# 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]	240725	73.160910
1	[Java, ByteCode, Member, Method]	13180	4.005653
2	[Java, ByteCode, Parameter]	13000	3.950947
3	[Git, Commit]	11850	3.601441
4	[File, Git]	10613	3.225493
5	[Java, ByteCode, Bound, ParameterizedType]	7166	2.177884
6	[Java, ByteCode, Bound]	7112	2.161472
7	[Java, ByteCode, Member, Field]	3490	1.060677
8	[Java, ByteCode, Bound, WildcardType]	2906	0.883189
9	[Java, Value, ByteCode, Annotation]	2821	0.857356
10	[Xml, Element]	2144	0.651602
11	[Java, ByteCode, Member, Constructor, Method]	2039	0.619691
12	[Xml, Text]	1436	0.436428
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.333095
14	[Java, ByteCode, Member, Method, Lambda]	936	0.284468
15	[Type, File, Java, ByteCode, ResolvedDuplicate	900	0.273527
16	[Type, File, Java, ByteCode, Class]	861	0.261674
17	[Type, File, Java, ByteCode, JavaType]	746	0.226724
18	[Java, Value, ByteCode, Primitive]	650	0.197547
19	[Java,ByteCode,GenericDeclaration,Member,M	569	0.172930
20	[Json, Key]	555	0.168675
21	[Value, Json, Scalar]	539	0.163812
22	[Type, File, Java, ByteCode, ExternalType]	396	0.120352
23	[Author, Git, Person]	290	0.088137
24	[Value, Array]	267	0.081146
25	[Committer, Git, Person]	244	0.074156
26	[Type, File, Java, ByteCode, Class, GenericDec	236	0.071725
27	[Value, Property]	202	0.061392
28	[Type, File, Java, ByteCode, Interface]	187	0.056833
29	[Java, Value, ByteCode, Class]	172	0.052274

# 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.

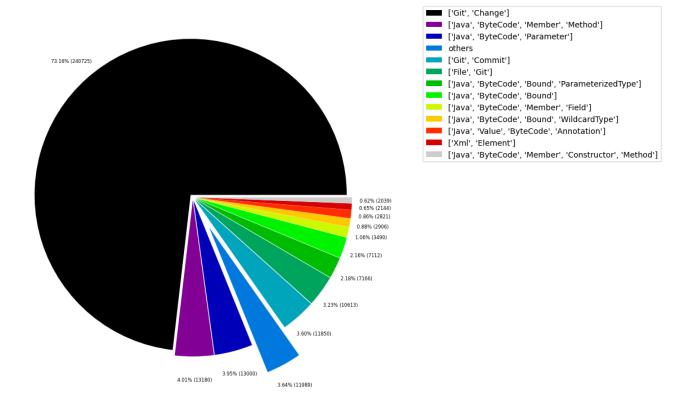


Table 1b - Lowest node count by label combination

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	0.000304
1	[File, TS, Scan]	1	0.000304
2	[File, Json]	2	0.000608
3	[File]	3	0.000912
4	[Java,ByteCode,GenericDeclaration,Member,C	4	0.001216
5	[Maven, Exclusion]	5	0.001520
6	[Value, Array, Json]	6	0.001824
7	[Type, File, Java, ByteCode, Void]	9	0.002735
8	[File, Maven, Xml, Pom, Document]	9	0.002735
9	[Java, ManifestSection]	9	0.002735
10	[File, Java, Manifest]	9	0.002735
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002735
12	[File, Java, ServiceLoader]	10	0.003039
13	[File, Java, Properties]	12	0.003647
14	[Maven, ExecutionGoal]	16	0.004863
15	[Maven, PluginExecution]	16	0.004863
16	[Xml, Attribute]	18	0.005471
17	[jQAssistant, Rule, Concept]	19	0.005774
18	[Maven, Plugin]	21	0.006382
19	[Maven, Configuration]	21	0.006382
20	[Git, Branch]	28	0.008510
21	[Type, File, Java, ByteCode, Enum]	28	0.008510
22	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008814
23	[Xml, Namespace]	36	0.010941
24	[Type, File, Java, ByteCode, Annotation]	43	0.013069
25	[File, Directory]	44	0.013372
26	[Type, File, Java, ByteCode, GenericDeclaratio	85	0.025833
27	[Artifact, Maven]	92	0.027961
28	[Type, File, Java, ByteCode, ExternalAnnotatio	93	0.028264
29	[Value, ManifestEntry]	120	0.036470

### 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.

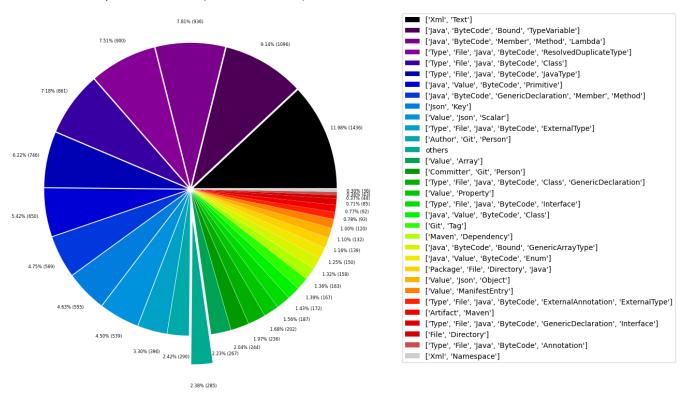


Table 1c - Highest node count by single label

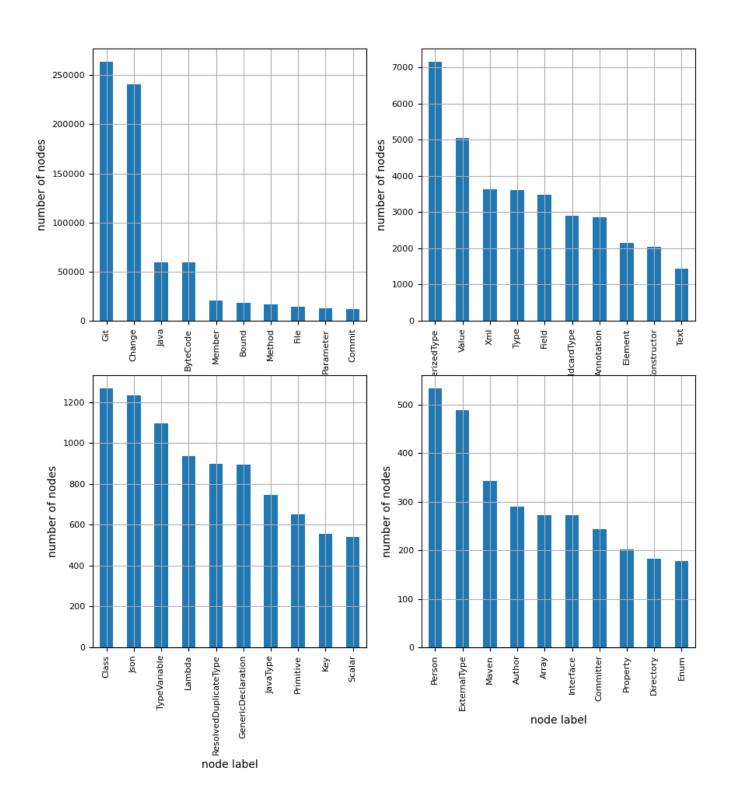
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodes With That Label	nodes With That Label Percent
0	Git	263917	80.209400
1	Change	240725	73.160910
2	Java	59250	18.007203
3	ByteCode	59062	17.950066
4	Member	20218	6.144635
5	Bound	18438	5.603659
6	Method	16728	5.083958
7	File	14464	4.395885
8	Parameter	13000	3.950947
9	Commit	11850	3.601441
10	ParameterizedType	7166	2.177884
11	Value	5059	1.537526
12	Xml	3643	1.107177
13	Туре	3613	1.098059
14	Field	3490	1.060677
15	WildcardType	2906	0.883189
16	Annotation	2864	0.870424
17	Element	2144	0.651602
18	Constructor	2043	0.620907
19	Text	1436	0.436428
20	Class	1269	0.385673
21	Json	1234	0.375036
22	TypeVariable	1096	0.333095
23	Lambda	936	0.284468
24	ResolvedDuplicateType	900	0.273527
25	GenericDeclaration	894	0.271704
26	JavaType	746	0.226724
27	Primitive	650	0.197547
28	Key	555	0.168675
29	Scalar	539	0.163812
30	Person	534	0.162293
31	ExternalType	489	0.148616
32	Maven	343	0.104244
33	Author	290	0.088137
34	Array	273	0.082970
35	Interface	272	0.082666
36	Committer	244	0.074156
37	Property	202	0.061392
38	Directory	183	0.055617
39	Enum	178	0.054098

# 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: 997713

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS CHANGE	240725	24.127680
1	MODIFIES	240725	24.127680
2	UPDATES	155355	15.571111
3	CREATES	58825	5.895984
4	DELETES	36675	3.675907
5	INVOKES	35593	3.567459
6	COMMITTED	23700	2.375433
7	DEPENDS_ON	21873	2.192314
8	OF_TYPE	21347	2.139593
9	DECLARES	20659	2.070636
10	OF_RAW_TYPE	17077	1.711614
11	HAS_PARENT	14409	1.444203
12	HAS	14094	1.412631
13	RETURNS	12547	1.257576
14	RENAMES	10130	1.015322
15	READS	9078	0.909881
16	HAS_ACTUAL_TYPE_ARGUMENT	8276	0.829497
17	HAS_NEW_NAME	6177	0.619116
18	OF_GENERIC_TYPE	5902	0.591553
19	RESOLVES_TO	5228	0.523998
20	SIMILAR	3968	0.397710
21	CONTAINS	3805	0.381372
22	WRITES	3803	0.381172
23	RETURNS_GENERIC	3543	0.355112
24	ANNOTATED_BY	2809	0.281544
25	REQUIRES	2173	0.217798
26	HAS_FIRST_CHILD	2144	0.214891
27	HAS_LAST_CHILD	2144	0.214891
28	HAS_ELEMENT	2126	0.213087
29	HAS_UPPER_BOUND	1521	0.152449

### 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.

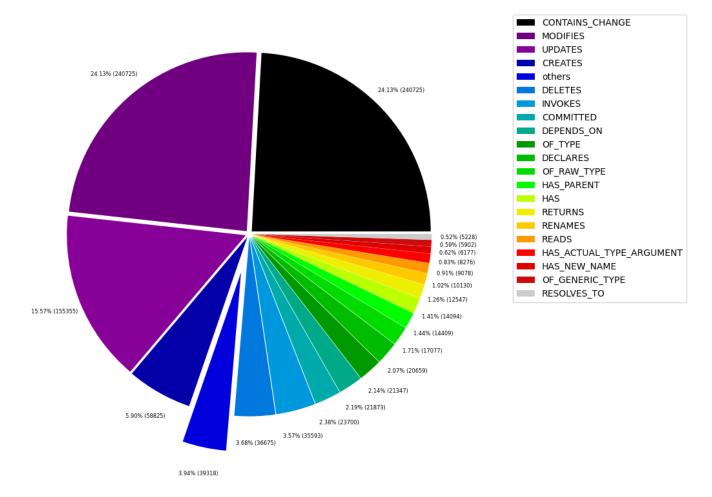


Table 2b - Lowest relationship count by type

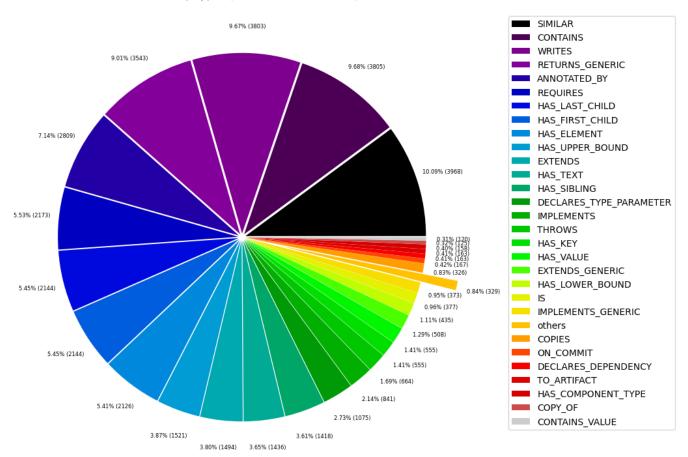
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodes With That Relationship Type	nodes With That Relationship Type Percent
0	HAS_PROPERTY	1	0.000100
1	THROWS_GENERIC	5	0.000501
2	EXCLUDES	5	0.000501
3	DESCRIBES	9	0.000902
4	HAS_ROOT_ELEMENT	10	0.001002
5	HAS_GOAL	16	0.001604
6	HAS_EXECUTION	16	0.001604
7	OF_NAMESPACE	18	0.001804
8	HAS_ATTRIBUTE	18	0.001804
9	INCLUDES_CONCEPT	19	0.001904
10	USES_PLUGIN	21	0.002105
11	REQUIRES_TYPE_PARAMETER	21	0.002105
12	IS_ARTIFACT	21	0.002105
13	HAS_CONFIGURATION	21	0.002105
14	REQUIRES_CONCEPT	28	0.002806
15	HAS_HEAD	28	0.002806
16	DECLARES_NAMESPACE	36	0.003608
17	HAS_DEFAULT	36	0.003608
18	CONTAINS_VALUE	120	0.012028
19	COPY_OF	125	0.012529
20	HAS_COMPONENT_TYPE	158	0.015836
21	TO_ARTIFACT	163	0.016337
22	DECLARES_DEPENDENCY	163	0.016337
23	ON_COMMIT	167	0.016738
24	COPIES	326	0.032675
25	IMPLEMENTS_GENERIC	373	0.037386
26	IS	377	0.037786
27	HAS_LOWER_BOUND	435	0.043600
28	EXTENDS_GENERIC	508	0.050916
29	HAS_VALUE	555	0.055627

### 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]	240725	11850	
1	[Git, Change]	MODIFIES	[File, Git]	240725	240725	
2	[Git, Change]	UPDATES	[File, Git]	155355	240725	
3	[Git, Change]	CREATES	[File, Git]	58825	240725	
4	[Git, Change]	DELETES	[File, Git]	36675	240725	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21578	13066	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	14400	11850	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	11850	290	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	11850	244	
9	[Git, Change]	RENAMES	[File, Git]	10130	240725	
10	[Java, ByteCode, Member, Method]	HAS	[Java, ByteCode, Parameter]	8316	13066	
11	[Java, ByteCode, Member, Method]	READS	[Java, ByteCode, Member, Field]	8138	13066	
12	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6360	13000	
13	[File, Git]	HAS_NEW_NAME	[File, Git]	6177	10613	
14	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3569	7112	
15	[Java, ByteCode, Bound, ParameterizedType]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3043	7166	
16	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, WildcardType]	2906	7166	
17	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound, ParameterizedType]	2655	13000	
18	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, TypeVariable]	2454	7166	
19	[Java, ByteCode, Member, Constructor, Method]	WRITES	[Java, ByteCode, Member, Field]	2426	2039	
20	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalAnnotatio	2340	2821	
21	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
22	[Java, ByteCode, Member, Constructor, Method]	INVOKES	[Java, ByteCode, Member, Constructor, Method]	2111	2039	
23	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2087	13066	
24	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Constructor, Method]	2032	13066	
25	[Java, ByteCode, Member, Constructor, Method]	HAS	[Java, ByteCode, Parameter]	2023	2039	
26	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound]	2023	13000	
27	[Java, ByteCode, Member, Method, Lambda]	INVOKES	[Java, ByteCode, Member, Method]	1882	936	
28	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound]	1802	7166	
29	[Java, ByteCode, Member, Constructor, Method]	INVOKES	[Java, ByteCode, Member, Method]	1797	2039	

# **Graph Density**

total\_number\_of\_nodes (vertices): 329035
total\_number\_of\_relationships (edges): 997713

- -> total directed graph density: 9.215581767645813e-06
- -> total directed graph density in percent: 0.0009215581767645812