Report

Wenjin Zhang

2021/3/05

Overview

Introduction

The COVID-19 pandemic has had a devastating effect on the economy in the United States. In early March 2020, the first lockdowns began and the stock market plunged. After this initial reaction, however, the market recovered.

I use some stock information downloaded by myself at the beginning of winter holidays through package **tidyquant**. And I assume an intial position of $250,000 in cash on 2020-07-01 and invest 6 stocks using investment advice before 1st July from ETF. I track the value of my portfolio using last close prices through the fall to 2020-11-30. At the end of 2020-12-01, I have a position of $495,201 in cash.

In shiny I not only display the result of our portfolio, but also show close prices of the indexes and some other stocks. This app ensure people to have a general impression on the change in stock market and make their own portfolio.

The goal of this project is to make us go from raw data to an interactive web app in R via the shiny package.

Click here go to shiny app

Data

I choose several datasets:
- The list of stocks containing symbol, company name, last sale price, market caption, ipo, sector and industry from 2020-07-01 to 2020-11-30.
- All stock prices from 2020-07-01 to 2020-11-30.
- Values of three indexes(S&P 500, Dow Jones Industrial Average and NASDAQ) from 2020-07-01 to 2021-03-10.

To be honest, I am really interested in the stock market and want to experience investment. So I downloaded the stock list and stock prices (2020-07-01 – 2020-11-30) through functions in tidyquant package at the beginning of winter holidays (Downloading these data cost me at least 5 hours). I choose stock data from 2020-07-01 is to find out: 1) Since I use online advice to invest, I can check whether their suggestions are reliable. 2) I try to keep an eye on the change in financial market after the outbreak of COVID-19. Unfortunatly, I find there is something wrong with functions in tidyquant when I try to download the latest data.
Hence, I will show how I downloaded the data before. And I will clean and display these datasets.

Preparation

- I use tq_exchange() to collect the name of all stocks from three stock exchanges (AMER, NASDAQ, NYSE). This dataset has 7535 rows (7535 stocks) and 7 columns.

symbol	company	last.sale.price	market.cap	ipo.year	sector	industry
GOED	1847 Goedeker Inc.	6.49	$39.66M	2020	Consumer Services	Home Furnishings
XXII	22nd Century Group, Inc	2.15	$298.55M	NA	Consumer Non-Durables	Farming/Seeds/Milling
FAX	Aberdeen Asia-Pacific Income Fund Inc	4.34	$1.07B	1986	n/a	n/a

- Then I use tq_get() to download price of all stocks form 2020-07-01 to 2020-11-30. When I continue to download data on 2020-12-01, the connection of tq_get() failes. This dataset contains 654984 rows and 8 columns.

symbol	date	open	high	low	close	volume	adjusted
BBH	2020-07-01	163.24	165.17	161.39	164.67	29000	164.67
BBH	2020-07-02	166.03	167.15	165.13	165.81	32600	165.81
BBH	2020-07-06	168.45	168.54	166.49	166.97	16700	166.97

- So, we search closing price on 2020-12-01 of the six stocks we are interested in one by one on **YAHOO** website.

date	ZM	CTVA	PINS	UBER	MRNA	PENN
2020-12-01	406.31	38.31	68.21	49.63	141.01	70.03

- I also download values of S&P 500, Dow Jones Industrial Average and NASDAQ to take an overall view of the stock market.

S&P 500

Date	Open	High	Low	Close	Adj.Close	Volume
2021/3/10	3891.99	3917.35	3885.73	3898.81	3898.81	5827250000
2021/3/9	3851.93	3903.76	3851.93	3875.44	3875.44	5496340000
2021/3/8	3844.39	3881.06	3819.25	3821.35	3821.35	5852240000

Dow Jones Industrial Average

Date	Open	High	Low	Close	Adj.Close	Volume
2021/3/10	31906.96	32389.50	31906.96	32297.02	32297.02	4091100
2021/3/9	31892.35	32150.32	31822.64	31832.74	31832.74	4586900
2021/3/8	31512.15	32148.04	31512.15	31802.44	31802.44	4969800

NASDAQ

Date	Open	High	Low	Close	Adj.Close	Volume
2020-07-01	10063.67	10197.19	10048.04	10154.63	10154.63	4595970000
2020-07-02	10268.67	10310.36	10194.06	10207.63	10207.63	4017530000
2020-07-06	10360.38	10462.05	10354.98	10433.65	10433.65	4529480000

- Finally I obtain two big datasets (ranging from 2020-07-01 to 2020-11-30) and six closing prices on 2020-12-01. And I get three datasets of three indexes (ranging from 2020-07-01 to 2021-03-10).

Cleaning

- We eliminate duplicated records of the stocks and stocks without full trading days.

Investment

In this part, I use online investment advice that was available at the beginning of July 2020 to make investment decisions and then track my investments through the fall until 1st December.

Online Advice

I search some big fund companies and look at their top ten holdings on **ETF** website.

Portfolio

From the website, I choose 6 companies that I am interested in and give them different weights. Assume an initial position of $250,000 in cash starting on 1 July 2020. After buying these 6 stocks, we still have $1.59 left. The holdings are listed in here:

company	share	amount	percent
Maderna	1218	75016.62	30 %
Zoom	193	50012.09	20 %
Uber	1643	49996.49	20 %
Corteva	924	24984.96	10 %
Pinterest	1074	25002.72	10 %
Penn	783	24985.53	10 %
Total		249998.41	100 %

Then we use donut chart to see the proportions of each sector.

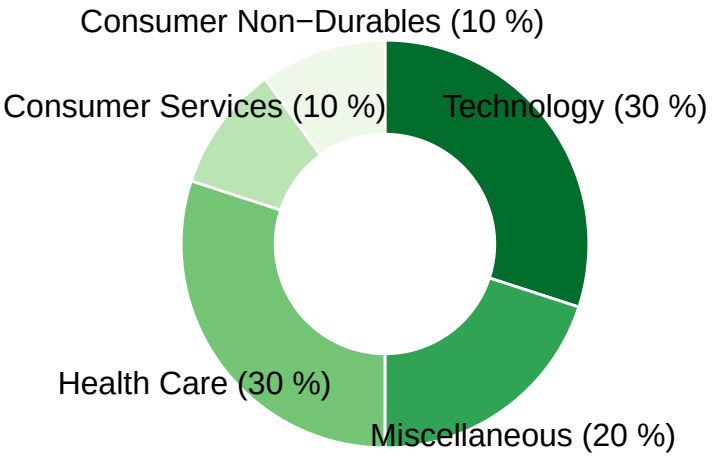

Figure 1: This shows the proportions of each sector.

As we can see from Figure1, the investment is concentrated in technology and health care sectors.

Result

Now I hold 1218 shares of Maderna, 193 shares of Zoom, 1643 shares of Uber, 924 shares of Corteva, 1074 shares of Pinterest and 783 shares of Penn National Gaming and $1.59 in cash. I track our investments using last close prices through the fall to 2020-12-01 and calculate the value of my portfolio every day.

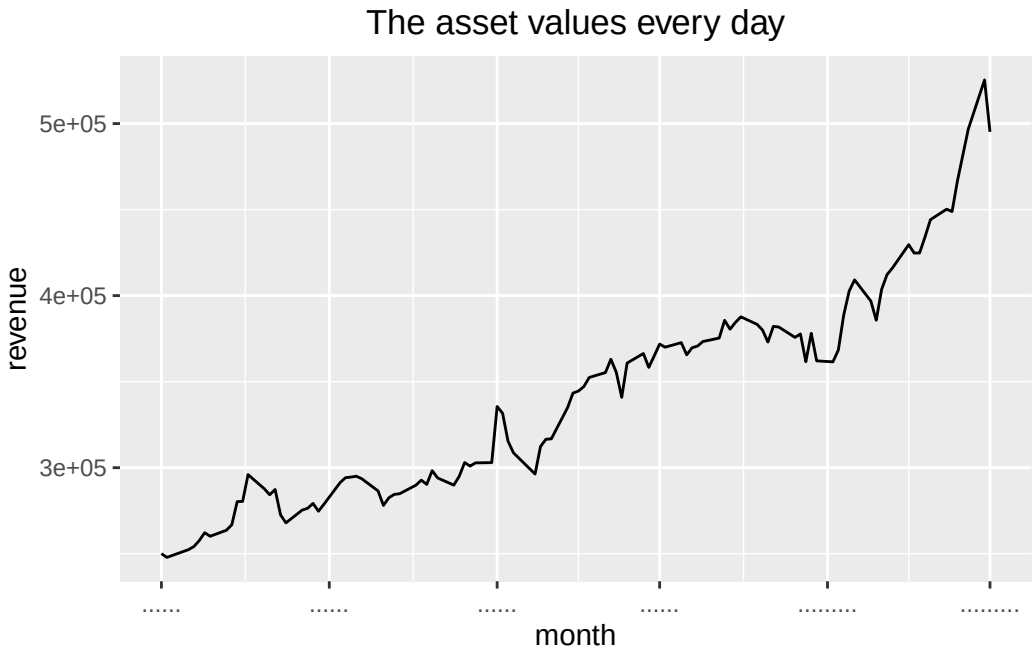

Figure 2: This shows the asset value from 2020-07-01 to 2020-12-01

date	total	ZMr	CTVAr	PINSr	UBERr	MRNAr	PENNr
2020-12-01	495201.2	78417.83	35398.44	73257.54	81542.09	171750.2	54833.49

Assume an initial position of $250,000 in cash starting on 1 July 2020. At the end of 2020-12-01, I have a position of $495,201 in cash.

Purpose and Function

In shiny I can not only display the result of our portfolio, but also show close prices of the indexes and some other stocks. This app ensure people to have a general impression on the change in stock market and make their own portfolio.
- You can choose from the sidebar to see the information of all stocks and my portfolio.
- Information of all stocks contains the index information from 2020-07-01 to 2021-03-10 and various prices of a single stock from 2020-07-01 to 2020-11-30. In control box you can choose which period and stock to display.
- I use tables and dynamic plots to show portfolio.
- I provide a way to build your own portfolio. You can choose 3 stocks and set different proportions, then track your portfolio.

Click here go to shiny app

Most Difficult Part

Getting access to high quality data is the most difficult part. Since the functions in tidyquant go wrong, I spend hours trying to use other packages to download data and failed. I can only download the latest values of indexes on YAHOO website.

Screen Shots

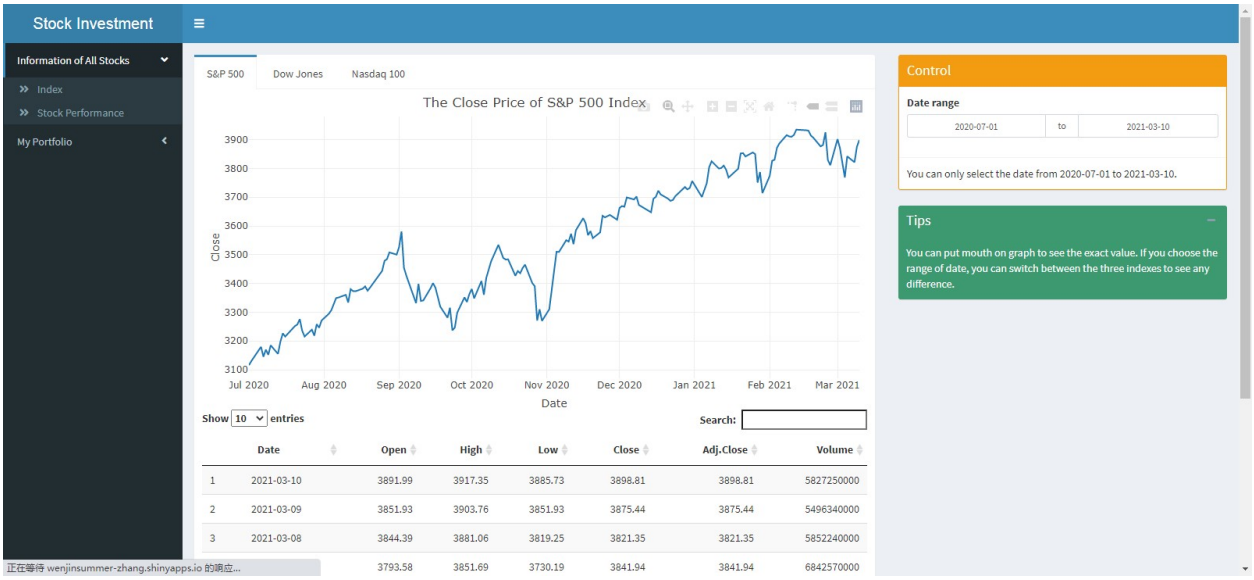

Figure 3: Screen Shot1

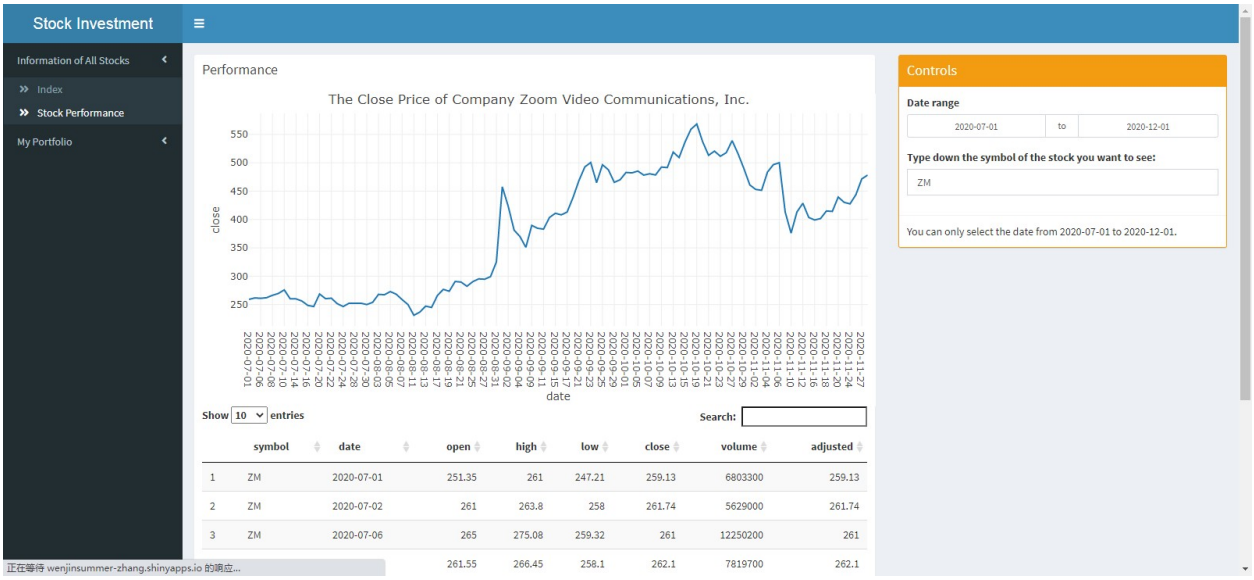

Figure 4: Screen Shot2

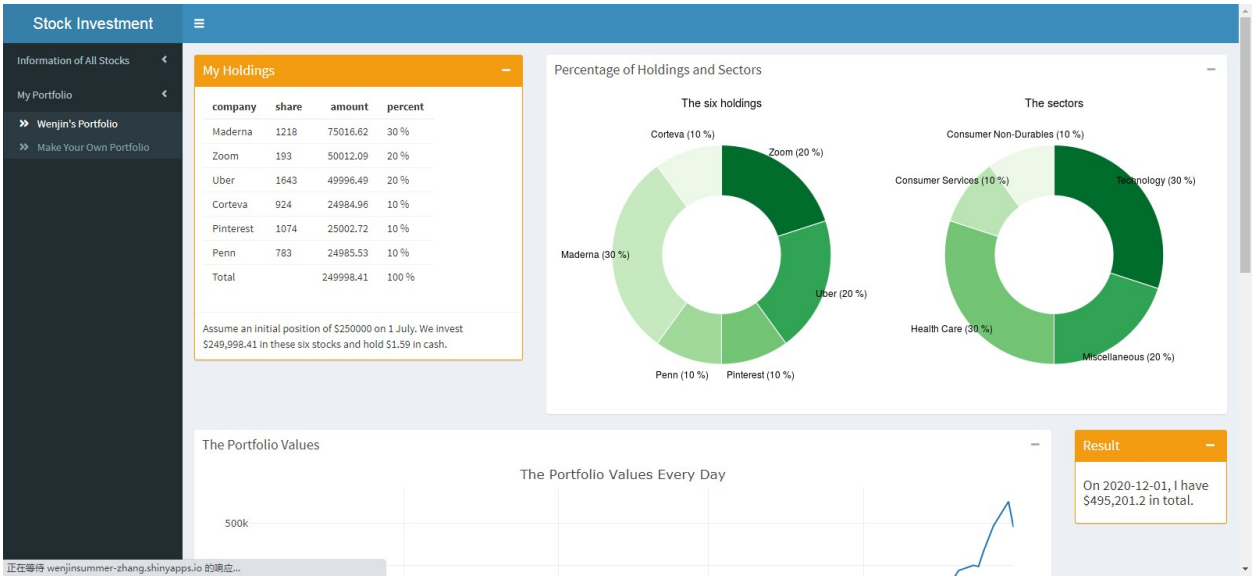

Figure 5: Screen Shot3

Figure 6: Screen Shot4

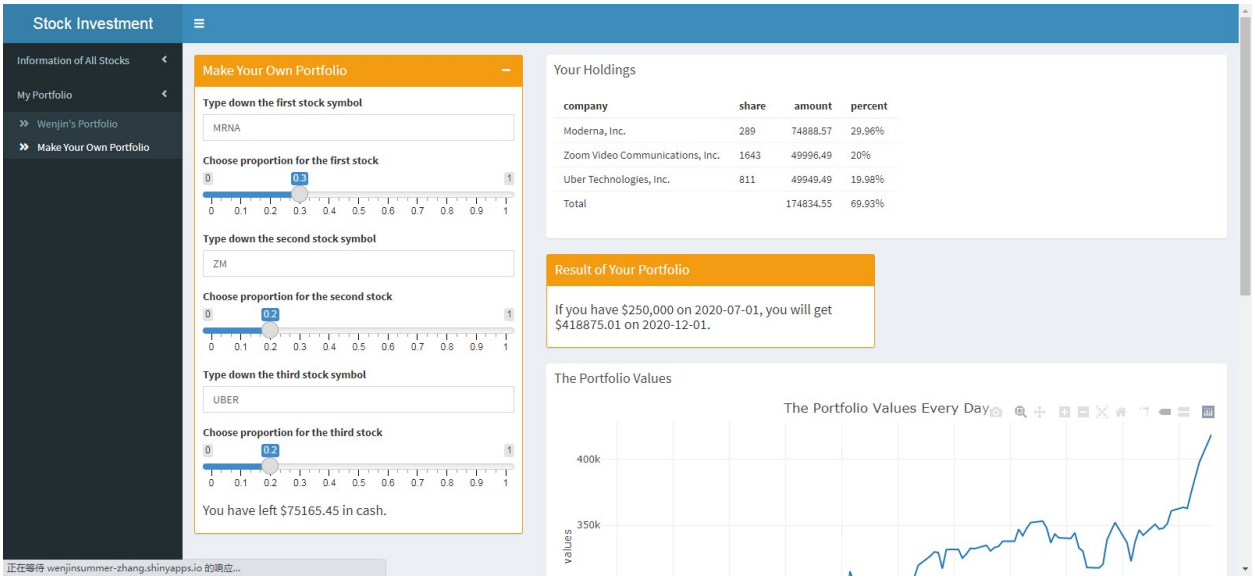

Figure 7: Screen Shot5

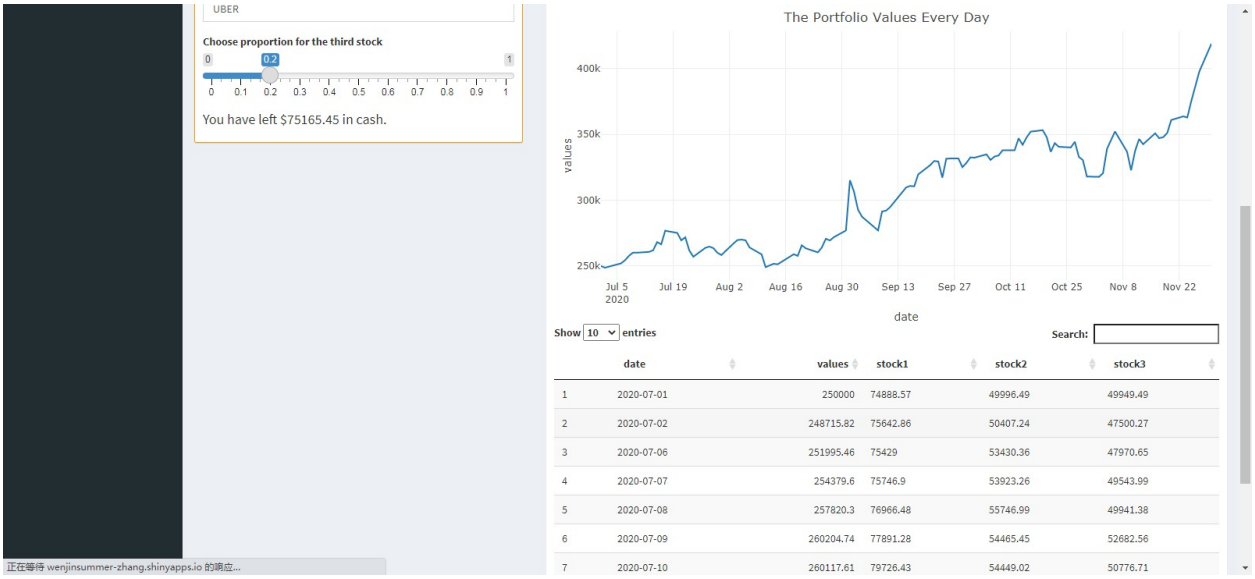

Figure 8: Screen Shot6

Inference

- The main packages I use:
  1.Matt Dancho and Davis Vaughan (2020). tidyquant: Tidy Quantitative Financial Analysis. R package version 1.0.2. https://CRAN.R-project.org/package=tidyquant
  2.Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2020). dplyr: A Grammar of Data Manipulation. R package version 1.0.2. https://CRAN.R-project.org/package=dplyr
  3.Alboukadel Kassambara (2020). ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.4.0. https://CRAN.R-project.org/package=ggpubr
  4.Stefan Milton Bache and Hadley Wickham (2014). magrittr: A Forward-Pipe Operator for R. R package version 1.5. https://CRAN.R-project.org/package=magrittr
  5.Hadley Wickham (2020). tidyr: Tidy Messy Data. R package version 1.1.2. https://CRAN.R-project.org/package=tidyr
- The data I use is downloaded from YAHOO.
- The online investment I use is from ETF.