

[REPORT]

**■ 과 목 명 : 시계열 자료 분석과 예측**

**■ 담당교수 : 김명석 교수님**

**■ 제 출 일 : 2022.10.10**

**■ 학과: 데이터 사이언스. 인공지능**

**■ 학번: A64037**

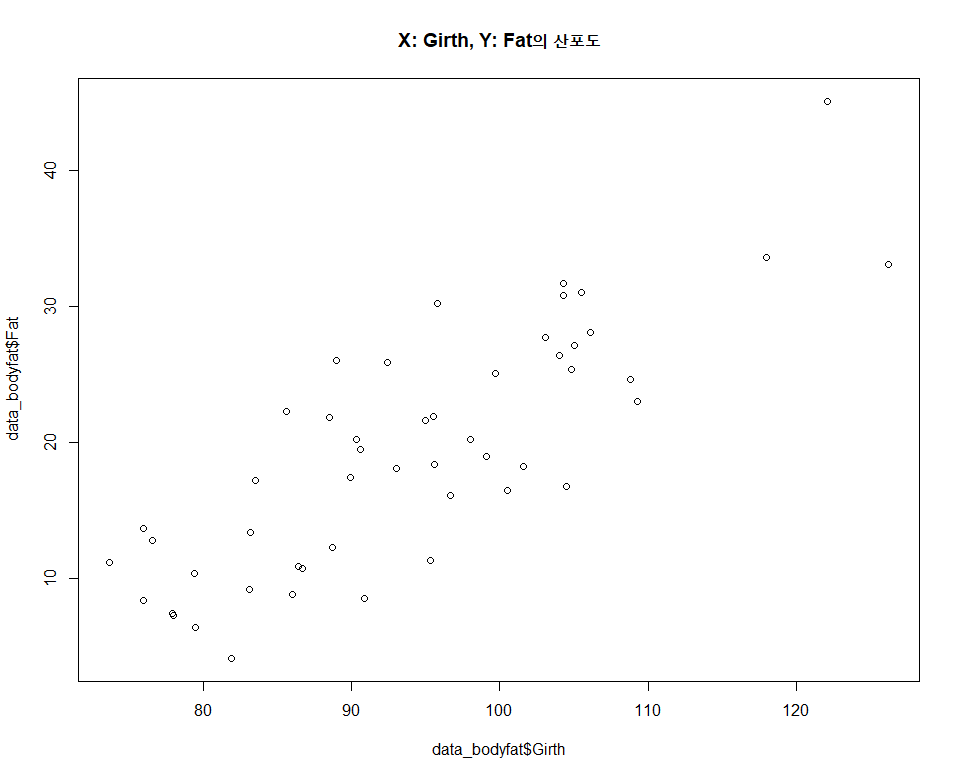
**■ 성명: 서승덕**

1.

1) 0.5313237

2) p-value = 0.02818로 귀무가설을 잘못 기각할 확률이 2.818%로 유의수준 1%에 비해 크기 때문에 귀무가설을 기각하지 못한다. 즉, 유의수준을 5%로 설정한다면 귀무가설을 기각하여, 이 상관계수가 의미가 있다고 판단하겠으나, 1%로 설정하였기 때문에 이 상관계수를 사용할 수 없다.

2.

1) 

2)

상관계수(r) : 0.8188484

결정계수(R^2): 0.6705

3)

상관계수의 제곱이 결정계수임. 음과 양이 없어지므로, 결정계수의 경우 상관 관계 분석이 아닌 정량화를 통한 회귀 분석에서 사용하는 수치이며, 상관계수의 경우 방향성이 있기 때문에 상관성을 나타냄.

4)

절편: -36.2397, 기울기: 0.5905

5)

Girth가 1 증가할 때마다 Fat이 0.5905증가

6)

기울기의 유의성 검정

귀무가설: 회귀직선의 의미가 없다.(직선의 기울기가 0이다.)

대립가설: 회귀직선의 의미가 있다.(직선의 기울기가 0이 아니다.)

7)

Estimate Std. Error t value Pr(>|t|)

data\_bodyfat$Girth 0.59053 0.05975 9.883 3.71e-13 \*\*\*

P-value가 3.71e-13로 유의수준 5프로보다 작음.

🡺 귀무가설 기각

8)

F검정

귀무가설: 모델이 유효하지 않다.

대립가설: 모델이 유효하다.

9)

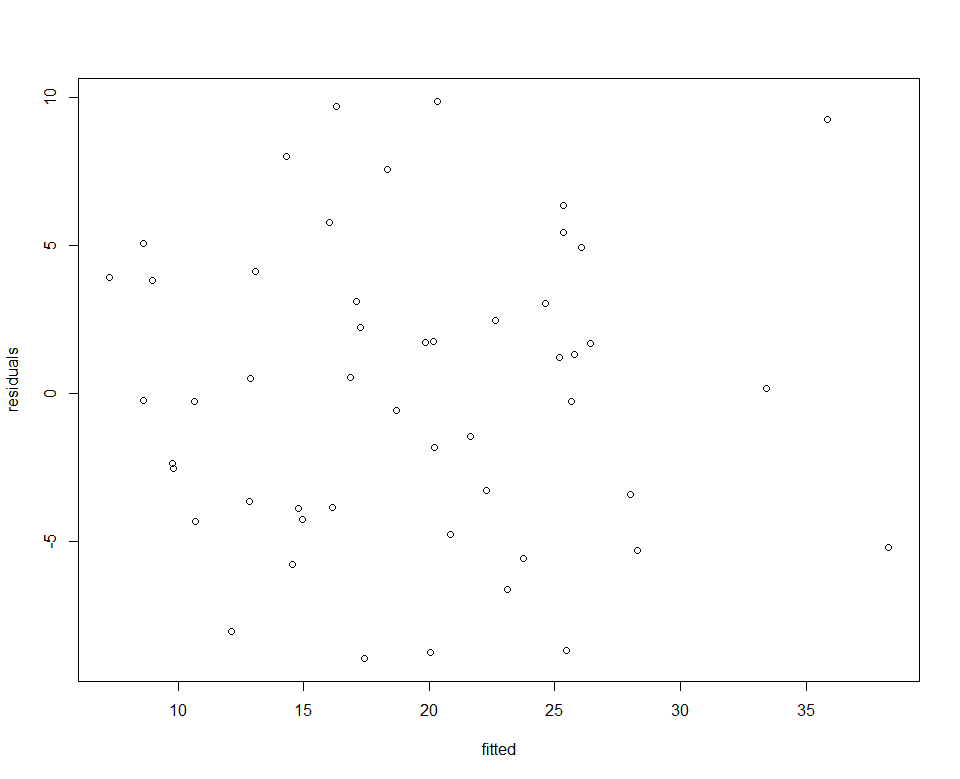
anova(model)

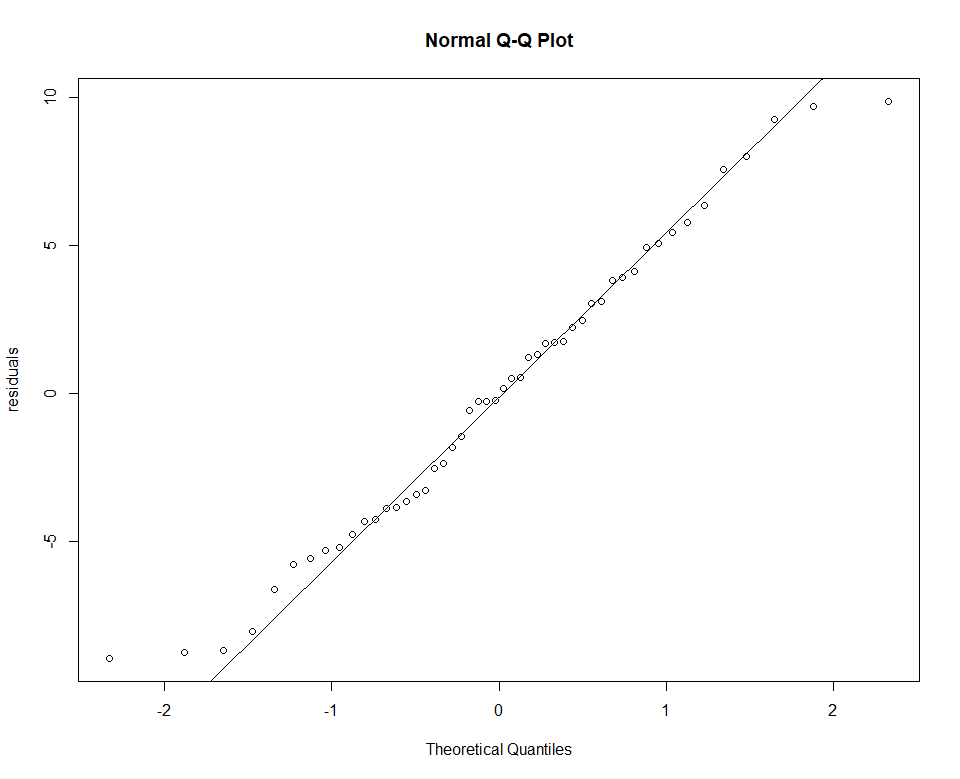
Response: data\_bodyfat$Fat

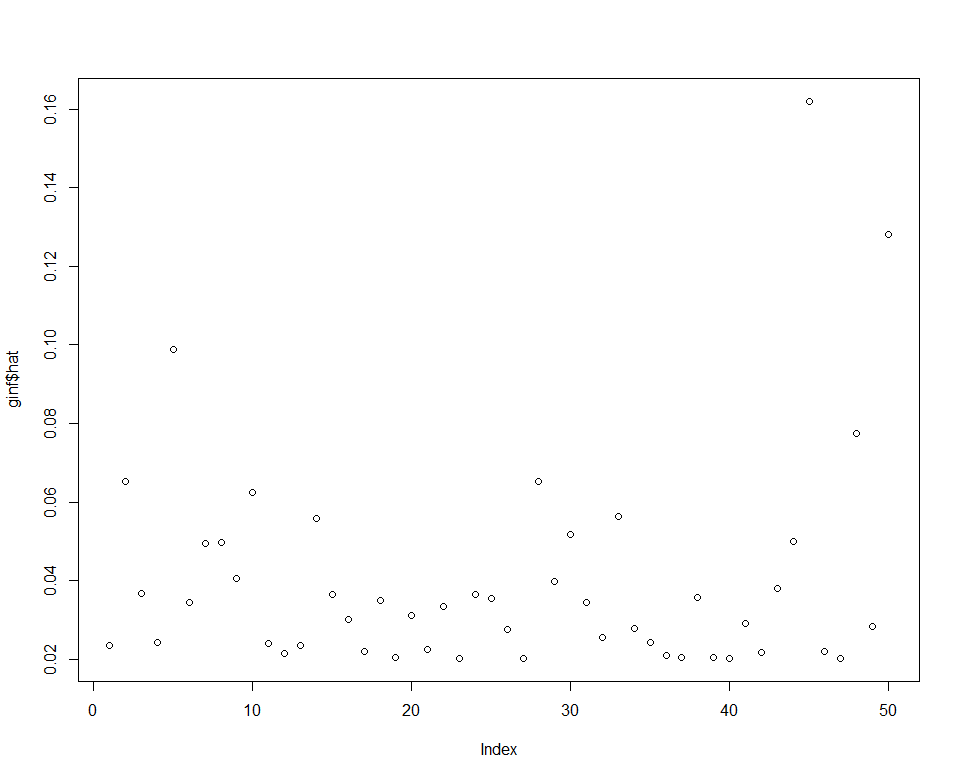
Df Sum Sq Mean Sq F value Pr(>F)

data\_bodyfat$Girth 1 2527.1 2527.12 97.681 3.714e-13 \*\*\*

5%기준으로 보았을 때 p-value = 3.714e-13로 귀무가설을 기각하게 됨. 즉 모델이 유효하다.

10)

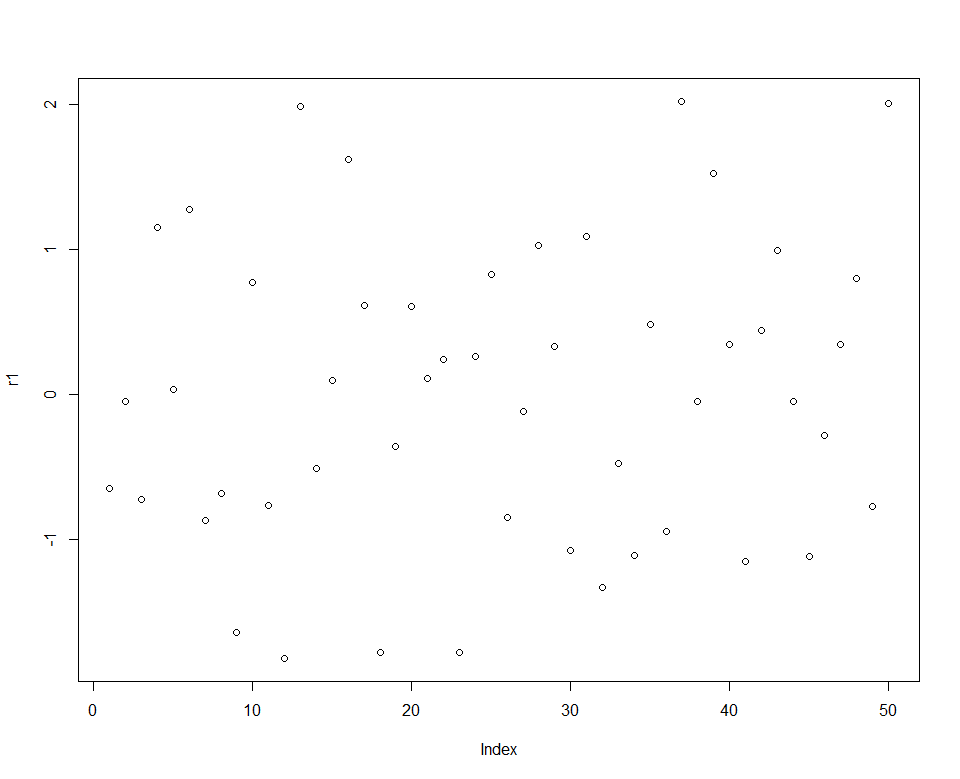
11)

이상치( leverage가 3/n 보다 크면 이상치, mean=0.04 🡺 0.12)

2개 존재

45 50

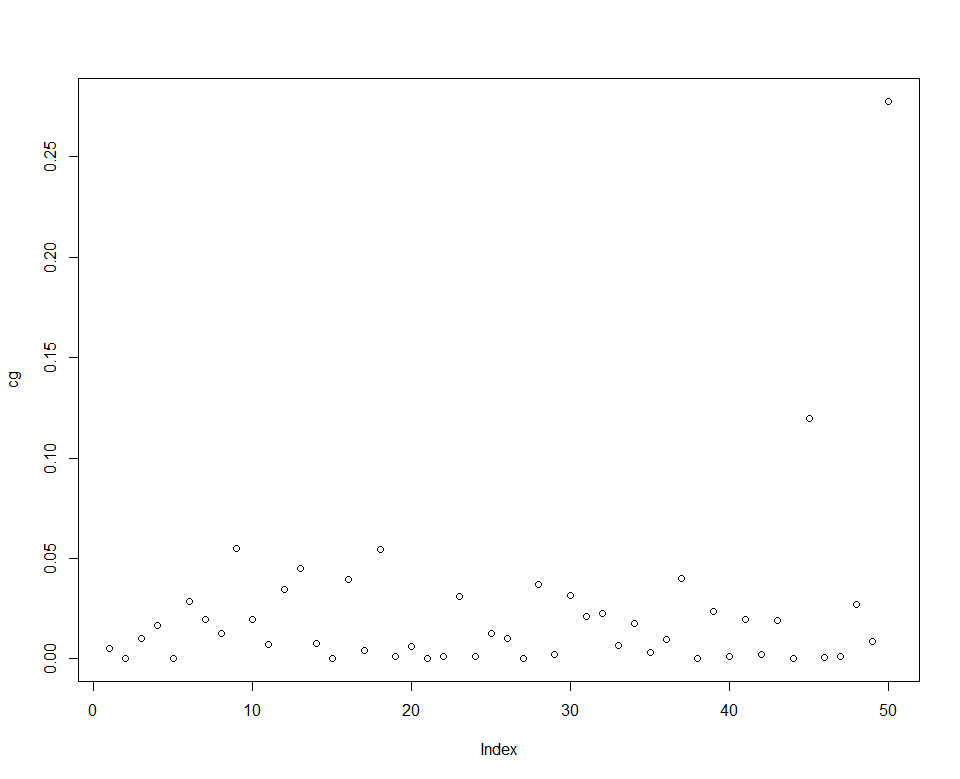
0.1620821 0.1280931

12)

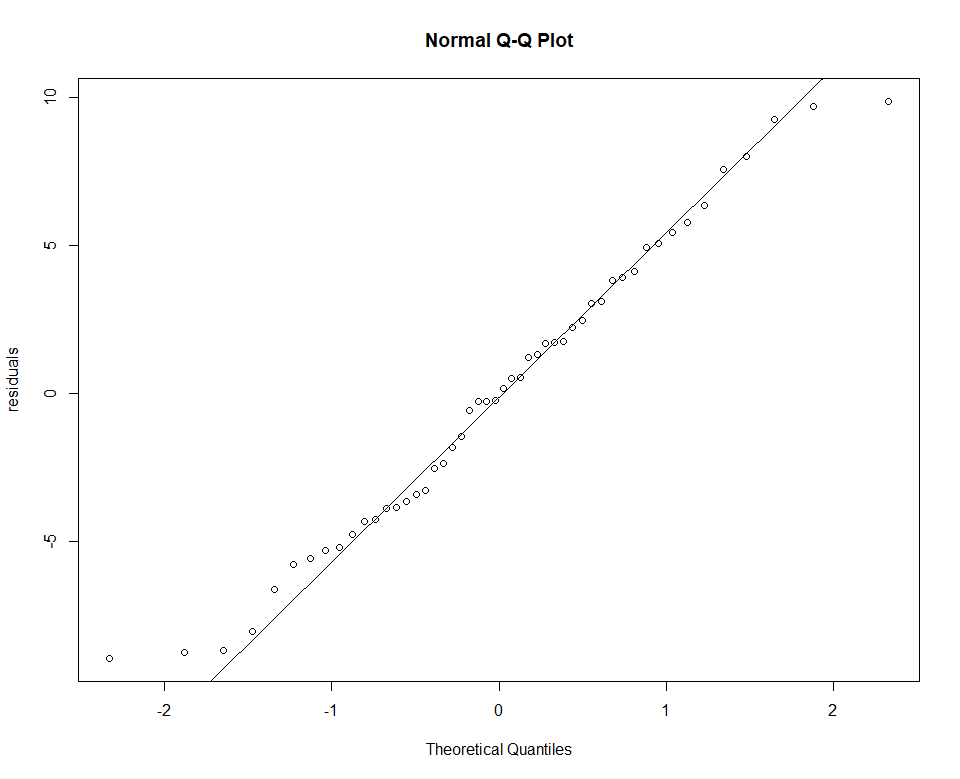
37, 50 번째 이상치(sdr의 절대값이 2 이상이면 이상치)

37 50

2.022098 2.004957

13)

영향치(Cook’s D가 1 이상이면 influential point) 없음.

14

15)

p-value = 0.8522로 유의수준 10%보다 높음. 🡺 H0: F=F0 🡺 표본의 모집단이 정규분포를 이루고 있다.

16)

p-value = 0.459 유의수준 10%보다 높음 🡺 H1: H0: 오차는 동분산

17)

p-value = 0.9247 유의수준 10%보다 높음 🡺 잔차들 사이에 자기상관관계가 없다. 독립적이다.