Bitcoin Exchange Rate

Shreekar Gonguntla Reddy
School of Information Technology
Illinois State University sgonugu@ilstu.edu

October 22, 2014

Abstract

Bitcoin is the first decentralized digital currency.a type of digital currency that is created and maintained electronically. Bitcoin is not a physical currency like dollars or euros, but it can be used in the same way to buy things electronically.

1 Introduction

Bitcoin was created by a software engineer **Satoshi Nakamoto** who proposed it as an electronic payment system based on mathematical proof. Bitcoins have a number of advantages, bitcoins are transferred directly person to person, so there would be not involvement of a bank or any third party, which results in low transfer fees. Also, it can be used anywhere across the world, as it is a global currency and does not depend on the geographic location of the user. The community responsible for authorizing the transactions and securing the networks are called miners.

There are several currency exchanges available to buy and sell bitcoin, like dollar, euro, yang, etc. Bitcoins are stored on a individuals personal devices like mobile or system on their digital wallets. Bitcoins can be used to purchase anything online, just like the way you use dollars, euro, etc. Bitcoins are the best ways for a business to minimize their transition fee and also it does not charge them any additional charges for accepting bitcoins.

Bitcoin price is calculated on the basis of supply and demand, when the demand increases the price increases, and when demand is low, then the price falls. If new bitcoins are created and introduced then they will be introduced with less price, and their price will rise only once the demand increases.

2 Exchange Rate Analysis

For our observation we are going to gather data from Quandl () for the last 30 days of bitcoin exchange rate and analyze the trends.

```
> # Extracting data from Quandl
>
> library(Quandl)
> mydata <- Quandl("BITCOIN/BITSTAMPUSD", authcode="PxxFfireRAiixyqysWeJ")
```

The above shows us how we can get the data from Quandl into R for our analysis. To get the data from Quandl API we need to know just two small things, the dataset name and the authorization code which is unique for each user. We try and load all the data into a data frame in R for our further analysis. The data extracted will contain all the rows and columns without any restrictions. We need to perform data cleaning operations to get the data that we want.

```
> #for no. of rows and columns
> dim(mydata)

[1] 1136    8

> #Displaying the data from Quandl
> head(mydata,10)

         Date   Open   High    Low  Close Volume (BTC) Volume (Currency)
1  2014-10-22 385.79 388.87 380.00 386.59     5470.030           2099827
2  2014-10-21 381.00 393.55 378.98 385.81     7953.323           3078104
3  2014-10-20 389.06 390.65 376.20 381.21    12081.799           4614454
4  2014-10-19 390.59 394.25 385.00 387.53     3242.374           1262683
5  2014-10-18 383.20 397.29 377.00 390.58     7074.825           2746148
6  2014-10-17 384.19 386.00 370.96 383.65    10507.826           3989422
7  2014-10-16 394.52 399.00 370.10 383.95    22776.515           8706162
8  2014-10-15 402.01 404.32 385.92 394.52    19147.690           7537827
9  2014-10-14 392.99 417.99 390.48 403.38    24822.162          10060958
10 2014-10-13 378.48 405.00 368.07 392.60    26083.746          10034403
   Weighted Price
1        383.8785
2        387.0212
3        381.9343
4        389.4317
5        388.1577
6        379.6620
7        382.2429
8        393.6677
9        405.3216
10       384.6995
```

The above command displays only first 10 rows from Quandl. The data that we get from Quandl is generally not the final data there might be many rows or columns that we wont need. Here, we need the data just for columns *Volume*

(BTC),Volume Currency,Weighted Price, rest columns we dont want the data and also we wnat to check for last 30 days, so we need to filter out the data and take only the lst 30 days data

2.1 Scrubing of Data

Generally all the data that you get is not gerenally useful to you. Similarly, in our case we need just last 30 days data and only the Volume BTC, Volume Currency and Weighted Price.

```
> #Restricting data to just the first 30 rows
> mydata <- mydata[1:30,]
> #Selecting only the required columns
> mydata1 <- mydata[,c(1,6:8)]
> #Renaming the column names
> colnames(mydata1) <- c("Date","Volume_BTC","Volume_Currency","Weighted_Price")
> #writing the data into csv file
> write.csv(mydata,"Bitcoin_exchange_rate.csv")
```

Using the above code we have selected the last 30 days data and also just for the columns we want. mydata1 is now holding the final data that we want for this analysis. Also we would want the data in csv format so that we can provide it as a support for our analysis and also help others in reproducing the result.

2.2 Exploring of Data

Exploring the data that we have gathered is as important as analyzing the data. We need to know the data we have and the format of the data and the data types and the amount of data we have.

```
> summary(mydata1)
```

```
      Date              Volume_BTC    Volume_Currency    Weighted_Price
 Min.   :2014-09-23   Min.   : 3242   Min.   : 1262683   Min.   :305.8
 1st Qu.:2014-09-30   1st Qu.:10438   1st Qu.: 3996427   1st Qu.:365.1
 Median :2014-10-07   Median :15578   Median : 6189163   Median :383.1
 Mean   :2014-10-07   Mean   :20315   Mean   : 7388456   Mean   :377.1
 3rd Qu.:2014-10-14   3rd Qu.:24357   3rd Qu.: 9541542   3rd Qu.:389.1
 Max.   :2014-10-22   Max.   :69538   Max.   :21937271   Max.   :428.9
```

The above code tells us the summary of the data. It tells us the max,min,mean values,etc. for the dataset and also other properties of the data.

```
> str(mydata1)
```

```
'data.frame':        30 obs. of  4 variables:
 $ Date           : Date, format: "2014-10-22" "2014-10-21" ...
 $ Volume_BTC     : num  5470 7953 12082 3242 7075 ...
 $ Volume_Currency: num  2099827 3078104 4614454 1262683 2746148 ...
 $ Weighted_Price : num  384 387 382 389 388 ...
```

The above code tells us the data types for different columns and the form in which the data is present. Here we see that the Date column is in Date data type while the other three are in number.

```
> head(mydata1,10)
```

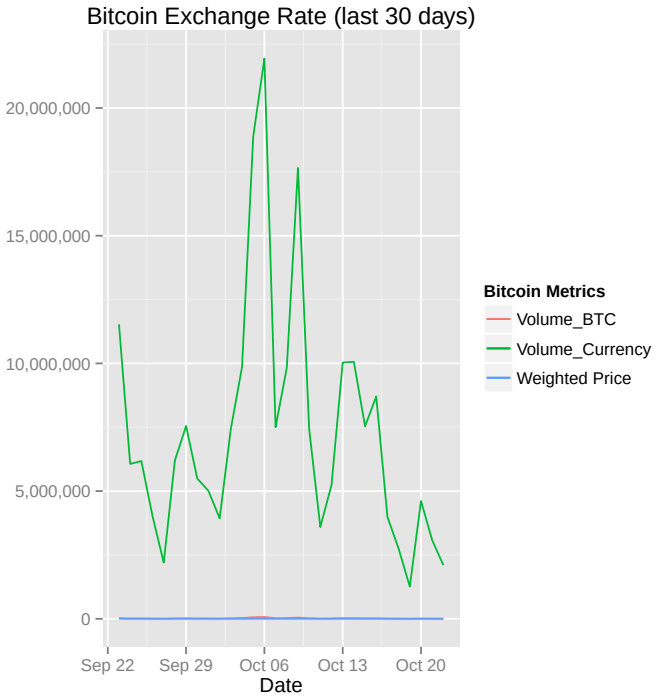
```
         Date Volume_BTC Volume_Currency Weighted_Price
1  2014-10-22   5470.030         2099827       383.8785
2  2014-10-21   7953.323         3078104       387.0212
3  2014-10-20  12081.799         4614454       381.9343
4  2014-10-19   3242.374         1262683       389.4317
5  2014-10-18   7074.825         2746148       388.1577
6  2014-10-17  10507.826         3989422       379.6620
7  2014-10-16  22776.515         8706162       382.2429
8  2014-10-15  19147.690         7537827       393.6677
9  2014-10-14  24822.162        10060958       405.3216
10 2014-10-13  26083.746        10034403       384.6995
```

The above code displayes the first 10 rows of data.

2.3 Plotting of Graph

Just getting tabular data is never enough for analysis. Analyzing requires proper understanding of the data, which we can usually get from visualizations.

```
> library(ggplot2)
> library(scales)
> #Plotting of Graph
> dd <- ggplot(data=mydata1, aes(x=Date)) + ylab(" ")+
+   geom_line(aes(y = Volume_BTC, colour = "Volume_BTC")) +
+   geom_line(aes(y = Volume_Currency, colour = "Volume_Currency")) +
+   geom_line(aes(y = Weighted_Price,color = "Weighted Price")) +
+   scale_y_continuous(labels=comma)+scale_colour_hue(name="Bitcoin Metrics")
> dd + ggtitle("Bitcoin Exchange Rate (last 30 days)")
```



The above graph shows the *Volume BTC, Volume Currency and Weighted price* trends for the last 30 days. We see that the Volume Currency values are fluctuaing in millions and the average value od Volume Currency is 7388456.22. The Volume BTC also has fluctuating trends with a average value of 20314.51, the average value of weigthed price is 377.08.

The above graph tells us the trends and the differences in the values of the exchange metrics. We can clearly make out from it that the volume usage of currency is very high compared to the volume usage of the BTC.