On the Origin of Citations

Rahul Subramanian

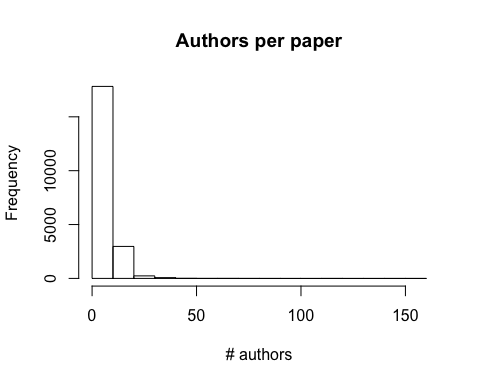
September 15, 2015

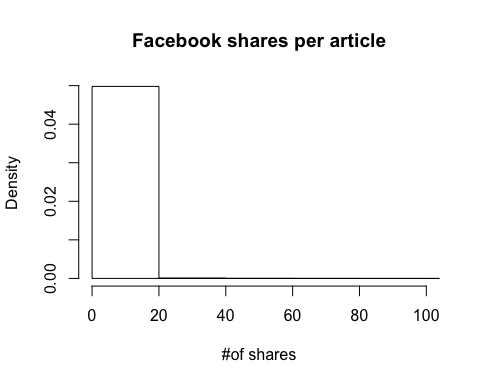
# Load the data

## Using read.delim

counts\_raw <- read.delim("data/counts-raw.txt.gz")  
counts\_norm <- read.delim("data/counts-norm.txt.gz")

# Data exploration

What's the distribution of all articles of our data set? 



The average number of facebook shares per article is 0.6524187 .