

## Project #2 – MongoDB / Document Stores

### Due Date: Sunday, May 18<sup>th</sup>, 2025

#### 1. Using the load() function inside the mongo shell, load the prep.js file

This will create a *students* collection in whatever database you are currently using. Be advised that this script contains a random component so that you will each get a slightly different (but same structure) version of the *students* collection. Here is an example of an object that could be from the *students* collection:

```
db.students.findOne()

{
  "_id" : ObjectId("558d08925e083d8cdd7be831"),
  "home_city" : "Kalamata",
  "first_name" : "Eirini",
  "hobbies" : [
    "skydiving",
    "guitar",
    "AD&D"
  ],
  "favourite_os" : "OS X",
  "laptop_cost" : 1506,
  "courses" : [
    {
      "course_code" : "P102",
      "course_title" : "Introduction to R",
      "course_status" : "Complete",
      "grade" : 10
    },
    {
      "course_code" : "S102",
      "course_title" : "Mathematical Statistics",
      "course_status" : "In Progress"
    },
    {

```

```

        "course_code" : "P201",
        "course_title" : "Advanced R",
        "course_status" : "In Progress"
    },
    {
        "course_code" : "S202",
        "course_title" : "Graph Theory",
        "course_status" : "Complete",
        "grade" : 7
    },
    {
        "course_code" : "M102",
        "course_title" : "Data Mining",
        "course_status" : "In Progress"
    }
]

```

This is a collection that tracks the performance of different students enrolled in classes

**2. For each of the following write a SINGLE command that will produce the desired result.**

Either use `db.students.find()` function for the simpler questions, or use `db.students.aggregate()` function (i.e. the aggregation pipeline). You can consult the excellent online documentation, in particular for the aggregation pipeline (<http://docs.mongodb.org/manual/core/aggregation-pipeline/>).

- How many students in your database are currently taking at least a class (i.e. have a *course\_status* of “In Progress”)?
- Produce a grouping of the documents that contains the name of each home city and the number of students enrolled from that home city.
- Which hobby or hobbies are the most popular?
- What is the GPA (ignoring dropped classes and in progress classes) of the best student?
- Which student has the largest number of grade 10's?
- Which class has the highest average GPA?

- Which class has been dropped the most number of times?
- Produce of a count of classes that have been COMPLETED by class type. The class type is found by taking the first letter of the course code so that M102 has type M. So I basically want how many courses have been completed in type M, how many of type S, how many of type P etc... (HINT: check out the \$substr function here: <http://docs.mongodb.org/manual/reference/operator/aggregation/substr/>)
- Produce a transformation of the documents so that the documents now have an additional boolean field called “hobbyist” that is true when the student has more than 3 hobbies and false otherwise.
- Produce a transformation of the documents so that the documents now have an additional field that contains the number of classes that the student has completed
- Produce a transformation of the documents in the collection so that they look like this:

```
{
  "_id" :
  ObjectId("558d08925e083d8cdd7be831"),
  "first_name" : "Eirini",
  "GPA" : 8.5
  "classesInProgress" : 3
  "droppedClasses" : 0
}
```

The GPA is the average grade of all the completed classes. The other two computed fields are the number of classes currently in progress and the number of classes dropped. No other fields should be in there. No other fields should be present.