

```
In [46]: import matplotlib.pyplot as plt
import matplotlib.font_manager as fm # 한글 폰트
import matplotlib
import os, warnings
import pandas as pd
import seaborn as sns

warnings.filterwarnings('ignore')

pd.set_option('display.max_rows', None)
pd.set_option('display.max_columns', None)

df1 = pd.read_csv('C:/Users/user/LikeLion_AI/data/CSV/today_corona_edit.csv') # 90
df2 = pd.read_csv('C:/Users/user/LikeLion_AI/data/CSV/Natl_GDP.csv') # 90
```

## 데이터 처리 과정

```
In [47]: df1
```

```
Out[47]:
```

	국가	확진자	치료중	사망자	완치	치명 (%)	완치 (%)	발생률	인구수	국토면적
0	중국	91629	512.0	4636	86481.0	5	94.0	64	1439323776	9640821
1	인도	29977861	662521.0	389302	28926038.0	1	97.0	21723	1380004385	3287263
2	미국	34419838	5034868.0	617463	28767507.0	2	84.0	103987	331002651	9833517
3	인도 네시 아	2004445	147728.0	54956	1801761.0	3	90.0	7328	273523615	1904569
4	파키 스탄	949838	33452.0	22034	894352.0	2	94.0	4300	220892340	796095
5	브라 질	17969806	1178597.0	502817	16288392.0	3	91.0	84540	212559417	8515770
6	나이 지리 아	167292	1377.0	2118	163797.0	1	98.0	812	206139589	923768
7	방글 라데 시	856304	57196.0	13626	785482.0	2	92.0	5200	164689383	143998
8	러시 아	5334204	326070.0	129801	4878333.0	2	92.0	36552	145934462	17125407
9	멕시 코	2478551	273052.0	231244	1974255.0	9	80.0	19224	128932753	1964375
10	일본	785969	22248.0	14439	749282.0	2	95.0	6214	126476461	377975
11	에티 오피 아	275318	15530.0	4286	255502.0	2	93.0	2395	114963588	1104300
12	필리 핀	1364239	55847.0	23749	1284643.0	2	94.0	12450	109581078	300000
13	이집 트	277797	55846.0	15898	206053.0	6	74.0	2715	102334404	1010408

	국가	확진자	치료중	사망자	완치	치명 (%)	완치 (%)	발생률	인구수	국토면적
14	베트남	13530	8008.0	69	5453.0	1	40.0	139	97338579	331210
15	콩고민주공화국	37926	9109.0	879	27938.0	2	74.0	423	89561403	2344858
16	터키	5375593	88626.0	49236	5237731.0	1	97.0	63738	84339067	783356
17	이란	3105620	262290.0	83101	2760229.0	3	89.0	36975	83992949	1648195
18	독일	3730599	31492.0	91007	3608100.0	2	97.0	44526	83783942	357022
19	태국	225365	35836.0	1693	187836.0	1	83.0	3229	69799978	513120
20	영국	4640507	208530.0	127981	4303996.0	3	93.0	68357	67886011	243610
21	프랑스	5757798	80626.0	110778	5566394.0	2	97.0	88210	65273511	640679
22	이탈리아	4253460	76853.0	127291	4049316.0	3	95.0	70350	60461826	301340
23	탄자니아	509	305.0	21	183.0	4	36.0	9	59734218	947303
24	남아프리카공화국	1832479	117004.0	58795	1656680.0	3	90.0	30897	59308690	1219090
25	미얀마	148617	11599.0	3265	133753.0	2	90.0	2731	54409800	676578
26	케냐	179293	52867.0	3461	122965.0	2	69.0	3334	53771296	580367
27	대한민국	151901	6078.0	2006	143817.0	1	95.0	2963	51269185	100412
28	콜롬비아	3968405	181876.0	100582	3685947.0	3	93.0	77991	50882891	1138914
29	스페인	3764651	123928.0	80689	3560034.0	2	95.0	80519	46754778	505990
30	우간다	72679	22205.0	680	49794.0	1	69.0	1589	45741007	241038
31	아르헨티나	4277395	277070.0	89490	3910835.0	2	91.0	94641	45195774	2780400
32	알제리	136294	37831.0	3641	94822.0	3	70.0	3108	43851044	2381741
33	수단	36347	3548.0	2737	30062.0	8	83.0	829	43849260	1861484
34	우크라이나	2229846	25674.0	52032	2152140.0	2	97.0	50987	43733762	603628
35	이라크	1292700	70911.0	16910	1204879.0	1	93.0	32139	40222493	438317

	국가	확진자	치료중	사망자	완치	치명 (%)	완치 (%)	발생률	인구수	국토면적
36	아프가니스탄	105755	36391.0	4293	65071.0	4	62.0	2717	38928346	652230
37	폴란드	2878840	153363.0	74829	2650648.0	3	92.0	76066	37846611	312685
38	캐나다	1409607	11032.0	26084	1372491.0	2	97.0	37348	37742154	9984670
39	모로코	526737	3595.0	9244	513898.0	2	98.0	14271	36910560	446550
40	사우디아라비아	475403	10584.0	7691	457128.0	2	96.0	13656	34813871	2149690
41	우즈베키스탄	106452	3571.0	718	102163.0	1	96.0	3181	33469203	447400
42	페루	2030611	NaN	190645	NaN	9	NaN	61586	32971854	1285216
43	앙골라	37748	5098.0	868	31782.0	2	84.0	1149	32866272	1246700
44	말레이시아	701019	62918.0	4477	633624.0	1	90.0	21659	32365999	329847
45	모잠비크	72577	1488.0	848	70241.0	1	97.0	2322	31255435	801590
46	가나	95059	1260.0	794	93005.0	1	98.0	3059	31072940	238533
47	예멘	6889	1576.0	1355	3958.0	20	58.0	231	29825964	527968
48	네팔	622640	53940.0	8772	559928.0	1	90.0	21370	29136808	147516
49	베네수엘라	262038	16708.0	2973	242357.0	1	93.0	9215	28435940	912050
50	마다가스카르	42137	765.0	903	40469.0	2	96.0	1522	27691018	587041
51	카메룬	80328	853.0	1313	78162.0	2	97.0	3026	26545863	475440
52	코트디부아르	48047	252.0	308	47487.0	1	99.0	1821	26378274	322463
53	호주	30366	184.0	910	29272.0	3	96.0	1191	25499884	7741220
54	니제르	5469	98.0	193	5178.0	4	95.0	226	24206644	1267000
55	대만	14080	5977.0	569	7534.0	4	54.0	591	23816775	36197
56	스리랑카	241820	37798.0	2633	201389.0	1	83.0	11293	21413249	65610

	국가	확진자	치료중	사망자	완치	치명 (%)	완치 (%)	발생률	인구수	국토면적
57	부르키나파소	13469	9.0	167	13293.0	1	99.0	644	20903273	274200
58	말리	14385	3825.0	524	10036.0	4	70.0	710	20250833	1240192
59	루마니아	1080282	2684.0	32391	1045207.0	3	97.0	56154	19237691	238391
60	말라위	34914	852.0	1171	32891.0	3	94.0	1825	19129952	118484
61	칠레	1522223	38470.0	31645	1452108.0	2	95.0	79630	19116201	756096
62	카자흐스탄	410523	20147.0	4249	386127.0	1	94.0	21863	18776707	2724900
63	잠비아	130631	18490.0	1691	110450.0	1	85.0	7106	18383955	752618
64	과테말라	280854	17419.0	8735	254700.0	3	91.0	15677	17915568	108889
65	에콰도르	446633	9821.0	21304	415508.0	5	93.0	25315	17643054	283561
66	시리아	25158	1563.0	1848	21747.0	7	86.0	1438	17500658	185180
67	네덜란드	1679542	NaN	17727	NaN	1	NaN	98019	17134872	41543
68	세네갈	42437	365.0	1158	40914.0	3	96.0	2534	16743927	196722
69	캄보디아	43446	5002.0	441	38003.0	1	88.0	2599	16718965	181035
70	차드	4947	5.0	174	4768.0	4	96.0	301	16425864	1284000
71	소말리아	14867	6888.0	775	7204.0	5	49.0	935	15893222	637657
72	짐바브웨	42195	3310.0	1685	37200.0	4	88.0	2839	14862924	390757
73	기니	23535	1085.0	168	22282.0	1	95.0	1792	13132795	245857
74	르완다	31435	4343.0	388	26704.0	1	85.0	2427	12952218	26338
75	베냉	8140	58.0	103	7979.0	1	98.0	671	12123200	112622
76	부룬디	5242	4461.0	8	773.0	0	15.0	441	11890784	27834
77	튀니지	385428	34658.0	14118	336652.0	4	87.0	32612	11818619	163610
78	볼리비아	422811	65678.0	16174	340959.0	4	81.0	36221	11673021	1098581
79	벨기에	1079640	37925.0	25141	1016574.0	2	94.0	93156	11589623	30528

	국가	확진자	치료중	사망자	완치	치명 (%)	완치 (%)	발생률	인구수	국토면적
80	아이티	17371	4319.0	385	12667.0	2	73.0	1523	11402528	27750
81	남수단	10786	157.0	115	10514.0	1	98.0	964	11193725	644329
82	도미니카공화국	317645	54238.0	3758	259649.0	1	82.0	29282	10847910	48670
83	체코	1666082	2640.0	30280	1633162.0	2	98.0	155578	10708981	78867
84	그리스	418548	6111.0	12559	399878.0	3	96.0	40156	10423054	131957
85	요르단	747000	6498.0	9671	730831.0	1	98.0	73213	10203134	89342
86	포르투갈	865806	28657.0	17068	820081.0	2	95.0	84910	10196709	92090
87	아제르바이잔	335521	959.0	4963	329599.0	2	98.0	33092	10139177	86600
88	스웨덴	1084636	NaN	14537	NaN	1	NaN	107398	10099265	450295
89	온두라스	254194	155560.0	6772	91862.0	3	36.0	25664	9904607	112492

```
In [49]: ## 위의 두 자료를 합치는 과정
target = list(df1['국가'])
country = list(df2['국가'])

gdp = list(df2['2021'])
idx_list = []
new_cty = []
gdp_list = []

for i in target:
    for idx, one in enumerate(country):
        if i == one:
            idx_list.append(idx)

for i in idx_list:
    new_cty.append(country[i])
    if gdp[i] == 'no data':
        gdp_list.append(0.0)
    else:
        gdp_list.append(float(gdp[i]))

for i in range(90):
    print(target[i], gdp_list[i])
```

```
중국 16642.318
인도 3049.704
미국 22675.271
인도네시아 1158.783
파키스탄 0.0
브라질 1491.772
나이지리아 514.049
```

방글라데시 352.908  
러시아 1710.734  
멕시코 1192.48  
일본 5378.136  
에티오피아 93.966  
필리핀 402.638  
이집트 394.284  
베트남 354.868  
콩고민주공화국 12.022  
터키 794.53  
이란 682.859  
독일 4319.286  
태국 538.735  
영국 3124.65  
프랑스 2938.271  
이탈리아 2106.287  
탄자니아 65.919  
남아프리카 공화국 329.529  
미얀마 76.195  
케냐 106.041  
대한민국 1806.707  
콜롬비아 295.61  
스페인 1461.552  
우간다 41.271  
아르헨티나 418.15  
알제리 151.459  
수단 35.827  
우크라이나 164.593  
이라크 190.733  
아프가니스탄 19.938  
폴란드 642.121  
캐나다 1883.487  
모로코 124.003  
사우디아라비아 804.921  
우즈베키스탄 61.203  
페루 225.918  
앙골라 66.493  
말레이시아 387.093  
모잠비크 13.957  
가나 74.26  
예멘 25.095  
네팔 36.084  
베네수엘라 42.53  
마다가스카르 14.746  
카메룬 44.893  
코트디부아르 70.991  
호주 1617.543  
니제르 15.899  
대만 759.104  
스리랑카 84.532  
부르키나파소 18.853  
말리 19.912  
루마니아 289.13  
말라위 9.268  
칠레 307.938  
카자흐스탄 187.836  
잠비아 18.955  
과테말라 81.402  
에콰도르 100.595  
시리아 0.0  
네덜란드 1012.598  
세네갈 27.927  
캄보디아 27.239  
차드 12.531  
소말리아 5.367  
짐바브웨 26.085  
기니 16.339  
르완다 10.633  
베냉 17.327

부룬디 3.244  
 튀니지 44.265  
 볼리비아 43.11  
 벨기에 578.996  
 아이티 22.431  
 남수단 4.461  
 도미니카 공화국 83.917  
 체코 276.109  
 그리스 209.857  
 요르단 44.979  
 포르투갈 257.391  
 아제르바이잔 49.914  
 스웨덴 625.948  
 온두라스 26.161

## 정확한 통계를 위해 중국은 삭제한다 ...!!

```

In [51]: gdp_list = pd.DataFrame(gdp_list)
df1['GDP'] = gdp_list
# df1 = df1.drop([0])
df3 = df1.sort_values(by='인구수', ascending=False).head(30)
# df3 = df3.drop([1])
df1
  
```

Out[51]:

	국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
1	인 도	29977861	662521.0	389302	28926038.0	1	97.0	21723	1380004385	3287263	3049.7
2	미 국	34419838	5034868.0	617463	28767507.0	2	84.0	103987	331002651	9833517	22675.2
3	인 도	2004445	147728.0	54956	1801761.0	3	90.0	7328	273523615	1904569	1158.7
4	네 시 아	949838	33452.0	22034	894352.0	2	94.0	4300	220892340	796095	0.0
5	파 키 스 탄	17969806	1178597.0	502817	16288392.0	3	91.0	84540	212559417	8515770	1491.7
6	브 라 질	167292	1377.0	2118	163797.0	1	98.0	812	206139589	923768	514.0
7	나 이 지 리 아	856304	57196.0	13626	785482.0	2	92.0	5200	164689383	143998	352.9
8	방 글 라 데 시	5334204	326070.0	129801	4878333.0	2	92.0	36552	145934462	17125407	1710.7

국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
9 멕시코	2478551	273052.0	231244	1974255.0	9	80.0	19224	128932753	1964375	1192.4
10 일본	785969	22248.0	14439	749282.0	2	95.0	6214	126476461	377975	5378.1
11 에티오피아	275318	15530.0	4286	255502.0	2	93.0	2395	114963588	1104300	93.9
12 필리핀	1364239	55847.0	23749	1284643.0	2	94.0	12450	109581078	300000	402.6
13 이집트	277797	55846.0	15898	206053.0	6	74.0	2715	102334404	1010408	394.2
14 베트남	13530	8008.0	69	5453.0	1	40.0	139	97338579	331210	354.8
15 콩고민주공화국	37926	9109.0	879	27938.0	2	74.0	423	89561403	2344858	12.0
16 터키	5375593	88626.0	49236	5237731.0	1	97.0	63738	84339067	783356	794.5
17 이란	3105620	262290.0	83101	2760229.0	3	89.0	36975	83992949	1648195	682.8
18 독일	3730599	31492.0	91007	3608100.0	2	97.0	44526	83783942	357022	4319.2
19 태국	225365	35836.0	1693	187836.0	1	83.0	3229	69799978	513120	538.7
20 영국	4640507	208530.0	127981	4303996.0	3	93.0	68357	67886011	243610	3124.6
21 프랑스	5757798	80626.0	110778	5566394.0	2	97.0	88210	65273511	640679	2938.2
22 이탈리아	4253460	76853.0	127291	4049316.0	3	95.0	70350	60461826	301340	2106.2
23 탄자니아	509	305.0	21	183.0	4	36.0	9	59734218	947303	65.9



	국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
24	남 아 프 리 카 공 화 국	1832479	117004.0	58795	1656680.0	3	90.0	30897	59308690	1219090	329.5
25	미 얀 마	148617	11599.0	3265	133753.0	2	90.0	2731	54409800	676578	76.1
26	케 냐	179293	52867.0	3461	122965.0	2	69.0	3334	53771296	580367	106.0
27	대 한 민 국	151901	6078.0	2006	143817.0	1	95.0	2963	51269185	100412	1806.7
28	콜 롬 비 아	3968405	181876.0	100582	3685947.0	3	93.0	77991	50882891	1138914	295.6
29	스 페 인	3764651	123928.0	80689	3560034.0	2	95.0	80519	46754778	505990	1461.5
30	우 간 다	72679	22205.0	680	49794.0	1	69.0	1589	45741007	241038	41.2
31	아 르 헨 티 나	4277395	277070.0	89490	3910835.0	2	91.0	94641	45195774	2780400	418.1
32	알 제 리	136294	37831.0	3641	94822.0	3	70.0	3108	43851044	2381741	151.4
33	수 단	36347	3548.0	2737	30062.0	8	83.0	829	43849260	1861484	35.8
34	우 크 라 이 나	2229846	25674.0	52032	2152140.0	2	97.0	50987	43733762	603628	164.5
35	이 라 크	1292700	70911.0	16910	1204879.0	1	93.0	32139	40222493	438317	190.7
36	아 프 가 니 스 탄	105755	36391.0	4293	65071.0	4	62.0	2717	38928346	652230	19.9

	국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
37	폴 란 드	2878840	153363.0	74829	2650648.0	3	92.0	76066	37846611	312685	642.1
38	캐 나 다	1409607	11032.0	26084	1372491.0	2	97.0	37348	37742154	9984670	1883.4
39	모 로 코	526737	3595.0	9244	513898.0	2	98.0	14271	36910560	446550	124.0
40	사 우 디 아 라 비 아	475403	10584.0	7691	457128.0	2	96.0	13656	34813871	2149690	804.9
41	우 즈 베 키 스 탄	106452	3571.0	718	102163.0	1	96.0	3181	33469203	447400	61.2
42	페 루	2030611	NaN	190645	NaN	9	NaN	61586	32971854	1285216	225.9
43	앙 골 라	37748	5098.0	868	31782.0	2	84.0	1149	32866272	1246700	66.4
44	말 레 이 시 아	701019	62918.0	4477	633624.0	1	90.0	21659	32365999	329847	387.0
45	모 잠 비 크	72577	1488.0	848	70241.0	1	97.0	2322	31255435	801590	13.9
46	가 나	95059	1260.0	794	93005.0	1	98.0	3059	31072940	238533	74.2
47	예 멘	6889	1576.0	1355	3958.0	20	58.0	231	29825964	527968	25.0
48	네 팔	622640	53940.0	8772	559928.0	1	90.0	21370	29136808	147516	36.0
49	베 네 수 엘 라	262038	16708.0	2973	242357.0	1	93.0	9215	28435940	912050	42.5

	국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
50	마다 가 스 카 르	42137	765.0	903	40469.0	2	96.0	1522	27691018	587041	14.7
51	카 메 룬	80328	853.0	1313	78162.0	2	97.0	3026	26545863	475440	44.8
52	코 트 디 부 아 르	48047	252.0	308	47487.0	1	99.0	1821	26378274	322463	70.9
53	호 주	30366	184.0	910	29272.0	3	96.0	1191	25499884	7741220	1617.5
54	니 제 르	5469	98.0	193	5178.0	4	95.0	226	24206644	1267000	15.8
55	대 만	14080	5977.0	569	7534.0	4	54.0	591	23816775	36197	759.1
56	스 리 랑 카	241820	37798.0	2633	201389.0	1	83.0	11293	21413249	65610	84.5
57	부 르 키 나 파 소	13469	9.0	167	13293.0	1	99.0	644	20903273	274200	18.8
58	말 리	14385	3825.0	524	10036.0	4	70.0	710	20250833	1240192	19.9
59	루 마 니 아	1080282	2684.0	32391	1045207.0	3	97.0	56154	19237691	238391	289.1
60	말 라 위	34914	852.0	1171	32891.0	3	94.0	1825	19129952	118484	9.2
61	칠 레	1522223	38470.0	31645	1452108.0	2	95.0	79630	19116201	756096	307.9
62	카 자 흐 스 탄	410523	20147.0	4249	386127.0	1	94.0	21863	18776707	2724900	187.8
63	잠 비 아	130631	18490.0	1691	110450.0	1	85.0	7106	18383955	752618	18.9

	국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
64	과테말라	280854	17419.0	8735	254700.0	3	91.0	15677	17915568	108889	81.4
65	에콰도르	446633	9821.0	21304	415508.0	5	93.0	25315	17643054	283561	100.5
66	시리아	25158	1563.0	1848	21747.0	7	86.0	1438	17500658	185180	0.0
67	네덜란드	1679542	NaN	17727	NaN	1	NaN	98019	17134872	41543	1012.5
68	세네갈	42437	365.0	1158	40914.0	3	96.0	2534	16743927	196722	27.5
69	캄보디아	43446	5002.0	441	38003.0	1	88.0	2599	16718965	181035	27.2
70	차드	4947	5.0	174	4768.0	4	96.0	301	16425864	1284000	12.5
71	소말리아	14867	6888.0	775	7204.0	5	49.0	935	15893222	637657	5.3
72	짐바브웨	42195	3310.0	1685	37200.0	4	88.0	2839	14862924	390757	26.0
73	기니	23535	1085.0	168	22282.0	1	95.0	1792	13132795	245857	16.3
74	르완다	31435	4343.0	388	26704.0	1	85.0	2427	12952218	26338	10.6
75	베냉	8140	58.0	103	7979.0	1	98.0	671	12123200	112622	17.3
76	부룬디	5242	4461.0	8	773.0	0	15.0	441	11890784	27834	3.2
77	튀니지	385428	34658.0	14118	336652.0	4	87.0	32612	11818619	163610	44.2
78	볼리비아	422811	65678.0	16174	340959.0	4	81.0	36221	11673021	1098581	43.1

	국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
79	벨기에	1079640	37925.0	25141	1016574.0	2	94.0	93156	11589623	30528	578.9
80	아티티	17371	4319.0	385	12667.0	2	73.0	1523	11402528	27750	22.4
81	남수단	10786	157.0	115	10514.0	1	98.0	964	11193725	644329	4.4
82	도미니카공화국	317645	54238.0	3758	259649.0	1	82.0	29282	10847910	48670	83.9
83	체코	1666082	2640.0	30280	1633162.0	2	98.0	155578	10708981	78867	276.1
84	그리스	418548	6111.0	12559	399878.0	3	96.0	40156	10423054	131957	209.8
85	요르단	747000	6498.0	9671	730831.0	1	98.0	73213	10203134	89342	44.9
86	포르투갈	865806	28657.0	17068	820081.0	2	95.0	84910	10196709	92090	257.3
87	아제르바이잔	335521	959.0	4963	329599.0	2	98.0	33092	10139177	86600	49.9
88	스웨덴	1084636	NaN	14537	NaN	1	NaN	107398	10099265	450295	625.9
89	온두라스	254194	155560.0	6772	91862.0	3	36.0	25664	9904607	112492	26.1



```
In [7]: %matplotlib inline
```

```
In [8]: # 그래프에서 마이너스 폰트 깨지는 문제에 대한 대처
matplotlib.rcParams['axes.unicode_minus'] = False
```

```
In [117]: ## 한글 폰트 적용 ( 안하면 한글 깨짐 )
f_name = fm.FontProperties(fname="C:/Windows/Fonts/malgunbd.ttf").get_name()
plt.rc('font', family=f_name)
```

## 인구수 하위 30개국 확진자 비율

인도의 인구수가 다른 나라들에 비해 매우 많으므로 일시적으로 제외시켰습니다.

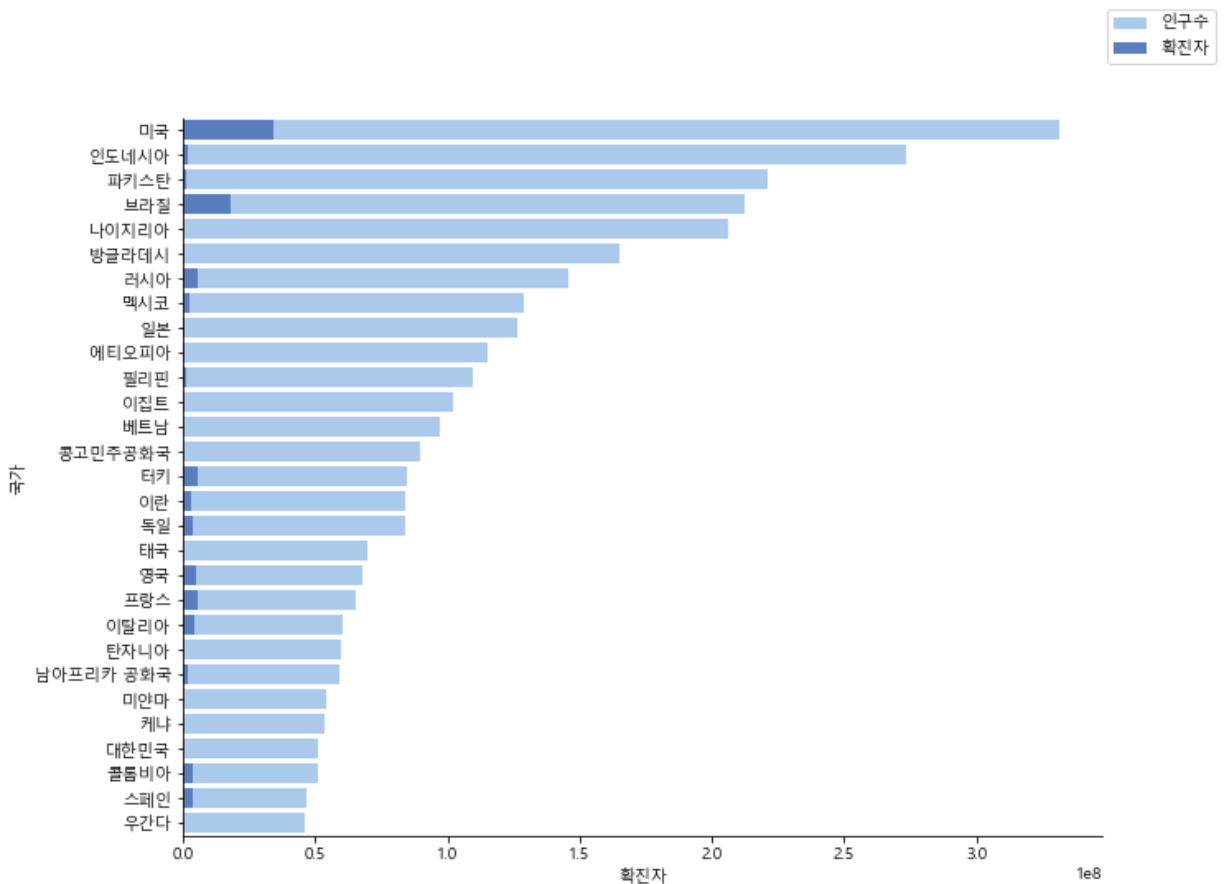
```
In [52]: fig = plt.figure(figsize=(10, 8))
fig.patch.set_facecolor('xkcd:white')

sns.set_color_codes('pastel')
sns.barplot(x='인구수', y='국가', data=df3, label='인구수', color='b')

sns.set_color_codes('muted')
sns.barplot(x='확진자', y='국가', data=df3, label='확진자', color='b')

fig.legend(ncol=1, loc="upper right")

sns.despine()
```



### 데이터 해석 :

이 데이터에서는 확진자가 눈에 띄게 보이는 몇몇 국가들이 있습니다.

- 미국의 경우에는 전 세계적으로 GDP가 가장 높으면서 확진자가 가장 많습니다. 도심지를 주변으로 사람들이 매우 밀집해있으며, 몇몇 보수적인 시민들의 단체행동으로 인해 확진자가 크게 늘어나게 되었습니다.
- 브라질의 경우에는 대통령이 적극적으로 방역활동을 지시하지 않은 것이 큰 문제점이었습니다. 심지어 대면활동과 경제활동을 장려하기까지 했고, 이로 인해 수많은 사람들이 죽어나갔습니다.
- 그리고 여러 유럽국가에서, 우선 처음으로 변이바이러스가 발생한 곳이 영국이었고, 그를 중심으로 코로나 바이러스가 크게 확산하기도 하였습니다.

# 인구수 하위 30개국 확진자 비율

```
In [53]: df4 = df1.sort_values(by='인구수', ascending=False).tail(30)

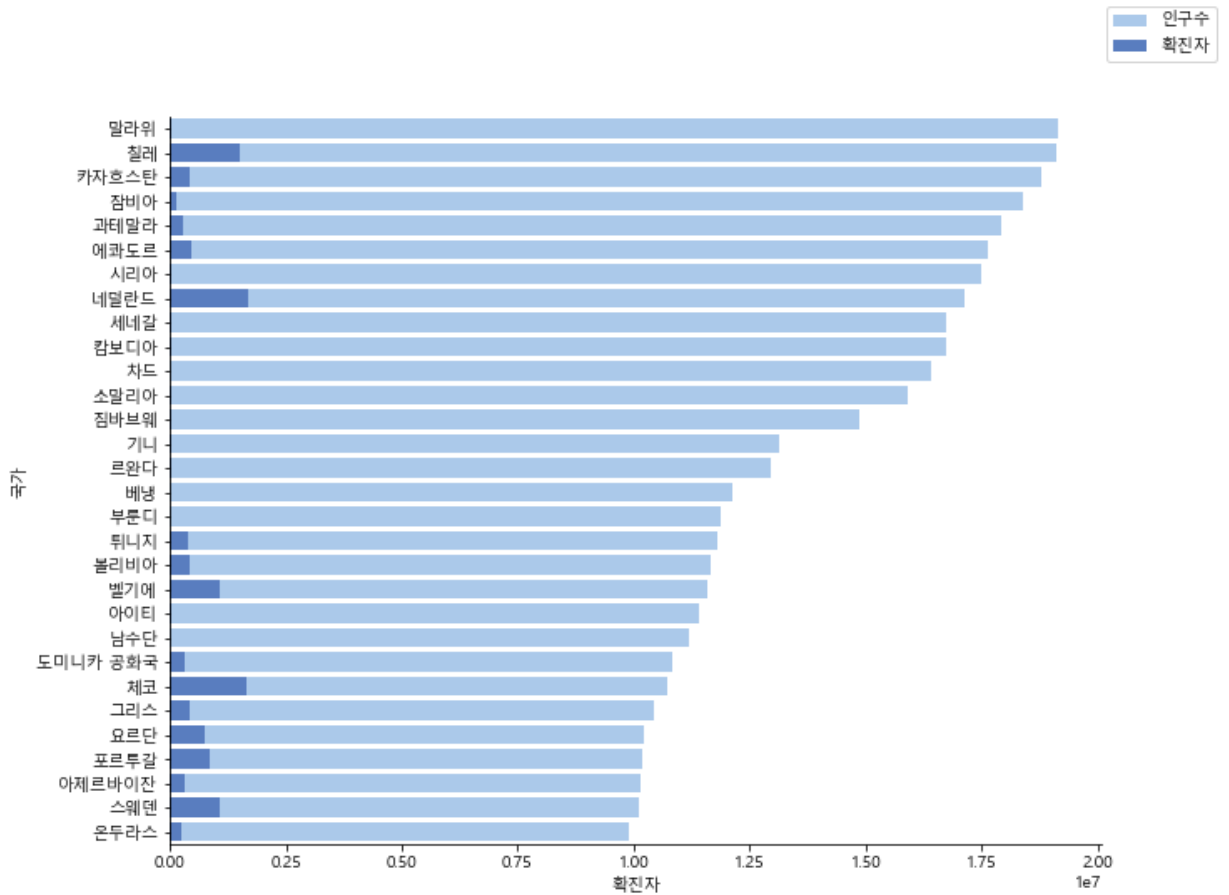
fig = plt.figure(figsize=(10, 8))
fig.patch.set_facecolor('xkcd:white')

sns.set_color_codes('pastel')
sns.barplot(x='인구수', y='국가', data=df4, label='인구수', color='b')

sns.set_color_codes('muted')
sns.barplot(x='확진자', y='국가', data=df4, label='확진자', color='b')

fig.legend(ncol=1, loc="upper right")

sns.despine()
```



```
In [54]: df1['인구밀도'] = df1['인구수']/df1['국토면적'] # 인구밀도 Column 추가
df1
```

Out[54]:

	국가	확진자	치료중	사망자	완치	치명률 (%)	완치률 (%)	발생률	인구수	국토면적	G
1	인도	29977861	662521.0	389302	28926038.0	1	97.0	21723	1380004385	3287263	3049.7
2	미국	34419838	5034868.0	617463	28767507.0	2	84.0	103987	331002651	9833517	22675.2
3	인도네시아	2004445	147728.0	54956	1801761.0	3	90.0	7328	273523615	1904569	1158.7

	국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
4	파 키 스 탄	949838	33452.0	22034	894352.0	2	94.0	4300	220892340	796095	0.0
5	브 라 질	17969806	1178597.0	502817	16288392.0	3	91.0	84540	212559417	8515770	1491.7
6	나 이 지 리 아	167292	1377.0	2118	163797.0	1	98.0	812	206139589	923768	514.0
7	방 글 라 데 시	856304	57196.0	13626	785482.0	2	92.0	5200	164689383	143998	352.9
8	러 시 아	5334204	326070.0	129801	4878333.0	2	92.0	36552	145934462	17125407	1710.7
9	멕 시 코	2478551	273052.0	231244	1974255.0	9	80.0	19224	128932753	1964375	1192.4
10	일 본	785969	22248.0	14439	749282.0	2	95.0	6214	126476461	377975	5378.1
11	에 티 오피 아	275318	15530.0	4286	255502.0	2	93.0	2395	114963588	1104300	93.9
12	필 리 핀	1364239	55847.0	23749	1284643.0	2	94.0	12450	109581078	300000	402.6
13	이 집 트	277797	55846.0	15898	206053.0	6	74.0	2715	102334404	1010408	394.2
14	베 트 남	13530	8008.0	69	5453.0	1	40.0	139	97338579	331210	354.8
15	콩 고 민 주 공 화 국	37926	9109.0	879	27938.0	2	74.0	423	89561403	2344858	12.0
16	터 키	5375593	88626.0	49236	5237731.0	1	97.0	63738	84339067	783356	794.5
17	이 란	3105620	262290.0	83101	2760229.0	3	89.0	36975	83992949	1648195	682.8



	국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
18	독일	3730599	31492.0	91007	3608100.0	2	97.0	44526	83783942	357022	4319.2
19	태국	225365	35836.0	1693	187836.0	1	83.0	3229	69799978	513120	538.7
20	영국	4640507	208530.0	127981	4303996.0	3	93.0	68357	67886011	243610	3124.6
21	프랑스	5757798	80626.0	110778	5566394.0	2	97.0	88210	65273511	640679	2938.2
22	이탈리아	4253460	76853.0	127291	4049316.0	3	95.0	70350	60461826	301340	2106.2
23	탄자니아	509	305.0	21	183.0	4	36.0	9	59734218	947303	65.9
24	남아프리카공화국	1832479	117004.0	58795	1656680.0	3	90.0	30897	59308690	1219090	329.5
25	미얀마	148617	11599.0	3265	133753.0	2	90.0	2731	54409800	676578	76.1
26	케냐	179293	52867.0	3461	122965.0	2	69.0	3334	53771296	580367	106.0
27	대한민국	151901	6078.0	2006	143817.0	1	95.0	2963	51269185	100412	1806.7
28	콜롬비아	3968405	181876.0	100582	3685947.0	3	93.0	77991	50882891	1138914	295.6
29	스페인	3764651	123928.0	80689	3560034.0	2	95.0	80519	46754778	505990	1461.5
30	우간다	72679	22205.0	680	49794.0	1	69.0	1589	45741007	241038	41.2
31	아르헨티나	4277395	277070.0	89490	3910835.0	2	91.0	94641	45195774	2780400	418.1

	국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
32	알제리	136294	37831.0	3641	94822.0	3	70.0	3108	43851044	2381741	151.4
33	수단	36347	3548.0	2737	30062.0	8	83.0	829	43849260	1861484	35.8
34	우크라이나	2229846	25674.0	52032	2152140.0	2	97.0	50987	43733762	603628	164.5
35	이라크	1292700	70911.0	16910	1204879.0	1	93.0	32139	40222493	438317	190.7
36	아프가니스탄	105755	36391.0	4293	65071.0	4	62.0	2717	38928346	652230	19.9
37	폴란드	2878840	153363.0	74829	2650648.0	3	92.0	76066	37846611	312685	642.1
38	캐나다	1409607	11032.0	26084	1372491.0	2	97.0	37348	37742154	9984670	1883.4
39	모로코	526737	3595.0	9244	513898.0	2	98.0	14271	36910560	446550	124.0
40	사우디아라비아	475403	10584.0	7691	457128.0	2	96.0	13656	34813871	2149690	804.9
41	우즈베키스탄	106452	3571.0	718	102163.0	1	96.0	3181	33469203	447400	61.2
42	페루	2030611	NaN	190645	NaN	9	NaN	61586	32971854	1285216	225.9
43	앙골라	37748	5098.0	868	31782.0	2	84.0	1149	32866272	1246700	66.4
44	말레이시아	701019	62918.0	4477	633624.0	1	90.0	21659	32365999	329847	387.0

	국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
45	모잠비크	72577	1488.0	848	70241.0	1	97.0	2322	31255435	801590	13.9
46	가나	95059	1260.0	794	93005.0	1	98.0	3059	31072940	238533	74.2
47	예멘	6889	1576.0	1355	3958.0	20	58.0	231	29825964	527968	25.0
48	네팔	622640	53940.0	8772	559928.0	1	90.0	21370	29136808	147516	36.0
49	베네수엘라	262038	16708.0	2973	242357.0	1	93.0	9215	28435940	912050	42.5
50	마다가스카르	42137	765.0	903	40469.0	2	96.0	1522	27691018	587041	14.7
51	카메룬	80328	853.0	1313	78162.0	2	97.0	3026	26545863	475440	44.8
52	코트디부아르	48047	252.0	308	47487.0	1	99.0	1821	26378274	322463	70.9
53	호주	30366	184.0	910	29272.0	3	96.0	1191	25499884	7741220	1617.5
54	니제르	5469	98.0	193	5178.0	4	95.0	226	24206644	1267000	15.8
55	대만	14080	5977.0	569	7534.0	4	54.0	591	23816775	36197	759.1
56	스리랑카	241820	37798.0	2633	201389.0	1	83.0	11293	21413249	65610	84.5
57	부르키나파소	13469	9.0	167	13293.0	1	99.0	644	20903273	274200	18.8
58	말리	14385	3825.0	524	10036.0	4	70.0	710	20250833	1240192	19.9

	국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
59	루 마 니 아	1080282	2684.0	32391	1045207.0	3	97.0	56154	19237691	238391	289.1
60	말 라 위	34914	852.0	1171	32891.0	3	94.0	1825	19129952	118484	9.2
61	칠 레	1522223	38470.0	31645	1452108.0	2	95.0	79630	19116201	756096	307.9
62	카 자 흐 스 탄	410523	20147.0	4249	386127.0	1	94.0	21863	18776707	2724900	187.8
63	잠 비 아	130631	18490.0	1691	110450.0	1	85.0	7106	18383955	752618	18.9
64	과 테 말 라	280854	17419.0	8735	254700.0	3	91.0	15677	17915568	108889	81.4
65	에 과 도 르	446633	9821.0	21304	415508.0	5	93.0	25315	17643054	283561	100.9
66	시 리 아	25158	1563.0	1848	21747.0	7	86.0	1438	17500658	185180	0.0
67	네 덜 란 드	1679542	NaN	17727	NaN	1	NaN	98019	17134872	41543	1012.9
68	세 네 갈	42437	365.0	1158	40914.0	3	96.0	2534	16743927	196722	27.9
69	캠 보 디 아	43446	5002.0	441	38003.0	1	88.0	2599	16718965	181035	27.2
70	차 드	4947	5.0	174	4768.0	4	96.0	301	16425864	1284000	12.9
71	소 말 리 아	14867	6888.0	775	7204.0	5	49.0	935	15893222	637657	5.9
72	짐 바 브 웨	42195	3310.0	1685	37200.0	4	88.0	2839	14862924	390757	26.0

	국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
73	기 니	23535	1085.0	168	22282.0	1	95.0	1792	13132795	245857	16.3
74	르 완 다	31435	4343.0	388	26704.0	1	85.0	2427	12952218	26338	10.6
75	베 냉	8140	58.0	103	7979.0	1	98.0	671	12123200	112622	17.3
76	부 룬 디	5242	4461.0	8	773.0	0	15.0	441	11890784	27834	3.2
77	튀 니 지	385428	34658.0	14118	336652.0	4	87.0	32612	11818619	163610	44.2
78	볼 리 비 아	422811	65678.0	16174	340959.0	4	81.0	36221	11673021	1098581	43.1
79	벨 기 에	1079640	37925.0	25141	1016574.0	2	94.0	93156	11589623	30528	578.9
80	아 이 티	17371	4319.0	385	12667.0	2	73.0	1523	11402528	27750	22.4
81	남 수 단	10786	157.0	115	10514.0	1	98.0	964	11193725	644329	4.4
82	도 미 니 카 공 화 국	317645	54238.0	3758	259649.0	1	82.0	29282	10847910	48670	83.9
83	체 코	1666082	2640.0	30280	1633162.0	2	98.0	155578	10708981	78867	276.1
84	그 리 스	418548	6111.0	12559	399878.0	3	96.0	40156	10423054	131957	209.8
85	요 르 단	747000	6498.0	9671	730831.0	1	98.0	73213	10203134	89342	44.9
86	포 르 투 갈	865806	28657.0	17068	820081.0	2	95.0	84910	10196709	92090	257.3

	국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
87	아 제 르 바 이 잔	335521	959.0	4963	329599.0	2	98.0	33092	10139177	86600	49.9
88	스 웨 덴	1084636	NaN	14537	NaN	1	NaN	107398	10099265	450295	625.9
89	온 두 라스	254194	155560.0	6772	91862.0	3	36.0	25664	9904607	112492	26.1

## GDP 정보가 없는 하위 2개국 삭제하기 !

```
In [144]: df5 = df1.sort_values(by='GDP', ascending=False)
# df5 = df5.drop([4])
# df5 = df5.drop([66])

df5
```

Out[144]:

	국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
2	미 국	34419838	5034868.0	617463	28767507.0	2	84.0	103987	331002651	9833517	22675.2
10	일 본	785969	22248.0	14439	749282.0	2	95.0	6214	126476461	377975	5378.1
18	독 일	3730599	31492.0	91007	3608100.0	2	97.0	44526	83783942	357022	4319.2
20	영 국	4640507	208530.0	127981	4303996.0	3	93.0	68357	67886011	243610	3124.6
1	인 도	29977861	662521.0	389302	28926038.0	1	97.0	21723	1380004385	3287263	3049.7
21	프 랑스	5757798	80626.0	110778	5566394.0	2	97.0	88210	65273511	640679	2938.2
22	이 탈 리 아	4253460	76853.0	127291	4049316.0	3	95.0	70350	60461826	301340	2106.2
38	캐 나 다	1409607	11032.0	26084	1372491.0	2	97.0	37348	37742154	9984670	1883.4
27	대 한 민 국	151901	6078.0	2006	143817.0	1	95.0	2963	51269185	100412	1806.7

	국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
8	러 시 아	5334204	326070.0	129801	4878333.0	2	92.0	36552	145934462	17125407	1710.7
53	호 주	30366	184.0	910	29272.0	3	96.0	1191	25499884	7741220	1617.5
5	브 라 질	17969806	1178597.0	502817	16288392.0	3	91.0	84540	212559417	8515770	1491.7
29	스 페 인	3764651	123928.0	80689	3560034.0	2	95.0	80519	46754778	505990	1461.5
9	멕 시 코	2478551	273052.0	231244	1974255.0	9	80.0	19224	128932753	1964375	1192.4
3	인 도	2004445	147728.0	54956	1801761.0	3	90.0	7328	273523615	1904569	1158.7
67	네 덜 란 드	1679542	NaN	17727	NaN	1	NaN	98019	17134872	41543	1012.5
40	사 우 디 아 라 비 아	475403	10584.0	7691	457128.0	2	96.0	13656	34813871	2149690	804.5
16	터 키	5375593	88626.0	49236	5237731.0	1	97.0	63738	84339067	783356	794.5
55	대 만	14080	5977.0	569	7534.0	4	54.0	591	23816775	36197	759.1
17	이 란	3105620	262290.0	83101	2760229.0	3	89.0	36975	83992949	1648195	682.8
37	폴 란 드	2878840	153363.0	74829	2650648.0	3	92.0	76066	37846611	312685	642.1
88	스 웨 덴	1084636	NaN	14537	NaN	1	NaN	107398	10099265	450295	625.5
79	벨 기 에	1079640	37925.0	25141	1016574.0	2	94.0	93156	11589623	30528	578.5
19	태 국	225365	35836.0	1693	187836.0	1	83.0	3229	69799978	513120	538.7

국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
6 나 이 지 리 아	167292	1377.0	2118	163797.0	1	98.0	812	206139589	923768	514.0
31 아 르 헨 티 나	4277395	277070.0	89490	3910835.0	2	91.0	94641	45195774	2780400	418.1
12 필 리 핀	1364239	55847.0	23749	1284643.0	2	94.0	12450	109581078	300000	402.6
13 이 집 트	277797	55846.0	15898	206053.0	6	74.0	2715	102334404	1010408	394.2
44 말 레 이 시 아	701019	62918.0	4477	633624.0	1	90.0	21659	32365999	329847	387.0
14 베 트 남	13530	8008.0	69	5453.0	1	40.0	139	97338579	331210	354.8
7 방 글 라 데 시	856304	57196.0	13626	785482.0	2	92.0	5200	164689383	143998	352.9
24 남 아 프 리 카 공 화 국	1832479	117004.0	58795	1656680.0	3	90.0	30897	59308690	1219090	329.5
61 칠 레	1522223	38470.0	31645	1452108.0	2	95.0	79630	19116201	756096	307.9
28 콜 롬 비 아	3968405	181876.0	100582	3685947.0	3	93.0	77991	50882891	1138914	295.6
59 루 마 니 아	1080282	2684.0	32391	1045207.0	3	97.0	56154	19237691	238391	289.1
83 체 코	1666082	2640.0	30280	1633162.0	2	98.0	155578	10708981	78867	276.1



	국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
86	포르투갈	865806	28657.0	17068	820081.0	2	95.0	84910	10196709	92090	257.3
42	페루	2030611	NaN	190645	NaN	9	NaN	61586	32971854	1285216	225.9
84	그리스	418548	6111.0	12559	399878.0	3	96.0	40156	10423054	131957	209.8
35	이라크	1292700	70911.0	16910	1204879.0	1	93.0	32139	40222493	438317	190.7
62	카자흐스탄	410523	20147.0	4249	386127.0	1	94.0	21863	18776707	2724900	187.8
34	우크라이나	2229846	25674.0	52032	2152140.0	2	97.0	50987	43733762	603628	164.5
32	알제리	136294	37831.0	3641	94822.0	3	70.0	3108	43851044	2381741	151.4
39	모로코	526737	3595.0	9244	513898.0	2	98.0	14271	36910560	446550	124.0
26	케냐	179293	52867.0	3461	122965.0	2	69.0	3334	53771296	580367	106.0
65	에콰도르	446633	9821.0	21304	415508.0	5	93.0	25315	17643054	283561	100.5
11	에티오피아	275318	15530.0	4286	255502.0	2	93.0	2395	114963588	1104300	93.9
56	스리랑카	241820	37798.0	2633	201389.0	1	83.0	11293	21413249	65610	84.5
82	도미니카공화국	317645	54238.0	3758	259649.0	1	82.0	29282	10847910	48670	83.9

	국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
64	과테말라	280854	17419.0	8735	254700.0	3	91.0	15677	17915568	108889	81.4
25	미얀마	148617	11599.0	3265	133753.0	2	90.0	2731	54409800	676578	76.1
46	가나	95059	1260.0	794	93005.0	1	98.0	3059	31072940	238533	74.2
52	코트디부아르	48047	252.0	308	47487.0	1	99.0	1821	26378274	322463	70.9
43	앙골라	37748	5098.0	868	31782.0	2	84.0	1149	32866272	1246700	66.4
23	탄자니아	509	305.0	21	183.0	4	36.0	9	59734218	947303	65.9
41	우즈베키스탄	106452	3571.0	718	102163.0	1	96.0	3181	33469203	447400	61.2
87	아제르바이잔	335521	959.0	4963	329599.0	2	98.0	33092	10139177	86600	49.9
85	요르단	747000	6498.0	9671	730831.0	1	98.0	73213	10203134	89342	44.9
51	카메룬	80328	853.0	1313	78162.0	2	97.0	3026	26545863	475440	44.8
77	튀니지	385428	34658.0	14118	336652.0	4	87.0	32612	11818619	163610	44.2
78	볼리비아	422811	65678.0	16174	340959.0	4	81.0	36221	11673021	1098581	43.1

	국 가	확진자	치료중	사망자	완치	치 명 (%)	완치 (%)	발생률	인구수	국토면적	G
49	베 네 수 엘 라	262038	16708.0	2973	242357.0	1	93.0	9215	28435940	912050	42.5
30	우 간 다	72679	22205.0	680	49794.0	1	69.0	1589	45741007	241038	41.2
48	네 팔	622640	53940.0	8772	559928.0	1	90.0	21370	29136808	147516	36.0
33	수 단	36347	3548.0	2737	30062.0	8	83.0	829	43849260	1861484	35.8
68	세 네 갈	42437	365.0	1158	40914.0	3	96.0	2534	16743927	196722	27.9
69	캠 보 디 아	43446	5002.0	441	38003.0	1	88.0	2599	16718965	181035	27.2
89	온 두 라스	254194	155560.0	6772	91862.0	3	36.0	25664	9904607	112492	26.1
72	짐 바 브 웨	42195	3310.0	1685	37200.0	4	88.0	2839	14862924	390757	26.0
47	예 멘	6889	1576.0	1355	3958.0	20	58.0	231	29825964	527968	25.0
80	아 이 티	17371	4319.0	385	12667.0	2	73.0	1523	11402528	27750	22.4
36	아 프 가 니 스 탄	105755	36391.0	4293	65071.0	4	62.0	2717	38928346	652230	19.9
58	말 리	14385	3825.0	524	10036.0	4	70.0	710	20250833	1240192	19.9
63	잠 비 아	130631	18490.0	1691	110450.0	1	85.0	7106	18383955	752618	18.9
57	부 르 키 나 파 소	13469	9.0	167	13293.0	1	99.0	644	20903273	274200	18.8

	국가	확진자	치료중	사망자	완치	치명 (%)	완치 (%)	발생률	인구수	국토면적	G
75	베냉	8140	58.0	103	7979.0	1	98.0	671	12123200	112622	17.3
73	기니	23535	1085.0	168	22282.0	1	95.0	1792	13132795	245857	16.3
54	니제르	5469	98.0	193	5178.0	4	95.0	226	24206644	1267000	15.8
50	마다가스카르	42137	765.0	903	40469.0	2	96.0	1522	27691018	587041	14.7
45	모잠비크	72577	1488.0	848	70241.0	1	97.0	2322	31255435	801590	13.9
70	차드	4947	5.0	174	4768.0	4	96.0	301	16425864	1284000	12.5
15	콩고민주공화국	37926	9109.0	879	27938.0	2	74.0	423	89561403	2344858	12.0
74	르완다	31435	4343.0	388	26704.0	1	85.0	2427	12952218	26338	10.6
60	말라위	34914	852.0	1171	32891.0	3	94.0	1825	19129952	118484	9.2
71	소말리아	14867	6888.0	775	7204.0	5	49.0	935	15893222	637657	5.3
81	남수단	10786	157.0	115	10514.0	1	98.0	964	11193725	644329	4.4
76	부룬디	5242	4461.0	8	773.0	0	15.0	441	11890784	27834	3.2

## GDP 상위 10개국, 하위 10개국 parameter 별 상관계수 비교 ( Heatmap )

< 확진자, 인구수, 국토면적, 사망자, GDP, 완치, 인구밀도 >

```

In [91]: gdp_top10 = df5.sort_values(by='GDP', ascending=False).head(10)
          gdp_low10 = df5.sort_values(by='GDP', ascending=False).tail(10)

          corr_1 = gdp_top10.loc[:, ['확진자', '인구수', '국토면적', '사망자', 'GDP', '완치', '인구밀도']]
          corr_2 = gdp_low10.loc[:, ['확진자', '인구수', '국토면적', '사망자', 'GDP', '완치', '인구밀도']]

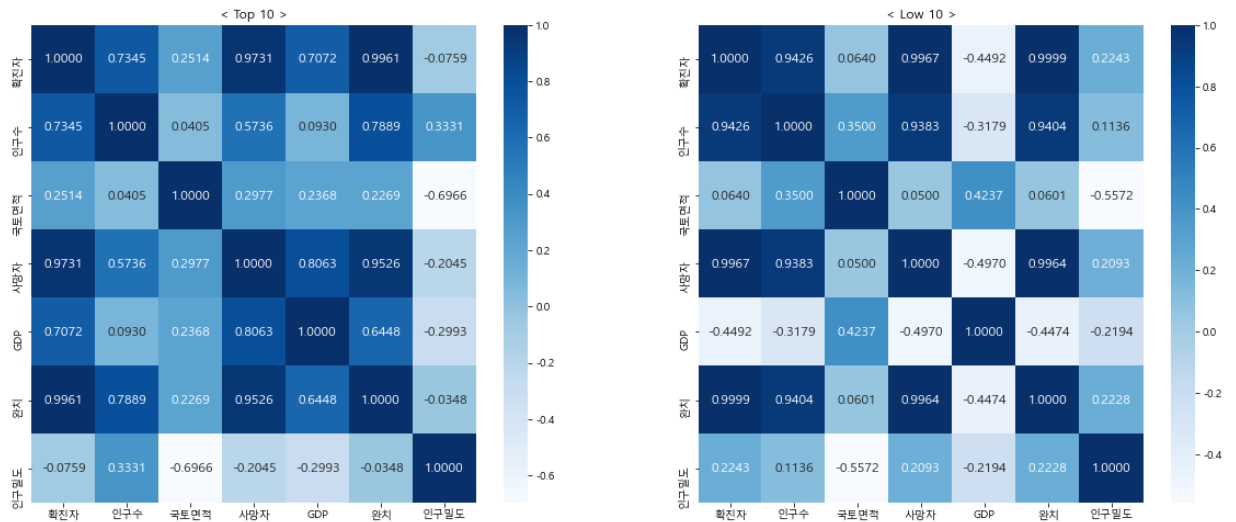
In [92]: fig = plt.figure(figsize=(20, 8))
          fig.patch.set_facecolor('xkcd:white')

          axis1 = fig.add_subplot(1, 2, 1)
          axis1 = sns.heatmap(corr_1, annot=True, annot_kws={'size':12}, fmt='.4f', cmap='Blue')
          plt.title('< Top 10 >')
          axis1.plot()

          axis2 = fig.add_subplot(1, 2, 2)
          axis2 = sns.heatmap(corr_2, annot=True, annot_kws={'size':12}, fmt='.4f', cmap='Blue')
          plt.title('< Low 10 >')
          axis2.plot()

          plt.show()

```



## 데이터 해석 :

우선, 기본적으로 7개의 parameter 간의 상관관계를 분석하는 것이기 때문에 Heatmap으로 표현 하긴 했지만 조금 복잡하게 보이긴합니다. 그래서 이 Heatmap은 훑어보듯이 넘어가고 대충 어떤 느낌인지만 설명하려고 합니다.

### Top 10 나라의 경우

인구밀도를 제외한 나머지의 Parameter에서 매우 높은 상관관계를 보이거나 낮은 정도의 상관관계를 보입니다. 인구밀도 기준으로 보았을 때, 인구수를 제외하고는 다 상관관계가 음수(-)의 값을 가집니다. 왜냐하면  $\text{인구밀도} = \text{인구수} / \text{국토면적}$  이기 때문에 당연한 결과라고 볼 수 있습니다. 같은 원리로, 인구밀도와 국토면적은 반비례관계이기 때문에 상관관계수가 -0.7로 상당히 낮은 것을 볼 수 있습니다. 나머지 parameter 간의 관계는 다음 축소된 Heatmap을 보면서 추가적으로 설명해보겠습니다.

### GDP가 낮은 하위 10개 나라의 경우

Low 10 국가의 경우 Top 10 국가와 비교해서 parameter 간의 관계가 극명하게 나뉜 것을 볼 수 있습니다. 아무래도 경제상황이 좋지 못하다 보니 이런 상관관계를 보이는 것이 아닐까 예상해볼 수

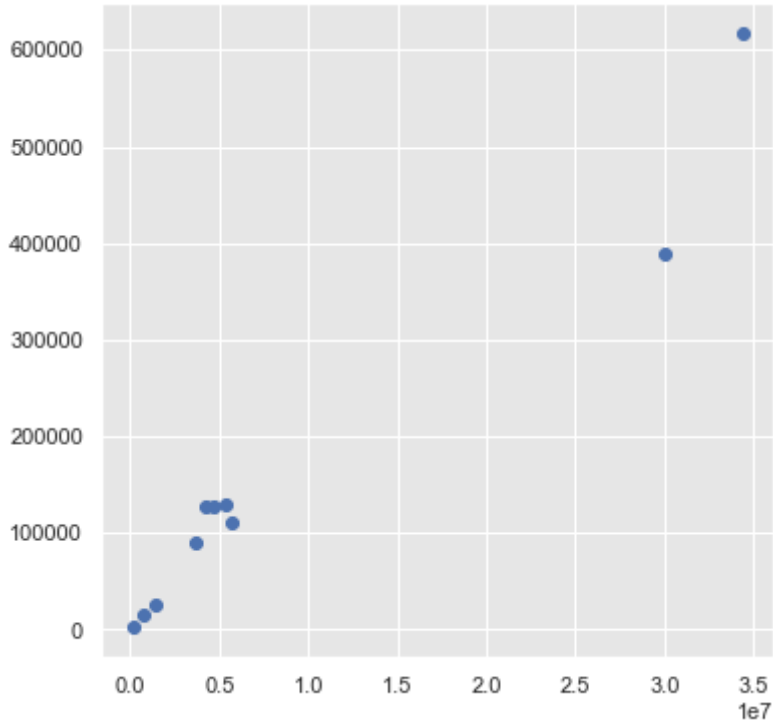
있습니다. 다음 Heatmap을 보면서 Top 10 국가와 어떻게 차이가 나는지 살펴보도록 하겠습니다.

## 상관관계를 시각화 하기 위한 **ScatterPlot**, 회귀직선

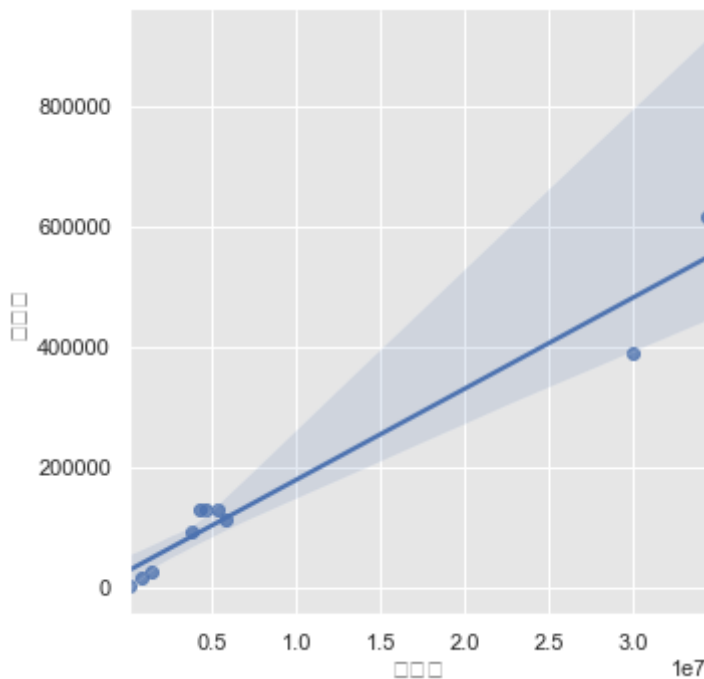
```
In [114]: fig = plt.figure(figsize=(6,6), facecolor='white')

plt.scatter(x='확진자',y='사망자',data=gdp_top10)
plt.show()

sns.set_style("darkgrid", {"axes.facecolor": ".9"})
sns.lmplot(x='확진자',y='사망자',data=gdp_top10)
```



Out[114]: <seaborn.axisgrid.FacetGrid at 0x188c1f7a880>



< 확진자, 인구수, 국토면적, 사망자, **GDP**, 완치 >

```
In [93]: gdp_top10 = df5.sort_values(by='GDP', ascending=False).head(10)
```

```
gdp_low10 = df5.sort_values(by='GDP', ascending=False).tail(10)
```

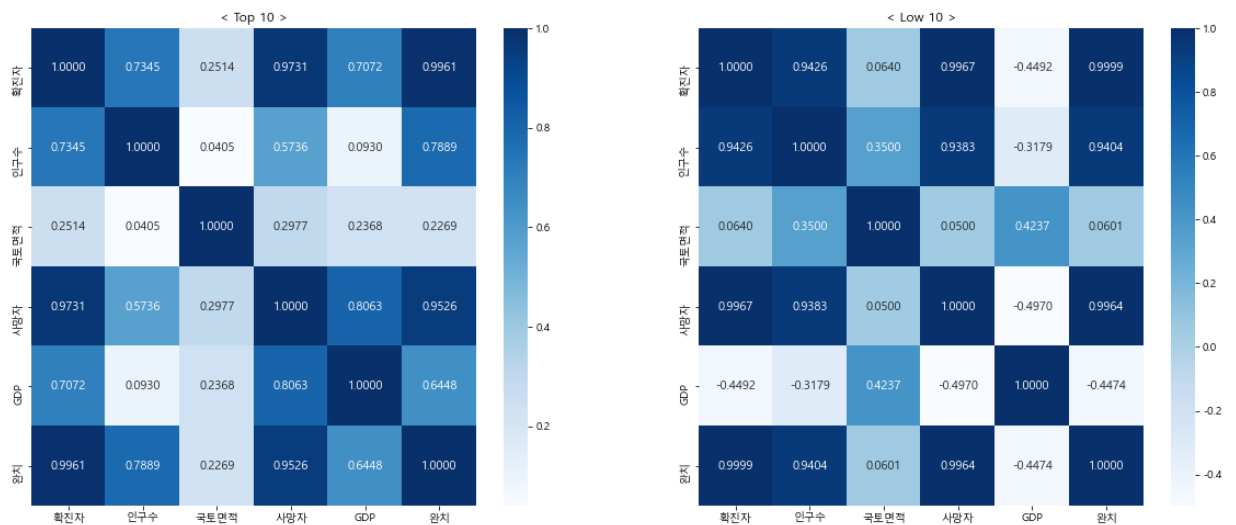
```
corr_1 = gdp_top10.loc[:, ['확진자', '인구수', '국토면적', '사망자', 'GDP', '완치']]
corr_2 = gdp_low10.loc[:, ['확진자', '인구수', '국토면적', '사망자', 'GDP', '완치']]
```

```
In [94]: fig = plt.figure(figsize=(20, 8))
fig.patch.set_facecolor('xkcd:white')

axis1 = fig.add_subplot(1, 2, 1)
axis1 = sns.heatmap(corr_1, annot=True, annot_kws={'size':11}, fmt='.4f', cmap='Blue')
plt.title('< Top 10 >')
axis1.plot()

axis2 = fig.add_subplot(1, 2, 2)
axis2 = sns.heatmap(corr_2, annot=True, annot_kws={'size':11}, fmt='.4f', cmap='Blue')
plt.title('< Low 10 >')
axis2.plot()

plt.show()
```



## 데이터 해석 :

인구밀도 parameter를 제거하고 나니 Heatmap이 더 명확해지면서 parameter 간의 상관관계를 더욱 파악하기 쉬워졌습니다.

### Top 10 나라의 경우

국토면적이 다른 parameter들과 비교해서 약한 상관관계를 가지는 것을 알 수 있습니다. 즉, 국토면적의 넓고 좁음이 확진자에 끼치는 영향이 크지 않다고 해석할 수 있을 것 같습니다.

그리고 국토면적 - 인구수, 인구수 - GDP의 관계가 약한 것을 볼 수 있는데 그렇다고 국토면적 - GDP의 관계는 그렇게 약하지는 않구나라는 것을 알 수 있습니다.

추가적으로, 국토면적이 넓다고 꼭 인구수가 많은 것은 아니며(상관계수 = 0.0405), GDP가 높다고 꼭 인구수가 많은 것은 아니라고(상관계수 = 0.0930) 판단할 수 있습니다.

### Low 10 나라의 경우

Low 10 국가와 Top 10 국가와의 차이가 극명하게 보이는 것 같습니다. 대표적으로, GDP와 다른 parameter와의 상관관계가 크지 않다는 점을 뽑을 수 있습니다. 세부적인 내용들을 설명해보겠습니다.

1. GDP와 Covid-19 환자(확진자, 사망자, 완치자)와의 관계 : Top 10 나라에서는 비교적 높은 상관관계를 가지는 것에 비해 Low 10 나라에서는 오히려 약한 음의 상관관계를 가지는 것을 알 수 있습니다. 조금 과장해서 말하면 반비례 관계를 가지므로, 이 나라들 사이에서는 GDP가 높을 수록 그렇지 않은 나라들에 비해 확진자가 더 적다고 판단할 수 있습니다. 이러한 결과가 나오는 이유는, GDP가 높은 선진국의 경우에 크게 발달한 도심지와 사람들의 높은 소비수준으로 인해 사람들 간의 전염이 빠르고 치명적이었습니다. 반면에 GDP가 낮은 나라의 경우에는 사람들이 한 지역에 많이 밀집해 있을 가능성이 낮고, 이러한 나라들에 있어서 GDP가 높은 것이 사람들 간의 적극적인 소비활동과 직접적인 접촉이 많이 생긴다는 것을 의미하지 않기 때문으로 해석할 수 있습니다. \\
2. GDP와 국토면적의 관계 : Top 10 나라의 상관계수(0.2368)에 비해 Low 10 나라의 상관계수(0.4237)가 조금 더 의미있는 수치를 가지는 것을 알 수 있습니다. 즉, Top 10 나라에 비해 국토면적이 넓을 수록 GDP가 더 높을 수 있다는 것을 판단할 수 있습니다. \\
3. 국토면적과 인구수의 관계 : Top 10 나라의 상관계수(0.0405)에 비해 Low 10 나라의 상관계수(0.3500)가 비교적 더 높습니다. 즉, 이 카테고리에서는 국토면적이 넓을수록 인구 수가 더 많을 수 있다는 것을 알 수 있습니다. \\
4. 인구수와 Covid-19 환자(확진자, 사망자, 완치자)와의 관계 : Top 10 나라에 비해 인구수 대비 Covid-19 환자의 상관계수가 상당히 높은 것을 알 수 있습니다(거의 1의 상관관계). 즉, 인구수가 많을수록 그에 비례해 확진자도 많다는 뜻입니다. 이 수치가 의미하는 것은 GDP가 높은 나라에서는 인구가 많아도 감염을 예방할 수 있는 수 많은 인적자원과 물자, 개인이 비축할 수 있는 보급품들이 많아 어느정도 예방이 가능했다는 뜻입니다. 반면에, GDP가 낮은 나라의 경우에는 이러한 물자들이 부족하기 때문에 Covid-19를 예방할 수 있는 시민들의 경제적인 능력과 국가적 능력이 부족하다고 해석할 수 있습니다.

< 확진자, 인구수, 사망자, GDP, 완치 > : 마지막으로 국토면적 항목을 삭제하고 출력해보았습니다.

```
In [115]: gdp_top10 = df5.sort_values(by='GDP', ascending=False).head(10)
gdp_low10 = df5.sort_values(by='GDP', ascending=False).tail(10)

corr_1 = gdp_top10.loc[:, ['확진자', '인구수', '사망자', '완치', 'GDP']].corr(method=
corr_2 = gdp_low10.loc[:, ['확진자', '인구수', '사망자', '완치', 'GDP']].corr(method=

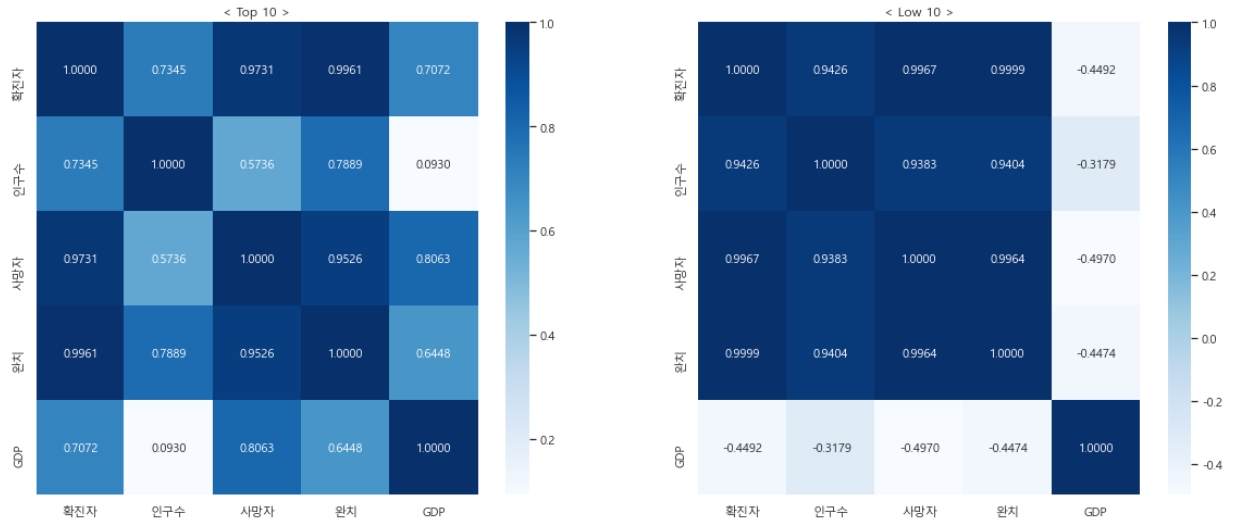
In [118]: fig = plt.figure(figsize=(20, 8))
fig.patch.set_facecolor('xkcd:white')

axis1 = fig.add_subplot(1, 2, 1)
axis1 = sns.heatmap(corr_1, annot=True, annot_kws={'size':11}, fmt='.4f', cmap='Blue:
plt.title('< Top 10 >')
axis1.plot()

axis2 = fig.add_subplot(1, 2, 2)
axis2 = sns.heatmap(corr_2, annot=True, annot_kws={'size':11}, fmt='.4f', cmap='Blue:
plt.title('< Low 10 >')
axis2.plot()

plt.show()
```





## GDP 상위 10개국, 하위 10개국 parameter 별 수치 비교 (중국 제외)

위에서 parameter 간의 상관관계를 비교했다면 이제는 수치를 비교해보자!

### < GDP 비교 >

```
In [146]: gdp_list = df5.sort_values(by='GDP', ascending=False)

top_names = list(gdp_list['국가'].head(10))
top_values = list(gdp_list['GDP'].head(10))
top_values = list(map(float, top_values))

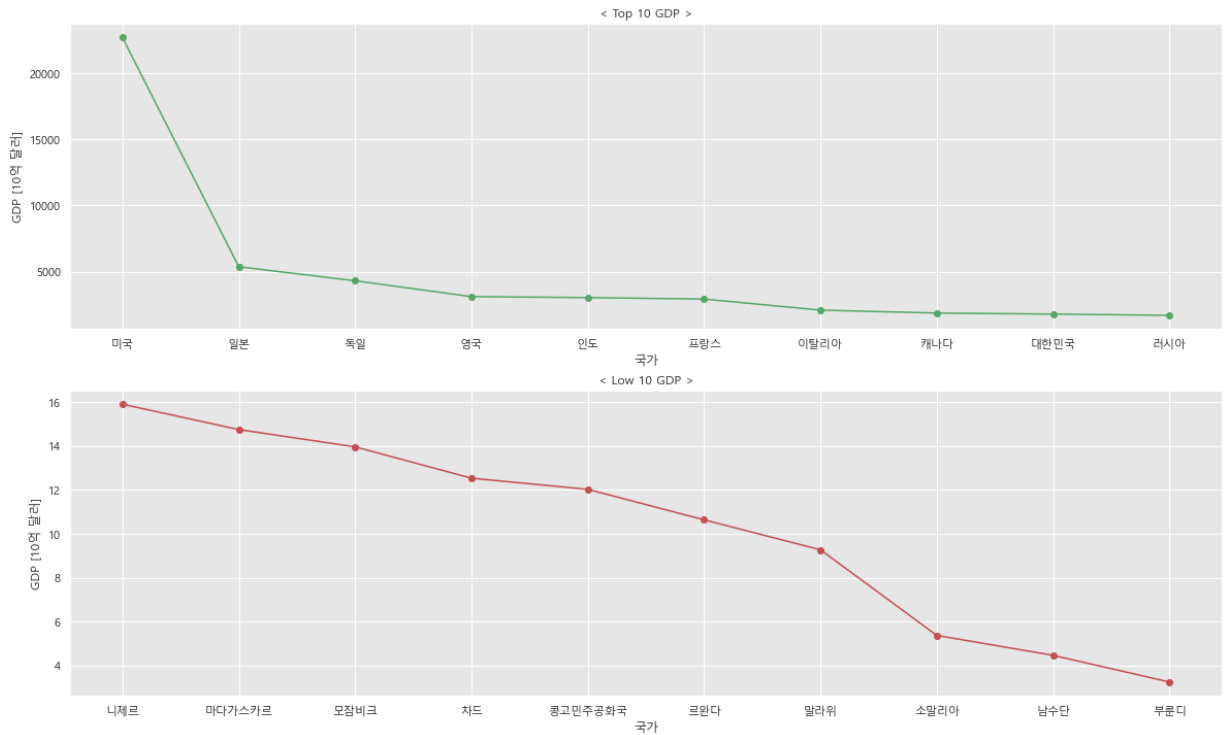
low_names = list(gdp_list['국가'].tail(10))
low_values = list(gdp_list['GDP'].tail(10))
low_values = list(map(float, low_values))

fig = plt.figure(figsize=(20, 12), facecolor='white')

axis1 = fig.add_subplot(2, 1, 1)
plt.plot(top_names, top_values, 'go-')
plt.title('< Top 10 GDP >')
plt.xlabel('국가')
plt.ylabel('GDP [10억 달러]')

axis2 = fig.add_subplot(2, 1, 2)
plt.plot(low_names, low_values, 'ro-')
plt.title('< Low 10 GDP >')
plt.xlabel('국가')
plt.ylabel('GDP [10억 달러]')

plt.show()
```



상위 10개국과 하위 10개국의 수치 차이가 상당히 심한 것을 볼 수 있습니다.

### < 확진자 & 인구수 비교 >

```
In [154]: top_names = list(gdp_list['국가'].head(10))
top_values_1 = list(gdp_list['확진자'].head(10))
top_values_1 = list(map(float, top_values_1))
top_values_2 = list(gdp_list['인구수'].head(10))
top_values_2 = list(map(float, top_values_2))

low_names = list(gdp_list['국가'].tail(10))
low_values_1 = list(gdp_list['확진자'].tail(10))
low_values_1 = list(map(float, low_values_1))
low_values_2 = list(gdp_list['인구수'].tail(10))
low_values_2 = list(map(float, low_values_2))

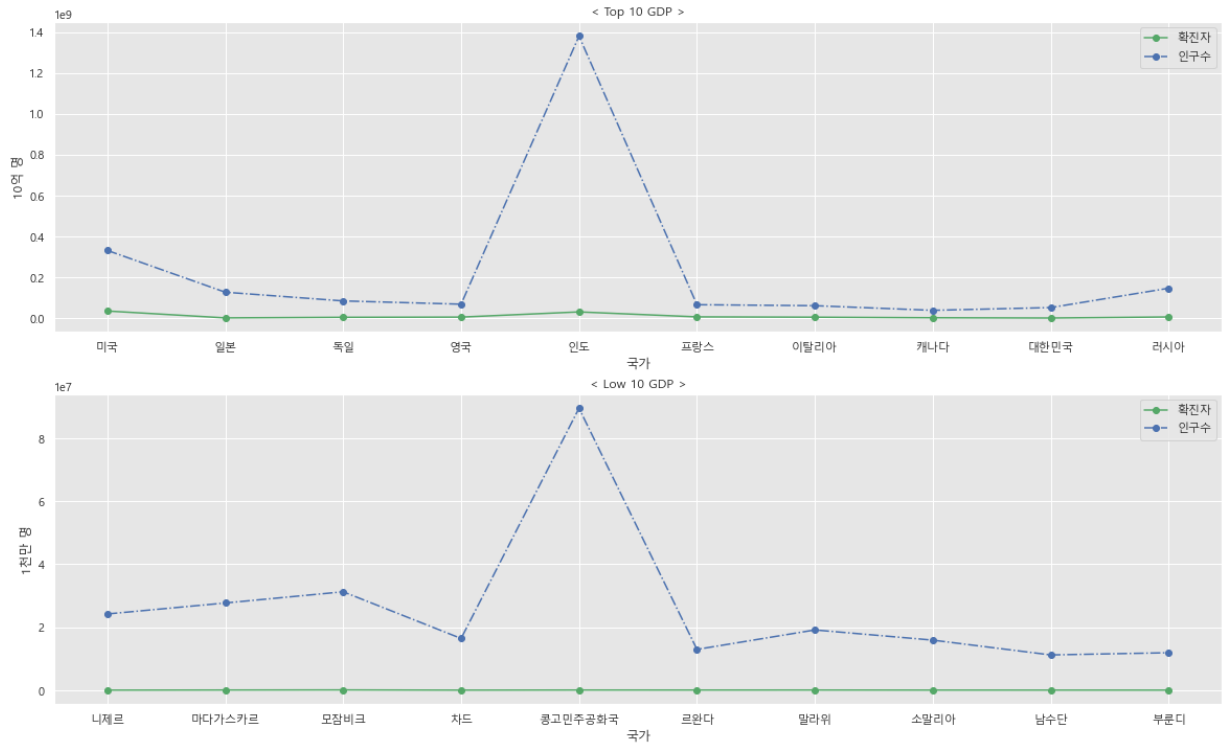
fig = plt.figure(figsize=(20, 12), facecolor='white')

axis1 = fig.add_subplot(2, 1, 1)
plt.plot(top_names, top_values_1, 'go-')
plt.plot(top_names, top_values_2, 'bo-.')
plt.title('< Top 10 GDP >')
plt.legend(['확진자', '인구수'])
plt.xlabel('국가')
plt.ylabel('10억 명')

axis2 = fig.add_subplot(2, 1, 2)
plt.plot(low_names, low_values_1, 'go-')
plt.plot(low_names, low_values_2, 'bo-.')
plt.title('< Low 10 GDP >')
plt.legend(['확진자', '인구수'])
plt.xlabel('국가')
plt.ylabel('1천만 명')

plt.show()

print('Top 10 인구수 :', top_values_2)
print('Top 10 확진자수 :', top_values_1)
print('Low 10 인구수 :', low_values_2)
print('Low 10 확진자수 :', low_values_1)
```



Top 10 인구수 : [331002651.0, 126476461.0, 83783942.0, 67886011.0, 1380004385.0, 65273511.0, 60461826.0, 37742154.0, 51269185.0, 145934462.0]  
 Top 10 확진자수 : [34419838.0, 785969.0, 3730599.0, 4640507.0, 29977861.0, 5757798.0, 4253460.0, 1409607.0, 151901.0, 5334204.0]  
 Low 10 인구수 : [24206644.0, 27691018.0, 31255435.0, 16425864.0, 89561403.0, 12952218.0, 19129952.0, 15893222.0, 11193725.0, 11890784.0]  
 Low 10 확진자수 : [5469.0, 42137.0, 72577.0, 4947.0, 37926.0, 31435.0, 34914.0, 14867.0, 10786.0, 5242.0]

## 확진자 비율이 낮은 관계로 확진자만 다시 출력

In [160]:

```
top_names = list(gdp_list['국가'].head(10))
top_values = list(gdp_list['확진자'].head(10))
top_values = list(map(float, top_values))

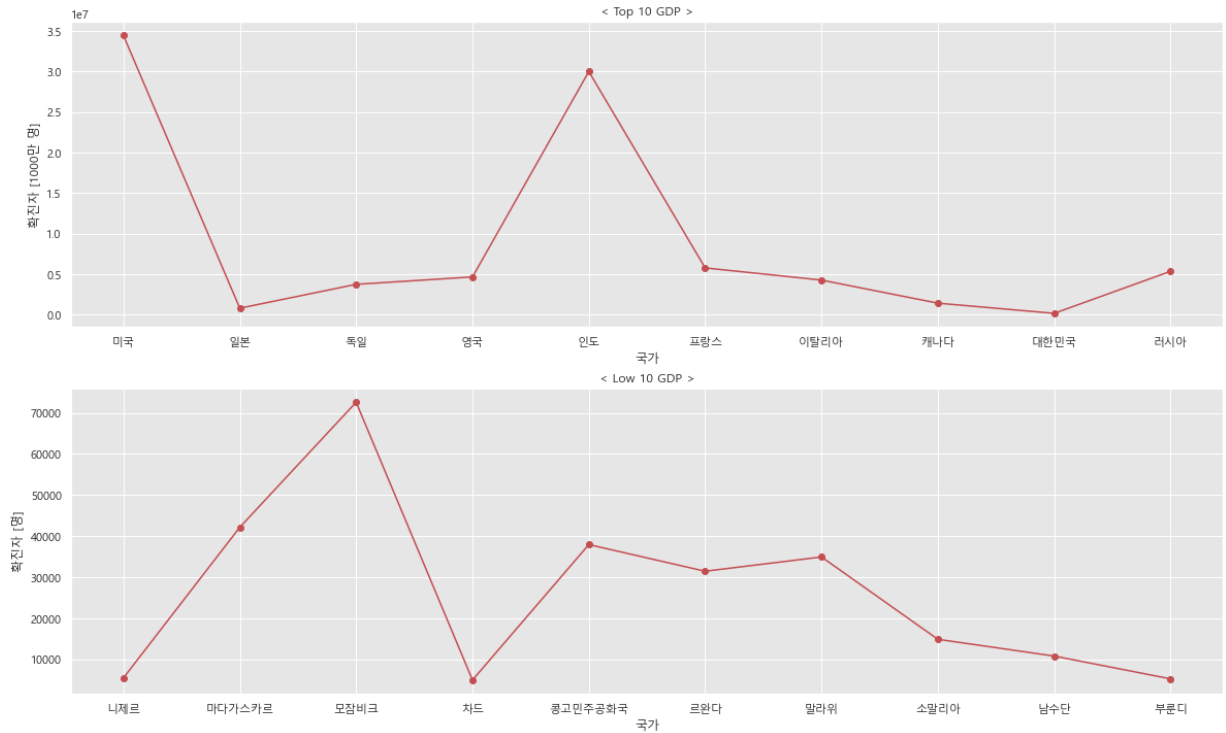
low_names = list(gdp_list['국가'].tail(10))
low_values = list(gdp_list['확진자'].tail(10))
low_values = list(map(float, low_values))

fig = plt.figure(figsize=(20, 12), facecolor='white')

axis1 = fig.add_subplot(2, 1, 1)
plt.plot(top_names, top_values, 'ro-')
plt.title('< Top 10 GDP >')
plt.xlabel('국가')
plt.ylabel('확진자 [1000만 명]')

axis2 = fig.add_subplot(2, 1, 2)
plt.plot(low_names, low_values, 'ro-')
plt.title('< Low 10 GDP >')
plt.xlabel('국가')
plt.ylabel('확진자 [명]')

plt.show()
```



## < 사망자 비교 >

```
In [161]: top_names = list(gdp_list['국가'].head(10))
top_values = list(gdp_list['사망자'].head(10))
top_values = list(map(float, top_values))

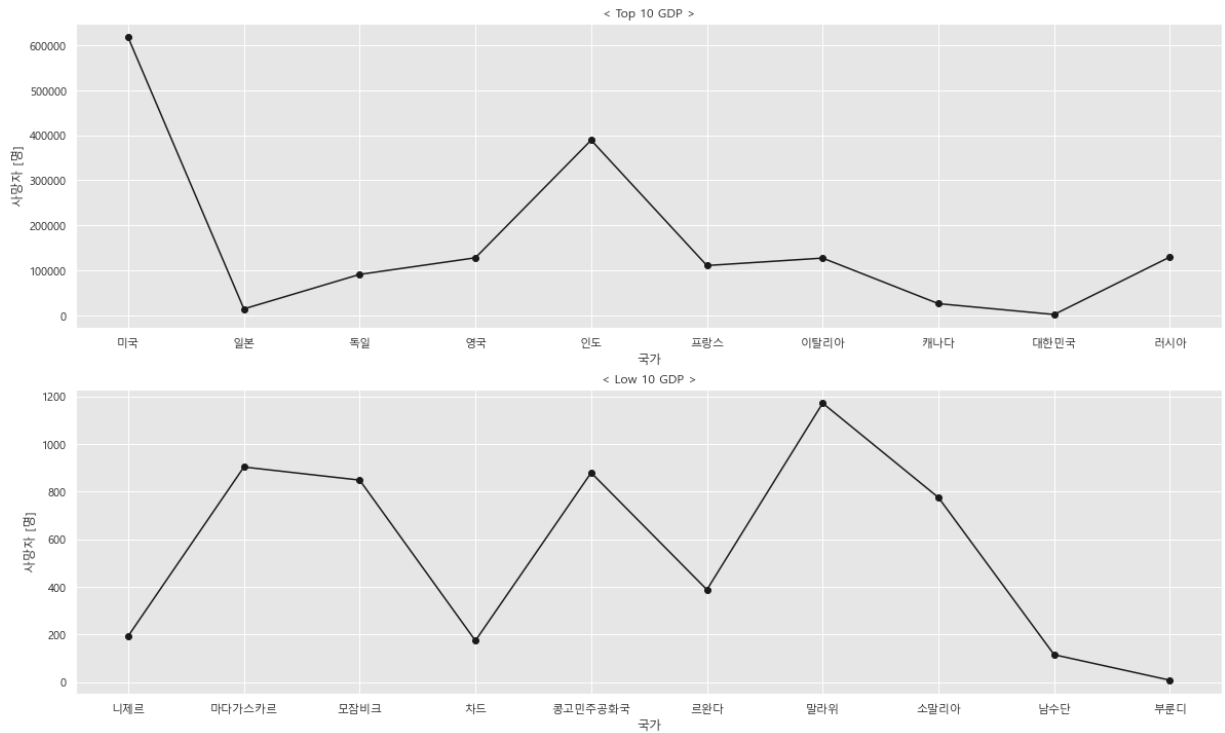
low_names = list(gdp_list['국가'].tail(10))
low_values = list(gdp_list['사망자'].tail(10))
low_values = list(map(float, low_values))

fig = plt.figure(figsize=(20, 12), facecolor='white')

axis1 = fig.add_subplot(2, 1, 1)
plt.plot(top_names, top_values, 'ko-')
plt.title('< Top 10 GDP >')
plt.xlabel('국가')
plt.ylabel('사망자 [명]')

axis2 = fig.add_subplot(2, 1, 2)
plt.plot(low_names, low_values, 'ko-')
plt.title('< Low 10 GDP >')
plt.xlabel('국가')
plt.ylabel('사망자 [명]')

plt.show()
```



## < 완치자 비교 >

```
In [157]: top_names = list(gdp_list['국가'].head(10))
top_values = list(gdp_list['완치'].head(10))
top_values = list(map(float, top_values))

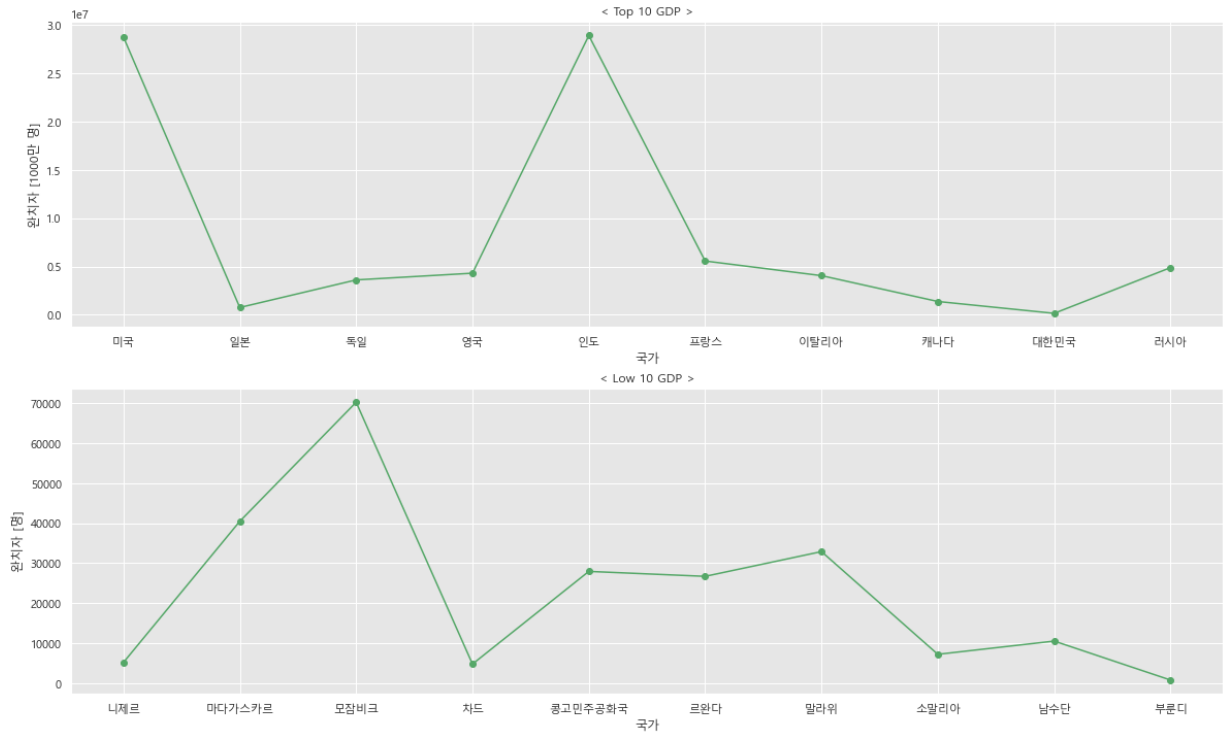
low_names = list(gdp_list['국가'].tail(10))
low_values = list(gdp_list['완치'].tail(10))
low_values = list(map(float, low_values))

fig = plt.figure(figsize=(20, 12), facecolor='white')

axis1 = fig.add_subplot(2, 1, 1)
plt.plot(top_names, top_values, 'go-')
plt.title('< Top 10 GDP >')
plt.xlabel('국가')
plt.ylabel('완치자 [1000만 명]')

axis2 = fig.add_subplot(2, 1, 2)
plt.plot(low_names, low_values, 'go-')
plt.title('< Low 10 GDP >')
plt.xlabel('국가')
plt.ylabel('완치자 [명]')

plt.show()
```



## < 인구밀도 비교 >

```
In [159]: top_names = list(gdp_list['국가'].head(10))
top_values = list(gdp_list['인구밀도'].head(10))
top_values = list(map(float, top_values))

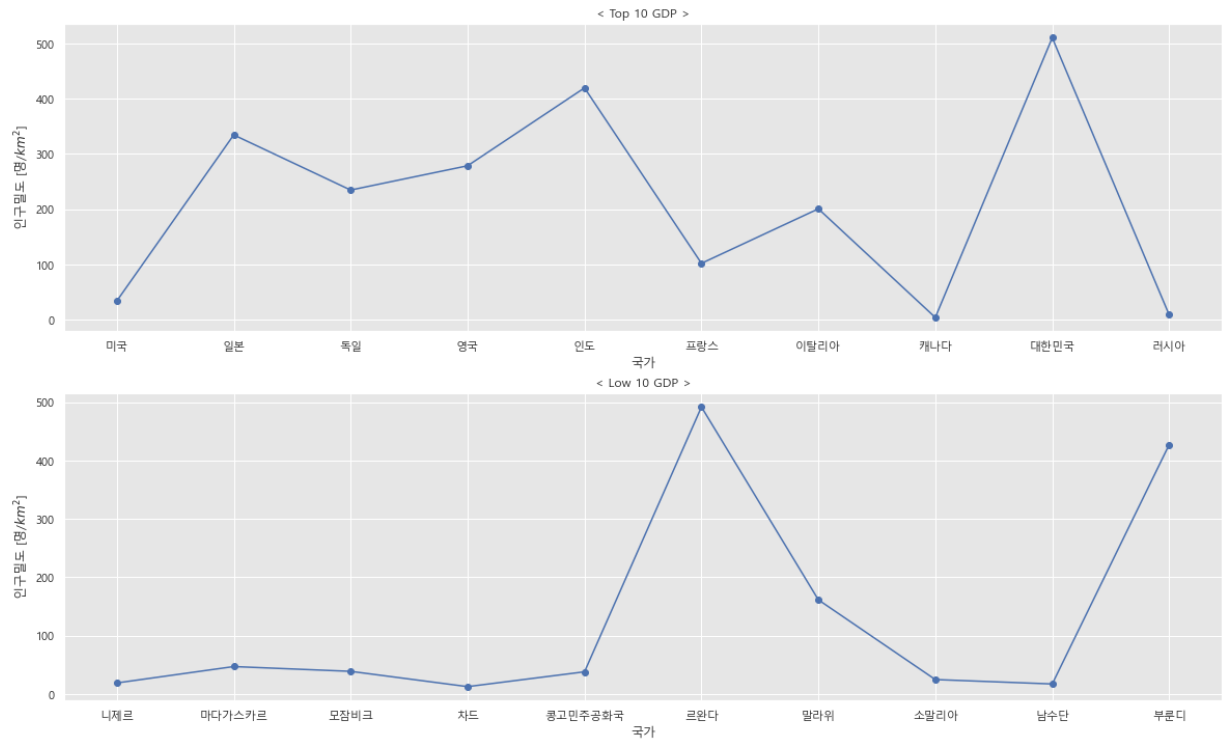
low_names = list(gdp_list['국가'].tail(10))
low_values = list(gdp_list['인구밀도'].tail(10))
low_values = list(map(float, low_values))

fig = plt.figure(figsize=(20, 12), facecolor='white')

axis1 = fig.add_subplot(2, 1, 1)
plt.plot(top_names, top_values, 'bo-')
plt.title('< Top 10 GDP >')
plt.xlabel('국가')
plt.ylabel('인구밀도 [명/$km^2$]')

axis2 = fig.add_subplot(2, 1, 2)
plt.plot(low_names, low_values, 'bo-')
plt.title('< Low 10 GDP >')
plt.xlabel('국가')
plt.ylabel('인구밀도 [명/$km^2$]')

plt.show()
```



## < 치명율 비교 >

```
In [163]: gdp_list_1 = df5.sort_values(by='GDP', ascending=False)

gdp_list_1['치명율'] = gdp_list_1['사망자']/gdp_list_1['확진자'] # 치명율 Column 추가

top_names = list(gdp_list_1['국가'].head(10))
top_values = list(gdp_list_1['치명율'].head(10))
top_values = list(map(float, top_values))

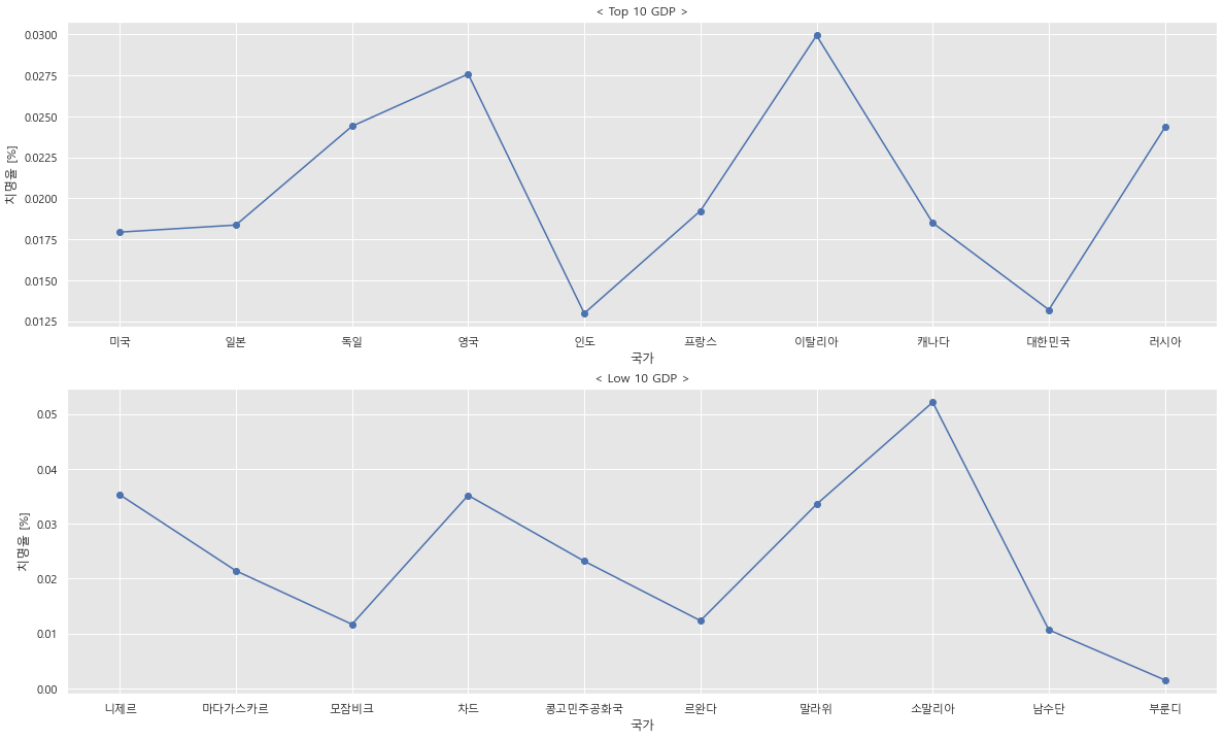
low_names = list(gdp_list_1['국가'].tail(10))
low_values = list(gdp_list_1['치명율'].tail(10))
low_values = list(map(float, low_values))

fig = plt.figure(figsize=(20, 12), facecolor='white')

axis1 = fig.add_subplot(2, 1, 1)
plt.plot(top_names, top_values, 'bo-')
plt.title('< Top 10 GDP >')
plt.xlabel('국가')
plt.ylabel('치명율 [%]')

axis2 = fig.add_subplot(2, 1, 2)
plt.plot(low_names, low_values, 'bo-')
plt.title('< Low 10 GDP >')
plt.xlabel('국가')
plt.ylabel('치명율 [%]')

plt.show()
```



In [ ]: