

FACULDADE SENAC PR

Apresentação e aplicação da ferramenta Jupyter Notebook

Mauricio Antonio Ferste

Fecomércio PR
Sesc Senac IFPD

Senac





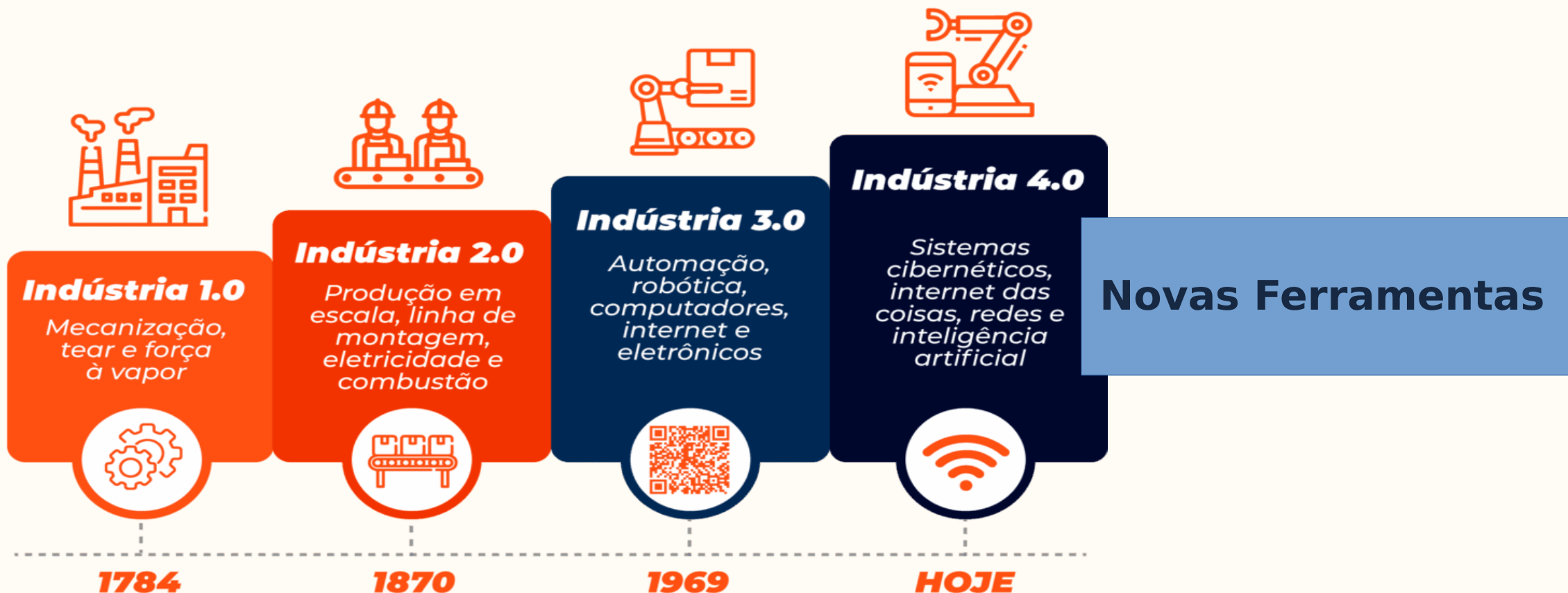
“Precisamos de diversidade de pensamentos no mundo para tratar os novos desafios.”

Tim Berners-Lee

<https://laidlawscholars.network/posts/we-need-diversity-of-thought-in-the-world-to-face-the-new-challenges>

Contexto da Aplicação do Notebook

Indústria 4.0



Indústria 4.0 - nova ênfase



Novas Ferramentas

PAAS (platform as a service), nuvem

Uso de IA no código

Colaborativo

Jupyter Notebook

Jupyter.org



[Try](#) [Install](#) [Get Involved](#) [Documentation](#) [News](#) [Social](#) [Governance](#) [Security](#) [About](#)



Free software, open standards, and web services for interactive computing across all programming languages

JupyterLab: A Next-Generation Notebook Interface

JupyterLab is the latest web-based interactive development environment for notebooks, code, and data. Its flexible interface allows users to configure and arrange workflows in data science, scientific computing, computational journalism, and machine learning. A modular design invites extensions to expand and enrich functionality.

Try it in your browser

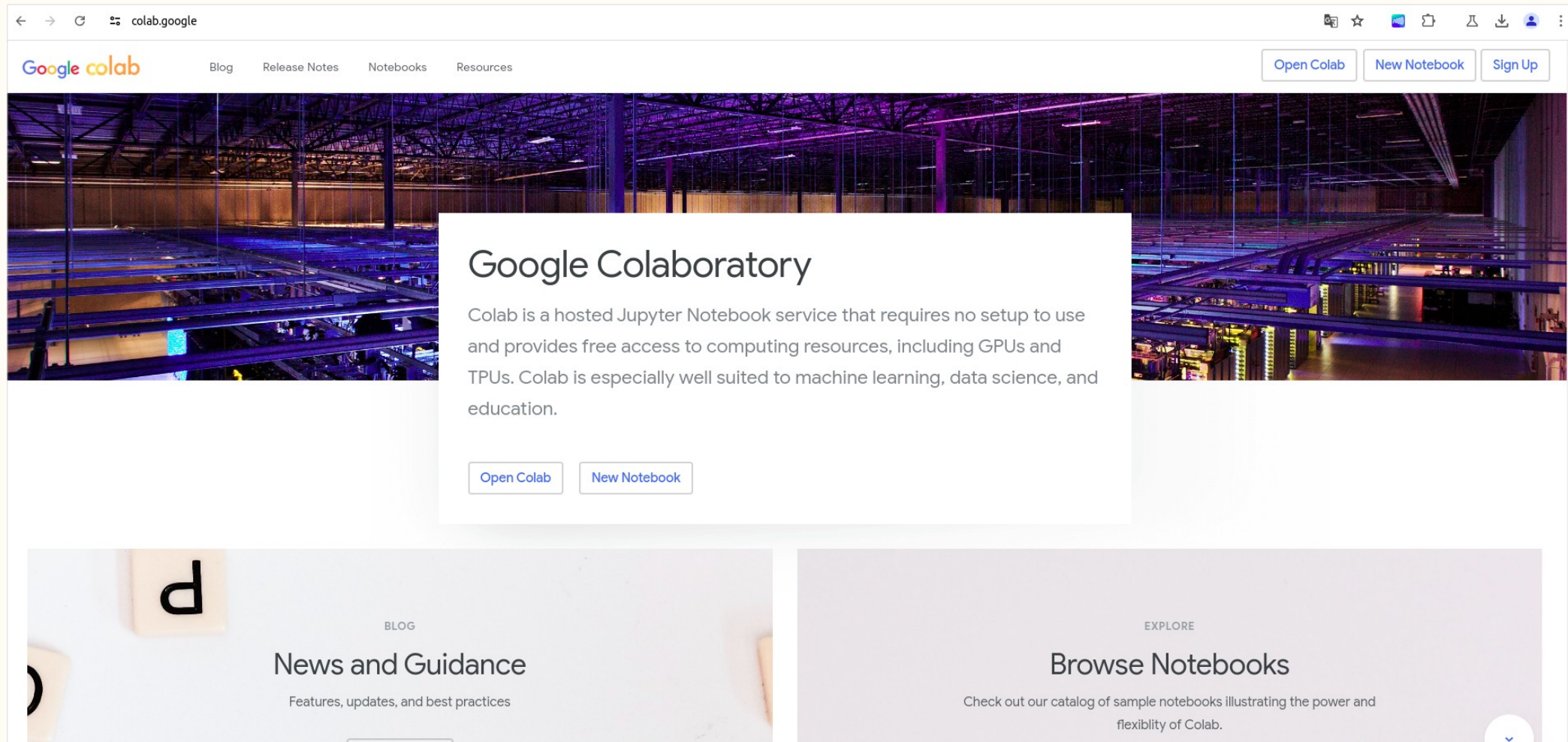
Install JupyterLab

Jupyter Notebook

Aplicação web de código aberto que permite criar e compartilhar documentos que contêm código, equações, visualizações e texto explicativo.

Suporta várias linguagens de programação e é amplamente utilizado em ciência de dados, aprendizado de máquina, pesquisa acadêmica e desenvolvimento de software.

Google Colab



The screenshot shows the Google Colab homepage in a web browser. The browser's address bar displays 'colab.google'. The page features a navigation bar with the Google Colab logo, links to 'Blog', 'Release Notes', 'Notebooks', and 'Resources', and three buttons: 'Open Colab', 'New Notebook', and 'Sign Up'. The main visual is a large banner with a server room background. A central white box contains the heading 'Google Colaboratory' and a paragraph describing the service as a hosted Jupyter Notebook with free access to GPUs and TPUs. Below this text are two buttons: 'Open Colab' and 'New Notebook'. At the bottom, there are two sections: 'News and Guidance' (labeled 'BLOG') and 'Browse Notebooks' (labeled 'EXPLORE').

colab.google

Google colab

Blog Release Notes Notebooks Resources

Open Colab New Notebook Sign Up

Google Colaboratory

Colab is a hosted Jupyter Notebook service that requires no setup to use and provides free access to computing resources, including GPUs and TPUs. Colab is especially well suited to machine learning, data science, and education.

Open Colab New Notebook

BLOG

News and Guidance

Features, updates, and best practices

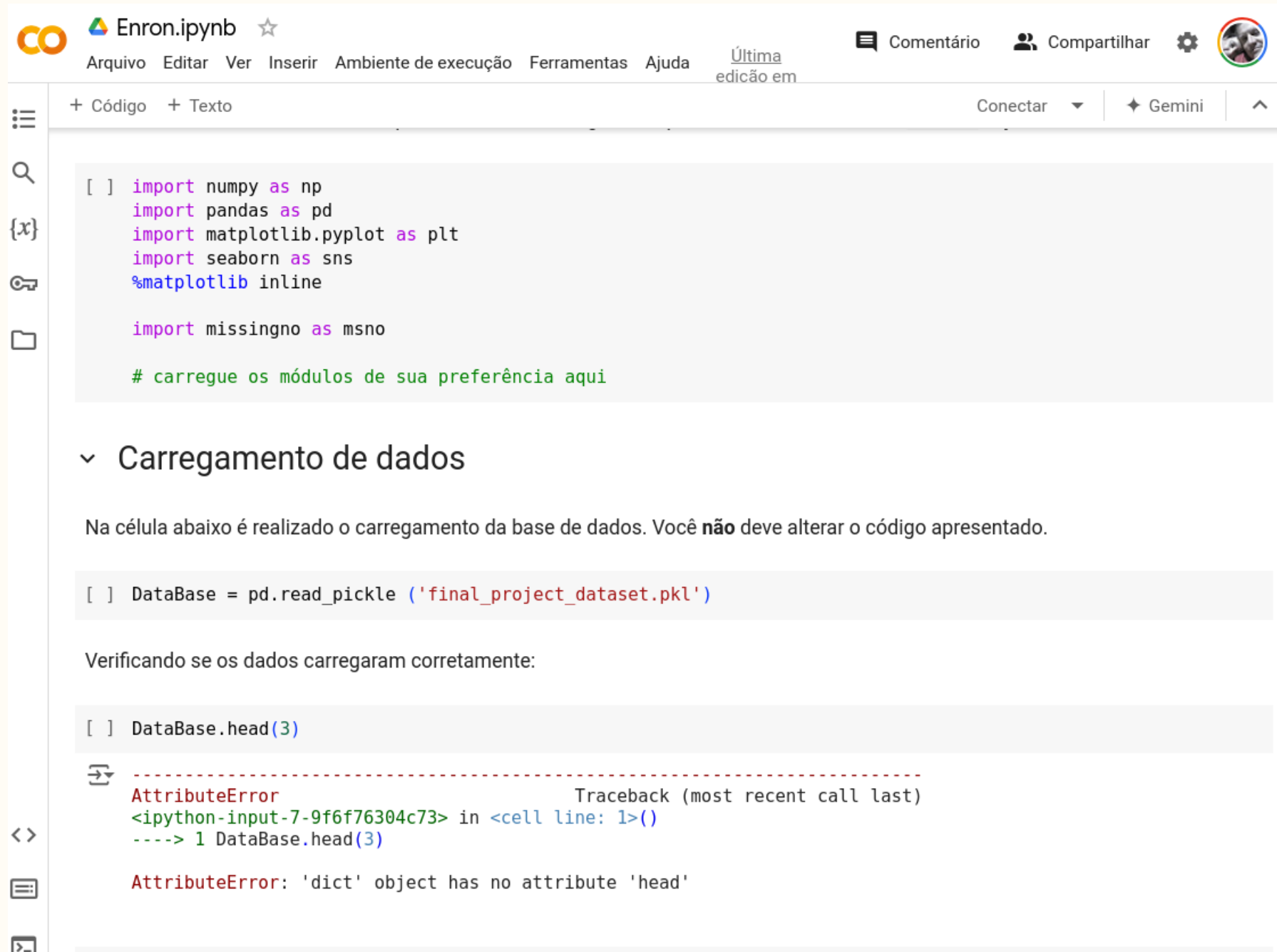
EXPLORE

Browse Notebooks

Check out our catalog of sample notebooks illustrating the power and flexibility of Colab.

Um Notebook

No fim teremos código !



The screenshot shows a Jupyter Notebook interface for a file named 'Enron.ipynb'. The top bar includes navigation links like 'Arquivo', 'Editar', 'Ver', 'Inserir', 'Ambiente de execução', 'Ferramentas', and 'Ajuda'. It also shows 'Comentário', 'Compartilhar', and a user profile icon. The left sidebar contains icons for file explorer, search, and other notebook functions. The main area displays a code cell with the following Python code:

```
[ ] import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline

import missingno as msno

# carregue os módulos de sua preferência aqui
```

Below the code cell, there is a section titled 'Carregamento de dados' (Data Loading). It contains a text block stating: 'Na célula abaixo é realizado o carregamento da base de dados. Você **não** deve alterar o código apresentado.' (In the cell below, the loading of the data base is performed. You **must not** change the presented code.)

Below this text is another code cell with the following code:

```
[ ] DataBase = pd.read_pickle ('final_project_dataset.pkl')
```

Below the code cell, there is a text block stating: 'Verificando se os dados carregaram corretamente:' (Checking if the data was loaded correctly:).

Below this text is a third code cell with the following code:

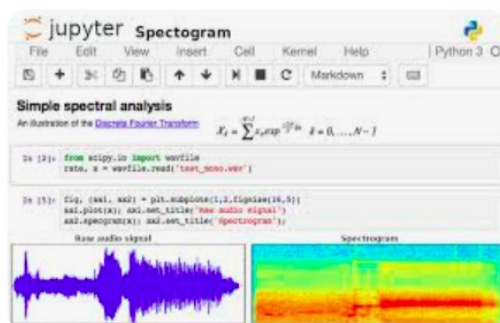
```
[ ] DataBase.head(3)
```

Below the code cell, there is an error message displayed in a light blue box:

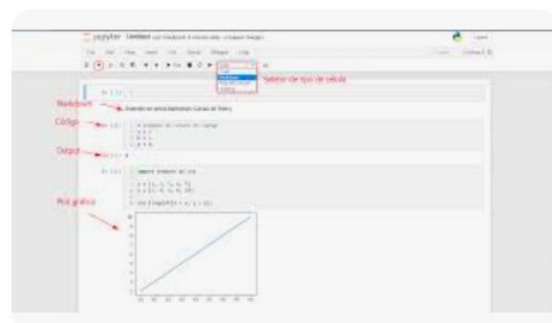
```
-----
AttributeError                                Traceback (most recent call last)
<ipython-input-7-9f6f76304c73> in <cell line: 1>()
----> 1 DataBase.head(3)

AttributeError: 'dict' object has no attribute 'head'
```

Várias Versões de Ferramenta



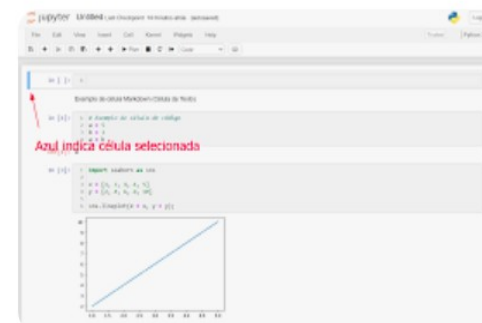
Phylos.net
Jupyter Notebook - Phylos.net



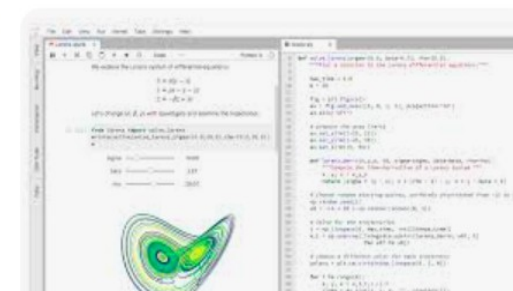
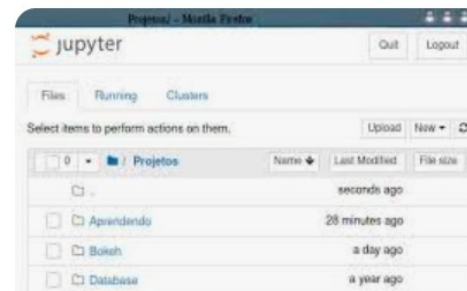
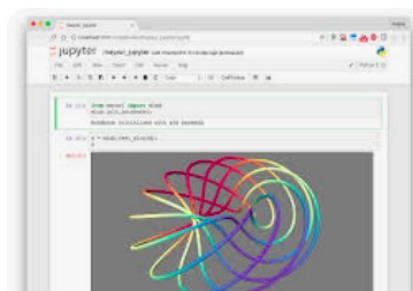
Alura
Jupyter Notebook: Exemplos de Códigos e C...



F. Schuch
Jupyter Notebook como uma Poderosa ...

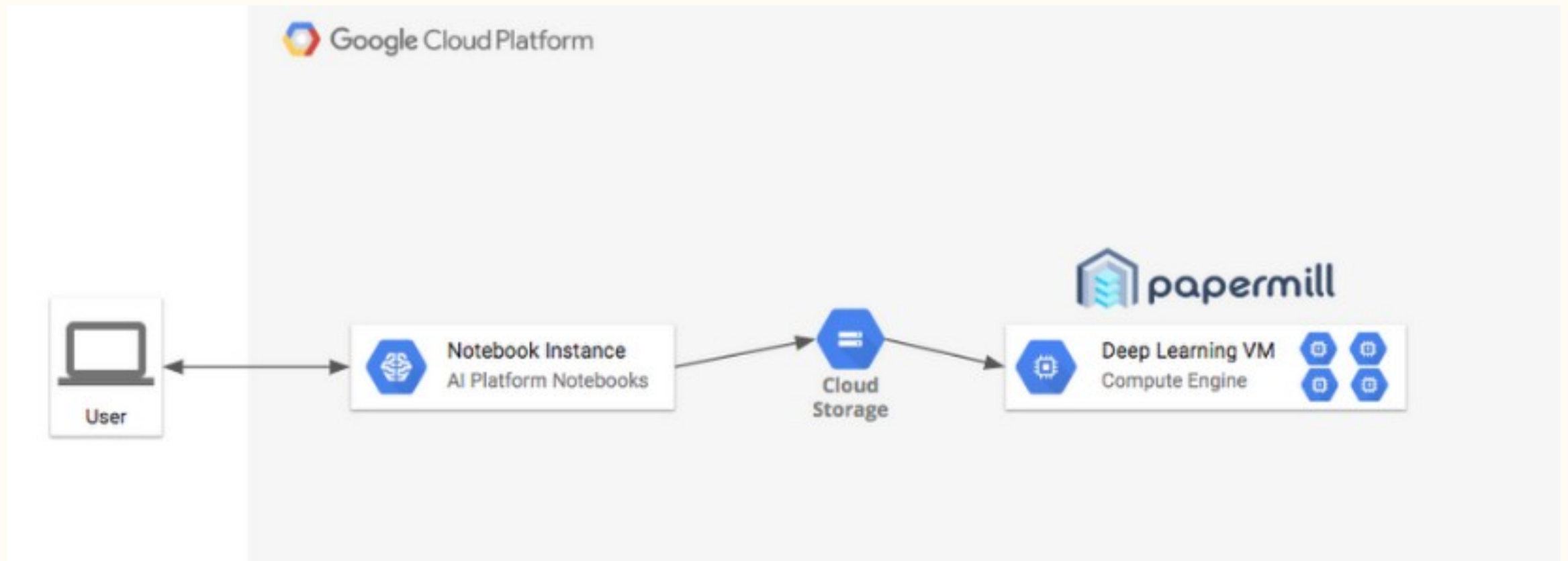


Alura
Jupyter Notebook: Exemplos de Código...

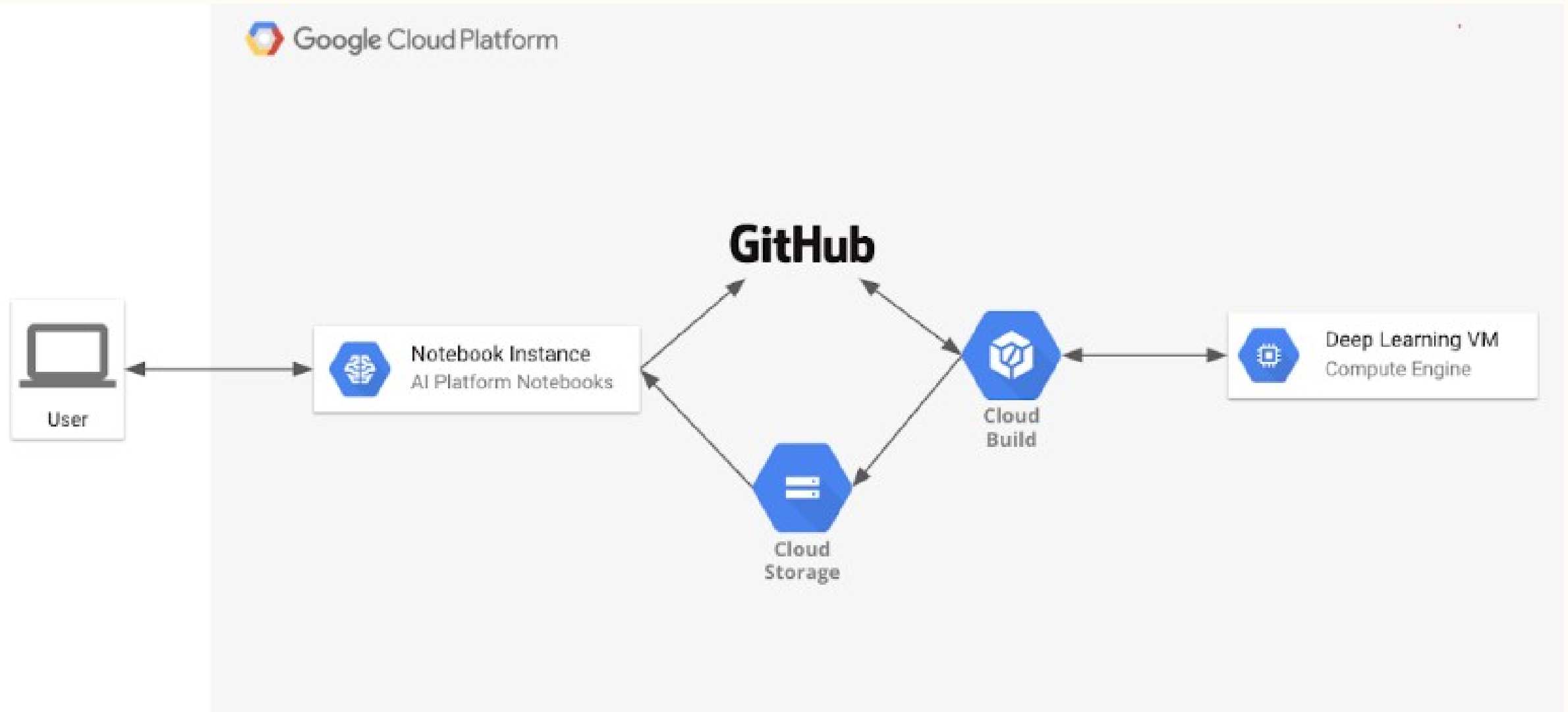


Conceitos

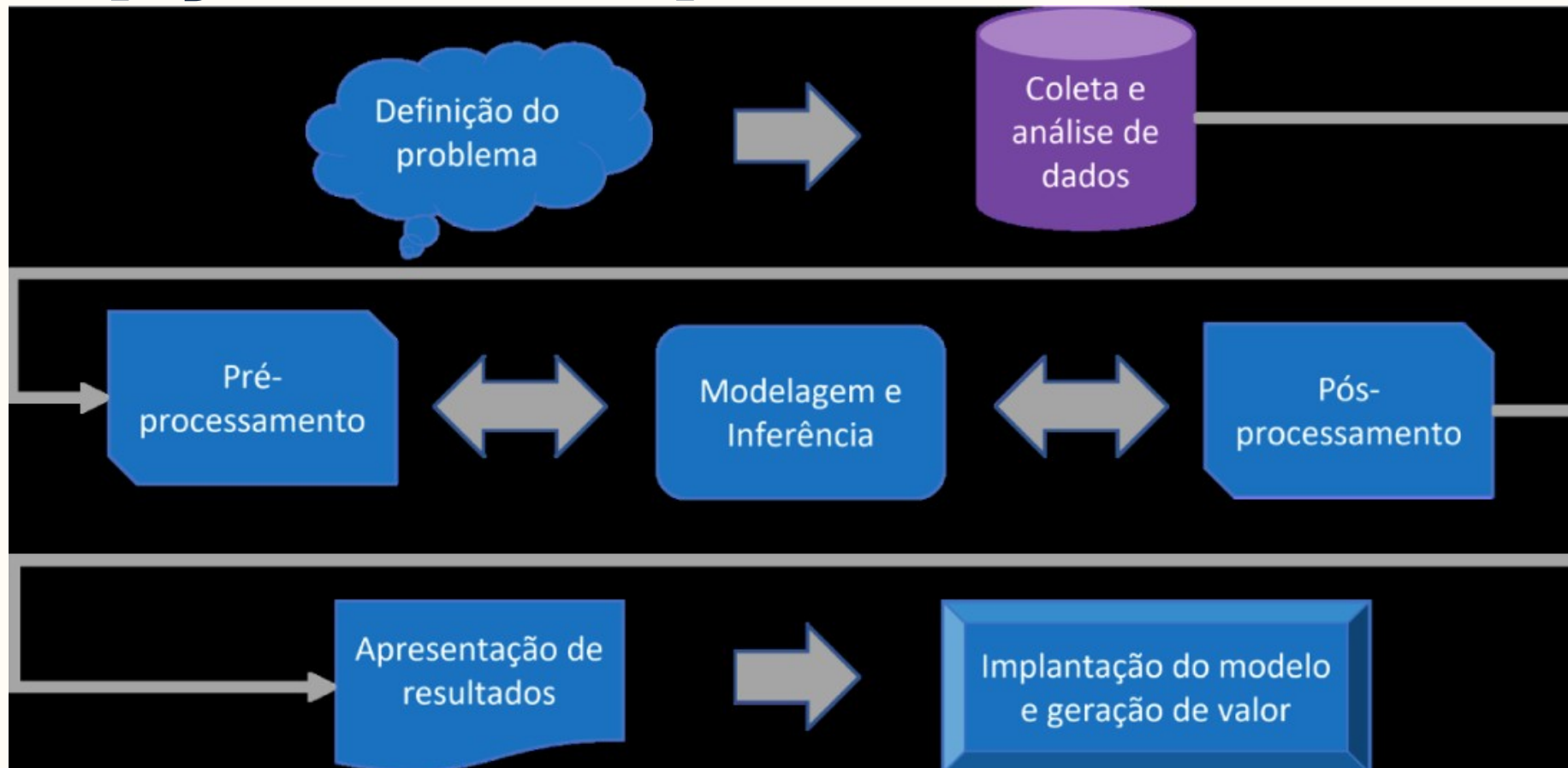
Arquitetura



Arquitetura

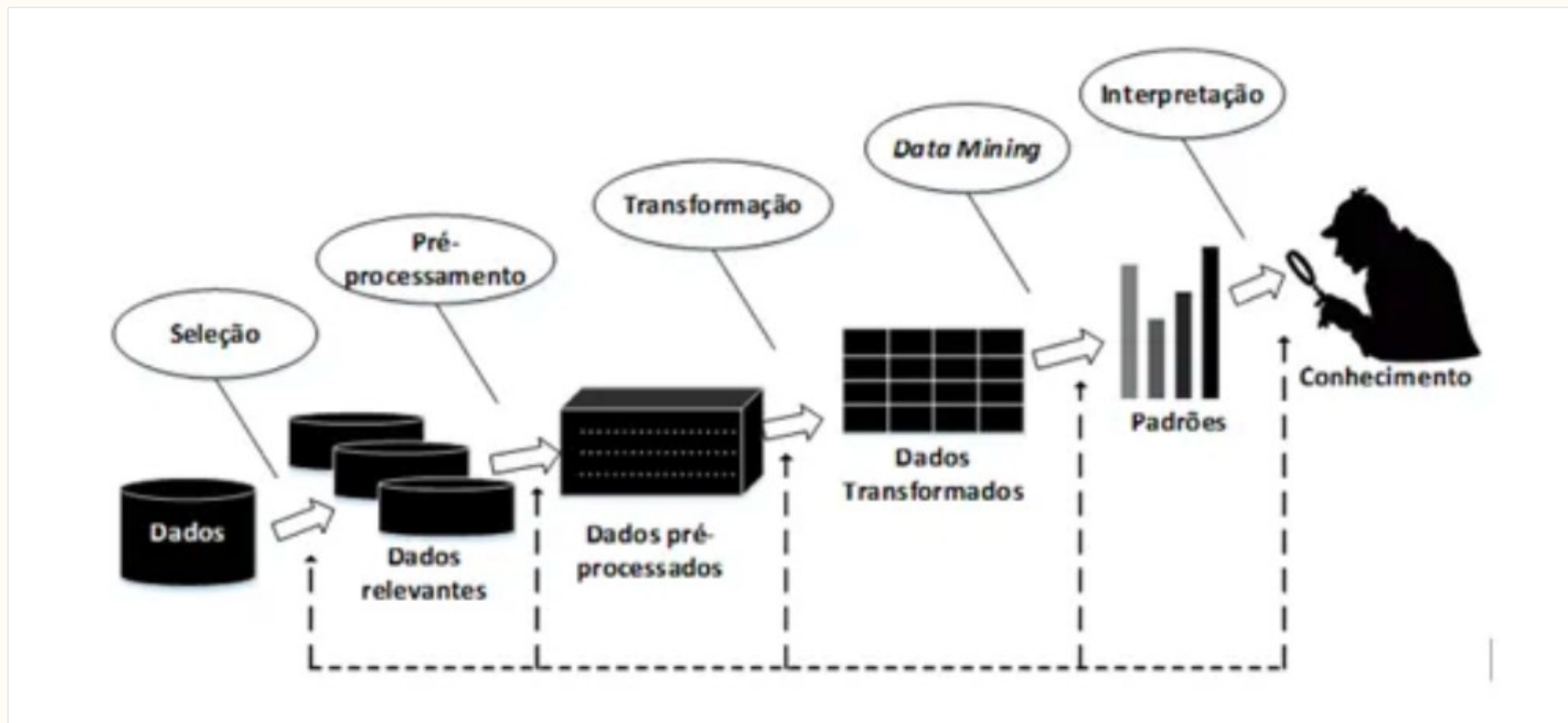


Jupyter → Pipeline (v1)



Fonte: DE DONNO (2019).

Jupyter → Pipeline (v2)



Contando uma história



Contando uma história

Texto

Código

Enron.ipynb

Arquivo Editar Ver Inserir Ambiente de execução Ferramentas Ajuda Última edição em 13 de junho

+ Código + Texto

Conectar + Gemini

```
#####
#####
#####
```

PERGUNTAS - HIPÓTESE LEVANTA

É possível encontrar uma relação das informações apenas por meio de preparação dos dados para análise e com as ações propostas nesse trabalho?

A resposta foi SIM. Fato disso é que a maioria do grupo considerado como POI, possui em REGISTROS que se enquadram como OUTLIERS, ou seja, se focarmos apenas em entender dados que tem um desvio padrão considerável no atributo analisado, já podemos entender melhor essa base de dados e encontrar correlação função dos dados de POI e os demais. 72% dos POIs, tem algum registro classificado como OUTLIER.

```
#####
#####
#####
```

Importando as bibliotecas

Na célula abaixo, as principais bibliotecas para análise de dados são importadas. Sinta-se a vontade para importar mais bibliotecas de acordo com sua demanda e/ou preferência, contudo, garanta que todos os comandos de `import` sejam feitos na célula abaixo.

```
[ ] import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline

import missingno as msno

# carregue os módulos de sua preferência aqui
```

Carregamento de dados

Na célula abaixo é realizado o carregamento da base de dados. Você **não** deve alterar o código apresentado.

```
[ ] DataBase = pd.read_pickle('final_project_dataset.pkl')
```

Verificando se os dados carregaram corretamente:

```
[ ] DataBase.head(3)
```

```
AttributeError                                Traceback (most recent call last)
<ipython-input-7-9f6f76304c73> in <cell line: 1>()
----> 1 DataBase.head(3)

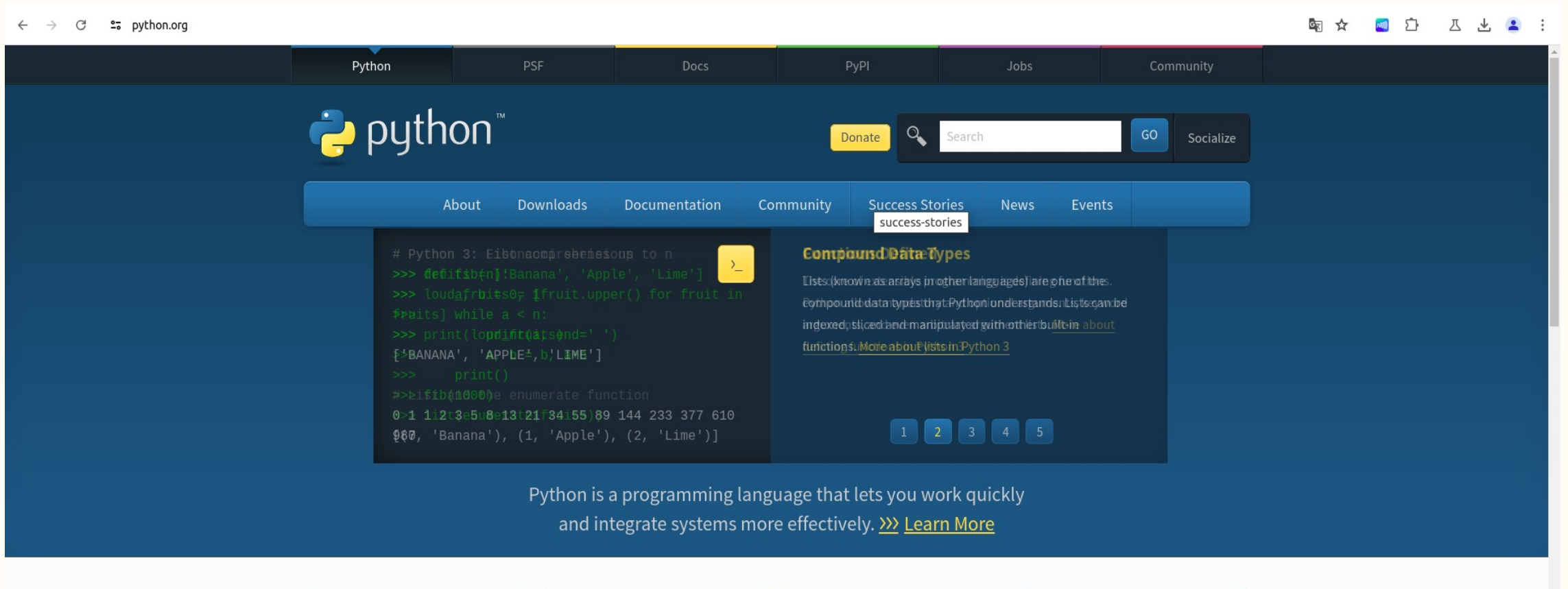
AttributeError: 'dict' object has no attribute 'head'
```

```
[ ] # Alterar os nomes das colunas usando o método set_axis
DataBase = DataBase.set_axis(['2.2 Bônus',
                              '2.6 RemuneraçãoAdiantados',
                              '2.7 RendaAdiantados',
                              '2.3 HonoráriosAdministradores',
                              '4.1 EndereçoEmail',
                              '3.3 ComprasAçõesEmExercício',
                              '2.8 Despesas',
                              '4.2 MsgEnviados',
                              '4.6 Msg_POI>PESSOA',
                              '4.5 Msg_PESSOA>POI',
                              '2.5 EmpréstimosAdiantamento',
                              '2.4 IncentivoDeLongoPrazo',
                              '2.9 Outros',
                              '3.2 POI',
```

Texto

Código

Python



Mas é melhor usar Python?

A boa linguagem é a que resolve ! Mas...existem muitos frameworks.

- TensorFlow
- OpenCV
- Scikit
-

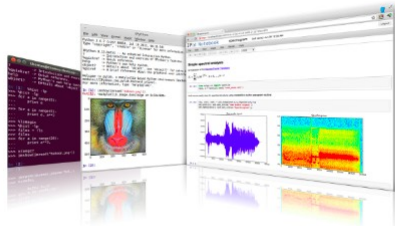
iPython

IP[y]: IPython Interactive Computing

[Install](#) · [Documentation](#) · [Project](#) · [Jupyter](#) · [News](#) · [Cite](#) · [Donate](#) · [Books](#)

IPython provides a rich architecture for interactive computing with:

- A powerful interactive shell.
- A kernel for [Jupyter](#).
- Support for interactive data visualization and use of [GUI toolkits](#).
- Flexible, [embeddable](#) interpreters to load into your own projects.
- Easy to use, high performance tools for [parallel computing](#).



To get started with IPython in the Jupyter Notebook, see our [official example collection](#). Our [notebook gallery](#) is an excellent way to see the many things you can do with IPython while learning about a variety of topics, from basic programming to advanced statistics or quantum mechanics.

To learn more about IPython, you can download our [talks and presentations](#), or read our [extensive documentation](#). IPython is open source (BSD license), and is used by a range of [other projects](#); add your project to that list if it uses IPython as a library, and please don't forget to [cite the project](#).

IPython supports Python 2.7 and 3.3 or newer. Our older 1.x series supports Python 2.6 and 3.2.

Jupyter and the future of IPython

IPython is a growing project, with increasingly language-agnostic components. IPython 3.x was the last monolithic release of IPython, containing the notebook server, qtconsole, etc. As of IPython 4.0, the language-agnostic parts of the project: the notebook format, message protocol, qtconsole, notebook web application, etc. have moved to new projects under the name [Jupyter](#). IPython itself is focused on interactive Python, part of which is providing a Python kernel for Jupyter.

Announcements

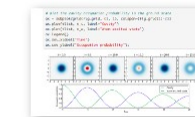
Loading

JUPYTERCON



NOTEBOOK
VIEWER

Share your notebooks



COMMUNITY

[Stack Overflow](#)
[Mailing list](#)
[File a bug](#)
[Reddit](#)

FOR DEVELOPERS

[Mailing list](#)
[Development Chat Room](#)
[Development information](#)
[Travis CI](#)
[Wiki](#)

BOOK



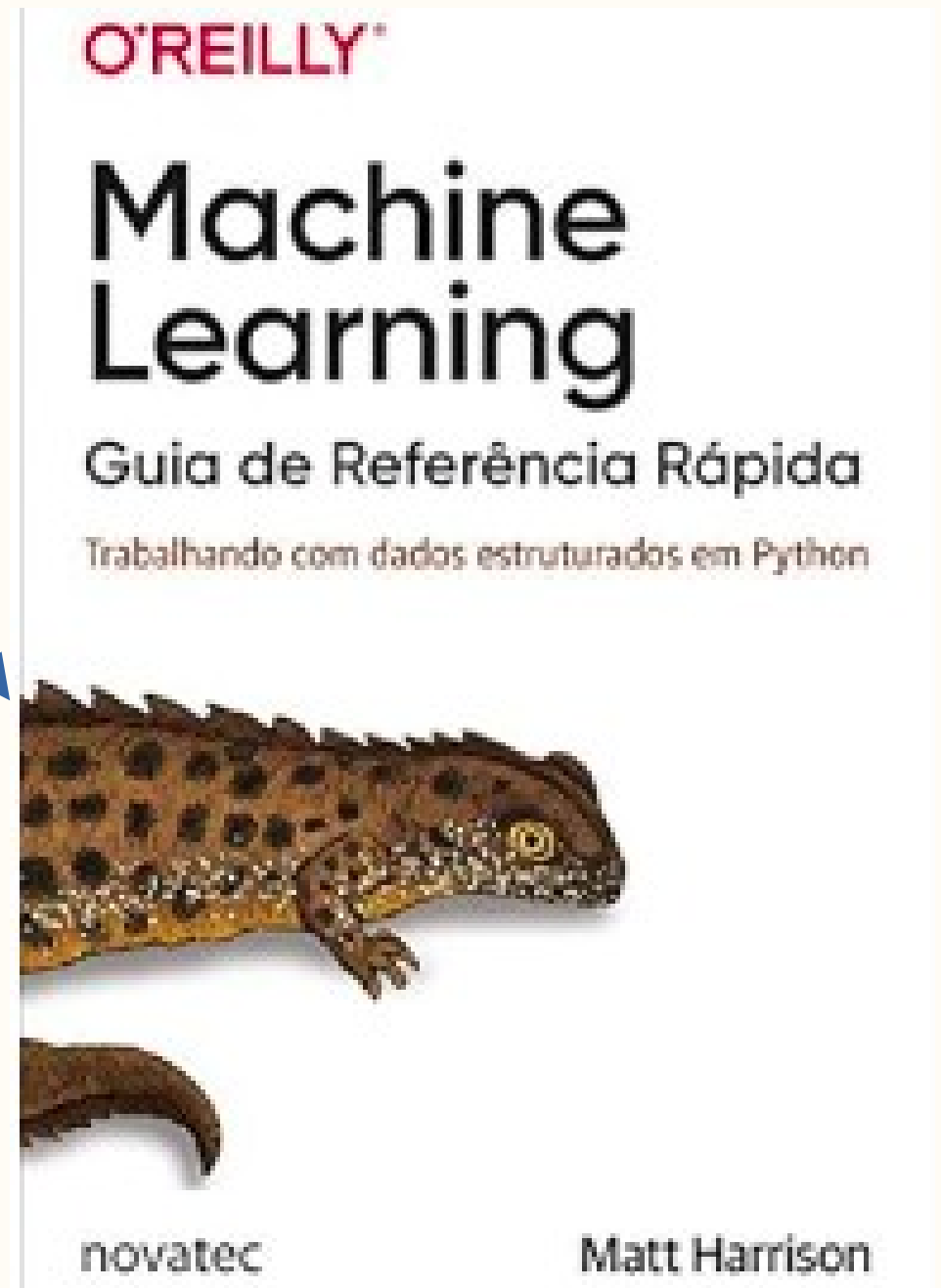
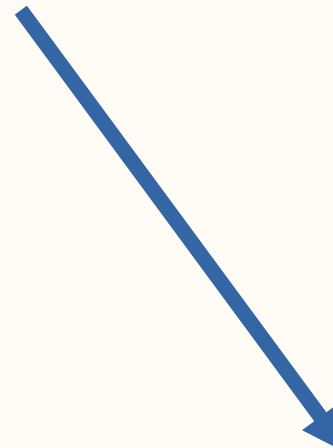
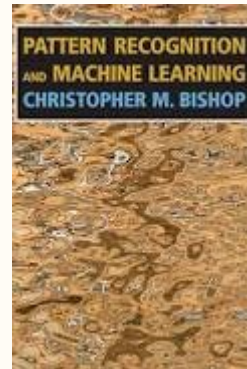
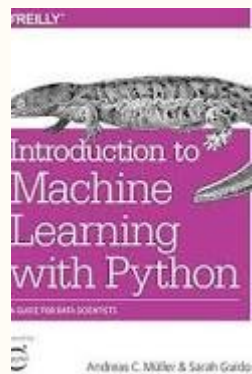
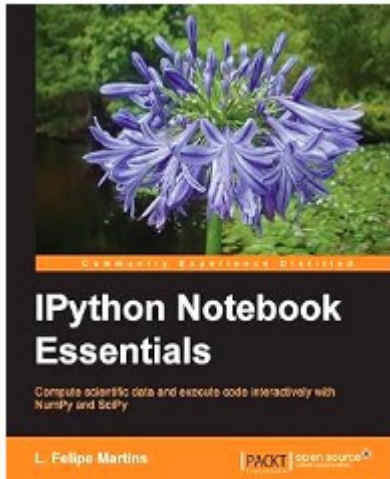
→ Jupyter

→ VsCode

→ Outras
IDEs...

Exemplos de Aplicações Acadêmicas

Livros

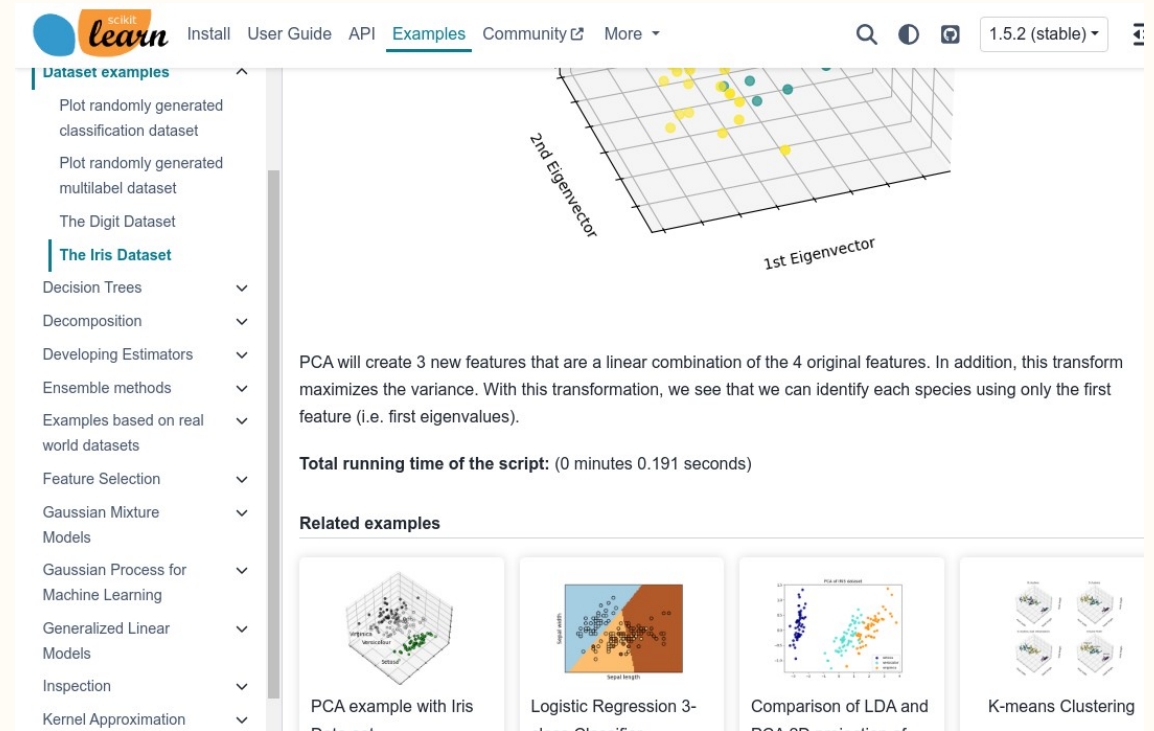
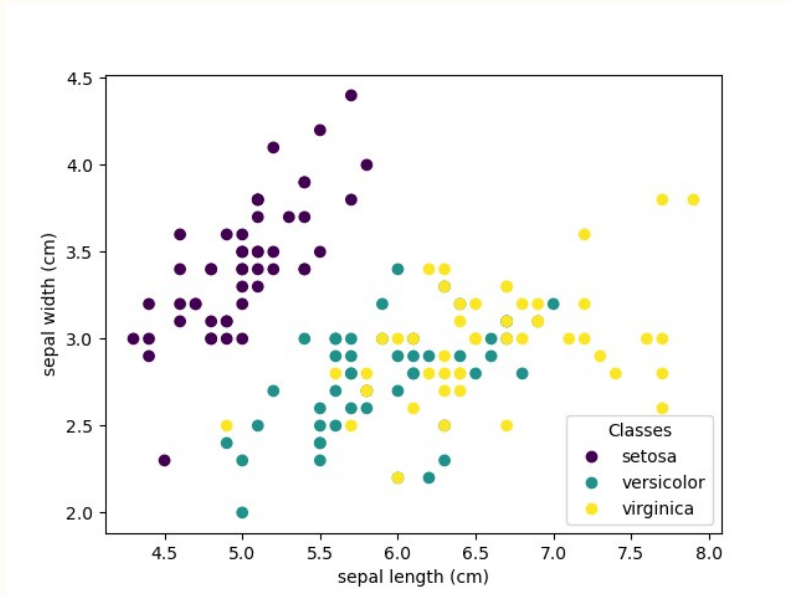


Case Enron - Sarbane-Oxley

A Enron foi uma companhia de energia e gás natural que se tornou um caso de referência mundial em fraudes contábeis. A empresa manipulou os números trimestrais e subornou instituições que deveriam fiscalizar as escrituras contábeis. A Enron contabilizava os ganhos de contratos de longo prazo na receita recorrente, otimizando os rendimentos projetados.

A Lei Sarbanes-Oxley (SOx) foi criada nos Estados Unidos em 2002 para combater fraudes financeiras e proteger os investidores, como resposta aos escândalos envolvendo empresas como a Enron.

Iris - Dataset



<https://github.com/1akshat/Iris-Dataset-Python-Notebook-Solution/blob/master/Iris.ipynb>

Fontes

Google Drive

<https://drive.google.com/drive/folders/1Zh7Nmrmyx-gyuPWdW7zS7flZxkhAEqNk?usp=sharing>

GitHub

<https://github.com/mferste/jupyteresenac>

Obrigado!

mauricio.ferste@docente.pr.senac.br

