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]	77092	77.962057
1	[Git, Commit]	10261	10.376805
2	[File, Git]	5224	5.282958
3	[Author, Git, Person]	1191	1.204442
4	[Git, Tag]	1165	1.178148
5	[Json, Key]	668	0.675539
6	[Json, Value, Scalar]	603	0.609805
7	[Committer, Git, Person]	371	0.375187
8	[NPM, Dependency]	330	0.333724
9	[Type, TS, Primitive]	291	0.294284
10	[Type, TS, Declared]	276	0.279115
11	[TS, ExternalDeclaration]	215	0.217426
12	[Type, TS, Literal]	136	0.137535
13	[Json, Value, Object]	133	0.134501
14	[Type, TS, Union]	119	0.120343
15	[Type, TS, ObjectMember]	101	0.102140
16	[NPM, Script]	91	0.092027
17	[TS, Property]	65	0.065734
18	[TS, Function]	47	0.047530
19	[Type, TS, FunctionParameter]	40	0.040451
20	[Type, Object, TS]	39	0.039440
21	[File, Directory]	34	0.034384
22	[Type, TS, Function]	34	0.034384
23	[TS, Parameter]	33	0.033372
24	[Package, File, Json, NPM]	29	0.029327
25	[Git, Branch]	29	0.029327
26	[TS, ExternalModule]	25	0.025282
27	[TS, Variable]	24	0.024271
28	[Value, TS, Literal]	20	0.020226
29	[jQAssistant, Rule, Concept]	19	0.019214

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>
```

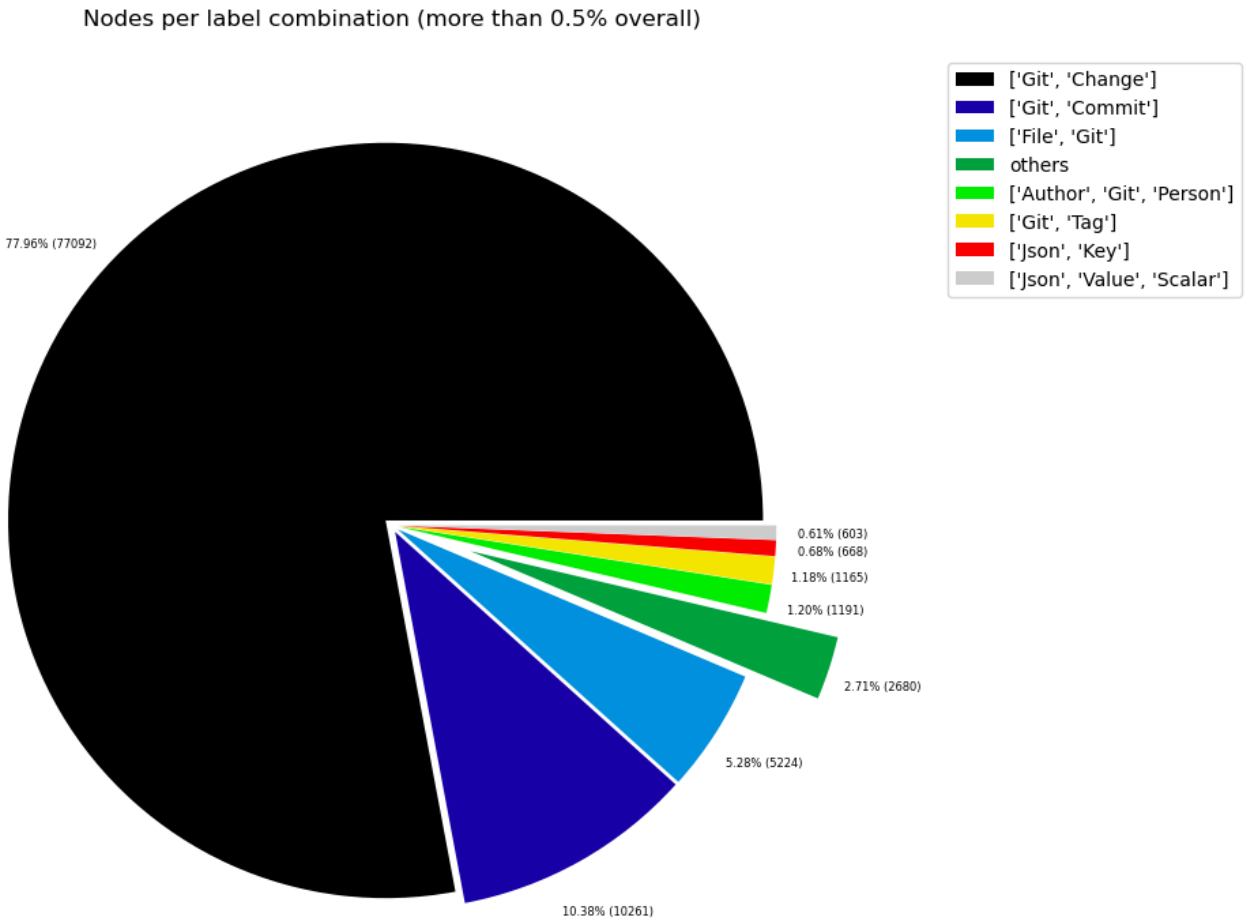


Table 1b - Lowest node count by label combination

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	0.001011
1	[File, TS, Scan]	1	0.001011
2	[TS, Method]	1	0.001011
3	[Value, TS, ObjectMember]	1	0.001011
4	[TS, Constructor]	1	0.001011
5	[TS, Class]	1	0.001011
6	[TS, Enum]	2	0.002023
7	[Value, Object, TS]	3	0.003034
8	[Type, TS, Tuple]	3	0.003034
9	[Value, TS, Function]	4	0.004045
10	[TS, TypeParameter]	4	0.004045
11	[Value, TS, Complex]	5	0.005056
12	[NPM, Engine]	6	0.006068
13	[Project, TS]	6	0.006068
14	[File, Local]	6	0.006068
15	[Value, TS, Call]	6	0.006068
16	[Value, TS, Member]	6	0.006068
17	[File, TS, Local, Module]	6	0.006068
18	[Type, TS, TypeParameterReference]	6	0.006068
19	[TS, EnumMember]	8	0.008090
20	[Type, TS, NotIdentified]	11	0.011124
21	[Json, Value, Array]	12	0.012135
22	[Value, TS, Declared]	13	0.013147
23	[TS, TypeAlias]	16	0.016181
24	[File, Directory, Local]	16	0.016181
25	[TS, Interface]	17	0.017192
26	[Type, TS, Intersection]	17	0.017192
27	[jQAssistant, Rule, Concept]	19	0.019214
28	[Value, TS, Literal]	20	0.020226
29	[TS, Variable]	24	0.024271

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>
```

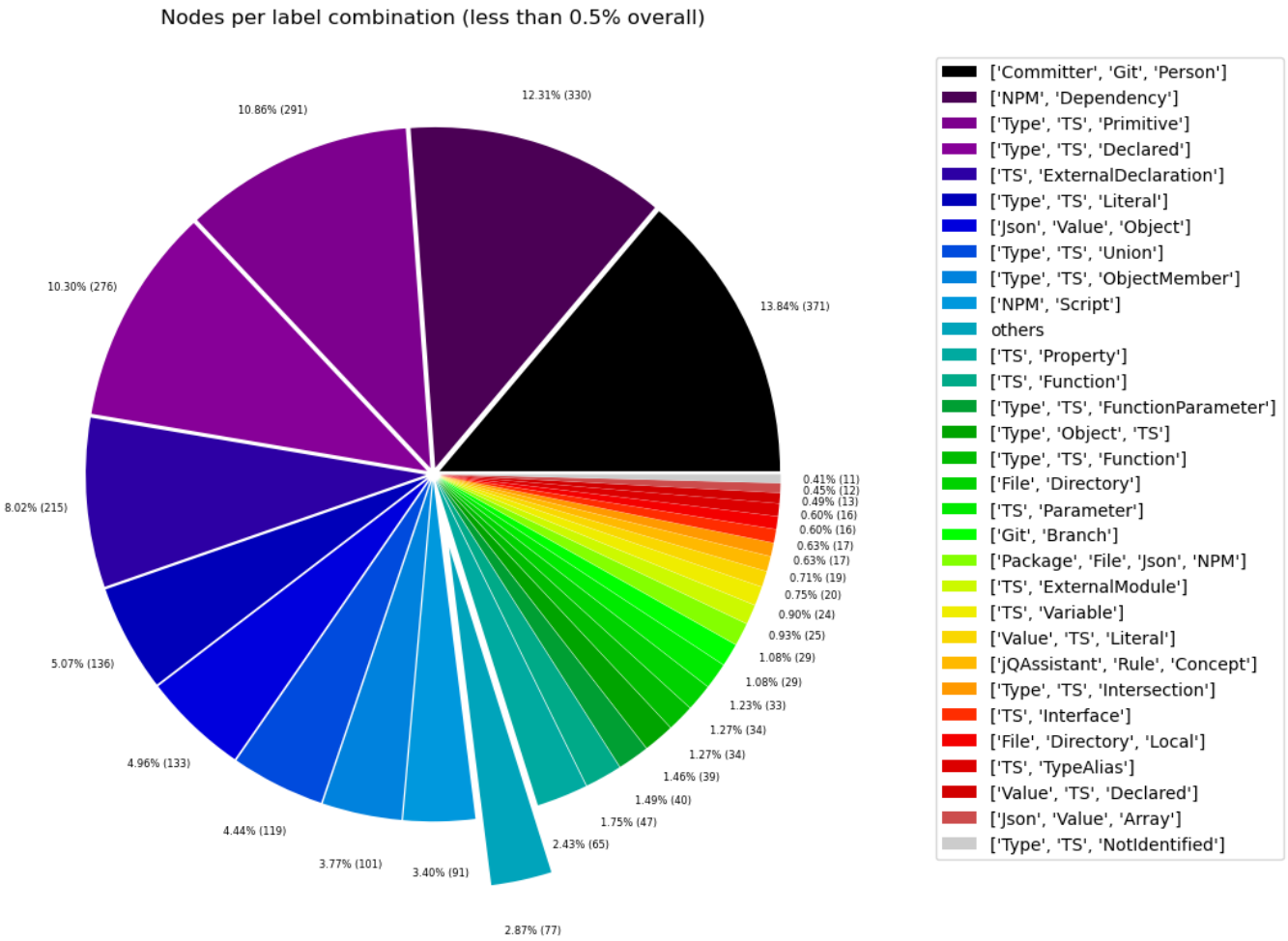


Table 1c - Highest node count by single label

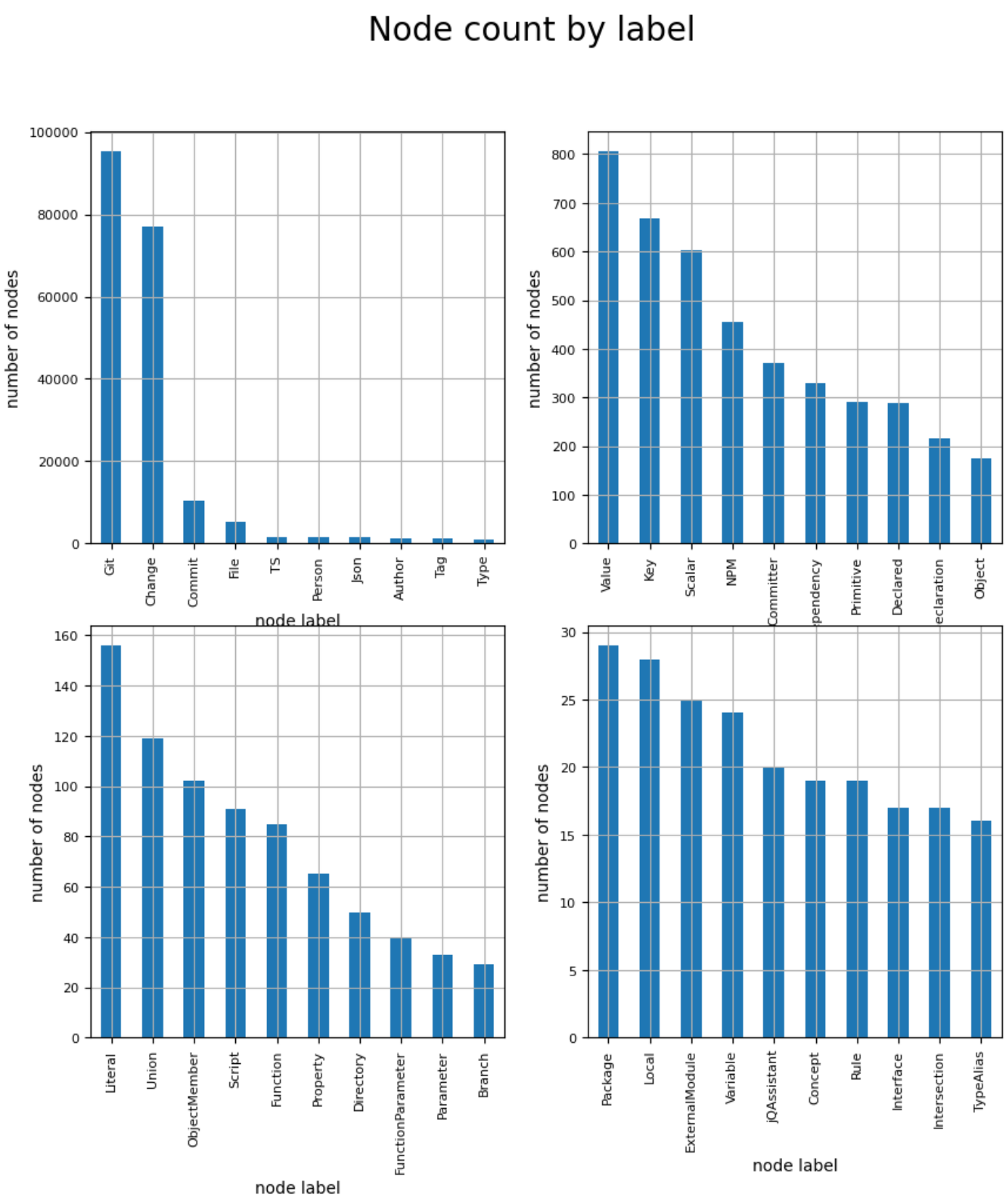
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	95333	96.408924
1	Change	77092	77.962057
2	Commit	10261	10.376805
3	File	5316	5.375996
4	TS	1603	1.621091
5	Person	1562	1.579629
6	Json	1445	1.461308
7	Author	1191	1.204442
8	Tag	1165	1.178148
9	Type	1073	1.085110
10	Value	806	0.815096
11	Key	668	0.675539
12	Scalar	603	0.609805
13	NPM	456	0.461146
14	Committer	371	0.375187
15	Dependency	330	0.333724
16	Primitive	291	0.294284
17	Declared	289	0.292262
18	ExternalDeclaration	215	0.217426
19	Object	175	0.176975
20	Literal	156	0.157761
21	Union	119	0.120343
22	ObjectMember	102	0.103151
23	Script	91	0.092027
24	Function	85	0.085959
25	Property	65	0.065734
26	Directory	50	0.050564
27	FunctionParameter	40	0.040451
28	Parameter	33	0.033372
29	Branch	29	0.029327
30	Package	29	0.029327
31	Local	28	0.028316
32	ExternalModule	25	0.025282
33	Variable	24	0.024271
34	jQAssistant	20	0.020226
35	Concept	19	0.019214
36	Rule	19	0.019214
37	Interface	17	0.017192
38	Intersection	17	0.017192
39	TypeAlias	16	0.016181

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: 276509
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	77092	27.880467
1	MODIFIES	77092	27.880467
2	UPDATES	51036	18.457265
3	COMMITTED	20522	7.421820
4	CREATES	18232	6.593637
5	HAS_PARENT	11289	4.082688
6	DELETES	10614	3.838573
7	RENAMES	2790	1.009009
8	HAS_NEW_NAME	1581	0.571772
9	ON_COMMIT	1165	0.421324
10	DEPENDS_ON	959	0.346824
11	HAS_KEY	668	0.241583
12	HAS_VALUE	668	0.241583
13	CONTAINS	593	0.214460
14	OF_TYPE	337	0.121877
15	EXPORTS	276	0.099816
16	REFERENCES	197	0.071245
17	DECLARES	186	0.067267
18	DECLARES_DEV_DEPENDENCY	169	0.061119
19	DECLARES_DEPENDENCY	161	0.058226
20	HAS_MEMBER	102	0.036888
21	HAS_TYPE_ARGUMENT	94	0.033995
22	DECLARES_SCRIPT	91	0.032910
23	RETURNS	82	0.029655
24	RESOLVES_TO	77	0.027847
25	HAS_PARAMETER	73	0.026401
26	CONTAINS_VALUE	51	0.018444
27	COPIES	43	0.015551
28	INITIALIZED_WITH	32	0.011573
29	IS_DESCRIBED_IN_NPM_PACKAGE	32	0.011573

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>
```

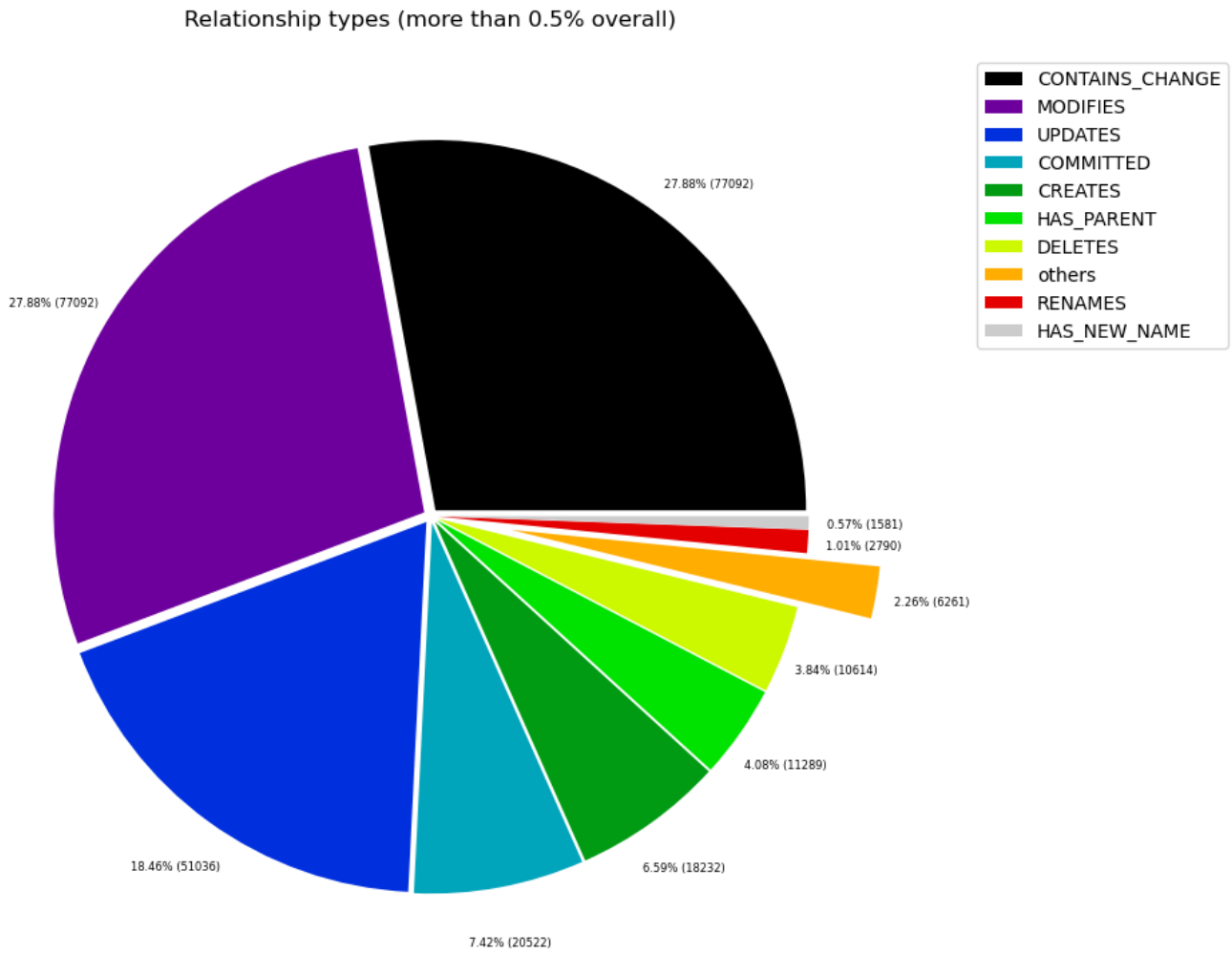


Table 2b - Lowest relationship count by type

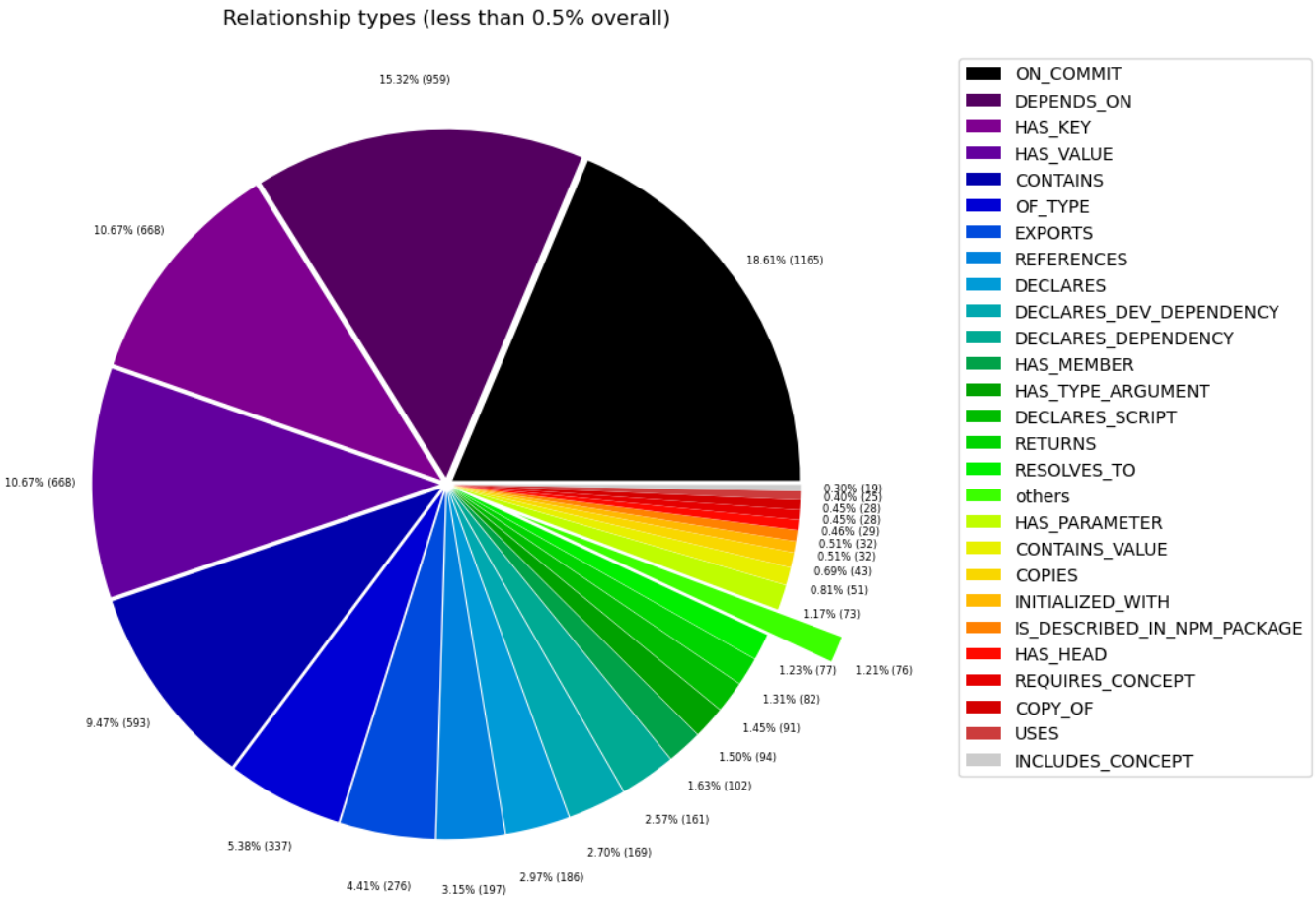
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	PROVIDED_BY_NPM_DEPENDENCY	1	0.000362
1	CONSTRAINED_BY	4	0.001447
2	REFERENCED_PROJECTS	5	0.001808
3	SIMILAR	6	0.002170
4	DECLARES_ENGINE	6	0.002170
5	EXTENDS	6	0.002170
6	HAS_ARGUMENT	6	0.002170
7	CALLS	6	0.002170
8	HAS_NPM_PACKAGE	6	0.002170
9	HAS_ROOT	6	0.002170
10	MEMBER	6	0.002170
11	PARENT	6	0.002170
12	HAS_CONFIG	6	0.002170
13	CONTAINS_PROJECT	6	0.002170
14	INCLUDES_CONCEPT	19	0.006871
15	USES	25	0.009041
16	REQUIRES_CONCEPT	28	0.010126
17	COPY_OF	28	0.010126
18	HAS_HEAD	29	0.010488
19	IS_DESCRIBED_IN_NPM_PACKAGE	32	0.011573
20	INITIALIZED_WITH	32	0.011573
21	COPIES	43	0.015551
22	CONTAINS_VALUE	51	0.018444
23	HAS_PARAMETER	73	0.026401
24	RESOLVES_TO	77	0.027847
25	RETURNS	82	0.029655
26	DECLARES_SCRIPT	91	0.032910
27	HAS_TYPE_ARGUMENT	94	0.033995
28	HAS_MEMBER	102	0.036888
29	DECLARES_DEPENDENCY	161	0.058226

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, Commit]	CONTAINS_CHANGE	[Git, Change]	77092	10261	
1	[Git, Change]	MODIFIES	[File, Git]	77092	77092	
2	[Git, Change]	UPDATES	[File, Git]	51036	77092	
3	[Git, Change]	CREATES	[File, Git]	18232	77092	
4	[Git, Commit]	HAS_PARENT	[Git, Commit]	11289	10261	
5	[Git, Change]	DELETES	[File, Git]	10614	77092	
6	[Author, Git, Person]	COMMITTED	[Git, Commit]	10261	1191	
7	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10261	371	
8	[Git, Change]	RENAMES	[File, Git]	2790	77092	
9	[File, Git]	HAS_NEW_NAME	[File, Git]	1581	5224	
10	[Git, Tag]	ON_COMMIT	[Git, Commit]	1165	1165	
11	[Json, Value, Object]	HAS_KEY	[Json, Key]	668	133	
12	[Json, Key]	HAS_VALUE	[Json, Value, Scalar]	552	668	
13	[TS, Function]	DEPENDS_ON	[TS, ExternalDeclaration]	285	47	
14	[TS, ExternalModule]	EXPORTS	[TS, ExternalDeclaration]	215	25	
15	[File, TS, Local, Module, Mark4ModuleWeaklyCon...	DEPENDS_ON	[TS, ExternalDeclaration]	192	2	
16	[Package, File, Json, NPM]	DECLARES_DEV_DEPENDENCY	[NPM, Dependency]	169	29	
17	[Package, File, Json, NPM]	DECLARES_DEPENDENCY	[NPM, Dependency]	161	29	
18	[Type, TS, Union]	CONTAINS	[Type, TS, Primitive]	147	119	
19	[Type, TS, Declared]	REFERENCES	[TS, ExternalDeclaration]	142	276	
20	[TS, Function]	DEPENDS_ON	[TS, ExternalModule]	131	47	
21	[Type, TS, Union]	CONTAINS	[Type, TS, Literal]	119	119	
22	[Json, Key]	HAS_VALUE	[Json, Value, Object]	104	668	
23	[Type, Object, TS]	HAS_MEMBER	[Type, TS, ObjectMember]	101	39	
24	[Package, File, Json, NPM]	DECLARES_SCRIPT	[NPM, Script]	91	29	
25	[Type, TS, Union]	CONTAINS	[Type, TS, Declared]	70	119	
26	[File, Directory]	CONTAINS	[File, Directory]	63	34	
27	[TS, Interface]	DECLARES	[TS, Property]	61	17	
28	[File, Directory]	CONTAINS	[Package, File, Json, NPM]	58	34	
29	[File, Git]	RESOLVES_TO	[Package, File, Json, NPM]	57	5224	

Graph Density

```
total_number_of_nodes (vertices): 98884
total_number_of_relationships (edges): 276509
-> total directed graph density: 2.8278841366530634e-05
-> total directed graph density in percent: 0.0028278841366530636
```