BTC Sports Betting

Paul Sztorc

April 13, 2016

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

```r
Use('Quandl')
```

```
## Loading required package: Quandl
## Loading required package: xts
## Loading required package: zoo
##
## Attaching package: 'zoo'
##
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```r
Btc_tx <- suppressWarnings( Quandl("BCHAIN/NTRAN") )


d2 <- t(Btc_tx)
b_tx <- as.numeric( d2[,2] )

# Format seems to be completely broken...
# ... fixing it.
Dates <- seq(from = as.Date("2016-04-12"), to = as.Date("2009-01-03"), by = "-1 day")
length(Dates) == length(b_tx) # TRUe
```

```
## [1] TRUE
```

```r
# Reassemble the data, formatted correctly
BTC_tx <- data.frame(  "date"=Dates,
                       "b.tx"=b_tx     )[-1,]

head(BTC_tx)
```

```
##         date   b.tx
## 2 2016-04-11 217378
## 3 2016-04-10 181662
## 4 2016-04-09 207555
## 5 2016-04-08 215441
## 6 2016-04-07 219915
## 7 2016-04-06 218585
```

```r
lapply(BTC_tx, class) # Date, numeric
```

```
## $date
## [1] "Date"
##
## $b.tx
## [1] "numeric"
```

```r
Use("XLConnect")
```

```
## Loading required package: XLConnect
## Loading required package: XLConnectJars
## XLConnect 0.2-9 by Mirai Solutions GmbH [aut],
##   Martin Studer [cre],
##   The Apache Software Foundation [ctb, cph] (Apache POI, Apache Commons
##     Codec, XML Commons External Components XML APIs),
##   Stephen Colebourne [ctb, cph] (Joda-Time Java library),
##   Metastuff, Ltd. [ctb, cph] (dom4j)
## http://www.mirai-solutions.com ,
## http://miraisolutions.wordpress.com
```

```r
Excel_Sheet <- "https://excelfantasyfootball.wordpress.com/2015/04/26/nfl-2015-schedule-free-excel-sprea
# Modified by me.

Data2 <- readNamedRegionFromFile(file = "C:/Users/Psztorc/Documents/GitHub/btc-sportsbet-research/nfl-20

SportsDates <- data.frame("date" = as.Date(Data2[,1]),
                          "gametype" = factor(Data2[,2]))

lapply(SportsDates, class) # Date, numeric
```

```
## $date
## [1] "Date"
##
## $gametype
## [1] "factor"
```

```r
SportsDates$daybefore <- SportsDates$date - 1
```

Merge and Plot

Time to combine the two datasets, and compare them.

I compare using a basic visualization, and a basic linear model.

```r
# BTC Transactions Near the 2015 Season
Set1 <- BTC_tx[50:300,]
Set2 <- SportsDates[,-1]
names(Set2)[2] <- "date"
All <- merge( Set1, Set2 , all = TRUE)

# Add a new factor, for days where there isn't a game tomorrow.
temp <- as.character(All[,3])
temp[ is.na(temp) ] <-  "None"
```
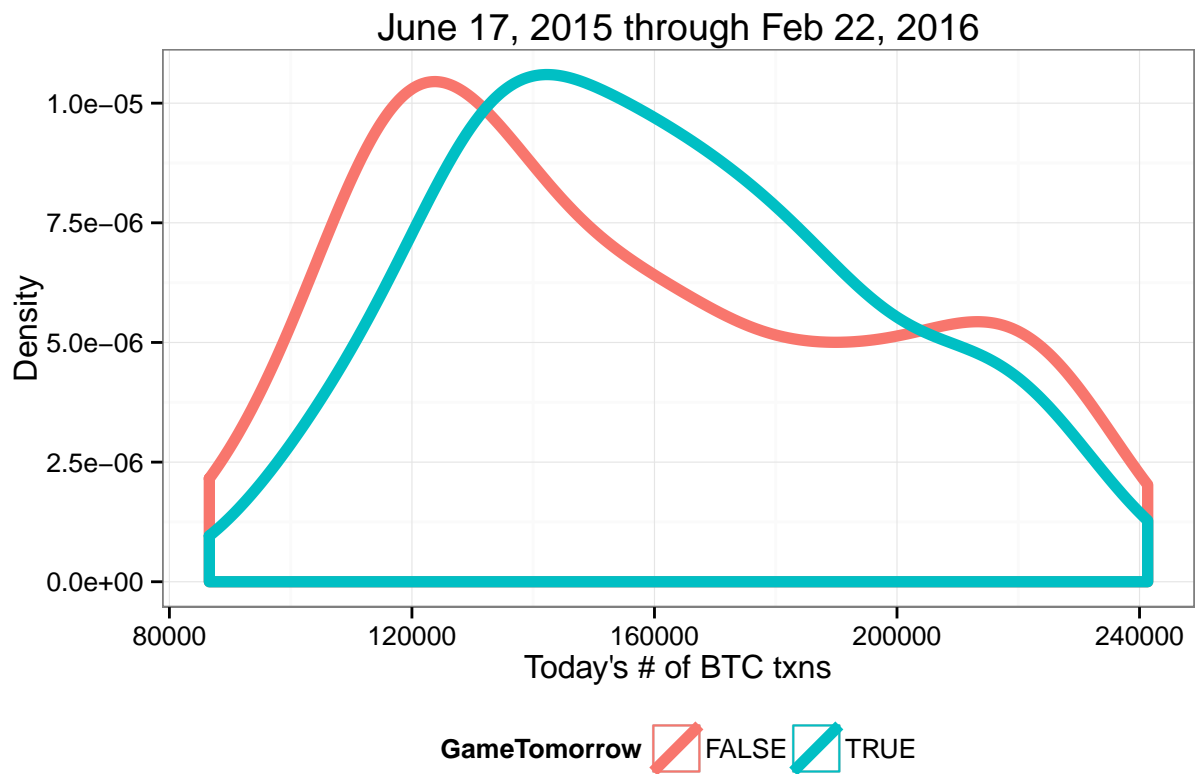
```
All[,3] <- as.factor(temp)

All$GameTomorrow <- All[,3] != "None"

Use('ggplot2')
```

```
## Loading required package: ggplot2
```

```
p1 <- ggplot(All, aes(x=b.tx, color=GameTomorrow))
p1 + geom_density(size=2) + theme_bw() + labs(x = "Today's # of BTC txns", y = "Density", title = "June
```

June 17, 2015 through Feb 22, 2016

GameTomorrow FALSE TRUE

```
m1 <- lm(b.tx ~ date + GameTomorrow, data = All)
summary(m1)
```

```
##
## Call:
## lm(formula = b.tx ~ date + GameTomorrow, data = All)
##
## Residuals:
##    Min    1Q Median    3Q    Max
## -66249 -17247  -3609  12135  97928
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)       -6.849e+06  3.870e+05 -17.698   <2e-16 ***
## date                4.189e+02  2.314e+01  18.102   <2e-16 ***
## GameTomorrowTRUE -8.610e+03  4.003e+03  -2.151   0.0324 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 26110 on 248 degrees of freedom
## Multiple R-squared:  0.5703, Adjusted R-squared:  0.5668
## F-statistic: 164.6 on 2 and 248 DF,  p-value: < 2.2e-16
```

The dates were chosen arbitrarily by me. I did not revise them (not once!), and therefore did no data dredging (you'll have to take my word for it)!

The results would appear to be that "having a football game tomorrow" is associated with about 8000 *fewer* Bitcoin transactions.

I have no idea who used what time zone, to calculate either of the two datasets that I merged! So this research is basically worthless although it was fun to do while I finished listening to the weekly Bitcoin Uncensored.