

Dokumentowe bazy danych – MongoDB

Ćwiczenie 1 - do wykonania podczas zajęć

Imię i nazwisko: Gracjan Filipek

Materiały:

Książki

Np.

- Shannon Bradshaw, Eoin Brazil, Kristina Chodorow, MongoDB: The Definitive Guide. Powerful and Scalable Data Storage, O'Reilly 2019
- Alex Giamas, Mastering MongoDB 4.x., Pact 2019

Dokumentacja

- <https://www.mongodb.com/docs/manual/reference/program/mongo/>

MongoDB University Courses

- <https://university.mongodb.com/courses/catalog>
- MongoDB Basics
 - <https://university.mongodb.com/courses/M001/about>
- The MongoDB Aggregation Framework
 - <https://university.mongodb.com/courses/M121/about>
- Data Modeling
 - <https://university.mongodb.com/courses/M320/about>

Zbiory danych

The MongoDB Atlas Sample Datasets

<https://www.mongodb.com/developer/article/atlas-sample-datasets/>

Yelp Dataset

www.yelp.com - serwis społecznościowy – informacje o miejscach/lokalach

- restauracje, kluby, hotele itd. (*businesses*),
- użytkownicy piszą recenzje (*reviews*) o miejscach i wystawiają oceny
- użytkownicy odwiedzają te miejsca - "meldują się" (*check-in*)
- Przykładowy zbiór danych zawiera dane z 5 miast: Phoenix, Las Vegas, Madison, Waterloo i Edinburgh.

Kolekcje:

- **42,153** businesses
- **320,002** business attributes
- **31,617** check-in sets
- **252,898** users
- **955,999** edge social graph
- **403,210** tips
- **1,125,458** reviews

business

```
{
  'type': 'business',
  'business_id': (encrypted business id),
  'name': (business name),
  'neighborhoods': [(hood names)],
  'full_address': (localized address),
  'city': (city),
  'state': (state),
  'latitude': latitude,
  'longitude': longitude,
  'stars': (star rating, rounded to half-stars),
  'review_count': review count,
  'categories': [(localized category names)]
  'open': True / False (corresponds to closed, not business hours),
  'hours': {
    (day_of_week): {
      'open': (HH:MM),
      'close': (HH:MM)
    },
    ...
  },
  'attributes': {
    (attribute_name): (attribute_value),
    ...
  },
}
```

review

```
{
  'type': 'review',
  'business_id': (encrypted business id),
  'user_id': (encrypted user id),
  'stars': (star rating, rounded to half-stars),
  'text': (review text),
  'date': (date, formatted like '2012-03-14'),
  'votes': {(vote type): (count)},
}
```

user

```
{
  'type': 'user',
  'user_id': (encrypted user id),
  'name': (first name),
  'review_count': (review count),
  'average_stars': (floating point average, like 4.31),
  'votes': {(vote type): (count)},
  'friends': [(friend user_ids)],
  'elite': [(years_elite)],
  'yelping_since': (date, formatted like '2012-03'),
  'compliments': {
    (compliment_type): (num_compliments_of_this_type),
    ...
  },
  'fans': (num_fans),
}
```

check-in

```
{
  'type': 'checkin',
  'business_id': (encrypted business id),
  'checkin_info': {
    '0-0': (number of checkins from 00:00 to 01:00 on all Sundays),
    '1-0': (number of checkins from 01:00 to 02:00 on all Sundays),
    ...
    '14-4': (number of checkins from 14:00 to 15:00 on all Thursdays),
    ...
    '23-6': (number of checkins from 23:00 to 00:00 on all Saturdays)
  }, # if there was no checkin for a hour-day block it will not be in the dict
}
```

tip

```
{
  'type': 'tip',
  'text': (tip text),
  'business_id': (encrypted business id),
  'user_id': (encrypted user id),
  'date': (date, formatted like '2012-03-14'),
  'likes': (count),
}
```

Narzędzia

- Konsola: mongosh, MongoDB Database Tools
- Studio3T
- MongoDB Compass
- DataGrip
- VS Code

Zadania

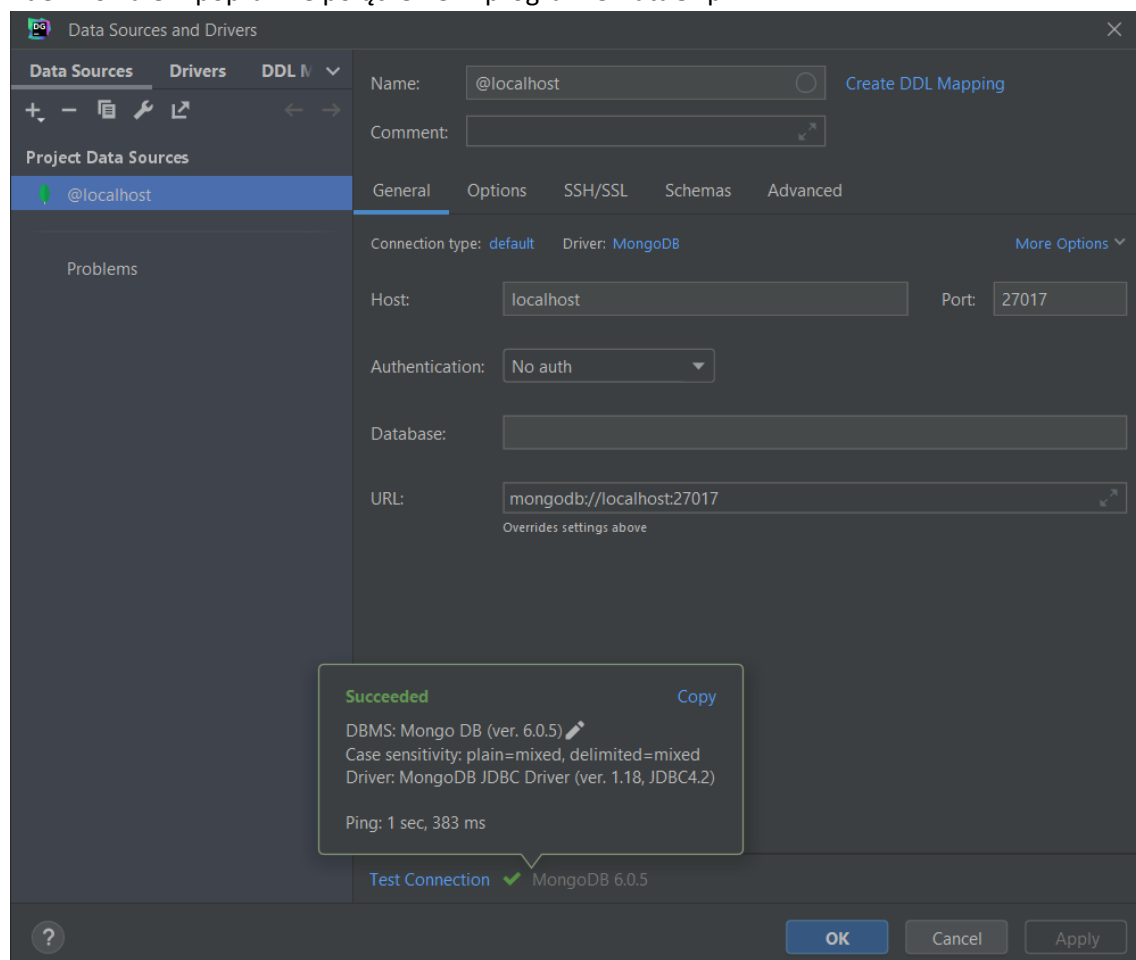
1. Połącz się serwerem MongoDB

- Można skorzystać z własnego/lokanego serwera MongoDB
- Można skorzystać własny klaster/bazę danych w serwisie MongoDB Atlas
 - <https://www.mongodb.com/atlas/database>
 - Serwis umożliwia założenie darmowego konta
- Połącz za pomocą konsoli mongosh
- Ewentualnie zdefiniuj połączenie w wybranym przez siebie narzędziu

.... Wyniki, zrzuty ekranów, kod, komentarz

Korzystam z własnego lokalnego serwera

Zdefiniowałem poprawne połączenie w programie DataGrip:

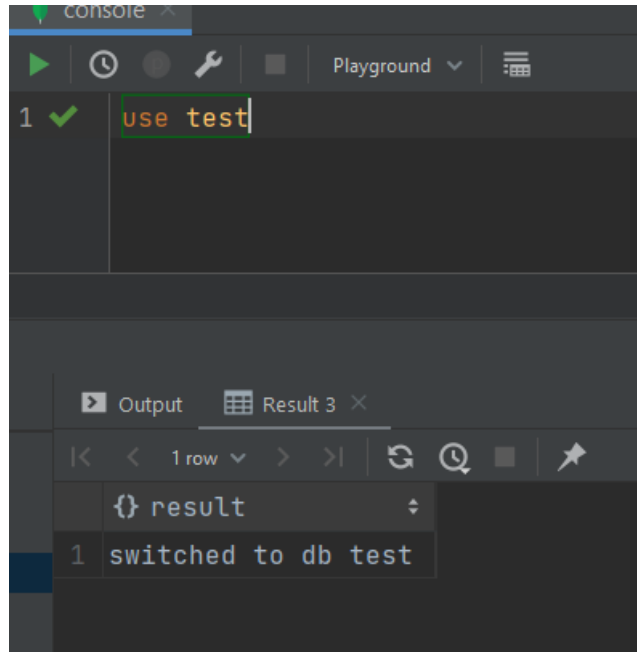


2. Stwórz bazę danych/kolekcję/dokument

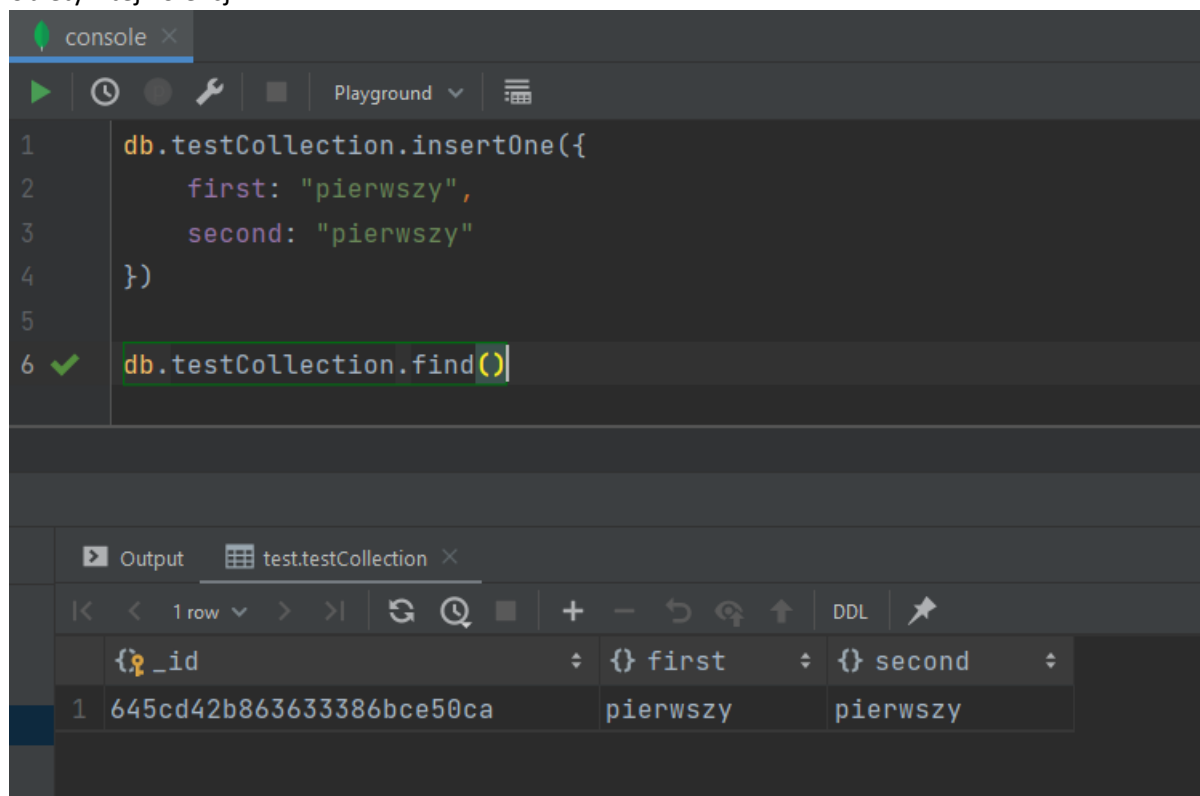
- może to być dowolna kolekcja, dowolny dokument – o dowolnej strukturze
- chodzi o przetestowanie działania połączenia

.... Wyniki, zrzuty ekranów, kod, komentarz

Stworzyłem bazę „test”



Następnie dodałem przykładowy obiekt do pustej kolekcji „testCollection” i wypisałem wszystkie obiekty z tej kolekcji



Jak widać wszystko działa poprawnie

3. Zaimportuj przykładowe zbory danych

- MongoDB Atlas Sample Dataset
 - <https://docs.atlas.mongodb.com/sample-data>
 - w przypadku importu z lokalnych plików można wykorzystać polecenie mongorestore
 - mongorestore <data-dump-folder>
 - np.

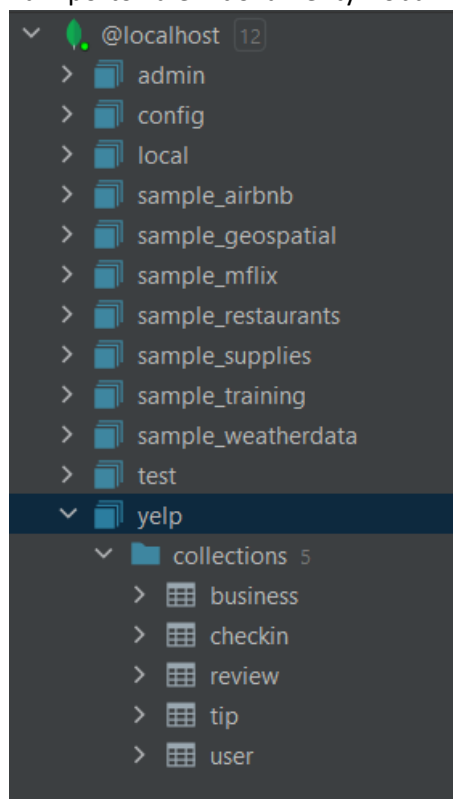
```
mongorestore samples
```

 - Oczywiście, w przypadku łączenia się zdalnym serwerem należy podać parametry połączenia oraz dane logowania
- Yelp Dataset
 - wykorzystaj komendę mongoimport
 - mongoimport --db <db-name> --collection <coll-name> --type json --file <file>
 - np.

```
mongoimport --db yelp --collection business --type json --file ./yelp_academic_dataset_business.json
```
 - można też wykorzystać np. narzędzie MongoDB Compass

.... Wyniki, zrzuty ekranów, kod, komentarz

Zaimportowałem dokumenty z obu źródeł



4. Zapoznaj się ze strukturą przykładowych zbiorów danych/kolekcji

- W bazach danych: MongoDB Atlas Sample Dataset
 - Skomentuj struktury użyte w dokumentach dla dwóch wybranych zbiorów (takich które wydają ci się najciekawsze)

- Np. Sample Analytics Dataset i Sample Training Dataset
- W bazie Yelp
 - Skomentuj struktury użyte w dokumentach bazy Yelp

.... Wyniki, zrzuty ekranów, kod, komentarz

5. Operacje CRUD

- <https://www.mongodb.com/docs/manual/crud/>
- stwórz nową bazę danych, jako nazwy bazy danych użyj swoich inicjałów
- stwórz kolekcję "student"
 - informacje o studentach, przedmiotach ocenach z przedmiotów itp.
 - zaproponuj strukturę dokumentu
 - wykorzystaj typy proste/podstawowe, dokumenty zagnieżdżone, tablice itp.
- wprowadź kilka przykładowych dokumentów
 - przetestuj operacje wstawiania, modyfikacji/usuwania dokumentów
- przetestuj operacje wyszukiwania dokumentów

.... Wyniki, zrzuty ekranów, kod, komentarz

W sprawozdaniu należy umieścić zrzuty ekranów (z kodem poleceń oraz z uzyskanymi wynikami). Dodatkowo należy dołączyć plik tekstowy (najlepiej z rozszerzeniem .js) zawierający kod poleceń

Punktacja za zadanie (razem 1pkt)