

Міністерство освіти і науки України
Черкаський національний університет імені Богдана Хмельницького

Навчально-науковий інститут інформаційних та освітніх технологій
Кафедра автоматизації та комп'ютерно-інтегрованих технологій

АНАЛІЗ І ВІЗУАЛІЗАЦІЯ ДАНИХ

Методичні рекомендації
для організації самостійної роботи студентів у форматі командного проєкту
для здобувачів вищої освіти першого (бакалаврського) рівня
за спеціальністю 014 Середня освіта (014.09 Інформатика)

Черкаси 2024

Укладач і автор: професор кафедри автоматизації та комп'ютерно-інтегрованих технологій, доктор педагогічних наук Луценко Г.В.

Рекомендовано вченою радою Черкаського національного університету імені Богдана Хмельницького (протокол № 3 від 24 жовтня 2024 року).

Аналіз і візуалізація даних: методичні рекомендації для організації самостійної роботи студентів у форматі командного проєкту для здобувачів вищої освіти першого (бакалаврського) рівня за спеціальністю 014 Середня освіта (014.09 Інформатика) [Електронне видання] / Г.В. Луценко – Черкаси: Черкаський національний університет імені Богдана Хмельницького, 2024. – 24 с.

Рецензенти:

Гриценко В.Г., доктор педагогічних наук, професор кафедри автоматизації та комп'ютерно-інтегрованих технологій Черкаського національного університету імені Богдана Хмельницького Гриценко В.Г.

Русіна Н.Г., кандидат педагогічних наук, доцент кафедри теорії та технології програмування факультету комп'ютерних наук та кібернетики Київського національного університету імені Тараса Шевченка

Методичні матеріали призначені для студентів денної форми навчання зі спеціальності 014 Середня освіта (014.09 Інформатика) ННІ інформаційних та освітніх технологій для виконання командного проєкту з аналізу даних, а також для викладачів ННІ інформаційних та освітніх технологій за кафедрою автоматизації та комп'ютерно-інтегрованих технологій.

ЗМІСТ

ПЕРЕДМОВА	4
1. ЗАГАЛЬНА ІНФОРМАЦІЯ	5
2. ОСНОВНІ ЕТАПИ РОБОТИ НАД ПРОЄКТОМ.....	6
3. ДОКУМЕНТУВАННЯ ПРОЄКТУ	7
3.1. Проміжний звіт 1: Збір і попередній аналіз даних	8
3.2. Проміжний звіт 2: Моделювання даних.....	9
3.3. Підсумковий звіт: Аналіз та візуалізація даних	9
3.4. Вимоги до підсумкового виступу та презентації результатів.....	10
4. ОЦІНЮВАННЯ ПРОЄКТУ	12
5. ПЛАН ВИКОНАННЯ КОМАНДНОГО ПРОЄКТУ	16
6. ПРИКЛАД ВИКОНАННЯ ПРОЄКТНОГО ЗАВДАННЯ З ВИКОРИСТАННЯМ GOOGLE ТАБЛИЦЬ.....	18
ДЖЕРЕЛА ТА РЕКОМЕНДОВАНА ЛІТЕРАТУРА	22

ПЕРЕДМОВА

Метою самостійної роботи студентів у форматі командного проєкту з дисципліни «Аналіз і візуалізація даних» є формування у студентів денної форми навчання зі спеціальності 014 Середня освіта (014.09 Інформатика) ННІ інформаційних та освітніх технологій навичок аналізу реальних наборів даних та розвиток вміння працювати в команді. Командний проєкт також спрямований на те, щоб майбутні учителі інформатики могли інтегрувати отримані теоретичні знання в практичні завдання, подібні до тих, з якими можуть зіткнутися в професійній діяльності.

Командний проєкт з аналізу і візуалізації даних є важливим елементом для вивчення курсу «Аналіз і візуалізація даних», адже його реалізація передбачає застосування сучасних методів збору, очищення та обробки даних, алгоритмів та статистичних методів для аналізу великих обсягів інформації та формування вміння інтерпретувати отримані результати, робити висновки, приймати рішення і пропонувати рекомендації щодо подолання негативних явищ.

Командний проєкт сприяє розвитку загальних компетентностей, зокрема, здатності співпрацювати з фахівцями з різних галузей, обговорювати ідеї та приймати рішення колективно. Реалізація проєктів групами студентів передбачає розподіл завдань відповідно до сильних сторін учасників команди. Студенти отримують можливість розвивати відповідальність за свою частину роботи та підтримувати загальний прогрес команди.

Завданням командного проєкту є удосконалення навичок презентувати результати дослідження в наочному форматі, що є критично важливим для ефективної комунікації з експертами з різних галузей. Студенти навчаються обирати відповідні інструменти та техніки для створення інформативних і зрозумілих візуалізацій; правильно інтерпретувати дані через графіки, діаграми та інші візуальні засоби; представляти складну інформацію таким чином, щоб вона була зрозумілою як для технічних фахівців, так і для широкої аудиторії.

Аналіз даних вимагає вміння критично оцінювати інформацію, знаходити закономірності, формулювати гіпотези та перевіряти їх. У процесі роботи над проєктом студенти навчаються ставити правильні питання і шукати на них відповіді через аналіз даних, розвивають здатність бачити проблеми в наборі даних (аномалії, пропуски) та вирішувати їх, формують аналітичне мислення, яке допомагає їм у майбутньому професійному житті.

1. ЗАГАЛЬНА ІНФОРМАЦІЯ

Командний проєкт з дисципліни «Аналіз і візуалізація даних» є обов'язковим завданням курсу, що реалізовується студентами під час годин, відведених на самостійну роботу.

Командний проєкт – це завдання, яке виконується групою студентів. Він передбачає розподіл ролей, спільну роботу над досягненням спільної мети та розвиток навичок командної взаємодії.

Для студентів, які планують стати вчителями інформатики, командні проєкти мають особливе значення. Вони допомагають розвинути навички, необхідні для ефективної роботи в освітньому середовищі. Професійний стандарт за професіями "Вчитель початкових класів закладу загальної середньої освіти", "Вчитель закладу загальної середньої освіти", "Вчитель з початкової освіти (з дипломом молодшого спеціаліста)" наголошує на важливості не лише предметних знань, а й на розвитку особистісних якостей, необхідних для ефективної педагогічної діяльності. Однією з таких якостей є здатність до командної роботи (Мінекономіки, 2020). Хоча в Стандарті немає прямого переліку всіх навичок командної роботи, з його положень можна виокремити окремі аспекти, ефективність формування яких підвищується в умовах проєктно орієнтованого навчання.

Серед ключових навичок виділимо власне роботу в команді, що передбачає вміння будувати довірчі стосунки, знаходити компроміси, спільно діяти для вирішення проблем. Студенти також навчаються планувати завдання та час завдань проєкту, розподіляти ролі й обов'язки, координувати дії інших членів команди. У рамках проєкту студенти зможуть взяти на себе роль лідера, що допоможе їм розвинути навички організації та мотивації інших.

Ефективна комунікація є основою будь-якої командної роботи. Реалізуючи проєкт, студенти навчаються чітко формулювати свої думки, слухати інших та вирішувати конфлікти.

Під час виконання проєкту студенти розвивають навички критичного мислення, аналізу інформації, пошуку рішень та прийняття обґрунтованих рішень. Командні проєкти стимулюють творчий підхід до вирішення завдань та пошуку нестандартних рішень, здатність генерувати нові ідеї.

Студенти отримують можливість застосувати свої теоретичні знання на практиці та розв'язувати реальні проблеми. Важливим аспектом командного проєкту є використання широкого спектру цифрових застосунків освітнього, організаційного та вузькоспеціалізованого спрямування.

Загальні і фахові компетентності

ЗК 2. Здатність до міжособистісної взаємодії, роботи в команді, спілкування з представниками інших професійних груп різного рівня.

ЗК 5. Здатність до генерування нових ідей, виявлення та розв'язання проблем, ініціативності та підприємливості.

ФК 7. Здатність орієнтуватися в інформаційному просторі, шукати, обробляти, організовувати, візуалізувати й критично оцінювати інформацію, оперувати нею у професійній діяльності.

ФК 15. Здатність демонструвати знання, розуміння та навички використання наукових фактів, теорій, принципів і методів вищої математики, дискретної математики, теорії ймовірностей та математичної статистики.

ФК 17. Здатність демонструвати використання методів математичного та інформаційного моделювання; здійснювати комп'ютерний експеримент.

Програмні результати навчання

ПРН 2. Демонструвати знання, розуміння й навички застосування вищої математики, дискретної математики, теорії ймовірностей та математичної статистики в обсязі, необхідному для користування математичним апаратом і методами в сфері інформатики та робототехніки.

ПРН 5. Аргументовано обирати й використовувати сучасні програмні й технічні засоби для реалізації методів та алгоритмів розв'язання професійних задач в сфері інформатики та робототехніки.

ПРН 9. Застосовувати освітні технології розвитку системного, критичного та проєктного мислення учнів, використовувати міжпредметні зв'язки та інтеграцію змісту різних освітніх галузей.

ПРН 14. Реалізовувати педагогічні дослідження на засадах академічної доброчесності та з використанням сучасних методів наукового пошуку.

ПРН 16. Шукати й обробляти дані, інформацію та контент, використовуючи різні стратегії; аналізувати, порівнювати, критично оцінювати та застосовувати їх.

2. ОСНОВНІ ЕТАПИ РОБОТИ НАД ПРОЄКТОМ

Тривалість проєкту – 6 тижнів. Команди формуються з 3-4 осіб, з урахуванням індивідуальних сильних сторін кожного учасника. Рекомендовано, щоб кожна команда мала представників із навичками роботи з цифровими технологіями і сервісами, планування, презентації тощо. Важливо, щоб учасники команд домовилися про розподіл ролей на початковому етапі роботи, включаючи лідера команди, що координуватиме роботу та відповідатиме за комунікацію.

Команди мають визначитися які цифрові інструменти команди будуть використовувати для комунікації між учасниками та співпраці. Учасники команди повинні регулярно зустрічатися (в оффлайн або в онлайн форматі) для обговорення прогресу роботи та вирішення проблем.

Тема проєкту має бути погоджена з викладачем для перевірки складності й релевантності. Вона повинна відображати актуальні проблеми освітньої сфери, для вирішення яких застосовується аналіз і візуалізація даних (наприклад, аналіз успішності учнів, вивчення патернів користувацької поведінки, соціологічні дослідження тощо). Рекомендовано обрати тему, яка буде корисною для майбутньої педагогічної діяльності студентів.

Студенти повинні самостійно зібрати або знайти необхідний набір даних для аналізу, використовуючи сервіси відкритих даних, наприклад, Kaggle, UCI Machine Learning Repository, Google Search Engine тощо. Важливо провести первинну обробку отриманого набору даних: очистити дані від шуму, перевірити наявність пропусків, усунути дублікати записів і підготувати дані до подальшого аналізу.

На етапі аналізу даних студенти повинні застосувати відповідні методи, які вони опановують в рамках курсу «Аналіз і візуалізація даних»: описову статистику, кореляційний аналіз, перевірку гіпотез, регресійний аналіз, кластеризацію тощо. Важливо аргументувати вибір кожного методу та обґрунтувати його застосування.

Ключовою частиною проєкту є візуалізація даних, яка має бути зрозумілою та інформативною. Це можуть бути графіки трендів, кореляцій, прогностичних моделей, 3D-візуалізації або інтерактивні графіки. Використання спеціалізованих цифрових інструментів для візуалізації даних, сприятиме ефективності її презентації. Особливу увагу слід приділити інтерпретації даних та формулюванню висновків, що відповідають цілям дослідження.

Захист проєкту відбувається у формі презентації результатів. Кожен учасник команди має представити свій внесок. Рекомендується використання мультимедійних засобів (слайди, діаграми, графіки) для наочного пояснення отриманих даних та висновків. Після презентації проєкту викладач та одногрупники можуть поставити питання, що стосуються виконаної роботи.

3. ДОКУМЕНТУВАННЯ ПРОЄКТУ

Весь процес виконання проєкту повинен бути належним чином задокументований. Документація проєкту – це невід'ємна частина будь-якого дослідження, особливо командного. Вона дозволяє не тільки продемонструвати

виконану роботу, але й забезпечує можливість повернутися до проекту в майбутньому, проаналізувати допущені помилки та досягнуті успіхи. Під час виконання проекту варто фіксувати відомості про проблеми, які виникли, рішення, які були знайдені. Обов'язково має фіксуватися розподіл завдань між членами команди.

Рекомендовано зберігати всі версії документів (проекти звітів, презентацій, кодів програм тощо) в одному зі хмарних сховищ, з обов'язковим доступом до них усіх учасників команди. Це допоможе простежити еволюцію проекту, відновити попередні версії за необхідності та запобігти втраті інформації. Усі використані дані мають зберігатися в організованому вигляді. У ході реалізації проектного завдання слід робити скріншоти інтерфейсів програмних сервісів, графіків, таблиць та інших візуальних елементів проекту. Це допоможе проілюструвати різні етапи роботи на етапі підготовки звіту та фінальної презентації.

У ході реалізації проекту передбачено два проміжні звіти (2 і 4 тижні, відповідно) та підсумковий звіт (6 тижень). Нижче наведено структуру звітів та питання, що мають в них розкриватися.

3.1. Проміжний звіт 1: Збір і попередній аналіз даних

Мета цього звіту — оцінити прогрес команди у зборі даних та виконанні початкового аналізу.

Зміст звіту:

1. Опис джерела даних:

- Звідки отримані дані (напр. відкриті джерела, власне зібрані дані).
- Короткий опис змінних та їх типів (числові, категоріальні, текстові).

2. Попередня обробка даних:

- Які дії вже проведено для очистки даних (виправлення пропусків, некоректних значень).
- Наявність пропусків, дублікати та інші проблеми.

3. Експлоративний аналіз даних (EDA):

- Обчислення базових статистичних показників: середнє значення, медіана, мінімум, максимум для основних змінних.
- Побудова попередніх візуалізацій (гістограми, діаграми розсіювання тощо).

4. План подальших дій:

- Які проблеми виявлені під час збору та аналізу даних.
- Наступні кроки щодо очищення та обробки даних, побудови моделей.

Формат звіту:

- Документ у форматі PDF або Google Docs.
- Додати таблиці з обчисленнями та графіки для візуалізації результатів попереднього аналізу.
- Обсяг: 3-5 сторінок.

3.2. Проміжний звіт 2: Моделювання даних

Мета цього звіту — оцінити, як команда підготувала дані для моделювання і які результати отримані на цьому етапі. Зміст звіту:

1. Очищення і підготовка даних:

- Опис процесу обробки пропусків та аномалій.
- Нормалізація або кодування змінних (якщо це було необхідно).

2. Побудова моделей:

- Які моделі було обрано для аналізу (лінійна регресія, кластеризація, класифікація тощо).
- Обґрунтування вибору моделей та опис, як підготовлені дані були використані в моделях.

3. Оцінка результатів моделей:

- Основні метрики для оцінки точності моделей.
- Порівняння результатів різних моделей, якщо було використано кілька підходів.

4. Проблеми і можливі покращення:

- Визначити, які складнощі виникли під час моделювання.
- Описати подальші кроки для покращення моделей або результатів.

Формат звіту:

- Документ у форматі PDF або Google Docs.
- Містити метрики моделей, таблиці з результатами та приклади графічних візуалізацій.
- Обсяг: 4-6 сторінок.

3.3. Підсумковий звіт: Аналіз та візуалізація даних

Метою підсумкового звіту є представлення комплексного аналізу та візуалізації даних з обґрунтованими висновками. Зміст звіту:

1. Вступ:

- Опис теми та цілей проєкту.
- Короткий огляд джерела даних і структури змінних.

2. Методи аналізу:

- Опис методів обробки і аналізу даних, які було застосовано.
- Обґрунтування вибору методів для аналізу та моделювання.
- Опис статистичних гіпотез та обраних методів перевірки їх валідності.

3. Результати аналізу:

- Висновки на основі експлоративного аналізу даних.
- Оцінка побудованих моделей та їх інтерпретація.
- Висновки за результатом перевірки гіпотез.

4. Візуалізації:

- Ключові графічні представлення результатів: тренди, кореляції, розподіли даних, результати моделювання.
- Оцінка інформативності візуалізацій для висновків.

5. Висновки:

- Основні інсайти, отримані під час аналізу.
- Рекомендації або висновки щодо отриманих результатів.

6. Командний внесок:

- Короткий опис внеску кожного члена команди у виконання проекту.

7. Формат звіту:

- Документ у форматі PDF або Google Docs.
- Обов'язкові візуалізації, таблиці з результатами аналізу та метриками.
- Обсяг: 8-12 сторінок.

Готуючи звіт, розбивайте інформацію на логічні блоки, використовуйте заголовки, підзаголовки, списки. Викладайте інформацію чітко і зрозуміло, уникайте зайвих деталей. Використовуйте графіки, діаграми, таблиці для наочного представлення отриманих результатів. Для дотримання вимог академічної доброчесності, посилайтесь на джерела інформації, які ви використовували. Дотримуйтесь вимог до оформлення, встановлених університетом.

3.4. Вимоги до підсумкового виступу та презентації результатів

1. Структура виступу:

Час для виступу команди — **10-15 хвилин**. Після презентації відводиться ще **5-7 хвилин** на запитання від аудиторії та викладача.

Структура виступу повинна включати наступні компоненти:

1. Вступ (1-2 хвилини):

- Представлення команди та теми проекту.
- Опис мети та основних завдань проекту.

- Короткий опис зібраних даних (джерело, обсяг даних, основні змінні).
- 2. Методи та інструменти (2-3 хвилини):**
- Огляд методів аналізу даних, які були використані (статистичний аналіз, моделювання, візуалізація тощо).
 - Перелік інструментів, використаних для виконання проєкту (Google Таблиці, Python, Excel, інші).
- 3. Основні результати аналізу (5-7 хвилин):**
- Представлення ключових результатів аналізу та моделювання.
 - Демонстрація трендів, кореляцій або висновків, отриманих в ході роботи з даними.
 - Оцінка точності моделей або аналізу (якщо були використані моделі).
 - Висновки на основі проведеного аналізу та рекомендації (якщо це доцільно).
- 4. Візуалізація даних (3-4 хвилини):**
- Пояснення побудованих графіків, діаграм, карт або інших візуалізацій.
 - Оцінка того, як візуалізація допомогла краще зрозуміти дані або результати аналізу.
- 5. Висновки та рефлексія (1-2 хвилини):**
- Короткий підсумок основних висновків.
 - Відображення командного процесу, труднощів, з якими зіткнулись, та навчання, яке отримали в процесі виконання проєкту.
- 6. Питання та відповіді (5-7 хвилин):**
- Готовність відповідати на запитання аудиторії та пояснювати деталі аналізу або обраних рішень.

Структура презентації:

1. Титульний слайд:

- Назва проєкту.
- Назви учасників команди.

2. Огляд проєкту:

- Мета проєкту, основні завдання та короткий опис даних.

3. Методи аналізу:

- Короткий опис методів та інструментів.

4. Основні результати:

- 2-3 слайди з результатами аналізу, підкріплені текстовими висновками та відповідними графіками.

5. Графічні візуалізації:

- 1-2 слайди з ключовими графіками або діаграмами, які найбільш чітко демонструють результати аналізу.

6. Висновки:

- Основні підсумки проєкту, що було виявлено або на що команда звернула особливу увагу.

7. Рефлексія та рекомендації:

- Короткий опис труднощів та висновків з процесу виконання проєкту, пропозиції щодо покращення (якщо є).

Загальні вимоги:

- Презентація повинна бути зрозумілою, лаконічною та візуально привабливою.
- Кількість слайдів: **7-12 слайдів**, щоб покрити основні моменти виступу, не перевантажуючи інформацією.
- Усі графіки та візуалізації повинні бути чіткими та інформативними, з правильно підписаними осями та заголовками.
- Мінімізувати кількість тексту на слайдах. Використовувати короткі марковані списки або схеми для пояснення основних моментів.

Підготовані звіти, презентація та супровідні матеріали (за потреби) мають бути прикріплені до відповідних завдань в середовищі Google Classroom з дотриманням визначених термінів.

4. ОЦІНЮВАННЯ ПРОЄКТУ

Оцінювання командного проєкту відбувається за наступними критеріями:

- Якість зібраних та підготовлених даних (20%).
- Коректність застосування методів аналізу (30%).
- Візуалізація та інтерпретація результатів (30%).
- Оформлення звіту та презентація проєкту (10%)
- Злагоженість командної роботи (10%).

Деталізуємо кожен із наведених критеріїв. Для кожного з них передбачено чотири рівні володіння: відмінний, добрий, задовільний та незадовільний.

1. Якість зібраних та підготовлених даних. Відповідно до цього критерію оцінюється чи були обрані джерела даних релевантними для досліджуваної проблеми. Також, оцінюється наскільки якісно команда впоралась із очищенням даних (видалення пропусків, корекція аномалій, усунення дублікатів), перетворенням даних (наприклад, кодування категоріальних змінних, нормалізація чи стандартизація числових змінних та підготовкою даних до

подальшої обробки. Цей критерій включає також оцінку даних на предмет відповідності гіпотезам або цілям дослідженням.

Детальний опис відповідності результатів складникам критерію наведено у таблиці 4.1.

Таблиця 4.1

Рівень	Опис
Відмінно	Дані зібрані з релевантних джерел, усі проблеми очищення даних вирішені. Пропуски та аномалії оброблені коректно. Набір даних чітко відповідає цілям проєкту.
Добре	Дані здебільшого відповідають цілям проєкту, виконано основні етапи очищення даних, але є незначні помилки або недоліки в обробці пропусків чи аномалій.
Задовільно	Дані частково відповідають цілям, але є проблеми з очищенням даних. Деякі пропуски або аномалії не оброблені належним чином.
Незадовільно	Дані не відповідають цілям проєкту, більшість проблем з даними не вирішені. Значні пропуски або помилки в підготовці даних.

2. Коректність застосування методів аналізу передбачає оцінювання відповідності обраних статистичних методів для вирішення поставленого завдання та обґрунтованості цього вибору.

У рамках перевірки критерію, звертається увага на формулювання статистичних гіпотез, відповідно до вимог предметної області, та коректність застосування статистичних тестів у частині виконання розрахунків із використанням відповідного програмного забезпечення. Важливим складником цього критерію є коректність інтерпретації отриманих результатів та відповідність поставленим гіпотезам.

Детальний опис відповідності результатів складникам критерію наведено у таблиці 4.2.

3. Візуалізація та інтерпретація результатів. У цьому критерії оцінюється чи представлені візуалізації (графіки, діаграми, таблиці) є чіткими, доречними та допомагають краще зрозуміти результати аналізу. Увага звертається на те, чи правильно інтерпретовані результати візуалізації та чи є логічний зв'язок між візуалізацією та висновками.

Таблиця 4.2

Рівень	Опис
Відмінно	Обрані та застосовані методи аналізу повністю відповідають поставленим гіпотезам та типу даних. Статистичні тести та моделі виконані правильно, результати коректно інтерпретовані.
Добре	Застосовані методи аналізу переважно правильні, але є невеликі помилки або спрощення в застосуванні методів чи інтерпретації результатів.
Задовільно	Частина методів була обрана або застосована некоректно. Можливі суттєві помилки в проведенні тестів або побудові моделей, що вплинуло на результати.
Незадовільно	Більшість методів аналізу була застосована неправильно, або взагалі не були використані адекватні методи для вирішення задачі.

Цей критерій включає також оцінювання чи надає команда достатньо контексту, щоб глядачі могли зрозуміти зміст графіків. Також, відзначається чи є візуалізації естетично привабливими, легко читабельними та з правильним використанням кольорів і елементів для підкреслення важливої інформації.

Детальний опис відповідності результатів складникам критерію наведено у таблиці 4.3.

Таблиця 4.3

Рівень	Опис
Відмінно	Візуалізації чіткі, доречні та допомагають інтерпретувати результати. Вибір типів графіків правильний, а їх дизайн привабливий та професійний. Усі результати добре пояснені.
Добре	Візуалізації здебільшого коректні та інформативні, але є кілька незначних помилок або проблем у дизайні. Пояснення результатів зрозумілі, але дещо поверхові.
Задовільно	Візуалізації частково допомагають зрозуміти результати, але деякі графіки або таблиці використані некоректно. Дизайн може бути не дуже зрозумілим. Є проблеми з поясненням результатів.
Незадовільно	Візуалізації не відповідають цілям дослідження або значно спрощують інформацію. Графіки некоректно відображають дані, або відсутні важливі пояснення.

4. Оформлення звіту та презентація проєкту. Відповідно до цього критерію оцінюється структурованість, послідовність та логічність звіту, зрозумілість та чіткість формату представлення інформації. Важливим складником є оцінювання відповідності оформлення звіту визначеним вимогам, дотримання норм ділового українського мовлення.

У частині оцінювання презентації, увага звертається на дотримання часових рамок виступу, наявність опису всіх ключових етапів проєкту, використання візуальних елементів, графіків, структур у зручному для сприйняття форматі.

Детальний опис відповідності результатів складникам критерію наведено у таблиці 4.4.

Таблиця 4.4

Рівень	Опис
Відмінно	Звіт чітко структурований, граматично правильний, оформлений відповідно до стандартів. Презентація повна, послідовна, графіки й результати чітко представлені, дотримано часу.
Добре	Звіт добре оформлений, але є незначні помилки або відсутні деякі деталі. Презентація здебільшого послідовна, але може містити пропущену інформацію або незначні помилки в оформленні.
Задовільно	Звіт має декілька значних помилок або не вистачає важливих деталей. Презентація неповна, відсутні ключові моменти або графіки.
Незадовільно	Звіт і презентація не структуровані, містять багато помилок, або відсутня велика частина важливої інформації.

5. Злагоженість командної роботи передбачає оцінювання розподілу ролей та обов'язків, комунікації та координації дій всередині команди, внеску кожного з учасників, відповідальності та своєчасності виконання завдань. Звертається увага також на здатність команди справлятися з внутрішніми конфліктами або різними думками в ході виконання завдань проєкту.

Детальний опис відповідності результатів складникам критерію наведено у таблиці 4.5.

Наведені критерії відображають чіткі орієнтири для студентів щодо того, які аспекти важливо враховувати для успішного виконання проєкту та отримання високих оцінок.

Таблиця 4.5

Рівень	Опис
Відмінно	Команда працювала злагоджено, чітко розподіливши ролі. Кожен учасник виконав свої обов'язки, проблеми вирішувалися спільно, постійна комунікація.
Добре	Загалом робота команди була злагодженою, але були незначні проблеми з комунікацією або розподілом ролей, які не сильно вплинули на результат.
Задовільно	Командна робота була частково організована, але відчувалася нестача координації, що вплинуло на процес або результат проєкту.
Незадовільно	Команда працювала незлагоджено, ролі були розподілені нерівномірно, або деякі учасники виконали менше обов'язків. Виникали постійні проблеми з комунікацією та координацією.

5. ПЛАН ВИКОНАННЯ КОМАНДНОГО ПРОЄКТУ

1. Формування команди та розподіл ролей

- Визначити склад команди (3-4 особи).
- Розподілити ролі в команді: лідер, аналітик даних, програміст, спеціаліст із візуалізації або інші.
- Узгодити способи спілкування та засоби для координації (чат, онлайн-дошка, зустрічі).

2. Вибір теми проєкту

- Обговорити можливі теми проєктів із використанням відкритих даних.
- Погодити тему з викладачем.
- Оцінити доступність даних та реалістичність виконання проєкту.

3. Збір та підготовка даних

- Знайти відкриті джерела даних (державні або міжнародні платформи, спеціалізовані ресурси).
- Завантажити та вивчити структуру завантажених наборів даних.
- Провести попередню обробку даних: очищення, усунення пропусків, нормалізація, перевірка наявності аномалій.
- Визначити необхідні перетворення даних (злиття таблиць, розділення даних у комірках, фільтрація, сортування, створення нових змінних).

4. Постановка завдання та вибір методів аналізу

- Сформулювати основні питання та гіпотези, які команда хоче дослідити.
- Вибрати відповідні методи аналізу даних (описова статистика, кореляційний аналіз, регресійні моделі, кластеризація, тощо).
- Розподілити задачі аналізу даних між учасниками команди відповідно до їхніх ролей.

5. Аналіз даних

- Виконати аналіз даних за допомогою обраних методів. Наприклад:
 - Оцінка кореляцій між змінними.
 - Побудова прогнозних моделей.
 - Кластеризація або сегментація даних.
- Проводити регулярні обговорення проміжних результатів у команді.
- Документувати отримані результати, проміжні й підсумкові результати, висновки, труднощі.

6. Візуалізація результатів

- Розробити та створити візуалізації (графіки, діаграми, таблиці) для демонстрації основних результатів дослідження.
- Оформити візуалізації так, щоб вони були інформативними та легко сприйнятними.

7. Оформлення проєкту

- Підготувати письмовий звіт про виконану роботу:
 - Вступ (опис проблеми та цілей проєкту).
 - Опис використаних даних і методів аналізу.
 - Основні результати та висновки.
 - Візуалізації та пояснення до них.
- Перевірити звіт на відповідність вимогам та коректність опису.

8. Підготовка презентації

- Розробити мультимедійну презентацію (слайди з візуалізаціями, діаграмами та ключовими висновками).
- Розподілити обов'язки під час презентації між учасниками команди.
- Підготувати короткий виклад основних результатів для доповіді.

9. Захист проєкту

- Провести презентацію перед викладачем і групою.
- Відповісти на запитання та аргументувати свої рішення щодо методів аналізу та висновків.
- Після презентації отримати зворотний зв'язок від викладача та одногрупників.

10. Оцінка та рефлексія

- Обговорити результати проєкту в команді: що вдалося, а що потребує покращення.
- Записати ключові висновки щодо командної роботи, застосування методів аналізу та презентації результатів для подальшого вдосконалення.

6. ПРИКЛАД ВИКОНАННЯ ПРОЄКТНОГО ЗАВДАННЯ З ВИКОРИСТАННЯМ GOOGLE ТАБЛИЦЬ

Тема проєкту: Перевірка гіпотези щодо різниці в середній успішності між студентами різної статі

Мета проєкту: Виконати статистичний аналіз, використовуючи інструменти Google Таблиць для обробки, аналізу та візуалізації даних.

Дані: Для аналізу використовуються дані з набору Student Performance Factors сервісу Kaggle.

Хід виконання роботи

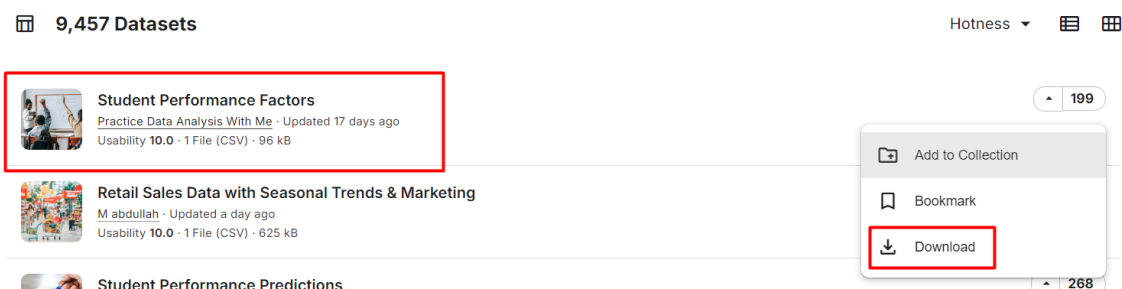
1. Завантаження даних з Kaggle

1.1. Зареєструйтеся в сервісі відкритих даних Kaggle. Це можна зробити через корпоративний акаунт Google.

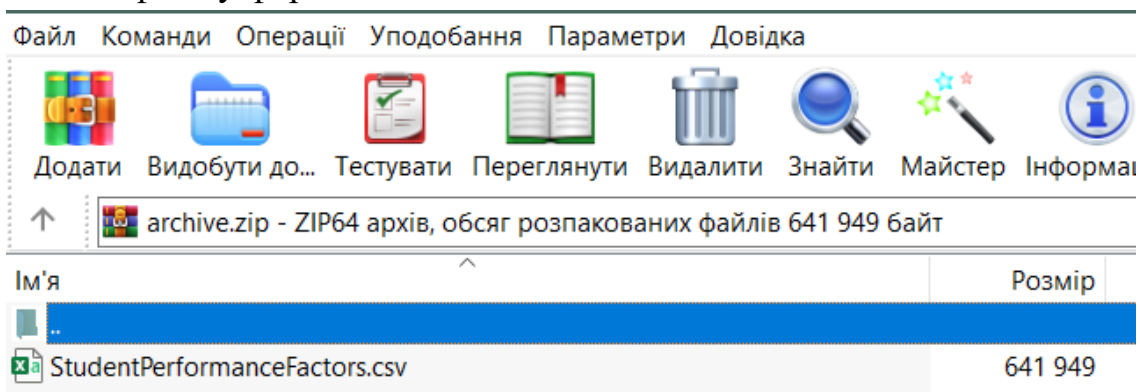
Після реєстрації обирається пункт Datasets та тип бази Education.

The screenshot shows the Kaggle website interface. On the left, there is a navigation menu with the following items: 'kaggle' (logo), 'Create', 'Home', 'Competitions', 'Datasets' (highlighted with a red box), 'Models', 'Code', 'Discussions', 'Learn', and 'More'. On the right, the main content area is titled 'Datasets' and includes a search bar, a description of the service, and buttons for '+ New Dataset' and 'Your Work'. Below this, there is another search bar for 'Search datasets' and a row of category filters: 'All datasets', 'Computer Science', 'Education' (highlighted with a red box), 'Classification', and 'Com'. At the bottom, there is a section for 'Trending Datasets'.

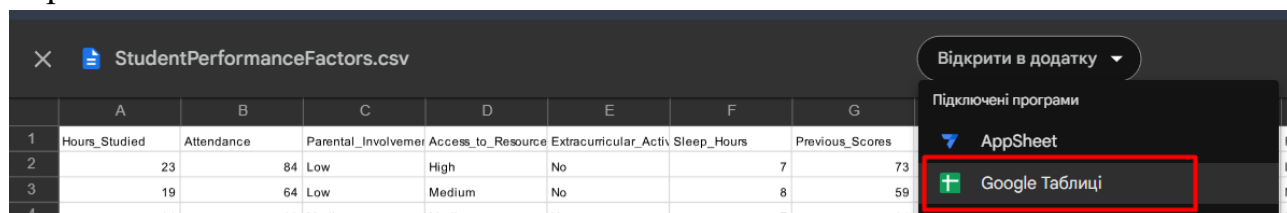
1.2. Обираєте набір даних Student Performance Factors



1.3. Обраний Dataset завантажується у вигляді архіву, що містить відповідний файл у форматі csv.



1.4. Файл з даними завантажується на Google Диск та відкривається за допомогою Google Таблиць. Отриманий файл використовується для подальшої обробки.



1.5. Копіюєте до нового аркушу стовпці з такими даними:

- **Стать** (Male, Female)
- **Успішність** (середній бал)

	A	B
1	Gender	Exam_Score
2	Male	67
3	Female	61
4	Male	74
5	Male	71
6	Female	70
7	Male	71
8	Male	67
9	Male	66
10	Male	69

2. Підготовка даних:

- Відфільтруйте дані за статтю, щоб отримати дві вибірки (для чоловіків і для жінок).
- Ви можете використовувати функцію FILTER, щоб створити окремі стовпці для успішності чоловіків та жінок.

D	E	F	G
Male	Exam	Female	Exam
Male	67	Female	61
Male	74	Female	70
Male	71	Female	64
Male	71	Female	65
Male	67	Female	67
Male	66	Female	69
Male	69	Female	66
Male	72	Female	63
Male	68	Female	65
Male	71	Female	66
Male	70	Female	63
Male	66	Female	61

3. Обчислення середнього та дисперсії:

Використовуйте функції AVERAGE для середнього та VAR.P для дисперсії кожної вибірки. Наведені в прикладі формули враховують кількість значень для кожної з категорій.

Приклад:

=AVERAGE(E2:E237) // Середня успішність чоловіків

=AVERAGE(G2:G164) // Середня успішність жінок

=VAR.P(E2:E237) // Дисперсія успішності чоловіків

=VAR.P(G2:G164) // Дисперсія успішності жінок

Середня успішність чоловіків	67,11440678
Середня успішність жінок	67,16101695
Дисперсія успішності чоловіків	12,06741956
Дисперсія успішності жінок	20,86390405
TTEST	0,9009672009

4. Т-тест для двох вибірок:

Для перевірки гіпотези щодо середньої успішності використовується функція TTEST для двох вибірок:

=TTEST(array1, array2, tails, type)

- array1 та array2 — це вибірки (успішність чоловіків та жінок).
- tails — кількість хвостів у тесті (1 для одностороннього тесту, 2 для двостороннього).
- type — тип тесту (1 — парний тест, 2 — для двох незалежних вибірок з рівними дисперсіями, 3 — для вибірок з нерівними дисперсіями).

Приклад для двостороннього тесту з нерівними дисперсіями:

=TTEST(E2:E237;G2:G164;2;3)

Результатом буде **p-значення**.

5. Інтерпретація результату:

Якщо отримане p-значення менше за обраний рівень значущості (наприклад, 0.05), то нульова гіпотеза відхиляється, і можна стверджувати, що є статистично значуща різниця між середньою успішністю чоловіків та жінок.

Якщо p-значення більше за 0.05, то нульова гіпотеза не відхиляється.

6. Звіт:

На основі результатів T-тесту студенти повинні зробити висновок:

- Прийняти або відхилити нульову гіпотезу.
- Описати, що означають результати для їхнього дослідження.

Приклад:

- **Н₀:** Немає різниці в середній успішності між чоловіками і жінками.
- **Результат:** p-значення = 0.03. Оскільки p-значення менше за 0.05, ми відхиляємо нульову гіпотезу і можемо стверджувати, що є різниця в середній успішності.

ДЖЕРЕЛА ТА РЕКОМЕНДОВАНА ЛІТЕРАТУРА

1. Положення про організацію освітнього процесу у Черкаському національному університеті імені Богдана Хмельницького. URL: https://drive.google.com/file/d/1CI95MVmJF3Hz8u3x_YIXzdUIrOEJRtWf/view
2. Василенко, О. А., & Сенча, І. В. (2011). Математично-статистичні методи аналізу в прикладних дослідженнях: навчальний посібник. Одеса: ОНАЗ ім. О.С. Попова.
3. Вишневецька, В. П. (2019). Основи математичної статистики з елементами хмаро орієнтованих технологій: Лабораторний практикум для майбутніх фахівців сфери фізичної культури та спорту. Київ: Видавництво НПУ імені М.П. Драгоманова.
4. Горват, А. А., Молнар, О. О., & Мінькович, В. В. (2019). Методи обробки експериментальних даних з використанням MS Excel: Навчальний посібник. Ужгород: Видавництво УжНУ "Говерла".
5. Грицюк, М. П., & Остапчук, О. П. (2008). Аналіз даних: Навчальний посібник. Рівне: НУВГП.
6. Дьячкова О.В. (2018). Комп'ютерний аналіз даних в MS Excel. Частина 1. Організація розрахунків і візуалізація даних. Харків: ХНУ ім. В.Н. Каразіна.
7. Жалдак, М. І., Кузьміна, Н. М., & Михалін, Г. О. (2020). Теорія ймовірностей і математична статистика: Підручник для студентів фізико-математичних та інформатичних спеціальностей педагогічних університетів. Видання четверте, доповнене. Київ: НПУ імені М.П. Драгоманова.
8. Кубай Д., Газін А., Горбаль А., Шульга Є., Шаповаленко Г. Відкритий посібник з відкритих даних. Український центр суспільних даних. Київ, 2016. Режим доступу: <http://socialdata.org.ua/manual/>.
9. Peck, R., Olsen, C., & Devore, J. L. (2015). Introduction to statistics and data analysis. Cengage Learning. URL: <https://www.spps.org/cms/lib/MN01910242/Centricity/Domain/859/Statistics%20Textbook.pdf>
10. Мозгульський, Є. З., & Бородай, Г. П. (2008). Методичні вказівки та завдання до розрахунково-графічної роботи з дисципліни "Теорія ймовірностей та математична статистика" для студентів економічних спеціальностей. Харків: Українська державна академія залізничного транспорту.
11. Огірко, О. І., & Галайко, Н. В. (2017). Теорія ймовірностей та математична статистика: навчальний посібник. Львів: ЛьвДУВС.
12. Оксанич, А. П., & Рилова, Н. В. (2018). Методичні вказівки щодо виконання лабораторних робіт з навчальної дисципліни «Обробка та аналіз даних» для

аспірантів зі спеціальності 122 – «Комп'ютерні науки» (третій освітньо-професійний рівень). Кременчук: КНУ імені М. Остроградського.

13. Руденко, В. М. (2012). Математична статистика. Навчальний посібник. Київ: Центр учбової літератури.

14. Сисоєва С.О., Кристопчук Т.Є. (2013). Методологія науково-педагогічних досліджень. Підручник. Рівне: Волинські обереги.

15. Швачич, Г. Г., Коноваленков, В. С., Соболенко, О. В., Заборова, Т. М., Христян, В. І., & Єгорцева, Є. Є. (2017). Навчальний посібник щодо вивчення дисципліни "Методи прикладного статистичного аналізу". Діпро: НМетАУ.

16. Відео-курс «Data Analysis Full Course Using Statistics». URL: <https://www.youtube.com/watch?v=Q-dOX4Y1fKE>

17. Bailey, B., Spence, D. J., & Sinn, R. (2013). Implementation of discovery projects in statistics. *Journal of Statistics Education*, 21(3).

18. Spence, D. J., Hix, S. L., Cooper, T. E., Sinn, R. (2013). Student Guide. Discovery Projects in Introductory Statistics. <https://faculty.ung.edu/djspence/NSF/PhaseIIProposalAbbreviated.pdf>

18. Spence, D. J., Hix, S.L. (2013). Instructor Guide Discovery Projects in Introductory Statistics.

https://faculty.ung.edu/DJSpence/NSF/materials/StatProjectInstructorGuide_Main.pdf

