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]	76390	77.941026
1	[Git, Commit]	10182	10.388736
2	[File, Git]	5185	5.290277
3	[Author, Git, Person]	1188	1.212121
4	[Git, Tag]	1140	1.163147
5	[Json, Key]	668	0.681563
6	[Json, Value, Scalar]	603	0.615243
7	[Committer, Git, Person]	371	0.378533
8	[NPM, Dependency]	330	0.336700
9	[Type, TS, Primitive, ExternalType]	285	0.290787
10	[Type, TS, Declared, ExternalType]	272	0.277523
11	[TS, ExternalDeclaration]	215	0.219365
12	[Type, TS, Literal, ExternalType]	136	0.138761
13	[Json, Value, Object]	133	0.135700
14	[Type, TS, Union, ExternalType]	117	0.119376
15	[Type, TS, ObjectMember, ExternalType]	98	0.099990
16	[NPM, Script]	91	0.092848
17	[TS, Property]	65	0.066320
18	[TS, Function]	47	0.047954
19	[Type, Object, TS, ExternalType]	38	0.038772
20	[Type, TS, FunctionParameter, ExternalType]	37	0.037751
21	[File, Directory]	34	0.034690
22	[TS, Parameter]	33	0.033670
23	[Type, TS, Function, ExternalType]	32	0.032650
24	[Package, File, Json, NPM]	29	0.029589
25	[TS, ExternalModule]	25	0.025508
26	[Git, Branch]	25	0.025508
27	[TS, Variable]	24	0.024487
28	[Value, TS, Literal]	20	0.020406
29	[jQAssistant, Rule, Concept]	19	0.019386

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>
```

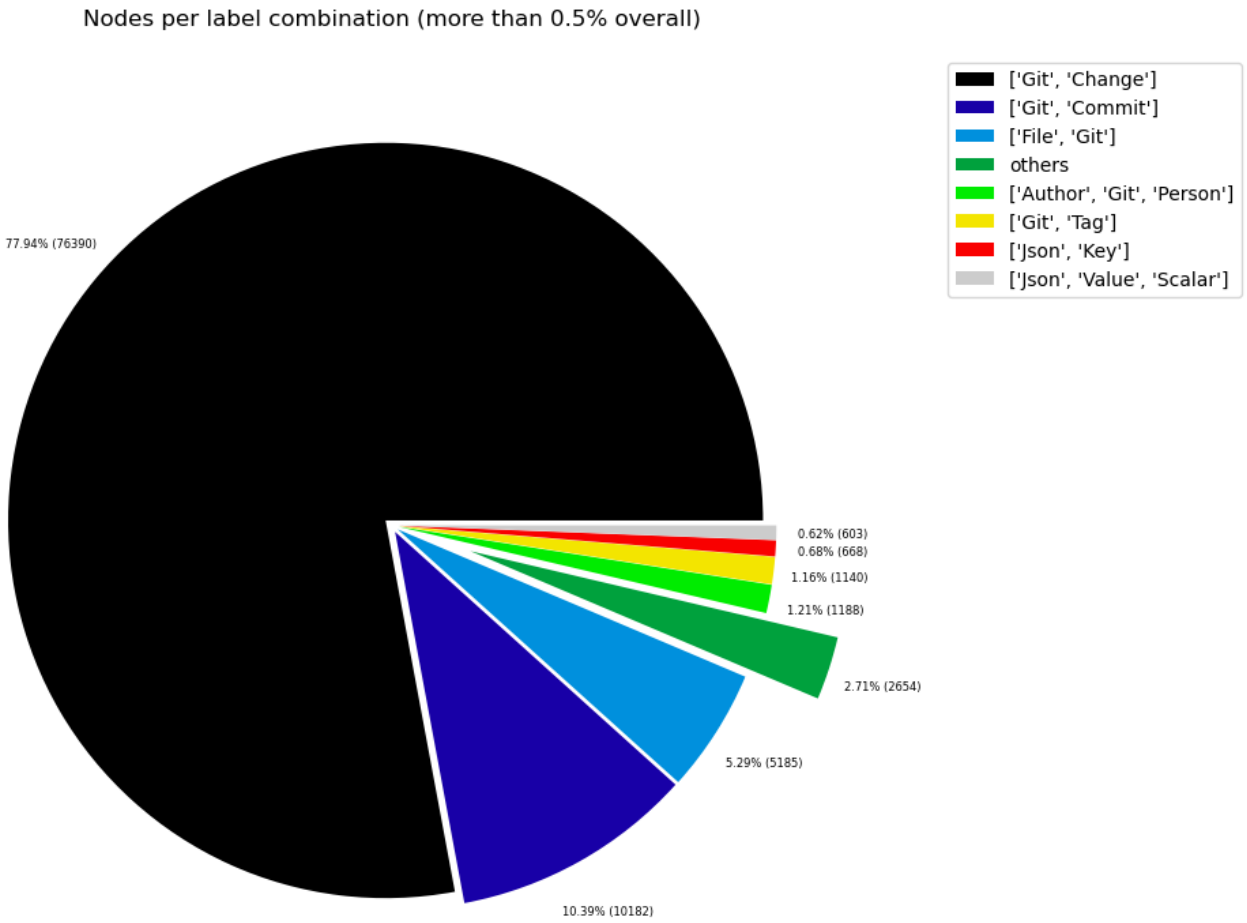


Table 1b - Lowest node count by label combination

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	0.001020
1	[File, TS, Scan]	1	0.001020
2	[TS, Method]	1	0.001020
3	[Value, TS, ObjectMember]	1	0.001020
4	[TS, Constructor]	1	0.001020
5	[TS, Class]	1	0.001020
6	[TS, Enum]	2	0.002041
7	[Value, Object, TS]	3	0.003061
8	[Type, TS, Tuple, ExternalType]	3	0.003061
9	[Value, TS, Function]	4	0.004081
10	[TS, TypeParameter]	4	0.004081
11	[Value, TS, Complex]	5	0.005102
12	[NPM, Engine]	6	0.006122
13	[Project, TS]	6	0.006122
14	[File, Local]	6	0.006122
15	[Value, TS, Call]	6	0.006122
16	[Value, TS, Member]	6	0.006122
17	[File, TS, Local, Module]	6	0.006122
18	[Type, TS, TypeParameterReference, ExternalType]	6	0.006122
19	[TS, EnumMember]	8	0.008162
20	[Type, TS, NotIdentified, ExternalType]	11	0.011223
21	[Json, Value, Array]	12	0.012244
22	[Value, TS, Declared]	13	0.013264
23	[TS, TypeAlias]	14	0.014284
24	[File, Directory, Local]	16	0.016325
25	[Type, TS, Intersection, ExternalType]	17	0.017345
26	[TS, Interface]	18	0.018365
27	[jQAssistant, Rule, Concept]	19	0.019386
28	[Value, TS, Literal]	20	0.020406
29	[TS, Variable]	24	0.024487

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>
```

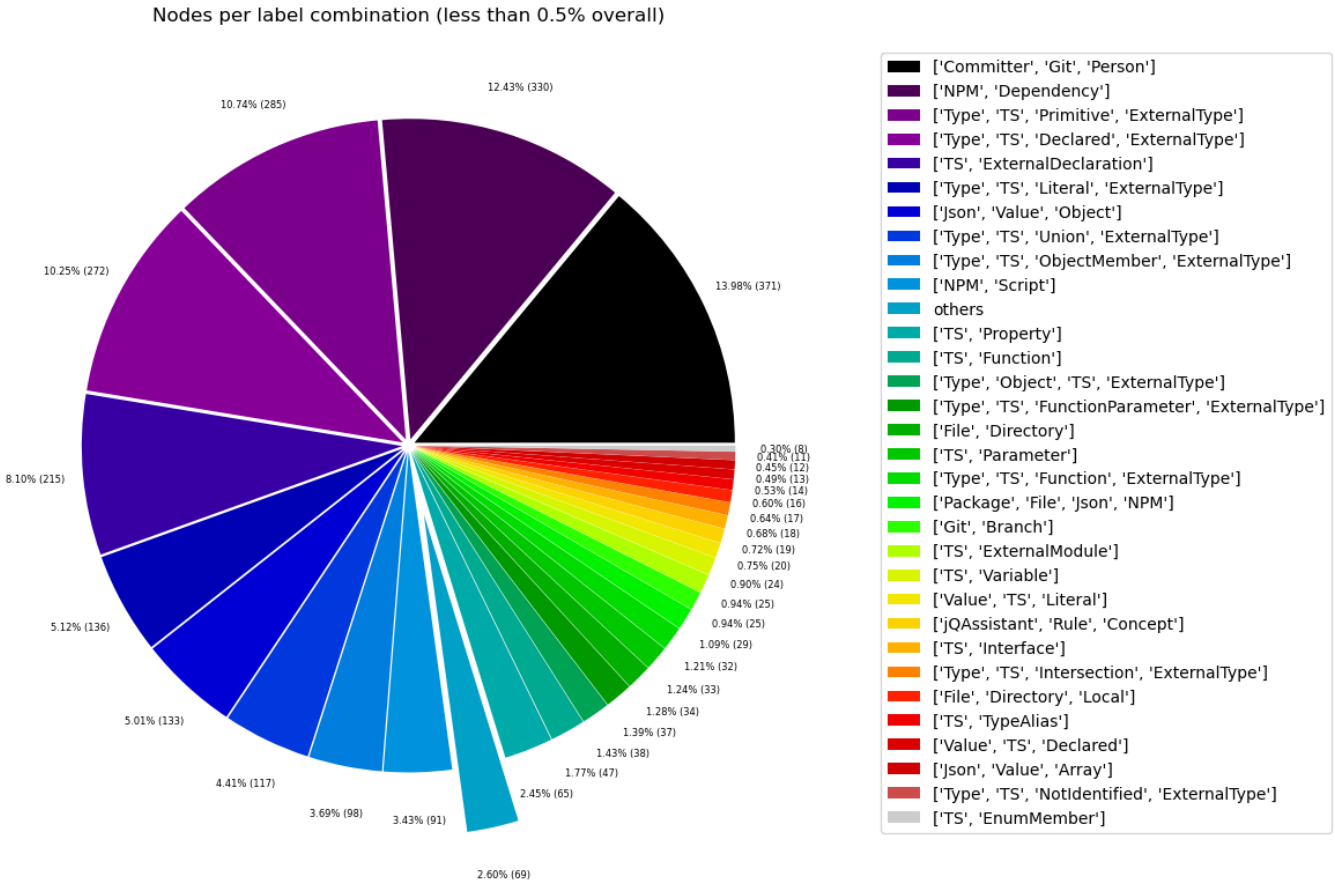


Table 1c - Highest node count by single label

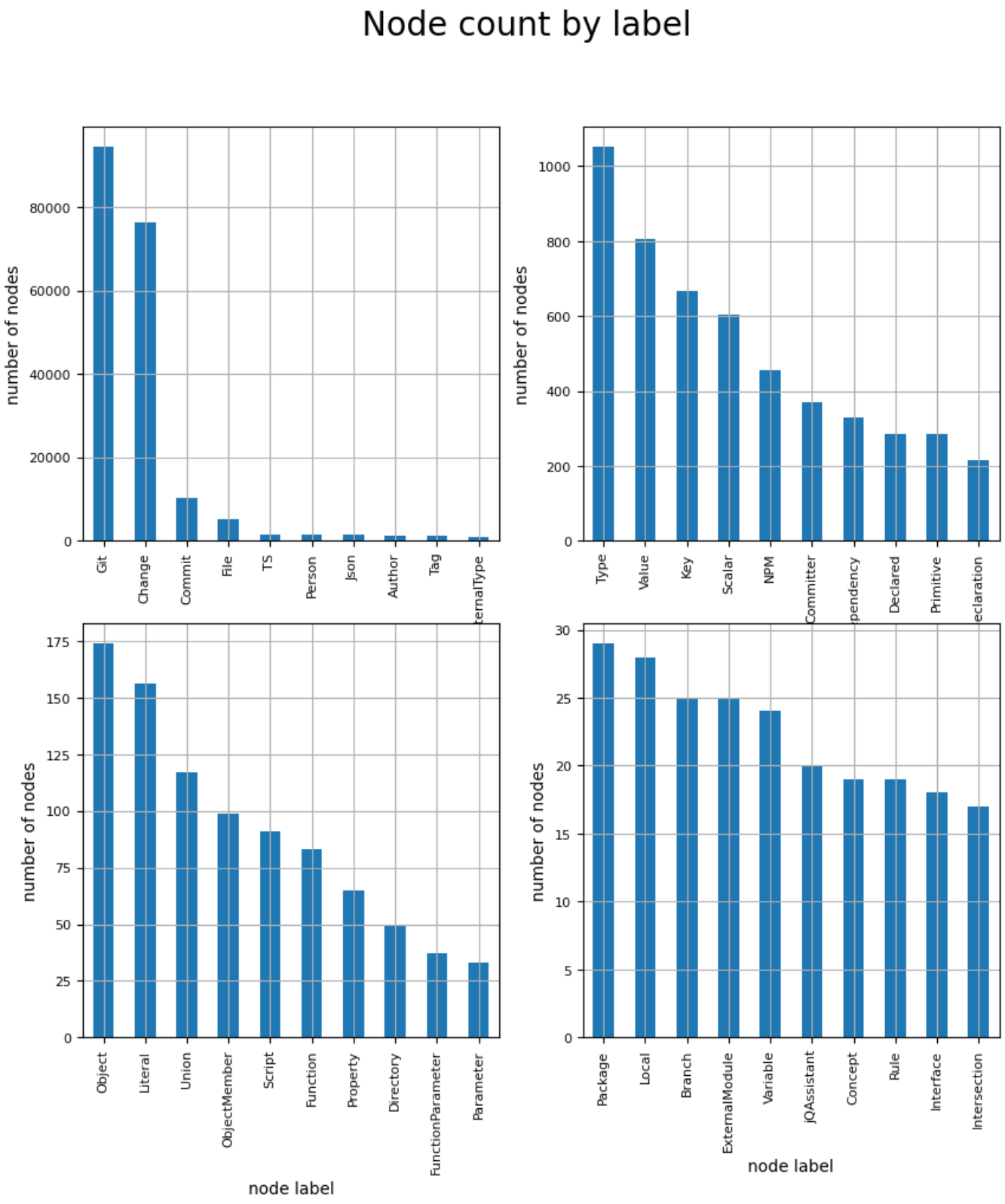
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	94481	96.399347
1	Change	76390	77.941026
2	Commit	10182	10.388736
3	File	5277	5.384144
4	TS	1581	1.613101
5	Person	1559	1.590654
6	Json	1445	1.474339
7	Author	1188	1.212121
8	Tag	1140	1.163147
9	ExternalType	1052	1.073360
10	Type	1052	1.073360
11	Value	806	0.822365
12	Key	668	0.681563
13	Scalar	603	0.615243
14	NPM	456	0.465259
15	Committer	371	0.378533
16	Dependency	330	0.336700
17	Declared	285	0.290787
18	Primitive	285	0.290787
19	ExternalDeclaration	215	0.219365
20	Object	174	0.177533
21	Literal	156	0.159167
22	Union	117	0.119376
23	ObjectMember	99	0.101010
24	Script	91	0.092848
25	Function	83	0.084685
26	Property	65	0.066320
27	Directory	50	0.051015
28	FunctionParameter	37	0.037751
29	Parameter	33	0.033670
30	Package	29	0.029589
31	Local	28	0.028569
32	Branch	25	0.025508
33	ExternalModule	25	0.025508
34	Variable	24	0.024487
35	jQAssistant	20	0.020406
36	Concept	19	0.019386
37	Rule	19	0.019386
38	Interface	18	0.018365
39	Intersection	17	0.017345

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: 274065
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	76390	27.872950
1	MODIFIES	76390	27.872950
2	UPDATES	50425	18.398920
3	COMMITTED	20364	7.430354
4	CREATES	18178	6.632733
5	HAS_PARENT	11210	4.090271
6	DELETES	10568	3.856020
7	RENAMES	2781	1.014723
8	HAS_NEW_NAME	1572	0.573587
9	ON_COMMIT	1140	0.415960
10	DEPENDS_ON	962	0.351012
11	HAS_KEY	668	0.243738
12	HAS_VALUE	668	0.243738
13	CONTAINS	589	0.214913
14	OF_TYPE	329	0.120045
15	EXPORTS	275	0.100341
16	REFERENCES	196	0.071516
17	DECLARES	185	0.067502
18	DECLARES_DEV_DEPENDENCY	169	0.061664
19	DECLARES_DEPENDENCY	161	0.058745
20	HAS_MEMBER	99	0.036123
21	HAS_TYPE_ARGUMENT	92	0.033569
22	DECLARES_SCRIPT	91	0.033204
23	RESOLVES_TO	81	0.029555
24	RETURNS	80	0.029190
25	HAS_PARAMETER	70	0.025541
26	CONTAINS_VALUE	51	0.018609
27	COPIES	43	0.015690
28	INITIALIZED_WITH	32	0.011676
29	COPY_OF	28	0.010217

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>
```

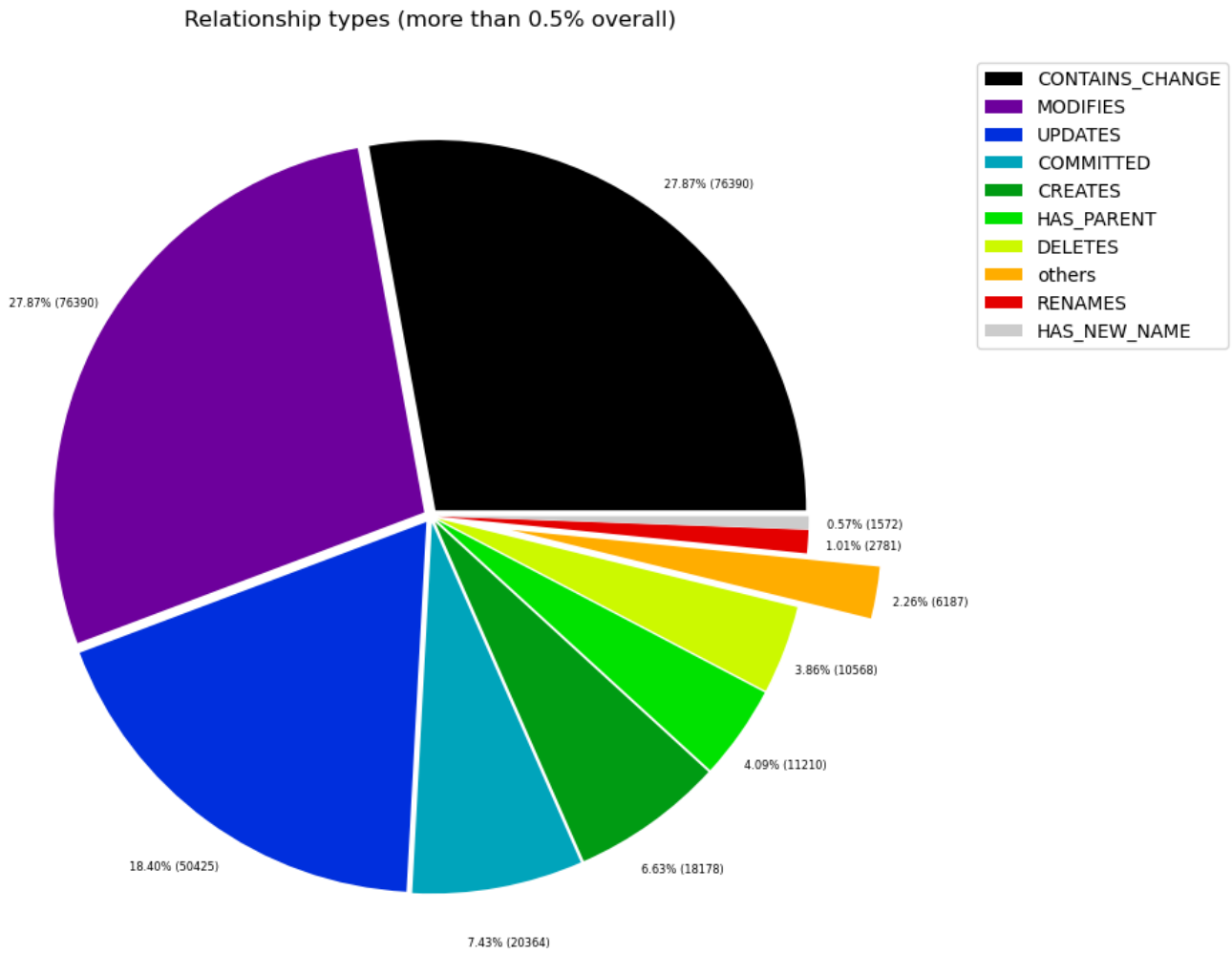


Table 2b - Lowest relationship count by type

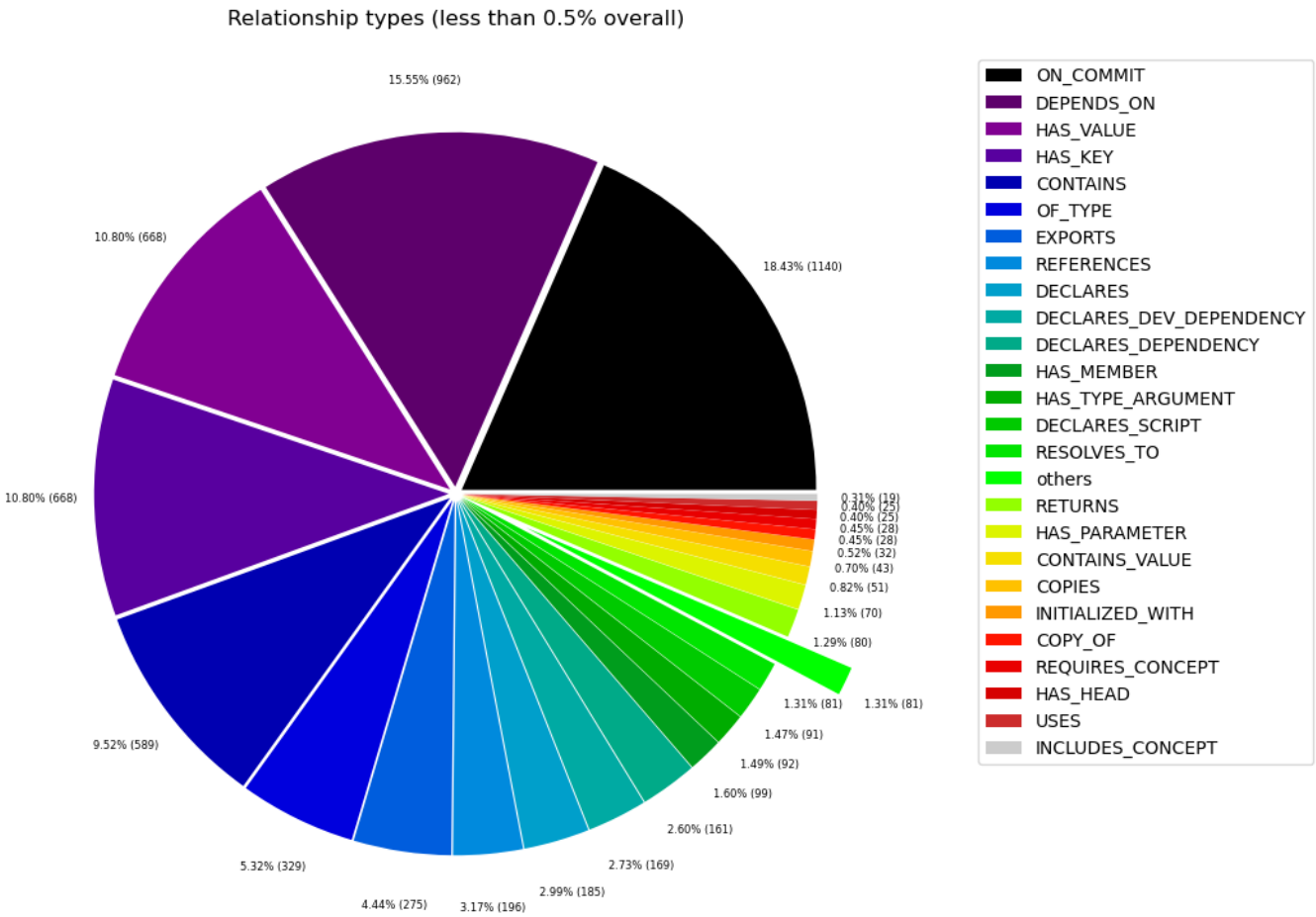
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	PROVIDED_BY_NPM_DEPENDENCY	1	0.000365
1	CONSTRAINED_BY	4	0.001460
2	REFERENCED_PROJECTS	5	0.001824
3	PARENT	6	0.002189
4	MEMBER	6	0.002189
5	HAS_ROOT	6	0.002189
6	HAS_NPM_PACKAGE	6	0.002189
7	HAS_CONFIG	6	0.002189
8	HAS_ARGUMENT	6	0.002189
9	DECLARES_ENGINE	6	0.002189
10	CONTAINS_PROJECT	6	0.002189
11	CALLS	6	0.002189
12	EXTENDS	7	0.002554
13	SIMILAR	10	0.003649
14	INCLUDES_CONCEPT	19	0.006933
15	USES	25	0.009122
16	HAS_HEAD	25	0.009122
17	REQUIRES_CONCEPT	28	0.010217
18	COPY_OF	28	0.010217
19	INITIALIZED_WITH	32	0.011676
20	COPIES	43	0.015690
21	CONTAINS_VALUE	51	0.018609
22	HAS_PARAMETER	70	0.025541
23	RETURNS	80	0.029190
24	RESOLVES_TO	81	0.029555
25	DECLARES_SCRIPT	91	0.033204
26	HAS_TYPE_ARGUMENT	92	0.033569
27	HAS_MEMBER	99	0.036123
28	DECLARES_DEPENDENCY	161	0.058745
29	DECLARES_DEV_DEPENDENCY	169	0.061664

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]	76390	10182	
1	[Git, Change]	MODIFIES	[File, Git]	76390	76390	
2	[Git, Change]	UPDATES	[File, Git]	50425	76390	
3	[Git, Change]	CREATES	[File, Git]	18178	76390	
4	[Git, Commit]	HAS_PARENT	[Git, Commit]	11210	10182	
5	[Git, Change]	DELETES	[File, Git]	10568	76390	
6	[Author, Git, Person]	COMMITTED	[Git, Commit]	10182	1188	
7	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10182	371	
8	[Git, Change]	RENAMES	[File, Git]	2781	76390	
9	[File, Git]	HAS_NEW_NAME	[File, Git]	1572	5185	
10	[Git, Tag]	ON_COMMIT	[Git, Commit]	1140	1140	
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	3	
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, ExternalType]	CONTAINS	[Type, TS, Primitive, ExternalType]	144	117	
19	[Type, TS, Declared, ExternalType]	REFERENCES	[TS, ExternalDeclaration]	141	272	
20	[TS, Function]	DEPENDS_ON	[TS, ExternalModule]	131	47	
21	[Type, TS, Union, ExternalType]	CONTAINS	[Type, TS, Literal, ExternalType]	119	117	
22	[Json, Key]	HAS_VALUE	[Json, Value, Object]	104	668	
23	[Type, Object, TS, ExternalType]	HAS_MEMBER	[Type, TS, ObjectMember, ExternalType]	98	38	
24	[Package, File, Json, NPM]	DECLARES_SCRIPT	[NPM, Script]	91	29	
25	[Type, TS, Union, ExternalType]	CONTAINS	[Type, TS, Declared, ExternalType]	69	117	
26	[File, Directory]	CONTAINS	[File, Directory]	63	34	
27	[TS, Interface]	DECLARES	[TS, Property]	61	18	
28	[File, Directory]	CONTAINS	[Package, File, Json, NPM]	58	34	
29	[File, Git]	RESOLVES_TO	[Package, File, Json, NPM]	57	5185	

Graph Density

```
total_number_of_nodes (vertices): 98010
total_number_of_relationships (edges): 274065
-> total directed graph density: 2.8531015481193525e-05
-> total directed graph density in percent: 0.0028531015481193527
```