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]	77121	77.955120
1	[Git, Commit]	10264	10.375013
2	[File, Git]	5237	5.293642
3	[Author, Git, Person]	1191	1.203882
4	[Git, Tag]	1165	1.177600
5	[Json, Key]	668	0.675225
6	[Json, Value, Scalar]	603	0.609522
7	[Committer, Git, Person]	371	0.375013
8	[NPM, Dependency]	330	0.333569
9	[Type, TS, Primitive]	291	0.294147
10	[Type, TS, Declared]	276	0.278985
11	[TS, ExternalDeclaration]	215	0.217325
12	[Type, TS, Literal]	136	0.137471
13	[Json, Value, Object]	133	0.134438
14	[Type, TS, Union]	119	0.120287
15	[Type, TS, ObjectMember]	101	0.102092
16	[NPM, Script]	91	0.091984
17	[TS, Property]	65	0.065703
18	[TS, Function]	47	0.047508
19	[Type, TS, FunctionParameter]	40	0.040433
20	[Type, Object, TS]	39	0.039422
21	[File, Directory]	34	0.034368
22	[Type, TS, Function]	34	0.034368
23	[TS, Parameter]	33	0.033357
24	[Git, Branch]	30	0.030324
25	[Package, File, Json, NPM]	29	0.029314
26	[TS, ExternalModule]	25	0.025270
27	[TS, Variable]	24	0.024260
28	[Value, TS, Literal]	20	0.020216
29	[jQAssistant, Rule, Concept]	19	0.019205

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>
```

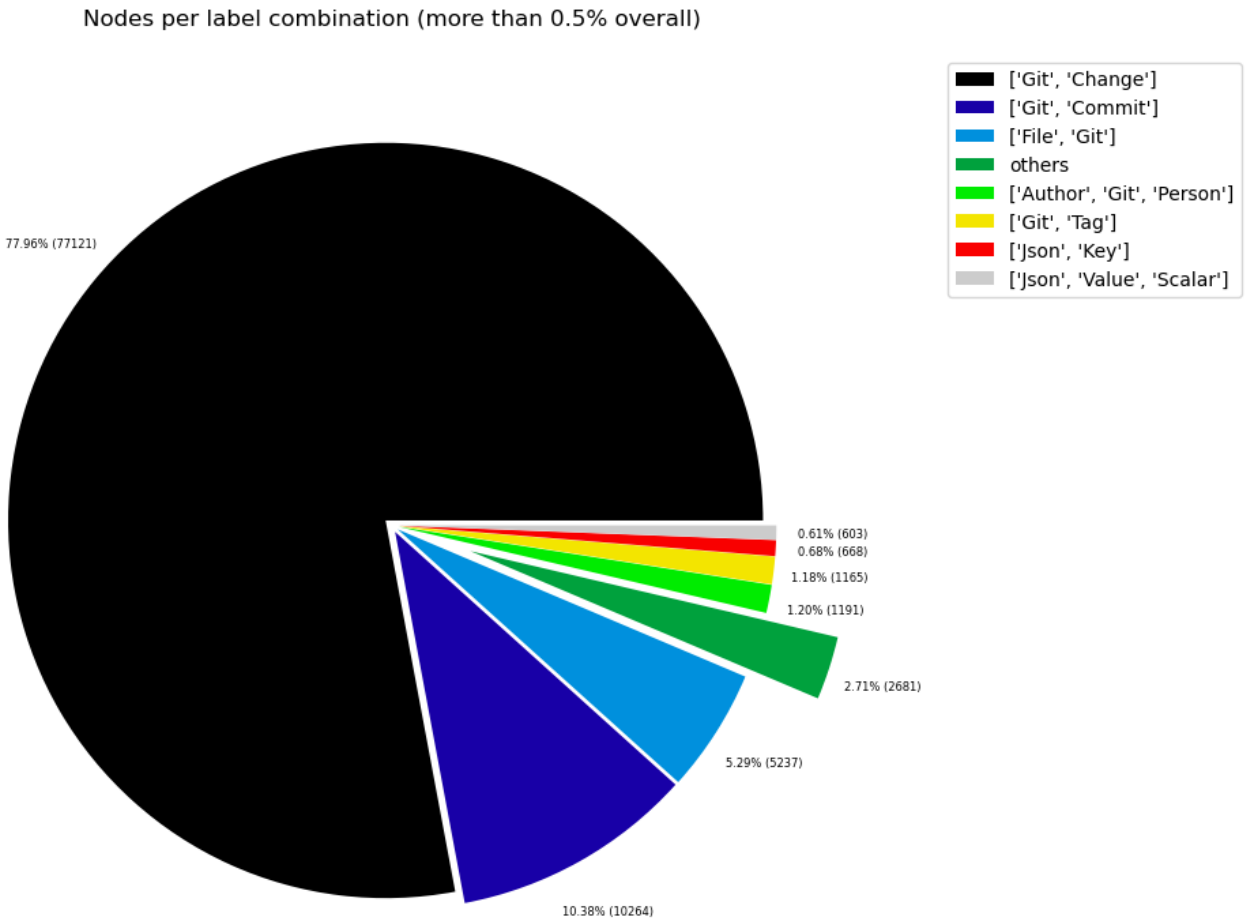


Table 1b - Lowest node count by label combination

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	0.002022
7	[Value, Object, TS]	3	0.003032
8	[Type, TS, Tuple]	3	0.003032
9	[Value, TS, Function]	4	0.004043
10	[TS, TypeParameter]	4	0.004043
11	[Value, TS, Complex]	5	0.005054
12	[NPM, Engine]	6	0.006065
13	[Project, TS]	6	0.006065
14	[File, Local]	6	0.006065
15	[Value, TS, Call]	6	0.006065
16	[Value, TS, Member]	6	0.006065
17	[File, TS, Local, Module]	6	0.006065
18	[Type, TS, TypeParameterReference]	6	0.006065
19	[TS, EnumMember]	8	0.008087
20	[Type, TS, NotIdentified]	11	0.011119
21	[Json, Value, Array]	12	0.012130
22	[Value, TS, Declared]	13	0.013141
23	[TS, TypeAlias]	16	0.016173
24	[File, Directory, Local]	16	0.016173
25	[TS, Interface]	17	0.017184
26	[Type, TS, Intersection]	17	0.017184
27	[jQAssistant, Rule, Concept]	19	0.019205
28	[Value, TS, Literal]	20	0.020216
29	[TS, Variable]	24	0.024260

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>
```

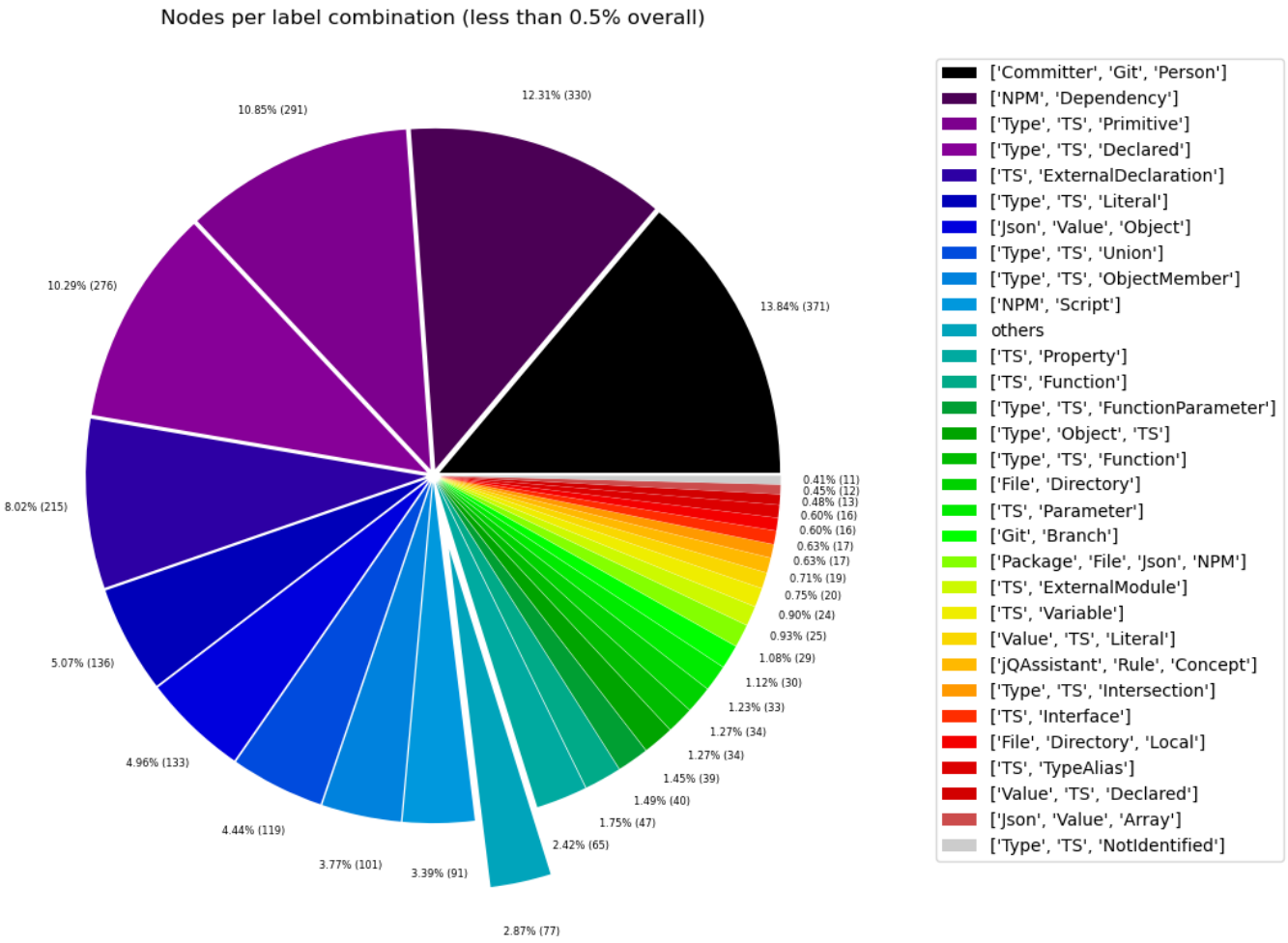


Table 1c - Highest node count by single label

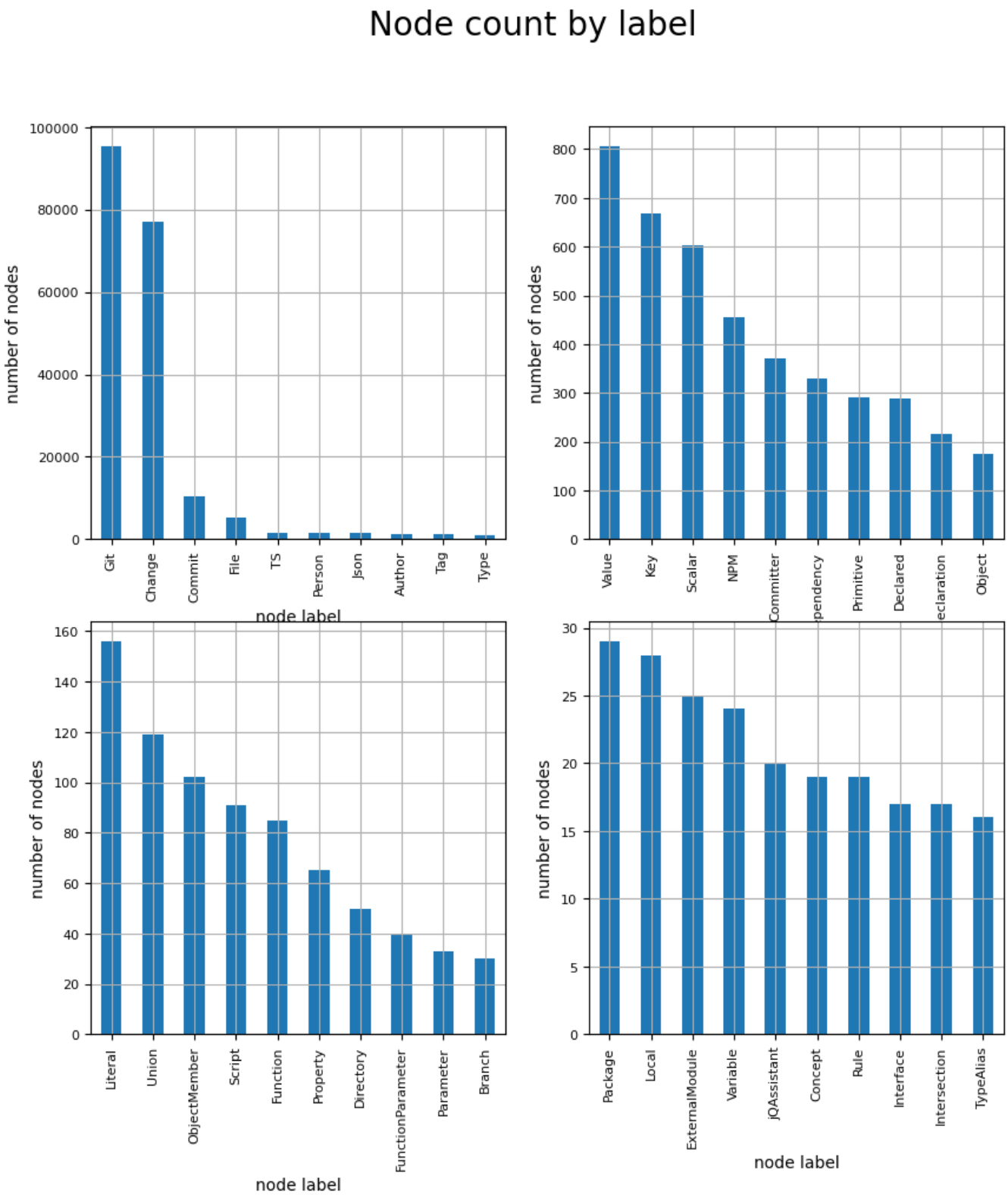
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	95379	96.410593
1	Change	77121	77.955120
2	Commit	10264	10.375013
3	File	5329	5.386637
4	TS	1603	1.620338
5	Person	1562	1.578894
6	Json	1445	1.460629
7	Author	1191	1.203882
8	Tag	1165	1.177600
9	Type	1073	1.084605
10	Value	806	0.814717
11	Key	668	0.675225
12	Scalar	603	0.609522
13	NPM	456	0.460932
14	Committer	371	0.375013
15	Dependency	330	0.333569
16	Primitive	291	0.294147
17	Declared	289	0.292126
18	ExternalDeclaration	215	0.217325
19	Object	175	0.176893
20	Literal	156	0.157687
21	Union	119	0.120287
22	ObjectMember	102	0.103103
23	Script	91	0.091984
24	Function	85	0.085919
25	Property	65	0.065703
26	Directory	50	0.050541
27	FunctionParameter	40	0.040433
28	Parameter	33	0.033357
29	Branch	30	0.030324
30	Package	29	0.029314
31	Local	28	0.028303
32	ExternalModule	25	0.025270
33	Variable	24	0.024260
34	jQAssistant	20	0.020216
35	Concept	19	0.019205
36	Rule	19	0.019205
37	Interface	17	0.017184
38	Intersection	17	0.017184
39	TypeAlias	16	0.016173

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: 276639
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	77121	27.877848
1	MODIFIES	77121	27.877848
2	UPDATES	51049	18.453291
3	COMMITTED	20528	7.420501
4	CREATES	18245	6.595238
5	HAS_PARENT	11292	4.081854
6	DELETES	10628	3.841830
7	RENAMES	2801	1.012511
8	HAS_NEW_NAME	1592	0.575479
9	ON_COMMIT	1165	0.421126
10	DEPENDS_ON	959	0.346661
11	HAS_KEY	668	0.241470
12	HAS_VALUE	668	0.241470
13	CONTAINS	593	0.214359
14	OF_TYPE	337	0.121819
15	EXPORTS	276	0.099769
16	REFERENCES	197	0.071212
17	DECLARES	186	0.067236
18	DECLARES_DEV_DEPENDENCY	169	0.061090
19	DECLARES_DEPENDENCY	161	0.058199
20	HAS_MEMBER	102	0.036871
21	HAS_TYPE_ARGUMENT	94	0.033979
22	DECLARES_SCRIPT	91	0.032895
23	RETURNS	82	0.029642
24	RESOLVES_TO	77	0.027834
25	HAS_PARAMETER	73	0.026388
26	CONTAINS_VALUE	51	0.018436
27	COPIES	43	0.015544
28	INITIALIZED_WITH	32	0.011567
29	IS_DESCRIBED_IN_NPM_PACKAGE	32	0.011567

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>
```

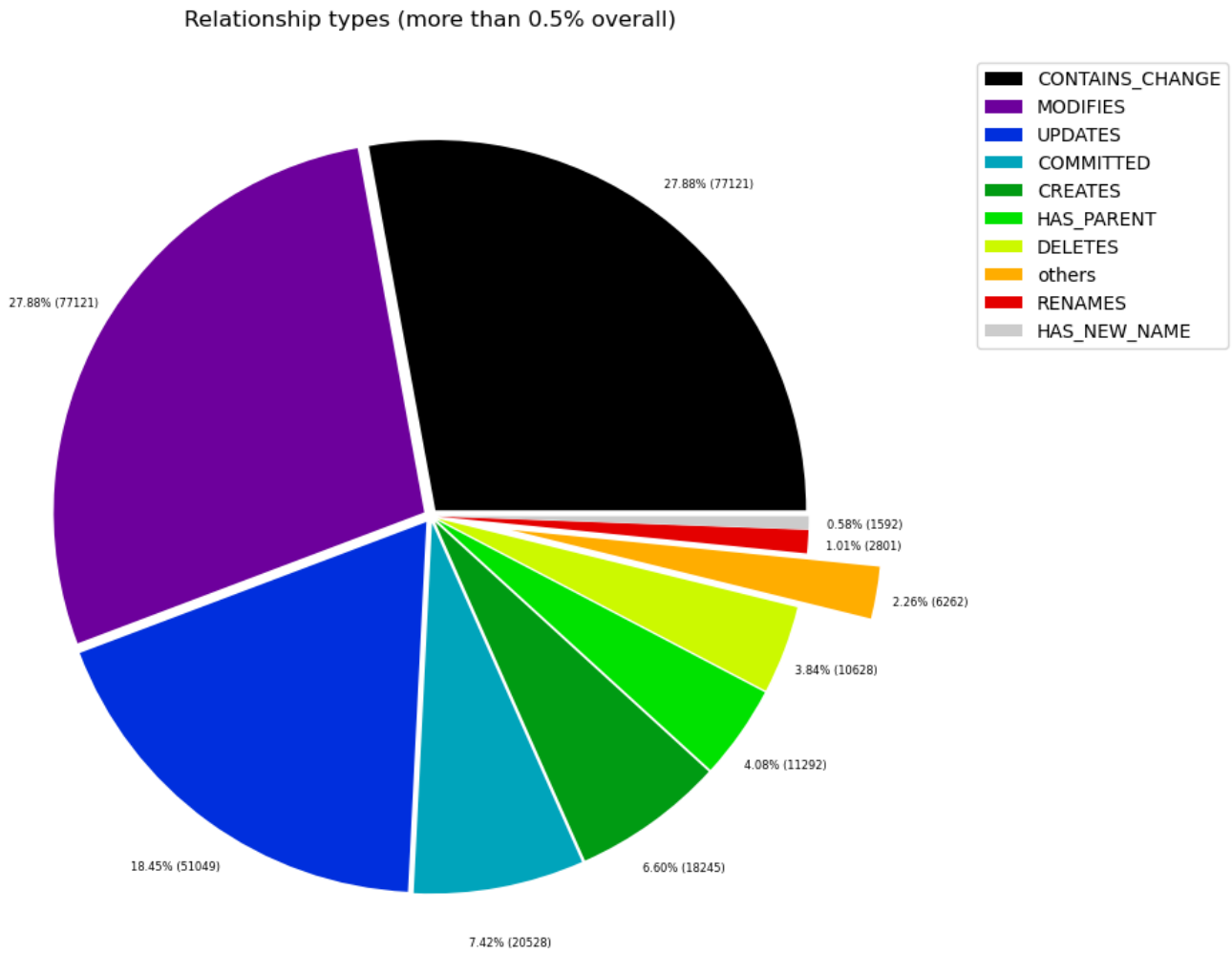


Table 2b - Lowest relationship count by type

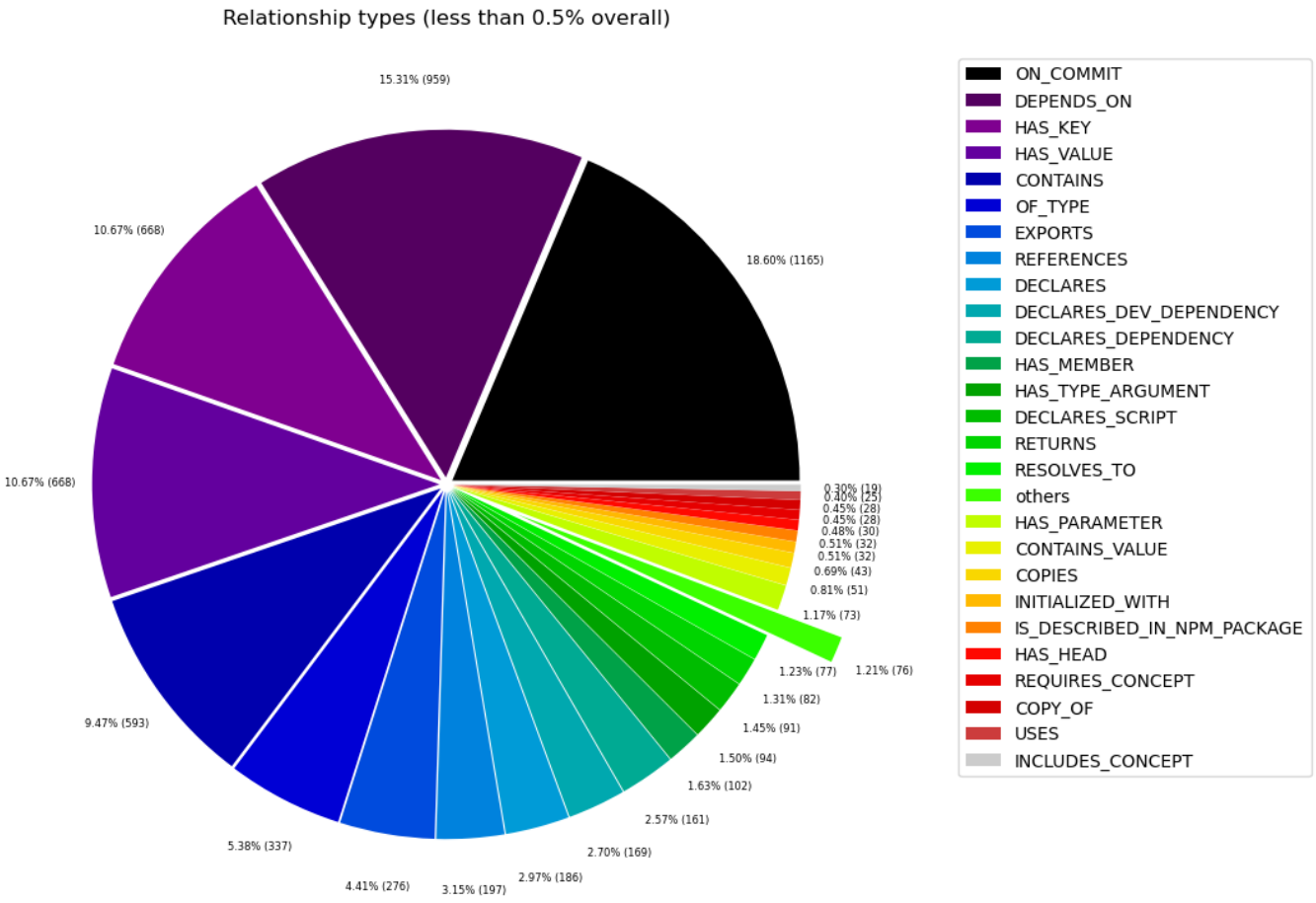
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	PROVIDED_BY_NPM_DEPENDENCY	1	0.000361
1	CONSTRAINED_BY	4	0.001446
2	REFERENCED_PROJECTS	5	0.001807
3	SIMILAR	6	0.002169
4	DECLARES_ENGINE	6	0.002169
5	EXTENDS	6	0.002169
6	HAS_ARGUMENT	6	0.002169
7	CALLS	6	0.002169
8	HAS_NPM_PACKAGE	6	0.002169
9	HAS_ROOT	6	0.002169
10	MEMBER	6	0.002169
11	PARENT	6	0.002169
12	HAS_CONFIG	6	0.002169
13	CONTAINS_PROJECT	6	0.002169
14	INCLUDES_CONCEPT	19	0.006868
15	USES	25	0.009037
16	REQUIRES_CONCEPT	28	0.010121
17	COPY_OF	28	0.010121
18	HAS_HEAD	30	0.010844
19	IS_DESCRIBED_IN_NPM_PACKAGE	32	0.011567
20	INITIALIZED_WITH	32	0.011567
21	COPIES	43	0.015544
22	CONTAINS_VALUE	51	0.018436
23	HAS_PARAMETER	73	0.026388
24	RESOLVES_TO	77	0.027834
25	RETURNS	82	0.029642
26	DECLARES_SCRIPT	91	0.032895
27	HAS_TYPE_ARGUMENT	94	0.033979
28	HAS_MEMBER	102	0.036871
29	DECLARES_DEPENDENCY	161	0.058199

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]	77121	10264	
1	[Git, Change]	MODIFIES	[File, Git]	77121	77121	
2	[Git, Change]	UPDATES	[File, Git]	51049	77121	
3	[Git, Change]	CREATES	[File, Git]	18245	77121	
4	[Git, Commit]	HAS_PARENT	[Git, Commit]	11292	10264	
5	[Git, Change]	DELETES	[File, Git]	10628	77121	
6	[Author, Git, Person]	COMMITTED	[Git, Commit]	10264	1191	
7	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10264	371	
8	[Git, Change]	RENAMES	[File, Git]	2801	77121	
9	[File, Git]	HAS_NEW_NAME	[File, Git]	1592	5237	
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	5237	

Graph Density

```
total_number_of_nodes (vertices): 98930
total_number_of_relationships (edges): 276639
-> total directed graph density: 2.8265832289740434e-05
-> total directed graph density in percent: 0.0028265832289740434
```