main.R

Gayatri Shahane

4/22/2020

#Loading Libraries  
library(tidyverse)  
library(readr)  
library(lubridate)  
  
#TASK1  
files <- list.files(path = "/Users/gayatrishahane/Desktop/GS\_CQuant/GS\_CQuant\_Solution/historicalPriceData", pattern = "\*.csv", full.names = TRUE)  
PriceData <- sapply(files, read\_csv, simplify=FALSE) %>%   
bind\_rows()  
PriceData <- PriceData %>% arrange(Date)

#TASK2  
view(PriceData)  
AvgMonth <- PriceData %>%   
 separate(Date, c("Date", "Time"), " ") %>%   
 separate(Date,c("Year", "Month", "Day"), "-") %>%  
 group\_by(SettlementPoint, Year, Month) %>%   
 summarise(AveragePrice = mean(Price))  
  
#TASK3  
write\_csv(AvgMonth,'AveragePriceByMonth.csv', col\_names = TRUE)

#TASK4  
HourlyVolatalityByYear <- PriceData %>%   
 filter(Price>0) %>%   
 filter(str\_detect(SettlementPoint, "^HB")) %>%   
 group\_by(SettlementPoint, year(Date)) %>%   
 summarise(HourlyVolatility = sd(log(Price))) %>%   
 rename(Year = `year(Date)`)  
  
#TASK5  
write\_csv(HourlyVolatalityByYear,'HourlyVolatilityByYear.csv', col\_names = TRUE)  
  
#TASK6  
MaxVolatality <- HourlyVolatalityByYear %>%   
 group\_by(Year) %>%   
 filter(HourlyVolatility == max(HourlyVolatility)) %>%  
 arrange(Year) %>%   
 write\_csv('MaxVolatilityByYear.csv')

#TASK7  
FormattedPrice <- PriceData %>%  
 mutate(Time = as.integer(hour(Date))+1, Date = date(Date)) %>%   
 spread(Time, Price) %>%   
 rename(Variable = SettlementPoint) %>%   
 select(c(2,1,3:26))  
  
#Naming the Columns  
 newcolnames <- colnames(FormattedPrice)  
 for (i in 3:length(newcolnames)){  
 newcolnames[i] = paste('X', newcolnames[i])  
 }  
 colnames(FormattedPrice) <- newcolnames  
   
#Writing the csv files   
 for(var1 in unique(FormattedPrice$Variable)){  
 uni <- FormattedPrice %>%   
 filter(Variable == var1)  
 uni %>%   
 write\_csv(paste0('spot\_',var1, '.csv'))  
 }  
#Creating new sub directory was causing path error. Hence, all the csv files are in the same main file.

#Bonus Mean Plots  
  
MeanPlot <- AvgMonth %>%   
 mutate(Day = '01') %>%   
 unite(Date, Year, Month, Day, sep = "-") %>%   
 mutate(Date = date(Date))  
  
#Average Prices for Settlement Hubs  
  
HubPlot <- MeanPlot %>%   
 filter(str\_detect(SettlementPoint, "^HB")) %>%   
 ggplot(mapping = aes(x= Date, y = AveragePrice, color = SettlementPoint)) +  
 geom\_line() +  
 theme(axis.text.x = element\_text(angle = 90)) +  
 scale\_x\_date(date\_breaks = "months" , date\_labels = "%Y-%m-%d") +  
 ggsave('SettlementHubAveragePriceByMonth.png')

## Saving 5 x 4 in image

#Average Prices for Load Zones  
  
LoadZonePlot <- MeanPlot %>%   
 filter(str\_detect(SettlementPoint, "^LZ")) %>%   
 ggplot(mapping = aes(x= Date, y = AveragePrice, color = SettlementPoint)) +  
 geom\_line() +  
 theme(axis.text.x = element\_text(angle = 90)) +  
 scale\_x\_date(date\_breaks = "months" , date\_labels = "%Y-%m-%d") +  
 ggsave('LoadZoneAveragePriceByMonth.png')

## Saving 5 x 4 in image

#Bonus Volatality Plot  
  
VolPLot1 <- ggplot(data = HourlyVolatalityByYear) +   
 geom\_point(mapping = aes(x = Year, y = HourlyVolatility, color = SettlementPoint))+  
 geom\_line(mapping = aes(x = Year, y = HourlyVolatility, color = SettlementPoint))+  
 ggtitle('Settlement Hub Hourly Volatility by Year')+  
 ggsave('SettlementHubHourlyVolatility.png')

## Saving 5 x 4 in image

#Bonus Hourly Shape Profile Computation  
  
HourlyShape <- PriceData %>%   
 mutate(Time = as.integer(hour(Date))+1, Date = date(Date)) %>%   
 mutate(Day = day(Date), Month = month(Date)) %>%   
 group\_by(SettlementPoint, Month, Day, Time) %>%  
 summarise(AvgPrice = mean(Price)) %>%   
 group\_by(SettlementPoint, Month, Day) %>%  
 mutate(NormalAvgPrice = AvgPrice/mean(AvgPrice)) %>%   
 select(SettlementPoint, Month, Day, Time, NormalAvgPrice)  
  
#Writing the csv files  
  
for(var2 in unique(HourlyShape$SettlementPoint)){  
 uni <- HourlyShape %>%   
 filter(SettlementPoint == var2)  
 uni %>%   
 write\_csv(paste0('profile\_',var2, '.csv'))  
}

#Bonus Analysis  
  
#Plot1  
ggplot(data = PriceData) +   
 geom\_point(mapping = aes(x = Date, y = Price, color = SettlementPoint)) +   
 facet\_wrap(~ SettlementPoint) +  
 ggsave('BonusAnalysis1.png')

## Saving 5 x 4 in image

