**Python for DA walk-through**

Runtime: *Python3.7.5*

**Real Python -** Using python’s built-in CSV to reduce need for uncalled need for external libraries. Follow the link below to get a comprehensive view of how the **csv** package works.  
[realpython.com/python-csv/](https://realpython.com/python-csv/) .

Checks:

* What is a CSV file >>> [#what-is-a-csv-file](https://realpython.com/python-csv/" \l "what-is-a-csv-file)
* Parsing CSV files with Python’s Built-in CSV Library >>> [#parsing-csv-files-with-pythons-built-in-csv-library](https://realpython.com/python-csv/" \l "parsing-csv-files-with-pythons-built-in-csv-library)

Extended Reading *(Optional)*:

* Parsing CSV files with the Pandas Library >>> [#parsing-csv-files-with-the-pandas-library](https://realpython.com/python-csv/" \l "parsing-csv-files-with-the-pandas-library).

You can also check out the [Python’s Built-in CSV Library docs.](https://docs.python.org/3/library/csv.html) to be sure no stone is left unturned.

**O’reilly - Python for Data Analysis**

To reduce complexity in learning, the NumPy understanding from this book is what you need to continue on your DA journey. [Download the e-book](https://www.pdfdrive.com/download.pdf?id=158189564&h=a3ad6e6b2504b95ec39a6c57d465ba5d&u=cache&ext=pdf) and select from the table of contents:

* Preliminaries >>> 1.1 to 1.3 *(If 1.4 is something you already know)*
* Python Language Basics, IPython and Jupiter Notebooks
* Built-In Data Structures, Functions, and Files
* NumPy Basics: Arrays and Vectorized Computation
* Getting Started with Pandas *(Quick walk through to ensure you’re leaving nothing behind)*
* Data Loading, Storage, and File Formats
* Data Cleaning and Preparation
* Data Wrangling: Join, Combine and Reshape
* Advanced Pandas
* Advanced Numpy

For all the Data you need, check [this repository](https://github.com/Adminixtrator/Machine_Learning), and try to follow or creatively improve the data preparation in [this notebook](https://github.com/Adminixtrator/Machine_Learning/blob/master/Abalone/Abalone.ipynb).

*\*\*Please note that you can dive into any other topic of interest, for now, visualization will be done with* ***Seaborn****, see next section\*\**

**Seaborn -** Data Visualization

As seen in the fore-mentioned e-book, **Matplotlib** is a good tool for visualization, but will be for now considered as a next option.  
Follow the [documentation](https://seaborn.pydata.org/tutorial.html) to understand how seaborn works and also *play* with it.

*\*\*After this section, please go back to the e-book to make comparison on how data is being visualized using either seaborn or matplotlib\*\**

©2021. Adminixtrator.