

DATA SCIENCE...

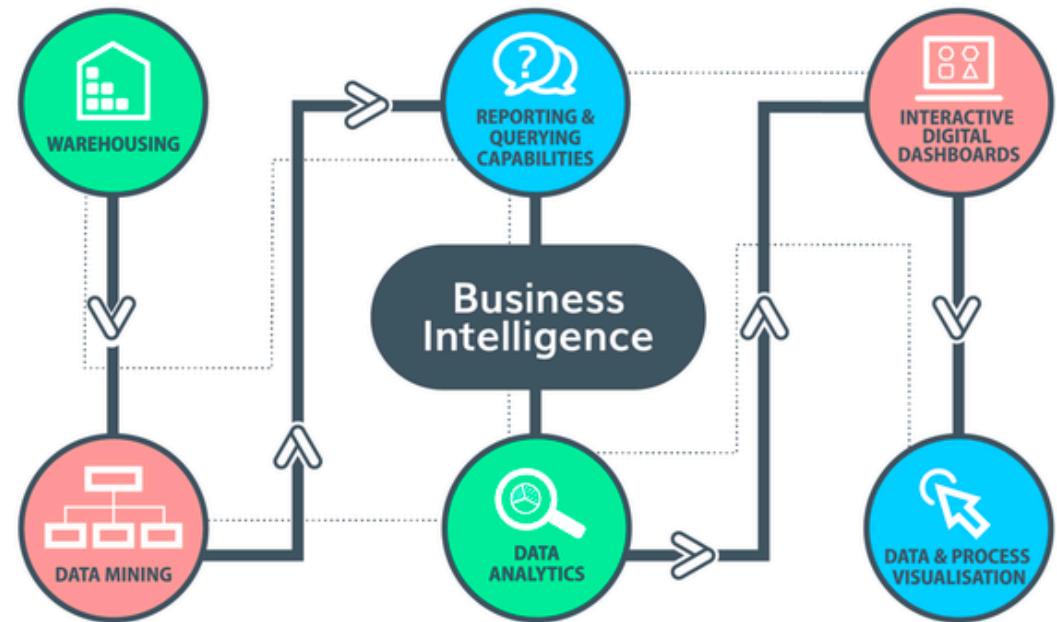
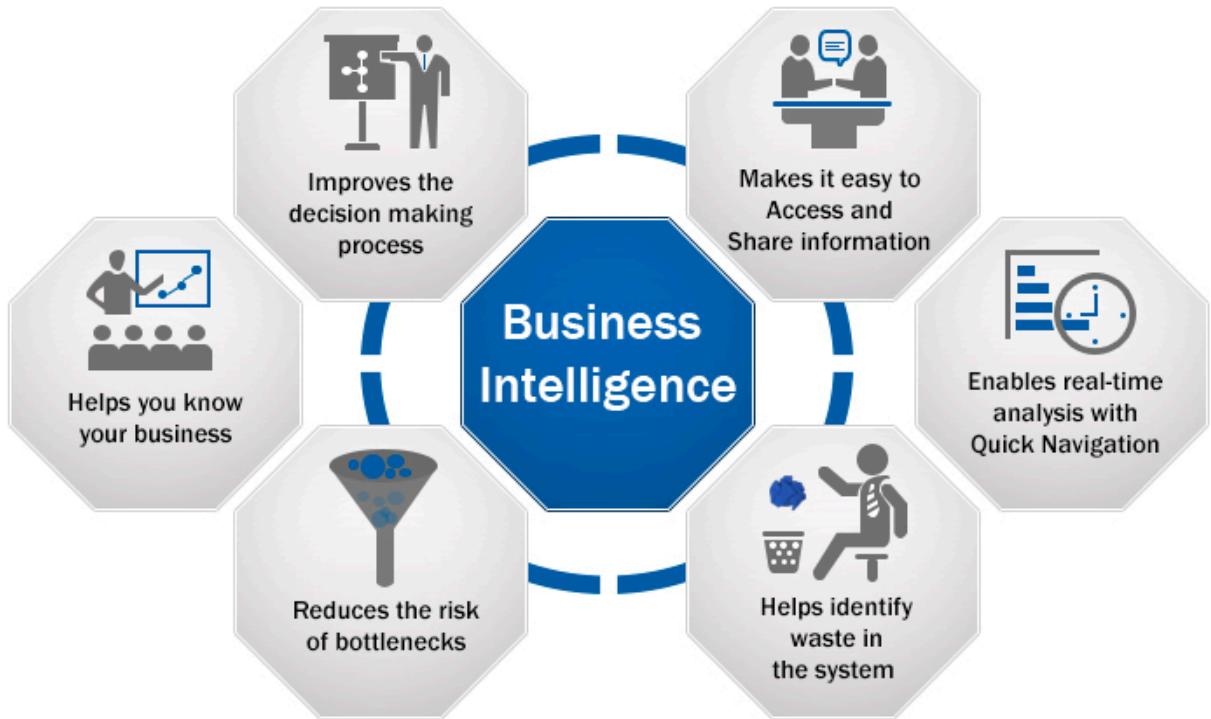
Data Scientist....

Desmistificando o trabalho mais sexy do século 21

1

2

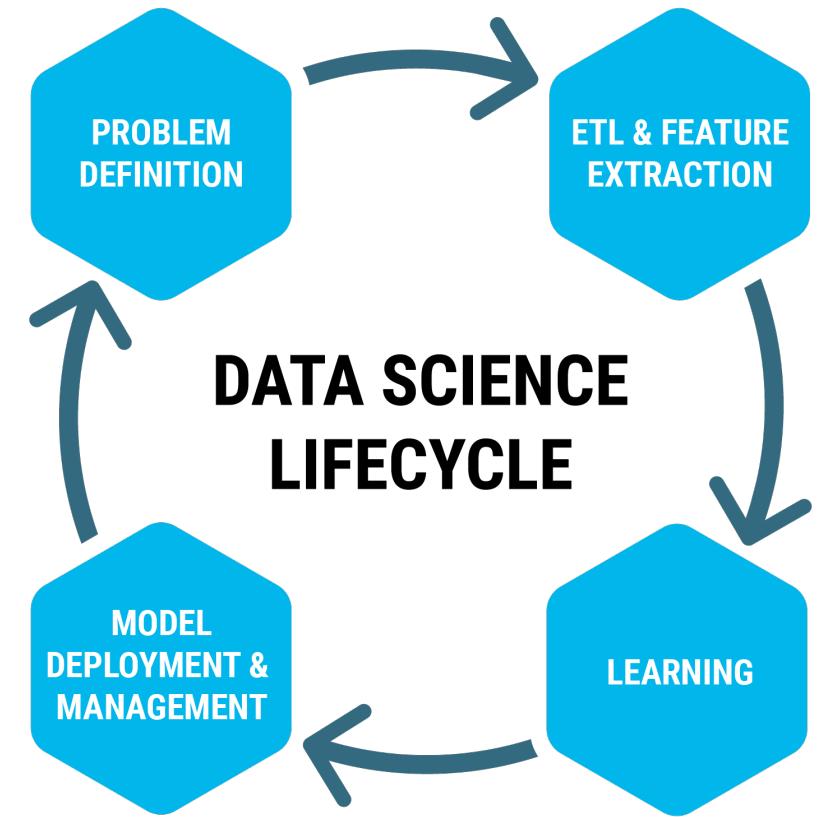
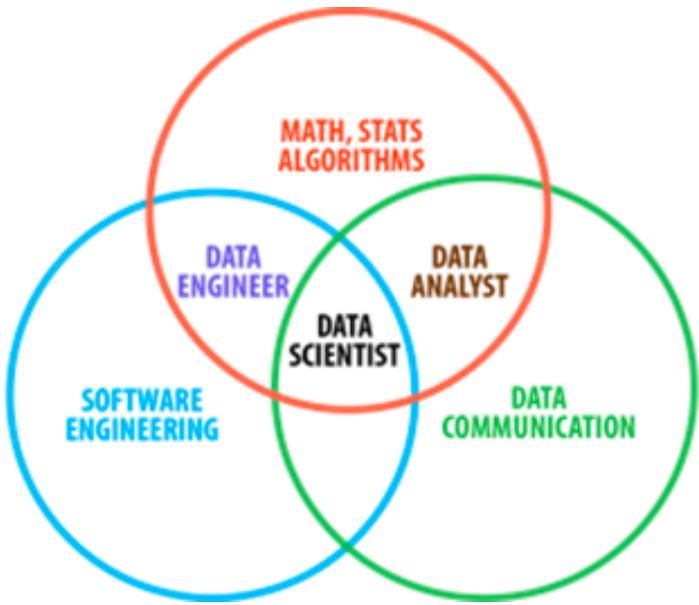
AUDIENCE INTROSPECTION



Business Intelligence

Existem Data Scientists/Cientistas de Dados ?

4





Prerequisite: Python for Data analysis

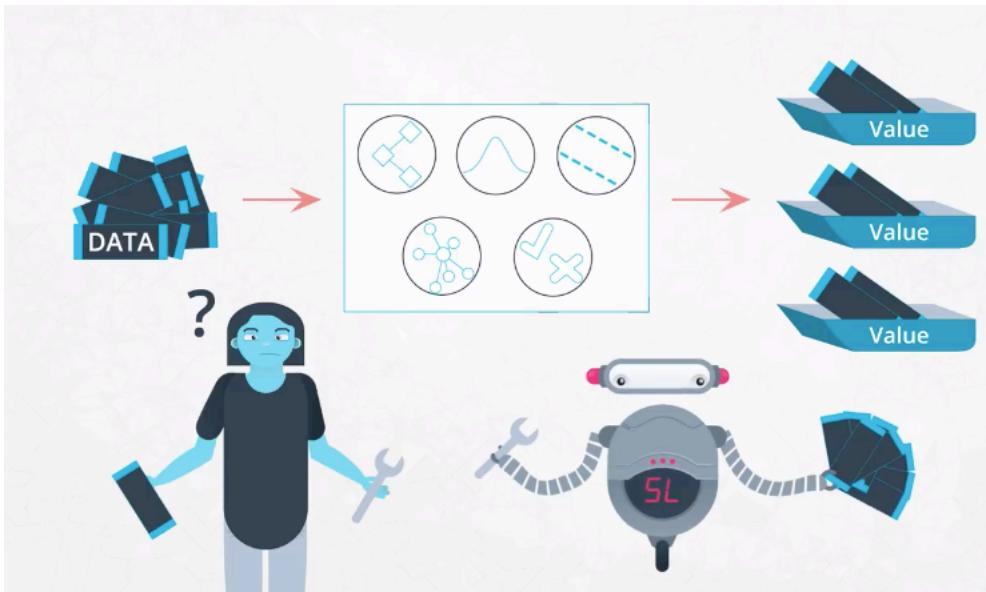
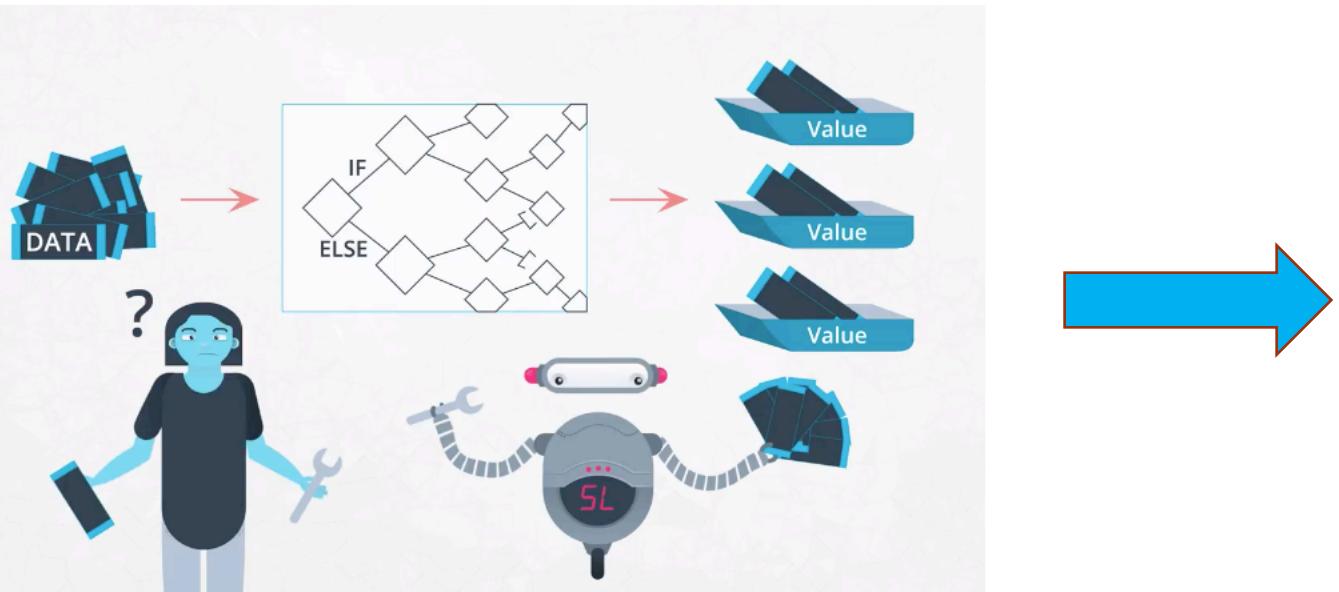
Prerequisite: SQL

Prerequisite: Data visualization

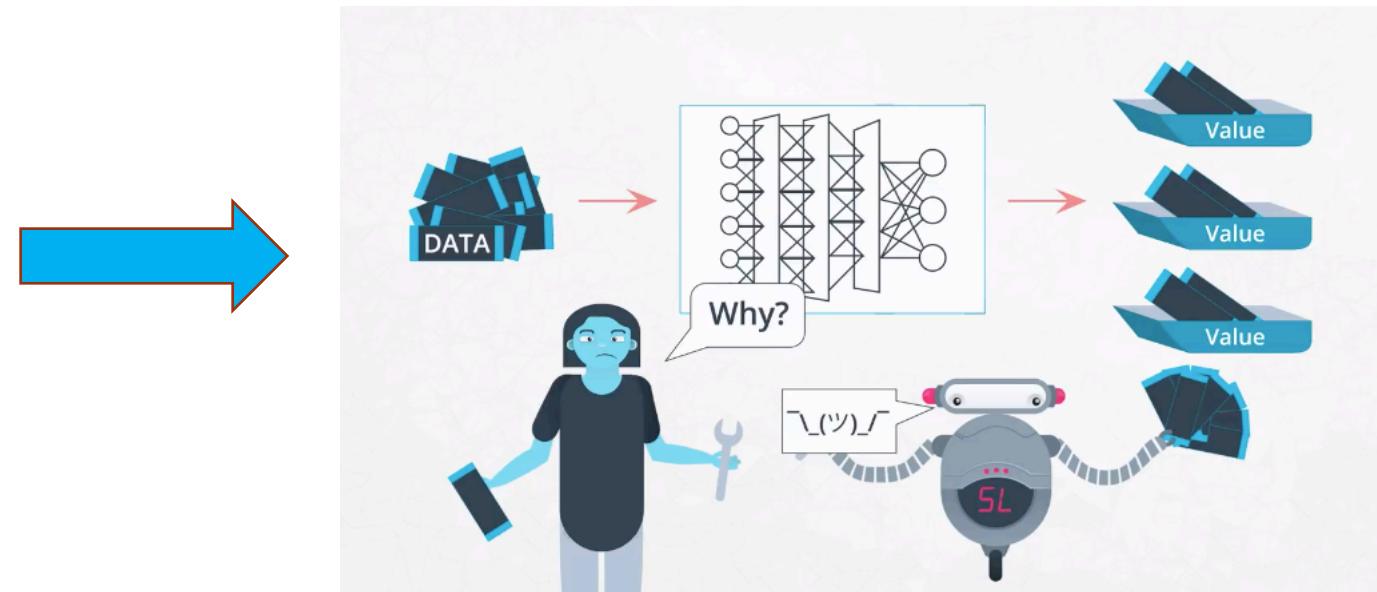
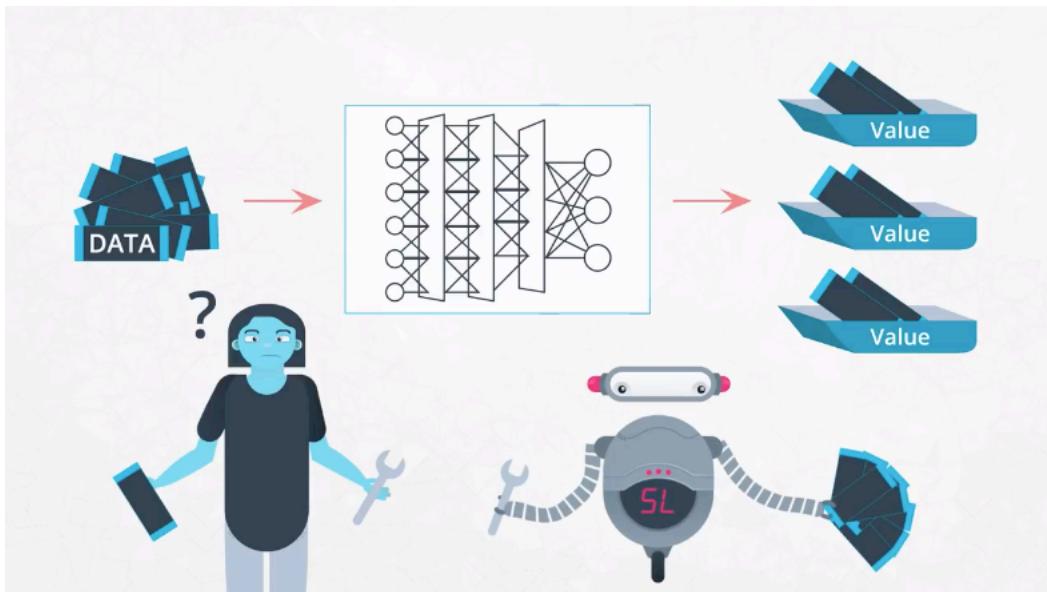
Prerequisite: Command line Essentials

Prerequisite: Git & GitHub/Bitbucket

Prerequisite: Linear Algebra



Computer Science perspective – Raw Data, Data Patterns, Probability Distributions



**Computer Science perspective – Prediction Models, Machine Learning,
Complex Implementations – Medical Diagnostic**

Supervised Learning

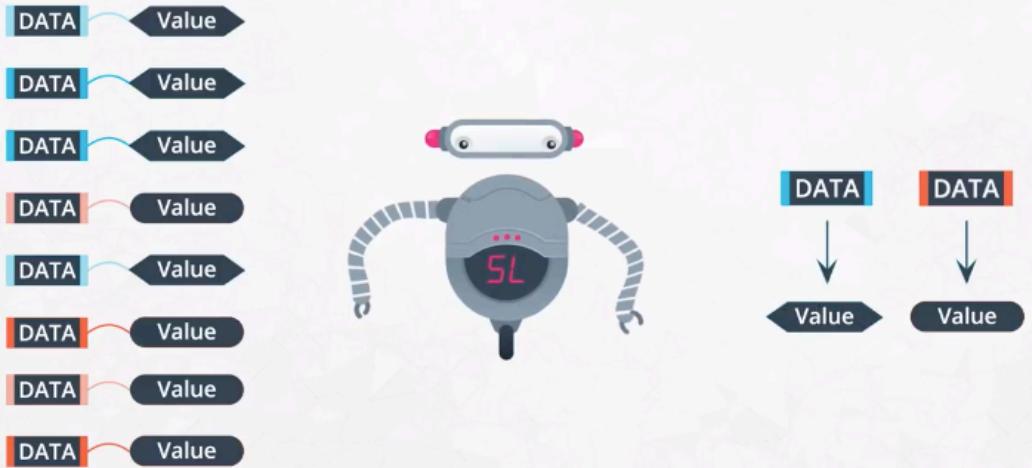
8

Unsupervised Learning

Reinforcement Learning

Types of Machine Learning

Supervised Learning



Supervised Learning

Classification

Categorical Outcomes



Regression

Numeric Outcomes



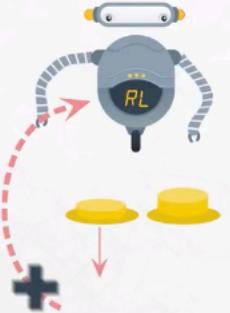
Supervised Learning – Classifying emails, or Races of animals, and predicting House Prices, Loans Outcomes or Physical characteristics

10



Unsupervised Learning – Grouping Items, Music Recommendation

Reinforcement Learning

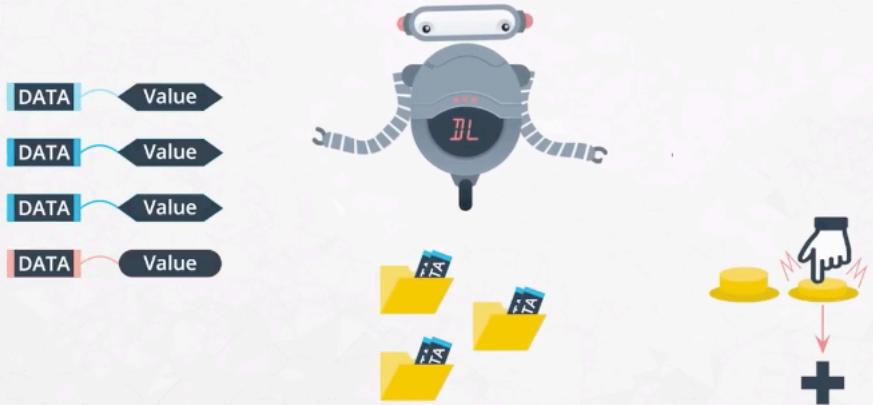


Reinforcement Learning

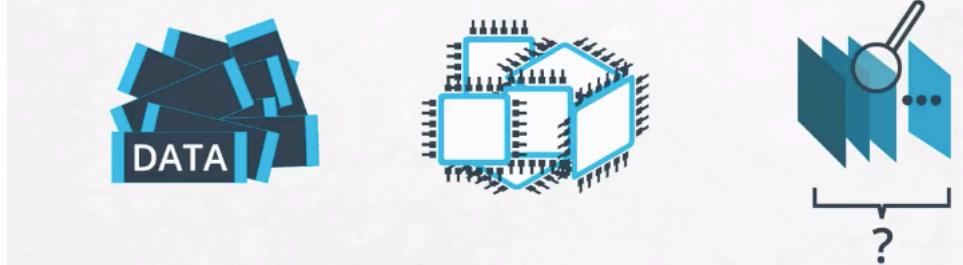


Reinforcement Learning – Learn on rewards for actions, self-driving vehicles, game agents

Deep Learning



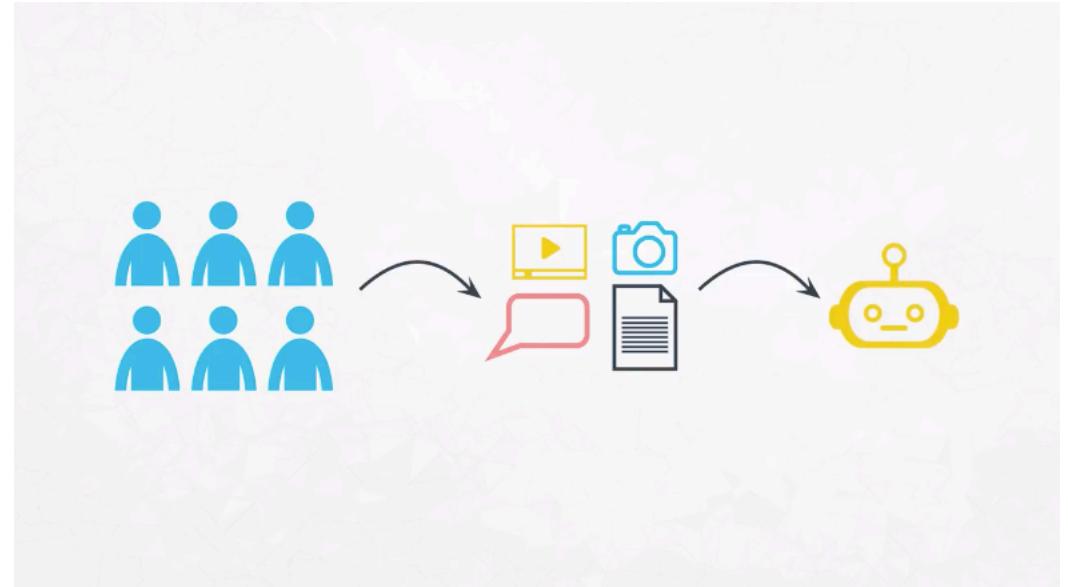
Barriers to Deep Learning



Deep Learning – Better ability to predictions, more accuracy

Barriers to Deep Learning – Amount of data, computing power, understanding of decisions and complexity

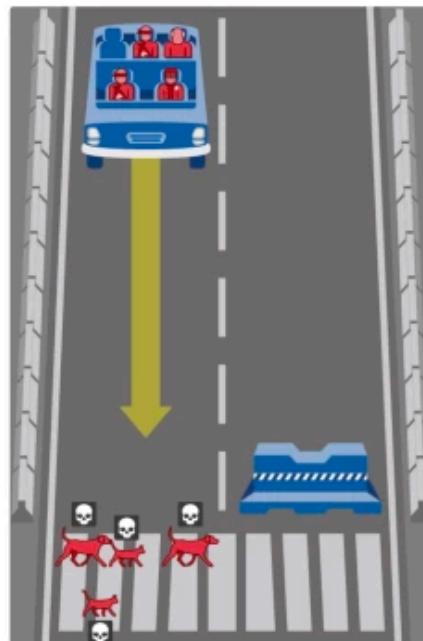
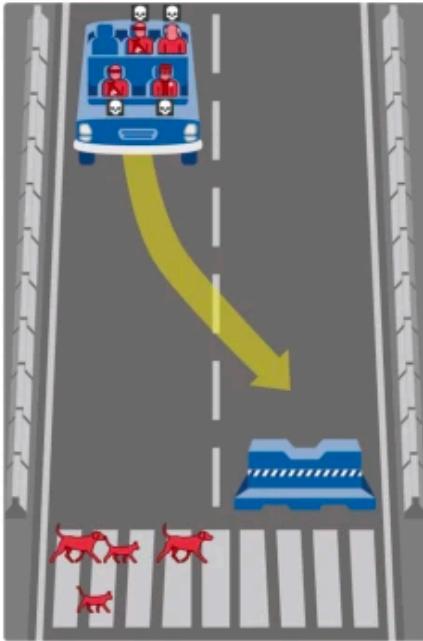
Ex: Google DeepMind to train Go-playing AI AlphaGo – 1,202 CPUs and 176 GPUs



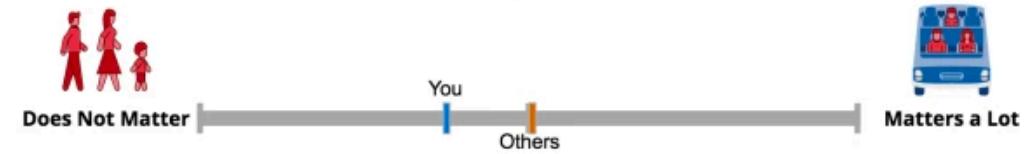
Ethics & Constraints in Machine Learning – Errors and biases

Ex: Best candidate for vacancy, image recognition.

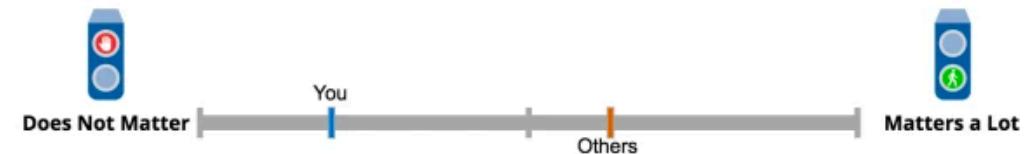
What should the self-driving car do?



Protecting Passengers



Upholding the Law



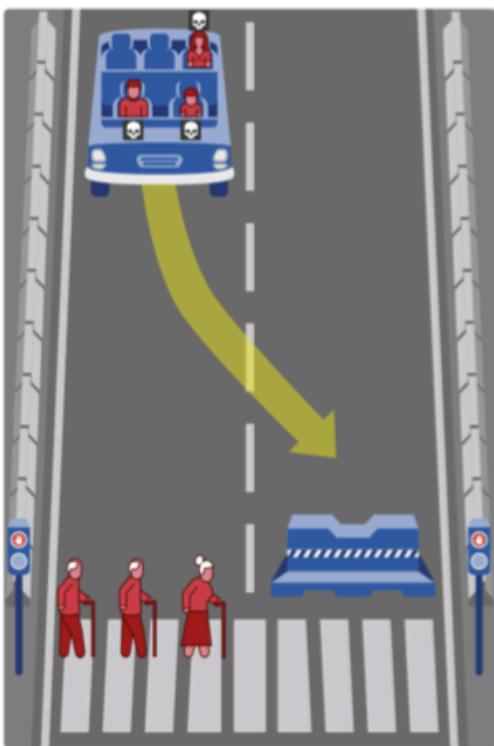
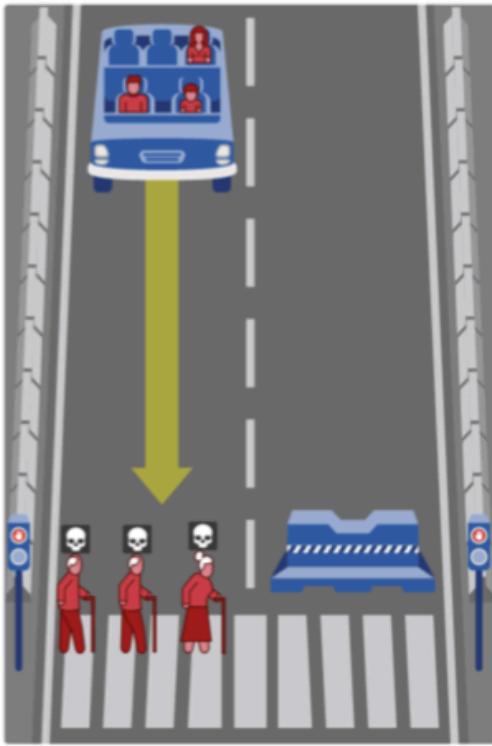
Avoiding Intervention



Ethics in Machine Learning

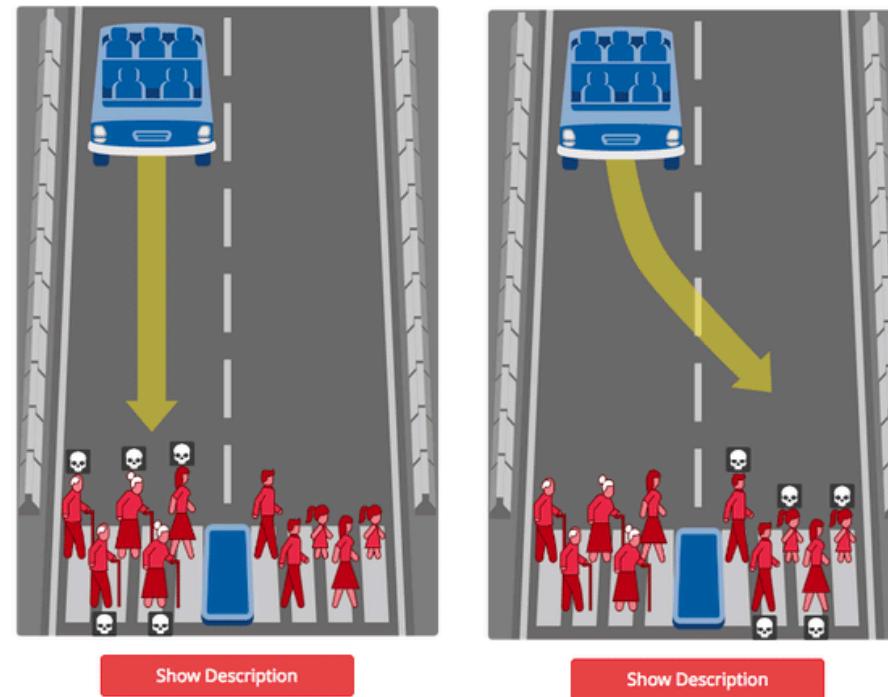
1. Old folks crash or hit puppies
2. Who's responsible ?

What should the self-driving car do?



What should the self-driving car do?

1 / 13



Show Description

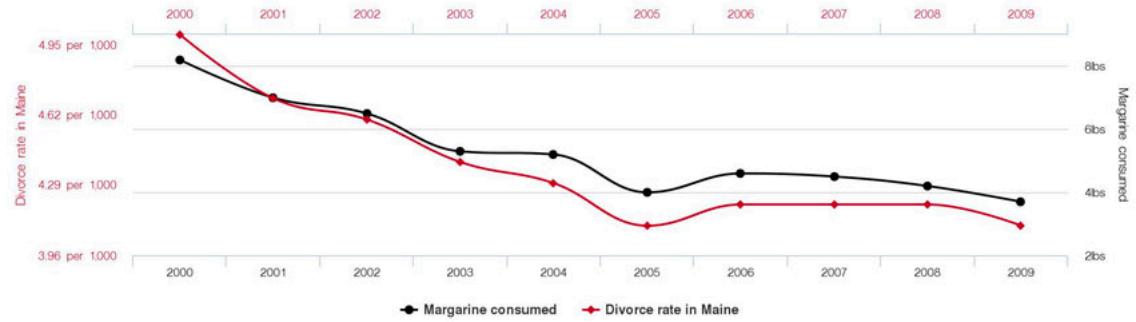
Ethics in Machine Learning

1. Old people or young people
2. Who's responsible?

Divorce rate in Maine

correlates with

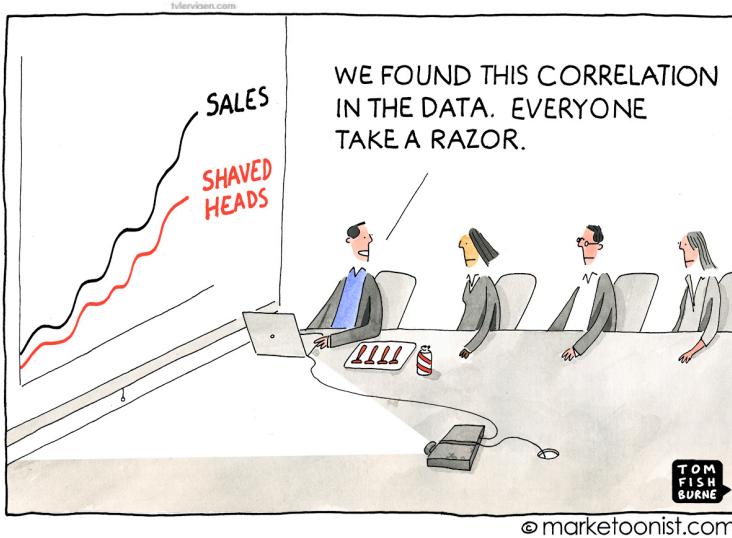
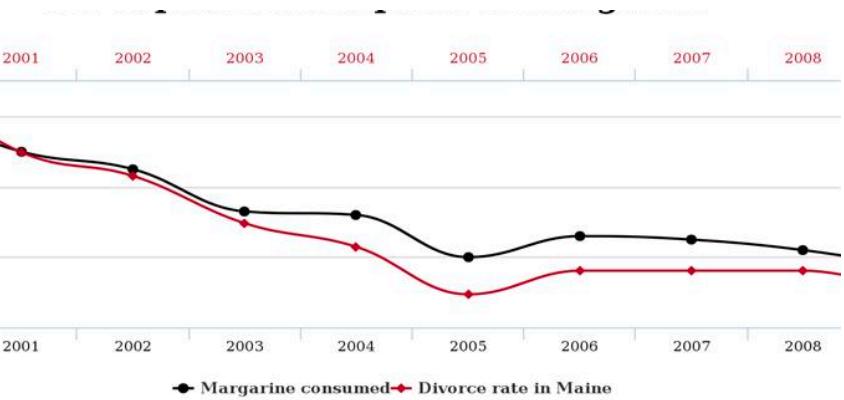
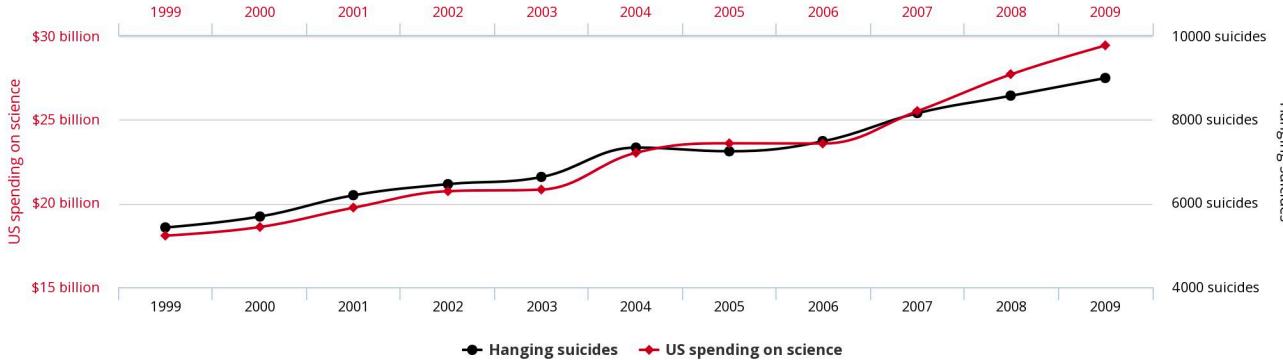
Per capita consumption of margarine



US spending on science, space, and technology

correlates with

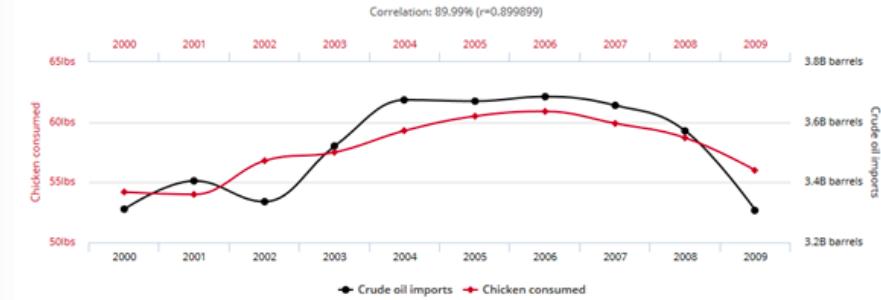
Suicides by hanging, strangulation and suffocation



Per capita consumption of chicken

correlates with

Total US crude oil imports



Data sources: U.S. Department of Agriculture and Dept. of Energy

Can Data trick you – Correlation does not Imply Causation

Ex: Iphones sales ~ Marandzas in Moz

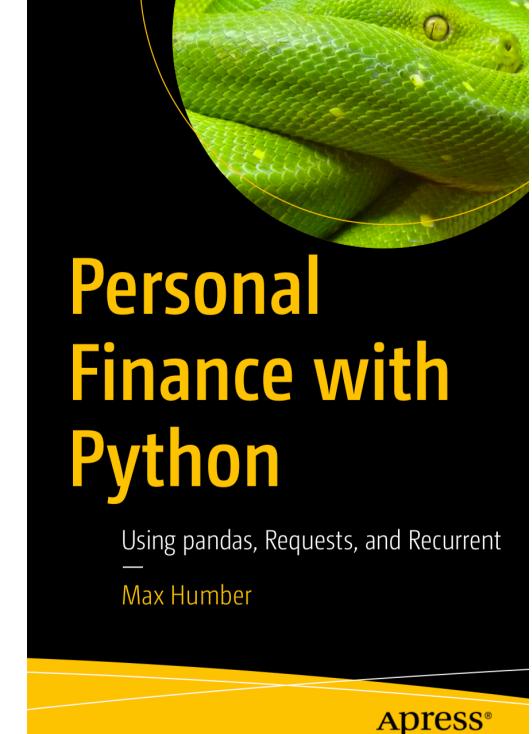
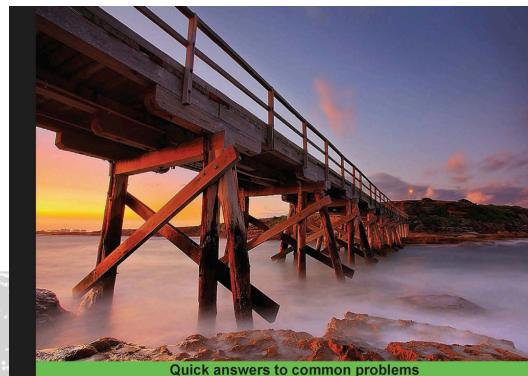
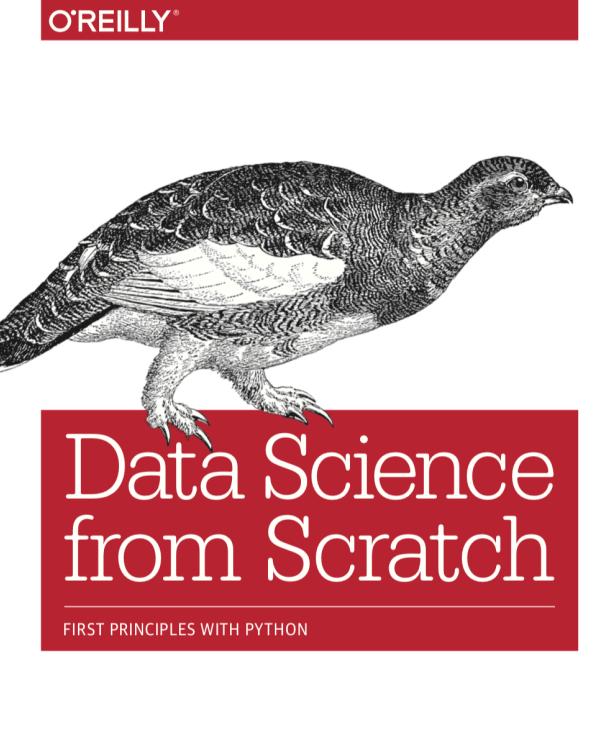
Learn Data Analysis with Python

Lessons in Coding

A.J. Henley
Dave Wolf

Apress®

Other resources...

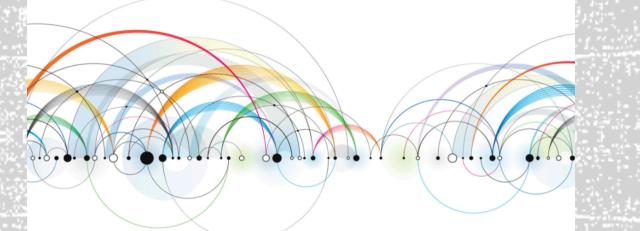


"A must-read resource for anyone who is serious about embracing the opportunity of big data."

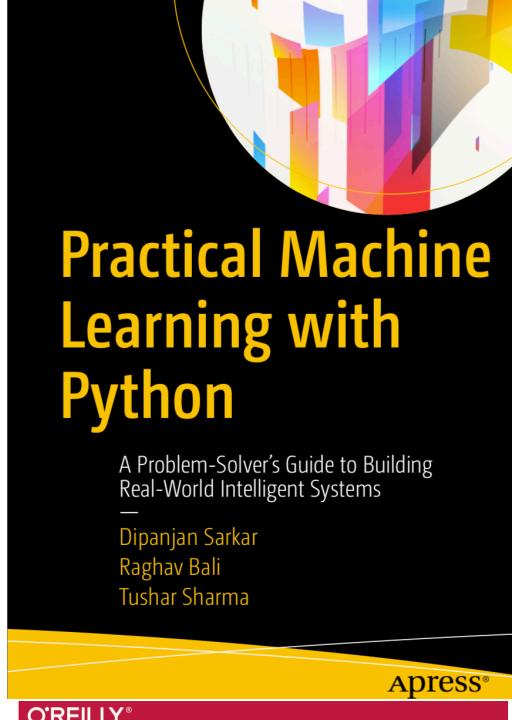
—Craig Vaughan, Global Vice President, SAP

Data Science for Business

What You Need to Know About Data Mining and Data-Analytic Thinking

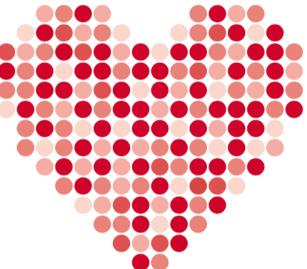


Foster Provost & Tom Fawcett



Learning to Love Data Science

Exploring Predictive Analytics, Machine Learning, Digital Manufacturing, and Supply Chain Optimization



Mike Barlow

Thank you for your patience

WHERE TO FIND ME



larslemos@gmail.com



[lars-lemos-28446152](#)



[lassinialbino](#)



github.com/larslemos

[@toplars](#)

