

Код та назва дисципліни українською мовою/ Назва дисципліни англійською мовою	2-ф11-6 Дослідницький аналіз даних / Exploratory Data Analysis (EDA)
Рекомендується для галузі знань	A – Освіта, E – Природничі науки, математика та статистика, F – Інформаційні технології, G – Інженерія, виробництво та будівництво
Кафедра	Статистики й теорії ймовірностей
П.І.П. НПП	Карнаух Євген Володимирович
Рівень ВО	Другий (магістерський)
КУРС, семестр (в якому буде викладатись)	1 курс, парний семестр
Мова викладання	Українська
Пререквізити (передумови вивчення дисципліни)	«Вища математика («Теорія ймовірностей та математична статистика») або «Теорія ймовірностей та математична статистика» або «Теорія ймовірностей» або «Теорія ймовірностей, ймовірнісні процеси та математична статистика»
Чому це цікаво/треба вивчати	Дослідницький аналіз даних (Exploratory Data Analysis, EDA) – це процес дослідження даних з метою виявлення важливих характеристик, закономірностей та аномалій перед глибшим аналізом чи моделюванням. EDA допомагає зрозуміти, як дані розподілені, чи містять вони пропущені значення, аномалії або кореляції між змінними.
Перелік тем з дисципліни	Тема 1. Джерела даних: відкриті набори, веб-скрапінг, API, бази даних Тема 2. Очищення та попередня обробка даних Тема 3. Описова статистика Тема 4. Одновимірний аналіз Тема 5. Аналіз залежностей між змінними Тема 6. Візуалізація взаємозв'язків Тема 7. Виявлення прихованих структур у даних Тема 8. Автоматичний вибір найкращих ознак
Як можна користуватися набутими знаннями і уміннями	Опанування дослідницького аналізу даних відкриває широкі кар'єрні перспективи у різних галузях, зокрема у сфері аналітики даних (Data Analyst, Business Analyst, Machine Learning Engineer, Financial Analyst, Marketing Data Analyst, Data Scientist). Володіння навичками збору, очищення, аналізу та візуалізації даних є важливим і затребуваним у бізнесі, фінансах, медицині, маркетингу та ІТ.
Очікувані результати навчання	Після опанування дисципліни студенти зможуть самостійно проводити дослідницький аналіз даних, використовувати методи візуалізації, працювати з великими масивами інформації та отримувати корисні нові знання для ухвалення рішень.
Інформаційне забезпечення	Відкриті набори даних Kaggle та UCI Machine Learning Repository, основна мова для реалізації EDA – Python Робоча програма навчальної дисципліни, інтернет-ресурси
Види навчальних занять	Лекції, практичні заняття
Вид семестрового контролю	Диференційований залік
Максимальна кількість здобувачів на семестр	без обмежень

Декан факультету _____ Олександр ХАМІНІЧ