Python Test

Price data from vendors often contains errors that need to be identified and investigated. There are three types of issues we are interested in:

- Missing values ('missing value')
- Stale values that are unchanged for more than 1 week ('stale value')
- Outliers, i.e. values that are far away from nearby values ('outlier')

Please write a tool containing the function below that takes a file path as a parameter and returns a collection of dates, associated values and issue type that should be investigated, in ascending date order.

```
def check_file_data(file_path: str) -> List[tuple[date, float, str]]:
    pass
```

The first priority is to detect *all* of the erroneous data. Avoiding false positives is secondary, less than 20 per series is acceptable.

We have supplied some CSVs of sample raw and clean prices for different instruments types. There are 2 columns in each file, Date and Price, with a header row. Dates are in the format dd/mm/yyyy. Plots of the files are on following pages.

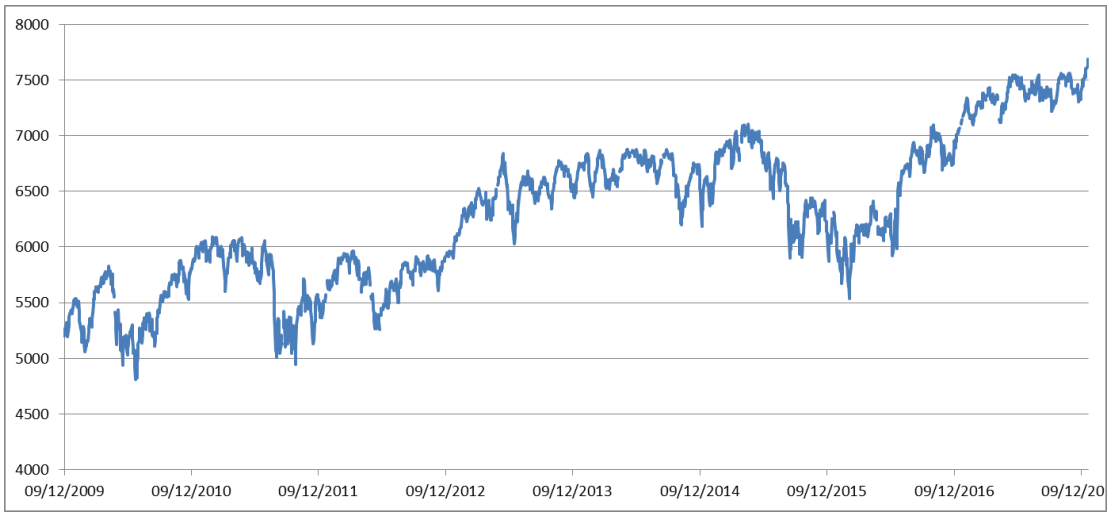
Use only core Python and standard library for this exercise. We would like to see well written, maintainable code and appropriate testing.

Number of issues in file:

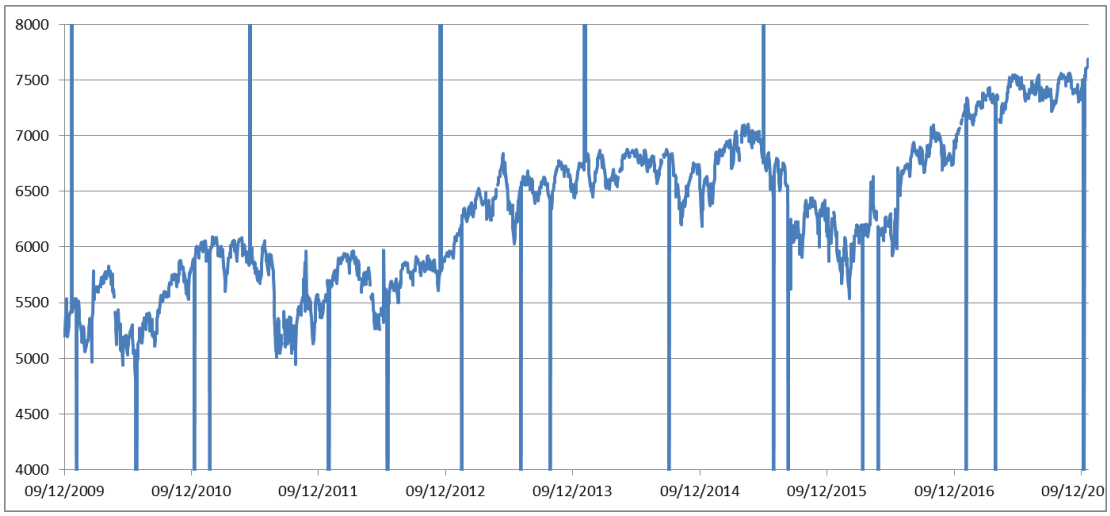
| | |
|---|---|
| Equity Index Clean | 0 |
| Equity Index Raw | 71 |
| Spot FX Clean | 0 |
| Spot FX Raw | 39 |
| Interest Rate Future Clean | 0 |
| Interest Rate Future Raw | 31 |

**Equity Index**

Clean

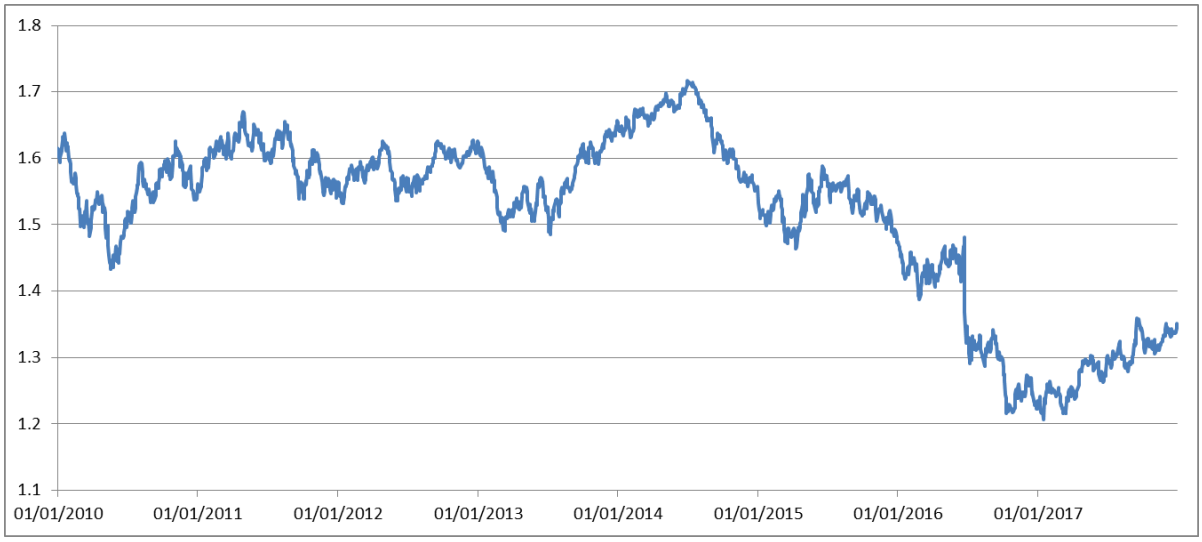



Raw

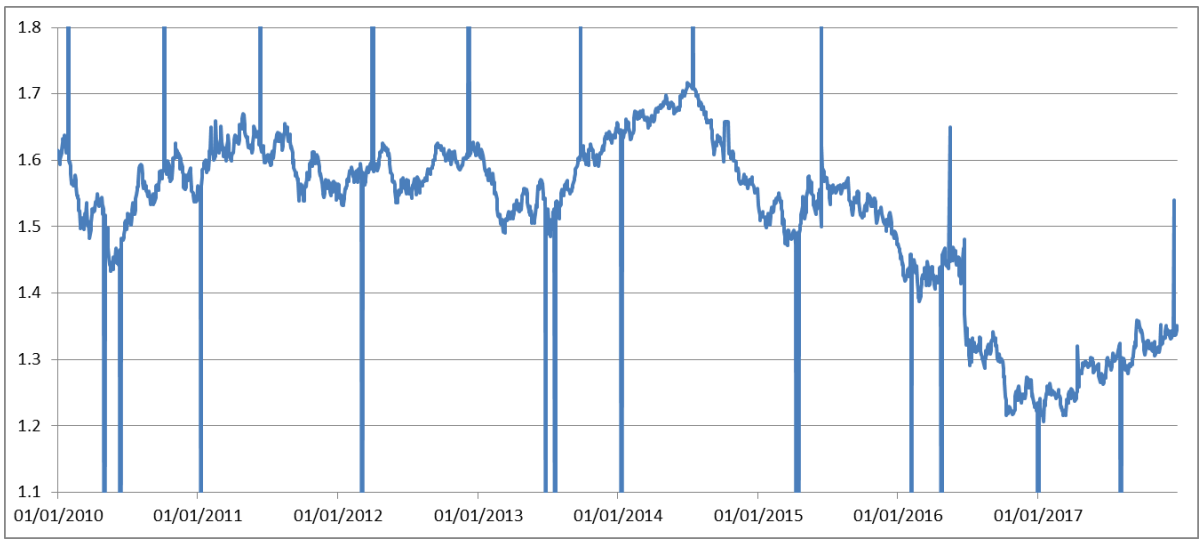


**Spot FX**

Clean




Raw



**Interest Rate Future Contract**

Clean


Raw