KIM IN SU

HAN SEO JIN

YOON JU HO

LEE SEUNG WOO

INTRO

1.

Why?
Literary Review,
Initial Setup

2.

DATA ANALYSIS

Methods,
Result S,
Discussions

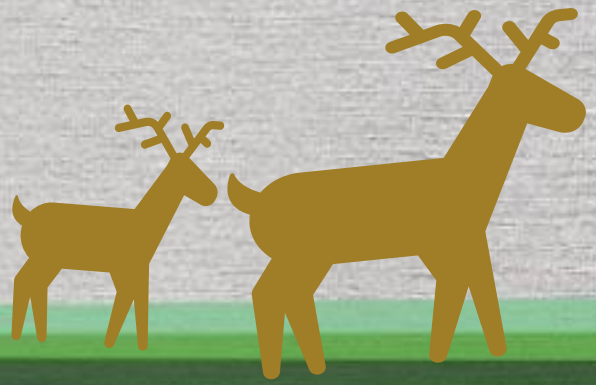
CONCLUSION

3.

Findings ,
Limitations,
Future Plan

1. INTRO

- Why Fine Dust?
- Literary Review
- Initial Setup



FINE DUST , A RISING ISSUE

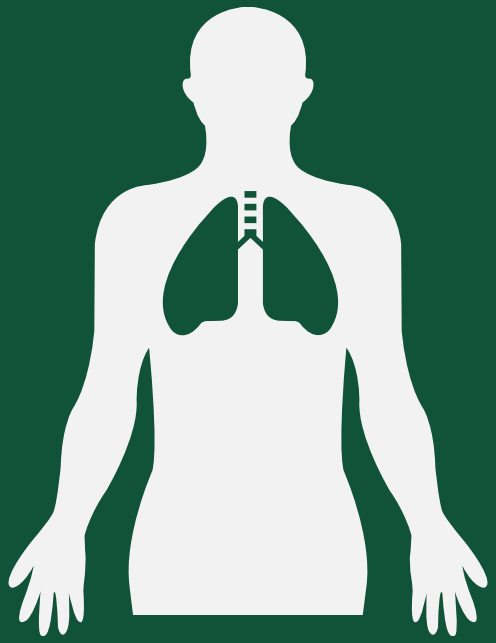
At Least 1 in 6 dies early From Fine dust.

(The Korean Herald,2015.4.20)

Heavy

SMog

in Seoul



Fine dust come to

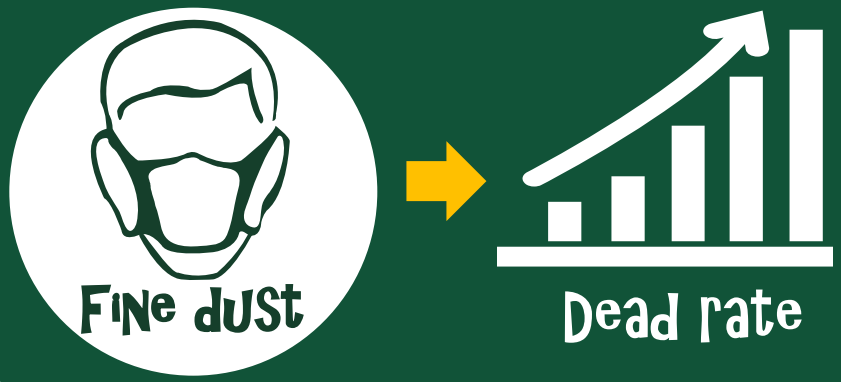
HUMAN body

AS deep as to LUNG

LITERARY REVIEW

Previous Study case 1
(In 1996, Tong and his Colleagues)

Previous Study case 2
(In 2015, professors from Inha University and Ajou University)



Our Data resource

KOSIS 국가통계포털
KOrea Statistical Information Service

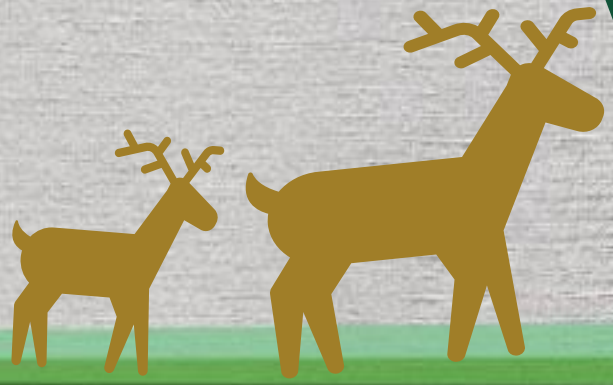
Main set of data
- $\mu g/m^3$ of fine dust level in Seoul (from 2005 to 2013)

Other Factor data set also in Seoul (from 2005 to 2013)


POTENTIAL
CAUSES
OF
FINE
DUST
Registered Cars
Bike roads
Subway Users
GRDP
Constructions
Nuclear power
Coal power
Green belt
Street trees

2. DATA ANALYSIS

- Methods
- Results
- Discussions



PROCESS OF ANALYZING

Setting Variables → Setting Hypothesis → Subset Selection Function → Regression Test (Six Factors) → Regression Test (Individually) → Results & Discussions

SETTING VARIABLES

- Dependent Variable:
  Level of Fine Dusts in Seoul

- Independent Variables:
  Set of Potential Factors (X1, X2, X3, ... , X10)

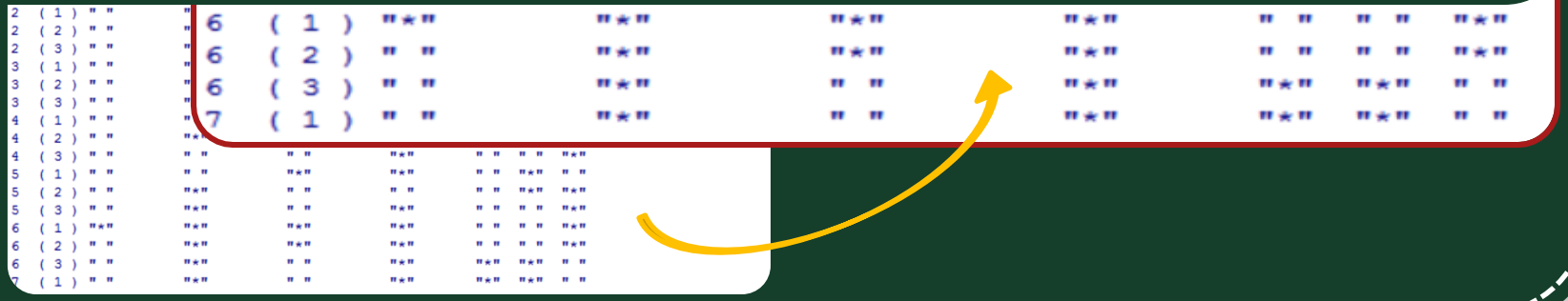
SETTING HYPOTHESIS

-Null Hypothesis:
"We expect to see no correlation in any of the factors to the fine dust level in Seoul from 2005 to 2013."

-Alternate Hypothesis:
"We expect to see a correlation in any of the factors to the fine dust level in Seoul from 2005 to 2013."

SUBSET SELECTION FUNCTION

- Using R's subset function to select much more effective factors.
- After using this function, We decided to select 6 factors.
- 6 factors : registration of car, length of bike road, number of bike road, number of subway users, area of green belt, amount of nuclear energy.



REGRESSION TEST(SIX FACTORS)

```
> summary(fit)

Call:
lm(formula = data$finedust ~ data$car_regist + data$bikeroad_lg +
    data$bikeroad_nu + data$sub_users + data$greenbelt + data$nuclearenergy)

Residuals:
        1         2         3         4         5
-0.011439  0.049317 -0.066126  0.128692 -0.141158  0.1646
        7         8         9
-0.145659  0.019987  0.001691

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)         1.108e+02  1.260e+01   8.798 0.012675
data$car_regist     2.363e+02  4.399e+01   5.372 0.032944
data$bikeroad_lg    2.096e+05  1.713e+04  12.237 0.006612
data$bikeroad_nu   -3.115e+05  4.499e+04  -6.923 0.020235
data$sub_users     -2.904e-01  8.994e-03 -32.281 0.000958
data$greenbelt     -1.196e+01  8.051e-01 -14.860 0.004498 **
data$nuclearenergy -3.178e+03  2.312e+02 -13.743 0.005253 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2147 on 2 degrees of freedom
Multiple R-squared:  0.9997,    Adjusted R-squared:  0.9989
F-statistic:  1170 on 6 and 2 DF,  p-value: 0.0008543
```

Top3 Factors:
SubWay Users, Area of Green Belt, Registered Cars.

Setting Variables → Setting Hypothesis → Subset Selection Function → Regression Test (Six Factors) → Regression Test (Individually) → Results & Discussions

REGRESSION TEST(INDIVIDUALLY_TOP3)

Green Belt (0.6795)

```
> summary(lm(data$finedust~data$greenbelt))

Call:
lm(formula = data$finedust ~ data$greenbelt)

Residuals:
   Min     1Q Median     3Q    Max
-5.924 -2.160 -1.280  3.479  4.544

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)     78.520      6.690  11.736 7.38e-06 ***
data$greenbelt -21.216      5.507  -3.853  0.00627 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.849 on 7 degrees of freedom
Multiple R-squared:  0.6795,    Adjusted R-squared:  0.6337
F-statistic: 14.84 on 1 and 7 DF,  p-value: 0.006271
```

SubWay Users (0.8521)

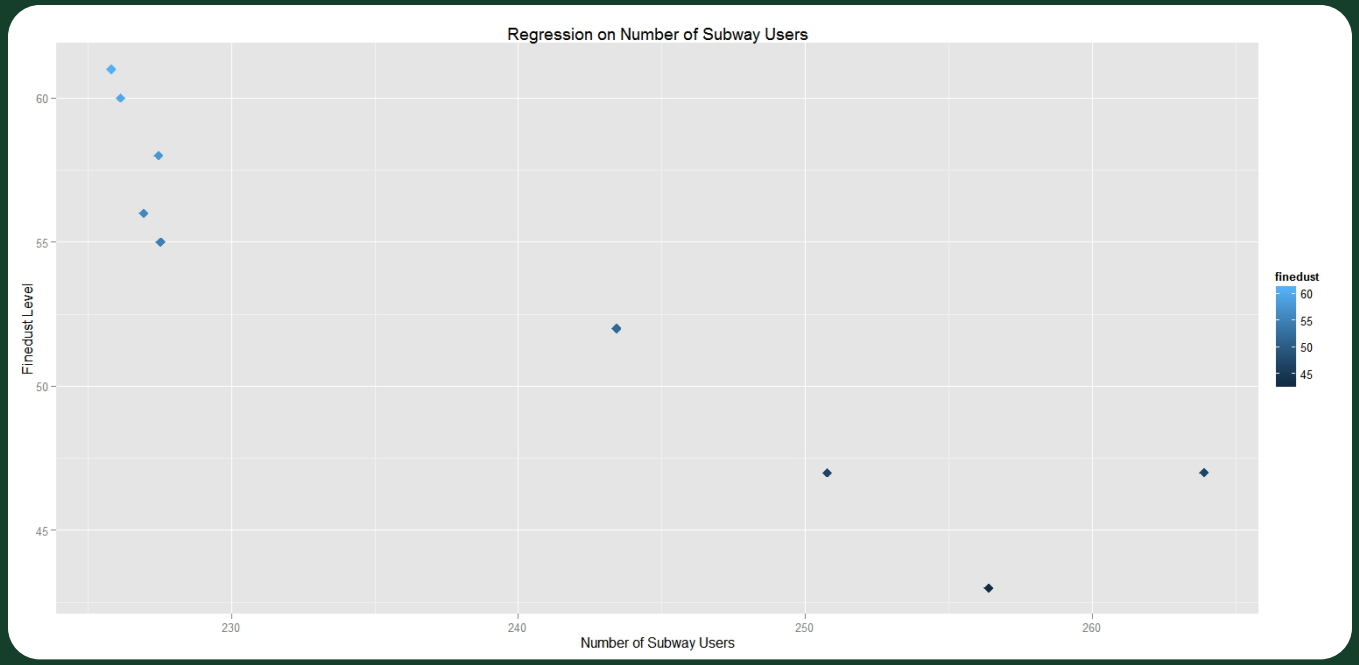
```
> summary(lm(data$finedust~data$sub_users))

Call:
lm(formula = data$finedust ~ data$sub_users)

Residuals:
    Min      1Q  Median      3Q     Max
-3.3568 -1.7905  0.4135  1.8966  3.5580

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)    145.92159   14.62586   9.977 2.17e-05 ***
data$sub_users  -0.38833    0.06116  -6.349 0.000385 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.615 on 7 degrees of freedom
Multiple R-squared:  0.8521,    Adjusted R-squared:  0.8309
F-statistic: 40.31 on 1 and 7 DF,  p-value: 0.0003855
```

Registered Cars (0.5602)

```
> summary(lm(data$finedust~data$car_regist))

Call:
lm(formula = data$finedust ~ data$car_regist)

Residuals:
   Min     1Q Median     3Q    Max
-6.476 -3.155  0.814  1.938  7.473

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)      276.29      74.71   3.698  0.00767 **
data$car_regist -762.64     255.38  -2.986  0.02033 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.508 on 7 degrees of freedom
Multiple R-squared:  0.5602,    Adjusted R-squared:  0.4974
F-statistic: 8.918 on 1 and 7 DF,  p-value: 0.02033
```

Setting Variables — Setting Hypothesis — Subset Selection Function — Regression Test (Six Factors) — Regression Test (Individually) — **Results & Discussions**

RESULTS & DISCUSSIONS (TOP 1 : NUMBER OF SUBWAY USERS)

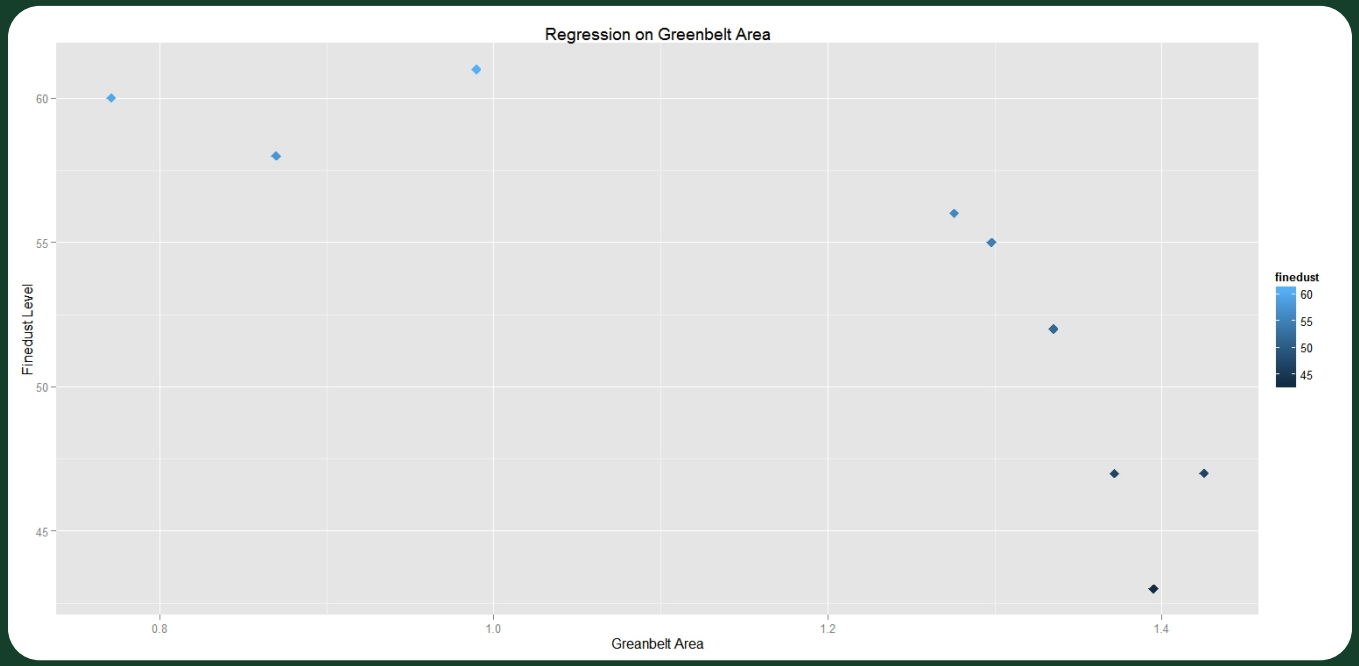


r^2 = 0.8521 p-value = 0.0003855

SubWay users dramatically increased from 2,277,298,000 to 2,619,529,000.
SubWay Expansion Policy.

Setting Variables → Setting Hypothesis → Subset Selection Function → Regression Test (Six Factors) → Regression Test (Individually) → **Results & Discussions**

RESULTS & DISCUSSIONS (TOP 2 : Green Belt Area)

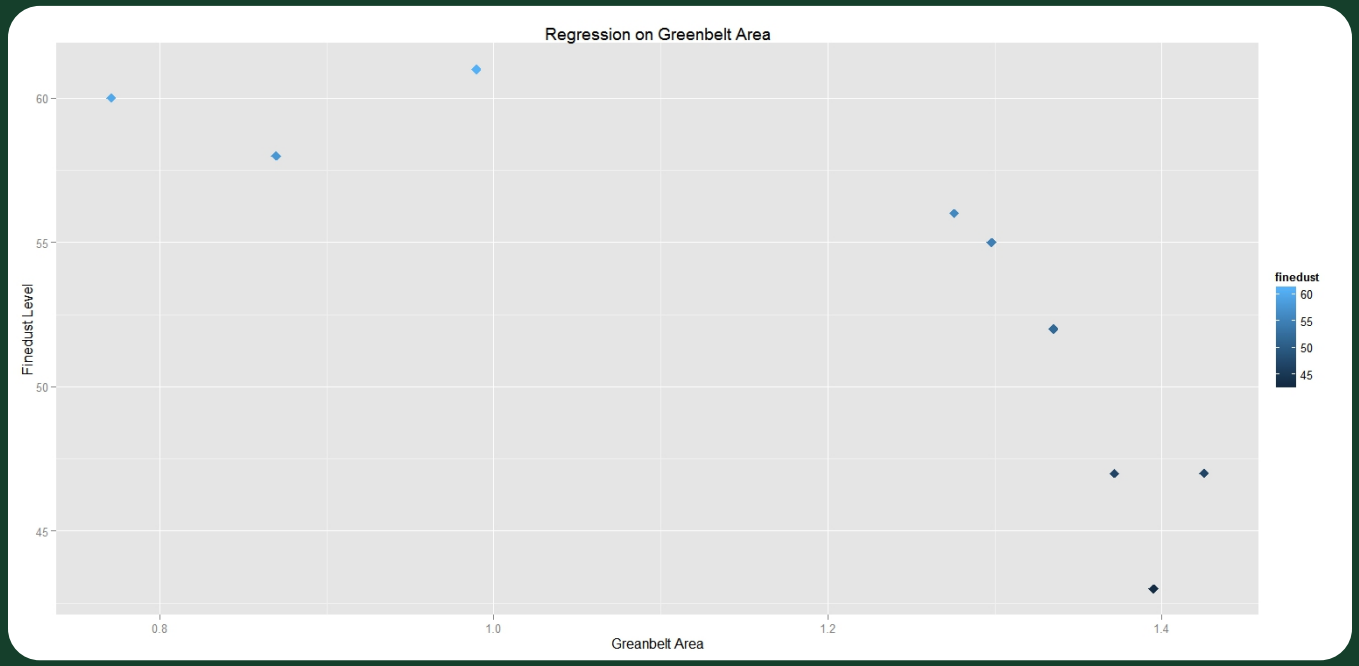


r^2 = 0.6795 p-value = 0.006271

Approximately 2% of air pollution will be cleansed by the green belt, according to study done by the department of rural development.

Setting Variables → Setting Hypothesis → Subset Selection Function → Regression Test (Six Factors) → Regression Test (Individually) → Results & Discussions

RESULTS & DISCUSSIONS (TOP 3 : Number of car registration)

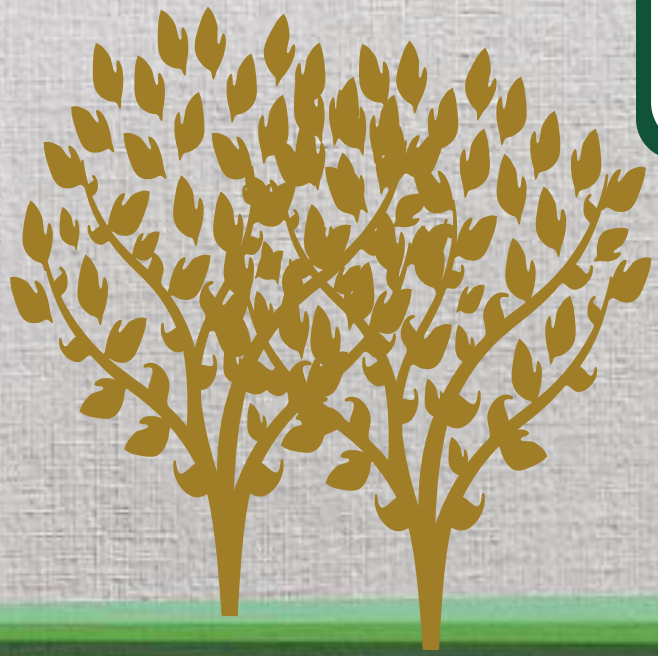


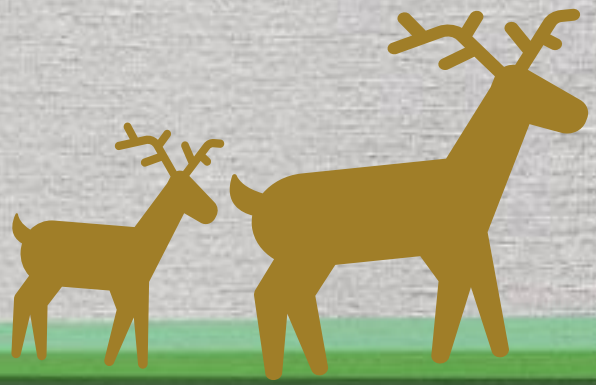
r^2 value = 0.5602 p-value = 0.02033

The span of the increase in registered cars < that of subway users.
Exhaustion gas regulation in action (e.g. Euro 6 from 2015)

3. CONCLUSION

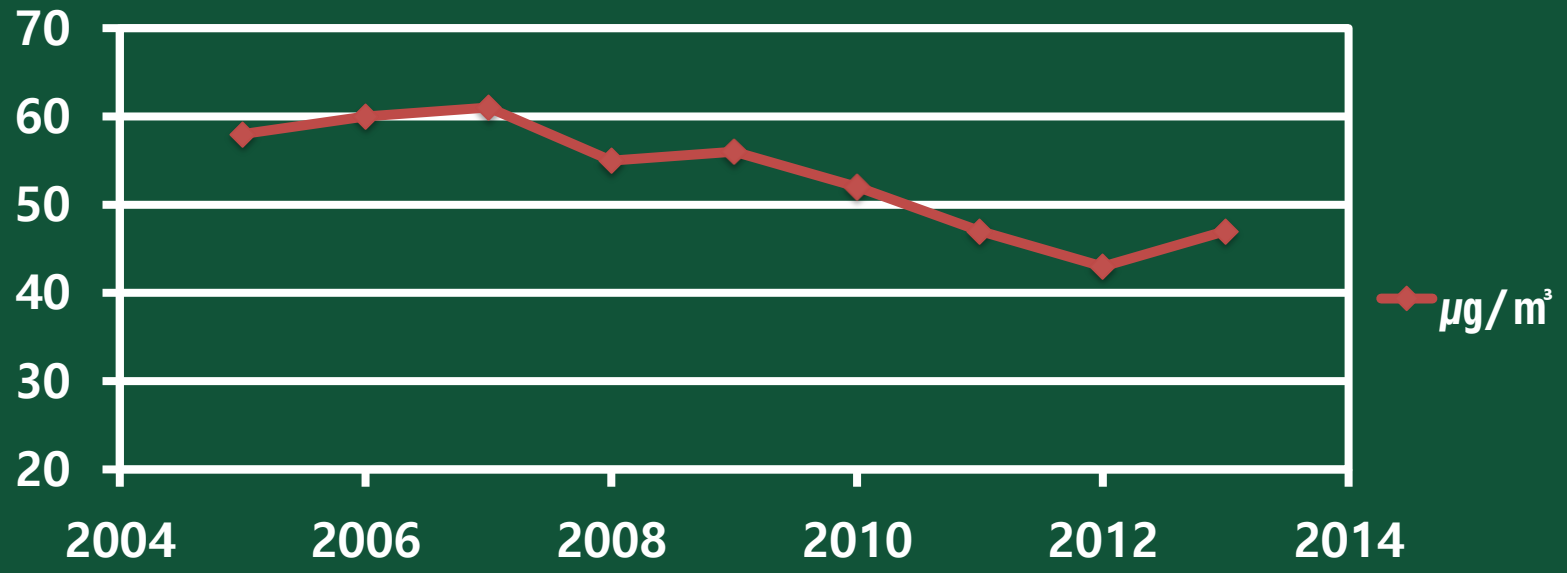
- Findings
- Limitations
- Future plan





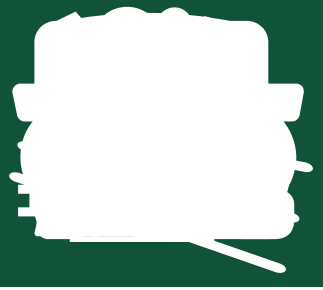
FINDING

Fine dust in Seoul



the Fine dust level in Seoul is rather decreasing throughout the years of 2005 to 2013, and that the subway users, the area of green belt, and the registered cars could be contributing to such phenomenon in regards to our findings in data analysis and research.

LIMITATION



(The population of Seoul)

We controlled For our independent variables by the population, But Still remain variables and other Factors We could not control For

There have possibility that other Factors NON-INCLUDING in our independent variables remain.

FUTURE STUDY

More studies and research need to be done to decrease the level of fine dust overstepping our limitation!

Thank you :)