**Q1: Download the data Consumer Price Index (CPI) data from Statistics Canada (table 18-10-0004-01, DOI: https://doi.org/10.25318/1810000401-eng). Tidy the table data and keep only the fields required to solve the questions below**

getwd() library(dplyr) library(tidyverse) library(ggplot2) library(zoo) #loading data

mydata = read.csv("consumerpriceindex.csv",strip.white=TRUE)

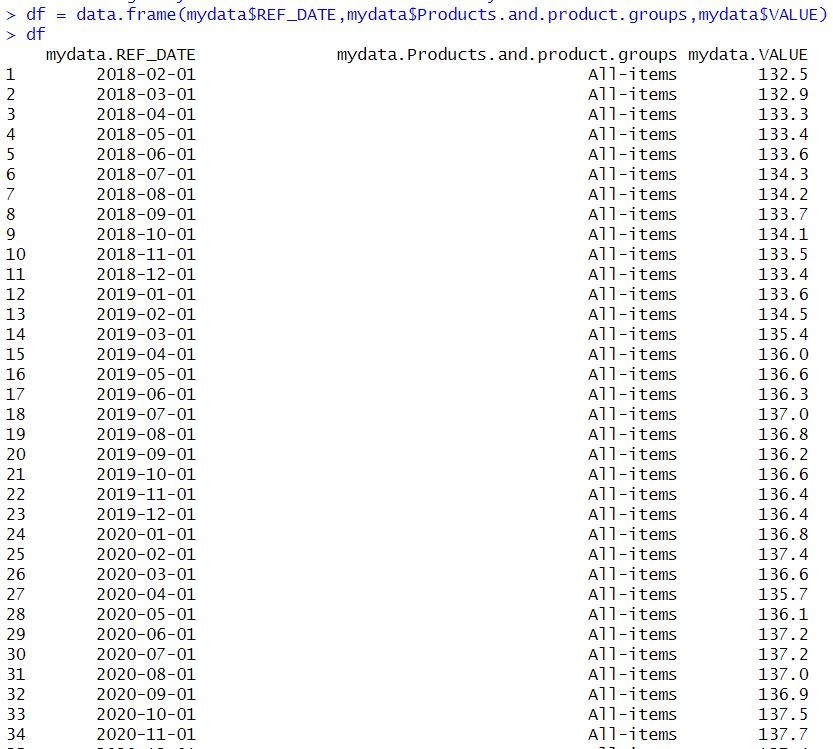
#Checking the type of date typeof(mydata$REF\_DATE) typeof(mydata$VALUE)

#Converting date to date format

mydata$REF\_DATE<- as.Date(paste0(mydata$REF\_DATE, "-01"))

# Selecting only columns relevant for the analysis

df = data.frame(mydata$REF\_DATE,mydata$Products.and.product.groups,mydata$VALUE) df



# The dataset used here is the consumer price index over the past 5 years i.e. from 2018February to 2023 February. Here the dataset is cleaned to format the REF\_DATE as Date format and selected only three columns ; REF\_DATE, Products.andproduct.groups and VALUE, relevant for the analysis.

**Q2: Quarterly (i.e every 3 months) change in the 1) overall CPI, and 2) CPI excluding food and energy; over the past 5 years (30%)**

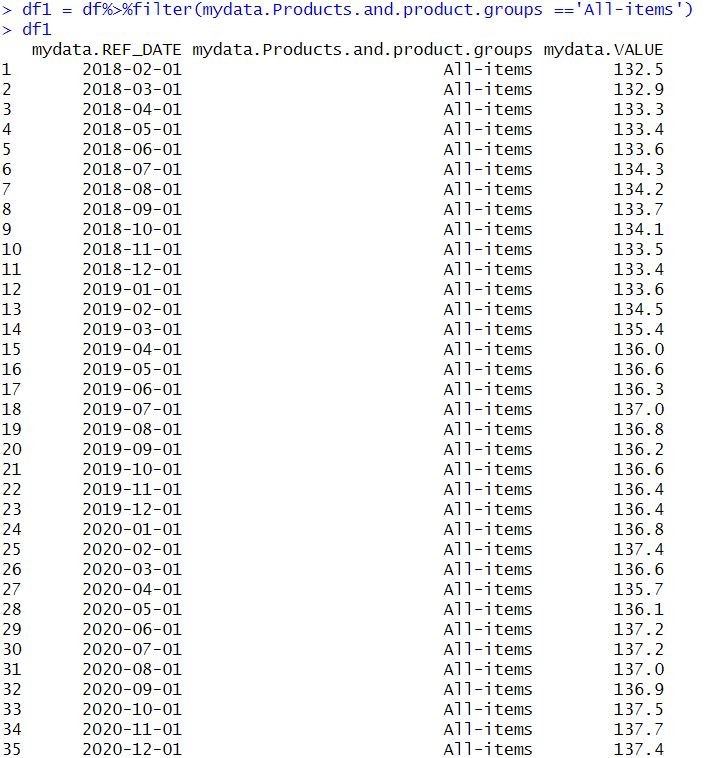


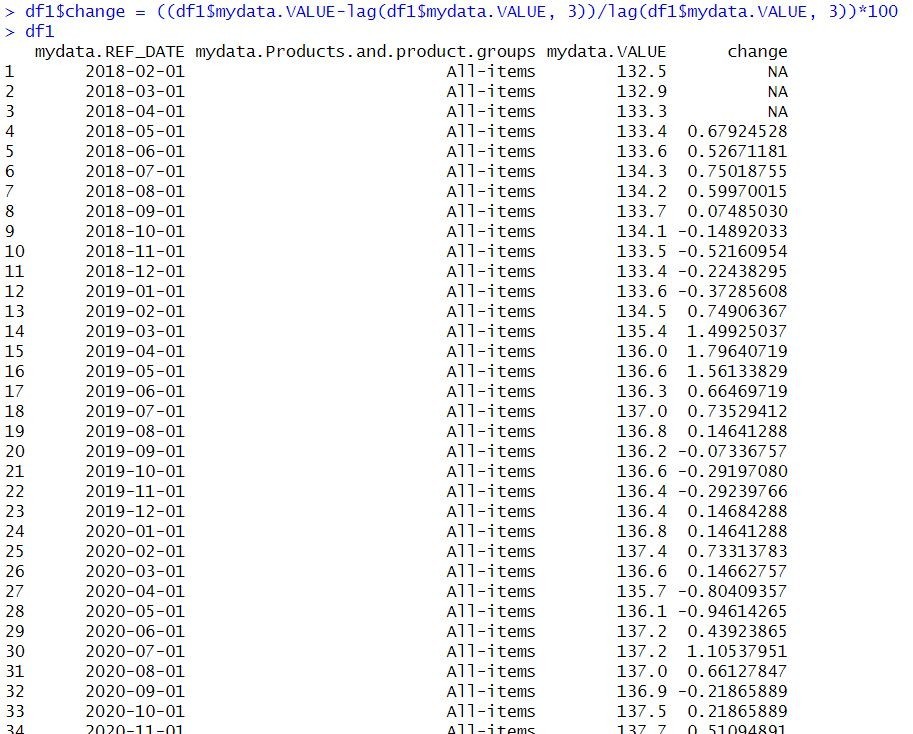
#Filtering the data for All items and calculating the change over 3 months(Quarterly)

df1 = df%>%filter (mydata.Products.and.product.groups =='All-items') df1

df1$change = ((df1$mydata.VALUE-lag(df1$mydata.VALUE, 3))/lag(df1$mydata.VALUE, 3))\*100

df1



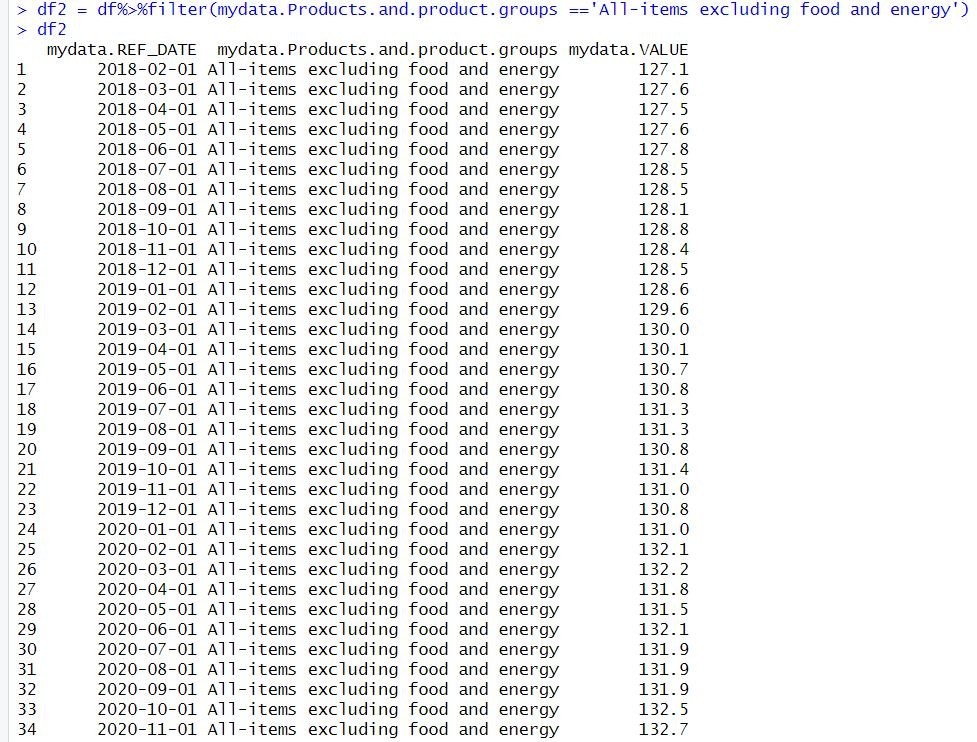


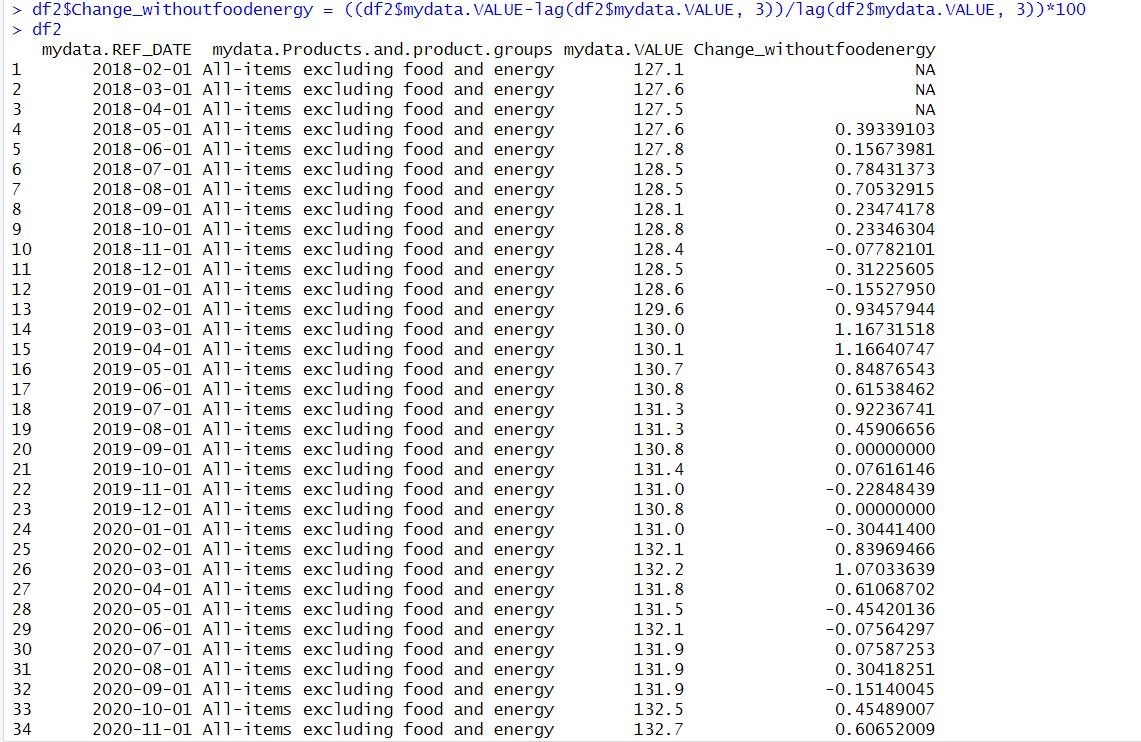
#Filtering the data for All items excluding food and energy and calculating the change over 3 months(Quarterly)

df2 = df%>%filter(mydata.Products.and.product.groups =='All-items excluding food and energy')

df2

df2$Change\_withoutfoodenergy = ((df2$mydata.VALUE- lag(df2$mydata.VALUE, 3))/lag(df2$mydata.VALUE, 3))\*100 df2





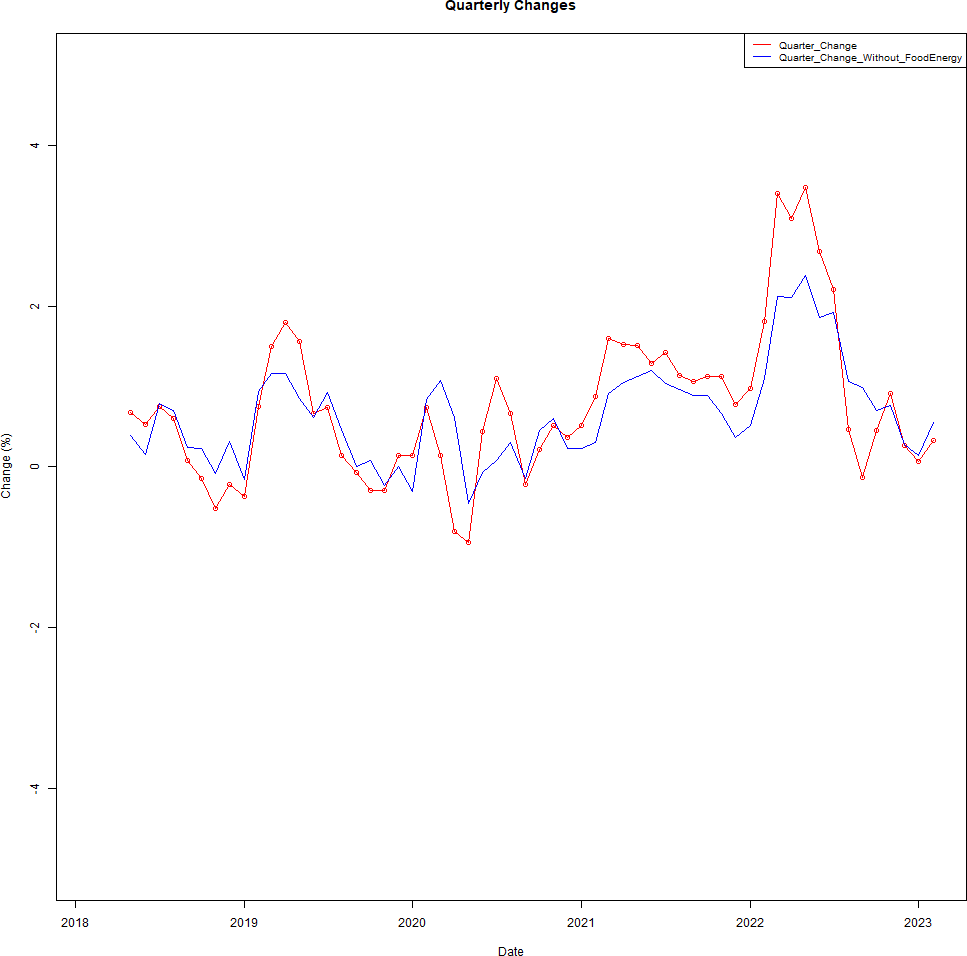
#Plotting the quarterly changes

png(filename="QuarterChanges.png", height=3000, width=3000, bg="white")

plot(df1$mydata.REF\_DATE, df1$Change,ylim = c(-20,20), type = "o", col = "red", xlab = "Date", ylab = "Change (%)", main = "Quarterly Changes") lines(df2$mydata.REF\_DATE, df2$Change\_withoutfoodenergy, col = "blue") legend("topright", legend = c("Quarter\_Change", "Quarter\_Change\_Without\_FoodEnergy"), col = c("red", "blue"), lty = 1, cex

= 0.8)

dev.off()



# Analysis

We can see that the quarterly changes for overall CPI and CPI excluding food and energy shows almost similar trends. Both shows great fluctuations over the last 5 years.

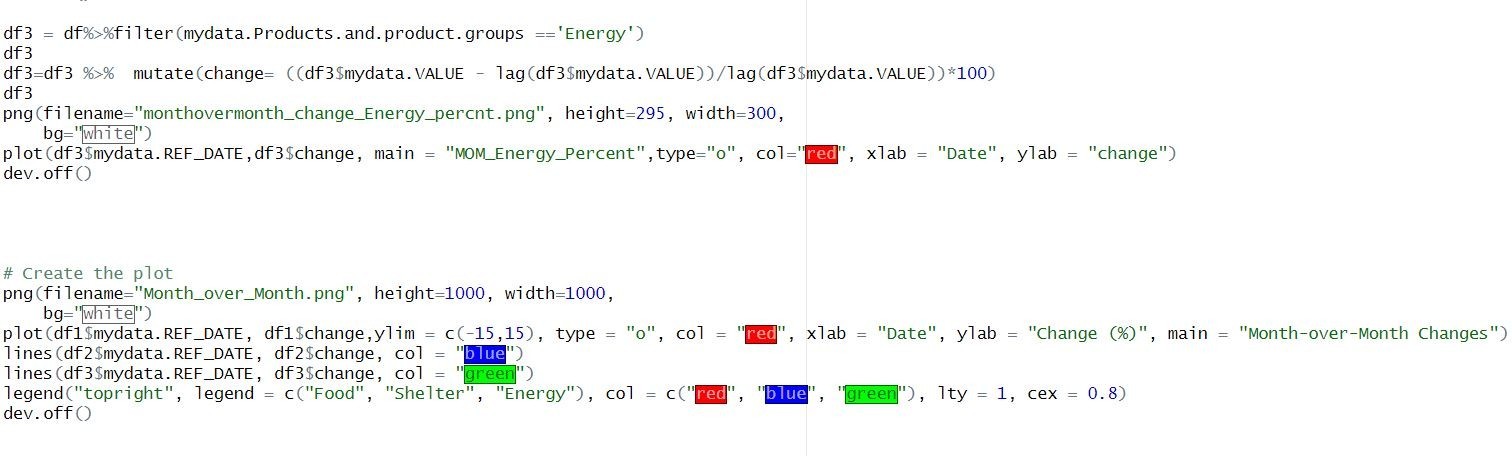
The change in overall CPI values started at nearly 0% around 2018 and then it decreased to below 0% and showed a rapid increase during 2019 and again it dropped significantly during 2020. It shows steady growth by the end of 2020 and by the beginning of 2021 it again increased. In 2022 the change in overall CPI significantly increased to nearly 3% at the beginning. By the end of 2022, the change decreased to below 0% and it again increased to above 0% by 2023.

The change in the overall CPI excluding food and energy followed a similar trend. In 2018 the change is nearly 0% but it decreased slightly by the end of 2018. The change then increased significantly to nearly 1% by the quarter of 2019. Then it decreased significantly by the beginning of 2020. The change showed steady increase with slight fluctuations throughout 2020 and 2021.

By the end of 2021 it decreased, then it increased rapidly to nearly 2% change by the mid of 2022. Then it decreased to nearly 0% change by 2023. But the change in Overall CPI is high almost every year than the Overall CPI without food and energy expect in very few times.

**Q3: Monthly price growth, in: 1) food 2) housing (shelter), and 3) energy (electricity, gas, fuels); over the past 18 months, showing the price index of each month as compared to: 1) last month (also called Month over month), and 2) Same month of last year (year over year) (50%).**





getwd() library(dplyr) library(tidyverse) library(ggplot2) library(zoo) #loading data

mydata = read.csv("consumerpriceindex.csv",strip.white=TRUE) #Checking the type of date

typeof(mydata$REF\_DATE) #Converting date to date format

mydata$REF\_DATE<- as.Date(paste0(mydata$REF\_DATE, "-01")) # Selecting only columns relevant for the analysis

df = data.frame(mydata$REF\_DATE,mydata$Products.and.product.groups,mydata

$VALUE)

#Converting the value to numeric format df$VALUE <- as.numeric(df$mydata.VALUE)

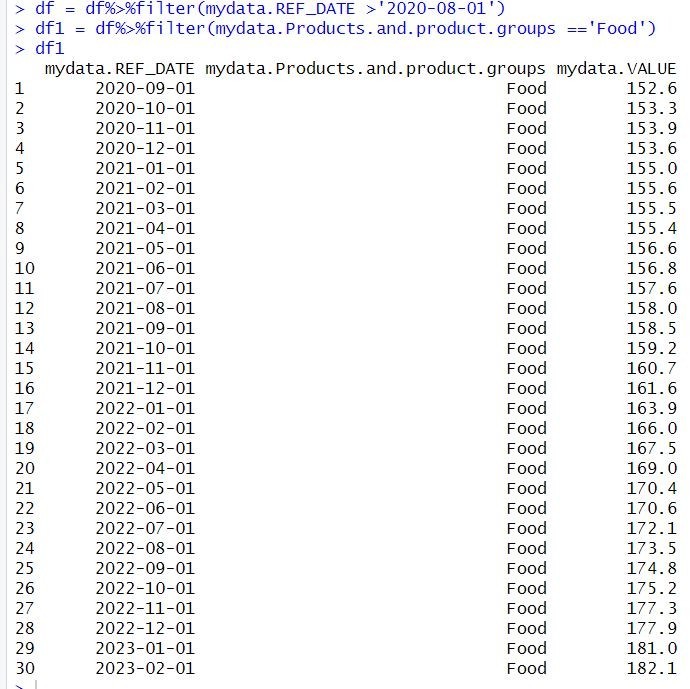
df = data.frame(mydata$REF\_DATE,mydata$Products.and.product.groups,mydata

$VALUE)

df

df = df%>%filter(mydata.REF\_DATE >'2020-08-01')

df1 = df%>%filter(mydata.Products.and.product.groups =='Food') df1

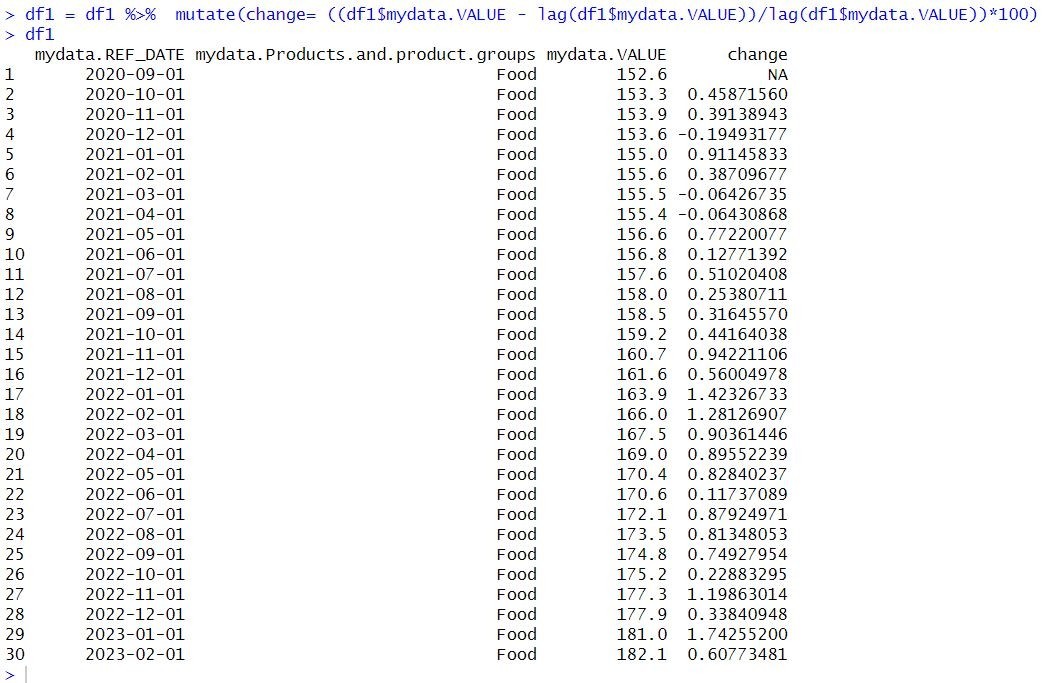


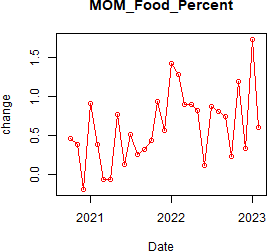
df1 = df1 %>% mutate(change= ((df1$mydata.VALUE - lag(df1$mydata.VALUE))/lag(df1$mydata.VALUE))\*100) df1

png(filename="monthovermonth\_change\_Food\_percnt.png", height=295, width=300,

bg="white") plot(df1$mydata.REF\_DATE,df1$change, main =

"MOM\_Food\_Percent",type="o", col="red", xlab = "Date", ylab = "change") dev.off()





df2 = df%>%filter(mydata.Products.and.product.groups =='Shelter') df2

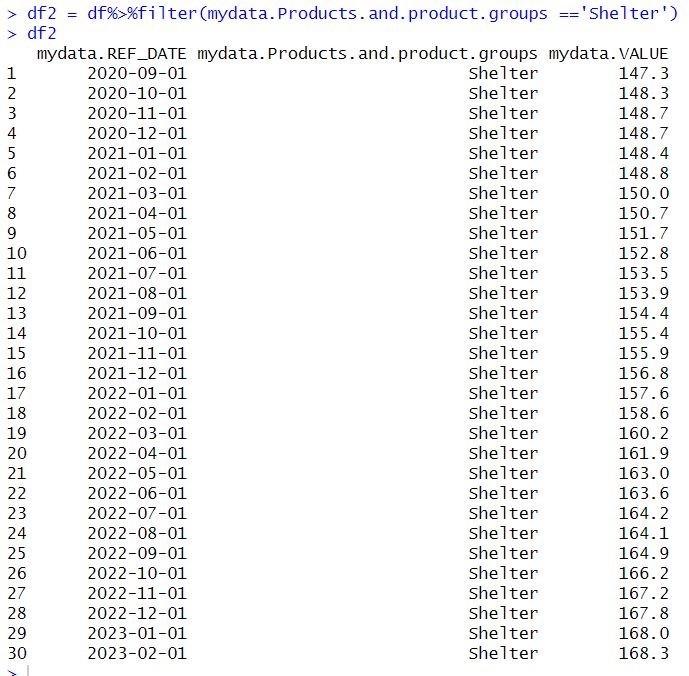
df2 = df2 %>% mutate(change= ((df2$mydata.VALUE - lag(df2$mydata.VALUE))/lag(df2$mydata.VALUE))\*100) df2

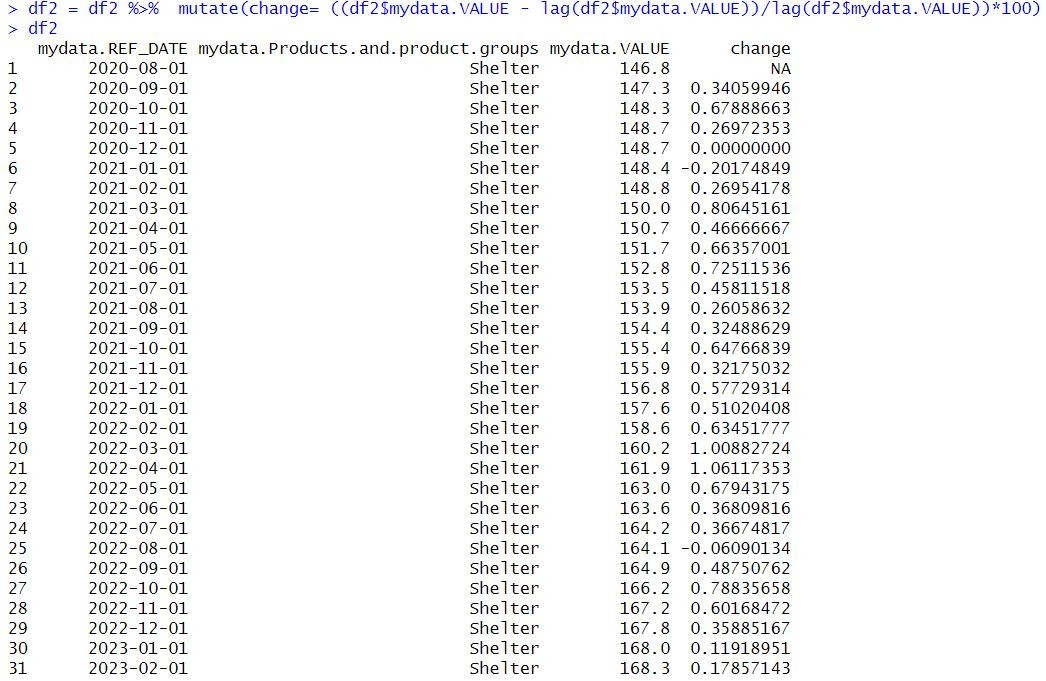
png(filename="monthovermonth\_change\_Shelter\_percent.png", height=295, width=300,

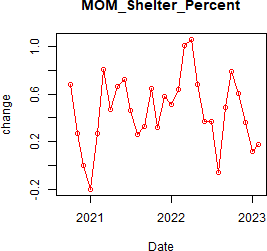
bg="white") plot(df2$mydata.REF\_DATE,df2$change,main =

"MOM\_Shelter\_Percent",type="o", col="red", xlab = "Date", ylab = "change")

dev.off()







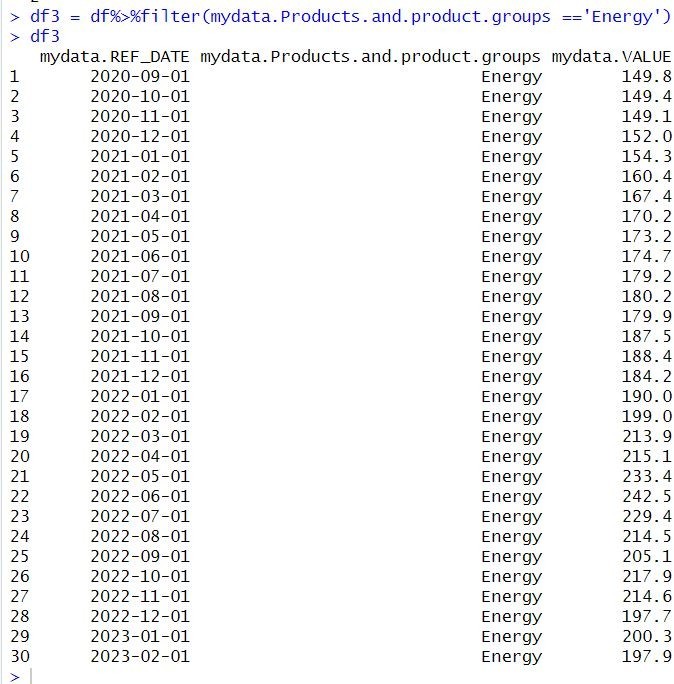
df3 = df%>%filter (mydata.Products.and.product.groups =='Energy') df3

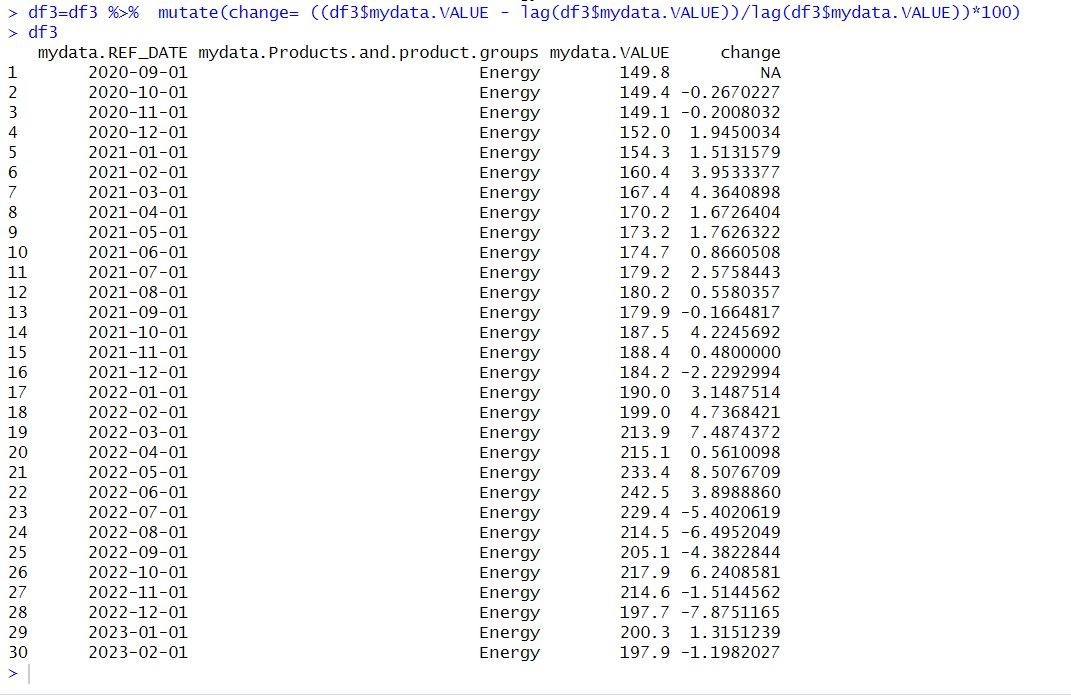
df3=df3 %>% mutate (change= ((df3$mydata.VALUE - lag(df3$mydata.VALUE))/lag(df3$mydata.VALUE))\*100) df3

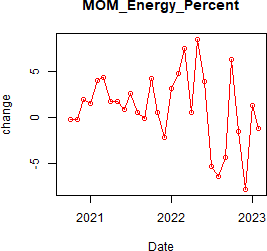
png (filename="monthovermonth\_change\_Energy\_percnt.png", height=295, width=300,

bg="white") plot(df3$mydata.REF\_DATE,df3$change, main =

"MOM\_Energy\_Percent",type="o", col="red", xlab = "Date", ylab = "change") dev.off()







# Create the plot

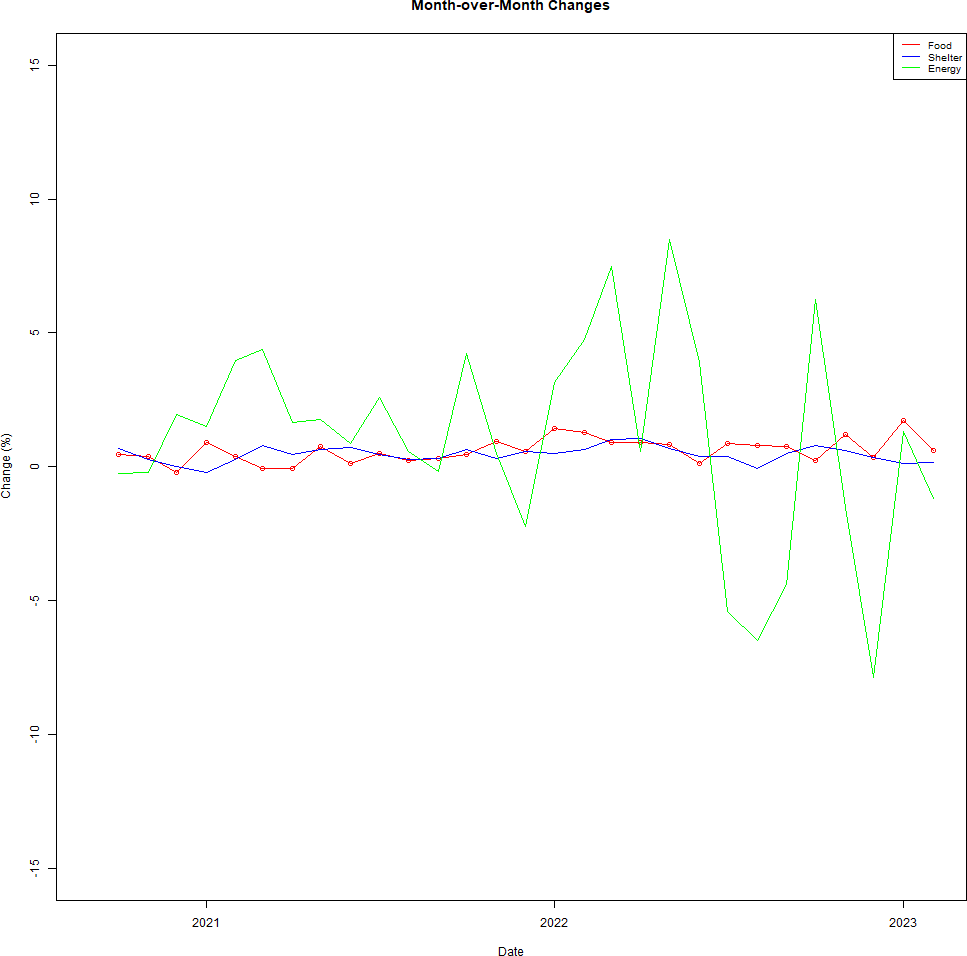
Png (filename="Month\_over\_Month.png", height=1000, width=1000, bg="white")

plot (df1$mydata.REF\_DATE, df1$change, ylim = c (-15,15), type = "o", col = "red", xlab = "Date", ylab = "Change (%)", main = "Month-over-Month Changes")

lines (df2$mydata.REF\_DATE, df2$change, col = "blue") lines (df3$mydata.REF\_DATE, df3$change, col = "green")

legend("topright", legend = c("Food", "Shelter", "Energy"), col = c("red", "blue", "green"), lty = 1, cex = 0.8)

dev.off()



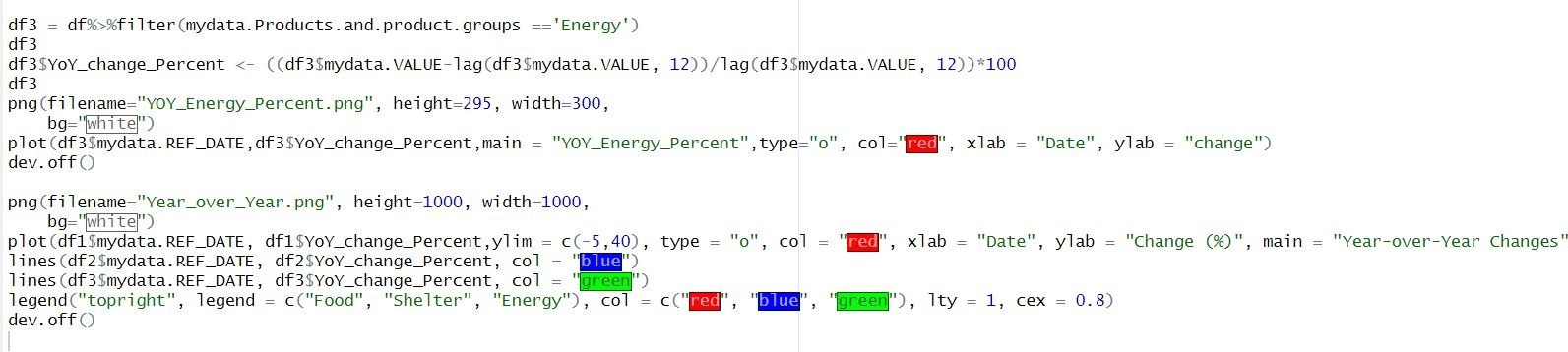
# Analysis

The line graph above describes the percentage change in the consumer price index month over month of food, shelter and energy in Canada over the last 30 months. For Food and Shelter we can see almost similar changes. In both cases the change in CPI is nearly 0% throughout the last 30 months. The change in Energy was consistently larger than the other two till the beginning of 2022. Then the change in energy fluctuates over the months of 2022 and of 2023.

The month over month percentage change in CPI of food was almost steady with slight increases and decreases periodically. The month over month percentage change in CPI of Shelter also showed a similar pattern.

But the month over month percentage change of energy begins at around 1% and then it shows slight variations till the last quarter of 2022. The year 2022 showed high variations in CPI between positive and negative changes. The change in CPI increased at the beginning of 2022 then it decreased and then it slumped to a negative change. By the middle of 2022, it again rose to a positive change and by 2023 beginning it decreased.

# Year Over Year Changes



getwd()

library(dplyr) library(tidyverse) library(ggplot2) library(zoo)

#loading data

mydata = read.csv("consumerpriceindex.csv",strip.white=TRUE)

#Checking the type of date typeof(mydata$REF\_DATE)

#Converting date to date format

mydata$REF\_DATE<- as.Date(paste0(mydata$REF\_DATE, "-01"))

# Selecting only columns relevant for the analysis df =

data.frame(mydata$REF\_DATE,mydata$Products.and.product.groups,mydata

$VALUE)

#Converting the value to numeric format df$VALUE <- as.numeric(df$mydata.VALUE)

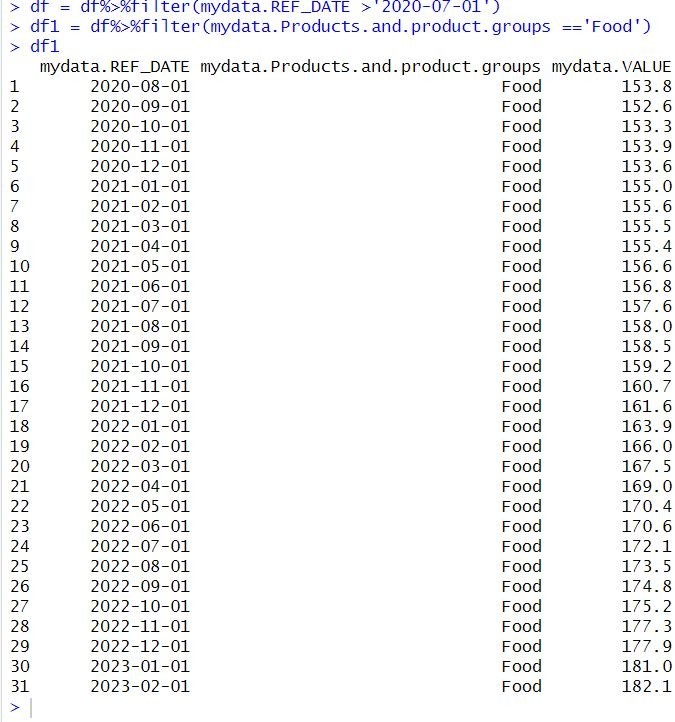
df = data.frame(mydata$REF\_DATE,mydata$Products.and.product.groups,mydata

$VALUE)

df

df = df%>%filter(mydata.REF\_DATE >'2020-08-01')

df1 = df%>%filter(mydata.Products.and.product.groups =='Food') df1

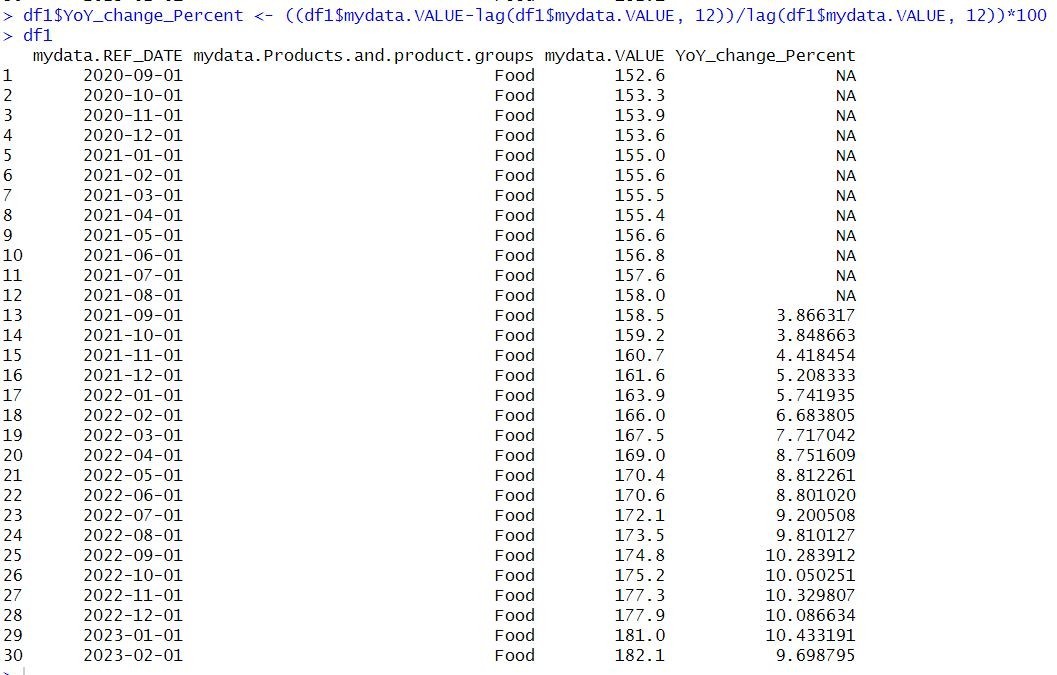


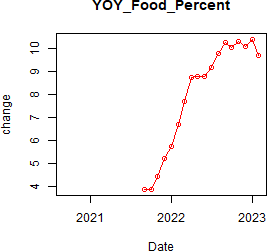
df1$YoY\_change\_Percent <- ((df1$mydata.VALUE-lag(df1$mydata.VALUE, 12))/lag(df1$mydata.VALUE, 12))\*100

df1

png(filename="YOY\_Food\_Percent.png", height=295, width=300, bg="white")

plot(df1$mydata.REF\_DATE,df1$YoY\_change\_Percent,main = "YOY\_Food\_Percent",type="o", col="red", xlab = "Date", ylab = "change") dev.off()





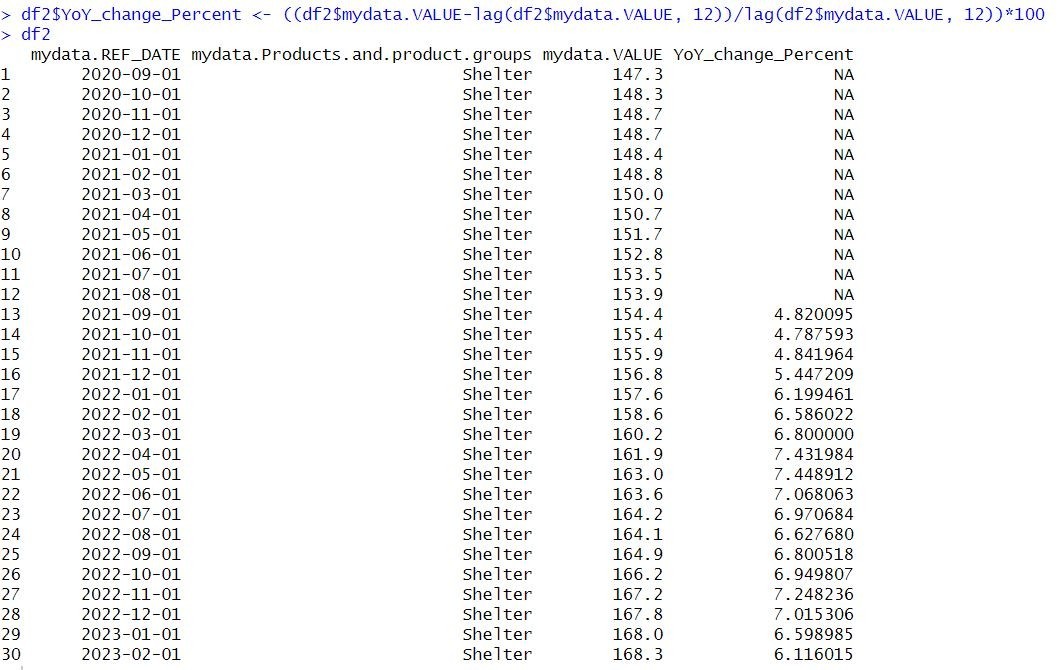
df2 = df%>%filter(mydata.Products.and.product.groups =='Shelter') df2

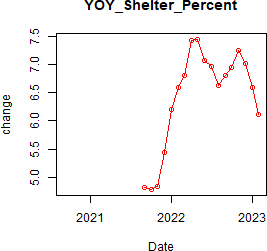
df2$YoY\_change\_Percent <- ((df2$mydata.VALUE-lag(df2$mydata.VALUE, 12))/lag(df2$mydata.VALUE, 12))\*100

df2

png(filename="YOY\_Shelter\_Percent.png", height=295, width=300, bg="white")

plot(df2$mydata.REF\_DATE,df2$YoY\_change\_Percent,main = "YOY\_Shelter\_Percent",type="o", col="red", xlab = "Date", ylab = "change") dev.off()





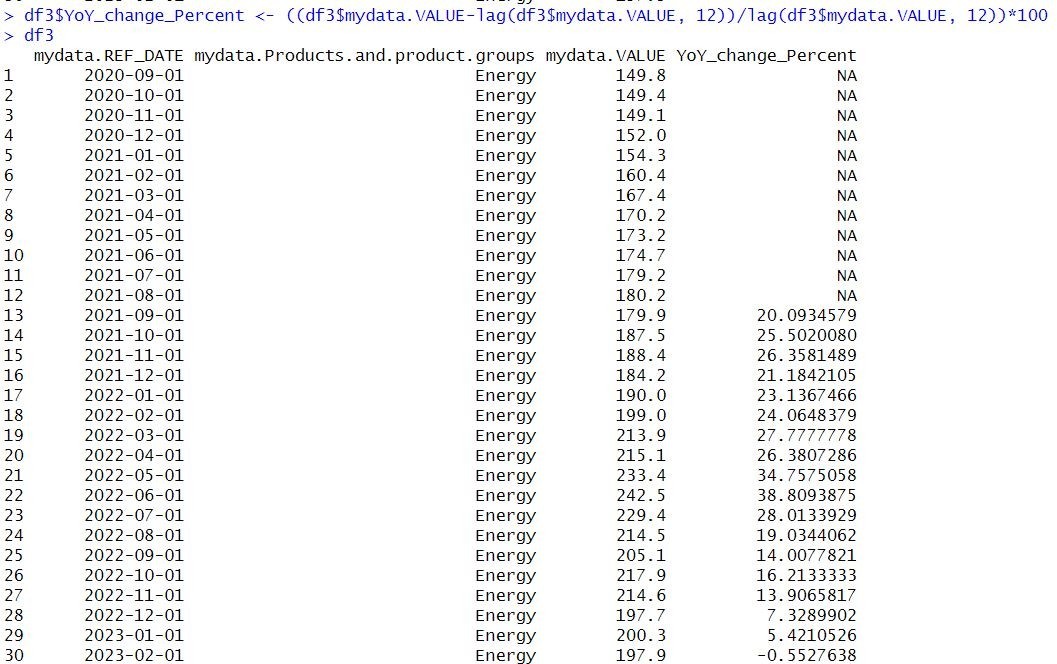
df3 = df%>%filter(mydata.Products.and.product.groups =='Energy') df3

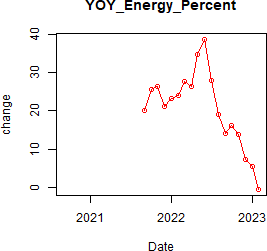
df3$YoY\_change\_Percent <- ((df3$mydata.VALUE-lag(df3$mydata.VALUE, 12))/lag(df3$mydata.VALUE, 12))\*100

df3

png(filename="YOY\_Energy\_Percent.png", height=295, width=300, bg="white")

plot(df3$mydata.REF\_DATE,df3$YoY\_change\_Percent,main = "YOY\_Energy\_Percent",type="o", col="red", xlab = "Date", ylab = "change") dev.off()





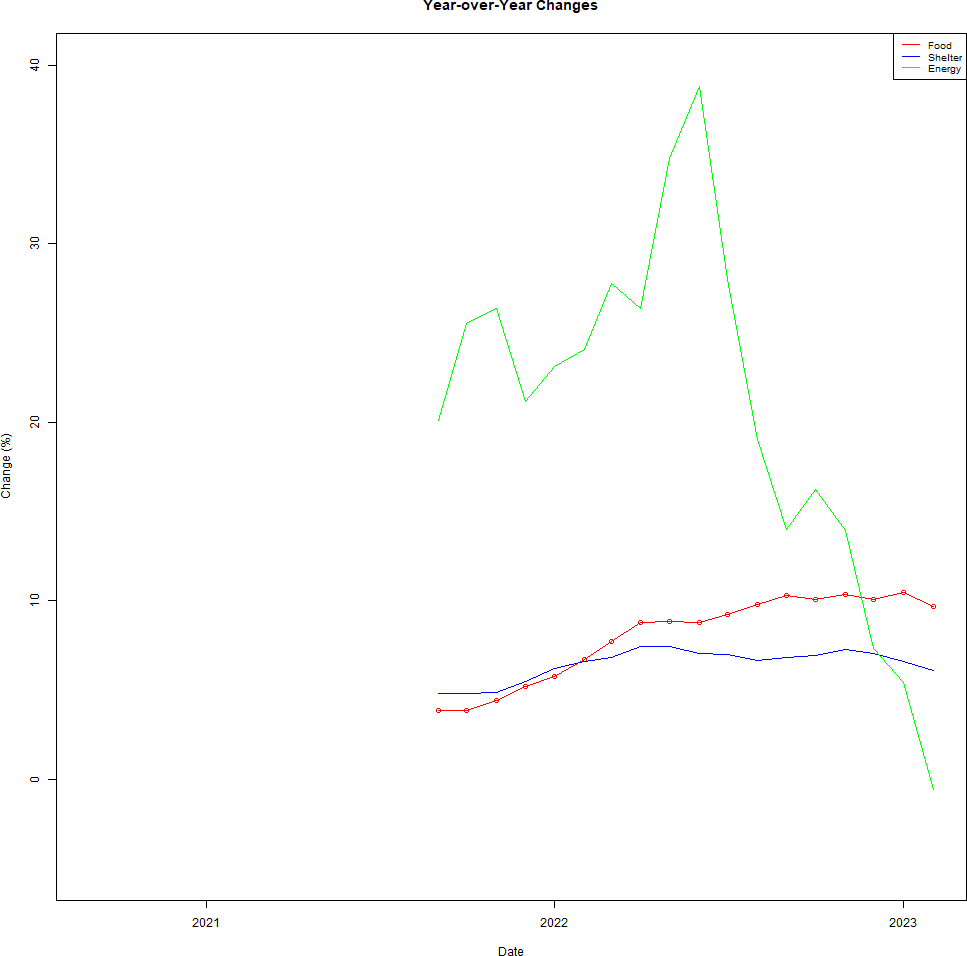
png(filename="Year\_over\_Year.png", height=1000, width=1000, bg="white")

plot(df1$mydata.REF\_DATE, df1$YoY\_change\_Percent,ylim = c(-15,15), type

= "o", col = "red", xlab = "Date", ylab = "Change (%)", main = "Year-over- Year Changes")

lines(df2$mydata.REF\_DATE, df2$YoY\_change\_Percent, col = "blue") lines(df3$mydata.REF\_DATE, df3$YoY\_change\_Percent, col = "green") legend("topright", legend = c("Food", "Shelter", "Energy"), col = c("red", "blue", "green"), lty = 1, cex = 0.8)

dev.off()



# Analysis

The plot clearly shows that the percentage change in Food and Shelter year over year over the last 18 months showed almost similar trend. The percentage change is always positive which indicates a rise in prices for food and shelter over the last 1.5 years as compared to the previous year. The

percentage change is higher for shelter at the beginning but the percentage change for food increases towards the end.

In the case of Energy, the percentage change is much higher as compared to Food and Shelter at the beginning. It shows a rapid increase in the middle of 2022 which indicates that there occurred a huge rise in price for energy in 2022 as compared to the same months in 2021. This may be due to external factors such as political conditions, war etc. After the middle of 2022 the percentage change in energy declines significantly. By the beginning of 2023 the change becomes negative which indicates a decrease in price.