

Ohio Digital Collections Analytics 2016: Report compiled for the Ohio Digitization Interest Group

Marcus Ladd, Miami University / Morgan State University, marcus.ladd@morgan.edu

Now in its third year, the purpose of this project is to establish benchmarks for the analytics of individual institutions' digital collections platforms. Although similar services are available by membership to certain organizations (e.g. Association of Research Libraries), this particular project provides a free alternative that focuses solely on Ohio libraries.

Similar trends to last year are observed, including the overall noticeable correlation between the number of items in the digital collections platform and a geographically broader base of users. A few of this year's most noticeable trends were found especially in the public libraries, where the time spent on digitization was very strongly correlated with an increase in users, page views, and sessions. Also worth noting among public libraries was the particularly strong correlations between the size of the holdings (number of items) and overall use (users, page views, and sessions). For academic institutions, the strongest predictor of greater use of the collections was the size of the university itself (as measured in population served).

Our sample size has been slightly smaller compared to last year, but still larger than 2014.

Year	n
2014	10
2015	15
2016	12

Means and ranges for key stats for ALL libraries and archives were as follows (n = 12):

	Minimum	Mean	Maximum	Standard deviation
Weekly hours digitization	1	44	150	42
Total items online	364	28,756	103,261	32,701
Sessions	648	28,757	93,805	30,352
Users	446	20,542	74,233	22,916
Page views	5,920	131,625	376,300	134,262
Average duration (sec)	175	298	660	135
Bounce rate	47	55	65	6
% of users from USA	45	80	100	16
% of users from Ohio	19	47	63	18
% of users from local	7	28	44	14

Means and ranges for key stats for PUBLIC libraries and archives were as follows (n = 6):

	Minimum	Mean	Maximum	Standard deviation
Weekly hours digitization	1	61	150	49
Total items online	364	14,390	62,631	22,321
Sessions	648	19,162	67,731	22,402
Users	446	13,522	48,272	16,064
Page views	5,920	101,948	315,552	99,479
Average duration (sec)	246	361	660	141
Bounce rate %	48	53	61	5
% of users from USA	84	88	93	3
% of users from Ohio	58	61	63	2
% of users from local	30	35	39	4

Means and ranges for key stats for ACADEMIC libraries and archives were as follows (n = 5):

	Minimum	Mean	Maximum	Standard deviation
Weekly hours digitization	1	30	60	25
Total items online	39,28	42,184	103,261	38,086
Sessions	27,21	46,149	93,805	34,715
Users	1,085	31,072	74,233	27,222
Page views	10,844	167,237	376,300	159,555
Average duration (sec)	175	205	269	38
Bounce rate %	47	57	65	8
% of users from USA	45	66	84	14
% of users from Ohio	19	33	52	14
% of users from local	7	22	44	16

CORRELATIONS BETWEEN STATISTICS

ALL (n = 12)

	Weekly Time	Pop.	Collections	Items	Sessions	Users	Pages	Duration	Bounce	% USA	% Ohio	% Local
Weekly Time	1	0.77	<u>-0.18</u>	<u>0.28</u>	<i>0.40</i>	<i>0.42</i>	<u>0.38</u>	-0.03	<u>0.25</u>	<u>0.37</u>	<i>0.48</i>	<i>0.48</i>
Pop		1	<u>-0.13</u>	-0.00	0.09	0.00	<u>0.20</u>	-0.03	<u>0.13</u>	<u>0.35</u>	<i>0.52</i>	<u>0.33</u>
Collections			1	<u>0.37</u>	<i>0.41</i>	<u>0.38</u>	<i>0.50</i>	<i>-0.56</i>	<u>0.32</u>	<i>0.44</i>	<u>-0.33</u>	<i>-0.60</i>
Items				1	<u>0.38</u>	<u>0.33</u>	<i>0.47</i>	<i>-0.29</i>	<u>0.18</u>	-0.07	-0.72	-0.73
Sessions					1	0.99	0.94	<u>-0.32</u>	0.06	<u>0.31</u>	0.01	<u>0.28</u>
Users						1	0.88	<u>-0.34</u>	<u>0.15</u>	<u>0.33</u>	<u>0.10</u>	<u>0.39</u>
Pages							1	<u>-0.28</u>	<u>-0.13</u>	<u>0.26</u>	<u>-0.22</u>	-0.03
Duration								1	<i>-0.62</i>	<u>0.39</u>	<u>0.28</u>	0.06
Bounce									1	<u>-0.23</u>	0.02	<u>0.16</u>
% USA										1	0.93	0.89
% Ohio											1	0.87
% Local												1

Correlations where $r \geq 0.7$ were considered 'strong' and are marked in **bold red**

Correlations where $0.4 \leq r < 0.7$ were considered 'moderate' and are marked in *blue italics*

Correlations where $0.1 \leq r < 0.4$ were considered 'weak' and are marked in underlined green

PUBLIC (n = 7)

	Weekly Time	Pop.	Collections	Items	Sessions	Users	Pages	Duration	Bounce	% USA	% Ohio	% Local
Weekly Time	1	0.79	<u>0.10</u>	0.87	0.79	0.78	0.85	<u>-0.26</u>	<u>0.15</u>	<i>0.48</i>	0.09	<i>0.55</i>
Pop		1	<u>0.16</u>	<i>0.54</i>	<i>.051</i>	<i>0.50</i>	<i>0.59</i>	<i>-0.46</i>	<i>0.56</i>	<u>0.28</u>	<i>0.41</i>	<u>0.25</u>
Collections			1	<u>0.15</u>	<u>0.36</u>	<u>0.38</u>	<u>0.39</u>	<i>-0.67</i>	<i>0.56</i>	<u>-0.10</u>	<i>0.67</i>	-0.99
Items				1	0.95	0.95	0.96	<u>-0.24</u>	<u>-0.18</u>	<u>0.34</u>	<u>0.33</u>	<u>0.33</u>
Sessions					1	1.00	0.99	<u>-0.26</u>	<u>-0.12</u>	0.07	<i>0.65</i>	-0.03
Users						1	0.99	<u>-0.28</u>	<u>-0.11</u>	0.08	<i>0.66</i>	-0.05
Pages							1	<u>-0.31</u>	-0.02	<u>0.15</u>	<i>0.60</i>	0.03
Duration								1	-0.71	<i>-0.49</i>	-0.74	1.00
Bounce									1	<u>0.31</u>	-0.01	<i>-0.62</i>
% USA										1	-0.90	0.97
% Ohio											1	-0.78
% Local												1

Correlations where $r \geq 0.7$ were considered 'strong' and are marked in **bold red**

Correlations where $0.4 \leq r < 0.7$ were considered 'moderate' and are marked in *blue italics*

Correlations where $0.1 \leq r < 0.4$ were considered 'weak' and are marked in underlined green

ACADEMIC (n = 6)

	Weekly Time	Pop.	Collections	Items	Sessions	Users	Pages	Duration	Bounce	% USA	% Ohio	% Local
Weekly Time	1	<u>0.25</u>	<u>0.32</u>	<u>0.38</u>	<i>0.47</i>	<i>0.58</i>	<u>0.23</u>	<u>-0.34</u>	0.97	0.71	<i>0.69</i>	<i>0.62</i>
Pop		1	0.74	<u>0.37</u>	0.81	0.71	0.95	<i>0.67</i>	-0.88	0.78	-0.04	0.05
Collections			1	0.84	<u>0.34</u>	<u>0.24</u>	<i>0.66</i>	0.81	<i>-0.65</i>	<i>0.54</i>	-0.98	-0.97
Items				1	<u>-0.14</u>	<u>-0.19</u>	<u>0.19</u>	<i>0.55</i>	0.09	<u>0.18</u>	-0.82	-0.86
Sessions					1	0.98	0.89	<u>0.11</u>	<u>-0.19</u>	0.93	0.76	0.81
Users						1	0.79	-0.05	0.04	0.93	0.88	0.92
Pages							1	<i>0.55</i>	-0.76	0.81	<u>0.19</u>	<u>0.27</u>
Duration								1	-0.89	<u>0.14</u>	-0.84	-0.80
Bounce									1	<u>0.13</u>	<i>0.50</i>	<i>0.43</i>
% USA										1	0.92	0.95
% Ohio											1	0.99
% Local												1

Correlations where $r \geq 0.7$ were considered 'strong' and are marked in **bold red**

Correlations where $0.4 \leq r < 0.7$ were considered 'moderate' and are marked in *blue italics*

Correlations where $0.1 \leq r < 0.4$ were considered 'weak' and are marked in underlined green

KEY

Weekly time (hours): Estimated employee time spent on digital collections

Pop. Served: Estimated population of primary service area Total

items: Total items available in digital collections:

Total collections: Total collections the items are organized into

Sessions: Number of times the digital collections website was accessed

Users: Number of individual users of the digital collections (measured by unique IP addresses)

Pages: Number of individual page (e.g. item) views

Avg. duration (sec): Average sessions duration, measured in sec

Bounce rate: Estimated rate of sessions where the user opens the website and immediately leaves without accessing the information in the digital collections

% USA: Percentage of users coming from USA IP addresses

% Ohio: Percentage of users coming from Ohio IP addresses

% Local: Percentage of users coming from the institution's city IP addresses