

2017240008 미디어학부 고유경
2019150411 통계학과 배진희
2017150560 행정학과 안선정
2016320118 컴퓨터학과 윤석준
2019150454 통계학과 이강은
2019150441 통계학과 이나운
2019150458 통계학과 이정민

회귀분석_5조_주제발표

여름철 날씨와 모기포집량의 관계



〈목차〉

1) 서론

1-1. 주제 선정이유

1-2. 데이터소개

1-3. 회귀 분석 과정 소개

2) 본론 : 회귀분석

2-1. 데이터 전처리

2-2. 회귀식 적합

2-3. 회귀 진단

3) 결론

3-1 회귀 분석 결과 해석

3-2 일상생활에서의 행동요령

1-1. 주제 선정 이유

여름철 날씨와 모기포집량 사이에는
관계가 있을까?



1-1. 주제 선정 이유 - 여름철 날씨와 모기 포집량의 관계

“기후변화에 따라 말라리아 발생 위험이 달라질 수 있다.”

[채수미,김동진,윤석준,신호성, 2014,, 기온과 지역특성이 말라리아 발생에 미치는 영향,보건사회연구34(1),436-455]

“온도가 모기 생태계에 공간적, 시간적으로 영향을 줄 수 있다.”

[Reisen WK. Effect of temperature on *Culex tarsalis* (Diptera: Culicidae) from the Coachella and San Joaquin Valleys of California. J Med Entomol. 1995;32:636-645]

“풍속과 습도가 높을 수록 중국얼룩날개모기의 분포가 증가하였다.”

[Ji Hun Ryu, Heung Chul Kim, Kwang Shik Choi. (2018). Potential impact of climate change on malaria vectors in Korea. 한국농약과학회 학술발표대회 논문집, (), 476-476.]

1-1. 주제 선정 이유 - 연구의 의의

기온, 습도 등을 통해 모기포집량을 예측하여
다음 날 야외활동 계획에 참고 가능

모기포집량이 높을 것이라고 예측되는 경우
미리 모기 퇴치제를 구비하여 모기 물림과 전염병 예방 가능



1-2. 데이터 소개 - 모기 포집량(반응변수)

데이터 출처: http://news.seoul.go.kr/welfare/mos_dmsnblt2

	일시	포집량	모기	기타
1	2015-05-01	2801	2783	18
2	2015-05-02	2747	2729	18
⋮	⋮	⋮	⋮	⋮
920	2019-10-31	185	185	-



- 서울시는 매년 4월 15일~10월 31일 사이의 모기 포집량을 사이트에 공개
- 2015년~2019년, 총 5개년치 5월1일~10월31일의 관측치 920개

1-2. 데이터 소개 - 기상청 기상관측데이터(설명변수)

데이터 출처: <https://data.kma.go.kr/data/grnd/selectAsosRltmList.do?pgmNo=36>

	일시	기온(°C)	강수량(mm)	풍속(m/s)	상대습도(%)	현지기압(hPa)	지면온도(°C)	...
1	2015-05-01	21	N/A	2.5	58.8	1003.1	25.1	...
2	2015-05-02	20.7	0.5	2.1	64.4	1001.7	23.7	...
:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:
920	2019-10-31	11.2	N/A	1.6	60.3	111.1	11.2	...

- 데이터 수(n): 2015년~2019년, 총 5개년치 5월1일~10월31일 관측치 920개
- 설명변수 : 평균&최소&최대 기온, 강수량, 풍속, 상대습도, 지면온도 등 기상관련 변수
 - 결측치(N/A) 처리, 변수 별 scale 조정 등 전처리 필요

1-3. 회귀 분석 과정 소개 - 회귀식 적합 과정

설명 변수 정리

방법: 후진제거법



다중회귀모형 수립

R, SAS 사용



회귀 진단



결과 해석

2) 본론 : 회귀 분석



2-1. 데이터 전처리

	일시	기온(°C)	강수량(mm)	풍속(m/s)	상대습도(%)	현지기압(hPa)	지면온도(°C)	...
1	2015-05-01	21	N/A → 0	2.5	58.8	1003.1	25.1	...
2	2017-05-12	19.3	0 → 0.01	1.8	65.3	1001.1	21.6	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
920	2019-10-31	11.2	N/A → 0	1.6	60.3	1101.1	11.2	...

- 결측치는 평균으로 대체, 모기 포집량은 월평균으로 대체
- 강수량
 - NA 는 0으로 대체
 - 기존의 0(극소량)은 0.01로 비가 온 날과 오지 않은 날 구분하여 처리

2-1. 데이터 전처리

10개의 날씨 변수 선택

1. average_temp (평균기온)
2. low_temp (최소기온)
3. precipitation (일강수량)
4. average_wind (평균풍속)
5. average_humidity (평균 상대습도)
6. low_humidity (최소 상대습도)
7. total_daylight (합계 일조시간)
8. average_ground_temp (평균 지면온도)
9. average_vapor_pressure (평균 증기압)
10. average_atmos_pressure (평균 현지기압)

2-2. 회귀식 적합 - 기본 모델

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_{10} X_{10} + \varepsilon$$

Y : 모기 포집량

X1. average_temp

X2. low_temp

X3. precipitation

X4. average_wind

X5. average_humidity

X6. low_humidity

X7. total_daylight

X8. average_ground_temp

X9. average_vapor_pressure

X10. average_atmos_pressure

2-2. 회귀식 적합 - 기본 모델 R output

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_{10} X_{10}$$

```
Call:
lm(formula = mogi ~ ., data = dat_nodummy)

Residuals:
    Min       1Q   Median       3Q      Max
-4154.0 -1224.4  -315.5   704.6 21383.4

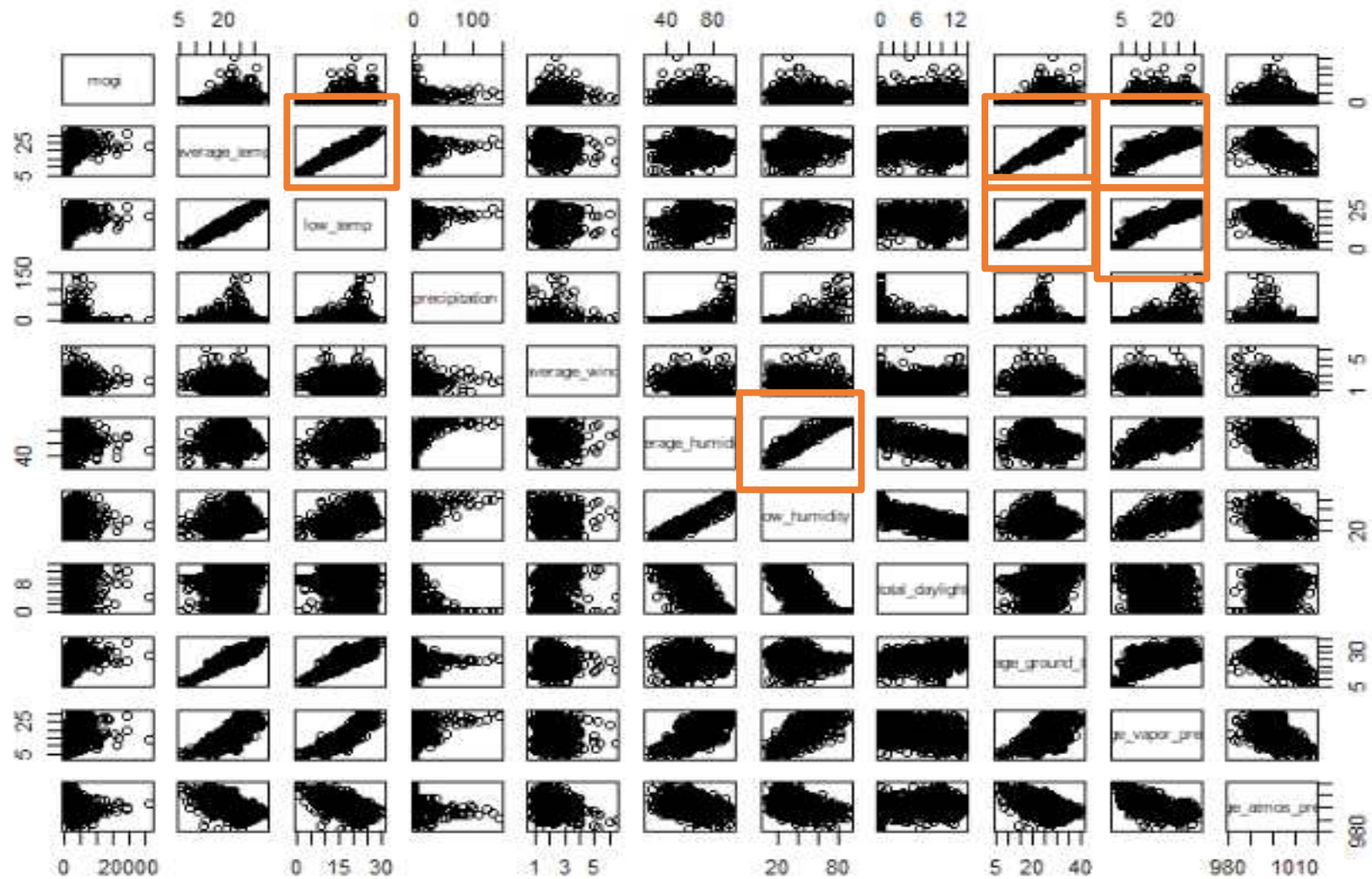
Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   -55556.594  18034.300   -3.081  0.002128 **
average_temp    122.354    102.874    1.189  0.234610
low_temp        88.466     81.840    1.081  0.279998
precipitation   -5.495      6.024   -0.912  0.361910
average_wind    838.236    119.305    7.026  4.17e-12 ***
average_humidity 77.547     18.330    4.231  2.57e-05 ***
low_humidity    -38.680     13.901   -2.783  0.005504 **
total_daylight  -118.657     33.815   -3.509  0.000472 ***
average_ground_temp 173.687     41.851    4.150  3.64e-05 ***
average_vapor_pressure -141.354     58.745   -2.406  0.016318 *
average_atmos_pressure 48.917     17.424    2.807  0.005100 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2219 on 909 degrees of freedom
Multiple R-squared:  0.2543,    Adjusted R-squared:  0.2461
F-statistic: 51 on 10 and 909 DF,  p-value: < 2.2e-16
```

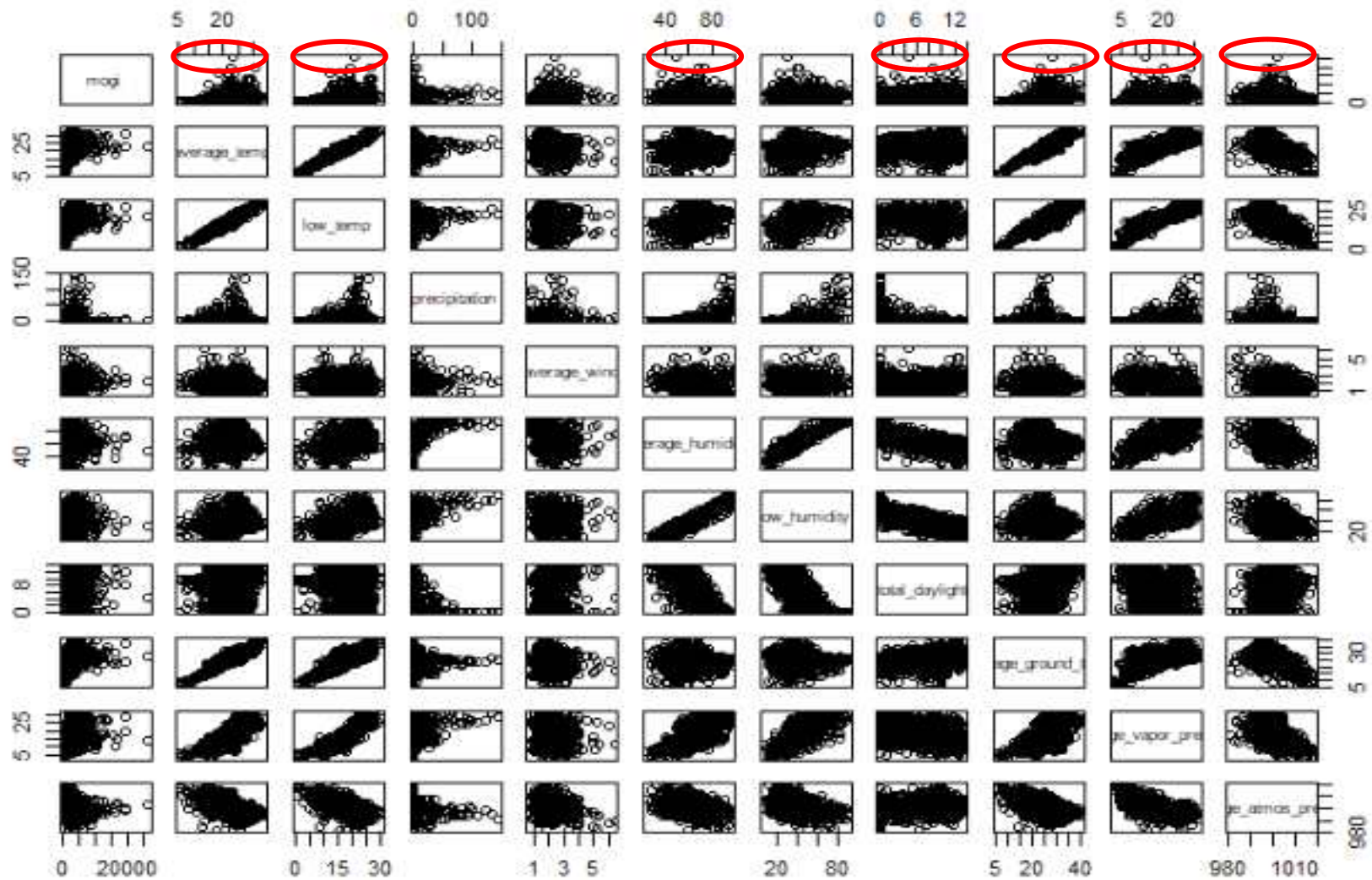
유의하지
않은 변수
존재

모델 설명력
0.2543

2-2. 회귀식 적합 - 기본 모델 산점도 행렬



2-2. 회귀식 적합 - 반응 변수 이상치 제거 필요



average_temp VS low_temp

(VIF)

47

VS

37

✓ 평균 온도 제거

```
> f(data1) # average_temp 제거한 모델
```

```
Call:
lm(formula = mogi ~ ., data = data)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-3500.3 -1006.1 -189.5   772.5  6829.9
```

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -27406.098  12691.420  -2.159  0.031085 *
low_temp       102.127    44.445    2.298  0.021802 *
precipitation   -2.221     4.225   -0.526  0.599235
average_wind    630.760    82.733    7.624  6.32e-14 ***
average_humidity  49.322    12.115    4.071  5.09e-05 ***
low_humidity   -32.629     9.727   -3.355  0.000828 ***
total_daylight -90.933    23.373   -3.891  0.000108 ***
average_ground_temp 157.660    26.486    5.953  3.80e-09 ***
average_vapor_pressure -51.308    37.229   -1.378  0.168490
average_atmos_pressure 23.609    12.278    1.923  0.054816 .
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 1550 on 887 degrees of freedom
Multiple R-squared:  0.3593,    Adjusted R-squared:  0.3528
F-statistic: 55.26 on 9 and 887 DF, p-value: < 2.2e-16
```

최저 온도 제거

```
> f(data1) # low_temp 제거한 모델
```

```
Call:
lm(formula = mogi ~ ., data = data)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-3676.7 -984.4 -213.3   802.6  6796.4
```

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -29108.084  12641.000  -2.303  0.021528 *
average_temp    208.535    55.661    3.747  0.000191 ***
precipitation   -2.570     4.180   -0.615  0.538816
average_wind    674.339    82.018    8.222  7.08e-16 ***
average_humidity  62.409    12.839    4.861  1.38e-06 ***
low_humidity   -28.458     9.764   -2.915  0.003651 **
total_daylight -102.752    21.118   -4.866  1.35e-06 ***
average_ground_temp 117.650    29.353    4.008  6.63e-05 ***
average_vapor_pressure -107.221    40.208   -2.667  0.007800 **
average_atmos_pressure 23.490    12.211    1.924  0.054717 .
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 1545 on 887 degrees of freedom
Multiple R-squared:  0.3655,    Adjusted R-squared:  0.359
F-statistic: 56.77 on 9 and 887 DF, p-value: < 2.2e-16
```


low_temp VS average_ground_temp

(VIF)

22

VS

8

✓ 최저 온도 제거

평균 지면 온도제거

```
Call:
lm(formula = mogi ~ ., data = data)
```

Residuals:

Min	1Q	Median	3Q	Max
-3461.7	-1011.1	-222.3	734.0	6856.8

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-25743.683	12701.283	-2.027	0.042975 *
precipitation	-3.370	4.205	-0.801	0.423176
average_wind	652.201	82.404	7.915	7.36e-15 ***
average_humidity	40.103	11.459	3.500	0.000489 ***
low_humidity	-33.429	9.744	-3.431	0.000630 ***
total_daylight	-114.597	21.032	-5.449	6.58e-08 ***
average_ground_temp	194.255	21.214	9.157	< 2e-16 ***
average_vapor_pressure	14.383	23.905	0.602	0.547551
average_atmos_pressure	22.491	12.298	1.829	0.067750 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1554 on 888 degrees of freedom
Multiple R-squared: 0.3554, Adjusted R-squared: 0.3496
F-statistic: 61.21 on 8 and 888 DF, p-value: < 2.2e-16

```
Call:
lm(formula = mogi ~ ., data = data)
```

Residuals:

Min	1Q	Median	3Q	Max
-3904.1	-1002.5	-206.9	745.1	6936.3

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-3594.4691	12275.8166	-0.293	0.76974
low_temp	261.2059	36.1944	7.217	1.14e-12 ***
precipitation	0.2248	4.2856	0.052	0.95818
average_wind	553.7570	83.2851	6.649	5.15e-11 ***
average_humidity	34.0100	12.0660	2.819	0.00493 **
low_humidity	-32.6368	9.9135	-3.292	0.00103 **
total_daylight	-25.7174	21.0427	-1.222	0.22197
average_vapor_pressure	-61.2444	37.9053	-1.616	0.10651
average_atmos_pressure	1.6742	11.9367	0.140	0.88849

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1580 on 888 degrees of freedom
Multiple R-squared: 0.3337, Adjusted R-squared: 0.3277
F-statistic: 55.58 on 8 and 888 DF, p-value: < 2.2e-16

low_humidity VS average_humidity

(VIF)

7

VS

9

✓ 최저 습도 제거

```
1. 300434      1. 100030      0. 340700
> data5<- data2[, -5] #low_humidity 제거
> f(data5) # 편향다.

Call:
lm(formula = mogi ~ ., data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-3536.1 -1043.3 -198.9   795.5  6870.0

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -21473.753  12716.496  -1.689   0.0916 .
precipitation    -4.974     4.204   -1.183   0.2371 .
average_wind    593.839    81.115    7.321 5.51e-13 ***
average_humidity  19.778     9.868    2.004   0.0453 *
total_daylight  -81.608    18.819   -4.337 1.61e-05 ***
average_ground_temp 196.228    21.334    9.198 < 2e-16 ***
average_vapor_pressure  -3.089    23.497   -0.131  0.8954
average_atmos_pressure  18.276    12.310    1.485   0.1380

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1563 on 889 degrees of freedom
Multiple R-squared:  0.3469,    Adjusted R-squared:  0.3418
F-statistic: 67.46 on 7 and 889 DF,  p-value: < 2.2e-16
```

평균 습도 제거

```
> data4<-data2[, -4] #average_humidity 제거
> f(data4)# average_humidity 제거하고 나니 low_humidity, low_temp, pr

Call:
lm(formula = mogi ~ ., data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-3829.1 -1036.2 -218.7   734.7  6942.4

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -12802.056  12227.707  -1.047   0.29540
precipitation    -1.067     4.180   -0.255   0.79854
average_wind    610.870    82.067    7.444 2.31e-13 ***
low_humidity    -15.798     8.393   -1.882   0.06012 .
total_daylight  -107.943    21.078   -5.121 3.73e-07 ***
average_ground_temp 153.737    17.888    8.594 < 2e-16 ***
average_vapor_pressure  60.017    20.162    2.977   0.00299 **
average_atmos_pressure  11.603    11.973    0.969   0.33274

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1564 on 889 degrees of freedom
Multiple R-squared:  0.3465,    Adjusted R-squared:  0.3414
F-statistic: 67.35 on 7 and 889 DF,  p-value: < 2.2e-16
```

2-2. 회귀식 적합 - Rain dummy 추가

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-2615.455	499.435	-5.237	2.04e-07	***
<u>precipitation</u>	<u>-5.087</u>	<u>4.192</u>	<u>-1.213</u>	<u>0.22528</u>	
average_wind	550.361	75.366	7.303	6.26e-13	***
average_humidity	16.222	5.882	2.758	0.00593	**
total_daylight	-78.075	18.301	-4.266	2.20e-05	***
average_ground_temp	183.013	9.424	19.419	< 2e-16	***

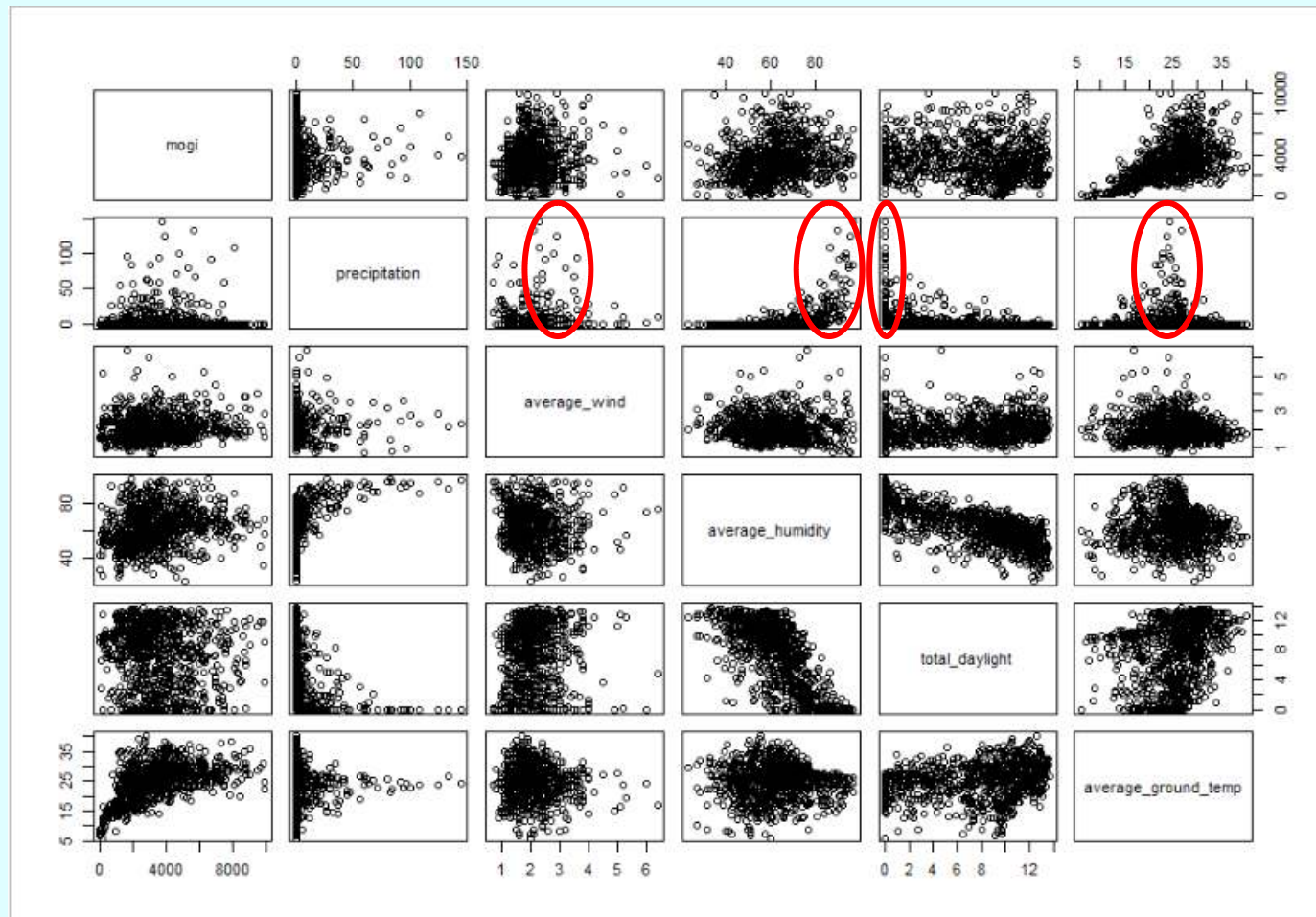
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1563 on 891 degrees of freedom
Multiple R-squared: 0.3453, Adjusted R-squared: 0.3416
F-statistic: 93.98 on 5 and 891 DF, p-value: < 2.2e-16

X3. precipitation

유의하지 않으므로 제거?

2-2. 회귀식 적합 - Rain dummy 추가



-> 강수에 대한 더미 변수 필요

2-2. 회귀식 적합 - 다중공선성 및 후진 제거법에 의한 변수 제거

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \gamma \delta_5 + \varepsilon$$

X_1 : *average_wind*

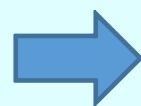
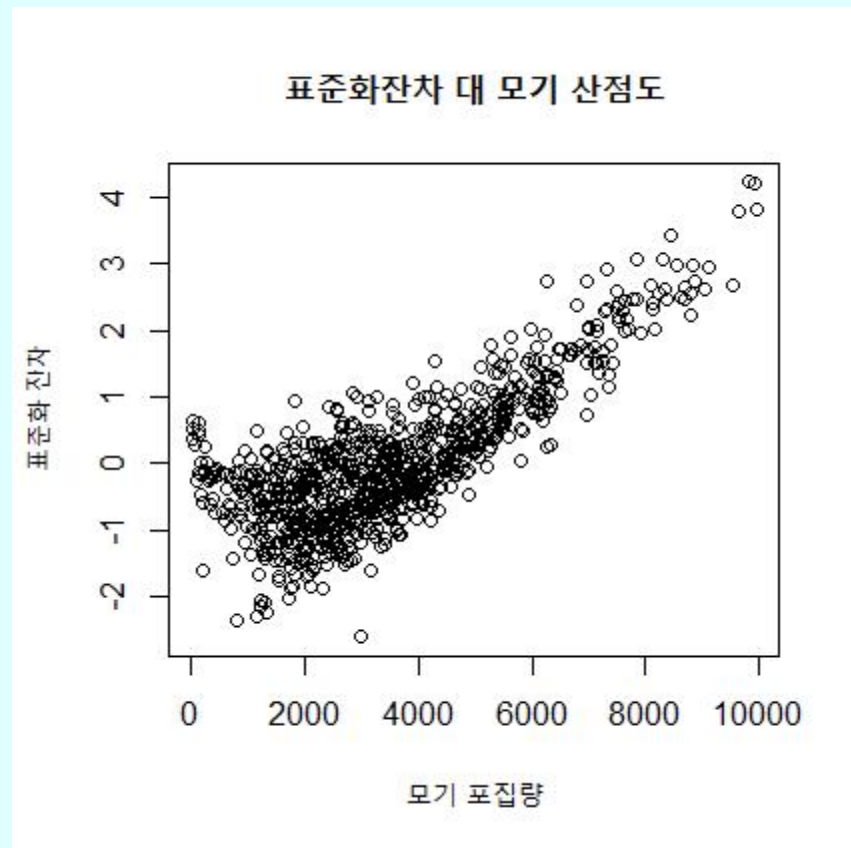
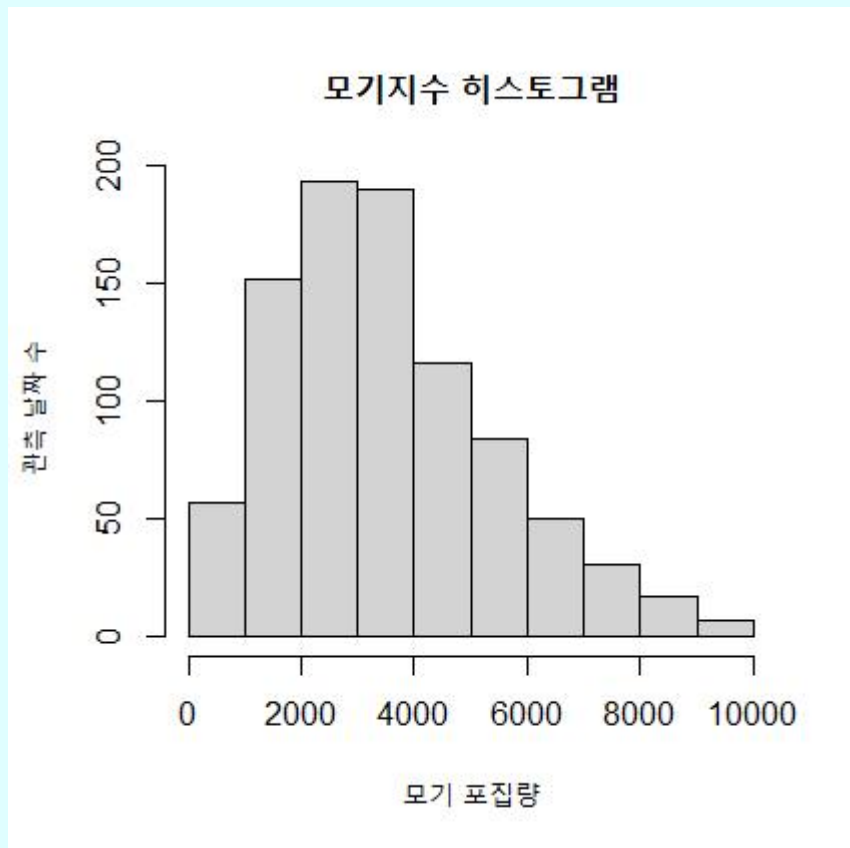
X_2 : *average_humidity*

X_3 : *total_daylight*

X_4 : *average_ground_temp*

δ_5 : *rain* $\begin{cases} 0, & \text{precipitation} = 0 \\ 1, & \text{precipitation} \neq 0 \end{cases}$

2-2. 회귀식 적합 - 변수 변환



로그 변환 필요

2-2. 회귀식 적합 - 로그 변환 이후 R output

$$\widehat{\log Y} = \widehat{\beta}_0 + \widehat{\beta}_1 X_1 + \widehat{\beta}_2 X_2 + \widehat{\beta}_3 X_3 + \widehat{\beta}_4 X_4 + \hat{\gamma} \delta_5 + \varepsilon$$

```
Call:
lm(formula = logmogi ~ ., data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-2.95454 -0.31697  0.05639  0.35463  1.41710

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    5.237454   0.168297   31.120 < 2e-16 ***
average_wind    0.192518   0.025871    7.441 2.35e-13 ***
average_humidity 0.007834   0.002008    3.902 0.000103 ***
total_daylight -0.038888   0.006581   -5.909 4.90e-09 ***
average_ground_temp 0.087054  0.003217   27.063 < 2e-16 ***
rrain          -0.129527   0.050089   -2.586 0.009869 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5341 on 891 degrees of freedom
Multiple R-squared:  0.4909,    Adjusted R-squared:  0.488
F-statistic: 171.8 on 5 and 891 DF,  p-value: < 2.2e-16
```


2-2. 회귀식 적합 - 계절 dummy 변수 추가

여름철에 모기 포집량이
증가한다는 자명한 사실



계절
Dummy 변수
추가

```
> f<-function(data){  
+   model<-lm(logmog~., dat=data)  
+   summary(model)} # 시간 단축용 함수입니다..  
> f(dat1) # dummy 추가해도 크게 유의해지지 않음..?
```

Call:

```
lm(formula = logmog ~ ., data = data)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-3.02864	-0.30768	0.04358	0.35130	1.35126

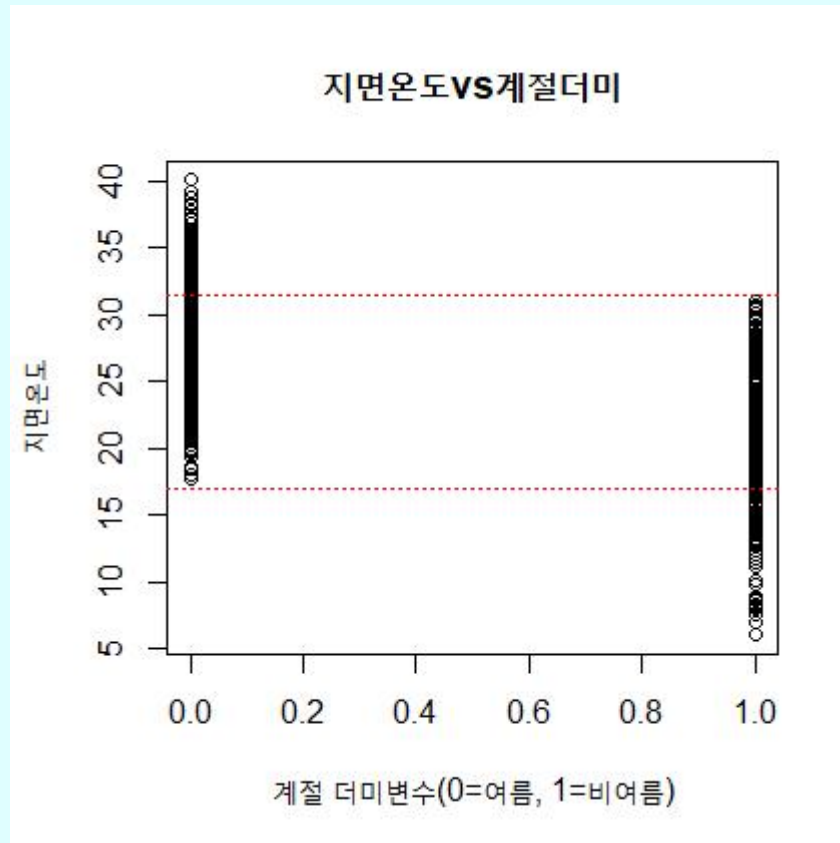
Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	5.551248	0.198640	27.946	< 2e-16 ***
dummy	-0.157858	0.053655	-2.942	0.003345 **
average_wind	0.192280	0.025761	7.464	2.00e-13 ***
average_humidity	0.006848	0.002027	3.378	0.000761 ***
total_daylight	-0.036246	0.006615	-5.480	5.55e-08 ***
average_ground_temp	0.078378	0.004354	18.002	< 2e-16 ***
rrain	-0.132438	0.049885	-2.655	0.008075 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5318 on 890 degrees of freedom
Multiple R-squared: 0.4958, Adjusted R-squared: 0.4924
F-statistic: 145.9 on 6 and 890 DF, p-value: < 2.2e-16

2-2. 회귀식 적합 - average_ground_temp dummy 변수 추가



지면 온도와 계절 더미 변수 관계성 있음

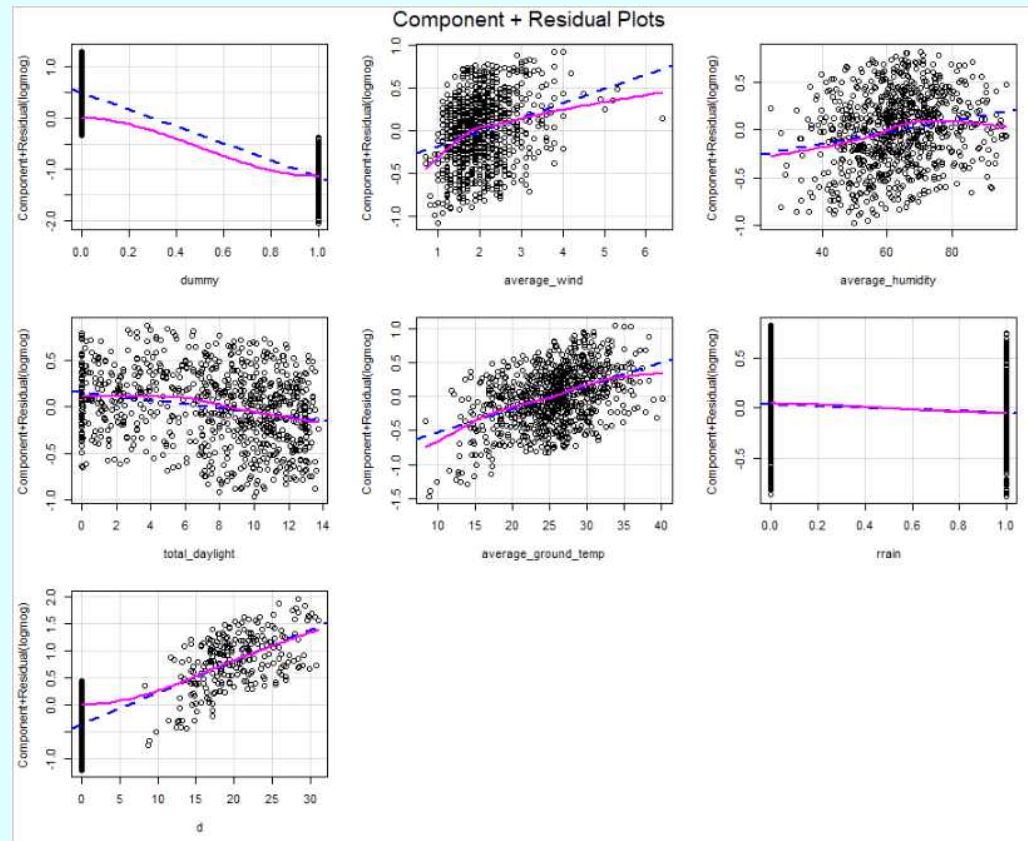
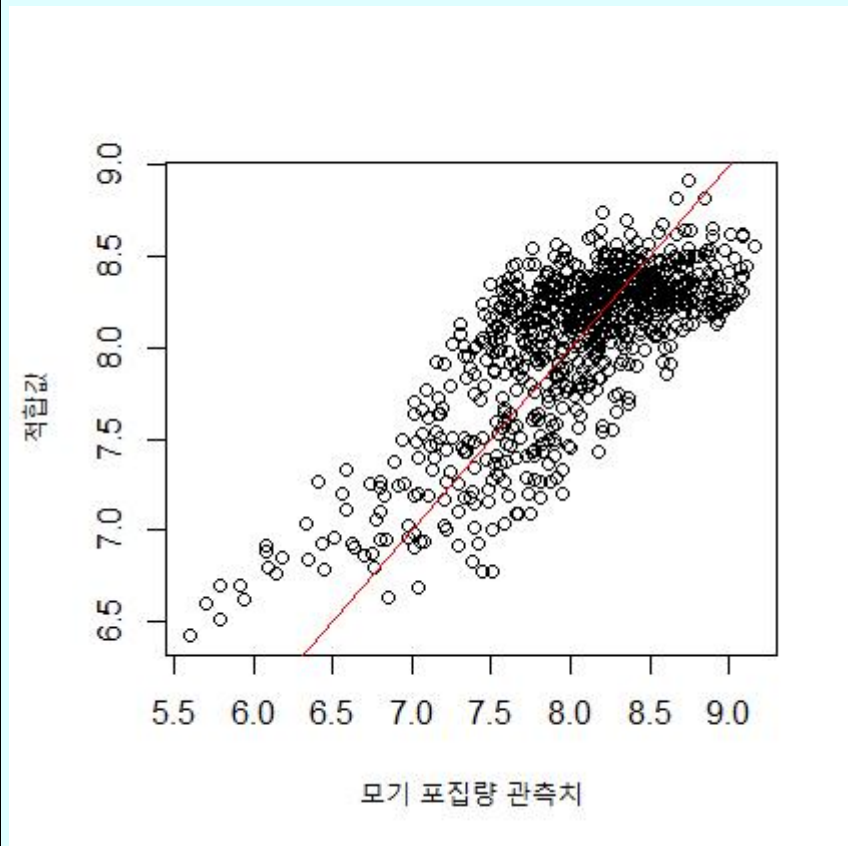
-> 지면 온도와 계절 더미변수에 대한 상호작용 변수 추가

2-2. 회귀식 적합 - 최종 회귀 모형

최종 회귀 모형

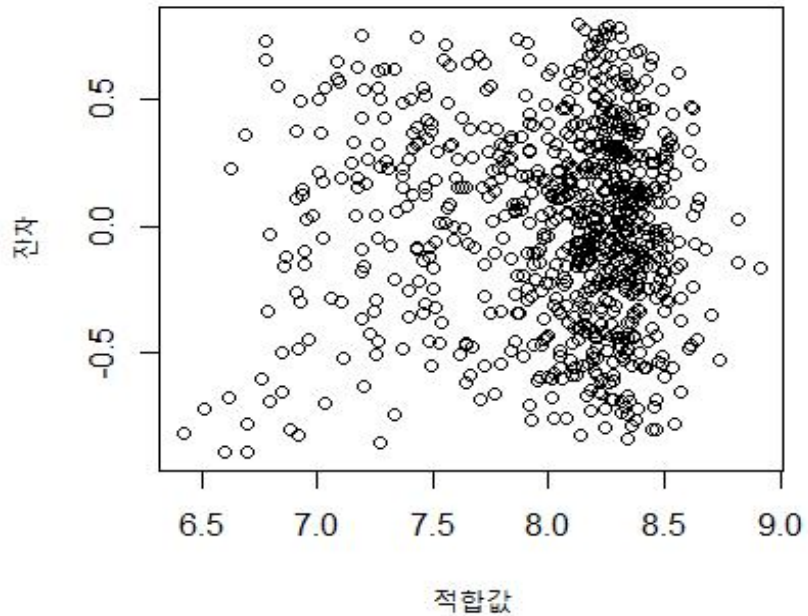
$$\widehat{\log y} = 6.77 - 1.64 \times \delta_{not-summer} + 1.6 \times \frac{x_{avg-wind}}{10} + 6 \times \frac{x_{avg-humi}}{1000} \\ - 2 \times \frac{x_{daylight}}{100} + 3.4 \times \frac{x_{avg-g-temp}}{100} - 8.8 \times \frac{\delta_{rain}}{100} + 6 \times \frac{\theta_{not-summer \times temp}}{100}$$

2-3 회귀 진단 - 선형성

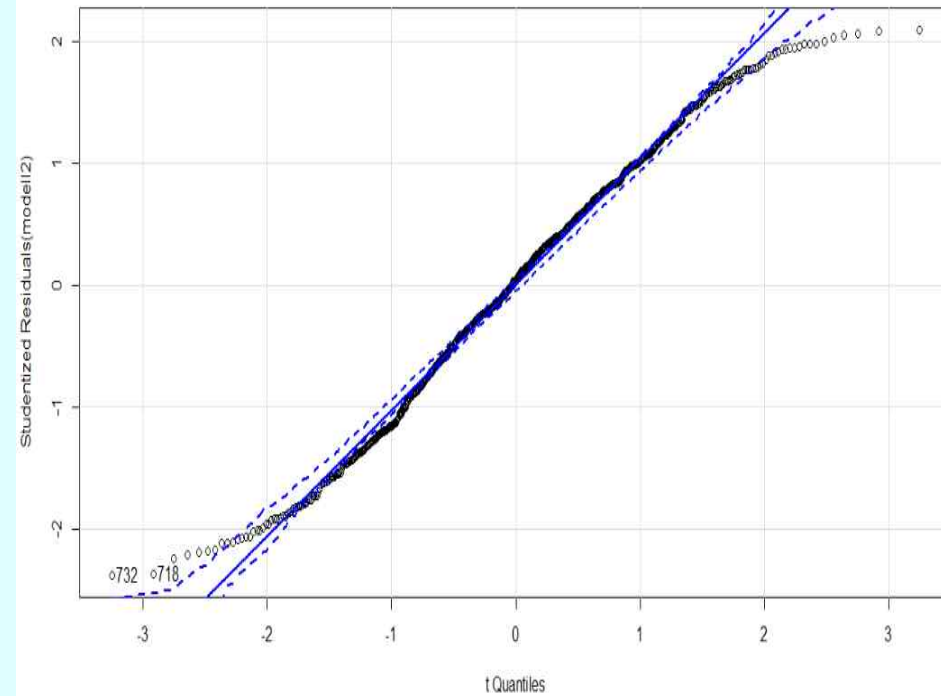


2-3 회귀 진단 - 잔차

잔차 대 적합값 플롯



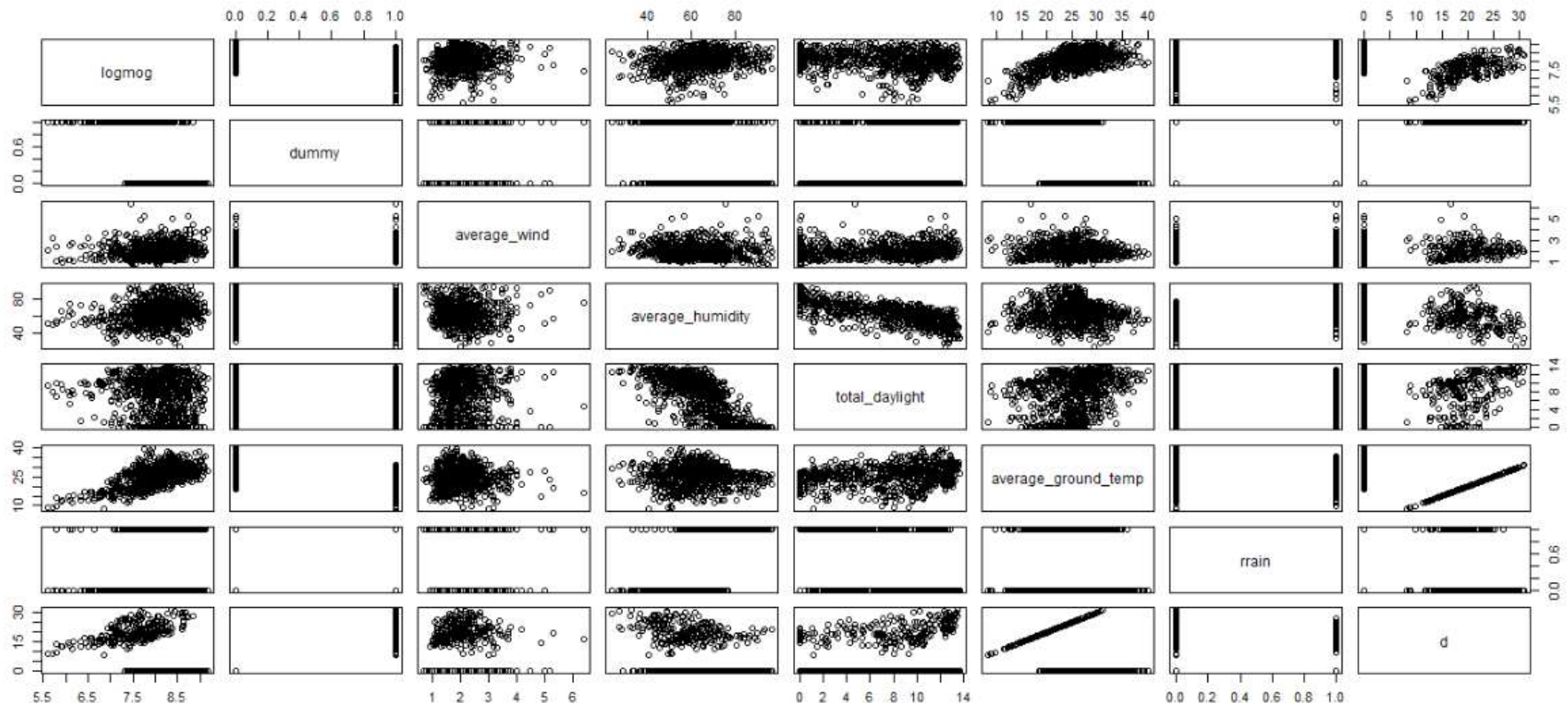
Q - Q plot



2-3. 회귀 진단 - 다중공선성

```
> vif(model12)
```

	dummy	average_wind	average_humidity	total_daylight
	30.100857	1.036820	2.381188	2.511777
average_ground_temp		rrain	d	
	3.252823	1.851649	22.972154	



3) 결론



3-1. 회귀 분석 결과 해석

변수 목록

$\delta_{not-summer}$	여름 / 비여름 에 따른 dummy 변수	$x_{daylight}$	일조량 (hr)
$x_{avg-wind}$	평균 풍속 (m/s)	$x_{avg-g-temp}$	평균 지면 온도 (°C)
$x_{avg-humi}$	평균 상대 습도 (%)	δ_{rain}	강수 여부에 따른 dummy 변수
		$\theta_{summer \times temp}$	여름 여부와 온도 간의 상호작용 변수

3-1. 회귀 분석 결과 해석

최종 회귀 모형

$$\widehat{\log y} = 6.77 - 1.64 \times \delta_{not-summer} + 1.6 \times \frac{x_{avg-wind}}{10} + 6 \times \frac{x_{avg-humi}}{1000} \\ - 2 \times \frac{x_{daylight}}{100} + 3.4 \times \frac{x_{avg-g-temp}}{100} - 8.8 \times \frac{\delta_{rain}}{100} + 6 \times \frac{\theta_{not-summer \times temp}}{100}$$

3-1. 회귀 분석 결과 해석

1) 여름이 아닐 때 ($\delta_{not-summer} = 1$)

$$\widehat{\log y} = 5.13 + 1.6 \times \frac{x_{avg-wind}}{100} + 6 \times \frac{x_{avg-humi}}{1000} - 2 \times \frac{x_{daylight}}{100} \\ + 4 \times \frac{x_{avg-g-temp}}{100} - 8.8 \times \frac{\delta_{rain}}{100}$$



절편항 : 5.13 -> 모기 약 169마리

평균풍속 1 m/s 증가 -> 모기 개체수 약 17.35% 증가

평균습도 1% 증가 -> 모기 개체수 약 0.6 증가

평균지면온도 1 °C 증가 -> 모기 개체수 약 4.08% 증가

일조량 1시간 증가 -> 모기 개체수 약 2.02% 감소

3-1. 회귀 분석 결과 해석

2) 여름 일 때 ($\delta_{not-summer} = 0$)

$$\widehat{\log y} = 6.77 + 1.6 \times \frac{x_{avg-wind}}{10} + 6 \times \frac{x_{avg-humi}}{1000} - 2 \times \frac{x_{daylight}}{100} \\ + 3.4 \times \frac{x_{avg-g-temp}}{100} - 8.8 \times \frac{\delta_{rain}}{100}$$



절편항 : 6.77 -> 모기 약 871마리

평균풍속 1 m/s 증가 -> 모기 개체수 약 17.35% 증가

평균습도 1% 증가 -> 모기 개체수 약 0.6 증가

평균지면온도 1 °C 증가 -> 모기 개체수 약 3.46% 증가

일조량 1시간 증가 -> 모기 개체수 약 2.02% 감소

3-1. 회귀 분석 결과 해석 - 회귀계수

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	6.777139	0.166580	40.684	< 2e-16	***
dummy	-1.641653	0.158895	-10.332	< 2e-16	***
average_wind	0.165857	0.019931	8.322	3.63e-16	***
average_humidity	0.006113	0.001559	3.921	9.56e-05	***
total_daylight	-0.022383	0.005023	-4.456	9.50e-06	***
average_ground_temp	0.034308	0.004444	7.720	3.39e-14	***
rrain	-0.087889	0.037337	-2.354	0.0188	*
d	0.059007	0.006619	8.914	< 2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3836 on 817 degrees of freedom
Multiple R-squared: 0.5732, Adjusted R-squared: 0.5695
F-statistic: 156.7 on 7 and 817 DF, p-value: < 2.2e-16

비여름

음의
상관관계

일조량

강수여부

평균 풍속

양의
상관관계

평균 상대 습도

평균 지면 온도

3-1. 회귀 분석 결과 해석 - 평균 풍속

$$\widehat{\log y} = 6.77 - 1.64 \times \delta_{not-summer} + 1.6 \times \frac{x_{avg-wind}}{10} + 6 \times \frac{x_{avg-humi}}{1000} \\ - 2 \times \frac{x_{daylight}}{100} + 3.4 \times \frac{x_{avg-g-temp}}{100} - 8.8 \times \frac{\delta_{rain}}{100} + 6 \times \frac{\theta_{not-summer \times temp}}{100}$$

풍속이 빠를수록 모기의 활동성은 증가한다.

- 일반적으로 바람이 강할 수록 모기의 활동에 방해를 한다.
- 하지만 바람이 빠를수록 동물의 체취(이산화탄소, 땀 냄새) 또한 더 멀리 퍼지게 되고 모기의 탐지가 더 수월해지기 때문에 결과적으로 양의 상관관계를 가진다.

3-1. 회귀 분석 결과 해석 - 평균상대습도

$$\widehat{\log y} = 6.77 - 1.64 \times \delta_{not-summer} + 1.6 \times \frac{x_{avg-wind}}{10} + 6 \times \frac{x_{avg-humi}}{1000} \\ - 2 \times \frac{x_{daylight}}{100} + 3.4 \times \frac{x_{avg-g-temp}}{100} - 8.8 \times \frac{\delta_{rain}}{100} + 6 \times \frac{\theta_{not-summer \times temp}}{100}$$

습도가 높을수록 모기는 활발하게 활동한다.

- 습도가 높을수록 모기는 더 활발히 산란하며, 부화시기를 앞당긴다.
- 습도가 높으면 땀이 잘 증발하지 않게되고 그 체취는 모기를 유인한다.

3-1. 회귀 분석 결과 해석 - 일조량

$$\widehat{\log y} = 6.77 - 1.64 \times \delta_{not-summer} + 1.6 \times \frac{x_{avg-wind}}{10} + 6 \times \frac{x_{avg-humi}}{1000} \\ - 2 \times \frac{x_{daylight}}{100} + 3.4 \times \frac{x_{avg-g-temp}}{100} - 8.8 \times \frac{\delta_{rain}}{100} + 6 \times \frac{\theta_{not-summer \times temp}}{100}$$

모기는 야행성이기 때문에 일조시간이 짧을수록 활동시간이 늘어난다.

- 모든 모기가 야행성은 아니지만 한국에서 흔히 볼 수 있는 모기 종들의 하루 활동 시간은 초저녁~새벽이다.
- 모기는 일반적으로 빛에 본능적으로 이끌리지만 너무 밝은 백색광의 경우에는 거부 반응을 보인다.

3-1. 회귀 분석 결과 해석 - 평균 지면 온도

$$\widehat{\log y} = 6.77 - 1.64 \times \delta_{not-summer} + 1.6 \times \frac{x_{avg-wind}}{10} + 6 \times \frac{x_{avg-humi}}{1000} \\ - 2 \times \frac{x_{daylight}}{100} + 3.4 \times \frac{x_{avg-g-temp}}{100} - 8.8 \times \frac{\delta_{rain}}{100} + 6 \times \frac{\theta_{not-summer \times temp}}{100}$$

모기는 일반적으로 높은 기온 환경에서 더 활발히 활동한다.

- 모기는 기온이 32 °C 이하에서는 기온이 높을수록 더 많은 활동성을 보인다.
- 32 °C 이상에서는 성충 모기는 하면에 들어간다. 하지만 그 만큼 실내로 피신하는 경향을 보인다.

3-1. 회귀 분석 결과 해석 - 강수여부

$$\widehat{\log y} = 6.77 - 1.64 \times \delta_{not-summer} + 1.6 \times \frac{x_{avg-wind}}{10} + 6 \times \frac{x_{avg-humi}}{1000} \\ - 2 \times \frac{x_{daylight}}{100} + 3.4 \times \frac{x_{avg-g-temp}}{100} - 8.8 \times \frac{\delta_{rain}}{100} + 6 \times \frac{\theta_{not-summer \times temp}}{100}$$

비가 오면 모기의 활동성이 떨어진다.

- 비가오면 교미,흡혈,산란 활동에 어려움이 생긴다.
- 장마철 10일 이상 강수가 지속될 경우 모기 개체수는 급감한다.

3-2. 일상생활에서의 행동 요령

1.

바람이 빠르게 부는 날에는 체취가 멀리 날아가 모기 유인 반경이 넓어지므로 실외 활동을 자제한다.

2.

흐린 날, 혹은 해가 짧은 날에는 모기가 오랫동안 활동하므로 이를 고려해 모기 퇴치 용품(전자 모기향, 모기향 등)을 사용한다.

3.

기온이 높은 날은 땀도 많이 나고, 모기의 활동성이 증가하므로 실외 활동을 자제하고, 35도 이상 고온의 날씨라면 실내로의 침입에 대비한다.

4.

일기예보를 참고하여 다음날 모기 개체 수가 많을 것 같다면 미리 모기 퇴치제, 관련 의약품을 구비한다.

데이터 출처 및 참고자료

1. 데이터 출처

모기 포집량

http://news.seoul.go.kr/welfare/mos_dmsnblt2

기상관측 데이터

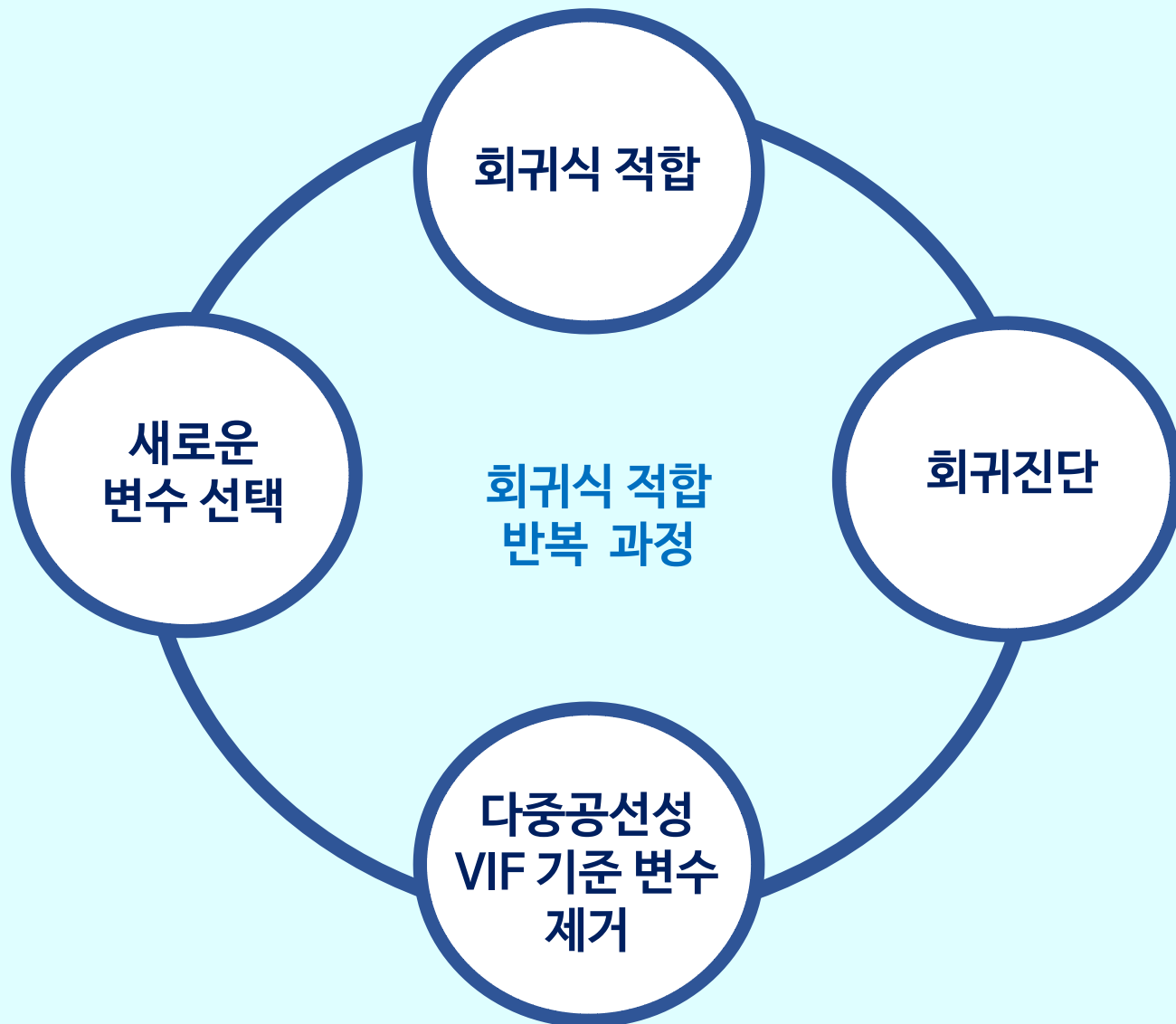
<https://data.kma.go.kr/data/grnd/selectAsosRltmList.do?pgmNo=36>

2. 참고자료

신호성, 채수미, 김동진, 윤석준. (2014). 기온과 지역특성이 말라리아 발생에 미치는 영향. 보건사회연구, 34(1), 436-455.

장진영. (2014). 모기 밀도와 기후 요인과의 연관성 : 경기도 고양시 내 2008-2012년 자료를 중심으로.

2-2. 회귀식 적합 - 후진 제거법 및 변수 선택



2-2. 회귀식 적합 - 다중공선성 및 후진 제거법에 의한 변수 제거

~~*X1. average_temp*~~

~~*X2. low_temp*~~

~~*X3. precipitation*~~

X4. average_wind

X5. average_humidity

~~*X6. low_humidity*~~

X7. total_daylight

X8. average_ground_temp

~~*X9. average_vapor_pressure*~~

~~*X10. average_atmos_pressure*~~

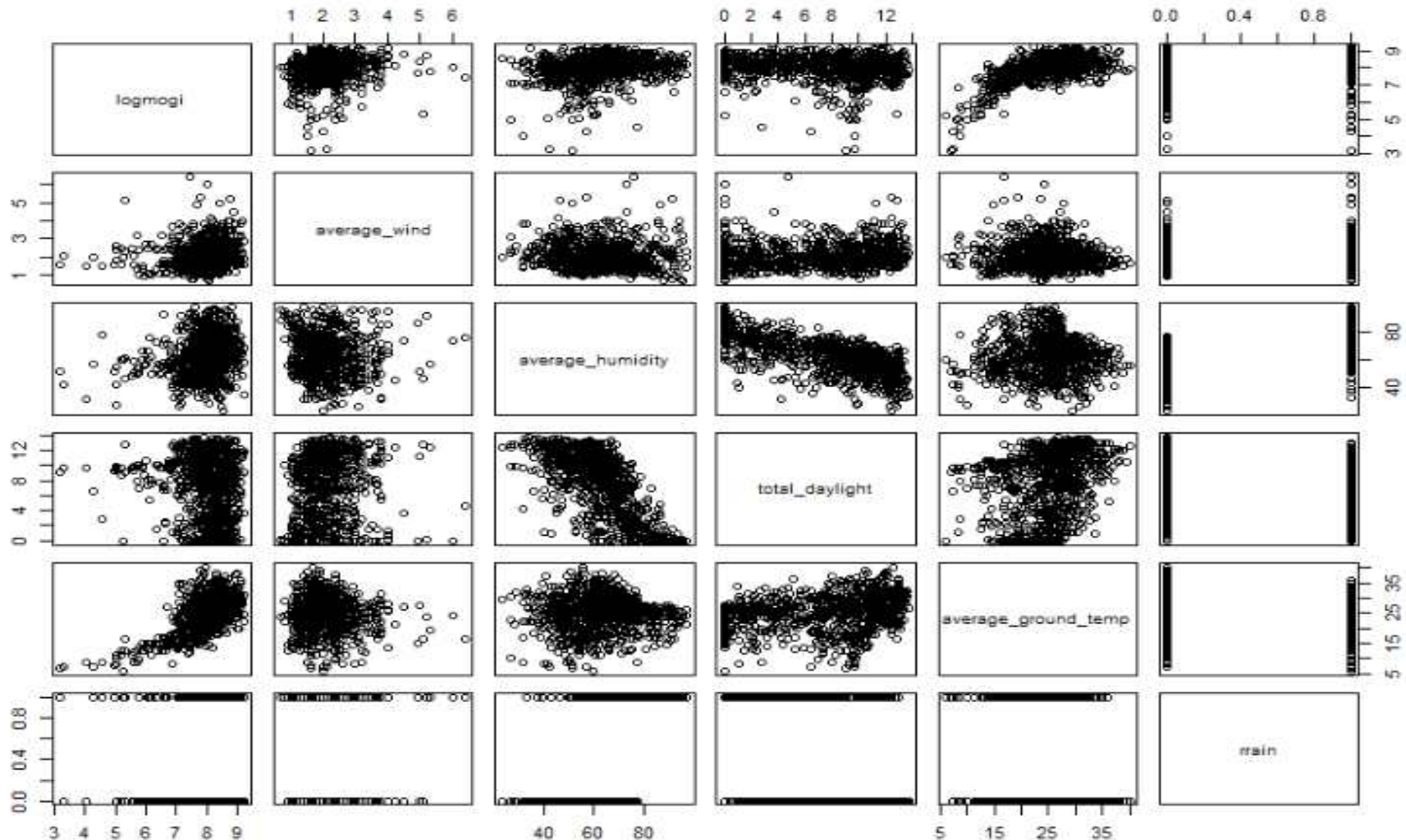
X1: average_wind

X2: average_humidity

X3: total_daylight

X4: average_ground_temp

2-2. 회귀식 적합 - 로그 변환 이후 산점도 행렬



2-2. 회귀식 적합 - 이후 R output

```
Call:
lm(formula = logmog ~ ., data = data13)

Residuals:
    Min       1Q   Median       3Q      Max
-0.89928 -0.26824  0.01665  0.27885  0.79808

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    6.777139   0.166580  40.684 < 2e-16 ***
dummy          -1.641653   0.158895 -10.332 < 2e-16 ***
average_wind     0.165857   0.019931   8.322 3.63e-16 ***
average_humidity  0.006113   0.001559   3.921 9.56e-05 ***
total_daylight  -0.022383   0.005023  -4.456 9.50e-06 ***
average_ground_temp 0.034308   0.004444   7.720 3.39e-14 ***
rrain           -0.087889   0.037337  -2.354  0.0188 *
d                0.059007   0.006619   8.914 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3836 on 817 degrees of freedom
Multiple R-squared:  0.5732,    Adjusted R-squared:  0.5695
F-statistic: 156.7 on 7 and 817 DF, p-value: < 2.2e-16
```