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]	76027	77.919669
1	[Git, Commit]	10137	10.389357
2	[File, Git]	5182	5.311004
3	[Author, Git, Person]	1186	1.215525
4	[Git, Tag]	1111	1.138658
5	[Json, Key]	668	0.684630
6	[Json, Value, Scalar]	603	0.618011
7	[Committer, Git, Person]	371	0.380236
8	[NPM, Dependency]	330	0.338215
9	[Type, TS, Primitive, ExternalType]	285	0.292095
10	[Type, TS, Declared, ExternalType]	272	0.278771
11	[TS, ExternalDeclaration]	215	0.220352
12	[Type, TS, Literal, ExternalType]	136	0.139386
13	[Json, Value, Object]	133	0.136311
14	[Type, TS, Union, ExternalType]	117	0.119913
15	[Type, TS, ObjectMember, ExternalType]	98	0.100440
16	[NPM, Script]	91	0.093265
17	[TS, Property]	65	0.066618
18	[TS, Function]	47	0.048170
19	[Type, Object, TS, ExternalType]	38	0.038946
20	[Type, TS, FunctionParameter, ExternalType]	37	0.037921
21	[File, Directory]	34	0.034846
22	[TS, Parameter]	33	0.033822
23	[Type, TS, Function, ExternalType]	32	0.032797
24	[Package, File, Json, NPM]	29	0.029722
25	[Git, Branch]	28	0.028697
26	[TS, ExternalModule]	25	0.025622
27	[TS, Variable]	24	0.024597
28	[Value, TS, Literal]	20	0.020498
29	[jQAssistant, Rule, Concept]	19	0.019473

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>
```

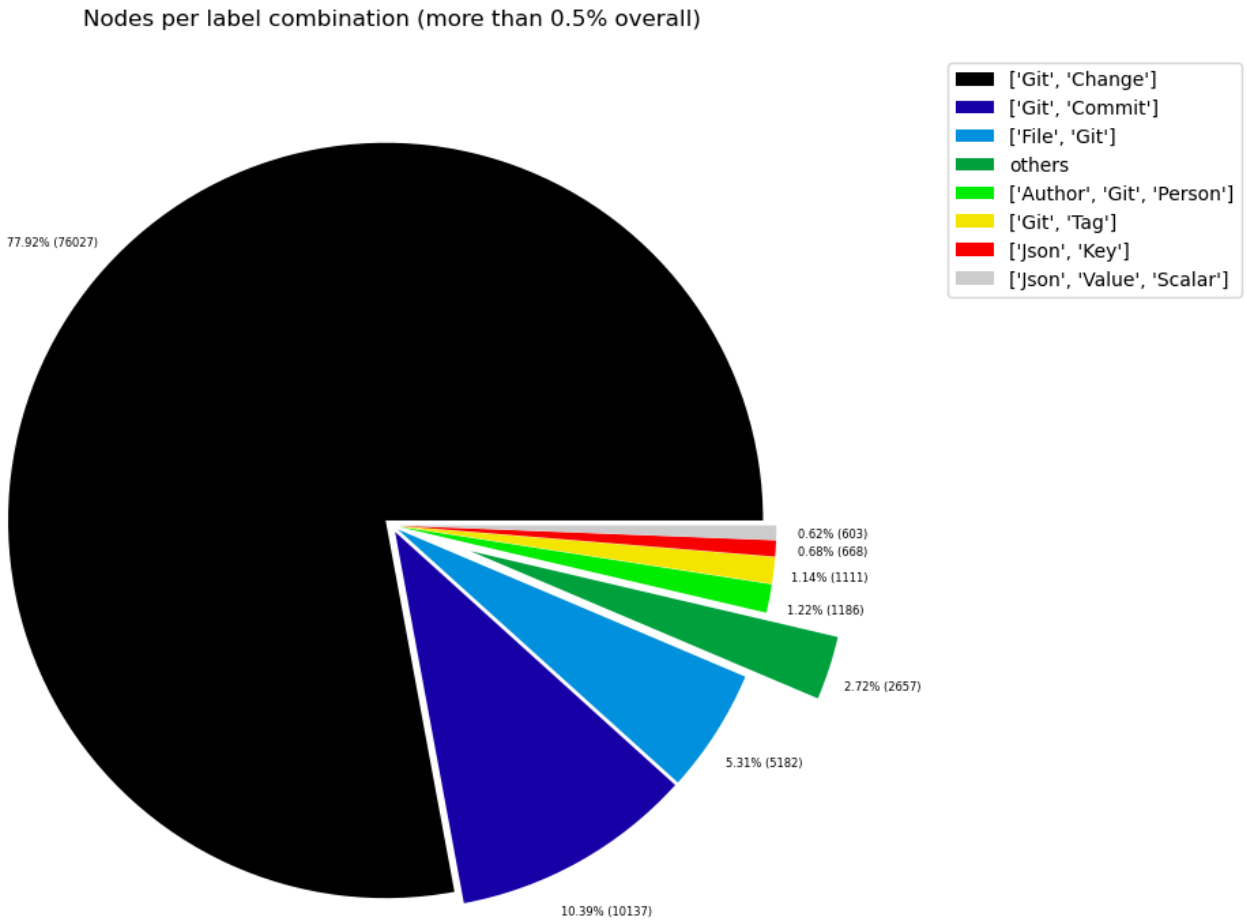


Table 1b - Lowest node count by label combination

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	0.001025
1	[File, TS, Scan]	1	0.001025
2	[TS, Method]	1	0.001025
3	[Value, TS, ObjectMember]	1	0.001025
4	[TS, Constructor]	1	0.001025
5	[TS, Class]	1	0.001025
6	[TS, Enum]	2	0.002050
7	[Value, Object, TS]	3	0.003075
8	[Type, TS, Tuple, ExternalType]	3	0.003075
9	[Value, TS, Function]	4	0.004100
10	[TS, TypeParameter]	4	0.004100
11	[Value, TS, Complex]	5	0.005124
12	[NPM, Engine]	6	0.006149
13	[Project, TS]	6	0.006149
14	[File, Local]	6	0.006149
15	[Value, TS, Call]	6	0.006149
16	[Value, TS, Member]	6	0.006149
17	[File, TS, Local, Module]	6	0.006149
18	[Type, TS, TypeParameterReference, ExternalType]	6	0.006149
19	[TS, EnumMember]	8	0.008199
20	[Type, TS, NotIdentified, ExternalType]	11	0.011274
21	[Json, Value, Array]	12	0.012299
22	[Value, TS, Declared]	13	0.013324
23	[TS, TypeAlias]	14	0.014349
24	[File, Directory, Local]	16	0.016398
25	[Type, TS, Intersection, ExternalType]	17	0.017423
26	[TS, Interface]	18	0.018448
27	[jQAssistant, Rule, Concept]	19	0.019473
28	[Value, TS, Literal]	20	0.020498
29	[TS, Variable]	24	0.024597

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>
```

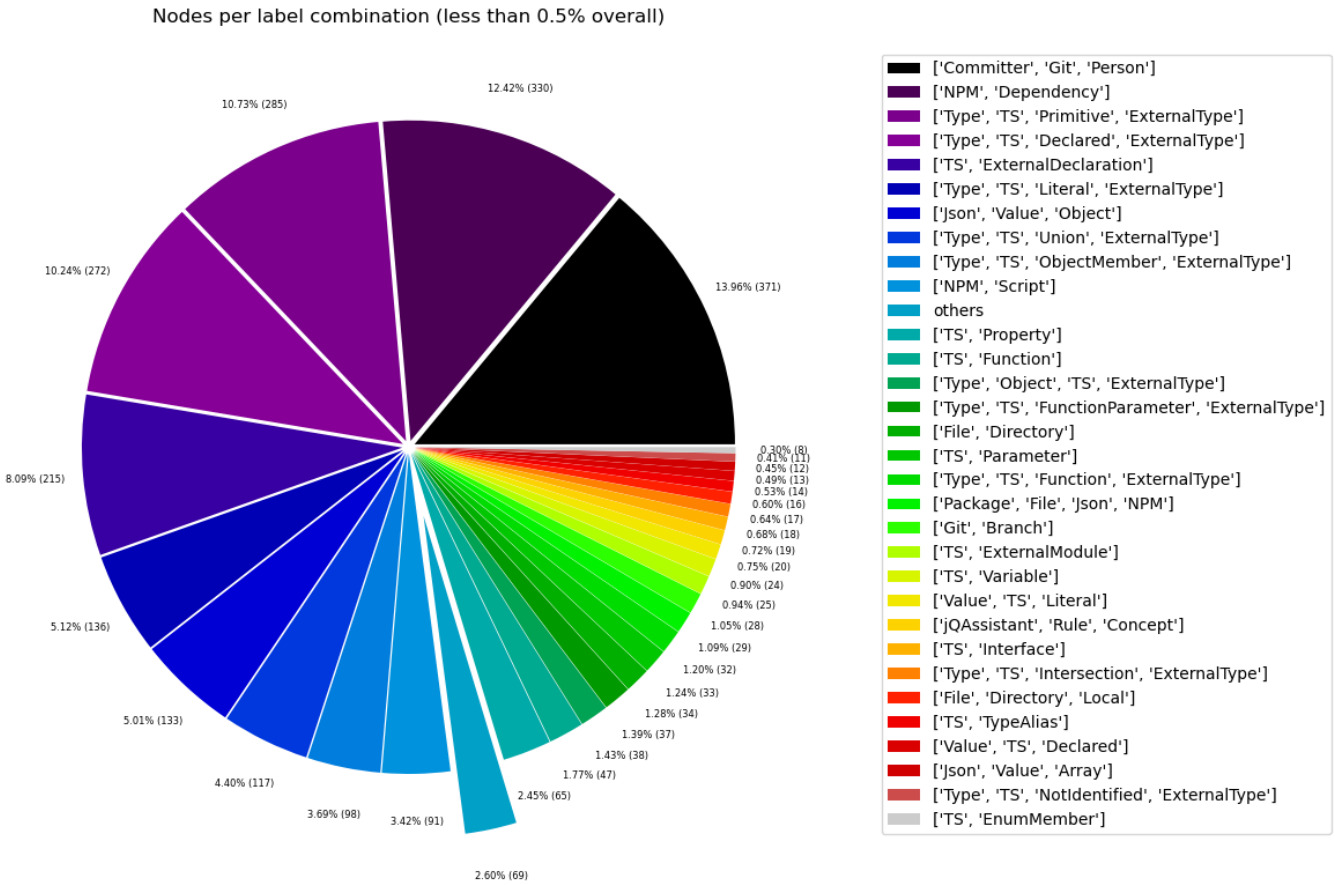


Table 1c - Highest node count by single label

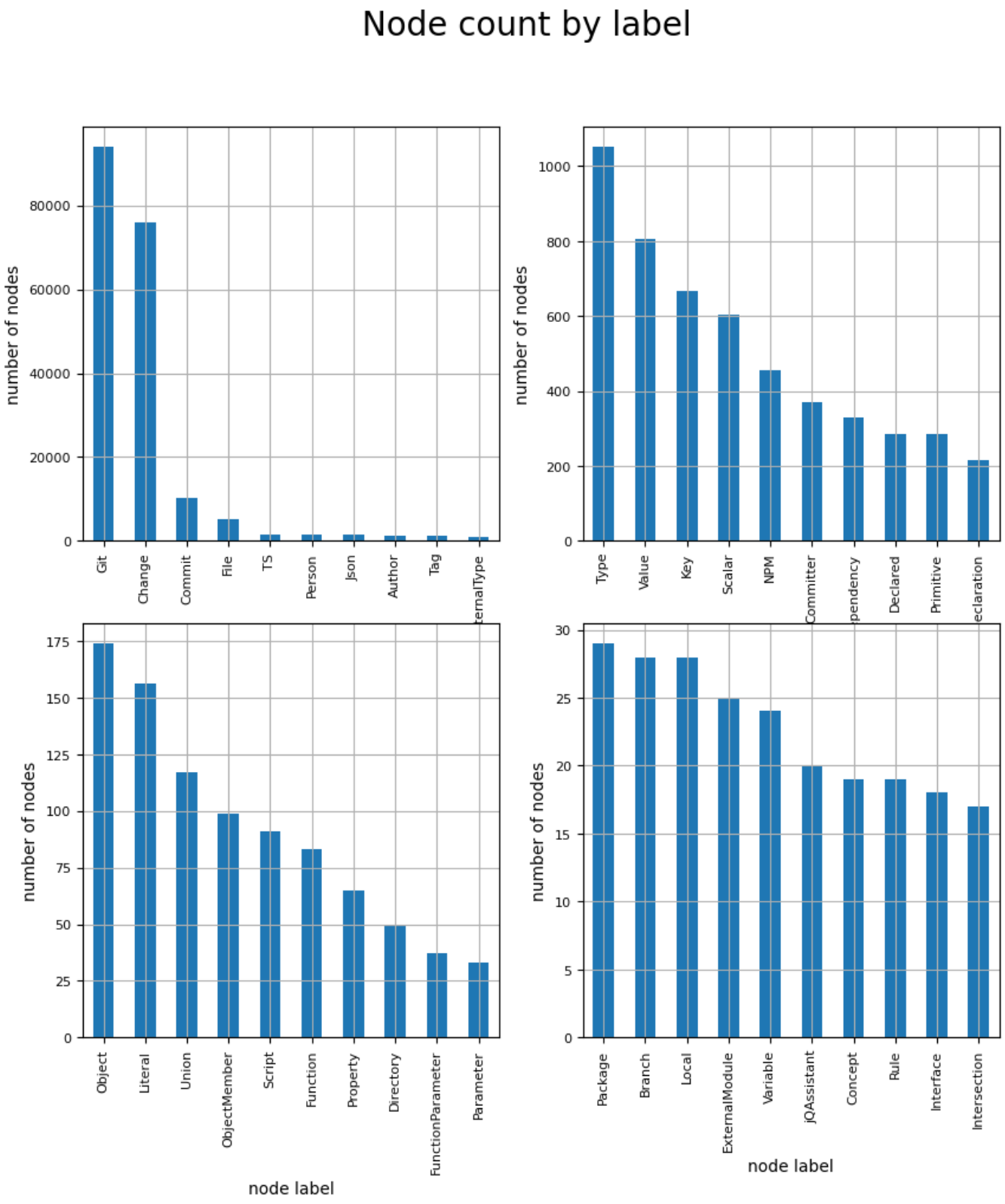
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	94042	96.383147
1	Change	76027	77.919669
2	Commit	10137	10.389357
3	File	5274	5.405295
4	TS	1581	1.620359
5	Person	1557	1.595761
6	Json	1445	1.480973
7	Author	1186	1.215525
8	Tag	1111	1.138658
9	ExternalType	1052	1.078189
10	Type	1052	1.078189
11	Value	806	0.826065
12	Key	668	0.684630
13	Scalar	603	0.618011
14	NPM	456	0.467352
15	Committer	371	0.380236
16	Dependency	330	0.338215
17	Declared	285	0.292095
18	Primitive	285	0.292095
19	ExternalDeclaration	215	0.220352
20	Object	174	0.178332
21	Literal	156	0.159884
22	Union	117	0.119913
23	ObjectMember	99	0.101465
24	Script	91	0.093265
25	Function	83	0.085066
26	Property	65	0.066618
27	Directory	50	0.051245
28	FunctionParameter	37	0.037921
29	Parameter	33	0.033822
30	Package	29	0.029722
31	Branch	28	0.028697
32	Local	28	0.028697
33	ExternalModule	25	0.025622
34	Variable	24	0.024597
35	jQAssistant	20	0.020498
36	Concept	19	0.019473
37	Rule	19	0.019473
38	Interface	18	0.018448
39	Intersection	17	0.017423

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: 272810
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	76027	27.868113
1	MODIFIES	76027	27.868113
2	UPDATES	50116	18.370294
3	COMMITTED	20274	7.431546
4	CREATES	18156	6.655181
5	HAS_PARENT	11162	4.091492
6	DELETES	10536	3.862029
7	RENAMES	2781	1.019391
8	HAS_NEW_NAME	1572	0.576225
9	ON_COMMIT	1111	0.407243
10	DEPENDS_ON	962	0.352626
11	HAS_KEY	668	0.244859
12	HAS_VALUE	668	0.244859
13	CONTAINS	589	0.215901
14	OF_TYPE	329	0.120597
15	EXPORTS	275	0.100803
16	REFERENCES	196	0.071845
17	DECLARES	185	0.067813
18	DECLARES_DEV_DEPENDENCY	169	0.061948
19	DECLARES_DEPENDENCY	161	0.059015
20	HAS_MEMBER	99	0.036289
21	HAS_TYPE_ARGUMENT	92	0.033723
22	DECLARES_SCRIPT	91	0.033357
23	RESOLVES_TO	80	0.029324
24	RETURNS	80	0.029324
25	HAS_PARAMETER	70	0.025659
26	CONTAINS_VALUE	51	0.018694
27	COPIES	43	0.015762
28	INITIALIZED_WITH	32	0.011730
29	COPY_OF	28	0.010264

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>
```

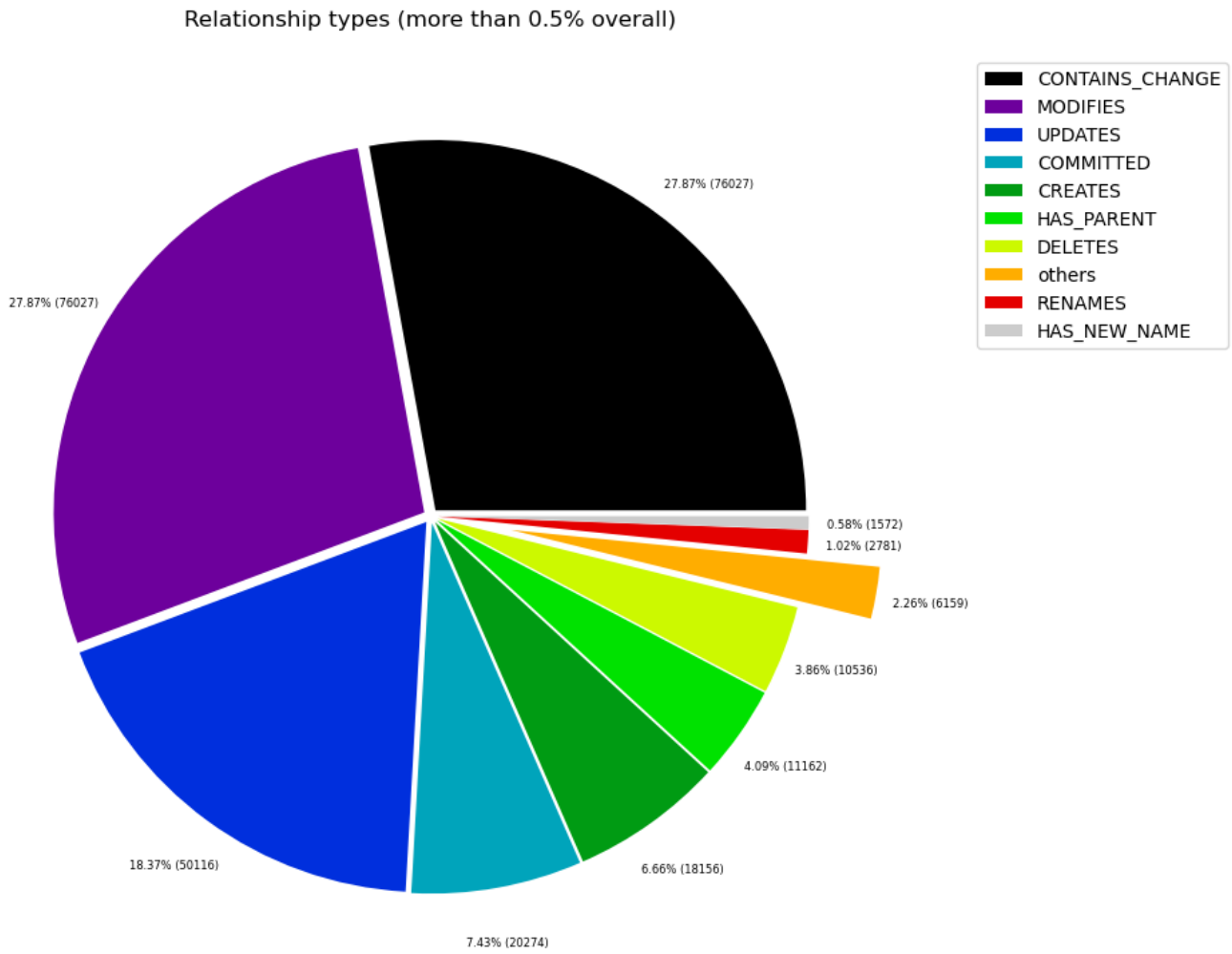


Table 2b - Lowest relationship count by type

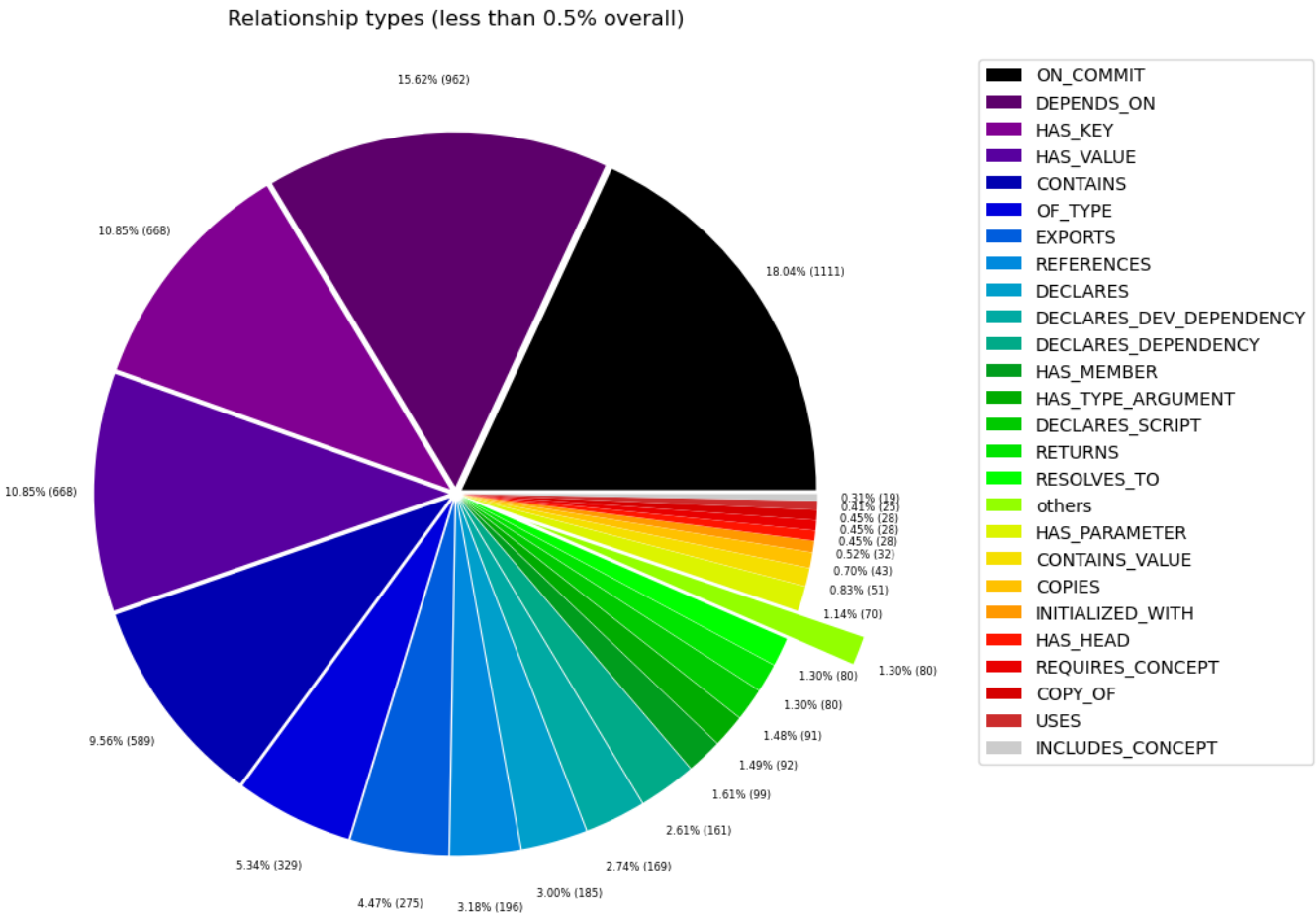
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONSTRAINED_BY	4	0.001466
1	REFERENCED_PROJECTS	5	0.001833
2	MEMBER	6	0.002199
3	HAS_ROOT	6	0.002199
4	HAS_NPM_PACKAGE	6	0.002199
5	HAS_CONFIG	6	0.002199
6	HAS_ARGUMENT	6	0.002199
7	DECLARES_ENGINE	6	0.002199
8	CONTAINS_PROJECT	6	0.002199
9	CALLS	6	0.002199
10	PARENT	6	0.002199
11	EXTENDS	7	0.002566
12	SIMILAR	10	0.003666
13	INCLUDES_CONCEPT	19	0.006965
14	USES	25	0.009164
15	REQUIRES_CONCEPT	28	0.010264
16	HAS_HEAD	28	0.010264
17	COPY_OF	28	0.010264
18	INITIALIZED_WITH	32	0.011730
19	COPIES	43	0.015762
20	CONTAINS_VALUE	51	0.018694
21	HAS_PARAMETER	70	0.025659
22	RETURNS	80	0.029324
23	RESOLVES_TO	80	0.029324
24	DECLARES_SCRIPT	91	0.033357
25	HAS_TYPE_ARGUMENT	92	0.033723
26	HAS_MEMBER	99	0.036289
27	DECLARES_DEPENDENCY	161	0.059015
28	DECLARES_DEV_DEPENDENCY	169	0.061948
29	DECLARES	185	0.067813

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]	76027	10137	
1	[Git, Change]	MODIFIES	[File, Git]	76027	76027	
2	[Git, Change]	UPDATES	[File, Git]	50116	76027	
3	[Git, Change]	CREATES	[File, Git]	18156	76027	
4	[Git, Commit]	HAS_PARENT	[Git, Commit]	11162	10137	
5	[Git, Change]	DELETES	[File, Git]	10536	76027	
6	[Author, Git, Person]	COMMITTED	[Git, Commit]	10137	1186	
7	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10137	371	
8	[Git, Change]	RENAMES	[File, Git]	2781	76027	
9	[File, Git]	HAS_NEW_NAME	[File, Git]	1572	5182	
10	[Git, Tag]	ON_COMMIT	[Git, Commit]	1111	1111	
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	3	
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, ExternalType]	CONTAINS	[Type, TS, Primitive, ExternalType]	144	117	
19	[Type, TS, Declared, ExternalType]	REFERENCES	[TS, ExternalDeclaration]	141	272	
20	[TS, Function]	DEPENDS_ON	[TS, ExternalModule]	131	47	
21	[Type, TS, Union, ExternalType]	CONTAINS	[Type, TS, Literal, ExternalType]	119	117	
22	[Json, Key]	HAS_VALUE	[Json, Value, Object]	104	668	
23	[Type, Object, TS, ExternalType]	HAS_MEMBER	[Type, TS, ObjectMember, ExternalType]	98	38	
24	[Package, File, Json, NPM]	DECLARES_SCRIPT	[NPM, Script]	91	29	
25	[Type, TS, Union, ExternalType]	CONTAINS	[Type, TS, Declared, ExternalType]	69	117	
26	[File, Directory]	CONTAINS	[File, Directory]	63	34	
27	[TS, Interface]	DECLARES	[TS, Property]	61	18	
28	[File, Directory]	CONTAINS	[Package, File, Json, NPM]	58	34	
29	[File, Git]	RESOLVES_TO	[Package, File, Json, NPM]	57	5182	

Graph Density

```
total_number_of_nodes (vertices): 97571
total_number_of_relationships (edges): 272810
-> total directed graph density: 2.8656505170003385e-05
-> total directed graph density in percent: 0.0028656505170003383
```