

Eric Heitman

Python Library: Pandas

Project Assignment #2: Write-Up Part 2

4/8/20

Python Library: Pandas

Background and Central Purpose:

A popular python library known as Pandas is helpful in the organizing of raw data. Amongst many additional functions, pandas is a Python library which primarily serves in the domains of data analysis and data manipulation. Within pandas, there are two major data structures: series and data frames which are one and two dimensional structures respectively. Both of these structures assist in Pandas central purpose, which revolves around presenting data in meaningful ways. Whether it be merging, formatting, aligning, or plotting different data sets, pandas is a powerful library which can assist a data scientist or programmer in all sorts of ways.

Practical Applications:

The realm of possibility of applications of Pandas are seemingly endless. During my research, I found many applications of Pandas in the realm of data science and analysis. For example, you could use pandas to take the seven day weather forecast from the web, and then produce that raw data into a meaningful data frame which can display the data in an effective way. Corresponding rows and columns can be labelled such as day of week, temperature, etc. to display to the user a live updated view of the weather forecast.

Other interesting ways of using Pandas would be to write Python code to interpolate raw data into a .csv file for further manipulation. Perhaps you have some colleagues who work exclusively in excel and need access to the data frame you are currently working on. In order to do this effectively, you could leverage one of the functions of the pandas library to export your pandas dataframe into a .csv file, and then forward this information to your colleagues.

Overview of Functions & Further References:

The amount of functions Pandas contains are extensive. Functions such as `dataframe.mean()` will average specified indices within your data frame to produce the mean result. Other functions of pandas include `pandas.read.csv`, which can read in a .csv file for further data manipulation. As you begin to work with pandas, you will begin to notice just how easy it is to use the library and find what you are looking for, no matter how large your dataset is. For further reference, please visit the official documentation page for Pandas at <https://pandas.pydata.org/pandas-docs/stable/> .