

**НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ І
ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ**

Факультет інформаційних технологій

НАВЧАЛЬНО-МЕТОДИЧНИЙ КОМПЛЕКС
з дисципліни

«Аналітика з R»

для підготовки фахівців

Галузь знань **05 «Соціальні та поведінкові науки»**

Спеціальність **051 «Економіка»**

Освітня програма **«Економічна кібернетика»**

Галузь знань **12 «Інформаційні технології»**

Спеціальність **121 "Інженерія програмного забезпечення"**

Освітня програма **"Інженерія програмного забезпечення"**

Спеціальність **122 «Комп'ютерні науки»**

Освітня програма **«Комп'ютерні науки»**

Ступінь вищої освіти – Бакалавр

Київ – 2020 р.

**НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ І
ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ**

Кафедра економічної кібернетики

“ЗАТВЕРДЖУЮ”

Декан факультету інформаційних технологій
_____ О.Г. Глазунова

(Підпис)

“ ___ ” _____ 20__ р.

РОЗГЛЯНУТО І СХВАЛЕНО

на засіданні кафедри економічної кібернетики
Протокол №11 від 19.06.2020 р.

Завідувач кафедри

_____ А.В. Скрипник

(Підпис)

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

«Аналітика з R»

Галузь знань **05 «Соціальні та поведінкові науки»**

Спеціальність **051 «Економіка»**

Освітня програма **«Економічна кібернетика»**

Галузь знань **12 «Інформаційні технології»**

Спеціальність **121 "Інженерія програмного забезпечення"**

Освітня програма **"Інженерія програмного забезпечення"**

Спеціальність **122 «Комп'ютерні науки»**

Освітня програма **«Комп'ютерні науки»**

Факультет інформаційних технологій

Розробники:

професор кафедри економічної кібернетики, д.е.н., доцент Жерліцин Д.М.

Київ – 2020 р.

1. Опис навчальної дисципліни

Аналітика з R

Галузь знань, напрям підготовки, спеціальність, ступень вищої освіти	
Освітній ступінь	<i>Бакалавр</i>
Галузь знань	05 Соціальні та поведінкові науки; 12 Інформаційні технології
Спеціальність	051 Економіка; 122 Комп'ютерні науки; 121 Інженерія програмного забезпечення
Освітня програма	Економічна кібернетика; Комп'ютерні науки; Інженерія програмного забезпечення
Характеристика навчальної дисципліни	
Вид	Вибіркова
Загальна кількість годин	105
Кількість кредитів ECTS	4
Кількість змістовних модулів	2
Курсовий проект (робота) (за наявності)	Не передбачено
Форма контролю	Залік
Показники навчальної дисципліни	
Форма навчання	<i>Денна</i>
Рік підготовки	3
Семестр	5
Лекційні заняття	30
Практичні, семінарські заняття	30
Лабораторні заняття	
Самостійна робота студента	45
Індивідуальні завдання	
Кількість тижневих аудиторних годин для денної форми навчання	<i>4 години на тиждень</i>

2. Мета, завдання та компетентності навчальної дисципліни

Мета: формування системи теоретичних знань і практичних навичок моделювання в середовищі R; опанування основних принципів роботи та базових методів і моделей у середовищі R.

Завдання вивчення курсу: вивчення основних принципів роботи в середовищі R, орієнтованих на застосування сучасних наукових методів, моделей та засобів інформаційних технологій; набуття вмінь створення, використання й адаптації моделей в R.

Засвоївши курс студент повинен:

знати

основні принципи організації середовища R,
функціонал роботи R,
основи статистичних розрахунків в R,
підходи до візуалізації в R,
принципи побудови моделей з R,
економетричні моделі,
використання моделей в управлінні економічними процесами;

вміти

працювати із даними в R;
графічно представляти дані в R;
застосувати функціонал роботи R для розв'язання практичних задач;
самостійно здійснювати підбір інформації необхідної для вирішення поставлених задач, аналізувати отримані результати;

володіти

методами аналізу статистичних даних,
використовувати сучасне програмне забезпечення.

Набуття компетентностей:

загальні компетентності (ЗК):

Освітня програма «Економічна кібернетика»

ЗК3. Здатність до абстрактного мислення, аналізу та синтезу..

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

Освітня програма "Інженерія програмного забезпечення"

ЗК2. Здатність до абстрактного мислення, аналізу, синтезу.

Освітня програма «Комп'ютерні науки»

ЗК01. Здатність до абстрактного мислення, аналізу та синтезу.

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

фахові (спеціальні) компетентності (ФК):

Освітня програма «Економічна кібернетика»

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

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

Освітня програма "Інженерія програмного забезпечення"

СК7. Володіння знаннями про інформаційні моделі даних, здатність створювати програмне забезпечення для зберігання, видобування та опрацювання даних.

СК9. Здатність оцінювати і враховувати економічні, соціальні, технологічні та екологічні чинники, що впливають на сферу професійної діяльності.

Освітня програма «Комп'ютерні науки»

СК6. Здатність до системного мислення, застосування методології системного аналізу для дослідження складних проблем різної природи, методів формалізації та розв'язування системних задач, що мають суперечливі цілі, невизначеності та ризики.

СК7. Здатність застосовувати теоретичні та практичні основи методології та технології моделювання для дослідження характеристик і поведінки складних об'єктів і систем, проводити обчислювальні експерименти з обробкою й аналізом результатів.

Програмні результати навчання:

Освітня програма «Економічна кібернетика»

ПРН12. Застосовувати набуті теоретичні знання для розв'язання практичних завдань та змістовно інтерпретувати отримані результати.

ПРН13. Ідентифікувати джерела та розуміти методологію визначення і методи отримання соціально-економічних даних, збирати та аналізувати необхідну інформацію, розраховувати економічні та соціальні показники.

Освітня програма "Інженерія програмного забезпечення"

ПРН01. Аналізувати, цілеспрямовано шукати і вибирати необхідні для вирішення професійних завдань інформаційно-довідникові ресурси і знання з урахуванням сучасних досягнень науки і техніки.

ПРН11. Вибирати вихідні дані для проектування, керуючись формальними методами опису вимог та моделювання.

Освітня програма «Комп'ютерні науки»

ПРН1. Застосовувати знання основних форм і законів абстрактно-логічного мислення, основ методології наукового пізнання, форм і методів вилучення, аналізу, обробки та синтезу інформації в предметній області комп'ютерних наук.

ПРН8. Використовувати методологію системного аналізу об'єктів, процесів і систем для задач аналізу, прогнозування, управління та проектування динамічних процесів в макроекономічних, технічних, технологічних і фінансових об'єктах.

3. Програма навчальної дисципліни

Змістовий модуль 1. Основи аналітичної роботи з R

Тема 1. Основи аналізу даних та сучасні інформаційні технології їх оброблення (6 годин).

Сучасні особливості підготовки та первинного оброблення аналітичної інформації. Пакети аналізу даних. Підходи щодо аналізу великих даних та представлення результатів аналізу. Застосування сучасних методів аналізу в управлінні соціально-економічними процесами.

Тема 2. Підготовка та оброблення даних з R. Робота з Data.frame (6 годин).

Основні принципи організації середовища R. Робота з командною консоллю інтерфейсу R. Робота з меню пакету R commander. Об'єкти, пакети, функції, пристрої. Типи даних мови R. Вектори і матриці. Фактори. Списки і таблиці. Імпортування даних в R. Представлення дати і часу. Часові ряди. Організація обчислень: функції, відгалуження, цикли. Робота с даними data.frame.

Тема 3. Описова статистика і нормальний закон розподілу в R (6 годин).

Пакет psych. Гіпотеза про рівність середніх двох генеральних сукупностей. Використання рангових критеріїв. Гіпотеза про однорідність дисперсії. Дисперсійний аналіз. Оцінка кореляції двох випадкових величин. Критерій χ^2 -квадрат. Тест Фішера.

Тема 4. Випадкові змінні та інструменти програмування в R (6 годин).

Оцінка вибірових параметрів з використанням спеціальних функцій. Використання вбудованих та власних функцій, додаткових пакетів. Аналіз викидів. Заповнення пропущених значень у таблицях даних. Відтворення результатів при використанні генератора випадкових чисел. Закони розподілу, реалізовані в R. Підбір закону і параметрів розподілу в R. Перевірка на нормальність розподілу.

Змістовий модуль 2. Аналітичні інструменти з мовою програмування R

Тема 5. Графічний аналіз з R (6 годин).

Діаграми розсіювання plot() і параметри графічних функцій. Гістограми і функція sdplot(). Діаграми розмахів. Кругові і стовпчикові діаграми. Діаграми Клівленда і одномірні діаграми розсіювання. Категоризовані графіки. Пакет ggplot2 та побудова складних діаграм.

Тема 6. Регресійний аналіз з R (9 годин).

Проста лінійна регресія. Поліноміальні і нелінійні моделі регресії. Множинна регресія. Моделі згладжування. Узагальнені моделі регресії. Логістична регресія. Коваріаційний аналіз. Індуктивні моделі. Моделювання структурних рівнянь.

Тема 7. Специфічні аналітичні інструменти, функції та пакети R (6 годин).

Створення та робота з інструментами створення звітів R Markdown. Створення звітів у форматі .html. Робота з форматом даних data.table. Бібліотека data.table. Відбір, фільтрування та групування даних data.table. Систематизація інших аналітичних пакетів R.

Назви змістових модулів і тем	Кількість годин					
	усього	у тому числі				
		л	п	лаб.	інд.	с. р.
1	2	3	4	5	6	7
Змістовий модуль 1. Основи аналітичної роботи з R						
Тема 1. Основи аналізу даних та сучасні інформаційні технології їх оброблення.	13	4	4			5
Тема 2. Підготовка та оброблення даних з R. Робота з data.frame.	13	4	4			5
Тема 3. Описова статистика і нормальний закон розподілу в R.	13	4	4			5
Тема 4. Випадкові змінні та інструменти програмування в R.	13	4	4			5
Разом за змістовим модулем 1	52	16	16			20
Змістовий модуль 2. Аналітичні інструменти з мовою програмування R						

1	2	3	4	5	6	7
Тема 5. Графічний аналіз з R.	16	4	4			8
Тема 6. Регресійний аналіз з R	18	5	5			8
Тема 7. Специфічні аналітичні інструменти, функції та пакети R	19	5	5			9
Разом за змістовим модулем 2	53	14	14			25
Усього годин	105	30	30			45

6. Теми лабораторних занять

№ з/п	Назва теми	Кількість годин
1.	Встановлення R. Ознайомлення з основними компонентами R	4
2.	Збирання та первинне оброблення даних	4
3.	Робота із таблицями. Аналіз статистичних даних	4
4.	Побудова та аналіз законів розподілу в R	4
5.	Побудова графіків в R	4
6.	Побудова регресійних моделей в R	5
7.	Робота з R Markdown	5
	Разом	30

7. Контрольні питання, комплекти тестів для визначення рівня засвоєння знань студентами.

Перелік питань для визначення рівня засвоєння знань студентами

1. Історія виникнення R.
2. Основні принципи організації середовища R.
3. Робота з командною консоллю інтерфейсу R.
4. Робота з меню пакету R commander.
5. Об'єкти, пакети, функції, пристрої.
6. Типи даних мови R.
7. Вектори і матриці.
8. Факторні змінні.
9. Списки і таблиці.
10. Data.Frame
11. Об'єднання та розділення таблиць Data.Frame.
12. Групування даних Data.Frame.
13. Бібліотека psych
14. Бібліотека

15. Імпортування даних в R.
16. Представлення дати і часу.
17. Часові ряди.
18. Організація обчислень: функції.
19. Організація обчислень: відгалуження.
20. Організація обчислень: цикли.
21. Векторизовані обчислення в R з використанням apply-функцій.
22. Оцінка вибірових параметрів з використанням спеціальних функцій.
23. Використання функцій summary () і додаткових пакетів.
24. Аналіз викидів.
25. Заповнення пропущених значень в таблицях даних.
26. Відтворення результатів при використанні генератора випадкових чисел.
27. Закони розподілу, реалізовані в R.
28. Підбір закону і параметрів розподілу в R.
29. Перевірка на нормальність розподілу.
30. Гіпотеза про рівність середніх двох генеральних сукупностей.
31. Використання рангових критеріїв.
32. Гіпотеза про однорідність дисперсій.
33. Дисперсійний аналіз.
34. Оцінка кореляції двох випадкових величин.
35. Критерій хі-квадрат.
36. Тест Фішера.
37. Діаграми розсіювання plot() і параметри графічних функцій.
38. Гістограми і функція qplot().
39. Бібліотека ggplot2.
40. Діаграми розмахів.
41. Кругові і стовпчикові діаграми.
42. Одномірні діаграми розсіювання.
43. Категоризовані графіки.
44. Проста лінійна регресія.
45. Поліноміальні моделі регресії.
46. Нелінійні моделі регресії.
47. Множинна регресія.
48. Моделі згладжування.
49. Узагальнені моделі регресії.
50. Моделі пробіт- і логіт- регресії.
51. Коваріаційний аналіз.
52. Звіти R Markdown
53. Робота з пакетом data.table

8. Методи навчання

В процесі викладання навчальної дисципліни за характером пізнавальної діяльності застосовуються переважно методи гейміфікації та пояснювально-ілюстративний, евристичний методи, а також частково кожен із зазначених методів залежно від видів робіт на занятті. (Табл. 1).

Таблиця 1

Класифікація методів навчання

Засади	Найменування	Характеристики
1. Джерело знань: слово образ досвід	словесні, наочні, практичні	
2. Етапи навчання	підготовка до вивчення нового матеріалу, вивчення нового матеріалу, закріплення вправ, контроль і оцінка	
3. Спосіб педагогічного керівництва	пояснення педагога, самостійна робота	керівництво: безпосереднє; опосередковане
4. Стиль викладання (пояснення)	інформаційно-повідомлювальний, пояснювальний, інструктивно-практичний, пояснювально-спонукальний	
4. Логіка навчання	індуктивні, дедуктивні, аналітичні, синтетичні	
5. Дидактичні цілі	організація навчальної діяльності, стимулювання і релаксація, контроль і оцінка	
6. Дидактичні завдання	методи оволодіння знаннями, методи формування умінь і навичок, застосування здобутих знань, умінь і навичок	
7. Характер пізнавальної діяльності	пояснювально- ілюстративні, репродуктивні проблемного викладу, частково-пошукові (евристичні), дослідницькі методи	репродуктивні Продуктивні

9. Форми контролю

Кожна з форм контролю має особливості й залежить від мети, змісту та характеру навчання. У процесі навчання дисципліни використовуються наступні форми контролю:

– **Поточний контроль:** усне опитування (індивідуальне, фронтальне, групове), комп'ютерне тестування, виконання практичних завдань на комп'ютері згідно програми;

– **Підсумковий контроль:** тестування та співбесіда за результатами роботи.

10. Розподіл балів, які отримують студенти. Оцінювання знань студента відбувається за 100-бальною шкалою і переводиться в національні оцінки згідно з табл. 1 «Положення про екзамени та заліки у НУБіП України» (наказ про уведення в дію від 27.12.2019 р. № 1371)

Поточний контроль		Рейтинг з навчальної роботи $R_{НР}$	Рейтинг з додаткової роботи $R_{ДР}$	Рейтинг штрафний $R_{ШТР}$	Підсумкова атестація <i>(залік)</i>	Загальна кількість балів
Змістовий модуль 1	Змістовий модуль 2					
0-100	0-100	0-70	0-20	0-5	0-30	0-100

Примітки.

Рейтинг з поточної роботи (навчальної, додаткової), штрафний рейтинг та показники підсумкової атестації визначаються відповідно п. 2 та п.4 Положення про екзамени та заліки у Національному університеті біоресурсів і природокористування України, що затверджене Вченою радою НУБіП України « 27 » грудня 2019 р. протокол №5.

Для визначення рейтингу студента (слухача) із засвоєння дисципліни $R_{дис}$ (до 100 балів) одержаний рейтинг з атестації (до 30 балів) додається до рейтингу студента (слухача) з навчальної роботи $R_{НР}$ (до 70 балів): $R_{дис} = R_{НР} + R_{ат}$.

Загальний рейтинг здобувача вищої освіти із засвоєння дисципліни у балах переводиться у національної оцінки наступним чином:

Шкала оцінювання: національна та ECTS

Рейтинг здобувача вищої освіти, бали	Оцінка національна за результати складання екзаменів заліків	
	екзаменів	заліків
90-100	відмінно	зараховано
74-89	добре	
60-73	задовільно	
0-59	незадовільно	не зараховано

11. Методичне забезпечення

Електронний навчальний курс, розроблений на базі платформи LMS Moodle, розміщений на навчальному порталі НУБіП України за адресою: <https://elearn.nubip.edu.ua/course/view?id=2455>

12. Рекомендована друкована література

Основна

1. *Економетрика з R* : навчальний посібник / А.В. Скрипник, Д.М. Жерліцин, Ю.О. Нам'ясенко. – Київ: ФОП Ямчинський О.В., 2020. – 248 с.
2. Роберт И. Кабаков. R в действии. Анализ и визуализация данных в программе R / пер. с англ. Полины А. Волковой. – М.: ДМК Пресс, 2014. – 588 с.: ISBN 978-5-947060-077-1.
3. Шитиков В.К., Мاستицкий С.Э. (2017) Классификация, регрессия и другие алгоритмы Data Mining с использованием R. 351 с. – Электронная книга, адрес доступа: <https://github.com/ranalytics/data-mining>
4. Роберт И. Кабаков. R в действии. Анализ и визуализация данных в программе R / пер. с англ. Полины А. Волковой. – М.: ДМК Пресс, 2014. – 588 с.: ISBN 978-5-947060-077-1

Допоміжна

1. Grant V. Farnsworth (2008) Econometrics in R. October 26, 2008. 50 p. <https://cran.r-project.org/doc/contrib/Farnsworth-EconometricsInR.pdf>
2. Machine Learning with R - Second Edition. by Brett Lantz. Publisher: Packt Publishing. Release Date: July 2015. ISBN: 9781784393908. <https://www.oreilly.com/library/view/machine-learning-with/9781784393908/>
3. Deborah Nolan, Duncan Temple Lang (2015) Data Science in R: A Case Studies Approach to Computational Reasoning and Problem Solving. 1st Edition. Chapman and Hall/CRC. Published April 21, 2015. 539 p. ISBN 9781482234817
4. XML and web technologies for data sciences with R / Nolan, Deborah; Lang, Duncan Temple. New York [u.a.] : Springer, 2014. 663 p. - ISBN 978-1-4614-7899-7
5. [Joseph Adler](#) (2012) R in a Nutshell, 2nd Edition. A Desktop Quick Reference. [O'Reilly Media](#). October 2012. 724 p
6. Winston Chang (2013) R Graphics Cookbook
7. Мэтлофф Норман. Искусство программирования на R. Погружение в большие данные. – СПб., Питер, 2019. – 416 с. (Серия «Библиотека программиста»). ISBN 978-5-4461-1101-5.

Допоміжна (Internet джерела)

8. <http://r-analytics.blogspot.com/> блог «Анализ и визуализация данных» С. Мاستицкого;
9. <http://www.algorithmist.ru/search/label/R> серія статей С. Едунова по реалізації в R різних алгоритмів ;
10. <http://rrus.wordpress.com/> блог «R»

11. <http://www.inp.nsk.su/~baldin/DataAnalysis/index.html> - матеріали Е. Балдіна;
12. <https://m7876.wiki.zoho.com/Introduction-to-R.html> «Введение в R»

13. Інформаційні ресурси

1. World Bank Open Data. Режим доступу: <https://data.worldbank.org>
2. World Economic Forum Режим доступу: <https://www.weforum.org/>
3. Державна служба статистики України. Режим доступу: <http://www.ukrstat.gov.ua>
4. Міністерство Фінансів України. Режим доступу: <https://www.minfin.gov.ua/>
5. Національний банк України. Режим доступу: <https://bank.gov.ua/control/uk/index>
6. Організація економічного співробітництва та розвитку. Режим доступу: <https://data.oecd.org>