

Stock Data Analysis

This file is using data generated from the Athena query to get the sotck's hourly max high price for each stock. The stocks the data is generated for are listed below.

Stocks

- Facebook (FB)
- Shopify (SHOP)
- Beyond Meat (BYND)
- Netflix (NFLX)
- Pinterest (PINS)
- Square (SQ)
- The Trade Desk (TTD)
- Okta (OKTA)
- Snap (SNAP)
- Datadog (DDOG)

Importing all necessary libraries

```
In [55]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns
import warnings
warnings.filterwarnings("ignore")
```

Importing the query result data generated from the athena query and named as results.csv

The query was to find the highest stock price per hour for each stock
The data is imported in a pandas dataframe object

```
In [56]: df = pd.read_csv('/Users/sera/Desktop/Baruch Books/CIS 9760 Big Data/Projects/project03/results.csv')
```

Calling the top 5 rows of the dataframe to make sure the data looks right

```
In [57]: df.head()
```

	Name	Hour	Time	Max High
0	BYND	9	2021-05-11 09:55:00-04:00	104.709999
1	BYND	10	2021-05-11 10:55:00-04:00	106.460999
2	BYND	11	2021-05-11 11:55:00-04:00	107.695000
3	BYND	12	2021-05-11 12:55:00-04:00	108.839996
4	BYND	13	2021-05-11 13:45:00-04:00	110.660004

Line chart showing the max high price per hour for each stock

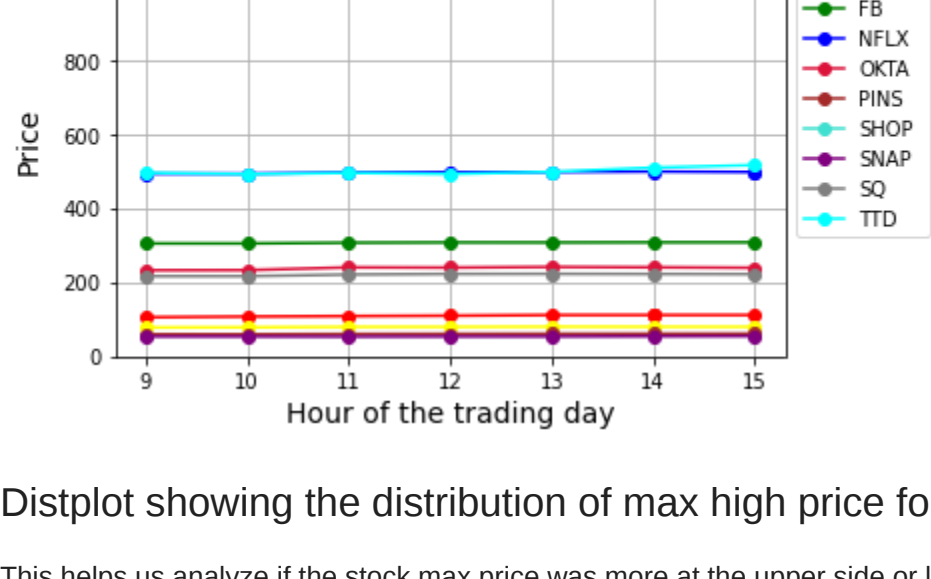
This helps us see the trend of the max high price per hour and compare the different stocks. Also if there is a sudden sharp increase or decrease it can be easily seen as well.

```
In [58]: bynd = df[df['Name']=='BYND'][['Hour','Max High']]
ddog = df[df['Name']=='DDOG'][['Hour','Max High']]
fb = df[df['Name']=='FB'][['Hour','Max High']]
nflx = df[df['Name']=='NFLX'][['Hour','Max High']]
okta = df[df['Name']=='OKTA'][['Hour','Max High']]
pins = df[df['Name']=='PINS'][['Hour','Max High']]
shop = df[df['Name']=='SHOP'][['Hour','Max High']]
snap = df[df['Name']=='SNAP'][['Hour','Max High']]
sq = df[df['Name']=='SQ'][['Hour','Max High']]
ttt = df[df['Name']=='TTD'][['Hour','Max High']]
```

```
In [59]: plt.figure()
plt.plot(bynd['Hour'],bynd['Max High'],color= 'red',marker='o',label='BYND')
plt.plot(ddog['Hour'],ddog['Max High'],color= 'yellow',marker='o',label='DDOG')
plt.plot(fb['Hour'],fb['Max High'],color= 'green',marker='o',label='FB')
plt.plot(nflx['Hour'],nflx['Max High'],color= 'blue',marker='o',label='NFLX')
plt.plot(okta['Hour'],okta['Max High'],color= 'crimson',marker='o',label='OKTA')
plt.plot(pins['Hour'],pins['Max High'],color= 'brown',marker='o',label='PINS')
plt.plot(shop['Hour'],shop['Max High'],color= 'turquoise',marker='o',label='SHOP')
plt.plot(snap['Hour'],snap['Max High'],color= 'purple',marker='o',label='SNAP')
plt.plot(sq['Hour'],sq['Max High'],color= 'gray',marker='o',label='SQ')
plt.plot(ttd['Hour'],ttd['Max High'],color= 'cyan',marker='o',label='TTD')

plt.title("Stock hourly highest price",fontsize=16)
plt.xlabel("Hour of the trading day",fontsize=14)
plt.ylabel("Price",fontsize=14)
plt.legend(bbox_to_anchor=(1.0, 1.0),loc='upper left')
plt.grid(True)
plt.show()

plt.show()
```

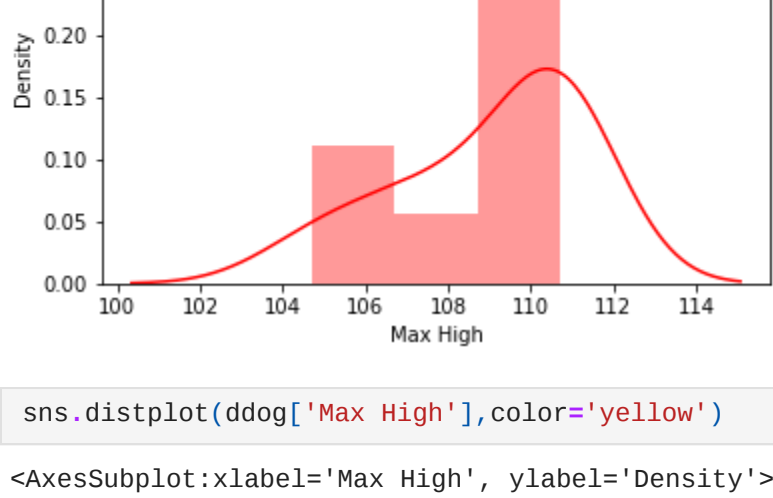


Distplot showing the distribution of max high price for each stock

This helps us analyze if the stock max price was more at the upper side or lower side

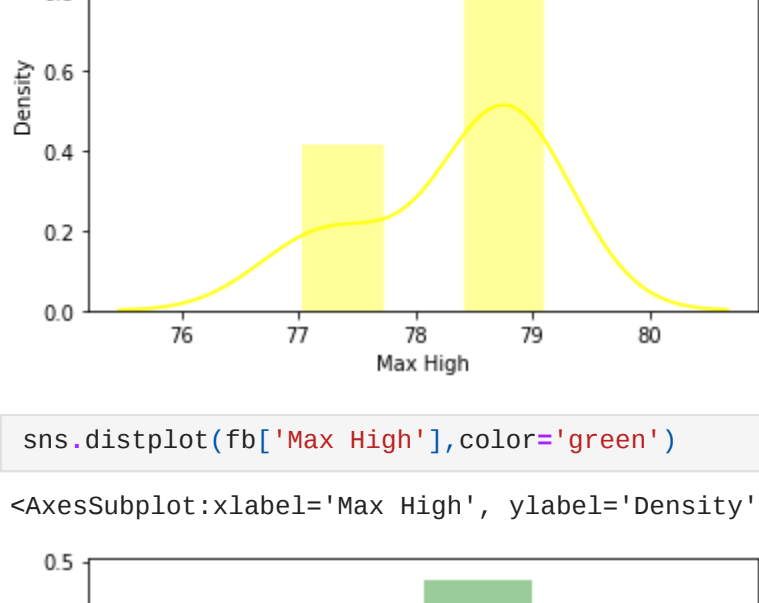
```
In [60]: sns.distplot(bynd['Max High'],color='red')
```

```
Out[60]: <AxesSubplot:xlabel='Max High', ylabel='Density'>
```



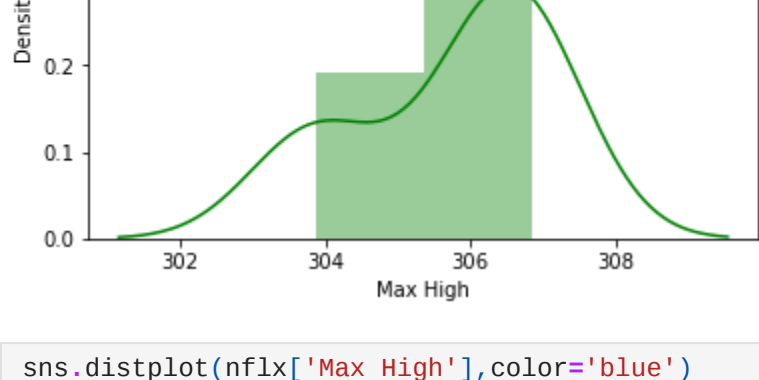
```
In [61]: sns.distplot(ddog['Max High'],color='yellow')
```

```
Out[61]: <AxesSubplot:xlabel='Max High', ylabel='Density'>
```



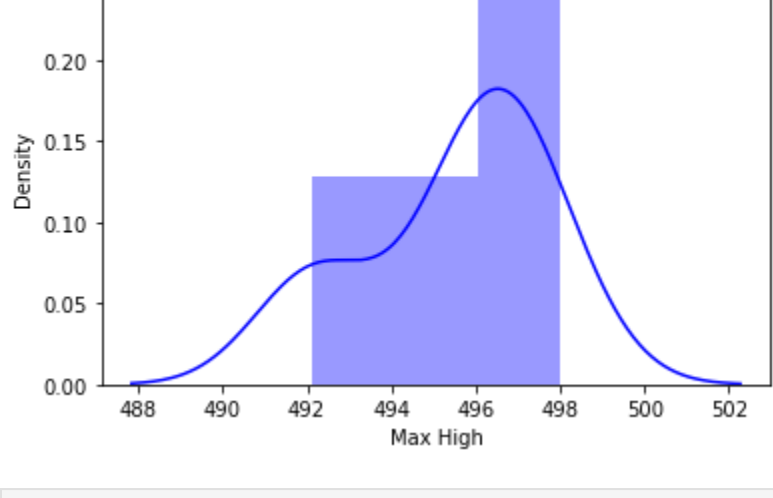
```
In [62]: sns.distplot(fb['Max High'],color='green')
```

```
Out[62]: <AxesSubplot:xlabel='Max High', ylabel='Density'>
```



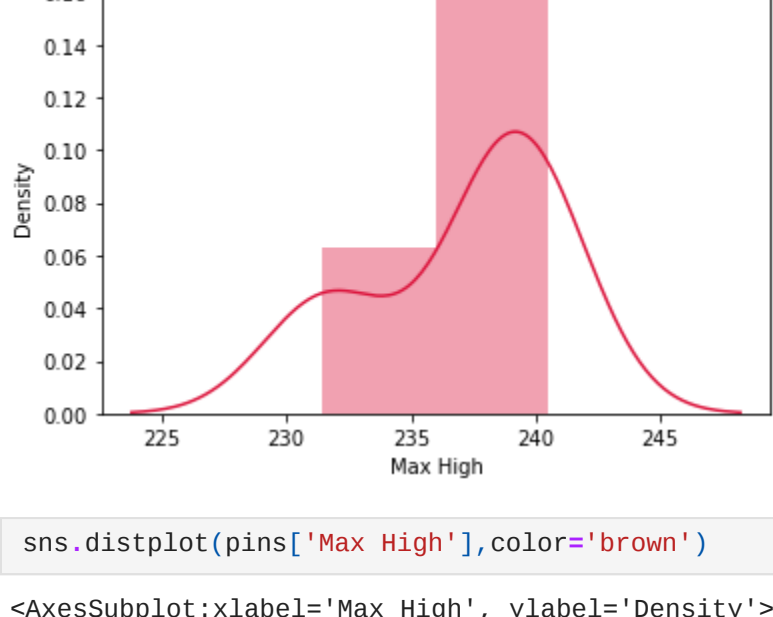
```
In [63]: sns.distplot(nflx['Max High'],color='blue')
```

```
Out[63]: <AxesSubplot:xlabel='Max High', ylabel='Density'>
```



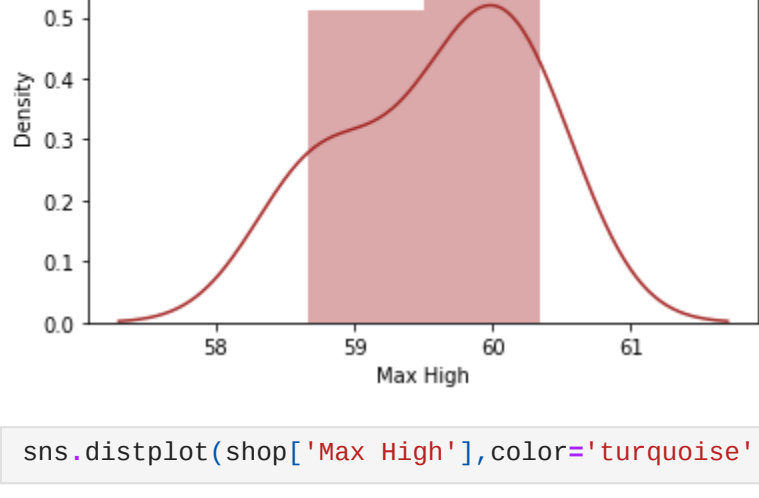
```
In [64]: sns.distplot(okta['Max High'],color='crimson')
```

```
Out[64]: <AxesSubplot:xlabel='Max High', ylabel='Density'>
```



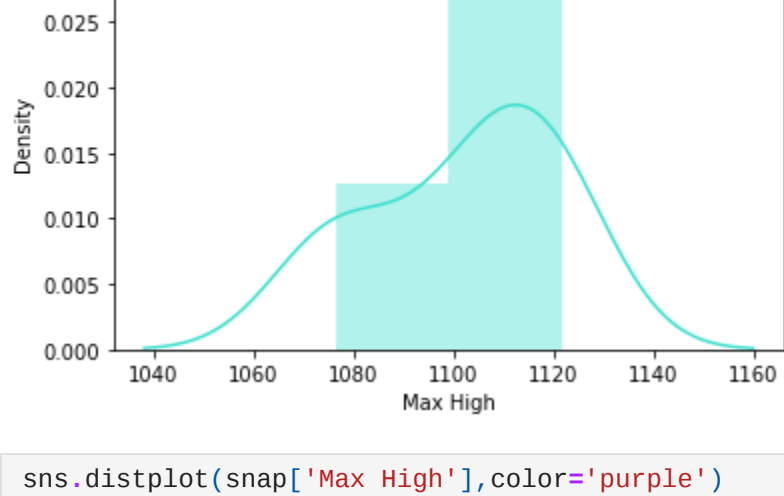
```
In [65]: sns.distplot(pins['Max High'],color='brown')
```

```
Out[65]: <AxesSubplot:xlabel='Max High', ylabel='Density'>
```



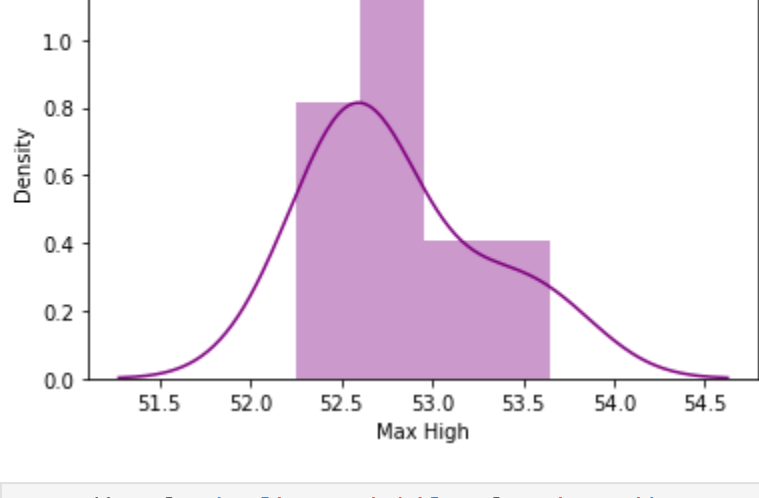
```
In [66]: sns.distplot(shop['Max High'],color='turquoise')
```

```
Out[66]: <AxesSubplot:xlabel='Max High', ylabel='Density'>
```



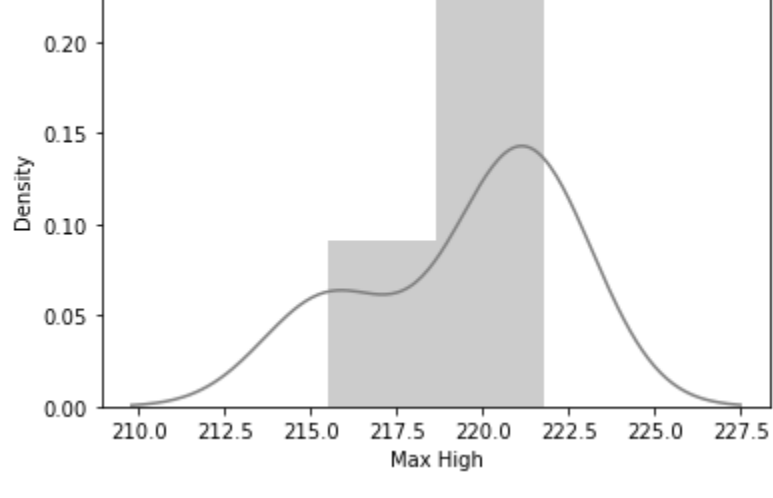
```
In [67]: sns.distplot(snap['Max High'],color='purple')
```

```
Out[67]: <AxesSubplot:xlabel='Max High', ylabel='Density'>
```



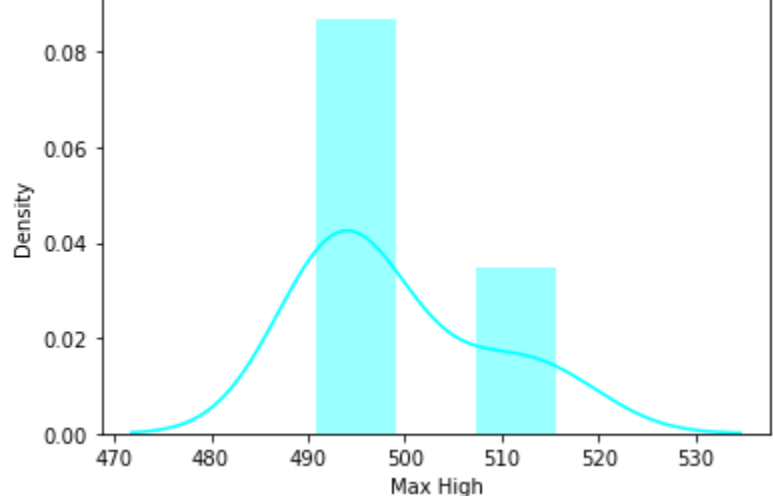
```
In [68]: sns.distplot(sq['Max High'],color='gray')
```

```
Out[68]: <AxesSubplot:xlabel='Max High', ylabel='Density'>
```



```
In [69]: sns.distplot(ttd['Max High'],color='cyan')
```

```
Out[69]: <AxesSubplot:xlabel='Max High', ylabel='Density'>
```



Thank you!!

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```