

DOI: <https://doi.org/10.32782/2524-0072/2023-51-15>

УДК 330.47

ВИКОРИСТАННЯ СТАТИСТИЧНИХ ФУНКЦІЙ EXCEL В АНАЛІТИЧНИХ ДОСЛІДЖЕННЯХ ВЕЛИКИХ ДАНИХ

USE OF EXCEL STATISTICAL FUNCTIONS IN ANALYTICAL RESEARCH OF BIG DATA

Талах Валентин Івановичкандидат економічних наук, доцент,
Луцький національний технічний університет
ORCID: <https://orcid.org/0000-0002-2185-7957>**Талах Тетяна Анатоліївна**кандидат економічних наук, доцент,
Луцький національний технічний університет
ORCID: <https://orcid.org/0000-0002-1947-8114>**Talakh Valentyn, Talakh Tetiana**
Lutsk National Technical University

У статті розглянуто специфіку використання статистичних функцій Excel у проведенні аналітичних досліджень великих даних. Розглядаються основні статистичні функції Excel, такі як середнє арифметичне, медіана, дисперсія, стандартне відхилення та інші. Досліджено особливості використання цих функцій для аналізу великих даних у різних сферах та бізнес-процесах, включаючи фінанси, економіку, маркетингові дослідження тощо. Розглянуто методи зведення даних, фільтрації, сортування та використання таблиць змісту для швидкого та ефективного аналізу великих наборів даних. Досліджено переваги та обмеження використання Excel для аналітичних досліджень великих даних. Розглянуто альтернативні інструменти та програми для аналізу великих даних. На основі проведених досліджень визначено, що Excel є потужним інструментом для аналізу даних у різних галузях економічної аналітики, в тому числі – для роботи з Великими даними. Визначено обмеження програми, які можуть вплинути на точність та швидкість обробки великих обсягів даних. Рекомендовано враховувати такі обмеження в процесі аналізу та розглядати альтернативні засоби обробки великих даних залежно від конкретних потреб дослідження.

Ключові слова: економічний аналіз, функції Excel, Великі дані, статистика, аналітичні дослідження.

The article examines the specifics of using Excel's statistical functions in conducting analytical studies of big data. Basic Excel statistical functions such as arithmetic mean, median, variance, standard deviation, and others are covered. The specifics of using these functions for analyzing big data in various fields and business processes, including finance, economics, marketing research, etc., have been studied. Methods of data summarization, filtering, sorting and use of tables of contents for fast and efficient analysis of large data sets are considered. The advantages and limitations of using Excel for big data analytics are explored. Alternative tools and programs for big data analysis are considered. Based on the conducted research, it was determined that Excel is a powerful tool for data analysis in various areas of economic analytics, including for working with Big Data. Limitations of the program that can affect the accuracy and speed of processing large volumes of data have been identified. It is recommended to consider such limitations in the analysis process and consider alternative means of processing big data depending on the specific needs of the study. The use of Excel statistical functions in analytical studies of big data becomes especially relevant in modern economic analysis, where data volumes are constantly and dynamically growing and becoming extremely large in volume. At the same time, Big Data plays a key role in various fields, including scientific research, technology, business processes, medicine, and others. At the same time, in recent years, the pace of data growth has become a significant challenge for analysts and researchers in all fields. Processing and analysis of large volumes of data require powerful tools and specific analysis methods that would ensure accuracy, speed and efficiency of analytical processes. In this aspect, Excel, as one of the most common software products for data analysis, has powerful statistical functions that can be used for Big Data analytical research. However, it is known that Excel has certain limitations regarding the amount of data that can be processed, and there may also be a problem with the performance of the software application itself when working with large data sets.

Keywords: economic analysis, Excel functions, Big Data, statistics, analytical research.



Постановка проблеми. Використання статистичних функцій Excel в аналітичних дослідженнях великих даних набуває особливої актуальності в сучасному економічному аналізі, де обсяги даних постійно та динамічно зростають й стають надзвичайно великими за обсягом. При цьому Великі дані (Big Data) відіграють ключову роль у різних галузях, включаючи наукові дослідження, технології, бізнес-процеси, медицину та інші. Водночас, за останні роки, темпи зростання обсягу даних стали суттєвим викликом для аналітиків та дослідників в усіх сферах. Обробка та аналіз великих обсягів даних вимагають потужних інструментів та специфічних методів аналізу, які б забезпечували точність, швидкість та ефективність аналітичних процесів. В даному аспекті Excel, як один з найпоширеніших програмних продуктів для аналізу даних, володіє потужними статистичними функціями, котрі можуть використовуватися для аналітичних досліджень Big Data. Разом з тим, відомо, що Excel має певні обмеження щодо обсягу даних, які можуть бути оброблені, а також можуть виникати проблема з продуктивністю самого програмного додатку при роботі з великими наборами даних.

Відповідно, дослідження можливостей статистичних функцій Excel може забезпечити підвищення ефективності використання даної програми для аналізу великих даних. При цьому розуміння принципів використання статистичних функцій Excel для аналізу великих обсягів даних мають велике значення для фахівців, які займаються економічною аналітикою та статистичними дослідженнями. Причиною цього є те, що такі принципи дозволяють економити час на проведення аналізу, а також значно спрощують сам аналітичний процес та підвищують достовірність результатів. Отже, на основі зазначених вище аспектів можна стверджувати про актуальність та важливість практичних досліджень способів використання Excel в сучасному аналітичному середовищі з великими обсягами даних.

Аналіз останніх досліджень і публікацій.

Як свідчить практика, дослідження та аналіз великих даних набувають все більшої актуальності в останні роки. При цьому їх обробка, зокрема, за допомогою використання статистичного інструментарію Excel виступає у якості потужного засобу для підвищення ефективності аналізу великих даних.

Важливий внесок у вивчення специфіки використання Excel в аналітичних дослідженнях великих даних здійснено у праці С. Бега,

в якій розглянуто особливості використання різних статистичних функцій Excel, таких як середнє значення, медіана, стандартне відхилення, кореляція та регресія, а також продемонстровано їх практичне застосування для аналізу великих даних. Крім того, автор розглянув приклади формування коду і визначив поради щодо використання даних функцій [8].

Крім того, використання Excel для аналізу великих даних розглянуто у роботі Т. О. Шмітківської, в якій детально описано специфіку статистичних функцій Excel і визначено їх застосування у сфері бізнес-моделювання. Також автором розглянуто практичні аспекти роботи з великими даними, котрі передбачають ефективне використання формул, оптимізацію обчислень та візуалізацію одержаних результатів [7].

Додатковою варто звернути увагу на дослідження М. Х. Гардінга, який визначив особливості статистичних функцій Excel та їх застосування для аналізу даних. В дослідженні визначено специфіку та роль великих даних в процесі формування цифрової економіки, а також розглянуто специфічні підходи до їх аналізу та обробки за допомогою різноманітного аналітичного інструментарію [10].

Також вагомий внесок в дослідження специфіки використання статистичних функцій Excel у контексті аналітичних досліджень великих даних здійснено в роботі Ю. Чалюк, яка визначила сучасні специфічні аспекти використання статистичних функцій для аналізу великих даних на прикладі аналітики японської федерації бізнесу Кейданрен [4].

Загалом варто відзначити також дослідження, присвячені дослідженню особливостей використання статистичних функцій Excel, як потужного інструменту аналізу великих даних, викладені в роботах таких дослідників, як Л. В. Вербівська [1], І. О. Дейнега [9], М. І. Дзямучич [2], В. А. Никифорак [3], Н. М. Чиж [6] та багатьох інших. Втім, зростаюча потреба забезпечення ефективного аналізу великих обсягів даних вимагає поглиблення досліджень у сфері використання статистичних функцій Excel для практичної реалізації поставлених аналітиками завдань.

Постановка завдання. Метою статті є дослідження та аналіз специфіки застосування статистичних функцій програмного додатку Excel для реалізації аналітичних завдань по обробці великих даних для забезпечення ефективності економічних досліджень в умовах формування цифрової економіки.

Виклад основного матеріалу дослідження. Специфіка сучасного етапу розвитку інформаційних систем полягає у постійному зростанні обсягів даних, які надходять в аналітичні дата-центри і потребують обробки. При цьому потік даних триває цілодобово в режимі онлайн і охоплює глобальні дані з усіх ринків. Особливістю такого аналізу є те, що вплив на діяльність окремих локальних підприємств або на функціонування регіональних ринків в даний час мають не лише місцеві, але й глобальні економічні тенденції, що потребує врахування все більшого обсягу різноманітних факторів впливу. Таке зростання обсягів даних потребує вдосконалення аналітичного інструментарію їх обробки. Зокрема, впровадження цифрових та інформаційних технологій в систему економічної аналітики дозволяє суттєво збільшити продуктивність та результативність опрацювання великих даних, що підвищує достовірність та точність прогнозів, що розробляються на основі такого аналізу. При цьому різниця в продуктивності обробки даних за допомогою автоматизованих систем є досить відчутною у порівнянні з традиційними підходами (рис. 1).

Одним з засобів автоматизованої обробки великих даних наразі виступає спеціалізований програваний додаток Excel, котрий містить в собі певний набір спеціалізованого інструментарію, котрий може застосовуватися в економічній аналітиці. Зокрема, статистичні функції Excel виступають потужним засо-

бом для аналітичної обробки великих даних. До числа найбільш застосовуваних з них варто віднести наступні:

1. Сумування даних – такі статистичні функції, як SUM, AVERAGE, MIN, MAX та COUNT можуть бути використані для швидкого підрахунку загальної суми, середнього значення, мінімального та максимального значення, а також кількості елементів у великих наборах даних, що виступає базовою основою їх опрацювання. При цьому, як свідчить практика, незважаючи на свою простоту, дані функції найбільш часто застосовуються в процесі економічної аналітики стосовно обробки вхідних даних про різноманітні бізнес-процеси.

2. Фільтрація даних – комплекс вбудованих функцій Excel, таких як AVERAGE.IF, SUM.IF, COUNT.IF тощо цілком може бути використаний в процесі фільтрації визначених даних на основі певних умов аналітичних досліджень. Це дозволяє швидко виділити визначені частини великих даних для їх подальшого аналізу та деталізованої обробки відповідно до поставленої мети.

3. Розрахунок статистичних показників – Excel надає аналітикам достатньо широкий вибір спеціалізованих статистичних функцій, таких як STDEV, VARIANCE, CORRELATION, REGRESSION та інші. Дані функції дозволяють обчислювати стандартні відхилення, коваріацію, коефіцієнт кореляції, а також дають змогу проводити регресійний аналіз та

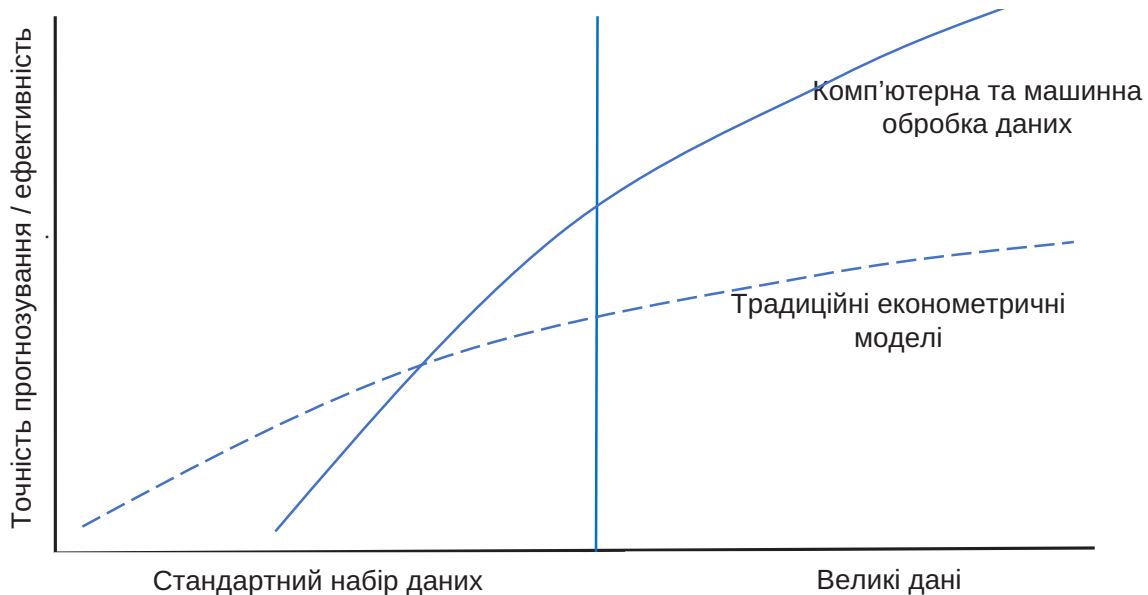


Рис. 1. Різниця обсягів обробки даних за допомогою традиційних та автоматизованих методів

Джерело: [10]

інші спеціалізовані статистичні розрахунки з великими обсягами даних.

4. Візуалізація даних – як відомо, Excel має ряд вбудованого функціоналу, необхідного для створення графіків та діаграм, котрі використовуються для візуалізації результатів аналізу великих даних. При цьому певні спеціалізовані графіки, діаграми розсіювання, гістограми тощо допомагають відображати особливості розподілу великих даних та наочно демонструють залежності між певними змінними.

5. Інтерактивні таблиці та звіти – Excel також має особливі вбудовані функції, які дозволяють створювати інтерактивні таблиці та спеціалізовані звіти для аналізу великих даних. Зокрема використання таких табличних функцій, як Сортування, Фільтрація, Групування та Сума даних, можна швидко формувати стислі аналітичні звіти за результатами обрахунків та здійснювати їх динамічний аналіз.

В даному аспекті варто особливо відзначити можливості, які надає для статистичної обробки великих даних функція STDEV (або STDEV.P, якщо дані представлені в повній вибірці). Зокрема, дана функція в Excel оптимізована для швидкого обчислення стандартного відхилення і використовує ефективні алгоритми та оптимізований код, що дозволяє обробляти великі обсяги даних без значного збільшення часу на обчислення. Функція STDEV має дві варіанти застосування – STDEV.S (або STDEV, STDEV.A) для розрахунку стандартного відхилення на основі вибірки та STDEV.P для розрахунку стандартного відхилення на основі повної групи даних. При цьому вибіркоче стандартне відхилення використовується, коли дані представлені як вибірка з загального обсягу, тоді як повне стандартне відхилення використовується, коли дані представлені повною групою без вибірки. Усе це дає змогу використовувати STDEV в Excel для обчислення стандартного відхилення навіть для великих обсягів даних. При цьому існує також можливість використання діапазонів даних, коли після визначення широкого діапазону функція автоматично обчислює стандартне відхилення для нього. Таким чином, застосування функції STDEV дозволяє здійснювати статистичний аналіз та порівняння даних, незалежно від їх обсягу.

Загалом важливо відзначити, що для обробки великих обсягів даних у програмному додатку Excel може знадобитися його спеці-

алізований додатковий функціонал, такий як використання таблиць зв'язків та фільтрація даних за допомогою Power Query, використання макросів тощо. Однак, як свідчить практика, вбудовані статистичні функції Excel наразі є відмінним початковим інструментом для обробки та аналізу великих даних, що дозволяє здійснювати широкий спектр статистичних обчислень та отримувати достовірні результати.

На даний час варто особливо відзначити переваги такого додатку Excel, як Power Query, який формує додаткові можливості для статистичного аналізу великих даних. Зокрема, до його специфічних особливостей обробки даних можна віднести наступні:

- Power Query дозволяє об'єднувати дані з різних джерел даних, таких як бази даних, CSV-файли, веб-сервіси тощо, що дозволяє збирати великі обсяги даних та обробляти їх як єдиний набір даних для подальшого аналізу;

- Power Query надає різноманітний практичний інструментарій для очищення даних шляхом видалення дублікатів, фільтрації дані за певними умовами, перейменування та перетворення стовпців, видалення анормальних значень та виконання інших операцій, призначених для забезпечення підготовки вибраних даних до статистичного аналізу;

- Power Query дозволяє застосовувати різноманітні перетворення та обчислення до даних шляхом додавання нових стовпців на основі існуючих, обчислення статистичних показників, проведення агрегації даних, застосування функцій до стовпців та рядків для отримання потрібної інформації для статистичного аналізу;

- Power Query може використовуватися для інтеграції з іншими спеціалізованими інструментами для статистичного аналізу, такими як Power Pivot або Power BI, що дає змогу створювати складні моделі даних для аналітичних обрахунків.

Загалом, можна стверджувати, що використання Excel у аналітичних дослідженнях великих даних дозволяє швидко та ефективно виконувати спеціалізовані розрахунки, фільтрацію, візуалізацію та аналіз даних на основі зручного інструментарію їх обробки. При цьому практична зручність та зрозумілість: Excel є його суттєвою перевагою, оскільки багато людей мають базові навички роботи з даною програмою, що полегшує її використання у дослідженнях великих даних [5]. Зокрема, до ключової переваги Excel в

процесі статистичного аналізу великих даних варто віднести вбудовані можливості фільтрації та сортування даних, за допомогою яких аналітики мають можливості по виявленню релевантних частин великого даних, що суттєво спрощує їх аналіз. Крім того, можливості інтеграції Excel з іншими інструментами економічної аналітики, такими як SQL-бази даних, Python, R тощо дозволяє широко використовувати їх для обробки та аналізу великих даних, навіть якщо вони виходять за межі можливостей самого Excel.

Висновки. Таким чином, приходимо до висновку, що використання статистичних функцій Excel в аналітичних дослідженнях великих даних має ряд переваг, які варто враховувати при використанні даного інструментарію. Зокрема, Excel дає змогу використовувати його для аналізу великих даних без потреби в додатковому навчанні персоналу або використанні складних спеціалізованих платних програм. Також Excel надає для аналітиків широкий набір спеціалізованих функцій та формул, які можна використовувати

для виконання статистичного аналізу вхідних даних шляхом використання функцій математичного моделювання, фінансових функцій тощо. Усе це дозволяє здійснювати різноманітні спеціалізовані аналітичні розрахунки у сфері обробки великих даних в межах одного програмного інструменту.

Разом з тим, варто зазначити, що Excel має свої функціональні обмеження у роботі саме з великими обсягами даних. Зокрема, програма може стикатися з обмеженнями продуктивності та масштабованості обробки, особливо при роботі з даними надзвичайно великих розмірів. Також, враховуючи постійне зростання складності та обсягів даних, для ефективного застосування Excel можуть знадобитися спеціалізовані інструменти та підходи, які виходять за межі його можливостей. Проте загалом можна стверджувати, що застосування статистичних функцій Excel у аналітичних дослідженнях великих даних на даний час виступає найбільш ефективним інструментом, який може спростити та полегшити процес обробки та аналізу даних.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. Вербівська Л. В. Концептуальні підходи щодо удосконалення діагностики мікросередовища в розробці стратегії підприємств. Особливості управління підприємствами в умовах глобальних викликів : монографія / за ред. О. В. Зибаревої. Чернівці : Технодрук, 2016. С. 98–113.
2. Дзямулич М. І., Шматковська Т. О., Борисюк О. В. Великі дані та їх роль у формуванні цифрової економіки. *Галицький економічний вісник*. 2021. Том 70. № 3. С. 16–21.
3. Никифорак В. А., Кобеля З. І., Вербівська Л. В. Організація виробництва : навч. посіб. Чернівці : Чернів. нац. ун-т, 2010. 407 с.
4. Чалюк Ю. О. Суспільство 5.0 у японській концепції кейданрен. *Mechanism of an Economic Regulation*. 2023. Вип. 1(99). С. 65–74.
5. Чалюк Ю. О. Цифрова конкурентоспроможність країн. *Інфраструктура ринку*. 2020. Вип. 50. С. 23–30.
6. Чиж Н. М., Дзямулич М. І., Урбан О. А. Управління інноваційною діяльністю банківських установ. *Економічний форум*. 2018. № 1. С. 284–289.
7. Шматковська Т. О., Стащук О. В., Дзямулич М. І. Великі дані та бізнес-моделювання економічних систем. *Ефективна економіка*. 2021. № 5.
8. Bag S., Wood L. C., Xu L., Dhamija P., Kayikci Y. Big data analytics as an operational excellence approach to enhance sustainable supply chain performance. *Resources, Conservation and Recycling*. 2020. Vol. 153. 104559.
9. Deineha, I., Maslov, A., Potryvaieva, N., Verbivska, L., Koliadych, O. Institutional Environment Tools for Small and Medium-Sized Enterprises Development. *Estudios de Economia Aplicada*, 2021, 39(3), 4798.
10. Harding M. H., Hersh J. Big Data in economics. *IZA World of Labor*, 2018. Vol. 451.

REFERENCES:

1. Verbivska, L. V. (2016) Kontseptualni pidkhody shchodo udoskonalennia diahnostryky mikroseredovishcha v rozrobtsi stratehii pidpriemstv. Osoblyvosti upravlinnia pidpriemstvamy v umovakh hlobalnykh vyklykiv [Conceptual approaches to improving diagnostics of the microenvironment in the development of enterprise strategy. Peculiarities of enterprise management in conditions of global challenges]. In O. V. Zybareva (Eds.). Chernivtsi: Technoprint. (in Ukrainian)
2. Dziamulych, M. I., Shmatkovska, T. O., & Borysiuk, O. V. (2021). Velyki dani ta yikh rol u formuvanni tsy-frovoi ekonomiky [Big data and its role in shaping the digital economy]. *Galytskyi ekonomichnyi visnyk*, vol. 70(3), pp. 16–21. (in Ukrainian)

3. Nykyforak, V. A., Kobelia, Z. I., & Verbivska, L. V. (2010). Organizatsia vyrobnytstva [Organization of production]. Chernivtsi: Chernivtsi national University. (in Ukrainian)
4. Chaliuk, Yu. O. (2023). Suspilstvo 5.0 u yaponskii kontseptsii keidanren [Society 5.0 in the Japanese concept of keidanren]. *Mechanism of an Economic Regulation*, vol. 1(99), pp. 65–74. (in Ukrainian)
5. Chaliuk, Yu. O. (2020). Tsyfrova konkurentospromozhnist kraiin [Digital competitiveness of countries]. *Infrastrukтура rynku*, vol. 50, pp. 23–30. (in Ukrainian)
6. Chyzh, N. M., Dziamulych, M. I., & Urban, O. A. (2018). Upravlinnia innovatsiinoiu diialnistiu bankivskykh ustanov [Management of innovative activities of banking institutions]. *Ekonomichnyi forum*, vol. 1, pp. 284–289. (in Ukrainian)
7. Shmatkovska, T. O., Stashchuk, O. V., & Dziamulych, M. I. (2021). Velyki dani ta bisnes-modeliuvannia ekonomichnykh system [Big data and business modeling of economic systems]. *Efektivna ekonomka*, vol. 5. (in Ukrainian)
8. Bag, S., Wood, L.C., Xu, L., Dhamija, P., & Kayikci, Y. (2020). Big data analytics as an operational excellence approach to enhance sustainable supply chain performance. *Resources, Conservation and Recycling*, vol. 153, 104559.
9. Deineha, I., Maslov, A., Potryvaieva, N., Verbivska, L., & Koliadych, O. (2021). Institutional Environment Tools for Small and Medium-Sized Enterprises Development. *Estudios de Economia Aplicada*, vol. 39(3), 4798.
10. Harding, M. H., & Hersh, J. (2018). Big Data in economics. *IZA World of Labor*, vol. 451.