

Uniwersytet Warszawski  
Wydział Nauk Ekonomicznych

Mikołaj Rymajdo  
Nr albumu: 432511

## **Analiza Szeregów Czasowych**

Praca zaliczeniowa

Praca wykonana pod kierunkiem  
dr Natalii Nehrebeckiej  
z Katedry Statystyki i Ekonometrii  
WNE UW

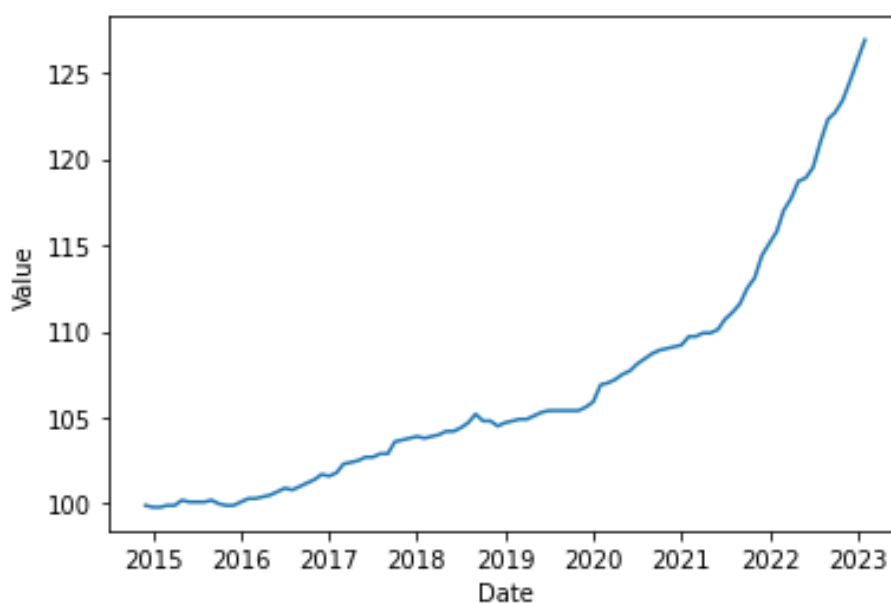
Warszawa, czerwiec 2023

## Rozdział I – Szereg niesezonowy

### 1.1. Opis danych

Szereg niesezonowy wykorzystany w pracy dotyczy zharmonizowanego wskaźnika cen konsumpcyjnych (HICP) instrumentów muzycznych w Polsce, gdzie rokiem bazowym jest 2015 (2015=100). Pobrane dane mają charakter miesięczny i nie są skorygowane sezonowo. Liczba obserwacji wynosi 99 - pierwsza z nich wystąpiła w grudniu 2014 roku, ostatnia w lutym 2023 roku. Dane zostały pobrane z bazy danych Eurostatu. Wykres opisanego szeregu znajduje się na Rys 1. Na podstawie wykresu można stwierdzić brak sezonowości w szeregu czasowym, a także widoczny jest wyraźny deterministyczny trend rosnący, ponieważ inflacja z założenia dąży do stałej wartości.

Rys 1. Wykres zharmonizowanego wskaźnika cen konsumpcyjnych (HICP) instrumentów muzycznych w Polsce w kolejnych miesiącach (2015=100)



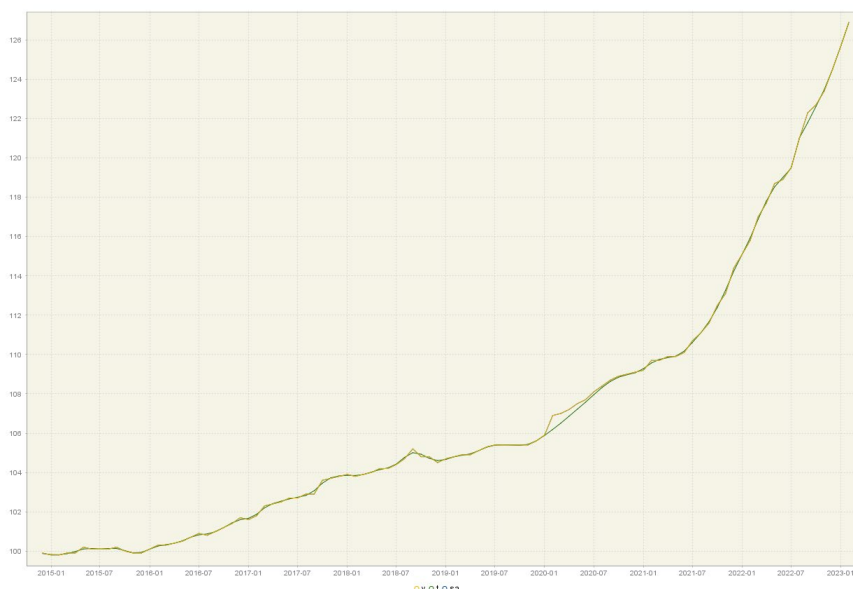
Źródło: Opracowanie własne na podstawie danych z Eurostatu

### 1.2. Dekompozycja szeregu

Na Rys. 2 dokonano dekompozycji szeregu czasowego w programie Demetra przy pomocy metody TramoSeats. Szereg odsezonowany nie różni się od rzeczywistego, co więcej brak sezonowości został potwierdzony testami diagnostycznymi. Niestety została wykryta 1 obserwacja odstająca w lutym 2020 roku. Być może jest to związane z wprowadzaniem restrykcji dotyczących pandemii COVID-19 w Europie. Wiele osób zaczęło spędzać więcej czasu w domu, co mogło doprowadzić do zwiększonego popytu na instrumenty muzyczne. Niektórzy początkujący muzycy nie posiadają własnych instrumentów, lecz ćwiczą na

sprzęcie ze szkół muzycznych. Demetra zdecydowała się na zlogarytmowanie szeregu, a więc dane na potrzeby dalszej analizy są w postaci logarytmicznej.

Rys 2. Dekompozycja szeregu czasowego w programie Demetra

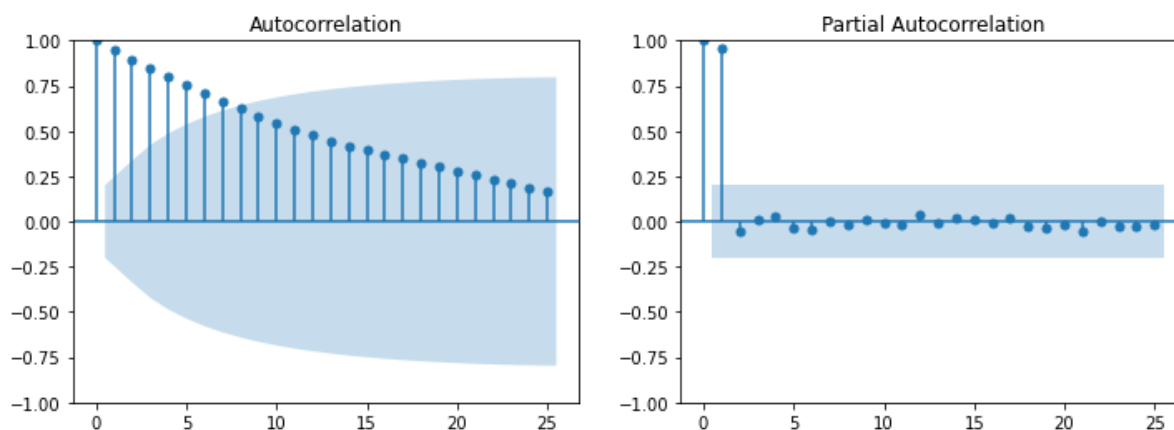


Źródło: Opracowanie własne na podstawie danych z Eurostatu

### 1.3. Model ARIMA

Na potrzeby modelowania wydzielono out-of-sample, która jest równa 4 ostatnim obserwacjom. Aby przeprowadzić estymację dla modelu ARIMA wymagana jest stacjonarność szeregu. Dokonano analizy wykresów ACF i PACF (Rys. 3) oraz wykonano rozszerzony test Dickeya-Fullera (Rys. 4) dla 5 rozszerzeń. Wyniki wskazują na istnienie autokorelacji oraz brak stacjonarności szeregu. Liczba rozszerzeń do wszystkich testów ADF została dobrana automatycznie tak, aby otrzymać nieskorelowane reszty.

Rys. 3. Funkcje ACF i PACF dla szeregu bazowego



Źródło: Opracowanie własne z wykorzystaniem języka Python

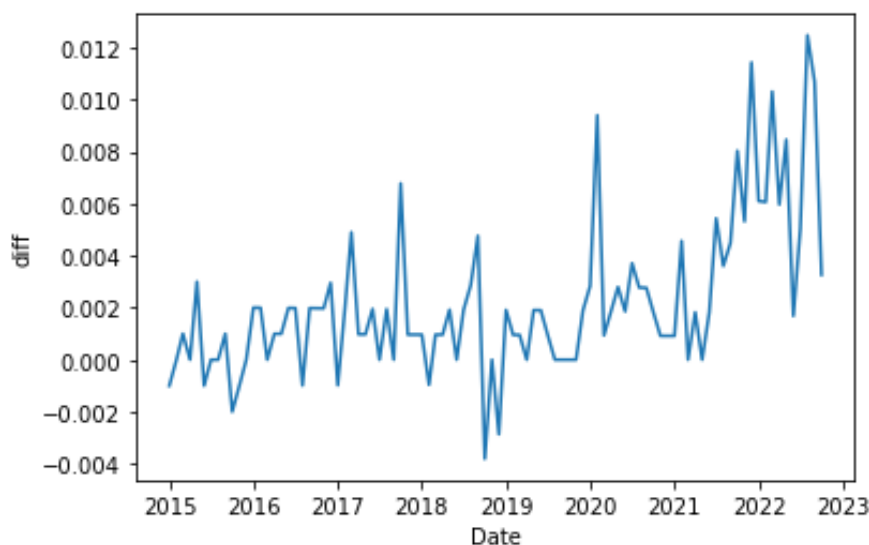
Rys. 4. Rozszerzony test Dickeya-Fullera dla szeregu bazowego

```
1. ADF : 2.16377160595637
2. P-Value : 0.9988513486379113
3. Num Of Lags : 5
4. Num Of Observations Used For ADF Regression and Critical Values Calculation : 89
5. Critical Values :
    1% : -3.506057133647011
    5% : -2.8946066061911946
    10% : -2.5844100201994697
Brak podstaw do odrzucenia H0, szereg jest niestacjonarny
```

Źródło: Opracowanie własne z wykorzystaniem języka Python

Podjęto decyzję o zastosowaniu różnicowania 1 stopnia. Rys. 5 przedstawia szereg po zróżnicowaniu. Po pierwszym różnicowaniu szereg dalej wydaje się nie być stacjonarny. Po 2021 roku występuje nietypowe odchylenie średniej wartości szeregu od zera. Brak stacjonarności potwierdza test ADF, którego statystyka wyniosła -0,505 dla 7 rozszerzeń (Rys. 6).

Rys. 5. Pierwsze różnice szeregu HICP instrumentów muzycznych w Polsce



Źródło: Opracowanie własne z wykorzystaniem języka Python

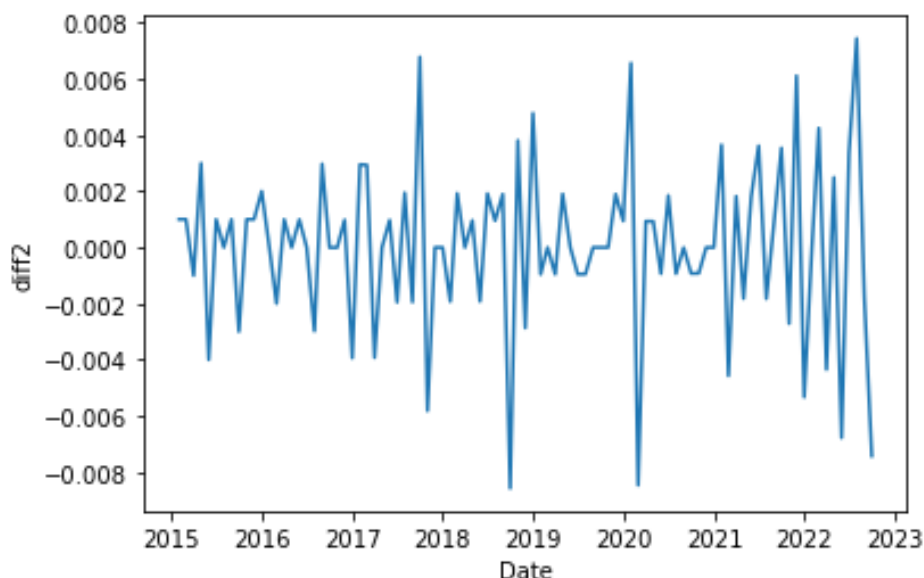
Rys. 6. Rozszerzony test Dickeya-Fullera dla szeregu zróżnicowanego jednokrotnie

```
1. ADF : -0.5054331428493557
2. P-Value : 0.8909518871262498
3. Num Of Lags : 7
4. Num Of Observations Used For ADF Regression and Critical Values Calculation : 86
5. Critical Values :
    1% : -3.5087828609430614
    5% : -2.895783561573195
    10% : -2.5850381719848565
Brak podstaw do odrzucenia H0, szereg jest niestacjonarny
```

Źródło: Opracowanie własne z wykorzystaniem języka Python

Następnie podjęto decyzję o zastosowaniu różnicowania 2 stopnia. Rys. 7 przedstawia szereg po kolejnym zróżnicowaniu. Szereg w tym momencie wydaje się być stacjonarny. Dokonano testu ADF, aby potwierdzić tę obserwację (Rys. 8). Dla 6 rozszerzeń otrzymaliśmy wiarygodny test, który pozwolił na nieodrzućcenie hipotezy zerowej o stacjonarności szeregu.

Rys. 7. Drugie różnice szeregu HICP instrumentów muzycznych w Polsce



Źródło: Opracowanie własne z wykorzystaniem języka Python

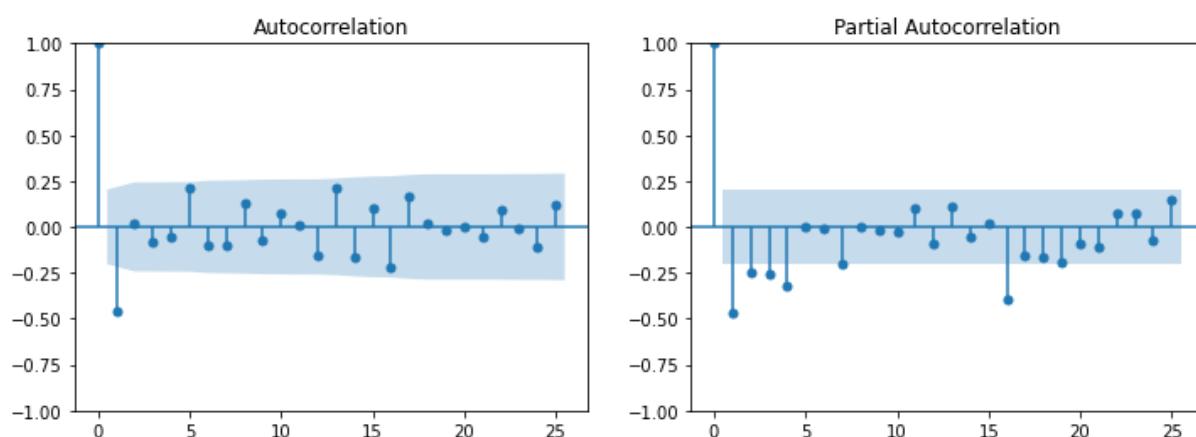
Rys. 8. Rozszerzony test Dickeya-Fullera dla szeregu zróżnicowanego dwukrotnie

```
1. ADF : -5.397912963572342
2. P-Value : 3.4215602313270926e-06
3. Num Of Lags : 6
4. Num Of Observations Used For ADF Regression and Critical Values Calculation : 86
5. Critical Values :
   1% : -3.5087828609430614
   5% : -2.895783561573195
  10% : -2.5850381719848565
Odrzucamy H0, szereg jest stacjonarny
```

Źródło: Opracowanie własne z wykorzystaniem języka Python

Ostatecznie szereg jest zintegrowany w stopniu drugim. Możemy przejść do wyznaczenia rzędów części autoregresyjnej oraz średniej ruchomej w modelu ARIMA na podstawie zaprezentowanych poniżej wykresów funkcji autokorelacji i autokorelacji cząstkowej (Rys. 9). W obydwu procesach wydaje się zasadne uwzględnienie pierwszego opóźnienia procesu MA. Przetestowano jednak wszystkie warianty modelu ARMIA, gdzie  $p$  i  $q$  jest  $\leq 5$ . Model ARIMA(0,2,1) osiągnął najniższe wartości kryteriów informacyjnych AIC oraz BIC wśród wszystkich modeli. Drugą i trzecią najniższą wartość AIC uzyskały modele (0,2,2) oraz (1,2,1). Porównanie wyżej wymienionych modeli znajduje się w Tabeli 1.

Rys. 9. Funkcje ACF i PACF dla szeregu zróżnicowanego dwukrotnie



Źródło: Opracowanie własne z wykorzystaniem języka Python

Tab. 1. Wartości kryteriów informacyjnych AIC oraz BIC najlepszych modeli

Lp.	p	d	q	AIC	BIC
1	0	2	1	-853.698	-846.101
2	0	2	2	-852.873	-842.743
3	1	2	1	-852.010	-841.880

Źródło: Opracowanie własne z wykorzystaniem języka Python

Sprawdzono wymienione 3 modele pod kątem istotności zmiennych i jedynie model ARIMA(0,2,1) posiadał wszystkie procesy istotne. Uznano więc model ARIMA(0,2,1) za najlepszy, a dalsza analiza dotyczy tego modelu. Jest to również model, który został wskazany przez Demetrę jako najodpowiedniejszy. Sprawdzono czy reszty modelu ARIMA(0,2,1) są białym szumem. Przeprowadzono test Ljunga-Boxa oraz Boxa-Pierce'a dla 24 opóźnień (Rys. 10). Obydwa testy ujawniły p-value większe niż 0.05, a więc reszty zgodnie z założeniem formalnym modelu przyjmują wartości losowe.

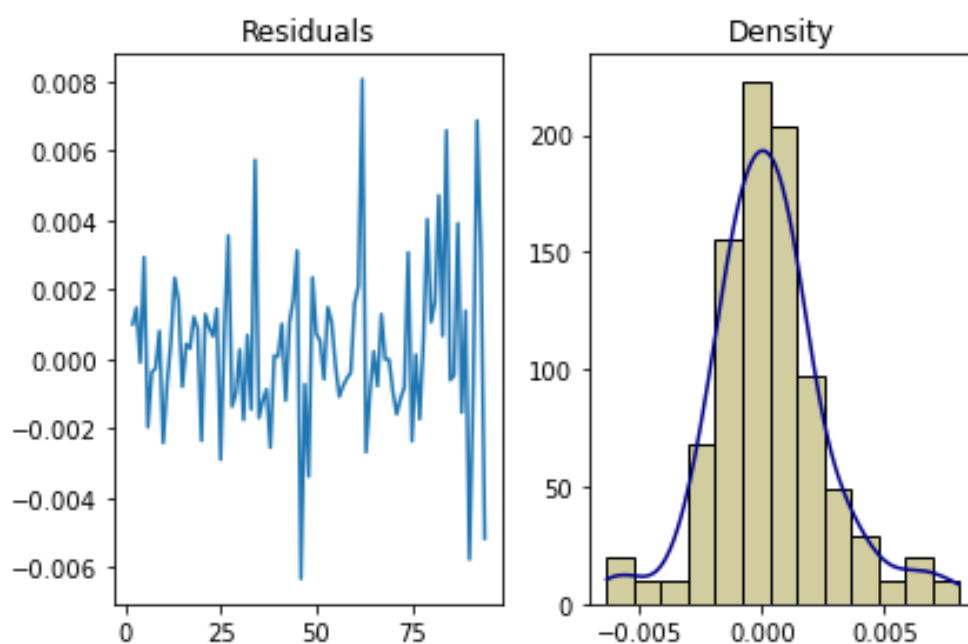
Rys. 10. Wyniki testu Ljunga-Boxa oraz Boxa-Pierce'a dla reszt modelu ARIMA (0,2,1)

	lb_stat	lb_pvalue	bp_stat	bp_pvalue
24	23.423265	0.494941	19.767	0.709946

Źródło: Opracowanie własne z wykorzystaniem języka Python

Dodatkowo przedstawiono reszty na wykresie rozrzutu oraz ich histogram (Rys. 11). Zgodnie z oczekiwaniami reszty oscylują wokół zera w sposób losowy, a rozkład reszt jest zbliżony do rozkładu normalnego.

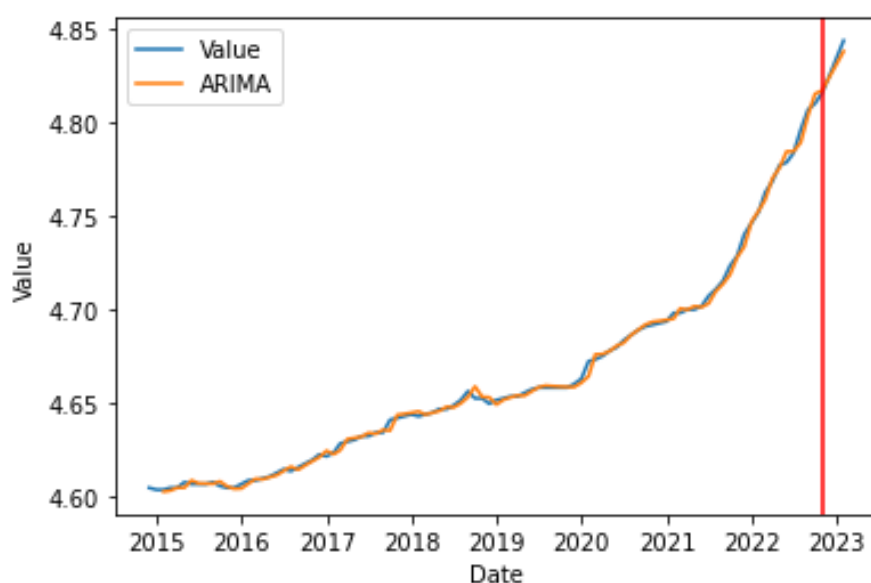
Rys 11. Wykres rozrzutu oraz histogram reszt modelu ARIMA(0,2,1)



Źródło: Opracowanie własne z wykorzystaniem języka Python

Wykonano prognozę dla modelu ARIMA(0,2,1), a jej wyniki przedstawiono na Rys. 12. Wartości prognozy są bardzo zbliżone do rzeczywistych wartości szeregu. Model przewidział trend rosnący o dużym nachyleniu. Teoretycznie model powinien być wykorzystywany jedynie do przewidywania na jeden okres, aczkolwiek przy dłuższym okresie również dobrze się sprawdził. Wartości dla pierwszej przewidzianej obserwacji są niemal identyczne i wynoszą odpowiednio 4.815431 (rzeczywista) i 4.816715 (prognoza).

Rys. 12. Wyniki prognozy dla modelu ARIMA(0,2,1)



Źródło: Opracowanie własne z wykorzystaniem języka Python

#### 1.4. Model ekstrapolacyjny

W tej sekcji użyjemy modelu liniowego Holta do utworzenia prognozy dla badanego szeregu czasowego. W poniższym modelu próba out-of-sample również wyniosła 4 obserwacje, aby wiarygodnie porównać prognozę między modelami. Utworzono dwa modele liniowe Holta, w których parametry szacowane są iteracyjnie. Pierwszy z nich uwzględnia trend addytywny, a drugi trend multiplikatywny. W Tabeli 2 przedstawione zostało porównanie modeli. Obydwa z nich mają bardzo zbliżone wartości błędów prognozy, parametrów oraz kryteriów informacyjnych. Jednakże dla trendu multiplikatywnego wszystkie wskaźniki są trochę korzystniejsze oraz wydaje się być bardziej zgodny z szeregiem, który przypomina funkcję wykładniczą, a więc został uznany jako lepszy.

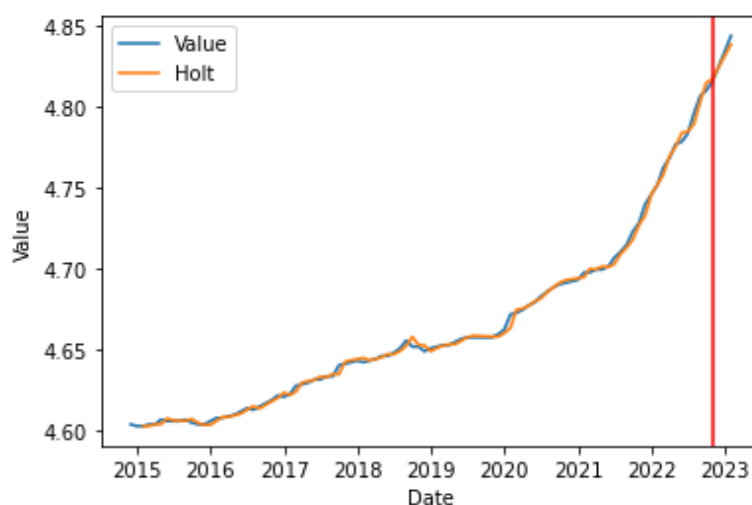
Tab. 2. Porównanie modeli Holta uwzględniających różne rodzaje trendu

	addytywny	multiplikatywny
Alpha	0.9714305	0.9714286
Beta	0.2312947	0.2312926
AIC	-1142.106	-1142.163
BIC	-1131.890	-1131.948
RMSE (out-of-sample)	0.003240	0.003100
MAE (out-of-sample)	0.002608	0.002491
MAPE (out-of-sample)	0.000539	0.000515

Źródło: Opracowanie własne z wykorzystaniem języka Python

Prognoza dla modelu Holta została przedstawiona na Rys. 13. Wartości prognozowane są bardzo bliskie rzeczywistym, a model poprawnie przedział silny trend rosnący.

Rys. 13. Wyniki prognozy dla modelu Holta z trendem multiplikatywnym



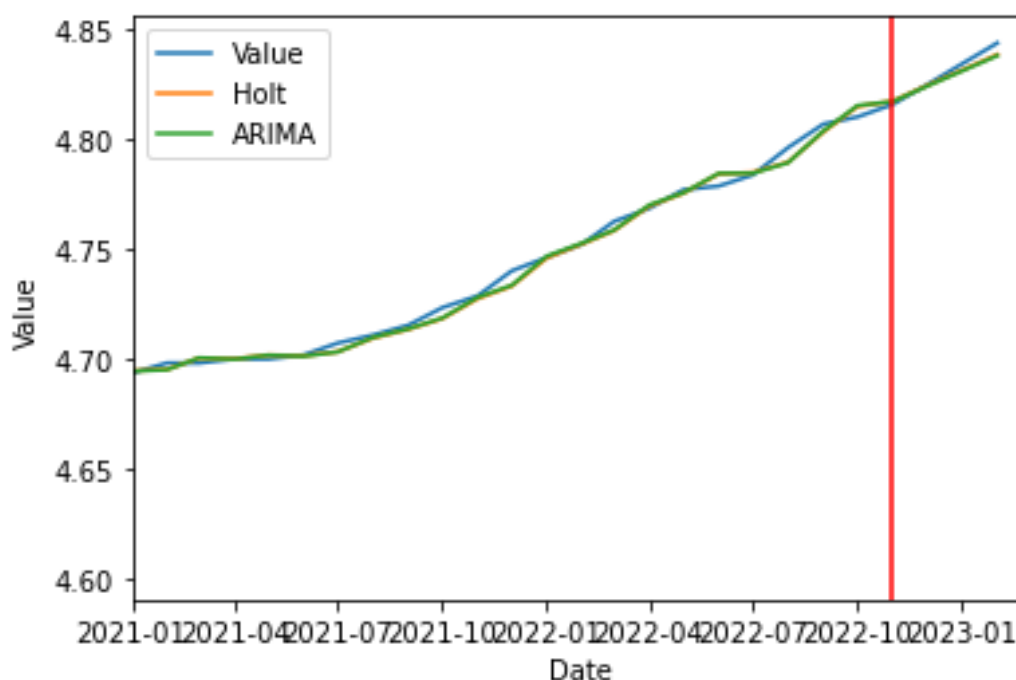
Źródło: Opracowanie własne z wykorzystaniem języka Python



## 1.5. Porównanie modeli

Wykonano porównanie modelu ARIMA(0,2,1) oraz modelu Holta uwzględniającego trend multiplikatywny. Na Rys. 14 widoczne jest porównanie modeli. Prognozy obydwu modeli są tak podobne, że nawet po zbliżeniu wykresu linie się pokrywają. Każdy z modeli bardzo dobrze poradził sobie z prognozowaniem przyszłych wartości szeregu.

Rys. 14. Porównanie modeli ARIMA oraz Holta



Źródło: Opracowanie własne z wykorzystaniem języka Python

Ostatecznie porównano modele przy pomocy wartości średnich błędów prognoz. Wyniki przedstawiono w Tabeli 3. Model Holta okazał się istotnie lepszy od modelu ARIMA, ponieważ wartości wszystkich jego błędów są niższe. Oznacza to, że prognozy modelu Holta są bliższe rzeczywistości, a jego dopasowanie do danych w próbie in-sample jest lepsze.

Tab. 3. Porównanie modelu ARIMA(0,2,1) oraz modelu Holta

	ARIMA	Holt
RMSE (out-of-sample)	0.003383	0.003100
MAE (out-of-sample)	0.002727	0.002491
MAPE (out-of-sample)	0.000564	0.000515
RMSE (in-sample)	0.002390	0.002350
MAE (in-sample)	0.001718	0.001665
MAPE (in-sample)	0.000367	0.000356

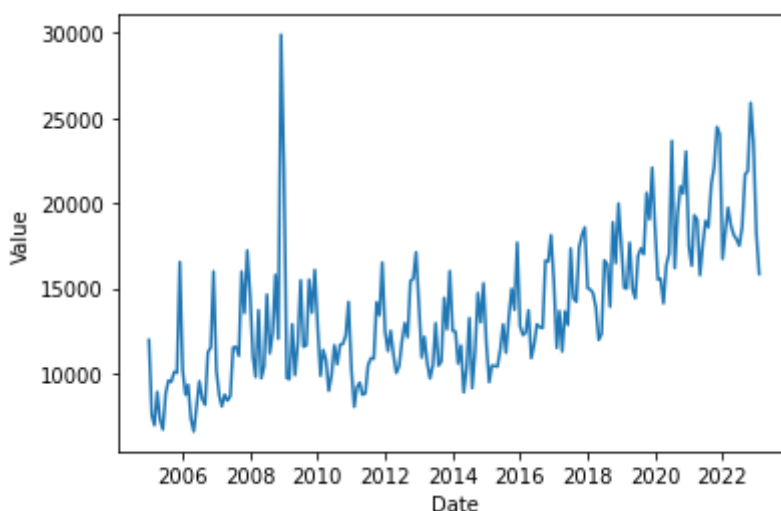
Źródło: Opracowanie własne z wykorzystaniem języka Python

## Rozdział II – Szereg sezonowy

### 2.1. Opis danych

Szereg sezonowy wykorzystany w pracy dotyczy mieszkań oddanych do użytkowania w Polsce. Pobrane dane mają charakter miesięczny i nie są skorygowane sezonowo. Liczba obserwacji wynosi 218 - pierwsza z nich wystąpiła w styczniu 2005 roku, a ostatnia w lutym 2023 roku. Dane zostały pobrane z bazy danych GUS. Wykres opisanego szeregu znajduje się na Rys 15. Na podstawie wykresu można stwierdzić obecność sezonowości w szeregu czasowym.

Rys. 15. Mieszkania oddawane do użytkowania w Polsce w kolejnych miesiącach



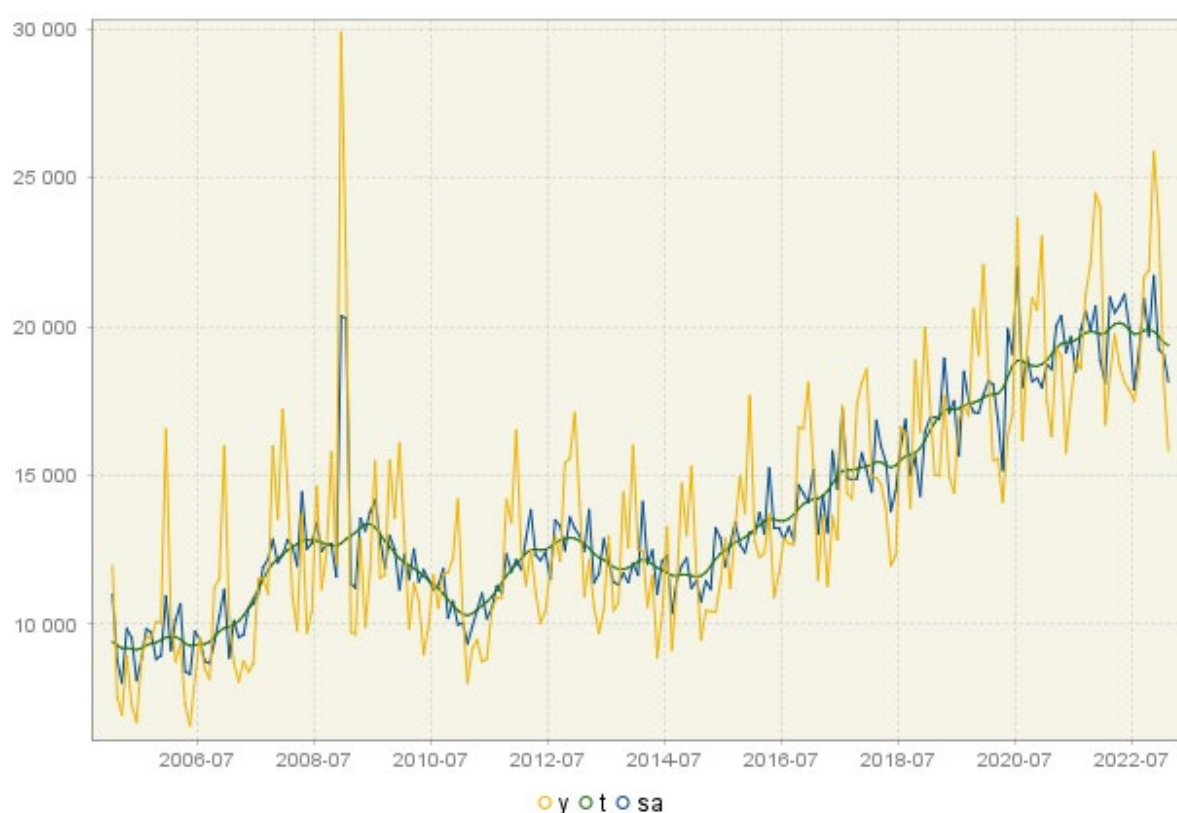
Źródło: Opracowanie własne na podstawie danych z Eurostatu

### 2.2. Dekompozycja szeregu

Na Rys. 16 dokonano dekompozycji szeregu czasowego w programie Demetra przy pomocy metody TramoSeats. Demetra wykryła cztery obserwacje odstające w miesiącach: 12-2008, 02-2009, 04-2020 oraz 07-2020. Obserwacje odstające mają charakter krótkotrwały. Pierwsze dwie z nich związane są z kryzysem finansowym w 2008 roku. Nagły wzrost liczby mieszkań oddanych do użytkowania mógł wynikać z obaw deweloperów przed widmem obniżonego popytu na mieszkania i chęcią jak najszybszej sprzedaży. Widoczny jest trend rosnący, który przejściowo został zaburzony przez kryzys 2008 roku. W wyniku kryzysu finansowego wiele osób straciło pracę lub miało trudności w spłacie kredytów hipotecznych. W rezultacie banki zaczęły doświadczać wysokiego wskaźnika niespłacalności kredytów i znacznie ograniczyły udzielanie nowych pożyczek. To spowodowało spadek popytu na nieruchomości, ponieważ mniej osób było w stanie otrzymać finansowanie na zakup domu.

Jednocześnie wystąpiła nadmierna podaż, ponieważ ludzie nie byli w stanie utrzymać swoich mieszkań. Doprowadziło to do zmniejszonego zapotrzebowania na mieszkania, a w rezultacie wystąpił przejściowy trend malejący oddanych mieszkań do użytkowania. Obserwacje odstające z 2020 roku prawdopodobnie związane są z pandemią COVID-19. W kwietniu 2020, zaraz po wybuchu pandemii w Europie, wystąpił niewielki dołek, a następnie w lipcu nadwyżka mieszkań oddanych do użytkowania. Być może na początku pandemii niepewna sytuacja wpłynęła na ograniczone możliwości pracy, a tym samym trudności z domykaniem projektów. W lipcu mogła nastąpić odłożona w czasie podaż.

Rys. 16. Dekompozycja szeregu czasowego w programie Demetra

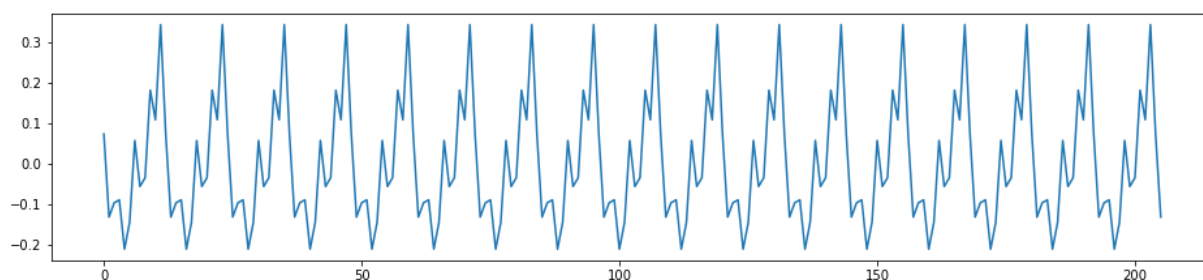


Źródło: Opracowanie własne na podstawie danych z Eurostatu

Sezonowość wydaje się być deterministyczna i addytywna, ponieważ wahania są stałe w czasie. Składnik sezonowy przedstawiony na Rys. 17 osiąga swoje minimum w maju, a maksimum w grudniu. Jest to logiczne, ponieważ deweloperom zależy, aby budowa skończyła się przed końcem roku zanim spadnie śnieg. Na wiosnę z kolei budowy raczej się rozpoczynają niż kończą, stąd pojawia się spadek oddanych mieszkań do użytkowania.

Demetra zdecydowała się na zlogarytmowanie szeregu, a więc na potrzeby dalszej analizy również przekształciłem dane do postaci logarytmicznej.

Rys. 17. Wahania składnika sezonowego

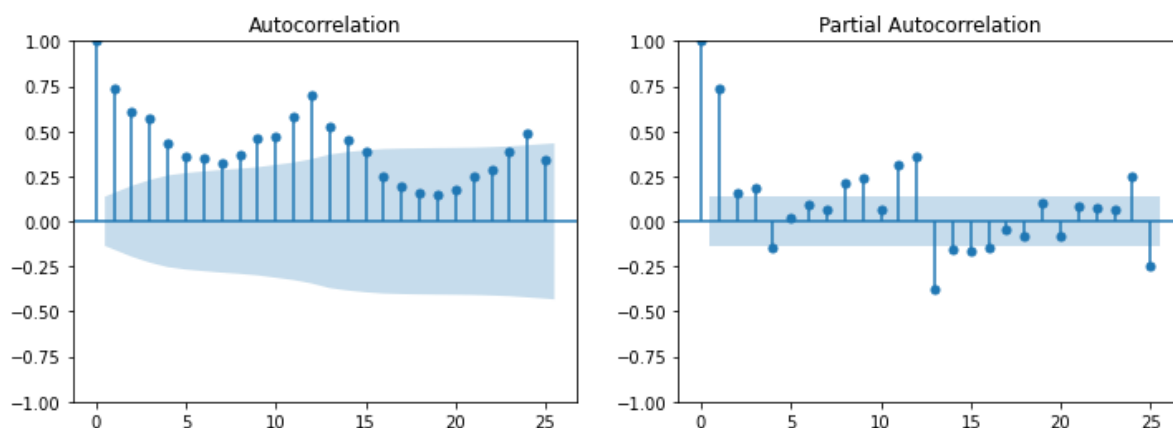


Źródło: Opracowanie własne na podstawie danych z Eurostatu

### 2.3. Model SARIMA

Na potrzeby modelowania wydzielono out-of-sample, które jest równe 12 ostatnim obserwacjom, czyli jednemu okresowi sezonowemu. Aby przeprowadzić estymację dla modelu SARIMA wymagana jest stacjonarność szeregu. W podstawowej wersji modelu występuje silna autokorelacja, która jest zależna od składnika sezonowego (Rys. 18). Wykonany rozszerzony test Dickeya-Fullera dla 15 rozszerzeń wskazuje na niestacjonarność szeregu (Rys 19).

Rys. 18. Funkcje ACF i PACF dla szeregu bazowego



Źródło: Opracowanie własne z wykorzystaniem języka Python

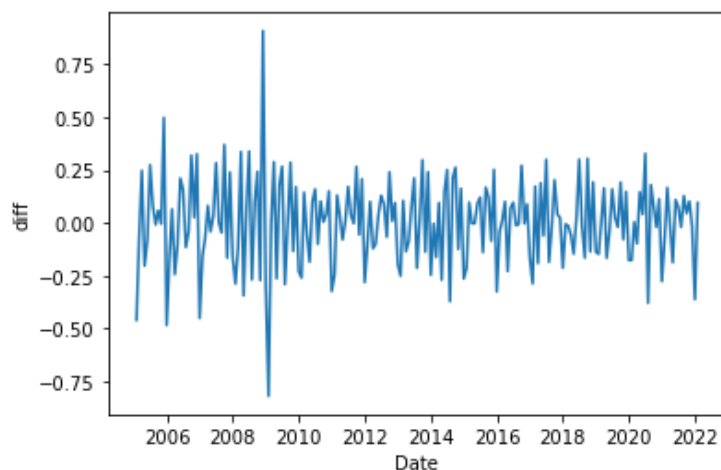
Rys. 19. Rozszerzony test Dickeya-Fullera dla szeregu bazowego

```
1. ADF : -0.8292582958887716
2. P-Value : 0.8103762471341615
3. Num Of Lags : 15
4. Num Of Observations Used For ADF Regression and Critical Values Calculation : 190
5. Critical Values :
    1% : -3.4652439354133255
    5% : -2.8768752281673717
    10% : -2.574944653739612
Brak podstaw do odrzucenia H0, szereg jest niestacjonarny
```

Źródło: Opracowanie własne z wykorzystaniem języka Python

W celu usunięcia trendu podjęto decyzję o zastosowaniu różnicowania 1 stopnia. Rys. 20 przedstawia szereg po zróżnicowaniu. Po pierwszym różnicowaniu szereg dalej wydaje się być stacjonarny. Stacjonarność potwierdza test ADF, którego statystyka wyniosła -3,88 dla 14 rozszerzeń (Rys. 21).

Rys. 20. Pierwsze różnice szeregu



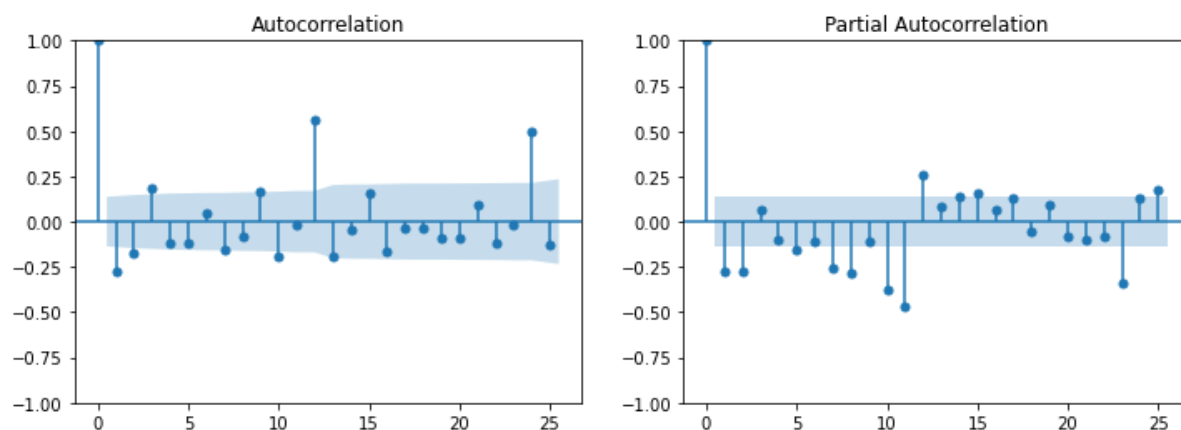
Źródło: Opracowanie własne z wykorzystaniem języka Python

Rys. 21. Rozszerzony test Dickeya-Fullera dla szeregu zróżnicowanego jednokrotnie

```
1. ADF : -3.879916192529138
2. P-Value : 0.002187726539798264
3. Num Of Lags : 14
4. Num Of Observations Used For ADF Regression and Critical Values Calculation : 190
5. Critical Values :
   1% : -3.4652439354133255
   5% : -2.8768752281673717
  10% : -2.574944653739612
Odrzucamy H0, szereg jest stacjonarny
```

Źródło: Opracowanie własne z wykorzystaniem języka Python

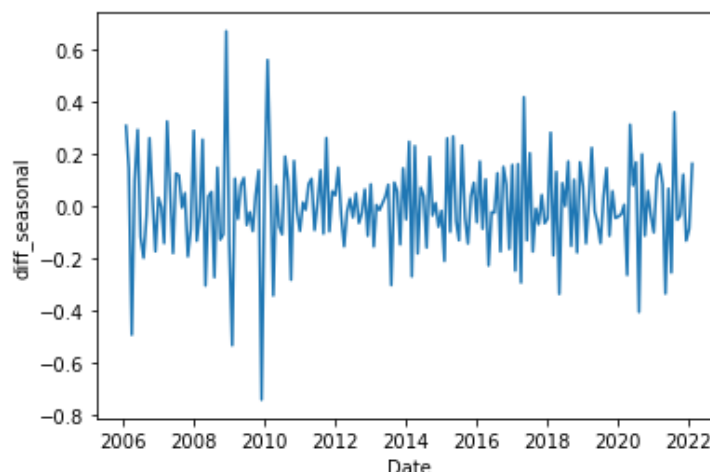
Rys. 22. Funkcje ACF i PACF dla szeregu zróżnicowanego



Źródło: Opracowanie własne z wykorzystaniem języka Python

Dokonano dodatkowo analizy korelogramów po pierwszym zróżnicowaniu (Rys. 22). Z wykresów wynika, że w szeregu dalej występuje silna sezonowość, ponieważ wielokrotności 12-tego opóźnienia, które powoli wygasają, są istotnie dodatnie. Podjęto decyzję o zastosowaniu różnicowania sezonowego, aby pozbyć się wahań sezonowych. Wykres zróżnicowanej zmiennej po odsezonowaniu znajduje się na Rys. 23.

Rys. 23. Pierwsze różnice odsezonowanego szeregu



Źródło: Opracowanie własne z wykorzystaniem języka Python

Rys. 24. Rozszerzony test Dickeya-Fullera dla zróżnicowanego odsezonowanego szeregu

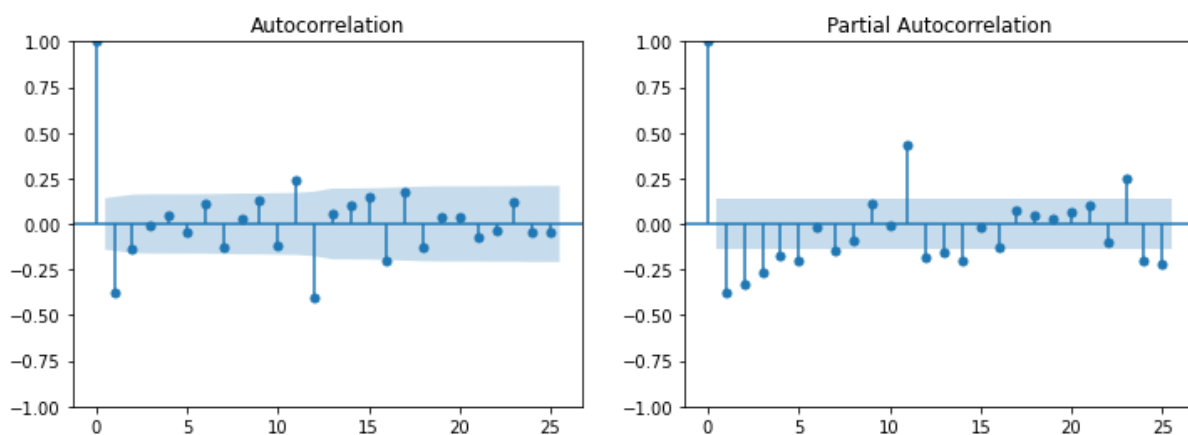
```
1. ADF : -4.756111235660604
2. P-Value : 6.593493882573579e-05
3. Num Of Lags : 13
4. Num Of Observations Used For ADF Regression and Critical Values Calculation : 179
5. Critical Values :
   1% : -3.4674201432469816
   5% : -2.877826051844538
  10% : -2.575452082332012
Odrzucamy H0, szereg jest stacjonarny
```

Źródło: Opracowanie własne z wykorzystaniem języka Python

Szereg w obecnej formie jest stacjonarny (Rys. 24) oraz nie występuje w nim sezonowość. Oznacza to, że model powinien przyjąć postać SARIMA (p, 1, q) (P, 1, Q, 12). Analizując korelogramy dla zróżnicowanego szeregu po odsezonowaniu (Rys. 25) widać wygasające wielokrotności 12 opóźnień dla funkcji PACF oraz jednorazowy wzrost dla 12 opóźnień funkcji ACF, a więc proces SMA powinien być równy 1. Dodatkowo wartości niesezonowe PACF zbiegają od dołu do zera, a ACF przyjmuje niską ujemną wartość tylko dla 1 opóźnienia, czyli proces MA również powinien być równy 1. Oznacza to, że poprawna

forma modelu SARIMA przyjmuje postać  $(0, 1, 1)(0, 1, 1)$ . Taką formę również zaproponowała Demetra jako najlepszą.

Rys. 25. Funkcje ACF i PACF dla odsezonowanego szeregu zróżnicowanego



Źródło: Opracowanie własne z wykorzystaniem języka Python

Zdecydowałem się dodatkowo na stworzenie rankingu najlepszych modeli, dla których wartości  $p$ ,  $d$ ,  $P$ ,  $Q$  są mniejsze bądź równe 2. W Tabeli 4 znajduje się 5 najlepszych modeli pod kątem wartości AIC. Wybrany model znalazł się na 4 miejscu, aczkolwiek pod kątem kryterium BIC wypadł najlepiej. Sprawdziłem każdy z wymienionych modeli i tylko model SARIMA(0,1,1)(0,1,1) dobrany manualnie charakteryzuje się istotnym procesem. Finalnie tenże model zostaje uznany za najlepszy.

Tab. 4. Wartości kryteriów informacyjnych AIC oraz BIC najlepszych modeli

Lp.	p	d	q	P	D	Q	AIC	BIC
1	2	1	2	2	1	2	270,612	-241.248
2	2	1	1	0	1	1	268,269	-251.955
3	2	1	1	0	1	2	266,802	-247.226
4	0	1	1	0	1	1	266,65	-256.862
5	2	1	1	2	1	1	266,559	-243.720

Źródło: Opracowanie własne z wykorzystaniem języka Python

Sprawdzono czy reszty modelu SARIMA(0,1,1)(0,1,1) są białym szumem. Przeprowadzono test Ljunga-Boxa oraz Boxa-Pierce'a dla 24 opóźnień (Rys. 26). Tylko test Boxa-Pierce'a ujawnił  $p$ -value większe niż 0.05, a więc w resztach może występować pewna autokorelacja.

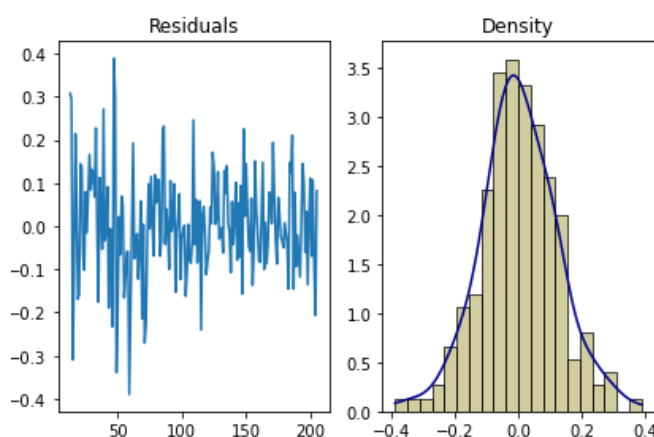
Rys. 26. Wyniki testu Ljunga-Boxa oraz Boxa-Pierce'a reszt modelu SARIMA(0,1,1)(0,1,1)

	lb_stat	lb_pvalue	bp_stat	bp_pvalue
24	38.71412	0.029255	35.973681	0.055211

Źródło: Opracowanie własne z wykorzystaniem języka Python

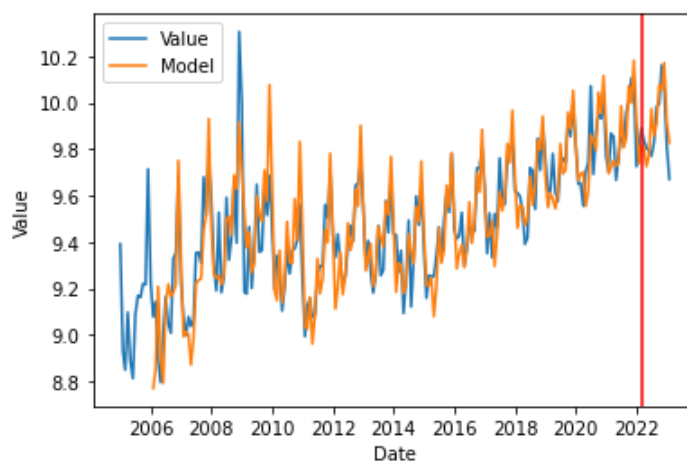
Dodatkowo przedstawiono reszty na wykresie rozrzutu oraz ich histogram (Rys. 27). Rozkład reszt przypomina normalny, aczkolwiek pierwsze kilkadziesiąt reszt charakteryzuje się wyższą wariancją. W obrębie obserwacji odstającej reszty zaczęły częściej przyjmować ujemne wartości. Ze względu, że jest to zjawisko przejściowe wynikające z obserwacji nietypowej model zostanie wykorzystany do prognozowania. Wykonano prognozę dla modelu SARIMA(0,1,1)(0,1,1), a jej wyniki przedstawiono na Rys. 28. Model poprawnie przewidział komponent sezonowy, a predykcja wydaje się całkiem dobra.

Rys 27. Wykres rozrzutu oraz histogram reszt modelu SARIMA(0,1,1)(0,1,1)



Źródło: Opracowanie własne z wykorzystaniem języka Python

Rys. 28. Wyniki prognozy dla modelu SARIMA(0,1,1)(0,1,1)



Źródło: Opracowanie własne z wykorzystaniem języka Python



## 2.4. Model ekstrapolacyjny

W tej sekcji użyjemy modelu Holta-Wintersa do utworzenia prognozy dla badanego szeregu czasowego. W poniższym modelu próba out-of-sample również wyniosła 12 obserwacji. Utworzono cztery modele Holta-Wintersa, w których parametry szacowane są iteracyjnie. Modele uwzględniają wszystkie możliwe kombinacje sezonowości i trendu, czyli postaci addytywne i multiplikatywne. W Tabeli 5 przedstawione zostało porównanie modeli. Na zielono oznaczono najniższe wartości statystyk porównawczych.

Tab. 5. Porównanie modeli Holta uwzględniających różne rodzaje trendu

	trend addytywny i sezonowość addytywna	trend addytywny i sezonowość multiplikatywna	trend multiplikatywny i sezonowość addytywna	trend multiplikatywny i sezonowość multiplikatywna
Alpha	0.2265641	0.2287080	0.2266675	0.2296789
Beta	0.000	0.000	0.000	0.000
Gamma	0.1562288	0.1767246	0.1563808	0.1766379
AIC	-873.889	-872.083	-873.830	-872.020
BIC	-820.643	-818.837	-820.584	-818.774
RMSE (out-of-sample)	0.103567	0.101797	0.103849	0.102046
MAE (out-of-sample)	0.091271	0.089932	0.091332	0.090054
MAPE (out-of-sample)	0.009235	0.009100	0.009241	0.009112

Źródło: Opracowanie własne z wykorzystaniem języka Python

Modele z trendem addytywnym są wyraźnie lepsze. Model z trendem addytywnym i sezonowością addytywną charakteryzuje się najniższymi wartościami AIC oraz BIC, a model z trendem addytywnym i sezonowością multiplikatywną najniższymi błędami prognozy. Celem modelu jest prognozowanie, a więc za najbardziej odpowiedni wybrano model drugi z najniższymi wartościami wszystkich błędów prognozy.

Sprawdzono model Holta-Wintersa z trendem addytywnym i sezonowością multiplikatywną pod kątem autokorelacji reszt. Zarówno test Ljunga-Boxa oraz Boxa-Pierce'a ujawnił p-value większe niż 0.05, a więc reszty zgodnie z założeniem formalnym modelu przyjmują wartości losowe.

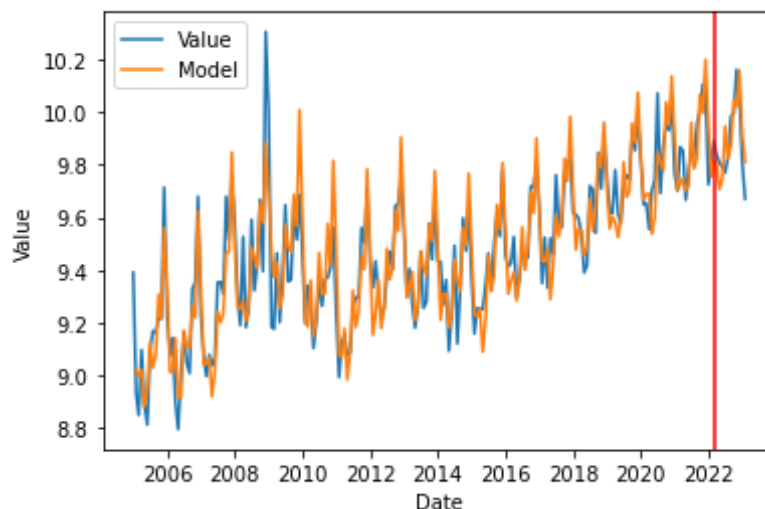
Rys. 29. Wyniki testu Ljunga-Boxa oraz Boxa-Pierce'a reszt modelu Holta-Wintersa

	lb_stat	lb_pvalue	bp_stat	bp_pvalue
24	33.480664	0.094331	30.941971	0.155463

Źródło: Opracowanie własne z wykorzystaniem języka Python

Przeprowadzono prognozę dla tego modelu, której wyniki widoczne są na Rys. 30. Prognoza wydaje się dobra i bardzo podobna do prognozy modelu SARIMA(0,1,1)(0,1,1).

Rys. 30. Wyniki prognozy dla modelu Holta-Wintersa

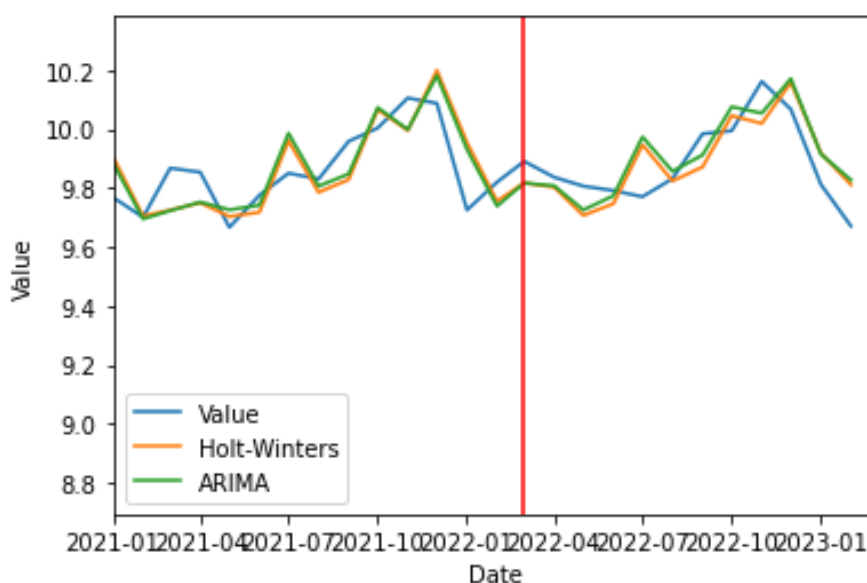


Źródło: Opracowanie własne z wykorzystaniem języka Python

## 2.5. Porównanie modeli

Porównano model SARIMA(0,1,1)(0,1,1) z modelem Holta-Wintersa uwzględniającym trend addytywny i sezonowość mutliplikatywną. Obydwa modele idą podobnym torem, aczkolwiek model ARIMA wydaje się być bliższy wartościom rzeczywistym.

Rys. 31. Porównanie modelu SARIMA(0,1,1)(0,1,1) oraz Holta-Wintersa



Źródło: Opracowanie własne z wykorzystaniem języka Python

Porównano modele przy pomocy wartości średnich błędów dopasowania. Wyniki przedstawiono w Tabeli 6. Model SARIMA okazał się być istotnie lepszy od modelu Holta-Wintersa do prognozowania, ponieważ wartości wszystkich jego błędów w out-of-sample są niższe. Model Holta-Wintersa z kolei charakteryzuje się lepszym dopasowaniem do danych w próbie in-sample.

Tab. 6. Porównanie modelu SARIMA(0,1,1)(0,1,1) oraz Holta-Wintersa

	SARIMA	Holt-Winters
RMSE (out-of-sample)	0.101715	0.101797
MAE (out-of-sample)	0.087743	0.089932
MAPE (out-of-sample)	0.008887	0.009100
RMSE (in-sample)	0.122597	0.111401
MAE (in-sample)	0.093903	0.087159
MAPE (in-sample)	0.009908	0.009187

Źródło: Opracowanie własne z wykorzystaniem języka Python

## ZAKOŃCZENIE

Celem pracy było porównanie jakości prognoz modeli z klasy ARIMA oraz modeli ekstrapolacyjnych. Uzyskane wyniki nie są jednoznaczne. W przypadku szeregu niesezonowego model liniowy Holta okazał się być lepszy od modelu z klasy ARIMA, a w przypadku danych sezonowych prognoza SARIMy była dokładniejsza.