Mandatory assignment #3 - Homophily

Nodes: 1490

Type 1(p) nodes: 732

Type 0 (q) nodes: 758

1-0 edges: 1.683

We first calculate the probability that a randomly picked edge is a “1-0”.

The total number of edges in the graph is 19.025. 0.498% of that is 9.474.

But in reality there are only 1.683 1-0 edges. This means that there must be some homophily present, that connects 1s more to other 1s and 0s to other 0s rather than 1s to 0s.

That homophily could be a difference in political opinion, since the 0s are liberals, and 1s are conservatives. The type of homophily present in this data is Value Homophily, as the context here revolves around politics and the blog(ger)’s oppinion. Based on the information presented to us in the *polblogs.txt* file, some blogs were manually labeled based on their posts and outgoing links, others were automatically fetched from blog directories. As stated in the document this creates quite a possibility for error in the automatic labels and especially the manual labels. This means that the number of nodes in both *p* and *q* might not necessarily be accurate.