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: 90040
```

	nodeLabels	nodesWithThatLabels	nodesWithThatLabelsPercent
0	[Git, Change]	70945	78.792759
1	[Git, Commit]	9776	10.857397
2	[File, Git]	5032	5.588627
3	[Author, Git, Person]	1180	1.310529
4	[Git, Tag]	1026	1.139494
5	[Committer, Git, Person]	370	0.410928
6	[Type, TS, Primitive, ExternalType]	291	0.323190
7	[Type, TS, Declared, ExternalType]	286	0.317637
8	[TS, ExternalDeclaration]	211	0.234340
9	[Type, TS, Literal, ExternalType]	136	0.151044
10	[Type, TS, Union, ExternalType]	120	0.133274
11	[Type, TS, ObjectMember, ExternalType]	98	0.108841
12	[TS, Property]	65	0.072190
13	[TS, Function]	47	0.052199
14	[Type, TS, Object, ExternalType]	38	0.042203
15	[Type, TS, FunctionParameter, ExternalType]	38	0.042203
16	[TS, Parameter]	33	0.036650
17	[Type, TS, Function, ExternalType]	33	0.036650
18	[TS, ExternalModule]	25	0.027765
19	[Git, Branch]	25	0.027765
20	[File]	25	0.027765
21	[TS, Variable]	24	0.026655
22	[TS, Literal, Value]	20	0.022212
23	[jQAssistant, Rule, Concept]	19	0.021102
24	[TS, Interface]	18	0.019991
25	[Type, TS, Intersection, ExternalType]	17	0.018880
26	[File, Local, Directory]	16	0.017770
27	[File, Directory]	16	0.017770
28	[TS, TypeAlias]	14	0.015549
29	[TS, Declared, Value]	13	0.014438

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>
```

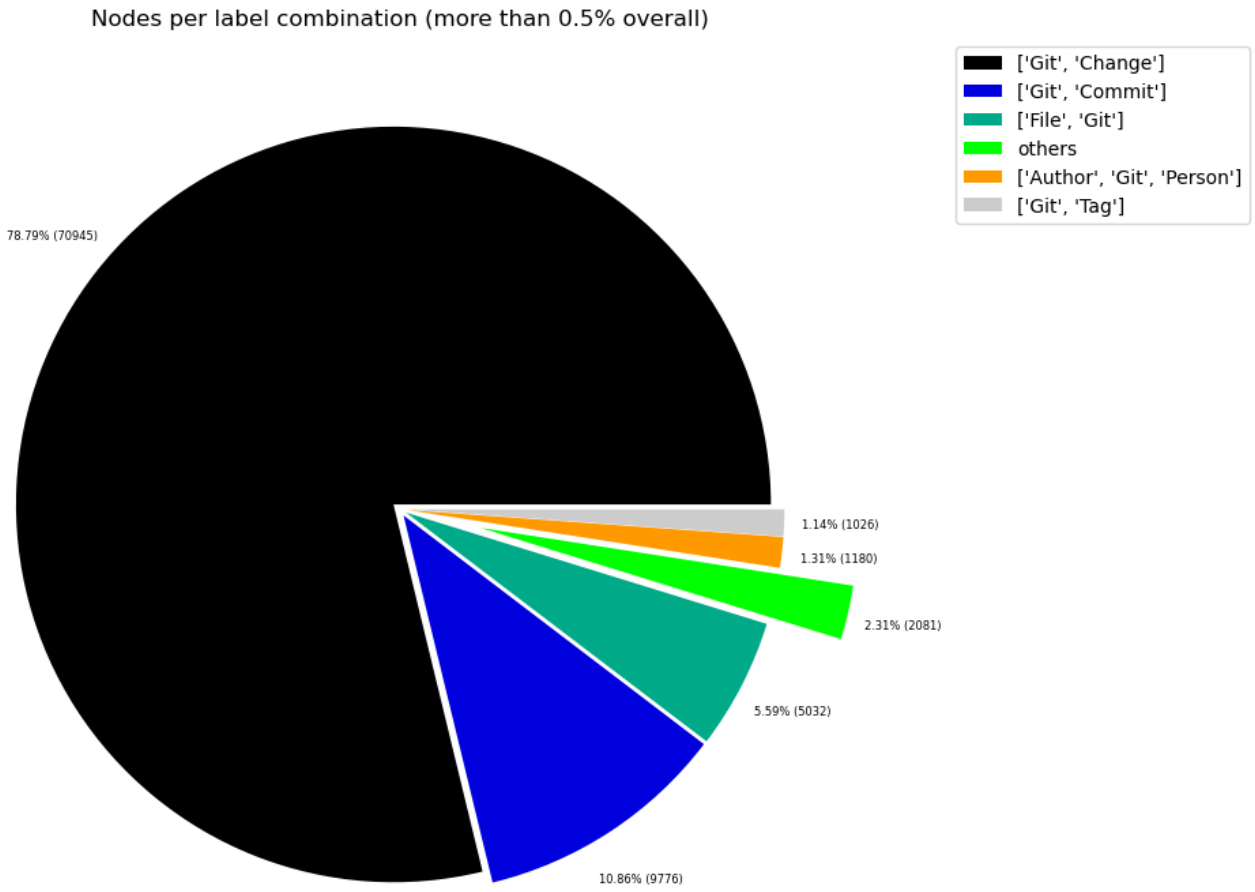


Table 1b - Lowest node count by label combination

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	0.001111
1	[File, TS, Scan]	1	0.001111
2	[TS, Class]	1	0.001111
3	[Repository, File, Git]	1	0.001111
4	[TS, Method]	1	0.001111
5	[TS, ObjectMember, Value]	1	0.001111
6	[TS, Constructor]	1	0.001111
7	[TS, Enum]	2	0.002221
8	[TS, Object, Value]	3	0.003332
9	[Type, TS, Tuple, ExternalType]	3	0.003332
10	[TS, Function, Value]	4	0.004442
11	[TS, TypeParameter]	4	0.004442
12	[TS, Value, Complex]	5	0.005553
13	[Project, TS]	6	0.006664
14	[File, Local]	6	0.006664
15	[Type, TS, TypeParameterReference, ExternalType]	6	0.006664
16	[TS, Value, Call]	6	0.006664
17	[File, TS, Local, Module]	6	0.006664
18	[TS, Value, Member]	6	0.006664
19	[TS, EnumMember]	8	0.008885
20	[Type, TS, NotIdentified, ExternalType]	11	0.012217
21	[TS, Declared, Value]	13	0.014438
22	[TS, TypeAlias]	14	0.015549
23	[File, Local, Directory]	16	0.017770
24	[File, Directory]	16	0.017770
25	[Type, TS, Intersection, ExternalType]	17	0.018880
26	[TS, Interface]	18	0.019991
27	[jQAssistant, Rule, Concept]	19	0.021102
28	[TS, Literal, Value]	20	0.022212
29	[TS, Variable]	24	0.026655

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>
```

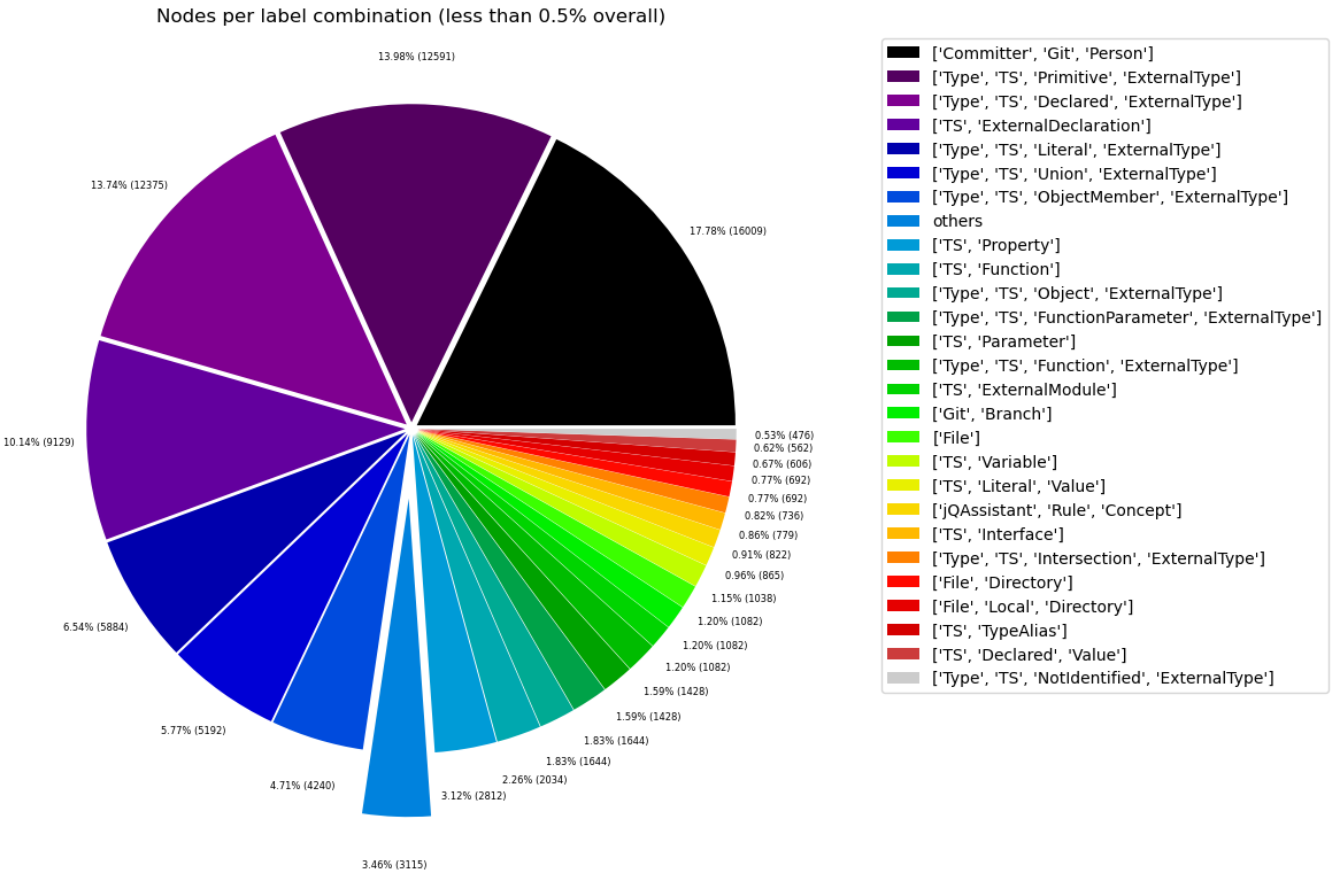


Table 1c - Highest node count by single label

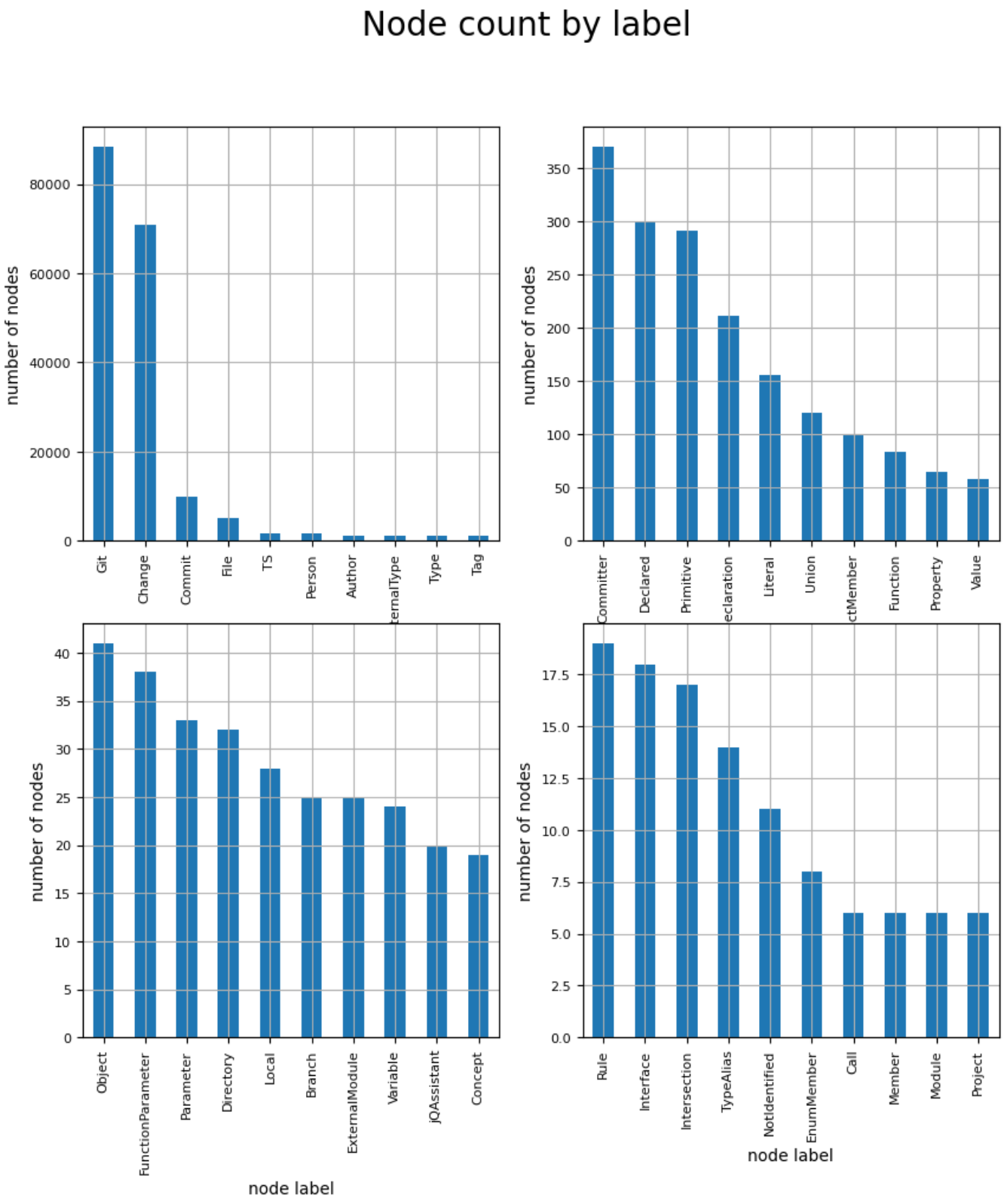
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	88355	98.128610
1	Change	70945	78.792759
2	Commit	9776	10.857397
3	File	5103	5.667481
4	TS	1602	1.779209
5	Person	1550	1.721457
6	Author	1180	1.310529
7	ExternalType	1077	1.196135
8	Type	1077	1.196135
9	Tag	1026	1.139494
10	Committer	370	0.410928
11	Declared	299	0.332075
12	Primitive	291	0.323190
13	ExternalDeclaration	211	0.234340
14	Literal	156	0.173256
15	Union	120	0.133274
16	ObjectMember	99	0.109951
17	Function	84	0.093292
18	Property	65	0.072190
19	Value	58	0.064416
20	Object	41	0.045535
21	FunctionParameter	38	0.042203
22	Parameter	33	0.036650
23	Directory	32	0.035540
24	Local	28	0.031097
25	Branch	25	0.027765
26	ExternalModule	25	0.027765
27	Variable	24	0.026655
28	jQAssistant	20	0.022212
29	Concept	19	0.021102
30	Rule	19	0.021102
31	Interface	18	0.019991
32	Intersection	17	0.018880
33	TypeAlias	14	0.015549
34	NotIdentified	11	0.012217
35	EnumMember	8	0.008885
36	Call	6	0.006664
37	Member	6	0.006664
38	Module	6	0.006664
39	Project	6	0.006664

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: 271602
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	70945	26.120942
1	MODIFIES	70945	26.120942
2	UPDATES	47213	17.383156
3	COMMITTED	19552	7.198769
4	CREATES	16563	6.098261
5	HAS_PARENT	10765	3.963520
6	DELETES	9857	3.629207
7	HAS_COMMIT	9776	3.599384
8	HAS_FILE	5032	1.852711
9	RENAMES	2688	0.989683
10	HAS_NEW_NAME	1541	0.567374
11	HAS_AUTHOR	1180	0.434459
12	HAS_TAG	1026	0.377759
13	ON_COMMIT	1026	0.377759
14	DEPENDS_ON	953	0.350881
15	CONTAINS	536	0.197348
16	HAS_COMMITTER	370	0.136229
17	OF_TYPE	330	0.121501
18	EXPORTS	271	0.099778
19	REFERENCES	198	0.072901
20	DECLARES	185	0.068114
21	HAS_MEMBER	99	0.036450
22	HAS_TYPE_ARGUMENT	99	0.036450
23	RETURNS	81	0.029823
24	HAS_PARAMETER	71	0.026141
25	INITIALIZED_WITH	32	0.011782
26	COPIES	29	0.010677
27	REQUIRES_CONCEPT	28	0.010309
28	RESOLVES_TO	27	0.009941
29	HAS_HEAD	26	0.009573

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>
```

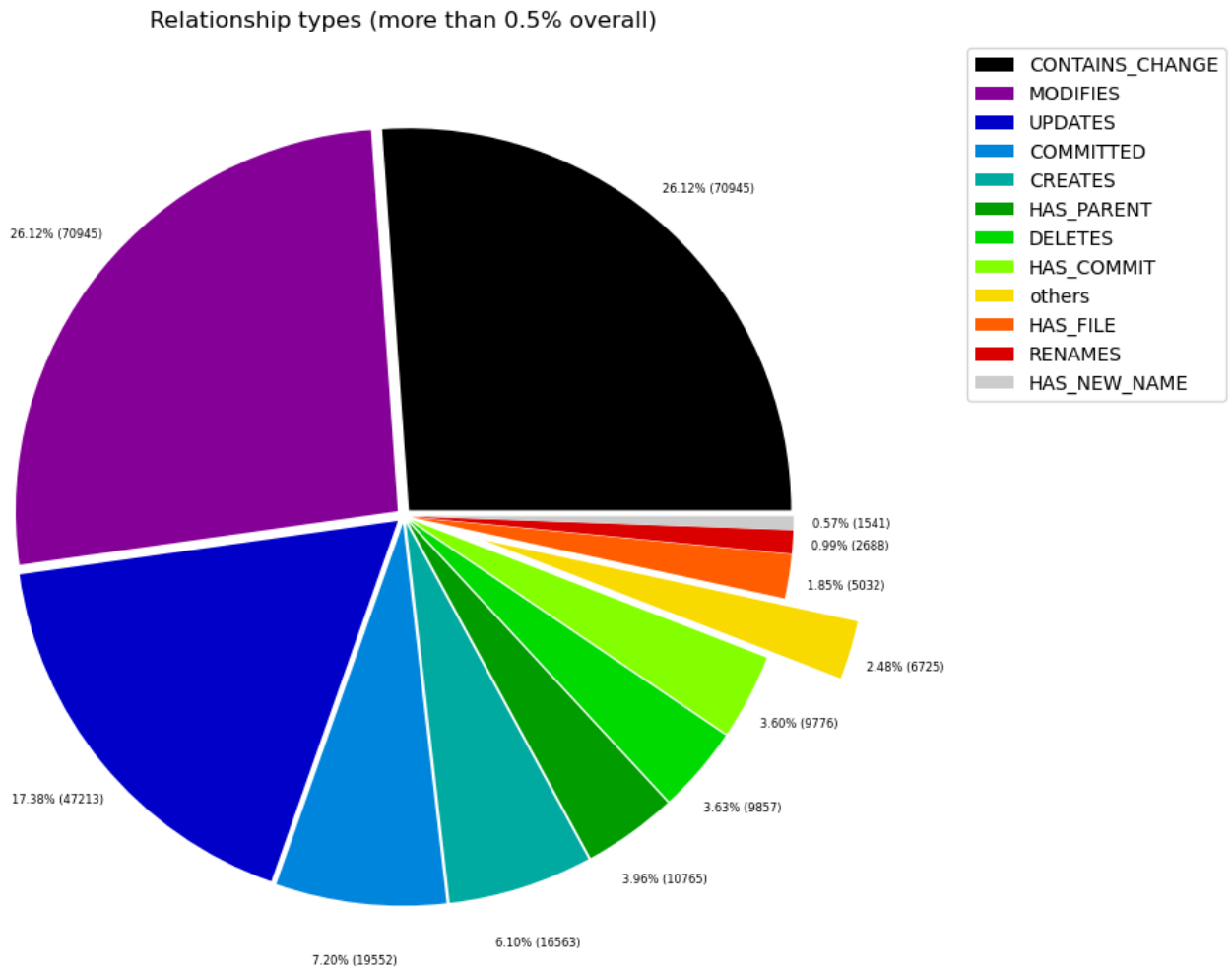


Table 2b - Lowest relationship count by type

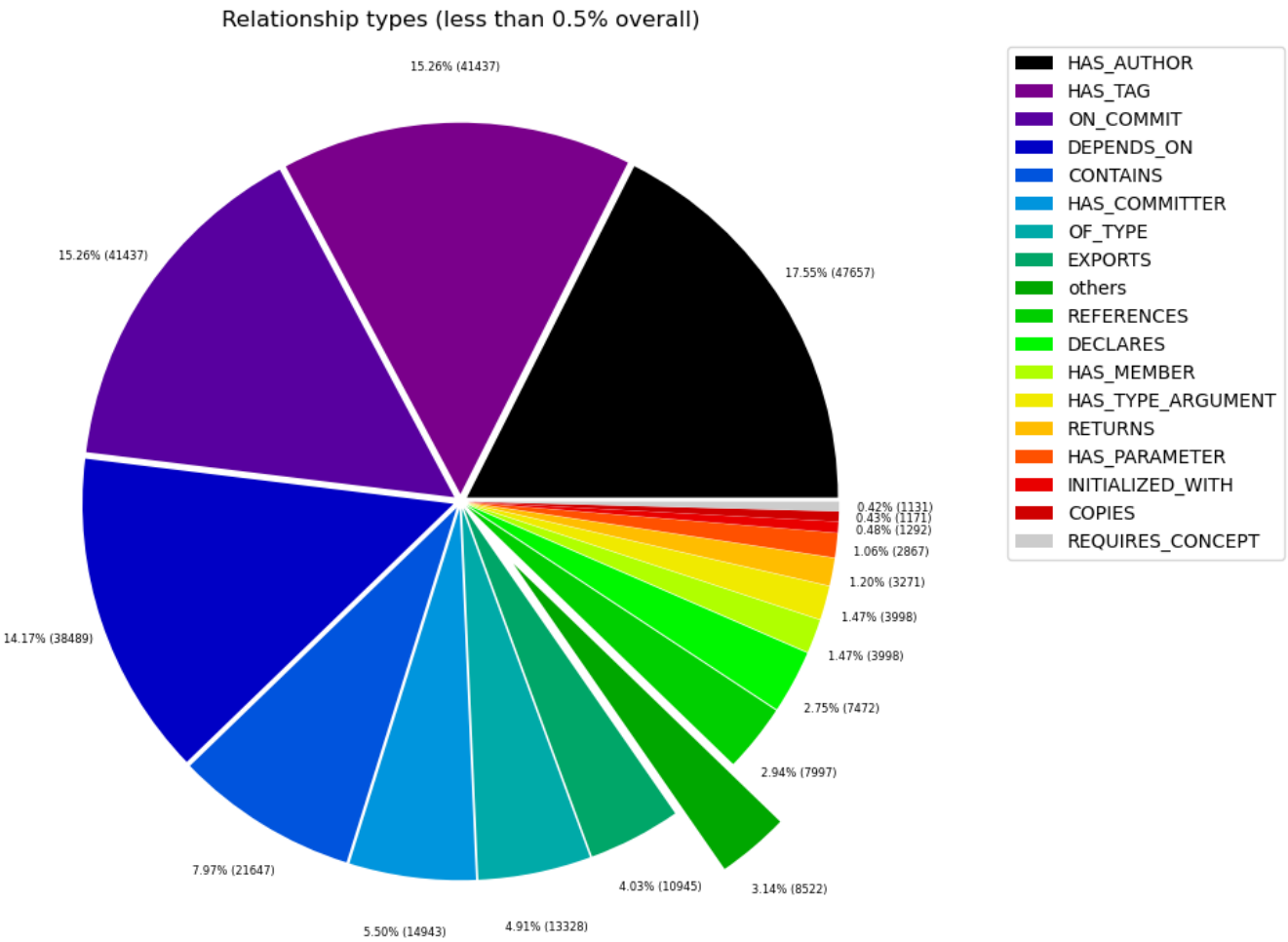
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONSTRAINED_BY	4	0.001473
1	REFERENCED_PROJECTS	5	0.001841
2	MEMBER	6	0.002209
3	HAS_ROOT	6	0.002209
4	HAS_CONFIG	6	0.002209
5	HAS_ARGUMENT	6	0.002209
6	CONTAINS_PROJECT	6	0.002209
7	CALLS	6	0.002209
8	PARENT	6	0.002209
9	EXTENDS	7	0.002577
10	SIMILAR	10	0.003682
11	INCLUDES_CONCEPT	19	0.006996
12	COPY_OF	21	0.007732
13	USES	25	0.009205
14	HAS_BRANCH	25	0.009205
15	HAS_HEAD	26	0.009573
16	RESOLVES_TO	27	0.009941
17	REQUIRES_CONCEPT	28	0.010309
18	COPIES	29	0.010677
19	INITIALIZED_WITH	32	0.011782
20	HAS_PARAMETER	71	0.026141
21	RETURNS	81	0.029823
22	HAS_TYPE_ARGUMENT	99	0.036450
23	HAS_MEMBER	99	0.036450
24	DECLARES	185	0.068114
25	REFERENCES	198	0.072901
26	EXPORTS	271	0.099778
27	OF_TYPE	330	0.121501
28	HAS_COMMITTER	370	0.136229
29	CONTAINS	536	0.197348

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	numberOfNodesWithSam
0	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	70945	9776	
1	[Git, Change]	MODIFIES	[File, Git]	70945	70945	
2	[Git, Change]	UPDATES	[File, Git]	47213	70945	
3	[Git, Change]	CREATES	[File, Git]	16563	70945	
4	[Git, Commit]	HAS_PARENT	[Git, Commit]	10765	9776	
5	[Git, Change]	DELETES	[File, Git]	9857	70945	
6	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	9776	1	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	9776	1180	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	9776	370	
9	[Repository, File, Git]	HAS_FILE	[File, Git]	5032	1	
10	[Git, Change]	RENAMES	[File, Git]	2688	70945	
11	[File, Git]	HAS_NEW_NAME	[File, Git]	1541	5032	
12	[Repository, File, Git]	HAS_AUTHOR	[Author, Git, Person]	1180	1	
13	[Repository, File, Git]	HAS_TAG	[Git, Tag]	1026	1	
14	[Git, Tag]	ON_COMMIT	[Git, Commit]	1026	1026	
15	[Repository, File, Git]	HAS_COMMITTER	[Committer, Git, Person]	370	1	
16	[TS, Function]	DEPENDS_ON	[TS, ExternalDeclaration]	280	47	
17	[TS, ExternalModule]	EXPORTS	[TS, ExternalDeclaration]	211	25	
18	[Type, TS, Union, ExternalType]	CONTAINS	[Type, TS, Primitive, ExternalType]	149	120	
19	[File, TS, Local, Module, Mark4ModuleWeaklyCon...	DEPENDS_ON	[TS, ExternalDeclaration]	148	1	
20	[Type, TS, Declared, ExternalType]	REFERENCES	[TS, ExternalDeclaration]	139	286	
21	[TS, Function]	DEPENDS_ON	[TS, ExternalModule]	129	47	
22	[Type, TS, Union, ExternalType]	CONTAINS	[Type, TS, Literal, ExternalType]	119	120	
23	[Type, TS, Object, ExternalType]	HAS_MEMBER	[Type, TS, ObjectMember, ExternalType]	98	38	
24	[Type, TS, Union, ExternalType]	CONTAINS	[Type, TS, Declared, ExternalType]	78	120	
25	[TS, Interface]	DECLARES	[TS, Property]	61	18	
26	[TS, Property]	OF_TYPE	[Type, TS, Union, ExternalType]	46	65	
27	[File, Directory]	CONTAINS	[File]	46	16	
28	[TS, Variable]	DEPENDS_ON	[TS, ExternalDeclaration]	44	24	
29	[Type, TS, Declared, ExternalType]	HAS_TYPE_ARGUMENT	[Type, TS, Declared, ExternalType]	43	286	

Graph Density

```
total_number_of_nodes (vertices): 90040
total_number_of_relationships (edges): 271602
-> total directed graph density: 3.3501697613249574e-05
-> total directed graph density in percent: 0.003350169761324957
```