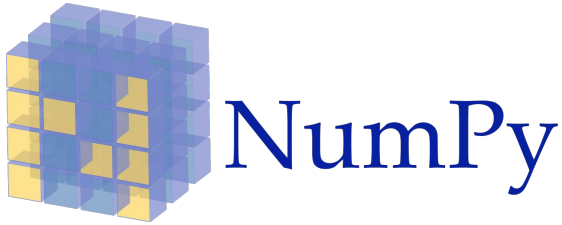
What Libraries Will We Use?

1. **Pandas**
2. **NumPy**
3. **Matplotlib**
4. **SkiKit-Learn**

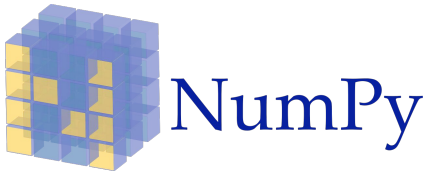


Extract, Transform, and Load

Data Exploration

Data Evaluation

Data Modeling



- Open source data analysis and manipulation tool
- Build on top of Python

- Provides fast, flexible, and expressive data structures
- Designed to make working with data easy and intuitive

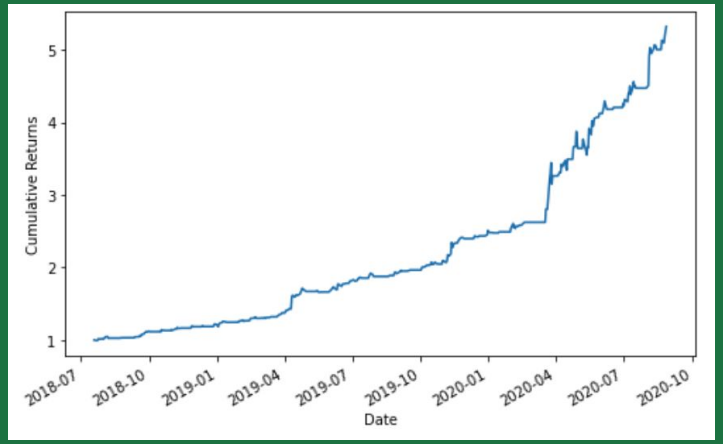
Good to use with:

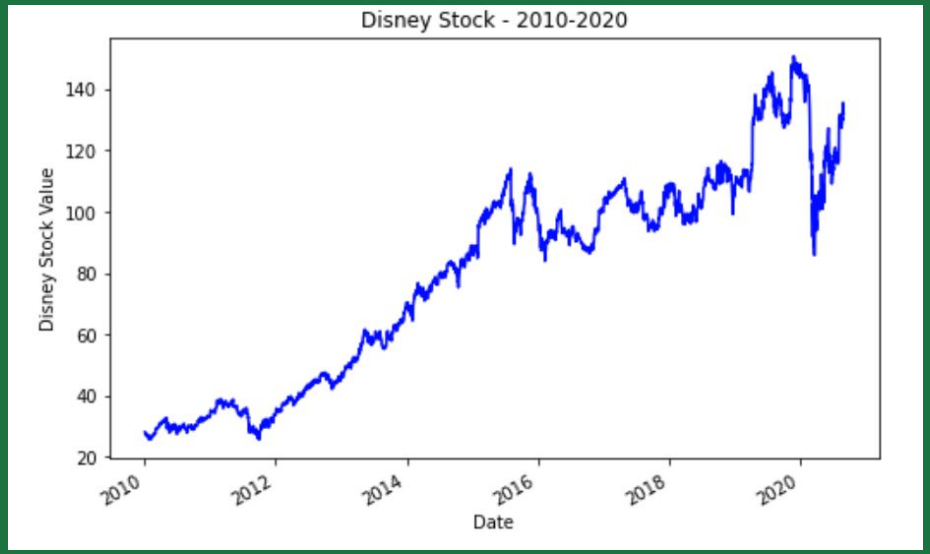
- SQL table or Excel spreadsheet
- Time series data
- Matrix data
- Observational / statistical datasets

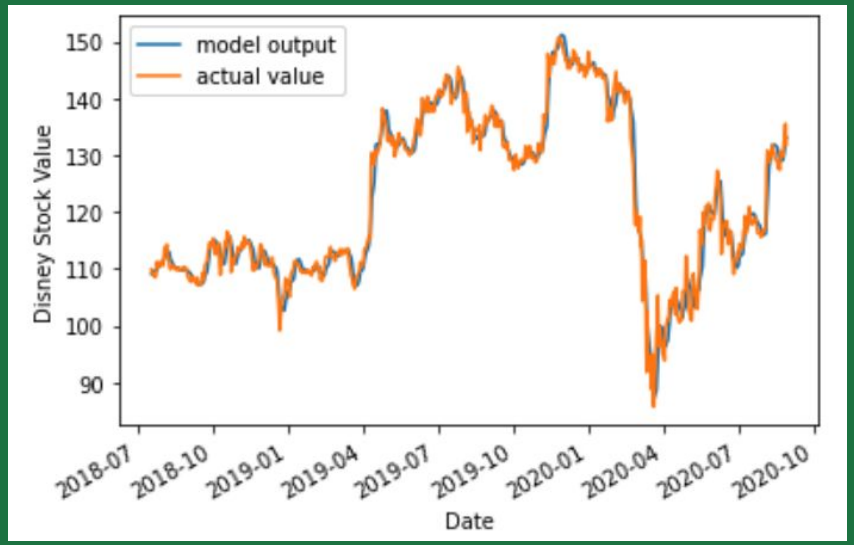
- Handle missing data
- Mutate size
- Align data
- Convert data into DataFrame objects
- Join datasets
- Pivot data sets
- Time series functionality

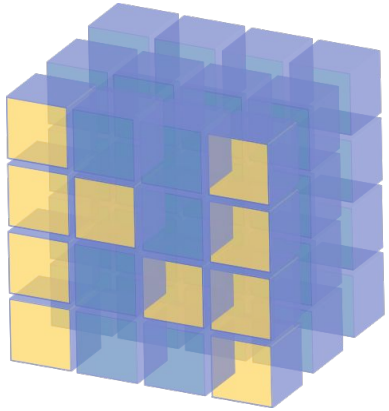
- Python 2D plotting library

- Generate plots, histograms, bar charts, scatterplots and more with a few lines of code


Cumulative Returns
Date
1
2
3
4
5
2018-07
2018-10
2019-01
2019-04
2019-07
2019-10
2020-01
2020-04
2020-07
2020-10

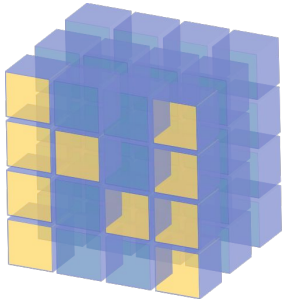

Disney Stock - 2010-2020
Disney Stock Value
Date
20
40
60
80
100
120
140
2010
2012
2014
2016
2018
2020


model output
actual value
Disney Stock Value
Date
90
100
110
120
130
140
150
2018-07
2018-10
2019-01
2019-04
2019-07
2019-10
2020-01
2020-04
2020-07
2020-10



NumPy

- Scientific computing with Python



NumPy

- Powerful n-dimension arrays
- Mathematical functions and more numerical computing tools

- Simple and efficient tools for predictive data analysis

1. **Classification**
2. **Regression**
3. **Clustering**
4. **Dimensionality reduction**
5. **Model selection**
6. **Preprocessing**

What Libraries Will We Use? ✓

Let's begin...