

Lab2_Modeling

2023-12-04

Imports

```
train <- read_csv("./data/processed/train_data.csv")
test<- read_csv("./data/processed/test_data.csv")
```

Basic Model

```
modell1 <- lm(total_sales ~ is_vegetarian + ingredeint_count + month_02 + month_03 + month_04 + month_05
summary(modell1)
```

```
##
## Call:
## lm(formula = total_sales ~ is_vegetarian + ingredeint_count +
##      month_02 + month_03 + month_04 + month_05 + month_06 + month_07 +
##      month_08 + month_09 + month_10 + month_11 + month_12, data = train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1441.09  -661.64   -17.75   571.82  1823.19
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2323.927    226.023   10.282 < 2e-16 ***
## is_vegetarianTRUE -372.624    132.092   -2.821  0.00514 **
## ingredeint_count    -9.181     31.516   -0.291  0.77103
## month_02    -143.517    222.773   -0.644  0.51997
## month_03    -111.773    225.985   -0.495  0.62128
## month_04      11.209    216.969    0.052  0.95884
## month_05       8.402    233.195    0.036  0.97128
## month_06     -77.125    207.834   -0.371  0.71086
## month_07     154.823    212.084    0.730  0.46602
## month_08    -105.595    217.218   -0.486  0.62727
## month_09     -62.290    219.794   -0.283  0.77709
## month_10    -186.236    207.811   -0.896  0.37096
## month_11     -36.863    212.176   -0.174  0.86220
## month_12    -187.832    214.465   -0.876  0.38191
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 777 on 270 degrees of freedom
## Multiple R-squared:  0.04588,    Adjusted R-squared:  -5.92e-05
## F-statistic: 0.9987 on 13 and 270 DF,  p-value: 0.4527
```

Advanced Model

```
model2 <- lm(total_sales ~ meat_count + ingredeint_count + alt_sauce + alt_cheese + month_02 + month_03  
summary(model2)
```

```
##  
## Call:  
## lm(formula = total_sales ~ meat_count + ingredeint_count + alt_sauce +  
##     alt_cheese + month_02 + month_03 + month_04 + month_05 +  
##     month_06 + month_07 + month_08 + month_09 + month_10 + month_11 +  
##     month_12, data = train)  
##  
## Residuals:  
##      Min       1Q   Median       3Q      Max   
## -1515.05  -587.40   -74.12   521.74  2075.87   
##  
## Coefficients:  
##              Estimate Std. Error t value Pr(>|t|)      
## (Intercept)    2373.95     233.30   10.176  <2e-16 ***  
## meat_count       59.93       64.26    0.933   0.3519      
## ingredeint_count -38.00      37.16   -1.022   0.3075      
## alt_sauceTRUE    140.97     109.52    1.287   0.1991      
## alt_cheeseTRUE  -201.91      96.38   -2.095   0.0371 *    
## month_02        -180.32     223.95   -0.805   0.4214      
## month_03        -95.95     227.05   -0.423   0.6729      
## month_04         14.71     218.12    0.067   0.9463      
## month_05        -10.95     235.00   -0.047   0.9629      
## month_06        -91.57     209.00   -0.438   0.6616      
## month_07         154.21     213.29    0.723   0.4703      
## month_08        -124.53     218.43   -0.570   0.5691      
## month_09        -102.18     221.62   -0.461   0.6451      
## month_10        -193.98     208.91   -0.928   0.3540      
## month_11         -33.31     213.35   -0.156   0.8761      
## month_12        -204.68     215.67   -0.949   0.3434      
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 780.8 on 268 degrees of freedom  
## Multiple R-squared:  0.04385,    Adjusted R-squared:  -0.009671   
## F-statistic: 0.8193 on 15 and 268 DF,  p-value: 0.6558
```