

# 2021\_2학기 회귀분석 기말프로젝트

성인근로자의 근무형태 · 수면시간이

건강 관련 삶의 지수에 미치는 영향

- 국민건강영양조사 자료를 이용하여 -

황 건 하



# CONTENTS

1



배경 설명 및 연구방법

2



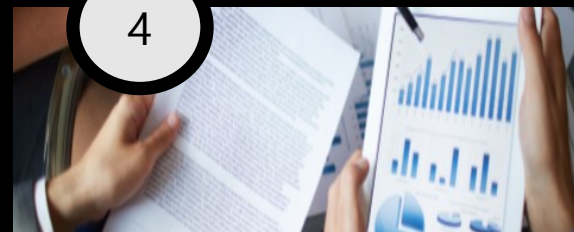
변수 선택 기준 · 선택 방법

3



회귀진단 · 모형확인

4



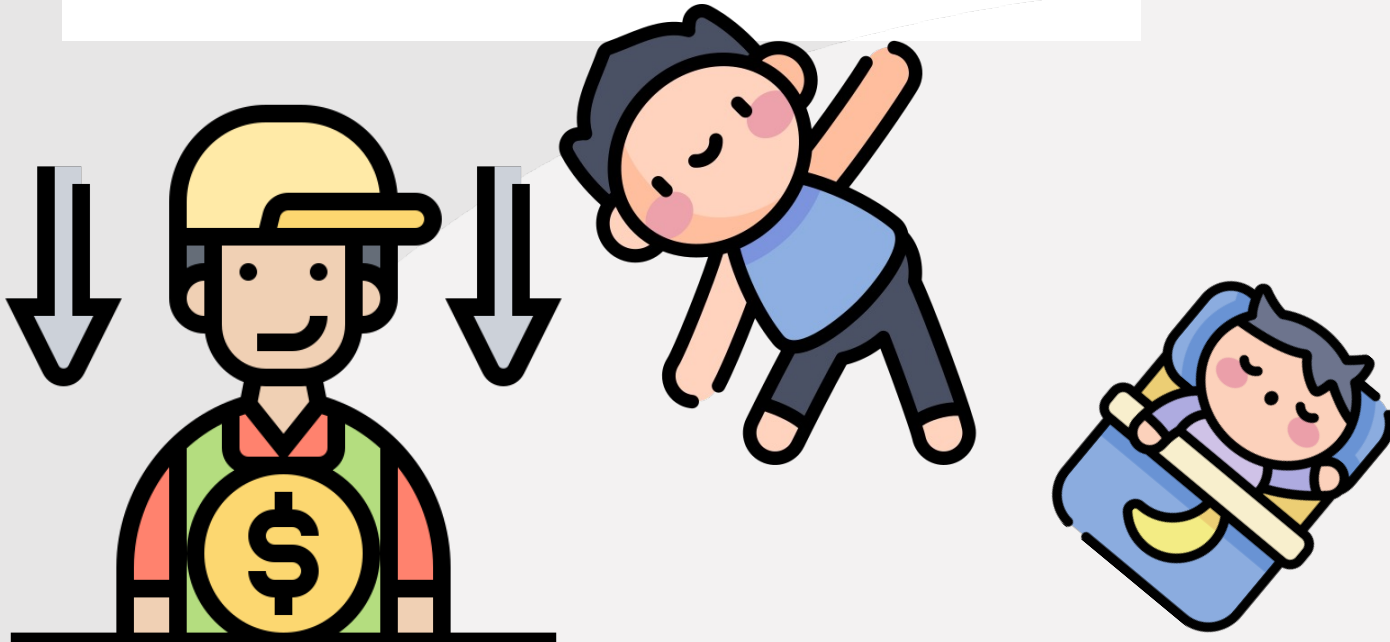
결론 및 정리

## 1. 배경설명 및 연구방법

### 배경

공공서비스 증가 ➡ 노동시간 24시간으로 확대  
야간근로자 증가

불규칙한 생활  
교대근무 ➡ 수면시간 · 운동시간 ?  
삶의 질 · 만성질환  
연구 필요



### 연구방법

R\_dplyr패키지 활용 ➡ 데이터 전처리

중선형회귀모형

데이터 추출 -> 변수 선택 · 기준 -> 회귀진단 ->  
모형 확인  
연구대상

최초 표본수 8,110

19세 미만대상자와 결측 데이터 제외

설문지 데이터가 누락된 개인도 데이터  
제외

표본수 142 (남자: 130 여자: 12)



표본수 135 (남자: 123 여자: 12)

# 종속변수

출처:한국보건경제정책학회

## HINT-8 지수 산출식

**HINT-8 지수 = 1-(0.073**  
**건강 관련 삶의 지수**

$$\begin{aligned} &+ 0.018 \times \text{CL2} + 0.072 \times \text{CL3} + 0.122 \times \text{CL4} \\ &+ 0.055 \times \text{PA2} + 0.116 \times \text{PA3} + 0.188 \times \text{PA4} \\ &+ 0.019 \times \text{VI23} + 0.070 \times \text{VI4} \\ &+ 0.004 \times \text{WO2} + 0.028 \times \text{WO3} + 0.036 \times \text{WO4} \\ &+ 0.012 \times \text{DE2} + 0.044 \times \text{DE3} + 0.098 \times \text{DE4} \\ &+ 0.014 \times \text{ME2} + 0.058 \times \text{ME3} + 0.109 \times \text{ME4} \\ &+ 0.020 \times \text{SL3} + 0.090 \times \text{SL4} \\ &+ 0.014 \times \text{HA2} + 0.068 \times \text{HA3} + 0.082 \times \text{HA4}) \end{aligned}$$

## 산출식 설명

항목	항목별 수준
CL: 계단 오르기	2: 수준 2인 경우 1, 나머지 경우 0
PA: 통증	3: 수준 3인 경우 1, 나머지 경우 0
VI: 활력	4: 수준 4인 경우 1, 나머지 경우 0
WO: 일하기	(활력 항목에서 수준 2 또는 3인 경우 VI23=1이고 나머지 경우 0이 됨)
DE: 우울	
ME: 기억하기	
SL: 수면	가장 좋은 상태는 11111111
HA: 행복	가장 나쁜 상태는 44444444

# 독립변수

\*결측값, 88,99,888,999 등 제외

BP16\_1 주중 하루 평균 수면 시간

EC\_wht\_23 주당 평균 근로시간

BE3\_32 걷기 지속 시간

BS3\_2 (성인) 하루평균 흡연량

EC\_wht\_23 주당 평균 근로시간

BE3\_87 여가\_증강도 신체활동 시간

sex 성별

0 -> 여자 , 1 -> 남자

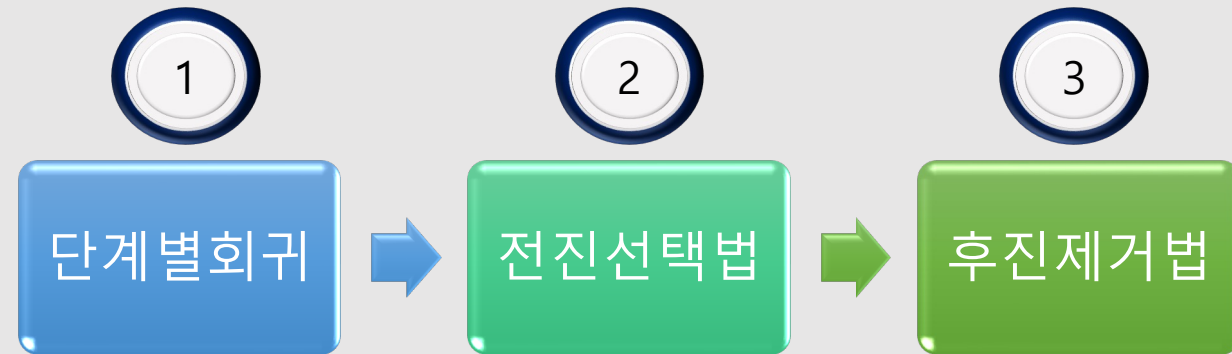
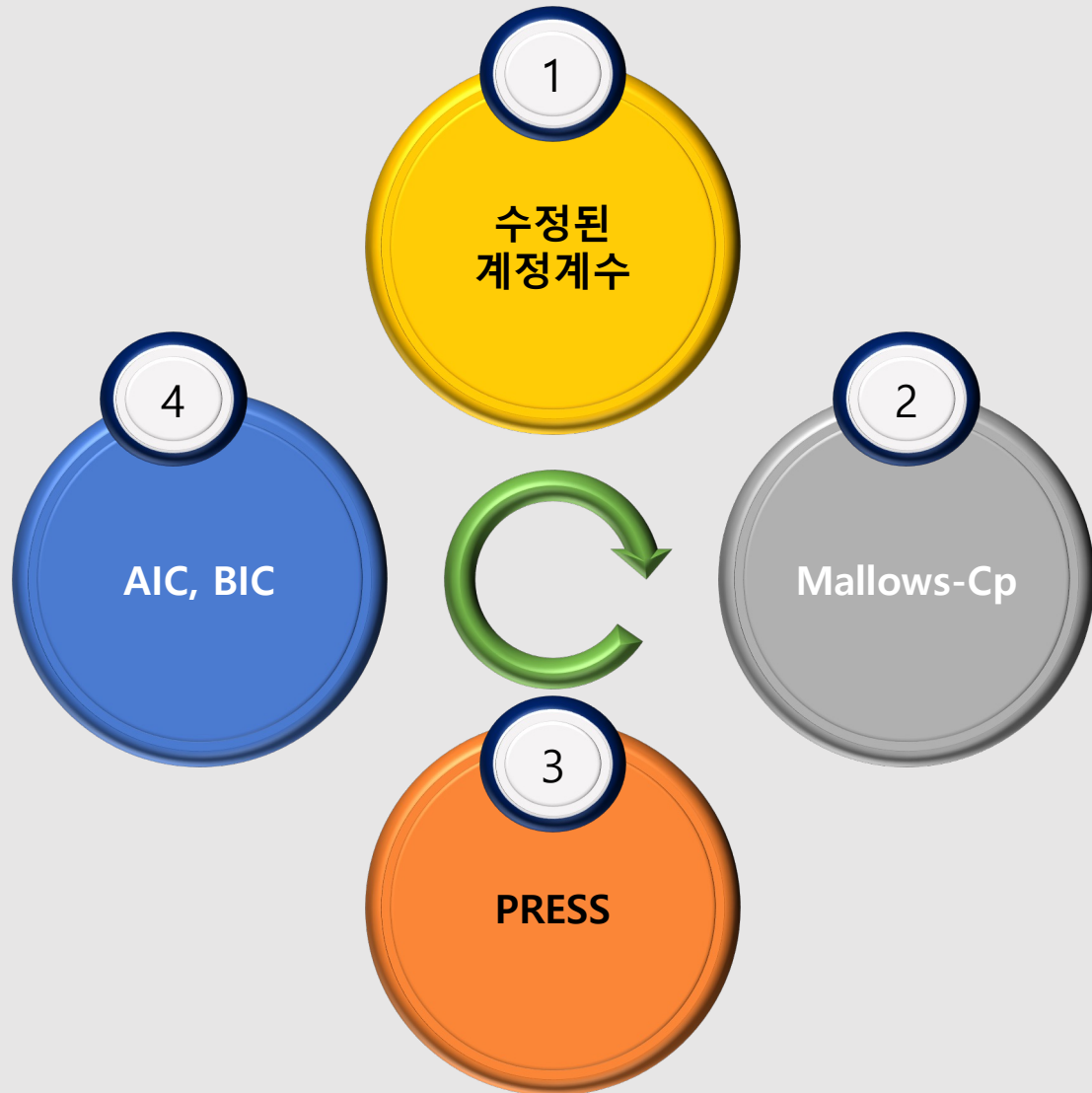
edu 교육수준

EC\_wht\_변형근로시간

0 -> 야간근무, 1 -> 주간+ 저녁근무

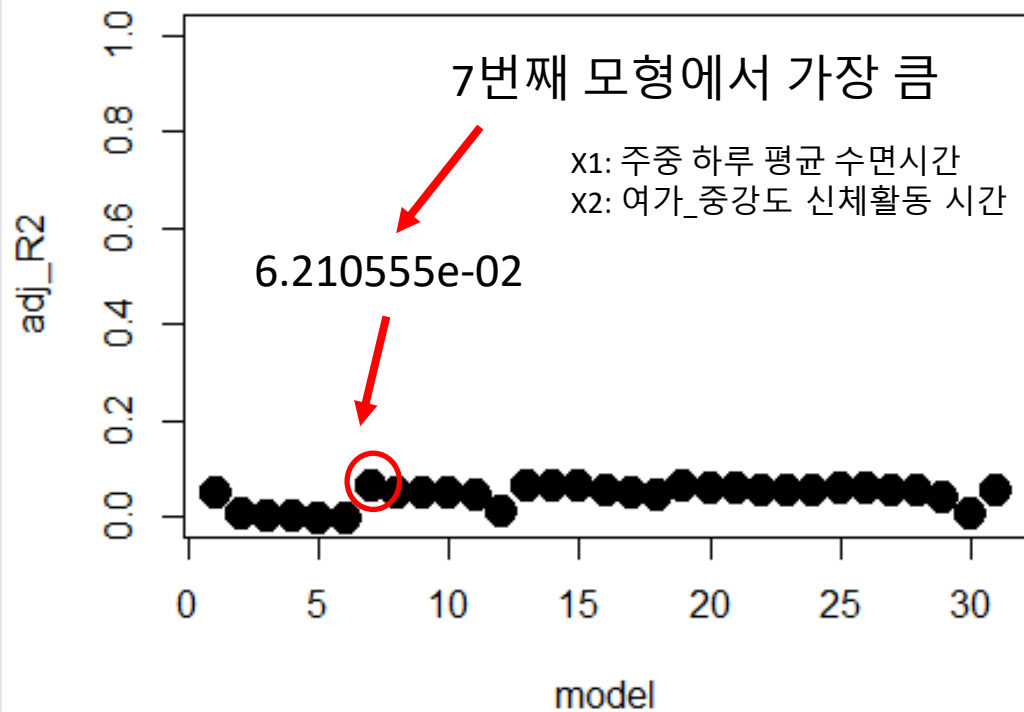
문항번호	변수유형	변수명	변수설명	내용
1-1	N(1)	LQ_1HT	HINT-8: 계단오르기	1. 나는 계단을 오르는 데 어려움이 전혀 없었다. 2. 나는 계단을 오르는 데 어려움이 약간 있었다. 3. 나는 계단을 오르는 데 어려움이 많이 있었다. 4. 나는 계단을 오를 수 없었다. 8. 비해당( 청소년, 소아) 9. 모름, 무응답
1-2	N(1)	LQ_2HT	HINT-8: 통증	1. 나는 전혀 통증이 없었다. 2. 나는 약한 통증이 있었다. 3. 나는 심한 통증이 있었다. 4. 나는 극심한 통증이 있었다. 8. 비해당( 청소년, 소아) 9. 모름, 무응답
1-3	N(1)	LQ_3HT	HINT-8: 기운	1. 나는 항상 기운이 있었다. 2. 나는 자주 기운이 있었다. 3. 나는 가끔 기운이 있었다. 4. 나는 전혀 기운이 없었다. 8. 비해당( 청소년, 소아) 9. 모름, 무응답
1-4	N(1)	LQ_4HT	HINT-8: 일하기	1. 나는 일을 하는 데 어려움이 전혀 없었다. 2. 나는 일을 하는 데 어려움이 약간 있었다. 3. 나는 일을 하는 데 어려움이 많이 있었다. 4. 나는 일을 할 수 없었다. 8. 비해당( 청소년, 소아) 9. 모름, 무응답
1-5	N(1)	LQ_5HT	HINT-8: 우울	1. 나는 전혀 우울하지 않았다. 2. 나는 가끔 우울했다. 3. 나는 자주 우울했다. 4. 나는 항상 우울했다. 8. 비해당( 청소년, 소아) 9. 모름, 무응답
1-6	N(1)	LQ_6HT	HINT-8: 기억	1. 나는 기억하는 데 어려움이 전혀 없었다. 2. 나는 기억하는 데 어려움이 약간 있었다. 3. 나는 기억하는 데 어려움이 많이 있었다. 4. 나는 전혀 기억을 할 수 없었다. 8. 비해당( 청소년, 소아) 9. 모름, 무응답
				1. 나는 잠을 자는 데 어려움이 전혀 없었다. 

## 2. 변수선택 기준·선택방법





## ① 수정된 결정계수와 MSE



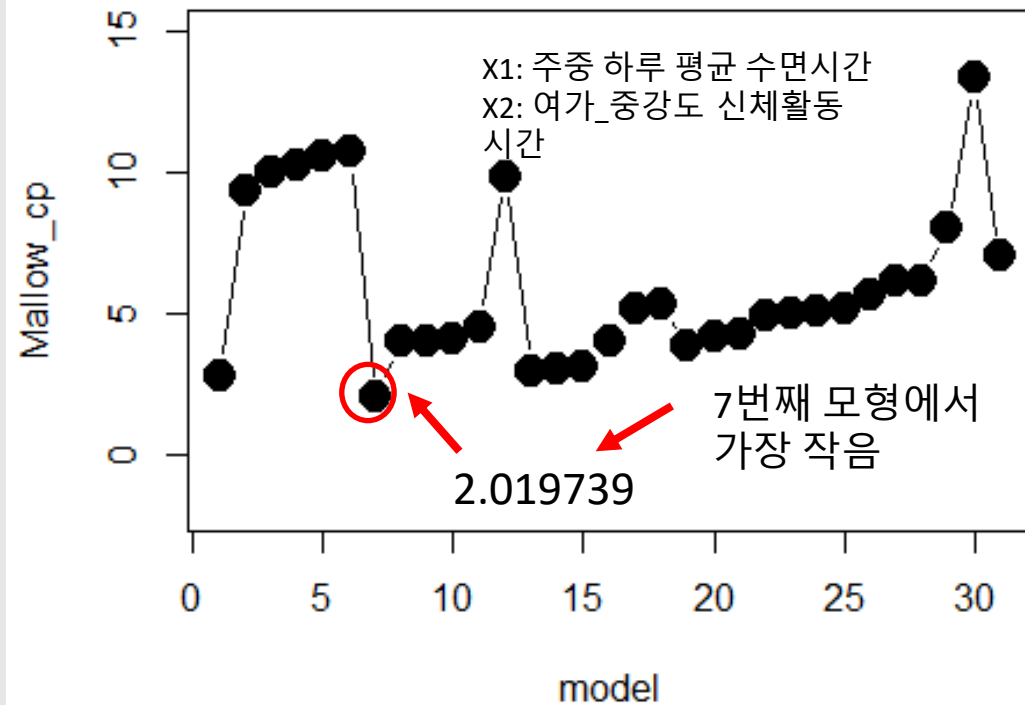
부분 F검정  
P-value : 0.0043 \*\*



귀무가설 기각

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

## ② Mallows-Cp



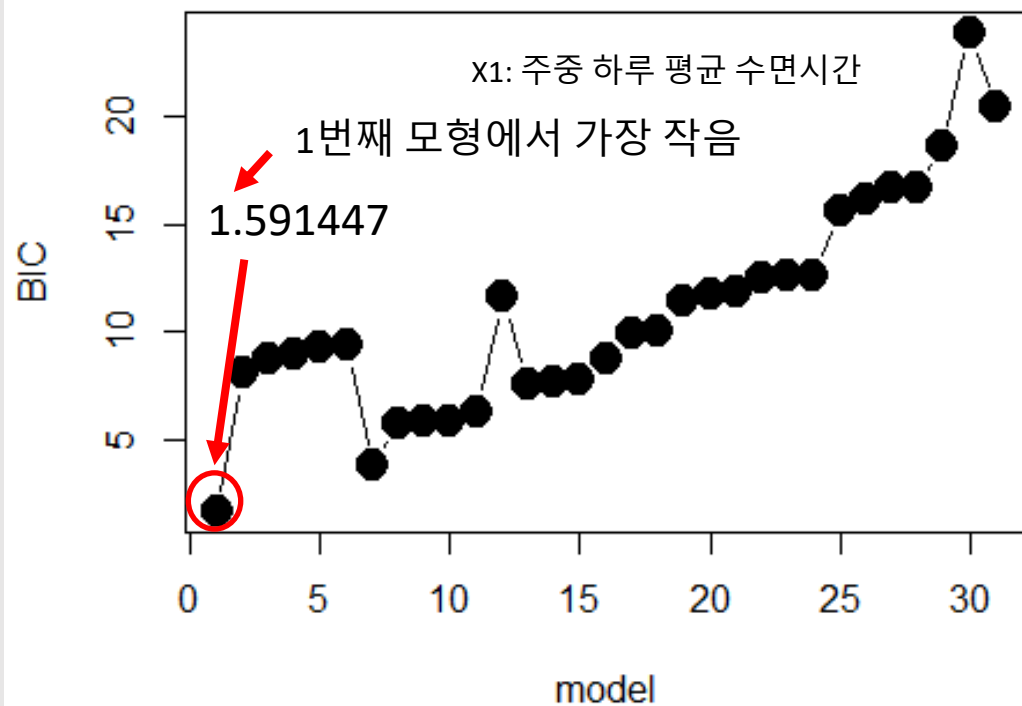
부분 F검정  
P-value : 0.0043 \*\*



귀무가설 기각

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

### ③ BIC



부분 F검정  
P-value : 0.00425 \*\*

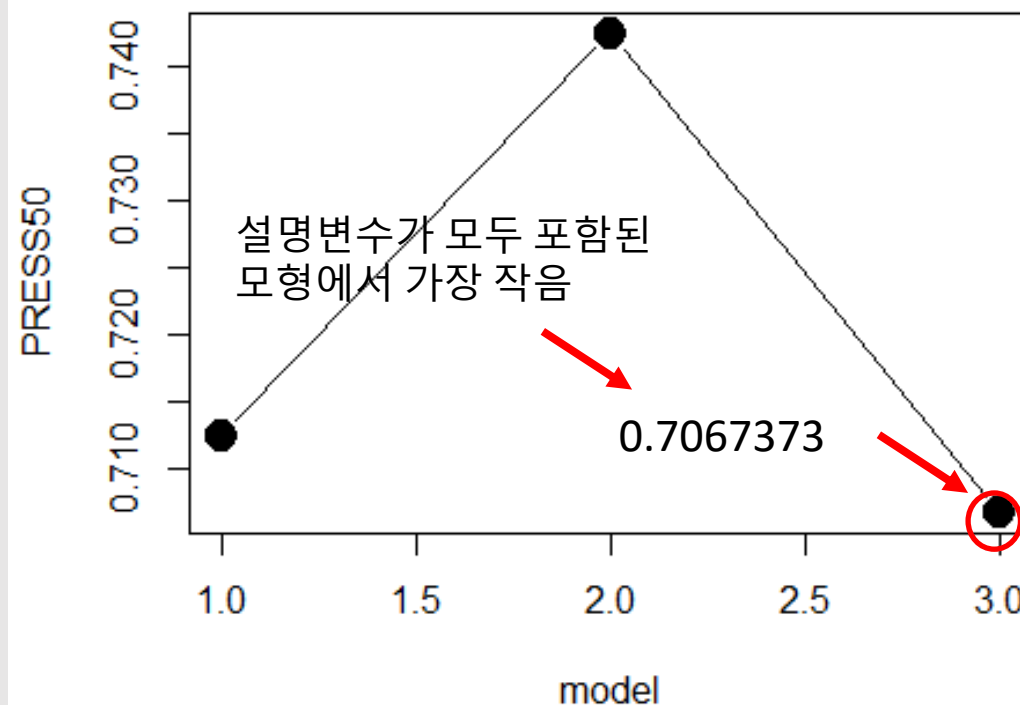


귀무가설 기각

$$Y = \beta_0 + \beta_1 x_1 + \varepsilon$$

### ④ PRESS\_p

x1: 주중 하루 평균 수면시간  
x2: 여가\_중강도 신체활동 시간



부분 F검정  
P-value : 0.0043 \*\*

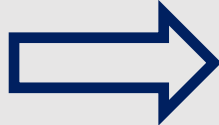


귀무가설 기각

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$



단계별 회귀 

전진 선택법 

후진제거법

```
lm(formula = HINT_8 ~ BP16_1 + BE3_87, data = health_data)
```

Coefficients:

(Intercept)	BP16_1	BE3_87
0.72635	0.01506	0.01282

Call:

```
lm(formula = HINT_8 ~ BP16_1 + BE3_87, data = health_data)
```

Coefficients:

(Intercept)	BP16_1	BE3_87
0.72635	0.01506	0.01282

Call:

```
lm(formula = HINT_8 ~ BP16_1 + BE3_87, data = health_data)
```

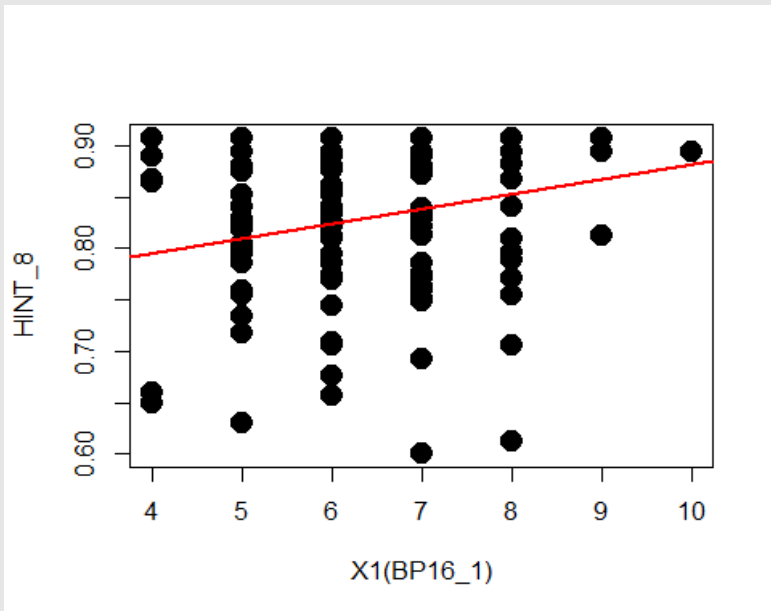
Coefficients:

(Intercept)	BP16_1	BE3_87
0.72635	0.01506	0.01282

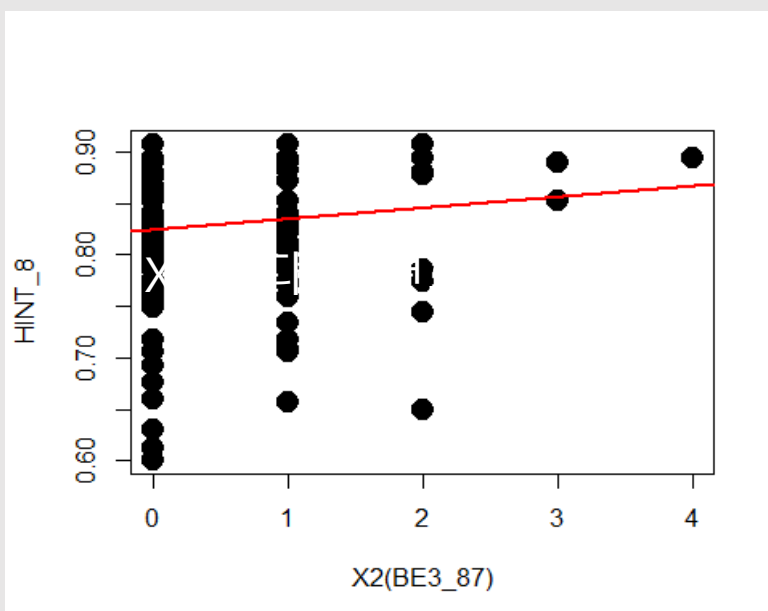
### 3. 회귀진단. 모형확인

편회귀그림

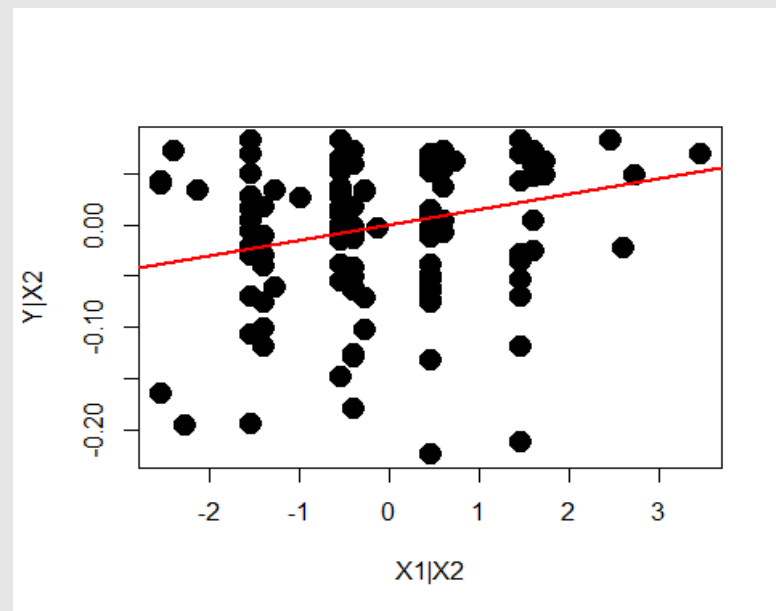
X1만을 포함한 모형의  
적합



X2만을 포함한 모형의  
적합



X1에 대한 편회귀그림



1.편회귀그림

다른 설명변수들의 영향이 제거된

Y와 Xj의 순수한 관계를 보여주는 그림

Pearson's product-moment correlation

```
data: x1.x2 and y.x2
t = 3.0743, df = 140, p-value = 0.002538
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.09049386 0.39964473
sample estimates:
cor
0.2514725
```

X1에 대한 편회귀상관계수

X1에 대한 회귀 기울기

```
Call:
lm(formula = y.x2 ~ x1.x2)

Coefficients:
(Intercept)          x1.x2
-4.575e-18         1.506e-02
```

### 3. 회귀진단. 모형확인

적합결여검정

가설:H0: 가정된 모형  $E[Y_i] = B_0 + \beta_1 X_{1i}$  VS H1: 가정된 모형은 옳지 않다

```
Model 1: Y ~ X1
Model 2: Y ~ factor(X1)
  Res.Df    RSS Df Sum of Sq    F Pr(>F)
1     140 0.69117
2     135 0.68999  5 0.0011767 0.046 0.9987
```

귀무가설을  
기각하지 못함



HINT\_8지수에 대한  
주중수면시간이 있는  
모형이 적절

### 2. 적합결여검정

가설:H0: 가정된 모형  $E[Y_i] = \mathbf{x}^T \boldsymbol{\beta}$  VS H1: 가정된 모형은 옳지 않다

가설:H0: 가정된 모형  $E[Y_i] = B_0 + \beta_1 X_{1i} + \beta_2 X_{2i}$  VS H1: 가정된 모형은 옳지 않다

```
Model 1: Y ~ X1 + X2
Model 2: Y ~ factor(X1) + factor(X2)
  Res.Df    RSS Df Sum of Sq    F Pr(>F)
1     139 0.67761
2     131 0.67184  8 0.0057668 0.1406 0.9972
```

귀무가설을  
기각하지 못함



HINT\_8지수에 대한 주중수면시간과  
여가 중강도 신체활동에 대한  
hint8지수가 있는 모형이 적절

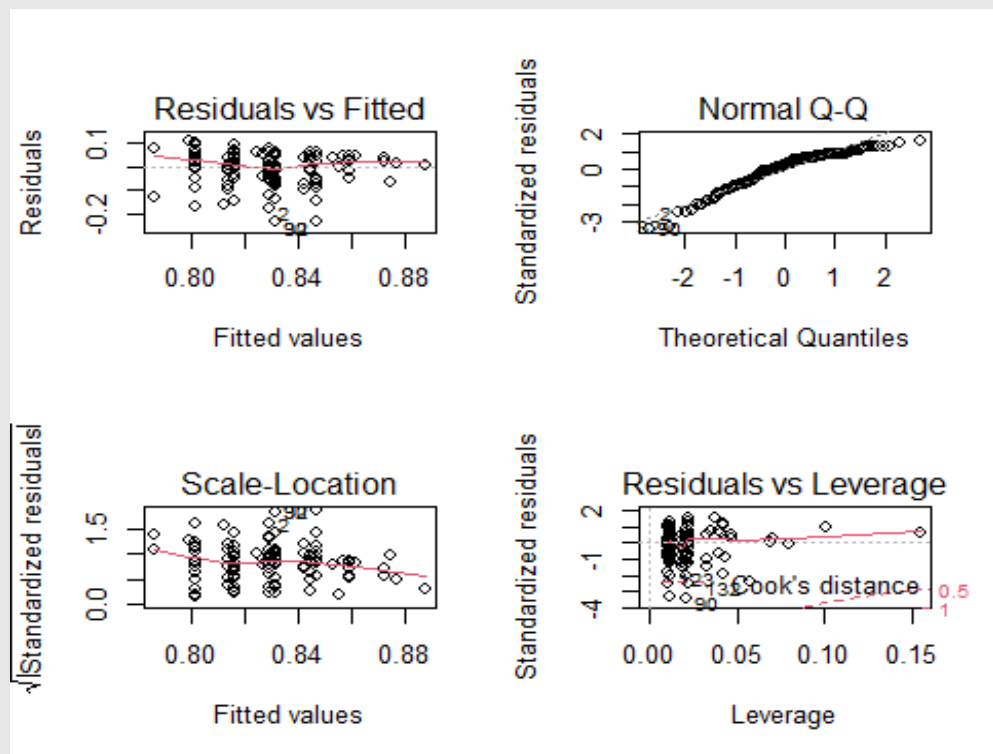
# 3. 회귀진단. 모형확인

변수변환

## 3. 변수변환

분산안정화 변환 / 정규성을 위한  
변환

### ① 변환하지 않는 모형



Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.726351	0.032935	22.054	< 2e-16 ***
x1	0.015063	0.004917	3.063	0.00263 **
x2	0.012821	0.007688	1.668	0.09763 .

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

Residual standard error: 0.06982 on 139 degrees of freedom  
Multiple R-squared: 0.07541, Adjusted R-squared: 0.06211  
F-statistic: 5.668 on 2 and 139 DF, p-value: 0.0043

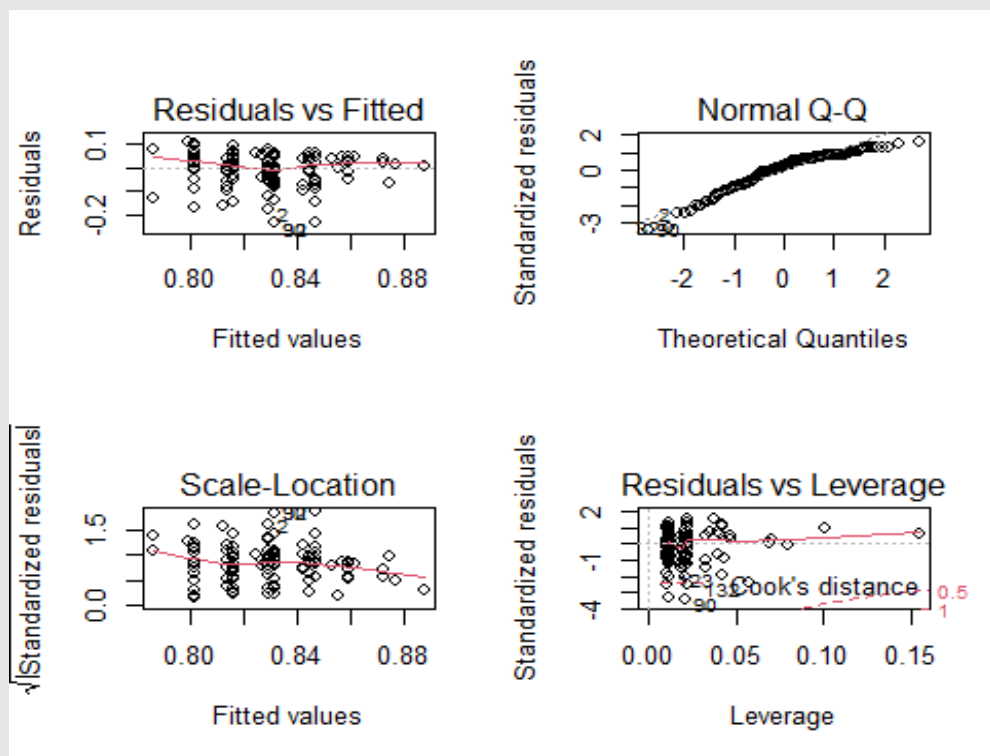
잔차분석한 결과 ➡ 변환전과 후가 크게 차이가 없음

그러므로 변환하지 않는 모형을 최종모형으로 선택

# 3. 회귀진단. 모형확인

변수변환

## ② logY 변환



## 3. 변수변환

분산안정화 변환 / 정규성을 위한 변환

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.726351	0.032935	22.054	< 2e-16 ***
x1	0.015063	0.004917	3.063	0.00263 **
x2	0.012821	0.007688	1.668	0.09763 .

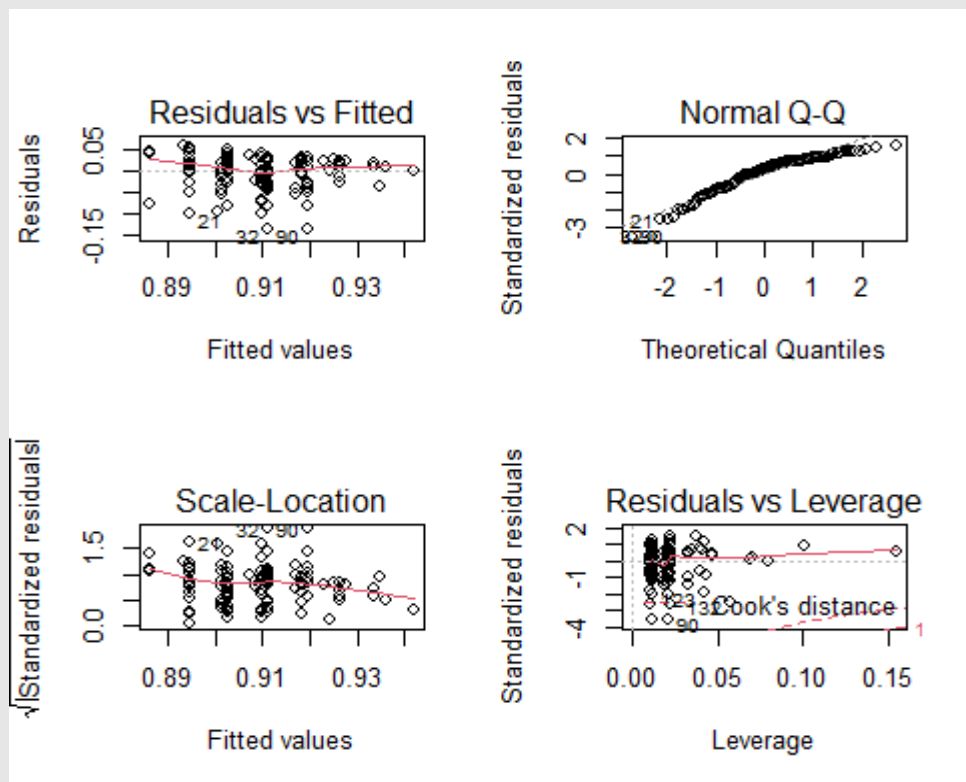
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.06982 on 139 degrees of freedom  
Multiple R-squared: 0.07541, Adjusted R-squared: 0.06211  
F-statistic: 5.668 on 2 and 139 DF, p-value: 0.0043

# 3. 회귀진단. 모형확인

변수변환

③  $\sqrt[3]{Y}$  변환



## 3. 변수변환

분산안정화 변환 / 정규성을 위한  
변환

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.852857	0.018572	45.922	< 2e-16 ***
x1	0.008324	0.002773	3.002	0.00318 **
x2	0.007097	0.004335	1.637	0.10389

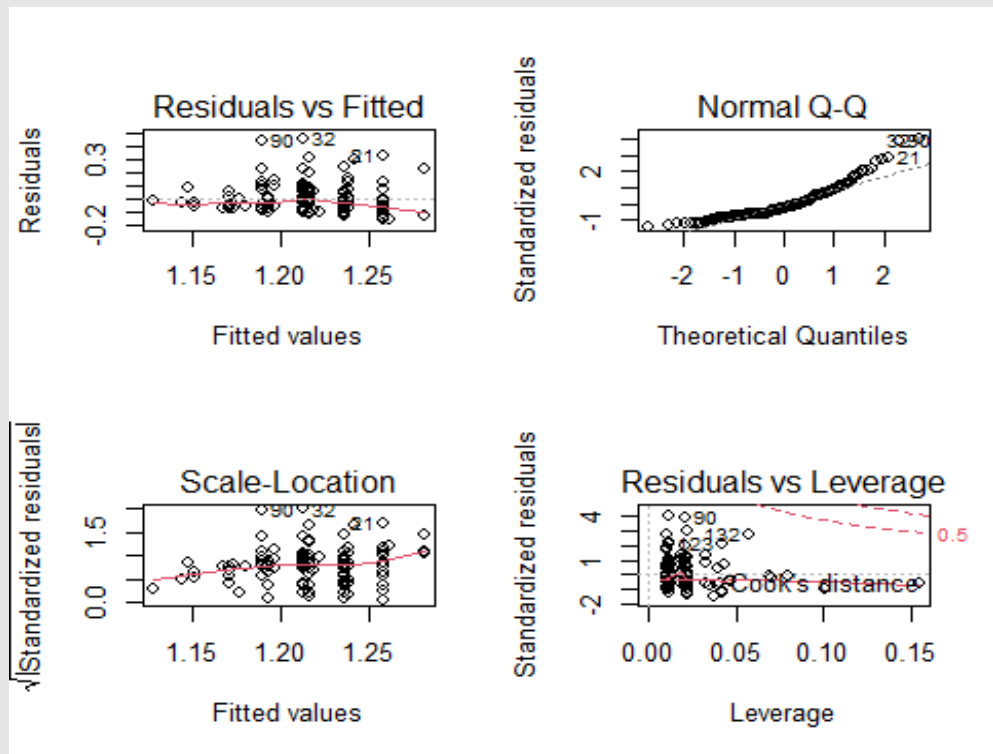
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.03937 on 139 degrees of freedom  
Multiple R-squared: 0.07269, Adjusted R-squared: 0.05935  
F-statistic: 5.448 on 2 and 139 DF, p-value: 0.005274

# 3. 회귀진단. 모형확인

변수변환

## ④ 1/Y 변환



## 3. 변수변환

분산안정화 변환 / 정규성을 위한  
변환

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	1.371145	0.054284	25.259	< 2e-16 ***
x1	-0.022681	0.008105	-2.798	0.00587 **
x2	-0.019522	0.012671	-1.541	0.12568

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

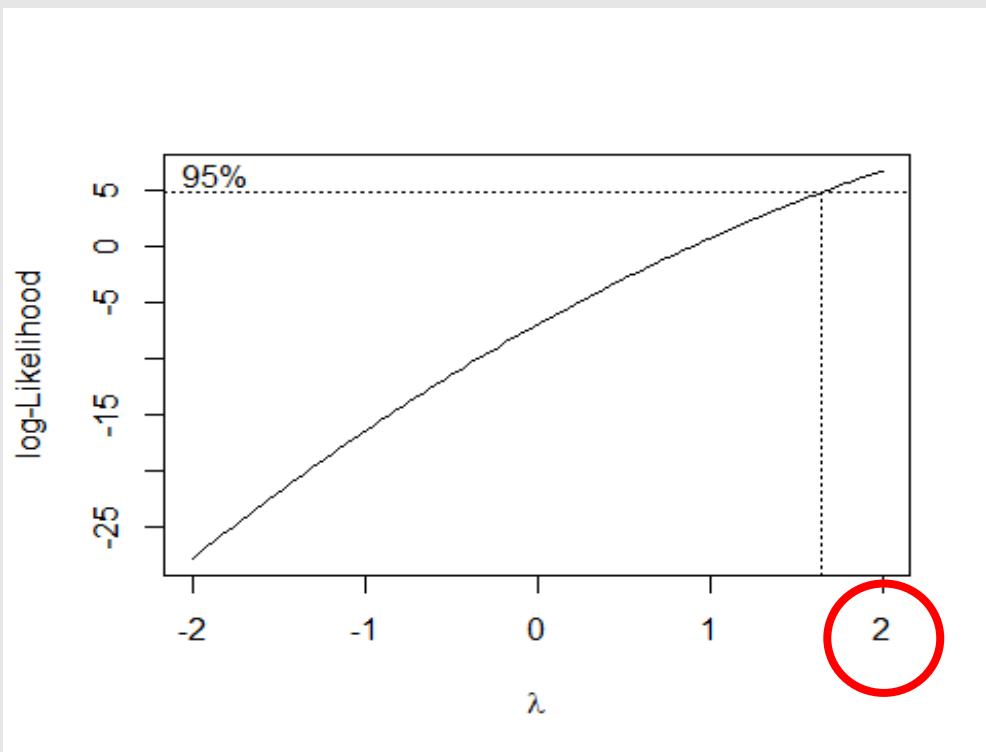
Residual standard error: 0.1151 on 139 degrees of freedom  
Multiple R-squared: 0.06401, Adjusted R-squared: 0.05054  
F-statistic: 4.753 on 2 and 139 DF, p-value: 0.01008



# 3. 회귀진단. 모형확인

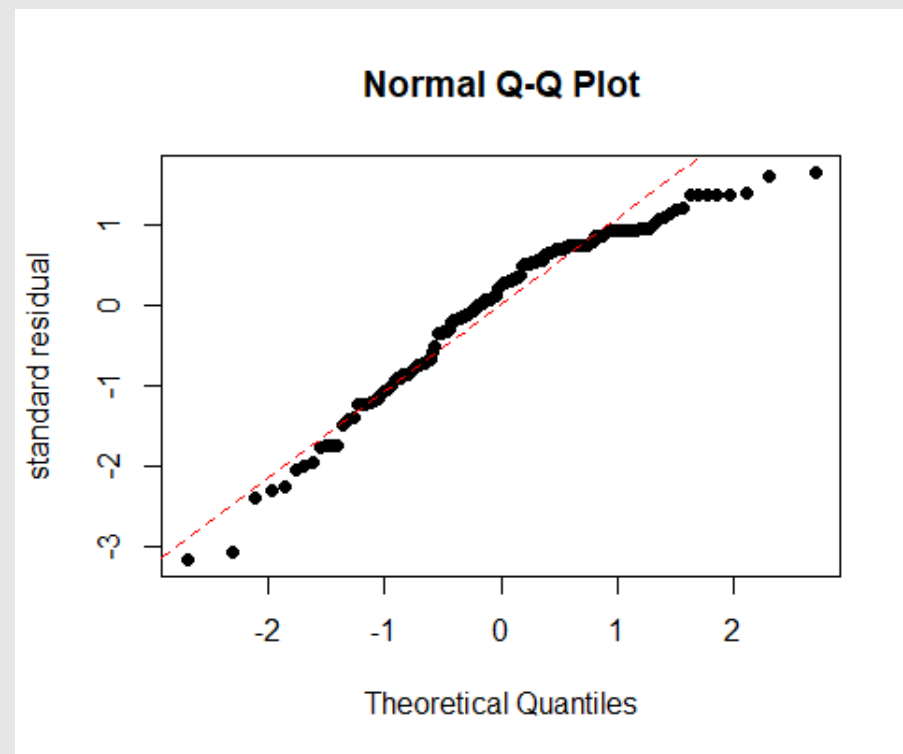
변수변환

## 5 Box-Cox 멱변환 2



## 3. 변수변환

분산안정화 변환 / 정규성을 위한 변환



잔차분석한 결과 ➡ 변환전과 후가 크게 차이가 없음  
그러므로 변환하지 않는 모형을 최종모형으로 선택

### 3. 회귀진단. 모형확인

다중공선성

### 4. 다중공선성

VIF  $\geq 5$ , 의심 / VIF  $\geq 10$ , 가능성  
있음

VIF

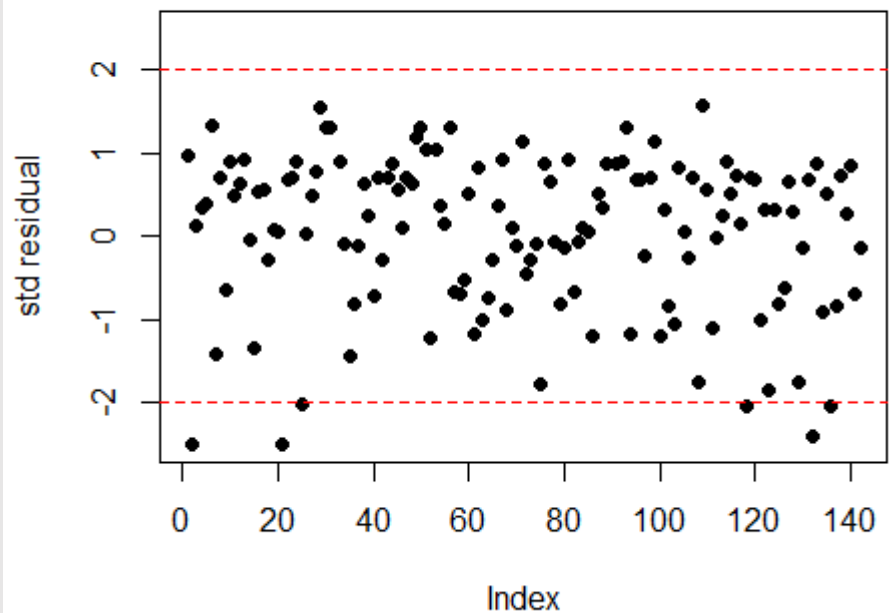
```
> vif(model12)
      x1      x2
1.008172 1.008172
```

다중공선성 X

### 3. 회귀진단· 모형확인

잔차분석 · 지렛값

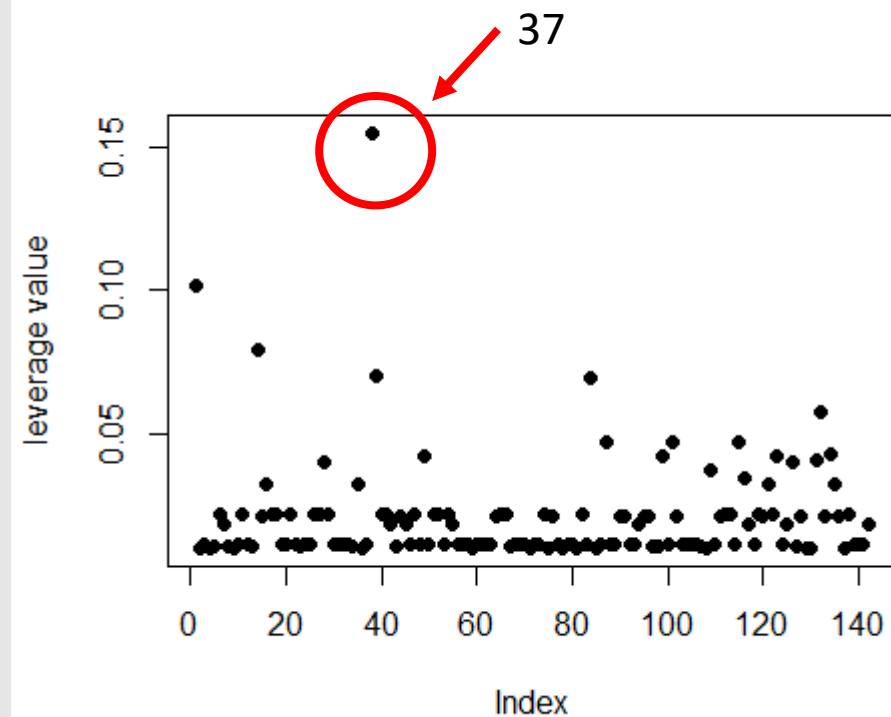
#### 잔차분석



5.잔차분석 · 지렛값

지렛값 기준:  $2p / n$

#### 지렛값



### 3. 회귀진단. 모형확인

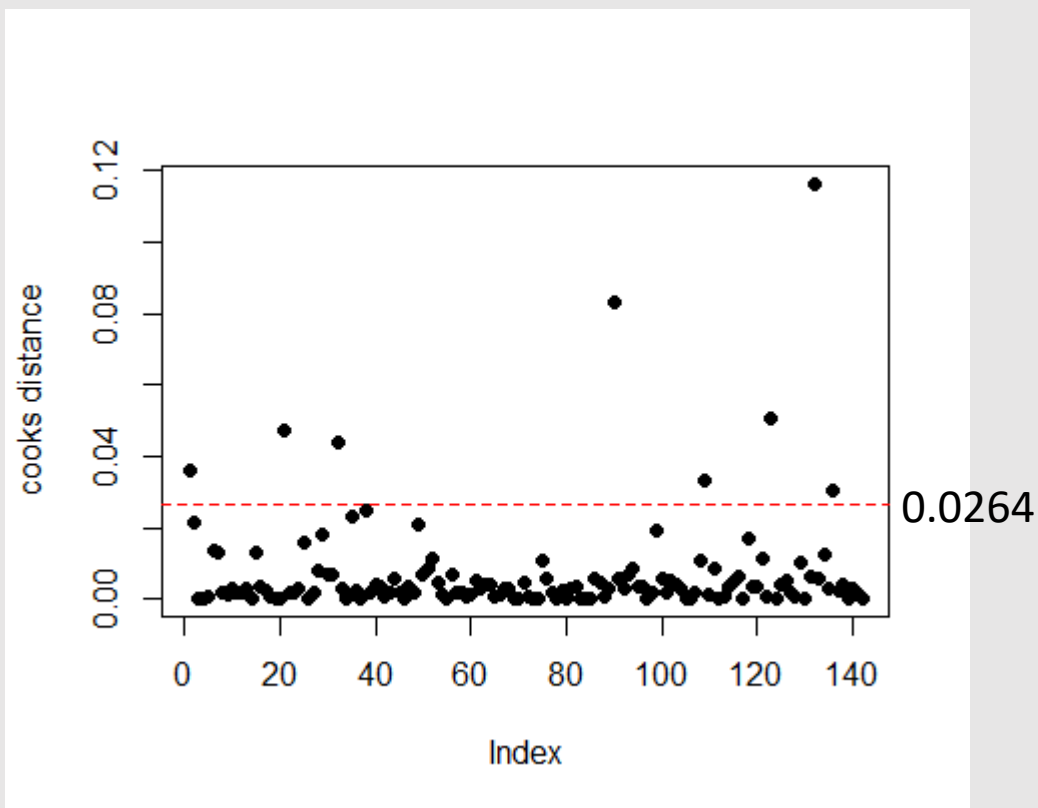
영향력 있는 측도(제거)

### 6.영향력 측도

Cook's distance  $\rightarrow 3.67 / n$

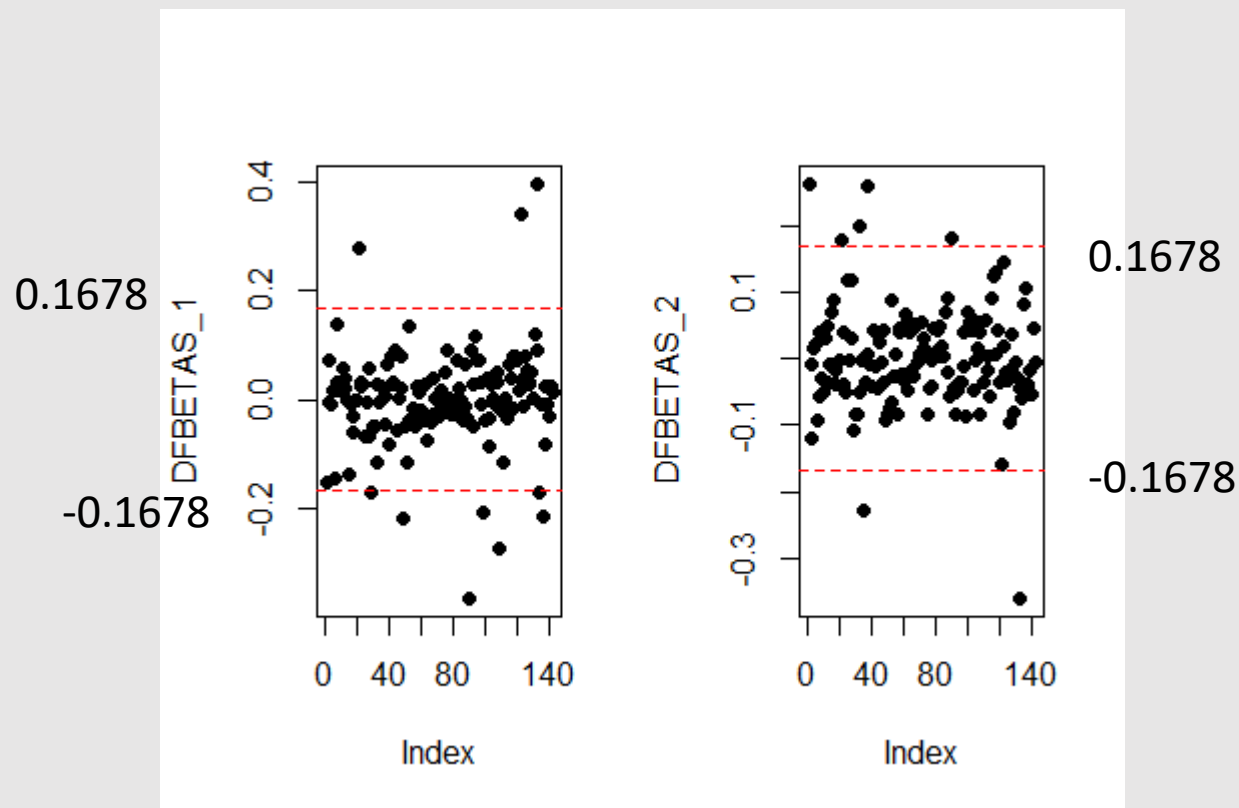
DFBETAS  $\rightarrow 2\sqrt{n}$

① Cook's distance  $\Rightarrow \beta^{\wedge}$  에 대한 영향력 측정



이상치(8개):  
1, 21, 38, 90, 109, 123, 132, 136

② DFBETAS  $\Rightarrow \beta^{\wedge}_i$  에 대한 영향력 측정

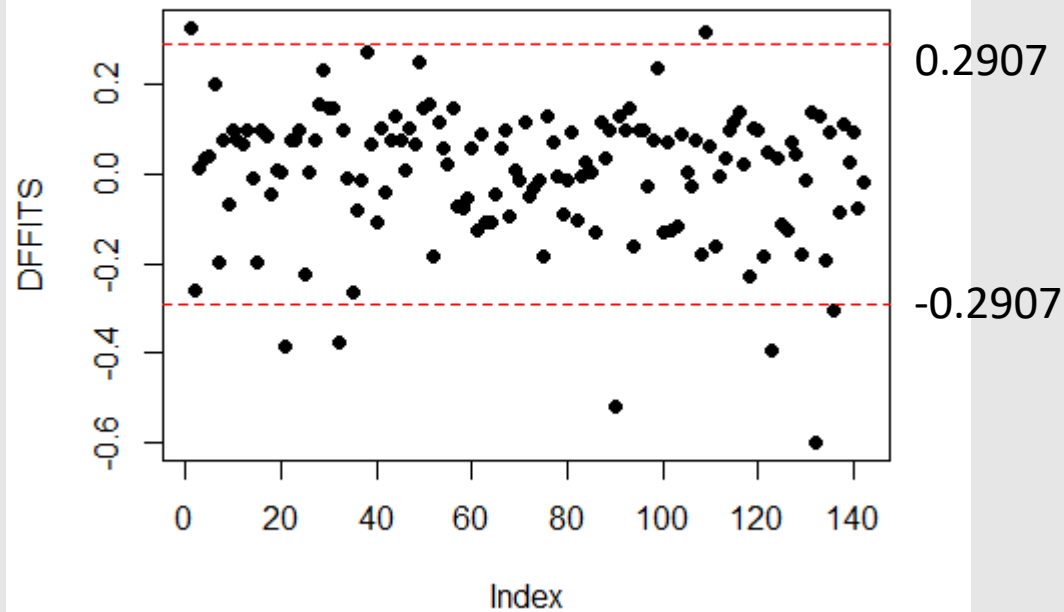


BETA1이상치(8개): 21, 38, 49, 90, 99, 109, 132, 136  
BETA2이상치(7개): 1, 21, 32, 35, 38, 90, 132

### 3. 회귀진단. 모형확인

영향력 있는 측도(제거)

③ DFFITS  $\Rightarrow Y^i$  에 대한 영향력 측정



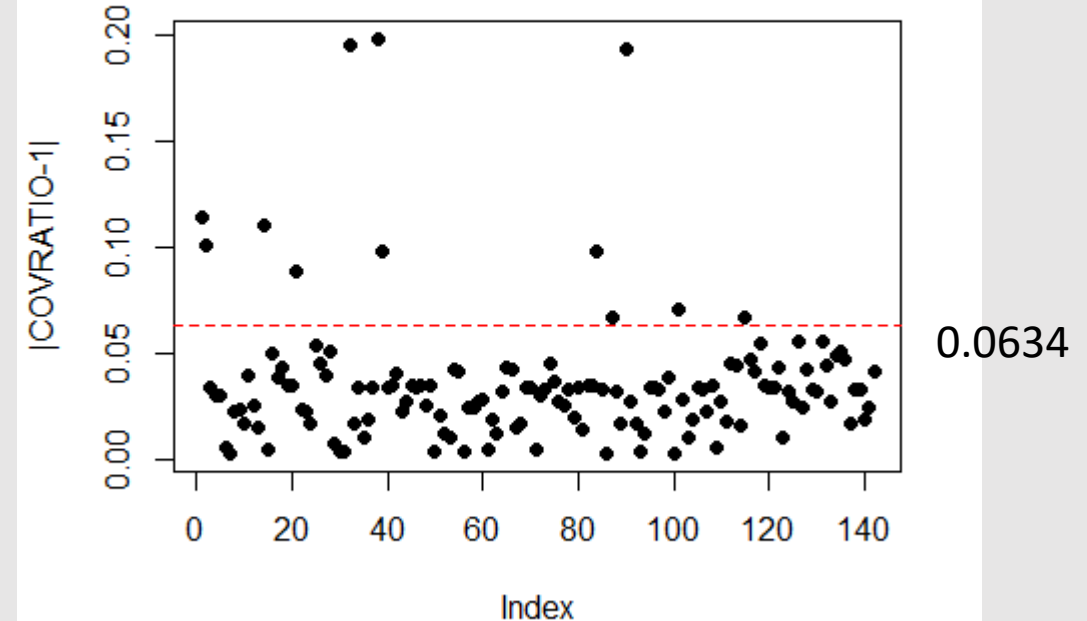
이상치(8개):  
1, 21, 32, 90, 109, 123, 132, 136

6.영향력 측도

DFFITS  $\rightarrow 2\sqrt{p/n}$

COVRATIO  $\rightarrow |COVRATIO-1| \geq 3p/n$

④ COVRATIO  $\Rightarrow \text{cov}(\beta^i)$ 에 대한 영향력 측정



이상치(12개):  
1, 2, 17, 21, 32, 35, 81, 86, 90, 93, 100, 118

### 3. 회귀진단. 모형확인

더빈-왓슨 검정

7.더빈-왓슨 검정

$D \sim 2(1-p^{\wedge})$   
독립 만족

$0 < D < 4$

$D \sim 2$ 이면, 오차항의

```
> dwtest(Final_model)
```

```
Durbin-Watson test
```

```
data: Final_model
```

```
DW = 2.4102, p-value = 0.9929
```

```
alternative hypothesis: true autocorrelation is greater than 0
```

귀무가설이 기각되지 못하므로 오차항은 서로 독립

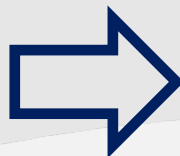
### 3. 회귀진단. 모형확인

최종 모형 확인

표본수 142 (남자: 130 여자: 12)



표본수 134 (남자: 122 여자: 12)



$$Y = B_0 + \beta_1 X_1 + B_2 x_2 + \varepsilon$$

### 8. 모형확인

train\_data -> 70%, test\_data -> 30%

PRESS ~ SSE ,  $R^2_{\text{predict}} \sim R^2$

```
> PRESS_Final$stat #PRESS
[1] 0.5271493
> SST <- sum((health_data$HINT_8-mean(health_data$HINT_8))^2)
> SSE <- sum(resid(Final_model)^2) #SSE
> SSR <- SST-SSE
> SSE
[1] 0.5072338
> 1-(SSE/SST) #R2
[1] 0.07910797
> 1-(PRESS_Final$stat/SST) #R2_predict
[1] 0.042951
```

PRESS ~ SSE

$R^2_{\text{predict}} \sim R^2$

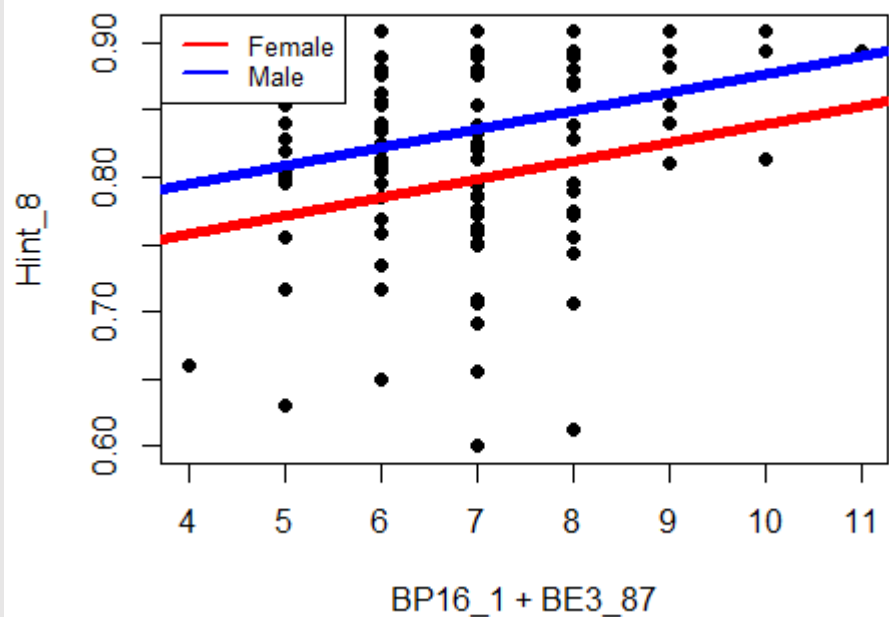


### 3. 변수형 설명변수 추가

변수형 설명변수

9. 설명변수(남녀, 근무형태, 교육수준)  
추가

성별과 HINT\_8지수의  
차이



Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.702300	0.036388	19.300	< 2e-16 ***
BP16_1	0.013702	0.005006	2.737	0.00706 **
BE3_87	0.010141	0.008059	1.258	0.21055
factor(sex)1	0.037125	0.021058	1.763	0.08023 .

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

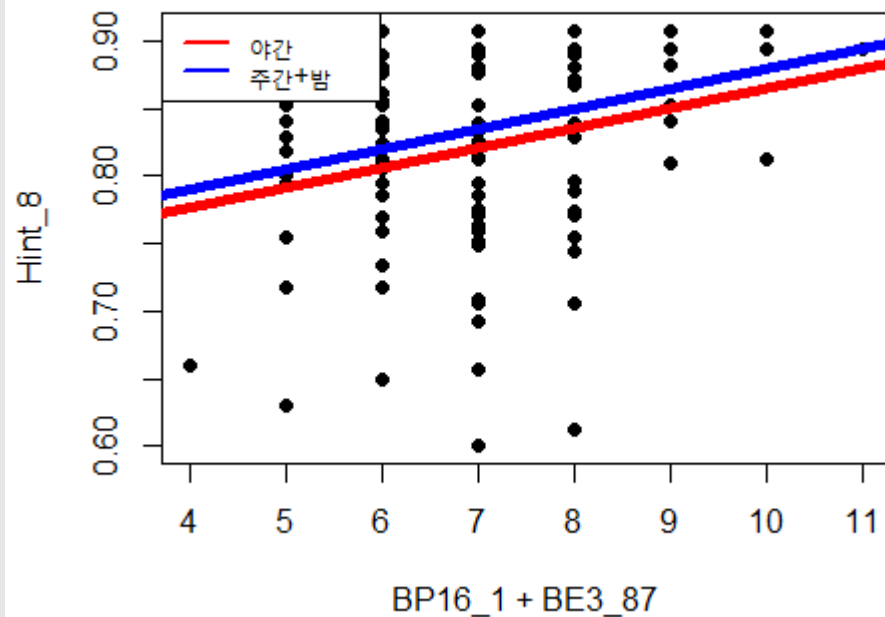
Residual standard error: 0.0691 on 131 degrees of freedom  
Multiple R-squared: 0.0928, Adjusted R-squared: 0.07202  
F-statistic: 4.467 on 3 and 131 DF, p-value: 0.005075

### 3. 변수형 설명변수 추가

변수형 설명변수

9. 설명변수(남녀, 근무형태, 교육수준)  
추가

근무형태과 HINT\_8지수의



Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	0.721840	0.051408	14.041	< 2e-16	***
BP16_1	0.014762	0.005027	2.936	0.00392	**
BE3_87	0.010230	0.008201	1.247	0.21450	
factor(EC_wht_5)1	0.007551	0.041069	0.184	0.85441	

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

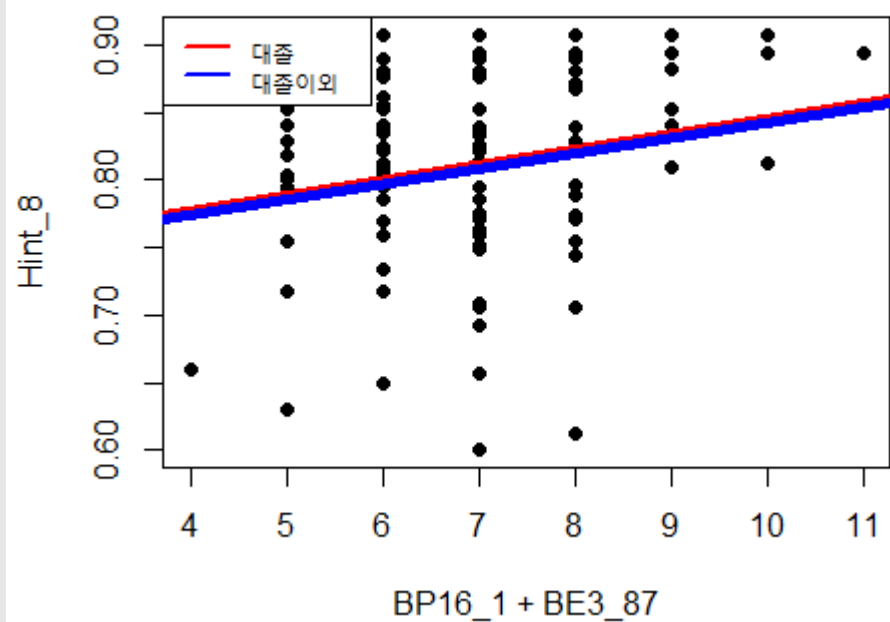
Residual standard error: 0.06991 on 131 degrees of freedom  
Multiple R-squared: 0.07151, Adjusted R-squared: 0.05025  
F-statistic: 3.363 on 3 and 131 DF, p-value: 0.02075

### 3. 변수형 설명변수 추가

변수형 설명변수

9. 설명변수(남녀, 근무형태, 교육수준)  
추가

교육수준과 HINT\_8지수의  
차이



Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.730600	0.034125	21.410	< 2e-16 ***
BP16_1	0.014730	0.005031	2.928	0.00402 **
BE3_87	0.010700	0.008252	1.297	0.19702
factor(edu)1	-0.002898	0.012196	-0.238	0.81253

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

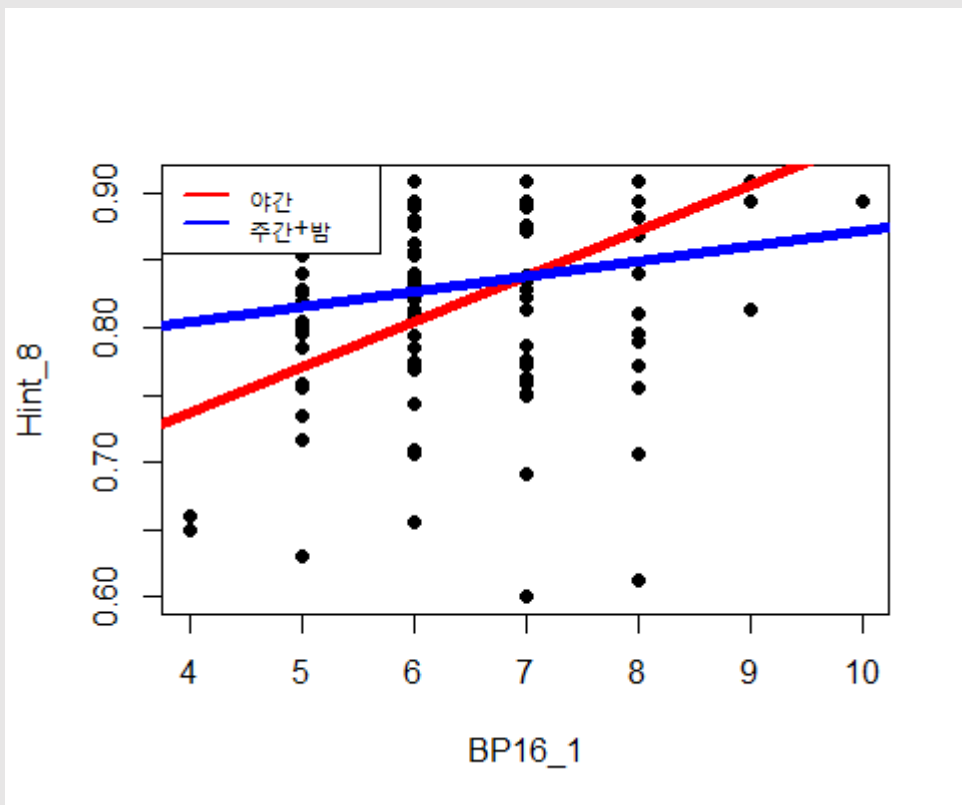
Residual standard error: 0.0699 on 131 degrees of freedom  
Multiple R-squared: 0.07167, Adjusted R-squared: 0.05042  
F-statistic: 3.371 on 3 and 131 DF, p-value: 0.02053

### 3. 변수형 설명변수 추가

변수형 설명변수

9. 설명변수(남녀, 근무형태, 교육수준)  
추가

HINT8지수에 대한 주중수면시간과 근무형태의 교호작용



Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.94064	0.20920	4.496	1.51e-05 ***
BP16_1	-0.01979	0.03241	-0.610	0.543
factor(EC_wht_5)1	-0.20986	0.21187	-0.990	0.324
BP16_1:factor(EC_wht_5)1	0.03519	0.03281	1.073	0.285
---				
Signif. codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.' 0.1 ' ' 1

Residual standard error: 0.07001 on 131 degrees of freedom  
Multiple R-squared: 0.06867, Adjusted R-squared: 0.04734  
F-statistic: 3.22 on 3 and 131 DF, p-value: 0.02493

교호작용 유의함.

근무형태에 따른 평균Hint\_8지수가 효과가 있음.

### 3. 변수형 설명변수 추가

변수형 설명변수

9. 설명변수(남녀, 근무형태, 교육수준)  
추가

#### Full모형에 대한 추정

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.6916377  0.0552456  12.519 < 2e-16 ***
BP16_1         0.0136621  0.0050468   2.707  0.00771 **
BE3_87         0.0099253  0.0082842   1.198  0.23308
factor(sex)1    0.0373685  0.0213834   1.748  0.08292 .
factor(edu)1   -0.0002992  0.0122580  -0.024  0.98057
factor(EC_wht_5)1 0.0112099  0.0410753   0.273  0.78536
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.06961 on 129 degrees of freedom
Multiple R-squared:  0.09334,    Adjusted R-squared:  0.0582
F-statistic: 2.656 on 5 and 129 DF,  p-value: 0.02548
```

## 4. 결론 및 정리

$$Y = B_0 + \beta_1 X_1 + B_2 x_2 + \varepsilon$$

Y: HINT\_8지수

X1: 주중 하루 평균 수면시간

X2: 여가\_중강도 신체활동 시간

$$Y = B_0 + \beta_1 X_1 + B_2 x_2 + \delta Z + \gamma X_1 Z + \varepsilon$$

Y: HINT\_8지수

X1: 주중 하루 평균 수면시간

X2: 여가\_중강도 신체활동 시간

$\delta$  : 근무형태에 따른 평균 HINT\_8 지수의 차이

$\gamma$  : 근무형태에 따른 평균 HINT\_8 지수의 기울기 차이

근무형태에 따른 평균Hint\_8지수가 효과가 있음.

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.726351	0.032935	22.054	< 2e-16 ***
x1	0.015063	0.004917	3.063	0.00263 **
x2	0.012821	0.007688	1.668	0.09763 .

---

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

Residual standard error: 0.06982 on 139 degrees of freedom  
Multiple R-squared: 0.07541, Adjusted R-squared: 0.06211  
F-statistic: 5.668 on 2 and 139 DF, p-value: 0.0043

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.94064	0.20920	4.496	1.51e-05 ***
BP16_1	-0.01979	0.03241	-0.610	0.543
factor(EC_wht_5)1	-0.20986	0.21187	-0.990	0.324
BP16_1:factor(EC_wht_5)1	0.03519	0.03281	1.073	0.285

---

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

Residual standard error: 0.07001 on 131 degrees of freedom  
Multiple R-squared: 0.06867, Adjusted R-squared: 0.04734  
F-statistic: 3.22 on 3 and 131 DF, p-value: 0.02493

# 삼성서울병원·성균관대·카이스트 연구팀, 교대 근무 간호사 수면패턴 분석



"수면패턴, 개인 일주기 리듬과 일치시 평균 수면시간  
비슷·주간졸음 감소"

"야간 근무 후 짧게, 주간 근무 후 길게 수면을 취하면  
주간 졸림이 완화"

"주간 근무와 야간 근무 후 동일한 시간을 자는 것보단  
야간 근무 후에 짧게 자고, 주간 근무 후에는 길게 자는  
것이 주간 졸림증을 완화시킬 수 있을 것 "

"불규칙한 수면질환으로 고통을 겪고 있을  
교대근무자들에게 도움이 될 수 있는 연구를 할 수 있어  
기뻐 "



# 한계점 및 느낀점

- 1 설명변수를 너무 적게 설정하여  $R^2$ 값의 유의미한 결과를 도출 할 수 없었다.
- 2 변수를 선택할 때 p-value 값을 기준으로 많이 한 것 같아서 변수가 유의미하지 않을 수도 있을 것이다.
- 3 많은 연구팀에서 수리모델을 이용해 개인 맞춤형 수면 패턴을 실시간으로 제공하고 있다. 나도 열심히 데이터분석 실력을 향상시켜서 정신적으로 고통받고 있는 그들에게 힘이 되주고 싶다.
- 4 데이터분석을 혼자서 하는것은 처음이었는데 나의 실력이 많이 부족하다는 것을 뼈빠지게 깨달았다. 이것을 발판 삼아 데이터 분석에 관심을 갖고 공부하기로 마음 먹었다.  
한 학기 동안 고생하신 박주영 교수님께 감사의 인사를 전합니다.

감사합니다