

# Python para Inteligência Artificial

Jean Phelipe de Oliveira Lima

jpdol.eng16@uea.edu.br

/jpdol

# Análise e Visualização de Dados

- NumPy;
- Pandas;
- Matplotlib;

# NumPy

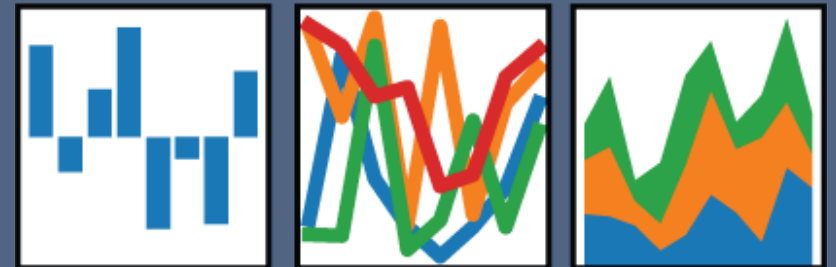
- NumPy é um pacote para a linguagem Python que suporta arrays e matrizes multidimensionais, possuindo uma larga coleção de funções matemáticas para trabalhar com estas estruturas.



# Pandas

- Biblioteca de software escrita para a linguagem de programação Python para manipulação e análise de dados. Em particular, oferece estruturas de dados e operações para manipular tabelas numéricas e séries temporais.

pandas

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


# Matplotlib

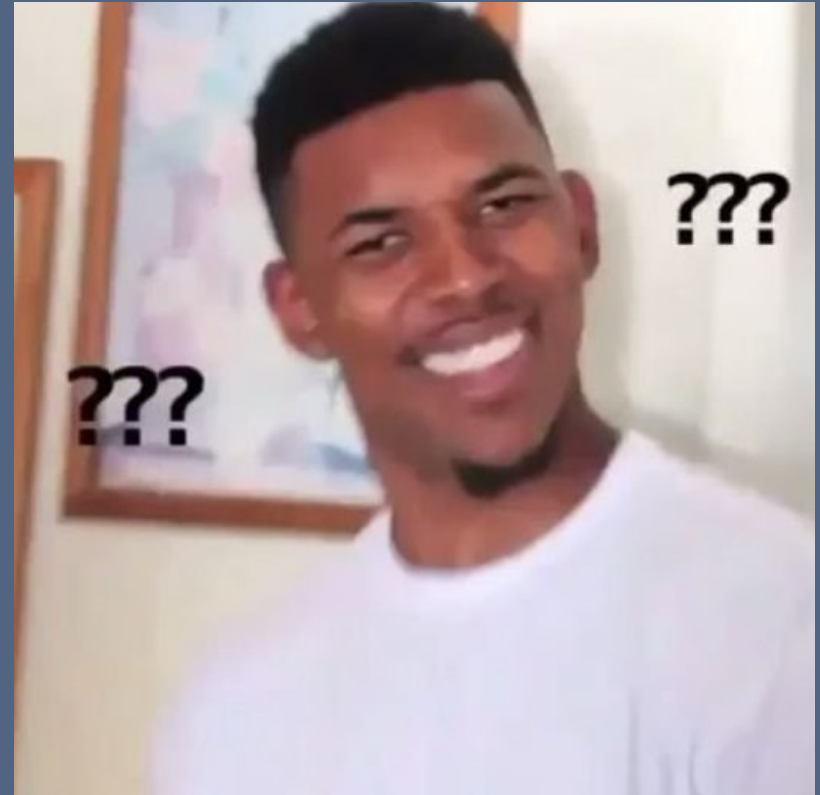
- Biblioteca de plotagem para a linguagem de programação Python e sua extensão de matemática numérica NumPy.



# Por que utilizar essas Bibliotecas?

- Explorar Dados
- Visualizar Dados
- Extrair Informações não triviais para o processamento humano
- Identificar correlações
- Etc

# Dados???



# Big Data

- Em tecnologia da informação, o termo big data refere-se a um grande conjunto de dados gerados e armazenados com os quais os aplicativos de processamento de dados tradicionais ainda não conseguem lidar em um tempo tolerável.



# # BIG DATA FACTS

HEKIMA



MAIS DE  
**12 TERABYTES**  
DE INFORMAÇÃO  
POR DIA



MAIS DE  
**3 BILHÕES**  
DE LIKES E  
COMENTÁRIOS  
POR DIA.

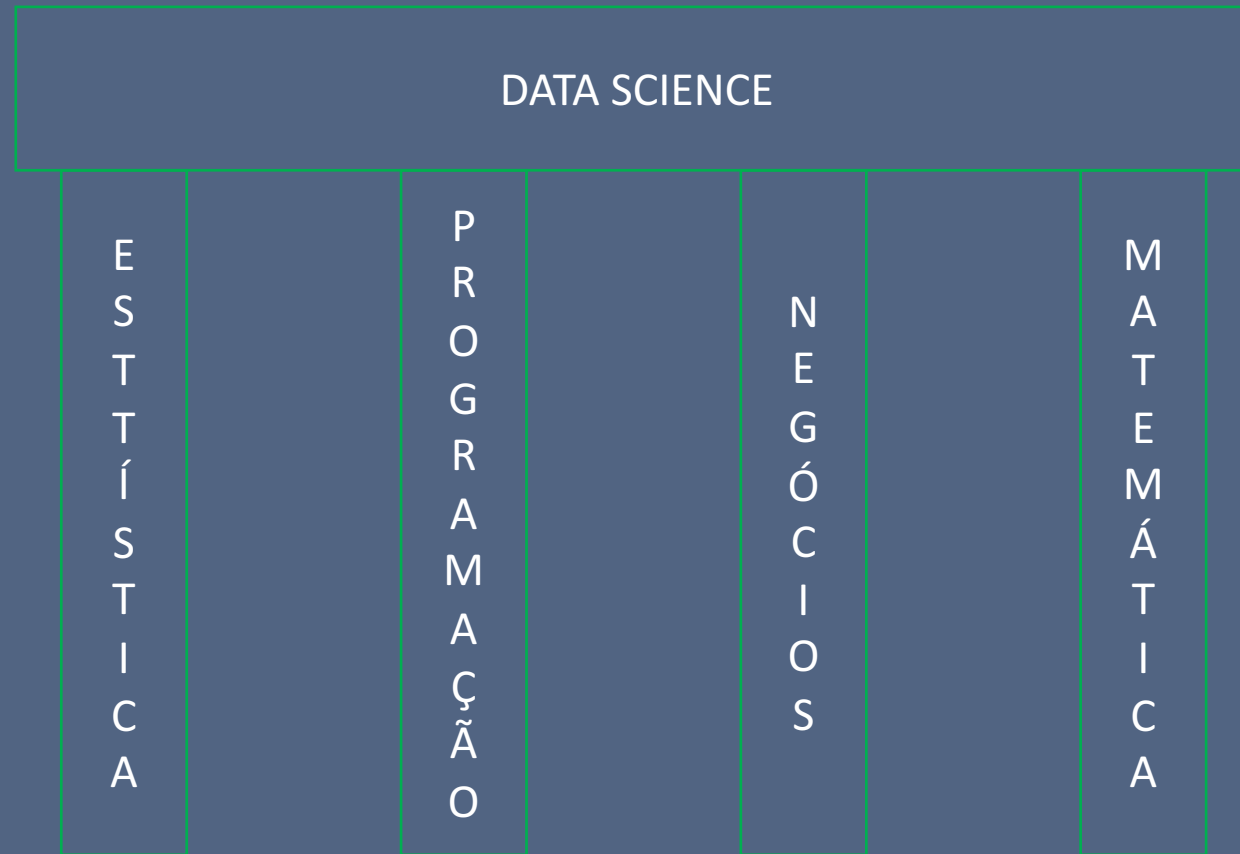


MAIS DE  
**24 PETABYTES**  
DE DADOS  
POR DIA

# Data Science

- É o estudo disciplinado dos dados e informações inerentes ao negócio e todas as visões que podem cercar um determinado assunto. É uma ciência que estuda as informações, seu processo de captura, transformação, geração e, posteriormente, análise de dados.

# Data Science



# Caso Walmart



# Caso Walmart

