

4장. 판다스 데이터프레임

학번	32233421	이름	이은지
----	----------	----	-----

※ 제공한 GNI2014.csv 데이터셋에 대해 다음 문제를 해결하시오. 컬럼에 대한 설명은 다음과 같다.

iso3	국가코드
country	국가명
continent	소속 대륙명
population	인구수
GNI	국민총소득

(각 문제별로 코드 & 실행결과화면을 캡처를 제출. 실행 결과가 길 경우에는 결과의 앞쪽 일부만 자라서 캡쳐해도 됨)

1. GNI2014.csv 파일을 읽어서 df에 저장하고 df 앞쪽 5개 행의 내용을 보이시오.

```
>>> df.head()
   iso3      country  continent  population      GNI
0  BMU    Bermuda  North America     67837  106140
1  NOR      Norway     Europe    4676305  103630
2  QAT       Qatar      Asia    833285  92200
3  CHE  Switzerland     Europe    7604467  88120
4  MAC  Macao SAR, China      Asia    559846  76270
```

데이터사이언스

2. df의 행의 개수와 열의 개수를 각각 보이시오

```
>>> df.shape[0]  
188  
>>> df.shape[1]  
5
```

3. 소속 대륙의 종류를 보이시오

```
>>> df['continent'].unique()  
array(['North America', 'Europe', 'Asia', 'Oceania', 'South America',  
       'Seven seas (open ocean)', 'Africa'], dtype=object)
```

4. 'Norway'의 인구수와 국민총생산을 보이시오.

```
>>> df.loc[[1],['population','GNI']]  
           population      GNI  
1        4676305    103630
```

5. 유럽 대륙에 속한 국가들의 목록을 보이시오.

```
>>> df.loc[df.continent=='Europe']  
           iso3    ...      GNI  
1        NOR    ...    103630  
3        CHE    ...    88120  
5        LUX    ...    75990  
7        SWE    ...    61610  
8        DNK    ...    61310  
11       NLD    ...    51890  
13       AUT    ...    49670
```

6. 인구수가 5000만명 이상인 국가들의 목록을 보이시오

```
>>> df.loc[df.population>=50000000]  
           iso3          country    ...  population  
           GNI  
9        USA      United States    ...    313973000  5  
5200  
16       DEU      Germany    ...    82329758  4  
7640  
21       GBR      United Kingdom    ...    62262000  4  
3430  
22       FRA      France    ...    64057792  4
```

데이터사이언스

7. 유럽 국가들의 평균 인구수를 보이시오.

```
>>> df2=df['population']
>>> df3=df2.loc[df.continent=='Europe']
>>> df3.mean()
np.float64(17772437.780487806)
```

8. 'Canada' 보다 국민총생산액이 많은 나라들의 평균 국민총생산액을 보이시오

```
>>> df2=df.loc[(df.GNI>df.loc[12,'GNI'], 'GNI')]
>>> df2.mean()
np.float64(74337.5)
```

9. North America 국가들에 대해 인구가 10% 늘었을 때의 인구수를 보이시오.

```
>>> df3=df2+df2*0.1
>>> df3
0          74620.7
9        345370300.0
12        36835928.8
38        340071.6
40        1441000.0
41        4368122.0
```

10. 유럽 국가들에 대해 국가별 1인당 생산액(총생산액/인구수)을 보이시오

```
>>> df2=df.loc[(df.continent=='Europe'), ['country','population','GNI']]
>>> df3=df2.GNI/df2.population
>>> df3
1      0.022161
3      0.011588
5      0.154522
```

데이터사이언스

11. 유럽 국가들에 대해 국가별 1인당 생산액(총생산액/인구수)을 'PNP'라는 이름의 컬럼으로 해서 df에 추가 한 후 앞쪽 5개의 행의 내용을 보이시오

```
>>> new_col=df2.GNI/df2.population
>>> df['PNP']=new_col
>>> df
   iso3              country  continent  population      GNI      PNP
0  BMU        Bermuda  North America     67837  106140      NaN
1  NOR          Norway      Europe    4676305  103630  0.022161
2  QAT           Qatar       Asia    833285  92200      NaN
3  CHE  Switzerland      Europe    7604467  88120  0.011588
4  MAC  Macao SAR, China       Asia    559846  76270      NaN
...  ...            ...      ...     ...  ...  ...

```