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`.

```
Total number of nodes: 93533
```

	nodeLabels	nodesWithThatLabels	nodesWithThatLabelsPercent
0	[Git, Change]	72341	77.342756
1	[Git, Commit]	9928	10.614436
2	[File, Git]	5089	5.440860
3	[Author, Git, Person]	1183	1.264794
4	[Git, Tag]	1068	1.141843
5	[Json, Key]	668	0.714186
6	[Json, Value, Scalar]	603	0.644692
7	[Committer, Git, Person]	371	0.396651
8	[NPM, Dependency]	330	0.352817
9	[Type, TS, Primitive, ExternalType]	285	0.304705
10	[Type, TS, Declared, ExternalType]	272	0.290806
11	[TS, ExternalDeclaration]	215	0.229865
12	[Type, TS, Literal, ExternalType]	136	0.145403
13	[Json, Value, Object]	133	0.142196
14	[Type, TS, Union, ExternalType]	117	0.125090
15	[Type, TS, ObjectMember, ExternalType]	98	0.104776
16	[NPM, Script]	91	0.097292
17	[TS, Property]	65	0.069494
18	[TS, Function]	47	0.050250
19	[Type, Object, TS, ExternalType]	38	0.040627
20	[Type, TS, FunctionParameter, ExternalType]	37	0.039558
21	[File, Directory]	34	0.036351
22	[TS, Parameter]	33	0.035282
23	[Type, TS, Function, ExternalType]	32	0.034213
24	[Package, File, Json, NPM]	29	0.031005
25	[TS, ExternalModule]	25	0.026729
26	[TS, Variable]	24	0.025659
27	[Git, Branch]	24	0.025659
28	[Value, TS, Literal]	20	0.021383
29	[jQAssistant, Rule, Concept]	19	0.020314

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>
```

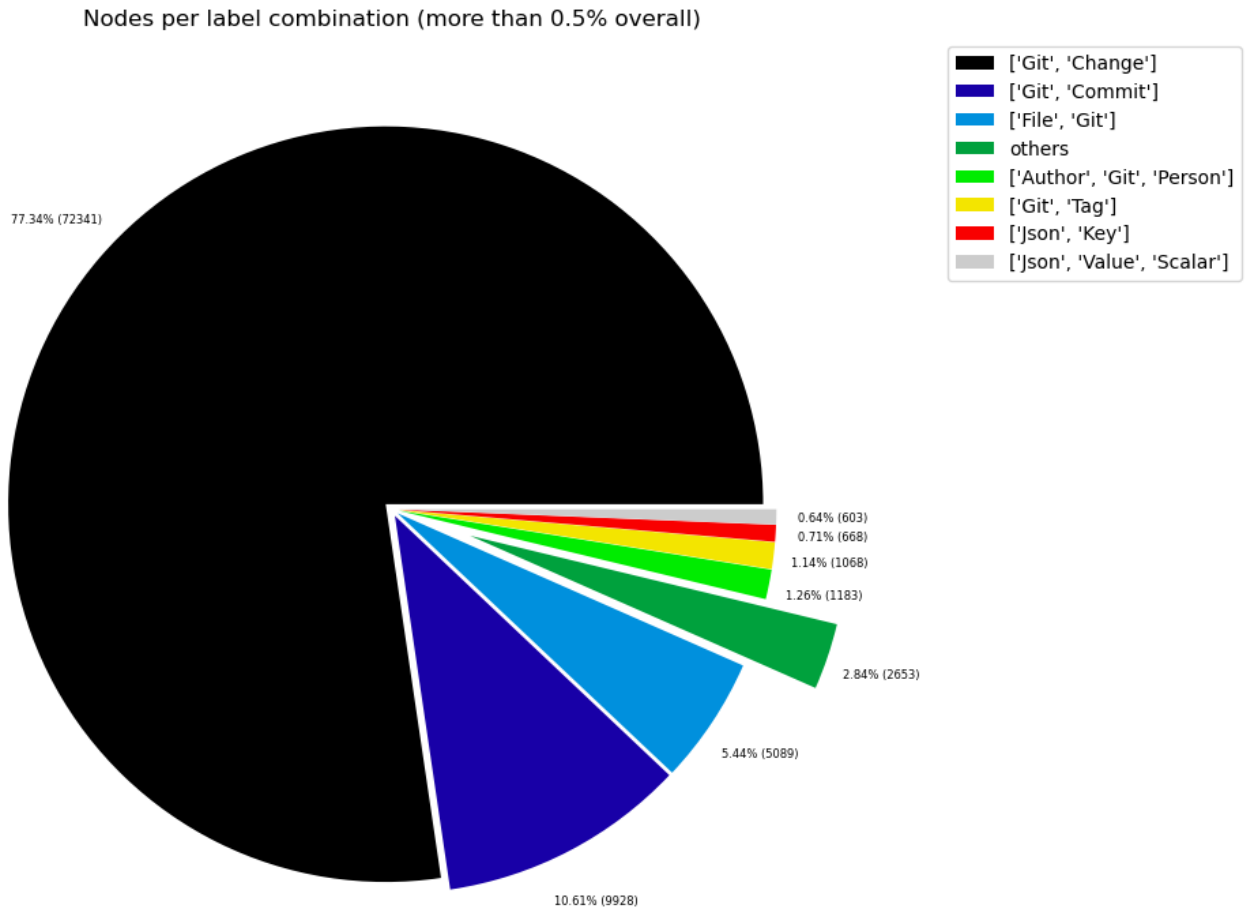


Table 1b - Lowest node count by label combination

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	0.001069
1	[File, TS, Scan]	1	0.001069
2	[TS, Method]	1	0.001069
3	[Value, TS, ObjectMember]	1	0.001069
4	[TS, Constructor]	1	0.001069
5	[TS, Class]	1	0.001069
6	[TS, Enum]	2	0.002138
7	[Value, Object, TS]	3	0.003207
8	[Type, TS, Tuple, ExternalType]	3	0.003207
9	[Value, TS, Function]	4	0.004277
10	[TS, TypeParameter]	4	0.004277
11	[Value, TS, Complex]	5	0.005346
12	[NPM, Engine]	6	0.006415
13	[Project, TS]	6	0.006415
14	[File, Local]	6	0.006415
15	[Value, TS, Call]	6	0.006415
16	[Value, TS, Member]	6	0.006415
17	[File, TS, Local, Module]	6	0.006415
18	[Type, TS, TypeParameterReference, ExternalType]	6	0.006415
19	[TS, EnumMember]	8	0.008553
20	[Type, TS, NotIdentified, ExternalType]	11	0.011761
21	[Json, Value, Array]	12	0.012830
22	[Value, TS, Declared]	13	0.013899
23	[TS, TypeAlias]	14	0.014968
24	[File, Directory, Local]	16	0.017106
25	[Type, TS, Intersection, ExternalType]	17	0.018175
26	[TS, Interface]	18	0.019245
27	[jQAssistant, Rule, Concept]	19	0.020314
28	[Value, TS, Literal]	20	0.021383
29	[TS, Variable]	24	0.025659

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>
```

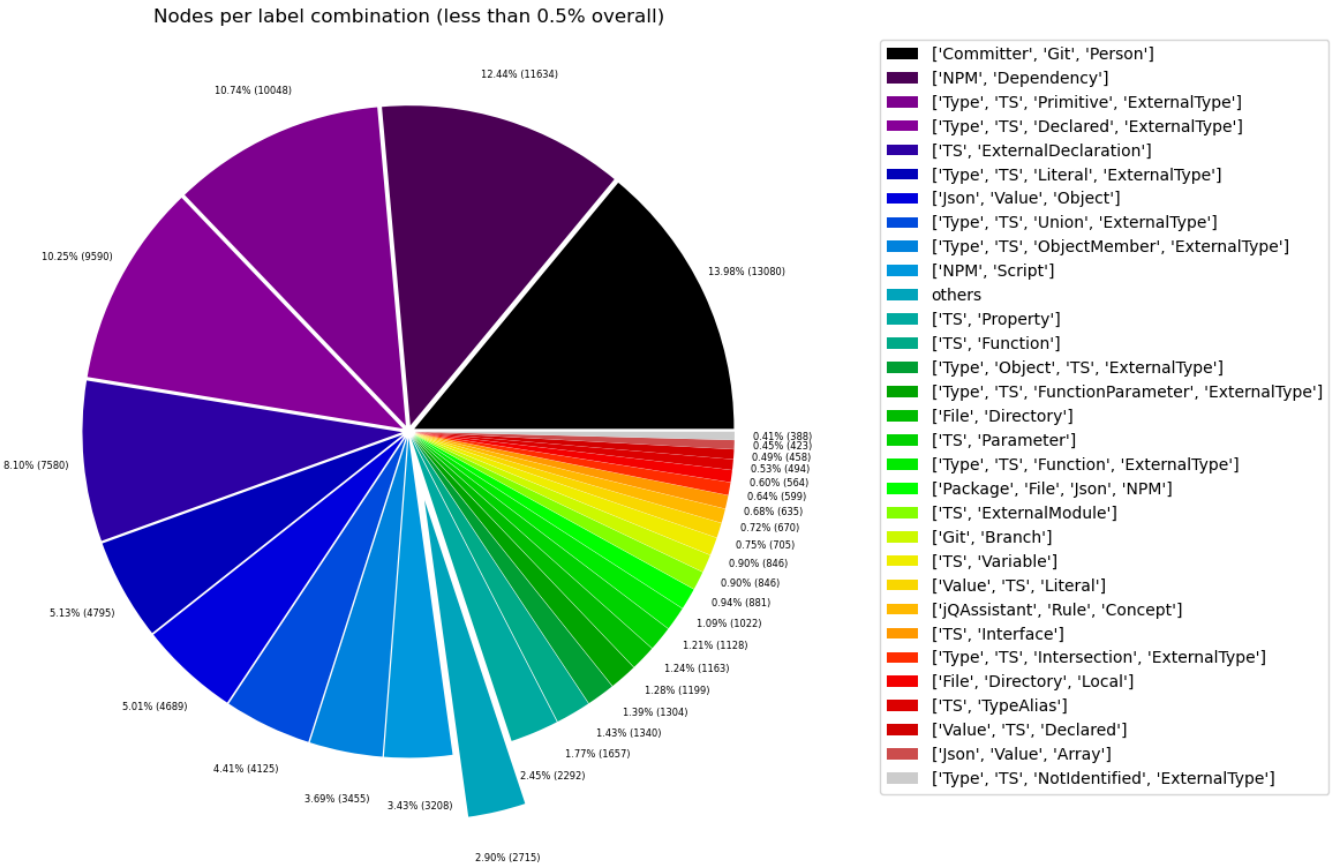


Table 1c - Highest node count by single label

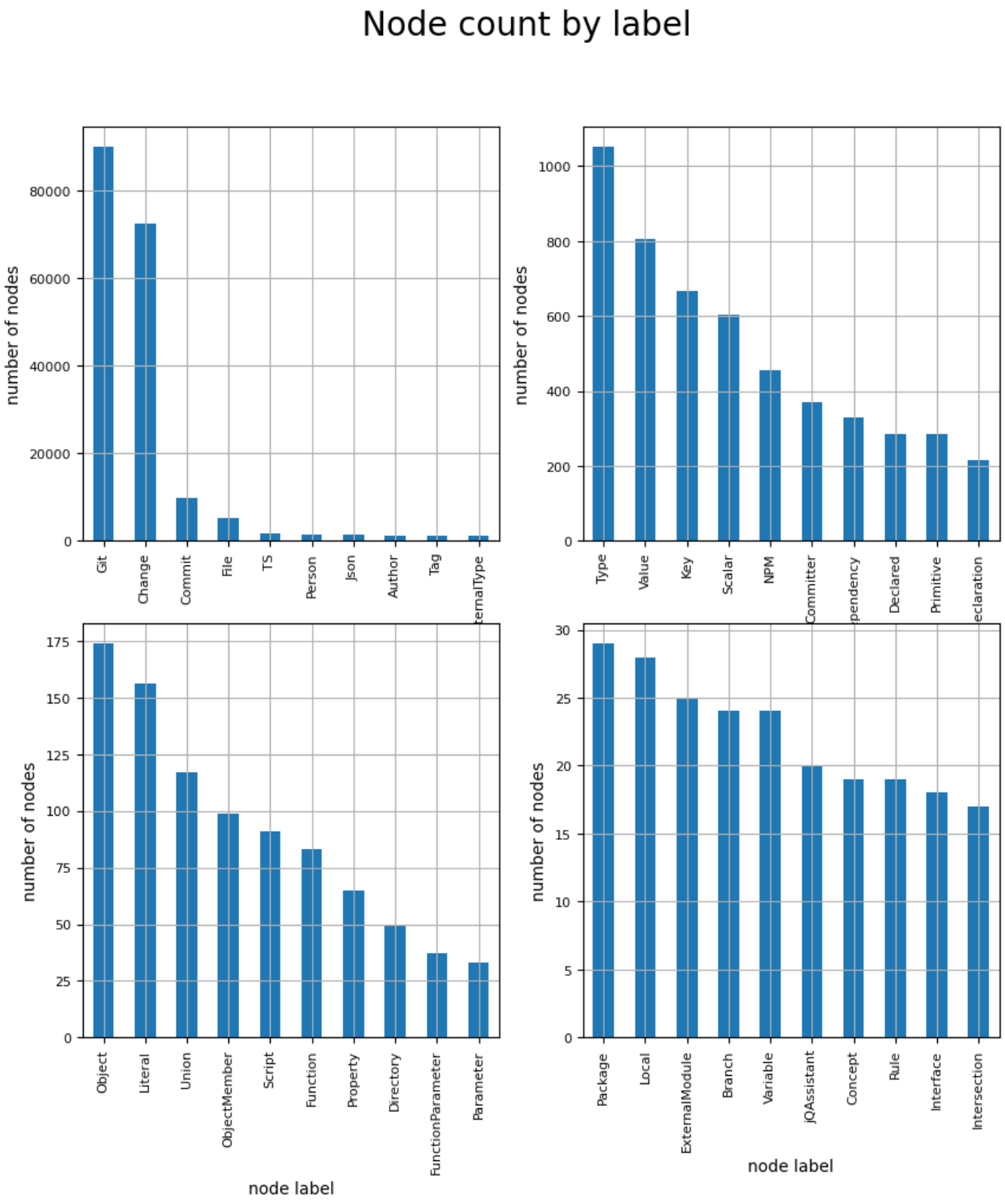
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	90004	96.227000
1	Change	72341	77.342756
2	Commit	9928	10.614436
3	File	5181	5.539221
4	TS	1581	1.690313
5	Person	1554	1.661446
6	Json	1445	1.544909
7	Author	1183	1.264794
8	Tag	1068	1.141843
9	ExternalType	1052	1.124737
10	Type	1052	1.124737
11	Value	806	0.861728
12	Key	668	0.714186
13	Scalar	603	0.644692
14	NPM	456	0.487528
15	Committer	371	0.396651
16	Dependency	330	0.352817
17	Declared	285	0.304705
18	Primitive	285	0.304705
19	ExternalDeclaration	215	0.229865
20	Object	174	0.186031
21	Literal	156	0.166786
22	Union	117	0.125090
23	ObjectMember	99	0.105845
24	Script	91	0.097292
25	Function	83	0.088739
26	Property	65	0.069494
27	Directory	50	0.053457
28	FunctionParameter	37	0.039558
29	Parameter	33	0.035282
30	Package	29	0.031005
31	Local	28	0.029936
32	ExternalModule	25	0.026729
33	Branch	24	0.025659
34	Variable	24	0.025659
35	jQAssistant	20	0.021383
36	Concept	19	0.020314
37	Rule	19	0.020314
38	Interface	18	0.019245
39	Intersection	17	0.018175

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: 260922
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	72341	27.725144
1	MODIFIES	72341	27.725144
2	UPDATES	48407	18.552288
3	COMMITTED	19856	7.609937
4	CREATES	16729	6.411495
5	HAS_PARENT	10934	4.190524
6	DELETES	9928	3.804969
7	RENAMES	2723	1.043607
8	HAS_NEW_NAME	1558	0.597113
9	ON_COMMIT	1068	0.409318
10	DEPENDS_ON	962	0.368693
11	HAS_KEY	668	0.256015
12	HAS_VALUE	668	0.256015
13	CONTAINS	589	0.225738
14	OF_TYPE	329	0.126091
15	EXPORTS	275	0.105395
16	REFERENCES	196	0.075118
17	DECLARES	185	0.070902
18	DECLARES_DEV_DEPENDENCY	169	0.064770
19	DECLARES_DEPENDENCY	161	0.061704
20	HAS_MEMBER	99	0.037942
21	HAS_TYPE_ARGUMENT	92	0.035260
22	DECLARES_SCRIPT	91	0.034876
23	RESOLVES_TO	80	0.030661
24	RETURNS	80	0.030661
25	HAS_PARAMETER	70	0.026828
26	CONTAINS_VALUE	51	0.019546
27	COPIES	36	0.013797
28	INITIALIZED_WITH	32	0.012264
29	COPY_OF	28	0.010731

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>
```

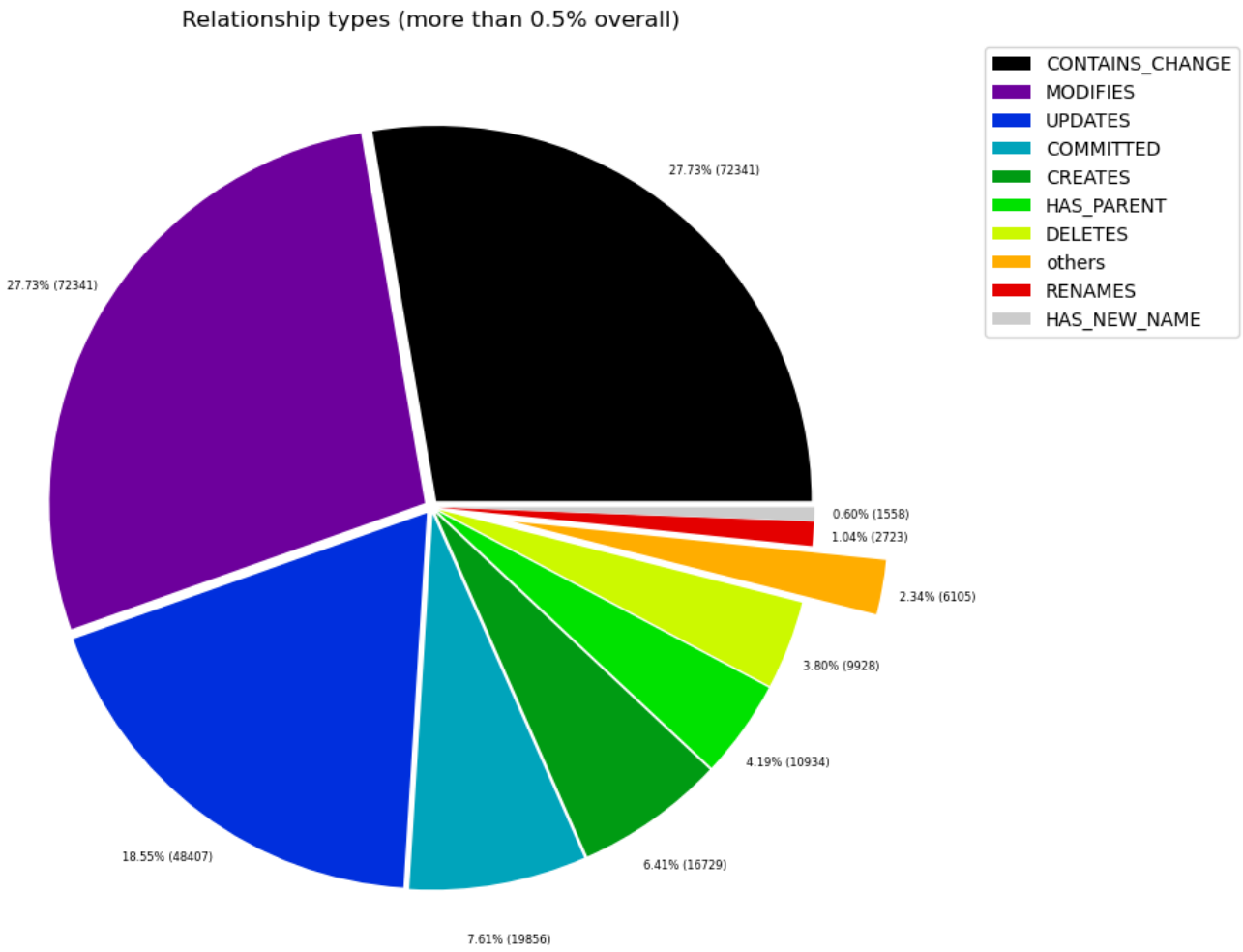


Table 2b - Lowest relationship count by type

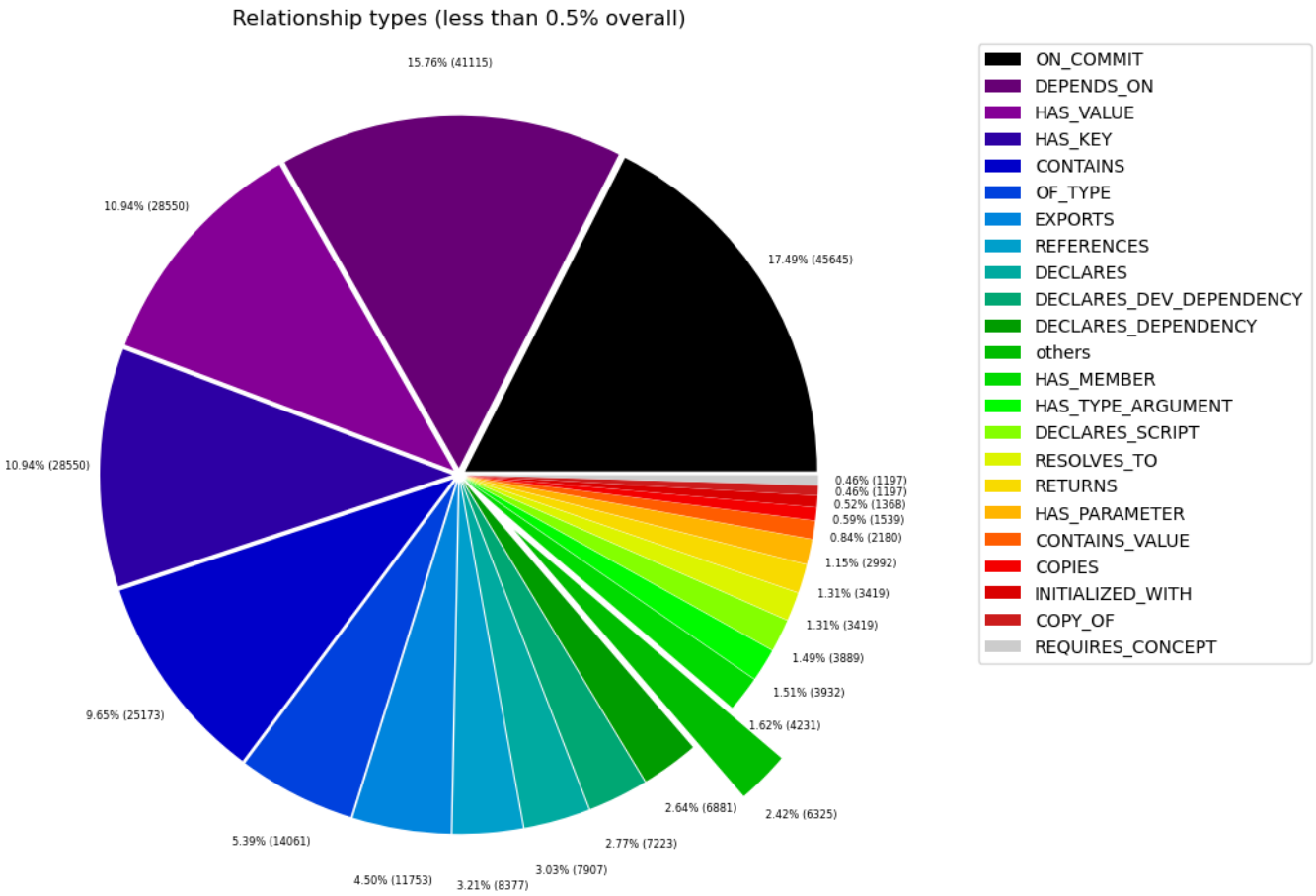
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONSTRAINED_BY	4	0.001533
1	REFERENCED_PROJECTS	5	0.001916
2	MEMBER	6	0.002300
3	HAS_ROOT	6	0.002300
4	HAS_NPM_PACKAGE	6	0.002300
5	HAS_CONFIG	6	0.002300
6	HAS_ARGUMENT	6	0.002300
7	DECLARES_ENGINE	6	0.002300
8	CONTAINS_PROJECT	6	0.002300
9	CALLS	6	0.002300
10	PARENT	6	0.002300
11	EXTENDS	7	0.002683
12	SIMILAR	10	0.003833
13	INCLUDES_CONCEPT	19	0.007282
14	HAS_HEAD	24	0.009198
15	USES	25	0.009581
16	REQUIRES_CONCEPT	28	0.010731
17	COPY_OF	28	0.010731
18	INITIALIZED_WITH	32	0.012264
19	COPIES	36	0.013797
20	CONTAINS_VALUE	51	0.019546
21	HAS_PARAMETER	70	0.026828
22	RETURNS	80	0.030661
23	RESOLVES_TO	80	0.030661
24	DECLARES_SCRIPT	91	0.034876
25	HAS_TYPE_ARGUMENT	92	0.035260
26	HAS_MEMBER	99	0.037942
27	DECLARES_DEPENDENCY	161	0.061704
28	DECLARES_DEV_DEPENDENCY	169	0.064770
29	DECLARES	185	0.070902

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]	72341	9928	
1	[Git, Change]	MODIFIES	[File, Git]	72341	72341	
2	[Git, Change]	UPDATES	[File, Git]	48407	72341	
3	[Git, Change]	CREATES	[File, Git]	16729	72341	
4	[Git, Commit]	HAS_PARENT	[Git, Commit]	10934	9928	
5	[Author, Git, Person]	COMMITTED	[Git, Commit]	9928	1183	
6	[Committer, Git, Person]	COMMITTED	[Git, Commit]	9928	371	
7	[Git, Change]	DELETES	[File, Git]	9928	72341	
8	[Git, Change]	RENAMES	[File, Git]	2723	72341	
9	[File, Git]	HAS_NEW_NAME	[File, Git]	1558	5089	
10	[Git, Tag]	ON_COMMIT	[Git, Commit]	1068	1068	
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	5089	

Graph Density

```
total_number_of_nodes (vertices): 93533
total_number_of_relationships (edges): 260922
-> total directed graph density: 2.98253544468272e-05
-> total directed graph density in percent: 0.00298253544468272
```