

Raport I statystyka stosowana
Analiza cen (USD) złota i srebra za uncję
od 04.01.2016 r. do 28.12.2018 r.

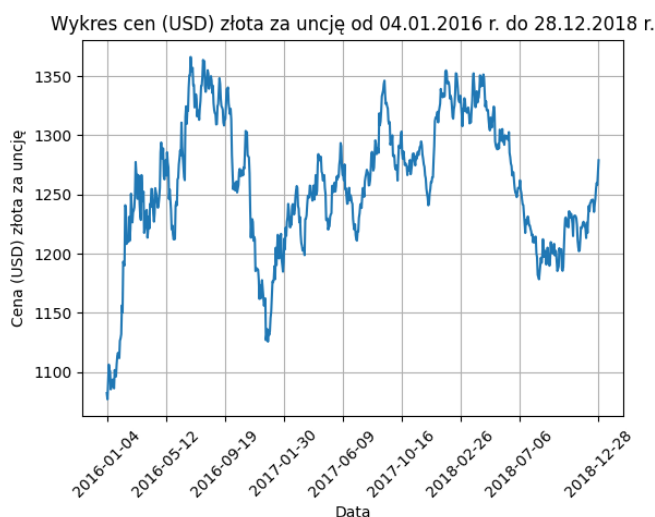
Paweł Stępień, 276 038

06.05.2024 rok

1 Wprowadzenie

Celem tego raportu jest analiza i porównanie cen dwóch metali szlachetnych: złota i srebra. Zarówno złoto, jak i srebro od wieków są uznawane za wartościowe kruszce, które odgrywają istotną rolę zarówno w dziedzinie inwestycji, jak i przemysłu. Ich ceny są często monitorowane przez inwestorów, ekonomistów i analityków jako wskaźniki zdrowia gospodarczego i stabilności finansowej. W niniejszym raporcie przeprowadzimy kompleksową analizę cen złota i srebra od 04.01.2016 r. do 28.12.2018 r. oraz zbadamy ich wzajemne relacje [3].

2 Wizualizacja danych



Rysunek 1: Wykres cen (USD) złota za uncję od 04.01.2016 r. do 28.12.2018 r.



Rysunek 2: Wykres cen (USD) srebra za uncję od 04.01.2016 r. do 28.12.2018 r.

Wykresy Rysunek 1 i Rysunek 2 przedstawiają ceny (USD) złota i srebra za uncję w latach 2016-2018 na rynku światowym. Został stworzony na podstawie danych historycznych, które zostały zebrane przez Londyńskie Stowarzyszenie Rynku Surowców (ang. LBMA - skrót od "London Bullion Market Association"). Jest to międzynarodowa organizacja, która nadzoruje i reguluje rynek złota i srebra w Londynie, jednym z najważniejszych rynków surowców na świecie. Dane to dzienne notowania cen metali szlachetnych, od 4 stycznia 2016 roku do dnia 28 grudnia 2018 roku, uwzględniając wszystkie dni robocze, wyłączając soboty, niedziele oraz święta [1], [2], [3].

Obserwując wykres, można zauważyć, że ceny obu metali podlegają wahanom w czasie. W pewnych okresach występują wyraźne zmiany cenowe, które są spowodowane czynnikami zewnętrznymi, takimi jak polityka gospodarcza, zmiany kursów walutowych czy globalne wydarzenia geopolityczne. W niektórych okresach ceny złota i srebra wydają się być skorelowane, co sugerują podobne trendy cenowe. Jednak w innych okresach można zaobserwować różnice w zachowaniu cen obu metali.

Na początku badanego okresu widać gwałtowny skok cen, by następnie zaobserwować znaczący spadek. Po tych wydarzeniach cena srebra jest stosunkowo stabilna, podczas gdy cena złota zaczyna stopniowo wzrastać. Zbliżając się do końca przedziału czasowego, złoto szybko traci na wartości, a następnie jego cena wzrasta. Natomiast cena srebra notuje jedno z najniższych wartości w latach 2016–2018.

3 Wykresy pudełkowe

3.1 Mediana

Definicja

Wartość cechy w szeregu uporządkowanym, powyżej i poniżej której znajduje się jednakowa liczba obserwacji. Mediana jest drugim kwartylem [4].

3.2 Pierwszy kwartył

Definicja

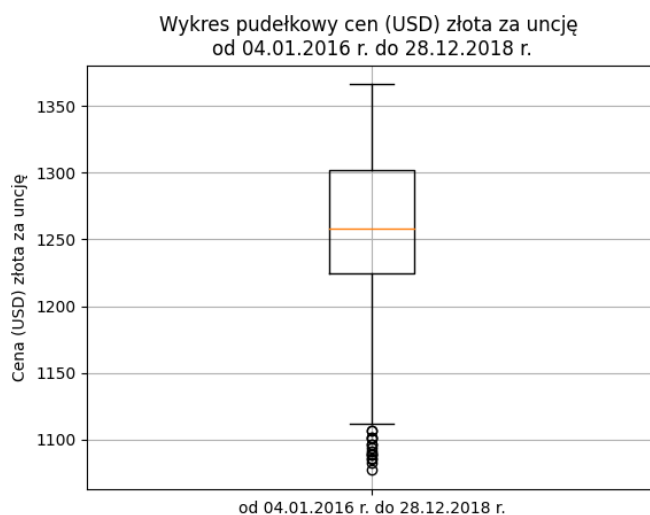
Kwantyl rzędu $1/4 = 25\%$ obserwacji jest położonych poniżej. Oznaczamy Q_1 [5].

3.3 Trzeci kwartył

Definicja

Kwantyl rzędu $3/4 = 75\%$ obserwacji dzieli zbiór obserwacji na dwie części odpowiednio po 75% położonych poniżej tego kwartyła i 25% położonych powyżej. Oznaczamy Q_3 [5].

3.4 Wykres pudełkowy dla złota

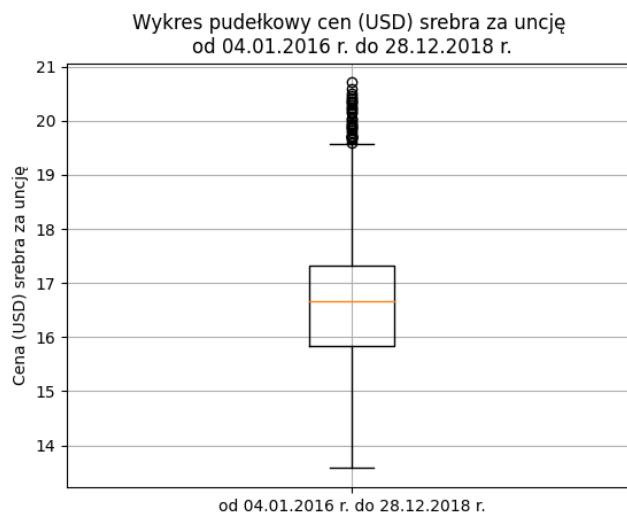


Rysunek 3: Wykres pudełkowy cen (USD) złota za uncję od 04.01.2016 r. do 28.12.2018 r.

Na Rysunku 3 widzimy wykres pudełkowy cen (USD) złota za uncję w latach 2016-2018. Mediana wyniosła 1258.15 USD. Pierwszy kwartył równa się

1224.28 USD, natomiast trzeci kwartył 1301.88 USD. Ponadto możemy zobaczyć, że linia wyznaczająca medianę znajduje się blisko środka pudełka, tworzonego przez pierwszy i trzeci kwartył. Świadczy to, że dane mają rozkład symetryczny. Dolny wąs wyniósł 1111.80 USD, natomiast górny równa się 1366.25 USD. Warto zwrócić także uwagę na dolne obserwacje odstające (dolne outliery). Występowanie wartości odstających powoduje, że wąsy są nierównomierne. Widzimy wiele punktów znajdujących się poniżej dolnego wąsa. Ich występowanie wynika z faktu, że przez bardzo krótki okres czasu, w styczniu 2016 roku, notowania cen złota były bardzo niskie. Następnie wystąpił gwałtowny wzrost cen i dużo wyższe ceny utrzymywały się już do końca naszych obserwacji [12], [13].

3.5 Wykres pudełkowy dla srebra



Rysunek 4: Wykres pudełkowy cen (USD) srebra za uncję od 04.01.2016 r. do 28.12.2018 r.

Na Rysunku 4 widzimy wykres pudełkowy cen (USD) srebra za uncję w latach 2016-2018. Mediana wyniosła 16.67 USD. Pierwszy kwartył równa się 15.83 USD, natomiast trzeci kwartył 17.33 USD. Ponadto możemy zobaczyć, że linia wyznaczająca medianę znajduje się blisko środka pudełka, tworzonego przez pierwszy i trzeci kwartył. Świadczy to, że dane mają rozkład symetryczny. Dolny wąs wyniósł 13.58 USD, natomiast górny równa się 19.58 USD. Warto zwrócić także uwagę na górne obserwacje odstające (górne outliery). Widzimy wiele punktów znajdujących się powyżej górnego wąsa. Ich występowanie wynika z faktu, że przez bardzo krótki okres czasu, od czerwca do października 2016 roku, notowania cen srebra były zdecydowanie wyższe niż w pozostałych okresach badanego przedziału [12], [13].

4 Porównanie średniej arytmetycznej i mediany

4.1 Średnia arytmetyczna

Definicja

Suma liczb podzielona przez ich liczbę. Średnią arytmetyczną oznaczamy przez \bar{x} [6].

Dla n liczb a_1, a_2, \dots, a_n jest to więc wyrażenie:

$$\bar{x} = \frac{a_1 + a_2 + \dots + a_n}{n}.$$

4.2 Odchylenie standardowe

Definicja

Odchylenie standardowe mówi, jak szeroko wartości jakiejś wielkości są rozrzucone wokół jej średniej. Im mniejsza wartość odchylenia tym obserwacje są bardziej skupione wokół średniej. Odchylenie standardowe oznaczamy σ [7]. Wyraża się wzorem:

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

4.3 Wnioski płynące z wartości średniej arytmetycznej, mediany i odchylenia standardowego

4.3.1 złoto

Mediana cen (USD) złota za uncję w latach 2016-2018 wyniosła 1258.15 USD. Średnia arytmetyczna równa się 1258.79 USD, natomiast odchylenie standardowe 56.48 USD.

Mediana i średnia arytmetyczna mają zbliżone wartości, oznacza to, że rozkład danych może być zbliżony do symetrycznego. Dodatkowo odchylenie standardowe jest małe, co sugeruje, że wartości danych są blisko średniej arytmetycznej. Zachodzi możliwość, że występujące wartości odstające, nie mają znaczącego wpływu na centralną tendencję danych.

4.3.2 srebro

Mediana cen (USD) srebra za uncję w latach 2016-2018 wyniosła 16.67 USD. Średnia arytmetyczna równa się 16.64 USD, natomiast odchylenie standardowe 1.44 USD.

Mediana i średnia arytmetyczna mają zbliżone wartości. Odchylenie standardowe jest małe. Możemy wysnuć podobne wnioski jak w przypadku cen złota w podpunkcie 4.3.1.

5 Dystrybuanty empiryczne

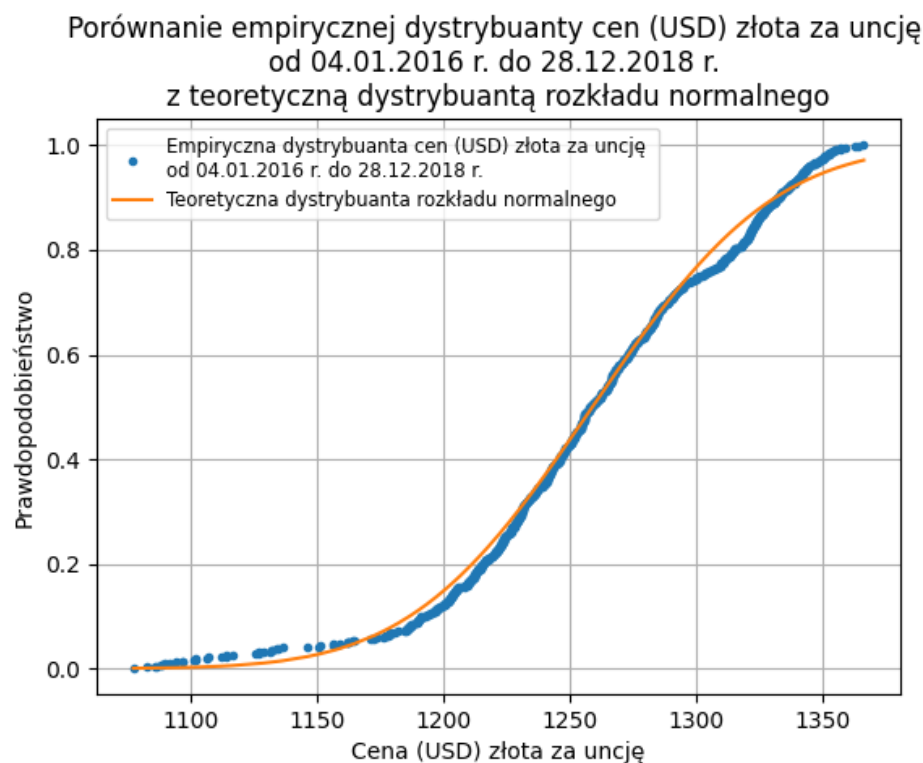
5.1 Definicja

Niech \mathbb{P} będzie rozkładem prawdopodobieństwa na prostej. Funkcję $F: \mathbb{R} \rightarrow \mathbb{R}$ daną wzorem:

$$F(t) = \mathbb{P}((-\infty, t])$$

nazywamy dystrybuantą rozkładu \mathbb{P} . W statystyce dystrybuanta rozkładu próby zwana jest dystrybuantą empiryczną [8].

5.2 Złoto



Rysunek 5: Porównanie empirycznej dystrybuanty cen złota (USD) za uncję od 04.01.2016 r. do 28.12.2018 r. z teoretyczną dystrybuantą rozkładu normalnego.

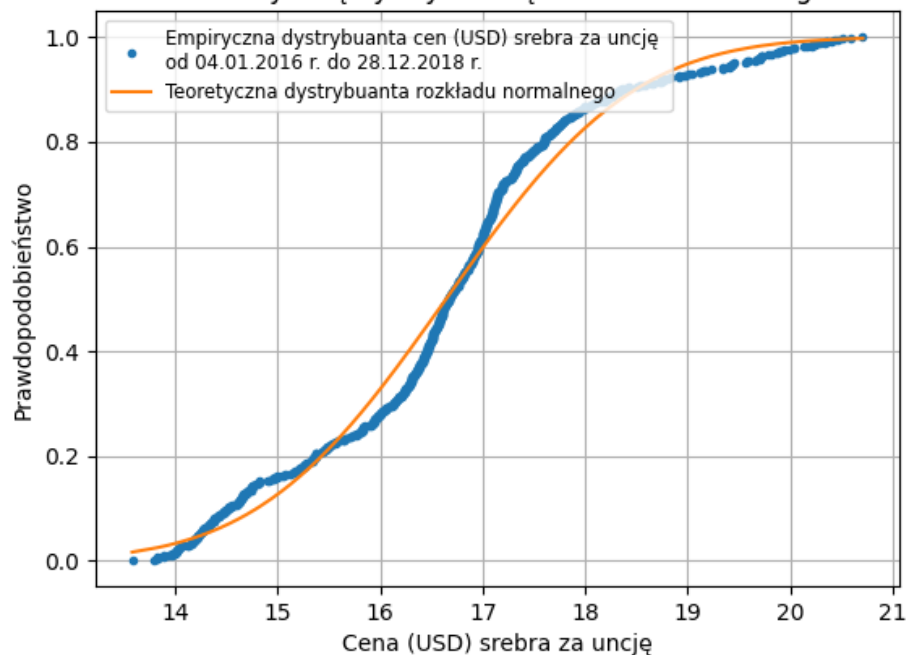
W trakcie analizy danych dotyczących cen (USD) złota za uncję w latach 2016-2018 możemy zauważyć, że dystrybuanta empiryczna cen złota jest bardzo

podobna do teoretycznej dystrybuanty rozkładu normalnego. Wynik ten sugeruje, że rozkład cen złota może być zbliżony do rozkładu normalnego w badanym okresie czasu.

Jednak warto zaznaczyć, że choć zaobserwowaliśmy podobieństwo rozkładu normalnego, musimy zachować ostrożność i przeprowadzić dodatkowe testy statystyczne, aby potwierdzić to założenie [9].

5.3 Srebro

Porównanie empirycznej dystrybuanty cen (USD) srebra za uncję od 04.01.2016 r. do 28.12.2018 r. z teoretyczną dystrybuantą rozkładu normalnego



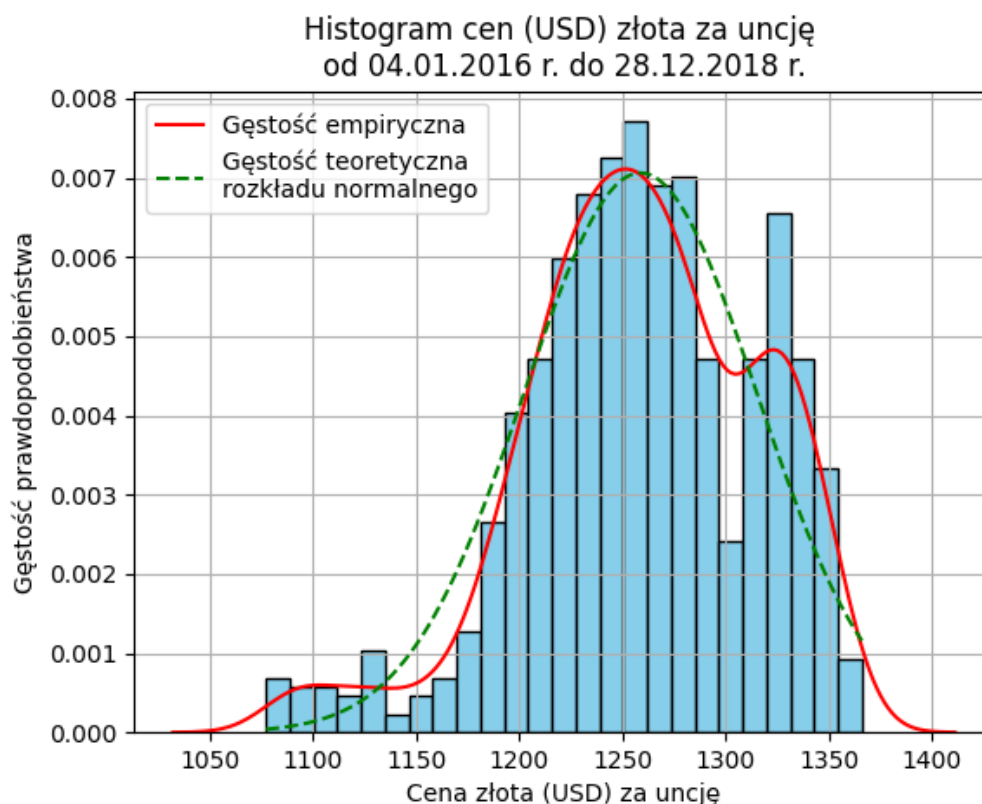
Rysunek 6: Porównanie empirycznej dystrybuanty cen srebra (USD) za uncję od 04.01.2016 r. do 28.12.2018 r. z teoretyczną dystrybuantą rozkładu normalnego.

Podczas analizowania danych dotyczących cen (USD) srebra za uncję w latach 2016-2018, widzimy, że empiryczna dystrybuanta cen srebra wykazuje pewne podobieństwo do teoretycznej dystrybuanty rozkładu normalnego.

Choć obserwowane podobieństwo jest znaczące, dystrybuanta empiryczna nie jest idealnie zbieżna z dystrybuantą teoretyczną. Sugeruje to, że rozkład cen srebra może mieć cechy zbliżone do rozkładu normalnego, ale mogą również występować pewne odstępstwa [9].

6 Histogramy i gęstość empiryczna

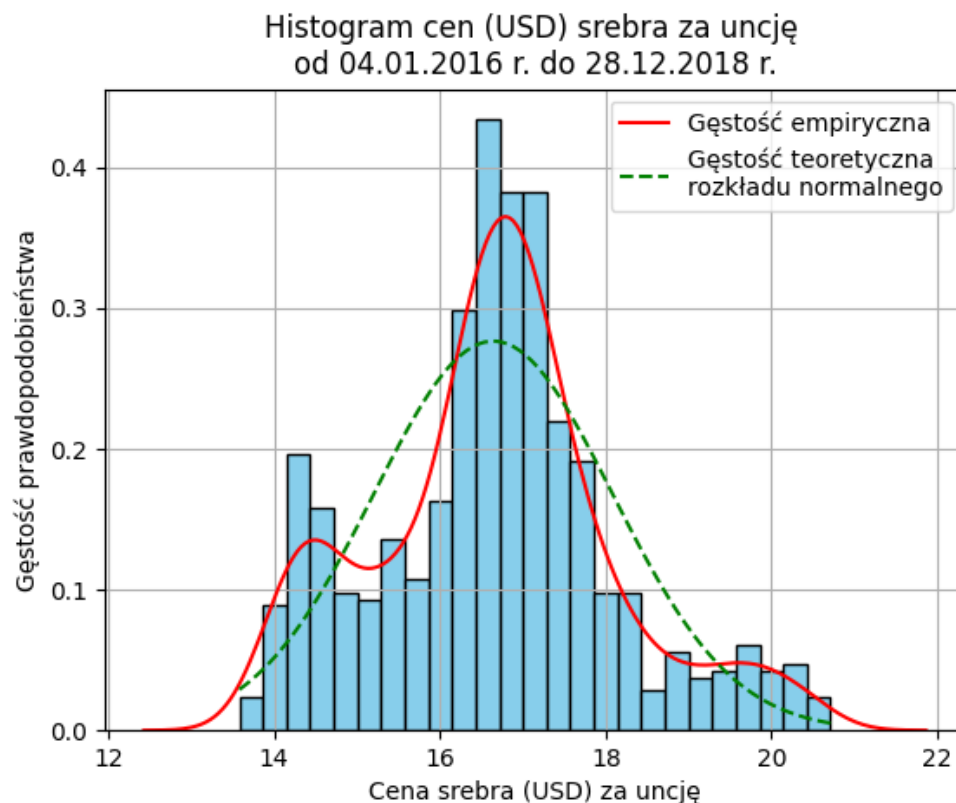
6.1 Złoto



Rysunek 7: Histogram cen złota (USD) za uncję od 04.01.2016 r. do 28.12.2018 r. oraz porównanie gęstości empirycznej z teoretyczną gęstością rozkładu normalnego.

Histogram przedstawia rozkład cen złota w analizowanym okresie czasu poprzez podział zakresu cen na 25 przedziałów wartości. Największe prawdopodobieństwo występuje dla cen oscylujących w okolicach 1250 USD za uncję. Histogram jest symetryczny z lekkimi anomaliasi po prawej stronie [10]. Empiryczna dystrybuanta cen złota wykazała wyraźne podobieństwo do teoretycznej dystrybuanty rozkładu normalnego. Takie rezultaty sugerują, że rozkład cen złota w analizowanym okresie czasu może być zbliżony do rozkładu normalnego. Jednak pomimo zaobserwowanego podobieństwa do rozkładu normalnego, konieczne jest zachowanie ostrożności i przeprowadzenie dodatkowych testów statystycznych w celu potwierdzenia tego założenia [9].

6.2 Srebro



Rysunek 8: Histogram cen srebra (USD) za uncję od 04.01.2016 r. do 28.12.2018 r. oraz porównanie gęstości empirycznej z teoretyczną gęstością rozkładu normalnego.

Histogram przedstawia rozkład cen srebra w analizowanym okresie czasu poprzez podział zakresu cen na 25 przedziałów wartości. Największe prawdopodobieństwo występuje dla cen oscylujących w przedziale 16.5 - 17.4 USD za uncję. Jest ono zdecydowanie wyższe niż dla pozostałych przedziałów. Histogram jest symetryczny z lekkimi anomaliami po lewej stronie [10]. Możemy zaobserwować różnice między empiryczną gęstością prawdopodobieństwa, a teoretyczną gęstością rozkładu normalnego. Choć obie gęstości są podobne, nie występuje idealne pokrycie między nimi [9].

7 Współczynnik korelacji Pearsona

7.1 Definicja

Współczynnik określający poziom zależności liniowej między zmiennymi losowymi. Został opracowany przez Karla Pearsona.

Wartość współczynnika korelacji mieści się w przedziale domkniętym $[-1,1]$. Im większa jego wartość bezwzględna, tym silniejsza jest zależność liniowa między zmiennymi. $r_{xy} = 0$ oznacza brak liniowej zależności między cechami, $r_{xy} = 1$ oznacza dokładną dodatnią liniową zależność między cechami, natomiast $r_{xy} = -1$ oznacza dokładną ujemną liniową zależność między cechami [11]. Współczynnik wyraża się wzorem:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

7.2 Wartość współczynnika dla badanych danych

W przypadku cen (USD) złota i srebra za uncję w latach od 04.01.2016 r. do 28.12.2018 r. współczynnik korelacji Pearsona wyniósł 0.61. Wartość ta wskazuje na umiarkowaną dodatnią zależność liniową między cenami złota, a srebra w tym okresie. Oznacza to, że w większości przypadków, gdy ceny złota wzrastają, ceny srebra również rosną, ale stopień zależności nie jest bardzo silny.

8 Podsumowanie

Analiza cen (USD) srebra i złota za uncję od 04.01.2016 r. do 28.12.2018 r. pozwoliła nam na lepsze zrozumienie dynamiki rynków surowców oraz ich wzajemnych relacji.

Już wizualizacja danych, pozwoliła wysnuć wnioski na temat korelacji cen obu metali szlachetnych. Współczynnik korelacji Pearsona między cenami złota, a srebra wyniósł 0.61, co wskazuje na umiarkowaną dodatnią zależność liniową między cenami złota a srebra w latach 2016-2018.

Dystrybucje dla obu surowców pokazały, że zarówno ceny srebra, jak i złota, wykazują tendencję zbliżania się do rozkładu normalnego. Dla złota jest to wyraźniej widoczne, podczas gdy srebro ma pewne odstępstwa. Empiryczne gęstości prawdopodobieństwa dla cen srebra i złota również wykazały duży stosunek do gęstości teoretycznych rozkładów normalnych. Ponownie dla cen srebra zauważono pewne różnice.

Zarówno histogram cen złota, jak i srebra są symetryczne, jednak oba posiadają także pewne anomalie.

Wykresy pudełkowe również pomogły zobaczyć, że dane mają rozkład symetryczny, dzięki znajdującej się idealnie po środku pudełka kresce oznaczającej medianę. Ponadto zaobserwowaliśmy obserwacje odstające. Dla złota punkty znajdujące się poniżej dolnego wąsa - dolne obserwacje odstające, natomiast

dla srebra, punkty znajdujące się powyżej górnego wąsa - górne obserwacje odstające.

W przypadku złota, wartość mediany i średniej arytmetycznej były bardzo zbliżone. Analogiczna sytuacja występowała również w kontekście cen srebra.

9 Bibliografia

Dane pochodzą ze strony Kaggle, która pozyskała je od LBMA.

Korzystałem z języka programowania Python do pomocy w obliczeniach i wizualizacji wykresów.

Literatura

- [1] Strona internetowa LBMA, <https://www.lbma.org.uk>, dostęp: 5 maj 2024 r.
- [2] Wikipedia LBMA, https://en.wikipedia.org/wiki/London_Bullion_Market_Association, dostęp: 5 maj 2024 r.
- [3] Kaggle, <https://www.kaggle.com/datasets/lbronchal/gold-and-silver-prices-dataset>, dostęp: 5 maj 2024 r.
- [4] Wikipedia mediana, <https://pl.wikipedia.org/wiki/Mediana>, dostęp: 5 maj 2024 r.
- [5] Wikipedia kwartył, <https://pl.wikipedia.org/wiki/Kwartył>, dostęp: 5 maj 2024 r.
- [6] Wikipedia średnia arytmetyczna, https://pl.wikipedia.org/wiki/%C5%9Arednia_arytmetyczna, dostęp: 5 maj 2024 r.
- [7] Wikipedia odchylenie standardowe, https://pl.wikipedia.org/wiki/Odchylenie_standardowe, dostęp: 5 maj 2024 r.
- [8] Wikipedia dystrybucja, <https://pl.wikipedia.org/wiki/Dystrybucja>, dostęp: 5 maj 2024 r.
- [9] Wikipedia rozkład normalny, https://pl.wikipedia.org/wiki/Rozk%C5%82ad_normalny, dostęp: 5 maj 2024 r.
- [10] Wikipedia histogram, <https://pl.wikipedia.org/wiki/Histogram>, dostęp: 5 maj 2024 r.
- [11] Wikipedia współczynnik korelacji Pearsona, https://pl.wikipedia.org/wiki/Wsp%C3%B3%C5%82czynnik_korelacji_Pearsona, dostęp: 5 maj 2024 r.
- [12] Wikipedia wykres pudełkowy, https://pl.wikipedia.org/wiki/Wykres_pudełkowy, dostęp: 5 maj 2024 r.

- [13] Strona o statystyce - wykres pudełkowy,
<https://pogotowiestatystyczne.pl/wykres-skrzynkowy-moc-informacji-na-jednym-rysunku/>,
dostęp: 5 maj 2024 r.