

مبانی علوم داده

على على آبادي - ٩٧٢٢٠۶١

تاریخ: ۲۷/۰۸/۲۷

گزارش تمرین سری ۲

#### ١. 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 کشور ها 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 کشور ها تا یک تاریخ خاص null و بعد از آن شروع به پر شدن کرده است. در این حالت این سلول ها را با و 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 •



مباني علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۲۷/۰۸/۲۷

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.

٣. تحليل داده ها

۳.۱ ایران

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

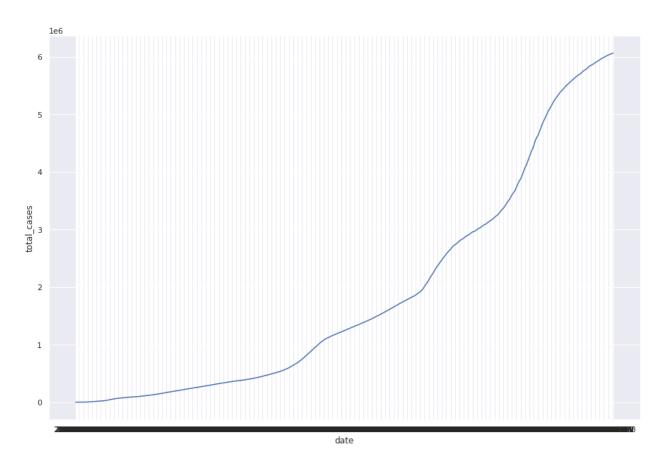


مباني علوم داده

على على آبادي - ٩٧٢٢٢٠۶١

تاریخ: ۲۷/۰۸/۲۷

گزارش تمرین سری ۲



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

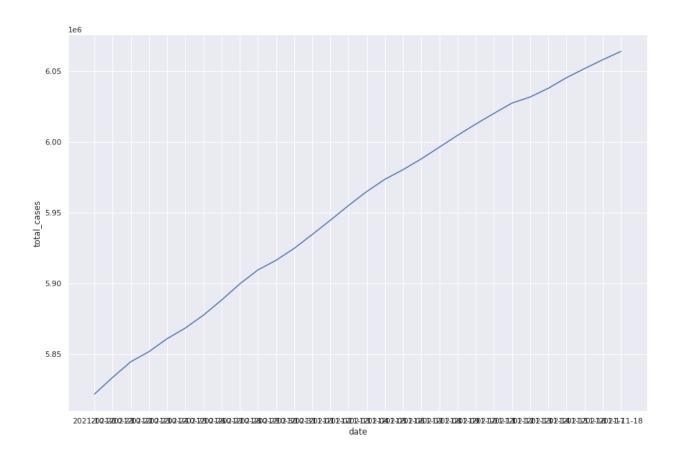


مباني علوم داده

على على آبادي - ٩٧٢٢٠۶١

تاریخ: ۲۷/۰۸/۲۷

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

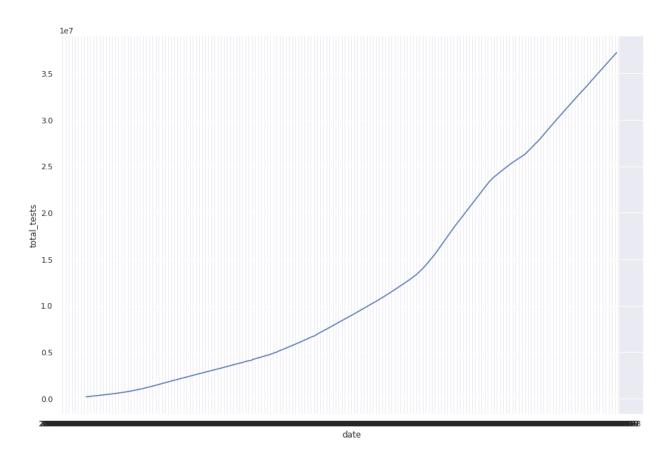


مباني علوم داده

على على آبادي - ٩٧٢٢٢٠۶١

تاریخ: ۲۷/۰۸/۲۷

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.

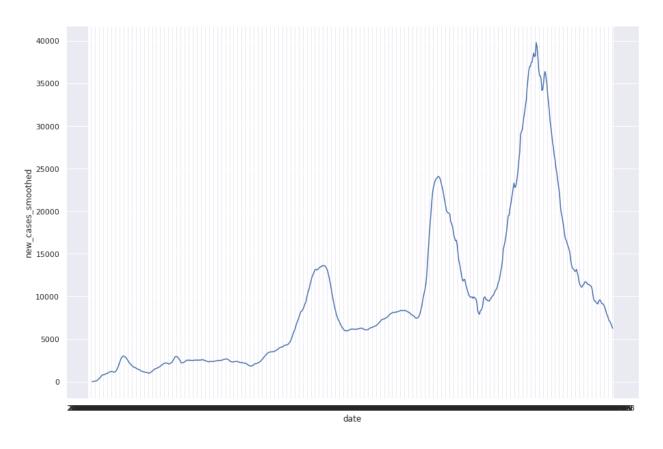


مبانی علوم داده

على على آبادي - ٩٧٢٢٢٠۶١

تاریخ: ۲۷/۰۸/۰۰

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 مشاهده می کنید و درباره کیفیت کنترل کرونا میتوان در این ۳ کشور بحث کرد.

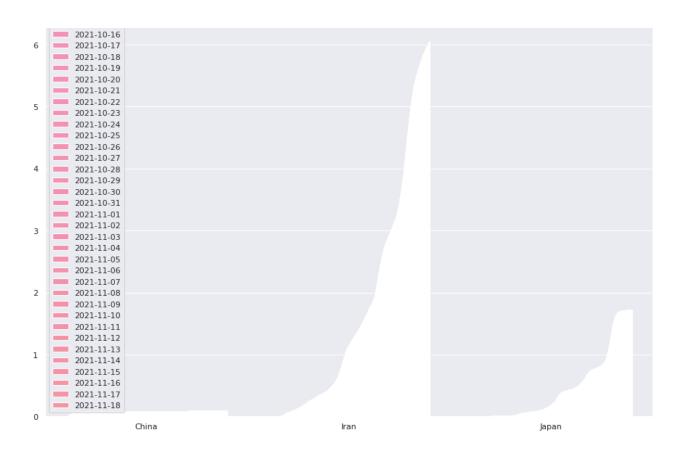


مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۲۷/۰۸/۲۷

گزارش تمرین سری ۲



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

٣.٢ تحليل تمام كشور ها

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

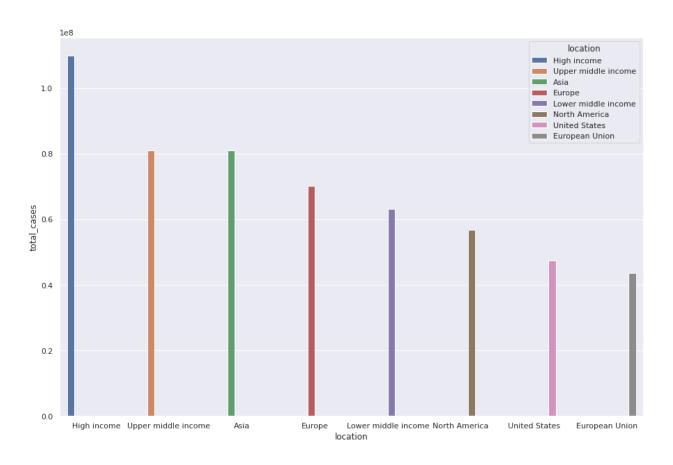


مباني علوم داده

تاریخ: ۲۷/۰۸/۲۷

على على آبادي - ٩٧٢٢٢٠۶١

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 کشور هایی که در پاکسازی داده حذف شده اند را با رنگ خاکستری مشاهده میکنید. این نمودار صرفا با دیتای روز 17-11-2021 کشیده شده که زمان نوشتن این گزارش است.

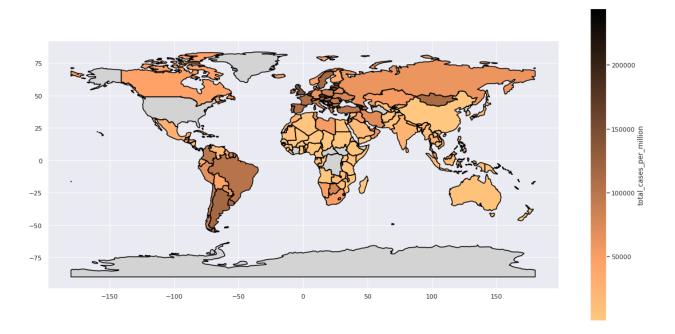


مبانی علوم داده

على على آبادي - ٩٧٢٢٢٠٤١

تاریخ: ۲۷/۰۸/۰۰

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 جاها 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.

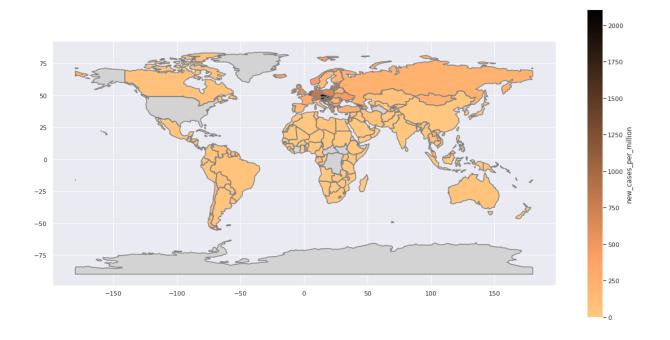


مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاريخ: ۲۷/۰۸

گزارش تمرین سری ۲



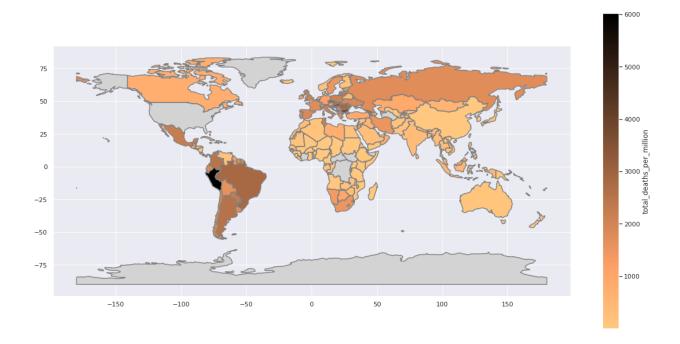


مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاریخ: ۲۷/۰۸

گزارش تمرین سری ۲



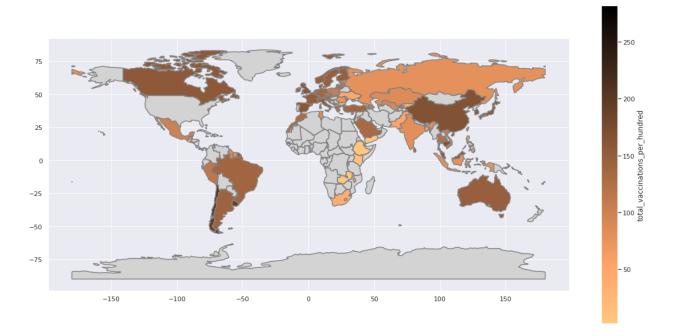


مبانی علوم داده

على على آبادي - ٩٧٢٢٢٠۶١

تاریخ: ۲۷/۰۸/۰۰

گزارش تمرین سری ۲



در نمودار بالا نقشه واکسن زدن بعضی از کشور ها را می بینید که به جهت توجیه نمودارهای بالا آورده شده است. همان طور که میبینید نمیتوانیم رفتار کرونا را در نقاط مختلف جهان صرفا 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.

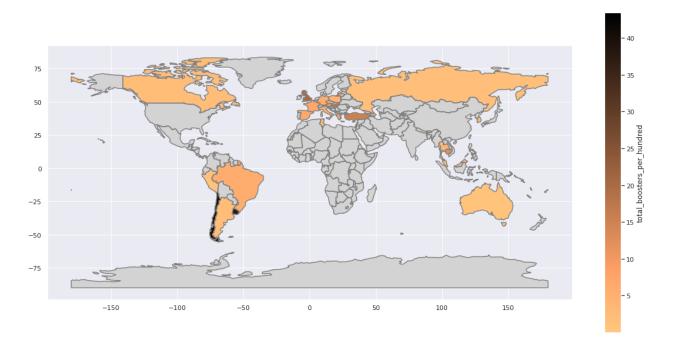


مباني علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاريخ: ۲۷/۰۸/۲۷

گزارش تمرین سری ۲



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

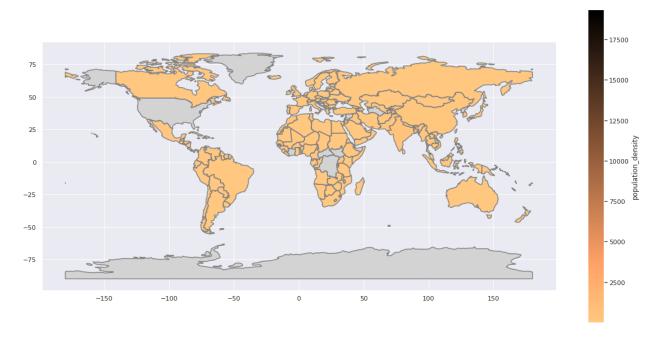


مباني علوم داده

تاریخ: ۲۷/۰۸/۰۰

گزارش تمرین سری ۲

على على آبادي - ٩٧٢٢٢٠۶١



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



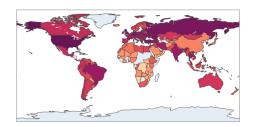
مبانی علوم داده

علی علی آبادی - ۹۷۲۲۲۰۶۱

تاريخ: ۲۷/۰۸/۲۷

گزارش تمرین سری ۲

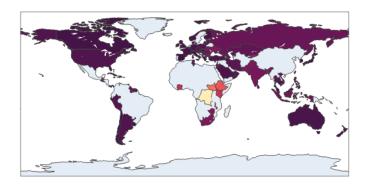
new cases over time



animation\_frame=2021-11-04

2020-02-24 2020-03-28 2020-04-30 2020-06-02 2020-07-05 2020-08-07 2020-08-09 2020-10-12 2020-11-14 2020-12-17 2021-01-19 2021-02-21 2021-03-26 2021-06-28 2021-06-31 2021-07-03 2021-08-05 2021-08-05 2021-08-07 2021-10-10 2021-11-12

این نمودار ها 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.



animation\_frame=2021-11-09



مبانی علوم داده

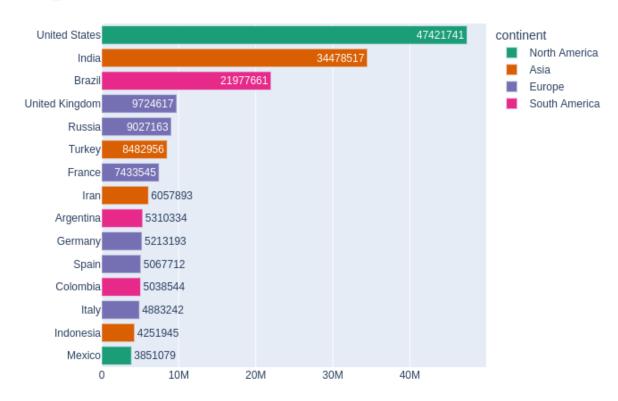
على على آبادي - ٩٧٢٢٢٠۶١

تاریخ: ۲۷/۰۸/۲۷

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





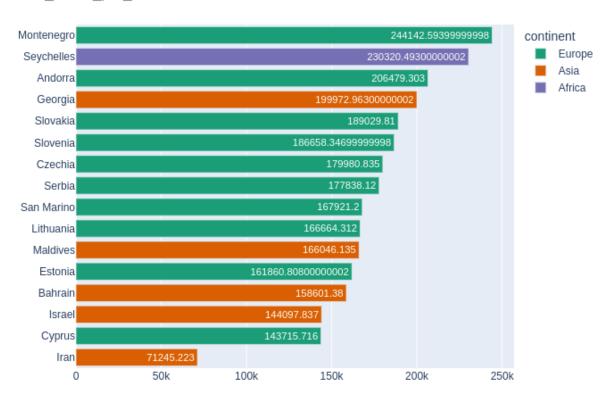
مبانی علوم داده

على على آبادي - ٩٧٢٢٠۶١

تاریخ: ۲۷/۰۸

گزارش تمرین سری ۲

#### total\_cases\_per\_million





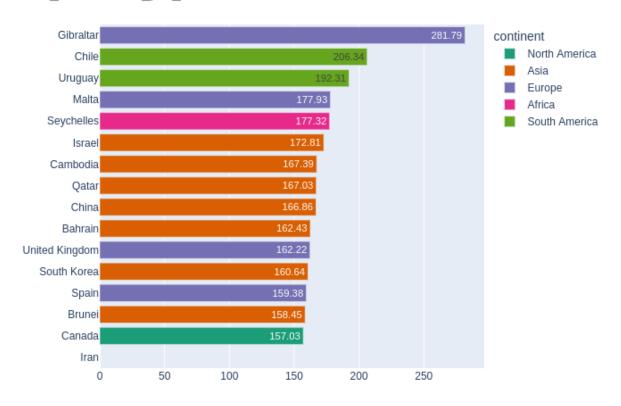
مباني علوم داده

على على آبادي - ٩٧٢٢٢٠۶١

تاریخ: ۲۷/۰۸/۲۷

گزارش تمرین سری ۲

#### total\_vaccinations\_per\_hundred





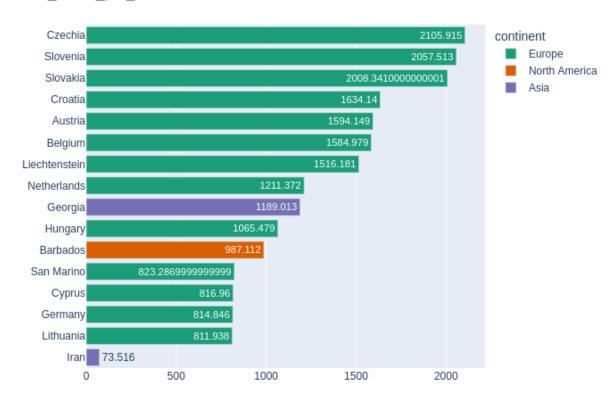
مبانی علوم داده

على على آبادي - ٩٧٢٢٢٠۶١

تاریخ: ۲۷/۰۸

گزارش تمرین سری ۲

#### new\_cases\_per\_million





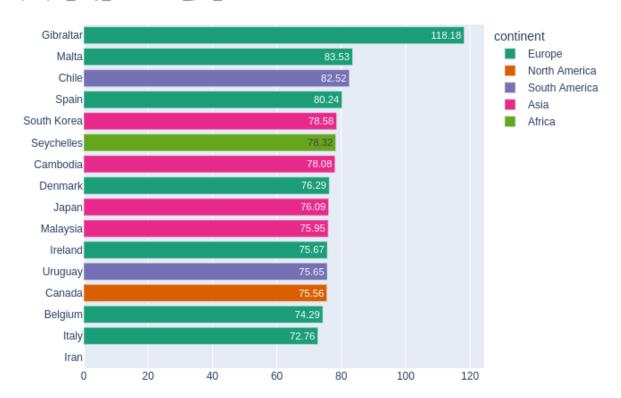
مبانی علوم داده

على على آبادي - ٩٧٢٢٢٠۶١

تاریخ: ۲۷/۰۸/۲۷

گزارش تمرین سری ۲

#### people\_fully\_vaccinated\_per\_hundred



در این نمودار مقدار ایران · نیست صرفا 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.



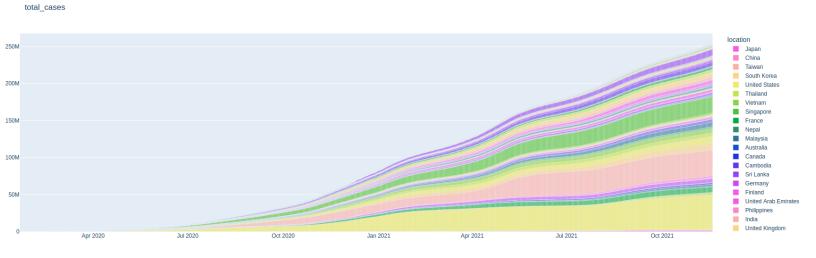
مبانی علوم داده

على على آبادي - ٩٧٢٢٢٠۶١

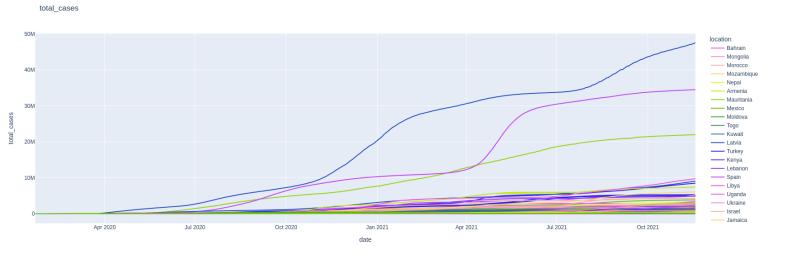
تاریخ: ۲۷/۰۸/۲۷

گزارش تمرین سری ۲

در ادامه نمودار total cases را به تفکیک کشور ها 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





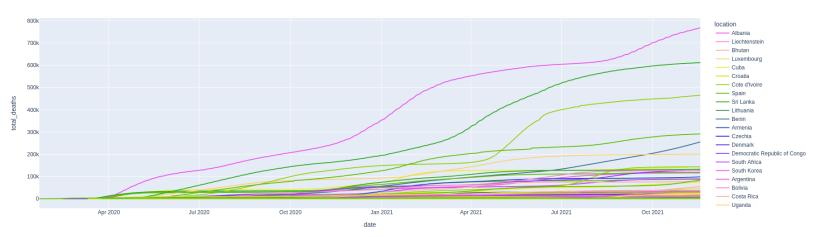
مباني علوم داده

على على آبادي - ٩٧٢٢٢٠۶١

تاریخ: ۲۷/۰۸/۲۷

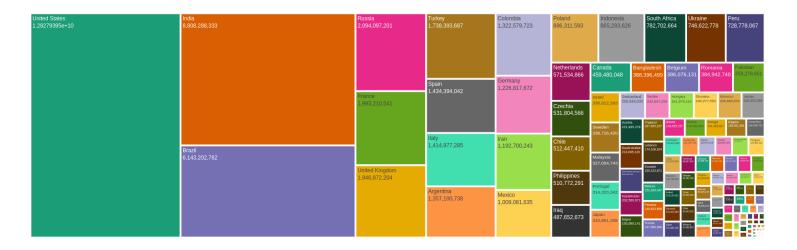
گزارش تمرین سری ۲

total\_deaths



در ادامه به نمودارهای درختی زیر توجه کنید.

total\_cases





مبانی علوم داده

على على آبادي - ٩٧٢٢٢٠۶١

گزارش تمرین سری ۲ تاریخ: ۲۰/۰۸/۲۷

total\_cases\_per\_million

Montenegro 56,247,363,68	Slovenia 40,613,763.84	Maldives 35,543,223.2	Lithuania 34,004,998.78	Qatar 31,613,293.74	France 29,500,964.08	Switzerland 28,746,968.67	North Macedonia 25,897,217.06		Moldova 22,787,339.29	Malta 20,494,382.87	Cape Verde Oman 17,829,550.04 17,736		9.34 Denn 9.34 17,44	nark Uki 13,351.11 17,		Albania 16,048,888.69
					Creatia	Description of the second	Colombia			Turkey 20.441.412.98	Paraguay 15,926,153.67	Belize 13,603,865.0	South Africa 13,035,920.	Suriname 13 13,011,214.9	Tunisia 99 12,396,942.	Azerbaijan 12,190,080.66
	Luxembourg	Georgia	Netherlands	Portugal 30,913,229.97	Croatia 29,114,235.73	Brazil 28,707,434.92	25,798,459.51	Uruguay 23,634,506.91	Jordan 22,234,872.89							
Czechia 49,587,574.04	39,801,277.21	35,533,638.98	33,280,832.55							United Arab Emirates 20,341,299.62	Belarus 15,915,129.05	Botswana 12,174,249.96	Bolivia 12,129,341.9	Mongolia 12,112,213.07	Canada 7 12,070,008.89	Iraq 19 11,842,165.08
				Spain 30,685,368.87	Armenia 29,020,815.82	United Kingdom 28,543,535.85	Lebanon 25,729,803.33	Poland 23,449,257.69	Peru 21,846,248.42		Germany 14.622.297.79	Libya 11,238,779.18	celand 10,043,132.57		valini Mexico 23,912.254 7,745,54	Kyrgyzstan 41.051 7,587,410.26
Bahrain 46,116,492,93 Siovakia	United States 38,532,167.8	Serbia 35,340,987.1	Belgium 33,189,911.07							Bulgaria 20,221,594.97		Dominican Republic	Malaysia 9,978,423.14	Singapore 7,038,872,305	Sid and Tiblings Guaterns 5,394.73 5,0034,040	Nepal 45.487 6,506,509.20
				Estonia 30,100,578.95	Cyprus 28,987,510.12	Kuwait 27,924,716.65	Hungary 25,116,259.61	Italy 23,439,399.76	Bosnia and Herzegovina 21,334,769.34	Romania 20,124,808.06	Russia 14,351,779.75		Cuba 9,746,379.903	Fij Finia 6,405,135.071 6,223	and Saudi Ar. 3,421.107 6,046,43	zabla Jamaica 35.055 5,852,781.69
	Israel	Panama	Sweden							20,124,808.06	Iran 14,027,021.47	Kazakhstan 10,664,233,68	Honduras	Barbados 4,40,4 6,401,347,859	A SEC NO. SEC. 235.	Depth Darks 2,000,000,000
44,001,034.47	36,251,468.42	34,194,011.39	33,140,861.27	Argentina 29,757,181.18	Liechtenstein 28,787,760.78	Chile 26,672,795.88	Austria 24,531,406.81	Latvia 22,859,388.72	Ireland 20.697.549.26	Palestine			9,600,293.012	India Prison 6,321,394.595	2.05.00 Sc 12015.50	Signal Si
										20,065,710.34	Greece 13,821,809.95		Ecuador 3,906,509.895	Morocco 5/146 6,310,564,277	2000 100 100 100 100 100 100 100 100 100	

people\_fully\_vaccinated\_per\_hundred

Israel 15,776.91	Bahrain 12,539.04	Denmark 11,065.27	Canada 9,851.22	Switzerland 9,512.93	Norway 9,012.97	Maldives 8,258.84	Hungary 7,609.72	Malaysia 6,867.62	Qatar 5,994.8	Romania 5,252.95	Australia 5,182.87	Panama 5,143.63	Argentina 5,115.15	Japan 4,993.57	Cyprus 4,929.47
								Poland 6.812.26	South Korea 5,922.36	New Zealand 4,863.5	Peru 4,399.47	Croatia 4,391.07	Kazakhstan 4,358.1	Azerbaijan 4,265.7	Mexico 4,174.8
	Uruguay 11,977.19	Ireland 10,802.06	Germany 9,771.69	France 9,505.24	Czechia 8,541.37		7,411.11	0,012.20	5,922.50						
Chile 13,744.16						Slovenia		Dominican Republic	Saudi Arabia	Sri Lanka 4,832.65	Trinidad and Tobago 4,068.23	Russia 3,662.21	Suriname Bulgar 2,884.66 2,844.	ia Albania 08 2,784.45	Thailand 2,769.62
						0.400.40		6,744.7	5,620.79				Tunisia North Macedon	IIIula	Zimbabwe Sweden 1,028,25 1,599,23
	United States 11,458.18	Belgium 10,774.2	Lithuania 9,666.38	Austria 9,407.73	Estonia 8,511.13					Montenegro 4,791.91	Morocco 3,923.52	Colombia 3,561.79	2,719.24 Netherlands	2,140.55	1,055.20
						Cambodia Slova 8,037.5 6,936	Slovakia Finland 6,936.78 6,678.43		Cuba 5,561.52				Brunei 1,514.57 2,658.62	1,349.94 1,340.8	1,303.42
Malta 13,440.79										Ecuador 4,678.2		Jordan 3,530.92		Mauritius Vietnam 1,107.29 1,103.24	Paraguay Georgia 1,126.39 1,054.96
	United Kingdom 11,427.33	Italy 10,117.95	Liechtenstein 9,518.67	Greece 9,273.91	Mongolia 8,383.98	Spain Portugal 6,896.37	Portugal	Singapore 6,382.81	Barbados 5,491.7		2	Lebanon Philippines 2,470.1 Philippines	Souti Atica BIS4 Belize 905.47	China dansica (32) 542.25 692.57	
							6,896.37			Serbia 4,495.71			Taiwan 1,414.83 2,430.5 Bhutan 1,349.99	Energlation   Fakistan   745.25	



مبانی علوم داده

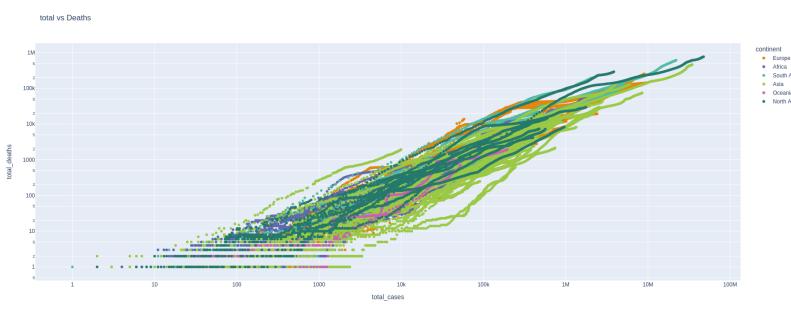
على على آبادي - ٩٧٢٢٢٠۶١

تاریخ: ۲۷/۰۸/۲۷

گزارش تمرین سری ۲

کشور های پیشتاز در زمینه واکسیناسیون در نمودار های بالا قابل دیدن هستند.

در انتها نیز نمودار total cases بر اساس total cases را مشاهده می کنیم.



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 <u>notebook</u> نیز در تحلیل ها و انتخاب نمودار ها استفاده شد.