Analysis

June 27, 2025

[1]:
```
### Import packages in the virtual environment
import pandas as pd   ## search the package "pandans"

import matplotlib.pyplot as plt    ## search the package "matplotlib"
import seaborn as sns     ## search the package "seaborn"
```

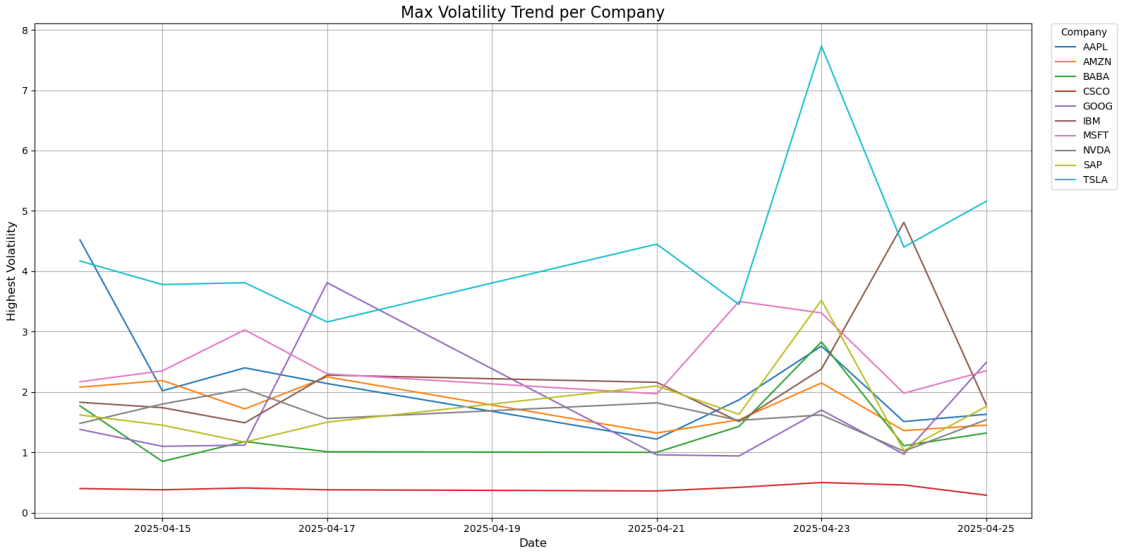
[2]:
```
data = pd.read_csv('/Users/donkor/Documents/CIS 4130 - Big Data Technologies/
 ↪Project2/Athena/results.csv')
```

[3]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 90 entries, 0 to 89
Data columns (total 5 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   company             90 non-null     object
 1   day                 90 non-null     object
 2   average volatility  90 non-null     float64
 3   highest volatility  90 non-null     float64
 4   lowest volatility   90 non-null     float64
dtypes: float64(3), object(2)
memory usage: 3.6+ KB
```

[4]:
```
data['day'] = pd.to_datetime(data['day'])

plt.figure(figsize=(16, 8))

for company in data['company'].unique():
    subset = data[data['company'] == company]
    plt.plot(subset['day'], subset['highest volatility'], label=company)

plt.title('Max Volatility Trend per Company', fontsize=16)
plt.xlabel('Date', fontsize=12)
plt.ylabel('Highest Volatility', fontsize=12)
plt.grid(True)
```

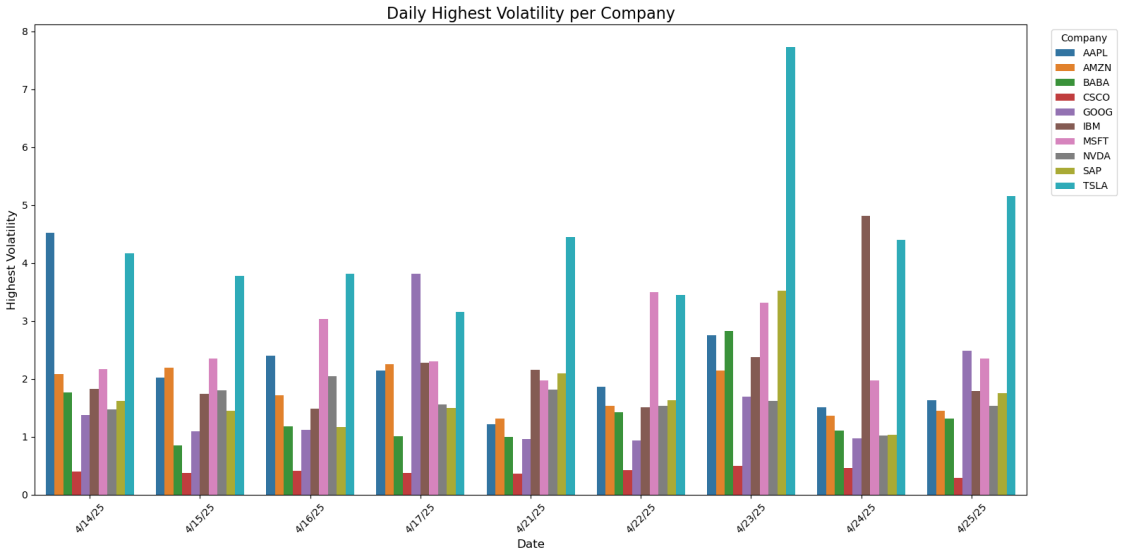
```
plt.legend(title='Company', bbox_to_anchor=(1.02, 1), loc='upper left',␣
 ↪borderaxespad=0)

plt.tight_layout()
plt.show()
```



Which company is the most volatile?

The company with the highest volatility is TSLA (Tesla).

[5]:
```
data['day'] = pd.to_datetime(data['day'])


data['formatted_day'] = data['day'].dt.strftime('%-m/%-d/%y')

# Plot
plt.figure(figsize=(16, 8))
sns.barplot(data=data, x='formatted_day', y='highest volatility', hue='company')

plt.title('Daily Highest Volatility per Company', fontsize=16)
plt.xlabel('Date', fontsize=12)
plt.ylabel('Highest Volatility', fontsize=12)
plt.xticks(rotation=45)
plt.legend(title='Company', bbox_to_anchor=(1.02, 1), loc='upper left')
plt.tight_layout()
plt.show()
```



[ ]:

Do the findings from this graph support your conclusion from the first graph?

Yes, the bar chart supports the line graph's conclusion that TSLA (Tesla) is the most volatile company.