# 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	nodes With That Labels	nodesWithThatLabelsPercent
0	[Git, Change]	83039	78.412653
1	[Git, Commit]	10742	10.143532
2	[File, Git]	5534	5.225685
3	[Git, Tag]	1384	1.306893
4	[Author, Git, Person]	1234	1.165250
5	[Json, Key]	668	0.630784
6	[Json, Value, Scalar]	603	0.569405
7	[Committer, Git, Person]	371	0.350331
8	[NPM, Dependency]	338	0.319169
9	[Type, TS, Primitive]	291	0.274788
10	[Type, TS, Declared]	276	0.260623
11	[TS, ExternalDeclaration]	214	0.202077
12	[Type, TS, Literal]	136	0.128423
13	[Json, Value, Object]	133	0.125590
14	[Type, TS, Union]	119	0.112370
15	[Type, TS, ObjectMember]	101	0.095373
16	[NPM, Script]	91	0.085930
17	[TS, Property]	65	0.061379
18	[TS, Function]	47	0.044381
19	[Type, TS, FunctionParameter]	40	0.037771
20	[Type, Object, TS]	39	0.036827
21	[Git, Branch]	39	0.036827
22	[File, Directory]	34	0.032106
23	[Type, TS, Function]	34	0.032106
24	[TS, Parameter]	33	0.031161
25	[Package, File, Json, NPM]	29	0.027384
26	[TS, Variable]	24	0.022663
27	[TS, ExternalModule]	23	0.021719
28	[Value, TS, Literal]	20	0.018886
29	[jQAssistant, Rule, Concept]	19	0.017941

# 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.

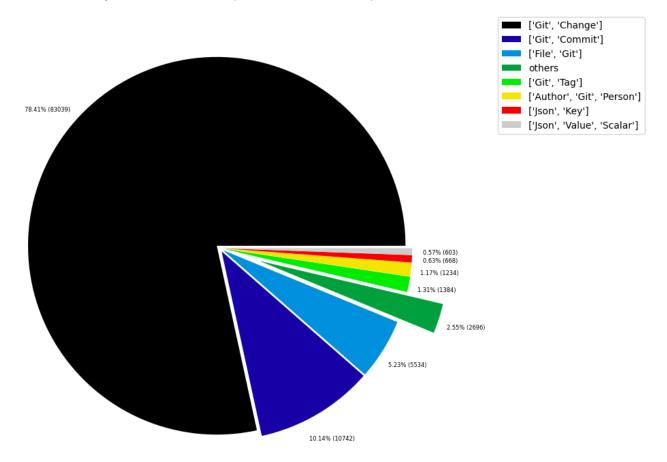


Table 1b - Lowest node count by label combination

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	nodes With That Labels	nodes With That Labels Percent
0	[Analyze, Task, jQAssistant]	1	0.000944
1	[File, TS, Scan]	1	0.000944
2	[TS, Method]	1	0.000944
3	[Repository, File, Git]	1	0.000944
4	[TS, Constructor]	1	0.000944
5	[Value, TS, ObjectMember]	1	0.000944
6	[TS, Class]	1	0.000944
7	[TS, Enum]	2	0.001889
8	[Value, Object, TS]	3	0.002833
9	[Type, TS, Tuple]	3	0.002833
10	[Value, TS, Function]	4	0.003777
11	[TS, TypeParameter]	4	0.003777
12	[Value, TS, Complex]	5	0.004721
13	[NPM, Engine]	6	0.005666
14	[Project, TS]	6	0.005666
15	[File, Local]	6	0.005666
16	[Value, TS, Call]	6	0.005666
17	[Value, TS, Member]	6	0.005666
18	[File, TS, Local, Module]	6	0.005666
19	[Type, TS, TypeParameterReference]	6	0.005666
20	[TS, EnumMember]	8	0.007554
21	[Type, TS, NotIdentified]	11	0.010387
22	[Json, Value, Array]	12	0.011331
23	[Value, TS, Declared]	13	0.012276
24	[TS, TypeAlias]	16	0.015109
25	[File, Directory, Local]	16	0.015109
26	[TS, Interface]	17	0.016053
27	[Type, TS, Intersection]	17	0.016053
28	[jQAssistant, Rule, Concept]	19	0.017941
29	[Value, TS, Literal]	20	0.018886

## 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.

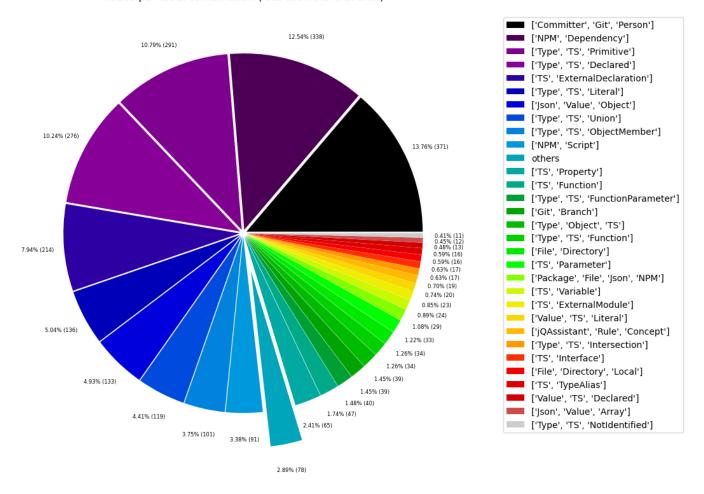


Table 1c - Highest node count by single label

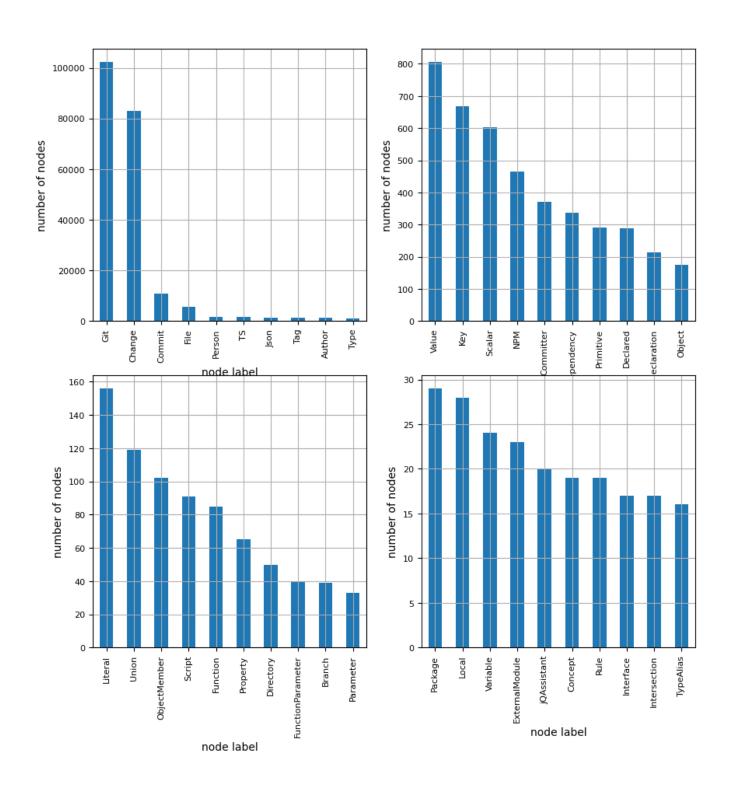
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodes With That Label	nodesWithThatLabelPercent	
0	Git	102344	96.642115	
1	Change	83039	78.412653	
2	Commit	10742	10.143532	
3	File	5627	5.313503	
4	Person	1605	1.515581	
5	TS	1600	1.510859	
6	Json	1445	1.364495	
7	Tag	1384	1.306893	
8	Author	1234	1.165250	
9	Туре	1073	1.013220	
10	Value	806	0.761095	
11	Key	668	0.630784	
12	Scalar	603	0.569405	
13	NPM	464	0.438149	
14	Committer	371	0.350331	
15	Dependency	338	0.319169	
16	Primitive	291	0.274788	
17	Declared	289	0.272899	
18	ExternalDeclaration	214	0.202077	
19	Object	175	0.165250	
20	Literal	156	0.147309	
21	Union	119	0.112370	
22	ObjectMember	102	0.096317	
23	Script	91	0.085930	
24	Function	85	0.080264	
25	Property	65	0.061379	
26	Directory	50	0.047214	
27	FunctionParameter	40	0.037771	
28	Branch	39	0.036827	
29	Parameter	33	0.031161	
30	Package	29	0.027384	
31	Local	28	0.026440	
32	Variable	24	0.022663	
33	ExternalModule	23	0.021719	
34	jQAssistant	20	0.018886	
35	Concept	19	0.017941	
36	Rule	19	0.017941	
37	Interface	17	0.016053	
38	Intersection	17	0.016053	
39	TypeAlias	16	0.015109	

# 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: 316451

	relationshipType	nodes With That Relationship Type	nodes With That Relationship Type Percent
0	CONTAINS_CHANGE	83039	26.240713
1	MODIFIES	83039	26.240713
2	UPDATES	54366	17.179911
3	COMMITTED	21484	6.789045
4	CREATES	19954	6.305558
5	DELETES	11969	3.782260
6	HAS_PARENT	11796	3.727591
7	HAS_COMMIT	10742	3.394522
8	HAS_FILE	5534	1.748770
9	RENAMES	3250	1.027015
10	HAS_NEW_NAME	1729	0.546372
11	HAS_TAG	1384	0.437350
12	ON_COMMIT	1384	0.437350
13	HAS_AUTHOR	1234	0.389950
14	DEPENDS_ON	961	0.303681
15	HAS_KEY	668	0.211091
16	HAS_VALUE	668	0.211091
17	CONTAINS	594	0.187707
18	HAS_COMMITTER	371	0.117238
19	OF_TYPE	337	0.106494
20	EXPORTS	283	0.089429
21	REFERENCES	197	0.062253
22	DECLARES	186	0.058777
23	DECLARES_DEV_DEPENDENCY	169	0.053405
24	DECLARES_DEPENDENCY	161	0.050877
25	HAS_MEMBER	102	0.032232
26	HAS_TYPE_ARGUMENT	94	0.029704
27	DECLARES_SCRIPT	91	0.028756
28	RETURNS	82	0.025912
29	COPIES	77	0.024332

# 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.

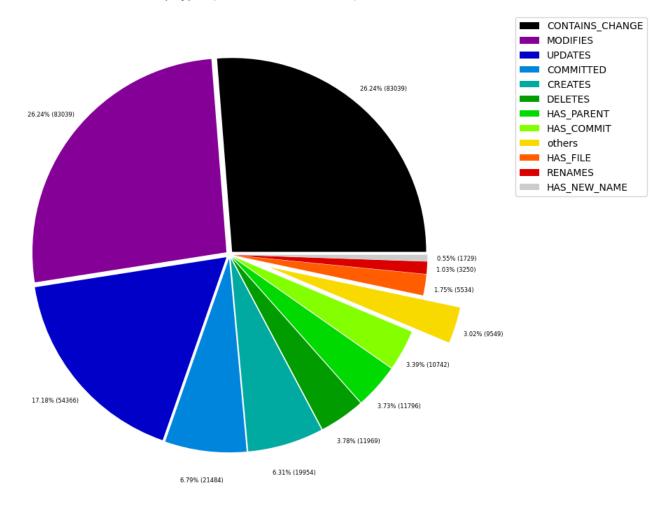


Table 2b - Lowest relationship count by type

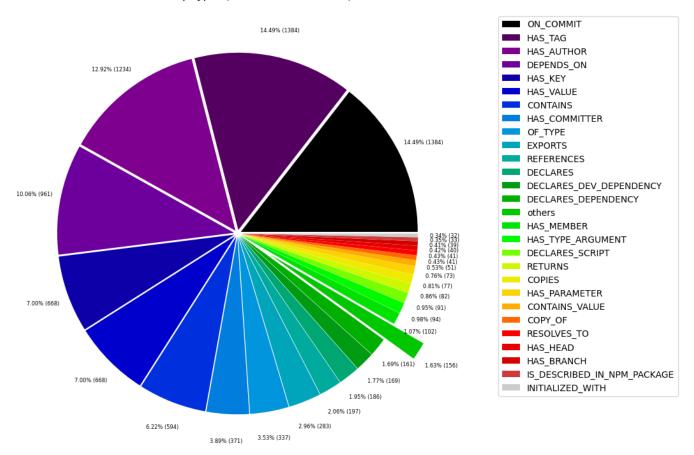
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent	
0	PROVIDED_BY_NPM_DEPENDENCY	1	0.000316	
1	IS_IMPLEMENTED_IN	2	0.000632	
2	CONSTRAINED_BY	4	0.001264	
3	REFERENCED_PROJECTS	5	0.001580	
4	CONTAINS_PROJECT	6	0.001896	
5	DECLARES_ENGINE	6	0.001896	
6	EXTENDS	6	0.001896	
7	HAS_ARGUMENT	6	0.001896	
8	CALLS	6	0.001896	
9	HAS_NPM_PACKAGE	6	0.001896	
10	HAS_ROOT	6	0.001896	
11	MEMBER	6	0.001896	
12	PARENT	6	0.001896	
13	HAS_CONFIG	6	0.001896	
14	SIMILAR	6	0.001896	
15	DECLARES_PEER_DEPENDENCY	8	0.002528	
16	INCLUDES_CONCEPT	19	0.006004	
17	USES	23	0.007268	
18	REQUIRES_CONCEPT	28	0.008848	
19	INITIALIZED_WITH	32	0.010112	
20	IS_DESCRIBED_IN_NPM_PACKAGE	33	0.010428	
21	HAS_BRANCH	39	0.012324	
22	HAS_HEAD	40	0.012640	
23	RESOLVES_TO	41	0.012956	
24	COPY_OF	41	0.012956	
25	CONTAINS_VALUE	51	0.016116	
26	HAS_PARAMETER	73	0.023068	
27	COPIES	77	0.024332	
28	RETURNS	82	0.025912	
29	DECLARES_SCRIPT	91	0.028756	

# 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	number Of Relationships	number Of Nodes With Same Labels As Source	numberOfNodes
0	[Git, Change]	MODIFIES	[File, Git]	83039	83039	
1	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	83039	10742	
2	[Git, Change]	UPDATES	[File, Git]	54366	83039	
3	[Git, Change]	CREATES	[File, Git]	19954	83039	
4	[Git, Change]	DELETES	[File, Git]	11969	83039	
5	[Git, Commit]	HAS_PARENT	[Git, Commit]	11796	10742	
6	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	10742	1	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	10742	1234	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10742	371	
9	[Repository, File, Git]	HAS_FILE	[File, Git]	5534	1	
10	[Git, Change]	RENAMES	[File, Git]	3250	83039	
11	[File, Git]	HAS_NEW_NAME	[File, Git]	1729	5534	
12	[Repository, File, Git]	HAS_TAG	[Git, Tag]	1384	1	
13	[Git, Tag]	ON_COMMIT	[Git, Commit]	1384	1384	
14	[Repository, File, Git]	HAS_AUTHOR	[Author, Git, Person]	1234	1	
15	[Json, Value, Object]	HAS_KEY	[Json, Key]	668	133	
16	[Json, Key]	HAS_VALUE	[Json, Value, Scalar]	552	668	
17	[Repository, File, Git]	HAS_COMMITTER	[Committer, Git, Person]	371	1	
18	[TS, Function]	DEPENDS_ON	[TS, ExternalDeclaration]	289	47	
19	[File, TS, Local, Module, Mark4ModuleWeaklyCon	DEPENDS_ON	[TS, ExternalDeclaration]	232	4	
20	[TS, ExternalModule]	EXPORTS	[TS, ExternalDeclaration]	214	23	
21	[Package, File, Json, NPM]	DECLARES_DEV_DEPENDENCY	[NPM, Dependency]	169	29	
22	[Package, File, Json, NPM]	DECLARES_DEPENDENCY	[NPM, Dependency]	161	29	
23	[Type, TS, Union]	CONTAINS	[Type, TS, Primitive]	147	119	
24	[Type, TS, Declared]	REFERENCES	[TS, ExternalDeclaration]	142	276	
25	[TS, Function]	DEPENDS_ON	[TS, ExternalModule]	132	47	
26	[Type, TS, Union]	CONTAINS	[Type, TS, Literal]	119	119	
27	[Json, Key]	HAS_VALUE	[Json, Value, Object]	104	668	
28	[Type, Object, TS]	HAS_MEMBER	[Type, TS, ObjectMember]	101	39	
29	[Package, File, Json, NPM]	DECLARES_SCRIPT	[NPM, Script]	91	29	

# **Graph Density**

total\_number\_of\_nodes (vertices): 105900
total\_number\_of\_relationships (edges): 316451

-> total directed graph density: 2.8217507762866432e-05

-> total directed graph density in percent: 0.002821750776286643