

**1) A list of the PageRank values obtained for each of the six vertices after 1, 10, and 100 iterations of the PageRank algorithm.**

Iteration 1 Results

Page rank of A is 0.2493055555555556  
Page rank of B is 0.1194444444444445  
Page rank of C is 0.1430555555555555  
Page rank of D is 0.13125  
Page rank of E is 0.213888888888889  
Page rank of F is 0.1430555555555555

Iteration 10 Results

Page rank of A is 0.2526524506150375  
Page rank of B is 0.1395550023626756  
Page rank of C is 0.151559179065915  
Page rank of D is 0.11906808910508244  
Page rank of E is 0.18734019939499408  
Page rank of F is 0.15154226794449255

Iteration 100 Results

Page rank of A is 0.2521271053751956  
Page rank of B is 0.1393061853185386  
Page rank of C is 0.1513064898667053  
Page rank of D is 0.11890782257353921  
Page rank of E is 0.18704590699931614  
Page rank of F is 0.1513064898667053

**2) The perplexity values obtained in each round until convergence for the in-links file for the WT2g**

ITERATION	PERPLEXITY VALUE
1	183810.999998184
2	80012.1601542053
3	86873.8575180254
4	72745.8322851183
5	75797.2689789137
6	69517.1536999668
7	71873.6652750149
8	68400.1608000733
9	70038.4246820118
10	68007.4902790218
11	69122.7318130402
12	67830.8608757711
13	68640.1974358948
14	67760.8021685689
15	68338.4092342628
16	67751.9522873413

17	68150.148869246
18	67751.252918314
19	68036.2805226325
20	67756.9846839845
21	67960.3544831705
22	67767.2140785136
23	67909.8728765902
24	67775.7548668072
25	67877.2373314033
26	67782.672490442
27	67855.1255438438
28	67788.712781962
29	67839.9928497459
30	67793.3846573715
31	67829.8509186332
32	67796.8717734973
33	67822.9143175032
34	67799.5822276017
35	67818.0969294241
36	67801.6111188303
37	67814.7991006918
38	67803.0965080022
39	67812.5136683289
40	67804.2043841914
41	67810.918701384
42	67805.0182905834
43	67809.8087648412
44	67805.6116069654
45	67809.0336376045
46	67806.0447655055
47	67808.4887410736
48	67806.3605277403
49	67808.1067700714
50	67806.5895895025
51	67807.8380222721
52	67806.7560360371
53	67807.6486675629
54	67806.8764662261
55	67807.5150567446
56	67806.9637083967

### 3.a) Top 50 document IDs sorted by In-links

WT21-B37-76	0.00269447441554888
WT21-B37-75	0.00153317909999839
WT25-B39-116	0.00146850141632585
WT23-B21-53	0.00137352493842574
WT24-B26-10	0.00127615674413019
WT24-B40-171	0.00124529321760879
WT23-B39-340	0.00124291315832404
WT23-B37-134	0.00120542141122597
WT08-B18-400	0.00114477728783794
WT13-B06-284	0.00113655800531197
WT13-B06-273	0.0010549245722372
WT01-B18-225	0.000955381964680431
WT04-B27-720	0.000940964388087425
WT24-B26-46	0.000862238220087864
WT23-B19-156	0.000825070709726481
WT04-B30-12	0.000816623256441777
WT25-B15-307	0.000797223535250271
WT07-B18-256	0.00077503741736185
WT24-B40-167	0.000707628733677399
WT14-B03-220	0.000698885966546419
WT18-B31-240	0.000694225761176821
WT14-B03-227	0.00068531619027251
WT04-B40-202	0.000684680858729142
WT08-B19-222	0.000649532567950543
WT23-B20-363	0.000639636012216146
WT27-B28-203	0.000627079718836486
WT13-B39-295	0.000621537510912944
WT13-B15-160	0.000619858571957634
WT12-B30-56	0.0006024553556204
WT10-B02-288	0.000584449403785491
WT14-B36-337	0.000561707322502672
WT21-B40-37	0.000547187988809291
WT21-B35-155	0.000545696979496517
WT23-B01-40	0.000540852129280064
WT08-B08-60	0.000536122308950785
WT22-B38-403	0.000533578786250681
WT13-B39-321	0.000532895454818477
WT04-B22-268	0.000532864118235418
WT14-B02-400	0.000532803706142627
WT18-B14-66	0.000529559574014968
WT06-B14-69	0.000519161976115176
WT23-B38-120	0.000518694140983382
WT06-B35-151	0.000516928125528944
WT10-B33-300	0.000516773966596665
WT14-B36-323	0.000515861862272522

WT14-B36-335	0.000515861862272522
WT14-B36-336	0.000515861862272522
WT14-B36-334	0.000515861862272522
WT06-B35-161	0.000515022164175948
WT27-B20-494	0.000514480870451751

### 3-b) Document IDs sorted by In-links

WT21-B37-76	2568
WT21-B37-75	1704
WT01-B18-225	1137
WT08-B19-222	1041
WT08-B18-400	990
WT21-B40-447	779
WT27-B34-57	630
WT27-B32-30	628
WT25-B15-307	605
WT27-B28-203	589
WT18-B40-82	576
WT21-B37-71	560
WT13-B15-160	484
WT23-B30-88	477
WT27-B28-177	461
WT13-B06-284	454
WT13-B06-273	452
WT13-B39-295	443
WT14-B36-337	417
WT10-B33-300	406
WT23-B23-51	400
WT23-B27-29	388
WT23-B30-105	379
WT14-B36-323	368
WT14-B36-335	368
WT14-B36-336	368
WT14-B36-334	368
WT23-B30-89	367
WT23-B19-156	364
WT12-B40-248	357
WT17-B34-504	356
WT17-B34-503	356
WT17-B34-506	356
WT17-B34-505	356
WT17-B34-500	356
WT17-B34-502	356
WT17-B34-501	356
WT17-B34-508	356

WT17-B34-507	356
WT17-B34-509	356
WT17-B34-498	356
WT17-B34-499	356
WT12-B40-235	355
WT12-B40-241	342
WT12-B40-254	334
WT04-B40-202	322
WT24-B26-2	320
WT04-B40-238	311

**3 c) Proportion of pages with no in-links 0.07806932120493333**

**3 d) Proportion of pages with no outlinks 0.3600165387272797**

**3 e) Proportion of pages whose PageRank is less than their initial, uniform values 0.7851434353765553**

#### **4) Analysis for the top 10 Document ids by Page rank in Lemur web interface.**

Sl no	Documents id	Page Rank	In-link count	Analysis
1	WT21-B37-76	0.00269447441554888	2568	<p><b>The Economist homepage</b>  This page is the homepage of The Economist magazine. It has the highest in-link count of 2568.  Since this is the homepage of a popular magazine, there are many in links from other web pages including important web pages with high page rank values.  This webpage is more likely to be the one which the user wants to see in response to a query.</p>
2	WT21-B37-75	0.00153317909999839	1704	<p><b>The Economist: Copyright Notice Page.</b>  This is the Copyright Notice Page of The Economist.  It has link from Economist home page which has the highest page rank. In addition, it has links from all the other webpages in The Economist .So, for this reason, Copyright notice page has a higher page rank with in-links of high quality and rank. This web page might not be an important result to user's search query. This is because, a user querying 'The Economist' in most cases does not mean the Copyright notice page. The might want to navigate to home page of The Economist and/or other articles in the Economist.</p>
3	WT25-B39-116	0.00146850141632585	169	<p><b>Security Assurance Requirements</b>  There are 169 in-links to this page.  Of all the pages that contribute to this webpage, WT25-B38-117 and WT25-B39-116 and has 4 and 169 in-links respectively. All other documents have 0 in-links. WT25-B38-117 is the 'Introduction and General Model' i.e. Part 1 and WT25-B39-116 is</p>

				the same document itself. (Part 3). The majority of the page rank contribution comes from the same page.
4	WT23-B21-53	0.00137352493842574	198	<b>Web Development Team</b> This page contains information about the web development team for <a href="http://www.sea.usscouts.org/">http://www.sea.usscouts.org/</a> This page includes links to the various members of the web development team who are also the members of sea exploring committee. So a user might want to go to this page to gather information about the team and also contact them when they have queries. In addition, the individual team member page has a link back to this page. All these contributed to its page rank.
5	WT24-B26-10	0.00127615674413019	291	<b>Information of Psychiatry Star project</b> <b>What is Psychiatry Star?</b> This page has links to Program Director and other people in this project. They have many team members and each of them have link back to this page. Anyone who visits the team member web page might be interested in Psychiatry Star and want to visit this page for more information. Also, the team member pages have a high page rank value. A user is likely to visit this page.
6	WT24-B40-171	0.00124529321760879	270	<b>The Evening News</b> on the Internet every weeknight This webpage has links to many news articles. Also news is a most common query term. A user is likely to visit this page.
7	WT23-B39-340	0.00124291315832404	274	<b>STREETLINK FINANCIAL REPORTS</b> This web page has financial reports of various companies and the page rank is high since this web page has link back from all the company reports. A user is likely to visit this page.
8	WT23-B37-134	0.00120542141122597	207	<b>Disclaimer and Copyright information of <a href="http://www.public.health.wa.gov.au">http://www.public.health.wa.gov.au</a>.</b> All the academic articles in this webpage have a disclaimer information. In addition, home page has a link to this page. Since this website is the Western Australian Government's public health web page, many people visit this website. But disclaimer and copyright information page may not be important for a user.
9	WT08-B18-400	0.00114477728783794	990	<b>General Disclaimer of tdbank.com</b> This page does not have any out-link and hence is a sink node. This page has a high page rank since all documents in this website would include disclaimer. This page may not be important for a user.
10	WT13-B06-284	0.00113655800531197	454	<b>Web Development page of <a href="http://www.lawlead.com">www.lawlead.com</a></b> This web page has just two out links. All the pages in the website has link to webmaster (credits) information of the given page. This page might not be important for the user. This page has a high page rank since all the pages in the website have link to the credits information.

**Summary:** There are few pages like home page and other articles which a user is likely to visit. But there are other pages like copyright information/Disclaimer, which has a higher page rank because all the webpages in the particular website has a link to these. Also some pages with high in-link count might have high Page rank value since some of the in-link pages have high page rank value.

According to the analysis,

**1) The webpages which are more likely to be the one which the user wants to see in response to a query:**

The Economist homepage, Security Assurance Requirements, What is Psychiatry Star? , The Evening News, STREETLINK FINANCIAL REPORTS

**2) The webpages which are likely to be the one which the user wants to see in response to a query:**

Web Development Team of <http://www.sea.usscouts.org/> and Web Development page of [www.lawlead.com](http://www.lawlead.com)

**3) Document that are least interesting to the user**

The Economist: Copyright Notice Page.

General Disclaimer of [tdbank.com](http://tdbank.com)

Disclaimer and Copyright information of <http://www.public.health.wa.gov.au>.

**5) Analysis for the top 10 Document ids by in-links in Lemur web interface.**

Si no	Document id	In link count	Page rank	Analysis
1	WT21-B37-76	2568	0.00269447441554888	<p><b>The Economist homepage</b></p> <p>This page is the homepage of The Economist magazine. It has the highest in-link count of 2568.</p> <p>Since this is the homepage of a popular magazine, there are many in links from other web pages including important web pages with high page rank values.</p> <p>This webpage is more likely to be the one which the user wants to see in response to a query.</p>
2	WT21-B37-75	1704	0.00153317909999839	<p><b>The Economist : Copyright Notice Page.</b></p> <p>This is the Copyright Notice Page of The Economist.</p> <p>It has link from Economist home page which has the highest page rank. In addition, it has links from all the other webpages in The Economist .So, for this reason, Copyright notice page has a higher page rank with in-links of high quality and rank. Even though, this web page might be important when considering The Economist, it might not be an important result to user's search query. This is because, a user querying 'economist' in most cases does not mean the Copyright notice page. The might want to navigate to home page of The Economist and other articles in the Economist.</p>

3	WT01-B18-225	1137	0.000955381964680431	Search page for <b>DRCNet Online Library of Drug Policy</b> . All the documents in the library has link to search page. Hence this page has high in link count. But this might be not important for a user query.
4	WT08-B19-222	1041	0.000649532567950543	<b>Terms and conditions in <a href="http://www.tdbank.ca">http://www.tdbank.ca</a></b> Many pages in the website has link to T&C including Q&A pages. Hence the in-link count is high. But this might be not important for a user query.
5	WT08-B18-400	990	0.00114477728783794	<b>General disclaimer for <a href="http://www.tdbank.ca">http://www.tdbank.ca</a>.</b> Many pages have link to Disclaimer. Hence the number of in-links is high. But this might not be important for a user query.
6	WT21-B40-447	779	0.000342636530268353	Archive page of <a href="http://www.economist.com">http://www.economist.com</a> . There are links from other pages like biographies to this page. This page might be important for a user query.
7	WT27-B34-57	630	0.000506413885037798	<b>Skeptics Society Message Board</b> has links from all feedback/responses in the message board. Hence have a high in-link count. But this might not be important for a user query.
8	WT27-B32-30	628	0.000500680146135429	<b>WWWBoard Frequently Asked Questions and Answers</b> FAQ link is there from all the feedback/responses in the message board. Hence it has a high in-link count.
9	WT25-B15-307	605	0.000797223535250271	<b>Intelligent information- search page.</b> Search page has links from all the articles. Hence the in-link count is high.
10	WT27-B28-203	589	0.000627079718836486	<a href="http://ada.computer.org">http://ada.computer.org</a> home page. All the pages in the website has link to the homepage. Hence in-link count is high.

**Summary:** Pages with high in-link count might or might not have a high Page rank value depending on the Page Rank of the incoming link. But high in-link count does not mean that the page is important. Some pages like search page, disclaimer, terms & conditions, FAQ pages have links from many pages but that does not mean that they are important nor are they likely to be the one which the user wants to see in response to a query.

E.g.: Intelligent information- search page, WWWBoard Frequently Asked Questions and Answers, DRCNet Online Library of Drug Policy Search page, The Economist: Copyright Notice Page.