



Pandas com Jupyter

Configuração

```
pip3 install jupyterlab
```

- instalado Jupyter Lab com pip3
- instalado pandas com pip3

```
pip3 install pandas
```

- rodando jupyter

```
jupyter-lab
```

Arquivo do projeto

- arquivo baixado do movie lens do group lens
- arquivo 'ratings.csv' carregado para pasta do projeto
- projeto rastreado no git

```
git init
git add -A
git commit -m "mensagem"
```

- projeto publicado no github

Comandos

```
import pandas as pd # importando o pandas, já instalado na máquina com o pip3
pd.read_csv('ratings.csv') # lendo o arquivo csv com o pandas
notas = pd.read_csv('ratings.csv') # colocando dentro de uma variável e lendo apenas as top 5 linhas
notas.head()
notas.shape # mostra o formato do dataframe em linhas e colunas
notas.columns = ['usuárioId', 'filmeId', 'notas', 'momento'] # alterando títulos das colunas
notas['notas'] # mostrando apenas uma coluna específica
notas['notas'].head()
notas['notas'].unique() # mostra os valores que apareceram pelo menos uma vez na coluna
notas['notas'].value_counts() # mostra, conta e ordena todos os valores que apareceram na coluna
notas['notas'].mean() # média dos valores dessa série (coluna)
```