실증분석연구회 2주차(20200112)

***지구온난화의 원인에 목축업/낙농업?***

# **서론**

## **자료 조사**

## 몇 년째, 심각해지고 있는 지구온난화. 과연 원인이 무엇인가?

## 기사: ['화석 연료, 사막화, 화학 비료'… 지구 온난화의 5가지 원인] 지구 온난화의 핵심 원인 중 하나로 목축업/낙농업으로 인한 화학 비료 사용 및 메탄 가스 배출이라고 지적하고 있다.

## 출처: <https://news.joins.com/article/19454330>

## **참고 연구 방향**

## 아래의 주소의 연구자의 지구온난화 연구 방법을 보면 알 수 있다시피, ‘지구 공전 궤도’, ‘태양의 온도’, ‘화산 활동’, ‘미세먼지’ , ‘산림벌채’, ‘오존’ 등은 지구 온난화와 큰 상관관계를 가지 않는다. 그의 예로, 그래프가 같은 방향으로 진행하고 있지 않다. 출처: 출처: <https://brunch.co.kr/@thyoon/1>

## 하지만, ‘온실 가스’의 경우 그래프가 지구온난화의 지표와 동일한 방향으로 진행되고 있음을 알 수 있다. 이와 같이, 이번 발제에서도 유제품 생산량을 통해서 지구 온도변화와 얼마나 유사한 추이를 갖는지 알아보자!

# **연구 계획**

## **주제**

지구온난화의 원인으로 주목되고 있는 목축업/낙농업이 얼마나 상관관계를 갖고 있는지 알아보자!

## **연구 방법**

### 유제품 생산량, 온실가스 배출량, 지구온도 변화 데이터 수집

### 각 데이터의 추이를 그래프로 표현

### 서로 어떤 상관관계를 갖는지 분석. 얼마나 비슷한 추이를 보이는지 비교! Ols 회귀분석을 이용해보자! Coefficient가 양수라면 서로 영향력을 갖는다는 것을 알 수 있겠다.

### Ex) 유제품 생산량이 증가하면, 온실 가스 배출량도 증가하는지, 지구 온도는 올라갔는지

## **연구 데이터**

### 유제품 생산량

### 온실가스 배출량

### 지구 평균 온도 추이

## **연구의 한계점**

## 다른 요소를 제외하지 않고 회귀 분석을 할 경우 왜곡될 수 있지 않을까? 통제변수?를 넣어야 하지 않을까 생각이 든다. 하지만, 아직 통제변수 다루는 법을 모르니… 일단 해보자! 결론적으로는, 목축업/낙농업이 지구온난화의 기여하는 부분이 있을 수 있지만, 그 정도를 왜곡할 수도 있다!

# **본론**

## **데이터 그래프로 표현**

### 유제품 생산량

### 

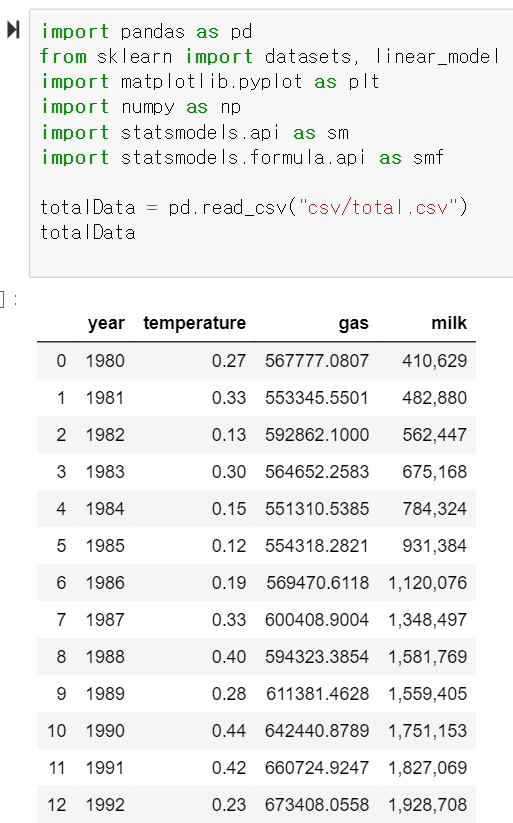
### 온실가스 배출량

### 지구 평균 온도

### 

## OLS 회귀분석

### 먼저, 각 데이터를 하나의 파일로 합친 후, 표로 나타낸다.



### OLS회귀 분석을 해준다.

### 회귀모형의 F값이 84.06 유의확률(Pr>F)은 0으로 통계적으로 유의한 모형이라 할 수 있다.

### R-squared가 0.75 즉, 75%의 설명력을 갖는다. 회귀모형의 유용성은 높은 편이라는 것을 알 수 있다.

### “[2] The condition number is large, 4.89e+06. This might indicate that there are strong multicollinearity or other numerical problems”를 보아하니, 다중공산성의 문제가 있는 것 같다. 다중공산성이란? 일부 독립변수(예측변수, 회귀자; regressor)가 다른 독립변수와 높은 상관관계를 가질 때 또는 그럴 때 발생하는 문제를 의미한다. 해결법? <https://m.blog.naver.com/PostView.nhn?blogId=cto_hwangga&logNo=220969542325&proxyReferer=https%3A%2F%2Fwww.google.com%2F> 여기를 참고해보자!

### OLS 회귀 분석 – milk

### 

* 이 표는 말썽쟁이 milk의 데이터를 좀 수정해서, 1990~2009년까지의 데이터로만 회귀분석을 돌려본 것인데… 보다시피 이상하게 나와요 ㅠㅠ 왜그런지는 저도 모르겠슴다….

# **결론**

## 그래프를 그려보아서 비교해보자!

## Milk graph를 python으로 그리는 것은 실패했지만! Greenhouse gas와 temperature은 그래프끼리 서로 비슷한 양상을 띄고 있음을 알 수 있다.

## 회귀식

## = -0.6103 + 0.00000149\*greenhousegas

## 이게 얼마나 유의미한가…는 솔직히 잘 모르겠음temperature자체가 30도 이런걸 나타나는게 아니라 작년 대비 얼마나 증가했는가 0.3증가함! 이걸 나타내는 거라서 greenhousegas의 베타값이 0.000000149인게 과연 큰 값인가?!

## 원래는 milk생산량에 대한 베타값을 구하고 싶었는데… jupyter notebook을 보면 알다시피 milk의 데이터만…이상하다!! 흑흑 ㅠㅠ 이상해요 왜 얘만 안되는지 모르겠네요 ㅠㅠ 우리 모두 다같이! 문제점을… 찾아보길 바라요...