FinanicalDataCorrelationAnalysis

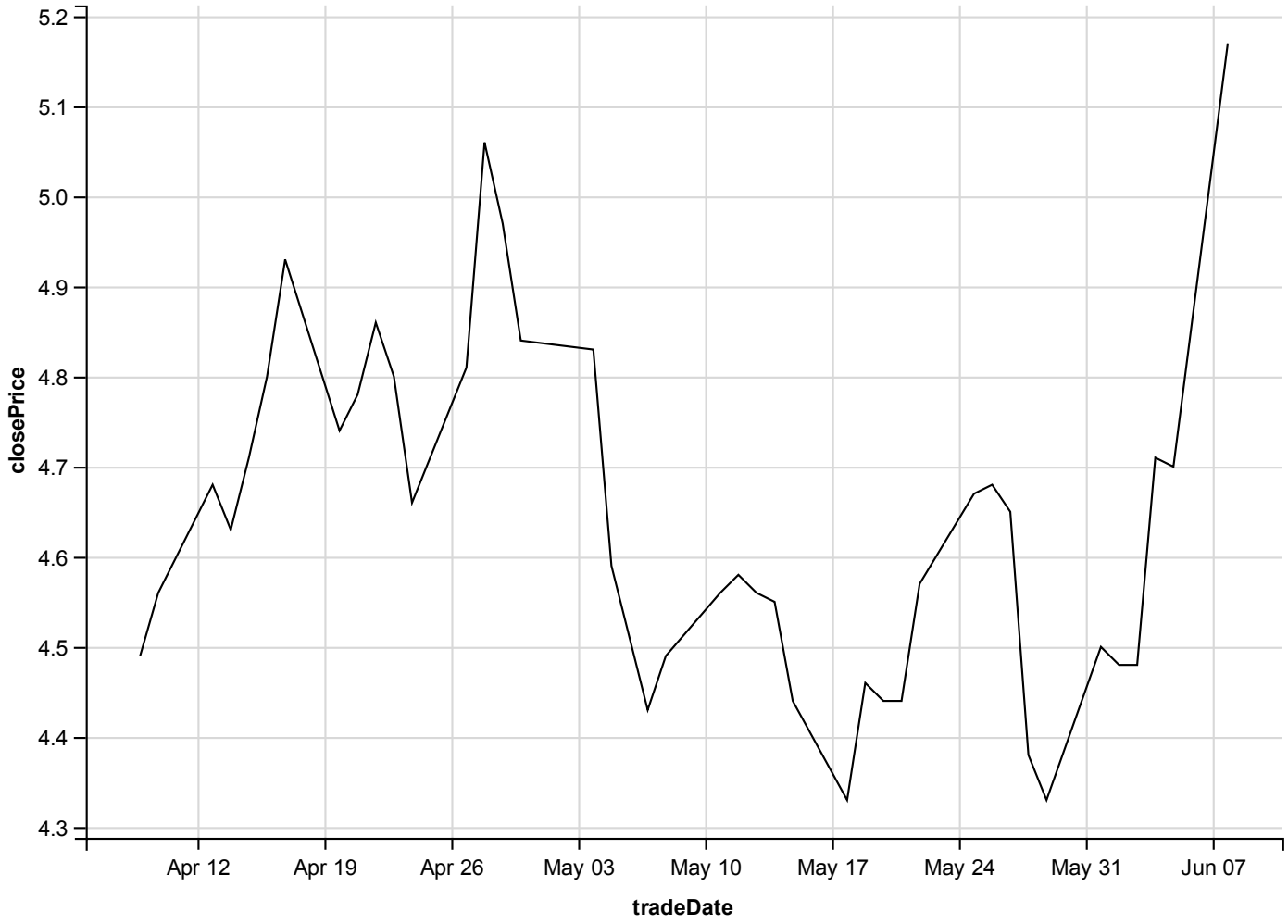
Zhi Zhang

Tuesday, June 09, 2015

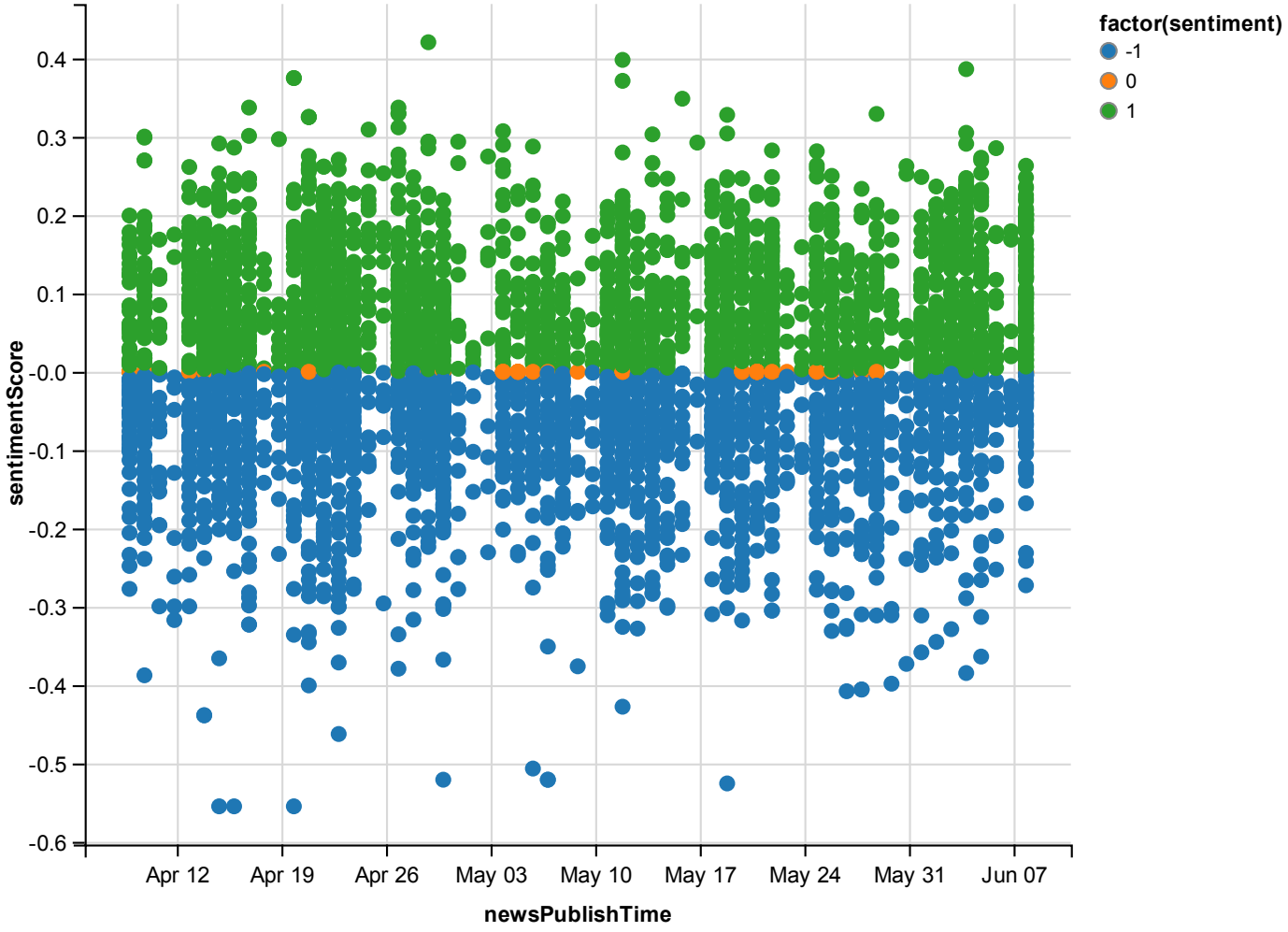
```
# include library
library(ggvis)
```

```
## Warning: package 'ggvis' was built under R version 3.1.3
```

```
# load market data
marketData <- read.csv(file="601988_mkt.csv", header=TRUE, sep=",")

# convert date string to date object
marketData[[5]] <- as.Date(marketData[[5]])

# plot market data
marketData %>% ggvis(~tradeDate, ~closePrice) %>% layer_lines()
```

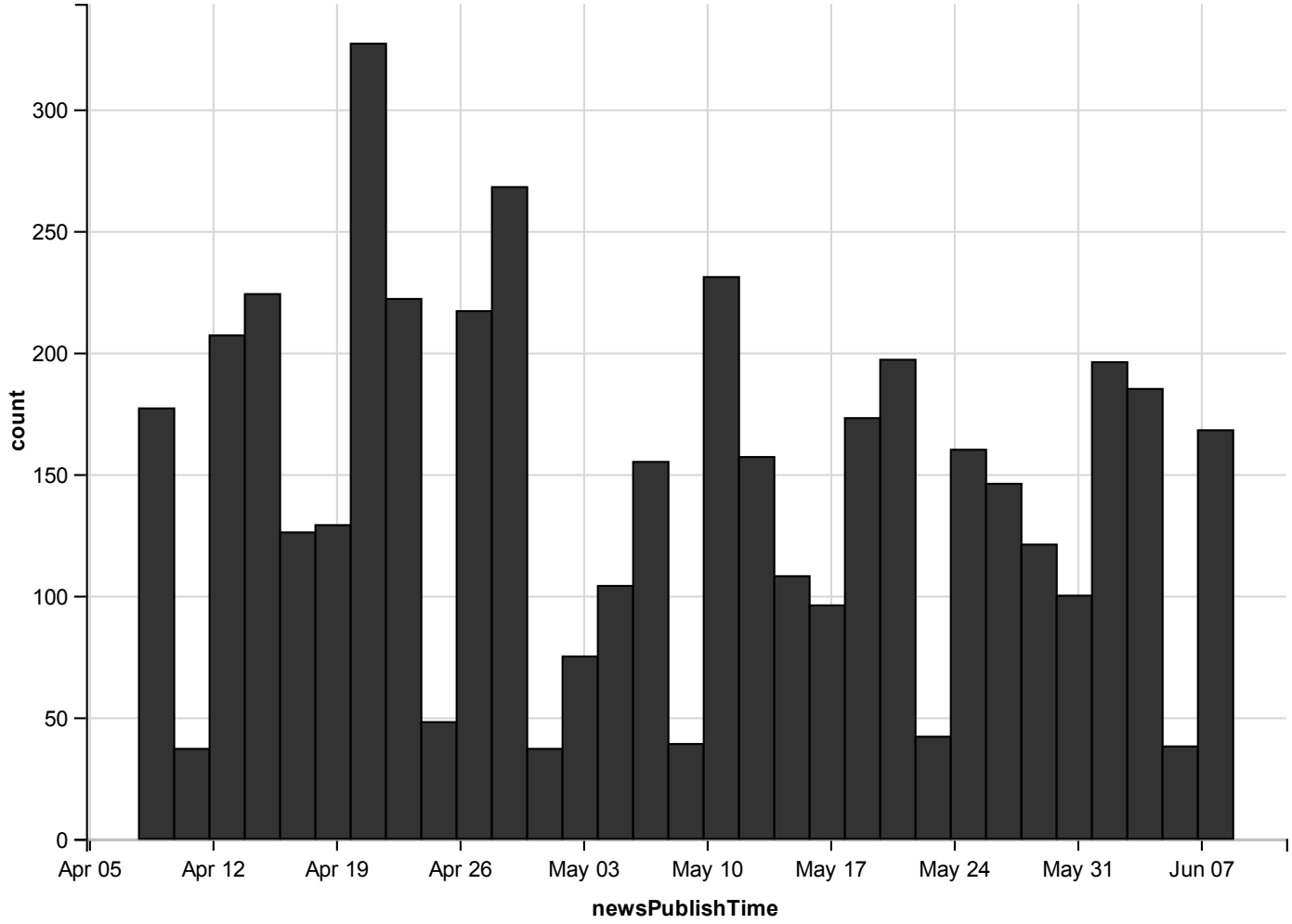


```
# load news data
newsData <- read.csv(file="601988_news.csv", header=TRUE, sep=",")

# convert date string to date object
newsData[[9]] <- as.Date(newsData[[9]])

# scatter plot news sentimental score data
newsData %>% ggvis(~newsPublishTime, ~sentimentScore) %>% layer_points(fill = ~factor(sen
timent))
```




```
# histograms of news
newsData %>% ggvis(~newsPublishTime) %>% layer_histograms()
```

```
## Guessing width = 2 days # range / 30
```


300
250
200
150
100
50
0
count
Apr 05
Apr 12
Apr 19
Apr 26
May 03
May 10
May 17
May 24
May 31
Jun 07
newsPublishTime