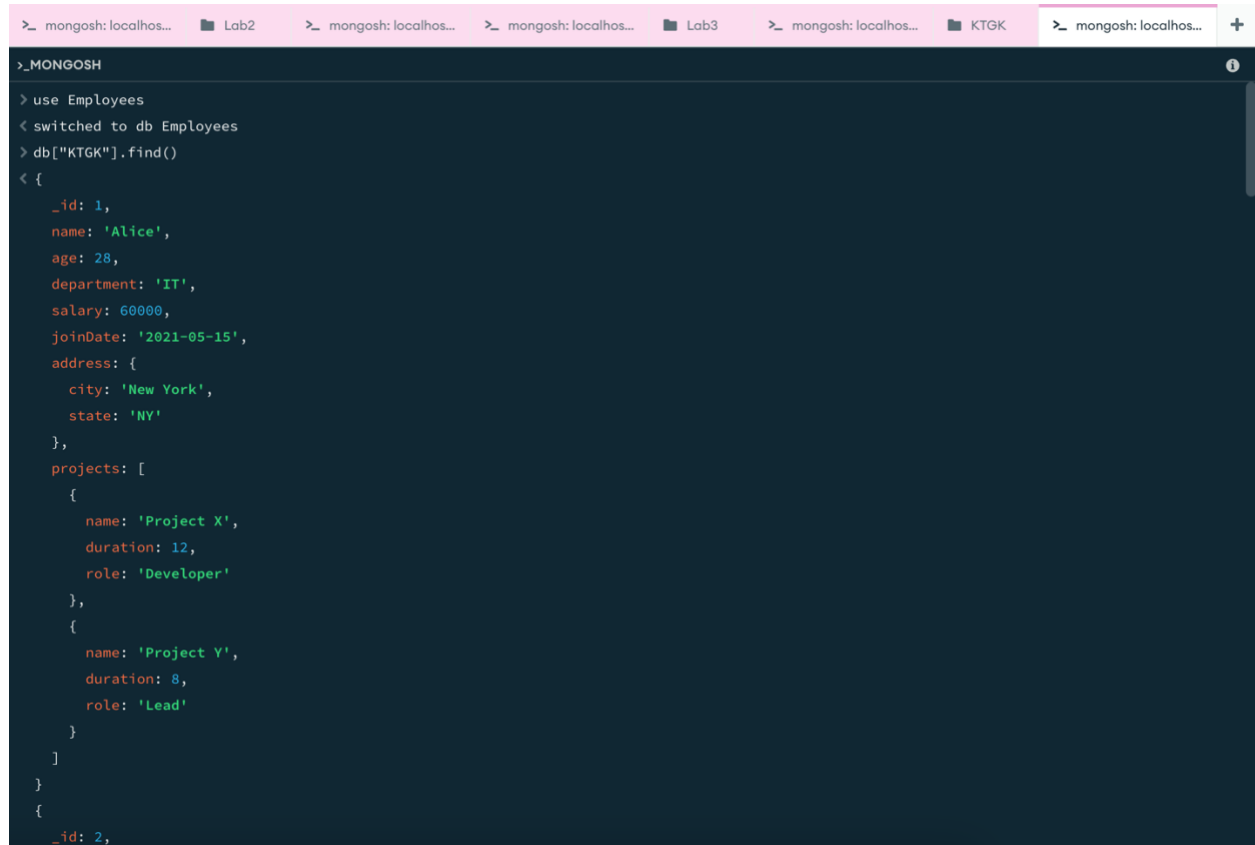


Họ và tên: Trần Tấn Phát  
MSSV: 2274802010644

KTGK: Nhập môn Phân tích Dữ liệu lớn

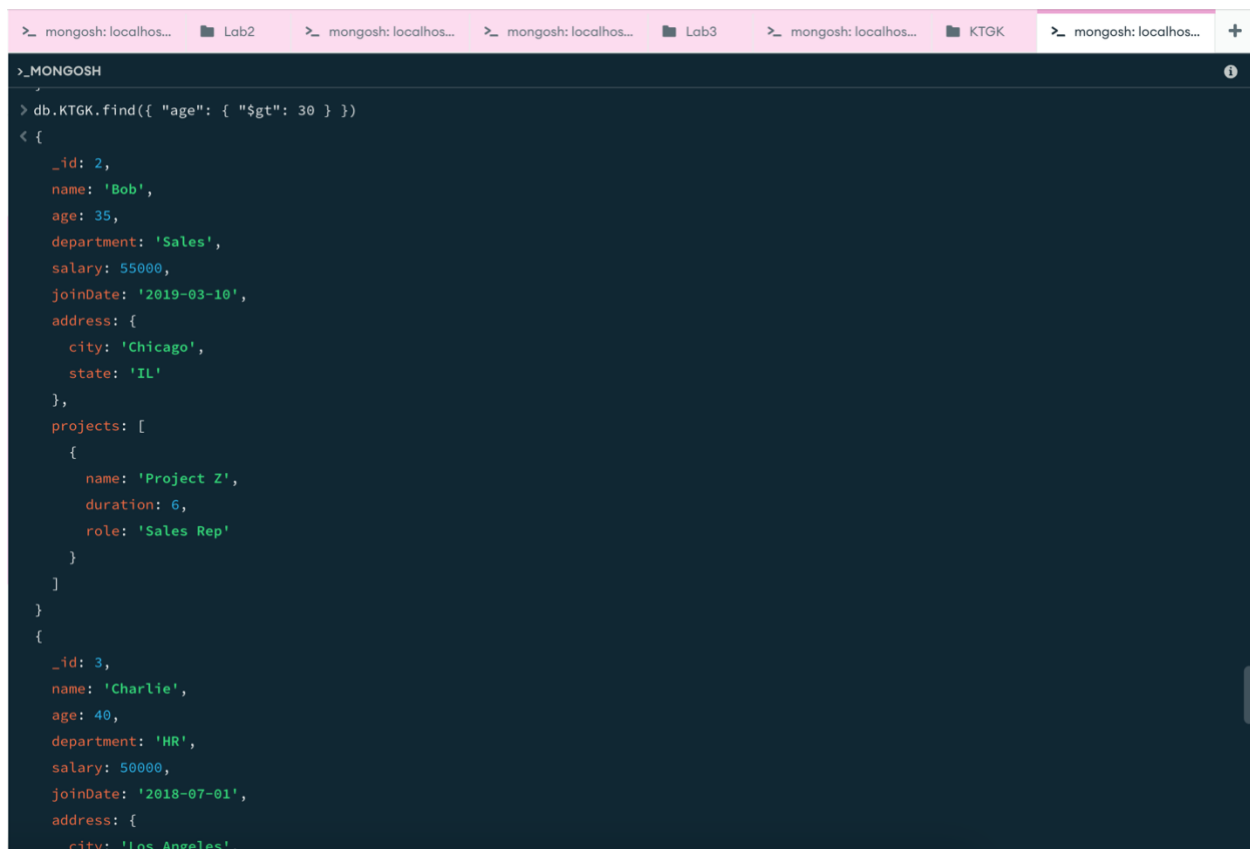
## Bài làm

### 1. Lấy tất cả dữ liệu có trong collection



```
> use Employees
< switched to db Employees
> db["KTGK"].find()
< {
  _id: 1,
  name: 'Alice',
  age: 28,
  department: 'IT',
  salary: 60000,
  joinDate: '2021-05-15',
  address: {
    city: 'New York',
    state: 'NY'
  },
  projects: [
    {
      name: 'Project X',
      duration: 12,
      role: 'Developer'
    },
    {
      name: 'Project Y',
      duration: 8,
      role: 'Lead'
    }
  ]
}
{
  _id: 2,
```

### 2. Tìm tất cả các nhân viên có tuổi lớn hơn 30



The screenshot shows a MongoDB terminal window with a dark background. The terminal title bar includes several tabs: 'mongosh: localhos...', 'Lab2', 'mongosh: localhos...', 'mongosh: localhos...', 'Lab3', 'mongosh: localhos...', 'KTGK', and 'mongosh: localhos...'. The terminal content shows a query being executed in the 'KTGK' database:

```
> db.KTGK.find({ "age": { "$gt": 30 } })
```

The results are displayed as a JSON array with two objects:

```
< {
  _id: 2,
  name: 'Bob',
  age: 35,
  department: 'Sales',
  salary: 55000,
  joinDate: '2019-03-10',
  address: {
    city: 'Chicago',
    state: 'IL'
  },
  projects: [
    {
      name: 'Project Z',
      duration: 6,
      role: 'Sales Rep'
    }
  ]
},
{
  _id: 3,
  name: 'Charlie',
  age: 40,
  department: 'HR',
  salary: 50000,
  joinDate: '2018-07-01',
  address: {
    city: 'Los Angeles',
```

3. Tìm tất cả các nhân viên thuộc phòng ban “IT”

```
>_ mongosh: localhos... Lab2 >_ mongosh: localhos... >_ mongosh: localhos... Lab3 >_ mongosh: localhos... KTGK >_ mongosh: localhos... +
>_MONGOSH
<
> db.KTGK.find({ "department": "IT" }).pretty()
< {
  _id: 1,
  name: 'Alice',
  age: 28,
  department: 'IT',
  salary: 60000,
  joinDate: '2021-05-15',
  address: {
    city: 'New York',
    state: 'NY'
  },
  projects: [
    {
      name: 'Project X',
      duration: 12,
      role: 'Developer'
    },
    {
      name: 'Project Y',
      duration: 8,
      role: 'Lead'
    }
  ]
}
{
  _id: 4,
  name: 'David',
  age: 30,
```

4. Tìm tất cả nhân viên có mức lương trên 50000.

```
>_ mongosh: localhos... Lab2 >_ mongosh: localhos... >_ mongosh: localhos... Lab3 >_ mongosh: localhos... KTGK >_ mongosh: localhos... +
>_MONGOSH
> db.KTGK.find({ "salary": { "$gt": 50000 } }).pretty()
< {
  _id: 1,
  name: 'Alice',
  age: 28,
  department: 'IT',
  salary: 60000,
  joinDate: '2021-05-15',
  address: {
    city: 'New York',
    state: 'NY'
  },
  projects: [
    {
      name: 'Project X',
      duration: 12,
      role: 'Developer'
    },
    {
      name: 'Project Y',
      duration: 8,
      role: 'Lead'
    }
  ]
}
{
  _id: 2,
  name: 'Bob',
  age: 35,
```

## 5. Tìm tất cả nhân viên có tên bắt đầu bằng chữ “A”

```
> db.KTGK.find({ "name": { "$regex": "A", "$options": "i" } })
< {
  _id: 1,
  name: 'Alice',
  age: 28,
  department: 'IT',
  salary: 60000,
  joinDate: '2021-05-15',
  address: {
    city: 'New York',
    state: 'NY'
  },
  projects: [
    {
      name: 'Project X',
      duration: 12,
      role: 'Developer'
    },
    {
      name: 'Project Y',
      duration: 8,
      role: 'Lead'
    }
  ]
}
```

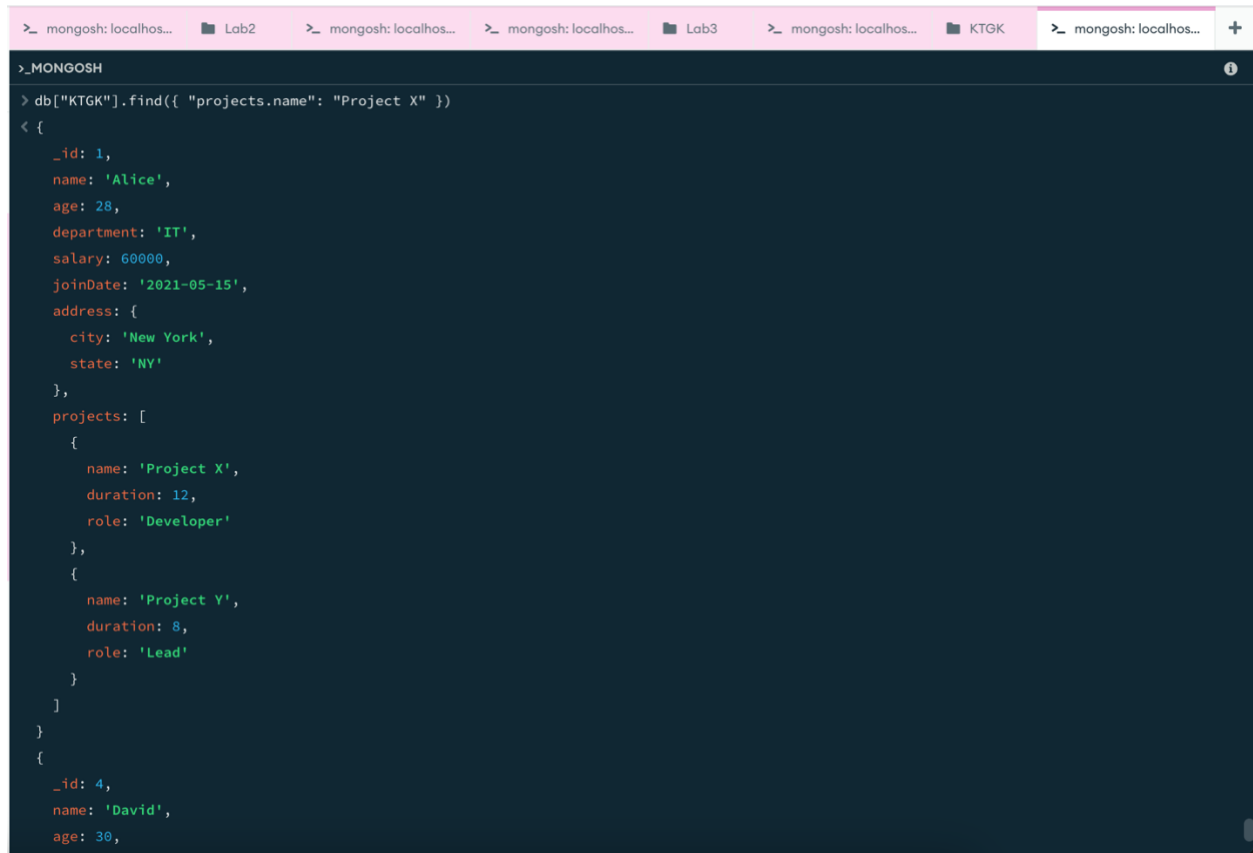
## 6. Sắp xếp nhân viên theo thứ tự tuổi tăng dần

```
>_ mongosh: localhos... Lab2 >_ mongosh: localhos... >_ mongosh: localhos... Lab3 >_ mongosh: localhos... KTGK >_ mongosh: localhos... +
>_MONGOSH ⓘ
> db.KTGK.find().sort({ "age": 1 }).pretty()
< {
  _id: 5,
  name: 'Eva',
  age: 25,
  department: 'Marketing',
  salary: 48000,
  joinDate: '2022-01-12',
  address: {
    city: 'Miami',
    state: 'FL'
  },
  projects: [
    {
      name: 'Project A',
      duration: 4,
      role: 'Coordinator'
    },
    {
      name: 'Project B',
      duration: 5,
      role: 'Analyst'
    }
  ]
},
{
  _id: 10,
  name: 'Jane',
  age: 27,
```

## 7. Đếm số nhân viên theo từng phòng ban

```
>_
> db.KTGK.aggregate([
  { $group: { _id: "$department", total_employees: { $count: {} } } }
])
< {
  _id: 'Finance',
  total_employees: 1
},
{
  _id: 'Marketing',
  total_employees: 2
},
{
  _id: 'HR',
  total_employees: 2
},
{
  _id: 'IT',
  total_employees: 3
},
{
  _id: 'Sales',
  total_employees: 2
}
```

## 8. Tìm các nhân viên đã tham gia dự án có tên “Project X”



The screenshot shows a MongoDB terminal window with the prompt `>_MONGOSH`. The command entered is `> db["KTGK"].find({ "projects.name": "Project X" })`. The output is a JSON array containing two documents. The first document is for an employee named Alice, and the second is for David. Alice's document includes a nested array of projects, with 'Project X' being the first project she worked on as a Developer.

```
>_MONGOSH
> db["KTGK"].find({ "projects.name": "Project X" })
< [
  {
    _id: 1,
    name: 'Alice',
    age: 28,
    department: 'IT',
    salary: 60000,
    joinDate: '2021-05-15',
    address: {
      city: 'New York',
      state: 'NY'
    },
    projects: [
      {
        name: 'Project X',
        duration: 12,
        role: 'Developer'
      },
      {
        name: 'Project Y',
        duration: 8,
        role: 'Lead'
      }
    ]
  },
  {
    _id: 4,
    name: 'David',
    age: 30,
```

## 9. Tìm nhân viên có tham gia ít nhất 2 dự án

```
>_ mongosh: localhos... Lab2 >_ mongosh: localhos... >_ mongosh: localhos... Lab3 >_ mongosh: localhos... KTGK >_ mongosh: localhos... +
>_MONGOSH
> db["KTGK"].find({ "projects.1": { $exists: true } })
< {
  _id: 1,
  name: 'Alice',
  age: 28,
  department: 'IT',
  salary: 60000,
  joinDate: '2021-05-15',
  address: {
    city: 'New York',
    state: 'NY'
  },
  projects: [
    {
      name: 'Project X',
      duration: 12,
      role: 'Developer'
    },
    {
      name: 'Project Y',
      duration: 8,
      role: 'Lead'
    }
  ]
}
{
  _id: 5,
  name: 'Eva',
  age: 25,
```

## 10. Tìm các nhân viên có mức lương cao nhất

```
> db["KTGK"].find().sort({ salary: -1 }).limit(1)
< {
  _id: 8,
  name: 'Helen',
  age: 45,
  department: 'Sales',
  salary: 80000,
  joinDate: '2015-04-22',
  address: {
    city: 'Chicago',
    state: 'IL'
  },
  projects: [
    {
      name: 'Project Z',
      duration: 9,
      role: 'Manager'
    },
    {
      name: 'Project C',
      duration: 3,
      role: 'Consultant'
    }
  ]
}
```

## 11. Tính mức lương trung bình giữa các phòng ban

```
>_ mongosh: localhos... Lab2 >_ mongosh: localhos... >_ mongosh: localhos... Lab3 >_ mongosh: localhos... KTGK >_ mongosh: localhos... +
>_MONGOSH
> db["KTGK"].aggregate([
  {
    $group: {
      _id: "$department",
      avgSalary: { $avg: "$salary" }
    }
  }
])
< {
  _id: 'Finance',
  avgSalary: 65000
}
{
  _id: 'Marketing',
  avgSalary: 49000
}
{
  _id: 'HR',
  avgSalary: 51000
}
{
  _id: 'IT',
  avgSalary: 67333.33333333333
}
{
  _id: 'Sales',
  avgSalary: 67500
}
Employees >
```

12. Tìm các nhân viên gia nhập sau ngày “2020-01-01”



```
>_ mongosh: localhos... Lab2 >_ mongosh: localhos... >_ mongosh: localhos... Lab3 >_ mongosh: localhos... KTGK >_ mongosh: localhos... +
>_MONGOSH
> db["KTGK"].find({
  joinDate: { $gt: "2020-01-01" }
})
< {
  _id: 1,
  name: 'Alice',
  age: 28,
  department: 'IT',
  salary: 60000,
  joinDate: '2021-05-15',
  address: {
    city: 'New York',
    state: 'NY'
  },
  projects: [
    {
      name: 'Project X',
      duration: 12,
      role: 'Developer'
    },
    {
      name: 'Project Y',
      duration: 8,
      role: 'Lead'
    }
  ]
}
{
  _id: 4,
```

13. Đếm số nhân viên theo thành phố (trong trường address)

```
>_ mongosh: localhos... Lab2 >_ mongosh: localhos... >_ mongosh: localhos... Lab3 >_ mongosh: localhos... KTGK >_ mongosh: localhos... +
>_MONGOSH
> db["KTGK"].aggregate([
  { $group: { _id: "$address.city", count: { $sum: 1 } } }
])
< {
  _id: 'New York',
  count: 2
}
{
  _id: 'San Francisco',
  count: 1
}
{
  _id: 'Boston',
  count: 2
}
{
  _id: 'Seattle',
  count: 1
}
{
  _id: 'Los Angeles',
  count: 1
}
{
  _id: 'Chicago',
  count: 2
}
{
  _id: 'Miami',
```

#### 14. Cập nhật tăng 10% lương cho tất cả nhân viên thuộc phòng “Sales”

```
> db["KTGK"].updateMany(
  { department: "Sales" },
  { $mul: { salary: 1.1 } }
)
< {
  acknowledged: true,
  insertedId: null,
  matchedCount: 2,
  modifiedCount: 2,
  upsertedCount: 0
}
```

#### 15. Xóa các nhân viên có tuổi nhỏ hơn 25

```
> db["KTGK"].deleteMany(
  { age: { $lt: 25 } }
)
< {
  acknowledged: true,
  deletedCount: 0
}
```

#### 16. Tìm các nhân viên có địa chỉ ở thành phố “New York”

```
>_ mongosh: localhos... Lab2 >_ mongosh: localhos... >_ mongosh: localhos... Lab3 >_ mongosh: localhos... KTGK >_ mongosh: localhos... +
>_MONGOSH
> db["KTGK"].find({ "address.city": "New York" })
< {
  _id: 1,
  name: 'Alice',
  age: 28,
  department: 'IT',
  salary: 60000,
  joinDate: '2021-05-15',
  address: {
    city: 'New York',
    state: 'NY'
  },
  projects: [
    {
      name: 'Project X',
      duration: 12,
      role: 'Developer'
    },
    {
      name: 'Project Y',
      duration: 8,
      role: 'Lead'
    }
  ]
}
{
  _id: 9,
  name: 'Ian',
  age: 38,
```

17. Lấy 3 nhân viên có mức lương cao nhất

```
>_ mongosh: localhos... Lab2 >_ mongosh: localhos... >_ mongosh: localhos... Lab3 >_ mongosh: localhos... KTGK >_ mongosh: localhos... +
>_MONGOSH
> db["KTGK"].find().sort({ salary: -1 }).limit(3)
< {
  _id: 8,
  name: 'Helen',
  age: 45,
  department: 'Sales',
  salary: 88000,
  joinDate: '2015-04-22',
  address: {
    city: 'Chicago',
    state: 'IL'
  },
  projects: [
    {
      name: 'Project Z',
      duration: 9,
      role: 'Manager'
    },
    {
      name: 'Project C',
      duration: 3,
      role: 'Consultant'
    }
  ]
}
{
  _id: 7,
  name: 'Grace',
  age: 32,
```

18. Tìm các nhân viên có tham gia ít nhất 1 dự án với thời gian thực hiện (duration) trên 6 tháng

```
>_ mongosh: localhos... Lab2 >_ mongosh: localhos... >_ mongosh: localhos... Lab3 >_ mongosh: localhos... KTGK >_ mongosh: localhos... +
>_MONGOSH
> db["KTGK"].find({ "projects.duration": { $gt: 6 } })
< {
  _id: 1,
  name: 'Alice',
  age: 28,
  department: 'IT',
  salary: 60000,
  joinDate: '2021-05-15',
  address: {
    city: 'New York',
    state: 'NY'
  },
  projects: [
    {
      name: 'Project X',
      duration: 12,
      role: 'Developer'
    },
    {
      name: 'Project Y',
      duration: 8,
      role: 'Lead'
    }
  ]
}
{
  _id: 4,
  name: 'David',
  age: 30,
```

## 19. Tính số dự án trung bình của nhân viên theo từng phòng ban

```
> db["KTGK"].aggregate([
  { $group: { _id: "$department", avgProjects: { $avg: { $size: "$projects" } } } }
])
< {
  _id: 'Sales',
  avgProjects: 1.5
}
{
  _id: 'Marketing',
  avgProjects: 1.5
}
{
  _id: 'Finance',
  avgProjects: 0
}
{
  _id: 'HR',
  avgProjects: 0.5
}
{
  _id: 'IT',
  avgProjects: 1.3333333333333333
}
```

## 20. Sắp xếp các nhân viên theo phòng ban, sau đó theo tuổi và cuối cùng mức lương tăng dần

```
>_ mongosh: locaishos... Lab2 >_ mongosh: locaishos... >_ mongosh: locaishos... Lab3 >_ mongosh: locaishos... KTGK >_ mongosh: locaishos... +
>_MONGOSH
> db["KTGK"].find().sort({ department: 1, age: 1, salary: 1 })
< {
  _id: 6,
  name: 'Frank',
  age: 29,
  department: 'Finance',
  salary: 65000,
  joinDate: '2020-11-30',
  address: {
    city: 'Boston',
    state: 'MA'
  },
  projects: []
}
{
  _id: 10,
  name: 'Jane',
  age: 27,
  department: 'HR',
  salary: 52000,
  joinDate: '2022-06-18',
  address: {
    city: 'Boston',
    state: 'MA'
  },
  projects: [
    {
      name: 'Project D',
      duration: 7,
      ...
    }
  ]
}
```