

TREINAMENTO
FORMAÇÃO
DATA
EXPERTS



Fatec



FASTER
MAKE SMARTER DECISIONS
FASTER

TREINAMENTO
FORMAÇÃO
DATA
EXPERTS

Análise de Dados com python



Fernando Barbieri
DataEX



O que é Python?

Uma das
linguagens de
programação mais
poderosas

Open Source e
grátis

Múltiplos casos de
uso



Características do Python?

Vários paradigmas

Interpretada
Multi-plataforma

Grande suporte de
bibliotecas e
pacotes



Por que Python?

Grande variedade
de pacotes e
bibliotecas

Grátis

Beginner friendly



Estruturas de Dados Python

LISTAS:

```
['Fernando', 'Maria', 'José']
```

Mutável, ordenada, permite duplicidade

TUPLAS:

```
('Fernando', 'Maria', 'José')
```

Imutável, ordenada, permite duplicidade

SETS:

```
{'Fernando', 'Maria', 'José'}
```

Imutável, não ordenada, não permite duplicidade



Pacotes para Análise de Dados

NumPy

Matplotlib

Pandas

SciPy

... entre muitos outros.



IDEs para Python

PyCharm

Visual Studio Code

Databricks

Jupyter

... entre muitas outras.



O que é Jupyter?

Interface e
bibliotecas para
documentos
interativos Python

Muito utilizado
para Data Science

Permite explorar,
processar e analisar
dados de maneira
conveniente

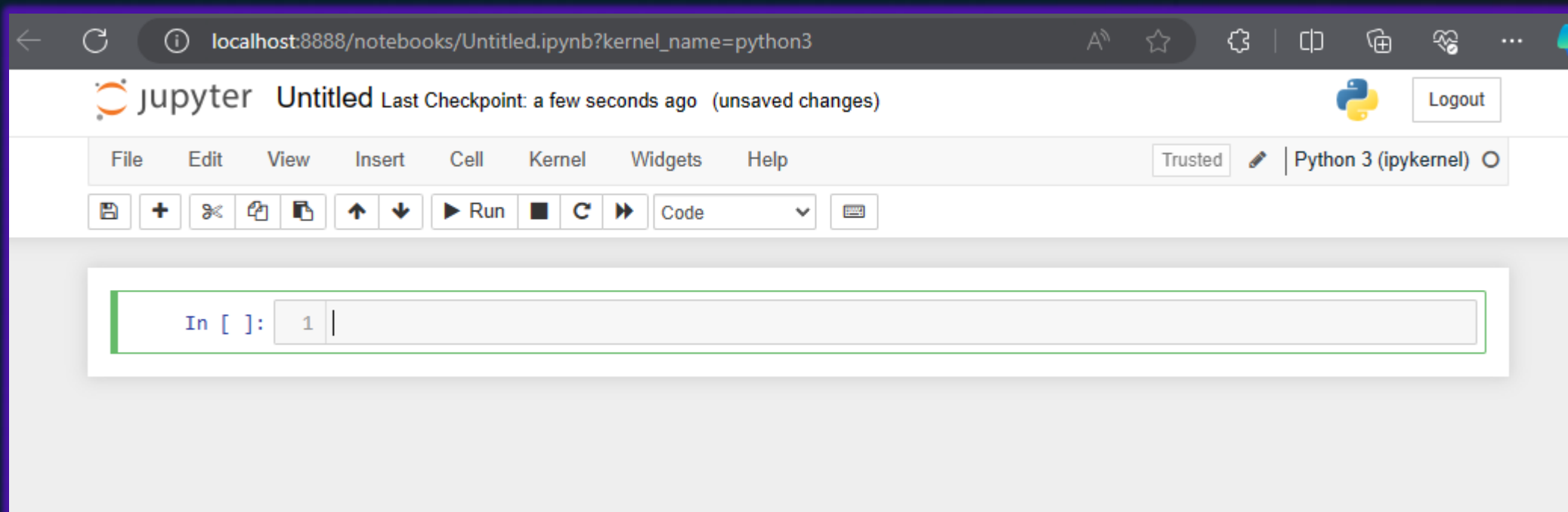


Jupyter Notebooks

Arquivos com extensão **.ipynb** e hospedados em um **servidor** de notebooks Jupyter, que podem ser modificados e executados usando uma **interface web**.



Jupyter Notebooks

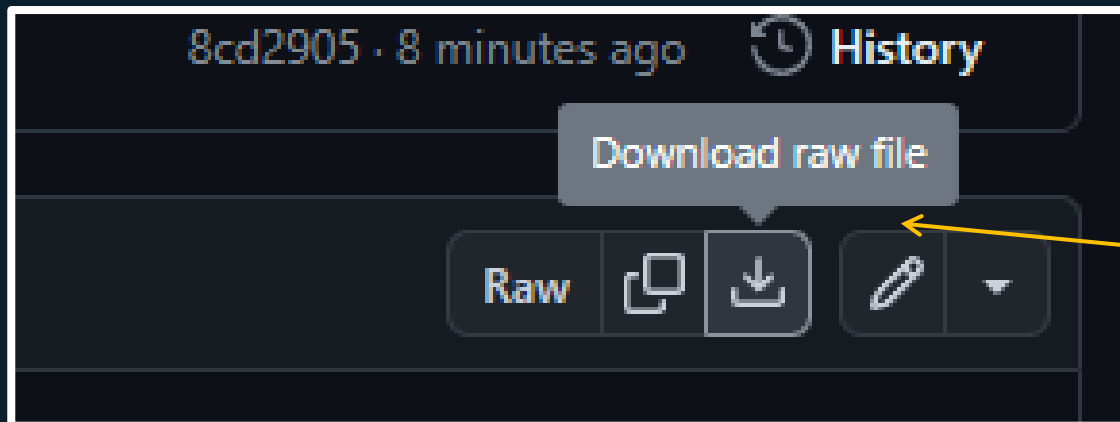


Arquivos Curso

github.com/ferbarbieri/fatec_dataexperts



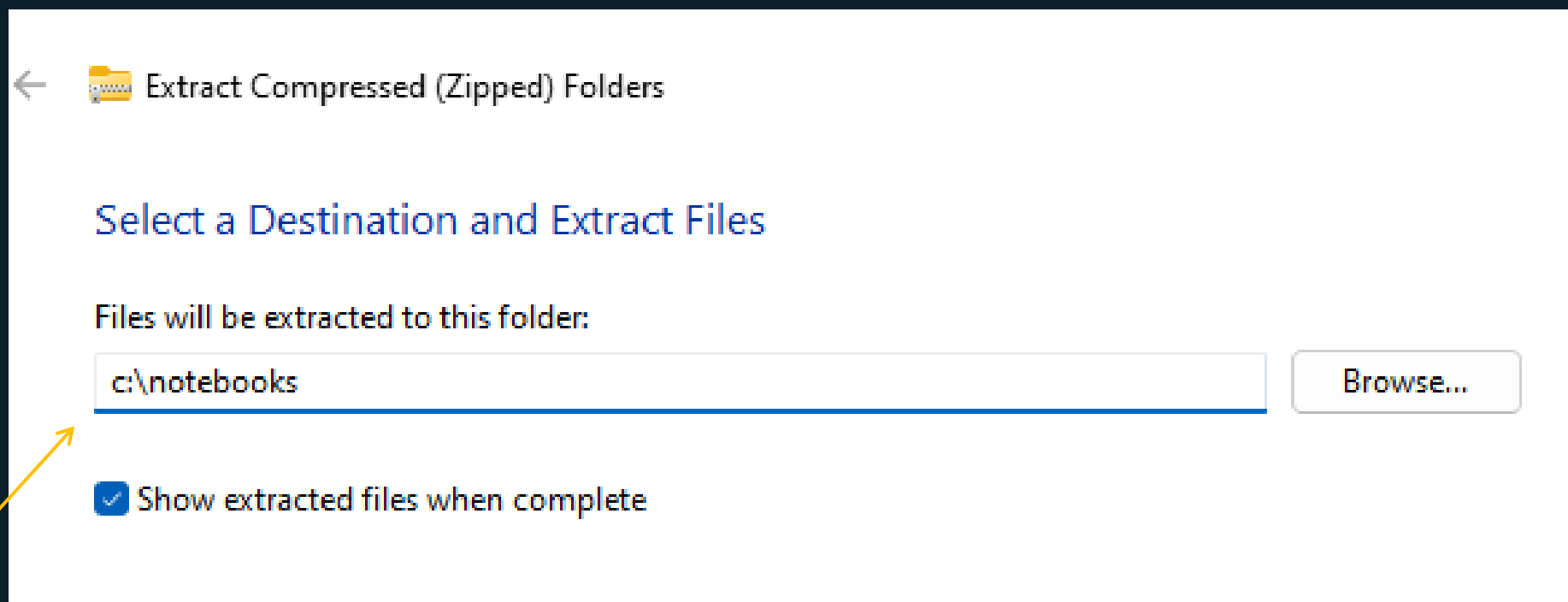
Baixa esse zip



Aqui



Extraia para C:\notebooks



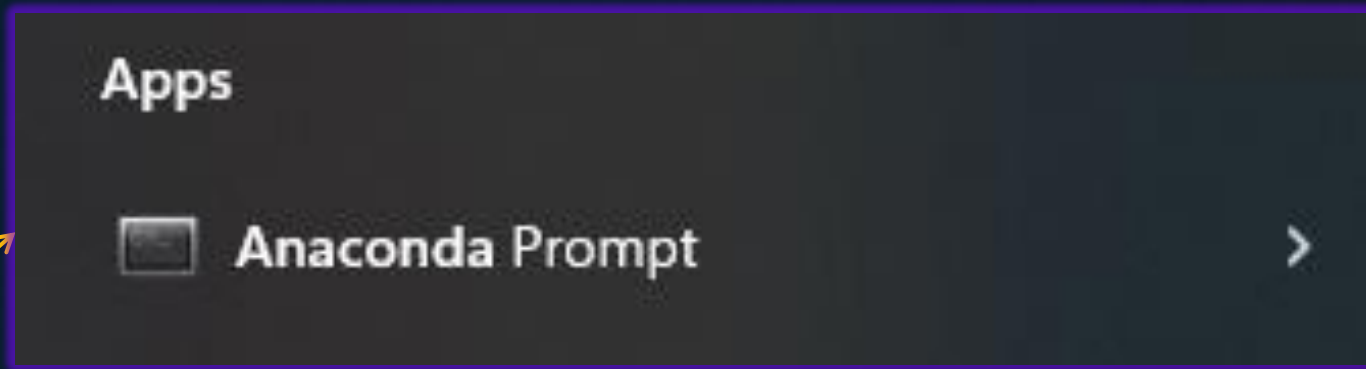
Aqui



Arquivos Curso

Abrir Anaconda Prompt

Esse



Arquivos Curso

Digite:

```
> cd \notebooks
```

```
> jupyter notebook .
```



TREINAMENTO
FORMAÇÃO
DATA
EXPERTS

Demo

Jupyter Notebooks





DESAFIO!

*Em um notebook Jupyter,
escreva a **Tabuada do 8**.*

Tempo: 3 minutos

Levante a mão quando terminar!

```
8 x 0: 0
8 x 1: 8
8 x 2: 16
8 x 3: 24
8 x 4: 32
8 x 5: 40
8 x 6: 48
8 x 7: 56
8 x 8: 64
8 x 9: 72
8 x 10: 80
```



O que é o Pandas

Biblioteca para
Manipulação de
Dados

Contém estruturas
de dados eficientes

Contém utilidades
para limpeza,
agregação e
visualização

https://pandas.pydata.org/docs/getting_started/intro_tutorials/



Data Frame

Um objeto bi-dimensional usado para representar uma estrutura de dados tabular.

Mutável

Index Based

Classe



Data Frame

Utilizado para:

Representação
Python de dados de
outros formatos

Análise de dados
exporatória

Agregação e
agrupamento de
dados

JSON
CSV
XML
PARQUET
Etc...

Transformação de
dados

Datasets para
Machine Learning



TREINAMENTO
FORMAÇÃO
DATA
EXPERTS

Demo

Pandas

https://pandas.pydata.org/docs/getting_started/intro_tutorials/





DESAFIO!

1- Carregue o arquivo: *TABELA_VENDAS.CSV*

2- Responda: Quais os modelos de *BICICLETAS* são vendidas na cor *VERMELHA*?

Tempo: 15 minutos

Levante a mão quando terminar!



TREINAMENTO
FORMAÇÃO
DATA
EXPERTS

Demo

+ Pandas : Análises e Gráficos



TREINAMENTO
FORMAÇÃO
DATA
EXPERTS

Demo

+ Pandas : Análise Preditiva



TREINAMENTO
FORMAÇÃO
DATA
EXPERTS

Demo

<https://www.kaggle.com/>

k

TREINAMENTO
FORMAÇÃO
DATA
EXPERTS

