7COM1079-0901-2024 - Team Research and Development Project

Final Report Title: Seasonal Impact on Fuel Inventories

**Group ID:**

**Dataset number:**

**Prepared by:**

[Name and ID of submitting student first],

[Name and ID of other group members]

**Table of Contents**

[1. Introduction 3](#_Toc185852033)

[1.1 Problem statement and research motivation 3](#_Toc185852034)

[1.2 The data set 3](#_Toc185852035)

[1.3 Research question 3](#_Toc185852036)

[1.4 Null hypothesis and alternative hypothesis (H0/H1) 3](#_Toc185852037)

[2. Background research 4](#_Toc185852038)

[2.1 Research Papers 4](#_Toc185852039)

[2.2 Why RQ is of interest 4](#_Toc185852040)

[3. Visualization 4](#_Toc185852041)

[3.1 RQ output of an R script 4](#_Toc185852042)

[3.2 Understanding the data 7](#_Toc185852043)

[3.3 Useful information for the data understanding 7](#_Toc185852044)

[4. Analysis 8](#_Toc185852045)

[4.1 Statistical test 8](#_Toc185852046)

[4.2 Null hypothesis is rejected /not rejected based on the p-value 8](#_Toc185852047)

[5. Evaluation – group’s experience at 7COM1079 8](#_Toc185852048)

[5.1 What Went Well 8](#_Toc185852049)

[5.2 Points for improvement 8](#_Toc185852050)

[5.3 Group’s time management 9](#_Toc185852051)

[5.4 Project’s overall judgement 9](#_Toc185852052)

[6. Conclusions 9](#_Toc185852053)

[6.1 Results 9](#_Toc185852054)

[6.2 Interpretation of the result 9](#_Toc185852055)

[6.3 Reasons and/or implications for future work, limitations 9](#_Toc185852056)

[7. References 10](#_Toc185852057)

[8. Appendices 11](#_Toc185852058)

[A. R code used for analysis and visualization 12](#_Toc185852059)

[B. GitHub log output 14](#_Toc185852060)

**List of Figures**

[Figure 1: US Gasoline Stocks 6](#_Toc185852141)

[Figure 2: East Coast Ethanol Stocks 7](#_Toc185852142)

[Figure 3: East Coast Ethanol Stocks vs US Gasoline Stocks 8](#_Toc185852143)

# 1. Introduction

## 1.1 Problem statement and research motivation

The rationale for dealing with the impact of seasonality on fuel inventory comes from the importance of establishing the tendencies in fuel stocks that are caused by the seasons and their influences on supply chain and markets. As the focus on fuel stock during peak seasons like winter heating requirement or summer vacation travel increases, this research seeks to find out relationship link between important fuel stock variables which includes ethanol and gasoline. Such knowledge will help policymakers and industries to enhance the inventory control and manage risks resulting from flu season fluctuations (Inchauspe, Li and Park, 2020).

## 1.2 The data set

Ethanol, gasoline, diesel, and crude oil stocks directly related to fuel inventories in the year 2004 are included within this dataset with stocks measured in thousands of barrels. It includes areas such as East Coast and the Mid-Atlantic and contains information for the whole country. The data set is weekly and time series based and is crucial for examining seasonal trends as well as co-movement patterns of fuel stocks variables to enhance inventory density and avoid supply rationing mishaps.

## 1.3 Research question

Is there correlation between East. Coast. Ethanol. Stocks...Thousand.Barrels and U.S..Gasoline. Stocks. Thousand.Barrels?

## 1.4 Null hypothesis and alternative hypothesis (H0/H1)

According to the null hypothesis, there is no flow-dependency between various fuel stocks in U.S. inventories as influenced by seasonal variations implying a negative co-relation. On the other hand, the alternative hypothesis argues that the changing of the seasons are a critical determinant of the patterns that observe in the relationships between fuel stocks, and strongly inform the U.S. fuel inventory (Lasslop and Kloster, 2015).

**null-hypothesis:** There is no significant relationship between seasonal patterns and correlations among fuel stocks on U.S. fuel inventories.

**alt-hypothesis:** Seasonal patterns and correlations among fuel stocks significantly influence U.S. fuel inventories.

# 2. Background research

## 2.1 Research Papers

Based on the literature review about seasonal effects on fuel stocks, the importance of the analysis of demand and supply changes experienced due to seasonal effects is underlined. The more recent research focuses on notable models of demand forecasts, inventory control approaches as well as market characteristics which can be vital to the adequate stock control of fuels and defense against probable fluctuations resulting from seasonal and geographical differences (Sánchez, Martín and Zhang, 2021). Some research conducted recently has focused on the seasonal effect on fuels, a factor that makes the demand prediction and stocking a challenging endeavor. Namwad et al., (2024) are used for formulating an inventory model under fuzzy seasonal demand forecasting involving the use of machine learning technique for improving the stock control of perishable items. Examining fuel markets of the Visegrad Group countries, Krawiec and Górska (2024) only partially identified that these markets exhibit robust seasonality which leads to the conclusion that regional differences play a crucial role in the fuel markets. Więcek and Kubek (2024) evaluated the impact of time series characteristics for the effectiveness of fuel demand forecasting and indicated changes to autoregressive models and Markov chains for modifying forecasting performance in the fuel industry distribution environment. Such studies highlight the need of using seasonal factors as well as other complex algorithms to enhance fuel stocks and supply chain operations.

## 2.2 Why RQ is of interest

The research question relates to an important issue to fill a research gap on ethanol and gasoline stocks and how East coast ethanol supply is related to U.S. gasoline supply. Previous work has mainly examined the ways spatial and temporal dynamics impact fuel stocks but has not sufficiently investigate the interactions between different fuels. This study seeks to undertake an analysis of these dynamics to outlay explanatory knowledge on inventory management and energy policy. Subsequent research could focus on employing models for determining stocking points so as to improve the supply chain network.

# 3. Visualization

## 3.1 RQ output of an R script

# Plotting a Histogram for the U.S. Gasoline Stocks with Bell Curve

ggplot(FuelData, aes(x = `U.S..Gasoline.Stocks..Thousand.Barrels.`)) +

geom\_histogram(aes(y = ..density..), binwidth = 500, fill = "lightblue", color = "black") +

geom\_density(color = "red", size = 1) +

stat\_function(fun = dnorm, args = list(mean = mean(FuelData$`U.S..Gasoline.Stocks..Thousand.Barrels.`),

sd = sd(FuelData$`U.S..Gasoline.Stocks..Thousand.Barrels.`)), color = "black", size = 1) +

labs(title = "Histogram of U.S. Gasoline Stocks",

x = "U.S. Gasoline Stocks (Thousand Barrels)",

y = "Density") +

theme\_minimal()

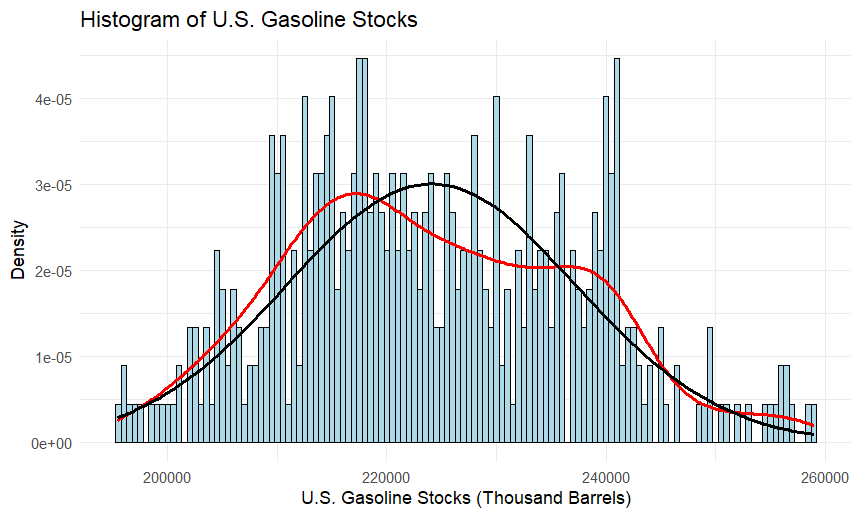


Figure 1: US Gasoline Stocks

# Plotting a Histogram for the East Coast Ethanol Stocks with Bell Curve

ggplot(FuelData, aes(x = `East.Coast.Ethanol.Stocks...Thousand.Barrels.`)) +

geom\_histogram(aes(y = ..density..), binwidth = 500, fill = "lightgreen", color = "black") +

geom\_density(color = "red", size = 1) +

stat\_function(fun = dnorm, args = list(mean = mean(FuelData$`East.Coast.Ethanol.Stocks...Thousand.Barrels.`),

sd = sd(FuelData$`East.Coast.Ethanol.Stocks...Thousand.Barrels.`)), color = "black", size = 1) +

labs(title = "Histogram of East Coast Ethanol Stocks",

x = "East Coast Ethanol Stocks (Thousand Barrels)",

y = "Density") +

theme\_minimal()

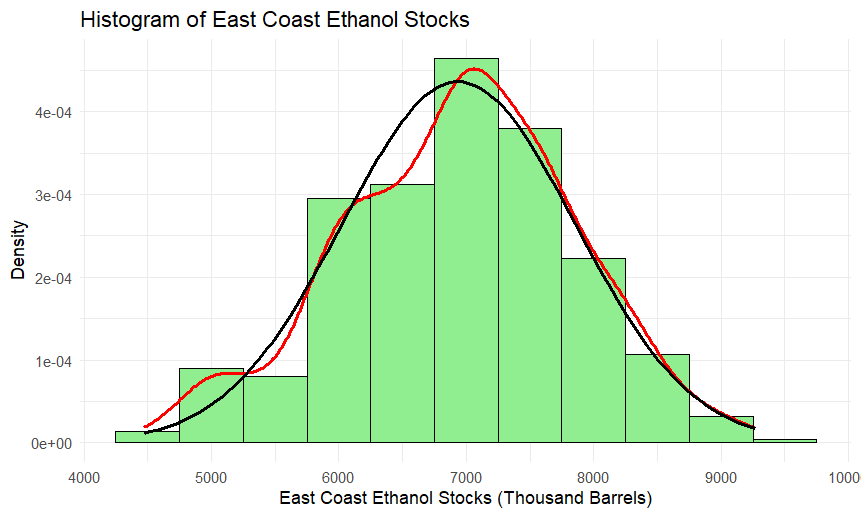


Figure 2: East Coast Ethanol Stocks

# Plotting a Scatter Plot of East Coast Ethanol Stocks vs U.S. Gasoline Stocks with Regression Line

ggplot(FuelData, aes(x = `East.Coast.Ethanol.Stocks...Thousand.Barrels.`,

y = `U.S..Gasoline.Stocks..Thousand.Barrels.`)) +

geom\_point(color = "blue") +

geom\_smooth(method = "lm", color = "red", se = FALSE) +

labs(title = "Scatter Plot of East Coast Ethanol Stocks vs U.S. Gasoline Stocks",

x = "East Coast Ethanol Stocks (Thousand Barrels)",

y = "U.S. Gasoline Stocks (Thousand Barrels)") +

theme\_minimal()

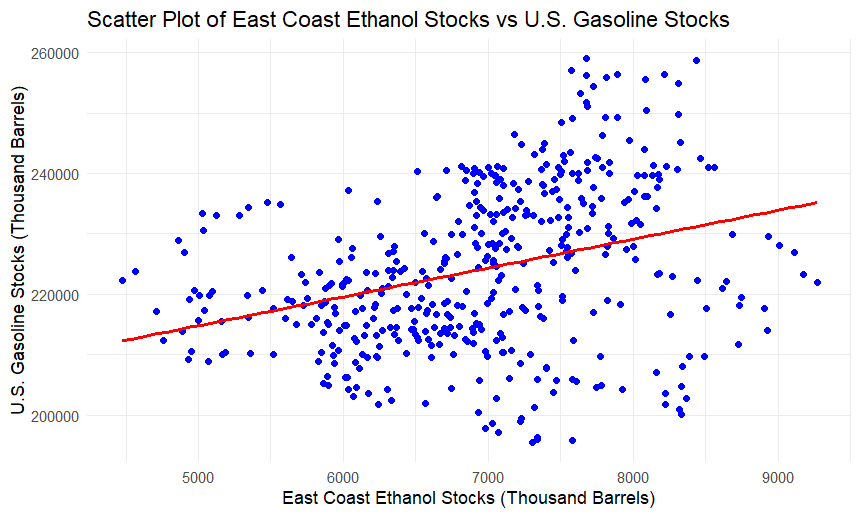


Figure 3: East Coast Ethanol Stocks vs US Gasoline Stocks

## 3.2 Understanding the data

The histograms given below depict U.S. Gasoline Stocks and East Coast Ethanol Stock, normality & density. Their pattern of relationship is further analyzed in the scatter plot which shows how the two increases in tandem. These visuals help achieve the research aim of exploring seasonal trend and relationship between fuel stocks for effective resource utilization.

## 3.3 Useful information for the data understanding

The histogram results suggest that both U.S. Gasoline and East Coast Ethanol stock data exhibit near-normal distributions with significant differences in density. Figure 3 is a scatter plot showing that there is a positive relationship of ethanol stocks with the gasoline stock levels. These viewpoints demonstrate useful correlations and reveal possible effects of seasonal fluctuation on fuel inventory.

# 4. Analysis

## 4.1 Statistical test

For this analysis, one-sample t-test is used because it determines if the observed mean of East Coast Ethanol Stocks or U.S. Gasoline Stocks is significantly different from hypothetical values of 5000 and 200000 respectively. The test is applicable to continuous data, and requires normal distribution, while the research is focused on mean deviations.

## 4.2 Null hypothesis is rejected /not rejected based on the p-value

The null hypothesis can be rejected according to p-value for both tests. In case of East Coast Ethanol Stocks, the p-value < 0.05 hence it can be concluded that the mean significantly deviates from the hypothesized value of 5000. In the same way, for the variable U.S. Gasoline Stocks is another variable which shows that p-value is less than 0.05 thus indicating that the mean is significantly different from 200,000. These results suggest that the observed data does not give evidence against the null hypothesis to reject the null hypothesis. Consequently, statistical data substantiate the hypothesis that the actual means are significantly different from the hypothesized values.

# 5. Evaluation – group’s experience at 7COM1079

## 5.1 What Went Well

Statistical testing and analysis at the end of the project were exhaustive and uncovered the variation in fuel inventories by seasons. It was also possible to coordinate to clean data and check hypotheses on data visualizations with high accuracy. The team used the right applications for analysis and provided significant information including the relationship between ethanol and gasoline stock.

## 5.2 Points for improvement

The major difficulties are related to the issues of data handling as well as the coordination of the group’s work. There was agreement about what could be done better: increasing their proficiency in using more sophisticated statistical methods and how the effectiveness of data displays could be augmented. Increased role clarity, and more effective individual characteristic task matching might result in increased efficiency in future work.

## 5.3 Group’s time management

Time was efficiently and coherently divided and followed in terms of data preprocessing, statistical testing and documentation. To that end, the meetings and timeline helped the process remain on time and track while stay on course on its significant goals and targets.

## 5.4 Project’s overall judgement

The project outcome was positive as all laid down goal and objectives were achieved and key conclusions on fuel inventory relations were derived. When it comes to the accomplishment of given objective, meaningful and effective utilization of tools and dedicated efforts, collaboration with the colleagues were some of the positive aspects that were possible to identify which will further enhance the faculty’s work in this particular domain.

# 6. Conclusions

## 6.1 Results

The analysis shows that East Coast Ethanol Stocks are positively related to U.S. Gasoline Stocks on the 0.6475 level. The one-sample t-tests showed that the null hypothesized means of both the fuel stock variables are statistically significant. These results of this study highlight how such seasonal fluctuations affect fuel inventory management and should be useful to practitioners regarding how resources can be better utilized and inventory controlled.

## 6.2 Interpretation of the result

The findings point out that ethanol stock fluctuations affect the gasoline stocks in support of the research question. This correlation indicates the importance of linking both fuel stock management to try and avoid such disruptions during the various seasons. In so doing, these insights will bring light to policymakers and stakeholders to extend the implementations of reliable policies to ensure the availability of fuel and reduce vulnerability to other forms of insecurity in the energy sector.

## 6.3 Reasons and/or implications for future work, limitations

The only limitation in this analysis is the oversimplification of results from a single data set and secondly, the low sensitivity towards outside variable like economic and geopolitical factors. New studies should include more measures, larger and longer databases, and prediction equations to further fine-tune understandings about seasonal effects on fuel stocks and other energy system changes.

# 7. References

‌Inchauspe, J., Li, J. and Park, J. (2020). Seasonal patterns of global oil consumption: Implications for long term energy policy. *Journal of Policy Modeling*, [online] 42(3), pp.536–556. doi: https://doi.org/10.1016/j.jpolmod.2019.12.005.

‌Krawiec, M. and Górska, A. (2024). Analysis of Seasonal Patterns in the Performance of Fuel Markets in the Visegrad Group. *Comparative Economic Research. Central and Eastern Europe*, [online] 27(2), pp.49–72. doi: https://doi.org/10.18778/1508-2008.27.12.

‌Lasslop, G. and Kloster, S. (2015). Impact of fuel variability on wildfire emission estimates. *Atmospheric Environment*, [online] 121, pp.93–102. doi: https://doi.org/10.1016/j.atmosenv.2015.05.040.

Namwad, R.S., Mishra, N.K., None Ranu and Jain, P. (2024). Optimizing Inventory Management with Seasonal Demand Forecasting in a Fuzzy Environment. *Journal Européen des Systèmes Automatisés*, [online] 57(4), pp.1091–1102. doi: https://doi.org/10.18280/jesa.570416.

‌Sánchez, A., Martín, M. and Zhang, Q. (2021). Optimal design of sustainable power-to-fuels supply chains for seasonal energy storage. *Energy*, [online] 234, pp.121300–121300. doi: https://doi.org/10.1016/j.energy.2021.121300.

Więcek, P. and Kubek, D. (2024). The Impact Time Series Selected Characteristics on the Fuel Demand Forecasting Effectiveness Based on Autoregressive Models and Markov Chains. *Energies*, [online] 17(16), pp.4163–4163. doi: https://doi.org/10.3390/en17164163.

‌

# 8. Appendices

## A. R code used for analysis and visualization

# Loading required libraries

library(tidyverse) # For data manipulation and visualization

library(lubridate) # For handling dates

library(ggplot2) # For visualizations

library(corrplot) # For correlation matrix visualization

# Loading the dataset

data <- read\_csv("transportation-fuels-inventory-beginning-2004-1.csv")

# Cleaning column names to make them syntactically valid

data <- data %>%

rename\_with(~make.names(.), everything())

# Converting the Date column to Date type

data <- data %>%

mutate(Date = mdy(Date)) %>% # Converting Date to proper date format

arrange(Date) # Arranging rows by Date

# Checking missing values in the data

sum(is.na(data))

# removing missing values from the data

data <- data %>% drop\_na()

# Printing the first few rows of the cleaned data

head(data)

# Checking the structure of the data

str(data)

# Printing the summary statistics of the data

summary(data)

# Printing the column names of the data

colnames(data)

# Selecting fuel stocks columns

FuelData <- data %>% select(Date, `East.Coast.Ethanol.Stocks...Thousand.Barrels.`, `U.S..Gasoline.Stocks..Thousand.Barrels.`)

# Plotting a Histogram for the U.S. Gasoline Stocks with Bell Curve

ggplot(FuelData, aes(x = `U.S..Gasoline.Stocks..Thousand.Barrels.`)) +

geom\_histogram(aes(y = ..density..), binwidth = 500, fill = "lightblue", color = "black") +

geom\_density(color = "red", size = 1) +

stat\_function(fun = dnorm, args = list(mean = mean(FuelData$`U.S..Gasoline.Stocks..Thousand.Barrels.`),

sd = sd(FuelData$`U.S..Gasoline.Stocks..Thousand.Barrels.`)), color = "black", size = 1) +

labs(title = "Histogram of U.S. Gasoline Stocks",

x = "U.S. Gasoline Stocks (Thousand Barrels)",

y = "Density") +

theme\_minimal()

# Plotting a Histogram for the East Coast Ethanol Stocks with Bell Curve

ggplot(FuelData, aes(x = `East.Coast.Ethanol.Stocks...Thousand.Barrels.`)) +

geom\_histogram(aes(y = ..density..), binwidth = 500, fill = "lightgreen", color = "black") +

geom\_density(color = "red", size = 1) +

stat\_function(fun = dnorm, args = list(mean = mean(FuelData$`East.Coast.Ethanol.Stocks...Thousand.Barrels.`),

sd = sd(FuelData$`East.Coast.Ethanol.Stocks...Thousand.Barrels.`)), color = "black", size = 1) +

labs(title = "Histogram of East Coast Ethanol Stocks",

x = "East Coast Ethanol Stocks (Thousand Barrels)",

y = "Density") +

theme\_minimal()

# Plotting a Scatter Plot of East Coast Ethanol Stocks vs U.S. Gasoline Stocks with Regression Line

ggplot(FuelData, aes(x = `East.Coast.Ethanol.Stocks...Thousand.Barrels.`,

y = `U.S..Gasoline.Stocks..Thousand.Barrels.`)) +

geom\_point(color = "blue") +

geom\_smooth(method = "lm", color = "red", se = FALSE) +

labs(title = "Scatter Plot of East Coast Ethanol Stocks vs U.S. Gasoline Stocks",

x = "East Coast Ethanol Stocks (Thousand Barrels)",

y = "U.S. Gasoline Stocks (Thousand Barrels)") +

theme\_minimal()

# Loading required libraries

library(tidyverse) # For data manipulation and visualization

library(lubridate) # For handling dates

library(ggplot2) # For visualizations

library(corrplot) # For correlation matrix visualization

# Loading the dataset

data <- read\_csv("transportation-fuels-inventory-beginning-2004-1.csv")

# Cleaning column names to make them syntactically valid

data <- data %>%

rename\_with(~make.names(.), everything())

# Converting the Date column to Date type

data <- data %>%

mutate(Date = mdy(Date)) %>% # Converting Date to proper date format

arrange(Date) # Arranging rows by Date

# Checking missing values in the data

sum(is.na(data))

# Removing missing values from the data

data <- data %>% drop\_na()

# Printing the first few rows of the cleaned data

head(data)

# Checking the structure of the data

str(data)

# Printing the summary statistics of the data

summary(data)

# Printing the column names of the data

colnames(data)

# Selecting fuel stocks columns

FuelData <- data %>% select(Date, `East.Coast.Ethanol.Stocks...Thousand.Barrels.`, `U.S..Gasoline.Stocks..Thousand.Barrels.`)

# Plotting a Histogram for the U.S. Gasoline Stocks with Bell Curve

ggplot(FuelData, aes(x = `U.S..Gasoline.Stocks..Thousand.Barrels.`)) +

geom\_histogram(aes(y = ..density..), binwidth = 500, fill = "lightblue", color = "black") +

geom\_density(color = "red", size = 1) +

stat\_function(fun = dnorm, args = list(mean = mean(FuelData$`U.S..Gasoline.Stocks..Thousand.Barrels.`),

sd = sd(FuelData$`U.S..Gasoline.Stocks..Thousand.Barrels.`)), color = "black", size = 1) +

labs(title = "Histogram of U.S. Gasoline Stocks",

x = "U.S. Gasoline Stocks (Thousand Barrels)",

y = "Density") +

theme\_minimal()

# Plotting a Histogram for the East Coast Ethanol Stocks with Bell Curve

ggplot(FuelData, aes(x = `East.Coast.Ethanol.Stocks...Thousand.Barrels.`)) +

geom\_histogram(aes(y = ..density..), binwidth = 500, fill = "lightgreen", color = "black") +

geom\_density(color = "red", size = 1) +

stat\_function(fun = dnorm, args = list(mean = mean(FuelData$`East.Coast.Ethanol.Stocks...Thousand.Barrels.`),

sd = sd(FuelData$`East.Coast.Ethanol.Stocks...Thousand.Barrels.`)), color = "black", size = 1) +

labs(title = "Histogram of East Coast Ethanol Stocks",

x = "East Coast Ethanol Stocks (Thousand Barrels)",

y = "Density") +

theme\_minimal()

# Plotting a Scatter Plot of East Coast Ethanol Stocks vs U.S. Gasoline Stocks with Regression Line

ggplot(FuelData, aes(x = `East.Coast.Ethanol.Stocks...Thousand.Barrels.`,

y = `U.S..Gasoline.Stocks..Thousand.Barrels.`)) +

geom\_point(color = "blue") +

geom\_smooth(method = "lm", color = "red", se = FALSE) +

labs(title = "Scatter Plot of East Coast Ethanol Stocks vs U.S. Gasoline Stocks",

x = "East Coast Ethanol Stocks (Thousand Barrels)",

y = "U.S. Gasoline Stocks (Thousand Barrels)") +

theme\_minimal()

# Testing if the mean of East Coast Ethanol Stocks is significantly different from a hypothetical value

TTest <- t.test(FuelData$`East.Coast.Ethanol.Stocks...Thousand.Barrels.`, mu = 5000)

TTest

# Testing if the mean of U.S. Gasoline Stocks is significantly different from a hypothetical value

TTestGasoline <- t.test(FuelData$`U.S..Gasoline.Stocks..Thousand.Barrels.`, mu = 200000)

TTestGasoline

## B. GitHub log output