

## АНАЛІЗ РЕЗУЛЬТАТІВ НАЦІОНАЛЬНОГО МУЛЬТИПРЕДМЕТНОГО ТЕСТУ З ВИКОРИСТАННЯМ ІНСТРУМЕНТІВ СТАТИСТИЧНОГО МОДЕЛЮВАННЯ В RSTUDIO

**Ясінський А. М.**

кандидат педагогічних наук, доцент,  
кафедра математичного моделювання  
Приватного вищого навчального закладу  
«Міжнародний економіко-гуманітарний університет  
імені академіка Степана Дем'янчука»  
ORCID ID: 0000-0002-1894-1314

**Соловей Л. Я.**

старший викладач  
кафедра інформаційних систем та обчислювальних методів  
Приватного вищого навчального закладу  
«Міжнародний економіко-гуманітарний університет  
імені академіка Степана Дем'янчука»  
ORCID ID: 0009-0001-2832-1741

**Колупаєв Б. Б.**

доктор фізико-математичних наук,  
професор кафедри інформаційних систем та обчислювальних методів  
Приватного вищого навчального закладу  
«Міжнародний економіко-гуманітарний університет  
імені академіка Степана Дем'янчука»  
ORCID ID: 0000-0002-0198-9872

**Ільків О. С.**

кандидат педагогічних наук, доцент,  
кафедра інформатики, кінезіології та кіберспорту  
Львівського державного університету фізичної культури  
імені Івана Боберського  
ORCID ID: 0000-0003-0864-4794

*У статті представлено результати емпіричного дослідження особливостей успішності учасників Національного мультипредметного тесту (далі – НМТ) із застосуванням інструментарію програмного середовища RStudio. Основною метою роботи є виявлення статистичних закономірностей у розподілі тестових балів з урахуванням демографічних та регіональних характеристик випускників, а також визначення можливих залежностей між віком, статтю та рівнем їхніх навчальних досягнень.*

*Для опрацювання та аналізу даних використано сучасні бібліотеки мови R, зокрема dplyr, ggplot2, corrplot та caret, що забезпечило можливість проведення комплексного дослідження: від очищення й структуризації даних до побудови описової статистики, візуалізації розподілів та застосування методів кластеризації для виокремлення типових профілів учасників.*

*У процесі дослідження сформовано вибіркові групи за регіональними показниками, віковими інтервалами та статтю, а також визначено середні та граничні показники результатів у предметних блоках НМТ. Проведений аналіз дав змогу не лише окреслити загальні тенденції успішності, а й виявити особливості, характерні для окремих категорій респондентів.*

*Побудовані візуалізації та статистичні моделі дали змогу не лише окреслити загальні тенденції успішності випускників, а й виокремити типові профілі учасників із подібними характеристиками.*

*Отримані результати мають як наукову, так і практичну цінність та можуть бути використані для подальшого моніторингу якості освіти, оптимізації процесу підготовки учнів до зовнішніх оцінювань, формування більш адресних освітніх стратегій та вдосконалення підходів до оцінювання навчальних досягнень. Крім того, результати дослідження становлять інтерес для освітніх аналітиків, педагогів та розробників освітньої політики, оскільки дають змогу враховувати регіональні, соціально-демографічні та когнітивні особливості учнівської молоді при плануванні освітніх реформ.*

**Ключові слова:** національний мультипредметний тест, RStudio, освітня статистика, кореляційний аналіз, кластеризація, середнє значення вибірки, візуалізація даних.

## Yasinsky A. M., Solovei L. Ya., Kolupaev B. B., Ilkiv O. S. Analysis of the results of the National Multi-Subject Test using statistical modeling tools in RSTUDIO

*The article presents the results of an empirical study of the characteristics of the success of participants in the National Multi-Subject Test (NMT) using the RStudio software environment. The main goal of the work is to identify statistical patterns in the distribution of test scores, taking into account the demographic and regional characteristics of graduates, as well as to determine possible dependencies between age, gender and the level of their educational achievements.*

*Modern R language libraries were used to process and analyze data, in particular dplyr, ggplot2, corplot and caret, which provided the opportunity to conduct a comprehensive study: from cleaning and structuring data to constructing descriptive statistics, visualizing distributions and applying clustering methods to identify typical profiles of participants.*

*During the study, sample groups were formed by regional indicators, age intervals and gender, and the average and limit indicators of results in the subject blocks of the NMT were determined. The analysis allowed not only to outline general trends in success, but also to identify features characteristic of individual categories of respondents.*

*The constructed visualizations and statistical models made it possible not only to outline general trends in the success of graduates, but also to identify typical profiles of participants with similar characteristics.*

*The results obtained have both scientific and practical value and can be used for further monitoring the quality of education, optimizing the process of preparing students for external assessments, forming more targeted educational strategies and improving approaches to assessing academic achievements. In addition, the results of the study are of interest to educational analysts, teachers and educational policy makers, as they allow taking into account regional, socio-demographic and cognitive characteristics of students when planning educational reforms.*

**Keywords:** National Multi-Subject Test, RStudio, educational statistics, correlation analysis, clustering, sample mean, data visualization.

**Вступ.** Сучасна система освіти в Україні зазнає глибоких трансформацій, спрямованих на забезпечення її якості, прозорості та відповідності європейським стандартам. Запровадження воєнного стану в Україні зумовило необхідність створення спрощеної моделі проведення державного оцінювання випускників закладів середньої освіти. У 2022 році впроваджено Національний мультипредметний тест (далі – НМТ), який став адаптаційним інструментом у нових умовах. НМТ не лише виконує функцію перевірки рівня знань випускників, а й слугує важливим джерелом статистичних даних для комплексного аналізу освітніх процесів.

Український центр оцінювання якості освіти (далі – УЦОЯО) на своєму офіційному сайті публікує звіти, зокрема «Офіційний звіт про результати НМТ у 2023 році», а багатомні матеріали за всі роки проведення ЗНО/НМТ наявні у відкритому доступі на веб-ресурсі УЦОЯО [1, с. 1–22].

Незважаючи на відкритість і доступність статистики результатів НМТ, у науковому обігу спостерігається недостатня кількість глибоких емпіричних досліджень, що базуються на сучасних методах обробки й аналізу

освітніх даних. Зокрема, обмежено досліджується зв'язок між демографічними характеристиками (вік, стать), регіональними відмінностями та результатами тестування. У практиці освітньої аналітики все ще недостатньо використовують інструменти автоматизованого статистичного аналізу, як-от програмне середовище RStudio, що дає змогу здійснювати якісне, масштабоване та прозоре опрацювання освітніх даних [2, с. 204].

Крім того, з огляду на впровадження елементів цифрової трансформації освіти та орієнтацію на дані як основу для ухвалення управлінських рішень, постає потреба в розробці та апробації методик глибокого аналізу результатів ЗНО/НМТ. Зокрема, актуальними є такі завдання, як побудова профілів успішності, ідентифікація закономірностей у вибірках різних років народження, виявлення відмінностей між хлопцями та дівчатами, а також візуалізація освітніх даних за допомогою теплових карт, кореляційних матриць та регресійних моделей.

**Методи та методики дослідження.** Для дослідження обрано проведення статистичного аналізу результатів НМТ із застосуванням RStudio. Отримані результати не лише

важливі з наукового боку, а й становлять практичну цінність для керівників закладів освіти, аналітиків та педагогів.

Зокрема, результати національного мультипредметного тесту є важливим джерелом інформації для виявлення освітніх тенденцій, регіональних відмінностей, впливу соціальних чинників на навчальні досягнення. Таке дослідження дає змогу отримати нові висновки щодо чинників, які впливають на успішність випускників, сформулювати рекомендації для підготовки учнів, а також слугувати основою для подальших досліджень у галузі педагогіки, психології та освітньої статистики.

**Результати.** У джерелі [3, с. 55–114] подано виклад правил формулювання та перевірки статистичних гіпотез, розкрито підходи до визначення рівнів статистичної значущості й схарактеризовано основні статистичні методи.

Авторами статті [4, с. 49–56] розроблено методiku проведення аналізу результатів педагогічного дослідження в середовищі RStudio, а також продемонстровано основні методи обробки даних, візуалізації та інтерпретації отриманих результатів. Крім того, у роботі детально проаналізовано методи попередньої обробки даних, нормалізації, описової статистики, кореляційного та регресійного аналізу, тестування статистичних гіпотез

У процесі проведення НМТ в базах даних УЦОЯО зібрано 835 547 знеособлених записів (2022 р. – 234 104; 2023 р. – 288 935; 2024 р. – 312 508), що відображають результати масштабного освітнього оцінювання; наразі у відкритому доступі наявні дані за всі три роки його проведення [1, с. 1–22].

Групою науковців Харківської академії неперервної освіти здійснено ґрунтовний аналіз на основі даних Українського центру оцінювання якості освіти за такими показниками НМТ 2024: середній бал НМТ (за шкалою 100–200 балів); відсоток учнів, які отримали 0 балів, менш ніж 120 балів, від 160 балів і більше, від 180 балів і більше та інші параметри. Електронне видання містить порівняльний аналіз респондентів Харківської області та загальноукраїнських показників [5, с. 1–20].

У межах дослідження акцентовано увагу на застосуванні описового, порівняльного та візуального аналізу результатів НМТ за 2022–2024 роки. Для візуалізації даних вико-

ристано бібліотеку `ggplot2`, що забезпечило наочність інтерпретації результатів та сприяло виявленню прихованих закономірностей.

Усі дані, що зберігаються в `csv`-файлах УЦОЯО, попередньо вивчені та проаналізовані щодо постановки проблемних завдань та побудови моделей коду, що опрацьовуватиметься в середовищі RStudio. Попередньо опрацьовано дані, зокрема вилучено змінні, що не мали аналітичної цінності для поставлених завдань. У сформованих масивах збережено інформацію щодо регіону проживання учасників НМТ, статі, року народження, повного переліку навчальних дисциплін та отриманих результатів за 200-бальною шкалою. Використання статистичного середовища R забезпечує широкі можливості для обробки, аналізу та візуалізації великих обсягів освітніх даних. Мова програмування R постає як ефективний інструмент не лише для реалізації базових методів описової статистики, а й для застосування сучасних алгоритмів машинного навчання, зокрема класифікації, кластеризації, кореляційного аналізу та побудови прогнозних моделей.

Актуальні методичні підходи до роботи з великими даними висвітлено в посібнику Р. Є. Майбороди [6, с. 589].

Дослідження даних ми розпочали з аналізу результатів тестування з української мови. Цю навчальну дисципліну складають усі учасники НМТ. Завдання дослідження – встановити регіональну структуру результатів та подати їх у графічному вигляді. Після зчитування даних проведено обчислення середніх значень за регіоном проведення тестування. Результати всіх років об'єднано в один датафрейм. Використовуючи бібліотеку `ggplot2`, побудовано графік залежності середнього бала з української мови від регіону (Рис. 1).

Під час проектування дослідження зроблено припущення, що в умовах повномасштабної війни рівень знань випускників продемонструє тенденцію до зниження в динаміці за роками. Найвищі результати з української мови протягом трьох років воєнного періоду зафіксовано в місті Києві, а також у Львівській, Волинській та Херсонській областях.

У 2024 році тестування з математики проходили всі учасники НМТ (312 508 осіб), проте значна їх частина не пододала порогового бар'єра, унаслідок чого результати були



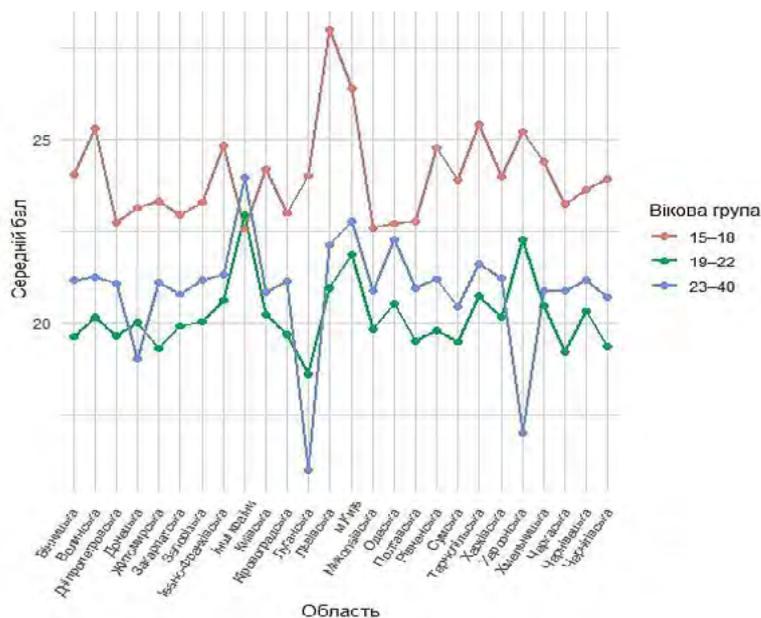
Для відстеження динаміки в результатах НМТ проведено обчислення, використовуючи результати 2023 та 2024 років. Структура розподілу результатів тестування з предмета «Історія України» у 2024 році однакова для обох статей, а от медіана дещо відрізняється і становить 142 бали для жінок і 141 бал для чоловіків. Значний позитивний ексцес для результатів з історії України у 2024 році вказує на те, що тест може потребувати перегляду рівня складності або балансування завдань, щоб уникнути занадто «гострих» крайніх значень.

Респонденти старшої групи віком 23–40 років продемонстрували дуже низький рівень підготовки з історії України в Донецькій, Луганській та Херсонській областях. Найкращі результати у вивченні історії України стабільно демонструють респонденти Львівської області та м. Київ.

**Висновки.** У статті розглянуто вплив таких факторів, як стать, вік, регіон проживання випускника на зміни в успішності результатів НМТ з окремих предметів.

Результати НМТ впродовж трьох років війни погіршуються, рівень знань майже синхронно знижується в усіх регіонах України. Найнижчі середні результати фіксуємо в Закарпатській, Дніпропетровській та Миколаївській областях.

Майже відсутня залежність у рівні вивчення математики залежно від статі. Унікальними щодо цього є Закарпатська область, де фіксуємо різницю в 5 балів



4. Джунь Й. В., Ясінський А. М., Соловей Л. Я. Опрацювання статистичних даних педагогічного дослідження засобами мови R. *Вісник Міжнародного економіко-гуманітарного університету імені академіка Степана Дем'янука. Серія: Педагогіка та психологія.* № 1/2025. Видавничий дім «Гельветика». 2025. С. 49–56.
5. Капустін І. В., Голтяй Т. М., Євтушенко С. С., Мілаш О. О. Аналіз результатів НМТ-2024 / за заг. ред. Л. Д. Покроєвої. Харків : Харківська академія неперервної освіти, 2025. 96 с. URL: <http://edu-post-diploma.kharkov.ua/?news=аналіз-результатів-нмт-2024-випускників> (дата звернення 13.08.2025).
6. Майборода Р. Є. Комп'ютерна статистика : підручник. Київ : ВПЦ «Київський університет», 2019. 589 с.

---

#### References

1. Ukrainyskiy tsentr otsiniuvannya yakosti osvity [Ukrainian Center for Education Quality Assessment]. (n. d.). [testportal.gov.ua](https://testportal.gov.ua) Retrieved from: [https://testportal.gov.ua/wp-content/uploads/2023/08/Pidsumky-provedennya\\_-NMT\\_2023-1.pdf](https://testportal.gov.ua/wp-content/uploads/2023/08/Pidsumky-provedennya_-NMT_2023-1.pdf) [in Ukrainian].
2. O. Ye., Kofanov, S. O., Solntsev, O. V., Zozulov. (2023). Prohramuvannya iz vykorystanniam R u statystychnykh ta marketynhovykh doslidzhenniakh [Programming using R in statistical and marketing research] KPI im. Ihoria Sikorskoho. Kyiv : KPI im. Ihoria Sikorskoho [in Ukrainian].
3. Moskalov, I. O., Lysenko, D. P. (2023). Zastosuvannya metodiv matematychnoi statystyky u psykholohopedahohichnykh doslidzhenniakh [Application of mathematical statistics methods in psychological and pedagogical research] Kyiv : NUOU [in Ukrainian].
4. Dzhun, Y. V., Yasynskiy, A. M., Solovei, L. Ia. (2025). Opratsiuvannya statystychnykh danykh pedahohichnoho doslidzhennia zasobamy movy R [Processing statistical data of pedagogical research using the R language] *Visnyk Mizhnarodnoho ekonomiko-humanitarnoho universytetu imeni akademika Stepana Demianchuka. Seriiia : Pedahohika ta psykholohiia.* №1/2025. Vydavnychiy dim «Helvetyka». (pp. 49–56). [in Ukrainian].
5. Kapustin, I. V. Analiz rezultativ NMT-2024. (2025). [Elektronne vydannia] / I. V. Kapustin, T. M. Holtiai, S. S. Yevtushenko, O. O. Milash; za zah. red. L. D. Pokroievoi. Kharkiv : Kharkivska akademiia neperervnoi osvity. 96 s. Retrieved from: <http://edu-post-diploma.kharkov.ua/?news=elektronne-vidannia-analiz-rezulta-15> [in Ukrainian].
6. Maiboroda, R. Ye. (2019). *Kompiuterna statystyka [Computer statistics]* K.: VPTs “Kyivskiy universytet” [in Ukrainian].

Стаття надійшла до редакції 11.11.2025

Стаття прийнята 29.11.2025

Статтю опубліковано 23.12.2025