

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ І
ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ

КАФЕДРА ЕКОНОМІЧНОЇ КІБЕРНЕТИКИ

«ЗАТВЕРДЖУЮ»

Декан факультету інформаційних технологій

_____ (О.Г. Глазунова)

«_____» _____ 2019

Розглянуто і схвалено

На засіданні кафедри економічної
кібернетики

Протокол №__ від _____ 2019 р.

Завідувач кафедри

_____ Скрипник А.В.

Робоча програма навчальної дисципліни

Аналітика з R

Спеціальність 051 Економіка (Економічна кібернетика)

Факультет інформаційних технологій

Розробники:

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

1. Опис навчальної дисципліни

Аналітика з Р

Галузь знань, напрям підготовки, спеціальність, ступень вищої освіти	
Ступень вищої освіти	<i>перший (бакалаврський) рівень</i>
Галузь знань	05 Соціальні та поведінкові науки
Спеціальність	051 Економіка
Освітньо-кваліфікаційна програма	Економічна кібернетика
Характеристика навчальної дисципліни	
Вид	Вибіркова
Загальна кількість годин	90
Кількість кредитів ECTS	2
Кількість змістовних модулів	2
Курсовий проект (робота) (за наявності)	
Форма контролю	Екзамен
Показники навчальної дисципліни	
Форма навчання	<i>Денна</i>
Рік підготовки	<i>3</i>
Семестр	<i>5</i>
Лекційні заняття	<i>30</i>
Практичні, семінарські заняття	<i>15</i>
Лабораторні заняття	
Самостійна робота студента	<i>45</i>
Індивідуальні завдання	
Кількість тижневих аудиторних годин для денної форми навчання	<i>3 години на тиждень</i>

Мета:

формування системи теоретичних знань і практичних навичок аналізу економічної інформації в середовищі R; опанування основних принципів роботи та базових аналітичних методів 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 аналітичних наборів даних 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 аналітичної обробки з R,
особливості економетричного моделювання в R,
використання інструментів R 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.

3. Програма навчальної дисципліни

Змістовий модуль 1. Основи роботи у середовищі R

Тема 1. Базові принципи аналізу даних.

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. Компоненти середовища та елементи синтаксису R.

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. Векторизовані обчислення в R з використанням apply-функцій.

Тема 3. Попередня обробка даних в R.

Оцінка вибірових параметрів з використанням спеціальних функцій. Використання функцій *summary()* і додаткових пакетів. Аналіз викидів. Заповнення пропущених значень в 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. Базові аналітичні інструменти середовища R

Тема 4. Графічний аналіз даних у середовищі R.

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

Тема 5. Data.frame та описова статистика в R

Пакет *psych*. Гіпотеза про рівність середніх двох генеральних сукупностей. Використання рангових критеріїв. Гіпотеза про однорідність дисперсій. Дисперсійний аналіз. Оцінка кореляції двох випадкових величин. Критерій хі-квадрат. Тест Фішера.

Тема 6. Економетричне моделювання в R

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

4. Структура навчальної дисципліни

Назви змістових модулів і тем	Кількість годин					
	усього	у тому числі				
		л	п	лаб.	інд.	с. р.
1	2	3	4	5	6	7
Модуль 1						
Змістовий модуль 1. Основи роботи у середовищі R						
Тема 1. Базові принципи аналізу даних.	14	5	2			7
Тема 2. Компоненти середовища та елементи синтаксису R.	14	5	2			7
Тема 3. Попередня обробка даних в R.	15	5	3			7
Разом за змістовим модулем 1	43	15	7			21
Модуль 2						
Змістовий модуль 2. Базові аналітичні інструменти середовища R						
Тема 4. Графічний аналіз даних у середовищі R.	15	5	2			8
Тема 5. Data.frame та описова статистика в R	16	5	3			8
Тема 6. Економетричне моделювання в R	16	5	3			8
Разом за змістовим модулем 2	47	15	15			24
Усього годин	90	30	8			45

6. Теми практичних занять

№ з/п	Назва теми	Кількість годин
1.	Тема 1. Базові принципи аналізу даних.	2
2.	Тема 2. Компоненти середовища та елементи синтаксису R.	2
3.	Тема 3. Попередня обробка даних в R.	3
4.	Тема 4. Графічний аналіз даних у середовищі R.	2
5.	Тема 5. Data.frame та описова статистика в R	3
6.	Тема 6. Економетричне моделювання в R	3
	Разом	15

8. Самостійна робота

№ з/п	Назва теми	Кількість годин
1.	Тема 1. Базові принципи аналізу даних.	7
2.	Тема 2. Компоненти середовища та елементи синтаксису R.	7
3.	Тема 3. Попередня обробка даних в R.	7
4.	Тема 4. Графічний аналіз даних у середовищі R.	8
5.	Тема 5. Data.frame та описова статистика в R	8
6.	Тема 6. Економетричне моделювання в R	8
	Разом	45

8. 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; модульні контрольні роботи
- Підсумковий контроль: екзамен

Розподіл балів, які отримують студенти. Оцінювання студента відбувається згідно положенням «Про екзамени та заліки у НУБіП України» від 27.02.2019 р. протокол № 7 з табл. 1.

	Поточний контроль		Рейтинг 3 навчальн ої роботи R _{НР}	Рейтинг 3 додатков ої роботи R _{др}	Рейтинг штрафни й R _{ШТР}	Підсумко ва атестація (іспит)	Загальн а кількіст ь балів
	Змістови й модуль 1	Змістови й модуль 2					
Кількіст ь балів	0-100	0-100	0-70	0-20	0-5	0-30	0-100
«Вага» виду роботи	0,35	0,35	0,7	0,2	0,05	0,3	1

Примітка. Відповідно до «Положення про кредитно-модульну систему навчання в НУБіП України», затвердженого ректором університету 03.04.2009 р., рейтинг студента з навчальної роботи R_{НР} стосовно вивчення певної дисципліни визначається за формулою

$$R_{\text{нр}} = \frac{0,7 \left(R_{\text{зМ}}^{(1)} \cdot K_{\text{зМ}}^{(1)} + R_{\text{зМ}}^{(2)} \cdot K_{\text{зМ}}^{(2)} \right)}{K_{\text{дис}}} + R_{\text{др}} - R_{\text{штр}}$$

де $R_{\text{зМ}}^{(1)}$, $R_{\text{зМ}}^{(2)}$ – рейтингові оцінки відповідно першого (1) і другого (2) змістових модулів за 100-бальною шкалою;

$K_{\text{зМ}}^{(1)}$, $K_{\text{зМ}}^{(2)}$ – кількість кредитів ECTS, передбачених робочим навчальним планом для відповідного змістового модуля;

$K_{\text{дис}} = K_{\text{зМ}}^{(1)} + K_{\text{зМ}}^{(2)}$ – кількість кредитів ECTS, передбачених робочим навчальним планом для дисципліни у поточному семестрі;

$R_{\text{др}}$ – рейтинг з додаткової роботи;

$R_{\text{штр}}$ – рейтинг штрафний.

Рейтинг з додаткової роботи $R_{\text{др}}$ додається до $R_{\text{нр}}$ і не може перевищувати 20 балів. Він визначається лектором і надається студентам рішенням кафедри за виконання робіт, які не передбачені навчальним планом, але сприяють підвищенню рівня знань студентів з дисципліни.

Рейтинг штрафний $R_{\text{штр}}$ не перевищує 5 балів і віднімається від $R_{\text{нр}}$ - Він визначається лектором і вводиться рішенням кафедри для студентів, які матеріал змістового модуля засвоїли невчасно, не дотримувалися графіка роботи, пропускали заняття тощо.

Оцінка національна	Оцінка ЄКТС	Визначення оцінки ЄКТС	Рейтинг студента, бали
Відмінно	A	ВІДМІННО - відмінне виконання лише з незначною кількістю помилок	90 -100
Добре	B	ДУЖЕ ДОБРЕ - вище середнього рівня з кількома помилками	82-89
	C	ДОБРЕ - в загальному правильна робота з певною кількістю грубих помилок	74-81

	Е	ДОСТАТНЬО – виконання задовольняє мінімальні критерії	60 – 63
Незадовільно	FX	НЕЗАДОВІЛЬНО – потрібно працювати перед тим, як отримати залік (позитивну оцінку)	35 – 59
	F	НЕЗАДОВІЛЬНО – необхідна серйозна подальша робота	01 – 34

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

10. Методичне забезпечення

1. Скрипник А.В., Галаєва Л.В., Долінська Є.Б. Математичні моделі та планування експерименту: Методичні розробки – К.: ТОВ «Аграр Медіа Груп», 2011. – 130 с.
2. Жадлун З.О., Галаєва Л.В., Шульга Н.Г. Теоретичні основи математичного моделювання економічних процесів. – К.: НАУ, 2009. – 63 с.
<http://elibrary.nubip.edu.ua/16946/>

11. Рекомендована література

Основні джерела:

1. Гнатюк В. Вступ до R на прикладах. Харківський національний економічний університет, 2010. 107 с.
2. Статистический анализ данных в системе R. Учебное пособие / А. Г. Буховец [и др.]; под ред. А. Г. Буховец. Воронеж: ВГАУ, 2010. 124 с. PDF
3. Хорник К. Часто задаваемые вопросы по R: R ЧаВо: пер. с англ., 2007. 7 с.
4. Шишкин В.А. Программное обеспечение эконометрических расчетов. Пермь: ПГНИУ, 2014. 273 с.
5. Яу Н. Искусство визуализации в бизнесе: Как представить сложную информацию простыми образами / пер. с англ. А. Кирова. Москва: Манн, Иванов и Фербер, 2013. 352 с.
6. Cohen Y., Cohen J.Y. Statistics and Data with R: An Applied Approach Through Examples. Wiley, 2008. 603 p.

Додаткові джерела:

1. Advances in Social Science Research Using R / ed. by H. D. Vinod. Springer, 2010. 205 p.
2. Albert J. Bayesian Computation With R. 2nd ed. Springer, 2009. 304 p.
3. Albert J., Rizzo M. R by Example. Springer, 2012. 374 p.
4. Basic R for Finance / D. Würtz [et al.]. Zurich: Rmetrics Association & Finance Online Publishing, 2010. 312 p.

5. Beyersmann J., Allignol A., Schumacher M. Competing Risks and Multistate Models with R. Springer, 2012. 249 p.
6. Carmona R. Statistical Analysis of Financial Data in R. 2nd ed. Springer, 2014. 595 p.
7. Curran J.M. Introduction to Data Analysis with R for Forensic Scientists. CRC Press, 2010. 317 p.
8. Demidenko E. Mixed Models: Theory and Applications with R. 2nd ed. Wiley, 2013. 754 p. (Wiley Series in Probability and Statistics).
9. Finch W.H., Bolin J.E., Kelley K. Multilevel Modeling Using R. CRC Press, 2014. 226 p.
10. Galecki A., Burzykowski T. Linear Mixed-Effects Models Using R: A Step-by-Step Approach. Springer, 2013. 625 p.
11. Horton N.J., Kleinman K. Using R for Data Management, Statistical Analysis, and Graphics. CRC Press, 2010. 296 p.
12. Kuhn M., Johnson K. Applied Predictive Modeling. Springer, 2013. 600 p.
13. Martinussen T., Scheike T.H. Dynamic Regression Models for Survival Data. Springer, 2006. 470 p.
14. Modern Actuarial Risk Theory: Using R / R. Kaas [et al.]. 2nd ed. Springer, 2008. 394 p.
15. Pfaff B. Financial Risk Modelling and Portfolio Optimization with R. Wiley, 2013. 374 p.
16. Tufféry S. Data Mining and Statistics for Decision Making. Wiley, 2011. 704 p.
17. Zivot E., Wang J. Modeling Financial Time Series with S-PLUS®. 2nd ed. Springer, 2006. 998 p.

Internet джерела:

1. <http://r-analytics.blogspot.com/> блог «Анализ и визуализация данных» С. Мастицкого;
2. <http://www.algorithmist.ru/search/label/R> серия статей С. Едунова по реализации в R различных алгоритмов;
3. <http://rrus.wordpress.com/> блог «R по-русски»;
4. <http://www.inp.nsk.su/~baldin/DataAnalysis/index.html> - материалы Е. Балдина;
5. <https://m7876.wiki.zoho.com/Introduction-to-R.html> «Введение в R» по-русски
6. http://ru.wikibooks.org/wiki/Язык_программирования_R материал из Вики-учебника;
7. <http://www.ibm.com/developerworks/ru/library/l-r1/index.html> перевод трех статей Метца и Хантинга «Статистическое программирование в R»;
8. <http://www.twirpx.com/library/comp/r/> раздел с материалами по R в библиотеке для студентов и аспирантов,