

SQL NA MATURZE

Jarosław Biszczuk, Agnieszka Samulska
Ośrodek Edukacji Informatycznej i Zastosowań Komputerów
w Warszawie
02-026 Warszawa, ul Raszyńska 8/10
fjaroslaw.biszczuk,agnieszka.samulska}@oeizk.waw.pl

Abstract. During the workshop, we will show how to support students in their preparation for the high school final exams in computer science in the field of searching and processing information in a relational database using SQL. We will present RDBMS SQLiteStudio supporting the learning of database issues and solve several practice questions consisting in filtering and organizing data.

1. Wstęp

Jednym z pierwszych zastosowań komputerów obok zadań obliczeniowych było gromadzenie i wyszukiwanie danych. Sposób, w jaki te dane są organizowane i składowane determinuje m.in. wybór algorytmów wyszukiwujących.

Najpopularniejsze sposoby organizowania danych jakie można znaleźć obecnie w systemach komputerowych, to:

- model hierarchiczny – przykładem może być system plików z folderami i plikami właściwymi.
- model grafowy (sieciowy) – w przeciwieństwie do modelu hierarchicznego nie posiada węzła nadrzędnego i może posiadać wielu rodziców (choć to określenie nie jest najlepsze, gdy połączenia węzłów są dwukierunkowe). Przykładem może być przechowywanie użytkowników portalu społecznościowego z listą znajomych, listą obserwujących.
- model relacyjny – zaproponował go E. F. Codd na początku lat siedemdziesiątych. Pomysł i nazwa "relacyjny" zaczerpnięty został z matematycznego pojęcia relacji. Do operowania danymi w tym modelu jest wykorzystywany język zapytań SQL (Structured Query Language). Sam język doczekał się wielu oficjalnych specyfikacji w dokumentach standaryzacyjnych ISO i ANSI.

Dodatkowo każda implementacja systemu zarządzania bazą danych (SZBD) wprowadza rozszerzenia.

W dalszym ciągu skupimy się na krótkiej charakterystyce narzędzi dostępnych zdającemu maturę oraz opiszemy zalety używania SZBD SQLiteStudio w szkole.

2. Środowiska na maturze

2.1. MS Access

Składnia SQL zaimplementowana w MS Access z pewnymi wyjątkami odpowiada standardowi ANSI SQL-89¹. Jest to najpopularniejsze środowisko wybierane na maturze z informatyki. W podstawie programowej dla 3-letniego liceum ogólnokształcącego jest zapis dotyczący treści nauczania: *uczeń stosuje metody wyszukiwania i przetwarzania informacji w relacyjnej bazie danych (język SQL)*. MS Access umożliwia tworzenie zapytań bez konieczności znajomości składni SQL.

2.2. LibreOffice Base

Składnik Base pakietu LibreOffice pozwala na tworzenie baz danych w oparciu o silnik HSQLDB. Jest on tworzony w języku Java i w obecnej wersji LibreOffice 7.0 używany jest HSQLDB w wersji 1.8.0 z roku 2008. Na stronie hsqldb.org/web/openoffice.html można znaleźć informacje, jak używać nowego SZBD. Program Base pozwala także na łączenie się z innymi SZBD, takimi jak MySQL, PostgreSQL, pliki MS Access itp. Systemy te muszą jednak zostać zainstalowane i skonfigurowane, a pliki formatu MS Access utworzone.

2.3. MySQL i phpMyAdmin

SZBD MySQL jest popularnym od wielu lat rozwiązaniem typu klient-serwer. Serwer MySQL odbiera na ogół przy pomocy protokołu TCP/IP zapytania SQL i zwraca tą samą drogą wyniki. Zapytania SQL są wysyłane przez serwer www (np. Apache), na którym są uruchamiane skrypty php z narzędzia phpMyAdmin.

Rozwiązanie to jest rzadko używane na maturze. Najprostszym sposobem zapoznania się z tym rozwiązaniem w MS Windows, jest instalacja pakietu XAMPP.

2.4. Moduł sqlite3 w Python

W standardowej instalacji Python 3 znajduje się biblioteka z implementacją SZBD SQLite - sqlite3. Według twórców projektu jest to obecnie najczęściej używany system zarządzania bazą danych. Biblioteka SQLite często jest używana przez

¹ <https://docs.microsoft.com/en-us/office/client-developer/access/desktop-database-reference/comparison-of-microsoft-access-sql-and-ansi-sql>

aplikacje, które wymagają składowania danych w pliku. Znajdziemy ją w przeglądarkach Chrome, Firefox, Safari, systemach Android, iOS, iPhone, Windows 10.

Artykuł *Baza danych w Pythonie*² poświęcony jest wykorzystaniu biblioteki `sqli-ite3`. Zaletą rozwiązywania zadań z wykorzystaniem tej biblioteki jest to, że uczniowie muszą wykazać się znajomością składni SQL.

3. SQLite

Zanim zaczniemy z uczniami tworzyć bazy danych w Pythonie, warto zapoznać ich ze środowiskiem SQLite. SZBD SQLite posiada wiele nakładek, które pozwalają na bezpośrednie zakładanie bazy danych i wykonywanie zapytań w języku SQL. Z wartych odnotowania wymienimy:

- klient wiersza poleceń <https://sqlite.org/cli.html>,
- graficzny interfejs <https://sqlitestudio.pl>,
- interfejs [www https://sqliteonline.com](https://sqliteonline.com)

Z najbardziej przyjaznym i z polskojęzycznym interfejsem spotkamy się w programie SQLiteStudio. Jest to aktywnie rozwijany projekt na licencji GNU GPL. Program nie potrzebuje instalacji. Do uruchomienia w systemie Windows (a także Linux, MacOS X) wystarczy pobranie i rozpakowanie odpowiedniego archiwum (~ 30 MB).

3.1. Import danych

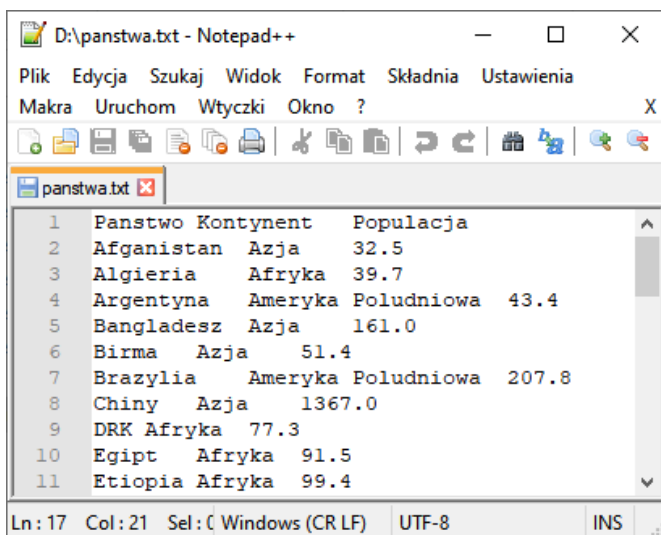
Proces tworzenia bazy danych zostanie przedstawiony na przykładzie fragmentu zadania 5. z egzaminu maturalnego z informatyki z czerwca 2020 roku, przeprowadzonego według nowej formuły (źródło: <https://cke.gov.pl>).

W pierwszym kroku należy stworzyć bazę danych oraz zaimportować dane z trzech plików tekstowych: `panstwa.txt`, `jezyki.txt` i `uzytkownicy.txt`. Przed wykonaniem tej czynności należy zapoznać się z konstrukcją danych. Z treści zadania wynika, że:

Pierwszy wiersz w każdym z plików jest wierszem nagłówkowym i zawiera nazwy pól. Dane w każdym wierszu oddzielone są znakami tabulacji.

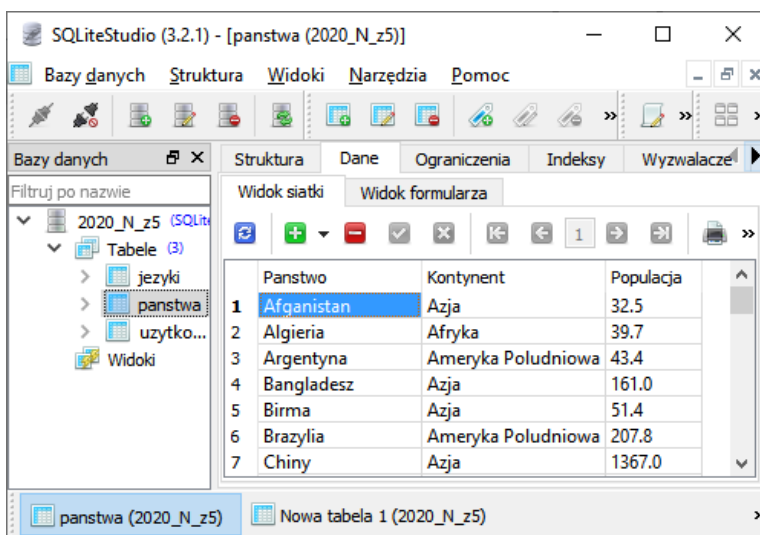
Istotne są również same dane, a w szczególności należy zwrócić uwagę na rodzaj separatora dziesiętnego, który należy zmienić z przecinka na kropkę. Należy to zrobić przed importem danych korzystając z np. edytora tekstu.

² https://www.oeiizk.waw.pl/wp-content/uploads/pliki/cyfrowa-szkola/Cyfrowa_Szkola_nr_3_na_WWW.pdf



Rysunek 1 Dane z separatorem dziesiętnym w postaci kropki

Mając przygotowane odpowiednio dane można przystąpić do realizacji. Po zapisaniu pustej bazy danych tworzone są trzy tabele o nazwach zgodnych z nazwami plików. Podczas importu danych do tabel należy zwrócić uwagę na sposób kodowania znaków, rodzaj separatora pól oraz czy pierwsza linia reprezentuje nazwy pól. Ostatnim krokiem będzie określenie typów pól.



Rysunek 2 Tabela z wczytanymi danymi

3.2. Wyszukiwanie i przetwarzanie informacji

Do wyszukiwania i przetwarzania informacji wykorzystujemy składnię języka SQL. Służą do tego celu edytor SQL. W zadaniu:

Utwórz zestawienie, które dla każdej rodziny językowej podaje, ile języków do niej należy. Posortuj zestawienie nierosnąco według liczby języków.
będzie to:

```
1. SELECT Jezyki.Rodzina,  
2.     Count(Jezyki.Jezyk) AS PoliczOfJezyk  
3. FROM Jezyki  
4. GROUP BY Jezyki.Rodzina  
5. ORDER BY Count(Jezyki.Jezyk) DESC;
```

Edytor SQL w programie SQLiteStudio ma kilka przydatnych funkcji, do których należy m.in. formatowanie kodu.

4. Podsumowanie

Podstawą sukcesu w rozwiązywaniu zadań bazodanowych jest umiejętność posługiwania się poleceniami języka SQL oraz znajomość zagadnień związanych z projektowaniem relacyjnych baz danych. Naukę warto rozpocząć od bardzo prostych przykładów, stopniowo zwiększając skalę trudności. Z roku na rok zadania maturalne dotyczące wyszukiwania i przetwarzania informacji są coraz trudniejsze, a znajomość języka SQL badana jest także w części teoretycznej. Dlatego warto do nauki w klasach z rozszerzoną informatyką wykorzystać SZBD SQLite ze względu na prostotę tego systemu. Niestety, korzystanie z tego narzędzia nie jest możliwe podczas egzaminu maturalnego – tam mamy do dyspozycji jedynie środowisko Python (z dostępnym w standardzie modulem sqlite3).

Literatura

1. Dokumentacja języka, <https://www.sqlite.org/>, ostatni dostęp 01.09.2020 roku.
2. Środowisko online, <https://sqliteonline.com/>, ostatni dostęp 01.09.2020 roku.
3. Serwis Centralnej Komisji Egzaminacyjnej, <http://cke.edu.pl>, ostatni dostęp 01.09.2020 roku.
4. Nield T., *Pierwsze kroki z SQL. Praktyczne podejście dla początkujących*, Helion, Gliwice 2016
5. Tutorial, <https://www.sqlitetutorial.net/>, ostatni dostęp 01.09.2020 roku.

6. Samulska A., *Baza danych w Pythonie*, OEliZK, kwartalnik edukacyjny W cyfrowej szkole Nr 1/2019.
7. Samulska A., *Matura z Pythonem*, Informatyka w Edukacji, Toruń 2019.