

СИЛАБУС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

1. Загальна інформація про навчальну дисципліну

Повна назва навчальної дисципліни	Програмне забезпечення статистичного аналізу
Повна офіційна назва закладу вищої освіти	Сумський державний університет
Повна назва структурного підрозділу	Навчально-науковий інститут бізнесу, економіки та менеджменту. Кафедра економічної кібернетики
Розробник(и)	Олійник Віктор Михайлович
Рівень вищої освіти	Перший рівень вищої освіти, НРК – 6 рівень, QF-LLL – 6 рівень, FQ-EHEA – перший цикл
Семестр вивчення навчальної дисципліни	16 тижнів протягом 6-го семестру
Обсяг навчальної дисципліни	Обсяг дисципліни становить 5 кред. ЄКТС, 150 год., з яких 1 кред. (30 год.) становить курсова робота, 64 год. становить контактна робота з викладачем (24 год. лекцій, 40 год. лабораторних занять), 56 годин становить самостійна робота
Мова викладання	Англійська

2. Місце навчальної дисципліни в освітній програмі

Статус дисципліни	Обов'язкова навчальна дисципліна для освітньої програми "Бізнес аналітика"
Передумови для вивчення дисципліни	Передумови для вивчення відсутні
Додаткові умови	Додаткові умови відсутні
Обмеження	Обмеження відсутні

3. Мета навчальної дисципліни

Метою навчальної дисципліни є формування системи теоретичних і практичних знань у галузі використання сучасних спеціалізованих програмних засобів при моделюванні та розв'язанні економіко-математичних задач.

4. Зміст навчальної дисципліни

Тема 1 Програмне забезпечення проведення статистичного аналізу даних. Основні статистичні характеристики

Інструменти для роботи з даними. Електронні таблиці. Операції над змінними і спостереженнями. Операції з датами. Операції з таблицями даних. Призначення звіту і робочої книги. Описові статистики. Модуль Descriptive statistics та робота з ним. Інструменти для роботи з даними. Способи організації даних у пакету Statistica. Операції з таблицями даних. Обмін даними з іншими додатками. Налаштування та формування звітів у статистичних пакетах. Побудова таблиць частот. Визначення змінних, їх групування. Розбиття вибірки на нерівні діапазони. Представлення результатів аналізу. Таблиці кростабуляції і таблиці прапорів і заголовків. Багатовимірні відгуки та їх використання при аналізі неструктурованих даних

Тема 2 Непараметричний аналіз даних

Кореляція між змінними, кореляційні матриці. Критерій Стюдента порівняння середніх. Особливості проведення та типи t-тестів у статистичних пакетах. Групування і однофакторний дисперсійний аналіз: модуль Breakdown & one-way ANOVA. Кореляція Спірмена, тау Кендала, Гамма. Критерій серій Вальда-Волфовіца. Таблиці частот 2x2.

Тема 3 Лінійне та нелінійне моделювання взаємозв'язків між даними

Лінійна регресійна модель при проведенні статистичних досліджень. Робота з модулем Multiple Regression. Оцінка якості моделі. Способи проведення регресійного аналізу, особливості представлення результатів. Прогнозування значень відгуків по отриманій регресійній моделі. Лінійна регресійна модель: її місце при проведенні статистичних досліджень. Показники, що дозволяють оцінити якість моделі та їх розрахунок у статистичних пакетах. Способи перевірки якості отриманої регресійної моделі. Фіксована нелінійна регресія. Аналіз залишків. Умови проведення нелінійного багатовимірного моделювання. Перетворення, що приводять до отриманні лінійних моделей. Основні процедури нелінійного аналізу даних у статистичних пакетах. Робота у модулі Fixed Nonlinear Regression. Моделі бінарних відгуків. Функція втрат та її визначення. Оцінювання моделей регресії, цільових функцій та функцій втрат. Шматково-лінійна регресія. Регресія, визначена користувачем.

Тема 4 Методи дискримінантного та кластерного аналізу даних»

Сфери застосування та основні поняття дискримінантного аналізу. Методи дискримінації даних. Загальні моделі аналізу дискримінанта. Основні припущення та співвідношення кластерного аналізу. Правила об'єднання об'єктів у кластери. Древа класифікації і їх властивості. Покроковий дискримінантний аналіз. Дискримінантна функція для двох груп. Загальні моделі аналізу дискримінанта. Альтернативні процедури класифікації. Властивості дерев класифікації. Двовходове об'єднання даних. Метод К-середніх. Діалогові вікна створення дерев класифікації при роботі з пакетом Statistica.

Тема 5 Загальна модель факторного аналізу. Метод аналізу головних компонент

Призначення та загальна модель факторного аналізу. Робота у модулі Factor Analysis: типи вхідних даних, способи обробки пропущених даних. Метод аналізу головних компонент. Модуль Principal Components & Classification Analysis

Тема 6 Методи аналізу і прогнозування часових рядів

Поняття динамічного ряду даних. Етапи їх аналізу динамічних рядів даних. Методи та параметри оцінювання динамічних рядів даних. Модель ковзаючого інтегрованого середнього. Модель інтервенції для АРПСС. Сезонна декомпозиція. Спектральний (Фур'є) аналіз. Аналіз розподілених лагів. Ідентифікація моделі часових рядів. Модель АРПС та її оцінювання. Експоненціальне згладжування і прогнозування даних. Сезонна декомпозиція. XII-місячне сезонне коректування динамічних рядів. Розподілений лаг Алмана. Спектральний (Фур'є) аналіз. Кросс-спектральний аналіз часових рядів.

5. Очікувані результати навчання навчальної дисципліни

Після успішного вивчення навчальної дисципліни здобувач вищої освіти зможе:

РН1	знати сучасне математичне програмне забезпечення, межі його використання і структуру математичних пакетів
РН2	структуру статистичних пакетів
РН3	організацію дослідження та моделювання економічних процесів і явищ за допомогою сучасних статистичних пакетів (на прикладі пакету Statistica)
РН4	висувати, перевіряти різні припущення на основі статистичних моделей
РН5	аналізувати отримані результати; представляти результати розрахунків у аналітичній, числовій та графічній формі

6. Роль навчальної дисципліни у досягненні програмних результатів

Програмні результати навчання, досягнення яких забезпечує навчальна дисципліна.

Для спеціальності 051 Економіка:

ПР12	Застосовувати набуті теоретичні знання для розв'язання практичних завдань та змістовно інтерпретувати отримані результати
ПР19	Використовувати інформаційні та комунікаційні технології для вирішення соціально-економічних завдань, підготовки та представлення аналітичних звітів
ПР32	Застосовувати сучасні інформаційні технології у соціально-економічних дослідженнях

7. Види навчальних занять та навчальної діяльності

7.1 Види навчальних занять

Тема 1. Програмне забезпечення проведення статистичного аналізу даних. Основні статистичні характеристики

Лк1 "Організація даних та робота з ними у статистичних пакетах" (денна)

Інструменти для роботи з даними. Структура електронної таблиці. Основні операції над змінними і спостереженнями. Описові статистики. Модуль Descriptive statistics та робота з ним. Побудова таблиць частот. Визначення змінних, їх групування. Розбиття вибірки на нерівні діапазони. Представлення результатів аналізу. Таблиці кростабуляції і таблиці прапорів і заголовків. Багатовимірні відгуки та їх використання при аналізі неструктурованих даних.

<p>Лб1 "Побудова таблиць кростабуляції у пакеті Statistica. Групування даних. Операції з датами." (денна)</p> <p>Створити таблицю з даними за кодами значень. Виконання операцій з датами. Трансформація та перекодування даних.</p>
<p>Лб2 "Побудова таблиць кростабуляції у пакеті Statistica. Групування даних. Операції з датами." (денна)</p> <p>Створення таблиць кростабуляції. Побудова гістограм для графічного представлення результатів.</p>
<p>Лб3 "Статистична обробка масивів даних в пакеті Statistica" (денна)</p> <p>Групування даних в пакеті STATISTICA, дослідження рядів розподілів, візуалізація результатів. Визначення основних статистичних характеристик початкових даних</p>
<p>Лб4 "Статистична обробка масивів даних в пакеті Statistica" (денна)</p> <p>Графічне та табличне представлення отриманих результатів. Частотний аналіз даних з рівними та нерівними діапазонами. Перевірка гіпотез про відповідність вибірки певному закону розподілу.</p>
<p>Тема 2. Непараметричний аналіз даних</p>
<p>Лк2 "Однофакторний дисперсійний аналіз даних" (денна)</p> <p>Описові статистики. Модуль Descriptive statistics та робота з ним. Кореляція між змінними, кореляційні матриці. Критерій Стюдента порівняння середніх.</p>
<p>Лк3 "Однофакторний дисперсійний аналіз даних" (денна)</p> <p>Особливості проведення та типи t-тестів у пакеті Statistica. Групування і однофакторний дисперсійний аналіз: модуль Breakdown & one-way ANOVA.</p>
<p>Лб5 "Оцінювання вибірок в пакеті Statistica" (денна)</p> <p>Отримання функції спостережень для наближеного визначення невідомого параметра за допомогою різних способів оцінювання: методом моментів, методом максимальної правдоподібності, методом порядкових статистик.</p>
<p>Лб6 "Оцінювання вибірок в пакеті Statistica" (денна)</p> <p>Перевірка гіпотез про рівність середніх у групах.</p>
<p>Лб7 "Перевірка граничних теорем засобами пакету Statistica" (денна)</p> <p>Перевірка основних співвідношень з області математичної статистики: закону великих чисел, стиснення розкиду середніх значень при збільшенні обсягу вибірки. Перевірка закону великих чисел та центральної граничної теореми</p>
<p>Тема 3. Лінійне та нелінійне моделювання взаємозв'язків між даними</p>

<p>Лк4 "Побудова лінійних регресійних моделей при статистичних дослідженнях. Показники оцінки якості моделей." (денна)</p> <p>Лінійна регресійна модель при проведенні статистичних досліджень. Робота з модулем Multiple Regression. Оцінка якості моделі. Способи проведення регресійного аналізу, особливості представлення результатів. Прогнозування значень відгуків по отриманій регресійній моделі.</p>
<p>Лк5 "Методи та умови проведення нелінійного багатовимірного моделювання взаємозв'язків" (денна)</p> <p>Умови проведення нелінійного багатовимірного моделювання. Перетворення, що приводять до отриманні лінійних моделей. Основні процедури нелінійного аналізу даних у пакеті Statistica. Робота у модулі Fixed Nonlinear Regression. Моделі бінарних відгуків.</p>
<p>Лб8 "Побудова та оцінювання регресійних моделей" (денна)</p> <p>Побудувати однофакторну модель залежності результативної ознаки Y від факторної ознаки X. Встановити вид апроксимуючої функції. Визначити силу і напрям зв'язку змінними. Отримати рівняння регресії. Оцінити адекватність моделі. Здійснити прогнозування значення результативної змінної при вказаному значенні факторної змінної.</p>
<p>Лб9 "Побудова та оцінювання регресійних моделей" (денна)</p> <p>Побудувати багатфакторну модель залежності результативної ознаки Y від факторних ознак X відповідно до варіантів завдань. Встановити вид апроксимуючої функції. Визначити силу і напрям зв'язку між результативною змінною і кожною факторною змінною. Побудувати рівняння регресії. Оцінити адекватність моделі.</p>
<p>Лб10 "Побудова та оцінювання регресійних моделей" (денна)</p> <p>Розробка одно- та множинних лінійних регресійних моделей, оцінка їх адекватності. Здійснити прогнозування значення результативної змінної при вказаному значенні факторної змінної. Побудова нелінійних регресійних моделей різних типів</p>
<p>Тема 4. Методи дискримінантного та кластерного аналізу даних»</p>
<p>Лк6 "Класифікація даних методами дискримінантного аналізу" (денна)</p> <p>Сфери застосування та основні поняття дискримінантного аналізу. Методи дискримінації даних. Загальні моделі аналізу дискримінанта</p>
<p>Лк7 "Основні методи кластерного аналізу." (денна)</p> <p>Основні припущення та співвідношення кластерного аналізу. Правила об'єднання об'єктів у кластери. Деревя класифікації і їх властивості.</p>
<p>Лб11 "Дискримінантний і кластерний аналіз даних" (денна)</p> <p>Можливості пакету Statistica для проведення дискримінантного та кластерного аналізу.</p>
<p>Лб12 "Дискримінантний і кластерний аналіз даних" (денна)</p> <p>Побудова дерев класифікації та класифікаційних функцій.</p>

Лб13 "Дискримінантний і кластерний аналіз даних" (денна) Проведення кластеризації даних різними методами.
Тема 5. Загальна модель факторного аналізу. Метод аналізу головних компонент
Лк8 "Загальна модель факторного аналізу" (денна) Призначення та загальна модель факторного аналізу. Робота у модулі Factor Analysis: типи вхідних даних, способи обробки пропущених даних.
Лк9 "Метод аналізу головних компонент" (денна) Метод аналізу головних компонент. Модуль Principal Components & Classification Analysis
Лб14 "Методи дослідження багатовимірних даних засобами пакету «Statistica»" (денна) Проведення множинного лінійного регресійного аналізу з покроковим виключенням регресорів.
Лб15 "Методи дослідження багатовимірних даних засобами пакету «Statistica»" (денна) Компонентний і факторний аналіз: виявити загальні для системи ознак латентні чинники (компоненти), впливом яких обумовлені варіації ознак і їх коваріації.
Лб16 "Методи дослідження багатовимірних даних засобами пакету «Statistica»" (денна) Зниження розмірності системи початкових ознак, обмежившись декількома першими головним компонентами, на долю яких доводиться не меншого 70% загальної дисперсії ознак
Тема 6. Методи аналізу і прогнозування часових рядів
Лк10 "Етапи аналізу динамічних рядів." (денна) Поняття динамічного ряду даних. Етапи їх аналізу динамічних рядів даних. Методи та параметри оцінювання динамічних рядів даних.
Лк11 "Етапи аналізу динамічних рядів." (денна) Модель ковзаючого інтегрованого середнього. Модель інтервенції для АРПСС.
Лк12 ". Методи та параметри прогнозування часових рядів ." (денна) Сезонна декомпозиція. Спектральний (Фур'є) аналіз. Аналіз розподілених лагів.
Лб17 "Аналіз часових рядів і їх прогнозування у пакеті Statistica" (денна) Оцінити та визначити структуру параметрів часового ряду даних за допомогою різних методів аналізу та перетворень даних.
Лб18 "Аналіз часових рядів і їх прогнозування у пакеті Statistica" (денна) Вилучення тренду, автокореляцій з часових рядів даних.

ЛБ19 "Аналіз часових рядів і їх прогнозування у пакеті Statistica" (денна) Проведення Фур'є-аналізу та використання різних методів згладжування даних. Вилучення періодичних складових часових рядів.
ЛБ20 "Аналіз часових рядів і їх прогнозування у пакеті Statistica" (денна) Аналіз моделей з інтервенцією даних. Проведення аналізу розподілених лагів. За введеними даними виконати прогнозування значень часового ряду на декілька періодів.

7.2 Види навчальної діяльності

НД1	Виконання віртуальних лабораторних робіт
НД2	Виконання та презентація результатів лабораторної роботи
НД3	Підготовка до лекцій
НД4	Підготовка до лабораторного заняття
НД5	Підготовка до поточного та підсумкового контролю

8. Методи викладання, навчання

Дисципліна передбачає навчання через:

МН1	Інтерактивні лекції
МН2	Пошукова лабораторна робота
МН3	Практико-орієнтоване навчання

Лекції надають студентам матеріали з оптимізаційних моделей з різних точок зору, що є основою для самостійного навчання здобувачів вищої освіти (РН 1). Лекції доповнюються лабораторними заняттями, що надають студентам можливість застосовувати теоретичні знання на практичних прикладах (РН 2, РН 3, РН 4, РН5).

Зміст лабораторних робіт спрямований на практико-орієнтоване навчання, що передбачає визначення студентами доцільності застосування відповідних методів статистичних досліджень та економіко-математичних моделей для проведення економічного аналізу діяльності соціально-економічних об'єктів (фірм, підприємств, банків) з метою оптимізації їх діяльності (РН 2, РН 3, РН 4, РН 5). Самостійному навчанню сприятиме підготовка до лекцій, лабораторних занять, а також для проходження тестів та виконання індивідуального завдання.

9. Методи та критерії оцінювання

9.1. Критерії оцінювання

Шкала оцінювання ECTS	Визначення	Чотирибальна національна шкала оцінювання	Рейтингова бальна шкала оцінювання
A	Відмінне виконання лише з незначною кількістю помилок	5 (відмінно)	$90 \leq RD \leq 100$

B	Вище середнього рівня з кількома помилками	4 (добре)	$82 \leq RD < 89$
C	Загалом правильна робота з певною кількістю помилок	4 (добре)	$74 \leq RD < 81$
D	Непогано, але зі значною кількістю недоліків	3 (задовільно)	$64 \leq RD < 73$
E	Виконання задовольняє мінімальні критерії	3 (задовільно)	$60 \leq RD < 63$
FX	Можливе повторне складання	2 (незадовільно)	$35 \leq RD < 59$
F	Необхідний повторний курс з навчальної дисципліни	2 (незадовільно)	$0 \leq RD < 34$

9.2 Методи поточного формативного оцінювання

МФО1	Самооцінка поточного тестування
МФО2	Захист презентацій та рефератів
МФО3	Проведення розрахунків

9.3 Методи підсумкового сумативного оцінювання

МСО1	Звіт за результатами виконання лабораторних робіт
МСО2	Поточні контрольні роботи (проміжний модульний контроль)
МСО3	Складання комплексного письмового модульного контролю
МСО4	Підсумковий контроль: екзамен
МСО5	Написання та захист курсової роботи

Контрольні заходи:

6 семестр		100 балів
МСО1. Звіт за результатами виконання лабораторних робіт		40
	10x4	40
МСО2. Поточні контрольні роботи (проміжний модульний контроль)		10
		10
МСО3. Складання комплексного письмового модульного контролю		10
		10
МСО4. Підсумковий контроль: екзамен		40
		40

Контрольні заходи в особливому випадку:

6 семестр		100 балів
МСО1. Звіт за результатами виконання лабораторних робіт		30

	5x6	30
МСО2. Поточні контрольні роботи (проміжний модульний контроль)		20
	2x10	20
МСО3. Складання комплексного письмового модульного контролю		10
		10
МСО4. Підсумковий контроль: екзамен		40
		40

Курсова робота:

6 семестр	100 балів
МСО5. Написання та захист курсової роботи	100
	100

Оцінювання знань студента підчас лабораторних занять має на меті контроль активності підчас виконання роботи (МФО2), перевірку рівня підготовленості студента до виконання конкретної роботи, для чого студент в аудиторії виконує практичні тестові завдання (МФО1), наявність ідеї, вміння обґрунтовувати свою позицію, здійснювати узагальнення інформації та робити висновки (МФО3).

10. Ресурсне забезпечення навчальної дисципліни

10.1 Засоби навчання

ЗН1	Комп'ютери, комп'ютерні системи та мережи
ЗН2	Інформаційно-комунікаційні системи
ЗН3	Бібліотечні фонди

10.2 Інформаційне та навчально-методичне забезпечення

Основна література	
1	А. А. Халафян, В. П. Боровиков Г. В. Калайдина. Теория вероятностей, математическая статистика и анализ данных: Основы теории и практика на компьютере. – ЛЕНАНД, 2017. – 440 с. ISBN: 978-5-9710-3040-9
2	В.П. Боровиков. Популярное введение в современный анализ данных и машинное обучение на Statistica. – 2018 – 354 с. ISBN 978-5-9912-0738-6
Допоміжна література	
1	А. А. Халафян. Промышленная статистика. Контроль качества, анализ процессов, планирование экспериментов в пакете STATISTICA. –М: URSS, 2017 – 480 с.
Інформаційні ресурси в Інтернеті	
1	Олійник В.М. Програмне забезпечення статистичного аналізу. 2020. https://classroom.google.com/u/1/c/NTM4MjQ5NzU1NDFa Код доступу : jnun5d