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: 95259
```

	nodeLabels	nodesWithThatLabels	nodesWithThatLabelsPercent
0	[Git, Change]	73974	77.655655
1	[Git, Commit]	9993	10.490347
2	[File, Git]	5102	5.355924
3	[Author, Git, Person]	1183	1.241877
4	[Git, Tag]	1080	1.133751
5	[Json, Key]	668	0.701246
6	[Json, Value, Scalar]	603	0.633011
7	[Committer, Git, Person]	371	0.389465
8	[NPM, Dependency]	330	0.346424
9	[Type, TS, Primitive, ExternalType]	285	0.299184
10	[Type, TS, Declared, ExternalType]	272	0.285537
11	[TS, ExternalDeclaration]	215	0.225700
12	[Type, TS, Literal, ExternalType]	136	0.142769
13	[Json, Value, Object]	133	0.139619
14	[Type, TS, Union, ExternalType]	117	0.122823
15	[Type, TS, ObjectMember, ExternalType]	98	0.102877
16	[NPM, Script]	91	0.095529
17	[TS, Property]	65	0.068235
18	[TS, Function]	47	0.049339
19	[Type, Object, TS, ExternalType]	38	0.039891
20	[Type, TS, FunctionParameter, ExternalType]	37	0.038841
21	[File, Directory]	34	0.035692
22	[TS, Parameter]	33	0.034642
23	[Type, TS, Function, ExternalType]	32	0.033593
24	[Package, File, Json, NPM]	29	0.030443
25	[Git, Branch]	27	0.028344
26	[TS, ExternalModule]	25	0.026244
27	[TS, Variable]	24	0.025194
28	[Value, TS, Literal]	20	0.020995
29	[jQAssistant, Rule, Concept]	19	0.019946

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>
```

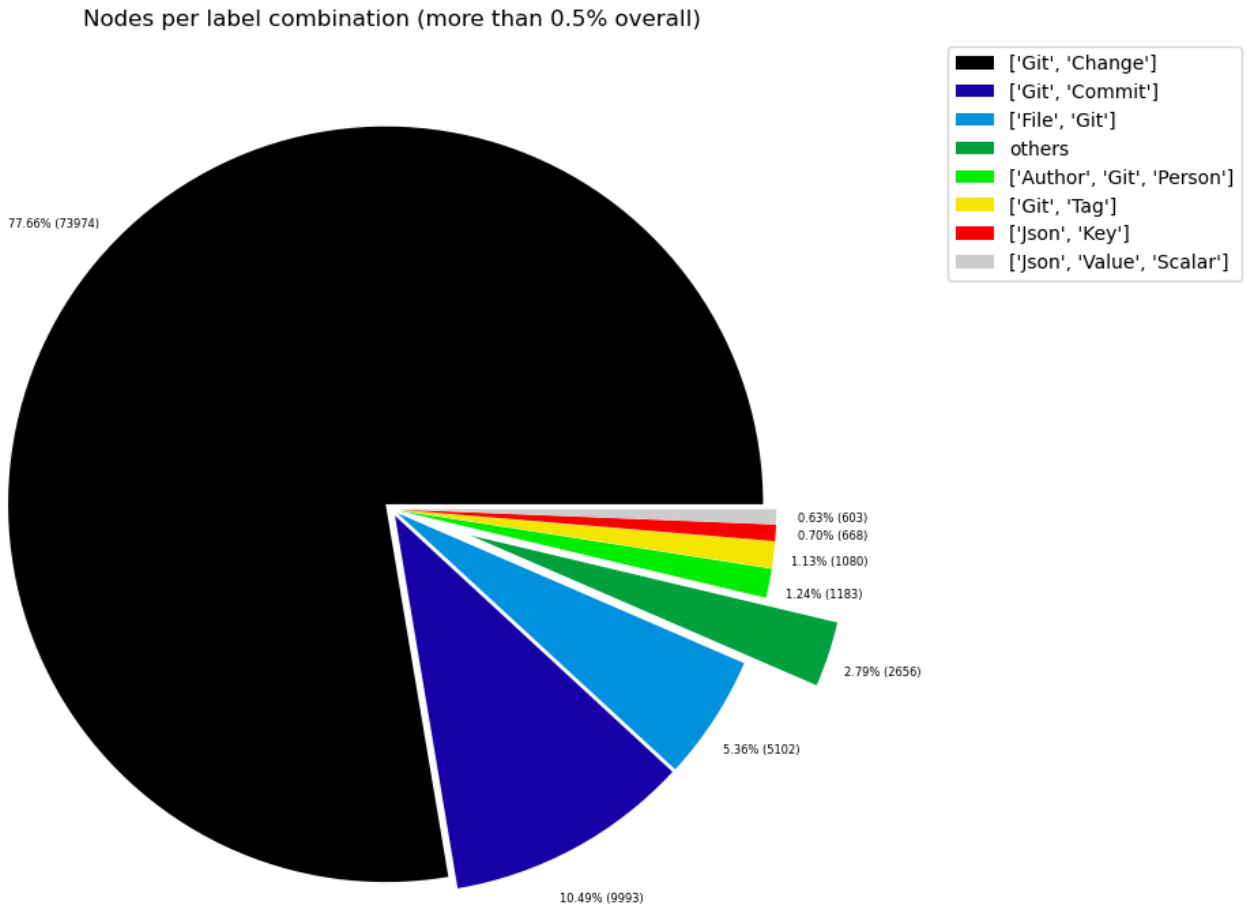


Table 1b - Lowest node count by label combination

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	0.001050
1	[File, TS, Scan]	1	0.001050
2	[TS, Method]	1	0.001050
3	[Value, TS, ObjectMember]	1	0.001050
4	[TS, Constructor]	1	0.001050
5	[TS, Class]	1	0.001050
6	[TS, Enum]	2	0.002100
7	[Value, Object, TS]	3	0.003149
8	[Type, TS, Tuple, ExternalType]	3	0.003149
9	[Value, TS, Function]	4	0.004199
10	[TS, TypeParameter]	4	0.004199
11	[Value, TS, Complex]	5	0.005249
12	[NPM, Engine]	6	0.006299
13	[Project, TS]	6	0.006299
14	[File, Local]	6	0.006299
15	[Value, TS, Call]	6	0.006299
16	[Value, TS, Member]	6	0.006299
17	[File, TS, Local, Module]	6	0.006299
18	[Type, TS, TypeParameterReference, ExternalType]	6	0.006299
19	[TS, EnumMember]	8	0.008398
20	[Type, TS, NotIdentified, ExternalType]	11	0.011547
21	[Json, Value, Array]	12	0.012597
22	[Value, TS, Declared]	13	0.013647
23	[TS, TypeAlias]	14	0.014697
24	[File, Directory, Local]	16	0.016796
25	[Type, TS, Intersection, ExternalType]	17	0.017846
26	[TS, Interface]	18	0.018896
27	[jQAssistant, Rule, Concept]	19	0.019946
28	[Value, TS, Literal]	20	0.020995
29	[TS, Variable]	24	0.025194

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>
```

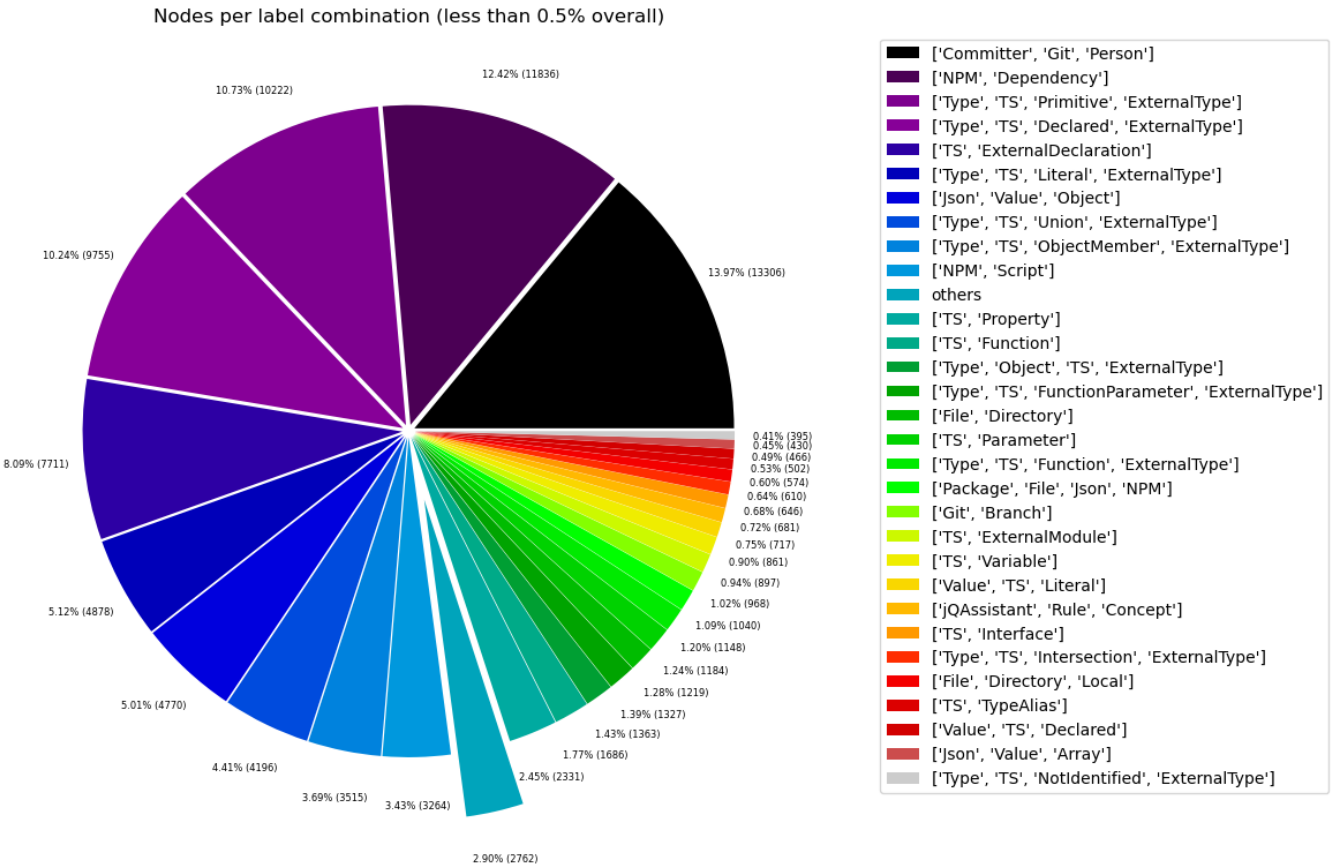


Table 1c - Highest node count by single label

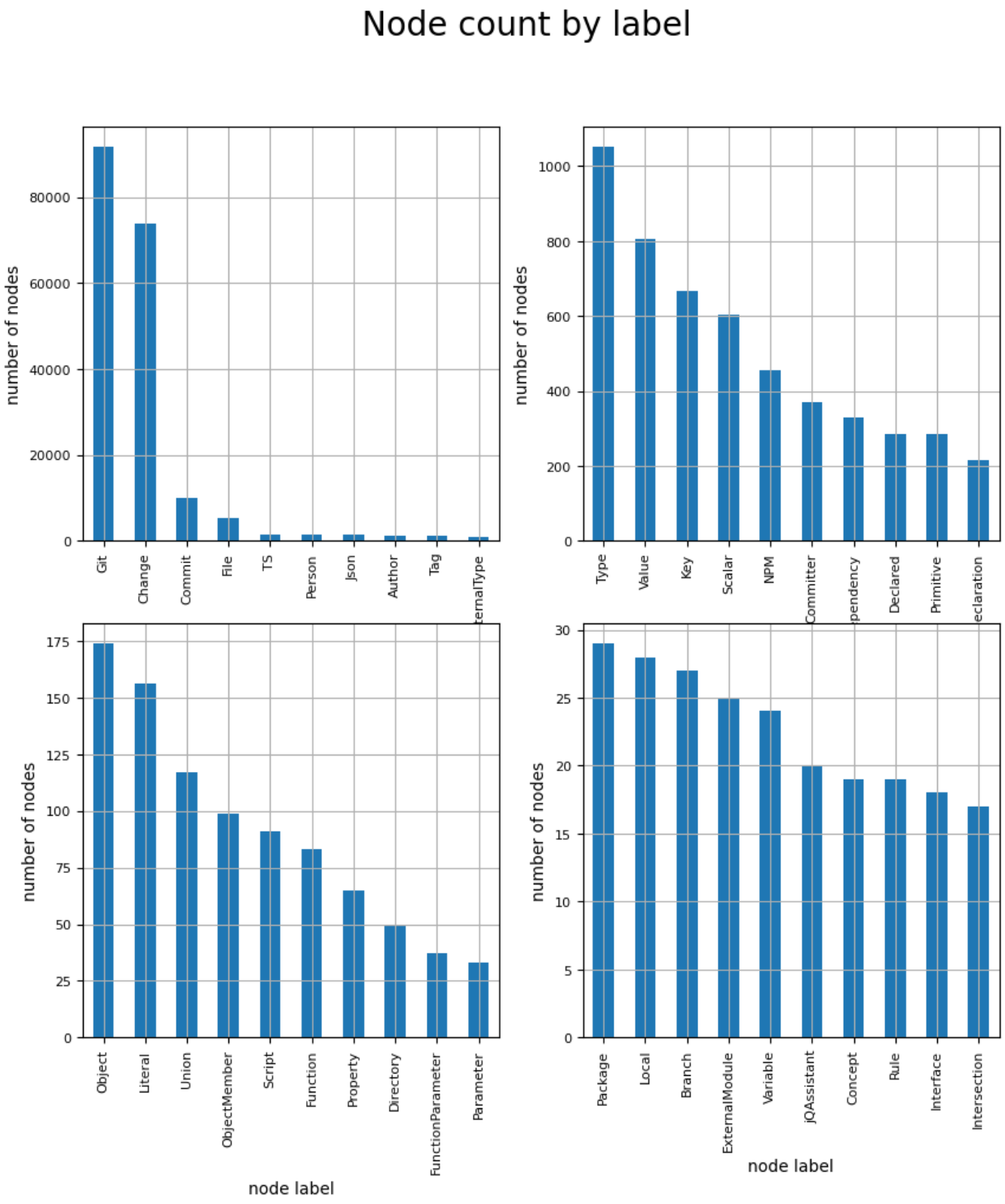
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	91730	96.295363
1	Change	73974	77.655655
2	Commit	9993	10.490347
3	File	5194	5.452503
4	TS	1581	1.659686
5	Person	1554	1.631342
6	Json	1445	1.516917
7	Author	1183	1.241877
8	Tag	1080	1.133751
9	ExternalType	1052	1.104358
10	Type	1052	1.104358
11	Value	806	0.846114
12	Key	668	0.701246
13	Scalar	603	0.633011
14	NPM	456	0.478695
15	Committer	371	0.389465
16	Dependency	330	0.346424
17	Declared	285	0.299184
18	Primitive	285	0.299184
19	ExternalDeclaration	215	0.225700
20	Object	174	0.182660
21	Literal	156	0.163764
22	Union	117	0.122823
23	ObjectMember	99	0.103927
24	Script	91	0.095529
25	Function	83	0.087131
26	Property	65	0.068235
27	Directory	50	0.052488
28	FunctionParameter	37	0.038841
29	Parameter	33	0.034642
30	Package	29	0.030443
31	Local	28	0.029394
32	Branch	27	0.028344
33	ExternalModule	25	0.026244
34	Variable	24	0.025194
35	jQAssistant	20	0.020995
36	Concept	19	0.019946
37	Rule	19	0.019946
38	Interface	18	0.018896
39	Intersection	17	0.017846

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: 266102
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	73974	27.799115
1	MODIFIES	73974	27.799115
2	UPDATES	49121	18.459463
3	COMMITTED	19986	7.510654
4	CREATES	17384	6.532833
5	HAS_PARENT	11005	4.135632
6	DELETES	10220	3.840633
7	RENAMES	2751	1.033814
8	HAS_NEW_NAME	1560	0.586241
9	ON_COMMIT	1080	0.405859
10	DEPENDS_ON	962	0.361516
11	HAS_KEY	668	0.251032
12	HAS_VALUE	668	0.251032
13	CONTAINS	589	0.221344
14	OF_TYPE	329	0.123637
15	EXPORTS	275	0.103344
16	REFERENCES	196	0.073656
17	DECLARES	185	0.069522
18	DECLARES_DEV_DEPENDENCY	169	0.063509
19	DECLARES_DEPENDENCY	161	0.060503
20	HAS_MEMBER	99	0.037204
21	HAS_TYPE_ARGUMENT	92	0.034573
22	DECLARES_SCRIPT	91	0.034197
23	RESOLVES_TO	80	0.030064
24	RETURNS	80	0.030064
25	HAS_PARAMETER	70	0.026306
26	CONTAINS_VALUE	51	0.019166
27	COPIES	43	0.016159
28	INITIALIZED_WITH	32	0.012025
29	COPY_OF	28	0.010522

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>
```

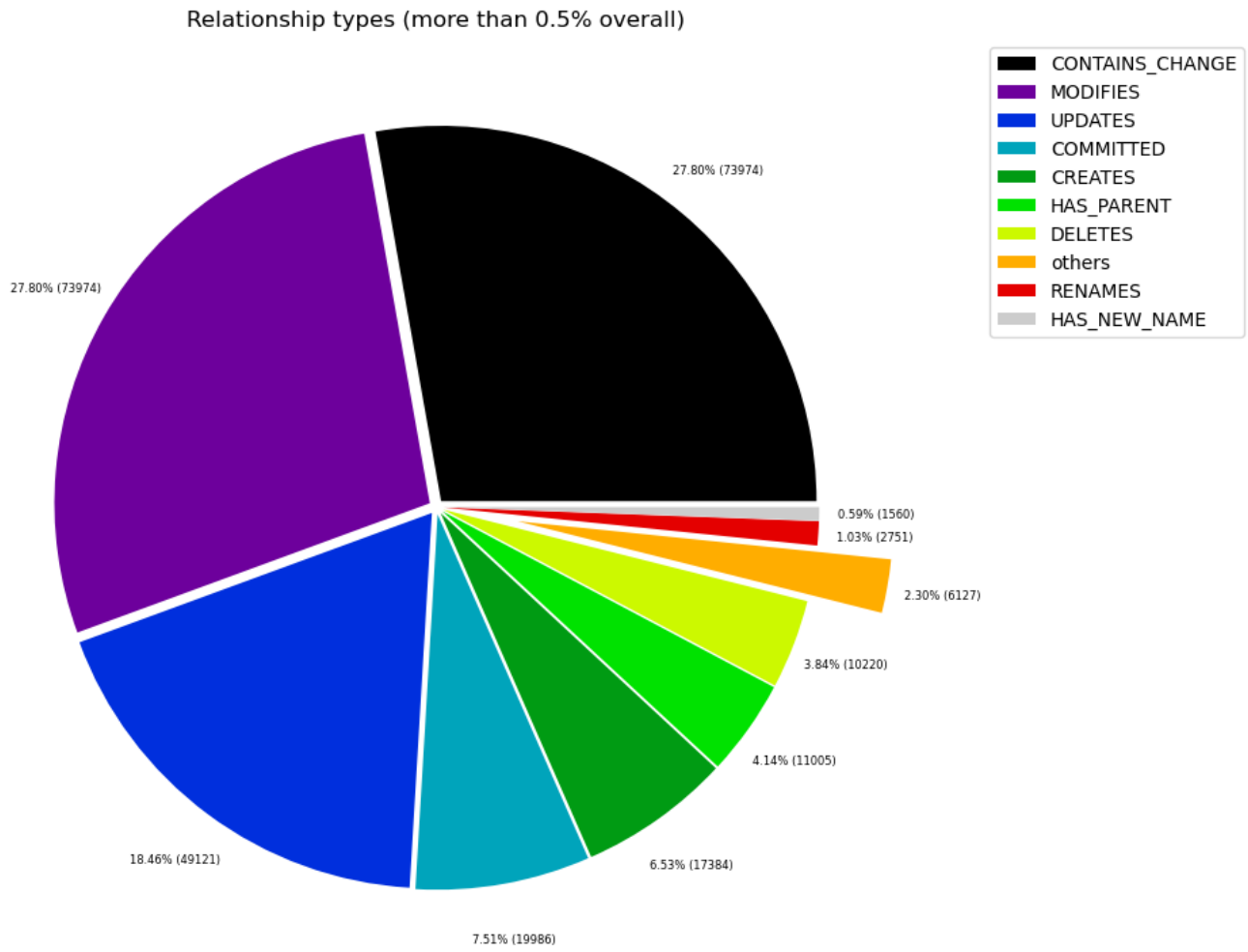


Table 2b - Lowest relationship count by type

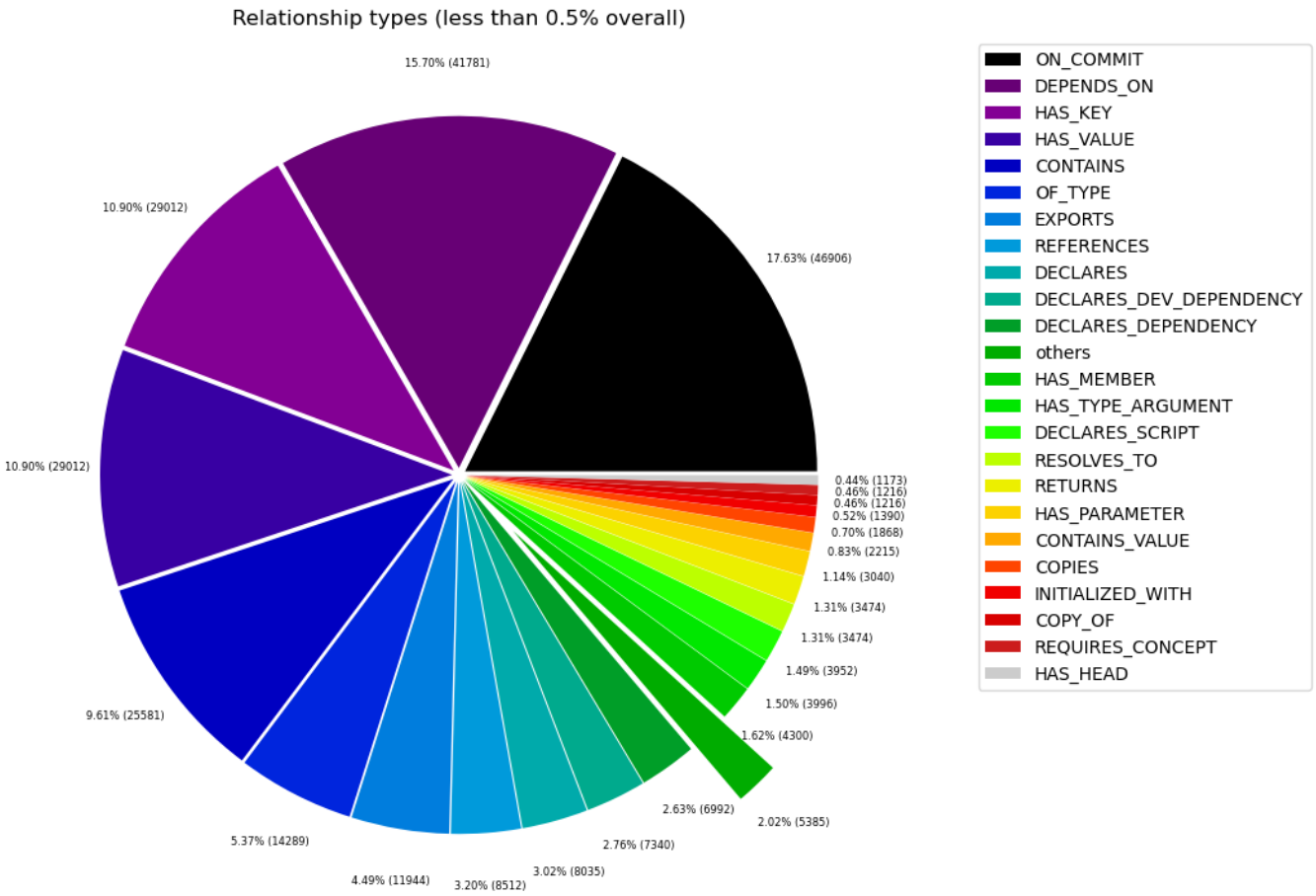
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONSTRAINED_BY	4	0.001503
1	REFERENCED_PROJECTS	5	0.001879
2	MEMBER	6	0.002255
3	HAS_ROOT	6	0.002255
4	HAS_NPM_PACKAGE	6	0.002255
5	HAS_CONFIG	6	0.002255
6	HAS_ARGUMENT	6	0.002255
7	DECLARES_ENGINE	6	0.002255
8	CONTAINS_PROJECT	6	0.002255
9	CALLS	6	0.002255
10	PARENT	6	0.002255
11	EXTENDS	7	0.002631
12	SIMILAR	10	0.003758
13	INCLUDES_CONCEPT	19	0.007140
14	USES	25	0.009395
15	HAS_HEAD	27	0.010146
16	REQUIRES_CONCEPT	28	0.010522
17	COPY_OF	28	0.010522
18	INITIALIZED_WITH	32	0.012025
19	COPIES	43	0.016159
20	CONTAINS_VALUE	51	0.019166
21	HAS_PARAMETER	70	0.026306
22	RETURNS	80	0.030064
23	RESOLVES_TO	80	0.030064
24	DECLARES_SCRIPT	91	0.034197
25	HAS_TYPE_ARGUMENT	92	0.034573
26	HAS_MEMBER	99	0.037204
27	DECLARES_DEPENDENCY	161	0.060503
28	DECLARES_DEV_DEPENDENCY	169	0.063509
29	DECLARES	185	0.069522

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]	73974	9993	
1	[Git, Change]	MODIFIES	[File, Git]	73974	73974	
2	[Git, Change]	UPDATES	[File, Git]	49121	73974	
3	[Git, Change]	CREATES	[File, Git]	17384	73974	
4	[Git, Commit]	HAS_PARENT	[Git, Commit]	11005	9993	
5	[Git, Change]	DELETES	[File, Git]	10220	73974	
6	[Author, Git, Person]	COMMITTED	[Git, Commit]	9993	1183	
7	[Committer, Git, Person]	COMMITTED	[Git, Commit]	9993	371	
8	[Git, Change]	RENAMES	[File, Git]	2751	73974	
9	[File, Git]	HAS_NEW_NAME	[File, Git]	1560	5102	
10	[Git, Tag]	ON_COMMIT	[Git, Commit]	1080	1080	
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	[Package, File, Json, NPM]	DECLARES_DEV_DEPENDENCY	[NPM, Dependency]	169	29	
16	[Package, File, Json, NPM]	DECLARES_DEPENDENCY	[NPM, Dependency]	161	29	
17	[File, TS, Local, Module, Mark4ModuleWeaklyCon...	DEPENDS_ON	[TS, ExternalDeclaration]	152	2	
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	5102	

Graph Density

```
total_number_of_nodes (vertices): 95259
total_number_of_relationships (edges): 266102
-> total directed graph density: 2.932517831578447e-05
-> total directed graph density in percent: 0.0029325178315784467
```