

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
ІМЕНІ ТАРАСА ШЕВЧЕНКА  
ФАКУЛЬТЕТ КОМП'ЮТЕРНИХ НАУК ТА КІБЕРНЕТИКИ  
Кафедра прикладної статистики**

**«ЗАТВЕРДЖУЮ»**

Заступник декана  
з навчальної роботи

\_\_\_\_\_ Людмила ОМЕЛЬЧУК  
«\_\_» \_\_\_\_\_ 20\_\_ року

## **РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ**

### **Візуалізація даних та елементи статистичного моделювання в R**

для студентів

галузь знань	<b>F Інформаційні технології</b>
спеціальність	<b>F4 Системний аналіз та наука про дані</b>
освітній рівень	<b>бакалавр</b>
освітня програма	<b>Системний аналіз</b>
вид дисципліни	<b>вільного вибору студента</b>

Форма навчання	<b>денна</b>
Навчальний рік	<b>2025/2026</b>
Семестр	<b>3</b>
Кількість кредитів ECTS	<b>3</b>
Мова викладання, навчання та оцінювання	<b>українська</b>
Форма заключного контролю	<b>залік</b>

Викладачі: **к.ф.-м.н, доцент Ганна ЛІВІНСЬКА**

Пролонговано: на 20\_\_/20\_\_ н.р. \_\_\_\_\_ (\_\_\_\_\_) «\_\_» 20\_\_ р.

на 20\_\_/20\_\_ н.р. \_\_\_\_\_ (\_\_\_\_\_) «\_\_» 20\_\_ р.

**КИЇВ – 2025**

Розробник: **Ганна ЛІВІНСЬКА** канд. фіз.-мат. н., доцент кафедри прикладної статистики

ЗАТВЕРДЖЕНО  
Зав. кафедри Прикладної Статистики  
\_\_\_\_\_ (Ірина РОЗОРА)

Протокол № \_\_ від «\_\_» \_\_\_\_\_ 20\_\_ р.

Схвалено. Гарант освітньо-професійної програми першого рівня вищої освіти "Системний аналіз"

\_\_\_\_\_ Михайло ШАРАПОВ

Протокол від «\_\_\_\_» \_\_\_\_\_ 20\_\_ року №\_\_

Голова науково-методичної комісії \_\_\_\_\_ Тетяна КАРНАУХ

(підпис)

**1. Мета дисципліни** – засвоєння студентами основних методів та засобів візуалізації даних, базового статистичного аналізу та моделювання за допомогою інструментів мови програмування R; ознайомлення з інструментами для створення інтерактивних візуалізацій та автоматизованого оформлення звітів у R Markdown з інтеграцією графіків та моделей; розвиток практичних навичок аналізу реальних даних, відтворення результатів досліджень та критичного мислення.

## **2. Попередні вимоги до опанування або вибору навчальної дисципліни**

- *Знати*: основи програмування.
- *Вміти*: розробляти програми на базовому рівні.
- *Володіти елементарними навичками*: елементарними навичками програмування.

## **3. Анотація навчальної дисципліни**

Навчальна дисципліна "Візуалізація даних та елементи статистичного моделювання в R" є складовою освітньо-професійної програми підготовки фахівців за першим (бакалаврським) рівнем вищої освіти *галузі знань* F "Інформаційні технології" зі спеціальності F4 "Системний аналіз та наука про дані", *освітньо-професійної програми* – "Системний аналіз". Дисципліна має такі розділи: 1. *Основи візуалізації даних з R*. 2. *Основи статистичного моделювання з R*. 3. *Елементи моделювання випадкових процесів з R*. Дисципліна є дисципліною вільного вибору студента. Використовує поняття з базового курсу програмування.

Викладається в 3-му семестрі, **обсяг – 90 год. (3 кредити ECTS)**, з них *лекції – 14 год., лабораторні – 14 год., консультації – 2 год., самостійна робота – 60 год.* У курсі передбачено **3 змістовні частини та 3 лабораторні роботи з тестовим контролем знань.**

Завершується дисципліна – **заліком.**

В результаті вивчення навчальної дисципліни студент повинен

– **знати**: базові принципи аналізу статистичної інформації, основні способи візуалізації даних різних типів, принципи побудови статистичних графіків та методи їх реалізації в середовищі R, базові підходи до статистичного моделювання з використанням R.

– **вміти**: користуватися широким спектром методів мови програмування R для візуалізації та різнопланового аналізу реальних даних за допомогою отриманих графіків, будувати базові статистичні моделі, готувати аналітичні звіти.

## **4. Завдання (навчальні цілі)**

набуття знань, умінь та навичок (компетенцій) на рівні новітніх досягнень у сфері базової обробки, візуалізації та первинного аналізу даних різної структури, набуття досвіду програмування мовою R при розв'язанні прикладних задач. Зокрема розвивати:

- володіння широким арсеналом методів та засобів візуалізації даних;
- здатність застосовувати знання для вирішення практичних задач;
- здатність оцінювати та забезпечувати якість виконуваних робіт;
- здатність вірно визначити тип та структуру даних та обрати інформативні методи їх візуалізації, проєктувати етапи необхідного аналізу, реалізовувати відповідні процедури візуалізації на комп'ютері за допомогою мови програмування R, надати трактовку отриманим результатам візуалізації.

В результаті вивчення навчальної дисципліни студент повинен мати:

**ЗК01.** Здатність до абстрактного мислення, аналізу та синтезу.

**ЗК02.** Здатність застосовувати знання у практичних ситуаціях.

**ЗК04.** Знання та розуміння предметної області та розуміння професійної діяльності.

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

**ЗК11.** Здатність генерувати нові ідеї (креативність).

**СК1.** Здатність використовувати системний аналіз як сучасну міждисциплінарну методологію, що базується на прикладних математичних методах та сучасних інформаційних

технологіях і орієнтована на вирішення задач аналізу і синтезу технічних, економічних, соціальних, екологічних та інших складних систем.

**СК4.** Здатність визначати основні чинники, які впливають на розвиток фізичних, економічних, соціальних процесів, виокремлювати в них стохастичні та невизначені показники, формулювати їх у вигляді випадкових або нечітких величин, векторів, процесів та досліджувати залежності між ними.

**СК7.** Здатність використовувати сучасні інформаційні технології для комп'ютерної реалізації математичних моделей та прогнозування поведінки конкретних систем а саме: об'єктно-орієнтований підхід при проектуванні складних систем різної природи, прикладні математичні пакети, застосування баз даних і знань.

## 5. Результати навчання за дисципліною

**ПРН12.** Застосовувати методи і засоби роботи з даними і знаннями, методи математичного, логіко-семантичного, об'єктного та імітаційного моделювання, технології системного і статистичного аналізу.

**ПРН14.** Розуміти і застосовувати на практиці методи статистичного моделювання і прогнозування, оцінювати вихідні дані.

<b>Результат навчання (РН)</b> (1. – знати; 2. – вміти; 3. – комунікація; 4. – автономність та відповідальність)		<b>Форми (та/або методи і технології) викладання та навчання</b>	<b>Методи оцінювання та пороговий критерій оцінювання (за необхідності)</b>	<b>Відсоток у підсумковій оцінці з дисципліни</b>
<b>Код</b>	<b>Результат навчання</b>			
РН.1.	Знати і розуміти основні підходи до візуалізації даних різної природи, інструменти побудови основних діаграм за допомогою середовища R.	Лекції, практичні заняття	Захист лабораторних робіт, поточне оцінювання	30%
РН.2.1	Вміти для реальних даних визначити їх тип та спектр методів їх візуалізації, побудувати коректні діаграми, проаналізувати отримані зображення, визначити задачі подальшого статистичного аналізу.	Лекції, практичні заняття, самостійна робота	Захист лабораторних робіт	20%
РН.2.2	Вміти будувати базові діаграми відповідно до типу даних за допомогою мови програмування R.	Практичні заняття, самостійна робота	Захист лабораторних робіт	20%
РН.2.3.	Вміти моделювати вибірку різної природи, прості стохастичні моделі за допомогою мови програмування R..	Лекції, практичні заняття, самостійна робота	Захист лабораторних робіт	20%
РН.2.4.	Вміти створювати статистичні звіти за допомогою R Markdown.	Практичні заняття, самостійна робота	Захист лабораторних робіт	10%
РН.3.1	Обґрунтовувати власний погляд на задачу, спілкуватися з колегами з питань проектування та розробки програм, складати письмові звіти.	Практичні заняття, самостійна робота	Захист лабораторних робіт, поточне оцінювання	10%

РН.4.1	Уміти організувати самостійну роботу та одержувати результат у рамках обмеженого часу.	Самостійна робота	Захист лабораторних робіт, поточне оцінювання	10%
--------	--	-------------------	---	-----

## 6 Співвідношення результатів навчання дисципліни з програмними результатами навчання

Результати навчання дисципліни	РН.1	РН.2.1	РН.2.2	РН.2.3	РН.2.4	РН.3.1	РН.4.1
Програмні результати навчання							
<i>(з опису освітньої програми)</i>							
<b>ПРН12.</b> Застосовувати методи і засоби роботи з даними і знаннями, методи математичного, логіко-семантичного, об'єктного та імітаційного моделювання, технології системного і статистичного аналізу.	+	+	+	+	+	+	+
<b>ПРН14.</b> Розуміти і застосовувати на практиці методи статистичного моделювання і прогнозування, оцінювати вихідні дані.	+	+		+		+	+

## 7. Схема формування оцінки

**7.1 Форми оцінювання студентів:** Контроль знань студентів здійснюється за кредитно-рейтинговою системою. Результати навчальної діяльності студентів оцінюються за 100-бальною шкалою.

**Семестрове оцінювання.** Робота в семестрі складається з 3-х змістовних частин. При виставленні балів за змістовну частину враховується: робота студентів на заняттях (до 10 балів), виконання лабораторних робіт та розуміння теоретичного матеріалу.

1. Лабораторна робота 1 + тематичний тест:– 20 балів.
2. Лабораторна робота 2 + тематичний тест:– 30 балів.
3. Лабораторна робота 3 + тематичний тест:– 20 балів.

Максимальна кількість балів які можуть бути отримані студентом: 100 балів.

Для отримання загальної позитивної оцінки з дисципліни кількість балів, набраних студентом протягом навчального семестру, має бути не меншою за 60.

*Результати навчання, які будуть оцінюватись:* РН 1., РН 2.1, РН 2.2, РН 2.3.

*Форма проведення і види завдань:* тестова, письмова.

*Види завдань:* 3 лабораторних роботи та тематичний тест за темою кожної з робіт.

*Підсумкова оцінка*  $100=3*10+2*20+30$ .

**Підсумковий контроль** проводиться у формі заліку – 100 балів. Залік виставляється за результатами роботи студента уздовж всього семестру і не передбачає додаткових заходів оцінювання для успішних студентів.

## 7.2 Організація оцінювання

**Терміни проведення форм оцінювання:**

1. Лабораторна робота 1 + тематичний тест: не пізніше 6 тижня семестру.
2. Лабораторна робота 2 + тематичний тест: не пізніше 10 тижня семестру.
3. Лабораторна робота 3 + тематичний тест: не пізніше 14 тижня семестру.

Студент має право на одне перескладання кожної лабораторної роботи із можливістю отримання максимально 80% початково визначених за цю контрольну роботу балів. Термін перескладання визначається викладачем.

У разі неякісного виконання лабораторної роботи, викладач має право не зарахувати лабораторну роботу, або знизити за неї бали.

Студент має право здавати лабораторні роботи після закінчення визначеного для них терміну, але з втратою двох балів за кожен тиждень, який пройшов з моменту закінчення терміну її здачі.

У випадку відсутності студентів з поважних причин відпрацювання та перескладання контрольних робіт здійснюються у відповідності до «Положення про організацію освітнього процесу» від 07.05.2018 року.

### 7.3 Шкала відповідності оцінок

<b>Зараховано / Passed</b>	60-100
<b>Незараховано / Failed</b>	0-59

### 8. Структура навчальної дисципліни. Тематичний план лекцій і практичних занять

№ п/п	Назва лекції	Кількість годин		
		лекції	практичні	с/р
<b>Частина 1. Основи візуалізації даних з R</b>				
1	<b>Тема 1. Вступ. Візуалізація інформації.</b> Переваги візуалізації даних. Особливості візуального представлення інформації. Попередня обробка даних. Типи даних. Способи подання даних. <i>Лабораторне заняття:</i> Ознайомлення з R, R-studio, бібліотеками та пакетами прикладних програм аналізу даних CRAN. Пакети R, призначені для візуалізації даних: stats, ggplot. <i>Самостійна робота:</i> Ознайомлення з R, R-studio, перші кроки та базові функції. Бібліотеки та пакети прикладних програм аналізу даних, CRAN.	2	2	8
2	<b>Тема 2. Мова програмування R.</b> Основи роботи в середовищі RStudio. Завантаження та встановлення пакетів. Основи роботи з даними: завантаження, очищення, структури даних. <i>Лабораторне заняття:</i> Базові операції в R. Робота зі структурами даних. Завантаження даних з файлів (CSV, Excel). Оформлення звіту в R Markdown. <i>Самостійна робота:</i> Опанування та реалізація розглянутих методів в R. Виконання лабораторної роботи №1.	2	2	8
3	<b>Тема 3. Основи візуалізації даних у R.</b> Основи роботи з графічними оболонками середовища R. Функція plot(). Графічні аргументи. Використання пакету ggplot2 для створення базових графіків. Типи візуалізацій: точкові, лінійні, гістограми, boxplot. <i>Лабораторне заняття:</i> Пакети R, призначені для візуалізації даних: stats, ggplot. Побудова базових графіків у ggplot2: налаштування осей, легенди, стилів. Інші графічні налаштування. <i>Самостійна робота:</i> Опанування та реалізація розглянутих методів в R. Виконання лабораторної роботи №1.	2	2	8
<b>Частина 2. Основи статистичного моделювання з R</b>				
4	<b>Тема 4. Основи статистичного аналізу.</b> Описові статистики, візуалізація розподілів даних. <i>Лабораторне заняття:</i> Отримання описових статистик та візуалізація результатів аналізу. Функції R для візуалізації розподілів. <i>Самостійна робота:</i> Опанування та реалізація розглянутих методів в R. Виконання лабораторної роботи №2.	2	2	8

5	<b>Тема 5. Моделювання випадкових вибірок.</b> Поняття випадкової вибірки, ймовірнісні закони розподілу, практичні приклади генерації даних. <i>Лабораторне заняття:</i> Функції R для генерації випадкових вибірок та аналізу їх властивостей. <i>Самостійна робота:</i> Опанування та реалізація розглянутих методів в R. Виконання лабораторної роботи №2.	2	2	8
6	<b>Тема 6. Випадкові розподіли: рівномірний, нормальний, експоненційний, їх властивості та застосування.</b> <i>Лабораторне заняття:</i> Функції R для роботи з розподілами випадкових величин. <i>Самостійна робота:</i> Опанування та реалізація розглянутих методів в R. Виконання лабораторної роботи №2.	2	2	10
<b>Частина 3. Елементи моделювання випадкових процесів з R</b>				
7	<b>Тема 7. Елементи теорії випадкових процесів. Найпростіші моделі: ланцюги Маркова.</b> <i>Лабораторне заняття:</i> Просте моделювання випадкових процесів: реалізація базових алгоритмів і аналіз результатів. Підготовка звіту у R Markdown. <i>Самостійна робота:</i> Опанування та реалізація розглянутих методів в R. Виконання лабораторної роботи №3.	2	4	10
<b>ВСЬОГО</b>		<b>14</b>	<b>14</b>	<b>60</b>

**Загальний обсяг 90 год.<sup>1</sup>**, в тому числі:

Лекцій – **14 год.**

Практичні – **14 год.**

Самостійна робота – **60 год.**

Консультації – **2 год.**

### **Умови лабораторних робіт:**

Деталізовані умови лабораторних робіт даються студентам на першому практичному занятті.

### **9. Рекомендовані джерела:**

#### **Основні:**

- 1) Майборода Р.Є. Комп'ютерна статистика – професійний старт, (2019), К.:ВПЦ “Київський університет”, 589 с. <http://probability.univ.kiev.ua/userfiles/mre/compsta1.pdf>
- 2) Гнатюк В. Вступ до R на прикладах / Віктор Гнатюк. Харків:ХНЕУ, 2010. 101 с.
- 3) Евелін Кларк, *Підручник R для початківців: вивчайте мову програмування R* (2024) <https://www.guru99.com/uk/r-tutorial.html>
- 4) Василенко Є. Прикладна аналітика для активістів природоохоронного руху (2020) [https://bookdown.org/geka/applied\\_analytics/](https://bookdown.org/geka/applied_analytics/)
- 5) A. Coghlan, *A Little Book of R For Multivariate Analysis* (2017).
- 6) P. Dalgaard, *Introductory Statistics with R*, 2-nd edition, Springer (2008).
- 7) J. Maindonald, W. John Braun, *Data Analysis and Graphics Using R – an Example-Based Approach*, Cambridge Univ. Press.(2010).
- 8) Wickham, Hadley. 2009. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <http://ggplot2.org>.
- 9) Hamilton, Nicholas. 2017. *Ggtern: An Extension to 'Ggplot2', for the Creation of Ternary Diagrams*. <https://CRAN.R-project.org/package=ggtern>.

<sup>1</sup> Загальна кількість годин, відведених на дану дисципліну згідно навчального плану.

- 10) Allaire, JJ, Jeffrey Horner, Vicent Marti, and Natacha Porte. 2015. *Markdown: 'Markdown' Rendering for R*. <https://CRAN.R-project.org/package=markdown>.
- 9) R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.

#### ***Додаткови:***

- 1) Kassambara, Alboukadel. 2016. *Ggcorrplot: Visualization of a Correlation Matrix Using 'Ggplot2'*. <https://CRAN.R-project.org/package=ggcorrplot>.
- 2) N. H. Timm, *Applied Multivariate Analysis*, Springer (2002).
- 3) Revelle, William. 2016. *Psych: Procedures for Psychological, Psychometric, and Personality Research*.
- 4) Evanston, Illinois: Northwestern University. <https://CRAN.R-project.org/package=psych>.
- 5) Wickham, Hadley, and Romain Francois. 2016. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.
- 6) Xie, Yihui. 2015. *Dynamic Documents with R and Knitr*. Chapman; Hall/CRC. <https://www.crcpress.com/Dynamic-Documents-with-R-and-knitr-Second-Edition/Xie/p/book/9781498716963>.
- 7) *Bookdown: Authoring Books and Technical Documents with R Markdown*. Chapman; Hall/CRC. 2016. <https://www.crcpress.com/bookdown-Authoring-Books-and-Technical-Documents-with-R-Markdown/Xie/p/book/9781138700109>.
- 8) Zeileis, Achim, and Gabor Grothendieck. 2005. "Zoo: S3 Infrastructure for Regular and Irregular Time Series." *Journal of Statistical Software* 14 (6): 1–27. <https://doi.org/10.18637/jss.v014.i06>.

#### ***Datasets***

<https://www.kaggle.com/datasets>

<https://vincentarelbundock.github.io/Rdatasets/datasets.html>