

# Prognozowanie szeregów czasowych - analiza porównawcza modeli (S)ARIMA, ETS i dekompozycji

Dominika Szulc

2026-02-04

## Spis treści

<b>1</b>	<b>Porównanie dokładności prognoz dla danych euretail</b>	<b>2</b>
1.1	Cel analizy . . . . .	2
1.2	Dane euretail . . . . .	2
1.3	Stacjonarność szeregu uczącego . . . . .	5
1.4	Modele (S)ARIMA . . . . .	9
1.5	Modele dekompozycji . . . . .	18
1.6	Algorytmy wygładzenia wykładniczego . . . . .	28
1.7	Prognozowanie . . . . .	36
1.8	Porównanie . . . . .	46
<b>2</b>	<b>Źródła</b>	<b>46</b>

# 1 Porównanie dokładności prognoz dla danych euretail

## 1.1 Cel analizy

Naszym celem jest porównanie dokładności prognoz skonstruowanych na bazie:

- modeli (S)ARIMA,
- modeli dekompozycji,
- algorytmów wygładzania wykładniczego

dla danych `euretail` z pakietu `fpp2`. Zbiór zawiera kwartalne wartości indeksu handlu detalicznego dla strefy euro w latach 1996 – 2011.

Jako domyślny poziom istotności przyjmujemy  $\alpha = 0.05$ .

## 1.2 Dane euretail

### 1.2.1 Typ i klasa danych

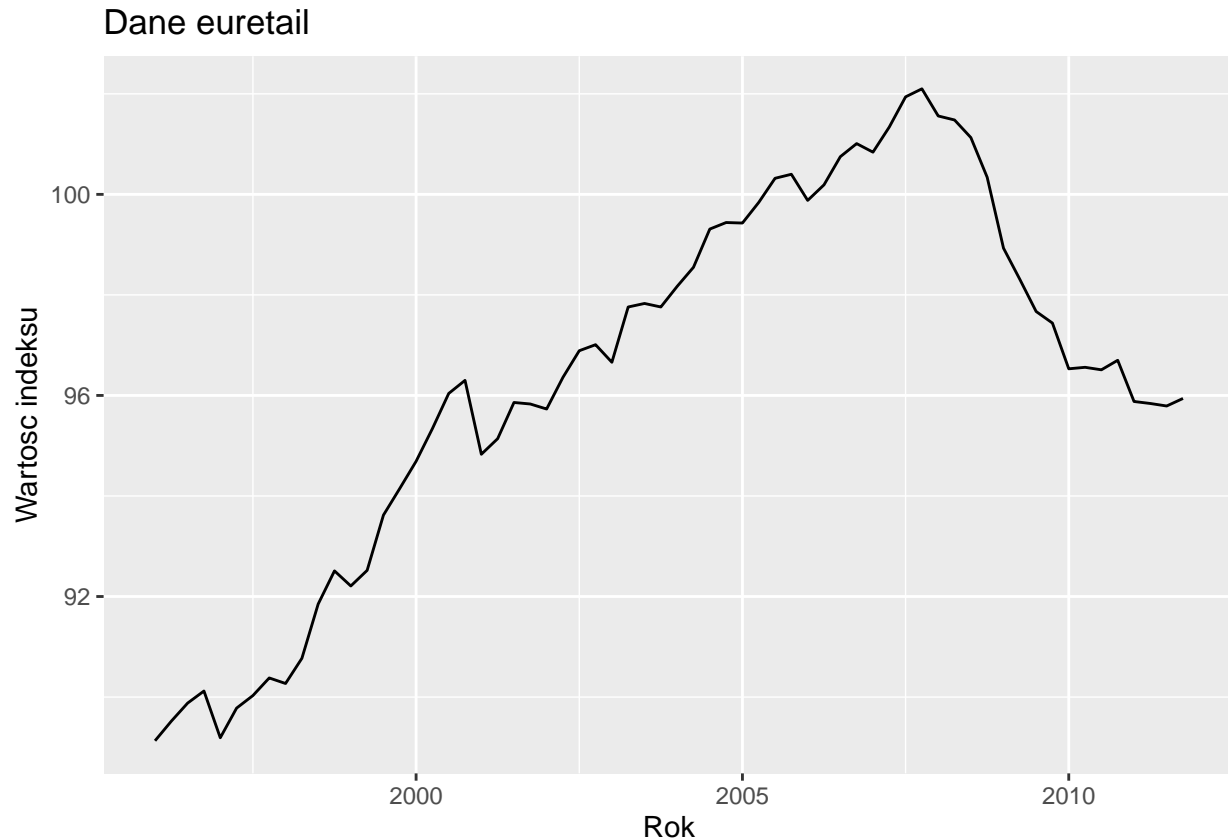
Zanim zaczniemy analizę, przyjrzymy się naszym danym.

```
##      Qtr1   Qtr2   Qtr3   Qtr4
## 1996  89.13  89.52  89.88  90.12
## 1997  89.19  89.78  90.03  90.38
## 1998  90.27  90.77  91.85  92.51
## 1999  92.21  92.52  93.62  94.15
## 2000  94.69  95.34  96.04  96.30
## 2001  94.83  95.14  95.86  95.83
## 2002  95.73  96.36  96.89  97.01
## 2003  96.66  97.76  97.83  97.76
## 2004  98.17  98.55  99.31  99.44
## 2005  99.43  99.84 100.32 100.40
## 2006  99.88 100.19 100.75 101.01
## 2007 100.84 101.34 101.94 102.10
## 2008 101.56 101.48 101.13 100.34
## 2009  98.93  98.31  97.67  97.44
## 2010  96.53  96.56  96.51  96.70
## 2011  95.88  95.84  95.79  95.94

## Time-Series [1:64] from 1996 to 2012: 89.1 89.5 89.9 90.1 89.2 ...
```

Są one klasy `ts`, co oznacza, że są szeregiem czasowym. Nie musimy dokonywać żadnych modyfikacji w tym kierunku.

### 1.2.2 Wykres



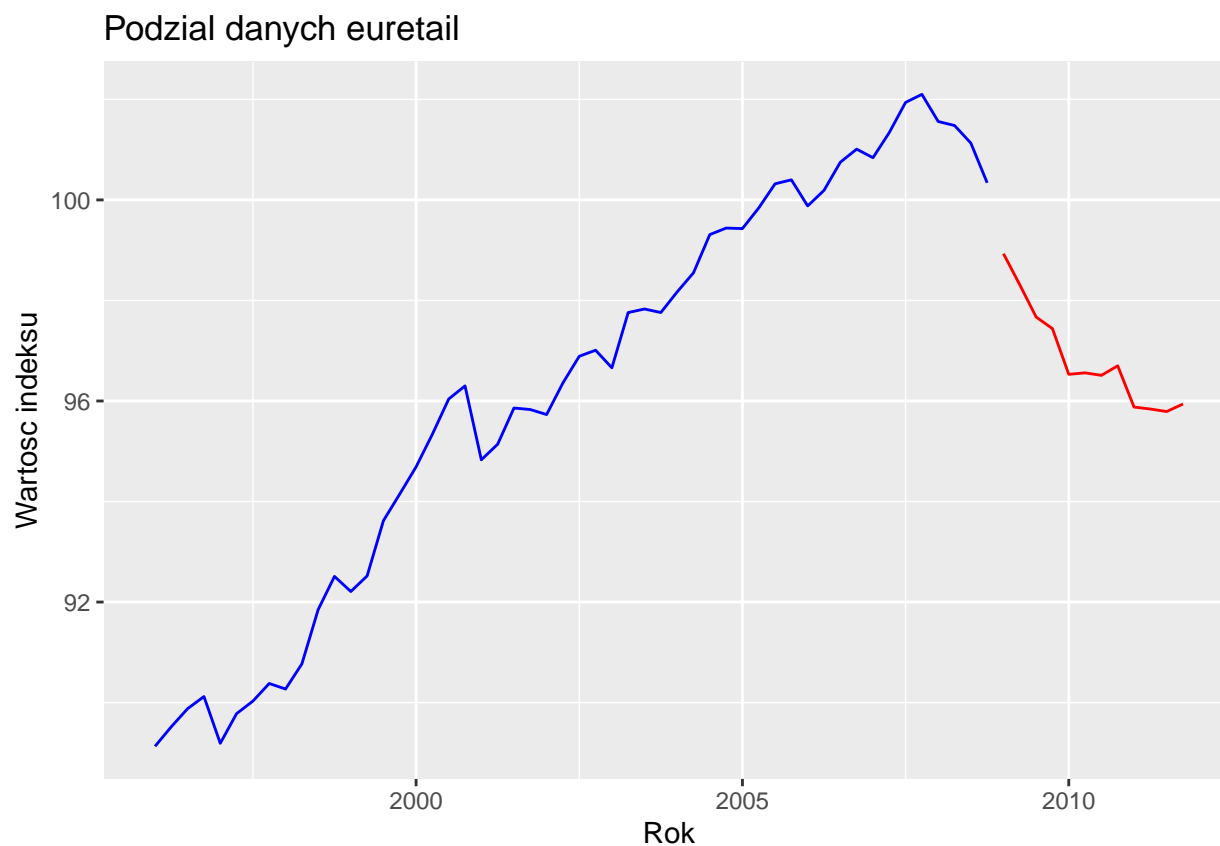
Rysunek 1: Wykres przedstawiający dane 'euretail' - kwartalne wartości indeksu handlu detalicznego w strefie euro.

Jak widać na wykresie 1 do około 2008 roku obserwujemy wyraźny trend wzrostowy, a następnie gwałtowny spadek spowodowany zapewne globalnym kryzysem finansowym, którego szczyt przypadł na lata 2008 – 2009. Sytuacja na rynku stabilizuje się dopiero w okolicach 2010 roku. Ponadto obserwujemy również sezonowość kwartalną - na początku prawie każdego roku wzrost do połowy roku, a potem spadek. Amplituda wahań sezonowych nie wydaje się być stała.

Obecność trendu powoduje brak stałej średniej, a sezonowość regularne cykle naruszające stacjonarność. Brak stałości amplitudy natomiast niestabilność wariancji. Możemy zatem stwierdzić, że szereg **euretail** nie jest szeregiem stacjonarnym. Dokładniejszych badań dokonamy po podziale danych na zbiór uczący i testowy.

### 1.2.3 Podział na zbiór uczący i testowy

Podzielimy teraz nasze dane na zbiór uczący i testowy w taki sposób, aby w zbiorze testowym znajdowało się ostatnich 12 kwartałów.

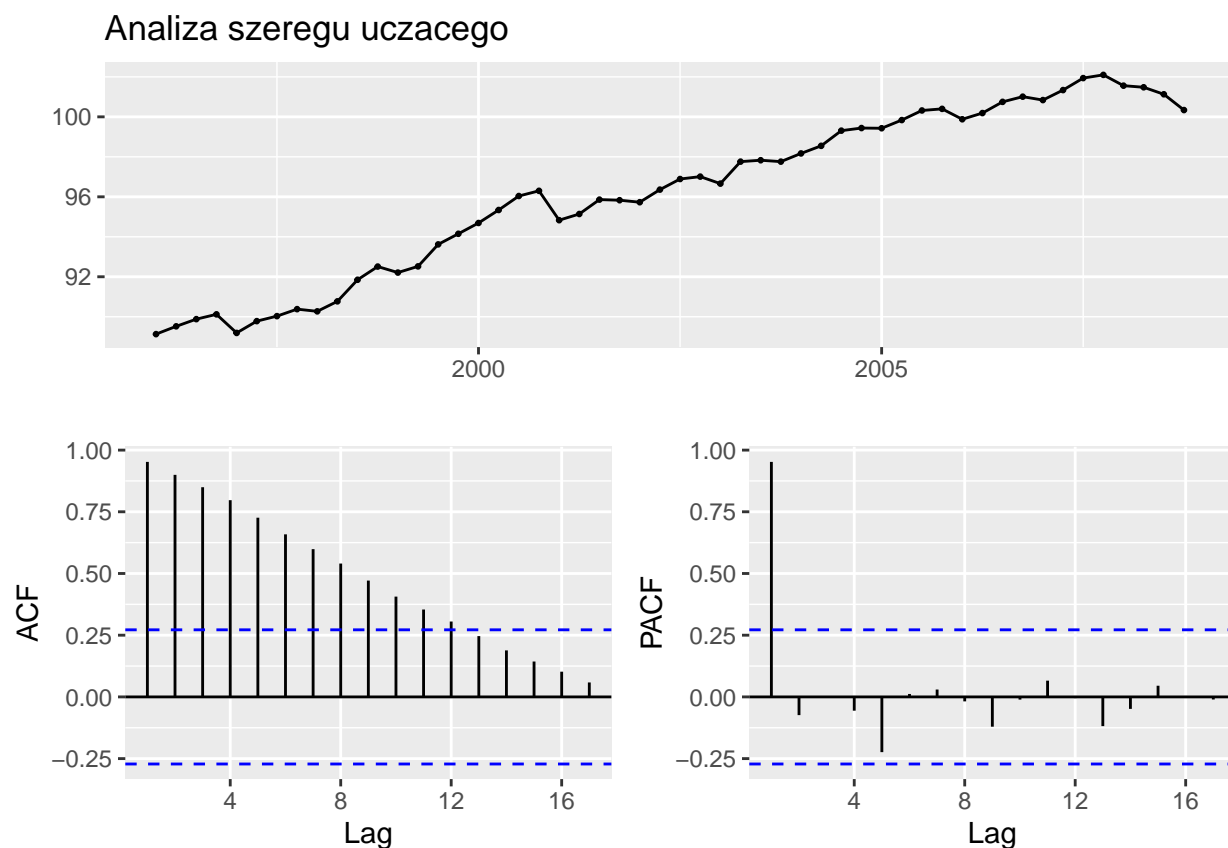


Rysunek 2: Wykres przedstawiający podział danych euretail na zbiór uczący (kolor niebieski) i testowy (kolor czerwony).

W dalszej części sprawozdania będziemy przekształcać zbiór (szereg) uczący.

### 1.3 Stacjonarność szeregu uczącego

Zanim zaczniemy dopasowywać modele do naszego szeregu uczącego, musimy doprowadzić go do postaci stacjonarnej.



Rysunek 3: Wykresy dla zbioru uczącego: szereg czasowy (góra), funkcja autokorelacji ACF (dół-lewo) oraz autokorelacji cząstkowej PACF (dół-prawo).

Jak zasugerowaliśmy wcześniej, nasze dane nie są szeregiem stacjonarnym. Wykresy ACF oraz PACF (4) potwierdzają to przypuszczenie:

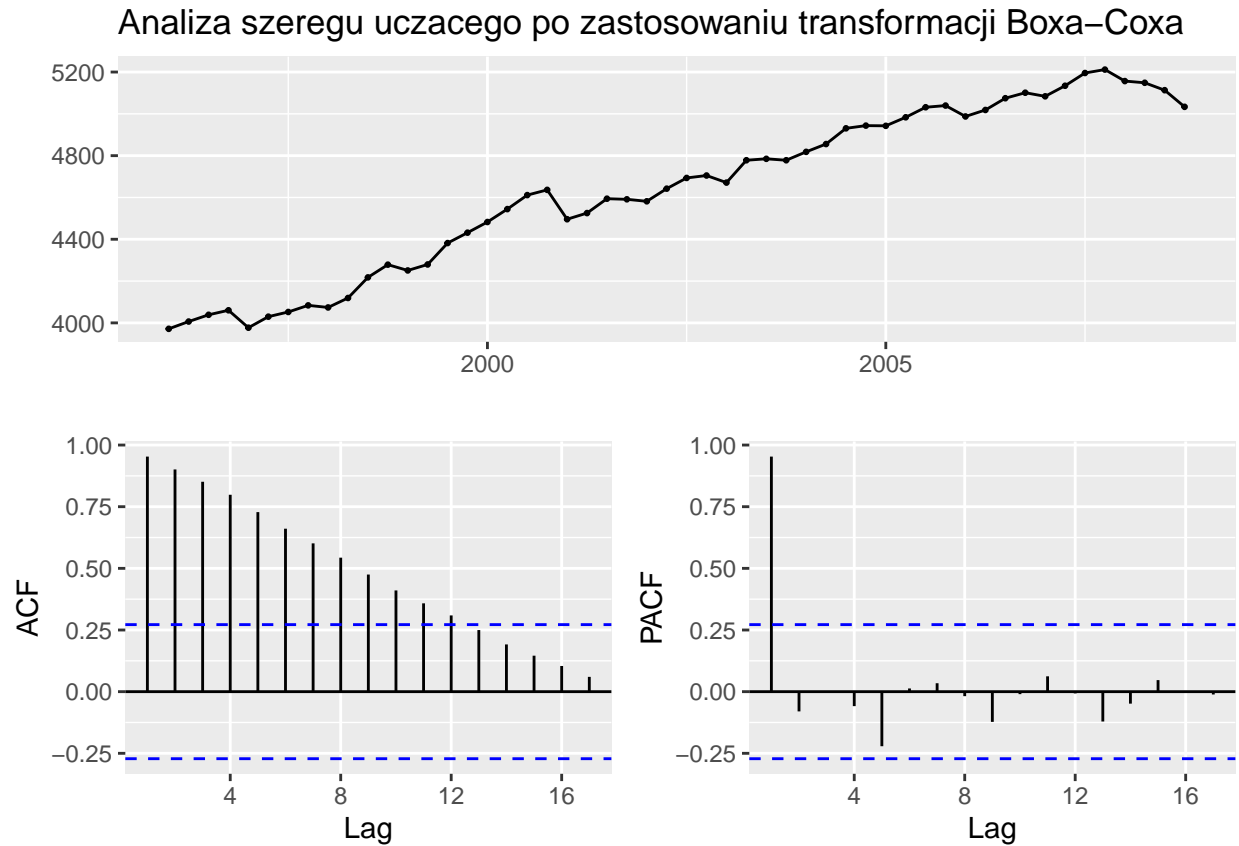
- dodatnie, powoli zanikające wartości funkcji ACF = deterministyczna składowa trendu w danych,
- duża, bliska 1 wartość funkcji PACF dla opóźnienia  $h = 1$  = obecność silnego trendu wzrostowego w danych.

Żaden z dwóch dolnych wykresów nie sugeruje obecności sezonowości w danych. Może być to spowodowane tym, że trend jest na tyle silny, że jak narazie nie widać efektu sezonowego.

### 1.3.1 Transformacja Boxa-Coxa

Korzystając z funkcji `BoxCox.lambda` dla zbioru uczącego, sprawdzamy, czy transformacja Boxa-Coxa, stabilizująca wariancję naszego szeregu, jest potrzebna. Wynik  $\lambda \neq 1$  sugeruje potrzebę zastosowania transformacji z podaną wartością  $\lambda$ .

Dla naszego szeregu dostajemy  $\lambda = 2$ , zatem stosujemy transformację Boxa-Coxa z tą wartością parametru.



Rysunek 4: Wykresy dla zbioru uczącego po zastosowaniu transformacji Boxa-Coxa z parametrem  $\lambda = 2$ : szereg czasowy (góra), funkcja autokorelacji ACF (dół-lewo) oraz autokorelacji cząstkowej PACF (dół-prawo).

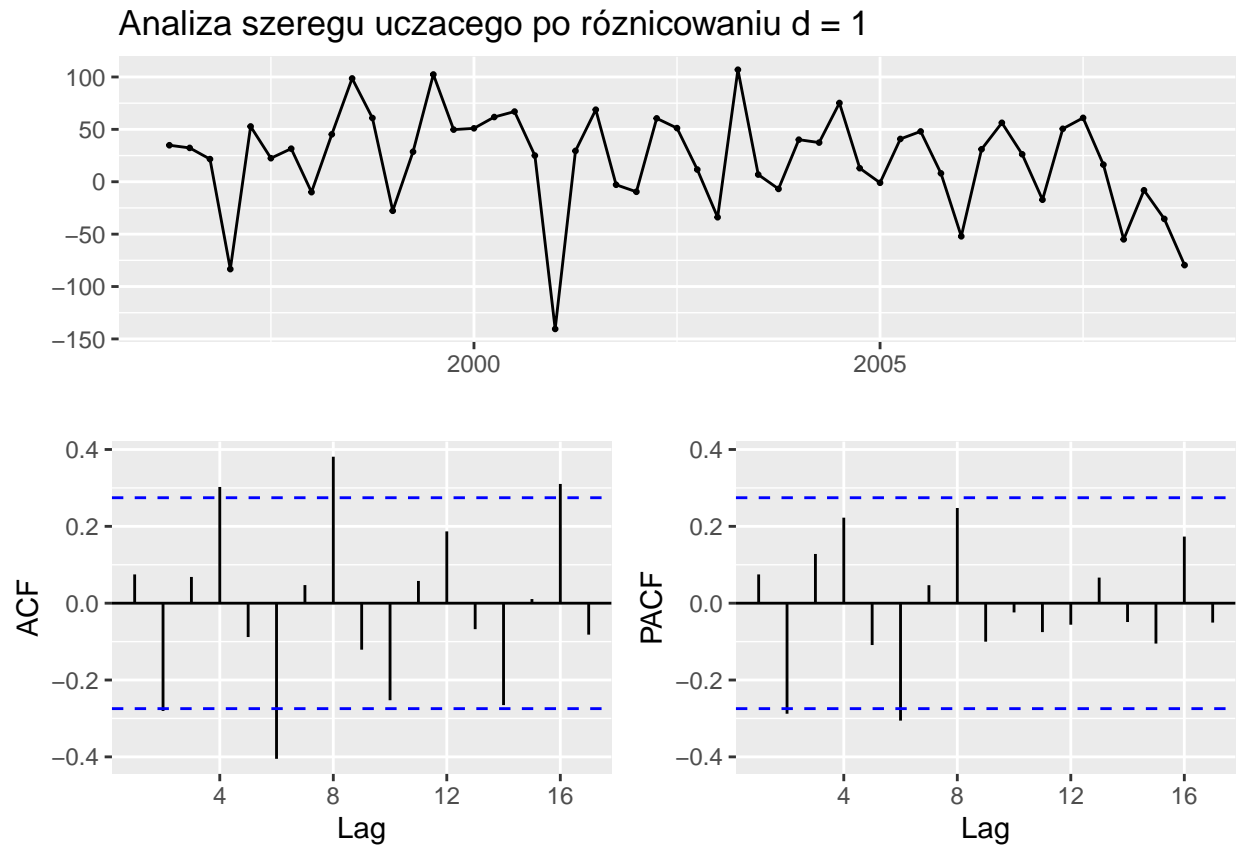
Porównując wykres 4 oraz wykres 4, zastosowanie transformacji Boxa-Coxa z parametrem  $\lambda = 2$  nie wpłynęło zauważalnie na strukturę funkcji ACF i PACF. Nadal na pierwszym planie jest wyraźny trend wzrostowy.

W dalszej części sprawozdania analizie będziemy poddawać stransformowane dane.

### 1.3.2 Różnicowanie

Aby usunąć trend z naszych danych zastosujemy różnicowanie. Skorzystamy z funkcji `ndiffs`, która pozwoli wyznaczyć sugerowaną liczbę potrzebnych różnicowań z opóźnieniem 1.

Wykonujemy  $d = 1$  różnicowań z opóźnieniem 1.



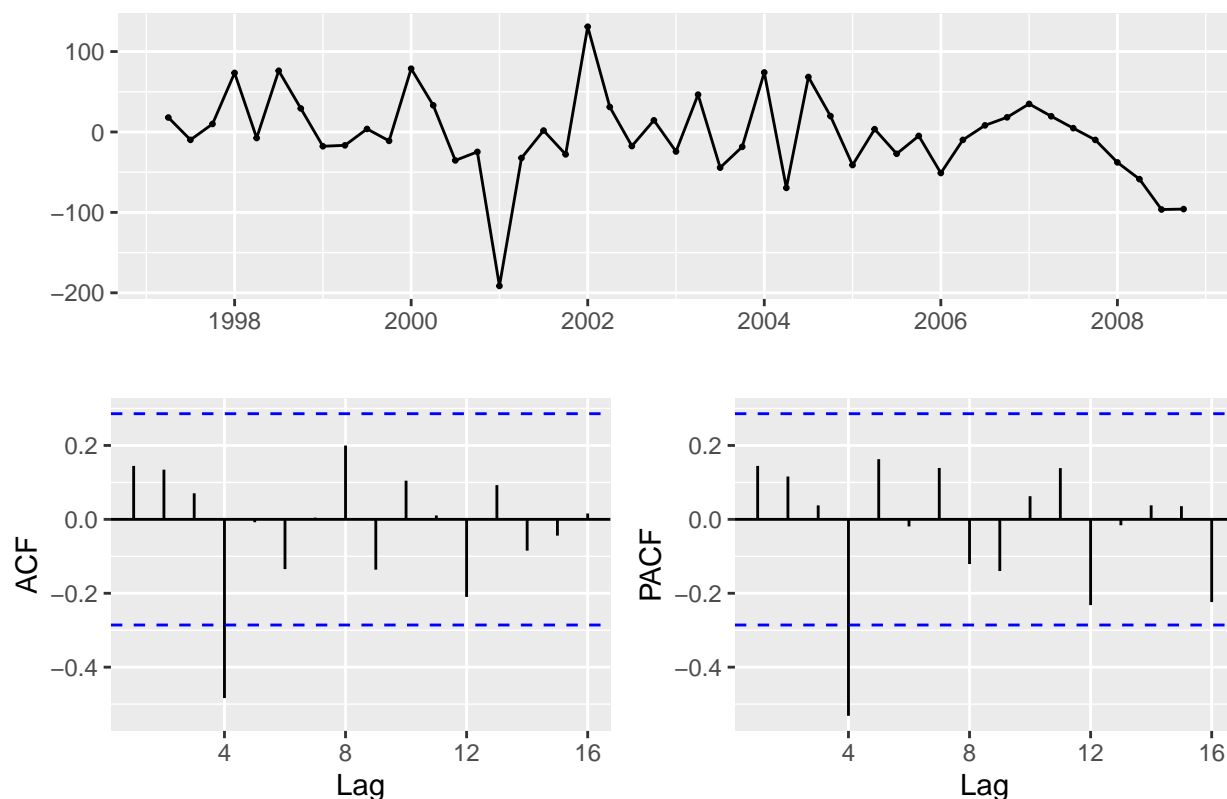
Rysunek 5: Wykresy dla zbioru uczącego Boxa-Coxa po zastosowaniu różnicowania z opóźnieniem  $d = 1$ : szereg czasowy (góra), funkcja autokorelacji ACF (dół-lewo) oraz autokorelacji cząstkowej PACF (dół-prawo).

Jak widać na wykresach 5 trend został skutecznie usunięty. Odsłoniło to nowe informacje o naszym szeregu. Na wykresie funkcji ACF widać cykliczne zanikanie i wystające słupki dla opóźnień 4, 8, 12 i 16 (dodatnia korelacja kwartalna) oraz 2, 6, 10 i 14 (ujemna korelacja w drugim kwartale) - obecność sezonowości kwartalnej.

Stosujemy funkcję `nsdiffs`, która pozwoli na wyznaczenie sugerowanej liczby różnicowań sezonowych z opóźnieniem 4.

Robimy zatem  $D = 1$  różnicowań z opóźnieniem sezonowym  $s = 4$ .

### Analiza szeregu uczącego po różnicowaniu $d = 1$ oraz $D = 1$



Rysunek 6: Wykresy dla zbioru uczącego Boxa-Coxa po zastosowaniu różnicowania z opóźnieniem  $d = 1$ ,  $D = 1$  oraz  $s = 4$ : szereg czasowy (góra), funkcja autokorelacji ACF (dół-lewo) oraz autokorelacji cząstkowej PACF (dół-prawo).

Jak widać na wykresach 6 sezonowość została skutecznie usunięta.

Możemy teraz swobodnie dobrać rzędy modeli ruchomej średniej oraz autoregresji.

#### 1.3.3 Model $MA(q)$

Rząd modelu ruchomej średniej identyfikujemy jako opóźnienie dla ostatniej statystycznie istotnej wartości funkcji autokorelacji ACF. Z wykresu 6 mamy  $q = 4$ .

#### 1.3.4 Model $AR(p)$

Rząd modelu autoregresji identyfikujemy jako opóźnienie dla ostatniej statystycznie istotnej wartości wykresie funkcji autokorelacji PACF. Z wykresu 6 mamy  $p = 4$ .



## 1.4 Modele (S)ARIMA

Dopasowywanie modeli do danych zaczynamy od modeli (S)ARIMA.

### 1.4.1 Dryf

W przypadku modeli (S)ARIMA ważne jest, by sprawdzić wartość średniej szeregu. Pozwala to stwierdzić, czy warto uwzględnić dryf w modelu.

Średnia szeregu uczącego Boxa-Coxa po zastosowaniu różnicowania z opóźnieniem  $d = 1$ ,  $D = 1$  oraz  $s = 4$  wynosi  $-3.9041872$ . Jest to wartość różna od zera, co sugeruje możliwość uwzględnienia dryfu w naszych modelach. W dokumentacji funkcji `Arima` z pakietu `forecast`, którą posłużymy się do tworzenia modeli, znajdujemy informację, że jeśli zastosowano więcej niż jedno różnicowanie (tak, jak w naszym przypadku), to dryf nie zostanie uwzględniony niezależnie od ustawień. Tworzymy zatem modele bez dryfu.

### 1.4.2 Model SARIMA(0, 1, 4)(0, 1, 0)<sub>4</sub>

Budujemy model SARIMA(0, 1, 4)(0, 1, 0)<sub>4</sub> z parametrem  $\lambda = 2$ .

```
set.seed(772)
model.ma <- Arima(uczący,
                  order = c(0, d, 4),
                  seasonal = list(order = c(0, D, 0), period = 4),
                  lambda = lambda)
model.ma

## Series: uczący
## ARIMA(0,1,4)(0,1,0)[4]
## Box Cox transformation: lambda= 2
##
## Coefficients:
##          ma1      ma2      ma3      ma4
##          0.2967  0.3566  0.4559 -0.6041
## s.e.    0.1819  0.1975  0.2009  0.1937
##
## sigma^2 = 1479:  log likelihood = -240.03
## AIC=490.06   AICc=491.53   BIC=499.31
```

#### 1.4.2.1 Istotność współczynników

Sprawdzamy teraz istotność współczynników tego modelu.

```
##
## z test of coefficients:
##
##      Estimate Std. Error z value Pr(>|z|)
## ma1  0.29666    0.18189  1.6310 0.102886
## ma2  0.35662    0.19750  1.8056 0.070974 .
## ma3  0.45589    0.20090  2.2692 0.023255 *
## ma4 -0.60405    0.19369 -3.1187 0.001816 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Współczynniki statystycznie istotne to `ma2`, `ma3` oraz `ma4`. Tworzymy zatem model uwzględniający tylko te zmienne.

```
## Series: uczący
## ARIMA(0,1,4)(0,1,0)[4]
## Box Cox transformation: lambda= 2
##
## Coefficients:
##      ma1      ma2      ma3      ma4
##      0  0.1916  0.2155  -0.8036
## s.e.    0  0.1673  0.1260   0.1567
##
## sigma^2 = 1566: log likelihood = -241.41
## AIC=490.82  AICc=491.77  BIC=498.22
```

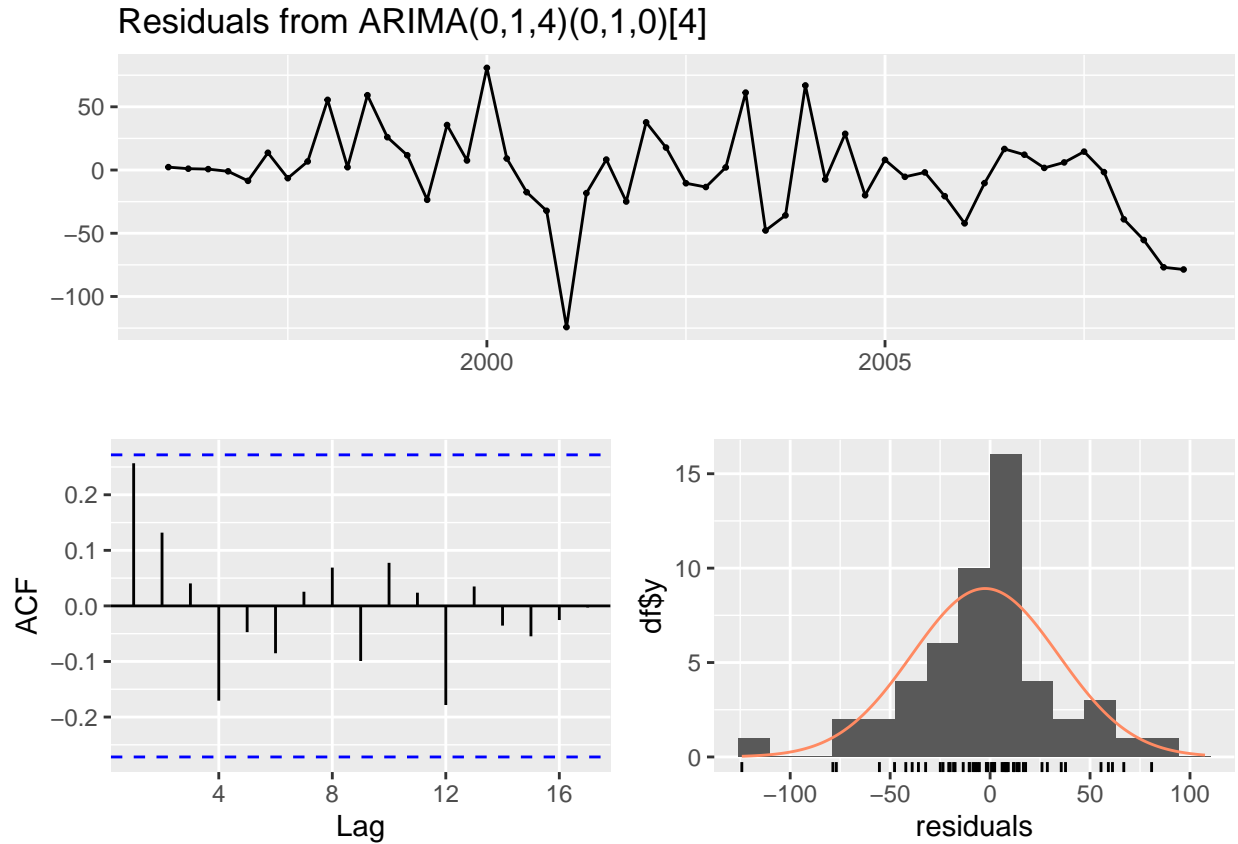
Sprawdzamy, którym model wypadł lepiej.

Tabela 1: Tabela porównująca wartości kryteriów informacyjnych AIC, AICc i BIC dla modelu SARIMA(0, 1, 4)(0, 1, 0)<sub>4</sub> oraz jego wersji z istotnymi współczynnikami.

Model	AIC	AICc	BIC
SARIMA(0,1,4)(0,1,0)[4]	490.0632	491.5266	499.3140
SARIMA(0,1,4)(0,1,0)[4] istotny	490.8163	491.7687	498.2169

Według kryteriów AIC i AICc model istotny jest minimalnie gorszy od zwykłego modelu. Wybieramy zatem podstawowy model SARIMA(0, 1, 4)(0, 1, 0)<sub>4</sub>.

**1.4.2.2 Diagnostyka** W celu weryfikacji poprawności modelu  $SARIMA(0, 1, 4)(0, 1, 0)_4$  przeprowadzimy analizę reszt.



Rysunek 7: Wykresy dla szeregu reszt modelu  $SARIMA(0, 1, 4)(0, 1, 0)_4$ : szereg czasowy (góra), funkcja autokorelacji ACF (dół-lewo) oraz histogram z krzywą gęstości rozkładu normalnego (dół-prawo).

```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(0,1,4)(0,1,0)[4]
## Q* = 7.3195, df = 4, p-value = 0.1199
##
## Model df: 4.    Total lags used: 8
```

Wartość poziomu krytycznego wskazuje, że nie ma podstaw do odrzucenia hipotezy o białoszumowości reszt tego modelu. Analiza wykresu ACF 7 również to potwierdza - brak istotnej autokorelacji. Model jest dobrze dopasowany do danych.

### 1.4.3 Model SARIMA(4,1,0)(0,1,0)<sub>4</sub>

Budujemy model SARIMA(4,1,0)(0,1,0)<sub>4</sub> z parametrem  $\lambda = 2$ .

```
set.seed(772)
model.ar <- Arima(uczający,
                  order = c(4, d, 0),
                  seasonal = list(order = c(0, D, 0), period = 4),
                  lambda = lambda)
model.ar

## Series: uczący
## ARIMA(4,1,0)(0,1,0)[4]
## Box Cox transformation: lambda= 2
##
## Coefficients:
##          ar1      ar2      ar3      ar4
##          0.1513  0.1829  0.0712 -0.5809
## s.e.      0.1179  0.1262  0.1216  0.1208
##
## sigma^2 = 1948: log likelihood = -243.43
## AIC=496.87   AICc=498.33   BIC=506.12
```

**1.4.3.1 Istotność współczynników** Sprawdamy teraz istotność współczynników tego modelu.

```
##
## z test of coefficients:
##
##      Estimate Std. Error z value Pr(>|z|)
## ar1  0.151250   0.117934  1.2825   0.1997
## ar2  0.182854   0.126248  1.4484   0.1475
## ar3  0.071231   0.121588  0.5858   0.5580
## ar4 -0.580880   0.120784 -4.8092 1.515e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Jedynym współczynnikiem statystycznie istotnym jest **ar4**. Tworzymy zatem model uwzględniający tylko tę zmienną.

```
## Series: uczący
## ARIMA(4,1,0)(0,1,0)[4]
## Box Cox transformation: lambda= 2
##
## Coefficients:
##      ar1  ar2  ar3      ar4
##      0    0    0  -0.5573
## s.e.    0    0    0   0.1267
##
## sigma^2 = 2005:  log likelihood = -245.59
## AIC=495.18   AICc=495.45   BIC=498.88
```

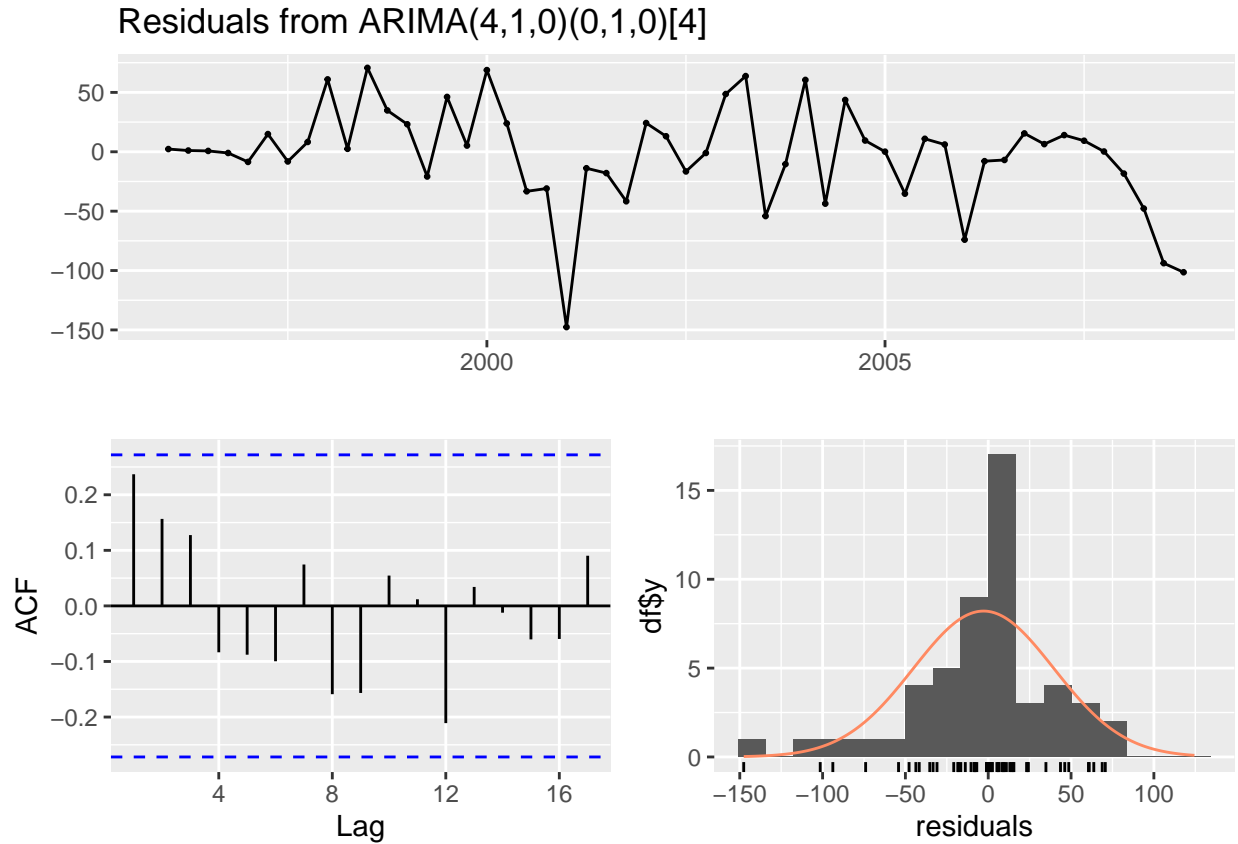
Sprawdzamy, którym model wypadł lepiej.

Tabela 2: Tabela porównująca wartości kryteriów informacyjnych AIC, AICc i BIC dla modelu SARIMA(4, 1, 0)(0, 1, 0)<sub>4</sub> oraz jego wersji z istotnymi współczynnikami.

Model	AIC	AICc	BIC
SARIMA(4,1,0)(0,1,0)[4]	496.8659	498.3293	506.1166
SARIMA(4,1,0)(0,1,0)[4] istotny	495.1753	495.4481	498.8756

Model SARIMA(4, 1, 0)(0, 1, 0)<sub>4</sub> istotny okazał się lepszy według wszystkich kryteriów informacyjnych.

**1.4.3.2 Diagnostyka** W celu weryfikacji poprawności modelu  $SARIMA(4, 1, 0)(0, 1, 0)_4$  istotnego przeprowadzimy analizę reszt.



Rysunek 8: Wykresy dla szeregu reszt modelu  $SARIMA(4, 1, 0)(0, 1, 0)_4$  istotnego: szereg czasowy (góra), funkcja autokorelacji ACF (dół-lewo) oraz histogram z krzywą gęstości rozkładu normalnego (dół-prawo).

```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(4,1,0)(0,1,0)[4]
## Q* = 8.827, df = 4, p-value = 0.06557
##
## Model df: 4.    Total lags used: 8
```

Wartość poziomu krytycznego wskazuje, że nie ma podstaw do odrzucenia hipotezy o białoszumowości reszt tego modelu. Analiza wykresu ACF 8 również to potwierdza - brak istotnej autokorelacja. Model jest dobrze dopasowany do danych.

#### 1.4.4 Model automatyczny

Budujemy model, korzystając z funkcji `auto.arima()` z argumentem  $\lambda = 2$ .

```
set.seed(772)
model.auto <- auto.arima(uczący,
                        lambda = lambda,
                        stepwise = FALSE)
model.auto

## Series: uczący
## ARIMA(1,0,3)(0,1,1)[4]
## Box Cox transformation: lambda= 2
##
## Coefficients:
##          ar1      ma1      ma2      ma3      sma1
##      0.9892  0.3095  0.2401  0.4171 -0.8596
## s.e.  0.0285  0.1590  0.1694  0.1811  0.1870
##
## sigma^2 = 1461: log likelihood = -244.62
## AIC=501.23  AICc=503.28  BIC=512.46
```

Automatycznie dopasowany model to SARIMA(1,0,3)(0,1,1)<sub>4</sub>.

**1.4.4.1 Istotność współczynników** Sprawdzamy teraz istotność współczynników tego modelu.

```
##
## z test of coefficients:
##
##      Estimate Std. Error z value Pr(>|z|)
## ar1  0.989192   0.028532 34.6698 < 2.2e-16 ***
## ma1  0.309483   0.159023  1.9462  0.05164 .
## ma2  0.240103   0.169446  1.4170  0.15649
## ma3  0.417059   0.181076  2.3032  0.02127 *
## sma1 -0.859557   0.186967 -4.5974 4.279e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Współczynniki statystycznie istotne to `ar1`, `ma1`, `ma3` i `sma1`. Tworzymy zatem model uwzględniający tylko te zmienne.

```
## Series: uczący
## ARIMA(1,0,3)(0,1,1)[4]
## Box Cox transformation: lambda= 2
##
## Coefficients:
##          ar1      ma1  ma2      ma3      sma1
##          0.9927  0.3040   0  0.3925  -0.8666
## s.e.    0.0212  0.1446   0  0.1957   0.1829
##
## sigma^2 = 1472:  log likelihood = -245.48
## AIC=500.97   AICc=502.39   BIC=510.32
```

Sprawdźmy, którym model wypadł lepiej.

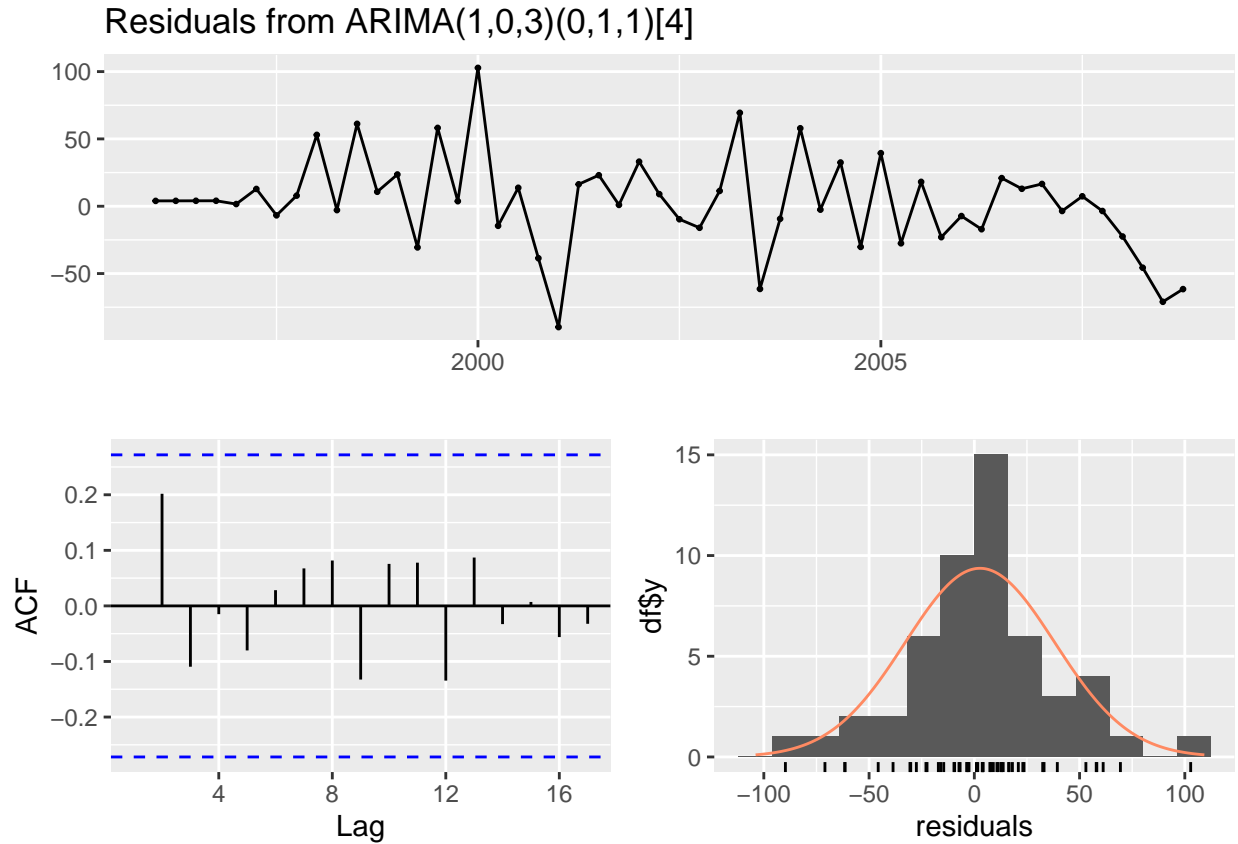
Tabela 3: Tabela porównująca wartości kryteriów informacyjnych AIC, AICc i BIC dla modelu SARIMA(1, 0, 3)(0, 1, 1)<sub>4</sub> oraz jego wersji z istotnymi współczynnikami.

Model	AIC	AICc	BIC
SARIMA(1,0,3)(0,1,1)[4]	501.2320	503.2808	512.4592
SARIMA(1,0,3)(0,1,1)[4] fixed	500.9662	502.3948	510.3222

Model istotny SARIMA(1, 0, 3)(0, 1, 1)<sub>4</sub> okazał się lepszy według wszystkich kryteriów informacyjnych.



**1.4.4.2 Diagnostyka** W celu weryfikacji poprawności modelu istotnego SARIMA(1,0,3)(0,1,1)<sub>4</sub> przeprowadzamy analizę reszt.



Rysunek 9: Wykresy dla szeregu reszt modelu SARIMA(1,0,3)(0,1,1)<sub>4</sub> istotnego: szereg czasowy (góra), funkcja autokorelacji ACF (dół-lewo) oraz histogram z krzywą gęstości rozkładu normalnego (dół-prawo).

```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(1,0,3)(0,1,1)[4]
## Q* = 4.1294, df = 3, p-value = 0.2478
##
## Model df: 5.    Total lags used: 8
```

Wartość poziomu krytycznego wskazuje, że nie ma podstaw do odrzucenia hipotezy o białoszumowości reszt tego modelu. Analiza wykresu ACF 9 również to potwierdza - brak istotnej autokorelacji. Model jest dobrze dopasowany do danych.

### 1.4.5 Porównanie

Na koniec porównamy wartości wszystkich kryteriów informacyjnych AIC, AICc i BIC dla wybranych lepszych modeli SARIMA. Postaramy się wybrać najlepszy z tej grupy modeli, który potem użyjemy do prognozowania naszego szeregu.

Tabela 4: Tabela porównująca wartości kryteriów informacyjnych AIC, AICc i BIC dla stworzonych modeli (S)ARIMA.

Model	AIC	AICc	BIC
SARIMA(0,1,4)(0,1,0)[4]	490.0632	491.5266	499.3140
SARIMA(4,1,0)(0,1,0)[4] fixed	495.1753	495.4481	498.8756
SARIMA(2, 0, 0)(1, 1, 0)[4]	501.2320	503.2808	512.4592

Jak widać w tabeli 4 według wszystkich kryteriów informacyjnych najlepszy okazał się model SARIMA(0, 1, 4)(0, 1, 0)<sub>4</sub>.

## 1.5 Modele dekompozycji

Teraz dopasujemy model oparty na dekompozycji. Skorzystamy w tym celu z funkcji `tslm`.

### 1.5.1 Model dekompozycji z trendem liniowym

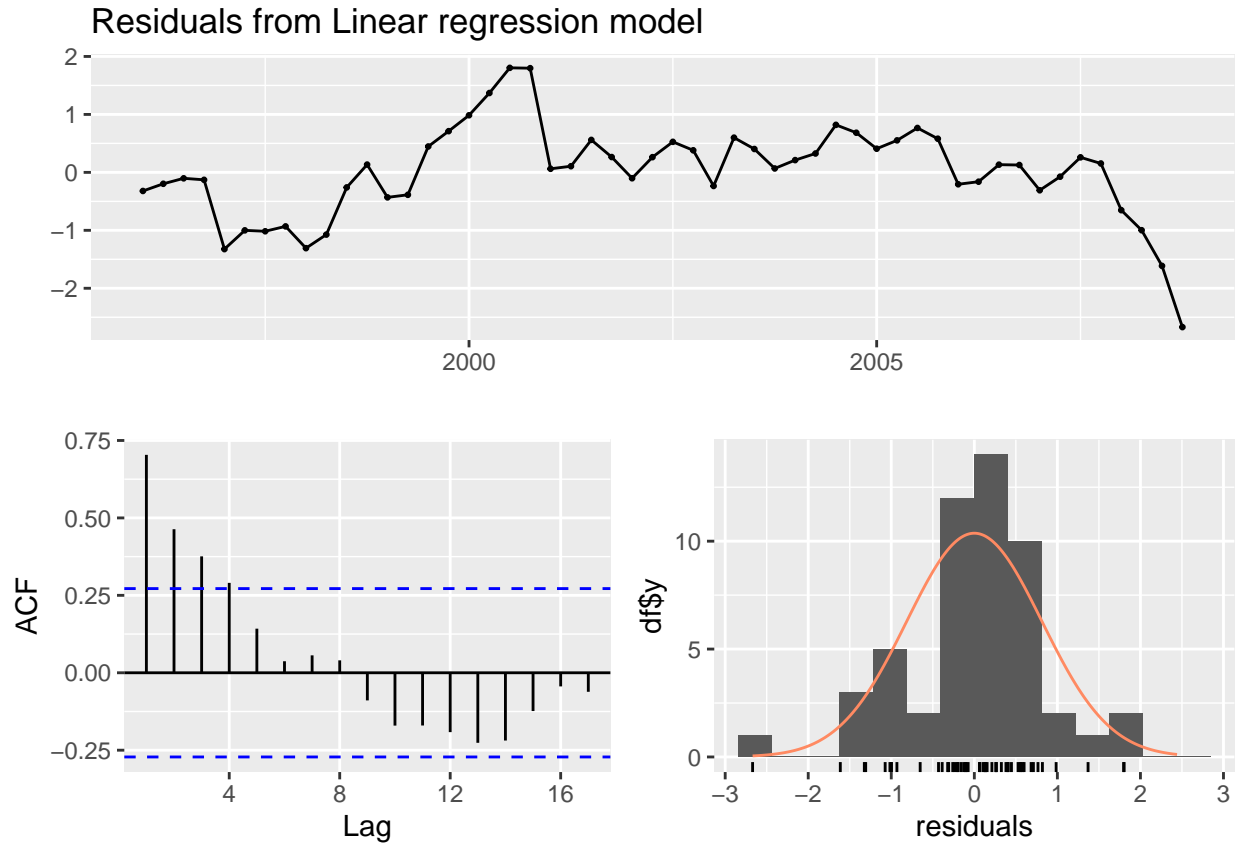
Zaczynamy od modeli dekompozycji z trendem liniowym.

```
set.seed(772)
model.lin <- tslm(uczący ~ trend)
sum.model.lin <- summary(model.lin)
sum.model.lin

##
## Call:
## tslm(formula = uczący ~ trend)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.6697 -0.3118  0.1166  0.4669  1.8035
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  89.185204   0.231233   385.69  <2e-16 ***
## trend         0.265856   0.007593    35.02  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8217 on 50 degrees of freedom
## Multiple R-squared:  0.9608, Adjusted R-squared:  0.96
## F-statistic: 1226 on 1 and 50 DF,  p-value: < 2.2e-16
```

Wartość współczynnika determinacji  $R^2$  wynosi 0.9608, co sugeruje, że model wyjaśnia 96.08 % zmienności naszych danych. Wartość poziomu krytycznego sugeruje, że parametr trendu jest statystycznie istotny w tym modelu.

**1.5.1.1 Diagnostyka** W celu weryfikacji poprawności modelu dekompozycji z trendem liniowym przeprowadzimy analizę reszt.



Rysunek 10: Wykresy dla szeregu reszt modelu dekompozycji z trendem liniowym: szereg czasowy (góra), funkcja autokorelacji ACF (dół-lewo) oraz histogram z krzywą gęstości rozkładu normalnego (dół-prawo).

```
##
## Breusch-Godfrey test for serial correlation of order up to 8
##
## data: Residuals from Linear regression model
## LM test = 38.232, df = 8, p-value = 6.82e-06
```

Wartość poziomu krytycznego wskazuje na odrzucenie hipotezy o białoszumowości reszt tego modelu. Analiza wykresu ACF 10 również to potwierdza - występuje istotna autokorelacja. Zatem model nie jest tak dobry, jak wydawało się na początku.

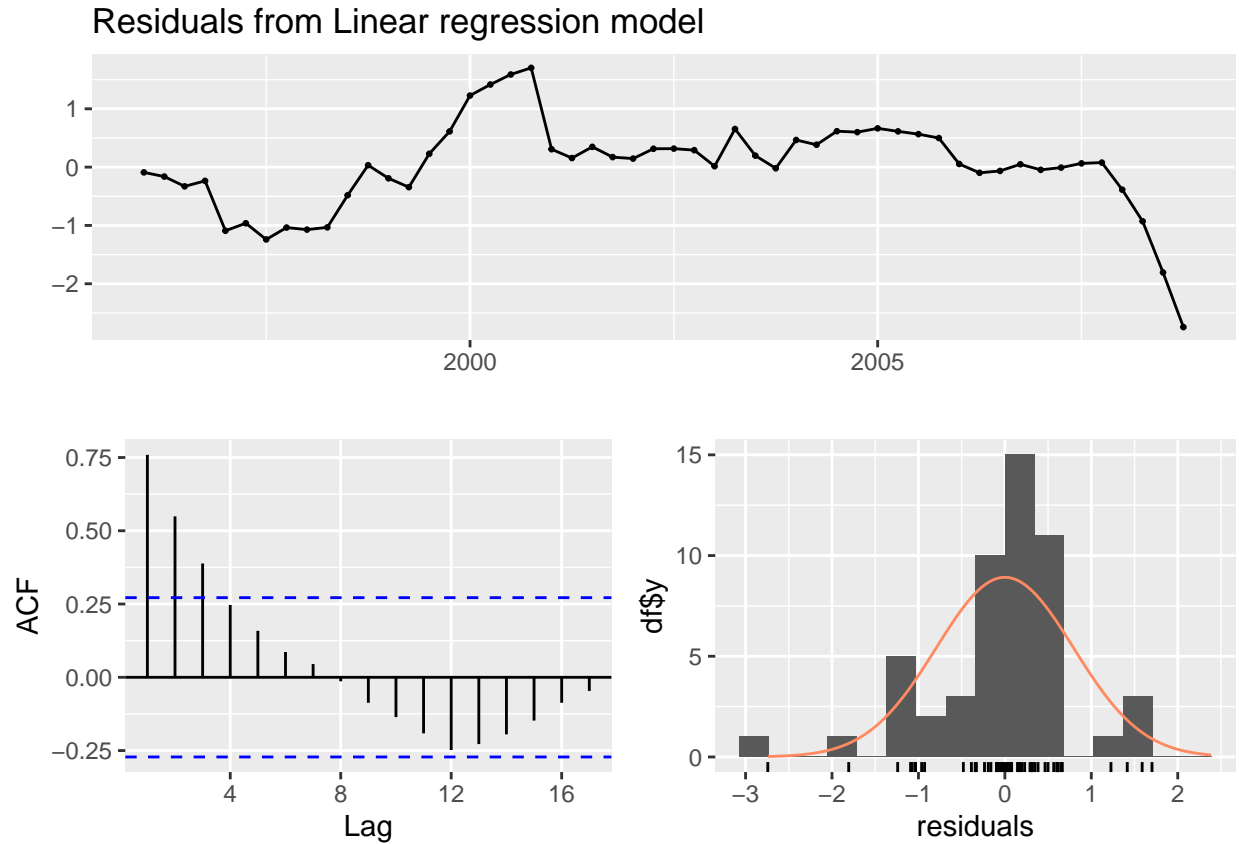
### 1.5.2 Model dekompozycji z trendem liniowym i sezonowością

Wiemy, że w naszych danych poza deterministycznym trendem wzrostowym obecna jest również sezonowość. Tworzymy zatem model z trendem liniowym oraz sezonowością.

```
##
## Call:
## tslm(formula = uczący ~ season + trend)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.74269 -0.25923  0.05894  0.40385  1.70192
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 88.955240   0.299310 297.201  <2e-16 ***
## season2      0.196394   0.324972   0.604   0.549
## season3      0.458942   0.325244   1.411   0.165
## season4      0.339952   0.325697   1.044   0.302
## trend        0.265144   0.007675  34.548  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8283 on 47 degrees of freedom
## Multiple R-squared:  0.9626, Adjusted R-squared:  0.9594
## F-statistic: 302.2 on 4 and 47 DF,  p-value: < 2.2e-16
```

Wartość współczynnika determinacji  $R^2$  wzrosła nieznacznie do poziomu 0.9626, co sugeruje, że model wyjaśnia 96.26 % zmienności naszych danych. Analiza istotności parametrów wskazuje jednak, że tylko parametr trendu jest statystycznie istotny. Skorygowany współczynnik determinacji ( $\text{Adjusted } R^2$ ) spadł z poziomu 0.96 do 0.9594, co potwierdza dodanie statystycznie nieistotnych zmiennych do modelu.

**1.5.2.1 Diagnostyka** W celu weryfikacji poprawności modelu dekompozycji z trendem liniowym i sezonowością przeprowadzimy analizę reszt.



Rysunek 11: Wykresy dla szeregu reszt modelu dekompozycji z trendem liniowym i sezonowością: szereg czasowy (góra), funkcja autokorelacji ACF (dół-lewo) oraz histogram z krzywą gęstości rozkładu normalnego (dół-prawo).

```
##
## Breusch-Godfrey test for serial correlation of order up to 8
##
## data: Residuals from Linear regression model
## LM test = 42.836, df = 8, p-value = 9.433e-07
```

Wartość poziomego krytycznego wskazuje na odrzucenie hipotezy o białoszumowości reszt tego modelu. Jego wartość jest jednak wyższa niż w przypadku modelu tylko z trendem liniowym. Analiza wykresu ACF 11 potwierdza występowanie istotnej autokorelacji. Ponadto histogram znacząco odbiega od gęstości rozkładu normalnego. Model nie jest dobry.

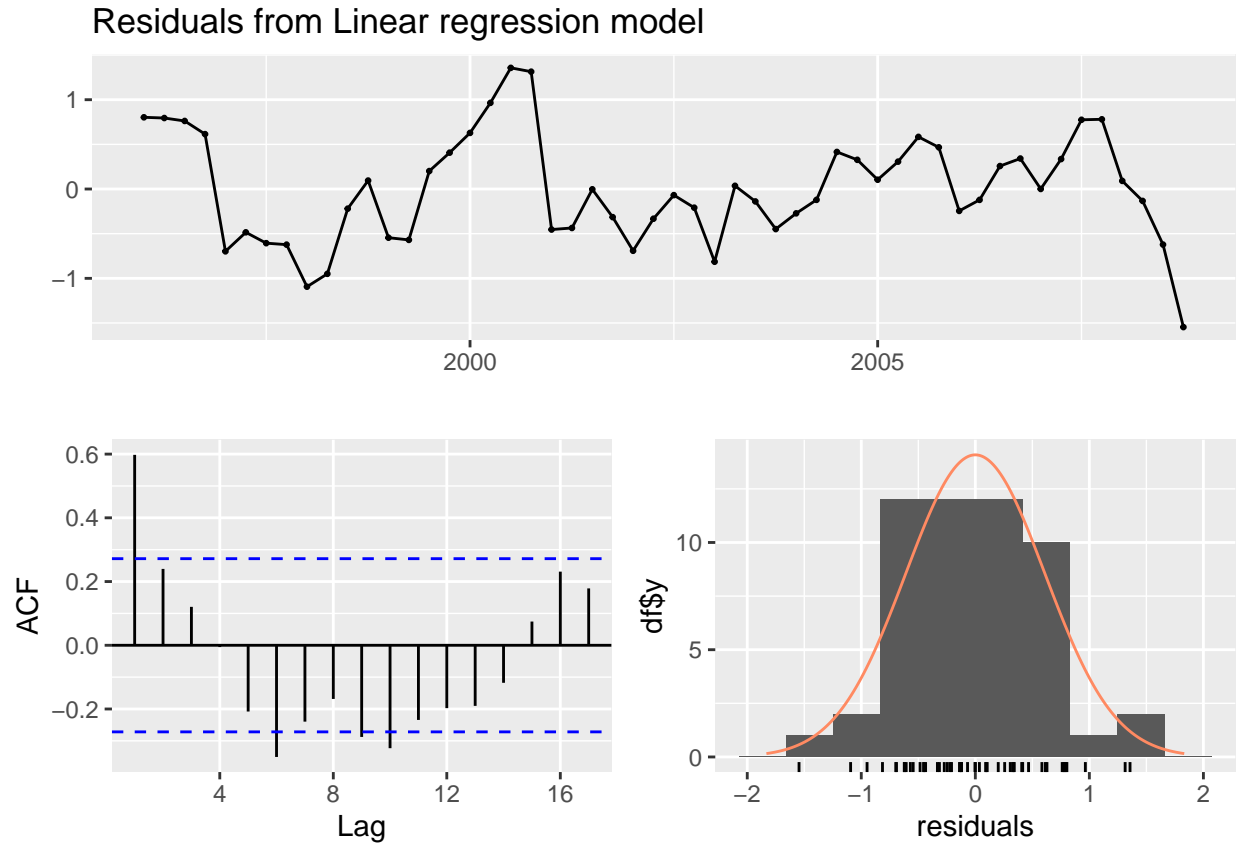
### 1.5.3 Model dekompozycji z trendem kwadratowym

Skoro dodanie sezonowości nie poprawiło jakości naszego modelu, spróbujemy zmienić postać trendu z liniowej na kwadratową.

```
##
## Call:
## tslm(formula = uczący ~ trend + I(trend^2))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.54579 -0.44967 -0.03575  0.40904  1.35661
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  87.923769   0.269472  326.282  < 2e-16 ***
## trend         0.406015   0.023457   17.309  < 2e-16 ***
## I(trend^2)   -0.002645   0.000429   -6.164  1.31e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.623 on 49 degrees of freedom
## Multiple R-squared:  0.9779, Adjusted R-squared:  0.977
## F-statistic: 1086 on 2 and 49 DF,  p-value: < 2.2e-16
```

Wartość współczynnika determinacji  $R^2$  wzrosła do poziomu 0.9779, co sugeruje, że model wyjaśnia 97.79 % zmienności naszych danych. Analiza istotności parametrów wskazuje, że parametry obu trendów są statystycznie istotne. Skorygowany współczynnik determinacji ( $\text{Adjusted } R^2$ ) wzrósł z poziomu 0.96 w modelu jedynie z trendem liniowym do 0.977, co potwierdza dodanie statystycznie istotnych zmiennych do modelu.

**1.5.3.1 Diagnostyka** W celu weryfikacji poprawności dekompozycji z trendem kwadratowym przeprowadzimy analizę reszt.



Rysunek 12: Wykresy dla szeregu reszt modelu dekompozycji z trendem kwadratowym: szereg czasowy (góra), funkcja autokorelacji ACF (dół-lewo) oraz histogram z krzywą gęstości rozkładu normalnego (dół-prawo).

```
##
## Breusch-Godfrey test for serial correlation of order up to 8
##
## data: Residuals from Linear regression model
## LM test = 29.653, df = 8, p-value = 0.0002434
```

Wartość poziomu krytycznego nadal wskazuje na odrzucenie hipotezy o białoszumowości reszt tego modelu, jednak jego wartość znacznie wzrosła. Analiza wykresu ACF 12 potwierdza występowanie istotnej autokorelacji. Model wydaje się być lepszy od poprzednich.

### 1.5.4 Model dekompozycji z trendem kwadratowym i sezonowością

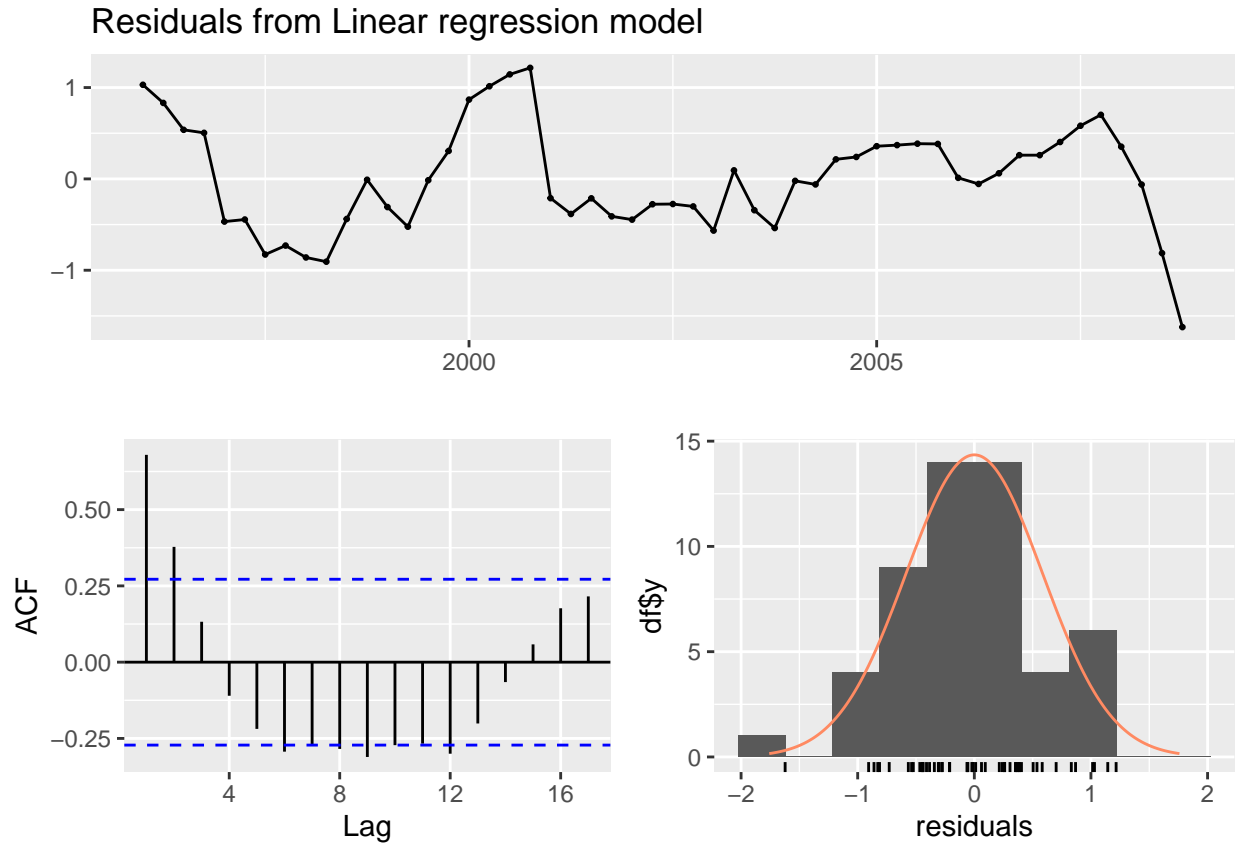
Spróbujmy jeszcze raz dodać sezonowość, tym razem do modelu z trendem kwadratowym.

```
##
## Call:
## tslm(formula = uczący ~ season + trend + I(trend^2))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.62221 -0.41715 -0.01786  0.37301  1.21568
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 87.6973451  0.3011220 291.235  < 2e-16 ***
## season2      0.1911090  0.2421452   0.789   0.4340
## season3      0.4536570  0.2423477   1.872   0.0676 .
## season4      0.3399519  0.2426833   1.401   0.1680
## trend        0.4052040  0.0232423  17.434  < 2e-16 ***
## I(trend^2)   -0.0026426  0.0004251  -6.217  1.37e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6172 on 46 degrees of freedom
## Multiple R-squared:  0.9797, Adjusted R-squared:  0.9775
## F-statistic: 443.2 on 5 and 46 DF,  p-value: < 2.2e-16
```

Wartość współczynnika determinacji  $R^2$  nieznacznie wzrosła do poziomu 0.9797, co sugeruje, że model wyjaśnia 97.97 % zmienności naszych danych. Analiza istotności parametrów wskazuje, że poza parametrami obu trendów statystycznie istotny jest również parametr `season3`. Skorygowany współczynnik determinacji (*Adjusted  $R^2$* ) również nieznacznie wzrósł z poziomu 0.977 w modelu jedynie z trendem kwadratowym do 0.9775, co potwierdza dodanie statystycznie istotnych zmiennych do modelu.



**1.5.4.1 Diagnostyka** W celu weryfikacji poprawności dekompozycji z trendem kwadratowym i sezonowością przeprowadzimy analizę reszt.



Rysunek 13: Wykresy dla szeregu reszt modelu dekompozycji z trendem kwadratowym i sezonowością: szereg czasowy (góra), funkcja autokorelacji ACF (dół-lewo) oraz histogram z krzywą gęstości rozkładu normalnego (dół-prawo).

```
##
## Breusch-Godfrey test for serial correlation of order up to 9
##
## data: Residuals from Linear regression model
## LM test = 35.553, df = 9, p-value = 4.758e-05
```

Wartość poziomu krytycznego nadal wskazuje na odrzucenie hipotezy o białoszumowości reszt tego modelu. Analiza wykresu ACF 13 potwierdza występowanie istotnej autokorelacji.

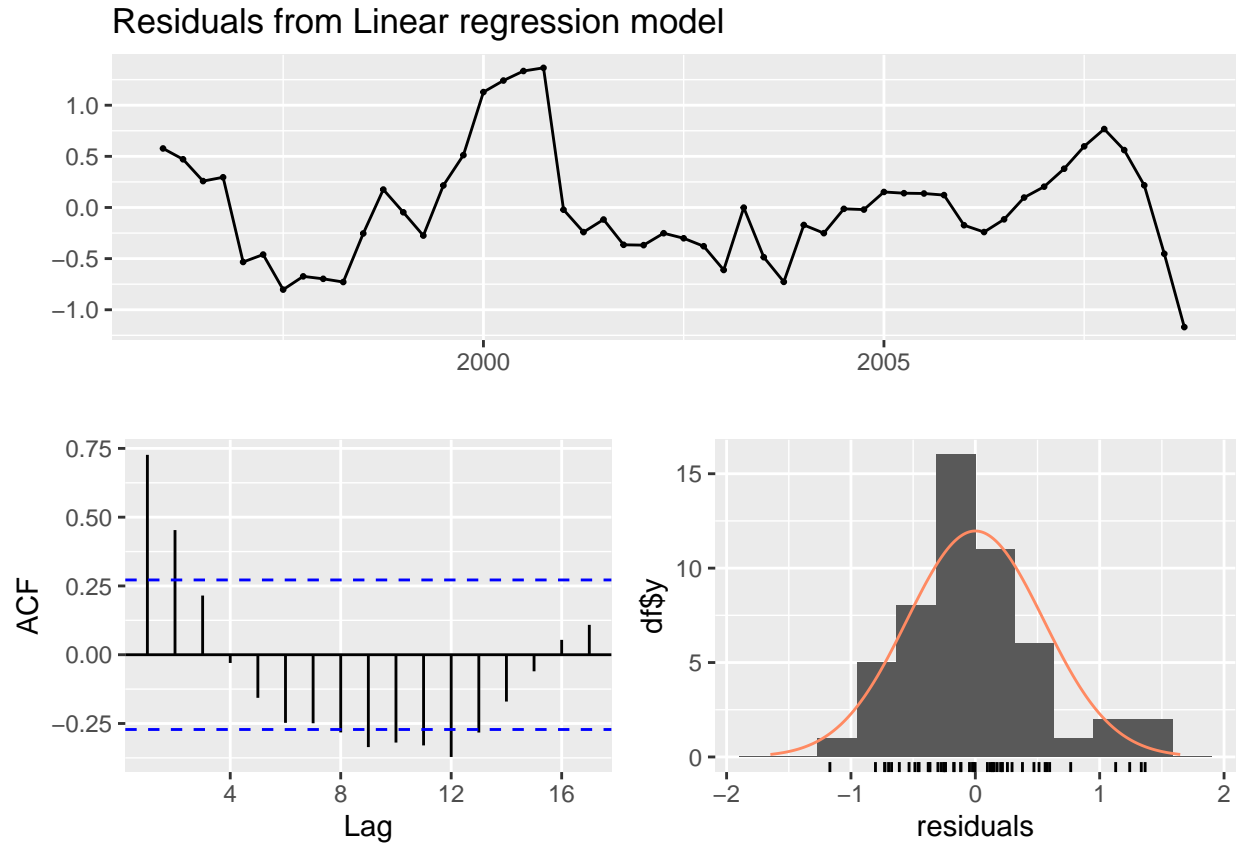
### 1.5.5 Model dekompozycji z trendem sześciennym i sezonowością

Na koniec stworzymy jeszcze model z trendem sześciennym i sezonowością. Nie poprawi to sytuacji z białoszumowością reszt, ale może pozwoli na lepsze dopasowanie trendu do danych.

```
##
## Call:
## tslm(formula = uczący ~ season + trend + I(trend^2) + I(trend^3))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.16942 -0.36561 -0.03433  0.22728  1.36538
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  8.828e+01  3.671e-01 240.453  < 2e-16 ***
## season2      2.120e-01  2.296e-01   0.924  0.3606
## season3      4.952e-01  2.302e-01   2.151  0.0368 *
## season4      4.024e-01  2.313e-01   1.740  0.0887 .
## trend        2.741e-01  5.687e-02   4.819 1.68e-05 ***
## I(trend^2)    3.479e-03  2.481e-03   1.402  0.1676
## I(trend^3)   -7.701e-05  3.079e-05  -2.501  0.0161 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5847 on 45 degrees of freedom
## Multiple R-squared:  0.9821, Adjusted R-squared:  0.9798
## F-statistic: 412.6 on 6 and 45 DF,  p-value: < 2.2e-16
```

Wartość współczynnika determinacji  $R^2$  wzrosła do poziomu 0.9821, co sugeruje, że model wyjaśnia 98.21 % zmienności naszych danych. Analiza istotności parametrów wskazuje, poza parametrami trendu liniowego oraz sześciennego statystycznie istotne są również parametry **season3** i **season4**. Skorygowany współczynnik determinacji (**Adjusted  $R^2$** ) nieznacznie wzrósł z poziomu 0.9775 w modelu jedynie z trendem kwadratowym do 0.9798, co potwierdza dodanie statystycznie istotnych zmiennych do modelu. Przeprowadzimy jeszcze analizę reszt.

**1.5.5.1 Diagnostyka** W celu weryfikacji poprawności dekompozycji z trendem sześciennym i sezonowością przeprowadzimy analizę reszt.



Rysunek 14: Wykresy dla szeregu reszt modelu dekompozycji z trendem sześciennym i sezonowością: szereg czasowy (góra), funkcja autokorelacji ACF (dół-lewo) oraz histogram z krzywą gęstości rozkładu normalnego (dół-prawo).

```
##
## Breusch-Godfrey test for serial correlation of order up to 10
##
## data: Residuals from Linear regression model
## LM test = 37.207, df = 10, p-value = 5.212e-05
```

Tak jak przypuszczaliśmy, wartość poziomu krytycznego nadal wskazuje na odrzucenie hipotezy o białoszumowości reszt tego modelu. Analiza wykresu ACF 14 potwierdza występowanie autokorelacji. Natomiast histogram reszt lepiej dopasowuje się do krzywej gęstości rozkładu normalnego. Wybieramy ten model, jako najlepszy model dekompozycji.

## 1.6 Algorytmy wygładzenia wykładniczego

Idea tych algorytmów opiera się na przypisaniu obserwacjom ze zbioru uczącego wykładniczo zanikających wag. Modele te służą od razu do tworzenia prognoz dla zadanych danych. Zatem w tej części raportu jedynie stworzymy nasze modele, a dopiero później przyjrzymy się prognozom.

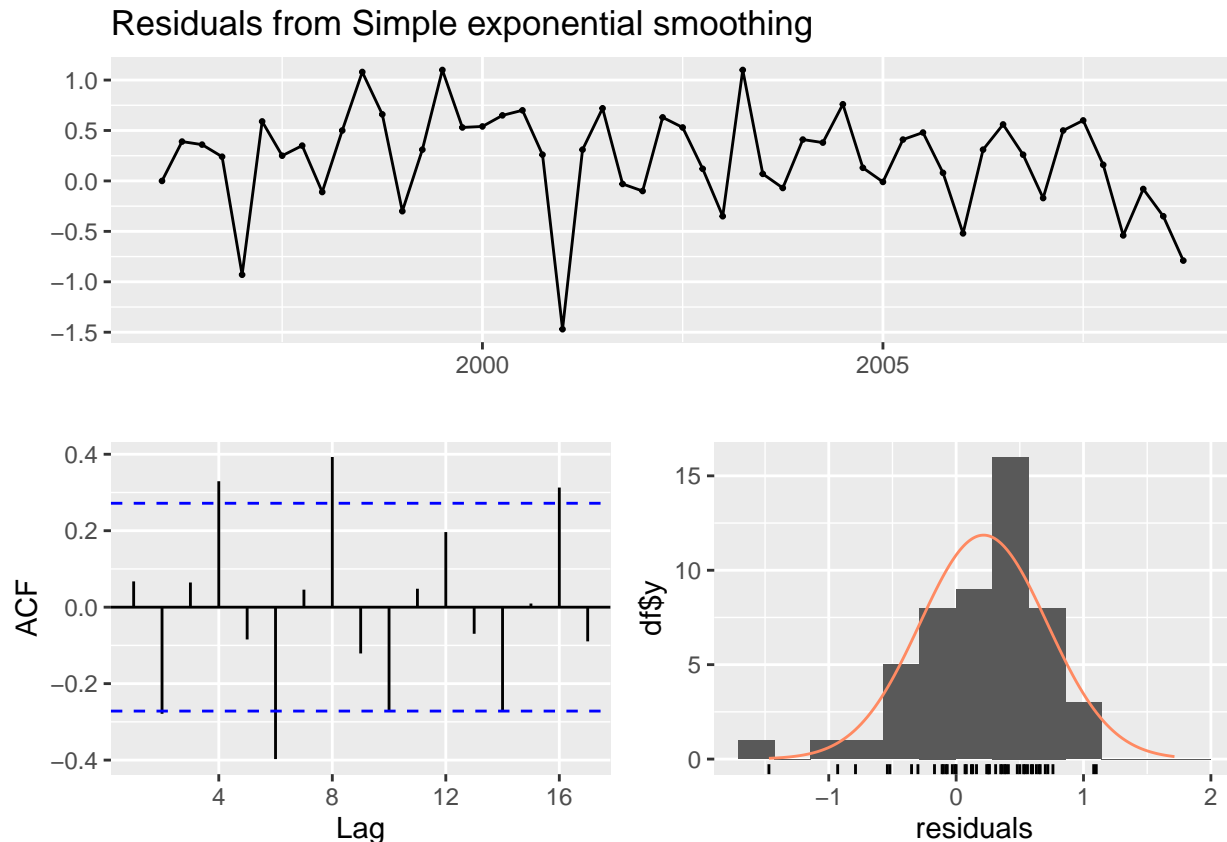
Na podstawie porównania wartości kryteriów informacyjnych wybierzemy model najlepiej dopasowany dla naszych danych.

### 1.6.1 Proste wygładzenie wykładnicze (algorytm SES)

Algorytm stosowany jest dla szeregów bez wyraźnego trendu i sezonowości. W danych `euretail` obecne są oba te parametry, więc stworzymy ten model jedynie kontrolnie.

```
set.seed(772)
model.ses <- ses(uczący, h = h, initial = "simple")
```

**1.6.1.1 Diagnostyka** W celu weryfikacji poprawności algorytmu SES przeprowadzimy analizę reszt.



Rysunek 15: Wykresy dla szeregu reszt modelu SES: szereg czasowy (góra), funkcja autokorelacji ACF (dół-lewo) oraz histogram z krzywą gęstości rozkładu normalnego (dół-prawo).

```
##
##  Ljung-Box test
##
## data:  Residuals from Simple exponential smoothing
## Q* = 31.25, df = 8, p-value = 0.0001268
##
## Model df: 0.   Total lags used: 8
```

Wartość poziomu krytycznego wskazuje na odrzucenie hipotezy o białoszumowości reszt tego modelu. Analiza wykresu ACF 15 również to potwierdza - występuje istotna autokorelacja. Możemy również zaobserwować sezonowość reszt tego modelu. Tak jak wspominałyśmy model ten nie jest odpowiedni dla tych danych.

## 1.6.2 Algorytm Holta

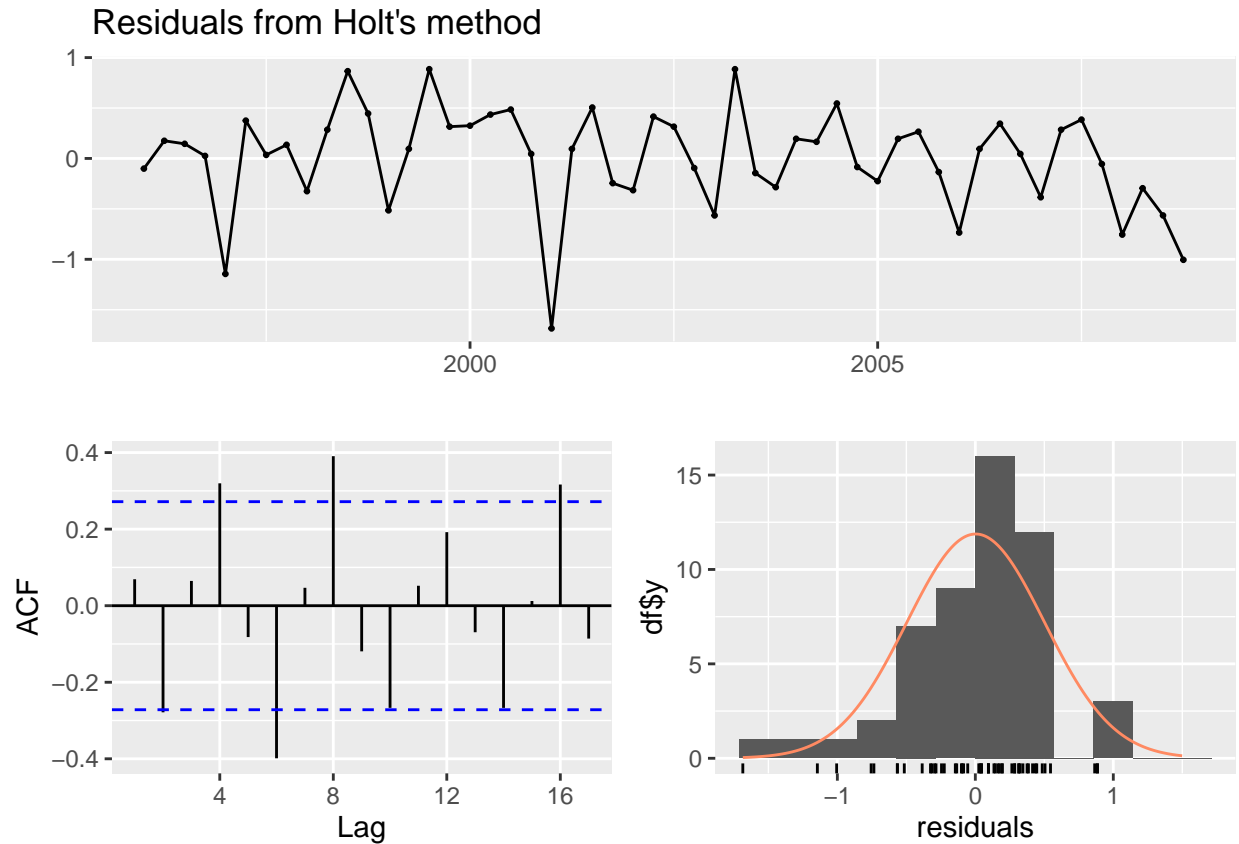
Metoda stosowana jest dla szeregów z wyraźnym trendem, ale bez sezonowości. W danych *euretail* sezonowość jest obecna, więc stworzymy ten model kontrolnie w dwóch wariantach:

- z domyślnym parametrem trendu,
- z trendem tłumionym.

### 1.6.3 Domyślny model Holta

```
set.seed(772)
model.holt <- holt(uczający, h = h)
```

**1.6.3.1 Diagnostyka** W celu weryfikacji poprawności modelu Holta z domyślnym parametrem trendu przeprowadzimy analizę reszt.



Rysunek 16: Wykresy dla szeregu reszt domyślnego modelu Holta: szereg czasowy (góra), funkcja autokorelacji ACF (dół-lewo) oraz histogram z krzywą gęstości rozkładu normalnego (dół-prawo).

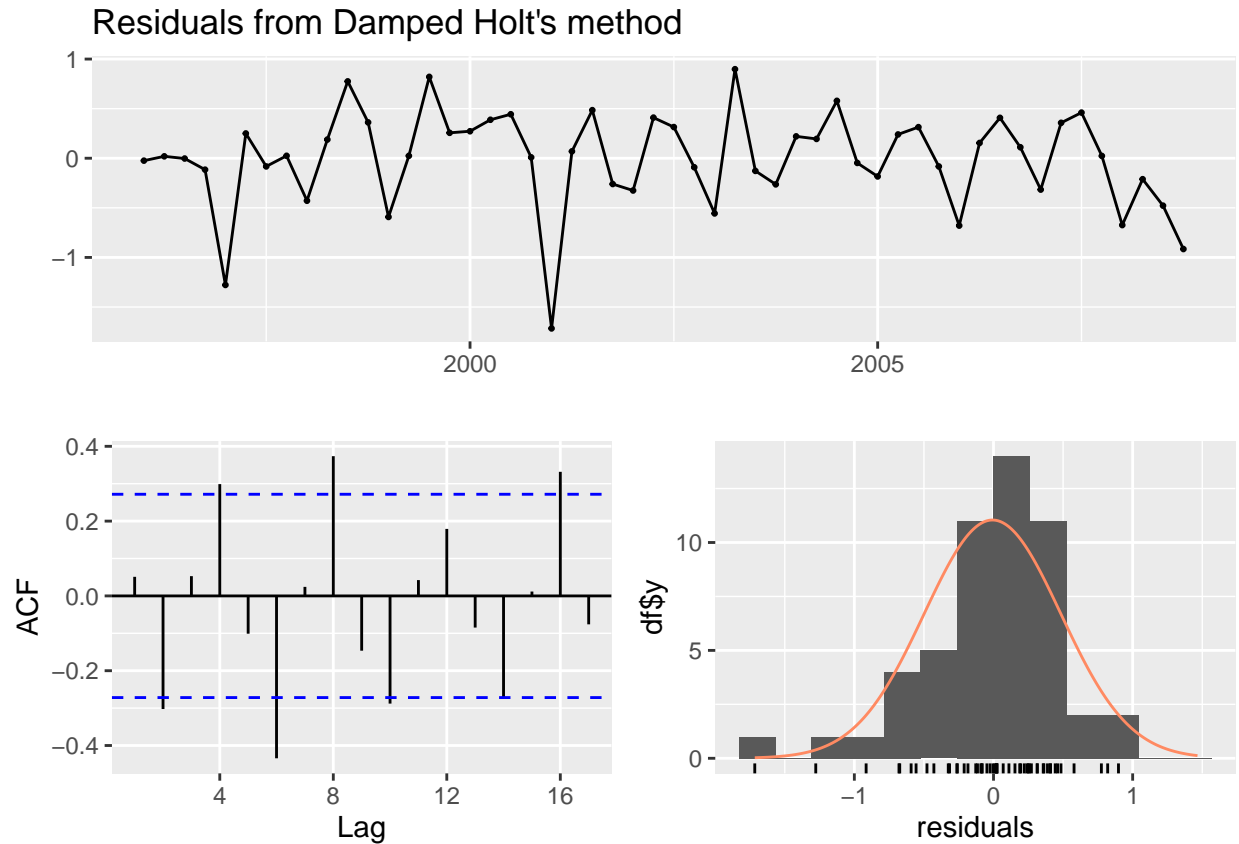
```
##
##  Ljung-Box test
##
## data:  Residuals from Holt's method
## Q* = 30.813, df = 8, p-value = 0.0001517
##
## Model df: 0.   Total lags used: 8
```

Wartość poziomu krytycznego wskazuje na odrzucenie hipotezy o białoszumowości reszt tego modelu. Analiza wykresu ACF 16 również to potwierdza - występuje istotna autokorelacja. Możemy również zaobserwować sezonowość reszt tego modelu. Tak, jak wspominaliśmy, model ten nie jest odpowiedni dla tych danych.

#### 1.6.4 Model Holta z trendem tłumionym

```
set.seed(772)
model.holt.d <- holt(uczący, h = h, damped = TRUE)
```

**1.6.4.1 Diagnostyka** W celu weryfikacji poprawności modelu Holta z trendem tłumionym przeprowadzimy analizę reszt.



Rysunek 17: Wykresy dla szeregu reszt modelu Holta z trendem tłumionym: szereg czasowy (góra), funkcja autokorelacji ACF (dół-lewo) oraz histogram z krzywą gęstości rozkładu normalnego (dół-prawo).

```
##
##  Ljung-Box test
##
## data:  Residuals from Damped Holt's method
## Q* = 31.722, df = 8, p-value = 0.0001045
##
## Model df: 0.   Total lags used: 8
```

Wartość poziomu krytycznego wskazuje na odrzucenie hipotezy o białoszumowości reszt tego modelu. Analiza wykresu ACF 16 również to potwierdza - występuje istotna autokorelacja. Możemy również zaobserwować sezonowość reszt tego modelu. Tak, jak wspominaliśmy, model ten nie jest odpowiedni dla tych danych i wypada gorzej niż model z domyślnym parametrem trendu.

### 1.6.5 Algorytm Holta-Wintersa

Metoda stosowana jest dla szeregów z wyraźnym trendem i sezonowością. W danych `euretail` obecne są oba te parametry, więc możemy już teraz przypuszczać, że ten algorytm będzie najlepszy. Stworzymy dwa warianty tej metody:

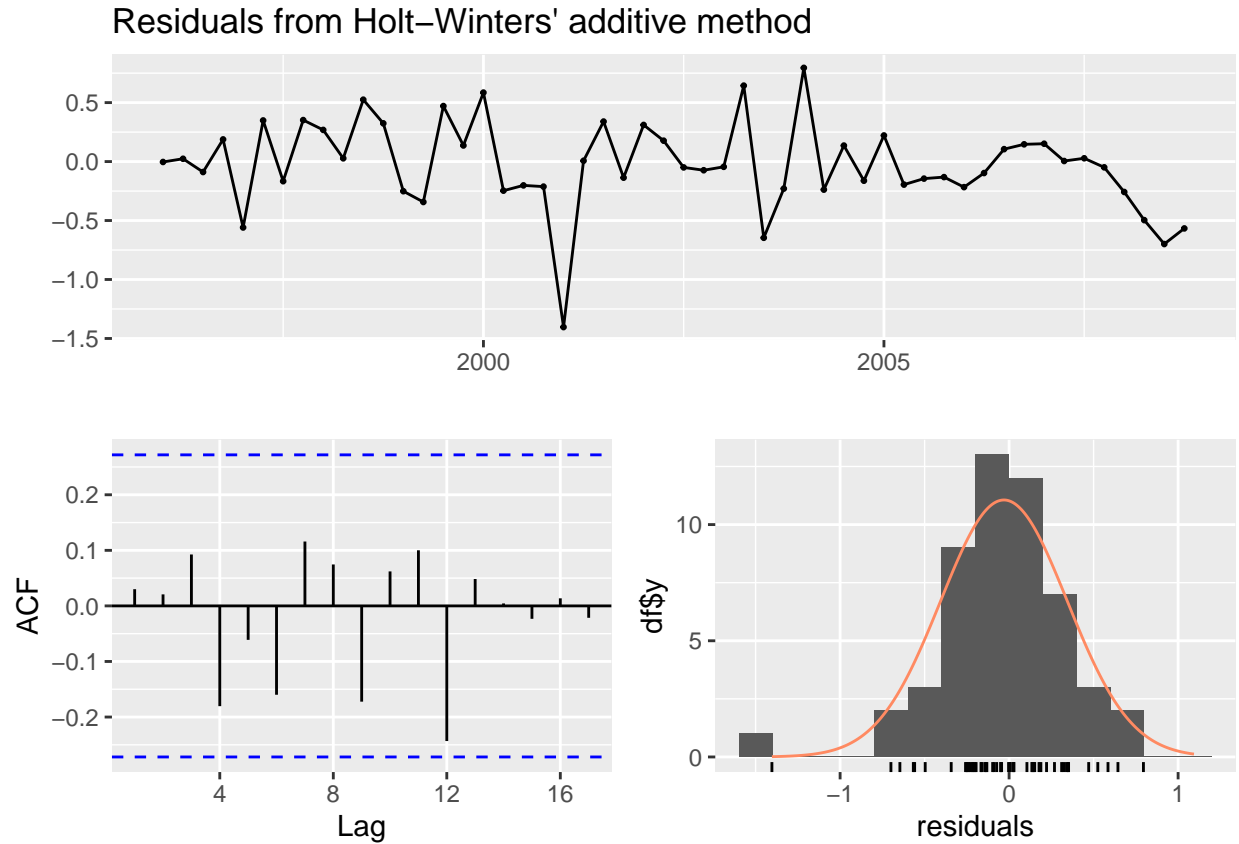
- z sezonowością addytywną,
- z sezonowością multiplikatywną.

### 1.6.6 Model Holta-Wintersa z sezonowością addytywną

```
set.seed(772)
model.hw.add <- hw(uczący, seasonal = "additive", h = h)
```



**1.6.6.1 Diagnostyka** W celu weryfikacji poprawności modelu Holta-Wintersa z sezonowością addytywną przeprowadzimy analizę reszt.



Rysunek 18: Wykresy dla szeregu reszt modelu Holta-Wintersa z sezonowością addytywną: szereg czasowy (góra), funkcja autokorelacji ACF (dół-lewo) oraz histogram z krzywą gęstości rozkładu normalnego (dół-prawo).

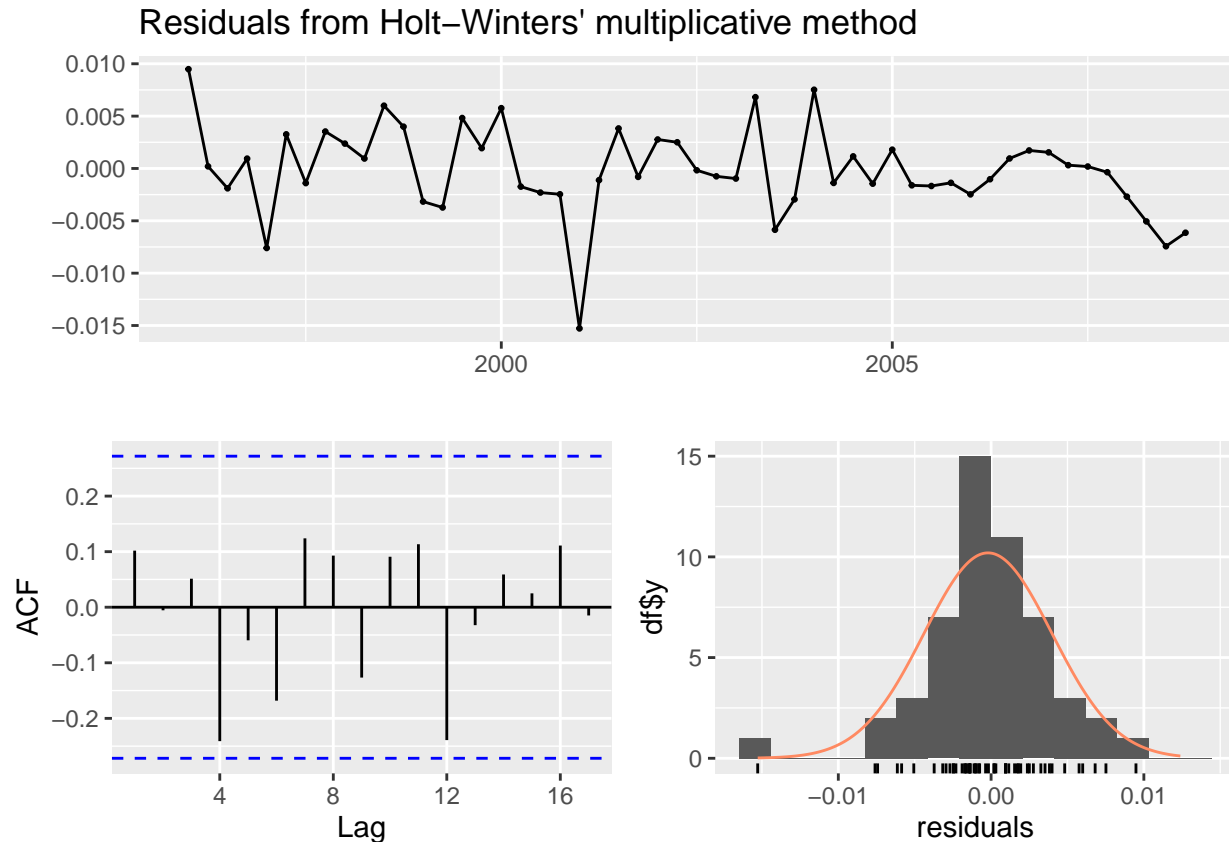
```
##
##  Ljung-Box test
##
## data:  Residuals from Holt-Winters' additive method
## Q* = 5.4445, df = 8, p-value = 0.7092
##
## Model df: 0.   Total lags used: 8
```

Wartość poziomu krytycznego wskazuje na brak podstaw do odrzucenia hipotezy o białoszumowości reszt tego modelu. Analiza wykresu ACF 18 również to potwierdza - brak istotnej autokorelacji. Dopasowanie histogramu do krzywej gęstości rozkładu normalnego sugeruje rozkład normalny reszt.

### 1.6.7 Model Holta-Wintersa z sezonowością multiplikatywną

```
set.seed(772)
model.hw.mult <- hw(uczający, seasonal = "multiplicative", h = h)
```

**1.6.7.1 Diagnostyka** W celu weryfikacji poprawności modelu Holta-Wintersa z sezonowością multiplikatywną przeprowadzimy analizę reszt.



Rysunek 19: Wykresy dla szeregu reszt modelu Holta-Wintersa z sezonowością multiplikatywną: szereg czasowy (góra), funkcja autokorelacji ACF (dół-lewo) oraz histogram z krzywą gęstości rozkładu normalnego (dół-prawo).

```
##
##  Ljung-Box test
##
## data:  Residuals from Holt-Winters' multiplicative method
## Q* = 7.5672, df = 8, p-value = 0.4768
##
## Model df: 0.   Total lags used: 8
```

Wartość poziomu krytycznego wskazuje na brak podstaw do odrzucenia hipotezy o białoszumowości reszt tego modelu. Analiza wykresu ACF 19 również to potwierdza - brak istotnej autokorelacji. Dopasowanie histogramu do krzywej gęstości rozkładu normalnego sugeruje rozkład normalny reszt.

### 1.6.8 Model automatyczny

Skorzystamy z funkcji `ets`, która automatycznie dobierze model ETS (opcja `model = "ZZZ"`) do naszych danych. Algorytm na podstawie kryterium AICc wskaże optymalny model, sprawdzając każdą dostępną opcję.

Wybrany w ten sposób model to ETS(A,A,A). Jest to dokładnie to ten sam model co algorytm Holta-Wintersa z sezonowością addytywną. Nie będziemy zatem poddawać modelu automatycznego dalszym analizom.

### 1.6.9 Porównanie

Porównamy teraz wartości kryteriów informacyjnych AIC, AICc i BIC dla modelu Holta-Wintersa z:

- sezonowością addytywną,
- sezonowością multiplikatywną.

Postaramy się wybrać najlepszy model.

Tabela 5: Tabela porównująca wartości kryteriów informacyjnych AIC, AICc i BIC dla algorytmów Holta-Wintersa z sezonowością addytywną i multiplikatywną.

Model	AIC	AICc	BIC
Holt-Winters add.	120.7754	125.0611	138.3366
Holt-Winters mult.	128.2209	132.5067	145.7821

Wartości wszystkich kryteriów informacyjnych (tabela 5) wskazują, że algorytm Holta-Wintersa z sezonowością addytywną jest najlepiej dopasowany do naszych danych. Wykorzystamy zatem ten algorytm do tworzenia prognoz naszego szeregu.

## 1.7 Prognozowanie

Wyznaczymy teraz prognozy dla zbioru testowego na podstawie dopasowanych modeli:

- SARIMA(0, 1, 4)(0, 1, 0)<sub>4</sub>,
- model dekompozycji z trendem sześciennym i sezonowością,
- algorytm Holta-Wintersa z sezonowością addytywną

oraz wybranej metody referencyjnej - sezonowej metody naiwnej, przyjmującej wartości z odpowiadającego kwartału poprzedniego roku.

Skonstruowane prognozy przedstawimy na wykresie i porównamy z wartościami rzeczywistymi ze zbioru testowego. Na koniec porównamy dokładność prognoz dla zbioru testowego i uczącego z uwzględnieniem wybranych miar oceny dokładności.

### 1.7.1 Funkcja prognoza

Tworzymy pomocniczą funkcję `prognoza`, przyjmującą argumenty:

- `model` - zadany model, na podstawie którego będziemy prognozować (domyślnie `NULL`),
- `zbiór.uczący` - zbiór uczący (domyślnie `NULL`, ponieważ argument jest potrzebny jedynie w przypadku metod naiwnych),
- `zbiór.testowy` - zbiór testowy, do którego będziemy porównywać naszą prognozę,
- `opóźnienie` - horyzont prognozy - liczba okresów, dla których wyznaczana jest prognoza (domyślnie długość zbioru testowego),
- `naiwna` - parametr decydujący o tworzeniu prognozy dla metody naiwnej (domyślnie `FALSE`),
- `naiwna.sezon` - parametr decydujący o tworzeniu prognozy dla sezonowej metody naiwnej (domyślnie `FALSE`),
- `wykres` - parametr decydujący o wygenerowaniu wykresu porównującego prognozę z wartościami rzeczywistymi (domyślnie `FALSE`),
- `error` - parametr decydujący o wywołaniu zestawienia dokładności - funkcja `accuracy` z pakietu `forecast` (domyślnie `FALSE`).

Funkcja tworzy prognozę dla wybranego modelu i w zależności od parametrów `wykres` i `error` zwraca wynik.

```

prognoza <- function(model = NULL, zbiór.uczący = NULL, zbiór.testowy,
                     opóźnienie = length(zbiór.testowy),
                     naiwna = FALSE, naiwna.sezon = FALSE,
                     wykres = FALSE, error = FALSE){

  prognoza <- forecast::forecast(model, h = opóźnienie)

  if (naiwna == TRUE){
    prognoza <- naive(x = zbiór.uczący, h = opóźnienie)
  }

  if (naiwna.sezon == TRUE){
    prognoza <- snaive(x = zbiór.uczący, h = opóźnienie)
  }

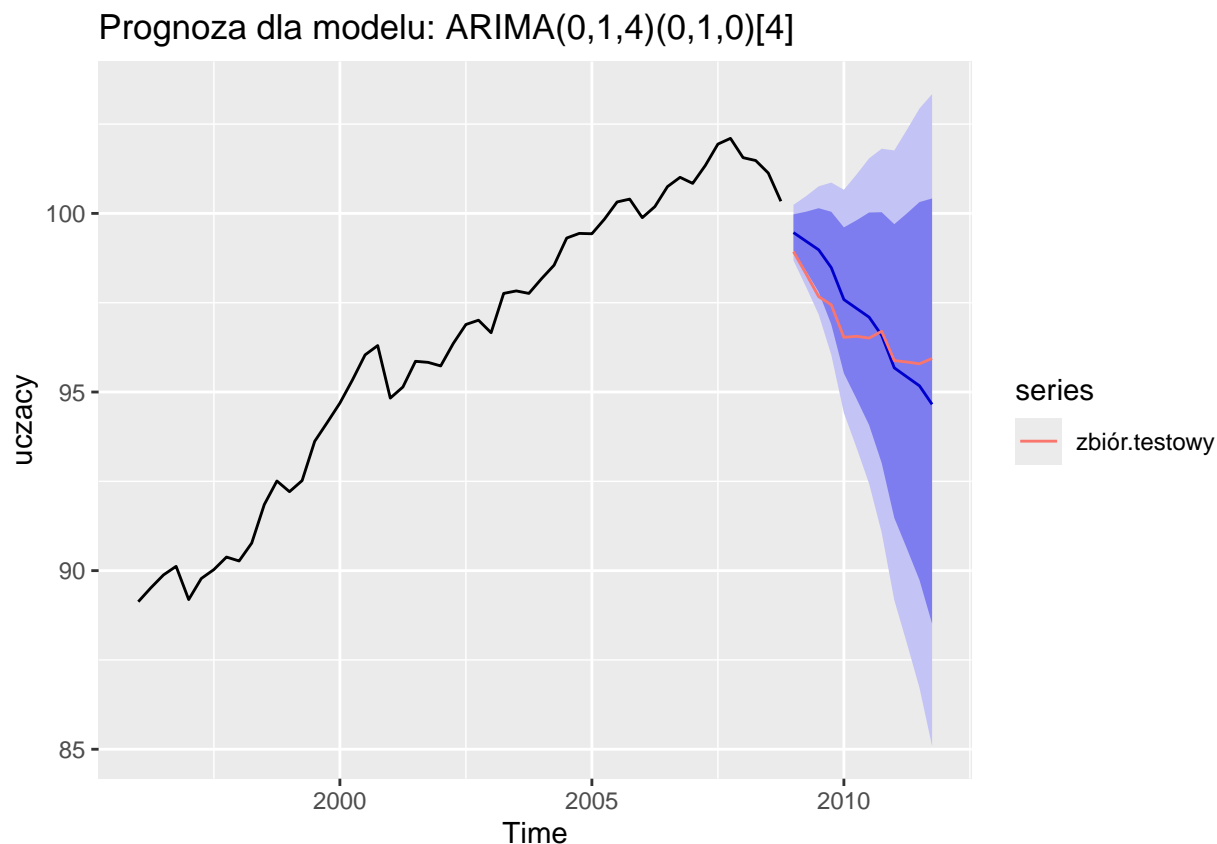
  prognoza.wykres <- autoplot(prognoza) +
    autolayer(zbiór.testowy) +
    ggtitle(paste("Prognoza dla modelu:", prognoza$method))
  prognoza.dokładność <- accuracy(prognoza, zbiór.testowy)

  if (wykres == TRUE){
    return(prognoza.wykres)
  } else if (error == TRUE){
    return(prognoza.dokładność)
  } else {
    return(prognoza)
  }
}

```

### 1.7.2 Model SARIMA(0, 1, 4)(0, 1, 0)<sub>4</sub>

Prognostowanie zaczynamy od modelu SARIMA(0, 1, 4)(0, 1, 0)<sub>4</sub>.



Rysunek 20: Wykres przedstawiający prognozy wyznaczone na podstawie modelu SARIMA(0, 1, 4)(0, 1, 0)<sub>4</sub> oraz przedziały ufności prognoz (80% i 95%). Czerwoną linią zaznaczono rzeczywiste wartości ze zbioru testowego.

Jak widać na wykresie 20 model dobrze przewidział tendencję spadkową zapoczątkowaną na końcu zbioru uczącego (granatowa linia). Rzeczywiste wartości indeksu ze zbioru testowego (czerwona linia) w całości mieszczą się w wyznaczonych przedziałach predykcji. Model można uznać za wiarygodny, chociaż wraz ze wzrostem horyzontu prognozy przedziały predykcji rozszerzają się, co sugeruje znaczną niepewność modelu w miarę oddalania się od ostatniej znanej obserwacji.

Przyjrzyjmy się teraz dokładności prognoz tego modelu na podstawie miar oceny dokładności.

Tabela 6: Tabela accuracy przedstawiająca wartości miar oceny dokładności prognoz dla modelu SARIMA(0, 1, 4)(0, 1, 0)<sub>4</sub>.

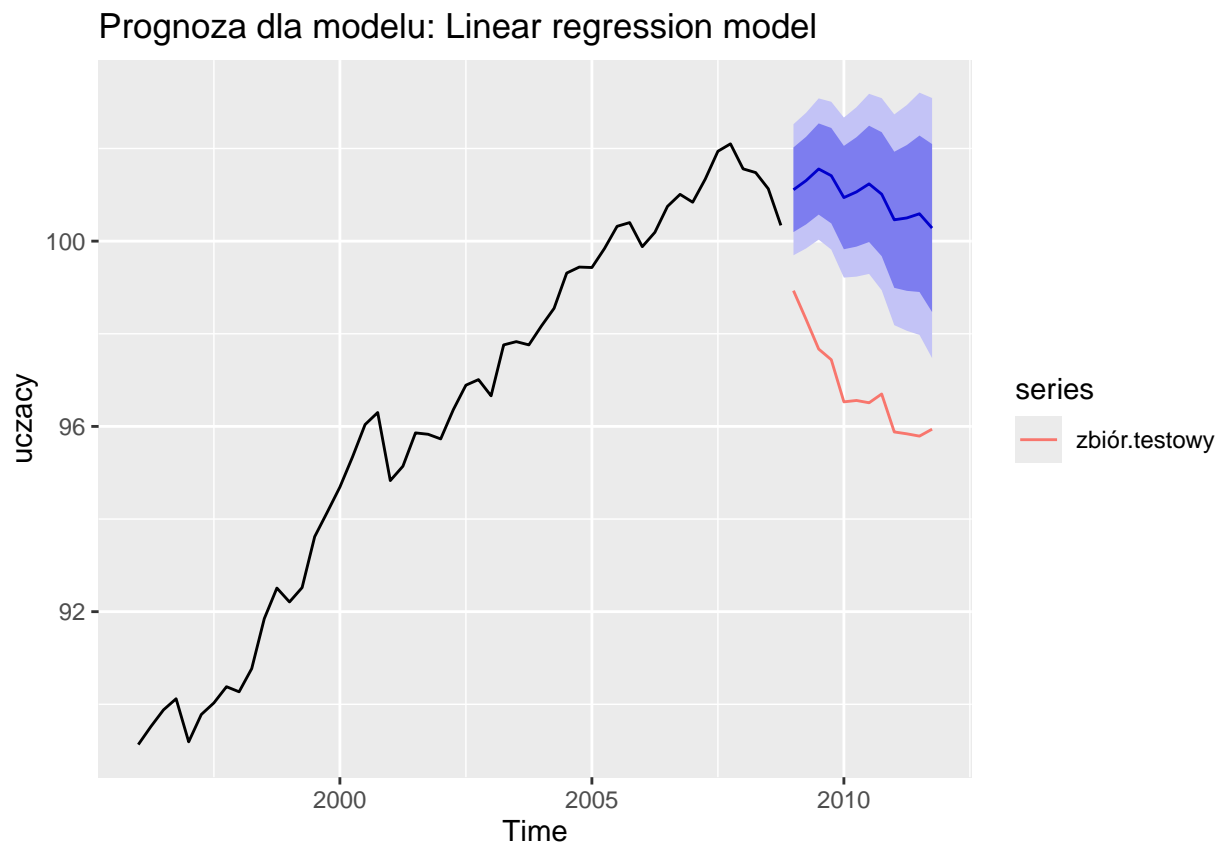
	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	- 0.0233541	0.3625593	0.2633495	- 0.0213779	0.2728761	0.2398629	0.0256259	NA
Test set	- 0.2987598	0.8290063	0.7379423	- 0.3033337	0.7612126	0.6721296	0.7156153	1.822908

Jak widać w tabeli 6, wartości wszystkich miar dokładności są większe dla zbioru testowego niż dla zbioru uczącego. Jest to zdecydowanie dobra wiadomość. Ponadto wartość średniego bezwzględnego błędu procentowego (MAPE) dla zbioru testowego wynosi zaledwie 0.76 %, co oznacza, że prognozowane wartości indeksu są minimalnie inne niż w rzeczywistości. Wartość średniego błędu bezwzględnego (MAE) wskazuje, że model na zbiorze testowym myli się jedynie o 0.74 punkta indeksu. Pierwiastek średniego błędu kwadratowego (RMSE) również przyjmuje niewielką wartość na zbiorze testowym (0.83). Średni błąd bezwzględny skalowany (MASE) przyjmuje wartości mniejsze od 1 dla obu zbiorów, oznacza to, że model jest lepszy od metody naiwnej.

Analiza dokładności naszego modelu wskazuje zatem na bardzo dobrą jakość prognoz, co pokrywa się w wnioskami na podstawie wykresu.

### 1.7.3 Model dekompozycji z trendem sześciennym i sezonowością

Przetestujemy teraz podejście oparte dekompozycji z trendem sześciennym i sezonowością.



Rysunek 21: Wykres przedstawiający prognozy wyznaczone na podstawie modelu dekompozycji z trendem sześciennym i sezonowością oraz przedziały ufności prognoz (80% i 95%). Czerwoną linią zaznaczono rzeczywiste wartości ze zbioru testowego.

Jak widać na wykresie 21 model okazał się nieodpowiedni do prognozowania w momencie, gdy trend w danych gwałtownie się zmienia. Zarówno prognoza punktowa, jak i przedziały predykcji, są całkowicie rozbieżne z wartościami rzeczywistymi. Model nie jest wiarygodny. Model ARIMA sprawdził się tu zdecydowanie lepiej - jest bardziej elastyczny i lepiej dopasowuje się do danych.

Przyjrzyjmy się teraz dokładności prognoz tego modelu na podstawie miar oceny dokładności.



Tabela 7: Tabela accuracy przedstawiająca wartości miar oceny dokładności prognoz dla modelu dekompozycji z trendem sześciennym i sezonowością.

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	0.00000	0.5439109	0.4210677	-	0.4409818	0.3835151	0.7265355	NA
				0.0032701				
Test set	-4.11452	4.1823785	4.1145196	-	4.2564512	3.7475701	0.5566286	9.300553
				4.2564512				

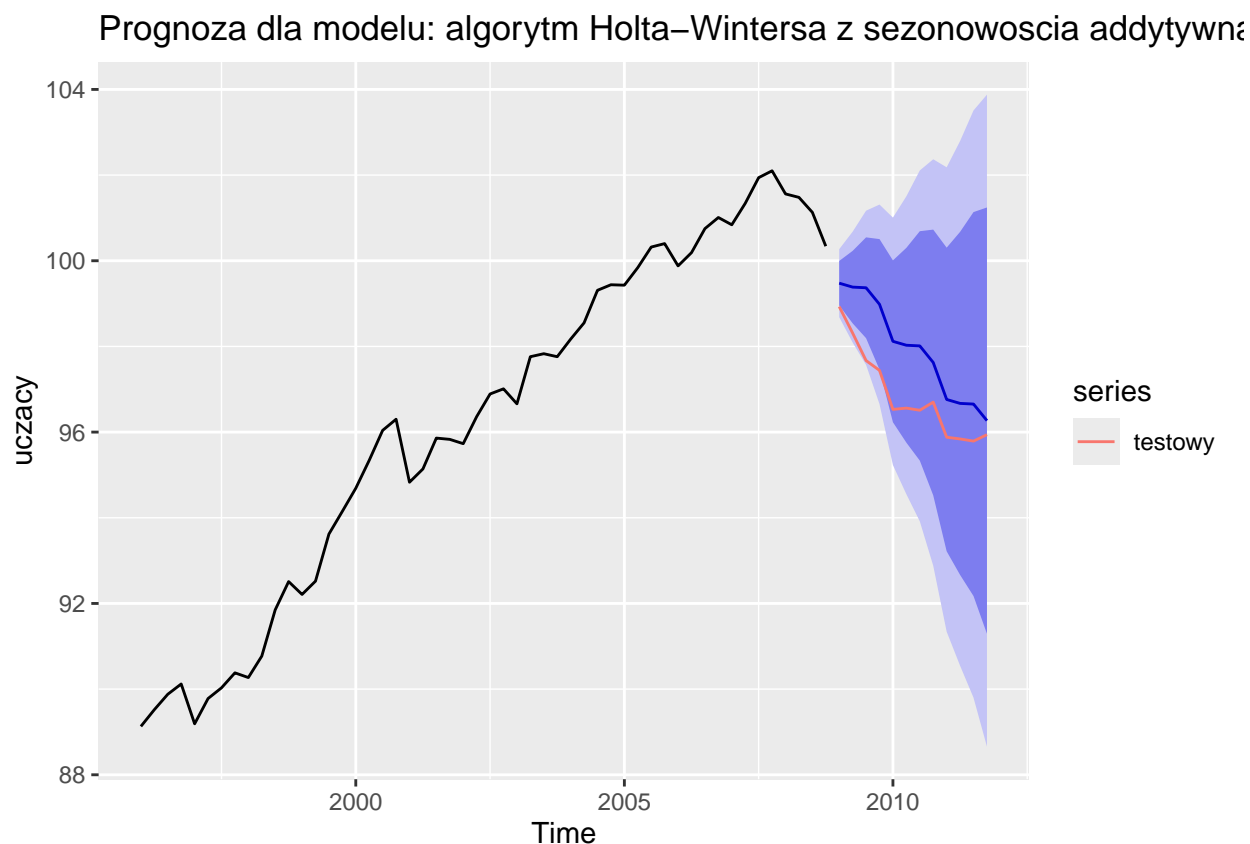
Jak widać w tabeli 7, wartości wszystkich miar dokładności są większe dla zbioru testowego niż dla zbioru uczącego. Jest to zdecydowanie dobra wiadomość, jednak błędy na zbiorze testowym są zdecydowanie większe od tych w poprzednim modelu.

Wartość średniego bezwzględnego błędu procentowego (MAPE) dla zbioru testowego wynosi 4.26 %, co oznacza, że prognozowane wartości indeksu są inne niż w rzeczywistości. Wartość średniego błędu bezwzględnego (MAE) wskazuje, że model na zbiorze testowym myli się o 4.11 punkta indeksu. Pierwiastek średniego błędu kwadratowego (RMSE) znacznie większą wartość na zbiorze testowym (4.18) niż w poprzednim modelu. Średni błąd bezwzględny skalowany (MASE) przyjmuje wartość większą od 1 dla zbioru testowego, oznacza to, że model jest gorszy od metody naiwnej.

Analiza dokładności naszego modelu wskazuje zatem na słabą jakość prognoz, co pokrywa się w wnioskami na podstawie wykresu.

#### 1.7.4 Algorytm Holta-Wintersa z sezonowością addytywną

Teraz do skonstruowania prognoz wykorzystamy algorytm Holta-Wintersa z sezonowością addytywną.



Rysunek 22: Wykres przedstawiający prognozy oraz przedziały ufności prognoz (80% i 95%) wyznaczone na podstawie modelu opartego na algorytmie Holta-Wintersa z sezonowością addytywną. Czerwoną linią zaznaczono rzeczywiste wartości ze zbioru testowego.

Jak widać na wykresie 22 model dobrze przewidział tendencję spadkową zapoczątkowaną na końcu zbioru uczącego (granatowa linia). Rzeczywiste wartości indeksu ze zbioru testowego (czerwona linia) w całości mieszczą się w wyznaczonych przedziałach predykcji, są jednak oddalone od punktowej predykcji. Model można uznać za wiarygodny, chociaż wraz ze wzrostem horyzontu prognozy przedziały predykcji znacząco się rozszerzają, co sugeruje znaczną niepewność modelu w miarę oddalania się od ostatniej znanej obserwacji. Prognoza wygląda na mniej dopasowaną, niż w przypadku modelu (S)ARIMA.

Przyjrzyjmy się teraz dokładności prognoz tego modelu na podstawie miar oceny dokładności.

Tabela 8: Tabela accuracy przedstawiająca wartości miar oceny dokładności prognoz dla modelu opartego na algorytmie Holta-Wintersa z sezonowością addytywną.

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	-	0.3725453	0.2736045	-	0.2852427	0.2492033	0.0300683	NA
	0.030613			0.0291033				
Test set	-	1.1844266	1.1042974	-	1.1398430	1.0058116	0.4881527	2.626489
	1.104297			1.1398430				

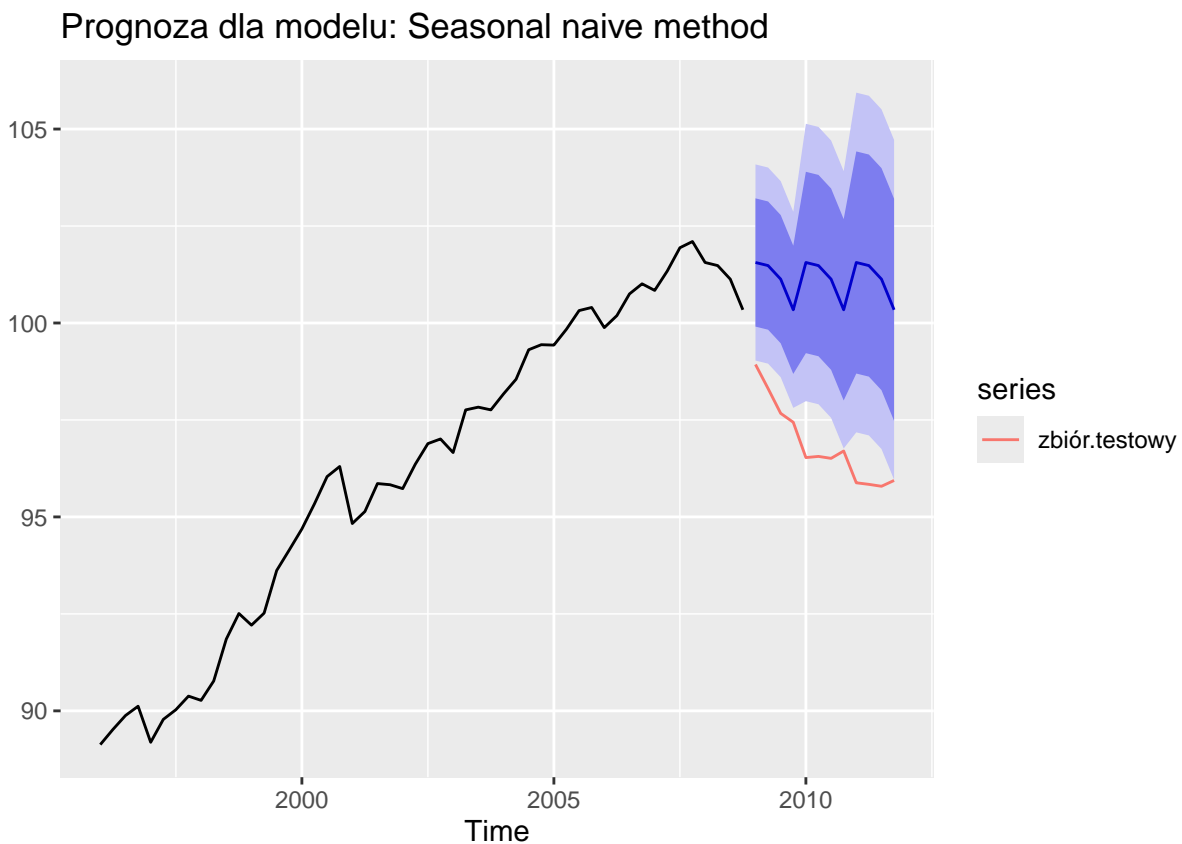
Jak widać w tabeli 8, wartości wszystkich miar dokładności są większe dla zbioru testowego niż dla zbioru uczącego. Jest to zdecydowanie dobra wiadomość, jednak błędy na zbiorze testowym są większe od tych w modelu (S)ARIMA.

Wartość średniego bezwzględnego błędu procentowego (MAPE) dla zbioru testowego wynosi 1.14 %, co oznacza, że prognozowane wartości indeksu są inne niż w rzeczywistości. Wartość średniego błędu bezwzględnego (MAE) wskazuje, że model na zbiorze testowym myli się o 1.1 punkta indeksu. Pierwiastek średniego błędu kwadratowego (RMSE) również przyjmuje dość małą wartość na zbiorze testowym (1.18). Średni błąd bezwzględny skalowany (MASE) przyjmuje wartość większą od 1 dla zbioru testowego, oznacza to, że model jest gorszy od metody naiwnej.

Analiza dokładności naszego modelu wskazuje zatem na średnią jakość prognoz, co pokrywa się w wnioskami na podstawie wykresu.

### 1.7.5 Sezonowa metoda naiwna

Na koniec prognozujemy nasze dane w oparciu o sezonową metodę naiwną.



Rysunek 23: Wykres przedstawiający prognozy oraz przedziały ufności prognoz (80% i 95%) wyznaczone na podstawie modelu opartego na sezonowej metodzie naiwnej. Czerwoną linią zaznaczono rzeczywiste wartości ze zbioru testowego.

Jak widać na wykresie 23 model okazał się nieodpowiedni do prognozowania naszych danych. Zarówno prognoza punktowa, jak i przedziały predykcji, są rozbieżne z wartościami rzeczywistymi. Warto zauważyć jednak, że początek i koniec dolnej granicy 95% przedziału predykcyjnego są zbliżone do pierwszej i ostatniej obserwacji w zbiorze testowym odpowiednio. Model nie jest jednak wiarygodny, czego można było się spodziewać, patrząc na ideę tej metody.

Przyjrzyjmy się teraz dokładności prognoz tego modelu na podstawie miar oceny dokładności.

Tabela 9: Tabela accuracy przedstawiająca wartości miar oceny dokładności prognoz dla modelu opartego na sezonowej metodzie naiwnej.

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	0.9554167	1.289827	1.097917	0.9935928	1.137069	1.000000	0.6908232	NA
Test set	-	4.410181	4.285833	-	4.435735	3.903605	0.4535898	9.780498
	4.2858333			4.4357355				

Jak widać w tabeli 9, wartości wszystkich miar dokładności są większe dla zbioru testowego niż dla zbioru uczącego. Jest to zdecydowanie dobra wiadomość, jednak błędy na obu zbiorach są zdecydowanie większe od tych w poprzednich modelach.

Wartość średniego bezwzględnego błędu procentowego (MAPE) dla zbioru testowego wynosi 4.26 %, co oznacza, że prognozowane wartości indeksu są inne niż w rzeczywistości. Wartość średniego błędu bezwzględnego (MAE) wskazuje, że model na zbiorze testowym myli się o 4.11 punkta indeksu. Pierwiastek średniego błędu kwadratowego (RMSE) ma znacznie większą wartość na zbiorze testowym (4.18) niż w modelu (S)ARIMA czy wykładzenia wykładniczego. Średni błąd bezwzględny skalowany (MASE) przyjmuje wartość większą od 1 dla zbioru testowego, oznacza to, że model jest gorszy od standardowej metody naiwnej.

Analiza dokładności naszego modelu wskazuje zatem na słabą jakość prognoz, co pokrywa się w wnioskami na podstawie wykresu.

## 1.8 Porównanie

Wybierając optymalny model będzie się kierować analizą dokładności prognoz dla zbioru testowego na podstawie wybranych miar oceny dokładności:

- RMSE,
- MAE,
- MAPE,
- MASE.

Tabela 10: Tabela przedstawiająca zestawienie dokładności prognoz dla zbioru testowego na podstawie wybranych miar oceny dokładności: RMSE, MAE, MAPE i MASE.

Model	RMSE	MAE	MAPE	MASE
SARIMA	0.8290063	0.7379423	0.7612126	0.6721296
Dekompozycja	4.1823785	4.1145196	4.2564512	3.7475701
Holt-Winters	1.1844266	1.1042974	1.1398430	1.0058116
Sezonowa naiwna	4.4101805	4.2858333	4.4357355	3.9036053

Jak widać w tabeli 10 model SARIMA okazał się bezkonkurencyjny - osiągnął najniższe wartości dla wszystkich analizowanych tu miar dokładności. Jako jedyny ma średni błąd bezwzględny skalowanego (MASE) mniejszy od 1, co oznacza, że tylko on jest lepszy od standardowej metody naiwej. Niedaleko tej wartości był model oparty na algorytmie Holta-Wintersa, jednak nawet on nie podołał nagłej zmianie trendu w zbiorze testowym. Dwa pozostałe modele przyjmują już wartości kilkukrotnie większe dla wszystkich przedstawionych miar dokładności.

Podsumowując, model SARIMA(0, 1, 4)(0, 1, 0)<sub>4</sub> z parametrem  $\lambda = 2$  okazał się najlepszy do prognozowania danych `euretail`.

## 2 Źródła

- [https://pl.wikipedia.org/wiki/Globalny\\_kryzys\\_finansowy](https://pl.wikipedia.org/wiki/Globalny_kryzys_finansowy)