

Prognozowanie PKB Republiki Czeskiej na podstawie danych historycznych

Jakub Kur

Zmienną prognozowaną w badaniu jest Produkt krajowy brutto wytworzony przez Czechy. Dane przyjęte do analizy pobrane zostały z bazy danych [Europejskiego Urzędu statystycznego \(EuroStat\)](#). Zestaw danych zawiera dane kwartalne o PKB (w milionach koron czeskich) w latach 1995 – 2020.

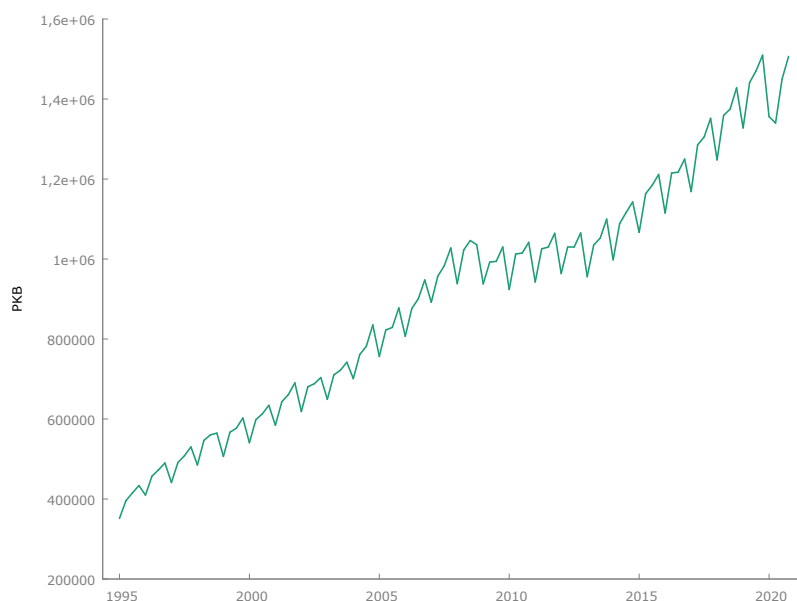
STATYSTYKI OPISOWE

Średnia	903990
Mediana	945160
Minimalna	351520
Maksymalna	1509400
Odchylenie standardowe	302260
Wsp. zmienności	0,3344
Skośność	0,0838
Kurtoza	-0,9371
Percentyl 5%	435550
Percentyl 95%	1437900
Zakres Q3-Q1	460450
Brakujące obs.	0

W badanym okresie Czechy wytwarzały PKB średnio na poziomie **903990 mln** koron czeskich. Przez połowę tego okresu wartość ta była nie więcej niż **945160 CZK**. Największy był w 4 (**1509400 CZK**) kwartale 2019 roku a najniższy w pierwszym badanym okresie (**351520 CZK**). PKB Czech różni się przeciętnie od wartości średniej o +/- **302260 CZK** i odchyła o **33%**. Ujemna kurtoza wskazuje na platykurtyczność (spłaszczenie rozkładu) i wskazuje na dużą liczbę wyników

skrajnych. Asymetria rozkładu jest umiarkowana. Skierowana prawostronnie. W 5% badanego czasu wartość nie przekraczała **435550 CZK** i w 95% nie była mniejsza od **1437900 CZK**. 50% środkowych obserwacji zawiera **460450 CZK**.

Wykres 1. PKB Republiki czeskiej. Lata 1995-2020 (dane kwartalne).



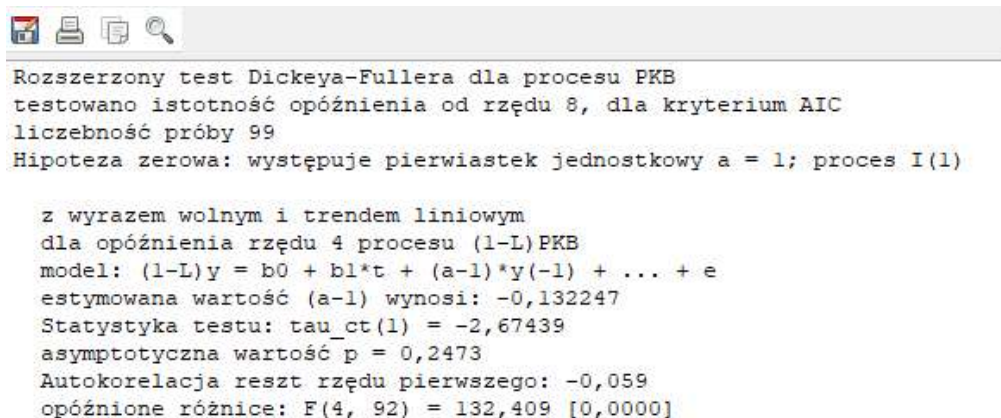
Źródło: Opracowanie własne na podstawie danych eurostat.

Rozpatrując jak kształtowały się dane w tym okresie zauważyć można występowanie sezonowości widoczne jest to na powyżej zbudowanym wykresie. Świadczą o tym regularne wahania na przestrzeni kwartałów.

TEST ADF

H_0 : niestacjonarność zmiennej

H_1 : stacjonarność zmiennej



```
Rozszerzony test Dickeya-Fullera dla procesu PKB
testowano istotność opóźnienia od rzędu 8, dla kryterium AIC
liczebność próby 99
Hipoteza zerowa: występuje pierwiastek jednostkowy a = 1; proces I(1)

z wyrazem wolnym i trendem liniowym
dla opóźnienia rzędu 4 procesu (1-L)PKB
model: (1-L)y = b0 + b1*t + (a-1)*y(-1) + ... + e
estymowana wartość (a-1) wynosi: -0,132247
Statystyka testu: tau_ct(1) = -2,67439
asymptotyczna wartość p = 0,2473
Autokorelacja reszt rzędu pierwszego: -0,059
opóźnione różnice: F(4, 92) = 132,409 [0,0000]
```

Autokorelacja reszt rzędu pierwszego:

$-0,059 < 0,1$ – Można zakładać, że opóźnienie do 8 rzędu jest właściwie dobrane i autokorelacja składnika zakłócającego w modelu pomocniczym testu ADF nie występuje.

Statystyka testu:

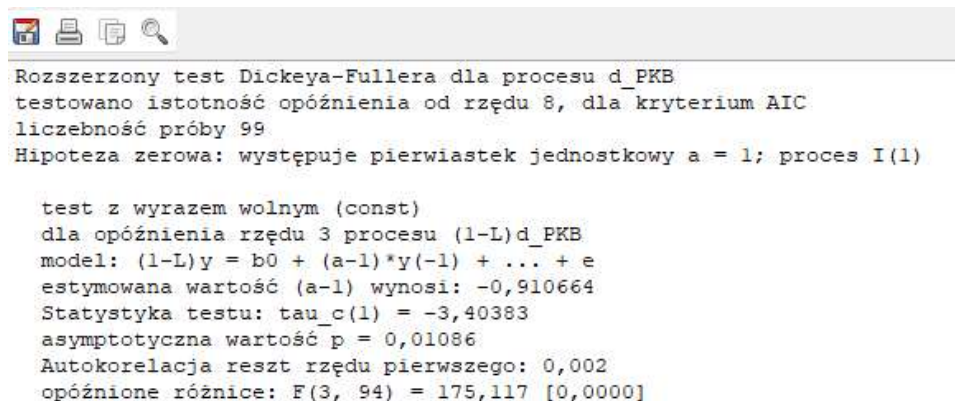
TAU = -2,674 [0,247]

$p = 0,247 > \alpha = 0,1$ – Brak podstaw do odrzucenia hipotezy zerowej mówiącej o niestacjonarności PKB czech.

Przeprowadzono test integracji pierwszego przyrostu:

H_0 : niestacjonarność pierwszego przyrostu

H_1 : stacjonarność pierwszego przyrostu



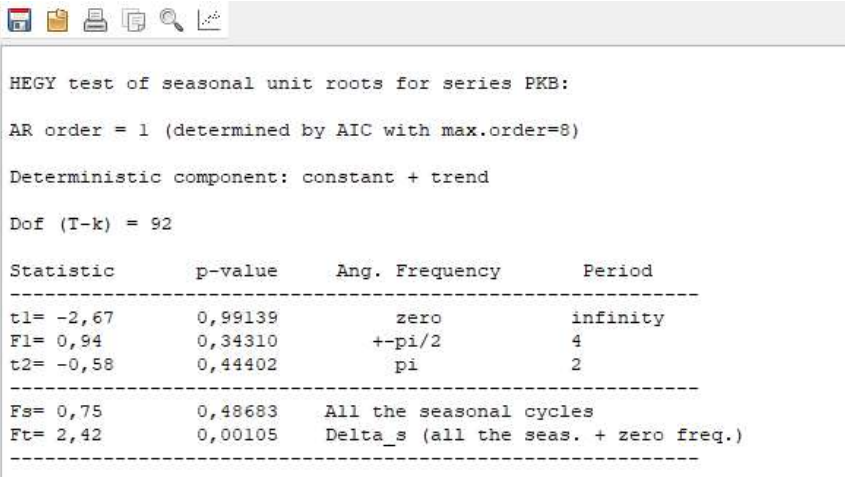
```
Rozszerzony test Dickeya-Fullera dla procesu d_PKB
testowano istotność opóźnienia od rzędu 8, dla kryterium AIC
liczebność próby 99
Hipoteza zerowa: występuje pierwiastek jednostkowy a = 1; proces I(1)

test z wyrazem wolnym (const)
dla opóźnienia rzędu 3 procesu (1-L)d_PKB
model: (1-L)y = b0 + (a-1)*y(-1) + ... + e
estymowana wartość (a-1) wynosi: -0,910664
Statystyka testu: tau_c(1) = -3,40383
asymptotyczna wartość p = 0,01086
Autokorelacja reszt rzędu pierwszego: 0,002
opóźnione różnice: F(3, 94) = 175,117 [0,0000]
```

$\tau = -3,404 [0,011]$

$p = 0,011 < \alpha = 0,1$ – Odrzucana jest hipoteza zerowa na rzecz alternatywnej. I przyrost zmiennej jest stacjonarny. Dlatego ΔPKB także jest szeregiem stacjonarnym. Wtedy $PKB \sim I(1)$ jest stacjonarne. Dlatego $d = 1$

TEST HEGY



HEGY test of seasonal unit roots for series PKB:

AR order = 1 (determined by AIC with max.order=8)

Deterministic component: constant + trend

Dof (T-k) = 92

Statistic	p-value	Ang. Frequency	Period
t1= -2,67	0,99139	zero	infinity
F1= 0,94	0,34310	$\pm\pi/2$	4
t2= -0,58	0,44402	π	2
Fs= 0,75	0,48683	All the seasonal cycles	
Ft= 2,42	0,00105	Delta_s (all the seas. + zero freq.)	

$t_1 = -2,67 [0,991] > \alpha$ – Nie można odrzucić H_a

$t_2 = -0,58 [0,343] > \alpha$ – Nie można odrzucić H_b

$F_1 = 0,94 [0,444] < \alpha$ – Odrzucamy H_0

Podsumowanie: Nie można odrzucić hipotez H_A i H_B natomiast można odrzucić hipotezę H_C .

Stacjonarną zmienną można otrzymać poprzez $y_t - y_{t-2}$. Ze względu na brak możliwości obliczenia zastosowano filtr $\Delta_4 y_t$. Dlatego można wyprowadzić wniosek, że $D = 1$.

MODEL SARIMA

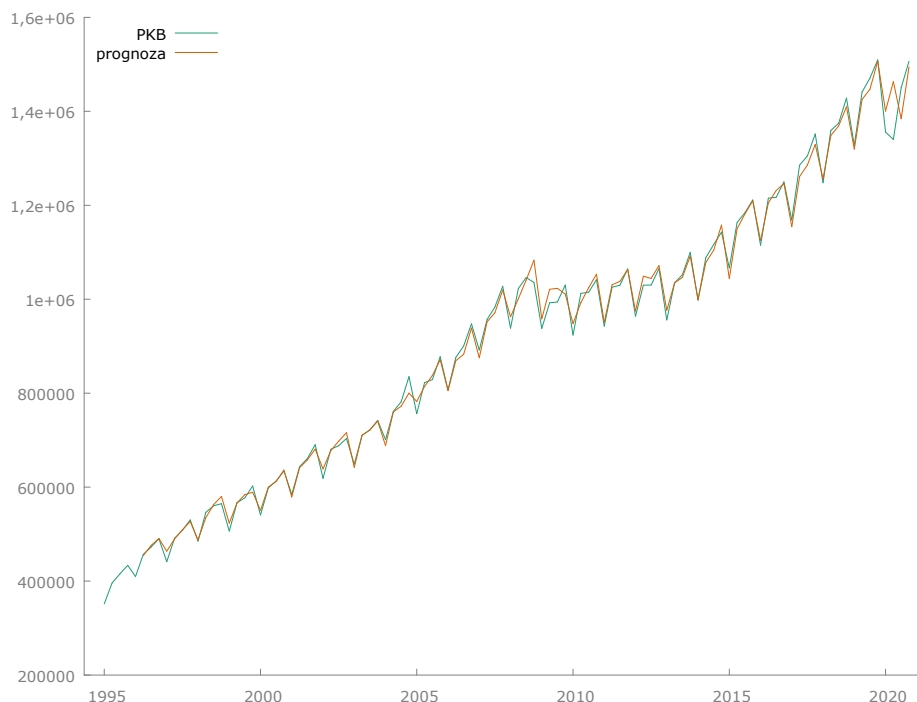
p	d	q	P	D	Q	AIC
1	1	1	1	1	1	2259,968
1	1	1	2	1	1	2261,166
1	1	1	1	1	2	2262,127
2	1	1	1	1	1	2261,967
1	1	2	1	1	1	2261,966

Na podstawie kryterium **AIC** wybrać można najlepszy model:

SARIMA (1,1,1) α (1,1,1)₄

PROGNOZOWANIE

EX POST



Wartości prognozowane pokrywają się z danymi zaobserwowanymi. Poniżej zestawiono miary dokładności prognoz ex post z wykorzystaniem 99 obserwacji.

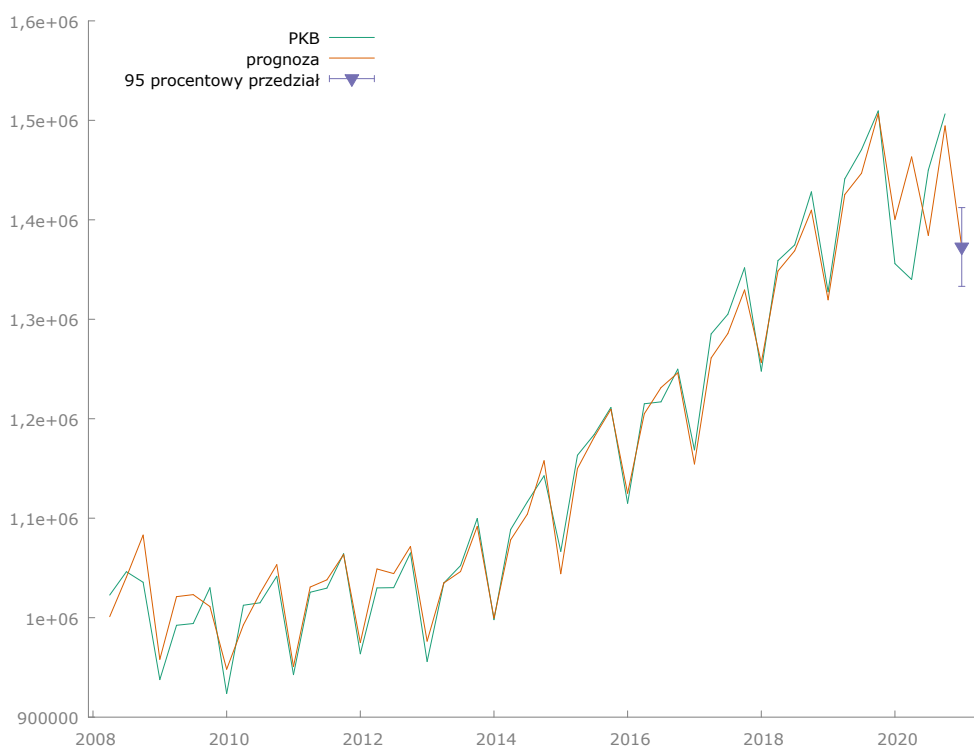
Miary dokładności prognoz ex post wykorzystano 99 obserwacji

Średni błąd predykcji	ME =	-560
Pierwiastek błędu średniokwadr.	RMSE =	20197
Średni błąd absolutny	MAE =	12949
Średni błąd procentowy	MPE =	-0,13699
Średni absolutny błąd procentowy	MAPE =	1,3507
Współczynnik Theila (w procentach)	I =	0,28309
Udział obciążoności predykc.	$I1^2/MSE =$	0,00076875
Udział niedost. elastyczności	$I2^2/MSE =$	0,0011044
Udział niezgodności kierunku	$I3^2/MSE =$	0,99813

- **ME** – średnia arytmetyczna błędów predykcji: prognozy są przeciętnie przeszacowane o **560** mln koron czeskich;
- **RMSE** – średni błąd prognozy ex post (pierwiastek z MSE): wartości prognoz różnią się od wartości prognozowanych średnio o +/- **20197** mln koron czeskich;
- **MAE** – średni błąd absolutny: wartości prognoz różnią się od wartości prognozowanych średnio o +/- **12949** mln koron czeskich;
- **MPE** – średni błąd procentowy: miara nie ma ekonometrycznego znaczenia; przeciętnie błąd prognozy stanowi **-0,13699%**;

- **MAPE** – średni absolutny błąd procentowy: wartości prognoz różnią się od wartości prognozowanych średnio o +/- **1,3507%**. Prognoza jest dopuszczalna, granica dopuszczalności prognozy wynosi 5%;
 - **I** – współczynnik Theila (w procentach): prognoza jest niedoszacowana o **0,28309%**, co jest wynikiem dopuszczalnym (granica dopuszczalności jest 10%).
1. **I²₁ (Um)** – obciążoność predykcji: błędy prognoz w **0,076875%** powstają na skutek obciążenia metody estymacji;
 2. **I²₂ (Us)**– niedostateczność elastyczności: błędy prognoz w **0,11044%** powstają na skutek niedostosowania elastyczności prognozy do faktycznych wahań zmiennej prognozowanej;
 3. **I²₃ (Uc)**– niewłaściwe prognozowanie punktów zwrotnych: błędy prognoz w **99,813%** powstają na skutek braku zgodności między kierunkami zmian zmiennej prognozowanej i jej prognoz.

EX ANTE



Wartość prognozowana PKB Republiki Czeskiej dla pierwszego kwartału 2021 roku wyniosła 1372694,27 mln koron czeskich. Jednocześnie odchylając się średnio od zmiennej prognozowanej o +/- 20197,405 mln koron czeskich. Stanowi to 0.01% wartości prognozy. Prognoza jest dopuszczalna. Przedział ufności z wartościami krańcowymi 1333108,08 i 1412280,45 mln koron czeskich z 95% prawdopodobieństwem. Pokrywa wartością prognozy dla PKB w Q1 2021.