Uniwersytet Warszawski

Wydział Nauk Ekonomicznych

Monika Zimolzak

# Analiza szeregów czasowych

Warszawa, lipiec 2023

**Spis treści**

[WSTĘP 3](#_Toc138024153)

[Szereg niesezonowy 4](#_Toc138024154)

[1.1. Opis szeregu 4](#_Toc138024155)

[1.2. Dekompozycja szeregu w programie Demetra 4](#_Toc138024156)

[1.3. Stacjonarność 5](#_Toc138024157)

[1.4. ARIMA 9](#_Toc138024158)

[1.5. Model ekstrapolacyjny 10](#_Toc138024159)

[Szereg sezonowy 12](#_Toc138024160)

[2.1. Opis szeregu 12](#_Toc138024161)

[2.2. Dekompozycja szeregu w programie Demetra 13](#_Toc138024162)

[2.3. Przygotowanie danych 14](#_Toc138024163)

[2.4. Stacjonarność 15](#_Toc138024164)

[2.5. Model SARIMA 21](#_Toc138024165)

[2.6. Model ekstrapolacyjny 25](#_Toc138024166)

[2.7. Porównanie modeli 26](#_Toc138024167)

[PODSUMOWANIE 27](#_Toc138024168)

## WSTĘP

W poniższej pracy przedstawiono wyniki analizy dwóch szeregów czasowych – niesezonowego jakim jest kurs bitcoina w stosunku do dolara amerykańskiego BTC/USD,   
oraz sezonowego, który dotyczy liczby nocy spędzonych przez turystów w różnego rodzaju obiektach noclegowych w Portugalii. Każdy z szeregów ma charakter miesięczny na przestrzeni lat 2015-2023.

Dla wskazanych szeregów zostanie dokonana dekompozycja, dopasowanie modelu z klasy ARIMA/SARIMA, przeprowadzenie modelu ekstrapolacyjnego oraz porównanie jakości prognoz modeli.

# Szereg niezezonowy

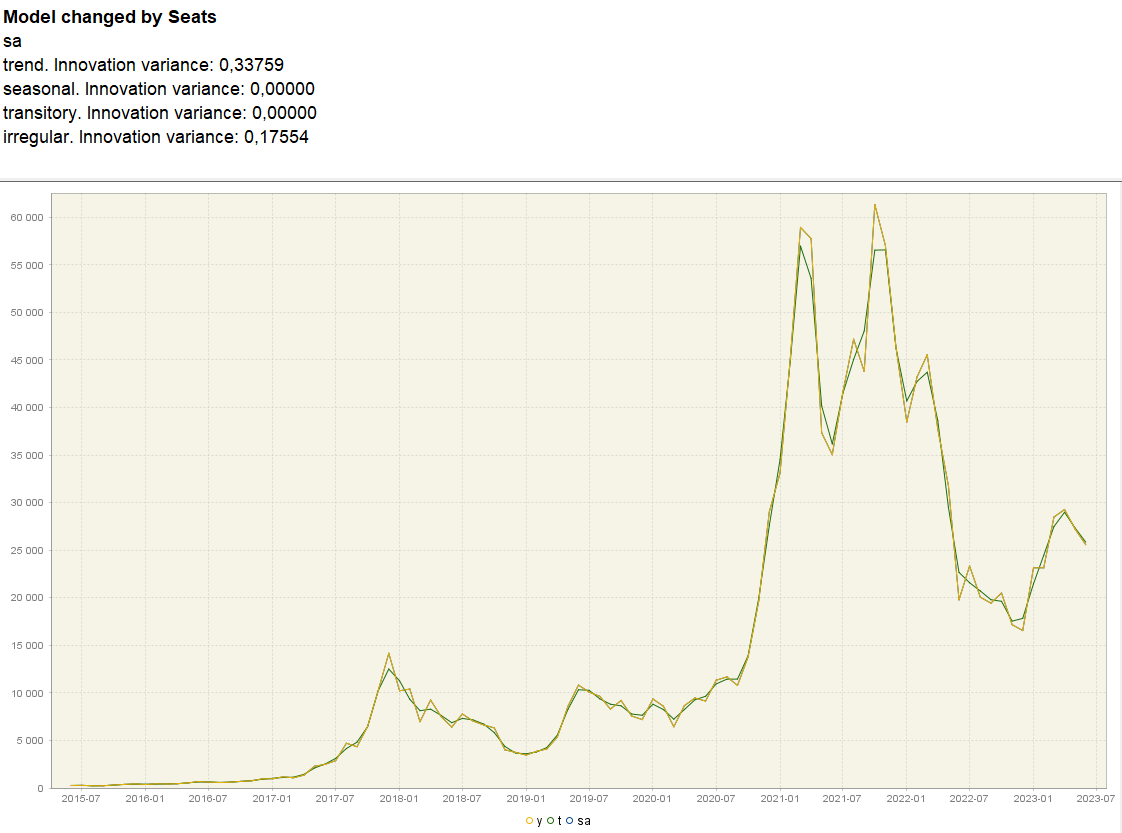
## Opis szeregu

Szereg danych niesezonowych wybrany do tej pracy to kurs bitcoina (BTC) w stosunku do dolara amerykańskiego (USD). Dane te obejmują wartości zamknięcia dla każdego pierwszego dnia miesiąca, począwszy od czerwca 2015 roku do czerwca 2023 (97 obserwacji). Zbiór danych pochodzi z baz danych Yahoo Finance i jest publicznie dostępny pod linkiem: <https://finance.yahoo.com/quote/BTC-USD/history?period1=1433116800&period2=1686873600&interval=1mo&filter=history&frequency=1mo&includeAdjustedClose=true>

## Dekompozycja szeregu w programie Demetra

Analizę szeregu rozpoczęto od programu Demetra, w którym dokonano dekompozycji szeregu za pomocą metody TRAMO/Seats. Proces ten potwierdził brak sezonowości oraz wyraźnego trendu – szereg odsezonowany oraz linia trendu nie odbiega od pierwotnych wartości szeregu. Szereg nie charakteryzuje się więc trendem, który mógłby zostać jednoznacznie zakwalifikowany jako addytywny lub multiplikatywny.

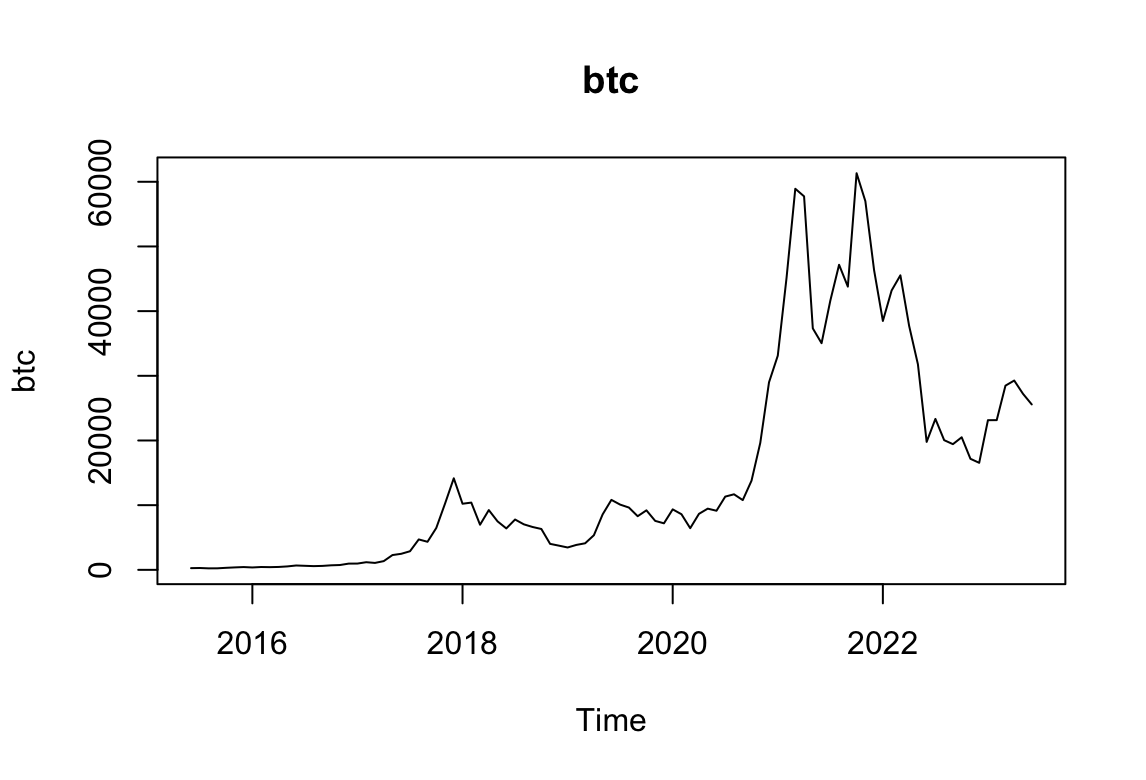
Wykres 1. Dekompozycja szeregu w programie Demetra

****

## Stacjonarność

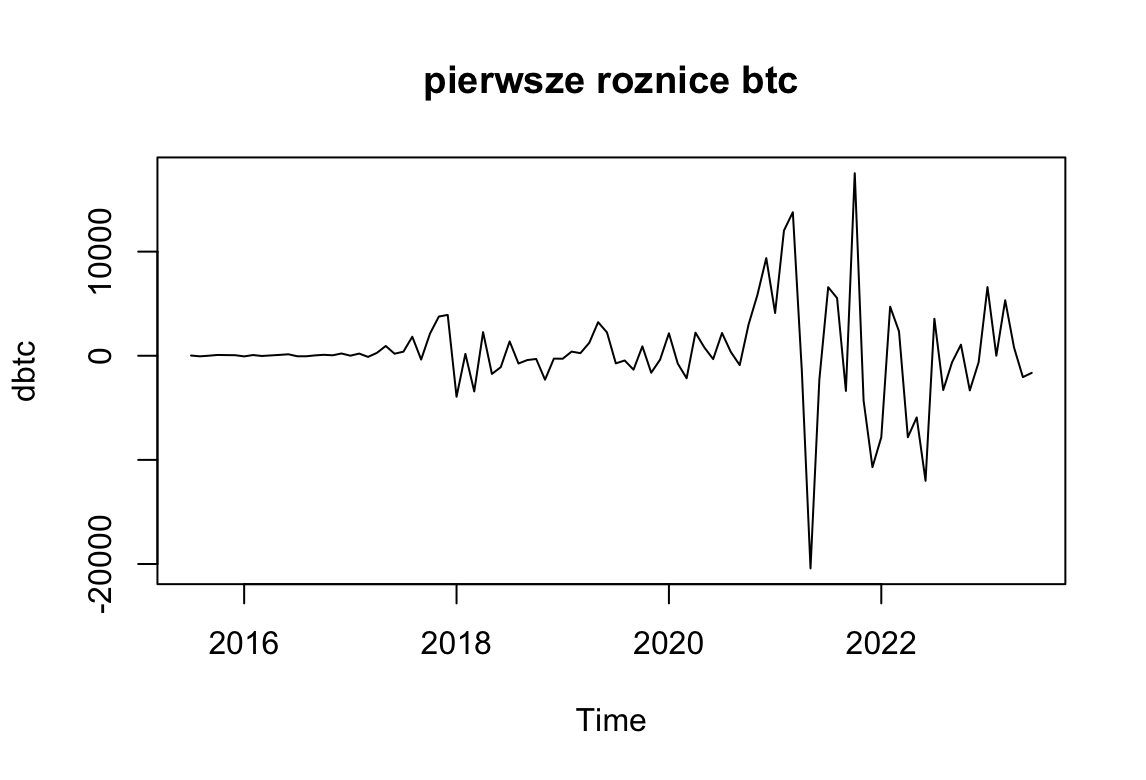
Rozpoczęto analizę szeregu czasowego, który prezentuje się następująco:

Wykres 2. Kurs BTC/USD – miesięczne wartości zamknięcia



Pierwszym krokiem było obliczenie pierwszych różnic w celu weryfikacji ewentualnego trendu oraz stacjonarności szeregu, co zostało zobrazowane na poniższym wykresie.

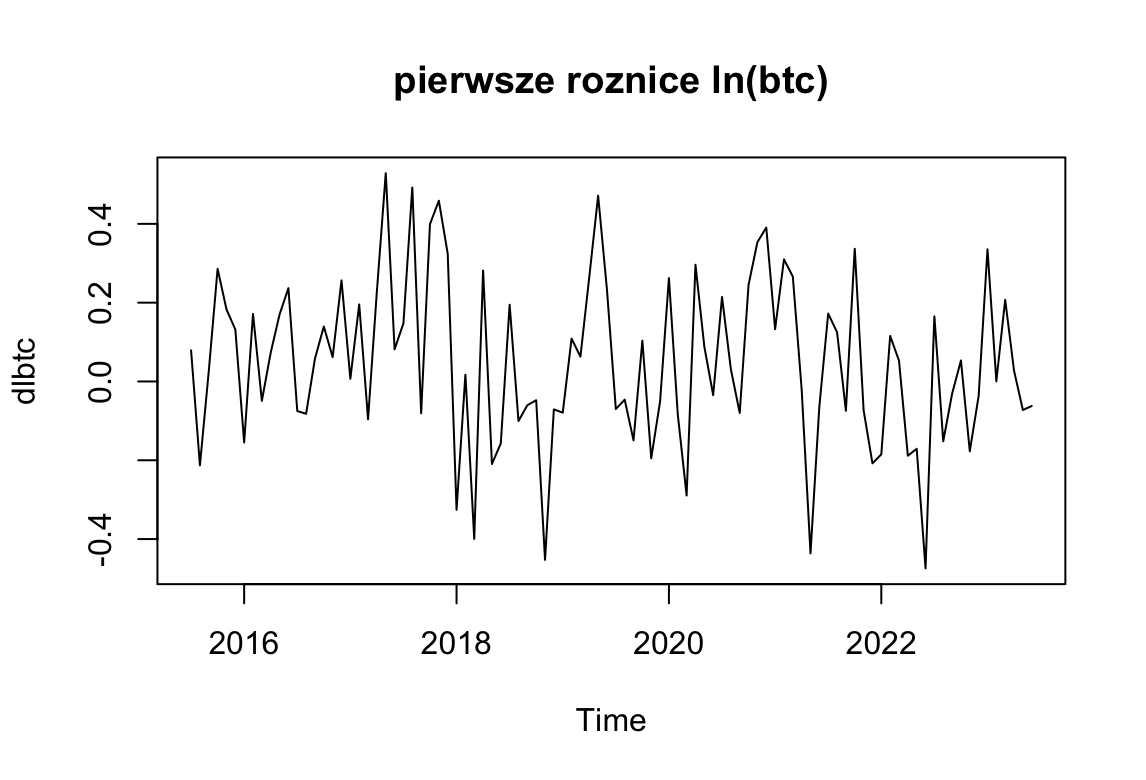
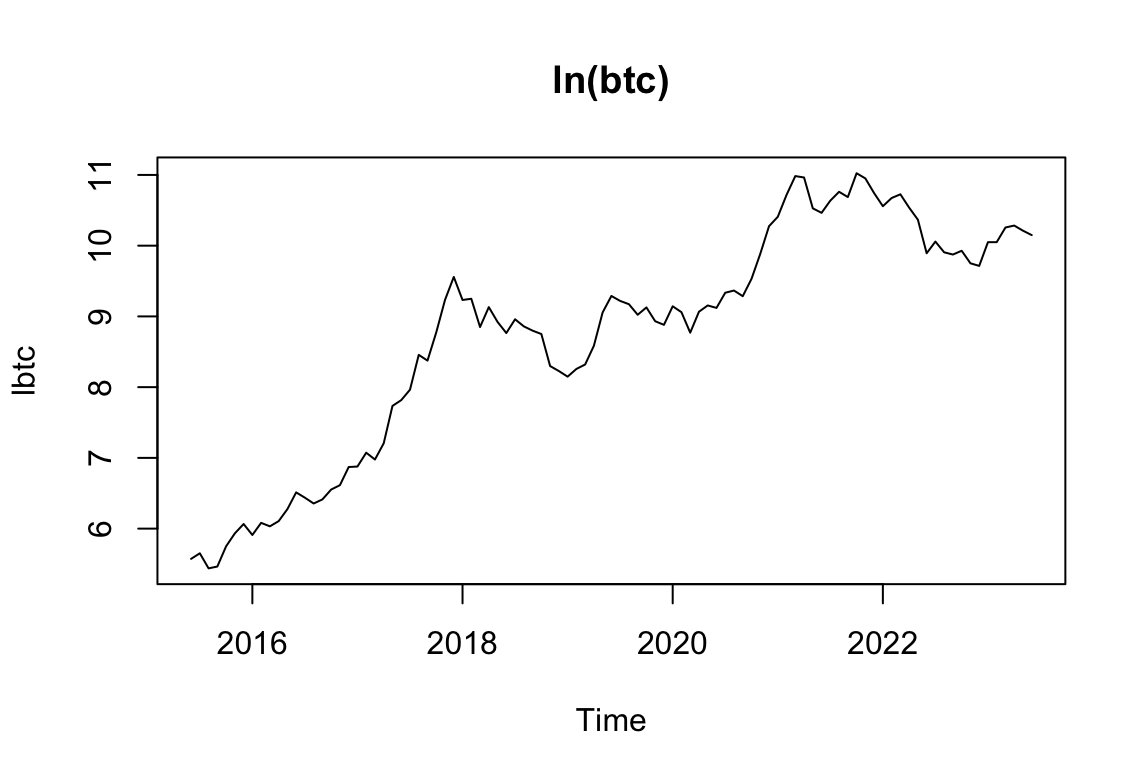
Wykres 3. Pierwsze różnice szeregu



Wykres dla danych podstawowych BTC wskazuje znaczne różnice między kolejnymi obserwacjami oraz odznacza się nieliniowym trendem, w związku z czym, podjęto decyzje o zastosowaniu transformacji logarytmicznej. Logarytmowanie może pomóc w przekształceniu niestabilności wariancji i nieliniowych trendów, co ułatwi analizę i modelowanie szeregu czasowego. Dodatkowo, logarytmowanie może pomóc w zmniejszeniu skali wartości.

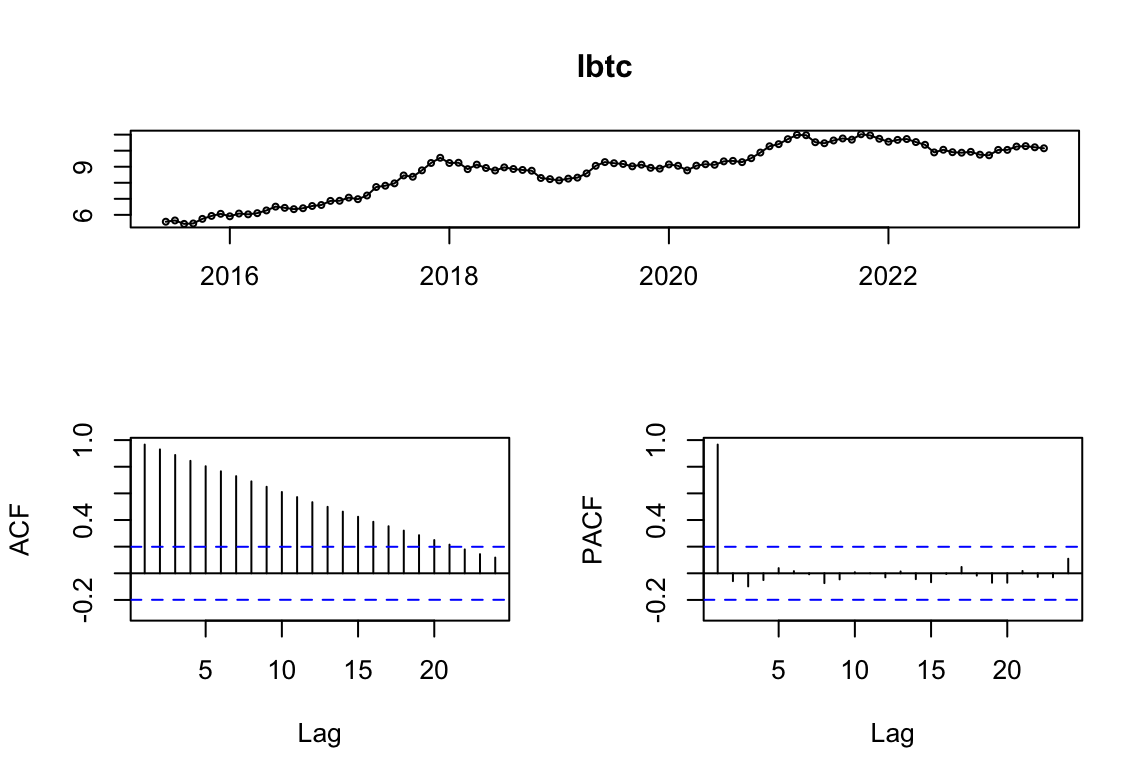
Poniższe wykresy wizualizują zmiany w danych po zastosowaniu transformacji logarytmicznej.

Wykres 4 i 5. Transformacja logarytmiczna szeregu oraz pierwsze różnice transformacji



Następnie przystąpiono do analizy korelogramów ACF i PACF dla szeregu lbtc.

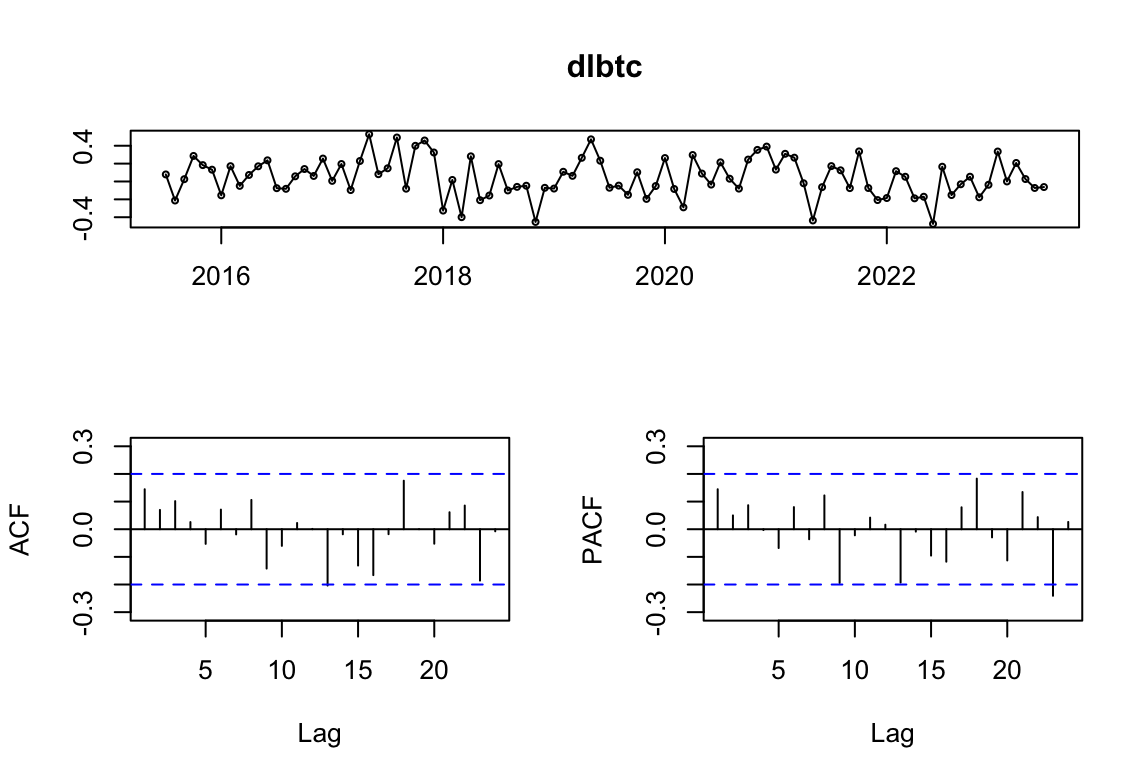
Wykres 6. Wykresy ACF i PACF dla logarytmu szeregu



Wartości ACF są bardzo wysokie dla opóźnień bliskich 0 i szybko maleją do zera, co wskazuje na obecność składowej autokorelacji w szeregu. Z kolei dla wykresu PACF pierwsza wartość jest znacznie większa od zera, a kolejne wartości szybko maleją i stają się nieistotne, może to wskazywać na istnienie autokorelacji jedynie w pierwszym opóźnieniu.

Weryfikacja ta nie dostarczyła więc wystarczająco dobrych wyników, dlatego zweryfikowano korelogramy dla pierwszych różnic.

Wykres 7. Wykresy ACF i PACF dla zróżnicowanego szeregu



Analiza ACF i PACF dla pierwszych różnic szeregu czasowego nie wykazuje żadnych wartości wykraczających poza obszar istotności, oznacza to, że nie występują znaczące zależności autokorelacyjne w danych po uwzględnieniu pierwszych różnic. Taki wynik może sugerować, że pierwsze różnicowanie szeregu czasowego było wystarczające do uzyskania stacjonarności i usunięcia występujących wcześniej zależności autokorelacyjnych.

W celu potwierdzenia wyników wyżej wykonanych analiz korelogramów, przeprowadzono testy stacjonarności dla danego szeregu czasowego - test ADF (Augmented Dickeya-Fullera) oraz test Phillipsa-Perrona dla różnych stopni różniczkowania, aby ocenić stacjonarność szeregu. Poniżej przedstawione zostały wyniku testowania:

augmentations adf p\_adf p\_bg.p.val.1 p\_bg.p.val.2 p\_bg.p.val.3 p\_bg.p.val.4 p\_bg.p.val.5

1 0 -1.898692 0.3525547 0.1864378 0.3925654 0.4974136 0.6629761 0.7057946

2 1 -1.829374 0.3781903 0.9181911 0.9628901 0.8653074 0.9457922 0.9230986

3 2 -2.132788 0.2659808 0.9037831 0.9927168 0.8528385 0.9389743 0.8704067

4 3 -2.156183 0.2573287 0.9692940 0.9800721 0.9978411 0.9997896 0.9374427

5 4 -1.912191 0.3475626 0.9850044 0.9965143 0.9990453 0.9999292 0.9479895

6 5 -1.806164 0.3867737 0.9561117 0.9982325 0.9990030 0.9992794 0.9982357

7 6 -1.751298 0.4070644 0.9916068 0.9966404 0.9887986 0.9974204 0.9980147

8 7 -2.070564 0.2889928 0.9224957 0.9339383 0.9815787 0.9957113 0.9991682

Wyniki dla lbtc potwierdzają analizę ACF i PACF. Dla różnych stopni różniczkowania nie ma wystarczających dowodów, aby odrzucić hipotezę o niestacjonarności szeregu czasowego. Wartości p są zwykle wyższe od poziomu istotności (0.05), co sugeruje, że szereg czasowy może być niestacjonarny. W przypadku testu ADF, żadna z wartości p nie jest niższa niż poziom istotności, a dla testu Phillipsa-Perrona również większość wartości p jest wysoka.

W związku z powyższym powtórzono te testy dla pierwszych różnic dlbtc:

augmentations adf p\_adf p\_bg.p.val.1 p\_bg.p.val.2 p\_bg.p.val.3 p\_bg.p.val.4 p\_bg.p.val.5

1 0 -8.322017 0.01000000 0.9585722 0.9364280 0.8207389 0.9169893 0.9106116

2 1 -5.945849 0.01000000 0.9710391 0.9987717 0.8348926 0.9264959 0.8944482

3 2 -4.557509 0.01000000 0.9978151 0.9776508 0.9973852 0.9996522 0.9543421

4 3 -4.041164 0.01000000 0.9364056 0.9965467 0.9972175 0.9996952 0.9630189

5 4 -3.943620 0.01000000 0.9717204 0.9990182 0.9993258 0.9997843 0.9993443

6 5 -3.312819 0.01895531 0.9458797 0.9957242 0.9868645 0.9970221 0.9985558

7 6 -3.183838 0.02461239 0.8810037 0.9550730 0.9838001 0.9963830 0.9993266

8 7 -2.591284 0.09851213 0.7634946 0.9437863 0.9870779 0.9973897 0.9984648

Wyniki wskazują, że dla różnych stopni różniczkowania szeregu czasowego istnieją silne dowody przemawiające za stacjonarnością szeregu.

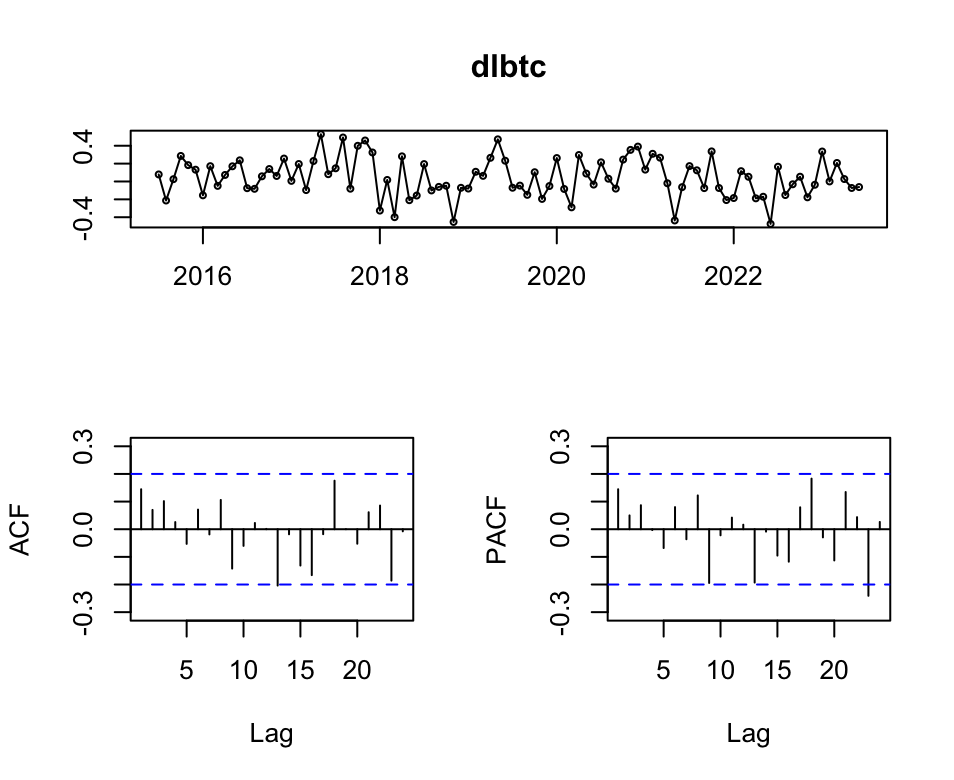
Dodatkowo zostało to zweryfikowane testem KPSS. Na podstawie wyników można stwierdzić, że wartość statystyki testowej (0.1979) jest mniejsza od wszystkich krytycznych wartości (dla istotności na poziomie 10%, krytyczna wartość wynosi 0.347, dla istotności na poziomie 5% - 0.463, dla istotności na poziomie 2.5% - 0.574, a dla istotności na poziomie 1% - 0.739). Oznacza to, że na każdym poziomie istotności można odrzucić hipotezę o niestacjonarności szeregu czasowego. Wskazuje to na silne dowody na stacjonarność szeregu.

## ARIMA

Następnie zweryfikowano czy dany szereg jest białym szumem przy wykorzystaniu testu Ljunga-Boxa oraz testu Boxa-Pearsa. Na podstawie wyników testu Boxa-Ljunga   
(p-value = 0.247) oraz Boxa-Pearce'a (p-value = 0.9547) dla szeregu czasowego dlbtc, możemy wnioskować, że szereg ten wykazuje cechy białego szumu. Brak istotnych autokorelacji w danych wskazuje na brak systematycznego wzorca czy zależności między kolejnymi obserwacjami. Otrzymane wartości p-value są wysokie, co sugeruje, że nie ma wystarczających dowodów na odrzucenie hipotezy zerowej o braku autokorelacji w szeregu czasowym.

W celu potwierdzenia, że analiza białego szumu jest najlepszą metodą dla wskazanego szeregu ponownie zweryfikowano diagramy ACF i PACF.

Wykres 8. Wykresy ACF i PACF dla zróżnicowanego szeregu



Żadne ze słupków funkcji autokorelacji (ACF) i funkcji autokorelacji częściowej (PACF) nie wychodzą poza obszar przedziału ufności, sugeruje to, że nie ma istotnych wartości autokorelacji i autokorelacji częściowej na danych opóźnieniach. Jest to dodatkowym potwierdzeniem występowania modelu białego szumu.

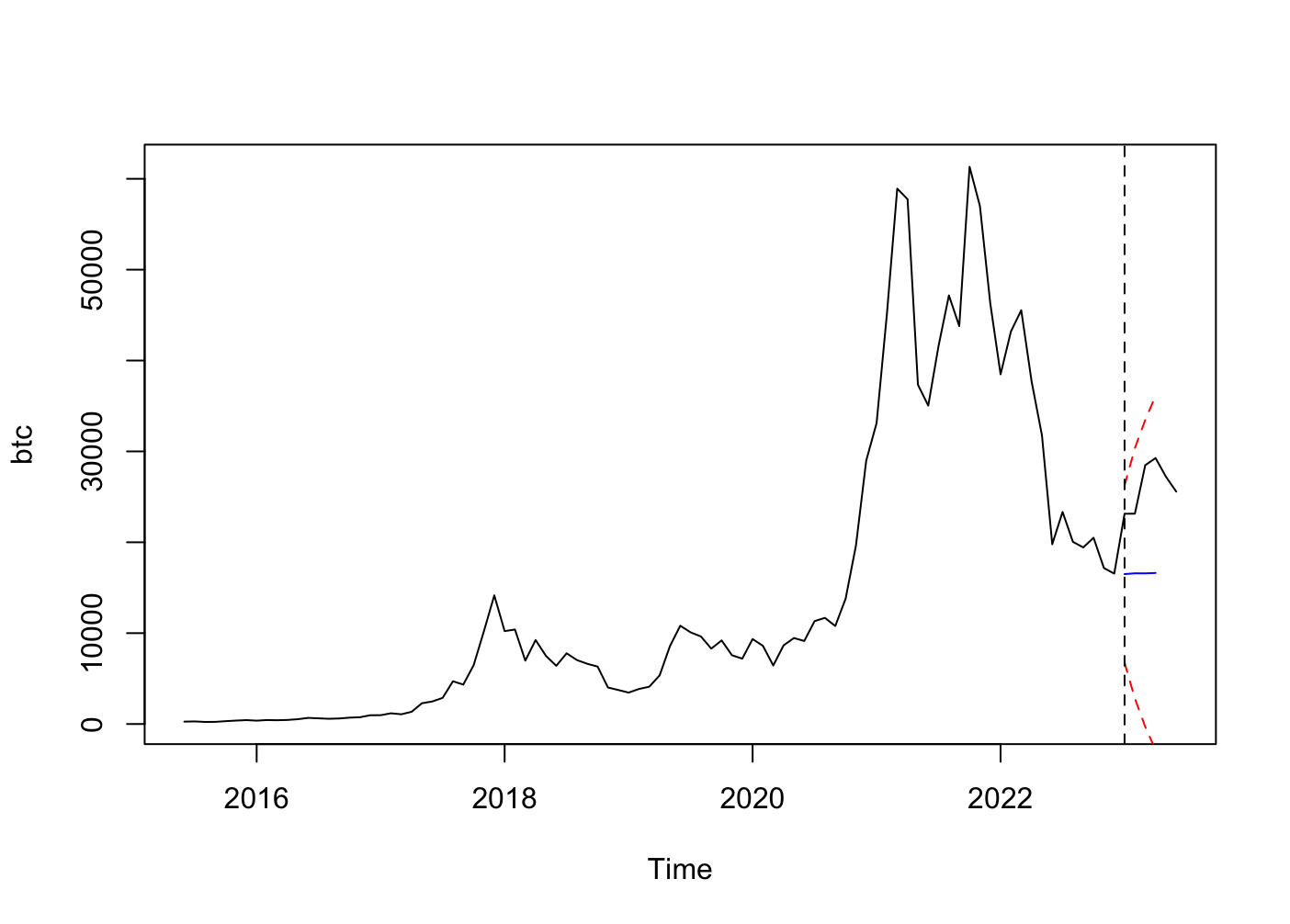
W przypadku białego szumu, próba zastosowania modelu ARIMA może nie być odpowiednia, ponieważ model ten zakłada istnienie zależności czasowych w danych.   
Przy braku tych zależności, model ARIMA może być nieefektywny lub nieodpowiedni. W związku z powyższym, na tym etapie ukończono analizę szeregu.

## Model ekstrapolacyjny

W celu przeprowadzenia prognozy modelu ekstrapolacyjnego wykorzystano metodę   
Holta-Wintersa. Jest to proste wygładzanie wykładnicze, które nie uwzględnia trendu ani efektów sezonowych. Prognozę dla tego modelu przeprowadzono dla 6 ostatnich miesięcy   
roku 2023 (6 ostatnich obserwacji).

Poniższy wykres przedstawia wyniki estymacji dla modelu Holta-Wintersa. Prognoza dla wybranego okresu nie jest zgodna z faktycznymi wartościami kursu. Przewidywania przy użyciu modelu ekstrapolacyjnego dla kursu BTC wskazują o wiele niższe wartości (niebieska linia prognozy) od faktycznego kursu ostatnich sześciu miesięcy.

Wykres 9. Prognozowanie metodą Holta-Wintersa



Następnie, w celu oceny jakości i skuteczności modelu prognozowanego wybrano miarę błędu prognozy. Błąd prognozy jest miarą różnicy między wartościami prognozowanymi a rzeczywistymi wartościami. Do oceny prognoz została wybrana miara błędu średniego bezwzględnego (MAE, Mean Absolute Error), ponieważ jest ona jedną z najpopularniejszych i najbardziej intuicyjnych miar błędu prognozy i dobrze dopasowuje się do danych (wartości w szeregu są co do modułu większe od 1). Dodatkowo MAE jest bardziej odporna na wartości odstające niż MSE (średni błąd kwadratowy), ponieważ kwadrat błędu w MSE powoduje większe znaczenie wartości odstających. MAE oblicza średnią wartość bezwzględną różnicy między prognozami a rzeczywistymi wartościami, co daje nam informację o przeciętnym błędzie prognozy bez względu na kierunek (nadprognozowanie lub podprognozowanie).

Ogólny błąd MAE wynosi 9493.961. MAE jest miarą przeciętnego bezwzględnego błędu między prognozowanymi wartościami a rzeczywistymi wartościami. Im niższa wartość MAE, tym lepsza jakość prognozy. W przypadku podanych danych, średni bezwzględny błąd wynosi 9493.961, co oznacza, że średnia różnica między prognozami a rzeczywistymi wartościami wynosi około 9493.961.

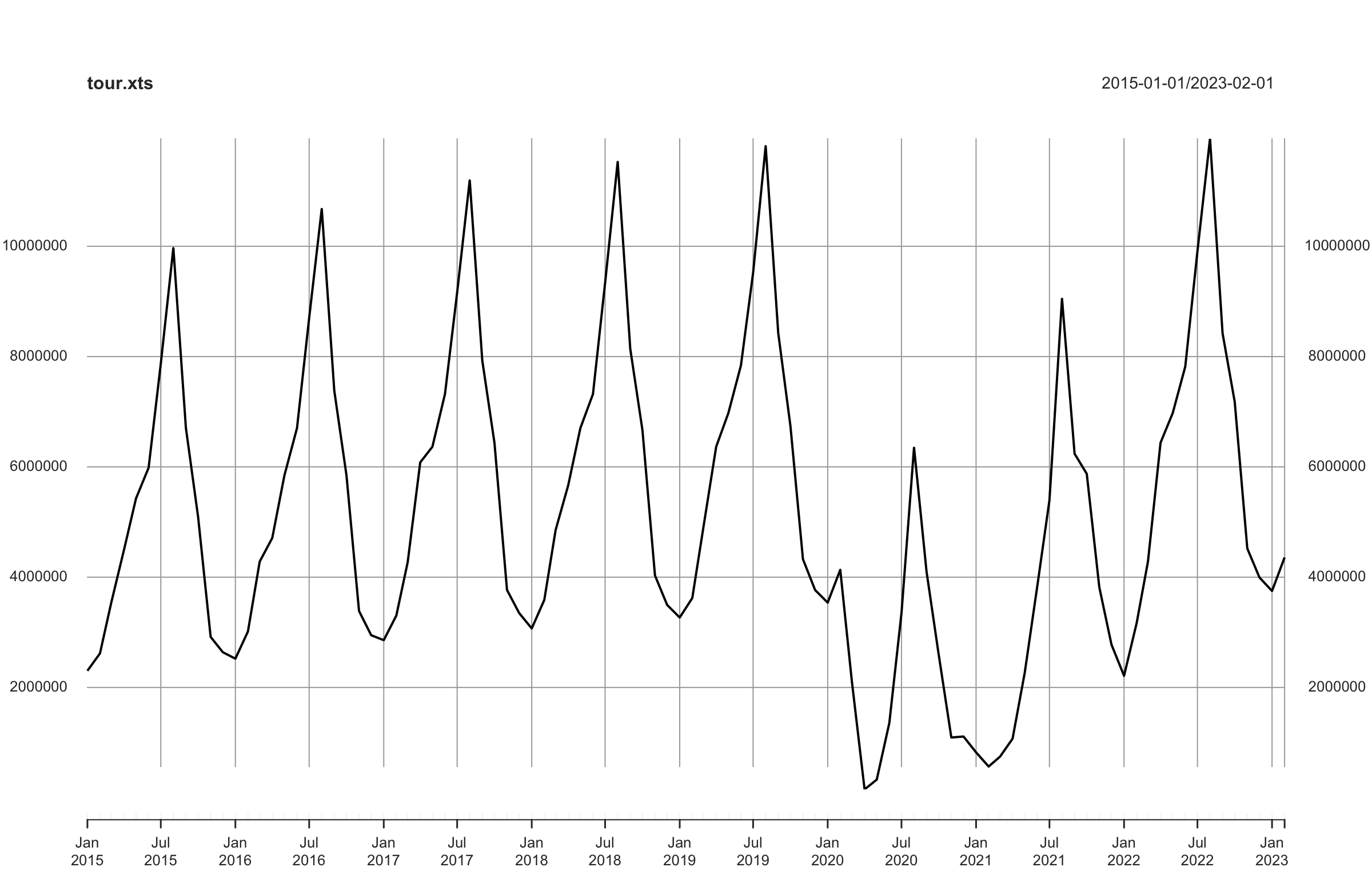
# Szereg sezonowy

## 2.1. Opis szeregu

Wybrany szereg sezonowy dotyczy liczby nocy spędzonych przez turystów w różnego rodzaju obiektach noclegowych, takich jak hotele czy pensjonaty. Dane zostały wybrane dla Portugalii, na przestrzeni lat 2015-2023. Analizowane dane mają charakter miesięczny,   
od stycznia 2015 do lutego 2023 roku (98 obserwacji), pozwalające na analizę sezonowości   
w ruchu turystycznym.

Zbiór danych pochodzi z Eurostatu i jest ogólnodostępny.

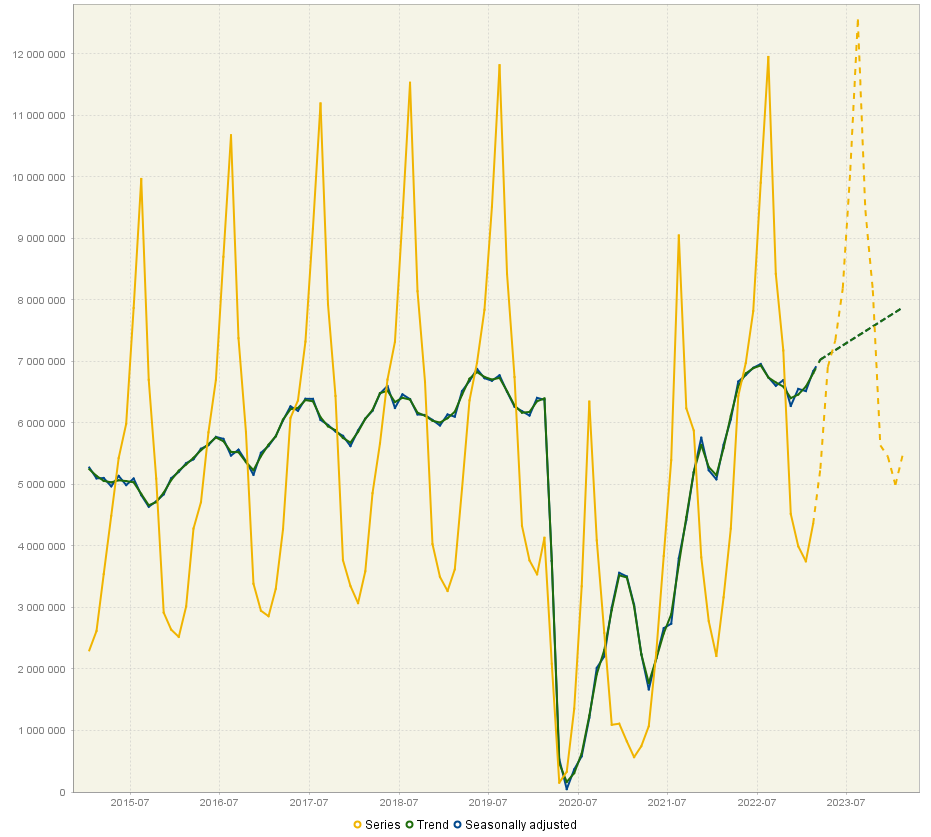
Wykres 10. Szereg czasowy ruchu turystycznego w Portugalii



## 2.2. Dekompozycja szeregu w programie Demetra

Analizę szeregu ponownie rozpoczęto od programu Demetra, w którym dokonano dekompozycji szeregu za pomocą metody TRAMO/Seats. Pozwoliło to na zauważenie sezonowego charakteru danych oraz nieliniowego trendu z jednym zaburzeniem wynikającym z pandemii COVID-19.

Wykres 11. Dekompozycja szeregu w programie Demetra



Trend nieliniowy w kontekście analizy szeregów czasowych odnosi się do wzorców zmian w danych, które nie są liniowe, czyli nie można ich opisać jednym stałym kierunkowym nachyleniem. Trend nieliniowy może być wynikiem różnych czynników, takich jak zmieniające się warunki gospodarcze, a w naszym przypadku pandemia COVID-19. W kontekście oceny trendu należy uwzględnić te zaburzenia jako czynnik tymczasowy. W takiej sytuacji, jeśli trend nieliniowy jest widoczny poza okresem zaburzeń, można uznać, że istnieje ogólny trend w danych.

## 2.3. Przygotowanie danych

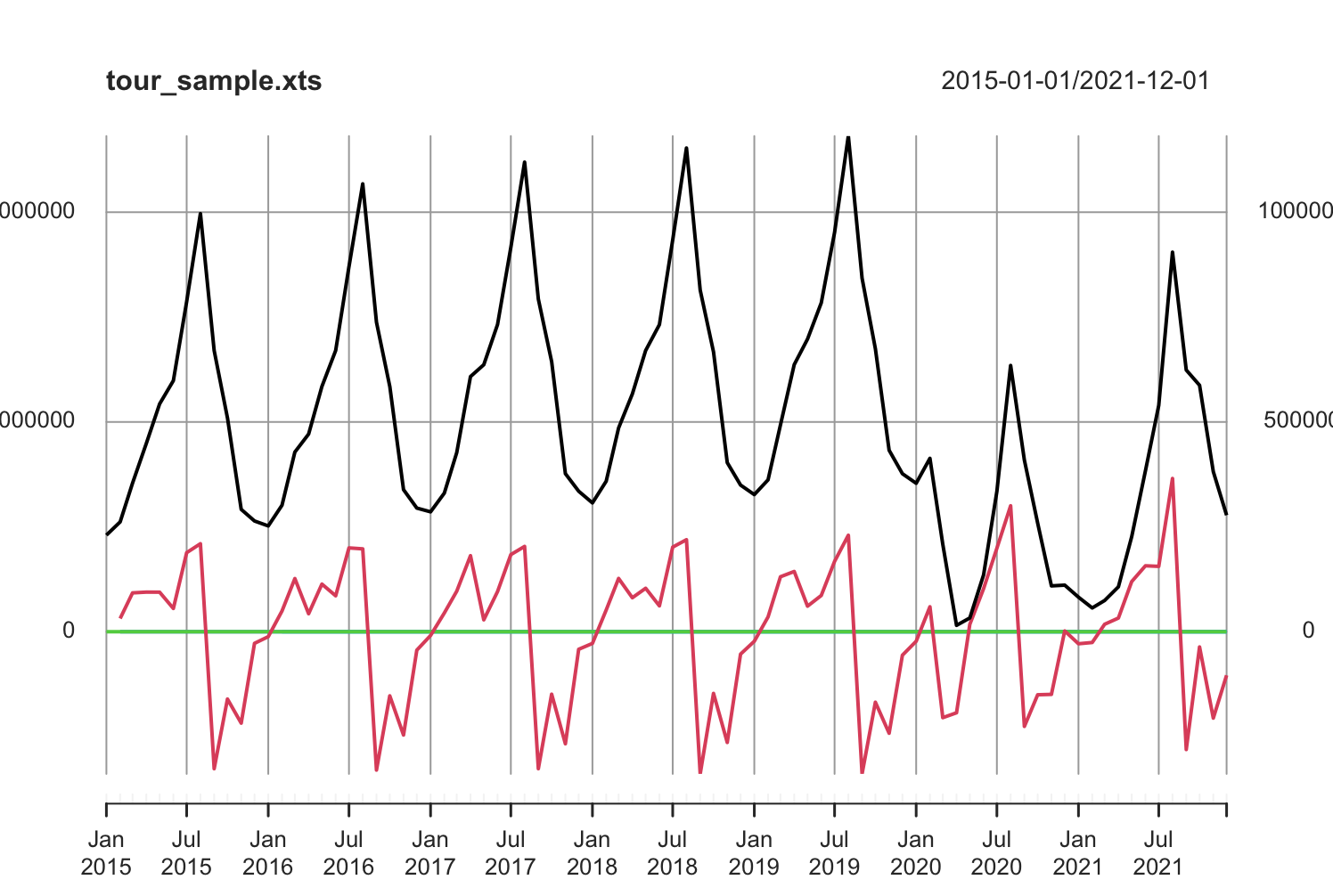
Proces analizy szeregu czasowego rozpoczęto od przygotowania próbki in-sample, która obejmuje wszystkie obserwacje do stycznia roku 2022. W tym kroku pozostawiono 12 miesięcy danych do przyszłej prognozy – out-of-sample.

Przygotowanie właściwej próbki in-sample jest kluczowym etapem analizy szeregów czasowych, ponieważ na podstawie niej konstruowany jest model i podejmowane są decyzje dotyczące prognozowania. Dalsze kroki analizy, takie jak identyfikacja struktury modelu, dobór parametrów, transformacje danych i ocena jakości prognoz, będą opierać się na wybranej próbce in-sample, aby zapewnić wiarygodność i trafność analizy.

## 2.4. Stacjonarność

Rozpoczęto analizę szeregu czasowego, który prezentuje się następująco:

Wykres 12. Szereg czasowy ruchu turystycznego dla próbki in-sample

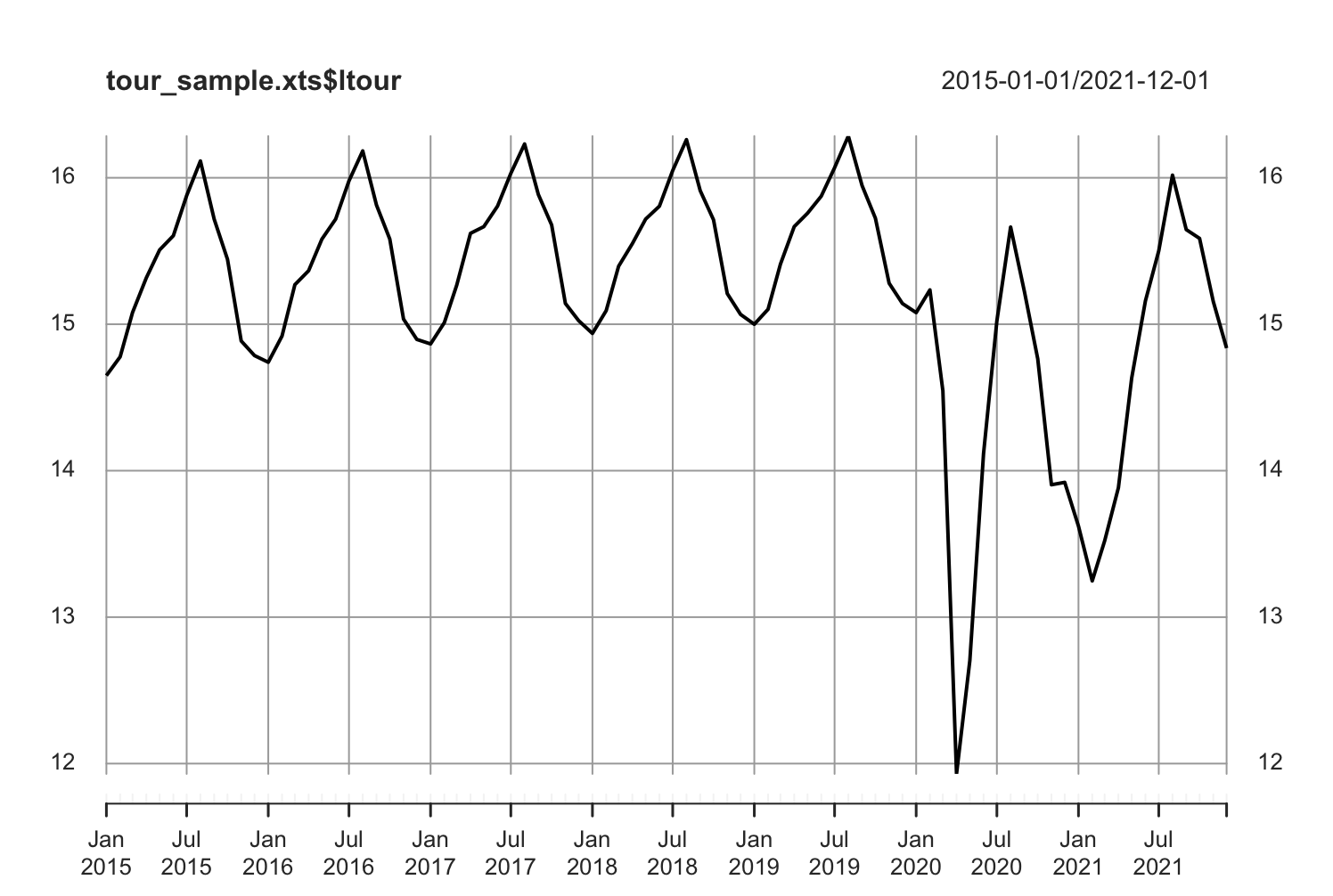
****

Na wykresie ponownie zauważalna jest sezonowość oraz trend nieliniowy.

Procedurę analizy szeregu rozpoczęto od obliczenia optymalnego parametru lambda dla transformacji Boxa-Coxa. Wartość lambda wyniosła w tym przypadku 0,7174473. Dla wartości optymalnego parametru lambda wynoszącej 0.717, można przypuszczać, że wprowadzenie logarytmowania danych może przynieść korzyści. Wartość lambda różna od 1 wskazuje na potrzebę transformacji danych w celu uzyskania bardziej liniowego lub symetrycznego rozkładu. Dodatkowo wykres nie wskazuje na występowanie stacjonarności. W związku z tym podjęto decyzje o zastosowaniu transformacji logarytmicznej w celu zmniejszenia wariancji.

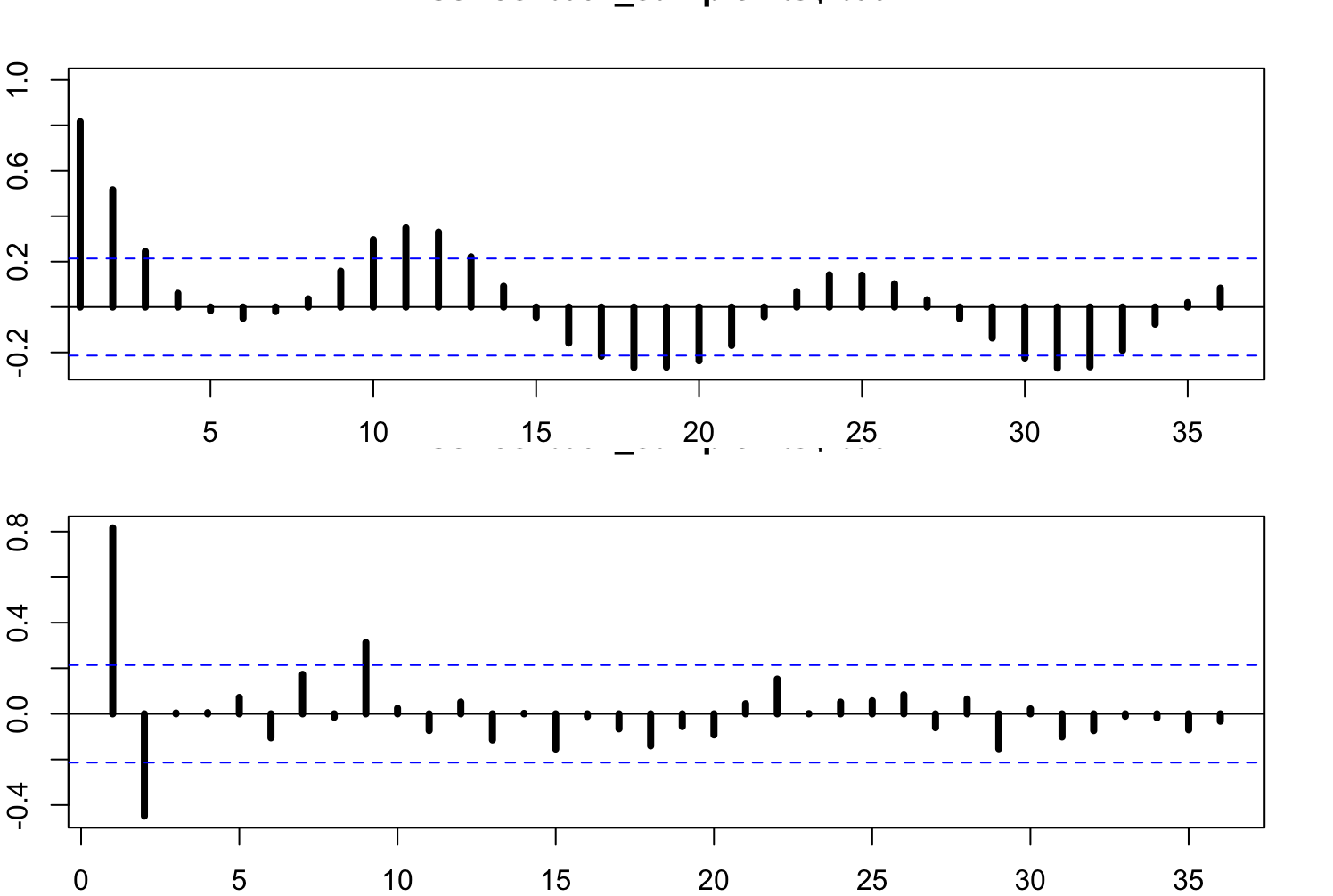
Poniższy wykres wizualizuje zmiany w danych po zastosowaniu transformacji logarytmicznej.

Wykres 13. Transformacja logarytmiczna szeregu czasowego

****

Następnie przystąpiono do analizy korelogramów ACF i PACF dla szeregu zlogarytmizowanego ltour.

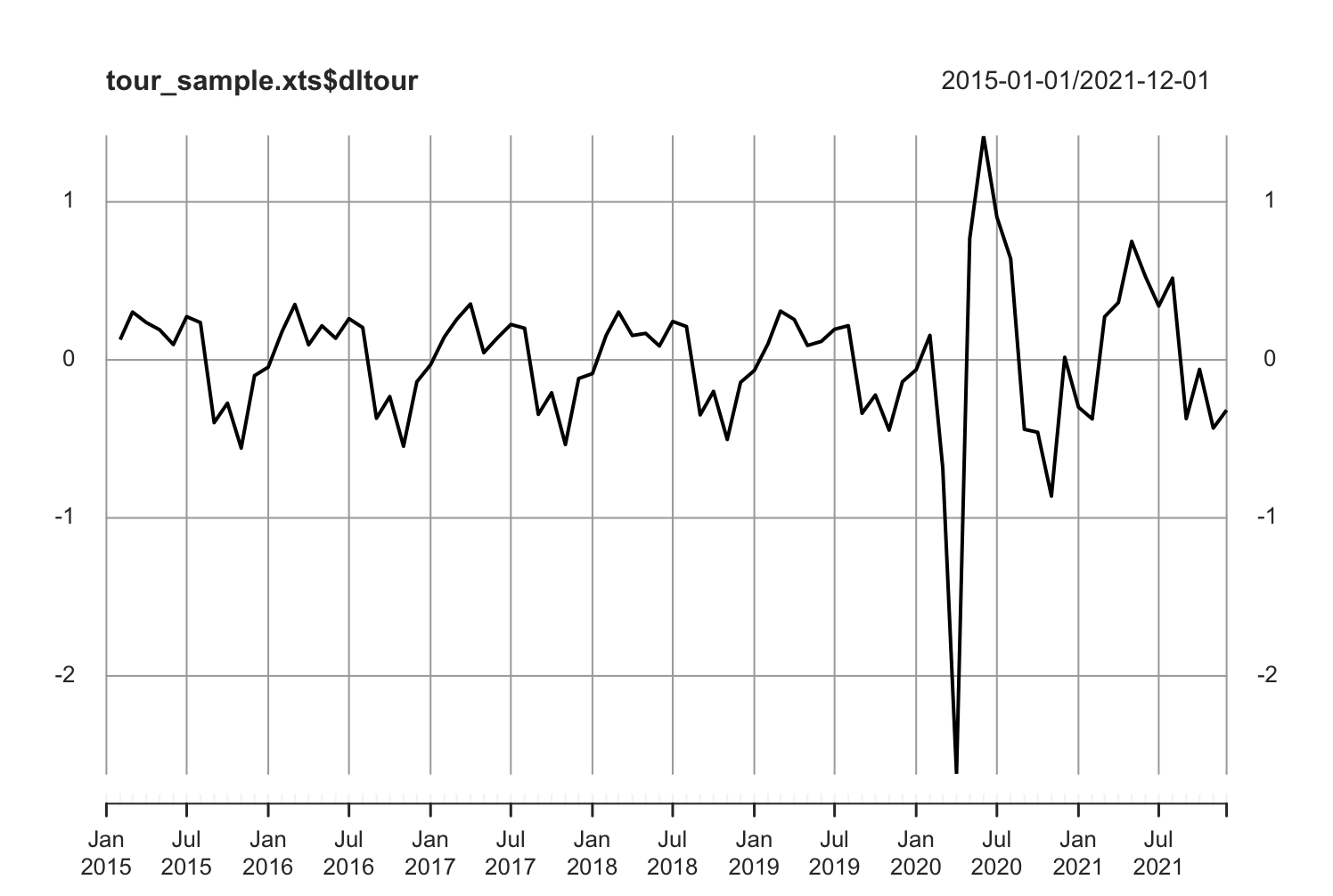
Wykres 14. Wykresy ACF i PACF dla zlogarytmizowanego szeregu



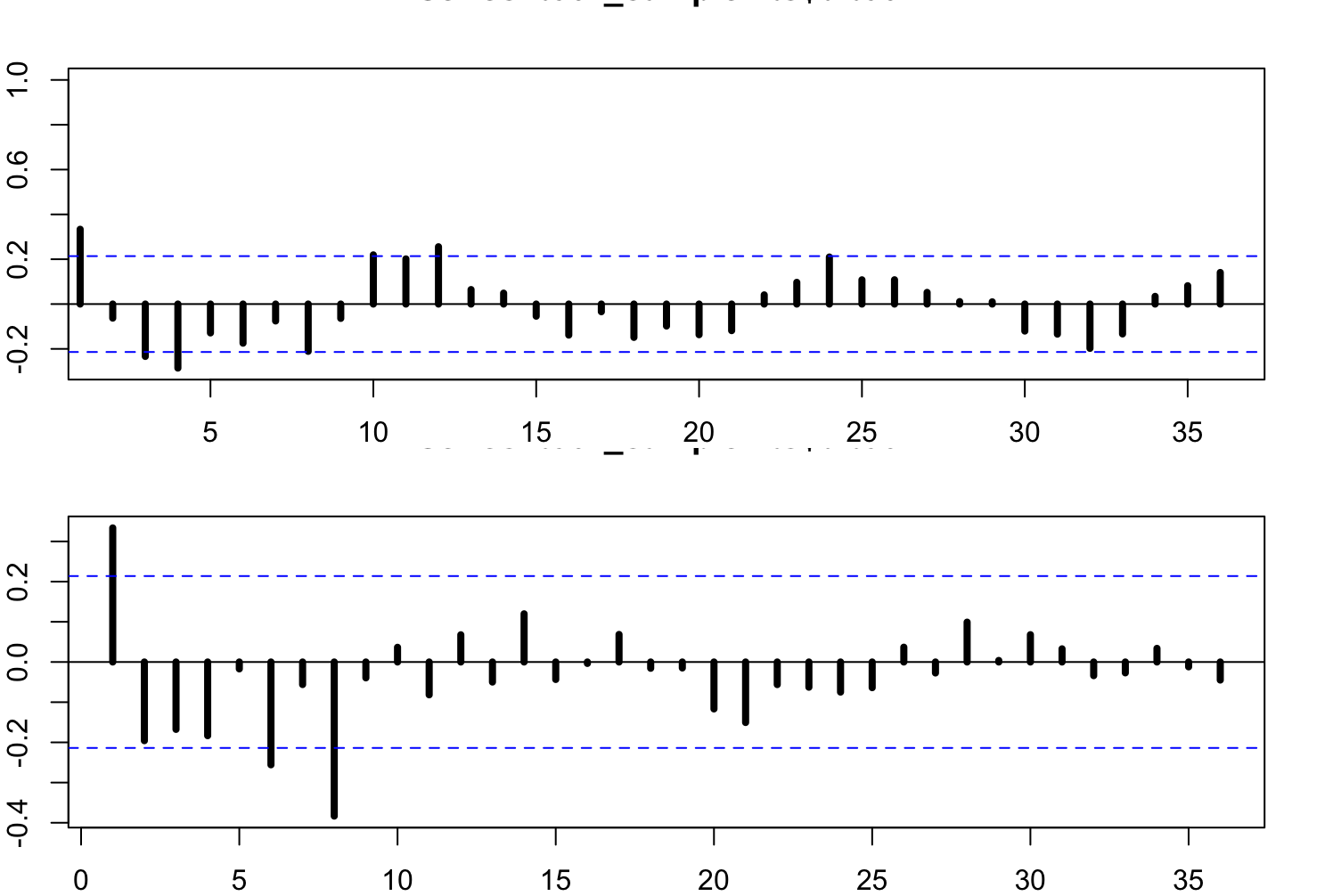
Wykres ACF ma formę sinusoidalną, co może wskazywać na obecność sezonowości w szeregu czasowym. Dodatkowo, pierwsze wartości ACF wykraczają poza obszar krytyczny, może to sugerować istnienie trendu w szeregu czasowym lub inną strukturę autokorelacji. Pierwsze wartości PACF również wykraczają poza obszar krytyczny, może to wskazywać na istotne korelacje na tych opóźnieniach, które nie są objaśniane przez wcześniejsze opóźnienia. Może to wskazywać na potrzebę uwzględnienia tych opóźnień w modelowaniu.

Weryfikacja ta nie dostarczyła więc wystarczająco dobrych wyników, dlatego wprowadzono pierwsze różnice i ponownie zweryfikowano korelogramy dla pierwszych różnic.

Wykres 15. Pierwsze różnice szeregu



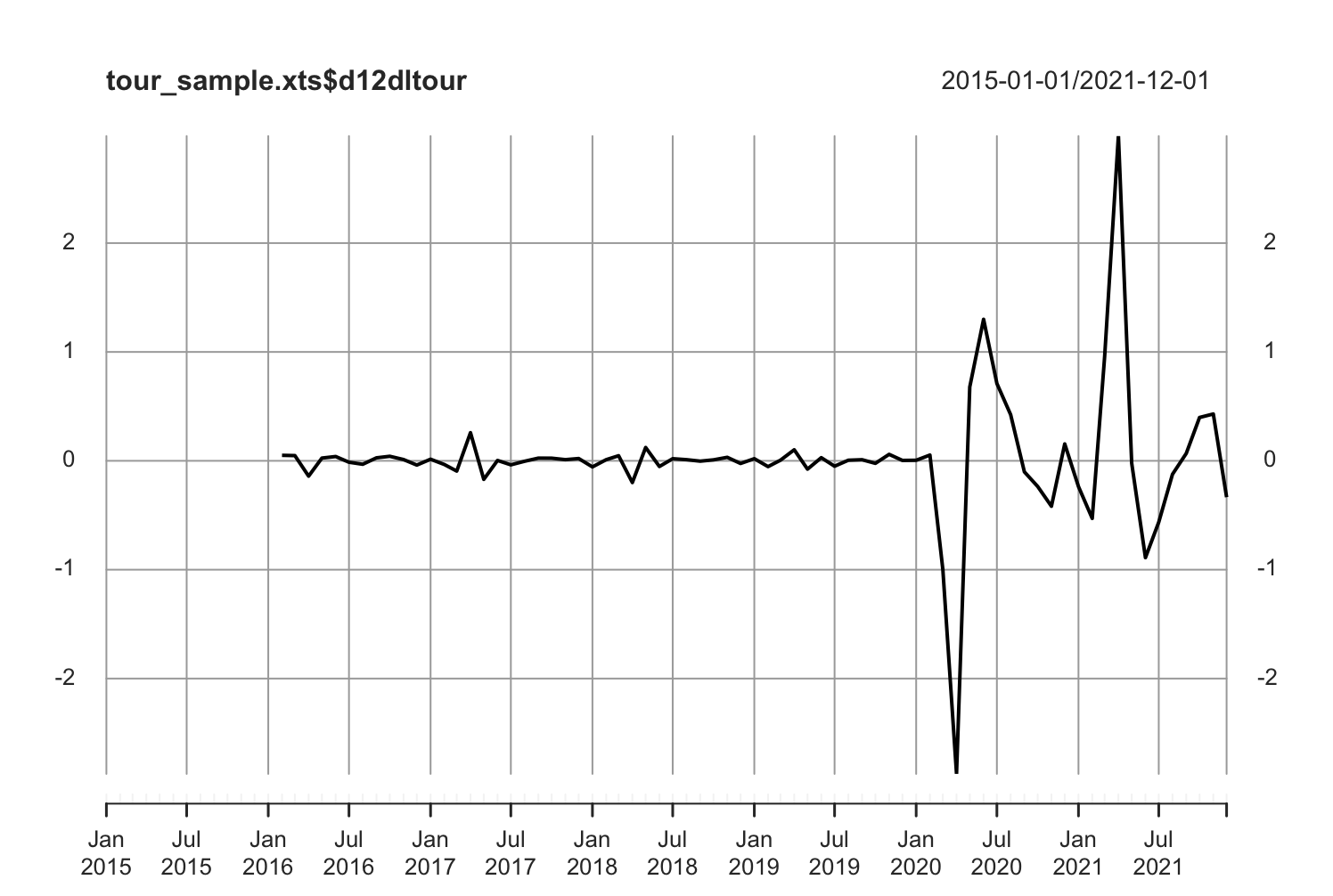
Wykres 16. Wykresy ACF i PACF dla pierwszych różnic szeregu



Falujące wzorce na wykresach ACF i PACF są oczekiwane w przypadku danych sezonowych. Wykresy dla pierwszych różnic wskazują na silną sezonowość, co potwierdzają wartości mieszczące się w obszarze krytycznym (z wyjątkiem tylko pierwszych słupków), może to wskazywać na obecność sezonowości w danych.

Co więcej, ACF wolno wygasa, zatem dodatkowo zróżnicowano szereg sezonowo – obliczono dwunaste różnice (ponieważ dane są miesięczne).

Wykres 17. Wykres dwunastu różnic szeregu



Kolejnym krokiem było formalne testowanie szeregu. Rozpoczęto od testu   
Dickeya, Haszy, Fullera (DHF) dla zmiennej logarytmowanej ltour. Mała wartość   
p-value = 0.00000000007942 wskazuje na odrzucenie hipotezy zerowej o braku autokorelacji w resztach. W tym przypadku, wartość p-value jest bardzo bliska zeru, co sugeruje, że istnieje silna podstawa do odrzucenia hipotezy zerowej. Wynik testu wskazuje na występowanie autokorelacji rzędu 1 w resztach modelu. Oznacza to, że reszty nie są niezależne od siebie, co może wskazywać na niedoszacowanie modelu lub brak uwzględnienia istotnych czynników w modelowaniu.

Ponieważ test Breuscha-Godfreya dla autokorelacji rzędu 1 wskazał istotną autokorelację w resztach modelu, warto przeprowadzić dalsze testy, takie jak Augmented Dickey-Fuller Test (ADHF), aby ocenić stacjonarność szeregów czasowych. Test ADHF służy do badania hipotezy o niestacjonarności szeregów czasowych.

Oto przedstawienie wyników testów BG dla różnych rzędów autokorelacji w formie tabeli:

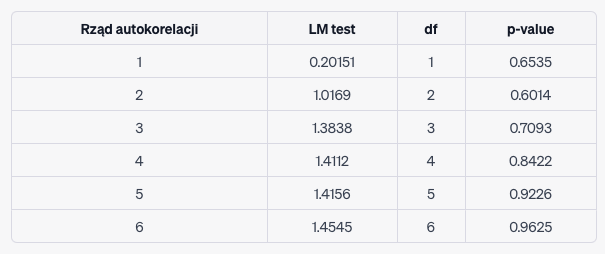


Tabela przedstawia wyniki testu Breusch-Godfrey dla różnych rzędów autokorelacji dla modelu 4 (dla k=4).

Analizując wyniki, można zauważyć, że dla wszystkich testowanych rzędów autokorelacji (od 1 do 6), wartości p są większe od ustalonego poziomu istotności (na przykład 0.05). Oznacza to, że nie ma wystarczających dowodów na istnienie autokorelacji o tych rzędach w resztach modelu 4. Wyniki te są zgodne z wcześniejszą interpretacją, że model 4 nie wykazuje istotnej autokorelacji.

Test ADHF potwierdził, że D=1 dla k=4.

Kolejno, przeprowadzono test Dickeya-Fullera (DF) dla sprawdzenia hipotezy czy zmienna jest zmienną niestacjonarną. Wyniki przedstawiono w formie tabeli:



Na podstawie wyników można zauważyć, że dla różnicy zerowej (augmentations = 0), wartość statystyki testowej DF (adf) wynosi -6.477415, a p-wartość (p\_adf) wynosi 0.01. Oznacza to, że istnieją silne dowody na stacjonarność szeregu czasowego.

Dla dalszych stopni różnicowania (augmentations = 1, 2, 3), wartości statystyki testowej DF są nadal ujemne, co wskazuje na stacjonarność szeregu czasowego. Jednakże, p-value (p\_adf) są wyższe od ustalonego poziomu istotności (0.01), co sugeruje, że dalsze różnicowanie może nie być konieczne.

Podsumowując, na podstawie wykonanych testów możemy stwierdzić uzyskanie szeregu stacjonarnego.

Ostatnim krokiem jest weryfikacja za pomocą testu KPSS, czy zmienna w modelu nie jest białym szumem. Podsumowując wyniki testów, można stwierdzić, że zmienna nie jest białym szumem. W teście Ljunga-Boxa otrzymano wartość statystyki równą 57.595 przy stopniach swobody równej 36, co wskazuje na istnienie autokorelacji w danych. Dodatkowo, p-value wynosi 0.0126, co oznacza, że istnieje statystycznie istotna autokorelacja.

W teście Boxa-Pierce'a otrzymano wartość statystyki równą 49.161 przy stopniach swobody równej 36. Chociaż p-wartość wynosi 0.07068, co sugeruje brak statystycznej istotności autokorelacji, warto zauważyć, że istnieje pewne odstępstwo od białego szumu.

## 2.5. Model SARIMA

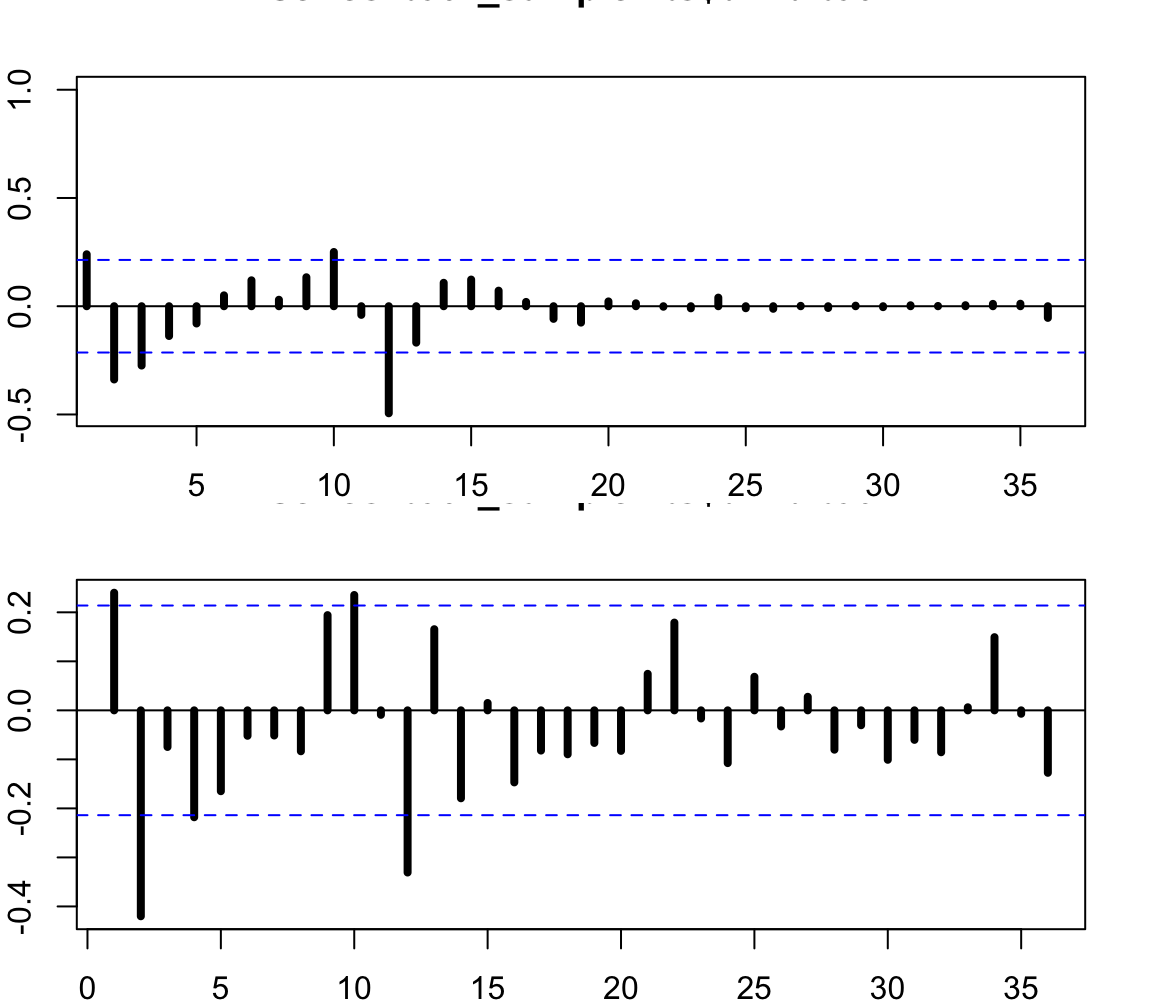
Weryfikacja stacjonarności modelu pozwoliła na przystąpienie do kolejnego etapu jakim jest modelowanie SARIMA. Prognoza będzie obejmowała okres out-of sample, czyli 12 ostatnich miesięcy.

Na podstawie pojedynczego zróżnicowania szeregu zostało stwierdzone, że d=1 w modelu SARIMA. Dodatkowo, po jednokrotnym zróżnicowaniu sezonowym pozbywamy się sezonowości, więc D=1.

**Identyfikacja**

Kolejnym etapem jest identyfikacja rzędów p,q,P, Q. W tym celu rozpoczęto od analizy korelogramów ACF i PACF:

Wykres 18. Wykresy ACF i PACF dla pierwszych różnic szeregu

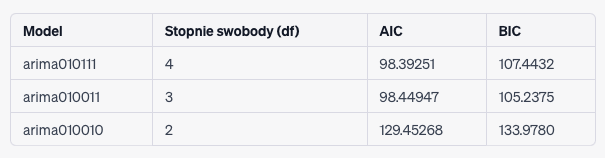


Korelogramy sugerują, że możemy mieć do czynienia z sezonowym procesem MA (malejąca PACF dla wielokrotności 12 opóźnienia). Z wykresu ACF zostaje odczytane Q=1, ponieważ tylko dla 12 widnieje istotnie różna od zera wartość słupka. Tą samą wartość wskazuje również PACF, więc P=1. Z kolei rzędy p i q wyznaczane są jak w modelu ARIMA – bierzemy pod uwagę pierwsze słupki wykresów, które wychodzą poza poziom istotności. Z wykresu PACF odczytujemy p=2, natomiast z wykresu ACF q=3. Wyniki wskazują, że potencjalnie efektywny model to: SARIMA(2,1,3)(1,1,1)12.

**Estymacja**

Estymacje modelowania rozpoczynamy od weryfikacji SARIMA(0,1,0)(1,1,1) w celu znalezienia najlepszych parametrów P i Q.

Przeprowadzono metodę od ogółu do szczegółu, a uzyskane wyniki przedstawia tabela poniżej:

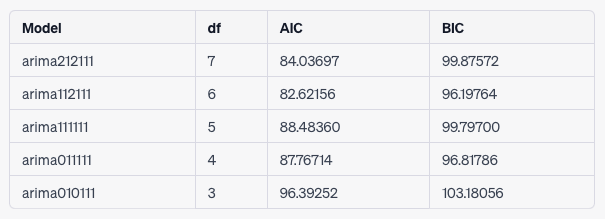
****

Dla kryterium informacyjnego Akaike (AIC), niższa wartość oznacza lepsze dopasowanie modelu. Na podstawie AIC można zauważyć, że model arima010111 ma najniższą wartość AIC (98.39251), co sugeruje, że ten model ma najlepsze dopasowanie do danych spośród trzech porównywanych modeli.

Dla kryterium informacyjnego Bayesa (BIC), również niższa wartość oznacza lepsze dopasowanie modelu. Tutaj również model arima010111 ma najniższą wartość BIC (107.4432), co wskazuje na jego najlepsze dopasowanie do danych.

Podsumowując, na podstawie wartości AIC i BIC można stwierdzić, że model arima010111 jest preferowany, ponieważ ma najniższe wartości obu kryteriów, co sugeruje najlepsze dopasowanie do danych.

Następnie przystąpiono do identyfikacji efektów regularnych – wprowadzono p=2 i q=3. Funkcja w R uniemożliwia wprowadzenie wartości q=3, dlatego estymacje modeli rozpoczęto od q=2. Poniższa tabela obrazuje uzyskane wyniki dla 5 różnych kombinacji parametrów p i q, również uzyskanych metodą od ogółu do szczegółu:



Na podstawie kryteriów AIC i BIC, model **arima112111** wydaje się być najlepszym modelem, który najlepiej dopasowuje się do danych przy uwzględnieniu liczby parametrów i złożoności modelu. Model arima011111 jest również dobrze dopasowany, ale ma nieco wyższą wartość AIC i BIC. Model arima010111 ma najgorsze dopasowanie do danych z uwagi na jego najwyższe wartości AIC i BIC.

Kolejnym etapem jest weryfikacja czy reszty modelu arima112111 są białym szumem. Dla testu Ljunga-Boxa, wartość statystyki testowej (X-squared) wynosi 11.97, a stopnie swobody (df) to 36. Wartość p-value wynosi 0.9999. Ponieważ wartość p-value jest znacznie większa niż zwykle przyjęty poziom istotności (0.05), nie ma podstaw do odrzucenia hipotezy o braku autokorelacji reszt. Oznacza to, że reszty modelu nie wykazują istotnej autokorelacji.

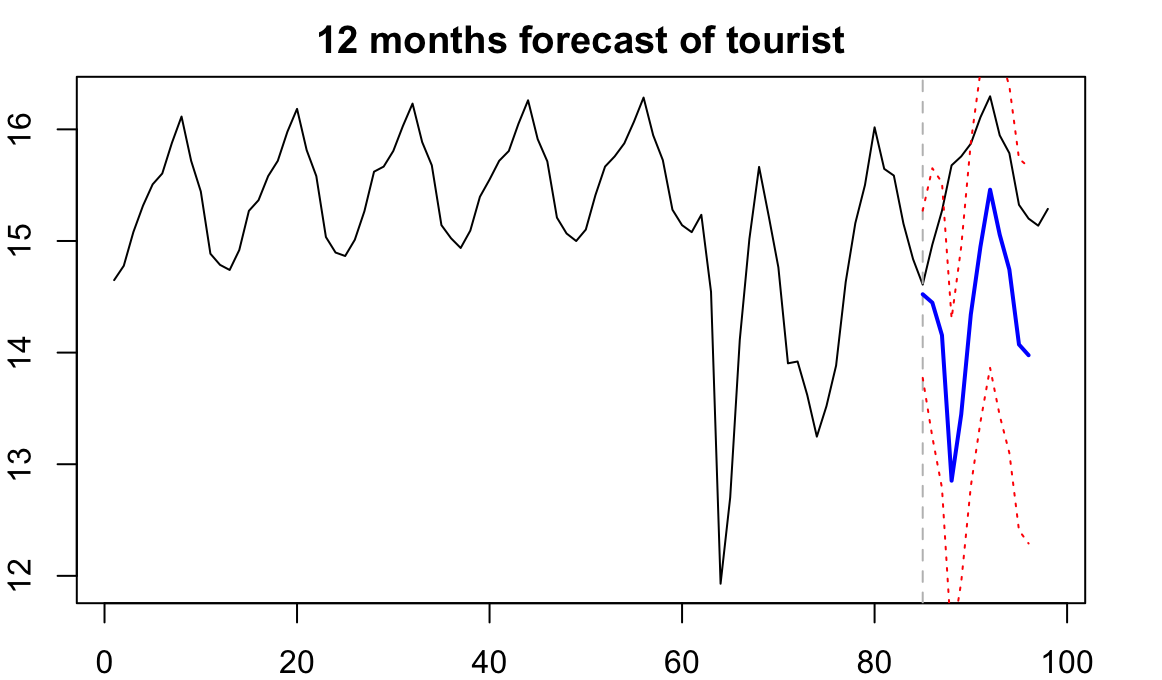
Podobnie, dla testu Boxa-Pierce'a, wartość statystyki testowej (X-squared) wynosi 9.6412, a stopnie swobody (df) to 36. Wartość p-value wynosi 1. Również tutaj wartość p-value jest dużo większa niż zwykle przyjęty poziom istotności, co wskazuje na brak autokorelacji reszt.

Wnioskując, na podstawie obu testów, możemy stwierdzić, że reszty modelu arima112111 są białym szumem.

**Prognozowanie**

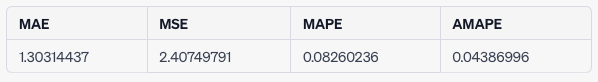
Prognozowanie przeprowadzono dla najlepszego wybranego modelu SARIMA(1,1,2)(1,1,1). Wyniki prognozy przedstawia poniższy wykres.

Wykres 19. Prognozowanie modelem SARIMA(1,1,2)(1,1,1)

****

Weryfikacja modelu wskazuje, że prognoza dla okresu out-of-sample (12 ostatnich miesięcy) nie pokrywa się z wartościami realnymi liczby nocy spędzonych przez trustów w Portugalii dla roku 2022. Jest to najprawdopodobniej spowodowane ponownym, nagłym wzrostem turystów odwiedzających ten kraj po zakończeniu pandemii COVID – 19.

W celu dodatkowej weryfikacji poprawności prognozowania obliczono średnie błędy prognozy SARIMA(1,1,2)(1,1,1). Pozwoli to na późniejsze porównanie poprawności prognozowania.



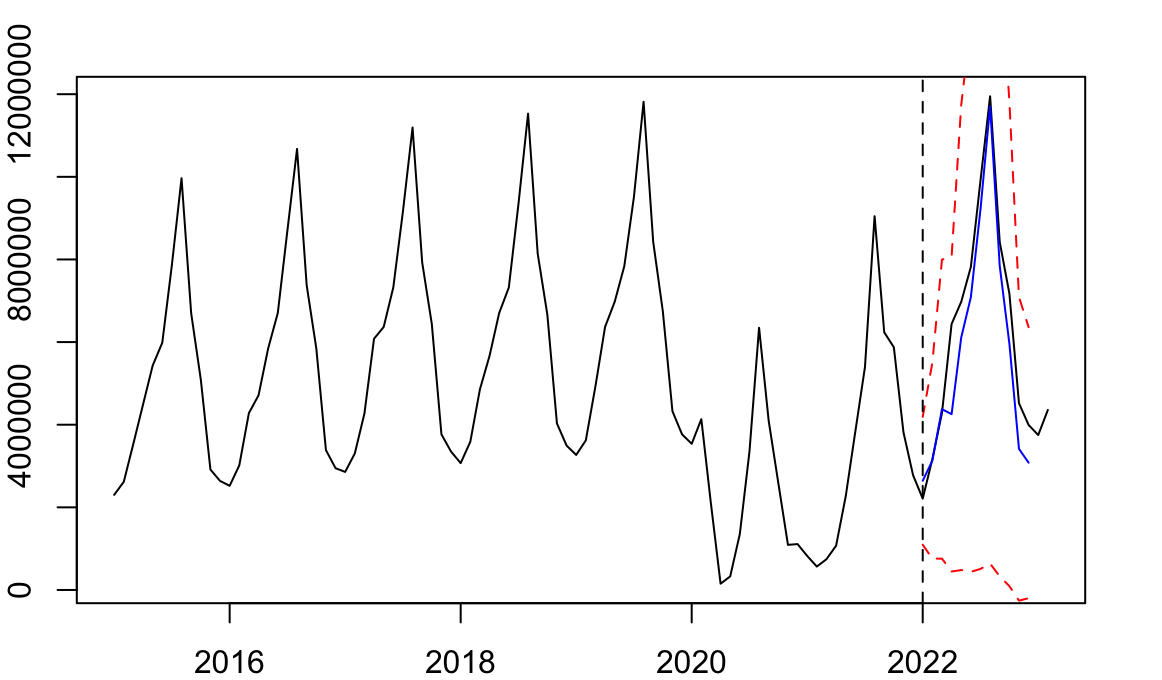
## 2.6. Model ekstrapolacyjny

W celu porównania prognozowania modelem SARIMA, wykonano również model ekstrapolacyjny Holta – Wintersa dla szeregów sezonowych. Ze względu na charakter sezonowości zweryfikowany w programie Demetra wybrano model multiplikatywny. Wartości składowe nie wahają się wokół stałej wartości średniej oraz rosną i maleją wraz ze wzrostem lub spadkiem poziomu szeregu.

W celu zachowania spójności prognoz z modelem SARIMA, oszacowań dokonano dla tej samej próbki in-sample a prognozy dla out-of sample.

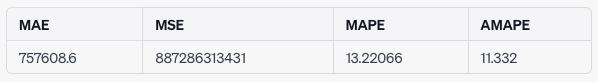
Poniższy wykres przedstawia prognozę przy wykorzystaniu multiplikatywnego modelu ekstrapolacyjnego Holta – Wintersa.

Wykres 20. Prognozowanie multiplikatywnym modelem ekstrapolacyjnym Holta – Wintersa



Weryfikacja wykresu wskazuje na prawidłowe dopasowanie prognoz do faktycznych wartości szeregu.

W celu dodatkowej weryfikacji poprawności prognozowania obliczono średnie błędy prognozy:



## 2.7. Porównanie modeli

Porównując wyniki modelu SARIMA(1,1,2)(1,1,1) z modelem multiplikatywnym Holta – Wintersa na podstawie wykresów jednoznacznie można stwierdzić, że model multiplikatywny lepiej estymuje wartości rzeczywiste niż najlepsza wybrana SARIMA.

Z kolei analiza błędów prognozy wskazuje odwrotny wynik prognozowania. Średnie błędy dla modelu multiplikatywnego są znacznie gorsze, co może wynikać z różnic w skali wartości między prognozami a rzeczywistymi danymi. Średnie błędy, takie jak MAE, MSE, MAPE i AMAPE, biorą pod uwagę różnice między prognozami a rzeczywistymi wartościami, niezależnie od skali danych. Dlatego nawet jeśli wydaje się, że model ekstrapolacyjny lepiej dopasowuje się do kształtu danych na wykresach, może wciąż generować większe średnie błędy ze względu na większe odchylenia wartości prognoz od rzeczywistych.

W związku z tym, mimo że model ekstrapolacyjny wydawał się lepiej prognozować na podstawie wizualnej oceny, analiza średnich błędów wskazuje na to, że model SARIMA(1,1,2)(1,1,1) okazał się lepszy. Mimo gorszego dopasowania graficznego, SARIMA generuje mniejsze średnie błędy, co oznacza, że prognozy tego modelu są bliższe rzeczywistym wartościom. Można również wysnuć hipotezę, że gdyby nie znaczne zaburzenie szeregu spowodowane pandemią COVID-19, mające bezpośredni wpływ na osłabienie turystyki, wybrany model SARIMA mógłby w doskonałym stopniu prognozować jej rzeczywiste wartości.

# PODSUMOWANIE

W powyższej pracy przeprowadzono analizę dwóch szeregów czasowych – niesezonowego oraz sezonowego. Dla każdego z nich dokonano dekompozycji oraz przeprowadzono prognozy wybranym modelem ekstrapolacyjnym. Szereg niesezonowy okazał się białym szumem, co uniemożliwiło dopasowanie modelu z klasy ARIMA. Dla szeregu sezonowego dopasowano natomiast najlepszy model SARIMA(1,1,2)(1,1,1).