

در این تمرین قصد داریم تا با استفاده از روش‌های مختلف مصورسازی داده به تحلیل داده بپردازیم.

نکته مهم در پیاده‌سازی تمرین این است که:

- در صورت لزوم تمیزسازی داده انجام شود و دلایل آن توضیح داده شود. برای مثال اگر بخش‌هایی از داده وجود ندارد، دلیل آن بیان شود و در صورت نیاز با روش مناسبی این مشکل برطرف گردد.
- برای انتخاب انواع مختلف نمودار و سایر فرض‌های انجام شده توضیحاتی ارائه نمایید.
- برای رسیدن به نتایج قابل اتکا لازمست feature های مناسبی انتخاب شوند تا اثر confounding factor ها به حداقل برسد. برای این مورد نیز در صورت نیاز دلایل خود را بیان کنید.
- در رسم نمودارها به رنگ‌ها، فونت‌ها، سایزها، تمیزی و... توجه شود. در صورتی که از زبان فارسی و فونت خاصی برای متن‌های به کار رفته در نمودارها استفاده کنید، نمره مثبت محسوب می‌شود.
- انتخاب نوع static یا interactive بودن نمودارها و انتخاب کتابخانه‌ها به عهده خودتان است.
- نمودارهای بدون تحلیل هیچ ارزشی ندارند و در هر بخش باید تحلیل خود را ارائه کنید.

## دریافت داده

با توجه به شیوع ویروس کرونا سعی داریم که داده‌های کرونا در جهان را بررسی کنیم. شما می‌توانید از [این لینک](#) داده‌های مورد نظر را دانلود کنید. توضیحات درباره feature های مختلف این داده را نیز می‌توانید از [اینجا](#) مطالعه کنید. سپس ستون‌هایی را که اطلاعات در هر میلیون دارند را حذف کنید.

## بخش اول - بررسی کرونا در ایران

۱. نمودار تعداد مبتلایان، تعداد فوتی‌ها و تعداد واکسینه شده‌های ایران را بصورت هفتگی در طول زمان در یک نمودار رسم کنید. در این نمودار تاریخ‌ها را به فرمت شمسی بنویسید و axis grid را هم به ازای هر ماه رسم کنید. اگر داده‌های مربوط به این هفته هنوز کامل نشده است، باید این هفته آخر را بصورت متمایزی نمایش دهید. سپس نمودار را تحلیل کنید.
۲. روند سختگیری در حوزه کرونا در ایران را در طول زمان بررسی کنید. این نمودار را با کتابخانه plotly و با اضافه کردن یک range slider نمایش دهید.
۳. فرض کنید ما زمانی که تعداد مبتلایان به کرونا افزایش یابد را به عنوان شروع یک پیک در نظر بگیریم. بررسی کنید که در زمان پیک کرونا و در شرایط عادی تر میزان سختگیری دولت چه تفاوتی داشته است؟ آیا میزان سختگیری کم باعث ایجاد موج جدید می‌شود؟

۴. قصد داریم سرعت واکسیناسیون در ایران را با ۵ کشور دلخواه مقایسه کنیم. نموداری ارائه دهید.

### بخش دوم - مقایسه کشورهای مختلف

۱. تاثیر ویژگی‌های تراکم جمعیت، میانگین سنی، وجود امکانات بهداشتی، تعداد تخت بیمارستان‌ها و شاخص پیشرفت انسانی را بر تعداد فوتی‌ها با رسم correlation heatmap بررسی کنید. در این نمودار باید color bar مشخص باشد و min آن 1- به رنگ قرمز و max آن 1+ به رنگ آبی باشد. همچنین بر روی color bar عنوان داشته باشد.
۲. بصورت ساده نمایش دهید که چند درصد مردم در جهان دز اول، دز دوم و یا دز booster واکسن را دریافت کرده اند.
۳. کدام کشورها بهترین و کدام کشورها بدترین عملکرد در مهار ویروس کرونا را داشته اند؟ (منظور از عملکرد، تعداد فوتی نسبت به کل جمعیت است)
۴. می‌خواهیم تاثیر واکسیناسیون بر تعداد فوتی‌ها را بررسی کنیم. برای این کار فرض کنید لازمست که اطلاعات ۱۰ کشور را بررسی کنیم. شما کدام کشورها را برای مقایسه انتخاب می‌کنید؟ چرا؟ نمودارهای مرتبط برای رسیدن به این نتیجه را رسم و تحلیل کنید. (در این نوع نمودارها توجه کنید که بهتر است از تعداد فوتی‌ها برای تحلیل استفاده کنیم و یا تعداد فوتی‌ها در جمعیت)
۵. رابطه بین وضعیت اقتصادی کشورها و تعداد افراد واکسینه شده را بررسی کنید. در این نمودار کشور ایران را بصورت متمایزی نمایش دهید.
۶. تعداد فوتی‌های شش ماه اخیر کشورهای مختلف را به نسبت جمعیت آن‌ها بر روی نقشه نمایش دهید. Color bar هم به نمودار اضافه کنید. (می‌توانید از لینک‌های ۱ و ۲ برای رسم این نمودار کمک بگیرید)
۷. مشابه نمودار قبل، تعداد کل افراد واکسینه شده در کشورها را بر روی نقشه نمایش دهید. در قسمت tool tip هر کشور مقدار جمعیت آن کشور را هم بنویسید.
۸. برای سوالهای مشابه با سوال ۶ و ۷ استفاده از coropleth map و یا از scatter plot (همانند مثال اول [این لینک](#)) را پیشنهاد می‌کنید؟ هر کدام برای چه مسائلی مناسب هستند؟



### بخش سوم - نتیجه گیری کلی و تحلیل

به این موارد با تحلیل داده‌های موجود پاسخ دهید:

۱. به نظر شما در مهار کرونا، واکسینه شدن موثر است؟ اگر جواب مثبت است، میزان سرعت واکسیناسیون و یا جمعیت واکسینه شده موثر خواهد بود؟ اگر واکسن موثر است پس چرا همچنان تعداد فوتی‌ها بالاست؟
۲. شما به عنوان یک تحلیلگر داده و بدون استفاده از هیچ گونه مدل پیش بینی کننده ای، انتظار دارید که در روزهای آتی در ایران موج بعدی شروع شود؟ اگر بله تقریباً از چه زمانی شروع می‌شود؟



## نکات تحویل

- مهلت ارسال این تمرین تا پایان روز جمعه ۲۱ آبان ماه خواهد بود.
- انجام این تمرین به صورت یک نفره می باشد.
- لطفا هر گونه فرض در حل سوالات را در گزارش خود ذکر کنید.
- لطفا گزارش ، فایل کدها و سایر ضمائم مورد نیاز را با فرمت زیر در سامانه مدیریت دروس بارگذاری نمائید.

HW3\_[Lastname]\_[StudentNumber].zip

- در صورت وجود سوال و یا ابهام می توانید از طریق رایانامه زیر با دستیار آموزشی در ارتباط باشید:  
[m.vatandoust@ut.ac.ir](mailto:m.vatandoust@ut.ac.ir)