MA615 Final Project

Chenghao Meng

2020/12/12

1 Introduction

COVID-19 has seriously influenced the stock market, making a wise investment in the US stock market has became a much harder thing for individual investors. According to BlueRocke Investment Advisors's analysis on 2020-6-26 , stocks in healthcare sector might have robust performance amid COVID-19.

During this pandemic, Remdesivir produced by Gilead Sciences **(Nasdaq:GILD)** has became a very promising cure for COVID-19. Meanwhile, on Nov. 9th, 2020, Pfizer **(NYSE:PFE)** and BioNTech **(Nasdaq: BNTX)** announced that their vaccine is more than 90% effective based on the study on the vaccine candidates .

These two campanies has been taking active roles in combating COVID-19, so the further stock analysis would be based on a portfolio consisted by Gilead Sciences **(Nasdaq:GILD)**, Pfizer **(NYSE:PFE)** and BioNTech **(Nasdaq: BNTX)** based on stock data from July 1st, 2020 to Dec.1st, 2020.

2 Stock Data Preparation and Exploration

2.1 Get Stock Data

To get stock data, we will use `tq_get` function in `tidyquant` package by specifying symbol.

```
# Get stock data
symbol <- c("GILD","PFE","BNTX")
dt <- symbol %>%
  tq_get(get="stock.prices",from = "2020-07-01", to = "2020-12-01")

# Show first 6 rows
dt2 <- as.data.frame(dt)
head(dt2)
```

```
##   symbol       date  open  high   low close  volume adjusted
## 1   GILD 2020-07-01 77.10 77.23 75.90 76.06 7786900 75.27133
## 2   GILD 2020-07-02 76.31 77.09 75.65 76.35 6381000 75.55833
## 3   GILD 2020-07-06 77.51 77.66 76.07 76.76 7056600 75.96408
## 4   GILD 2020-07-07 76.49 76.96 76.31 76.42 5055700 75.62760
## 5   GILD 2020-07-08 76.70 77.00 74.93 75.61 6293500 74.82600
## 6   GILD 2020-07-09 75.11 75.84 73.95 74.71 5199800 73.93533
```

After getting the price data of those stocks, we would also like to take a glance at the dividends of these three stocks. The data below indicates that **BNTX** has not given dividends during the given time window.

```
# Get Dividend data
dt_div <- tq_get(symbol,get="dividends",from = "2020-07-01", to = "2020-12-01")
dt_div
```

```
## # A tibble: 3 x 3
##   symbol date       value
##   <chr>  <date>     <dbl>
## 1 GILD   2020-09-14 0.68
## 2 PFE    2020-07-30 0.342
## 3 PFE    2020-11-05 0.342
```

2.2 Exploration on the Stock Data

2.2.1 Trend of Closing Price

To explore the dataset we aquired, we will do a summary first on each stock with given time window.

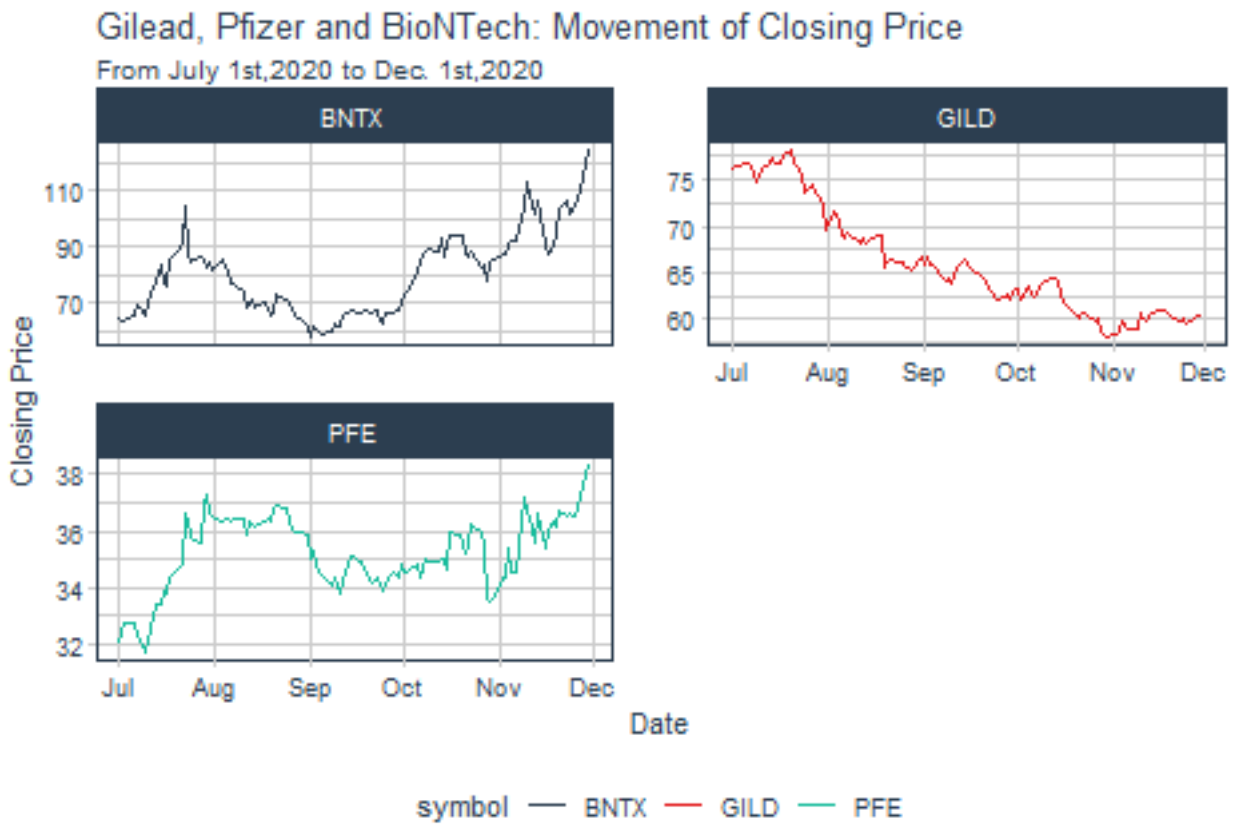
```
# Summary of GILD and PFE
sum_price <- dt %>%
  select(symbol,date,close) %>%
  group_by(symbol) %>%
  summarise(`Min Closing Price` = min(close),
            `Max Closing Price` = max(close),
            `Average Closing Price` = mean(close))
# Transform into data frame
sum_price2 <- as.data.frame(sum_price)
head(sum_price2)
```

```
##   symbol Min Closing Price Max Closing Price Average Closing Price
## 1   BNTX          57.81000            124.24              80.05057
## 2   GILD          58.15000             78.08              66.00274
## 3    PFE          31.74573             38.31              35.23087
```

The chart indicates that the minimum closing price of Gilead, BioNTech and Pfizer are $58.15, $57.81 and $31.75 respectively, and the maximum closing price of Gilead, BioNTech and Pfizer are $78.08, $124.24 and $38.31 respectively.

Then, we would like to draw line charts to demonstrate the movement of closing price between July 1st,2020 and Dec.1st, 2020.

```
# Closing price movement
dt %>% ggplot(aes(x = date, y = close, color = symbol)) +
    geom_line() +
    facet_wrap(~ symbol, ncol = 2, scales = "free_y") +
    theme_tq() +
    scale_color_tq() +
    labs(title = "Gilead, Pfizer and BioNTech: Movement of Closing Price",
         subtitle = "From July 1st,2020 to Dec. 1st,2020",
         x = "Date",y="Closing Price")
```



From the charts above, we can see that there are upper trends on BioNTech's and Pfizer's closing price. However, the movement of Gilead is showing a downward trend, which might due to the undesired result of test on Remdesivir.

2.2.2 Individual Stock Performance

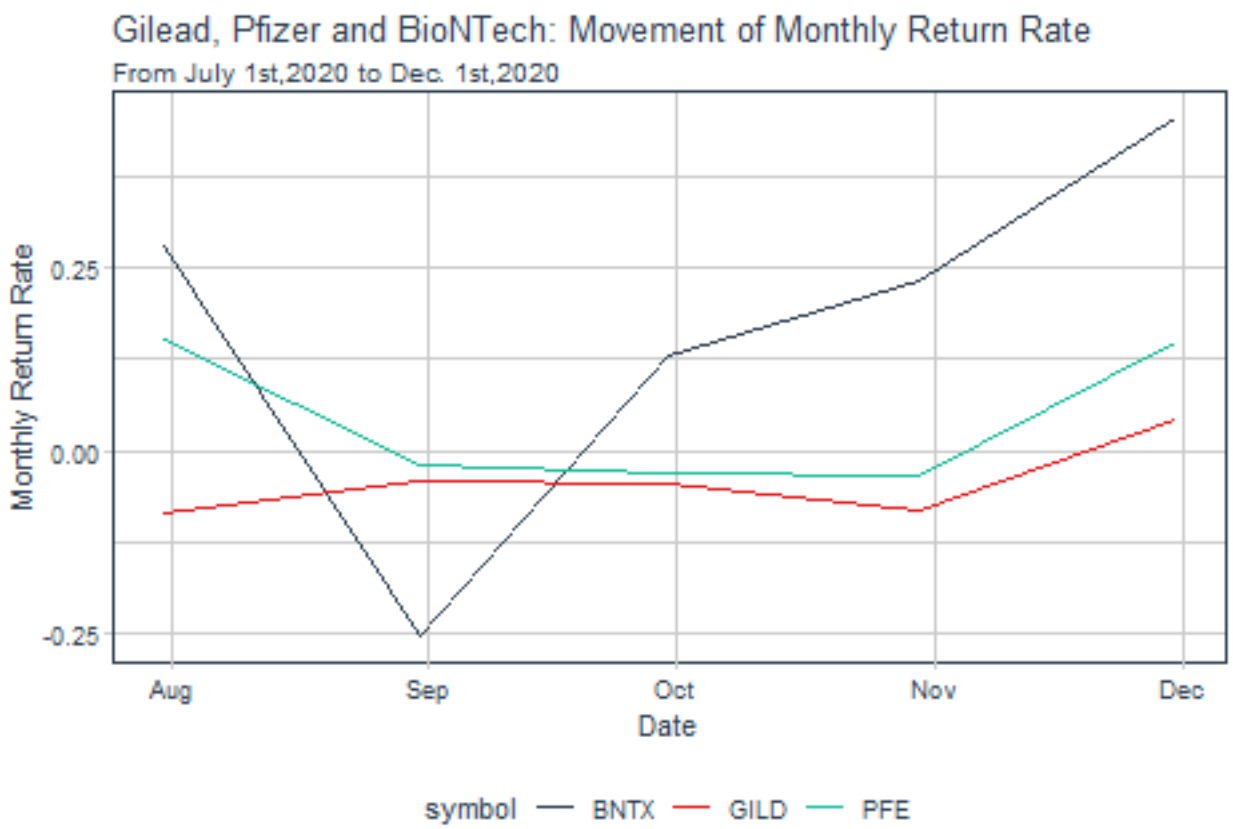
Before establishing a portfolio, we would like to assess the monthly return rate of these three stocks based on their adjusted price using `tq_transmute`.

```
# Compute monthly return rate
stock_returns_monthly <- dt %>%
    group_by(symbol) %>%
    tq_transmute(select     = adjusted,
                 mutate_fun = periodReturn,
                 period     = "monthly",
                 col_rename = "monthly.return")
```

Then, we visualize the movement of the monthly return rate of these three individual stocks: **GILD**,**PFE** and **BNTX**.

```
# Visualize: Monthly Return Rate
stock_returns_monthly %>%
  ggplot(aes(x = date, y = monthly.return, color = symbol)) +
    geom_line() +
    theme_tq() +
    scale_color_tq() +
    labs(title = "Gilead, Pfizer and BioNTech: Movement of Monthly Return Rate",
         subtitle = "From July 1st,2020 to Dec. 1st,2020",
```

```
        x = "Date",y="Monthly Return Rate")
```



The plot above indicates that the monthly return rate of **BNTX** is higher than the rest of two stocks all the time after mid-September. So, it would be better to make **BNTX** take the main proportion in the following portfolio.

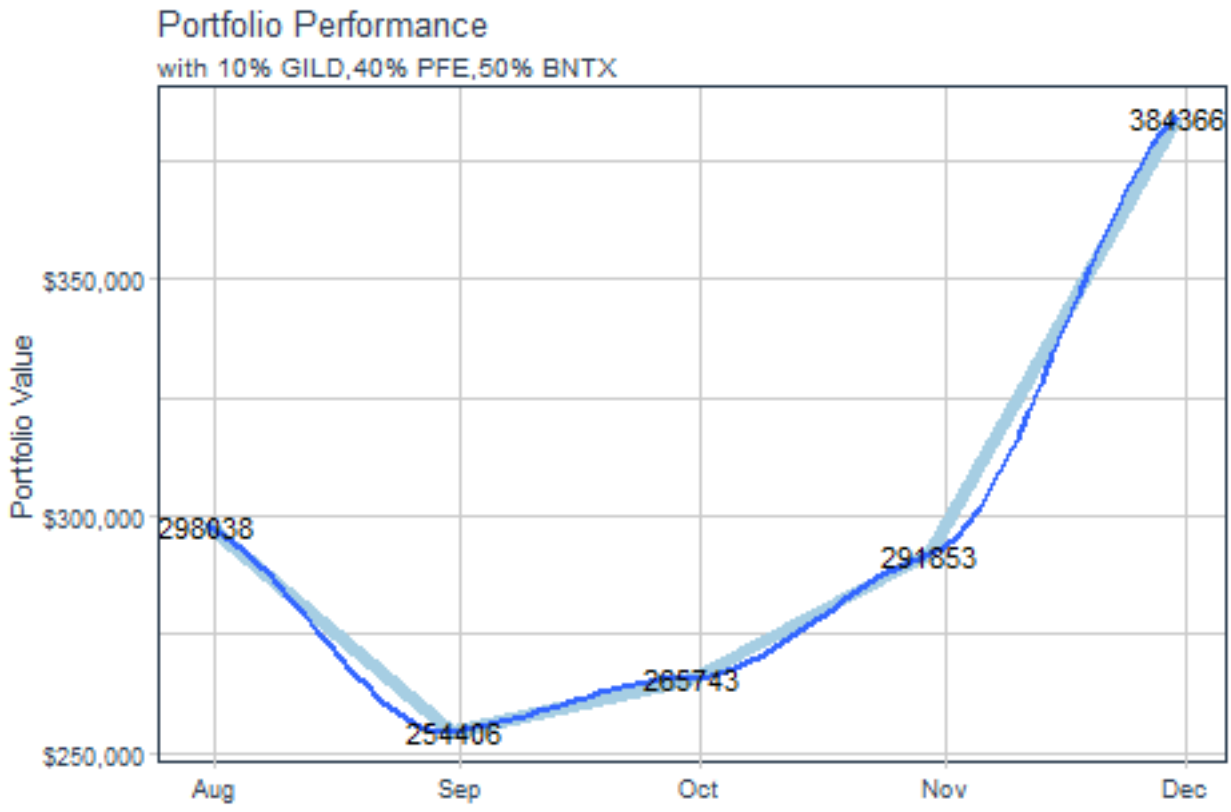
3 Portfolio: Establishment and Analysis

3.1 Portfolio Establishment

Since **BNTX** outperformed the rest of two stocks on monthly return rate, we set **BNTX**'s position to 50% for the retrospective analysis, and 10% to **GILD**, 40% to **PFE**. Meanwhile, the inital position on July 1st, 2020 would be $250,000 as required.

```
# Compute investment result with specific initial position
initial_position <- 250000
wts <- c(0.1,0.4,0.5) # proportion of GILD,PFE,BNTX
portfolio_growth_monthly <- stock_returns_monthly %>%
    tq_portfolio(assets_col   = symbol,
                 returns_col  = monthly.return,
                 weights      = wts,
                 col_rename   = "investment.growth",
                 wealth.index = TRUE) %>%
    mutate(investment.growth = investment.growth * initial_position)
```

```
# Visualize the performance
portfolio_growth_monthly %>%
    ggplot(aes(x = date, y = investment.growth)) +
    geom_line(size = 2, color = palette_light()[[5]]) +
    labs(title = "Portfolio Performance",
         subtitle = "with 10% GILD,40% PFE,50% BNTX",
         x = "", y = "Portfolio Value") +
    geom_smooth(method = "loess") +
    geom_text(aes(label = round(investment.growth)),nudge_x=0.1,nudge_y=0.1)+
    theme_tq() +
    scale_color_tq() +
    scale_y_continuous(labels = scales::dollar)
```




```
# Compute total rate of return
total_return_rate <-
(portfolio_growth_monthly$investment.growth[5]-initial_position)/initial_position
cat(paste("The total rate of return between July 1st,2020 and Dec.1st,2020 is ",
          round(total_return_rate*100,2), "%", sep=''))
```

```
## The total rate of return between July 1st,2020 and Dec.1st,2020 is 53.75%
```

The plot above indicates that the portfolio value on Nov.30th is around $384,366, and the **total rate of return** is about **53.75%**.

3.2 Result: Position Change

After creating a portofolio, we would like to know what will happen if we change the position of individual stocks in the portofolio.

Now, we would like to see the portfolio value of the following portfolios:

- (1) 10% GILD,40% PFE,50% BNTX
- (2) 20% GILD,20% PFE,60% BNTX
- (3) 30% GILD,30% PFE,40% BNTX

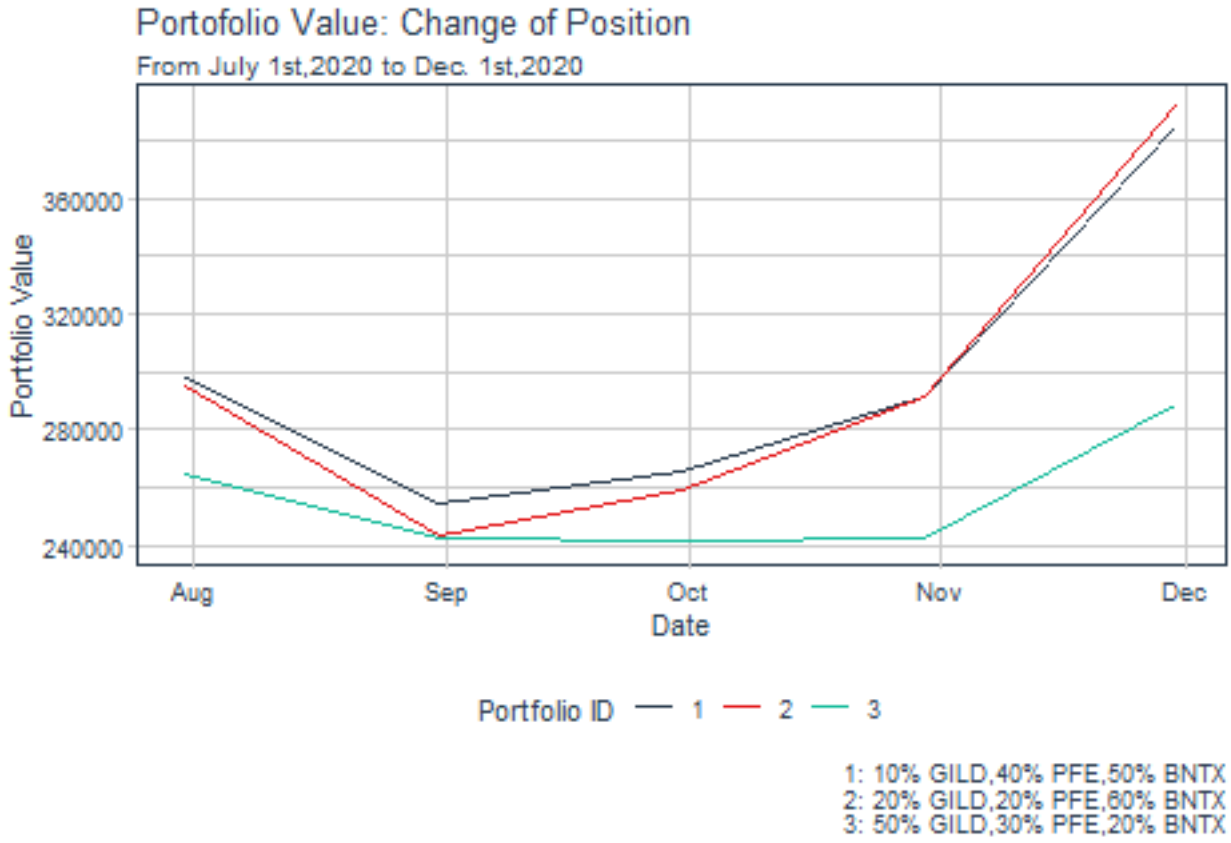
```
stock_returns_monthly_multi <- stock_returns_monthly %>%
    tq_repeat_df(n = 3)

# Set different positions
weights <- c(
    0.10, 0.40, 0.50,
    0.20, 0.20, 0.60,
    0.50, 0.30, 0.20)
# Creat weight table
weights_table <-  tibble(symbol) %>%
    tq_repeat_df(n = 3) %>%
    bind_cols(tibble(weights)) %>%
    group_by(portfolio)
```

```
# Compute the portfolio value of multi portfolio
portfolio_growth_monthly_multi <-
  stock_returns_monthly_multi %>%
    tq_portfolio(assets_col   = symbol,
                 returns_col  = monthly.return,
                 weights      = weights_table,
                 col_rename   = "investment.growth",
                 wealth.index = TRUE) %>%
    mutate(investment.growth = investment.growth * initial_position)
```

```
# Portfolio rate of return
portfolio_returns_monthly_multi <-
  stock_returns_monthly_multi %>%
    tq_portfolio(assets_col  = symbol,
                 returns_col = monthly.return,
                 weights     = weights_table,
                 col_rename  = "portfolio.returns")
```

```
# Plot the change
portfolio_growth_monthly_multi %>%
  ggplot(aes(x = date, y = investment.growth,
             col=factor(portfolio))) +
    geom_line() +
    theme_tq() +
    scale_color_tq() +
    labs(title = "Portofolio Value: Change of Position",
         caption = "1: 10% GILD,40% PFE,50% BNTX\n2: 20% GILD,20% PFE,60% BNTX\n3: 50% GILD,30% PFE,20%
         col="Portfolio ID",
         subtitle = "From July 1st,2020 to Dec. 1st,2020",
         x = "Date",y="Portfolio Value")
```



After changing the position of the stocks, the rate of return does not change significantly when **BNTX** takes the major position in the portfolio.

3.3 Beat the Baseline: S&P 500

Moreover, we would like to compare the performance of the three portfolios above with the performance of a baseline portfolio. Then, we will select a S&P 500 ETF(IVE) to be the baseline.

```
# S&P500 ETF(IVE)
baseline <- "IVE" %>%
    tq_get(get  = "stock.prices",
           from = "2020-07-01",
           to   = "2020-12-01") %>%
    tq_transmute(select     = adjusted,
                 mutate_fun = periodReturn,
                 period     = "monthly",
                 col_rename = "base.monthly.return")
```

Now, we join the monthly return data of the baseline with the monthly returns of our portfolios, and use the Capital Asset Pricing Model (CAPM) model with `tq_performance` in `tidyquant` to compare them.

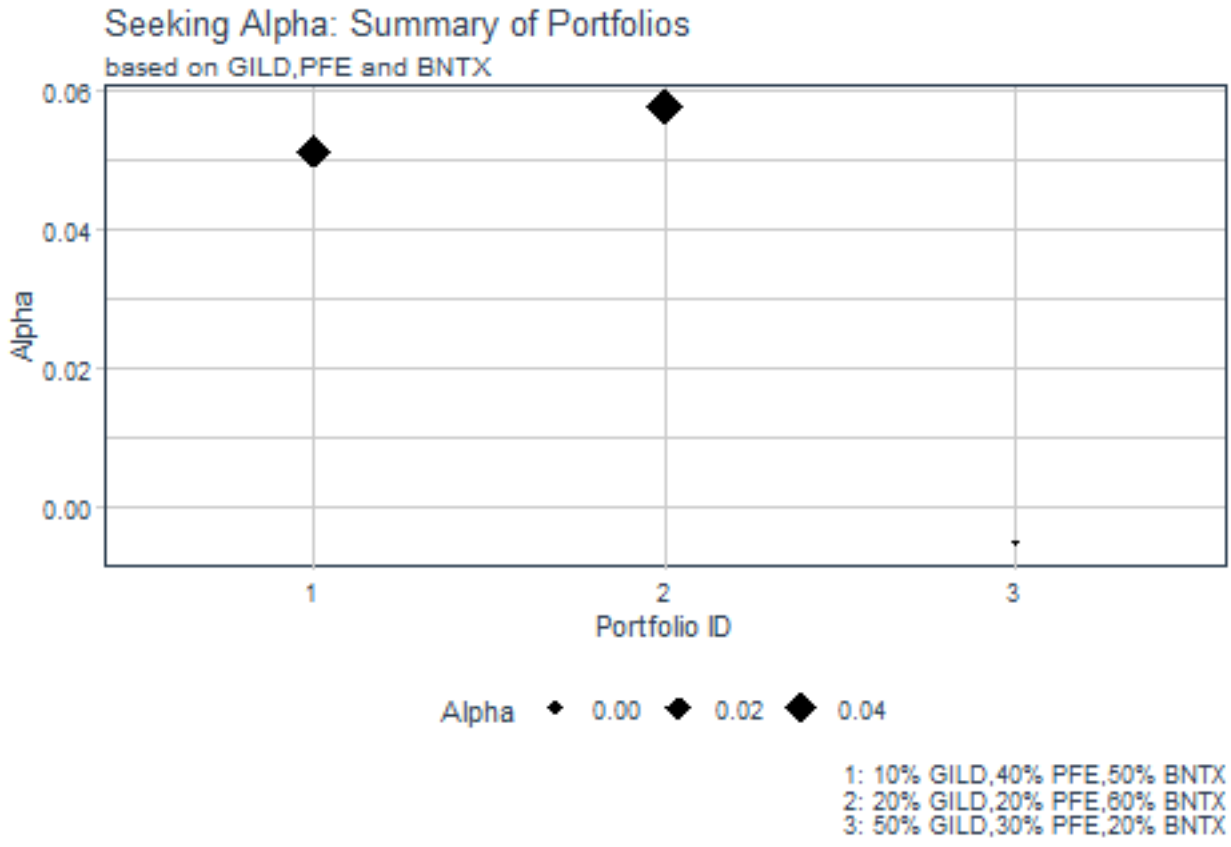
```
# Join the baseline
beat_base <- left_join(portfolio_returns_monthly_multi,
                                    baseline,
                                    by = "date")
```

```
beat_base_df <- beat_base %>%
  tq_performance(Ra = portfolio.returns,
                 Rb = base.monthly.return,
                 performance_fun = table.CAPM)
```

The plot below shows that Portfolio 2 has the highest alpha, which means the excess return of Portfolio 2 is the highest with 20% GILD,20% PFE and 60% BNTX, compared with the baseline S&P 500 ETF.

As a result, this active investment to build a portfolio has a positive outcome comparing to S&P 500.

```
# Visualize: CAPM
beat_base_df %>%
  ggplot(aes(x=factor(portfolio),y=Alpha,size=Alpha))+
  geom_point(shape=18) +
  theme_tq() +
  scale_color_tq() +
  labs(title = "Seeking Alpha: Summary of Portfolios",
       subtitle = "based on GILD,PFE and BNTX",
       caption = "1: 10% GILD,40% PFE,50% BNTX\n2: 20% GILD,20% PFE,60% BNTX\n3: 50% GILD,30% PFE,20% BI
       x="Portfolio ID",y="Alpha")
```



Reference

[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] Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019). "Welcome to the tidyverse." Journal of Open Source Software, 4(43), 1686. doi: 10.21105/joss.01686.

[3] Zvi Bodie and Alex Kane and Alan Marcus(2011). Investments. New York, McGraw-Hill/Irwin