Food Inflation Study

Jessica Woods

2023-12-20

```
knitr::opts_chunk$set(echo = TRUE)

library(glmnet)
```

```
## Loading required package: Matrix
```

```
## Loaded glmnet 4.1-8
```

```
library(MASS)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following object is masked from 'package:MASS':
##
##     select
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
# library(emergency_back_up_brain)
# library(coffee)
```

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

In this R document I will be specifically just cleaning data and correlating data points that are relavant to modeling and saving those back to a new CSV file to be visually assessed in Python and modeled in R.

Below is the loaded CSV files needed to start cleaning for modeling and analysis work. The link to get the food inflation information is https://microdata.worldbank.org/index.php/catalog/4483 from December 2019

to December 2023 to complete 4 years of data on internation food inflation. "Monthly food price estimates by product and market 25 countries, 1353 markets, 2007/01/01-2023/12/01, version 2023/12/11" -The World Bank of Microdata website. There are two other factors in the analysis work, that is conflict (gun violence and war) and the US stock Market prices from December 2019 to December 2023.

The current HO: "The occurrence of conflict in key food-producing regions, coupled with fluctuations in the US stock market, significantly influences international food inflation rates. Higher instances of conflict and volatility in the US stock market are expected to correlate with increased food inflation on a global scale." The current Alternative HO : "The impact of conflict and US stock market fluctuations on international food inflation rates may not exhibit a significant correlation."

```r
# Food inflation data
food_data <- read.csv("Food_inflation_2019_2023.csv", header=TRUE, sep=",")
# time to clean this up
clean_food_data <- subset(food_data, select = -X)
clean_food_data <-na.omit(clean_food_data)
clean_food_data <- subset(clean_food_data, select = -Market)
clean_food_data <-subset(clean_food_data, select = -Currency)
```

Narrowing down the top most conflicted countries in the past 4 years according to Wikipedia https://en.wikipedia.org/wiki/List_of_ongoing_armed_conflicts sited. Mexico, Ukraine, Afghanistan, Syria (Syrian Arab Republic), Ethiopia, Yemen

```r
conflicted_countries_food_data <- clean_food_data %>%
  filter(Country %in% c("Afghanistan", "Syrian Arab Republic", "Mexico", "Yemen", "Ukraine","Ethiopia"))
str(conflicted_countries_food_data)
```

```
## 'data.frame':    145236 obs. of  8 variables:
##  $ Country: chr  "Afghanistan" "Afghanistan" "Afghanistan" "Afghanistan" ...
##  $ Region : chr  "Badakhshan" "Badakhshan" "Badakhshan" "Badakhshan" ...
##  $ Product: chr  "bread" "bread" "bread" "bread" ...
##  $ Date   : chr  "2019-12-01" "2019-12-01" "2020-01-01" "2020-01-01" ...
##  $ Open   : num  49.8 49.8 49.8 49.8 49.9 ...
##  $ High   : num  49.9 49.9 49.9 49.9 50 ...
##  $ Low    : num  49.7 49.7 49.7 49.7 49.8 ...
##  $ Close  : num  49.8 49.8 49.9 49.9 49.9 ...
```

Afghanistan data on Food Inflation open and close, the difference between the two. Its important to look at the distribution of the data to do further analysis and cleaning of the data.

```r
data_afghanistan <- conflicted_countries_food_data[conflicted_countries_food_data$Country == "Afghanistan
open_values_afghanistan <- data_afghanistan$Open
closed_values_afghanistan <- data_afghanistan$Close
difference_afghanistan <- open_values_afghanistan - closed_values_afghanistan

# Create a sequence for x-axis (assuming you want a sequence of numbers as x-axis)
x <- seq(length(open_values_afghanistan))
x2 <-seq(length(closed_values_afghanistan))
# Plotting Open values for Afghanistan
plot(x, open_values_afghanistan, type = "l", xlab = "Data Points", ylab = "Open Values", main = "Open Va
```
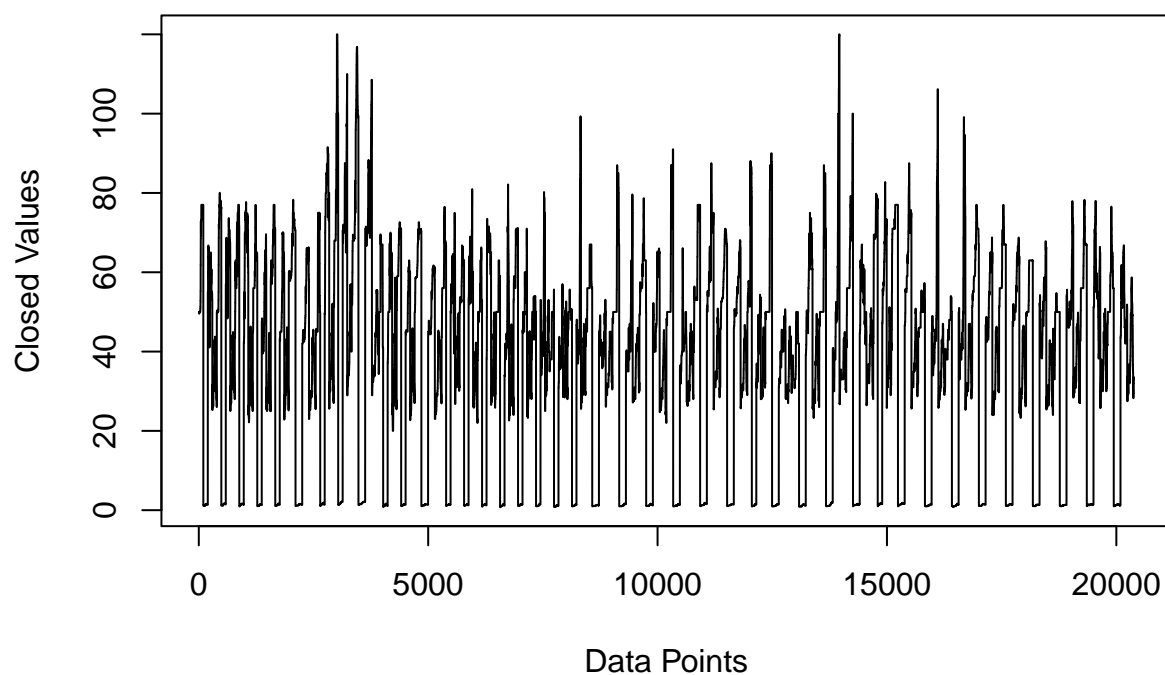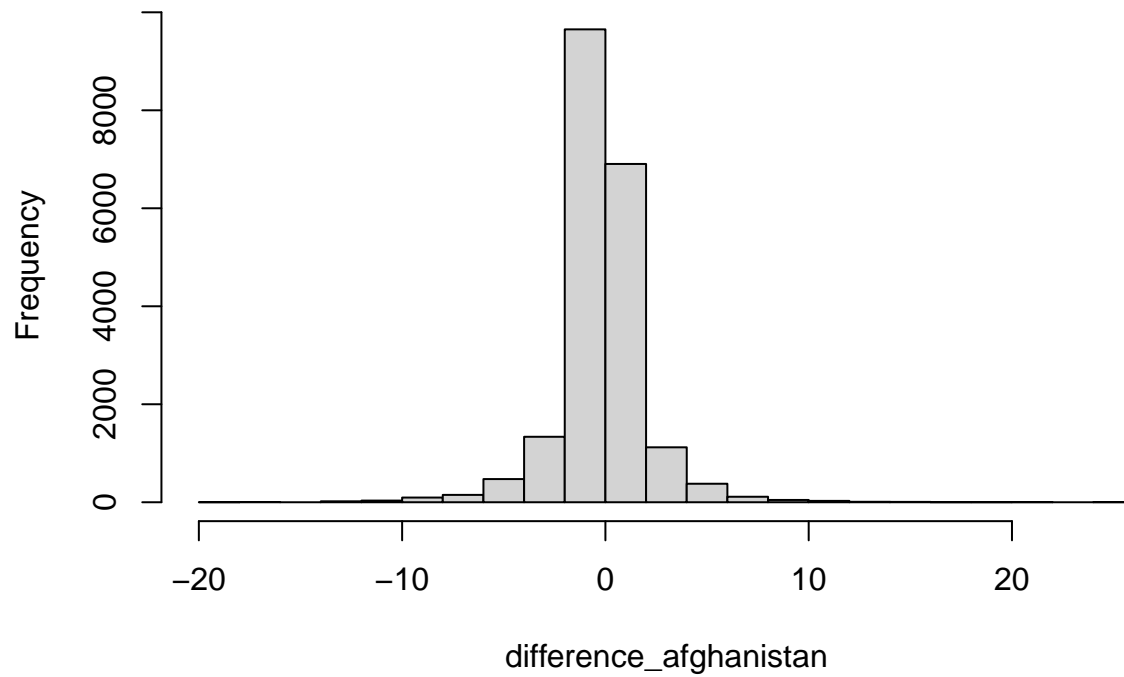
Open Values for Afghanistan
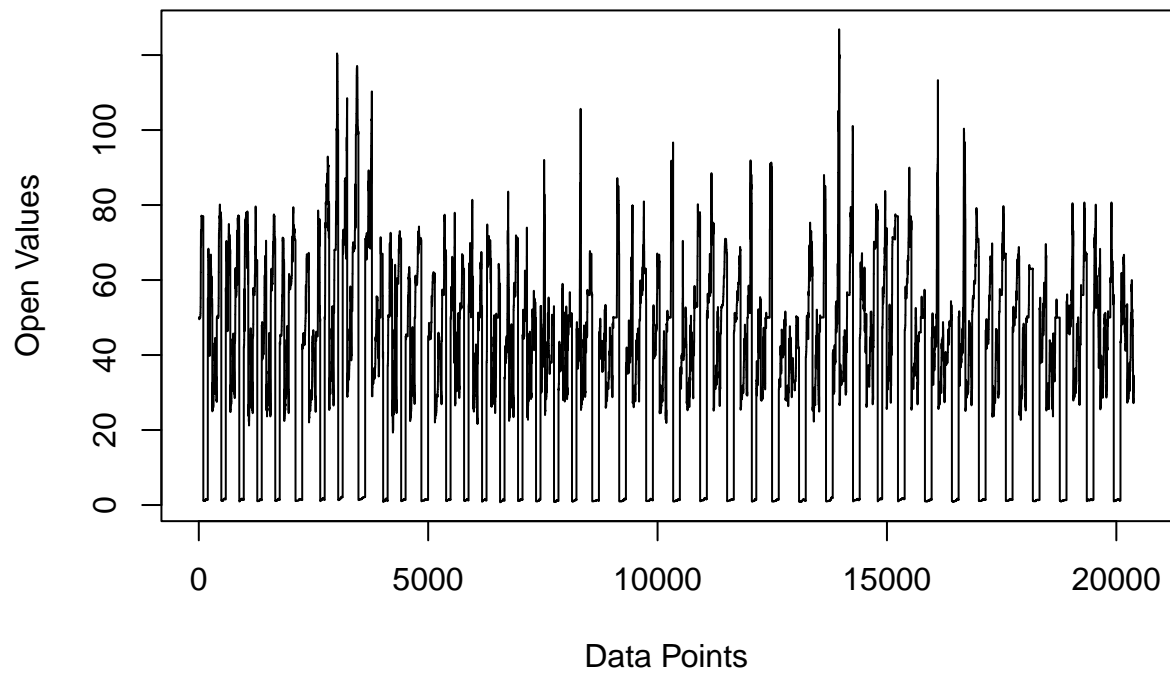

```
plot(x2, closed_values_afghanistan, type = "l", xlab = "Data Points", ylab = "Closed Values", main = "C
```
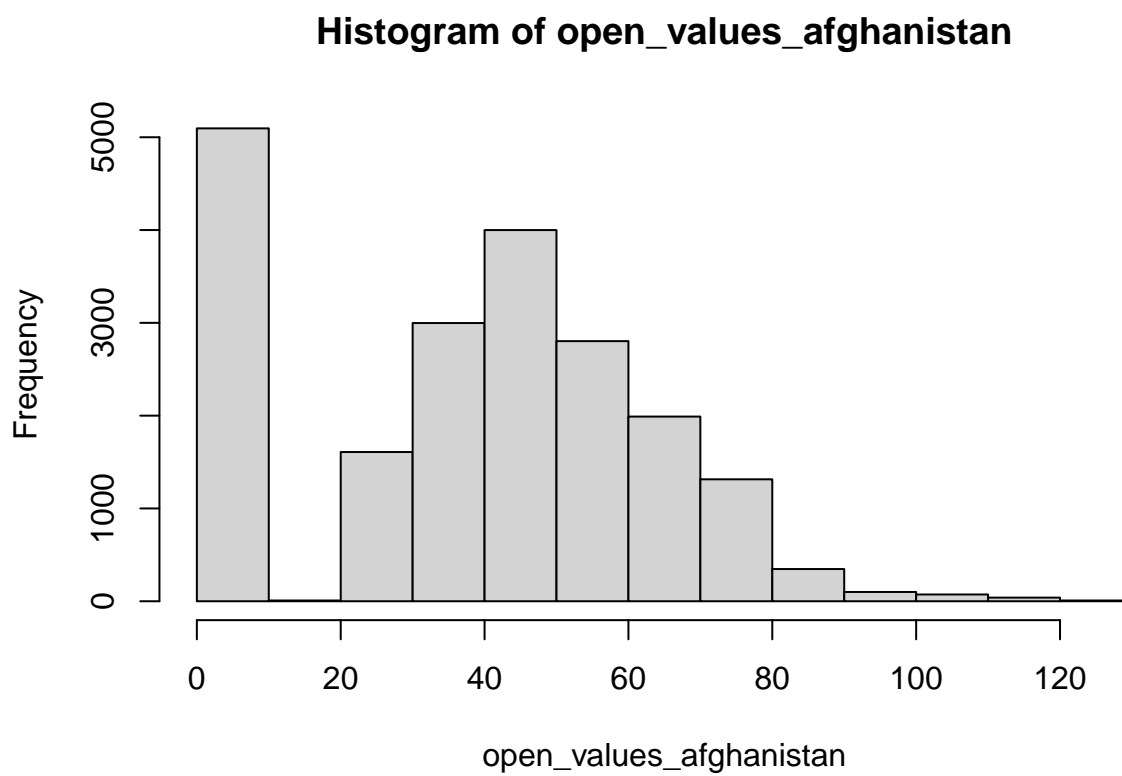
Closed Values for Afghanistan


```
#seems to be normally distributed, this helps narrow down appropriate models to examine Afghanistan spe
hist(difference_afghanistan)
```
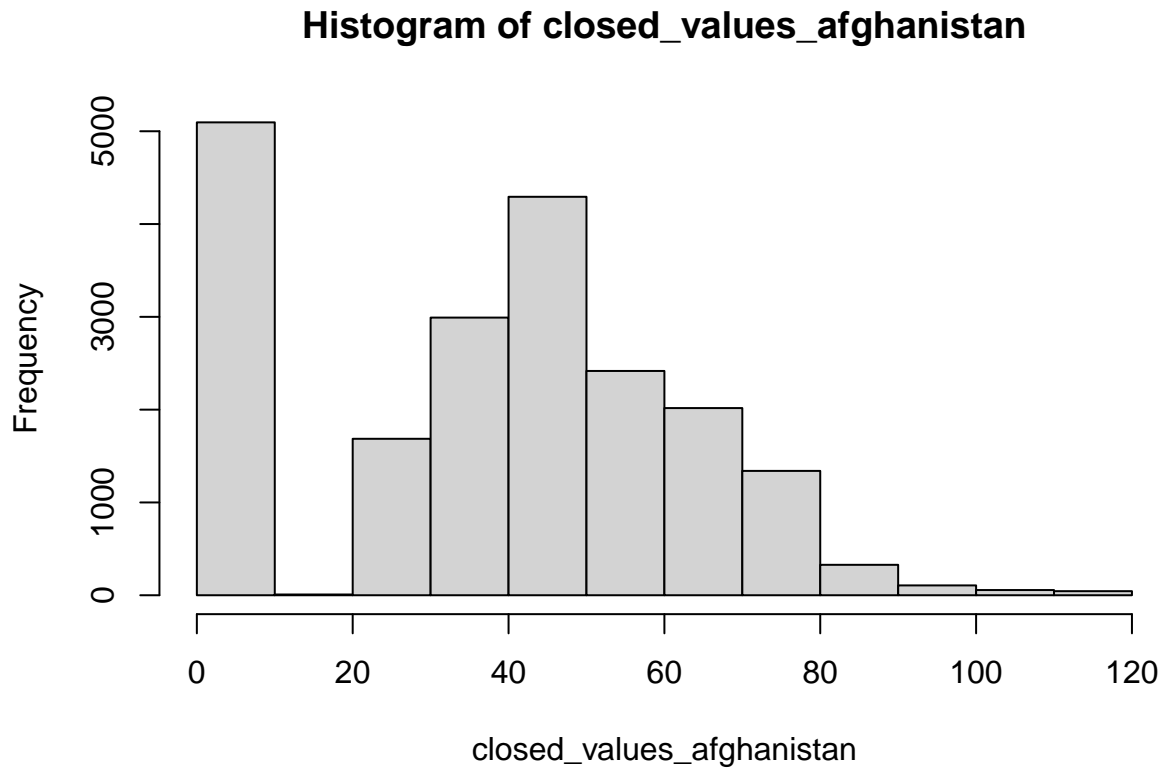
Histogram of difference_afghanistan


```
# lots of zero values in both that might effect the ability to show a normal distribution of the data.H
hist(open_values_afghanistan)
```

Histogram of open_values_afghanistan


```
hist(closed_values_afghanistan)
```

Histogram of closed_values_afghanistan

Below is the conflict data which is from https://acleddata.com/data-export-tool/. This requires a access key to get this data, it seems to be the most up to date data on conflicts. That coincides with the Dates from the Food inflation.

```
# Conflict data
```

Below is the the US Stock Market Data which is from https://www.ers.usda.gov/data-products/wheat-data/ and Yahoo Stocks for rought rice https://finance.yahoo.com/quote/ZR%3DF/history?period1=1575158400&period2=1703030400&interval=1d&filter=history&frequency=1d&includeAdjustedClose=true.

```
# US Stock Market Data
```