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]	77924	77.994195
1	[Git, Commit]	10305	10.314283
2	[File, Git]	5340	5.344810
3	[Git, Tag]	1197	1.198078
4	[Author, Git, Person]	1192	1.193074
5	[Json, Key]	668	0.668602
6	[Json, Value, Scalar]	603	0.603543
7	[Committer, Git, Person]	371	0.371334
8	[NPM, Dependency]	330	0.330297
9	[Type, TS, Primitive]	291	0.291262
10	[Type, TS, Declared]	276	0.276249
11	[TS, ExternalDeclaration]	215	0.215194
12	[Type, TS, Literal]	136	0.136123
13	[Json, Value, Object]	133	0.133120
14	[Type, TS, Union]	119	0.119107
15	[Type, TS, ObjectMember]	101	0.101091
16	[NPM, Script]	91	0.091082
17	[TS, Property]	65	0.065059
18	[TS, Function]	47	0.047042
19	[Type, TS, FunctionParameter]	40	0.040036
20	[Type, Object, TS]	39	0.039035
21	[File, Directory]	34	0.034031
22	[Type, TS, Function]	34	0.034031
23	[TS, Parameter]	33	0.033030
24	[Git, Branch]	30	0.030027
25	[Package, File, Json, NPM]	29	0.029026
26	[TS, ExternalModule]	25	0.025023
27	[TS, Variable]	24	0.024022
28	[Value, TS, Literal]	20	0.020018
29	[jQAssistant, Rule, Concept]	19	0.019017

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.

```
<Figure size 640x480 with 0 Axes>
```

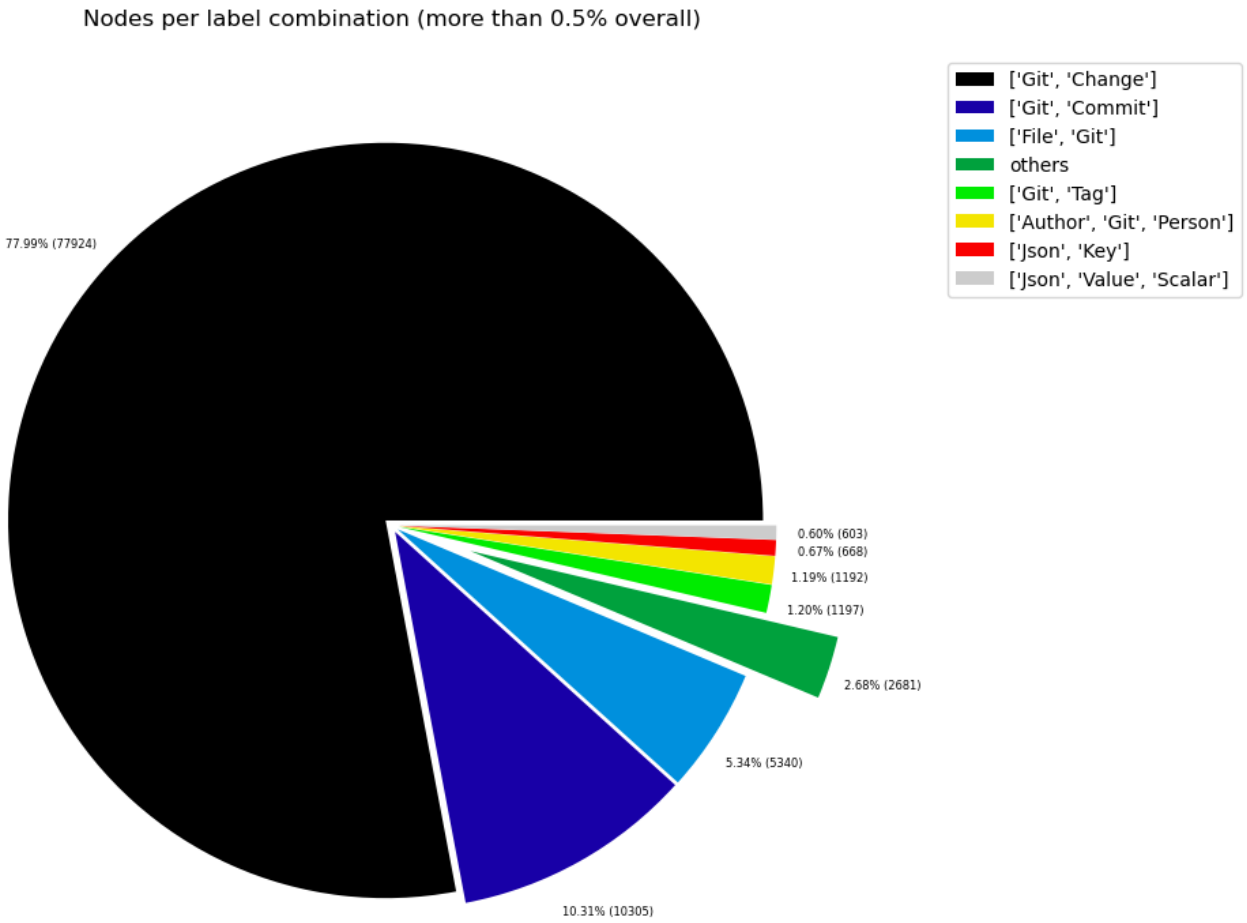


Table 1b - Lowest node count by label combination

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	0.001001
1	[File, TS, Scan]	1	0.001001
2	[TS, Method]	1	0.001001
3	[Value, TS, ObjectMember]	1	0.001001
4	[TS, Constructor]	1	0.001001
5	[TS, Class]	1	0.001001
6	[TS, Enum]	2	0.002002
7	[Value, Object, TS]	3	0.003003
8	[Type, TS, Tuple]	3	0.003003
9	[Value, TS, Function]	4	0.004004
10	[TS, TypeParameter]	4	0.004004
11	[Value, TS, Complex]	5	0.005005
12	[NPM, Engine]	6	0.006005
13	[Project, TS]	6	0.006005
14	[File, Local]	6	0.006005
15	[Value, TS, Call]	6	0.006005
16	[Value, TS, Member]	6	0.006005
17	[File, TS, Local, Module]	6	0.006005
18	[Type, TS, TypeParameterReference]	6	0.006005
19	[TS, EnumMember]	8	0.008007
20	[Type, TS, NotIdentified]	11	0.011010
21	[Json, Value, Array]	12	0.012011
22	[Value, TS, Declared]	13	0.013012
23	[TS, TypeAlias]	16	0.016014
24	[File, Directory, Local]	16	0.016014
25	[TS, Interface]	17	0.017015
26	[Type, TS, Intersection]	17	0.017015
27	[jQAssistant, Rule, Concept]	19	0.019017
28	[Value, TS, Literal]	20	0.020018
29	[TS, Variable]	24	0.024022

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.

```
<Figure size 640x480 with 0 Axes>
```

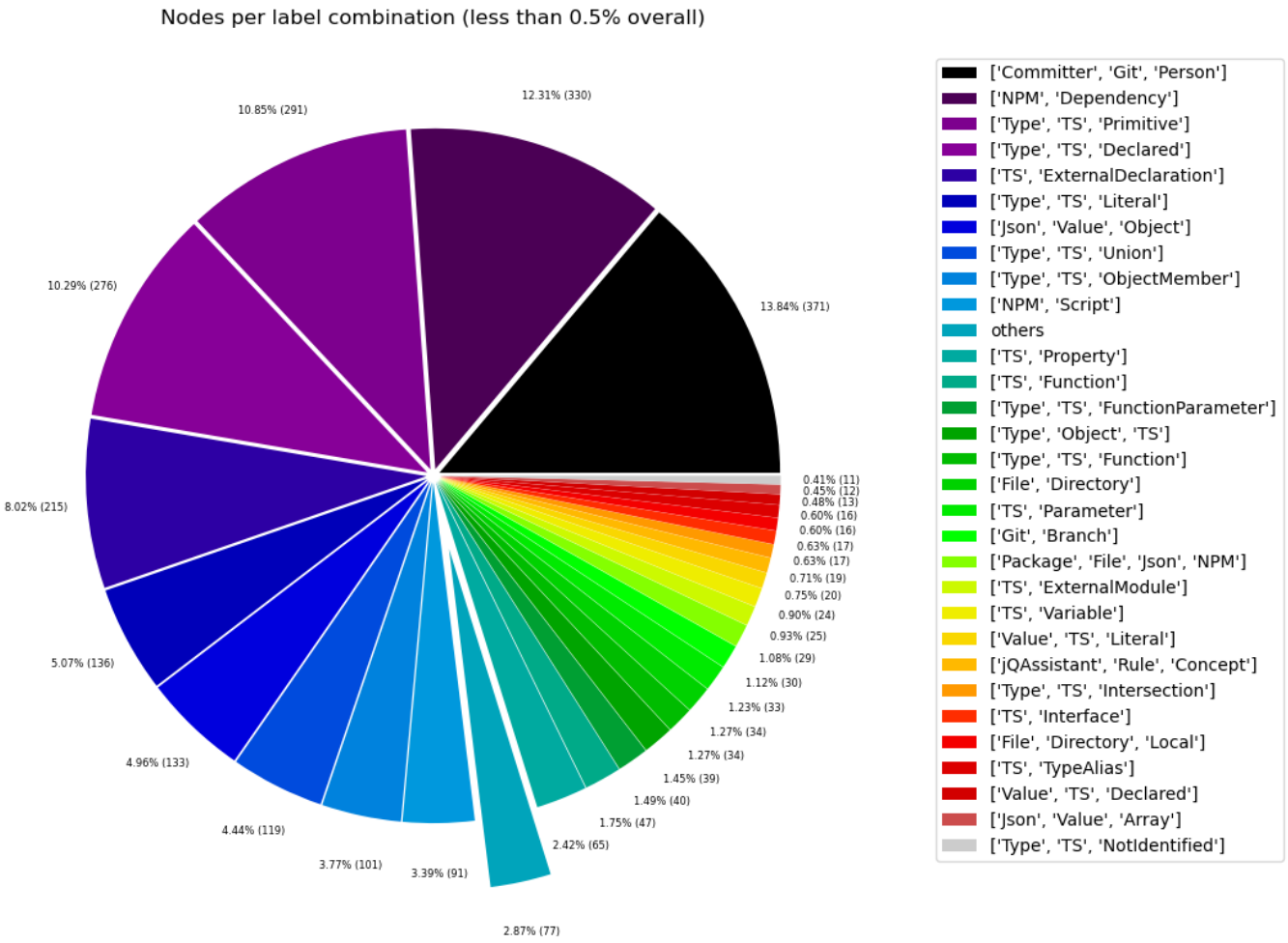


Table 1c - Highest node count by single label

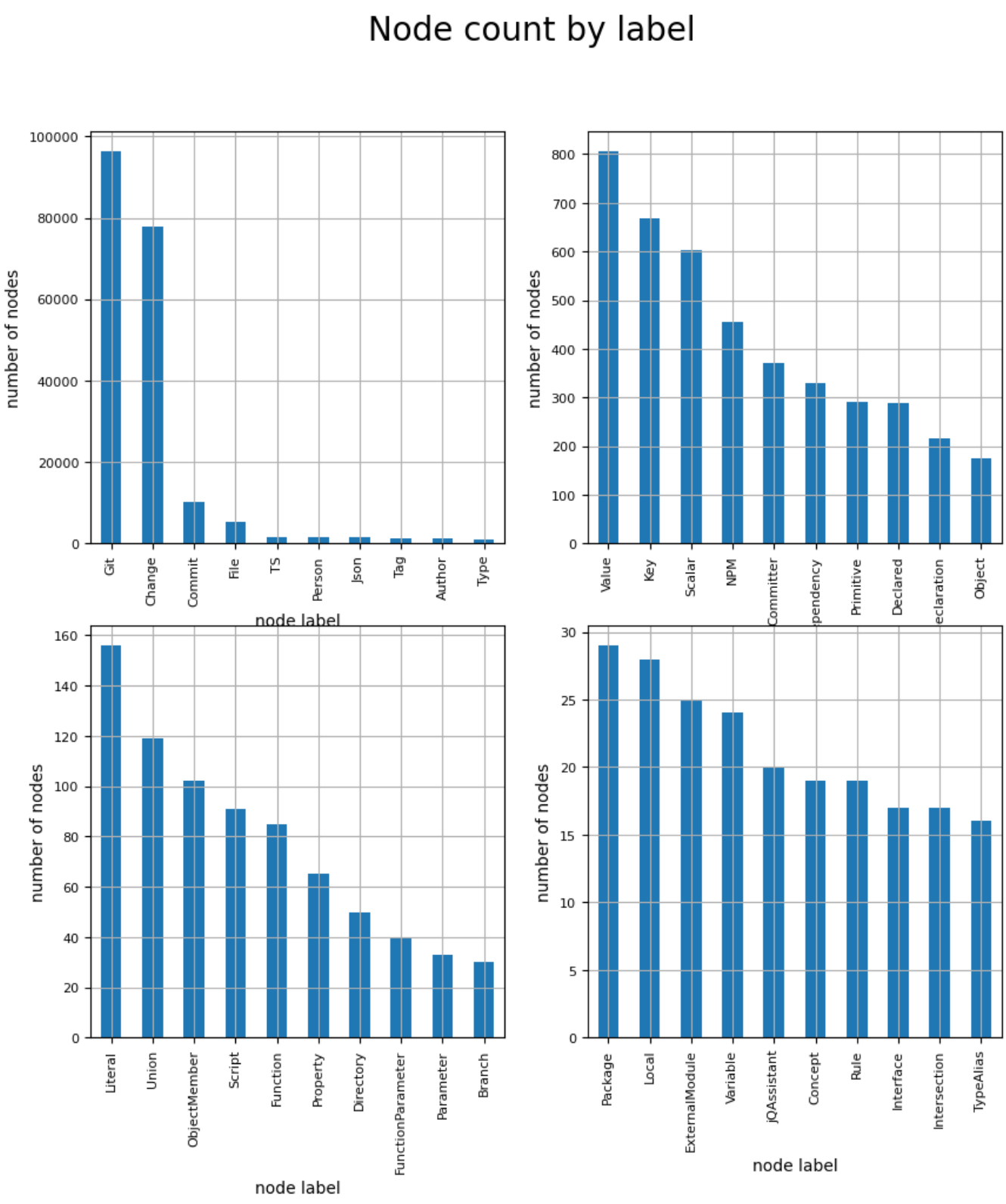
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	96359	96.445801
1	Change	77924	77.994195
2	Commit	10305	10.314283
3	File	5432	5.436893
4	TS	1603	1.604444
5	Person	1563	1.564408
6	Json	1445	1.446302
7	Tag	1197	1.198078
8	Author	1192	1.193074
9	Type	1073	1.073967
10	Value	806	0.806726
11	Key	668	0.668602
12	Scalar	603	0.603543
13	NPM	456	0.456411
14	Committer	371	0.371334
15	Dependency	330	0.330297
16	Primitive	291	0.291262
17	Declared	289	0.289260
18	ExternalDeclaration	215	0.215194
19	Object	175	0.175158
20	Literal	156	0.156141
21	Union	119	0.119107
22	ObjectMember	102	0.102092
23	Script	91	0.091082
24	Function	85	0.085077
25	Property	65	0.065059
26	Directory	50	0.050045
27	FunctionParameter	40	0.040036
28	Parameter	33	0.033030
29	Branch	30	0.030027
30	Package	29	0.029026
31	Local	28	0.028025
32	ExternalModule	25	0.025023
33	Variable	24	0.024022
34	jQAssistant	20	0.020018
35	Concept	19	0.019017
36	Rule	19	0.019017
37	Interface	17	0.017015
38	Intersection	17	0.017015
39	TypeAlias	16	0.016014

Chart 1c - Highest node count by label

Shows the 40 labels with the highest number of nodes.

```
<Figure size 640x480 with 0 Axes>
```



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: 279481
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	77924	27.881681
1	MODIFIES	77924	27.881681
2	UPDATES	51502	18.427729
3	COMMITTED	20610	7.374383
4	CREATES	18526	6.628715
5	HAS_PARENT	11336	4.056090
6	DELETES	10807	3.866810
7	RENAMES	2911	1.041573
8	HAS_NEW_NAME	1647	0.589307
9	ON_COMMIT	1197	0.428294
10	DEPENDS_ON	959	0.343136
11	HAS_KEY	668	0.239014
12	HAS_VALUE	668	0.239014
13	CONTAINS	593	0.212179
14	OF_TYPE	337	0.120581
15	EXPORTS	276	0.098754
16	REFERENCES	197	0.070488
17	DECLARES	186	0.066552
18	DECLARES_DEV_DEPENDENCY	169	0.060469
19	DECLARES_DEPENDENCY	161	0.057607
20	HAS_MEMBER	102	0.036496
21	HAS_TYPE_ARGUMENT	94	0.033634
22	DECLARES_SCRIPT	91	0.032560
23	RETURNS	82	0.029340
24	RESOLVES_TO	77	0.027551
25	HAS_PARAMETER	73	0.026120
26	CONTAINS_VALUE	51	0.018248
27	COPIES	43	0.015386
28	INITIALIZED_WITH	32	0.011450
29	IS_DESCRIBED_IN_NPM_PACKAGE	32	0.011450

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.

```
<Figure size 640x480 with 0 Axes>
```

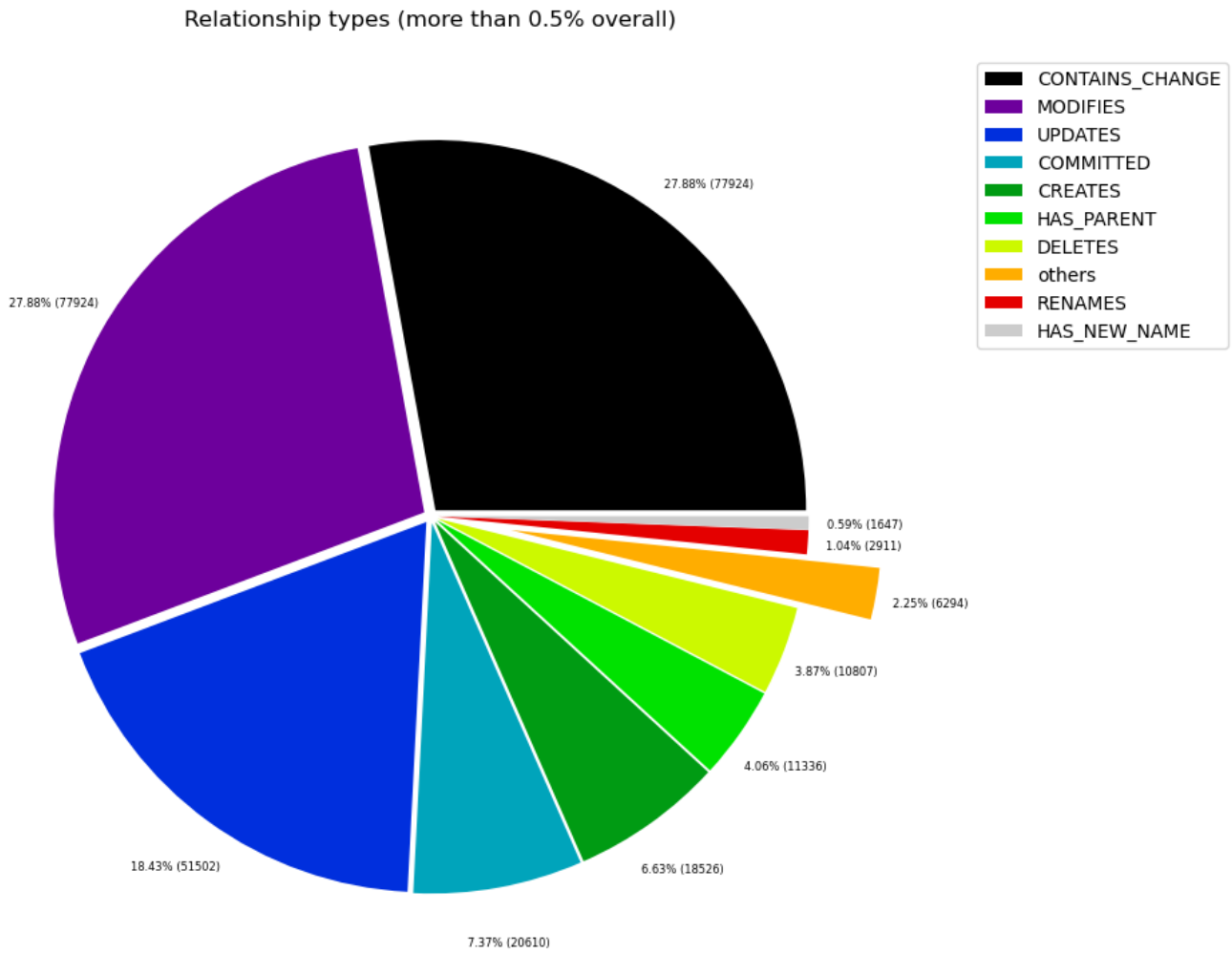


Table 2b - Lowest relationship count by type

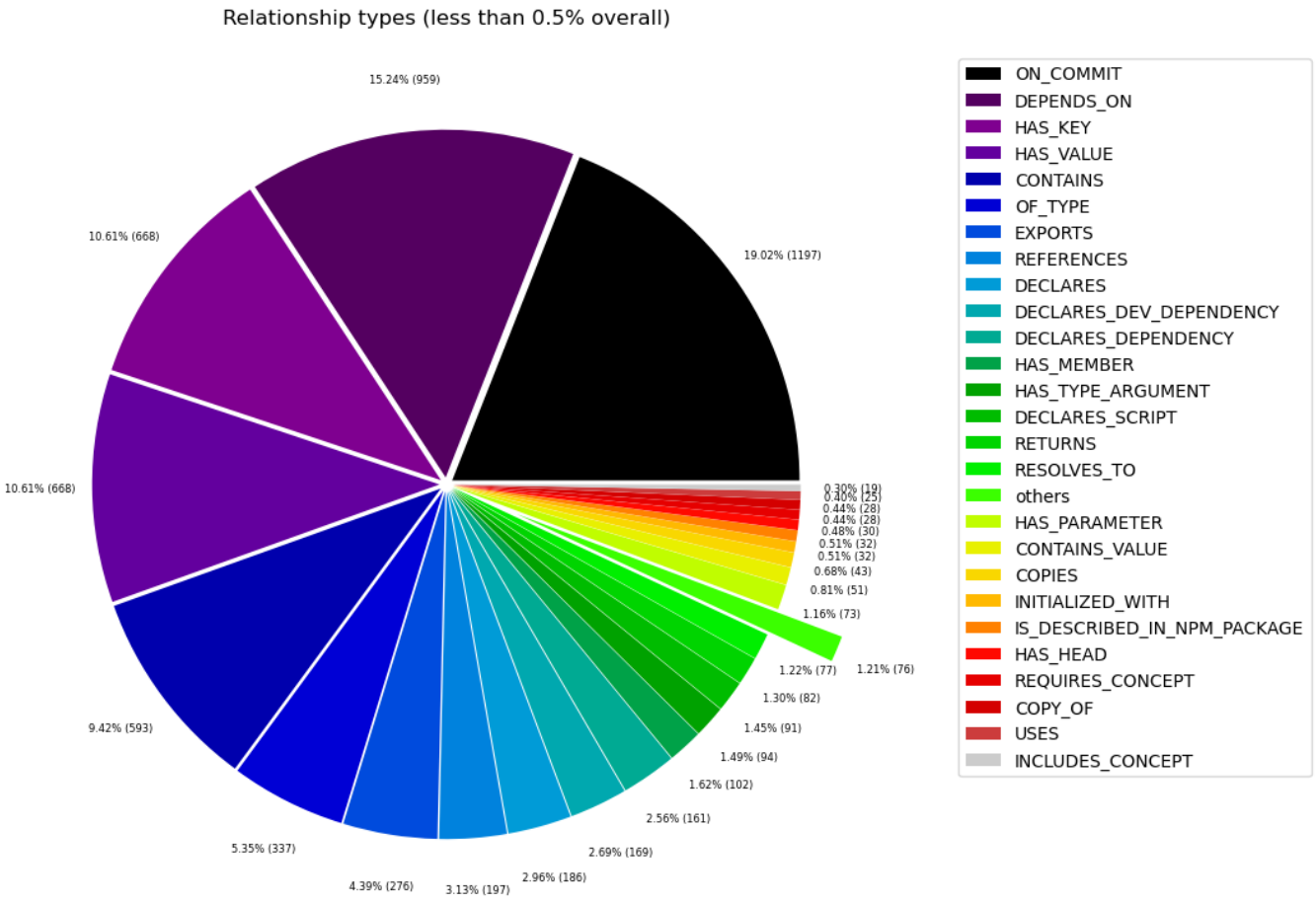
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	PROVIDED_BY_NPM_DEPENDENCY	1	0.000358
1	CONSTRAINED_BY	4	0.001431
2	REFERENCED_PROJECTS	5	0.001789
3	SIMILAR	6	0.002147
4	DECLARES_ENGINE	6	0.002147
5	EXTENDS	6	0.002147
6	HAS_ARGUMENT	6	0.002147
7	CALLS	6	0.002147
8	HAS_NPM_PACKAGE	6	0.002147
9	HAS_ROOT	6	0.002147
10	MEMBER	6	0.002147
11	PARENT	6	0.002147
12	HAS_CONFIG	6	0.002147
13	CONTAINS_PROJECT	6	0.002147
14	INCLUDES_CONCEPT	19	0.006798
15	USES	25	0.008945
16	REQUIRES_CONCEPT	28	0.010019
17	COPY_OF	28	0.010019
18	HAS_HEAD	30	0.010734
19	IS_DESCRIBED_IN_NPM_PACKAGE	32	0.011450
20	INITIALIZED_WITH	32	0.011450
21	COPIES	43	0.015386
22	CONTAINS_VALUE	51	0.018248
23	HAS_PARAMETER	73	0.026120
24	RESOLVES_TO	77	0.027551
25	RETURNS	82	0.029340
26	DECLARES_SCRIPT	91	0.032560
27	HAS_TYPE_ARGUMENT	94	0.033634
28	HAS_MEMBER	102	0.036496
29	DECLARES_DEPENDENCY	161	0.057607

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.

```
<Figure size 640x480 with 0 Axes>
```



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	numberOfNodesWithSameLabelsAsSource	numberOfNodes
0	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	77924	10305	
1	[Git, Change]	MODIFIES	[File, Git]	77924	77924	
2	[Git, Change]	UPDATES	[File, Git]	51502	77924	
3	[Git, Change]	CREATES	[File, Git]	18526	77924	
4	[Git, Commit]	HAS_PARENT	[Git, Commit]	11336	10305	
5	[Git, Change]	DELETES	[File, Git]	10807	77924	
6	[Author, Git, Person]	COMMITTED	[Git, Commit]	10305	1192	
7	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10305	371	
8	[Git, Change]	RENAMES	[File, Git]	2911	77924	
9	[File, Git]	HAS_NEW_NAME	[File, Git]	1647	5340	
10	[Git, Tag]	ON_COMMIT	[Git, Commit]	1197	1197	
11	[Json, Value, Object]	HAS_KEY	[Json, Key]	668	133	
12	[Json, Key]	HAS_VALUE	[Json, Value, Scalar]	552	668	
13	[TS, Function]	DEPENDS_ON	[TS, ExternalDeclaration]	285	47	
14	[TS, ExternalModule]	EXPORTS	[TS, ExternalDeclaration]	215	25	
15	[File, TS, Local, Module, Mark4ModuleWeaklyCon...	DEPENDS_ON	[TS, ExternalDeclaration]	192	2	
16	[Package, File, Json, NPM]	DECLARES_DEV_DEPENDENCY	[NPM, Dependency]	169	29	
17	[Package, File, Json, NPM]	DECLARES_DEPENDENCY	[NPM, Dependency]	161	29	
18	[Type, TS, Union]	CONTAINS	[Type, TS, Primitive]	147	119	
19	[Type, TS, Declared]	REFERENCES	[TS, ExternalDeclaration]	142	276	
20	[TS, Function]	DEPENDS_ON	[TS, ExternalModule]	131	47	
21	[Type, TS, Union]	CONTAINS	[Type, TS, Literal]	119	119	
22	[Json, Key]	HAS_VALUE	[Json, Value, Object]	104	668	
23	[Type, Object, TS]	HAS_MEMBER	[Type, TS, ObjectMember]	101	39	
24	[Package, File, Json, NPM]	DECLARES_SCRIPT	[NPM, Script]	91	29	
25	[Type, TS, Union]	CONTAINS	[Type, TS, Declared]	70	119	
26	[File, Directory]	CONTAINS	[File, Directory]	63	34	
27	[TS, Interface]	DECLARES	[TS, Property]	61	17	
28	[File, Directory]	CONTAINS	[Package, File, Json, NPM]	58	34	
29	[File, Git]	RESOLVES_TO	[Package, File, Json, NPM]	57	5340	

Graph Density

```
total_number_of_nodes (vertices): 99910
total_number_of_relationships (edges): 279481
-> total directed graph density: 2.7998754815235382e-05
-> total directed graph density in percent: 0.0027998754815235383
```