

Problem 2: 6210422036 ธนัท เอี่ยมปริตติ

ให้ใช้ independent variable เดียวคือ Car Age.

```
reg <- lm(FuelConsumption ~ CarAge, data=d)
summary(reg)

checkresiduals(reg)
```

```
Call:
lm(formula = FuelConsumption ~ CarAge, data = d)

Residuals:
    Min       1Q   Median       3Q      Max
-0.9940 -0.4436 -0.1507  0.3553  1.0316

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  26.8323     0.4998  53.688 1.22e-13 ***
CarAge       -1.2255     0.1502  -8.161 9.88e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7221 on 10 degrees of freedom
Multiple R-squared:  0.8695,    Adjusted R-squared:  0.8564
F-statistic: 66.61 on 1 and 10 DF,  p-value: 9.881e-06
```

2.1 จงเขียน regression equation ที่ได้

$\text{FuelConsumption} = 26.8323 - 1.2255(\text{CarAge})$

2.2 จงทำ diagnostic checks ตรวจสอบ regression model ที่ได้ (ให้ครบทุก assumptions)

```
Call:
lm(formula = FuelConsumption ~ CarAge, data = d)

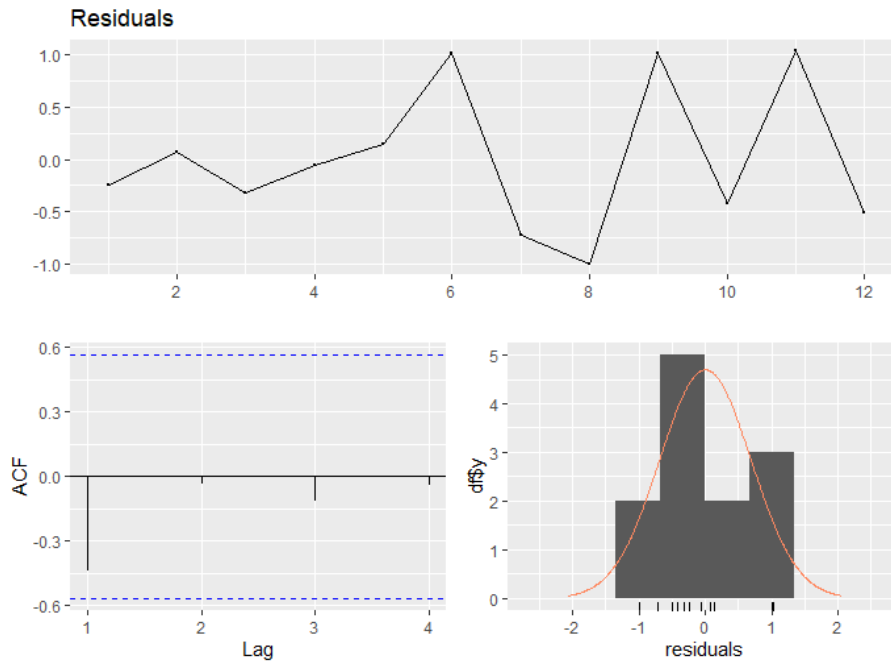
Residuals:
    Min       1Q   Median       3Q      Max
-0.9940 -0.4436 -0.1507  0.3553  1.0316

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  26.8323     0.4998  53.688 1.22e-13 ***
CarAge       -1.2255     0.1502  -8.161 9.88e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7221 on 10 degrees of freedom
Multiple R-squared:  0.8695,    Adjusted R-squared:  0.8564
F-statistic: 66.61 on 1 and 10 DF,  p-value: 9.881e-06
```

1. เครื่องหมายของ coefficient ของ Car Age = -1.2255 เมื่ออายุของรถมากขึ้น ประสิทธิภาพ ค่อย ๆ ลดลง ซึ่งก็น่าจะสอดคล้องกับการที่มี fuel consumption ลดลงเพราะ output ที่น้อยลง
2. จากผลของการ run regression ด้วย R บอกว่า coefficient ของ Car Age นั้น significant ที่จะไม่เท่ากับ 0
3. เนื่องจากเป็น simple regression เราสามารถเช็คความสามารถในการอธิบาย variance ด้วยค่า R squared = 0.8695 ซึ่งถือว่าสูง

4. Check residual assumption



- 4.1) ค่า mean ของ residual น่าจะอยู่แถว ๆ 0
- 4.2) มี constant variance (แต่อาจจะไม่แน่ใจช่วง 0-5 ที่มี variance แคบกว่านิดหน่อย)
- 4.3) จาก ACF plot จะเห็นว่าไม่มี serial correlation
- 4.4) error มีการกระจายตัวแบบ normal distribution

2.3 ถ้าอายุของรถเป็น 3.2 ปี จงพยากรณ์อัตราการใช้น้ำมัน พร้อมทั้ง 90% Prediction Interval (PI)

Point Forecast	Lower Bound of 90% PI	Upper Bound of 90% PI

```
> newdata <- data.frame(CarAge=3.2)
> fc <- forecast(reg, newdata=newdata, level=90)
> fc
   Point Forecast    Lo 90    Hi 90
1      22.91053  21.54754  24.27352
```

Forecasting Fuel Consumption (Car Age = 3.2)

PointForecast = 22.91053
 Lower Bound 90% PI = 21.54754
 Upper Bound 90% PI = 24.27352