Task 1

Since our graph is not a connected graph, we could not compute the path length.

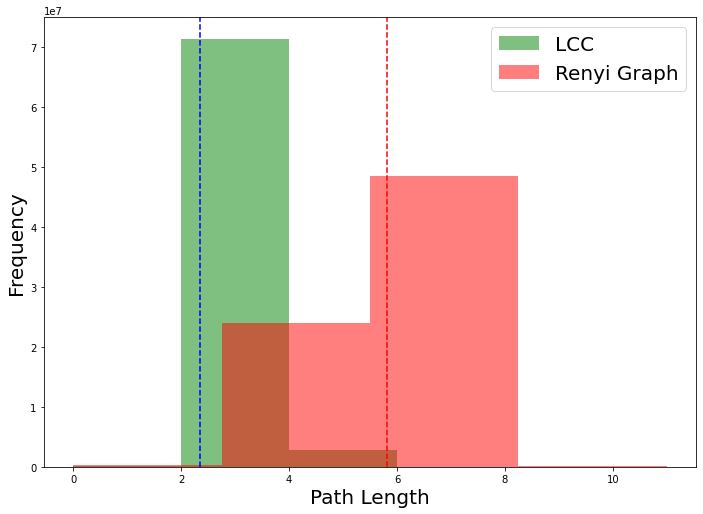
Therefore we processed (proceeded) by extracting the largest connected components from both graph Glastonbury & Erdos Renyi.

The largest connected components were extracted in the form of list of nodes, thus we constructed a list of edges of all possible combinations of pair of nodes in that list.

After sorting the list of edges, we checked the existence of the edge in our graph before affecting it to a dictionary or a list of tuples in order to construct a new graph (largestcc\_G, largestcc\_rnd)

Once our new graph created, we computed the path length as well as the clustering coefficient using Networks library.

**Findings:**

In the illustrated figure, we can notice that x-axis of the graph shows the shortest path length between every two nodes and Y-axis represent the frequency of those shortest path length ,the number of links between any two hashtags varies between two to four interactions. Thus, 

For the Glatonbury network with x nodes and x edges, the average path length is ***l = 2.34*** and the clustering coefficient is ***C = 0.55***, thus nearly all hashtags are linked within a short number of steps to all other hashtags and are densely connected, while the average path length and the clustering coefficient of a random graph of the same size and average degree are ***lrand* = 5.82** and ***Crand* = 4 × 10−**3, respectively, thus nearly all hashtags are linked within a big number of steps to all other hashtags and are sparsely connected.

**Glastonburry** shows that the clustering parameter is much larger than that of a random network while the average path length is smaller, in this case does our **Glastonburry** have the characteristic of small word phenomenon?