



БАЗИ ДАНИХ

Робоча програма навчальної дисципліни (Силабус)

Реквізити навчальної дисципліни

Рівень вищої освіти	<i>Перший (бакалаврський)</i>
Галузь знань	<i>11 Математика і статистика</i>
Спеціальність	<i>113 Прикладна математика</i>
Освітня програма	<i>Наука про дані та математичне моделювання</i>
Статус дисципліни	<i>Нормативна</i>
Форма навчання	<i>очна(денна)</i>
Рік підготовки, семестр	<i>3 курс, весняний семестр</i>
Обсяг дисципліни	<i>3</i>
Семестровий контроль/ контрольні заходи	<i>Екзамен/МКР (1)</i>
Розклад занять	<i>Лекції – 1 раз на тиждень (18 лекцій) Лабораторні заняття – 1 раз на 2 тижні (9 занять)</i>
Мова викладання	<i>Українська</i>
Інформація про керівника курсу / викладачів	<i>Лектор: Шияк Богдан Антонович Лабораторні: Ковальчук-Химюк Людмила Олександрівна, Шияк Богдан Антонович</i>
Розміщення курсу	<i>Слак-канали #dbis-course та #dbis-practice</i>

Програма навчальної дисципліни

1. Опис навчальної дисципліни, її мета, предмет вивчення та результати навчання

Дисципліна «Бази даних» посідає одне з чільних місць в підготовці фахівців з прикладної математики й відноситься до професійно-орієнтованих дисциплін. Це визначається тим, що для широкого класу автоматизованих систем бази даних є системоутворюючим фактором.

Метою кредитного модуля є формування у студентів здатностей:

- розуміти основні сучасні інформаційні технології обробки даних;*
- розуміти місце і роль баз даних в інформаційних системах обробки інформації, методи проектування і управління базами даних;*
- розуміти основні моделі структур даних (списки, ієрархії, відношення, мережеві структури);*
- розуміти класифікацію систем управління базами даних (СУБД), як реляційних так і нереляційних (NoSQL);*
- знати та вміти застосовувати на практиці транзакції в реляційних СУБД;*
- розуміти обмеження та переваги нереляційних СУБД;*
- розуміти основні шаблони та підходи роботи з базами даних*
- працювати з різними видами СУБД з використанням мови програмування Python*

2. Пререквізити та постреквізити дисципліни (місце в структурно-логічній схемі навчання за відповідною освітньою програмою)

Дисципліна базується на знаннях і навичках, які студенти одержали при вивченні наступних дисциплін: "Програмування на мові Python", "Вступ до баз даних та інформаційних систем" та "Аналіз даних".

Дисципліна забезпечує засвоєння студентами наступних дисциплін: «Бази даних та інформаційні системи» навчального плану підготовки бакалаврів за спеціальністю 113 Прикладна математика.

3. Зміст навчальної дисципліни

Розділ 1. Робота з реляційними СУБД

1. Робота з транзакціями в реляційних СУБД
2. Міграція та рефакторинг баз даних
3. Моделювання реляційних баз даних. ERD
4. Багаторівнева архітектура. Шаблони роботи з даними
5. Об'єктно-реляційне відображення

Розділ 2. Основи розгортання інформаційних систем

1. Основи контейнеризації
2. Розгортання системи на Heroku

Розділ 3. Робота з нереляційними СУБД

1. Види нереляційних СУБД
2. CAP/PACELC теореми
3. Робота з MongoDB
4. Робота з Elasticsearch

4. Навчальні матеріали та ресурси

Базові джерела:

- Матеріали курсу, що публікуються викладачами у слак каналі курсу
- Принципи The Twelve-Factor App (<https://12factor.net/uk/>)
- Буй Д.Б., Сільвейструк Л.М. Формалізація моделі "Сутність-зв'язок" — Київ: Видавничо-поліграфічний центр "Київський університет", 2011. — 175 с. (<http://csc.knu.ua/uk/library/books/bui-silvestruk-33.pdf>)

Додаткові джерела:

- *Designing Data-Intensive Applications*, by Martin Kleppmann, O'Reilly Media, Inc., 2017 (<https://dataintensive.net/>)
- *Patterns of enterprise application architecture*, by Martin Fowler, with Dave Rice, Matthew Foemmel, Edward Hieatt, Robert Mee, and Randy Stafford, Addison-Wesley, 2003
- Документація по PostgreSQL (<https://www.postgresql.org/docs/current/index.html>)
- Документація по Flyway (<https://flywaydb.org/documentation/>)
- Документація по SQLAlchemy (<https://docs.sqlalchemy.org/en/14/orm/tutorial.html>)
- *Evolutionary Database Design*, by Pramod Sadalage, Martin Fowler (<https://martinfowler.com/articles/evodb.html>)

- 2014.05.27 Андрей Солнцев - Про агильную разработку базы данных (LiquiBase) (https://www.youtube.com/watch?v=A3XoEp_3V88)
- The SQLAlchemy Session - In Depth (<https://www.youtube.com/watch?v=uAtaKr5HOdA>)
- Consistency Tradeoffs in Modern Distributed Database System Design, by Daniel Abadi (<http://www.cs.umd.edu/~abadi/papers/abadi-pacelc.pdf>)

Навчальний контент

5. Методика опанування навчальної дисципліни (освітнього компонента)

Лекції

Розділ 1. Робота з реляційними СУБД

Лекція 1. Узагальнення SQL (SQL, DDL, DML). Індекси. Розрізи даних.

Лекція 2. Транзакції та їх властивості. Рівні ізоляції транзакцій. Поняття Deadlock.

Лекція 3. Міграція/рефакторинг БД. Принципи роботи Flyway

Лекція 4. Моделювання реляційних баз даних. Концептуальна, логічна та фізична ERD

Лекція 5. Багаторівнева архітектура. Принципи. Проблеми, які вона вирішує.

Лекція 6. Шаблони роботи з базами даних. Transaction Script. Domain model. Service Layer. Table Data Gateway. Row Data Gateway.

Лекція 7. OPM шаблони роботи з базами даних. Active Record. Session.

Лекція 8. Принципи роботи Flask та SQLAlchemy.

Розділ 2. Основи розгортання інформаційних систем

Лекція 9. Робота з Docker.

Лекція 10. Розгортання системи на Heroku.

Розділ 3. Робота з нереляційними СУБД

Лекція 11. Класифікація нереляційних СУБД. Ключ-значення СУБД. Документ-орієнтовані СУБД. Колоночні СУБД. Графові СУБД.

Лекція 12. CAP теорема. PACELC теорема. Компроміси при виборі технології бази даних.

Лекція 13. Принципи роботи документно орієнтованих СУБД на прикладі MongoDB.

Лекція 14. Робота з розподіленими базами даних на прикладі MongoDB.

Лекція 15. Повнотекстовий пошук. Реалізація в PostgreSQL. Реалізація в Elasticsearch.

Лекція 16. Робота з Elasticsearch, як з основною базою даних для системи.

Лекція 17. BigData бази даних. Принципи роботи Spark. Принципи роботи HBase.

Лабораторні роботи

Передбачено 4 лабораторні роботи.

Лабораторна робота 1.

Написати програму, що завантажує результати ЗНО з <https://zno.testportal.com.ua/opendata> за декілька років у таблицю (1 таблицю) в реляційній базі даних. Структуру таблиці (колонки та їх типи) студенти мають визначити на основі датасету.

Програма має поновлювати свою роботу у разі помилки (наприклад помилки в роботі програми, розриву мережевого з'єднання або помилки в роботі СУБД). Програма не має породжувати дублікати записів. Студенти мають продумати та продемонструвати сценарій "падіння" бази, та те як програма поновлює свою роботу.

Виконати запити, що повертають порівняльну статистику за кілька років (див. Варіанти завдань). Результат запиту має бути записаний у CSV-файл (засобами обраного стеку технологій). Але студент має бути готовим виконати запити з клієнта до БД.

Лабораторна робота 2.

Змінити структуру створеної у лабораторній роботі 1 бази, так щоб вона відповідала щонайменше 3 нормальній формі. Імена сутностей мають відображати об'єкти реального світу.

Міграції мають дозволяти, як створити базу з нуля, так і мігрувати існуючу базу (з лабораторній роботі 1).

При міграції бази дані не мають бути втрачені.

Лабораторна робота 3.

Студент самостійно обирає доменну область для моделювання.

Написати веб-додаток, що реалізує CRUD операції для обраних сутностей. В моделі мають бути принаймні зв'язки один-до-багатьох. Також додаток має містити можливість пошуку по декількох полях.

Система має бути розгорнута на Хероку.

Система має мати веб-інтерфейс (веб-сторінка). При розробці необхідно використати OPM для роботи з РСУБД.

Стек студент може обрати самостійно, але погодити з викладачем. Рекомендований стек: Bootstrap, Flask, SQLAlchemy, PostgreSQL.

Лабораторна робота 4.

Написати програму, що завантажує результати ЗНО з <https://zno.testportal.com.ua/opendata> за декілька років у таблицю (1 таблицю) в реляційній базі даних. Структуру документів студенти мають визначити на основі датасету.

Програма має поновлювати свою роботу у разі помилки (наприклад помилки в роботі програми, розриву мережевого з'єднання або помилки в роботі СУБД). Програма не має породжувати дублікати записів. Студенти мають продумати та продемонструвати сценарій "падіння" бази, та те як програма поновлює свою роботу.

Виконати запити, що повертають порівняльну статистику за кілька років (див. Варіанти завдань). Результат запиту має бути записаний у CSV-файл (засобами обраного стеку технологій). Але студент має бути готовим виконати запити з клієнта до БД.

6. Самостійна робота студента/аспіранта

До самостійного опрацювання виносяться:

- Вивчення синтаксису мови запитів MongoDB (<https://docs.mongodb.com/manual/reference/sql-comparison/>)
- Вивчення синтаксису мови запитів Elasticsearch (<https://www.elastic.co/guide/en/elasticsearch/reference/current/getting-started.html>)

Політика та контроль

7. Політика навчальної дисципліни (освітнього компонента)

- **Відвідування лекцій та лабораторних занять.** Відсутність на лекціях та практичних заняттях без поважної причини не допускається.
- **Правила поведінки на заняттях.** На лекційних та практичних заняттях студенти мають вимкнути мобільні телефони або увімкнути їх на беззвучний режим. Забороняється користуватися телефонами та іншими засобами зв'язку для пошуку інформації на гугл-диску викладача, в Інтернеті тощо.
- **Правила захисту лабораторних робіт.** Усі комп'ютерні практикуми оформляються і здаються студентами у вигляді звітів – doc або pdf-файлів. Захист практикумів відбувається у вигляді усної співбесіди зі студентом за результатами оформленого звіту, також студенту можуть бути поставлені запитання зі списку контрольних питань.
- **Правила призначення заохочувальних та штрафних балів.** За проходження онлайн курсів, погоджених з викладачем, можуть бути нараховані бонусні бали, але не більше ніж 10% від максимального балу за семестр (відповідно до РСО). Штрафні бали призначаються за несвоєчасне виконання лабораторних робіт (відповідно до РСО)
- **Політика щодо академічної доброчесності.** Згідно з Кодексом честі студента КПІ, при виконанні лабораторних чи контрольних робіт **забороняється** користуватися чужими виконаними лабораторними чи контрольними роботами та їх фрагментами. У випадку виявлення плагіату лабораторна робота може бути оцінена від 0 до 1/3 сумарної кількості запланованих за неї балів. Під час виконання контрольних робіт забороняється користуватись будь-якими сторонніми джерелами.

8. Види контролю та рейтингова система оцінювання результатів навчання (РСО)

Вказуються всі види контролю та бали за кожен елемент контролю, наприклад:

Поточний контроль: МКР, лабораторні роботи

Календарний контроль: провадиться двічі на семестр як моніторинг поточного стану виконання вимог силабусу.

Семестровий контроль: екзамен

Умови допуску до семестрового контролю: *зараховані щонайменше 3 лабораторних роботи та семестровий рейтинг рівний або більше 30 балів (з 50-и за роботу в семестрі).*

Таблиця відповідності рейтингових балів оцінкам за університетською шкалою:

Кількість балів	Оцінка
100-95	Відмінно
94-85	Дуже добре
84-75	Добре
74-65	Задовільно
64-60	Достатньо
Менше 60	Незадовільно
Не виконані умови допуску	Не допущено

9. Додаткова інформація з дисципліни (освітнього компонента)

- *Курсом передбачена можливість зарахування сертифікатів проходження дистанційних чи онлайн курсів за відповідною тематикою.*

Робочу програму навчальної дисципліни (силабус):

Складено асистент Шияк Б.А.

Ухвалено кафедрою *прикладної математики* (протокол № 13 від 16.06.2022)

Погоджено Методичною комісією факультету *прикладної математики* (протокол № 9 від 24.06.2022)