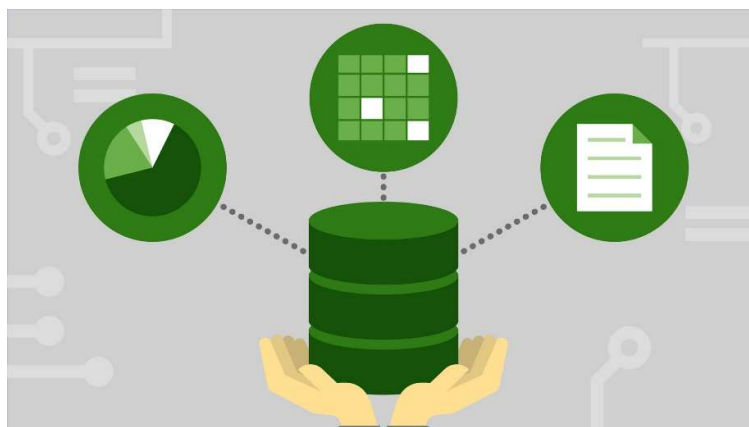


به نام خدا



دانشگاه تهران  
پردیس دانشکده‌های فنی  
دانشکده برق و کامپیوتر



## آزمایشگاه پایگاه داده

دستورکار شماره ۷ و ۸

# MongoDB

مهلت تحویل : ۱۴۰۱/۰۳/۲۰

مجتبی بنائی

## دستور کار شماره ۷ و ۸

هدف اصلی از این دستورکار، آشنایی با مانگودی بی به عنوان یکی از رایجترین دیتابیس‌های غیررابطه‌ای (NoSQL) دنیاست.

با توجه به رواج استفاده از این دیتابیس در شرکتهای مختلف، کار با مانگو دی بی را در دو بخش انجام خواهیم داد.

### دستورکار شماره ۷

در بخش اول که همان دستور کار شماره ۷ خواهد بود، به آشنایی با مانگودی بی به کمک فیلم‌های آموزشی سایت رسمی مانگودی به آدرس

<https://university.mongodb.com/courses/M001/about>

خواهیم پرداخت.

بنابراین، با مرور فیلم‌های آموزشی مانگودی بی و انجام بخشی از تمرینات انجام شده آن، از دستوراتی که نوشته‌اید و خروجی‌هایی که تولید کرده‌اید (انتخاب آنها بر عهده خودتان خواهد بود)، اسکرین شات گرفته، برای هر یک، توضیحی کوتاه بدهید تا گزارش کار جلسه ۷ ام، آماده شود. (شش کوئری و خروجی آنها برای گزارش کار هفتم کافی است.)

دقت کنید که مجموع طول دوره آموزشی فوق، هشت ساعت است اما در همان حدی از فیلم‌ها را مشاهده و به صورت عملی کار کنید که آماده انجام تمرین و دستور کار شماره ۸ شوید. (هر چند توصیه می‌کنم شش فصل این دوره را حتما مشاهده و به صورت عملی انجام دهید)

### دستورکار شماره ۸

در این تمرین، فقط به ذخیره اطلاعات صد هزار کاربر در قالب جی‌سان در مانگودی بی و کار با این داده‌ها می‌پردازیم.

#### دریافت اطلاعات

برای ایجاد صد هزار کاربر از امکانات ایجاد کاربر تصادفی سایت randomuser.me استفاده می‌کنیم. این سایت که ملیت ایرانی را هم پشتیبانی می‌کند، از طریق API، یک تا پنج هزار کاربر تصادفی با ملیت و مشخصات داده شده تولید کرده، به ما در قالب جی‌سان برمی‌گرداند. نمونه‌ای از داده‌های تولید شده توسط این سایت را در زیر مشاهده می‌کنید:

```
{
  "results": [
    {
      "gender": "female",
      "name": {
        "title": "Ms",
        "first": "نیایش",

```

```

    "last": "نژاد سلطانی"
  },
  "location": {
    "street": {
      "number": 8962,
      "name": "استقلال میدان"
    },
    "city": "یزد",
    "state": "تهران",
    "country": "Iran",
    "postcode": 59373,
    "coordinates": {
      "latitude": "7.5752",
      "longitude": "162.0874"
    },
    "timezone": {
      "offset": "-5:00",
      "description": "Eastern Time (US & Canada), Bogota, Lima"
    }
  },
  "email": "nyysh.sltnynjd@example.com",
  "login": {
    "uuid": "c5e3eaf3-80e9-4693-8eb6-aa3c64c04480",
    "username": "happyostrich714",
    "password": "qwert1",
    "salt": "b5JyKob5",
    "md5": "a240f2718337bc2cfee83b01b8b492d4",
    "sha1": "0d2a39cb9bbec1893731f529875a9809c7d6478f",
    "sha256": "20a123ea18446e5a515d406df7ad5931d7af4d82a31a30d3dffd096d536c914a"
  },
  "dob": {
    "date": "1955-12-06T15:34:18.796Z",
    "age": 65
  },
  "registered": {
    "date": "2015-11-17T00:53:50.137Z",
    "age": 5
  },
  "phone": "037-02250625",
  "cell": "0932-136-2921",
  "id": {
    "name": "",
    "value": "None"
  },
  "picture": {

```

```

        "large": "https://randomuser.me/api/portraits/women/41.jpg",
        "medium": "https://randomuser.me/api/portraits/med/women/41.jpg",
        "thumbnail": "https://randomuser.me/api/portraits/thumb/women/41.jpg"
    },
    "nat": "IR"
  }
],
"info": {
  "seed": "b4fc8f16a480868c",
  "results": 1,
  "page": 1,
  "version": "1.3"
}
}

```

کافی است همین جی‌سان دریافت شده را در مانگو دی بی ذخیره کنید (البته اطلاعات موجود در فیلد results و آن هم به ازای هر کاربر در یک سند جداگانه ذخیره خواهد شد).

برای دریافت اطلاعات کاربران (که برای این تمرین حداقل نیاز به صد هزار کاربر خواهیم داشت)، می‌توانید از کد پایتون زیر استفاده کنید (البته کد را باید اصلاح کنید که این تعداد کاربر دریافت و در مانگو ذخیره شود - هر یک دقیقه ۵۰۰۰ کاربر را می‌توانید دریافت کنید بنابراین بین هر دو دستور، وقفه لازم را تولید کنید):

```

import requests, json

url = "https://randomuser.me/api/?nat=ir "

querystring = {"nat": "ir", "results": "5000"}

response = requests.request("GET", url, params=querystring)
data = json.loads(response.text)
print(data)

```

### دانلود مانگودی بی و ساخت کالکشن کاربران

مانگو دی بی را نصب کرده<sup>۱</sup> و کالکشن users را درون دیتابیس metadata (این دیتابیس هم باید ایجاد شود) بسازید. می‌توانید از خط فرمان مانگودی بی یا ابزارهای گرافیکی رایج مانند MongoDB Compass یا Robo 3T<sup>۲</sup> برای این منظور استفاده کنید. در مرحله بعد، کوئری‌های زیر را نوشته، از خروجی آنها اسکرین شات بگیرید و توضیحات لازم را در پایین هر عکس، بنویسید تا گزارش کار ۸ شما نیز آماده شود. نیازی به ضمیمه کردن کدها و کوئری‌ها نیست.

<sup>۱</sup> <https://bit.ly/2XWSqM7>

<sup>۲</sup> <https://robomongo.org/>

## گام اول

در گام اول، با فراخوانی آدرس <https://randomuser.me/api/?nat=ir> اطلاعات چندین کاربر تصادفی را دریافت و در مانگودی بی به صورت دستی ذخیره کنید و بررسی کنید چه فیلدهایی توسط خود مانگو به صورت خودکار، به داده‌ها افزوده می‌شود. سپس با استفاده از کتابخانه `pymongo`<sup>1</sup> کدهای دریافت اطلاعات فوق را به گونه‌ای تغییر دهید که همزمان با دریافت اطلاعات کاربران، آنها را در مانگو هم ذخیره کنید. با دستور `count`، مطمئن شوید که صد هزار داده، دریافت شده باشند.

## گام دوم - دستورات اصلی

در گام دوم به دستورات پایه مانگودی بی می‌پردازیم و به سوالات زیر پاسخ می‌دهیم:

1. پنج کوئری برای یافتن اطلاعات از بین این صد هزار سند، به دلخواه خودتان طراحی کنید مثلاً یک کوئری می‌تواند این باشد: نام و نام خانوادگی کاربرانی را پیدا کنید که بالای 40 سال سن داشته و ساکن اهواز باشند. یا تمام کاربرانی را بیابید که در یک ماه اخیر در سایت ما ثبت نام کرده‌اند
2. تاریخ ثبت شده برای تولد (dob) و زمان ثبت نام به میلادی است. می‌خواهیم در کنار این تاریخ‌ها، یک فیلد جدید اضافه کنیم با نام `year_persian` که سال متناظر تقویم فارسی تولد و ثبت نام را هم داشته باشیم. از دستور `update` و یک فرمول ساده تبدیل سال میلادی به شمسی، استفاده کرده، این تغییر را اعمال کنید.
3. ذخیره پسورد به شکل خام در اطلاعات کاربر، کاری غیرحرفه‌ای است. می‌خواهیم این مشکل را برطرف کنیم. چه راه حلی برای حل مساله پیشنهاد می‌کنید؟ راه حل را ابتدا روی یک سند خاص، امتحان کنید و مطمئن شوید بعد از اعمال تغییرات، آن کاربر خاص را با داشتن یوزنیم و پسورد، می‌توانید پیدا کنید. سپس تمام کاربران را به روز رسانی کنید.

## گام سوم - دستورات تجمعی و آماری (Aggregate Functions)

1. می‌خواهیم یک کمپین تبلیغاتی برای سایت ایجاد کنیم. برای این منظور باید کاربران را به سه گروه سنی تقسیم کنیم: نوجوانان / جوانان و افراد میانسال به بالا. دسته اول سنی کمتر از 16، دسته دوم سنی بین 16 تا 30 و دسته سوم، بالاتر از 30 سال خواهند داشت. دستوری بنویسید که تعداد هر گروه را برگرداند.
2. تعداد کاربران هر استان را به تفکیک لازم داریم. چگونه این اطلاعات را تولید می‌کنید؟

<sup>1</sup> <https://pymongo.readthedocs.io/en/stable/tutorial.html>