wp_ico_analysis

Michael Rosenberg and Michael McCaffrey

11/10/2018

```
setwd(dir = "~/Dropbox/crypto_analyses/")
wp_ico_frame <- read.csv("wp_ico.csv",header = TRUE)
```

Data Sanity Checks

```
dim(wp_ico_frame)
```

```
## [1] 475   5
```

```
#more lines than expected; delete statistics rows
colnames(wp_ico_frame)
```

```
## [1] "Project"                      "Amount.Raised.in.ICO...M."
## [3] "ICO.Close.Date"               "Page.count"
## [5] "Amount.Raised.per.Page...M."
```

```
shit_row_start <- 440
shit_row_vec <- seq(from = shit_row_start,to = dim(wp_ico_frame)[1],
                    by = 1)
wp_ico_frame <- wp_ico_frame[-shit_row_vec,]
dim(wp_ico_frame)
```

```
## [1] 439   5
```

EDA

```
#cleanup from dollar to numeric
wp_ico_frame$amount_raised_m <- as.numeric(gsub('[$,]', '',
                                          wp_ico_frame$Amount.Raised.in.ICO...M.))
hist(wp_ico_frame$amount_raised_m,
     main = "Distribution of Amount Raised ($M)",
     ylab = "Count",
     xlab = "Amount Raised ($M)",
     col = "blue")
```

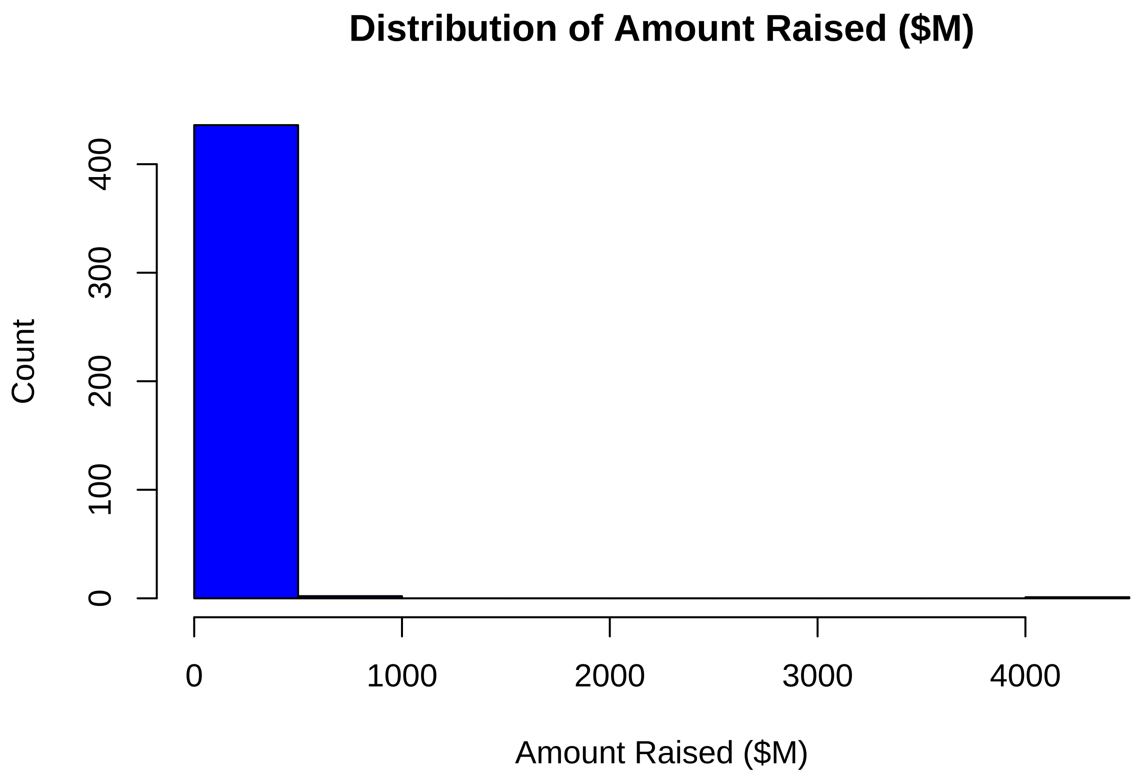

Figure 1: Distribution of Amount Raised (In Millions).

Looks very right skewed. Let us log it.

```
wp_ico_frame$log_amt_raised_m <- log(wp_ico_frame$amount_raised_m)
hist(wp_ico_frame$log_amt_raised_m,
     main = "Distribution of Log-Amount Raised",
     ylab = "Count",
     xlab = "Log-Amount Raised",
     col = "blue")
```

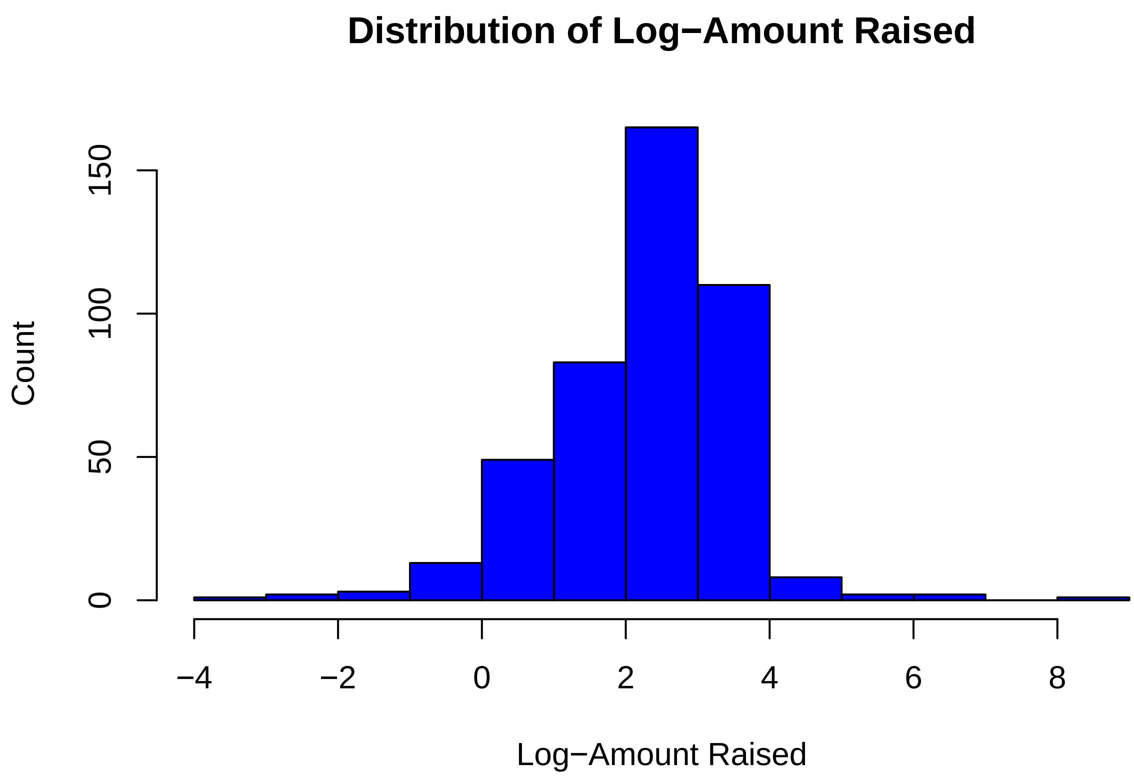

Figure 2: Distribution of log-*Ammount Raised.*

```r
wp_ico_frame$Page.count <- as.numeric(wp_ico_frame$Page.count)
hist(wp_ico_frame$Page.count,
     main = "Distribution of Page Count",
     ylab = "Frequency",
     xlab = "Page Count",
     col = "blue")
```

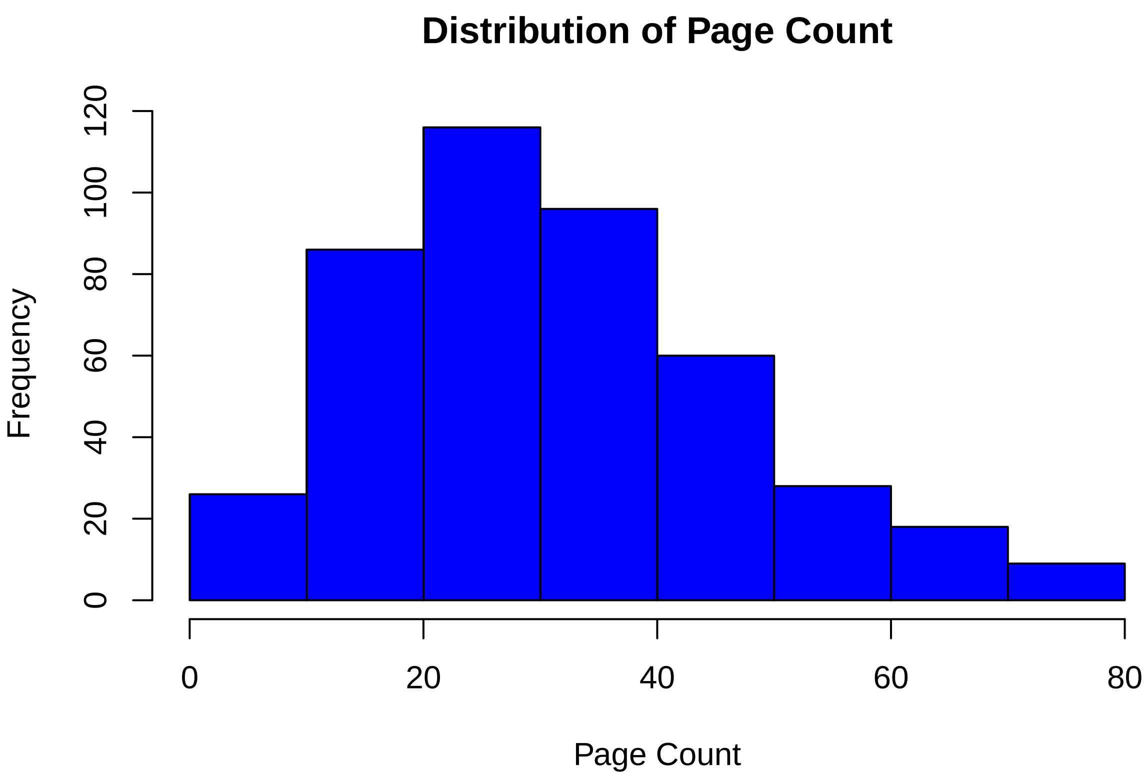

Figure 3: Distribution of Page Count.

```
plot(x = wp_ico_frame$Page.count,
     y = wp_ico_frame$log_amt_raised_m,
     pch = pch_lev,
     main = "Log-Amount Raised on Page Count",
     ylab = "Log-Amount Raised (M)",
     xlab = "Page Count")
```

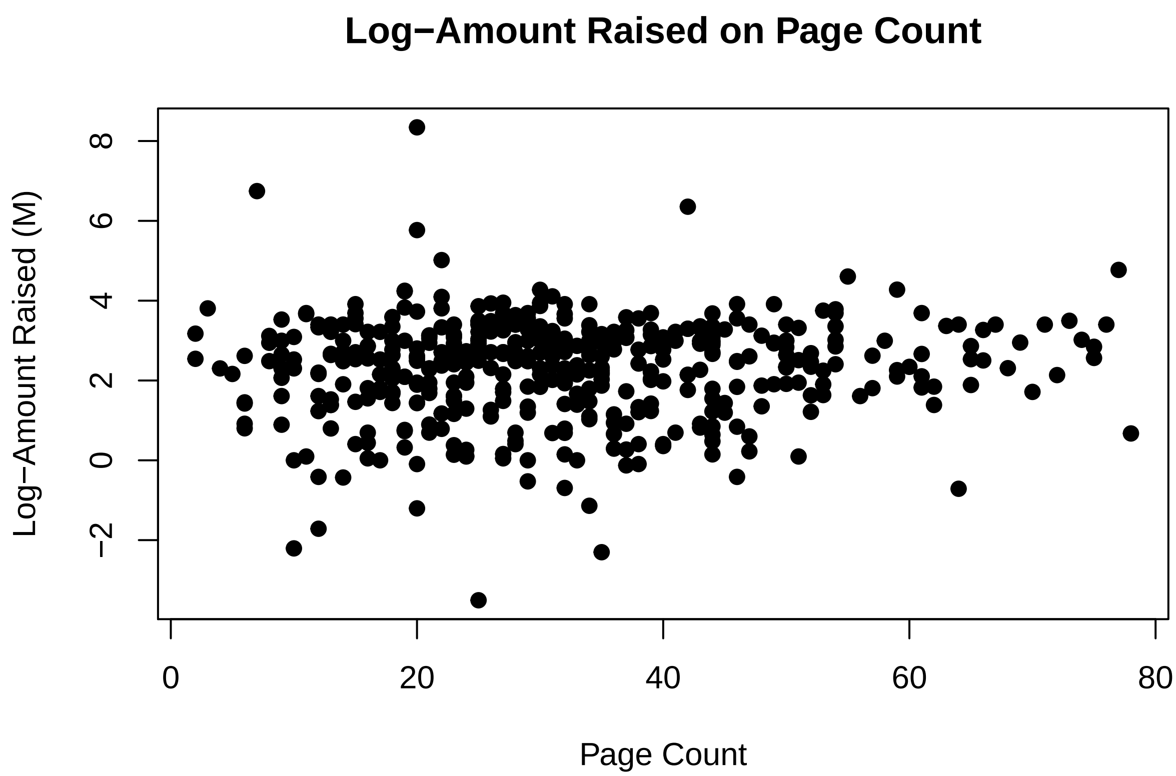

Figure 4: log-*Amount Raised on Page Count.*

Initial Modeling

```
init_mod_lm <- lm(log_amt_raised_m ~ Page.count,data = wp_ico_frame)
summary(init_mod_lm)
```

```
##
## Call:
## lm(formula = log_amt_raised_m ~ Page.count, data = wp_ico_frame)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.7656 -0.6371  0.2164  0.8033  6.1101
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.127792   0.137356  15.491   <2e-16 ***
## Page.count  0.005250   0.003877   1.354    0.176
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.265 on 437 degrees of freedom
## Multiple R-squared:  0.004177,   Adjusted R-squared:  0.001898
## F-statistic: 1.833 on 1 and 437 DF,  p-value: 0.1765
```

```
page_count_coef <- coefficients(init_mod_lm)["Page.count"]
#log(var) ~ page_count -> var ~ exp(page_count)
page_count_mul = exp(page_count_coef)
```

```r
amount_raised_increase <- (page_count_mul - 1) * percent_lev
amount_raised_increase
```

```
## Page.count
##   0.526341
```

This suggests to me that there isn't a strong global effect of page count on amount.

```r
page_cutoff_of_interest <- 55
wp_ico_frame$above_cutoff <- wp_ico_frame$Page.count >= page_cutoff_of_interest
```

```r
cutoff_mod_lm <- lm(log_amt_raised_m ~ above_cutoff,data = wp_ico_frame)
summary(cutoff_mod_lm)
```

```
##
## Call:
## lm(formula = log_amt_raised_m ~ above_cutoff, data = wp_ico_frame)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.7733 -0.7171  0.2181  0.7924  6.0761
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       2.26677    0.06298  35.990   <2e-16 ***
## above_cutoffTRUE  0.34205    0.21994   1.555    0.121
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.264 on 437 degrees of freedom
## Multiple R-squared:  0.005504,   Adjusted R-squared:  0.003228
## F-statistic: 2.419 on 1 and 437 DF,  p-value: 0.1206
```

```r
cutoff_coef <- coefficients(cutoff_mod_lm)["above_cutoffTRUE"]
cutoff_impact_mul <- exp(cutoff_coef)
impact_percent_increase = (cutoff_impact_mul - 1) * percent_lev
impact_percent_increase
```

```
## above_cutoffTRUE
##         40.78357
```

While the cutoff effect is not statistically significant, we can see by the percent effect that it implies a lot of financial significance. This suggests that the base effect is very strong, we just don't have enough data to claim statistical significance in this context.

Robust to Time

We're wondering if there are particular time effects that are confounding this process. In particular, it might just be the case that all the heavy whitepapers were released early in the year, which would suggest that the burndown on the cryptocurrency craze is really inform the ICO-generating process. Let's plot that.

```r
wp_ico_frame$month_of_close <- month(
                                   as.POSIXct(wp_ico_frame$ICO.Close.Date,
                                              format = "%m/%d/%y"))
```

```r
#get color vector
col_set <- colorRampPalette(c("blue","red"))
num_months <- length(unique(wp_ico_frame$month_of_close))
col_map_vec <- col_set(num_months)
col_application <- function(month,col_map_vec){
    return(col_map_vec[month])
}
col_application(1,col_map_vec)
```

```
## [1] "#0000FF"
```

```r
col_row_vec <- sapply(wp_ico_frame$month_of_close,
                      col_application,
                      col_map_vec = col_map_vec)
#then plot
plot(x = wp_ico_frame$Page.count,
     y = wp_ico_frame$log_amt_raised_m,
     pch = pch_lev,
     col = col_row_vec,
     main = "Log-Amount Raised on Page Count\n(Conditioned on Month)",
     ylab = "Log-Amount Raised (M)",
     xlab = "Page Count")
legend("topright",legend = unique(wp_ico_frame$month_of_close),
       col = col_map_vec,title = "Month")
```

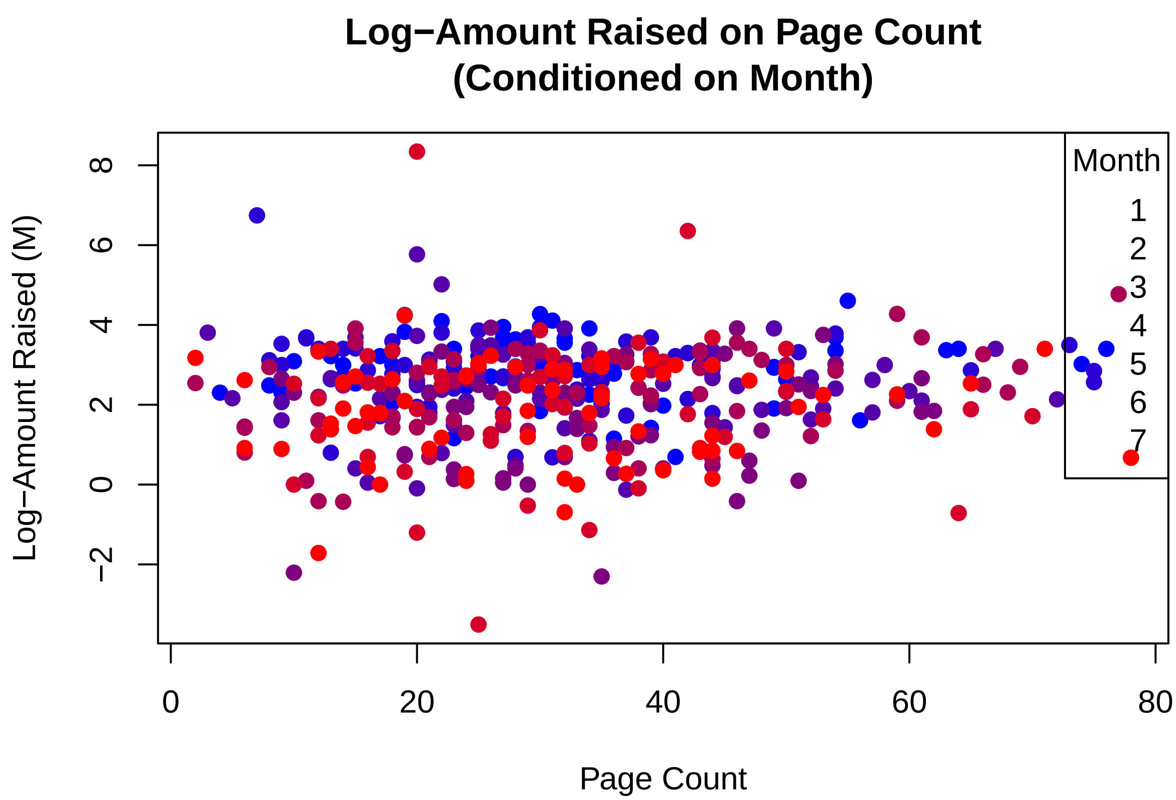

Figure 5: log-Amount Raised ($M) on page count, conditioned on month. earlier months are colored in blue, while later months are colored in red.

```r
month_conditoned_lm <- lm(log_amt_raised_m ~ above_cutoff + as.factor(month_of_close),
                          data = wp_ico_frame)
```

```r
summary(month_conditoned_lm)
```

```
##
## Call:
## lm(formula = log_amt_raised_m ~ above_cutoff + as.factor(month_of_close),
##     data = wp_ico_frame)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.5403 -0.6548  0.1123  0.7602  6.3091
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 2.88134    0.13835  20.827  < 2e-16 ***
## above_cutoffTRUE            0.29425    0.20803   1.414   0.1579
## as.factor(month_of_close)2 -0.04393    0.20799  -0.211   0.8328
## as.factor(month_of_close)3 -0.53860    0.20615  -2.613   0.0093 **
## as.factor(month_of_close)4 -1.18367    0.20087  -5.893 7.67e-09 ***
## as.factor(month_of_close)5 -0.59097    0.20535  -2.878   0.0042 **
## as.factor(month_of_close)6 -0.84755    0.21234  -3.992 7.71e-05 ***
## as.factor(month_of_close)7 -1.12722    0.20156  -5.592 3.98e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.189 on 431 degrees of freedom
## Multiple R-squared:  0.1321, Adjusted R-squared:  0.118
## F-statistic: 9.368 on 7 and 431 DF,  p-value: 7.961e-11
```