library(readxl)

> data <- read\_excel("Documents/11-Linear Regression:Linear Mixed Model/data.xlsx")

> View(data)

> number <- data$No.

> province <- data$Province

> plant\_name <- data$`Plant name`

> unit <- data$`Number of units`

> capacity <- data$`Total plant installed capacity (MW)`

> type <- data$`Fuel type`

> SO2 <- data$`SO2 emissions (Mg)`

> NOx <- data$`NOx emissions (Mg)`

> PM <- data$`PM emissions (Mg)`

> df <- data.frame(number,province, plant\_name, unit, capacity, type, SO2, NOx, PM)

> model\_SO2\_unit <- lm(SO2~unit, data = df)

> summary(model\_SO2\_unit)

Call:

lm(formula = SO2 ~ unit, data = df)

Residuals:

Min 1Q Median 3Q Max

-2483.9 -229.2 -113.1 4.6 6182.6

Coefficients:

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

(Intercept) -11.771 21.038 -0.56 0.576

unit 133.534 7.674 17.40 <2e-16 \*\*\*

---

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

Residual standard error: 573.5 on 2595 degrees of freedom

Multiple R-squared: 0.1045, Adjusted R-squared: 0.1041

F-statistic: 302.8 on 1 and 2595 DF, p-value: < 2.2e-16

> summary(model\_SO2\_unit)$coef

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

(Intercept) -11.77144 21.03837 -0.5595226 5.758534e-01

unit 133.53371 7.67435 17.4000023 3.170445e-64

> model\_SO2\_capacity <- lm(SO2 ~ capacity, data = df)

> summary(model\_SO2\_capacity)

Call:

lm(formula = SO2 ~ capacity, data = df)

Residuals:

Min 1Q Median 3Q Max

-1825.5 -107.5 -82.4 -10.1 5971.0

Coefficients:

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

(Intercept) 96.15260 11.48862 8.369 <2e-16 \*\*\*

capacity 0.53151 0.01552 34.247 <2e-16 \*\*\*

---

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

Residual standard error: 503 on 2595 degrees of freedom

Multiple R-squared: 0.3113, Adjusted R-squared: 0.311

F-statistic: 1173 on 1 and 2595 DF, p-value: < 2.2e-16

> summary(model\_SO2\_capacity)$coef

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

(Intercept) 96.1526039 11.48862120 8.369377 9.338275e-17

capacity 0.5315064 0.01551995 34.246654 2.051908e-212

> table(type)

type

Biomass Coal Gas & Oil Others

76 2099 111 311

> res <- model.matrix(~type)

> head(res)

(Intercept) typeCoal typeGas & Oil typeOthers

1 1 0 1 0

2 1 1 0 0

3 1 1 0 0

4 1 1 0 0

5 1 1 0 0

6 1 1 0 0

> model\_SO2\_type <- lm(SO2 ~ type, data = df)

> model\_SO2\_type

Call:

lm(formula = SO2 ~ type, data = df)

Coefficients:

(Intercept) typeCoal typeGas & Oil typeOthers

52.97 290.38 -28.09 92.29

> summary(model\_SO2\_type)

Call:

lm(formula = SO2 ~ type, data = df)

Residuals:

Min 1Q Median 3Q Max

-343.2 -304.9 -145.3 7.3 6837.8

Coefficients:

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

(Intercept) 52.97 68.65 0.771 0.440

typeCoal 290.38 69.88 4.155 3.36e-05 \*\*\*

typeGas & Oil -28.09 89.11 -0.315 0.753

typeOthers 92.29 76.58 1.205 0.228

---

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

Residual standard error: 598.5 on 2593 degrees of freedom

Multiple R-squared: 0.02561, Adjusted R-squared: 0.02448

F-statistic: 22.72 on 3 and 2593 DF, p-value: 1.625e-14

> summary(model\_SO2\_type)$coef

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

(Intercept) 52.96579 68.65309 0.7714990 4.404815e-01

typeCoal 290.37571 69.88493 4.1550548 3.357385e-05

typeGas & Oil -28.08561 89.10856 -0.3151842 7.526472e-01

typeOthers 92.28791 76.58353 1.2050620 2.282893e-01

> table(province)

province

Anhui Beijing Chongqing Fujian Gansu Guangdong Guangxi Guizhou Hainan

65 1 33 57 29 111 37 26 11

Hebei Heilongjiang Henan Hubei Hunan Inner Mongol Jiangsu Jiangxi jilin

128 111 133 72 41 155 244 43 51

Liaoning Ningxia Qinghai Shaanxi Shandong Shanghai Shanxi Sichuan Tianjin

89 39 11 101 397 24 164 53 22

Xinjiang Yunnan Zhejiang

97 42 210

> res\_SO2\_province <- model.matrix(~province)

> head(res\_SO2\_province)

(Intercept) provinceBeijing provinceChongqing provinceFujian provinceGansu provinceGuangdong provinceGuangxi

1 1 1 0 0 0 0 0

2 1 0 0 0 0 0 0

3 1 0 0 0 0 0 0

4 1 0 0 0 0 0 0

5 1 0 0 0 0 0 0

6 1 0 0 0 0 0 0

provinceGuizhou provinceHainan provinceHebei provinceHeilongjiang provinceHenan provinceHubei provinceHunan

1 0 0 0 0 0 0 0

2 0 0 0 0 0 0 0

3 0 0 0 0 0 0 0

4 0 0 0 0 0 0 0

5 0 0 0 0 0 0 0

6 0 0 0 0 0 0 0

provinceInner Mongol provinceJiangsu provinceJiangxi provincejilin provinceLiaoning provinceNingxia provinceQinghai

1 0 0 0 0 0 0 0

2 0 0 0 0 0 0 0

3 0 0 0 0 0 0 0

4 0 0 0 0 0 0 0

5 0 0 0 0 0 0 0

6 0 0 0 0 0 0 0

provinceShaanxi provinceShandong provinceShanghai provinceShanxi provinceSichuan provinceTianjin provinceXinjiang

1 0 0 0 0 0 0 0

2 0 0 0 0 0 1 0

3 0 0 0 0 0 1 0

4 0 0 0 0 0 1 0

5 0 0 0 0 0 1 0

6 0 0 0 0 0 1 0

provinceYunnan provinceZhejiang

1 0 0

2 0 0

3 0 0

4 0 0

5 0 0

6 0 0

> model\_SO2\_province <- lm(SO2 ~ province, data = df)

> model\_SO2\_province

Call:

lm(formula = SO2 ~ province, data = df)

Coefficients:

(Intercept) provinceBeijing provinceChongqing provinceFujian provinceGansu

377.38 -360.08 223.05 -197.09 337.40

provinceGuangdong provinceGuangxi provinceGuizhou provinceHainan provinceHebei

-66.50 333.44 2280.60 -136.05 -243.13

provinceHeilongjiang provinceHenan provinceHubei provinceHunan provinceInner Mongol

-154.58 -246.22 -156.89 -26.75 415.25

provinceJiangsu provinceJiangxi provincejilin provinceLiaoning provinceNingxia

-211.56 54.19 -24.98 39.85 177.68

provinceQinghai provinceShaanxi provinceShandong provinceShanghai provinceShanxi

-98.26 -134.84 -292.52 -71.59 -154.02

provinceSichuan provinceTianjin provinceXinjiang provinceYunnan provinceZhejiang

-92.70 -242.70 308.56 -167.32 -292.53

> summary(model\_SO2\_province)

Call:

lm(formula = SO2 ~ province, data = df)

Residuals:

Min 1Q Median 3Q Max

-2638.8 -192.3 -79.5 28.3 6388.5

Coefficients:

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

(Intercept) 377.38 64.45 5.855 5.38e-09 \*\*\*

provinceBeijing -360.08 523.63 -0.688 0.491723

provinceChongqing 223.05 111.07 2.008 0.044727 \*

provinceFujian -197.09 94.30 -2.090 0.036702 \*

provinceGansu 337.40 116.04 2.908 0.003674 \*\*

provinceGuangdong -66.50 81.16 -0.819 0.412644

provinceGuangxi 333.44 107.02 3.116 0.001855 \*\*

provinceGuizhou 2280.60 120.58 18.913 < 2e-16 \*\*\*

provinceHainan -136.05 169.42 -0.803 0.422009

provinceHebei -243.13 79.14 -3.072 0.002149 \*\*

provinceHeilongjiang -154.58 81.16 -1.905 0.056940 .

provinceHenan -246.22 78.64 -3.131 0.001762 \*\*

provinceHubei -156.89 88.91 -1.765 0.077744 .

provinceHunan -26.75 103.64 -0.258 0.796322

provinceInner Mongol 415.25 76.79 5.408 6.97e-08 \*\*\*

provinceJiangsu -211.56 72.53 -2.917 0.003568 \*\*

provinceJiangxi 54.19 102.15 0.531 0.595786

provincejilin -24.98 97.21 -0.257 0.797187

provinceLiaoning 39.85 84.78 0.470 0.638383

provinceNingxia 177.68 105.25 1.688 0.091508 .

provinceQinghai -98.26 169.42 -0.580 0.561962

provinceShaanxi -134.84 82.63 -1.632 0.102840

provinceShandong -292.52 69.53 -4.207 2.68e-05 \*\*\*

provinceShanghai -71.59 124.12 -0.577 0.564113

provinceShanxi -154.02 76.16 -2.022 0.043253 \*

provinceSichuan -92.70 96.17 -0.964 0.335190

provinceTianjin -242.70 128.17 -1.894 0.058395 .

provinceXinjiang 308.56 83.30 3.704 0.000216 \*\*\*

provinceYunnan -167.32 102.88 -1.626 0.103990

provinceZhejiang -292.53 73.76 -3.966 7.50e-05 \*\*\*

---

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

Residual standard error: 519.6 on 2567 degrees of freedom

Multiple R-squared: 0.2728, Adjusted R-squared: 0.2646

F-statistic: 33.21 on 29 and 2567 DF, p-value: < 2.2e-16

> summary(model\_SO2\_province)$coef

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

(Intercept) 377.38154 64.45388 5.8550630 5.376246e-09

provinceBeijing -360.08154 523.62581 -0.6876696 4.917230e-01

provinceChongqing 223.05483 111.07214 2.0081977 4.472686e-02

provinceFujian -197.09206 94.29563 -2.0901506 3.670244e-02

provinceGansu 337.39777 116.04168 2.9075568 3.673889e-03

provinceGuangdong -66.50136 81.16035 -0.8193823 4.126444e-01

provinceGuangxi 333.44008 107.01590 3.1157995 1.854745e-03

provinceGuizhou 2280.59538 120.58217 18.9132054 8.574295e-75

provinceHainan -136.05427 169.41799 -0.8030686 4.220095e-01

provinceHebei -243.12919 79.14487 -3.0719515 2.148952e-03

provinceHeilongjiang -154.58064 81.16035 -1.9046324 5.693954e-02

provinceHenan -246.21988 78.64224 -3.1308861 1.762406e-03

provinceHubei -156.89126 88.90847 -1.7646380 7.774352e-02

provinceHunan -26.75227 103.63587 -0.2581372 7.963217e-01

provinceInner Mongol 415.25266 76.78822 5.4077654 6.970359e-08

provinceJiangsu -211.55531 72.53263 -2.9166917 3.568347e-03

provinceJiangxi 54.19288 102.14726 0.5305368 5.957857e-01

provincejilin -24.98350 97.20598 -0.2570161 7.971870e-01

provinceLiaoning 39.84992 84.78410 0.4700164 6.383832e-01

provinceNingxia 177.68000 105.25275 1.6881269 9.150833e-02

provinceQinghai -98.26336 169.41799 -0.5800054 5.619619e-01

provinceShaanxi -134.83896 82.63089 -1.6318227 1.028395e-01

provinceShandong -292.51580 69.53041 -4.2070193 2.676399e-05

provinceShanghai -71.59404 124.11906 -0.5768175 5.641134e-01

provinceShanxi -154.02178 76.16315 -2.0222612 4.325290e-02

provinceSichuan -92.70041 96.17278 -0.9638945 3.351896e-01

provinceTianjin -242.70427 128.17324 -1.8935642 5.839522e-02

provinceXinjiang 308.55867 83.29534 3.7043928 2.163685e-04

provinceYunnan -167.31725 102.87653 -1.6263889 1.039897e-01

provinceZhejiang -292.53487 73.75743 -3.9661749 7.503349e-05

> SO2\_model\_1 <- lmer(SO2 ~ type + unit + capacity +(1|province), data = df, REML = FALSE)

> summary(SO2\_model\_1)

Linear mixed model fit by maximum likelihood ['lmerMod']

Formula: SO2 ~ type + unit + capacity + (1 | province)

Data: df

AIC BIC logLik deviance df.resid

38835.2 38882.1 -19409.6 38819.2 2589

Scaled residuals:

Min 1Q Median 3Q Max

-4.8777 -0.3069 -0.0073 0.1971 11.8565

Random effects:

Groups Name Variance Std.Dev.

province (Intercept) 153743 392.1

Residual 173598 416.7

Number of obs: 2597, groups: province, 30

Fixed effects:

Estimate Std. Error t value

(Intercept) 83.32583 88.71368 0.939

typeCoal 25.44723 51.50993 0.494

typeGas & Oil -201.96170 65.29817 -3.093

typeOthers 79.86000 55.56651 1.437

unit 51.63875 6.21195 8.313

capacity 0.43736 0.01482 29.517

Correlation of Fixed Effects:

(Intr) typeCl typG&O typOth unit

typeCoal -0.544

typeGas&Oil -0.440 0.764

typeOthers -0.498 0.877 0.700

unit -0.106 -0.039 -0.025 -0.062

capacity 0.031 -0.083 -0.059 0.018 -0.407