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]	83111	78.414001
1	[Git, Commit]	10750	10.142466
2	[File, Git]	5538	5.225021
3	[Git, Tag]	1388	1.309558
4	[Author, Git, Person]	1234	1.164261
5	[Json, Key]	668	0.630248
6	[Json, Value, Scalar]	603	0.568922
7	[Committer, Git, Person]	371	0.350033
8	[NPM, Dependency]	338	0.318898
9	[Type, TS, Primitive]	291	0.274554
10	[Type, TS, Declared]	276	0.260402
11	[TS, ExternalDeclaration]	214	0.201906
12	[Type, TS, Literal]	136	0.128314
13	[Json, Value, Object]	133	0.125484
14	[Type, TS, Union]	119	0.112275
15	[Type, TS, ObjectMember]	101	0.095292
16	[NPM, Script]	91	0.085857
17	[TS, Property]	65	0.061327
18	[TS, Function]	47	0.044344
19	[Git, Branch]	41	0.038683
20	[Type, TS, FunctionParameter]	40	0.037739
21	[Type, Object, TS]	39	0.036796
22	[File, Directory]	34	0.032078
23	[Type, TS, Function]	34	0.032078
24	[TS, Parameter]	33	0.031135
25	[Package, File, Json, NPM]	29	0.027361
26	[TS, Variable]	24	0.022644
27	[TS, ExternalModule]	23	0.021700
28	[Value, TS, Literal]	20	0.018870
29	[jQAssistant, Rule, Concept]	19	0.017926

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>
```

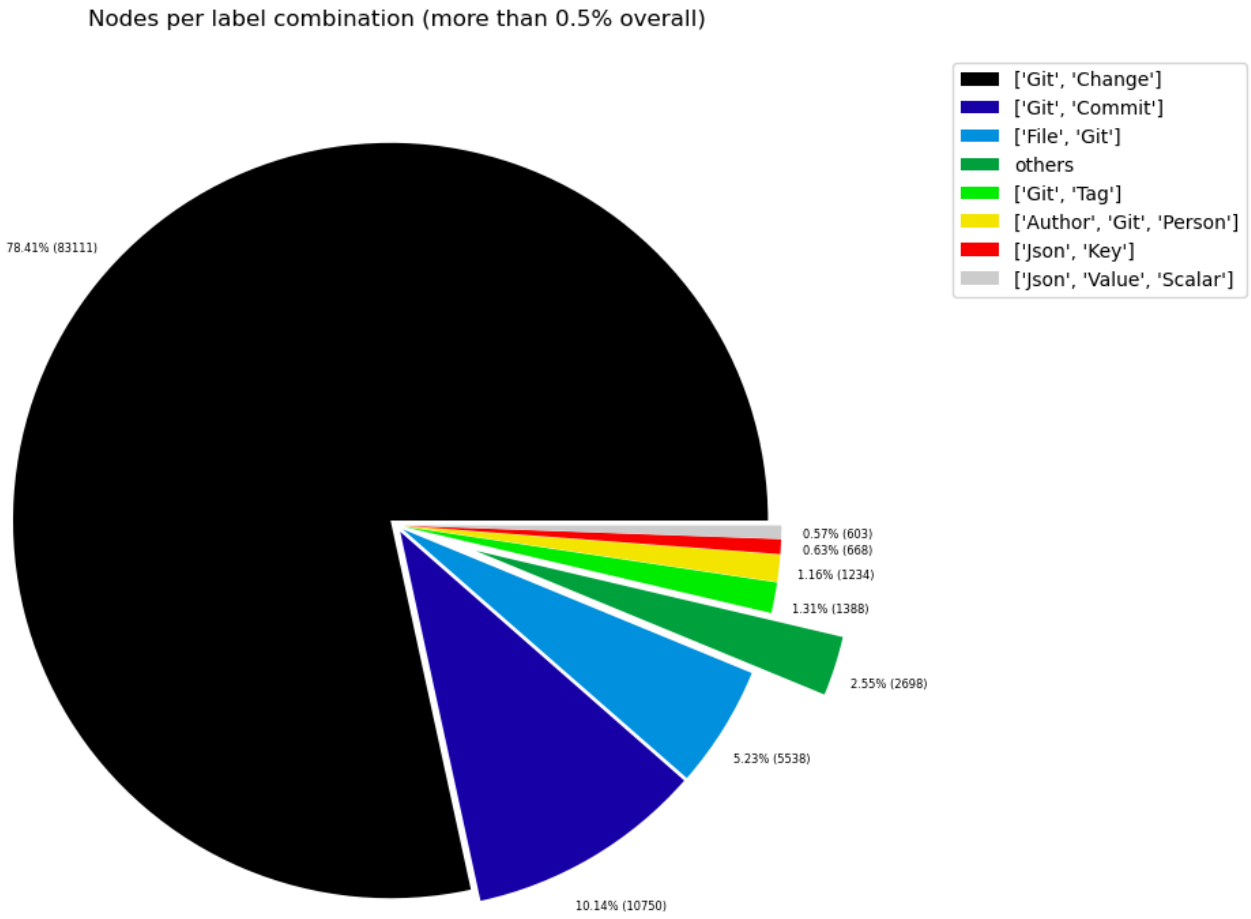


Table 1b - Lowest node count by label combination

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	0.000943
1	[File, TS, Scan]	1	0.000943
2	[TS, Method]	1	0.000943
3	[Repository, File, Git]	1	0.000943
4	[TS, Constructor]	1	0.000943
5	[Value, TS, ObjectMember]	1	0.000943
6	[TS, Class]	1	0.000943
7	[TS, Enum]	2	0.001887
8	[Value, Object, TS]	3	0.002830
9	[Type, TS, Tuple]	3	0.002830
10	[Value, TS, Function]	4	0.003774
11	[TS, TypeParameter]	4	0.003774
12	[Value, TS, Complex]	5	0.004717
13	[NPM, Engine]	6	0.005661
14	[Project, TS]	6	0.005661
15	[File, Local]	6	0.005661
16	[Value, TS, Call]	6	0.005661
17	[Value, TS, Member]	6	0.005661
18	[File, TS, Local, Module]	6	0.005661
19	[Type, TS, TypeParameterReference]	6	0.005661
20	[TS, EnumMember]	8	0.007548
21	[Type, TS, NotIdentified]	11	0.010378
22	[Json, Value, Array]	12	0.011322
23	[Value, TS, Declared]	13	0.012265
24	[TS, TypeAlias]	16	0.015096
25	[File, Directory, Local]	16	0.015096
26	[TS, Interface]	17	0.016039
27	[Type, TS, Intersection]	17	0.016039
28	[jQAssistant, Rule, Concept]	19	0.017926
29	[Value, TS, Literal]	20	0.018870

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>
```

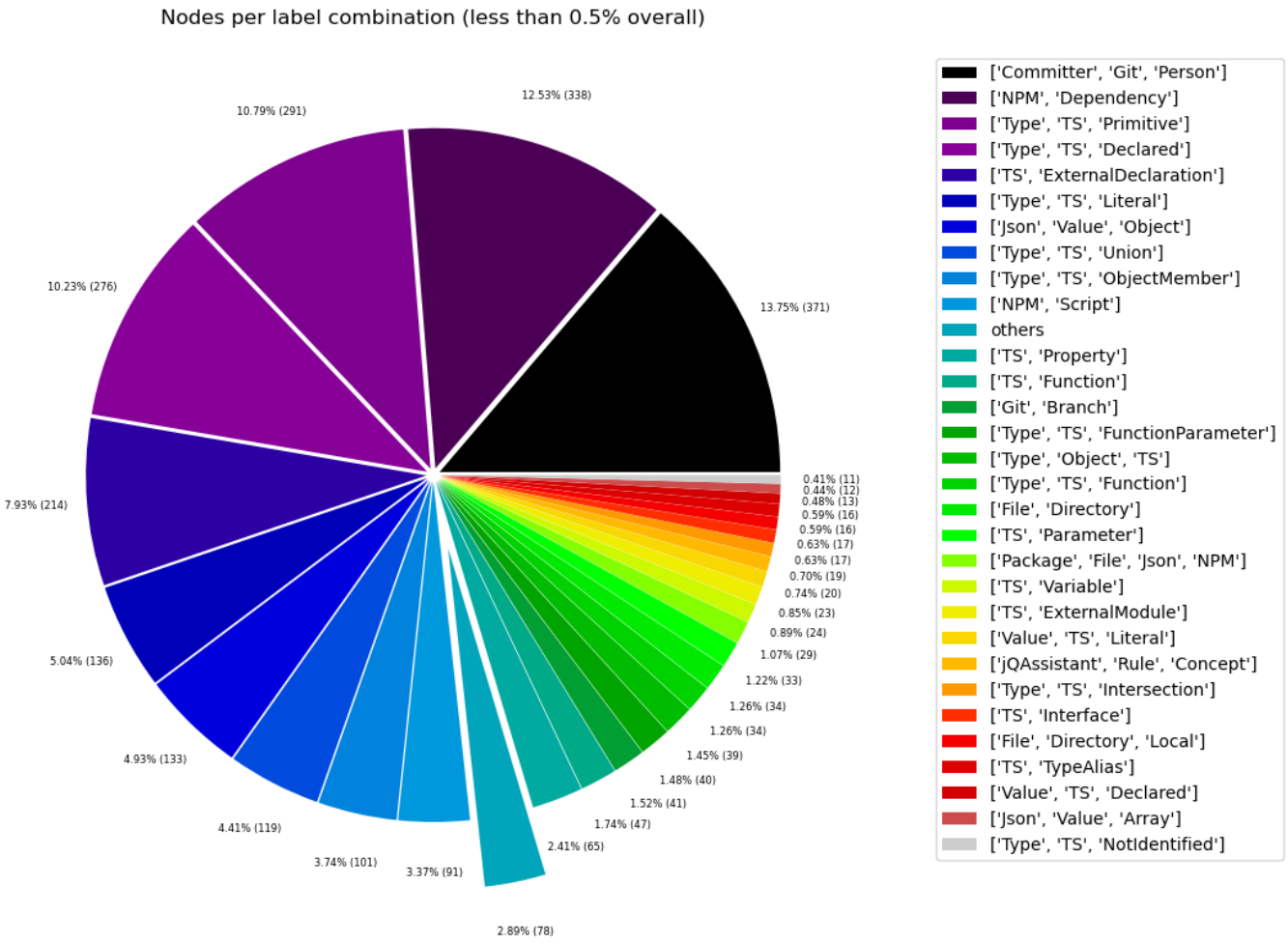


Table 1c - Highest node count by single label

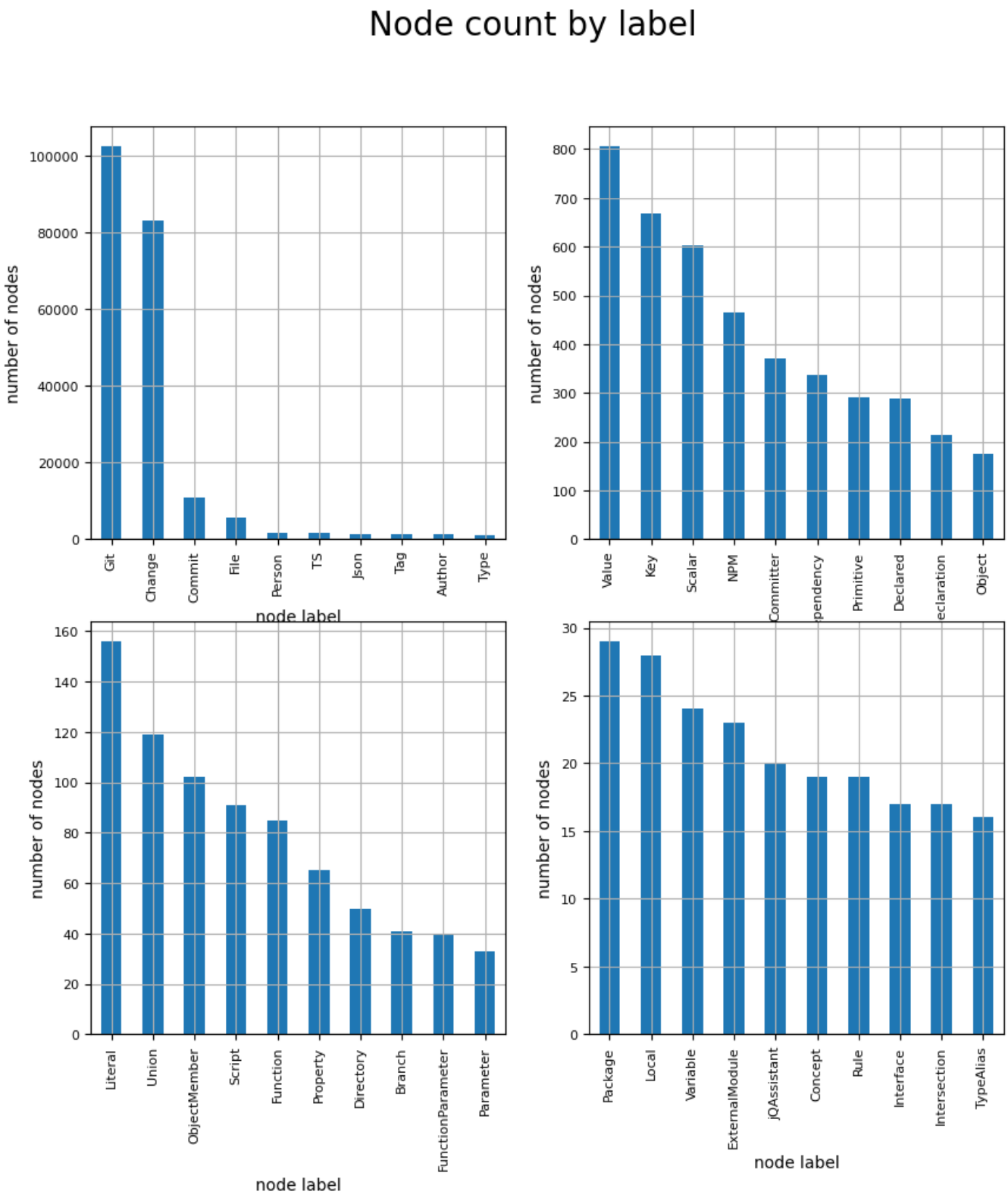
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	102434	96.644967
1	Change	83111	78.414001
2	Commit	10750	10.142466
3	File	5631	5.312765
4	Person	1605	1.514294
5	TS	1600	1.509576
6	Json	1445	1.363336
7	Tag	1388	1.309558
8	Author	1234	1.164261
9	Type	1073	1.012360
10	Value	806	0.760449
11	Key	668	0.630248
12	Scalar	603	0.568922
13	NPM	464	0.437777
14	Committer	371	0.350033
15	Dependency	338	0.318898
16	Primitive	291	0.274554
17	Declared	289	0.272667
18	ExternalDeclaration	214	0.201906
19	Object	175	0.165110
20	Literal	156	0.147184
21	Union	119	0.112275
22	ObjectMember	102	0.096235
23	Script	91	0.085857
24	Function	85	0.080196
25	Property	65	0.061327
26	Directory	50	0.047174
27	Branch	41	0.038683
28	FunctionParameter	40	0.037739
29	Parameter	33	0.031135
30	Package	29	0.027361
31	Local	28	0.026418
32	Variable	24	0.022644
33	ExternalModule	23	0.021700
34	jQAssistant	20	0.018870
35	Concept	19	0.017926
36	Rule	19	0.017926
37	Interface	17	0.016039
38	Intersection	17	0.016039
39	TypeAlias	16	0.015096

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: 316715
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	83111	26.241574
1	MODIFIES	83111	26.241574
2	UPDATES	54434	17.187061
3	COMMITTED	21500	6.788438
4	CREATES	19958	6.301564
5	DELETES	11969	3.779107
6	HAS_PARENT	11804	3.727010
7	HAS_COMMIT	10750	3.394219
8	HAS_FILE	5538	1.748575
9	RENAMES	3250	1.026159
10	HAS_NEW_NAME	1729	0.545917
11	HAS_TAG	1388	0.438249
12	ON_COMMIT	1388	0.438249
13	HAS_AUTHOR	1234	0.389625
14	DEPENDS_ON	961	0.303427
15	HAS_KEY	668	0.210915
16	HAS_VALUE	668	0.210915
17	CONTAINS	594	0.187550
18	HAS_COMMITTER	371	0.117140
19	OF_TYPE	337	0.106405
20	EXPORTS	283	0.089355
21	REFERENCES	197	0.062201
22	DECLARES	186	0.058728
23	DECLARES_DEV_DEPENDENCY	169	0.053360
24	DECLARES_DEPENDENCY	161	0.050834
25	HAS_MEMBER	102	0.032206
26	HAS_TYPE_ARGUMENT	94	0.029680
27	DECLARES_SCRIPT	91	0.028732
28	RETURNS	82	0.025891
29	COPIES	77	0.024312

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>
```

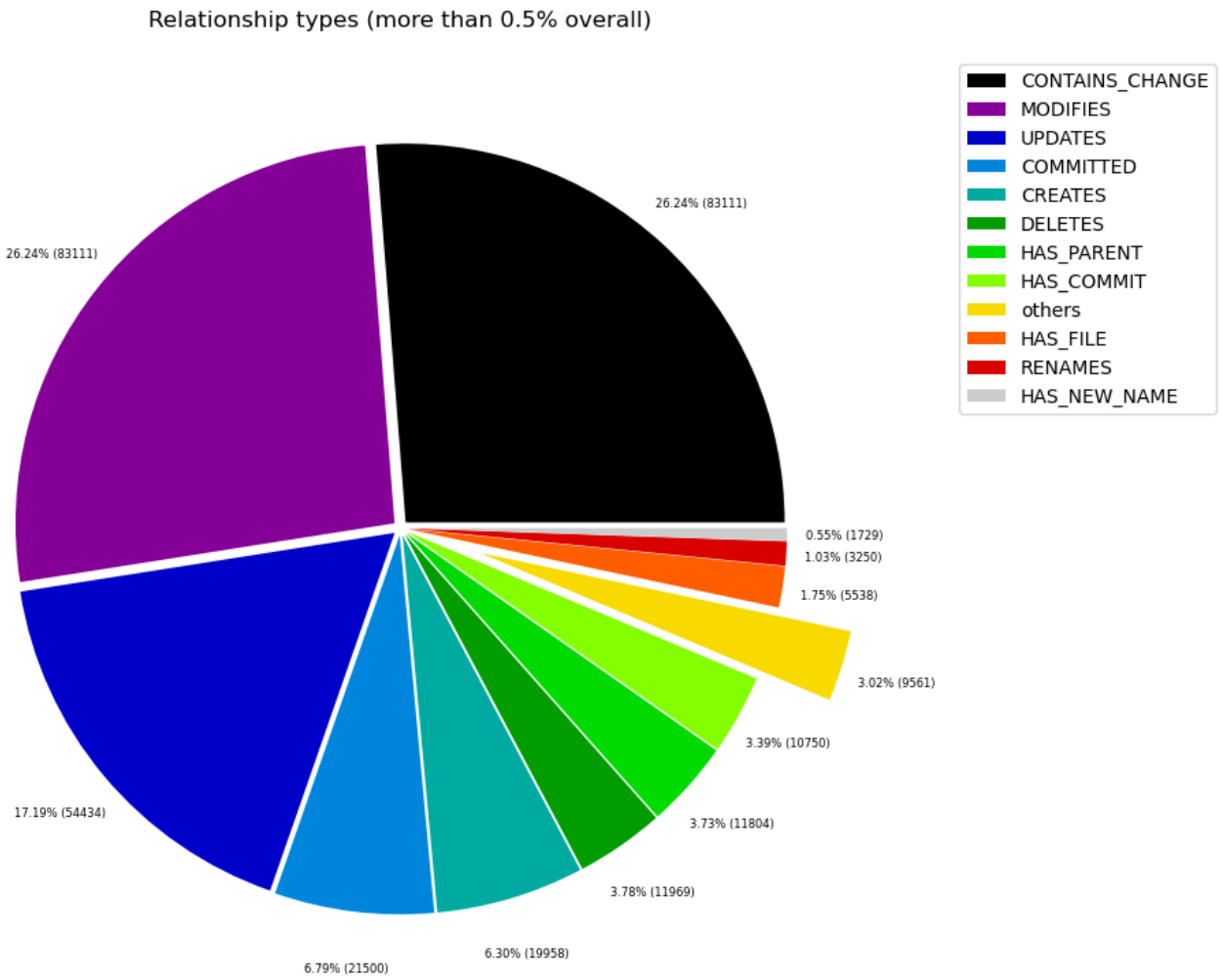


Table 2b - Lowest relationship count by type

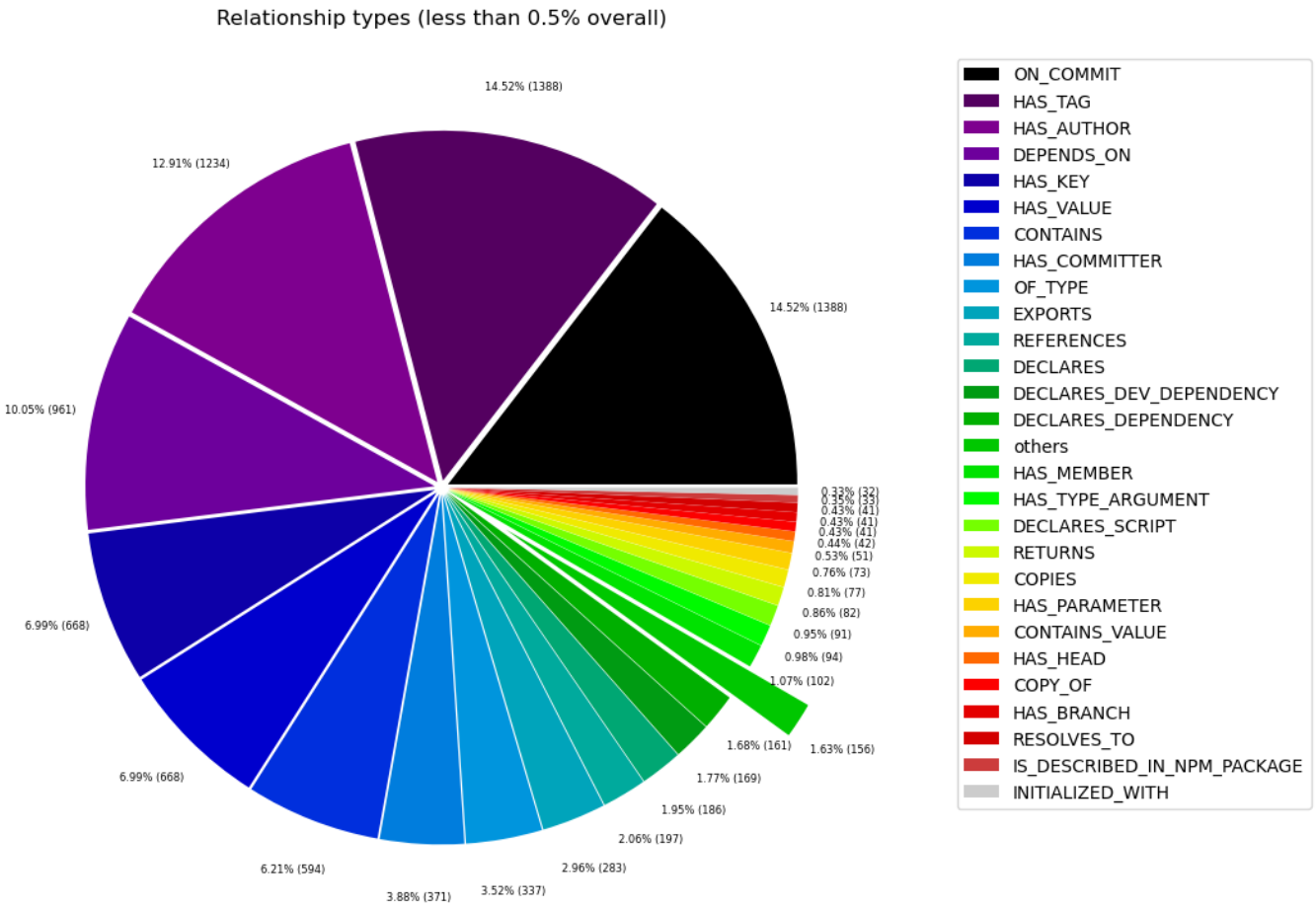
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	PROVIDED_BY_NPM_DEPENDENCY	1	0.000316
1	IS_IMPLEMENTED_IN	2	0.000631
2	CONSTRAINED_BY	4	0.001263
3	REFERENCED_PROJECTS	5	0.001579
4	CONTAINS_PROJECT	6	0.001894
5	DECLARES_ENGINE	6	0.001894
6	EXTENDS	6	0.001894
7	HAS_ARGUMENT	6	0.001894
8	CALLS	6	0.001894
9	HAS_NPM_PACKAGE	6	0.001894
10	HAS_ROOT	6	0.001894
11	MEMBER	6	0.001894
12	PARENT	6	0.001894
13	HAS_CONFIG	6	0.001894
14	SIMILAR	6	0.001894
15	DECLARES_PEER_DEPENDENCY	8	0.002526
16	INCLUDES_CONCEPT	19	0.005999
17	USES	23	0.007262
18	REQUIRES_CONCEPT	28	0.008841
19	INITIALIZED_WITH	32	0.010104
20	IS_DESCRIBED_IN_NPM_PACKAGE	33	0.010419
21	RESOLVES_TO	41	0.012945
22	HAS_BRANCH	41	0.012945
23	COPY_OF	41	0.012945
24	HAS_HEAD	42	0.013261
25	CONTAINS_VALUE	51	0.016103
26	HAS_PARAMETER	73	0.023049
27	COPIES	77	0.024312
28	RETURNS	82	0.025891
29	DECLARES_SCRIPT	91	0.028732

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]	83111	83111	
1	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	83111	10750	
2	[Git, Change]	UPDATES	[File, Git]	54434	83111	
3	[Git, Change]	CREATES	[File, Git]	19958	83111	
4	[Git, Change]	DELETES	[File, Git]	11969	83111	
5	[Git, Commit]	HAS_PARENT	[Git, Commit]	11804	10750	
6	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	10750	1	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	10750	1234	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10750	371	
9	[Repository, File, Git]	HAS_FILE	[File, Git]	5538	1	
10	[Git, Change]	RENAMES	[File, Git]	3250	83111	
11	[File, Git]	HAS_NEW_NAME	[File, Git]	1729	5538	
12	[Repository, File, Git]	HAS_TAG	[Git, Tag]	1388	1	
13	[Git, Tag]	ON_COMMIT	[Git, Commit]	1388	1388	
14	[Repository, File, Git]	HAS_AUTHOR	[Author, Git, Person]	1234	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): 105990
total_number_of_relationships (edges): 316715
-> total directed graph density: 2.8193107400839486e-05
-> total directed graph density in percent: 0.0028193107400839488
```