

Spis treści

O autorce	9
O korektorze merytorycznym	10
Podziękowania	11
Wprowadzenie	13
Rozdział 1. Źródła danych	17
Źródła danych	17
Narzędzia do łączenia się ze źródłami danych i edycji zapytań SQL	18
Relacyjne bazy danych	19
Hurtownie danych	23
Pytania dotyczące źródła danych	25
Wprowadzenie do bazy danych Farmer's Market	27
Uwaga dotycząca terminologii stosowanej w uczeniu maszynowym	28
Ćwiczenia	29
Rozdział 2. Instrukcja SELECT	30
Instrukcja SELECT	30
Podstawowa składnia zapytania SELECT	30
Wybieranie kolumn i ograniczanie liczby zwracanych wierszy	31
Sortowanie wyników za pomocą klauzuli ORDER BY	33
Wprowadzenie do prostych obliczeń inline	35
Więcej przykładów obliczeń typu inline — zaokrąglanie	37
Więcej przykładów obliczeń inline — konkatencja łańcuchów znaków	39
Ocena wyniku zapytania	41
Podsumowanie instrukcji SELECT	44
Ćwiczenia	45

Rozdział 3. Klauzula WHERE	46
Klauzula WHERE	46
Filtrowanie wyników z zapytania SELECT	46
Filtrowanie według wielu warunków	49
Wielokolumnowe filtrowanie warunkowe	54
Więcej sposobów filtrowania	55
BETWEEN	55
IN	56
LIKE	57
IS NULL	58
Ostrzeżenie na temat porównań z wartościami NULL	58
Filtrowanie za pomocą podzapytań	60
Ćwiczenia	61
Rozdział 4. Instrukcja CASE	62
Składnia instrukcji CASE	62
Tworzenie flag binarnych za pomocą CASE	65
Grupowanie wartości ciągłych za pomocą CASE	66
Kodowanie wartości kategoryjnych za pomocą CASE	69
Podsumowanie instrukcji CASE	70
Ćwiczenia	71
Rozdział 5. Złączenia w SQL-u	73
Relacje w bazie danych i złączenia	73
Pułapka często pojawiająca się podczas filtrowania połączonych danych	82
Złączenie więcej niż dwóch tabel	85
Ćwiczenia	87
Rozdział 6. Agregacja wyników na potrzeby analizy	89
Składnia GROUP BY	89
Statystyki podsumowujące grupy	90
Obliczenia wewnątrz funkcji agregujących	93
Minimum i maksimum	98
COUNT i COUNT DISTINCT	99
Średnia	101
Filtrowanie za pomocą HAVING	102
Instrukcje CASE wewnątrz funkcji agregujących	103
Ćwiczenia	106
Rozdział 7. Funkcje okienkowe i podzapytania	107
ROW_NUMBER	108
RANK i DENSE_RANK	111
NTILE	112
Agregujące funkcje okienkowe	113
LAG i LEAD	118
Ćwiczenia	121

Rozdział 8. Funkcje związane z datą i czasem	123
Ustawianie wartości w polu typu datetime	124
EXTRACT i DATE_PART	125
DATE_ADD i DATE_SUB	126
DATEDIFF	128
TIMESTAMPDIFF	128
Funkcje do obsługi dat w obliczeniach na zagregowanych danych i funkcjach okienkowych	129
Ćwiczenia	135
Rozdział 9. Eksploracyjna analiza danych w języku SQL	137
Eksploracyjna analiza danych z użyciem języka SQL	138
Eksploracja tabeli product	138
Eksploracja potencjalnych wartości w kolumnach	141
Badanie zmian w czasie	143
Eksploracja wielu tabel naraz	145
Stan a sprzedaż	148
Ćwiczenia	152
Rozdział 10. Tworzenie zbiorów danych na potrzeby raportów analitycznych	153
Wymagania stawiane zbiorom danych do analizy	154
Korzystanie z własnych analitycznych zbiorów danych — wspólne wyrażenia tablicowe i widoki	159
Wykorzystanie SQL-a do tworzenia bardziej zaawansowanych raportów	163
Ćwiczenia	167
Rozdział 11. Bardziej zaawansowane zapytania	168
Operator UNION	168
Samozłączenie w celu określenia, czy dana wartość ustanowiła rekord	172
Nowi i powracający klienci według tygodni	176
Podsumowanie	179
Ćwiczenia	180
Rozdział 12. Tworzenie zbiorów danych na potrzeby uczenia maszynowego	181
Zbiory danych dla modeli szeregów czasowych	182
Zbiory danych do klasyfikacji binarnej	184
Tworzenie zbioru danych	186
Poszerzanie zbioru cech	189
Inżynieria cech	192
Kolejne kroki	195
Ćwiczenia	196
Rozdział 13. Przykłady tworzenia analitycznych zbiorów danych	197
Jakie czynniki wpływają na sprzedaż świeżych produktów?	197
Jak zmienia się sprzedaż w zależności od kodu pocztowego klienta, odległości od targowiska oraz danych demograficznych?	209
Jak rozkład cen produktów wpływa na sprzedaż?	215

Rozdział 14. Przechowywanie i modyfikowanie danych	222
Przechowywanie zbiorów danych w postaci tabel i widoków	222
Dodawanie kolumny ze znacznikiem czasu	225
Dodawanie wierszy i aktualizowanie wartości w tabelach	226
Korzystanie z SQL-a wewnątrz skryptów	229
Na zakończenie	230
Ćwiczenia	231
Dodatek A. Odpowiedzi	233