

Project Info

#	Project	Age (yr)	#Contributor	#LOC	#Commit	#Issue	#Bug	#PL	Main PL	%Main PL
1	Airavata	9	102	1316K	9761	3629	1673	5	Java	74.8
2	Ambari	12	232	1480K	24,614	25,386	17,961	7	Java	44.7
3	Archiva	9	49	225K	8,782	2,012	1,165	3	Java	81.6
4	Arrow	7	952	1038K	12,384	16,953	5,576	9	C/C++	50.6
5	Atlas	6	191	332K	3,697	4,582	2,759	3	Java	60.2
6	Avro	14	311	225K	3,403	3,565	1,523	8	Java	41.7
7	Beam	7	1350	1674K	37,126	14,551	6,008	8	Java	66.0
8	Bigtop	12	207	294K	2,892	3,733	1,957	4	Python	41.9
9	CarbonData	7	222	380K	4,962	4,343	2,405	5	Scala	58.9
10	CloudStack	10	715	1324K	35,454	10,342	7,861	5	Java	66.5
11	Cocoon	14	73	398K	13,161	2,365	1,726	4	Java	76.5
12	CouchDB	14	221	162K	12,900	3,292	1,878	5	Erlang	76.4
13	Falcon	10	69	190K	2,227	2,343	1,169	4	Java	86.5
14	Flink	9	1629	2272K	32,037	28,170	11,003	6	Java	86.1
15	Guacamole-client	7	93	94K	6,223	1,637	856	3	Java	61.7
16	Harmony	13	24	1599K	956	6,676	4,305	4	Java	72.9
17	HAWQ	8	107	1741K	1,764	1,825	910	4	C/C++	70.2
18	Hudi	6	346	304K	3,387	4,400	1,177	3	Java	81.6
19	Ignite	8	414	1795K	28,292	17,333	7,242	7	Java	78.8
20	Impala	7	249	886K	10,397	11,400	6,238	4	C/C++	51.9
21	Isis(Causeway)	10	63	695K	21,281	3,080	870	4	Java	77.8
22	James-project	8	109	695K	13,961	3,714	1,375	4	Java	87.9
23	JSPWiki	14	28	127K	9,220	1,054	583	2	Java	77.2
24	jUDDI	9	18	251K	2,996	1,018	583	3	Java	80.4
25	Kafka	12	1163	723K	10,420	12,799	6,808	3	Java	74.7
26	Kudu	7	173	490K	9,800	3,226	1,728	7	C/C++	79.9
27	Kylin	8	283	337K	8,773	5,052	2,546	5	Java	58.9
28	Mahout	9	65	214K	4,506	2,126	827	5	Java	80.5
29	Metron	7	80	333K	1,453	2,349	974	7	Java	59.4
30	MXNet	8	980	463K	11,893	1,404	375	2	C/C++	48.5
31	NiFi	8	545	1138K	7,719	10,137	4,427	7	Java	87.3
32	OFBiz	14	38	1039K	24,427	12,646	4,740	4	Java	46.0
33	OpenMeetings	8	21	138K	3,527	2,743	1,804	2	Java	82.3
34	Qpid	14	58	1018K	14,182	8,485	5,097	3	Java	64.2
35	Qpid-dispatch	9	42	172K	3,471	2,349	1,386	3	Python	40.5
36	Ranger	9	156	747K	3,997	3,788	2,460	4	Java	69.3
37	Rave	9	27	108K	1,555	1,286	270	4	Java	69.5
38	REEF	8	99	198K	3,875	2,066	501	6	Java	52.8
39	Samza	8	209	182K	2,512	2,748	1,467	4	Java	87.2
40	Shindig	14	42	169K	4,786	1,979	1,032	3	Java	62.1
41	Spark	9	2738	1107K	34,239	39,087	15,508	5	Scala	67.1
42	Storm	9	506	396K	10,517	3,818	1,773	6	Java	83.5
43	Stratos	9	81	529K	8,951	1,662	768	5	Java	56.3
44	Submarine	4	72	177K	971	1,281	268	5	Java	62.1
45	Subversion	22	186	866K	60,209	4,537	3,247	5	C/C++	66.7
46	Tapestry-5	11	50	439K	5,695	2,726	1,350	5	Java	51.2
47	Thrift	14	558	329K	6,654	5,590	3,058	13	C/C++	33.1
48	TinkerPop	8	222	668K	18,955	2,751	787	6	Java	69.7
49	TrafficServer	13	440	668K	14,101	5,103	3,026	5	C/C++	81.4
50	Trafodion	8	200	2540K	7,254	3,337	2,341	7	C/C++	62.3
51	Usergrid	9	119	427K	10,954	1,341	350	8	Java	66.3
52	VCL	5	19	176K	2,626	1,122	466	3	Perl	47.0
53	Zeppelin	8	507	307K	5,167	5,731	2,921	5	Java	60.8
54	ZooKeeper	14	250	178K	2,433	4,371	2,299	5	Java	75.3

In the table, where (1) Age is the time interval (in years) from the creation time to the last update time of a project; (2) #Contributor is the number of contributors (including anonymous contributors), collected through the GitHub API; (3) #LOC is the number of lines of source code in the latest version of the project, calculated by the tool CLOC¹; (4) #Commit is the total number of commits, collected from the GitHub API. (5) #Issue is the total number of issues collected through the JIRA REST API; (6) #Bug is the number of issues of type Bug; and (7) #PL is the number of PLs used in the project, according to criterion C4.

¹<https://github.com/AIDanial/cloc>