

PROIECT PREVIZIUNI ECONOMICE

Bonaț Paula-Mihaela

Statistică și Previziune Economică, anul III

Învățământ cu frecvență

Cuprins

Cronograma.....	3
Statistici descriptive	5
Sezonalitate	7
Ciclicitate	10
Testul ADF.....	12
Corelograma.....	13
Testul de cauzalitate Granger	13
Testul Eagle-Granger	14
Previziuni	15
Previziunea naivă	15
Metoda valorilor trecute.....	15
Metode de netezire	16
Pe baza mediei mobile	16
Netezire exponențială.....	16
Previziune pe baza trendului	17
Valori drummy + trend(timp)	17

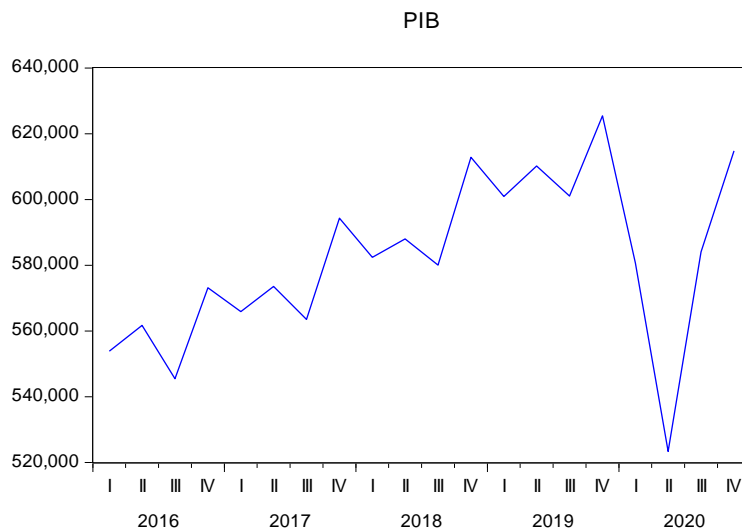
Baza de date: PIB (variabilă dependentă), EXPORT și TAXE PE PRODUSE pentru Franța

Perioada analizată: 20016Q1-2020Q4

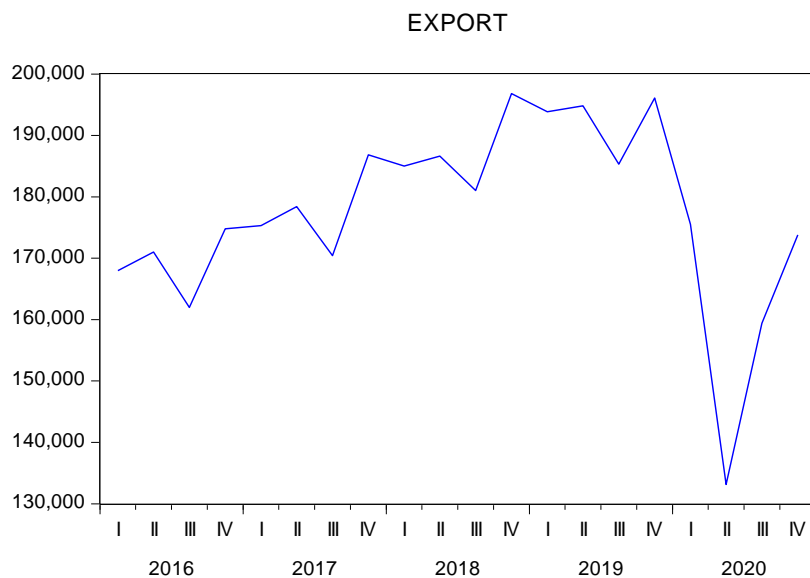
Sursa: Eurostat

Soft-urile utilizate: Excel și EViews

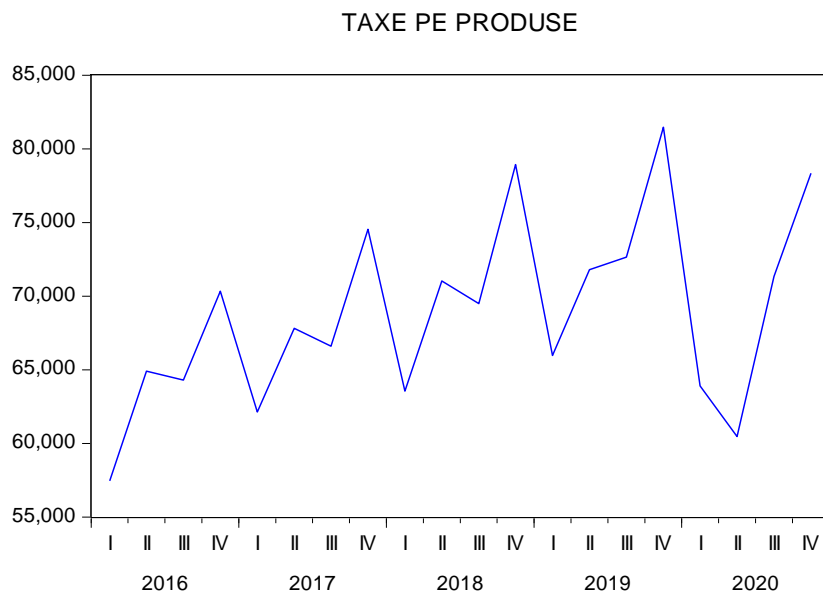
Cronograma



Mai sus este redată evoluția trimestrială a PIB-ului Franței în perioada 2016-2020. După cum se poate observa seria are o tendință crescătoare în perioada 2016Q1-2019Q4, după care are loc o prăbușire în 2020Q1 din cauza pandemiei provocată de virusul SARS-COV2. Se mai poate observa componenta sezonieră care înregistrează minime în trimestrele Q3 și maxime în trimestrele Q4, precum și prezența componentei aleatoare.



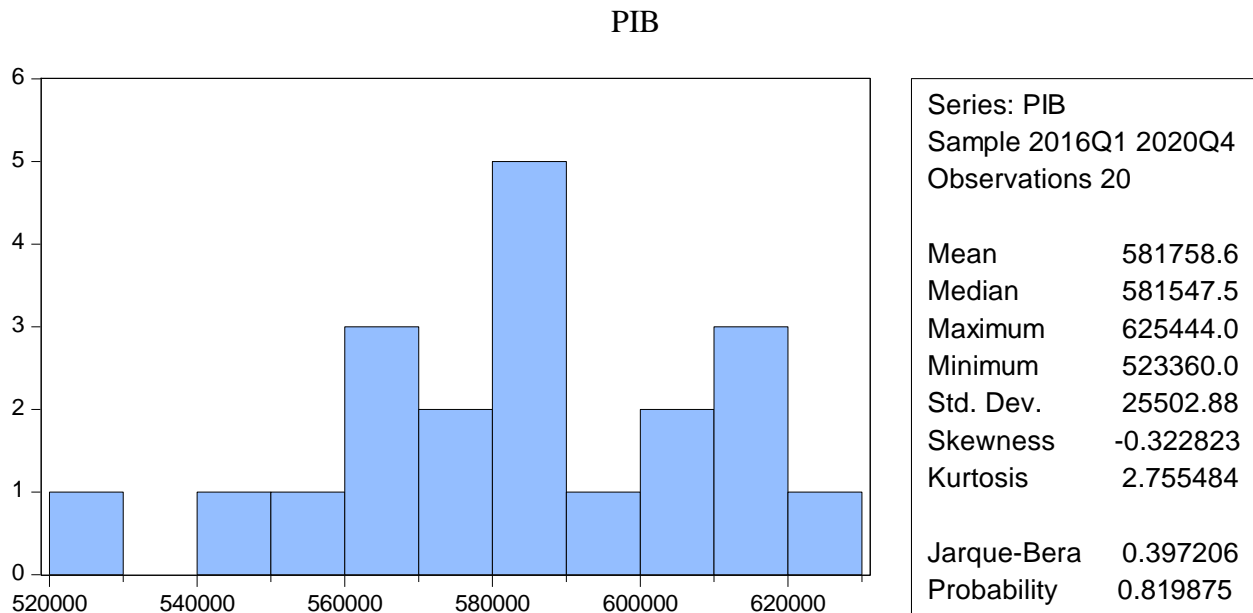
Mai sus este redată evoluția trimestrială a exporturilor în Franța în perioada 2016-2020. După cum se poate observa seria are o tendință crescătoare în perioada 2016Q1-2019Q4, după care are loc o prăbușire în 2020Q1 din cauza pandemiei provocată de virusul SARS-COV2. Se mai poate remarca și prezența componentei sezoniere care înregistrează minime în trimestrele Q3 și maxime în trimestrele Q4, precum și prezența componentei aleatoare.



Mai sus este redată evoluția trimestrială a taxelor pe produs în Franța în perioada 2016-2020. După cum se poate observa seria are o tendință ușor crescătoare în perioada 2016Q1-2019Q4, după care

are loc o scădere în 2020Q1 din cauza pandemiei provocată de virusul SARS-COV2. Se mai poate remarca și prezența componentei sezoniere care înregistrează minime în trimestrele Q1 și maxime în trimestrele Q4, precum și prezența componentei aleatoare.

Statistici descriptive



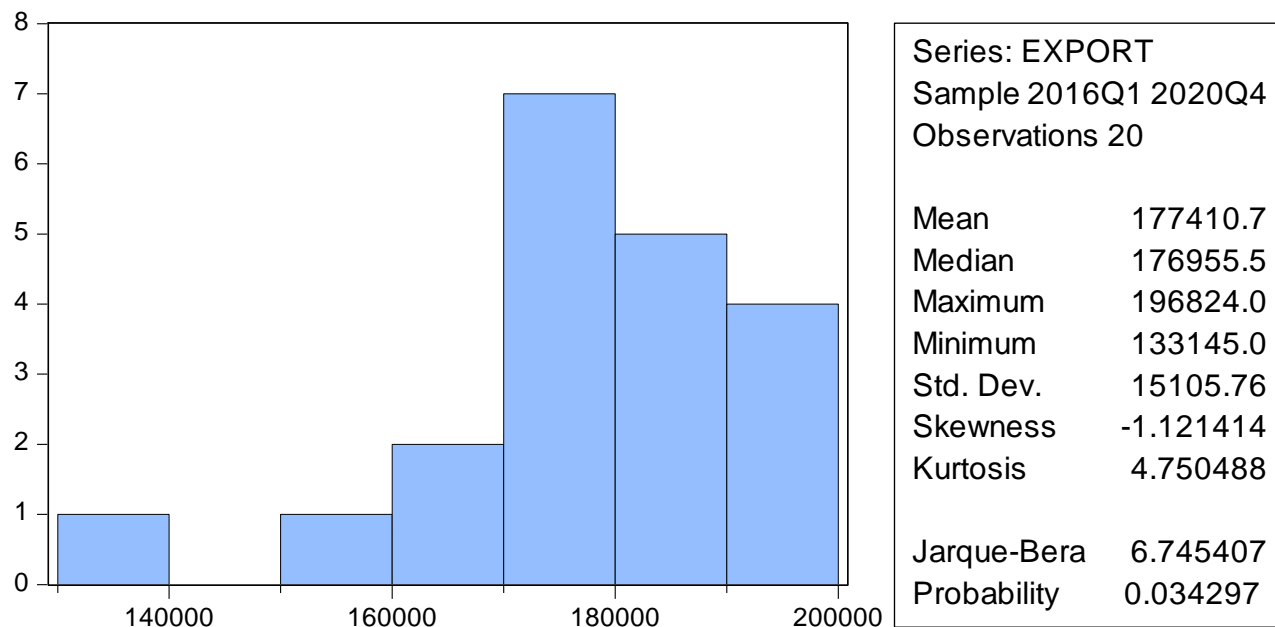
PIB-ul înregistrează o valoare medie de 581758.6 și o valoare mediană de 581547.5. Valoarea minimă din serie este de 523360.0 și valoarea maximă din serie este de 625444.0, reprezentând intervalul în care se găsesc restul valorilor. Abaterea medie pătratică este de 25502.88 și reprezintă abaterea medie a valorilor de la medie. Seria are o asimetrie negativă (de stânga) ($-0.323 < 0$) și este platikurtică ($2.755 < 3$). Testul Jarque-Bera indică dacă seria urmează o distribuție normală sau nu.

H_0 : Seria urmează o distribuție normală.

H_1 : Seria nu urmează o distribuție normală.

Probabilitatea = $0.8198 > 0.05$, deci se acceptă ipoteza nulă cu un prag de semnificație de 5%.

EXPORT



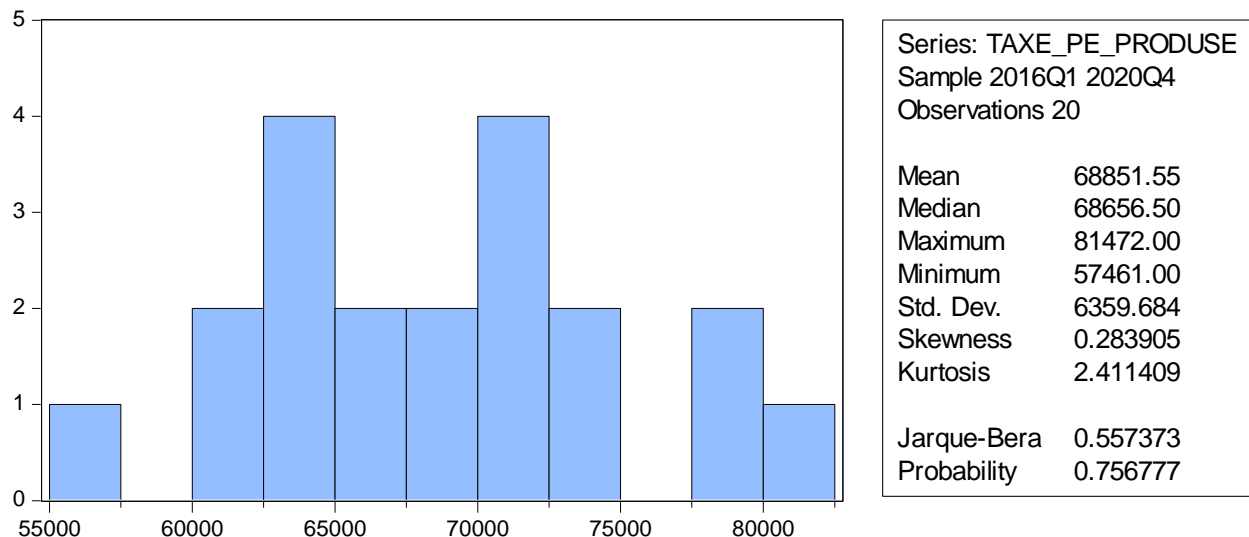
EXPORTUL înregistrează o valoare medie de 177410.7 și o valoare mediană de 176955.5. Valoarea minimă din serie este de 133145.0 și valoarea maximă din serie este de 196824.0, reprezentând intervalul în care se găsesc restul valorilor. Abaterea medie pătratică este de 15105.76 și reprezintă abaterea medie a valorilor de la medie. Seria are o asimetrie negativă (de stânga) ($-1.12 < 0$) și este leptokurtică ($4.75 > 3$). Testul Jarque-Bera indică dacă seria urmează o distribuție normală sau nu.

H_0 : Seria urmează o distribuție normală.

H_1 : Seria nu urmează o distribuție normală.

Probabilitatea = $0.03 > 0.01$, deci se acceptă ipoteza nulă cu un prag de semnificație de 1%.

TAXE PE PRODUSE



TAXELE PE PRODUSE înregistrează o valoare medie de 68851.55 și o valoare mediană de 68656.50. Valoarea minimă din serie este de 57461.00 și valoarea maximă din serie este de 81472.00, reprezentând intervalul în care se găsesc restul valorilor. Abaterea medie pătratică este de 6359.684 și reprezintă abaterea medie a valorilor de la medie. Seria are o asimetrie pozitivă (de dreapta) ($0.28 > 0$) și este platikurtică ($2.41 < 3$). Testul Jarque-Bera indică dacă seria urmează o distribuție normală sau nu.

H_0 : Seria urmează o distribuție normală.

H_1 : Seria nu urmează o distribuție normală.

Probabilitatea = $0.756 > 0.05$, deci se acceptă ipoteza nulă cu un prag de semnificație de 5%.

Sezonalitate

Variabilele au sezonalitate, așadar se vor sezonaliza în prealabil folosind mediile mobile și modelul multiplicativ (deoarece seriile au tendință).

PIB

Date: 01/03/22 Time: 23:52
Sample: 2016Q1 2020Q4
Included observations: 20
Ratio to Moving Average
Original Series: PIB
Adjusted Series: PIBSA

Scaling Factors:

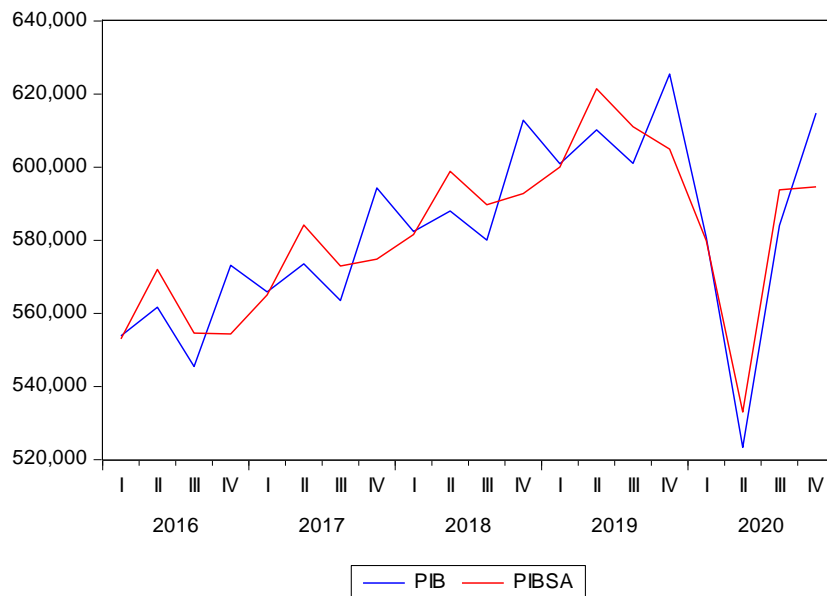
1	1.001482
2	0.981916
3	0.983579
4	1.033887

În urma procesului se obțin coeficienții sezonalityi aferenți PIB-ului.

Interpretări:

În trimestrul II, seria înregistrează valori cu 0.02% mai scăzute în medie decât media.

În trimestrul IV, seria înregistrează valori de 1.03 ori mai mari în medie decât media.



Seria desezonalizată este surprinsă în graficul de mai sus.

Observăm că variabila are o sezonalitya foarte mică.

EXPORT

Date: 01/03/22 Time: 23:56
Sample: 2016Q1 2020Q4
Included observations: 20
Ratio to Moving Average
Original Series: EXPORT
Adjusted Series: EXPORTSA

Scaling Factors:

1	1.021817
---	----------

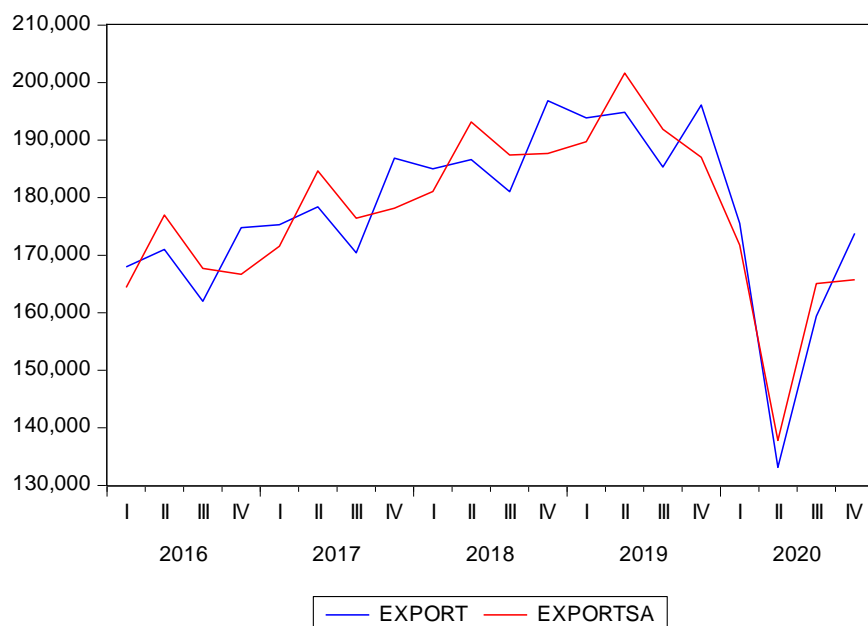
2	0.966200
3	0.965881
4	1.048664

În urma procesului se obțin coeficienții sezonaliții aferenți EXPORTULUI.

Interpretări:

În trimestrul II, seria înregistrează valori cu 0.04% mai scăzute în medie decât media.

În trimestrul IV, seria înregistrează valori de 1.04 ori mai mari în medie față de medie.



Seria desezonalizată este surprinsă în graficul de mai sus. În trimestrele I și IV valorile observate sunt peste tendință, iar valorile din trimestrele II și III sunt sub medie.

TAXE PE PRODUSE

Date: 01/03/22 Time: 23:59

Sample: 2016Q1 2020Q4

Included observations: 20

Ratio to Moving Average

Original Series:

TAXE_PE_PRODUSE

Adjusted Series: TAXE_PESA

Scaling Factors:

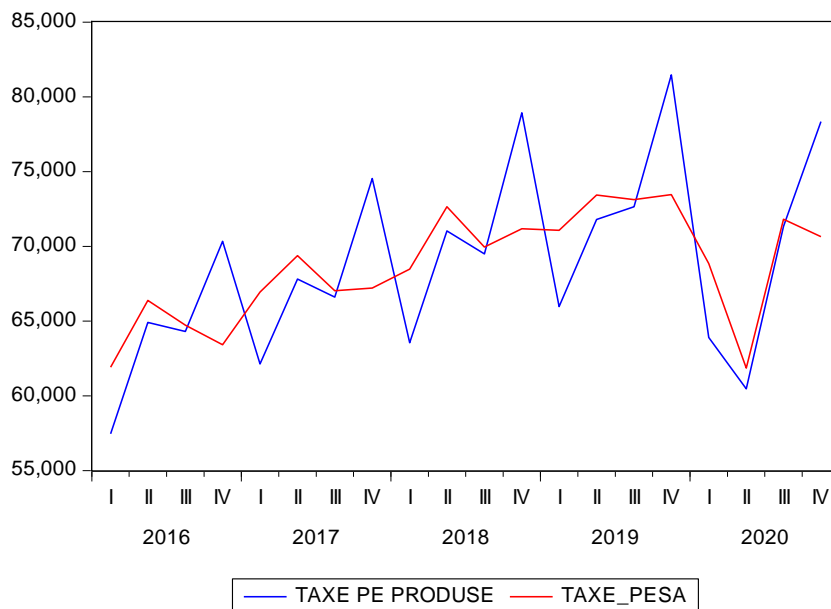
1	0.928233
2	0.977656
3	0.993525
4	1.109118

În urma procesului se obțin coeficienții sezonality aferenți TAXELOR PE PRODUSE.

Interpretări:

În trimestrul I, seria înregistrează valori cu 0.08% mai scăzute în medie decât media.

În trimestrul IV, seria înregistrează valori de 1.1 ori mai mari în medie față de medie.

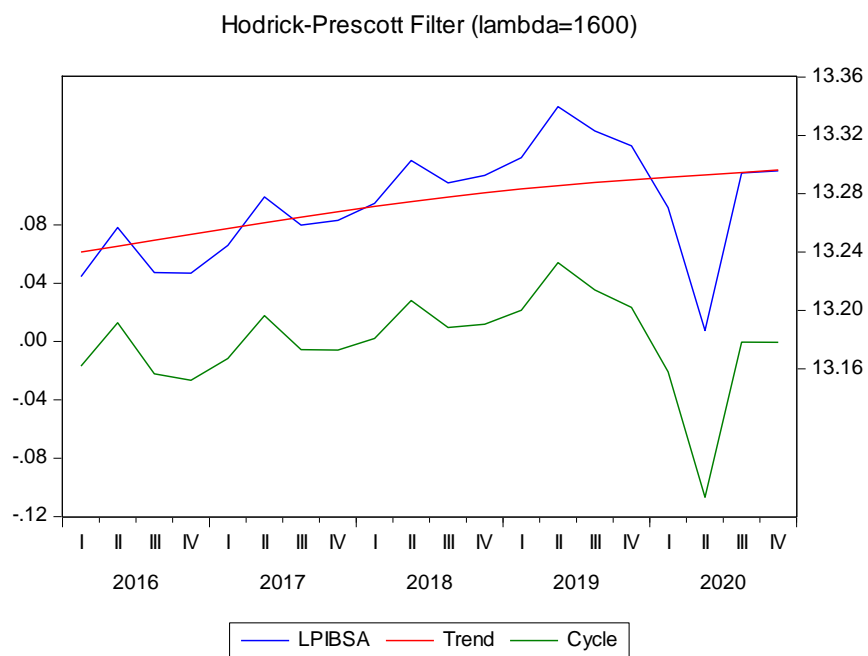


Seria desezonalizată este surprinsă în graficul de mai sus. În trimestrele I și II valorile observate sunt sub medie, valorile din trimestrele III se apropie de cele din tendință, iar valorile din trimestrele IV sunt peste medie.

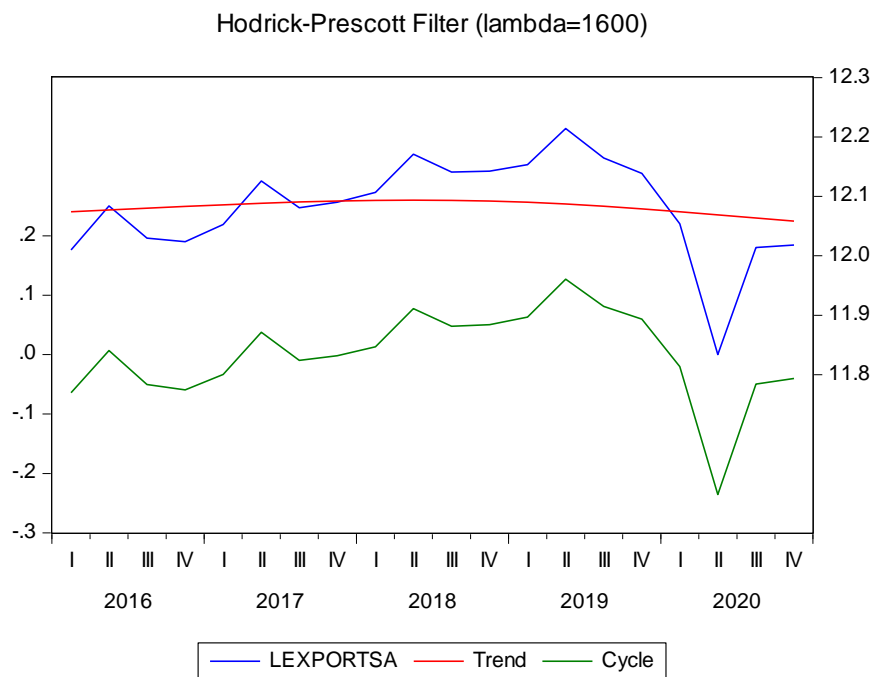
Ciclicitate

Pentru separarea tendinței și ciclului, se va folosi filtrul Hodrick-Prescott.

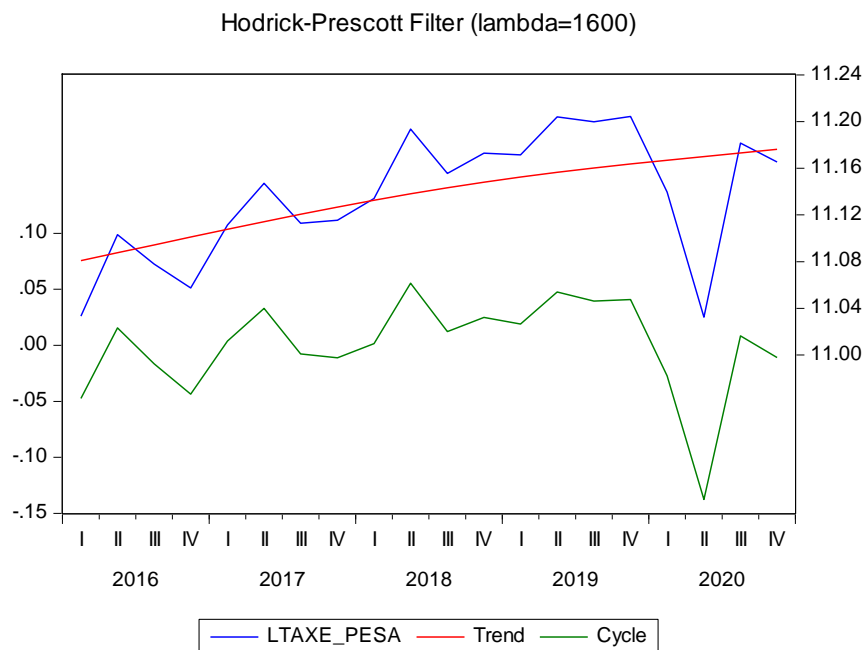
PIB



EXPORT



TAXE PE PRODUSE



Testul ADF

Testul este utilizat pentru a verifica dacă seria are rădăcină unitate (are trend aleator).

Vom aplica testul pentru variabila PIB.

Null Hypothesis: LPIBSA has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 0 (Automatic - based on SIC, maxlag=4)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.754299	0.2285
Test critical values: 1% level	-4.532598	
5% level	-3.673616	
10% level	-3.277364	

H_0 : Seria are rădăcină unitate.

H_1 : Seria este staționară.

Probabilitatea $0.2285 > 0.05$, deci ipoteza nulă se acceptă cu un prag de semnificație de 5%, rezultă că seria are trend.

Pentru a staționariza seria se vor face diferențele de ordin I.

Null Hypothesis: DLPIBSA has a unit root

Exogenous: Constant

Lag Length: 0 (Automatic - based on SIC, maxlag=3)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-5.155130	0.0007
Test critical values: 1% level	-3.857386	
5% level	-3.040391	
10% level	-2.660551	

Prob<0.05

Pentru d=1, ipoteza nulă se respinge cu un prag de semnificație de 5% și seria este staționară.

























Corelograma

Corelograma pentru seria staționarizată indică semnificativitatea coeficienților de autocorelare. După cum se poate observa, coeficienții sunt nesemnificativi. Nu există autocorelare între valorile variabilei.

Date: 01/04/22 Time: 00:46

Sample: 2016Q1 2020Q4

Included observations: 19

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.211	-0.211	0.9839	0.321
		2 -0.229	-0.286	2.2170	0.330
		3 -0.144	-0.303	2.7328	0.435
		4 0.235	0.043	4.1958	0.380
		5 -0.191	-0.273	5.2325	0.388
		6 -0.044	-0.178	5.2920	0.507
		7 0.024	-0.144	5.3110	0.622
		8 0.186	-0.016	6.5642	0.584
		9 -0.139	-0.126	7.3371	0.602
		10 -0.073	-0.171	7.5713	0.671
		11 0.074	-0.064	7.8414	0.727
		12 0.071	-0.099	8.1323	0.775

Testul de cauzalitate Granger

Testul de cauzalitate Granger va fi aplicat pe variabilele DLPIBSA si DLEXPORTSA grupate

Pairwise Granger Causality Tests
Date: 01/04/22 Time: 01:11
Sample: 2016Q1 2020Q4
Lags: 2

Null Hypothesis:	Obs	F-Statistic	Prob.
DLEXPORTSA does not Granger Cause DLPIBSA	17	14.1647	0.0007
DLPIBSA does not Granger Cause DLEXPORTSA		11.8733	0.0014

Prob < 0.05, ipoteza nulă se respinge în ambele cazuri cu un prag de semnificație de 5%.

În urma testului se observă că pe termen lung avem o cauzalitate bidirecțională, atât dispre export înspre pib, cât și dispre pib înspre export. O modificare actuală în export va influența o modificare în pib și viceversa.

Corelograma încrucișată

Date: 01/04/22 Time: 01:13
Sample: 2016Q1 2020Q4
Included observations: 19
Correlations are asymptotically consistent approximations

DLPIBSA,DLEXPORTS...	DLPIBSA,DLEXPORTS...	i	lag	lead
		0	0.9829	0.9829
		1	-0.3355	-0.1209
		2	-0.1307	-0.1361
		3	-0.0457	-0.0765
		4	-0.0320	-0.0397
		5	0.0473	0.0276
		6	-0.0024	0.0117
		7	-0.0366	-0.0098
		8	-0.0587	-0.0646
		9	0.0605	0.0335
		10	0.0110	0.0262
		11	0.0043	0.0100
		12	-0.1597	-0.1470

Cea mai mare valoare se regăsește pentru coeficientul de rang 0. Deci o modificare actuală în PIB se va resimți în aceeași perioadă pentru export și invers, adică relația este contemporană. Legătura dintre cele două variabile este directă și de intensitate mare. Asta înseamnă că atunci când o variabilă crește va aduce cu sine și creșterea celeilalte.

Testul Eagle-Granger

Metoda testează dacă variabilele sunt cointegrate

H₀: Variabilele nu sunt cointegrate.

H₁: Variabilele sunt cointegrate.

Date: 01/04/22 Time: 01:19
Series: DLPIBSA DLEXPORTSA
Sample (adjusted): 2016Q2 2020Q4
Included observations: 19 after adjustments
Null hypothesis: Series are not cointegrated
Cointegrating equation deterministics: C @TREND
Automatic lags specification based on Schwarz criterion (maxlag=3)

Dependent	tau-statistic	Prob.*	z-statistic	Prob.*
DLPIBSA	-5.776087	0.0048	-153.7264	0.0000
DLEXPORTSA	-5.302533	0.0109	-93.23919	0.0000

Probabilitățile sunt mai mici decât pragul de semnificație de 5% pentru ambele teste (tau-statistic și z-statistic), deci ipoteza nulă se acceptă, iar seriile sunt cointegrate.

Previziuni

(Toate previziunile sunt realizate în Excel)

Previziunea naivă

- Presupune că perioada următoare va înregistra valoarea curentă
- Se adaptează repede
- (Ce s-a întâmplat ieri se întâmplă și azi)

Previziunea naivă am realizat-o asupra variabilei dependente, PIB

Valoarea previzionată pentru anul 2011 trimestrul I a fost: 614775

Rezultatele sunt:

MAE: 21792

MSE: 757222993.9

MAPE: 3.762444526

- Valorile previzionate deviază în medie cu 3.76% față de valorile reale.

Metoda valorilor trecute

- Previziunea se calculează ca și medie aritmetică a valorilor trecute
- Este posibilă previziunea doar pe următoarea perioadă

Previziunea pe baza valorilor trecute am realizat-o asupra variabilei dependente, PIB
Valoarea previzionată pentru anul 2011 trimestrul I a fost: 581758.55

Rezultatele sunt:

MAE: 21547.83788

MSE: 709322270.3

MAPE: 3.680105909

- Valorile previzionate deviază în medie cu 3.68% față de valorile reale.

În urma celor două previziune, se poate observa că erorile mai scăzute sunt prezente în metoda valorilor trecută, deci e mai eficientă.

Metode de netezire

Pe baza mediei mobile

- Previzionează valoarea din $t+1$ pe baza celor mai recente k valori.
- k – setat de utilizator în funcție de necesitățile analizei. De ex. dacă valorile cele mai recente sunt considerate relevante, k va fi redus. La fel în cazul schimbărilor de nivel. O valoare mai mare va nivela mai bine fluctuațiile.
- (Excel) Meniul DATA → Data Analysis → Moving Average

Pentru $k=2$

MAPE: 1.881222263

- Valorile previzionate deviază în medie cu 1.88% față de valorile reale.

Pentru $k=4$

MAPE: 2.597336228

- Valorile previzionate deviază în medie cu 2.59% față de valorile reale.

Dintre cele două, metoda cea mai eficientă este aceea cu $k=2$, deoarece are eroarea relativă mai mică.

Netezire exponențială

- Caz special de medie ponderată în care utilizatorul setează doar ponderea celei mai recente valori trecute, restul ponderilor fiind calculate automat.
- α – constanta de netezire cu valori între 0 și 1.
- (Excel) Meniul DATA → Data Analysis → Exponential Smoothing

Pentru $\alpha=0.3$

MAPE: 3.163381943

- Valorile previzionate deviază în medie cu 3.16% față de valorile reale.

Pentru $\alpha=0.7$

MAPE: 3.321662423

- Valorile previzionate deviază în medie cu 3.32% față de valorile reale.

Dintre cele două, metoda cea mai eficientă este aceea cu $\alpha=0.3$, deoarece eroarea relativă e mai mică, astfel se poate afirma faptul că valorile trecute sunt mai importante decât valorile recente.

Previziune pe baza trendului

Pentru trend am folosit Moving Average cu $k=4$, deoarece datele sunt trimestriale. Am obținut valorile pentru fiecare trimestru, iar apoi am facut media acestora și am obținut: 583450.0735

Am extras sezonabilitatea și coeficienții medii ai sezonaliității pentru a estima valorile pentru anii 2016-2020 și pentru a previziona pentru anul 2021.

Interpretare coeficienții medii ai sezonaliității:

- Sub impactul sezonaliității, în trimestrele I PIB-ul a fost în medie 2472.375 unități mai mic.
- Sub impactul sezonaliității, în trimestrele IV, PIB-ul a fost în medie cu 22345.65 unități mai mare.

Pentru 2016-2020:

MAPE: 1.994947333

Datele estimate variază într-un procent de 1.99% în medie față de valorile reale.

Previziunile pentru anul 2021:

- Trimestrul I: 580977.6985
- Trimestrul II: 574682.26
- Trimestrul III: 580642.886
- Trimestrul IV: 605795.7235

Valori dummy + trend(timp)

Datele sunt trimestriale, perioada este egală cu 4 de unde rezultă că vom avea 4-1 valori dummy.

Trimestrul I este bază de comparație.

timpul $T = 1, 2, 3 \dots$

Pentru a forma regresia:

(Excel) Meniul DATA \rightarrow Data Analysis \rightarrow Regression

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	4	5333425925	1.33E+09	2.847385	0.061234872
Residual	15	7024110398	4.68E+08		
Total	19	12357536323			

H_0 : Nu există legătură între variabile. (Variabilele sunt independente).

Sig. $F = 0.06 < 0.10$ (pragul de semnificație), deci ipoteza nulă se respinge. Modelul este semnificativ.

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	561118.6563	12366.11244	45.37551	1.79E-17	534760.9115	587476.401	534760.9115	587476.401
X Variable 1 - D2	-7122.99375	13712.81476	-0.51944	0.611033	-36351.16655	22105.17905	-36351.16655	22105.17905
X Variable 2 - D3	-5400.7875	13792.61843	-0.39157	0.700882	-34799.05779	23997.48279	-34799.05779	23997.48279
X Variable 3 - D4	22146.41875	13924.60822	1.590452	0.132584	-7533.181114	51826.01861	-7533.181114	51826.01861
X Variable 4 - T	1736.59375	855.3818834	2.030197	0.060459	-86.60957661	3559.797077	-86.60957661	3559.797077

H_0 : Coef = 0

P.value > 0.5 pentru toți coeficienții, ipoteza nulă se acceptă și rezultă - coeficienții nu sunt semnificativi. Putem afirma că nu există influențe sezoniere semnificative.

În concluzie, din cele studiate știm că previziunea cea mai eficientă este aceea cu eroarea cea mai mică. Pentru acest set de date, eroarea cea mai mică, de 1.88%, a fost în cazul **previziunii prin metoda de netezire pe baza mediilor mobile**.