Чернівецький національний університет імені Юрія Федьковича
Факультет математики та інформатики
Кафедра прикладної математики та інформаційних технологій

“ЗАТВЕРДЖУЮ”
Декан факультету
математики та
інформатики

_____ Ольга МАРТИНЮК

“ 28 ” серпня 2026 року

РОБОЧА ПРОГРАМА
навчальної дисципліни

Прикладний статистичний аналіз даних
(назва навчальної дисципліни)
Вибіркова
(вказати: обов’язкова / вибіркова)
Освітньо-професійна програма Технології програмування та комп’ютерне моделювання
(назва програми)
Спеціальність F1 Прикладна математика
(вказати: код, назва)
Галузь знань F Інформаційні технології

Рівень вищої освіти перший (бакалаврський)

Факультет математики та інформатики
(назва факультету/інституту, на якому здійснюється підготовка фахівців за вказаною освітньо-професійною програмою)

Мова навчання українська
(вказати: на якій мові читається дисципліна)

Чернівці – 2026

Робоча програма навчальної дисципліни "*Прикладний статистичний аналіз даних"* складена відповідно до освітньо-професійної програми «Технології програмування та комп'ютерне моделювання» підготовки здобувачів вищої освіти за спеціальністю 113 Прикладна математика (галузь знань 11 Математика та статистика). Програма є уніфікованою для обох форм навчання і реалізує вибіркові освітні компоненти "Прикладний статистичний аналіз даних" (повна форма навчання, на базі повної загальної середньої освіти) та "Прикладний статистичний аналіз даних" (на базі диплома молодшого бакалавра/ спеціаліста) в обсязі 3,0 кредитів ЄКТС (90 годин).

Розробник: Унгурян Галина Михайлівна канд. – фіз. мат. наук, доцент, асистент кафедри прикладної математики та інформаційних технологій

Викладач, що забезпечує читання даної навчальної дисципліни:
Унгурян Галина Михайлівна, канд. – фіз. мат. наук, асистент кафедри прикладної математики та інформаційних технологій

Погоджено з гарантом ОП _____ Василь МАЦЕНКО

Затверджено на засіданні кафедри прикладної математики та інформаційних технологій
Протокол № 2 від "28" серпня 2026 року

Завідувач кафедри _____ **Ярослав БІГУН**

Схвалено методичною радою факультету математики та інформатики
Протокол № 1 від "28" серпня 2026 року

Голова методичної ради
факультету математики та інформатики _____ **Тоня ФРАТОВЧАН**

Мета навчальної дисципліни «Прикладний статистичний аналіз даних»: сприяти оволодінню студентами теоретичними основами математичної статистики та методами їх практичної реалізації у R для моделювання, візуалізації та аналізу складних наборів даних, а також формувати навички критичної інтерпретації результатів статистичних досліджень та обґрунтованого прийняття рішень на їх основі.

Пререквізити: теорія ймовірностей, програмування, основи математичного моделювання та системного аналізу.

Результати навчання: у результаті вивчення навчальної дисципліни студент буде

знати:

- відповідний теоретичний матеріал в межах програми курсу;
- основи статистики та принципи прикладного аналізу даних;
- структуру, синтаксис та основні принципи роботи в середовищі R;
- основні графічні оболонки та структури мови програмування R, методи графічного відображення даних та результатів аналізу за допомогою ggplot2;
- основні ймовірнісні розподіли та способи роботи з ними у R;
- методи описової статистики та перевірки статистичних гіпотез.

вміти:

- застосовувати практичні методи статистики для аналізу даних, використовувати R для обробки;
- моделювання та візуалізації даних, створювати інформативні графіки та діаграми з використанням ggplot2;
- виконувати описовий та індуктивний статистичний аналіз даних, знаходити коефіцієнти регресійних моделей та проводити їхню діагностику;
- інтерпретувати результати аналізу та приймати обґрунтовані висновки на основі даних.

Знання, які студент отримає в результаті вивчення даної дисципліни, відіграватимуть важливу роль у процесі його професійного формування та зростання, а також вони є необхідною основою для вивчення деяких інших фахових дисциплін.

Студент повинен оволодіти програмним матеріалом, здати колоквіуми, виконати домашні та контрольні завдання.

Вивчення навчальної дисципліни передбачає формування та розвиток у студентів компетентностей, передбачених відповідним стандартом вищої освіти України (за наявності) або ОП, програмні результати навчання якої

відповідають вимогам Національної рамки кваліфікацій (за відсутності стандарту):

- загальних:

ЗК05. Здатність проведення досліджень на відповідному рівні.

ЗК07. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

ЗК08. Знання та розуміння предметної області та розуміння професійної діяльності;

ЗК10. Навички у використанні інформаційних і комунікаційних технологій;

- фахових:

ФК03. Здатність обирати та застосовувати математичні методи для розв'язання прикладних задач, моделювання, аналізу, проектування, керування, прогнозування, прийняття рішень.

ФК09. Здатність до проведення математичного і комп'ютерного моделювання, аналізу та обробки даних, обчислювального експерименту, розв'язання формалізованих задач за допомогою спеціалізованих програмних засобів.

Результати вивчення дисципліни деталізують такі **програмні результати** навчання ОП:

ПРН03. Формалізувати задачі, сформульовані мовою певної предметної галузі; формулювати їх математичну постановку та обирати раціональний метод вирішення; розв'язувати отримані задачі аналітичними та чисельними методами, оцінювати точність та достовірність отриманих результатів.

ПРН14. Виявляти здатність до самонавчання та продовження професійного розвитку.

ПРН15. Уміти організувати власну діяльність та одержувати результат у рамках обмеженого часу

ПРН21. Досліджувати математичні моделі процесів якісними й аналітичними методами, застосовувати програмне забезпечення для їх комп'ютерного моделювання.

Опис навчальної дисципліни
Загальна інформація

Форма навчання	Рік підготовки	Семестр	Кількість		Кількість годин						Вид підсум-кового контролю
			кредитів	годин	лекції	практичні	семінарські	лабораторні	самостійна робота	індивідуальні завдання	
Денна	1	1	3	90	15	–	–	30	45	–	залік
Заочна	–	–	–	–	–	–	–	–	–	–	–

Структура навчальної дисципліни

Назви змістових модулів і тем	Кількість годин											
	денна форма						Заочна форма					
	усього	у тому числі					усього	у тому числі				
		л	п	лаб	інд	с.р.		л	п	лаб	інд	с.р.
1	2	3	4	5	6	7	8	9	10	11	12	13
Теми лекційних занять	**Змістовий модуль 1. Обробка, класифікація та візуалізація даних в R. Статистичні методи та аналіз даних в R.**											
Тема 1. Робота з R та RStudio, встановлення R-пакетів. Основні поняття та робота з даними.	6	1	–	2	–	3	–	–	–	–	–	–
Тема 2. Вектори та робота з ними, функції apply(), supply(), order(). Векторна арифметика.	6	1	–	2	–	3	–	–	–	–	–	–
Тема 3. Пакет даних Tidyverse. Матриці, масиви, фактори. Індексація.	6	1	–	2	–	3	–	–	–	–	–	–
Тема 4. Імпорт даних. Шляхи та робочий каталог. Пакети readr та readxl	6	1	–	2	–	3	–	–	–	–	–	–
Тема 5. Data table. Фрейми даних	6	1	–	2	–	3	–	–	–	–	–	–
Тема 6. Візуалізація	6	1	–	2	–	3	–	–	–	–	–	–

даних. Можливості пакету ggplot2.												
Тема 7. Типи змінних, види графіків, щільність, нормальний розподіл.	6	1	–	2	–	3	–	–	–	–	–	–
Тема 8. Діаграми розсіювання, facet_wrap, графіки часових рядів, ridge plots.	6	1	–	2	–	3	–	–	–	–	–	–
Тема 9. Кодування даних за допомогою візуальних підказок. Принципи побудови якісного графіку, діаграми нахилів.	6	1	–	2	–	3	–	–	–	–	–	–
Разом за змістовим модулем 1	54	9	–	18	–	27	–	–	–	–	–	–
Теми лекційних занять	**Змістовий модуль 2. Аналіз та дослідження даних, підготовка статистичних звітів.**											
Тема 10. Описова статистика в R. Емпіричні комулятивні функції розподілу. Поняття перцентиля, квартиля. Стратифікація.	6	1	–	2	–	3	–	–	–	–	–	–

Тема 11. Описова статистика за допомогою спеціальних пакетів. Функції пакету dplyr.	6	1	–	2	–	3	–	–	–	–	–	–
Тема 12. Поняття ймовірності. Моделювання методом Монте Карло для категоріальних змінних. Побудова гістограм, P-P, Q-Q діаграм.	6	1	–	2	–	3	–	–	–	–	–	–
Тема 13. Функція щільності ймовірності. Метод Монте-Карло для неперервних змінних.	6	1	–	2	–	3	–	–	–	–	–	–
Тема 14. Поняття випадкової величини. Статистичний висновок. IQR.	6	1	–	2	–	3	–	–	–	–	–	–
Тема 15. Лінійна регресія. Множинна лінійна регресія. Діагностика моделі.	6	1	–	2	–	3	–	–	–	–	–	–
Разом за змістовим модулем 2	36			–	–	18	–	–	–	–	–	–
Усього годин	90	15	–	30	–	45	–	–	–	–	–	–

Тематика лекційних занять з переліком питань

№	Назва теми з основними питаннями
1	**Робота з R та RStudio, встановлення R-пакетів. Основні поняття та робота з даними.** 1. Мова програмування R та сфери її застосування. 2. Призначення RStudio та основні елементи інтерфейсу. 3. R-пакети: призначення, встановлення та підключення. 4. Основні типи даних у R. 5. Основні структури даних у R та їхні особливості. **6.** Функції str(), summary(), head() та їх призначення.
2	**Вектори та робота з ними, функції apply(), supply(), order(). Векторна арифметика.** 1. Поняття вектора в R, його основні властивості та особливості. 2. Створення векторів за допомогою функцій c(), seq() та rep(). 3. Способи доступу до елементів вектора та вибір елементів за заданими умовами. 4. Векторна арифметика та виконання арифметичних операцій над векторами. 5. Функції apply() та sapply(): призначення та особливості використання. 6. Функція order() та її застосування для сортування елементів вектора. 7. Основні статистичні операції над векторами та пошук елементів за заданими умовами.
3	**Пакет даних Tidyverse. Матриці, масиви, фактори. Індексація.** 1. Поняття Tidyverse та основні пакети, що входять до його складу. 2. Матриці та масиви в R: призначення та відмінності від векторів. 3. Створення матриць за допомогою функції matrix() та зміна їх розмірності. 4. Фактори (factor): призначення та використання у статистичному аналізі даних.

	5. Індексація в R та способи доступу до елементів векторів, матриць і масивів.
4	**Імпорт даних. Шляхи та робочий каталог. Пакети readr та readxl** 1. Поняття імпорту даних у R та основні способи завантаження даних. 2. Робочий каталог у R: призначення та особливості використання. 3. Функції перегляду та зміни робочого каталогу в R. 4. Пакет readr та його функції для імпорту CSV- і текстових файлів. 5. Пакет readxl та його можливості для імпорту даних з Excel-файлів.
5	**Data table. Фрейми даних** 1. Особливості організації та структури даних у форматі data.frame. 2. Рядки та стовпці фрейму даних: призначення та вимоги до їх структури. 3. Перейменування стовпців фрейму даних і додавання нових змінних. 4. Способи об'єднання фреймів даних за рядками та стовпцями. 5. Основні відмінності між data.frame і data.table.
6	**Візуалізація даних. Можливості пакету ggplot2.** 1. Призначення візуалізації даних у статистичному аналізі та основні завдання, які вона допомагає розв'язувати. 2. Основні типи графіків для відображення статистичних даних та особливості їх застосування. 3. Основні компоненти графіків у пакеті ggplot2. 4. Функції ggplot(), aes() та geom_point(): призначення та особливості використання. 5. Побудова лінійних графіків, гістограм і діаграм розмаху за допомогою ggplot2. 6. Налаштування оформлення графіків: заголовки, підписи осей, кольори та теми.
7	**Типи змінних, види графіків, щільність, нормальний розподіл.** 1. Призначення візуалізації даних у статистичному аналізі. 2. Основні типи графіків та особливості їх застосування.

	3. Компоненти графіків у пакеті ggplot2. 4. Призначення та використання функцій ggplot(), aes() і geom_point(). 5. Побудова лінійних графіків, гістограм і діаграм розмаху засобами ggplot2. 6. Налаштування зовнішнього вигляду графіків: заголовки, підписи осей, кольори та теми.
8	**Діаграми розсіювання, facet_wrap, графіки часових рядів, ridge plots.** 1. Діаграма розсіювання та її застосування для дослідження залежностей між кількісними змінними. 2. Визначення напрямку, сили та характеру зв'язку між змінними за діаграмою розсіювання. 3. Функція facet_wrap() у пакеті ggplot2 та її використання для порівняння груп даних. 4. Графіки часових рядів: особливості побудови та аналізу змін показників у часі. 5. Підготовка даних для побудови графіків часових рядів засобами ggplot2. 6. Ridge plots: призначення та використання для порівняння розподілів даних між групами. 7. Побудова та налаштування ridge plots за допомогою пакету ggridges.
9	**Кодування даних за допомогою візуальних підказок. Принципи побудови якісного графіку, діаграми нахилів.** 1. Типові помилки під час побудови графіків та способи їх уникнення. 2. Роль масштабу осей, підписів, заголовків і легенди у створенні якісної візуалізації. 3. Діаграма нахилів (slope chart): призначення та особливості використання для аналізу змін у даних. 4. Побудова діаграми нахилів за допомогою ggplot2 та інтерпретація отриманих результатів.
10	**Описова статистика в R. Емпіричні комулятивні функції розподілу. Поняття перцентиля, квартиля. Стратифікація.**

	1. Основні показники описової статистики та їх обчислення в R. 2. Показники центральної тенденції та мінливості даних: призначення і відмінності. 3. Емпірична кумулятивна функція розподілу та її використання для аналізу вибірки. 4. Перцентилі та квартилі: призначення та визначення за допомогою функції quantile(). 5. Міжквартильний розмах: оцінювання розкиду даних та виявлення викидів. 6. Стратифікація даних: призначення та аналіз статистичних показників для окремих груп спостережень у R.
11	**Описова статистика за допомогою спеціальних пакетів. Функції пакету dplyr.** 1. Переваги використання спеціалізованих пакетів порівняно з базовими функціями R для статистичного аналізу даних. 2. Призначення та використання функцій filter() і select() пакету dplyr. 3. Створення нових змінних на основі наявних стовпців таблиці за допомогою функції mutate(). 4. Групування та узагальнення даних за допомогою функцій group_by() і summarise().
12	**Поняття ймовірності. Моделювання методом Монте Карло для категоріальних змінних. Побудова гістограм, P-P, Q-Q діаграм.** 1. Поняття ймовірності випадкової події та основні підходи до її визначення. 2. Метод Монте-Карло: принципи та особливості моделювання випадкових подій. 3. Моделювання випадкових спостережень для категоріальних змінних у R та оцінювання ймовірностей їхніх значень. 4. Гістограма: призначення та відмінності від стовпчикової діаграми для категоріальних даних. 5. P-P діаграма: призначення та використання для порівняння емпіричного й теоретичного розподілів. 6. Q-Q діаграма: призначення та оцінювання відповідності вибіркових

	даних нормальному розподілу.
13	**Функція щільності ймовірності. Метод Монте-Карло для неперервних змінних.** 1. Метод Монте-Карло та його застосування для моделювання неперервних випадкових величин. 2. Генерування випадкових значень із рівномірного, нормального та експоненціального розподілів у R. 3. Наближене обчислення математичного сподівання та ймовірності подій методом Монте-Карло. 4. Вплив кількості повторень моделювання на точність результатів методу Монте-Карло.
14	**Поняття випадкової величини. Статистичний висновок. IQR.** 1. Поняття випадкової величини та відмінності між дискретними й неперервними випадковими величинами. 2. Закон розподілу випадкової величини та основні характеристики його опису. 3. Статистичний висновок: сутність, основні етапи формування та інтерпретація результатів. 4. Описова статистика та статистичний висновок: основні відмінності. 5. Міжквартильний розмах (IQR): поняття та способи обчислення за вибірковими даними. 6. Виявлення потенційних викидів за допомогою IQR та визначення меж їх виявлення.
15	**Лінійна регресія. Множинна лінійна регресія. Діагностика моделі.** 1. Сутність лінійної регресії та її застосування для дослідження залежностей між змінними. 2. Інтерпретація коефіцієнтів регресії та вільного члена моделі. 3. Метод найменших квадратів та його використання для оцінювання параметрів лінійної регресії. 4. Множинна лінійна регресія: відмінності від простої лінійної регресії та оцінювання впливу кількох незалежних змінних на результат.

Тематика практичних занять з переліком питань

Навчальним планом та робочою програмою з навчальної дисципліни «Прикладний статистичний аналіз даних» виконання практичних завдань не передбачено.

Теми лабораторних занять

№	Назва теми	Зміст роботи та перелік завдань
1	**Основні структурні одиниці середовища програмування R та операції з ними.**	1. Основні структурні одиниці мови програмування R. 2. Поняття змінної та способи її створення в R. 3. Правила іменування змінних у R. 4. Основні типи даних у R та функції для визначення типу значень. 5. Арифметичні оператори в R та порядок виконання операцій. 6. Оператори порівняння та логічні оператори: призначення, відмінності й особливості використання. 7. Функції class(), typeof(), length() та str() для дослідження властивостей об’єктів у R.
2	**Фрейми даних**	1. Поняття фрейму даних (data.frame) та його призначення в R. 2. Створення фреймів даних із векторів різних типів.

		3. Способи доступу до окремих рядків, стовпців та елементів фрейму даних. 4. Додавання нових стовпців і зміна значень наявних стовпців фрейму даних. 5. Основні функції для перегляду структури, розмірності та вмісту фрейму даних. 6. Відбір рядків за заданими умовами та сортування записів за значеннями стовпців. 7. Об’єднання фреймів даних за рядками та стовпцями за допомогою функцій rbind() і cbind().
3		1. Основні джерела даних для імпорту інформації в середовище R. 2. Поширені формати файлів для збереження статистичних даних. 3. Визначення та зміна робочого каталогу в RStudio. 4. Абсолютні та відносні шляхи до файлів: особливості та відмінності. 5. Імпорт даних із CSV-файлів за допомогою функцій read.csv() та read_csv(). 6. Основні параметри імпорту

		даних: роздільники, кодування, десятковий знак і назви стовпців. 7. Перевірка правильності імпорту та виявлення можливих помилок у завантажених даних.
4	**Робота з пакетом даних ggplot2**	1. Призначення пакета ggplot2 та його можливості для візуалізації статистичних даних. 2. Підключення пакета ggplot2 та використання вбудованих наборів даних для побудови графіків. 3. Призначення та особливості використання функцій ggplot() і aes() під час створення графіків. 4. Побудова різних типів графіків за допомогою геометричних шарів geom_point(), geom_line() та geom_bar(). 5. Налаштування кольору, форми та розміру графічних елементів залежно від значень змінних. 6. Оформлення графіків за допомогою функцій labs(), scale_*() та theme(). 7. Відображення кількох груп даних на одному графіку та налаштування легенди.
5	**Візуалізація та аналіз даних.**	1. Роль візуалізації у

	Побудова різних видів графіків.	дослідженні та аналізі статистичних даних. 2. Вибір типу графіка залежно від структури даних і мети дослідження. 3. Призначення та особливості використання гістограм, діаграм розмаху і стовпчикових діаграм. 4. Дослідження взаємозв'язку між числовими змінними за допомогою діаграми розсіювання. 5. Особливості побудови графіків для порівняння кількох груп спостережень. 6. Виявлення аномальних значень, тенденцій і закономірностей засобами графічного аналізу. 7. Можливості мови R та пакета ggplot2 для побудови й налаштування статистичних графіків.
6	**Описова статистика та аналіз розподілу даних.**	1. Основні характеристики вибірки та їх використання для опису статистичних даних. 2. Обчислення середнього арифметичного, медіани та моди засобами R. 3. Показники мінливості даних та їх статистична інтерпретація. 4. Визначення мінімального, максимального значень і розмаху вибірки.

		5. Аналіз форми та особливостей розподілу даних за допомогою гістограми. 6. Визначення симетричності та асиметричності розподілу за його графічною формою. 7. Використання функцій summary(), sd(), var(), quantile() та IQR() для описового статистичного аналізу в R.
7	**Поняття ймовірності у мові R.**	1. Поняття ймовірності випадкової події та її основні властивості. 2. Класичне, статистичне та геометричне трактування ймовірності. 3. Моделювання випадкових експериментів у R за допомогою функцій sample() та replicate(). 4. Обчислення відносної частоти настання випадкової події за результатами багаторазових випробувань. 5. Дослідження збіжності експериментальної ймовірності до теоретичного значення зі збільшенням кількості випробувань. 6. Використання функцій set.seed() та sample() для імітаційного моделювання

		в R. 7. Графічне дослідження зміни відносної частоти події залежно від кількості проведених випробувань.
8	**Метод Монте-Карло та Центральна гранична теорема.**	1. Основна ідея методу Монте-Карло та його застосування для розв'язання статистичних задач. 2. Вплив кількості повторень випадкового експерименту на точність результатів моделювання. 3. Оцінювання ймовірності випадкової події за допомогою генерації випадкових чисел у R. 4. Центральна гранична теорема: основні положення та умови застосування. 5. Наближення розподілу вибіркових середніх до нормального зі збільшенням обсягу вибірки. 6. Зміна форми розподілу вибіркових середніх залежно від обсягу вибірки. 7. Моделювання вибірок у R, дослідження розподілу їх середніх та графічна демонстрація центральної граничної теореми.

9	**Дослідження статистичних методів, кореляції, регресії та центральної граничної теореми в середовищі R**	1. Основні статистичні методи дослідження взаємозв'язків між змінними та аналізу вибіркових даних. 2. Поняття кореляції та інтерпретація додатного, від'ємного і нульового коефіцієнтів кореляції. 3. Відмінності між кореляційним і причинно-наслідковим зв'язками. 4. Обчислення коефіцієнта кореляції Пірсона та перевірка його статистичної значущості за допомогою функцій cor() і cor.test() у R. 5. Побудова лінійної регресійної моделі за допомогою функції lm() та оцінювання впливу незалежної змінної на залежну. 6. Оцінювання якості регресійної моделі та інтерпретація її основних діагностичних показників. 7. Імітаційне моделювання в R для демонстрації центральної граничної теореми та дослідження розподілу вибіркових середніх.

Індивідуальні науково-дослідні завдання (ІНДЗ)

Навчальним планом та робочою програмою з навчальної дисципліни «Прикладний статистичний аналіз даних» виконання ІНДЗ не передбачено.

Завдання для самостійної роботи студентів

№	Назва теми	Завдання для самостійної роботи	Кількість годин
1	**Попередня обробка та очищення статистичних даних у R**	1. Завантажити набір даних із відкритого джерела або використати вбудований набір даних R. Провести перевірку його структури та типів змінних. 2. Виявити пропущені значення за допомогою функцій is.na() та colSums(). Запропонувати способи їх обробки. 3. Знайти дублікати записів та вилучити їх із набору даних. 4. Виконати перевірку даних на наявність некоректних значень і підготувати очищений набір для подальшого аналізу.	3
2	**Об'єднання та перетворення наборів даних у R**	1. Створити або завантажити два набори даних, які містять спільний ідентифікатор. 2. Виконати об'єднання таблиць за допомогою функцій merge(), left_join() та inner_join(). 3.Дослідити, як змінюється кількість записів залежно від способу об'єднання таблиць. 4. Виконати перетворення структури даних за допомогою функцій pivot_longer() та pivot_wider() пакета tidyr.	3
3	**Статистичне дослідження категоріальних даних у R**	1. Сформувати таблицю частот для категоріальної змінної за допомогою функцій table() та prop.table(). 2. Побудувати таблиці спряженості	3

		для двох категоріальних змінних. 3. Визначити найчастіше та найрідше представлені категорії у вибірці. 4. Побудувати стовпчикові діаграми для порівняння частот і відносних частот категорій за допомогою ggplot2.	
4	**Вибіркові характеристики та моделювання статистичних вибірок у R**	1. Сформувати декілька випадкових вибірок різного обсягу з одного набору даних за допомогою функції sample(). 2. Обчислити вибіркові середні, дисперсії та стандартні відхилення для кожної вибірки. 3. Дослідити, як зміна обсягу вибірки впливає на отримані статистичні характеристики. 4. Порівняти результати вибіркового аналізу з характеристиками всієї генеральної сукупності.	3
5	**Перевірка статистичних гіпотез у середовищі R**	1. Сформулювати нульову та альтернативну статистичні гіпотези для заданої дослідницької задачі. 2. Виконати перевірку гіпотези про середнє значення за допомогою функції t.test(). 3. Провести порівняння середніх значень двох незалежних вибірок. 4. Проаналізувати отримане p-значення та сформулювати висновок щодо прийняття або відхилення нульової гіпотези.	3
6	**Довірчі інтервали та**	1. Дослідити поняття точкової та інтервальної оцінок статистичних	3

	оцінювання статистичних параметрів у R	параметрів. 2. Побудувати довірчий інтервал для середнього значення вибірки. 3. Дослідити вплив обсягу вибірки та рівня довіри на ширину довірчого інтервалу. 4. Провести порівняння довірчих інтервалів для декількох незалежних вибірок та інтерпретувати результати.	
7	**Непараметричні методи статистичного аналізу в R**	1. Дослідити відмінності між параметричними та непараметричними статистичними методами. 2. Виконати порівняння двох незалежних вибірок за допомогою критерію Вілкоксона (wilcox.test()). 3. Провести перевірку відповідності вибіркових даних заданому розподілу за допомогою критерію Колмогорова – Смирнова (ks.test()). 4.Порівняти результати параметричного та непараметричного аналізу для одного набору даних.	3
8	**Аналіз часових рядів та прогнозування в R**	1. Завантажити набір даних, що містить спостереження за певним показником у часі. 2. Перетворити дані у формат часового ряду за допомогою функції ts(). 3. Побудувати графік часового ряду та дослідити наявність тенденції, сезонності й випадкових коливань. 4. Виконати просте прогнозування майбутніх значень за допомогою	3

		функції forecast() або методів пакета forecast.	
9	**Аналіз залежностей між категоріальними змінними в R**	1. Сформувати таблицю спряженості для двох категоріальних змінних. 2. Обчислити частоти та відносні частоти для комбінацій категорій. 3. Перевірити наявність статистично значущого зв'язку між змінними за допомогою критерію хі-квадрат (chisq.test()). 4.Побудувати графічне представлення результатів та сформулювати висновки щодо взаємозв'язку змінних.	3
10	**Створення статистичного звіту за результатами аналізу даних у R**	1. Обрати набір статистичних даних та визначити мету його дослідження. 2. Провести попередню обробку даних і виконати розрахунок основних описових статистичних показників. 3. Побудувати декілька графіків для відображення основних закономірностей у даних. 4. Підготувати статистичний звіт із таблицями, графіками, результатами обчислень та висновками за допомогою R Markdown або Quarto.	3
11	**Логічні операції, умови та цикли в мові програмування R**	1. Які логічні оператори використовуються в R та як вони застосовуються для перевірки умов? 2. Як реалізувати розгалуження за допомогою конструкцій if, else та ifelse()? 3. Як організувати повторення обчислень за допомогою циклів for,	3

		while та repeat? 4. Створити програму для обробки вектора числових значень із використанням умовних конструкцій і циклів.	
12	**Функції користувача та програмування в R**	1. Що таке функція користувача та з яких основних частин вона складається? 2. Як створити власну функцію з одним або декількома аргументами? 3. Для чого використовуються аргументи за замовчуванням та оператор return()? 4. Створити власну функцію для обчислення основних статистичних характеристик числового вектора.	3
13	**Аналіз багатовимірних статистичних даних у R**	1. Що таке багатовимірні статистичні дані та які особливості їх аналізу? 2. Як обчислити матрицю кореляцій для декількох числових змінних? 3. Як побудувати кореляційну матрицю та представити її у вигляді теплової карти (heatmap)? 4. Провести дослідження взаємозв'язків між декількома змінними та сформулювати висновки за результатами аналізу.	3
14	**Перевірка статистичних розподілів та критерії узгодженості в R**	1. Які методи використовуються для перевірки відповідності вибіркових даних теоретичному розподілу? 2. Як перевірити нормальність розподілу за допомогою критерію Шапіро – Вілка (shapiro.test())?	3

		3. Для чого використовується критерій хі-квадрат (chisq.test()) під час аналізу статистичних даних? 4. Провести перевірку відповідності вибірки нормальному розподілу та порівняти результати статистичних критеріїв із графічним аналізом.	
15	**Прогнозування та оцінювання результатів регресійного аналізу в R**	1. Що таке прогнозування на основі регресійної моделі та які його основні завдання? 2. Як отримати прогнозні значення залежної змінної за допомогою функції predict()? 3. Як побудувати довірчі та прогнозні інтервали для результатів регресійної моделі? 4. Провести прогнозування значень досліджуваного показника та порівняти прогнозні результати з фактичними даними.	3

Методи навчання

У процесі вивчення навчальної дисципліни використовуються інноваційні освітні технології: лекції та їх презентації, практичні та лабораторні заняття, групова робота при розв’язуванні студентами практичних завдань, онлайн-тестування та опитування. Використання системи MOODLE, електронні підручники і ресурси репозитарію ЧНУ.
Система контролю та оцінювання

Методи контролю:

- усне опитування з теоретичних питань за темами навчальної дисципліни;
- перевірка засвоєння основних понять, методів і засобів статистичного аналізу даних мовою R;
- тестування з використанням системи дистанційного навчання Moodle;

- перевірка правильності виконання програмного коду та результатів статистичних обчислень;
- оцінювання вміння застосовувати функції та пакети R для обробки, аналізу й візуалізації даних;
- аналіз самостійності виконання практичних завдань та обґрунтованості отриманих результатів.

Форми контролю:

- поточний контроль (усне опитування, тестування, перевірка виконання лабораторних робіт);
- контроль самостійної роботи студентів (виконання практичних завдань, опрацювання теоретичного матеріалу, підготовка звітів);
- модульний контроль (виконання модульних контрольних робіт, тестових завдань із програмування мовою R);
- підсумковий контроль — залік.

Критерії оцінювання поточного та підсумкового контролю
Розподіл балів, які отримують студенти

<table>
<tr><td colspan="9">Поточне оцінювання (аудиторна та самостійна робота)</td><td>Кількість балів</td><td>Сумарна кількість балів</td></tr>
<tr><td colspan="5">Змістовний модуль №1</td><td colspan="4">Змістовний модуль №2</td><td rowspan="3">30</td><td rowspan="3">100</td></tr>
<tr><td>Т1</td><td>Т2</td><td>Т3</td><td>Т4</td><td>Т5</td><td>Т6</td><td>Т7</td><td>Т8</td><td>Т9</td></tr>
<tr><td>5</td><td>5</td><td>5</td><td>5</td><td>5</td><td>6</td><td>7</td><td>6</td><td>6</td></tr>
</table>

Шкала оцінювання: національна та ЄКТС

<table>
<tr><td rowspan="2">Оцінка за національною шкалою</td><td colspan="2">Оцінка за шкалою ECTS</td></tr>
<tr><td>Оцінка (бали)</td><td>Пояснення за розширеною шкалою</td></tr>
<tr><td>Відмінно</td><td>A (90-100)</td><td>відмінно</td></tr>
<tr><td rowspan="2">Добре</td><td>B (80-89)</td><td>дуже добре</td></tr>
<tr><td>C (70-79)</td><td>добре</td></tr>
<tr><td rowspan="2">Задовільно</td><td>D (60-69)</td><td>задовільно</td></tr>
<tr><td>E (50-59)</td><td>достатньо</td></tr>
<tr><td rowspan="2">Незадовільно</td><td>FX (35-49)</td><td>(незадовільно)</td></tr>
<tr><td>F (1-34)</td><td>(незадовільно)</td></tr>
</table>

Перелік питань для самоконтролю та підсумкового контролю навчальних досягнень студентів

1. Середовище R та RStudio.
2. Встановлення та підключення пакету у R, і чим відрізняються функції install.packages() та library()?
3. Вектори у R та способи його створення.
4. Для чого використовуються функції apply(), sapply() та order()?
5. Основні пакети, що входять до складу Tidyverse та для чого вони використовуються?
6. Чим відрізняються матриця, масив і фактор у R?
7. Робочі каталоги, його перегляди та зміни в R?
8. Переваги пакетів readr та readxl для імпорту даних.
9. Чим відрізняється структура data.table від data.frame?
10. Основні операції над фреймами даних у R?
11. Основні компоненти графіка в ggplot2.
12. Принцип пошарового створення графіків у ggplot2.
13. Основні типи змінних, що використовуються у статистичному аналізі.
14. Властивості нормального розподілу і його візуалізація.
15. Для яких задач використовується діаграма розсіювання?
16. Призначення функції facet_wrap() у ggplot2?
17. Візуальні характеристики (колір, форма, розмір тощо) для кодування даних.
18. Які принципи необхідно враховувати під час побудови якісного графіка?
19. Поняття квартиля та перцентиля.
20. Для чого застосовується стратифікація даних у статистичному аналізі?
21. Основні функції пакета dplyr використання для відбору та групування даних.
22. Чим відрізняються функції mutate() та summarise()?
23. Суть моделювання методом Монте-Карло.
24. Для чого використовуються P-P та Q-Q діаграми?
25. Функція щільності ймовірності та її основні властивості.
26. Як застосовується метод Монте-Карло для моделювання неперервних випадкових величин?
27. Поняття випадкової величини та її основні види.
28. Поняття міжквартильного розмаху (IQR) і приципи його використання.

Зарахування результатів неформальної освіти

Зарахування результатів неформальної освіти У межах поточного контролю можуть визнаватися результати неформальної освіти за умови наявності сертифікату або освітньої декларації про результати неформальної освіти з питань, що відповідає тематиці курсу відповідно до Порядку

визнання у Чернівецькому національному університеті імені Юрія Федьковича результатів навчання, здобутих шляхом неформальної та/або інформальної освіти уведеному в дію наказом ректора № 422 від 27.11.2024 р.

Студентам можуть бути зараховані додаткові бали, отримані через неформальну освіту, до загальної суми балів, набраної з начальної дисципліни, за умови, що результати з теми, за якою відбувалося навчання, відповідають тематиці курсу «Прикладний статистичний аналіз даних» та отримані на ліцензійних платформах.

Можливі напрями неформальної освіти:

- проходження онлайн-курсів із програмування мовою R та використання середовища RStudio;
- опанування курсів зі статистичного аналізу та обробки даних;
- проходження навчальних програм із використання пакетів tidyverse, dplyr, ggplot2;
- вивчення методів візуалізації даних, кореляційного та регресійного аналізу;
- участь у практичних тренінгах і вебінарах із застосування R для аналізу статистичних даних.

Підставою для розгляду питання про зарахування результатів можуть бути сертифікати, свідоцтва, документи про проходження навчання, виконані практичні завдання або інші підтвердження набутих компетентностей.

Рішення щодо зарахування результатів неформальної та інформальної освіти, визначення відповідності набутих результатів навчання програмним результатам дисципліни та обсягу їх зарахування приймається в установленому закладом вищої освіти порядку.

Рекомендована література

1. Wickham H., Çetinkaya-Rundel M., Grolemund G. R for Data Science: Import, Tidy, Transform, Visualize, and Model. 2nd ed. O'Reilly Media, 2023.

2. Kabacoff R. I. R in Action: Data Analysis and Graphics with R and Tidyverse. 3rd ed. Manning Publications, 2022.

3. Dalgaard P. Introductory Statistics with R. 2nd ed. Springer, 2008. 364 p.

Допоміжна література та електронні ресурси

1. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

2. Wickham H. ggplot2: Elegant Graphics for Data Analysis. 3rd ed. Springer, 2016.

3. Wickham H., François R., Henry L., Müller K., Vaughan D. dplyr: A Grammar of Data Manipulation.

Інформаційні ресурси:

1. Сайт дистанційного навчання факультету математики та інформатики Чернівецького національного університету імені Юрія Федьковича:

https://moodle.chnu.edu.ua/course/view.php?id=8862
2. Персональна сторінка викладача:
https://amit.chnu.edu.ua/pro-kafedru/personalii/unhurian-halyna-mykhailivna/

Політика академічної доброчесності

Здобувач зобов'язаний своєчасно та якісно виконувати всі отримані завдання; за необхідності, з метою з'ясування всіх незрозумілих під час самостійної та індивідуальної роботи питань, відвідувати консультації викладача. Учасники освітнього процесу зобов'язані дотримуватись політики щодо академічної доброчесності при вивченні освітнього компонента регламентовано чинними документами ЧНУ:

- «Етичний кодекс Чернівецького національного університету імені Юрія Федьковича» https://www.chnu.edu.ua/universytet/normatyvni-dokumenty/etychnyi-kodeks chernivetskoho-natsionalnoho-universytetu-imeni-yuriia-fedkovycha/
- «Положенням про виявлення та запобігання академічного плагіату у Чернівецькому національному університеті імені Юрія Федьковича» https://www.chnu.edu.ua/universytet/normatyvni-dokumenty/polozhennia-pro-vyiavlennia-ta zapobihannia-akademichnomu-plahiatu-u-chernivetskomu-natsionalnomu-universyteti-imeni yuriia-fedkovycha/
- «Політика використання штучного інтелекту в Чернівецькому національному університету імені Юрія Федьковича» https://www.chnu.edu.ua/universytet/normatyvni dokumenty/polityka-vykorystannia-shtuchnoho-intelektu-v-chernivetskomu-natsionalnomu universyteti-imeni-yuriia-fedkovycha/

Наявність академічного плагіату в студентських роботах є підставою для виставлення негативної оцінки. Списування студентів під час проходження тестування є підставою для дострокового припинення його складання та виставлення негативної оцінки.

Використання здобувачами освіти інструментів штучного інтелекту (ШІ) під час виконання навчальних завдань:

- *допускається за умови* дотримання принципів академічної доброчесності (здобувач освіти несе відповідальність за зміст, достовірність і результати поданої роботи, а у разі використання ШІ зобов’язаний зазначити факт і спосіб його застосування);
- *не допускається* використання ШІ як заміни самостійного виконання завдань або з метою порушення вимог академічної доброчесності;
- *додаткові обмеження чи вимоги* щодо використання ШІ можуть визначатися викладачем в освітньому процесі для конкретних видів навчальної діяльності відповідно до особливостей дисципліни.