

# 지구 온난화 현상 현황 파악 및 원인 분석

온실가스(CO<sub>2</sub>, CH<sub>4</sub>)와 외부데이터 관계 분석을 중심으로

2019112486 고혜영 산업시스템공학과  
2016110906 조성택 경제학과

# 목차



## 주제 설명

- 1) 주제 선정 이유
- 2) 용어 설명
- 3) 온실가스



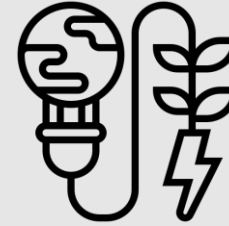
## CO2

- 1) 데이터 설명 & 전처리
- 2) CO2 배출 현황
- 3) EDA + 외부데이터 연계



## CH4

- 1) 데이터 설명 & 전처리
- 2) CO2 배출 현황
- 3) EDA + 외부데이터 연계



## 재생에너지

- 1) 데이터 설명 & 전처리
- 2) EDA + 외부데이터 연계



## 한계점 & 보완점

# 주제 선정 이유

## "올해 '수능한파' 없어...14일부터 평년 보다 '따뜻'

기사등록 : 2021-11-11 13:30

가 + 가 -



## 국내 탄소배출권 '금융화'의 길로



한광덕 기자

+ 구독



[서울=뉴스핌] 최현민 기자 = 오는 18일 대학수학능력시험(수능)은 평년보다 따뜻한 날씨속에 치러질 전망이다.

기상청은 11일 "지난 8일 시작된 추위가 토요일인 오는 13일 오전까지 이어지겠다"며 "오후부터 기온이 회복되면서 14일에는 평년보다 높은 기온을 보일 것"이라고 예측했다.

증권사들 내달부터 자기매매  
유동성 불어넣어 가격상승 유도  
탄소저감 등 친환경 투자 기대  
선물도 도입되면 '투자상품'으로

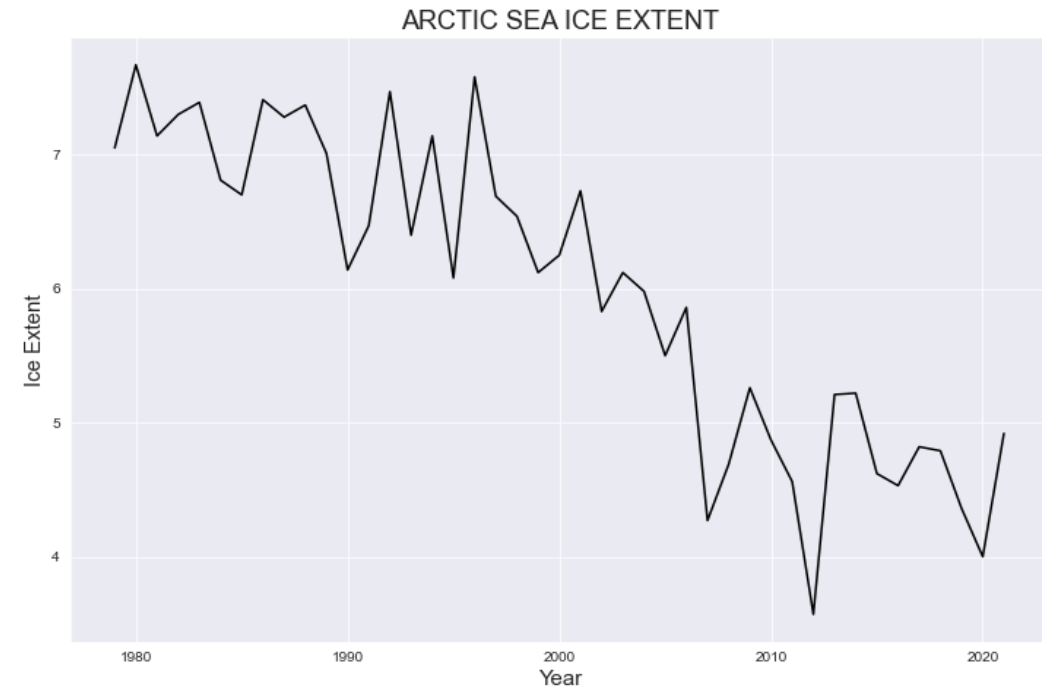
## [이제는 ESG 경영] 정부, ESG 인프라 확충 지원 나선다!

✎ 이정윤 기자 | ⌚ 입력 2021.10.18 06:00 | 💬 댓글 0

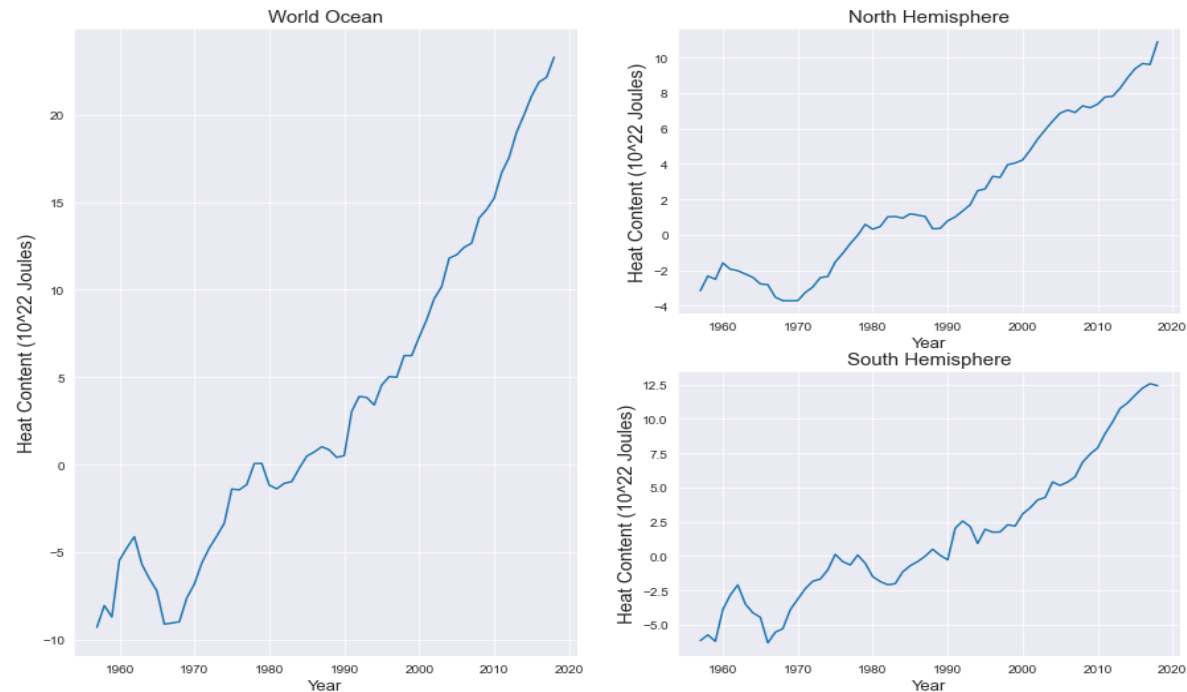
- 최근 **탄소배출권 규제**, **전기차** 등 친환경 사업들이 주목받고 있음.
- UN을 포함한 다양한 범국가기관들에서 **기후변화의 심각성**을 강조하고 있음

# 기후변화란?

- '기후변화'란 지구 온난화를 포함하여 **해수면 상승, 산악 빙하 축소, 식물 개화시기 변화** 등 온난화로 발생한 다양한 결과들을 의미함



<표1> 북극 해빙 규모 (Seaborn)

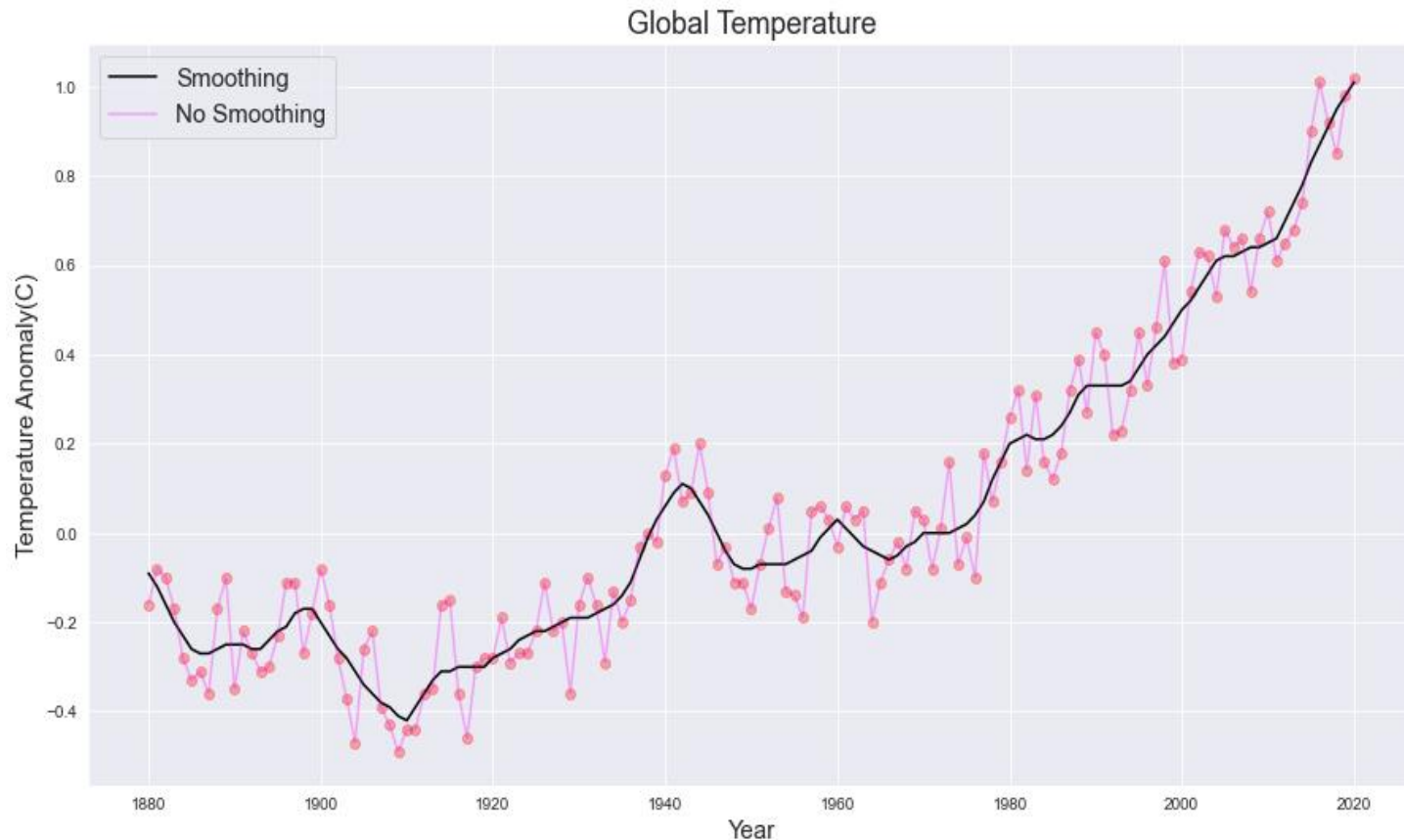


<표2> 세계 및 북반구와 남반구 해수 온도 (Matplotlib)

- 북극의 해빙 면적은 단기적으로는 증가, 감소를 반복하지만, **장기적으로는 감소**
- 해수 온도의 경우, 북반구와 남반구 모두 **꾸준히 상승**하는 추세를 보임

# 지구 온난화란?

- '지구 온난화'는 지구의 장기적인 온도 상승을 의미함



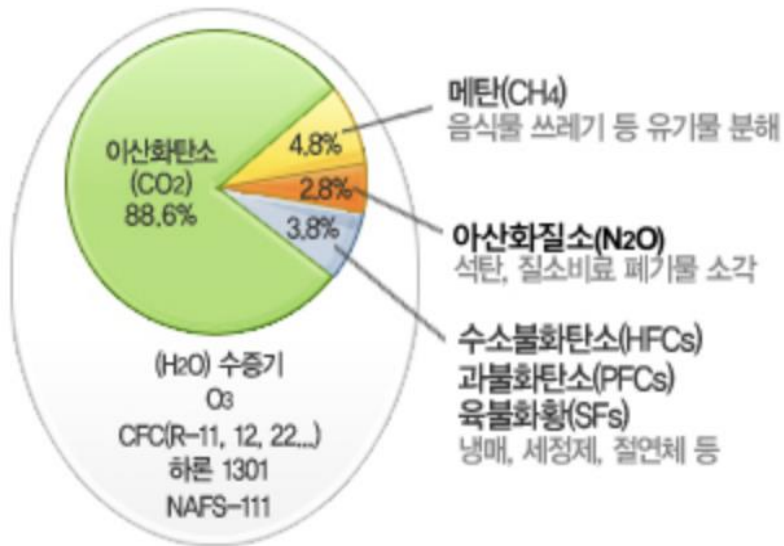
<표3> 연도별 표면 온도 상승량 (Seaborn)

➤ 표 해석: 1.02 °C(2020), -0.16 °C(1880)  
= 140년간 표면 온도가 1.18 상승했다

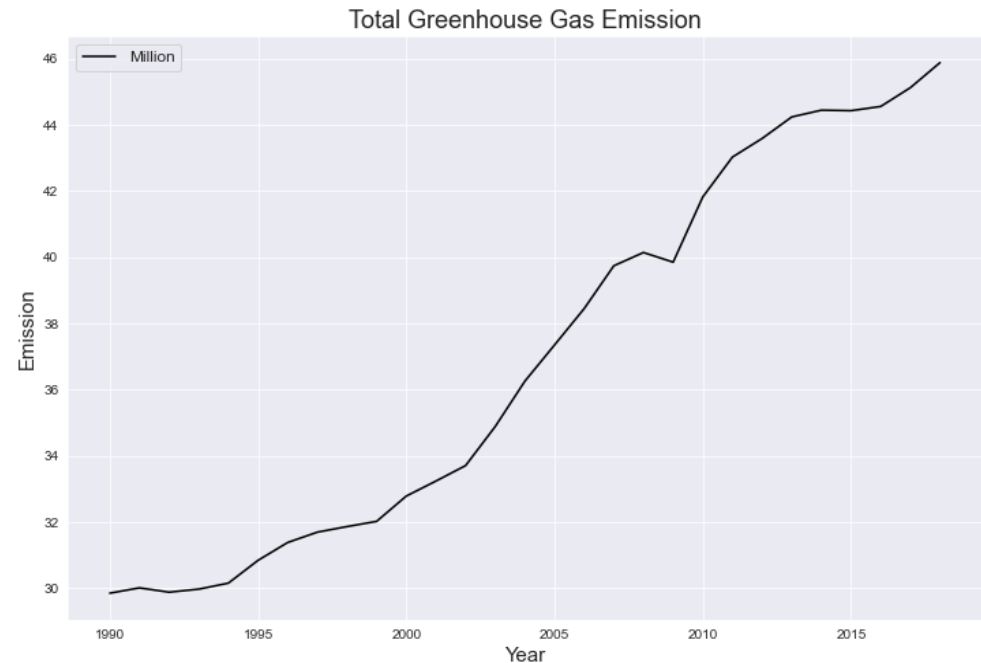
➤ 1910년 이후에는 꾸준히 상승함을  
알 수 있다

# 지구 온난화란?

- 지구 온난화의 원인은 온실가스의 증가 때문인데 교토의정서에 따르면 온실가스에는 다음과 같이 6가지가 포함됨



<그림1> 온실가스 6종



<표4> 세계 온실가스 배출량 (Seaborn)

➢ 온실가스 배출량 또한 꾸준히 상승함을 알 수 있음

해당 프로젝트에서는 6가지의 온실가스들 중에서 주요 온실가스인 CO<sub>2</sub>와 CH<sub>4</sub>에 집중하여 배출 현황과 원인을 분석해보고자 함

<그림1> 출처: GS칼텍스 미디어허브  
(<https://gscltexmediahub.com/energy/about-greenhouse-gas/>)

CO2 배출량

# CO2 데이터 설명 및 전처리 요약

- 데이터명: CO2 Emission (kt)
- World Bank data
- 국가별 연도에 따른 CO2 배출량
- Wide Form Pannel 형태

	Country Name	Country Code	Indicator Name	Indicator Code	1960	1961	1962	1963	1964
0	Aruba	ABW	CO2 emissions (kt)	EN.ATM.CO2E.KT	11092.67500	11576.71900	12713.48900	12178.10700	11840.74300
1	Africa Eastern and Southern	AFE	CO2 emissions (kt)	EN.ATM.CO2E.KT	118545.90130	123758.90330	128093.89780	132810.33250	144345.35240
2	Afghanistan	AFG	CO2 emissions (kt)	EN.ATM.CO2E.KT	414.37100	491.37800	689.39600	707.73100	839.74300
3	Africa Western and Central	AFW	CO2 emissions (kt)	EN.ATM.CO2E.KT	8760.46300	9376.51900	9710.21600	11540.04900	13985.93800
4	Angola	AGO	CO2 emissions (kt)	EN.ATM.CO2E.KT	550.05000	454.70800	1180.77400	1151.43800	1224.77800

- 핵심적인 전처리 내용
- 일부 변수만 사용
- 결측치 처리 : bfill , ffill

[bfill]

	1966	1967	1968	1969	1970	1971	1972	1973
0	NaN	NaN	NaN	NaN	10.2469	10.4531	10.6570	10.8551
1	NaN	NaN	NaN	NaN	384746.7182	265076.7091	278192.0872	296153.8053
2	NaN	NaN	NaN	NaN	10202.0000	10201.5000	9170.5900	9403.5400
3	NaN	NaN	NaN	NaN	143714.5355	145309.3393	149181.1477	161770.8276
4	NaN	NaN	NaN	NaN	23376.6000	14500.0000	15590.0000	16651.7000



	1966	1967	1968	1969	1970	1971	1972	1973
0	10.2469	10.2469	10.2469	10.2469	10.2469	10.4531	10.657	10.8551
1	384746.7182	384746.7182	384746.7182	384746.7182	384746.7182	265076.7091	278192.0872	296153.8053
2	10202.0	10202.0	10202.0	10202.0	10202.0	10201.5	9170.59	9403.54
3	143714.5355	143714.5355	143714.5355	143714.5355	143714.5355	145309.3393	149181.1477	161770.8276
4	23376.6	23376.6	23376.6	23376.6	23376.6	14500.0	15590.0	16651.7

[ffill]

	2016	2017	2018	2019	2020
	NaN	NaN	NaN	NaN	NaN
563050.0	564380.0	570040.0	NaN	NaN	
78150.0	79610.0	81510.0	NaN	NaN	
383250.0	386400.0	392130.0	NaN	NaN	
37400.0	37450.0	35520.0	NaN	NaN	



	2016	2017	2018	2019	2020
	14.8981	14.8981	14.8981	14.8981	14.8981
563050.0000	564380.0000	570040.0000	570040.0000	570040.0000	
78150.0000	79610.0000	81510.0000	81510.0000	81510.0000	
383250.0000	386400.0000	392130.0000	392130.0000	392130.0000	
37400.0000	37450.0000	35520.0000	35520.0000	35520.0000	

- Country Name 컬럼에서 소득 수준에 따라 그룹화된 그룹명, UN, EU, SAARC, G8, G20과 같은 초국가적 조직명을 제외한 country\_list 리스트를 생성

```
country_list=['Afghanistan','Angola','Albania','Argentina','Armenia','Australia',
'Austria','Azerbaijan','Burundi','Belgium','Benin','Burkina Faso','Bangladesh','Bulgaria',
'Bahrain','Bosnia and Herzegovina','Belarus','Belize','Bolivia','Brazil','Barbados','Brunei Darussalam',
'Bhutan','Botswana','Central African Republic','Canada','Switzerland','Chile','China','Cameroon',
'Congo','Colombia','Comoros','Cabo Verde','Costa Rica','Cuba','Cyprus','Czech Republic','Germany',
'Denmark','Dominican Republic','Algeria','Ecuador','Egypt','Spain','Estonia','Ethiopia','Finland','Fiji',
'France','Gabon','United Kingdom','Georgia','Ghana','Guinea','Guatemala','Guyana','Hong Kong',
'Honduras','Croatia','Haiti','Hungary','Indonesia','India','Ireland','Iran','Iraq','Iceland','Israel',
'Italy','Jamaica','Jordan','Japan','Kazakhstan','Kenya','Cambodia','Korea, Rep.','Kuwait','Lebanon','Liberia',
'Libya','Sri Lanka','Lesotho','Lithuania','Luxembourg','Latvia','Macao','Morocco','Moldova','Madagascar',
'Maldives','Mexico','Macedonia','Mali','Malta','Myanmar','Montenegro','Mongolia','Mozambique','Mauritania',
'Mauritius','Malawi','Malaysia','Namibia','Niger','Nigeria','Nicaragua','Netherlands',
'Norway','Nepal','New Zealand','Oman','Pakistan','Panama','Peru','Philippines','Papua New Guinea',
'Poland','Puerto Rico','Portugal','Paraguay','Qatar','Romania','Russian Federation','Rwanda','Saudi Arabia',
'Sudan','Senegal','Singapore','Solomon Islands','Sierra Leone','El Salvador','Somalia','Serbia','Slovenia',
'Sweden','Swaziland','Syrian Arab Republic','Chad','Thailand','Tajikistan','Turkmenistan','Togo',
'Trinidad and Tobago','Tunisia','Turkey','Tanzania','Uganda','Ukraine','Uruguay','United States','Uzbekistan',
'Vietnam','Yemen, Rep.','Congo, Dem. Rep.','Zambia','Zimbabwe']
```



# World Bank Data의 공통적 전처리

1. 분석에 필요하지 않은 열(Column)제거

```
co2_meta_data.drop(columns=['SpecialNotes', 'TableName', 'Unnamed: 5'], axis=1, inplace=True)
co2_value_data.drop(columns=['Indicator Name', 'Indicator Code'], axis=1, inplace=True)
```

2. 시계열 데이터 값의 결측치 처리

*#결측치 처리 (bfill)*

```
co2_value_data.iloc[:, 2:] = co2_value_data.iloc[:, 2:].fillna(method='bfill', axis=1)
f_co2_value_data = co2_value_data
```

*#결측치 처리 (ffill)*

```
f_co2_value_data.iloc[:, 2:] = f_co2_value_data.iloc[:, 2:].fillna(method='ffill', axis=1)
final_co2_value_data = f_co2_value_data
```

3. Metadata데이터와 API value데이터를 병합(merge)

```
co2_mergedData = pd.merge(co2_meta_data, final_co2_value_data, on=["country_code"], how="inner")
co2_mergedData
```

4. 병합한 파일에서 시계열 데이터값이 음수인 행을 제거

```
minus_list=[]
for col in cols:
    idx=year_data_co2[year_data_co2[col]<0].index
    if (len(idx))>0:
        minus_list.append(idx)
```

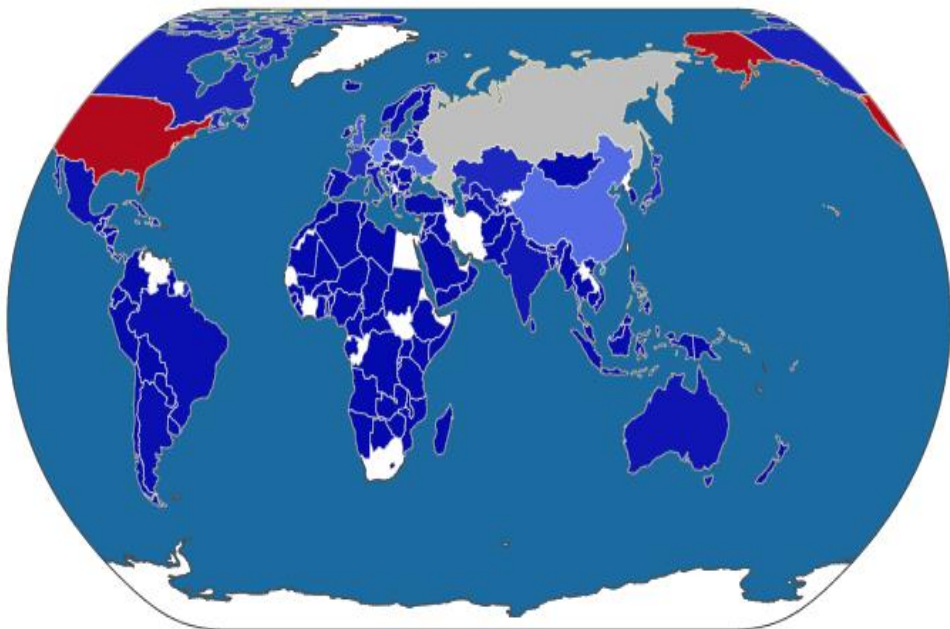
```
co2_mergedData.drop(minus_list[0], inplace=True)
```

5. Country Name이 개별 국가에 해당하는 데이터가 아닌 데이터를 제거

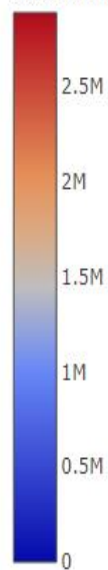
```
co2_mergedData=co2_mergedData[co2_mergedData['country_name'].isin(country_list)]
```

# 세계 CO2 배출량 비교(1960년, 2018년)

World Map of CO2 emission in the Year 1960

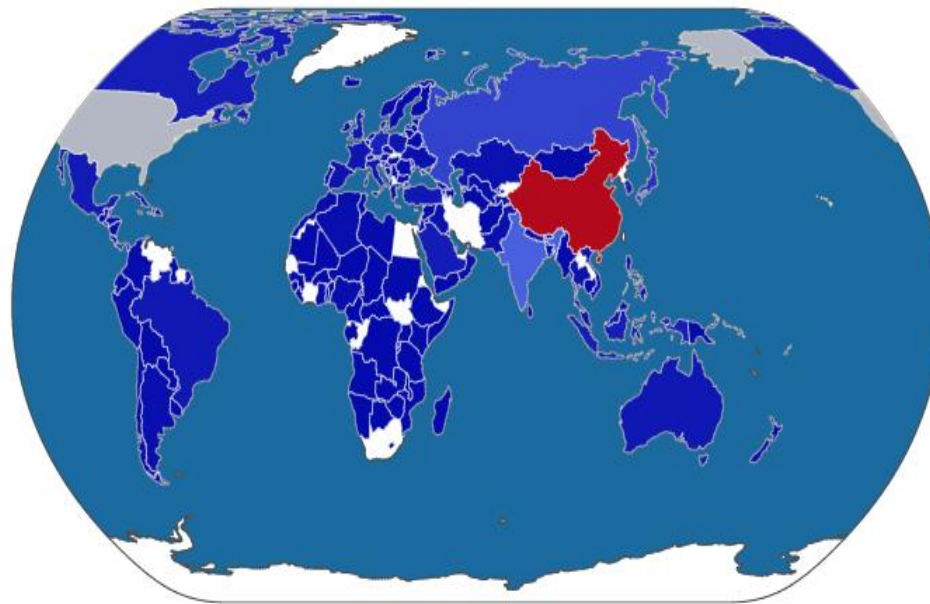


CO2 emission

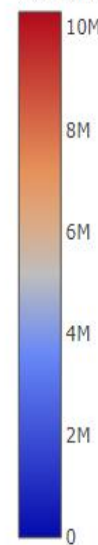


<그림2> 1960년 CO2 배출량 지도 (Plotly)

World Map of CO2 emission in the Year 2018



CO2 emission



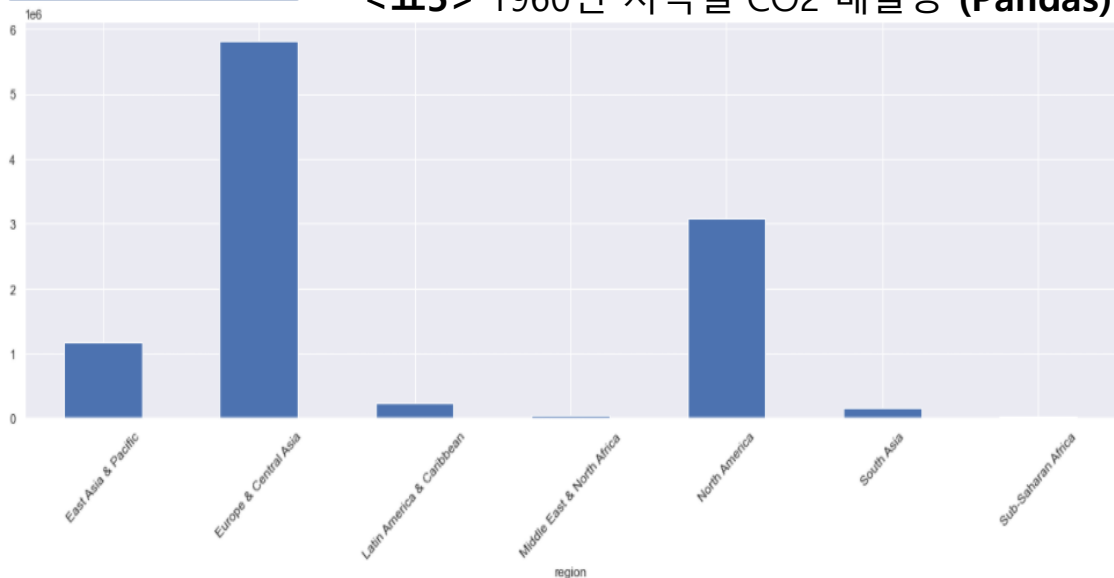
<그림3> 2018년 CO2 배출량 지도 (Plotly)

1960년대에 비해 **2018년에 CO2배출량이 확연히 많음**을 알 수 있음

# 지역별 CO2배출량 비교(1960년,2018년)

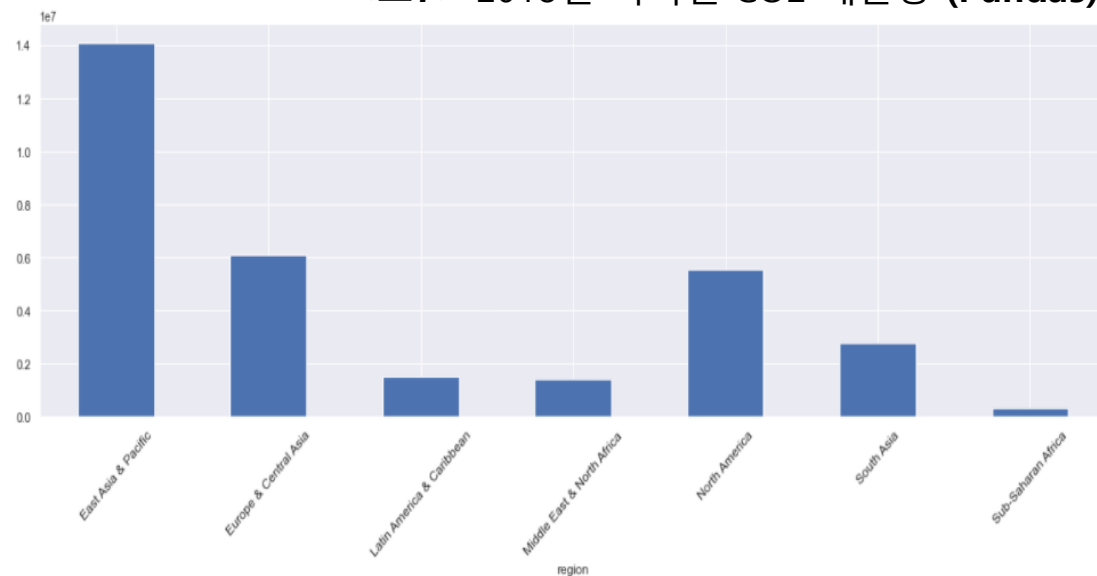
1960년

<표5> 1960년 지역별 CO2 배출량 (Pandas)



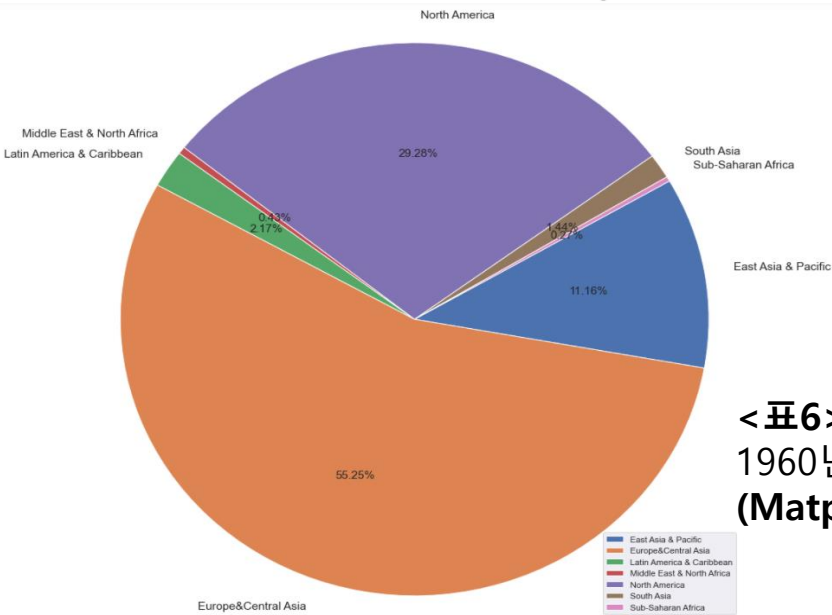
2018년

<표7> 2018년 지역별 CO2 배출량 (Pandas)

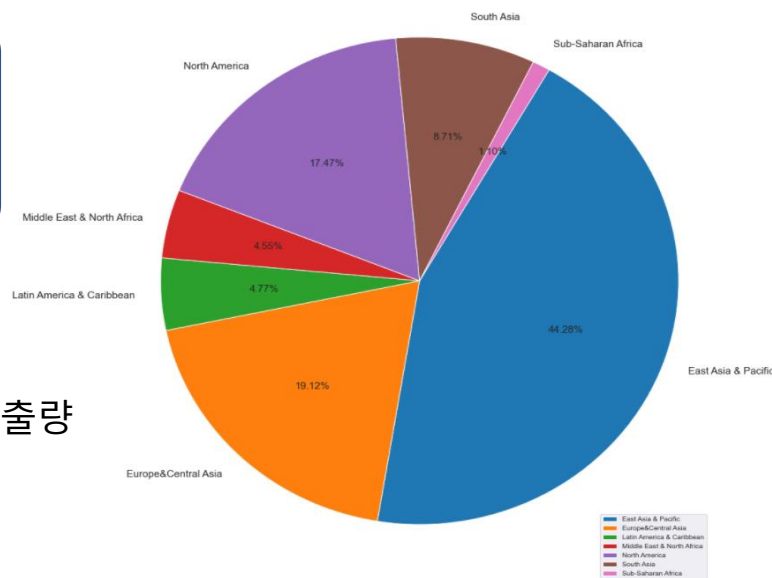


1960년: 유럽&중앙 아시아 55.25%( 1위 )  
2018년: 동아시아&태평양 국가 44.28% ( 1위 )

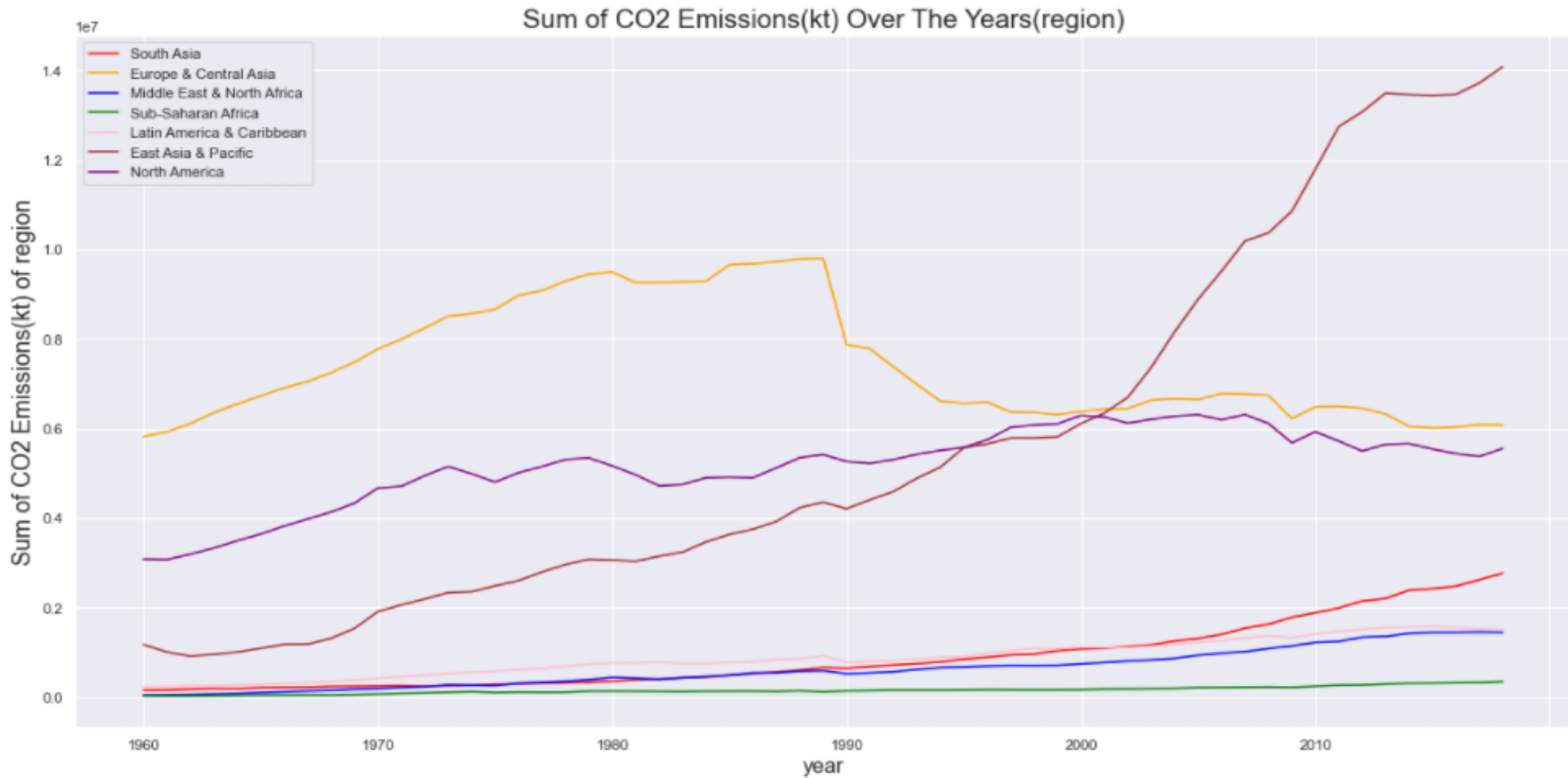
<표6>  
1960년 지역별 CO2 배출량  
(Matplotlib)



<표8>  
2018년 지역별 CO2 배출량  
(Matplotlib)

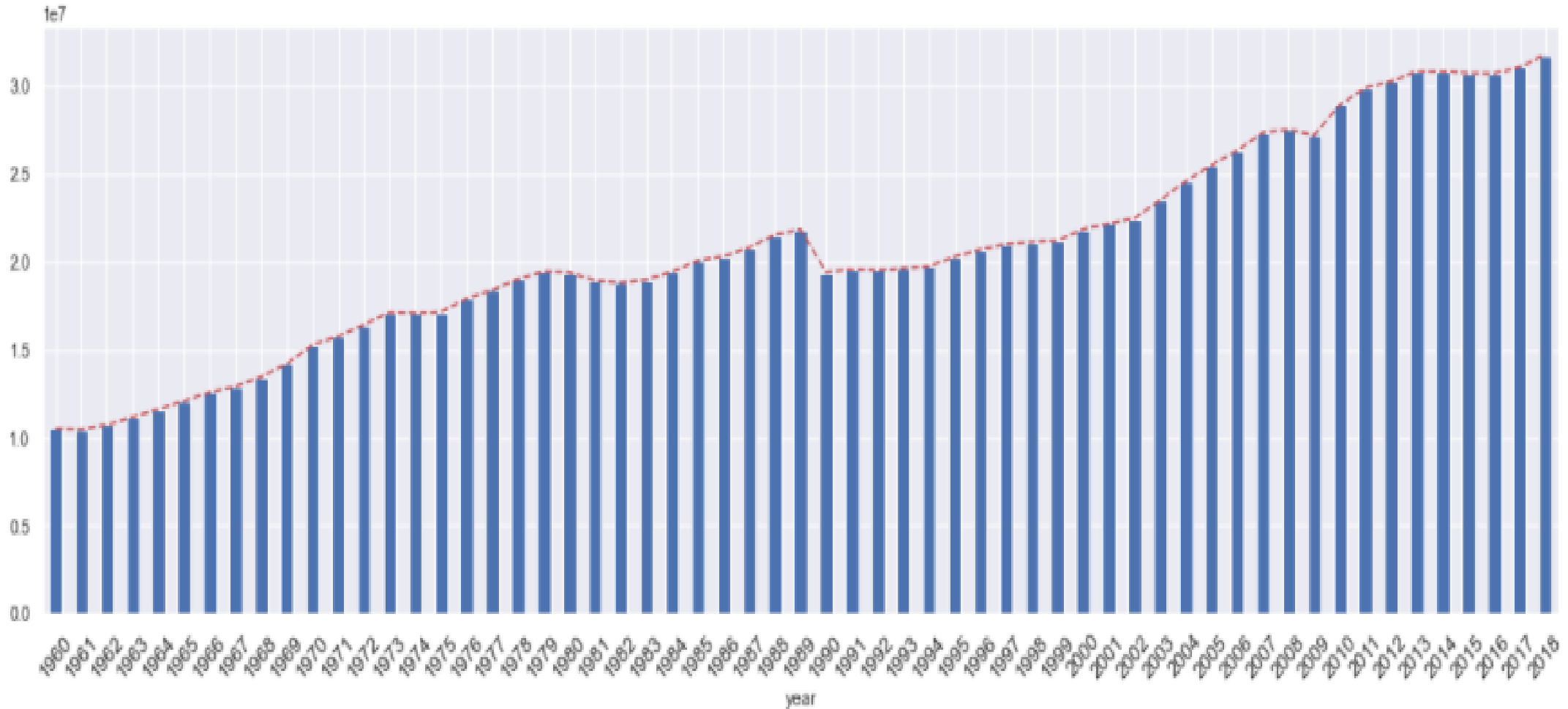


# 지역별 CO2 배출량 합 추이



<표9> 지역별 CO2 배출량 합 추이 (Pandas)

# 세계 CO2 배출량 합 추이



<표10> 세계 CO2 배출량 합 추이 (Pandas)

기간에 따라 증가하는 형태를 보임

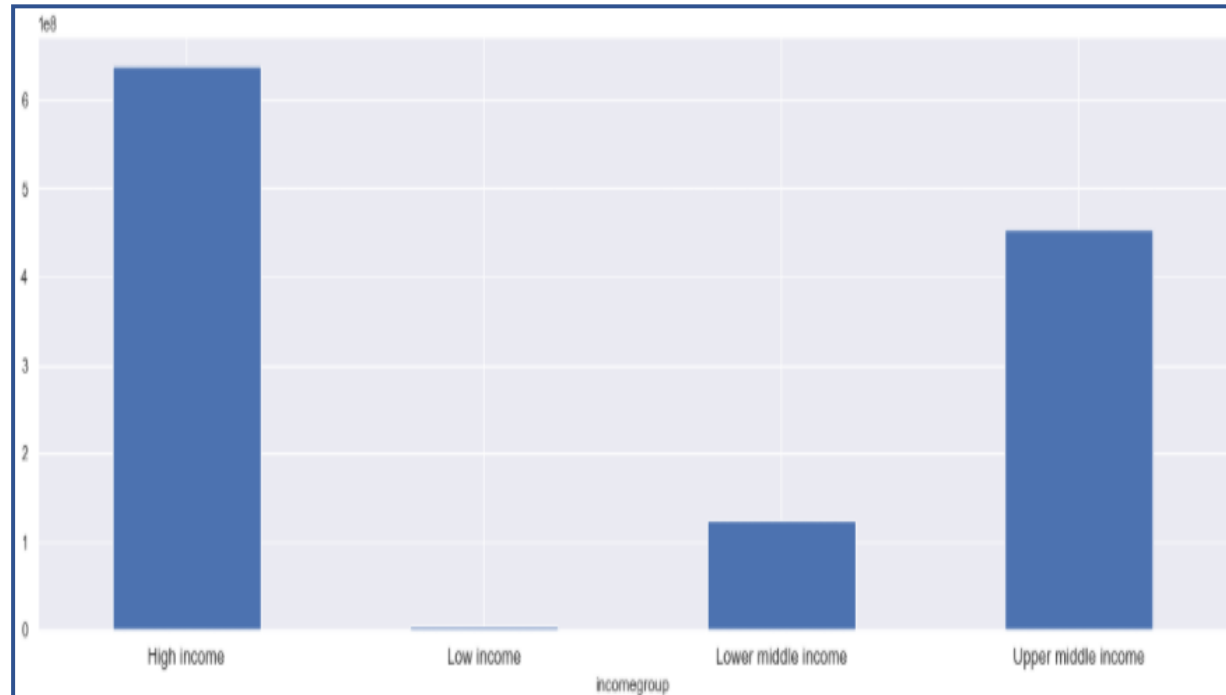
# CO2 데이터 EDA

우리 데이터 변수 중 '소득' 과 CO2 배출량 사이에 어떤 관계가 있을까?

year	country_name	country_code	region	incomegroup	value
1960	China	CHN	East Asia & Pacific	Upper middle income	7.807263e+05
1960	India	IND	South Asia	Lower middle income	1.205820e+05
1960	Japan	JPN	East Asia & Pacific	High income	2.327812e+05
1960	Russian Federation	RUS	Europe & Central Asia	Upper middle income	1.448637e+06
1960	United States	USA	North America	High income	2.890696e+06
...	...	...	...	...	...
2018	China	CHN	East Asia & Pacific	Upper middle income	1.031346e+07
2018	India	IND	South Asia	Lower middle income	2.434520e+06
2018	Japan	JPN	East Asia & Pacific	High income	1.106150e+06
2018	Russian Federation	RUS	Europe & Central Asia	Upper middle income	1.607550e+06
2018	United States	USA	North America	High income	4.981300e+06

295 rows × 5 columns

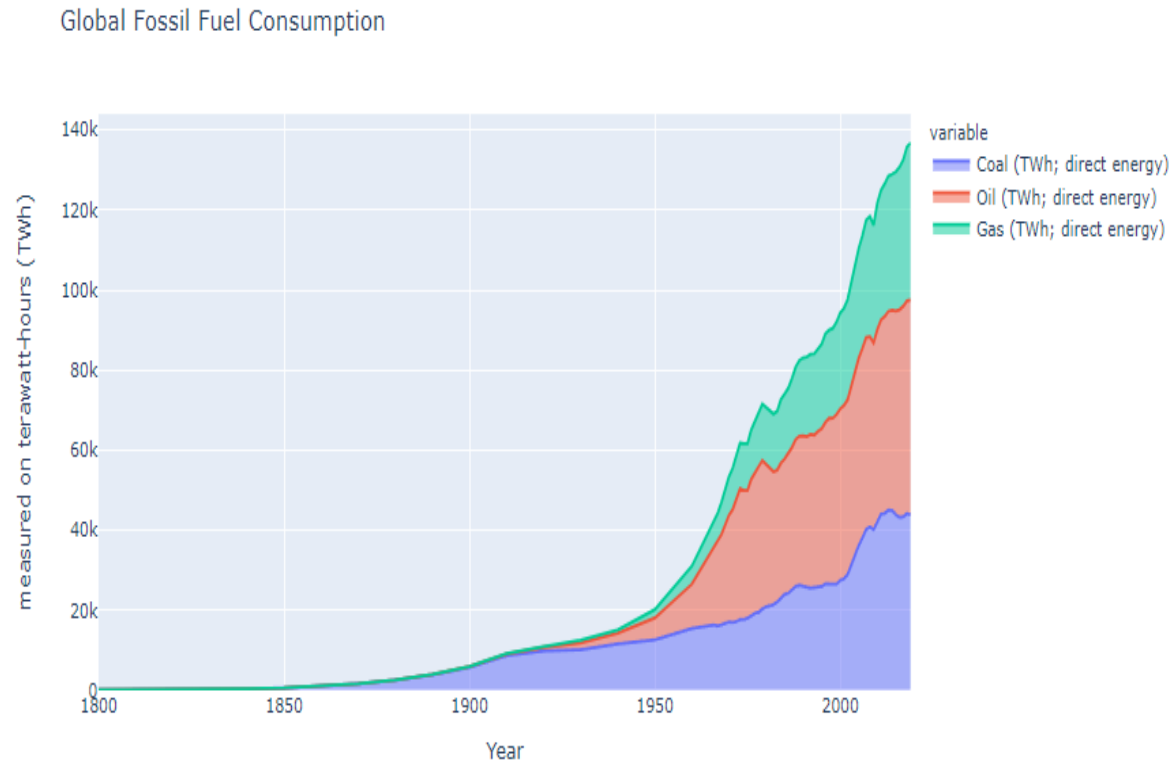
소득이 높은 국가일수록 CO2 배출량이 많음을 알 수 있음



<표11> 국가 소득수준 별 CO2 배출량 (Pandas)

# CO2 데이터 EDA

CO2 배출 주 요인인 화석연료 사용량과 어떤 관계가 있을까?



<표12> 세계 화석연료 사용량 (Plotly)

화석연료 사용량과 CO2 배출량 사이 상관관계가 높음을 알 수 있음

```
corr_co2_fossil=final_co2_fossil[['CO2','coal','oil','gas']].corr(method='pearson')  
corr_co2_fossil
```

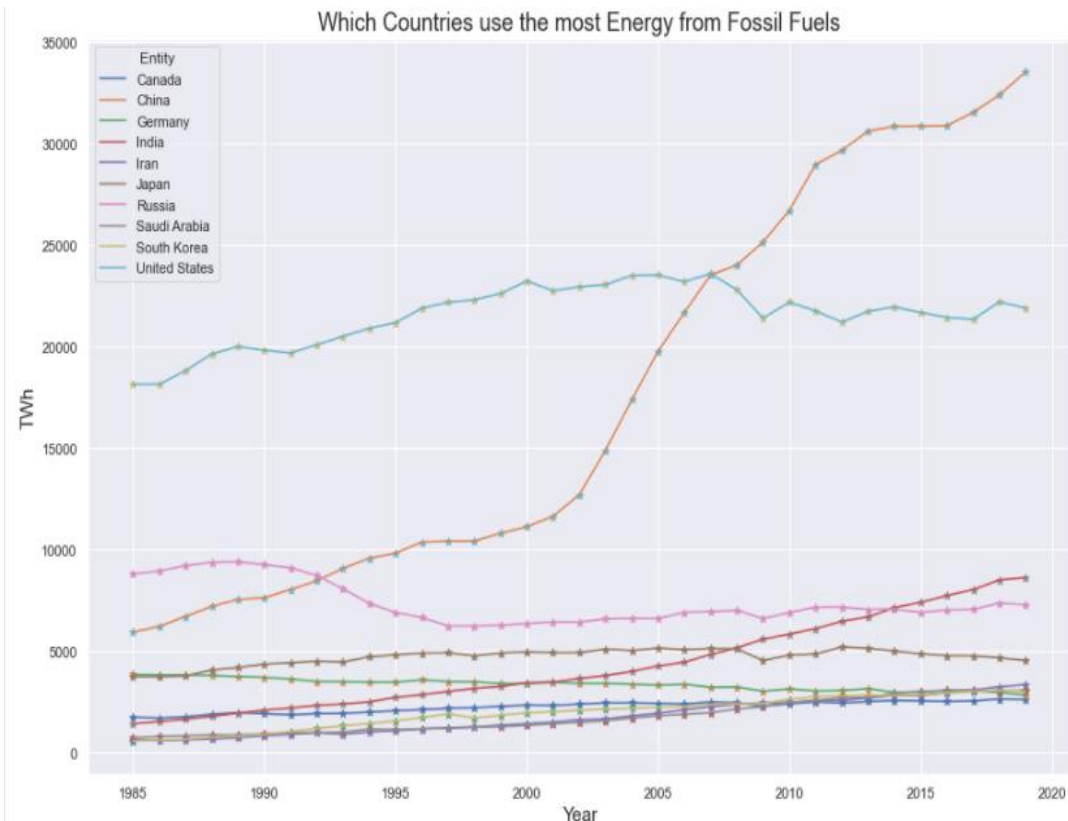
	CO2	coal	oil	gas
1965	11402575	16140	18109	6303
1966	11988321	16324	19496	6869
1967	12413351	16061	20891	7375
1968	13087346	16301	22675	8044
1969	13976005	16799	24577	8833
1970	15272289	17059	26708	9614
1971	15913710	16966	28205	10292
1972	16680140	17154	30378	10861
1973	17582029	17668	32746	11378
1974	17563629	17682	32272	11656
1975	17641960	18025	31948	11661
1976	18580873	18688	34030	12356
1977	19183226	19241	35215	12761
1978	19921141	19458	36426	13294
1979	20459807	20364	37024	14119
1980	20396427	20858	35577	14239
1981	19879068	21150	34296	14394
1982	19815965	21384	33198	14469



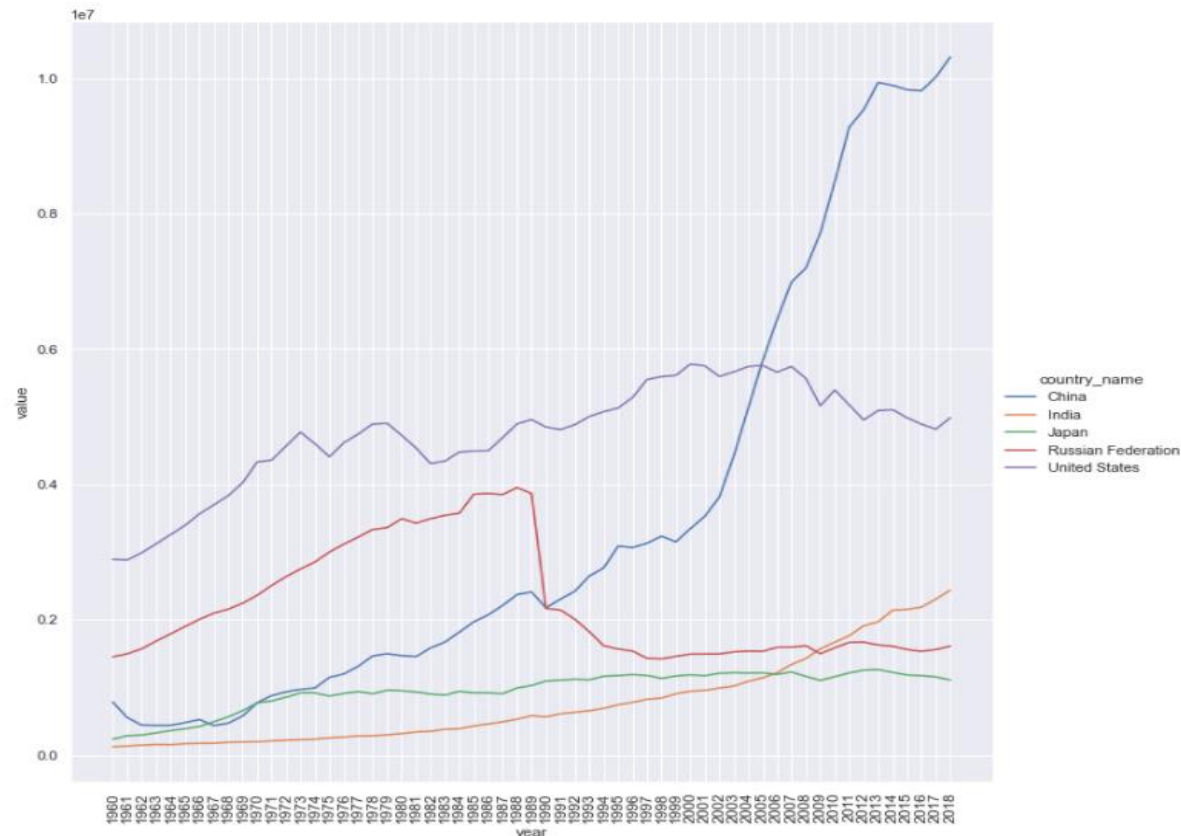
<표13>  
화석연료와 CO2 배출량 상관계수  
(Seaborn)



# CO2 데이터 EDA



<표14> 국가별 화석연료 사용량 (Pandas)



<표15> CO2 배출량 Top5 국가 (Seaborn)

화석연료 사용량 Top5: China, United States, India, Russia, Japan

≈

CO2 배출량 TOP5: China, United States, India, Russia, Japan

화석연료 사용량이 많은 국가일수록 CO2 배출량이 높음



# CH<sub>4</sub> 배출량

# CH4 데이터 설명 및 전처리 요약

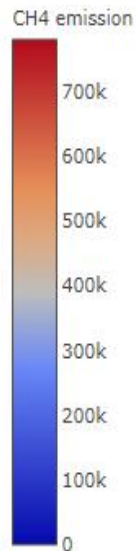
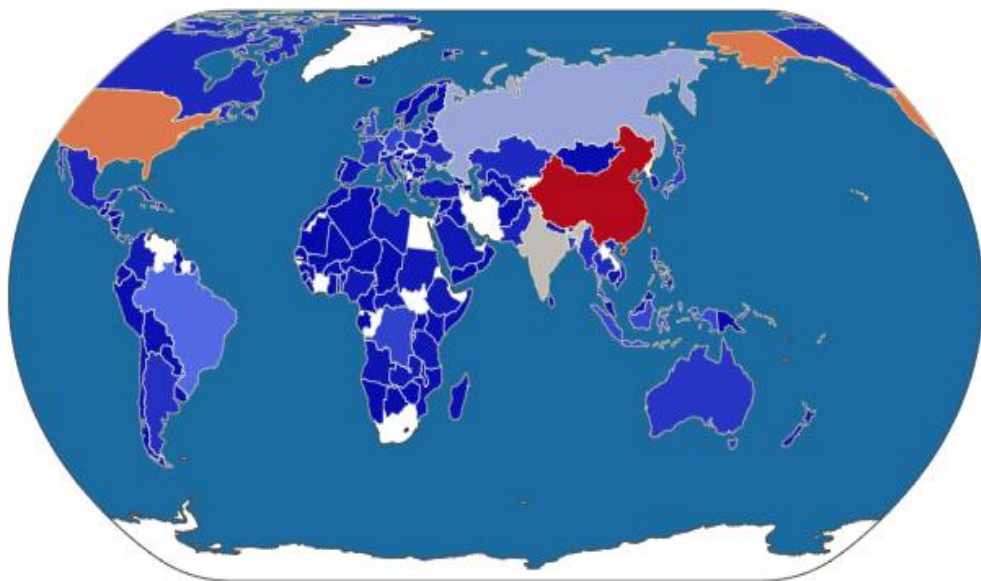
데이터 형태가 동일하므로

CH4 배출량 데이터에 대해서도

CO2 배출량 데이터와 동일한 방식의 전처리

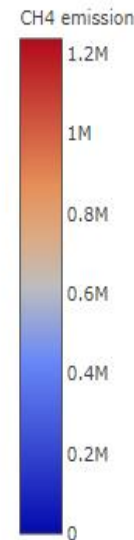
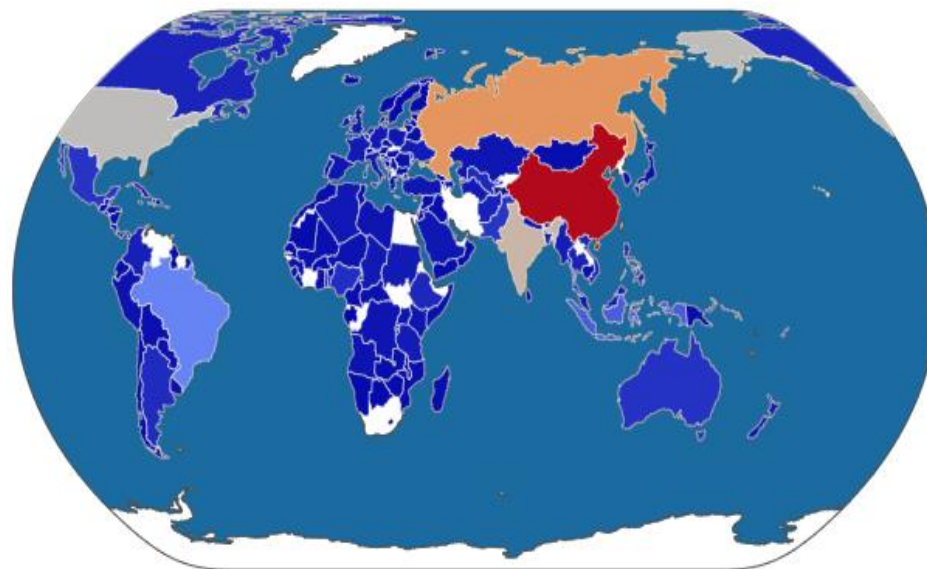
# 세계 CH4 배출량 비교(1970년, 2018년)

World Map of CH4 emission in the Year 1970



<그림4> 1970년 CH4 배출량 지도 (Plotly)

World Map of CH4 emission in the Year 2018



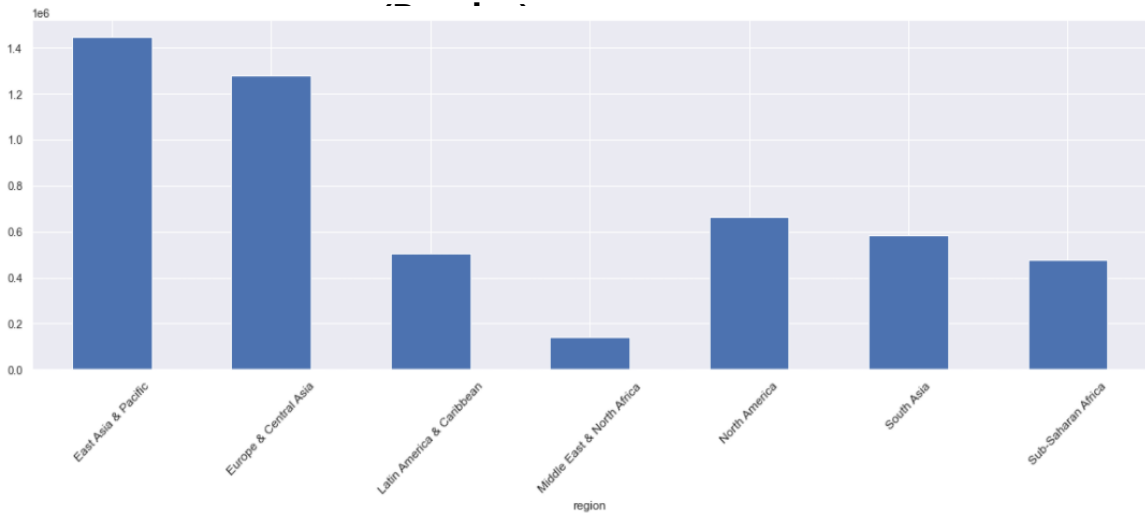
<그림5> 2018년 CH4 배출량 지도 (Plotly)

1970년대에 비해 2018년에 CO2배출량이 확연히 많음을 알 수 있음

# 지역별 CH4배출량 비교 (1970년, 2018년)

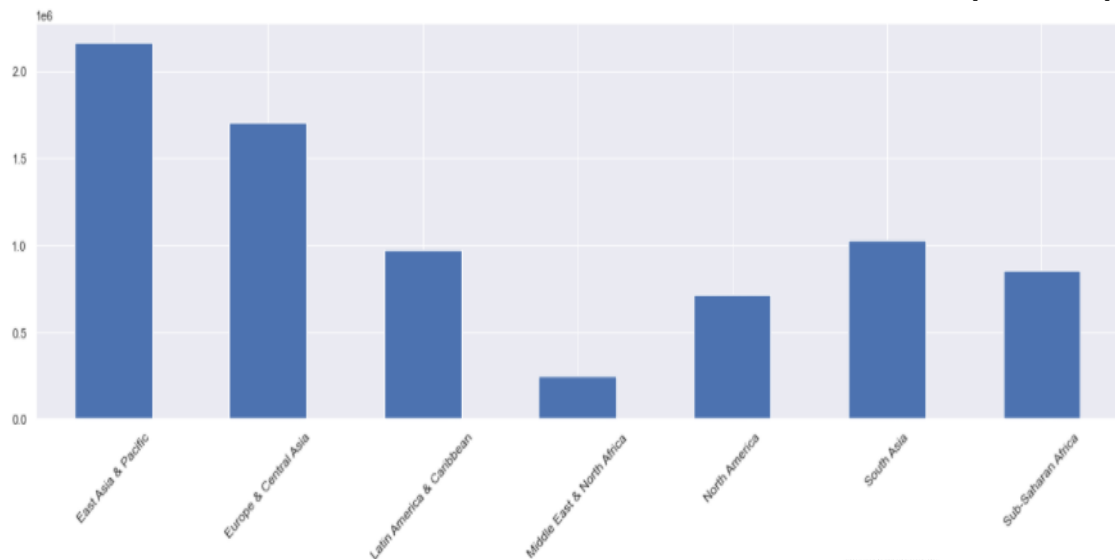
1970년

<표16> 1970년 지역별 CH4 배출량



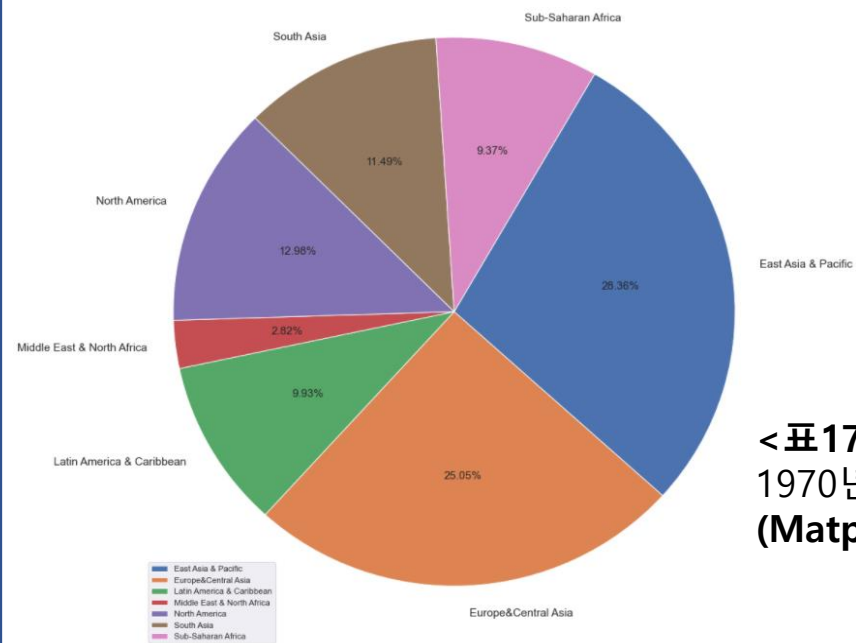
2018년

<표18> 2018년 지역별 CH4 배출량 (Pandas)

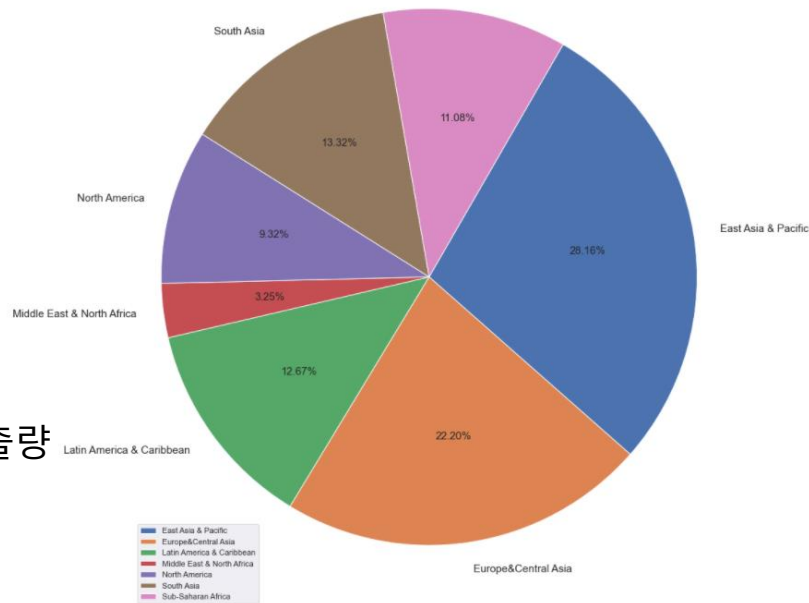


지역별 배출량 비율은  
두 연도 사이에 큰 차이가 없음

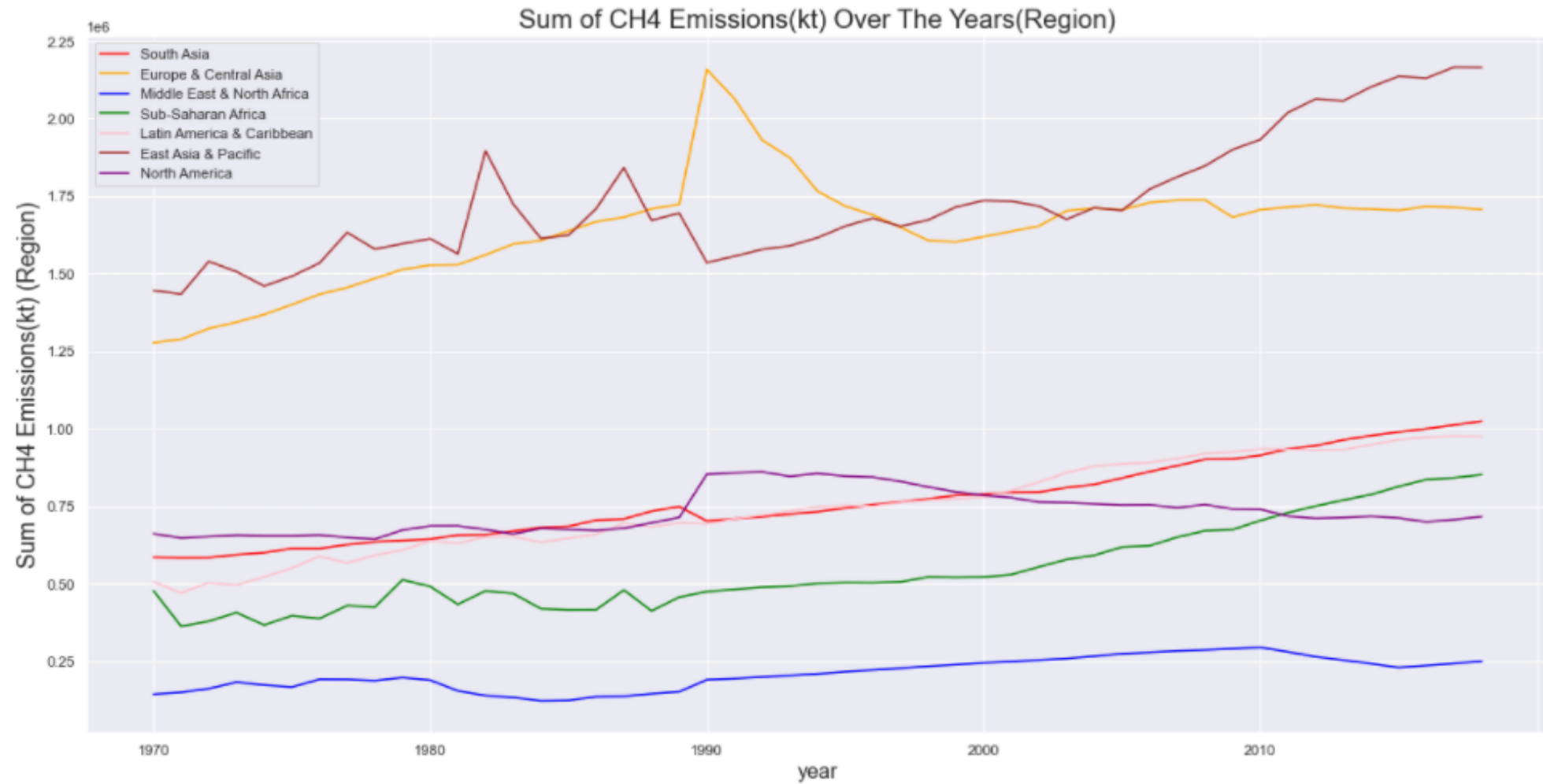
<표17>  
1970년 지역별 CH4 배출량  
(Matplotlib)



<표19>  
2018년 지역별 CH4 배출량  
(Matplotlib)

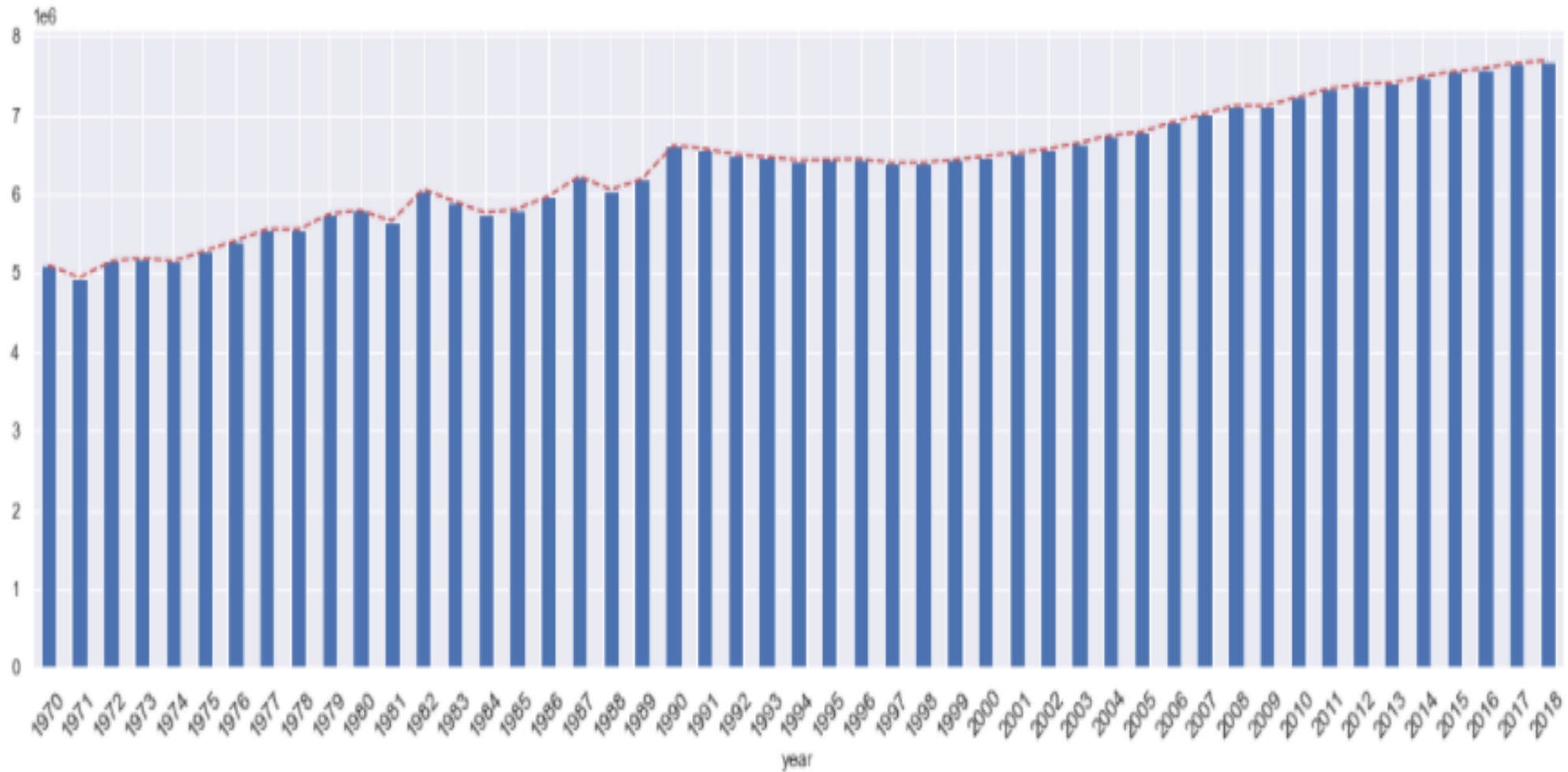


# 지역별 CH4 배출량 합 추이



<표20> 지역별 CH4 배출량 합 추이 (Pandas)

# 세계 CH4 배출량 합 추이



<표21> 세계 CH4 배출량 합 추이 (Pandas)

기간에 따라 증가하는 형태를 보임

# CH4 데이터 EDA

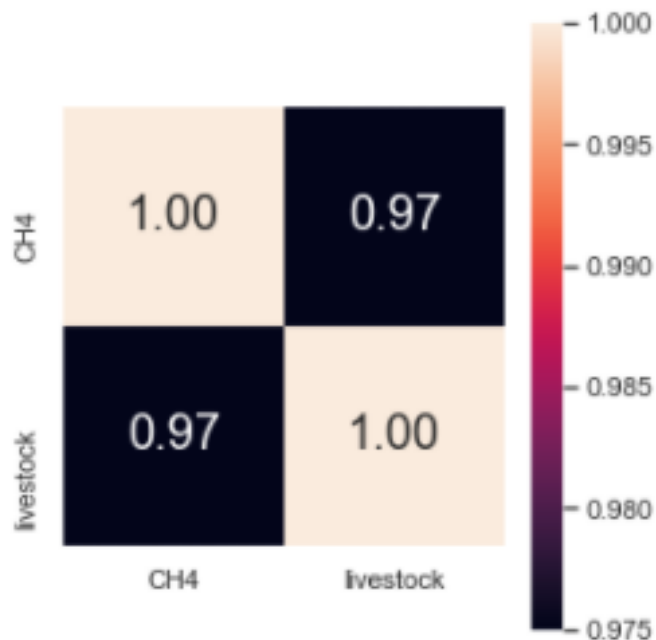
CH4 배출 주 요인인 가축 생산량과 어떤 관계가 있을까?

```
corr_ch4_livestock=final_ch4_livestock[['CH4','livestock']].corr(method='pearson')  
corr_ch4_livestock
```

	CH4	livestock
1980	5997921	45565944
1981	5858004	45953128
1982	6275398	45914058
1983	6126518	47159800
1984	5973200	48483781
1985	6025375	49306824
1986	6184970	50983632
1987	6449419	50940348
1988	6276209	51339867
1989	6416172	51557846
1990	6898360	53028463
1991	6875120	53553615
1992	6810430	52690686

데이터 클리닝 작업

가축 생산량과 CH4 배출량 사이 상관관계가 높음을 알 수 있음



<표22>  
가축 생산량과 CH4 배출량 상관계수  
(Seaborn)

# 재생에너지



# 재생에너지 데이터 설명 및 전처리 요약

- **데이터명:** Renewable Energy Consumption (%)
  - **데이터:** World Bank data
  - 국가별 연도에 따른 전체 에너지 대비 재생에너지 사용 비율

	Country Name	Country Code	Indicator Name	Indicator Code	1960	1961	1962	1963	1964	1965	...	2011	2012
0	Aruba	ABW	Renewable energy consumption (% of total final...	EG.FEC.RNEW.ZS	NaN	NaN	NaN	NaN	NaN	NaN	...	5.661800	6.855900
1	Africa Eastern and Southern	AFE	Renewable energy consumption (% of total final...	EG.FEC.RNEW.ZS	NaN	NaN	NaN	NaN	NaN	NaN	...	62.293842	61.109554
2	Afghanistan	AFG	Renewable energy consumption (% of total final...	EG.FEC.RNEW.ZS	NaN	NaN	NaN	NaN	NaN	NaN	...	11.559100	14.404900

- **데이터명:** Modern Renewable Energy Consumption
  - **데이터:** Our World In Data ( ourworldindata.org )
  - 재생에너지 종류별 세계 생산량

	Entity	Code	Year	Wind Generation - TWh	Solar Generation - TWh	Geo Biomass Other - TWh	Hydro Generation - TWh
4947	World	OWID_WRL	1965	0.000000	0.000000	17.985232	923.197924
4948	World	OWID_WRL	1966	0.000000	0.000000	19.806007	983.817037
4949	World	OWID_WRL	1967	0.000000	0.000000	19.986571	1005.742652
4950	World	OWID_WRL	1968	0.000000	0.000000	22.120045	1059.289008
4951	World	OWID_WRL	1969	0.000000	0.000000	23.257964	1121.743221

## 전처리 과정

1. 분석 시 필요없는 열(컬럼) 삭제
2. 결측치 처리-bfill,ffill(CO2데이터 전처리 시 적용한 bfill,ffill 방법과 동일)
3. 음수값 삭제
4. country\_name 열의 값이 개별국가(country\_list리스트 안의 요소)인 행만 추출

전처리 코드는 CO2데이터 전처리 시 작성한 코드와 동일

## 전처리 과정

1. Entity 열에서 "World" 변수인 행을 모두 추출
2. 재생에너지 종류를 리스트로 저장
3. 연도를 x 축으로 하여 시각화

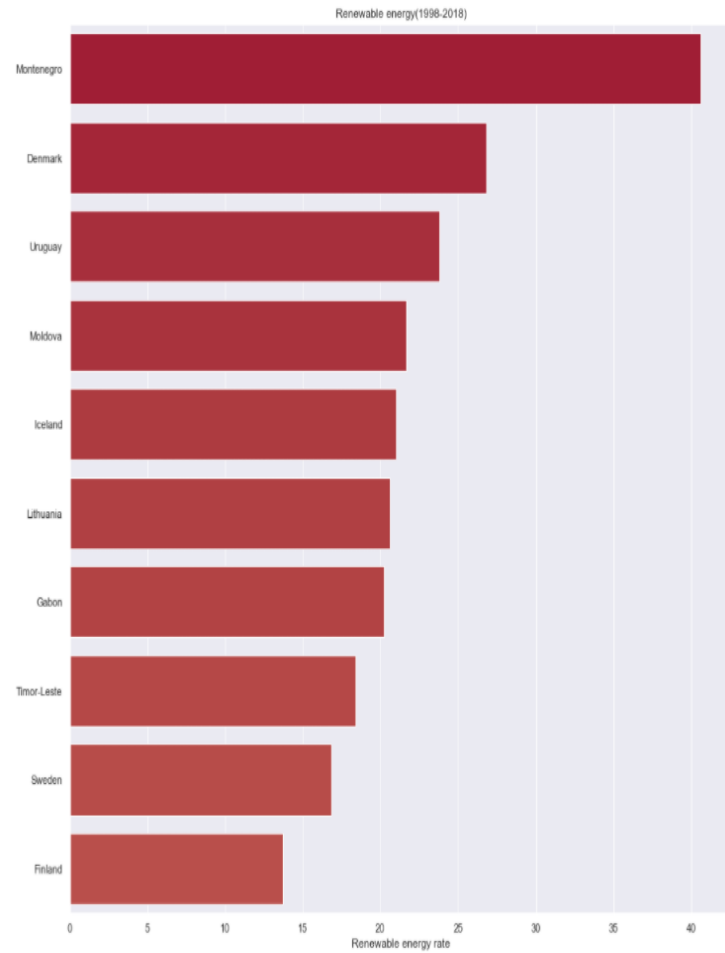
```
global_renewable = pd.read_csv('./dataset/modern-renewable-energy-consumption.csv')
global_renewable = global_renewable[global_renewable['Entity']=='World']
```

```
renewable_y = ['Hydro Generation - TWh', 'Wind Generation - TWh',
               'Solar Generation - TWh', 'Geo Biomass Other - TWh']
```

```
fig = px.area(global_renewable, x="Year", y=renewable_y, title='Renewable Energy Generation, World',
               labels={'value': 'measured on terawatt-hours (TWh)'})
```

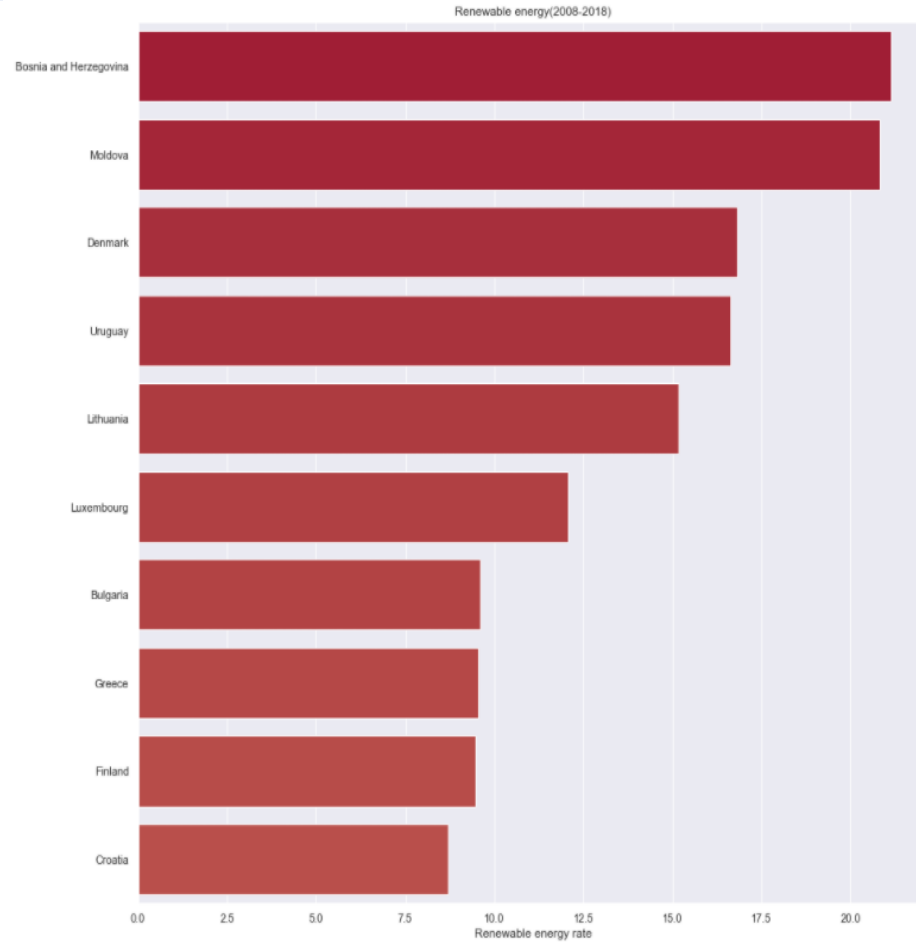
```
fig.show()
```

# 재생에너지 EDA



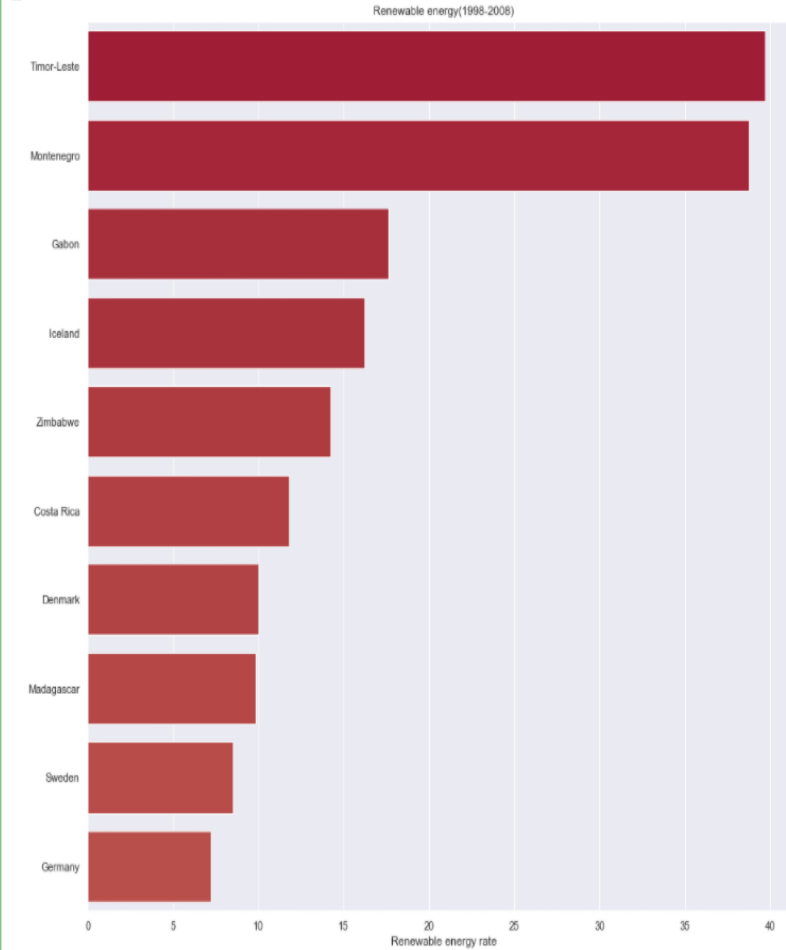
[1998-2018]

1. 몬테네그로 2. 덴마크 3. 우루과이 4. 몰도바 5. 아이슬란드



[2008-2018]

1. 보스니아 2. 몰도바 3. 덴마크 4. 우루과이 5. 리투아니아



[1998-2008]

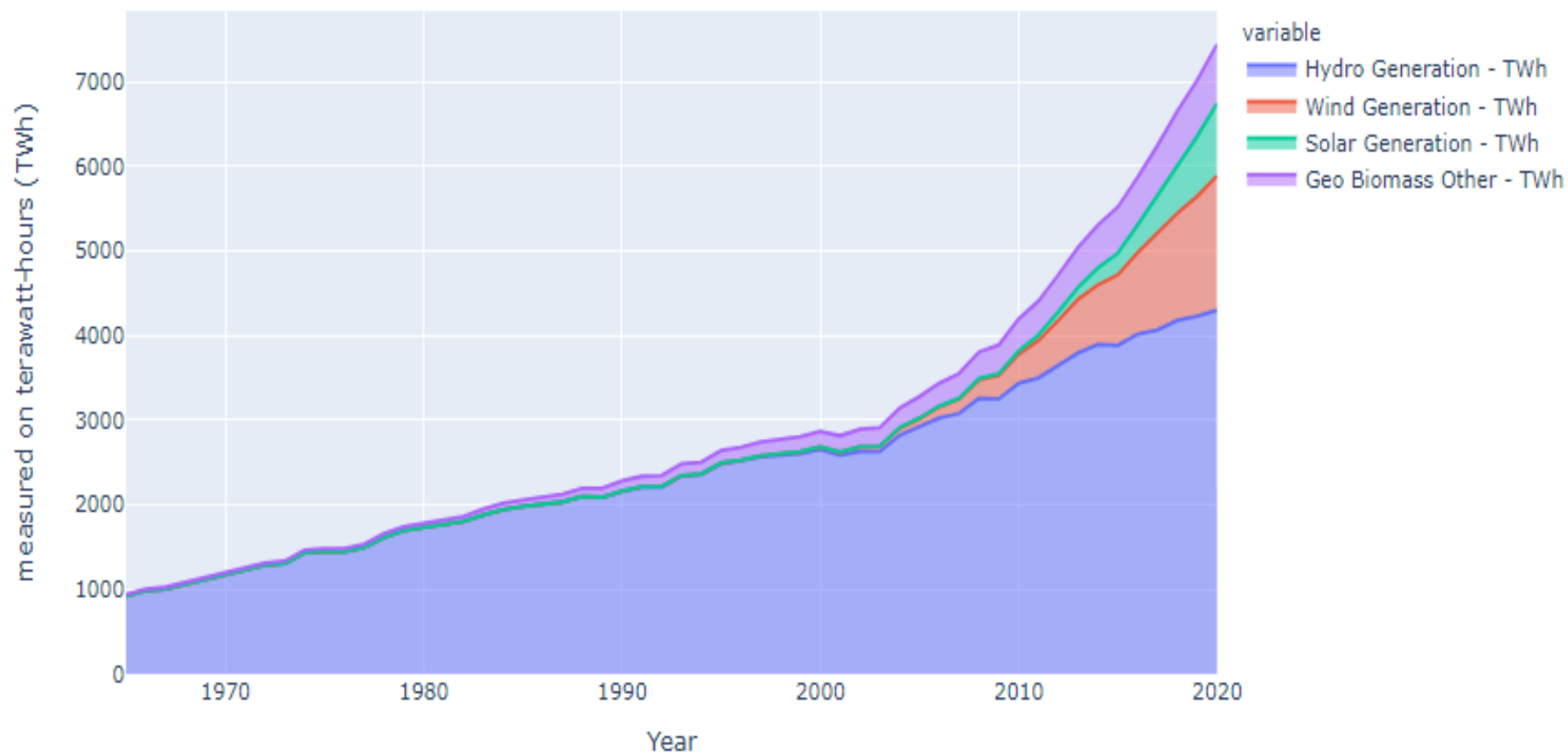
1. 동티모르 2. 몬테네그로 3. 가봉 4. 아이슬란드 5. 짐바브웨

<표23> 기간 차에 따른 재생에너지 사용 비율 증가량 (Matplotlib)

대체로 북유럽, 동유럽 및 다수의 아프리카 국가에서 높은 재생에너지 사용률을 보임  
자연 및 지형적 특성 때문일 것이라 유추

# 재생에너지 EDA

Renewable Energy Generation, World



<표24> 세계 재생에너지 생산량 (Plotly)

세계적으로 수력에너지 생산량이 가장 많고  
풍력 생산량이 급격히 증가하고 있음을 알 수 있다.

# 한계점 및 보완사항

# 한계점 및 보완할 부분



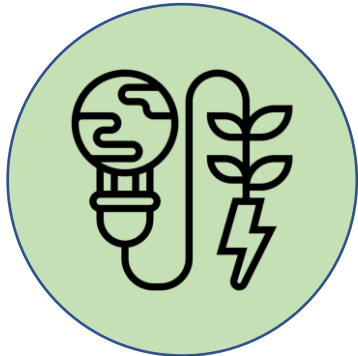
## 결측치 처리 문제

CO<sub>2</sub>, CH<sub>4</sub>에 대해 **bfill**, **ffill**(각 국가의 지난 해/ 다음 해의 데이터)로 결측치 처리

### 이유

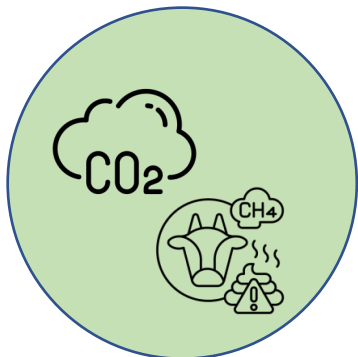
- 상당 수의 국가에서 연도별 CO<sub>2</sub>/CH<sub>4</sub>배출량 변화가 불규칙적 & 기하급수적 변화 (평균치 사용하기 모호함)
- 결측치 수가 많아 모두 제거할 경우 문제가 발생

⚠ 결측치 처리방법에 대해 더 바람직한 고민이 필요



## 재생에너지 사용비율 증가

- World Bank 데이터는 전체 에너지 사용량 대비 재생에너지 사용 비율을 나타냄
- 기간 사이 사용 비율 차분값으로는 실질적인 사용량이 얼마나 증가했는지 제대로 파악이 어려움



## CO<sub>2</sub>, CH<sub>4</sub> 배출량에 대해 상관관계만 분석

- Graner Test등의 인과관계 모델도 진행해볼 수 있음
- 더 나아가 변수를 추가하여 회귀 분석도 진행해볼 수 있음

감사합니다.