

Tabela 8.1. Wykaz pakietów programu R wykorzystanych w części C

Nazwa pakietu	Opis zastosowania
readxl	obsługa plików programu MS Excel
lmtest	diagnostyka modeli regresji liniowej
sandwich	wyznaczanie odpornych estymatorów błędów standardowych
nortest, PowerR	testy normalności rozkładu
strucchange	identyfikacja zmian strukturalnych modelu
Metrics	błędy predykcji
skedastic	testy heteroskedastyczności składnika losowego
tseries, forecast	testowanie stacjonarności szeregów czasowych
performance	ocena jakości modelu
InformationValue	ocena jakości modelu logitowego
OptimalCutpoints	wyznaczanie optymalnych punktów odcięcia

Wydruk 8.1. Statystyki opisowe zmiennych z przykładu 1.1

```
dane <- read_excel("Dane_koniunktura.xlsx")
summary(dane)
```

rok	KP	KB
Min. :1997-01-01 00:00:00	Min. :-18.60	Min. :-16.000
1st Qu.:2002-09-23 12:00:00	1st Qu.: 3.15	1st Qu.: 0.400
Median :2008-06-16 00:00:00	Median : 10.15	Median : 4.200
Mean :2008-06-16 01:44:20	Mean : 9.67	Mean : 4.857
3rd Qu.:2014-03-08 18:00:00	3rd Qu.: 17.45	3rd Qu.: 10.100
Max. :2019-12-01 00:00:00	Max. : 28.60	Max. : 23.100

  

CP	poliSLD	poliPO	poliPiS
Min. : 96.6	Min. :0.0000	Min. :0.0000	Min. :0.0000
1st Qu.:100.3	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000
Median :102.4	Median :0.0000	Median :1.0000	Median :0.0000
Mean :102.9	Mean :0.2101	Mean :0.5254	Mean :0.2645
3rd Qu.:104.5	3rd Qu.:0.0000	3rd Qu.:1.0000	3rd Qu.:1.0000
Max. :113.0	Max. :1.0000	Max. :1.0000	Max. :1.0000

Do oszacowania parametrów modelu (1.1) z wykorzystaniem KMNK wystarczy jedno polecenie `lm`. Wyniki estymacji uzyskane poleceniem `summary` przedstawiamy na wydruku 8.2. Analogicznie do wydruku 1.2 z części A zaznaczono pogrubioną czcionką najważniejsze elementy: wartości współczynników, empiryczne poziomy istotności

$p$ , współczynnik determinacji  $R^2$  i test Walda (statystyka  $F$ ). Zwróćmy uwagę, że liczba gwiazdek w wierszu związanym z daną zmienną, świadcząca o jej istotności statystycznej, pojawia się przy innych poziomach istotności niż w programie gretl.

**Wydruk 8.2.** Wyniki estymacji parametrów modelu oczekiwań koniunkturalnych w programie R

```

kmnk <- lm(data = dane, formula = KP ~ KB + CP + poliPiS + poliSLD)
summary(kmnk)

Call:
lm(formula = KP ~ KB + CP + poliPiS + poliSLD, data = dane)

Residuals:
    Min       1Q   Median       3Q      Max
-17.1595  -4.3492   0.1716   4.5905  14.7062

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -95.21985    11.50697  -8.275 5.88e-15 ***
KB           0.64626     0.04945  13.068 < 2e-16 ***
CP           0.98299     0.11196   8.780 < 2e-16 ***
poliPiS      -0.76298     0.97434  -0.783 0.434272
poliSLD       3.88437     0.98479   3.944 0.000102 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.198 on 271 degrees of freedom
Multiple R-squared:  0.5751, Adjusted R-squared:  0.5689
F-statistic: 91.71 on 4 and 271 DF, p-value: < 2.2e-16

```

Przypomnijmy, że korzystając z programu gretl, należało pamiętać, iż zmienna *poliPO* została wybrana jako bazowa i nie należy dodawać jej *explicite* do modelu. W R, tak jak i w gretlu, można wpisać wszystkie zmienne, a program przy wykryciu liniowej współzależności pomiędzy zmiennymi automatycznie dokonuje wyboru kategorii bazowej<sup>2</sup>. Łatwo też, poleceniem `plot`, można uzyskać sześć wykresów (wybieranych parametrem *which*), pozwalających na wzrokową ocenę jakości dopasowania modelu czy identyfikację obserwacji odstających. Przykładowo na rysunku 8.1 widzimy reszty modelu – dla idealnego dopasowania układałyby się one dokładnie na linii poziomej. Przy niektórych punktach znajdują się numery obserwacji, dla których oszacowania najbardziej odbiegają od wartości faktycznych.

<sup>2</sup> Wystarczy wykonać polecenie:

```
summary(lm(data=dane, formula=KP ~ KB + CP + poliPiS + poliSLD + poliPO))
```