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]	249837	73.731072
1	[Java, ByteCode, Member, Method]	13204	3.896721
2	[Java, ByteCode, Parameter]	13029	3.845076
3	[Git, Commit]	12283	3.624918
4	[File, Git]	10736	3.168373
5	[Java, ByteCode, Bound, ParameterizedType]	7176	2.117757
6	[Java, ByteCode, Bound]	7118	2.100641
7	[Java, ByteCode, Member, Field]	3505	1.034384
8	[Java, ByteCode, Bound, WildcardType]	2908	0.858199
9	[Java, Value, ByteCode, Annotation]	2822	0.832819
10	[Xml, Element]	2144	0.632730
11	[Java, ByteCode, Member, Method, Constructor]	2050	0.604989
12	[Xml, Text]	1436	0.423788
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.323448
14	[Java, ByteCode, Member, Method, Lambda]	942	0.278000
15	[Type, File, Java, ByteCode, ResolvedDuplicate...	899	0.265310
16	[Type, File, Java, ByteCode, Class]	864	0.254981
17	[Type, File, Java, ByteCode, JavaType]	748	0.220747
18	[Java, Value, ByteCode, Primitive]	652	0.192416
19	[Java, ByteCode, GenericDeclaration, Member, M...	569	0.167921
20	[Json, Key]	560	0.165265
21	[Value, Json, Scalar]	544	0.160543
22	[Type, File, Java, ByteCode, ExternalType]	396	0.116866
23	[Author, Git, Person]	298	0.087945
24	[Value, Array]	267	0.078796
25	[Committer, Git, Person]	252	0.074369
26	[Type, File, Java, ByteCode, Class, GenericDec...	236	0.069648
27	[Value, Property]	202	0.059614
28	[Type, File, Java, ByteCode, Interface]	188	0.055482
29	[Java, Value, ByteCode, Class]	172	0.050760

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']
- ['Java', 'ByteCode', 'Member', 'Method']
- ['Java', 'ByteCode', 'Parameter']
- ['Git', 'Commit']
- others
- ['File', 'Git']
- ['Java', 'ByteCode', 'Bound', 'ParameterizedType']
- ['Java', 'ByteCode', 'Bound']
- ['Java', 'ByteCode', 'Member', 'Field']
- ['Java', 'ByteCode', 'Bound', 'WildcardType']
- ['Java', 'Value', 'ByteCode', 'Annotation']
- ['Xml', 'Element']
- ['Java', 'ByteCode', 'Member', 'Method', 'Constructor']

Pie slice labels:
- 73.73% (249837)
- 0.60% (2050)
- 0.63% (2144)
- 0.83% (2822)
- 0.86% (2908)
- 1.03% (3505)
- 2.10% (7118)
- 2.12% (7176)
- 3.17% (10736)
- 3.55% (12037)
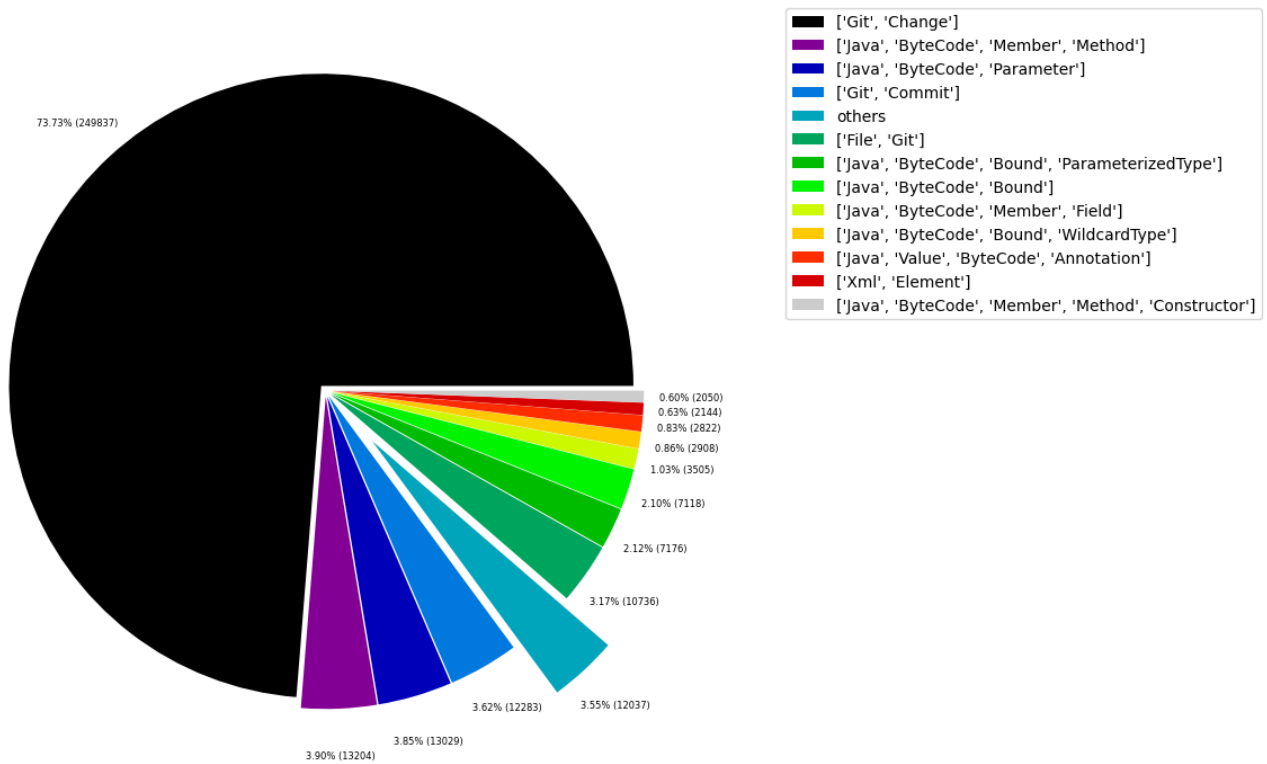- 3.62% (12283)
- 3.85% (13029)
- 3.90% (13204)

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	0.000295
1	[Repository, File, Git]	1	0.000295
2	[File, Json]	2	0.000590
3	[File]	3	0.000885
4	[Java, ByteCode, GenericDeclaration, Member, M...	4	0.001180
5	[Maven, Exclusion]	5	0.001476
6	[Value, Array, Json]	6	0.001771
7	[Type, File, Java, ByteCode, Void]	9	0.002656
8	[File, Maven, Xml, Pom, Document]	9	0.002656
9	[Java, ManifestSection]	9	0.002656
10	[File, Java, Manifest]	9	0.002656
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002656
12	[File, Java, ServiceLoader]	10	0.002951
13	[File, Java, Properties]	12	0.003541
14	[Maven, ExecutionGoal]	16	0.004722
15	[Maven, PluginExecution]	16	0.004722
16	[Xml, Attribute]	18	0.005312
17	[jQAssistant, Rule, Concept]	19	0.005607
18	[Maven, Plugin]	21	0.006197
19	[Maven, Configuration]	21	0.006197
20	[Type, File, Java, ByteCode, Enum]	28	0.008263
21	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008558
22	[Git, Branch]	34	0.010034
23	[Xml, Namespace]	36	0.010624
24	[Type, File, Java, ByteCode, Annotation]	43	0.012690
25	[File, Directory]	44	0.012985
26	[Type, File, Java, ByteCode, GenericDeclaratio...	85	0.025085
27	[Artifact, Maven]	92	0.027151
28	[Type, File, Java, ByteCode, ExternalAnnotatio...	93	0.027446
29	[Value, ManifestEntry]	120	0.035414

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:
- ['Xml', 'Text']
- ['Java', 'ByteCode', 'Bound', 'TypeVariable']
- ['Java', 'ByteCode', 'Member', 'Method', 'Lambda']
- ['Type', 'File', 'Java', 'ByteCode', 'ResolvedDuplicateType']
- ['Type', 'File', 'Java', 'ByteCode', 'Class']
- ['Type', 'File', 'Java', 'ByteCode', 'JavaType']
- ['Java', 'Value', 'ByteCode', 'Primitive']
- ['Java', 'ByteCode', 'GenericDeclaration', 'Member', 'Method']
- ['Json', 'Key']
- ['Value', 'Json', 'Scalar']
- ['Type', 'File', 'Java', 'ByteCode', 'ExternalType']
- others
- ['Author', 'Git', 'Person']
- ['Value', 'Array']
- ['Committer', 'Git', 'Person']
- ['Type', 'File', 'Java', 'ByteCode', 'Class', 'GenericDeclaration']
- ['Value', 'Property']
- ['Type', 'File', 'Java', 'ByteCode', 'Interface']
- ['Java', 'Value', 'ByteCode', 'Class']
- ['Git', 'Tag']
- ['Maven', 'Dependency']
- ['Java', 'ByteCode', 'Bound', 'GenericArrayType']
- ['Java', 'Value', 'ByteCode', 'Enum']
- ['Package', 'File', 'Directory', 'Java']
- ['Value', 'Json', 'Object']
- ['Value', 'ManifestEntry']
- ['Type', 'File', 'Java', 'ByteCode', 'ExternalAnnotation', 'ExternalType']
- ['Artifact', 'Maven']
- ['Type', 'File', 'Java', 'ByteCode', 'GenericDeclaration', 'Interface']
- ['File', 'Directory']
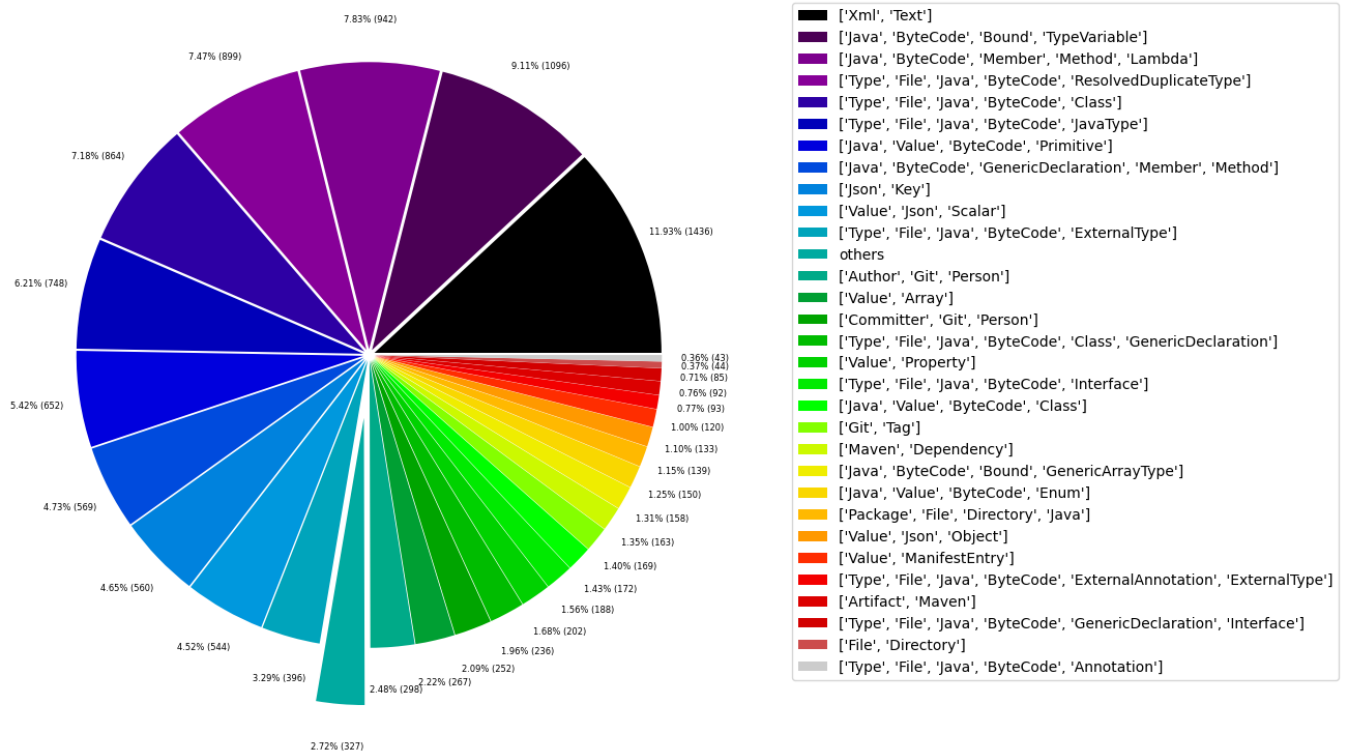- ['Type', 'File', 'Java', 'ByteCode', 'Annotation']

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.
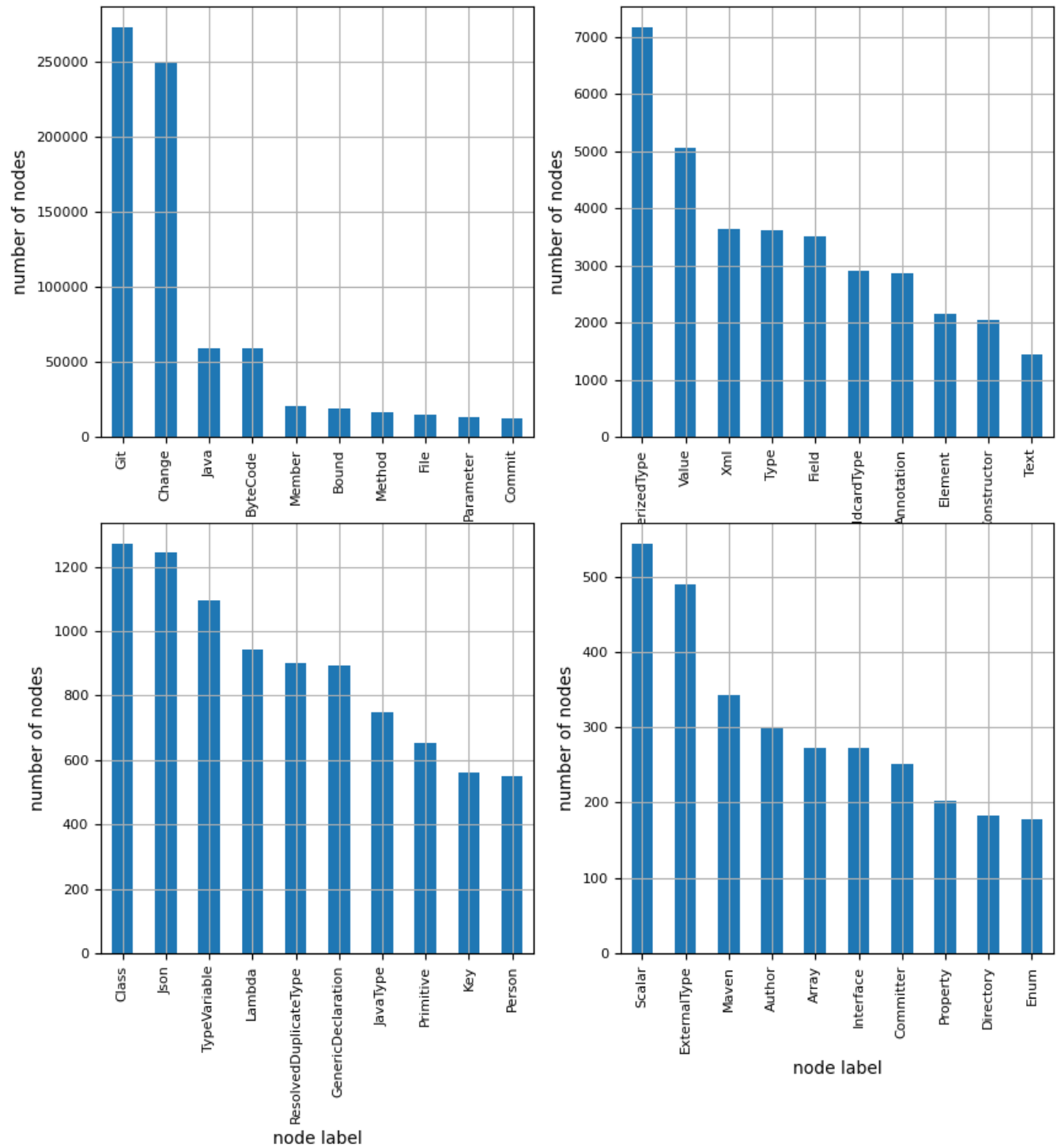
	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	273610	80.746881
1	Change	249837	73.731072
2	Java	59361	17.518423
3	ByteCode	59173	17.462941
4	Member	20274	5.983196
5	Bound	18456	5.446674
6	Method	16769	4.948812
7	File	14592	4.306343
8	Parameter	13029	3.845076
9	Commit	12283	3.624918
10	ParameterizedType	7176	2.117757
11	Value	5068	1.495651
12	Xml	3643	1.075110
13	Type	3618	1.067732
14	Field	3505	1.034384
15	WildcardType	2908	0.858199
16	Annotation	2865	0.845509
17	Element	2144	0.632730
18	Constructor	2054	0.606170
19	Text	1436	0.423788
20	Class	1272	0.375388
21	Json	1245	0.367420
22	TypeVariable	1096	0.323448
23	Lambda	942	0.278000
24	ResolvedDuplicateType	899	0.265310
25	GenericDeclaration	894	0.263834
26	JavaType	748	0.220747
27	Primitive	652	0.192416
28	Key	560	0.165265
29	Person	550	0.162314
30	Scalar	544	0.160543
31	ExternalType	489	0.144312
32	Maven	343	0.101225
33	Author	298	0.087945
34	Array	273	0.080567
35	Interface	273	0.080567
36	Committer	252	0.074369
37	Property	202	0.059614
38	Directory	183	0.054006
39	Enum	178	0.052531

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: 1051898
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	249837	23.751067
1	MODIFIES	249837	23.751067
2	UPDATES	162807	15.477451
3	CREATES	60062	5.709869
4	DELETES	37341	3.549869
5	INVOKES	35700	3.393865
6	COMMITTED	24566	2.335398
7	DEPENDS_ON	21931	2.084898
8	OF_TYPE	21394	2.033847
9	DECLARES	20719	1.969678
10	OF_RAW_TYPE	17093	1.624967
11	HAS_PARENT	14918	1.418198
12	HAS	14125	1.342811
13	RETURNS	12578	1.195743
14	HAS_COMMIT	12283	1.167699
15	HAS_FILE	10736	1.020631
16	RENAMES	10373	0.986122
17	READS	9164	0.871187
18	HAS_ACTUAL_TYPE_ARGUMENT	8288	0.787909
19	HAS_NEW_NAME	6208	0.590171
20	OF_GENERIC_TYPE	5906	0.561461
21	RESOLVES_TO	5198	0.494154
22	SIMILAR	3977	0.378078
23	WRITES	3818	0.362963
24	CONTAINS	3814	0.362583
25	RETURNS_GENERIC	3545	0.337010
26	ANNOTATED_BY	2810	0.267136
27	REQUIRES	2174	0.206674
28	HAS_FIRST_CHILD	2144	0.203822
29	HAS_LAST_CHILD	2144	0.203822

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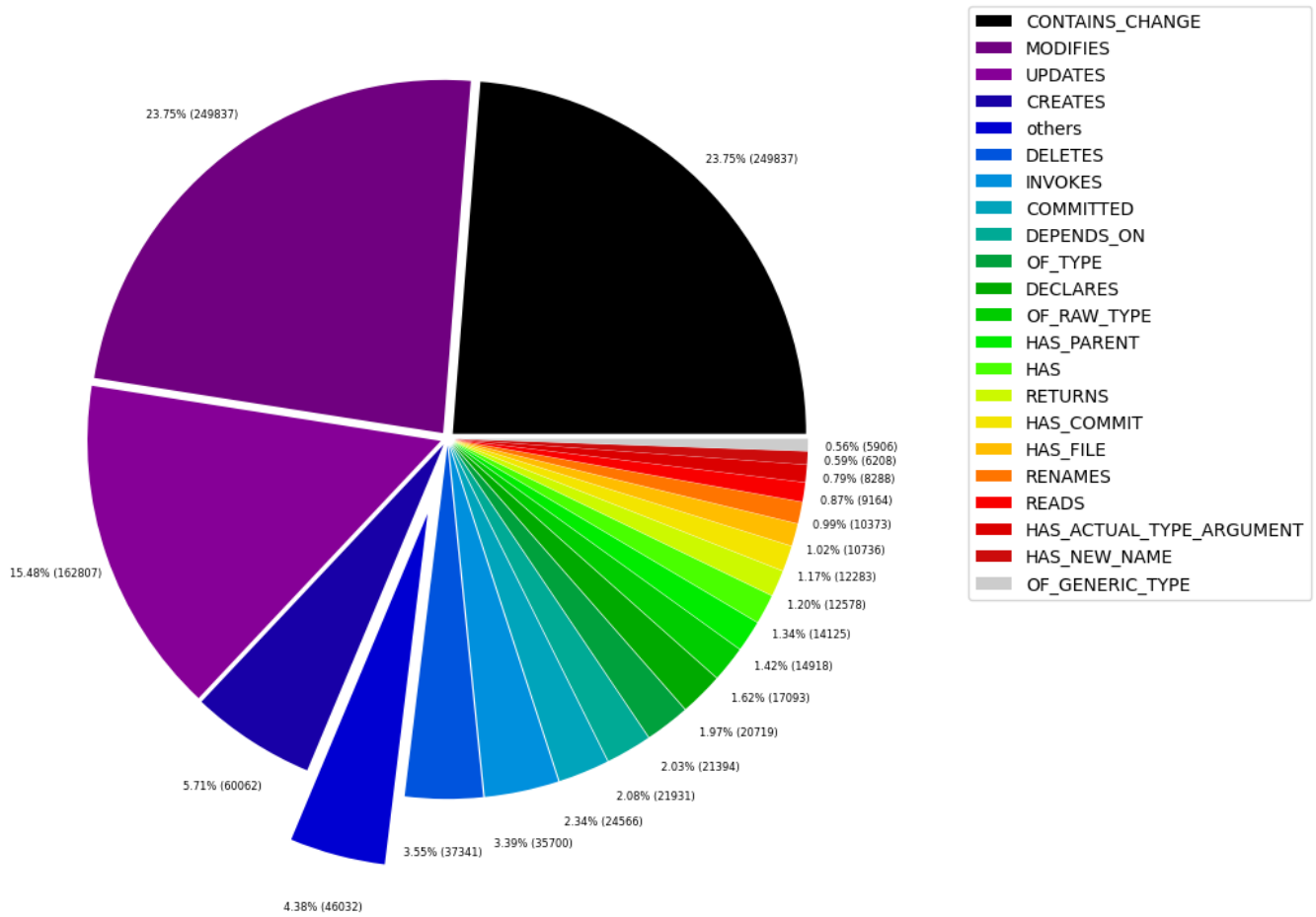
Relationship types (more than 0.5% overall)

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.
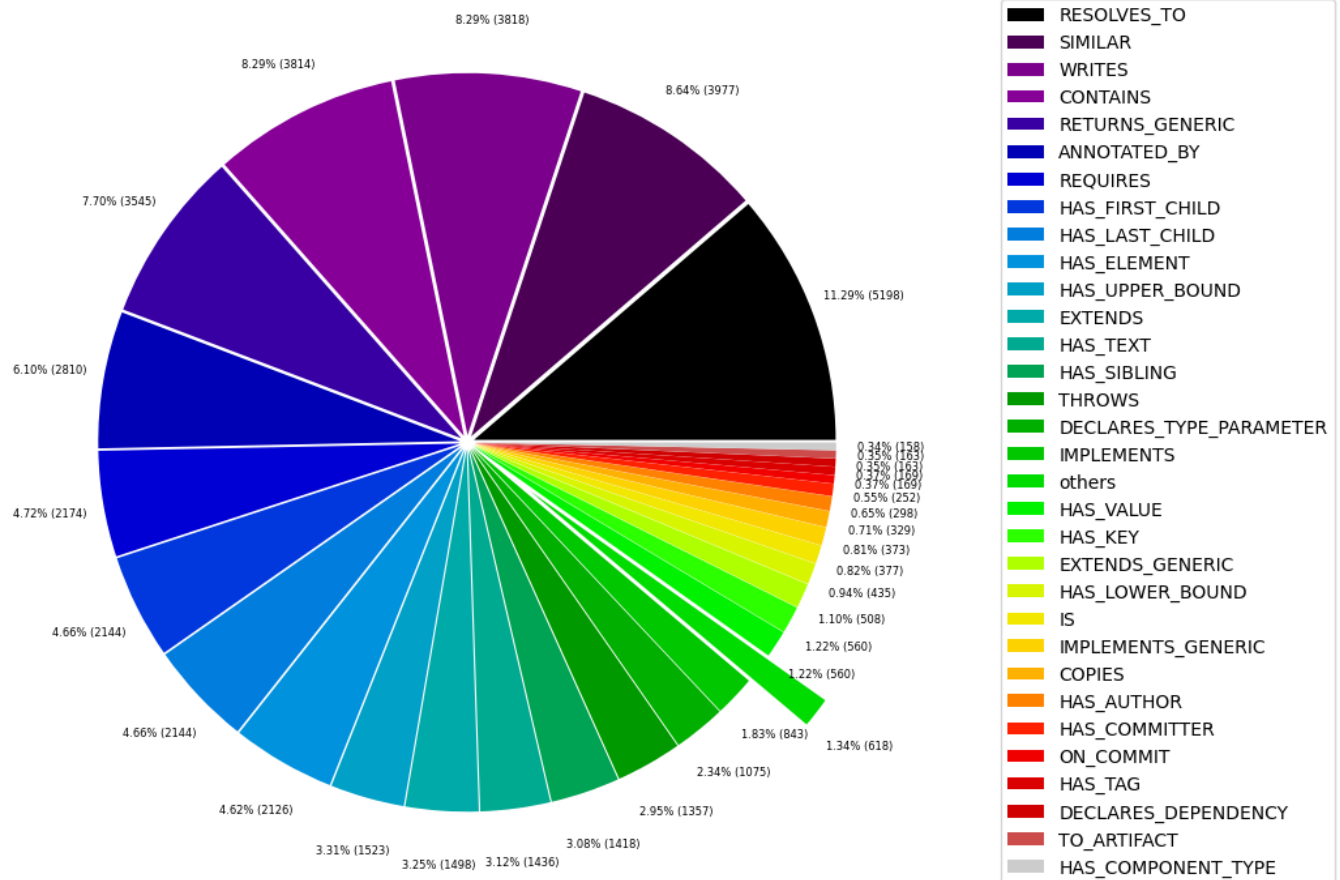
	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	HAS_PROPERTY	1	0.000095
1	THROWS_GENERIC	5	0.000475
2	EXCLUDES	5	0.000475
3	DESCRIBES	9	0.000856
4	HAS_ROOT_ELEMENT	11	0.001046
5	HAS_GOAL	16	0.001521
6	HAS_EXECUTION	16	0.001521
7	OF_NAMESPACE	18	0.001711
8	HAS_ATTRIBUTE	18	0.001711
9	INCLUDES_CONCEPT	19	0.001806
10	USES_PLUGIN	21	0.001996
11	REQUIRES_TYPE_PARAMETER	21	0.001996
12	IS_ARTIFACT	21	0.001996
13	HAS_CONFIGURATION	21	0.001996
14	REQUIRES_CONCEPT	28	0.002662
15	HAS_BRANCH	34	0.003232
16	HAS_HEAD	35	0.003327
17	DECLARES_NAMESPACE	36	0.003422
18	HAS_DEFAULT	36	0.003422
19	CONTAINS_VALUE	121	0.011503
20	COPY_OF	126	0.011978
21	HAS_COMPONENT_TYPE	158	0.015020
22	TO_ARTIFACT	163	0.015496
23	DECLARES_DEPENDENCY	163	0.015496
24	ON_COMMIT	169	0.016066
25	HAS_TAG	169	0.016066
26	HAS_COMMITTER	252	0.023957
27	HAS_AUTHOR	298	0.028330
28	COPIES	329	0.031277
29	IMPLEMENTS_GENERIC	373	0.035460

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>
```

Relationship types (less than 0.5% overall)

Legend:
- RESOLVES_TO
- SIMILAR
- WRITES
- CONTAINS
- RETURNS_GENERIC
- ANNOTATED_BY
- REQUIRES
- HAS_FIRST_CHILD
- HAS_LAST_CHILD
- HAS_ELEMENT
- HAS_UPPER_BOUND
- EXTENDS
- HAS_TEXT
- HAS_SIBLING
- THROWS
- DECLARES_TYPE_PARAMETER
- IMPLEMENTS
- others
- HAS_VALUE
- HAS_KEY
- EXTENDS_GENERIC
- HAS_LOWER_BOUND
- IS
- IMPLEMENTS_GENERIC
- COPIES
- HAS_AUTHOR
- HAS_COMMITTER
- ON_COMMIT
- HAS_TAG
- DECLARES_DEPENDENCY
- TO_ARTIFACT
- HAS_COMPONENT_TYPE

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	numberOfNodesWithS
0	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	249837	12283	
1	[Git, Change]	MODIFIES	[File, Git]	249837	249837	
2	[Git, Change]	UPDATES	[File, Git]	162807	249837	
3	[Git, Change]	CREATES	[File, Git]	60062	249837	
4	[Git, Change]	DELETES	[File, Git]	37341	249837	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21813	13110	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	14909	12283	
7	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	12283	1	
8	[Author, Git, Person]	COMMITTED	[Git, Commit]	12283	298	
9	[Committer, Git, Person]	COMMITTED	[Git, Commit]	12283	252	
10	[Repository, File, Git]	HAS_FILE	[File, Git]	10736	1	
11	[Git, Change]	RENAMES	[File, Git]	10373	249837	
12	[Java, ByteCode, Member, Method]	HAS	[Java, ByteCode, Parameter]	8360	13110	
13	[Java, ByteCode, Member, Method]	READS	[Java, ByteCode, Member, Field]	8218	13110	
14	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6371	13029	
15	[File, Git]	HAS_NEW_NAME	[File, Git]	6208	10736	
16	[Type, File, Java, ByteCode, Class]	DECLARES	[Java, ByteCode, Member, Method]	5109	814	
17	[Type, File, Java, ByteCode, Class]	DEPENDS_ON	[Type, File, Java, ByteCode, JavaType]	4480	814	
18	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3570	7118	
19	[Java, ByteCode, Bound, ParameterizedType]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3053	7176	
20	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, WildcardType]	2908	7176	
21	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound, ParameterizedType]	2657	13029	
22	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, TypeVariable]	2454	7176	
23	[Java, ByteCode, Member, Method, Constructor]	WRITES	[Java, ByteCode, Member, Field]	2438	2050	
24	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalAnnotatio...]	2341	2822	
25	[Type, File, Java, ByteCode, Class]	DECLARES	[Java, ByteCode, Member, Field]	2203	814	
26	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
27	[Java, ByteCode, Member, Method, Constructor]	INVOKES	[Java, ByteCode, Member, Method, Constructor]	2117	2050	
28	[Type, File, Java, ByteCode, Class, GenericDec...]	DECLARES	[Java, ByteCode, Member, Method]	2116	231	
29	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2091	13110	

Graph Density

```
total_number_of_nodes (vertices): 338849
total_number_of_relationships (edges): 1051898
-> total directed graph density: 9.16141368990192e-06
-> total directed graph density in percent: 0.0009161413689901921
```