

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДОНЕЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ВАСИЛЯ СТУСА
ЕКОНОМІЧНИЙ ФАКУЛЬТЕТ
КАФЕДРА МАРКЕТИНГУ ТА БІЗНЕС-АНАЛІТИКИ

І. Г. Гевлич

**ПАКЕТИ ПРИКЛАДНИХ ПРОГРАМ
У БІЗНЕС-АНАЛІТИЦІ**

Конспект лекцій

Вінниця
2024

УДК 005.52:338.2]:004.42(075.8)

П 13

*Рекомендовано до друку Вченою радою економічного факультету
Донецького національного університету імені Василя Стуса
(протокол № 4 від 25.10.2024)*

Укладач:

Гевлич І. Г., канд. техн. наук, доцент, доцент кафедри маркетингу та бізнес-аналітики Донецького національного університету імені Василя Стуса.

Рецензенти:

Фурман Т. Ю., канд. пед. наук, доцент, доцент кафедри маркетингу та бізнес-аналітики Донецького національного університету імені Василя Стуса;

Панченко І. В., канд. екон. наук, доцент, доцент кафедри менеджменту та поведінкової економіки Донецького національного університету імені Василя Стуса.

П 13 Пакети прикладних програм у бізнес-аналітиці: конспект лекцій / уклад. І. Г. Гевлич. Вінниця: ДонНУ імені Василя Стуса, 2024. 84 с.

Конспект лекцій з дисципліни «Пакети прикладних програм у бізнес-аналітиці» містить загальні положення щодо таких тем курсу: «Сучасні пакети прикладних програм та їх характеристика»; «Описова статистика та оцінка параметрів»; «Перевірка параметричних гіпотез»; «Регресійний та дисперсійний аналіз»; «Використання інструменту зведених таблиць пакету Microsoft Excel»; «Групування даних, підведення підсумків, консолідація даних у MS Excel»; «Візуалізація даних»; «Розв'язання оптимізаційних задач»; «Автоматизований статистичний аналіз». Їх розгляд дає змогу сформулювати результати навчання, зазначені у силабусі та освітньо-професійних програмах за спеціальністю 051 Економіка. Конспект лекцій побудовано на основі практико-орієнтованого підходу, рекомендовано для студентів ОС «Бакалавр» усіх форм навчання та викладачів.

УДК 005.52:338.2]:004.42(075.8)

© Гевлич І. Г., 2024

© ДонНУ імені Василя Стуса, 2024

ЗМІСТ

ВСТУП.....	5
Тема 1. Сучасні пакети прикладних програм та їх характеристика	7
1.1. Характеристика сучасних пакетів обробки інформації.....	7
1.2. ППП для вирішення фахових економічних завдань.....	9
1.3. Функціональний інструментарій ППП STATISTICA і EXCEL	13
ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ	16
Тема 2. Описова статистика та оцінка параметрів.....	18
2.1. Поняття генеральної сукупності і вибірки	18
2.2. Основні показники описової статистики	20
ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ	22
Тема 3. Перевірка параметричних гіпотез	23
3.1. Визначення термінів	23
3.2. Процес перевірки гіпотези	24
3.3. Параметричні і непараметричні тести.....	25
ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ	29
Тема 4. Регресійний та дисперсійний аналіз	31
4.1. Етапи регресійного аналізу	31
4.2. Припущення коректного застосування лінійної регресії	32
4.3. Процес проведення однофакторного дисперсійного аналізу	40
ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ	42
Тема 5. Використання інструменту зведених таблиць пакету Microsoft Excel.....	43
5.1. Використання зведених таблиць для аналізу даних у великих обсягах інформації	43
5.2. Ключові елементи зведеної таблиці в Excel	44
5.3. Робота зі зведеними таблицями	45
ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ	55
Тема 6. Групування даних, підведення підсумків, консолідація даних у MS Excel. Візуалізація даних	56
6.1. Групування та розгрупування даних у зведеній таблиці	56
6.2. Фільтрування даних у зведеній таблиці.....	57
6.3. Створення зведеної діаграми	61
ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ	63
Тема 7. Розв'язання оптимізаційних задач	64
7.1. Поняття оптимізаційної задачі.....	64

7.2. Розв’язання оптимізаційних задач за допомогою надбудови Microsoft Excel. Розв’язувач	65
7.3. Розв’язання оптимізаційних задач за допомогою VBA (Visual Basic for Applications)	68
ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ	75
Тема 8. Автоматизований статистичний аналіз	76
8.1. Аналіз із використанням мови програмування Basic	76
8.2. Аналіз із використанням VBA (Visual Basic for Applications)	78
ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ	79
СПИСОК ВИКОРИСТАНИХ ПОСИЛАНЬ	81

ВСТУП

Дисципліна «Пакети прикладних програм у бізнес-аналітиці» формує загальні, спеціальні компетентності та програмні результати навчання зі спеціальності 051 Економіка. Вона сприяє розвитку аналітичного мислення здобувачів освіти, формує вміння автоматизованого статистичного аналізу явищ і процесів в економіці і бізнесі, використання прикладних програм у статистичному аналізі.

Мета вивчення навчальної дисципліни полягає у формуванні теоретичних знань використання комп'ютерних технологій та пакетів прикладних програм статистичної обробки й аналізу даних, практичних навичок розв'язання економічних завдань за їх допомогою, створення моделей соціально-економічних явищ із погляду оптимізації процесів, зокрема:

- збирати, аналізувати та інтерпретувати дані статистичної інформації, використовувати її під час побудови економічних моделей;
- здійснювати пошук інформації відповідно до завдання, аналізувати її та представляти у вигляді, що дає змогу вирішувати економічні задачі;
- проводити статистичний аналіз соціально-економічних явищ та бізнес-процесів автоматизованим способом.

Вивчення навчальної дисципліни передбачає формування та розвиток у здобувачів вищої освіти компетентностей та програмних результатів навчання відповідно до освітньої програми «Бізнес-економіка».

Компетентності, які формуються у здобувача за результатами вивчення навчальної дисципліни:

Загальні компетентності (ЗК):

ЗК4. Здатність застосовувати знання у практичних ситуаціях.

ЗК7. Навички використання інформаційних і комунікаційних технологій.

ЗК8. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

Спеціальні компетентності (СК):

СК7. Здатність застосовувати комп'ютерні технології та програмне забезпечення з обробки даних для вирішення економічних завдань, аналізу інформації та підготовки аналітичних звітів.

СК9. Здатність прогнозувати на основі стандартних теоретичних та економічних моделей соціально-економічні процеси.

СК10. Здатність використовувати сучасні джерела економічної, соціальної, управлінської, облікової інформації для складання службових документів та аналітичних звітів.

СК12. Здатність самостійно виявляти проблеми економічного характеру при аналізі конкретних ситуацій, пропонувати способи їх вирішення.

СК15. Здатність використовувати пакети фахових прикладних програм для аналізу, моделювання та прогнозування соціально-економічних явищ та бізнес-процесів.

Результати навчання, які формує навчальна дисципліна:

РНД 1. Розуміти основні принципи та методи статистичної обробки та аналізу даних.

РНД 2. Вміти збирати, аналізувати та інтерпретувати статистичні дані, створювати моделі соціально-економічних явищ на їх основі автоматизованим способом.

РНД 3. Вміти використовувати пакети прикладних програм для збору, обробки та аналізу статистичної інформації.

РНД 4. Вміти представляти результати аналізу у наочному вигляді для вирішення економічних задач, критично оцінювати дані та результати аналізу, робити обґрунтовані висновки.

РНД 5. Вміти ідентифікувати та вирішувати проблеми, що виникають при статистичній обробці даних, ефективно комунікувати з зацікавленими особами щодо результатів аналізу та використання даних.

Результати навчання, які формує навчальна дисципліна за освітньою програмою:

ПРН 13. Ідентифікувати джерела та розуміти методологію визначення і методи отримання соціально-економічних даних, збирати та аналізувати необхідну інформацію, розраховувати економічні та соціальні показники.

ПРН 19. Використовувати інформаційні та комунікаційні технології для вирішення соціально-економічних завдань, підготовки та представлення аналітичних звітів.

ПРН 21. Вміти абстрактно мислити, застосовувати аналіз та синтез для виявлення ключових характеристик економічних систем різного рівня, а також особливостей поведінки їх суб'єктів.

ПРН 25. Визначати необхідні комп'ютерні програми та засоби візуальної аналітики для обробки великих масивів даних з метою виявлення нових закономірностей та тенденцій.

ТЕМА 1. СУЧАСНІ ПАКЕТИ ПРИКЛАДНИХ ПРОГРАМ ТА ЇХ ХАРАКТЕРИСТИКА

1.1. Характеристика сучасних пакетів обробки інформації.

1.2. ППП для вирішення фахових економічних завдань.

1.3. Функціональний інструментарій ППП STATISTICA і EXCEL.

1.1. Характеристика сучасних пакетів обробки інформації

Пакети прикладних програм (ППП) – це програмне забезпечення, призначене для виконання специфічних завдань користувачів. Вони забезпечують автоматизацію певних процесів і підвищують ефективність роботи у різних сферах діяльності. Для економістів такі пакети є незамінними інструментами для аналізу даних, побудови моделей, прогнозування, обробки фінансової інформації та звітності.

Розрізняють такі **типи ППП**: загального призначення, методо-орієнтовані, проблемно-орієнтовані, глобальних мереж, організації адміністрування обчислювального процесу (див. табл. 1.1).

Таблиця 1.1 – Класифікація ППП

Тип ППП	Приклади	Основні функції
1. Загального призначення		
1.1. Текстові редактори	Notepad, Sublime Text	Введення, редагування та збереження текстових даних для створення простих текстових документів
1.2. Текстові процесори	Microsoft Word, Google Docs	Форматування, редагування та створення комплексних документів з інтеграцією тексту, графіки та таблиць
1.3. Графічні редактори	– растрові редактори: Adobe Photoshop, GIMP; – векторні редактори: Adobe Illustrator, CorelDRAW; – 3D редактори: Blender, Autodesk 3ds Max	Створення, редагування та обробка графічних зображень у растровому, векторному та тривимірному форматах
1.4. Системи управління базами даних (СУБД)	MySQL, Microsoft Access	Створення, управління та маніпулювання даними в організованих табличних структурах
1.5. Електронні таблиці	Microsoft Excel, Google Sheets	Збереження, обробка та аналіз числових даних з використанням формул та функцій
1.6. Настільні видавничі системи	Adobe InDesign, QuarkXPress	Автоматизація процесу верстання поліграфічних видань з інтеграцією тексту та графіки
1.7. Експертні системи	IBM Watson, CLIPS	Аналіз даних у базах знань для надання рекомендацій та рішень на основі спеціалізованих знань

Тип ППП	Приклади	Основні функції
1.8. Інтегровані системи діловодства	Microsoft SharePoint, OpenText	Автоматизація робочих процесів керівників з управління документами, комунікацією та координацією підрозділів
1.9. Системи відеомонтажу	Adobe Premiere Pro, Final Cut Pro	Цифрова обробка відеоматеріалів, монтаж, створення відеоефектів та додавання звуку і титрів
1.10. Системи автоматизованого перекладу	Google Translate, SDL Trados Studio	Переклад тексту з однієї мови на іншу з використанням електронних словників та програм автоматичного перекладу
2. Методо-орієнтовані		
2.1. Математичне програмування	MATLAB, LINDO	Оптимізація задач з використанням лінійних, нелінійних та динамічних моделей для знаходження найкращих рішень у межах заданих обмежень
2.2. Мережеве планування і управління	Microsoft Project, Primavera	Планування та контроль проєктів за допомогою методів критичного шляху (CPM) та методів оцінки й перегляду програм (PERT) для оптимізації часу та ресурсів
2.3. Теорія масового обслуговування	AnyLogic, Arena	Моделювання та аналіз систем обслуговування, як-от черги, для оптимізації використання ресурсів та підвищення ефективності обслуговування
2.4. Математична статистика	R, SPSS	Аналіз даних за допомогою статистичних методів для виявлення закономірностей, прогнозування та прийняття рішень на основі статистичних моделей
3. Проблемно-орієнтовані		
3.1. Бухгалтерські системи	BAS, QuickBooks	Автоматизація бухгалтерського обліку та звітності, управління фінансовими операціями та підготовка податкових і фінансових звітів
3.2. Фінансові аналітичні системи	Tableau, SAP Financials	Аналіз фінансових даних, контроль та прогнозування фінансової ситуації, підтримка прийняття рішень у банківських та біржових структурах
3.3. Системи управління підприємством (ERP-системи)	SAP ERP, Oracle ERP Cloud	Інтеграція та управління всіма основними бізнес-процесами підприємства, включно з виробництвом, постачанням, фінансами, HR та ін.
3.4. Системи бізнес-аналітики (BI-системи)	Power BI, QlikView	Збір, обробка та аналіз даних для підтримки прийняття управлінських рішень та стратегічного планування
4. ППП глобальних мереж		
4.1. Засоби доступу та навігації	Google Chrome, Mozilla Firefox	Забезпечення доступу до вебресурсів і навігації в інтернеті, включно з переглядом вебсторінок та пошуком інформації
4.2. Електронна пошта	Microsoft Outlook, Gmail	Надсилання, отримання та управління електронними повідомленнями, включно з підтримкою вкладень і організацією поштових скриньок

Тип ППП	Приклади	Основні функції
4.3. Організація телеконференцій	Microsoft Teams, Zoom	Доступ та участь у групових дискусіях та обговореннях через електронні конференції та форуми
4.4. Редактори для створення вебсторінок	WordPress, Adobe Dreamweaver	Інструменти для розробки, редагування та публікації вебсторінок з інтеграцією тексту, графіки та мультимедійних елементів
5. ППП для організації адміністрування обчислювального процесу		
5.1. Управління серверними ресурсами	Microsoft System Center, VMware vSphere	Координація та контроль доступу до серверів, управління ресурсами й забезпечення безперебійної роботи серверної інфраструктури
5.2. Моніторинг та оптимізація роботи систем	Nagios, Zabbix	Відстеження продуктивності систем, виявлення проблем та застосування заходів для покращення ефективності і стабільності роботи ІТ-інфраструктури
5.3. Адміністрування мереж	Cisco Network Assistant, SolarWinds	Налаштування, управління та забезпечення безпеки комп'ютерних мереж, включно з локальними та глобальними мережевими структурами

1.2. ППП для вирішення фахових економічних завдань

У фаховій діяльності економістам доводиться вирішувати завдання:

- прогнозування економічних показників: ВВП, інфляції, безробіття, інших макроекономічних показників для різних періодів часу;
- оцінки впливу політик на економіку та зайнятість: фіскальної (зміни податків, державних витрат), монетарної (зміни ставок відсотка) тощо;
- аналізу ринкових даних: вивчення ринкової долі, вартості та попиту на продукти або послуги для стратегії ціноутворення та маркетингових рішень;
- вивчення споживчої поведінки: аналізу зв'язку між доходами, цінами та попитом на різні товари і послуги;
- економічного аналізу інвестицій: визначення рентабельності та ризиків інвестиційних можливостей, оцінки впливу різних факторів на дохідність;
- оцінки ефективності бізнес-стратегій: аналізу впливу різних стратегій розвитку підприємства на прибутковість та фінансовий стан;
- моделювання фінансових ризиків: вивчення залежностей між фінансовими інструментами, прогнозування коливань цін на акції, валюти, товари тощо;
- аналізу впливу соціальних факторів: вивчення впливу освіти, здоров'я, демографічних чинників на економічний зріст та розвиток;
- вивчення довіри та економічної поведінки: аналізу взаємозв'язків між рівнем довіри, споживчою поведінкою та ринковими динаміками;
- оцінки ефективності господарських політик: аналізу впливу різних економічних реформ, інвестиційних проєктів на економіку країни або регіону.

Характеристика ППП та мов програмування, залучених до вирішення цих та інших фахових завдань, представлена у табл. 1.2.

Таблиця 1.2 – Характеристика ППП та мов програмування, залучених до вирішення фахових завдань економістами

Програмні продукти / мови	Переваги використання	Недоліки використання
Microsoft Excel	Широкий набір вбудованих функцій для обробки даних, включно зі складними статистичними, математичними та фінансовими	Значна вартість
	Широка інтеграція та сумісність з іншими продуктами Microsoft Office та сторонніми додатками	Можливі проблеми з продуктивністю та стабільністю під час обробки великих наборів даних
	Потужні інструменти візуалізації результатів аналізу через створення графіків, діаграм, інших візуальних елементів	Обмежена повторюваність аналізу через відсутність вбудованих засобів для документування та автоматизації процесу аналізу
LibreOffice	Безкоштовність, доступність для всіх користувачів	Обмежена функціональність Libre Office Calc у порівнянні з Excel
	Відкритий вихідний код, що сприяє швидкому виправленню помилок та додаванню нових функцій	Недостатня продуктивність під час обробки великих наборів даних або виконання складних обчислень
	Сумісність з різними форматами, включно з Microsoft Excel (XLSX)	Макроси, створені в Excel на мові VBA, можуть не працювати коректно в LibreOffice Calc
EViews	Інтуїтивний інтерфейс і зручність використання	Значна вартість
	Широкий набір інструментів для аналізу часових рядів, включно з розширеними методами прогнозування та моделювання	Інтеграція з іншими програмами та пакетами статистичного аналізу може бути обмеженою, порівняно з R чи Python
	Спеціалізація для економетричних та статистичних аналізів	Обмежені можливості обробки великих наборів даних
Gretl	Безкоштовність, доступність для всіх користувачів	Обмежена офіційна документація та технічна підтримка, порівняно з комерційними продуктами
	Спеціалізація на економетричному аналізі, включно з широким набором статистичних методів та інструментів для аналізу часових рядів, аналізу панельних даних та регресійного аналізу	Відсутність або менша розвиненість просунутих функцій та інструментів, доступних у комерційних програмах, зокрема в EViews або Stata
	Інтеграція з іншими статистичними пакетами, як-от R, Octave, Stata і Python	Менша зручність (інтуїтивність) інтерфейсу, порівняно з комерційними програмами
Stata	Широкий набір інструментів для економетричного аналізу, включно з регресійним аналізом, аналізом панельних	Значна вартість

Програмні продукти / мови	Переваги використання	Недоліки використання
	даних, часовими рядами та багатовимірною статистикою	
	Інтуїтивний інтерфейс і зручність використання	Значна вартість базового пакету та додаткових інструментів (toolboxes)
	Активна спільнота користувачів, добре розвинена документація, включно з навчальними матеріалами, курсами, підтримкою від розробників	Обмеження під час роботи з великими наборами даних, порівняно зі спеціалізованими інструментами для великих даних
Matlab	Потужні обчислювальні інструменти, включно з інструментами для лінійної алгебри, статистики, чисельного аналізу та оптимізації	Інтеграція з іншими програмами та пакетами статистичного аналізу може бути обмеженою, порівняно з R чи Python
	Гнучкість та розширюваність – створення власних скриптів та функцій, велика кількість додаткових пакетів (toolboxes)	Складність навчання для користувачів, які не мають досвіду програмування або роботи з подібними інструментами
	Потужні засоби для візуалізації даних через створення графіків, діаграм, інтерактивну візуалізацію	Обмежена інтеграція з деякими іншими інструментами
Octave	Безкоштовність, доступність для всіх користувачів	Менш розвинені інструменти візуалізації, порівняно з MATLAB
	Висока сумісність з MATLAB, що дає змогу використання більшості скриптів і функцій MATLAB без значних змін	Менша продуктивність порівняно з MATLAB, під час виконання великих обчислювальних задач, обробки великих наборів даних
	Широкий набір інструментів для чисельного аналізу, статистики та лінійної алгебри	Менша кількість доступних додаткових пакетів (toolboxes), менше ресурсів для навчання і підтримки
Scilab	Безкоштовність, доступність для всіх користувачів	Обмежена інтеграція з іншими програмами та пакетами для статистичного аналізу, обмежена підтримка для імпорту та експорту даних, порівняно з іншими інструментами
	Широкий набір інструментів для чисельного аналізу, лінійної алгебри, статистики та оптимізації	Менша кількість доступних додаткових пакетів (toolboxes), менше ресурсів для навчання і підтримки
	Деяка сумісність з MATLAB, що дає змогу використання багатьох скриптів і функцій MATLAB з невеликими змінами або без змін	Менша зручність (інтуїтивність) інтерфейсу, порівняно з комерційними програмами
R	Безкоштовність, доступність для всіх користувачів	Складність навчання для користувачів, які не мають досвіду програмування або роботи з подібними інструментами
	Широкий набір інструментів (пакетів) для різноманітних статистичних аналізів, включно з регресійним аналізом, аналізом часових рядів, кластерним аналізом	Менша продуктивність, порівняно з MATLAB, за умови виконання великих обчислювальних задач, обробки великих наборів даних

Програмні продукти / мови	Переваги використання	Недоліки використання
	Велика та активна спільнота користувачів та розробників, яка вносить зміни, розвиває нові пакети, надає підтримку через форуми, блоги та інші ресурси	Відсутність єдиного стандартного інтерфейсу користувача, що створює різниці в якості інтерфейсу між різними пакетами та розробниками
Gauss	Спеціалізація на економетричному аналізі	Значна вартість
	Широкий набір вбудованих функцій та методів для виконання різних видів аналізу, включно з регресійним аналізом, аналізом часових рядів, панельних даних	Менша спільнота користувачів, менша кількість доступних ресурсів для навчання, підтримки та розвитку
	Висока швидкість обчислень	Відсутність безкоштовної альтернативи
Ox	Спеціалізація на економетричному аналізі	Значна вартість
	Широкий набір вбудованих функцій та методів для виконання різних видів аналізу, включно з регресійним аналізом, аналізом часових рядів, панельних даних	Менша кількість розширень та додаткових пакетів, що обмежує функціональність, порівняно з іншими програмами
	Інтуїтивний інтерфейс і зручність використання	Менша спільнота користувачів, менша кількість доступних ресурсів для навчання, підтримки та розвитку
Python	Широкий спектр бібліотек, як-от NumPy, Pandas, SciPy та StatsModels, які забезпечують інструменти для обробки та аналізу даних, включно з економетричним аналізом	Менша ефективність у виконанні обчислень, порівняно зі спеціалізованими мовами програмування, що може бути проблемою під час обробки великих обсягів даних
	Інтеграція з іншими інструментами та мовами програмування	Складність навчання для користувачів, які не мають досвіду програмування або роботи з подібними інструментами
	Велика та активна спільнота користувачів і розробників	Обмежена підтримка для економетричних методів, порівняно зі спеціалізованими інструментами

Узагальнюючи переваги та недоліки програмних засобів для вирішення фахових економічних завдань, можна зробити такі висновки:

1. Перевагами найбільш універсальних програмних засобів (Microsoft Excel та LibreOffice) є легкість використання, доступність, зручність для невеликих обсягів даних та простих аналізів, а недоліками – обмежені можливості для складних економетричних моделей, обробки великих обсягів даних та автоматизації процесів.

2. Програмні засоби для обробки даних (Matlab, Octave, Scilab) мають переваги – потужні математичні інструменти, широкі можливості для чисельного аналізу та моделювання, але і недоліки – високу вартість для Matlab, круту криву навчання для Octave та Scilab, менш розвинені інструменти візуалізації та меншу кількість доступних ресурсів, порівняно з R чи Python.

3. Спеціалізовані програмні засоби (EViews, Gretl, Stata, Gauss, Ox) мають переваги спеціалізації на економетриці, наявності великого набору статистичних та економетричних методів, але їх недоліками є висока вартість, обмежена універсальність, менша кількість ресурсів та інтеграцій з іншими інструментами, порівняно з R та Python.

4. Залучення для вирішення фахових економічних завдань мов програмування (R, Python) має перевагами безкоштовність, відкритий вихідний код, широкий вибір бібліотек для аналізу даних та економетричних методів, велику й активну спільноту користувачів та розробників. Водночас є і недоліки – нижча швидкість виконання, порівняно з компільованими мовами, та крута крива навчання для початківців.

1.3. Функціональний інструментарій ППП STATISTICA і EXCEL

STATISTICA – це потужний інструмент для аналізу даних, який пропонує широкий спектр можливостей для користувачів з різним рівнем підготовки та широко використовується в різних галузях діяльності.

Основні характеристики STATISTICA:

1. *Інтуїтивний інтерфейс* як можливість освоїти програму без глибоких знань у програмуванні, настроювані панелі інструментів та меню для оптимізації робочого процесу.

2. *Широкий спектр статистичних методів*: підтримка основних методів статистичного аналізу, включно з описовою статистикою, кореляційним аналізом, регресійним аналізом, аналізом дисперсії (ANOVA), факторним аналізом тощо; просунутими методами аналізу, як-от багатовимірний аналіз, аналіз виживання, кластерний аналіз та часо-рядний аналіз.

3. *Візуалізація даних*: різноманітні графічні можливості для представлення даних, включаючи гістограми, діаграми розсіювання, графіки часорядних даних, тривимірні графіки тощо; інтерактивні графіки, які дають змогу користувачам змінювати параметри та миттєво бачити результати.

4. *Обробка великих даних*: можливість роботи з великими наборами даних завдяки оптимізованим алгоритмам та підтримці багатопотоковості; інтеграція з базами даних та можливість імпорту даних з різних джерел, включно з текстовими файлами, електронними таблицями та реляційними базами даних.

5. *Аналіз даних та моделювання*: підтримка прогнозних моделей, як-от лінійна та нелінійна регресія, дерева рішень, нейронні мережі та методи машинного навчання; інструменти для перевірки гіпотез та оцінки ймовірностей.

6. *Автоматизація та програмування*: можливість створення макросів для автоматизації повторюваних завдань; вбудована мова програмування для розширеної обробки даних та створення користувацьких функцій.

7. *Інтеграція з іншими системами*: легка інтеграція з іншими програмними продуктами та платформами, як-от Microsoft Excel, SPSS, SAS тощо; підтримка різних форматів даних для експорту та імпорту, включно з CSV, XML та базами даних ODBC.

8. *Висока надійність та безпека*: постійні оновлення і технічна підтримка від виробника; надійні алгоритми та методи забезпечення конфіденційності та захисту даних.

STATISTICA використовується у таких галузях:

– *бізнес та фінанси* – для аналізу ринкових тенденцій, оцінки ризиків, розробки фінансових моделей та прийняття стратегічних рішень;

– *дослідження* – у медичних дослідженнях, біостатистиці, фармацевтиці та біотехнологіях для аналізу експериментальних даних і клінічних випробувань;

– *соціальні науки* – для аналізу соціальних опитувань, психологічних досліджень та вивчення демографічних даних;

– *освіта* – в ЗВО та дослідницьких центрах для навчання статистичних методів і аналізу дослідницьких даних.

Microsoft Excel – це потужний і багатофункціональний інструмент для роботи з електронними таблицями, широко використовується в різних галузях, включно з бізнесом, наукою, освітою та ін.

Основні характеристики Microsoft Excel:

1. *Інтуїтивний інтерфейс* зі вкладками та стрічкою інструментів; налаштовані панелі інструментів та можливість створення власних макросів.

2. *Робота з великими обсягами даних*: підтримка великих обсягів даних у електронних таблицях; використання функцій сортування та фільтрації для швидкого пошуку потрібної інформації.

3. *Формули та функції*: понад 400 вбудованих функцій для математичних, статистичних, логічних та текстових операцій; можливість створення користувацьких формул для спеціалізованих обчислень.

4. *Графічні можливості*: створення різноманітних графіків та діаграм для візуалізації даних, включно з лінійними, стовпчастими, круговими, гістограмами та іншими типами; інтерактивні графіки, що дають змогу користувачам змінювати параметри та миттєво бачити результати.

5. *Аналіз даних*: інструменти для аналізу даних, як-от зведені таблиці та зведені діаграми, що допомагають швидко агрегувати та узагальнювати великі набори даних; аналіз даних за допомогою додаткових інструментів, як-от аналіз «що, якщо», пошук рішень, аналіз даних тощо.

6. *Співпраця та спільна робота*: можливість спільної роботи в режимі реального часу через хмарні сервіси, як-от OneDrive та SharePoint; інструменти для коментування та відстеження змін, що спрощують співпрацю між користувачами.

7. *Інтеграція з іншими програмами*: легка інтеграція з іншими продуктами Microsoft Office, як-от Word, PowerPoint та Access; підтримка імпорту та експорту даних у різних форматах, включно з CSV, XML та базами даних ODBC.

8. *Автоматизація задач*: використання мови програмування VBA (Visual Basic for Applications) для автоматизації повторюваних завдань та створення користувачьких функцій і макросів; можливість запису макросів для автоматизації часто виконуваних завдань.

9. *Безпека та захист даних*: засоби захисту даних за допомогою паролів, шифрування та управління правами доступу; можливість налаштування різних рівнів доступу для різних користувачів.

Microsoft Excel **використовується** у таких галузях:

- *бізнес та фінанси* – для бюджетування, фінансового моделювання, аналізу витрат та доходів, складання звітності та управління проектами;
- *дослідження* – для аналізу експериментальних даних, обробки статистичних даних та моделювання наукових експериментів;
- *освіта* – у ЗВО для викладання математичних та статистичних дисциплін, обробки та аналізу дослідницьких даних;
- *логістика* – для відстеження запасів, планування поставчань, управління складом та аналізу ефективності логістичних операцій;
- *маркетинг* – для аналізу ринкових тенденцій, планування маркетингових кампаній, аналізу ефективності рекламних заходів та вивчення поведінки споживачів.

Порівняльна характеристика розглянутих ППП представлена у табл. 1.3.

Таблиця 1.3 – Порівняльна характеристика ефективності використання STATISTICA та Microsoft Excel для обробки інформації

Характеристика	STATISTICA	Microsoft Excel
<i>Основне призначення</i>	Статистичний аналіз даних, передбачення та моделювання	Робота з електронними таблицями, обчислення та базовий аналіз даних
<i>Інтерфейс</i>	Спеціалізований для статистичних досліджень	Загальний офісний інтерфейс, зручний для широкого кола користувачів
<i>Статистичні функції</i>	Широкий спектр вбудованих статистичних функцій	Основні статистичні функції

Характеристика	STATISTICA	Microsoft Excel
<i>Аналіз даних</i>	Поглиблений аналіз даних, включно з регресійним аналізом, факторним аналізом, кластерним аналізом	Базовий аналіз даних з використанням формул та зведених таблиць
<i>Моделювання</i>	Підтримка складного статистичного моделювання	Обмежені можливості для моделювання
<i>Графіки та візуалізація</i>	Високоякісна візуалізація даних з розширеними можливостями налаштування	Стандартні графіки та діаграми, легко налаштовуються
<i>Обробка великих наборів даних</i>	Оптимізований для обробки великих обсягів даних	Підтримує великі таблиці, але обробка може бути повільнішою
<i>Автоматизація</i>	Підтримка скриптів для автоматизації завдань	Підтримка макросів (VBA) для автоматизації
<i>Співпраця</i>	Менш розвинені інструменти для спільної роботи	Можливість спільної роботи через хмарні сервіси (OneDrive, SharePoint)
<i>Інтеграція</i>	Інтеграція з іншими статистичними програмами	Інтеграція з іншими продуктами Microsoft Office
<i>Вартість</i>	Зазвичай вища, ліцензійне програмне забезпечення	Доступний у складі Office 365 або окремо
<i>Навчання та підтримка</i>	Потребує спеціалізованого навчання	Багато доступних ресурсів для навчання
<i>Гнучкість використання</i>	Висока для статистичних задач	Висока для загальних офісних задач
<i>Цільова аудиторія</i>	Дослідники, аналітики, статистики	Загальні користувачі, бізнес-користувачі, аналітики

За результатами порівняння можна зробити висновок, що STATISTICA підходить для користувачів, які потребують поглибленого статистичного аналізу та моделювання і готові інвестувати час у навчання та адаптацію до спеціалізованого інтерфейсу, а Microsoft Excel є універсальним інструментом, зручним для широкого кола задач від базового аналізу до складних обчислень, з великим потенціалом для автоматизації та інтеграції з іншими офісними програмами.

ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ

1. Які основні типи пакетів прикладних програм (ППП) ви знаєте?
2. Чим текстові редактори відрізняються від текстових процесорів?
3. Чим растрові графічні редактори відрізняються від векторних?
4. Які функції виконують системи управління базами даних (СУБД)?
5. Для чого використовуються електронні таблиці?
6. Які завдання вирішують системи автоматизованого перекладу?
7. Які є основні відмінності між методо-орієнтованими та проблемно-орієнтованими PPP?
8. Які основні завдання можуть вирішувати PPP для фахових економічних завдань?

9. Які переваги та недоліки має використання Microsoft Excel для економічного аналізу?
10. Чим спеціалізовані програмні засоби відрізняються від універсальних?
11. Які основні можливості надає ППП STATISTICA?
12. Як функціональний інструментарій Excel може бути корисним для економістів?
13. Якими є обмеження мов програмування Python та R для економічного аналізу?
14. Які можливості пропонує Microsoft Excel для створення інтерактивних звітів?
15. У чому полягають переваги використання мови програмування R для економетричного аналізу?
16. Як системи управління проектами допомагають в організації економічної діяльності?
17. Які переваги для зберігання та аналізу економічних даних мають хмарні обчислення?
18. Які функції виконують аналітичні системи бізнес-аналітики (BI)?
19. Якими є основні відмінності між локальними та веборієнтованими ППП?
20. Які існують підходи до інтеграції різних ППП у єдину інформаційну систему?

ТЕМА 2. ОПИСОВА СТАТИСТИКА ТА ОЦІНКА ПАРАМЕТРІВ

2.1. *Поняття генеральної сукупності і вибірки.*

2.2. *Основні показники описової статистики.*

2.1. *Поняття генеральної сукупності і вибірки*

Описова статистика – це розділ статистики, який займається обробкою емпіричних даних, їх систематизацією, наочним представленням у вигляді графіків та таблиць, а також їх кількісним описом через основні статистичні показники.

Описова статистика використовується в різних галузях, включно з економікою, соціологією, медициною, маркетингом. Вона допомагає аналізувати дані, робити висновки та приймати обґрунтовані рішення на основі цих даних. Основна *мета* описової статистики полягає в тому, щоб надати зрозуміле та узагальнене уявлення про дані, використовуючи різноманітні методи та засоби, а основними поняттями, які дають змогу аналізувати дані та робити висновки, є *генеральна сукупність і вибірка*.

Генеральна сукупність – це повна сукупність усіх можливих спостережень або елементів, які цікавлять дослідника.

Характеристиками генеральної сукупності є:

1) *параметри* – статистичні величини, які описують генеральну сукупність, як-от середнє (μ), дисперсія (σ^2), стандартне відхилення (σ) тощо;

2) *розмір сукупності* – кількість елементів у генеральній сукупності (N).

Вибірка – це підмножина генеральної сукупності, яка використовується для аналізу і формулювання висновків про всю сукупність. Вибірка має бути *репрезентативною*, тобто відображати основні характеристики генеральної сукупності, щоб отримані результати були достовірними.

Характеристиками вибірки є:

1) *статистичні показники* – величини, які обчислюються на основі вибірки і використовуються для оцінки параметрів генеральної сукупності, як-от вибіркоче середнє (\bar{x}), вибіркова дисперсія (s^2), вибіркоче стандартне відхилення (s) тощо;

2) *розмір вибірки* – кількість елементів у вибірці (n).

Взаємозв'язок між генеральною сукупністю та вибіркою проявляється у таких аспектах:

1. *Розгляд вибірки як представника генеральної сукупності.*

Оскільки обстеження всієї генеральної сукупності часто є неможливим або неефективним через великі витрати часу і ресурсів, вибірка використовується як представник генеральної сукупності. Для забезпечення достовірності результатів вибірка повинна бути *випадковою та репрезентативною*.

2. Оцінка параметрів генеральної сукупності.

За допомогою описової статистики на основі вибірових даних обчислюються статистичні показники для оцінки параметрів генеральної сукупності, наприклад, вибірове середнє (\bar{x}) використовується для оцінки середнього значення генеральної сукупності (μ).

3. Визначення помилки вибірки.

Різниця між значеннями параметрів генеральної сукупності та їх оцінками на основі вибірки називається помилкою вибірки. Її виникнення може бути зумовлено випадковими коливаннями в процесі відбору вибірки.

4. Визначення довірчих інтервалів.

Для врахування невизначеності, пов'язаної з використанням вибірки, використовуються довірчі інтервали, які надають діапазон значень, у якому з певною ймовірністю знаходиться істинне значення параметра генеральної сукупності.

Основні поняття та методи описової статистики представлені у табл. 2.1.

Таблиця 2.1 – Основні поняття та методи описової статистики

Характеристики	Показники	Опис
1. Характеристики центральної тенденції	Середнє арифметичне	Середнє значення всіх даних
	Медіана	Центральне значення у впорядкованому ряді даних
	Мода	Найбільш часте значення в наборі даних
2. Характеристики розсіювання	Розмах	Різниця між максимальним і мінімальним значеннями
	Дисперсія	Середнє квадратичне відхилення значень від середнього
	Стандартне відхилення	Корінь квадратний з дисперсії
	Коефіцієнт варіації	Відношення стандартного відхилення до середнього, виражене у відсотках
3. Форма розподілу	Асиметрія	Міра асиметрії розподілу даних
	Ексцес	Міра «гостроти» або «площинності» розподілу даних
4. Графічне представлення даних	Гістограми	Графічне зображення розподілу частот
	«Ящики з вусами»	Графічне представлення розподілу, яке показує медіану, квартилі та екстремальні значення
	Кругові діаграми	Графічне представлення часток у вигляді кругового сектору
	Стовпчасті діаграми	Графічне представлення даних у вигляді стовпців різної висоти
5. Ряди розподілу	Частотні таблиці	Таблиці, що показують, як часто зустрічаються різні значення в наборі даних
	Кумулятивні частотні розподіли	Таблиці або графіки, що показують накопичену частоту значень

2.2. Основні показники описової статистики

Основні статистичні показники, які використовуються для опису набору даних:

1. Міри центральної тенденції:

а) *середнє значення* як характеристика положення для вибіркового розподілу, наприклад, середнє арифметичне:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{x_1 + x_2 + \dots + x_n}{n};$$

б) *медіана* як величина, що розташована в середині ряду величин, проранжованих у зростальному або спадному порядку. Медіана є квантилем порядку 1/2.

Медіаною функції розподілу F називається таке число \tilde{x} , що:

$$F(\tilde{x})=1/2, \text{ або } P(X<\tilde{x})=P(X>\tilde{x})=1/2,$$

тобто ймовірність того, що випадкова величина матиме значення більше або менше за медіану, однакова і дорівнює 1/2;

в) *мода* m_x як точка максимуму емпіричної функції щільності розподілу.

Як характеристику центру моду можна використовувати лише для розподілів із симетричною кривою щільності розподілу. В окремих випадках можуть спостерігатися багатомодальні розподіли. В реальних ситуаціях для достатньо великої кількості даних це зазвичай свідчить про неоднорідність вибірки, тобто досліджувана вибірка може розглядатися як суміш декількох однорідних вибірок. Існують також розподіли, зокрема рівномірний, які не мають моди.

Співвідношення між середнім арифметичним, медіаною та модою розподілу залежить від знака коефіцієнта асиметрії розподілу. Якщо він додатний, то $m_x < m_e < \bar{x}$, у протилежному випадку $\bar{x} < m_e < m_x$. Якщо ж коефіцієнт асиметрії дорівнює нулю, то ці три показники центру розподілу є рівними.

2. Міри мінливості:

а) *стандартне відхилення (чи дисперсія)* як середньоквадратичне відхилення від середнього значення. Інакше кажучи, це математичне сподівання квадрата відхилення цієї змінної від її очікуваного значення (її математичного сподівання). Отже, дисперсія є вимірюванням величини розпорошеності значень цієї змінної, беручи до уваги всі її значення і їх ймовірності або ваги;

б) *мінімальне та максимальне значення змінної*;

в) *розмах* як різниця між найбільшим та найменшим із сукупності числових значень; він є однією з найпростіших мір розсіяння (розкиду) набору числових значень, дає інформацію про ширину інтервалу, в якому зосереджений весь набір числових даних, геометрично – це ширина відрізка, в якому розташовуються всі значення.

г) *коефіцієнт ексцесу* як числова характеристика розподілу ймовірностей дійсної випадкової величини. Коефіцієнт ексцесу характеризує «крутість», тобто стрімкість підвищення кривої розподілу у порівнянні з нормальною кривою;

д) *коефіцієнт асиметрії* як числова характеристика розподілу ймовірностей дійсної випадкової величини.

Описова статистика використовує такі *методи представлення даних*: *табличне представлення, графічне представлення, розрахунок статистичних показників*, адже резюме про вибірку та спостереження, які були зроблені, можуть бути кількісними (резюмуюча статистика) та візуальними (прості графіки). Ці резюме можуть бути основою початкового опису даних, частиною більш обширного статистичного аналізу або самодостатніми результатами конкретного дослідження.

Аналіз даних під час цього може бути одновимірним чи двовимірним.

Одновимірний аналіз передбачає, що описують розподіл однієї змінної, включно з його центральною тенденцією (у тому числі середнім значенням, медіаною і модою) та дисперсією (включно з діапазоном і квантилями набору даних та заходами з розповсюдження, як-от дисперсія і стандартне відхилення). Форма розподілу може також бути описана за допомогою індексів, як-от дисперсія випадкової величини та стандартне відхилення. Особливості розподілу змінної також можуть бути зображеними у графічному або табличному форматі, включно з гістограмами і діаграмою «стовбур – листя».

Коли модель складається з більш ніж однієї змінної, описова статистика може бути використана для опису відносин між парами змінних (*двовимірний аналіз*). У цьому випадку описові статистичні дані включають в себе: перехресні таблиці та таблиці спряженості, графічне представлення за допомогою діаграми розсіювання, кількісні показники залежності, описи умовних розподілів.

Кількісні показники залежності включають коефіцієнти кореляції (наприклад, коефіцієнт Пірсона або Спірмана) і коваріації. Нахил у регресійному аналізі також відображає відношення між змінними. Нестандартизований нахил вказує на одиницю заміни змінної для критерію зміни змінних в одній системі вимірювання. Стандартизований схил вказує, що це зміна в стандартних одиницях вимірювання. Вкрай нерівномірні дані часто можна перетворити за допомогою логарифмів. Використання логарифмів робить графіки більш симетричними, і вони виглядають більш схожими на нормальний розподіл, завдяки чому їх легше інтерпретувати.

ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ

1. Що таке генеральна сукупність у контексті описової статистики?
2. Які основні характеристики має генеральна сукупність?
3. Як визначаються параметри генеральної сукупності?
4. Навіщо використовується вибірка у статистичних дослідженнях?
5. Які основні характеристики вибірки ви знаєте?
6. Як вибірка повинна бути відібрана, щоб бути репрезентативною?
7. Яке значення має випадковість під час відбору вибірки?
8. Що таке помилка вибірки і як вона виникає?
9. Що таке довірчий інтервал і як він використовується в оцінці параметрів генеральної сукупності?
10. Як можна оцінити середнє значення генеральної сукупності за допомогою вибіркових даних?
11. Що таке середнє арифметичне і як воно обчислюється?
12. Як визначається медіана і у яких випадках вона використовується?
13. Що таке мода і коли її доцільно використовувати для аналізу даних?
14. Які основні міри розсіювання ви знаєте і як вони обчислюються?
15. Що таке дисперсія і чому вона важлива у статистиці?
16. Як визначається стандартне відхилення і що воно показує?
17. Що таке коефіцієнт варіації і як його інтерпретувати?
18. Що таке асиметрія розподілу і як її визначити?
19. Що таке ексцес і як він характеризує розподіл даних?
20. Які методи графічного представлення даних ви знаєте і для чого вони використовуються?

ТЕМА 3. ПЕРЕВІРКА ПАРАМЕТРИЧНИХ ГІПОТЕЗ

3.1. Визначення термінів.

3.2. Процес перевірки гіпотези.

3.3. Параметричні і непараметричні тести.

3.1. Визначення термінів

Перевірка статистичних гіпотез – це клас базових задач у математичній статистиці, що полягають у перевірці статистичних гіпотез на основі даних спостереження за процесом, який моделюється за допомогою множини випадкових величин. Перевірка статистичних гіпотез є *методом статистичного висновування*.

Альтернативний метод перевірки статистичних гіпотез полягає у визначенні множини статистичних моделей, по одній для кожної гіпотези кандидата, після чого використовуються техніки відбору моделі, аби вибрати ту, яка підходить найбільше. Найбільш загальні техніки відбору моделей ґрунтуються на інформаційному критерії Акаїке або коефіцієнті Баєса.

Протилежністю такого аналізу вибірки може бути *розвідувальний аналіз* вибірки, який може не мати наперед визначених гіпотез.

Основні терміни, що використовуються під час перевірки гіпотез, представлені у табл. 3.1.

Таблиця 3.1 – Визначення термінів, які використовуються під час перевірки параметричних гіпотез

№	Терміни	Визначення
1	Гіпотеза	Припущення про вигляд або параметри розподілу певної сукупності, бувають основні та альтернативні
2	Нульова гіпотеза (H ₀)	Гіпотеза, яку перевіряють і щодо якої роблять висновок про прийняття або відхилення
3	Альтернативна гіпотеза (H ₁)	Твердження, яке приймається, якщо нульова гіпотеза відхиляється
4	Рівень значущості (α)	Ймовірність помилки першого роду, тобто ймовірність відхилення нульової гіпотези, коли вона є істинною
5	Критична область	Область значень статистики, за якої нульова гіпотеза відхиляється. Межі критичної області визначаються рівнем значущості
6	Статистика тесту	Числове значення, обчислене на основі вибірових даних, яке використовується для перевірки гіпотези
7	Критичне значення	Значення, яке розділяє критичну область та область прийняття нульової гіпотези. Залежить від рівня значущості та типу розподілу статистики тесту
8	p-значення	Ймовірність отримати значення статистики тесту, яке є щонайменше таким екстремальним, як спостережуване, за умови, що нульова гіпотеза є істинною
9	Потужність тесту	Ймовірність правильно відхилити нульову гіпотезу, коли альтернативна гіпотеза є істинною. Висока потужність тесту означає високу ймовірність виявити ефект

№	Терміни	Визначення
10	Помилка першого роду	Відхилення нульової гіпотези, коли вона є істинною. Ймовірність такої помилки дорівнює рівню значущості (α)
11	Помилка другого роду	Прийняття нульової гіпотези, коли вона є хибною. Ймовірність такої помилки позначається як β
12	Двосторонній тест	Тест, що перевіряє гіпотезу про те, що параметр відрізняється від певного значення в будь-яку сторону (наприклад, $H_0 : \mu = \mu_0$ проти $H_1 : \mu \neq \mu_0$)
13	Односторонній тест	Тест, що перевіряє гіпотезу про те, що параметр відрізняється в конкретну сторону (наприклад, $H_0 : \mu \leq \mu_0$ проти $H_1 : \mu > \mu_0$)

3.2. Процес перевірки гіпотези

Процес перевірки гіпотези у статистиці включає такі ключові етапи:

1. Формулювання нульової та альтернативної гіпотез

Нульова гіпотеза (H_0) – це гіпотеза, яку ми перевіряємо і щодо якої робимо висновок про прийняття або відхилення. Зазвичай вона формулюється як твердження про відсутність ефекту або різниці. Альтернативна гіпотеза (H_1) – це гіпотеза, яка приймається, якщо нульова гіпотеза відхиляється. Вона виражає наявність ефекту або різниці. Наприклад:

$H_0 : \mu = \mu_0$ (середнє значення генеральної сукупності дорівнює μ_0);

$H_1 : \mu \neq \mu_0$ (середнє значення генеральної сукупності не дорівнює μ_0).

2. Вибір рівня значущості (α)

Рівень значущості (α) – це ймовірність помилки першого роду, тобто ймовірність відхилення нульової гіпотези, коли вона є істинною. Найчастіше використовується значення $\alpha = 0,05$, що означає 5 % ризику відхилити правильну нульову гіпотезу.

3. Вибір відповідного статистичного тесту

Вибір тесту залежить від типу даних, розподілу і того, яка гіпотеза перевіряється, наприклад:

- *t-тест* – для порівняння середніх значень двох вибірок;
- *z-тест* – для перевірки середнього значення за відомої дисперсії;
- *F-тест* – для порівняння дисперсій двох вибірок;
- *χ^2 -тест* – для перевірки незалежності або однорідності категоріальних даних.

4. Збір даних та обчислення статистики тесту

Необхідно зібрати дані відповідно до дизайну експерименту або вибірки. Потім обчислюється значення статистики тесту на основі цих даних. Формули для обчислення залежать від вибраного тесту.

5. Визначення критичної області та критичного значення

На основі рівня значущості (α) визначається критична область та відповідне критичне значення для обраного статистичного тесту. Критична область – це область значень статистики тесту, за якої нульова гіпотеза відхиляється.

6. Порівняння статистики тесту з критичним значенням

Значення статистики тесту порівнюється з критичним значенням. Якщо статистика тесту потрапляє в критичну область, то нульова гіпотеза відхиляється.

7. Визначення р-значення

Альтернативний підхід до ухвалення рішення полягає у використанні р-значення (ймовірності отримати значення статистики тесту, яке є щонайменше таким екстремальним, як спостережуване, за умови, що нульова гіпотеза є істинною). Якщо р-значення менше рівня значущості (α), то нульова гіпотеза відхиляється.

8. Прийняття рішення

На основі порівняння статистики тесту з критичним значенням або р-значення з рівнем значущості приймається рішення про прийняття або відхилення нульової гіпотези.

9. Інтерпретація результатів

Важливо пояснити, що означає прийняття або відхилення нульової гіпотези з погляду практичних наслідків.

Ці етапи формують систематичний підхід до перевірки параметричних гіпотез у статистичному аналізі.

3.3. Параметричні і непараметричні тести

Параметричні тести – це статистичні тести, які роблять певні припущення щодо параметрів популяції, з якої була взята вибірка. Ці тести є більш потужними за своїми можливостями, порівняно з непараметричними тестами, але вимагають дотримання певних умов, як-от нормальний розподіл даних і однорідність дисперсій.

Приклади параметричних тестів:

1) t-тест використовується для порівняння середніх значень двох вибірок, щоб визначити, чи є статистично значуща різниця між ними. Існують кілька видів t-тестів:

– *одновибірковий t-тест* порівнює середнє значення вибірки з відомим середнім значенням популяції:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

де \bar{x} – вибіркоче середнє;

μ – середнє значення популяції;

s – стандартне відхилення вибірки;

n – розмір вибірки;

– *двовибірковий t-тест для незалежних вибірок* порівнює середні значення двох незалежних вибірок:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

де \bar{x}_1, \bar{x}_2 – вибіркові середні;

s_1^2, s_2^2 – вибіркові дисперсії;

n_1, n_2 – розміри вибірок;

– *парний t-тест* порівнює середні значення двох залежних вибірок (наприклад, до і після впливу):

$$t = \frac{\bar{d}}{s_d/\sqrt{n}}$$

де \bar{d} – середнє значення різниць між парними спостереженнями;

s_d – стандартне відхилення різниць;

n – кількість пар;

2) z-тест використовується для перевірки гіпотез щодо середнього значення за відомої дисперсії популяції. Він схожий на t-тест, але застосовується, коли розмір вибірки великий (зазвичай $n > 30$) і дисперсія популяції відома:

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

де \bar{x} – вибіркоче середнє;

μ – середнє значення популяції;

σ – стандартне відхилення популяції;

n – розмір вибірки;

3) f-тест (ANOVA) використовується для порівняння дисперсій двох або більше вибірок. Найбільш поширеною формою f-тесту є однофакторний дисперсійний аналіз (ANOVA), який використовується для порівняння середніх значень трьох або більше груп:

$$F = \frac{\text{Міжгрупова дисперсія}}{\text{Внутрішньогрупова дисперсія}},$$

де міжгрупова дисперсія визначається як різниця між середніми значеннями груп, а внутрішньогрупова дисперсія – як різниця всередині груп;

4) χ^2 -тест використовується для перевірки гіпотез щодо розподілу категоріальних даних:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i},$$

χ де O_i – спостережувані частоти;

E_i – очікувані частоти.

Найчастіше застосовується для перевірки незалежності та відповідності:

– *тест незалежності* χ^2 перевіряє, чи є дві категоріальні змінні незалежними;

– *тест відповідності* χ^2 перевіряє, чи відповідає спостережуваний розподіл теоретично очікуваному;

5) кореляційний аналіз використовується для визначення ступеня зв'язку між двома кількісними змінними. Найбільш поширеним є коефіцієнт кореляції Пірсона (r). Формула для коефіцієнта кореляції Пірсона:

$$r = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2 \sum(Y_i - \bar{Y})^2}},$$

де X_i, Y_i – значення змінних;

\bar{X}, \bar{Y} – їх середні значення;

6) регресійний аналіз використовується для моделювання залежності між однією залежною змінною і однією або більше незалежними змінними. Лінійна регресія є найбільш простою формою регресійного аналізу.

Модель лінійної регресії виглядає так:

$$Y = a + bX + \epsilon,$$

де Y – залежна змінна;

X – незалежна змінна;

a – перетин з віссю ординат;

b – коефіцієнт нахилу;

ϵ – випадкова похибка.

Непараметричні тести – це статистичні методи, які не роблять припущень щодо розподілу популяції, з якої взята вибірка. Вони використовуються, коли дані не відповідають вимогам параметричних тестів, наприклад, не мають нормального розподілу або мають різні дисперсії. Непараметричні тести є менш потужними, але більш гнучкими.

Приклади непараметричних тестів:

1) тест Манна–Уїтні (U-тест) використовується для порівняння двох незалежних вибірок, щоб визначити, чи відрізняються їх медіани. Це непараметричний аналог двовибіркового t-тесту.

Процедура тесту:

- а) об'єднати дві вибірки;
- б) впорядкувати всі значення за зростанням;
- в) присвоїти ранги всім значенням;
- г) обчислити суму рангів для кожної вибірки;
- д) використовувати отримані ранги для обчислення U-статистики;

2) тест Вілкоксона (парний тест) використовується для порівняння двох залежних вибірок. Це непараметричний аналог парного t-тесту.

Процедура тесту:

- а) обчислити різниці між парними спостереженнями;
- б) впорядкувати абсолютні значення різниць;
- в) присвоїти ранги абсолютним значенням різниць;
- г) обчислити суму рангів для позитивних і негативних різниць;
- д) використовувати отримані ранги для обчислення W-статистики;

3) критерій Крускала–Уоліса використовується для порівняння трьох або більше незалежних вибірок. Це непараметричний аналог однофакторного дисперсійного аналізу (ANOVA).

Процедура тесту:

- а) об'єднати всі вибірки;
- б) впорядкувати всі значення за зростанням;
- в) присвоїти ранги всім значенням;
- г) обчислити суму рангів для кожної вибірки;
- д) використовувати отримані ранги для обчислення H-статистики;

4) тест Фрідмана використовується для порівняння трьох або більше залежних вибірок. Це непараметричний аналог повторних вимірів ANOVA.

Процедура тесту:

- а) присвоїти ранги всім значенням у кожній вибірці;
- б) обчислити суму рангів для кожної вибірки;
- в) використовувати отримані ранги для обчислення F-статистики Фрідмана;

5) **тест К-серій Вальда–Вольфовіца** використовується для перевірки незалежності двох вибірок. Він оцінює, чи є дві вибірки з різних популяцій або з однієї.

Процедура тесту:

- а) об'єднати дві вибірки;
- б) впорядкувати всі значення за зростанням;
- в) присвоїти мітки кожному значенню залежно від вибірки;
- г) порахувати кількість серій для кожної вибірки;
- д) використовувати отримані серії для обчислення тестової статистики;

6) **тест Колмогорова–Смірнова** використовується для перевірки рівності двох розподілів або для перевірки відповідності вибірки теоретичному розподілу.

Процедура тесту:

- а) побудувати емпіричні функції розподілу для кожної вибірки;
- б) визначити максимальну відстань між емпіричними функціями розподілу;
- в) використовувати отриману відстань для обчислення К-статистики;

7) **тест знаків** використовується для перевірки медіани однієї вибірки або для порівняння медіан двох залежних вибірок.

Процедура тесту:

- а) обчислити знаки різниць між парними спостереженнями;
- б) порахувати кількість позитивних і негативних знаків;
- в) використовувати отримані знаки для обчислення тестової статистики.

ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ

1. Що таке гіпотеза в контексті статистичного аналізу?
2. Що таке рівень значущості (α) і яку роль він відіграє в перевірці гіпотез?
3. Як визначається статистика тесту і для чого вона використовується?
4. Що таке потужність тесту і чому вона важлива?
5. Чим відрізняється помилка першого роду від помилки другого роду?
6. У яких випадках використовується двосторонній тест, а у яких – односторонній?
7. Як правильно сформулювати нульову та альтернативну гіпотези?
8. Які фактори впливають на вибір рівня значущості (α)?
9. Як визначити критичну область та критичне значення для обраного тесту?
10. Який альтернативний підхід використовується для ухвалення рішення на основі р-значення?
11. Як інтерпретувати результати перевірки гіпотези з практичного погляду?

12. Чим параметричні тести відрізняються від непараметричних тестів і в яких випадках треба використовувати кожен з них?
13. Як обчислюється статистика для одновибіркового t-тесту?
14. Яка формула використовується для двовибіркового t-тесту для незалежних вибірок?
15. Як проводиться парний t-тест і в яких випадках він використовується?
16. Які особливості має z-тест і в яких ситуаціях він застосовується?
17. Поясніть, як використовується χ^2 -тест для перевірки гіпотез щодо розподілу категоріальних даних.
18. Що таке кореляційний аналіз і як обчислюється коефіцієнт кореляції Пірсона?
19. Як проводиться регресійний аналіз і що означають його основні параметри?
20. У яких випадках треба використовувати непараметричні тести замість параметричних?

ТЕМА 4. РЕГРЕСІЙНИЙ ТА ДИСПЕРСІЙНИЙ АНАЛІЗ

4.1. Етапи регресійного аналізу.

4.2. Припущення коректного застосування лінійної регресії.

4.3. Процес проведення однофакторного дисперсійного аналізу.

4.1. Етапи регресійного аналізу

Регресійний аналіз – це статистичний метод, що використовується для моделювання відносин між залежною змінною та однією або більше незалежними змінними.

Основні етапи проведення регресійного аналізу представлені такими:

1. Формулювання моделі

На цьому етапі визначаються змінні, які будуть включені до моделі. Залежна змінна (Y) є тією, яку ми намагаємось пояснити або передбачити, тоді як незалежні змінні (X) є тими, що використовуються для пояснення варіацій у залежній змінній. *Приклад:*

$$Y = a + bX + \epsilon,$$

де Y – залежна змінна;

X – незалежна змінна;

a – перетин з віссю ординат;

b – коефіцієнт нахилу;

ϵ – випадкова похибка.

2. Збір та підготовка даних

На цьому етапі збираються дані, необхідні для аналізу. Дані повинні бути перевірені на наявність пропущених значень, аномалій і в належний спосіб масштабовані та перетворені, якщо це необхідно.

3. Оцінка параметрів моделі

Для оцінки параметрів моделі (коефіцієнтів регресії) зазвичай використовується **метод найменших квадратів (МНК)**. Цей метод мінімізує суму квадратів різниць між спостережуваними значеннями залежної змінної та передбаченими значеннями.

4. Перевірка припущень моделі

Для коректного застосування лінійної регресії важливо перевірити виконання таких *припущень*:

– *лінійність*: відношення між незалежною та залежною змінними повинно бути лінійним;

– *незалежність*: спостереження повинні бути незалежними;

– *гомоскедастичність*: варіація залишків повинна бути постійною по всьому діапазону значень незалежної змінної;

- *нормальність залишків*: залишки повинні мати нормальний розподіл;
- *відсутність мультиколінеарності*: незалежні змінні не повинні бути висококорельованими.

5. Оцінка адекватності моделі

Для оцінки адекватності моделі використовуються такі статистичні *критерії*:

- *R-квадрат (R^2)*: частка варіації залежної змінної, яка пояснюється незалежними змінними;
- *f-тест*: тест на значущість моделі загалом;
- *t-тести для окремих коефіцієнтів*: тести на значущість окремих коефіцієнтів регресії.

6. Інтерпретація результатів

На цьому етапі інтерпретуються отримані коефіцієнти регресії та статистичні показники. Важливо зрозуміти, які змінні мають значний вплив на залежну змінну і наскільки добре модель описує дані.

Приклад: якщо b дорівнює 0,5, це означає, що зростання незалежної змінної на 1 одиницю призводить до зростання залежної змінної на 0,5 одиниці.

7. Прогнозування

Після оцінки моделі її можна використовувати для прогнозування значень залежної змінної на основі нових значень незалежних змінних.

8. Верифікація та валідація моделі

Останній етап включає перевірку моделі на нових даних або за допомогою методів перехресної валідації, щоб переконатися, що модель має достатню прогностичну здатність і не є перенавченою.

Ці етапи формують цілісний процес проведення регресійного аналізу, який дає змогу дослідникам аналізувати і моделювати залежності між змінними, робити висновки та прогнози на основі наявних даних.

4.2. Припущення коректного застосування лінійної регресії

Лінійна регресія є одним з найпоширеніших методів статистичного аналізу, який використовується для моделювання залежності між однією залежною змінною та однією або більше незалежними змінними.

Для коректного застосування лінійної регресії необхідно виконувати декілька ключових припущень. Якщо ці припущення порушені, результати регресійного аналізу можуть бути некоректними або вводити в оману.

Основні припущення, які повинні бути виконані, представлені у табл. 4.1.

Таблиця 4.1 – Основні припущення, які повинні бути виконані для коректного застосування лінійної регресії

Припущення	Опис	Пояснення	Перевірка
Лінійність	Відношення між незалежною змінною (або змінними) і залежною змінною є лінійним	Зміна залежної змінної може бути адекватно описана лінійною функцією незалежної змінної	Побудова діаграми розсіювання між незалежною та залежною змінними. Якщо точки розподілені вздовж прямої лінії, це свідчить про лінійність
Незалежність	Спостереження незалежні одне від одного	Значення одного спостереження не впливає на значення іншого	Перевірка незалежності спостережень за допомогою тесту Дурбіна–Уотсона для часових рядів, який перевіряє автокореляцію залишків
Гомоскедастичність	Варіація залишків є постійною для всіх значень незалежної змінної	Дисперсія залишків не змінюється у разі зміни значень незалежної змінної	Побудова графіка залишків проти передбачених значень. Якщо точки рівномірно розподілені навколо горизонтальної лінії, то гомоскедастичність наявна. У разі порушення може використовуватись тест Бреуша–Пагана
Нормальність залишків	Залишки мають нормальний розподіл	Розподіл залишків (різниць між спостереженими та передбаченими значеннями) є нормальним	Побудова гістограми або Q-Q графіка залишків; використання тесту Шапіро–Вілکا для перевірки нормальності
Відсутність мультиколінеарності	Незалежні змінні не повинні бути високорельованими між собою	Висока кореляція між незалежними змінними ускладнює визначення унікального впливу кожної змінної на залежну змінну	Розрахунок коефіцієнтів кореляції між незалежними змінними або використання фактора інфляції дисперсії (VIF – Variance Inflation Factor). Якщо VIF перевищує 10, це свідчить про мультиколінеарність

Приклади порушення припущень:

1. *Нелінійність*: якщо відношення між змінними не є лінійним, застосування лінійної регресії може привести до неточних передбачень. У таких випадках можна використовувати поліноміальну або іншу нелінійну регресію.

2. *Автокореляція*: у випадку часових рядів, якщо значення змінних корельовані між собою у часі, то моделі, що враховують часову структуру (наприклад, авторегресійні моделі), можуть бути більш доречними.

3. *Гетероскедастичність*: якщо варіація залишків змінюється, то можна використовувати методи робастної регресії або перетворення змінних для стабілізації варіації.

4. *Порушення нормальності*: у випадках серйозного відхилення від нормальності залишків можуть бути застосовані непараметричні методи або перетворення змінних.

Тож виконання розглянутих припущень є критично важливим для забезпечення достовірності та точності результатів лінійного регресійного аналізу. В разі порушення припущень треба використовувати відповідні коригуючі методи або альтернативні моделі, щоб забезпечити адекватність аналізу даних.

Побудовані лінійні регресійні моделі в економіці зустрічаються не так часто. Їх використовують як окремий випадок для зручності і наочності розгляду досліджуваного економічного явища чи процесу. Найчастіше зустрічаються моделі, які відображають економічні процеси у вигляді нелінійної залежності.

Під час оцінювання параметрів регресії, нелінійних за пояснюючими змінними, використовується метод заміни змінних, суть якого полягає в заміні нелінійних пояснюючих змінних новими лінійними змінними, внаслідок чого нелінійна регресія зводиться до лінійної.

Якщо нелінійні моделі внутрішньо лінійні, то вони за допомогою відповідних перетворень можуть бути приведені до лінійного вигляду (наприклад, логарифмуванням і заміною змінних).

До перетвореної інформації та нової регресії може бути застосований звичайний МНК.

Наведемо кілька прикладів реалізації регресії на мові VBA MS Excel:

1. Лінійна регресія

```
Sub Кнопка2_Клік() 'Лінійна регресія
  Dim a, b, x, y As Double
  Dim sum1, sum2, sum3, sum4, sum5 As Double
  Dim r, av_y, av_y1 As Double
  Dim n As Integer
  n = Range("I2").Value
  sum1 = 0
  sum2 = 0
  sum3 = 0
  sum4 = 0
  sum5 = 0
  For i = 3 To n + 2
    x = Cells(i, 2).Value
    y = Cells(i, 3).Value
    sum1 = sum1 + x
    sum2 = sum2 + y
    sum3 = sum3 + y * x
    sum4 = sum4 + (x * x)
  Next i
  av_x = sum1 / n
```

```

av_y = sum2 / n
b = (sum3 * n - sum2 * sum1) / (sum4 * n - sum1 * sum1)
a = (sum2 - b * sum1) / n
Cells(4, 9).Value = (a)
Cells(5, 9).Value = (b)
Cells(6, 9) = «y = « & Str(a) & « + « & Str(b) & « x»
sum1 = 0
For i = 3 To n + 2
    x = Cells(i, 2).Value
    y = a + b * x
    sum1 = sum1 + y
    Cells(i, 4).Value = y
Next i
av_y1 = sum1 / n
sum1 = 0
sum2 = 0
For i = 3 To n + 2
    Rem x = Cells(i, 2).Value
    y = Cells(i, 3).Value
    y1 = Cells(i, 4).Value
    sum1 = sum1 + (y - av_y) ^ 2
    sum2 = sum2 + (y1 - av_y1) ^ 2
Next i
r = Sqr(sum2) / Sqr(sum1)
Cells(7, 9).Value = r
End Sub

```

2. Показова регресія

```

Sub s_r() 'Показова регресія
    Dim a, b, x, y As Double
    Dim sum1, sum2, sum3, sum4, sum5, sum6 As Double
    Dim r, av_y, av_y1 As Double
    Dim n As Integer
    n = Range(«I2»).Value
    sum1 = 0
    sum2 = 0
    sum3 = 0
    sum4 = 0
    sum5 = 0
    sum6 = 0
    For i = 3 To n + 2
        x = Cells(i, 2).Value
        y = Cells(i, 3).Value
        sum5 = sum5 + x
        sum6 = sum6 + y
    
```

```

x = Log(x)
y = Log(y)
sum1 = sum1 + x
sum2 = sum2 + y
sum3 = sum3 + y * x
sum4 = sum4 + (x * x)
Next i
av_x = sum5 / n
av_y = sum6 / n
b = (sum3 * n - sum2 * sum1) / (sum4 * n - sum1 * sum1)
a = (sum2 - b * sum1) / n
a = Exp(a)
Cells(4, 7).Value = (a)
Cells(5, 7).Value = (b)
Cells(6, 7) = «y = « & Format(a, «0.###») & « x^» & Format(b, «0.###»)
sum1 = 0
For i = 3 To n + 2
  x = Cells(i, 2).Value
  y = a * (x) ^ (b)
  sum1 = sum1 + y
  Cells(i, 4).Value = y
Next i
av_y1 = sum1 / n
sum1 = 0
sum2 = 0
For i = 3 To n + 2
  Rem x = Cells(i, 2).Value
  y = Cells(i, 3).Value
  y1 = Cells(i, 4).Value
  sum1 = sum1 + (y - av_y) ^ 2
  sum2 = sum2 + (y1 - av_y1) ^ 2
Next i
r1 = Sqr(sum2) / Sqr(sum1)
Cells(7, 7).Value = r1
End Sub

```

3. Експоненційна регресія

```

Sub e_r() 'експоненційна регресія
  Dim a, b, x, y, lny As Double
  Dim sum1, sum2, sum3, sum4, sum5 As Double
  Dim r, av_y, av_y1 As Double
  Dim n As Integer
  n = Range(«I2»).Value
  sum1 = 0
  sum2 = 0

```

```

sum3 = 0
sum4 = 0
sum5 = 0
For i = 3 To n + 2
    x = Cells(i, 2).Value
    y = Cells(i, 3).Value
    lny = Log(y)
    sum1 = sum1 + x
    sum2 = sum2 + lny
    sum3 = sum3 + lny * x
    sum4 = sum4 + (x * x)
    sum5 = sum5 + y
Next i
av_y = sum5 / n
b = (sum3 * n - sum2 * sum1) / (sum4 * n - sum1 * sum1)
a = (sum2 - b * sum1) / n
a = Exp(a)
Cells(4, 9).Value = (a)
Cells(5, 9).Value = (b)
Cells(6, 9) = «y = « & Format(a, «0.###») & « * exp( « & Format(b, «0.###») & « * x )»
sum1 = 0
For i = 3 To n + 2
    x = Cells(i, 2).Value
    y = a * Exp(b * x)
    sum1 = sum1 + y
    Cells(i, 5).Value = y
Next i
av_y1 = sum1 / n
sum1 = 0
sum2 = 0
For i = 3 To n + 2
    y = Cells(i, 3).Value
    y1 = Cells(i, 5).Value
    sum1 = sum1 + (y - av_y) ^ 2
    sum2 = sum2 + (y1 - av_y1) ^ 2
Next i
r2 = Sqr(sum2) / Sqr(sum1)
Cells(7, 9).Value = r2
End Sub

```

4. Гіперболічна регресія

```

Sub Кнопка2_Клік() 'Гіперболічна регресія
    Dim a, b, x, y, xx As Double
    Dim sum1, sum2, sum3, sum4 As Double
    Dim r, av_y, av_y1 As Double

```

```

Dim n As Integer
n = Range(«I2»).Value
sum1 = 0
sum2 = 0
sum3 = 0
sum4 = 0
For i = 3 To n + 2
    x = Cells(i, 2).Value
    y = Cells(i, 3).Value
    xx = 1# / x 'заміна
    sum1 = sum1 + xx * y
    sum2 = sum2 + xx
    sum3 = sum3 + y
    sum4 = sum4 + (xx * xx)
Next i
b = (n * sum1 - sum2 * sum3) / (n * sum4 - sum2 * sum2)
a = (1 / n) * sum3 - (1 / n) * b * sum2
Cells(4, 9).Value = (a)
Cells(5, 9).Value = (b)
Cells(6, 9) = «y = « & Format(a, «0.###») & « + « & Format(b, «0.###») & «/x»
sum1=0
For i = 3 To n + 2
    x = Cells(i, 2).Value
    y = a + b / x
    Cells(i, 5).Value = y
    sum1 = sum1 + y
Next i
av_y1 = sum1 / n
sum1 = 0
sum2 = 0
For i = 3 To n + 2
    y = Cells(i, 3).Value
    y1 = Cells(i, 5).Value
    sum1 = sum1 + (y - av_y) ^ 2
    sum2 = sum2 + (y1 - av_y1) ^ 2
Next i
r2 = Sqr(sum2) / Sqr(sum1)
Cells(7, 9).Value = r2
End Sub

```

5. Степенева регресія

```

Sub Кнопка2_Клік()'Степенева регресія
Dim a, b, x, y, lny As Double
Dim sum1, sum2, sum3, sum4, sum5 As Double
Dim r, av_y, av_y1 As Double

```

```

Dim n As Integer
n = Range("I2").Value
sum1 = 0
sum2 = 0
sum3 = 0
sum4 = 0
sum5 = 0
For i = 3 To n + 2
    x = Cells(i, 2).Value
    y = Cells(i, 3).Value
    lny = Log(y)
    sum1 = sum1 + x * lny
    sum2 = sum2 + x
    sum3 = sum3 + lny
    sum4 = sum4 + (x * x)
    sum5 = sum5 + y
Next i
av_y = sum5 / n
b = (n * sum1 - sum2 * sum3) / (n * sum4 - sum2 * sum2)
a = (1 / n) * sum3 - (1 / n) * b * sum2
b = Exp(b)
a = Exp(a)
Cells(4, 9).Value = (a)
Cells(5, 9).Value = (b)
Cells(6, 7) = "y = " & Str(a) & " * " & Str(b) & " ^ x "
sum1 = 0
For i = 3 To n + 2
    x = Cells(i, 2).Value
    y = a * (b ^ (x))
    sum1 = sum1 + y
    Cells(i, 4).Value = y
Next i
av_y1 = sum1 / n
sum1 = 0
sum2 = 0
For i = 3 To n + 2
    y = Cells(i, 3).Value
    y1 = Cells(i, 4).Value
    sum1 = sum1 + (y - av_y) ^ 2
    sum2 = sum2 + (y1 - av_y1) ^ 2
Next i
r = Sqr(sum2) / Sqr(sum1)
Cells(7, 8).Value = r
End Sub

```

4.3. Процес проведення однофакторного дисперсійного аналізу

Однофакторний дисперсійний аналіз (ANOVA, Analysis of Variance) – це статистичний метод, який використовується для порівняння середніх значень трьох або більше груп.

Мета ANOVA – визначити, чи існують статистично значущі відмінності між середніми значеннями цих груп. ANOVA допомагає зрозуміти, чи є варіації у вибірках значущими або вони викликані випадковими відхиленнями.

Процес проведення однофакторного дисперсійного аналізу проходить такі етапи:

1. Формулювання гіпотез

У однофакторному ANOVA формулюються дві гіпотези:

Нульова гіпотеза (H_0): середні значення всіх груп рівні:

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k.$$

Альтернативна гіпотеза (H_1): принаймні одна пара середніх значень груп є різною:

$$H_1 : \text{Не всі } \mu_i \text{ рівні.}$$

2. Розрахунок загального середнього

Обчислюється загальне середнє всіх спостережень у всіх групах:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} X_{ij},$$

де n – загальна кількість спостережень;

k – кількість груп;

n_i – кількість спостережень в i -й групі;

X_{ij} – спостереження j -го елемента в i -й групі.

3. Розрахунок міжгрупової (факторної) дисперсії (SSB)

Міжгрупова дисперсія визначається як сума квадратів відхилень середніх значень груп від загального середнього, зважена на кількість спостережень у кожній групі:

$$SSB = \sum_{i=1}^k n_i (\bar{X}_i - \bar{X})^2,$$

де \bar{X}_i – середнє значення в i -й групі.

4. Розрахунок внутрішньогрупової (залишкової) дисперсії (SSW)

Внутрішньогрупова дисперсія визначається як сума квадратів відхилень спостережень у кожній групі від середнього значення цієї групи:

$$SSW = \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2.$$

5. Розрахунок загальної дисперсії (SST)

Загальна дисперсія є сумою міжгрупової та внутрішньогрупової дисперсій:

$$SST = SSB + SSW.$$

6. Обчислення середньоквадратичних значень

Середньоквадратичні значення обчислюються для міжгрупової (MSB) та внутрішньогрупової (MSW) дисперсій:

$$MSB = \frac{SSB}{k - 1},$$

$$MSW = \frac{SSW}{N - k}.$$

7. Розрахунок F-статистики

F-статистика обчислюється як відношення середньоквадратичних значень міжгрупової та внутрішньогрупової дисперсій:

$$F = \frac{MSB}{MSW}.$$

8. Визначення критичного значення та ухвалення рішення

На основі обчисленого значення F та рівня значущості (α) визначається критичне значення F з таблиці F-розподілу. Якщо обчислене значення F перевищує критичне значення, нульова гіпотеза відхиляється.

Приклад застосування однофакторного дисперсійного аналізу

Дослідження ефективності трьох різних методів навчання математики. Групи студентів навчаються за різними методами, і їхні результати оцінюються.

1. *Формулюються гіпотези:*

H₀: середні результати трьох груп рівні.

H₁: принаймні один метод навчання приводить до різних результатів.

2. *Збираються дані* про результати студентів у кожній групі.

3. *Розраховуються* загальне середнє, міжгрупова та внутрішньогрупова дисперсії.

4. Обчислюється F -статистика.

5. На основі критичного значення F ухвалюється рішення щодо відхилення або прийняття нульової гіпотези.

Тож однофакторний дисперсійний аналіз (ANOVA) є потужним інструментом для порівняння середніх значень трьох або більше груп. Він дає змогу визначити, чи є відмінності між групами статистично значущими, що важливо для прийняття обґрунтованих рішень у багатьох галузях дослідження.

ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ

1. Що таке регресійний аналіз і для чого він використовується?
2. Які основні етапи проведення регресійного аналізу ви знаєте?
3. Які змінні називаються залежними і незалежними у регресійному аналізі?
4. Як перевіряються дані на наявність пропущених значень і аномалій?
5. Що таке метод найменших квадратів (МНК)?
6. Які статистичні критерії використовуються для оцінки адекватності моделі?
7. Як інтерпретувати значення F -тесту у регресійному аналізі?
8. Що означає значення коефіцієнта регресії $b = 0.5$?
9. Що таке верифікація та валідація моделі?
10. Які основні припущення повинні бути виконані для коректного застосування лінійної регресії?
11. Як перевіряється припущення лінійності у регресійній моделі?
12. Що таке гомоскедастичність і як її перевірити?
13. Як перевірити нормальність розподілу залишків?
14. Що таке мультиколінеарність і як її виявити?
15. Що таке однофакторний дисперсійний аналіз (ANOVA) і для чого він використовується?
16. Які гіпотези формулюються у однофакторному ANOVA?
17. Як обчислюється загальне середнє у однофакторному ANOVA?
18. Що таке міжгрупова (факторна) дисперсія (SSB) і як вона розраховується?
19. Як обчислюється внутрішньогрупова (залишкова) дисперсія (SSW)?
20. Що таке загальна дисперсія (SST) і як вона визначається?

ТЕМА 5. ВИКОРИСТАННЯ ІНСТРУМЕНТУ ЗВЕДЕНИХ ТАБЛИЦЬ ПАКЕТУ MICROSOFT EXCEL

5.1. Використання зведених таблиць для аналізу даних у великих обсягах інформації.

5.2. Ключові елементи зведеної таблиці в Excel.

5.3. Робота зі зведеними таблицями.

5.1. Використання зведених таблиць для аналізу даних у великих обсягах інформації

Зведені таблиці – це потужний інструмент для обчислення, узагальнення та аналізу даних, який дає змогу переглядати порівняння, закономірності та тенденції у даних. Їх робота трохи відрізняється залежно від того, яку платформу ви використовуєте для запуску Excel: Інтернет, macOS або Офіс для iOS.

Використання зведених таблиць для аналізу даних у великих обсягах інформації надає такі **переваги**:

1. Зведення та узагальнення даних:

Зведені таблиці дають змогу об'єднувати та узагальнювати великі обсяги даних з різних джерел. Наприклад, швидко узагальнити інформацію компанії щодо продажів за місяцями, клієнтами або продуктами.

2. Автоматичне підсумування та розрахунки:

Зведені таблиці автоматично виконують підсумування, підрахунок кількості, знаходять середнє значення, мінімальні та максимальні значення та інші розрахунки без необхідності ручного введення формул.

3. Гнучкість в аналізі:

Зведені таблиці дають змогу легко змінювати параметри аналізу, наприклад, додавати або видаляти поля, змінювати рівні деталізації, перетягувати дані між рядками та стовпцями. Це забезпечує швидке адаптування аналізу під конкретні запити чи потреби.

4. Візуалізація та інтерпретація даних:

Використання зведених таблиць сприяє легшій інтерпретації даних через їх структурування. До того ж, на основі зведених таблиць можна швидко створювати зведені діаграми, які ще більше полегшують розуміння складних даних.

5. Фільтрація та сортування:

Зведені таблиці надають потужні інструменти для фільтрації даних за різними критеріями та їх сортування даних за зростанням чи спаданням, що допомагає швидко знаходити потрібну інформацію.

6. Ефективність та економія часу:

Зведені таблиці значно підвищують ефективність роботи з великими обсягами даних. Вони автоматизують багато ручних процесів, що зменшує ймовірність помилок та економить час.

7. Прийняття рішень на основі даних:

Зведені таблиці допомагають швидко отримати аналітичну інформацію, що є основою для прийняття обґрунтованих рішень. Наприклад, аналіз продажів за регіонами може допомогти виявити, де потрібно посилити маркетингові зусилля.

Отже, використання зведених таблиць є критично важливим для ефективного аналізу великих обсягів даних, оскільки вони спрощують процеси зведення, аналізу, візуалізації та інтерпретації інформації, що підтримує прийняття ефективних управлінських рішень.

5.2. Ключові елементи зведеної таблиці в Excel

Зведена таблиця складається з таких **ключових елементів**:

1. Джерело даних:

Зведена таблиця базується на певному наборі даних (діапазон комірок, таблиця або зовнішнє джерело даних).

2. Поля зведеної таблиці:

– *рядки*: поля, що перетягуються в цю область, утворюють заголовки рядків зведеної таблиці. Це дає змогу групувати дані за певними категоріями або критеріями (наприклад, категорії продуктів, регіони, дати);

– *стовпці*: поля, що перетягуються в цю область, формують заголовки стовпців зведеної таблиці. Це допомагає розподіляти дані за різними категоріями в стовпцях (наприклад, продажі за кожен місяць або квартал);

– *значення*: поля, що перетягуються в цю область, відображаються як числові дані, які можна підсумовувати, підраховувати, обчислювати середнє значення тощо (наприклад, сума продажів, кількість товарів, середній дохід);

– *фільтри*: поля, що перетягуються в цю область, дають змогу застосовувати фільтри до всієї зведеної таблиці. Це дає можливість аналізувати дані тільки за певними критеріями (наприклад, фільтрування за певним роком або певною категорією продуктів).

3. Робота з полями:

– *перетягування полів*: щоб створити зведену таблицю, користувач перетягує відповідні поля з панелі полів зведеної таблиці до області рядків, стовпців, значень та фільтрів. Це дає змогу налаштовувати вигляд та структуру таблиці відповідно до потреб аналізу;

– *налаштування полів*: користувач може змінювати налаштування полів (наприклад, сортування та фільтрацію, обирати типи підсумків (сума, середнє значення, мінімум, максимум тощо)), а також налаштовувати формат відображення даних.

4. Динамічність та інтерактивність:

– *змінюваність*: зведена таблиця дає змогу легко змінювати структуру даних, користувач може переміщувати поля між різними областями, додавати або видаляти поля, що миттєво змінює вигляд та зміст таблиці;

– *фільтрація*: можна застосовувати різноманітні фільтри, щоб зосередитись на конкретних даних (наприклад, фільтрування за конкретним періодом або певною категорією продуктів);

– *розгортання та згортання даних*: користувач може розгортати або згорнути деталі певних груп даних для детальнішого аналізу або для більш компактного вигляду таблиці.

5. Візуалізація даних:

На основі зведених таблиць можна швидко створювати зведені діаграми, які допомагають візуалізувати дані та роблять аналіз більш наочним: стовпчасті діаграми, графіки, кругові діаграми тощо.

6. Приклади взаємодії елементів:

Наприклад, якщо у вас є дані про продажі за різними регіонами та періодами, ви можете перетягнути поле «Регіон» у область рядків, поле «Місяць» у область стовпців, а поле «Сума продажів» у область значень. Це створить таблицю, яка покаже суму продажів для кожного регіону за кожен місяць. Додатково можна застосувати фільтр за роком, щоб аналізувати дані тільки за обраний рік.

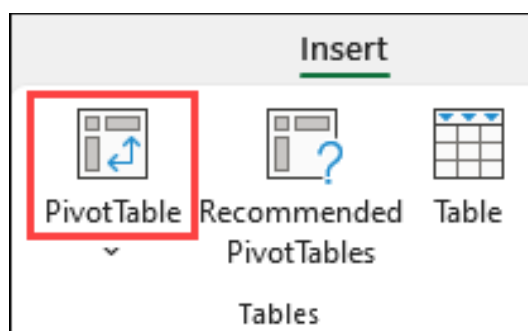
5.3. Робота зі зведеними таблицями

Створення зведеної таблиці

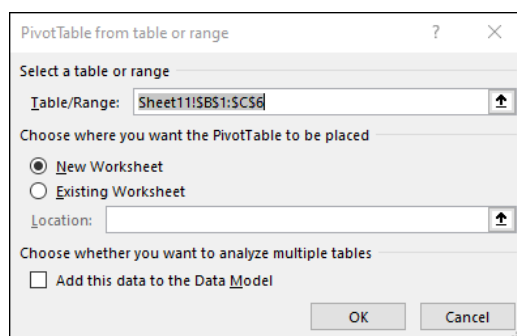
1. Виберіть клітинки, з яких потрібно створити зведену таблицю.

Примітка: ваші дані мають бути організовані в стовпці з одним рядком заголовка.

2. Виберіть «Вставити» > «Зведена таблиця».



3. Це створює зведену таблицю на основі наявної таблиці або діапазону.



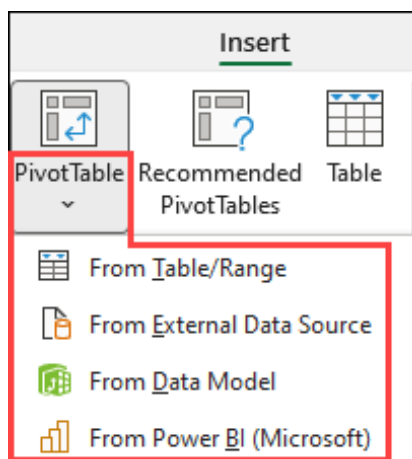
Примітка: вибір «Додати ці дані до моделі даних» додає таблицю або діапазон, який використовується для цієї зведеної таблиці, до моделі даних книги.

4. Виберіть місце для розміщення звіту зведеної таблиці. Виберіть «**Новий робочий аркуш**», щоб розмістити зведену таблицю на новому аркуші або на наявному аркуші, виберіть, де має відобразитися нова зведена таблиця.

5. Виберіть **ОК**.

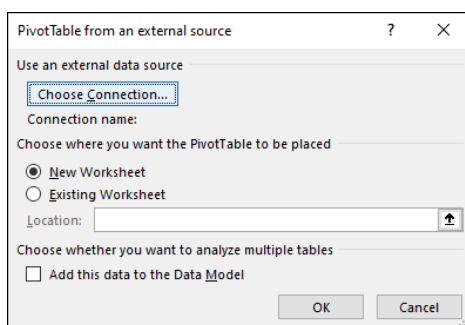
Зведені дані з інших джерел

Натиснувши стрілку вниз на кнопці, ви можете вибрати інші можливі джерела для своєї зведеної таблиці. Окрім використання наявної таблиці чи діапазону, є три інші джерела, з яких ви можете вибрати одне для заповнення своєї зведеної таблиці.



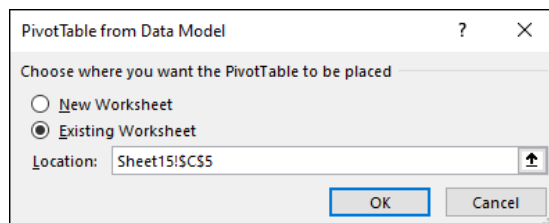
Примітка: залежно від IT-налаштувань вашої організації ви можете побачити назву вашої організації в списку. Наприклад, «З Power BI (Microsoft)».

Отримати із зовнішнього джерела даних



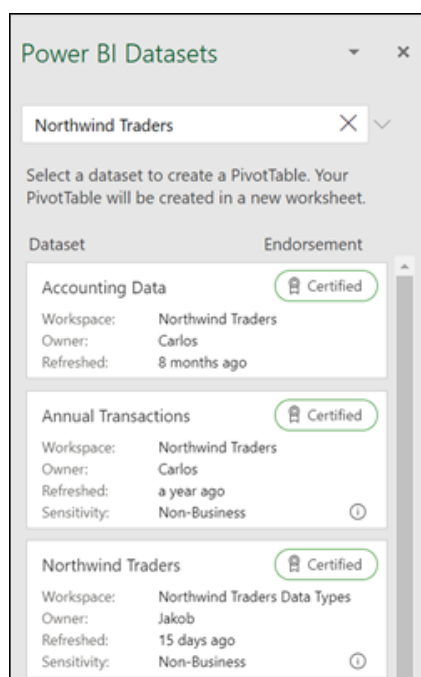
Отримати з моделі даних

Використовуйте цей параметр, якщо ваша робоча книга містить модель даних і ви хочете створити зведену таблицю з кількох таблиць, розширити зведену таблицю за допомогою спеціальних показників або працювати з дуже великими наборами даних.



Отримати з Power BI

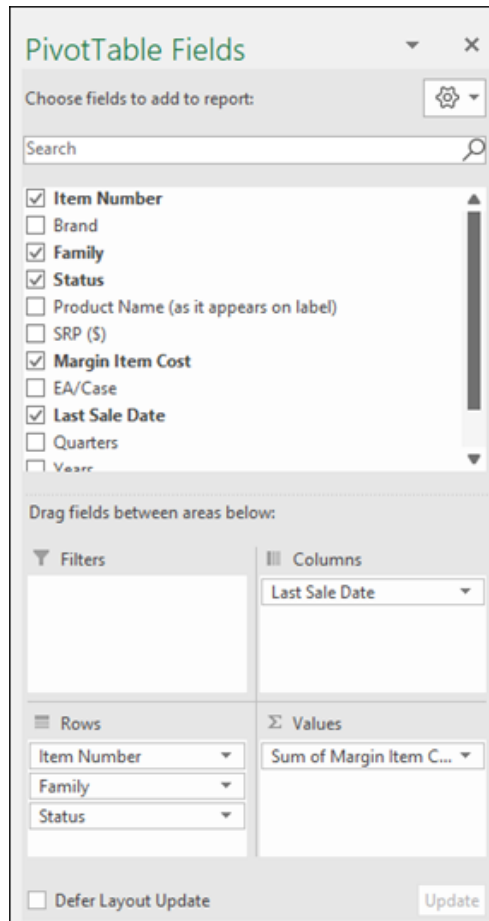
Використовуйте цей параметр, якщо ваша організація використовує Power BI і ви хочете знайти та підключитися до схвалених хмарних наборів даних, до яких у вас є доступ.



Створення вашої зведеної таблиці

1. Щоб додати поле до зведеної таблиці, установіть прапорець біля імені поля на панелі **полів зведених таблиць**.

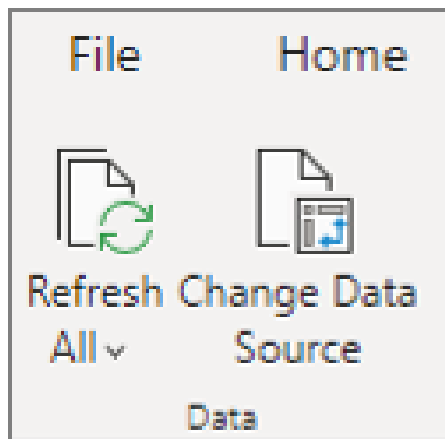
Примітка: вибрані поля додаються до областей за замовчуванням: нечислові поля додаються до **рядків**, ієрархії дат і часу додаються до **стовпців**, а числові поля додаються до **значень**.



2. Щоб перемістити поле з однієї області в іншу, перетягніть поле до цільової області.

Оновлення зведених таблиць

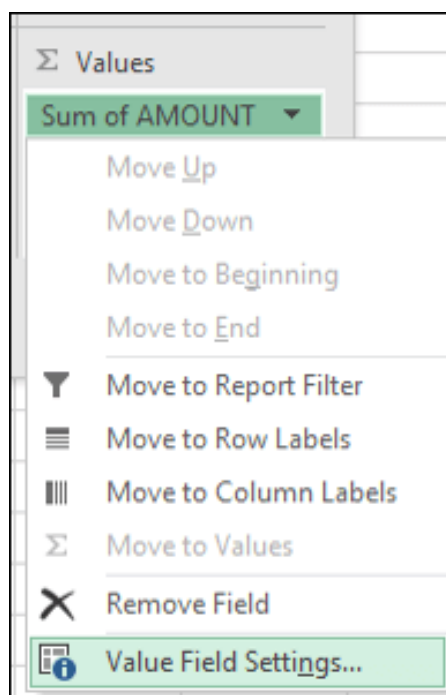
Якщо ви додаєте нові дані до джерела даних зведеної таблиці, будь-які зведені таблиці, створені на основі цього джерела даних, потрібно оновити. Щоб оновити лише одну зведену таблицю, клацніть правою кнопкою миші будь-де в діапазоні зведеної таблиці та виберіть **«Оновити»**. Якщо у вас є кілька зведених таблиць, спочатку виберіть будь-яку клітинку в будь-якій зведеній таблиці, а потім на стрічці перейдіть до **пункту Аналіз зведеної таблиці** > виберіть стрілку під кнопкою **«Оновити»**, а потім виберіть **«Оновити все»**.



Робота зі значеннями зведеної таблиці

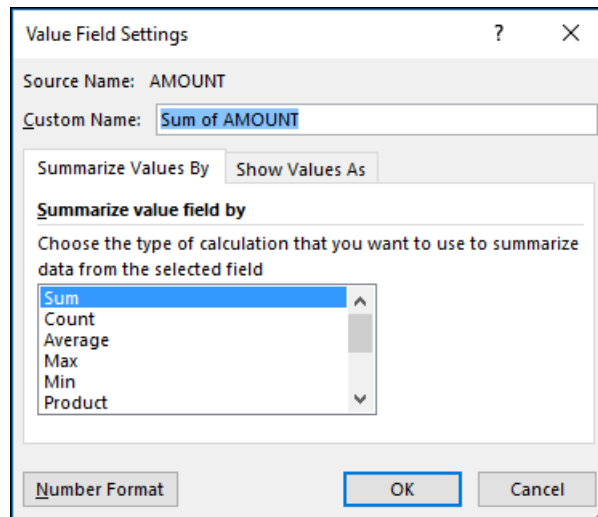
Узагальнити значення за...

За замовчуванням поля зведеної таблиці, розміщені в області **значень**, відображаються як **СУМА**. Якщо Excel інтерпретує ваші дані як текст, дані відображаються як **COUNT**. Ось чому так важливо переконатися, що ви не змішуєте типи даних для полів значень. Ви можете змінити обчислення за замовчуванням, спершу вибравши стрілку праворуч від назви поля, а потім виберіть параметр «**Параметри поля значення**».



Далі змініть обчислення в розділі «**Підсумувати значення за**». Зауважте, що коли ви змінюєте метод обчислення, Excel автоматично додає його в розділ «**Настроюване ім'я**», наприклад, «Сума імені поля», але ви можете змінити його. Якщо вибрати «**Формат числа**», можна змінити формат числа для всього поля.

Порада: оскільки зміна обчислень у розділі «Підсумувати значення за» змінює ім'я поля зведеної таблиці, краще не перейменовувати поля зведеної таблиці, доки ви не завершите налаштування зведеної таблиці. Одна хитрість полягає в тому, щоб скористатися «Знайти й замінити» (Ctrl+H) > «Знайти що» > «Сума», а потім «Замінити на» > залишити порожнім, щоб замінити все відразу замість повторного введення вручну.



Показати значення як...

Замість того, щоб використовувати обчислення для узагальнення даних, ви також можете відобразити їх у відсотках від поля. У наступному прикладі ми змінили суми витрат домогосподарства, щоб відображати їх як **% від загальної суми** замість суми значень.

AMOUNT	MONTH	January	February	March	Grand Total
Entertainment		5.10%	6.38%	6.13%	17.61%
Grocery		12.00%	12.25%	13.27%	37.52%
Household		8.93%	11.49%	10.21%	30.63%
Transportation		3.78%	5.87%	4.59%	14.24%
Grand Total		29.81%	35.99%	34.20%	100.00%

AMOUNT	MONTH	January	February	March	Grand Total
Entertainment		5.10%	6.38%	6.13%	17.61%
Grocery		12.00%	12.25%	13.27%	37.52%
Household		8.93%	11.49%	10.21%	30.63%
Transportation		3.78%	5.87%	4.59%	14.24%
Grand Total		29.81%	35.99%	34.20%	100.00%

Відкривши діалогове вікно «Параметри поля значень», ви можете зробити вибір на вкладці «Показати значення як».

Відображення значення як обчислення, так і у відсотках

Двічі перетягніть елемент у розділ «Значення», а потім установіть параметри «Підсумувати значення за» та «Показати значення як» для кожного з них.

Поради та підказки щодо формату даних:

- використовуйте чисті табличні дані для отримання найкращих результатів;
- упорядкуйте дані в стовпцях, а не в рядках;
- переконайтеся, що всі стовпці мають заголовки з одним рядком унікальних непорожніх міток для кожного стовпця, уникайте подвійних рядків заголовків або об'єднаних клітинок;

	A	B	C	D	E	F
1	Year	Category	Product	Sales	Rating	
2	2017	Components	Chains	\$20,000	75%	
3	2015	Clothing	Socks	\$ 3,700	22%	
4	2017	Clothing	Bib-Shorts	\$ 4,000	22%	
5	2015	Clothing	Shorts	\$13,300	56%	
6	2017	Clothing	Tights	\$36,000	100%	
7	2015	Components	Handlebars	\$ 2,300	35%	
8	2016	Clothing	Socks	\$ 2,300	28%	
9	2016	Components	Brakes	\$ 3,400	36%	
10	2016	Bikes	Mountain Bikes	\$ 6,300	40%	
11	2017	Components	Brakes	\$ 5,400	38%	
12	2016	Accessories	Helmets	\$17,000	90%	
13	2016	Accessories	Lights	\$21,600	90%	
14	2016	Accessories	Locks	\$29,800	90%	

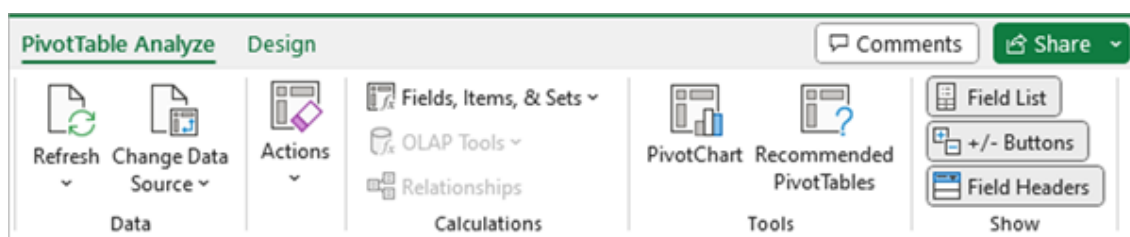
– відформатуйте дані як таблицю Excel (виберіть будь-де в даних, а потім виберіть «Вставити» > «Таблиця» на стрічці);

– якщо у вас є складні або вкладені дані, використовуйте Power Query, щоб перетворити їх (наприклад, щоб скасувати зведення даних), щоб вони були організовані в стовпці з одним рядком заголовка.

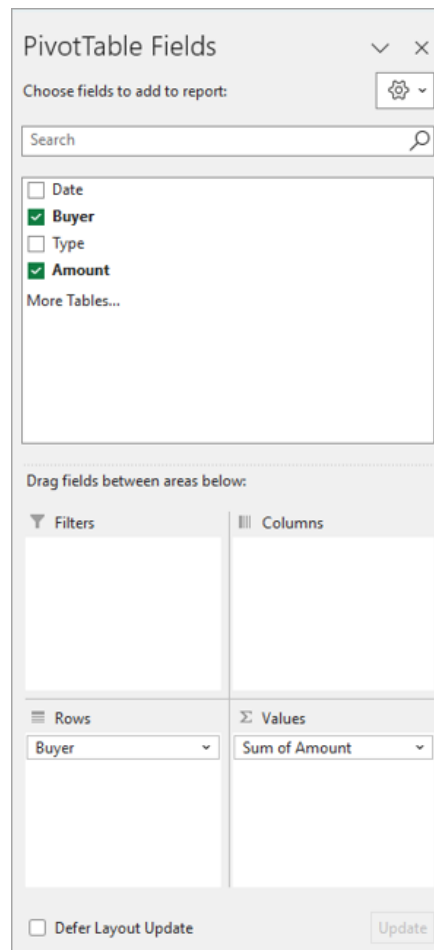
Упорядкування полів у зведеній таблиці

Після створення зведеної таблиці ви побачите список полів. Ви можете змінити дизайн зведеної таблиці, додавши та впорядкувавши її поля.

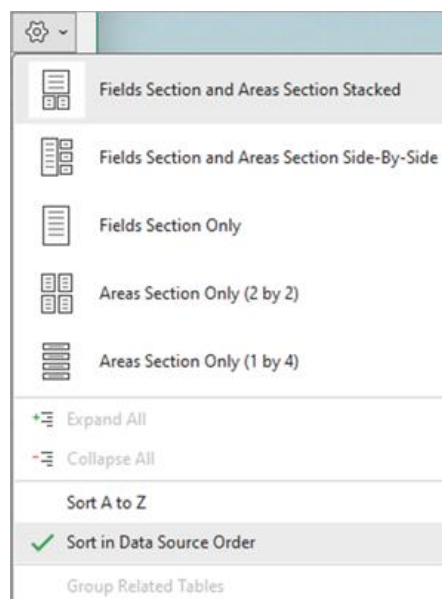
Список полів має з'явитися, коли ви клацнете будь-де у зведеній таблиці. Якщо ви клацнете всередині зведеної таблиці, але не бачите списку полів, відкрийте його, клацнувши будь-де у зведеній таблиці. Потім відобразіть **інструменти зведеної таблиці** на стрічці та натисніть «Аналіз» > «Список полів».



У списку полів є розділ полів, у якому ви вибираєте поля, які потрібно відобразити у зведеній таблиці, і розділ «Області» (внизу), у якому ви можете впорядкувати ці поля так, як вам потрібно.



Порада: якщо ви хочете змінити спосіб відображення розділів у списку полів, натисніть кнопку **Інструменти**, а потім виберіть потрібний макет.



Додавання, зміна порядку та видалення полів у списку полів

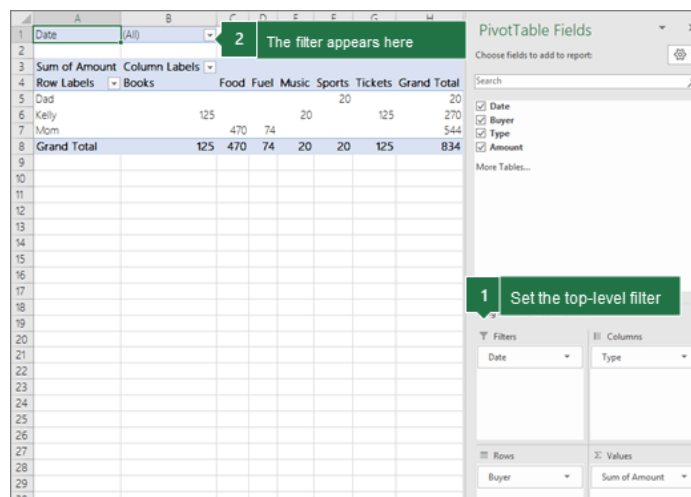
Використовуйте розділ полів у списку полів, щоб додати поля до зведеної таблиці, установивши прапорці біля імен полів, щоб розмістити ці поля в області за замовчуванням у списку полів.

Примітка: зазвичай нечислові поля додаються до області «Рядки», числові поля додаються до області «Значення», а ієрархії дат і часу онлайн-аналітичної обробки (OLAP) додаються до області «Стовпці».

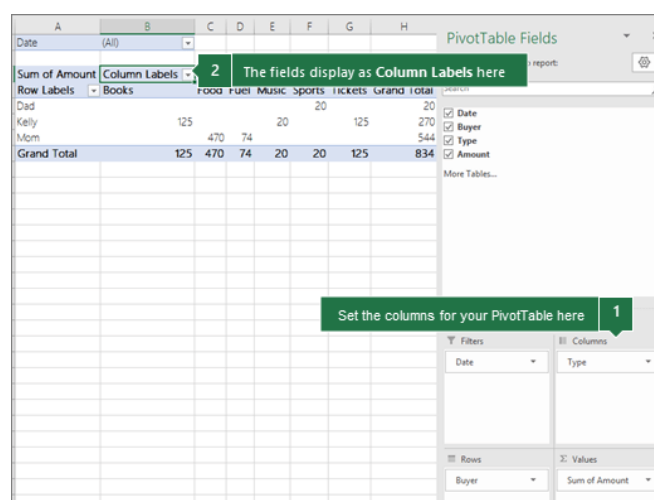
Використовуйте розділ областей (унизу) списку полів, щоб змінити порядок полів, як вам потрібно, перетягуючи їх між чотирма областями.

Поля, які ви розміщуєте в різних областях, відображаються у зведеній таблиці так:

– **Поля області фільтрів** відображаються як фільтри звіту верхнього рівня над зведеною таблицею:

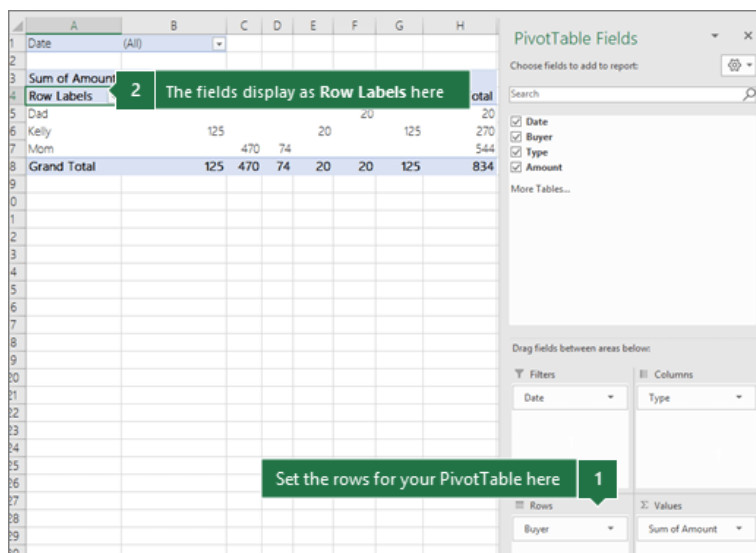


– **Поля області стовпців** відображаються як мітки стовпців у верхній частині зведеної таблиці:



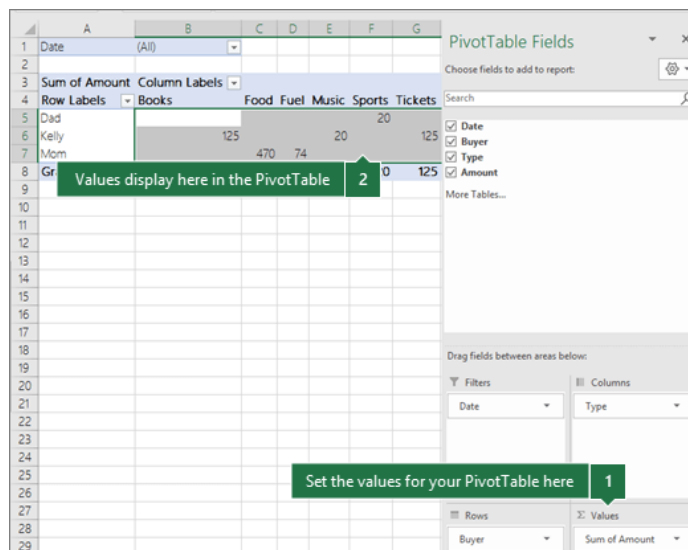
Залежно від ієрархії полів стовпці можуть бути вкладені в стовпці, які знаходяться вище за позицією.

– *Поля області рядків* відображаються як **підписи рядків** у лівій частині зведеної таблиці:



Залежно від ієрархії полів, рядки можуть бути вкладені в рядки, які є вищими за позицією.

– *Поля області значень* відображаються як зведені числові значення у зведеній таблиці:



Якщо у вас більше одного поля в області, ви можете змінити порядок, перетягнувши поля в точне положення, яке вам потрібно.

Щоб видалити поле зі зведеної таблиці, перетягніть поле з розділу його областей. Ви також можете видалити поля, натиснувши стрілку вниз поруч із полем і вибравши **Видалити поле**.

ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ

1. Як зведені таблиці допомагають узагальнювати дані з різних джерел?
2. Які основні переваги використання зведених таблиць для аналізу великих обсягів даних?
3. Як зведені таблиці допомагають у фільтрації та сортуванні даних?
4. Чому зведені таблиці підвищують ефективність роботи з даними?
5. Як використання зведених таблиць сприяє прийняттю обґрунтованих рішень?
6. Які ключові елементи складають зведену таблицю?
7. Що таке «джерело даних» для зведеної таблиці?
8. Які типи полів зведеної таблиці існують і які їх функції?
9. Що означає область «рядки» у зведеній таблиці?
10. Що означає область «стовпці» у зведеній таблиці?
11. Що означає область «значення» у зведеній таблиці?
12. Що означає область «фільтри» у зведеній таблиці?
13. Як здійснюється перетягування полів у зведеній таблиці?
14. Які налаштування можна змінювати для полів зведеної таблиці?
15. Які можливості динамічності та інтерактивності мають зведені таблиці?
16. Як створити зведену таблицю в Excel?
17. Які дані мають бути організовані для створення зведеної таблиці?
18. Як вибрати місце для розміщення звіту зведеної таблиці?
19. Як оновити зведену таблицю після додавання нових даних до джерела?
20. Як налаштувати відображення значень у зведеній таблиці, щоб показати їх у відсотках від загальної суми?

ТЕМА 6. ГРУПУВАННЯ ДАНИХ, ПІДВЕДЕННЯ ПІДСУМКІВ, КОНСОЛІДАЦІЯ ДАНИХ У MS EXCEL. ВІЗУАЛІЗАЦІЯ ДАНИХ

6.1. Групування та розгрупування даних у зведеній таблиці.

6.2. Фільтрування даних у зведеній таблиці.

6.3. Створення зведеної діаграми.

6.1. Групування та розгрупування даних у зведеній таблиці

Групування даних у зведеній таблиці допомагає показати підмножину даних для аналізу. Наприклад, ви можете згрупувати громіздкий список полів дати й часу у зведеній таблиці за кварталами та місяцями.

Формування даних групи

1. У зведеній таблиці клацніть значення правою кнопкою миші та виберіть «Групувати».

2. У полі «Групування» встановіть прапорці «Початок о» та «Закінчення о» та за потреби відредагуйте значення.

3. У розділі **До** виберіть період часу. Для числових полів введіть число, яке визначає інтервал для кожної групи.

4. Виберіть **ОК**.

Групування вибраних елементів

1. Утримуйте **Ctrl** і виберіть два або більше значень.

2. Клацніть правою кнопкою миші та виберіть «Група».

Групування за датою та часом

Завдяки групуванню часу зв'язки між пов'язаними з часом полями автоматично виявляються та групуються разом, коли ви додаєте рядки полів часу до своїх зведених таблиць. Згрупувавши їх разом, ви можете перетягнути групу до своєї зведеної таблиці та почати аналіз.

Country	(All)	
Row Labels		Sum of Order Amount
Amy Dodsworth		75048.04
7/15/2006		2490.5
7/31/2006		1873.8
10/10/2006		5275.71
10/21/2006		88.5
12/25/2006		166

Country	(All)	
Row Labels		Sum of Order Amount
Amy Dodsworth		75048.04
Qtr1		
Jan		6660.62
Feb		20418.34
Mar		5401.05
Qtr2		
Apr		10881.61
May		555.6
Jun		3482.5

Як назвати групу

1. Виберіть групу.
2. Виберіть **Аналіз > Параметри поля**. На вкладці «Аналіз зведеної таблиці» в розділі «Активне поле» натисніть «Параметри поля».
3. Змініть **настроюване ім'я** на потрібне, а потім виберіть **ОК**.


Як розгрупувати згруповані дані

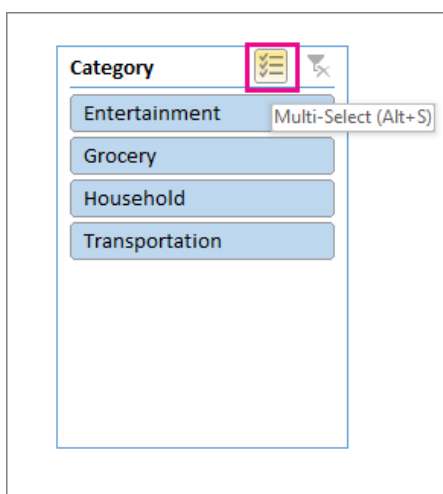
1. Клацніть правою кнопкою миші будь-який елемент у групі.
2. Виберіть **Розгрупувати**.

6.2. Фільтрування даних у зведеній таблиці


Зведені таблиці чудово підходять для створення докладних підсумків із великих наборів даних.

Фільтруванням даних за допомогою зрізу

1. Виберіть будь-яку комірку у зведеній таблиці, а потім на вкладці «Аналіз зведеної таблиці» виберіть  **Вставити слайсер**.
2. Виберіть поля, для яких ви хочете створити зрізи, і виберіть **ОК**.
3. Excel розмістить один роздільник для кожного зробленого вами вибору на робочому аркуші, але ви самі можете впорядкувати їх і задати розмір, який буде найкращим для вас.
4. Натисніть кнопки зрізу, щоб вибрати елементи, які потрібно відобразити у зведеній таблиці.



Фільтрування даних вручну

1. Щоб відобразити автофільтр, виберіть  стрілку спадного меню фільтра, яка змінюється залежно від макета звіту.

Компактний план

Row Labels	Sum of Sales
Accessories	68400
Bikes	6300
Clothing	66000
Components	32100
Grand Total	172800

Поле «Значення» знаходиться в області «Рядки»

Column Labels	Accessories	Bikes	Clothing	Componen	Grand Total
Sum of Sales	68400	6300	66000	32100	172800

Поле «Значення» знаходиться в області «Стовпці»

Табличний макет

Category	Sum of Sales
Accessories	68400
Bikes	6300
Clothing	66000
Components	32100
Grand Total	172800

Відображає назву поля значень у верхньому лівому куті

- Щоб відфільтрувати, створивши умовний вираз, виберіть «**Фільтри міток**», а потім створіть фільтр міток.
- Щоб відфільтрувати за значеннями, виберіть «**Фільтри значень**», а потім створіть фільтр значень.
- Щоб відфільтрувати за певними мітками рядків (у компактному макеті) або мітками стовпців (у табличному макеті), зніміть прапорець «**Вибрати все**», а потім установіть прапорці біля елементів, які потрібно відобразити. Ви також можете фільтрувати, ввівши текст у поле **пошуку**.
- Виберіть **ОК**.

Порада: ви також можете додати фільтри до поля «Фільтр» зведеної таблиці. Це дає можливість створювати окремі аркуші зведеної таблиці для кожного елемента в полі фільтра.

Як показати 10 верхніх або нижніх елементів:

- Щоб відобразити автофільтр, виберіть стрілку спадного меню фільтра, яка змінюється залежно від макета звіту.

Компактний план

Row Labels	Sum of Sales
Accessories	68400
Bikes	6300
Clothing	66000
Components	32100
Grand Total	172800

Поле «Значення» знаходиться в області «Рядки»

Column Labels	Accessories	Bikes	Clothing	Component	Grand Total
Sum of Sales	68400	6300	66000	32100	172800

Поле «Значення» знаходиться в області «Стовпці»

Табличний макет

Category	Sum of Sales
Accessories	68400
Bikes	6300
Clothing	66000
Components	32100
Grand Total	172800

Відображає назву поля значень у верхньому лівому куті

2. Виберіть **Фільтри значень > 10 найкращих**.
3. У першому полі виберіть «**Вгорі**» або «**Внизу**».
4. У другому полі введіть число.
5. У третьому полі виконайте такі дії:
 - щоб відфільтрувати за кількістю елементів, виберіть «**Пункти**»;
 - щоб відфільтрувати за відсотками, виберіть «**Відсоток**»;
 - щоб відфільтрувати за сумою, виберіть «**Сума**».
6. У четвертому полі виберіть поле «**Значення**».

Використання фільтра звіту

Використовуючи фільтр звіту, ви можете швидко відобразити інший набір значень у зведеній таблиці. Елементи, які ви вибрали у фільтрі, відображаються у зведеній таблиці, а невибрані елементи будуть приховані. Якщо ви хочете відображати сторінки фільтрів (набір значень, які відповідають вибраним елементам фільтра звіту) на окремих аркушах, ви можете вказати цей параметр.

Як додати фільтр звіту

1. Клацніть будь-де всередині зведеної таблиці.

З'явиться панель **«Поля зведеної таблиці»**.

2. У списку полів зведеної таблиці клацніть поле в області та виберіть

Перемістити до фільтра звіту.

Ви можете повторити цей крок, щоб створити кілька фільтрів звіту. Фільтри звітів відображаються над зведеною таблицею для легкого доступу.

Щоб змінити порядок полів, в області **фільтрів** ви можете або перетягнути поля в потрібне місце, або двічі клацнути на полі та вибрати **«Перемістити вгору»** чи **«Перемістити вниз»**. Порядок фільтрів звіту буде відповідно відображено у зведеній таблиці.

Як відобразити фільтри звітів у рядках або стовпцях

1. Клацніть зведену таблицю або пов'язану зведеною таблицю зведеної діаграми.

2. Клацніть правою кнопкою миші будь-де у зведеній таблиці, а потім виберіть **Параметри зведеної таблиці**.

3. На вкладці **«Макет»** укажіть такі параметри:

а) в області **«Фільтр звітів»** у списку **«Упорядкувати поля»** виконайте одну з таких дій:

– щоб відобразити фільтри звіту в рядках зверху вниз, виберіть **«Вниз»**, потім **«Нагору»**;

– щоб відобразити фільтри звіту в стовпцях зліва направо, виберіть **«Нагору»**, потім **«Вниз»**;

б) у полі **«Фільтрувати поля за стовпцем»** введіть або виберіть кількість полів для відображення, перш ніж перейти до іншого стовпця чи рядка (на основі параметра **«Упорядкувати поля»**, вказаного на попередньому кроці).

Як вибрати елементи у фільтрі звіту

1. У зведеній таблиці клацніть стрілку спадного меню поруч із фільтром звіту.

2. Установіть прапорці біля елементів, які потрібно відобразити у звіті. Щоб вибрати всі елементи, установіть прапорець біля **«Вибрати все»**.

Фільтр звіту тепер відображає відфільтровані елементи.

Як відобразити сторінки фільтрів звітів на окремих аркушах

1. Клацніть будь-де в зведеній таблиці (або пов'язаній зведеній таблиці зведеної діаграми), де є один або кілька фільтрів звіту.

2. Натисніть **Аналіз зведеної таблиці** (на стрічці) > **Параметри** > **Показати сторінки фільтра звіту**.

3. У діалоговому вікні «Показати сторінки фільтра звіту» виберіть поле фільтра звіту, а потім натисніть **ОК**.

Як створити фільтр за вибором

1. У зведеній таблиці виберіть один або кілька елементів у полі, яке потрібно відфільтрувати.

2. Клацніть правою кнопкою миші елемент із виділення, а потім натисніть **«Фільтр»**.

3. Виконайте одну з таких дій:

– щоб відобразити вибрані елементи, натисніть **«Зберегти лише вибрані елементи»**;

– щоб приховати вибрані елементи, натисніть **«Приховати вибрані елементи»**.

Порада: ви можете знову відобразити приховані елементи, видаливши фільтр. Клацніть правою кнопкою миші інший елемент у цьому ж полі, натисніть **«Фільтр»**, а потім **«Очистити фільтр»**.

Параметри фільтрації

Якщо ви хочете застосувати кілька фільтрів до поля або якщо ви не хочете відображати кнопки фільтрів у зведеній таблиці, можна ввімкнути або вимкнути ці та інші параметри фільтрації:

1. Клацніть у будь-якому місці зведеної таблиці, щоб відобразити вкладки зведеної таблиці на стрічці.

2. На вкладці **«Аналіз зведеної таблиці»** натисніть **«Параметри»**, а потім виконайте такі дії:

а) у діалоговому вікні **«Параметри зведеної таблиці»** клацніть вкладку **«Підсумки та фільтри»**;

б) в області **«Фільтри»** встановіть або зніміть прапорець **«Дозволити кілька фільтрів на поле»** залежно від того, що вам потрібно;

в) перейдіть на вкладку **«Відображення»**, а потім установіть або зніміть прапорець **«Показати підписи полів і фільтри»**, щоб відобразити або приховати підписи полів і спадні списки фільтрів.

6.3. Створення зведеної діаграми

Візуалізація зведених таблиць у Excel допомагає швидко і ефективно аналізувати та презентувати дані. Завдяки гнучким можливостям налаштування **зведені діаграми** є потужним інструментом для прийняття обґрунтованих рішень.

Excel підтримує такі **типи діаграм** для візуалізації даних зведених таблиць:

– **гістограми** використовуються для відображення розподілу даних за категоріями;

- *лінійні діаграми* використовуються для показу тенденцій даних за певний період;
- *кругові діаграми* використовуються для візуалізації частки кожної категорії від загальної суми;
- *стовпчасті діаграми* використовуються для порівняння даних між різними групами;
- *діаграми з областями* використовуються для показу зміни величин з плином часу.

Як створити зведену діаграму

1. Виберіть клітинку у зведеній таблиці. Переконайтеся, що активна клітинка знаходиться всередині зведеної таблиці.
2. Перейдіть на вкладку «**Вставка**».
3. Виберіть тип діаграми: натисніть «**Зведена діаграма**» (**PivotChart**). З'явиться вікно з варіантами типів діаграм.
4. Виберіть місце для діаграми: на тому ж листі або на новому.
5. Натисніть **ОК**.
6. Налаштування полів: виберіть поля для відображення у діаграмі, перетягуючи їх у відповідні області.

Налаштування зведеної діаграми

Після створення зведеної діаграми ви можете налаштувати її, щоб покращити візуалізацію, шляхом:

- додавання заголовків і підписів осей для кращого розуміння діаграми;
- зміни типу діаграми;
- форматування елементів діаграми: кольорів, шрифтів та стилів;
- додавання ліній тренду для відображення загальної тенденції даних;
- додавання даних через панель полів.

Використання зведених діаграм для аналізу даних

Зведені діаграми допомагають аналізувати дані різними способами, зокрема:

- порівнянням даних між різними категоріями або періодами;
- виявленням тенденцій та аномалій у даних;
- узагальненням великих обсягів даних в одному графічному вигляді;
- ефектної презентації даних на нарадах або в звітах.

Приклади використання зведених діаграм

- *аналіз продажів*: візуалізація обсягів продажів за місяцями для виявлення сезонних коливань;

- *фінансовий аналіз*: побудова діаграми для аналізу витрат та доходів за різні періоди;
- *аналіз ефективності*: порівняння продуктивності різних відділів компанії.

ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ

1. Як згрупувати дані у зведеній таблиці за датою і часом?
2. Як можна згрупувати числові дані у зведеній таблиці?
3. Як розгрупувати згруповані дані у зведеній таблиці?
4. Як вибрати кілька значень для групування у зведеній таблиці?
5. Які типи даних можна групувати у зведених таблицях?
6. Як змінити інтервал для групування числових даних?
7. Які існують методи фільтрування даних у зведених таблицях?
8. Як вставити зрізи для швидкого фільтрування даних?
9. Як працюють автофільтри у зведених таблицях?
10. Як створити фільтри міток у зведеній таблиці?
11. Як створити фільтри значень у зведеній таблиці?
12. Як відфільтрувати дані вручну у зведеній таблиці?
13. Як відобразити верхні або нижні 10 елементів у зведеній таблиці?
14. Як використовувати фільтр звіту у зведеній таблиці?
15. Як змінити порядок полів у фільтрі звіту?
16. Як застосувати кілька фільтрів до одного поля у зведеній таблиці?
17. Що таке зведена діаграма і як вона відрізняється від зведеної таблиці?
18. Як створити зведену діаграму на основі зведеної таблиці?
19. Як вибрати дані для відображення у зведеній діаграмі?
20. Які типи зведених діаграм існують в Excel?

ТЕМА 7. РОЗВ'ЯЗАННЯ ОПТИМІЗАЦІЙНИХ ЗАДАЧ

7.1. *Поняття оптимізаційної задачі.*

7.2. *Розв'язання оптимізаційних задач за допомогою надбудови Microsoft Excel Розв'язувач.*

7.3. *Розв'язання оптимізаційних задач за допомогою VBA (Visual Basic for Applications).*

7.1. *Поняття оптимізаційної задачі*

Задача оптимізації – це задача знаходження точки (точок) екстремуму, або декількох екстремумів заданої функції.

Формальне визначення:

Нехай задано деяку множину X із n -вимірного евклідового простору і функцію $f(x)$, визначену на X . Необхідно знайти точки мінімуму значень функції $f(x)$ на X . Або:

$$f(x) \rightarrow \min, x \in X,$$

де $f(x)$ – цільова функція, X – допустима множина, кожна точка x цієї множини – допустима точка задачі.

Також задачу оптимізації можна сформулювати як пошук максимуму (максимумів) цільової функції:

$$f(x) \rightarrow \max, x \in X.$$

Ця задача еквівалентна попередній задачі мінімізації цільової функції зі знаком мінус, в тому сенсі, що їх множини розв'язків збігаються.

Розв'язки задачі можна розділити на дві множини:

Глобальні (глобального мінімуму) – це такі допустимі точки x^* , в яких цільова функція має найменше значення на всій допустимій області.

Локальні (локального мінімуму) – це такі допустимі точки x^* , в яких цільова функція приймає найменше значення в деякому околі.

Для їх вирішення застосовуються економіко-математичні методи (методи економетрики).

Класичні задачі економетрики

1. *Транспортна задача:* мінімізація витрат на перевезення товарів від джерел до призначених місць з урахуванням обмежень на обсяги перевезень.

Застосування: логістика (постачання сировини, збут продукції, перевезення, транспортування працівників).

2. *Задача комівояжера*: знаходження найкоротшого маршруту, який проходить через усі точки (міста) і повертається в початковий пункт.

Застосування: маршрутизація транспорту, логістика, телекомунікації.

3. *Задача розміщення*: вибір місць розташування об'єктів (заводів, складів), щоб задовольнити попит та мінімізувати загальні витрати.

Застосування: планування локацій філій, підприємств, складів, магазинів.

4. *Задача управління запасами*: максимізація прибутку або мінімізація витрат під час управління рівнем запасів товарів, сировини, готової продукції.

Застосування: логістика, управління постачанням, роздрібна та оптова торгівля.

5. *Задача про раціон*: оптимізація складу їжі для досягнення певних цілей, як-от забезпечення необхідних споживчих речовин, мінімізація витрат, врахування дієтичних обмежень, задоволення інших факторів, наприклад, особисті вподобання або заборони.

Застосування: медицина, фітнес, тваринництво, харчова промисловість, громадське харчування.

6. *Задача про розкрій матеріалів*: вирізання деталей із заготовок стандартного розміру з мінімальними втратами та оптимальним використанням матеріалу.

Застосування: меблеве виробництво, будівництво, текстильна, автомобільна промисловість, електроніка, упаковка, друкарство.

7. *Задача про призначення*: ефективний розподіл ресурсів (зазвичай людських) для виконання конкретних завдань чи робіт з урахуванням ступеню підготовки (компетентностей) працівників.

Застосування: управління кадрами, виконання проєктів.

7.2. Розв'язання оптимізаційних задач за допомогою надбудови Microsoft Excel. Розв'язувач

Надбудова «Розв'язувач» (Solver) в Excel – це інструмент для вирішення складних оптимізаційних задач. Вона використовується для пошуку оптимального рішення, яке відповідає певним умовам або обмеженням, наприклад, максимізація прибутку чи мінімізація витрат.

Як встановити та активувати «Розв'язувач» в Excel

1. Запуск Excel: відкрийте програму Excel.
2. Перехід до меню «Файл»: у верхньому лівому куті натисніть на вкладку «Файл».
3. Вибір «Параметри»: у нижній частині меню виберіть «Параметри».
4. Надбудови: у вікні параметрів виберіть розділ «Надбудови».

5. Увімкнення «Розв'язувача»: у нижньому меню виберіть «Надбудови Excel» і натисніть «Перейти».

6. Вибір «Розв'язувача»: у вікні «Надбудови» знайдіть та поставте галочку біля «Розв'язувач» і натисніть «ОК».

7. Використання «Розв'язувача»: після цього на вкладці «Дані» у групі «Аналіз» з'явиться кнопка «Розв'язувач».

Основні налаштування «Розв'язувача»:

1. Цільова функція: необхідно задати функцію, яку потрібно мінімізувати, максимізувати або наблизити до конкретного значення.

2. Змінні клітинки – це комірки, значення яких змінює «Розв'язувач» для пошуку оптимального рішення.

3. Обмеження: ви можете додавати умови, яких повинні дотримуватися змінні (наприклад, їх значення має бути більше або менше певного числа).

Приклад використання «Розв'язувача» для розв'язання транспортної задачі

Транспортна задача – це класична задача оптимізації, яка полягає в мінімізації вартості перевезення товарів від постачальників до споживачів, враховуючи обмеження на обсяги виробництва та потреби.

Формулювання задачі:

Маємо 3 постачальників і 3 споживачів. Потрібно перевезти товари від постачальників до споживачів, мінімізувавши загальні витрати на транспортування.

Постачальники мають певні обсяги товарів, які можуть постачати:

Постачальник 1: 20 одиниць

Постачальник 2: 30 одиниць

Постачальник 3: 25 одиниць

Споживачі мають певні потреби в товарах:

Споживач 1: 15 одиниць

Споживач 2: 35 одиниць

Споживач 3: 25 одиниць

Вартість перевезення товарів від кожного постачальника до кожного споживача подана у вигляді таблиці:

Постачальник / Споживач	Споживач 1	Споживач 2	Споживач 3
Постачальник 1	8	6	10
Постачальник 2	9	12	13
Постачальник 3	14	9	16

Кроки для розв'язання транспортної задачі в Excel за допомогою «Розв'язувача»:

1. Введення даних в Excel

Створіть таблицю в Excel з вартістю перевезення, обсягами постачання та потребами:

Постачальник / Споживач	Споживач 1	Споживач 2	Споживач 3	Постачання
Постачальник 1				20
Постачальник 2				30
Постачальник 3				25
Потреби	15	35	25	

2. Налаштування «Розв'язувача»

1) Цільова функція: потрібно мінімізувати загальні витрати на транспортування.

Цільова функція – це сума добутків перевезених одиниць на вартість перевезення.

Додайте комірки для кількості товарів, що будуть перевезені між постачальниками і споживачами.

Додайте формулу для обчислення загальних витрат: =SUMPRODUCT (масив_кількостей, масив_вартостей).

2) Змінні клітинки – це комірки, що містять кількість товарів, що будуть перевезені від постачальників до споживачів. Ці комірки надбудова використовує для виведення кількості товарів які перевезені у кожному напрямі.

3) Обмеження:

Кожен постачальник повинен постачати товар у межах своїх можливостей (обсяг постачання).

Кожен споживач повинен отримати точно стільки товару, скільки йому потрібно.

Установіть обмеження на те, що значення в комірках для кількості товарів повинні бути додатними (≥ 0).

3. Розв'язання задачі

1. Перейдіть до вкладки «Дані» і натисніть кнопку «Розв'язувач».

2. У полі «Установити цільову комірку» виберіть комірку, в якій обчислюються загальні витрати.

3. У полі «Рівно» виберіть «Min», оскільки нам потрібно мінімізувати загальні витрати.

4. У полі «Змінюючи клітинки змінних» виберіть діапазон клітинок, що відповідають кількості товарів, які постачаються від кожного постачальника до кожного споживача.

5. Додайте обмеження:

Для кожного постачальника: сума товарів, що постачається, не може перевищувати наявний обсяг постачання.

Для кожного споживача: сума отриманих товарів має дорівнювати його потребам.

6. Встановіть обмеження, що кількість перевезених товарів повинна бути не від'ємною.

7. Натисніть **«Вирішити»**.

4. Результат

Після того, як «Розв'язувач» знайде рішення, ви побачите оптимальні кількості товарів, які потрібно перевезти від кожного постачальника до кожного споживача для мінімізації загальних витрат.

Висновок:

Використання «Розв'язувача» для розв'язання транспортної задачі допомагає швидко знаходити оптимальні рішення, враховуючи різні обмеження на постачання та потреби.

7.3. Розв'язання оптимізаційних задач за допомогою VBA (Visual Basic for Applications)

Розв'язання оптимізаційних задач за допомогою VBA (Visual Basic for Applications) у Microsoft Excel є ефективним для автоматизації складних обчислень і алгоритмів.

Основні етапи використання VBA для розв'язання оптимізаційних задач:

1. Підготовка даних в Excel

Наприклад, це можуть бути вихідні дані для задачі оптимізації.

2. Запуск редактора VBA

Натисніть **Alt + F11**, щоб відкрити редактор VBA. Створіть новий модуль для запису коду.

3. Написання коду

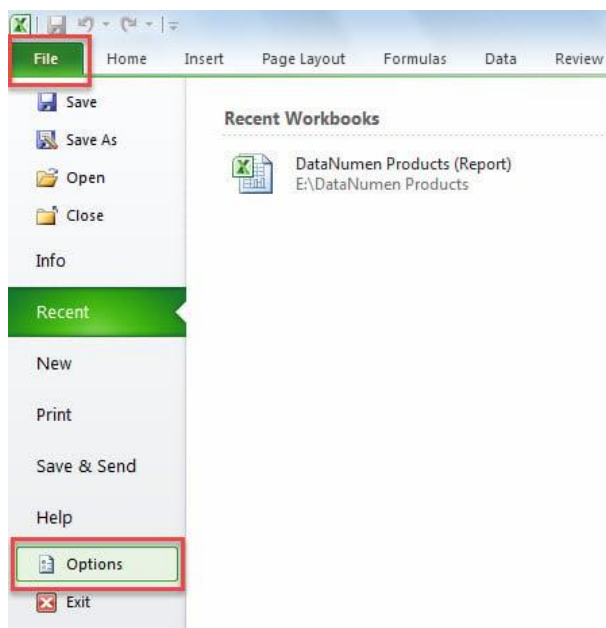
Напишіть VBA код, який виконує необхідні розрахунки та оптимізацію. Це може бути цикл, який перебирає можливі варіанти рішень і обирає найкращий.

4. Запуск коду:

Крок 1: Показати стрічку «Розробник»:

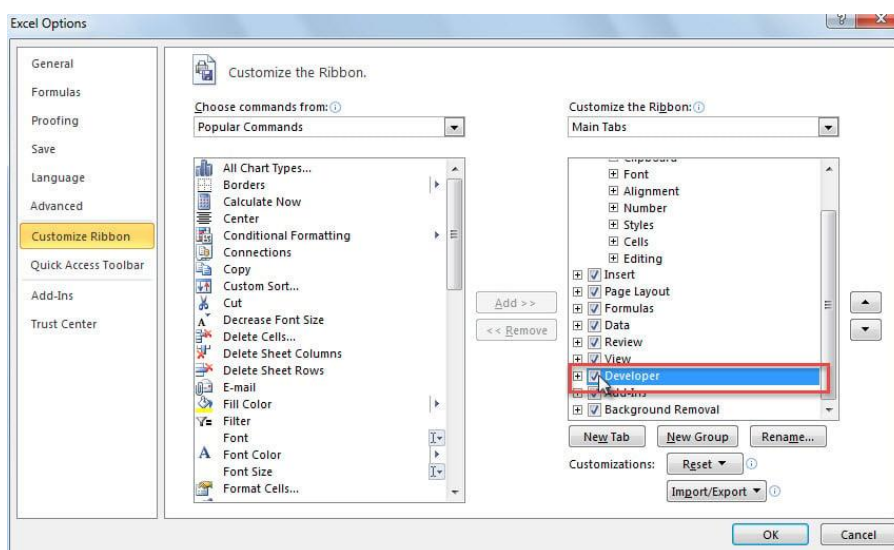
1. Запустіть програму Excel.

2. Перейдіть до меню **«Файл»** і натисніть **«Параметри»**.



3. У спливаючому вікні «**Параметри Excel**» перейдіть на вкладку «**Налаштувати стрічку**».

4. Далі знайдіть з правого боку і поставте прапорець біля опції «**Розробник**».

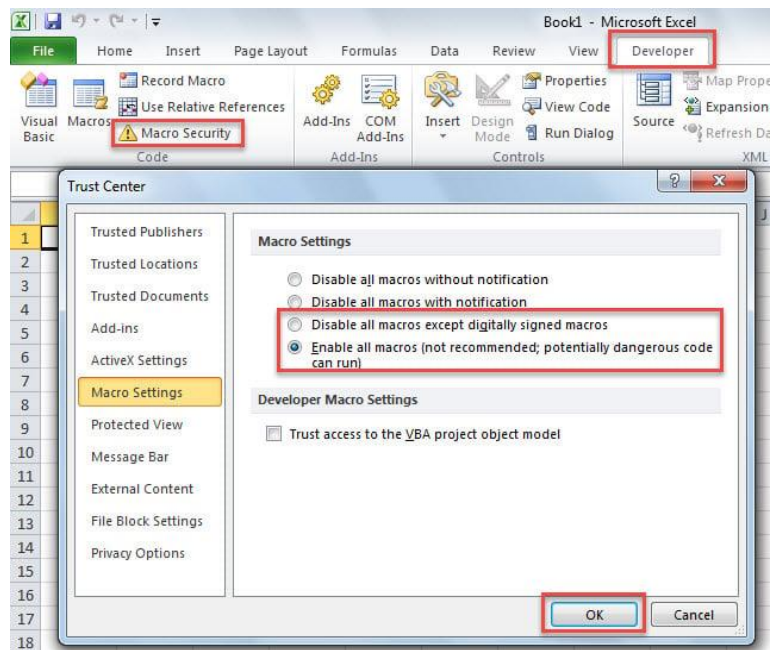


5. Натисніть **ОК**, щоб увімкнути зміни.

Крок 2: Змініть налаштування безпеки макросів.

Оскільки Excel постачається з автоматичним вимкненням макросів у налаштуваннях безпеки, спочатку потрібно змінити налаштування, виконавши такі дії:

1. Спочатку в головному вікні Excel перейдіть на стрічку «**Розробник**».
2. Потім натисніть кнопку «**Macro Security**» у групі «**Code**».
3. Далі у новому вікні виберіть «**Вимкнути всі макроси, крім цифрових підписаних макросів**» або «**Увімкнути всі макроси**» на ваш смак.

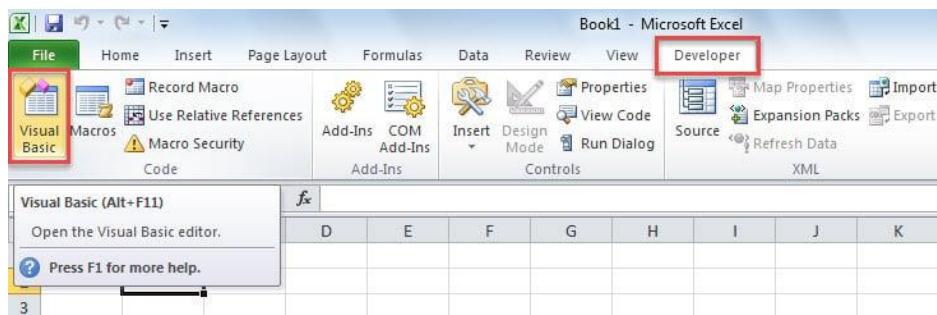


4. Натисніть **ОК**, щоб зберегти зміни.

Крок 3: Запустіть редактор Excel VBA.

Редактор Excel VBA – це місце, куди ви вставляєте код. Ви можете отримати доступ до нього двома способами:

1. На вкладці «Розробник» натисніть кнопку «**Visual Basic**» у групі «**Код**».



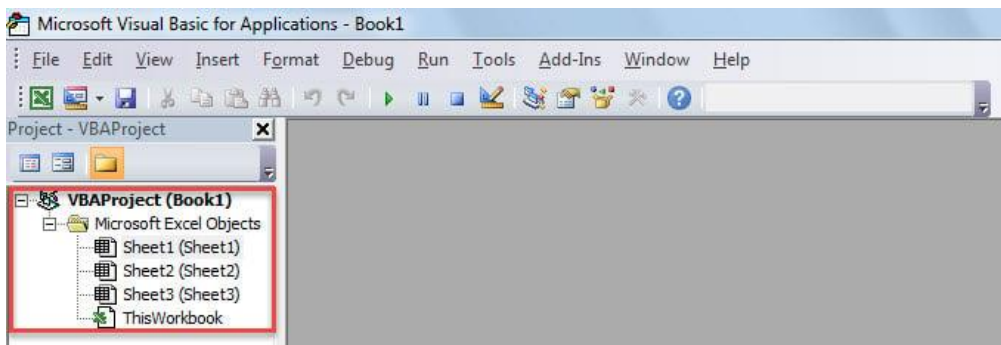
2. Натисніть клавіші швидкого доступу «**Alt + F11**», щоб відкрити вікно «**Microsoft Visual Basic for Applications**».

Крок 4: Додайте код VBA.

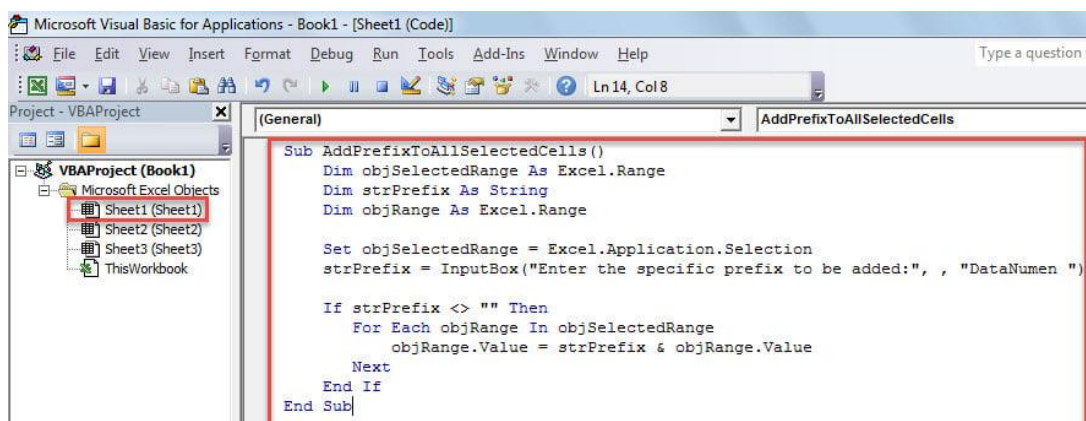
Після входу в редактор Excel VBA вам потрібно додати свій код до проекту або до модуля:

1. *Введіть код у наявний проект*

У розгорнутій папці «**Об'єкти Microsoft Excel**» ви можете побачити наявні проекти, включно з проектами для кожного аркуша та проектом для поточної книги під назвою «**ThisWorkbook**»:



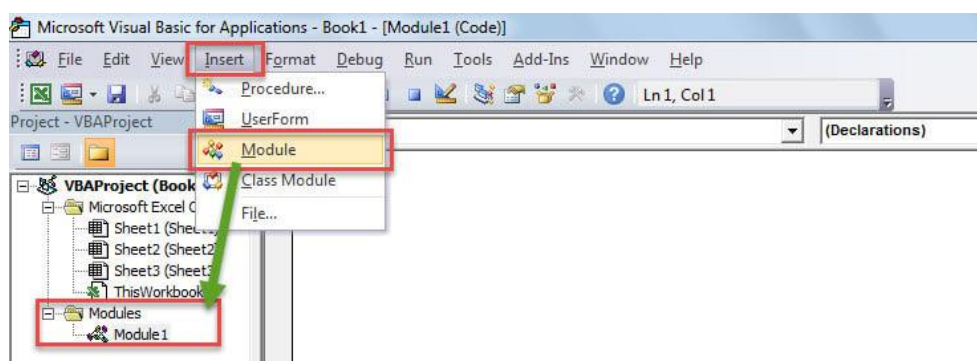
- 1) двічі клацніть на проєкті, щоб відкрити його вікно;
- 2) вставте в нього код.



2. Помістіть код у модуль

Ви можете додати код до модуля, виконавши такі дії:

- 1) натисніть кнопку «**Вставити**» на панелі інструментів;
- 2) зі спадного списку виберіть «**Модуль**»;

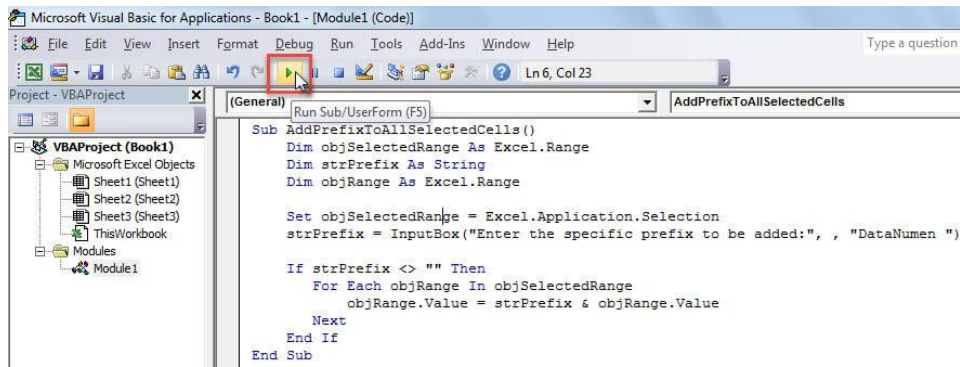


- 3) помістіть свій код у новий модуль.

Крок 5: Запустіть код VBA:

1. Запустіть макрос у редакторі VBA.

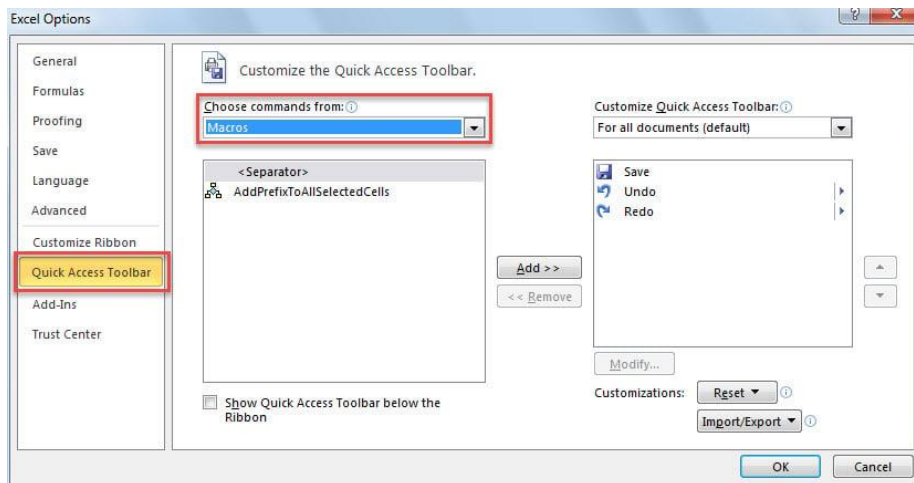
Щоб запустити VBA у вікні «Microsoft Visual Basic for Applications», натисніть кнопку клавіші «**F5**» або піктограму «**Виконати**» на панелі інструментів.



2. Запустіть макрос через панель швидкого доступу (QAT).

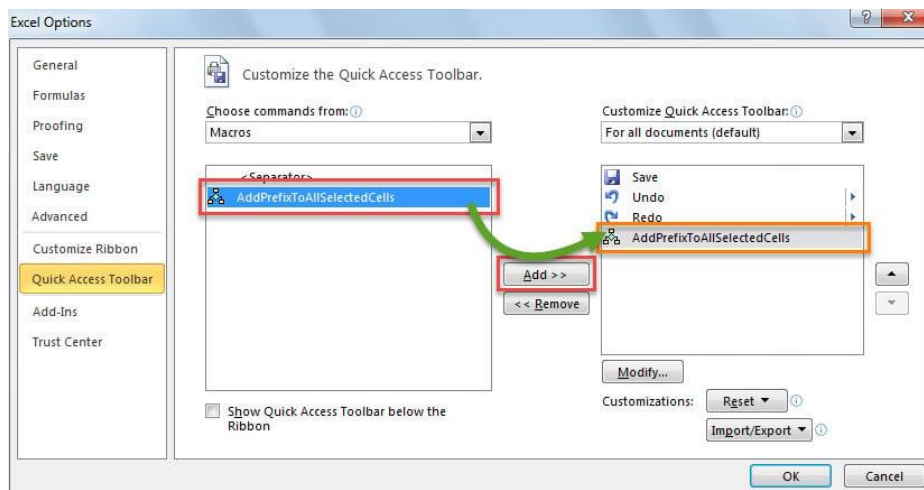
Ви можете додати макрос на панель швидкого доступу та запустити його через QAT, виконавши такі дії:

- 1) перейдіть до «Файл» > «Параметри»;
- 2) у спливаючому вікні перейдіть на вкладку «Панель швидкого доступу»;
- 3) виберіть «Макроси» зі списку «Виберіть команди з»;



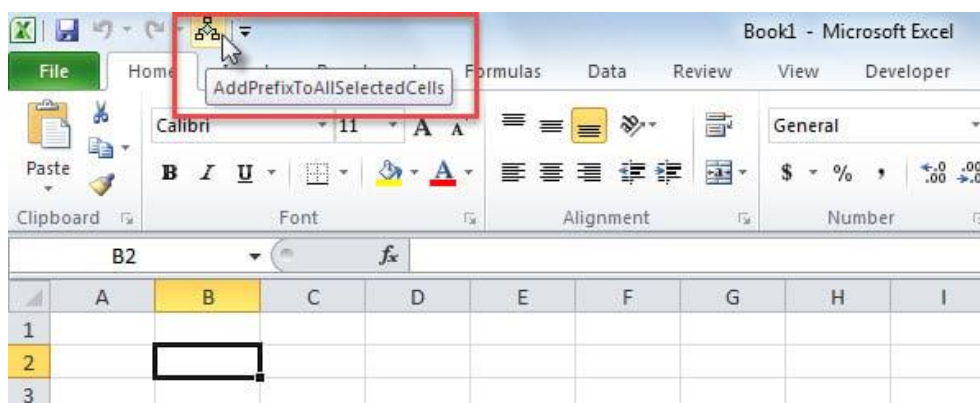
4) виберіть макрос ліворуч;

5) натисніть кнопку «Додати» в центрі;



6) натисніть **ОК**, щоб зберегти зміни на панелі швидкого доступу;

7) в головному вікні Excel, щоб запустити макрос, натисніть кнопку макросу на панелі швидкого доступу, як показано на скриншоті:



3. Запустіть макрос через стрічку

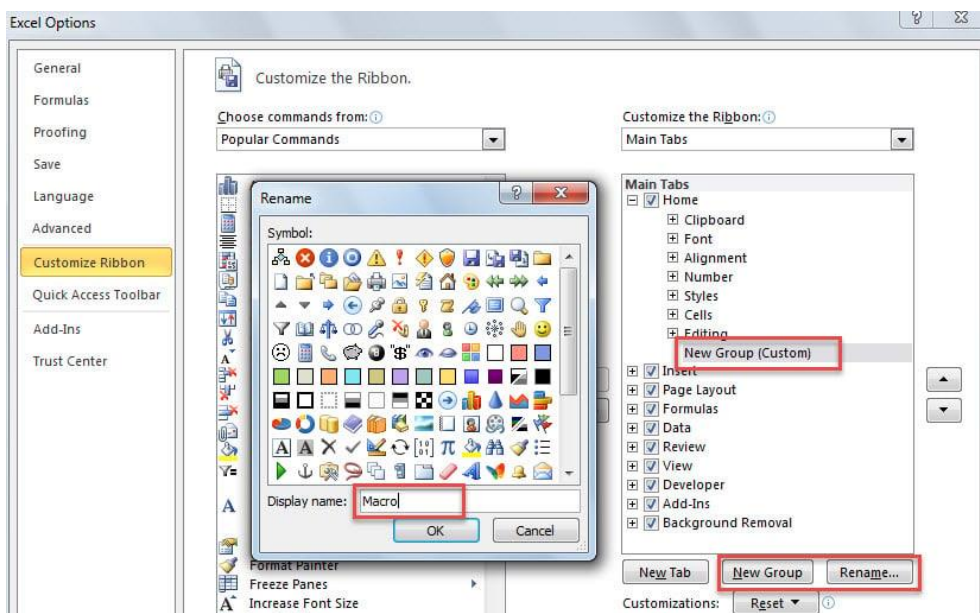
Як і панель швидкого доступу, ви можете додавати та запускати макрос у стрічці:

1) відкрийте меню «Файл» і натисніть «Параметри»;

2) у «Параметрах Excel» перейдіть до «Налаштування стрічки»;

3) створіть нову групу на вкладці «Домашня сторінка» від «Нова група»;

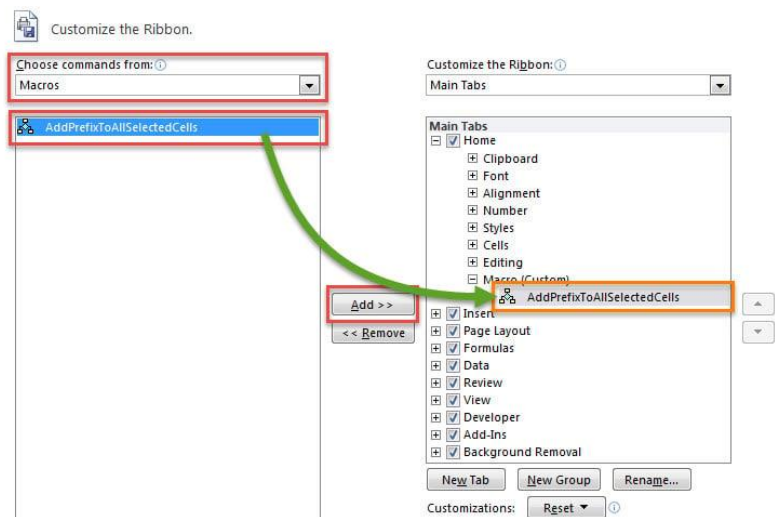
4) за бажанням ви можете перейменувати цю групу за допомогою «Перейменувати...»;



5) виберіть «Макроси» у списку «Виберіть команди з»;

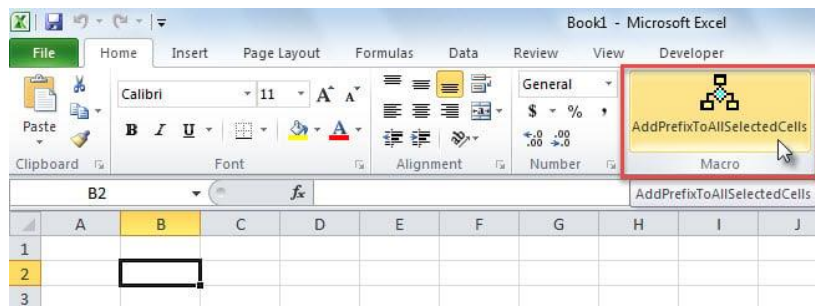
6) подайте потрібний макрос у лівій частині;

7) натисніть кнопку «Додати», щоб додати макрос до нової групи;



8) натисніть **ОК**, щоб зберегти параметри;

9) щоб запустити макрос, натисніть на додану кнопку макросу на стрічці.



Приклад VBA коду для простої оптимізаційної задачі

Sub Optimize()

Dim bestValue As Double

Dim bestX As Double

Dim bestY As Double

Dim x As Double

Dim y As Double

Dim currentValue As Double

bestValue = -1E+308 ' Початкове значення (мінімальне можливе число)

For x = 1 To 100

For y = 1 To 100

currentValue = x * y - (x^2 - y^2) ' Ваше оптимізаційне рівняння

If currentValue > bestValue Then

bestValue = currentValue

bestX = x

bestY = y

End If

Next y

Next x

```
MsgBox «Найкраще значення: « & bestValue & « при X=« & bestX & « та Y=« & bestY  
End Sub
```

Цей код ілюструє просту оптимізаційну задачу, де ми шукаємо найкращі значення для змінних x і y , які максимізують функцію $xy - (x^2 - y^2)$.

ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ

1. Як формулюється задача мінімізації цільової функції?
2. Як формулюється задача максимізації цільової функції?
3. У чому полягає еквівалентність задач мінімізації та максимізації цільової функції?
4. Що таке глобальний мінімум цільової функції?
5. Що таке локальний мінімум цільової функції?
6. Як визначається допустима множина X у задачі оптимізації?
7. Як визначається цільова функція $f(x)$ у задачі оптимізації?
8. Які економіко-математичні методи застосовуються для розв'язання задач оптимізації?
9. Що таке задача комівояжера і де вона застосовується?
10. Що таке задача розміщення і як вона використовується в плануванні?
11. У чому полягає задача управління запасами?
12. Як формулюється задача про раціон?
13. Що таке задача про розкрій матеріалів і де вона застосовується?
14. Яке основне застосування задачі про призначення?
15. Які основні етапи використання VBA для розв'язання оптимізаційних задач?
16. Як відкрити редактор VBA в Excel?
17. Як підготувати дані в Excel для задачі оптимізації?
18. Який код VBA необхідний для запуску простого оптимізаційного алгоритму?
19. Як змінити налаштування безпеки макросів у Excel?
20. Як запустити макрос у редакторі VBA?

ТЕМА 8. АВТОМАТИЗОВАНИЙ СТАТИСТИЧНИЙ АНАЛІЗ

8.1. Аналіз із використанням мови програмування Basic.

8.2. Аналіз із використанням VBA (Visual Basic for Applications).

8.1. Аналіз із використанням мови програмування Basic

Автоматизований статистичний аналіз – це процес використання комп'ютерних програм для виконання статистичних розрахунків, аналізу даних та створення звітів.

Мова програмування **Basic** (Beginners' All-purpose Symbolic Instruction Code) є однією з перших мов програмування, яка має простий синтаксис і використовується для навчання програмуванню та розробки простих програм.

Основні переваги використання мови Basic для автоматизованого статистичного аналізу:

1. **Простота використання:** Basic має простий та зрозумілий синтаксис, що робить його доступним для новачків у програмуванні.

2. **Автоматизація обчислень:** Basic дає змогу автоматизувати обчислення основних статистичних показників, як-от середнє, медіана, стандартне відхилення тощо.

3. **Робота з файлами:** Basic може працювати з файлами, що допомагає зчитувати дані з текстових файлів, виконувати аналіз та зберігати результати.

4. **Створення програм:** Мова Basic дає змогу створювати програми, які можуть виконувати складні обчислення і аналіз даних автоматично.

Приклади використання мови Basic для статистичного аналізу (коди):

1. Обчислення основних статистичних показників: середньої, медіани, стандартного відхилення набору даних:

```
10 DIM data(100)
20 FOR i = 1 TO 100
30 INPUT «Введіть значення:», data(i)
40 NEXT i

50 sum = 0
60 FOR i = 1 TO 100
70 sum = sum + data(i)
80 NEXT i
90 mean = sum / 100
100 PRINT «Середнє:»; mean

110 REM Обчислення медіани
120 FOR i = 1 TO 100
130 FOR j = i + 1 TO 100
```

```

140 IF data(i) > data(j) THEN SWAP data(i), data(j)
150 NEXT j
160 NEXT i
170 IF 100 MOD 2 = 0 THEN
180 median = (data(50) + data(51)) / 2
190 ELSE
200 median = data(50)
210 END IF
220 PRINT «Медіана:»; median

230 REM Обчислення стандартного відхилення
240 sum = 0
250 FOR i = 1 TO 100
260 sum = sum + (data(i) - mean) ^ 2
270 NEXT i
280 stdDev = SQR(sum / 99)
290 PRINT «Стандартне відхилення:» ; stdDev

```

2. Робота з файлами для зчитування даних із файла і збереження результатів аналізу:

```

10 OPEN «дані.txt» FOR INPUT AS #1
20 DIM data(100)
30 FOR i = 1 TO 100
40 INPUT #1, data(i)
50 NEXT i
60 CLOSE #1

70 sum = 0
80 FOR i = 1 TO 100
90 sum = sum + data(i)
100 NEXT i
110 mean = sum / 100

120 OPEN «результати.txt» FOR OUTPUT AS #2
130 PRINT #2, «Середнє:»; mean
140 CLOSE #2

```

Тож використання мови програмування Basic для автоматизованого статистичного аналізу може бути корисним для новачків у програмуванні та для виконання базових статистичних розрахунків. Basic дає змогу автоматизувати обчислення, працювати з файлами та створювати прості програми для аналізу даних.

8.2. Аналіз із використанням VBA (Visual Basic for Applications)

Використання VBA у Microsoft Excel допомагає автоматизувати процеси аналізу даних, зменшуючи потребу в ручних розрахунках та знижуючи ризик помилок.

Основні переваги використання VBA для автоматизованого статистичного аналізу:

1. *Автоматизація рутинних завдань:* VBA дає змогу автоматизувати повторювані завдання, як-от обчислення основних статистичних показників (середнє, медіана, стандартне відхилення тощо).

2. *Обробка великих обсягів даних:* VBA може обробляти великі набори даних, що значно прискорює аналіз і підвищує ефективність.

3. *Створення макросів:* Макроси VBA можуть виконувати складні обчислення і аналіз даних автоматично, що допомагає зосередитися на інтерпретації результатів.

4. *Побудова графіків та діаграм:* VBA може автоматично створювати графіки і діаграми, що ілюструють результати аналізу.

5. *Створення звітів:* VBA дає змогу створювати автоматизовані звіти, що містять результати аналізу, таблиці, графіки та інші візуалізації.

Приклади використання VBA для статистичного аналізу (коди):

1. Обчислення основних статистичних показників: середньої, медіани, стандартного відхилення набору даних:

```
Sub StatAnalysis()
```

```
Dim ws As Worksheet
```

```
Set ws = ThisWorkbook.Sheets(«Data»)
```

```
Dim dataRange As Range
```

```
Set dataRange = ws.Range(«A1:A100») ‘ Задайте діапазон даних
```

```
Dim mean As Double
```

```
mean = Application.WorksheetFunction.Average(dataRange)
```

```
Dim median As Double
```

```
median = Application.WorksheetFunction.Median(dataRange)
```

```
Dim stdDev As Double
```

```
stdDev = Application.WorksheetFunction.StDev(dataRange)
```

```
‘ Виведення результатів на аркуш
```

```
ws.Range(«C1»).Value = «Середнє»
```

```
ws.Range(«C2»).Value = mean
```

```
ws.Range(«D1»).Value = «Медіана»
```

```
ws.Range(«D2»).Value = median
```

```
ws.Range(«E1»).Value = «Стандартне відхилення»  
ws.Range(«E2»).Value = stdDev  
End Sub
```

2. Побудова графіків і діаграм (гістограм):

```
Sub CreateHistogram()  
    Dim ws As Worksheet  
    Set ws = ThisWorkbook.Sheets(«Data»)  
  
    Dim chartObj As ChartObject  
    Set chartObj = ws.ChartObjects.Add(Left:=300, Width:=400, Top:=50, Height:=300)  
  
    With chartObj.Chart  
        .SetSourceData Source:=ws.Range(«A1:A100») ‘ Задайте діапазон даних  
        .ChartType = xlColumnClustered  
        .HasTitle = True  
        .ChartTitle.Text = «Гістограма»  
        .Axes(xlCategory).HasTitle = True  
        .Axes(xlCategory).AxisTitle.Text = «Значення»  
        .Axes(xlValue).HasTitle = True  
        .Axes(xlValue).AxisTitle.Text = «Частота»  
    End With  
End Sub
```

Тож використання VBA для автоматизованого статистичного аналізу в Excel дає змогу значно підвищити ефективність і точність аналізу даних. Це особливо корисно під час роботи з великими наборами даних або виконання складних обчислень, які вимагають багато часу та ресурсів за ручного виконання. Автоматизація за допомогою VBA допомагає зосередитися на інтерпретації результатів і прийнятті рішень на основі даних.

ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ

1. Які основні цілі автоматизованого статистичного аналізу?
2. Які переваги має мова програмування Basic для автоматизованого статистичного аналізу?
3. Що таке синтаксис мови Basic?
4. Як мова Basic дає змогу працювати з файлами для зчитування та збереження даних?
5. Як обчислити середнє значення набору даних за допомогою Basic?
6. Як реалізувати обчислення медіани в програмі на Basic?
7. Як обчислити стандартне відхилення за допомогою мови Basic?

8. Які основні кроки необхідні для створення програми на Basic для статистичного аналізу?
9. Як реалізувати сортування масиву даних у програмі на Basic?
10. Як працювати з циклами в Basic для обробки масивів даних?
11. Як використовувати умови та оператори порівняння в Basic для аналізу даних?
12. Як зберігати результати статистичного аналізу в текстовому файлі за допомогою Basic?
13. Як можна обробляти пропущені значення в наборі даних за допомогою Basic?
14. Які можливості має Basic для візуалізації даних?
15. Які основні можливості VBA для автоматизованого статистичного аналізу в Excel?
16. Як відкрити редактор VBA в Excel і створити новий модуль?
17. Як написати VBA код для обчислення середнього значення даних у Excel?
18. Як використовувати функції VBA для обчислення медіани та стандартного відхилення в Excel?
19. Як створити макрос у VBA для автоматизації статистичного аналізу?
20. Як побудувати графік або діаграму за допомогою VBA в Excel?

СПИСОК ВИКОРИСТАНИХ ПОСИЛАНЬ

Основна література

1. Гевлич І. Г. Аналіз даних засобами MS Excel. Методичні вказівки до виконання лабораторних робіт та завдань з самостійної роботи студентів. Вінниця: ДонНУ імені Василя Стуса, 2022. 54 с.
2. Гевлич І. Г. Інформаційні технології в економіці і управлінні: конспект лекцій. Вінниця: ДонНУ імені Василя Стуса, 2023. 132 с.
3. Гороховатський В. О., Творошенко І. С. Методи інтелектуального аналізу та оброблення даних: навч. посіб. М-во освіти і науки України, Харків. нац. ун-т радіоелектроніки. Харків: ХНУРЕ, 2021. 92 с.
4. Дейна А. Ю., Фурман Т. Ю. Методичні рекомендації до виконання лабораторних робіт з дисципліни «Пакети прикладних програм у бізнес-аналітиці». Вінниця: ДонНУ імені Василя Стуса, 2020. 65 с.
5. Козьменко О. В., Кузьменко О. В. Економіко-математичні методи і моделі: економетрика. Суми: Університетська книга, 2023. 406 с.
6. Основи статистичного моделювання: навч. посіб. / за заг. ред. С. В. Чугаєвської, Н. В. Ковтун. Житомир: Видавництво ПП «Рута», 2022. 604 с.
7. Перегуда О. В., Капустян О. А., Курилко О. Б. Статистична обробка даних: навч. посіб. Київ: КНУ імені Тараса Шевченка, 2022. 103 с.

Допоміжна література

1. Бандоріна Л. М., Лозовська Л. І., Савчук Л. М. Моделювання економіки: навч. посіб. Дніпро: УДУНТ, 2022. 154 с.
2. Гевлич І. Г. Викладання програмування та чисельного моделювання при підготовці економістів. *Економіка і організація управління*. 2021. № 3(43). С. 131–137.
3. Гевлич І. Г. Інформаційні технології в економетриці для вирішення фахових завдань. *Економіка і організація управління*. 2023. № 3(51). С. 34–42. DOI: 10.31558/2307-2318.2023.3.4.
4. Гевлич І. Г. Інформаційні технології при викладанні економетрики. *Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення (випуск 80): міжнародна науково-практична інтернет-конференція (19–20 вересня 2023 р.)*. 2023. URL: <http://www.konferenciaonline.org.ua/ua/article/id-1252/>
5. Гевлич І. Г., Левченко Н. К. Кореляційно-регресійний аналіз ефективності управління витратами на підприємстві ПрАТ «Вінницький завод «Маяк»». *Modern ways of development of science and the latest theories: The XIII International Scientific and Practical Conference*. (December 11–13, 2023, Madrid, Spain). P. 57–61.

URL: <https://eu-conf.com/wp-content/uploads/2023/11/MODERN-WAYS-OF-DEVELOPMENT-OF-SCIENCE-AND-THE-LATEST-THEORIES.pdf>

Інформаційні ресурси в мережі Інтернет

1. BASIC Programming Resources. URL: <https://www.qbasic.net/>
2. Statistica. Офіційний сайт. URL: <https://www.statista.com/markets/>
3. Vintage BASIC. URL: <http://www.vintage-basic.net/download.html>
4. Visual Basic для додатків. URL: <https://learn.microsoft.com/ru-ru/office/vba/api/overview/>
5. Державна служба статистики України. URL: www.ukrstat.gov.ua/
6. Інститут економіки і прогнозування НАН України. URL: <http://www.ief.org.ua>
7. Офіційний сайт Statisticstimes. GDP Indicators. URL: <http://statisticstimes.com/economy/world-statistics.php>
8. Офіційний сайт Євростату. URL: <https://ec.europa.eu/>
9. Офіційний сайт Світового Банку. URL: <https://databank.worldbank.org/>
10. Служба підтримки Microsoft. URL: <https://support.microsoft.com/uk-ua>

ДЛЯ ПОДАТОК

Навчальне видання

Гевлич Іван Геннадійович

**ПАКЕТИ ПРИКЛАДНИХ ПРОГРАМ
У БІЗНЕС-АНАЛІТИЦІ**

Конспект лекцій

Редактор О. А. Солдатова
Технічний редактор Т. О. Важеніна-Гопрак

Підписано до друку 16.12.2024
Формат 60×84/16. Папір офсетний.
Друк – цифровий. Умовн. друк. арк. 4,88.
Тираж 30. Зам. 88.

Донецький національний університет імені Василя Стуса
21021, м. Вінниця, 600-річчя, 21
Свідоцтво про внесення суб'єкта видавничої справи
до Державного реєстру
серія ДК № 5945 від 15.01.2018