Stock_Market_Analysis

Ubanna Dan-Ekeh

September 23, 2020

1 STOCK ANALYSIS ON SQUARE, PAYPAL, AND GOLDMAN SACHS

On the 4th of September, 2020, CNBC published an article on the rise of fintech disruptors - PayPal and Square surpassed Bank of America and Goldman Sachs respectively.

Here is the link to the article by CNBC - https://www.cnbc.com/2020/09/04/disruptors-paypal-and-square-surpass-wall-street-giants-including-goldman-sachs-in-market-cap.html.

This is brief analysis of the stocks of Paypal, Square, and Goldman Sachs from January 2016 - September 2020.

Data is gotten from yahoo finance api

1.0.1 Libraries/Dependencies used

[2]:
```
import pandas_datareader.data as web
import datetime
import matplotlib.pyplot as plt
%matplotlib inline
```

```
import seaborn as sns
sns.set_style("dark")
import pandas as pd
import numpy as np
from pandas.plotting import scatter_matrix
from mpl_finance import candlestick_ohlc
from matplotlib.dates import DateFormatter, date2num, WeekdayLocator,␣
 ↪DayLocator, MONDAY
import warnings
warnings.filterwarnings('ignore')
warnings.simplefilter('ignore')
```

1.0.2 Load the data from yahoo api

[3]:
```
start = datetime.datetime(2016, 1, 1)
end = datetime.datetime(2020, 9, 22)
```

[4]:
```
square = web.DataReader("SQ", 'yahoo', start, end)
goldman = web.DataReader("GS", 'yahoo', start, end)
paypal = web.DataReader("PYPL", 'yahoo', start, end)
```

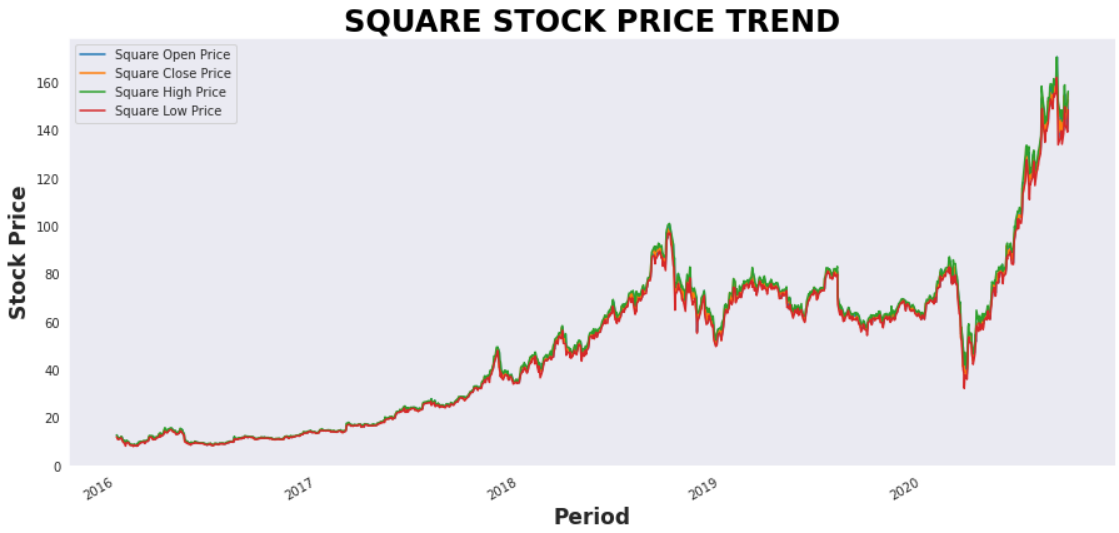
1.0.3 Explore the trend of each company's stock price over time

[5]:
```
# First five rows of the data for square
square.head()
```

[5]:

Date	High	Low	Open	Close	Volume	Adj Close
2016-01-04	12.90	12.050	12.75	12.16	2751500	12.16
2016-01-05	12.34	11.500	12.20	11.51	2352800	11.51
2016-01-06	11.64	11.015	11.50	11.52	1850600	11.52
2016-01-07	11.37	11.000	11.13	11.16	1636000	11.16
2016-01-08	11.54	11.200	11.25	11.31	587300	11.31

[6]:
```
square['Open'].plot(label = 'Square Open Price', figsize=(15, 7))
square['Close'].plot(label = 'Square Close Price')
square['High'].plot(label = 'Square High Price')
square['Low'].plot(label = 'Square Low Price')
plt.legend()
plt.title('SQUARE STOCK PRICE TREND', fontsize = 24, fontweight = 'bold',color␣
 ↪= 'black')
plt.xlabel('Period', fontsize = '17', fontweight = 'bold')
plt.ylabel('Stock Price', fontsize = '17', fontweight = 'bold')
plt.show()
```

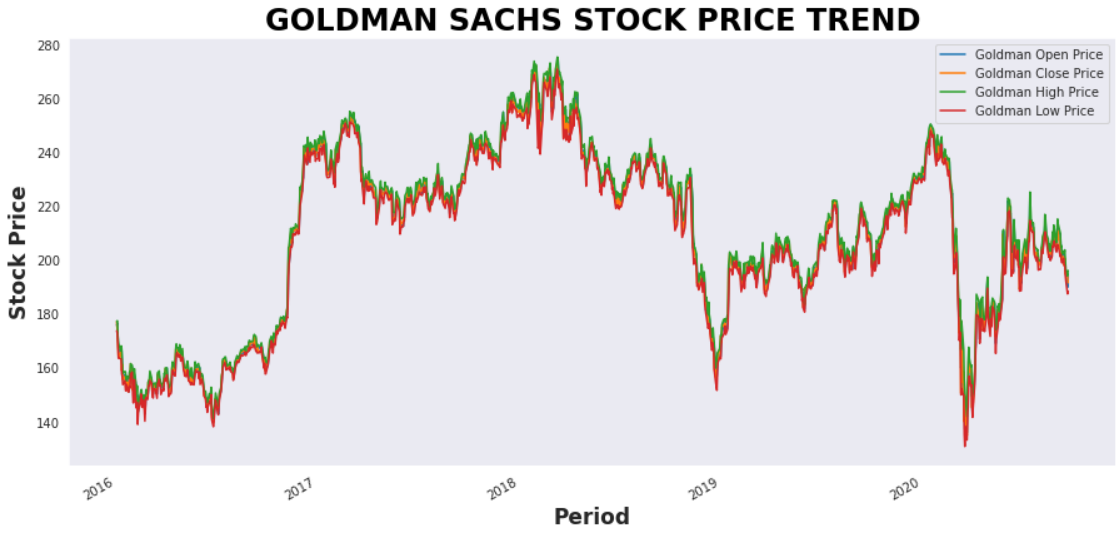


[7]:
```
# First five rows of the data for Goldman Sachs
goldman.head()
```

[7]:

Date	High	Low	Open	Close	Volume	Adj Close
2016-01-04	177.190002	173.759995	175.789993	177.139999	3745500.0	163.456665
2016-01-05	177.500000	172.919998	176.710007	174.089996	4521600.0	160.642242
2016-01-06	172.020004	169.100006	171.309998	169.839996	5539400.0	156.720551
2016-01-07	169.500000	163.600006	166.669998	164.619995	5687900.0	151.903793
2016-01-08	168.419998	163.630005	166.750000	163.940002	4929800.0	151.276291

[8]:
```
goldman['Open'].plot(label = 'Goldman Open Price', figsize=(15, 7))
goldman['Close'].plot(label = 'Goldman Close Price')
goldman['High'].plot(label = 'Goldman High Price')
goldman['Low'].plot(label = 'Goldman Low Price')
plt.legend()
plt.title('GOLDMAN SACHS STOCK PRICE TREND', fontsize = 24, fontweight =␣
↪'bold',color = 'black')
plt.xlabel('Period', fontsize = '17', fontweight = 'bold')
plt.ylabel('Stock Price', fontsize = '17', fontweight = 'bold')
plt.show()
```

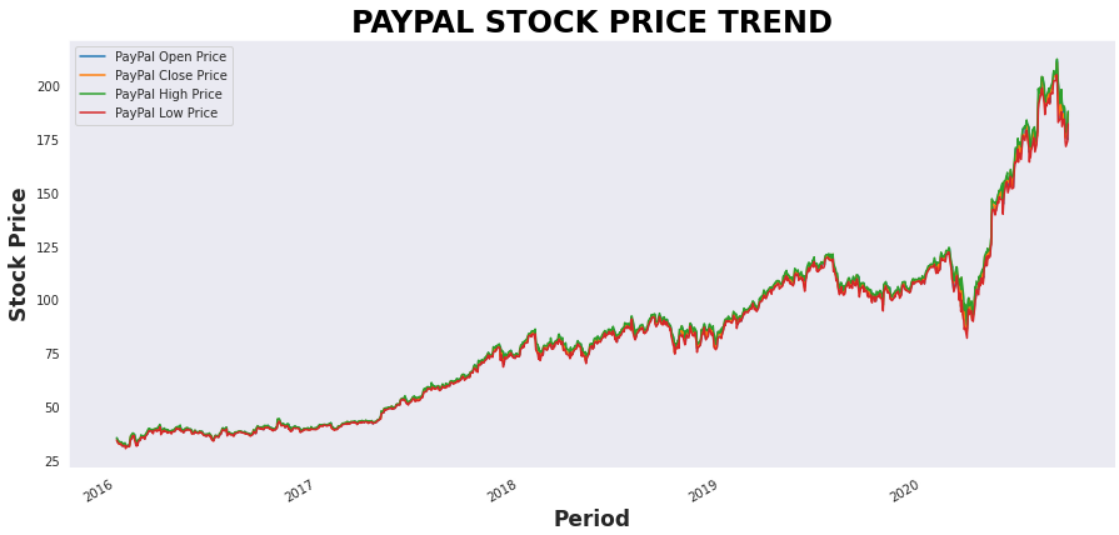


[9]:
```
# First five rows of the data for Paypal
paypal.head()
```

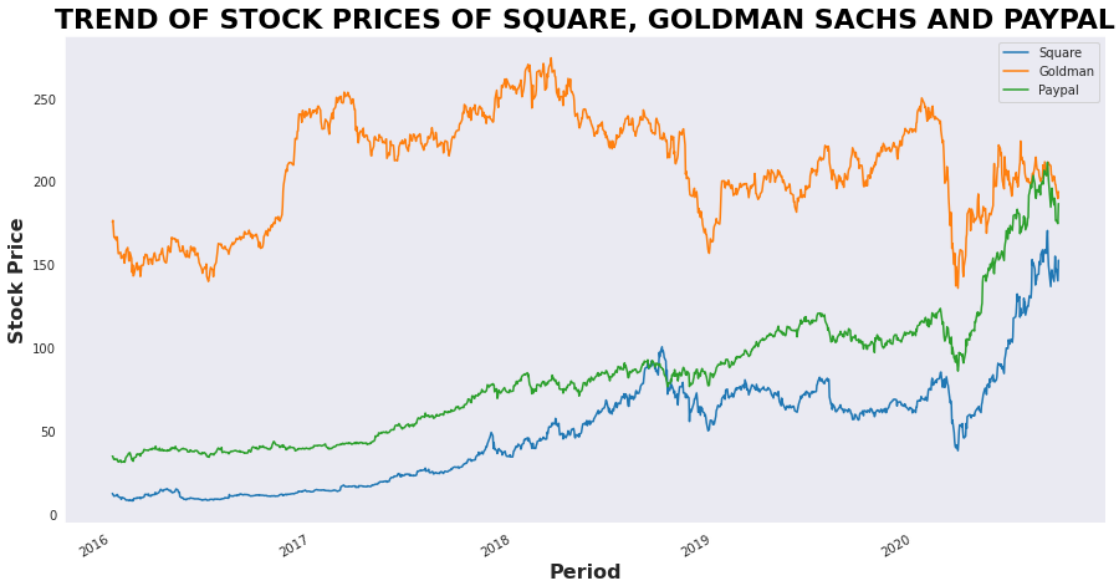
[9]:

Date	High	Low	Open	Close	Volume	Adj Close
2016-01-04	35.560001	34.279999	35.130001	34.750000	12287700	34.750000
2016-01-05	34.980000	33.860001	34.980000	34.310001	11227700	34.310001
2016-01-06	34.009998	33.209999	33.700001	33.980000	8441300	33.980000
2016-01-07	34.160999	33.020000	33.150002	33.130001	11041100	33.130001
2016-01-08	33.880001	32.630001	33.459999	32.689999	7848800	32.689999

[10]:
```
paypal['Open'].plot(label = 'PayPal Open Price', figsize=(15, 7))
paypal['Close'].plot(label = 'PayPal Close Price')
paypal['High'].plot(label = 'PayPal High Price')
paypal['Low'].plot(label = 'PayPal Low Price')
plt.legend()
plt.title('PAYPAL STOCK PRICE TREND', fontsize = 24, fontweight = 'bold',color␣
↪= 'black')
plt.xlabel('Period', fontsize = '17', fontweight = 'bold')
plt.ylabel('Stock Price', fontsize = '17', fontweight = 'bold')
plt.show()
```



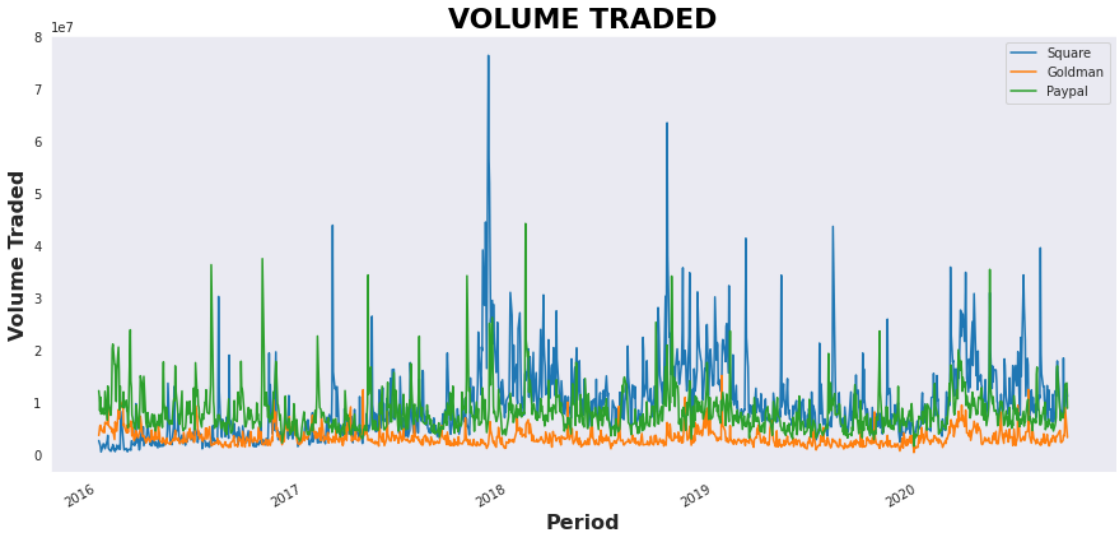
[11]:
```
square['Open'].plot(label="Square", figsize=(15, 8))
goldman["Open"].plot(label="Goldman")
paypal["Open"].plot(label="Paypal")
plt.xlabel('Period', fontsize = '16', fontweight = 'bold')
plt.ylabel('Stock Price', fontsize = '16', fontweight = 'bold')
plt.title("TREND OF STOCK PRICES OF SQUARE, GOLDMAN SACHS AND PAYPAL", fontsize␣
 ↪= 22, fontweight = 'bold',color = 'black')

plt.legend()
plt.show()
```



In the current year, 2020, we can see a common trend in the stock prices for the 3 companies. During the beginning of the year, all 3 stock prices dipped and later rose, with PayPal stock price rising to match Goldman Sachs stock price.

1.0.4 Explore the volume traded of each company's stock price over time

[12]:
```
square['Volume'].plot(label='Square', figsize=(15, 7))
goldman["Volume"].plot(label='Goldman')
paypal['Volume'].plot(label='Paypal')
plt.xlabel('Period', fontsize = '16', fontweight = 'bold')
plt.ylabel("Volume Traded", fontsize = '16', fontweight = 'bold')
plt.title("VOLUME TRADED", fontsize = 22, fontweight = 'bold',color = 'black')

plt.legend()
plt.show()
```



Explore the reason Square had very high volume traded between 2017 - 2018

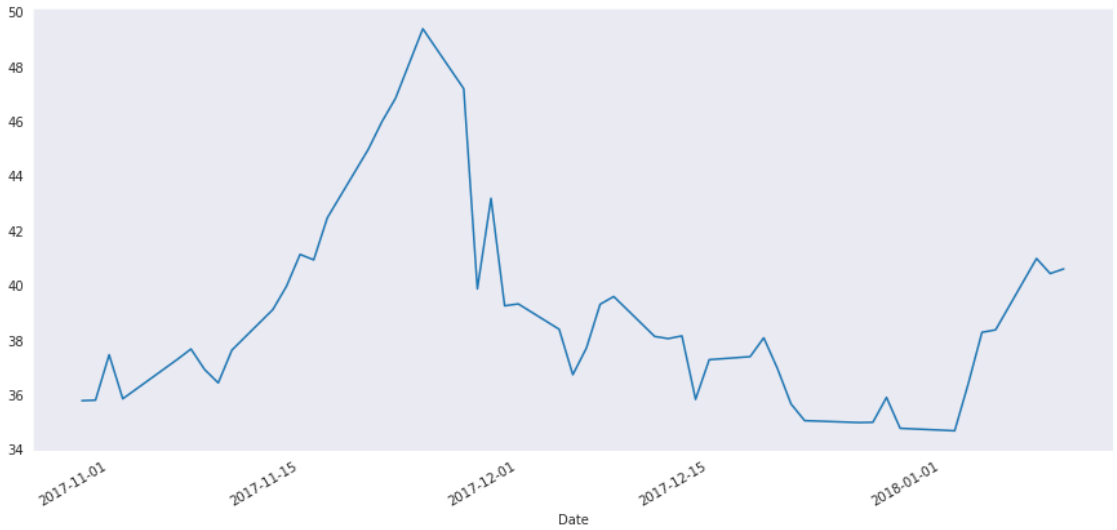
Check for the exact date this happened

[13]:
```
square.iloc[[square['Volume'].argmax()]]
```

[13]:

Date	High	Low	Open	Close	Volume	Adj Close
2017-11-27	47.400002	40.369999	47.209999	41.02	76437500	41.02

[14]:
```
square.iloc[460: 510]['Open'].plot(figsize=(15, 7))
plt.show()
```



Following an article by CNBC on the 3rd Nov 2017. Analysts and Traders were betting big that Square will pop on earnings, hence, the reason for the positive large volume traded.

Link to article below:

https://www.cnbc.com/2017/11/03/traders-are-betting-big-on-valeant-and-square-before-earnings-next-week.html

1.0.5 Explore which company is more valuable via each company's stock price * volume

Calculation is to explore total money traded for the duration for vizualization purposes. This does not represent market capitalization for each company

[15]:
```
square['Total Traded'] = square['Open'] * square['Volume']
goldman['Total Traded'] = goldman['Open'] * goldman['Volume']
paypal['Total Traded'] = paypal['Open'] * paypal['Volume']
```

[16]:
```
square.head()
```

[16]:

Date	High	Low	Open	Close	Volume	Adj Close	Total Traded
2016-01-04	12.90	12.050	12.75	12.16	2751500	12.16	3.508162e+07
2016-01-05	12.34	11.500	12.20	11.51	2352800	11.51	2.870416e+07
2016-01-06	11.64	11.015	11.50	11.52	1850600	11.52	2.128190e+07
2016-01-07	11.37	11.000	11.13	11.16	1636000	11.16	1.820868e+07
2016-01-08	11.54	11.200	11.25	11.31	587300	11.31	6.607125e+06

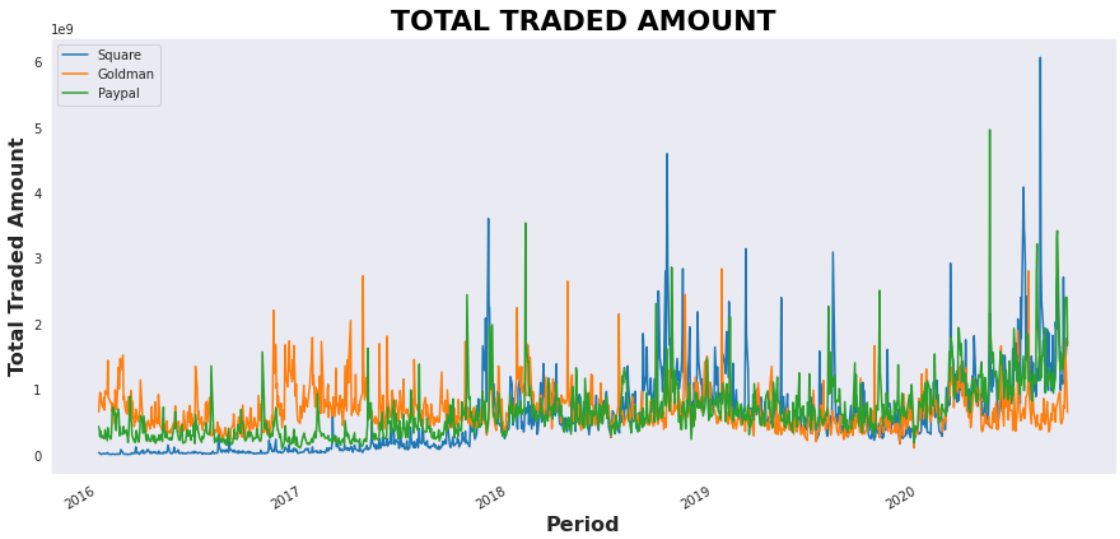
[17]:
```
goldman.head()
```

[17]:
```
                 High         Low        Open       Close     Volume  \
Date
2016-01-04  177.190002  173.759995  175.789993  177.139999  3745500.0
2016-01-05  177.500000  172.919998  176.710007  174.089996  4521600.0
2016-01-06  172.020004  169.100006  171.309998  169.839996  5539400.0
2016-01-07  169.500000  163.600006  166.669998  164.619995  5687900.0
2016-01-08  168.419998  163.630005  166.750000  163.940002  4929800.0

             Adj Close  Total Traded
Date
2016-01-04  163.456665  6.584214e+08
2016-01-05  160.642242  7.990120e+08
2016-01-06  156.720551  9.489546e+08
2016-01-07  151.903793  9.480023e+08
2016-01-08  151.276291  8.220442e+08
```

[18]:
```
paypal.head()
```

[18]:
```
                High        Low       Open      Close    Volume  Adj Close  \
Date
2016-01-04  35.560001  34.279999  35.130001  34.750000  12287700  34.750000
2016-01-05  34.980000  33.860001  34.980000  34.310001  11227700  34.310001
2016-01-06  34.009998  33.209999  33.700001  33.980000   8441300  33.980000
2016-01-07  34.160999  33.020000  33.150002  33.130001  11041100  33.130001
2016-01-08  33.880001  32.630001  33.459999  32.689999   7848800  32.689999

            Total Traded
Date
2016-01-04  4.316669e+08
2016-01-05  3.927449e+08
2016-01-06  2.844718e+08
2016-01-07  3.660125e+08
2016-01-08  2.626208e+08
```

[19]:
```
square['Total Traded'].plot(label="Square", figsize=(15, 7))
goldman['Total Traded'].plot(label="Goldman")
paypal['Total Traded'].plot(label="Paypal")

plt.xlabel('Period', fontsize = '16', fontweight = 'bold')
plt.ylabel("Total Traded Amount", fontsize = '16', fontweight = 'bold')
plt.title("TOTAL TRADED AMOUNT", fontsize = 22, fontweight = 'bold',color =␣
 ↪'black')

plt.legend()
plt.show()
```



Explore the reason Square had very high total traded amount in 2020

Check for the exact date this happened

[20]:
```
square["Total Traded"].argmax()
```

[20]:
```
1155
```

[21]:
```
square.iloc[[square['Total Traded'].argmax()]]
```

[21]:
```
                 High         Low        Open       Close    Volume  \
Date
2020-08-05  158.429993  146.199997  153.160004  146.550003  39629700

             Adj Close  Total Traded
Date
2020-08-05  146.550003  6.069685e+09
```

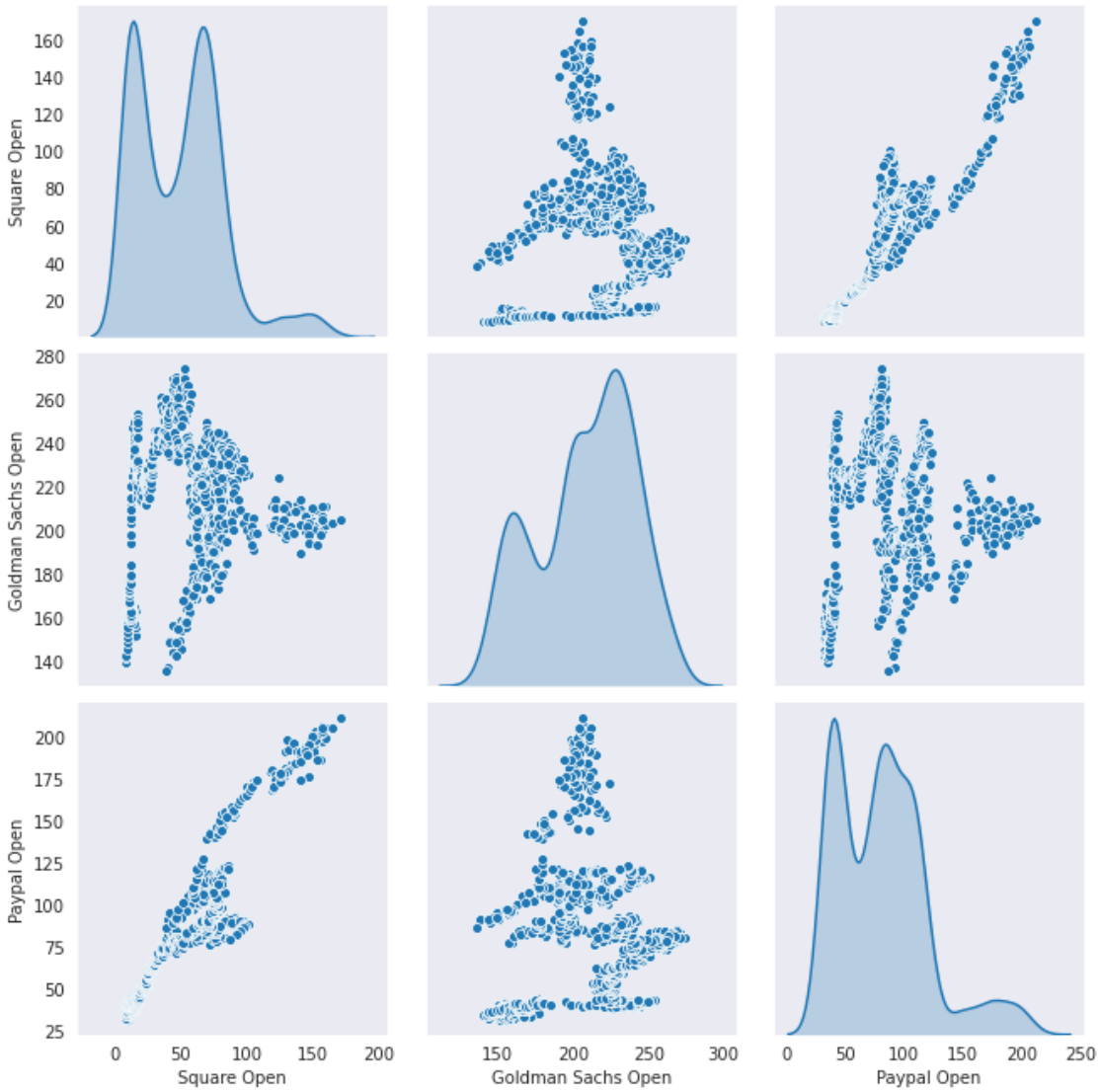
Link to article: https://www.fool.com/investing/2020/08/12/up-120-in-2020-is-it-too-late-to-buy-square-stock/

Excerpts from article: Square (NYSE:SQ) stock recently soared to an all-time high after its second-quarter numbers soundly beat analysts' estimates. The fintech company's revenue surged 64% annually (70% after excluding the divestment of Caviar) to USD 1.92 billion, clearing the consensus forecast by USD 790 million. Its adjusted EBITDA dipped 7% to USD 97.9 million, but earnings of USD 0.18 per share still beat expectations by USD 0.23.

1.0.6 Scatter Matrix to check for linear correlation between Sqaure, PayPal, and Goldman Sachs stock prices

[22]:
```
payment_comp = pd.concat([square['Open'], goldman['Open'], paypal['Open']],␣
↪axis = 1)

payment_comp.columns = ['Square Open', 'Goldman Sachs Open', 'Paypal Open']
```

[23]:
```
#sns.pairplot(payment_comp)
sns.pairplot(payment_comp, diag_kind = 'kde',
             size = 3)
plt.show()
```



From the above scatter matrix, we can say that there is a linear correlation between Square and

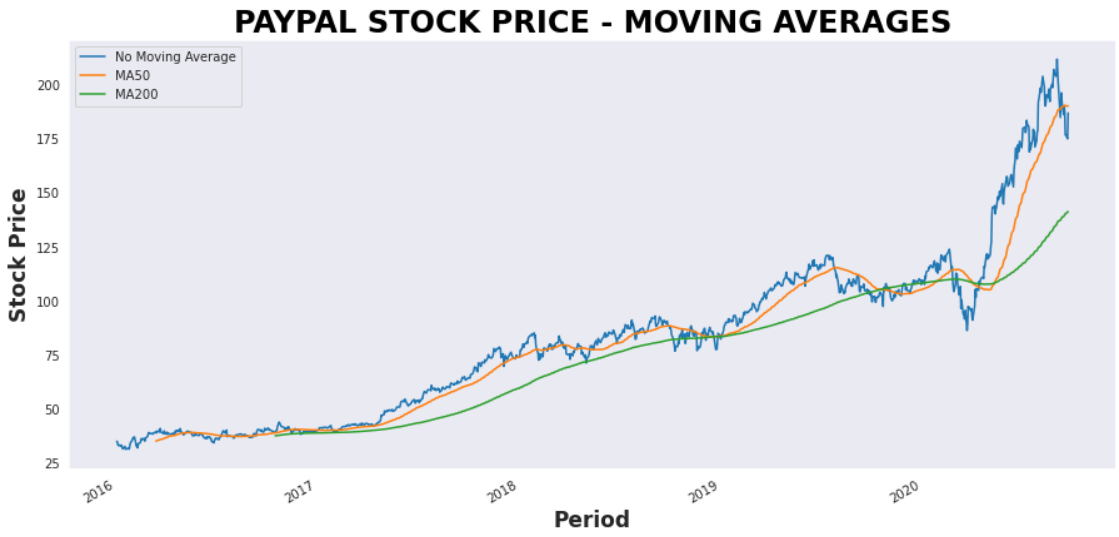
PayPal and less correlation between Goldman Sachs and the other companies.

This may be attributed to the fact that Goldman Sachs is a traditional bank and Sqaure and PayPal are non-traditional fintech organisation who employ the use of the latest technology for their business operations

1.0.7 Explore PayPal Stock Prices

Moving averages are a simple and common type of smoothing used in time series analysis and time series forecasting. Calculating a moving average involves creating a new series where the values are comprised of the average of raw observations in the original time series

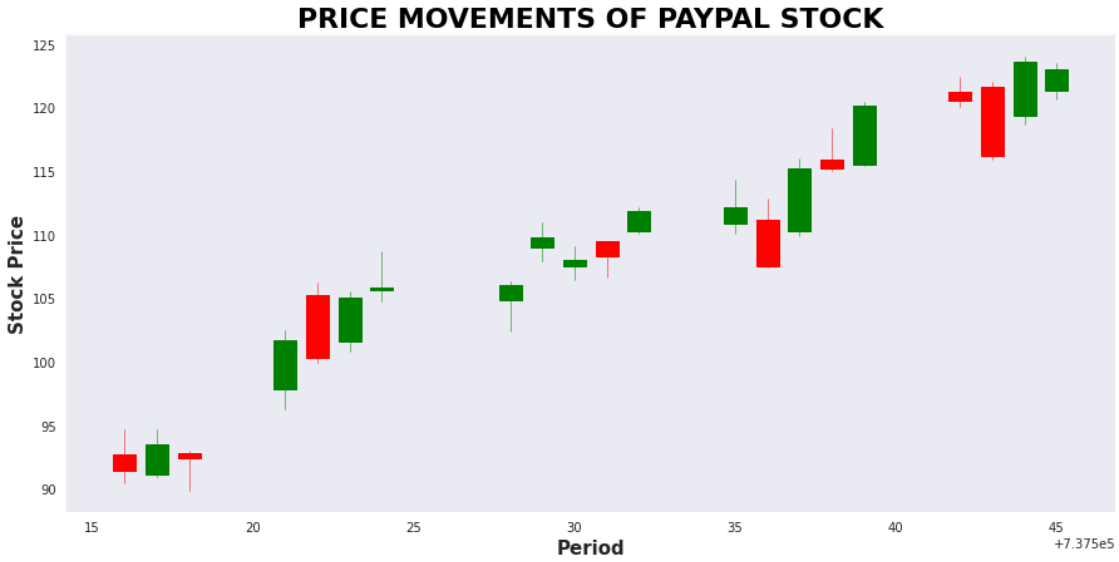
[24]:
```
paypal['Open'].plot(label="No Moving Average", figsize=(15,7))
paypal['MA50'] = paypal['Open'].rolling(50).mean()
paypal['MA50'].plot(label="MA50")
paypal['MA200'] = paypal['Open'].rolling(200).mean()
paypal['MA200'].plot(label="MA200")
plt.title('PAYPAL STOCK PRICE - MOVING AVERAGES', fontsize = 24, fontweight =␣
↪'bold',color = 'black')
plt.xlabel('Period', fontsize = '17', fontweight = 'bold')
plt.ylabel('Stock Price', fontsize = '17', fontweight = 'bold')
plt.legend()
plt.show()
```



Candle stick visualization of price movements

[25]:
```
paypal_reset = paypal.loc['2020-04': '2020-04'].reset_index()

paypal_reset['date_ax'] = paypal_reset['Date'].apply(lambda date:␣
↪date2num(date))
```

```
paypal_values = [tuple(vals) for vals in paypal_reset[['date_ax', 'Open',␣
 ↪'High', 'Low', 'Close']].values]

modays = WeekdayLocator(MONDAY)
alldays = DayLocator()
weekFormatter = DateFormatter('%b %d')
dayFormatter = DateFormatter('%d')

fig, ax = plt.subplots(figsize=(15, 7))
candlestick_ohlc(ax, paypal_values, width=0.7, colorup="g", colordown="r")
plt.xlabel('Period', fontsize = '15', fontweight = 'bold')
plt.ylabel('Stock Price', fontsize = '15', fontweight = 'bold')
plt.title('PRICE MOVEMENTS OF PAYPAL STOCK', fontsize = 22, fontweight =␣
 ↪'bold',color = 'black')
plt.show()
```



1.0.8 Daily Percentage Change of Square, PayPal, and Goldman Sachs for analysing volatility of stock price

$r_t = \frac{p_t}{p_{t-1}} - 1$

[26]:
```
square['Returns'] = (square['Close']/square['Close'].shift(1)) - 1

square.head()
```

[26]:

Date	High	Low	Open	Close	Volume	Adj Close	Total Traded
2016-01-04	12.90	12.050	12.75	12.16	2751500	12.16	3.508162e+07

```
2016-01-05  12.34  11.500  12.20  11.51  2352800      11.51  2.870416e+07
2016-01-06  11.64  11.015  11.50  11.52  1850600      11.52  2.128190e+07
2016-01-07  11.37  11.000  11.13  11.16  1636000      11.16  1.820868e+07
2016-01-08  11.54  11.200  11.25  11.31   587300      11.31  6.607125e+06

              Returns
Date
2016-01-04       NaN
2016-01-05 -0.053454
2016-01-06  0.000869
2016-01-07 -0.031250
2016-01-08  0.013441
```

[27]:
```
goldman['Returns'] = (goldman['Close']/goldman['Close'].shift(1)) - 1

goldman.head()
```

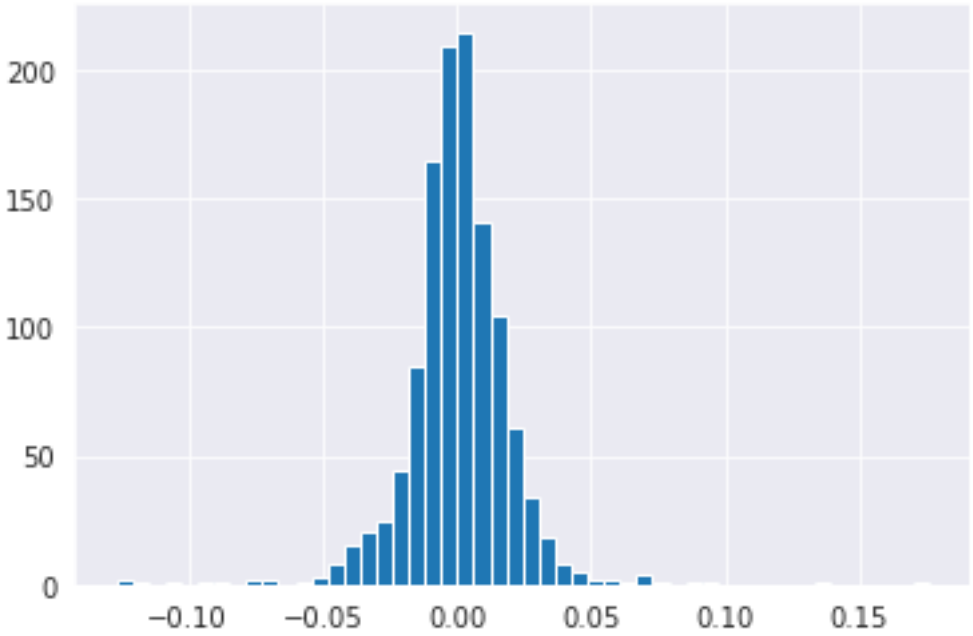
[27]:
```
                 High         Low        Open       Close     Volume  \
Date
2016-01-04  177.190002  173.759995  175.789993  177.139999  3745500.0
2016-01-05  177.500000  172.919998  176.710007  174.089996  4521600.0
2016-01-06  172.020004  169.100006  171.309998  169.839996  5539400.0
2016-01-07  169.500000  163.600006  166.669998  164.619995  5687900.0
2016-01-08  168.419998  163.630005  166.750000  163.940002  4929800.0

              Adj Close  Total Traded   Returns
Date
2016-01-04  163.456665  6.584214e+08       NaN
2016-01-05  160.642242  7.990120e+08 -0.017218
2016-01-06  156.720551  9.489546e+08 -0.024413
2016-01-07  151.903793  9.480023e+08 -0.030735
2016-01-08  151.276291  8.220442e+08 -0.004131
```

[28]:
```
paypal['Returns'] = (paypal['Close']/paypal['Close'].shift(1)) - 1

paypal.head()
```

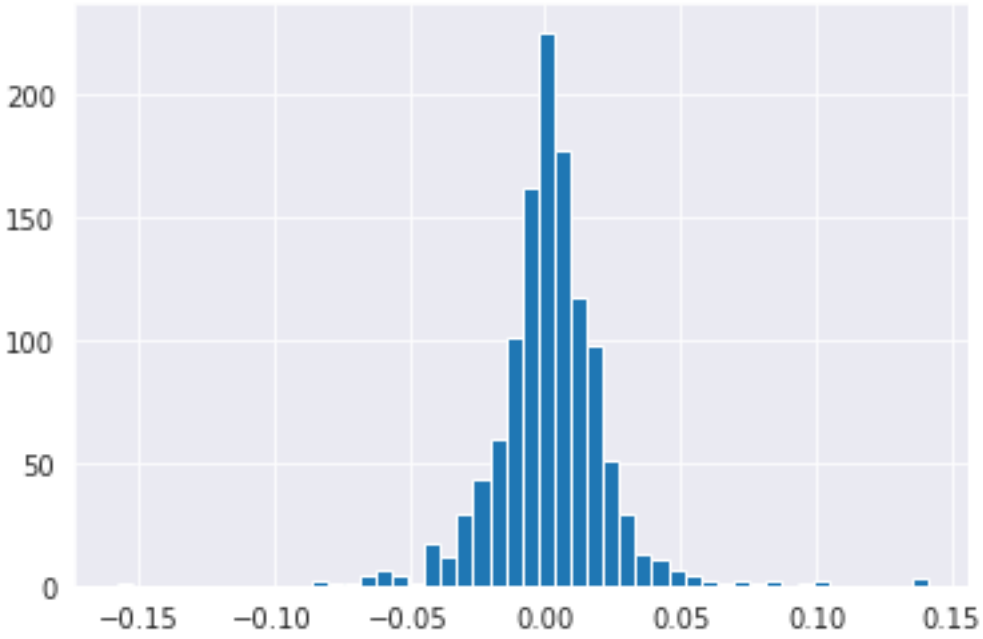
[28]:
```
                 High        Low       Open      Close    Volume  Adj Close  \
Date
2016-01-04  35.560001  34.279999  35.130001  34.750000  12287700  34.750000
2016-01-05  34.980000  33.860001  34.980000  34.310001  11227700  34.310001
2016-01-06  34.009998  33.209999  33.700001  33.980000   8441300  33.980000
2016-01-07  34.160999  33.020000  33.150002  33.130001  11041100  33.130001
2016-01-08  33.880001  32.630001  33.459999  32.689999   7848800  32.689999

            Total Traded  MA50  MA200   Returns
Date
```

```
2016-01-04  4.316669e+08   NaN    NaN       NaN
2016-01-05  3.927449e+08   NaN    NaN -0.012662
2016-01-06  2.844718e+08   NaN    NaN -0.009618
2016-01-07  3.660125e+08   NaN    NaN -0.025015
2016-01-08  2.626208e+08   NaN    NaN -0.013281
```

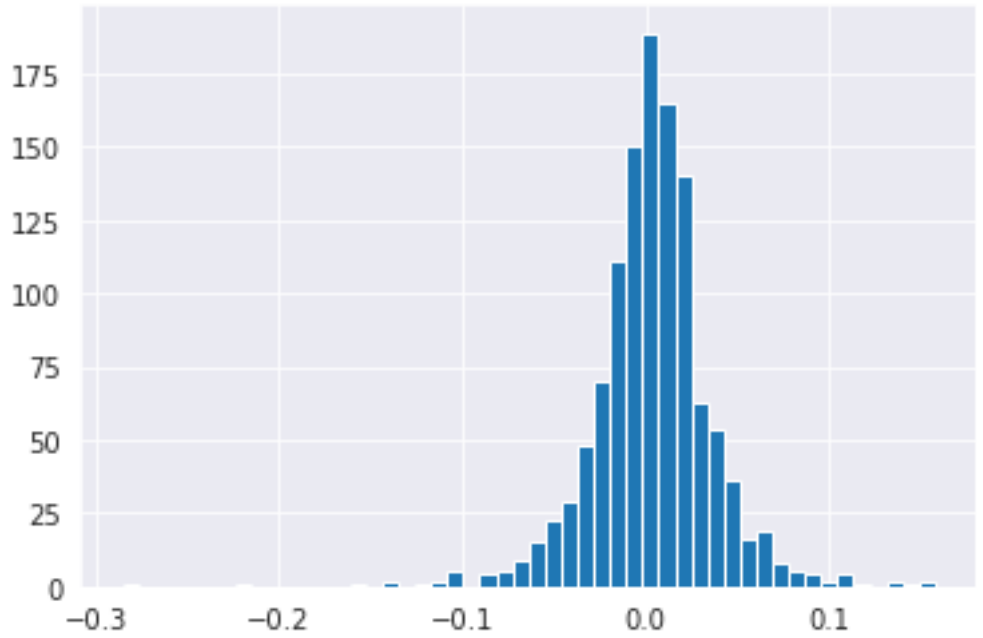
[29]:
```
goldman['Returns'].hist(bins=50)
plt.show()
```



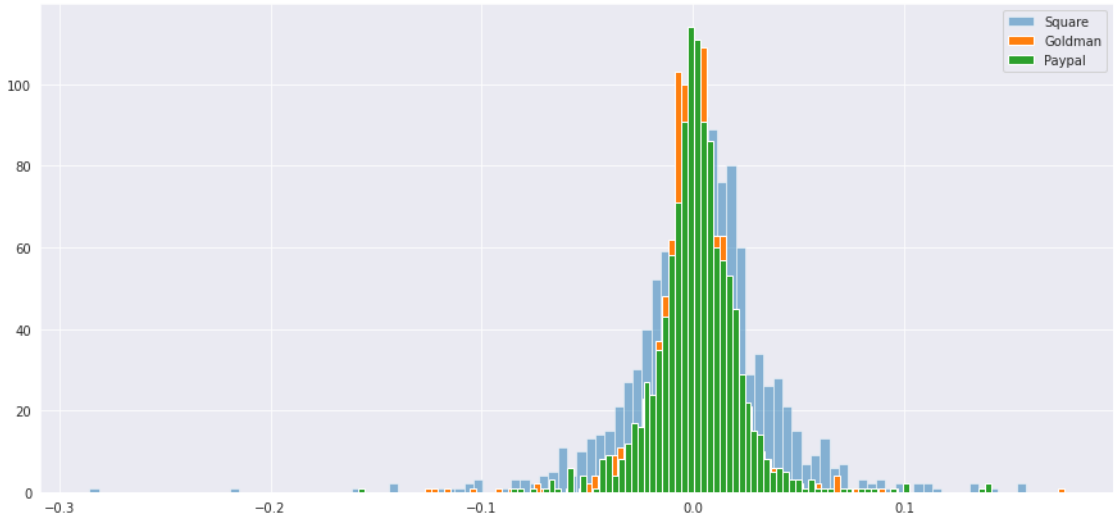
[30]:
```
paypal['Returns'].hist(bins=50)
plt.show()
```



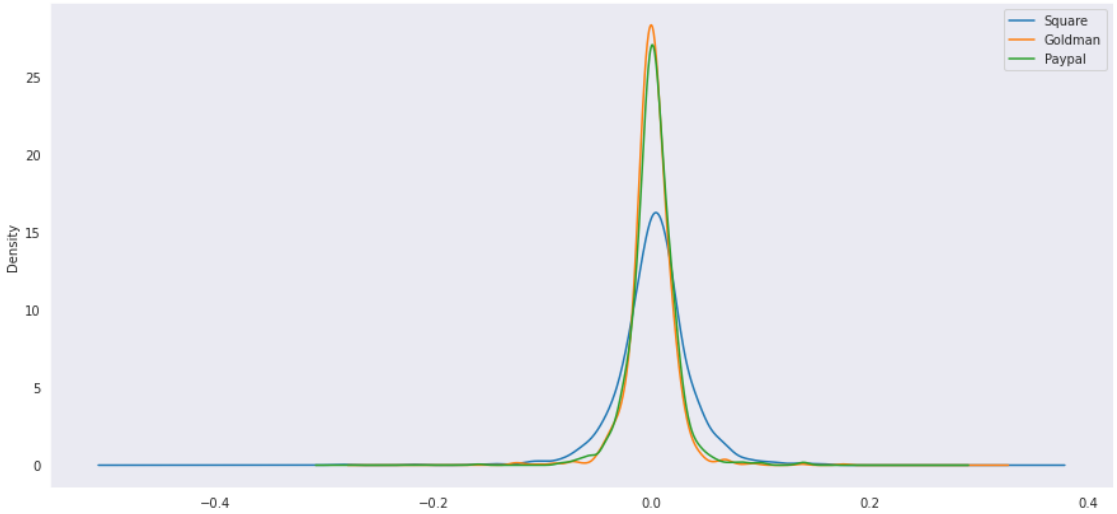
[31]:
```
square['Returns'].hist(bins=50)
plt.show()
```



[32]:
```
square['Returns'].hist(bins=100, label='Square', alpha=0.5, figsize=(15,7))
goldman['Returns'].hist(bins=100, label='Goldman')
paypal['Returns'].hist(bins=100, label="Paypal")
plt.legend()
plt.show()
```



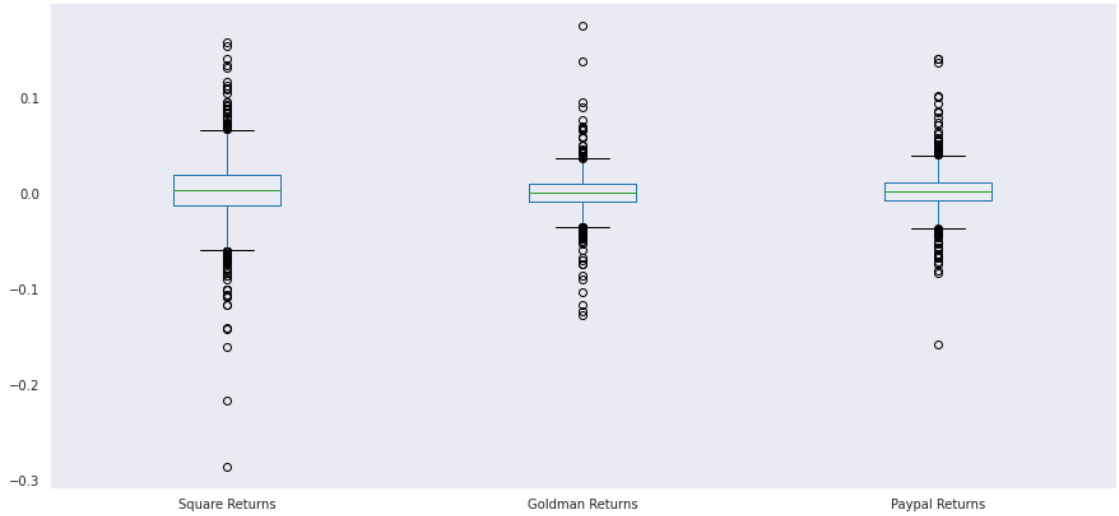
[33]:
```
square['Returns'].plot(kind='kde', label='Square', figsize=(15,7))
goldman['Returns'].plot(kind='kde', label='Goldman')
paypal['Returns'].plot(kind='kde', label='Paypal')
plt.legend()
plt.show()
```



The KDE curve of Square is fatter/wider than that of Goldman Sachs and PayPal meaning that Square's stock is most volatile amongst the three.

We can also say that the KDE curve of PayPal and Goldman Sachs are very close to each other

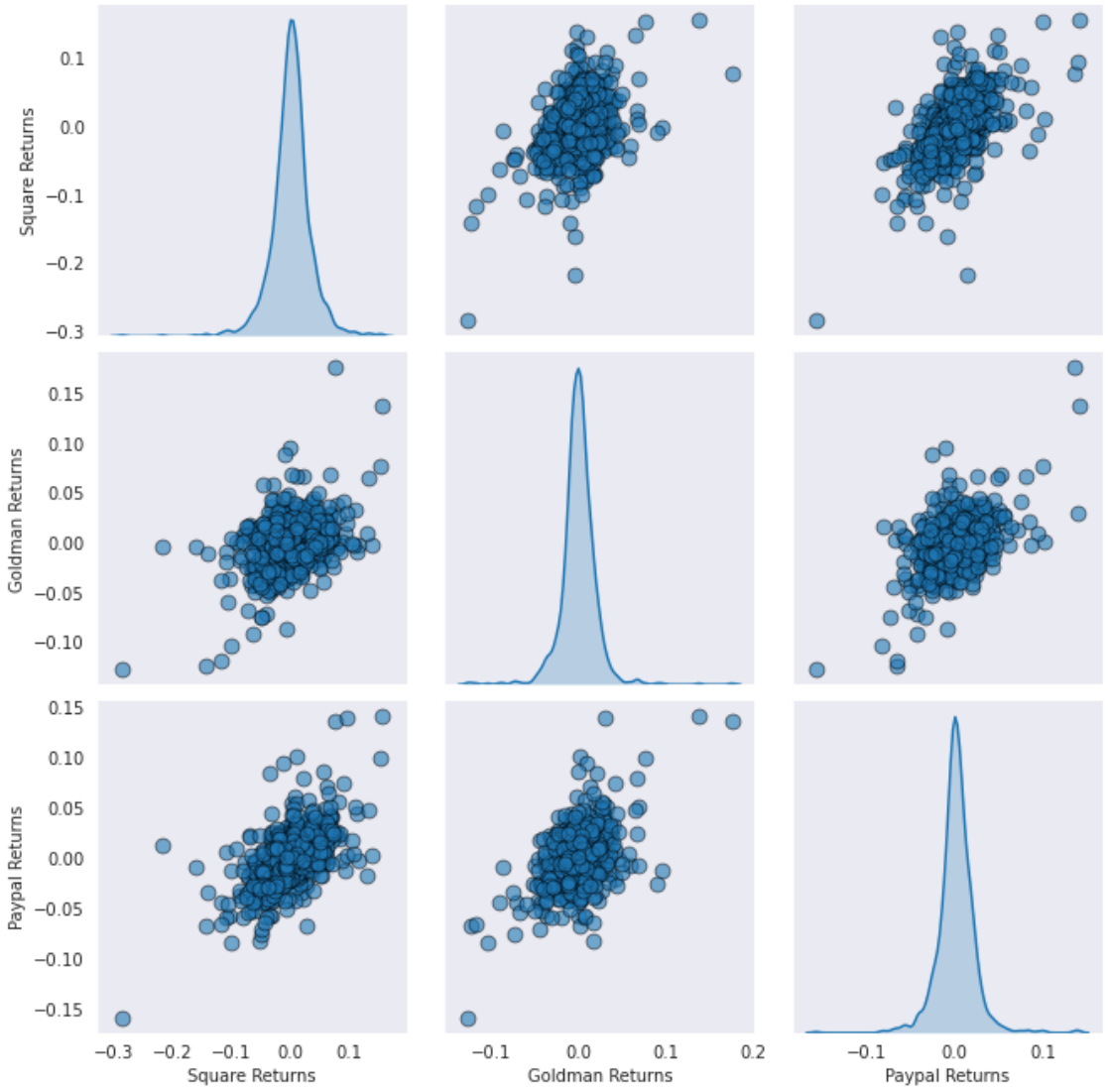
[34]:
```
box_df = pd.concat([square['Returns'], goldman['Returns'], paypal['Returns']],␣
↪axis=1)
box_df.columns=['Square Returns', 'Goldman Returns', 'Paypal Returns']

box_df.plot(kind='box', figsize= (15,7))
plt.show()
```



From the above box plot, We can also say that the daily percentage change/return for Square is more spread out than that of Goldman Sachs and PayPal. Square is more volatile than Goldman Sachs and PayPal

1.0.9 Scatter Matrix to check for linear correlation between Sqaure, PayPal, and Goldman Sachs stock daily return

[35]:
```
#scatter_matrix(box_df, figsize=(8,8), hist_kwds={'bins': 50}, alpha=0.25)
sns.pairplot(box_df, diag_kind = 'kde',
             plot_kws = {'alpha': 0.6, 's': 80, 'edgecolor': 'k'},
             size = 3)
plt.show()
```



From the above scatter matrix, we can say that there is a little correlation between Square, Goldman Sachs, and PayPal volatility/stability

1.0.10 Cumulative Return

$i_t = (1 + r_t)i_{t-1} = (1 + \frac{p_t}{p_{t-1}} - 1)i_{t-1} = \frac{p_t}{p_{t-1}}\ i_{t-1}$

[36]:
```
square['Cumulative Return'] = (1 + square['Returns']).cumprod()
```

[37]:
```
square.head()
```

[37]:

Date	High	Low	Open	Close	Volume	Adj Close	Total Traded
2016-01-04	12.90	12.050	12.75	12.16	2751500	12.16	3.508162e+07

```
2016-01-05  12.34  11.500  12.20  11.51  2352800      11.51  2.870416e+07
2016-01-06  11.64  11.015  11.50  11.52  1850600      11.52  2.128190e+07
2016-01-07  11.37  11.000  11.13  11.16  1636000      11.16  1.820868e+07
2016-01-08  11.54  11.200  11.25  11.31   587300      11.31  6.607125e+06

             Returns  Cumulative Return
Date
2016-01-04       NaN               NaN
2016-01-05 -0.053454          0.946546
2016-01-06  0.000869          0.947368
2016-01-07 -0.031250          0.917763
2016-01-08  0.013441          0.930099
```

[38]:
```
goldman['Cumulative Return'] = (1 + goldman['Returns']).cumprod()
```

[39]:
```
goldman.head()
```

[39]:
```
                 High         Low        Open       Close     Volume  \
Date
2016-01-04  177.190002  173.759995  175.789993  177.139999  3745500.0
2016-01-05  177.500000  172.919998  176.710007  174.089996  4521600.0
2016-01-06  172.020004  169.100006  171.309998  169.839996  5539400.0
2016-01-07  169.500000  163.600006  166.669998  164.619995  5687900.0
2016-01-08  168.419998  163.630005  166.750000  163.940002  4929800.0

             Adj Close  Total Traded   Returns  Cumulative Return
Date
2016-01-04  163.456665  6.584214e+08       NaN                NaN
2016-01-05  160.642242  7.990120e+08 -0.017218           0.982782
2016-01-06  156.720551  9.489546e+08 -0.024413           0.958790
2016-01-07  151.903793  9.480023e+08 -0.030735           0.929321
2016-01-08  151.276291  8.220442e+08 -0.004131           0.925483
```

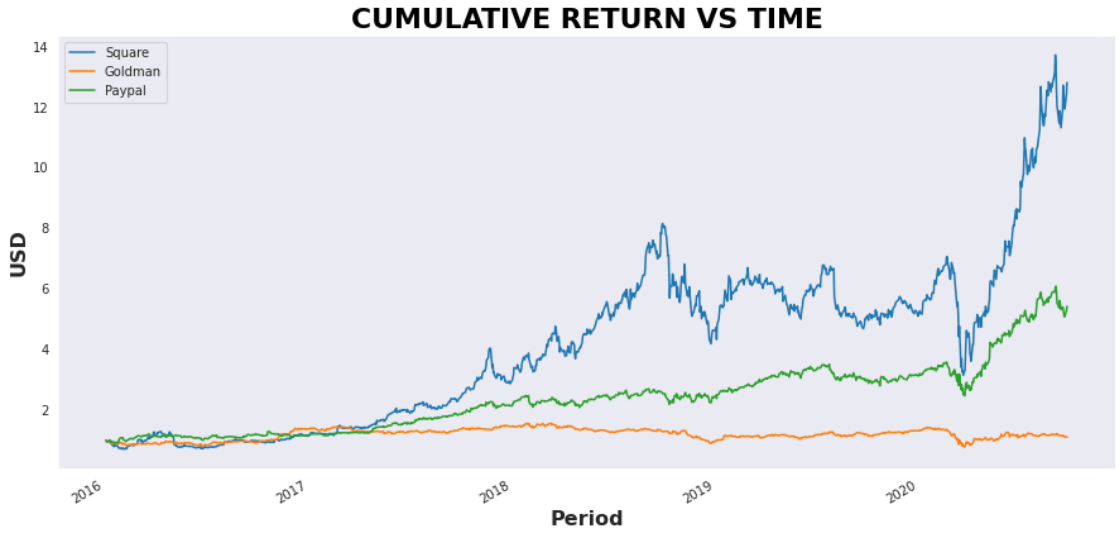
[40]:
```
paypal['Cumulative Return'] = (1 + paypal['Returns']).cumprod()
```

[41]:
```
paypal.head()
```

[41]:
```
                High        Low       Open      Close    Volume  Adj Close  \
Date
2016-01-04  35.560001  34.279999  35.130001  34.750000  12287700  34.750000
2016-01-05  34.980000  33.860001  34.980000  34.310001  11227700  34.310001
2016-01-06  34.009998  33.209999  33.700001  33.980000   8441300  33.980000
2016-01-07  34.160999  33.020000  33.150002  33.130001  11041100  33.130001
2016-01-08  33.880001  32.630001  33.459999  32.689999   7848800  32.689999

            Total Traded  MA50  MA200   Returns  Cumulative Return
Date
```

```
2016-01-04  4.316669e+08   NaN    NaN       NaN                NaN
2016-01-05  3.927449e+08   NaN    NaN -0.012662           0.987338
2016-01-06  2.844718e+08   NaN    NaN -0.009618           0.977842
2016-01-07  3.660125e+08   NaN    NaN -0.025015           0.953381
2016-01-08  2.626208e+08   NaN    NaN -0.013281           0.940719
```

[42]:
```
square['Cumulative Return'].plot(label="Square", figsize=(15, 7))
goldman['Cumulative Return'].plot(label="Goldman")
paypal['Cumulative Return'].plot(label="Paypal")
plt.xlabel('Period', fontsize = '16', fontweight = 'bold')
plt.ylabel('USD', fontsize = '16', fontweight = 'bold')
plt.title("CUMULATIVE RETURN VS TIME", fontsize = 22, fontweight = 'bold',color␣
↪= 'black')
plt.legend()
plt.show()
```



Which Stock showed the highest return for a 1 USD investment before 01-01-2016 to 22-09-2020?

Answer - Square

Then, followed by PayPal, with Goldman Sachs having the lowest return.