feature_visualisations

May 10, 2024

1 Feature Visualisations

[ ]:
```
import sys

sys.path.insert(1, "/Users/simon/Documents/II/Dissertation/")
%load_ext autoreload
%autoreload 2
import seaborn as sns
import matplotlib.pyplot as plt
import pandas as pd
from src.misc import load_processed_dataset

pd.set_option("display.max_columns", None)
import warnings
```

```
The autoreload extension is already loaded. To reload it, use:
  %reload_ext autoreload
```

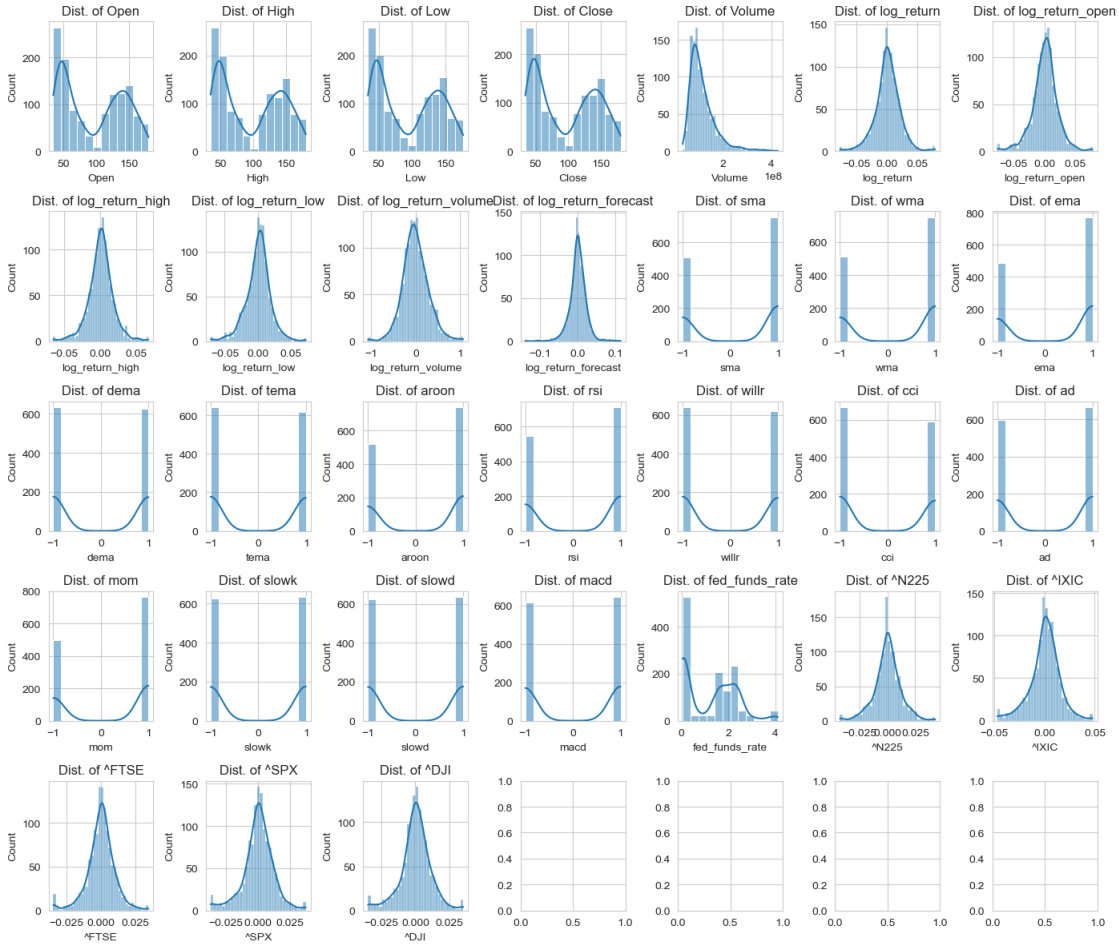
[ ]:
```
df = load_processed_dataset("aapl")
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 1259 entries, 2018-01-02 05:00:00 to 2022-12-30 05:00:00
Data columns (total 31 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   Open                 1259 non-null   float64
 1   High                 1259 non-null   float64
 2   Low                  1259 non-null   float64
 3   Close                1259 non-null   float64
 4   Volume               1259 non-null   int64
 5   log_return           1259 non-null   float64
 6   log_return_open      1259 non-null   float64
 7   log_return_high      1259 non-null   float64
 8   log_return_low       1259 non-null   float64
 9   log_return_volume    1259 non-null   float64
 10  log_return_forecast  1259 non-null   float64
 11  sma                  1259 non-null   int64
```

```
 12  wma                 1259 non-null   int64
 13  ema                 1259 non-null   int64
 14  dema                1259 non-null   int64
 15  tema                1259 non-null   int64
 16  aroon               1259 non-null   int64
 17  rsi                 1259 non-null   int64
 18  willr               1259 non-null   int64
 19  cci                 1259 non-null   int64
 20  ad                  1259 non-null   int64
 21  mom                 1259 non-null   int64
 22  slowk               1259 non-null   int64
 23  slowd               1259 non-null   int64
 24  macd                1259 non-null   int64
 25  fed_funds_rate      1259 non-null   float64
 26  ^N225               1259 non-null   float64
 27  ^IXIC               1259 non-null   float64
 28  ^FTSE               1259 non-null   float64
 29  ^SPX                1259 non-null   float64
 30  ^DJI                1259 non-null   float64
dtypes: float64(16), int64(15)
memory usage: 314.8 KB
```

[ ]:
```
warnings.filterwarnings("ignore", "is_categorical_dtype")
warnings.filterwarnings("ignore", "use_inf_as_na")

rows = 5
cols = 7
fig, axs = plt.subplots(nrows=rows, ncols=cols, figsize=(14, 12))
for i, col in enumerate(df.columns):
    sns.histplot(df[col], kde=True, ax=axs[i // cols][i % cols])
    axs[i // cols][i % cols].set_title(f"Dist. of {col}")
fig.tight_layout()
plt.show()
```



[ ]:
```
fig, axs = plt.subplots(nrows=rows, ncols=cols, figsize=(14, 12))
for i, col in enumerate(df.columns):
    sns.boxenplot(df[col], ax=axs[i // cols][i % cols])
    axs[i // cols][i % cols].set_title(f"Box Plot of {col}")
fig.tight_layout()
plt.show()
```

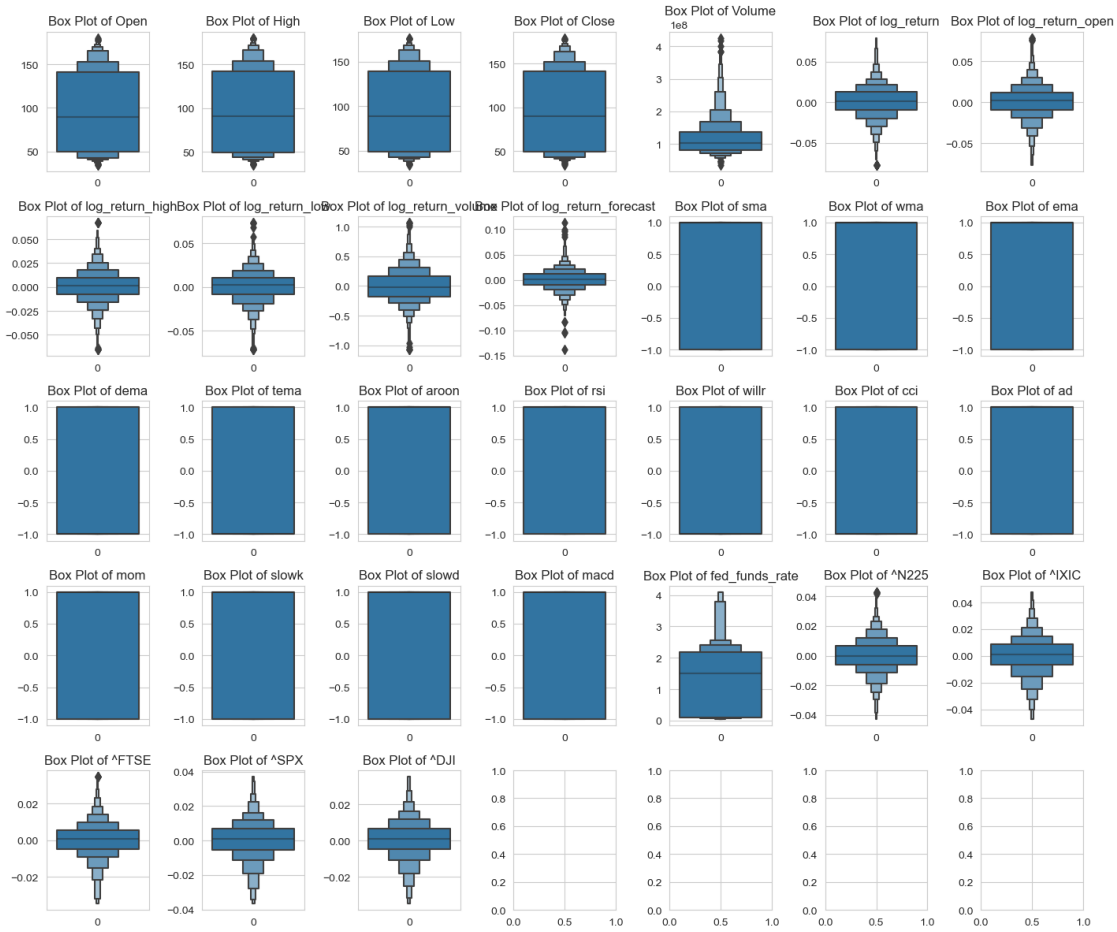
```
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
```

```
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
```

```
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
```

```
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
```

```
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
/Users/simon/anaconda3/envs/proj/lib/python3.9/site-
packages/seaborn/categorical.py:486: FutureWarning: Series.__getitem__ treating
keys as positions is deprecated. In a future version, integer keys will always
be treated as labels (consistent with DataFrame behavior). To access a value by
position, use `ser.iloc[pos]`
  if np.isscalar(data[0]):
```


Box Plot of Open
Box Plot of High
Box Plot of Low
Box Plot of Close
Box Plot of Volume
1e8
Box Plot of log_return
Box Plot of log_return_open
Box Plot of log_return_high
Box Plot of log_return_low
Box Plot of log_return_volume
Box Plot of log_return_forecast
Box Plot of sma
Box Plot of wma
Box Plot of ema
Box Plot of dema
Box Plot of tema
Box Plot of aroon
Box Plot of rsi
Box Plot of willr
Box Plot of cci
Box Plot of ad
Box Plot of mom
Box Plot of slowk
Box Plot of slowd
Box Plot of macd
Box Plot of fed_funds_rate
Box Plot of ^N225
Box Plot of ^IXIC
Box Plot of ^FTSE
Box Plot of ^SPX
Box Plot of ^DJI