Vaccine Stock Analysis during the COVID-19 Pandemic

May 14, 2022

Vaccine Stock Analysis during the COVID-19 Pandemic

[70]:
```
import numpy as np
import pandas as pd

from functools import reduce
```

Covid Dataset

[71]:
```
df = pd.read_csv("us_covid19_cases.csv")
```

[72]:
```
df.columns
```

[72]:
```
Index(['iso_code', 'continent', 'location', 'date', 'total_cases', 'new_cases',
       'new_cases_smoothed', 'total_deaths', 'new_deaths',
       'new_deaths_smoothed', 'total_cases_per_million',
       'new_cases_per_million', 'new_cases_smoothed_per_million',
       'total_deaths_per_million', 'new_deaths_per_million',
       'new_deaths_smoothed_per_million', 'reproduction_rate', 'icu_patients',
       'icu_patients_per_million', 'hosp_patients',
       'hosp_patients_per_million', 'weekly_icu_admissions',
       'weekly_icu_admissions_per_million', 'weekly_hosp_admissions',
       'weekly_hosp_admissions_per_million', 'total_tests', 'new_tests',
       'total_tests_per_thousand', 'new_tests_per_thousand',
       'new_tests_smoothed', 'new_tests_smoothed_per_thousand',
       'positive_rate', 'tests_per_case', 'tests_units', 'total_vaccinations',
       'people_vaccinated', 'people_fully_vaccinated', 'total_boosters',
       'new_vaccinations', 'new_vaccinations_smoothed',
       'total_vaccinations_per_hundred', 'people_vaccinated_per_hundred',
       'people_fully_vaccinated_per_hundred', 'total_boosters_per_hundred',
       'new_vaccinations_smoothed_per_million',
       'new_people_vaccinated_smoothed',
       'new_people_vaccinated_smoothed_per_hundred', 'stringency_index',
       'population', 'population_density', 'median_age', 'aged_65_older',
       'aged_70_older', 'gdp_per_capita', 'extreme_poverty',
       'cardiovasc_death_rate', 'diabetes_prevalence', 'female_smokers',
       'male_smokers', 'handwashing_facilities', 'hospital_beds_per_thousand',
       'life_expectancy', 'human_development_index',
       'excess_mortality_cumulative_absolute', 'excess_mortality_cumulative',
       'excess_mortality', 'excess_mortality_cumulative_per_million'],
```

```
      dtype='object')
```

[73]:
```
covid_df = df[['date', 'new_cases', 'new_deaths', 'new_tests',␣
 ↪'new_vaccinations']]
```

[74]:
```
covid_df.head()
```

[74]:

	date	new_cases	new_deaths	new_tests	new_vaccinations
0	1/22/2020	NaN	NaN	NaN	NaN
1	1/23/2020	0.0	NaN	NaN	NaN
2	1/24/2020	1.0	NaN	NaN	NaN
3	1/25/2020	0.0	NaN	NaN	NaN
4	1/26/2020	3.0	NaN	NaN	NaN

[59]:
```
covid_df.dtypes
```

[59]:
```
date                 object
new_cases           float64
new_deaths          float64
new_tests           float64
new_vaccinations    float64
dtype: object
```

Stock Datasts

[56]:
```
df2 = pd.read_csv("moderna.csv")
moderna_df = df2[['date', 'close', 'volume', 'rsi']]

df3 = pd.read_csv("jnj.csv")
jnj_df = df3[['date', 'close', 'volume', 'rsi']]

df4 = pd.read_csv("astra-zeneca.csv")
astra_zeneca_df = df4[['date', 'close', 'volume', 'rsi']]

df5 = pd.read_csv("biontech.csv")
biontech_df = df5[['date', 'close', 'volume', 'rsi']]

df6 = pd.read_csv("novavax.csv")
novavax_df = df6[['date', 'close', 'volume', 'rsi']]

df7 = pd.read_csv("pfizer.csv")
pfizer_df = df7[['date', 'close', 'volume', 'rsi']]
```

Merge Datasets

[57]:
```
stock_df = [moderna_df, jnj_df, astra_zeneca_df, biontech_df, novavax_df,␣
 ↪pfizer_df]

stock_merged = reduce(lambda left, right: pd.merge(left, right, on=['date'],␣
 ↪how='outer'), stock_df)
```

```
stock_merged.columns = ['date', 'moderna_closing_price', 'moderna_volume',␣
 ↪'moderna_rsi', 'jnj_closing_price', 'jnj_volume', 'jnj_rsi',␣
 ↪'astra_zeneca_closing_price', 'astra_zeneca_volume', 'astra_zeneca_rsi',␣
 ↪'biontech_closing_price', 'biontech_volume', 'biontech_rsi',␣
 ↪'novavax_closing_price', 'novavax_volume', 'novavax_rsi',␣
 ↪'pfizer_closing_price', 'pfizer_volume', 'pfizer_rsi']
```

[58]:
```
stock_merged.head()
```

[58]:

	date	moderna_closing_price	moderna_volume	moderna_rsi	jnj_closing_price	jnj_volume	jnj_rsi	astra_zeneca_closing_price	astra_zeneca_volume	astra_zeneca_rsi	biontech_closing_price	biontech_volume	biontech_rsi	novavax_closing_price	novavax_volume	novavax_rsi	pfizer_closing_price	pfizer_volume	pfizer_rsi
0	2022-04-08	160.84	5454415.0	45.310068	182.12	7144703.0	58.007217	71.14	9082865.0	62.388493	170.26	886206.0	47.024802	60.63	3603222.0	41.168775	55.17	23128622.0	55.231783
1	2022-04-07	159.00	5720873.0	45.059303	181.76	7385291.0	57.800809	71.01	8310168.0	62.271561	169.11	1505572.0	46.875633	59.50	4367650.0	40.924147	55.16	36292543.0	55.220421
2	2022-04-06	154.62	7401800.0	44.465273	182.23	9991790.0	58.166612	69.07	5596805.0	60.474451	166.65	2275843.0	46.560180	62.44	5162016.0	41.362681	52.87	31718155.0	52.515858
3	2022-04-05	162.05	6641095.0	45.279139	177.61	7279617.0	55.450641	67.05	6953316.0	58.455160	180.82	1878969.0	48.174997	65.23	6661778.0	41.779037	51.24	21027857.0	50.427343
4	2022-04-04	172.54	5908675.0	46.455614	176.47	6595724.0	54.740074	66.67	3987163.0	58.060152	186.24	2350714.0	48.809551	75.29	2642969.0	43.319900	50.94	20491602.0	50.030928

[49]:
```
stock_merged.dtypes
```

[49]:
```
date                        object
moderna_closing_price      float64
moderna_volume             float64
```

```
moderna_rsi                    float64
jnj_closing_price              float64
jnj_volume                     float64
jnj_rsi                        float64
astra_zeneca_closing_price     float64
astra_zeneca_volume            float64
astra_zeneca_rsi               float64
biontech_closing_price         float64
biontech_volume                float64
biontech_rsi                   float64
novavax_closing_price          float64
novavax_volume                 float64
novavax_rsi                    float64
pfizer_closing_price           float64
pfizer_volume                  float64
pfizer_rsi                     float64
dtype: object
```

[67]:
```
covid_df['date'] = pd.to_datetime(covid_df['date'])
stock_merged['date'] = pd.to_datetime(stock_merged['date'])
covid_df['new_vaccinations'] = covid_df['new_vaccinations'].fillna(0) # The NaN␣
 ↪value in this column cannot be simply dropped since we also need to analyze␣
 ↪the data before the vaccination process.
df_merged = pd.merge(covid_df, stock_merged, on='date').dropna()
```

[69]:
```
df_merged.head(10)
```

[69]:

	date	new_cases	new_deaths	new_tests	new_vaccinations	\
27	2020-03-02	23.0	5.0	515.0	0.0	
28	2020-03-03	19.0	1.0	620.0	0.0	
29	2020-03-04	33.0	4.0	891.0	0.0	
30	2020-03-05	77.0	1.0	1203.0	0.0	
31	2020-03-06	53.0	2.0	1523.0	0.0	
32	2020-03-09	75.0	1.0	2399.0	0.0	
33	2020-03-10	188.0	6.0	3481.0	0.0	
34	2020-03-11	365.0	5.0	4833.0	0.0	
35	2020-03-12	439.0	10.0	8891.0	0.0	
36	2020-03-13	633.0	8.0	11732.0	0.0	

	moderna_closing_price	moderna_volume	moderna_rsi	jnj_closing_price	\
27	29.88	33084026.0	64.837624	140.020004	
28	27.91	17599114.0	61.075243	135.589996	
29	27.49	11817666.0	60.313849	143.479996	
30	28.01	14669976.0	60.929202	142.009995	
31	29.61	21097488.0	62.742982	142.029999	
32	24.29	14124076.0	54.205308	136.440002	
33	22.34	13019516.0	51.580159	141.639999	
34	23.61	14825434.0	53.089965	131.800003	

```
35                 22.30      11305347.0    51.402919         125.410004
36                 21.30      11853052.0    50.161307         134.289993

    jnj_volume  ...  astra_zeneca_rsi  biontech_closing_price  \
27  11508200.0  ...         40.456508                   36.60
28  13662500.0  ...         39.738102                   38.48
29  10560500.0  ...         48.025686                   39.19
30  11339200.0  ...         47.549072                   37.12
31  12239100.0  ...         45.738435                   38.09
32  13848600.0  ...         41.449773                   33.48
33  12698100.0  ...         45.268621                   33.96
34  17763400.0  ...         41.437368                   32.17
35  21539200.0  ...         36.278622                   28.55
36  20084200.0  ...         39.344643                   30.93

    biontech_volume  biontech_rsi  novavax_closing_price  novavax_volume  \
27         185100.0     50.791524                  12.02      14261700.0
28         297900.0     52.850212                  10.78       9957300.0
29         127100.0     53.598311                  11.32      10977400.0
30          77000.0     51.182375                  12.87      14057100.0
31         279200.0     52.212347                  12.48      13250800.0
32         190200.0     47.365959                  10.02       8328800.0
33         327500.0     47.879960                  10.65      12365400.0
34         119000.0     46.164460                  10.51       9052100.0
35         184100.0     42.986160                   9.29       5488400.0
36         197600.0     45.503257                   8.41       6115400.0

    novavax_rsi  pfizer_closing_price  pfizer_volume  pfizer_rsi
27    62.899608             33.092979     42034469.0   37.171087
28    59.621290             32.542694     46174475.0   35.568345
29    60.535314             34.535103     38712155.0   44.422057
30    62.989154             33.643265     35096303.0   41.798627
31    61.999390             33.225807     40931036.0   40.651919
32    56.304878             32.030361     43183856.0   37.634936
33    57.329068             32.817837     40548329.0   40.598216
34    57.025971             30.521822     65350213.0   35.569873
35    54.465420             28.481974     62731445.0   31.979150
36    52.722981             31.034157     60553038.0   39.744867

[10 rows x 23 columns]
```

[75]:
```
df_merged.dtypes
```

[75]:
```
date                          datetime64[ns]
new_cases                            float64
new_deaths                           float64
new_tests                            float64
new_vaccinations                     float64
```

```
moderna_closing_price              float64
moderna_volume                     float64
moderna_rsi                        float64
jnj_closing_price                  float64
jnj_volume                         float64
jnj_rsi                            float64
astra_zeneca_closing_price         float64
astra_zeneca_volume                float64
astra_zeneca_rsi                   float64
biontech_closing_price             float64
biontech_volume                    float64
biontech_rsi                       float64
novavax_closing_price              float64
novavax_volume                     float64
novavax_rsi                        float64
pfizer_closing_price               float64
pfizer_volume                      float64
pfizer_rsi                         float64
dtype: object
```

Data Analysis

Q1: How did each pharmaceutical stock perform during the pandemic, and is there a trend between performance and COVID-19 cases?

[ ]:

Q2: Is there a positive or negative relationship between COVID-19 cases and pharmaceutical stock prices of the aforementioned companies? Could a rise in COVID-19 cases be used as a factor to predict a rise in pharmaceutical stock prices?

[ ]:

Q3: Assuming there is a pattern/relationship, what are the nuances that explain any breaks from the pattern between COVID-19 cases and stock prices?

[ ]: