


**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**  
**Львівський національний університет імені Івана Франка**  
**Механіко-математичний факультет**  
**Кафедра математичної статистики і диференціальних рівнянь**

**Затверджено**

На засіданні кафедри математичної  
статистики  
і диференціальних рівнянь  
механіко-математичного  
факультету  
Львівського національного  
університету  
імені Івана Франка  
(протокол №1 від 28.08.2020 р.)

В.о. завідувача кафедри:

 \_\_\_\_\_ Олег БУГРІЙ

**Силабус з навчальної дисципліни**

**“ Аналітичні сховища даних та хмарні технології ”,**  
**що викладається в межах ОПШ “ Статистичний аналіз даних ”**  
**першого (бакалаврського) рівня вищої освіти для здобувачів з**  
**спеціальності 112 - Статистика**

Львів 2020 р.

<b>Назва дисципліни</b>	Аналітичні сховища даних та хмарні технології
<b>Адреса викладання дисципліни</b>	Головний корпус ЛНУ ім. І. Франка м. Львів, вул. Університетська 1, 79000
<b>Факультет та кафедра, за якою закріплена дисципліна</b>	Механіко-математичний факультет Кафедра математичної статистики і диференціальних рівнянь
<b>Галузь знань, шифр та назва спеціальності</b>	11 - Математика та статистика 112 – Статистика
<b>Викладачі дисципліни</b>	Бугрій О.М., доктор фізико-математичних наук, доцент, в.о. завідувача кафедри математичної статистики і диференціальних рівнянь
<b>Контактна інформація викладачів</b>	<a href="mailto:oleh.buhrii@lnu.edu.ua">oleh.buhrii@lnu.edu.ua</a> , <a href="http://new.mmf.lnu.edu.ua/employee/buhrii_o_m">http://new.mmf.lnu.edu.ua/employee/buhrii_o_m</a> ;
<b>Консультації з питань навчання по дисципліні відбуваються</b>	Консультації в день проведення лекцій/практичних занять (за попередньою домовленістю). Головний корпус ЛНУ ім. І. Франка, каб. 267. м. Львів, вул. Університетська, 1
<b>Інформація про дисципліну</b>	Дисципліна “ Аналітичні сховища даних та хмарні технології ” є нормативною дисципліною з спеціальності 112 – Статистика для освітньої програми “Статистичний аналіз даних”, яка викладається в 8-му семестрі в обсязі 3-ох кредитів (за Європейською Кредитно-Трансферною Системою ECTS).
<b>Коротка анотація дисципліни</b>	Курс спрямований на формування у студентів професійних компетентностей, розвиток системи знань про роботу з аналітичними сховищами даних та хмарними технологіями, навичок та готовність до впровадження отриманих знань у реальних проектах на платформі Microsoft Azure.
<b>Мета та цілі дисципліни</b>	<i>Мета:</i> формування у студентів теоретичної та практичної бази знань щодо роботи з аналітичними сховищами даних та хмарними технологіями, вивчення базових засад хмарних середовищ та їх структури, набуття навичок представлення результатів аналізу
<b>Література для вивчення дисципліни</b>	1) Пасічник В.В., Шаховська Н.Б. Сховища даних: Навчальний посібник. – Львів: «Магнолія 2006», 2008. – 496 с. 2) W. H. Inmon, Building the Data Warehouse. Third Edition, 2002 3) Mark Beckner, Quick Start Guide to Azure Data Factory, Azure Data Lake Server, and Azure Data Warehouse, 2019 4) Matt How, The Modern Data Warehouse in Azure: Building with Speed and Agility on Microsoft’s Cloud Platform, 2020
<b>Обсяг курсу</b>	Загальний обсяг: 90 годин. Аудиторних занять: 52 год., з них 26 год. лекційних та 26 годин лабораторних робіт. Самостійна робота: 38 год.
<b>Очікувані результати навчання</b>	У результаті вивчення даного курсу студент буде: <b>знати:</b> різноманітні аспекти зберігання даних; <b>вміти:</b>

	<p>використовувати клаудні сервіси для роботи з даними</p> <p>В результаті засвоєння матеріалу даного курсу студент набуде таких <b>загальних (ЗК) і спеціальних (фахових) (СК) компетентностей:</b></p> <p>ЗК-2. Здатність застосовувати знання у практичних ситуаціях.</p> <p>ЗК-6. Навички використання інформаційних і комунікаційних технологій.</p> <p>ЗК-7. Здатність вчитися і оволодівати сучасними знаннями.</p> <p>ЗК-8. Здатність до пошуку, обробки та аналізу інформації з різних джерел.</p> <p>СК-2. Здатність застосовувати у професійній діяльності знання та навички в галузях теорії ймовірностей, математичної статистики, теорії випадкових процесів.</p> <p>СК-7. Здатність робити якісні висновки з кількісних даних.</p> <p>СК-8. Уміння працювати з інформаційними базами даних.</p> <p>СК-9. Здатність розробляти експериментальні та спостережувальні дослідження та аналізувати дані цих досліджень.</p> <p>СК-10. Здатність проводити дослідження ймовірно-статистичних моделей та інтерпретувати одержані результати.</p> <p>СК-11. Здатність використання обчислювальної техніки, спеціалізованих мов програмування та програмних засобів для розв'язання задач і здобуття додаткової інформації.</p> <p>СК-12. Здатність застосовувати ймовірно-статистичні методи в міждисциплінарному контексті.</p> <p>СК-13. Здатність подавати статистичні процедури та результати їхнього застосування у формі, придатній для цільової аудиторії, до якої звертаються, як усно, так і письмово.</p> <p>і здобуде такі <b>результати навчання (РН):</b></p> <p>РН-12. Вміти збирати та обробляти дані, застосовувати статистичні процедури для аналізу даних за допомогою обчислювальної техніки та програмних засобів.</p> <p>РН-14. Володіти сучасними інформаційними технологіями для створення презентацій, роботи з базами даних, пошуку інформації та обміну нею.</p>
<b>Ключові слова</b>	Аналітичні сховища даних, хмарні технології, Microsoft Azure, Apache Hadoop, Apache Spark, Azure SQL Database, Azure Synapse Analytics, Azure Data Lake
<b>Формат курсу</b>	Очний
<b>Теми</b>	Див. <b>Схема курсу</b>
<b>Підсумковий контроль, форма</b>	Іспит у кінці семестру
<b>Пререквізити</b>	Для вивчення даного курсу студенти потрібні базові знання з програмування
<b>Навчальні методи та техніки, які будуть використовуватися під час викладання курсу</b>	Інформаційні методи (лекція, бесіда, ілюстрація, демонстрація); дедуктивні методи на основі узагальнень; евристичні методи (проблемна лекція); інтерактивні методи (дискусія)
<b>Необхідне обладнання</b>	Мультимедійне обладнання, комп'ютерний клас, доступ до мережі Internet, програми та сервіси MS Teams, платформа Microsoft Azure

<p><b>Критерії оцінювання (окремо для кожного виду навчальної діяльності)</b></p>	<p>Оцінювання проводиться упродовж семестру за 100-бальною шкалою. Бали нараховуються за такими видами робіт з наступним співвідношенням:</p> <ul style="list-style-type: none"> <li>• лабораторні роботи: 32% семестрових балів, максимальна кількість балів – 32.</li> <li>• самостійна робота: 18% семестрових балів, максимальна кількість балів – 18.</li> <li>• Іспит: 50% семестрових балів, максимальна кількість балів – 50.</li> </ul> <p>Підсумкова максимальна кількість балів 100.</p> <p><b>Академічна доброчесність:</b> очікується, що роботи студентів будуть оригінальними дослідженнями чи міркуваннями. Списування та втручання в роботу інших студентів становлять, але не обмежують, приклади можливої академічної недоброчесності. Виявлення ознак академічної недоброчесності в написанні завдань є підставою для її незарахування викладачем, незалежно від масштабів плагіату чи обману. Жодні форми порушення академічної доброчесності не толеруються.</p> <p><b>Відвідування занять</b> є важливою складовою навчання. Очікується, що всі студенти відвідають усі лекції та лабораторні заняття курсу. Студенти повинні інформувати викладача про неможливість відвідати заняття. У будь-якому випадку студенти зобов'язані дотримуватися термінів визначених для виконання всіх видів робіт, передбачених курсом.</p> <p><b>Література.</b> Уся література, яку студенти не зможуть знайти самостійно, буде надана викладачем виключно в освітніх цілях без права її передачі третім особам. Студенти заохочуються до використання також й іншої літератури та джерел, яких немає серед рекомендованих.</p> <p><b>Політика виставлення балів.</b> Враховуються бали, набрані при поточному контролі та бали за виконання лабораторних робіт. При цьому обов'язково враховуються присутність на заняттях та активність студента під час практичного заняття; недопустимість пропусків та запізнь на заняття, користування мобільним телефоном, планшетом чи іншими мобільними пристроями під час заняття в цілях не пов'язаних з навчанням; списування та плагіат; несвоєчасне виконання поставленого завдання і т. ін.</p> <p><b>Оцінювання лабораторних робіт:</b> студенти виконують 12 лабораторних робіт. Оцінювання включає ставлення студента під час проведення лабораторної роботи в аудиторії та рівень компетенції при захисті звіту по виконаній роботі (0-2,5 бали за одну роботу). У підсумку, всі набрані бали множаться на коефіцієнт (1,067) для переведення у 32-ох бальну шкалу. У підсумку максимальна кількість балів студента за лабораторний практикум - 32.</p> <p>Бали оцінювання лабораторних робіт нараховуються за наступним співвідношенням:</p> <p><b>2.5</b> – студент в повному обсязі володіє навчальним матеріалом, має повне розуміння досліджуваної проблеми, надає правильні відповіді на запитання по темі, має свої ідейні міркування щодо реалізації даної проблеми ;</p> <p><b>2.0</b> – студент достатньо розуміє розглянутий матеріал, демонструє робочі варіанти схем, встановлює різні функціональні режими з достатнім обґрунтуванням (або з несуттєвими недоліками);</p> <p><b>1</b> – студент не достатньо розуміє приведені ним у звіті результати, вагається та надає неточні/неконкретні відповіді на запитання по темі,</p>
---	--

	<p>відсутні адекватні висновки;</p> <p><b>0,5</b> – студент погано розуміє наведені у звіті результати, у більшості надає помилкові відповіді на питання по роботі, не здатний вибрати робочі режими схем;</p> <p><b>0</b> - студент безвідповідально ставився до виконання роботи, представлений ним звіт не відповідає вимогам, студент виявляє нульовий рівень компетентності та зовсім не засвоїв розглянутий матеріал.</p> <p><b>Оцінювання самостійної роботи:</b> студенти самостійно завершують виконання лабораторних робіт і пишуть звіт до кожної лабораторної роботи. У підсумку максимальна кількість балів студента за самостійну роботу - 18.</p> <p>Бали оцінювання домашнього завершення виконання лабораторних робіт та написання звіту нараховуються за наступним співвідношенням:</p> <p><b>2</b> – звіт цілком і повністю відображає індивідуальне завдання студента, містить правильні висновки, ілюстрований (за потреби) відповідними графіками і таблицями, які правильно відображають суть виконаного завдання, студент має повне розуміння розглянутої теми, надає правильні відповіді на запитання по темі, код програми функціонує відповідно до завдання;</p> <p><b>1,5</b> – звіт в достатній мірі відображає індивідуальне завдання студента, містить допустимі висновки, ілюстрований (за потреби) відповідними графіками і таблицями, які частково відображають суть виконаного завдання, студент достатньо розуміє принципи написаного ним коду програми, присутні неточності та незначні помилки у відповідях на запитання по темі, код програми функціонує відповідно до завдання;</p> <p><b>1</b> – звіт містить загальні формулювання завдання, висновки нечіткі, необхідні ілюстрації чи таблиці відсутні, студент не досить добре розуміє розглянутий матеріал та представлений код програми, надає неточні/неконкретні відповіді на запитання по темі, код програми функціонує з помірними недоліками;</p> <p><b>0,5</b> – звіт не містить формулювання завдання, висновки необґрунтовані чи неповні, необхідні ілюстрації чи таблиці відсутні, студент погано розуміє розглянутий матеріал та представлений код програми, студент в більшості надає помилкові відповіді на питання по темі, код програми не функціонує належним чином;</p> <p><b>0</b> - звіт відсутній/не відповідає темі, студент зовсім не засвоїв розглянутий матеріал, написаний ним код програми не відповідає темі/не функціонує взагалі.</p> <p><b>Оцінювання іспиту (3 питання, 17 балів за кожне, але не більше 50 балів)</b> — за результатами написаних студентом письмових відповідей на 3 питання.</p> <p>Бали оцінювання кожного екзаменаційного завдання нараховуються за наступним співвідношенням:</p> <p><b>17-13</b> - розглянута тема відтворюється в повному обсязі, правильно, обґрунтовано, логічно, містить аналіз і систематизацію, аргументовані висновки. Засвідчено глибоке володіння матеріалом. Наведено приклади використання теоретичного матеріалу. Можуть бути присутні несуттєві описки та невідповідності;</p> <p><b>13-10</b> - відтворюється значна частина розглянутої теми. Виявлено знання і розуміння основних положень навчальної дисципліни, проте присутні неточності та/або невідповідності основній темі. Приклади використання</p>
--	--

	<p>теоретичного матеріалу відсутні;  10-7 - відстежується загальне розуміння розглянутої теми. Виявлені множинні неточності та невідповідності, пояснення наведених формул відсутні чи частково помилкові;  7-3 – студент погано розуміє розглянуту тему. Виявлені суттєві неточності та невідповідності. Наведені факти майже не відповідають темі;  3-0 – студент взагалі не розуміє розглянуту тему. Тему не розкрито, кількість викладеного матеріалу не відповідає загальним нормам. Наведені приклади не стосуються теми.</p> <p><b>Критерії оцінювання результатів неформальної освіти:</b>  Нарахування балів відбувається за публікацію студентом тез доповідей на конференціях, наукових статей, за участь студента у діяльності наукових гуртків, семінарів, круглих столів, конкурсів, участь у заходах неформальної освіти, за отримання сертифікатів про проходження навчання на різних освітніх платформах (Coursera, Prometheus тощо), курсах провідних ІТ компаній за тематикою навчальної дисципліни. Кількість балів визначається відсотком покриття результатів відповідної активності до вимог результатів навчання з навчальної дисципліни.</p>
<p><b>Питання для контрольного оцінювання</b></p>	<ol style="list-style-type: none"> <li>1. Аналітичні сховища даних.</li> <li>2. Проектування сховищ даних.</li> <li>3. Хмарні технології.</li> <li>4. Етапи роботи з аналітичним сховищем даних.</li> <li>5. Виконання процесу ETL (екстракція, трансформація, завантаження) в аналітичних сховищах даних.</li> <li>6. OLAP та які її основні характеристики.</li> <li>7. Технології Big Data в аналітичних сховищах даних.</li> <li>8. Apache Hadoop для обробки даних в аналітичних сховищах.</li> <li>9. Особливості використання Apache Spark.</li> <li>10. Використання Azure SQL Database для аналітики.</li> <li>11. Використання Azure HDInsight для обробки великих даних.</li> <li>12. Використання Azure Synapse Analytics.</li> <li>13. Переваги та недоліки використання стрімової обробки даних у хмарних технологіях.</li> <li>14. Стратегії забезпечення безпеки даних в аналітичних сховищах даних у хмарних технологіях.</li> <li>15. Інструменти, які використовуються для візуалізації даних у хмарних технологіях.</li> </ol>
<p><b>Опитування</b></p>	<p>Анкету-оцінку з метою оцінювання якості курсу буде надано по завершенню курсу.</p>

## Схема курсу

Тижні	Лекційний курс		Лабораторні заняття		К-сть год СР	Література
	Назва теми	К-сть год	Назва теми	К-сть год		
1	Основні поняття про аналітичні сховища даних та хмарні технології	2	Ознайомлення з Microsoft Azure	2	3	[1]-[4]
2	Архітектура аналітичних сховищ даних	2	Впровадження ETL процесів для імпорту даних в аналітичне сховище даних	2	3	[1]-[4]
3	Проектування та моделювання аналітичних сховищ даних	2	Створення бази даних для аналітичного сховища даних	2	3	[1]-[4]
4	Хмарні технології для аналізу даних	2	Налаштування хмарного середовища для аналізу даних	2	3	[1]-[4]
5	Робота з великими обсягами даних та Big Data	2	Використання хмарних сервісів для створення звітів та дашбордів	2	3	[1]-[4]
6	Впровадження Apache Hadoop та Apache Spark	2	Використання Apache Hadoop для обробки великих обсягів даних	2	3	[1]-[4]
7	Використання Azure SQL Database для аналітики	2	Використання Apache Spark для аналізу великих даних	2	3	[1]-[4]
8	Використання Azure HDInsight для обробки великих даних	2	Налаштування та робота з Azure SQL Database	2	3	[1]-[4]
9	Використання Azure Synapse Analytics	2	Робота з Azure Synapse Analytics	2	3	[1]-[4]
10	Стрімова обробка даних. Використання Azure Data Lake	2	Робота з Azure Data Lake Storage	2	3	[1]-[4]
11	Безпека та конфіденційність даних	2	Використання Azure Stream Analytics для обробки стрімових даних	2	3	[1]-[4]
12	Візуалізація та представлення результатів аналізу	2	Використання інструментів для візуалізацій та створення аналітичних звітів	2	3	[1]-[4]
13	Підсумкове заняття	2	Підсумкове заняття	2	2	[1]-[4]
	<b>Разом:</b>	<b>26</b>		<b>26</b>	<b>38</b>	