Need for Linear mixed effect model

Chandan Pandey

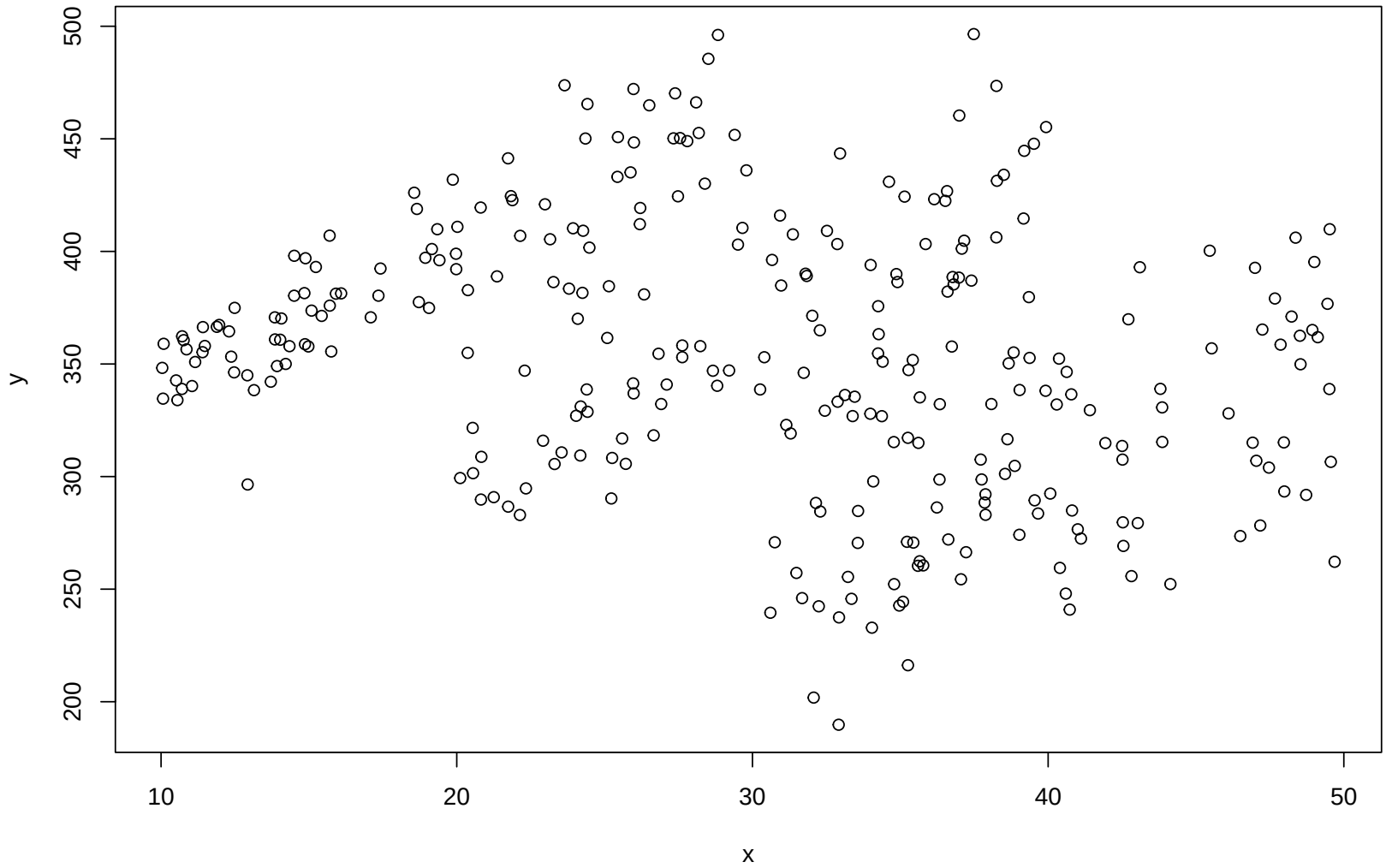
19/10/2022

When do we need GLMM.

- Some data are very messy and complex.
- eg., effect of N_2 on grass growth; Many factor can influence
  - Species difference.
  - Local Area variable.
  - Species composition.
  - Other Nutrient factors.
  - And many many more.
- Impossible to account for all the predictors.

Simpson's Paradox

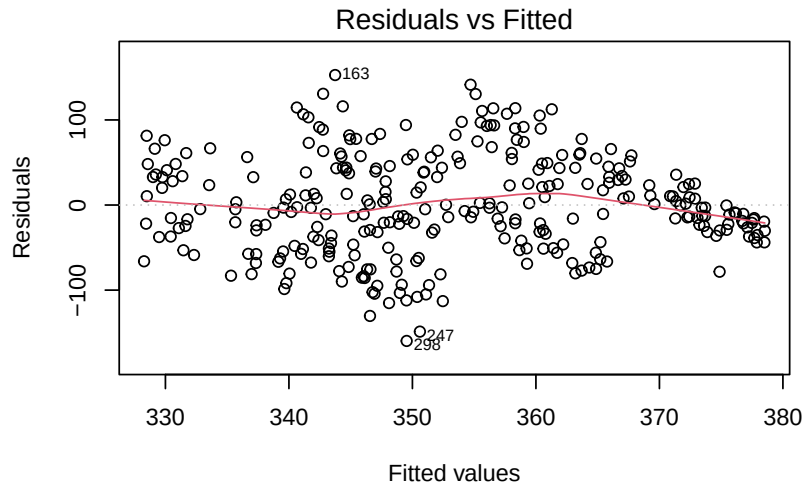
What is your intuition about relation between X and Y ?

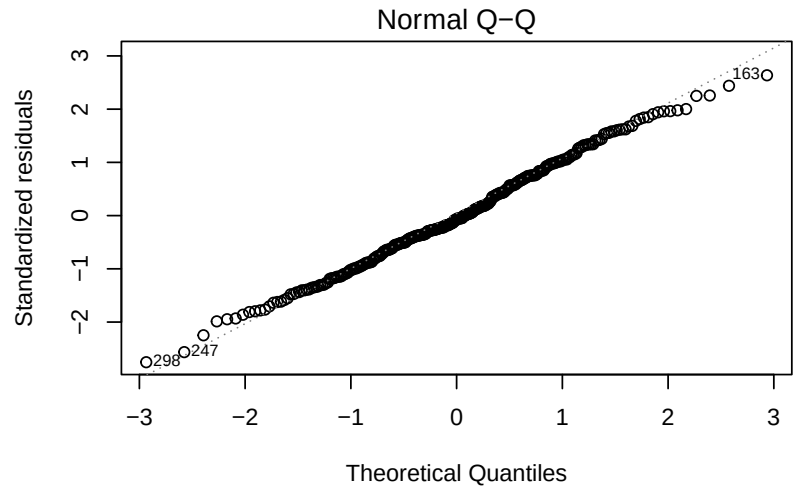


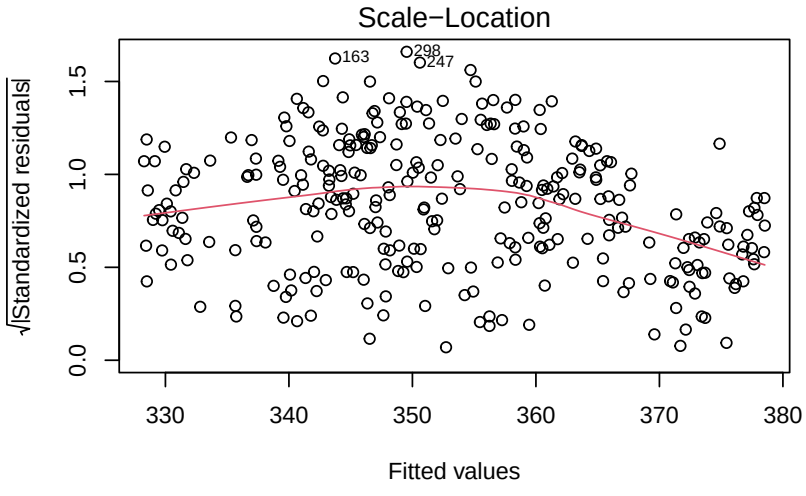
Let plot this model with simple linear model

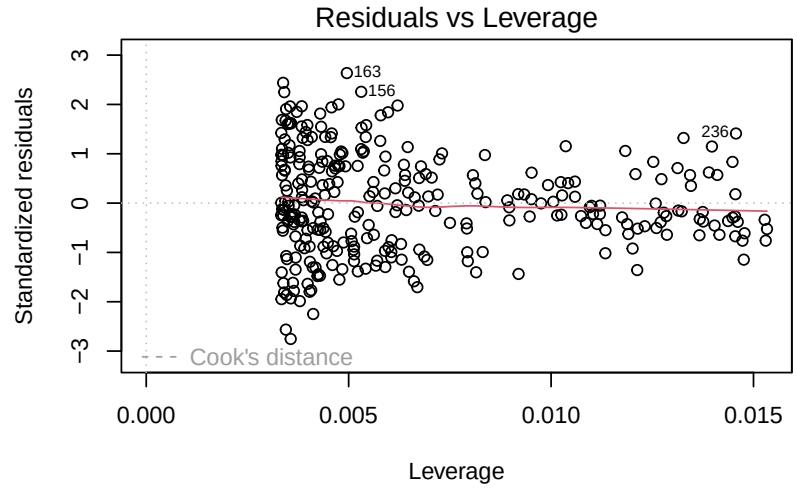
```
##                Estimate     SE t_value          P
## (Intercept)      391.26 10.120    38.7 3.90e-118
## x                 -1.27  0.317    -4.0  8.08e-05
```

Let plot this model with simple linear model

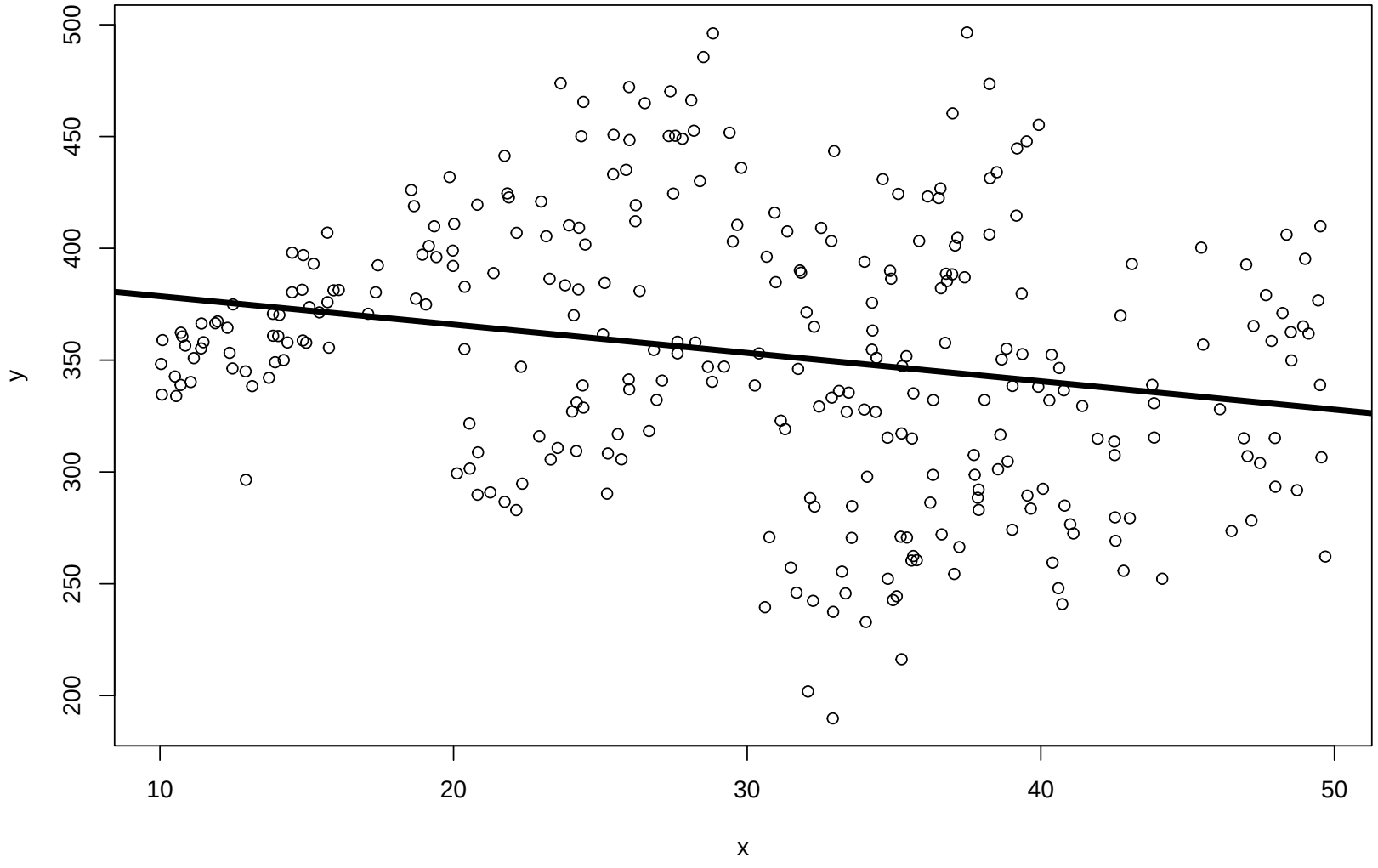








Let plot this model with simple linear model



Now let me group these variable

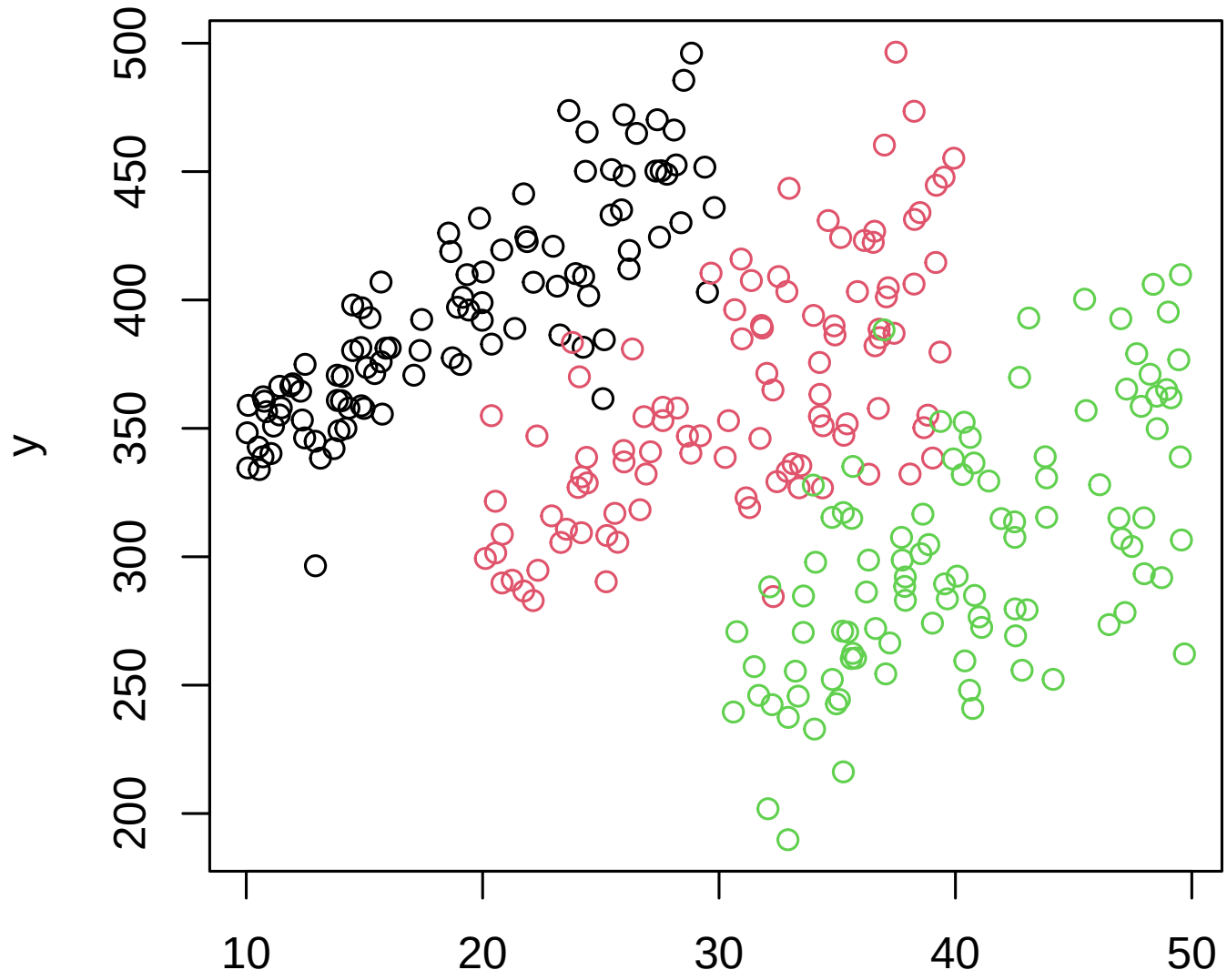
I have lied about dataset here a bit

These value belongs to 3 groups. Think these groups as

- Three different species
- Three trials
- Plots
- communities
- subspecies
- and many more

What will say on relation of x and Y ?

Here is plot

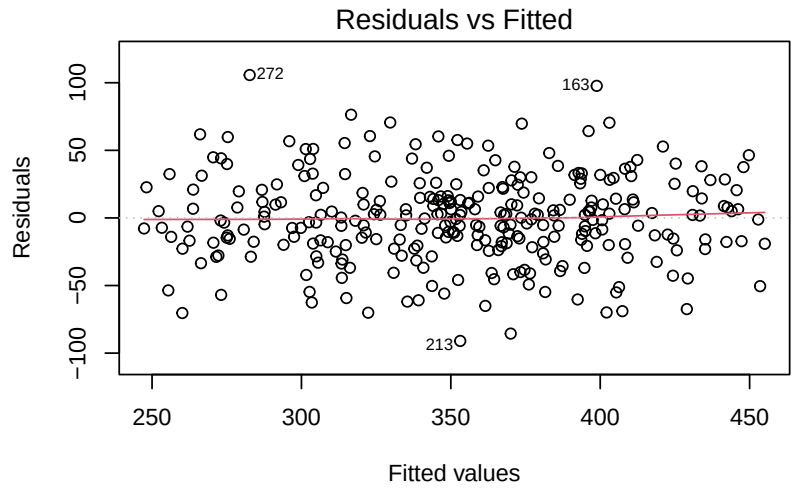


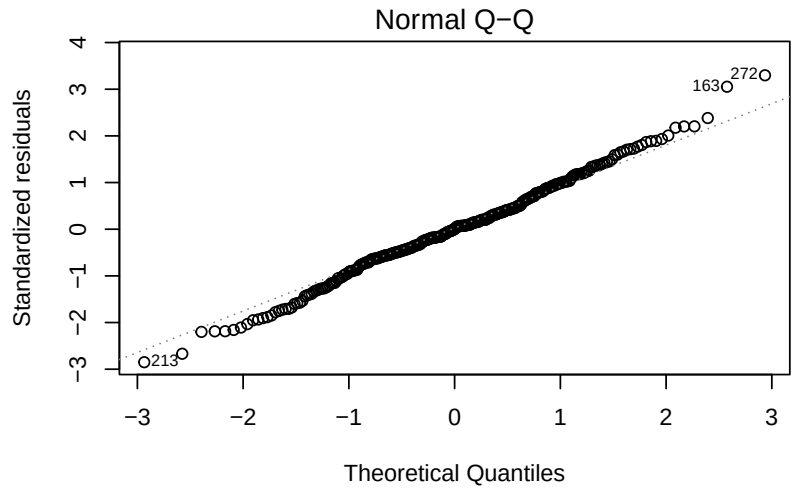
One way to address this problem

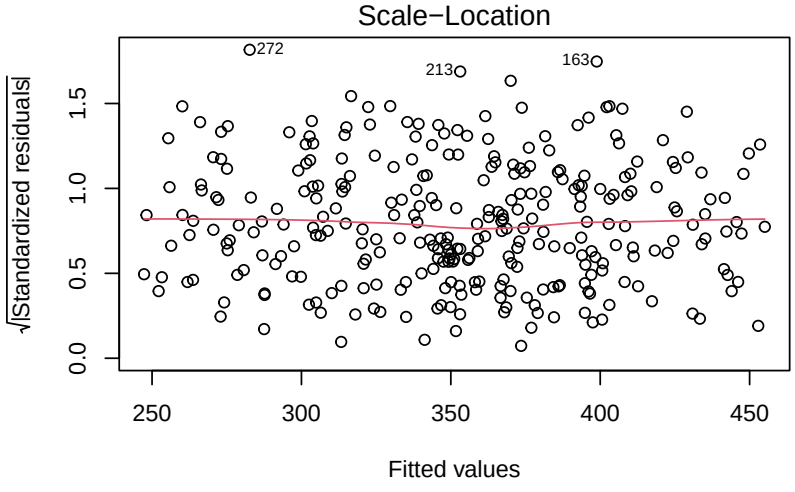
Model with two predictor but there is problem with that as well.
We will discuss this latter but first model with two predictor.

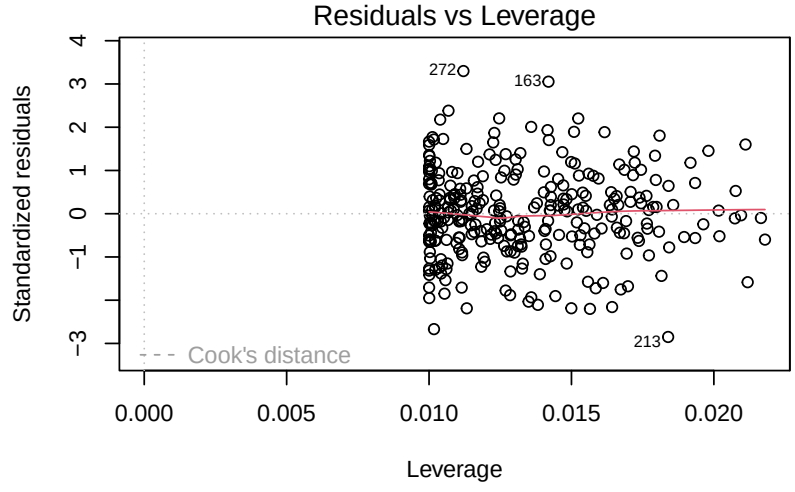
```
##               Df Sum Sq Mean Sq F value  Pr(>F)
## x              1  53932   53932    51.9 4.8e-12 ***
## group          2 698493  349247   336.4 < 2e-16 ***
## Residuals    296 307325    1038
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

## (Intercept)           x      groupB      groupC
##      289.86         5.55      -98.90     -212.26
```

Model validation









Models performance

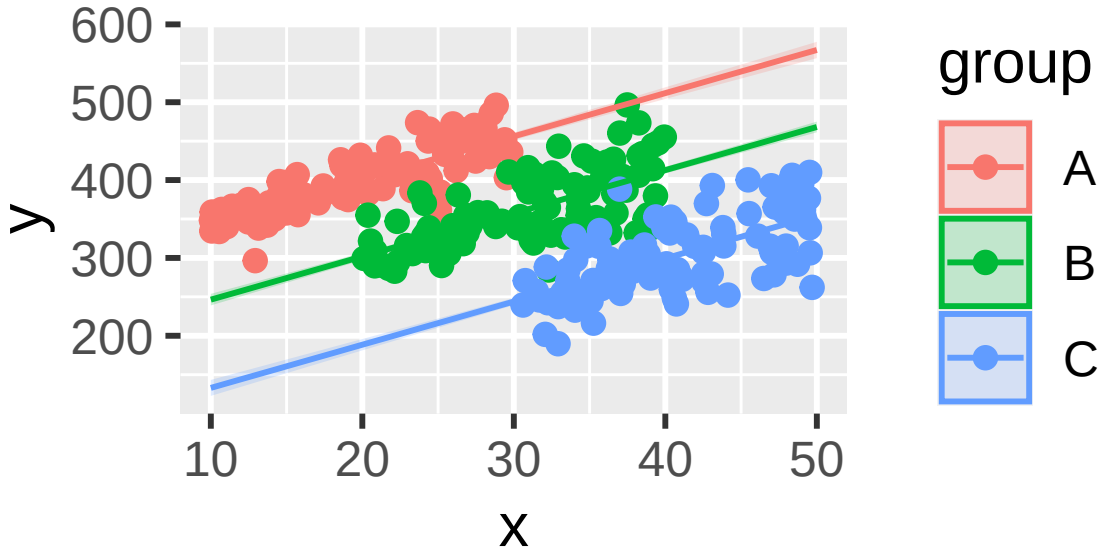
Let us compare the two models.

```
##
## Model selection based on AICc:
##
##          K AICc Delta_AICc AICcWt Cum.Wt    LL
## grouped 5 2941          0      1      1 -1465
## Single  3 3293        352      0      1 -1643
```

```
##     model R_squared
## 1  Single    0.0477
## 2 grouped    0.7071
```

Model with two predictor have improved the model but what is the trend

Plot the trend with two predictors



Q:

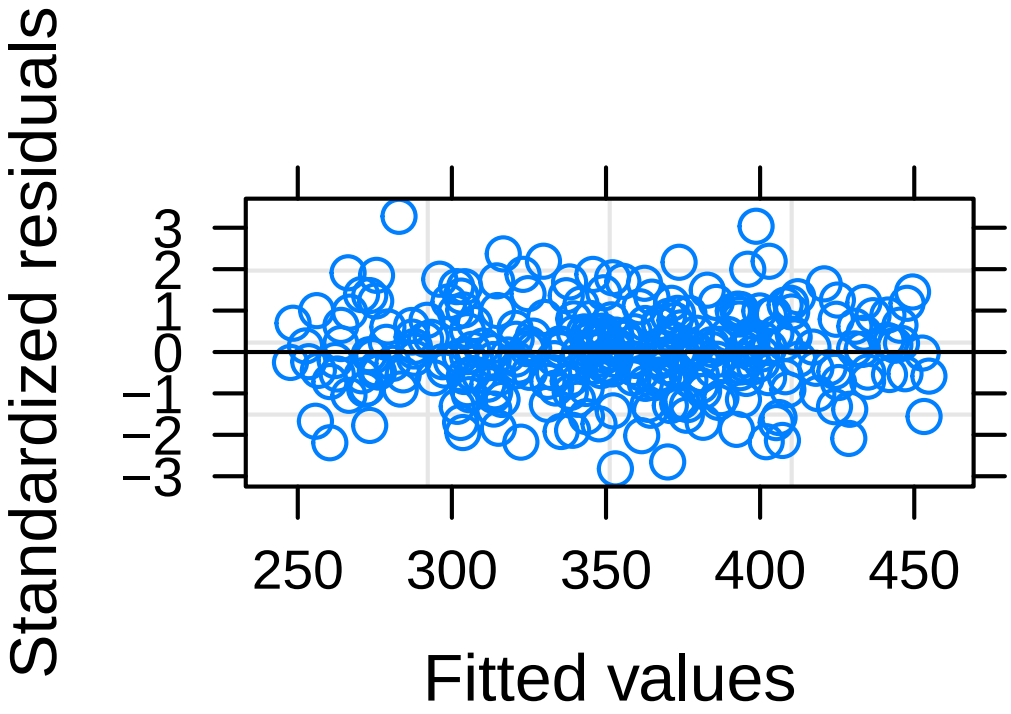
- What is the correct trend ?
- What was real question this case ? can we generalized this model.

What is the answer

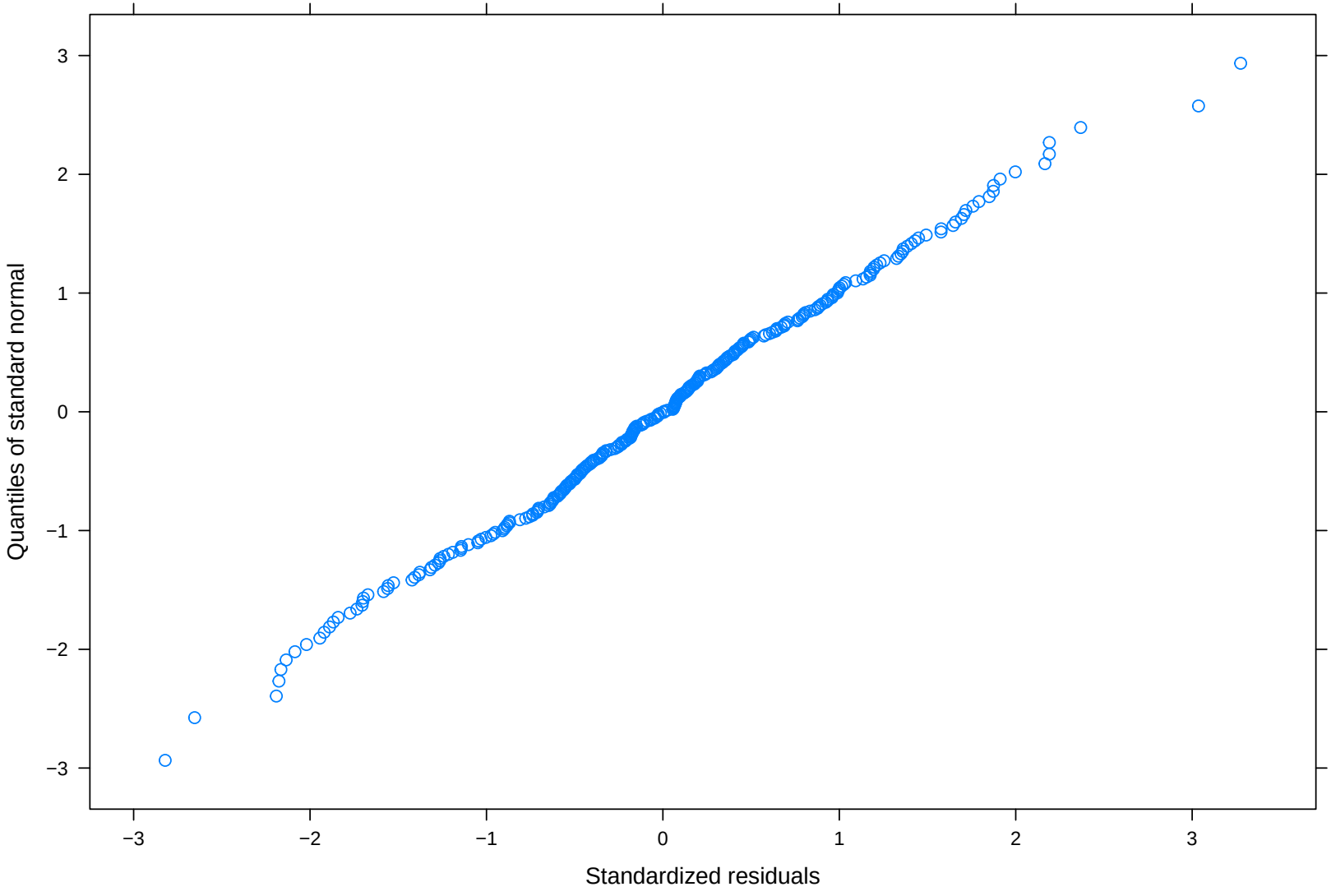
Linear Mixed effect model

Let model with LMM (1)

```
mod_LM <- lme(y~x, random =~1|group,
              data = simps_para,method = "REML")
plot(mod_LM)
```



Let model with LMM (2)



Model summary

```
## Linear mixed-effects model fit by REML
##   Data: simps_para
##    AIC  BIC logLik
##   2952 2967  -1472
##
## Random effects:
##  Formula: ~1 | group
##         (Intercept) Residual
## StdDev:         106     32.2
##
## Fixed effects:  y ~ x
##             Value Std.Error  DF t-value p-value
## (Intercept) 186.8      62.0 296    3.01  0.0028
## x             5.5       0.3 296   17.26  0.0000
##  Correlation:
##   (Intr)
## x -0.155
##
```

Ploting the model

```
plot(y~x,data=simps_para)
abline(a=mod_LM$coefficients$fixed[1],
       b=mod_LM$coefficients$fixed[2],
       col="red",lwd=5)
```