On the origin of citations

Silvia Kariuki

September 15, 2015

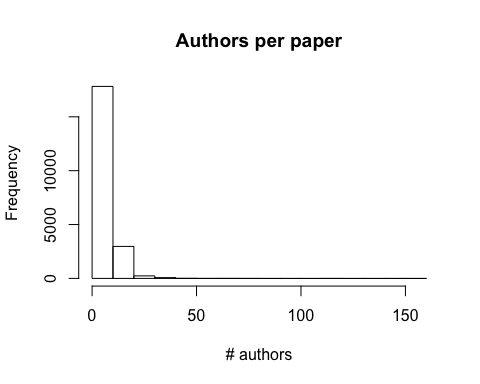
# Load the data

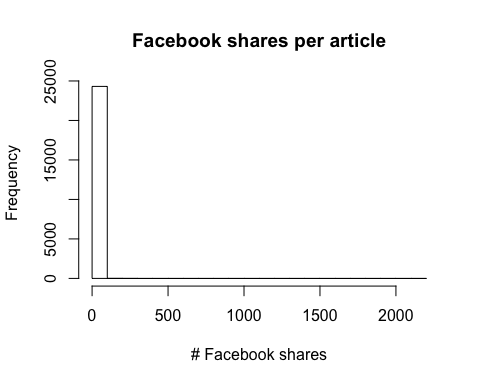
## using read.delim

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 authors in all articles of our data set?





The average number of Facebook shares per paper in the data set is 0.6524187