Bitcoin price and UTXO

Daran Thach

3/17/2021

## R Markdown

The blockchain universe can be divided into two main architectures for record-keeping of output transactions: 1) UTXO-Based: Bitcoin, Bitcoin Cash, Litecoin, Dash, and 2) Account-Based: Ethereum, Ethereum Classic. Beyond being a critical architecture of some of the main blockchains in the market, UTXOs are also a unique source of intelligence, and thus may help us understand unique characteristics of buyers and sellers. Using blockchain.com data, I used R to import, find descriptive stats, cleaned, prepared UTXO/price data of BTC for the last 2 years, in order to compile correlation analysis between the two.

knitr::opts\_chunk$set(echo = TRUE)  
  
#import the bitcoin excel values  
btcdata <- read.csv("btc\_utxo\_price.csv")  
  
#find descriptive stats  
summary(btcdata)

## endweek utxo\_average price\_usdavg   
## 01/07/19: 1 Min. :48914834 Min. : 3468   
## 01/07/20: 1 1st Qu.:50271451 1st Qu.: 6262   
## 01/14/20: 1 Median :51775368 Median : 7270   
## 01/15/19: 1 Mean :55305002 Mean : 7252   
## 01/21/19: 1 3rd Qu.:61309328 3rd Qu.: 8853   
## 01/21/20: 1 Max. :65817352 Max. :11624   
## (Other) :99

#find variance and price samples  
var(btcdata)

## Warning in var(btcdata): NAs introduced by coercion

## endweek utxo\_average price\_usdavg  
## endweek NA NA NA  
## utxo\_average NA 3.433538e+13 7734145830  
## price\_usdavg NA 7.734146e+09 4747903

head(btcdata$price\_usdavg)

## [1] 9150.75 8485.76 8258.40 7081.33 7015.15 7988.84

#convert factor to date data  
as.Date.character(btcdata$endweek, tryFormats = c("%m/%d/%Y"))

## [1] "0018-03-18" "0018-03-26" "0018-04-02" "0018-04-09" "0018-04-16"  
## [6] "0018-04-23" "0018-04-30" "0018-05-07" "0018-05-14" "0018-05-21"  
## [11] "0018-05-28" "0018-06-04" "0018-06-11" "0018-06-18" "0018-06-25"  
## [16] "0018-07-02" "0018-07-09" "0018-07-16" "0018-07-23" "0018-07-30"  
## [21] "0018-08-06" "0018-08-13" "0018-08-20" "0018-08-27" "0018-09-03"  
## [26] "0018-09-10" "0018-09-17" "0018-09-24" "0018-10-01" "0018-10-08"  
## [31] "0018-10-15" "0018-10-22" "0018-10-29" "0018-11-05" "0018-11-12"  
## [36] "0018-11-19" "0018-11-26" "0018-12-03" "0018-12-10" "0018-12-17"  
## [41] "0018-12-24" "0018-12-31" "0019-01-07" "0019-01-15" "0019-01-21"  
## [46] "0019-01-28" "0019-02-04" "0019-02-11" "0019-02-18" "0019-02-25"  
## [51] "0019-03-04" "0019-03-11" "0019-03-18" "0019-03-25" "0019-04-01"  
## [56] "0019-04-08" "0019-04-15" "0019-04-22" "0019-04-29" "0019-05-06"  
## [61] "0019-05-13" "0019-05-20" "0019-05-27" "0019-06-03" "0019-06-10"  
## [66] "0019-06-17" "0019-06-24" "0019-07-01" "0019-07-08" "0019-07-15"  
## [71] "0019-07-22" "0019-07-29" "0019-08-05" "0019-08-12" "0019-08-19"  
## [76] "0019-08-26" "0019-09-02" "0019-09-09" "0019-09-16" "0019-09-23"  
## [81] "0019-09-30" "0019-10-07" "0019-10-14" "0019-10-21" "0019-10-28"  
## [86] "0019-11-04" "0019-11-11" "0019-11-18" "0019-11-25" "0019-12-02"  
## [91] "0019-12-09" "0019-12-16" "0019-12-23" "0019-12-30" "0019-12-31"  
## [96] "0020-01-07" "0020-01-14" "0020-01-21" "0020-01-28" "0020-02-04"  
## [101] "0020-02-11" "0020-02-18" "0020-02-25" "0020-03-03" "0020-03-04"

# Visualize time series and plots  
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 3.6.3

library(tidyr)

## Warning: package 'tidyr' was built under R version 3.6.3

library(dplyr)

## Warning: package 'dplyr' was built under R version 3.6.3

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

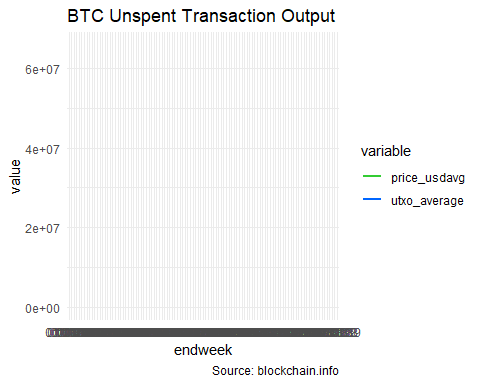
## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

# Multiple ts plot change aesthetics and labels  
chart <- btcdata %>%  
 select(endweek, utxo\_average, price\_usdavg) %>%  
 gather(key = "variable", value = "value", -endweek)  
head(chart, 3)

## endweek variable value  
## 1 03/18/18 utxo\_average 52547089  
## 2 03/26/18 utxo\_average 51456131  
## 3 04/02/18 utxo\_average 51224774

g <- ggplot(chart, aes(x = endweek, y = value)) +   
 geom\_line(aes(color = variable), size = 1) +  
 scale\_color\_manual(values = c("#33CC33", "#0066FF")) +  
 theme\_minimal()  
g <- g + labs(title = "BTC Unspent Transaction Output",   
 caption = "Source: blockchain.info")  
g

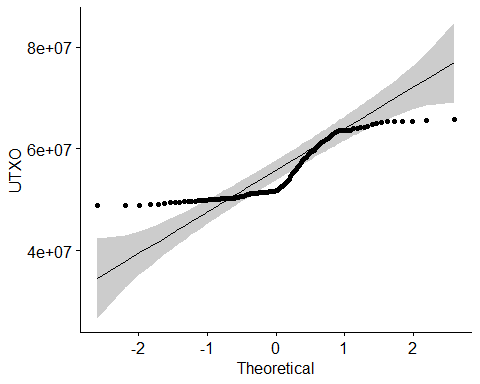
## geom\_path: Each group consists of only one observation. Do you need to adjust  
## the group aesthetic?



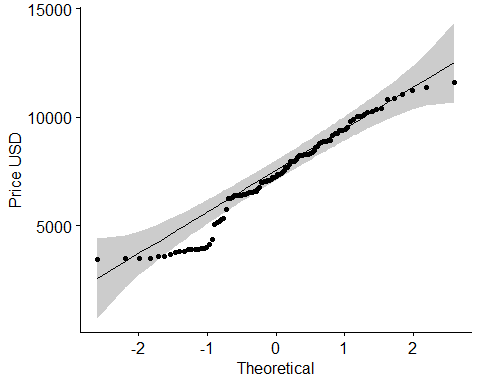
# Correlate the ts plots by determining linearity and nml dist  
library(ggpubr)

## Warning: package 'ggpubr' was built under R version 3.6.3

# plot for utxo  
ggqqplot(btcdata$utxo\_average, ylab = "UTXO")



# plot for price  
ggqqplot(btcdata$price\_usdavg, ylab = "Price USD")

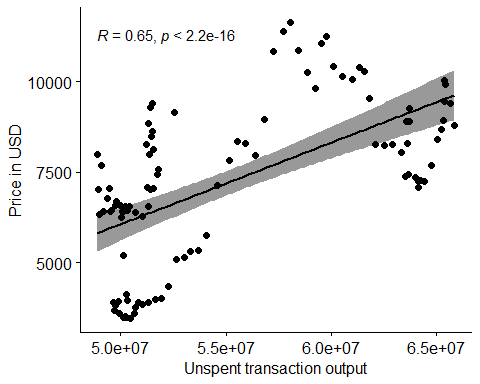


# Use Spearman rank correlation coefficient  
rank <-cor.test(btcdata$utxo\_average, btcdata$price\_usdavg, method = "spearman")  
rank

##   
## Spearman's rank correlation rho  
##   
## data: btcdata$utxo\_average and btcdata$price\_usdavg  
## S = 68344, p-value < 2.2e-16  
## alternative hypothesis: true rho is not equal to 0  
## sample estimates:  
## rho   
## 0.6457392

ggscatter(btcdata, x = "utxo\_average", y = "price\_usdavg",   
 add = "reg.line", conf.int = TRUE,   
 cor.coef = TRUE, cor.method = "spearman",  
 xlab = "Unspent transaction output", ylab = "Price in USD")

## `geom\_smooth()` using formula 'y ~ x'

 # End of R script The above outputs generated time series plots upon changing aesthetics/labels, then applied statistical linearity testing, and Spearman rank correlation. Revealed UTXO moderately correlated to price of BTC.