

# Sprawozdanie

Jakub Kaźmierczyk

2025-06-01

## Spis treści

<b>1</b>	<b>Wprowadzenie</b>	<b>2</b>
1.1	Opis projektu . . . . .	2
1.2	Zmienne . . . . .	2
1.2.1	Zmienna objaśniana . . . . .	2
1.2.2	Zmienne objaśniające . . . . .	2
1.3	Źródła . . . . .	3
<b>2</b>	<b>Wczytywanie danych</b>	<b>4</b>
<b>3</b>	<b>Podstawowe statystyki</b>	<b>5</b>
3.1	Zmienna objaśniana . . . . .	5
3.2	Zmienne objaśniające . . . . .	5
3.3	Macierze korelacji . . . . .	7
3.3.1	Macierz korelacji przed usunięciem zmiennych . . . . .	7
3.3.2	Macierz korelacji po usunięciu zmiennych . . . . .	8
<b>4</b>	<b>Identyfikacja niestacjonarnych zmiennych objaśniających</b>	<b>9</b>
4.1	Sprawdzenie niestacjonarności zmiennych . . . . .	9
4.2	Usunięcie niestacjonarności . . . . .	9
4.3	Ponowne sprawdzenie niestacjonarności zmiennych . . . . .	13
4.4	Sprawdzenie korelacji po usunięciu niestacjonarności . . . . .	14
4.5	Usunięcie zmiennych o zerowej wariancji . . . . .	14
4.5.1	Przed usunięciem . . . . .	14
4.5.2	Po usunięciu . . . . .	15

---

<b>5</b>	<b>Metoda doboru zmiennych</b>	<b>16</b>
5.1	Metoda Hellwiga . . . . .	16
<b>6</b>	<b>Tworzenie modelu ekonometrycznego</b>	<b>17</b>
<b>7</b>	<b>Niby TEST</b>	<b>18</b>
7.1	Testowanie autokorelacji . . . . .	19
7.2	Badanie heteroskedastyczności . . . . .	20
7.3	Testowanie współliniowości (VIF) . . . . .	22
7.4	Testowanie stabilności parametrów (TEST CHOWA) . . . . .	22
7.5	Testowanie stabilności postaci analitycznej (TEST RESET) . . . . .	23
7.6	Badanie efektu katalizy . . . . .	24
7.7	Badanie koincydencji . . . . .	24
<b>8</b>	<b>Podsumowanie wyników</b>	<b>26</b>

# 1 Wprowadzenie

## 1.1 Opis projektu

Projekt ma na celu budowę kompleksowego modelu ekonometrycznego służącego do analizy i prognozowania rentowności 10-letnich polskich obligacji skarbowych. Model zostanie zbudowany na podstawie szeregów czasowych, co umożliwi głębszą analizę dynamicznych zależności ekonomicznych.

## 1.2 Zmienne

### 1.2.1 Zmienna objaśniana

**CLOSE** - rentowność 10-letnich polskich obligacji skarbowych

### 1.2.2 Zmienne objaśniające

**XAUUSD** - cena złota w dolarze amerykańskim

**S&P500** - ETF 500 największych notowanych na giełdzie amerykańskich spółek

**PMI** - wskaźnik aktywności przemysłowej

---

**WIG20** - 20 największych notowanych na giełdzie polskich spółek

**OIL** - cena ropy naftowej za baryłkę

**UNEMPLOYMENT** - stopa bezrobocia w Polsce

**USDPLN** - kurs dolara amerykańskiego wyrażony w złotych

**INFLATION** - inflacja rok do roku

### **1.3 Źródła**

[www.stooq.com](http://www.stooq.com)

---

## 2 Wczytywanie danych

```
data_all <- read_excel("data.xlsx")
data_all <- data_all[, -c(1, 3, 4)]
data_all[] <- lapply(data_all, function(col) {
  na.approx(col, na.rm = FALSE)
})

n <- nrow(data_all)
train_size <- floor(0.8 * n)

data <- data_all[1:train_size, , drop = FALSE]

Y <- data["CLOSE"]
X <- data[, !names(data) %in% "CLOSE", drop = FALSE]
```

---

## 3 Podstawowe statystyki

### 3.1 Zmienna objaśniana

```
##      CLOSE
##  Min.    : 1.843
## 1st Qu.: 3.457
##  Median : 5.495
##   Mean   : 5.610
## 3rd Qu.: 6.269
##   Max.   :13.288
```

Wartości zmiennej objaśnianej wachają się pomiędzy 13,288 a 1,149. Mediana wynosi 5,461 a średnia 5,347.

### 3.2 Zmienne objaśniające

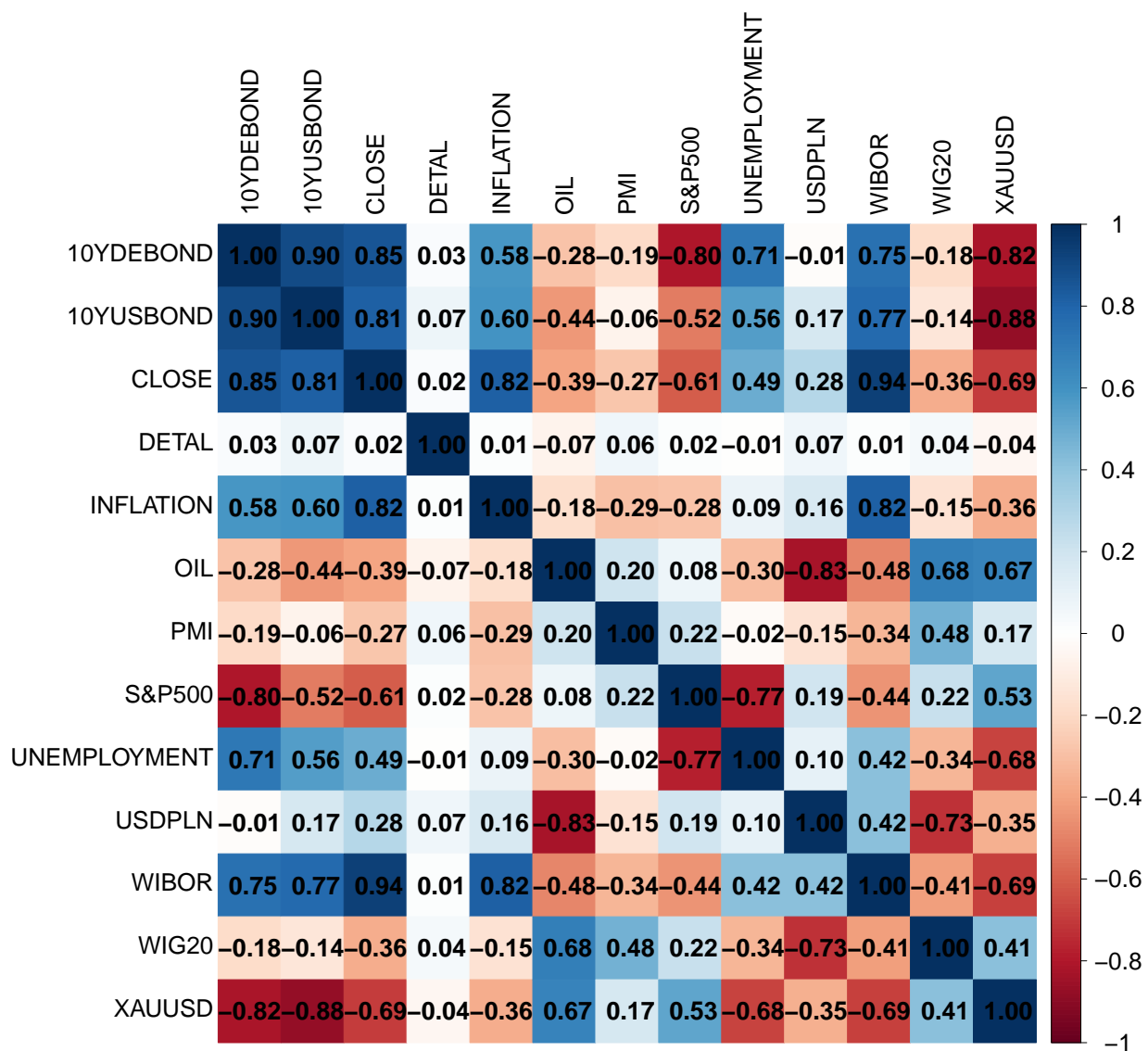
```
##      INFLATION      10YUSBOND      XAUUSD      USDPLN
##  Min.    :-0.01600   Min.    :1.455   Min.    : 255.8   Min.    :2.060
## 1st Qu.: 0.01000   1st Qu.:2.337   1st Qu.: 416.2   1st Qu.:3.084
##  Median : 0.02250   Median :3.385   Median :1024.5   Median :3.509
##   Mean   : 0.02711   Mean    :3.461   Mean    : 921.2   Mean    :3.476
## 3rd Qu.: 0.04000   3rd Qu.:4.480   3rd Qu.:1292.5   3rd Qu.:3.910
##   Max.   : 0.11600   Max.    :6.667   Max.    :1825.3   Max.    :4.644
##
##      WIBOR      10YDEBOND      WIG20      S&P500
##  Min.    : 1.560   Min.    :-0.7010   Min.    :1023   Min.    : 735.1
## 1st Qu.: 2.062   1st Qu.: 0.7907   1st Qu.:1789   1st Qu.:1154.7
##  Median : 4.175   Median : 3.1740   Median :2268   Median :1366.2
##   Mean   : 5.572   Mean    : 2.6779   Mean    :2182   Mean    :1578.4
## 3rd Qu.: 6.143   3rd Qu.: 4.1895   3rd Qu.:2462   3rd Qu.:1972.2
##   Max.   :20.520   Max.    : 5.5390   Max.    :3878   Max.    :3230.8
##
##      UNEMPLOYMENT      PMI      DETAL      OIL
```

---

```
## Min.      :0.0500   Min.      :38.30   Min.      : -10.7000   Min.      : 18.57
## 1st Qu.:0.1030   1st Qu.:48.38   1st Qu.: -0.5000   1st Qu.: 37.32
## Median :0.1245   Median :51.15   Median :  0.5000   Median : 58.28
## Mean    :0.1288   Mean    :50.58   Mean    :  0.5221   Mean    : 60.99
## 3rd Qu.:0.1590   3rd Qu.:53.20   3rd Qu.:  1.5000   3rd Qu.: 80.75
## Max.    :0.2070   Max.    :56.90   Max.    : 10.8000   Max.    :140.00
##                                     NA's      :8
```

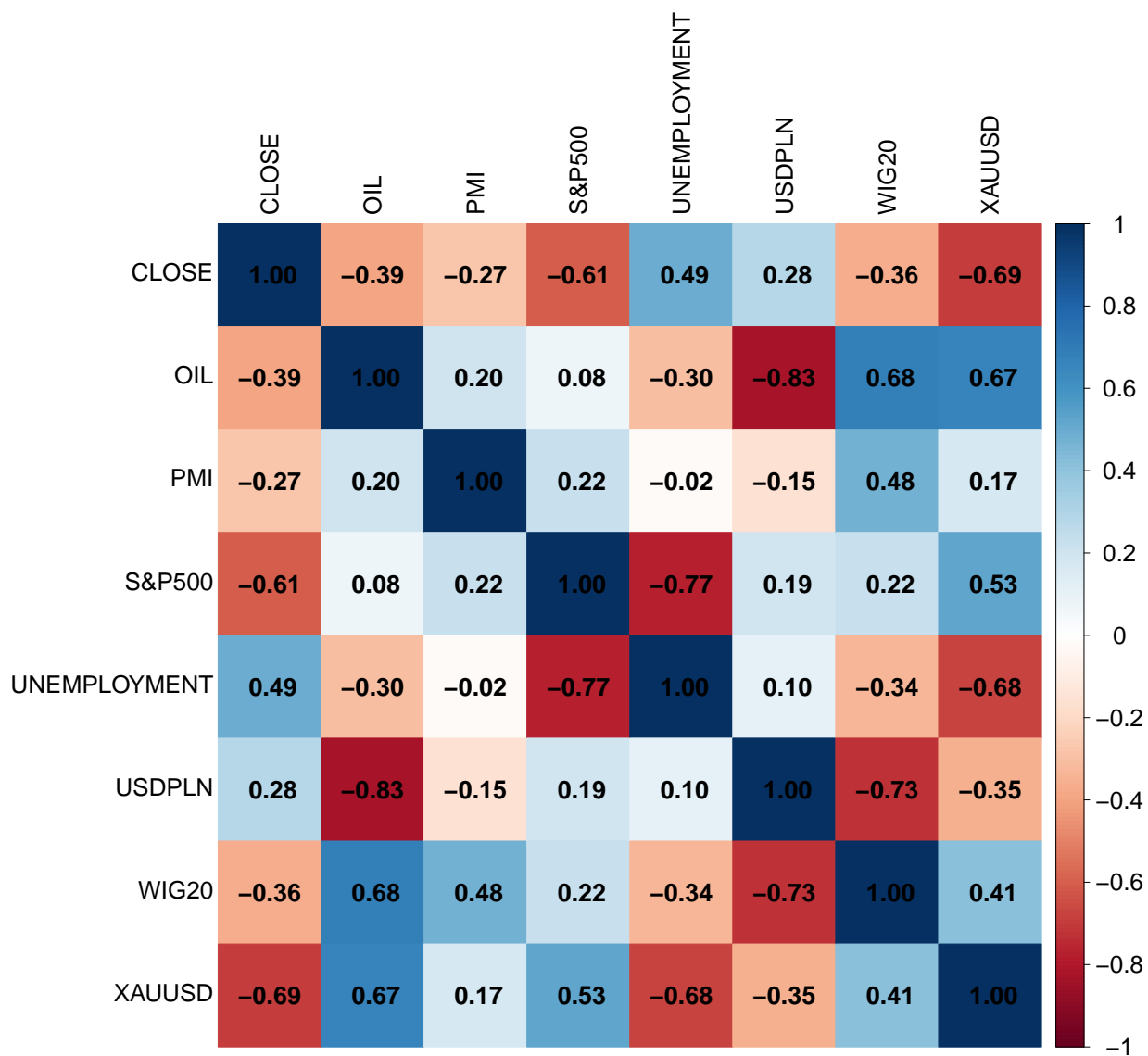
### 3.3 Macierze korelacji

#### 3.3.1 Macierz korelacji przed usunięciem zmiennych



Z 11 zmiennych objaśniających wybrałem 7, których wartość bezwzględna korelacji nie przekracza 0.7.

### 3.3.2 Macierz korelacji po usunięciu zmiennych





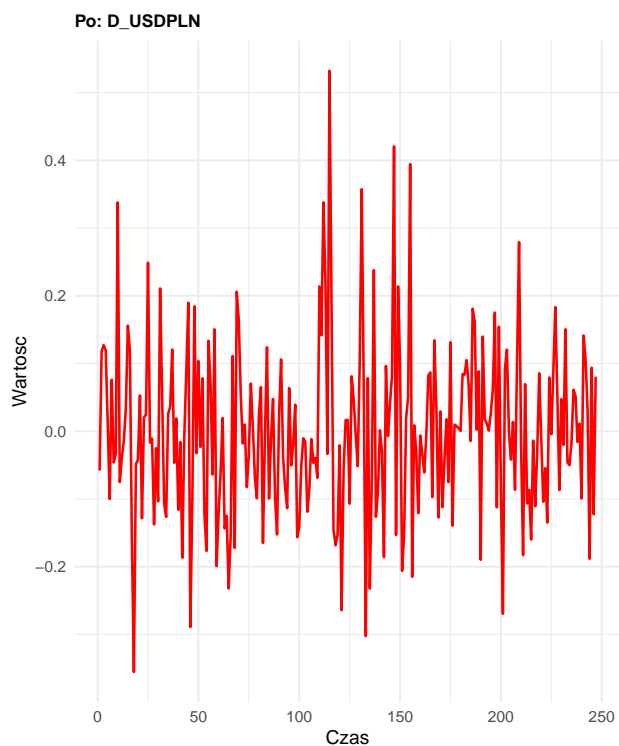
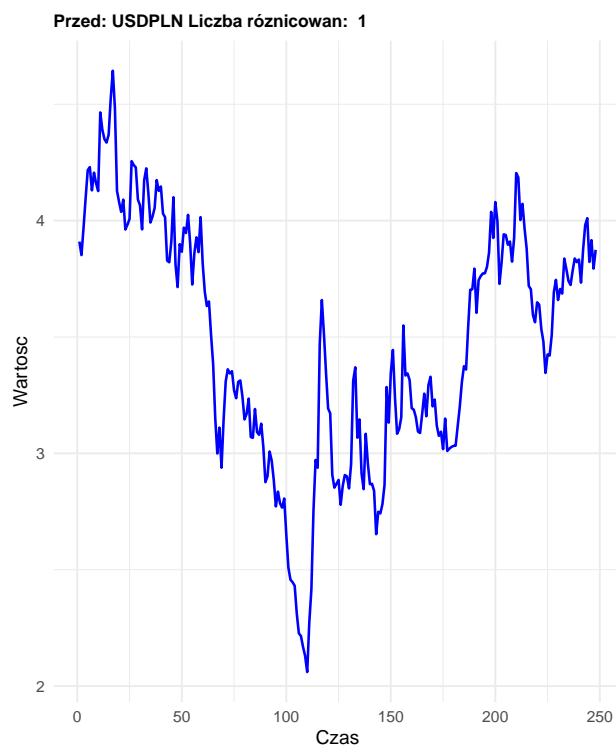
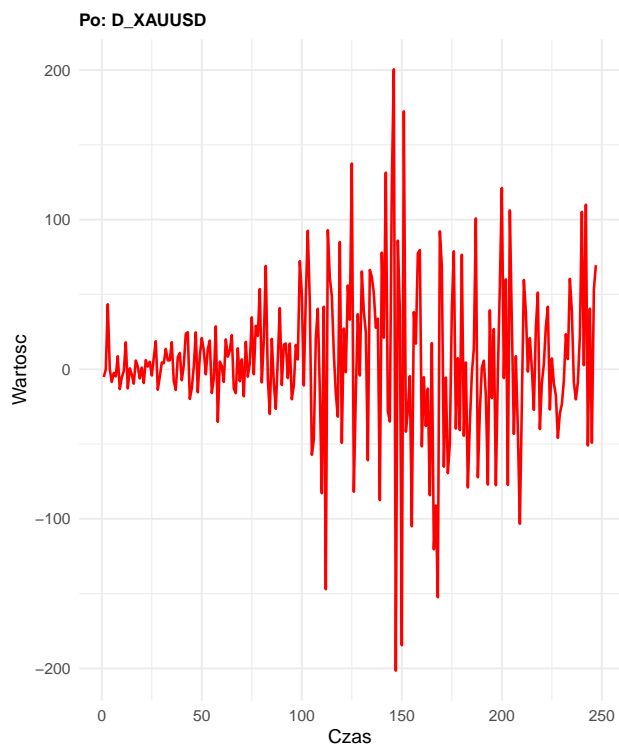
## 4 Identyfikacja niestacjonarnych zmiennych objaśniających

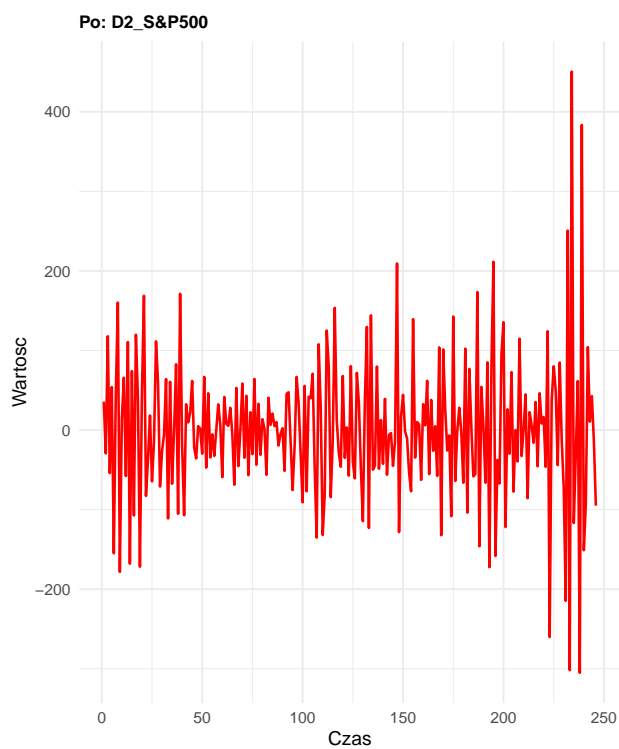
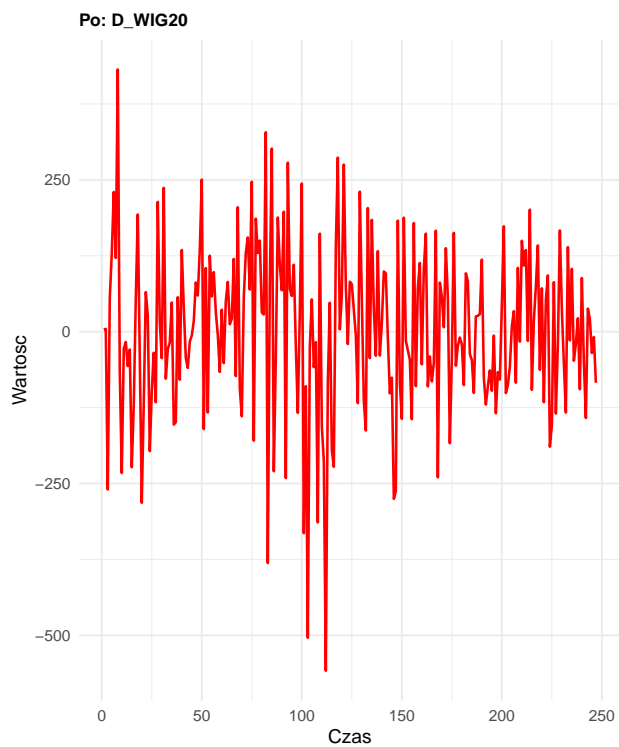
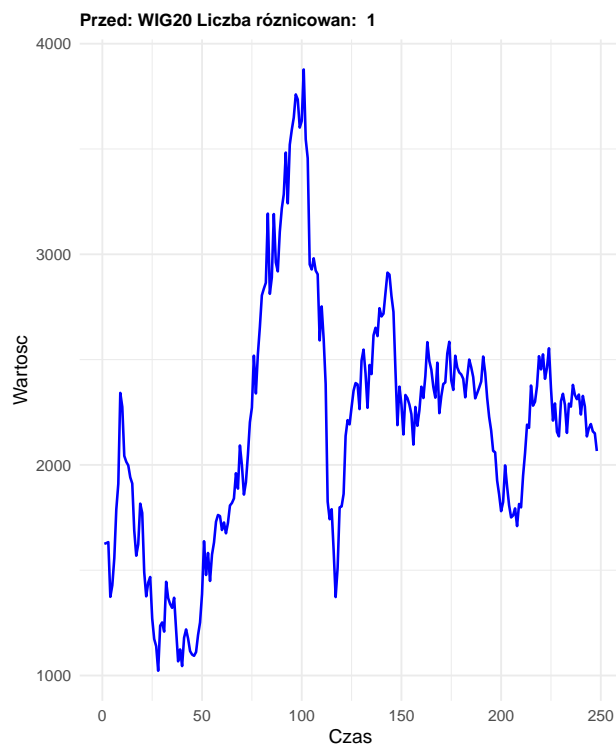
### 4.1 Sprawdzenie niestacjonarności zmiennych

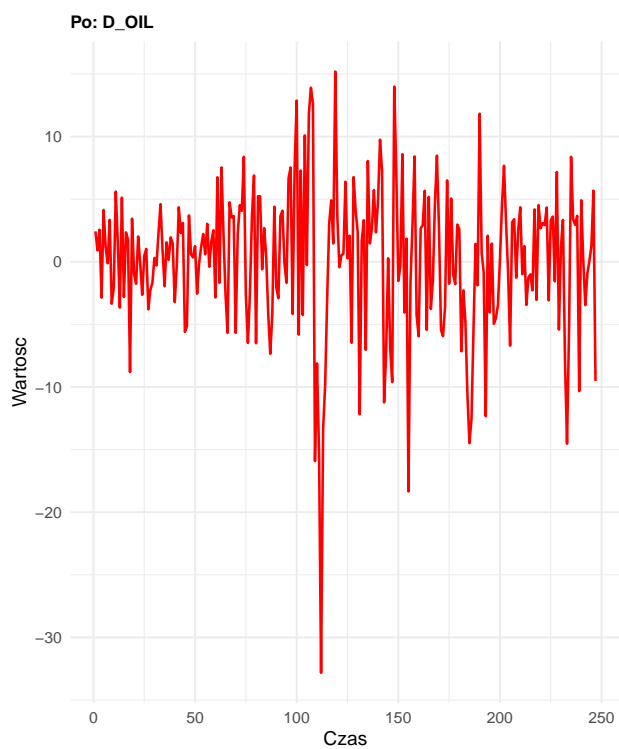
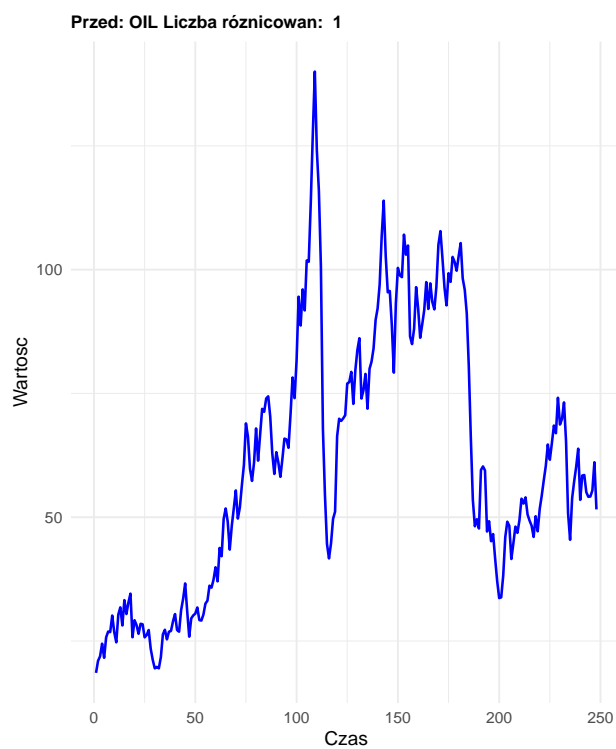
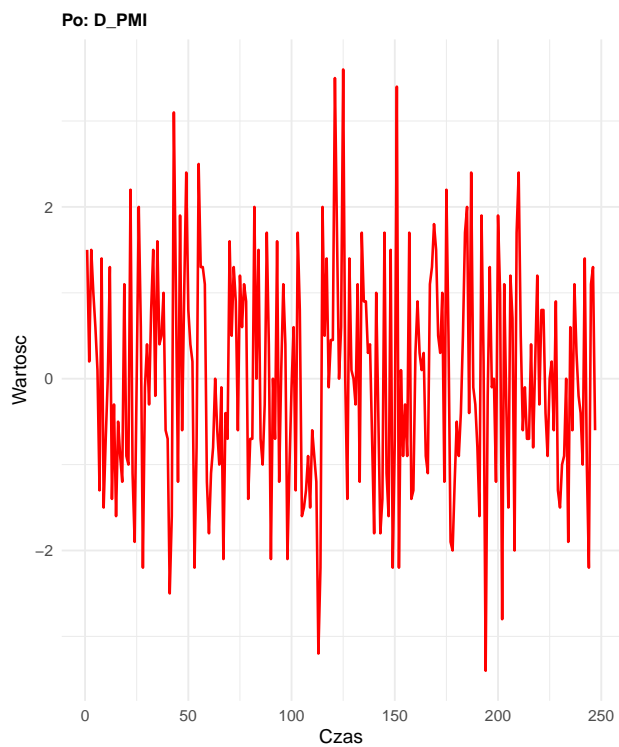
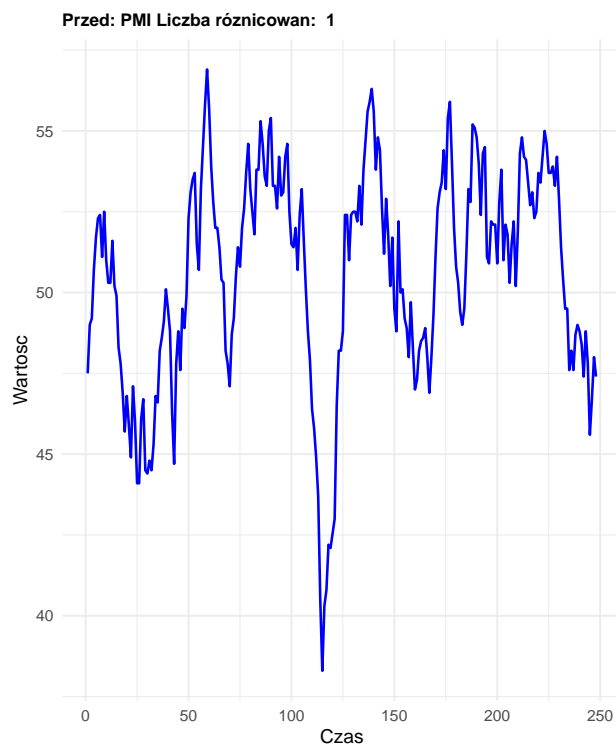
Zmienna	Stacjonarnosc
CLOSE	Niestacjonarna
XAUUSD	Niestacjonarna
USDPLN	Niestacjonarna
WIG20	Niestacjonarna
S&P500	Niestacjonarna
PMI	Niestacjonarna
OIL	Niestacjonarna
UNEMPLOYMENT	Niestacjonarna

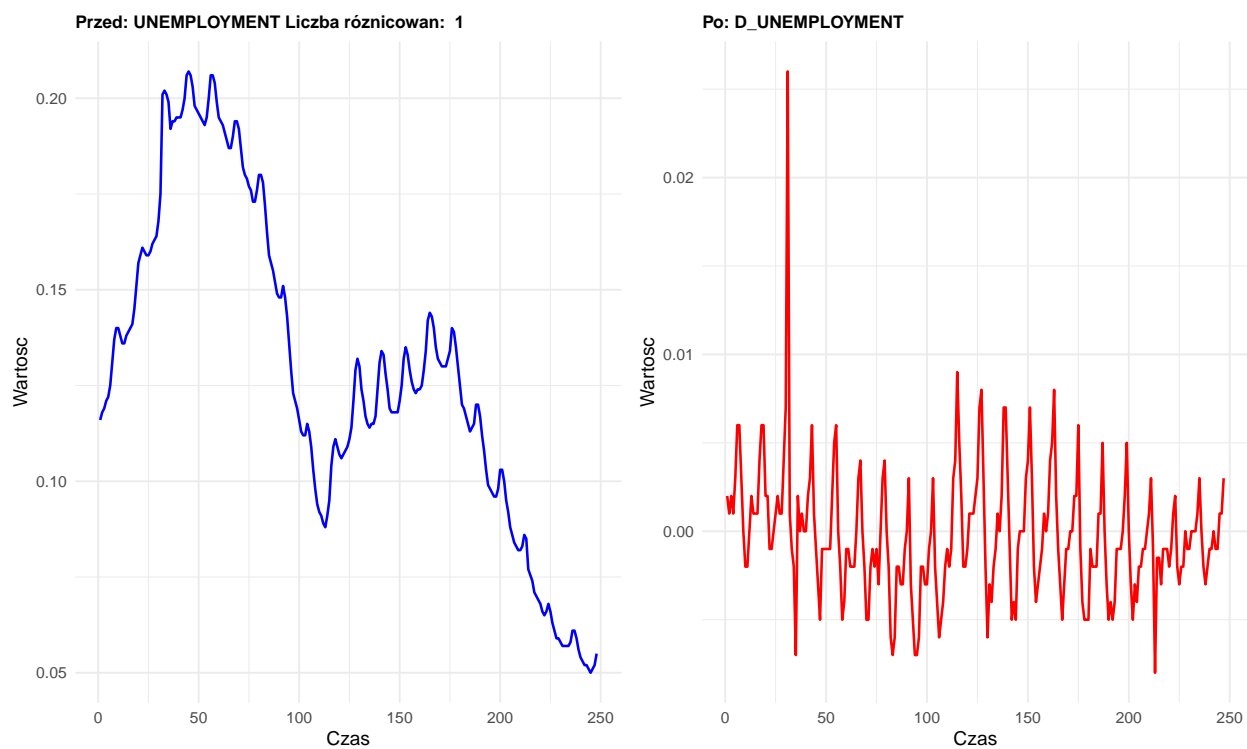
### 4.2 Usunięcie niestacjonarności







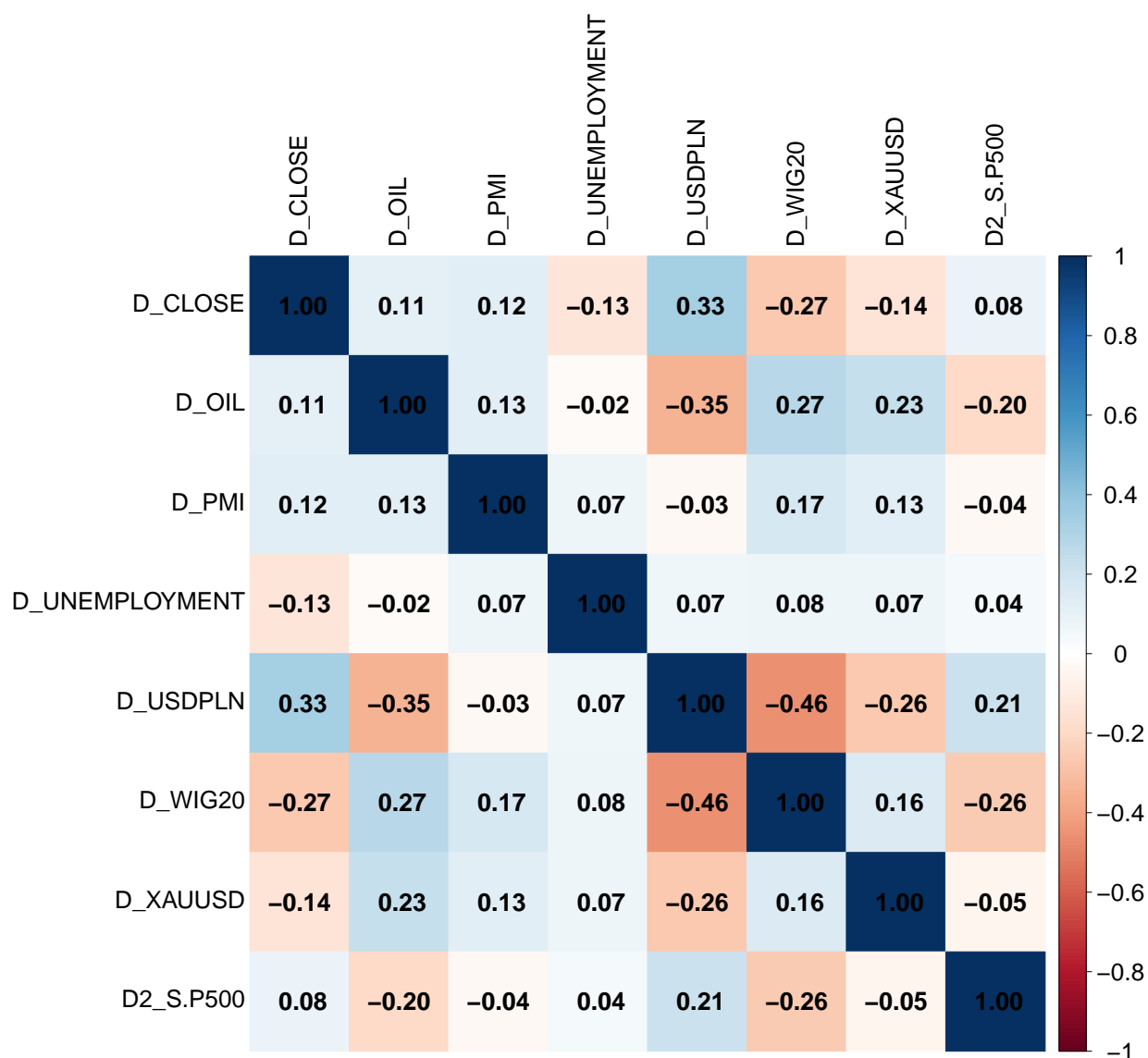




#### 4.3 Ponowne sprawdzenie niestacjonarności zmiennych

Zmienna	Stacjonarnosc
D_CLOSE	Stacjonarna
D_XAUUSD	Stacjonarna
D_USDPLN	Stacjonarna
D_WIG20	Stacjonarna
D2_S.P500	Stacjonarna
D_PMI	Stacjonarna
D_OIL	Stacjonarna
D_UNEMPLOYMENT	Stacjonarna

## 4.4 Sprawdzenie korelacji po usunięciu niestacjonarności



## 4.5 Usunięcie zmiennych o zerowej wariancji

### 4.5.1 Przed usunięciem

D\_CLOSE - Współczynnik zmienności: -1479.303 %, Wariancja: 0.1246856

D\_XAUUSD - Współczynnik zmienności: 1006.021 %, Wariancja: 2638.734

D\_USDPLN - Współczynnik zmienności: -27121.85 %, Wariancja: 0.01641273

---

D\_WIG20 - Współczynnik zmienności: 6463.945 %, Wariancja: 19022.24  
D2\_S.P500 - Współczynnik zmienności: 57653.87 %, Wariancja: 8239.136  
D\_PMI - Współczynnik zmienności: 63313.66 %, Wariancja: 1.656016  
D\_OIL - Współczynnik zmienności: 3422.796 %, Wariancja: 35.01729  
D\_UNEMPLOYMENT - Współczynnik zmienności: -1377.121 %, Wariancja: 1.283612e-05

#### **4.5.2 Po usunięciu**

D\_CLOSE - Współczynnik zmienności: -1479.303 %, Wariancja: 0.1246856  
D\_XAUUSD - Współczynnik zmienności: 1006.021 %, Wariancja: 2638.734  
D\_USDPLN - Współczynnik zmienności: -27121.85 %, Wariancja: 0.01641273  
D\_WIG20 - Współczynnik zmienności: 6463.945 %, Wariancja: 19022.24  
D2\_S.P500 - Współczynnik zmienności: 57653.87 %, Wariancja: 8239.136  
D\_PMI - Współczynnik zmienności: 63313.66 %, Wariancja: 1.656016  
D\_OIL - Współczynnik zmienności: 3422.796 %, Wariancja: 35.01729

---

## 5 Metoda doboru zmiennych

### 5.1 Metoda Hellwiga

Zmienne składowe w najlepszej kombinacji:

D\_USDPLN

D\_WIG20

D\_PMI

Pojemność Hellwiga dla tej kombinacji: 0.1297



---

## 6 Tworzenie modelu ekonometrycznego

```
formula_modelu <- reformulate(best_hellwig_vars, response = "D_CLOSE")

model <- lm(formula_modelu, data = data_stationary)

print(summary(model))
```

Call:

```
lm(formula = formula_modelu, data = data_stationary)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-1.09467	-0.16183	-0.00217	0.18815	1.28861

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.0226434	0.0208685	-1.085	0.278978
D_USDPLN	0.6970193	0.1837255	3.794	0.000187 ***
D_WIG20	-0.0004625	0.0001732	-2.671	0.008076 **
D_PMI	0.0440821	0.0165181	2.669	0.008129 **

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3273 on 242 degrees of freedom

Multiple R-squared: 0.1515, Adjusted R-squared: 0.141

F-statistic: 14.41 on 3 and 242 DF, p-value: 1.149e-08

H0 takie ze ... p wynoszace xyz oznacza...

---

## 7 Niby TEST

## TEORIA: Testy normalności sprawdzają czy reszty mają rozkład normalny.

## H0: Reszty mają rozkład normalny

## H1: Reszty nie mają rozkładu normalnego

## Poziom istotności:  $\alpha = 0.05$

## 1. TEST SHAPIRO-WILKA:

## Statystyka W = 0.9719

## p-value = 1e-04

## Wniosek: Odrzucamy H0 - reszty nie są normalne

## 2. TEST JARQUE-BERA:

## Statystyka JB = 34.9526

## p-value = 0

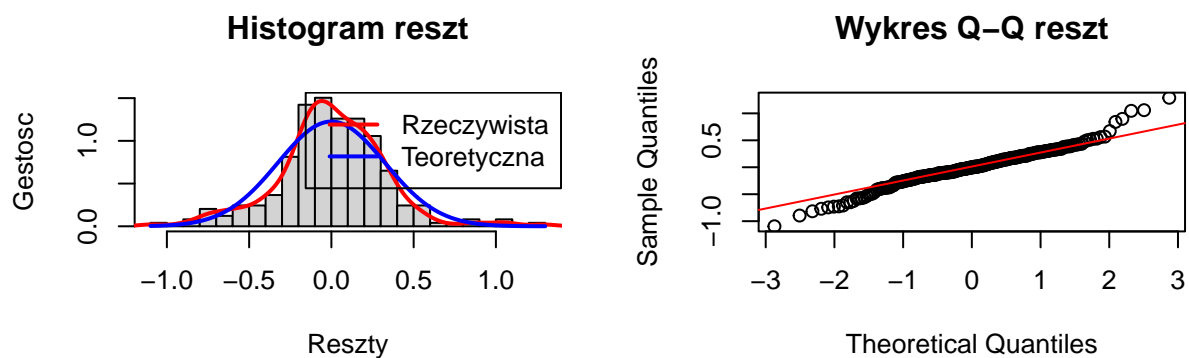
## Wniosek: Odrzucamy H0 - reszty nie są normalne

## 3. TEST ANDERSON-DARLING:

## Statystyka A = 1.6762

## p-value = 3e-04

## Wniosek: Odrzucamy H0 - reszty nie są normalne



## 7.1 Testowanie autokorelacji

## TEORIA: Autokorelacja oznacza korelację między resztami w różnych okresach.

## H0: Brak autokorelacji reszt

## H1: Występuje autokorelacja reszt

## 1. TEST DURBINA-WATSONA:

## Statystyka DW = 1.5324

## p-value = 0

## Wniosek: Odrzucamy H0 - występuje autokorelacja

## 2. TEST LJUNGA-BOXA:

---

```
## Statystyka LB = 31.6143

## p-value = 5e-04

## Wniosek: Odrzucamy H0 - występuje autokorelacja

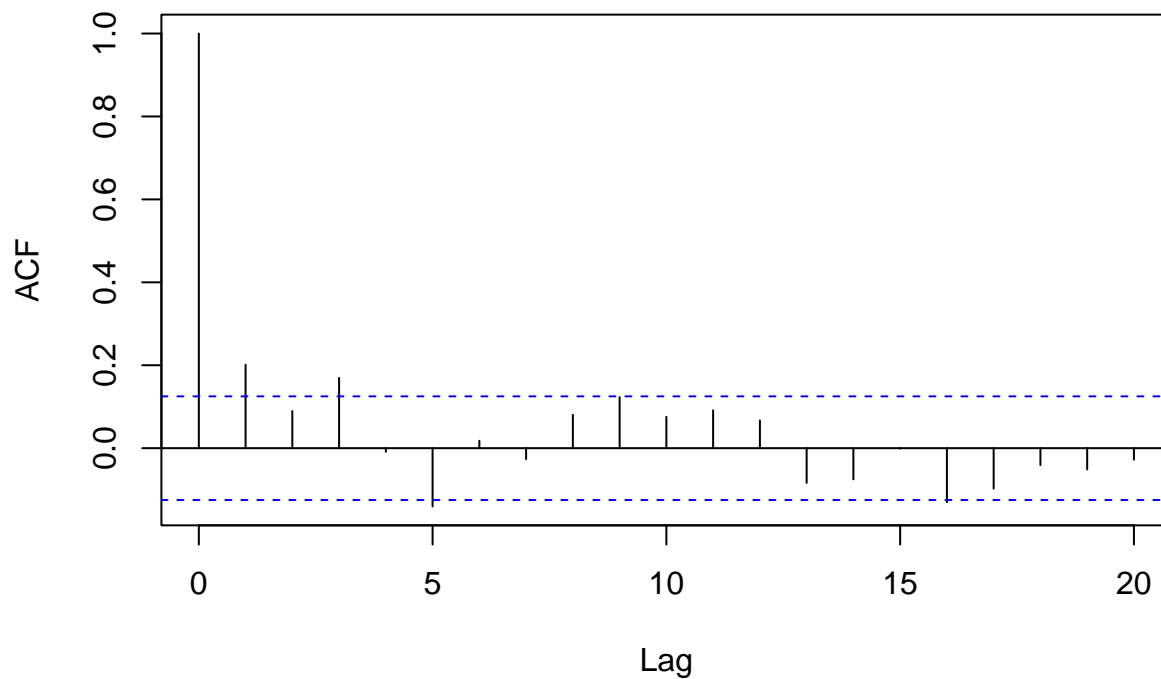
## 3. TEST BREUSCHA-GODFREYA:

## Statystyka LM = 10.6555

## p-value = 0.0049

## Wniosek: Odrzucamy H0 - występuje autokorelacja
```

### Funkcja autokorelacji reszt



## 7.2 Badanie heteroskedastyczności

```
## TEORIA: Heteroskedastyczność oznacza niestabilną wariancję składnika losowego.
```

---

## H0: Homoskedastyczność (stała wariancja)

## H1: Heteroskedastyczność (niestała wariancja)

## 1. TEST BREUSCHA-PAGANA:

## Statystyka BP = 0.1821

## p-value = 0.9804

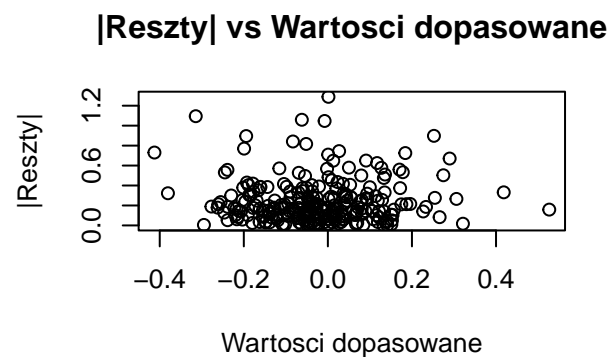
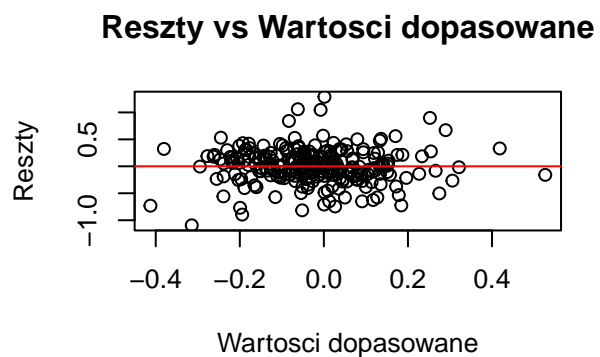
## Wniosek: Nie ma podstaw do odrzucenia H0 - homoskedastyczność

## 3. TEST GOLDFELDA-QUANDTA:

## Statystyka GQ = 1.1587

## p-value = 0.2115

## Wniosek: Nie ma podstaw do odrzucenia H0 - homoskedastyczność



---

### 7.3 Testowanie współliniowości (VIF)

## TEORIA: Współliniowość oznacza wysoką korelację między zmiennymi objaśniającymi.

## VIF > 10: poważna współliniowość

## VIF > 5: umiarkowana współliniowość

## VIF < 5: brak problemów ze współliniowością

## WSPÓŁCZYNNIKI VIF:

## D\_USDPLN : 1.267 - OK

## D\_WIG20 : 1.305 - OK

## D\_PMI : 1.034 - OK

##

## WNIOSEK: Brak problemów ze współliniowością

### 7.4 Testowanie stabilności parametrów (TEST CHOWA)

## TEORIA: Test Chowa sprawdza czy parametry modelu są stabilne w czasie.

## H0: Parametry są stabilne (brak przełomu strukturalnego)

## H1: Parametry nie są stabilne (występuje przełom strukturalny)

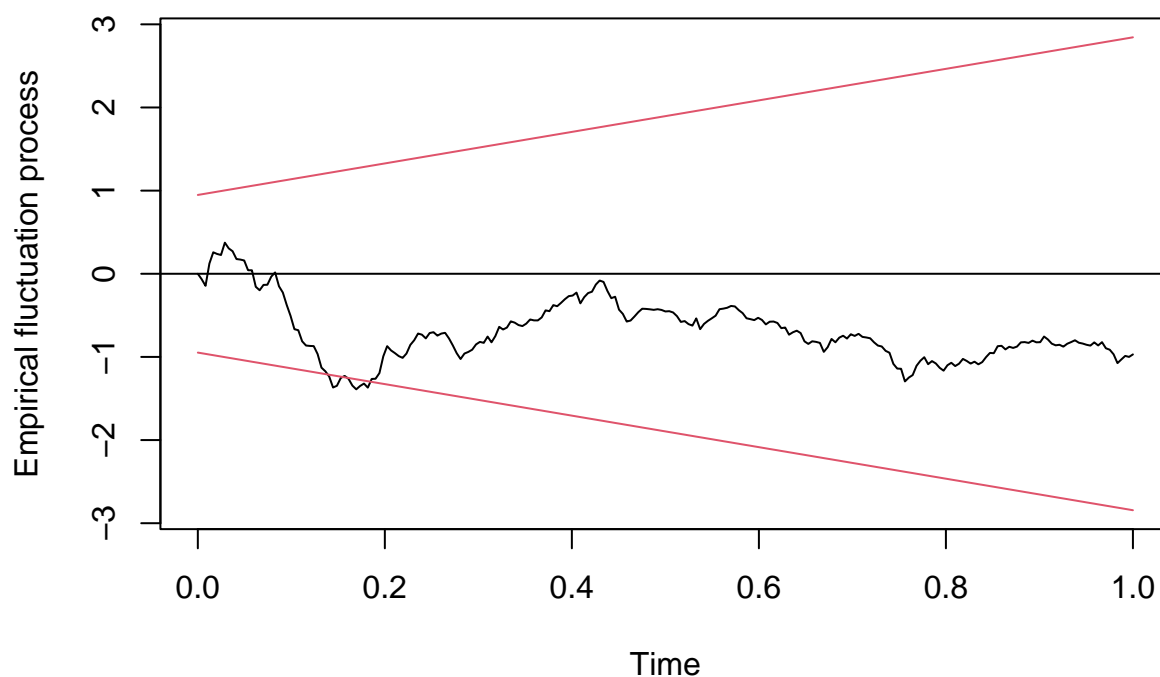
## TEST CHOWA (punkt przełomu w obserwacji 123 ):

## Statystyka F = 2.9766

## p-value = 0.02

## Wniosek: Odrzucamy H0 - brak stabilności parametrów

### Test CUSUM stabilnosci parametrów



### 7.5 Testowanie stabilności postaci analitycznej (TEST RESET)

## TEORIA: Test RESET sprawdza czy postać funkcyjna modelu jest poprawna.

## H0: Model ma poprawną postać funkcyjną

## H1: Model ma niepoprawną postać funkcyjną

## 1. TEST RESET RAMSEYA:

## Statystyka F = 1.5663

## p-value = 0.2109

## Wniosek: Nie ma podstaw do odrzucenia H0 - poprawna postać modelu

## 2. TEST LICZBY SERII:

---

## TEORIA: Test sprawdza czy reszty są losowo rozłożone.

## H0: Reszty są losowo rozłożone

## H1: Reszty wykazują systematyczne wzorce

## Statystyka = -1.6602

## p-value = 0.0969

## Wniosek: Nie ma podstaw do odrzucenia H0 - reszty są losowe

## 7.6 Badanie efektu katalizy

## TEORIA: Efekt katalizy - jedna zmienna wpływa na siłę oddziaływania innej.

## Sprawdzamy czy interakcje między zmiennymi są istotne.

## TEST F DLA INTERAKCJI:

## Statystyka F = 0.9552

## p-value = 0.4328

## Wniosek: Brak istotnego efektu katalizy

##

## WSPÓŁCZYNNIKI INTERAKCJI:

## D\_USDPLN:D\_WIG20 : p-value = 0.9471

## D\_USDPLN:D\_PMI : p-value = 0.8729

## D\_WIG20:D\_PMI : p-value = 0.121

## D\_USDPLN:D\_WIG20:D\_PMI : p-value = 0.5999

## 7.7 Badanie koincydencji

## TEORIA: Koincydencja - zmienna objaśniająca ma wpływ jedynie w określonych okresach.

## Sprawdzamy stabilność parametrów w różnych podokresach.



---

## ANALIZA STABILNOŚCI PARAMETRÓW W PODOKRESACH:

## Współczynniki determinacji:

## Okres 1 (obs. 1- 82 ):  $R^2 = 0.2761$

## Okres 2 (obs. 83 - 164 ):  $R^2 = 0.1815$

## Okres 3 (obs. 165 - 246 ):  $R^2 = 0.0457$

## PORÓWNIANIE PARAMETRÓW W PODOKRESACH:

## Parametr (Intercept) :

## Okres 1: -0.0094

## Okres 2: -0.0175

## Okres 3: -0.0265

## \*\*\* MOŻLIWA KOINCYDENCJA - duże różnice między okresami \*\*\*

##

## Parametr D\_USDPLN :

## Okres 1: 1.3833

## Okres 2: 0.244

## Okres 3: 0.4964

## \*\*\* MOŻLIWA KOINCYDENCJA - duże różnice między okresami \*\*\*

##

## Parametr D\_WIG20 :

## Okres 1: -9e-04

## Okres 2: -5e-04

## Okres 3: 0

## \*\*\* MOŻLIWA KOINCYDENCJA - duże różnice między okresami \*\*\*

##

## Parametr D\_PMI :

## Okres 1: 0.0912

## Okres 2: 0.025

## Okres 3: 0.0063

## \*\*\* MOŻLIWA KOINCYDENCJA - duże różnice między okresami \*\*\*

---

## 8 Podsumowanie wyników

## WYNIKI TESTÓW DIAGNOSTYCZNYCH:

##		Test	Statystyka	p_value	Wynik
## 1	Normalność (Jarque-Bera)	34.953	0	NIESPEŁNIONE	
## 2	Autokorelacja (Ljung-Box)	31.614	0	NIESPEŁNIONE	
## 3	Heteroskedastyczność (Breusch-Pagan)	0.182	0.98	SPEŁNIONE	
## 4	Współliniowość (max VIF)	1.305	N/A	SPEŁNIONE	
## 5	Stabilność (Chow)	2.977	0.02	NIESPEŁNIONE	
## 6	Postać modelu (RESET)	1.566	0.211	SPEŁNIONE	

##

## === OGÓLNA OCENA MODELU ===

## Spełnione założenia: 3 / 6

## Niespełnione założenia: 3 / 6

## MODEL WYMAGA ISTOTNYCH POPRAWEK - niespełnia kluczowych założeń

##

## === REKOMENDACJE ===

## • Rozważ transformację zmiennych (logarytmowanie) ze względu na brak normalności reszt

## • Dodaj zmienne opóźnione lub rozważ model ARIMA ze względu na autokorelację

## • Rozważ model ze zmiennymi strukturalnymi ze względu na niestabilność parametrów

##

## =====

## KONIEC WERYFIKACJI MODELU

## =====