**ACADEMIA DE STUDII ECONOMICE BUCUREŞTI**

**FACULTATEA DE CIBERNETICĂ, STATISTICĂ ŞI INFORMATICĂ ECONOMICĂ**

**PROIECT ECONOMETRIE**

**Coordonator științific Studenți**

Conf. Univ. Dr. Adriana Ana Maria DAVIDESCU ENACHE ANDRA-ROXANA

FLIŢĂ-VASILE ADRIAN

GLEJARU COSTIN-COSMIN

**București**

**2022**

Cuprins

INTRODUCERE 4

APLICATIA 1 – MODELUL DE REGRESIE SIMPLĂ5

* 1. Literatura review5

1.2. Metodologia cercetării ……………………………………………………………………………………………………………. 6

1.3. Evoluția economică a fiecărei variabilă …………………………………………………………………………………. 10

1.4. Măsurarea intensității dintre variabile ………………………………………………………………………………….. 13

1.5. Estimare parametri, testare model și interpretare ………………………………………………………………. 14

1.5.1. Indicatori de bonitate ……………………………………………………………………………………………………….. 15

1.5.2. Testarea semnificației parametrilor -Testul T …………………………………………………………………… 15

1.5.3. Testarea validității modelului- Testul Fisher ………………………………………………………………………. 16

1.5.4. Normalitatea distribuției erorilor aleatoare ……………………………………………………………………….. 16

1.5.5. Testarea autocorelării erorilor - Testul Durbin-Watson ……………………………………………………... 18

1.5.6. Homoscedasticitatea erorilor aleatoare …………………………………………………………………………….. 19

1.6.Prognoza ……………………………………………………………………………………………………………………………….. 20

CONCLUZII …………………………………………………………………………………………………………………………………………… 22

APLICAȚIA 2 - MODEL DE REGRESIE MULTIPLĂ 23

2.1.Literatura review …………………………………………………………………………………………………………………… 23

2.2.Metodologia cercetării ………………………………………………………………………………………………………….. 23

2.3.Evoluția economică a fiecărei variabile …………………………………………………………………………………. 27

2.4.Măsurarea intensității dintre variabile ………………………………………………………………………………….. 30

2.5.Estimare parametri, testare model și interpretare ……………………………………………………………….. 31

2.5.1.Indicatori de bonitate ………………………………………………………………………………………………………... 32

2.5.2.Testarea semnificației parametrilor - Testul T ……………………………………………………………………. 32

2.5.3.Testarea validității modelului- Testul Fisher ………………………………………………………………………. 33

2.5.4. Normalitatea distribuției erorilor aleatoare ………………………………………………………………………. 34

2.5.5.Testarea autocorelării erorilor - Testul Durbin-Watson ………………………………………………………. 35

2.5.6.Homoscedasticitatea erorilor aleatoare ……………………………………………………………………………… 36

2.5.7.Necoliniaritatea variabilelor explicative ……………………………………………………………………………… 37

2.6.Prognoză ………………………………………………………………………………………………………………………………. 40

CONCLUZII ……………………………………………………………………………………………………………………………………….. 42

APLICAȚIA 3 - MODEL DE REGRESIE MULTIPLĂ + DUMMY …………………………………………………………………… 43

3.1.Literatura review …………………………………………………………………………………………………………………… 43

3.2.Metodologia cercetării ………………………………………………………………………………………………………….. 43

3.3.Evoluția economică a fiecărei variabile …………………………………………………………………………………. 47

3.4.Măsurarea intensității dintre variabile ………………………………………………………………………………….. 50

3.5.Estimare parametri, testare model și interpretare ……………………………………………………………….. 51

3.5.1.Indicatori de bonitate ………………………………………………………………………………………………………... 52

3.5.2.Testarea semnificației parametrilor - Testul T ……………………………………………………………………. 52

3.5.3.Testarea validității modelului- Testul Fisher ………………………………………………………………………. 53

3.5.4. Normalitatea distribuției erorilor aleatoare ………………………………………………………………………. 54

3.5.5.Testarea autocorelării erorilor - Testul Durbin-Watson ………………………………………………………. 55

3.5.6.Homoscedasticitatea erorilor aleatoare ……………………………………………………………………………… 56

3.5.7.Necoliniaritatea variabilelor explicative ……………………………………………………………………………… 58

3.6.Prognoză ………………………………………………………………………………………………………………………………. 61

CONCLUZII ………………………………………………………………………………………………………………………………………….. 63

APLICAȚIA 4 - MODEL CU DATE DE TIP PANEL ………………………………………………………………………………….. 64

4.1.Literatura review ………………………………………………………………………………………………………………... 64

4.2.Metodologia cercetării ……………………………………………………………………………………………………….. 65

4.3.Normalitatea distribuției erorilor aleatoare …………………………………………………………………………. 74

4.3. Prognoza ……………………………………………………………………………………………………………………………. 75

CONCLUZII ……………………………………………………………………………………………………………………………………….. 76

Bibliografie ……………………………………………………………………………………………………………………………………….. 77

### INTRODUCERE

Urbanizarea se referă la trecerea populației din zonele rurale, la scăderea considerabilă a proporției oamenilor care locuiesc în zonele rurale și la modurile în care situațiile se adaptează la această schimbare. Este în principal procesul prin care orașele se formează și devin mai mari pe măsură ce mai mulți oameni încep să trăiască și să lucreze în zonele centrale.

Urbanizarea este relevantă pentru o serie de discipline precum geografie, sociologie, arhitectură, dar mai ales pentru econometrie. Prin urmare, urbanizarea poate fi cuantificată fie în ceea ce privește nivelul de dezvoltare urbană raportat la populația totală, fie ca ritmul cu care crește proporția urbană a populației.

În cadrul acestui proiect, vom realiza 3 aplicații ce au în vedere diverse modele de regresie. În cazul nostru, vom lua ca variabilă dependentă, indicele de performanță logistică, care este mai cunoscut sub numele de infrastructură, iar rolurile variantelor independente va fi jucat ulterior de PIB(produs intern brut) per capita(PPP) și de înscrierea la învățământul superior(procent din total). Numărul total de date este de 104, reprezentând diferite țări de pe glob.

Prima aplicație va fi reprezentată de un model de regresie simplă, care va prezenta 2 variabile: indicele de performanță logistică jucând rolul de variabilă endogenă, iar indicele de PIB per capita fiind una exogenă.

A doua aplicație va ilustra un model de regresie multiplă, și vom lua din nou variabilele din cadrul primei aplicații, cărora se va adăuga procentul de înscriere la învățământul superior.

A treia aplicație va conține un model pe baza de date de tip panel, în care vom trata cum rata de antreprenoriat influențează diferite variante precum șomajul sau migrarea.

Structura fiecărei aplicații este similară, fiind formată dintr-o parte teoretică – ce prezintă informații dobândite în cadrul unor sesiuni de cercetare; o parte aplicativă – în care vom lua informațiile dobândite și le vom prelucra, cu scopul de a analiza rezolutele; și concluziile, în cadrul cărora vom prezenta un rezumat al celor doi pași realizați anterior.

### APLICAŢIA 1 – MODELUL DE REGRESIE SIMPLĂ

* 1. **Literatura review**

“Urbanizarea a fost mult timp asociată cu dezvoltarea și progresul uman, dar studii recente au arătat că mediile urbane pot duce, de asemenea, la inegalități semnificative la probleme de sănătate. Această lucrare se referă la impactul negativ al urbanizării asupra națiunilor dezvoltate, cât și asupra celor în curs de dezvoltare și asupra populațiilor bogate și sărace din acele națiuni, abordând probleme asociate cu problemele de sănătate publică din zonele urbane. Discuția din această lucrare va fi de interes pentru factorii de decizie. Lucrare susține politici care îmbunătățesc condițiile socio-economice ale populației din mediul urban și promovează o sănătate mai bună a acestora. În plus, această discuție încurajează oamenii și națiunile bogate să devină mai bine informate cu privire la provocările care pot apărea atunci când urbanizarea are loc în regiunile lor fără sprijinul social și infrastructura necesară”. (Md Abdul Kuddus, Elizabeth Tynan, Emma McBryde, în Public Health Review, “Urbanization: a problem for the rich and the poor?”, Abstract, 2020, pagina 1)

Urbanizarea se referă la mișcarea în masă a populațiilor din mediul rural în mediul urban și la schimbările fizice ulterioare în mediul urban. În 2019, conform unui studiu realizat de ONU(Organizația Națiunilor Unite) a estimat ca mai mult din jumătate din populația lumii (4,2 miliarde de oameni), trăiesc acum în zone urbane, iar până în anul 2041, această cifra va ajunge la 6 miliarde de oameni.

### Metodologia cercetării

Ecuația modelului de regresie unifactorial folosit în cadrul acestei aplicații este:

****

**Fig. 1. - Ecuația modelului de regresie unifactorial**

Modelul va fi implementat în EViews, folosind date extrase de pe site-ul THE WORLD BANK, secțiunea DATA( <https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD> ), dar și secțiunea LPI (<https://lpi.worldbank.org/international/global?sort=asc&order=Country>). Aceste date vor avea o periodicitate anuală. Prin urmare, variabilele folosite în cadrul acestei aplicații:

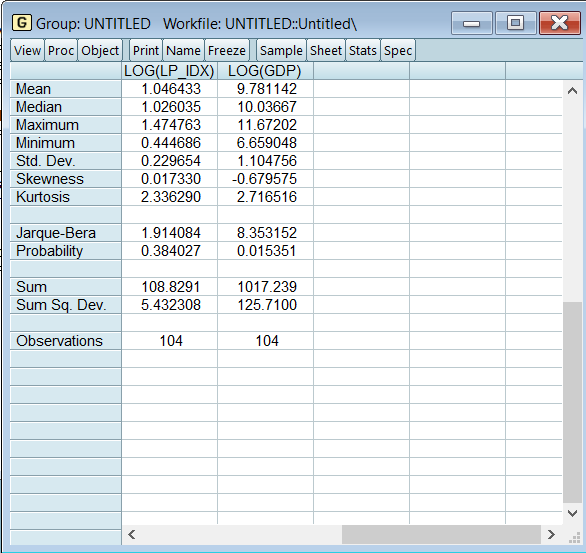
1. Variabila dependenta, indicele de performanță logistic, mai precis infrastructura, având valori între (1,5), unde (1 – foarte puțin dezvoltat, 5 – foarte dezvoltat).
2. Variabila independenta, PIB (produs intern brut) per capita(PPP).

Datele pentru variabilele precizate mai sus se pot vizualiza în tabelul următor:

|  |  |  |
| --- | --- | --- |
| Tari | PIB per capita | Indice de performanță logistic |
| Afghanistan | 2082.635648 | 1.81 |
| Albania | 13554.89895 | 2.29 |
| Algeria | 11909.92982 | 2.42 |
| Argentina | 23293.39863 | 2.77 |
| Armenia | 13020.39632 | 2.48 |
| Australia | 50100.22467 | 3.97 |
| Austria | 57059.53894 | 4.18 |
| Bahrain | 47503.06194 | 2.72 |
| Bangladesh | 4547.480016 | 2.39 |
| Belarus | 19430.27877 | 2.44 |
| Belgium | 52623.55967 | 3.98 |
| Benin | 3236.670417 | 2.5 |
| Bhutan | 11620.9159 | 1.91 |
| Bosnia and Herzegovina | 14916.78501 | 2.42 |
| Brazil | 15020.43822 | 2.93 |
| Brunei Darussalam | 61839.11757 | 2.46 |
| Bulgaria | 22957.40013 | 2.76 |
| Burkina Faso | 2168.662213 | 2.43 |
| Burundi | 779.8081757 | 1.95 |
| Cambodia | 4259.203119 | 2.14 |
| Cameroon | 3800.971267 | 2.57 |
| Canada | 50239.99105 | 3.75 |
| Chile | 24740.49615 | 3.21 |
| China | 15497.35999 | 3.75 |
| Colombia | 14865.89324 | 2.67 |
| Costa Rica | 21022.76519 | 2.49 |
| Croatia | 28960.38457 | 3.01 |
| Cyprus | 40476.38627 | 2.89 |
| Czech Republic | 41134.08988 | 3.46 |
| Denmark | 57456.61183 | 3.96 |
| Ecuador | 11839.34324 | 2.72 |
| Egypt, Arab Rep. | 11639.23909 | 2.82 |
| El Salvador | 8822.550774 | 2.25 |
| Estonia | 36158.81965 | 3.1 |
| Finland | 49706.59709 | 4 |
| France | 46620.68448 | 4 |
| Georgia | 14595.63064 | 2.38 |
| Germany | 55142.31878 | 4.37 |
| Ghana | 5442.872147 | 2.44 |
| Greece | 29652.93088 | 3.17 |
| Guinea | 2559.724766 | 1.56 |
| Honduras | 5816.747324 | 2.47 |
| Hungary | 31862.88251 | 3.27 |
| Iceland | 58317.43934 | 3.19 |
| India | 6675.361294 | 2.91 |
| Indonesia | 11644.75854 | 2.89 |
| Iran, Islamic Rep. | 13795.31747 | 2.77 |
| Ireland | 84665.60623 | 3.29 |
| Israel | 40690.01183 | 3.33 |
| Italy | 43119.35319 | 3.85 |
| Japan | 42386.39981 | 4.25 |
| Jordan | 10253.81745 | 2.72 |
| Kazakhstan | 26157.65965 | 2.55 |
| Korea, Rep. | 42487.04195 | 3.73 |
| Kuwait | 51690.56896 | 3.02 |
| Kyrgyz Republic | 5256.397922 | 2.38 |
| Lao PDR | 7775.131876 | 2.44 |
| Latvia | 30811.07179 | 2.98 |
| Lesotho | 2677.436314 | 1.96 |
| Lithuania | 36365.36049 | 2.73 |
| Luxembourg | 117245.2789 | 3.63 |
| Madagascar | 1630.189357 | 2.16 |
| Malaysia | 28239.50008 | 3.15 |
| Malta | 44482.23968 | 2.9 |
| Mauritania | 5318.085954 | 2.26 |
| Mauritius | 22741.35132 | 2.8 |
| Mexico | 20258.81045 | 2.85 |
| Moldova | 12659.98366 | 2.02 |
| Mongolia | 12317.44092 | 2.1 |
| Montenegro | 21547.28254 | 2.57 |
| Morocco | 7625.735753 | 2.43 |
| Myanmar | 4575.13099 | 1.99 |
| Nepal | 3864.322124 | 2.19 |
| Netherlands | 57901.09701 | 4.21 |
| New Zealand | 43930.88142 | 3.99 |
| Niger | 1229.223782 | 2 |
| North Macedonia | 16725.81473 | 2.47 |
| Norway | 69710.48451 | 3.69 |
| Oman | 29125.45574 | 3.16 |
| Pakistan | 4853.266557 | 2.2 |
| Poland | 31978.53115 | 3.21 |
| Portugal | 34931.7846 | 3.25 |
| Qatar | 93154.2349 | 3.38 |
| Romania | 29248.80932 | 2.91 |
| Russian Federation | 28681.78446 | 2.78 |
| Rwanda | 2138.899699 | 2.76 |
| Saudi Arabia | 48735.4605 | 3.11 |
| Senegal | 3388.946311 | 2.22 |
| Serbia | 17736.0068 | 2.6 |
| Singapore | 100581.158 | 4.06 |
| Slovak Republic | 31530.92023 | 3 |
| Slovenia | 38917.04712 | 3.26 |
| South Africa | 14209.09383 | 3.19 |
| Spain | 40686.98547 | 3.84 |
| Sri Lanka | 13169.0947 | 2.49 |
| Sweden | 53553.31245 | 4.24 |
| Switzerland | 71671.51508 | 4.02 |
| Togo | 2110.82642 | 2.23 |
| Tunisia | 11658.72682 | 2.1 |
| Turkey | 27946.36185 | 3.21 |
| United Kingdom | 47863.2227 | 4.03 |
| United States | 63064.41841 | 4.05 |
| Uruguay | 23588.15139 | 2.43 |
| Uzbekistan | 7252.985197 | 2.57 |

**Tabelul 1. - Datele specifice variabi****lelor în funcție de fiecare țară**

### Evoluția economică a fiecărei variabile



**Fig. 2. - Statistică descriptivă pentru indicele de performanță**

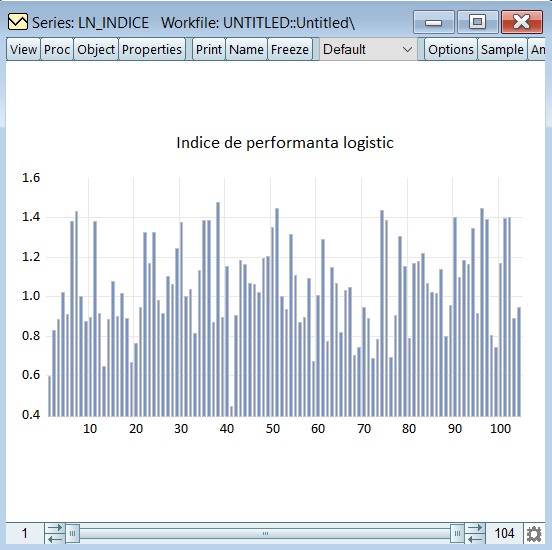
Având in vedere output-ul de mai sus, putem considera:

Variabila independenta, PIB per capita, urmează o distribuție asimetrică negativă (asimetrie de stânga), astfel predomină valorile mici, iar coeficientul de asimetrie(Skewness) are o valoare de aproximativ – 0,6795. Coeficientul de aplatizare (Kurtosis) are o valoarea e 2.7165 și arată că are o distribuție leptocurtică. Minimul seriei este de 6.659, iar maximul seriei este 11.6720.

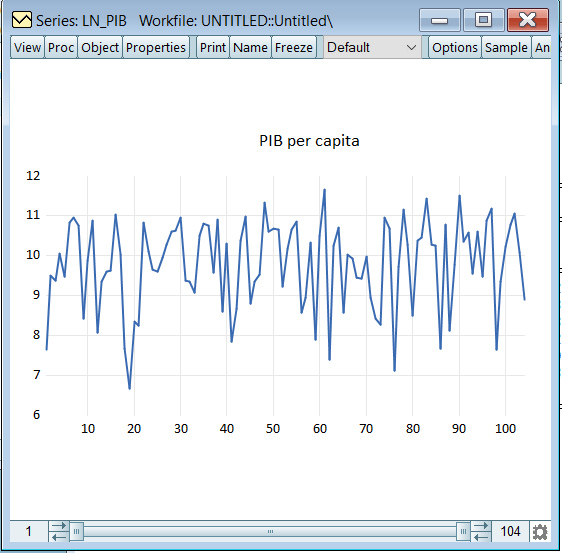
În anul 2018, în medie, PIB-ul per capita atinge o valoare de 9.7811.

Variabila dependentă, indicele de performanta logistic, urmează o distribuție asimetrică pozitivă (asimetrie de dreapta), astfel predomină valorile mici, iar coeficientul de asimetrie(Skewness) are o valoare aproximativ 0,0173. Coeficientul de aplatizare (Kurtosis) are valoarea 2.3362 și arată faptul că are o distribuție leptocurtică. Minimul seriei este de 0,4468, iar maximul este 1.4748.

În anul 2018, în medie, indicele de performanță logistic are o valoare de 1.0464.



**Fig. 3. - Situația indicelui de performanță logistic**

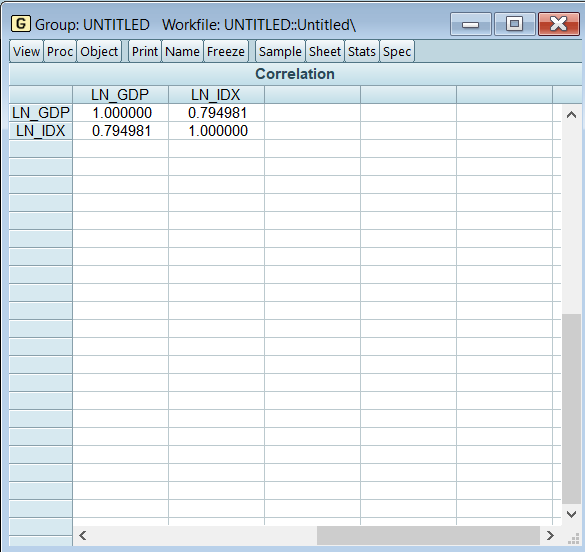


**Fig. 4. - Situația produsului intern brut**

Analizând graficele de mai sus, deducem următoarele:

1. În ceea ce privește indicelui de performanță logistic valorile sunt foarte variate, dar în principal predomină valorile de până la 1.4; astfel reiese ca indicele se învârte în jurul de 1.3 cu mici variații.
2. PIB per capita, de asemenea, are valori foarte variat, cel mai mic nivel ar fi sub 7, într-o singură țară; pe când cea mai mare se apropie de 12 într-o singura țară.
   1. **Măsurarea intensității dintre variabile**

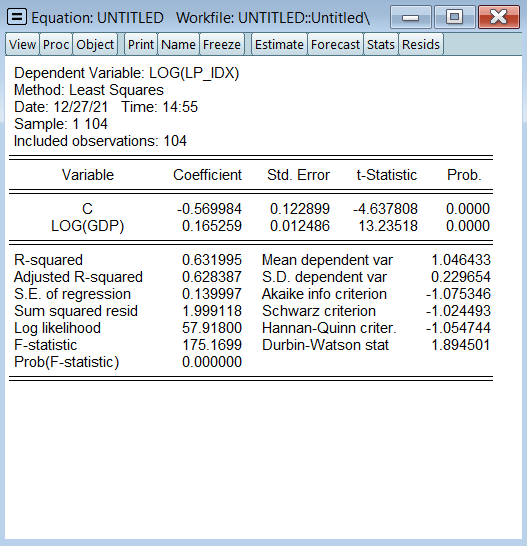
Înainte de a aplica metoda celor mai mici pătrate, am logaritm datele pentru a nu întâmpina probleme, în ceea ce privește diverși coeficienți. De asemenea, am calculat intensitatea și tipul legăturii dintre variabila independentă(PIB per capita), și variabila dependentă(indicele de performanță logistic). Prin urmare, am calculat coeficientul de corelație Pearson, folosindu-ne de EViews, iar output-ul se poate observa mai jos:



**Fig. 5. - Coeficient de corelație Pearson**

În figura 5, se observă că între cele două variabile avem o corelație de 0.794, o corelație puternică și pozitivă. Astfel, o creștere a nivelului indicelui de performanță logistic va duce la o creștere a PIB per capita.

### Estimare parametri, testare model și interpretare



**Fig. 6. - Modelul de regresie unifactorial**

Ecuația generală a modelului de regresie este următoarea:

Y = + \* X

Particularizăm această ecuație prin faptul că preluam din figura 6 coeficienții = 0.16525 și = -0.569984. Ecuația pentru modelul nostru devenind astfel:

Y = 0.16525 - 0.569984 \* X

Notăm Indicele de performanță logistic cu IPL și PIB per capita cu PPP. Astfel, ecuația rescrisă devine:

IPL = 0.16525 - 0.569984 \* PPP

### Indicatori de bonitate

1. **Coeficientul de determinație**

Coeficientul de corelație R Squared indică cât din variabila dependentă Indicele de performanță logistic este influențată de variabila independentă PIB per capita. Așadar, R Squared = 0.576453 arată că PIB per capita explică 57.6453% din modificarea indicelui de performanță logistică, restul de 42.3547% fiind alte variabile care nu sunt incluse în model.

1. **Coeficientul de determinație ajustat**

Adjusted R Squared are aceeași semnificație ca R Squared. Astfel, PIB per capita reprezintă 57.2300% din modificarea speranței de viață.

1. **Eroarea standard**

Standard Error of regression reprezintă cât se abate de la medie valorile introduse în model. În modelul nostru avem o eroare standard de doar 0.012486.

### Testarea semnificației parametrilor -Testul T

Pentru a putea testa validitatea modelului statistic utilizat pentru analiza acestui topic, și pentru a dovedi semnificația parametrului β, vom folosi testul T pentru un model bilateral și formulăm următoarele ipoteze:

: (parametrul nu este semnificativ din punct de vedere statistic, ceea ce rezultă în invaliditatea modelului);

: (parametrul este semnificativ din punct de vedere statistic, ceea ce rezultă în validitatea modelului); iar toate datele necesare sunt luate pentru un interval de încredere de 95%, rezultând că

Conform analizei pe care am realizat-o în punctele anterioare, știm că t calculat are valoarea aproximativ egală cu 13.23518. Următoarea variabilă de care avem nevoie este t critic, corespunzător lui și numărului de observații, care este egal cu 104. Pentru a putea respinge ipoteza care infirmă validitatea aplicației, trebuie ca |Întrucât || are valoarea egală cu 13.23518, iar are valoarea 1.66 și că 13.23518> 1.66, putem respinge ipoteza nulă și să acceptăm ipoteza alternativă, ceea ce înseamnă că modelul nostru este semnificativ din punct de vedere statistic.

### Testarea validității modelului- Testul Fisher

Pentru a testa validitatea modelului utilizat în cadrul acestui proiect, vom formula 2 ipoteze:

: modelul nu este valid din punct de vedere statistic,

: modelul este valid din punct de vedere statistic.

Pentru a reuși să analizăm aceste ipoteze, ne vom folosi de testul Fisher, pentru o probabilitate de 95%, cu având 102 de grade de libertate (102 de observații, din care scădem 2 pentru eliminarea erorilor), pentru un model de regresie simplă, cu o singură variabilă independentă.

Conform rezultatelor analizei realizate în Fig. 6, aflăm că are valoarea aproximativ egală cu 175.1699. Valoarea lui pentru o probabilitate de 95%, un model de regresie simplă, cu 102 de grade de libertate, are valoarea 3.936.

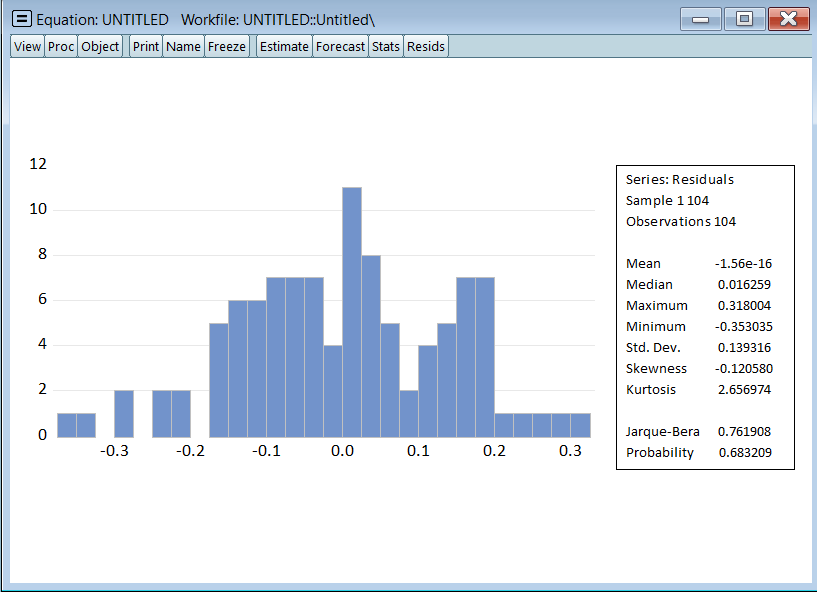
În cazul în care > , vom putea respinge ipoteza nulă, ceea ce înseamnă că modelul ales este valid. În urma comparării acestor valori (175.1699> 3.936), rezultă că respingerea lui este posibilă, și putem astfel accepta ipoteza , ceea ce înseamnă că modelul nostru este valid din punct de vedere statistic. O altă variabilă care confirmă validitatea acestui model este Prob(F-statistic), care are valoarea 0, care este mai mică decât pragul de 5% ales de noi.

### Normalitatea distribuției erorilor aleatoare

În această etapă vom verifica ipotezele de normalitate ale erorilor aleatoare, aplicând testul Jarque–Bera. Ipotezele testului sunt următoarele:

: erorile aleatoare urmează o distribuție normală

: erorile aleatoare nu urmează o distribuție normală



**Fig. 7. - Rezultatul testului Jarque-Bera**

Din figura de mai sus observăm că valoarea testului Jarque-Bera este de 0.761908, cu o probabilitate de 0.683209. Cum modelul are un nivel de încredere de 95%, probabilitatea testului trebuie să fie mai mare de 5%(0.05). Astfel 68.3209% este mai mare de pragul de 5%, așadar acceptăm ipoteza nulă, erorile aleatoare urmează o distribuție normală.

De asemenea, valoarea tabelată a testului Jarque-Bera este 3.841. Având în vedere că JB statistic are valoarea de 0.761908 care este mai mică decât valoarea tabelată, putem concluziona faptul că erorile sunt normal distribuite.

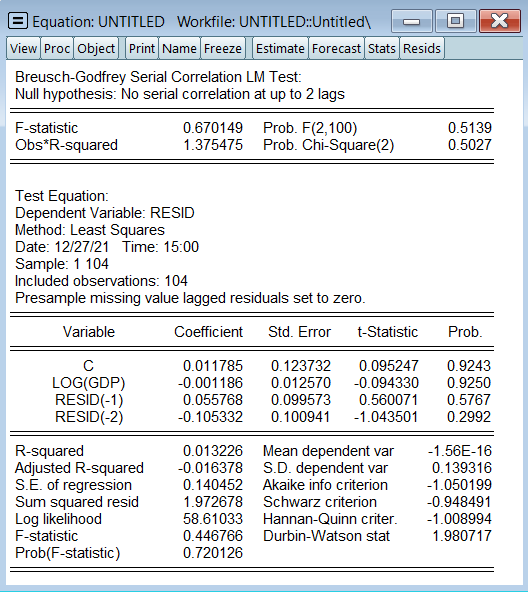
### Testarea autocorelării erorilor - Testul Durbin-Watson

În cadrul acestui pas, vom analiza condiția de autocorelare a erorilor modelului ales de noi pentru realizarea acestei aplicații. Pentru acest lucru, ne vom folosi de testul Durbin-Watson și de ipotezele corespunzătoare acestuia:

: erorile nu sunt autocorelate,

: erorile sunt autocorelate; unde vom lua probabilitatea implicită de 95%, luând valoarea de 0.05.

Conform figurii 5, rezultatul este reprezentat de indicele calculat luând o valoare egală cu 1.894501. Pentru a putea verifica dacă erorile sunt corelate, vom avea nevoie de figura numărul 8.



**Fig. 8. - Rezultatul testului Breush-Godfrey privind corelația**

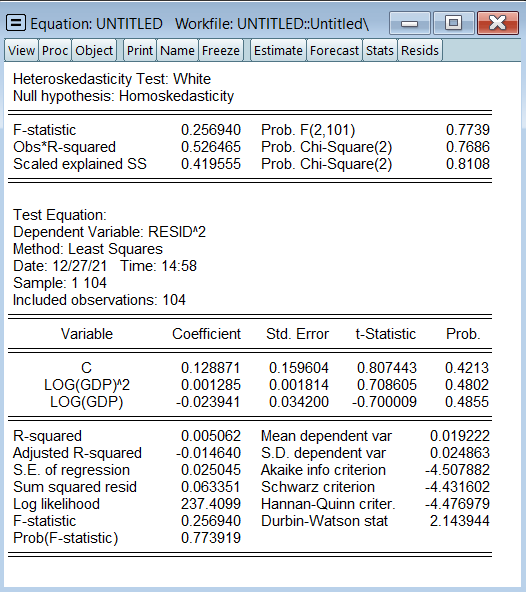
În figura 8, observăm că CHI-Square(2) are valoarea egala cu 0.5027. Pentru a putea demonstra ca erorile nu sunt autocorelate, rezultatul CHI-Square(2) trebuie sa fie mai mare sau egal cu 0.05, ceea ce în modelul de față este adevărat.

De asemenea, cunoaștem că valorile testului Durbin-Watson se situează în intervalul 0-4. O valoare de 1.908, situată în intervalul (0,2) ne indică faptul că avem parte de autocorelare pozitivă.

### Homoscedasticitatea erorilor aleatoare

: erorile aleatoare sunt homoscedastice

: erorile aleatoare sunt heteroscedastice



**Fig. 9. - Rezultatul testului White**

Prin urmare, vom compara:

Obs\*R-Squared = 0.5264 < Chi Squared critic = 3.841

Prob(F-statistic) > .

F critic = 3.936 (valoare estimativă pentru 100 de observații)> F statistic = 0.256940

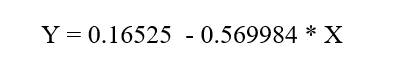
Obs \* R – Squared = 0.526465

Prob(Chi-Square(2)) = 0.7686 > 0.05

De aici, reiese faptul că respingem ipoteza alternativă și acceptăm ipoteza nulă, prin urmare erorile noastre homoschedastice.

### Prognoza

În modelul ales de noi pentru prima aplicație, pe termen scurt putem observa că rata de performanță logistică – infrastructura - contribuie în mod negativ la bunăstarea oamenilor, deoarece conduce la scăderea PIB – ului pe capita, cum se poate vedea și în ecuația următoare.



**Fig. 10. - Ecuația modelului de regresie unifactorial**

Pe termen lung însă, indicele de performanță logistic poate avea efect pozitiv în ceea ce privește PIB-ul per capita, deoarece prin dezvoltarea infrastructurii pot fi generate locuri noi de muncă, ceea ce va genera în timp, creșterea PIB – ului.

### 

**Fig. 11. Prognoza**

### Observații:

1. Bias-ul unui estimator reprezintă diferența dintre valoarea așteptată a estimatorului și valoarea adevărată a parametrului prognozat. Un estimator cu bias zero sau fără bias este un aspect pozitiv, acesta indică o prognoză sănătoasă. În cazul de față, se poate observa că ceea ce am prognozat este foarte corect.
2. Covarianța este măsura de variație comună a două variabile aleatorii. În cazul de față avem o covarianță de 88.57%.
3. De asemenea, theil inequality are o valoare de 0.0649. Cu cât este mai apropiată de zero această valoare, cu atât prognoza este aproape perfectă, iar în cazul nostru prognoza este aproape perfectă.

### Concluzii

Pentru această aplicație, am pornit de la ipoteza “Cum poate influența indicele de performanță logistic PIB-ul per capita?”. Pentru a putea testa această ipoteză am realizat un model de regresie folosind datele privind indicele de performanță logistic și PIB per capita pentru 104 diferita tari, de pe suprafața pământului, reprezentând astfel peste jumătate din numărul total de țări.

Am aplicat testul T pentru verificarea semnificației parametrilor(aceștia au ieșit semnificativi), iar testul F a ilustrat validitatea modelului. Modelul a respectat toate cerințele (homoschedasticitate și neautocorelarea reziduurilor, precum și distribuția normală a reziduurilor).

Având in vedere cele enunțate mai sus putem garanta cu o probabilitate de 95% ca evoluția indicelui influențează într-o maniera pozitiva PIB-ul per capita.

### APLICAȚIA 2 - MODEL DE REGRESIE MULTIPLĂ

* 1. **Literatura review**

“Educația este larg acceptată ca o resursă fundamentală, atât pentru indivizi, cât și pentru societate. Într-adevăr, în majoritatea țărilor, în prezent, educația de bază este percepută nu doar ca un drept, ci și ca o datorie – se așteaptă de obicei guvernele să asigure accesul la educația de bază în timp ce cetățenii sunt adesea obligați prin lege să atingă educația până la un anumit nivel de bază.”(Max Roser, Esteban Ortiz-Ospina în OurWorldInData Global Education, 2020,p3).

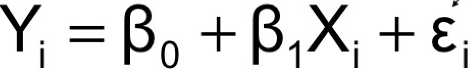
Educația este o investiție valoroasă, atât pentru individ, cât și pentru colectiv. Cea mai obișnuită modalitate de a măsura profiturile educației este de a studia modul în care îmbunătățesc rezultatele individuale pe piața muncii – de obicei, încercând să măsoare efectul educației asupra salariilor.

Pe de altă parte, urbanizarea este una dintre cele mai importante metode de dezvoltare economică a unui țări, la care contribuie cu un rol foarte important și educația. Educația și urbanizarea conduc la dezvoltarea infrastructurii și a PIB per capita.

|  |  |  |  |
| --- | --- | --- | --- |
| Tari | PIB per capita | Studii superioare | Indice de dezvoltare a infrastructurii |
| Afghanistan | 2082.635648 | 9.686420441 | 1.81 |
| Albania | 13554.89895 | 54.96133041 | 2.29 |
| Algeria | 11909.92982 | 51.36566925 | 2.42 |
| Argentina | 23293.39863 | 91.6006012 | 2.77 |
| Armenia | 13020.39632 | 54.57048035 | 2.48 |
| Australia | 50100.22467 | 107.8080521 | 3.97 |
| Austria | 57059.53894 | 86.68834686 | 4.18 |
| Bahrain | 47503.06194 | 50.48189926 | 2.72 |
| Bangladesh | 4547.480016 | 20.5650692 | 2.39 |
| Belarus | 19430.27877 | 87.42909241 | 2.44 |
| Belgium | 52623.55967 | 78.90116882 | 3.98 |
| Benin | 3236.670417 | 12.51698971 | 2.5 |
| Bhutan | 11620.9159 | 15.55451012 | 1.91 |
| Bosnia and Herzegovina | 14916.78501 | 43.4591217 | 2.42 |
| Brazil | 15020.43822 | 53.25735092 | 2.93 |
| Brunei Darussalam | 61839.11757 | 31.21179962 | 2.46 |
| Bulgaria | 22957.40013 | 71.52152252 | 2.76 |
| Burkina Faso | 2168.662213 | 6.501870155 | 2.43 |
| Burundi | 779.8081757 | 4.057660103 | 1.95 |
| Cambodia | 4259.203119 | 13.68776035 | 2.14 |
| Cameroon | 3800.971267 | 14.26727009 | 2.57 |
| Canada | 50239.99105 | 70.11302185 | 3.75 |
| Chile | 24740.49615 | 90.8963089 | 3.21 |
| China | 15497.35999 | 50.60443878 | 3.75 |
| Colombia | 14865.89324 | 55.32749176 | 2.67 |
| Costa Rica | 21022.76519 | 55.31010056 | 2.49 |
| Croatia | 28960.38457 | 67.65422821 | 3.01 |
| Cyprus | 40476.38627 | 81.33708191 | 2.89 |
| Czech Republic | 41134.08988 | 63.76876831 | 3.46 |
| Denmark | 57456.61183 | 81.18334961 | 3.96 |
| Ecuador | 11839.34324 | 47.59777069 | 2.72 |
| Egypt, Arab Rep. | 11639.23909 | 38.90494919 | 2.82 |
| El Salvador | 8822.550774 | 29.37192917 | 2.25 |
| Estonia | 36158.81965 | 70.36543274 | 3.1 |
| Finland | 49706.59709 | 90.2617569 | 4 |
| France | 46620.68448 | 67.54351807 | 4 |
| Georgia | 14595.63064 | 60.33444977 | 2.38 |
| Germany | 55142.31878 | 70.34264374 | 4.37 |
| Ghana | 5442.872147 | 15.6917696 | 2.44 |
| Greece | 29652.93088 | 142.8520355 | 3.17 |
| Guinea | 2559.724766 | 7.921060085 | 1.56 |
| Honduras | 5816.747324 | 26.16422081 | 2.47 |
| Hungary | 31862.88251 | 50.30653 | 3.27 |
| Iceland | 58317.43934 | 73.1032486 | 3.19 |
| India | 6675.361294 | 28.06055069 | 2.91 |
| Indonesia | 11644.75854 | 36.31103897 | 2.89 |
| Iran, Islamic Rep. | 13795.31747 | 62.78507996 | 2.77 |
| Ireland | 84665.60623 | 77.28288269 | 3.29 |
| Israel | 40690.01183 | 61.47687149 | 3.33 |
| Italy | 43119.35319 | 64.29161835 | 3.85 |
| Japan | 42386.39981 | 64.10384369 | 4.25 |
| Jordan | 10253.81745 | 34.41532135 | 2.72 |
| Kazakhstan | 26157.65965 | 53.98825073 | 2.55 |
| Korea, Rep. | 42487.04195 | 95.864151 | 3.73 |
| Kuwait | 51690.56896 | 54.36413956 | 3.02 |
| Kyrgyz Republic | 5256.397922 | 41.26702118 | 2.38 |
| Lao PDR | 7775.131876 | 14.96675968 | 2.44 |
| Latvia | 30811.07179 | 93.02448273 | 2.98 |
| Lesotho | 2677.436314 | 10.19626045 | 1.96 |
| Lithuania | 36365.36049 | 73.7303009 | 2.73 |
| Luxembourg | 117245.2789 | 18.59046936 | 3.63 |
| Madagascar | 1630.189357 | 5.350490093 | 2.16 |
| Malaysia | 28239.50008 | 45.1253891 | 3.15 |
| Malta | 44482.23968 | 59.25175858 | 2.9 |
| Mauritania | 5318.085954 | 5.002089977 | 2.26 |
| Mauritius | 22741.35132 | 41.5999794 | 2.8 |
| Mexico | 20258.81045 | 41.52280045 | 2.85 |
| Moldova | 12659.98366 | 55.77304077 | 2.02 |
| Mongolia | 12317.44092 | 65.59544373 | 2.1 |
| Montenegro | 21547.28254 | 56.08097076 | 2.57 |
| Morocco | 7625.735753 | 35.93592834 | 2.43 |
| Myanmar | 4575.13099 | 18.81594086 | 1.99 |
| Nepal | 3864.322124 | 12.41086006 | 2.19 |
| Netherlands | 57901.09701 | 87.09784698 | 4.21 |
| New Zealand | 43930.88142 | 82.98303223 | 3.99 |
| Niger | 1229.223782 | 4.414350033 | 2 |
| North Macedonia | 16725.81473 | 43.11650085 | 2.47 |
| Norway | 69710.48451 | 83.01725769 | 3.69 |
| Oman | 29125.45574 | 38.03620911 | 3.16 |
| Pakistan | 4853.266557 | 8.9584198 | 2.2 |
| Poland | 31978.53115 | 68.62010193 | 3.21 |
| Portugal | 34931.7846 | 65.66265869 | 3.25 |
| Qatar | 93154.2349 | 17.86992073 | 3.38 |
| Romania | 29248.80932 | 51.01253891 | 2.91 |
| Russian Federation | 28681.78446 | 84.5842514 | 2.78 |
| Rwanda | 2138.899699 | 6.725719929 | 2.76 |
| Saudi Arabia | 48735.4605 | 68.03988647 | 3.11 |
| Senegal | 3388.946311 | 12.76259041 | 2.22 |
| Serbia | 17736.0068 | 67.15866089 | 2.6 |
| Singapore | 100581.158 | 88.88645172 | 4.06 |
| Slovak Republic | 31530.92023 | 45.36722946 | 3 |
| Slovenia | 38917.04712 | 77.11312103 | 3.26 |
| South Africa | 14209.09383 | 23.80195999 | 3.19 |
| Spain | 40686.98547 | 91.11257172 | 3.84 |
| Sri Lanka | 13169.0947 | 19.63023949 | 2.49 |
| Sweden | 53553.31245 | 72.46116638 | 4.24 |
| Switzerland | 71671.51508 | 61.37744904 | 4.02 |
| Togo | 2110.82642 | 13.51175022 | 2.23 |
| Tunisia | 11658.72682 | 31.74651909 | 2.1 |
| Turkey | 27946.36185 | 113.2170715 | 3.21 |
| United Kingdom | 47863.2227 | 61.38288879 | 4.03 |
| United States | 63064.41841 | 88.29917908 | 4.05 |
| Uruguay | 23588.15139 | 97.44525146 | 2.43 |
| Uzbekistan | 7252.985197 | 10.07635021 | 2.57 |

### Metodologia cercetării

Ecuația modelului de regresie multifuncțional este:



**Fig. 12. - Ecuația modelului de regresie multifactorial**

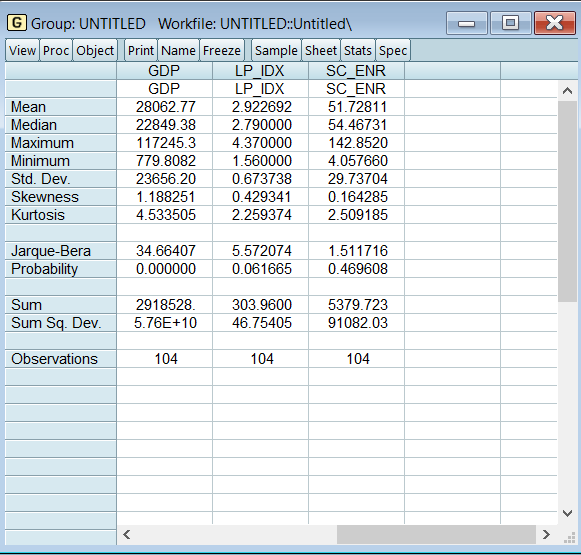
**Tabelul 2. Datele specifice variabilelor în funcție de fiecare țară**

Modelul va fi implementat în EViews, folosind date extrase de pe site-ul THE WORLD BANK, secțiunea DATA( <https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD> ), ( <https://data.worldbank.org/indicator/SE.TER.ENRR> ), dar și secțiunea LPI (<https://lpi.worldbank.org/international/global?sort=asc&order=Country>). Aceste date vor avea o periodicitate anuală. Prin urmare, variabilele folosite în cadrul acestei aplicații:

1. Variabila dependenta, indicele de performanță logistic, mai precis infrastructura, având valori între (1,5), unde (1 – foarte puțin dezvoltat, 5 – foarte dezvoltat).
2. Variabila independenta, PIB (produs intern brut) per capita(PPP).
3. Înscriere la învățământ superior, raport din total.

Datele pentru variabilele precizate mai sus se pot vizualiza în tabelul următor:

### Evoluția economică a fiecărei variabile

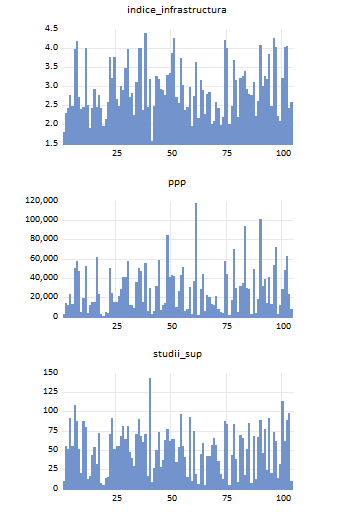


**Fig. 13. - Statistica descriptivă pentru indicele de performanță logistică,**

**PIB per capita, înrolarea la învățământul superior.**

Având în vedere output-ul de mai sus, concluziile sunt:

1. Variabila independentă PIB per capita formează o distribuție asimetrică pozitivă (asimetrie de dreapta), astfel predomină valorile mici, iar coeficientul de asimetrie( Skewness ) este 1.1882. Coeficientul de aplatizare ( Kurtosis) are valoarea de 4.5335 și arată faptul că are o distribuție leptocurtică. Minimul seriei este 779, iar maximul 117245.
2. În anul 2018, în medie, PIB per capita atinge o valoare de 28062.77 .
3. Variabila dependentă indicele de performanță logistic urmează o distribuție asimetrică pozitivă, astfel predomină valorile mici, iar coeficientul de asimetrie (Skewness) are o valoare de aproximativ 0.4293. Coeficientul de aplatizare (Kurtosis) are valoarea 2.25 și arată faptul că are o distribuție leptocurtică. Minimul seriei este 1.56, iar maximul e 4.37.
4. În anul 2018, indicele de performanță logistic a fost, în medie, 2.92.
5. Variabila independentă procentul de înscriere la învățământul superior urmează o distribuție asimetrică pozitivă, astfel predomină valorile mici, iar coeficientul de asimetrie (Skewness) are o valoare de aproximativ 0.1642. Coeficientul de aplatizare (Kurtosis) are valoarea 2.50 și arată faptul că are o distribuție leptocurtică. Minimul seriei este 4.06, iar maximul e 142.85.
6. În anul 2018, procentul de înscriere la învățământul superior mediu a fost de 51.71.



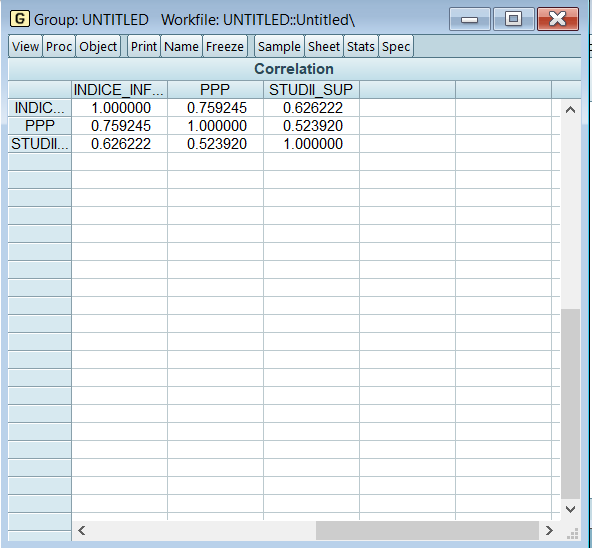
### Fig. 14. - Graficele bară ale variabilelor

### Analizând graficele de mai sus, deducem faptul că:

1. În ceea ce privește indicele de performanță logistic, valorile sunt foarte variate, dar în principal predomină valorile până în 4.37, astfel reiese că indicele de performanță logistic se învârte în jurul valorii de 3.00.
2. În ceea ce privește PIB per capita, valorile sunt de asemenea foarte variate, atingând un maxim de 117245, iar în general se învârte în jurul valorii de 20.000.
3. În graficul ratei de înscriere la învățământul superior, putem observa că există o valoare de 142, dar în general se învârte în jurul valorii de 50.

### Măsurarea intensității dintre variabile

Înainte de a aplica metoda celor mai mici pătrate, am calculat intensitatea și tipul legăturii dintre variabilele independente (rata de înscriere la învățământul superior și PIB per capita) și variabila dependentă (indicele de performanță logistic). Prin urmare, am calculat coeficientul de corelație Pearson folosindu-ne de EViews, iar output-ul generat se poate observa mai jos:

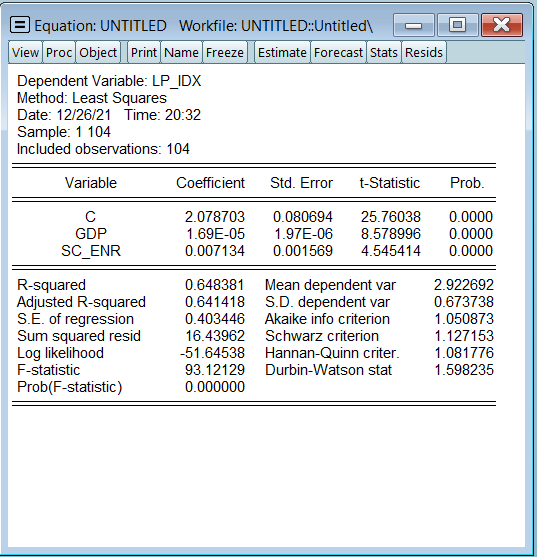


**Fig. 15. - Coeficientul de corelație Pearson**

Din figura de mai sus observăm către variabila independentă PIB per capita și indicele de performanță logistic avem un coeficient de corelație puternic de 0.7592, dar între ratei de înscriere la învățământul superior și indicele de performanță logistic se observă o relație puternică de 0.6262. Între cele doua variabile independente avem o relație, de asemenea, tot pozitivă de 0.5239.

### Estimare parametri, testare model și interpretare

Am aplicat metoda celor mai mici pătrate în EViews și am realizat o estimare a parametrilor modelului econometric de regresie multifactorială, output care se poate observa în figura care urmează.



**Fig. 16. - Model de regresie multifactorial**

Ecuația generală a modelului de regresie este următoarea:

Y=+\*X1+ \*X2

Particularizăm această ecuație prin faptul că preluăm din figură coeficienții = 2.0787, = 0.0000169 și =0.0071. Ecuația pentru modelul nostru devenind astfel:

Y=2.0787+0.0000169\*(X1) + 0.0071\*(X2)

### . Indicatori de bonitate

**1.Coeficientul de determinație**

Coeficientul de corelație R Squared indică cât din variabila dependentă indicele de performanță logistică este influențată de variabilele independente PIB per capita și rata de înscriere la învățământul superior. Așadar, R Squared = 0.648381 arată că PIB per capita împreună cu rata de înscriere la învățământul superior explică 64.8381% din modificarea indicelui de performanță logistică, restul de 36.1619% fiind alte variabile care nu sunt incluse în model.

**2.Coeficientul de determinație ajustat**

Adjusted R Squared are aceeași semnificație ca R Squared. Astfel, PIB per capita și rata de înscriere la învățământul superior reprezintă 64.1418% din modificarea indicelui de performanță logistică.

**3.Eroarea standard**

Standard Error of regression reprezintă cu cât se abat de la medie valorile introduse în model. În modelul nostru avem o eroare standard de 0.001569 pentru rata de înscriere la învățământul superior și o eroare standard de 0.00000167 pentru PIB per capita.

### Testarea semnificației parametrilor - Testul T

Pentru a putea testa validitatea modelului statistic utilizat pentru analiza acestui topic, și pentru a dovedi semnificația parametrului β, vom folosi testul T pentru un model bilateral și formulăm următoarele ipoteze:

= 0; (parametrul nu este semnificativ din punct de vedere statistic, ceea ce rezultă în invaliditatea modelului);

: 0; (parametrul este semnificativ din punct de vedere statistic, ceea ce rezultă în validitatea modelului); iar toate datele necesare sunt luate pentru un interval de încredere de 95%, rezultând că

Conform calculelor obținute în cadrul figuri 15(figura de sus), știm că t calculat pentru PIB per capita este aproximativ egal cu 8.5788, rata de înscriere la învățământul superior este aproximativ egal cu valoarea 4.5454, valoarea liberă are o valoarea a t-ului 25.7603. Următoarea variabilă de care avem nevoie este t critic, corespunzător lui și numărului de observații, care este egal cu 104. Acest lucru rezultă că t critic are valoarea 1.66.

Pentru a putea respinge ipoteza care infirmă validitatea aplicației, trebuia ca |. În urma calculelor, observăm că t0calc= 8.5788, t1calc= 4.5454, t2calc = 25.7603, pe când

### = 1.66. Din toate acestea, conform valorii de adevăr a acestor inegalități, putem respinge ipoteza nulă și să acceptăm ipoteza alternativă, ceea ce înseamnă că modelul nostru este semnificativ din punct de vedere statistic.

### Testarea validității modelului- Testul Fisher

Pentru a testa validitatea modelului utilizat în cadrul acestui proiect, vom formula 2 ipoteze:

: modelul nu este valid din punct de vedere statistic,

: modelul este valid din punct de vedere statistic.

Pentru a reuși să analizăm aceste ipoteze, ne vom folosi de testul Fisher, pentru o probabilitate de 95%, cu având 102 de grade de libertate (102 de observații, din care scădem 2 pentru eliminarea erorilor), pentru un model de regresie simplă, cu o singură variabilă independentă.

Conform rezultatelor analizei realizate în Fig. 6, aflăm că are valoarea aproximativ egală cu 93.1212. Valoarea lui pentru o probabilitate de 95%, un model de regresie simplă, cu 102 de grade de libertate, are valoarea 3.09 .

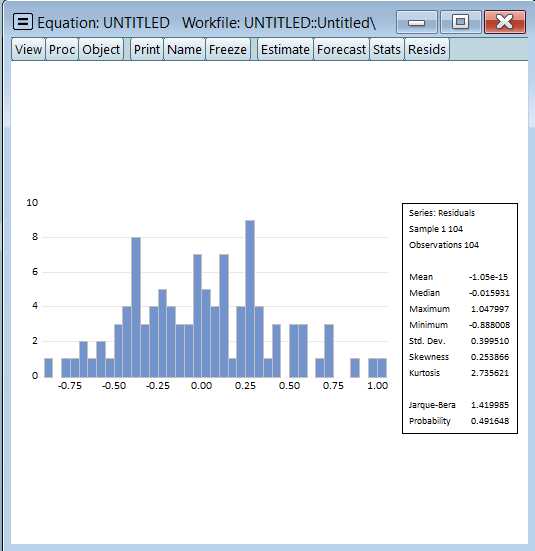
În cazul în care > , vom putea respinge ipoteza nulă, ceea ce înseamnă că modelul ales este valid. În urma comparării acestor valori (93.1212> 3.09), rezultă că respingerea lui este posibilă, și putem astfel accepta ipoteza , ceea ce înseamnă că modelul nostru este valid din punct de vedere statistic. O altă variabilă care confirmă validitatea acestui model este Prob(F-statistic), care are valoarea 0, care este mai mică decât pragul de 5% ales de noi.

### 2.5.4. Normalitatea distribuției erorilor aleatoare

În această etapă vom verifica ipotezele de normalitate ale erorilor aleatoare, aplicând testul Jarque – Bera. Ipotezele testului sunt următoarele:

: erorile aleatoare urmează o distribuție normală

: erorile aleatoare nu urmează o distribuție normală



**Fig. 17. - Rezultatul testului Jarque-Bera**

Din figura anterioara, observăm că valoarea testului Jarque-Bara este 1.4199, cu o probabilitate de 0.4916. Cum modelul are un nivel de încredere de 95%, probabilitatea testului trebuie să fie mai mare de 5%(0.05). Astfel, 49.16% este mai mare decât pragul de 0.05, acceptăm ipoteza nulă, erorile aleatoare urmează o distribuție normală.

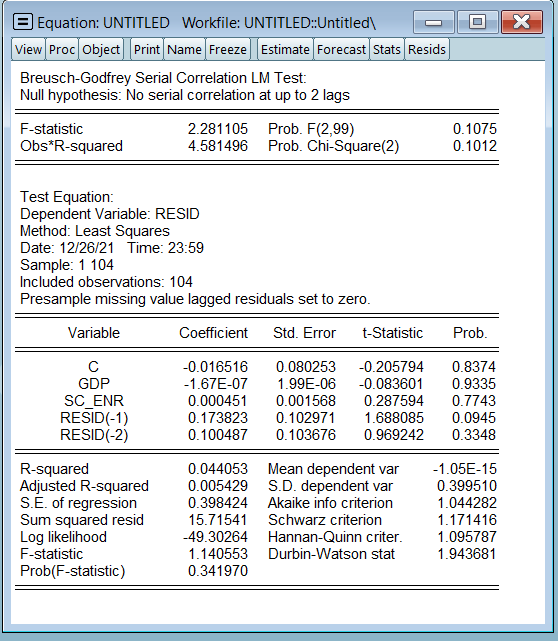
Cum valoarea tabelată a testului Jarque-Bara este de 5.991, iar Jarque-Bera statistic are valoarea 1.4199, putem concluziona că erorile sunt normal distribuite

### 

### Testarea autocorelării erorilor - Testul Durbin-Watson

În cadrul acestui pas, vom analiza condiția de autocorelare a erorilor modelului ales de noi pentru realizarea acestei aplicații. Pentru acest lucru, ne vom folosi de testul Durbin-Watson și de ipotezele corespunzătoare acestuia:

: erorile nu sunt autocorelate,

 : erorile sunt autocorelate; unde vom lua probabilitatea implicită de 95%, luând valoarea de 0.05.

**Fig. 18. – Rezultatul testului Breush-Godfrey privind corelația**

Conform figurii 16, rezultatul este reprezentat de indicele calculat luând o valoare egală cu 1.5982. Pentru a putea verifica dacă erorile sunt corelate, vom avea nevoie de figura numărul prezentată mai sus.

În figura anterioară, observăm că CHI-Square(2) are valoarea egala cu 0.1012. Pentru a putea demonstra ca erorile nu sunt autocorelate, rezultatul CHI-Square(2) trebuie sa fie mai mare sau egal cu 0.05, ceea ce în modelul de față este adevărat.

De asemenea, cunoaștem că valorile testului Durbin-Watson se situează în intervalul 0-4. O valoare de 1.9436, situată în intervalul (0,2) ne indică faptul că avem parte de autocorelare pozitivă, și că reziduurile sunt independente.

### Homoscedasticitatea erorilor aleatoare

: erorile aleatoare sunt homoschedastice

: erorile aleatoare sunt heteroscedastice

Prin urmare, vom compara:

Obs\*R-Squared = 9.8054 Prob(F-statistic) > .

Prob(Chi-Square(2)) = 0.0809 > 0.05

De aici, reiese faptul că respingem ipoteza alternativă și acceptăm ipoteza nulă, prin urmare erorile noastre homoschedastice.

### 

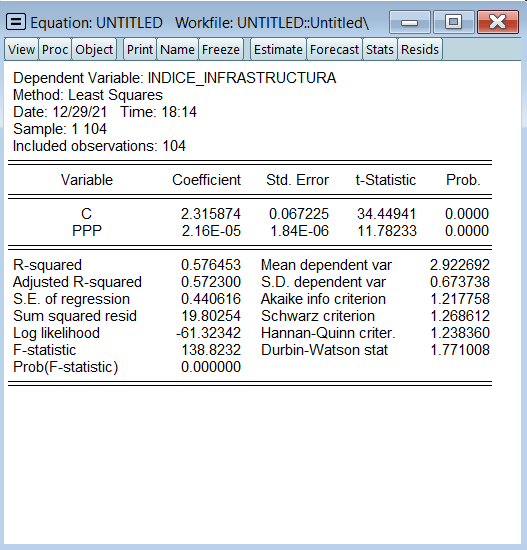
**Fig. 19. – Rezultatul testului White**

#### **Necoliniaritatea variabilelor explicative**

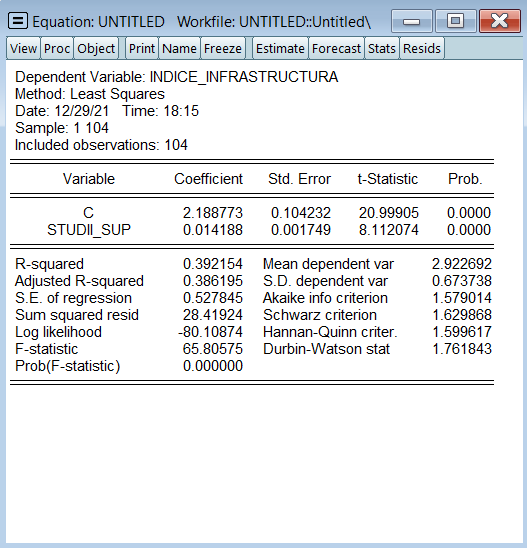
Pentru a verifica necoliniaritatea variabilelor explicative, am creat două modele de regresie unifactorială între indicele de performanță logistic și PIB per capita pe de-o parte și între indicele de performanță logistic și rata de înscriere la învățământul superior pe de altă parte. În figura 19 se regăsește modelul de regresie multifactorială, iar în figura 20 și în figura 21 se regăsesc cele două modele de regresie simplă precizate mai sus.

### C:\Users\costi\Desktop\proiect econometrie\aplicatia 3 next to came\a prima.png

**Fig. 20. – Model de regresie multifactorial**



**Fig. 21. – Model de regresie unifactorial**



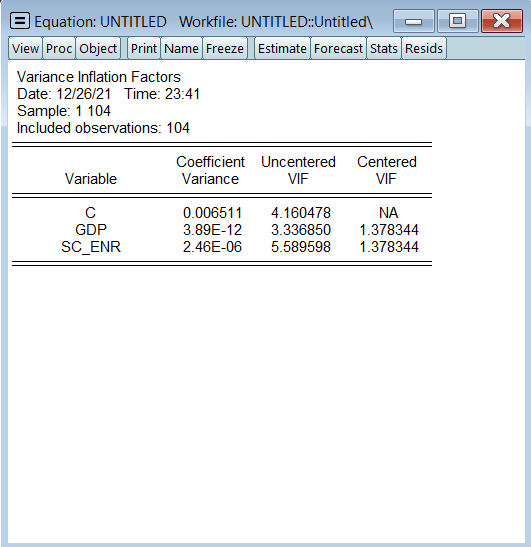
**Fig. 22. – Model de regresie unifactorial**

Pentru a afla dacă avem multicoliniaritate, ne vom folosi de regresiile auxiliare. Astfel, vom compara R-Squared din modelul de regresie multifactorial cu fiecare R-Squared din modele unifactoriale. Aceste verificări s pot observa mai jos:

R multifactorial = 0.64 < R unifactorial = 0.57 (fals) = nu induce multicoliniaritate

R multifactorial = 0.64 < R unifactorial = 0.39 (fals) = nu induce multicoliniaritate

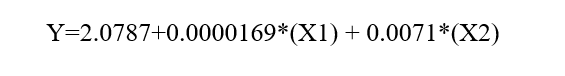
Așadar, putem observa că variabilele alese nu induc multicoliniaritate.



**Fig. 23. – Criteriul factorilor de inflație**

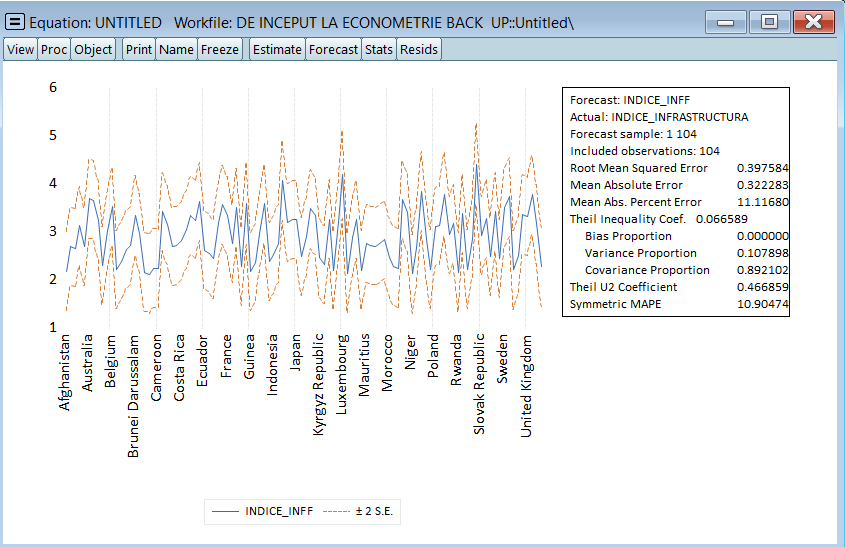
Din figura de mai sus se poate observa că multicoliniaritate este redusă și nu avem nevoie de corectare.

### Prognoză

 În cadrul acestei model, alcătuit din indicele de performanță logistică – ce reprezintă infrastructura, PIB – ul per capita și rata de înscriere la învățământul superior conduc la evoluția pe termen lung a unei țări.

**Fig. 24. – Ecuația modelului de regresie multifactorial**

Privind ecuația obținută în cadrul rezolvării modelului, putem prognoza că pentru cele 104 țări incluse în modelul nostru, infrastructura va avea un rol foarte important în creșterea PIB-ului pe cap de locuitor dar și a ratei de școlarizare în învățământul superior.

****

**Fig. 25. – Prognoză**

Observații:

1. Bias-ul unui estimator reprezintă diferența dintre valoarea așteptată a estimatorului și valoarea adevărată a parametrului prognozat. Un estimator cu bias zero sau fără bias este un aspect pozitiv, acesta indică o prognoză sănătoasă.
2. Covarianța este măsura de variație comună a două variabile aleatorii. În cazul de față avem o covarianță de 89.21%.
3. De asemenea, theil inequality are o valoare de 0.06655. Cu cât este mai apropiată de zero această valoare, cu atât prognoza este aproape perfectă.

### Concluzii

Pentru această aplicație, am pornit de la ipoteza “Cum poate influența indicele de performanță logistic PIB-ul per capita dar și rata de înscriere la învățământul superior?”. Pentru a putea testa această ipoteză am realizat un model de regresie folosind datele privind indicele de performanță logistic, PIB per capita, dar și rata de înscriere la învățământul superior pentru 104 diferita tari, de pe suprafața pământului, reprezentând astfel peste jumătate din numărul total de țări.

Am aplicat testul T pentru verificarea semnificației parametrilor(aceștia au ieșit semnificativi), iar testul F a ilustrat validitatea modelului. Modelul a respectat toate cerințele (homoschedasticitate și neautocorelarea reziduurilor, precum și distribuția normală a reziduurilor).

Având in vedere cele enunțate mai sus putem garanta cu o probabilitate de 95% ca evoluția indicelui influențează într-o maniera pozitiva PIB per capita dar și rata de înscriere la învățământul universitar.

### APLICAȚIA 3 - MODEL DE REGRESIE MULTIPLĂ + DUMMY

* 1. **Literatura review**

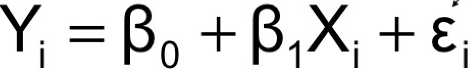
“Educația este larg acceptată ca o resursă fundamentală, atât pentru indivizi, cât și pentru societate. Într-adevăr, în majoritatea țărilor, în prezent, educația de bază este percepută nu doar ca un drept, ci și ca o datorie – se așteaptă de obicei guvernele să asigure accesul la educația de bază în timp ce cetățenii sunt adesea obligați prin lege să atingă educația până la un anumit nivel de bază.”(Max Roser, Esteban Ortiz-Ospina în OurWorldInData Global Education, 2020,p3).

Educația este o investiție valoroasă, atât pentru individ, cât și pentru colectiv. Cea mai obișnuită modalitate de a măsura profiturile educației este de a studia modul în care îmbunătățesc rezultatele individuale pe piața muncii – de obicei, încercând să măsoare efectul educației asupra salariilor.

Pe de altă parte, urbanizarea este una dintre cele mai importante metode de dezvoltare economică a unui țări, la care contribuie cu un rol foarte important și educația. Educația și urbanizarea conduc la dezvoltarea infrastructurii și a PIB per capita.

### Metodologia cercetării

Ecuația modelului de regresie multifuncțional este:



**Fig. 26. - Ecuația modelului de regresie multifactorial**

Modelul va fi implementat în EViews, folosind date extrase de pe site-ul THE WORLD BANK, secțiunea DATA( <https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD> ), ( <https://data.worldbank.org/indicator/SE.TER.ENRR> ), dar și secțiunea LPI (<https://lpi.worldbank.org/international/global?sort=asc&order=Country>). Aceste date vor avea o periodicitate anuală. Prin urmare, variabilele folosite în cadrul acestei aplicații:

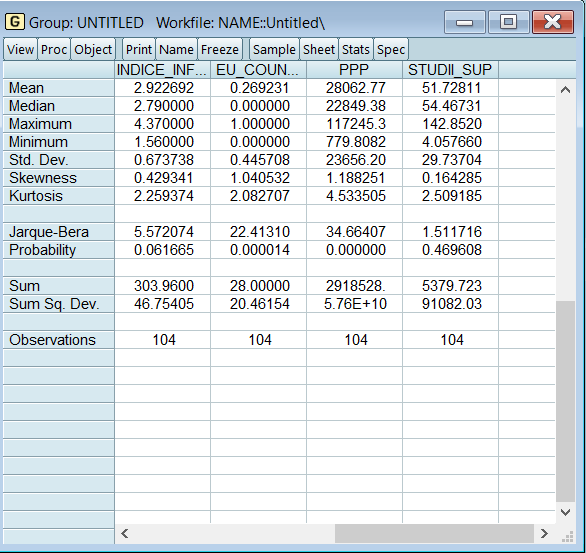
1. Variabila dependenta, indicele de performanță logistic, mai precis infrastructura, având valori între (1,5), unde (1 – foarte puțin dezvoltat, 5 – foarte dezvoltat).
2. Variabila independenta, PIB (produs intern brut) per capita(PPP).
3. Înscriere la învățământ superior, raport din total.
4. Variabila de tip Dummy care face referire la faptul dacă țara este membră Uniunea Europeană sau nu

Datele pentru variabilele precizate mai sus se pot vizualiza în tabelul următor:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Tari | ppp | eu\_country | studii\_sup | indice\_infrastructura |
| Afghanistan | 2082.635648 | 0 | 9.686420441 | 1.81 |
| Albania | 13554.89895 | 0 | 54.96133041 | 2.29 |
| Algeria | 11909.92982 | 0 | 51.36566925 | 2.42 |
| Argentina | 23293.39863 | 0 | 91.6006012 | 2.77 |
| Armenia | 13020.39632 | 0 | 54.57048035 | 2.48 |
| Australia | 50100.22467 | 0 | 107.8080521 | 3.97 |
| Austria | 57059.53894 | 1 | 86.68834686 | 4.18 |
| Bahrain | 47503.06194 | 0 | 50.48189926 | 2.72 |
| Bangladesh | 4547.480016 | 0 | 20.5650692 | 2.39 |
| Belarus | 19430.27877 | 0 | 87.42909241 | 2.44 |
| Belgium | 52623.55967 | 1 | 78.90116882 | 3.98 |
| Benin | 3236.670417 | 0 | 12.51698971 | 2.5 |
| Bhutan | 11620.9159 | 0 | 15.55451012 | 1.91 |
| Bosnia and Herzegovina | 14916.78501 | 0 | 43.4591217 | 2.42 |
| Brazil | 15020.43822 | 0 | 53.25735092 | 2.93 |
| Brunei Darussalam | 61839.11757 | 0 | 31.21179962 | 2.46 |
| Bulgaria | 22957.40013 | 1 | 71.52152252 | 2.76 |
| Burkina Faso | 2168.662213 | 0 | 6.501870155 | 2.43 |
| Burundi | 779.8081757 | 0 | 4.057660103 | 1.95 |
| Cambodia | 4259.203119 | 0 | 13.68776035 | 2.14 |
| Cameroon | 3800.971267 | 0 | 14.26727009 | 2.57 |
| Canada | 50239.99105 | 0 | 70.11302185 | 3.75 |
| Chile | 24740.49615 | 0 | 90.8963089 | 3.21 |
| China | 15497.35999 | 0 | 50.60443878 | 3.75 |
| Colombia | 14865.89324 | 0 | 55.32749176 | 2.67 |
| Costa Rica | 21022.76519 | 0 | 55.31010056 | 2.49 |
| Croatia | 28960.38457 | 1 | 67.65422821 | 3.01 |
| Cyprus | 40476.38627 | 1 | 81.33708191 | 2.89 |
| Czech Republic | 41134.08988 | 1 | 63.76876831 | 3.46 |
| Denmark | 57456.61183 | 1 | 81.18334961 | 3.96 |
| Ecuador | 11839.34324 | 0 | 47.59777069 | 2.72 |
| Egypt, Arab Rep. | 11639.23909 | 0 | 38.90494919 | 2.82 |
| El Salvador | 8822.550774 | 0 | 29.37192917 | 2.25 |
| Estonia | 36158.81965 | 1 | 70.36543274 | 3.1 |
| Finland | 49706.59709 | 1 | 90.2617569 | 4 |
| France | 46620.68448 | 1 | 67.54351807 | 4 |
| Georgia | 14595.63064 | 0 | 60.33444977 | 2.38 |
| Germany | 55142.31878 | 1 | 70.34264374 | 4.37 |
| Ghana | 5442.872147 | 0 | 15.6917696 | 2.44 |
| Greece | 29652.93088 | 1 | 142.8520355 | 3.17 |
| Guinea | 2559.724766 | 0 | 7.921060085 | 1.56 |
| Honduras | 5816.747324 | 0 | 26.16422081 | 2.47 |
| Hungary | 31862.88251 | 1 | 50.30653 | 3.27 |
| Iceland | 58317.43934 | 0 | 73.1032486 | 3.19 |
| India | 6675.361294 | 0 | 28.06055069 | 2.91 |
| Indonesia | 11644.75854 | 0 | 36.31103897 | 2.89 |
| Iran, Islamic Rep. | 13795.31747 | 0 | 62.78507996 | 2.77 |
| Ireland | 84665.60623 | 1 | 77.28288269 | 3.29 |
| Israel | 40690.01183 | 0 | 61.47687149 | 3.33 |
| Italy | 43119.35319 | 1 | 64.29161835 | 3.85 |
| Japan | 42386.39981 | 0 | 64.10384369 | 4.25 |
| Jordan | 10253.81745 | 0 | 34.41532135 | 2.72 |
| Kazakhstan | 26157.65965 | 0 | 53.98825073 | 2.55 |
| Korea, Rep. | 42487.04195 | 0 | 95.864151 | 3.73 |
| Kuwait | 51690.56896 | 0 | 54.36413956 | 3.02 |
| Kyrgyz Republic | 5256.397922 | 0 | 41.26702118 | 2.38 |
| Lao PDR | 7775.131876 | 0 | 14.96675968 | 2.44 |
| Latvia | 30811.07179 | 1 | 93.02448273 | 2.98 |
| Lesotho | 2677.436314 | 0 | 10.19626045 | 1.96 |
| Lithuania | 36365.36049 | 1 | 73.7303009 | 2.73 |
| Luxembourg | 117245.2789 | 1 | 18.59046936 | 3.63 |
| Madagascar | 1630.189357 | 0 | 5.350490093 | 2.16 |
| Malaysia | 28239.50008 | 0 | 45.1253891 | 3.15 |
| Malta | 44482.23968 | 1 | 59.25175858 | 2.9 |
| Mauritania | 5318.085954 | 0 | 5.002089977 | 2.26 |
| Mauritius | 22741.35132 | 0 | 41.5999794 | 2.8 |
| Mexico | 20258.81045 | 0 | 41.52280045 | 2.85 |
| Moldova | 12659.98366 | 0 | 55.77304077 | 2.02 |
| Mongolia | 12317.44092 | 0 | 65.59544373 | 2.1 |
| Montenegro | 21547.28254 | 0 | 56.08097076 | 2.57 |
| Morocco | 7625.735753 | 0 | 35.93592834 | 2.43 |
| Myanmar | 4575.13099 | 0 | 18.81594086 | 1.99 |
| Nepal | 3864.322124 | 0 | 12.41086006 | 2.19 |
| Netherlands | 57901.09701 | 1 | 87.09784698 | 4.21 |
| New Zealand | 43930.88142 | 0 | 82.98303223 | 3.99 |
| Niger | 1229.223782 | 0 | 4.414350033 | 2 |
| North Macedonia | 16725.81473 | 0 | 43.11650085 | 2.47 |
| Norway | 69710.48451 | 0 | 83.01725769 | 3.69 |
| Oman | 29125.45574 | 0 | 38.03620911 | 3.16 |
| Pakistan | 4853.266557 | 0 | 8.9584198 | 2.2 |
| Poland | 31978.53115 | 1 | 68.62010193 | 3.21 |
| Portugal | 34931.7846 | 1 | 65.66265869 | 3.25 |
| Qatar | 93154.2349 | 0 | 17.86992073 | 3.38 |
| Romania | 29248.80932 | 1 | 51.01253891 | 2.91 |
| Russian Federation | 28681.78446 | 1 | 84.5842514 | 2.78 |
| Rwanda | 2138.899699 | 0 | 6.725719929 | 2.76 |
| Saudi Arabia | 48735.4605 | 0 | 68.03988647 | 3.11 |
| Senegal | 3388.946311 | 0 | 12.76259041 | 2.22 |
| Serbia | 17736.0068 | 0 | 67.15866089 | 2.6 |
| Singapore | 100581.158 | 0 | 88.88645172 | 4.06 |
| Slovak Republic | 31530.92023 | 1 | 45.36722946 | 3 |
| Slovenia | 38917.04712 | 1 | 77.11312103 | 3.26 |
| South Africa | 14209.09383 | 0 | 23.80195999 | 3.19 |
| Spain | 40686.98547 | 1 | 91.11257172 | 3.84 |
| Sri Lanka | 13169.0947 | 0 | 19.63023949 | 2.49 |
| Sweden | 53553.31245 | 1 | 72.46116638 | 4.24 |
| Switzerland | 71671.51508 | 0 | 61.37744904 | 4.02 |
| Togo | 2110.82642 | 0 | 13.51175022 | 2.23 |
| Tunisia | 11658.72682 | 0 | 31.74651909 | 2.1 |
| Turkey | 27946.36185 | 0 | 113.2170715 | 3.21 |
| United Kingdom | 47863.2227 | 0 | 61.38288879 | 4.03 |
| United States | 63064.41841 | 0 | 88.29917908 | 4.05 |
| Uruguay | 23588.15139 | 0 | 97.44525146 | 2.43 |
| Uzbekistan | 7252.985197 | 0 | 10.07635021 | 2.57 |

**Tabelul 3. Datele specifice variabilelor în funcție de fiecare țară**

* 1. **Evoluția economică a fiecărei variabile**



**Fig. 27. - Statistica descriptivă pentru indicele de performanță logistică,**

**PIB per capita, înrolarea la învățământul superior, membra Uniunea Europeana.**

Având în vedere output-ul de mai sus, concluziile sunt:

1. Variabila independentă PIB per capita formează o distribuție asimetrică pozitivă (asimetrie de dreapta), astfel predomină valorile mici, iar coeficientul de asimetrie( Skewness ) este 1.1882. Coeficientul de aplatizare ( Kurtosis) are valoarea de 4.5335 și arată faptul că are o distribuție leptocurtică. Minimul seriei este 779, iar maximul 117245.
2. În anul 2018, în medie, PIB per capita atinge o valoare de 28062.77 .
3. Variabila dependentă indicele de performanță logistic urmează o distribuție asimetrică pozitivă, astfel predomină valorile mici, iar coeficientul de asimetrie (Skewness) are o valoare de aproximativ 0.4293. Coeficientul de aplatizare (Kurtosis) are valoarea 2.25 și arată faptul că are o distribuție leptocurtică. Minimul seriei este 1.56, iar maximul e 4.37.
4. În anul 2018, indicele de performanță logistic a fost, în medie, 2.92.
5. Variabila independentă procentul de înscriere la învățământul superior urmează o distribuție asimetrică pozitivă, astfel predomină valorile mici, iar coeficientul de asimetrie (Skewness) are o valoare de aproximativ 0.1642. Coeficientul de aplatizare (Kurtosis) are valoarea 2.50 și arată faptul că are o distribuție leptocurtică. Minimul seriei este 4.06, iar maximul e 142.85.
6. În anul 2018, procentul de înscriere la învățământul superior mediu a fost de 51.71.
7. Variabila independenta care face referire la faptul că țara europeana pozitivă (asimetrie de dreapta), astfel predomină valorile mici, iar coeficientul de asimetrie( Skewness ) are o valoare aproximativă egală cu 1.040. Coeficientul de aplatizare (Kurtosis) are valoarea 2.082 și arată faptul că are o distribuție leptocurtică. Minimul seriei este 0, iar maximul e 1.
8. În anul 2018, în medie, numărul de țări care fac parte din Uniunea Europeană din model a fost de 0.28.

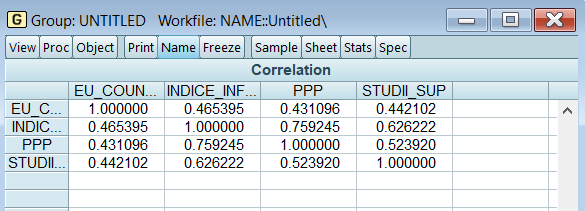
### 

### Fig. 28. - Graficele bară ale variabilelor

### Analizând graficele de mai sus, deducem faptul că:

1. În ceea ce privește indicele de performanță logistic, valorile sunt foarte variate, dar în principal predomină valorile până în 4.37, astfel reiese că indicele de performanță logistic se învârte în jurul valorii de 3.00.
2. În ceea ce privește PIB per capita, valorile sunt de asemenea foarte variate, atingând un maxim de 117245, iar în general se învârte în jurul valorii de 20.000.
3. În graficul ratei de înscriere la învățământul superior, putem observa că există o valoare de 142, dar în general se învârte în jurul valorii de 50.
4. În graficul dedicat țărilor membrilor Uniunii Europene, majoritatea valorilor se învârt în jurul valorii 0.
   1. **Măsurarea intensității dintre variabile**

Înainte de a aplica metoda celor mai mici pătrate, am calculat intensitatea și tipul legăturii dintre variabilele independente (rata de înscriere la învățământul superior și PIB per capita) și variabila dependentă (indicele de performanță logistic). Prin urmare, am calculat coeficientul de corelație Pearson folosindu-ne de EViews, iar output-ul generat se poate observa mai jos:

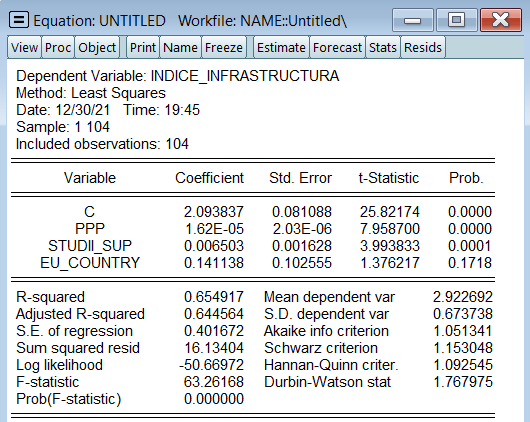


**Fig. 29. - Coeficientul de corelație Pearson**

Din figura de mai sus observăm către variabila independentă PIB per capita și indicele de performanță logistic avem un coeficient de corelație puternic de 0.7592, dar între ratei de înscriere la învățământul superior și indicele de performanță logistic se observă o relație puternică de 0.6262. Între cele doua variabile independente avem o relație, de asemenea, tot pozitivă de 0.5239.

Observăm că între rata de înscriere la învățământul superior și valoarea dummy dedicată țărilor membre Uniunii Europene există un coeficient pozitiv de 0.4421. Tot o dată, între valoarea dummy a țărilor membre Uniunii Europene și PIB per capita, respectiv indicele de performanță logistic există un coeficient pozitiv având valorile 0.4310, respectiv 0.4653.

### Estimare parametri, testare model și interpretare

Am aplicat metoda celor mai mici pătrate în EViews și am realizat o estimare a parametrilor modelului econometric de regresie multifactorială, output care se poate observa în figura care urmează.

**Fig. 30. - Model de regresie multifactorial**

Ecuația generală a modelului de regresie este următoarea:

Y= +\* X1 + \* X2 + \* X3

Particularizăm această ecuație prin faptul că preluăm din figură coeficienții = 2.093 = 0.0000162 și =0.0065, = 0.1411. Ecuația pentru modelul nostru devenind astfel:

Y= 2.093 + 0.0000162 \* X1 + 0.0065 \* X2 + 0.1411\* X3

### Indicatori de bonitate

**1.Coeficientul de determinație**

Coeficientul de corelație R Squared indică cât din variabila dependentă indicele de performanță logistică este influențată de variabilele independente PIB per capita, rata de înscriere la învățământul superior și variabila dummy destinată țărilor membre Uniunii Europene. Așadar, R Squared = 0.654917 arată că PIB per capita împreună, rata de înscriere la învățământul superior și variabila dummy destinată țărilor membre Uniunii Europene explică 65.4917% din modificarea indicelui de performanță logistică, restul de 36.1619% fiind alte variabile care nu sunt incluse în model.

**2.Coeficientul de determinație ajustat**

Adjusted R Squared are aceeași semnificație ca R Squared. Astfel, PIB per capita, rata de înscriere la învățământul superior și variabila dummy destinată țărilor membre Uniunii Europene reprezintă 64.4564% din modificarea indicelui de performanță logistică.

**3.Eroarea standard**

Standard Error of regression reprezintă cu cât se abat de la medie valorile introduse în model. În modelul nostru avem o eroare standard de 0.001628 pentru rata de înscriere la învățământul superior, o eroare standard de 0.00000203 pentru PIB per capita, și o eroare de 0.10255 pentru variabila de tip dummy destinată țărilor membre Uniunii Europene.

### Testarea semnificației parametrilor - Testul T

Pentru a putea testa validitatea modelului statistic utilizat pentru analiza acestui topic, și pentru a dovedi semnificația parametrului β, vom folosi testul T pentru un model bilateral și formulăm următoarele ipoteze:

= 0; (parametrul nu este semnificativ din punct de vedere statistic, ceea ce rezultă în invaliditatea modelului);

: 0; (parametrul este semnificativ din punct de vedere statistic, ceea ce rezultă în validitatea modelului); iar toate datele necesare sunt luate pentru un interval de încredere de 95%, rezultând că

Conform calculelor obținute în cadrul figuri 15(figura de sus), știm că t calculat pentru PIB per capita este aproximativ egal cu 7.95887, rata de înscriere la învățământul superior este aproximativ egal cu valoarea 3.9938, variabila de tip dummy destinată țărilor Uniunii Europene are o valoare aproximativ egala cu 1.3762,valoarea liberă are o valoarea a t-ului 25.8217. Următoarea variabilă de care avem nevoie este t critic, corespunzător lui și numărului de observații, care este egal cu 104. Acest lucru rezultă că t critic are valoarea 1.29.

Pentru a putea respinge ipoteza care infirmă validitatea aplicației, trebuia ca |. În urma calculelor, observăm că t0calc= 7.95887, t1calc= 3.9938, t2calc = 1.3762, t3calc = 25.8217, pe când = **1.29**. Din toate acestea, conform valorii de adevăr a acestor inegalități, putem respinge ipoteza nulă și să acceptăm ipoteza alternativă, ceea ce înseamnă că modelul nostru este semnificativ din punct de vedere statistic.

### Testarea validității modelului- Testul Fisher

Pentru a testa validitatea modelului utilizat în cadrul acestui proiect, vom formula 2 ipoteze:

: modelul nu este valid din punct de vedere statistic,

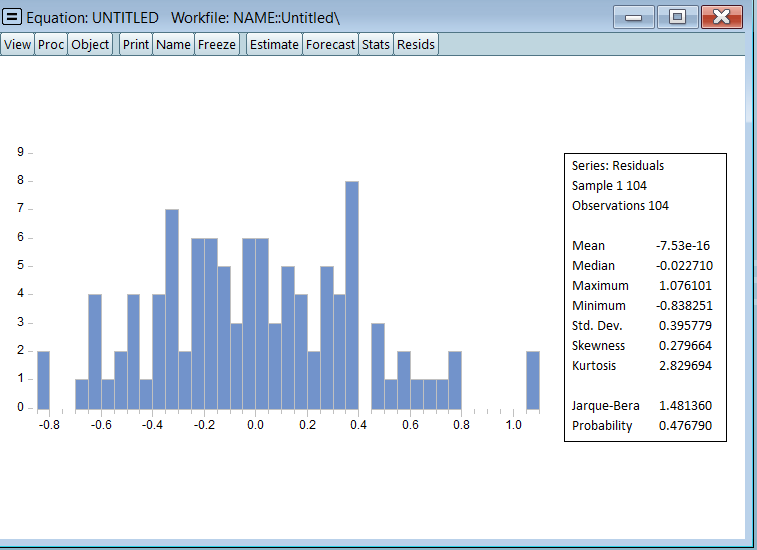
: modelul este valid din punct de vedere statistic.

Pentru a reuși să analizăm aceste ipoteze, ne vom folosi de testul Fisher, pentru o probabilitate de 95%, cu având 102 de grade de libertate (102 de observații, din care scădem 2 pentru eliminarea erorilor), pentru un model de regresie simplă, cu o singură variabilă independentă.

Conform rezultatelor analizei realizate în Fig. 6, aflăm că are valoarea aproximativ egală cu 63.261. Valoarea lui pentru o probabilitate de 95%, un model de regresie simplă, cu 102 de grade de libertate, are valoarea 3.09.

În cazul în care > , vom putea respinge ipoteza nulă, ceea ce înseamnă că modelul ales este valid. În urma comparării acestor valori (63.261> 3.09), rezultă că respingerea lui este posibilă, și putem astfel accepta ipoteza , ceea ce înseamnă că modelul nostru este valid din punct de vedere statistic. O altă variabilă care confirmă validitatea acestui model este Prob(F-statistic), care are valoarea 0, care este mai mică decât pragul de 5% ales de noi.

### Normalitatea distribuției erorilor aleatoare

În această etapă vom verifica ipotezele de normalitate ale erorilor aleatoare, aplicând testul Jarque – Bera. Ipotezele testului sunt următoarele:

: erorile aleatoare urmează o distribuție normală

: erorile aleatoare nu urmează o distribuție normală

**Fig. 31. - Rezultatul testului Jarque-Bera**

Din figura anterioara, observăm că valoarea testului Jarque-Bara este 1.4813, cu o probabilitate de 0.4769. Cum modelul are un nivel de încredere de 95%, probabilitatea testului trebuie să fie mai mare de 5% (0.05). Astfel, 47.69% este mai mare decât pragul de 0.05, acceptăm ipoteza nulă, erorile aleatoare urmează o distribuție normală.

Cum valoarea tabelată a testului Jarque-Bara este de 5.991, iar Jarque-Bera statistic are valoarea 1.4813, putem concluziona că erorile sunt normal distribuite.

### Testarea autocorelării erorilor - Testul Durbin-Watson

În cadrul acestui pas, vom analiza condiția de autocorelare a erorilor modelului ales de noi pentru realizarea acestei aplicații. Pentru acest lucru, ne vom folosi de testul Durbin-Watson și de ipotezele corespunzătoare acestuia:

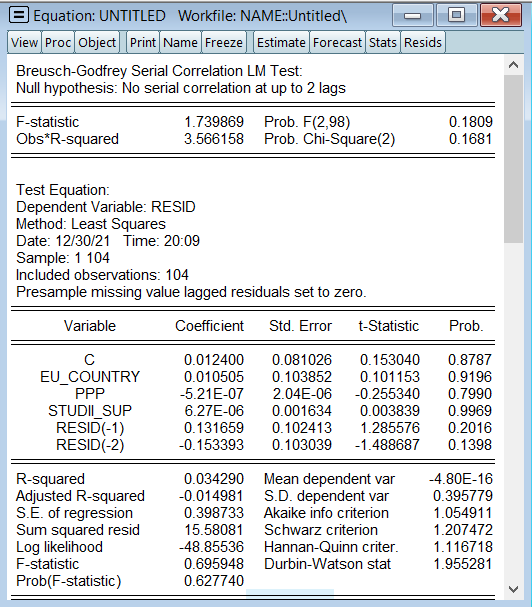
: erorile nu sunt autocorelate,

: erorile sunt autocorelate; unde vom lua probabilitatea implicită de 95%, luând valoarea de 0.05.

Conform figurii 27, rezultatul este reprezentat de indicele calculat luând o valoare egală cu 1.7679. Pentru a putea verifica dacă erorile sunt corelate, vom avea nevoie de figura numărul prezentată mai sus.

În figura anterioară, observăm că CHI-Square(2) are valoarea egala cu 0.1681. Pentru a putea demonstra ca erorile nu sunt autocorelate, rezultatul CHI-Square(2) trebuie sa fie mai mare sau egal cu 0.05, ceea ce în modelul de față este adevărat.

De asemenea, cunoaștem că valorile testului Durbin-Watson se situează în intervalul 0-4. O valoare de 1.95528, situată în intervalul (0,2) ne indică faptul că avem parte de autocorelare pozitivă, și că reziduurile sunt independente.



**Fig. 32. - Rezultatul testului Breush-Godfrey privind corelația**

### Homoscedasticitatea erorilor aleatoare

: erorile aleatoare sunt homoschedastice

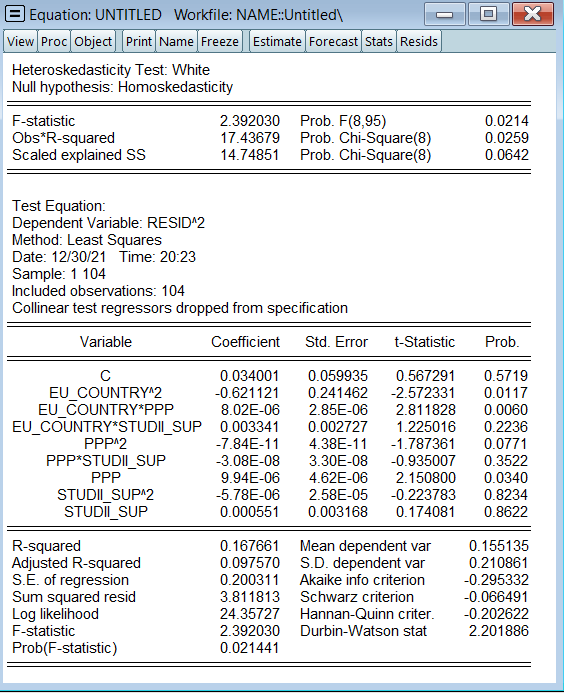
: erorile aleatoare sunt heteroscedastice

Prin urmare, vom compara:

Obs\*R-Squared = 17.43679 Prob(F-statistic) > .

Prob(Chi-Square(2)) = 0.0642 > 0.05

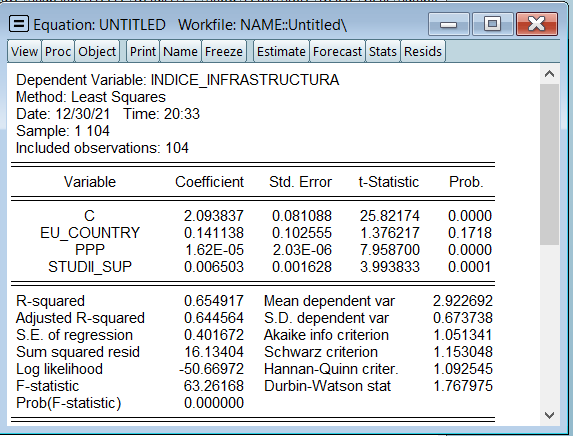
De aici, reiese faptul că respingem ipoteza alternativă și acceptăm ipoteza nulă, prin urmare erorile noastre homoschedastice.



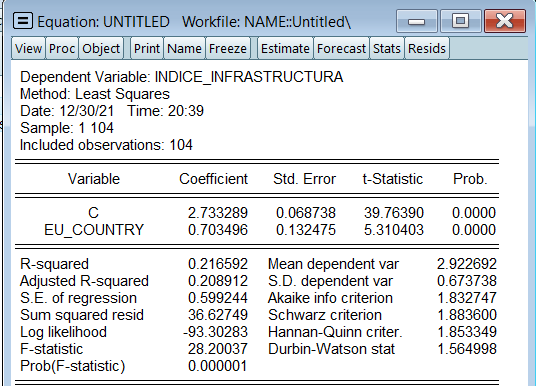
**Fig. 33. – Rezultatul testului White**

#### **Necoliniaritatea variabilelor explicative**

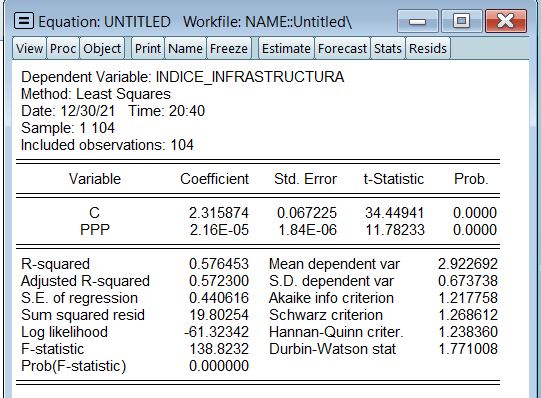
Pentru a verifica necoliniaritatea variabilelor explicative, am creat două modele de regresie unifactorială între indicele de performanță logistic și PIB per capita pe de-o parte și între indicele de performanță logistic și rata de înscriere la învățământul superior pe de altă parte. În figura 19 se regăsește modelul de regresie multifactorială, iar în figura 20 și în figura 21 se regăsesc cele două modele de regresie simplă precizate mai sus.



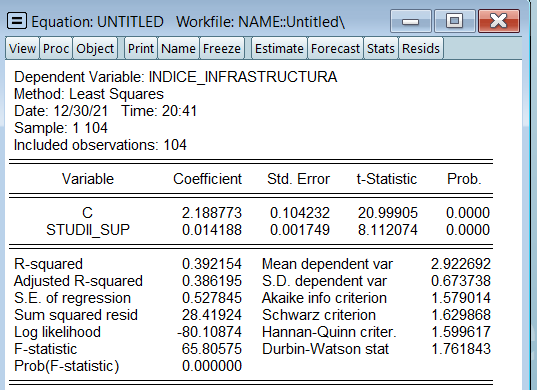
**Fig. 34. – Model de regresie multifactorial**



**Fig. 35. – Model de regresie unifactorial**



**Fig. 36. – Model de regresie unifactorial**



**Fig. 37. – Model de regresie unifactorial**

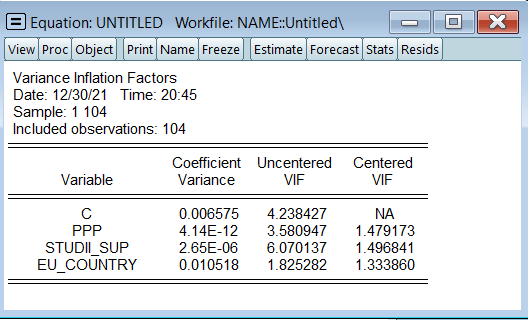
Pentru a afla dacă avem multicoliniaritate, ne vom folosi de regresiile auxiliare. Astfel, vom compara R-Squared din modelul de regresie multifactorial cu fiecare R-Squared din modele unifactoriale. Aceste verificări s pot observa mai jos:

R multifactorial = 0.6549 < R unifactorial = 0.57 (fals) = nu induce multicoliniaritate

R multifactorial = 0.6549 < R unifactorial = 0.39 (fals) = nu induce multicoliniaritate

R multifactorial = 0.6549 < R unifactorial = 0.21 (fals) = nu induce multicoliniaritate

Așadar, putem observa că variabilele alese nu induc multicoliniaritate.



**Fig. 38. – Criteriul factorilor de inflație**

Din figura de mai sus se poate observa că multicoliniaritate este redusă și nu avem nevoie de corectare.

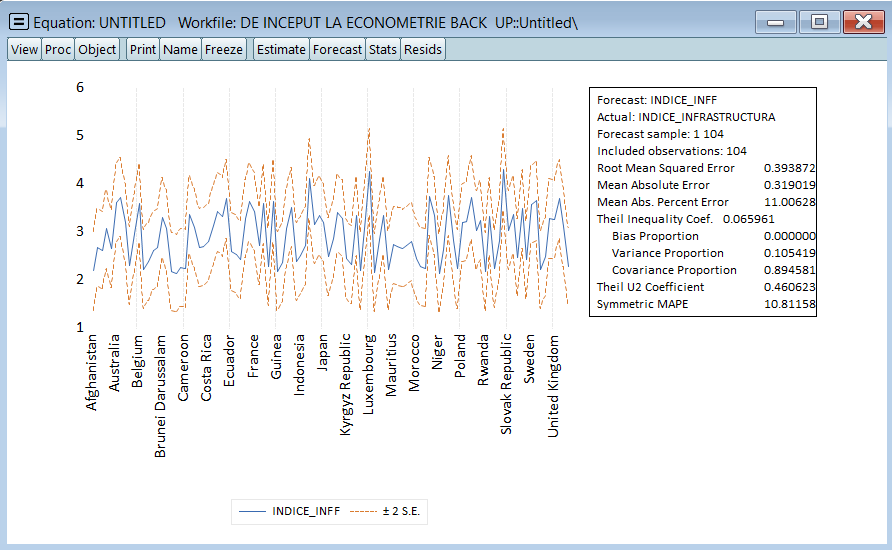
### Prognoză

În cadrul acestei model, alcătuit din indicele de performanță logistică – ce reprezintă infrastructura, PIB – ul per capita și rata de înscriere la învățământul superior conduc la evoluția pe termen lung a unei țări.



**Fig. 39. – Ecuația modelului de regresie multifactorial**

Privind ecuația obținută în cadrul rezolvării modelului, putem prognoza că pentru cele 104 țări incluse în modelul nostru, infrastructura va avea un rol foarte important în creșterea PIB-ului pe cap de locuitor dar și a ratei de școlarizare în învățământul superior în țările membre uniunii europene.



**Fig. 40. – Prognoză**

Observații:

1. Bias-ul unui estimator reprezintă diferența dintre valoarea așteptată a estimatorului și valoarea adevărată a parametrului prognozat. Un estimator cu bias zero sau fără bias este un aspect pozitiv, acesta indică o prognoză sănătoasă.
2. Covarianța este măsura de variație comună a două variabile aleatorii. În cazul de față avem o covarianță de 94.5%.
3. De asemenea, theil inequality are o valoare de 0.017. Cu cât este mai apropiată de zero această valoare, cu atât prognoza este aproape perfectă.

### Concluzii

Pentru această aplicație, am pornit de la ipoteza “Cum poate influența indicele de performanță logistic PIB-ul per capita dar și rata de înscriere la învățământul superior?”. Pentru a putea testa această ipoteză am realizat un model de regresie folosind datele privind indicele de performanță logistic, PIB per capita, dar și rata de înscriere la învățământul superior pentru 104 diferita tari, de pe suprafața pământului, reprezentând astfel peste jumătate din numărul total de țări.

Am aplicat testul T pentru verificarea semnificației parametrilor(aceștia au ieșit semnificativi), iar testul F a ilustrat validitatea modelului. Modelul a respectat toate cerințele (homoscedasticitatea și neautocorelarea reziduurilor, precum și distribuția normală a reziduurilor).

Având in vedere cele enunțate mai sus putem garanta cu o probabilitate de 95% ca evoluția indicelui influențează într-o maniera pozitiva PIB per capita dar și rata de înscriere la învățământul universitar în țările membre uniunii europene.

### APLICAȚIA 4 - MODELE CU DATE DE TIP PANEL

* 1. **Literatura review**

### Uniunea Europeană (UE) este un parteneriat politic și economic care reprezintă o formă unică de cooperare între statele membre, de cooperare între țări suverane. UE reprezintă cea mai recentă etapă a unui proces de integrare început după cel de-al Doilea Război Mondial, inițial de către șase țări din Europa de Vest, pentru a promova interdependența și pentru a face de neconceput un alt război în Europa. În prezent, UE este formată din 27 de state membre, inclusiv majoritatea țărilor din Europa Centrală și de Est, și a contribuit la promovarea păcii, stabilitatea și prosperitatea economică pe întreg continentul european.(Congressional Research Service,”The European Union: Questions and Answers”, Kristin Archick,2021,pagina I).

Uniunea Europeana a asigurat, încă de la începutul ei, o mai bună stabilitate economică, care a evoluat constant de a lungul timpului, impunându-se altfel și o monedă numită euro. Piața unică este principalul motor economic al UE, permițând libera circulație a mărfurilor, serviciilor și a persoanelor.

De asemenea, UE are un rol foarte important în ceea ce privește previzionare și analizarea economică a țărilor membre, ajutând astfel la soluționarea unor probleme care ar putea apărea pe parcurs.

În modelul ce urmează, vom analiza cum unele date precum rata de antreprenoriat, numărul de șomeri, rata de emigrare, eficiența guvernamentală și numărul de zile necesare pentru a porni o afacere determină evoluția economică a 25 de țări membre a Uniunii Europene.

### Metodologia cercetării

### În cadrul acestei aplicații, ne-am gândit să abordăm un model în care este analizată perioada 2012-2018, în 25 de țări de pe teritoriul continentului Europa.

### În acest model, ca variabile independente am considerat că ar fi relevante date precum rata numărului de șomeri, numărul de oameni care emigrează, eficiența guvernamentală, și nu în ultimul rând, am ales numărul necesar de zile pentru a începe o afacere.

Ca și variabilă dependentă, a fost aleasă rata de antreprenoriat, deoarece cu ajutorul acesteia putem realiza un mod valid și foarte ușor de interpretat.

Modelul de regresie lineară simplă, pentru datele de tip panel se scrie:



**Fig. 41. - Ecuația modelului de regresie pentru date de tip panel**

### 

**Fig. 42. - Model de regresie pentru date de tip panel.**

### După cum se poate observa în figura anterioară, rata de antreprenoriat are un efect pozitiv în ceea ce privește celelalte variabile ale modelului. Privind datele prezentate mai sus, creșterea ratei de antreprenoriat are un efect pozitiv în ceea ce privește emigrarea, deoarece aceasta scade. De asemenea, observăm că și numărul de șomeri se diminuează, o dată cu creșterea ratei de antreprenoriat ceea ce conduce la ideea ca populația unei țări nu mai trebuie să părăsească țara pentru a-și găsi un loc de muncă.

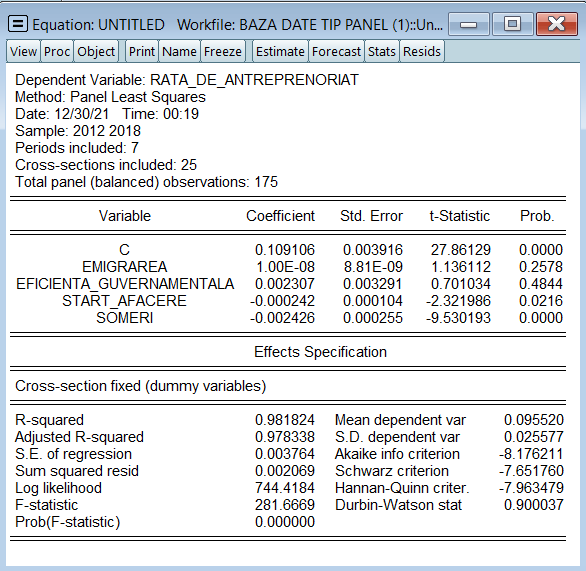
Mai putem observa faptul că creșterea ratei de antreprenoriat conduce la dezvoltarea de alte afaceri, ceea ce reprezintă un lucru pozitiv pentru economia unei țări. După cum ne este prezentat în figura anterioară, creșterea numărului de întreprinderi conduce și la o mai bună organizare a guvernului, în ceea ce privește deciziile guvernamentale, datorată creșterii numărului total de banii din taxe și impozite pe care le generează o firmă.

Dar putem observa că și creșterea ratei de antreprenoriat determina creșterea și numărului de zile necesare pentru a deschide o afacere.

Având aceste date, putem concluziona, că modelul prezentat este un model valid, și în același timp un model logic, care reprezintă ramură economică a unei țări.

În cele ce urmează, ne-am pus problema ce tip de model trebuie să folosim.

Pentru asta, am încercat în primul rând modelul cu efecte fixe, care arata în felul următor.



**Fig. 43. - Model de regresie pentru date de tip panel cu efecte fixe.**

**1.Coeficientul de determinație**

Coeficientul de corelație R Squared indică cât din variabila dependentă rata de antreprenoriat este influențată de variabilele independente Așadar, R Squared = 0.9818 arată emigrarea, eficiența guvernamentală, numărul de zile necesare pentru a începe o afacere și numărul de șoferi explică 98.18% din modificarea ratei de antreprenoriat, restul de 1.82% fiind alte variabile care nu sunt incluse în model.

**2.Coeficientul de determinație ajustat**

Adjusted R Squared are aceeași semnificație ca R Squared. Astfel, emigrarea, eficiența guvernamentală, numărul de zile necesare pentru a începe o afacere și numărul de șoferi reprezintă 98.78% din modificarea ratei de antreprenoriat.

**3.Eroarea standard**

Standard Error of regression reprezintă cu cât se abat de la medie valorile introduse în model. În modelul nostru avem o eroare standard de 8.81E-9 pentru emigrare, o eroare standard de 0.003291 pentru eficiența guvernamentala, o eroare standard pentru numărul de zile necesare pentru începutul unei afaceri de 0.000104, dar nu în ultimul rând, pentru numărul de șoferi o eroare de standard 0.000255, ce reprezintă erori foarte mici.

În continuare, am testat tipul de model cu efecte random, pentru a putea observa diferențele obținute între modele.

### 

**Fig. 44. - Model de regresie pentru date de tip panel cu efecte random.**

**1.Coeficientul de determinație**

Coeficientul de corelație R Squared indică cât din variabila dependentă rata de antreprenoriat este influențată de variabilele independente Așadar, R Squared = 0.476731 arată emigrarea, eficiența guvernamentală, numărul de zile necesare pentru a începe o afacere și numărul de șoferi explică 47.6731% din modificarea ratei de antreprenoriat, restul de 53.3269% fiind alte variabile care nu sunt incluse în model.

**2.Coeficientul de determinație ajustat**

Adjusted R Squared are aceeași semnificație ca R Squared. Astfel, emigrarea, eficiența guvernamentală, numărul de zile necesare pentru a începe o afacere și numărul de șoferi reprezintă 46.4419% din modificarea ratei de antreprenoriat.

**3. Eroarea standard**

Standard Error of regression reprezintă cu cât se abat de la medie valorile introduse în model. În modelul nostru avem o eroare standard de 8.46E-9 pentru emigrare, o eroare standard de 0.002942 pentru eficiența guvernamentala, o eroare standard pentru numărul de zile necesare pentru începutul unei afaceri de 0.000102, dar nu în ultimul rând, pentru numărul de șoferi o eroare de standard 0.000248, ce reprezintă erori foarte mici.

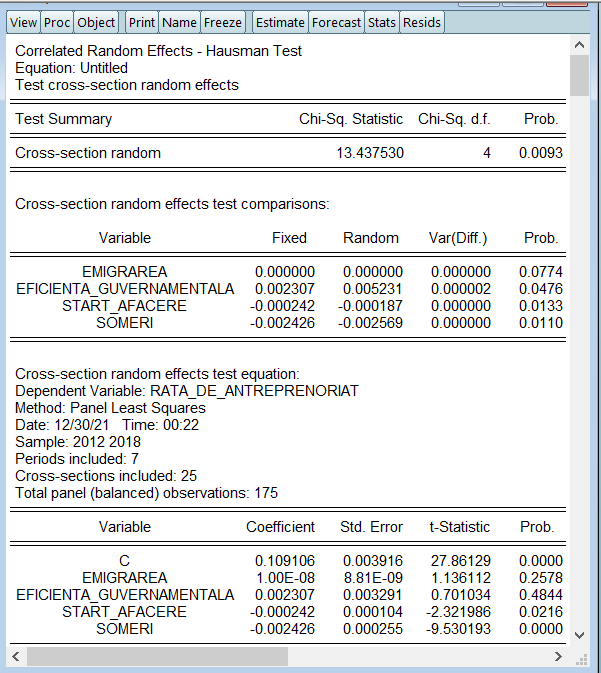
După cum se poate observa, între cele doua modele există diferențe semnificative între valorile lui R-Squared și Adjusted R-Squared. În ceea ce privește erorile standard, între cele modele există diferențe foarte mici, de aceea vom realiza încă un test pentru a putea determina ce model de regresie a datelor vom alege.

### 

**Fig. 45. - Testare redundanță**

În urmă testului de redundanță, putem observa că probabilitatea redundanței este mai mică decât 5%, adică 0.05, ceea ce înseamnă că respingem ipoteza nulă și acceptăm ipoteza alternativă.

Ultimul test pe care o să-l realizăm în cadrul acestei aplicații este constituit de testul Hausman.

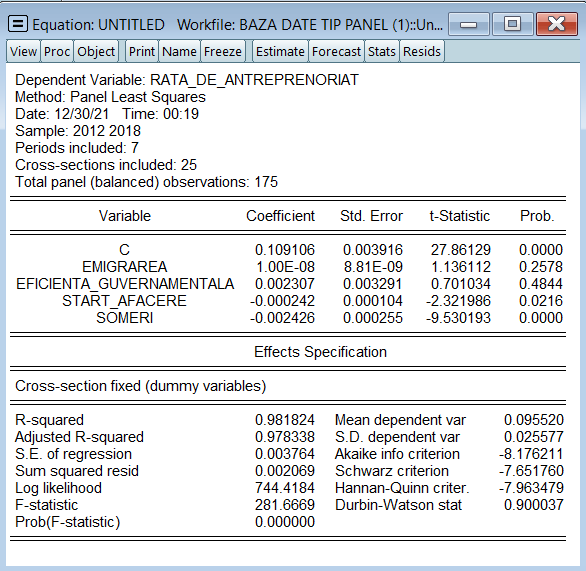


### C:\Users\costi\Desktop\proiect econometrie\aplicatia 3 super buna\hausman2.png

### Fig. 46. - Testul Hausman

În urma testului Hausman, putem observa că probabilitatea obținută, care are valoarea 0.0093, este mai mică decât valoarea de 0.05, de aceea respingem ipoteza nulă. Respingând ipoteza nulă se alege un model cu ipoteza alternativă, preferând un model cu efecte fixe.

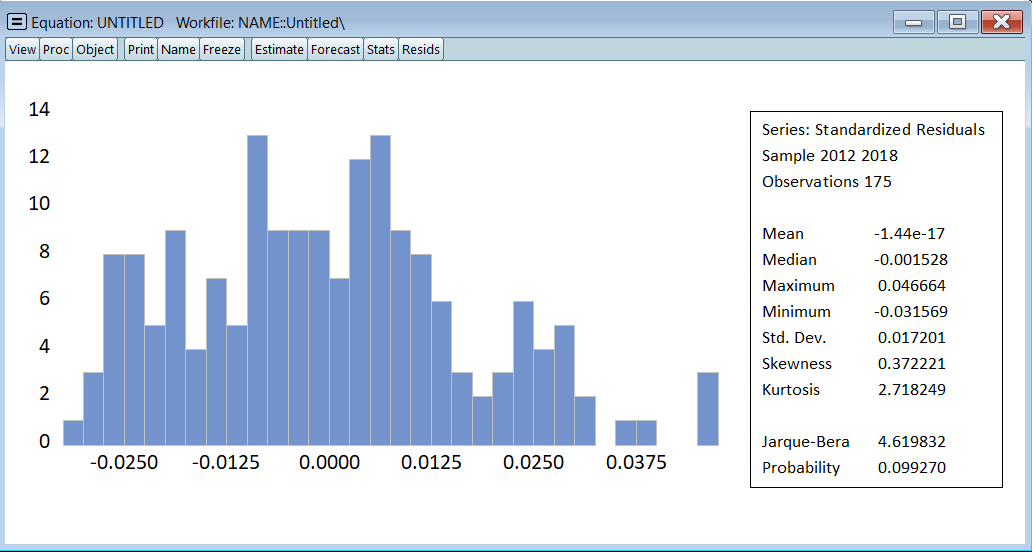
De aceea, putem concluziona că în cadrul modelului nostru, este preferat un model cu date fixe, pe care-l vom prezenta ulterior.



**Fig. 47. - Model de regresie pentru date de tip panel cu efecte fixe.**

### Având în vedere că modelul de regresie relevant pentru aplicația noastră este cu efecte fixe, putem concluziona că o creștere a ratei de antreprenoriat își pune amprenta în economia unei țări, care duce la evitarea migrației populației în alte țări pentru a munci, de asemenea, reduce numărul de șomeri, dar și o creștere eficienței guvernamentale, determinate de contribuțiile pe care le aduc companiile guvernului. De asemenea, creșterea ratei de antreprenoriat ar putea conduce, într-o proporție foarte mică la scăderea numărului de zile necesare pentru a începe o afacere.

* 1. **Normalitatea distribuției variabilelor aleatoare**

În această etapă vom verifica ipotezele de normalitate ale erorilor aleatoare, aplicând testul Jarque–Bera. Ipotezele testului sunt următoarele:

: erorile aleatoare urmează o distribuție normală

: erorile aleatoare nu urmează o distribuție normală

**Fig. 48. - Rezultatul testului Jarque-Bera**

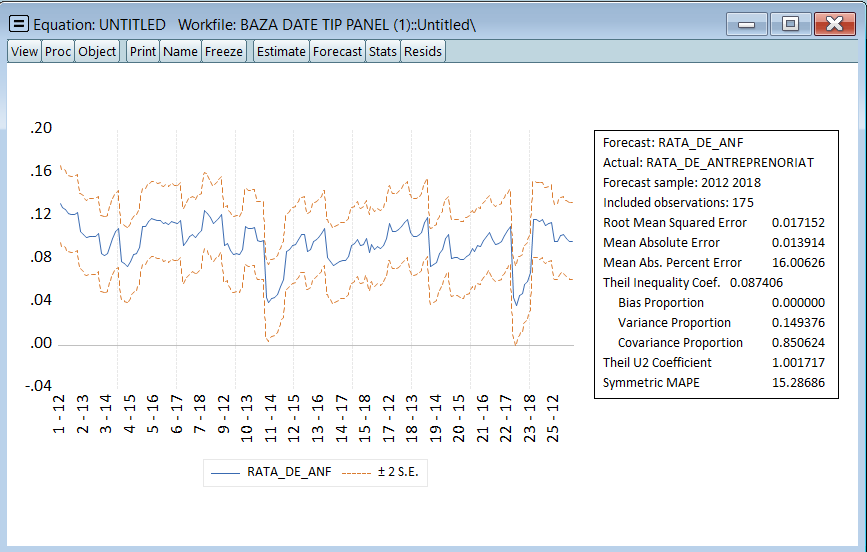
Din figura de mai sus observăm că valoarea testului Jarque-Bera este de 4.6198, cu o probabilitate de 0.0992. Cum modelul are un nivel de încredere de 95%, probabilitatea testului trebuie să fie mai mare de 5%(0.05). Astfel 9.92% este mai mare de pragul de 5%, așadar acceptăm ipoteza nulă, erorile aleatoare urmează o distribuție normală.

Având în vedere că JB statistic are valoarea de 4.6198, care este mai mică decât valoarea tabelată, putem concluziona faptul că erorile sunt normal distribuite.

### Prognoza

Pe termen lung, putem estima că rata de antreprenoriat va conduce la dezvoltarea și mai mult a economiei unei țări, prin generarea de locuri de muncă, o dată cu extinderea companiilor, va stopa emigrarea în alte zone pentru a munci.

De asemenea pe termen lung putem observa că și guvernul va avea de câștigat pe seama creșterii ratei de antreprenoriat, deoarece extinderea companiei, va genera mai multe venituri pentru acesta. Astfel, deciziile guvernamentale, vor fi redirecționate și către ultimul subiectul din modelul nostru, și anume, numărul de zile necesare pentru a porni o afacere, care va scădea cu timpul.



**Fig. 49. - Prognoză**

Observații:

1. Bias-ul unui estimator reprezintă diferența dintre valoarea așteptată a estimatorului și valoarea adevărată a parametrului prognozat. Un estimator cu bias zero sau fără bias este un aspect pozitiv, acesta indică o prognoză sănătoasă.
2. Covarianța este măsura de variație comună a două variabile aleatorii. În cazul de față avem o covarianță de 85.06%.
3. De asemenea, theil inequality are o valoare de 0.0847. Cu cât este mai apropiată de zero această valoare, cu atât prognoza este aproape perfectă.

### CONCLUZII

Pentru această aplicație, am pornit de la ipoteza “Cum poate rata de antreprenoriat emigrarea, numărul de șomeri, deciziile guvernamentale și numărul de zile necesare pentru a porni o afacere?”.

În acest model am putut observa că rata de antreprenoriat își rezervă un loc foarte important în economia unei țări prin generarea de venituri pentru guvern, de locuri de muncă pentru populația unei țări, ceea ce conduce la scăderea numărului de șomeri și a emigrării.

De asemenea, am mai putut observa că o posibilă creștere a ratei de antreprenoriat determină și scăderea numărului de zile necesare pentru a-ți putea deschide o afacere.

### Bibliografie

<https://ro.wikipedia.org/wiki/Urbanizare>

<https://publichealthreviews.biomedcentral.com/articles/10.1186/s40985-019-0116-0>

<https://ourworldindata.org/global-education>

<http://finsys.rau.ro/docs/21-msc-pop.pdf>

<https://faculty.washington.edu/heagerty/Books/Biostatistics/TABLES/t-Tables/>

https://www.itl.nist.gov/div898/handbook/eda/section3/eda3673.htm

<https://www.econstor.eu/bitstream/10419/49919/1/668828234.pdf>

<https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=hlth_hlye&lang=en>

<https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Unemployment_rate_2008-2018_(%25)_new.png>

<https://govdata360.worldbank.org/indicators/h580f9aa5?country=ITA&indicator=388&viz=line_chart&years=1996,2020>

<https://data.oecd.org/entrepreneur/enterprises-by-business-size.htm>

<https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD>

<https://lpi.worldbank.org/international/global?sort=asc&order=Country>

<https://data.worldbank.org/indicator/SE.TER.ENRR>

<https://docs.google.com/spreadsheets/d/1XVPQXvXLtdzOgqlCa6Ji-gLops5FfqDiVl4GLaV-Xro/edit?usp=sharing> – setul de date pentru aplicația 3

<http://www.biokin.com/tools/f-critical.html>

<https://www.investopedia.com/terms/d/durbin-watson-statistic.asp>

<https://sgp.fas.org/crs/row/RS21372.pdf>