

# Explorative DataViz Short Paper

*Ji, Ye*

*10/15/2017*

## Assignment Requirement

This assignment builds off the last assignment, resulting in a single document either in PDF or HTML form exported from an R markdown document. You must use the data exclusively from your portfolio proposal! Further, this document should have:

- Five data visualizations of most distinct graph types, made with ggplot2; This can include the three visualizations from last week, though you can replace them if you want. Note that graph requirements always include:
  - Correct usage of all visual encodings;
  - Appropriate data sourcing;
  - Proper labeling of ALL visual encodings, as well as an appropriate title and subtitle;
  - The code you used to generate each graphic, right above the graph (R Markdown should make this easy).
- A written narrative for each graph, from three to five sentences, describing the context of the graph and how it informs about your policy topic. You can write this as a narrative.

Note that, over time, you are expected to tackle progressively more varied and ambitious graph types. At this stage in the course, the standard is still not very high. That said, you should be able to clearly articulate the value of the graphs you created so far - meaning they need to make sense and have some value-add.

Please submit the URL of the file on a Git repository (it can be one git repo you keep adding to, or a stand alone repository).

## Introduction

I am dedicating my data visualization portfolio and 10-week effort to digital currency. There are several advantages in doing so:

+ Massive and Robust data source: Blockchain is a public transaction log that exists as a distributed database, validated by its powerful algorithm.

+ Fitting data types: Blockchain data includes time and IP addresses, and can be updated frequently, which means it is useful for both the static and dynamic data visualizations.

+ Well-supported: I can save the entire blockchain file locally, use community-maintained tools to parse it, and keep the data up-to-date.

+ Policy-setting: Digital currencies are being used as store of value to fight inflation in turbulent economies such as Zimbabwe and Latin American countries. Central Banks are issuing experimental policies in response to the dissemination of digital currency and blockchain technology.

While I'm waiting for my local blockchain file (150 GB and growing) to finish download, I will explore some secondary data sources published by major Bitcoin exchanges and platforms to better understand the data.

## Graph one

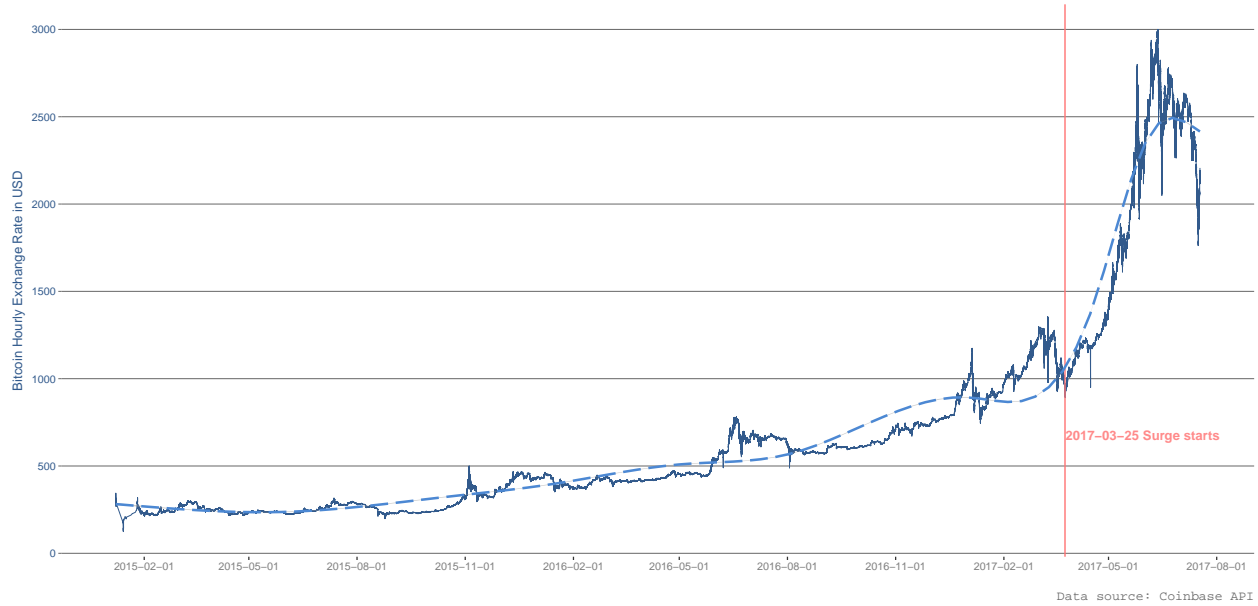
```
line <- ggplot(data=coinbase_lim) +
  geom_line(size=.25,
    aes(unix_timestamp, price),
    color="#325a8c") +
  geom_vline(xintercept = as.POSIXct("2017-03-25"),
    colour="#ff7575",
    size=.75,
    alpha=.75) +
  geom_smooth(aes(unix_timestamp,price),
    span=2,
    linetype="longdash",
    color="#4c88d3",
    alpha=.5) +
  scale_x_datetime(name = "",
    date_breaks = "3 month") +
  scale_y_continuous(name = "Bitcoin Hourly Exchange Rate in USD",
    breaks = c(0, 500, 1000, 1500, 2000, 2500, 3000)) +
  ggtitle("Bitcoin Daily Price",
    subtitle = "experienced the greatest surge in history this year") +
  labs(caption = "Data source: Coinbase API") +
  annotate(geom="text",
    x=as.POSIXct("2017-05-30"),
    y=675,
    label="2017-03-25 Surge starts",
    colour="#ff7575",
    fontface="bold",
    alpha=.85) +
  theme(panel.grid.major.y = element_line( size=.1, color="#666666"),
    panel.grid.major.x = element_blank(),
    panel.background = element_blank(),
    plot.title = element_text(size=18,
      family = "Helvetica",
      colour = "#3a3a3a",
      face = "bold"),
    plot.subtitle = element_text(size=12, family = "mono", colour = "#666666"),
    axis.title.y.right = element_text(color="#85bb65"),
    axis.text.y.right = element_text(color="#85bb65"),
    axis.text.y = element_text(color="#325a8c"),
    axis.title.y = element_text(color="#325a8c"),
    axis.ticks.y = element_line( size=.25, color="#666666"),
    axis.text.x = element_text(color="#7f7f7f"),
    legend.position = "none",
    plot.caption = element_text(size=11, family = "mono", colour = "#666666"),
    plot.margin = unit(c(2,2,2,2), "cm"))

line

## `geom_smooth()` using method = 'gam'
```

## Bitcoin Daily Price

experienced the greatest surge in history this year



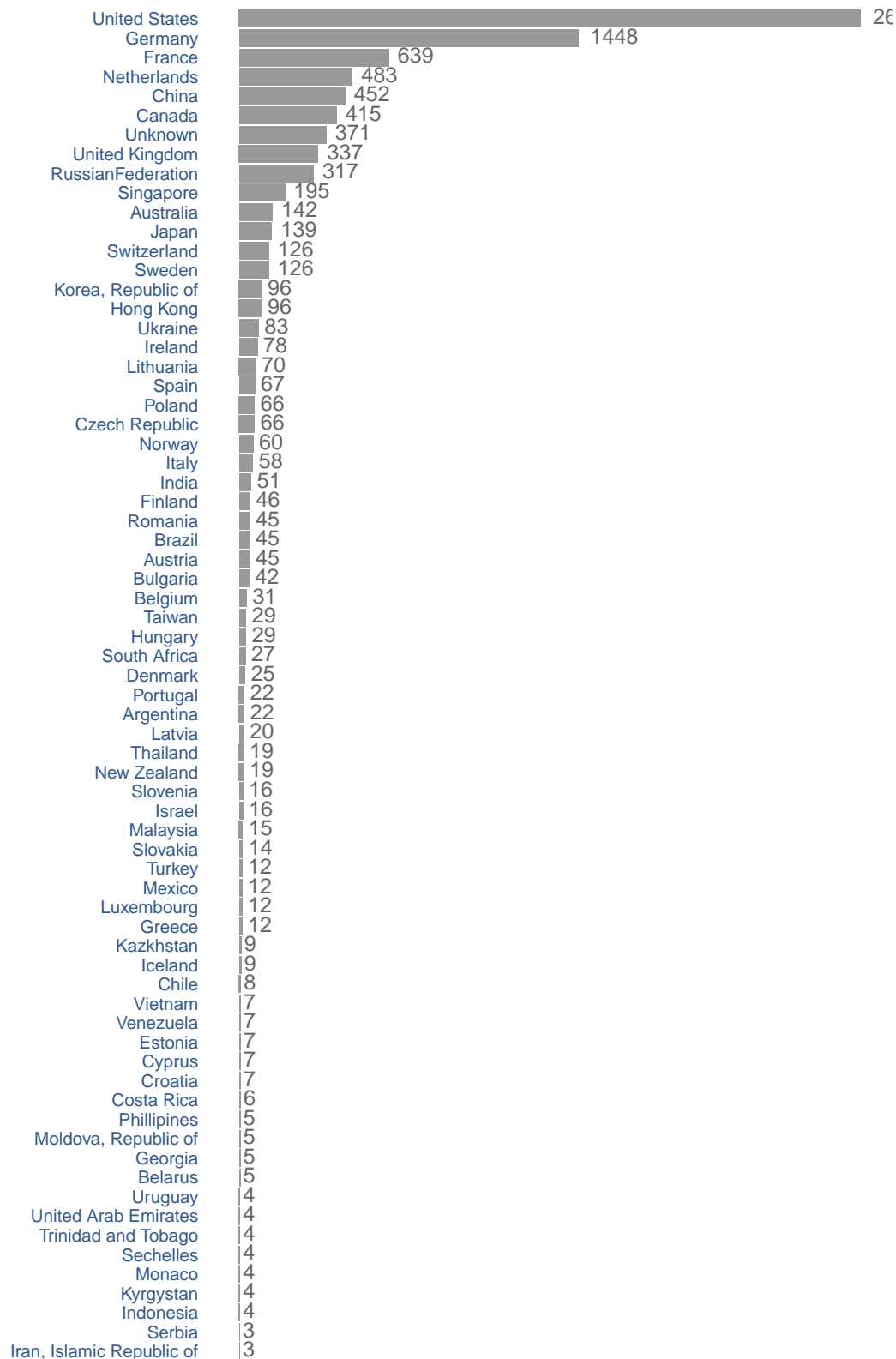
## Graph two

```
hist <- ggplot(data=nodes, aes(reorder(factor(country), node), node)) +
  geom_col(width=0.9, position=position_dodge(width=5), fill = "#999999") +
  geom_text(aes(label=node),
    color="#666666",
    position=position_dodge(width=0.9),
    vjust=0.35,
    hjust=-.25) +
  ggtitle("Bitcoin Nodes Distribution by Countries",
    subtitle = "Reachable nodes as of Oct 15, 2017") +
  ylab("") +
  xlab("") +
  labs(caption = "Data source: BitNodes.21.co") +
  theme(panel.grid.major.y = element_blank(),
    panel.grid.major.x = element_blank(),
    panel.background = element_blank(),
    plot.title = element_text(size=18,
      family = "Helvetica",
      colour = "#3a3a3a",
      face = "bold"),
    plot.subtitle = element_text(size=12, family = "mono", colour = "#666666"),
    axis.text.y = element_text(color="#325a8c"),
    axis.title.y = element_text(color="#325a8c"),
    axis.ticks.y = element_blank(),
    axis.ticks.x = element_blank(),
    axis.text.x = element_blank(),
    legend.position = "none",
    plot.caption = element_text(size=11, family = "mono", colour = "#666666"),
    plot.margin = unit(c(2,2,2,2), "cm")) +
  scale_x_discrete(expand=c(0,0)) +
  coord_flip()
```

hist

# Bitcoin Nodes Distribution by Countries

Reachable nodes as of Oct 15, 2017



Data source: BitNodes.21.co

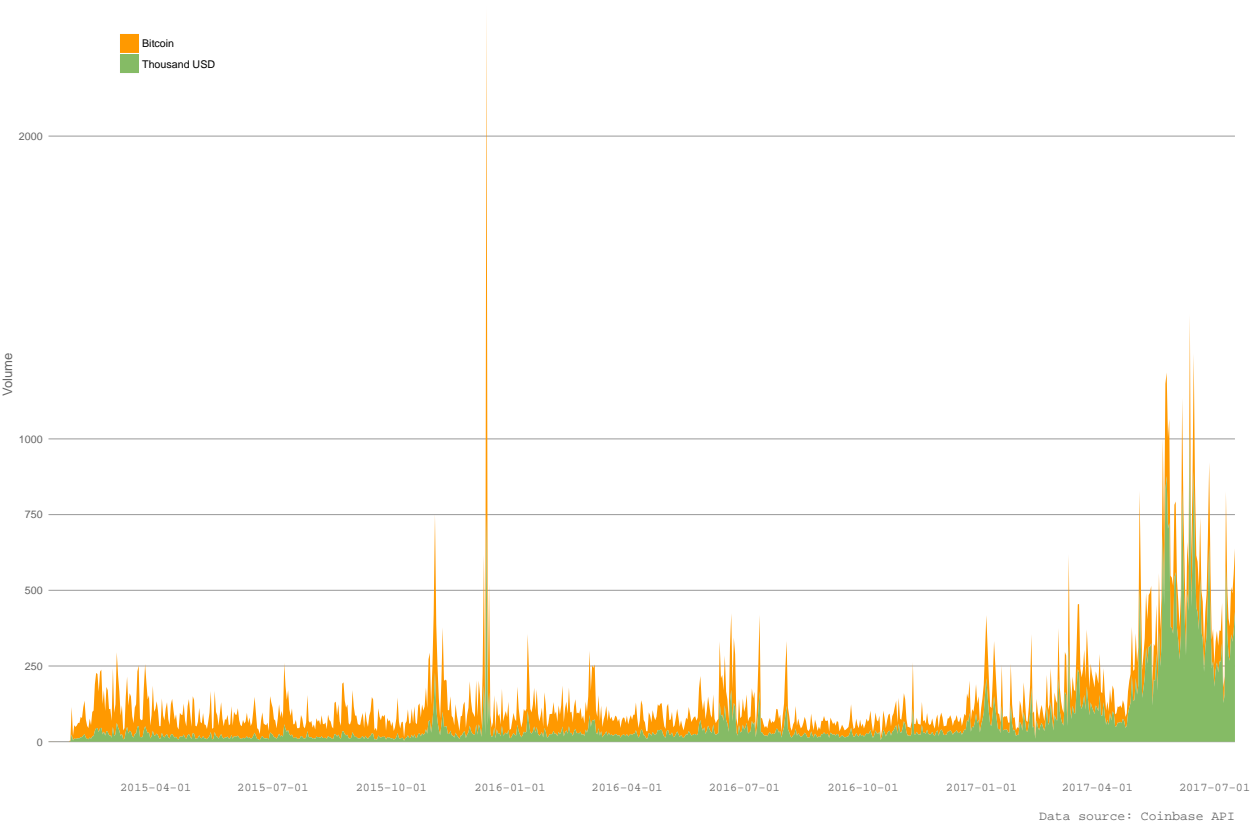
## Graph Three

```
# note: color #85bb65 is nicknamed as "dollar bill green".
#       color #FF9900 is the orange from bitcoin logo.
area <- ggplot(data=agg, aes(x = as.POSIXct(date), y = volume, fill = as.factor(unit))) +
  geom_area(position = "stack") +
  scale_x_datetime(date_breaks = "3 months", expand=c(0,0)) +
  scale_y_continuous(breaks = c(0, 250, 500, 750, 1000, 2000, 3000)) +
  scale_fill_manual(values = alpha(c("#FF9900", "#85bb65"), 0.2)) +
  ggtitle("Bitcoin Daily Trading Volume",
    subtitle = "in number of bitcoin & Thousands of USD") +
  labs(caption = "Data source: Coinbase API",
    x = "",
    y = "Volume") +
  theme(panel.grid.major.y = element_line( size=.1, color="#999999"),
    panel.grid.major.x = element_blank(),
    panel.background = element_blank(),
    plot.title = element_text(size=18,
      family = "Helvetica",
      colour = "#3a3a3a",
      face = "bold"),
    plot.subtitle = element_text(size=12, family = "mono", colour = "#666666"),
    axis.text.y = element_text(color="#666666"),
    axis.title.y = element_text(color="#666666"),
    axis.ticks.y = element_blank(),
    axis.ticks.x = element_blank(),
    axis.text.x = element_text(size=10, family = "mono", colour = "#666666"),
    legend.title = element_blank(),
    legend.position = c(0.1, 0.9),
    plot.caption = element_text(size=11, family = "mono", colour = "#666666"),
    plot.margin = unit(c(2,2,2,2), "cm"))
```

area

Bitcoin Daily Trading Volume

in number of bitcoin & Thousands of USD



## Graph Four



## Graph Five