Midterm 2

Yanlin Li

Read in the data

```
library(tidyverse)
library(tidymodels)
library(car)

beijing <- read_csv("beijing.csv")
```

Exercise 1

```
m1 <- lm(SO2 ~ as.factor(month) + TEMP + PRES + DEWP + RAIN + wd + WSPM, data = beijing)
summary(m1)
```

```
Call:
lm(formula = SO2 ~ as.factor(month) + TEMP + PRES + DEWP + RAIN +
    wd + WSPM, data = beijing)

Residuals:
   Min     1Q Median     3Q    Max
-46.07 -10.34  -3.15   5.14 369.93

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)        585.59581   22.13774  26.452  < 2e-16 ***
as.factor(month)2   -6.05484    0.54688 -11.072  < 2e-16 ***
as.factor(month)3  -11.53889    0.57517 -20.062  < 2e-16 ***
as.factor(month)4  -33.53894    0.68180 -49.192  < 2e-16 ***
as.factor(month)5  -39.75737    0.79914 -49.750  < 2e-16 ***
```

```
as.factor(month)6  -52.26176    0.91097 -57.369  < 2e-16 ***
as.factor(month)7  -58.11545    0.97784 -59.433  < 2e-16 ***
as.factor(month)8  -57.08227    0.96524 -59.138  < 2e-16 ***
as.factor(month)9  -48.63633    0.85637 -56.794  < 2e-16 ***
as.factor(month)10 -40.71187    0.70798 -57.504  < 2e-16 ***
as.factor(month)11 -26.09438    0.57222 -45.602  < 2e-16 ***
as.factor(month)12  -6.99803    0.52777 -13.260  < 2e-16 ***
TEMP                -0.00214    0.02741  -0.078 0.937777
PRES                -0.52554    0.02174 -24.177  < 2e-16 ***
DEWP                 0.37997    0.02150  17.673  < 2e-16 ***
RAIN                -0.21118    0.13549  -1.559 0.119078
wdENE               -2.65011    0.52376  -5.060 4.22e-07 ***
wdESE               -1.34770    0.60842  -2.215 0.026760 *
wdN                 -6.36749    0.70254  -9.064  < 2e-16 ***
wdNE                -3.73345    0.50869  -7.339 2.19e-13 ***
wdNNE               -4.16745    0.63549  -6.558 5.54e-11 ***
wdNNW               -7.67187    0.70900 -10.821  < 2e-16 ***
wdNW                -6.64408    0.56998 -11.657  < 2e-16 ***
wdS                  1.85917    0.64462   2.884 0.003927 **
wdSE                -1.28733    0.67860  -1.897 0.057830 .
wdSSE                0.76882    0.73503   1.046 0.295579
wdSSW                2.46929    0.56153   4.397 1.10e-05 ***
wdSW                 0.72196    0.51848   1.392 0.163793
wdW                  1.29017    0.55974   2.305 0.021174 *
wdWNW               -3.37404    0.56671  -5.954 2.65e-09 ***
wdWSW                2.12853    0.55268   3.851 0.000118 ***
WSPM                -2.51058    0.11290 -22.237  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 19.88 on 34279 degrees of freedom
  (753 observations deleted due to missingness)
Multiple R-squared:  0.312, Adjusted R-squared:  0.3114
F-statistic: 501.4 on 31 and 34279 DF,  p-value: < 2.2e-16
```