УДК 621.316.13: 519.85

ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ СТАТИСТИЧНОГО АНАЛІЗУ

ДУБРОВІН В.І. канд. техн. наук, професор, професор кафедри програмних засобів Національного університету «Запорізька політехніка», Запоріжжя, Україна, e-mail: vdubrovin@gmail.com;

ДЕЙНЕГА Л.Ю. старший викладач кафедри програмних засобів Національного університету «Запорізька політехніка», Запоріжжя, Україна, e-mail: deynega.larisa@gmail.com;

ЯЦЕНКО А.К. студентка кафедри програмних засобів Національного університету «Запорізька політехніка», Запоріжжя, Україна, e-mail: nastya.yatsenko.zp@gmail.com

Мета роботи. *Аналіз існуючого програмного забезпечення для виконання статистичного аналізу для подальшого використання у рамках вибору необхідного програмного засобу для проведення обробки даних.*

Методи дослідження. *Для проведення огляду програмного забезпечення було проведено аналіз наукових статей та відкритих джерел щодо програмного забезпечення статистичного аналізу.*

Отримані результати. *Вибір відповідного статистичного програмного забезпечення є ключовим рішенням у сфері аналізу даних, з численними опціями, які задовольняють різноманітні потреби. У цій статті міститься вичерпний огляд п'яти провідних статистичних програмних засобів: IBM SPSS Statistics, RStudio, Stata, Minitab і Python. У цій роботі виявлено ключові відомості про можливості, функції та придатність кожного інструменту для різних аналітичних завдань.*

Цей огляд робить висновок, що вибір статистичного програмного забезпечення повинен узгоджуватися з конкретними вимогами проекту, складністю даних і досвідом користувача. Дослідники та аналітики повинні враховувати свої аналітичні цілі та переваги при виборі найбільш підходящого інструменту. Крім того, для прийняття обґрунтованих рішень у цій динамічній сфері важливо бути в курсі нових тенденцій в аналізі даних і машинному навчанні.

Наукова новизна. *Проведений аналіз виявив можливості та застосування найбільш популярного програмного забезпечення для вирішення задач статистичного аналізу. Робота надає вичерпний огляд сучасних тенденцій та інновацій у галузі програмного забезпечення для статистичного аналізу, пропонуючи читачам глибше розуміння наявних інструментів.*

Практична цінність. *Проведений аналіз дозволить обрати програмне забезпечення для вирішення конкретної задачі статистичного аналізу на основі його характеристик та наявних вимог. Ця робота допомагає визначити практичні переваги програмного забезпечення для статистичного аналізу і сприяє впровадженню цих інструментів у різних сферах діяльності, забезпечуючи покращення процесів аналізу та прийняття рішень.*

Ключові слова: статистичний аналіз, програмне забезпечення, обробка даних, огляд програмних засобів, IBM SPSS Statistics, RStudio, Stata, Minitab, Python.

I. ВСТУП

У сучасному світі електротехніка та електроенергетика відіграють фундаментальну роль у забезпеченні життєво важливих функцій суспільства. Від електричної енергії залежать багато аспектів нашої повсякденної діяльності – від освітлення та обігріву до виробництва та технологічних процесів. В умовах постійно зростаючих потреб в енергоресурсах та стрімких технологічних змін, забезпечення надійності, ефективності та безпеки систем електропостачання стає ключовим завданням.

Саме тут статистичний аналіз входить у гру як потужний інструмент, що дозволяє оцінити, аналізувати та оптимізувати різноманітні аспекти роботи електротехнічних систем та компонентів. Статистичний аналіз, що базується на обробці великих обсягів даних, надає нам цінні інсайти про продуктивність, надійність та ефективність систем електропостачання, а також допомагає запобігати аварійним ситуаціям та покращувати якість електроенергії.

Завдяки статистичному аналізу, ми можемо визначити та керувати факторами, що впливають на функціонування систем, включаючи розподіл навантаження, стійкість до відмов, а також підвищення надійності компонентів. Цей підхід дозволяє здійснювати обґрунтовані рішення щодо модернізації та підвищення продуктивності систем електропостачання.

Статистичний аналіз також використовується для передбачення можливих аварій та визначення ризиків у роботі систем електропостачання. Це допомагає операторам та інженерам бути випереджувальними у запобіганні проблем та покращенні надійності мереж.

Слугуючи як міст між теорією та практикою, статистичний аналіз є невід'ємною частиною сучасної електротехніки та електроенергетики. Він дозволяє

© Дубровін В.І., Дейнега Л.Ю., Яценко А.К., 2023
DOI 10.15588/1607-6761-2023-3-3

нам проникнути в глибину даних, виявити закономірності та тренди, на основі яких можна розробляти ефективніші стратегії управління енергоресурсами та забезпечення надійності систем електропостачання.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ

Й. Мьолтофт у своїй роботі [1] розглядає використання методів статистичного аналізу даних випробувань на термін служби електронних компонентів. Детально розглянуто методи аналізу з використанням припущень про постійний рівень ризику (експоненціальний розподіл), розподілу Вейбулла та змішаних розподілів Вейбулла. Обговорюються методи, які використовуються для аналізу даних Вейбулла, коли самі дані неоднорідні через як видалення тестових зразків під час тестування, так і через непродовження нагляду за тестованими компонентами. Зрештою, увагу приділено ефекту двох або більше механізмів відмови, які можуть створювати S-подібні візерунки, коли дані наносяться на міліметровий папір Вейбулла.

Автори [2] розглядають окремі статистичні методи, що дозволяють ефективно оцінити вплив обраних факторів на ефективність процесу монтажу RF-систем. Розгляд проведено на прикладі досліджень впливу розміру пустот, що виникають у зоні пайки ВЧ-силових транзисторів, на значення вибраних електричних параметрів цих пристроїв у розглянутому електронному модулі. Було описано вибрані інструменти якості та проведено додаткові аналізи за допомогою методології шести сигм та програми Minitab.

III. МЕТА ДОСЛІДЖЕННЯ

Метою цього дослідження є аналіз існуючого програмного забезпечення для виконання статистичного аналізу, яке може бути використано для різноманітних завдань, що стосуються оцінки надійності компонентів систем, прогнозування подій та відмов, а також діагностики стану обладнання. Серед цих завдань важливо визначити та систематизувати можливості доступного програмного забезпечення, щоб забезпечити належне вирішення питань, пов'язаних із надійністю, передбаченням подій та виявленням проблем у функціонуванні системи. Крім того, це дослідження спрямоване на визначення наявних інструментів та методів, які можуть бути корисними для інженерів, дослідників та фахівців, що працюють у галузі технічної діагностики та управління надійністю систем.

IV. ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ДЛЯ СТАТИСТИЧНОГО АНАЛІЗУ

Статистичний аналіз відіграє важливу роль в електротехніці та електроенергетиці, допомагаючи в розумінні та оптимізації різних аспектів електричних пристроїв та компонентів. Він дозволяє здійснювати об'єктивний аналіз та оцінку надійності, тривалості служби та відмовостійкості електротехнічних систем. Покладаючися на статистичні методи, інженери можуть проводити дослідження та експерименти, щоб визначити фактори, які впливають на функціонування електротехнічних систем.

Так, наприклад, статистичний аналіз може використовуватись у таких областях (рис. 1):

Рисунок 1. Застосування статистичного аналізу

1. Оцінка надійності та довговічності компонентів та систем [3]. Статистичний аналіз може бути використаний для оцінки надійності та довговічності електричних компонентів, таких як дроти, ізоляція, напівпровідники, реле та інші. Це дозволяє визначити ймовірність відмови та прогнозувати термін служби.

2. Аналіз споживання електроенергії [4]. Статистичний аналіз може допомогти в аналізі та прогнозуванні споживання електроенергії у мережах різних масштабів – від домашніх мереж до промислових підприємств. Це може допомогти у плануванні енергетичних ресурсів та ефективної їх утилізації.

3. Прогнозування подій та відмов [5]. Статистичні методи дозволяють створювати моделі для прогнозування можливих подій, таких як навантаження в мережі, короткі замикання або відмови обладнання. Це допомагає операторам енергосистем більш ефективно планувати технічне обслуговування та реагувати на проблеми.

4. Діагностика та обслуговування [6]. Статистичний аналіз може допомогти у розробці методів діагностики стану обладнання на основі аналізу даних про його роботу. Це сприяє більш ефективному обслуговуванню, ремонту та заміні деталей.

5. Аналіз ризиків та безпеки [7]. Статистичний аналіз може бути використаний для оцінки ризиків, пов'язаних з різними аспектами електротехніки та електроенергетики. Це дозволяє розробляти заходи

для забезпечення безпеки працівників, обладнання та навколишнього середовища.

6. Дослідження параметрів мережі [8]. Статистичні методи можуть допомогти в аналізі параметрів електроенергетичних мереж, таких як втрати енергії, напруга, струми, реактивна потужність та інші. Це допомагає в оптимізації роботи мережі та підвищення ефективності.

Одними з основних програмних засобів для вирішення задач статистичного аналізу є наступні: IBM SPSS Statistics, RStudio, Stata, Minitab, [9] та Python. Кожен з цих інструментів володіє своїми унікальними перевагами та можливостями, які роблять їх популярними в різних галузях та серед різних користувачів. Далі будуть розглянуті можливості та використання наведених програмних засобів.

У табл. 1 наведено загальну характеристику зазначених програмних продуктів.

Таблиця 1. Загальні характеристика програмних засобів для статистичного аналізу

Основні характеристики	IBM SPSS Statistics	RStudio	Stata	Minitab	Python
Початковий випуск програми	1968	1992	1985	1972	1991
Розробник(и)	Норман Х. Ні, Дейл Х. Бент, К. Хадлай Халл	Джозеф Аллер	Вільям Гулд	Барбара Ф. Райан, Томас А. Раян молодший та Браян Л. Джойнер	Ґвідо ван Россум
Мова програмування	Java	C, Fortran, R	C	C, Java, C++	Python
Ліцензія	Пропрієтарна	GNU з вільним відкритим вихідним кодом	Пропрієтарна	Пропрієтарна	Безкоштовна, відкритий код
Головне призначення	Статистичний аналіз, чисельний аналіз	Аналітика даних, статистичний аналіз, машинне навчання	Статистичний аналіз, чисельний аналіз	Статистичний аналіз	Наука про дані

IBM SPSS Statistics. IBM SPSS Statistics – це популярний статистичний аналізовий пакет, який використовується для обробки та аналізу даних в різних галузях, включаючи науку, бізнес, медицину та соціальні науки. Програма дозволяє користувачам виконувати різноманітні статистичні обчислення, візуалізувати дані та створювати звіти.

Серед можливостей SPSS Statistics розробники виділяють наступні [10]:

1. Розширена статистика: однофакторне/багатофакторне моделювання для отримання більш точних висновків при аналізі складних взаємозв'язків.

2. Регресія: прогнозування категоричних результатів та застосовування процедури нелінійної регресії.

3. Настроювані таблиці: легка організація великої кількості даних.

4. Точні тести: більш точні висновки при малих вибірках.

5. Техніка повторної вибірки: оцінка розподілу вибірки оцінювача шляхом повторної вибірки із заміною вихідної вибірки.

6. Відсутні значення: виявлення шаблонів відсутніх даних, оцінка підсумкової статистики та заповнення відсутніх значень.

7. Підготовка даних: оптимізація підготовки даних, щоб пришвидшити підготовку до аналізу та поліпшити точність висновків.

8. Категорії: прогнозування результатів та виявлення зв'язків, використовуючи категоріальні дані.

9. Прогнозування: створення часових рядів прогнозів.

10. Дерева рішень: використання класифікації та дерев рішень, щоб полегшити визначення груп та зв'язків та передбачити результати.

11. Комплексні вибірки: аналіз статистичних даних та інтерпретація результатів опитування складних вибірок.

12. Нейромережі: виявлення складних зв'язків та вдосконалення прогнозних моделей.

IBM SPSS Statistics найчастіше використовується в предметних областях, пов'язаних з соціальними на-

уками, економікою та маркетингом. Його основний фокус лежить на соціальних науках та аналізі даних у сферах, пов'язаних зі збором і обробкою інформації про людську поведінку та суспільні явища.

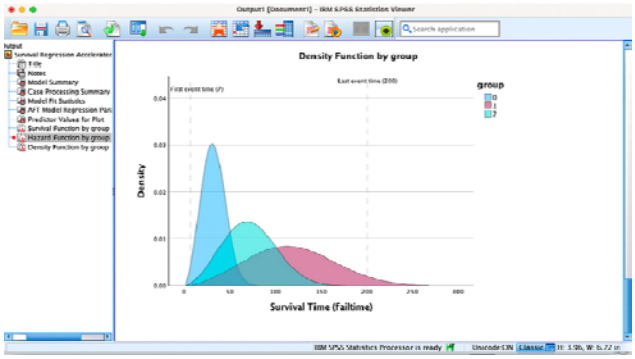

Рисунок 2. IBM SPSS Statistics [10]

RStudio. RStudio – це інтегроване середовище розробки (IDE) для мови програмування R, яка широко використовується для статистичного аналізу та обробки даних [11].

Серед можливостей RStudio можна виділити наступні:

1. Інтерфейс для написання коду: RStudio надає зручний інтерфейс для розробки, збереження та виконання скриптів на мові R. Скрипти дозволяють створювати послідовність команд для обробки та аналізу даних.

2. Робота з даними: середовище дає можливість обробляти та імпортувати дані з різних джерел, таких, як включаючи текстові файли, бази даних, Excel, CSV та інші.

3. Візуалізація: такі бібліотеки, як ggplot2, надають інструменти для побудови діаграм та графіків, що дозволить покращити представлення даних.

4. Статистичний аналіз: бібліотечні розширення середовища дають змогу працювати із задачами статистичного аналізу, включаючи дескриптивну статистику, t-тести, аналіз дисперсії, кореляцію, регресію, кластерний аналіз і багато інших.

5. Підготовка даних: такі бібліотеки, як dplyr та tidyr, дозволяють підготувати дані для подального аналізу, надаючи інструменти для фільтрації, сортування, групування, об'єднання та перетворення даних.

6. Моделі даних: RStudio надає інструменти для роботи з моделями даних, виконування прогнозування та класифікації, оцінки рівень надійності моделей і оптимізації їх параметрів.

7. Звіти: використовуючи R Markdown можна створювати звіти та презентації, вбудовувати код R, результати аналізу та графіку безпосередньо в документи Word, PDF, HTML.

8. Спільна робота та обмін даними: RStudio Server і RStudio Connect дозволяють працювати над проєктами у команді, обмінюючись даними та кодом.

R і RStudio є універсальними інструментами для статистичного аналізу та обробки даних і можуть бути використані в різних предметних областях, де статистика та аналіз даних є важливою частиною досліджень.

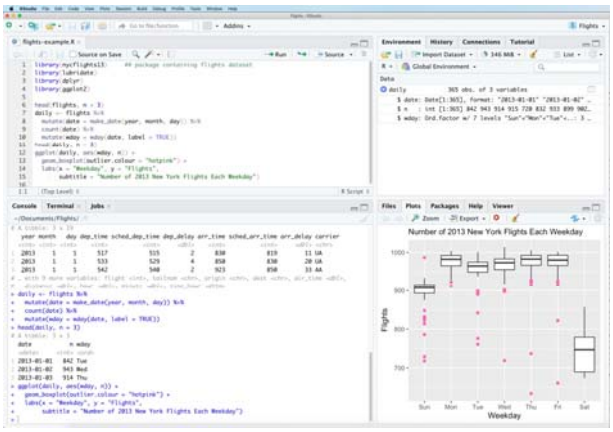

Рисунок 3. RStudio [12]

Stata. Stata - це потужне статистичне програмне забезпечення, яке використовується для проведення статистичного аналізу та досліджень даних в різних галузях науки, бізнесу та громадського сектора [13].

Можливості Stata включають:

1. Робота з даними: Stata підтримує імпорт даних з різних джерел, включаючи текстові файли, Excel, бази даних SQL, SAS, SPSS і багато інших форматів.

2. Візуалізація: Stata надає засоби для візуалізації даних через графіки, гістограми, діаграми розсіювання та інші методи візуалізації.

3. Дескриптивна статистика: надані інструменти для обчислення дескриптивних статистик, таких як середнє значення, медіана, стандартне відхилення, мода, квартилі та інші.

4. Інференційна статистика: програмне забезпечення надає можливість проведення t-тестів, аналізу дисперсії, регресійного аналізу, логістичної регресії, аналізу виживання, аналізу часових рядів та багато інших. Також можна виконувати операції для обчислення довірчих інтервалів та p-значень.

5. Моделі даних: Stata надає інструментарій для побудови та роботи зі статистичними моделями, включаючи лінійні та нелінійні регресійні моделі, дискримінантний аналіз, аналіз даних панельного типу, а також робити обчислення ефектів за допомогою інструментів для обробки даних з панельними даними.

6. Інтерфейс для роботи зі скриптами: є можливість використання команд та скриптів для автоматизації аналізу та створення звітів.

7. Спільна робота та обмін даними: Stata під-

тримує спільну роботу з іншими користувачами, дозволяючи обмінюватися скриптами, даними і результатами аналізу.

Stata відомий своєю потужністю та розширеними можливостями для статистичного аналізу та обробки даних, що робить його корисним інструментом для дослідження та аналізу даних у багатьох наукових галузях і сферах діяльності.

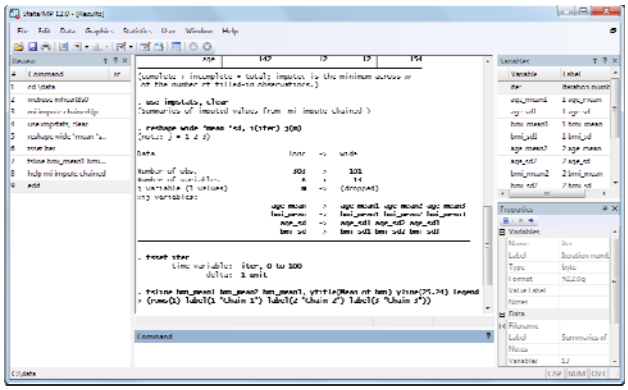

Рисунок 4. Stata [13]

Minitab. Minitab є пакетом програм обробки статистичних даних [14]. Серед його можливостей можна виділити наступні:

1. Збір і імпорт даних: Minitab дозволяє легко імпортувати дані з різних джерел, таких як Excel, бази даних SQL, текстові файли і інші.

2. Візуалізація: Minitab надає інструменти для візуалізації даних.

3. Статистичний аналіз: програмне забезпечення дозволяє вирішувати задачі статистичного аналізу, включаючи дескриптивну статистику (середнє значення, медіана, стандартне відхилення та ін.), t-тести, аналіз дисперсії, регресійний аналіз, аналіз виживання, байєсівський аналіз, аналіз ефекту від випадкових факторів та багато інших.

4. Моделі даних: Minitab дозволяє створювати та оцінювати статистичні моделі, включаючи лінійну та нелінійну регресію, логістичну регресію, аналіз часових рядів і багато інших.

5. Планування експериментів: наявний інструментарій для планування, проведення, а також оптимізації процесу проведення експериментів.

6. Автоматизація завдань: Minitab підтримує автоматизацію рутинних завдань за допомогою мови програмування Minitab.

Minitab найчастіше застосовується в галузях, де важливу роль відіграє контроль якості та процесів виробництва. Це програмне забезпечення допомагає підприємствам та організаціям досягати високої якості продукції, вдосконалювати виробничі процеси та зменшувати ризики дефектів. Науковці, інженери та фахівці з якості використовують Minitab для статистичного аналізу даних та розробки ефективних стратегій управління якістю [15].

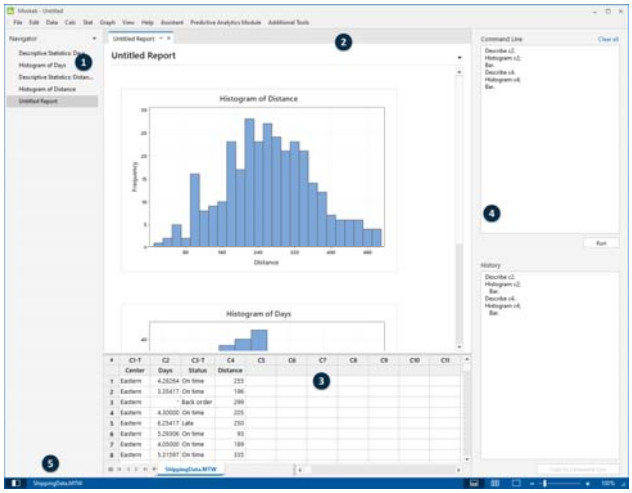

Рисунок 5. Minitab [14]

Python. Python – це мова програмування, яка отримала широку популярність в області статистичного аналізу і обробки даних завдяки своїй гнучкості та потужним бібліотекам для наукових обчислень [16].

Бібліотеки Python надають наступні можливості:

1. Обробка та аналіз даних: такі бібліотеки, як NumPy, pandas, SciPy дозволяють виконувати операції з даними, статистичні обчислення; працювати з масивами та таблицями.

2. Візуалізація: існує велика кількість бібліотек для візуалізації даних. Зокрема, такі бібліотеки, як Matplotlib, Seaborn та Plotly дозволяють легко представляти дані у вигляді графіків та діаграм.

3. Статистичний аналіз: Python має численні бібліотеки для статистичного аналізу, включаючи StatsModels (для виконання статистичного моделювання), scikit-learn (для машинного навчання та статистичного класифікації), SciPy (для статистичних тестів та операцій) і багато інших.

4. Машинне навчання: для виконання завдань статистичного аналізу, у Python є можливість використання бібліотек машинного навчання таких, як TensorFlow, Keras, PyTorch, і scikit-learn.

5. Інтеграція з іншим програмним забезпеченням: Python може бути інтегрований з іншими мовами програмування, такими як R, що розширює можливості аналізу даних. Також є можливість взаємодії з іншими програмами та бібліотеками за допомогою різних інтерфейсів.

6. Спільнота та підтримка: Python має велику та активну спільноту користувачів та розробників, тож є можливість знайти велику кількість необхідної документації.

Python, як і RStudio, є універсальним інструме-

нтом для вирішення задач статистичного аналізу. Python володіє багатьма бібліотеками та інструментами для статистичного аналізу, такими як NumPy, pandas, SciPy, scikit-learn, StatsModels та інші, що робить його потужним інструментом для роботи з даними у різних галузях науки та промисловості.

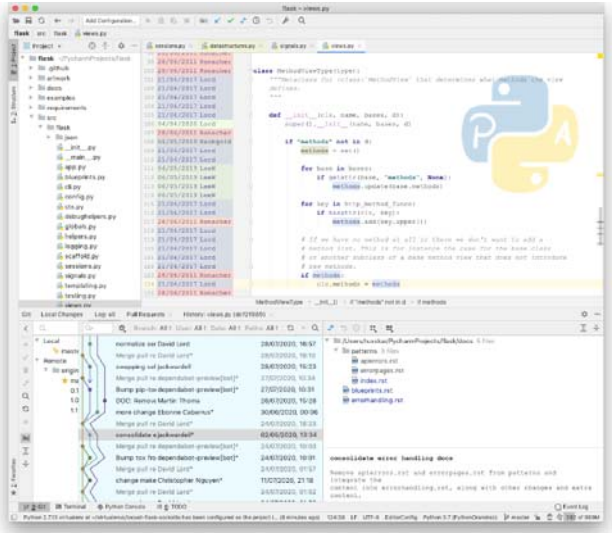

Рисунок 5. PyCharm, IDE для Python [17]

За результатами проведеного огляду можна скласти детальну порівняльну таблицю (табл. 2).

Можна побачити, що розглянуте програмне забезпечення має великий функціонал та дозволяє ефективно вирішувати задачі статистичного аналізу. Кожна з розглянутих програм є доволі потужним інструментом.

Підсумовуючи, вибір статистичного програмного забезпечення повинен ґрунтуватися на конкретних вимогах проєкту, складності даних і досвіді користувача. IBM SPSS Statistics є надійним вибором для користувачів, яким потрібен простий та дружній інтерфейс. IBM SPSS Statistics обслуговує спільноту соціальних наук та економіки, тоді як Minitab використовується у контролі якості та виробництві.

RStudio, Python та Stata є універсальними інструментами у питання вирішення задач статистичного аналізу. RStudio не має собі рівних у своїй гнучкості та аналітичних можливостях, що робить його популярним інструментом для статистиків і дослідників даних. Універсальність Python, його природа з відкритим кодом і надійні бібліотеки роблять його чудовим вибором для тих, хто хоче інтегрувати аналіз даних із програмуванням і машинним навчанням.

Таблиця 2. Загальні характеристика програмних засобів для статистичного аналізу

Характеристика	IBM SPSS Statistics	RStudio	Stata	Minitab	Python
Дескриптивна статистика	+	+	+	+	+
Інференційна статистика	+	+	+	+	+
Робота з моделями даних	+	+	+	+	+
Візуалізація	+	+	+	+	+
Створення звітів	+	+	+	+	±
Робота зі скриптами	-	+	+	+	+
Використання нейромереж машинного навчання	+	±	-	-	+
Автоматизація завдань	±	-	+	+	+
Інтеграція з іншими програмами	±	+	±	±	+
Спільна робота	±	±	+	±	±

V. ВИСНОВОК

Завдання вибору програмного забезпечення для проведення статистичного аналізу є надзвичайно важливим для дослідників, адже може неабияк впливати на результативність та вартість процесу. У даній ро-

боті було проведено огляд можливостей популярних програм для роботи зі статистикою: IBM SPSS Statistics, RStudio, Stata, Minitab, та Python. Усі розглянуті інструменти мають свої переваги та можуть бути обрані в залежності від вимог дослідника.

Так, IBM SPSS Statistics та Stata мають дружній користувацький інтерфейс та досить потужні можливості вирішення задач статистики. Minitab буде чудовим вибором для контролю якості та планування експериментів.

RStudio та Python варіанти з відкритим кодом, такі як RStudio та Python, пропонують економічно ефективні, гнучкі та настроювані рішення.

Зараз існують безліч інструментів для статистичного аналізу, і вибір програмного забезпечення залежить від конкретних потреб дослідження чи бізнесу. Важливо відзначити, що користувачі мають доступ до широкого спектру інструментів, які можуть полегшити їхню роботу і допомогти досягти результатів. Незалежно від обраного програмного забезпечення, статистичний аналіз залишається невід'ємною складовою успішного наукового дослідження та прийняття рішень у бізнесі.

REFERENCE

[1] Møltoft, J. Statistical analysis of data from electronic component lifetests (a tutorial paper). *Active and Passive Elec. Comp*., 1987, Vol. 12, pp. 259-279.

[2] Górecki K., Kowalke W., (2022). Application of Statistical Methods to Analyze the Quality of Electronic Circuits Assembly. *Applied Sciences*, 12, 12694.

[3] Elshahhat, A, Abu El Azm, WS. Statistical reliability analysis of electronic devices using generalized progressively hybrid censoring plan. *Qual Reliab Eng Int.*, 2022, 38, 1112–1130. DOI: https://doi.org/10.1002/qre.3058

[4] Xiaoqing, W., Nianping L., Wenjie Z. Statistical Analyses of Energy Consumption Data in Urban Office Buildings of Changsha, *Procedia Engineering*, 2015 121, 1158-1163, DOI: https://doi.org/10.1016/j.proeng.2015.09.125

[5] Dudek G., Piotrowski P., Baczy´nski D. Intelligent Forecasting and Optimization in Electrical Power Systems: Advances in Models and Applications. *Energies*, 2023, 16, 3024. DOI: https://doi.org/10.3390/en16073024

[6] Khwaja H. A., Gupta S. P., Kumar V. A Statistical Approach for Fault Diagnosis in Electrical Machines, *IETE Journal of Research*, 2010, 56:3, 146-155, DOI: 10.4103/0377-2063.67099

[7] Rychlik, Igor & Rydén, Jesper. Probability and risk analysis. An introduction for engineers. 2006. DOI: 10.1007/978-3-540-39521-8.

[8] Rofii F., Naba A., Dharmawan H.A., Hunaini F. Analysis of Electrical Power Quality Disturbances Based on Empirical Mode Decomposition and Statistical Parameters, *IOP Conf. Series: Materials Science and Engineering*, 2020, 846, 012050, DOI:10.1088/1757-899X/846/1/012050

[9] Abatan, S. M., Olayemi, M. S. (2014). The Role of Statistical Software in Data Analysis. *International Journal of Applied Research and Studies*, 2014, 3, 8, 1-15.

[10] IBM SPSS Statistics [Electronic resource] – Access mode: https://www.ibm.com/products/spss-statistics

[11] RStudio IDE [Electronic resource] – Access mode: https://posit.co/products/open-source/rstudio/

[12] RStudio [Electronic resource] – Access mode: https://uk.wikipedia.org/wiki/RStudio

[13] Stata [Electronic resource] – Access mode: https://www.stata.com/

[14] Minitab [Electronic resource] – Access mode: https://www.minitab.com/en-us/

[15] Rodriguez, D. Using Minitab to achieve Statistical Quality Control [Electronic resource] / D. Rodriguez – Access mode: https://www.invensislearning.com/blog/statistical-quality-control-using-minitab/

[16] Python [Electronic resource] – Access mode: https://www.python.org/

[17] PyCharm [Electronic resource] – Access mode: https://www.jetbrains.com/pycharm/

Стаття надійшла до редакції *03.10.2023*

STATISTICAL ANALYSIS SOFTWARE

DUBROVIN V.I. Ph.D., Associate professor, Associate professor of the software tools department of the National University "Zaporizhzhia Polytechnic", Zaporizhzhia, Ukraine, e-mail: vdubrovin@gmail.com;

DEINEHA L.Y. Senior Lecturer of the software tools department of the National University "Zaporizhzhia Polytechnic", Zaporizhzhia, Ukraine, e-mail: deynega.larisa@gmail.com;

YATSENKO A.K. Student of the software tools department of the National University "Zaporizhzhia Polytechnic", Zaporizhzhia, Ukraine, e-mail: nastya.yatsenko.zp@gmail.com

Purpose. *Analysis of existing software to perform statistical analysis for further use as part of the selection of the necessary software for data processing.*

Methodology. *To conduct a software review, an analysis of scientific articles and open sources on statistical analysis software was conducted.*

Findings. *Choosing the right statistical software is a key decision in the field of data analysis, with numerous options to meet a variety of needs. This article provides a comprehensive overview of five leading statistical software tools: IBM SPSS Statistics, RStudio, Stata, Minitab, and Python. This paper reveals key insights into the capabilities, functions, and suitability of each tool for various analytical tasks.*

This review concludes that the choice of statistical software should be consistent with specific project requirements, data complexity, and user experience. Researchers and analysts should consider their analytical goals and preferences when choosing the most appropriate tool. In addition, to make informed decisions in this dynamic field, it is important to stay abreast of new trends in data analysis and machine learning.

Originality. *The conducted analysis revealed the possibilities and application of the most popular software for solving problems of statistical analysis. The work provides a comprehensive overview of current trends and innovations in the field of software for statistical analysis, offering readers a deeper understanding of existing tools.*

Practical value. *The conducted analysis will allow to choose software for solving a specific task of statistical analysis based on its characteristics and existing requirements. This work helps to identify the practical benefits of statistical analysis software and promotes the implementation of these tools in various fields of activity, providing improvements in analysis and decision-making processes.*

Keywords: statistical analysis, software, data processing, software review, IBM SPSS Statistics, RStudio, Stata, Minitab, Python

REFERENCES

[1] Møltoft, J., (1987). Statistical analysis of data from electronic component lifetests (a tutorial paper). *Active and Passive Elec. Comp*., Vol. 12, pp. 259-279.

[2] Górecki K., Kowalke W., (2022). Application of Statistical Methods to Analyze the Quality of Electronic Circuits Assembly. *Applied Sciences*, 12, 12694.

[3] Elshahhat, A, Abu El Azm, WS. (2022). Statistical reliability analysis of electronic devices using generalized progressively hybrid censoring plan. Qual Reliab Eng Int., 38, 1112–1130. DOI: https://doi.org/10.1002/qre.3058

[4] Xiaoqing, W., Nianping L., Wenjie Z. (2015). Statistical Analyses of Energy Consumption Data in Urban Office Buildings of Changsha, Procedia Engineering, 121, 1158-1163, DOI: https://doi.org/10.1016/j.proeng.2015.09.125

[5] Dudek G., Piotrowski P., Baczy´nski D. (2023). Intelligent Forecasting and Optimization in Electrical Power Systems: Advances in Models and Applications. Energies, 16, 3024. DOI: https://doi.org/10.3390/en16073024

[6] Khwaja H. A., Gupta S. P., Kumar V. (2010). A Statistical Approach for Fault Diagnosis in Electrical Machines, IETE Journal of Research, 56:3, 146-155, DOI: 10.4103/0377-2063.67099

[7] Rychlik, Igor & Rydén, Jesper. (2006). Probability and risk analysis. An introduction for engineers. DOI: 10.1007/978-3-540-39521-8.

[8] Rofii F., Naba A., Dharmawan H.A., Hunaini F. (2020). Analysis of Electrical Power Quality Disturbances Based on Empirical Mode Decomposition and Statistical Parameters, IOP Conf. Series: Materials Science and Engineering, 846, 012050, DOI:10.1088/1757-899X/846/1/012050

[9] Abatan, S. M., Olayemi, M. S. (2014). The Role of Statistical Software in Data Analysis. *International Journal of Applied Research and Studies*, 3, 8, 1-15.

[10] IBM SPSS Statistics [Electronic resource] – Access mode: https://www.ibm.com/products/spss-statistics

[11] RStudio IDE [Electronic resource] – Access mode: https://posit.co/products/open-source/rstudio/

[12] IBM SPSS Statistics [Electronic resource] – Access mode: https://uk.wikipedia.org/wiki/RStudio

[13] Stata [Electronic resource] – Access mode: https://www.stata.com/

[14] Minitab [Electronic resource] – Access mode: https://www.minitab.com/en-us/

[15] Rodriguez, D. Using Minitab to achieve Statistical Quality Control [Electronic resource] / D. Rodriguez – Access mode: https://www.invensislearning.com/blog/statistical-quality-control-using-minitab/

[16] Python [Electronic resource] – Access mode: https://www.python.org/

[17] PyCharm [Electronic resource] – Access mode: https://www.jetbrains.com/pycharm/