بسمه تعالی

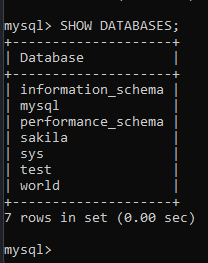
با عرض سلام و ادب و احترام خدمت شما و همکاران محترم

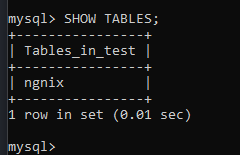
لینک های گیت‌هاب:

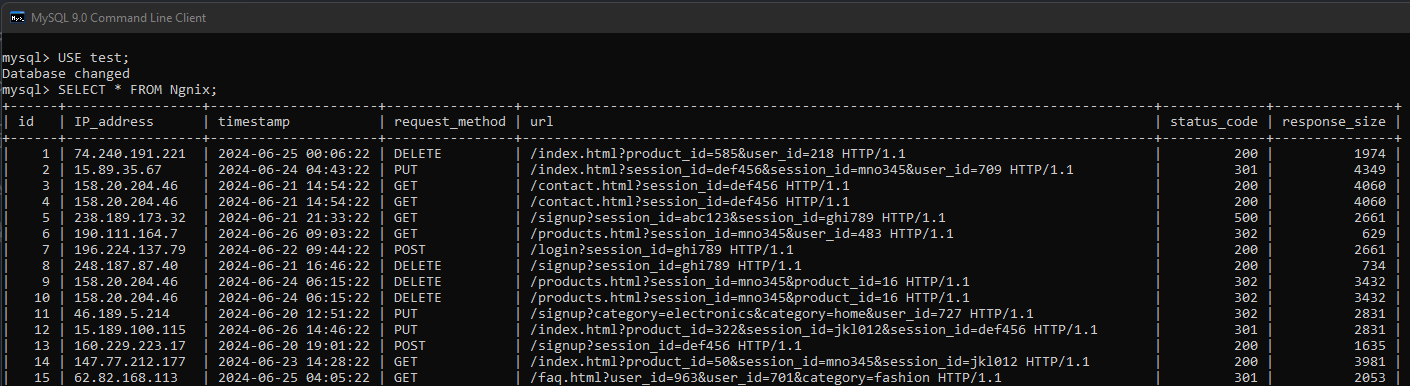
سایت بنده جهت نمایش رزومه و پروفایل گیت‌هاب: demo-amin.ir

گزارش عملکرد بررسی دیتاست Ngnix در قالب فایل .txt به صورت زیر است:

1. این برنامه به 2 صورت function base و class base نوشته شده است.
2. هر دو صورت در گیت‌هاب بارگزاری شده و قابل دسترسی هستند.
3. خروجی دیتابیس به صورت فایل هایی با فرمت های .csv و .xlsx تهیه شده است.

تصاویری از دیتابیس:



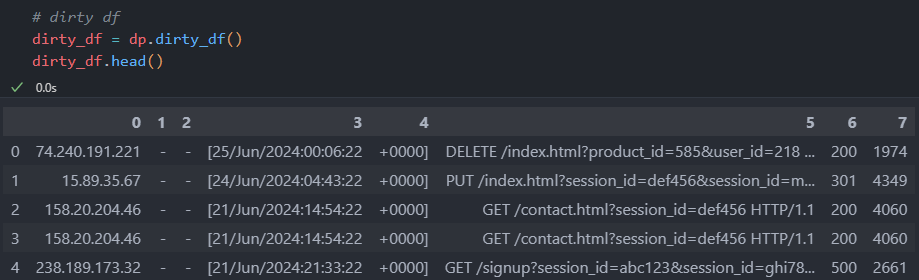


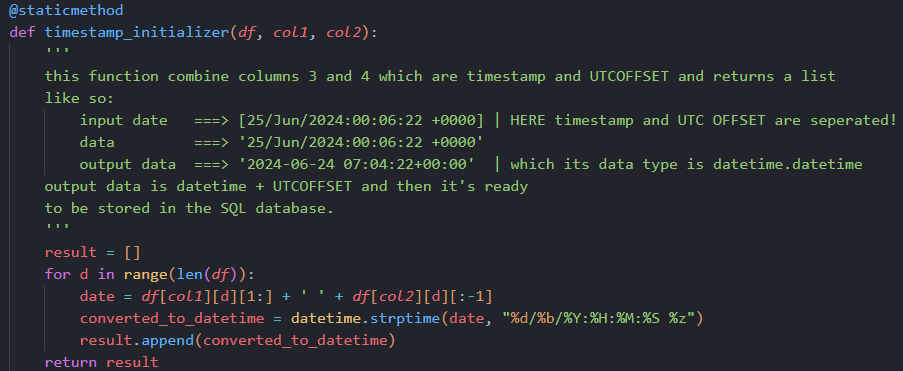
در اینجا فقط در مورد نحوه عملکرد در برنامه class base صحبت خواهیم کرد.

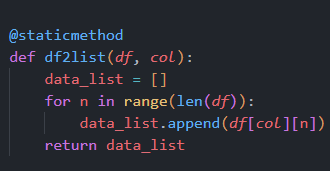
نام کلاس ها:

کلاس های DataProcessor و DataBase

بررسی کلاس DataProcessor:

این کلاس با دریافت فایل داده آنرا می‌خواند، تبدیل به دیتا فریم می‌کند. برای ستون timestamp نیاز به تغییرات هست و چون طبق خوانش زیر قسمتی از timestamp در ستون 4 قرار دارد بایستی آنها را یکی کرد. این قسمت OFFSET UTC نام دارد.

این موارد به کمک متد استاتیک زیر انجام می‌شود.

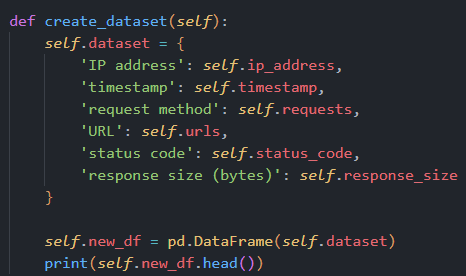


متد بعدی df2list هست که ستون ها را استخراج

و آنها را به لیست تبدیل می‌کند. این متد را برای

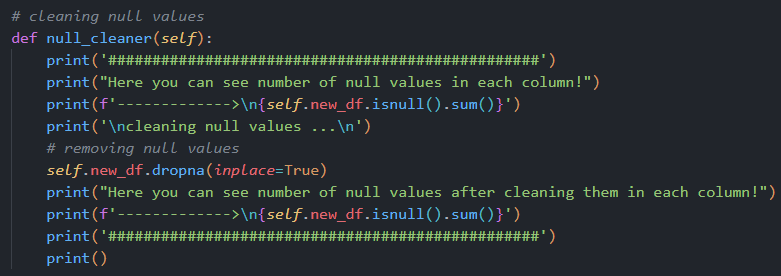
ستون های ip\_address و status\_code و

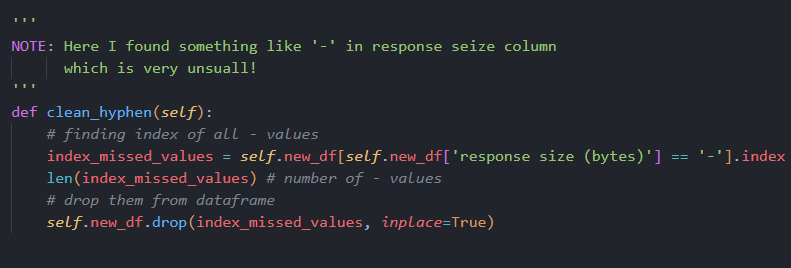
response\_size استفاده کردیم.

سپس به کمک متد create\_database دیتابیس جدید را ایجاد کردیم که ستون timestamp و دیگر ستون ها مثل url بدون مشکل حاضر هستند.

حالا که دیتابیس جدید را ساختیم باید دیتا را از null values و missed values بررسی کنیم.

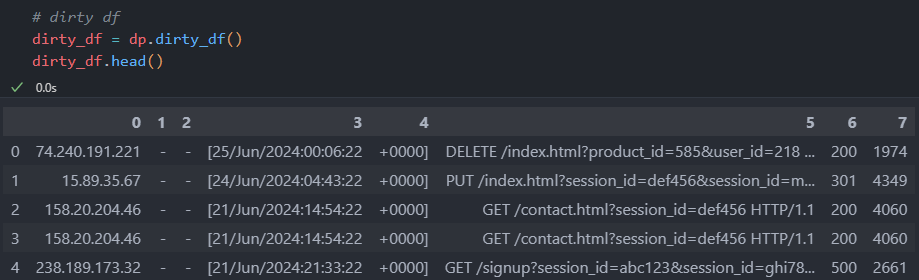
در این متد تعداد مقدار های ازدست رفته و نال هر ستون را نمایش دادیم:



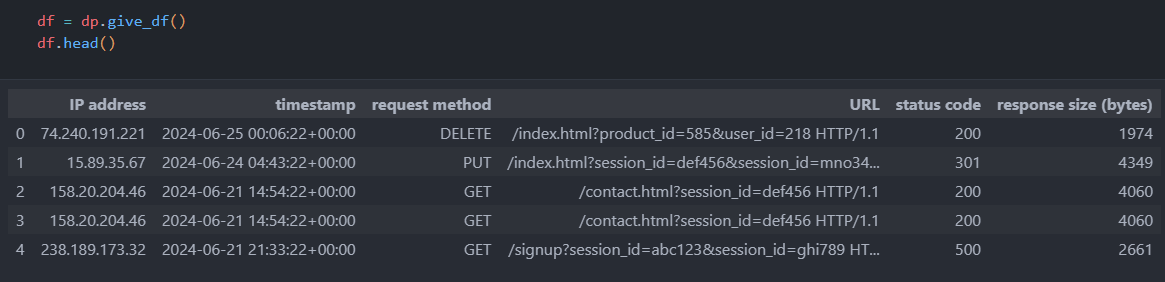
نکته بسیار مهم: هنگام وارد کردن دیتا به دیتابیس MySQL به نکته جالبی برخورد کردم، آن نکته این بود که در ستون response\_size مواردی مثل hyphen یا همان – که شبیه علامت تفریق هستند دیده شد و چون دیتابیس انتظار دریافت integer داشت ارور گرفتم. پس به عقب برگشتم و این مورد را به صورت زیر حل کردم. عملکرد من به اینصورت بود که index جاهایی که – یا hyphen مشاهده میشد را پیدا و حذفشان کردم.

نهایتا هم به کمک متد loader() تمامی متد ها را فراخوانی کردم. در آخر با متد give\_df() می‌توان به دیتافریم پاک سازی شده و آماده کار دست یافت.

تصاویری از دیتابیس کثیف و تمیز:

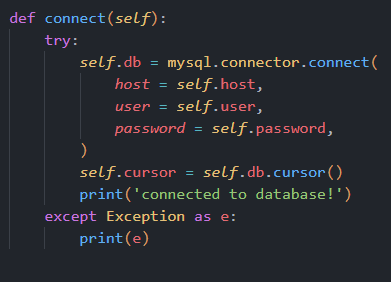


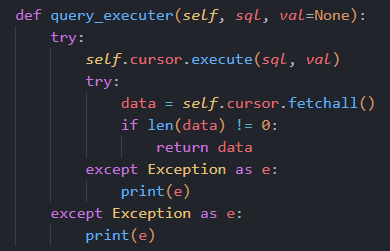
دیتابیس کثیف و پردازش نشده



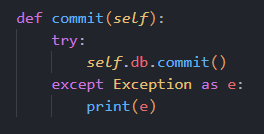
**دیتابیس تمیز و پردازش شده**

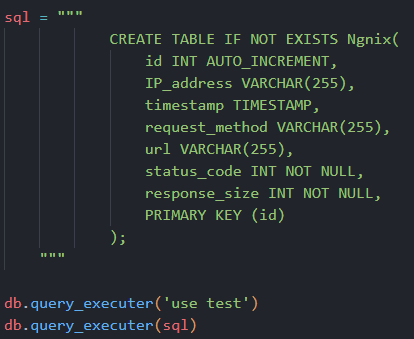
بررسی کلاس DataBase:

 این کلاس ابتدا به صورت زیر و با دریافت اطلاعاتی که مشاهده می‌کنید سعی می‌کند به دیتابیس متصل شود.

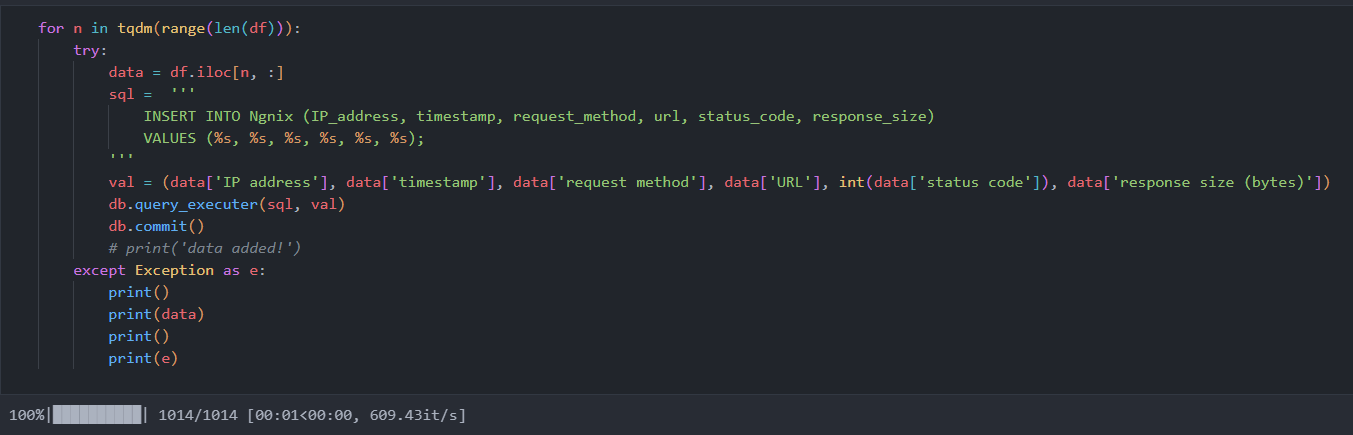
ایجاد یک کرسر یا cursor برای اجرای کوئری ها.

نکته: وقتی دیتا را insert می‌کنیم باید حتما بعدش آنرا commit() هم کنیم:



ساخت جدول:

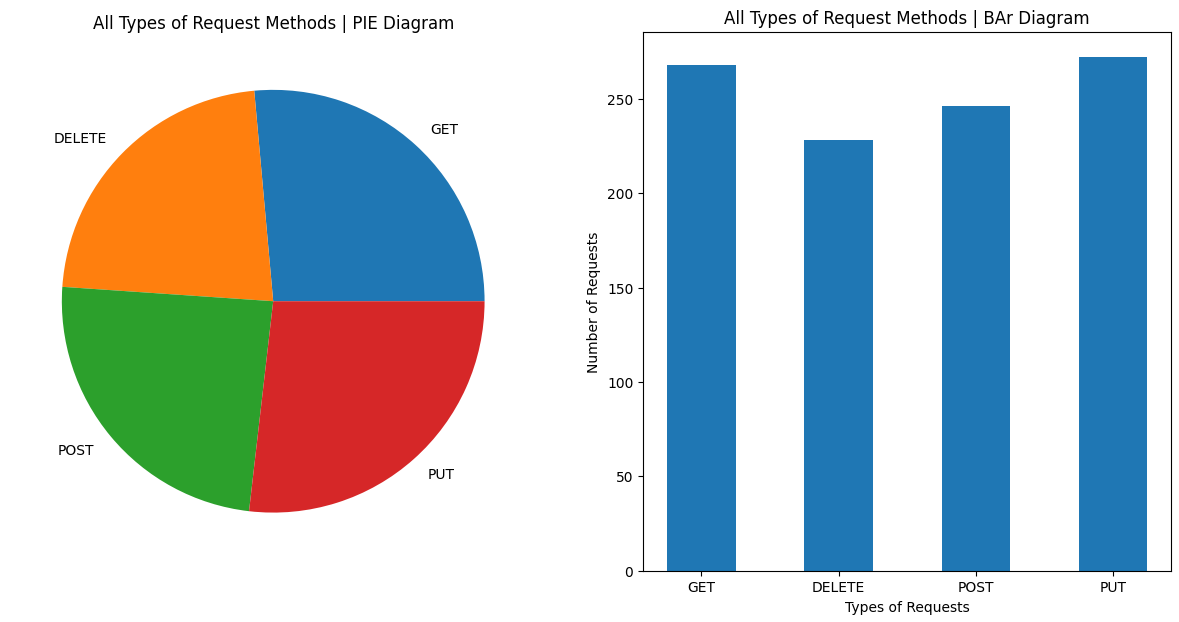
وارد کردن دیتا به دیتابیس:



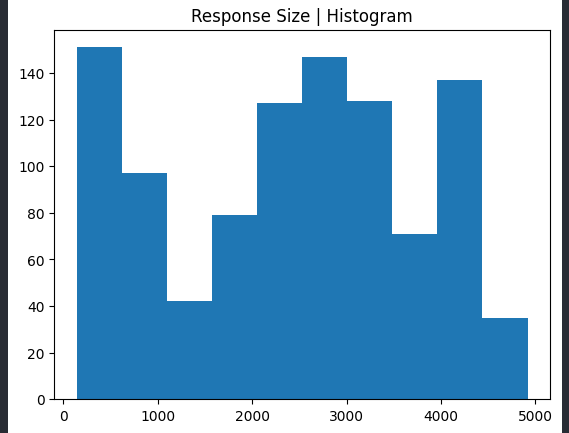
استفاده از کتابخانه tqdm برای نشان‌ دادن مقدار پیشرفت

قسمت رسم نمودار ها:

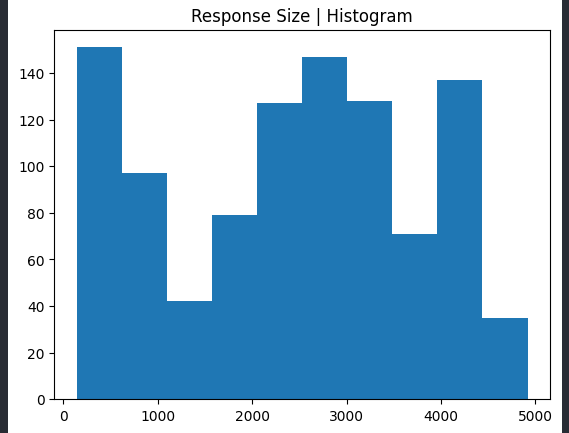
تمامی داده ها و اعداد استفاده شده برای ترسیم نمودار ها از دیتابیس خوانده شده و کاری با دیتافریم تمیز شده نداشتیم.

1.  نمایش request method های مختلف براساس تعدادشان بر روی 2 نمودار میله‌ای و پای

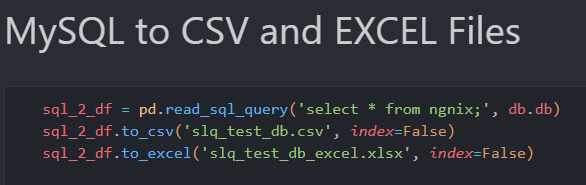
ترسیم مقادیر response size روی نمودار هیستوگرام



ترسیم مقادیر status code با هیستوگرام:



--------------------------------------------------------------------------

ذخیره سازی فایل های .csv و .xlsx

با تشکر فراوان از توجه شما.