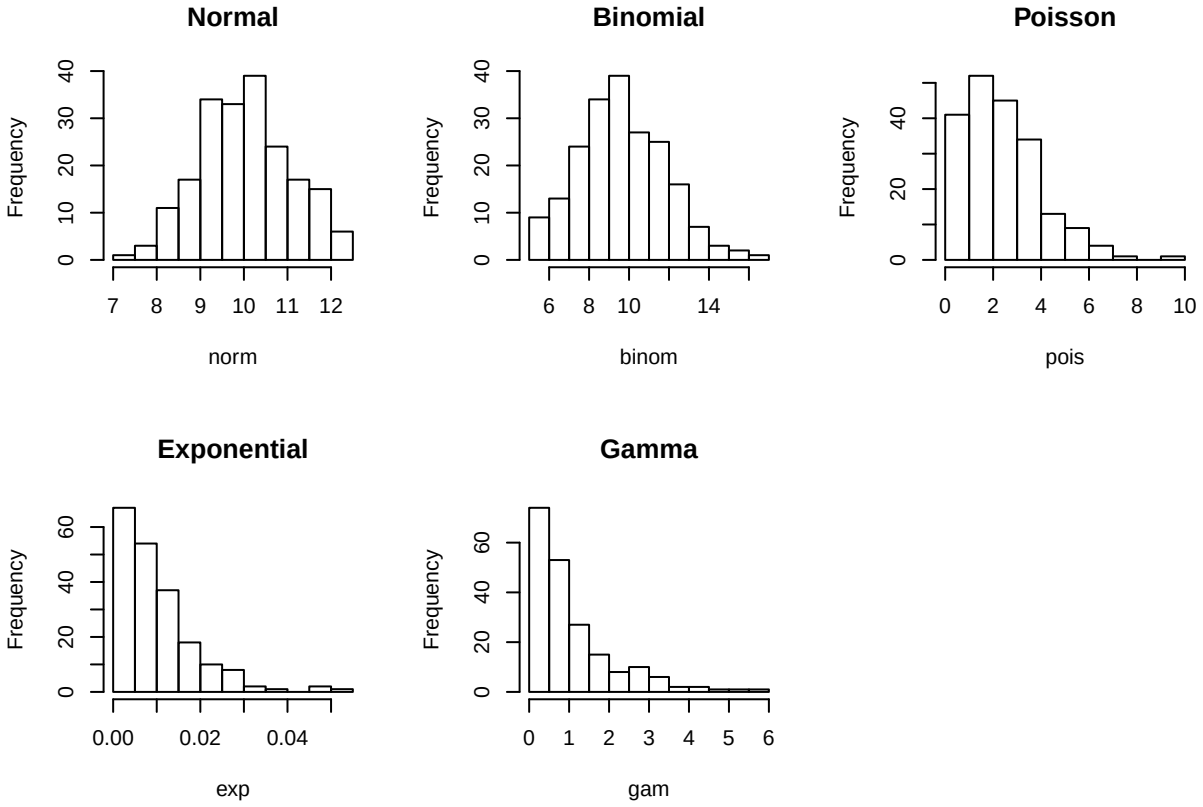
Data Distributions

Matt Chapman

Known probability distributions

```
norm <- rnorm(n = 200, mean = 10, sd = 1)
binom <- rbinom(n = 200, size = 20, prob = 0.5)
pois <- rpois(n = 200, lambda = 3)
exp <- rexp(n = 200, rate = 100)
gam <- rgamma(n = 200, rate = 1, scale = 1, shape = 1)

par(mfrow = c(2,3))
hist(norm, main="Normal")
hist(binom, main="Binomial")
hist(pois, main="Poisson")
hist(exp, main="Exponential")
hist(gam, main="Gamma")
```

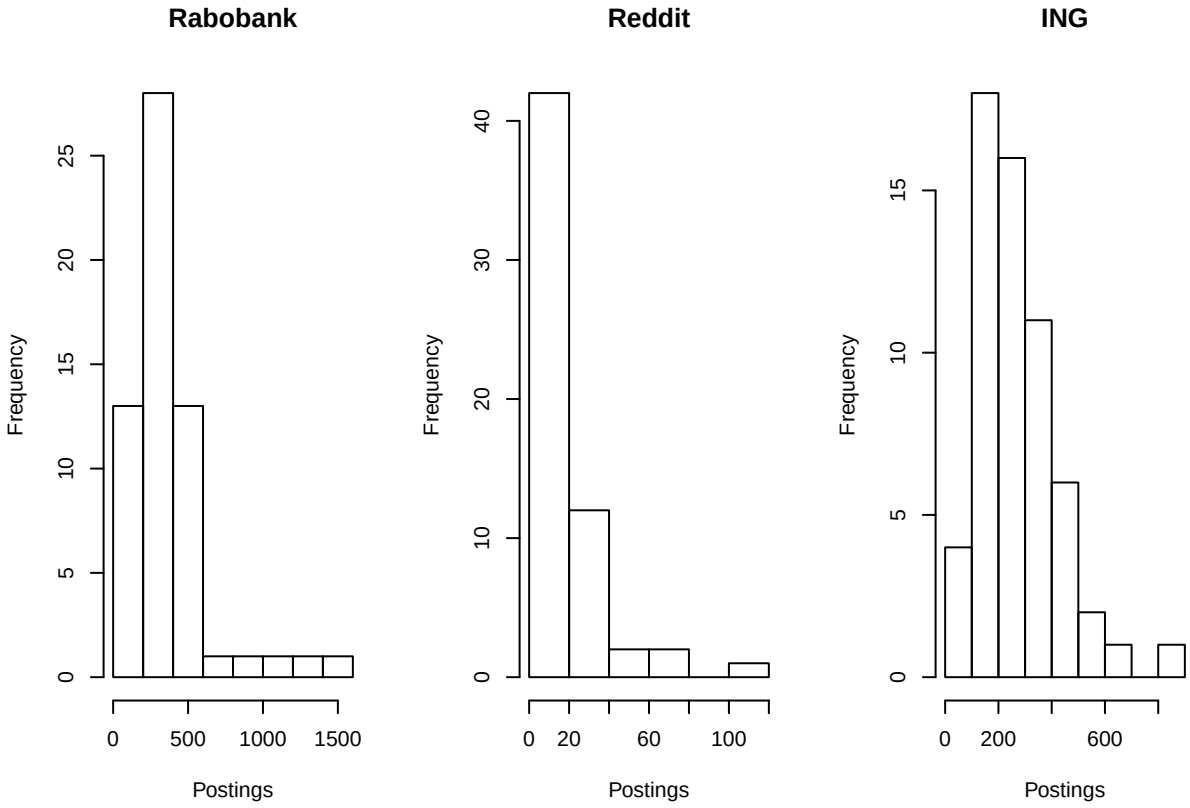


Histogram plots of rabobank, reddit, ing

```
source('~/Repos/master-thesis/Experiments/R/ChangeDetection/experiment.R')
par(mfrow = c(1,3))
```

```r
rabodata <- ProcessData(rabobank)
redditdata <- ProcessData(reddit)
ingdata <- ProcessData(ing)

hist(rabodata$Freq, main="Rabobank", xlab = "Postings")
hist(redditdata$Freq, main="Reddit", xlab = "Postings")
hist(ingdata$Freq, main="ING", xlab = "Postings")
```



Example Experiment Output