

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Центральноукраїнський державний педагогічний університет імені Володимира Винниченка

Факультет математики, природничих наук та технологій
Кафедра математики та методики її навчання

«ЗАТВЕРДЖУЮ»

В.о. завідувача кафедри



Ботузова Ю.В.

«4» серпня 2022 року



РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

ОК 11 Практикум з аналізу даних великих масивів

Рівень вищої освіти: другий (магістерський)

Галузь знань: 11 Математика та статистика

Спеціальність: 112 Статистика

Освітня програма: Статистика (Фінансова, страхова та комп'ютерна статистика)

Форма навчання: денна

2022 – 2023 навчальний рік

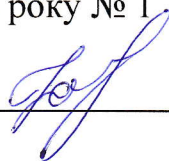
Робоча програма навчальної дисципліни Практикум з аналізу даних великих масивів розроблена на основі освітньо-професійної програми Статистика (Фінансова, страхова та комп'ютерна статистика), навчального плану підготовки здобувачів вищої освіти освітнього ступеня магістр за спеціальністю 112 Статистика

Розробники: Акбаш Катерина Сергіївна, доцент кафедри математики та методики її навчання, кандидат фізико-математичних наук

Робочу програму обговорено та схвалено на засіданні кафедри математики та методики її навчання

Протокол від «4» серпня 2022 року № 1

В.о. завідувача кафедри _____ д.пед.н., доцент Ботузова Ю.В.



Робоча програма навчальної дисципліни **Практикум з аналізу даних великих масивів** для студентів спеціальності 112 Статистика за другим (магістерським) рівнем вищої освіти. – ЦДПУ імені В. Винниченка, 2022.

© Акбаш К.С., 2022 рік

© ЦДПУ імені В.Винниченка, 2022 рік

1. ПОЯСНЮВАЛЬНА ЗАПИСКА

1.1. Опис навчальної дисципліни

| Найменування показників | Галузь знань, спеціальність, рівень вищої освіти | Характеристика навчальної дисципліни | |
|--|--|--------------------------------------|-----------------------|
| | | денна форма навчання | заочна форма навчання |
| Кількість кредитів – 3 | Галузь знань <u>11 Математика та статистика</u> (шифр і назва) | Нормативна | |
| Індивідуальне навчально-дослідне завдання _____ (назва) | Спеціальність: <u>112 Статистика</u> (шифр і назва) | Рік підготовки | |
| | | 1-й | -й |
| Загальна кількість годин – 90 | Освітня програма: <u>Статистика (Фінансова, страхова та комп'ютерна статистика)</u> (шифр і назва) | Семестр | |
| | | 2-й | -й |
| Тижневих годин для денної форми навчання: аудиторних – 40 самостійної роботи студента – 50 | Рівень вищої освіти: <u>другий (магістерський)</u> | Лекції | |
| | | 0 год. | год. |
| | | Практичні, семінарські | |
| | | 0 год. | год. |
| | | Лабораторні | |
| | | 40 год. | год. |
| | | Самостійна робота | |
| | | 50 год. | год. |
| | | Індивідуальні завдання: | |
| | | год. | |
| Вид контролю: | | | |
| диф. залік | | | |

1.2. Мета та завдання навчальної дисципліни

Метою викладання дисципліни є формування цілісного представлення про сучасні проблеми аналізу та обробки великих даних, допомогти оволодіти досвідом розробки та аналізу концептуальних та теоретичних моделей прикладних задач аналізу великих даних.

Завданнями вивчення навчальної дисципліни є:

- Закріплення принципів роботи з великими даними різного роду;
- Опанувати стандартні інструменти роботи з великими даними;
- Опанувати методи представлення результатів обробки великих даних.

У результаті вивчення навчальної дисципліни у студента мають бути сформовані такі *компетентності*:

| загальні | фахові |
|---|---|
| <p>ЗК1. Здатність до абстрактного мислення, аналізу та синтезу.</p> <p>ЗК3. Здатність проведення теоретичних та прикладних досліджень на відповідному рівні.</p> <p>ЗК6. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.</p> | <p>ФК1. Здатність аналізувати предметні області, формувати, аналізувати та моделювати процеси їх функціонування.</p> <p>ФК2. Знання на рівні новітніх досягнень, необхідні для дослідницької та/або інноваційної діяльності у сферах статистики та інформаційних технологій з їх практичними застосуваннями.</p> <p>ФК7. Здатність самостійно розробляти проекти шляхом творчого застосування існуючих та генерування нових математичних та статистичних підходів.</p> <p>ФК8. Здатність до розвитку нових та удосконалення існуючих статистичних методів аналізу, моделювання, прогнозування, розв'язування нових проблем у нових галузях знань.</p> <p>ФК10. Здатність до самоосвіти та підвищення кваліфікації на основі інноваційних підходів у сфері статистики та інформаційних технологій.</p> |

1.3. Очікувані програмні результати навчання:

ПРН1. Знання методології аналізу та моделювання прикладної предметної області, виявлення інформаційних потреб і збору вихідних даних для статистичного дослідження процесів її функціонування.

ПРН5. Володіння статистичними методами аналізу, прогнозування та оцінки параметрів моделей, способами інтерпретації числових даних та принципами функціонування процесів реального світу.

ПРН12. Застосовувати методології обґрунтованого вибору парадигм і спеціалізоване програмне забезпечення для вирішення прикладних статистичних завдань.

ПРН13. Інтегрувати знання з різних галузей для вирішення теоретичних та/або практичних задач і проблем.

2. ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Змістовий модуль 1. Методи аналізу законів розподілів ймовірностей випадкових величин

Тема 1. Аналіз розподілів залишкового часу життя людей за даними International Database on Longevity

Реєстрація в базі International Database on Longevity. Можливості пакету Statgraphics для знаходження розподілів залишкового часу життя людей. Групування даних в інтервальні розподіли засобами статистичного пакету SPSS. Розрахунок імовірностей для різних інтервалів. Інтерпретація та візуалізація результатів дослідження.

Тема 2. Застосування непараметричних тестів для проведення гендерного аналізу та аналізу відмінностей між країнами за даними International Database on Longevity

Використання критерію Манна-Уїтні для проведення гендерного аналізу. Застосування критерію Курскала Уолліса для проведення аналізу за країнами. Інтерпретація та візуалізація результатів дослідження.

Тема 3. Робота з вибірковими мікроданими державної служби статистики України

Створення таблиць спряженості різного виду (з номінальними, порядковими та кількісними змінними). Дослідження взаємозв'язку між змінними. Інтерпретація та візуалізація результатів дослідження.

Тема 4. Моделювання часових рядів на основі ARIMA моделей з передаточними функціями

Збір щоденних даних (наприклад, криптовалют). Побудова ARIMA на основі даних без дослідження впливу факторів. Побудова ARIMA на основі даних з аналізом факторів. Введення найбільш значущого фактору в модель (передаточна функція). Порівняння моделей.

Тема 5. Створення аналітичних звітів з інфографікою

Створення коротких аналітичних звітів до кожної тем 1-4 практикуму. Презентація звітів.

2. СТРУКТУРА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

| Назви розділів і тем | Кількість годин | | | | | |
|--|-----------------|--------------|---|-----------|-----|-----------|
| | денна форма | | | | | |
| | усього | у тому числі | | | | |
| | | л | п | лаб. | інд | СРС |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Змістовий модуль 1. Методи аналізу законів розподілів ймовірностей випадкових величин | | | | | | |
| Тема 1. Аналіз розподілів залишкового часу життя людей за даними International Database on Longevity | 18 | | | 8 | | 10 |
| Тема 2. Застосування непараметричних тестів для проведення гендерного аналізу та аналізу відмінностей між країнами за даними International Database on Longevity | 18 | | | 8 | | 10 |
| Тема 3. Робота з вибірковими мікроданими державної служби статистики України | 18 | | | 8 | | 10 |
| Тема 4. Моделювання часових рядів на основі ARIMA моделей з передаточними функціями | 18 | | | 8 | | 10 |
| Тема 5. Створення аналітичних звітів з інфографікою | 18 | | | 8 | | 10 |
| Усього годин | 90 | | | 40 | | 50 |

4. ОПИС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

4.1. Теми лекційних занять

4.2. Теми семінарських (практичних) занять

4.3. Теми лабораторних занять

4.3.1 денна форма навчання

| № з/п | Назва теми | Кількість годин |
|-------|--|-----------------|
| 1 | Аналіз розподілів залишкового часу життя людей за даними International Database on Longevity | 8 |
| 2 | Застосування непараметричних тестів для проведення гендерного аналізу та аналізу відмінностей між країнами за даними International Database on Longevity | 8 |
| 3 | Робота з вибірковими мікроданими державної служби статистики України | 8 |
| 4 | Моделювання часових рядів на основі ARIMA моделей з передаточними функціями | 8 |
| 5 | Створення аналітичних звітів з інфографікою | 8 |
| | Разом | 40 |

4.4. Завдання для самостійної роботи

4.4.1 денна форма навчання

| № з/п | Назва теми | Кількість годин |
|-------|--|-----------------|
| 1 | Аналіз відомих математичних моделей залишкового часу життя людей. Робота з науковими статтями | 10 |
| 2 | Дослідження відмінностей у моделях залишкового часу життя жінок та чоловіків. Робота з науковими статтями | 10 |
| 3 | Таблиці спряженості. Коефіцієнти асоціації, колігації. Дослідження залежностей між змінними у таблицях спряженості | 10 |
| 4 | ARIMA моделі з передаточними функціями | 10 |
| 5 | Сучасні методи відображення інфорграфіки. Ресурс Map Chart для створення інтерактивних мап | 10 |
| | Разом | 50 |

4.5. Індивідуальне навчально-дослідне завдання

Не передбачено.

4.6. Методи навчання

Під час вивчення дисципліни «Практикум з аналізу даних великих масивів» передбачено комплексне використання різноманітних методів організації і здійснення навчально-пізнавальної діяльності студентів та методів стимулювання

і мотивації їх навчання, що сприяє розвитку творчих засад особистості майбутнього фахівця, з урахуванням індивідуальних особливостей учасників освітнього процесу.

З метою формування професійних компетентностей широко впроваджуються інноваційні методи навчання. Це – комп'ютерна підтримка освітнього процесу, впровадження інтерактивних методів навчання.

Бали студенти накопичують на основі виконаних завдань у вигляді короткого аналітичного звіту до тем 1-4 практикуму.

4.7. Засоби діагностики результатів навчання здобувачів освіти.

Порядок та критерії виставлення балів

Навчальна дисципліна «Практикум з аналізу даних великих масивів» оцінюється за модульно–рейтинговою системою.

Контрольні заходи здійснюються з дотриманням вимог об'єктивності, індивідуального підходу, системності, всебічності.

Використовуються такі методи контролю (усний, письмовий), які мають сприяти підвищенню мотивації студентів до навчально-пізнавальної діяльності, зокрема, поточний контроль та підсумковий контроль.

Завданням поточного контролю є перевірка розуміння та засвоєння певної частини учбового матеріалу, рівня сформованості навичок, умінь самостійно опрацьовувати навчальний матеріал, здатності осмислити зміст теми.

Об'єктами поточного контролю знань студента є систематичність та активність роботи на заняттях; виконання завдань для самостійної роботи. Оцінюванню можуть підлягати: рівень знань, продемонстрований у комплексних аналітичних звітах; активність при обговоренні питань практичного заняття.

У разі невиконання завдань поточного контролю студент має право скласти їх індивідуально до останнього практичного заняття за дозволом завідувача кафедри. Порядок такого контролю регламентований викладачем.

Завданням підсумкового контролю є перевірка розуміння студентом програмного матеріалу в цілому, логіки та взаємозв'язків між окремими темами, здатності творчого використання накопичених знань, уміння сформулювати своє ставлення до певної проблеми навчальної дисципліни тощо.

4.9. Схема нарахування балів, які отримують студенти

Залік (I семестр)

| Поточне тестування та самостійна робота | | | | Сума |
|---|----|----|----|------|
| Розділ 1 | | | | |
| T1 | T2 | T3 | T4 | 100 |
| 20 | 20 | 20 | 20 | |

T1, T2 ... T4 – теми розділів.

Шкала оцінювання: національна та ЄКТС

| Сума балів за всі види навчальної діяльності | Оцінка за національною шкалою | |
|--|--|--------------|
| | для екзамену, курсового проєкту (роботи), практики | для заліку |
| 90-100 | відмінно | зараховано |
| 82-89 | добре | |
| 74-81 | | |
| 64-73 | | |
| 60-63 | задовільно | незараховано |
| 35-59 | незадовільно | |
| 1-34 | незадовільно | |

5. НАВЧАЛЬНО-МЕТОДИЧНЕ ЗАБЕЗПЕЧЕННЯ

5.1. Рекомендована література

Основна

1. Бахрушин В.Є. Методи аналізу даних : навчальний посібник для студентів / В.Є. Бахрушин. – Запоріжжя : КПУ, 2011. – 268 с. ISBN 978-966-414-103-8
2. Горват А.А., Молнар О.О., Мінькович В.В. Методи обробки експериментальних даних з використанням MS Excel: Навчальний посібник. Ужгород: Видавництво УжНУ "Говерла", 2019. – 160 с.
3. Грицюк П.М., Остапчук О.П. Аналіз даних: Навчальний посібник.– Рівне: НУВГП, 2008. – 218 с.
4. Лук'яненко І. Г., Жук В. М. Аналіз часових рядів. Частина перша : Побудова ARIMA, ARCH/GARCH моделей з використанням пакета E.Views 6.0. Практичний посібник для роботи в комп'ютерному класі / І. Г. Лук'яненко, В. М. Жук. – К. : НаУКМА; Аграр Медіа Груп, 2013. – 187 с.
5. Лупан І.В., Авраменко О.В., Акбаш К.С. Комп'ютерні статистичні пакети: навчально-методичний посібник. – 2-е вид. Кіровоград: «КОД» 2015.
6. Майборода Р. Є. Комп'ютерна статистика : підручник / Р. Є. Майборода. – К. : ВПЦ "Київський університет", 2019. – 589 с. ISBN 978-966-933-010-9
7. Оленко А.Я. Комп'ютерна статистика: Навчальний посібник. – К.: Видавничо-поліграфічний центр «Київський університет», 2007. – 174 с.

Допоміжна література

1. Karin Aarssen & Laurens de Haan (1994) On the maximal life span of humans, *Mathematical Population Studies*, 4:4, 259-281, DOI: 10.1080/08898489409525379
2. Rootzén, Holger & Zholud, Dmitrii. (2017). Human life is unlimited - but short. *Extremes*. 20. 10.1007/s10687-017-0305-5.
3. Modig K, Andersson T, Vaupel J, Rau R, Ahlbom A. How long do centenarians survive? Life expectancy and maximum lifespan. *J Intern Med*. 2017 Aug;282(2):156-163. doi: 10.1111/joim.12627. Epub 2017 May 24. PMID: 28470872.
4. Thatcher AR. The long-term pattern of adult mortality and the highest attained age. *J R Stat Soc Ser A Stat Soc*. 1999;162(Pt. 1):5-43. doi: 10.1111/1467-985x.00119. PMID: 12294994.

6. ПОЛІТИКА ЩОДО АКАДЕМІЧНОЇ ДОБРОЧЕСНОСТІ

Політика щодо академічної доброчесності формується на основі дотримання принципів академічної доброчесності відповідно до Законів України «Про освіту», «Про вищу освіту», «Про наукову і науково-технічну діяльність», «Про авторське право і суміжні права», «Про видавничу справу», з урахуванням норм Положення «Про академічну свободу та академічну доброчесність в Центральнорукраїнському державному педагогічному університеті імені Володимира Винниченка» (затверджене вченою радою, протокол №2 від 30.09.2019; №10 від 07.02.2022).