

## Task 5: Forecasting

```
In [1]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
```

**Camden National Bank** is a publicly traded bank with headquarters in Maine. It is one of Lobster Land's creditors and its data has been used to predict the bank's share price at end of 2020.

```
In [4]: cac = pd.read_csv('CAC.csv') # Camden National Bank
```

```
In [5]: cac.head()
```

Out[5]:

	Date	Open	High	Low	Close	Adj Close	Volume
0	2019-12-09	44.070000	44.070000	43.650002	43.810001	42.186695	61900
1	2019-12-10	43.869999	44.209999	43.570000	44.139999	42.504471	59200
2	2019-12-11	44.310001	44.410000	44.040001	44.209999	42.571873	51100
3	2019-12-12	44.340000	45.320000	44.240002	45.070000	43.400009	98600
4	2019-12-13	45.040001	45.110001	44.279999	44.709999	43.053341	37000

```
In [6]: cac.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 252 entries, 0 to 251
Data columns (total 7 columns):
#   Column      Non-Null Count  Dtype
---  -
0   Date        252 non-null    object
1   Open        252 non-null    float64
2   High        252 non-null    float64
3   Low         252 non-null    float64
4   Close       252 non-null    float64
5   Adj Close   252 non-null    float64
6   Volume      252 non-null    int64
dtypes: float64(5), int64(1), object(1)
memory usage: 13.9+ KB
```

Convertin data for time series

```
In [7]: cac['Date'] = pd.to_datetime(cac['Date'])
```

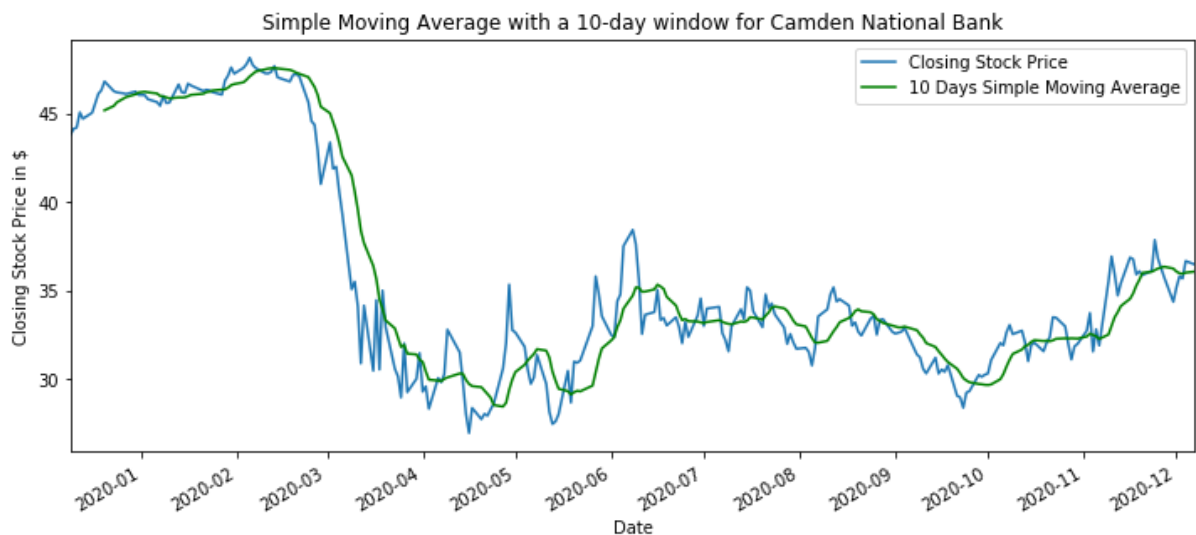
```
In [8]: cac.set_index('Date', inplace=True)
```

```
In [10]: cac.info()
```

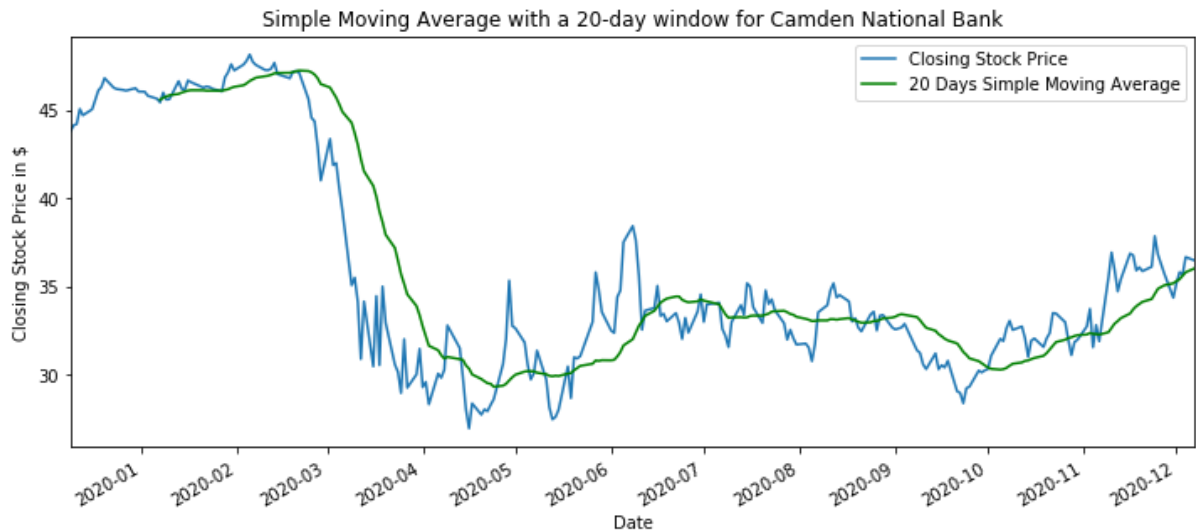
```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 252 entries, 2019-12-09 to 2020-12-07
Data columns (total 6 columns):
 #   Column      Non-Null Count  Dtype
---  ---
 0   Open        252 non-null    float64
 1   High        252 non-null    float64
 2   Low         252 non-null    float64
 3   Close       252 non-null    float64
 4   Adj Close   252 non-null    float64
 5   Volume      252 non-null    int64
dtypes: float64(5), int64(1)
memory usage: 13.8 KB
```

Predicting Share Price at the end of December 2020 using Simple Moving Averages

```
In [12]: cac['Close'].plot(title = 'Simple Moving Average with a 10-day window fo
r Camden National Bank', fontsize = 10,
                        figsize=(12,5), label = 'Closing Stock Price').auto
scale(axis='x', tight = True);
cac.rolling(window = 10).mean()['Close'].plot(color = 'green', label =
'10 Days Simple Moving Average');
plt.legend();
plt.xlabel('Date')
plt.ylabel('Closing Stock Price in $');
```



```
In [11]: cac['Close'].plot(title = 'Simple Moving Average with a 20-day window fo
r Camden National Bank', fontsize = 10,
                        figsize=(12,5), label = 'Closing Stock Price').auto
scale(axis='x', tight = True);
cac.rolling(window = 20).mean()['Close'].plot(color = 'green', label =
'20 Days Simple Moving Average');
plt.legend();
plt.xlabel('Date')
plt.ylabel('Closing Stock Price in $');
```



Camden National Bank's publicly-available share price data has been obtained for the year 2019. Lobster Land wishes to forecast the share-price at the end of the year 2020, and for the same simple moving averages has been used. It is a mathematical tool used to gauge trend changes or determine prices, mainly in stocks, based on historical data. It shows a "true" average over time by smoothening out any unusual changes. A 20-day period over the 10-day period has been chosen as the optimal  $k$ -value because longer  $k$ -values help in smoothening out sudden fluctuations and are much more stabilized. Longer  $k$ -values are better for long-term trends and effects.

It can be seen from the plot that in until March, 2020, the share prices were high. However, it declined after March, and the prices kept fluctuating until October. It then saw a gradual increase from 30 dollars in October to approximately 37 dollars in December. So, the bank's share price has been predicted to be 37 dollars at the end of December, 2020.