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]	84160	76.705737
1	[Git, Commit]	10851	9.889900
2	[File, Git]	5589	5.093968
3	[Git, Tag]	1457	1.327950
4	[Author, Git, Person]	1236	1.126524
5	[Type, TS, Primitive]	811	0.739168
6	[Json, Key]	668	0.608834
7	[Json, Value, Scalar]	603	0.549591
8	[Type, TS, Declared]	598	0.545034
9	[TS, ExternalDeclaration]	444	0.404674
10	[Committer, Git, Person]	371	0.338140
11	[NPM, Dependency]	338	0.308062
12	[Type, TS, ObjectMember]	318	0.289834
13	[Type, TS, Literal]	274	0.249731
14	[Type, TS, Union]	246	0.224211
15	[TS, Property]	137	0.124866
16	[Json, Value, Object]	133	0.121220
17	[Value, TS, Literal]	124	0.113017
18	[Type, Object, TS]	109	0.099346
19	[TS, Function]	109	0.099346
20	[NPM, Script]	91	0.082940
21	[Value, TS, ObjectMember]	88	0.080206
22	[Type, TS, FunctionParameter]	80	0.072914
23	[TS, Parameter]	76	0.069268
24	[Type, TS, Function]	70	0.063800
25	[File, Directory, Local]	64	0.058331
26	[TS, Variable]	59	0.053774
27	[File, TS, Local, Module]	46	0.041926
28	[Git, Branch]	46	0.041926
29	[TS, Interface]	37	0.033723

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>
```

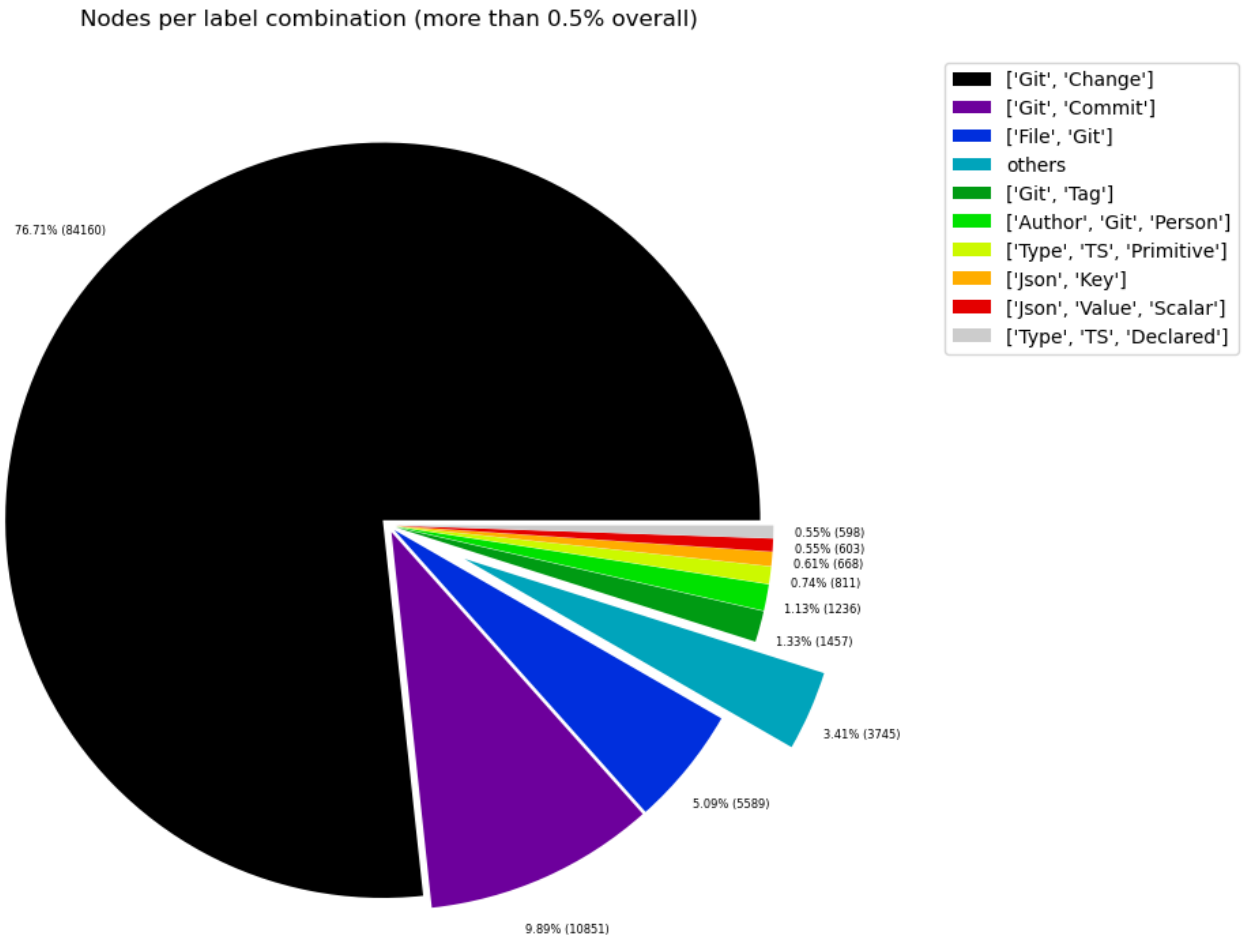


Table 1b - Lowest node count by label combination

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	0.000911
1	[Repository, File, Git]	1	0.000911
2	[Value, TS, Null]	1	0.000911
3	[TS, Constructor]	2	0.001823
4	[TS, Class]	2	0.001823
5	[TS, Enum]	4	0.003646
6	[TS, Method]	4	0.003646
7	[Value, Array, TS]	5	0.004557
8	[Type, TS, Tuple]	6	0.005469
9	[NPM, Engine]	6	0.005469
10	[TS, TypeParameter]	8	0.007291
11	[Value, TS, Complex]	11	0.010026
12	[Type, TS, TypeParameterReference]	12	0.010937
13	[Json, Value, Array]	12	0.010937
14	[Value, TS, Function]	13	0.011849
15	[Value, TS, Call]	14	0.012760
16	[Value, TS, Member]	14	0.012760
17	[TS, EnumMember]	16	0.014583
18	[jQAssistant, Rule, Concept]	19	0.017317
19	[Type, TS, NotIdentified]	23	0.020963
20	[Value, Object, TS]	28	0.025520
21	[File, Local]	28	0.025520
22	[File, TS, Scan]	29	0.026431
23	[Package, File, Json, NPM]	29	0.026431
24	[Value, TS, Declared]	30	0.027343
25	[TS, TypeAlias]	32	0.029166
26	[TS, ExternalModule]	33	0.030077
27	[Project, TS]	33	0.030077
28	[Type, TS, Intersection]	34	0.030989
29	[File, Directory]	35	0.031900

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>
```

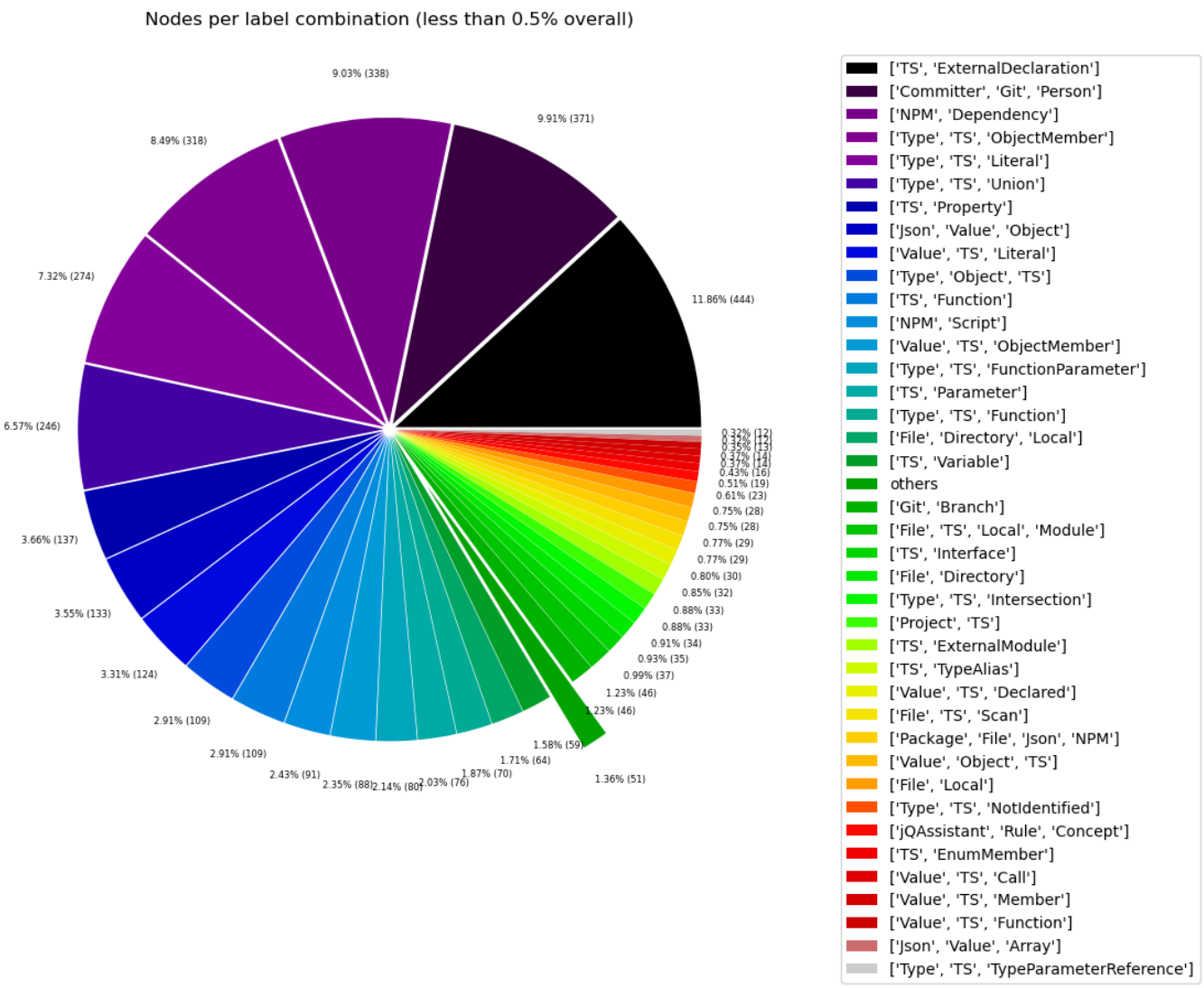


Table 1c - Highest node count by single label

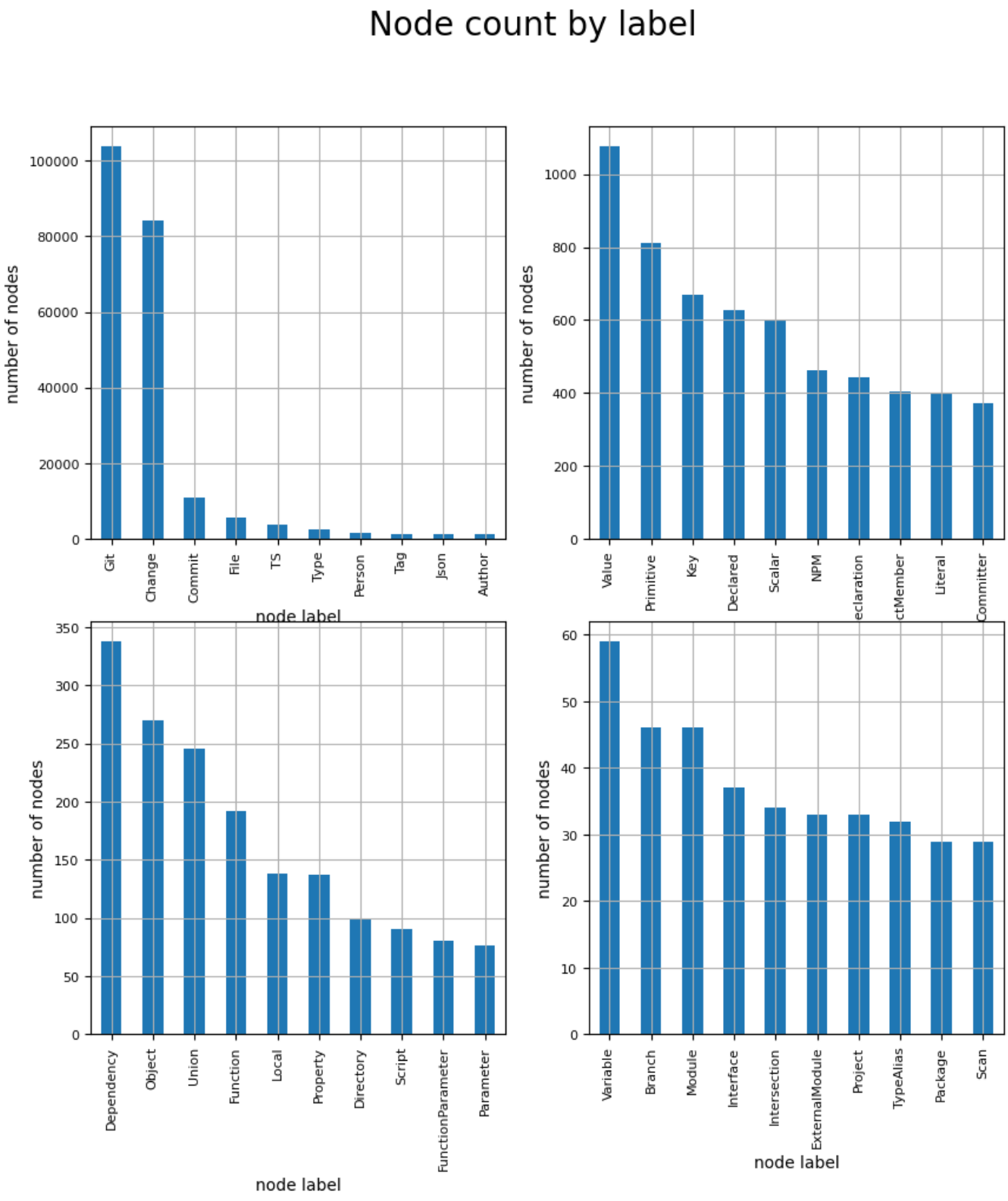
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	103711	94.525055
1	Change	84160	76.705737
2	Commit	10851	9.889900
3	File	5821	5.305419
4	TS	3980	3.627481
5	Type	2581	2.352394
6	Person	1607	1.464664
7	Tag	1457	1.327950
8	Json	1445	1.317013
9	Author	1236	1.126524
10	Value	1076	0.980696
11	Primitive	811	0.739168
12	Key	668	0.608834
13	Declared	628	0.572376
14	Scalar	603	0.549591
15	NPM	464	0.422902
16	ExternalDeclaration	444	0.404674
17	ObjectMember	406	0.370040
18	Literal	398	0.362748
19	Committer	371	0.338140
20	Dependency	338	0.308062
21	Object	270	0.246085
22	Union	246	0.224211
23	Function	192	0.174994
24	Local	138	0.125777
25	Property	137	0.124866
26	Directory	99	0.090231
27	Script	91	0.082940
28	FunctionParameter	80	0.072914
29	Parameter	76	0.069268
30	Variable	59	0.053774
31	Branch	46	0.041926
32	Module	46	0.041926
33	Interface	37	0.033723
34	Intersection	34	0.030989
35	ExternalModule	33	0.030077
36	Project	33	0.030077
37	TypeAlias	32	0.029166
38	Package	29	0.026431
39	Scan	29	0.026431

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: 324526
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	84160	25.933207
1	MODIFIES	84160	25.933207
2	UPDATES	55224	17.016818
3	COMMITTED	21702	6.687292
4	CREATES	20144	6.207207
5	DELETES	12089	3.725125
6	HAS_PARENT	11915	3.671509
7	HAS_COMMIT	10851	3.343646
8	HAS_FILE	5589	1.722204
9	RENAMES	3297	1.015943
10	DEPENDS_ON	1823	0.561742
11	HAS_NEW_NAME	1756	0.541097
12	HAS_TAG	1457	0.448962
13	ON_COMMIT	1457	0.448962
14	HAS_AUTHOR	1236	0.380863
15	CONTAINS	1199	0.369462
16	OF_TYPE	1030	0.317386
17	HAS_KEY	668	0.205839
18	HAS_VALUE	668	0.205839
19	EXPORTS	651	0.200600
20	REFERENCES	489	0.150681
21	DECLARES	410	0.126338
22	HAS_MEMBER	406	0.125106
23	HAS_COMMITTER	371	0.114321
24	HAS_TYPE_ARGUMENT	202	0.062245
25	RETURNS	183	0.056390
26	DECLARES_DEV_DEPENDENCY	169	0.052076
27	DECLARES_DEPENDENCY	161	0.049611
28	HAS_PARAMETER	155	0.047762
29	RESOLVES_TO	103	0.031739

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>
```

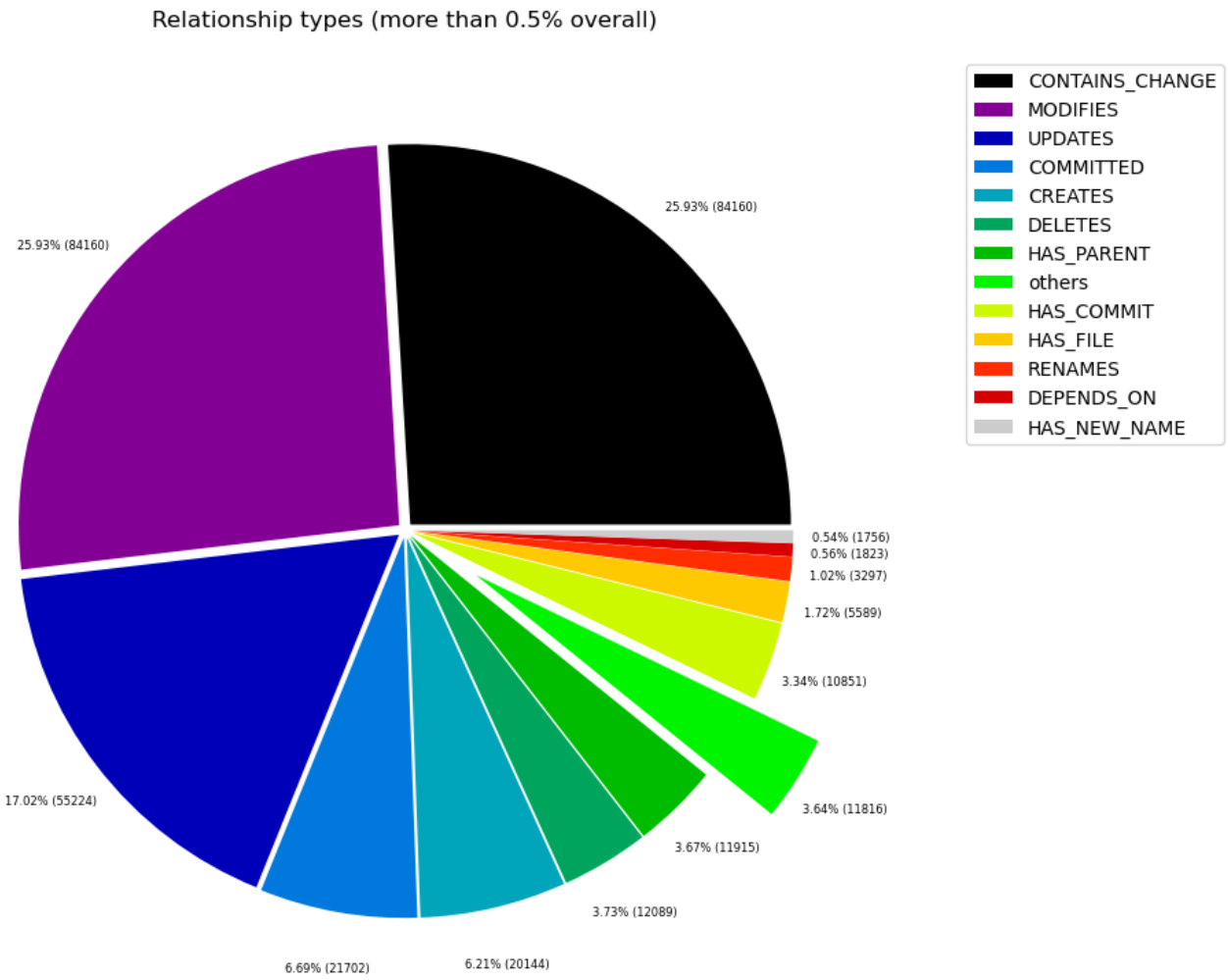


Table 2b - Lowest relationship count by type

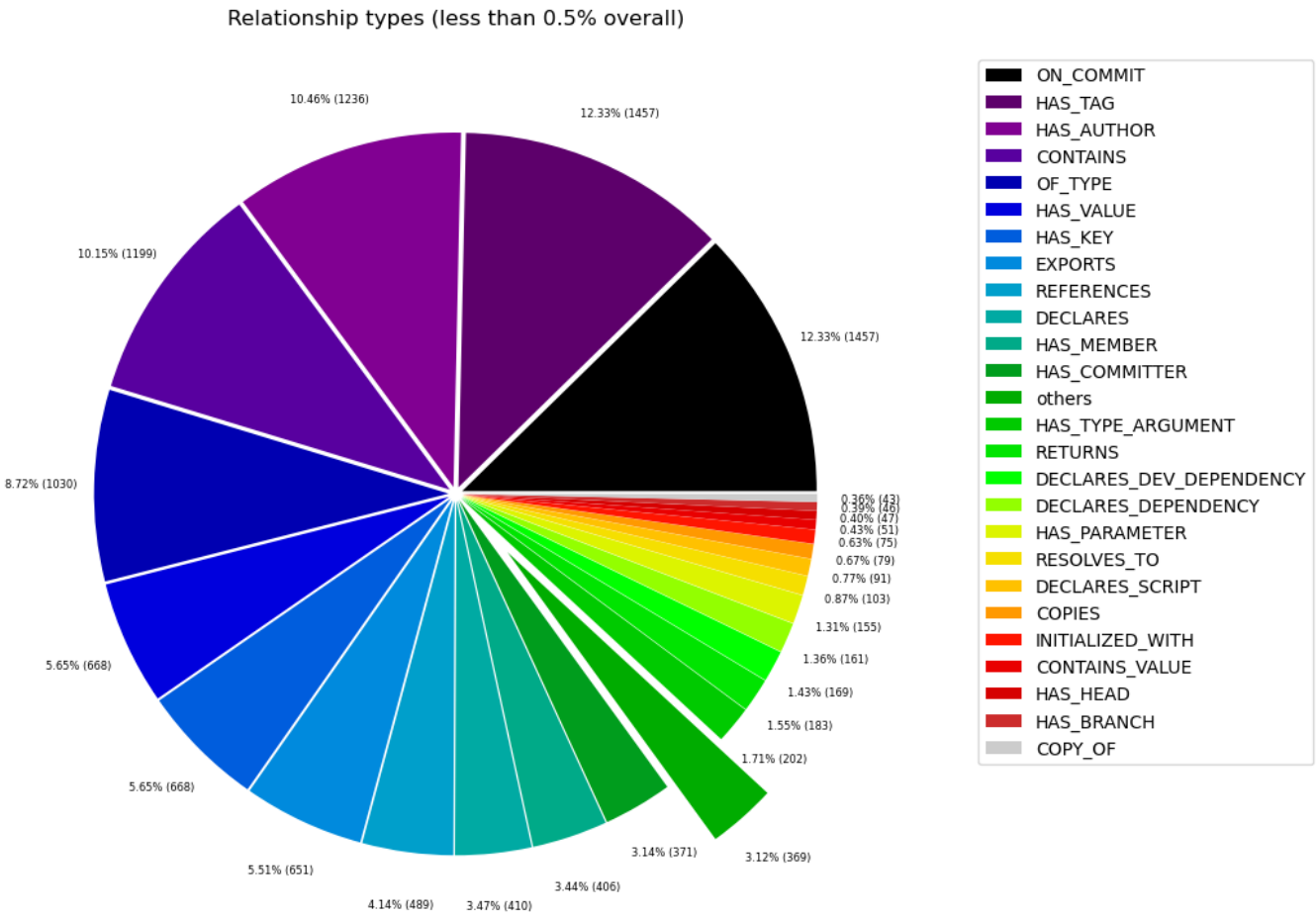
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	HAS	1	0.000308
1	REFERENCED_PROJECTS	5	0.001541
2	DECLARES_ENGINE	6	0.001849
3	SIMILAR	8	0.002465
4	DECLARES_PEER_DEPENDENCY	8	0.002465
5	CONSTRAINED_BY	8	0.002465
6	EXTENDS	12	0.003698
7	PARENT	14	0.004314
8	MEMBER	14	0.004314
9	HAS_ARGUMENT	14	0.004314
10	CALLS	14	0.004314
11	INCLUDES_CONCEPT	19	0.005855
12	PROVIDED_BY_NPM_DEPENDENCY	20	0.006163
13	REQUIRES_CONCEPT	28	0.008628
14	CONTAINS_PROJECT	33	0.010169
15	HAS_CONFIG	33	0.010169
16	IS_DESCRIBED_IN_NPM_PACKAGE	33	0.010169
17	HAS_ROOT	33	0.010169
18	HAS_NPM_PACKAGE	33	0.010169
19	USES	33	0.010169
20	COPY_OF	43	0.013250
21	HAS_BRANCH	46	0.014175
22	HAS_HEAD	47	0.014483
23	CONTAINS_VALUE	51	0.015715
24	INITIALIZED_WITH	75	0.023111
25	COPIES	79	0.024343
26	DECLARES_SCRIPT	91	0.028041
27	RESOLVES_TO	103	0.031739
28	HAS_PARAMETER	155	0.047762
29	DECLARES_DEPENDENCY	161	0.049611

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, Commit]	CONTAINS_CHANGE	[Git, Change]	84160	10851	
1	[Git, Change]	MODIFIES	[File, Git]	84160	84160	
2	[Git, Change]	UPDATES	[File, Git]	55224	84160	
3	[Git, Change]	CREATES	[File, Git]	20144	84160	
4	[Git, Change]	DELETES	[File, Git]	12089	84160	
5	[Git, Commit]	HAS_PARENT	[Git, Commit]	11915	10851	
6	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	10851	1	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	10851	1236	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10851	371	
9	[Repository, File, Git]	HAS_FILE	[File, Git]	5589	1	
10	[Git, Change]	RENAMES	[File, Git]	3297	84160	
11	[File, Git]	HAS_NEW_NAME	[File, Git]	1756	5589	
12	[Repository, File, Git]	HAS_TAG	[Git, Tag]	1457	1	
13	[Git, Tag]	ON_COMMIT	[Git, Commit]	1457	1457	
14	[Repository, File, Git]	HAS_AUTHOR	[Author, Git, Person]	1236	1	
15	[Json, Value, Object]	HAS_KEY	[Json, Key]	668	133	
16	[TS, Function]	DEPENDS_ON	[TS, ExternalDeclaration]	580	109	
17	[Json, Key]	HAS_VALUE	[Json, Value, Scalar]	552	668	
18	[TS, ExternalModule]	EXPORTS	[TS, ExternalDeclaration]	444	33	
19	[Repository, File, Git]	HAS_COMMITTER	[Committer, Git, Person]	371	1	
20	[Type, Object, TS]	HAS_MEMBER	[Type, TS, ObjectMember]	318	109	
21	[File, TS, Local, Module, Mark4ModuleWeaklyCon...	DEPENDS_ON	[TS, ExternalDeclaration]	308	1	
22	[Type, TS, Union]	CONTAINS	[Type, TS, Primitive]	303	246	
23	[Type, TS, Declared]	REFERENCES	[TS, ExternalDeclaration]	288	598	
24	[Type, TS, Union]	CONTAINS	[Type, TS, Literal]	238	246	
25	[Type, TS, ObjectMember]	OF_TYPE	[Type, TS, Primitive]	173	318	
26	[Package, File, Json, NPM]	DECLARES_DEV_DEPENDENCY	[NPM, Dependency]	169	29	
27	[Package, File, Json, NPM]	DECLARES_DEPENDENCY	[NPM, Dependency]	161	29	
28	[TS, Function]	DEPENDS_ON	[TS, ExternalModule]	148	109	
29	[Type, TS, Union]	CONTAINS	[Type, TS, Declared]	145	246	

Graph Density

```
total_number_of_nodes (vertices): 109718
total_number_of_relationships (edges): 324526
-> total directed graph density: 2.6958622057942908e-05
-> total directed graph density in percent: 0.0026958622057942907
```