

ZHVILLIMI I KOMPETENCAVE TË HULUMTUESVE PËRMES PROGRAMIT R PËR ANALIZËN E TË DHËNAVE

Ilir Palla¹, Eljona Tasho², Silvia Çobani³

*^{1,2,3} Departamenti i Matematikës dhe i Fizikës,
Fakulteti i Shkencave Natyrore dhe Shkencave Humane,
Universiteti “Fan S. Noli”, Korçë
ipalla@unkorce.edu.al*

Abstrakt

Proceset tradicionale manuale të analizës së të dhënave nuk arrijnë më të përmbushin kërkesat për saktësi, efikasitet dhe prezantim të qartë të rezultateve kërkimore. Për këtë arsye trajnimi i studentëve, pedagogëve dhe hulumtuesve në përdorimin e programeve kompjuterike për përpunimin dhe analizimin e të dhënave është bërë i domosdoshëm. Ky studim ka për qëllim hetimin mbi efektivitetin e trajnimit “Analiza e të dhënave me SPSS, R dhe EViews”, të mbajtur në Universitetin “Fan S. Noli”, Korçë. Për të matur këtë efektivitet, u përdorën pyetësorë para dhe pas trajnimit duke përdorur metodën e matjes në shkallën Likert. Analiza statistikore e të dhënave të mbledhura evidentoi një ndryshim statistikisht të rëndësishëm ndërmjet nivelit të aftësive para trajnimit dhe atij pas trajnimit në përdorimin e programit R. Analiza e të dhënave theksoi gjithashtu rëndësinë e përdorimit të programit R në kërkimin shkencor nga ana e studiuesve, si dhe evidentoi interesin e tyre të vazhdueshëm për thellimin e njohurive lidhur me këtë program.

Fjalët kyçe: *Programi R, kompetenca kërkimore, teste statistikore, shkalla Likert.*

DEVELOPING RESEARCHERS' COMPETENCES THROUGH THE R PROGRAM FOR DATA ANALYSIS

Abstract

Traditional manual data analysis processes no longer meet the requirements for accuracy, efficiency and clear presentation of research results. Therefore training students, lecturers and researchers in the use of computer programs for data processing and analysis has become necessary. This study aims to investigate the effectiveness of the training “Data

analysis with SPSS, R and EViews”, held at the “Fan. S. Noli” University, Korça. To measure this effectiveness, pre-training and post-training questionnaires were used using the Likert scale measurement method. Statistical analysis of the collected data showed a statistically significant difference between pre-training and post-training skills in using the R program. Data analysis also showed the importance of using R for researchers’ scientific research and their further desire to deepen their knowledge about this program.

Keywords: *R program, research competence, statistical tests, Likert scale.*

1. HYRJJE

Në kontekstin e epokës së informacionit, rritja eksponenciale e sasisë së të dhënave ka bërë që përdorimi i mjeteve kompjuterike për analizën dhe përpunimin e tyre të jetë jo vetëm i dobishëm, por thelbësor. Proceset tradicionale manuale të analizës së të dhënave nuk arrijnë më të përbushin kërkesat për saktësi, efikasitet dhe prezantim të qartë të rezultateve. Në këtë kuadër, gjuha e programimit *R* dhe mjedisi i saj statistik përbëjnë një instrument të fuqishëm për trajtimin dhe vizualizimin e të dhënave (Palla I., Hoxha S., 2024), duke u bërë një ndër mjetet kryesore në praktikën moderne të analizës statistikore.

R është një platformë softuerike dhe një gjuhë programimi e gjithanshme që lejon përdoruesit të krijojnë funksione dhe paketa të personalizuara për të përbushur kërkesat e tyre specifike. *R* është një gjuhë programimi dhe mjedis softuerik që përdoret kryesisht për llogaritjet statistikore, analizën e të dhënave, modele të avancuara dhe paraqitjen grafike. Komuniteti i përdoruesve të *R* ka krijuar një ekosistem të gjerë paketash që u shërbejnë fushave si ekonomia, mjekësia, bioinformatika etj., duke e bërë atë të gjithanshëm dhe të fuqishëm për kërkime shkencore dhe aplikacione të drejtuara nga të dhënat.

Përdorimi i programit *R* nga studiuesit në Shqipëri, është në rritje dhe po fiton rëndësi në analizën e të dhënave dhe realizimin e kërkimeve shkencore. Megjithatë, krahasuar me përdoruesit e *R* në vendet jashtë Shqipërisë, vihet re një diferencë në nivelin e përhapjes dhe thellësisë së përdorimit. Në kontekstin ndërkombëtar, *R* përdoret gjerësisht dhe në mënyrë më të avancuar, duke përfshirë aplikime të sofistikuar në

fusha të ndryshme shkencore dhe industri, si dhe mbështetet nga një komunitet i madh dhe burime të shumta trajnimi (Robinson, D., 2017). Ndërkohë, në Shqipëri, megjithëse ekziston interes në rritje dhe një bazë përdoruesish në zgjerim, përdorimi i *R* mbetet ende në fazë zhvillimi, me nevojë për më shumë trajnim, burime dhe integrim në praktikën kërkimore bashkëkohore. Po përmëndim disa prej tyre: analiza e regresit linear (Palla I., 2022), testim hipotezash me metodën bootstrap (Palla I., Ekonomi E., 2017), performanca e metodes Bootstrap me blloqe bazuar te mbetjet (Margo Zeqo L., 2023). Për të përmirësuar cilësinë e analizave dhe saktësinë e rezultateve, trajtimi i vlerave të mungura në seritë kohore mund të realizohet me imputacion të shumëfishtë, duke përdorur programin *R* dhe paketat *mice*, *Hmisc*, *mi* dhe *Amelia*, çka mundëson rikuperim të besueshëm të të dhënave dhe përforcon bazën për analiza statistikore të besueshme. (Milo Tasho E., Margo Zeqo L., 2022), modele të avancuara SEM Modelimi i Ekuacionit Struktural (Jashiku E., Palla I. 2025), testim hipotezash për krahasimin e metodave të ndryshme edukimi (Karasaliu A., Palla I. 2025). Studime të mëparshme kanë trajtuar analiza të serive kohore dhe parashikime ekonomike, duke përfshirë gjetjen e modeleve ARIMA dhe ETS, si dhe përdorimin e metodave të parashikimit në seritë kohore (Gjika, E., Ferrja, A., Kamberi, A., 2019). Në të njëjtën kohë, është shqyrtuar ndikimi i faktorëve makroekonomikë dhe makrofinanciarë mbi kreditë me probleme në sistemin bankar shqiptar (Nerjaku, S., Sinaj, V., 2019).

Duke qenë një program me burim të hapur, *R* gjithashtu lejon riprodhueshmërinë në analizën e të dhënave, e cila është thelbësore në kërkimin shkencor. Aksesueshmëria e programit *R*, e kombinuar me kapacitetet e tij të avancuara statistikore dhe funksionalitetet e vizualizimit, e bën atë një mjet të përdorshëm dhe miqësor për përdoruesit, të përshtatshëm si për fillestarët, ashtu edhe për ekspertët në fushat statistikore dhe të analizës së të dhënave. Një arsye tjetër e rëndësishme për përdorimin e *R*-së lidhet me mjetet falas të zhvilluara për të rritur riprodhueshmërinë e hulumtimit, të cilat janë ndërtuar mbi *R Markdown* dhe *knitr* (Xie, 2015; Xie et al., 2018a).

Për të lehtësuar përdorimin e gjuhës së programimit *R*, është zhvilluar RStudio, një mjedis i integruar zhvillimi (Integrated Development Environment – IDE), i cili ofron një ndërfaqe të strukturuar dhe funksionale. Ky mjedis synon të përmirësojë përvojën e përdoruesit

duke mundësuar akses më të lehtë në funksionalitetet e R , çka e bën të përshtatshëm si për përdoruesit fillestarë, ashtu edhe për ata me përvojë të avancuar në analizën statistikore dhe programim.

Mjedisi RStudio përbëhet nga disa dritare kryesore: dritarja e skriptit (sipër majtas), ku mund të shkruhen, modifikohen dhe ruhen koleksionet e komandave; dritarja e komandave (poshtë majtas), ku ekzekutohen komandat; dritarja e hapësirës së punës (sipër djathtas), ku mund të vizualizohen të dhënat dhe objektet aktuale në kujtesën e R -së; dhe dritarja skedarë/grafiqe/paketa/ndihmë (poshtë djathtas), ku përdoruesit mund të hapin skedarë, të shikojnë grafikë, të instalojnë dhe ngarkojnë paketa, ose të përdorin funksionin e ndihmës (Torfs & Brauer, 2014).

2. QËLLIMI I PUNIMIT

Ky studim, i realizuar në kuadër të projektit të financuar nga Universiteti “Fan S. Noli”, me titull *Teknologjia dhe analiza e të dhënave në kërkimin shkencor: Aplikime në SPSS, R dhe Eviews*, synoi të vlerësonte efektivitetin e trajnimit praktik në programin R për përmirësimin e kompetencave kërkimore të studentëve dhe stafit akademik, si dhe të identifikonte sfidat ekzistuese dhe masat që Universiteti mund të ndërmarrë për të forcuar aftësitë kërkimore të studentëve dhe pedagogëve.

3. METODOLOGJIA E PËRDORUR

Pjesëmarrësit në studim ishin pedagogë e studentë nga Universiteti “F. S. Noli” i Korçës, mësues nga shkollat parauniversitare dhe punonjës të administratës së universitetit, të cilët morën pjesë në një sesion praktik trajnimi dyditor në programin R .

Trajnimi kishte si qëllim zhvillimin e aftësive të pjesëmarrësve për:

- Procedurat për instalimin dhe njohjen fillestare me ambientin e programimit R dhe mjedisin zhvillues RStudio;
- Përdorimin e funksioneve bazë në R ;
- Ndërtimin dhe importimin e kornizave të të dhënave;
- Analizën përmbledhëse dhe vizualizimin grafik të të dhënave;
- Ndërtimin e modeleve të regresionit linear;
- Kryerjen e testeve statistikore për verifikimin e hipotezave statistikore dhe vlerësimin e rezultateve.

Gjatë zhvillimit të këtij projekti, u identifikua një mungesë e konsiderueshme e kompetencave kërkimore në përdorimin e programeve kompjuterike për analizimin e të dhënave. Ky hulumtim synon të plotësojë këtë boshllëk të identifikuar duke ofruar një pasqyrë empirike mbi rolin e trajnimit praktik në *R* për përmirësimin e aftësive kërkimore. Për të arritur këtë objektivi, u organizua një trajnim dy javor në Universitetin “Fan S. Noli”, i titulluar “Analiza e të dhënave me SPSS, R dhe EViews”, pjesë e të cilit ishte edhe një sesion trajnimi dyditor për programin *R*. Bazuar në përmbajtjen e programit mësimor të statistikës, trajnimi kishte për qëllim të mbështeste pedagogët në aktivitetet e tyre kërkimore, studentët e ciklit të dytë (niveli Master), si dhe individët e tjerë të interesuar, duke i ndihmuar të zhvillonin dhe të thellonin kompetencat në fushën e analizës së të dhënave.

Për të vlerësuar efektivitetin e trajnimit, u përdor një pyetësor anonim para dhe pas sesionit, për të matur nivelet e njohurive të vetë-raportuara të pjesëmarrësve në gjashtë fusha kryesore: futja e të dhënave, statistika përshkruese, vizualizimi i të dhënave, analiza e regresionit dhe testimi i hipotezave statistikore. Të dhënat u analizuan vetëm për qëllime studimi dhe u trajtuan me konfidencialitet të plotë. Rezultatet e trajnimit ishin premtuese dhe demonstuan përmirësimin e aftësive të pjesëmarrësve në përdorimin e programit *R* për të kryer prezantimin dhe analizën statistikore të të dhënave sipas fushës së tyre të studimit. Rezultatet treguan një ndryshim statistikisht të rëndësishëm në të gjitha fushat e përmendura më sipër, duke konfirmuar kështu se pjesëmarrësit performuan më mirë pas trajnimit sesa më parë.

Seksioni i parë i pyetësorit përfshinte pesë pyetje mbi karakteristikat demografike dhe akademike të pjesëmarrësve në projekt, të tilla si gjinia, mosha, niveli i studimit, njohuri paraprake mbi programet kompjuterike dhe fusha e studimit.

Seksioni i dytë përfshin perceptimet e pjesëmarrësve para dhe pas trajnimit në lidhje me aftësinë e tyre mbi përdorimin e mjedisit të programit *R*, importimin e të dhënave, statistikat përshkruese, vizualizimin e të dhënave, instalimin e paketave, përpunimin e të dhënave dhe interpretimin e rezultateve statistikore të nxjerra nga programi *R*.

Në seksionin e tretë të pyetësorit, pjesëmarrësit dhanë perceptimet e tyre mbi aftësitë e përdorimit të Rstudio, rëndësinë e përdorimit të këtij programi për fushën e tyre akademike dhe interesin e tyre për të thelluar njohuritë e tyre rreth analizës së të dhënave me R. Për seksionin e dytë dhe të tretë u përdor shkalla Likert (Likert, 1932) me 5 pikë.

Analiza e të dhënave të mbledhura në studim u realizua me programin R (version 4.3.1). Në hapin e parë, grupi i të dhënave u ngarkua në R në formën e një skedari CSV, ndërsa nëngrupet e çështjeve u organizuan në korniza të veçanta të dhënash për lehtësinë e përpunimit dhe analizës.

4. REZULTATET

4.1. Karakteristikat demografike dhe akademike të pjesëmarrësve

Në total pyetësorin e plotësuan 33 pjesëmarrës, ku 21% ishin meshkuj dhe 79% ishin femra, 59% ishin të moshës 20-30 vjeç, 24% të moshës 31-40 vjeç, 15% të moshës 41-50 vjeç dhe 3% të moshës mbi 50 vjeç.

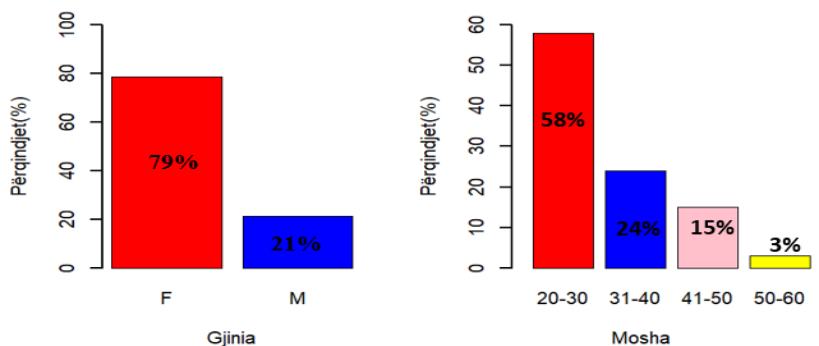


Figura 1. Diagrami me shtylla: majtas gjinia, djathtas moshja e të anketuarëve

Shpërndarja e pjesëmarrësve sipas fushave të studimit ishte: Shkenca Natyrore 61%, Shkenca Humane 12%, Shkenca Ekonomike 12%, Edukim dhe Pedagogji 9%, dhe Shkenca Kompjuterike dhe Inxhinierike dhe Shkenca Shoqërore secila 3%.

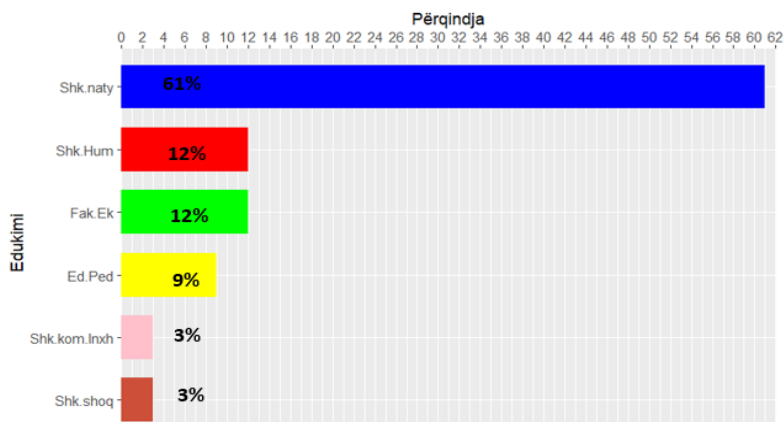


Figura 2. Shpërndarja e të anketuarëve sipas fushave të studimit

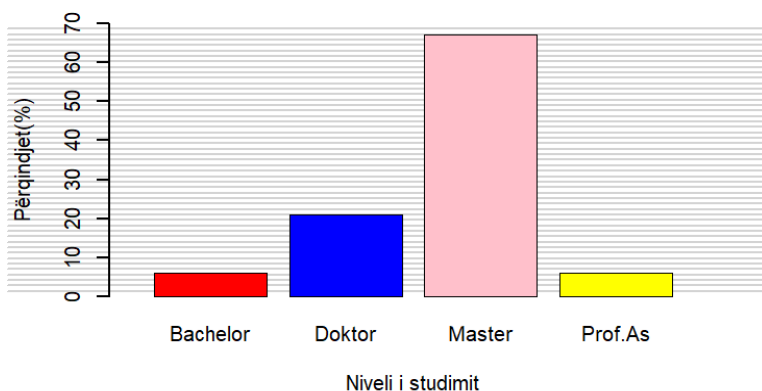


Figura 3. Shpërndarja e të anketuarëve sipas nivelit të studimit

Një përmbledhje të përgjigjeve mbi pyetjen “Keni përdorur më parë ndonjë program kompjuterik për analizimin e të dhënave?”, pyetje në të cilën mund të zgjidhje disa alternativa pasqyrohet në tabelën e mëposhtme:

Tabela1. Përqindja e përdorimit të programeve kompjuterike

Programet kompjuterike	Përdorimi i mëparshëm nga pjesmarrsit
SPSS	27%
Excel	9%
Excel dhe SPSS	46%
Excel, SPSS dhe Python	3%
Excel dhe EViews	6%
Excel, SPSS, Python dhe R	3%

Nga analizimi i përgjigjeve nxjerrim konkluzionin se kanë përdorur për analizimin e të dhënave 67% excelin, 77% SPSS, 6% EViews dhe vetëm 3% programin R.

4.2 Analiza statistikore e perceptimeve

Mesataret e vlerësimeve të perceptimeve mbi njohuritë në programin R, para trajnimit dhe pas trajnimit jepen në tabelën 2 të mëposhtme:

Tabela 2. Vlerësimet mesatare të perceptimeve

Pyetja	Para trajnimit	Pas trajnimit
P.6.1: Di të përdor mjedisin e punës në RStudio	1.31	3.70
P.6.2: Di të importoj të dhëna në R	1.39	3.91
P.6.3: Gjej statistikat përshkruese të të dhënave	1.21	3.61
P.6.4: Di të vizualizoj të dhënat me R	1.15	3.70
P.6.5: Di të instaloj dhe përdor paketa në R	1.21	3.45
P.6.6: Di të gjej regresin linear me R	1.30	3.67
P.6.7: Di të testoj hipoteza statistikore me R	1.31	3.58

Rezultatet e nxjerra tregojnë një përmirësim të dukshëm të kompetencave kompjuterike për analizimin e të dhënave me R. Për të konfirmuar këtë tezë përdorim testin “Wilcoxon Signed Rank Test” (David F. Bauer,1972) dhe (Hollander & Wolfe,1973). Duke aplikuar testin Wilcoxon Signed Rank Test, p-vlera ka rezultuar $p=.000<0.05$, që tregon se kemi një ndryshim domethënës midis aftësive të pjesmarrëseve para trajnimit dhe pas trajnimit.

Për më tepër, rezultatet për pohimin P.7.1 (“Ndihejm i sigurtë në aftësitë e mia për të analizuar të dhënat me *R*”) tregojnë se shumica e pjesëmarrësve kanë një nivel mesatar deri të lartë sigurie në aftësitë e tyre për të analizuar të dhënat me *R*. Konkretisht, 46% e vlerësojnë veten “shumë mirë” dhe 33% “mirë”. Nga ana tjetër, Figura 5 tregon se përdorimi i *R* perceptohet si shumë i rëndësishëm për fushën akademike, ku 43% e vlerësojnë “shumë” dhe 30% “mesatarisht”. Këto rezultate sugjerojnë se, megjithëse vetëbesimi është përgjithësisht pozitiv, ekziston ende hapësirë për përmirësim, ndërkohë që rëndësia e programit *R* është gjerësisht e pranuar nga pjesëmarrësit.

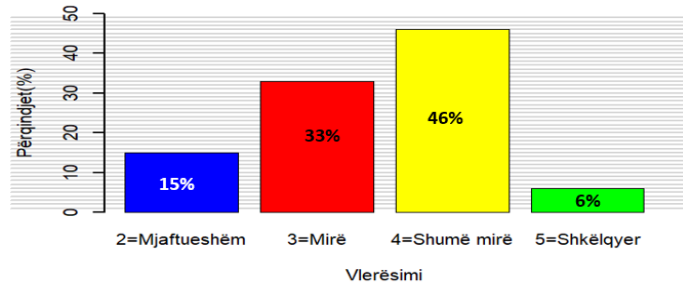


Figura 4. Diagrami mbi vetëvlerësimin

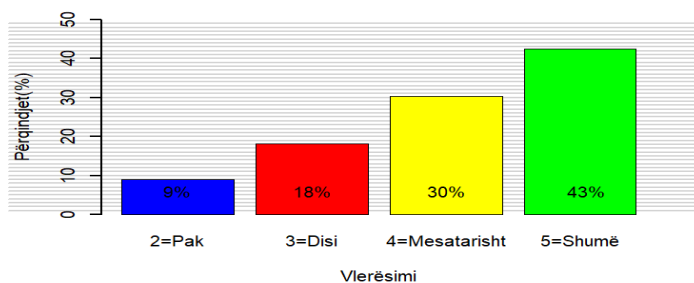


Figura 5. Perceptimi mbi rëndësinë e e përdorimit të programit *R* për fushën e tyre akademike

5. PËRFUNDIME & REKOMANDIME

- Studimi evidentoi mangësitë që ekzistojnë në njohuritë e studiuesëve në përdorimin e programeve kompjuterike për analizimin e të dhënave.

- Rezultatet e studimit demonstrojnë suksesin e trajnimit në ndërtimin e kompetencave kompjuterike për analizimin e të dhënave me *R*.
- Përdorimi i programit *R* është shumë i rëndësishëm në fushën e tyre akademike për përpunimin, analizën përshkruese dhe analizën statistikore të të dhënave.
- Për shkak të ndihmës që jep vet programi *R*, lehtësisë së gjuhës dhe mundësisë që është pa pagesë, ky program është një program që po përdoret më së shumti nga kërkuesit shkencorë.
- Rekomandohet të organizohen trajnime të tjera për integrimin e teknologjisë për të përmirësuar kompetencat kërkimore përmes trajnimit praktik të *R-së*.
- Në trajnimet e ardhshme mund të përfshihen module më të avancuara specifike, si analiza e serive kohore, modeli i ekuacionit strukturor (SEM), analiza e komponentit kryesor (PCA), cluster analysis, analiza e vlerave të munguara, etj.

Kufizime të studimit:

- Të dhënat nuk janë zgjedhje e rastit, por janë përfituar nga pjesmarrsit që janë regjistruar me dëshirë në trajnim.
- Rezultatet janë vlerësime të perceptimeve të vetë deklaruara, nuk janë matje që vijnë nga testime kontrolli njohurish.

REFERENCAT

[1] David F. Bauer (1972). “*Constructing confidence sets using rank statistics*”. *Journal of the American Statistical Association* 67, 687–690. Doi:10.1080/01621459.1972.10481279.

[2] Foundation, T. R. (2019). The r project for statistical computing. Retrieved from <https://www.r-project.org>

[3] Jashiku E., Palla I. (2025), “*The Impact of Tourism on Quality of Life: Community Perception in the Albanian Area Around Ohrid Lake*”. *Conferencii*, Vol.12, 2025, 06 September 2025, E-ISSN: 2753-6408.

[https://conferencii.com/arc/Conferencii_2025_12\(39\).pdf](https://conferencii.com/arc/Conferencii_2025_12(39).pdf)

[4] Gjika E., Ferrja A., Kamberi A., (2019), “*A Study on the Efficiency of Hybrid Models in Forecasting Precipitations and Water Inflow Albania Case Study*”, *Advances in Science Technology and Engineering Systems Journal*. DOI: 10.25046/aj040129

[5] Karasaliu A., Palla I. (2025), “*Cultivating words: integrating immersive technologies and translation in agricultural settings*”, International Journal of Evaluation and Research in Education (IJERE), 2025-12-04, p-ISSN: 2252-8822, e-ISSN: 2620-5440.

<http://doi.org/10.11591/ijere.v14i6.35451>

[6] Likert, R. (1932). A Technique for the Measurement of Attitudes. Archives of Psychology. 140, 1-55.

https://legacy.voteview.com/pdf/Likert_1932.pdf

[7] Margo Zeqo L. (2023), *Some empirical results using block bootstrap in estimating the coefficients of a periodic autoregressive model*, Balcan Journal of Interdisciplinary Research, Vol. 9, No.2, 2023, E-ISBN 2411-9725/ISSN 2410-759X, pp.34-41.

<https://iipcccl.org/wp-content/uploads/2023/09/04-1.pdf>.

[8] Myles Hollander and Douglas A. Wolfe (1973). Nonparametric Statistical Methods. New York: John Wiley & Sons. Pages 27–33 (one-sample), 68–75 (two-sample).

[9] Nerjaku S., Sinaj V., (2019), “A VAR Model for Non-Performing-Loans in Albania”, European Modern Studies Journal, Vol 8 No 4. DOI: 10.59573/emsj.8(4).2024.20

[10] Palla I., Hoxha S., “*Graphical presentation of data using R*”, Third international Conference: “Research, Application and Educational Methods”, Science, Technology and Society in 21st Century, 1-2 March 2024, Korçë, Albania. ISBN (print): 9789928473196.

<https://unkorce.edu.al/wp-content/uploads/2024/07/Punimet-RAEM>

[11] Palla I., (2023), “*Linear regression analysis in R*”, International Conference on Natural Sciences, Mathematics and Technology, Durrës, Albania. Proceedings Book, pg. 191, ISBN: 9789928267726.

<https://uamd.edu.al/wp-content/uploads/2023/06/Proceeding-Book-ICNSMT-2023-1.pdf>

[12] Palla, I., (2022), “*The Comparison of Some Methods in Analysis of Linear Regression Using R Software*”, European Journal of Formal Sciences and Engineering, Vol. 5 No.2, page 38–49. ISSN 2601-8675 (Online).

DOI: 10.26417/ejef.v3i3.p22-31.

https://revistia.com/files/articles/ejfe_v5_i2_22/Palla.pdf

[13] Palla I., Ekonomi E., (2017), “*The Bootstrap Methods to Test the Equality of Two Means*”. EUROPIAN ACADEMIC RESEARCH. Volume V. Issue 8. Impact Factor: 3.4546. ISSN 2286-4822. www.euacademic.org

- [14] Tasho Milo E., Zeqo Margo L, (2022) “Comparison methods of estimating missing data in real data time series” Asian-European Journal of Mathematics , <https://doi.org/10.1142/S1793557122502436>
- [15] Tyson S. Barrett (2019), “*Six Reasons to Consider Using R in Psychological*”, DOI: 10.31234/osf.io/8mb6d.
- [16] Robinson, D. (2017, October). The impressive growth of r. StackOverflow Blog.
<https://stackoverflow.blog/2017/10/10/impressive-growth-r/>
- [17] Torfs, P., Brauer, C., (2014), “A (very) short introduction to R”, Hydrology and Quantitative Water Management Group Wageningen University, Netherlands.
- [18] Xie, Y. (2015). Dynamic documents with R and knitr (2nd ed.). Boca Raton, Florida: Chapman; Hall/CRC. <https://yihui.name/knitr/>