



ДЕРЖАВНИЙ ТОРГОВЕЛЬНО-ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ

Факультет інформаційних технологій
Кафедра цифрової економіки та системного аналізу

СИЛАБУС (SYLLABUS)

Дисципліна «Інструментальні засоби бізнес-аналітики /
Business intelligence tools»

ІНФОРМАЦІЯ ПРО ВИКЛАДАЧА

Викладач	Роскладка Андрій Анатолійович
Науковий ступінь	Доктор економічних наук
Вчене звання	Професор
Посада	Завідувач кафедри цифрової економіки та системного аналізу
Адреса кафедри	м.Київ, вул. Кіото 19, каб. Б 516
E-mail	prikmath@bigmir.net
Консультації	Відповідно до графіку індивідуальних консультацій на сайті кафедри

ПОЛІТИКА АКАДЕМІЧНОЇ ДОБРОЧЕСНОСТІ

<https://knu-te.edu.ua/file/MzEyMQ==/c12a9f74e87d9154696ca0f761da2e5c.pdf>

Дотримання академічної доброчесності передбачає:

- самостійне виконання навчальних завдань, завдань поточного та підсумкового контролю результатів навчання (для осіб з особливими освітніми потребами ця вимога застосовується з урахуванням їхніх індивідуальних потреб і можливостей);
- посилання на джерела інформації у разі використання не авторських ідей, розробок, тверджень, відомостей і т.п.;
- дотримання норм законодавства про авторське право і суміжні права;
- надання достовірної інформації про результати власної наукової діяльності, використанні методики досліджень і джерела інформації.

Порушенням академічної доброчесності вважається:

- академічний плагіат – оприлюднення (частково або повністю) наукових (творчих) результатів, отриманих іншими особами, як результатів власного дослідження (творчості) та/або відтворення опублікованих текстів (оприлюднених творів мистецтва) інших авторів без зазначення авторства;
- самоплагіат – оприлюднення (частково або повністю) власних раніше опублікованих наукових результатів як нових наукових результатів;
- фабрикація – вигадкування даних чи фактів, що використовуються в наукових дослідженнях;
- фальсифікація – свідомо зміна чи модифікація вже наявних даних, що стосуються наукових досліджень.

За порушення академічної доброчесності здобувачі освіти можуть бути притягнені до академічної відповідальності:

- повторне проходження оцінювання (модульний контроль, іспит, залік тощо);
- повторне проходження відповідного освітнього компонента освітньо-професійної програми;
- відрахування з Університету;
- позбавлення наданих університетом пільг;
- відмова у присудженні відповідного ступеня вищої освіти;

ПОЛІТИКА ЩОДО ВІДВІДУВАННЯ ЗАНЯТЬ

- відвідування занять є обов'язковим;
- за об'єктивних причин (наприклад, хвороба, міжнародне стажування та ін.) навчання може відбутись в он-лайн формі за погодженням із викладачем дисципліни.

ОПИС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Назва дисципліни / тип дисципліни	Інструментальні засоби бізнес-аналітики /вибіркова
Навчальний рік	2023-2024
Факультет	Факультет інформаційних технологій
Курс	4
Семестр	7-8
Освітній ступінь	Бакалавр
Галузь знань	12 «Інформаційні технології»
Спеціальність	126 «Інформаційні системи та технології»
Загальна характеристика	Кількість годин –180 Кількість кредитів – 6 Види занять: лекції, практичні, самостійна робота. Співвідношення аудиторних годин і годин самостійної роботи - 48/132 Мова викладання – українська Форма викладання – очна
Підсумковий контроль	Екзамен
Програмне забезпечення	Пакет Microsoft Office
Обладнання	Проектор, комп'ютерна техніка із встановленим програмним забезпеченням та доступом до мережі Інтернет.
Необхідні попередні дисципліни	Навчальні дисципліни «Математичний аналіз», «Дискретна математика».
Методика вивчення	Методика вивчення дисципліни полягає у набутті студентами знань теоретичного і практично-прикладного характеру під час лекцій, практичних занять, самостійної роботи та вивчення першоджерел і навчально-методичної літератури.
Мета і завдання	Метою вивчення дисципліни «Інструментальні засоби бізнес-аналітики» є формування у майбутніх фахівців необхідного рівня фундаментальних знань та набуття практичних навичок використання мови програмування R та її бібліотек для реалізації процесів збору, первинної обробки, трансформації, аналізу та візуалізації даних, виявлення нових властивостей та закономірностей у бізнес-інформації, створення моделей і програмних систем бізнес-аналітики. Завданням вивчення дисципліни «Інструментальні засоби бізнес-аналітики» є надання студентам ґрунтовних знань в глибокого аналізу даних у процесі цифрової трансформації бізнесу.
Місце дисципліни в освітньо-професійній програмі	
Загальні компетентності	КЗ 1. Здатність до абстрактного мислення, аналізу та синтезу. КЗ 2. Здатність застосовувати знання у практичних ситуаціях. КЗ 6. Здатність до пошуку, оброблення та узагальнення інформації з різних джерел.
Фахові компетентності (результати навчання)	КС 1. Здатність аналізувати об'єкт проектування або функціонування та його предметну область. КС 4. Здатність проектувати, розробляти та використовувати засоби реалізації інформаційних систем, технологій та інфокомунікацій

(методичні, інформаційні, алгоритмічні, технічні, програмні та інші). КС 6. Здатність використовувати сучасні інформаційні системи та технології (виробничі, підтримки прийняття рішень, інтелектуального аналізу даних та інші), методики й техніки кібербезпеки під час виконання функціональних завдань та обов'язків. КС 13. Здатність проводити обчислювальні експерименти, порівнювати результати експериментальних даних і отриманих рішень.

ТЕМАТИКА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Тема 1. Моделювання даних у бізнес-аналітиці

Основні поняття бізнес-аналітики. Аналітичні дані. Види аналітики. Актуальні бізнес-задачі аналітики даних. Знайомство з R. Доступ до інсталяції програмного середовища. Редагування та виконання коду. Принципи пакетної організації обчислень. Робота з великими масивами даних бізнесу. Графічний інтерфейс GUI для R. RStudio. Робота з консоллю. Термінал RStudio. Редагування та виконання коду. Автоматичне створення функцій. Коментування коду. Історія команд. Робочий простір Workspace. Введення даних з клавіатури. Імпорт аналітичних даних з текстового файлу, файлу Excel, XML-файлу. Добування бізнес-даних із веб-сторінок. Імпорт даних із систем управління базами даних.

Тема 2. Управління даними в R.

Синтаксис мови R. Змінні. Типи об'єктів бізнес-аналізу: вектори, матриці, масиви даних, таблиці даних, фактори, списки. Робота з Data Frame. Перетворення даних (Reshape, Split, Combine). Робота із пропущеними даними. Перетворення типів. Сортування даних. Об'єднання наборів даних. Поділ множини даних на складові. Функції R. Числові і текстові функції. Умовні оператори. Цикли. Функції користувача. Експорт результатів аналізу із середовища R.

Тема 3. Описова та індуктивна бізнес-аналітика.

Описова статистика даних бізнес-процесів. Оцінка вибірових параметрів. Найбільш поширені помилки при виконанні статистичного аналізу. Таблиці частот та спряженості. Кореляції. Перевірка статистичної значущості кореляцій. Візуалізація кореляцій. Закони розподілу випадкових даних: біноміальний, рівномірний, експонентний, нормальний розподіл, розподіл Пуассона. Підгонка розподілів. Аналіз викидів бізнес-даних. Заповнення пропущених значень у таблицях даних. Відновлюваність результатів спостережень з використанням генератора випадкових чисел. Основні кроки побудови та верифікації моделей. Статистична перевірка гіпотез. Класичні методи і критерії статистики. Тест Шапіро-Уїлка. Рангові критерії Уїлкінсона-Манна-Уїтні. Тести Стьюдента для залежних і незалежних вибірок даних. Непараметричні тести міжгрупових різниць. Рандомізація, бустреп та оцінка статистичної потужності вибірки. Оцінка кореляції. Тест хі-квадрат (Пірсона). Тест Колмогорова-Смірнова. Критерій Фішера. Точний тест Фішера. Критерії Мак-Немара та Кохрана-Мантеля-Хензеля. Дисперсійний аналіз (ANOVA). Однофакторний дисперсійний та коваріаційний аналіз. Лінійні моделі дисперсійного аналізу. Дисперсійний аналіз за Краскелом-Уолісом. Багатовимірний дисперсійний аналіз.

Тема 4. Візуалізація даних та результатів бізнес-аналізу

Створення графіки в R. Робота з діаграмами показників бізнес-процесу. Прості методи візуалізації бізнес-даних: точкові діаграми, стовпчикові діаграми, кругові та секторні діаграми, спінограми, гістограми, діаграми розмахів, діаграми ядерної оцінки функції щільності, скрипкові діаграми. Графічні параметри візуалізації даних: розташування діаграм на сторінці, системи координат, колір та прозорість, налаштування відображення графіка, легенда, заголовки графіка та осей, розмітка осей, підписи та анотації, фасети, графічні теми.

Розширена графіка для просунутої бізнес-аналітики. Чотири графічні системи R: Base, Grid, Lattice, ggplot2. Пакет lattice. Умовні змінні. Функції для зміни формату даних. Пакет ggplot2. Інтерактивна графіка. Взаємодія з діаграмами: ідентифікація точок. Пакет playwith. Пакет latticist. Створення інтерактивної графіки за допомогою пакета iplots. Пакет rggobi. Діаграми розсіювання, бульбашкові діаграми, лінійні графіки, корелограми, мозаїчні діаграми, діаграми Клівленда. Категоризовані графіки.

Тема 5. Прогностична бізнес-аналітика.

Регресійний бізнес-аналіз. Моделі регресії при різних видах функції втрат. МНК-регресія: проста лінійна, поліноміальна, множинна лінійна регресія. Діагностика та відбір регресійних моделей. Мультиколінеарність. Нестандартні спостереження за реальними даними: викиди, точки високої напруженості, впливові спостереження. Способи коригування нестандартних спостережень: видалення спостережень, перетворення змінних, додавання або видалення змінних. Узагальнені лінійні моделі і функція glm(). Логістична регресія в R. Пуасонівська регресія. Моделі згладжування. Узагальнені моделі регресії. Моделі пробіт-та логіт-регресії. Вибір оптимальної регресійної моделі бізнес-даних. Порівняння моделей. Критерії вибору моделей оптимальної складності. Крос-валідація. Відносна важливість. Прогнозування даних в R. Порівняння ефективності різних моделей при прогнозуванні.

Тема 6. Розширена бізнес-аналітика

Задачі Data Mining в R. Бінарні матриці та асоціативні правила. Пошук логічних закономірностей у даних. Бінарні класифікатори. Дискримінантний аналіз. Моделі класифікації. Іриси Фішера і метод k- neighbors. Класифікаційні дерева, основні принципи їх роботи та імплементація в R. Наївний класифікатор Байеса. Кластерний аналіз. Бібліотека NbClust. Багатомірне шкалювання для візуалізації кластерів. Ієрархічна кластеризація. Алгоритм побудови дендрограми. Самоорганізовані карти Кохонена в задачах кластеризації та сегментації бізнес-даних. Макет призначеного для користувача інтерфейсу. Додавання керуючих віджетів. Пакети gWidgets та fgui. Основні підходи до розробки веб-програм. Пакет shiny. Поширення створених веб-додатків. Способи сканування бізнес-даних. CSS-селектори. Web scraping з пакетом rvest. PhantomJS і збереження вмісту динамічних веб-сторінок. Збір інформації за допомогою API. Просторовий аналіз даних і створення картограм. Моделі просторових даних. Векторні та растрові дані. Проста карта: використання растрового рисунка і підрахунок відстаней. Аналіз просторового розміщення даних бізнес-аналізу. Особливості аналізу та візуалізації просторових даних. Використання сервісів картографічної системи Google Maps. Побудова картограм з використанням R.

Перелік навчальних робіт студентів та оцінки їх у балах з дисципліни «Інструментальні засоби бізнес-аналітики»

Назва теми	Кількість годин				Форма контролю
	Усього годин / кредитів	з них:			
		лекції	лабораторні заняття / МК	самостійна робота студентів	
Тема 1. Моделювання даних у бізнес-аналітиці	48	8	14	26	Т, ІДЗ, П, Пр
Тема 2. Управління даними в R	44	6	14	24	Т, ІДЗ, П, Пр, МК
Тема 3. Описова та індуктивна бізнес-аналітика	46	8	8	30	Т, ІДЗ, П, Пр
Тема 4. Візуалізація даних та	39	6	8	25	Т, ІДЗ,

результатів бізнес-аналізу					П, Пр
Тема 5. Прогностична бізнес-аналітика	36	6	10	20	Т, ІДЗ, Пр, П, МК
Тема 6. Розширена бізнес-аналітика	12	2	0	10	Т, Пр
Разом	225/7.5	36	54	135	х
Підсумковий контроль – екзамен					

Примітка: Т – тестування; ІДЗ – виконання індивідуальних домашніх завдань; П – перевірка індивідуальних завдань; Пр. – презентація індивідуального завдання; МК – модульний контроль

КОНТРОЛЬ ТА КРИТЕРІЇ ОЦІНЮВАННЯ ЗНАНЬ СТУДЕНТІВ

При вивченні дисципліни використовуються наступні форми контролю знань студентів: поточний; модульний; підсумковий.

Поточний контроль передбачає перевірку теоретичних питань, самостійної роботи, практичних робіт та усне опитування по кожній практичній роботі. По даному виду контролю оцінювання знань здійснюється у відповідності до бального розподілу наведеного в попередній таблиці.

Модульний контроль передбачає виконання модульної контрольної роботи. Всі завдання оцінюються в 20 балів. Перше завдання (теоретичне) – 4 бали, друге завдання (практичне) – 8 балів, третє завдання (практичне) – 8 балів.

Формою підсумкового контролю є екзамен. Екзаменаційна оцінка (100 балів) є результатом виконання двох теоретичних питань (2 x 20 балів = 40 балів) та практичного завдання (60 балів).

Результуюча оцінка з дисципліни визначається як середня від балів набраних протягом семестру та отриманих на іспиті.

СПИСОК РЕКОМЕНДОВАНИХ ДЖЕРЕЛ

Основні

1. Etaati L. Advanced Analytics with Power BI and R. - Auckland: RADACAD Systems Limited, 2017. - 179 p.
2. Hafner S. An Introduction to R for Beginners. - Hafner Consulting LLC, 2019. - 360 p.
3. Kabacoff R. R in Action. Data analysis and graphics with R. – Manning: Shelter island, 2015. – 608 p
4. *Roskladka A., Ivanova O., Kulazhenko V. Data Scientist: a glance into the future // Зовнішня торгівля: економіка, фінанси, право. 2019. № 3. С. 109-120*
5. Негрей М., Гнот Т. Аналітика з R: навч. посібник / М. Негрей, Т. Гнот. – Київ: Компрінт, 2020. – 236 с.