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]	83687	78.276526
1	[Git, Commit]	10893	10.188753
2	[File, Git]	5610	5.247306
3	[Git, Tag]	1530	1.431084
4	[Author, Git, Person]	1241	1.160768
5	[Json, Key]	668	0.624813
6	[Json, Value, Scalar]	603	0.564015
7	[Committer, Git, Person]	370	0.346079
8	[NPM, Dependency]	338	0.316148
9	[Type, TS, Primitive]	291	0.272186
10	[Type, TS, Declared]	276	0.258156
11	[TS, ExternalDeclaration]	215	0.201100
12	[Type, TS, Literal]	136	0.127207
13	[Json, Value, Object]	133	0.124401
14	[Type, TS, Union]	119	0.111306
15	[Type, TS, ObjectMember]	101	0.094470
16	[NPM, Script]	91	0.085117
17	[TS, Property]	65	0.060798
18	[TS, Function]	47	0.043961
19	[Type, TS, FunctionParameter]	40	0.037414
20	[Type, Object, TS]	39	0.036479
21	[Git, Branch]	35	0.032737
22	[File, Directory]	34	0.031802
23	[Type, TS, Function]	34	0.031802
24	[TS, Parameter]	33	0.030867
25	[Package, File, Json, NPM]	29	0.027125
26	[TS, Variable]	24	0.022448
27	[Value, TS, Literal]	20	0.018707
28	[jQAssistant, Rule, Concept]	19	0.017772
29	[Type, TS, Intersection]	17	0.015901

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>
```

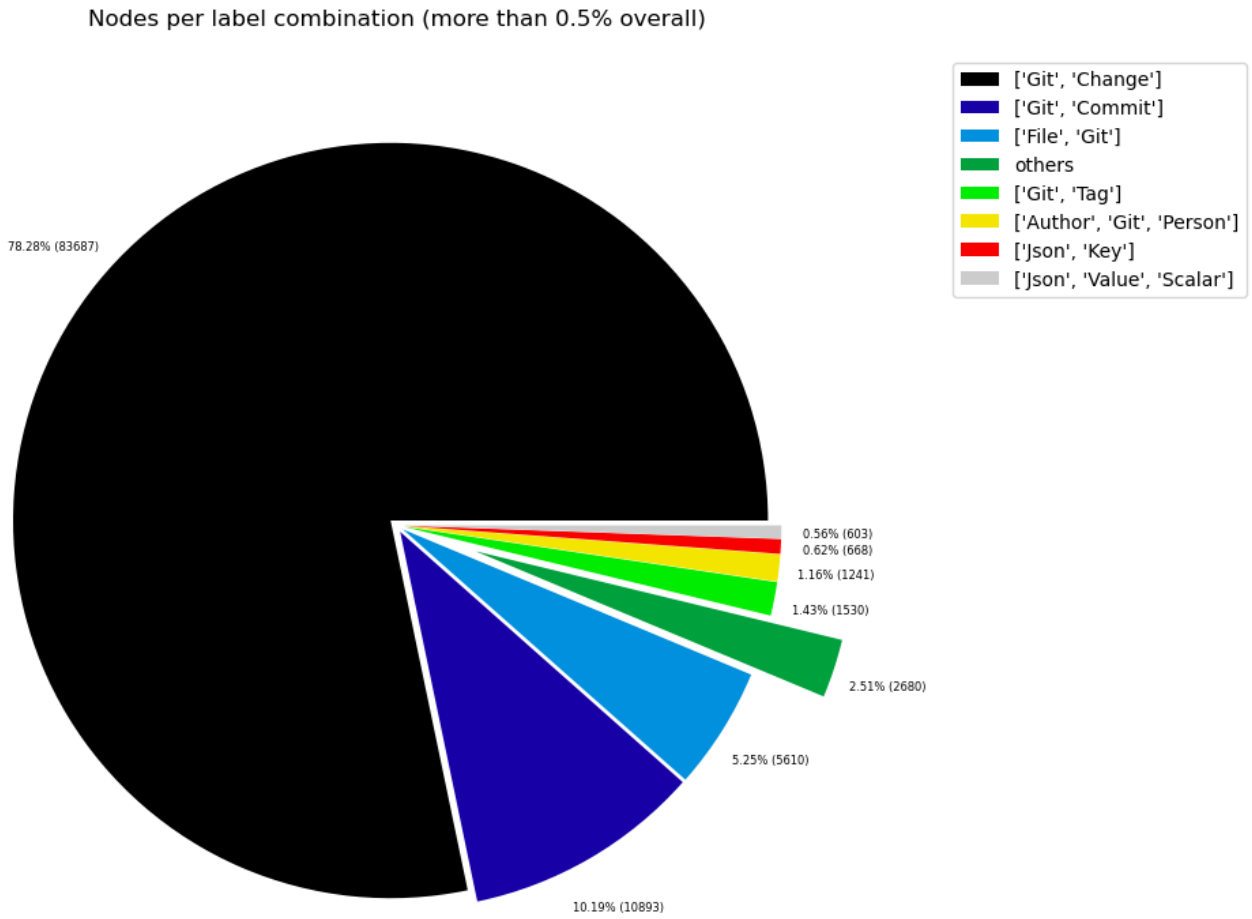


Table 1b - Lowest node count by label combination

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	0.000935
1	[File, TS, Scan]	1	0.000935
2	[TS, Method]	1	0.000935
3	[Repository, File, Git]	1	0.000935
4	[TS, Constructor]	1	0.000935
5	[Value, TS, ObjectMember]	1	0.000935
6	[TS, Class]	1	0.000935
7	[TS, Enum]	2	0.001871
8	[Value, Object, TS]	3	0.002806
9	[Type, TS, Tuple]	3	0.002806
10	[Value, TS, Function]	4	0.003741
11	[TS, TypeParameter]	4	0.003741
12	[Value, TS, Complex]	5	0.004677
13	[NPM, Engine]	6	0.005612
14	[Project, TS]	6	0.005612
15	[File, Local]	6	0.005612
16	[Value, TS, Call]	6	0.005612
17	[Value, TS, Member]	6	0.005612
18	[File, TS, Local, Module]	6	0.005612
19	[Type, TS, TypeParameterReference]	6	0.005612
20	[TS, EnumMember]	8	0.007483
21	[Type, TS, NotIdentified]	11	0.010289
22	[TS, ExternalModule]	11	0.010289
23	[Json, Value, Array]	12	0.011224
24	[Value, TS, Declared]	13	0.012160
25	[TS, TypeAlias]	16	0.014966
26	[File, Directory, Local]	16	0.014966
27	[Type, TS, Intersection]	17	0.015901
28	[TS, Interface]	17	0.015901
29	[jQAssistant, Rule, Concept]	19	0.017772

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>
```

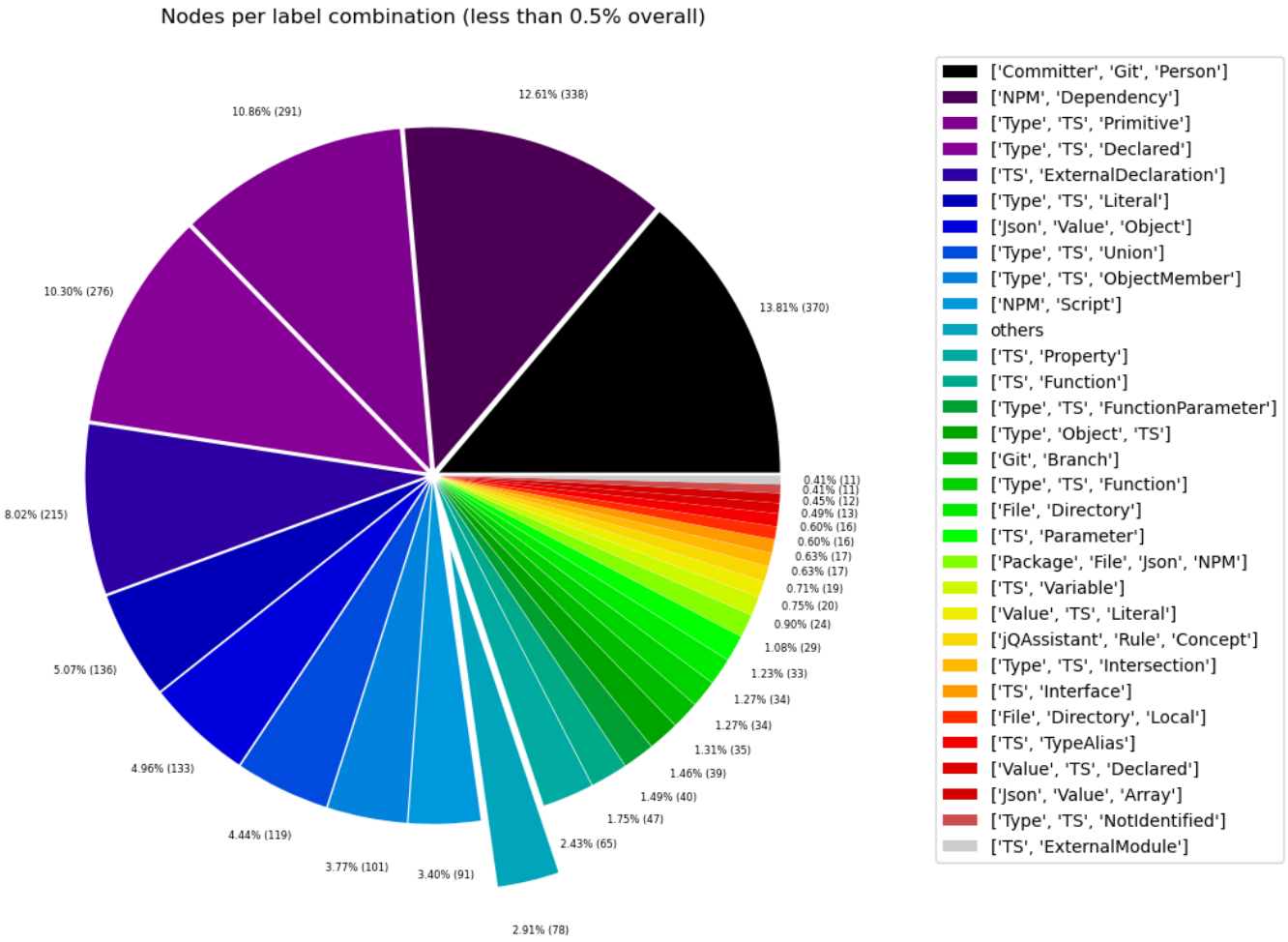


Table 1c - Highest node count by single label

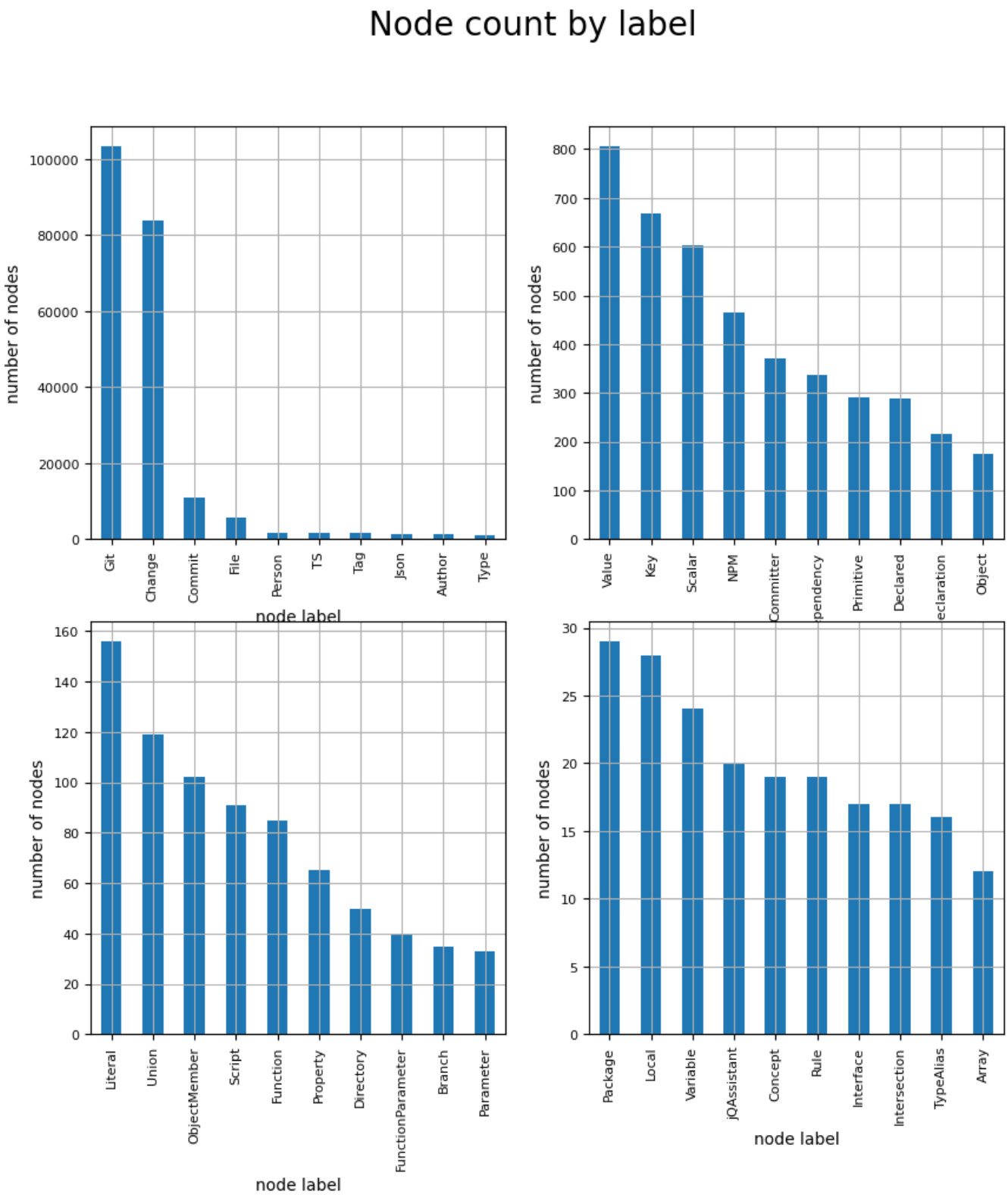
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	103367	96.684189
1	Change	83687	78.276526
2	Commit	10893	10.188753
3	File	5703	5.334294
4	Person	1611	1.506847
5	TS	1589	1.486269
6	Tag	1530	1.431084
7	Json	1445	1.351579
8	Author	1241	1.160768
9	Type	1073	1.003629
10	Value	806	0.753891
11	Key	668	0.624813
12	Scalar	603	0.564015
13	NPM	464	0.434002
14	Committer	370	0.346079
15	Dependency	338	0.316148
16	Primitive	291	0.272186
17	Declared	289	0.270316
18	ExternalDeclaration	215	0.201100
19	Object	175	0.163686
20	Literal	156	0.145914
21	Union	119	0.111306
22	ObjectMember	102	0.095406
23	Script	91	0.085117
24	Function	85	0.079505
25	Property	65	0.060798
26	Directory	50	0.046767
27	FunctionParameter	40	0.037414
28	Branch	35	0.032737
29	Parameter	33	0.030867
30	Package	29	0.027125
31	Local	28	0.026190
32	Variable	24	0.022448
33	jQAssistant	20	0.018707
34	Concept	19	0.017772
35	Rule	19	0.017772
36	Interface	17	0.015901
37	Intersection	17	0.015901
38	TypeAlias	16	0.014966
39	Array	12	0.011224

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: 319365
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	83687	26.204186
1	MODIFIES	83687	26.204186
2	UPDATES	55501	17.378548
3	COMMITTED	21786	6.821662
4	CREATES	19609	6.139997
5	HAS_PARENT	11956	3.743679
6	DELETES	11844	3.708609
7	HAS_COMMIT	10893	3.410831
8	HAS_FILE	5610	1.756611
9	RENAMES	3267	1.022967
10	HAS_NEW_NAME	1751	0.548275
11	HAS_TAG	1530	0.479076
12	ON_COMMIT	1530	0.479076
13	HAS_AUTHOR	1241	0.388584
14	DEPENDS_ON	887	0.277739
15	HAS_KEY	668	0.209165
16	HAS_VALUE	668	0.209165
17	CONTAINS	594	0.185994
18	HAS_COMMITTER	370	0.115855
19	OF_TYPE	337	0.105522
20	EXPORTS	309	0.096754
21	REFERENCES	197	0.061685
22	DECLARES	186	0.058241
23	DECLARES_DEV_DEPENDENCY	169	0.052918
24	DECLARES_DEPENDENCY	161	0.050413
25	HAS_MEMBER	102	0.031938
26	HAS_TYPE_ARGUMENT	94	0.029433
27	DECLARES_SCRIPT	91	0.028494
28	RETURNS	82	0.025676
29	COPIES	79	0.024737

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>
```

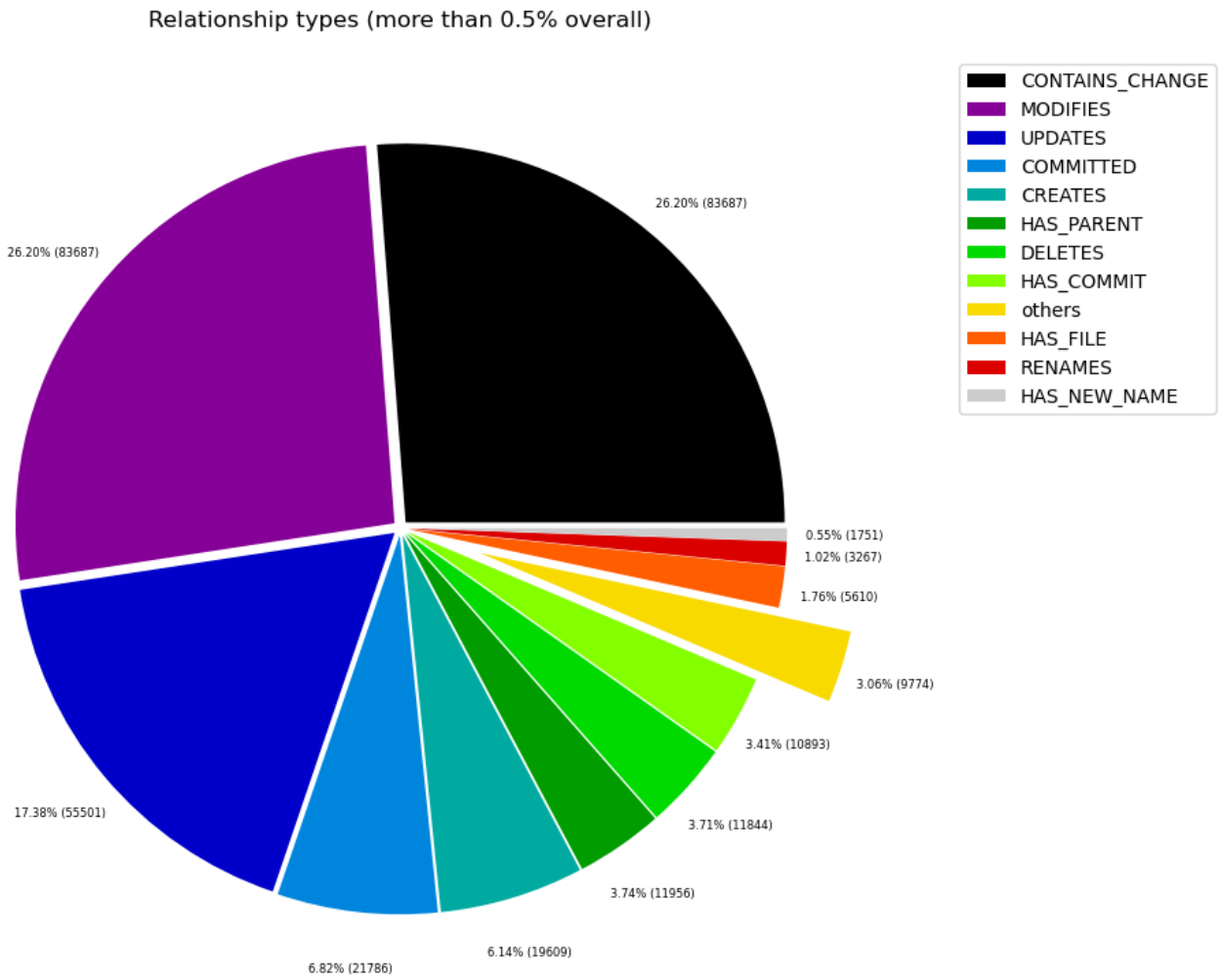


Table 2b - Lowest relationship count by type

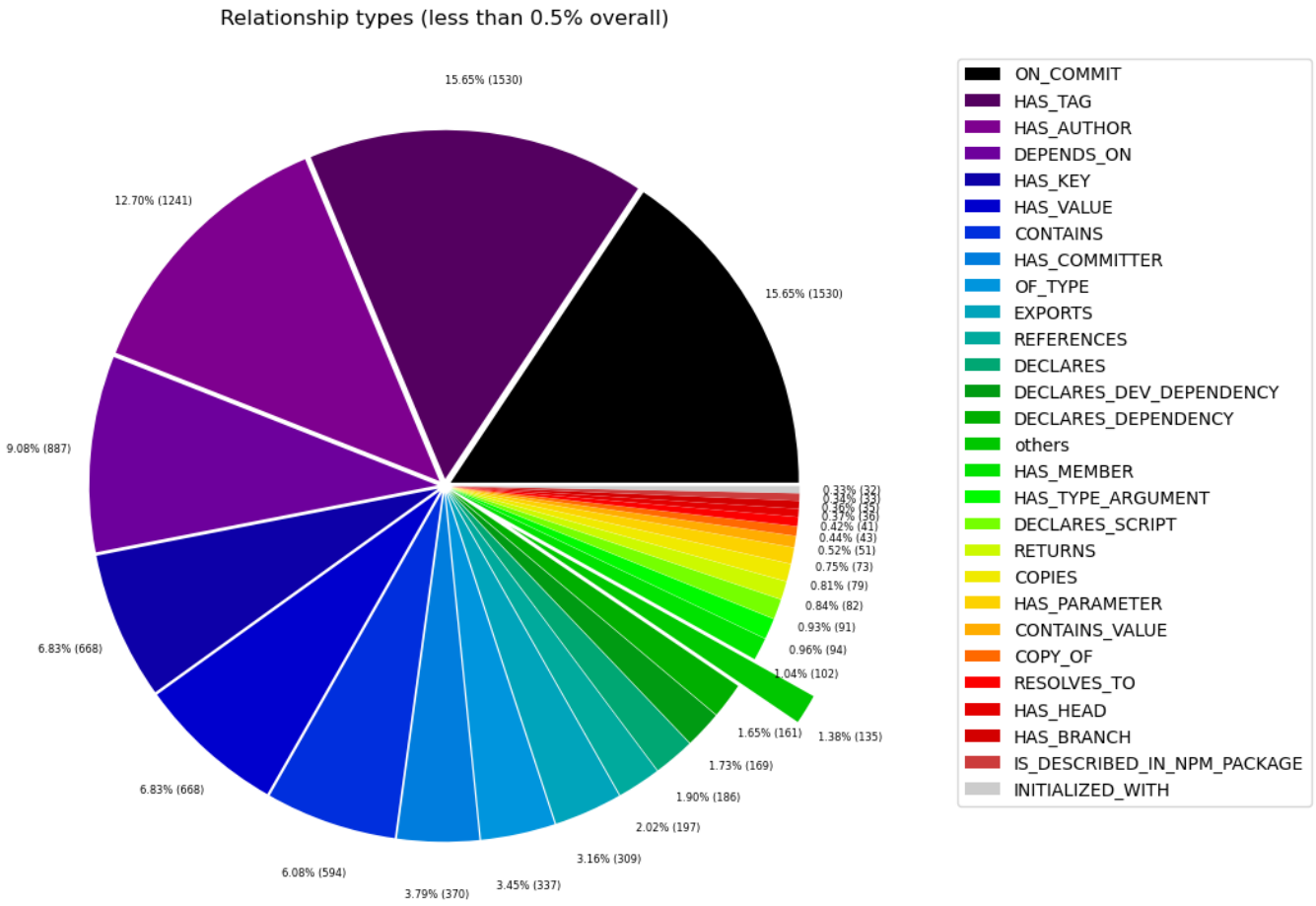
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONSTRAINED_BY	4	0.001252
1	REFERENCED_PROJECTS	5	0.001566
2	MEMBER	6	0.001879
3	HAS_ROOT	6	0.001879
4	HAS_NPM_PACKAGE	6	0.001879
5	HAS_CONFIG	6	0.001879
6	HAS_ARGUMENT	6	0.001879
7	EXTENDS	6	0.001879
8	DECLARES_ENGINE	6	0.001879
9	CONTAINS_PROJECT	6	0.001879
10	CALLS	6	0.001879
11	PARENT	6	0.001879
12	DECLARES_PEER_DEPENDENCY	8	0.002505
13	USES	11	0.003444
14	INCLUDES_CONCEPT	19	0.005949
15	REQUIRES_CONCEPT	28	0.008767
16	INITIALIZED_WITH	32	0.010020
17	IS_DESCRIBED_IN_NPM_PACKAGE	33	0.010333
18	HAS_BRANCH	35	0.010959
19	HAS_HEAD	36	0.011272
20	RESOLVES_TO	41	0.012838
21	COPY_OF	43	0.013464
22	CONTAINS_VALUE	51	0.015969
23	HAS_PARAMETER	73	0.022858
24	COPIES	79	0.024737
25	RETURNS	82	0.025676
26	DECLARES_SCRIPT	91	0.028494
27	HAS_TYPE_ARGUMENT	94	0.029433
28	HAS_MEMBER	102	0.031938
29	DECLARES_DEPENDENCY	161	0.050413

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	numberOfNodesWithSame
0	[Git, Change]	MODIFIES	[File, Git]	83687	83687	
1	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	83687	10893	
2	[Git, Change]	UPDATES	[File, Git]	55501	83687	
3	[Git, Change]	CREATES	[File, Git]	19609	83687	
4	[Git, Commit]	HAS_PARENT	[Git, Commit]	11956	10893	
5	[Git, Change]	DELETES	[File, Git]	11844	83687	
6	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	10893	1	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	10893	1241	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10893	370	
9	[Repository, File, Git]	HAS_FILE	[File, Git]	5610	1	
10	[Git, Change]	RENAMES	[File, Git]	3267	83687	
11	[File, Git]	HAS_NEW_NAME	[File, Git]	1751	5610	
12	[Repository, File, Git]	HAS_TAG	[Git, Tag]	1530	1	
13	[Git, Tag]	ON_COMMIT	[Git, Commit]	1530	1530	
14	[Repository, File, Git]	HAS_AUTHOR	[Author, Git, Person]	1241	1	
15	[Json, Value, Object]	HAS_KEY	[Json, Key]	668	133	
16	[Json, Key]	HAS_VALUE	[Json, Value, Scalar]	552	668	
17	[Repository, File, Git]	HAS_COMMITTER	[Committer, Git, Person]	370	1	
18	[TS, Function]	DEPENDS_ON	[TS, ExternalDeclaration]	293	47	
19	[File, TS, Local, Module]	DEPENDS_ON	[TS, ExternalDeclaration]	236	6	
20	[TS, ExternalModule]	EXPORTS	[TS, ExternalDeclaration]	215	11	
21	[Package, File, Json, NPM]	DECLARES_DEV_DEPENDENCY	[NPM, Dependency]	169	29	
22	[Package, File, Json, NPM]	DECLARES_DEPENDENCY	[NPM, Dependency]	161	29	
23	[Type, TS, Union]	CONTAINS	[Type, TS, Primitive]	147	119	
24	[Type, TS, Declared]	REFERENCES	[TS, ExternalDeclaration]	143	276	
25	[Type, TS, Union]	CONTAINS	[Type, TS, Literal]	119	119	
26	[Json, Key]	HAS_VALUE	[Json, Value, Object]	104	668	
27	[Type, Object, TS]	HAS_MEMBER	[Type, TS, ObjectMember]	101	39	
28	[Package, File, Json, NPM]	DECLARES_SCRIPT	[NPM, Script]	91	29	
29	[Git, Change]	COPIES	[File, Git]	79	83687	

Graph Density

```
total_number_of_nodes (vertices): 106912
total_number_of_relationships (edges): 319365
-> total directed graph density: 2.7940776621212913e-05
-> total directed graph density in percent: 0.0027940776621212912
```