



Rescaling of visits and time in gemiusAudience results

The high level of non-cookie page views on FDIM publications affects the statistics based on user identification - visits and time – by making them lower than they would be without non-cookie phenomenon. Therefore, since October 2012 the visits and time in gemiusAudience statistics (gemiusExplorer files and fdim.dk's toplists) will be rescaled proportionally to the average number of cookie page views per visit and time on a given node, in a given period.

The whole process of rescaling has been done in two steps. Rescaling as multiplication by a coefficient, and adjustment of values between parent and children nodes.

1. Rescaling

Rules of rescaling:

- Rescaling is done only for nodes with **PV > 1000**, in a given period of time
- **Visits** are rescaled the following way:
 $vis_all = vis_cookie * (PV_all/PV_cookie)$
vis_all - visits after rescaling
vis_cookie – visits calculated the standard way, based on cookie page views
PV_all - all page views (both cookie and non-cookie page views)
PV_cookie – only cookie page views
- **Time** is rescaled the following way:
 $time_all = time_cookie * (PV_all/PV_cookie)$
time_all - time after rescaling
time_cookie – time calculated in a standard way, based on vis_cookie

2. Adjustment

Rescaled data must be consistent regarding parent-child node relation. The following conditions have to be fulfilled:

- **Time**
Rule: sum of time on children nodes must be equal to time on the parent node
In practice: after rescaling of time of all nodes, the time of children nodes on each level is summed and this value replaces the rescaled time of the parent node.
Example:



node name	status	regular	rescaled	adjusted	adjusted vs. rescaled	
		time	time	time	time	time (%)
abc.dk Photography	parent	44 423 846	62 431 752	62 699 791	268 039	0,43%
abc.dk Photography Gallery	child	33 437 022	46 886 786	46 886 786	0	0,00%
abc.dk Photography News	child	10 986 824	15 813 005	15 813 005	0	0,00%

Figure 1. Adjusted time of the parent node “Photography” is a sum of time on the children nodes “Gallery” and “News” (15 813 005 + 46 886 786 = 62 699 791)

In October 2012, monthly results there were about 500 parent nodes (both public and un-public) corrected this way.

- **Visits**

Rule: visits on any of the children nodes cannot be higher than visits on parent node

In practice: If rescaled visits of any of the children node are higher than the rescaled visits of the parent node, the highest value of visits of the children nodes replaces the rescaled visits of the parent node.

Example:

node name	status	regular	rescaled	adjusted	adjusted vs. rescaled	
		visits	visits	visits	visits	visits %
abc.dk Photography	parent	209 723	294 738	298 938	4 200	1,43%
abc.dk Photography Gallery	child	167 175	234 420	234 420	0	0,00%
abc.dk Photography News	child	207 701	298 938	298 938	0	0,00%

Figure 2. Adjusted visits of the parent node “Photography” are equal to the rescaled visits of node “News”

In October 2012 monthly results there were only 2 nodes with visits increased this way.

- **Visits vs. page views**

There is no need for correction of page views for nodes with visits adjusted after rescaling.

Values of rescaled visits on a given node are always lower than page views on this node. This is because PV_all is multiplied by (vis_cookie/PV_cookie) <= 1. And PV_all on a child node is always lower than PV_all of the parent node, so vis_all of a child node that is assigned to the parent node cannot exceed the PV_all on the parent node.

Warsaw, 20.11.2012