

Group Assignment 1

Simon Yu

2025-10-01

Table of contents

Question 1	3
Question 2	3
Question 3	3
Question 4	3
Question 5	4
Question 6	4
Question 7	5
Question 8	5
Question 9	5
Question 10	6

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
```

```
cmg_data <- mutate(cmg_data, daily_return = (prccd - lag(prccd)) / lag(prccd))
```

```
cmg_data <- filter(cmg_data, !is.na(daily_return))
```

Question 2

```
#2. Add a new column named 10-day momentum indicator
```

```
cmg_data <- mutate(cmg_data, momentum_10 = prccd - lag(prccd, 10))
```

```
cmg_data <- filter(cmg_data, !is.na(momentum_10))
```

Question 3

```
#3. Add a new column named daily_range
```

```
cmg_data <- mutate(cmg_data, daily_range = prchd - prcld)
```

Question 4

```
#4. Add a new column named MFV
```

```
cmg_data <- mutate(cmg_data,
```

```
mfv = ((prccd - prcld) - (prchd - prccd)) * cshtd / (prchd - prcld))
```

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

head(cmg_data_metrics, 10)
```

A data.frame: 10 × 5

	datadate <date>	daily_return <dbl>	momen- tum_10 <dbl>	daily_range <dbl>	mfv <dbl>
1	2020-09-17	-0.031993807	-153.66	42.3600	-103314.12
2	2020-09-18	-0.008071032	-113.25	46.4200	-292595.13
3	2020-09-21	-0.007371573	-107.21	27.8099	280004.81
4	2020-09-22	0.023256970	-65.17	27.2700	123610.04
5	2020-09-23	0.011185990	-72.64	50.4911	-363145.20
6	2020-09-24	-0.015307716	-79.21	30.4063	-22833.32
7	2020-09-25	0.012381231	-52.97	36.0900	117739.00
8	2020-09-28	0.015146645	-3.84	31.4400	51534.89
9	2020-09-29	-0.001377291	-15.75	21.5700	94370.66
10	2020-09-30	-0.014188220	-22.16	37.6300	-144418.84

Question 5

```
library(lubridate)
```

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

Question 6

```
#6. 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

	date <date>	month <dbl>	year <dbl>
1	2020-09-17	9	2020
2	2020-09-18	9	2020
3	2020-09-21	9	2020
4	2020-09-22	9	2020
5	2020-09-23	9	2020
6	2020-09-24	9	2020
7	2020-09-25	9	2020
8	2020-09-28	9	2020
9	2020-09-29	9	2020
10	2020-09-30	9	2020

Question 7

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

sum(trade_volume_2023_06$cshtd)
```

5392605

Question 8

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

[1] 0.000789406

Question 9

```
#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

```
#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"
```