

به نام خدا



مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲

۱. خلاصه داده ها

داده های این سری تمرین شامل آمار مربوط به پاندمی کرونا بود. در این داده ها ۶۷ ستون و ۱۳۴۲۳۵ ردیف وجود دارد که روزانه در حال آپدیت شدن و بزرگ تر شدن است. با مطالعه داده ها متوجه می شویم که در بعضی کشورها آمار کامل منتشر نمی شود. چون هدف تمرین مطالعه حالت های خاص نیست در ادامه دیتا های نامناسب را پاک و تمیز میکنیم.

۲. پیش پردازش داده ها

۲.۱ ابتدا داده هایی را که میتوانیم منطقی پر کنیم را با اعداد پر میکنیم. برای مثال ستون `total_new_cases` برای بعضی از کشورها تا یک تاریخ خاص `null` و بعد از آن شروع به پر شدن کرده است. در این حالت این سلول ها را با ۰ پر میکنیم. برای مثال دیگر به `total_cases_per_million` اشاره میکنیم که جاهای خالی را با توجه به ستون `population` پر کردیم. لیست ستون های پر شده در زیر آورده شده است.

● `total_deaths`

● `total_cases`

● `total_tests`

● `total_vaccinations`

● `people_fully_vaccinated`

به نام خدا



مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲

● `total_cases_per_million`

● `total_deaths_per_million`

● `total_tests_per_thousand`

● `people_vaccinated_per_hundred`

● `people_fully_vaccinated_per_hundred`

۲.۲ در ادامه بر روی کشور های مختلف راه می رویم و کشور هایی که داده های خالی زیاد دارند را حذف می کنیم. یعنی کشور هایی که بیش تر از ۵۰ درصد داده هایشان خالی است حذف می شوند. بعد از این کار ها تعداد کشور ها از ۲۲۳ به ۱۴۴ و تعداد داده ها از ۱۳۴۲۳۵ به ۹۱۰۲۳ رسید.

۳. تحلیل داده ها

۳.۱ ایران

ابتدا داده های مربوط به ایران را جدا می کنیم و یک سری تحلیل های اولیه روی آن می دهیم و سپس تمام کشور ها را با هم مقایسه می کنیم. در ادامه ابتدا نمودار `total cases` در طول زمان را می کشیم.

به نام خدا

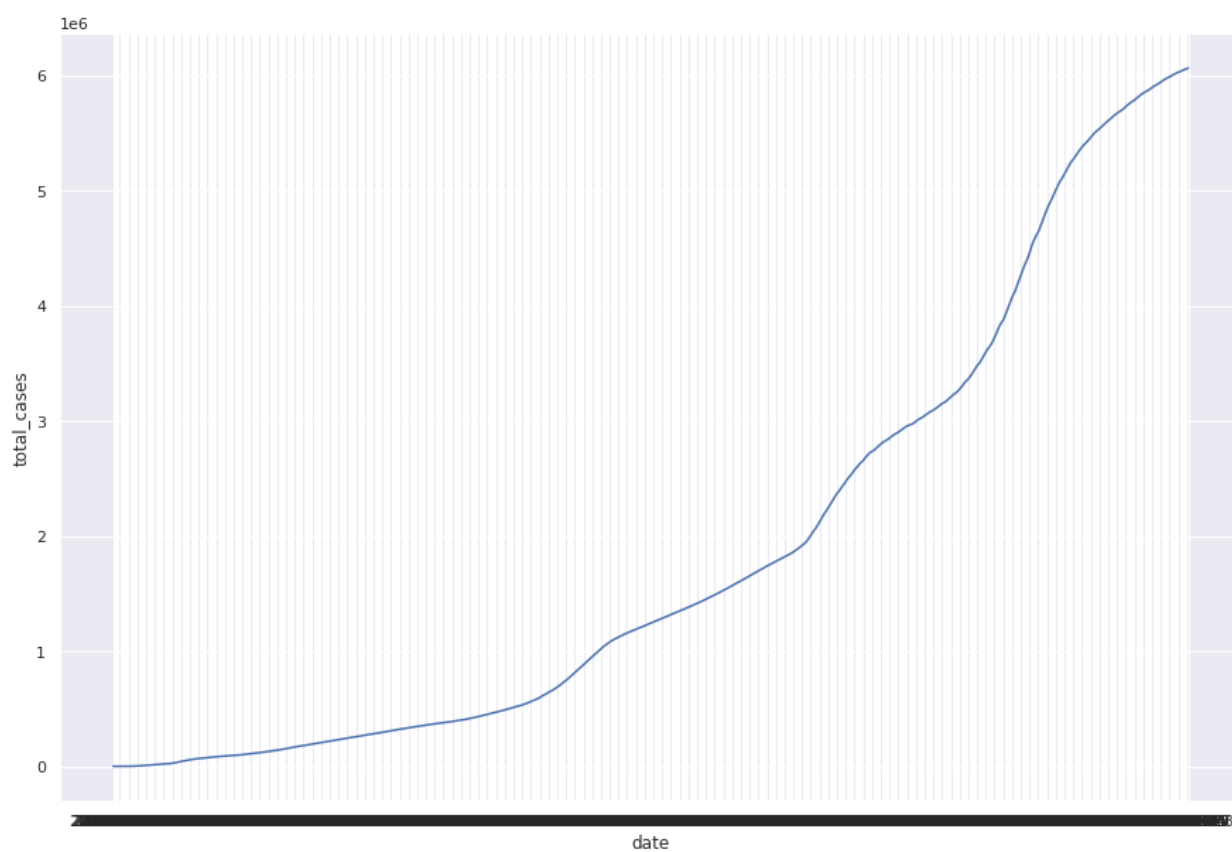


مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲



نمودار زیر همان نمودار بالا است اما فقط ۳۰ روز آخر را مورد بررسی قرار می دهیم.

به نام خدا

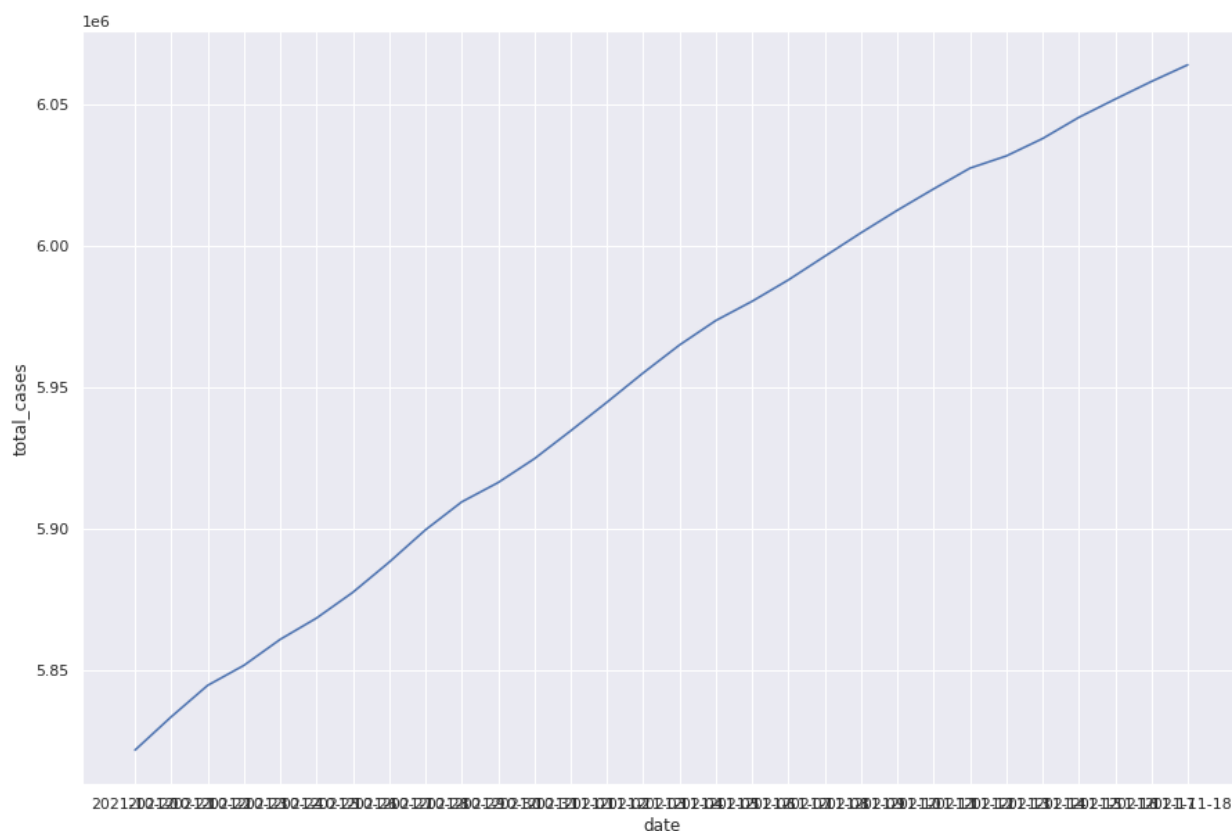


مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲



همان طور که مشاهده میکنید شیب نمودار تغییر خیلی محسوسی ندارد ولی کم کم در حال کم شدن است.

یعنی تعداد افراد مبتلا هر روز دارد کاهش می یابد.

در نمودار زیر total test را مشاهده میکنیم. این نمودار به ما نشان خواهد داد (با توجه به شیب که تقریباً

ثابت است) یک سری از افراد وجود دارند که به صورت مرتب تست می دهند.

به نام خدا

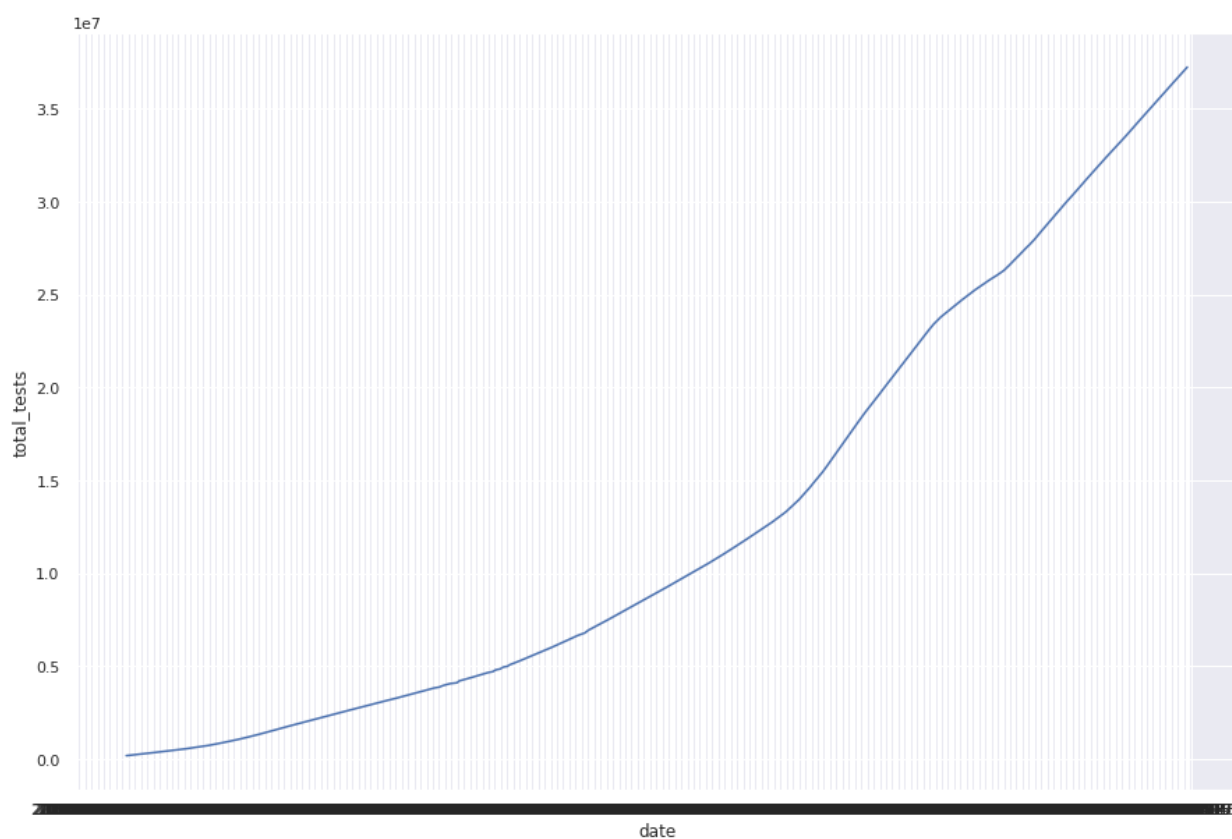


مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲



در نمودار بعد تعداد new cases smoothed را نمایش می دهیم که پیک های کرونا در تاریخ های مختلف را به خوبی نشان می دهد.

به نام خدا

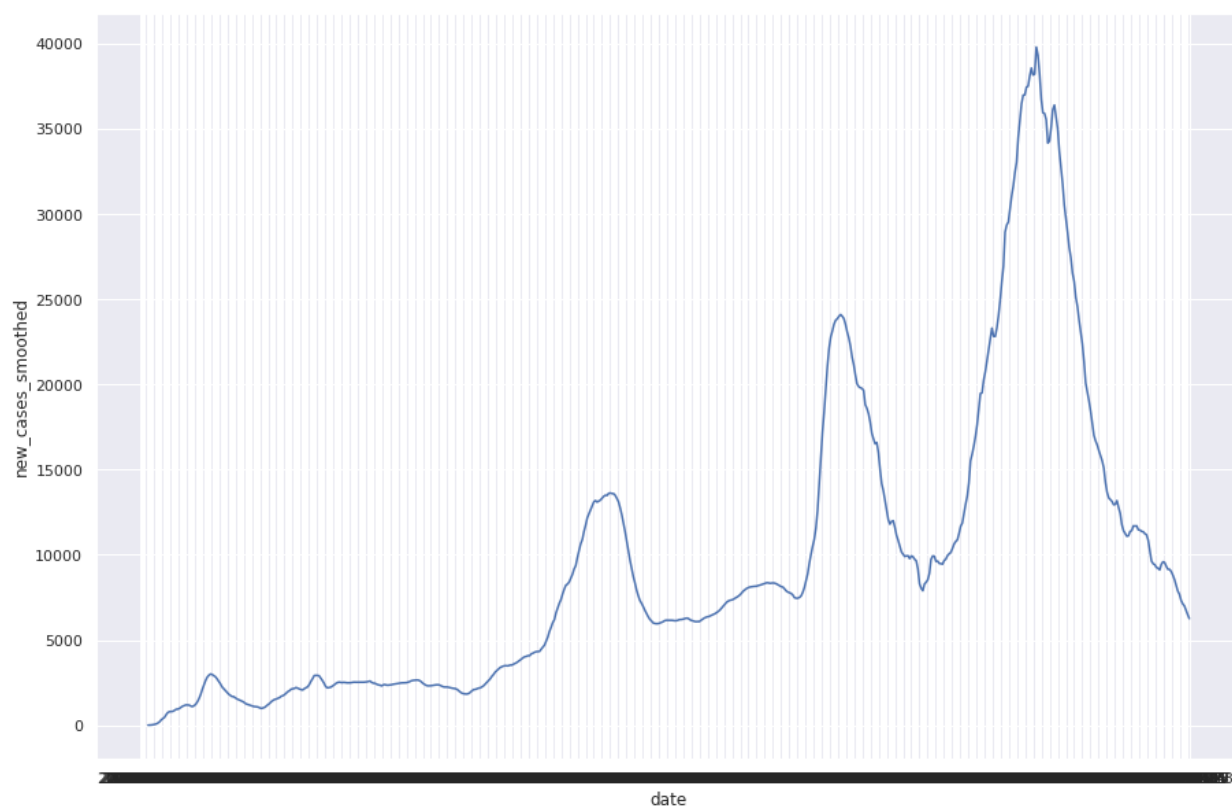


مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲



در ادامه ایران را با ۲ کشور چین به عنوان اولین کشور درگیر با کرونا و ژاپن به عنوان یک کشور موفق در کنترل کرونا (بر اساس اخبار خارج از این تحقیق) بررسی میکنیم. در نمودار پایین ۳ نمودار total cases گذر زمان را در این ۳ کشور مشاهده می کنید و درباره کیفیت کنترل کرونا میتوان در این ۳ کشور بحث کرد.

به نام خدا

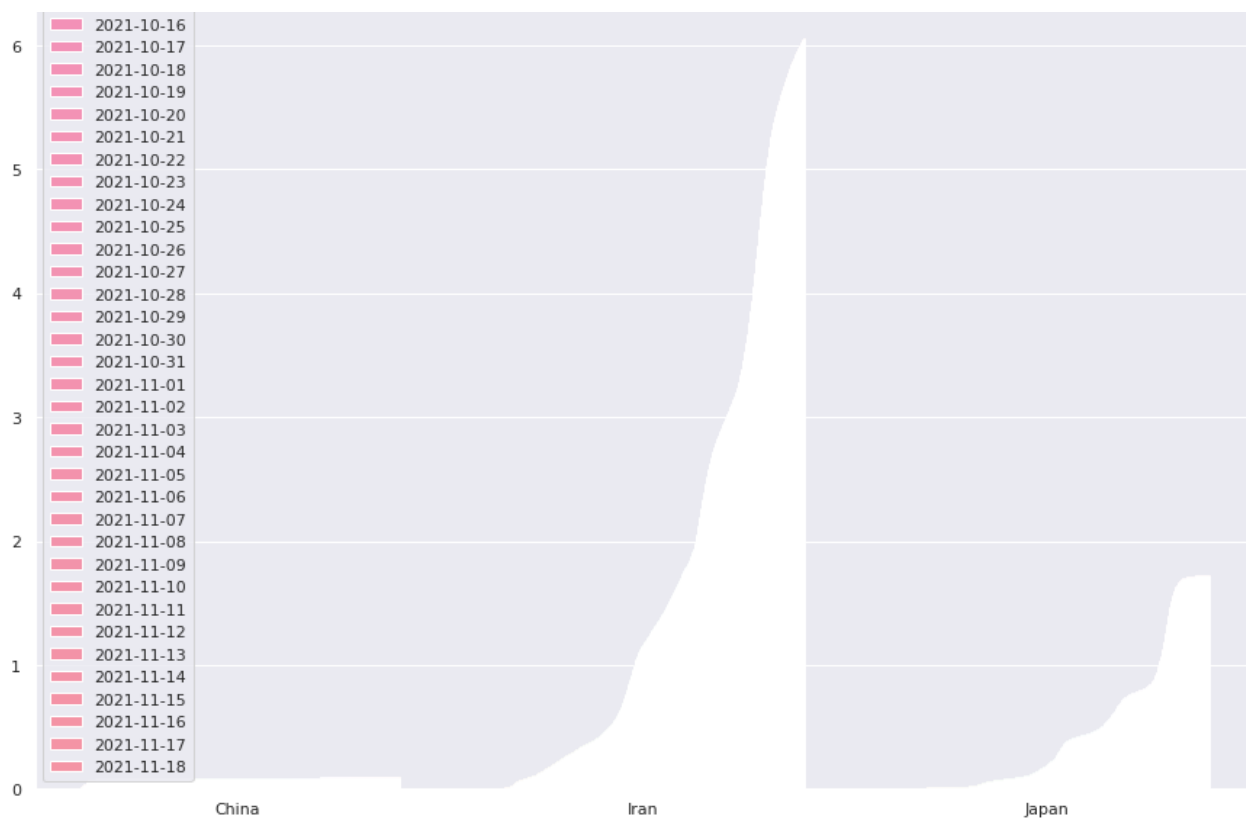


مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲



تاکید می شود که نمودار بالا total است و شیب نمودار بیشتر مورد توجه باید باشد.

۳.۲ تحلیل تمام کشور ها

نمودار زیر ۸ کشور (یا منطقه) با بیشترین تعداد ابتلا نسبت به بقیه کشور ها رسم میکنیم.

به نام خدا

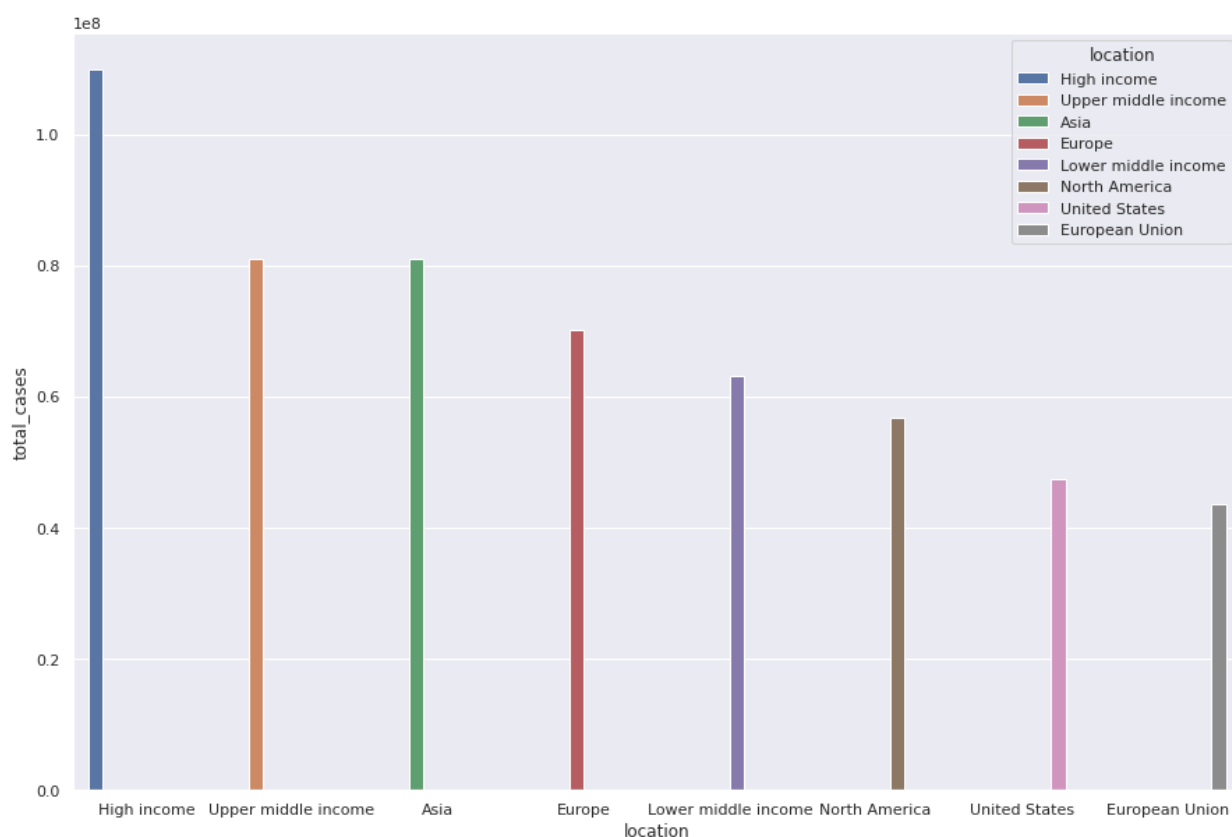


مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲



نقشه زیر نشان دهنده تعداد کل ابتلا در هر ۱ میلیون نسبت به هم در سرتاسر جهان است. در این عکس بعضی از کشورهایی که در پاکسازی داده حذف شده اند را با رنگ خاکستری مشاهده میکنید. این نمودار صرفاً با دیتای روز 2021-11-17 کشیده شده که زمان نوشتن این گزارش است.

به نام خدا

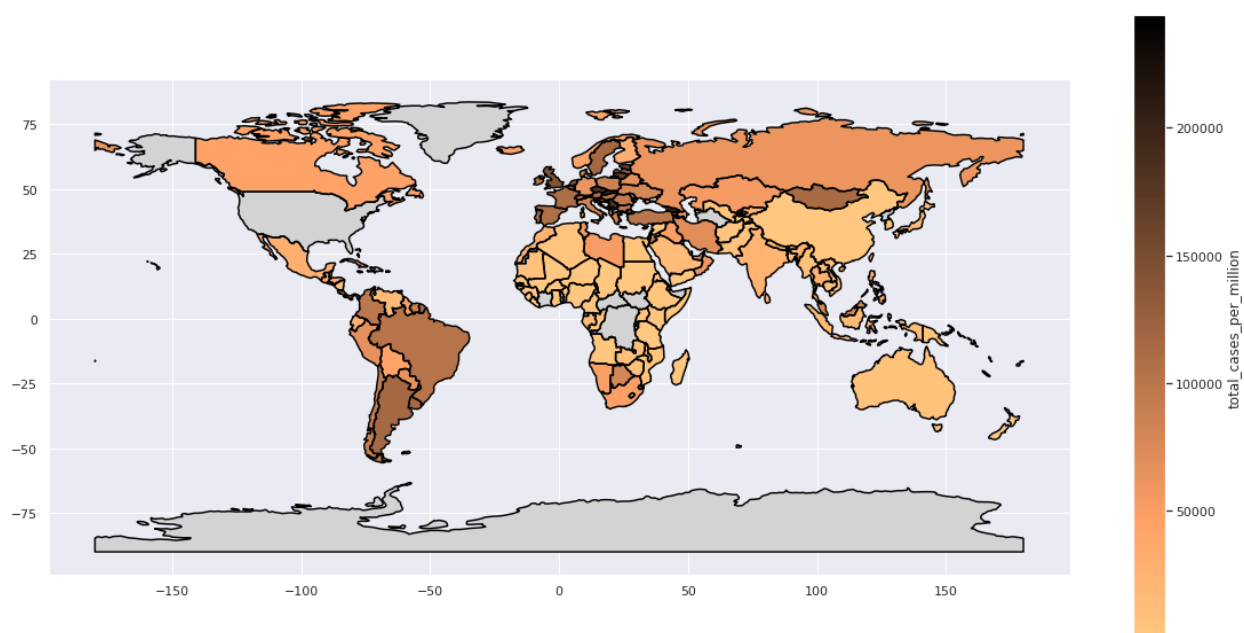


مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲



همان طور که مشاهده میکنید این نمودار خبر خارجی (اخبار خارج از این تحقیق) افزایش درگیری کرونا در اروپا را تایید می کند. چون اروپا رنگ های پر رنگ تری دارند پس یعنی افراد بیشتری در یک میلیون نسبت به بقیه جا های دنیا کرونا دارند. همچنین مشاهده می کنیم اوضاع در آمریکا جنوبی هم نسبت به بقیه جاها خیلی خوب نیست. در ادامه بقیه نمودار ها برای نشان دادن همین مورد کشیده شده اند. سمت راست نمودارها موردی که نمودار بر اساس آن کشیده است را نوشته شده است.

به نام خدا

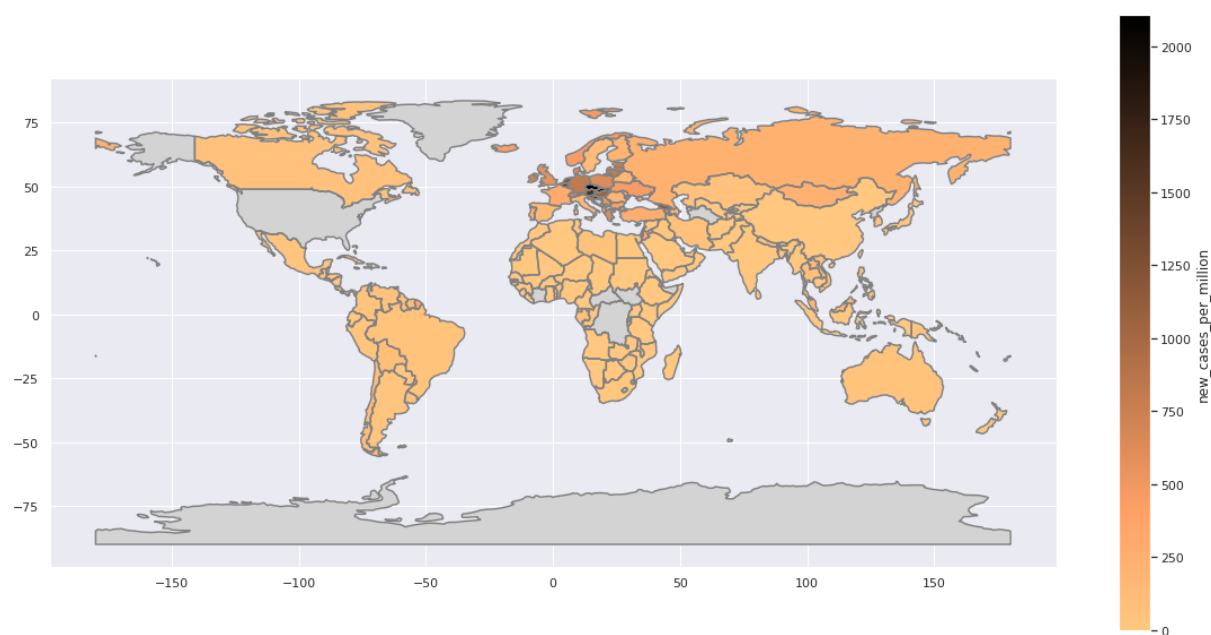


مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲



به نام خدا

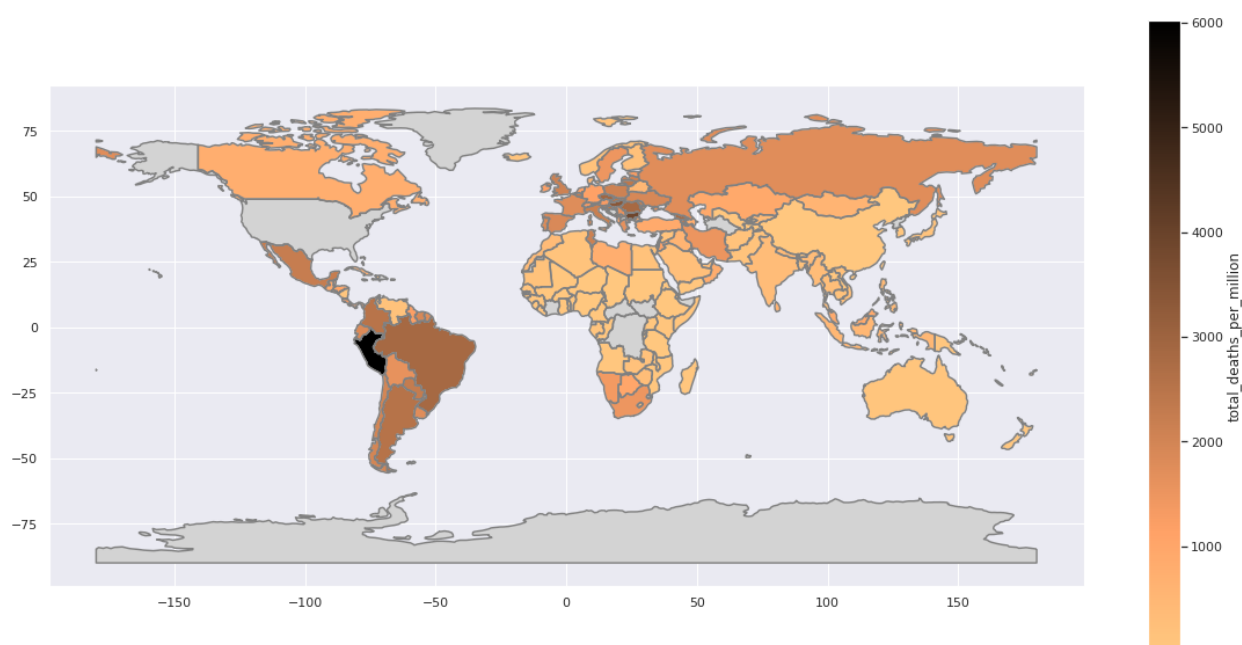


مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲



به نام خدا

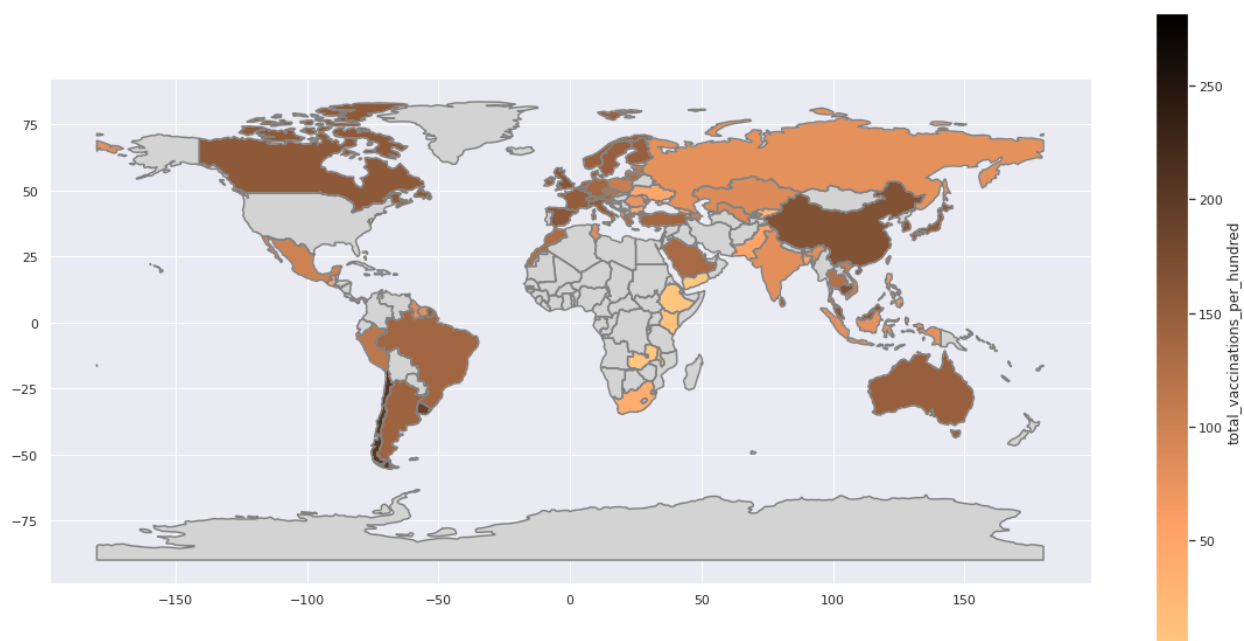


مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲



در نمودار بالا نقشه واکسن زدن بعضی از کشورها را می بینید که به جهت توجیه نمودارهای بالا آورده شده است. همان طور که میبینید نمیتوانیم رفتار کرونا را در نقاط مختلف جهان صرفاً بر اساس درصد واکسن شده ها و نشده ها توجیه کنیم. نمودار بعدی هم تعداد واکسن های booster را نشان می دهد. در این میان کشور شیلی به نظر پیشتاز در زمینه واکسن زدن است.

به نام خدا

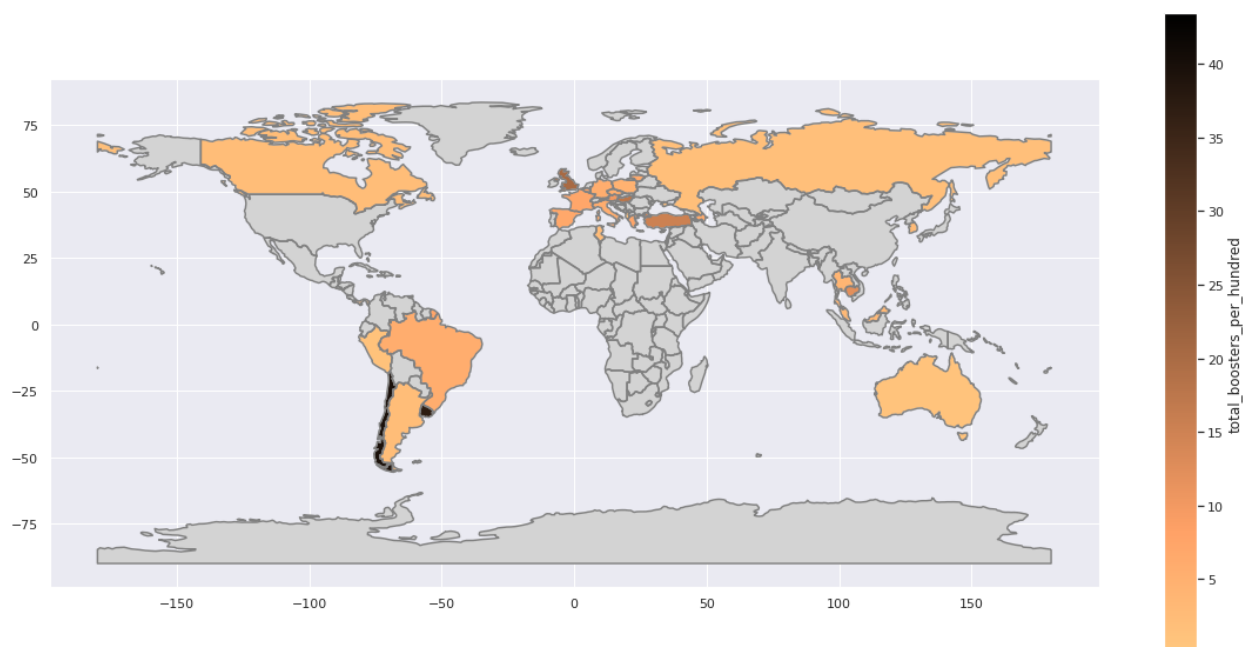


مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲



در نمودار بعد یک عامل مهم دیگر یعنی چگالی جمعیت را مشاهده میکنیم که در اکثر کشورها یکسان است.

به نام خدا

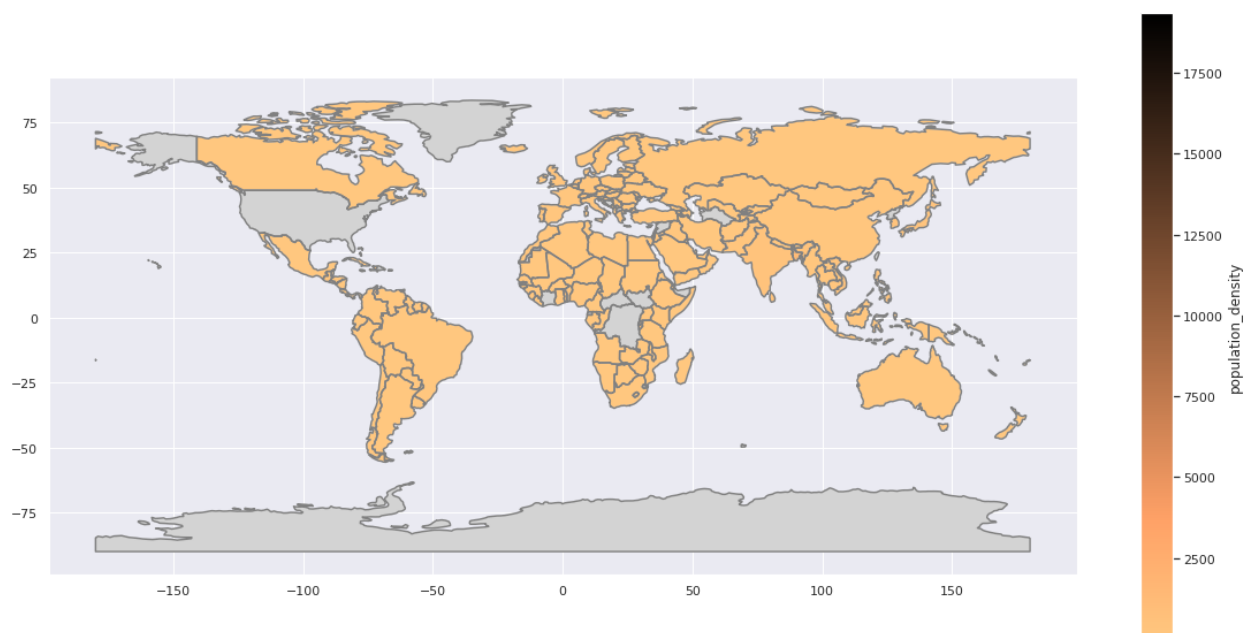


مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲



در مجموعه نمودارهای کشیده شده به صورت فیلم در زمان های مختلف تعداد new cases در کشورهای مختلف جهان نسبت به هم را مشاهده کردیم. در این مجموعه نشان داده می شود که کرونا به ترتیب از کدام کشورها شروع شد و به کدام کشورها در طی زمان رفت و دنیا را فرا گرفت. در اینجا صرفاً یک صفحه از این مجموعه را می آوریم.

به نام خدا



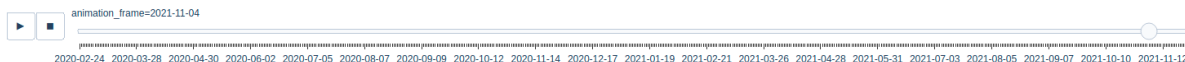
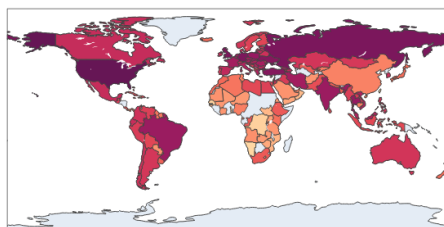
مبانی علوم داده

علی علی آبادی - ۹۷۲۲۰۶۱

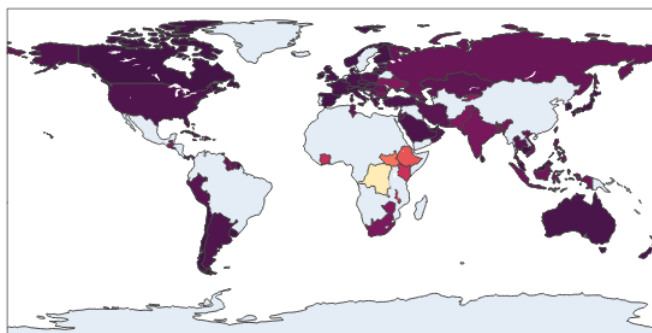
تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲

new cases over time



این نمودارها را برای people fully vaccinated per hundred هم کشیدیم که یک صفحه آن بدین صورت است. مشاهده کردیم که واکسیناسیون ابتدا به صورت گسترده در قاره آمریکا شروع شد و چطور در جهان پخش شد.



به نام خدا



مبانی علوم داده

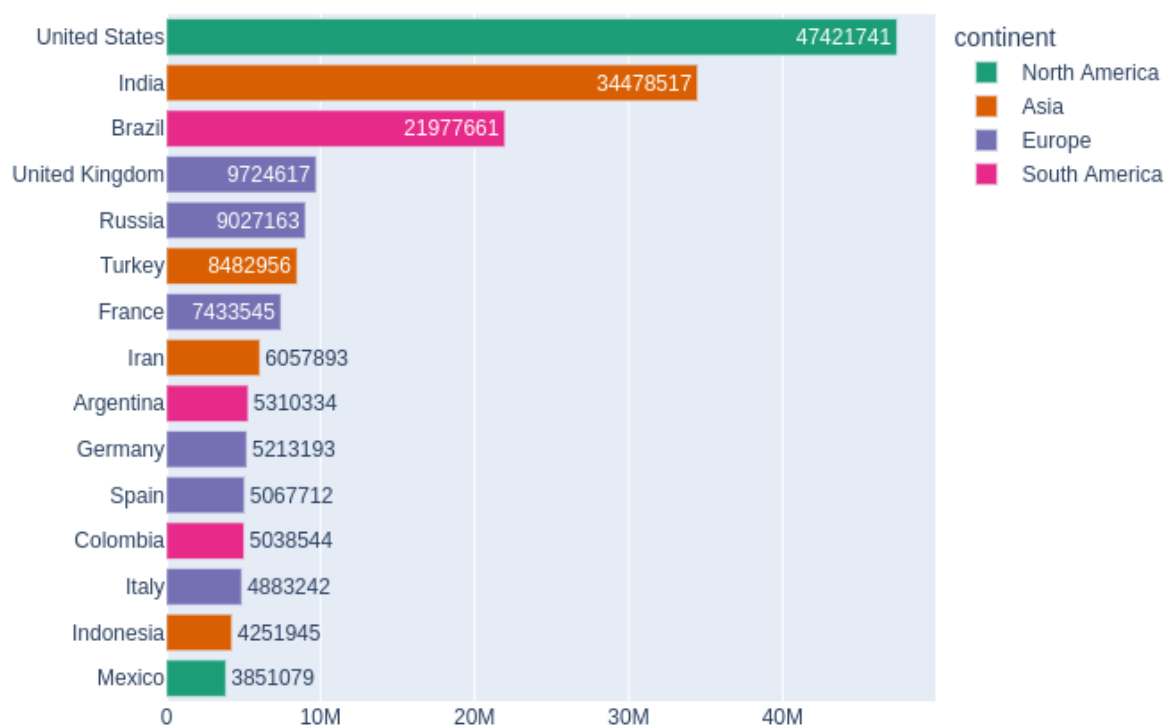
علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲

در ادامه به این نمودار ها توجه کنید. در این نمودار ها ۱۵ کشور اول در جهان به علاوه ایران اگر در جمع ۱۵ کشور نبوده آورده شده است. رنگ هر کدام نشان دهنده قاره آن کشور است. در این نمودار ها هم می بینیم که کشور های اروپایی زیادی وجود دارند.

total_cases



به نام خدا



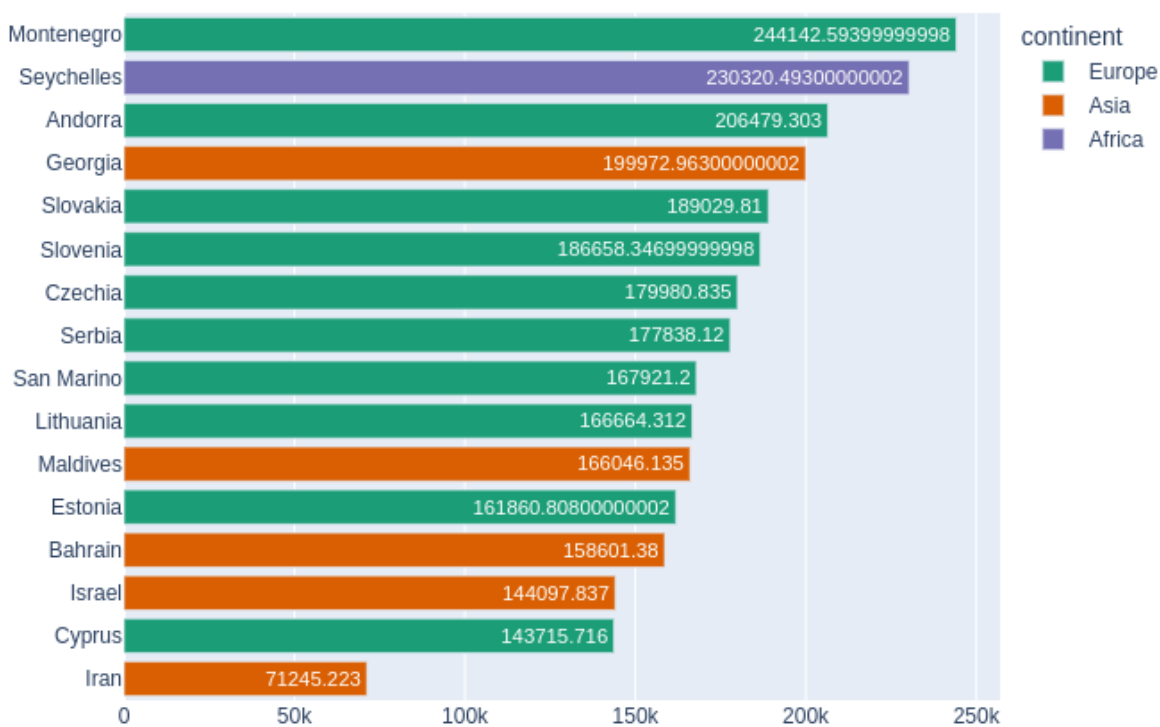
مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲

total_cases_per_million



به نام خدا



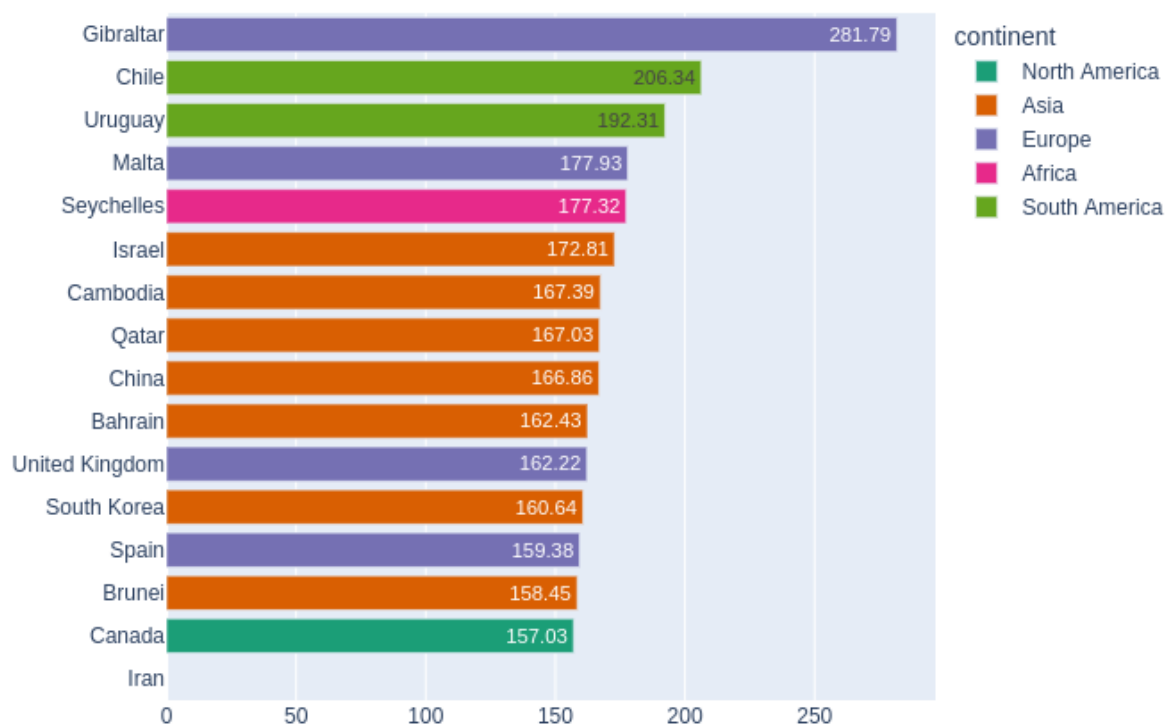
مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲

total_vaccinations_per_hundred



به نام خدا



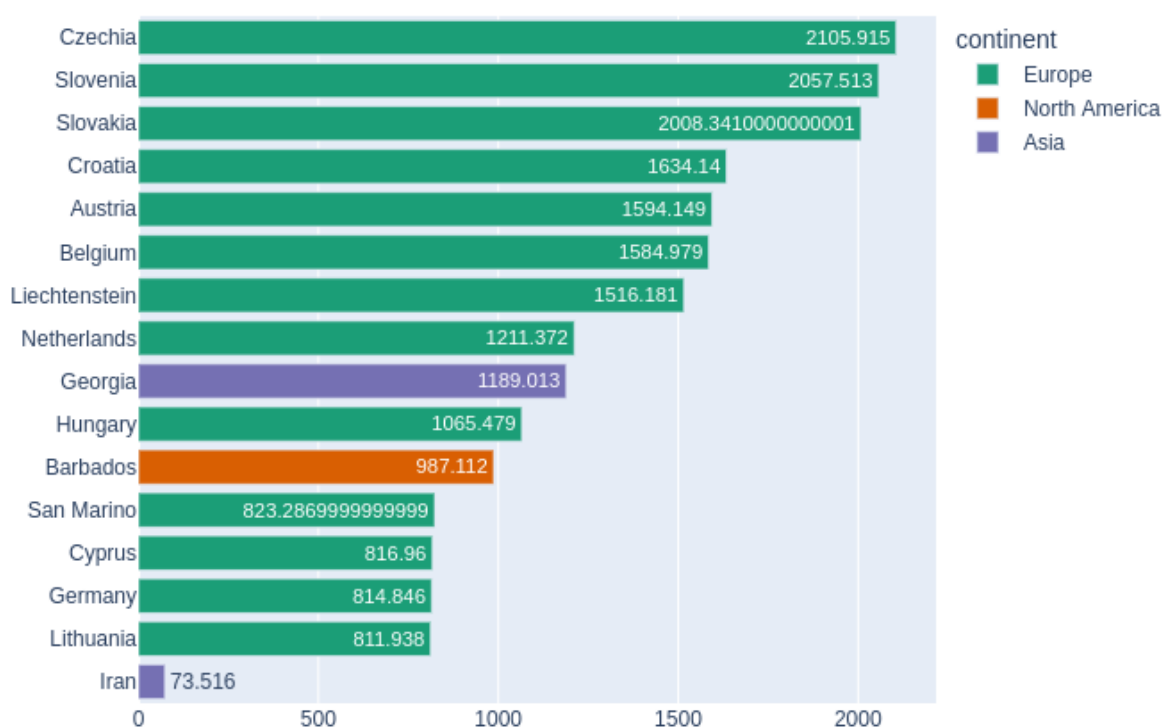
مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲

new_cases_per_million



به نام خدا



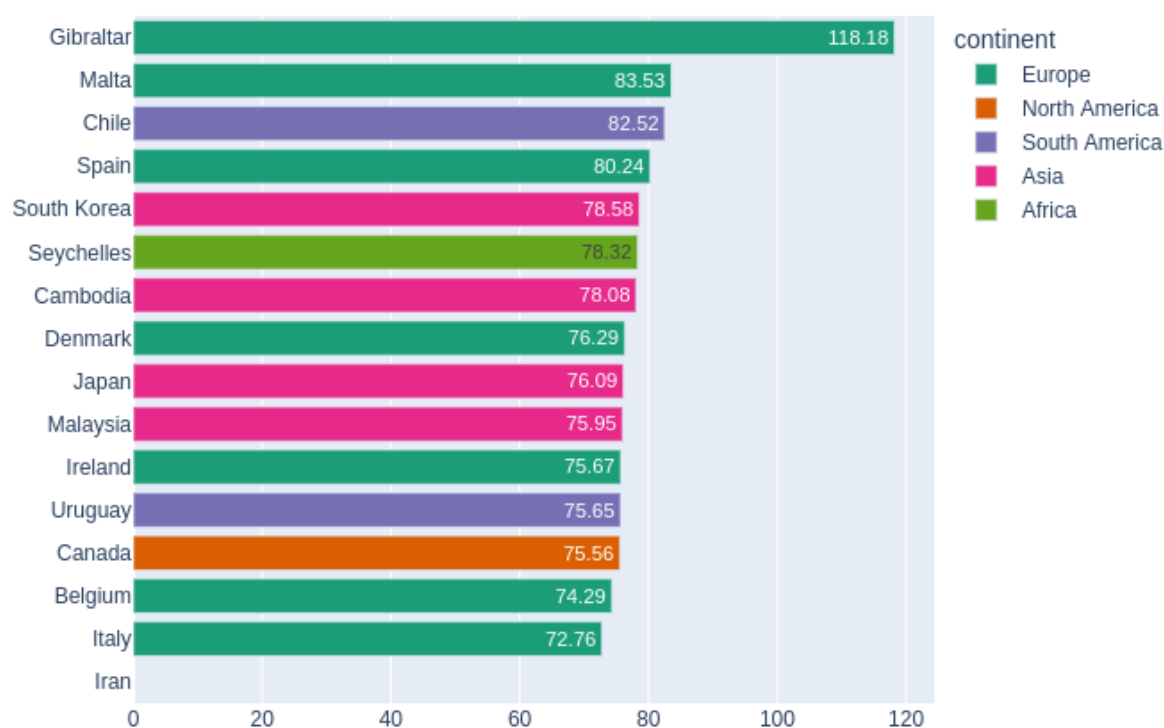
مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲

people_fully_vaccinated_per_hundred



در این نمودار مقدار ایران ۰ نیست صرفاً مقداری وجود ندارد (قابل انتظار است همان طور که در اخبار خارجی درصد واکسیناسیون جمعی در ایران گفته نمی شود). همان طور که حدس زدیم و در نمودار جغرافیایی دیدیم شیلی در زمینه واکسیناسیون پیشتاز است.

به نام خدا



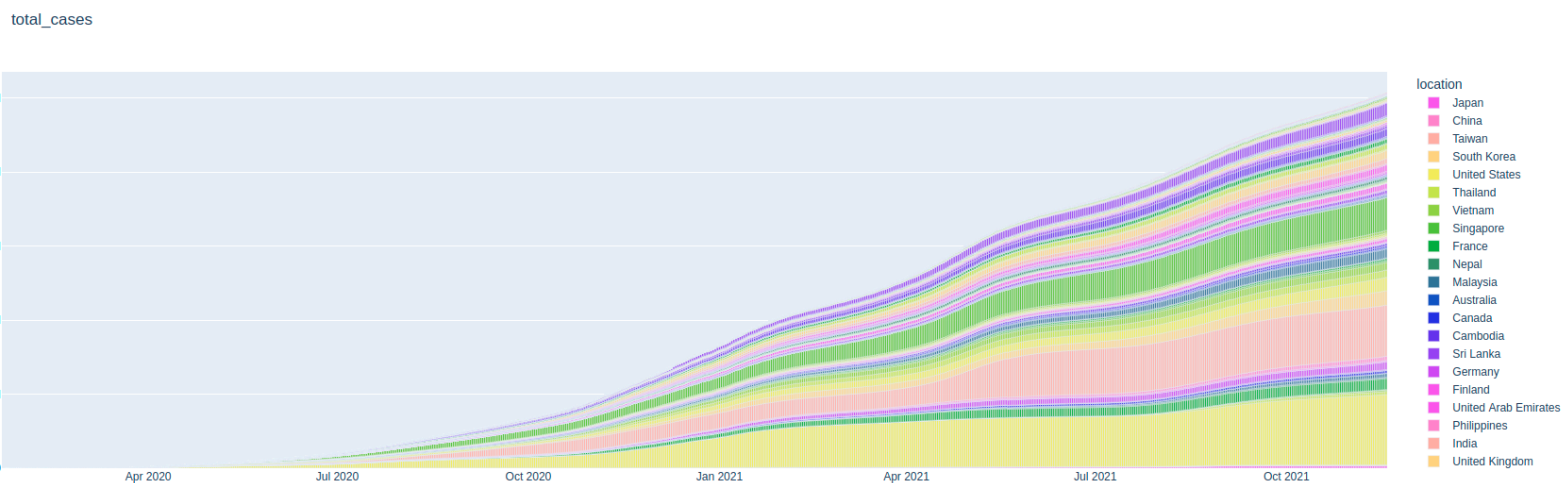
مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

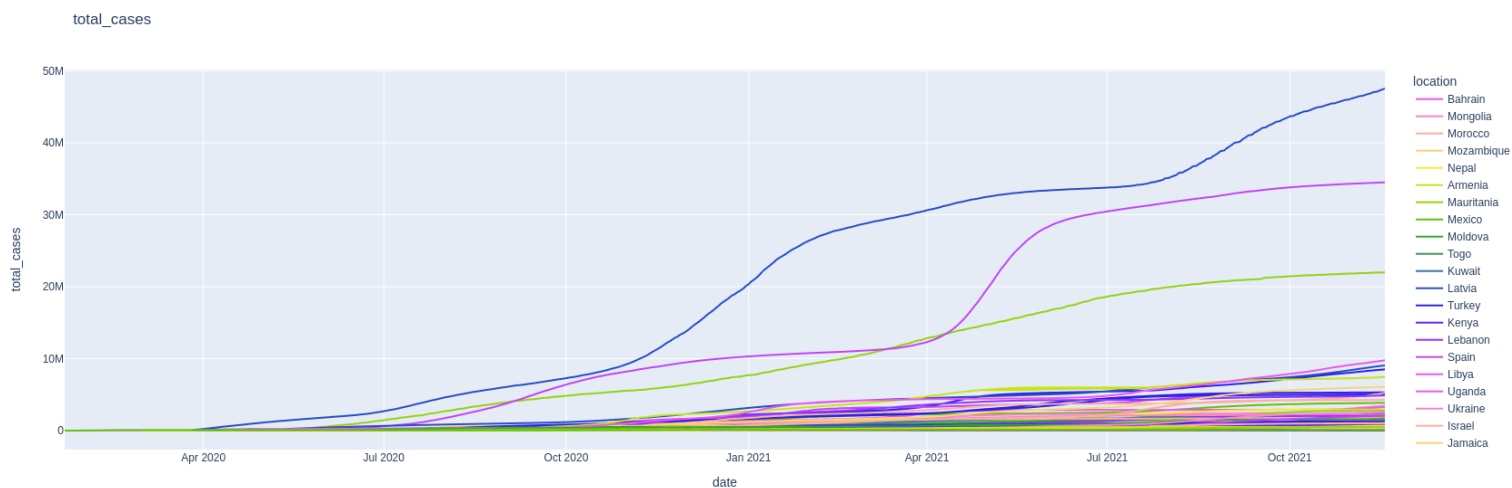
تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲

در ادامه نمودار total cases را به تفکیک کشورها در گذر زمان میبینیم که هر کشور چه سهمی از آن را دارد.



به صورت خطی همین اطلاعات را برای total cases و total deaths مشاهده می کنیم



به نام خدا

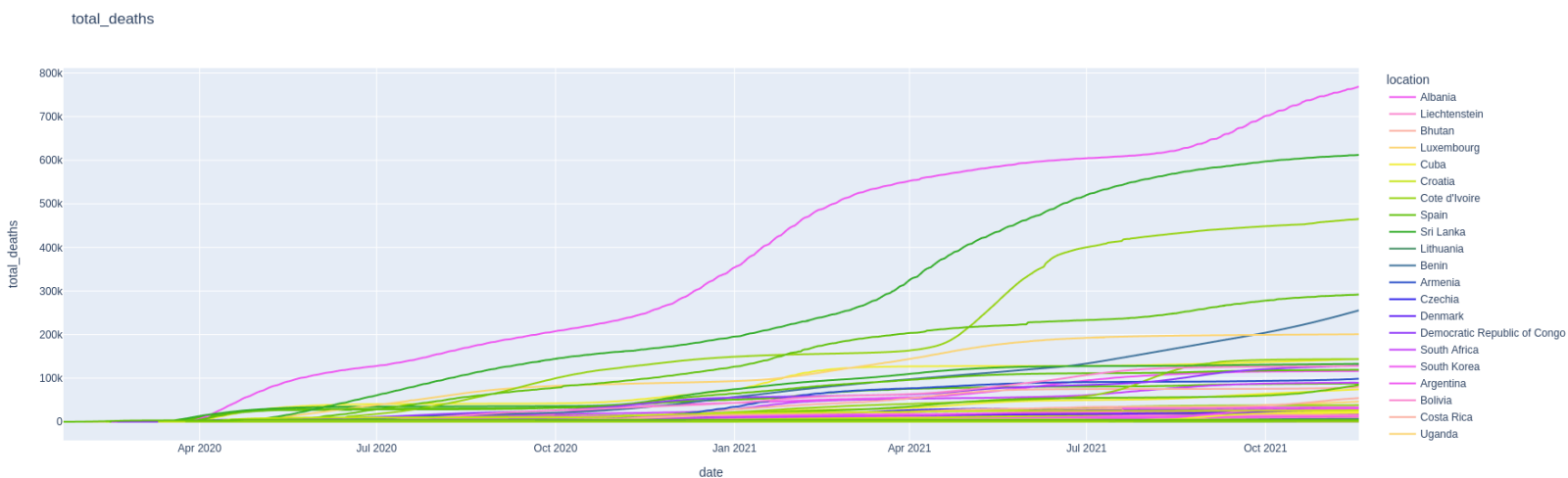


مبانی علوم داده

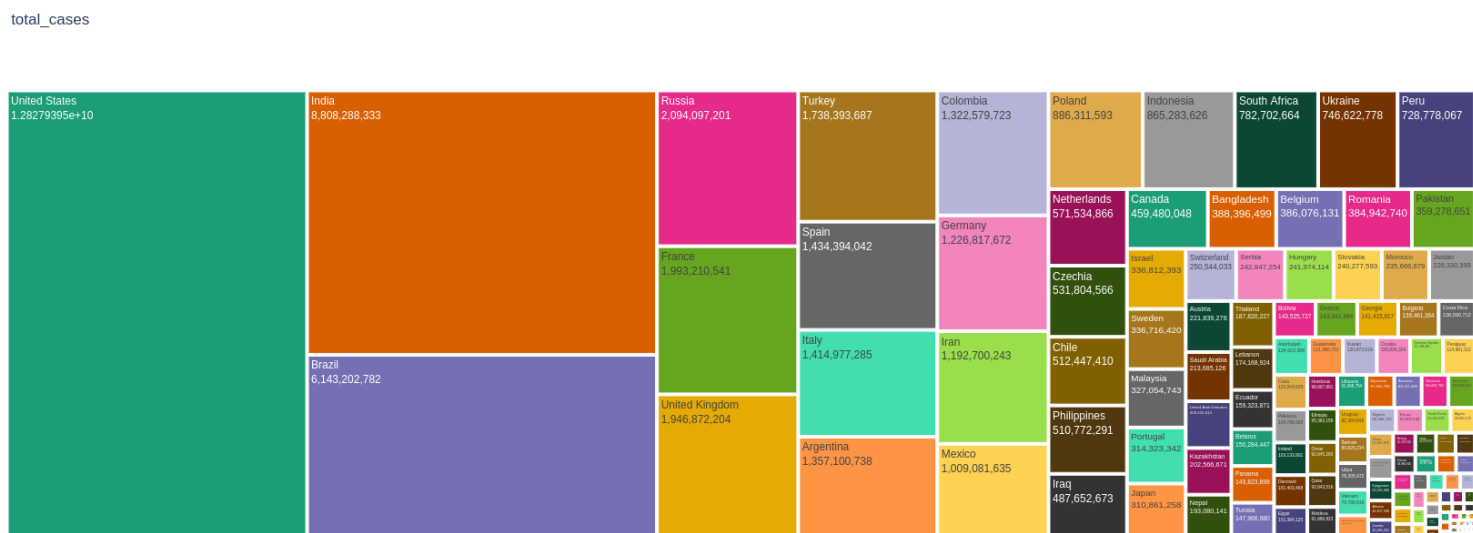
علی علی آبادی - ۹۷۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲



در ادامه به نمودارهای درختی زیر توجه کنید.



بہ نام خدا



مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲

total cases per million



people_fully_vaccinated_per_hundred



به نام خدا



مبانی علوم داده

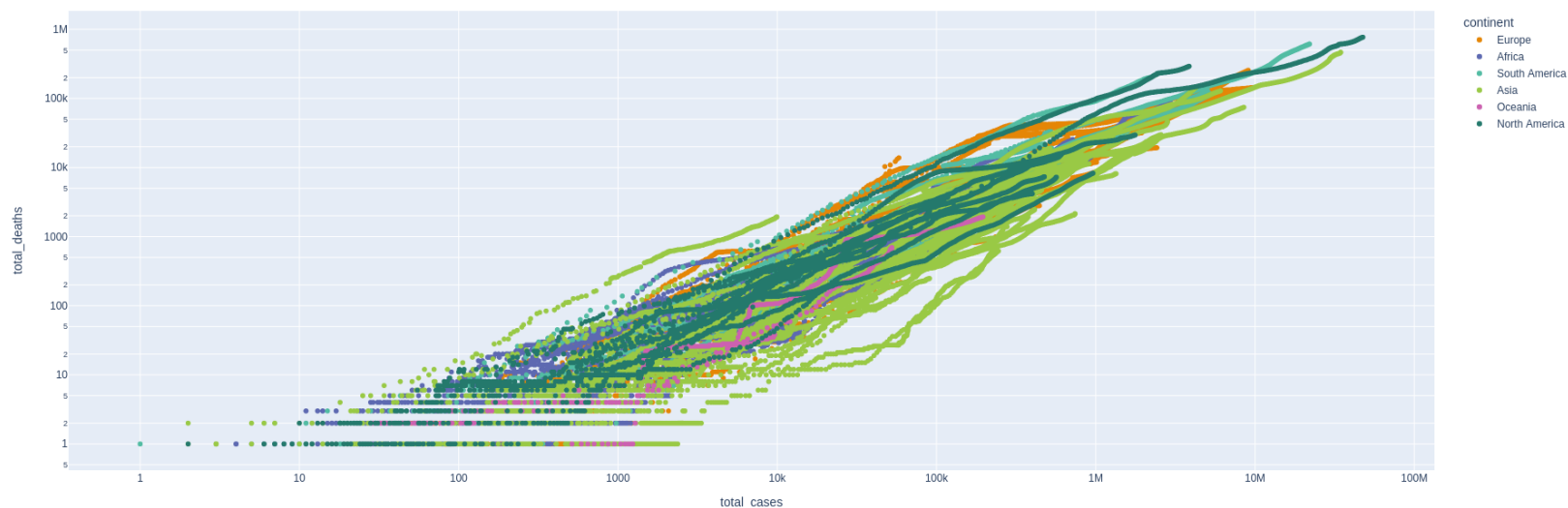
علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۰۰/۰۸/۲۷

گزارش تمرین سری ۲

کشورهای پیشتاز در زمینه واکسیناسیون در نمودارهای بالا قابل دیدن هستند. در انتها نیز نمودار total deaths بر اساس total cases را مشاهده می کنیم.

total vs Deaths



این داده ها بر اساس قاره به صورت رنگی تفکیک شده اند.

۴. منابع و لینک ها

محتوی به صورت با کیفیت و قابل مشاهده در [notebook](#) گذاشته شده است که به پیوست ارسال شده.

از این [notebook](#) نیز در تحلیل ها و انتخاب نمودارها استفاده شد.