**Hurtownie danych – Spr. 1. (dwa etapy)**

PWr. WIiT, Data: 13.03.2025 r.

|  |  |  |
| --- | --- | --- |
| Student | Email: 272675@student.pwr.edu.pl | Ocena |
| Indeks | 272675 |  |
| Imię | Jakub |
| Nazwisko | Gańko |

Zestaw składa się z 1 zadania realizowanego w dwóch etapach oraz dodatkowych zadań z podstaw DML. Jeżeli nie potrafisz rozwiązać zadania, to próbuj podać, chociaż częściowe rozwiązanie lub uzasadnienie przyczyny braku rozwiązania. Pamiętaj o podaniu nr. indeksu oraz imienia i nazwiska.

SZBD: **MS SQL 2022/2019**

Baza danych: **Szpital**

Zad. 1.

Proszę utworzyć bazę danych, która umożliwi gromadzenie i przetwarzanie danych niezbędnych dla funkcjonowania szpitala. Rozpatrywany wycinek rzeczywistości (obszar dziedzinowy) obejmuje następujące określenia:

1. Szpital
2. Oddział
3. Lekarz
4. Pacjent
5. Choroba
6. Diagnoza

**Etap I**

Pierwszy etap procesu opracowania bazy danych dotyczy analizy obszaru dziedzinowego w zakresie wyżej wymienionych pojęć i obejmuje:

1. Charakterystykę **dziedziny** (słowny opis kluczowych aspektów dziedziny)
2. Definicję **podstawowych pojęć** (określenie semantyki obiektów), które mają być reprezentowane w bazie danych
3. Identyfikację **klas obiektów** dziedziny oraz ich własności
4. Identyfikację i specyfikację **reguł dziedzinowych** (biznesowych)
5. Definicję **słownika danych** dla lekarza, pacjenta, choroby
6. **Model konceptualny danych** oraz jego weryfikacja względem reguł biznesowych
7. Wykaz źródeł (strony Internetowe, książki, dokumentacje techniczne, itd.)

Poszczególne elementy mogą być opracowane (co najmniej jedna pozycja) z wykorzystanie sztucznej inteligencji (LLM)

**Zasady specyfikacji składowych dokumentacji:**

2. Definicja podstawowych pojęć (słownik dziedzinowy)

Tabela . Słownik pojęć dziedzinowych

|  |  |  |
| --- | --- | --- |
| Lp. | Nazwa | Definicja (semantyka) |
| 1. |  |  |
| 2. |  |  |
| … |  |  |

3. Specyfikacja własności klas obiektów

Tabela . Specyfikacja klas obiektów

|  |  |  |  |
| --- | --- | --- | --- |
| Lp. | Nazwa klasy | Zbiór własności klasy | Klucze kandydujące |
| 1. |  |  |  |
| 2. |  |  |  |
| … |  |  |  |

4. Specyfikacja reguł biznesowych

Tabela . Reguły biznesowe

|  |  |
| --- | --- |
| Lp. | Reguła |
| Reg/01 |  |
| Reg/02 |  |
| … |  |

Przykład:

Reg/01 – Szpital składa się z co najmniej jednego oddziału

Reg/02 – Oddział …

5. Słownik danych

Tabela . Słownik atrybutów klas

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1. | Nazwa klasy | | | |
| Lp. | Atrybut | Typ danych | Znaczenie | Ograniczenia dziedzinowe |
| 1.1 |  |  |  |  |
| 1.2 |  |  |  |  |
| … |  |  |  |  |
| 2. | Nazwa klasy | | | |
| Lp. | Atrybut | Typ danych | Znaczenie | Ograniczenia dziedzinowe |
| 2.1 |  |  |  |  |
| 2.2 |  |  |  |  |
| … |  |  |  |  |

6. Model konceptualny danych – diagram klas UML

Uwaga:

* Proszę nie modyfikować i nie usuwać definicji sprawozdania
* Dyskusja i ocena realizacji zadań pierwszego etapu odbędzie się na zajęciach lab. w dniu 20.03.2025 lub 21.03.2025

**Rozwiązanie:**

1. Charakterystykę **dziedziny** (słowny opis kluczowych aspektów dziedziny)

Obszar dziedziny: Zarządzanie szpitalem – gromadzenie i przetwarzanie informacji niezbędnych do funkcjonowania placówki medycznej, w tym**:**

* **Struktury organizacyjnej szpitala (podział na oddziały).**
* **Personelu medycznego (lekarze).**
* **Pacjentów (dane osobowe, historia diagnoz i chorób).**
* **Procesu leczenia (diagnozowanie chorób przez lekarzy).**

1. Słownik pojęć

|  |  |  |
| --- | --- | --- |
| Lp. | Nazwa | Definicja (semantyka) |
| 1. | Szpital | Placówka medyczna; może mieć wiele oddziałów, w których pracują lekarze i przyjmowani są pacjenci. |
| 2. | Oddział | |  | | --- | |  |  |  | | --- | | Wydzielona część szpitala, skupiająca się na określonej dziedzinie medycyny | |
| 3. | Lekarz | Osoba posiadająca odpowiednie wykształcenie i uprawnienia do leczenia pacjentów, wystawiania diagnoz. |
| 4. | Pacjent | |  | | --- | |  |  |  | | --- | | Osoba korzystająca z usług szpitala; przechowuje się o niej dane osobowe, historię leczenia i diagnoz. | |
| 5. | |  | | --- | |  |   Choroba | Jednostka chorobowa (np. zapalenie płuc, nadciśnienie), często klasyfikowana kodem (np. ICD-10). |
| 6. | Diagnoza | |  | | --- | |  |  |  |  |  | | --- | --- | --- | | |  | | --- | |  |  |  | | --- | | Akt rozpoznania choroby u pacjenta przez lekarza; zawiera informacje o dacie i charakterystyce rozpoznania. | | |

3. Specyfikacja własności klas obiektów

Tabela . Specyfikacja klas obiektów

|  |  |  |  |
| --- | --- | --- | --- |
| Lp. | Nazwa klasy | Zbiór własności klasy | Klucze kandydujące |
| 1. | Szpital | KodSzpitala, Nazwa, Adres, Typ (publiczny/prywatny) | KodSzpitala |
| 2. | Oddział | KodOddziału, NazwaOddziału, KodSzpitala | KodOddziału, (KodSzpitala,NazwaOddziału) |
| 3. | Lekarz | NrLicencji, Imie, Nazwisko, Specjalizacja, KodOddziału | NrLicencji |
| 4. | Pacjent | PESEL, Imie, Nazwisko, DataUrodzenia, Adres | PESEL |
| 5. | Choroba | KodICD, NazwaChoroby, Opis | KodICD |
| 6. | Diagnoza | NrDiagnozy, DataRozpoznanai, OpisDiagnozy, KodICD, PESEL, NrLIcencji | NrDiagnozy |

* 1. Reguły biznesowe

Tabela . Reguły biznesowe

|  |  |
| --- | --- |
| Lp. | Reguła |
| Reg/01 | Każdy Szpital musi mieć co najmniej jeden Oddział. |
| Reg/02 | Każdy Oddział należy dokładnie do jednego Szpitala. |
| Reg/03 | Lekarz jest identyfikowany unikalnym Numerem Licencji. Numer ten musi być jedyny w całym systemie. |
| Reg/04 | Pacjent jest identyfikowany unikalnym numerem PESEL, który musi być jedyny w całym systemie. |
| Reg/05 | KodICD w tabeli Choroba jest unikalny i służy do klasyfikacji choroby w skali całego systemu. |
| Reg/06 | |  | | --- | |  |  |  | | --- | | Diagnoza jest wystawiana przez dokładnie jednego lekarza i dotyczy dokładnie jednego pacjenta. | |
| Reg/07 | Data rozpoznania w diagnozie nie może być datą przyszłą. |
| Reg/08 | Każdy pacjent powinien mieć co najmniej jedną diagnozę w czasie swojego pobytu w szpitalu). |
| Reg/09 | Nazwa oddziału w obrębie tego samego szpitala nie może się dublować (np. dwa oddziały o tej samej nazwie w jednym szpitalu są niedozwolone). |
| Reg/10 | Jeśli lekarz jest przypisany do wielu oddziałów, należy to odwzorować relacją N:M (tabela pośrednia). |
| Reg/11 | Każdy oddział musi mieć przypisanego co najmniej jednego lekarza (chyba że to nowo utworzony oddział jeszcze nieuruchomiony – zależnie od wymagań). |
| Reg/12 | Każdy lekarz musi być przypisany do dokładnie jednego oddziału. |
| Reg/13 | Szpital może być prywatny lub publiczny |

5. Słownik danych

Tabela . Słownik atrybutów klas

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1. | Szpital | | | | | | | |
| Lp. | Atrybut | | Typ danych | | Znaczenie | | Ograniczenia dziedzinowe | |
| 1.1 | KodSzpitala | | |  | | --- | |  |  |  | | --- | | VARCHAR(20) | | | Unikalny kod/numer placówki | | Klucz główny, niepusty, unikalny | |
| 1.2 | Nazwa | | |  | | --- | |  |  |  | | --- | | VARCHAR(100) | | | |  | | --- | |  |  |  | | --- | | Nazwa szpitala | | | niepuste | |
| 1.3 | Adres | | |  | | --- | |  |  |  | | --- | | VARCHAR(200) | | | |  | | --- | |  |  |  | | --- | | Adres placówki | | | niepuste | |
| 1.4 | Typ | | VARCHAR(20) | | Rodzaj szpitala (publiczny/prywatny) | | wartość z ograniczonego zbioru {„publiczny”, „prywatny”} | |
| 2. | Oddział | | | | | | | |
| Lp. | Atrybut | | Typ danych | | Znaczenie | | Ograniczenia dziedzinowe | |
| 2.1 | KodOddzialu | | VARCHAR(20) | | Kod/numer oddziału | | Klucz główny, niepusty, unikalny (w skali całego systemu lub danego szpitala) | |
| 2.2 | NazwaOddzialu | | VARCHAR(100) | | Nazwa oddziału (np. „Chirurgia”, „Kardiologia”) | | niepuste | |
| 2.3 | KodSzpitala | | VARCHAR(20) | | Powiązanie z klasą Szpital | | Klucz obcy -> Szpital(KodSzpitala), niepuste | |
| 3. | Lekarz | | | | | | |
| Lp. | Atrybut | Typ danych | | Znaczenie | | Ograniczenia dziedzinowe | |
| 3.1 | NrLicencji | VARCHAR(20) | | Numer licencji lekarskiej (PK) | | Klucz główny, niepusty, unikalny | |
| 3.2 | Imie | VARCHAR(50) | | |  | | --- | |  |  |  | | --- | | Imię lekarza | | | niepuste | |
| 3.3 | Nazwisko | VARCHAR(50) | | Nazwisko lekarza | | niepuste | |
| 3.4 | Specjalizacja | VARCHAR(100) | | Określenie specjalizacji medycznej | | niepuste | |
| 3.5 | KodOddziału | VARCHAR(20) | | |  | | --- | |  |  |  | | --- | | Jeśli lekarz pracuje w jednym oddziale bezpośrednio | | | Klucz obcy -> Oddzial(KodOddzialu). | |
| 4. | Pacjent | | | | | | |
| Lp. | Atrybut | Typ danych | | Znaczenie | | Ograniczenia dziedzinowe | |
| 4.1 | PESEL | VARCHAR(11) | | Numer ewidencyjny pacjenta (PK) | | Klucz główny, niepusty, unikalny | |
| 4.2 | Imie | VARCHAR(50) | | Imie pacjenta | | Niepuste | |
| 4.3 | Nazwisko | VARCHAR(50) | | Nazwisko pacjenta | | Niepuste | |
| 4.4 | DataUrodzenia | DATE | | Data urodzenia | | Niepuste, <= data bieżąca | |
| 4.5 | Adres | VARCHAR(200) | | Adres pacjenta | | może być puste lub wymagane | |
| 5. | Choroba | | | | | | |
| Lp. | Atrybut | Typ danych | | Znaczenie | | Ograniczenia dziedzinowe | |
| 5.1 | KodICD | VARCHAR(10) | | Kod choroby wg klasyfikacji ICD (PK) | | Klucz główny, niepusty, unikalny | |
| 5.2 | NazwaChoroby | VARCHAR(100) | | Nazwa choroby (np. „Kaszel”) | | niepuste | |
| 5.3 | Opis | VARCHAR(50) | | Dodatkowe informacje o chorobie (opcjonalnie objawy, krótki opis) | | Może być puste | |
| 6. | Diagnoza | | | | | | |
| Lp. | Atrybut | Typ danych | | Znaczenie | | Ograniczenia dziedzinowe | |
| 6.1 | NrDiagnozy | VARCHAR(20) | | Numer diagnozy (PK, np. „DX0001234”) | | Klucz główny, niepusty, unikalny | |
| 6.2 | DataRozpoznania | DATE | | Data wystawienia diagnozy | | niepuste, <= bieżąca data | |
| 6.3 | OpisDiagnozy | VARCHAR(500) | | Notatki lekarza, komentarz do diagnozy | | może być puste | |
| 6.4 | KodICD | VARCHAR(10) | | Kod choroby, klucz obcy | | niepuste (choroba musi być przypisana) | |
| 6.5 | PESEL | VARCHAR(11) | | Numer pacjenta, klucz obcy | | niepuste (diagnoza dotyczy pacjenta) | |
| 6.6 | NrLicencji | VARCHAR(20) | | Numer licencji lekarza, klucz obcy | | niepuste (diagnoza musi mieć lekarza, który ją wystawił) | |

6.Diagram w UML

Obraz zawierający tekst, zrzut ekranu, diagram, Czcionka

Zawartość wygenerowana przez sztuczną inteligencję może być niepoprawna.

**---------------------------------------------------------------------------------------------------------------------------------**

**Etap II**

Etap ten obejmuje następujące zadania:

1. Implementację konceptualnego modelu danych w systemie MS SQL 2019 w postaci skryptu SQL (definicje tabel wraz z ograniczeniami)
2. Przygotowanie danych testowych, które pozwolą zweryfikować poprawność zdefiniowania ograniczeń dziedzinowych. Wykonanie testów (bez testów wydajnościowych) poprzez wykonanie próby wprowadzenia kilku rekordów do każdej tabeli, sprawdzając w ten sposób poprawność implementacji (zarówno poprawne dane, jak i niezgodne z obowiązującymi regułami – komentując i wyjaśniając uzyskane komunikaty z systemu SZBD). Dane testowe należy przygotować w postaci skryptu SQL (zbiór poleceń INSERT INTO ...)

**Rozwiązanie:**

----------------------------------------------------------------------------------------------------------------------------

**Etap III**

Podstawy SQL

Źródło danych:

Baza danych: AdventureWorks - SalesOrderHeader

**Zad. 1.**

Proszę utworzyć wykaz zamówień złożonych w pierwszym roku rejestracji zamówień (Identyfikator, Rok, Kwota zamówienia) . Przykładowy wynik zapytania przedstawiony jest poniżej w tabeli 5.

Zapytanie SQL + fragment wyniku (4 rekordy z ?)

**Tabela 5. Fragment wyników zapytania zad. 1.**

Tabela . Wynik zapytania zad. 1.

|  |  |  |
| --- | --- | --- |
| Identyfikator | Rok | Kwota |
| 45266 | 2012 | 27605.63 |
| 45267 | 2012 | 3899.68 |
| 45268 | 2012 | 944.62 |
| 45269 | 2012 | 2280.14 |
| … | … | … |

Rek.: 4/?

**Rozwiązanie:**

----------------------------------------------------------------------------

**Zad. 2.**

Proszę wyznaczyć liczbę klientów, którzy mają liczbę zamówień:

1. mniejszą niż 10
2. w przedziale 10 - 19
3. co najmniej 20

Przykładowy wynik zapytania przedstawiony jest poniżej w tabeli 6.:

Tabela 6. Fragment wyników zapytania zad. 2. . – wariant A

|  |  |
| --- | --- |
| Grupa | Liczba klientów |
| 0-9 | 19002 |
| 10-19 | 103 |
| 20 ... | 14 |

Rek.: 3/3

lub

Tabela 7. Fragment wyników zapytania zad. 2. – wariant B

|  |  |  |  |
| --- | --- | --- | --- |
| Grupa | 0-9 | 10-19 | 20 ... |
|  | 19002 | 103 | 14 |

Rek.: 1

**Rozwiązanie: Zapytanie SQL + wynik zapytania**

----------------------------------------------------------------------------

**Zad. 3.**

Ustalić, jakie czynniki mają wpływ na liczbę dokonanych zakupów. Przykładowy wynik zapytania przedstawiony jest poniżej w tabeli 2.3.

Źródło danych: SalesOrderHeaderSalesReason, ?

**Tabela 2.3. Fragment wyników zapytania zad. 3**

|  |  |
| --- | --- |
| Czynnik | Dotyczy |
| Price | 17473 |
| On Promotion | 3515 |
| Manufacturer | 1746 |
| … | … |

Rek.: 3/?

**Rozwiązanie: Zapytanie SQL + fragment wyniku (4 rekordy z ?)**

**Wnioski:**

*Uwaga:*

* Sprawozdanie bez wniosków końcowych nie będzie sprawdzane i tym samym ocena jest negatywna!
* Plik ze sprawozdaniem powinien mieć nazwę **Spr01HD-index-Nazwisko - 2025** i format **pdf**
* **Kompletne sprawozdanie należy przekazać w terminie do 26.03.2025 godz. 20:00**
* Oprogramowanie MS SQL 2022/2019
* Model konceptualny danych należy przygotować wykorzystując narzędzie wspierające język UML np. Visual Paradigm
* Wnioski powinny zawierać ocenę możliwości wykorzystania sztucznej inteligencji