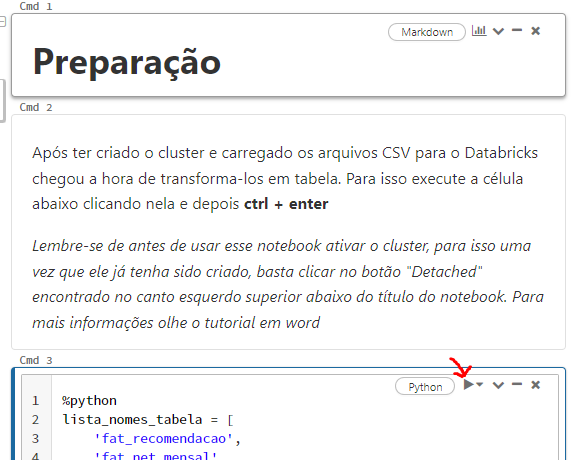


Caso você fique muito tempo inativo o cluster pode desconectar do Databricks se isso acontecer basta repetir essa etapa

Depois de conectado ao invés de “Terminated” ou “Connect” aparecerá o nome do cluster selecionado:



1. Agora basta criar as tabelas que foram carregadas anteriormente em formato CSV dentro do Databricks. Para isso rode o código da primeira célula do notebook, traduzindo, basta ir à seção de “Preparação” e clicar no símbolo “play“ que está sendo destacado com a setinha vermelha na imagem abaixo. Esse processo levará aproximadamente 30 segundos e depois... pronto! As tabelas já estarão criadas e você poderá seguir para o passo a passo que está destacado no próprio notebook



1. Agora você já tem tudo que precisa para concluir essa etapa do processo! No notebook constam as questões que você irá resolver nessa etapa e instruções para aprender um pouco mais sobre o mundo XP. Esse notebook será depois usado na entrevista técnica e você poderá apresentar sua solução, caso você não consiga avançar muito não se preocupe, o entrevistador vai te guiar e tentar te ajudar com as dúvidas que você teve, o mais importante para nós é que você saia desse processo com mais conhecimento do que entrou

**Atenção:**

O cluster pode acabar desconectando do notebook sozinho caso isso aconteça basta repetir a etapa 16 para reconectar o cluster.

Por ser uma conta gratuita outro ponto que pode acontecer é o cluster desligar depois de 2h, se isso acontecer, você precisará recriar o cluster, subir os arquivos e criar as tabelas novamente, basta seguir todas as etapas a partir da 8, assim você ganhará mais 2h.