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]	80737	78.280556
1	[Git, Commit]	10512	10.192170
2	[File, Git]	5448	5.282243
3	[Git, Tag]	1284	1.244934
4	[Author, Git, Person]	1203	1.166398
5	[Json, Key]	668	0.647676
6	[Json, Value, Scalar]	603	0.584654
7	[Committer, Git, Person]	371	0.359712
8	[NPM, Dependency]	330	0.319960
9	[Type, TS, Primitive]	291	0.282146
10	[Type, TS, Declared]	276	0.267603
11	[TS, ExternalDeclaration]	215	0.208459
12	[Type, TS, Literal]	136	0.131862
13	[Json, Value, Object]	133	0.128953
14	[Type, TS, Union]	119	0.115379
15	[Type, TS, ObjectMember]	101	0.097927
16	[NPM, Script]	91	0.088231
17	[TS, Property]	65	0.063022
18	[TS, Function]	47	0.045570
19	[Type, TS, FunctionParameter]	40	0.038783
20	[Type, Object, TS]	39	0.037813
21	[File, Directory]	34	0.032966
22	[Type, TS, Function]	34	0.032966
23	[TS, Parameter]	33	0.031996
24	[Git, Branch]	31	0.030057
25	[Package, File, Json, NPM]	29	0.028118
26	[TS, ExternalModule]	25	0.024239
27	[TS, Variable]	24	0.023270
28	[Value, TS, Literal]	20	0.019391
29	[jQAssistant, Rule, Concept]	19	0.018422

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>
```

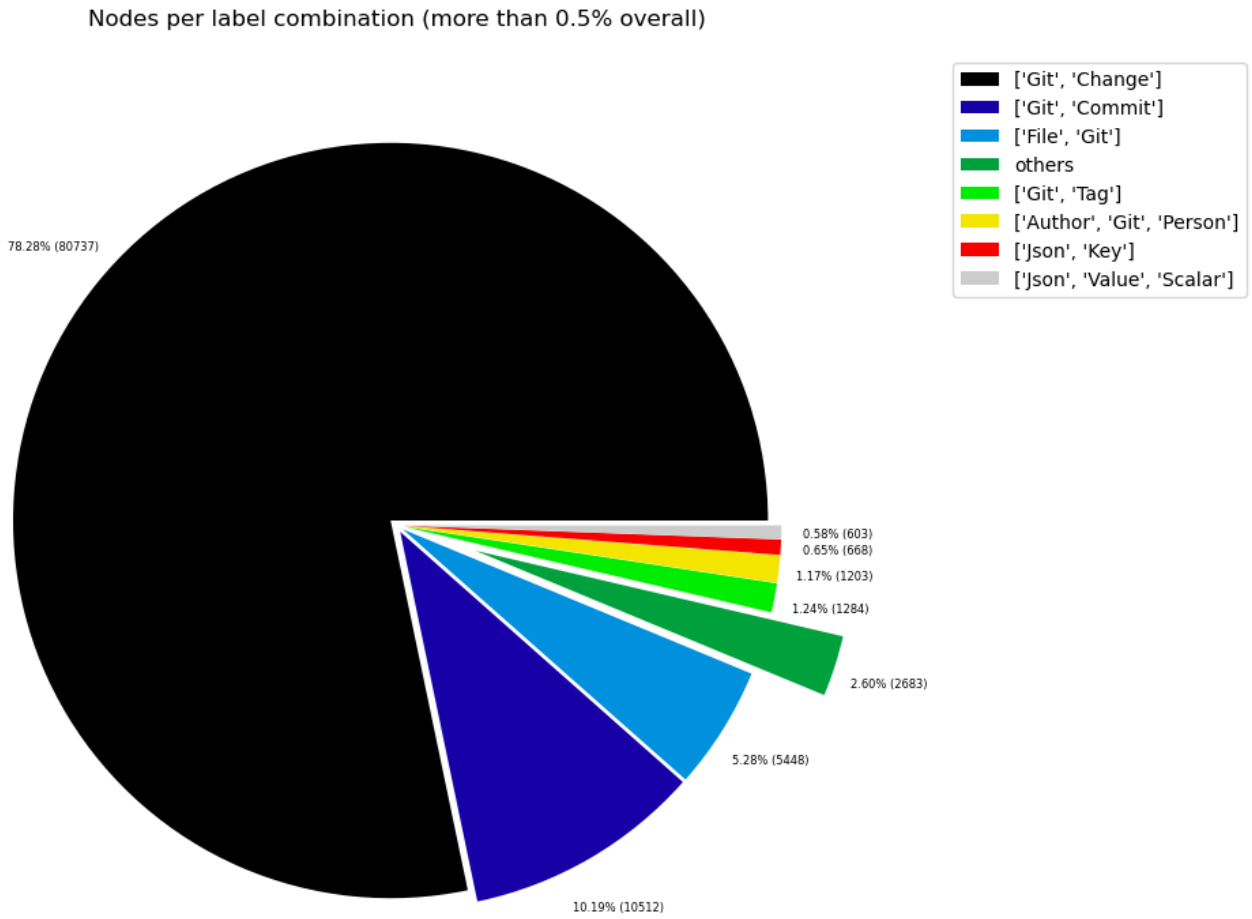


Table 1b - Lowest node count by label combination

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	0.001939
8	[Value, Object, TS]	3	0.002909
9	[Type, TS, Tuple]	3	0.002909
10	[Value, TS, Function]	4	0.003878
11	[TS, TypeParameter]	4	0.003878
12	[Value, TS, Complex]	5	0.004848
13	[NPM, Engine]	6	0.005817
14	[Project, TS]	6	0.005817
15	[File, Local]	6	0.005817
16	[Value, TS, Call]	6	0.005817
17	[Value, TS, Member]	6	0.005817
18	[File, TS, Local, Module]	6	0.005817
19	[Type, TS, TypeParameterReference]	6	0.005817
20	[TS, EnumMember]	8	0.007757
21	[Type, TS, NotIdentified]	11	0.010665
22	[Json, Value, Array]	12	0.011635
23	[Value, TS, Declared]	13	0.012604
24	[TS, TypeAlias]	16	0.015513
25	[File, Directory, Local]	16	0.015513
26	[TS, Interface]	17	0.016483
27	[Type, TS, Intersection]	17	0.016483
28	[jQAssistant, Rule, Concept]	19	0.018422
29	[Value, TS, Literal]	20	0.019391

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>
```

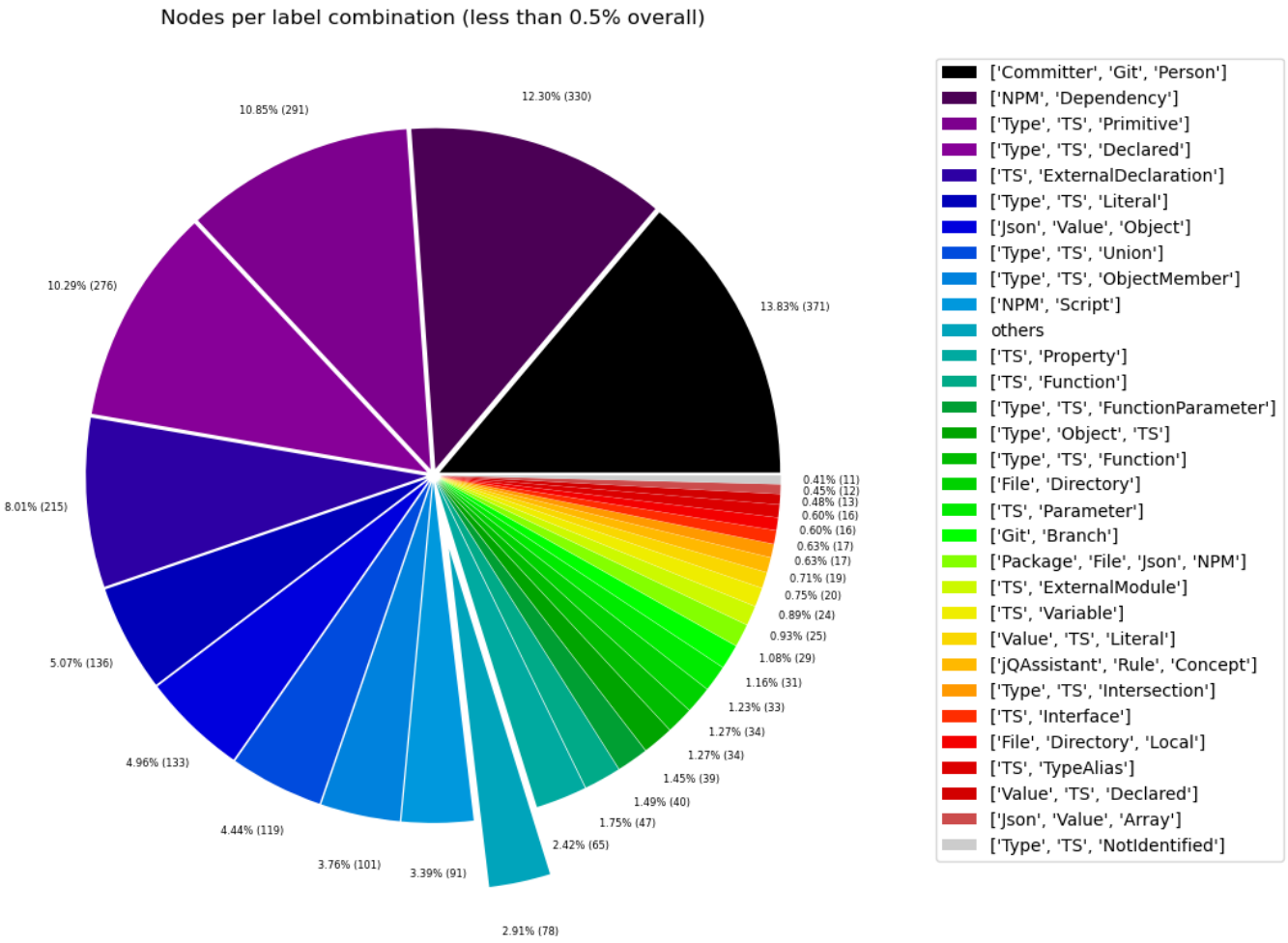


Table 1c - Highest node count by single label

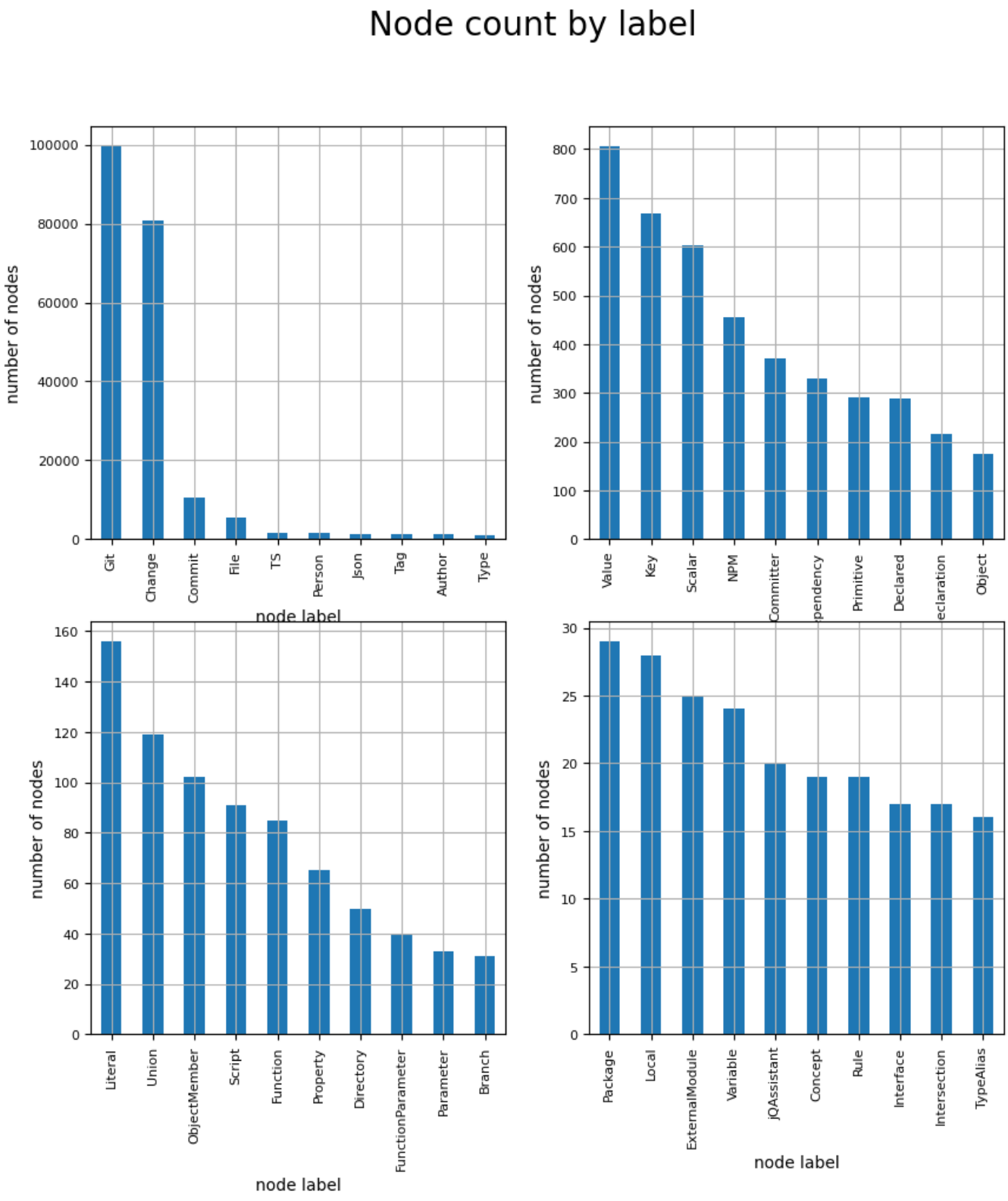
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	99587	96.557040
1	Change	80737	78.280556
2	Commit	10512	10.192170
3	File	5541	5.372414
4	TS	1603	1.554228
5	Person	1574	1.526111
6	Json	1445	1.401036
7	Tag	1284	1.244934
8	Author	1203	1.166398
9	Type	1073	1.040354
10	Value	806	0.781477
11	Key	668	0.647676
12	Scalar	603	0.584654
13	NPM	456	0.442126
14	Committer	371	0.359712
15	Dependency	330	0.319960
16	Primitive	291	0.282146
17	Declared	289	0.280207
18	ExternalDeclaration	215	0.208459
19	Object	175	0.169676
20	Literal	156	0.151254
21	Union	119	0.115379
22	ObjectMember	102	0.098897
23	Script	91	0.088231
24	Function	85	0.082414
25	Property	65	0.063022
26	Directory	50	0.048479
27	FunctionParameter	40	0.038783
28	Parameter	33	0.031996
29	Branch	31	0.030057
30	Package	29	0.028118
31	Local	28	0.027148
32	ExternalModule	25	0.024239
33	Variable	24	0.023270
34	jQAssistant	20	0.019391
35	Concept	19	0.018422
36	Rule	19	0.018422
37	Interface	17	0.016483
38	Intersection	17	0.016483
39	TypeAlias	16	0.015513

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: 308048
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	80737	26.209227
1	MODIFIES	80737	26.209227
2	UPDATES	52795	17.138563
3	COMMITTED	21024	6.824910
4	CREATES	19562	6.350309
5	HAS_PARENT	11552	3.750065
6	DELETES	11551	3.749740
7	HAS_COMMIT	10512	3.412455
8	HAS_FILE	5448	1.768556
9	RENAMES	3171	1.029385
10	HAS_NEW_NAME	1718	0.557705
11	HAS_TAG	1284	0.416818
12	ON_COMMIT	1284	0.416818
13	HAS_AUTHOR	1203	0.390524
14	DEPENDS_ON	959	0.311315
15	HAS_KEY	668	0.216849
16	HAS_VALUE	668	0.216849
17	CONTAINS	594	0.192827
18	HAS_COMMITTER	371	0.120436
19	OF_TYPE	337	0.109399
20	EXPORTS	276	0.089596
21	REFERENCES	197	0.063951
22	DECLARES	186	0.060380
23	DECLARES_DEV_DEPENDENCY	169	0.054862
24	DECLARES_DEPENDENCY	161	0.052265
25	HAS_MEMBER	102	0.033112
26	HAS_TYPE_ARGUMENT	94	0.030515
27	DECLARES_SCRIPT	91	0.029541
28	RETURNS	82	0.026619
29	HAS_PARAMETER	73	0.023698

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>
```

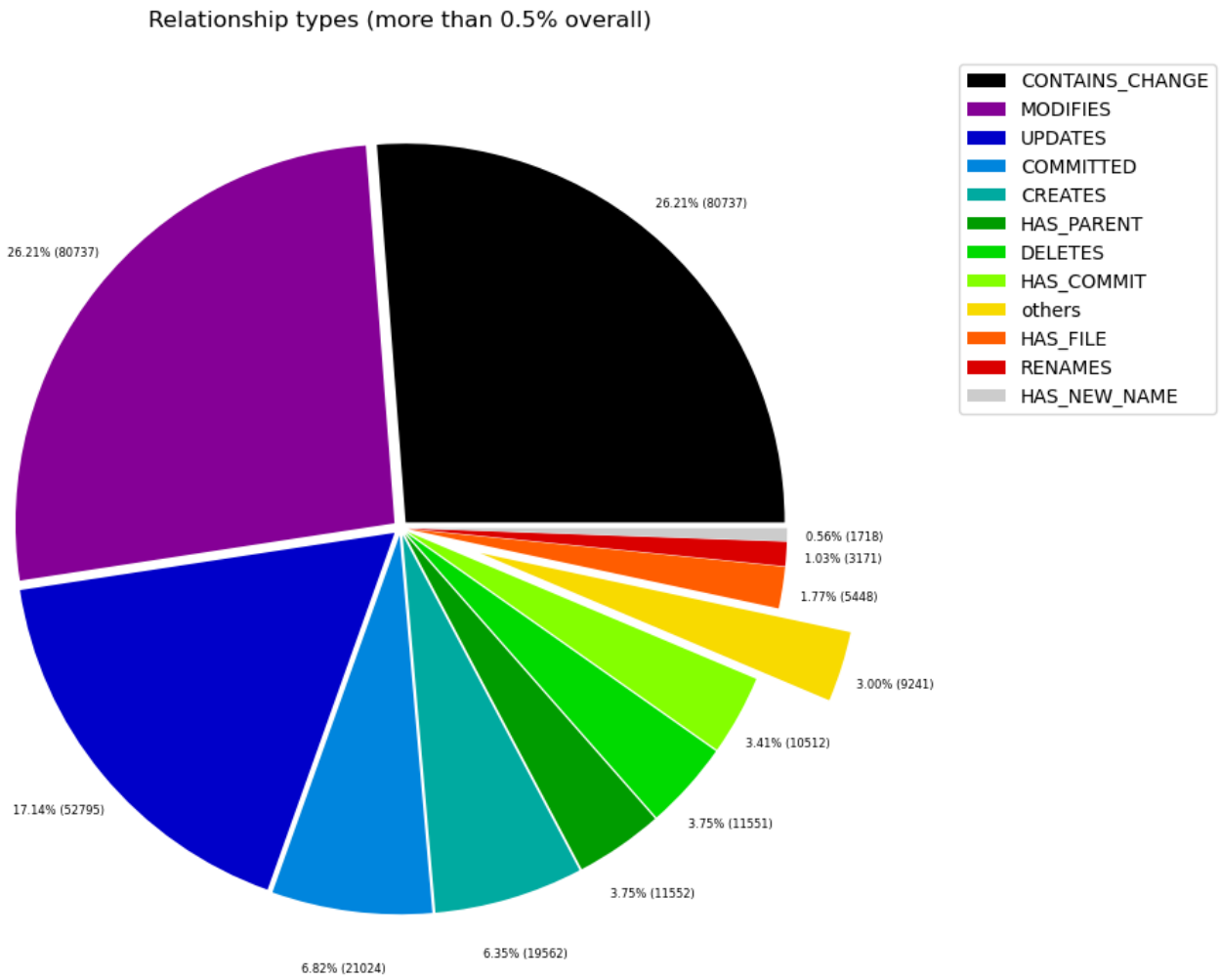


Table 2b - Lowest relationship count by type

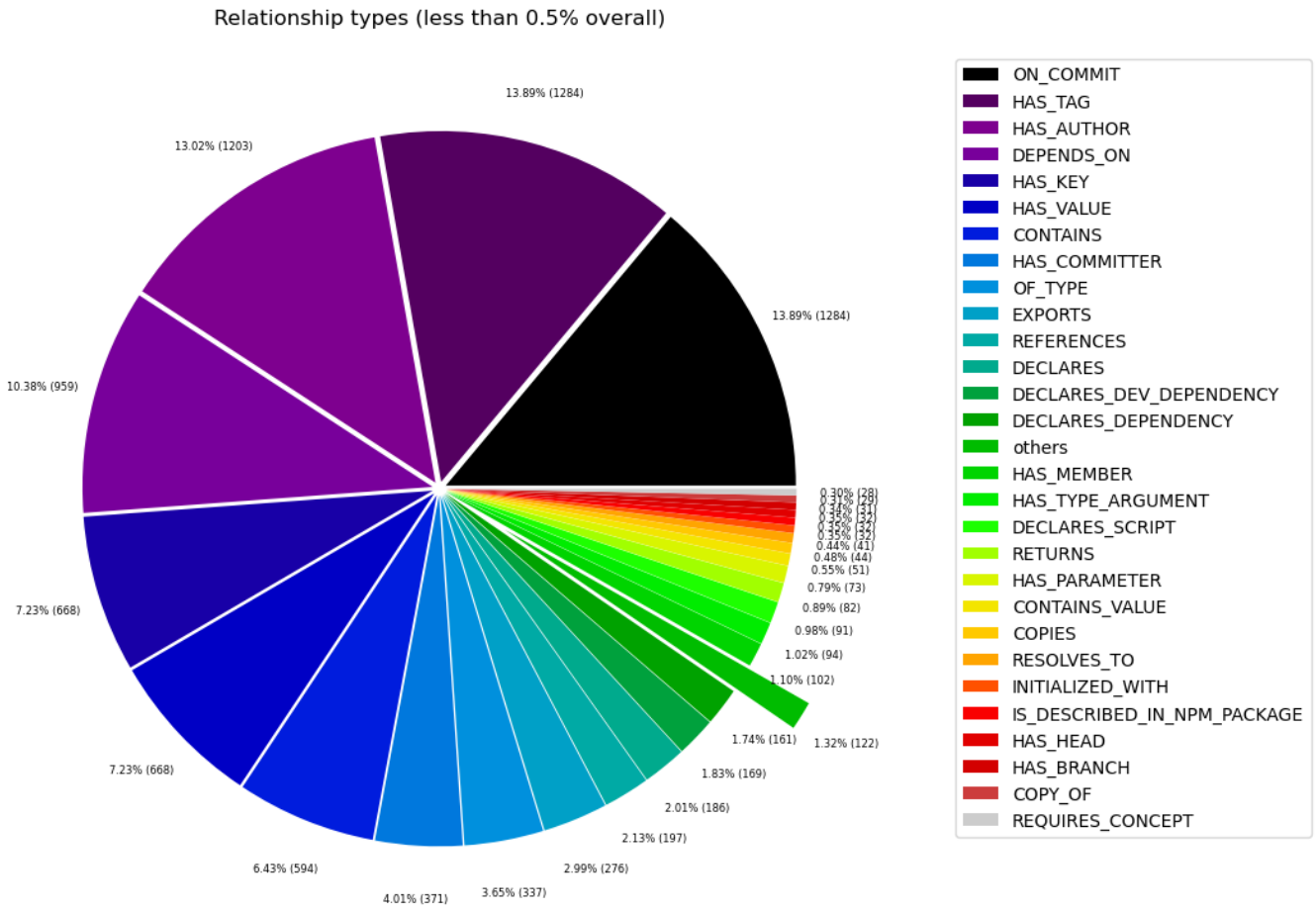
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	PROVIDED_BY_NPM_DEPENDENCY	1	0.000325
1	IS_IMPLEMENTED_IN	2	0.000649
2	CONSTRAINED_BY	4	0.001298
3	REFERENCED_PROJECTS	5	0.001623
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	0.006168
16	USES	25	0.008116
17	REQUIRES_CONCEPT	28	0.009089
18	COPY_OF	29	0.009414
19	HAS_BRANCH	31	0.010063
20	IS_DESCRIBED_IN_NPM_PACKAGE	32	0.010388
21	INITIALIZED_WITH	32	0.010388
22	HAS_HEAD	32	0.010388
23	RESOLVES_TO	41	0.013310
24	COPIES	44	0.014283
25	CONTAINS_VALUE	51	0.016556
26	HAS_PARAMETER	73	0.023698
27	RETURNS	82	0.026619
28	DECLARES_SCRIPT	91	0.029541
29	HAS_TYPE_ARGUMENT	94	0.030515

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, Change]	MODIFIES	[File, Git]	80737	80737	
1	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	80737	10512	
2	[Git, Change]	UPDATES	[File, Git]	52795	80737	
3	[Git, Change]	CREATES	[File, Git]	19562	80737	
4	[Git, Commit]	HAS_PARENT	[Git, Commit]	11552	10512	
5	[Git, Change]	DELETES	[File, Git]	11551	80737	
6	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	10512	1	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	10512	1203	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10512	371	
9	[Repository, File, Git]	HAS_FILE	[File, Git]	5448	1	
10	[Git, Change]	RENAMES	[File, Git]	3171	80737	
11	[File, Git]	HAS_NEW_NAME	[File, Git]	1718	5448	
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): 103138
total_number_of_relationships (edges): 308048
-> total directed graph density: 2.89591088454182e-05
-> total directed graph density in percent: 0.0028959108845418202
```