Data Curation

Arturo Aguilar, Diego Hernandez & Mariano Herrera



Data Sources Examples

Yahoo Finance!

Quandl

Banxico

<u>Inegi</u>

World Bank

UC Irvine Machine Learning Repository

Kaggle datasets

Amazon's AWS datasets

OpenDataMonitor

Built-in Functions & Methods

Python Built in Functions

Python String Methods

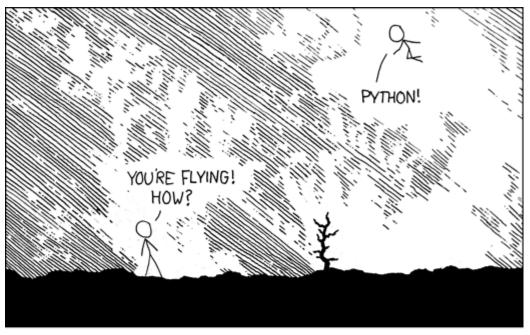
Python List/Array Methods

Python Dictionary Methods

Python Tuple Methods

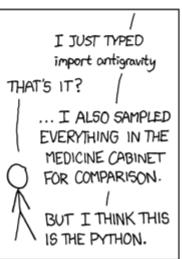
Python Set Methods

Understand how Import makes it easy









Source: https://xkcd.com/353/

Numpy – Tool for Mathematical and Scientific Calculations

What is NumPy?

NumPy quickstart

Array objects

Routines

Pandas – Data Analysis Tool

10 minutes to pandas

Intro to data structures

Cookbook

Input/output

<u>Series</u>

DataFrame

Data Visualization

Some libraries for visualization in Python:

Matplotlib

<u>Seaborn</u>

Folium

Bokeh

Plotly

<u>Altair</u>

Source: https://datapane.com/u/leo/reports/python-visualisation-guide-814e8638/

Some Financial Libraries

mlfinlab

Pyfolio

QuantLib

TA-Lib

<u>yfinance</u>

Disclaimer

The opinions expressed in this presentation and on the previous slides are solely those of the presenter.