

Lab Exercise: Pandas Basics – Power prices.xlsx

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 Power prices

- Use pandas read_excel for the file Power prices.xlsx
- sheet_name “Power prices”
- Index_col on ‘Name’ column
- set parse_dates=True
- skip the first 3 rows with skiprows=3
- When loaded, rename the index from ‘Name’ to ‘Date’

e.g.

```
df = pd.read_excel('Power prices.xlsx', sheet_name="Power prices",  
skiprows=3, index_col='Name', parse_dates=True)
```

```
df.index.rename('Date', inplace=True)
```

2 – Explore the data

- display the first 5 rows of the DataFrame
- how many rows and columns are in the DataFrame? (try df.shape)

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

- df.describe().transpose()

4 – Display the first 5 rows of the DataFrame showing just the columns beginning with ‘US’

5 – Do a quick plot of 2 columns

e.g.

```
df[ ['US Baseload Electricity 23', 'US Baseload Electricity 24'] ].plot()
```

6 – What was the maximum value for **ITA Baseload 2Y** in the years 2018 through to 2020