

Сергій Горобець,

кандидат педагогічних наук, доцент,
доцент кафедри прикладної математики та інформатики,
Житомирський державний університет імені Івана Франка,
ORCID ID 0000-0001-7639-9714
sgs_cat@yahoo.com

Ольга Горобець,

кандидат економічних наук, доцент,
доцент кафедри екологічної безпеки та економіки природокористування,
Житомирський національний агроекологічний університет,
ORCID ID 0000-0002-4526-0581
o.vikt.gorobets@gmail.com

ОСОБЛИВОСТІ ВИКЛАДАННЯ ДИСЦИПЛІНИ «БАЗИ ДАНИХ» ДЛЯ ЗДОБУВАЧІВ ВИЩОЇ ОСВІТИ ЗА СПЕЦІАЛЬНІСТЮ «СТАТИСТИКА»

У статті обґрунтовано, що в сучасному світі застосування баз даних має між-дисциплінарний характер, що висуває перед вищою освітою завдання підготовки фахівців, у т.ч., у сфері статистики, які володіють компетенціями щодо проєктування, розробки, налаштування, впровадження, експлуатації та супроводження баз даних. Метою статті визначено: поширення досвіду викладання дисципліни «Бази даних» для здобувачів вищої освіти за спеціальністю І12 «Статистика».

В статті розглянуто зміст дисципліни «Бази даних», що складається з двох тематичних модулів і викладається майбутнім статистикам в Житомирському державному університеті імені Івана Франка. Наведено можливості використання систем управління базами даних для здійснення статистичного аналізу даних. В першому модулі запропоновано розглянути підходи щодо аналізу предметної області, побудови схеми даних та її нормалізації. Другий модуль присвячено вивченню мови програмування SQL, яка надає можливість проводити статистичний аналіз великих масивів даних.

Для підготовки майбутніх фахівців у сфері статистики рекомендовано акцентувати увагу здобувачів вищої освіти на вивченні команд мови SQL, які дозволяють проводити вибірку, сортування, впорядкування та агрегацію даних.

З'ясовано, що збирання, обробка та аналіз значних обсягів інформації потребує від виконавця творчого, комплексного підходу та володіння основними методами розумової діяльності (аналізом, синтезом, порівнянням, абстрагуванням та узагальненням). Для успішного створення ефективних систем статистичної обробки інформації студенти повинні ознайомитись з теоретичними основами побудови баз даних та оволодіти практичними навичками роботи з програмними засобами, які дозволяють проєктувати, створювати та використовувати бази даних.

Ключові слова: бази даних, майбутні статистики, мова SQL, статистична обробка даних.

Сергей Горобец, Ольга Горобец. Особенности преподавания дисциплины «Базы данных» для соискателей высшего образования по специальности «Статистика»

В статье обосновано, что в современном мире применение баз данных имеет междисциплинарный характер, что ставит перед высшим образованием задачу подготовки специалистов, в т.ч. в сфере статистики, которые обладают компетенциями по проектированию, разработке, настройке, внедрению, эксплуатации и сопровождению баз данных. Целью статьи определено: распространение опыта преподавания дисциплины «Базы данных» для соискателей высшего образования по специальности 112 «Статистика».

В статье рассмотрены содержание дисциплины «Базы данных», состоящей из двух тематических модулей, которая преподаётся будущим статистикам в Житомирском государственном университете имени Ивана Франко. Приведены возможности использования систем управления базами данных для статистического анализа данных. В первом модуле предложено рассмотреть подходы к анализу предметной области, построению схемы данных и её нормализации. Второй модуль посвящён изучению языка программирования SQL, который позволяет проводить статистический анализ больших массивов данных.

Для подготовки будущих специалистов в сфере статистики рекомендуется акцентировать внимание соискателей высшего образования на изучении команд языка SQL, позволяющих проводить выборку, сортировку, упорядочение и агрегацию данных.

Выяснено, что сбор, обработка и анализ значительных объемов информации требует от исполнителя творческого, комплексного подхода и владение основными методами умственной деятельности (анализ, синтез, сравнение, абстрагирование и обобщение). Для успешного создания эффективных систем статистической обработки информации студенты должны ознакомиться с теоретическими основами построения баз данных и овладеть практическими навыками работы с программными средствами, которые позволяют проектировать, создавать и использовать базы данных.

Ключевые слова: *базы данных, будущие статистики, язык SQL, статистическая обработка данных.*

Serhiy Horobets, Olga Horobets. Peculiarities of teaching the discipline «Database» for students of higher education in the specialty «Statistics»

The article substantiates that in modern world the use of databases is interdisciplinary trend, which puts before higher education the task of training specialists, including in the field of statistics, who have competencies in the design, development, configuration, implementation, operation and maintenance of databases. The purpose of the article is to disseminate the experience of teaching the discipline «Databases» for applicants for higher education in the specialty 112 «Statistics».

The article considers the content of the discipline «Databases», which consists of two thematic modules and is taught to the prospective statisticians at Zhytomyr Ivan Franko State University. Possibilities of using database management systems for statistical data analysis are presented. In the first module it is offered to consider approaches to the analysis of the subject area, construction of the data scheme and its normalization. The second module is devoted to the study of the SQL programming language, which provides an opportunity to perform statistical analysis of large data sets.

To train prospective specialists in the field of statistics, it is recommended to focus the attention of higher education students on the study of SQL language commands, which allow sampling, sorting, organizing and aggregating data.

It was found that the collection, processing and analysis of significant amounts of information require a creative, comprehensive approach and mastery of basic methods of mental activity

(analysis, synthesis, comparison, abstraction and generalization). To successfully create effective systems of statistical information processing, students must become familiar with the theoretical foundations of building databases and master the practical skills of working with software that allows you to design, create and use databases.

Key words: *databases, prospective statistics, SQL language, statistical data processing.*

Постановка проблеми в загальному вигляді та її зв'язок із важливими науковими і практичними завданнями. Характерною ознакою сучасного світу є постійне збільшення обсягів інформації. Було підраховано, що у світі щодня з'являється близько $2,5 \cdot 10^{18}$ байт нової інформації. Фактично інформація стає тим чинником, що визначає ефективність будь-якої сфери діяльності.

При цьому постійно зростають вимоги до надійності та довгостроковості зберігання інформації, швидкості і зручності її обробки. Це призвело до того, що в наш час більшість операцій не може бути виконана вручну, що вимагає застосування сучасних комп'ютерних технологій, до яких можна віднести використання баз даних та систем управління ними. Бази даних широко використовуються в різноманітних сферах практичної людської діяльності, наприклад, для створення автоматизованих систем обліку чи документообігу, керування вмістом Інтернет-сайтів, розробки електронних каталогів, геоінформаційних систем, керування транспортними потоками чи складським господарством, здійснення банківської і страхової діяльності [1].

Отже, в сучасному світі застосування баз даних має міждисциплінарний характер, саме тому важливим завданням вищої освіти є підготовка фахівців, які володіють компетенціями щодо проектування, розробки, налаштування, впровадження, експлуатації та супроводження баз даних. Вивчення цих питань є досить актуальним для майбутніх фахівців у сфері статистики, оскільки в своїй професійній діяльності вони будуть мати справу з даними великого обсягу, виявляти в них певні закономірності та прогнозувати поведінку й тенденції розвитку економічних, соціальних та інших явищ.

При розробці навчальних матеріалів з дисципліни «Бази даних» для здобувачів вищої освіти спеціальності 112 «Статистика» виникає ряд проблем, зокрема: великий обсяг навчального матеріалу та обмежена кількість годин, що відводиться на його вивчення; міждисциплінарний характер баз даних та необхідність врахування специфіки майбутньої професійної діяльності здобувачів; несформована культура роботи з інформацією, недостатній практичний досвід використання інформаційних технологій та необхідність упорядкування великих обсягів даних.

Таким чином, вирішення зазначених проблем та підготовка майбутніх фахівців, які володіють інструментами інформаційно-комп'ютерних технологій для статистичного аналізу великих масивів даних, є досить актуальним.

Аналіз основних досліджень і публікацій із зазначеної проблеми. Досить багато уваги у вітчизняній і зарубіжній науковій літературі присвячено практичним аспектам організації баз даних: їх проектуванню, створенню, адмініструванню тощо. Ці питання досліджують А. Ю. Берко, О. Н. Верес, Г. А. Гайна, В. І. Гайдаржи, М. Грабер, М. К. Дж. Дейт, І. В. Изварін, Л. А. Павленко, В. В. Пасічник, В. А. Резніченко, В. Д. Руденко, О. В. Тарасов, Г. Хансен, Н. О. Харів, Д. Эймор, Н. Б. Шаховська та ін. Зокрема, багато уваги приділяється використанню мови SQL для роботи з сучасними системами керування базами даних (роботи Н. Р. Балик, М. Ю. Лосева, В. І. Мандзюка, О. В. Тарасова, В. В. Федька тощо).

Досить широко висвітлюються в літературі питання щодо управління бізнес-процесами, організації документообігу в економічних системах, а також аналізу даних, отри-

маних в процесі проведення досліджень соціально-економічних систем. Ці питання розкривають у своїх наукових працях Л. М. Малярець, В. С. Пономаренко, І. О. Золотарьова, С. В. Мінухін.

Методичні підходи щодо розробки системи навчання проектування баз даних розробляють Т. А. Бороненко, О. А. Козлов, А. Ю. Кравцова, А. А. Кузнецов, М. П. Лапчик, Н. В. Софронова, О. Г. Смолянинова, В. Г. Хоменко та ін. Проте питання щодо особливостей викладання дисципліни «Бази даних» під час підготовки майбутніх фахівців у сфері статистики у науковій літературі висвітлені недостатньо і потребують додаткового дослідження.

Метою статті є поширення досвіду викладання дисципліни «Бази даних» для здобувачів вищої освіти спеціальності 112 «Статистика».

Виклад основного матеріалу з обґрунтуванням отриманих наукових результатів.

Вивченню теоретичних основ побудови баз даних, розв'язанню практичних задач їх проектування, розвитку умінь використовувати мову SQL для доступу та управління даними, оволодінню навичками роботи із поширеними пакетами систем управління базами даних присвячена дисципліна «Бази даних», яка викладається в Житомирському державному університеті імені Івана Франка для студентів 2 курсу спеціальності 112 «Статистика». Згідно з навчальним планом дисципліна «Бази даних» включає в себе два основних модулі: «Основи проектування баз даних» та «Мова SQL та реляційні операції обробки даних».

Перший модуль містить теми, присвячені вивченню основних понять та визначень теорії баз даних, розгляду різних моделей даних та підходів щодо концептуального моделювання. На перших лекційних заняттях вивчаються основні та похідні операції реляційної алгебри, поняття та принципи побудови реляційних систем управління базами даних (СУБД). Студенти вчаться проектувати реляційні бази даних методом нормалізації відношень.

Опанування тем першого модулю формує у здобувачів вищої освіти уявлення про внутрішню організацію баз даних. Студенти набувають практичних навичок щодо побудови оптимальних логічних структур, призначених для збереження та обробки даних.

Після аналізу предметної області, виокремлення логічних сутностей, визначення функціональних залежностей між ними здобувачі вищої освіти повинні розробити та представити реляційну модель даних.

Отримана модель даних підлягає оптимізації шляхом чітко визначених дій, які отримали назву «нормалізація відношень». Шість кроків нормалізації дозволяють розробнику довести модель бази даних до найбільш раціонального вигляду, видалити повтори та надмірність даних. Нормалізація призводить до представлення даних у атомарному вигляді, що, своєю чергою, в подальшому дозволяє спростити обробку та аналіз даних, що зберігаються.

Особливу увагу під час проектування приділяють опису зв'язків між окремими логічними сутностями та групами даних. Правильно побудовані зв'язки дозволяють контролювати цілісність введеної інформації та зменшити обсяг даних, що зберігаються. Потрібно зазначити, що зв'язки будуються між так званими ключовими полями. Первинні ключі дозволяють однозначно ідентифікувати кожен запис бази даних, а вторинні (зовнішні) ключі дають можливість використовувати ці записи за умови мінімального збільшення об'єму даних, що зберігаються.

Опанування майбутніми фахівцями в галузі статистики правильного підходу до проектування структури бази даних дає можливість не тільки побудувати оптимальну логічну схему, а й отримати зручне представлення даних для подальшого їх зберігання, обробки та аналізу.

Отримані практичні навички допомагають здобувачам вищої освіти скласти чітке уявлення про найкращу структуру впорядкування та організації даних, сформувавши корисні навички щодо комплексного підходу в роботі з великими масивами інформації.

Другий модуль дисципліни «Бази даних» містить теми, які присвячені історії розвитку та загальній структурі мови SQL (Structured Query Language), питанням використання мови SQL для відображення інформації бази даних, опануванню прийомів доступу до даних засобами мови SQL, вивченню команд модифікації даних, розмежування прав користувачів та основ адміністрування систем управління базами даних.

Засвоєння тем другого модулю надає майбутнім статистикам можливість оперувати збереженими даними з метою отримання нової узагальненої інформації. Під час виконання практичних та лабораторних робіт, запланованих у другому модулі, студенти набувають навичок адміністрування та обробки даних за власними алгоритмами.

Потрібно зазначити, що знання та ефективне використання мови SQL представляє особливу цінність для проведення статистичних узагальнень даних. Ця мова належить до групи декларативних мов програмування, що дозволяє студентам писати програмні інструкції для SQL-сервера, не вдаючись у подробиці фізичної організації даних і не маючи уяви про дію внутрішніх алгоритмів та механізмів системи управління базою даних. Такий підхід значно пришвидшує процес навчання та опанування таким потужним інструментом, як мова SQL.

Так, при створенні запитів на вибірку даних, найбільш поширеною є команда SELECT. Використання різноманітних умов в комбінації з можливостями фільтрувати дані, групувати їх за певними ознаками, встановлювати діапазони значень, інтерактивно взаємодіяти з інструкцією запиту через введення параметрів – все це суттєво розширює можливості використання звичайної інструкції SELECT [2].

Певна частина потреб у статистичних обчисленнях може бути реалізована за допомогою окремих агрегатних функцій SQL та їх комбінацій. Так, окрім простих команд по визначенню найбільшого (MAX), найменшого (MIN) та середнього (AVG) значень у заданій множині можна визначати суму (SUM) та кількість елементів (COUNT) у групі з урахуванням повторів значень або без них (COUNT DISTINCT).

Паралельно з вищевказаними інструкціями можна застосовувати інструкцію групування (GROUP BY) з додатковими аргументами (column-expression) та інструкціями (WITH CUBE, WITH ROLLUP, GROUPING SETS, CUBE або ROLLUP).

Можливості вбудованих функцій дозволяють проводити різноманітні злиття (з'єднання) даних (FULL JOIN, LEFT JOIN, RIGHT JOIN, INNER JOIN), які зберігаються у різних логічних сутностях і, при потребі, навіть у різних базах даних. Отримані після вибірки та обробки дані можна впорядковувати за допомогою інструкцій виду ORDER BY за певними умовами.

Комбінація вищевказаних інструкцій суттєво розширює можливості використання мови SQL у питаннях статистичної обробки значних обсягів даних.

Окремо потрібно звернути увагу на можливість виконання арифметичних операцій в межах інструкції мовою SQL. За допомогою таких математичних обчислень можна обчислювати більшість статистичних показників, які не представлені окремими командами у мові SQL. Знаючи формулу для обчислення того чи іншого статистичного показника, можна спробувати використати запит, процедуру, тригер чи курсор (або їх комбінації) для програмування необхідного алгоритму пошуку значення цього показника.

Цікаво, що навіть відносно простий та широко поширений програмний продукт MS Excel має вбудовану можливість звертатись до зовнішніх баз даних за допомогою макросів

на мові Visual Basic for Applications, в середину яких можна вбудовувати інструкції мовою SQL. Така взаємодія відкриває широкі можливості використання вбудованих функцій MS Excel для статистичного аналізу даних спільно із зовнішніми серверами баз даних [3].

Висновки та перспективи подальшого дослідження. Отже, збирання, обробка та аналіз значних обсягів інформації потребує від виконавця творчого, комплексного підходу та володіння основними методами розумової діяльності – аналізом, синтезом, порівнянням, абстрагуванням та узагальненням. Для успішного створення ефективних систем статистичної обробки інформації студенти повинні ознайомитись з теоретичними основами побудови баз даних та оволодіти практичними навичками роботи з програмними засобами, які дозволяють проєктувати, створювати та використовувати бази даних. Для підготовки майбутніх фахівців у сфері статистики рекомендується приділити більше уваги вивченню мови програмування запитів SQL, оскільки вона містить команди, які дозволяють здійснювати обробку даних, в т.ч., визначати різні статистичні показники.

Перспективами подальших досліджень є вивчення можливостей адаптації курсу «Бази даних» до умов дистанційного навчання.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Хоменко В.Г. Моделювання змісту курсу «Принципи побудови і захист інформації баз даних». URL: <http://oaji.net/articles/2014/690-1396609558.pdf>
2. Документация по SQL: Преобразование «Статистическая обработка». URL: <https://docs.microsoft.com/ru-ru/sql/integration-services/data-flow/transformations/aggregate-transformation?view=sql-server-ver15>
3. SQL запрос из Excel VBA. URL: http://moonexcel.com.ua/sql-%D0%B7%D0%B0%D0%BF%D1%80%D0%BE%D1%81-%D0%B8%D0%B7-excel-vba_ru

REFERENCES

1. Khomenko V.H. Modeliuvannia zmistu kursu «Pryntsypy pobudovy i zakhyst informatsii baz danykh». URL: <http://oaji.net/articles/2014/690-1396609558.pdf> [in Ukrainian].
2. Dokumentatsiya po SQL: Preobrazovanye «Statysticheskaia obrabotka». URL: <https://docs.microsoft.com/ru-ru/sql/integration-services/data-flow/transformations/aggregate-transformation?view=sql-server-ver15> [in Russian].
3. SQL zapros yz Excel VBA. URL: http://moonexcel.com.ua/sql-%D0%B7%D0%B0%D0%BF%D1%80%D0%BE%D1%81-%D0%B8%D0%B7-excel-vba_ru [in Russian].