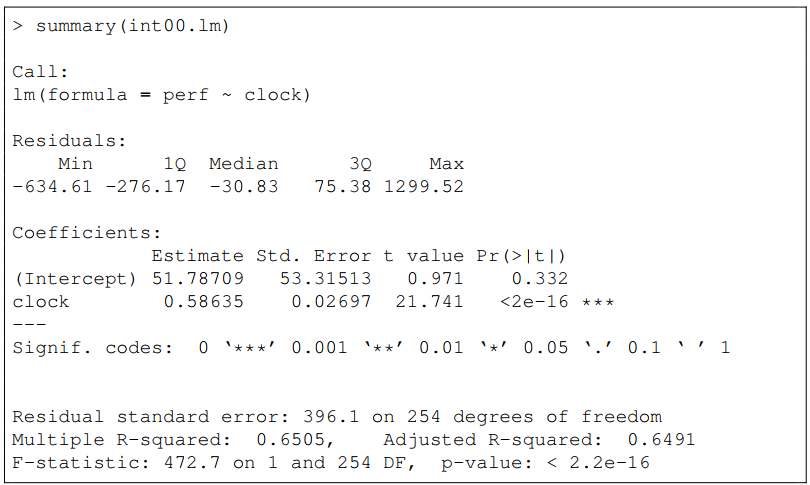
# Cách đánh giá 1 model:



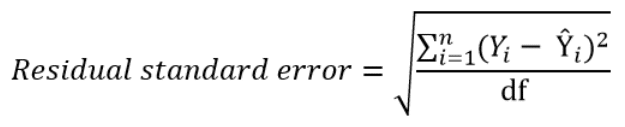
Residual và coefficent

Nếu mô hình fit tốt thì giá trị của residual nên phân bổ xung quanh số 0

Median của residual nên nằm gần 0

Std nên nhỏ hơn ít nhất 5 lần estimate(trừ Intercept)

Residual standard error:



Nếu residual phân phối chuẩn, thì 1Q và 3Q nên gấp 1,5 lần Residual standard error.

3 cái này khác nhau chỗ nào?

( https://stats.stackexchange.com/questions/110999/r-confused-on-residual-terminology)

* Root mean square error
* residual sum of squares
* residual standard error
* mean squared error
* test error

Multiple R-square:

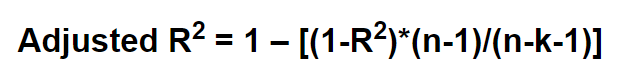
Multiple R-squared value is a number between 0 and 1

It is a statistical measure of how well the model describes the measured data.

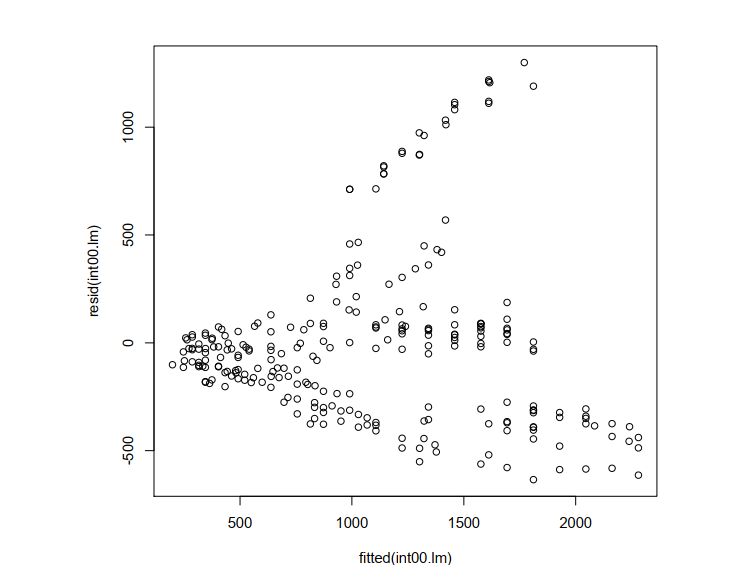
Được tính bằng ()

The reported *R*2 of 0.6505 for this model means that the model explains 65.05 percent of the data’s  
variation.

Adjusted R^2:

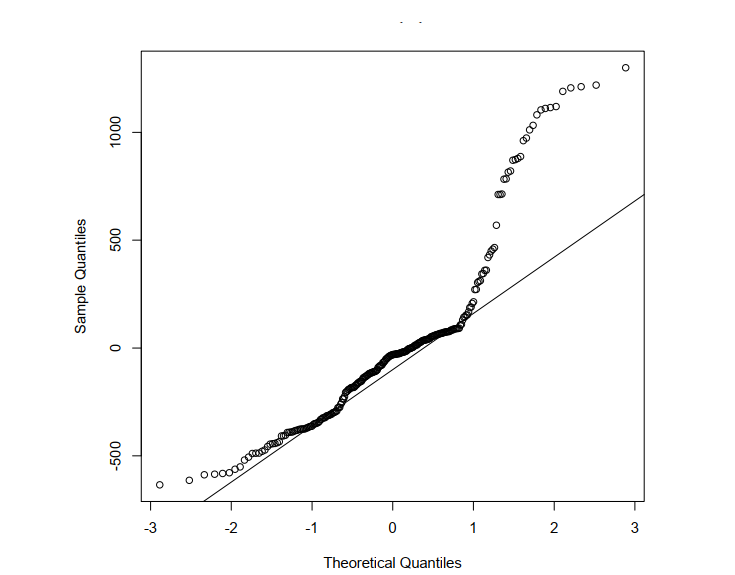


# Residual Analysis:



Trong biểu đồ này ta thấy được residual có xu hướng tăng khi mà di chuyển sang bên phải, có thể thấy được rằng dùng biến “clock” không miêu tả hết được data. Điều này ko có nghĩa là biến clock vô dụng mà ta nên xây dựng thêm

Q-Q Plot:



Nếu residual mà phân phối chuẩn thì các điểm sẽ nằm trên 1 đường thẳng, mà trong mô hình ta thấy cả đầu và đuôi đều bị tách ra, đây là dấu hiệu của **KHÔNG PHÂN PHỐI CHUẨN.** Điều này thể hiện rằng chỉ riêng biến clock thôi thì ko đủ để explain data.