

Magdalena Sudół, Magdalena Ryś

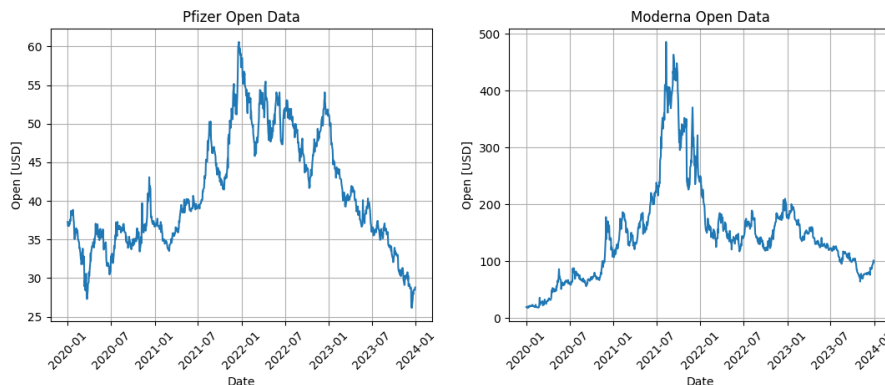
# Analiza cen akcji producentów szczepionek przeciwko COVID-19

kwiecień 2024

## 1. Cel raportu

Celem poniższego raportu jest przeprowadzenie analizy cen akcji producentów szczepionek przeciwko COVID-19 na rynku finansowym. Dane, na których opiera się raport, zostały pobrane ze strony [www.finance.yahoo.com](http://www.finance.yahoo.com) i dotyczą dwóch wiodących koncernów farmaceutycznych produkujących szczepionki przeciwko koronawirusowi - Pfizer oraz Moderna. Zebrane dane obejmują notowania akcji tych firm odnotowywane codziennie w okresie od 1 stycznia 2020 roku (daty zbliżonej do początku pandemii) do 31 grudnia 2023 roku.

Nasze obliczenia oparliśmy na danych na temat kursów otwarcia, czyli cen akcji, po jakich inwestorzy realizują transakcje zaraz po otwarciu się sesji notowań. Poniżej zamieszczono wyżej wymienione dane w formie wykresów:



Rysunek 1. Wykresy danych dot. cen akcji firm farmaceutycznych

Już na powyższych wykresach możemy zaobserwować nagły wzrost cen akcji w połowie badanego okresu, czyli w momencie masowych szczepień przeciwko COVID-19. Przeprowadzona przez nas analiza powyższych danych ma na celu zidentyfikowanie ewentualnych wzajemnych zależności pomiędzy cenami akcji różnych producentów szczepionek. Odkrycie takich korelacji może przyczynić się do lepszego zrozumienia dynamiki rynku farmaceutycznego w kontekście pandemii.

## 2. Analiza danych

W tej części zajmiemy się miarami położenia (średnia arytmetyczna, mediana, kwantyle) oraz rozproszenia danych (rozstęp międzykwartylowy), a także wariancją i odchyleniem standardowym. Obliczone wartości wykorzystamy do wyciągnięcia wniosków na temat notowań akcji firm farmaceutycznych w okresie pandemii oraz po jej zakończeniu. Wyniki zaokrąglamy do 2 miejsc po przecinku.

## 2.1. Średnia arytmetyczna

Analizę danych rozpoczniemy od obliczenia średnich arytmetycznych kursów otwarcia, czyli ich przeciętnych poziomów. Aby obliczyć średnią arytmetyczną, wykorzystamy wzór:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$$

gdzie  $n$  jest liczbą obserwacji, a  $x_i$   $i$ -tą obserwacją.

Wstawiając do powyższego wzoru pobrane dane, otrzymujemy wynik 40,96 dla firmy Pfizer oraz 147,19 dla firmy Moderna.

## 2.2. Mediana

Mediana, czyli wartość środkowa, dzieli uporządkowany zbiór danych na dwie części w taki sposób, że powyżej i poniżej tej wartości znajduje się jednakowa liczba danych. Inaczej można ją określić drugim kwantylem – 50% obserwacji jest położonych poniżej tej wartości. Medianę obliczamy według wzoru:

$$x_{med} = \begin{cases} x_{((n+1)/2)} & \text{gdy } n \text{ jest nieparzyste} \\ \frac{1}{2} \cdot (x_{(n/2)} + x_{(n/2+1)}) & \text{gdy } n \text{ jest parzyste} \end{cases}$$

gdzie  $n$  jest liczbą obserwacji, a  $x_{(i)}$  elementem na  $i$ -tej pozycji w uporządkowanym zbiorze danych.

Mediana obserwacji dotyczących kursów otwarcia firmy Pfizer wynosi 39,15, natomiast firmy Moderna – 137,68.

## 2.3. Kwantyle

Kwantyle dzielą zbiór danych na 4 części o jednakowej liczbie obserwacji. Pierwszy kwantyl ( $Q_1$ ) to wartość, poniżej której znajduje się 25% obserwacji, drugi kwantyl (mediana;  $Q_2$ ) to wartość, poniżej której znajduje się 50% obserwacji, a trzeci kwantyl – wartość, poniżej której znajduje się 75% obserwacji. Kwantyle obliczamy, znajdując medianę i dzieląc w ten sposób zbiór danych na dwie części – obserwacje poniżej i powyżej mediany. Następnie obliczamy mediany obydwu zbiorów. Mediana pierwszego ze zbiorów to pierwszy kwantyl, a drugiego ze zbiorów – trzeci kwantyl.

Kwantyle  $Q_1$  i  $Q_3$  dla danych firmy Pfizer wynoszą odpowiednio 35,49 oraz 47,41, a dla danych firmy Moderna – 82,05 oraz 173,51.

## 2.4. Różnica międzykwantylowa

Różnica międzykwantylowa (IQR) jest różnicą pomiędzy trzecim oraz pierwszym kwantylem. Pomiedzy tymi kwantylami znajdziemy 50% obserwacji.

Różnica międzykwantylowa kursów otwarcia firmy Pfizer wynosi 11,92, natomiast firmy Moderna – 91,46.

## 2.5. Wariancja

Wariancja informuje o zróżnicowaniu obserwacji – o tym, w jakim stopniu są skoncentrowane wokół wartości oczekiwanej. Im większa wariancja, tym bardziej zróżnicowane wyniki. Wariancję obliczamy ze wzoru:

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

gdzie  $n$  – liczba obserwacji,  $\bar{x}$  – średnia arytmetyczna obserwacji,  $x_i$  –  $i$ -ta obserwacja.

Wariancja dla kursów otwarcia firmy Pfizer wynosi 53,92, natomiast dla firmy Moderna – 7585,40.

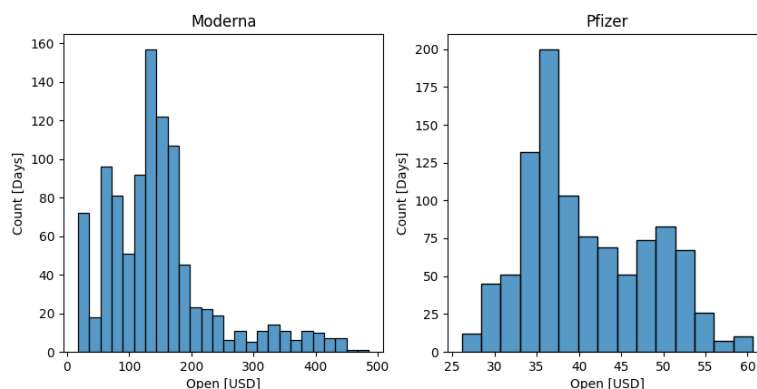
## 2.6. Odchylenie standardowe

Odchylenie standardowe jest pierwiastkiem z wariancji i również informuje o skoncentrowaniu obserwacji w stosunku do wartości oczekiwanej.

Dla danych firmy Pfizer odchylenie standardowe jest równe 7,34, a dla firmy Moderna – 87,09.

## 2.7. Histogramy

Histogramy są wykresami, które przedstawiają, jak często pojawiają się dane wartości w zbiorze danych. Na osi x umieszczone są wartości, w naszym przypadku kursów otwarcia, podzielone na przedziały, natomiast na osi y – jak często pojawiały się wartości należące do każdego z przedziałów, w naszym przypadku – przez ile dni je odnotowywano.



Rysunek 2. Histogramy

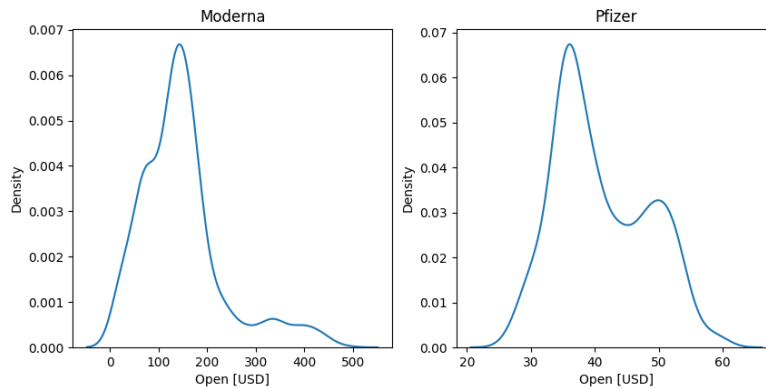
Po lewej stronie przedstawiono histogram dla cen akcji firmy Moderna. Akcje przyjmują wartości od około 18 USD do 485 USD. Najczęściej przyjmowane są wartości z przedziału 120 – 140 USD - wartości takie odnotowywano przez około 160 dni. Najrzadziej natomiast przyjmowane są wartości powyżej 460 USD – zostały odnotowane jedynie dwa razy.

Po prawej stronie przedstawiono histogram dla cen akcji firmy Pfizer. Akcje przyjmują wartości od około 26 USD do 60 USD. Najczęściej przyjmowane są wartości z przedziału 35 – 37, 50 USD - wartości takie odnotowywano przez około 200 dni. Najrzadziej natomiast przyjmowane są wartości z przedziału 56 – 58 USD – zostały odnotowane jedynie kilka razy.

## 2.8. Wykresy gęstości

Funkcje gęstości prawdopodobieństwa są funkcjami, na podstawie których można obliczyć prawdopodobieństwo zajścia danego zdarzenia losowego (całka z funkcji gęstości, obliczona na odpowiednim zbiorze  $A$ , jest równa prawdopodobieństwu  $P(A)$ ). Funkcje takie są nieujemne, a wartość całki z funkcji gęstości po całej przestrzeni  $R$  (przypadek jednowymiarowy) jest równa 1.

Poniżej przedstawiono wykresy gęstości dla cen akcji firmy Moderna oraz Pfizer. Wnioski, jakie możemy wyciągnąć, patrząc na wykresy, pokrywają się z wnioskami z analizy histogramów. W przypadku firmy Moderna, odnotowujemy największe prawdopodobieństwo zajścia takiego zdarzenia, że kurs otwarcia znajdzie się pomiędzy 130 a 150 USD. W przypadku firmy Pfizer, najbardziej prawdopodobne jest przyjęcie przez kurs otwarcia wartości od około 36 do 38 USD.

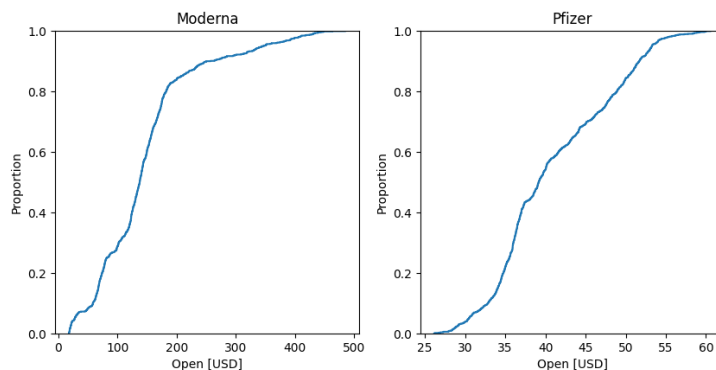


Rysunek 3. Wykresy gęstości

## 2.9. Wykresy dystrybuant

Dystrybuanty są funkcjami, które wskazują, jakie jest prawdopodobieństwo, że została przyjęta wartość mniejsza od  $x$  – naszym przypadku, jakie jest prawdopodobieństwo, że ceny akcji osiągnęły wartość  $x$  lub mniejszą. Funkcja dystrybuanty przyjmuje wartości pomiędzy 0 i 1 (wynika to z faktu, że prawdopodobieństwo zajścia jakiegokolwiek zdarzenia należy do przedziału  $[0, 1]$ , gdzie 0 to wartość dla zdarzenia niemożliwego, natomiast 1 – dla zdarzenia pewnego), jest niemalejąca i prawostronnie ciągła, a jej granice są równe 0 w  $-\infty$  oraz 1 w  $\infty$ .

Przedstawione poniżej wykresy to dystrybuanty kursów otwarcia firmy Moderna (po lewej) oraz Pfizer (po prawej). Można z nich odczytać, że prawdopodobieństwo, że kurs otwarcia firmy Moderna przyjmuje wartość mniejszą niż 18 USD jest równe 0. Pomędzy wartościami 50 a 90 oraz 100 a 200 USD następuje najszybszy wzrost wartości dystrybuanty. Są to wartości najczęściej przyjmowane, jednocześnie zróżnicowane – niemalże nie zdarza się, żeby kursy otwarcia przyjęły dokładnie takie same wartości z dokładnością do kilku miejsc po przecinku. Dla każdej wartości  $x$  powyżej 485 dystrybuanta przyjmuje wartość 1, co oznacza, że jest zdarzeniem pewnym, że ceny akcji przyjęły wartość mniejszą niż  $x$ . W przypadku firmy Pfizer prawdopodobieństwo, że kurs otwarcia przyjmuje wartość mniejszą niż 26 USD, jest równe 0. Najszybszy wzrost wartości dystrybuanty następuje pomiędzy 34 a 37,5 USD, a dla każdej wartości powyżej 60 USD dystrybuanta przyjmuje wartość 1.



Rysunek 4. Wykresy dystrybuant

### 3. Porównanie wyników

#### 3.1. Korelacja

Analiza korelacji pozwala nam na określenie, czy dwie zmienne losowe są ze sobą powiązane. Jeśli korelacja istnieje, to znaczy, że między zmiennymi jest jakaś relacja. Wartość współczynnika korelacji jest liczbą  $r$  z przedziału  $[-1, 1]$ , gdzie:

- $r < 0$  oznacza korelację ujemną, czyli gdy jedna zmienna rośnie, to druga maleje.
- $r > 0$  oznacza korelację dodatnią, czyli zmienne rosną i maleją jednocześnie.

Możemy podzielić ją również ze względu na jej siłę:

- $|r| < 0.2$  – brak korelacji między zmiennymi
- $0.2 < |r| < 0.4$  – słaba korelacja
- $0.4 < |r| < 0.7$  – umiarkowana korelacja
- $0.7 < |r| < 0.9$  – dość silna korelacja
- $0.9 < |r|$  – bardzo silna korelacja

##### 3.1.1. Współczynnik korelacji Pearsona

Współczynnik korelacji Pearsona pozwala nam zbadać liniową zależność między danymi, innymi słowy, czy dwie wartości rosną i maleją proporcjonalnie do siebie nawzajem. Współczynnik ten obliczamy ze wzoru:

$$r_{XY} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2} \sqrt{\sum (Y_i - \bar{Y})^2}}$$

Współczynnik korelacji Pearsona dla danych dot. cen akcji firm Pfizer i Moderna wyniósł  $r_{X,Y} \approx 0.56$ , co oznacza umiarkowaną korelację liniową powyższych zmiennych.

##### 3.1.2. Współczynnik korelacji Spearmana

Współczynnik korelacji Spearmana wykrywa natomiast dowolną monotoniczną zależność, w związku z czym jest on ogólniejszy niż współczynnik korelacji Pearsona. Współczynnik ten obliczamy ze wzoru:

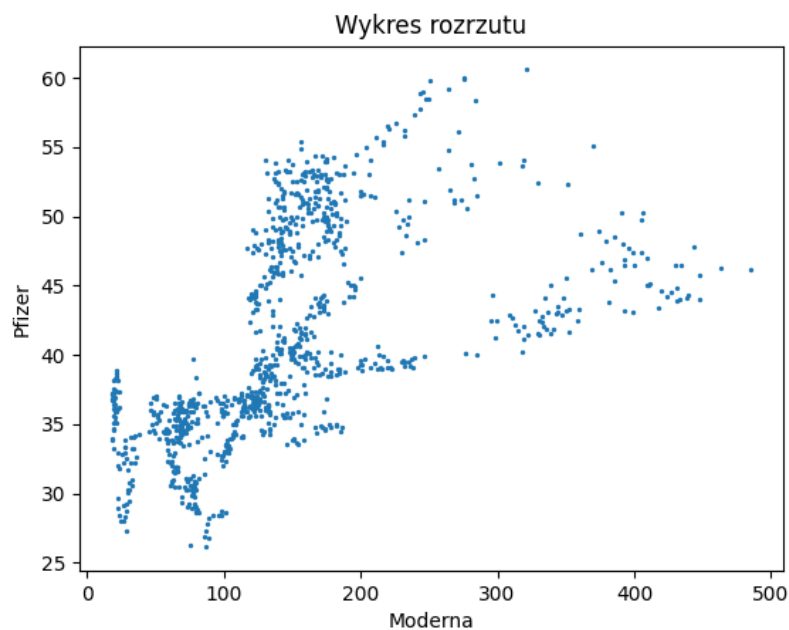
$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

gdzie  $d_i$  to różnica między rangami  $X_i$  i  $Y_i$  (numerami pozycji w przypadku rosnącego uszeregowania zbioru danych).

Współczynnik korelacji Spearmana dla danych dot. cen akcji firm Pfizer i Moderna wyniósł  $r_{X,Y} \approx 0.73$ , co oznacza dość silną zależność między powyższymi zmiennymi.

##### 3.1.3. Wykres rozrzutu

Wykres rozrzutu jest graficzna metoda przedstawienia korelacji między dwoma zmiennymi, polegający na naniesieniu na wykres odpowiadających sobie punktów  $(X, Y)$ .

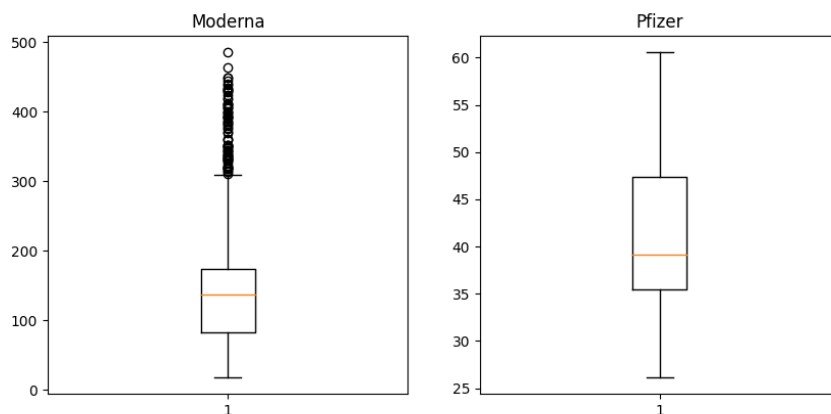


Rysunek 5. Wykres rozrzutu

Jak widać na powyższym wykresie, zmienne wykazują pewną zależność względnie monotoniczną, jednak nie jest ona jednoznacznie liniowa, co potwierdzają również wartości współczynników korelacji Pearsona i Spearmana.

### 3.2. Wykresy pudełkowe

Badane zbiory danych możemy graficznie zaprezentować za pomocą wykresów pudełkowych, ukazujących wizualnie kwartyle, wartości odstające i inne wiadomości dot. badanego rozkładu.



Rysunek 6. Wykres pudełkowy

Na wyżej zaprezentowanych wykresach możemy zaobserwować, że obliczone w sekcjach 2.2 *Mediana* i 2.3 *Kwartyle* wartości pokrywają się z tymi oznaczonymi na wykresie. Ponadto możemy w przypadku danych dot. firmy Moderna możemy zaobserwować liczne wartości odstające powyżej górnego wąsa. Nie pojawiają się one natomiast w przypadku danych dotyczących firmy Pfizer. Możemy również zauważyć, że w przypadku firmy Pfizer mamy do czynienia z wykresem prawostronnie asymetrycznym, natomiast w dla firmy Moderna jest on stosunkowo symetryczny.

Porównując oba wykresy możemy zauważyć, że mimo dosyć dużego współczynnika korelacji obliczonego w sekcji 3.1 *Korelacja*, rozkład danych znacząco się różni. Przede wszystkim, podstawową różnicą jest tu położenie - dla Moderny zakres danych to mniej więcej [18, 485], natomiast dla Pfizera [26, 60], co może wynikać z różnych zakresów działań dwóch firm w branży farmaceutycznej poza produkcją szczepionek przeciwko COVID-19. Ponadto zauważalne są też różnice dotyczące rozproszenia tych dwóch zbiorów oraz ich symetrii. Mogą one wynikać z różnorodnych czynników społeczno - gospodarczych, a także indywidualnych czynników dotyczących poszczególnych firm.

#### 4. Podsumowanie

Porównując wyniki obydwu firm, zaprezentowane w sekcji 2. *Analiza danych*, można zaobserwować, że średnia arytmetyczna, mediana, kwartyle oraz rozstęp międzykwartylowy cen akcji firmy Moderna przyjmują znacznie większe wartości niż w przypadku cen akcji firmy Pfizer. Może być to spowodowane z jednej strony stosunkowo dużą liczbą wartości odstających, widocznych na wykresie pudełkowym powyżej górnego wąsa, jak również faktem, że kursy otwarcia firmy Moderna znajdują się w znacznie większym zakresie ([18, 485] w porównaniu do [26, 60] dla Pfizera). Wariancja oraz odchylenie standardowe wskazują na znacznie większy rozrzut danych w przypadku firmy Moderna, co znajduje także odzwierciedlenie na histogramach oraz wykresie rozrzutu. Wykresy gęstości wskazują na największe prawdopodobieństwo przyjęcia przez ceny akcji wartości pomiędzy 130 a 150 USD dla firmy Moderna oraz pomiędzy 36 a 38 USD dla firmy Pfizer - oznacza to, że najbardziej prawdopodobne wartości dla firmy Moderna są niemalże 4 razy wyższe. Może to wynikać z czynników wspomnianych w sekcji 3.2. *Wykresy pudełkowe*. Współczynnik korelacji Pearsona wykazał umiarkowaną korelację liniową cen akcji obydwu firm, natomiast współczynnik korelacji Spearmana - dość silną zależność.

Pierwszą dopuszczoną do użytku szczepionką była szczepionka firmy Pfizer (zatwierdzona przez Komisję Europejską 21.12.2020, według informacji na stronie [www.gov.pl](http://www.gov.pl)). Szczepionka firmy Moderna została natomiast zatwierdzona 07.01.2021 i miała zostać dostarczana do Polski (i innych krajów) w okresie od pierwszego do trzeciego kwartału 2021 roku (informacje ze strony [www.gov.pl](http://www.gov.pl)). Właśnie w tym okresie firma odnotowywała największe ceny swoich akcji, według wykresu w sekcji 1. *Cel raportu*. Również firma Pfizer odnotowywała największe wartości kursów otwarcia w okresie dostarczania krajom na całym świecie szczepionek przeciwko COVID-19.

Na początku zestawienia, w styczniu 2020 roku, ceny akcji firmy Pfizer wynosiły 37,29 USD, natomiast Moderny - 19,57 USD. 31 grudnia 2023 roku, a zatem już po zakończeniu pandemii, kursy otwarcia firmy Pfizer były na poziomie 28,78 USD, natomiast firmy Moderna - 98,61 USD. Oznacza to, że obecnie kursy otwarcia firmy Pfizer znajdują się na podobnym poziomie, co na początku pandemii, natomiast firmy Moderna - osiągają wyższe wartości. Obydwie firmy odnotowywały najwyższe ceny akcji w okresie, w którym zapotrzebowanie na szczepionki na COVID-19 było największe.