

Sprawozdanie z projektu nr 1

Zaawansowane systemy baz danych

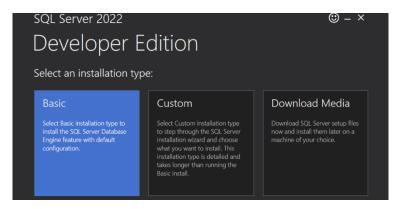
- 1. Case study sporządzone w oddzielnym pliku: case_study_311615.pdf
- 2. Wybór systemu:

Do projektu wybrałem system zarządzania relacyjnymi bazami danych MS SQL Server. Zdecydowałem się na taki wybór, ponieważ posiadam doświadczenie z kilku projektów studenckich, gdzie każdorazowo korzystałem z tej bazy danych. Co więcej mam już ponad rok doświadczenia komercyjnego w zawodzie, gdzie korzystamy głównie z tej bazy danych. Poza tym system wydaje się bardzo intuicyjny, dostępna jest dokładna instrukcja instalacji, a społeczeństwo zgromadzone wokół tego systemu jest bardzo zgrane i rozwiązania typowych problemów można łatwo znaleźć na forach takich jak StackOverflow, czy Reddit.

Proces instalacji:

SQL Server pobrałem z oficjalnej strony Microsoft'u: https://www.microsoft.com/pl-pl/sql-server-downloads

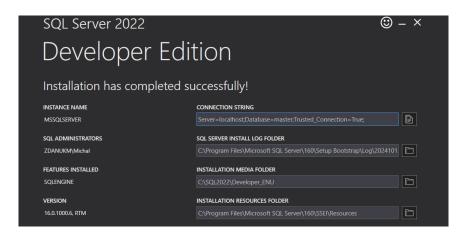
Z edycji Developer oraz Express wybrałem wersję Developer, ponieważ daje one pełną funkcjonalność za darmo pod warunkiem użytku własnego (nieprodukcyjnego).



Zrzut ekranu nr 1: wybrany typ instalacji Basic

Wybrany typ instalacji to Basic – podstawowa konfiguracja jest szybka do zainstalowania i wystarczająca na potrzeby projektu.

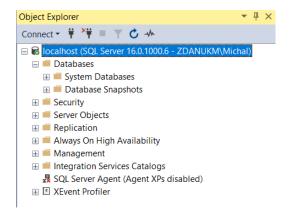
Z perspektywy czasu okazało się to błędem, co opiszę przy tworzeniu użytkowników i wnioskach



Zrzut ekranu nr 2: pomyślnie zainstalowany SQL Server

Server dostępny jest pod adresem localhost, baza systemowa to master.

Następnie zainstalowałem narzędzie graficzne do obsługi SSMS (SqlServerManagementStudio).



Zrzut ekranu nr 3: podłączenie się do serwera poprzez SSMS

Opis biznesowy systemu znajduje się w case_study_311615.pdf.
 Dokładniejszy opis ze strony bazodanowej:

System zakłada użytkowników, którzy mają rolę dostępną z: Klient schroniska, Pracownik schroniska, Administrator.

<u>Uwaga ogólna:</u> jeżeli w opisie kolumny nie ma jasnego określenia: <u>opcjonalne</u> oznacza to, że pole jest wymagalne (not null).

System będzie się składał z poniższych tabel:

Users:

- Id (unikalny identyfikator)
- FirstName (pole tekstowe)

- LastName (pole tekstowe)
- Roleld (odwołanie do tabeli UserRoles) określa rolę użytkownika

UserCredentials:

- Id (unikalny identyfikator)
- Email (pole tekstowe)
- PhoneNumber (pole tekstowe)
- DateOfBirth (data)
- Userld (odwołanie do tabeli *Users*) określa, którego właściciela szczegółowych danych

UserRoles:

- Id (unikalny identyfikator)
- Name (pole tekstowe) planowe role to: Client, Employee, Admin

FamiliarizationVisits:

- Id (unikalny identyfikator)
- DateOfVisit (data)
- Dogld (odwołanie do tabeli Dogs) każda wizyty dotyczy jednego psa
- Visitorld (odwołanie do tabeli *Users*) użytkownik z rolą Client, który zgłosił chęć wizyty
- Employeeld (odwołanie do tabeli *Users*) użytkownik z rolą Employee, który przypisany jest do przeprowadzenia wizyty
- VisitStatusId (odwołanie do tabeli VisitStatuses) określa stan wizyty

VisitStatuses:

- Id (unikalny identyfikator)
- Name (pole tekstowe) planowe statusy to: zaplanowana, odbyta, odwołana

AdoptionCards:

- Id (unikalny identyfikator)
- RegistrationDate (data rejestracji)
- Userld (odwołanie do tabeli *Users*) określa właściciela karty adopcyjnej

AdoptionRecords:

- Id (unikalny identyfikator)
- LastUpdateDate (data) data wykonania ostatniej zmiany na rekordzie adopcyjnym
- RejectionReason (pole tekstowe, <u>opcjonalne</u>) w przypadku odrzucenia należy wypełnić powód, poinformować użytkownika o przyczynie, w przyjętych adopcjach pole zostaje puste

- Dogld (odwołanie do tabeli Dogs) określa psa, która podlega procesowi adopcji
- AdoptionCardId (odwołanie do tabeli *AdoptionCards*) określa, do której karty adopcyjnej należy rekord adopcyjny
- AdoptionStatusId (odwołanie do tabeli AdoptionStatuses) określa stan adopcji

AdoptionStatuses:

- Id (unikalny identyfikator)
- Name (pole tekstowe) planowe statusy to: złożony wniosek, zaakceptowana, zrealizowana, odrzucona

Dogs:

- Id (unikalny identyfikator)
- Name (pole tekstowe)
- Breed (pole tekstowe)
- DogStatusId (odwołanie do tabeli *DogStatuses*) określa stan psa w systemie
- ShelterId (odwołanie do tabeli *Shelters*) powiązanie z schroniskiem, w którym znajduje się pies

DogStatuses:

- Id (unikalny identyfikator)
- Name (pole tekstowe) planowe statusy to: zarejestrowany (świeżo dodany pies do systemu), dostępny (status, w którym pies jest dostępny dla klientów), adoptowany (pies niedostępny dla klientów, nie można edytować jego danych), uśpiony (opis ten sam jak do adoptowany)

DogCharacteristics:

- Id (unikalny identyfikator)
- Weight (pole liczbowe)
- Description (pole tekstowe, opcjonalne)
- Dogld (odwołanie do tabeli Dogs) określa, którego psa dotyczy charakterystyka

Shelters:

- Id (unikalny identyfikator)
- Name (pole tekstowe)
- AddressId (odwołanie do tabeli *Adresses*) powiązanie z adresem, gdzie znajduje się schronisko

Addresses:

- Id (unikalny identyfikator)

- Street (pole tekstowe)
- City (pole tekstowe)
- PostalCode (pole tekstowe)
- Additional Description (pole tekstowe, opcjonalne)

MedicalCards:

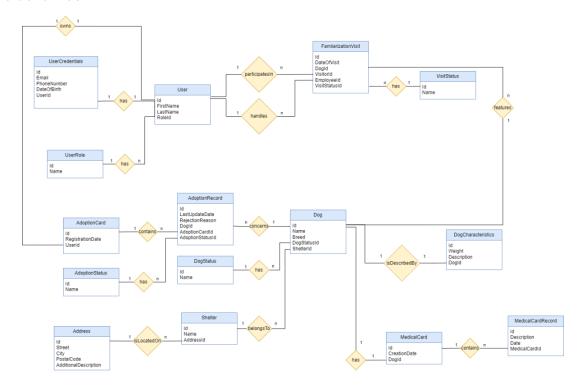
- Id (unikalny identyfikator)
- CreationDate (data) data założenia karty medycznej
- Dogld (odwołanie do tabeli *Dogs*) powiązanie z psem, którego dotyczy karta medyczna

MedicalCardRecords:

- Id (unikalny identyfikator)
- MedicalCardId (odwołanie do tabeli MedicalCards) określa, do której karty medycznej należy rekord medyczny
- Description (pole tekstowe)
- Date (data)

Diagram ERD:

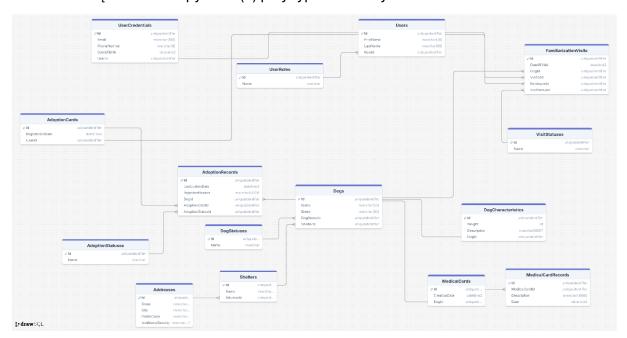
Diagram ERD został sporządzony przy pomocy darmowego edytora diagramów online lucidchart.com



Zrzut ekranu nr 4: diagram ERD

Schemat logiczny bazy danych:

Schemat został wykonany przy użyciu aplikacji webowej: drawsql.app. Pola nullowalne oznaczone są znakiem zapytania (?) przy typie kolumny.



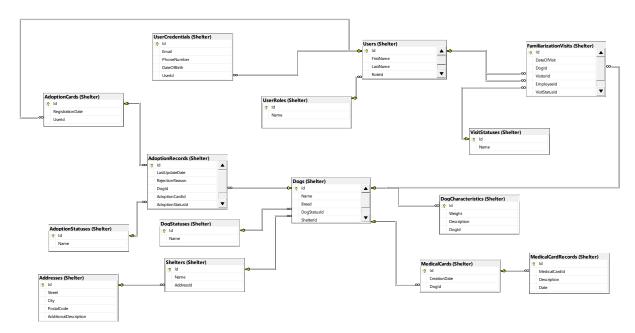
Zrzut ekranu nr 5: wstępny schemat logiczny bazy danych

Argumentacja decyzji:

Nie zdecydowałem się na zastosowanie kluczy kompozytowych pomimo kilku bardzo widocznych miejsc, gdzie można byłoby wykorzystać ich zalety (przykładowo tabela Adresses mogłaby mieć klucz kompozytowy, który składałby się z kolumn Street, City, PostalCode). Zdecydowanie preferuję korzystanie z unikalnych kluczy Id w każdej tabeli – jest to moim zdaniem o wiele wygodniejsze i bardziej intuicyjne rozwiązanie. Pozwala to na pisanie czystszego kodu backendowego, gdzie wszystkie encje mogą być abstrakcją bazowej encji, która ma zawsze posiadać swój identyfikator. Kolejnym plusem jest wykonywanie zapytać GET do REST API – załóżmy, że mamy endpoint do przeglądania listy encji oraz usługę GET do pobrania szczegółów konkretnej encji. Bez unikalnego identyfikatora musielibyśmy budować nasze zapytania na około tzn. podając wszystkie dopasowania kolumn, aby uzyskać rekord. Zastosowanie Id pozwala zatem na prostsze zapytania i spełnianie postulatów RESTa. Zdaję sobie sprawę, że klucze kompozytowe zamiast identyfikatorów pozwoliłyby mi zabezpieczyć się przed dodaniem dwóch takich samych adresów (kombinacji wymienionych wyżej trzech kolumn), nie mniej jednak tę logikę/walidację biznesową wolę odsunąć do warstwy domenowej/aplikacyjnej systemu, niż używać w bazie. Bardzo starałem się uzyskać jak najczystszą postać 3NF bazy, jednak zdaje mi się, że w jednym miejscu lekko nagiąłem normę. Pozostawiłem kolumnę Breed w tabeli Dogs. Tę kolumnę można by przenieść do DogCharacteristics, jednak chciałbym by te najbardziej bazowe informacje o psach wymagały select'a tylko do jednej tabeli bez żadnych JOINów.

Skrypt, którym wygenerowałem bazę danych wraz z konfiguracją kluczy oraz indeksów dołączam w pliku: *SheltersDB_init_sql.sql*.

Poniżej zrzut ekranu wygenerowanego schematu bazy przez SSMS (dostępny również w pliku ssms_db_diagram_311615.png):



Zrzut ekranu nr 6: wygenerowany diagram przez SSMS

Wypełnianie bazy danymi:

Skorzystałem z dedykowanego, darmowego narzędzia, które pozwala na relatywnie proste dodawanie rekordów imitujących prawdziwe dane: https://www.mockaroo.com/

Część danych bardziej specjalistycznych takich jak opis medycznego rekordu, czy opis psa zostały wygenerowane przy pomocy sztucznej inteligencji: https://chatgpt.com

Jedyną tabelą, gdzie pozwoliłem sobie zaseedować mniej niż 4 rekordy to tabela UserRoles (posiadam 3 rekordy), ponieważ chciałem zostać przy opracowanych założeniem i logice biznesowej. Za to zdecydowałem się znacząco przekroczyć próg minimalny do lepszej prezentacji danych i lepszych możliwości prezentacji dalej. Tabela Dogs, DogCharacteristics oraz MedicalCards posiadają po 14 rekordów, a tabela MedicalCardRecords, aż 20. To, aby zobrazować różne kombinacje danych.

Skrypt, który służy seedowaniu bazy dołączam w pliku SeedDB_with_fake_data.sql.

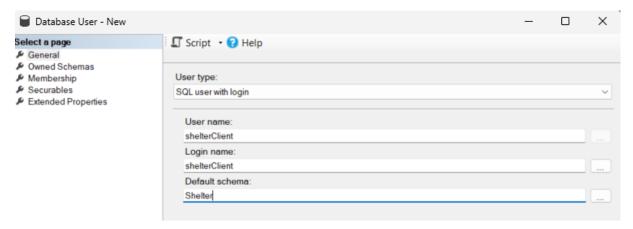
Tworzenie użytkowników:

Skorzystałem z możliwości dodawania użytkowników i konfiguracji ich uprawnień poprzez graficzny interfejs jaki dostarcza nam SSMS. Do dodania każdego użytkownika należy podać login, hasło oraz defaultowy schemat. Tutaj przydał się utworzony dla projektu schemat Shelter, ponieważ dzięki temu bezpiecznie możemy udostępnić tylko interesujący nas fragment systemu bazodanowego.

Zgodnie z założeniami biznesowymi dodałem trzy loginy i trzech użytkowników do bazy:

- shelterClient
- shelterEmployee
- shelterAdmin

<u>Użytkownik shelterClient:</u>



Zrzut ekranu nr 7: tworzenie użytkownika shelterClient

W tym momencie spróbowałem się zalogować na stworzonego użytkownika (użytkownik był widoczny na liście userów bazy), ale otrzymałem komunikat błędu:

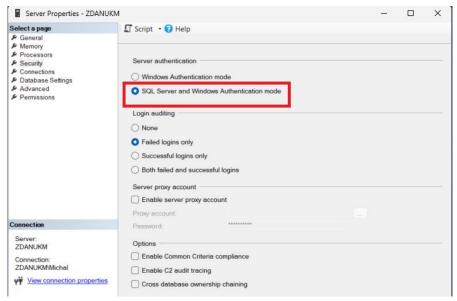
A connection was successfully established with the server, but then an error occurred during the login process. (provider: Shared Memory Provider, error: 0 - Na drugim końcu potoku nie ma żadnego procesu.) (Microsoft SQL Server, Error: 233)

Okazało się, że przy bazowej konfiguracji zabrakło mi ustawienia autentykacji serwera jako Mixed Authentication. Posiadałem logowanie tylko poprzez Windows Authentication.

Problem oraz rozwiązanie znalazłem na forum:

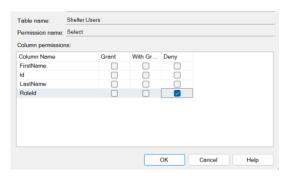
https://stackoverflow.com/questions/8075784/error-message-provider-shared-memory-provider-error-0-no-process-is-on-th

Zalogowałem się ponownie jako Admin przez Windows Authentication i w ustawieniach serwera przestawiłem ustawienie autentykacji. Po restarcie serwera logowanie zaczęło działać.

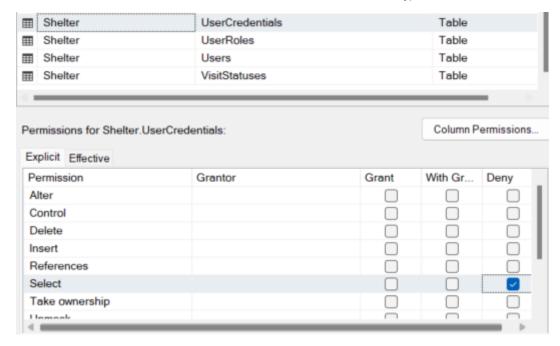


Zrzut ekranu nr 8: poprawiona konfiguracja serwera

Użytkownik klienta otrzymał prawa odczytu (SELECT) do prawie wszystkich tabel. Zablokowana mu została tabela UserCredentials, ponieważ zawiera wrażliwe dane jak email, czy numer telefonu. Poza tym zablokowałem mu dostęp do kolumny Roleld w tabeli Users, aby nie móc sprawdzać roli innych użytkowników (ze względów bezpieczeństwa). Poza tym użytkownik ma możliwość zapytań INSERT tylko do dwóch tabel: AdoptionRecords, aby mógł składać prośby o adopcję oraz FamiliarizationVisit, aby móc umawiać się na wizyty zapoznawcze.



Zrzut ekranu nr 9: zablokowanie dostępu do RoleId

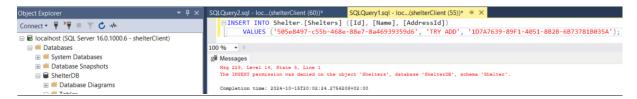


Zrzut ekranu nr 10: zablokowanie dostępu do tabeli UserCredentials

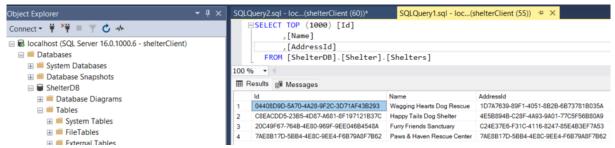
Poniżej prezentuję testy wykazujące poprawne ograniczenie uprawnień:



Zrzut ekranu nr 11: zablokowana próba dostępu do tabeli UserCredentials



Zrzut ekranu nr 12: zablokowana próba dodania danych do tabeli Shelters

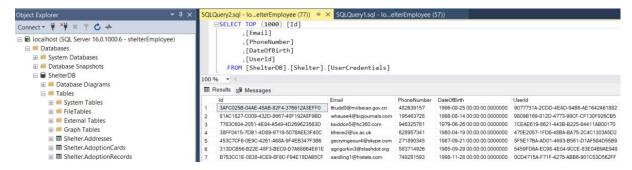


Zrzut ekranu nr 13: poprawny odczyt z dozwolonej tabeli Shelters

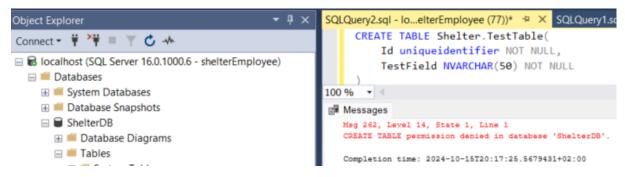
<u>Użytkownik shelterEmployee:</u>

Użytkownik pracownika powinien mieć dostęp SELECT do wszystkich tabel, bez ograniczeń (pracownik dodaje innych użytkowników do bazy, zatem musi dodać im dane poufne do tabeli UserCredentials). Pracownik ma również prawa UPDATE oraz INSERT do prawie wszystkich tabel. Wyjątkiem są tabele słownikowe (system jest zdefiniowany i zmian tutaj może dokonać tylko administrator). Zablokowana jest również możliwość tworzenia i wprowadzania zmian do tabel (to rola administratora).

Poniżej prezentuję testy wykazujące poprawne ograniczenie uprawnień:



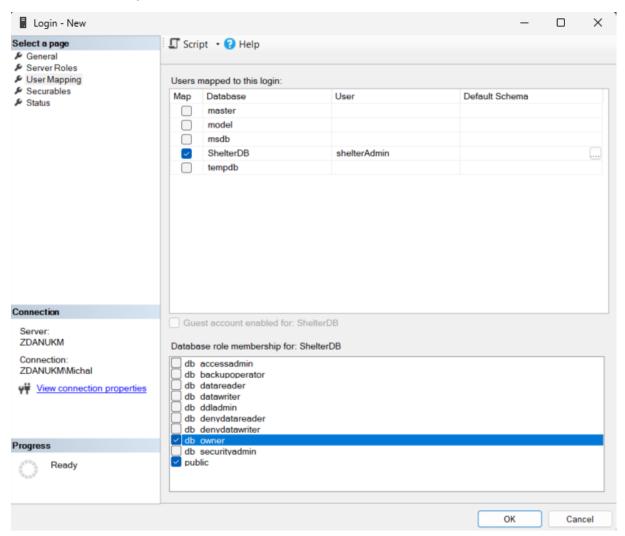
Zrzut ekranu nr 14: dostęp odczytu do tabeli UserCredentials



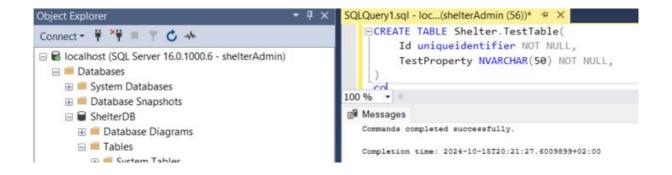
Zrzut ekranu nr 15: zablokowane tworzenie tabeli

<u>Użytkownik shelterAdmin:</u>

Użytkownik administratora ma pełen dostęp do bazy. Ustawiłem mu uprawnienia db_ownera do bazy ShelterDB.



Zrzut ekranu nr 16: nadanie uprawnień użytkownikowi admina



Zrzut ekranu nr 17: pomyślne utworzenie tabeli przez admina

Zapytania SQL:

1) lista użytkowników z rolą klienta, którym przynajmniej raz odrzucono wniosek o adopcję, wraz ze szczegółami ich ostatniej odrzuconej adopcji i związanego z nią psa. (zapytanie przydatne pracownikom/analitykom pod kątem analizy statystyk), dane posortowane względem daty ostatniego rekordu adopcyjnego

Wynik zapytania:



Zrzut ekranu nr 18: wyniki zapytania

2) zapytanie zwraca pogrupowane wyniki wizyt zapoznawczych dokonanych przez pracowników. Wyniki posortowane od największej liczby wizyt zapoznawczych do najmniejszej. Poza tym zwrócone są szczegóły ostatnich wizyt prowadzonych przez pracowników (data wizyty oraz jaki pies uczestniczył w wizycie).

Wynik zapytania:

	EmployeeFirstName	EmployeeLastName	SupervisedVisitsCount	LastVisitDate	LastVisitedDog
1	Addison	Grigorkin	3	2024-08-27 00:00:00.0000000	Vassily
2	Gigi	Scrymgeour	1	2024-06-12 00:00:00.0000000	Burton

Zrzut ekranu nr 19: wyniki zapytania

Kod sql z zapytaniami znajduje się w załączonym pliku Queries.sql.

Perspektywy:

Udostępnione widoki klientom:

 ClientAvailableDogsView – perspektywa umożliwia klientom przeglądanie dostępnych psów w schroniskach wraz z informacjami o ich rasie, statusie zdrowotnym oraz lokalizacją schroniska, w którym się znajdują



Zrzut ekranu nr 20: wyniki zwrócone z widoku ClientAvailableDogsView

 ClientAdoptionHistoryView – perspektywa umożliwia klientom schroniska przeglądanie historii adopcji, w tym statusu wniosków, przyczyny odrzucenia, czy informacji o zaakceptowanej adopcji.

	Userld	ClientFirstName	ClientLastName	DogName	RequestLastUpdate	AdoptionStatus	RejectionReason
1	9077731A-2CDD-4EAD-94B6-AE1642A61882	Timofei	Frudd	Vassily	2024-09-08 00:00:00.0000000	Realised	NULL
2	0CD4715A-F71F-4275-ABB6-901C53D562FF	Son	Ardling	Hanson	2024-10-06 00:00:00.0000000	Realised	NULL
3	9809B169-812D-4773-99CF-CF130F926CB5	Waylon	Hause	Earlie	2024-10-12 00:00:00.0000000	Rejected	Adoption request was denied due to the home not
4	9077731A-2CDD-4EAD-94B6-AE1642A61882	Timofei	Frudd	Jilleen	2024-09-05 00:00:00.0000000	Accepted	NULL
5	470E2057-1FD6-48BA-BA75-2C4C1303A5D2	Karylin	Freire	Belva	2024-10-02 00:00:00.00000000	Realised	NULL

Zrzut ekranu nr 21: wyniki zwrócone z widoku ClientAdoptionHistoryView

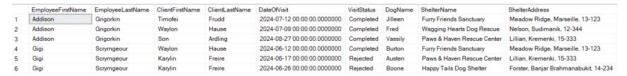
Udostępnione widoki pracownikom:

1) EmployeeDogManagementView - perspektywa dostarcza listę psów przypisanych do danego schroniska, wraz z ich aktualnym statusem, charakterystyką oraz ostatnim wpisem w karcie medycznej.



Zrzut ekranu nr 22: wyniki zwrócone z widoku EmployeeDogManagementView

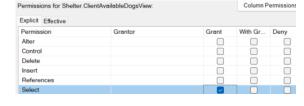
2) *EmployeeVisitSupervisionView* - perspektywa umożliwia pracownikom śledzenie i nadzorowanie wizyt zapoznawczych, którymi się zajmują.



Zrzut ekranu nr 23: wyniki zwrócone z widoku EmployeeVisitSupervisionView

Nadanie uprawnień klientowi:

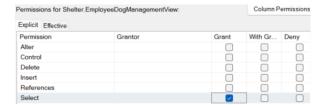




Zrzut ekranu nr 24: nadanie uprawnień do

widoków udostępnionych dla klienta

Nadanie uprawnień pracownikom:





Zrzut ekranu nr 25: nadanie uprawnień do widoków udostępnionych dla pracownika Skrypt do utworzenia widoków znajduje się w załączonym pliku: *Views_script.sql*.

Indeksy:

Zdecydowałem się na wprowadzenie kilku, najbardziej istotnych indeksów, dla wybranych kolumn:

```
-- indeksy do tabli Users

CREATE INDEX idx_users_lastname ON Shelter.Users(LastName);

-- indeksy do tabli AdoptionRecords

CREATE INDEX idx_adoptionrecords_card_status ON Shelter.AdoptionRecords(AdoptionCardId, AdoptionStatusId);

CREATE INDEX idx_adoptionrecords_dogid ON Shelter.AdoptionRecords(DogId);

-- indeksy do tabli Dogs

CREATE INDEX idx_dogs_name ON Shelter.Dogs(Name);

CREATE INDEX idx_dogs_breed_shelterid ON Shelter.Dogs(Breed, ShelterId);

-- indeksy do tabli FamiliarizationVisits

CREATE INDEX idx_familiarizationvisits_date ON Shelter.FamiliarizationVisits(DateOfVisit);

CREATE INDEX idx_familiarizationvisits_visitorid ON Shelter.FamiliarizationVisits(VisitorId);

-- indeksy do tabli MedicalCards

CREATE INDEX idx_medicalcards_dogid ON Shelter.MedicalCards(DogId);
```

Zrzut ekranu nr 26: zakładanie indeksów na wybranych kolumnach

Skrypt do zakładania indeksów znajduje się w pliku: Indexes.sql.

Uzasadnienie:

- *idx_users_lastname* nazwisko często jest wykorzystywane do filtrowania lub wyszukiwania użytkowników, więc indeks powinien przyspieszyć te operacje.
- idx_adoptionrecords_card_status AdoptionCardId i AdoptionStatusId mogą być często używane w parze. Adresuje on zapytania o rekordy adopcyjne konkretnego użytkownika z uwzględnieniem filtrowania po statusie.
- idx_adoptionrecords_dogid ułatwia szybkie wyszukiwanie rekordów adopcji związanych z konkretnym psem. Jest to istotne, biorąc pod uwagę, że zapytania mogą być skoncentrowane na poszczególnych psach (analizowanie historii konkretnego psa).
- idx_dogs_name ułatwia szybkie wyszukiwanie psów według ich imienia.
- idx_dogs_breed_shelterid wspiera zapytania dotyczące psów określonej rasy w konkretnym schronisku, co jest przydatne dla użytkownika, gdyby chciał zweryfikować dostępność konkretnej rasy w okolicznym schronisku.
- idx_medicalcards_dogid pozwala na szybsze wyszukiwanie kart medycznych przypisanych do konkretnego psa, co jest kluczowe dla zarządzania informacjami medycznymi.

Indeksy na jedną kolumnę warto stosować w przypadku kolumn na bazie, których często wykonywane są zapytania, bądź stosujemy warunki filtrujące z użyciem WHERE albo sortujemy rekordy przez ORDER BY, również, gdy łączymy tabelę na bazie kolumny przy JOIN. Indeksy wielokolumnowe (kompozytowe) przydatne są zwłaszcza dla zapytań, które korzystają z wielu warunków jednocześnie.

Wnioski:

W tym projekcie zdecydowanie zbytnio pospieszyłem się z konfiguracją środowiska przez co natrafiłem na problemy z dodawaniem użytkowników do bazy. Proces konfiguracji nie

powinien być traktowany mniej poważnie niż proces analizy, czy projektowania. Poza tym mój pierwotny projekt bazy (podział logiczny i rozmieszczenie danych) nie był optymalny, przez co wielokrotnie musiałem poprawiać diagramy i poświęciłem na to znaczną część czasu, powinienem był bardziej przemyśleć problematykę na początku projektu oraz bardziej zastanowić się jakie dane są potrzebne użytkowników, aby móc zrealizować ich rzeczywiste potrzeby. Przyznam, że pierwszy raz miałem do czynienia z dodawaniem różnych użytkowników do bazy oraz konfigurowania im dedykowanych praw dostępu, co zapewnia wyższy poziom bezpieczeństwa – uważam to za wartościową umiejętność i nową wiedzę zyskaną przy realizacji tego projektu.