# Overview in General

This file contains a general overview of the data in the graph including node labels and relationships types.

#### References

- jqassistant
- Neo4j Python Driver

#### **Node Labels**

#### Table 1a - Highest node count by label combination

Lists the 30 label combinations with the highest number of nodes. The labels with the lowest node count are listed in table 1b. The total list would sum up to the total number of labels (100%).

The whole table can be found in the CSV report Node\_label\_combination\_count.

	nodeLabels	nodesWithThatLabels	nodesWithThatLabelsPercent
0	[Git, Change]	256169	74.178566
1	[Java, ByteCode, Member, Method]	13204	3.823467
2	[Java, ByteCode, Parameter]	13029	3.772793
3	[Git, Commit]	12399	3.590364
4	[File, Git]	10780	3.121552
5	[Java, ByteCode, Bound, ParameterizedType]	7176	2.077946
6	[Java, ByteCode, Bound]	7118	2.061151
7	[Java, ByteCode, Member, Field]	3505	1.014939
8	[Java, ByteCode, Bound, WildcardType]	2908	0.842066
9	[Java, Value, ByteCode, Annotation]	2822	0.817163
10	[Xml, Element]	2144	0.620836
11	[Java, ByteCode, Member, Method, Constructor]	2049	0.593327
12	[Xml, Text]	1436	0.415821
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.317367
14	[Java, ByteCode, Member, Method, Lambda]	942	0.272774
15	[Type, File, Java, ByteCode, ResolvedDuplicate	899	0.260322
16	[Type, File, Java, ByteCode, Class]	864	0.250187
17	[Type, File, Java, ByteCode, JavaType]	748	0.216598
18	[Java, Value, ByteCode, Primitive]	652	0.188799
19	$[{\tt Java,ByteCode,GenericDeclaration,Member,M}$	569	0.164765
20	[Json, Key]	560	0.162159
21	[Value, Json, Scalar]	544	0.157525
22	[Type, File, Java, ByteCode, ExternalType]	396	0.114669
23	[Author, Git, Person]	298	0.086292
24	[Value, Array]	267	0.077315
25	[Committer, Git, Person]	252	0.072971
26	[Type, File, Java, ByteCode, Class, GenericDec	236	0.068338
27	[Value, Property]	202	0.058493
28	[Type, File, Java, ByteCode, Interface]	188	0.054439
29	[Java, Value, ByteCode, Class]	172	0.049806

# Chart 1a - Highest node count by label combination

Values under 0.5% will be grouped into "others" to get a cleaner plot. The group "others" is then broken down in Chart 1b.

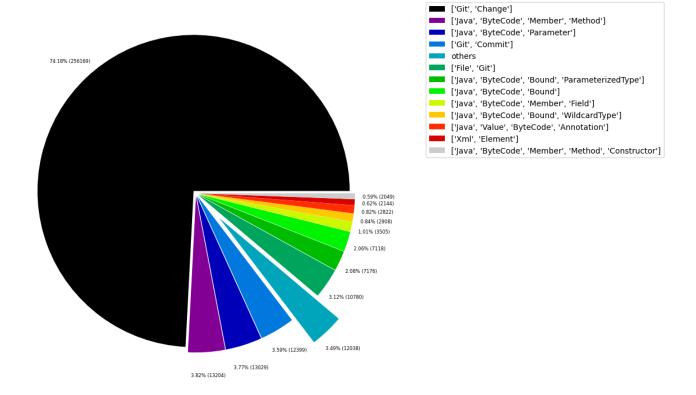


Table 1b - Lowest node count by label combination

Lists the 30 label combinations with the lowest number of nodes until they reach 0.5% of the total node count, which are shown above.

	nodeLabels	nodesWithThatLabels	nodesWithThatLabelsPercent	
0	[Analyze, Task, jQAssistant]	1	0.000290	
1	[Repository, File, Git]	1	0.000290	
2	[File, Json]	2	0.000579	
3	[File]	3	0.000869	
4	[Java, ByteCode, GenericDeclaration, Member, M	4	0.001158	
5	[Maven, Exclusion]	5	0.001448	
6	[Value, Array, Json]	6	0.001737	
7	[Type, File, Java, ByteCode, Void]	9	0.002606	
8	[File, Maven, Xml, Pom, Document]	9	0.002606	
9	[Java, ManifestSection]	9	0.002606	
10	[File, Java, Manifest]	9	0.002606	
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002606	
12	[File, Java, ServiceLoader]	10	0.002896	
13	[File, Java, Properties]	12	0.003475	
14	[Maven, ExecutionGoal]	16	0.004633	
15	[Maven, PluginExecution]	16	0.004633	
16	[Xml, Attribute]	18	0.005212	
17	[jQAssistant, Rule, Concept]	19	0.005502	
18	[Maven, Plugin]	21	0.006081	
19	[Maven, Configuration]	21	0.006081	
20	[Type, File, Java, ByteCode, Enum]	28	0.008108	
21	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008397	
22	[Git, Branch]	35	0.010135	
23	[Xml, Namespace]	36	0.010424	
24	[Type, File, Java, ByteCode, Annotation]	43	0.012451	
25	[File, Directory]	44	0.012741	
26	[Type, File, Java, ByteCode, GenericDeclaratio	85	0.024613	
27	[Artifact, Maven]	92	0.026640	
28	[Type, File, Java, ByteCode, ExternalAnnotatio	93	0.026930	
29	[Value, ManifestEntry]	120	0.034748	

### Chart 1b - Lowest node count by label combination

Shows the lowest (less than 0.5% overall) node count label combinations. Therefore, this plot breaks down the "others" slice of the pie chart above. Values under 0.01% will be grouped into "others" to get a cleaner plot.

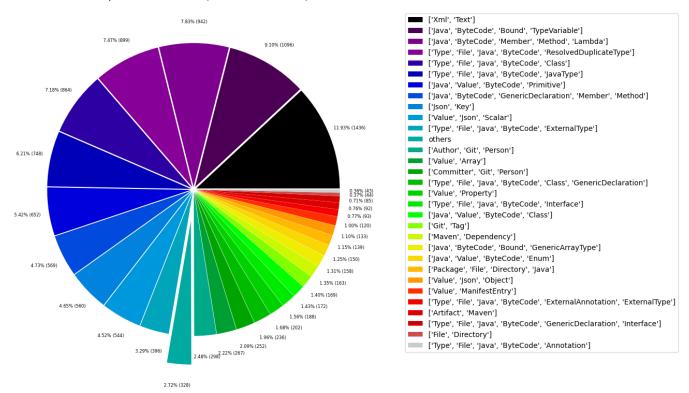


Table 1c - Highest node count by single label

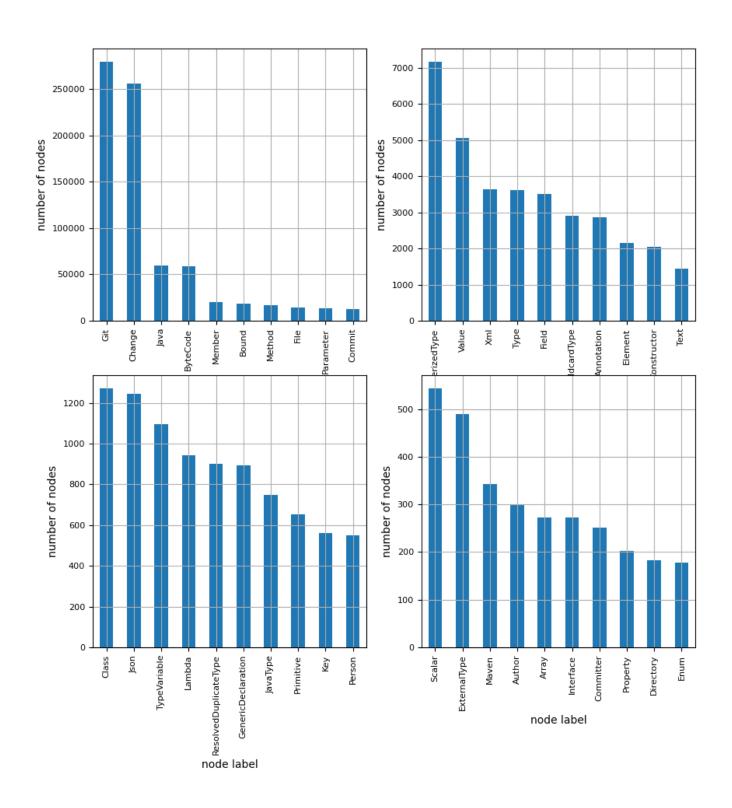
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodes With That Label	nodes With That Label Percent
0	Git	280103	81.109107
1	Change	256169	74.178566
2	Java	59360	17.188808
3	ByteCode	59172	17.134369
4	Member	20273	5.870430
5	Bound	18456	5.344283
6	Method	16768	4.855491
7	File	14636	4.238130
8	Parameter	13029	3.772793
9	Commit	12399	3.590364
10	ParameterizedType	7176	2.077946
11	Value	5068	1.467535
12	Xml	3643	1.054899
13	Туре	3618	1.047660
14	Field	3505	1.014939
15	WildcardType	2908	0.842066
16	Annotation	2865	0.829615
17	Element	2144	0.620836
18	Constructor	2053	0.594485
19	Text	1436	0.415821
20	Class	1272	0.368332
21	Json	1245	0.360513
22	TypeVariable	1096	0.317367
23	Lambda	942	0.272774
24	ResolvedDuplicateType	899	0.260322
25	GenericDeclaration	894	0.258875
26	JavaType	748	0.216598
27	Primitive	652	0.188799
28	Key	560	0.162159
29	Person	550	0.159263
30	Scalar	544	0.157525
31	ExternalType	489	0.141599
32	Maven	343	0.099322
33	Author	298	0.086292
34	Array	273	0.079052
35	Interface	273	0.079052
36	Committer	252	0.072971
37	Property	202	0.058493
38	Directory	183	0.052991
39	Enum	178	0.051543

# Chart 1c - Highest node count by label

Shows the 40 labels with the highest number of nodes.

# Node count by label



# Relationship Types

Table 2a - Highest relationship count by type

Lists the 30 relationship types with the highest number of occurrences. The whole table can be found in the CSV report Relationship\_type\_count .

Total number of relationships: 1071674

	voletien ehin Tune	nedeeWithThetDeletienehinTime	nodes With The tDeletions him Tome Devector	
_			nodesWithThatRelationshipTypePercent	
0	CONTAINS_CHANGE	256169	23.903631	
1	MODIFIES	256169	23.903631	
2	UPDATES	167868	15.664092	
3	CREATES	60933	5.685778	
4	DELETES	37874	3.534097	
5	INVOKES	35669	3.328344	
6	COMMITTED	24798	2.313950	
7	DEPENDS_ON	21931	2.046425	
8	OF_TYPE	21394	1.996316	
9	DECLARES	20718	1.933237	
10	OF_RAW_TYPE	17093	1.594981	
11	HAS_PARENT	15065	1.405745	
12	HAS	14125	1.318031	
13	RETURNS	12578	1.173678	
14	HAS_COMMIT	12399	1.156975	
15	HAS_FILE	10780	1.005903	
16	RENAMES	10506	0.980335	
17	READS	9164	0.855111	
18	HAS_ACTUAL_TYPE_ARGUMENT	8288	0.773370	
19	HAS_NEW_NAME	6220	0.580400	
20	OF_GENERIC_TYPE	5906	0.551100	
21	RESOLVES_TO	5189	0.484196	
22	SIMILAR	3977	0.371102	
23	WRITES	3818	0.356265	
24	CONTAINS	3814	0.355892	
25	RETURNS_GENERIC	3545	0.330791	
26	ANNOTATED_BY	2810	0.262207	
27	REQUIRES	2174	0.202860	
28	HAS_FIRST_CHILD	2144	0.200061	
29	HAS_LAST_CHILD	2144	0.200061	

### Chart 2a - Highest relationship count by type

Values under 0.5% will be grouped into "others" to get a cleaner plot. The group "others" is then broken down in the second chart.

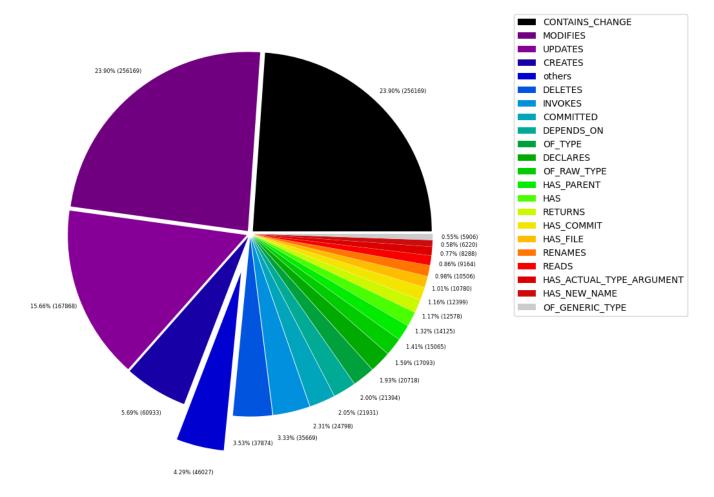


Table 2b - Lowest relationship count by type

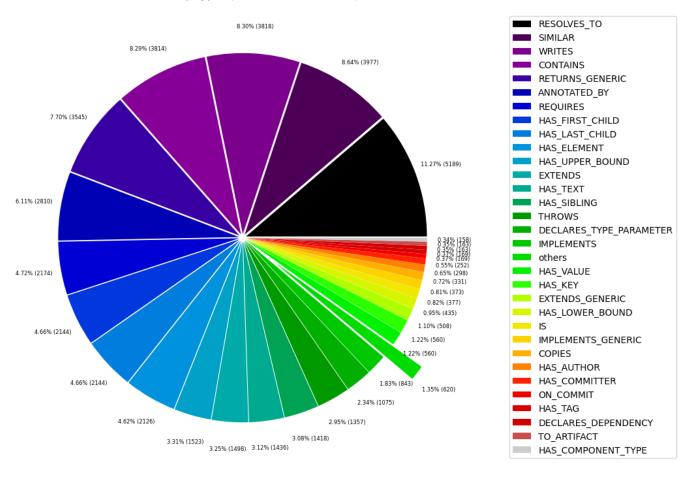
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodes With That Relationship Type	nodesWithThatRelationshipTypePercent	
0	HAS_PROPERTY	1	0.000093	
1	THROWS_GENERIC	5	0.000467	
2	EXCLUDES	5	0.000467	
3	DESCRIBES	9	0.000840	
4	HAS_ROOT_ELEMENT	11	0.001026	
5	HAS_GOAL	16	0.001493	
6	HAS_EXECUTION	16	0.001493	
7	OF_NAMESPACE	18	0.001680	
8	HAS_ATTRIBUTE	18	0.001680	
9	INCLUDES_CONCEPT	19	0.001773	
10	USES_PLUGIN	21	0.001960	
11	REQUIRES_TYPE_PARAMETER	21	0.001960	
12	IS_ARTIFACT	21	0.001960	
13	HAS_CONFIGURATION	21	0.001960	
14	REQUIRES_CONCEPT	28	0.002613	
15	HAS_BRANCH	35	0.003266	
16	DECLARES_NAMESPACE	36	0.003359	
17	HAS_HEAD	36	0.003359	
18	HAS_DEFAULT	36	0.003359	
19	CONTAINS_VALUE	121	0.011291	
20	COPY_OF	126	0.011757	
21	HAS_COMPONENT_TYPE	158	0.014743	
22	TO_ARTIFACT	163	0.015210	
23	DECLARES_DEPENDENCY	163	0.015210	
24	ON_COMMIT	169	0.015770	
25	HAS_TAG	169	0.015770	
26	HAS_COMMITTER	252	0.023515	
27	HAS_AUTHOR	298	0.027807	
28	COPIES	331	0.030886	
29	IMPLEMENTS_GENERIC	373	0.034805	

### Chart 2b - Lowest relationship count by type

Shows the lowest (less than 0.5% overall) relationship types. This plot breaks down the "others" slice of the pie chart above. Values under 0.01% will be grouped into "others" to get a cleaner plot.

#### Relationship types (less than 0.5% overall)



# Node labels with their relationships

# Table 3a - Highest relationship count by node labels and relationship type

Lists the 30 node labels and their relationship types with the highest number of occurrences.

	sourceLabels	relationType	targetLabels	numberOfRelationships	number Of Nodes With Same Labels As Source	numberOfNodesWithS
0	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	256169	12399	
1	[Git, Change]	MODIFIES	[File, Git]	256169	256169	
2	[Git, Change]	UPDATES	[File, Git]	167868	256169	
3	[Git, Change]	CREATES	[File, Git]	60933	256169	
4	[Git, Change]	DELETES	[File, Git]	37874	256169	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21794	13110	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	15056	12399	
7	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	12399	1	
8	[Author, Git, Person]	COMMITTED	[Git, Commit]	12399	298	
9	[Committer, Git, Person]	COMMITTED	[Git, Commit]	12399	252	
10	[Repository, File, Git]	HAS_FILE	[File, Git]	10780	1	
11	[Git, Change]	RENAMES	[File, Git]	10506	256169	
12	[Java, ByteCode, Member, Method]	HAS	[Java, ByteCode, Parameter]	8360	13110	
13	[Java, ByteCode, Member, Method]	READS	[Java, ByteCode, Member, Field]	8218	13110	
14	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6371	13029	
15	[File, Git]	HAS_NEW_NAME	[File, Git]	6220	10780	
16	[Type, File, Java, ByteCode, Class]	DECLARES	[Java, ByteCode, Member, Method]	5105	814	
17	[Type, File, Java, ByteCode, Class]	DEPENDS_ON	[Type, File, Java, ByteCode, JavaType]	4480	814	
18	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3570	7118	
19	[Java, ByteCode, Bound, ParameterizedType]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3053	7176	
20	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, WildcardType]	2908	7176	
21	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound, ParameterizedType]	2657	13029	
22	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, TypeVariable]	2454	7176	
23	[Java, ByteCode, Member, Method, Constructor]	WRITES	[Java, ByteCode, Member, Field]	2438	2049	
24	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalAnnotatio	2341	2822	
25	[Type, File, Java, ByteCode, Class]	DECLARES	[Java, ByteCode, Member, Field]	2203	814	
26	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
27	[Java, ByteCode, Member, Method, Constructor]	INVOKES	[Java, ByteCode, Member, Method, Constructor]	2117	2049	
28	[Type, File, Java, ByteCode, Class, GenericDec	DECLARES	[Java, ByteCode, Member, Method]	2116	231	
29	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2091	13110	

# **Graph Density**

total\_number\_of\_nodes (vertices): 345341

total\_number\_of\_relationships (edges): 1071674

- -> total directed graph density: 8.986025929830454e-06
- -> total directed graph density in percent: 0.0008986025929830453