

Sprawozdanie

Jakub Kaźmierczyk

2025-05-31

Spis treści

1	Wprowadzenie	2
1.1	Opis projektu	2
1.2	Zmienna objaśniana	2
1.3	Zmienne objaśniające	2
1.4	Źródła	2
2	Wczytywanie danych	3
3	Podstawowe statystyki	4
3.1	Zmienna objaśniana	4
3.2	Zmienne objaśniające	4
3.3	Macierz korelacji	5
4	Identyfikacja niestacjonarnych zmiennych objaśniających	6
4.1	Sprawdzenie niestacjonarności zmiennych	6
4.2	Usunięcie niestacjonarności	6
4.3	Ponowne sprawdzenie niestacjonarności zmiennych	11
4.4	Usunięcie zmiennych o zerowej wariancji	11
4.5	Po usunięciu	12
5	Metoda doboru zmiennych	13
5.1	Metoda Hellwiga	13

1 Wprowadzenie

1.1 Opis projektu

cos tam cos tam

1.2 Zmienna objaśniana

CLOSE - zmienna objaśniana będzie rentowność 10-letnich polskich obligacji skarbowych

1.3 Zmienne objaśniające

XAUUSD - cena złota w dolarze amerykańskim

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

PMI - cos tam

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 r/r wobec miesiąca odpowiadającego z roku temu

1.4 Źródła

www.stooq.com

2 Wczytywanie danych

```
data_all <- read_excel("data.xlsx")

data_all <- data_all[, c("CLOSE", "INFLATION", "XAUUSD", "USDPLN", "WIG20", "S&P500",
                        "UNEMPLOYMENT", "PMI", "OIL")]

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]
data_test <- data_all[(train_size + 1):n, , drop = FALSE]

Y <- data["CLOSE"]
X <- data[,c("INFLATION", "XAUUSD", "USDPLN", "WIG20", "S&P500", "UNEMPLOYMENT",
            "PMI", "OIL")]
```

3 Podstawowe statystyki

3.1 Zmienna objaśniana

```
##      CLOSE
##  Min.    :1.149
## 1st Qu.:3.129
##  Median :5.112
##   Mean   :4.467
## 3rd Qu.:5.778
##   Max.   :8.337
```

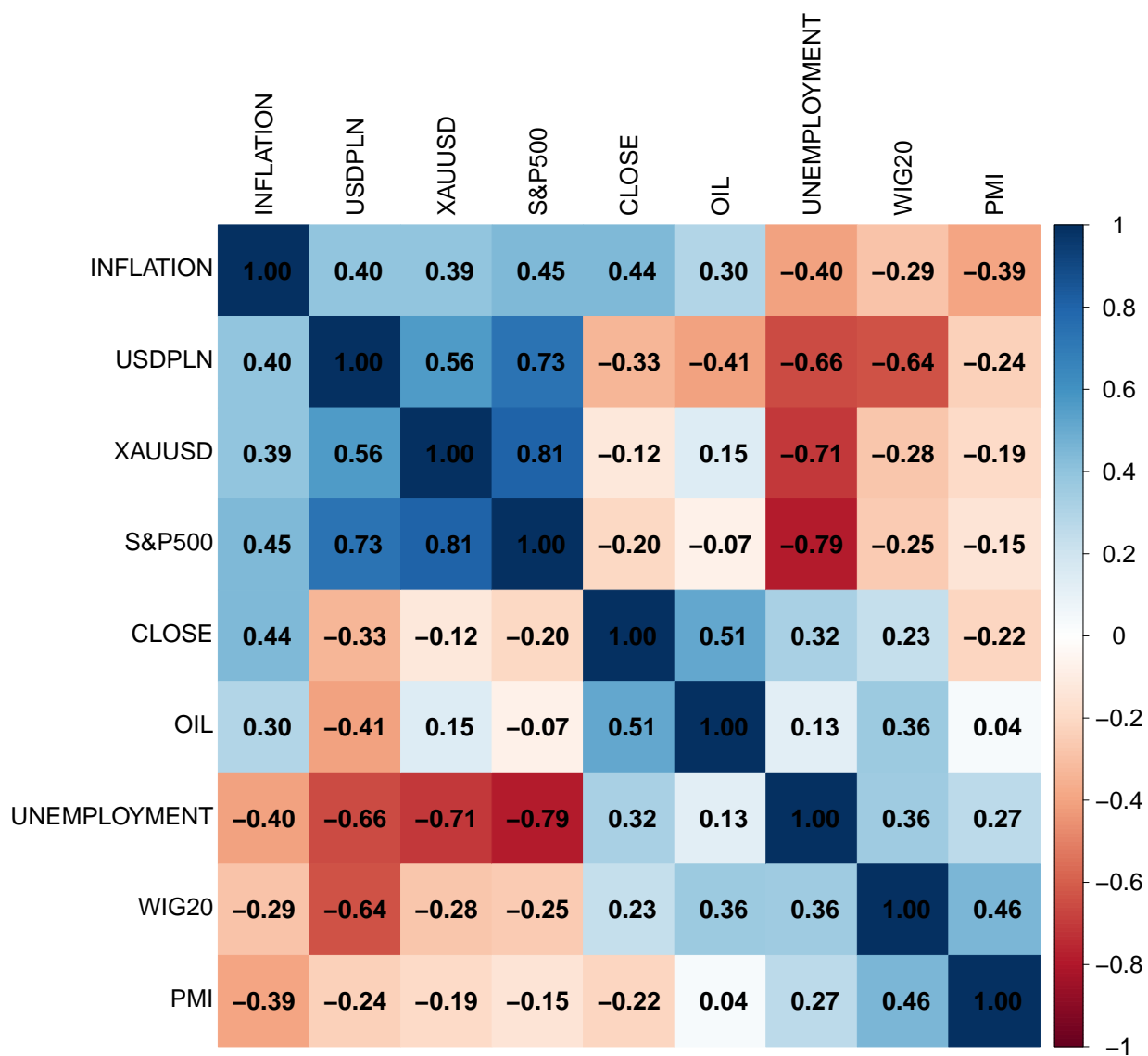
Wartości zmiennej objaśnianej wachają się pomiędzy 1,149 a 8,337. Mediana wynosi 5,112 a średnia 4,467.

3.2 Zmienne objaśniające

```
##      INFLATION      XAUUSD      USDPLN      WIG20
##  Min.    :-0.016   Min.    : 415.6   Min.    :2.060   Min.    :1372
## 1st Qu.: 0.013   1st Qu.:1056.3   1st Qu.:3.084   1st Qu.:2035
##  Median : 0.026   Median :1295.7   Median :3.509   Median :2306
##   Mean   : 0.034   Mean    :1353.6   Mean    :3.481   Mean    :2323
## 3rd Qu.: 0.043   3rd Qu.:1712.0   3rd Qu.:3.912   3rd Qu.:2487
##   Max.    : 0.184   Max.    :3288.4   Max.    :4.957   Max.    :3878
##
##      S&P500      UNEMPLOYMENT      PMI      OIL
##  Min.    : 735.1   Min.    :0.0480   Min.    :31.90   Min.    : 18.84
## 1st Qu.:1310.5   1st Qu.:0.0580   1st Qu.:48.20   1st Qu.: 54.16
##  Median :1983.6   Median :0.1015   Median :51.15   Median : 70.11
##   Mean   :2396.9   Mean    :0.0990   Mean    :50.43   Mean    : 71.68
## 3rd Qu.:3058.3   3rd Qu.:0.1265   3rd Qu.:53.30   3rd Qu.: 88.08
##   Max.    :6040.5   Max.    :0.1940   Max.    :59.40   Max.    :140.00
```

Z 11 zmiennych objaśniających wybrałem 8, których wartość bezwzględna korelacji nie przekracza 0.7.

3.3 Macierz korelacji



4 Identyfikacja niestacjonarnych zmiennych objaśniających

4.1 Sprawdzenie niestacjonarności zmiennych

Zidentyfikowane zmienne niestacjonarne:

CLOSE

INFLATION

XAUUSD

USDPLN

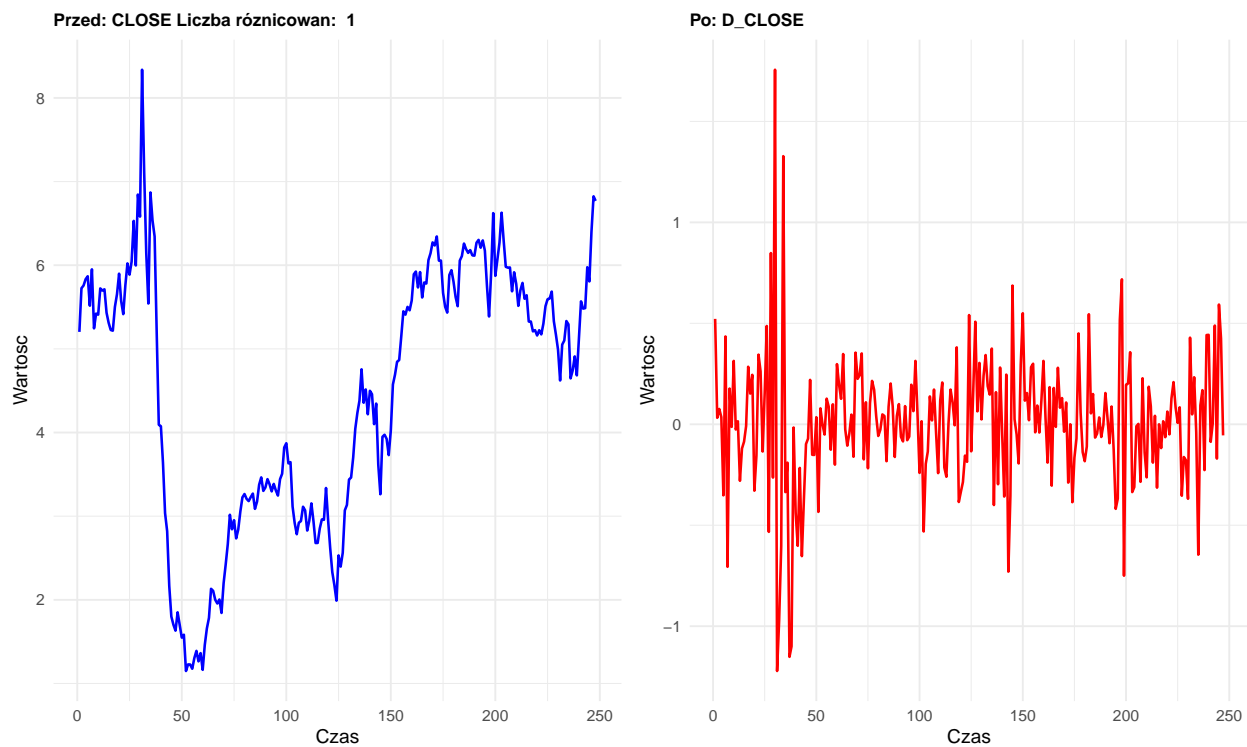
WIG20

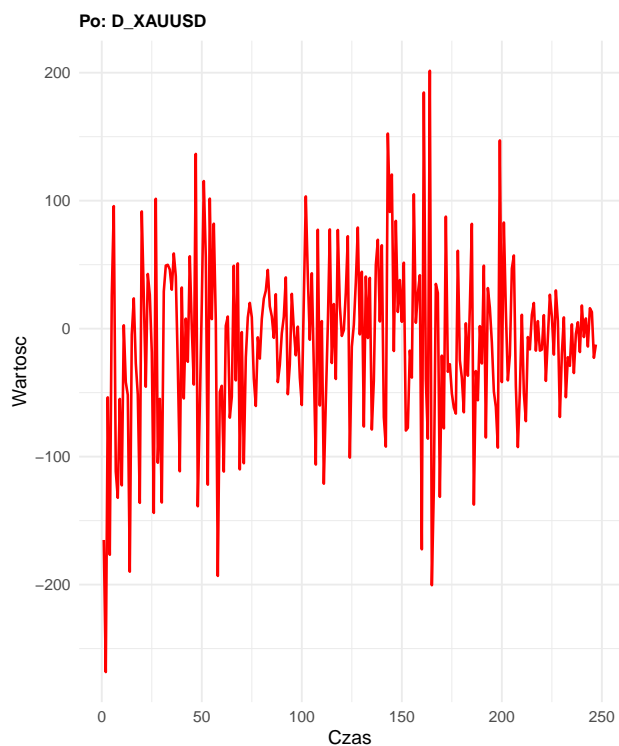
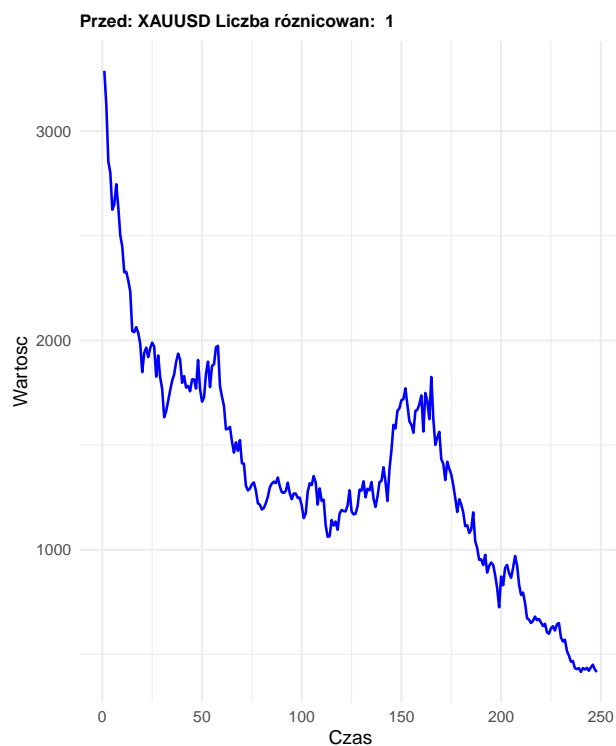
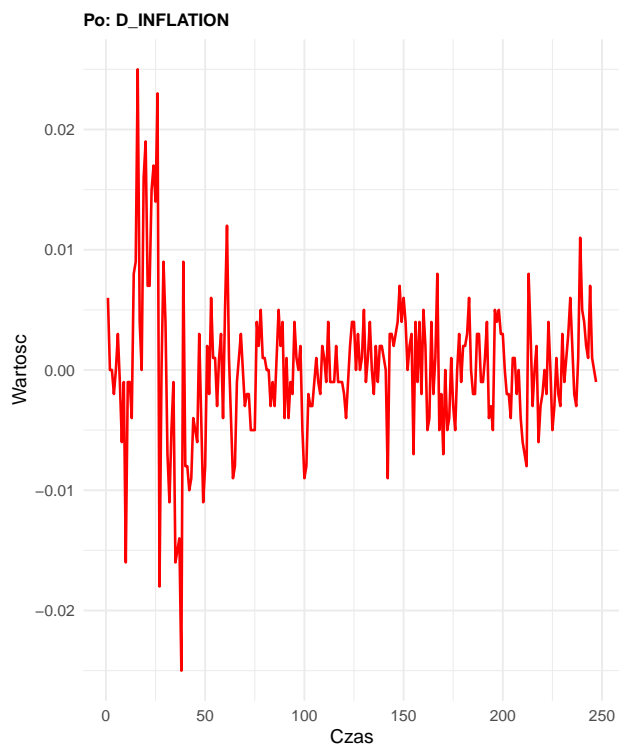
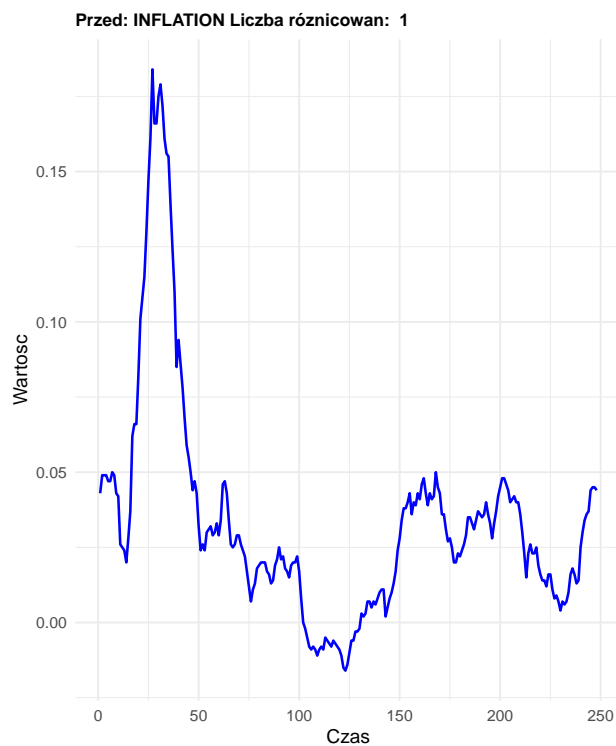
S&P500

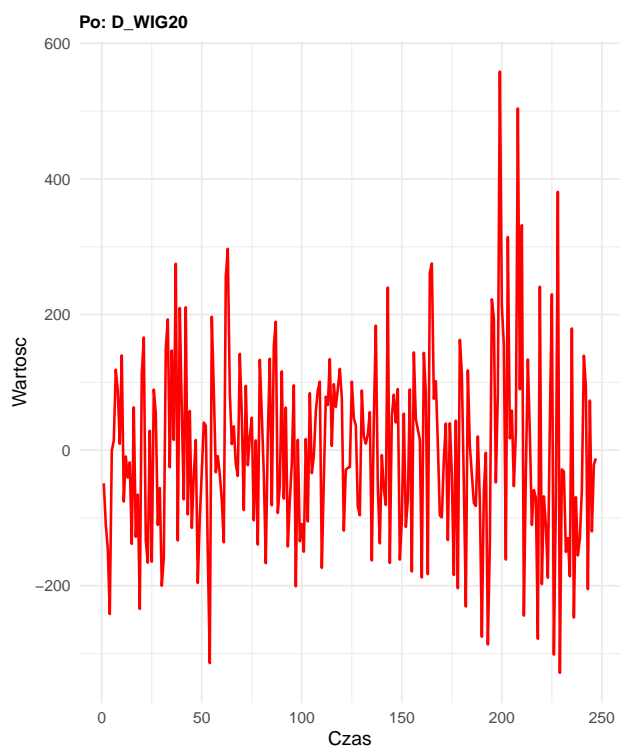
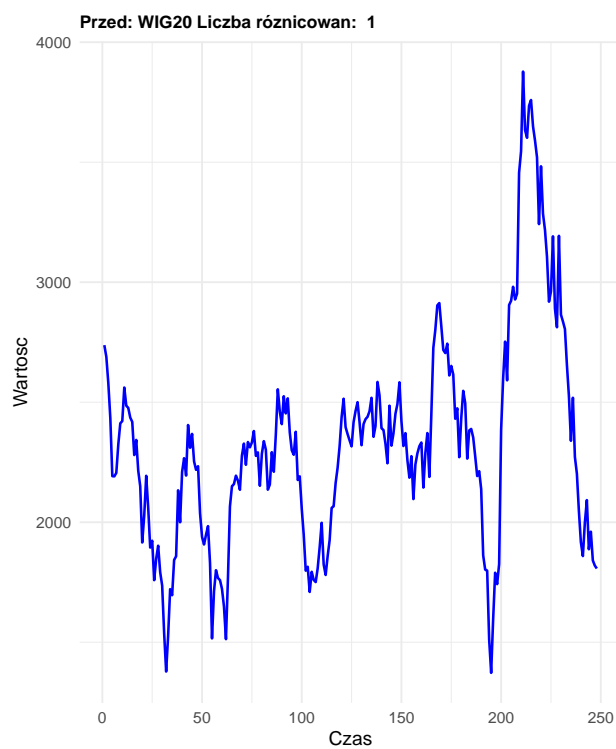
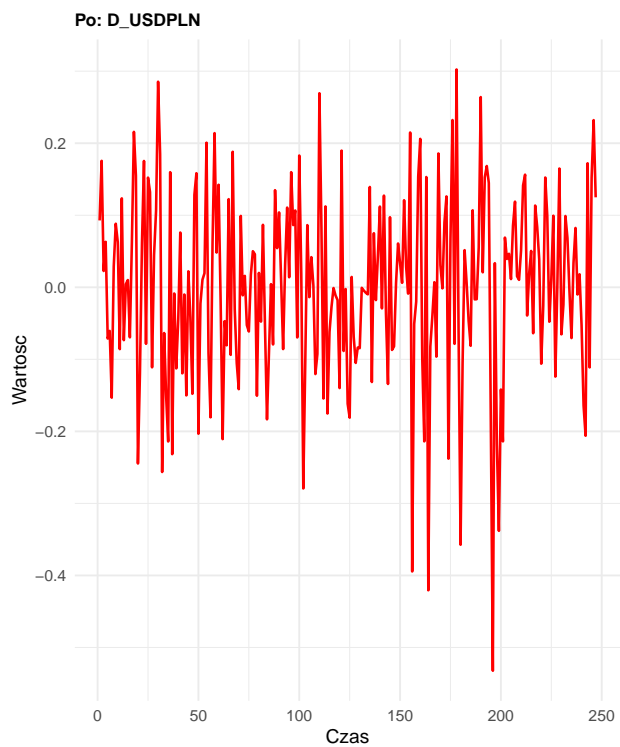
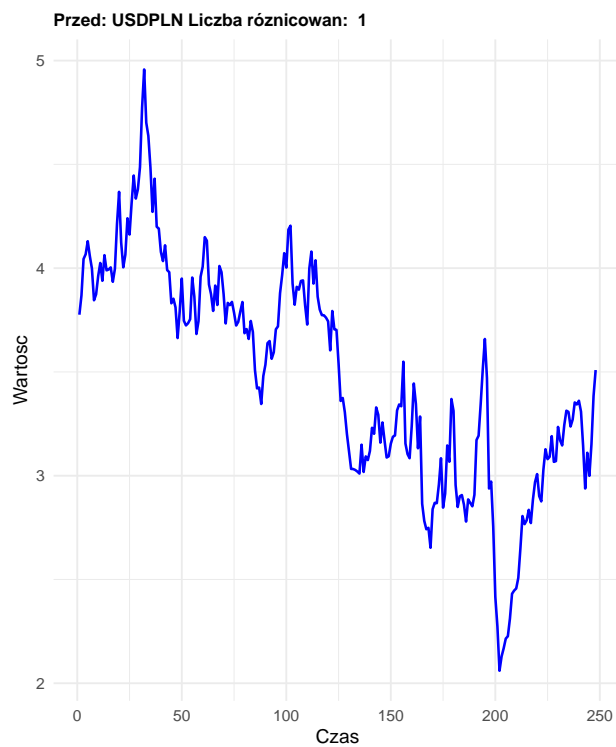
UNEMPLOYMENT

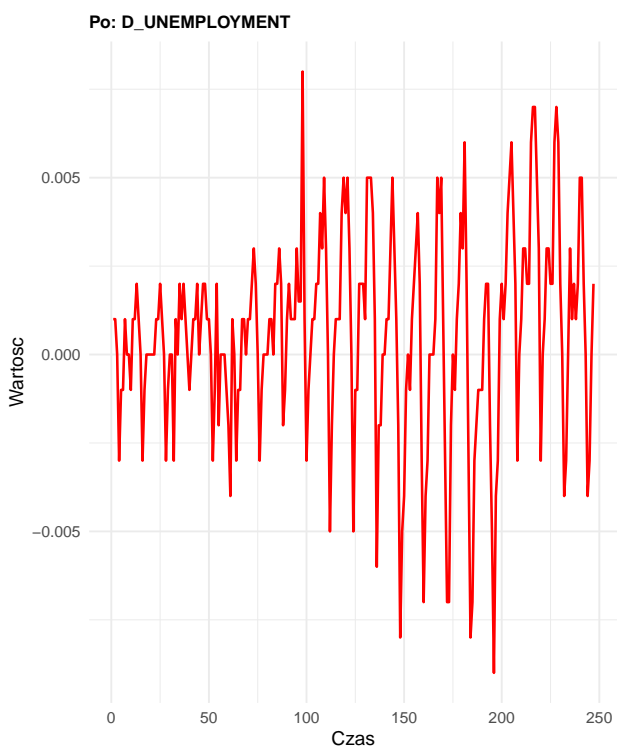
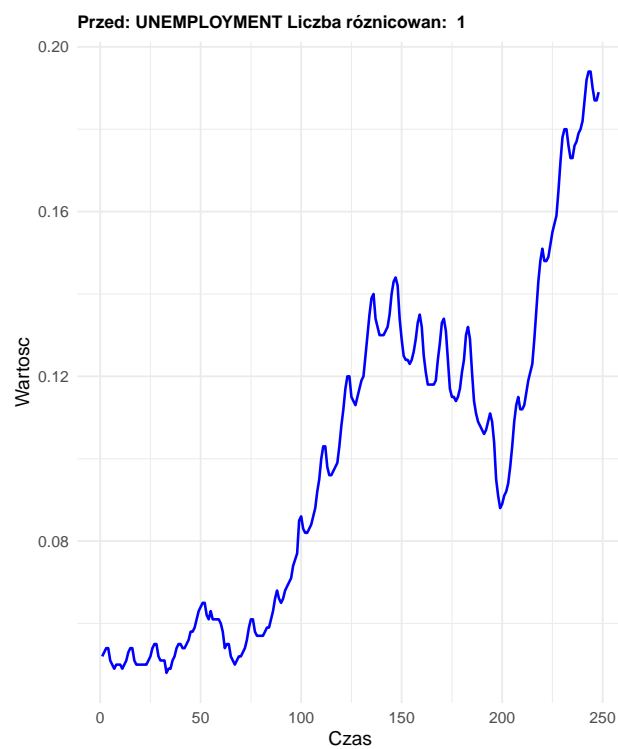
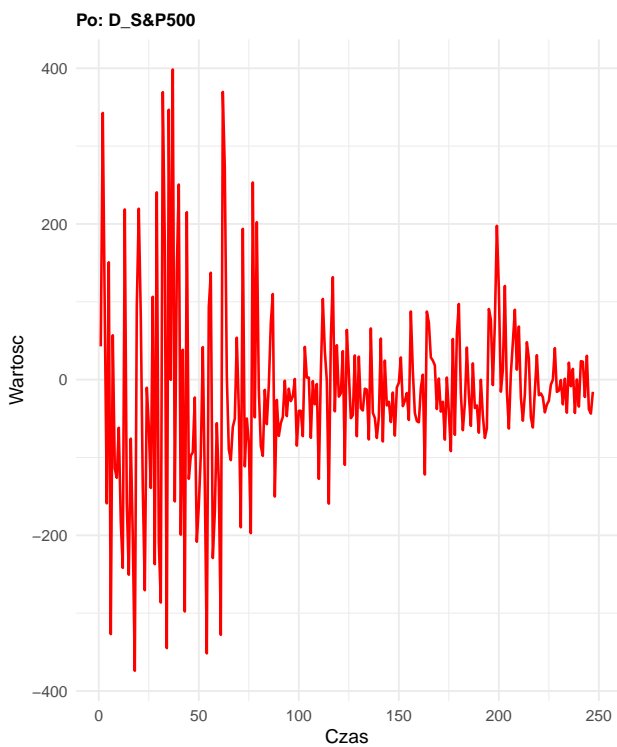
OIL

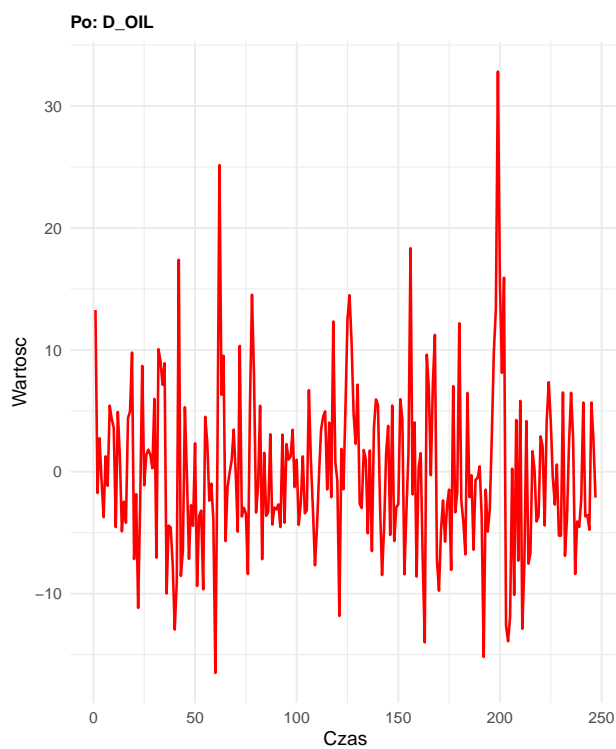
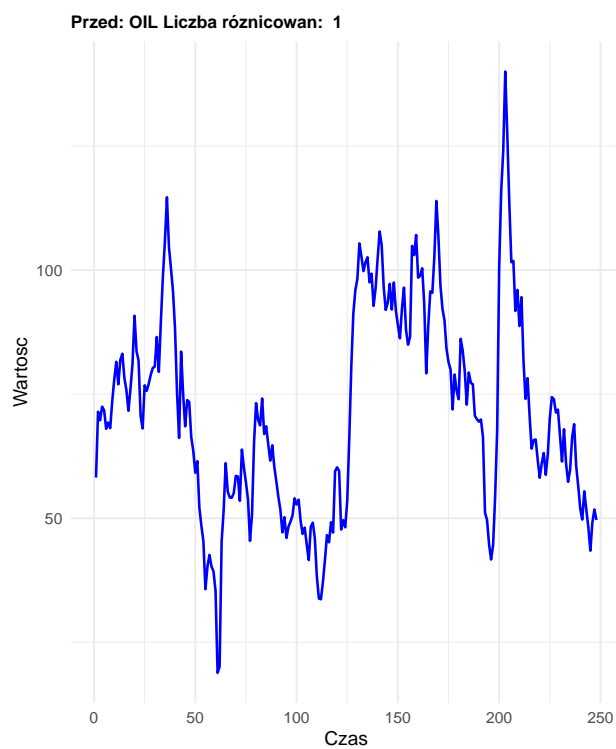
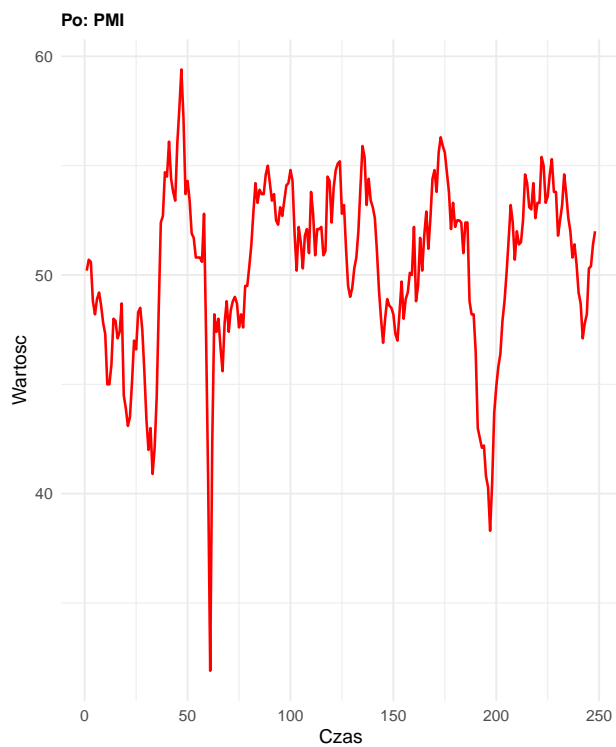
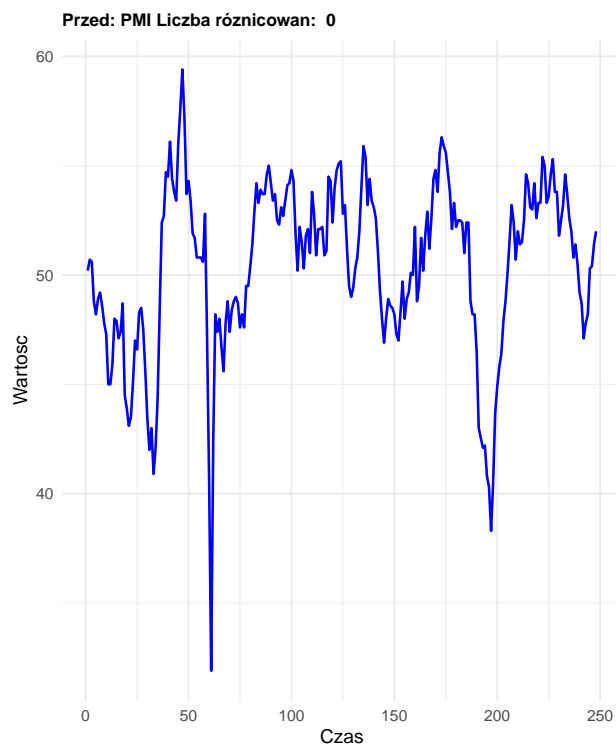
4.2 Usunięcie niestacjonarności











4.3 Ponowne sprawdzenie niestacjonarności zmiennych

D_CLOSE : Stacjonarna

D_INFLATION : Stacjonarna

D_XAUUSD : Stacjonarna

D_USDPLN : Stacjonarna

D_WIG20 : Stacjonarna

D_S.P500 : Stacjonarna

D_UNEMPLOYMENT : Stacjonarna

PMI : Stacjonarna

D_OIL : Stacjonarna

4.4 Usunięcie zmiennych o zerowej wariancji

4.4.1 ogólnie taki jest podpunkt u

4.4.2 tory ale nie usuwałem nic tylko wypisałem

D_CLOSE - Współczynnik zmienności: 5229.174 %, Wariancja: 0.1099147

D_INFLATION - Współczynnik zmienności: 143705.6 %, Wariancja: 3.384958e-05

D_XAUUSD - Współczynnik zmienności: -581.7211 %, Wariancja: 4577.532

D_USDPLN - Współczynnik zmienności: -11956.17 %, Wariancja: 0.01666743

D_WIG20 - Współczynnik zmienności: -3688.86 %, Wariancja: 19320.54

D_S.P500 - Współczynnik zmienności: -653.8075 %, Wariancja: 13902.73

D_UNEMPLOYMENT - Współczynnik zmienności: 512.851 %, Wariancja: 8.091513e-06

PMI - Współczynnik zmienności: 7.771052 %, Wariancja: 15.355

D_OIL - Współczynnik zmienności: -19693.51 %, Wariancja: 46.47127

4.5 Po usunięciu

D_CLOSE - Współczynnik zmienności: 5229.174 %, Wariancja: 0.1099147

D_XAUUSD - Współczynnik zmienności: -581.7211 %, Wariancja: 4577.532

D_WIG20 - Współczynnik zmienności: -3688.86 %, Wariancja: 19320.54

D_S.P500 - Współczynnik zmienności: -653.8075 %, Wariancja: 13902.73

PMI - Współczynnik zmienności: 7.771052 %, Wariancja: 15.355

D_OIL - Współczynnik zmienności: -19693.51 %, Wariancja: 46.47127

5 Metoda doboru zmiennych

5.1 Metoda Hellwiga

Zmienne składowe w najlepszej kombinacji:

D_XAUUSD

D_WIG20

D_S.P500

PMI

D_OIL

Pojemność Hellwiga dla tej kombinacji: 0.1257