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%).
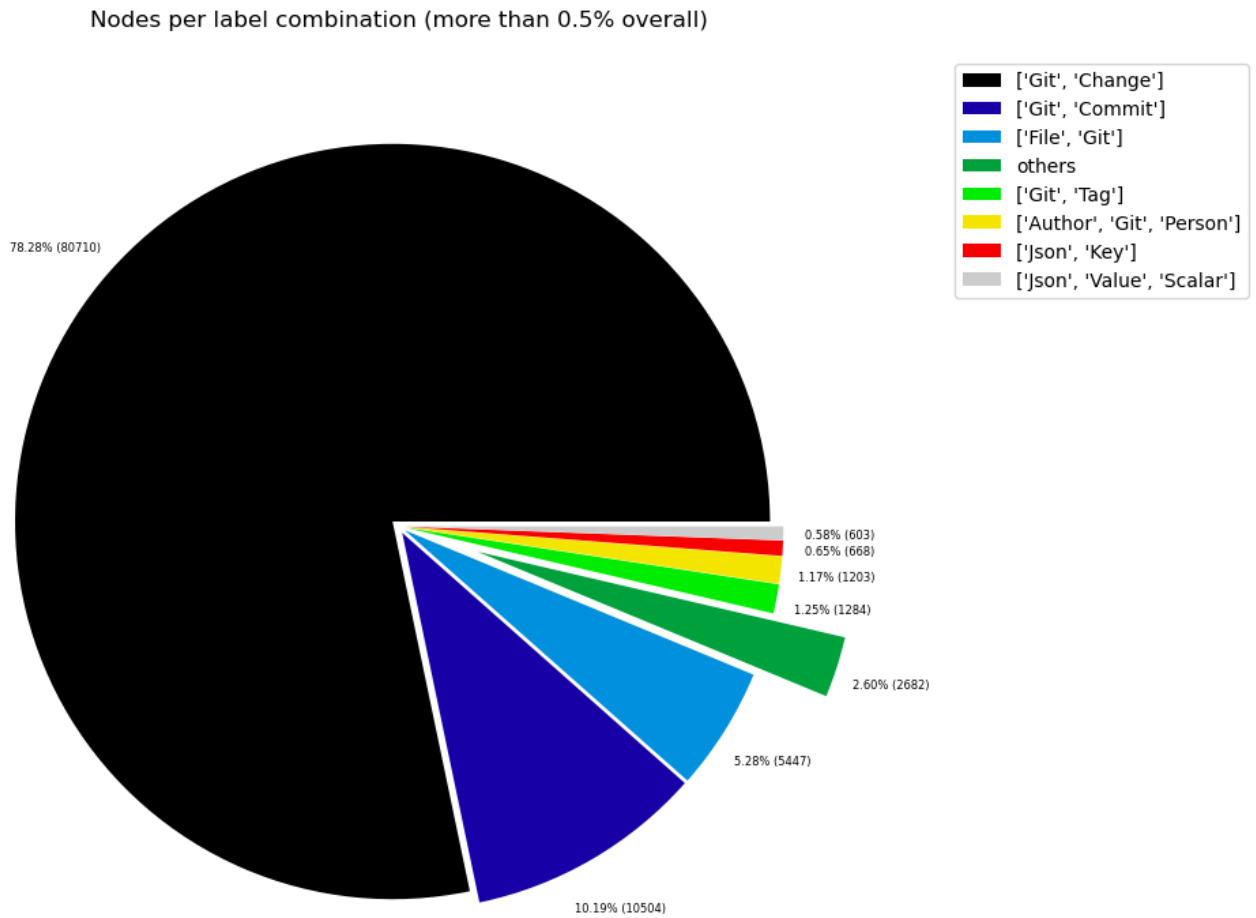
The whole table can be found in the CSV report `Node_label_combination_count`.

	nodeLabels	nodesWithThatLabels	nodesWithThatLabelsPercent
0	[Git, Change]	80710	78.282461
1	[Git, Commit]	10504	10.188068
2	[File, Git]	5447	5.283169
3	[Git, Tag]	1284	1.245381
4	[Author, Git, Person]	1203	1.166817
5	[Json, Key]	668	0.647908
6	[Json, Value, Scalar]	603	0.584863
7	[Committer, Git, Person]	371	0.359841
8	[NPM, Dependency]	330	0.320074
9	[Type, TS, Primitive]	291	0.282248
10	[Type, TS, Declared]	276	0.267699
11	[TS, ExternalDeclaration]	215	0.208533
12	[Type, TS, Literal]	136	0.131909
13	[Json, Value, Object]	133	0.129000
14	[Type, TS, Union]	119	0.115421
15	[Type, TS, ObjectMember]	101	0.097962
16	[NPM, Script]	91	0.088263
17	[TS, Property]	65	0.063045
18	[TS, Function]	47	0.045586
19	[Type, TS, FunctionParameter]	40	0.038797
20	[Type, Object, TS]	39	0.037827
21	[File, Directory]	34	0.032977
22	[Type, TS, Function]	34	0.032977
23	[TS, Parameter]	33	0.032007
24	[Git, Branch]	30	0.029098
25	[Package, File, Json, NPM]	29	0.028128
26	[TS, ExternalModule]	25	0.024248
27	[TS, Variable]	24	0.023278
28	[Value, TS, Literal]	20	0.019398
29	[jQAssistant, Rule, Concept]	19	0.018429

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>
```

Nodes per label combination (more than 0.5% overall)

Legend:
- ['Git', 'Change']
- ['Git', 'Commit']
- ['File', 'Git']
- others
- ['Git', 'Tag']
- ['Author', 'Git', 'Person']
- ['Json', 'Key']
- ['Json', 'Value', 'Scalar']

78.28% (80710)
0.58% (603)
0.65% (668)
1.17% (1203)
1.25% (1284)
2.60% (2682)
5.28% (5447)
10.19% (10504)

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	0.000970
1	[File, TS, Scan]	1	0.000970
2	[TS, Method]	1	0.000970
3	[Repository, File, Git]	1	0.000970
4	[TS, Constructor]	1	0.000970
5	[Value, TS, ObjectMember]	1	0.000970
6	[TS, Class]	1	0.000970
7	[TS, Enum]	2	0.001940
8	[Value, Object, TS]	3	0.002910
9	[Type, TS, Tuple]	3	0.002910
10	[Value, TS, Function]	4	0.003880
11	[TS, TypeParameter]	4	0.003880
12	[Value, TS, Complex]	5	0.004850
13	[NPM, Engine]	6	0.005820
14	[Project, TS]	6	0.005820
15	[File, Local]	6	0.005820
16	[Value, TS, Call]	6	0.005820
17	[Value, TS, Member]	6	0.005820
18	[File, TS, Local, Module]	6	0.005820
19	[Type, TS, TypeParameterReference]	6	0.005820
20	[TS, EnumMember]	8	0.007759
21	[Type, TS, NotIdentified]	11	0.010669
22	[Json, Value, Array]	12	0.011639
23	[Value, TS, Declared]	13	0.012609
24	[TS, TypeAlias]	16	0.015519
25	[File, Directory, Local]	16	0.015519
26	[TS, Interface]	17	0.016489
27	[Type, TS, Intersection]	17	0.016489
28	[jQAssistant, Rule, Concept]	19	0.018429
29	[Value, TS, Literal]	20	0.019398

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>
```

Nodes per label combination (less than 0.5% overall)

Legend:
- ['Committer', 'Git', 'Person']
- ['NPM', 'Dependency']
- ['Type', 'TS', 'Primitive']
- ['Type', 'TS', 'Declared']
- ['TS', 'ExternalDeclaration']
- ['Type', 'TS', 'Literal']
- ['Json', 'Value', 'Object']
- ['Type', 'TS', 'Union']
- ['Type', 'TS', 'ObjectMember']
- ['NPM', 'Script']
- others
- ['TS', 'Property']
- ['TS', 'Function']
- ['Type', 'TS', 'FunctionParameter']
- ['Type', 'Object', 'TS']
- ['Type', 'TS', 'Function']
- ['File', 'Directory']
- ['TS', 'Parameter']
- ['Git', 'Branch']
- ['Package', 'File', 'Json', 'NPM']
- ['TS', 'ExternalModule']
- ['TS', 'Variable']
- ['Value', 'TS', 'Literal']
- ['jQAssistant', 'Rule', 'Concept']
- ['Type', 'TS', 'Intersection']
- ['TS', 'Interface']
- ['File', 'Directory', 'Local']
- ['TS', 'TypeAlias']
- ['Value', 'TS', 'Declared']
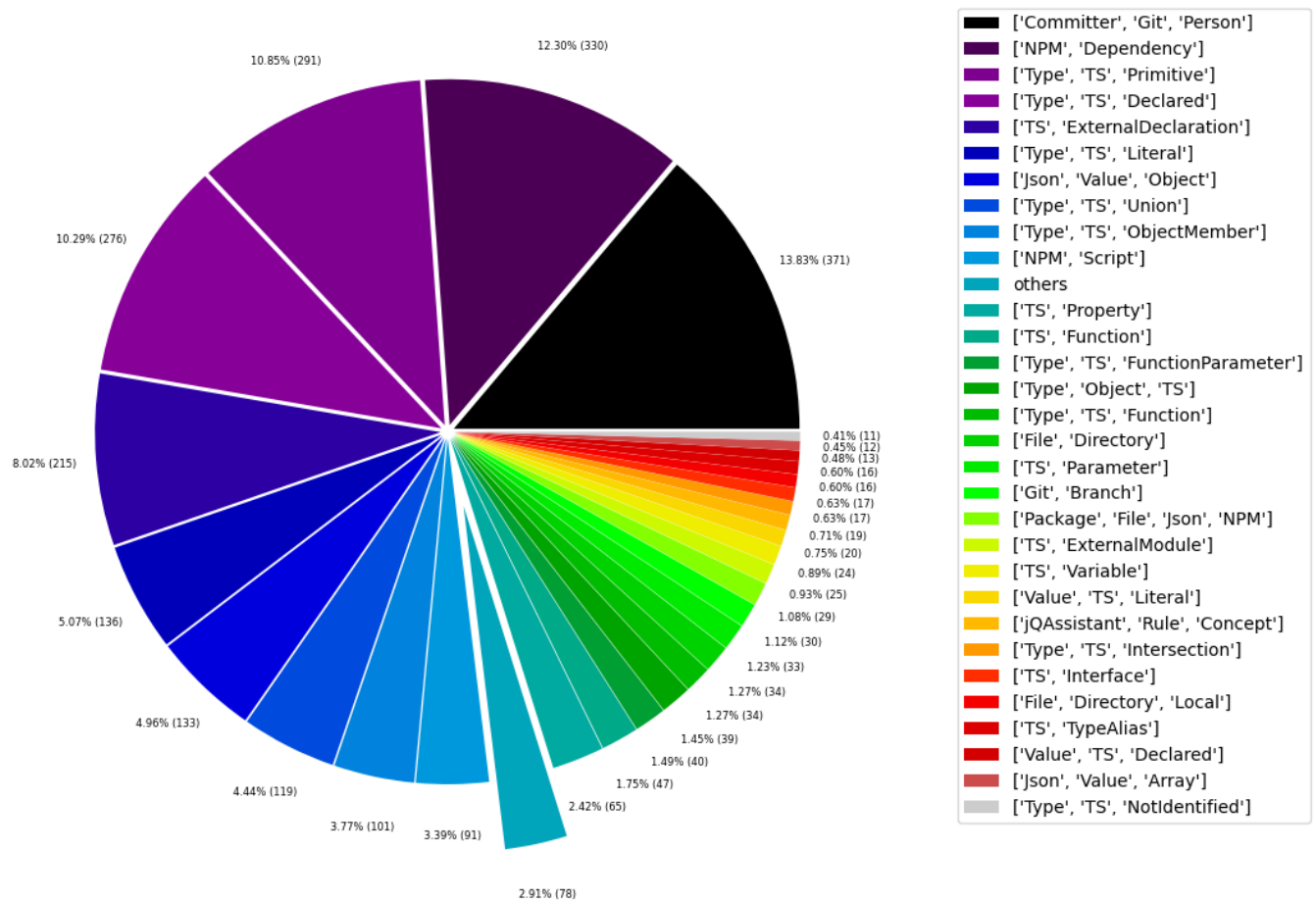- ['Json', 'Value', 'Array']
- ['Type', 'TS', 'NotIdentified']

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.
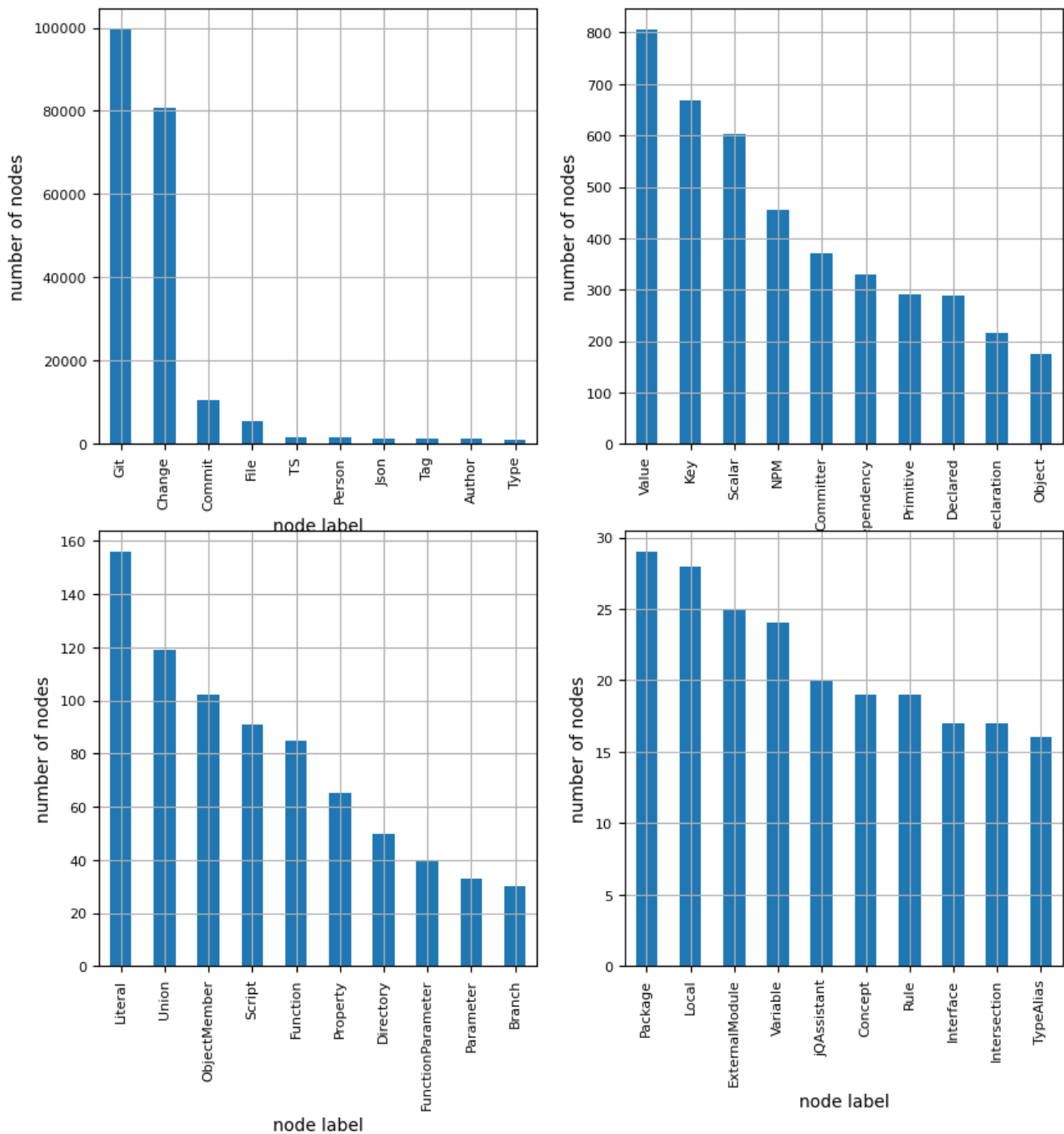
	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	99550	96.555805
1	Change	80710	78.282461
2	Commit	10504	10.188068
3	File	5540	5.373372
4	TS	1603	1.554786
5	Person	1574	1.526658
6	Json	1445	1.401538
7	Tag	1284	1.245381
8	Author	1203	1.166817
9	Type	1073	1.040727
10	Value	806	0.781758
11	Key	668	0.647908
12	Scalar	603	0.584863
13	NPM	456	0.442285
14	Committer	371	0.359841
15	Dependency	330	0.320074
16	Primitive	291	0.282248
17	Declared	289	0.280308
18	ExternalDeclaration	215	0.208533
19	Object	175	0.169736
20	Literal	156	0.151308
21	Union	119	0.115421
22	ObjectMember	102	0.098932
23	Script	91	0.088263
24	Function	85	0.082443
25	Property	65	0.063045
26	Directory	50	0.048496
27	FunctionParameter	40	0.038797
28	Parameter	33	0.032007
29	Branch	30	0.029098
30	Package	29	0.028128
31	Local	28	0.027158
32	ExternalModule	25	0.024248
33	Variable	24	0.023278
34	jQAssistant	20	0.019398
35	Concept	19	0.018429
36	Rule	19	0.018429
37	Interface	17	0.016489
38	Intersection	17	0.016489
39	TypeAlias	16	0.015519

Chart 1c - Highest node count by label

Shows the 40 labels with the highest number of nodes.

```
<Figure size 640x480 with 0 Axes>
```

Node count by label

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: 307932
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	80710	26.210332
1	MODIFIES	80710	26.210332
2	UPDATES	52786	17.142096
3	COMMITTED	21008	6.822285
4	CREATES	19561	6.352376
5	HAS_PARENT	11544	3.748880
6	DELETES	11534	3.745632
7	HAS_COMMIT	10504	3.411143
8	HAS_FILE	5447	1.768897
9	RENAMES	3171	1.029773
10	HAS_NEW_NAME	1718	0.557915
11	HAS_TAG	1284	0.416975
12	ON_COMMIT	1284	0.416975
13	HAS_AUTHOR	1203	0.390671
14	DEPENDS_ON	959	0.311432
15	HAS_KEY	668	0.216931
16	HAS_VALUE	668	0.216931
17	CONTAINS	594	0.192900
18	HAS_COMMITTER	371	0.120481
19	OF_TYPE	337	0.109440
20	EXPORTS	276	0.089630
21	REFERENCES	197	0.063975
22	DECLARES	186	0.060403
23	DECLARES_DEV_DEPENDENCY	169	0.054882
24	DECLARES_DEPENDENCY	161	0.052284
25	HAS_MEMBER	102	0.033124
26	HAS_TYPE_ARGUMENT	94	0.030526
27	DECLARES_SCRIPT	91	0.029552
28	RETURNS	82	0.026629
29	HAS_PARAMETER	73	0.023707

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>
```
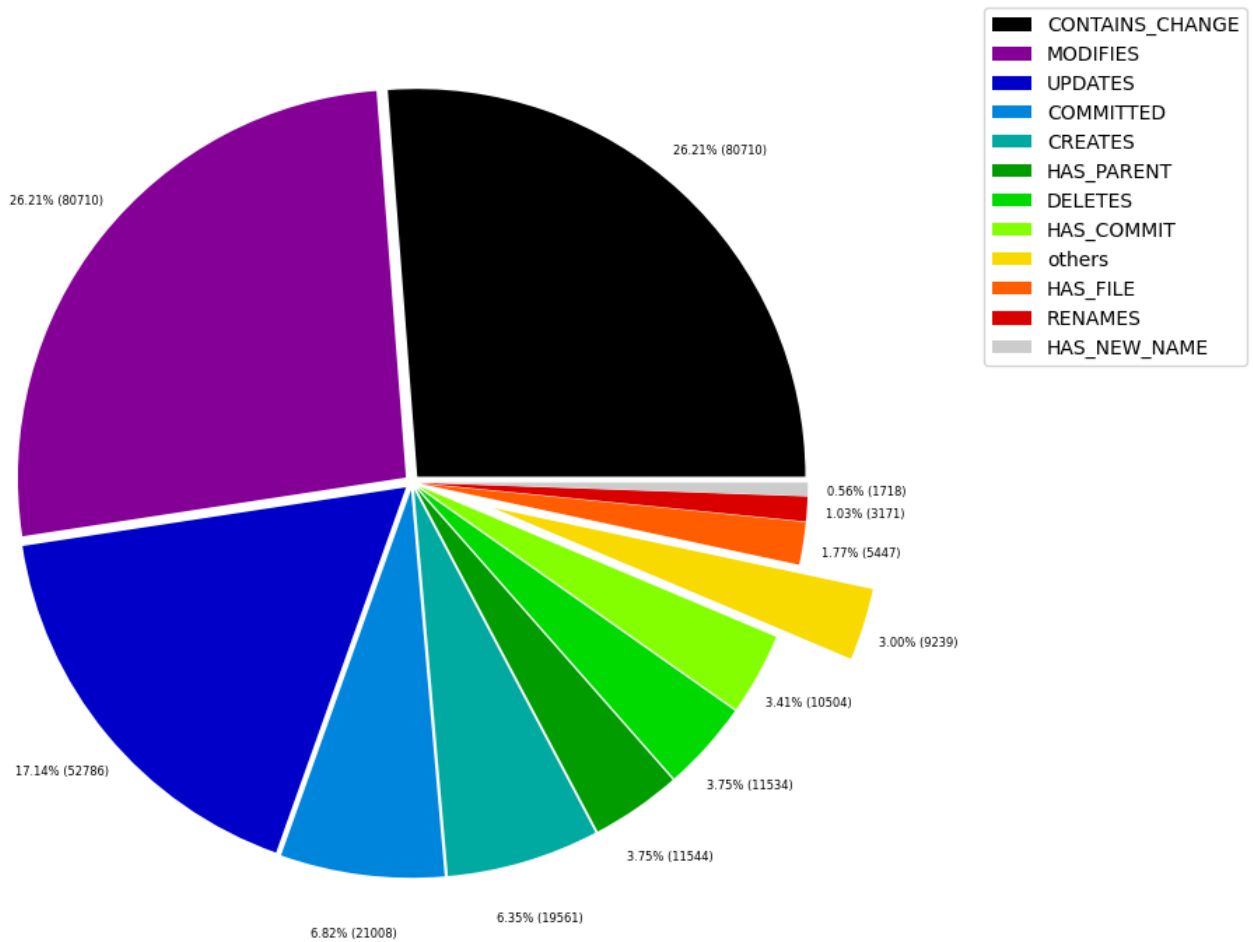
Relationship types (more than 0.5% overall)

Legend:
- CONTAINS_CHANGE
- MODIFIES
- UPDATES
- COMMITTED
- CREATES
- HAS_PARENT
- DELETES
- HAS_COMMIT
- others
- HAS_FILE
- RENAMES
- HAS_NEW_NAME

26.21% (80710)
26.21% (80710)
17.14% (52786)
6.82% (21008)
6.35% (19561)
3.75% (11544)
3.75% (11534)
3.41% (10504)
3.00% (9239)
1.77% (5447)
1.03% (3171)
0.56% (1718)

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.
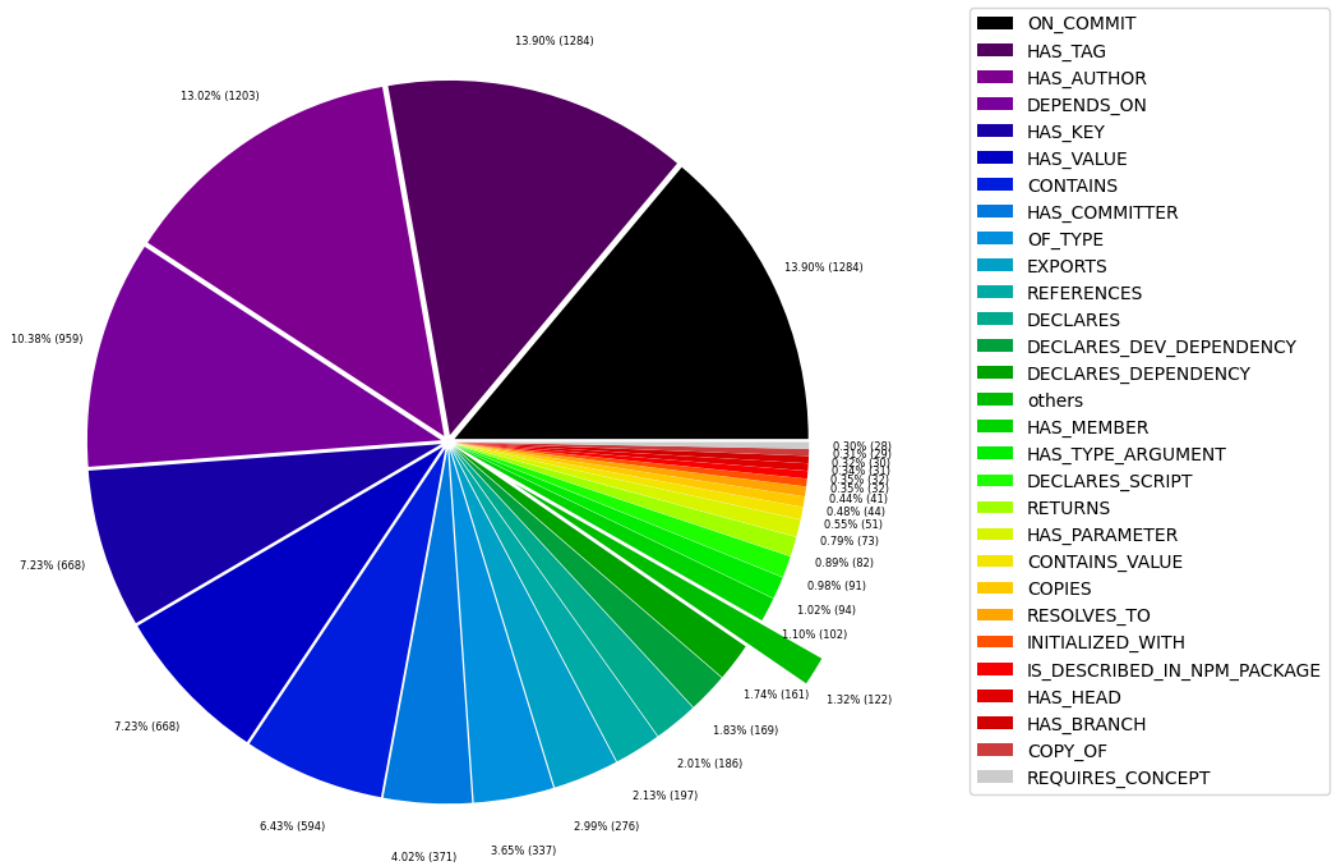
	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	PROVIDED_BY_NPM_DEPENDENCY	1	0.000325
1	IS_IMPLEMENTED_IN	2	0.000649
2	CONSTRAINED_BY	4	0.001299
3	REFERENCED_PROJECTS	5	0.001624
4	CONTAINS_PROJECT	6	0.001948
5	DECLARES_ENGINE	6	0.001948
6	EXTENDS	6	0.001948
7	HAS_ARGUMENT	6	0.001948
8	CALLS	6	0.001948
9	HAS_NPM_PACKAGE	6	0.001948
10	HAS_ROOT	6	0.001948
11	MEMBER	6	0.001948
12	PARENT	6	0.001948
13	HAS_CONFIG	6	0.001948
14	SIMILAR	6	0.001948
15	INCLUDES_CONCEPT	19	0.006170
16	USES	25	0.008119
17	REQUIRES_CONCEPT	28	0.009093
18	COPY_OF	29	0.009418
19	HAS_BRANCH	30	0.009742
20	HAS_HEAD	31	0.010067
21	IS_DESCRIBED_IN_NPM_PACKAGE	32	0.010392
22	INITIALIZED_WITH	32	0.010392
23	RESOLVES_TO	41	0.013315
24	COPIES	44	0.014289
25	CONTAINS_VALUE	51	0.016562
26	HAS_PARAMETER	73	0.023707
27	RETURNS	82	0.026629
28	DECLARES_SCRIPT	91	0.029552
29	HAS_TYPE_ARGUMENT	94	0.030526

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>
```

Legend:
- ON_COMMIT
- HAS_TAG
- HAS_AUTHOR
- DEPENDS_ON
- HAS_KEY
- HAS_VALUE
- CONTAINS
- HAS_COMMITTER
- OF_TYPE
- EXPORTS
- REFERENCES
- DECLARES
- DECLARES_DEV_DEPENDENCY
- DECLARES_DEPENDENCY
- others
- HAS_MEMBER
- HAS_TYPE_ARGUMENT
- DECLARES_SCRIPT
- RETURNS
- HAS_PARAMETER
- CONTAINS_VALUE
- COPIES
- RESOLVES_TO
- INITIALIZED_WITH
- IS_DESCRIBED_IN_NPM_PACKAGE
- HAS_HEAD
- HAS_BRANCH
- COPY_OF
- REQUIRES_CONCEPT

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, Change]	MODIFIES	[File, Git]	80710	80710	
1	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	80710	10504	
2	[Git, Change]	UPDATES	[File, Git]	52786	80710	
3	[Git, Change]	CREATES	[File, Git]	19561	80710	
4	[Git, Commit]	HAS_PARENT	[Git, Commit]	11544	10504	
5	[Git, Change]	DELETES	[File, Git]	11534	80710	
6	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	10504	1	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	10504	1203	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10504	371	
9	[Repository, File, Git]	HAS_FILE	[File, Git]	5447	1	
10	[Git, Change]	RENAMES	[File, Git]	3171	80710	
11	[File, Git]	HAS_NEW_NAME	[File, Git]	1718	5447	
12	[Repository, File, Git]	HAS_TAG	[Git, Tag]	1284	1	
13	[Git, Tag]	ON_COMMIT	[Git, Commit]	1284	1284	
14	[Repository, File, Git]	HAS_AUTHOR	[Author, Git, Person]	1203	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]	371	1	
18	[TS, Function]	DEPENDS_ON	[TS, ExternalDeclaration]	285	47	
19	[TS, ExternalModule]	EXPORTS	[TS, ExternalDeclaration]	215	25	
20	[File, TS, Local, Module, Mark4ModuleWeaklyCon...	DEPENDS_ON	[TS, ExternalDeclaration]	192	2	
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]	142	276	
25	[TS, Function]	DEPENDS_ON	[TS, ExternalModule]	131	47	
26	[Type, TS, Union]	CONTAINS	[Type, TS, Literal]	119	119	
27	[Json, Key]	HAS_VALUE	[Json, Value, Object]	104	668	
28	[Type, Object, TS]	HAS_MEMBER	[Type, TS, ObjectMember]	101	39	
29	[Package, File, Json, NPM]	DECLARES_SCRIPT	[NPM, Script]	91	29	

Graph Density

```
total_number_of_nodes (vertices): 103101
total_number_of_relationships (edges): 307932
-> total directed graph density: 2.8968985061318354e-05
-> total directed graph density in percent: 0.0028968985061318354
```