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]	78729	78.006658
1	[Git, Commit]	10413	10.317460
2	[File, Git]	5403	5.353427
3	[Git, Tag]	1229	1.217724
4	[Author, Git, Person]	1197	1.186017
5	[Json, Key]	668	0.661871
6	[Json, Value, Scalar]	603	0.597467
7	[Committer, Git, Person]	371	0.367596
8	[NPM, Dependency]	330	0.326972
9	[Type, TS, Primitive]	291	0.288330
10	[Type, TS, Declared]	276	0.273468
11	[TS, ExternalDeclaration]	215	0.213027
12	[Type, TS, Literal]	136	0.134752
13	[Json, Value, Object]	133	0.131780
14	[Type, TS, Union]	119	0.117908
15	[Type, TS, ObjectMember]	101	0.100073
16	[NPM, Script]	91	0.090165
17	[TS, Property]	65	0.064404
18	[TS, Function]	47	0.046569
19	[Type, TS, FunctionParameter]	40	0.039633
20	[Type, Object, TS]	39	0.038642
21	[File, Directory]	34	0.033688
22	[Type, TS, Function]	34	0.033688
23	[TS, Parameter]	33	0.032697
24	[Git, Branch]	32	0.031706
25	[Package, File, Json, NPM]	29	0.028734
26	[TS, ExternalModule]	25	0.024771
27	[TS, Variable]	24	0.023780
28	[Value, TS, Literal]	20	0.019816
29	[jQAssistant, Rule, Concept]	19	0.018826

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>
```

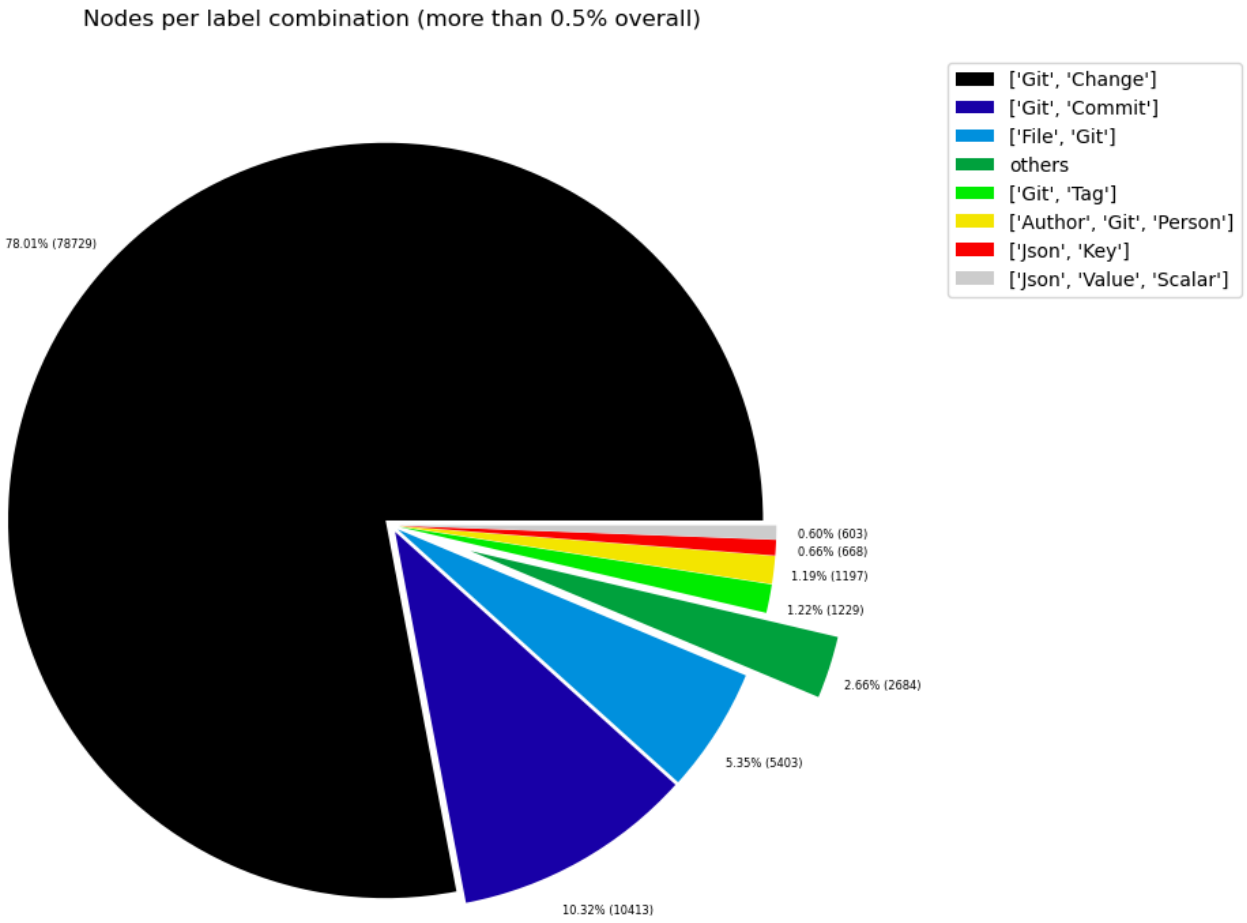


Table 1b - Lowest node count by label combination

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	0.000991
1	[File, TS, Scan]	1	0.000991
2	[TS, Method]	1	0.000991
3	[Repository, File, Git]	1	0.000991
4	[TS, Constructor]	1	0.000991
5	[Value, TS, ObjectMember]	1	0.000991
6	[TS, Class]	1	0.000991
7	[TS, Enum]	2	0.001982
8	[Value, Object, TS]	3	0.002972
9	[Type, TS, Tuple]	3	0.002972
10	[Value, TS, Function]	4	0.003963
11	[TS, TypeParameter]	4	0.003963
12	[Value, TS, Complex]	5	0.004954
13	[NPM, Engine]	6	0.005945
14	[Project, TS]	6	0.005945
15	[File, Local]	6	0.005945
16	[Value, TS, Call]	6	0.005945
17	[Value, TS, Member]	6	0.005945
18	[File, TS, Local, Module]	6	0.005945
19	[Type, TS, TypeParameterReference]	6	0.005945
20	[TS, EnumMember]	8	0.007927
21	[Type, TS, NotIdentified]	11	0.010899
22	[Json, Value, Array]	12	0.011890
23	[Value, TS, Declared]	13	0.012881
24	[TS, TypeAlias]	16	0.015853
25	[File, Directory, Local]	16	0.015853
26	[TS, Interface]	17	0.016844
27	[Type, TS, Intersection]	17	0.016844
28	[jQAssistant, Rule, Concept]	19	0.018826
29	[Value, TS, Literal]	20	0.019816

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>
```

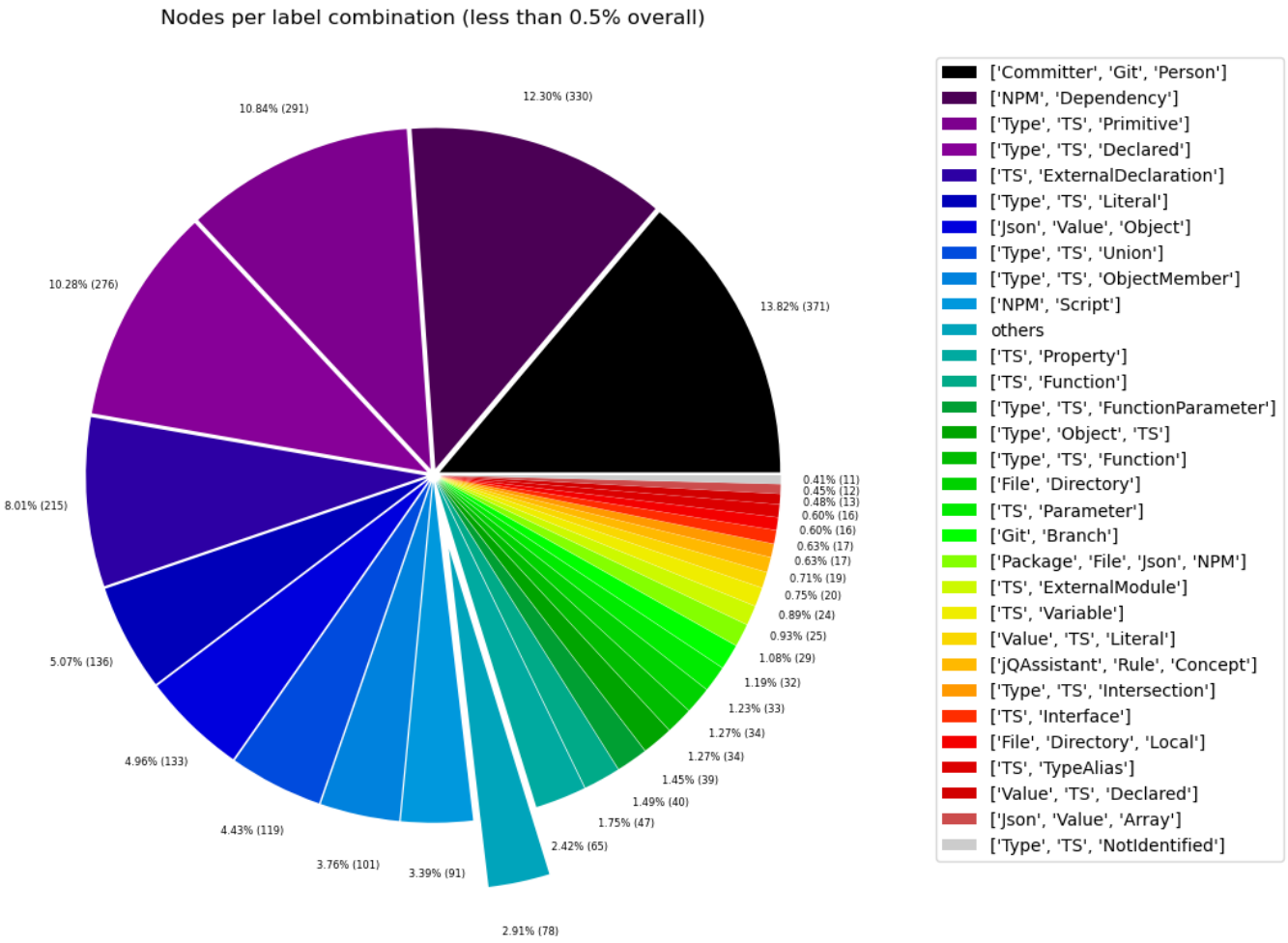


Table 1c - Highest node count by single label

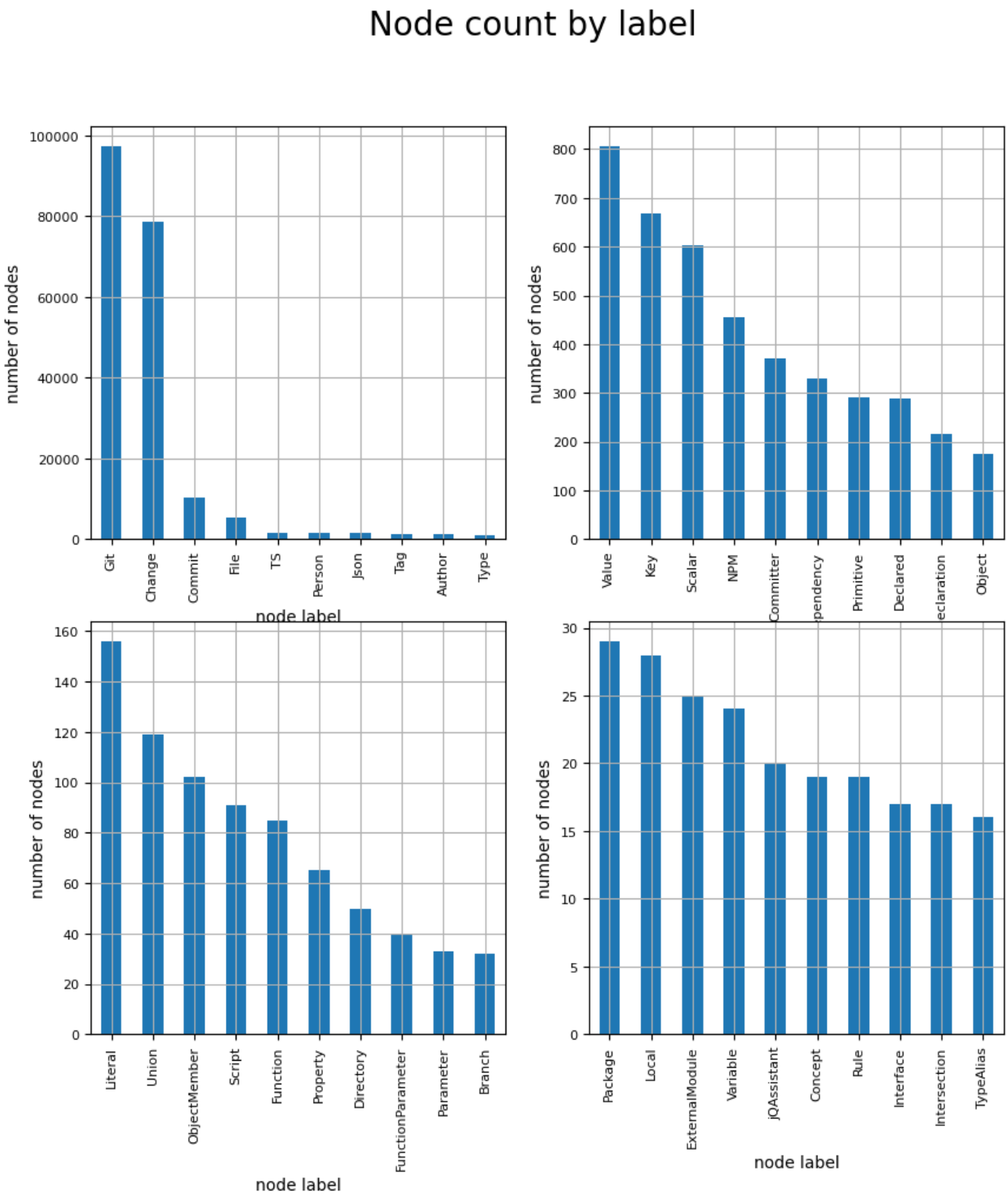
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	97375	96.481581
1	Change	78729	78.006658
2	Commit	10413	10.317460
3	File	5496	5.445574
4	TS	1603	1.588292
5	Person	1568	1.553614
6	Json	1445	1.431742
7	Tag	1229	1.217724
8	Author	1197	1.186017
9	Type	1073	1.063155
10	Value	806	0.798605
11	Key	668	0.661871
12	Scalar	603	0.597467
13	NPM	456	0.451816
14	Committer	371	0.367596
15	Dependency	330	0.326972
16	Primitive	291	0.288330
17	Declared	289	0.286348
18	ExternalDeclaration	215	0.213027
19	Object	175	0.173394
20	Literal	156	0.154569
21	Union	119	0.117908
22	ObjectMember	102	0.101064
23	Script	91	0.090165
24	Function	85	0.084220
25	Property	65	0.064404
26	Directory	50	0.049541
27	FunctionParameter	40	0.039633
28	Parameter	33	0.032697
29	Branch	32	0.031706
30	Package	29	0.028734
31	Local	28	0.027743
32	ExternalModule	25	0.024771
33	Variable	24	0.023780
34	jQAssistant	20	0.019816
35	Concept	19	0.018826
36	Rule	19	0.018826
37	Interface	17	0.016844
38	Intersection	17	0.016844
39	TypeAlias	16	0.015853

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: 301061
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	78729	26.150514
1	MODIFIES	78729	26.150514
2	UPDATES	52068	17.294834
3	COMMITTED	20826	6.917535
4	CREATES	18719	6.217677
5	HAS_PARENT	11446	3.801887
6	DELETES	10930	3.630493
7	HAS_COMMIT	10413	3.458767
8	HAS_FILE	5403	1.794653
9	RENAMES	2988	0.992490
10	HAS_NEW_NAME	1685	0.559687
11	HAS_TAG	1229	0.408223
12	ON_COMMIT	1229	0.408223
13	HAS_AUTHOR	1197	0.397594
14	DEPENDS_ON	959	0.318540
15	HAS_KEY	668	0.221882
16	HAS_VALUE	668	0.221882
17	CONTAINS	594	0.197302
18	HAS_COMMITTER	371	0.123231
19	OF_TYPE	337	0.111937
20	EXPORTS	276	0.091676
21	REFERENCES	197	0.065435
22	DECLARES	186	0.061781
23	DECLARES_DEV_DEPENDENCY	169	0.056135
24	DECLARES_DEPENDENCY	161	0.053478
25	HAS_MEMBER	102	0.033880
26	HAS_TYPE_ARGUMENT	94	0.031223
27	DECLARES_SCRIPT	91	0.030226
28	RETURNS	82	0.027237
29	HAS_PARAMETER	73	0.024248

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>
```

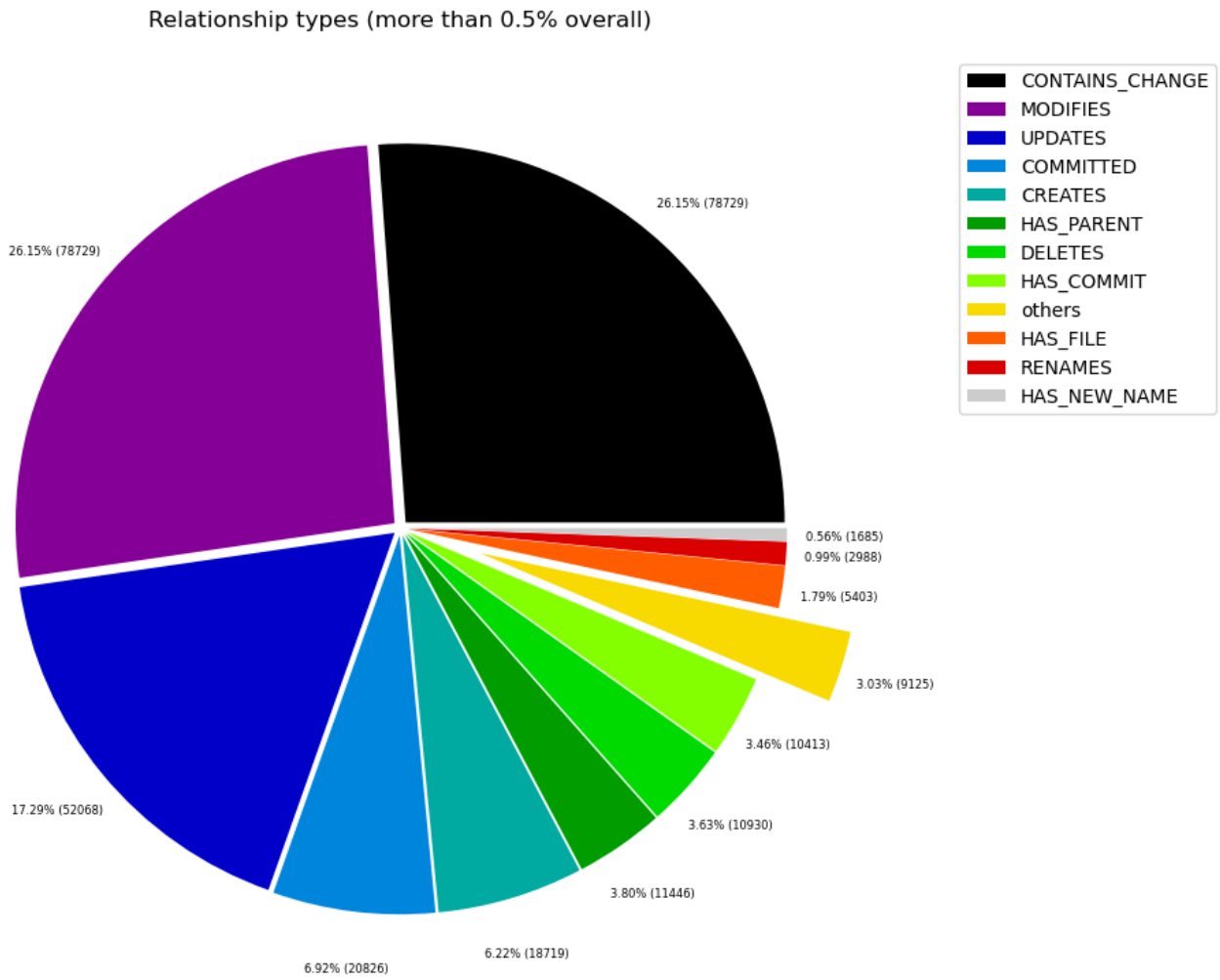


Table 2b - Lowest relationship count by type

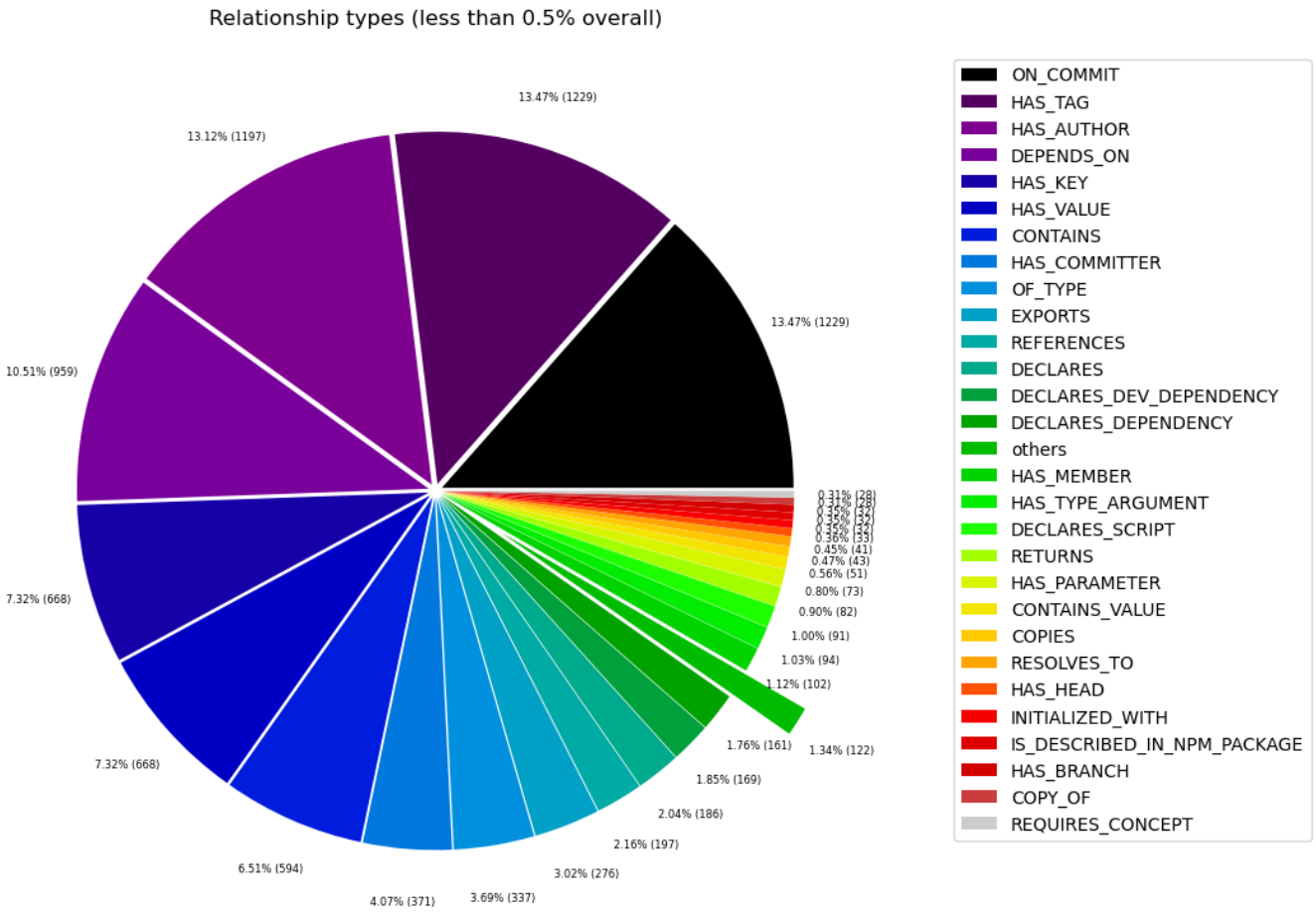
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	PROVIDED_BY_NPM_DEPENDENCY	1	0.000332
1	IS_IMPLEMENTED_IN	2	0.000664
2	CONSTRAINED_BY	4	0.001329
3	REFERENCED_PROJECTS	5	0.001661
4	CONTAINS_PROJECT	6	0.001993
5	DECLARES_ENGINE	6	0.001993
6	EXTENDS	6	0.001993
7	HAS_ARGUMENT	6	0.001993
8	CALLS	6	0.001993
9	HAS_NPM_PACKAGE	6	0.001993
10	HAS_ROOT	6	0.001993
11	MEMBER	6	0.001993
12	PARENT	6	0.001993
13	HAS_CONFIG	6	0.001993
14	SIMILAR	6	0.001993
15	INCLUDES_CONCEPT	19	0.006311
16	USES	25	0.008304
17	REQUIRES_CONCEPT	28	0.009300
18	COPY_OF	28	0.009300
19	INITIALIZED_WITH	32	0.010629
20	HAS_BRANCH	32	0.010629
21	IS_DESCRIBED_IN_NPM_PACKAGE	32	0.010629
22	HAS_HEAD	33	0.010961
23	RESOLVES_TO	41	0.013619
24	COPIES	43	0.014283
25	CONTAINS_VALUE	51	0.016940
26	HAS_PARAMETER	73	0.024248
27	RETURNS	82	0.027237
28	DECLARES_SCRIPT	91	0.030226
29	HAS_TYPE_ARGUMENT	94	0.031223

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]	78729	78729	
1	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	78729	10413	
2	[Git, Change]	UPDATES	[File, Git]	52068	78729	
3	[Git, Change]	CREATES	[File, Git]	18719	78729	
4	[Git, Commit]	HAS_PARENT	[Git, Commit]	11446	10413	
5	[Git, Change]	DELETES	[File, Git]	10930	78729	
6	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	10413	1	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	10413	1197	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10413	371	
9	[Repository, File, Git]	HAS_FILE	[File, Git]	5403	1	
10	[Git, Change]	RENAMES	[File, Git]	2988	78729	
11	[File, Git]	HAS_NEW_NAME	[File, Git]	1685	5403	
12	[Repository, File, Git]	HAS_TAG	[Git, Tag]	1229	1	
13	[Git, Tag]	ON_COMMIT	[Git, Commit]	1229	1229	
14	[Repository, File, Git]	HAS_AUTHOR	[Author, Git, Person]	1197	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): 100926
total_number_of_relationships (edges): 301061
-> total directed graph density: 2.955647793333656e-05
-> total directed graph density in percent: 0.0029556477933336558
```