

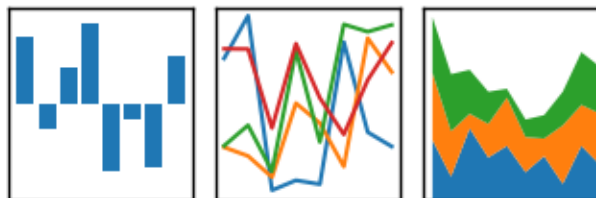


LISBON
SCHOOL OF
ECONOMICS &
MANAGEMENT
UNIVERSIDADE DE LISBOA

matplotlib

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



CARLOS J. COSTA

SM StatsModels
Statistics in Python



Exercicio

- Importar dados do ficheiro factbook.csv
- Criar dataframe só com país
- Qual o país com maior número de palavras no nome

StatsModels

- Utilizando a biblioteca statsmodels (<http://www.statsmodels.org/dev/examples/>), criar uma regressão utilizando os dados do ficheiro factbook.csv
- Identifique possíveis variáveis dependentes e variáveis independentes
- Atenção, porque poderá ter que fazer manipulação de dados

Scikit.learn

- Utilizando os dados do ficheiro `factbook.csv` e a biblioteca `scikit.learn` (<https://scikit-learn.org/stable/>) criar uma regressão.
- Identifique possíveis variáveis dependentes e variáveis independentes
- Atenção, porque poderá ter que fazer manipulação de dados

Matplotlib

- Utilizando os dados do ficheiro `factbook.csv` e a biblioteca `matplotlib` (<https://matplotlib.org/>), crie gráficos que permitam dar informação mais clara dos dados dos países.

Bibliografia

- <https://pandas.pydata.org/>
- https://pandas.pydata.org/pandas-docs/stable/getting_started/10min.html
- <https://scikit-learn.org/>
- <https://scikit-learn.org/stable/index.html>
- <https://www.statsmodels.org/stable/index.html>