

# Spis użytych technologii

sobota, 18 stycznia 2025 21:01

Do realizacji projektu zostały użyte programy:

- RStudio,
- VS Code,
- OneNote,
- Visual Studio 2022,
- Dbeaver.

A także języki programowania wraz z bibliotekami:

- Python:
  - sqlalchemy,
  - json,
  - random,
  - faker,
  - csv,
  - pandas,
- R:
  - DBI,
  - RMariaDB,
  - dplyr,
  - dbplyr,
  - ggplot2,
  - knitr,
  - lubridate,
  - RColorBrewer.

# Lista plików

sobota, 18 stycznia 2025 21:01

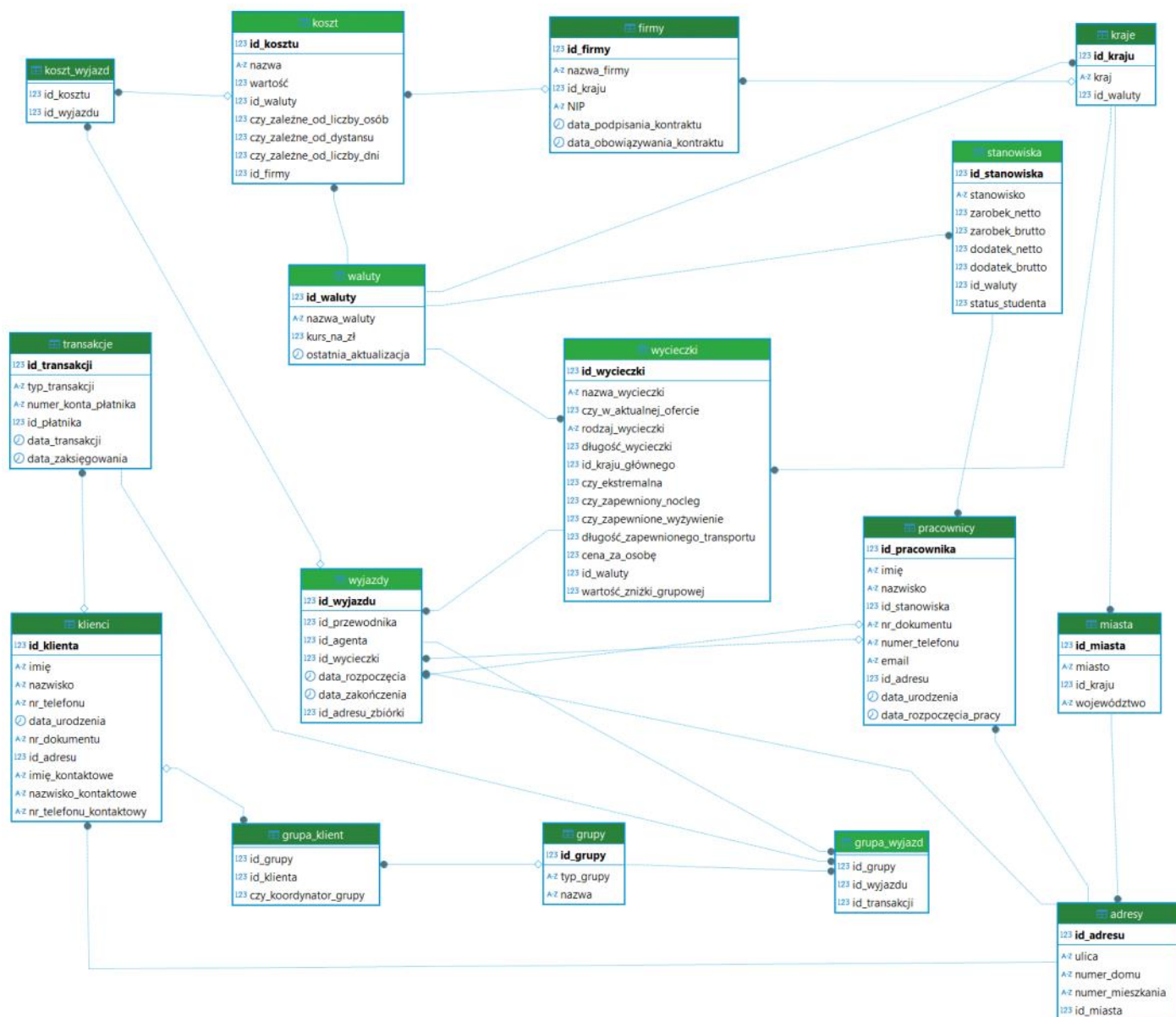
Poniżej znajduje się lista plików użytych do realizacji projektu:

- tworzenie.py
  - Plik zawiera skrypt w języku Python generujący wszystkie puste tabele.
- struktura.json
  - Plik zawiera opisy kolumn napisane w SQL dostosowane do formatu JSON.
- warunki\_bazy.py
  - W pliku znajduje się skrypt języka Python tworzący wyzwalacze dla tabel, dla których ograniczenie "CHECK" nie może zadziałać, a dla których wpisywanie lub edycja danych musi być ograniczona.
- \_\_init\_\_.py
  - Plik wspomagający czytanie funkcji z poszczególnych plików do pliku głównego tworzącego bazę.
- czysta\_baza.py
  - W pliku znajduje się skrypt bazujący na innych plikach, który doprowadza bazę do pustego stanu gotowego na generację danych.
- GENERACJA.py
  - Plik zawierający skrypt generujący dane do poszczególnych tabel opierający się na poniższych plikach:
    - aleja.txt
    - generał\_marszałek.txt
    - imiona.csv
    - kraje\_waluty.csv
    - miasta.txt
    - nazwa\_wycieczki.txt
    - plac\_skwer.txt
    - rodzaj\_wycieczki.txt
- Analiza-grupa6.Rmd
  - Plik dokonujący analizy danych zawartych w bazie oraz tworzący raport na tej podstawie.
- Analiza-grupa6.html
  - Plik zawierający raport z części analitycznej.

# Schemat projektu bazy danych

sobota, 18 stycznia 2025

20:51



# Lista zależności funkcyjnych

sobota, 18 stycznia 2025 18:27

Poniżej znajduje się lista zależności funkcyjnych dla poszczególnych tabel. Nazwy tabel (zmienne relacyjne) są wyszczególnione kolorem fioletowym. W każdym poszczególnym wierszu nienagłówkowym w tabelce znajdują się wypisane atrybuty, a także zbiory atrybutów, z którymi istnieją zależności funkcyjne (tzn. np. dla atrybutu `id_a` i zbioru `{pierwszy, drugi}` istnieją zależności funkcyjne `id_a->pierwszy` oraz `id_a->drugi`). Wszystkie zależności funkcyjne zostały wypisane dla potencjalnej zawartości całej bazy.

- pracownicy:**

Wszystkie={`id_pracownika`, `imię`, `nazwisko`, `id_stanowiska`, `nr_dokumentu`, `numer_telefonu`, `email`, `id_adresu`, `data_urodzenia`, `data_rozpoczęcia_pracy`}

Atrybuty	Zbiory atrybutów, z którymi są w zależności funkcyjnej	Wyjaśnienie
<code>id_pracownika</code>	<code>wszystkie\{id_pracownika}</code>	atrybut <code>id_pracownika</code> jest kluczem głównym (jest unikatowym), zatem jego jednoelementowy zbiór jest nadkluczem
<code>nr_dokumentu</code>	<code>wszystkie\{nr_dokumentu}</code>	atrybut <code>nr_dokumentu</code> jest unikatowy, zatem jego jednoelementowy zbiór jest nadkluczem
<code>numer_telefonu</code>	<code>wszystkie\{numer_telefonu}</code>	atrybut <code>numer_telefonu</code> jest unikatowy, zatem jego jednoelementowy zbiór jest nadkluczem
<code>email</code>	<code>wszystkie\{email}</code>	atrybut <code>email</code> jest unikatowy, zatem jego jednoelementowy zbiór jest nadkluczem

- stanowiska:**

Wszystkie={`id_stanowiska`, `stanowisko`, `zarobek_netto`, `zarobek_brutto`, `id_waluty`, `status_studenta`}

Atrybuty	Zbiory atrybutów, z którymi są w zależności funkcyjnej	Wyjaśnienie
<code>id_stanowiska</code>	<code>wszystkie\{id_stanowiska}</code>	atrybut <code>id_stanowiska</code> jest kluczem głównym (jest unikatowym), zatem jego jednoelementowy zbiór jest nadkluczem

- wycieczki:**

Wszystkie={`id_wycieczki`, `czy_w_aktualnej_ofercie`, `rodzaj_wycieczki`, `id_kraju_głównego`, `czy_ekstremalna`, `czy_zapewniony_nocleg`, `czy_zapewnione_wyżywienie`, `długość_zapewnionego_transportu`, `cena_za_osobę`, `id_waluty`, `wartość_zniżki_grupowej`}

Atrybuty	Zbiory atrybutów, z którymi są w zależności funkcyjnej	Wyjaśnienie
<code>id_wycieczki</code>	<code>wszystkie\{id_wycieczki}</code>	atrybut <code>id_wycieczki</code> jest kluczem głównym (jest unikatowym), zatem jego jednoelementowy zbiór jest nadkluczem

- klienci:**

Wszystkie={`id_klienta`, `imię`, `nazwisko`, `nr_telefonu`, `data_urodzenia`, `nr_dokumentu`, `id_adresu`, `imię_kontaktowe`, `nazwisko_kontaktowe`, `nr_telefonu_kontaktowy`}

Atrybuty	Zbiory atrybutów, z którymi są w zależności funkcyjnej	Wyjaśnienie
<code>id_klienta</code>	<code>wszystkie\{id_klienta}</code>	atrybut <code>id_klienta</code> jest kluczem głównym (jest unikatowym), zatem jego jednoelementowy zbiór jest nadkluczem
<code>nr_dokumentu</code>	<code>wszystkie\{nr_dokumentu}</code>	atrybut <code>nr_dokumentu</code> jest unikatowy, zatem jego jednoelementowy zbiór jest nadkluczem

- grupy:**

Wszystkie={`id_grupy`, `typ_grupy`, `nazwa`}

Atrybuty	Zbiory atrybutów, z którymi są w zależności funkcyjnej	Wyjaśnienie
<code>id_grupy</code>	<code>wszystkie\{id_grupy}</code>	atrybut <code>id_grupy</code> jest kluczem głównym (jest unikatowym), zatem jego jednoelementowy zbiór jest nadkluczem

- grupa\_klient:**

Wszystkie={`id_grupy`, `id_klienta`, `czy_koordinator_grupy`}

Atrybuty	Zbiory atrybutów, z którymi są w zależności funkcyjnej	Wyjaśnienie
----------	--	-------------

Tabela nie posiada zależności funkcyjnych

- grupa\_wyjazd:**

Wszystkie={`id_grupy`, `id_wyjazdu`, `id_transakcji`}

Atrybuty	Zbiory atrybutów, z którymi są w zależności funkcyjnej	Wyjaśnienie
----------	--	-------------

Tabela nie posiada zależności funkcyjnych

- klienci:**

Wszystkie={id\_kraju, kraj}

Atrybuty	Zbiory atrybutów, z którymi są w zależności funkcyjnej	Wyjaśnienie
id_kraju	{kraj}	atrybut id_kraju jest kluczem głównym (jest unikatowym), zatem jego jednoelementowy zbiór jest nadkluczem
kraj	{id_kraju}	atrybut kraj jest unikatowy, zatem jego jednoelementowy zbiór jest nadkluczem

- miasta:**

Wszystkie={id\_miasta, miasto, id\_kraju, województwo}

Atrybuty	Zbiory atrybutów, z którymi są w zależności funkcyjnej	Wyjaśnienie
id_miasta	wszystkie\{id_miasta}	atrybut id_miasta jest kluczem głównym (jest unikatowym), zatem jego jednoelementowy zbiór jest nadkluczem

- adresy:**

Wszystkie={id\_adresu, ulica, numer\_domu, numer\_mieszkania, id\_miasta}

Atrybuty	Zbiory atrybutów, z którymi są w zależności funkcyjnej	Wyjaśnienie
id_adresu	wszystkie\{id_adresu}	atrybut id_adresu jest kluczem głównym (jest unikatowym), zatem jego jednoelementowy zbiór jest nadkluczem

- firmy:**

Wszystkie={id\_firmy, nazwa\_firmy, id\_kraju, NIP, data\_podpisania\_kontraktu, data\_obowiązywania\_kontraktu}

Atrybuty	Zbiory atrybutów, z którymi są w zależności funkcyjnej	Wyjaśnienie
id_firmy	wszystkie\{id_firmy}	atrybut id_firmy jest kluczem głównym (jest unikatowym), zatem jego jednoelementowy zbiór jest nadkluczem

- koszt:**

Wszystkie={id\_kosztu, nazwa, wartość, id\_waluty, czy\_zależne\_od\_liczby\_osób, czy\_zależne\_od\_dystansu, czy\_zależne\_od\_liczby\_dni, id\_firmy}

Atrybuty	Zbiory atrybutów, z którymi są w zależności funkcyjnej	Wyjaśnienie
id_kosztu	wszystkie\{id_kosztu}	atrybut id_kosztu jest kluczem głównym (jest unikatowym), zatem jego jednoelementowy zbiór jest nadkluczem

- koszt\_wyjazd:**

Wszystkie={id\_kosztu, id\_wyjazdu}

Atrybuty	Zbiory atrybutów, z którymi są w zależności funkcyjnej	Wyjaśnienie
----------	--	-------------

Tabela nie posiada zależności funkcyjnych

- transakcje:**

Wszystkie={id\_transakcji, typ\_transakcji, numer\_konta\_płatnika, id\_płatnika, data\_transakcji, data\_zaksięgowania}

Atrybuty	Zbiory atrybutów, z którymi są w zależności funkcyjnej	Wyjaśnienie
id_transakcji	wszystkie\{id_transakcji}	atrybut id_transakcji jest kluczem głównym (jest unikatowym), zatem jego jednoelementowy zbiór jest nadkluczem

- wyjazdy:**

Wszystkie={id\_wyjazdu, id\_przewodnika, id\_agenta, id\_wycieczki, data\_rozpoczęcia, data\_zakończenia, id\_adresu\_zbiórki}

Atrybuty	Zbiory atrybutów, z którymi są w zależności funkcyjnej	Wyjaśnienie
id_wyjazdu	wszystkie\{id_wyjazdu}	atrybut id_wyjazdu jest kluczem głównym (jest unikatowym), zatem jego jednoelementowy zbiór jest nadkluczem

- waluty:**

Wszystkie={id\_waluty, nazwa\_waluty, kurs\_na\_zł, ostatnia\_aktualizacja}

Atrybuty	Zbiory atrybutów, z którymi są w zależności funkcyjnej	Wyjaśnienie
id_waluty	wszystkie\{id_waluty}	atrybut id_waluty jest kluczem głównym (jest unikatowym),

		zatem jego jednoelementowy zbiór jest nadkluczem
<b>nazwa_waluty</b>	wszystkie\{nazwa_waluty}	atrybut nazwa_waluty jest unikatowy, zatem jego jednoelementowy zbiór jest nadkluczem

Ponieważ wszystkie zależności w każdej tabeli zaczynają się od nadkluczy, to każda tabela jest EKNF, co oznacza, że cała baza również jest EKNF.

# Najtrudniejsze elementy projektu

sobota, 18 stycznia 2025 20:52

Naszym zdaniem najtrudniejszymi w realizacji projektu było:

- zaprojektowanie bazy tak, aby spełniała wszystkie wymagania (rozległość bazy utrudniała znormalizowanie jej, a także wydłużała czas testów),
- stworzenie programu generującego dane, tak aby zachował on przy tym spójność między tabelami, np. żeby ten sam wyjazd odbywał się na terenie tego samego kraju, w jakim obowiązuje dana waluta,
- podczas analizy dopasowanie rozmiaru wykresu do danych oraz wygenerowanie w pełni zautomatyzowanych odpowiedzi.