МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ОДЕСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ І.І. МЕЧНИКОВА
Кафедра математичного забезпечення комп'ютерних систем

"ЗАТВЕРДЖУЮ"
Проректор з науково-педагогічної роботи
_____
"____" _____ 20___ р.

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

ОК7 «Системи оперативної аналітичної обробки даних»
(назва навчальної дисципліни)

Рівень вищої освіти *Другий (магістерський)*

Галузь знань *12 – Інформаційні технології*

Спеціальність *126 – Інформаційні системи та технології*
(код і назва спеціальності)

Освітньо-професійна програма *Інформаційні системи та технології*
(назва ОПП/ОНП)

ОНУ – 2022

Робоча програма навчальної дисципліни «Системи оперативної аналітичної обробки даних». – Одеса: ОНУ, 2022. – 10 с.

Розробники:
Малахов Є.В., д-р техн. наук, професор, зав. кафедри МЗКС

Робоча програма затверджена на засіданні кафедри математичного забезпечення комп'ютерних систем

Протокол № 1 від. "25" 08 2022р.

Завідувач кафедри _____ (підпис) (Євгеній МАЛАХОВ) (Ім'я ПРІЗВИЩЕ)

Погоджено із гарантом ОПП «Інформаційні системи та технології»

_____ (підпис) (Євгеній МАЛАХОВ) (Ім'я ПРІЗВИЩЕ)

Схвалено навчально-методичною комісією (НМК) з ІТ спеціальностей факультету МФІТ

Протокол № 1 від. "31" 08 2022р.

Голова НМК _____ (підпис) (Алла РАЧИНСЬКА) (Ім'я ПРІЗВИЩЕ)

Переглянуто та затверджено на засіданні кафедри _____

Протокол № 1 від. "29" 08 2023р.

Завідувач кафедри _____ (підпис) (Є.В. Малахов) (Ім'я ПРІЗВИЩЕ)

Переглянуто та затверджено на засіданні кафедри _____

Протокол № ___ від. "____" _____ 20__ р.

Завідувач кафедри _____ (підпис) (_____) (Ім'я ПРІЗВИЩЕ)

1. Опис навчальної дисципліни

Найменування показників	Галузь знань, спеціальність, спеціалізація, рівень вищої освіти	Характеристика навчальної дисципліни	
		денна форма навчання	***заочна форма навчання***
Загальна кількість: кредитів – 3,5 годин – 105 змістових модулів – 5	Галузь знань 12 – Інформаційні технології Спеціальність 126 – Інформаційні системи та технології Рівень вищої освіти: *Другий (магістерський)*	*Обов'язкова*	
		Рік підготовки:	
		1-й	-й
		Семестр	
		2-й	-й
		Лекції	
		14 год.	6 год.
		Практичні, семінарські	
		год.	год.
		Лабораторні	
		16 год.	6 год.
		Самостійна робота	
		75 год.	93 год.
		у т.ч. ІНДЗ*: 30 год.	
		Форма підсумкового контролю: залік	

* – за наявності

2. Мета та завдання навчальної дисципліни

Метою є вивчення основних понять і принципів планування, проектування, побудови, заповнення і супроводу сховищ даних.

Завдання:

– ознайомлення з основами теорії інформацйних сховищ;
– вивчення засобів побудови багатовимірних структур даних;
– вивчення мови MDX;
– підготовка до виконання дипломних проектів та кваліфікаційних робіт, тематика яких пов'язана з розв'язаням аналітичних та прогнозних задач.

Процес вивчення дисципліни спрямований на формування елементів наступних **компетентностей**:

а) загальних (ЗК): —

в) спеціальних фахових (СК):

СК04. Здатність розробляти математичні, інформаційні та комп'ютерні моделі об'єктів і процесів інформатизації.

СК05. Здатність використовувати сучасні технології аналізу даних для оптимізації процесів в інформаційних системах.

СКМ04. Здатність застосовувати сучасні моделі та методи нечіткого логічного виводу на підставі форм представлення знань і способів організації пошуку рішень; проектувати та розробляти експертні систем.

СКМ07. Здатність проводити інформаційний аналіз та створювати багатовимірні моделі предметних областей.

Програмні результати навчання:

РН08. Розробляти моделі інформаційних процесів та систем різного класу, використовувати методи моделювання, формалізації, алгоритмізації та реалізації моделей з використанням сучасних комп'ютерних засобів.

РН09. Розробляти і використовувати сховища даних, здійснювати аналіз даних для підтримки прийняття рішень.

РНМ04. Використовувати сучасні нечіткі моделі, методи та засоби штучного інтелекту в системах прийняття рішень, застосовувати інтелектуальні алгоритми з використанням нечітких моделей для розв'язання задач штучного інтелекту.

РНМ08. Створювати оптимізовані пайплайни для підготовки даних до подальшого зберігання та обробки.

Очікувані результати навчання. У результаті вивчення навчальної дисципліни студент повинен

знати: *основні поняття і термінологію технології сховищ даних, технології та сфери застосування інструментів OLAP і методів видобутку даних (Data Mining), ролі сховищ даних у системах автоматизації прийняття рішень.*

вміти: *проектувати вітрини даних і сховища даних, застосовувати сучасні інструменти побудови сховищ даних і аналітичної обробки даних.*

3. Зміст навчальної дисципліни

Змістовний модуль 1. Поняття OLTP й OLAP систем.

Тема 1. Поняття OLTP й OLAP систем. Еволюція сховищ даних (DWH). Концепція DWH. Визначення DWH. Переваги технології DWH. Порівняння OLTP і DWH. Проблеми розробки і супроводу DWH.
Література: [1, 2, 10, 12].

Змістовний модуль 2. Архітектура DWH.

Тема 1. Архітектура DWH. Основні компоненти DWH та їхнє призначення. Інформаційні потоки в DWH.
Література: [1, 2, 5, 12].

Тема 2. Інструменти і технології DWH. СУБД для DWH і особливості її архітектури. Використання метаданных в DWH. Інструменти керування й адміністрування.
Література: [1, 5, 8, 9, 13].

Тема 3. Вітрини даних: загальна характеристика і зв'язок з DWH. Архітектура вітрин даних. Проблеми підтримки вітрин даних.
Література: [1, 5, 8, 9, 13].

Змістовний модуль 3. Проектування DWH.

Тема 1. Завдання проектування DWH. Моделювання вимірів. Схеми "зірка", "сніжинка", "зірка-сніжинка". Переваги цих схем. Порівняння моделей DM та ER.
Література: [1, 3, 8, 9].

Тема 2. Дев'ятикрокова методологія проектування БД для DWH. Суть, Опис і характеристика всіх етапів. Реалізація різних схем. Критерії оцінки вимірів DWH.
Література: [1, 5, 8, 9, 13].

Змістовний модуль 4. OLAP і розробка даних.

Тема 1. OLAP і розробка даних. Поняття OLAP системи, її призначення й функції. Еталонне тестування інструментів OLAP. Області застосування OLAP. Основні вимоги до OLAP-систем. Переваги OLAP-систем. Варіанти подання багатомірних даних. Основні аналітичні операції.
Література: [1, 2, 7, 12].

Тема 2. Інструменти OLAP. Правила Кодда для оцінки OLAP-систем.
Література: [1, 5, 7, 11, 12].

Тема 3. Розширення SQL для підтримки OLAP. Операції DECODE, CUBE, MOVINGAVG, MOVINGSUM, RANK...WHEN, RATIONTOREPORT, TERTILE, CREATE MACRO. Приклади.
Література: [1, 3, 6, 7].

Тема 4. Основні поняття технології розробки даних (data mining). Приклади додатків. Методи розробки даних. Інструменти розробки даних.
Література: [1, 7, 13].

Змістовний модуль 5. Створення DWH та обробка за допомогою MS Power BI або зведених таблиць MS Excel.

ІНДЗ *(самостійна робота)*

4. Структура навчальної дисципліни

Назви змістових модулів і тем	Кількість годин									
	Денна форма					Заочна форма				
	Усього	у тому числі				Усього	у тому числі			
		л	п	лаб	ср		л	п	лаб	ср
1	2	3	4	5	6	7	8	9	10	11
Змістовий модуль 1. Поняття OLTP й OLAP систем.										
Тема 1.	5	1			4	7	1		1	5
Змістовий модуль 2. Архітектура DWH.										
Тема 1.	5	1		4	3	8	1		1	8
Тема 2.	6	1			4	11				10
Тема 3.	7	1			4	11				10
Змістовий модуль 3. Проектування DWH.										
Тема 1.	9	2		6	4	12	2		2	10
Тема 2.	11	2			6	12				10
Змістовий модуль 4. OLAP і розробка даних.										
Тема 1.	7	2		6	4	11	2		2	10
Тема 2.	7	2			4	11				10
Тема 3.	9	1			6	11				10
Тема 4.	9	1			6	11				10
Всього годин без ІНДЗ	**75**	**14**		**16**	**45**		**6**		**6**	**93**
Змістовий модуль 5. Створення DWH та обробка за допомогою MS Power BI або зведених таблиць MS Excel.										
ІНДЗ*	30	–	–	–	30	–	–	–	–	–
Всього годин	**105**	**14**		**16**	**75**	**105**	**6**		**6**	**93**

* – за наявності

5. Теми семінарських занять

Семінарські заняття не передбачені

6. Теми практичних занять

Практичні заняття не передбачені

7. Теми лабораторних занять

№ з/п	Назва теми	Кількість годин
1	Проектування вітрини даних окремого ділового процесу в конкретній прикладній області	4
2	Проектування моделі сховища даних.	6
3	Реалізація проекту засобами MS SQL Analytical Services	6
	Разом	**16**

8. Самостійна робота

№ з/п	Назва теми / види завдань	Кількість годин
1	Поняття OLTP й OLAP систем. [1]	4
2	Архітектура DWH. Інформаційні потоки в DWH. [1]	4
3	Інструменти й технології DWH. [1]	4
4	Вітрини даних. [1]	3
5	Проектування DWH. [1]	10
6	OLAP і розробка даних. [1]	6
7	Інструменти OLAP. Правила Кодда для оцінки OLAP-систем. [1]	4
8	Мова MDX. [1]	10
	Разом (без ІНДЗ)	**45**

До самостійної роботи відноситься:

[1] – підготовка до лекцій, практичних, семінарських, лабораторних занять;

8.1. Індивідуальне навчально-дослідне завдання

Індивідуальне завдання полягає в проектуванні та створенні сховища даних для аналізу певної предметної області.

Студент повинен:

– дати опис предметної області;
– дати перелік аналітичних задач, які можуть чи повинні розв’язуватись над цією предметною областю;
– визначити структуру сховища даних та створити його за допомогою SSAS;
– реалізувати розв’язання зазначених аналітичних задач та виконати експорт результатів розв’язання до зведених таблиць MS Excel.

Примітка: студенти заочної форми ІНДЗ не виконують.

9. Методи навчання

Лекції з використанням мультимедійного презентаційного матеріалу.

10. Методи контролю

Під час **захисту індивідуального завдання** студент повинен:

– роз’яснити створену структуру сховища даних;
– продемонструвати та роз’яснити розв’язання зазначених аналітичних задач.

Під час **підсумкового контролю** студент повинен відповісти на 2 запитання екзаменатора з переліку, наведеному у п. 12.

10.1. Критерії оцінювання на підсумковому контролі:

1. Відповідь повинна бути повною і короткою. Вона не повинна мати в собі матеріал, що не відноситься до сутті питання.

2. Чітко формулювати твердження, вправно застосовувати необхідні формули і знання основних питань програми.
3. Відповіді, що мають помилкові твердження оцінюються виходячи з близькості відповіді до правильної.
4. Пропуски в обґрунтуванні тверджень враховуються і це призводить до зменшення кількості балів.
5. Малі недоліки, неточності при викладенні матеріалу, зменшують кількість балів.
6. Незнання і нерозуміння основної ідеї теоретичного питання або задачі призводить до зняття до 90 % балів.
7. Якщо відповідь на питання відсутня то виставляється нуль балів.

11. Питання для підсумкового контролю

1. Поняття та архітектура системи підтримки прийняття рішень.
2. Поняття: сховище даних, вітрина даних, операційні БД.
3. Поняття і модель даних OLAP. Засоби OLAP.
4. Способи аналітичної обробки даних.
5. Категорії даних у СД. Інформаційні потоки в СД.
6. Структура OLAP-кубу. Таблиця фактів. Таблиці вимірів. Проектування вимірів. Розгортання проекту. Зв’язки атрибутів. Створення користувацьких ієрархій. Зв’язки вимірів. Визначення обчислень.
7. Операції, виконувані над гіперкубом
8. Архітектура OLAP-систем. Шар добування, перетворення і завантаження даних. Шар зберігання даних. Шар аналізу даних.
9. Клієнтські OLAP-засоби. Серверні OLAP-засоби.
10. Технічні аспекти багатомірного зберігання даних: MOLAP, ROLAP, HOLAP.
11. Можливості служби SSAS.
12. Можливості Integration Services для роботи з OLAP. Злиття даних з різнорідних сховищ даних. Заповнення сховищ даних і вітрин даних. Очищення і стандартизація даних.
13. Архітектура служб SSIS. Пакет SSIS.
14. Введення в MDX. Область дії та основне призначення. Базові терміни мови. Опис вісей результуючого набору. Функції та вираження мови MDX.
15. MDX. Навігація по дереву вимірів.
16. MDX. Використання зовнішніх бібліотек.
17. Перетворення рядка до типів MDX.
18. MDX. Функції, що визначаються користувачем. Упорядкування на основі агрегатів. Обчислення ковзного середнього.
19. MDX. Операції над множинами. Іменовані множини.

12. Розподіл балів, які отримують студенти

Змістовий модуль №1	Змістовий модуль № 2			Змістовий модуль №3		Змістовий модуль № 4				Змістовий модуль № 5	
Т1	Т1	Т2	Т3	Т1	Т2	Т1	Т2	Т3	Т4	Т1	
8	8	8	10	8	10	8	8	10	8	14	100

Т1, Т2 ... – теми змістових модулів,
КР- курсова робота, ІНДЗ –індивідуальне навчально-дослідне завдання

Шкала оцінювання: національна та ECTS

Сума балів за всі види навчальної діяльності	Оцінка ECTS	Оцінка за національною шкалою	
		для екзамену, курсового проекту (роботи), практики	для заліку
90 – 100	A	відмінно	зараховано
85-89	B	добре	
75-84	C		
70-74	D	задовільно	
60-69	E		
35-59	FX	незадовільно з можливістю повторного складання	не зараховано з можливістю повторного складання
0-34	F	незадовільно з обов’язковим повторним вивченням дисципліни	не зараховано з обов’язковим повторним вивченням дисципліни

13. Навчально-методичне забезпечення

Конспект лекцій; комплекс навчально-методичного забезпечення дисципліни; нормативні документи; презентаційні матеріали.

14. Рекомендована література

14.1. Основна література

1. Reis J., Housley M., Fundamentals of Data Engineering. Plan and Build Robust Data Systems. – O'Reilly Media, 2022. – 446 p.
2. Inmon W.H., Krishnan K., Building the Unstructured Data Warehouse: Architecture, Analysis, and Design. – Technics Publications, 2010. – 216 p.
3. Codd E. F., Codd S. B., Salley C. T. Providing OLAP (On-Line Analytical Processing) to User-Analysts: An IT Mandate. - E. F. Codd & Associates, 1993.
4. Jukic N., Database Systems: Introduction to Databases and Data Warehouses. – Prospect Press, 2016. – 400 p.
5. Wrembel R., Koncilia C., Data Warehouses and OLAP: Concepts, Architectures and Solutions. – IGI Global, 2006. – 332 p.
6. Kimball R., Ross M., Becker B., Mundy J., Thornthwaite W., The Kimball Group Reader: Relentlessly Practical Tools for Data Warehousing and Business Intelligence Remastered Collection. 2nd Edition. – Wiley, 2015. – 912 p.
7. Rainardi V., Building a Data Warehouse: With Examples in SQL Server (Expert's Voice). – Apress, 2008. – 539 p.
8. Bhatia P., Data Mining and Data Warehousing: Principles and Practical Techniques. – Cambridge University Press, 2019. – 506 p.

9. Gorelik A., The Enterprise Big Data Lake: Delivering the Promise of Big Data and Data Science. – O'Reilly Media, 2019. – 224 p.
10. Corr L., Agile Data Warehouse Design: Collaborative Dimensional Modeling, from Whiteboard to Star Schema. – DecisionOne Press, 2011 – 328 p.
11. Crickard P., Data Engineering with Python. – Birmingham: Packt Publishing, 2020. – 357 p.

14.2. Допоміжна література

12. Спирли Э. Корпоративные хранилища данных. Планирование, разработка, реализация. Том 1. – М.: Издательский дом «Вильямс», 2001.
13. Ларсон Б. Разработка бизнес-аналитики в Microsoft SQL Server 2005 / СПб.: Питер, 2008 г. – 684 с.
14. Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. Методы и модели анализа данных: OLAP и Data Mining – СПб.: БХВ-Петербург, 2004. – 336 с.: ил.
15. Коннолли Т., Бегг К. Базы данных. Проектирование, реализация и сопровождение. Теория и практика. 3-е издание. – М.: Издательский дом «Вильямс», 2003.
16. Полубояров В.В. Использование MS SQL Server Analysis Services 2008 для построения хранилищ данных. [Электронный ресурс] / ННТУ, 2008. – Метод доступа: http://www.intuit.ru/department/database/mssqlsas2008/. – 488 с.
17. Некрасов В. 30 идей применения OLAP. [Электронный ресурс] / Intersoft Lab. – Метод доступа: http://www.iso.ru/rus/document5808.phtml.

15. Електронні інформаційні ресурси

1. Analysis Services | Microsoft SQL Server 2012 – Режим доступу: http://www.microsoft.com/en-us/sqlserver/solutions-technologies/business-intelligence/analysis.aspx
2. AdventureWorks Databases – 2019, 2016, 2012, 2008R2 and 2008 – Режим доступу: http://msftdbprodsamples.codeplex.com/releases/view/93587