# MongoDB Homework

## Instructions / Hints (how to import the dataset)

* You can use the mongoimport linux-command to import it [Note: it’s not a mongo-command, you have to type into the linux terminal, and **not** to the mongo shell]

## Dataset #1 - Movies

Contains basic information about movie titles.

#### Sample

{ "\_id" : "tt0118799", "runtimeMinutes" : 116, "title" : "Life Is Beautiful", "year" : 1997, "genres" : [ "Comedy", "Drama", "Romance" ] }  
{ "\_id" : "tt0209144", "runtimeMinutes" : 113, "title" : "Memento", "year" : 2000, "genres" : [ "Mystery", "Thriller" ] }  
{ "\_id" : "tt0245429", "runtimeMinutes" : 125, "title" : "Spirited Away", "year" : 2001, "genres" : [ "Adventure", "Animation", "Family" ] }  
{ "\_id" : "tt0468569", "runtimeMinutes" : 152, "title" : "The Dark Knight", "year" : 2008, "genres" : [ "Action", "Crime", "Drama" ] }  
{ "\_id" : "tt1675434", "runtimeMinutes" : 112, "title" : "The Intouchables", "year" : 2011, "genres" : [ "Biography", "Comedy", "Drama" ] }

#### Fields

* \_id – alphanumeric unique identifier
* title – title of the movie
* year – release year of a title
* runtimeMinutes – runtime of the title, in minutes
* genres – includes up to three genres associated with the title

#### Import

mongoimport --db test --collection movies --file movies.json

## Datase #2 - Crew

Writers and directors associated with the movie titles.

#### Sample

#### { "\_id" : "nm0000005", "name" : "Ingmar Bergman", "birthYear" : 1918, "deathYear" : 2007, "directed" : [ "tt0038468", "tt0038675", … ], "written" : [ "tt0036914", "tt0038468", … ] }

#### { "\_id" : "nm0000033", "name" : "Alfred Hitchcock", "birthYear" : 1899, "deathYear" : 1980, "directed" : [ "tt0016127", "tt0016230", … ], "written" : [ "tt0015493", "tt0015508", … ] }

#### { "\_id" : "nm0000040", "name" : "Stanley Kubrick", "birthYear" : 1928, "deathYear" : 1999, "directed" : [ "tt0045758", "tt0048254", … ], "written" : [ "tt0048254", "tt0049406", … ] }

#### { "\_id" : "nm0000041", "name" : "Akira Kurosawa", "birthYear" : 1910, "deathYear" : 1998, "directed" : [ "tt0034333", "tt0036400", … ], "written" : [ "tt0034333", "tt0036400", … ] }

#### { "\_id" : "nm0000217", "name" : "Martin Scorsese", "birthYear" : 1942, "directed" : [ "tt0063803", "tt0066416", … ], "written" : [ "tt0063803", "tt0064748", … ] }

#### Fields

* \_id – alphanumeric unique identifier
* name – full name of the person
* birthYear – the person’s year of birth in YYYY format. May be missing if unknown or undisclosed.
* deathYear – the person’s year of death in YYYY format, if applicable. Missing if the person is alive or if the date is unknown or undisclosed.
* directed – identifiers of all movies where this person is credited as director. May be empty.
* written – identifiers of all movies where this person is credited as writer. May be empty.

#### Import

mongoimport --db test --collection crew --file crew.json

## Exercises:

1. Insert a new document into the crew collection, which contains your name and birth year in the name and birthYear fields. Write a query to show the existence of this record!
2. List all movies that were released after 2000
3. List all the family dramas (where genres includes both “Family” and “Drama”)
4. Get the number of all comedies that were released this year (2021).
5. Find the title of the longest movie.
6. List the names of “one-shot” directors (people with a single director credit to their name). List only the names and no other fields!
7. Find the list of writer-directors, i.e. people who have at least one credit in each category (not necessarily on the same movie).
8. Find the list of people who contributed to title tt0071853 (Monty Python and the Holy Grail) either as writer or director. List each person only once even if they have credits in both categories.
9. Mark people where date of birth or date of death is supposedly missing from the database. Find people who have either no birthDate or have no deathDate and were born before 1900 and add a new Boolean field missingDate with value true to their documents.  
   Do not modify the documents of other people.

## Extra Exercises (For extra points, not mandatory):

1. Find movies that were codirections, i.e. have more than one directors. It is enough to list their ids.
2. Show the list of movies’ titles and their director(s)’ name.

## **What to send back?**

* Document all the steps that you accomplish. Create screenshots of your work and put them into the documentation.
* Write into the document all the commands that you run and the screenshots of the query results!