



PUC Minas

# Preparando Ambiente - Spark

Diego Roberto Gonçalves de Pontes

# Apache Spark

- O Apache Spark roda localmente para testes, mas para os nossos experimentos vamos utilizar um computador em nuvem para todos poderem utilizar.
- O recurso em nuvem que vai ser utilizado tem um período de testes de 14 dias sem a necessidade de cadastrar com cartão de crédito.

- A databricks já tem em seus clusters o Spark instalado e configurado.
- Acessar o link para fazer cadastro:
  - <https://www.databricks.com/try-databricks#account>
  - Basta realizar o cadastro e depois acessar o login
  - <https://community.cloud.databricks.com/login.html?tuuid=e4292078-e760-4b82-af92-7ae3aa90aeea>



databricks

Sign In to Databricks  
Community Edition

[Forgot Password?](#)

Sign In

[New to Databricks? Sign Up.](#)

Create your Databricks account 1/2

Sign up with your work email to elevate your trial with expert assistance and more.

First name

Last name

Email

Company

Title

Phone (Optional)

Country  
Brazil


What do you want to build and run with Databricks? (Optional)  
Please choose all options that are relevant


By submitting, I agree to the processing of my personal data by Databricks in accordance with our [Privacy Policy](#). I understand I can [update my preferences](#) at any time.


Continue





PUC Minas


 databricks


 New


 Workspace

 Recents


 Search


 Catalog


 Workflows


 Compute

Machine Learning


 Experiments

 Collapse menu

 You're using Databricks Community Edition. Upgrade for unlimited clusters and collaboration features.

Upgrade now 


## Get started

 Import and transform data

Create a table by uploading local files, or create a pipeline for continuous data ingestion and transformation.


Create table

Create pipeline

 Notebook

Create a new notebook for data analysis, transformation, and machine learning.


Create notebook

 AutoML

Accelerate the training of ML models for efficient discovery and iteration.


Start AutoML

## Recents



**No recent items**

Start exploring and your recently viewed items will show up here.

 PUC Minas

# Passos

- Criar um cluster;
- Realizar upload de arquivos;
- Criar notebook;
- Executar códigos.

# Notebooks

- Notebooks Python são ferramentas interativas que combinam código, texto, e visualizações em um único documento. Eles são amplamente usados em data science, machine learning, educação e pesquisa para experimentar e documentar análises de dados de forma clara e colaborativa. Um dos ambientes mais populares para trabalhar com notebooks Python é o Jupyter Notebook.



**PUC Minas**