Capstone Modeling

Qiaojuan Tu

11/3/2021

```r
#clean the environment
rm(list = ls())

#Load packages
library(data.table)
library(Metrics)
library(readxl)
library(ggplot2)

#Load Data
Stimulus_by_states <- read_excel("Checks_Amount.xlsx")
CPI <- read_excel("CPIAUCSL.xlsx") #national CPI
master <- merge(CPI, Stimulus_by_states)
CPI_by_States <- read_excel("CPI_by_States.xlsx")

#manipulate the data
South <- subset(Stimulus_by_states, Region == "South")
West  <- subset(Stimulus_by_states, Region == "West")
Midwest <- subset(Stimulus_by_states, Region == "Midwest")
Northeast <- subset(Stimulus_by_states, Region == "Northeast")

South_CPI <- subset(CPI_by_States, Region == "South")
West_CPI <- subset(CPI_by_States, Region == "West")
Midwest_CPI <- subset(CPI_by_States, Region == "Midwest")
Northeast_CPI <- subset(CPI_by_States, Region == "Northeast")

#For South Region
sum(South$Total_r1)/17
```

```
## [1] 6127664
```

```r
sum(South$Total_r2)/17
```

```
## [1] 3212521
```

```r
sum(South$Total_r3)/17
```

```
## [1] 8932882
```

```r
South_CPI_R1 <- South_CPI[1:12,]
South_CPI_R1$Stimulus <- sum(South$Total_r1)/17
South_CPI_R2 <- South_CPI[13:14,]
South_CPI_R2$Stimulus <- sum(South$Total_r2)/17
South_CPI_R3 <- South_CPI[15:18,]
South_CPI_R3$Stimulus <- sum(South$Total_r3)/17
```

```
South_Final <- rbind(South_CPI_R1, South_CPI_R2, South_CPI_R3)

#For West Region
sum(West$Total_r1)/13
```

[1] 4753503

```
sum(West$Total_r2)/13
```

[1] 2483845

```
sum(West$Total_r3)/13
```

[1] 6878794

```
West_CPI_R1 <- West_CPI[1:12,]
West_CPI_R1$Stimulus <- sum(West$Total_r1)/13
West_CPI_R2 <- West_CPI[13:14,]
West_CPI_R2$Stimulus <- sum(West$Total_r2)/13
West_CPI_R3 <- West_CPI[15:18,]
West_CPI_R3$Stimulus <- sum(West$Total_r3)/13

West_Final <- rbind(West_CPI_R1, West_CPI_R2, West_CPI_R3)

#For Midwest Region
sum(Midwest$Total_r1)/12
```

[1] 4850485

```
sum(Midwest$Total_r2)/12
```

[1] 2550871

```
sum(Midwest$Total_r3)/12
```

[1] 6874461

```
Midwest_CPI_R1 <- Midwest_CPI[1:12,]
Midwest_CPI_R1$Stimulus <- sum(Midwest$Total_r1)/12
Midwest_CPI_R2 <- Midwest_CPI[13:14,]
Midwest_CPI_R2$Stimulus <- sum(Midwest$Total_r2)/12
Midwest_CPI_R3 <- Midwest_CPI[15:18,]
Midwest_CPI_R3$Stimulus <- sum(Midwest$Total_r3)/12

Midwest_Final <- rbind(Midwest_CPI_R1, Midwest_CPI_R2, Midwest_CPI_R3)

#For  Northeast Region
sum(Northeast$Total_r1)/9
```

[1] 5049879

```
sum(Northeast$Total_r2)/9
```

[1] 2574754

```
sum(Northeast$Total_r3)/9
```

[1] 7024241

```
Northeast_CPI_R1 <- Northeast_CPI[1:12,]
Northeast_CPI_R1$Stimulus <- sum(Northeast$Total_r1)/9
Northeast_CPI_R2 <- Northeast_CPI[13:14,]
Northeast_CPI_R2$Stimulus <- sum(Northeast$Total_r2)/9
Northeast_CPI_R3 <- Northeast_CPI[15:18,]
Northeast_CPI_R3$Stimulus <- sum(Northeast$Total_r3)/9


Northeast_Final <- rbind(Northeast_CPI_R1, Northeast_CPI_R2, Northeast_CPI_R3)


#Put together
Master <- rbind(South_Final, West_Final, Midwest_Final, Northeast_Final)
View(Master)
```

```
#create the linear model
fit <- glm(CPI~Stimulus, data = Master, family = "gaussian")
plot(fit)
```


Residuals vs Fitted

Normal Q–Q

Std. Pearson resid.

Theoretical Quantiles
glm(CPI ~ Stimulus)

Scale−Location

√|Std. Pearson resid.|

Predicted values
glm(CPI ~ Stimulus)

Residuals vs Leverage

Std. Pearson resid.

Leverage
glm(CPI ~ Stimulus)