Wrangling DJIA data

1) reading the data
2) isolating microsoft returns

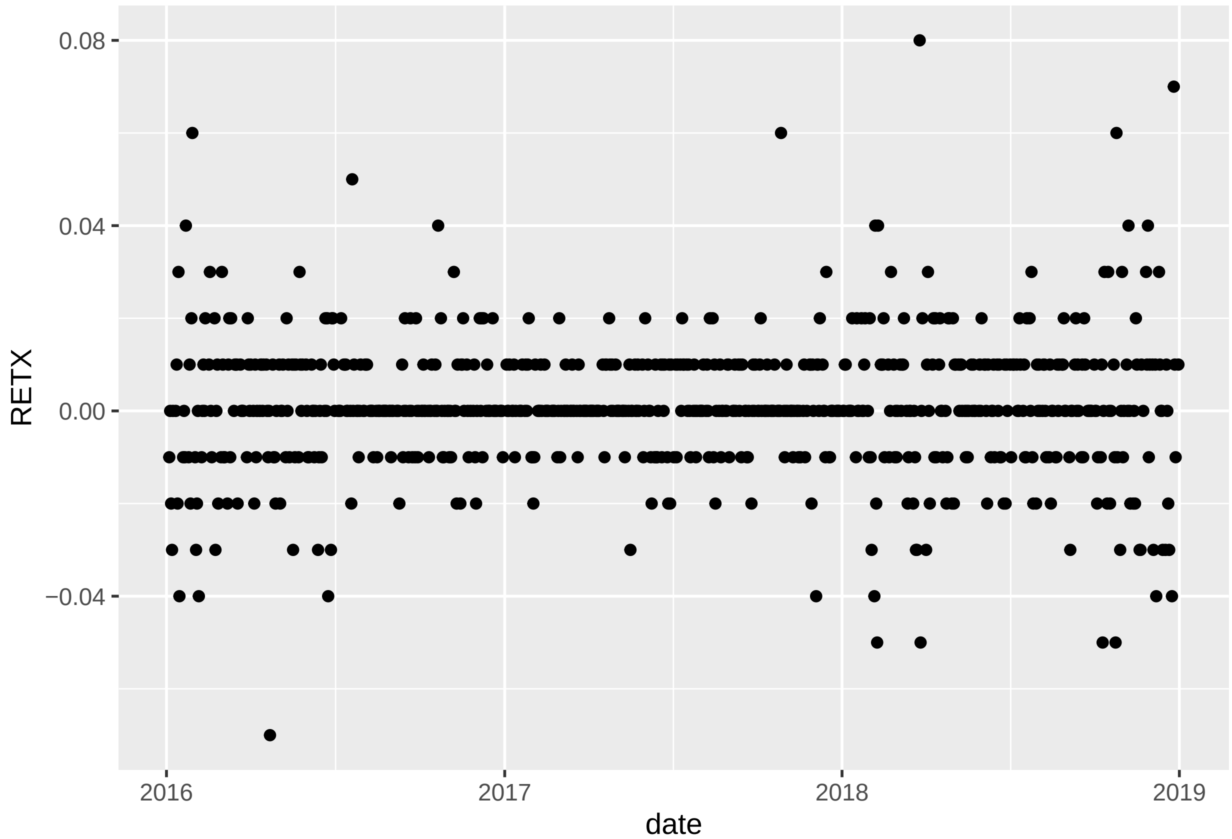
```r
library(tidyverse)
```

```
## -- Attaching packages ------------------------------------------------------------- tidyverse 1.3.0 -
```

```
## v ggplot2 3.2.1     v purrr   0.3.3
## v tibble  2.1.3     v dplyr   0.8.4
## v tidyr   1.0.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0
```

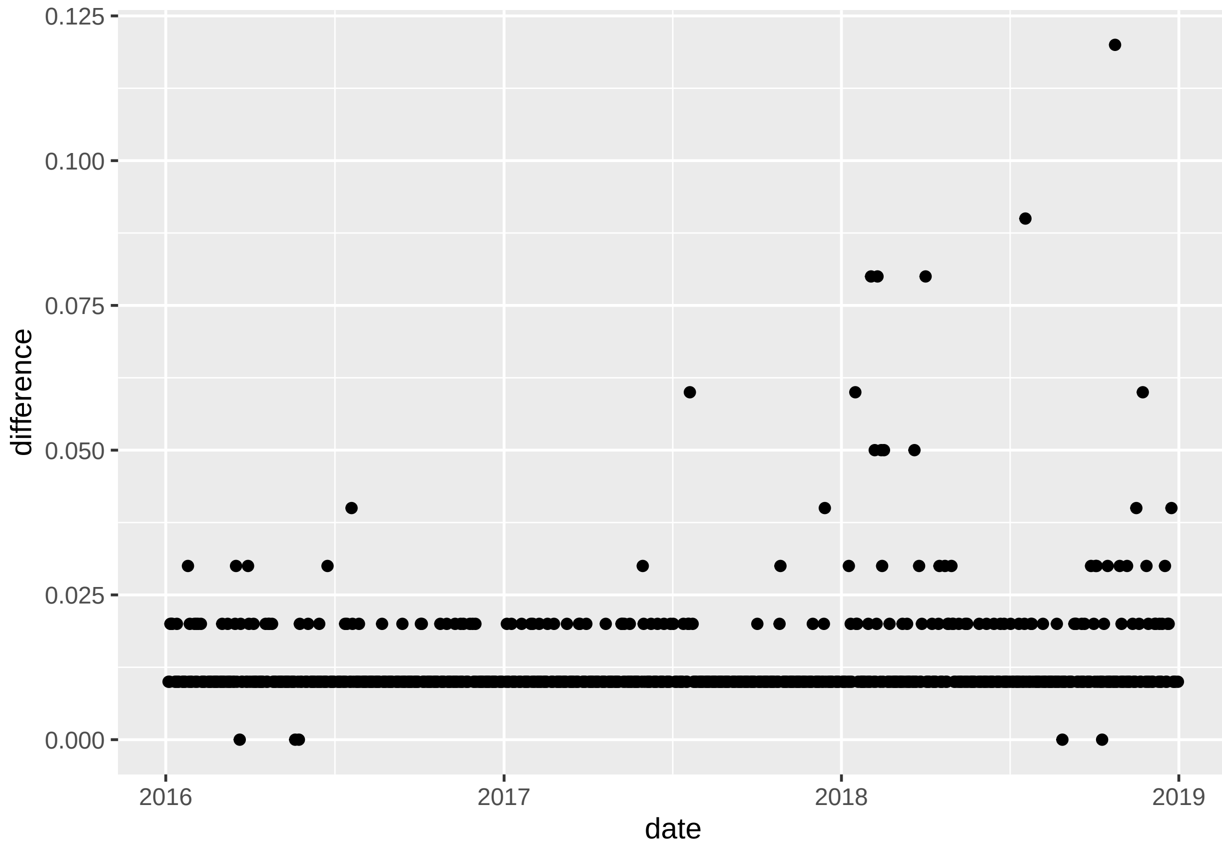
```
## -- Conflicts ---------------------------------------------------------------- tidyverse_conflicts() -
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
MSFT_Data <- please_work %>%
  filter(TSYMBOL == "MSFT")


ggplot(MSFT_Data, mapping = aes(date, RETX)) + geom_point()
```



3) calculating and plotting bid ask spreads

```
library(tidyverse)
bid_ask <- MSFT_Data %>%
  mutate(difference = ASK - BID)


ggplot(bid_ask) + geom_point(mapping = aes(date,difference))
```



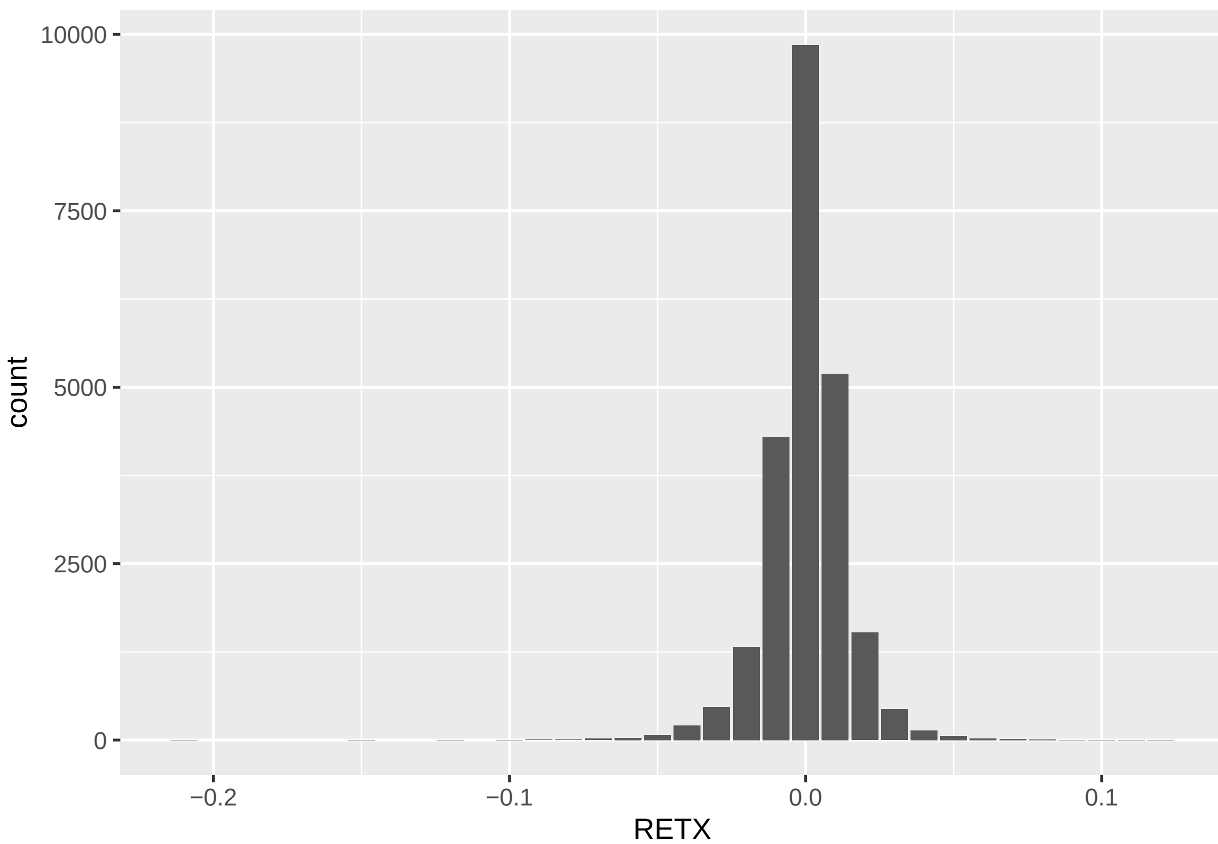

```r
mean(bid_ask$difference)
```

```
## [1] 0.01310345
```

4) calculating overeactions

```r
library(tidyverse)
library(ggplot2)
return_data <- please_work %>%
  select(date,TICKER,RETX) %>%
  filter(RETX != "C")




ggplot(return_data, mapping = aes(RETX)) + geom_histogram(stat = "count")
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```


10000
7500
5000
2500
0
count
−0.2
−0.1
0.0
0.1
RETX