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]	81481	78.304951
1	[Git, Commit]	10591	10.178173
2	[File, Git]	5493	5.278888
3	[Git, Tag]	1312	1.260860
4	[Author, Git, Person]	1216	1.168602
5	[Json, Key]	668	0.641962
6	[Json, Value, Scalar]	603	0.579496
7	[Committer, Git, Person]	371	0.356539
8	[NPM, Dependency]	338	0.324825
9	[Type, TS, Primitive]	291	0.279657
10	[Type, TS, Declared]	276	0.265242
11	[TS, ExternalDeclaration]	214	0.205658
12	[Type, TS, Literal]	136	0.130699
13	[Json, Value, Object]	133	0.127816
14	[Type, TS, Union]	119	0.114361
15	[Type, TS, ObjectMember]	101	0.097063
16	[NPM, Script]	91	0.087453
17	[TS, Property]	65	0.062466
18	[TS, Function]	47	0.045168
19	[Type, TS, FunctionParameter]	40	0.038441
20	[Type, Object, TS]	39	0.037480
21	[Git, Branch]	35	0.033636
22	[File, Directory]	34	0.032675
23	[Type, TS, Function]	34	0.032675
24	[TS, Parameter]	33	0.031714
25	[Package, File, Json, NPM]	29	0.027870
26	[TS, Variable]	24	0.023065
27	[TS, ExternalModule]	23	0.022103
28	[Value, TS, Literal]	20	0.019220
29	[jQAssistant, Rule, Concept]	19	0.018259

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>
```

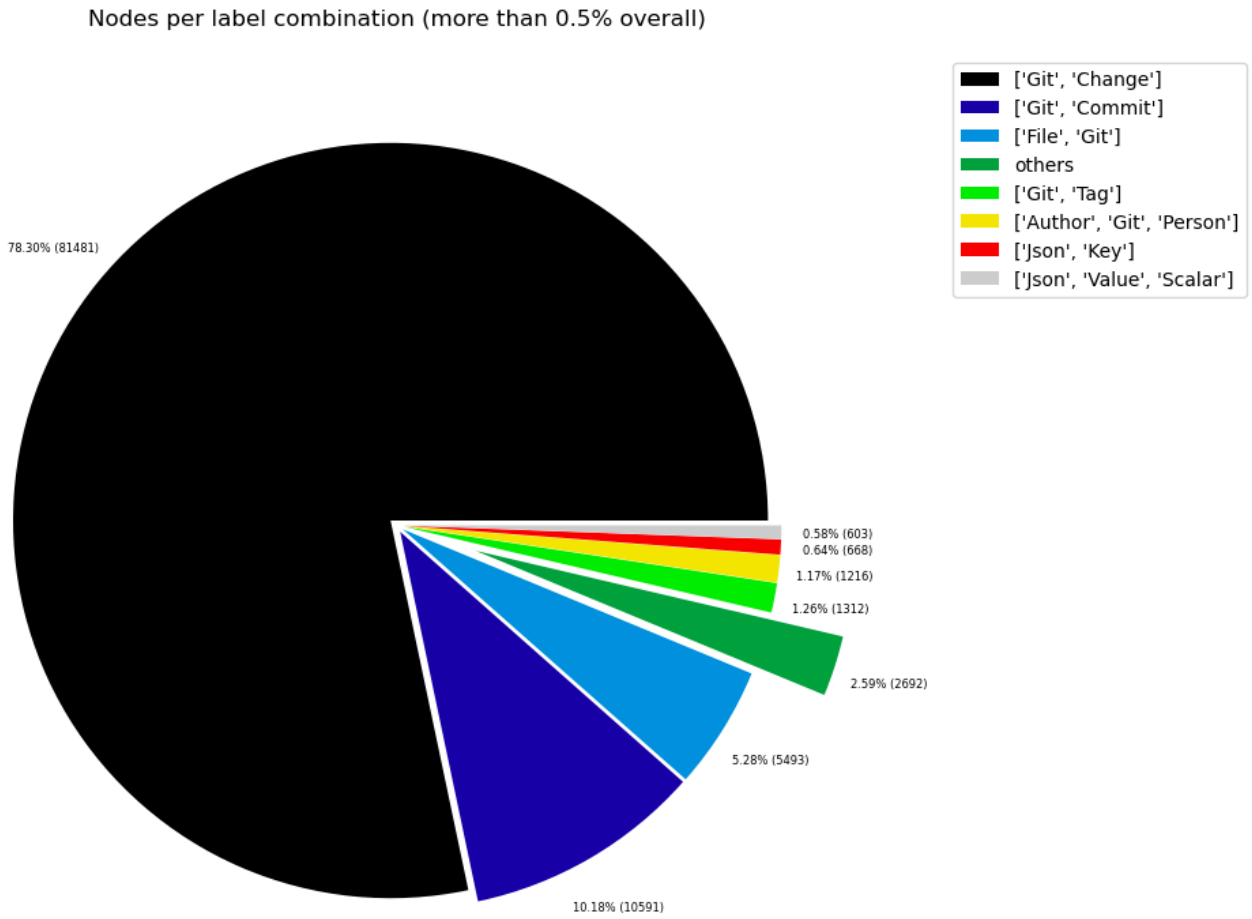


Table 1b - Lowest node count by label combination

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	0.000961
1	[File, TS, Scan]	1	0.000961
2	[TS, Method]	1	0.000961
3	[Repository, File, Git]	1	0.000961
4	[TS, Constructor]	1	0.000961
5	[Value, TS, ObjectMember]	1	0.000961
6	[TS, Class]	1	0.000961
7	[TS, Enum]	2	0.001922
8	[Value, Object, TS]	3	0.002883
9	[Type, TS, Tuple]	3	0.002883
10	[Value, TS, Function]	4	0.003844
11	[TS, TypeParameter]	4	0.003844
12	[Value, TS, Complex]	5	0.004805
13	[NPM, Engine]	6	0.005766
14	[Project, TS]	6	0.005766
15	[File, Local]	6	0.005766
16	[Value, TS, Call]	6	0.005766
17	[Value, TS, Member]	6	0.005766
18	[File, TS, Local, Module]	6	0.005766
19	[Type, TS, TypeParameterReference]	6	0.005766
20	[TS, EnumMember]	8	0.007688
21	[Type, TS, NotIdentified]	11	0.010571
22	[Json, Value, Array]	12	0.011532
23	[Value, TS, Declared]	13	0.012493
24	[TS, TypeAlias]	16	0.015376
25	[File, Directory, Local]	16	0.015376
26	[TS, Interface]	17	0.016337
27	[Type, TS, Intersection]	17	0.016337
28	[jQAssistant, Rule, Concept]	19	0.018259
29	[Value, TS, Literal]	20	0.019220

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>
```

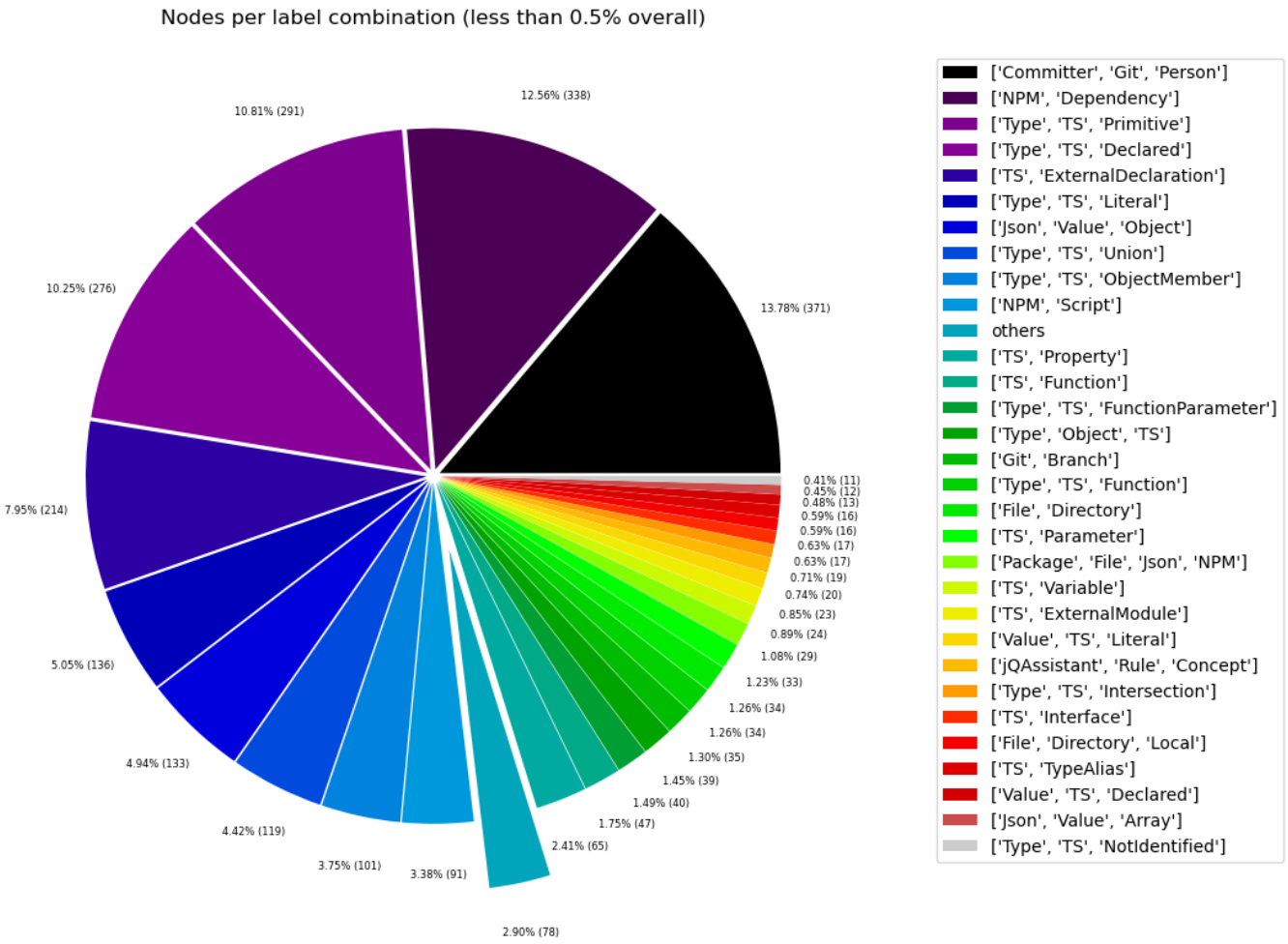


Table 1c - Highest node count by single label

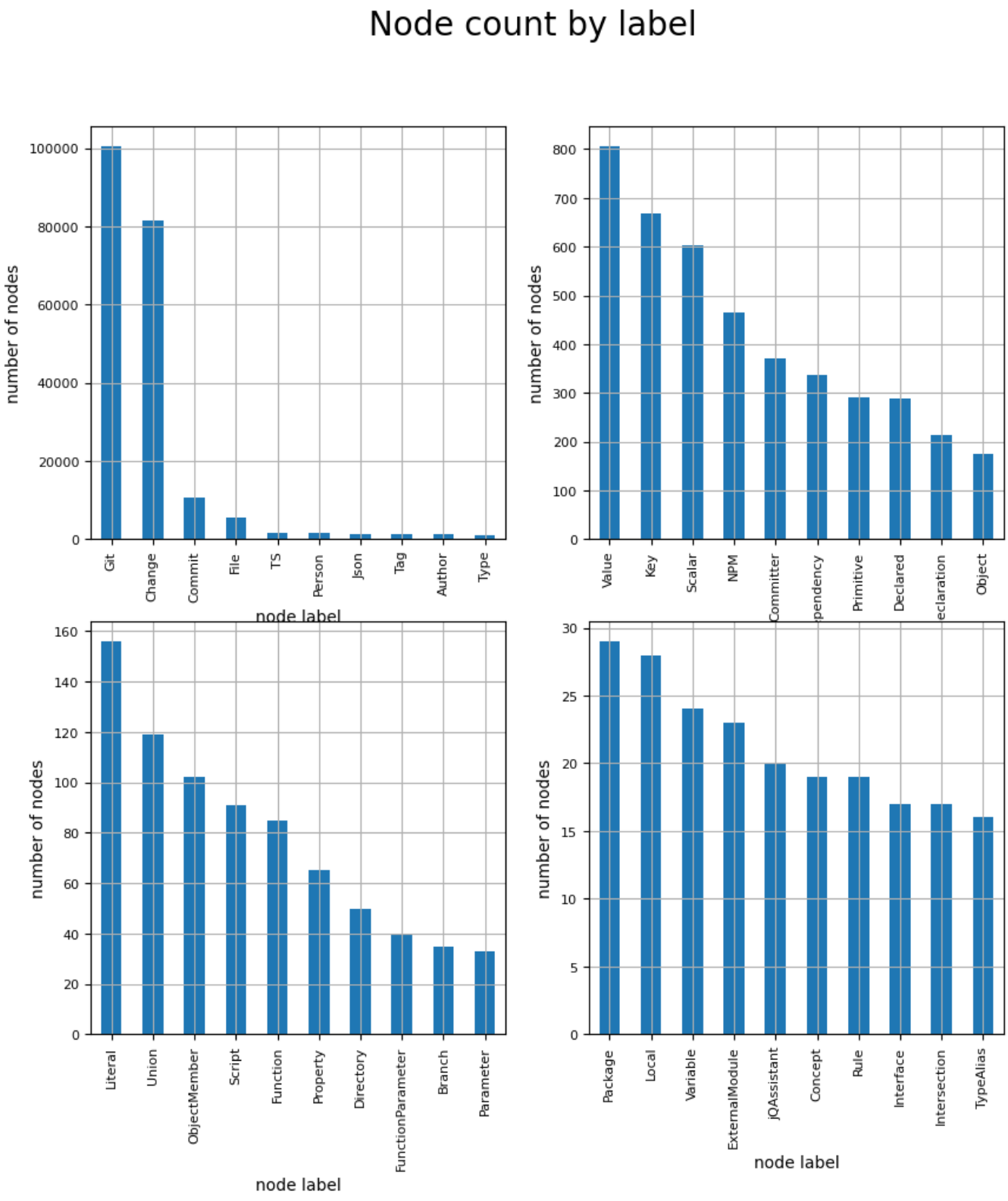
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	100500	96.582609
1	Change	81481	78.304951
2	Commit	10591	10.178173
3	File	5586	5.368263
4	TS	1600	1.537634
5	Person	1587	1.525140
6	Json	1445	1.388675
7	Tag	1312	1.260860
8	Author	1216	1.168602
9	Type	1073	1.031176
10	Value	806	0.774583
11	Key	668	0.641962
12	Scalar	603	0.579496
13	NPM	464	0.445914
14	Committer	371	0.356539
15	Dependency	338	0.324825
16	Primitive	291	0.279657
17	Declared	289	0.277735
18	ExternalDeclaration	214	0.205658
19	Object	175	0.168179
20	Literal	156	0.149919
21	Union	119	0.114361
22	ObjectMember	102	0.098024
23	Script	91	0.087453
24	Function	85	0.081687
25	Property	65	0.062466
26	Directory	50	0.048051
27	FunctionParameter	40	0.038441
28	Branch	35	0.033636
29	Parameter	33	0.031714
30	Package	29	0.027870
31	Local	28	0.026909
32	Variable	24	0.023065
33	ExternalModule	23	0.022103
34	jQAssistant	20	0.019220
35	Concept	19	0.018259
36	Rule	19	0.018259
37	Interface	17	0.016337
38	Intersection	17	0.016337
39	TypeAlias	16	0.015376

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: 310830
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	81481	26.214008
1	MODIFIES	81481	26.214008
2	UPDATES	53296	17.146350
3	COMMITTED	21182	6.814658
4	CREATES	19660	6.325001
5	DELETES	11722	3.771193
6	HAS_PARENT	11637	3.743847
7	HAS_COMMIT	10591	3.407329
8	HAS_FILE	5493	1.767204
9	RENAMES	3197	1.028536
10	HAS_NEW_NAME	1726	0.555287
11	HAS_TAG	1312	0.422096
12	ON_COMMIT	1312	0.422096
13	HAS_AUTHOR	1216	0.391211
14	DEPENDS_ON	961	0.309172
15	HAS_KEY	668	0.214908
16	HAS_VALUE	668	0.214908
17	CONTAINS	594	0.191101
18	HAS_COMMITTER	371	0.119358
19	OF_TYPE	337	0.108419
20	EXPORTS	283	0.091047
21	REFERENCES	197	0.063379
22	DECLARES	186	0.059840
23	DECLARES_DEV_DEPENDENCY	169	0.054371
24	DECLARES_DEPENDENCY	161	0.051797
25	HAS_MEMBER	102	0.032815
26	HAS_TYPE_ARGUMENT	94	0.030242
27	DECLARES_SCRIPT	91	0.029276
28	RETURNS	82	0.026381
29	HAS_PARAMETER	73	0.023486

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>
```

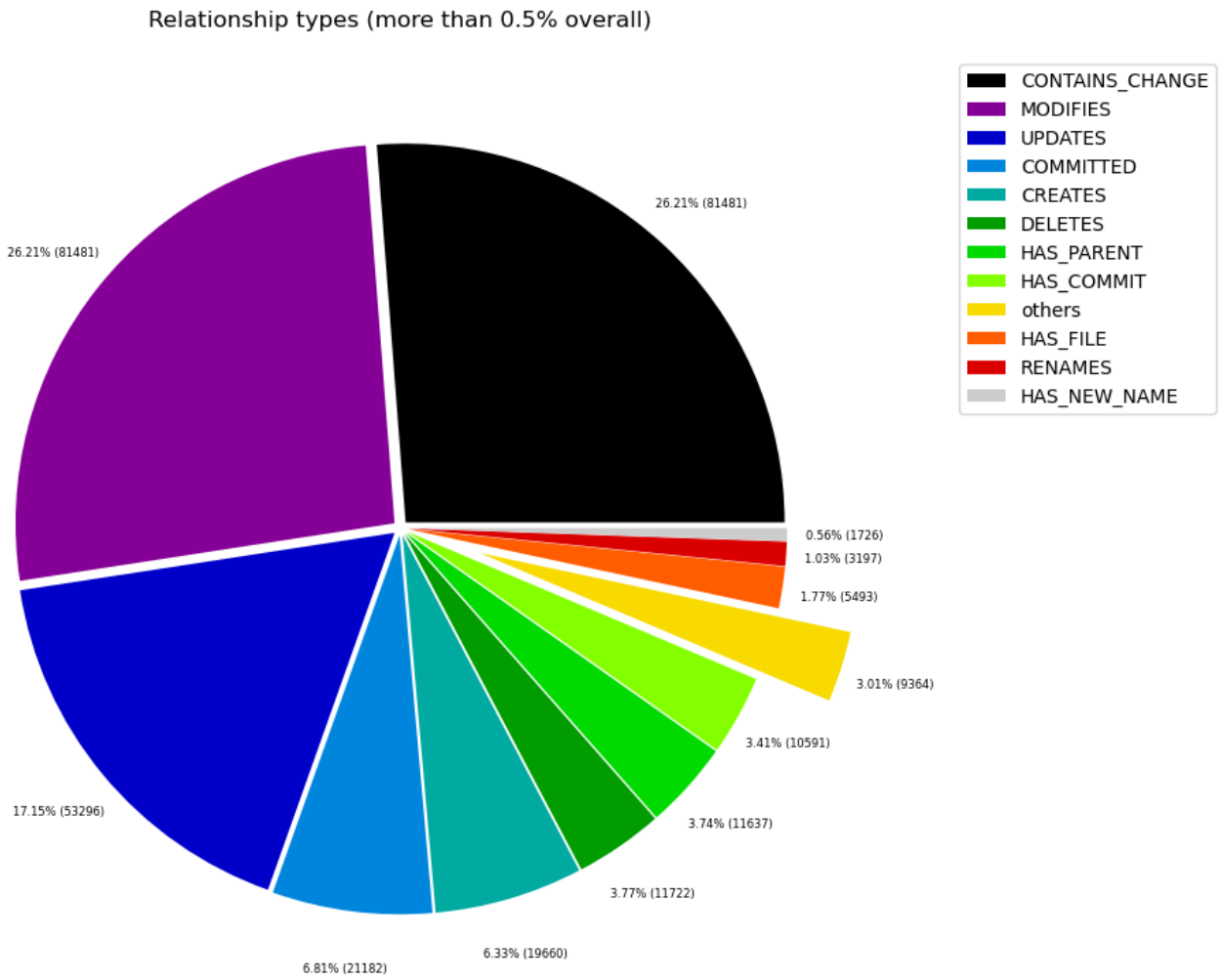


Table 2b - Lowest relationship count by type

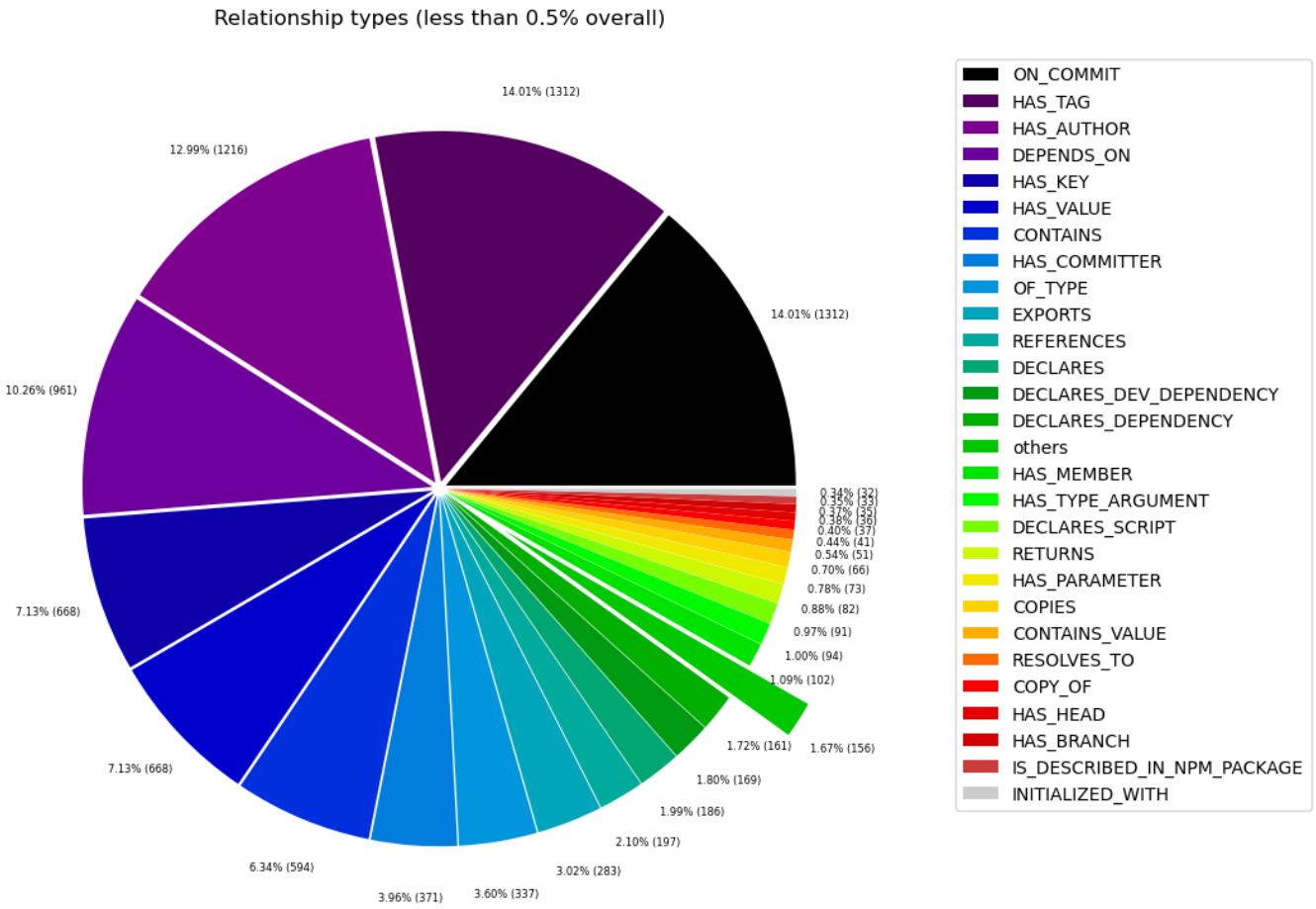
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	PROVIDED_BY_NPM_DEPENDENCY	1	0.000322
1	IS_IMPLEMENTED_IN	2	0.000643
2	CONSTRAINED_BY	4	0.001287
3	REFERENCED_PROJECTS	5	0.001609
4	CONTAINS_PROJECT	6	0.001930
5	DECLARES_ENGINE	6	0.001930
6	EXTENDS	6	0.001930
7	HAS_ARGUMENT	6	0.001930
8	CALLS	6	0.001930
9	HAS_NPM_PACKAGE	6	0.001930
10	HAS_ROOT	6	0.001930
11	MEMBER	6	0.001930
12	PARENT	6	0.001930
13	HAS_CONFIG	6	0.001930
14	SIMILAR	6	0.001930
15	DECLARES_PEER_DEPENDENCY	8	0.002574
16	INCLUDES_CONCEPT	19	0.006113
17	USES	23	0.007400
18	REQUIRES_CONCEPT	28	0.009008
19	INITIALIZED_WITH	32	0.010295
20	IS_DESCRIBED_IN_NPM_PACKAGE	33	0.010617
21	HAS_BRANCH	35	0.011260
22	HAS_HEAD	36	0.011582
23	COPY_OF	37	0.011904
24	RESOLVES_TO	41	0.013190
25	CONTAINS_VALUE	51	0.016408
26	COPIES	66	0.021233
27	HAS_PARAMETER	73	0.023486
28	RETURNS	82	0.026381
29	DECLARES_SCRIPT	91	0.029276

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]	81481	81481	
1	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	81481	10591	
2	[Git, Change]	UPDATES	[File, Git]	53296	81481	
3	[Git, Change]	CREATES	[File, Git]	19660	81481	
4	[Git, Change]	DELETES	[File, Git]	11722	81481	
5	[Git, Commit]	HAS_PARENT	[Git, Commit]	11637	10591	
6	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	10591	1	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	10591	1216	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10591	371	
9	[Repository, File, Git]	HAS_FILE	[File, Git]	5493	1	
10	[Git, Change]	RENAMES	[File, Git]	3197	81481	
11	[File, Git]	HAS_NEW_NAME	[File, Git]	1726	5493	
12	[Repository, File, Git]	HAS_TAG	[Git, Tag]	1312	1	
13	[Git, Tag]	ON_COMMIT	[Git, Commit]	1312	1312	
14	[Repository, File, Git]	HAS_AUTHOR	[Author, Git, Person]	1216	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): 104056
total_number_of_relationships (edges): 310830
-> total directed graph density: 2.8707333037059398e-05
-> total directed graph density in percent: 0.0028707333037059396
```