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]	82069	78.318335
1	[Git, Commit]	10666	10.178549
2	[File, Git]	5502	5.250551
3	[Git, Tag]	1360	1.297846
4	[Author, Git, Person]	1227	1.170924
5	[Json, Key]	668	0.637471
6	[Json, Value, Scalar]	603	0.575442
7	[Committer, Git, Person]	371	0.354045
8	[NPM, Dependency]	338	0.322553
9	[Type, TS, Primitive]	291	0.277701
10	[Type, TS, Declared]	276	0.263386
11	[TS, ExternalDeclaration]	214	0.204220
12	[Type, TS, Literal]	136	0.129785
13	[Json, Value, Object]	133	0.126922
14	[Type, TS, Union]	119	0.113562
15	[Type, TS, ObjectMember]	101	0.096384
16	[NPM, Script]	91	0.086841
17	[TS, Property]	65	0.062029
18	[TS, Function]	47	0.044852
19	[Type, TS, FunctionParameter]	40	0.038172
20	[Type, Object, TS]	39	0.037218
21	[Git, Branch]	37	0.035309
22	[File, Directory]	34	0.032446
23	[Type, TS, Function]	34	0.032446
24	[TS, Parameter]	33	0.031492
25	[Package, File, Json, NPM]	29	0.027675
26	[TS, Variable]	24	0.022903
27	[TS, ExternalModule]	23	0.021949
28	[Value, TS, Literal]	20	0.019086
29	[jQAssistant, Rule, Concept]	19	0.018132

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>
```

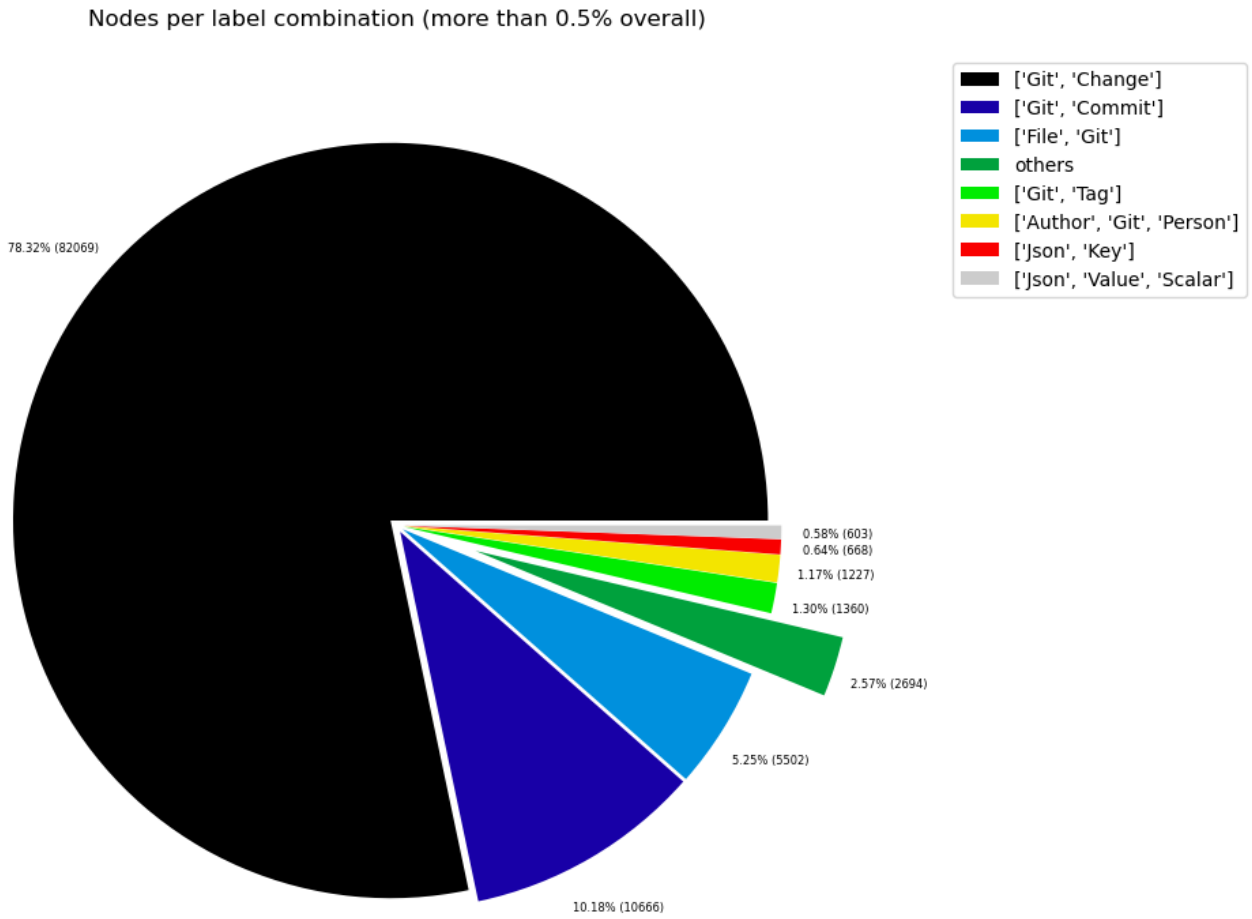


Table 1b - Lowest node count by label combination

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	0.000954
1	[File, TS, Scan]	1	0.000954
2	[TS, Method]	1	0.000954
3	[Repository, File, Git]	1	0.000954
4	[TS, Constructor]	1	0.000954
5	[Value, TS, ObjectMember]	1	0.000954
6	[TS, Class]	1	0.000954
7	[TS, Enum]	2	0.001909
8	[Value, Object, TS]	3	0.002863
9	[Type, TS, Tuple]	3	0.002863
10	[Value, TS, Function]	4	0.003817
11	[TS, TypeParameter]	4	0.003817
12	[Value, TS, Complex]	5	0.004771
13	[NPM, Engine]	6	0.005726
14	[Project, TS]	6	0.005726
15	[File, Local]	6	0.005726
16	[Value, TS, Call]	6	0.005726
17	[Value, TS, Member]	6	0.005726
18	[File, TS, Local, Module]	6	0.005726
19	[Type, TS, TypeParameterReference]	6	0.005726
20	[TS, EnumMember]	8	0.007634
21	[Type, TS, NotIdentified]	11	0.010497
22	[Json, Value, Array]	12	0.011452
23	[Value, TS, Declared]	13	0.012406
24	[TS, TypeAlias]	16	0.015269
25	[File, Directory, Local]	16	0.015269
26	[TS, Interface]	17	0.016223
27	[Type, TS, Intersection]	17	0.016223
28	[jQAssistant, Rule, Concept]	19	0.018132
29	[Value, TS, Literal]	20	0.019086

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>
```

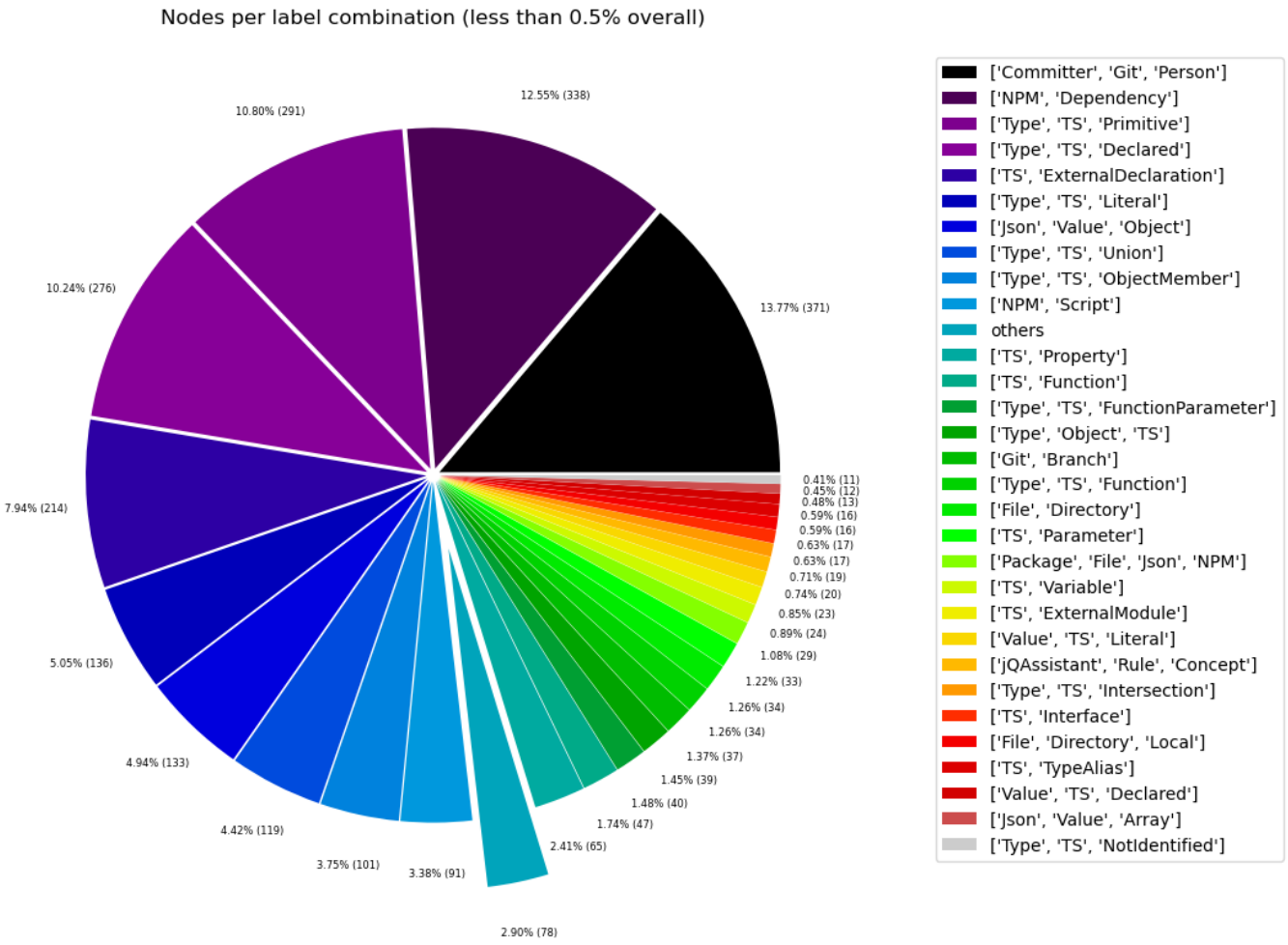


Table 1c - Highest node count by single label

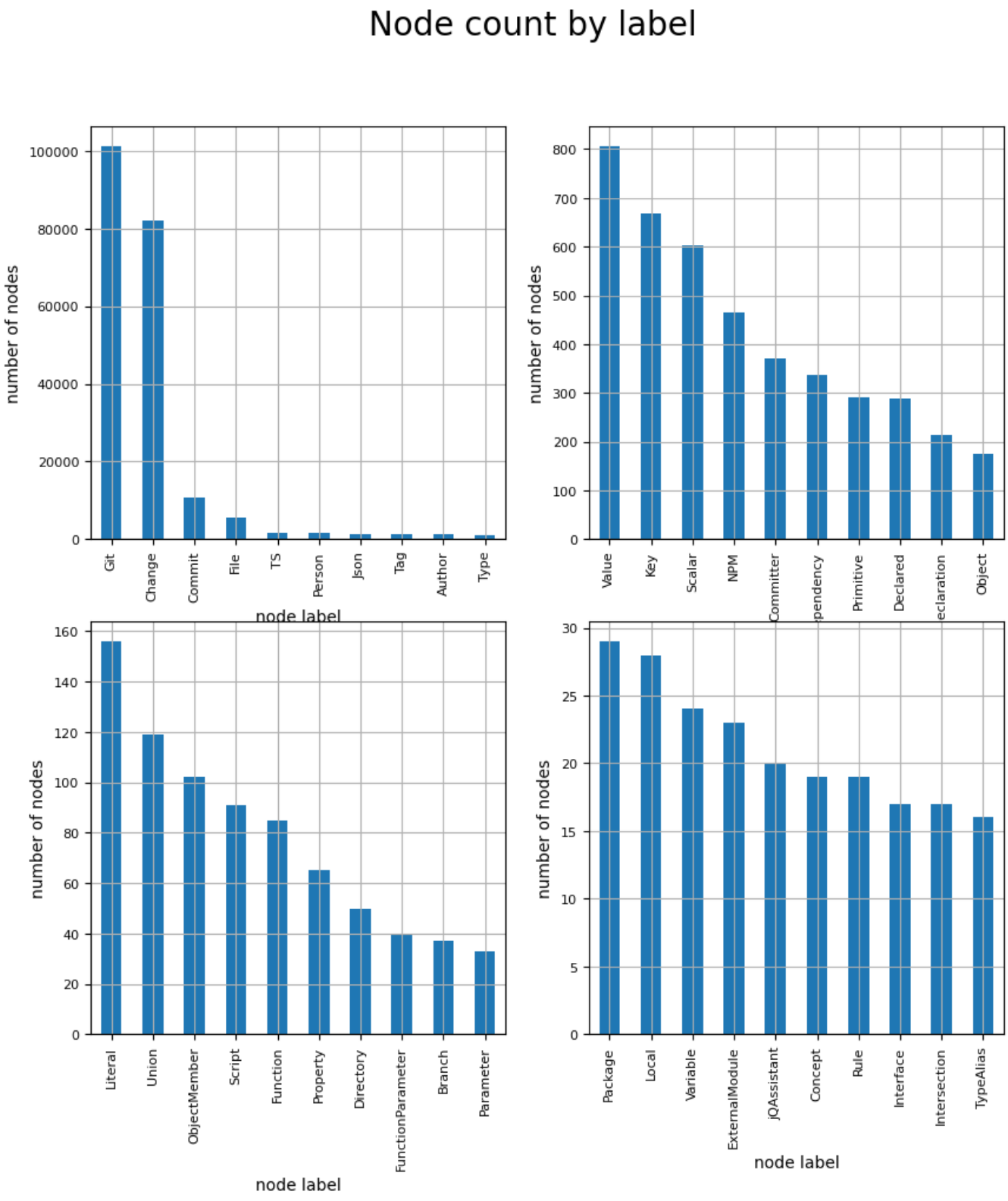
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	101233	96.606514
1	Change	82069	78.318335
2	Commit	10666	10.178549
3	File	5595	5.339301
4	TS	1600	1.526878
5	Person	1598	1.524969
6	Json	1445	1.378962
7	Tag	1360	1.297846
8	Author	1227	1.170924
9	Type	1073	1.023962
10	Value	806	0.769165
11	Key	668	0.637471
12	Scalar	603	0.575442
13	NPM	464	0.442795
14	Committer	371	0.354045
15	Dependency	338	0.322553
16	Primitive	291	0.277701
17	Declared	289	0.275792
18	ExternalDeclaration	214	0.204220
19	Object	175	0.167002
20	Literal	156	0.148871
21	Union	119	0.113562
22	ObjectMember	102	0.097338
23	Script	91	0.086841
24	Function	85	0.081115
25	Property	65	0.062029
26	Directory	50	0.047715
27	FunctionParameter	40	0.038172
28	Branch	37	0.035309
29	Parameter	33	0.031492
30	Package	29	0.027675
31	Local	28	0.026720
32	Variable	24	0.022903
33	ExternalModule	23	0.021949
34	jQAssistant	20	0.019086
35	Concept	19	0.018132
36	Rule	19	0.018132
37	Interface	17	0.016223
38	Intersection	17	0.016223
39	TypeAlias	16	0.015269

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: 313021
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	82069	26.218369
1	MODIFIES	82069	26.218369
2	UPDATES	53788	17.183512
3	COMMITTED	21332	6.814878
4	CREATES	19709	6.296383
5	DELETES	11772	3.760770
6	HAS_PARENT	11714	3.742241
7	HAS_COMMIT	10666	3.407439
8	HAS_FILE	5502	1.757710
9	RENAMES	3200	1.022296
10	HAS_NEW_NAME	1727	0.551720
11	HAS_TAG	1360	0.434476
12	ON_COMMIT	1360	0.434476
13	HAS_AUTHOR	1227	0.391986
14	DEPENDS_ON	961	0.307008
15	HAS_KEY	668	0.213404
16	HAS_VALUE	668	0.213404
17	CONTAINS	594	0.189764
18	HAS_COMMITTER	371	0.118522
19	OF_TYPE	337	0.107661
20	EXPORTS	283	0.090409
21	REFERENCES	197	0.062935
22	DECLARES	186	0.059421
23	DECLARES_DEV_DEPENDENCY	169	0.053990
24	DECLARES_DEPENDENCY	161	0.051434
25	HAS_MEMBER	102	0.032586
26	HAS_TYPE_ARGUMENT	94	0.030030
27	DECLARES_SCRIPT	91	0.029072
28	RETURNS	82	0.026196
29	HAS_PARAMETER	73	0.023321

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>
```

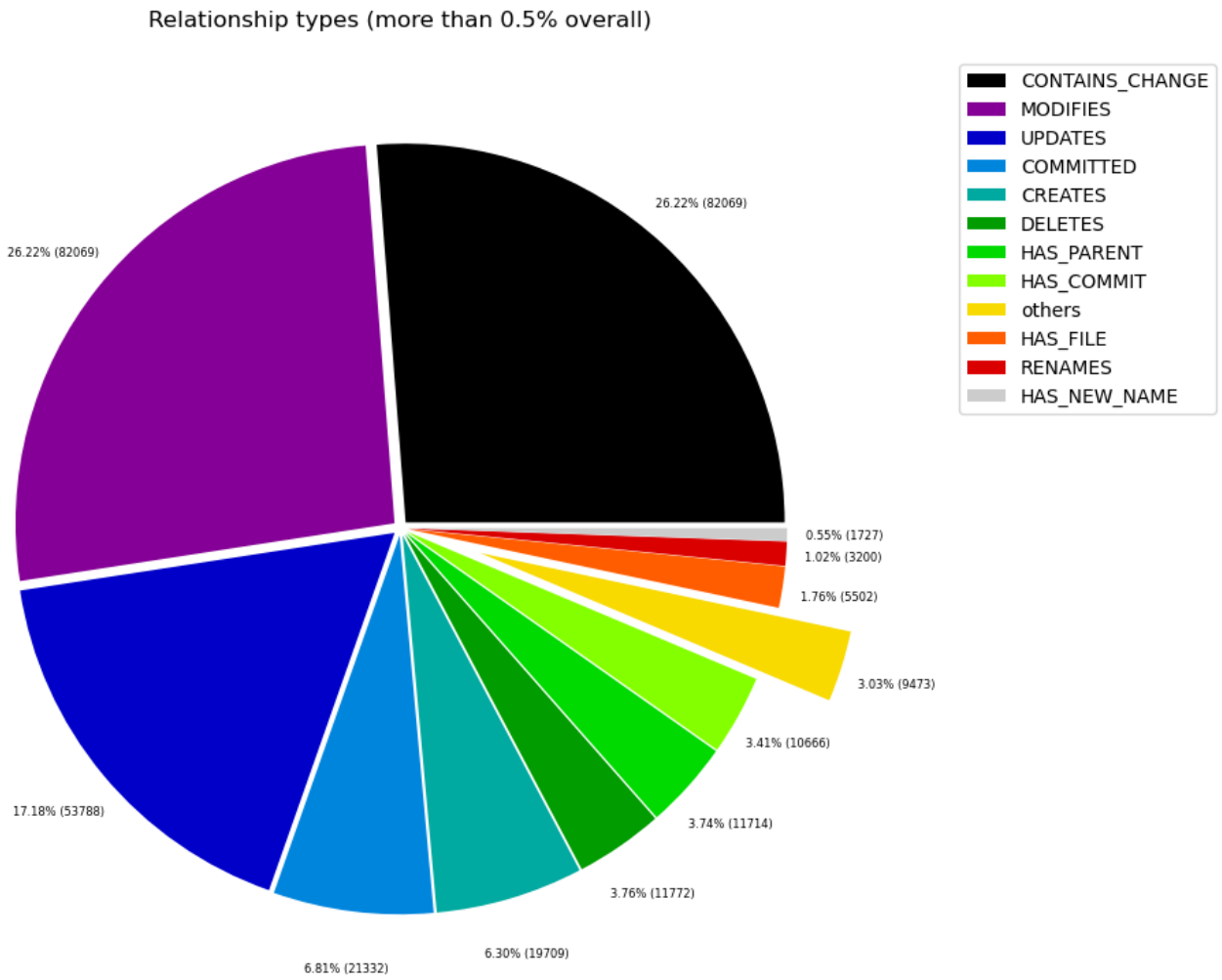


Table 2b - Lowest relationship count by type

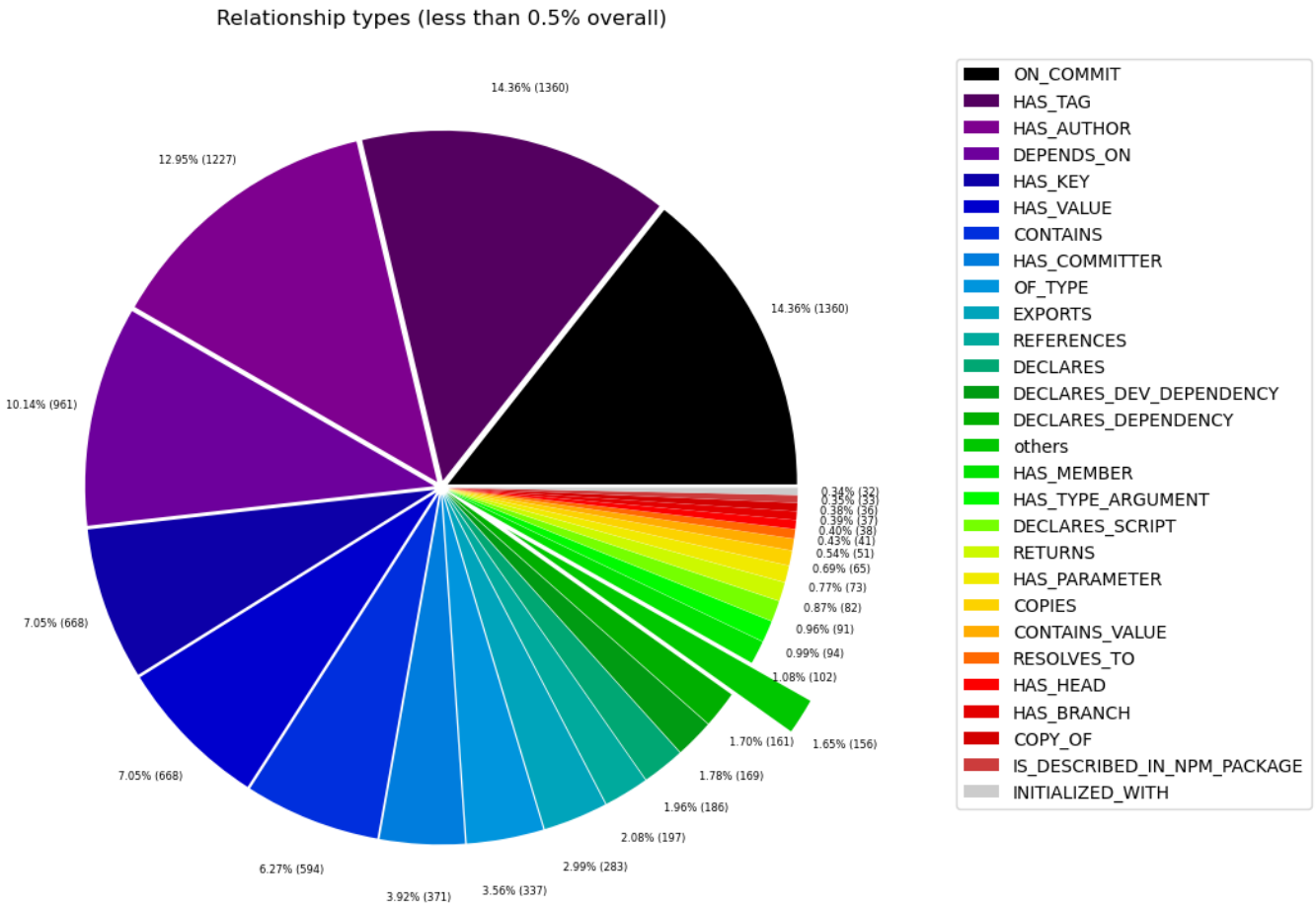
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	PROVIDED_BY_NPM_DEPENDENCY	1	0.000319
1	IS_IMPLEMENTED_IN	2	0.000639
2	CONSTRAINED_BY	4	0.001278
3	REFERENCED_PROJECTS	5	0.001597
4	CONTAINS_PROJECT	6	0.001917
5	DECLARES_ENGINE	6	0.001917
6	EXTENDS	6	0.001917
7	HAS_ARGUMENT	6	0.001917
8	CALLS	6	0.001917
9	HAS_NPM_PACKAGE	6	0.001917
10	HAS_ROOT	6	0.001917
11	MEMBER	6	0.001917
12	PARENT	6	0.001917
13	HAS_CONFIG	6	0.001917
14	SIMILAR	6	0.001917
15	DECLARES_PEER_DEPENDENCY	8	0.002556
16	INCLUDES_CONCEPT	19	0.006070
17	USES	23	0.007348
18	REQUIRES_CONCEPT	28	0.008945
19	INITIALIZED_WITH	32	0.010223
20	IS_DESCRIBED_IN_NPM_PACKAGE	33	0.010542
21	COPY_OF	36	0.011501
22	HAS_BRANCH	37	0.011820
23	HAS_HEAD	38	0.012140
24	RESOLVES_TO	41	0.013098
25	CONTAINS_VALUE	51	0.016293
26	COPIES	65	0.020765
27	HAS_PARAMETER	73	0.023321
28	RETURNS	82	0.026196
29	DECLARES_SCRIPT	91	0.029072

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]	82069	82069	
1	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	82069	10666	
2	[Git, Change]	UPDATES	[File, Git]	53788	82069	
3	[Git, Change]	CREATES	[File, Git]	19709	82069	
4	[Git, Change]	DELETES	[File, Git]	11772	82069	
5	[Git, Commit]	HAS_PARENT	[Git, Commit]	11714	10666	
6	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	10666	1	
7	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10666	371	
8	[Author, Git, Person]	COMMITTED	[Git, Commit]	10666	1227	
9	[Repository, File, Git]	HAS_FILE	[File, Git]	5502	1	
10	[Git, Change]	RENAMES	[File, Git]	3200	82069	
11	[File, Git]	HAS_NEW_NAME	[File, Git]	1727	5502	
12	[Repository, File, Git]	HAS_TAG	[Git, Tag]	1360	1	
13	[Git, Tag]	ON_COMMIT	[Git, Commit]	1360	1360	
14	[Repository, File, Git]	HAS_AUTHOR	[Author, Git, Person]	1227	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): 104789
total_number_of_relationships (edges): 313021
-> total directed graph density: 2.8506652864160158e-05
-> total directed graph density in percent: 0.002850665286416016
```