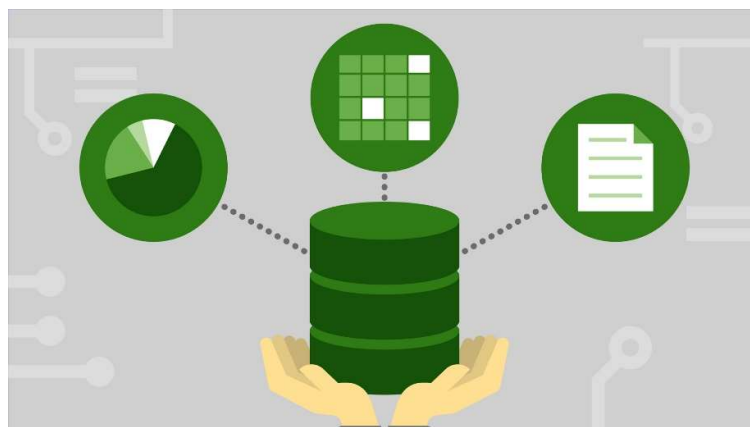


به نام خدا



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده برق و کامپیوتر



آزمایشگاه پایگاه داده

دستورکار شماره ۹

مهلت تحویل :

۱۴۰۱/۰۳/۲۰

مجتبی بنائی

دستور کار شماره ۹ - کار با ردیس

در بسیاری از کاربردهای تجاری، نیاز به ذخیره موقت داده‌ها در حافظه داریم که جلوی مراجعه مکرر به دیتابیس را گرفته، سرعت پاسخگویی را افزایش دهیم.

مثلا فرض کنید ده خبر برتر یک سایت خبری که در صفحه اول آن قرار گرفته است، میتواند در حافظه قرار گیرد و هر یک ربع، آپدیت شود. در این فاصله تمام درخواستهای مشاهده صفحه اول، به جای مراجعه به دیتابیس، از حافظه که بسیار سریعتر است پاسخ داده می شود.

ردیس به عنوان یک دیتابیس مقیم در حافظه، این نقش را دنیای امروز به عنوان یک ابزار دم دستی و بسیار کار راه انداز انجام می دهد. دو دستور اصلی آن یعنی `set` و `get` دقیقا اصلی ترین نیاز ما را برآورده میکنند: ذخیره یک متغیر با مقدار آن در حافظه و برگرداندن آن در صورت نیاز.

ردیس را دانلود و اجرا کنید. بعد از دانلود، سرور ردیس را باید اجرا کنید و در یک پنجره کنسول دیگر، خط فرمان یا `cli` آنرا اجرا کنید. برای ویندوز، نسخه به روز آن در دسترس نیست. برای اجرای ردیس در ویندوز از راهنمای زیر استفاده کنید:

<https://github.com/ServiceStack/redis-windows>

البته نسخه `buster` ردیس^۱ در داکر کمتر از چهل مگابایت است و اگر با داکر کار میکنید با این نسخه می‌توانید به راحتی کار کنید.

آنچه باید انجام دهید

۱. آموزش بسیار خلاصه اما مفیدی که در این آدرس^۲ قرار گرفته است مطالعه و دستورات آنرا انجام دهید و از هر یک، اسکرین شاتی بگیرید. (کار با خط فرمان)

۲. کدی که برای درج توئیت‌های سهام‌یاب نوشته‌اید را مجدداً با خواسته‌های زیر اجرا کنید و این دفعه به جای درج توئیت‌ها در الستیک سرچ، آنها را در ردیس با توجه به نیازمندیهای زیر، ذخیره کنید.

می‌خواهیم ده هشتگ برتر روز جاری و نیز هر ساعت از بیست و چهار ساعت گذشته را در ردیس ذخیره کنیم تا بدون نیاز به مراجعه به دیتابیس و تنها از طریق حافظه (استفاده از ردیس)، برترین (پرتکرارتری) هشتگ‌های روز گذشته و هر ساعت از شبانه روز را به دست آوریم. (این کار را در دنیای واقعی و برای جلوگیری از مراجعه مداوم به دیتابیس‌های اصلی برای گرفتن آمار لحظه‌ای داده‌ها، معمولا همیشه انجام می‌دهیم)

برای انجام این کار، نیاز به ایجاد متغیرهایی در ردیس به شکل زیر داریم (متغیرهایی از نوع `Sorted List` یا `Zlist`):

Hashtags:1401-02-31

Hashtags:1401-02-31:01

Hashtags:1401-02-31:13

که اولی برای ذخیره ده هشتگ برتر روز سی و یکم اردیبهشت، دومی برای ساعت ۰ تا ۱ صبح همین روز، سومی برای ساعت دوازده تا سیزده و الی آخر در نظر گرفته شده‌اند. هرکدام از این کلیدها، لیستی مرتب (z-list) از هشتگ‌ها را ذخیره میکنند که Score آنها (کلید مرتب سازی)، همان تعداد تکرار آنها در آن بازه خاص خواهد بود.

^۱ https://hub.docker.com/_/redis?tab=tags

^۲ yun.ir/5d2ew5 or <https://virgool.io/@vahiind/gyxrxnxtcju5>

بنابراین برای انجام خواسته این سوال، با دریافت هر توثیت، بعد از استخراج هشتگ‌های آن، تعداد تکرار هشتگ در ساعت فعلی و نیز بیست و چهار ساعت گذشته به روز رسانی کنید.

نکته: به ازای هر روز، بیست و پنج کلید در ردیس از نوع ZLIST ایجاد کنید یعنی برای هر ساعت از بیست و چهار ساعت، یک کلید و برای روز جاری هم یک کلید که جمعاً می‌شود ۲۵ کلید. بعد از دریافت هر هشتگ، این کلیدها را به روز رسانی کنید- دستور incr).

این تمرین را با پایتون انجام دهید. از این راهنما می‌توانید برای انجام این کار استفاده کنید:

<https://realpython.com/python-redis/>

البته بهتر است ابتدا در خط فرمان ردیس، چند هشتگ را به صورت دستی در لیست مرتب درج کرده و پس از بررسی‌های لازم و گرفتن جواب، آنها را به کد پایتون منتقل نمایید.

برنامه ساده‌ای به زبان پایتون، به صورت جداگانه بنویسید که وقتی اجرا شود، هشتگ‌های برتر ساعت جاری و نیز روز جاری را چاپ کند. البته می‌توانید ساعت و روز مورد نظر را هم به صورت پارامترهای خط فرمان از کاربر بگیرید و اگر هیچ پارامتری داده نشود، ساعت و روز جاری در نظر گرفته شود.

بعد از انجام این کار، به سوالات زیر هم پاسخ دهید:

- اگر بخواهیم بعد از اتمام هر روز، کلیدهای ساعتی روز قبل به صورت خودکار حذف شوند، چه دستوری باید هنگام ایجاد کلیدها اجرا کنیم؟
 - اگر بخواهیم داده‌های ردیس بر روی دیسک هم ذخیره شوند، (با فرکانس هر ده دقیقه یک بار)، چه دستوری استفاده می‌کنیم؟
 - اگر برنامه پایتون ما به صورت ناگهانی بسته شود و آنرا دوباره اجرا کنیم، تغییر خاصی در برنامه باید داده شود یا نه تا بتوان به نتایج ردیس اعتماد کرد؟ اگر ردیس به هر دلیلی متوقف و دوباره استارت شود، چه کاری باید انجام شود تا داده‌های هشتگ‌های موجود در ردیس برای یک هفته گذشته، قابل استفاده باشد؟ فقط توضیح دهید.
- فصل دو تا چهار کتاب The Little Redis Book برای ایده گرفتن انجام بخش دوم دستور کار می‌تواند منبع مناسبی باشد. موفق باشید.