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]	83571	76.609495
1	[Git, Commit]	10815	9.914105
2	[File, Git]	5583	5.117933
3	[Git, Tag]	1457	1.335631
4	[Author, Git, Person]	1237	1.133957
5	[Type, TS, Primitive]	811	0.743443
6	[Json, Key]	668	0.612355
7	[Json, Value, Scalar]	603	0.552770
8	[Type, TS, Declared]	598	0.548186
9	[TS, ExternalDeclaration]	450	0.412515
10	[Committer, Git, Person]	370	0.339179
11	[NPM, Dependency]	338	0.309844
12	[Type, TS, ObjectMember]	318	0.291510
13	[Type, TS, Literal]	274	0.251176
14	[Type, TS, Union]	246	0.225508
15	[TS, Property]	137	0.125588
16	[Json, Value, Object]	133	0.121921
17	[Value, TS, Literal]	124	0.113671
18	[Type, Object, TS]	109	0.099920
19	[TS, Function]	109	0.099920
20	[NPM, Script]	91	0.083420
21	[Value, TS, ObjectMember]	88	0.080670
22	[Type, TS, FunctionParameter]	80	0.073336
23	[TS, Parameter]	76	0.069669
24	[Type, TS, Function]	70	0.064169
25	[File, Directory, Local]	64	0.058669
26	[TS, Variable]	59	0.054085
27	[File, TS, Local, Module]	46	0.042168
28	[Git, Branch]	40	0.036668
29	[TS, Interface]	37	0.033918

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>
```

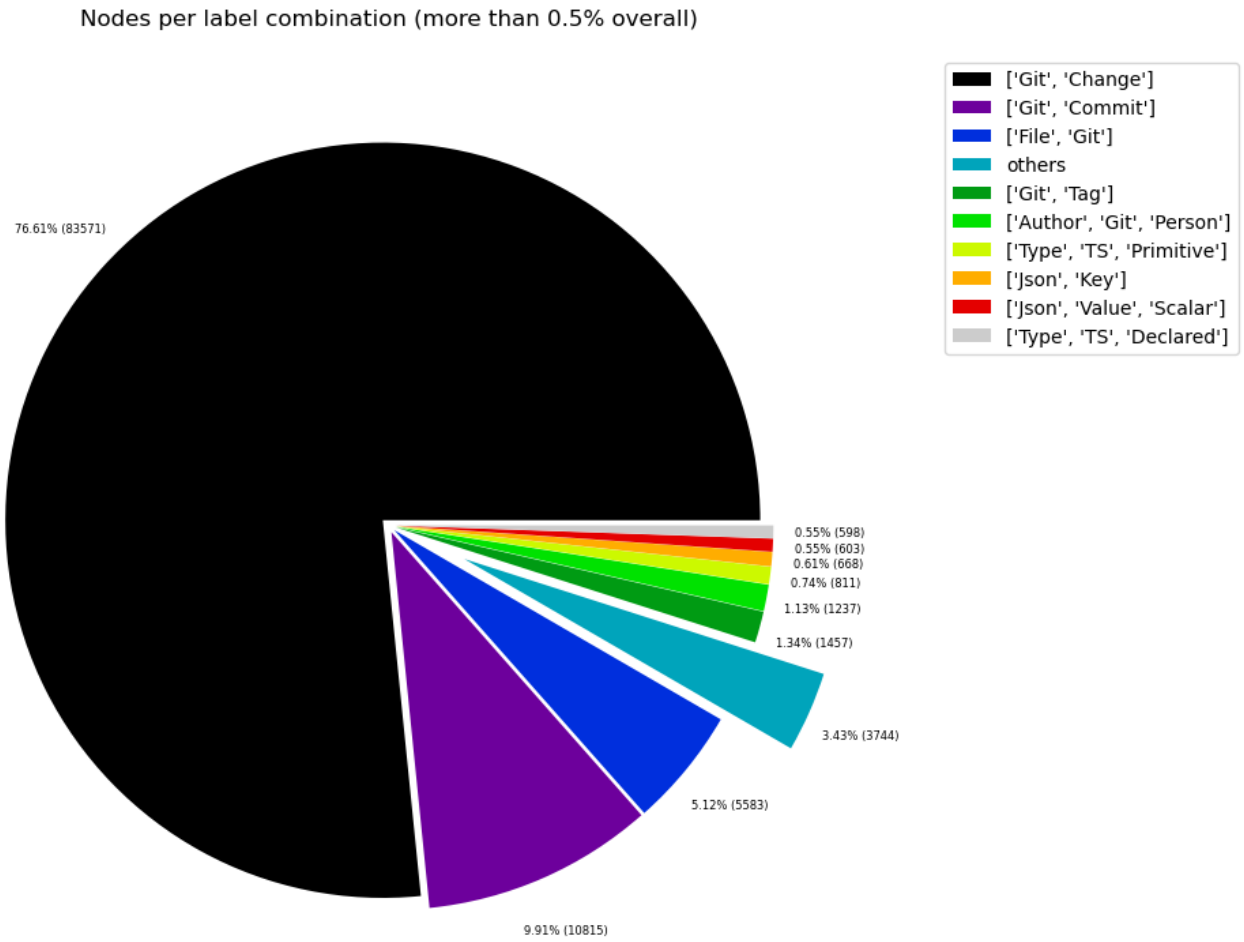


Table 1b - Lowest node count by label combination

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	0.000917
1	[Repository, File, Git]	1	0.000917
2	[Value, TS, Null]	1	0.000917
3	[TS, Constructor]	2	0.001833
4	[TS, Class]	2	0.001833
5	[TS, Enum]	4	0.003667
6	[TS, Method]	4	0.003667
7	[Value, Array, TS]	5	0.004583
8	[Type, TS, Tuple]	6	0.005500
9	[NPM, Engine]	6	0.005500
10	[TS, TypeParameter]	8	0.007334
11	[Value, TS, Complex]	11	0.010084
12	[Type, TS, TypeParameterReference]	12	0.011000
13	[Json, Value, Array]	12	0.011000
14	[Value, TS, Function]	13	0.011917
15	[Value, TS, Call]	14	0.012834
16	[Value, TS, Member]	14	0.012834
17	[TS, EnumMember]	16	0.014667
18	[jQAssistant, Rule, Concept]	19	0.017417
19	[Type, TS, NotIdentified]	23	0.021084
20	[Value, Object, TS]	28	0.025668
21	[File, Local]	28	0.025668
22	[File, TS, Scan]	29	0.026584
23	[Package, File, Json, NPM]	29	0.026584
24	[Value, TS, Declared]	30	0.027501
25	[TS, TypeAlias]	32	0.029334
26	[TS, ExternalModule]	33	0.030251
27	[Project, TS]	33	0.030251
28	[Type, TS, Intersection]	34	0.031168
29	[File, Directory]	35	0.032084

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>
```

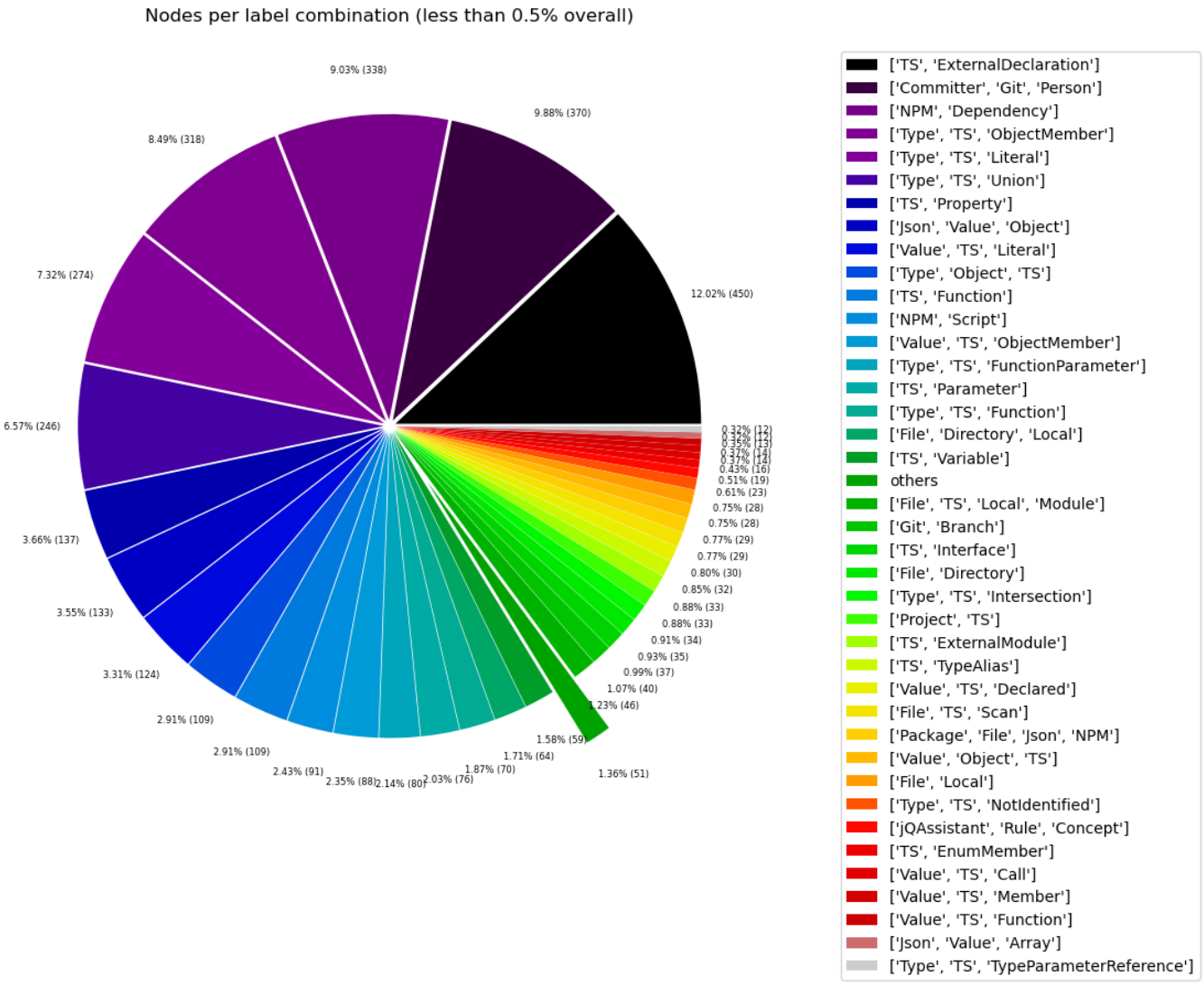


Table 1c - Highest node count by single label

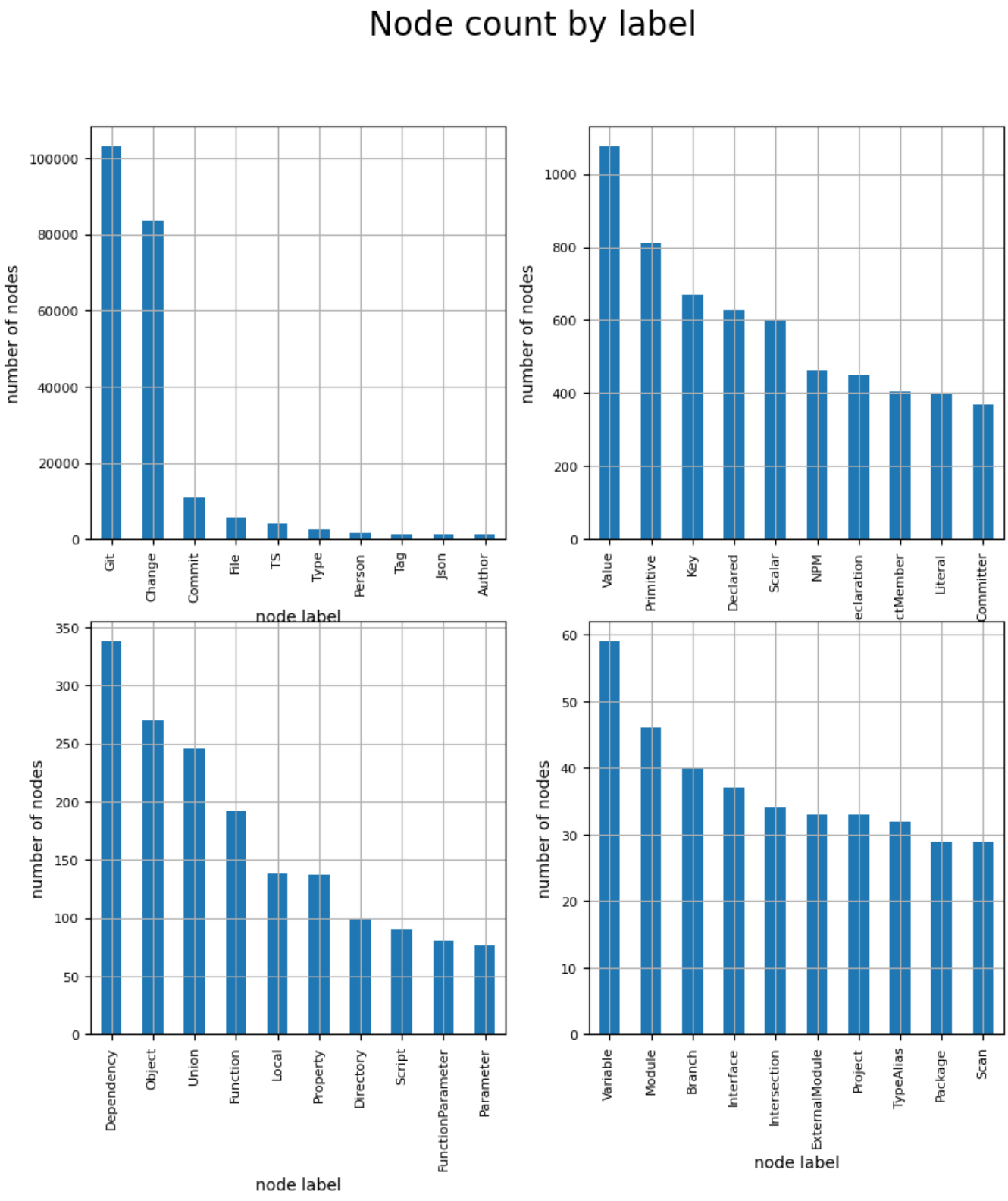
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	103074	94.487886
1	Change	83571	76.609495
2	Commit	10815	9.914105
3	File	5815	5.330608
4	TS	3986	3.653964
5	Type	2581	2.366001
6	Person	1607	1.473136
7	Tag	1457	1.335631
8	Json	1445	1.324631
9	Author	1237	1.133957
10	Value	1076	0.986369
11	Primitive	811	0.743443
12	Key	668	0.612355
13	Declared	628	0.575687
14	Scalar	603	0.552770
15	NPM	464	0.425349
16	ExternalDeclaration	450	0.412515
17	ObjectMember	406	0.372180
18	Literal	398	0.364846
19	Committer	370	0.339179
20	Dependency	338	0.309844
21	Object	270	0.247509
22	Union	246	0.225508
23	Function	192	0.176006
24	Local	138	0.126505
25	Property	137	0.125588
26	Directory	99	0.090753
27	Script	91	0.083420
28	FunctionParameter	80	0.073336
29	Parameter	76	0.069669
30	Variable	59	0.054085
31	Module	46	0.042168
32	Branch	40	0.036668
33	Interface	37	0.033918
34	Intersection	34	0.031168
35	ExternalModule	33	0.030251
36	Project	33	0.030251
37	TypeAlias	32	0.029334
38	Package	29	0.026584
39	Scan	29	0.026584

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: 322562
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	83571	25.908508
1	MODIFIES	83571	25.908508
2	UPDATES	54841	17.001693
3	COMMITTED	21630	6.705688
4	CREATES	19959	6.187648
5	DELETES	12041	3.732926
6	HAS_PARENT	11873	3.680843
7	HAS_COMMIT	10815	3.352844
8	HAS_FILE	5583	1.730830
9	RENAMES	3270	1.013759
10	DEPENDS_ON	1845	0.571983
11	HAS_NEW_NAME	1751	0.542841
12	HAS_TAG	1457	0.451696
13	ON_COMMIT	1457	0.451696
14	HAS_AUTHOR	1237	0.383492
15	CONTAINS	1199	0.371711
16	OF_TYPE	1030	0.319318
17	HAS_KEY	668	0.207092
18	HAS_VALUE	668	0.207092
19	EXPORTS	659	0.204302
20	REFERENCES	489	0.151599
21	DECLARES	410	0.127107
22	HAS_MEMBER	406	0.125867
23	HAS_COMMITTER	370	0.114707
24	HAS_TYPE_ARGUMENT	202	0.062624
25	RETURNS	183	0.056733
26	DECLARES_DEV_DEPENDENCY	169	0.052393
27	DECLARES_DEPENDENCY	161	0.049913
28	HAS_PARAMETER	155	0.048053
29	RESOLVES_TO	103	0.031932

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>
```

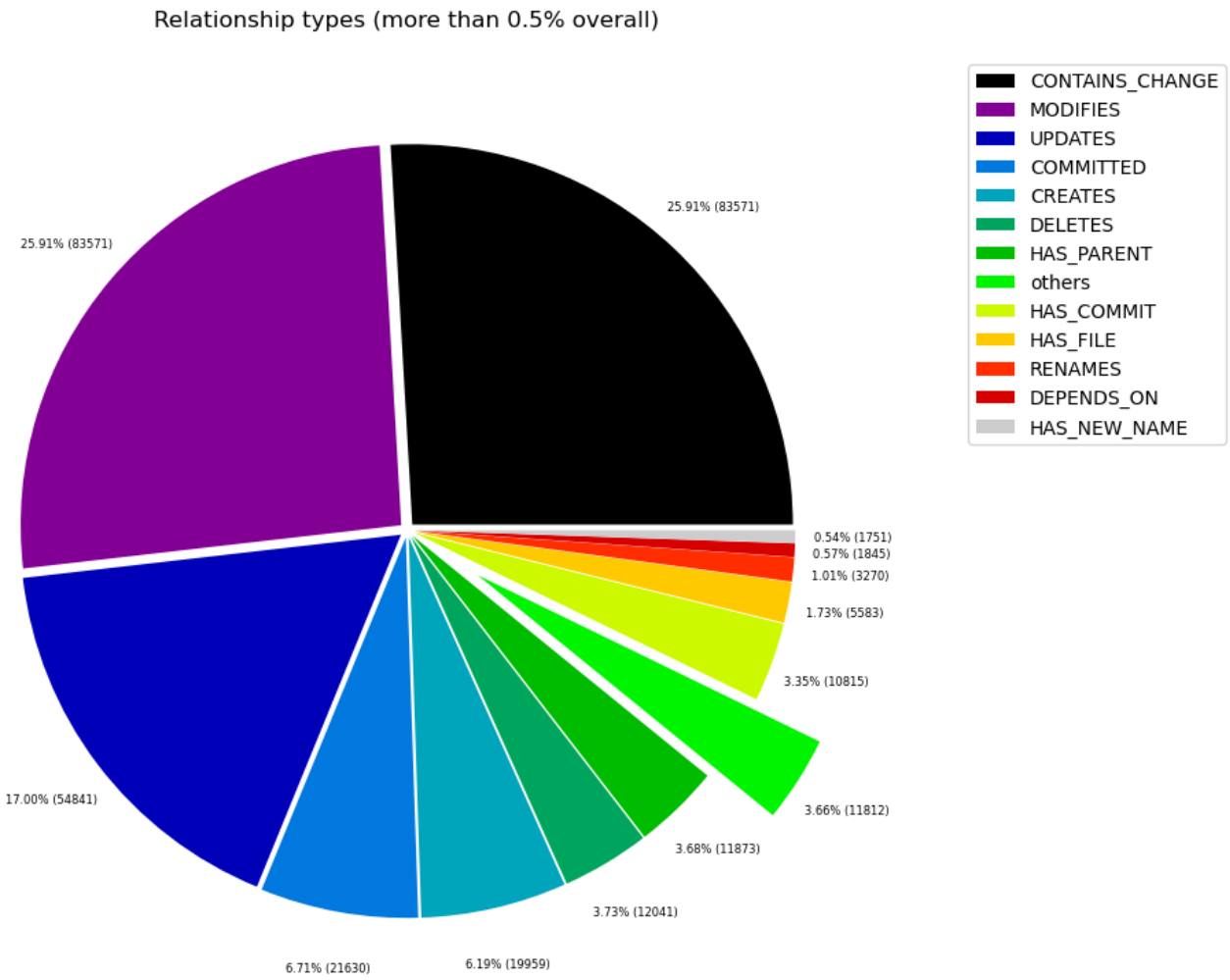


Table 2b - Lowest relationship count by type

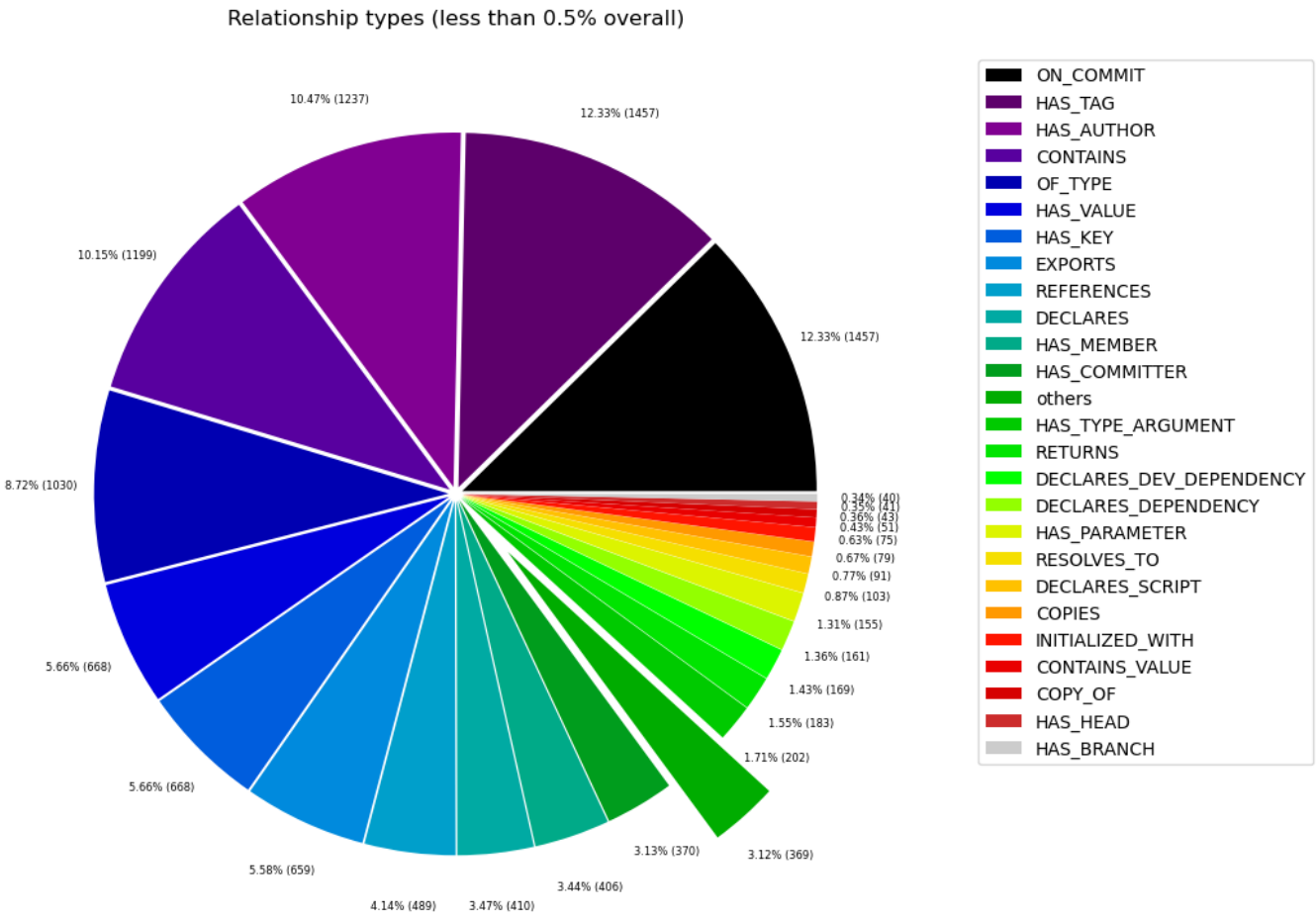
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	HAS	1	0.000310
1	REFERENCED_PROJECTS	5	0.001550
2	DECLARES_ENGINE	6	0.001860
3	SIMILAR	8	0.002480
4	DECLARES_PEER_DEPENDENCY	8	0.002480
5	CONSTRAINED_BY	8	0.002480
6	EXTENDS	12	0.003720
7	PARENT	14	0.004340
8	MEMBER	14	0.004340
9	HAS_ARGUMENT	14	0.004340
10	CALLS	14	0.004340
11	INCLUDES_CONCEPT	19	0.005890
12	PROVIDED_BY_NPM_DEPENDENCY	20	0.006200
13	REQUIRES_CONCEPT	28	0.008681
14	CONTAINS_PROJECT	33	0.010231
15	HAS_CONFIG	33	0.010231
16	IS_DESCRIBED_IN_NPM_PACKAGE	33	0.010231
17	HAS_ROOT	33	0.010231
18	HAS_NPM_PACKAGE	33	0.010231
19	USES	33	0.010231
20	HAS_BRANCH	40	0.012401
21	HAS_HEAD	41	0.012711
22	COPY_OF	43	0.013331
23	CONTAINS_VALUE	51	0.015811
24	INITIALIZED_WITH	75	0.023251
25	COPIES	79	0.024491
26	DECLARES_SCRIPT	91	0.028212
27	RESOLVES_TO	103	0.031932
28	HAS_PARAMETER	155	0.048053
29	DECLARES_DEPENDENCY	161	0.049913

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]	83571	10815	
1	[Git, Change]	MODIFIES	[File, Git]	83571	83571	
2	[Git, Change]	UPDATES	[File, Git]	54841	83571	
3	[Git, Change]	CREATES	[File, Git]	19959	83571	
4	[Git, Change]	DELETES	[File, Git]	12041	83571	
5	[Git, Commit]	HAS_PARENT	[Git, Commit]	11873	10815	
6	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	10815	1	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	10815	1237	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10815	370	
9	[Repository, File, Git]	HAS_FILE	[File, Git]	5583	1	
10	[Git, Change]	RENAMES	[File, Git]	3270	83571	
11	[File, Git]	HAS_NEW_NAME	[File, Git]	1751	5583	
12	[Repository, File, Git]	HAS_TAG	[Git, Tag]	1457	1	
13	[Git, Tag]	ON_COMMIT	[Git, Commit]	1457	1457	
14	[Repository, File, Git]	HAS_AUTHOR	[Author, Git, Person]	1237	1	
15	[Json, Value, Object]	HAS_KEY	[Json, Key]	668	133	
16	[TS, Function]	DEPENDS_ON	[TS, ExternalDeclaration]	588	109	
17	[Json, Key]	HAS_VALUE	[Json, Value, Scalar]	552	668	
18	[TS, ExternalModule]	EXPORTS	[TS, ExternalDeclaration]	450	33	
19	[Repository, File, Git]	HAS_COMMITTER	[Committer, Git, Person]	370	1	
20	[Type, Object, TS]	HAS_MEMBER	[Type, TS, ObjectMember]	318	109	
21	[File, TS, Local, Module, Mark4ModuleWeaklyCon...	DEPENDS_ON	[TS, ExternalDeclaration]	312	1	
22	[Type, TS, Union]	CONTAINS	[Type, TS, Primitive]	303	246	
23	[Type, TS, Declared]	REFERENCES	[TS, ExternalDeclaration]	288	598	
24	[Type, TS, Union]	CONTAINS	[Type, TS, Literal]	238	246	
25	[Type, TS, ObjectMember]	OF_TYPE	[Type, TS, Primitive]	173	318	
26	[Package, File, Json, NPM]	DECLARES_DEV_DEPENDENCY	[NPM, Dependency]	169	29	
27	[Package, File, Json, NPM]	DECLARES_DEPENDENCY	[NPM, Dependency]	161	29	
28	[TS, Function]	DEPENDS_ON	[TS, ExternalModule]	148	109	
29	[Type, TS, Union]	CONTAINS	[Type, TS, Declared]	145	246	

Graph Density

```
total_number_of_nodes (vertices): 109087
total_number_of_relationships (edges): 322562
-> total directed graph density: 2.710635910920788e-05
-> total directed graph density in percent: 0.002710635910920788
```