A Comparison of Oil Stocks vs. Commodity Price

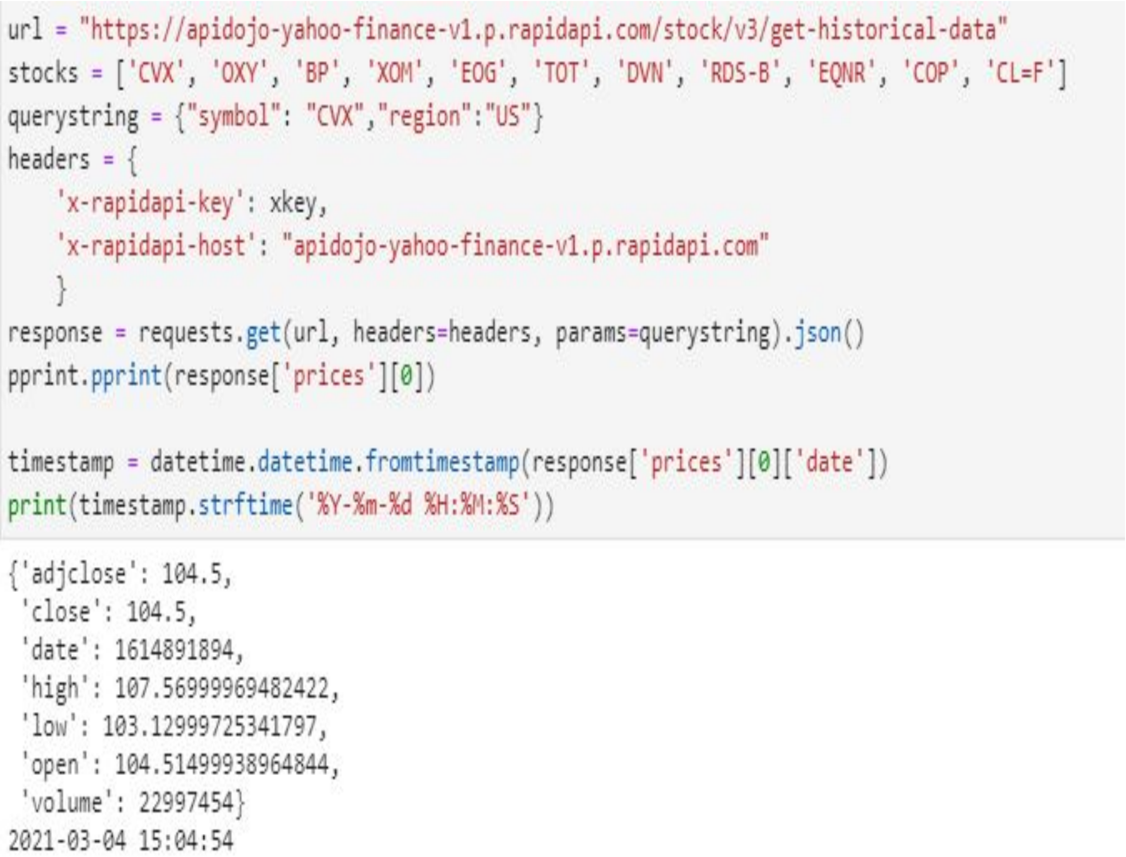
Producers
Lydia Jones
Tyler Nguyen
Freddy Rangel
Michael Thomas

GOALS

How can we utilize the Yahoo! Finance API to decide a good time to buy oil stocks?

Yahoo! Finance API:

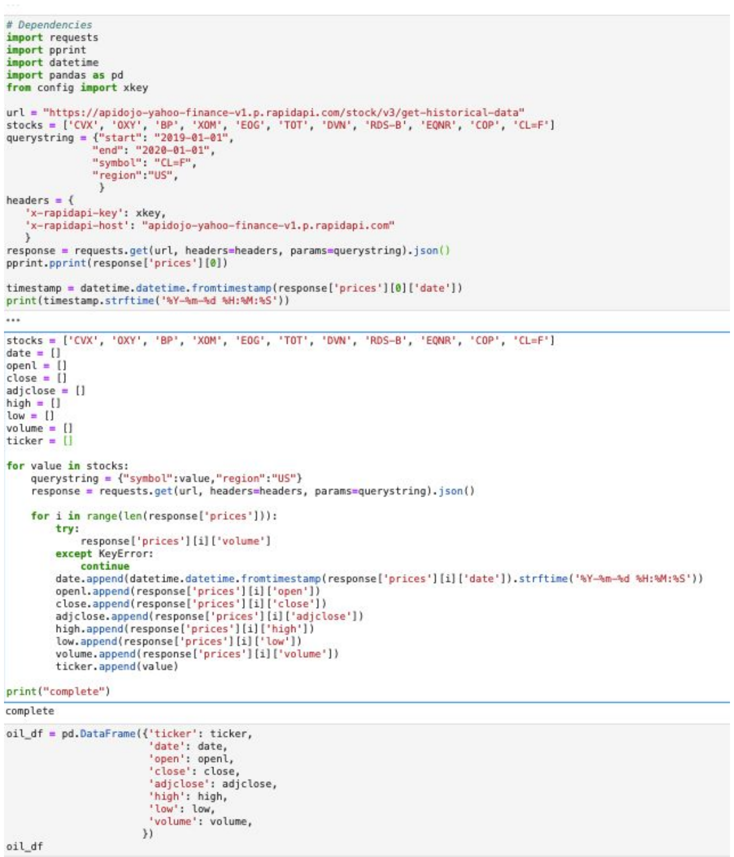
Yahoo API:Initially, Yahoo Finance API was selected for the project since it contained most of the required data. However, the historical data was not as malleable as needed for this project. The API key was provided by rapidapi.com and the tickers for the query from yahoo finances.



```
url = "https://apidojo-yahoo-finance-v1.p.rapidapi.com/stock/v3/get-historical-data"
stocks = ['CVX', 'OXY', 'BP', 'XOM', 'EOG', 'TOT', 'DVN', 'RDS-B', 'EQNR', 'COP', 'CL=F']
querystring = {"symbol": "CVX","region":"US"}
headers = {
    'x-rapidapi-key': xkey,
    'x-rapidapi-host': "apidojo-yahoo-finance-v1.p.rapidapi.com"
    }
response = requests.get(url, headers=headers, params=querystring).json()
pprint.pprint(response['prices'][0])

timestamp = datetime.datetime.fromtimestamp(response['prices'][0]['date'])
print(timestamp.strftime('%Y-%m-%d %H:%M:%S'))
```

```
{'adjclose': 104.5,
 'close': 104.5,
 'date': 1614891894,
 'high': 107.56999969482422,
 'low': 103.12999725341797,
 'open': 104.51499938964844,
 'volume': 22997454}
2021-03-04 15:04:54
```

How much cleaning did we need to do? What Libraries did we use?



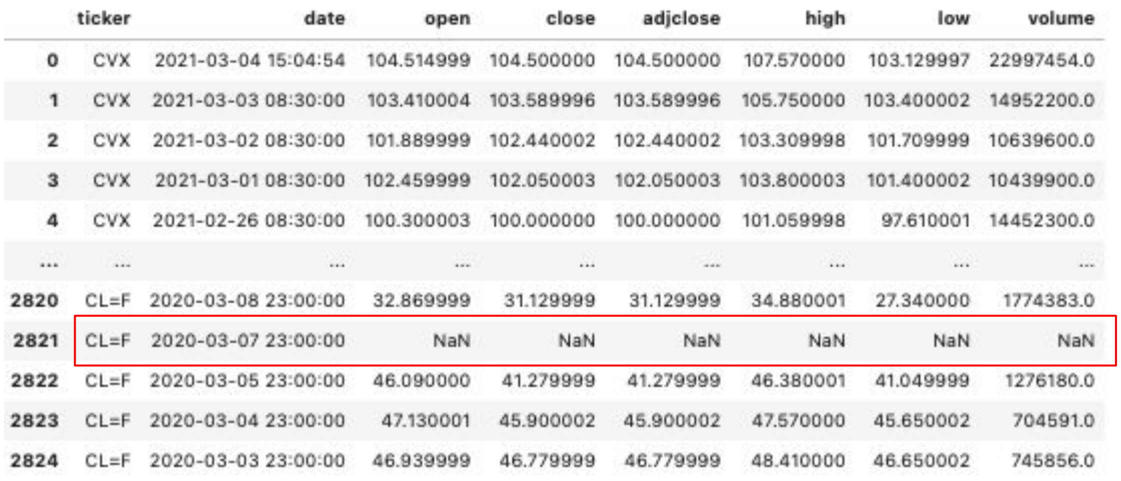
```
# Dependencies
import requests
import pprint
import datetime
import pandas as pd
from config import xkey

url = "https://apidojo-yahoo-finance-v1.p.rapidapi.com/stock/v3/get-historical-data"
stocks = ['CVX', 'OXY', 'BP', 'XOM', 'EOG', 'TOT', 'DVN', 'RDS-B', 'EQNR', 'COP', 'CL=F']
querystring = {"start": "2019-01-01",
               "end": "2020-01-01",
               "symbol": "CL=F",
               "region":"US",
                }
headers = {
    'x-rapidapi-key': xkey,
    'x-rapidapi-host': "apidojo-yahoo-finance-v1.p.rapidapi.com"
    }
response = requests.get(url, headers=headers, params=querystring).json()
pprint.pprint(response['prices'][0])

timestamp = datetime.datetime.fromtimestamp(response['prices'][0]['date'])
print(timestamp.strftime('%Y-%m-%d %H:%M:%S'))
```

```
...
```

```
stocks = ['CVX', 'OXY', 'BP', 'XOM', 'EOG', 'TOT', 'DVN', 'RDS-B', 'EQNR', 'COP', 'CL=F']
date = []
openl = []
close = []
adjclose = []
high = []
low = []
volume = []
ticker = []

for value in stocks:
    querystring = {"symbol":value,"region":"US"}
    response = requests.get(url, headers=headers, params=querystring).json()

    for i in range(len(response['prices'])):
        try:
            response['prices'][i]['volume']
        except KeyError:
            continue
        date.append(datetime.datetime.fromtimestamp(response['prices'][i]['date']).strftime('%Y-%m-%d %H:%M:%S'))
        openl.append(response['prices'][i]['open'])
        close.append(response['prices'][i]['close'])
        adjclose.append(response['prices'][i]['adjclose'])
        high.append(response['prices'][i]['high'])
        low.append(response['prices'][i]['low'])
        volume.append(response['prices'][i]['volume'])
        ticker.append(value)

print("complete")
```

```
complete
```

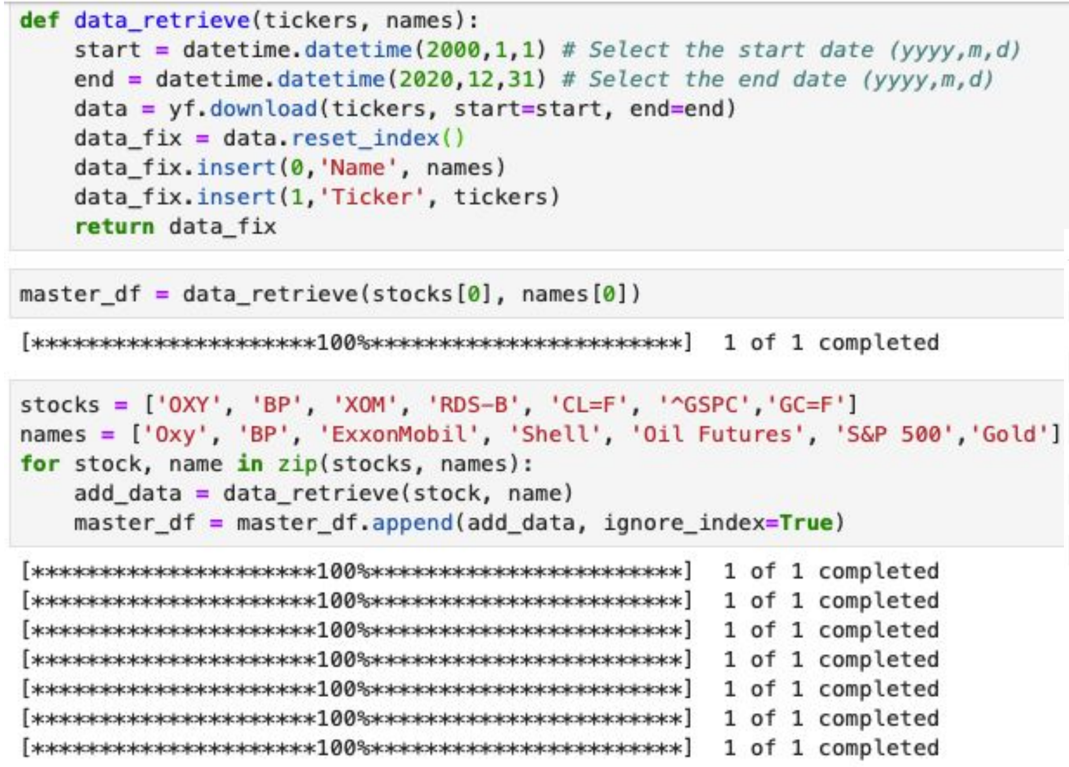
```
oil_df = pd.DataFrame({'ticker': ticker,
                       'date': date,
                       'open': openl,
                       'close': close,
                       'adjclose': adjclose,
                       'high': high,
                       'low': low,
                       'volume': volume,
                      })
oil_df
```



	ticker	date	open	close	adjclose	high	low	volume
0	CVX	2021-03-04 15:04:54	104.514999	104.500000	104.500000	107.570000	103.129997	22997454.0
1	CVX	2021-03-03 08:30:00	103.410004	103.589996	103.589996	105.750000	103.400002	14952200.0
2	CVX	2021-03-02 08:30:00	101.889999	102.440002	102.440002	103.309998	101.709999	10639600.0
3	CVX	2021-03-01 08:30:00	102.459999	102.050003	102.050003	103.800003	101.400002	10439900.0
4	CVX	2021-02-26 08:30:00	100.300003	100.000000	100.000000	101.059998	97.610001	14452300.0
...	...	...	...	...	...	...	...	...
2820	CL=F	2020-03-08 23:00:00	32.869999	31.129999	31.129999	34.880001	27.340000	1774383.0
2821	CL=F	2020-03-07 23:00:00	NaN	NaN	NaN	NaN	NaN	NaN
2822	CL=F	2020-03-05 23:00:00	46.090000	41.279999	41.279999	46.380001	41.049999	1276180.0
2823	CL=F	2020-03-04 23:00:00	47.130001	45.900002	45.900002	47.570000	45.650002	704591.0
2824	CL=F	2020-03-03 23:00:00	46.939999	46.779999	46.779999	48.410000	46.650002	745856.0

* NaN values are due to market closes on the weekends

How much cleaning did we need to do? What data did we drop?



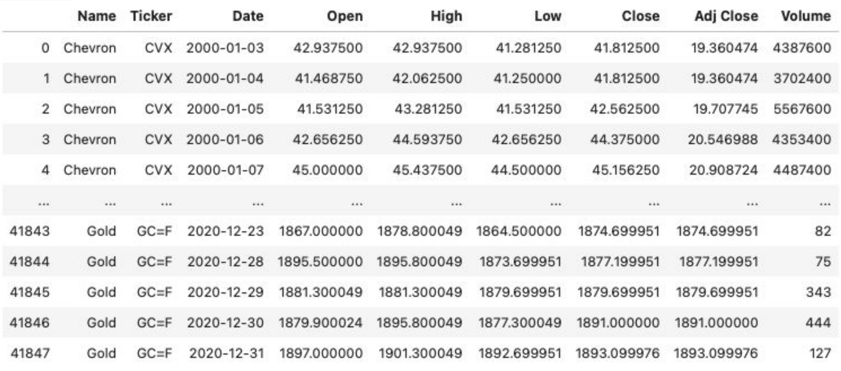
```
def data_retrieve(tickers, names):
    start = datetime.datetime(2000,1,1) # Select the start date (yyyy,m,d)
    end = datetime.datetime(2020,12,31) # Select the end date (yyyy,m,d)
    data = yf.download(tickers, start=start, end=end)
    data_fix = data.reset_index()
    data_fix.insert(0,'Name', names)
    data_fix.insert(1,'Ticker', tickers)
    return data_fix
```

```
master_df = data_retrieve(stocks[0], names[0])
```

```
[*********************100%***********************]  1 of 1 completed
```

```
stocks = ['OXY', 'BP', 'XOM', 'RDS-B', 'CL=F', '^GSPC','GC=F']
names = ['Oxy', 'BP', 'ExxonMobil', 'Shell', 'Oil Futures', 'S&P 500','Gold']
for stock, name in zip(stocks, names):
    add_data = data_retrieve(stock, name)
    master_df = master_df.append(add_data, ignore_index=True)
```

```
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
```



	Name	Ticker	Date	Open	High	Low	Close	Adj Close	Volume
0	Chevron	CVX	2000-01-03	42.937500	42.937500	41.281250	41.812500	19.360474	4387600
1	Chevron	CVX	2000-01-04	41.468750	42.062500	41.250000	41.812500	19.360474	3702400
2	Chevron	CVX	2000-01-05	41.531250	43.281250	41.531250	42.562500	19.707745	5567600
3	Chevron	CVX	2000-01-06	42.656250	44.593750	42.656250	44.375000	20.546988	4353400
4	Chevron	CVX	2000-01-07	45.000000	45.437500	44.500000	45.156250	20.908724	4487400
...	...	...	...	...	...	...	...	...	...
41843	Gold	GC=F	2020-12-23	1867.000000	1878.800049	1864.500000	1874.699951	1874.699951	82
41844	Gold	GC=F	2020-12-28	1895.500000	1895.800049	1873.699951	1877.199951	1877.199951	75
41845	Gold	GC=F	2020-12-29	1881.300049	1881.300049	1879.699951	1879.699951	1879.699951	343
41846	Gold	GC=F	2020-12-30	1879.900024	1895.800049	1877.300049	1891.000000	1891.000000	444
41847	Gold	GC=F	2020-12-31	1897.000000	1901.300049	1892.699951	1893.099976	1893.099976	127

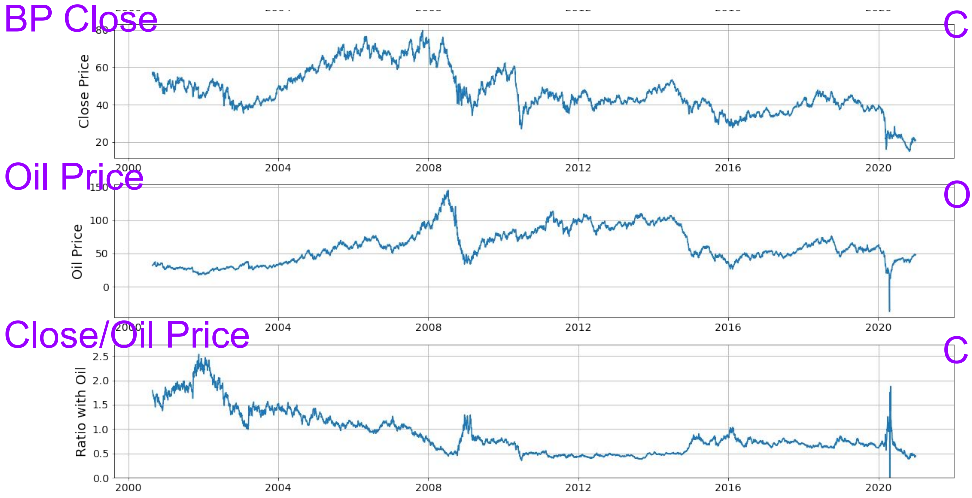
Did we use any new features/concepts/code not previously introduced in class?

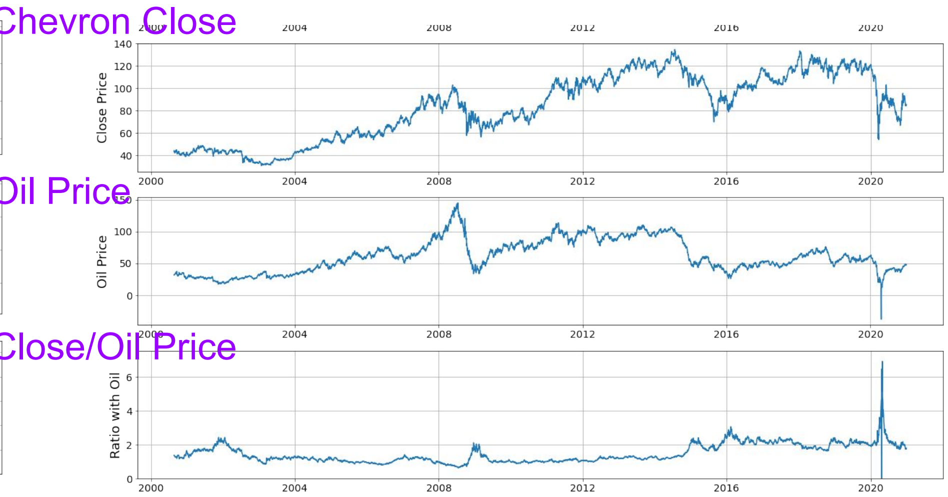
- yfinance wrapper is our new territory.
- Struggles to understand the syntax and manipulation.
- With some trials and errors, able to figure out how it works
- Pull data and generate graphs by defining functions, and the use of some libraries:
    - Pandas
    - Matplolib
    - Datetime

QUESTIONS...

1. **To what extent is oil price a proxy for the stock price for an oil company?**

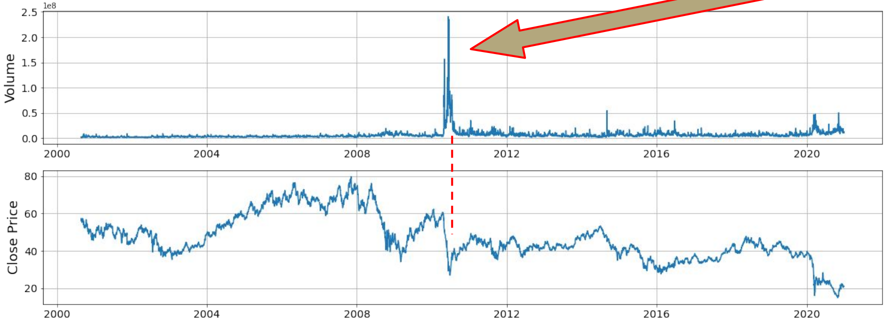
Answer: Kind of, but the correlation factor is very low. By creating an index of close price to oil price, you can see when the 2 values are in balance, or when the index fluctuates and they are out of balance. Some companies have a much correlation with oil price, which might reflect good management.





2. What connection is there between the volumes activity and stock price?

Trendlines of BP



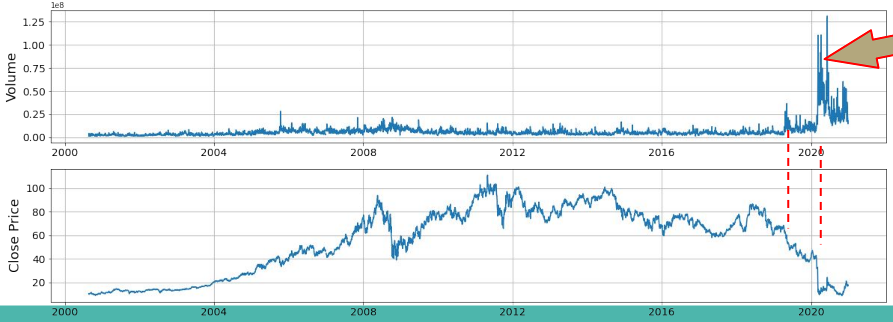
The New York Times
NEW YORK, SUNDAY, DECEMBER 26, 2010

Deepwater Horizon's Final Hours

Missed Signals. Indecision. Failed Defenses. Acts of Valor.

Answer: Greatest changes in volumes traded seems to occur due to big headlines ...

Trendlines of Oxy



Forbes

Oxy Struggles To Cope With The Impacts Of Its Acquisition Of Anadarko

David Blackmon Senior Contributor

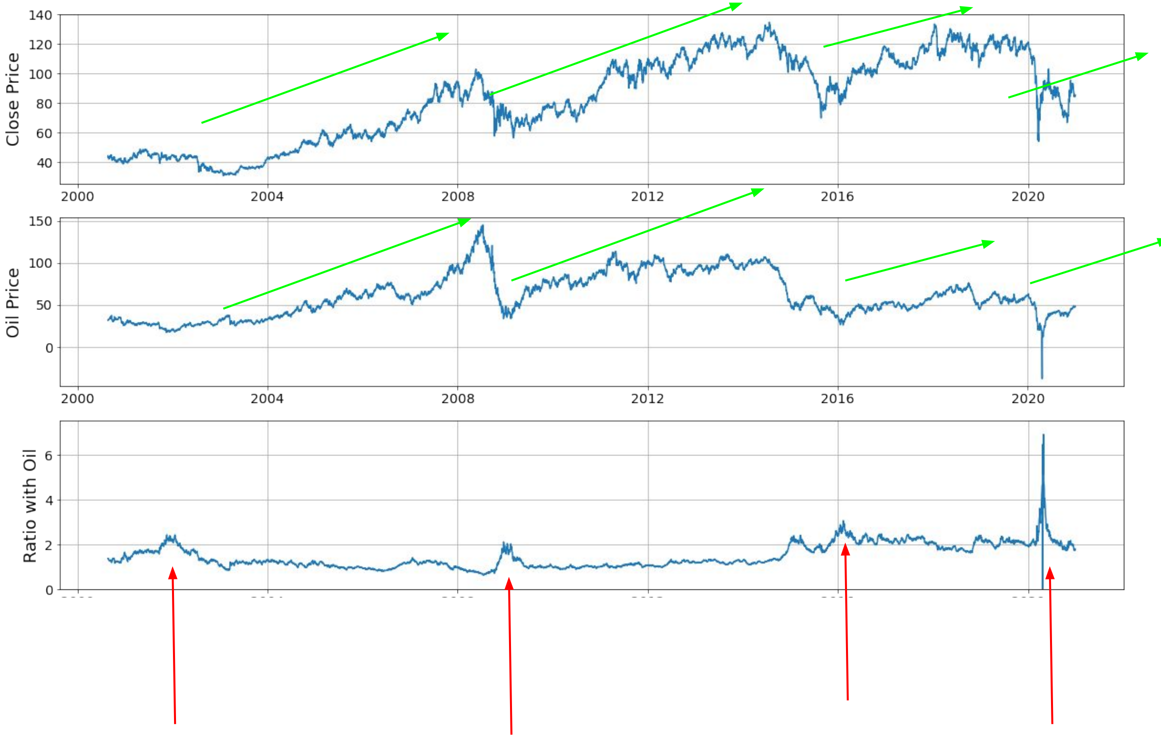
Energy

3. Can a ratio of stock price to oil price provide valuable insight? Perhaps opportunities to buy/sell?

Based on our observations it would difficult at best to look at oil prices alone and know when to sell stocks.

With each spike/dip in both oil prices and stock, there is a spike in the ratio (usually there is an offset between an event and the reaction).

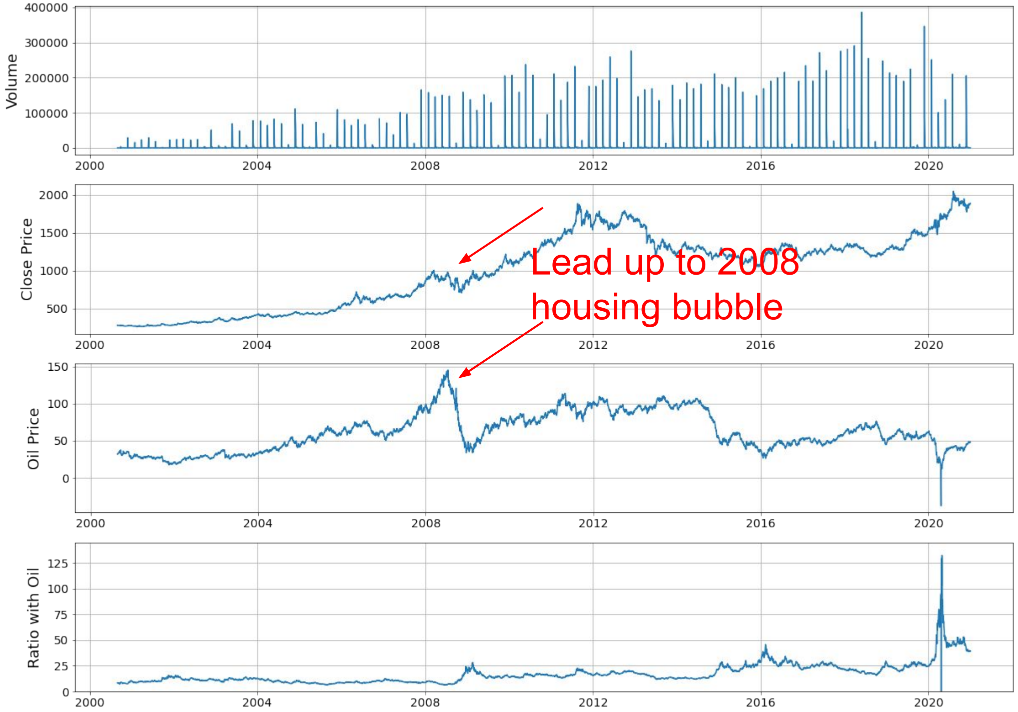
Waiting until the ratio levels out may indicate a better time to buy.



4. How does oil price compare with gold? How does oil do relative to inflation indicators?

- Both oil price and gold increase due to inflation.
- Gold is a good investment to hedge against inflation, while oil isn't necessarily.
- During a recession, the economy contracts and oil and gas markets can slump.
- Oil and Gas markets are also more volatile due to technological advances (fracking) or geopolitical events (Iraq War).

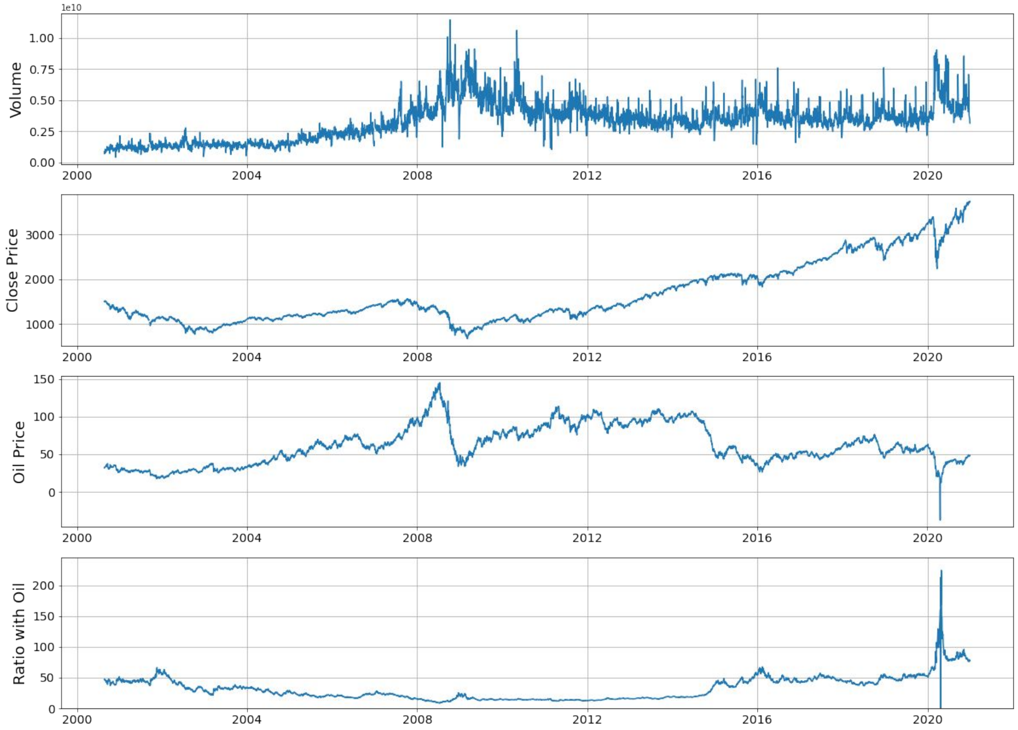
Trendlines of Gold



5. How does oil price or a typical stock ticker compare with the S&P 500?

Oil is less stable, but S&P is of course affected by a lot of the same global events.

Trendlines of S&P 500



Final Thoughts ....

- Is there a max price you would buy **insert stock_*_ that would seem to ensure long term (5-10 year) gains?

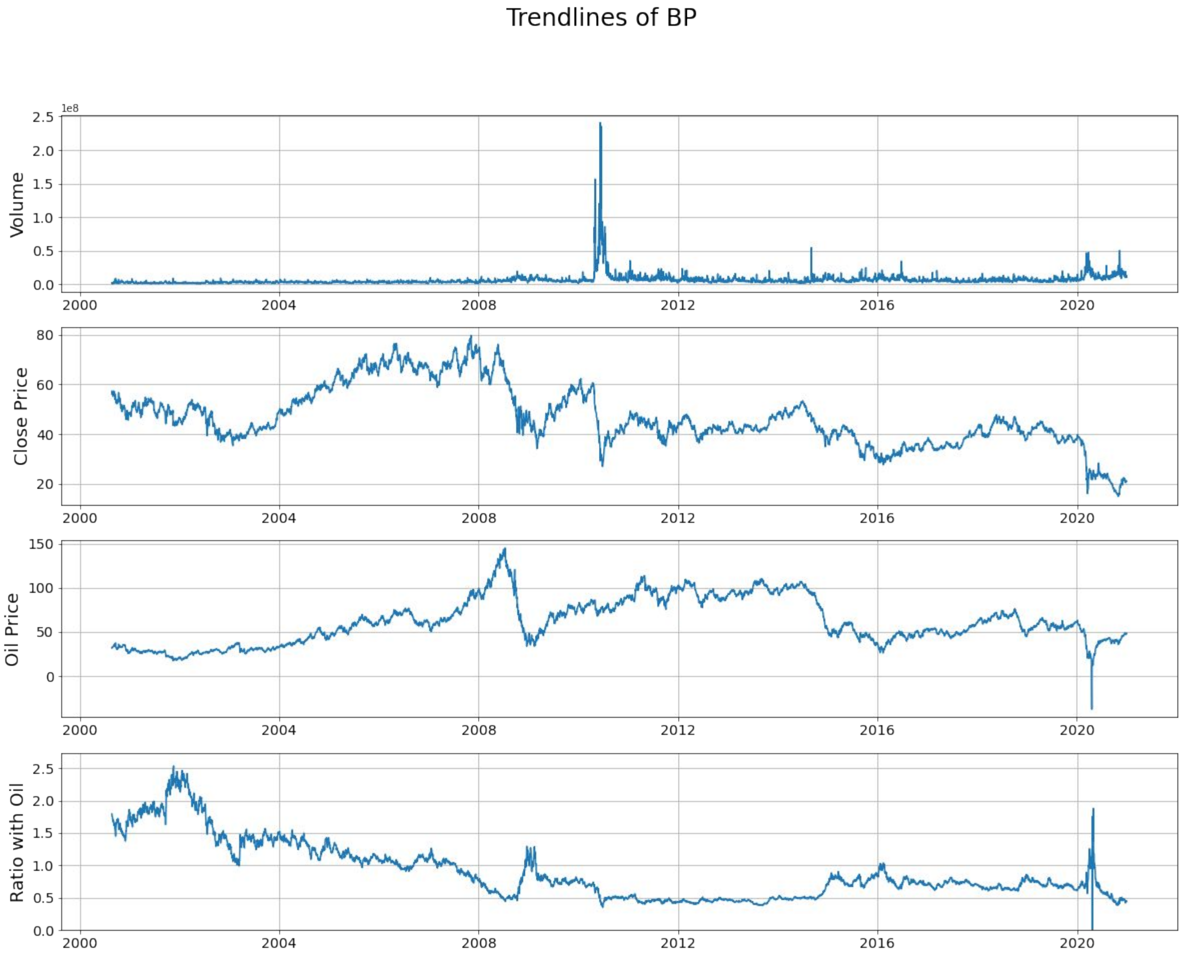
- What world events shaped these graphs?

- Is investing in oil just waiting around for global crises?... (haha, but really)

THANKS!

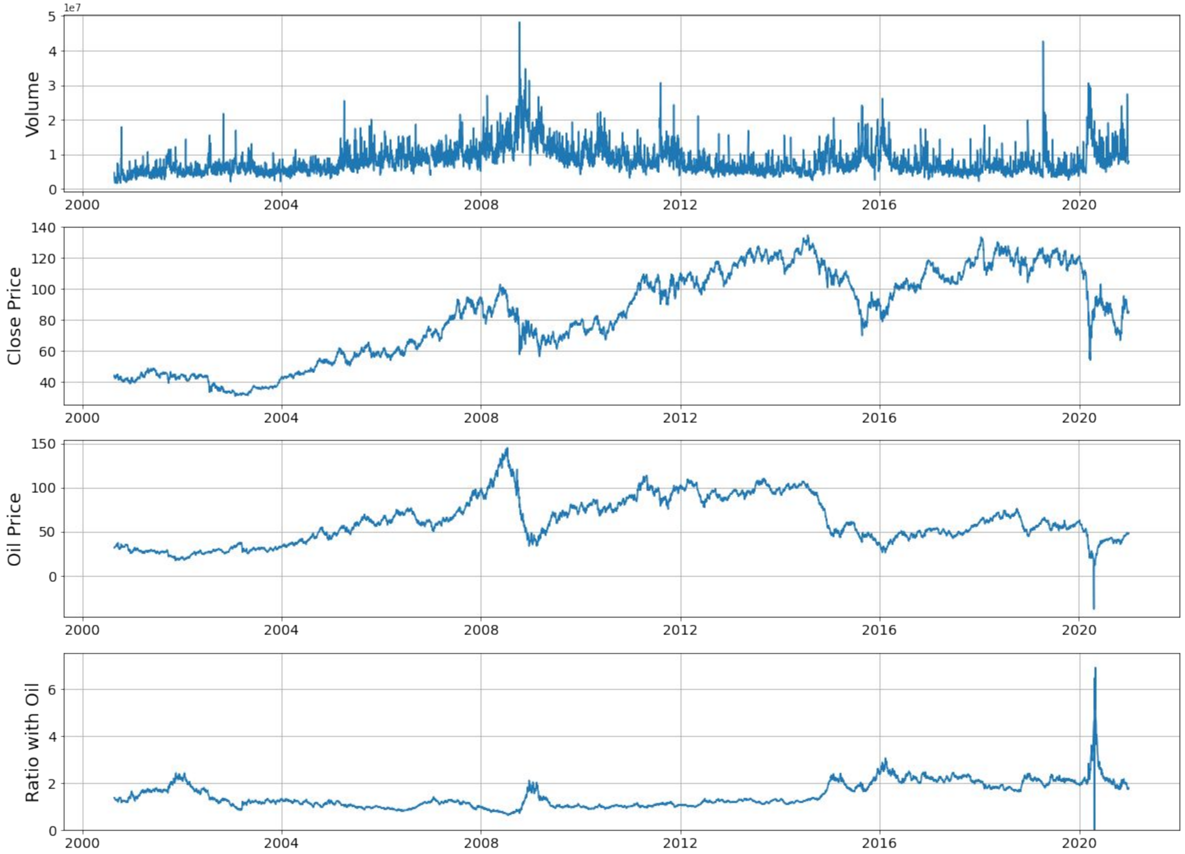
Back-Up Slides

BP

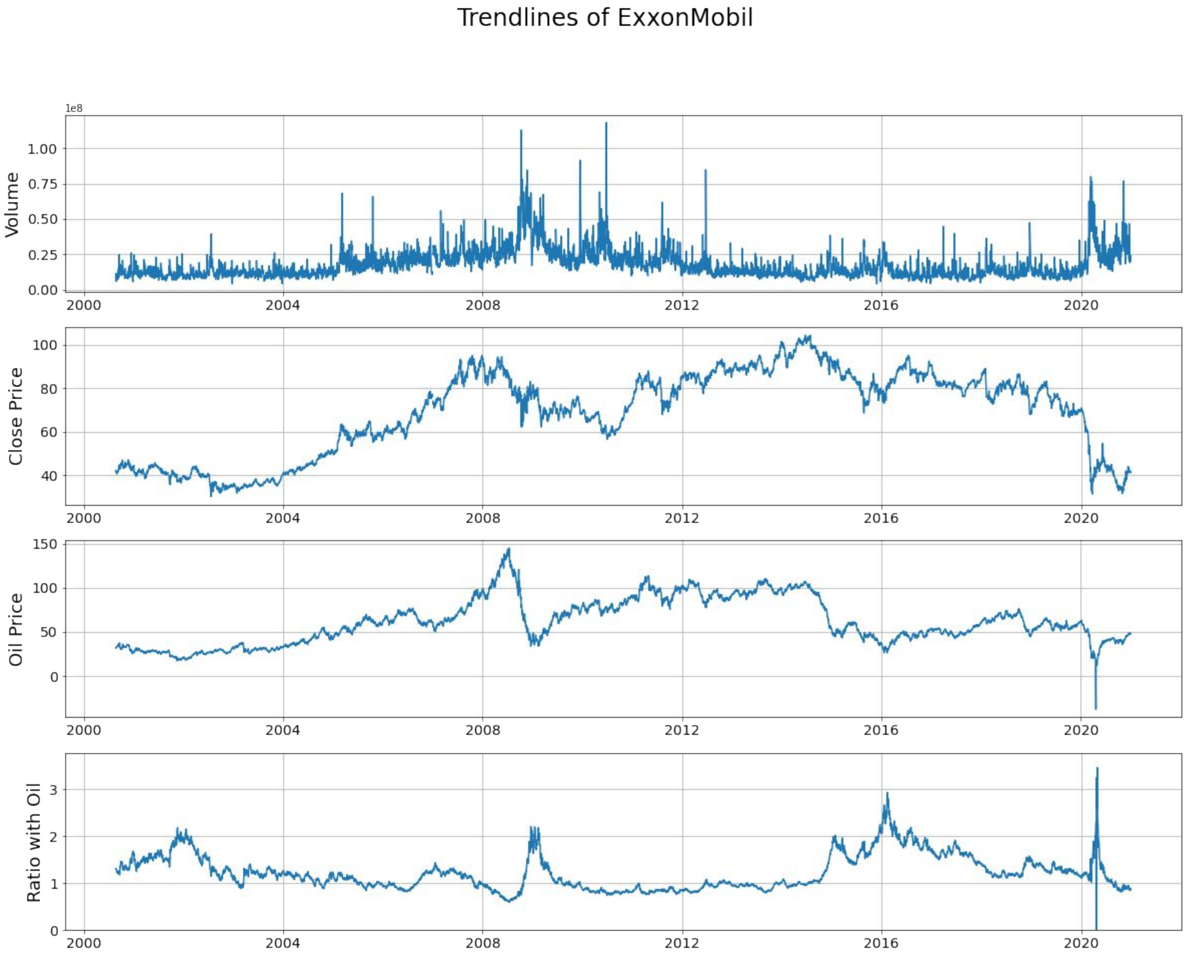

Trendlines of BP
1e8
Volume
0.0
0.5
1.0
1.5
2.0
2.5
Close Price
20
40
60
80
Oil Price
0
50
100
150
Ratio with Oil
0.0
0.5
1.0
1.5
2.0
2.5
2000
2004
2008
2012
2016
2020

Chevron

Trendlines of Chevron

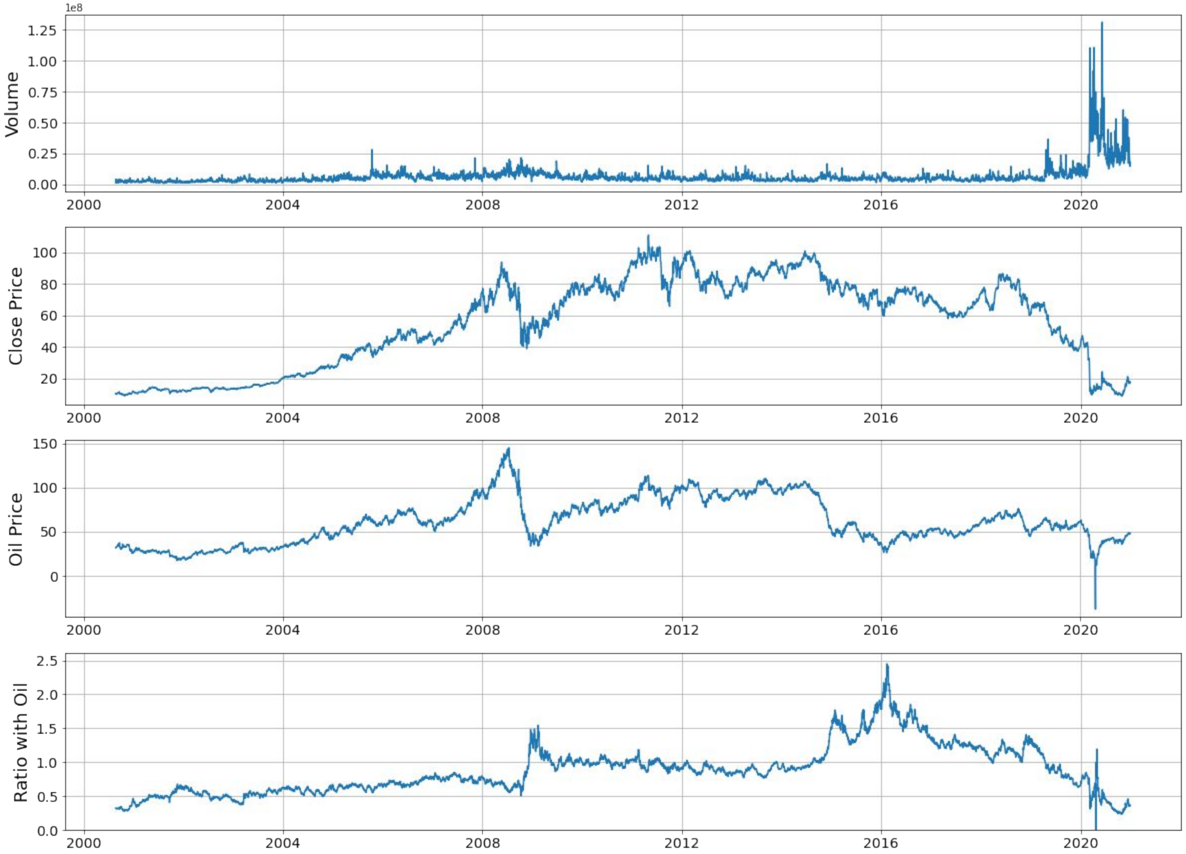


ExxonMobil

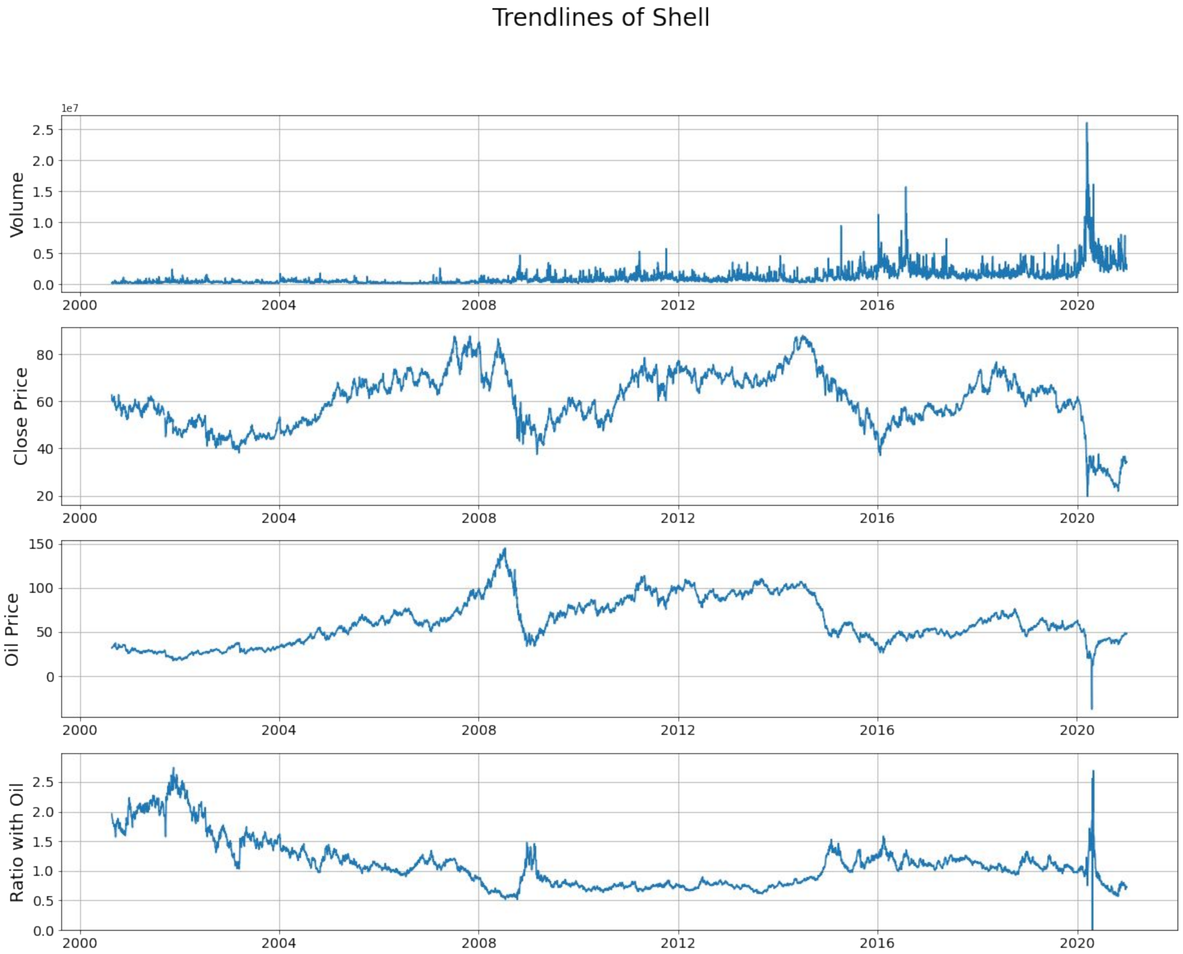

Trendlines of ExxonMobil
1e8
Volume
0.00
0.25
0.50
0.75
1.00
Close Price
40
60
80
100
Oil Price
0
50
100
150
Ratio with Oil
0
1
2
3
2000
2004
2008
2012
2016
2020

Oxy

Trendlines of Oxy

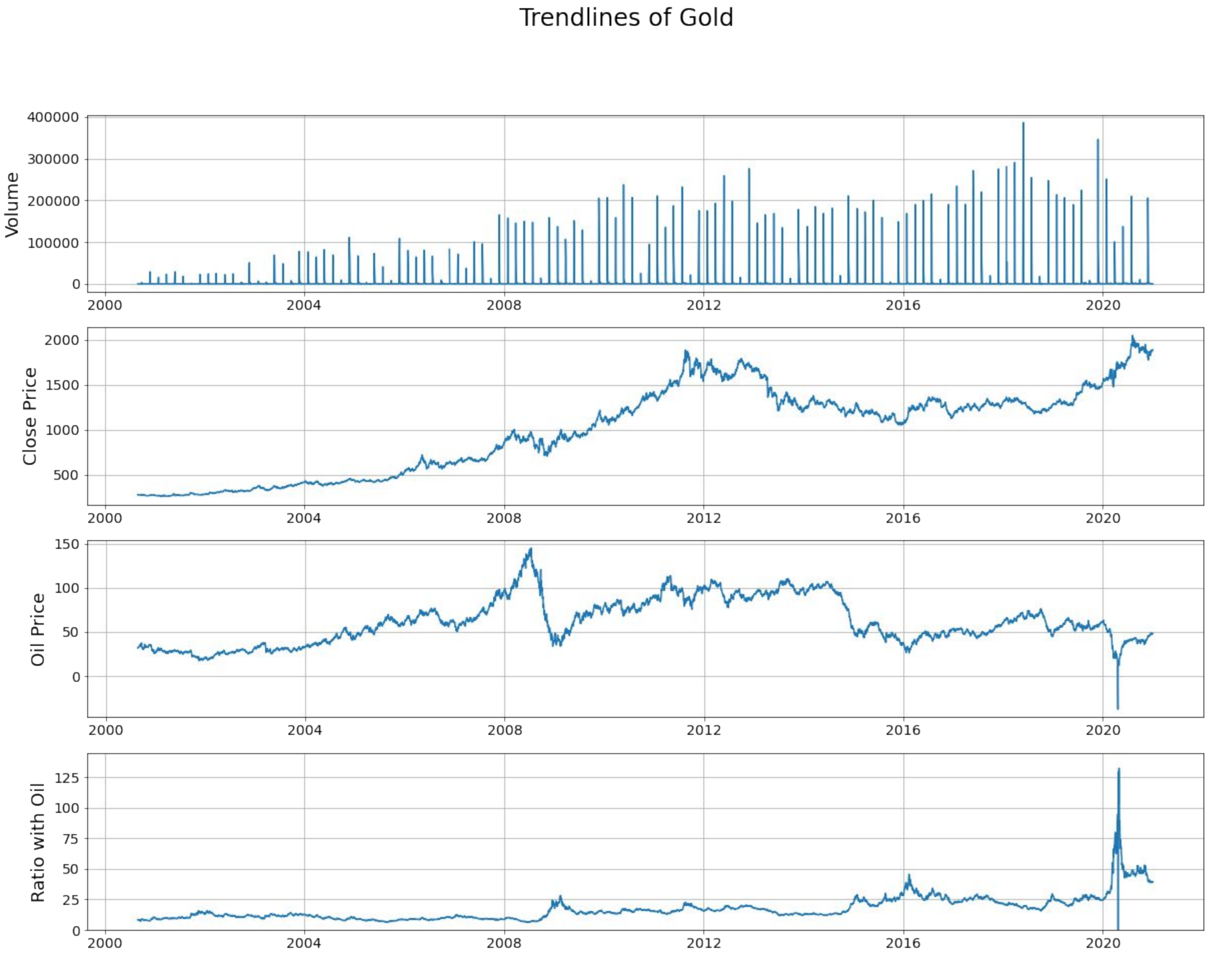


Shell

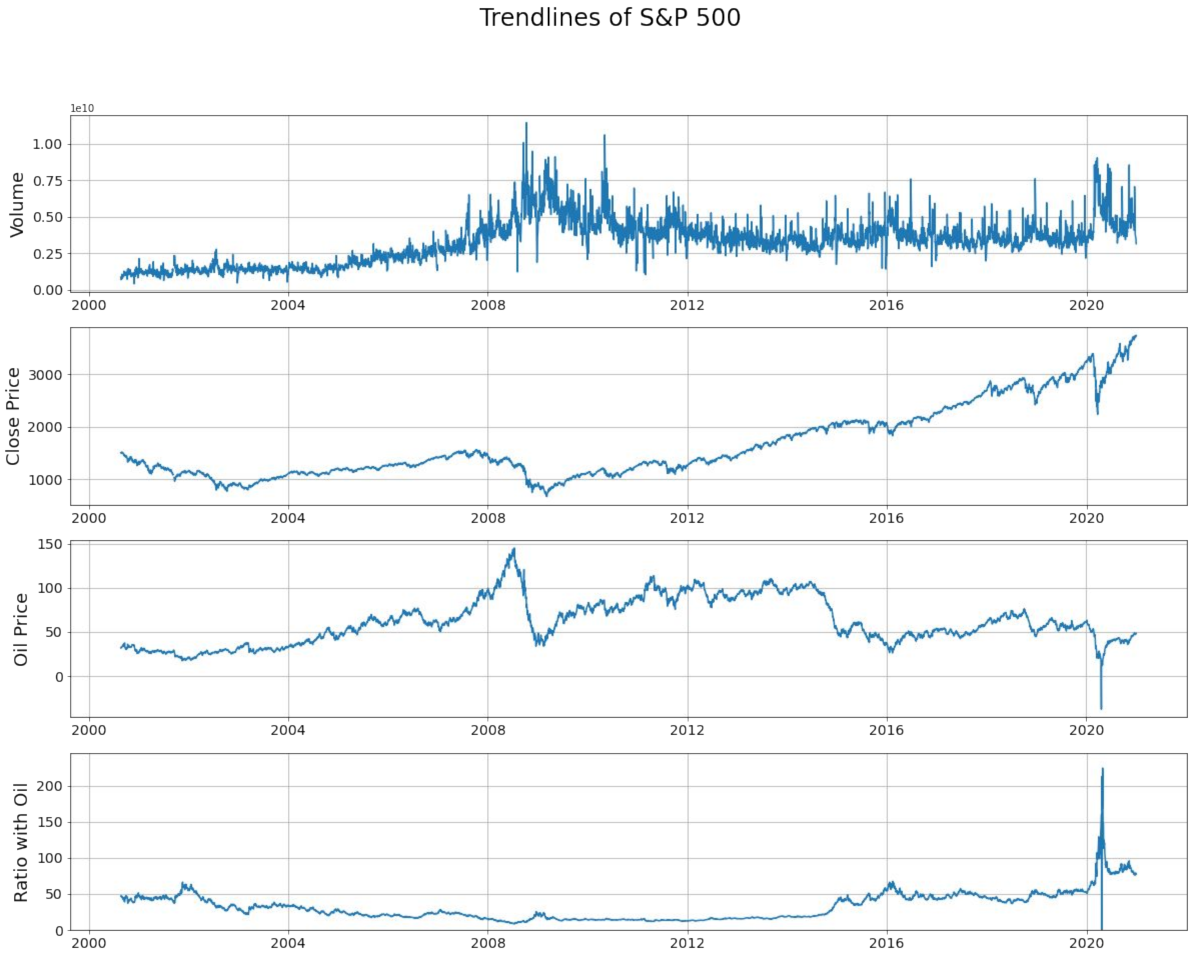


Trendlines of Shell
1e7
Volume
2.5
2.0
1.5
1.0
0.5
0.0
Close Price
80
60
40
20
Oil Price
150
100
50
0
Ratio with Oil
2.5
2.0
1.5
1.0
0.5
0.0
2000
2004
2008
2012
2016
2020

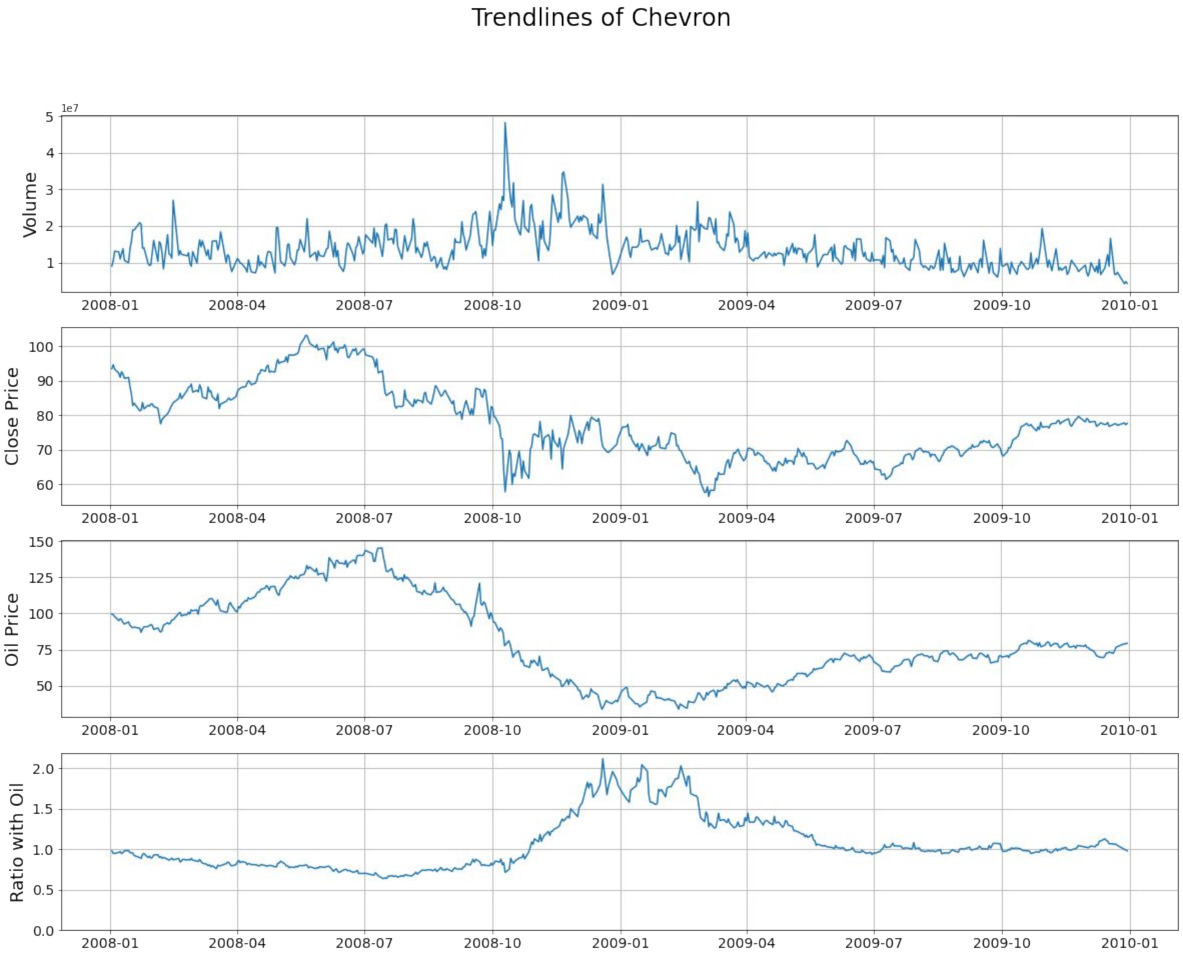
Gold


Trendlines of Gold
Volume
0
100000
200000
300000
400000
Close Price
500
1000
1500
2000
Oil Price
0
50
100
150
Ratio with Oil
0
25
50
75
100
125
2000
2004
2008
2012
2016
2020

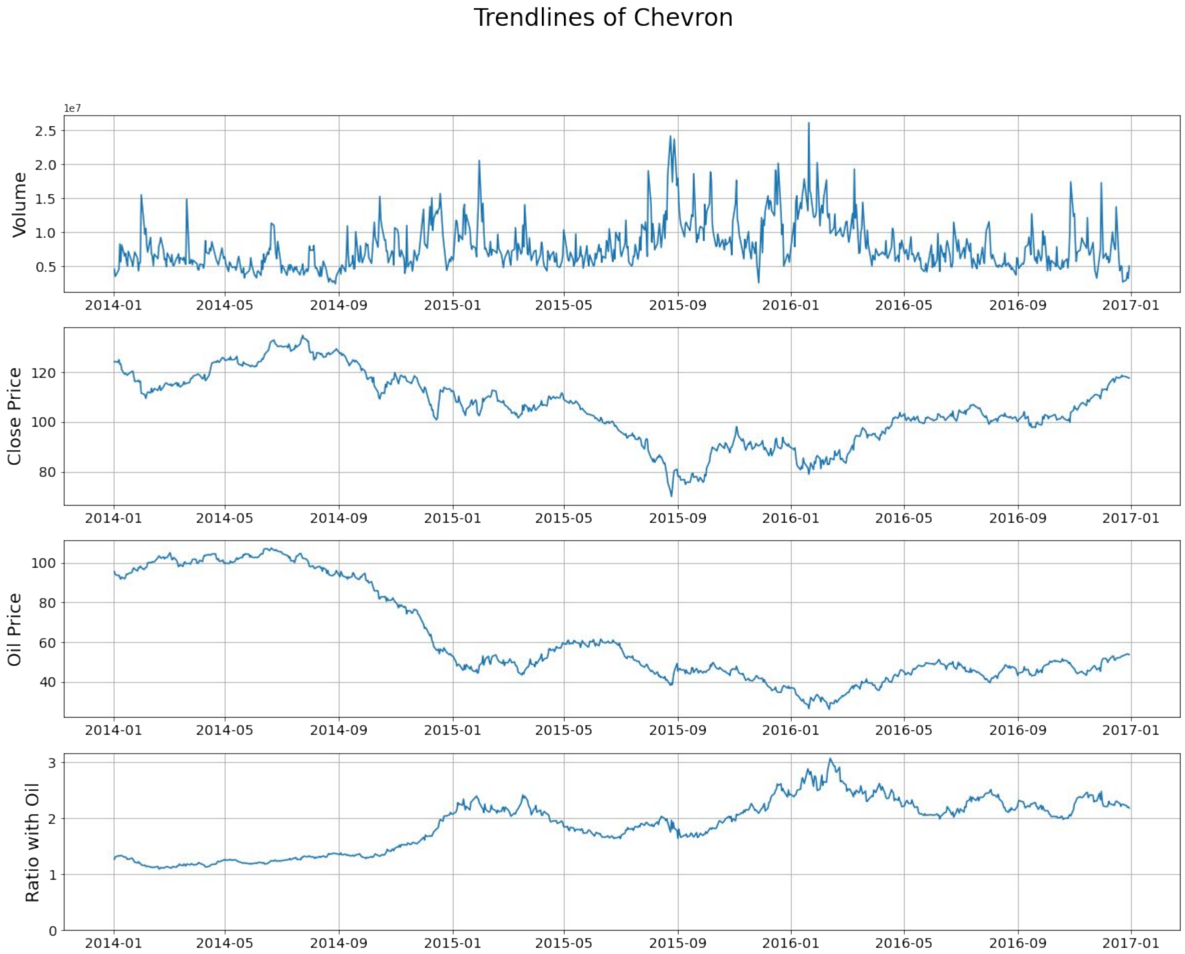
S&P 500


Trendlines of S&P 500
1e10
Volume
1.00
0.75
0.50
0.25
0.00
Close Price
3000
2000
1000
Oil Price
150
100
50
0
Ratio with Oil
200
150
100
50
0
2000
2004
2008
2012
2016
2020

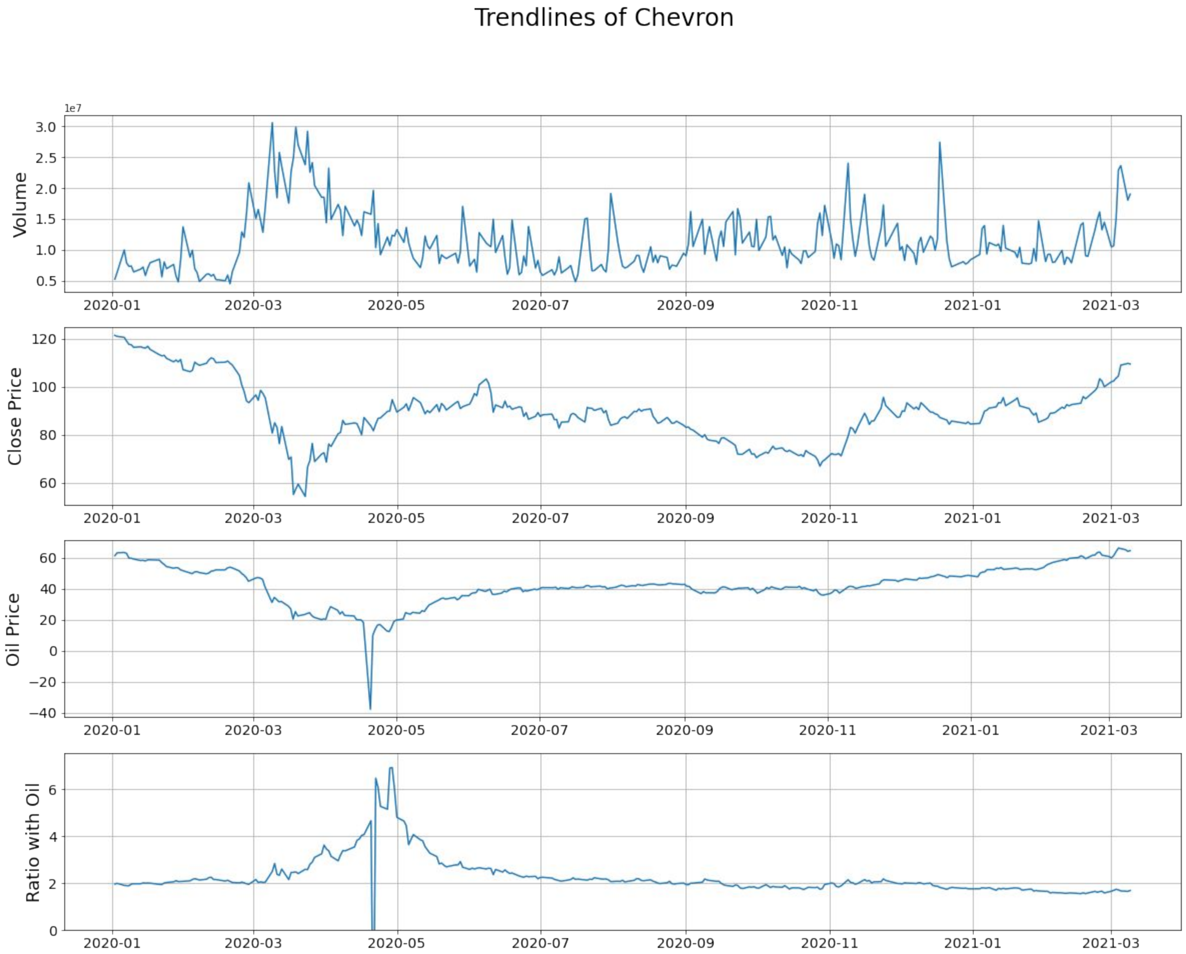
2008 Housing Bubble (Chevron)



2014 Oil Crisis (Chevron)



2020 COVID and Saudi/Russia Price War (Chevron)


Trendlines of Chevron
1e7
Volume
Close Price
Oil Price
Ratio with Oil