

МЕТОДИЧНІ ОСНОВИ ПОБУДОВИ АТРИБУТИВНИХ БАЗ ДАНИХ ГЕОГРАФІЧНИХ ІНФОРМАЦІЙНИХ СИСТЕМ (НА ПРИКЛАДІ ДЕМОГРАФІЧНИХ АТРИБУТИВ)

Сергєєва Н.П.

Львівський національний університет імені Івана Франка

Здійснено систематизацію демографічних показників для побудови атрибутивної реляційної бази даних географічної інформаційної системи (ГІС). Розроблено структуру бази даних ГІС «Демографічні показники Львівської області». Охарактеризовано можливості здійснення операцій селекції до побудованої бази даних. Наведено приклади реалізації атрибутивних запитів до розробленої бази даних ГІС у табличній і графічній формах.

Ключові слова: атрибутивні дані, база даних ГІС, географічна інформаційна система, операції селекції.

Постановка проблеми. Характерною особливістю сучасного розвитку людства є перехід до інформаційного суспільства, провідною ознакою якого є інформатизація, що сприяє пришвидшенню обробки великих обсягів інформації, підвищенню продуктивності праці, можливості моделювання та прогнозування соціально-економічних процесів, оптимізації процесів прийняття управлінських рішень тощо. Інформатизація у сфері географії представлена геоінформаційним картографуванням, створенням географічних інформаційних систем (ГІС) окремих регіонів, геоінформаційним аналізом. Найважливішою складовою частиною ГІС є інформаційне забезпечення. Воно представлено базами даних ГІС, які включають такі підблоки як атрибутивний і просторовий. Значення властивостей, параметри і характеристики просторових об'єктів у ГІС зберігаються в атрибутивних таблицях і називаються атрибутами. Атрибутивні таблиці формують атрибутивні бази даних, що є основою, на якій здійснюється подальше геоінформаційне моделювання та ГІС-аналіз. Тому дуже важливим завданням є розробка структури атрибутивних таблиць і визначення можливостей подальшого їх аналізу.

Аналіз останніх джерел і публікацій. У сучасній літературі з геоінформатики часто містяться визначення поняття баз даних ГІС, розглядається структура таких баз даних, аналізуються джерела атрибутивних даних ГІС. Серед публікацій із даної тематики особливо слід виділити праці О. Світличного, С. Плотницького [4], В. Волощина, П. Короля [1], Н. Кирильцевої, Ю. Ігнатова [2].

Виділення не вирішених раніше частин загальної проблеми. Хоча теоретичні основи ГІС і атрибутивних баз даних ГІС часто розглядаються у сучасній літературі, методика розробки структури, побудови та аналізу таких баз даних на конкретних прикладах не висвітлюється.

Мета статті. Головною метою даної статті є висвітлення методичних основ розробки структури тематичної бази даних ГІС «Демографічні показники Львівської області», а також можливостей подальшої роботи з нею.

Виклад основного матеріалу. Для управління атрибутивними даними та їх обробки використовують системи управління базами даних (СУБД), вбудовані в ГІС, оскільки логічно взаємопов'язані сукупності атрибутивних таблиць є, по суті, базами даних. Такі таблиці називаються реляційними (від слова «relation» – «відношення»), ГІС працюють саме з такими таблицями. Атрибутивні дані ГІС організовані у зрозумілому табличному вигляді з колонками (стовпцями) і рядками (записами) [4].

Джерелами атрибутивних даних для ГІС можуть бути матеріали моніторингових досліджень, дані спостережень і вимірювань, обліково-статистичні матеріали державної статистики, стандартні звітні форми різних державних, комерційних та громадських організацій, наукові звіти й публікації, текстові джерела, фотозображення, ілюстрації тощо [1]. Атрибутивну інформатизацію в суспільній географії розглянемо на прикладі демографічних атрибутів. Джерелами даних у цьому випадку можуть бути демографічні щорічники, бюлетені державної служби статистики, а також дані загальнодержавних переписів населення.

Під *структурою атрибутивної бази даних ГІС* розуміємо спосіб організації бази даних за записами і всередині кожного запису – за інформаційними полями. Нижче розглянемо структуру запису таблиці реляційної бази даних (РБД).

Сукупність використовуваних типів даних з усіма їхніми властивостями називається структурою атрибутивної таблиці шару ГІС [2]. Розробка структури бази даних є дуже важливим етапом її проектування. Перш за все, необхідно систематизувати усю сукупність демографічних показників. Для цього на першому етапі виділимо два основних блоки демографічних даних: дані про склад населення та дані про рух населення. Дані про склад населення представлені такими підблоками: дані про загальну кількість населення, про урботоральний склад населення, про статевий склад населення, про загальний віковий склад населення, про детальний віковий склад населення, та дані про працездатний склад населення. Дані про рух населення представлені наступними підблоками: дані про природний рух населення, про механічний рух населення, та про соціальний рух населення. Показники у межах кожного підблоку даних поділимо на абсолютні, відносні та середні, серед відносних величин виокремимо відносні величини динаміки, структури, координації, інтенсивності, порівняння та диференціації.

Дані про склад населення включають такі демографічні показники: а) *дані про загальну кількість населення:* абсолютні величини – загальна кількість населення у базовому періоді та загальна кількість населення у досліджуваному періоді; відносні величини: динаміки – темп росту населення, інтенсивності – загальна щільність населення, порівняння – відношення до кількості населення обласного центру; середні величини: середній вік населення; б) *дані про урботоральний склад населення:* абсолютні – кількість міського та кількість сільського населення; відносні величини: структури – частка міського та частка сільського населення, координа-

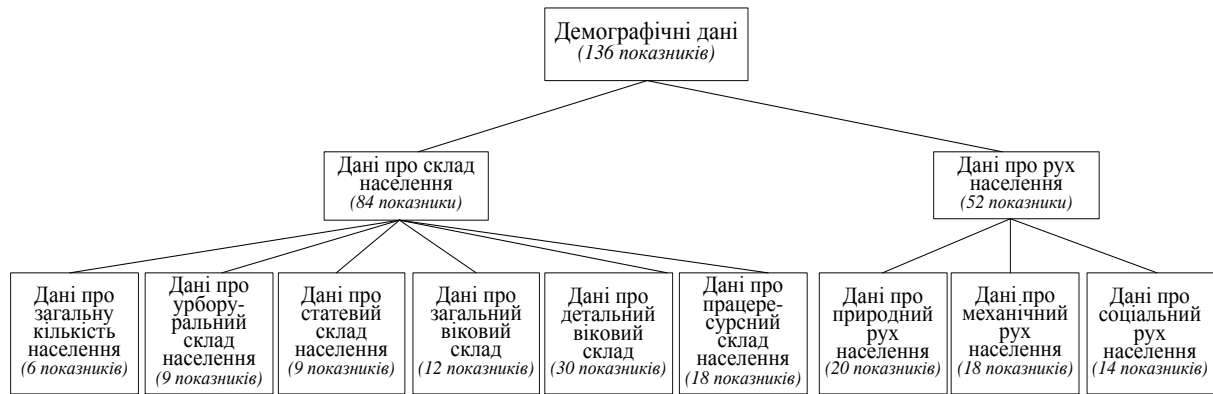


Рис. 1. Структура таблиці реляційної бази даних демографічних показників

ції – навантаження міського населення на сільське, інтенсивності – щільність міського та щільність сільського населення; середні: середній вік міського та середній вік сільського населення;

в) дані про статевий склад населення: абсолютні: кількість жінок і кількість чоловіків; відносні величини: структури – частка жінок і частка чоловіків, координації – навантаження чоловіків на жінок, порівняння – відношення до кількості жінок та відношення до кількості чоловіків обласного центру; середні величини: середній вік чоловіків і середній вік жінок; г) дані про загальний віковий склад населення: кількість осіб допрацездатного віку, кількість осіб працездатного віку та кількість осіб після працездатного віку. Відносні величини даного підблоку включають: величини структури – частки осіб до працездатного, працездатного та після працездатного віку; координації – навантаження осіб допрацездатного віку на осіб працездатного віку та навантаження осіб післяпрацездатного віку на осіб працездатного віку; порівняння – відношення до кількості осіб допрацездатного віку обласного центру, відношення до кількості осіб працездатного віку обласного центру та відношення осіб післяпрацездатного віку обласного центру; диференціації – коефіцієнт концентрації відносно щільності населення; г) дані про детальний віковий склад населення: включають такі абсолютні величини як кількість населення у віці 0–4, 5–9, 10–14, 15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–69, 60–64, 65–69, 70 і більше років. До відносних величин структури відносяться частки населення у віці 0–4, 5–9, 10–14, 15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–69, 60–64, 65–69, 70 і більше років; д) дані про працересурсний склад населення: кількість економічно активного населення, кількість зайнятих, кількість безробітних, кількість економічно неактивного населення та кількість пенсіонерів. Відносні величини представлені: величини структури – частка економічно активного населення, коефіцієнти зайнятості та безробіття, а також частки економічно неактивного населення та пенсіонерів; координації – навантаження безробітних на зайнятих, навантаження економічно неактивного населення на економічно активне населення; відносні величини порівняння – відношення до кількості економічно активного населення обласного центру, відношення до кількості зайнятих в обласному центрі, відношення до кількості безробітних в обласному центрі, відношення до кількості економічно неактивного населення обласного центру, відношення до кількості пенсіонерів в обласному центрі; диференціації – коефіцієнт концентрації відносно щільності населення.

Дані про рух населення включають такі показники: а) дані про природний рух населення: кількість осіб, які народилися за рік, кількість осіб, які померли за рік, природний приріст населення, кількість осіб, які народилися за рік у містах, кількість осіб, які померли за рік у містах, природний приріст населення у містах, кількість осіб, які народилися за рік у селах, кількість осіб, які померли за рік у селах, природний приріст населення у селах. Відносні величини представлені величинами інтенсивності та диференціації. До перших належать: загальні коефіцієнти народжуваності, смертності та природного приросту, коефіцієнти народжуваності, смертності та природного приросту в містах і коефіцієнти народжуваності, смертності та природного приросту в селах. До відносних величин диференціації належать коефіцієнти концентрації народжених і померлих відносно щільності населення; б) дані про механічний рух населення представлені абсолютними величинами та відносними величинами інтенсивності. Абсолютні величини: кількість осіб, які прибули за рік, кількість осіб, які вибули за рік, сальдо міграції, внутрішньорегіональна міграція (кількість прибулих, кількість вибулих, сальдо), міжрегіональна міграція (кількість прибулих, кількість вибулих, сальдо), міждержавна міграція (кількість прибулих, кількість вибулих, сальдо). Відносні величини: коефіцієнти прибуття, вибуття, міграційного приросту, внутрішнього регіонального прибуття, внутрішнього регіонального вибуття, внутрішньорегіонального міграційного приросту, міжрегіонального прибуття, міжрегіонального вибуття, міжрегіонального міграційного приросту, міждержавного прибуття, міждержавного вибуття та коефіцієнт міждержавного міграційного приросту; в) дані про соціальний рух населення. Абсолютні дані включають кількість шлюбів, кількість розлучень у містах, кількість шлюбів у містах і кількість розлучень у містах, кількість шлюбів у селах і кількість розлучень у селах. Відносні величини інтенсивності включають коефіцієнти шлюбності, розлучуваності, шлюбності у містах, розлучуваності у містах, шлюбності у селах, розлучуваності у селах; відносні величини диференціації – коефіцієнти концентрації шлюбів і розлучень відносно щільності населення.

Для того, щоб записати значення демографічних показників у таблицю реляційної бази даних, введемо систему імен назв інформаційних полів для демографічних даних. Перш за все, у системі демографічних показників виділимо такі рівні: демографічні дані – блок демографічних даних – підблок демографічних даних – демографічні по-

казники. Таким чином, назва кожного демографічного показника складатиметься з абрєвіатури DD (від demographic data) та трьох чисел, перше з яких вказує на приналежність показника до певного блока даних, друге – до певного підблока даних, третє – порядковий номер на найнижчому рівні – рівні демографічних показників. Приклади назв даних подано у Таблиці 1. На основі системи даних демографічних показників побудуємо атрибутивну реляційну базу даних «Демографічні показники Львівської області» у вигляді реляційної таблиці,

що містить 136 стовпців та 21 рядок, кожний комірка таблиці відповідає значенню демографічного показника у певному районі області та м. Львові. Фрагмент отриманої таблиці наведемо в Таблиці 1.

Оскільки суспільно-географічні дослідження часто ґрунтуються на проведенні аналізу великої кількості статистичного матеріалу, часто виникає необхідність вибору саме тої інформації, яка необхідна для конкретного дослідження. Такий вибір з однієї чи кількох баз геоданих здійснюється за допомогою операції селекції. Операція селекція – одна з найбільш поширених операцій реляційної алгебри, яка полягає у виборі з записів відношення R підмножини $r = Selection(R, Condition)$, записи якої задовольняють умову.

За допомогою операції селекції можна побудувати запит. Запити є одним із основних інструментів будь-якого ГІС-пакета. Користувач за допомогою запитів формулює вимоги до інформації, яку необхідно отримати із загального масиву доступних даних і подати в певному вигляді.

Запит формулюється за допомогою мови запитів SQL. SQL-запитом називається певний набір операторів SQL, за допомогою яких формується вибірка інформації з бази даних (у даному випадку – з атрибутивних таблиць ГІС).

Вибірка виконується за допомогою оператора SELECT, синтаксис якого наступний:

```
SELECT <список атрибутів таблиці
реляційної бази даних>
FROM <список таблиць>
WHERE <умова вибірки>
```

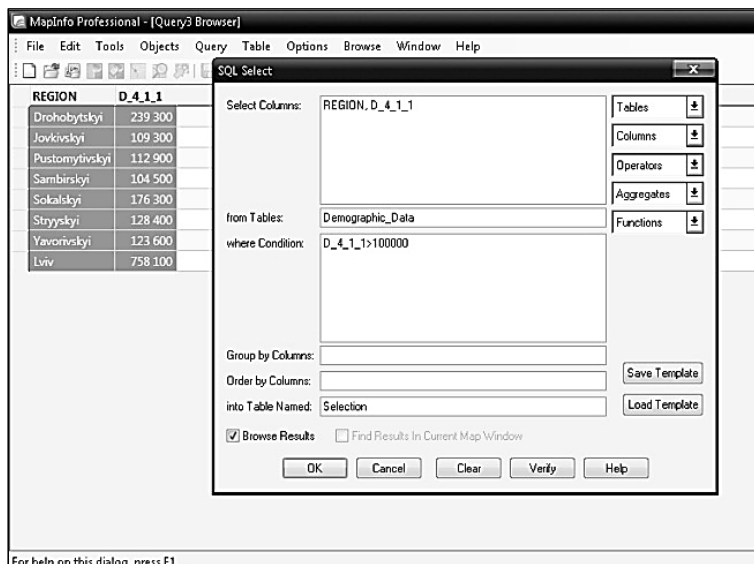


Рис. 2. Результат простого атрибутивного запиту, виконаного у середовищі ГІС-пакету «MapInfo Professional» та його відображення у табличній формі

Таблиця 1

Фрагмент таблиці реляційної бази даних демографічних показників Львівської області (побудована за даними [3])

№ п/п	Назва одиниці адміністративно-територіального устрою (REGION)	Загальна кількість населення у базовому періоді, тис. осіб (DD_4_1_1)	Загальна кількість населення у досліджуваному періоді, осіб (DD_4_1_2)	Темп росту населення, % (DD_4_1_3)	Загальна щільність населення, осіб/км ² (DD_4_1_4)
1.	Бродівський район	60,2	60,5	-5,9	51,8
2.	Буський район	46,4	46,5	0,3	54,6
3.	Городоцький район	69,3	69,2	-1,8	95,5
4.	Дрогобицький район*	239,3	229,1	-1,6	108,9
5.	Жидачівський район	71,1	71,5	-9,4	71,4
6.	Жовківський район	109,3	109,7	1,3	84,4
7.	Золочівський район	70,0	70,2	-2,2	63,8
8.	Кам'янка-Бузький район	57,4	57,4	-1,0	66,1
9.	Миколаївський район*	92,1	92,1	0,6	132,1
10.	Мостиський район	57,4	57,5	-0,2	67,9
11.	Перемишлянський район	40,1	40,2	-9,3	43,6
12.	Пустомитівський район	112,9	113,0	7,0	119,3
13.	Радехівський район	48,6	48,6	-4,4	42,5
14.	Самбірський район*	104,5	104,3	-3,3	146,5
15.	Сколівський район	47,4	47,5	2,6	32,2
16.	Сокальський район*	176,3	176,2	-2,1	110,8
17.	Старосамбірський район	78,2	78,2	1,5	62,8
18.	Стрийський район*	128,4	126,6	-2,3	155,6
19.	Турківський район	50,3	50,5	-1,0	42,2
20.	Яворівський район	123,6	123,5	4,4	80,1
21.	місто Львів	758,1	750,3	-0,9	4433,5

* Разом із містами обласного підпорядкування

Також оператор може містити й інші елементи: ORDER BY <список атрибутів>, GROUP BY <список атрибутів>, HAVING <умова>, UNION <вираз із оператором SELECT>.

Елемент WHERE дозволяє фільтрувати дані таблиці, а також задавати правила об'єднання даних у таблицях, якщо для побудови запиту використовується декілька таблиць.

Оскільки в ролі вхідних даних використовується атрибутивна таблиця, то до неї можливо здійснювати лише атрибутивні запити, і результатом їхнього виконання буде атрибутивна вибірка – також у вигляді таблиці.

Розглянемо приклади операцій запитів до атрибутивної реляційної бази даних та їх суспільно-географічний зміст на основі бази даних «Демографічні показники Львівської області» за допомогою ГИС MapInfo Professional». Для коректної роботи геоінформаційних систем із базою даних всі назви демографічних показників скорочено і подано англійською мовою.

Наведемо приклад простого SQL-запиту. Його суспільно-географічний зміст полягає в тому, що за його допомогою вибираються райони області, де загальна кількість населення в базовому періоді є більшою за 100 тис. осіб (рис. 1):

```
SELECT REGION, DD_4_1_1
FROM DEMOGRAPHIC_DATA
WHERE DD_4_1_1 > 100000,
```

де REGION – назви одиниць адміністративно-територіального поділу (атрибут таблиці РБД), DEMOGRAPHIC_DATA – назва таблиці демографічних показників,

DD_4_1_1 – загальна кількість населення у базовому періоді.

Список літератури:

1. Волошин В. У. Геоінформаційне картографічне картографування засобами ГИС MapInfo Professional / В. У. Волошин, П. П. Король. – Луцьк: ФОП Захарчук В. М., 2013. – 280 с.
2. Кирильцева Н. А. Атрибутивные данные ГИС и SQL запросы / Н. А. Кирильцева, Ю. М. Игнатов. – Кемерово, 2012. – 59 с.
3. Населення Львівської області. Демографічний щорічник / Головне управління статистики у Львівській області. Відділ статистики населення. – Львів, 2012. – 100 с.
4. Світличний О. О. Основи геоінформатики / О. О. Світличний, С. В. Плотницький. – Суми, 2006. – 295 с.

Сергеева Н.П.

Львовский национальный университет имени Ивана Франко

МЕТОДИЧЕСКИЕ ОСНОВЫ ПОСТРОЕНИЯ АТРИБУТИВНЫХ БАЗ ДАННЫХ ГЕОГРАФИЧЕСКИХ ИНФОРМАЦИОННЫХ СИСТЕМ (НА ПРИМЕРЕ ДЕМОГРАФИЧЕСКИХ АТРИБУТОВ)

Аннотация

Произведена систематизация демографических показателей для построения атрибутивной реляционной базы данных географической информационной системы (ГИС). Разработана структура базы данных ГИС «Демографические показатели Львовской области». Охарактеризованы возможности совершения операций селекции к построенной базе данных. Приведены примеры реализации атрибутивных запросов к разработанной базе данных ГИС в табличной и графической формах.

Ключевые слова: атрибутивные данные, база данных ГИС, географическая информационная система, операции селекции.

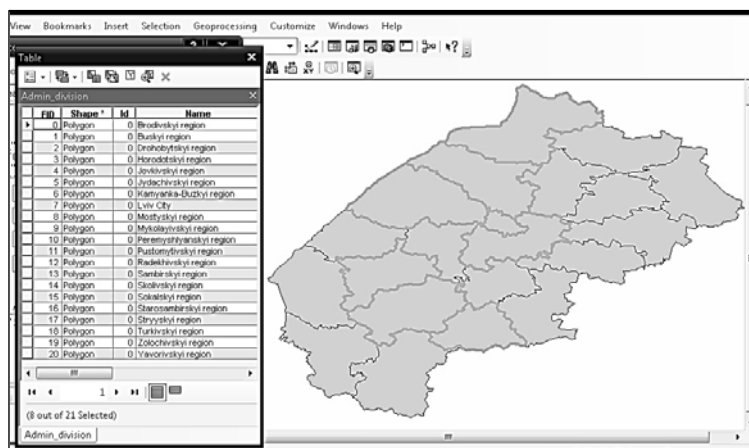


Рис. 3. Відображення результату простого атрибутивного запиту, виконаного в середовищі ГИС-пакету «ArcGIS», на картосхемі адміністративно-територіального поділу Львівської області

Якщо ж атрибутивна таблиця прив'язана до шару просторових об'єктів (наприклад до шару ареальних об'єктів, який відображає адміністративно-територіальний поділ області), то при виборі записів у таблиці будуть виділені й об'єкти в шарі, які відповідають вибраним записам (рис. 2).

Висновки. Проведене дослідження показує, що дуже важливим етапом суспільно-географічного дослідження є розробка структури таблиць реляційної бази даних ГИС. Структура та зміст визначаються цілями наукового дослідження. При вдалій розробці структури таблиці значно спрощується пошук необхідної інформації, а застосування операцій селекції дає змогу отримати необхідну змістовну суспільно-географічну інформацію в табличному або картографічному вигляді.

Sergieieva N.P.

Ivan Franko National University of Lviv

**THE METHODOICAL BASES OF THE GEOGRAPHIC INFORMATION SYSTEMS
ATTRIBUTIVE DATABASES CONSTRUCTION
(BASED ON THE DEMOGRAPHIC ATTRIBUTES)**

Summary

The demographic indicators for the geographic information system (GIS) attributive relational database are systematized. The structure of the GIS database «Demographic indicators of Lviv Oblast» is designed. The possibilities of realization of selection operations in relation to the constructed database are characterized. The examples of attributive queries realization regarding the designed GIS database in tabular and graphical forms are shown.

Keywords: attributive data, GIS database, geographic information system, selection operations.