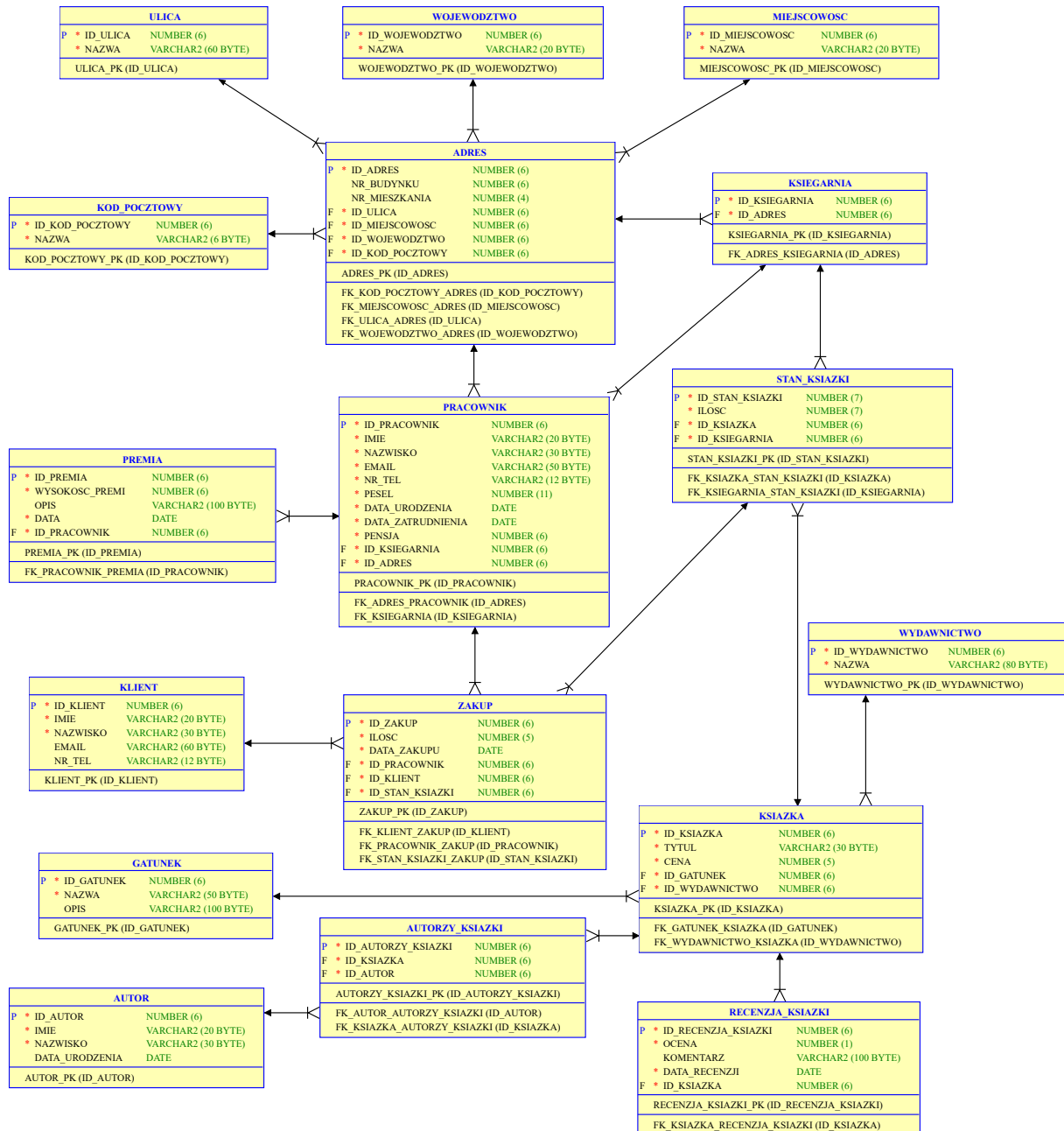


Bazy Danych 2 - Projekt	
Studia: Stacjonarne 1 stopnia	Kierunek: Informatyka
Data Wykonania: 30.05.2022	Temat projektu: Sieć Księgarni

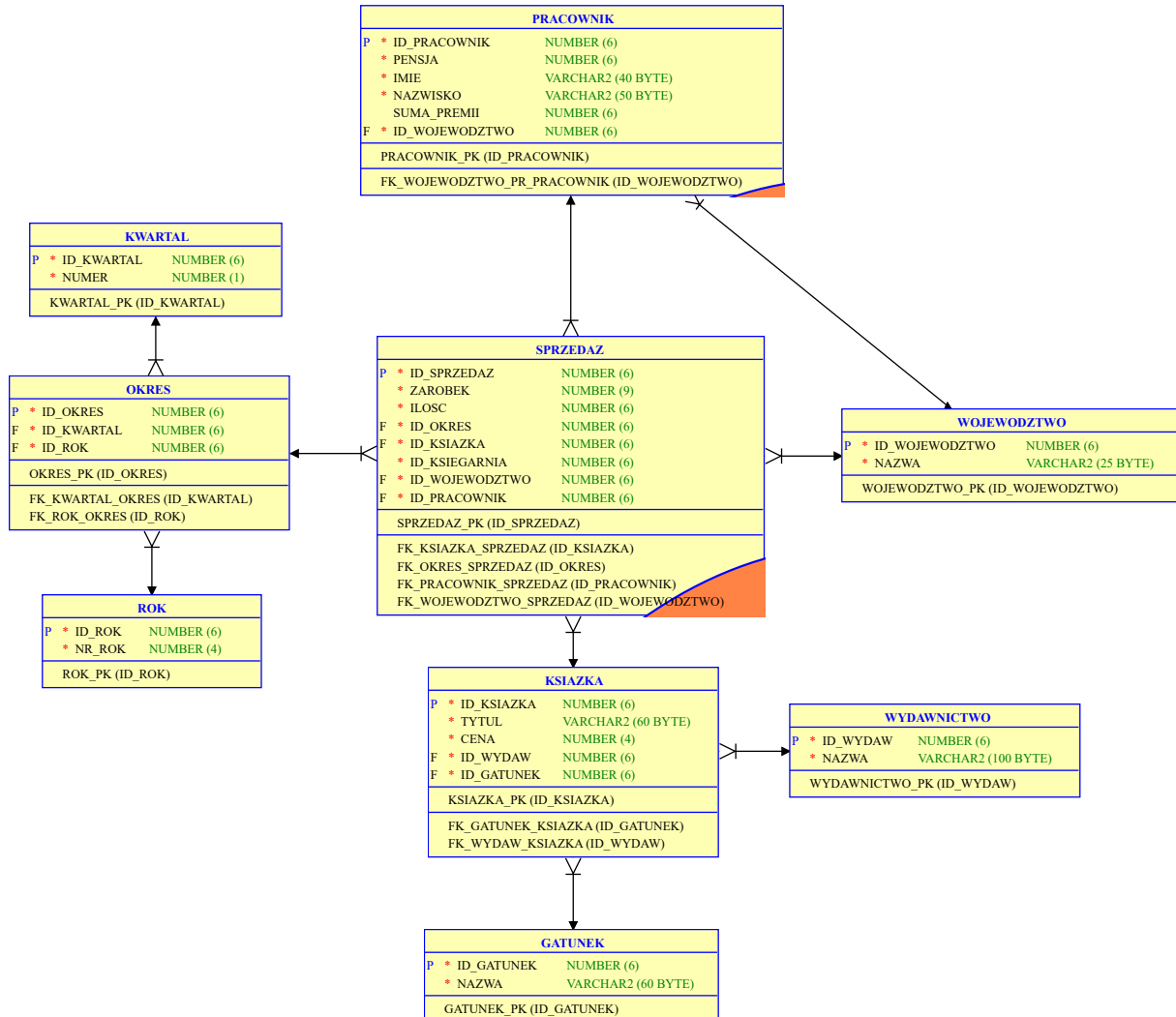
1. Opis problemu

Moim zadaniem było utworzyć bazę danych dużej sieci księgarni. Główną transakcją na której oparłem moją bazę jest zakup książki. Baza pozwala na zakup wybranej książki w wybranej księgarni przez klienta. Sprzedaż obsługuje pracownik, który przypisany jest do konkretnej księgarni. Baza przechowuje również recenzje klientów na temat konkretnej książki. Baza zawiera również inne tabele potrzebne do szczegółowego przeprowadzania transakcji takie jak adres czy słowniki takie jak gatunek, czy miejscowość. Każda książka może mieć więcej niż jednego autora dlatego zostały wykorzystane dwie tabele: autorzy_książki, oraz autor. Opisy wszystkich tabel znajdują się w dalszej części sprawozdania. Do pracy użyłem bazy danych Oracle w wersji 11g oraz programu do obsługi bazy sqldeveloper.

2.1. Schemat bazy danych



2.2. Schemat hurtowni danych



3.1. Opis tabel bazy danych

- **ulica** – Słownik zawierający listę ulic.
- **miescowosc** – Słownik zawierający listę miejscowości.
- **wojewodztwo** – Słownik zawierający listę województw.
- **kod_pocztowy** – Słownik zawierający listę kodów pocztowych .
- **adres** – Tabela zawierająca adresy pracowników oraz księgarni. Zawiera w sobie klucze obce do wyżej wymienionych słowników.
- **ksiegarnia** – Tabela reprezentująca księgarnie jako budynki.
- **pracownik** – Tabela zawiera listę pracowników.
- **premia** – Tabela zawierająca listę wypłaconych premii dla poszczególnych pracowników.
- **klient** – Tabela zawierająca listę klientów.
- **wydawnictwo** - Słownik zawierający listę wydawnictw.
- **gatunek** - Słownik zawierający listę gatunków książek.
- **ksiazka** – Tabela określa książkę jako konkretny „tytuł”, nie określa poszczególnych egzemplarzy książek. Posiada klucze obce słowników „wydawnictwo” oraz „gatunek”.
- **autor** – Słownik zawierający listę autorów.
- **autorzy_ksiazki** – Tabela stworzona do połączenia tabel autor oraz książka w celu eliminacji relacji „wiele do wielu”. Dzięki temu zabiegowi możliwe jest przypisanie wielu autorów do jednej książki.
- **recenzja_ksiazki** – Tabela zawierająca opinie użytkowników na temat konkretnych książek. Ocena w systemie gwiazdek od 1 do 5 oraz komentarz.
- **stan_ksiazki** – Tabela, która ma za zadanie przechowywać informacje o ilości konkretnych książek w konkretnych księgarniach. Pozwala sprawdzić czy konkretny egzemplarz jest w konkretnej księgarni.
- **zakup** – Tabela, która przechowuje transakcje kupna książek przez klientów. Przechowuje informacje o ilości zakupionych książek, pracowniku, który dokonał sprzedaży a także o dacie transakcji.

3.2. Opis tabel hurtowni danych

- **sprzedaz** – Tabela zawierająca wszystkie transakcje.
- **pracownik** – Tabela zawierająca dane na temat pracowników, w tym wyliczoną sumę premii.
- **kwartal** – Słownik zawierający cztery kwartały .
- **okres** – Tabela zawierająca okresy tzn. czas w jakim była wykonana transakcja.
- **rok** – Słownik zawierający lata w których doszło do transakcji.
- **województwo** – Słownik województw.
- **ksiazka** – Tabela zawierająca dane dotyczące książek .
- **wydawnictwo** – Słownik zawierający listę wydawnictw.
- **gatunek** – Słownik zawierający listę gatunków książek.

4. Zasilenie bazy danymi

Do generowania danych napisałem program w języku Python, który utworzył pliki tekstowe zawierające dane. Program Python zawarty jest w moim archiwum. Do importu danych do bazy użyłem sqloader. W archiwum znajdują się pliki konfiguracyjne, dane oraz skrypt bat ładujący te dane.

5. Opis procesu transformacji bazy do hurtowni

Do transformacji danych z mojej bazy danych do hurtowni, w większości wykorzystałem skrypty sql tworzące tabele, następnie wyeksportowałem je do plików csv po czym zaimportowałem je do hurtowni przy użyciu sqloadera. Wyjątkiem są tu tabele rok, kwartał oraz okres, które stworzyłem w Excelu. Wszystkie pliki konfiguracyjne wraz z danymi oraz skryptem bat znajdują się w archiwum.

6. Porównanie wykonania zapytań

Pomiar czasu wykonałem w programie DBeaver z ustawionym limitem odczytywanych wierszy na maksimum, aby mieć pewność, że zostaną wyświetlone wszystkie wyniki. Gwarantuje nam to czas wykonania całego zapytania. Poniżej widnieją wyniki dla wybranych zapytań.

-Rollup (Zapytanie nr.3)

- baza: 130 ms
- hurtownia: 75 ms

-Cube (Zapytanie nr.3)

- baza: 230 ms
- hurtownia: 190 ms

-Partycje Obliczeniowe (Zapytanie nr.1)

- baza: 120 ms
- hurtownia: 50 ms

-Okna czasowe (Zapytanie nr.2)

- baza: 70 ms
- hurtownia: 10 ms

-Funkcje rankingowe (Zapytanie nr.1)

- baza: 60 ms
- hurtownia: 5 ms

7. Uwagi i Wnioski

Udało mi się wykonać wszystkie zdania zarówno z bazy danych jak i hurtowni. Hurtownie zaprojektowałem opierając się o schemat „płatka śniegu”. Czas wykonania tych samych zapytań dla hurtowni danych był krótszy niż dla bazy danych. Jest to spowodowane mniejszą ilością połączonych tabel w hurtowni w zapytaniu jak i ogólną mniejszą złożonością hurtowni. Hurtownia przez to, że wykonuje znacznie szybciej zapytania, może przechowywać dużo więcej danych historycznych, oraz umożliwia ich szybszą analizę.