











경기대학교 회귀분석2 응용통계학과

나옥경 교수님

2조 201713110 오선영 201913015 이동훈 201913020 채윤아



01 서론

분석배경 및 목적

02 **E**E

1)문제에 대한 진술

2)잠재적으로 적절한 변수들의 선택+ 데이터 수집

3)모형설정(선형회귀모형)

03 본론

4)적합 방법 선택 및 모수 추정

5) 회귀진단-> 모형의 비판과 선택

6) 최종 모형선택 -> 예측 또는 문제해결

04 결론



분석배경 | 우리나라현황 | 주제선정





지구온난화의 가속화

온실가스를 배출할 수 있는 권리는 유한한 자원

전세계적으로 온실가스 배출량 규제 실행 탄소배출을 줄이는데 효과적인 방법은?



향후 인류가 배출가능한 온실가스 총량은 8~9년의 여유분만 남음 탄소를 줄이는데 초점을 맞추는데 있어 탄소배출을 줄이는데 효과적인 방법을 알아보기 위해 이번 분석을 실행하였다.



우리나라 현황

우리나라는 탄소배출을 줄이기 위해 탄소포인트제를 활용하고 있다. 탄소포인트제: 탄소배출 감축율에 따라 탄소포인트를 부여하고 추후에 인센티브지급





탄소포인트제 운영 체계



전기,상수도,도시가스 사용량 절약 >









주제선정

주제: 어느 분야가 탄소배출감소에 유의미한 영향을 미치는지 알아보고자 한다.

01 탄소포인트제

탄소배출을 효과적으로 줄이기 위해서 우리나라가 시행하고 있는 정책은 탄소포인트제 이다.

02 탄소포인트제 장려

탄소포인트제를 통해 탄소배출을 효과적으로 줄이려면 탄소배출을 많이 하고 있는 분야에 적극적으로 장려해야 할 것이다.



데이터 전처리 | 잠재적으로 선택한 변수들

회귀분석 절차



- 01 문제에 대한 진술
- 02 잠재적으로 적절한 변수들의 선택+ 데이터 수집
- 03 모형설정(선형회귀모형)
- 04 적합 방법 선택 및 모수 추정 -> 모형 적합(최소제곱법 or능형회귀추정)
- 05 회귀진단-> 모형의 비판과 선택
- 06 최종 모형선택 -> 예측 또는 문제해결



자료수집및출처

출처	자료원 명	연도	산출지표
국가온실가스 종합관리시스템	공공부문 배출량 통계	2018	기준배출량, 온실가스 배출량, 외부사업 감축량, 탄소포인트 감축량, 온실가스 감축량, 온실 가스 감축률
한국환경공단	한국환경공단_탄소포인트제 전기사용량 _20200910	2018	전기사용량
한국환경공단	한국환경공단_탄소포인트제 도시가스 사용량_20200910	2018	지역난방사용량
한국환경공단	한국환경공단_탄소포인트제 상수도 사 용량_20200910	2018	상수도사용량
한국환경공단	한국환경공단_탄소포인트제_전국 참여 자 현황_20200831	2018	지자체별 탄소포인트제 가입 세대 수, 가입자수

데이터 전처리- 결측치 처리

```
결측치 확인
모든 데이터를 합친 tot3에서
공공부문 배출량 통계의 기준배출량과
한국환경공단_탄소포인트제 상수도 사용량의 상수도사용량에서
결측치가 있음을 확인했다.
> table(is.na(tot3$상수도.사용량))
FALSE TRUE
  182 1
> table(is.na(tot3$기준배출량.tonCO..eq.))
FALSE
     TRUE
  175
         8
```

데이터 전처리- 결측치 처리

결측치 제거

filter를 이용하여 결측치를 제거한다.

```
> tot3 = tot3 %>% filter(!is.na(tot3$기준배출량.tonCO..eq.))
> tot3 = tot3 %>% filter(!is.na(tot3$상수도.사용량))
> table(is.na(tot3))
FALSE
```

2088

추출 변수 설명

한국환경공단_탄소포인트제_전국 참여자 현황에서는 개별단독, 개별아파트(개인), 개별아파트(상업), 상업법인, 공공기관, 학교,아파트(단지) 가입자 수 등 여러 변수들이 있습니다.

하지만, 이 변수들을 대표적으로 설명할 수 있는 변수로 합계(가입자수)를 추출하였습니다.

한국환경공단_탄소포인트제 전기, 도시가스, 상수도 사용량 데이터에서는 2018년, 지자체 변수들을 기준으로 각각 변수들을 추출하였습니다.

공공부문 배출량 통계에서는 기관명이라는 변수를 다른 데이터의 변수 이름인 지자체라는 변수로 변경 하여 기준배출량, 온실가스 배출량, 탄소포인트 감축량, 온실가스 감축량, 온실가스 감축률이라는 변수 들을 추출하였습니다.

여기서 외부사업 감축량이라는 변수를 제거하였습니다. 이유는 변수에 해당하는 데이터값이 0인 값들이 많아 회귀분석에 영향을 미칠 수 있다고 판단하였습니다.

탄소포인트제 이용 R 코드

```
data5 = read.csv("C:/Users/user/Desktop/공공부문 배출량 통계_20221122205157440.csv",header=T)
library(dplyr)
                                                                                                    summary(data5)
data1 = read.csv("C:/Users/user/Desktop/한국환경공단_탄소포인트제_전국 참여자 현황_20200831.csv",header=T)
                                                                                                    str(data5)
str(data1)
                                                                                                    data5[,3]
summary(data1)
                                                                                                    colnames(data5)[c(3)] = c('지자체')
data1 = data1[,-c(2:12)]
data1
                                                                                                    data5 = data5[,-c(1,2,10,11)]
|colnames(data1)[c(1)] = c('지자체')
                                                                                                    str(data5)
str(data1)
                                                                                                    tot = inner_join(data1,data2, by="지자체")
data2 = read.csv("C:/Users/user/Desktop/한국환경공단_탄소포인트제 전기 사용량_20200910.csv",header=T)
                                                                                                    tot1 = inner_join(tot,data3,by="지자체")
data2
                                                                                                    tot2 = left_join(tot1,data4,by="지자체")
data2 = data2 %>% filter(연도 == 2018)
data2 = data2[,-1]
                                                                                                    tot3 = left_join(tot2,data5, by="지자체")
str(data2)
                                                                                                    tot3
                                                                                                    summary(tot3)
data3 = read.csv("C:/Users/user/Desktop/한국환경공단_탄소포인트제 도시가스 사용량_20200910.csv",header=T)
summary(data3)
data3 = data3 %>% filter(연도 == 2018)
                                                                                                    tot3 = tot3 %>% filter(!is.na(tot3$기준배출량.tonCO..eq.))
data3 = data3[,-1]
                                                                                                    tot3 = tot3 %>% filter(!is.na(tot3$상수도.사용량))
str(data3)
                                                                                                    summary(tot3)
                                                                                                    Inrow(tot3)
data4 = read.csv("C:/Users/user/Desktop/한국환경공단_탄소포인트제 상수도 사용량_20200910.csv",header=T)
                                                                                                    str(tot3)
str(data4)
data4 = data4 %>% filter(연도 == 2018)
data4 = data4[,-1]
                                                                                                    tot4 = tot3[,-c(1,3,9)]
str(data4)
                                                                                                    str(tot4)
```

잠재적으로 선택한 변수들

탄소배출 감소에 유의미한 영향을 미치는지 알아보기 위해 탄소포인트제를 이용하고 있는 분야들을 살펴보고자 한다. 따라서 잠재적으로 선택한 변수는 다음과 같다.

```
> str(tot4)
'data.frame': 174 obs. of 9 variables:
$ 합계..가입자수.
                          : int 5411 1265 1935 1316 1579 8706 4007 779 1682 2126 ...
$ 전기.사용량
                           : num 40959267 16398978 6869821 5535720 7550114 ...
$ 도시가스.사용량
                          : num 57953 6829 33105 32148 4188 ...
$ 상수도.사용량
                           : num 564605 145132 214415 153044 33632 ...
$ 기준배출량.tonCO..eq.
                           : int 4944 4751 10365 3451 5338 14953 3191 8172 5728 12836 ...
$ 온실가스.배출량.tonCO..eq. : int 5313 4000 9285 4136 4962 12910 2435 7225 6122 9595 ...
$ 탄소포인트.감축량.tonCO..eq.: int 494 21 972 0 205 1495 319 442 248 0 ...
$ 온실가스.감축량.tonCO..eq. : int 125 772 2052 -685 581 3538 1075 1389 426 3241 ...
$ 온실가스.감축률...
                          : num 2.53 16.25 19.8 -19.85 10.88 ...
```

탄소포인트제 이용 R 회귀분석 결과

```
> summary(reg)
Call:
lm(formula = 탄소포인트.감축량.tonCO..eq. ~ ., data = tot4)
Residuals:
   Min
           10 Median 30
                               Max
-666.09 -62.21 31.33 92.97 309.49
Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
(Intercept)
                       3.134e+01 5.288e+01 0.593 0.554
합계..가입자수.
                      1.224e-03 3.473e-03 0.352
                                                  0.725
전기.사용량
                     -4.189e-08 3.888e-07 -0.108
                                                  0.914
도시가스.사용량 -7.905e-06 5.986e-06 -1.321
                                                  0.188
               1.764e-05 1.684e-05 1.047
상수도.사용량
                                                  0.297
기준배출량.tonCO..eq. -6.102e-01 4.367e-02 -13.974
                                                 <2e-16 ***
온실가스.배출량.tonCO..eq. 6.202e-01 4.185e-02 14.820
                                                 <2e-16 ***
온실가스.감축량.tonCO..eq. 6.607e-01 4.289e-02 15.404
                                                 <2e-16 ***
온실가스.감축률... -2.900e+00 2.017e+00 -1.438
                                                  0.152
Signif. codes: 0 (***, 0.001 (**, 0.01 (*, 0.05 (., 0.1 (, 1
Residual standard error: 158.7 on 165 degrees of freedom
Multiple R-squared: 0.8797, Adjusted R-squared: 0.8739
F-statistic: 150.8 on 8 and 165 DF, p-value: < 2.2e-16
```



발표를들어주셔서 스 감사합니다



탄소배출을 효과적으로 줄이는 방법