

## ROI on Hand Picked Stocks 2007-2020

Janis Corona

2/17/2020

```
portfolio <- read.csv('all_portfolio_prices.csv', header=TRUE,
na.strings=c('', ' '),
row.names=1)

portfolio$Date <- row.names(portfolio)

Vol <- grep('Volume', colnames(portfolio))
close <- grep('Close', colnames(portfolio))
Close <- portfolio[,close]
Volume <- portfolio[,Vol]
colnames(Close)
```

## [1] "TGT.Close"	"FTR.Close"	"UBSI.Close"	"HD.Close"
"JPM.Close"			
## [6] "XOM.Close"	"CVX.Close"	"NSANY.Close"	"GNBT.Close"
"MGM.Close"			
## [11] "TEVA.Close"	"HST.Close"	"FCAU.Close"	"WFC.Close"
"WWE.Close"			
## [16] "INO.Close"	"QSR.Close"	"GRPN.Close"	"SCE.PB.Close"
"FFIN.Close"			
## [21] "GOOG.Close"	"WM.Close"	"ONCY.Close"	"S.Close"
"GM.Close"			
## [26] "F.Close"	"ASCCY.Close"	"ARWR.Close"	"COST.Close"
"AAL.Close"			
## [31] "JWN.Close"	"CSSEP.Close"	"NUS.Close"	"AMC.Close"
"ADDYY.Close"			
## [36] "KSS.Close"	"MSFT.Close"	"LUV.Close"	"HMC.Close"
"PCG.Close"			
## [41] "DLTR.Close"	"KGJI.Close"	"NKE.Close"	"AMZN.Close"
"ROST.Close"			
## [46] "TMUS.Close"	"WMT.Close"	"TJX.Close"	"TM.Close"
"PBYI.Close"			
## [51] "T.Close"	"JNJ.Close"	"C.Close"	"EPD.Close"
"VZ.Close"			
## [56] "HRB.Close"	"NFLX.Close"	"AAP.Close"	"HOFT.Close"
"SIG.Close"			
## [61] "SDC.Close"	"RRGB.Close"	"M.Close"	"JBLU.Close"
"YELP.Close"			

Remove NAs from the data. The `colSums(is.na(Close))` isn't returning the columns with NAs, so this must be done manually.

```
Close_noNAs <- Close[, -c(9,13,17,18,25,27,32,34,46,50,61,65)]
Close_noNAs$SCE.PB.Close <- as.numeric(Close_noNAs$SCE.PB.Close)
```

Add in a value of the portfolio column for each day's closing price of all stock that don't have NAs.

```
Close_noNAs$DailyValue <- rowSums(Close_noNAs, na.rm=TRUE)
```

Add in a daily change column of the portfolio closing prices.

```
dayVal <- as.data.frame(Close_noNAs$DailyValue)
colnames(dayVal) <- 'previousDayValue'
zero <- as.data.frame(as.numeric(dayVal$previousDayValue[1]))
colnames(zero) <- 'previousDayValue'
prevDay <- rbind(zero, dayVal)
Close_noNAs$prevDay <- prevDay[1:3303, 1]
dailyChange <- as.data.frame(Close_noNAs$DailyValue - Close_noNAs$prevDay)
colnames(dailyChange) <- 'dailyChange'

Close1 <- cbind(Close_noNAs, dailyChange)
```

Add a column that gives the return in dollars on initial dollars invested.

```
Close1$ROI_dollars <- Close1$DailyValue - Close1$DailyValue[1]
```

Add some date fields to look at the values by date, day of the week, month, and year in analyzing this data.

```
Close1$Date <- as.Date.character(row.names(Close1))
Close1$DayOfWeek <- weekdays(as.Date(Close1$Date))

month <- month(as.Date(Close1$Date))
Month <- month.abb[month]
Close1$Month <- Month
```

Add in the year of the Date column.

```
Year <- year(as.Date(Close1$Date))

Close1$Year <- Year

Close1$MonthYear <- paste(Close1$Month, Close1$Year, sep='-')
Close1$MonthYear <- as.factor(Close1$MonthYear)
```

Add in some unemployment information as a column to see how the portfolio is doing by date.

```
ue <- read.delim('BLS_unemploymentRates2007-2020.txt', sep=',', header=TRUE,
                na.strings=c('', ' '))
UE <- ue[, -14] #remove the empty 'Annual' column
```

Use tidyr to gather the month fields with their respective unemployment rates per month.

```
gatherMonths <- gather(UE, 'UE_Month', 'UE_monthlyRate', 2:13)

gatherMonths$MonthYear <- paste(gatherMonths$UE_Month, gatherMonths$Year,
sep='-')
gatherMonths$MonthYear <- as.factor(gatherMonths$MonthYear)

UE2 <- gatherMonths[, 3:4]
Close2 <- merge(Close1, UE2, by.x='MonthYear', by.y='MonthYear')

write.csv(Close2, 'ROI_UE_2007_2020.csv', row.names=FALSE)
```