

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
## -1571.08  -596.58    6.15   558.45  1716.67
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2385.920    224.191   10.642 < 2e-16 ***
## is_vegetarianTRUE -438.454    107.696  -4.071 6.2e-05 ***
## ingredeint_count     2.151     32.479   0.066  0.947
## month_02    -323.918    219.296  -1.477   0.141
## month_03     -66.244    230.621  -0.287   0.774
## month_04     -73.868    205.646  -0.359   0.720
## month_05    -122.265    222.340  -0.550   0.583
## month_06    -121.026    207.991  -0.582   0.561
## month_07      11.454    205.659   0.056   0.956
## month_08    -159.743    216.041  -0.739   0.460
## month_09    -235.442    205.646  -1.145   0.253
## month_10    -305.420    201.417  -1.516   0.131
## month_11     -85.851    212.991  -0.403   0.687
## month_12    -283.812    212.926  -1.333   0.184
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 753.1 on 262 degrees of freedom
## Multiple R-squared:  0.08712,    Adjusted R-squared:  0.04183
## F-statistic: 1.923 on 13 and 262 DF,  p-value: 0.02783
```

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   
## -1398.09  -633.62    38.06   534.01  1926.44   
##  
## Coefficients:  
##              Estimate Std. Error t value Pr(>|t|)      
## (Intercept)    2256.57     237.68   9.494 < 2e-16 ***  
## meat_count      170.03      55.54   3.061 0.00243 **  
## ingredeint_count -15.50     35.36  -0.438 0.66144      
## alt_sauceTRUE    189.21     132.75   1.425 0.15527      
## alt_cheeseTRUE  -276.82     109.88  -2.519 0.01236 *  
## month_02        -329.84     215.84  -1.528 0.12769      
## month_03        -122.10     227.65  -0.536 0.59216      
## month_04         -41.42     202.57  -0.204 0.83813      
## month_05        -81.09     219.40  -0.370 0.71197      
## month_06        -97.20     204.75  -0.475 0.63537      
## month_07         42.47     202.85   0.209 0.83432      
## month_08        -162.68     212.84  -0.764 0.44537      
## month_09        -234.47     202.48  -1.158 0.24792      
## month_10        -303.50     198.39  -1.530 0.12727      
## month_11        -30.32     209.96  -0.144 0.88527      
## month_12        -282.85     209.62  -1.349 0.17840      
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 741.4 on 260 degrees of freedom  
## Multiple R-squared:  0.1221, Adjusted R-squared:  0.07143   
## F-statistic: 2.41 on 15 and 260 DF, p-value: 0.002726
```