**Instytut Informatyki  
Kolegium Nauk Przyrodniczych  
Uniwersytet Rzeszowski**

**Przedmiot:**

**Sieci semantyczne**

**Dokumentacja techniczna projektu:**

***Baza danych o klubach piłki nożnej***

**Wykonał: Albert Mazur**

**Prowadzący: dr. W. Kozioł**

**Rzeszów 2023**

Spis treści

[1. Nazwa bazy danych i opis 2](#_Toc138421450)

[2. Opis źródeł danych 2](#_Toc138421451)

[3. Zdjęcie całej bazy danych 2](#_Toc138421452)

[4. Opis struktury bazy danych 4](#_Toc138421453)

[5. Opis programu 5](#_Toc138421455)

[6. Opis węzłów 11](#_Toc138421456)

[7. Opis relacji 13](#_Toc138421457)

[8. Nietrywialne zapytania 14](#_Toc138421458)

[9. Możliwości użycia 19](#_Toc138421459)

[10. Podsumowanie 19](#_Toc138421460)

[11. Literatura 19](#_Toc138421461)

# Nazwa bazy danych i opis

Temat: Baza danych o klubach piłki nożnej

Opis: Baza zawiera informacje o klubach z 25 lig z Europy, stadionach, piłkarzach, trenerach i ich kontraktach. Zawiera również informacje takie jak puchary zdobyła każda reprezentacja oraz klub piłkarki w danym sezonie.

# Opis źródeł danych

Źródłem była strona która zawiera mnóstwo informacji o piłce <https://www.transfermarkt.pl>

# Zdjęcie całej bazy danych

**Obraz zawierający fajerwerki

Opis wygenerowany automatycznie**

# Opis struktury bazy danych

# Obraz zawierający tekst, zrzut ekranu, krąg, Czcionka Opis wygenerowany automatycznie

W bazie danych znajdują się węzły reprezentujące różne encje, takie jak kraje (Kraj), trofea (Trofeum), kluby (Klub), stadiony (Stadion), trenerzy (Trener) oraz zawodnicy (Zawodnik). Każdy węzeł ma określone właściwości, które opisują daną encję. Węzły reprezentujące kraje mają właściwość „nazwa”, węzły reprezentujące kluby mają właściwość „nazwa”, a węzły reprezentujące trofea mają właściwość „nazwa” i „sezon”, a zawodnik mają właściwość „Imię i nazwisko”, „pozycja”, „data urodzenia”. Węzły stadiony mają właściwości „nazwa” i „pojemność”. Węzły trenerzy mają właściwości „imię i nazwisko” i „wiek”

W bazie danych istnieją również relacje pomiędzy węzłami, które opisują różne zależności. Oto opis kilku z tych relacji:

Relacja "wygrali": Relacja pomiędzy węzłem reprezentującym kraj (Kraj) a węzłem reprezentującym trofeum (Trofeum). Ta relacja opisuje fakt, że kraj wygrał konkretne trofeum w określonym sezonie. Posiada również atrybut "sezon", który określa konkretny sezon, w którym trofeum zostało wygrane.

Relacja "nalezy": Relacja pomiędzy węzłem reprezentującym klub (Klub) a węzłem reprezentującym kraj (Kraj). Ta relacja wskazuje, że klub należy do określonego kraju. Posiada atrybut "nazwa", który opisuje przynależność klubu do konkretnej ligi.

Relacja "nalezy": Relacja pomiędzy węzłem reprezentującym stadion (Stadion) a węzłem reprezentującym klub (Klub). Ta relacja wskazuje, że stadion należy do określonego klubu.

Relacja "pochodzi": Relacja pomiędzy węzłem reprezentującym kraj (Kraj) a węzłem reprezentującym zawodnika (Zawodnik) lub trenera (Trener). Ta relacja wskazuje, że zawodnik lub trener pochodzi z określonego kraju.

Relacja "kontrakt": Relacja pomiędzy węzłem reprezentującym zawodnika (Zawodnik) lub trenera (Trener) a węzłem reprezentującym klub (Klub). Ta relacja opisuje fakt, że zawodnik lub trener ma kontrakt z danym klubem. Posiada atrybuty "od\_kiedy" i "do\_kiedy", które określają daty rozpoczęcia i zakończenia kontraktu. Dodatkowo zawodnik ma wpisaną atrybut „wartość”.

# Opis programu

1. Importowanie potrzebnych bibliotek:

Obraz zawierający tekst, Czcionka, zrzut ekranu

Opis wygenerowany automatycznie

Ten fragment importuje wymagane biblioteki. requests jest używane do wysyłania zapytań HTTP, BeautifulSoup do przetwarzania HTML, a GraphDatabase z pakietu neo4j do komunikacji z bazą danych neo4j.

1. Ustawienie połączenia z bazą danych:

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

Tutaj ustawione są połączenie do bazy danych neo4j za pomocą danych uwierzytelniających. Stworzono również plik 'database.cypher', do którego zapisujesz zapytania cypher. Oraz są utworzone zmienne domain do strony <https://www.transfermarkt.pl>.

1. Funkcje dodające dane do bazy danych:

Obraz zawierający tekst, zrzut ekranu, Czcionka

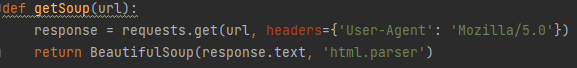
Opis wygenerowany automatycznie

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

Funkcje takie jak „add\_dataCountry”, „add\_dataPlayer”, „add\_dataCouch”, „add\_dataStadium”, „add\_dataClub”, „add\_dataTrophist” są odpowiedzialne za tworzenie odpowiednich węzłów i relacji w bazie danych neo4j. Każda z nich przyjmuje różne parametry związane z krajem, zawodnikiem, trenerem, stadionem, klubem i trofeami, które następnie dodaje do bazy danych.

1. Funkcja pobierająca stronę HTML:



Ta funkcja pobiera zawartość strony HTML dla danego adresu URL. Tworzy ona zapytanie GET do adresu URL, a następnie przekształca odpowiedź na obiekt BeautifulSoup, który umożliwia łatwe przetwarzanie struktury HTML.

1. Funkcje scrapujące dane:

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Czcionka

Opis wygenerowany automatycznie

Obraz zawierający tekst, zrzut ekranu, oprogramowanie

Opis wygenerowany automatycznie

Obraz zawierający tekst, zrzut ekranu, oprogramowanie

Opis wygenerowany automatycznie

Obraz zawierający tekst, zrzut ekranu

Opis wygenerowany automatycznie

Obraz zawierający tekst, zrzut ekranu

Opis wygenerowany automatycznie

Obraz zawierający tekst, zrzut ekranu, oprogramowanie

Opis wygenerowany automatycznie

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Czcionka

Opis wygenerowany automatycznie

Funkcje takie jak „getTrophist”, „getNationality”, „getCouch”, „getPlayers”, „getClubs” są odpowiedzialne za scrapowanie danych z witryny transfermarkt.pl. Każda z nich przyjmuje różne parametry, które są używane do generowania odpowiednich adresów URL, a następnie scrapuje potrzebne dane i zapisuje je do bazy danych.

1. Pętla przetwarzająca dane dla wszystkich lig:

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Czcionka

Opis wygenerowany automatycznie

Ta pętla przetwarza dane dla wszystkich lig, wywołując funkcję getClubs dla każdej z nich.

1. Zamykanie połączenia z bazą danych:

Obraz zawierający tekst, Czcionka, zrzut ekranu, typografia

Opis wygenerowany automatycznie

Na końcu skryptu zamykane jest połączenie z bazą danych i plik z zapytaniami cypher.

# Opis węzłów

Liczba wszystkich węzłów oraz relacji:

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

1. Właściwości węzła Klub:

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

1. Właściwości węzła Kraj:

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

1. Właściwości węzła Stadion:

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

1. Właściwości węzła Trener:

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

1. Właściwości węzła Trofeum:

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

1. Właściwości węzła Zawodnik:

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

# Opis relacji

1. Reakcja kontrakt
   1. Dla trenera

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

* 1. Dla zawodnika

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

1. Reakcja należy

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

1. Reakcja pochodzi

Obraz zawierający tekst, zrzut ekranu, Czcionka, czarne

Opis wygenerowany automatycznie

1. Reakcja wygrali

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

# Nietrywialne zapytania

1. Znajdź kraj, który wygrał Mistrzostwa Świata w sezonie 2022.

Obraz zawierający tekst, zrzut ekranu

Opis wygenerowany automatycznie

1. Znajdź ile pucharów wygrały każdy klub z Anglii.

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

Obraz zawierający zrzut ekranu, tekst, miejsce parkingowe/przestrzeń, astronomia

Opis wygenerowany automatycznie

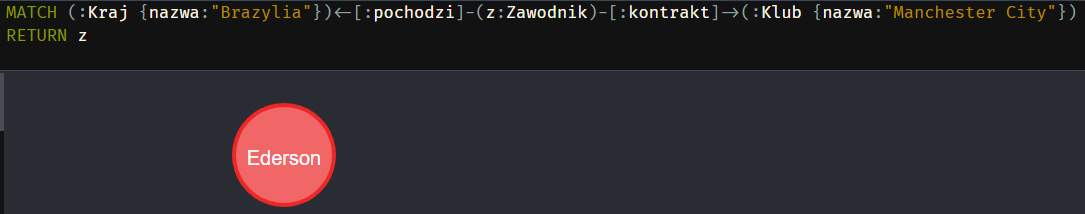
1. Znajdź wszystkich trenerów z Hiszpanii.



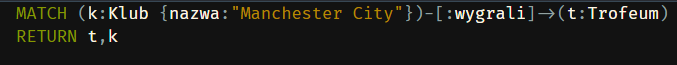
Obraz zawierający zrzut ekranu, Oprogramowanie multimedialne, oprogramowanie, Oprogramowanie graficzne

Opis wygenerowany automatycznie

1. Znajdź wszystkich zawodników z Brazylii grających w klubie Manchester City.



1. Znajdź wszystkie trofea wygrane przez klub Manchester City.



Obraz zawierający zrzut ekranu, tekst, diagram, krąg

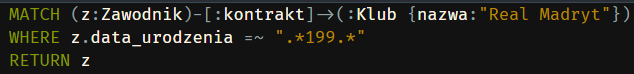
Opis wygenerowany automatycznie

1. Znajdź wszystkie kluby, które wygrały Mistrzostwa Anglii 

Obraz zawierający zrzut ekranu, tekst, Oprogramowanie multimedialne, oprogramowanie

Opis wygenerowany automatycznie

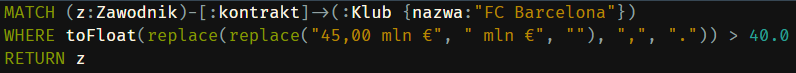
1. Znajdź wszystkich zawodników z Realu Madryt urodzonych między rokiem 1990 a 1999.



Obraz zawierający zrzut ekranu, Oprogramowanie multimedialne, oprogramowanie, tekst

Opis wygenerowany automatycznie

1. Znajdź wszystkich zawodników w klubie FC Barcelony, których wartość rynkowa przekracza 40 milionów euro.



Obraz zawierający zrzut ekranu, Oprogramowanie multimedialne, oprogramowanie

Opis wygenerowany automatycznie

1. Znajdź wszystkich bramkarzy grający w Seria A.



Obraz zawierający zrzut ekranu, Oprogramowanie multimedialne, oprogramowanie, Oprogramowanie graficzne

Opis wygenerowany automatycznie

1. Znajdź wszystkie kluby, które wygrały jakiekolwiek trofeum w sezonie 21/22.



Obraz zawierający zrzut ekranu, Oprogramowanie multimedialne, Oprogramowanie graficzne, oprogramowanie

Opis wygenerowany automatycznie

# Możliwości użycia

Baza może być wykorzystywana w celu wyświetlania danych na stronie wiązaną z tą tematyka. Czerpania informacji o zawodnikach oraz klucha piłkarskich o danymi z nimi powiązanymi. Baza może być świetną bazą dla statystyk.

# Podsumowanie

Podsumowując powyższe informację tą bazę można rozbudować o kolejne informacje takie jak na przykład mecze i dane związane z danym meczem. O kolejne relację tak na przykład przynależności do federacji piłkarskiej.

# Literatura

<https://app.diagrams.net> (29.05.2023r.)

<https://neo4j.com> (29.05.2023r.)