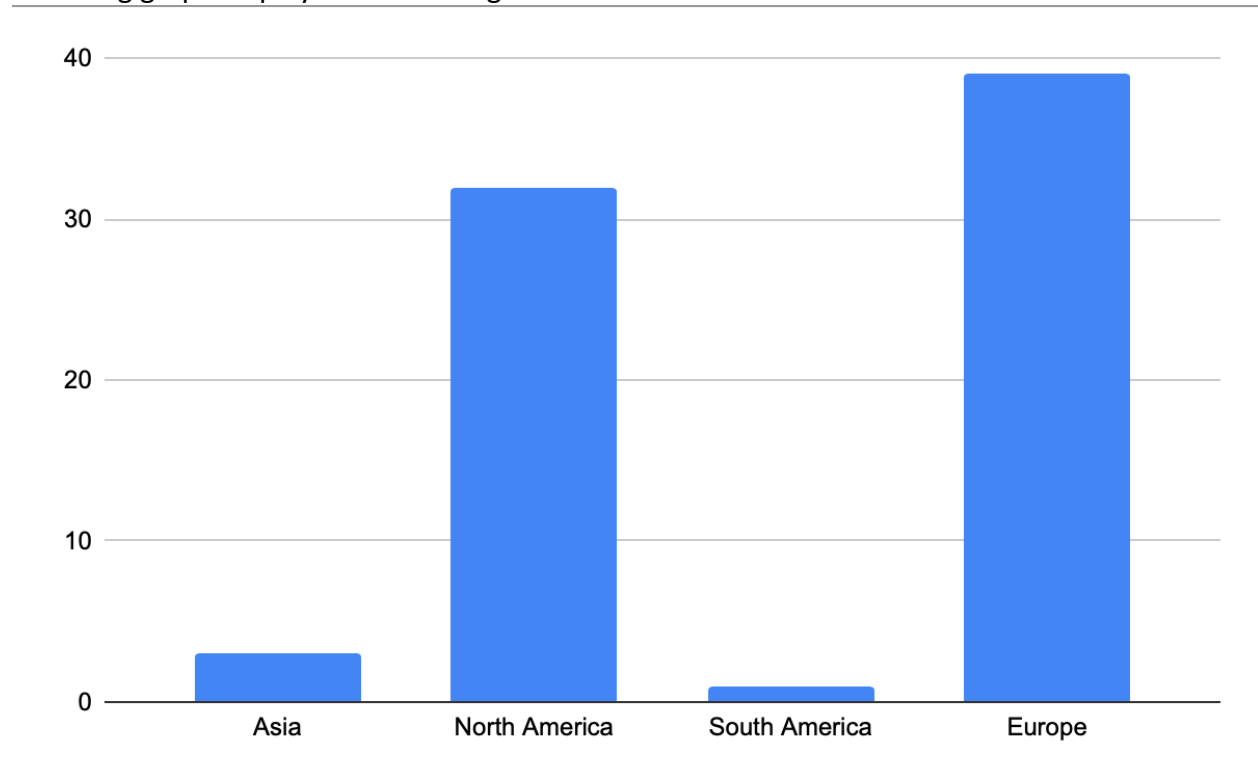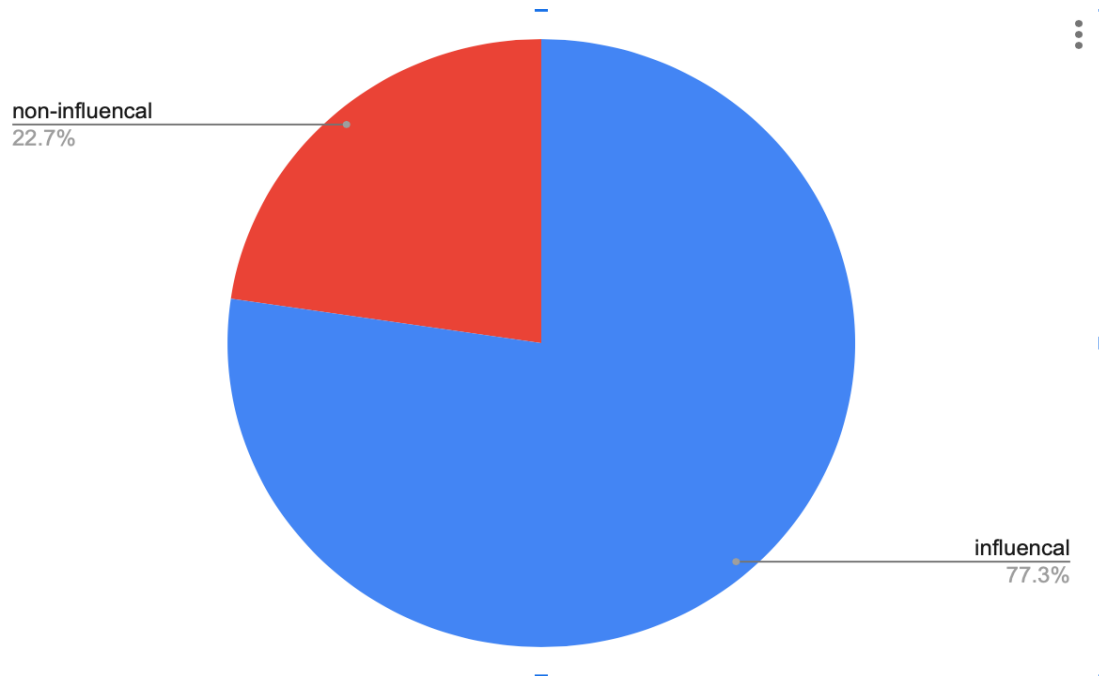To complete task 3, I wrote a variety of function to collect various data about the data set. The first thing I collected was the total number of data points in the set. I found that to be 75. I then wrote a function to find the distribution of nodes amongst the continents. The function worked by running through each JSONObject and keeping a tally for each continent. The following graph displays these findings:

I also calculated the total amount, and consequently the percentages, of nodes for various Boolean values. Firstly, the number of nodes from the data set which were influential. The function worked by keeping a running tally of influential and then subtracting that tally from the total number in the set to get the non-influential. The following data displays these findings.

non-influencal
22.7%

influencal
77.3%

The other Boolean value which I investigated via programming was originality. The following pie graph displays those findings The function for this worked very similarly to the one for influential vs non-influential.

non-original
44.0%

original
56.0%