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]	82209	78.328601
1	[Git, Commit]	10677	10.173028
2	[File, Git]	5503	5.243249
3	[Git, Tag]	1373	1.308192
4	[Author, Git, Person]	1228	1.170036
5	[Json, Key]	668	0.636469
6	[Json, Value, Scalar]	603	0.574537
7	[Committer, Git, Person]	371	0.353488
8	[NPM, Dependency]	338	0.322046
9	[Type, TS, Primitive]	291	0.277264
10	[Type, TS, Declared]	276	0.262972
11	[TS, ExternalDeclaration]	214	0.203899
12	[Type, TS, Literal]	136	0.129581
13	[Json, Value, Object]	133	0.126722
14	[Type, TS, Union]	119	0.113383
15	[Type, TS, ObjectMember]	101	0.096233
16	[NPM, Script]	91	0.086705
17	[TS, Property]	65	0.061932
18	[TS, Function]	47	0.044782
19	[Type, TS, FunctionParameter]	40	0.038112
20	[Type, Object, TS]	39	0.037159
21	[Git, Branch]	36	0.034301
22	[File, Directory]	34	0.032395
23	[Type, TS, Function]	34	0.032395
24	[TS, Parameter]	33	0.031442
25	[Package, File, Json, NPM]	29	0.027631
26	[TS, Variable]	24	0.022867
27	[TS, ExternalModule]	23	0.021914
28	[Value, TS, Literal]	20	0.019056
29	[jQAssistant, Rule, Concept]	19	0.018103

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>
```

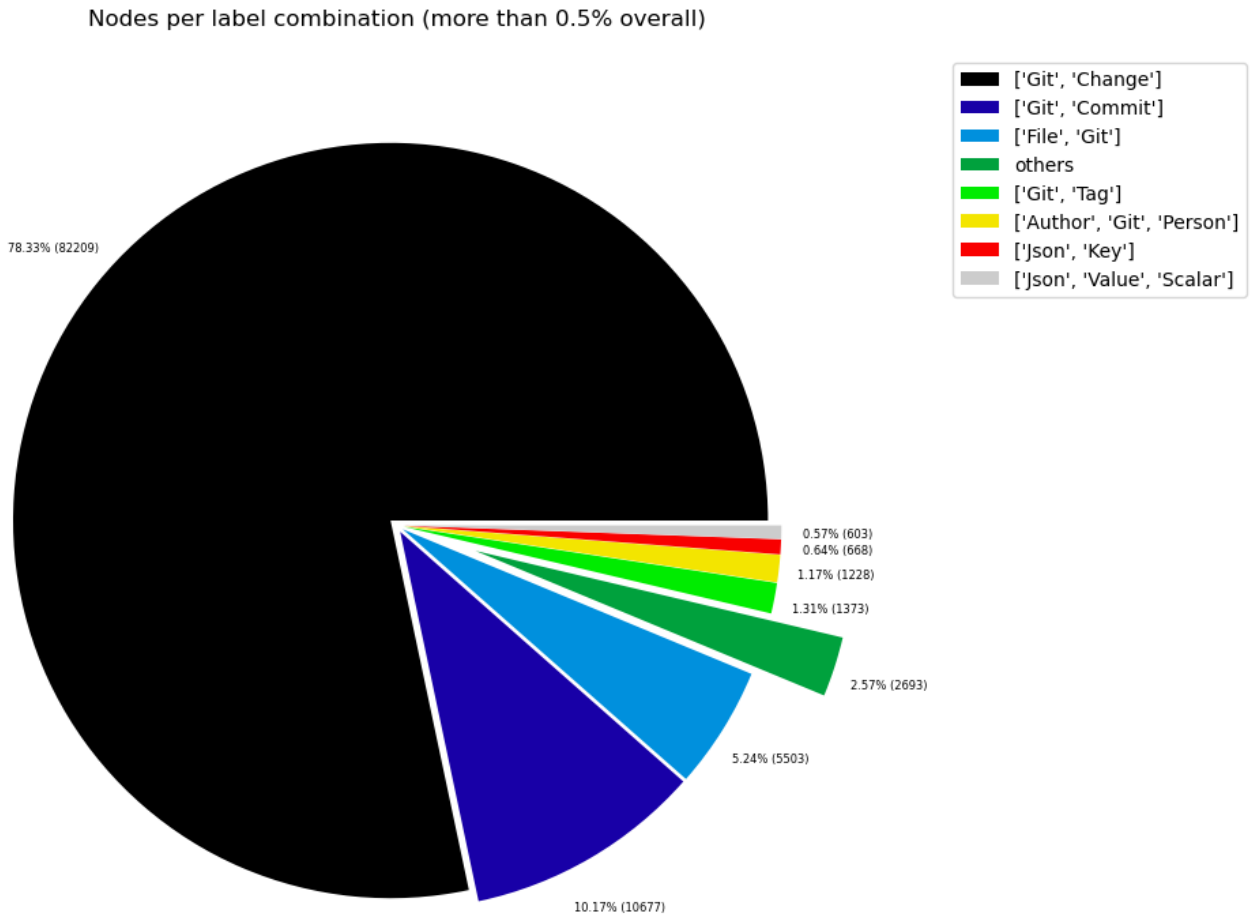


Table 1b - Lowest node count by label combination

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	0.000953
1	[File, TS, Scan]	1	0.000953
2	[TS, Method]	1	0.000953
3	[Repository, File, Git]	1	0.000953
4	[TS, Constructor]	1	0.000953
5	[Value, TS, ObjectMember]	1	0.000953
6	[TS, Class]	1	0.000953
7	[TS, Enum]	2	0.001906
8	[Value, Object, TS]	3	0.002858
9	[Type, TS, Tuple]	3	0.002858
10	[Value, TS, Function]	4	0.003811
11	[TS, TypeParameter]	4	0.003811
12	[Value, TS, Complex]	5	0.004764
13	[NPM, Engine]	6	0.005717
14	[Project, TS]	6	0.005717
15	[File, Local]	6	0.005717
16	[Value, TS, Call]	6	0.005717
17	[Value, TS, Member]	6	0.005717
18	[File, TS, Local, Module]	6	0.005717
19	[Type, TS, TypeParameterReference]	6	0.005717
20	[TS, EnumMember]	8	0.007622
21	[Type, TS, NotIdentified]	11	0.010481
22	[Json, Value, Array]	12	0.011434
23	[Value, TS, Declared]	13	0.012386
24	[TS, TypeAlias]	16	0.015245
25	[File, Directory, Local]	16	0.015245
26	[TS, Interface]	17	0.016198
27	[Type, TS, Intersection]	17	0.016198
28	[jQAssistant, Rule, Concept]	19	0.018103
29	[Value, TS, Literal]	20	0.019056

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>
```

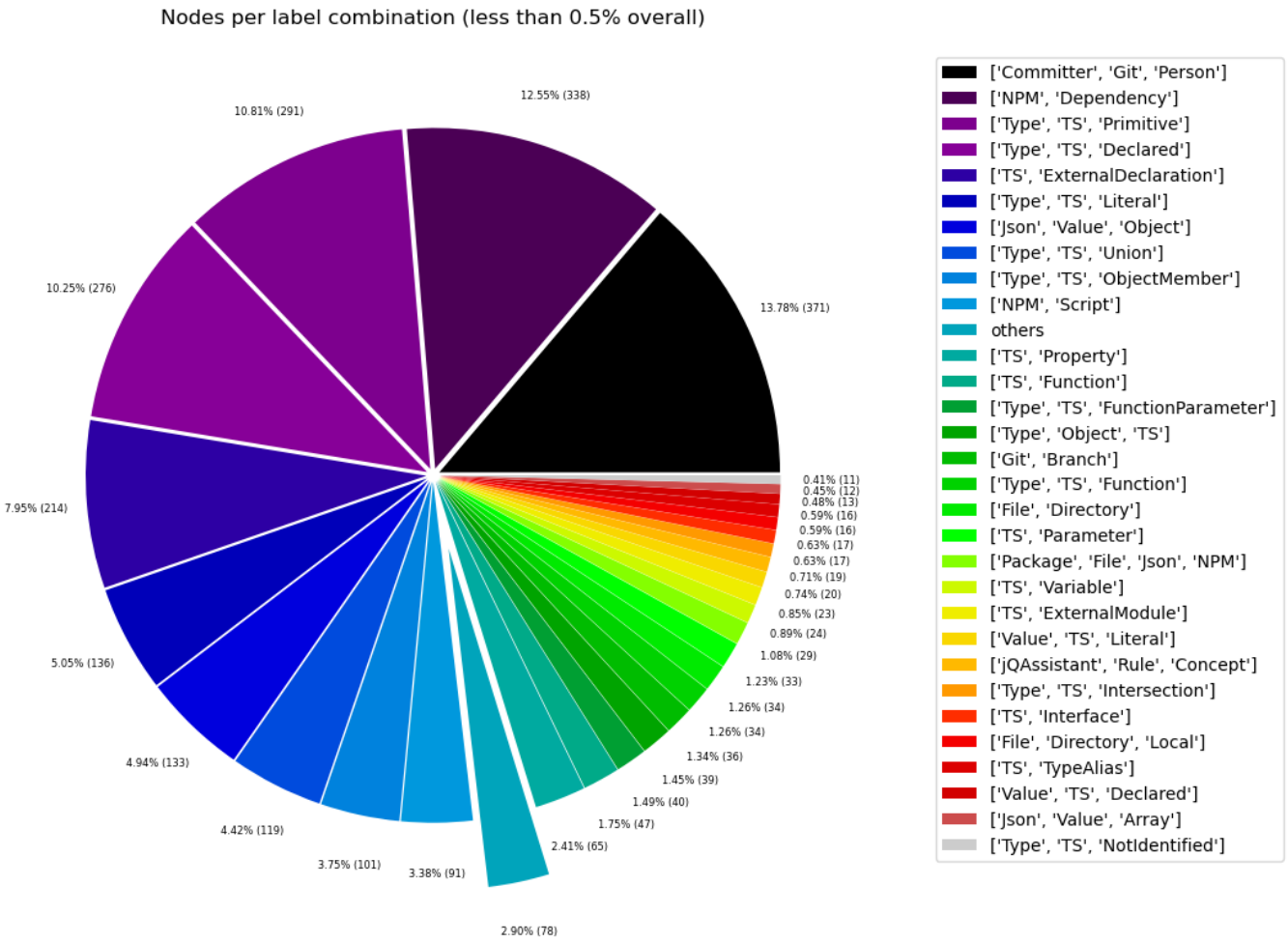


Table 1c - Highest node count by single label

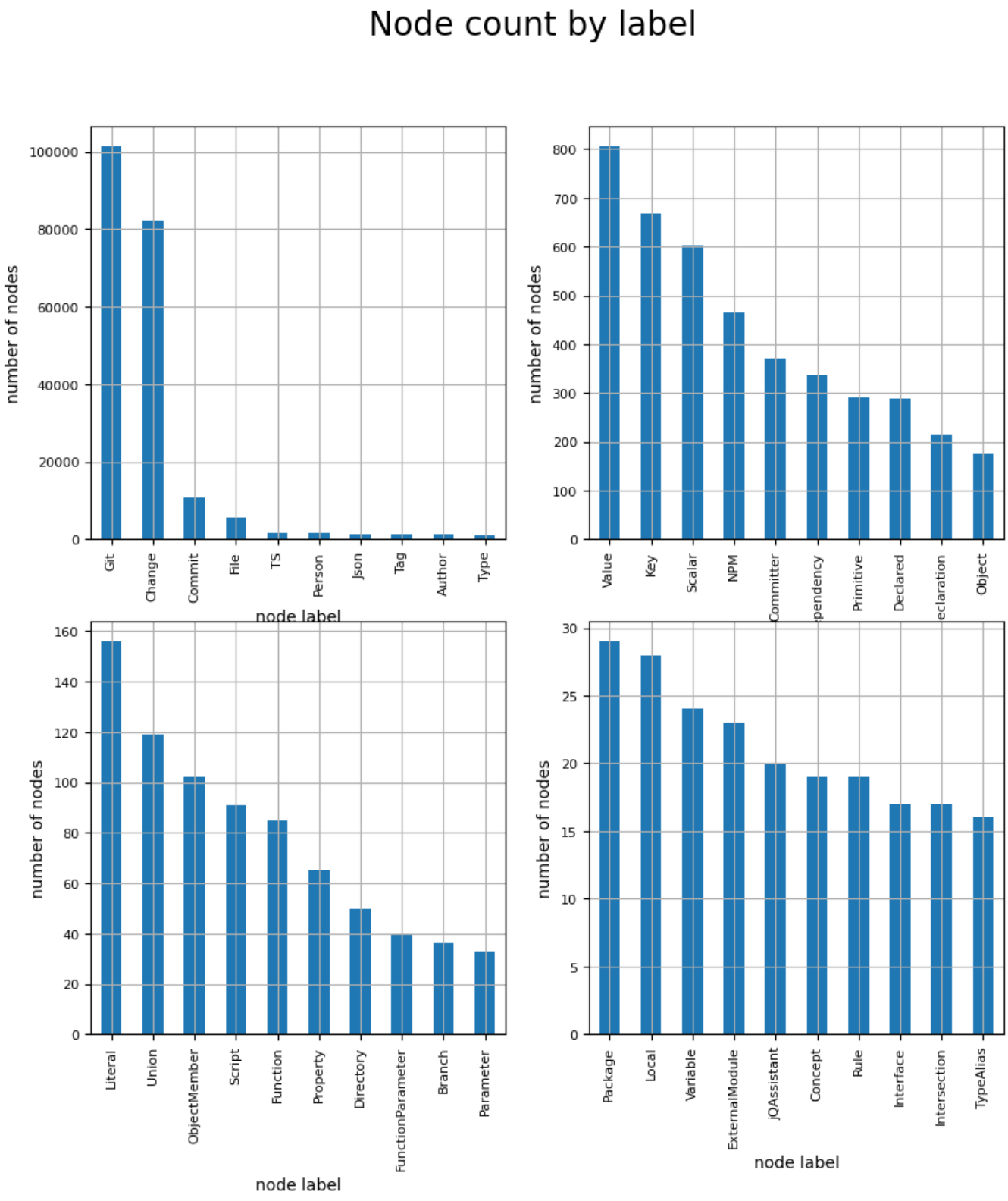
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	101398	96.611849
1	Change	82209	78.328601
2	Commit	10677	10.173028
3	File	5596	5.331860
4	TS	1600	1.524477
5	Person	1599	1.523525
6	Json	1445	1.376794
7	Tag	1373	1.308192
8	Author	1228	1.170036
9	Type	1073	1.022353
10	Value	806	0.767955
11	Key	668	0.636469
12	Scalar	603	0.574537
13	NPM	464	0.442098
14	Committer	371	0.353488
15	Dependency	338	0.322046
16	Primitive	291	0.277264
17	Declared	289	0.275359
18	ExternalDeclaration	214	0.203899
19	Object	175	0.166740
20	Literal	156	0.148637
21	Union	119	0.113383
22	ObjectMember	102	0.097185
23	Script	91	0.086705
24	Function	85	0.080988
25	Property	65	0.061932
26	Directory	50	0.047640
27	FunctionParameter	40	0.038112
28	Branch	36	0.034301
29	Parameter	33	0.031442
30	Package	29	0.027631
31	Local	28	0.026678
32	Variable	24	0.022867
33	ExternalModule	23	0.021914
34	jQAssistant	20	0.019056
35	Concept	19	0.018103
36	Rule	19	0.018103
37	Interface	17	0.016198
38	Intersection	17	0.016198
39	TypeAlias	16	0.015245

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: 313514
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	82209	26.221796
1	MODIFIES	82209	26.221796
2	UPDATES	53922	17.199232
3	COMMITTED	21354	6.811179
4	CREATES	19714	6.288076
5	DELETES	11773	3.755175
6	HAS_PARENT	11728	3.740822
7	HAS_COMMIT	10677	3.405590
8	HAS_FILE	5503	1.755265
9	RENAMES	3200	1.020688
10	HAS_NEW_NAME	1727	0.550853
11	HAS_TAG	1373	0.437939
12	ON_COMMIT	1373	0.437939
13	HAS_AUTHOR	1228	0.391689
14	DEPENDS_ON	961	0.306525
15	HAS_KEY	668	0.213069
16	HAS_VALUE	668	0.213069
17	CONTAINS	594	0.189465
18	HAS_COMMITTER	371	0.118336
19	OF_TYPE	337	0.107491
20	EXPORTS	283	0.090267
21	REFERENCES	197	0.062836
22	DECLARES	186	0.059327
23	DECLARES_DEV_DEPENDENCY	169	0.053905
24	DECLARES_DEPENDENCY	161	0.051353
25	HAS_MEMBER	102	0.032534
26	HAS_TYPE_ARGUMENT	94	0.029983
27	DECLARES_SCRIPT	91	0.029026
28	RETURNS	82	0.026155
29	HAS_PARAMETER	73	0.023284

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>
```

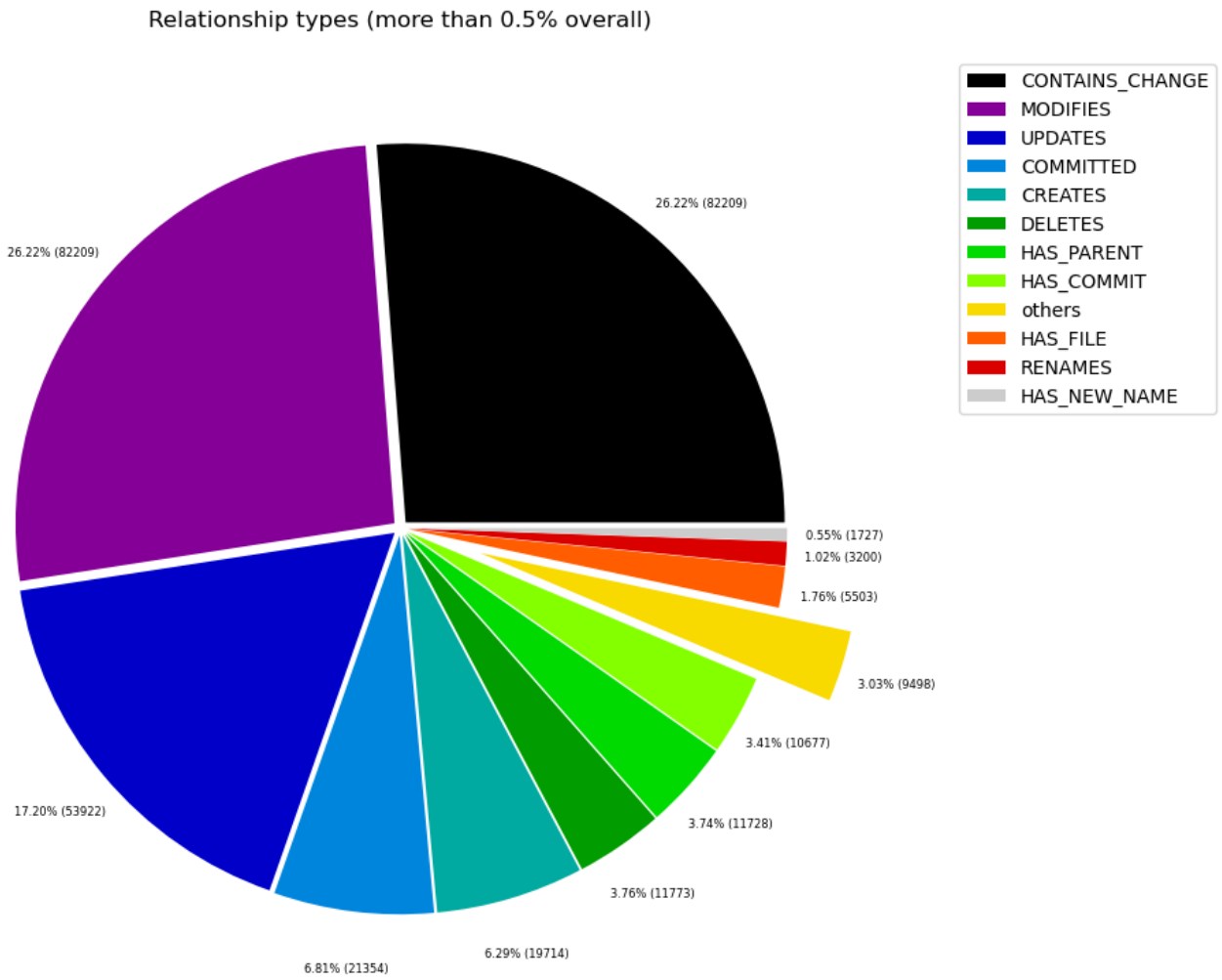


Table 2b - Lowest relationship count by type

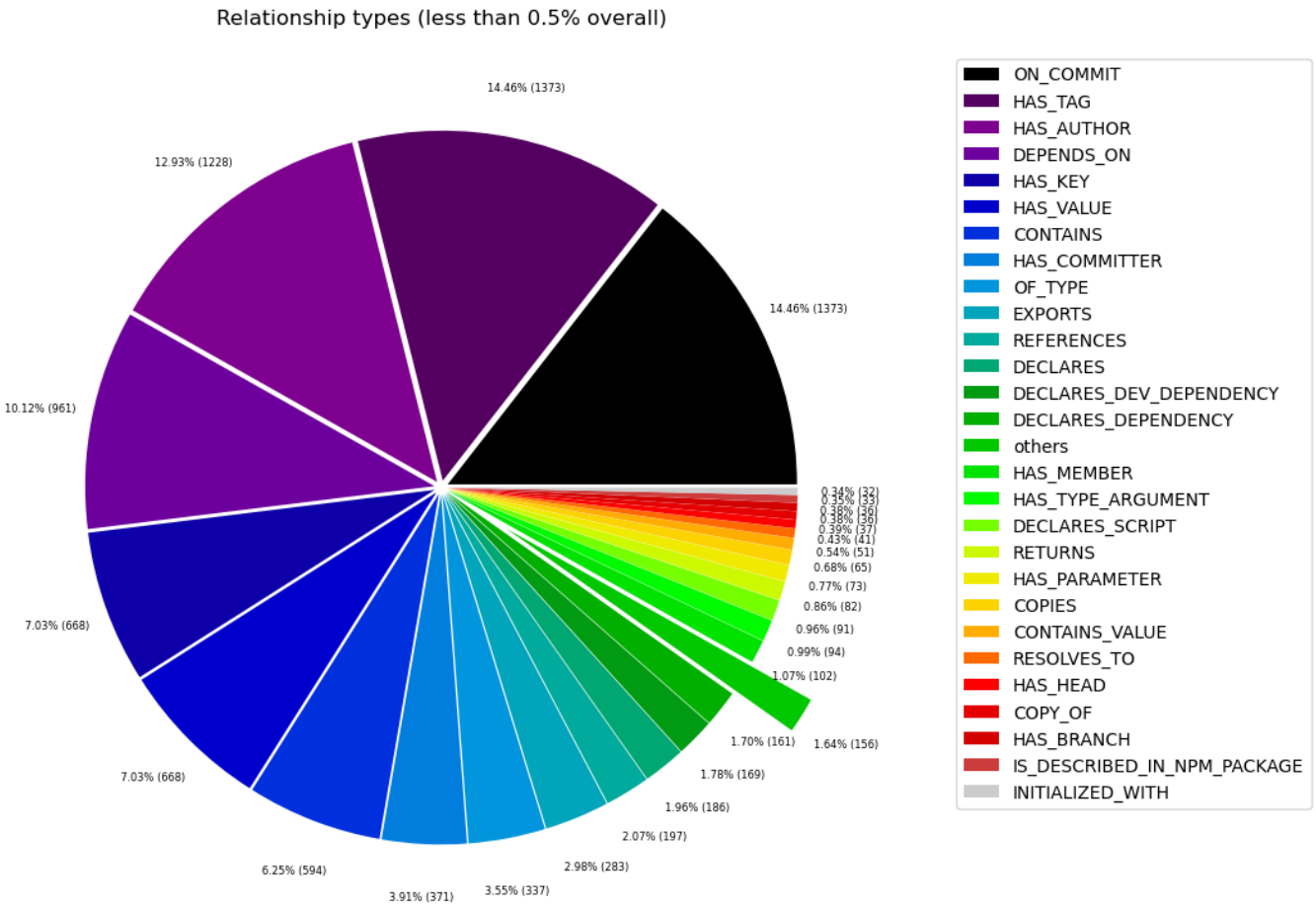
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	PROVIDED_BY_NPM_DEPENDENCY	1	0.000319
1	IS_IMPLEMENTED_IN	2	0.000638
2	CONSTRAINED_BY	4	0.001276
3	REFERENCED_PROJECTS	5	0.001595
4	CONTAINS_PROJECT	6	0.001914
5	DECLARES_ENGINE	6	0.001914
6	EXTENDS	6	0.001914
7	HAS_ARGUMENT	6	0.001914
8	CALLS	6	0.001914
9	HAS_NPM_PACKAGE	6	0.001914
10	HAS_ROOT	6	0.001914
11	MEMBER	6	0.001914
12	PARENT	6	0.001914
13	HAS_CONFIG	6	0.001914
14	SIMILAR	6	0.001914
15	DECLARES_PEER_DEPENDENCY	8	0.002552
16	INCLUDES_CONCEPT	19	0.006060
17	USES	23	0.007336
18	REQUIRES_CONCEPT	28	0.008931
19	INITIALIZED_WITH	32	0.010207
20	IS_DESCRIBED_IN_NPM_PACKAGE	33	0.010526
21	HAS_BRANCH	36	0.011483
22	COPY_OF	36	0.011483
23	HAS_HEAD	37	0.011802
24	RESOLVES_TO	41	0.013078
25	CONTAINS_VALUE	51	0.016267
26	COPIES	65	0.020733
27	HAS_PARAMETER	73	0.023284
28	RETURNS	82	0.026155
29	DECLARES_SCRIPT	91	0.029026

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]	82209	82209	
1	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	82209	10677	
2	[Git, Change]	UPDATES	[File, Git]	53922	82209	
3	[Git, Change]	CREATES	[File, Git]	19714	82209	
4	[Git, Change]	DELETES	[File, Git]	11773	82209	
5	[Git, Commit]	HAS_PARENT	[Git, Commit]	11728	10677	
6	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	10677	1	
7	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10677	371	
8	[Author, Git, Person]	COMMITTED	[Git, Commit]	10677	1228	
9	[Repository, File, Git]	HAS_FILE	[File, Git]	5503	1	
10	[Git, Change]	RENAMES	[File, Git]	3200	82209	
11	[File, Git]	HAS_NEW_NAME	[File, Git]	1727	5503	
12	[Repository, File, Git]	HAS_TAG	[Git, Tag]	1373	1	
13	[Git, Tag]	ON_COMMIT	[Git, Commit]	1373	1373	
14	[Repository, File, Git]	HAS_AUTHOR	[Author, Git, Person]	1228	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): 104954
total_number_of_relationships (edges): 313514
-> total directed graph density: 2.846184747447383e-05
-> total directed graph density in percent: 0.002846184747447383
```