

ВИЩИЙ НАВЧАЛЬНИЙ ЗАКЛАД УКООПСПІЛКИ
«ПОЛТАВСЬКИЙ УНІВЕРСИТЕТ ЕКОНОМІКИ І ТОРГІВЛІ»

Інститут економіки, управління та інформаційних технологій
Кафедра математичного моделювання та соціальної інформатики

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ О. О. Ємець
підпис ініціали, прізвище

«__» _____ 20__ р.

РОБОЧА ПРОГРАМА

навчальної дисципліни "Аналіз даних та прикладні пакети статистичної обробки"

освітня програма/спеціалізація «Комп'ютерні науки»

спеціальність 122 "Комп'ютерні науки"

галузь знань 12"Інформаційні технології"

ступінь вищої освіти бакалавр

Полтава 2019

Укладач: Олексійчук Юрій Федорович, доцент кафедри математичного моделювання та соціальної інформатики, к. ф.-м. н.

Робоча програма навчальної дисципліни "Аналіз даних та прикладні пакети статистичної обробки" схвалена та рекомендована до використання в освітньому процесі на засіданні кафедри математичного моделювання та соціальної інформатики

Протокол від « ___ » _____ 20__ року № ___

ПОГОДЖЕНО:

Гарант освітньої програми «Комп'ютерні науки»
спеціальності 122 "Комп'ютерні науки"
ступеня бакалавр

_____ О. О. Ємець _____
підпис ініціали, прізвище

« _____ » _____ 20__ року

Розділ 1. Загальна характеристика навчальної дисципліни

Таблиця 1. Загальна характеристика навчальної дисципліни «Аналіз даних та прикладні пакети статистичної обробки»

| |
|---|
| 1. Місце в структурно-логічній схемі підготовки базовими дисциплінами є: ТЙМС |
| 2. Кількість кредитів за ЄКТС <u>4</u> |
| 3. Кількість модулів <u>1</u> |
| 4. Вибіркова дисципліна |
| 5. Курс – <u>4</u> |
| 6. Семестр – <u>8</u> |
| 7. Кількість годин: <u>120</u> |
| - лекції: <u>16</u> |
| - лабораторні: <u>32</u> |
| - самостійна робота: <u>72</u> |
| Вид підсумкового контролю: ПМК |

Розділ 2. Перелік компетентностей та програмні результати навчання з навчальної дисципліни

Предметом дисципліни «Аналіз даних та прикладні пакети статистичної обробки» є комп'ютерні технології статистичної обробки даних за допомогою стандартного табличного процесора Excel і спеціалізованих прикладних пакетів.

Метою вивчення дисципліни «Аналіз даних та прикладні пакети статистичної обробки» є формування уявлення про методи статистичної обробки даних із використанням комп'ютерних технологій для розв'язування практичних задач; виробити навички вибору відповідних статистичних моделей та методів для наявних даних.

Основним завданням дисципліни є набуття студентам систематизованих знань з основних аналітичних і статистичних методів обробки даних за допомогою відповідних програмних засобів та формування відповідних вмінь та навичок.

Таблиця 2. Перелік компетентностей та програмні результати навчання

| № | Компетентності, якими повинен оволодіти студент | Програмні результати навчання |
|--------------------------------|--|--|
| Загальні компетентності | | |
| | ЗК1. Здатність до абстрактного мислення, аналізу та синтезу. | Здобувати систематичні знання в галузі комп'ютерних наук, аналізувати проблеми з погляду сучасних наукових парадигм |
| | ЗК3. Знання та розуміння предметної області та розуміння професійної діяльності. | Реалізовувати засвоєні поняття, концепції, теорії та методи в інтелектуальній і практичній діяльності в галузі комп'ютерних наук |

| | | |
|--|---|---|
| | ЗК7. Здатність до пошуку, оброблення та аналізу інформації з різних джерел. | Знання методів, способів і технологій збору інформації з різних джерел, контент-аналізу документів, аналізу та обробки даних. |
| Спеціальні (фахові, предметні) компетентності | | |
| | СК2. Здатність до виявлення статистичних закономірностей недетермінованих явищ, застосування методів обчислювального інтелекту, зокрема статистичної, нейромережевої та нечіткої обробки даних, методів машинного навчання та генетичного програмування тощо. | Знання методів опрацювання емпіричних даних, методів обчислювального інтелекту, елементів машинного навчання. Використовувати сучасні середовища для розв'язування задач статистичної обробки експериментальних даних. |

Розділ 3. Програма навчальної дисципліни

Програма навчальної дисципліни затверджена Вченою радою Вищого навчального закладу Укоопспілки «Полтавський університет економіки і торгівлі», протоколвід«28» лютого 2018 року №3.

Розділ 4. Тематичний план навчальної дисципліни

| Назва теми (лекції) та питання теми (лекції) | Назва теми та питання лабораторного заняття | Завдання самостійної роботи у розрізі тем | Інформаційні джерела (порядковий номер за переліком) |
|---|---|---|--|
| Модуль 1. Основи аналізу даних | | | |
| <u>Тема 1. Основи аналізу даних</u> <u>Лекція 1. Підготовка вхідної інформації</u> 1. Предмет «Аналіз даних і прикладні пакети статистичної обробки». 2. Розділи аналізу даних. 3. Підготовка вхідної | | | [1-4,5-8] |

| | | | |
|---|---|--|-----------|
| інформації. 3.1. Класифікація змінних 3.2. Групування даних | | | |
| <u>Тема 2. Попередня обробка даних</u> <u>Лекція 2. Попередня обробка</u> 1. Квантиль та процентні точки 2. Основні статистичні характеристики 3. Характеристики векторних величин 4. Перевірка стохастичності вибірки 5. Аномальні спостереження | <u>Лабораторна робота</u> 1. Видалення аномальних спостережень <u>Лабораторна робота</u> 2. Перевірка стохастичності вибірки <u>Лабораторна робота</u> 3. Основні статистичні характеристики | 1. Видалення аномальних спостережень. Векторний випадок | [1-2,7-8] |
| Модуль 2. Прикладні пакети для статистичної обробки даних | | | |
| <u>Тема 3. Загальний огляд статистичних пакетів</u> <u>Лекція 3. Огляд статистичних пакетів.</u> 1. MS Excel 2. Спеціалізовані статистичні пакети 3. Python | <u>Лабораторна робота</u> 4. Знайомство з дистрибутивом Anaconda та мовою програмування Python | 1. Мова програмування R 2. IDE для Python | [1-4] |
| <u>Тема 4. Аналіз одновимірних даних з використанням статистичних пакетів</u> <u>Лекція 4. Розвідувальний аналіз. Гіпотези</u> 1. Сімейства розподілів типу зсув-масштабу 2. Пробіт- та ймовірнісний графіки 3. Інші графічні методи 4. Спостереження за двома змінними 5. Статистичні гіпотези та їх | <u>Лабораторна робота</u> 5. Робота з даними в Python <u>Лабораторна робота</u> 6. МКР №1. Аналіз даних в Python | 1. Графічні можливості Python 2. Перевірка статистичних гіпотез | [6-7] |

| | | | |
|---|--|--|-------|
| перевірка | | | |
| Модуль 3. Кореляційний та регресійний аналіз | | | |
| <u>Тема 5. Кореляційний аналіз</u> <u>Лекція 5.</u> Кореляційний аналіз 1. Види залежностей 2. Коефіцієнт кореляції 3. Властивості коефіцієнта кореляції 4. Приклади 5. Частинна та множинна кореляція 6. Кореляція якісних змінних | <u>Лабораторна робота 7.</u> Коефіцієнт кореляції <u>Лабораторна робота 8.</u> Кореляція якісних змінних <u>Лабораторна робота 9.</u> Частинна кореляція | 1. Кореляція номінальних змінних 2. Застосування Python для кореляційного аналізу | [1-6] |
| <u>Тема 6. Регресійний аналіз</u> <u>Лекція 6.</u> Регресійний аналіз 1. Поняття регресії 2. Лінійна регресія 2.1. Парна лінійна регресія 2.2. Множинна лінійна регресія 3. Нелінійна регресія 4. Оцінка адекватності регресії | <u>Лабораторна робота 10-11.</u> Парна лінійна регресія <u>Лабораторна робота 12.</u> Обчислення параметрів парної лінійної регресії засобами MS Excel <u>Лабораторна робота 13.</u> Парна нелінійна регресія <u>Лабораторна робота 14.</u> Множинна регресія | 1. Застосування Python для регресійного аналізу | [6-8] |
| Модуль 4. Аналіз даних з використанням статистичних пакетів | | | |
| <u>Тема 7. Аналіз двовимірних даних</u> <u>Лекція 7.</u> Дисперсійний аналіз 1. Однофакторний дисперсійний аналіз 2. Багатофакторний дисперсійний аналіз | | 1. Приклади застосування однофакторного дисперсійного аналізу 2. Приклади застосування багатофакторного дисперсійного аналізу | [7-8] |
| <u>Тема 8. Класифікація даних</u> <u>Лекція 8.</u> Дискримінантний та кластерний аналізи 1. Поняття дискримінантного аналізу 2. Кластерний аналіз | <u>Лабораторна робота 15.</u> Метод k-найближчих сусідів <u>Лабораторна робота 16.</u> МКР №2. Регресійний аналіз | 1. Застосування Python для класифікації даних | [7-8] |

Розділ 5. Система поточного та підсумкового контролю знань студентів

| Форма навчальної роботи | Вид навчальної роботи | Бали |
|-----------------------------|---|------|
| 1. Аудиторна 1.1. Лекція | ▪ Відвідування всіх лекцій лекції та лабораторних | 20 |
| 1.2. Практичні заняття | ▪ Підготовка до лабораторної роботи та її виконання (3x14=48) | 48 |
| 2. Підсумковий контроль. | МКР№1 | 16 |
| | МКР№2 | 16 |
| Усього за семестр | | 100 |

Розділ 6. Програмне забезпечення комп'ютерної підтримки освітнього процесу

Використовується наступне програмне забезпечення:

1. Дистрибутив Anaconda та мова програмування Python 3.x.
2. MS Excel або інший табличний редактор.

Розділ 7. Рекомендовані джерела інформації Основна література

1. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: исследование зависимостей: Справ. изд. – М.: Финансы и статистика, 1985. – 487 с.
2. Боровиков В.П. STATISTICA: искусство анализа данных на компьютере. – СПб.: Питер, 2001. – 656 с.
3. Боровиков В.П. Программа STATISTICA для студентов и инженеров. 2-е изд. – М.: Компьютер-пресс, 2001. – 301 с.
4. Гайдышев И. Анализ и обработка данных: специальный справочник. — СПб: Питер, 2001. — 752 с.
5. Лялин В. С. Статистика: теория и практика в Excel. — М.: Финансы и статистика; ИНФРА-М, 2010. — 448 с.
6. Минько А.А. Статистический анализ в MS Excel. – М.: Издательский дом «Вильямс», 2004. – 448 с.
7. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере / По. ред. В.Э. Фигурнова. – М.: ИНФРА-М, Финансы и статистика, 1995. – 384 с.

8. Халафян А.А. STATISTICA 6. Статистический анализ данных. 3-е изд. Учебник. – М.: ООО «Бином-Пресс», 2008. – 512 с.

Додаткова література

9. Берк К., Кэйри П. Анализ данных с помощью Microsoft Excel.: Пер. с англ. – М.: Издательский дом «Вильямс», 2005. – 560 с.
- 10.Єріна А.М. Статистичне моделювання та прогнозування: Навч. посібник. – К.: КНЕУ, 2001. – 170 с.
- 11.Комп'ютерні методи в сільському господарстві та біології: Навч. посібник / О.М. Царенко, Ю.А. Злобін, В.Г. Складар, С.М. Панченко. – Суми: Університетська книга, 2000. – 203 с.
- 12.Корнелл П. Анализ данных в Excel. Просто как дважды два. — М.: Эксмо, 2007. — 224 с.
- 13.Левин Д., Стефан Д., Кребиль Т., Беренсон М. Статистика для менеджеров с использованием Microsoft Excel, 4-е изд.: Пер. с англ. – М.: Издательский дом «Вильямс», 2004. – 1312 с.
- 14.Зюков М. Є. Аналіз даних: Навчально-методичний посібник. — Полтава: РВВ ПУСКУ, 2009. — 41 с.