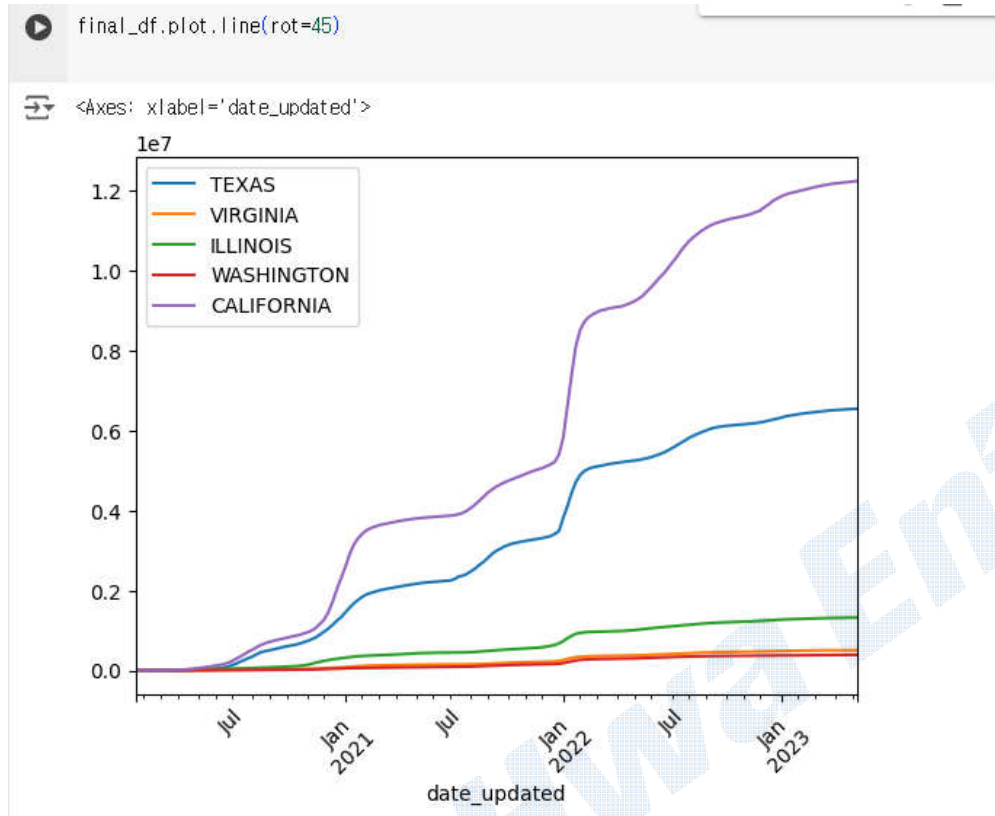


12주차 과제 : 미국 5개주 인구비례 확진자 수 비율 변동

학번 : 20238247

이름 : 전경용



미국의 5개주의 각 인구수는 약 아래와 같은 인구수와 비율을 나타냅니다.

캘리포니아 39,000,000명 / 1

텍사스 : 30,000,000명 / 0.77

일리노이 : 12,600,000명 / 0.32

버지니아 : 8,650,000명 / 0.22

워싱턴 : 7,730,000명 / 0.2

위 그래프에서와 같이 인구수가 매우 높은 캘리포니아와 텍사스의 경우

계속해서 확진자 수 비율이 상승하고 높은 반면

일리노이, 버지니아, 워싱턴 주의 경우 캘리포니아, 텍사스와 같이 급속한 증가 추세를 보이지 않고

안정적으로 완만한 곡선을 그리며 증가 추세가 잘 유지 되고 있습니다.

그중 워싱턴의 경우 인구수 대비 확진자 수는 가장 낮은 비율로 지속적으로 잘 유지 하고 있습니다.

코드 진행과정 캡처 :

✓ 0초 [4] `import pandas as pd`  
`raw_ca_df = pd.read_csv('/content/drive/MyDrive/12week/ca-coviddata.csv')`  
`raw_ca_df.head()`

⇄ 숨겨진 출력 표시

다음 단계: `raw_ca_df` 변수로 코드 생성

☐ 추천 차트 보기

✓ 0초 ▶ #ca 캘리포니아 확진자 분석  
`ca_total_cases = raw_ca_df[['tot_cases', 'date_updated']]`  
`ca_dateindex_df = ca_total_cases.set_index('date_updated')`  
`ca_dateindex_df.head()`

⇄

	<code>tot_cases</code>
<code>date_updated</code>	
01/23/2020	0
01/30/2020	2
02/06/2020	8
02/13/2020	16
02/20/2020	30

다음 단계: `ca_dateindex_df` 변수로 코드 생성

☐ 추천 차트 보기

[10] #tx 텍사스 확진자 분석  
`raw_tx_df = pd.read_csv('/content/drive/MyDrive/12week/tx-coviddata.csv')`  
`tx_total_cases = raw_tx_df[['tot_cases', 'date_updated']]`  
`tx_dateindex_df = tx_total_cases.set_index('date_updated')`  
`tx_dateindex_df.head()`

⇄

	<code>tot_cases</code>
<code>date_updated</code>	
01/23/2020	0
01/30/2020	0
02/06/2020	0
02/13/2020	0
02/20/2020	0

다음 단계: `tx_dateindex_df` 변수로 코드 생성

☐ 추천 차트 보기



### #va 버지니아 확진자 분석

```
raw_va_df = pd.read_csv('/content/drive/MyDrive/12week/va-coviddata.csv')
va_total_cases = raw_va_df[['tot_cases', 'date_updated']]
va_dateindex_df = va_total_cases.set_index('date_updated')
va_dateindex_df.head()
```



tot_cases	
date_updated	
01/23/2020	0
01/30/2020	0
02/06/2020	0
02/13/2020	0
02/20/2020	0



다음 단계: [va\\_dateindex\\_df 변수로 코드 생성](#)

[추천 차트 보기](#)

### [18] #il 일리노이 확진자 분석

```
raw_il_df = pd.read_csv('/content/drive/MyDrive/12week/il-coviddata.csv')
il_total_cases = raw_il_df[['tot_cases', 'date_updated']]
il_dateindex_df = il_total_cases.set_index('date_updated')
il_dateindex_df.head()
```



tot_cases	
date_updated	
01/23/2020	0
01/30/2020	1
02/06/2020	2
02/13/2020	2
02/20/2020	2



다음 단계: [il\\_dateindex\\_df 변수로 코드 생성](#)

[추천 차트 보기](#)

[13] #wa 워싱턴 확진자 분석

```
raw_wa_df = pd.read_csv('/content/drive/MyDrive/12week/wa-coviddata.csv')
wa_total_cases = raw_wa_df[['tot_cases', 'date_updated']]
wa_dateindex_df = wa_total_cases.set_index('date_updated')
wa_dateindex_df.head()
```



tot_cases	
date_updated	
01/23/2020	1
01/30/2020	1
02/06/2020	1
02/13/2020	1
02/20/2020	1



다음 단계:

[wa\\_dateindex\\_df 변수로 코드 생성](#)

[추천 차트 보기](#)

[19] #5개도시의 데이터프레임

```
ca_population = 38_965_193
tx_population = 30_000_000
va_population = 8_654_542
il_population = 12_600_000
wa_population = 7_730_000
```

```
tx_rate = round((tx_population / ca_population), 2)
va_rate = round((va_population / ca_population), 2)
il_rate = round((il_population / ca_population), 2)
wa_rate = round((wa_population / ca_population), 2)
print(tx_rate)
print(va_rate)
print(il_rate)
print(wa_rate)
```



```
0.77
0.22
0.32
0.2
```

```
[20] #index 를 datetime으로 변환
ca_dateindex_df.index = pd.to_datetime(ca_dateindex_df.index)
#ny_dateindex_df.index = ny_dateindex_df.index.astype('datetime64[ns]')
tx_dateindex_df.index = pd.to_datetime(ca_dateindex_df.index)
va_dateindex_df.index = pd.to_datetime(ca_dateindex_df.index)
il_dateindex_df.index = pd.to_datetime(ca_dateindex_df.index)
wa_dateindex_df.index = pd.to_datetime(ca_dateindex_df.index)
wa_dateindex_df.index.dtype
```

dtype('<M8[ns]')

```
[21] #최종 dataframe
final_df = pd.concat([tx_dateindex_df * tx_rate,
                      va_dateindex_df * va_rate,
                      il_dateindex_df * il_rate,
                      wa_dateindex_df * wa_rate, ca_dateindex_df], axis=1)
final_df.columns = ['TEXAS', 'VIRGINIA', 'ILLINOIS', 'WASHINGTON', 'CALIFORNIA']
final_df.head()
```

	TEXAS	VIRGINIA	ILLINOIS	WASHINGTON	CALIFORNIA
date_updated					
2020-01-23	0.0	0.0	0.00	0.2	0
2020-01-30	0.0	0.0	0.32	0.2	2
2020-02-06	0.0	0.0	0.64	0.2	8
2020-02-13	0.0	0.0	0.64	0.2	16
2020-02-20	0.0	0.0	0.64	0.2	30

다음 단계:  ☒ 추천 차트 보기

