

Dokumentowe bazy danych – MongoDB

Ćwiczenie 1 - do wykonania podczas zajęć

Imię i nazwisko:

Adrian Żerebiec

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

```
C:\Users\adria>mongosh
Current Mongosh Log ID: 647747b601cd86b3ef1a32fc
Connecting to:   mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+1.9.1
Using MongoDB:   6.0.6
Using Mongosh:   1.9.1

For mongosh info see: https://docs.mongodb.com/mongodb-shell/

-----
The server generated these startup warnings when booting
2023-05-31T11:50:37.831+02:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
-----

test> use TEST
switched to db TEST
TEST> show
MongoshInvalidInputError: [COMMON-10001] 'undefined' is not a valid argument for "show".
TEST> show tables
DANE
TEST>
(To exit, press Ctrl+C again or Ctrl+D or type .exit)
TEST>

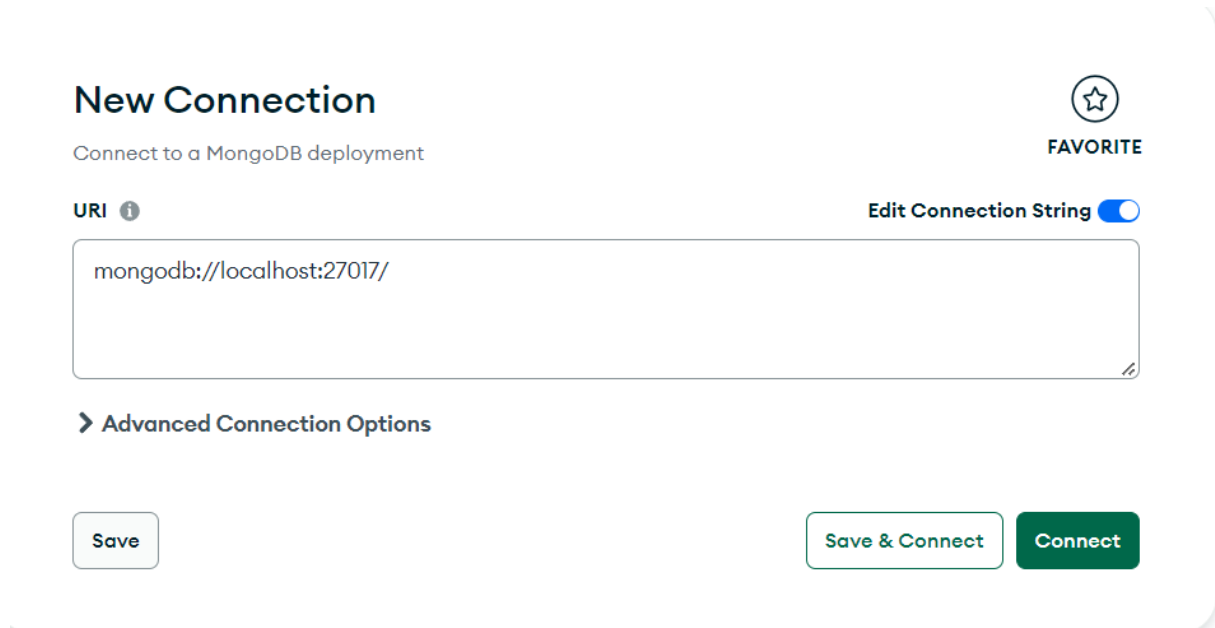
C:\Users\adria>mongosh
Current Mongosh Log ID: 647748a6d7fc4247a3e9c4b6
Connecting to:   mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+1.9.1
Using MongoDB:   6.0.6
Using Mongosh:   1.9.1

For mongosh info see: https://docs.mongodb.com/mongodb-shell/

-----
The server generated these startup warnings when booting
2023-05-31T11:50:37.831+02:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
-----

test> use TEST
switched to db TEST
TEST> show tables
DANE
TEST>
```

Połączenie odbyło się z pomocą programu MongoDBCompass



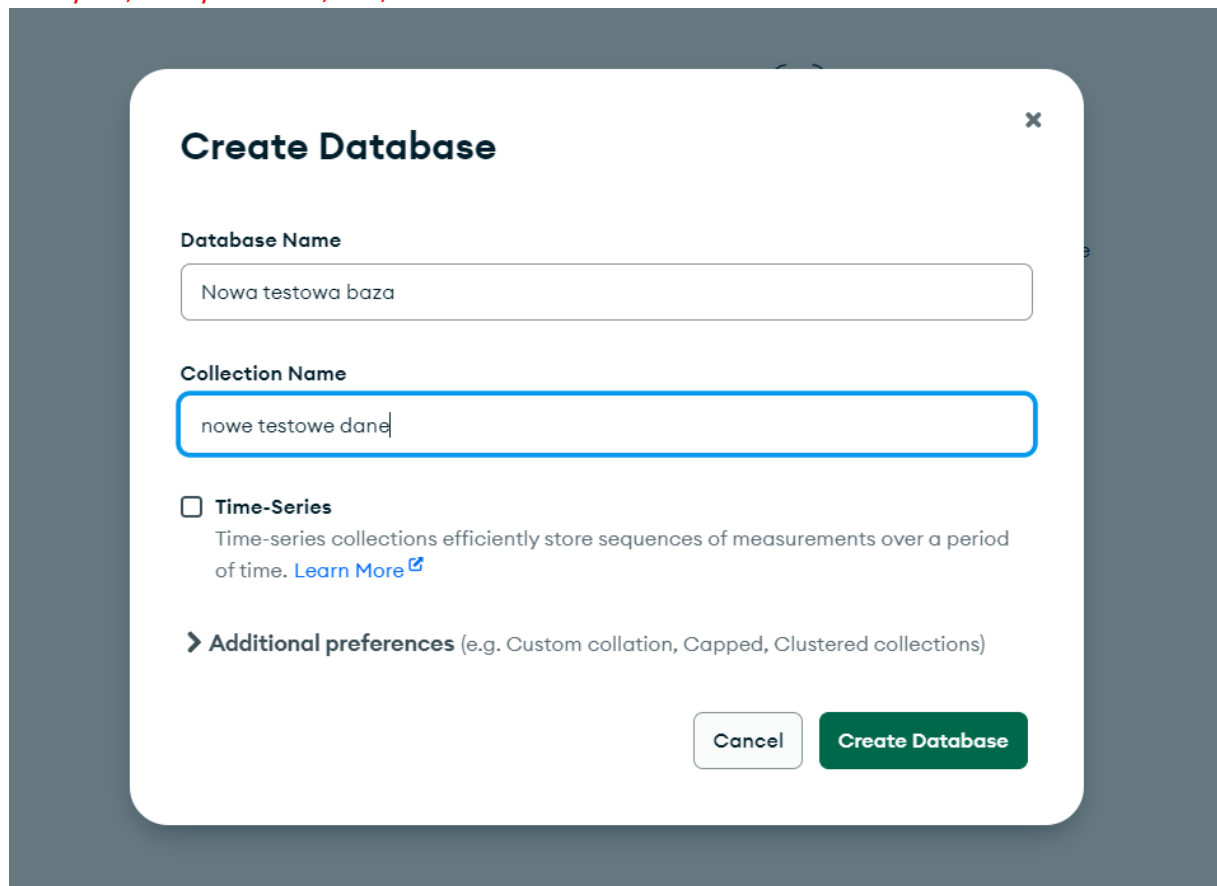
The screenshot shows the 'New Connection' window in MongoDB Compass. At the top, it says 'New Connection' and 'Connect to a MongoDB deployment'. There is a 'FAVORITE' toggle with a star icon. Below this, the 'URI' field is populated with 'mongodb://localhost:27017/'. To the right of the URI field is a toggle for 'Edit Connection String'. Below the URI field is a link for 'Advanced Connection Options'. At the bottom, there are three buttons: 'Save', 'Save & Connect', and 'Connect'.

Po kliknięciu connect przechodzimy do naszego lokalnego hosta

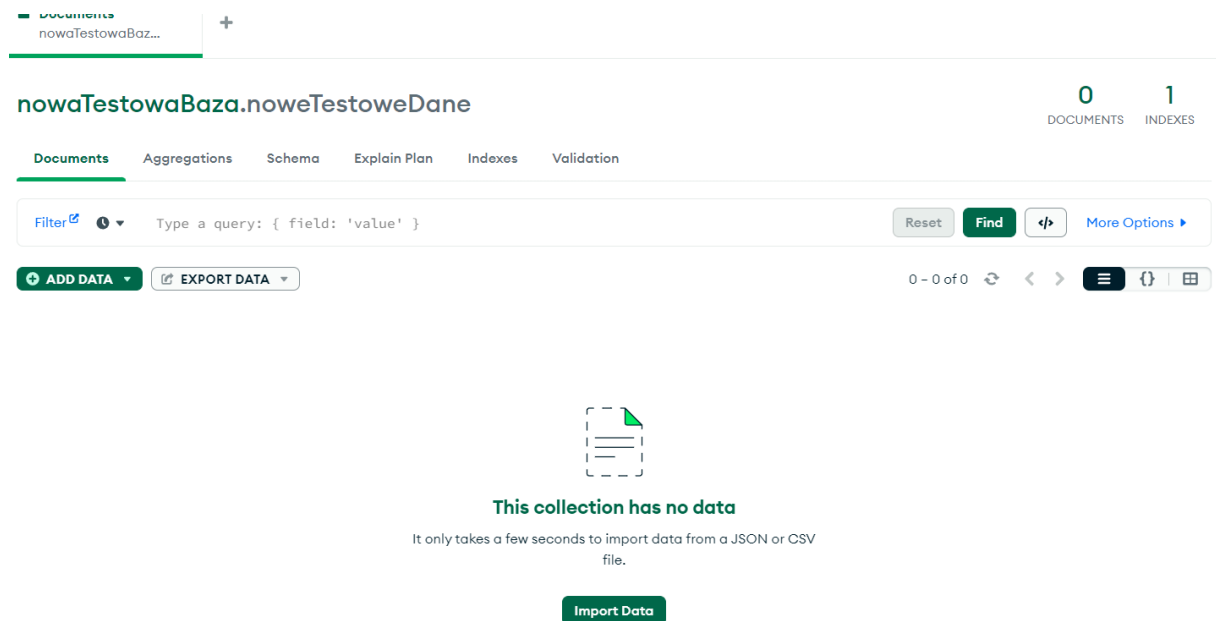
2. Stwórz bazę danych/kolekcję/dokument

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

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



The screenshot shows the 'Create Database' dialog box. It has a title bar with a close button. The 'Database Name' field contains 'Nowa testowa baza'. The 'Collection Name' field contains 'nowe testowe dane'. Below these fields is a checkbox for 'Time-Series' which is unchecked. Below the checkbox is a description: 'Time-series collections efficiently store sequences of measurements over a period of time. [Learn More](#)'. At the bottom, there is a link for 'Additional preferences (e.g. Custom collation, Capped, Clustered collections)'. At the bottom right, there are two buttons: 'Cancel' and 'Create Database'.



```
C:\Users\adria>mongosh
Current Mongosh Log ID: 64774b189a0ed539298f513c
Connecting to:  mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+1.9.1
Using MongoDB:  6.0.6
Using Mongosh:  1.9.1

For mongosh info see: https://docs.mongodb.com/mongosh-shell/

-----
The server generated these startup warnings when booting
  2023-05-31T11:50:37.831+02:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
-----

test> use nowaTestowaBaza
switched to db nowaTestowaBaza
nowaTestowaBaza> show tables
nowaTestoweDane
nowaTestoweDane2
nowaTestowaBaza> 
```

Widzimy, że tabele się dodały i połączenie działa

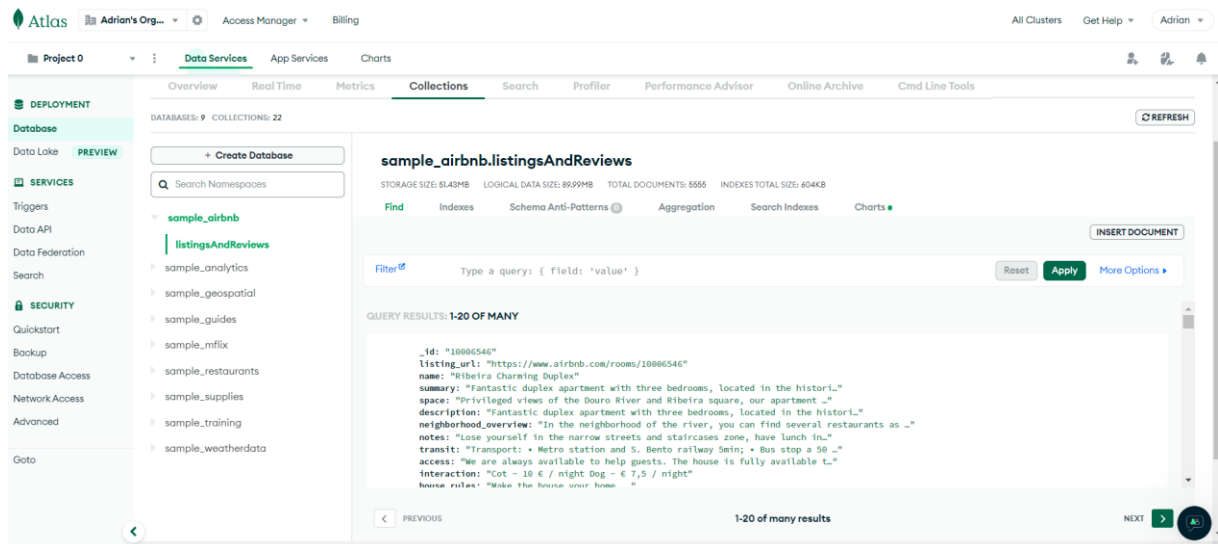
3. Zaimportuj przykładowe zbory danych

- 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

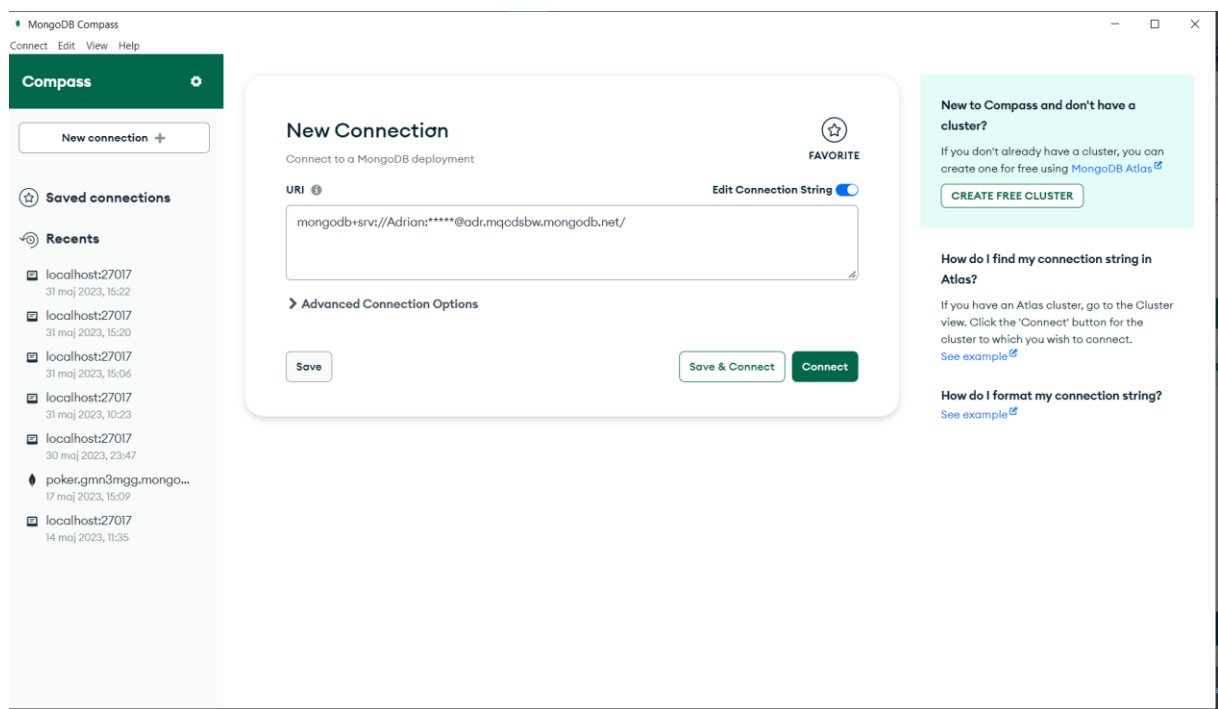
- można też wykorzystać np. narzędzie MongoDB Compass

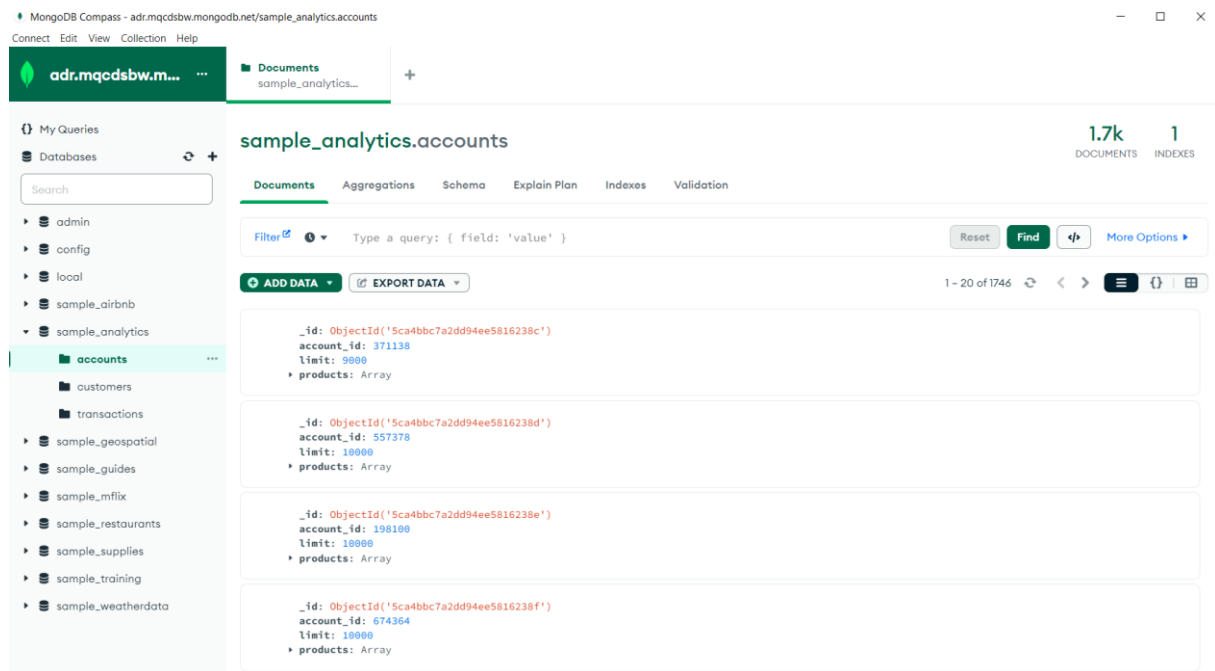
.... Wyniki, zrzuty ekranów, kod, komentarz

W celu wykonania ćwiczenia korzystam z MongoDB Atlas aby tam utworzyć bazę i dodać przykładowe dane. Jak widzimy dane są już wprowadzone, dla potwierdzenia sprawdziłem jeszcze połączenie w MongoDB Compass



Z połączeniem nie było problemów, wystarczy skorzystać z odpowiedniego URL





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

W Sample Analytics Dataset widzimy 3 różne kolekcje, wypełnione różnorodnymi danymi. Wykorzystywane są do tego zwykłe pola, tablice oraz obiekty. W transactions widzimy także spore zagnieżdżenie danych gdyż znajdują się one w tablicy, następnie mamy obiekty w których dopiero mamy dane.

W Sample Training Dataset podobnie występuje zagnieżdżenie obiektów w tablicach. Dobrze widać to w przypadku zdjęć w kolekcji companies gdzie mamy obiekt w którym jest zagnieżdżona tablica a w niej kolejna tablica.

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 celu stworzenia bazy korzystam z MongoDB Compass

Learn More'. At the bottom, there is a section for 'Additional preferences' with a chevron icon and a description '(e.g. Custom collation, Capped, Clustered collections)'. At the very bottom are two buttons: 'Cancel' and 'Create Database'." data-bbox="114 151 881 577"/>

Create Database

Database Name

AZ

Collection Name

student

☐ **Time-Series**

Time-series collections efficiently store sequences of measurements over a period of time. [Learn More](#)

➤ **Additional preferences** (e.g. Custom collation, Capped, Clustered collections)

Cancel Create Database

Z pomocą MongoDb Playgroud dodałem dane

```
/* global use, db */
// MongoDB Playground
// To disable this template go to Settings | MongoDB | Use Default Template
For Playground.
// Make sure you are connected to enable completions and to be able to run a
playground.
// Use Ctrl+Space inside a snippet or a string literal to trigger completions.
// The result of the last command run in a playground is shown on the results
panel.
// By default the first 20 documents will be returned with a cursor.
// Use 'console.log()' to print to the debug output.
// For more documentation on playgrounds please refer to
// https://www.mongodb.com/docs/mongodb-vscode/playgrounds/

// Select the database to use.
use('AZ');
```

```
// Insert a few documents into the sales collection.
db.getCollection('students').insertMany([
  {
    name: "Marcin Głowacz",
    address: "Polna 37",
    faculty: "IET",
    index: "412345",
    grades: [5, 5, 5, 5, 5],
    subjects: ["Bazy danych", "Programowanie", "Sieci komputerowe", "Systemy
operacyjne", "Wprowadzenie do informatyki"]
  },
  {
    name: "Marcin Głowacki",
    address: "Leśna 37",
    faculty: "MS",
    index: "412125",
    grades: [5, 3, 5, 4, 2],
    subjects: ["Algebra", "Matematyka dyskretna", "Sieci komputerowe",
"Analiza matematyczna", "Wprowadzenie do informatyki"]
  }
]);

// Run a find command to view items sold on April 4th, 2014.

// Print a message to the output window.

// Here we run an aggregation and open a cursor to the results.
// Use '.toArray()' to exhaust the cursor to return the whole result set.
// You can use '.hasNext().next()' to iterate through the cursor page by page.
db.getCollection('AZ').aggregate([
]);
```

The screenshot shows the MongoDB Compass interface for the 'AZ.students' collection. The top bar indicates 0 documents and 1 index. The 'Documents' tab is active. A search bar at the top right contains the query '{ field: 'value' }'. Below the search bar, there are buttons for 'ADD DATA' and 'EXPORT DATA'. The main area displays two documents:

```
{
  "_id": ObjectId('6477591399d8914ec97b5553'),
  "name": "Marcin Głowacz",
  "address": "Polna 37",
  "faculty": "IET",
  "index": "412345",
  "grades": Array,
  "subjects": Array
}
```

```
{
  "_id": ObjectId('6477591399d8914ec97b5554'),
  "name": "Marcin Głowacki",
  "address": "Leśna 37",
  "faculty": "MS",
  "index": "412125",
  "grades": Array
    0: 5
    1: 3
    2: 5
    3: 4
    4: 2
  "subjects": Array
}
```

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)