Some statistics about Pagerank & SERP

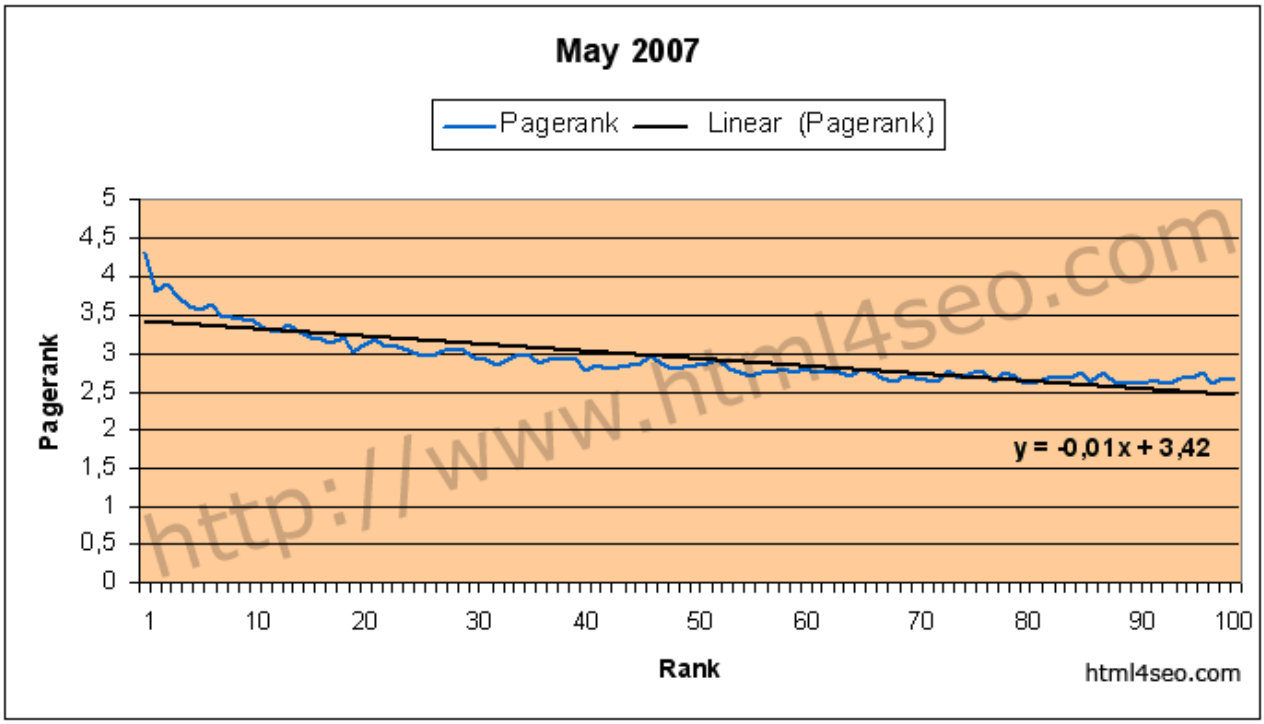
The Pagerank tracked

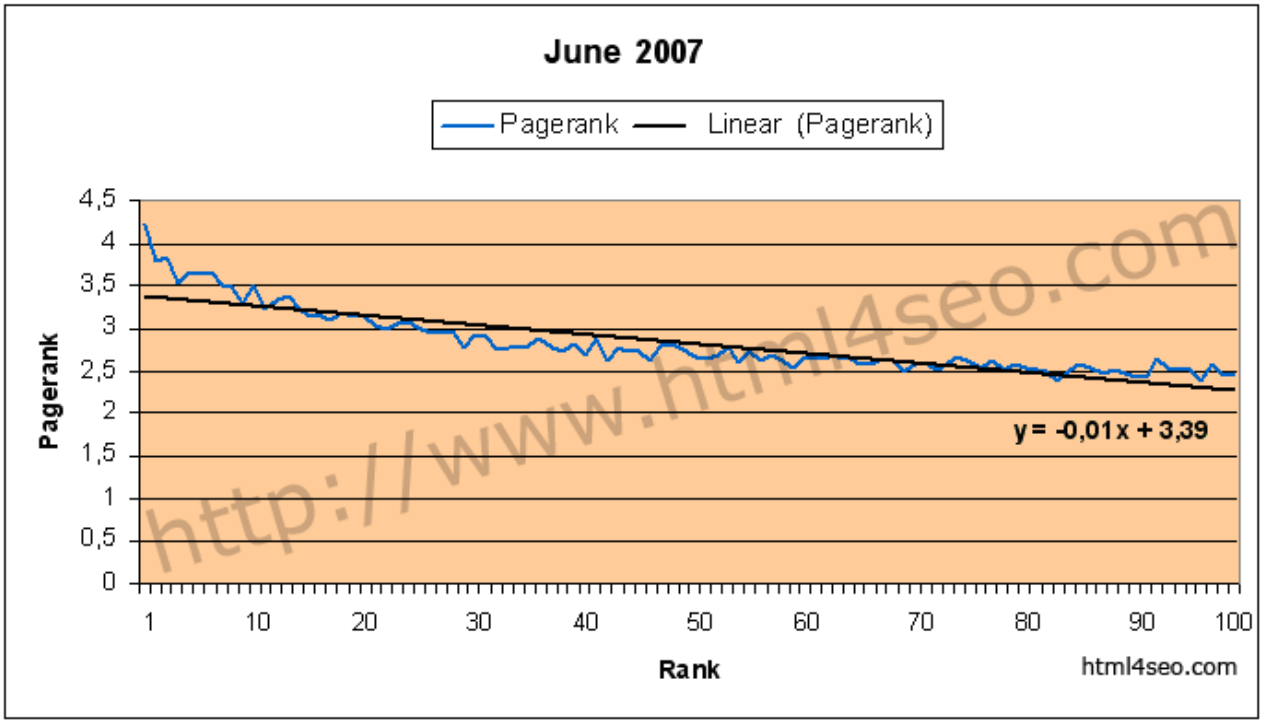
For several months, the HTML4SEO team has done thousands of SEO audits over the Google SERP. In each one were harvested the Pagerank, the number of backlinks and the HTML content for the 100 first of the Google SERP. Based on these results, the HTML4SEO team publishes this short statistical study (without any pretention), which should bring some conclusions based on real facts. We will try to limit technical data to the absolute minimum.

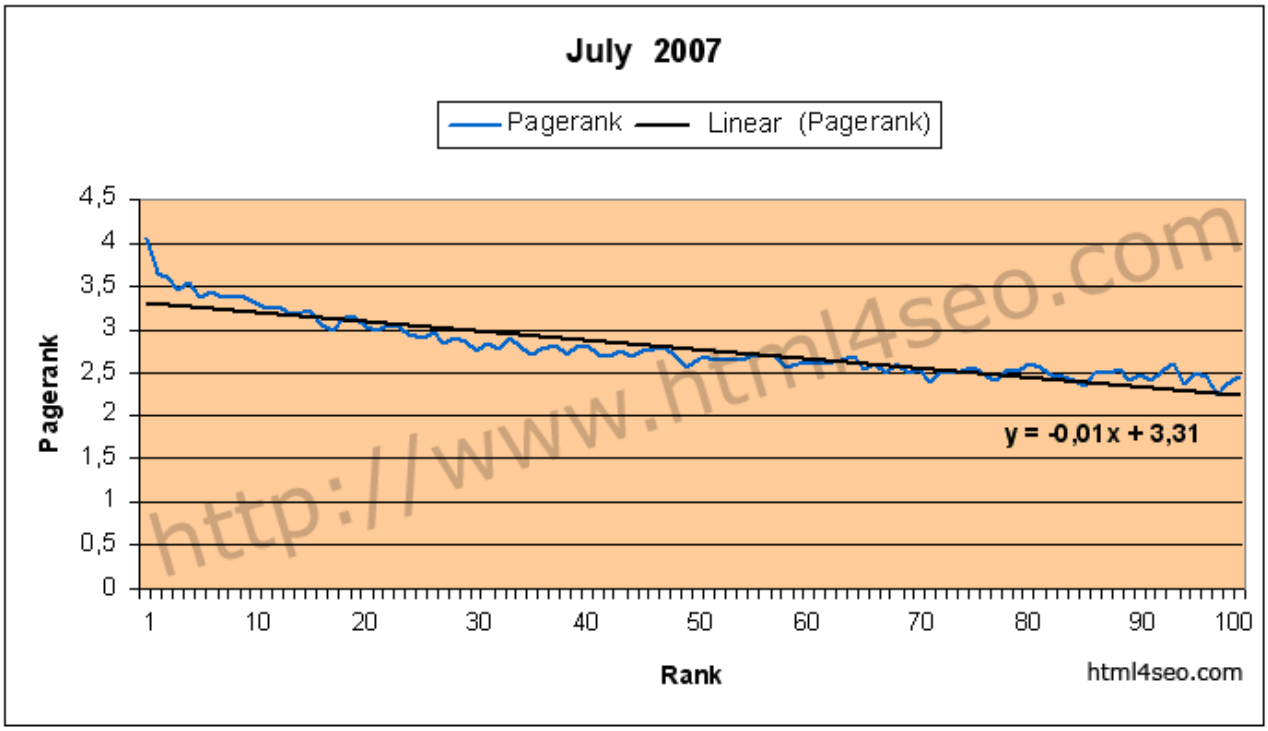
The study is restricted to the Pagerank, although we could extend it to the backlinks and the HTML content pages listed in the Google SERP. We will soon publish other studies on these 2 other major SEO criteria.

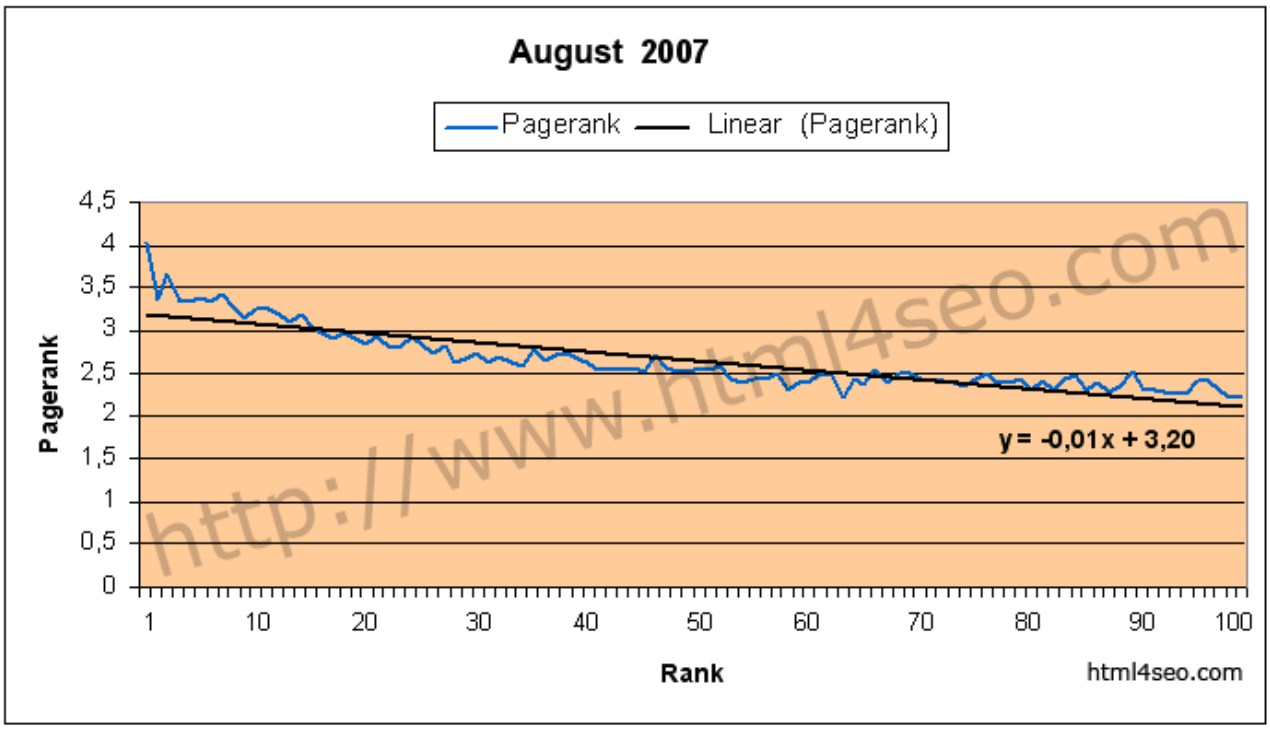
Pagerank Evolution in the Google SERP in the last 6 months

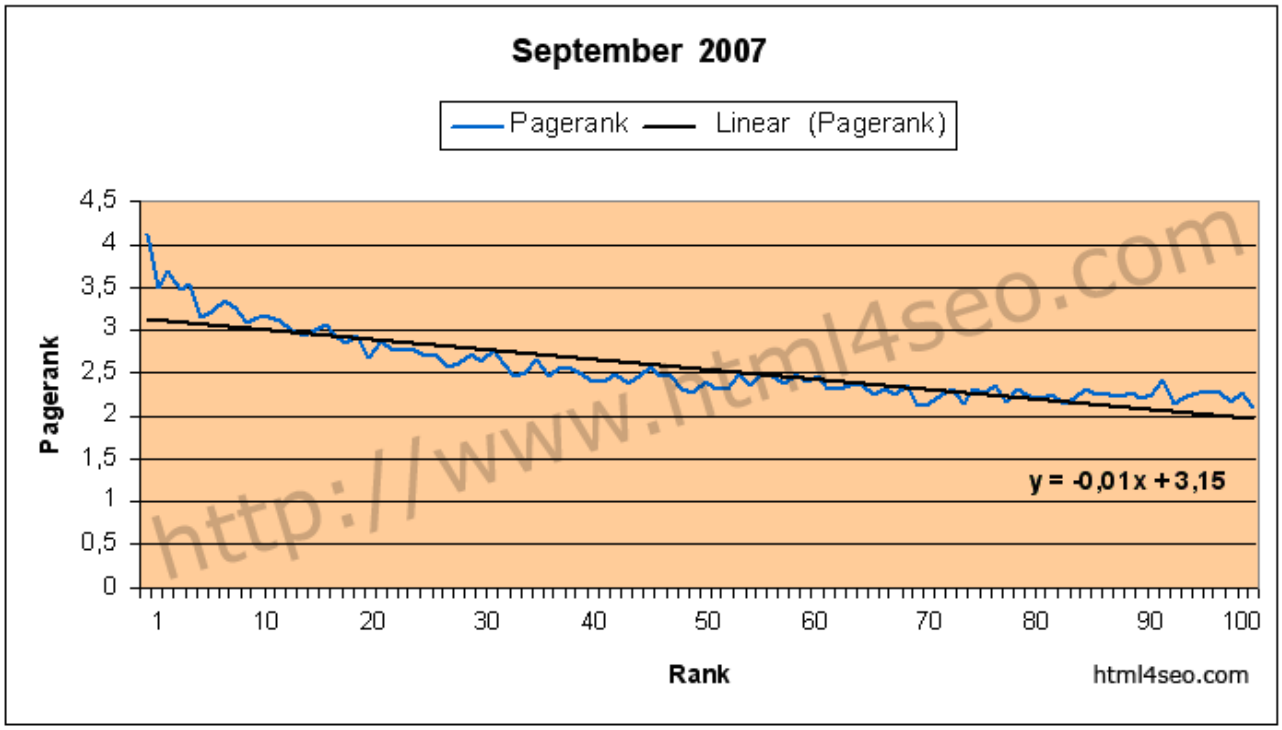
Pagerank Graphics

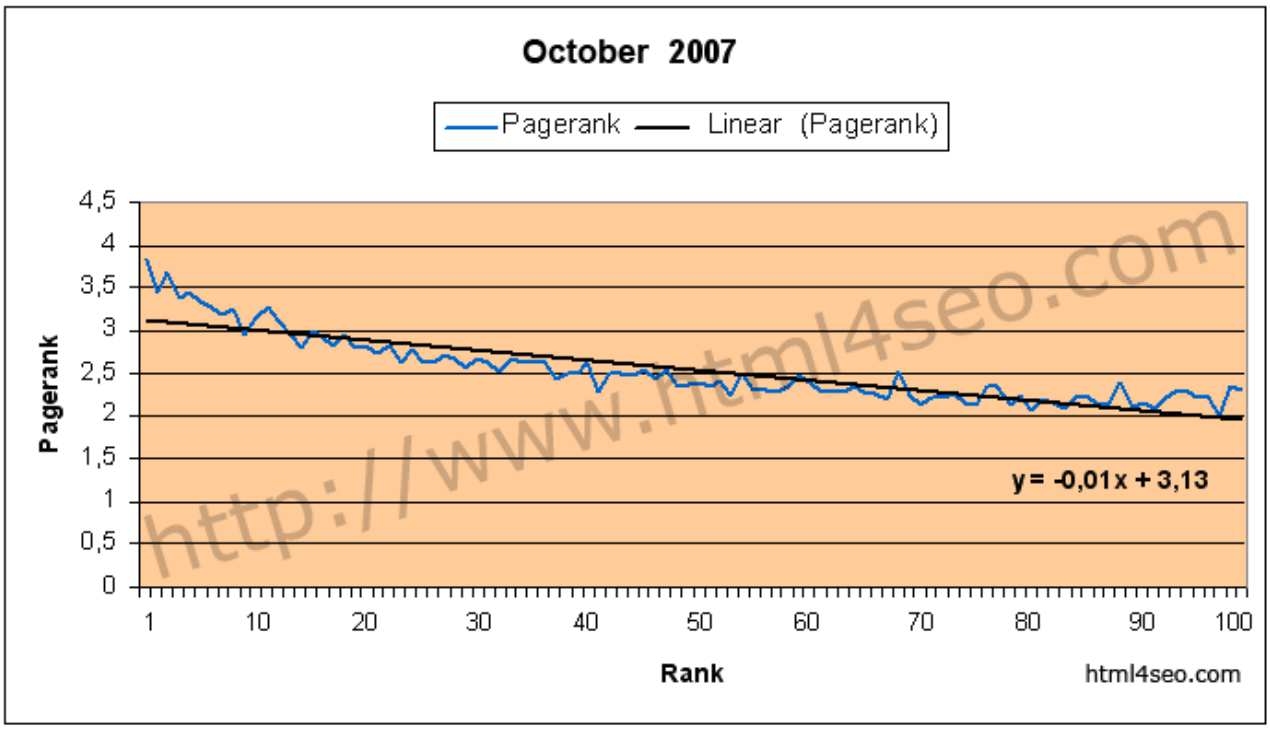



June 2007
Pagerank
Linear (Pagerank)
Pagerank
Rank
y = -0,01x + 3,39
html4seo.com
http://www.html4seo.com


July 2007
Pagerank
Linear (Pagerank)
Pagerank
Rank
y = -0,01x + 3,31
html4seo.com
http://www.html4seo.com


August 2007
Pagerank
Linear (Pagerank)
Pagerank
Rank
y = -0,01x + 3,20
html4seo.com
http://www.html4seo.com


September 2007
Pagerank
Linear (Pagerank)
Pagerank
Rank
y = -0,01x + 3,15
html4seo.com
http://www.html4seo.com



Rank/Pagerank Correlation

The most commonly used correlation measure ρ_{XY} between 2 variables X and Y is the Bravais-Pearson correlation coefficient often called correlation coefficient

link http://en.wikipedia.org/wiki/Correlation

The correlation coefficient ρ_{XY} varies between -1 and 1 and can be interpreted as follows.

If ρ_{XY} = 1 (resp. -1) then X and Y are considered as perfectly positively (or negatively) correlated. Knowing the value of X, we can infer the value of Y. The points (X, Y) are aligned on a line with a positive (or negative) slope.

If ρ_{XY} = 0 then X and Y are independents or dependents, but not linearly (the relationship between X and Y is more complex).

In the present case, X and Y are respectively the Rank (in the Google SERP) and the Pagerank. Below, data harvested from the 1000 first audits (Google SERP with more than 100 results) made each month from May to October 2007.

2007	correlation coef.
May	-0,85
June	-0,88
July	-0,90
August	-0,88
September	-0,86
October	-0,88
Mean	**-0,88**
Std. Dev.	**0,02**

Comments :

1st observation, the correlation coefficient between Rank and Pagerank is close to 1 in absolute terms, therefore, very strong. The Rank and the Pagerank are strongly correlated. A good pagerank improves ranking undoubtly.

2nd observation, the correlation coefficient between Rank and Pagerank is very stable over 6 months, around -0.88 with a standard deviation of 2%. The Pagerank weight in the Google algorithm has probably not changed. And the Rank depends on the Pagerank.

Those who said "Pagerank is dead!" or "Pagerank is just marketing" have missed the target. The pagerank is neither dead nor reduced to marketing, it has a significant part in the Google algorithm.

Linear interpolation Rank/Pagerank

Let us now have a look at the linear regression

link http://en.wikipedia.org/wiki/Linear_regression

	Linear regression	A Slope	B Y-intercept
may	y = -0,01x + 3,42	-0,01	3,42
june	y = -0,01x + 3,39	-0,01	3,39
july	y = -0,01x + 3,31	-0,01	3,31
august	y = -0,01x + 3,20	-0,01	3,20
september	y = -0,01x + 3,15	-0,01	3,15
october	y = -0,01x + 3,13	-0,01	3,13

Values have a 1/100 precision.

Comments :

1st observation, the A slope

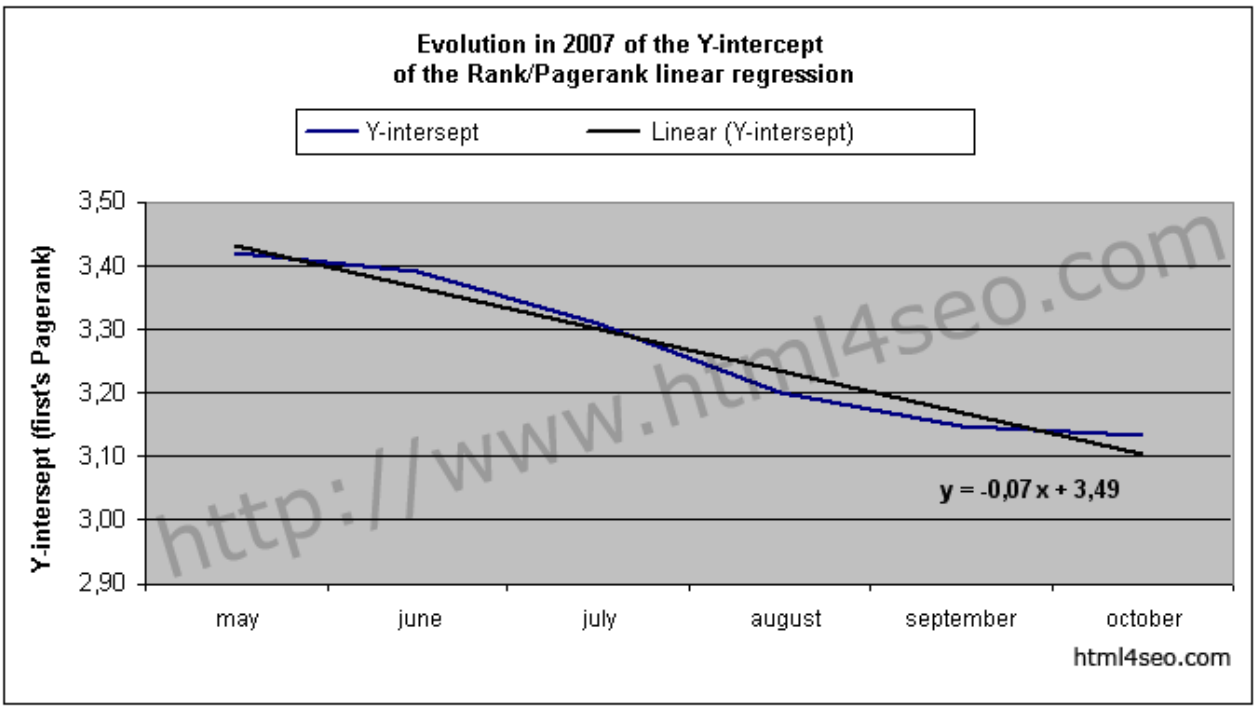
link http://en.wikipedia.org/wiki/Linear_function

is constant with a precision of 1/100 with a loss of 1 point of pagerank from the 1st to the 100th rank. This observation does not change along the months.

2nd observation, the B Y-intercept

link http://en.wikipedia.org/wiki/Linear_function

(linear interpolation of the 1st 's pagerank) decreases. In other words, it appears that the requirements in Pagerank to reach the top of the Google SERP decrease month after month.



The study of the distribution date (from May to October 2007) of the Y-intercept of the linear regression gives us a correlation coefficient of -0.98, which is a very strong correlation (close dependence) between these 2 variables. The linear regression is y = -0.07x +3.49, which reduces by 1 point of pagerank every 14 months the requirements of pagerank to reach the top SERP.

These results have to be taken with care. These data are just statistics, they only indicate general trends. Nevertheless, they are interesting.

What may have caused this loss of pagerank in the Google SERP?

As the world wide web grows, the relevant web pages with low pagerank are more numerous. In the huge quantity of web pages available, there are always many quality-distinguished web pages despite their weakness in Pagerank/Backlinks. This raises the question of the web page semantic relevance according to Google ?

Or maybe, the track for suspicious (black hat) backlinks began a few months ago. The backlinks considered as reliable are now fewer and fewer. And then, many web pages, better targeted semantic content (internally: HTML content and externally: Backlinks), reach the top 100. And this Pagerank evolution of the Google SERP is possible because the Pagerank is a global measure, non-sectorally focused.

Or, the recalibration of Pagerank (hypothesis suggested by some) began several months ago. And the recent visible Pagerank update would confirm this hypothesis. Note that a recalibration should involve a global loss (quasi-uniform) of Pagerank.

There are probably other hypothesis ...

The Pagerank of November 2007

It seems that the recent Pagerank changes are more important than the previous ones. Some have gained up to 3 points of Pagerank while others have lost up to 3. In addition, changes in Pagerank do not seem to be unidirectional. Given the strong correlation Rank/Pagerank (see above), those who lost may worry. Because this Pagerank loss will surely be accompanied by a ranking loss in the Google SERP and then by a web traffic loss.

We will try to have a look next month at this indicator evolution.

If you were out of laught reading some parts of this article, it may be beauce of our sense of translation. Feel free to indicate us any mistake.

HTML4SEO Team, November 2007

http://www.html4seo.com