Jeremy Long

Professor De los Santos

ECON 8710

2/4/2022

Forecasting Challenge 1

Method

Python- I started by creating a python program to analyze the data. I made a plot of the existing data plotting sales over time in months. It seemed to be mostly linear though increasing more towards the end of the dataset. I found that the First 2 months of the year were consistently lower in sales than the others, and that the last month was significantly higher. Knowing this, I created dummy variables to indicate whether a given month was one of the "First Two" months of the year or the "Last" month. I populated a data frame with these dummies, the months, and sales.

R Studio- I performed a few regressions in R Studio to find the best way to predict the data. The model I chose featured the following regressors: 'Months,' 'Months^2,' Dummy for 'First 2' months, Dummy for the 'Last' month, an interaction between 'Months' and the 'First 2' Dummy, and an interaction between 'Months^2' and the 'Last' month. The summary of my regression is at the bottom of the page.

Excel- I entered the months and dummy variables for 2021 into an excel sheet to predict the next year. I had a column that executed the regression and yielded the results shown below.

```
Call:
lm(formula = dat$Sales ~ dat$Months + I(dat$Months^2) + dat$M12 +
    dat$First2 + I(dat$Months^2):dat$M12 + dat$Months:dat$First2)

Residuals:
     Min       1Q   Median       3Q      Max
-13.2104  -2.1796   0.1103   2.7309   9.8284

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)              9.759e+01  1.144e+00  85.280  < 2e-16 ***
dat$Months               4.990e-01  3.452e-02  14.454  < 2e-16 ***
I(dat$Months^2)          8.849e-04  2.284e-04   3.874 0.000165 ***
dat$M12                  2.243e+01  1.951e+00  11.496  < 2e-16 ***
dat$First2              -1.296e+01  1.822e+00  -7.116 5.65e-11 ***
I(dat$Months^2):dat$M12  8.555e-04  1.922e-04   4.452 1.75e-05 ***
dat$Months:dat$First2   -8.351e-02  2.278e-02  -3.666 0.000351 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.163 on 137 degrees of freedom
Multiple R-squared:  0.9809,    Adjusted R-squared:   0.98
F-statistic:  1171 on 6 and 137 DF,  p-value: < 2.2e-16
```

Regression data from R Studio

date	sales
2021m1	163.4811
2021m2	164.1541
2021m3	190.0648
2021m4	190.8248
2021m5	191.5867
2021m6	192.3503
2021m7	193.1156
2021m8	193.8827
2021m9	194.6516
2021m10	195.4223
2021m11	196.1947
2021m12	240.2184

Forecast for the following year