

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Львівський національний університет імені Івана Франка
Механіко-математичний факультет
Кафедра математичної статистики і диференціальних рівнянь



Затверджено

На засіданні
кафедри математичної статистики і
диференціальних рівнянь
механіко-математичного факультету
Львівського національного університету
імені Івана Франка
(протокол № 1 від 22.06.2023 р.)

Завідувач кафедри: Олег БУГРІЙ

Силабус з навчальної дисципліни
“ Основи програмування в R ”,
що викладається в межах ОПШ “ Статистичний аналіз даних ”
першого (бакалаврського) рівня вищої освіти для здобувачів з
спеціальності 112 - Статистика

Львів 2023 р.

Назва дисципліни	Основи програмування в R
Адреса викладання дисципліни	Головний корпус ЛНУ ім. І. Франка м. Львів, вул. Університетська 1, 79000
Факультет та кафедра, за якою закріплена дисципліна	Механіко-математичний факультет Кафедра математичної статистики і диференціальних рівнянь
Галузь знань, шифр та назва спеціальності	11 - Математика та статистика 112 - Статистика
Викладачі дисципліни	Холявка О.Т., кандидат фізико-математичних наук, асистент кафедри математичної статистики і диференціальних рівнянь
Контактна інформація викладачів	oksana.kholiavka@lnu.edu.ua, https://new.mmf.lnu.edu.ua/employee/kholyavka_o_t
Консультації з питань навчання по дисципліні відбуваються	Консультації в день проведення лекцій/практичних занять (за попередньою домовленістю). Головний корпус ЛНУ ім. І. Франка, каб. 267. м. Львів, вул. Університетська, 1
Сторінка курсу	https://new.mmf.lnu.edu.ua/course/r-112-bak23
Інформація про дисципліну	Дисципліна “Основи програмування в R” є вибірковою дисципліною зі спеціальності 112 – Статистика для освітньої програми “Статистичний аналіз даних”, яка викладається в 4-му семестрі в обсязі 4-х кредитів (за Європейською Кредитно-Трансферною Системою ECTS).
Коротка анотація дисципліни	R — це функціональна мова для статистичної обробки даних і створення візуалізацій. Вона допомагає здійснювати наукові відкриття, отримувати маркетингові інсайти й орієнтуватися в big data. Курс розроблено для ознайомлення студентів з основами мови програмування R.
Мета та цілі дисципліни	<i>Мета:</i> ознайомити студентів з програмуванням на R. <i>Цілі:</i> навчити студентів використовувати базові структури мови R для написання простих програм, написання власних функцій, ознайомлення зі структурами даних, з якими може працювати R.
Література для вивчення дисципліни	1) Hadley Wickham and Garrett Grolemund. <i>R for Data Science</i> . O’Reilly Medi, 2016. 2) Graham J. Williams. <i>The Essentials of Data Science: Knowledge Discovery Using R</i> . O’Reilly Medi, 2017. 3) Braun W.J., Murdoch D.J. <i>A First Course in Statistical Programming with R</i> . Cambridge University Press, 2007. 4) Hadley Wickham and Jennifer Bryan. <i>R Packages R Packages: Organize, Test, Document, and Share Your Code 2nd Edition</i> . O’Reilly Medi, 2023. 5) Brett Lantz. <i>Machine Learning with R</i> . Packt Publishing, 2013. 6) Julia Silge and David Robinson. <i>Text Mining with R: A Tidy Approach</i> . O’Reilly Medi, 2017. 7) G. James. <i>An Introduction to Statistical Learning: with Applications in R</i> . Springer, 2013.

	8) Højsgaard S., Edwards D., Lauritzen S. <i>Graphical Models with R</i> . Springer, 2012.
Обсяг курсу	Загальний обсяг: 120 годин. Аудиторних занять: 48 год., з них 32 год. лекційних та 16 годин лабораторних занять. Самостійної роботи: 72 год.
Очікувані результати навчання	У результаті вивчення даного курсу студент буде: знати: основи програмування на R, базові типи даних та пакети в R; вміти: використовувати базові структури мови R для написання простих програм, писати власні функції, аналізувати дані.
Ключові слова	Код, операція, умовні оператори, оператори типу «apply», R-studio, аналіз даних
Формат курсу	Очний
Теми	Див. Схема курсу
Підсумковий контроль, форма	Залік
Пререквізити	Для вивчення даного курсу студенти потрібні базові знання з: - Інформатики; - Програмування.
Навчальні методи та техніки, які будуть використовуватися під час викладання курсу	Інформаційні методи (лекція, бесіда, ілюстрація, демонстрація); дедуктивні методи на основі узагальнень; евристичні методи (проблемна лекція); інтерактивні методи (дискусія)
Необхідне обладнання	Для проведення лекційних занять: комп'ютер (мінімальні характеристики: процесор Intel Core i3, 4ГБ оперативної пам'яті), доступ до мережі Internet, засоби мультимедіа (в т.ч. проектор). Для проведення практичних/лабораторних занять: комп'ютер (мінімальні характеристики: процесор Intel Core i3, 4ГБ оперативної пам'яті), доступ до мережі Internet. Необхідне програмне забезпечення включає в себе ОС Windows 10, програмні додатки (MS Teams, MS Excel, Jupyter Notebook з вбудованим компілятором мови програмування Python).
Критерії оцінювання (окремо для кожного виду навчальної діяльності)	Оцінювання проводиться за 100-бальною шкалою. Бали нараховуються за наступним співвідношенням: • Змістовий модуль 1: 40% семестрової оцінки за контрольну роботу та виконання домашніх завдань, максимальна кількість балів 40. • Змістовий модуль 2: 40% семестрової оцінки за контрольну роботу та виконання домашніх завдань, максимальна кількість балів 40. • Контрольне тестування: 20% семестрової оцінки, максимальна кількість балів 20. Академічна доброчесність: Очікується, що роботи студентів будуть оригінальними дослідженнями чи міркуваннями. Списування та втручання в роботу інших студентів становлять, але не обмежують, приклади можливої академічної недоброчесності. Виявлення ознак академічної недоброчесності в написанні завдань є підставою для її незарахування викладачем, незалежно від масштабів плагіату чи обману.

Жодні форми порушення академічної доброчесності не толеруються.

Відвідання занять є важливою складовою навчання. Очікується, що всі студенти відвідають усі лекції та практичні/лабораторні заняття курсу. Студенти повинні інформувати викладача про неможливість відвідати заняття. У будь-якому випадку студенти зобов'язані дотримуватися термінів визначених для виконання всіх видів робіт, передбачених курсом.

Література. Уся література, яку студенти не зможуть знайти самостійно, буде надана викладачем виключно в освітніх цілях без права її передачі третім особам. Студенти заохочуються до використання також й іншої літератури та джерел, яких немає серед рекомендованих.

Політика виставлення балів. Враховуються бали, набрані при поточному контролі та бали підсумкового тестування. При цьому обов'язково враховуються присутність на заняттях та активність студента під час практичного заняття; недопустимість пропусків та запізнь на заняття; користування мобільним телефоном, планшетом чи іншими мобільними пристроями під час заняття в цілях не пов'язаних з навчанням; списування та плагіат; несвоєчасне виконання поставленого завдання і т. ін.

Оцінювання лабораторних робіт (2 змістових модулі містять по 4 практичних роботи кожен, загалом 8 практичних робіт, максимальна кількість балів: 64) відбувається шляхом оцінки роботи студента під час проведення лабораторної роботи в аудиторії (0-2 балів за одну роботу) та захисту написаної студентом вдома лабораторної роботи (0-4 балів за одну роботу). До 2 балів студент отримує за коректне налаштування необхідного програмного забезпечення.

Бали оцінювання аудиторного виконання практичних робіт нараховуються за наступним співвідношенням:

- 2 – студент в повному обсязі володіє навчальним матеріалом, має повне розуміння розглянутої теми, надає правильні відповіді на запитання по темі, код програми функціонує відповідно до завдання;
- 1.5 – студент достатньо розуміє розглянутий матеріал та принципи написаного ним коду програми, присутні неточності та незначні помилки у відповідях на запитання по темі, код програми функціонує відповідно до завдання;
- 1 – студент не досить добре розуміє розглянутий матеріал та написаний ним код програми, вагається та надає неточні/не конкретні відповіді на запитання по темі, код програми функціонує з помірними недоліками;
- 0.5 – студент погано розуміє розглянутий матеріал та написаний ним код програми, студент в більшості надає помилкові відповіді на питання по темі, код програми не функціонує належним чином;
- 0 - студент зовсім не засвоїв розглянутий матеріал, написаний ним код програми не відповідає темі/не функціонує взагалі.

Бали оцінювання домашнього завершення виконання практичних робіт та наданого звіту нараховуються за наступним співвідношенням:

- 4 – звіт цілком і повністю відображає індивідуальне завдання студента, містить правильні висновки, ілюстрований (за потреби) відповідними графіками і таблицями які правильно відображають суть виконаного завдання, студент має повне розуміння розглянутої теми, надає правильні відповіді на запитання по темі, код програми функціонує відповідно до завдання;

	<p>3 – звіт в достатній мірі відображає індивідуальне завдання студента, містить допустимі висновки, ілюстрований (за потреби) відповідними графіками і таблицями які частково відображають суть виконаного завдання, студент достатньо розуміє принципи написаного ним коду програми, присутні неточності та незначні помилки у відповідях на запитання по темі, код програми функціонує відповідно до завдання;</p> <p>2 – звіт містить загальні формулювання завдання, висновки нечіткі, необхідні ілюстрації чи таблиці відсутні, студент не досить добре розуміє розглянутий матеріал та представлений код програми, надає неточні/не конкретні відповіді на запитання по темі, код програми функціонує з помірними недоліками;</p> <p>1 – звіт не містить формулювання завдання, висновки необґрунтовані чи неповні, необхідні ілюстрації чи таблиці відсутні, студент погано розуміє розглянутий матеріал та представлений код програми, студент в більшості надає помилкові відповіді на питання по темі, код програми не функціонує належним чином;</p> <p>0 – звіт відсутній/не відповідає темі, студент зовсім не засвоїв розглянутий матеріал, написаний ним код програми не відповідає темі/не функціонує взагалі.</p> <p>Оцінювання контрольного тестування (завдання з тематики кожного змістового модуля) відбувається шляхом оцінки письмових відповідей студента на поставлені запитання. Відсотки нарахування балів оцінювання відповіді на кожне запитання нараховуються за наступним співвідношенням: 75-100% – тема відтворюється в повному обсязі, правильно, обґрунтовано, логічно; 50-75% – відтворюється значна частина розглянутої теми, проте присутні неточності та/або невідповідності; 25-50% – виявлено множинні неточності та невідповідності, пояснення відсутні чи частково помилкові; 0-25% – тему майже не розкрито, кількість викладеного матеріалу не відповідає загальним нормам обраного виду роботи.</p> <p>Критерії оцінювання результатів неформальної освіти: Нарахування балів відбувається за публікацію студентом тез доповідей на конференціях, наукових статей, за участь студента у діяльності наукових гуртків, семінарів, круглих столів, конкурсів, участь у заходах неформальної освіти, за отримання сертифікатів про проходження навчання на різних освітніх платформах (Coursera, Prometheus тощо), курсах провідних ІТ компаній за тематикою навчальної дисципліни. Кількість балів визначається відсотком покриття результатів відповідної активності до вимог результатів навчання з навчальної дисципліни</p>
Питання для контрольного тестування	Типи даних, оператори, синтаксис функції, пакети в R, датафрейми, списки, матриці, цикли.
Опитування	Анкету-оцінку з метою оцінювання якості курсу буде надано по завершенню курсу.

**Схема курсу “Основи програмування в R”
для студентів спеціальності 112 - Статистика**

Тижні	Лекційний курс		Практичні заняття		К-сть год СР	Літе- ратура
	Назва теми	К-сть год	Назва теми	К-сть год		
1	2	3	4	5	6	7
1	Вступ у мову програмування R. Історичні аспекти мови R. Особливості, переваги і недоліки мови R.	2			4	[1]-[8], Сайт курсу
2	Робота зі змінними та типами даних. Розуміння змінних. Числові, символічні, логічні та факторні типи даних. Присвоєння та операції.	2	Налаштування середовища розробки. Опрацювання основних концепцій роботи зі змінними та типами даних в мові програмування R.	2	5	[1]-[8], Сайт курсу
3	Структури даних в R. Вектори: створення, індексація та операції. Матриці та масиви. Списки: створення та доступ до елементів.	2			5	[1]-[8], Сайт курсу
4	Датафрейми в R. Вступ до датафреймів. Створення та редагування. Імпорт та експорт даних в R.	2	Робота з векторами, матрицями та масивами: створення, індексація та використання операцій. Робота зі списками: Створення, індексація та маніпуляції. Робота з фреймами даних: створення, індексація та фільтрація даних.	2	5	[1]-[8], Сайт курсу
5	Умовні конструкції. Організація циклів в R. If, else, if-else. Цикли for та while. Інтерполяція і згладжування в R.	2			5	[1]-[8], Сайт курсу
6	Функції в R. Створення функцій. Параметри та значення	2	Написання програм із використанням циклів. Робота з функціями в R.	2	4	[1]-[8], Сайт курсу

	повернення. Область видимості змінних.					
7	Робота з пакетами. Вступ в пакети R. Встановлення та завантаження пакетів. Вивчення загальноживаних пакетів.	2			5	[1]-[8], Сайт курсу
8	Очищення та підготовка даних. Обробка відсутніх даних. Видалення дублікатів. Трансформація та маніпуляція даними.	2	Вправи для встановлення, завантаження та використання пакетів R. Основні техніки та інструменти, які допоможуть зробити дані готовими для подальшого використання.	2	5	[1]-[8], Сайт курсу
9	Розширений аналіз даних (EDA). Описова статистика. Візуалізація даних за допомогою ggplot2. Таблиці та графіки зведення.	2			5	[1]-[8], Сайт курсу
10	Статистичний аналіз з R. Введення в статистичні тести. Проведення t-тестів та хі- квадрат тестів. Розуміння кореляції та регресії.	2	Описова статистика в R. Інтервальні та точкові оцінки. Тестування гіпотез.	2	4	[1]-[8], Сайт курсу
11	Data Mining з R. Техніки кластеризації. Видобуток асоціативних правил. Вступ в машинне навчання з R.	2			5	[1]-[8], Сайт курсу
12	Отримання інформації з веб-сайтів за допомогою R. Введення в web scraping. Використання бібліотеки rvest та інших. Видобуток даних з веб- сайтів.	2	Data Mining з R. Web scraping.	2	4	[1]-[8], Сайт курсу
13	Створення інтерактивних панелей у Shiny. Введення в Shiny.	2			5	[1]-[8], Сайт курсу

	Створення інтерактивних веб-застосунків. Розгортання додатків Shiny.					
14	Аналіз часових рядів. Основи часових рядів. Аналіз та візуалізація часових рядів. Прогнозування часових рядів в R.	2	Аналіз часових рядів.	2	4	[1]-[8], Сайт курсу
15	Версійний контроль за допомогою Git та GitHub. Введення в Git. Налаштування репозиторію GitHub. Колаборативне програмування з версійним контролем.	2			4	[1]-[8], Сайт курсу
16	Підсумкове заняття.	2	Робота з Git та GitHub.	2	3	[1]-[8], Сайт курсу
	Разом	32		16	72	
	Викладач: Холявка О.Т.		Викладач: Холявка О.Т.			