#### 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 .

Total number of nodes: 321900

	nodeLabels	nodesWithThatLabels	nodesWithThatLabelsPercent
0	[Git, Change]	233759	72.618515
1	[Java, ByteCode, Member, Method]	13195	4.099099
2	[Java, ByteCode, Parameter]	12968	4.028580
3	[Git, Commit]	11776	3.658279
4	[File, Git]	10597	3.292016
5	[Java, ByteCode, Bound, ParameterizedType]	7154	2.222429
6	[Java, ByteCode, Bound]	7094	2.203790
7	[Java, ByteCode, Member, Field]	3490	1.084188
8	[Java, ByteCode, Bound, WildcardType]	2902	0.901522
9	[Java, Value, ByteCode, Annotation]	2813	0.873874
10	[Xml, Element]	2144	0.666045
11	[Java, ByteCode, Member, Constructor, Method]	2044	0.634980
12	[Xml, Text]	1436	0.446101
13	[Java, ByteCode, Bound, TypeVariable]	1094	0.339857
14	[Java, ByteCode, Member, Method, Lambda]	931	0.289220
15	[Type, File, Java, ByteCode, ResolvedDuplicate	899	0.279279
16	[Type, File, Java, Class, ByteCode]	858	0.266542
17	[Type, File, Java, ByteCode, JavaType]	746	0.231749
18	[Java, Value, ByteCode, Primitive]	647	0.200994
19	[Java, ByteCode, Member, Method, GenericDeclar	567	0.176142
20	[Json, Key]	555	0.172414
21	[Value, Json, Scalar]	539	0.167443
22	[Type, File, Java, ByteCode, ExternalType]	396	0.123020
23	[Author, Git, Person]	289	0.089779
24	[Value, Array]	267	0.082945
25	[Committer, Git, Person]	243	0.075489
26	[Type, File, Java, Class, ByteCode, GenericDec	236	0.073315
27	[Value, Property]	202	0.062752
28	[Type, File, Java, ByteCode, Interface]	186	0.057782
29	[Java, Value, Class, ByteCode]	170	0.052811

## 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.

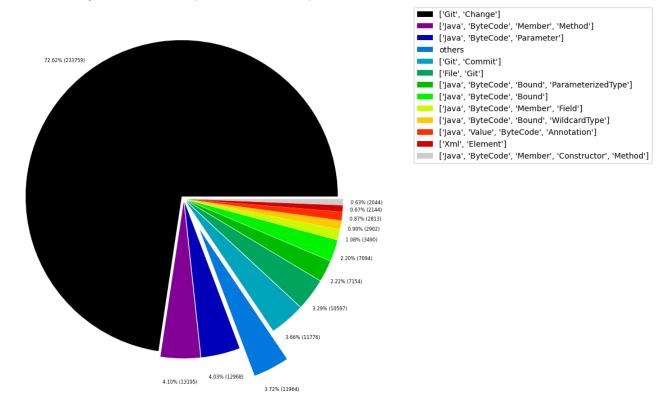


Table 1b - Lowest node count by label combination

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	0.000311
1	[File, TS, Scan]	1	0.000311
2	[File, Json]	2	0.000621
3	[File]	3	0.000932
4	[Java, ByteCode, Member, Constructor, Method, $\dots$	4	0.001243
5	[Maven, Exclusion]	5	0.001553
6	[Value, Array, Json]	6	0.001864
7	[Type, File, Java, ByteCode, Void]	9	0.002796
8	[File, Maven, Xml, Pom, Document]	9	0.002796
9	[Java, ManifestSection]	9	0.002796
10	[File, Java, Manifest]	9	0.002796
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002796
12	[File, Java, ServiceLoader]	10	0.003107
13	[File, Java, Properties]	12	0.003728
14	[Maven, ExecutionGoal]	16	0.004970
15	[Maven, PluginExecution]	16	0.004970
16	[Xml, Attribute]	18	0.005592
17	[jQAssistant, Rule, Concept]	19	0.005902
18	[Maven, Plugin]	21	0.006524
19	[Maven, Configuration]	21	0.006524
20	[Git, Branch]	24	0.007456
21	[Type, File, Java, ByteCode, Enum]	28	0.008698
22	[Type, File, Java, ByteCode, PrimitiveType]	29	0.009009
23	[Xml, Namespace]	36	0.011184
24	[Type, File, Java, ByteCode, Annotation]	43	0.013358
25	[File, Directory]	44	0.013669
26	[Type, File, Java, ByteCode, GenericDeclaratio	85	0.026406
27	[Artifact, Maven]	92	0.028580
28	[Type, File, Java, ByteCode, ExternalAnnotatio	93	0.028891
29	[Value, ManifestEntry]	120	0.037279

#### 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.

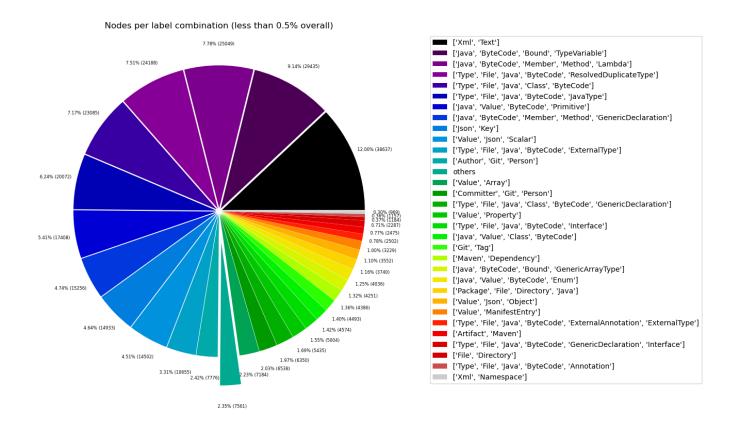


Table 1c - Highest node count by single label

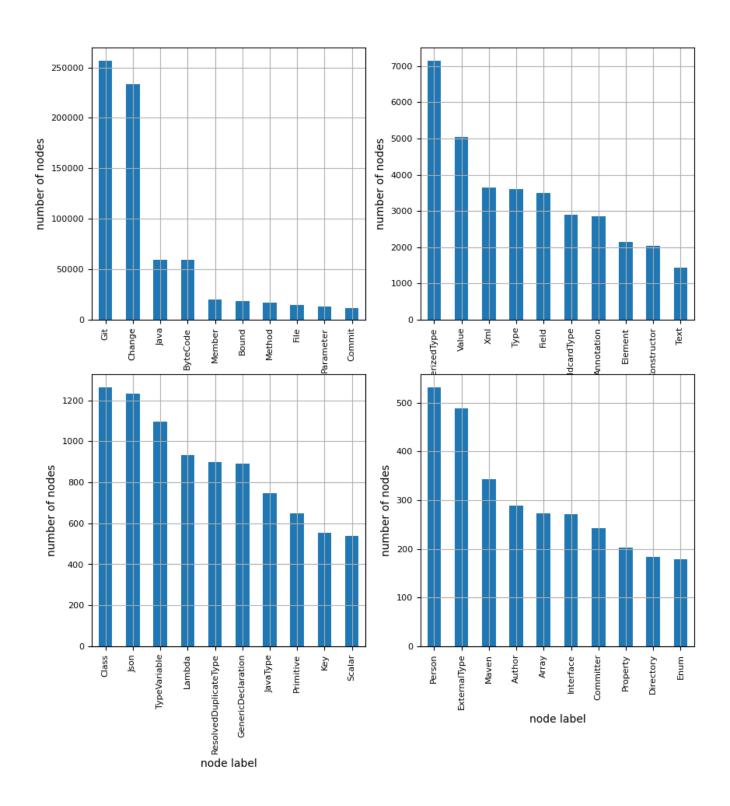
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodes With That Label	nodes With That Label Percent
0	Git	256855	79.793414
1	Change	233759	72.618515
2	Java	59177	18.383660
3	ByteCode	58989	18.325256
4	Member	20231	6.284871
5	Bound	18402	5.716682
6	Method	16741	5.200683
7	File	14443	4.486797
8	Parameter	12968	4.028580
9	Commit	11776	3.658279
10	ParameterizedType	7154	2.222429
11	Value	5046	1.567568
12	Xml	3643	1.131718
13	Туре	3608	1.120845
14	Field	3490	1.084188
15	WildcardType	2902	0.901522
16	Annotation	2856	0.887232
17	Element	2144	0.666045
18	Constructor	2048	0.636222
19	Text	1436	0.446101
20	Class	1264	0.392669
21	Json	1234	0.383349
22	TypeVariable	1094	0.339857
23	Lambda	931	0.289220
24	ResolvedDuplicateType	899	0.279279
25	GenericDeclaration	892	0.277105
26	JavaType	746	0.231749
27	Primitive	647	0.200994
28	Key	555	0.172414
29	Scalar	539	0.167443
30	Person	532	0.165269
31	ExternalType	489	0.151911
32	Maven	343	0.106555
33	Author	289	0.089779
34	Array	273	0.084809
35	Interface	271	0.084188
36	Committer	243	0.075489
37	Property	202	0.062752
38	Directory	183	0.056850
39	Enum	178	0.055297

## 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: 975767

	relationshipType	nodes With That Relationship Type	nodes With That Relationship Type Percent
0	CONTAINS_CHANGE	233759	23.956436
1	MODIFIES	233759	23.956436
2	UPDATES	150078	15.380516
3	CREATES	57924	5.936253
4	DELETES	35673	3.655893
5	INVOKES	35515	3.639701
6	COMMITTED	23552	2.413691
7	DEPENDS_ON	21820	2.236190
8	OF_TYPE	21292	2.182078
9	DECLARES	20670	2.118334
10	OF_RAW_TYPE	17041	1.746421
11	HAS_PARENT	14310	1.466539
12	HAS	14057	1.440610
13	RETURNS	12524	1.283503
14	RENAMES	9916	1.016226
15	READS	9069	0.929423
16	HAS_ACTUAL_TYPE_ARGUMENT	8261	0.846616
17	HAS_NEW_NAME	6165	0.631811
18	OF_GENERIC_TYPE	5892	0.603833
19	RESOLVES_TO	5238	0.536808
20	SIMILAR	3955	0.405322
21	CONTAINS	3797	0.389130
22	WRITES	3796	0.389027
23	RETURNS_GENERIC	3538	0.362587
24	ANNOTATED_BY	2801	0.287056
25	REQUIRES	2172	0.222594
26	HAS_FIRST_CHILD	2144	0.219725
27	HAS_LAST_CHILD	2144	0.219725
28	HAS_ELEMENT	2126	0.217880
29	HAS_UPPER_BOUND	1519	0.155672

#### 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.

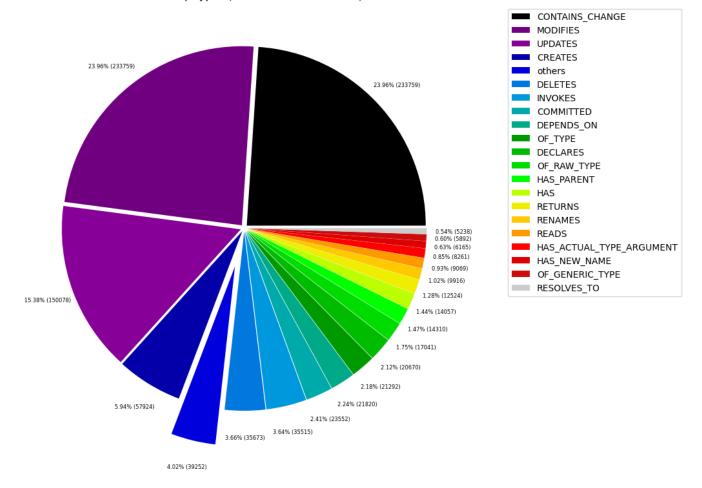


Table 2b - Lowest relationship count by type

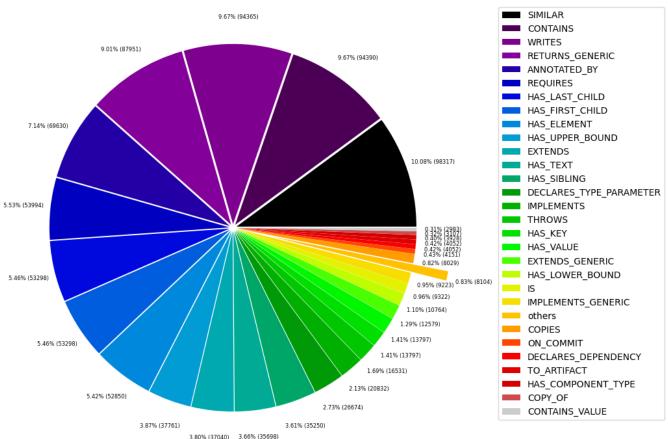
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	$nodes With That Relationship Type \\ nodes With That Relationship Type Per Company \\ nodes With That Relationship Type \\ nodes With That $	
0	HAS_PROPERTY	1	0.000102
1	THROWS_GENERIC	5	0.000512
2	EXCLUDES	5	0.000512
3	DESCRIBES	9	0.000922
4	HAS_ROOT_ELEMENT	11	0.001127
5	HAS_GOAL	16	0.001640
6	HAS_EXECUTION	16	0.001640
7	OF_NAMESPACE	18	0.001845
8	HAS_ATTRIBUTE	18	0.001845
9	INCLUDES_CONCEPT	19	0.001947
10	USES_PLUGIN	21	0.002152
11	REQUIRES_TYPE_PARAMETER	21	0.002152
12	IS_ARTIFACT	21	0.002152
13	HAS_CONFIGURATION	21	0.002152
14	HAS_HEAD	24	0.002460
15	REQUIRES_CONCEPT	28	0.002870
16	DECLARES_NAMESPACE	36	0.003689
17	HAS_DEFAULT	36	0.003689
18	CONTAINS_VALUE	120	0.012298
19	COPY_OF	125	0.012810
20	HAS_COMPONENT_TYPE	158	0.016192
21	TO_ARTIFACT	163	0.016705
22	DECLARES_DEPENDENCY	163	0.016705
23	ON_COMMIT	167	0.017115
24	COPIES	323	0.033102
25	IMPLEMENTS_GENERIC	371	0.038021
26	IS	375	0.038431
27	HAS_LOWER_BOUND	433	0.044375
28	EXTENDS_GENERIC	506	0.051857
29	HAS_VALUE	555	0.056878

#### 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.





#### 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	number Of Relation ships	number Of Nodes With Same Labels As Source	numberOfNodesWithS
0	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	233759	11776	
1	[Git, Change]	MODIFIES	[File, Git]	233759	233759	
2	[Git, Change]	UPDATES	[File, Git]	150078	233759	
3	[Git, Change]	CREATES	[File, Git]	57924	233759	
4	[Git, Change]	DELETES	[File, Git]	35673	233759	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21579	13086	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	14301	11776	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	11776	289	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	11776	243	
9	[Git, Change]	RENAMES	[File, Git]	9916	233759	
10	[Java, ByteCode, Member, Method]	HAS	[Java, ByteCode, Parameter]	8298	13086	
11	[Java, ByteCode, Member, Method]	READS	[Java, ByteCode, Member, Field]	8146	13086	
12	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6335	12968	
13	[File, Git]	HAS_NEW_NAME	[File, Git]	6165	10597	
14	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3554	7094	
15	[Java, ByteCode, Bound, ParameterizedType]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3034	7154	
16	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, WildcardType]	2902	7154	
17	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound, ParameterizedType]	2650	12968	
18	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, TypeVariable]	2453	7154	
19	[Java, ByteCode, Member, Constructor, Method]	WRITES	[Java, ByteCode, Member, Field]	2423	2044	
20	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalAnnotatio	2336	2813	
21	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
22	[Java, ByteCode, Member, Constructor, Method]	INVOKES	[Java, ByteCode, Member, Constructor, Method]	2109	2044	
23	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2087	13086	
24	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Constructor, Method]	2033	13086	
25	[Java, ByteCode, Member, Constructor, Method]	HAS	[Java, ByteCode, Parameter]	2021	2044	
26	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound]	2020	12968	
27	[Java, ByteCode, Member, Method, Lambda]	INVOKES	[Java, ByteCode, Member, Method]	1864	931	
28	[Java, ByteCode, Member, Constructor, Method]	INVOKES	[Java, ByteCode, Member, Method]	1793	2044	
29	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound]	1792	7154	

## **Graph Density**

total\_number\_of\_nodes (vertices): 321900
total\_number\_of\_relationships (edges): 975767

-> total directed graph density: 9.416847170946739e-06

-> total directed graph density in percent: 0.0009416847170946739