

Lab Exercise: Pandas Basics

Overview

These are a short series of exercises to experiment with loading data into a pandas DataFrame and selecting subsets of the data.

Pandas Basics

1 – Load in some data for GOOGL stock prices

- Use pandas `read_csv` for the file <http://neueda.conygre.com/pydata/GOOGL.csv>
- Index on 'Date' column
- set `parse_dates=True`
- display the first 5 rows of the DataFrame
- how many rows and columns are in the DataFrame? (try `df.shape`)

2 – Display the results of describe and transpose on the data e.g.

- `df.describe().transpose()`

3 – Display the first 5 rows of the DataFrame showing just the 'Open', 'High' and 'Low' columns

4 – Display the data for Feb 2005 only

5 – What was the maximum value for Adj. Close in the years 2014 through to 2016

6 – Display the mean, max, min and median of Adjusted Open, High, Low and Close for the entire data set for GOOGL.

- You could try this with the pandas `DataFrame.agg()` function, it takes a list of aggregate operations e.g. (you'll need to adjust this to your requirements)

```
df.agg(['mean', 'min', 'max', 'median'])
```