Group Assignment 1

Simon Yu

2025-10-01

Table of contents

List of Figures

List of Tables

```
library(dplyr)
```

```
data <- read.csv("compustat_food_bev.csv")

# filter data for Chipotle (SMG)
cmg_data <- filter(data, tic == "CMG")

# convert datadate to Date type
cmg_data$datadate <- as.Date(cmg_data$datadate, format = "%d/%m/%Y")
```

Question 1

```
#1. Add a new column named daily_return
#   to store daily return value

#2. Drop the rows whose daily return values are NA
cmg_data <- mutate(cmg_data, daily_return = (prccd - lag(prccd)) / lag(prccd))
cmg_data <- filter(cmg_data, !is.na(daily_return))
```

Question 2

```
#1. Add a new column named volume_change
#   to store volume change value

#2. Drop the rows whose volume_change values are NA
cmg_data <- mutate(cmg_data, volume_change = (cshtrd - lag(cshtrd)))
cmg_data <- filter(cmg_data, !is.na(volume_change))
```

Question 3

```
#1. Add a new column named momentum_10
#   to store 10-day momentum value

#2. Drop the rows whose momentum_10 values are NA
cmg_data <- mutate(cmg_data, momentum_10 = prccd - lag(prccd, 10))
cmg_data <- filter(cmg_data, !is.na(momentum_10))
```

Question 4

```r
# Add a new column named MFV
cmg_data <- mutate(cmg_data,
                mfv = ((prccd - prcld) - (prchd - prccd)) * cshtrd / (prchd - prcld))
```

```r
# show the table containing only the new 4 metrics columns
cmg_data_metrics <- filter(cmg_data[ ,
                 c("datadate", "daily_return", "volume_change", "momentum_10", "mfv")])

head(cmg_data_metrics)
```

A data.frame: 6 × 5

	datadate <date>	daily_return <dbl>	volume_change <int>	momentum_10 <dbl>	mfv <dbl>
1	2020-09-18	-0.008071032	331185	-113.25	-292595.13
2	2020-09-21	-0.007371573	-474653	-107.21	280004.81
3	2020-09-22	0.023256970	-6320	-65.17	123610.04
4	2020-09-23	0.011185990	227749	-72.64	-363145.20
5	2020-09-24	-0.015307716	-290140	-79.21	-22833.32
6	2020-09-25	0.012381231	-5553	-52.97	117739.00

Question 5

```r
library(lubridate)
```

```r
# Add a new column named month
cmg_data <- mutate(cmg_data, month = month(datadate))
```

Question 6

```r
# Add a new column named year
cmg_data <- mutate(cmg_data, year = year(datadate))
```

```
# show the table containing only the new 2 date columns
cmg_data_dates <- filter(cmg_data[ , c("datadate", "month", "year")])
head(cmg_data_dates, 10)
```

A data.frame: 10 × 3

	datadate <date>	month <dbl>	year <dbl>
1	2020-09-18	9	2020
2	2020-09-21	9	2020
3	2020-09-22	9	2020
4	2020-09-23	9	2020
5	2020-09-24	9	2020
6	2020-09-25	9	2020
7	2020-09-28	9	2020
8	2020-09-29	9	2020
9	2020-09-30	9	2020
10	2020-10-01	10	2020

Question 7

```
# Calculate the total trade volume (cshtrd) for June 2023
trade_volume_2023_06 <- filter(cmg_data, year == 2023 & month == 6)

sum(trade_volume_2023_06$cshtrd)
```

5392605

Question 8

```
# Calculate the mean of daily return over the period
mean_daily_return <- mean(cmg_data$daily_return)
print(mean_daily_return)
```

```
[1] 0.0008331754
```

Question 9

```
# Find the date for maximum high price over the period
max_high_price <- max(cmg_data$prchd)

date_max_high_price <- filter(cmg_data, prchd == max_high_price)

print(date_max_high_price$datadate)
```

```
[1] "2023-07-19"
```

Question 10

```
# Find the date for largest daily return over the period
max_daily_return <- max(cmg_data$daily_return)
date_max_daily_return <- filter(cmg_data, daily_return == max_daily_return)
print(date_max_daily_return$datadate)
```

```
[1] "2022-07-27"
```