### **Jhonathan Ferroni – SP3158403**

### **1. O que é o Jupyter Notebook?**

* O Jupyter Notebook é uma aplicação web interativa que permite criar e compartilhar documentos que contêm código, visualizações, texto explicativo, equações matemáticas e outros tipos de conteúdo multimídia.
* Ele suporta várias linguagens de programação, mas é mais amplamente usado com Python, R e Julia.
* O nome "Jupyter" vem de três linguagens principais: Julia, Python e R.

### **2. Para que serve?**

* É amplamente utilizado em ciência de dados, aprendizado de máquina, análise de dados e desenvolvimento de software.
* Permite combinar código executável com explicações detalhadas, gráficos e visualizações interativas, facilitando a apresentação e o compartilhamento de projetos.
* Comumente usado para experimentação e prototipagem rápida de ideias de programação e análise.
* É uma ferramenta excelente para documentação técnica ou apresentações de pesquisa e projetos.

### **3. Como instalar e começar a programar?**

#### **Instalação**

* A maneira mais simples de instalar o Jupyter Notebook é por meio de uma distribuição do Python como o Anaconda. O Anaconda vem com o Jupyter e várias outras bibliotecas prontas para uso.
* Você também pode instalá-lo diretamente com o pip:

pip install notebook

#### **Como iniciar**

* Após a instalação, você pode iniciar o Jupyter Notebook pelo terminal ou prompt de comando com:

jupyter notebook

* Isso abrirá uma interface web onde você pode criar novos notebooks e começar a programar. O Jupyter Notebooks organiza seu trabalho em células que podem conter código, texto ou gráficos.

### **4. Exemplo de aplicação**

#### **Cálculo de estatísticas básicas usando Python no Jupyter:**

import numpy as np  
  
# Gerando um array de números aleatórios  
data = np.random.randn(1000)  
  
# Calculando estatísticas  
media = np.mean(data)  
desvio\_padrao = np.std(data)  
mediana = np.median(data)  
  
print(f"Média: {media}")  
print(f"Desvio Padrão: {desvio\_padrao}")  
print(f"Mediana: {mediana}")

Neste exemplo, o Jupyter é usado para gerar dados aleatórios, calcular estatísticas como média, desvio padrão e mediana, e apresentar os resultados.

### **5. Exemplo de aplicação real**

Uma aplicação muito comum do Jupyter Notebook é em projetos de análise de dados e aprendizado de máquina. Um exemplo real disso é o uso do Jupyter para análises de grandes bases de dados publicamente compartilhados em plataformas como **Kaggle**.

#### **Predição de preços: aplicação no Kaggle**

Uma aplicação real do Jupyter Notebook pode ser vista no **Housing Prices Competition** no Kaggle, onde os desenvolvedores usam notebooks para construir modelos de machine learning e prever preços de casas com base em diversos dados históricos. O código a seguir é um exemplo de um projeto que utiliza o Jupyter Notebook para análise de dados e treinamento de modelos.

import pandas as pd

from sklearn.model\_selection import train\_test\_split

from sklearn.ensemble import RandomForestRegressor

from sklearn.metrics import mean\_squared\_error

# Carregando o dataset de preços de casas

data = pd.read\_csv("house\_prices.csv")

# Separando as features (características) e o target (preço)

X = data.drop(columns=["SalePrice"])

y = data["SalePrice"]

# Dividindo o dataset em treino e teste

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

# Criando o modelo de Regressão RandomForest

model = RandomForestRegressor()

# Treinando o modelo

model.fit(X\_train, y\_train)

# Fazendo previsões

predictions = model.predict(X\_test)

# Avaliando o modelo

mse = mean\_squared\_error(y\_test, predictions)

rmse = mse \*\* 0.5

print(f"Root Mean Squared Error: {rmse}")

**Fonte:**

Kaggle - House Prices: Advanced Regression Techniques

https://www.kaggle.com/c/house-prices-advanced-regression-techniques/overview

Link para o Notebook de exemplo no GitHub:

House Price Prediction on GitHub

https://github.com/ageron/handson-ml2/blob/master/02\_end\_to\_end\_machine\_learning\_project.ipynb

**Explicação da aplicação:**

- Esse projeto carrega um conjunto de dados de preços de casas, treina um modelo de aprendizado de máquina (Random Forest) para prever os preços, e depois avalia o desempenho do modelo.

- O Jupyter Notebook é ideal para visualizar os dados com gráficos interativos e acompanhar o fluxo do código, explicações e resultados.

**Aplicação no GitHub:**

No GitHub, vários desenvolvedores publicam projetos similares, onde notebooks são usados para documentar e compartilhar todas as etapas da análise e modelagem. Um exemplo real de aplicação desse tipo está no repositório do livro **Hands-On Machine Learning with Scikit-Learn and TensorFlow**:

Hands-On Machine Learning - GitHub Repository

https://github.com/ageron/handson-ml2