

НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ

ІНСТИТУТ СОЦІОЛОГІЇ

УХВАЛЕНО

Вченою радою Інституту
соціології НАН України
«28» травня 2024 р.
Протокол № 5

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

АНАЛІЗ ДАНИХ У ПРОГРАМНОМУ СЕРЕДОВИЩІ R

галузь знань	05 Соціальні та поведінкові науки
спеціальність	054 Соціологія
освітній рівень	третій освітньо-науковий (PhD доктор філософії)
освітньо-наукова програма	Соціологія
вид дисципліни	Вибіркова

Форма навчання	денна
Навчальний рік	_____
Семестр	
Кількість кредитів ECTS	3
Мова викладання, навчання та оцінювання	українська
Форма заключного контролю	Залік

Викладачі: PhD в галузі соціології, н.с. Москотіна Р.Ю.

КИЇВ – 2024

Розробник:

Москотіна Р.Ю., PhD в галузі соціології, науковий співробітник відділу методології та методів соціології

Викладач: PhD в галузі соціології Москотіна Р.Ю.

1. Мета – ознайомити аспірантів з основами роботи в R та зі статистичним аналізом соціальних даних у цьому програмному середовищі.

2. Попередні вимоги до опанування або вибору навчальної дисципліни (за наявності):

1. Знати

- методи математичної статистики.

2. Вміти

- застосовувати дескриптивний аналіз до соціальних даних;
- застосовувати тести для перевірки статистичних гіпотез та коефіцієнти кореляції для різних типів шкал змінних;
- працювати з методами багатовимірною аналізу даних (регресійний, факторний, кластерний аналіз).

3. Володіти навичками

- роботи з іншими пакетами статистичного аналізу даних (наприклад, SPSS).

3. Анотація дисципліни: Навчальна дисципліна «Аналіз даних у програмному середовищі R» має на меті ознайомити аспірантів з мовою програмування R та навчити проводити статистичний аналіз даних з її використанням, зважаючи на те, що вона має ширші можливості для аналізу даних, порівняно з іншими програмами завдяки існуванню безлічі додаткових пакетів, котрі дозволяють застосовувати саме ті методи, які потрібні досліднику, а також представляти результати аналізу даних у вигляді таблиць та візуалізацій, писати аналітичні звіти тощо. R широко використовується у сучасних соціологічних дослідженнях, оволодіння нею і підвищує конкурентну спроможність аспірантів і полегшує їх входження до міжнародних дослідницьких проєктів.

4. Завдання (навчальні цілі).

Основними завданнями навчальної дисципліни є:

- ознайомити аспірантів з типами даних в R;
- навчити робити в R одновимірний та двовимірний аналіз даних, розраховувати коефіцієнти кореляції;
- навчити за допомогою інструментів R здійснювати перевірку статистичних гіпотез;
- навчити аспірантів здійснювати в R багатовимірний аналіз даних (регресійний, факторний, кластерний аналіз).

5. Результати навчання за дисципліною:

Результат навчання (1. знати; 2. вміти; 3. комунікація; 4. автономність та відповідальність)		Форми (та/або методи і технології) викладання і навчання	Методи оцінювання та пороговий критерій оцінювання	Відсоток у підсумковій оцінці з дисципліни
Код	Результат навчання			
1	2	3	4	5
ЗК01	Розуміти принципи побудови мови програмування R як інструменту для аналізу даних	<i>Лекція, семінар:</i> Тема 1	Участь у дискусії на семінарі, виконання практичного завдання, участь у дискусійній частині лекції	10%
ЗК02	Вміти обробляти та аналізувати дані з використанням мови програмування R	<i>Лекції, семінари:</i> Теми 2-6	Участь у дискусії на семінарах, виконання практичних завдань, участь у дискусійній частині лекцій	30%
СК01	Вміти використовувати здобуті навички з аналізу даних в R під час виконання оригінальних соціологічних досліджень	<i>Лекції, семінари, індивідуальні практичні завдання:</i>	Участь у дискусії на семінарах, виконання практичних завдань, участь у дискусійній частині лекцій, оцінка індивідуальних	30%

		Теми 2-6	практичних завдань	
СК02	Розуміти логіку аналізу даних з використанням R, викладену в україномовних та англійськомовних наукових публікаціях та самостійно презентувати власні результати в наукових текстах	<i>Лекції, семінари, індивідуальні практичні завдання:</i> Теми 2-6	Участь у дискусії на семінарах, виконання практичних завдань, участь у дискусійній частині лекцій, оцінка індивідуальних практичних завдань	30%

6. Співвідношення результатів навчання дисципліни із програмними результатами навчання

Результати навчання дисципліни (код) Програмні результати навчання (назва)	Загальні компетентності		Спеціальні компетентності	
РН07. Застосовувати сучасні інструменти і технології пошуку, оброблення та аналізу інформації, зокрема статистичні методи аналізу даних великого обсягу та/або складної структури, спеціалізовані бази даних та інформаційні системи	ЗК01	ЗК02	СК02	СК03

7. Схема формування оцінки.

ВИМОГИ І КРИТЕРІЇ ОЦІНЮВАННЯ

Види робіт	Кількість балів за один вид робіт	Критерії оцінювання
Участь у дискусійній частині лекції	2	Активна участь, аргументовані, змістовні судження і зауваження
	1	Фрагментарна участь у дискусії
	0	Відсутність на лекції
Участь у дискусії на семінарському занятті	2	Активна участь у дискусіях; фахова постановка запитань, аргументовані і логічні зауваження
	1	Фрагментарна участь, окремі репліки, запитання
	0	Відсутність на семінарському занятті
Виконання практичних завдань під час семінарських занять	7-9	Завдання виконано коректно, продемонстровано засвоєння відповідної теми
	4-6	Завдання виконано, але є певні неточності
	1-3	Завдання виконано, але неточностей дуже багато
	0	Завдання не виконано
Індивідуальні практичні завдання за темою семінарського заняття	8-11	Завдання виконано вчасно. Високий рівень опанування матеріалу
	4-7	Завдання виконано вчасно, проте наявні певні неточності і пробіли в знаннях
	1-3	Завдання виконано, але наявні серйозні неточності і пробіли у знаннях
	0	Завдання не виконано

При викладанні дисципліни використовуються наступні методи контролю знань аспірантів:

- усний контроль (опитування – індивідуальні або фронтальні, дискусії);
- письмовий контроль (виконання практичних завдань під час семінарських занять, виконання індивідуальних практичних завдань);
- самоконтроль знань аспірантів і взаємоконтроль.

Поточний контроль проводиться за результатами роботи аспірантів на практичних заняттях шляхом оцінювання індивідуальних практичних завдань.

Об'єктом оцінювання знань аспірантів є програмний матеріал навчальної дисципліни, який перевіряється під час поточного і підсумкового контролю з метою встановлення рівня і якості засвоєння знань, формування необхідних для майбутньої професійної діяльності навичок і вмінь.

Поточний контроль передбачає оцінювання успішності аспірантів під час аудиторних занять та виконання ними окремих індивідуальних завдань. При поточному контролі оцінюється: участь у дискусіях на лекційних заняттях, активність роботи аспірантів на семінарських заняттях, якість виконання практичних завдань під час семінарських занять, результати виконання ними індивідуальних практичних завдань.

Контроль на семінарських заняттях – оцінювання відповідей студентів, оцінка виконання практичних завдань.

7.1 Форми оцінювання:

- підсумкове оцінювання у формі ЗАЛІКУ:

Семестровий контроль проводиться у формі заліку або шляхом накопичення балів за поточним контролем по змістовним модулям.

Залік виставляється за результатами роботи аспіранта впродовж усього семестру і не передбачає додаткових заходів оцінювання для успішних студентів.

- умови допуску до заліку:

У разі отримання оцінки «незадовільно» (нижче 59 балів) аспірант повинен додатково виконати індивідуальні завдання для підвищення рівня своїх знань і повторно перескласти підсумковий контроль. Слухач має право на два перескладання: викладачеві та комісії. При цьому максимальна підсумкова оцінка після перескладання може бути лише «задовільно».

7.2 Організація оцінювання: (обов'язково зазначається порядок організації передбачених робочою навчальною програмою форм оцінювання із зазначенням орієнтовного графіку оцінювання).

УМОВИ ВИЗНАЧЕННЯ НАВЧАЛЬНОГО РЕЙТИНГУ

Форми оцінювання	Кількість	Максимум балів	Разом
Участь у дискусійній частині лекції	6	2	12
Участь у дискусії на семінарському занятті	6	2	12
Виконання практичних завдань під час семінарських занять	6	9	54
Індивідуальні практичні завдання за темою семінарського заняття	2	11	22
ВСЬОГО			100

7.3 Шкала відповідності оцінок

Відмінно / Excellent	90-100
Добре / Good	75-89
Задовільно / Satisfactory	60-74
Незадовільно / Fail	0-59
Зараховано / Passed	60-100
Не зараховано / Fail	0-59

8. Структура навчальної дисципліни

Тематичний план занять з дисципліни

№ з/п	Номер і назва теми	Кількість годин		
		лекції	семінари	сам. робота
1	Тема 1. Знайомство з R. Типи даних в R	2	2	6
2	Тема 2. Одновимірний та двовимірний аналіз в R. Перетворення змінних	2	2	6
3	Тема 3. Кореляційний аналіз в R. Перевірка статистичних гіпотез	2	2	8
4	Тема 4. Регресійний аналіз в R	2	2	6
5	Тема 5. Кластерний в R	2	2	6
6	Тема 6. Експлораторний і конфірматорний факторний аналіз в R	2	2	8
ВСЬОГО за дисципліною 90 годин		12	12	66

Загальний обсяг **90 год**, в тому числі (вибрати необхідне):

Лекцій – **12 год**.

Семінари – **12 год**.

Самостійна робота – **66 год**.

Зміст навчальної дисципліни

№ з/п	Номер і назва теми	Кількість годин	
		Структура	Література
1	2	3	4
1	Тема 1. Знайомство з R. Типи даних в R	1. Вектори та матриці в R 2. Числові, логічні, символічні, факторні змінні 3. Датафрейми (таблиці даних). Імпорт масиву даних SPSS в R у вигляді датафрейму 4. Списки в R	1, 7
2	Тема 2. Одновимірний та двовимірний аналіз в R. Перетворення змінних	1. Побудова одно- та двовимірних відсоткових розподілів в R 2. Міри центральної тенденції та міри варіації 3. Перетворення змінних в R	1, 7
3	Тема 3. Кореляційний аналіз в R. Перевірка статистичних гіпотез	1. Коефіцієнти зв'язку для двох змінних 2. Перевірка статистичних гіпотез про рівність двох і більше розподілів 3. Перевірка статистичних гіпотез про нормальність розподілу	4, 6, 7, 8
4	Тема 4. Регресійний аналіз в R	1. Парна лінійна регресія в R 2. Множинна лінійна регресія в R 3. Бінарна логістична регресія в R	1, 2, 7, 11
5	Тема 5. Кластерний в R	1. Ієрархічний кластерний аналіз в R 2. Кластерний аналіз методом k-середніх в R	5, 7
6	Тема 6. Експлораторний і конфірмаційний факторний аналіз в R	1. Метод головних компонент в R 2. Експлораторний факторний аналіз (метод головних факторів) в R 3. Конфірмаційний факторний аналіз в R	3, 7, 9, 10
Підсумковий контроль			Залік

9. Рекомендовані джерела:

Основна:

- Гнатюк В. (2010). *Вступ до R на прикладах*. Харків, ХНЕУ. 107 с. URL: <https://sociology.knu.ua/sites/default/files/library/eloadopen/hnatyuk-r-book-ua.pdf>
- Bevans, R. (2023, June 22). *Linear Regression in R | A Step-by-Step Guide & Examples*. Scribbr. URL: <https://www.scribbr.com/statistics/linear-regression-in-r/>
- Confirmatory Factor Analysis (CFA) in R with lavaan*. (n.d.). URL: <https://stats.oarc.ucla.edu/r/seminars/rcfa/>
- Correlation test between two variables in R - Easy guides - Wiki - STHDA*. (n.d.). <http://www.sthda.com/english/wiki/correlation-test-between-two-variables-in-r>
- Finnstats. (2021, April 20). *Cluster analysis in R | R-Bloggers*. R-bloggers. <https://www.r-bloggers.com/2021/04/cluster-analysis-in-r/>
- Goyal, A. (2023, October 20). *Parametric and Non-Parametric Tests: The Complete guide*. Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2021/06/hypothesis-testing-parametric-and-non-parametric-tests-in-statistics/>
- Hothorn, T., & Everitt, B. S. (2005). *A handbook of statistical analyses using R*. URL: <https://vcarrion.people.uic.edu/handbookrstarters.pdf>
- Zach. (2022, June 11). *How to test for normality in R (4 methods)*. Statology. <https://www.statology.org/test-for-normality-in-r/>

9. Zhang, S. (2023, November 13). *Factor Analysis | R Programming for Psychometrics*.
URL: https://bookdown.org/sz_psyc490/r4psychometrics/factor-analysis.html

Додаткова:

10. Brown, T. A., & Moore, M. T. (2012). Confirmatory factor analysis. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 361–379). The Guilford Press.
11. Hilbe, J. M. (2016). *Practical guide to logistic regression*. CRC Press.

Додаткові ресурси (за наявності):

CRAN (тут можна встановити R): <https://cran.r-project.org/>
Встановити RStudio: <https://posit.co/download/rstudio-desktop/>
Rbloggers: <https://www.r-bloggers.com/>
Bookdown: <https://bookdown.org/>
Rpubs: <https://rpubs.com/>