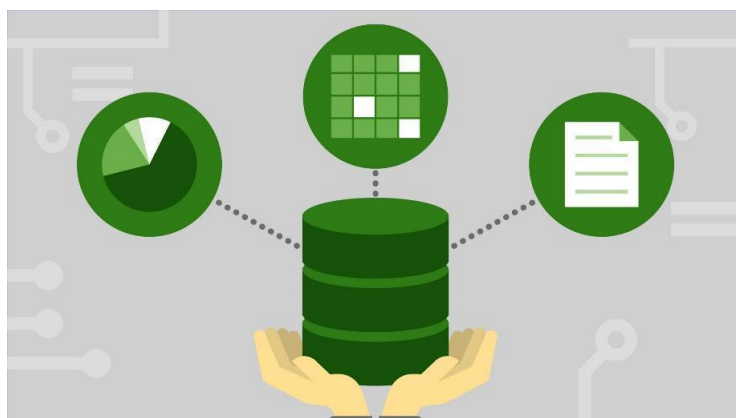


به نام خدا



دانشگاه تهران  
پردیس دانشکده‌های فنی  
دانشکده برق و کامپیوتر



## آزمایشگاه پایگاه داده

دستور کار شماره ۵

**MongoDB**

فروردین ۱۴۰۲

مجتبی بنائی

## مقدمه

در این تمرین قصد داریم از داده‌های مربوط به سازمان بهداشت جهانی (WHO) مربوط به COVID-19 استفاده کنیم. این مجموعه داده شامل داده‌هایی در مورد تعداد موارد گزارش شده، مرگ و میر و بهبودی برای هر کشور و منطقه و همچنین تاریخ گزارش است. به طور خاص، مجموعه داده شامل متغیرهای زیر است:

Date_reported	Country_code	Country	WHO_region	New_cases	Cumulative_cases	New_deaths	Cumulative_deaths
---------------	--------------	---------	------------	-----------	------------------	------------	-------------------

این مجموعه داده به روزرسانی‌های روزانه در مورد پیشرفت ویروس در نقاط مختلف جهان را نشان می‌دهد.

## جمع‌آوری داده

بخشی از پروژه شما شامل دریافت مجموعه داده از سایت WHO<sup>1</sup> است. برای دریافت این داده، شما می‌توانید از دو روش استفاده کنید: یا به صورت مستقیم از سایت WHO فایل CSV را دانلود کنید، یا از طریق کد پایتون مجموعه داده را دانلود کرده و به دیتابیس خود اضافه کنید. استفاده از روش دوم، یعنی دانلود مجموعه داده از طریق کد پایتون، به عنوان نمره امتیازی برای شما منظور خواهد شد. به این منظور، لازم است که از کتابخانه `pymongo`<sup>2</sup> برای ذخیره داده در مانگودی‌بی و `requests` برای دریافت داده استفاده کنید. مطمئن شوید که پس از اضافه شدن داده به کمک کد، نام ستون‌ها در دیتابیس با نام ستون‌ها در مجموعه داده اصلی همخوانی داشته باشد؛ در صورتی که اینطور نبود، ستون‌های ناهماهنگ را `rename` کنید.

در مرحله بعد، به دلخواه خود مانگودی‌بی را راه‌اندازی کرده و به آن وصل شوید. سپس، کالکشن `COVID_19` را ایجاد کرده و اطلاعات مربوطه را در آن اضافه کنید. پس از اضافه شدن، با استفاده از دستور `count` مطمئن شوید که تعداد رکوردها ۲۸۳۴۵۲ می‌باشد.

<sup>1</sup><https://covid19.who.int/WHO-COVID-19-global-data.csv>

<sup>2</sup><https://pymongo.readthedocs.io/en/stable/tutorial.html>

پیش از شروع به کار با مانگودی بی، می توانید کتاب <sup>3</sup>The Little Mongo را به عنوان یک راهنمای سریع برای کار با مانگو در این تمرین کنار خود داشته باشید.

برای دستوراتی که در ادامه پیاده سازی می کنید، لازم است که دستور نوشته شده، خروجی و زمان اجرا را به دستیار مربوطه نشان دهید.

## دستورات اصلی

در این قسمت، به دستورات پایه مانگودی بی می پردازیم:

۱- تمام اطلاعات موجود برای کشور ایران را نمایش دهید. (Country\_code = IR)

۲- تعداد افرادی که در تاریخ 08-04-2023 در کشور ایران فوت شدند (New\_deathes)

۳- تاریخ هایی که در کشور ایران بیشتر از ۲۰۰ نفر فوت شده اند.

۴- تاریخ ثبت شده به صورت میلادی می باشد، می خواهیم یک فیلد جدید با نام shamsi\_date اضافه کنیم. برای این منظور از دستور update و یک فرمول ساده تبدیل سال میلادی به شمسی استفاده کنید.

۵- تمامی رکوردهایی که مقدار New\_cases در آن ها منفی است یا مقدار Cumulative\_cases به ازای آن ها وجود ندارد را حذف کنید.

## دستورات تجمعی و آماری (Aggregate Function)

در این قسمت، به دستورات پیشرفته تر مانگودی بی می پردازیم:

۱- تاریخی را به ازای هر کشور پیدا کنید که اولین مورد کرونا گزارش شده باشد.

۲- میانگین تعداد موارد جدید در روز را برای هر کشور محاسبه کنید و سپس به صورت نزولی سورت کنید.

۳- تعداد روزهای بین اولین و آخرین مورد گزارش شده را برای هر کشور به دست آورید.

<sup>3</sup> <https://openmymind.net/mongodb.pdf>

۴- کشوری را پیدا کنید که مقدار Cumulative\_cases آن در یک روز خاص، بیشینه باشد.

۵- ۱۰ کشور با بیشترین میزان مرگ در میلیون نفر را گزارش دهید.