Qualitative Analysis:

* Page rank is a measure of determining the importance of a web page. Page rank is a measure based on a random surfer, where a person randomly clicking on links will arrive at a particular page.
* There are two variations by which the top 10 links are generated, one, by the number of in links and another by the Page rank algorithm.

The ranking based on the in link count does not take into the consideration the quality of a page. This proves to be inaccurate because of spam links. The page rank computed based on in link count also does not take into consideration the sink links or links that may form a loop.

* For **G1**, the link ‘**International\_Standard\_Book\_Number**’ is ranked number 1 by both the Page rank algorithm and by the count of number of in links. This is because this page does not contain any sink links within the crawled links and in link method proves to be efficient.
* Contrasting this for the link ‘**Tropical\_cyclone**’, it is ranked as second by the in link count but does not come in the top 10 computed by the page rank algorithm. This is because this page consists of 8 **sink links** among the 1000 links crawled which are the follows:

1. Geometric\_terms\_of\_location - Sink link of tc
2. Tropical\_Depression\_(EP) - SL
3. Tropical\_Storm\_(album) - SL
4. One\_If\_by\_Clam,\_Two\_If\_by\_Sea
5. The\_Checks
6. Allen\_Press
7. Dictionary.com
8. Wayback\_Machine

These sink links won’t be considered while calculating the page rank using in link count but simply ranked because it consists of 355 in links. But sink links have a considerable affect while computing the Page rank which is computed by the Page Rank algorithm.

* The remaining links follow the same rule and the above explains the difference in the page ranking mechanism.
* The difference in the ranking by Page Rank algorithm is influenced by both the number of in links and the sink nodes as follows;

The link ‘International\_Standard\_Book\_Number’ in G1 is ranked first while ‘Wayback\_Machine’ is ranked tenth by the page ranking algorithm. Both the links does not contain any sink links among the links crawled.

But here the ranking is affected by the in links, i.e. The number of inlinks to ‘International\_Standard\_Book\_Number’ is 433 while for ‘Wayback\_Machine’ is 244. Hence both the sink links and in links are taken into consideration in the Page Rank algorithm.

* Similarly, for graph G2, the link ‘Docklands\_Light\_Railway’ is ranked 5 by the count of in links but ranked 16 by PageRank algorithm. This is because the page contains one sink link, ‘Transport\_and\_Works\_Act\_1992’ which makes it to be ranked in a moderate position of 13 by PageRank algorithm.
* In order to calculate the page rank accurately, the user search logs have to be considered, and also user context, interest, current events and other factors are to be considered for pagerank in order to retrieve a relevant webpage based on the search for the user.

