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]	80711	78.281912
1	[Git, Commit]	10505	10.188840
2	[File, Git]	5447	5.283066
3	[Git, Tag]	1284	1.245357
4	[Author, Git, Person]	1203	1.166794
5	[Json, Key]	668	0.647896
6	[Json, Value, Scalar]	603	0.584852
7	[Committer, Git, Person]	371	0.359834
8	[NPM, Dependency]	330	0.320068
9	[Type, TS, Primitive]	291	0.282242
10	[Type, TS, Declared]	276	0.267693
11	[TS, ExternalDeclaration]	215	0.208529
12	[Type, TS, Literal]	136	0.131907
13	[Json, Value, Object]	133	0.128997
14	[Type, TS, Union]	119	0.115419
15	[Type, TS, ObjectMember]	101	0.097960
16	[NPM, Script]	91	0.088261
17	[TS, Property]	65	0.063044
18	[TS, Function]	47	0.045585
19	[Type, TS, FunctionParameter]	40	0.038796
20	[Type, Object, TS]	39	0.037826
21	[File, Directory]	34	0.032977
22	[Type, TS, Function]	34	0.032977
23	[TS, Parameter]	33	0.032007
24	[Git, Branch]	30	0.029097
25	[Package, File, Json, NPM]	29	0.028127
26	[TS, ExternalModule]	25	0.024248
27	[TS, Variable]	24	0.023278
28	[Value, TS, Literal]	20	0.019398
29	[jQAssistant, Rule, Concept]	19	0.018428

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>
```

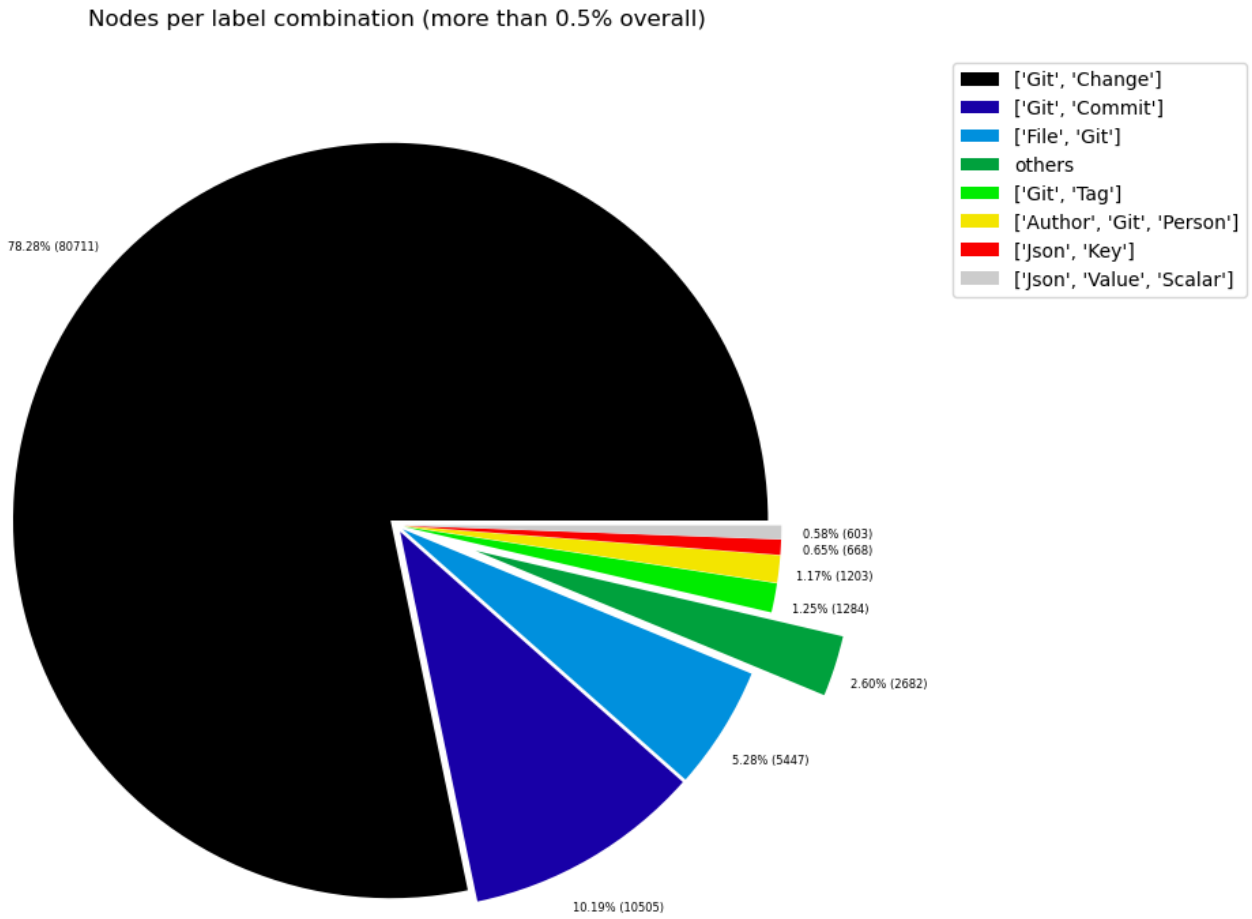


Table 1b - Lowest node count by label combination

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	0.000970
1	[File, TS, Scan]	1	0.000970
2	[TS, Method]	1	0.000970
3	[Repository, File, Git]	1	0.000970
4	[TS, Constructor]	1	0.000970
5	[Value, TS, ObjectMember]	1	0.000970
6	[TS, Class]	1	0.000970
7	[TS, Enum]	2	0.001940
8	[Value, Object, TS]	3	0.002910
9	[Type, TS, Tuple]	3	0.002910
10	[Value, TS, Function]	4	0.003880
11	[TS, TypeParameter]	4	0.003880
12	[Value, TS, Complex]	5	0.004850
13	[NPM, Engine]	6	0.005819
14	[Project, TS]	6	0.005819
15	[File, Local]	6	0.005819
16	[Value, TS, Call]	6	0.005819
17	[Value, TS, Member]	6	0.005819
18	[File, TS, Local, Module]	6	0.005819
19	[Type, TS, TypeParameterReference]	6	0.005819
20	[TS, EnumMember]	8	0.007759
21	[Type, TS, NotIdentified]	11	0.010669
22	[Json, Value, Array]	12	0.011639
23	[Value, TS, Declared]	13	0.012609
24	[TS, TypeAlias]	16	0.015518
25	[File, Directory, Local]	16	0.015518
26	[TS, Interface]	17	0.016488
27	[Type, TS, Intersection]	17	0.016488
28	[jQAssistant, Rule, Concept]	19	0.018428
29	[Value, TS, Literal]	20	0.019398

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>
```

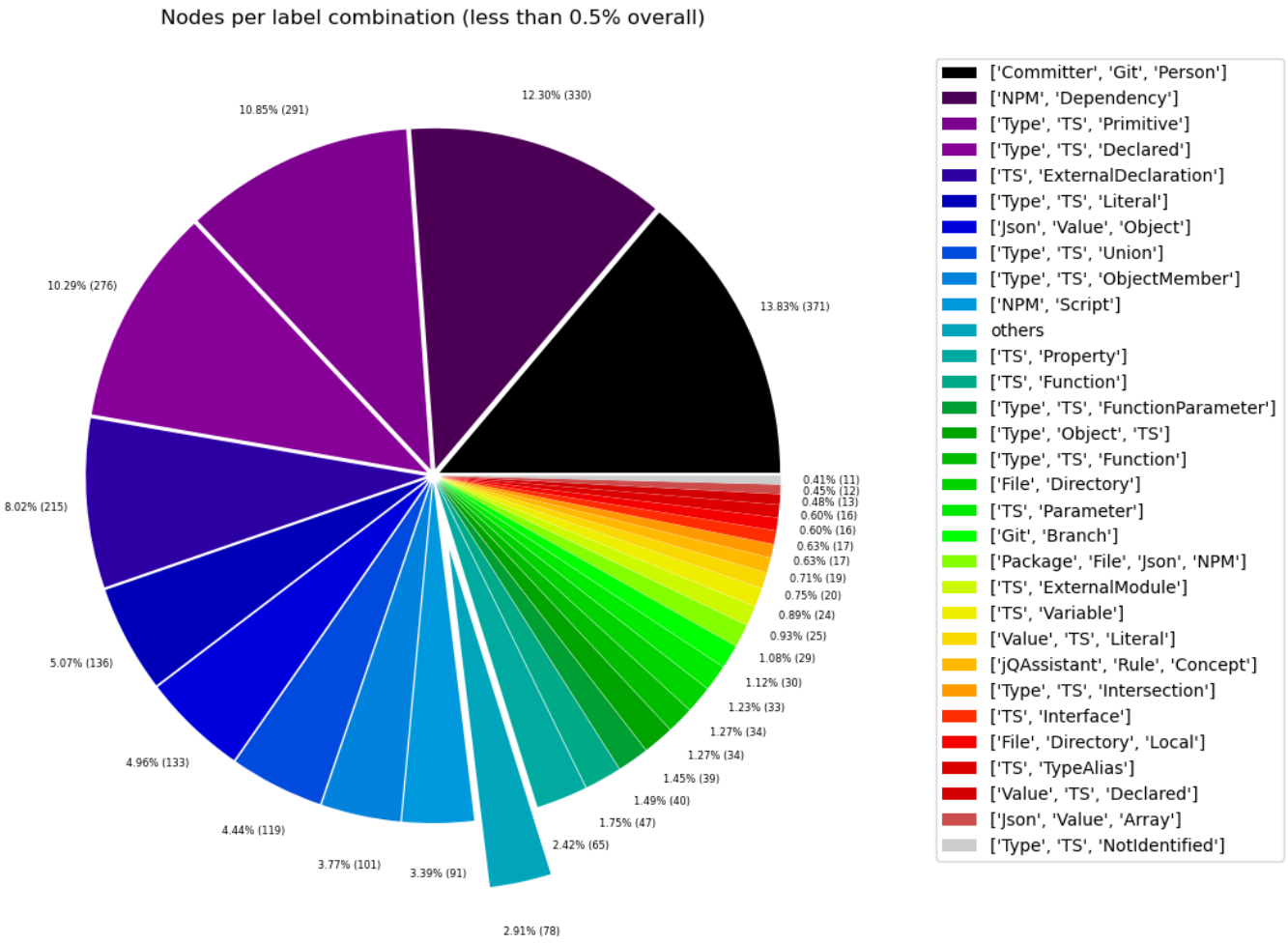


Table 1c - Highest node count by single label

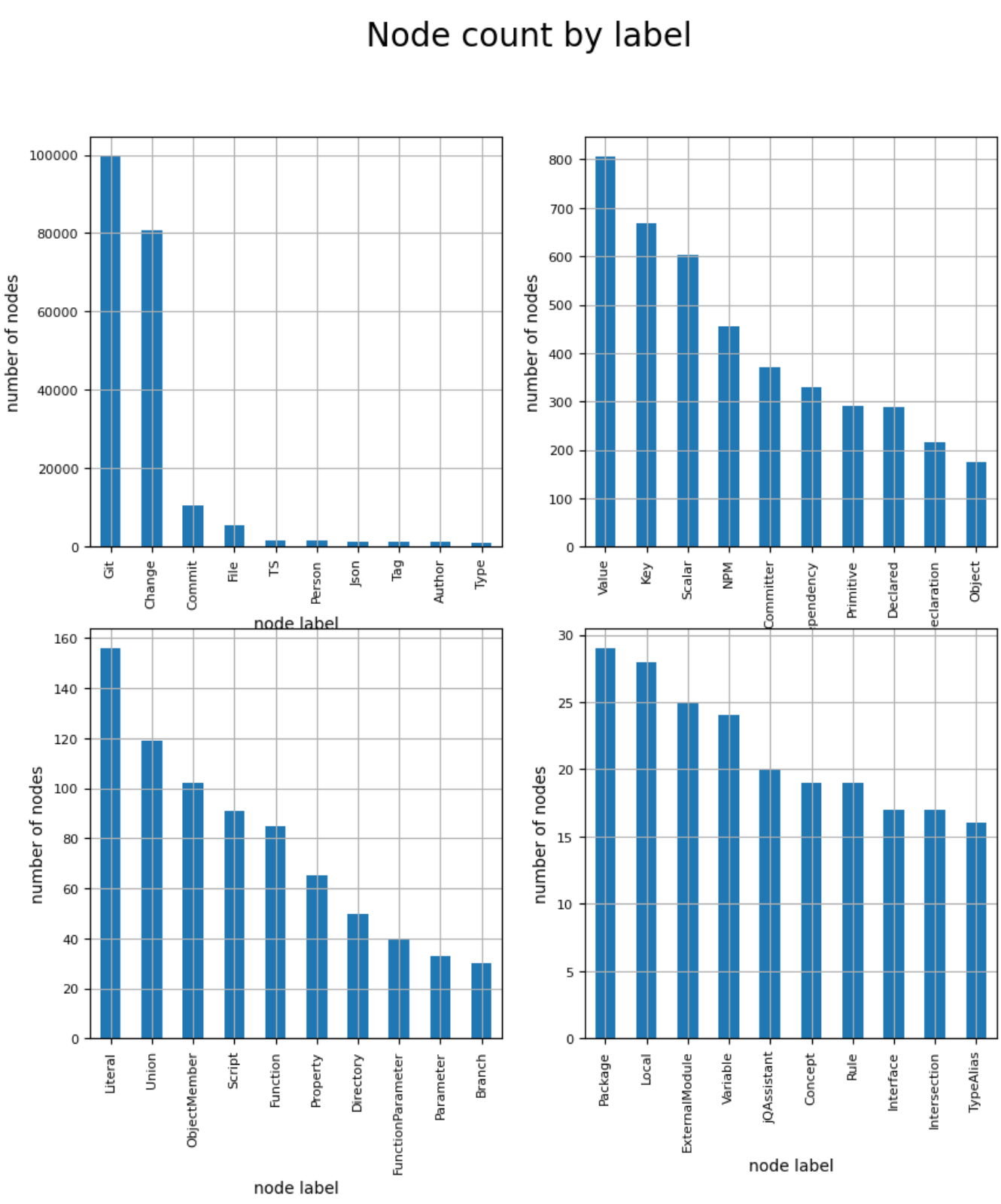
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	99552	96.555871
1	Change	80711	78.281912
2	Commit	10505	10.188840
3	File	5540	5.373268
4	TS	1603	1.554756
5	Person	1574	1.526629
6	Json	1445	1.401511
7	Tag	1284	1.245357
8	Author	1203	1.166794
9	Type	1073	1.040707
10	Value	806	0.781743
11	Key	668	0.647896
12	Scalar	603	0.584852
13	NPM	456	0.442276
14	Committer	371	0.359834
15	Dependency	330	0.320068
16	Primitive	291	0.282242
17	Declared	289	0.280302
18	ExternalDeclaration	215	0.208529
19	Object	175	0.169733
20	Literal	156	0.151305
21	Union	119	0.115419
22	ObjectMember	102	0.098930
23	Script	91	0.088261
24	Function	85	0.082442
25	Property	65	0.063044
26	Directory	50	0.048495
27	FunctionParameter	40	0.038796
28	Parameter	33	0.032007
29	Branch	30	0.029097
30	Package	29	0.028127
31	Local	28	0.027157
32	ExternalModule	25	0.024248
33	Variable	24	0.023278
34	jQAssistant	20	0.019398
35	Concept	19	0.018428
36	Rule	19	0.018428
37	Interface	17	0.016488
38	Intersection	17	0.016488
39	TypeAlias	16	0.015518

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: 307939
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	80711	26.210061
1	MODIFIES	80711	26.210061
2	UPDATES	52787	17.142031
3	COMMITTED	21010	6.822780
4	CREATES	19561	6.352232
5	HAS_PARENT	11545	3.749119
6	DELETES	11534	3.745547
7	HAS_COMMIT	10505	3.411390
8	HAS_FILE	5447	1.768857
9	RENAMES	3171	1.029749
10	HAS_NEW_NAME	1718	0.557903
11	HAS_TAG	1284	0.416966
12	ON_COMMIT	1284	0.416966
13	HAS_AUTHOR	1203	0.390662
14	DEPENDS_ON	959	0.311425
15	HAS_KEY	668	0.216926
16	HAS_VALUE	668	0.216926
17	CONTAINS	594	0.192895
18	HAS_COMMITTER	371	0.120478
19	OF_TYPE	337	0.109437
20	EXPORTS	276	0.089628
21	REFERENCES	197	0.063974
22	DECLARES	186	0.060402
23	DECLARES_DEV_DEPENDENCY	169	0.054881
24	DECLARES_DEPENDENCY	161	0.052283
25	HAS_MEMBER	102	0.033123
26	HAS_TYPE_ARGUMENT	94	0.030526
27	DECLARES_SCRIPT	91	0.029551
28	RETURNS	82	0.026629
29	HAS_PARAMETER	73	0.023706

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>
```

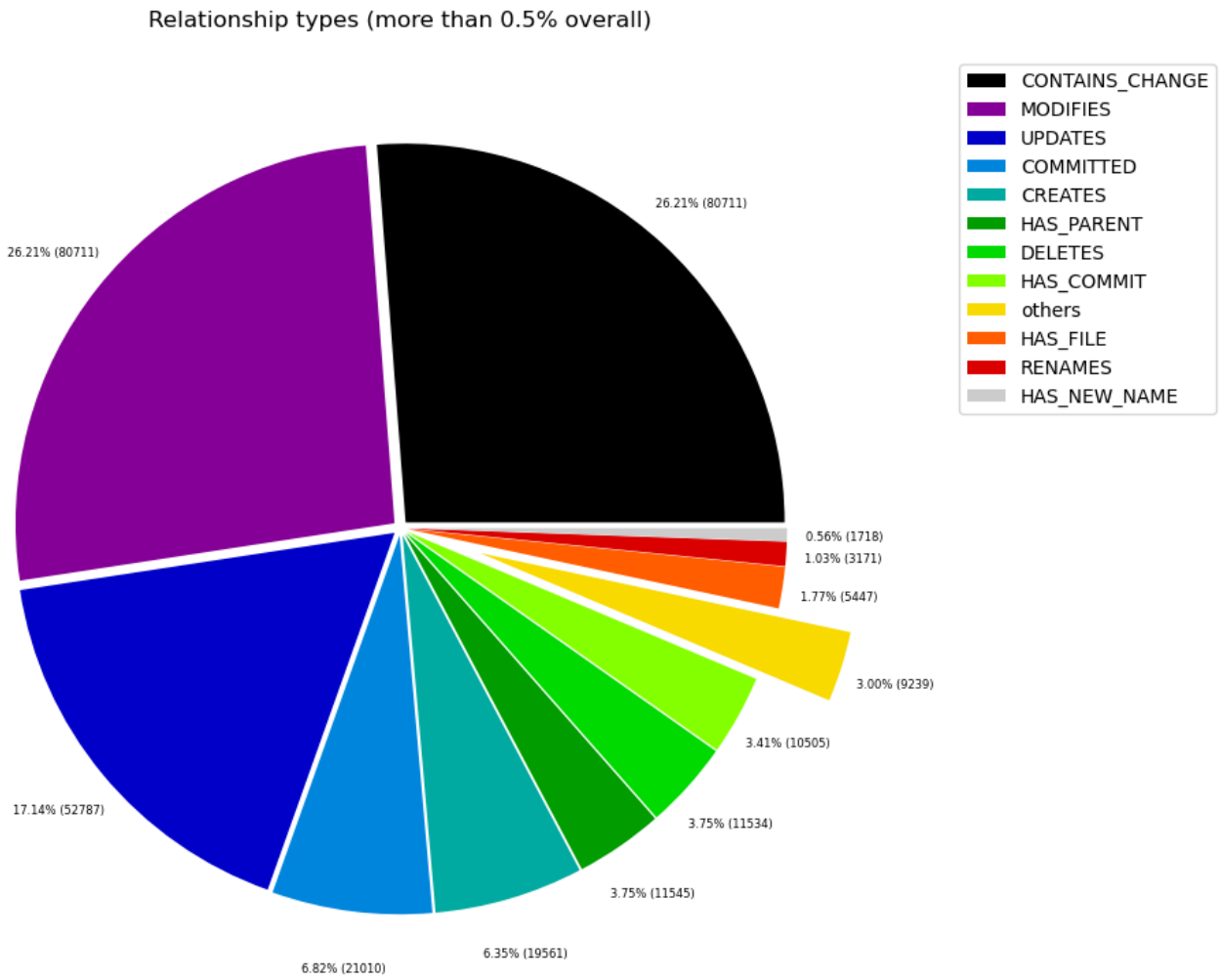


Table 2b - Lowest relationship count by type

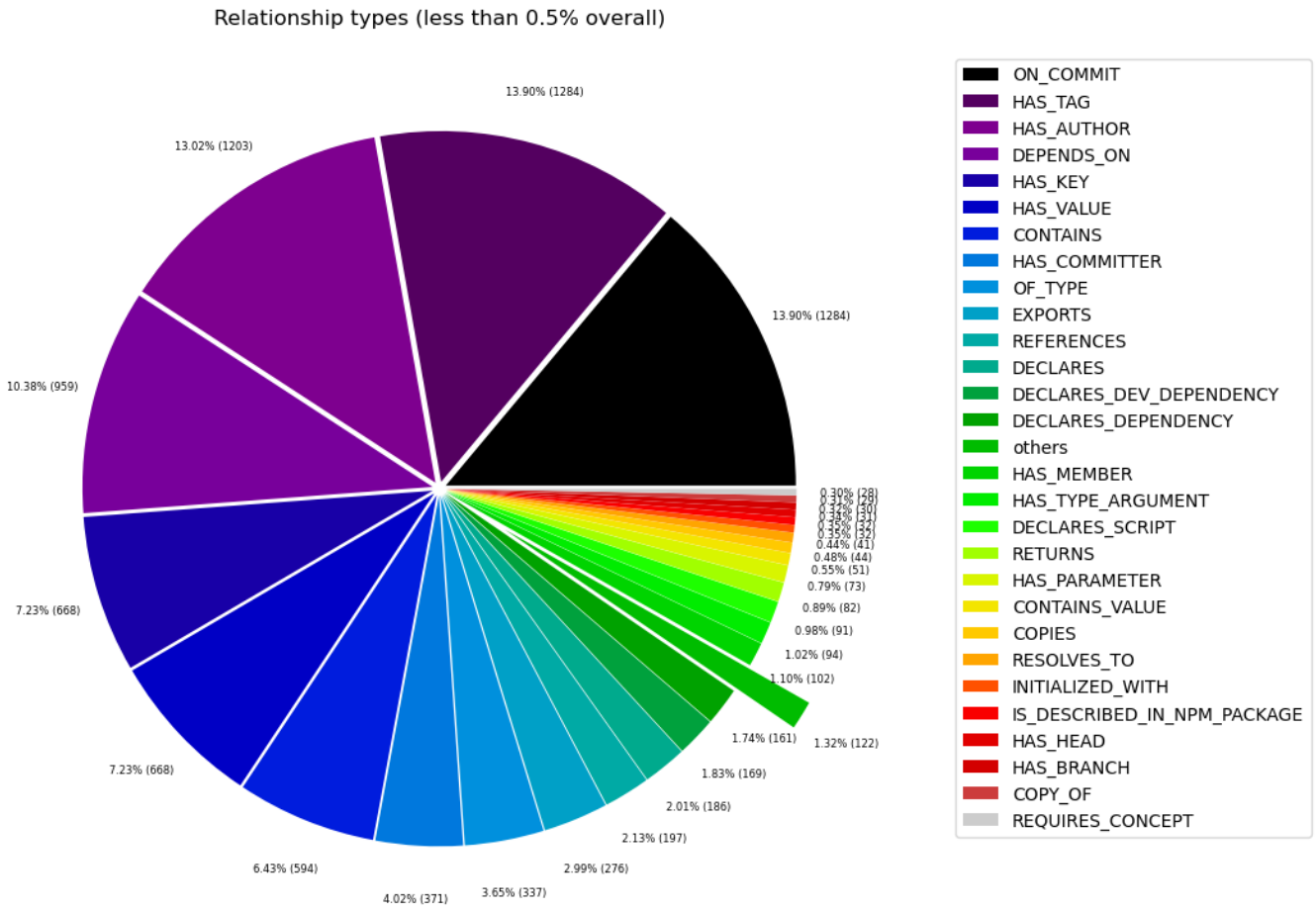
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	PROVIDED_BY_NPM_DEPENDENCY	1	0.000325
1	IS_IMPLEMENTED_IN	2	0.000649
2	CONSTRAINED_BY	4	0.001299
3	REFERENCED_PROJECTS	5	0.001624
4	CONTAINS_PROJECT	6	0.001948
5	DECLARES_ENGINE	6	0.001948
6	EXTENDS	6	0.001948
7	HAS_ARGUMENT	6	0.001948
8	CALLS	6	0.001948
9	HAS_NPM_PACKAGE	6	0.001948
10	HAS_ROOT	6	0.001948
11	MEMBER	6	0.001948
12	PARENT	6	0.001948
13	HAS_CONFIG	6	0.001948
14	SIMILAR	6	0.001948
15	INCLUDES_CONCEPT	19	0.006170
16	USES	25	0.008118
17	REQUIRES_CONCEPT	28	0.009093
18	COPY_OF	29	0.009417
19	HAS_BRANCH	30	0.009742
20	HAS_HEAD	31	0.010067
21	IS_DESCRIBED_IN_NPM_PACKAGE	32	0.010392
22	INITIALIZED_WITH	32	0.010392
23	RESOLVES_TO	41	0.013314
24	COPIES	44	0.014289
25	CONTAINS_VALUE	51	0.016562
26	HAS_PARAMETER	73	0.023706
27	RETURNS	82	0.026629
28	DECLARES_SCRIPT	91	0.029551
29	HAS_TYPE_ARGUMENT	94	0.030526

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	numberOfNodesT
0	[Git, Change]	MODIFIES	[File, Git]	80711	80711	
1	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	80711	10505	
2	[Git, Change]	UPDATES	[File, Git]	52787	80711	
3	[Git, Change]	CREATES	[File, Git]	19561	80711	
4	[Git, Commit]	HAS_PARENT	[Git, Commit]	11545	10505	
5	[Git, Change]	DELETES	[File, Git]	11534	80711	
6	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	10505	1	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	10505	1203	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10505	371	
9	[Repository, File, Git]	HAS_FILE	[File, Git]	5447	1	
10	[Git, Change]	RENAMES	[File, Git]	3171	80711	
11	[File, Git]	HAS_NEW_NAME	[File, Git]	1718	5447	
12	[Repository, File, Git]	HAS_TAG	[Git, Tag]	1284	1	
13	[Git, Tag]	ON_COMMIT	[Git, Commit]	1284	1284	
14	[Repository, File, Git]	HAS_AUTHOR	[Author, Git, Person]	1203	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]	285	47	
19	[TS, ExternalModule]	EXPORTS	[TS, ExternalDeclaration]	215	25	
20	[File, TS, Local, Module, Mark4ModuleWeaklyCon...	DEPENDS_ON	[TS, ExternalDeclaration]	192	2	
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]	131	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): 103103
total_number_of_relationships (edges): 307939
-> total directed graph density: 2.8968519687391168e-05
-> total directed graph density in percent: 0.002896851968739117
```