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]	259035	74.296146
1	[Java, ByteCode, Member, Method]	13284	3.810103
2	[Java, ByteCode, Parameter]	13029	3.736964
3	[Git, Commit]	12706	3.644322
4	[File, Git]	10810	3.100513
5	[Java, ByteCode, ParameterizedType, Bound]	7176	2.058213
6	[Java, ByteCode, Bound]	7118	2.041577
7	[Java, ByteCode, Member, Field]	3511	1.007021
8	[Java, ByteCode, Bound, WildcardType]	2908	0.834070
9	[Java, Value, ByteCode, Annotation]	2822	0.809403
10	[Xml, Element]	2144	0.614940
11	[Java, ByteCode, Member, Constructor, Method]	2068	0.593142
12	[Xml, Text]	1436	0.411872
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.314354
14	[Java, ByteCode, Member, Method, Lambda]	942	0.270183
15	[Type, File, Java, ByteCode, ResolvedDuplicate...	877	0.251540
16	[Type, File, Java, ByteCode, Class]	810	0.232323
17	[Java, Value, ByteCode, Primitive]	652	0.187006
18	[Type, File, Java, ByteCode, JavaType]	644	0.184711
19	[Java, ByteCode, Member, Method, GenericDeclar...	569	0.163200
20	[Json, Key]	560	0.160619
21	[Value, Json, Scalar]	544	0.156030
22	[Type, File, Java, ByteCode, ExternalType]	380	0.108991
23	[Author, Git, Person]	298	0.085472
24	[Value, Array]	267	0.076581
25	[Committer, Git, Person]	252	0.072278
26	[Type, File, Java, ByteCode, Class, GenericDec...	236	0.067689
27	[Value, Property]	202	0.057937
28	[Type, File, Java, ByteCode, Interface]	188	0.053922
29	[Java, Value, ByteCode, Class]	172	0.049333

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>
```

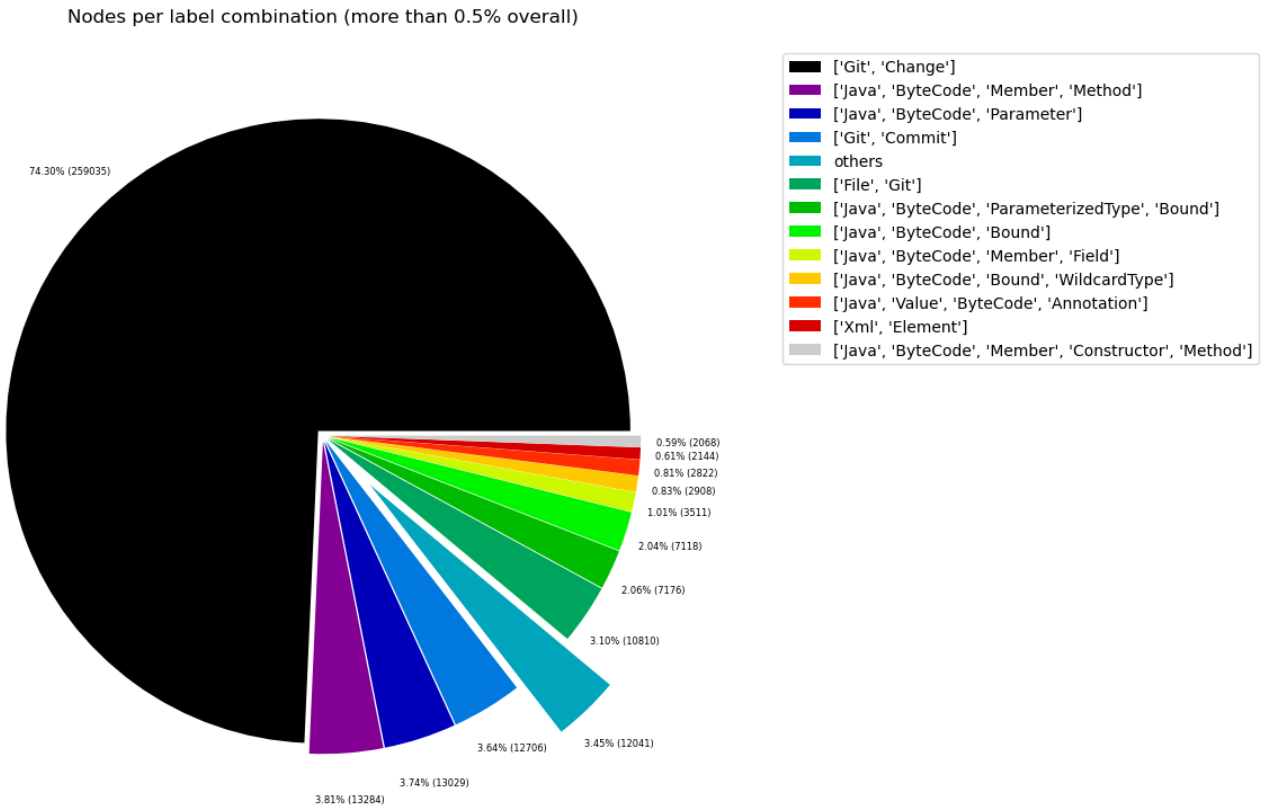


Table 1b - Lowest node count by label combination

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	0.000287
1	[Repository, File, Git]	1	0.000287
2	[File, Json]	2	0.000574
3	[File]	3	0.000860
4	[Java, ByteCode, Member, Constructor, Method, ...	4	0.001147
5	[Maven, Exclusion]	5	0.001434
6	[Value, Json, Array]	6	0.001721
7	[File, Maven, Xml, Pom, Document]	9	0.002581
8	[Type, File, Java, ByteCode, Void]	9	0.002581
9	[Java, ManifestSection]	9	0.002581
10	[File, Java, Manifest]	9	0.002581
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002581
12	[File, Java, ServiceLoader]	10	0.002868
13	[File, Java, Properties]	12	0.003442
14	[Maven, ExecutionGoal]	16	0.004589
15	[Maven, PluginExecution]	16	0.004589
16	[Type, File, Java, ByteCode, Throwable, Extern...	16	0.004589
17	[Xml, Attribute]	18	0.005163
18	[jQAssistant, Rule, Concept]	19	0.005450
19	[Maven, Plugin]	21	0.006023
20	[Maven, Configuration]	21	0.006023
21	[Type, File, Java, ByteCode, Throwable, Resolv...	22	0.006310
22	[Type, File, Java, ByteCode, Enum]	28	0.008031
23	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008318
24	[Xml, Namespace]	36	0.010325
25	[Git, Branch]	38	0.010899
26	[Type, File, Java, ByteCode, Annotation]	43	0.012333
27	[File, Directory]	44	0.012620
28	[Type, File, Java, ByteCode, Class, Throwable]	54	0.015488
29	[Type, File, Java, ByteCode, GenericDeclaratio...	85	0.024380

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>
```

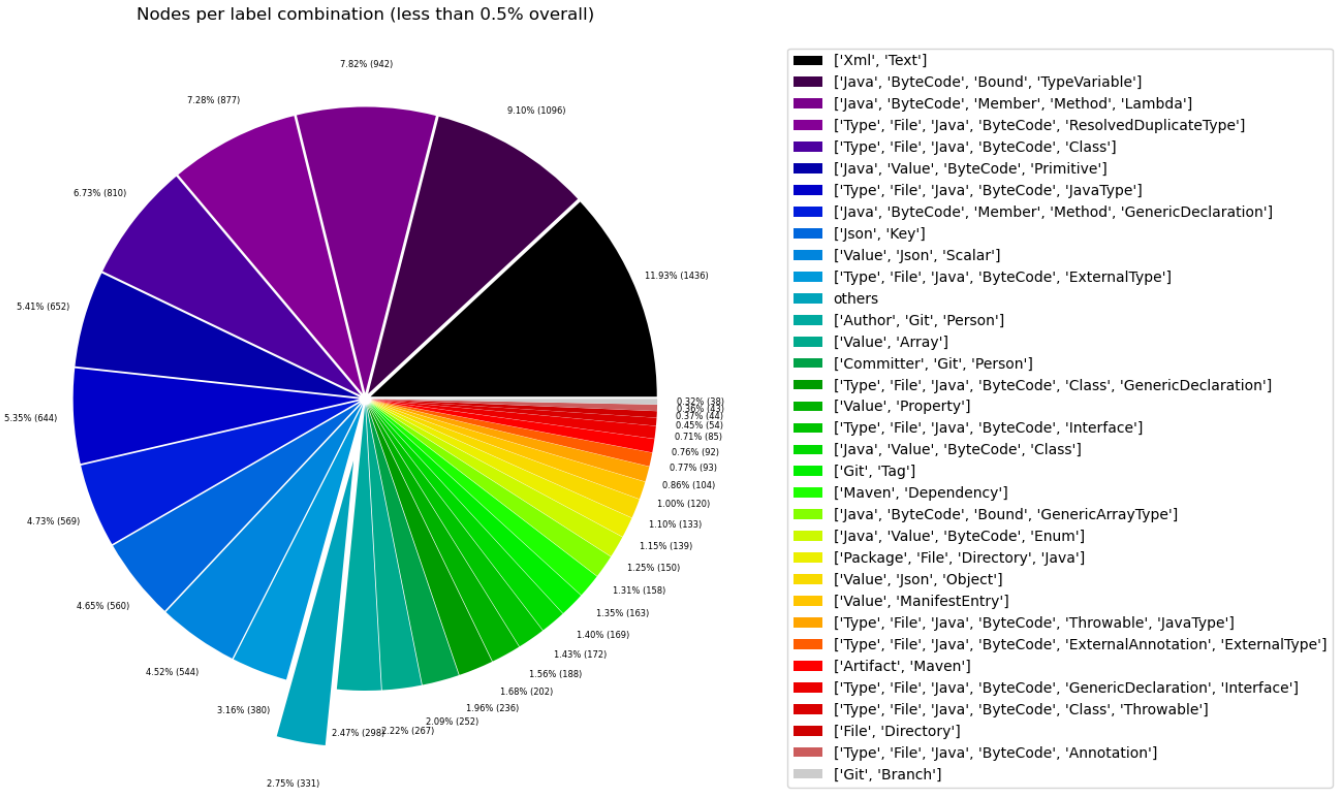


Table 1c - Highest node count by single label

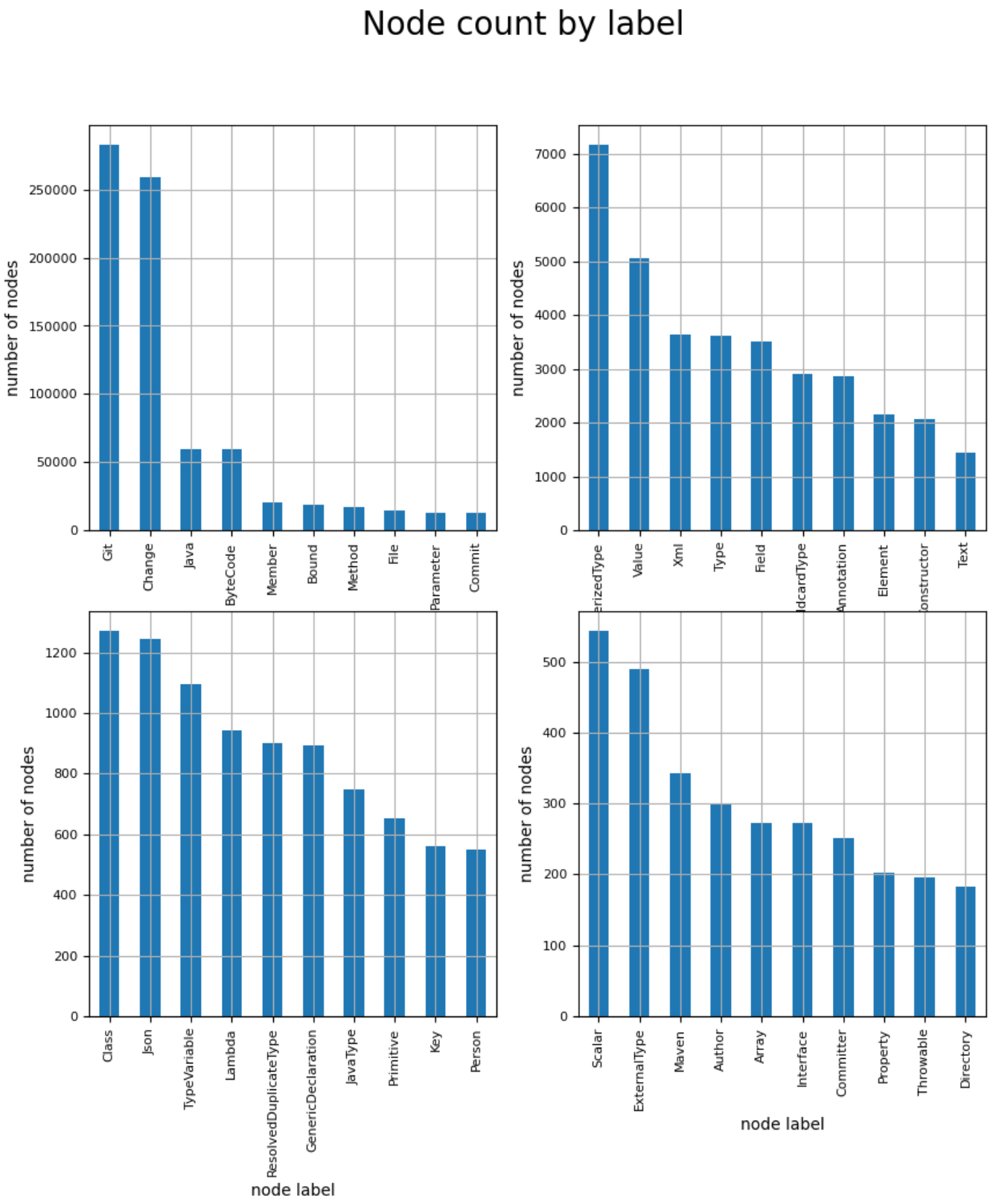
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	283309	81.258389
1	Change	259035	74.296146
2	Java	59465	17.055689
3	ByteCode	59277	17.001767
4	Member	20378	5.844797
5	Bound	18456	5.293531
6	Method	16867	4.837775
7	File	14666	4.206487
8	Parameter	13029	3.736964
9	Commit	12706	3.644322
10	ParameterizedType	7176	2.058213
11	Value	5068	1.453598
12	Xml	3643	1.044881
13	Type	3618	1.037711
14	Field	3511	1.007021
15	WildcardType	2908	0.834070
16	Annotation	2865	0.821736
17	Element	2144	0.614940
18	Constructor	2072	0.594289
19	Text	1436	0.411872
20	Class	1272	0.364834
21	Json	1245	0.357090
22	TypeVariable	1096	0.314354
23	Lambda	942	0.270183
24	ResolvedDuplicateType	899	0.257850
25	GenericDeclaration	894	0.256416
26	JavaType	748	0.214541
27	Primitive	652	0.187006
28	Key	560	0.160619
29	Person	550	0.157750
30	Scalar	544	0.156030
31	ExternalType	489	0.140254
32	Maven	343	0.098379
33	Author	298	0.085472
34	Array	273	0.078302
35	Interface	273	0.078302
36	Committer	252	0.072278
37	Property	202	0.057937
38	Throwable	196	0.056217
39	Directory	183	0.052488

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: 1082640
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	259035	23.926236
1	MODIFIES	259035	23.926236
2	UPDATES	170412	15.740412
3	CREATES	61193	5.652202
4	DELETES	37976	3.507722
5	INVOKES	36005	3.325667
6	COMMITTED	25412	2.347225
7	DEPENDS_ON	21931	2.025696
8	OF_TYPE	21394	1.976095
9	DECLARES	20823	1.923354
10	OF_RAW_TYPE	17093	1.578826
11	HAS_PARENT	15387	1.421248
12	HAS	14125	1.304681
13	HAS_COMMIT	12706	1.173613
14	RETURNS	12578	1.161790
15	HAS_FILE	10810	0.998485
16	RENAMES	10546	0.974100
17	READS	9166	0.846634
18	HAS_ACTUAL_TYPE_ARGUMENT	8288	0.765536
19	HAS_NEW_NAME	6237	0.576092
20	OF_GENERIC_TYPE	5906	0.545518
21	RESOLVES_TO	5253	0.485203
22	SIMILAR	3977	0.367343
23	WRITES	3818	0.352656
24	CONTAINS	3814	0.352287
25	RETURNS_GENERIC	3545	0.327440
26	ANNOTATED_BY	2810	0.259551
27	REQUIRES	2174	0.200805
28	HAS_FIRST_CHILD	2144	0.198034
29	HAS_LAST_CHILD	2144	0.198034

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>
```

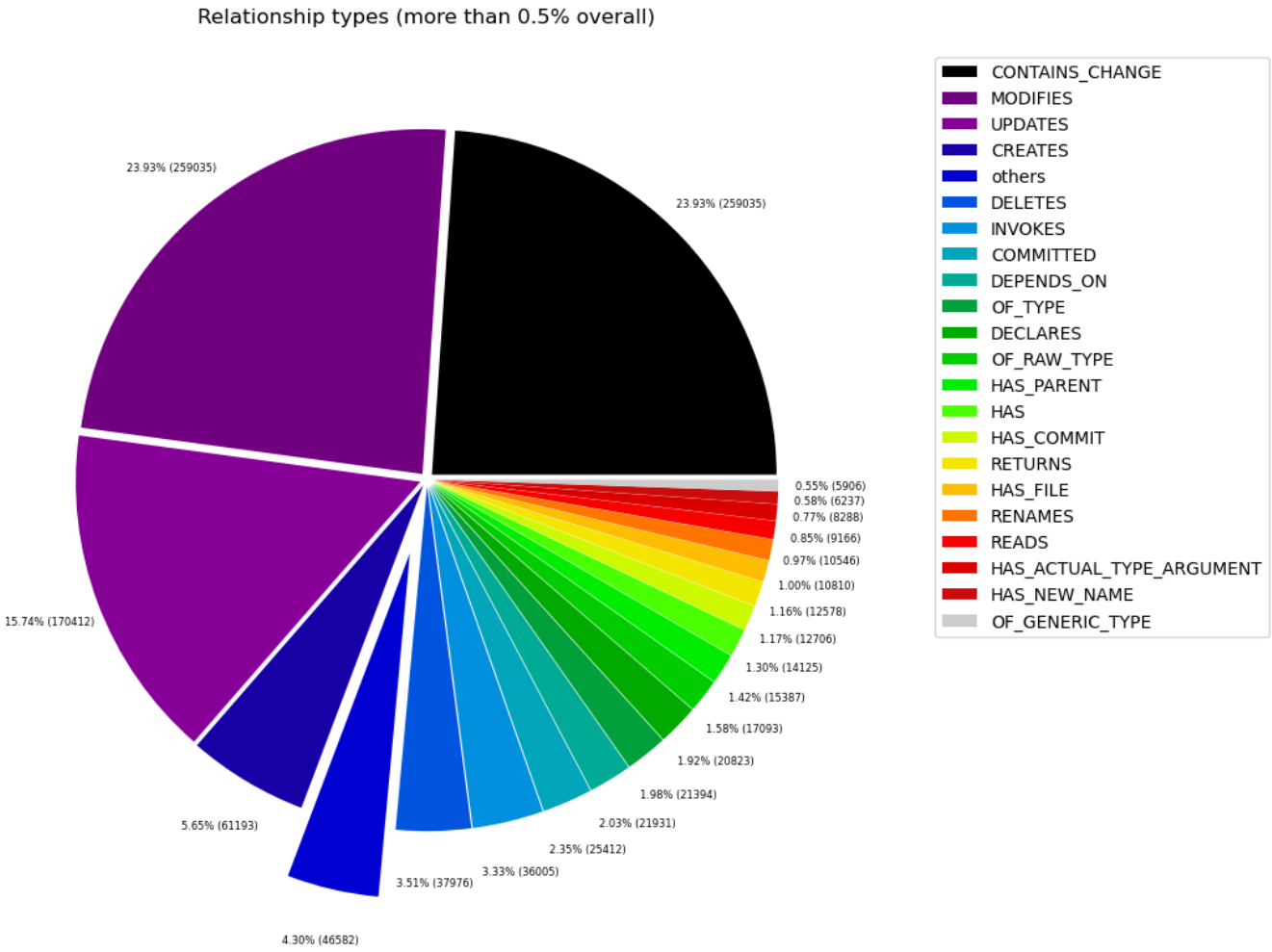


Table 2b - Lowest relationship count by type

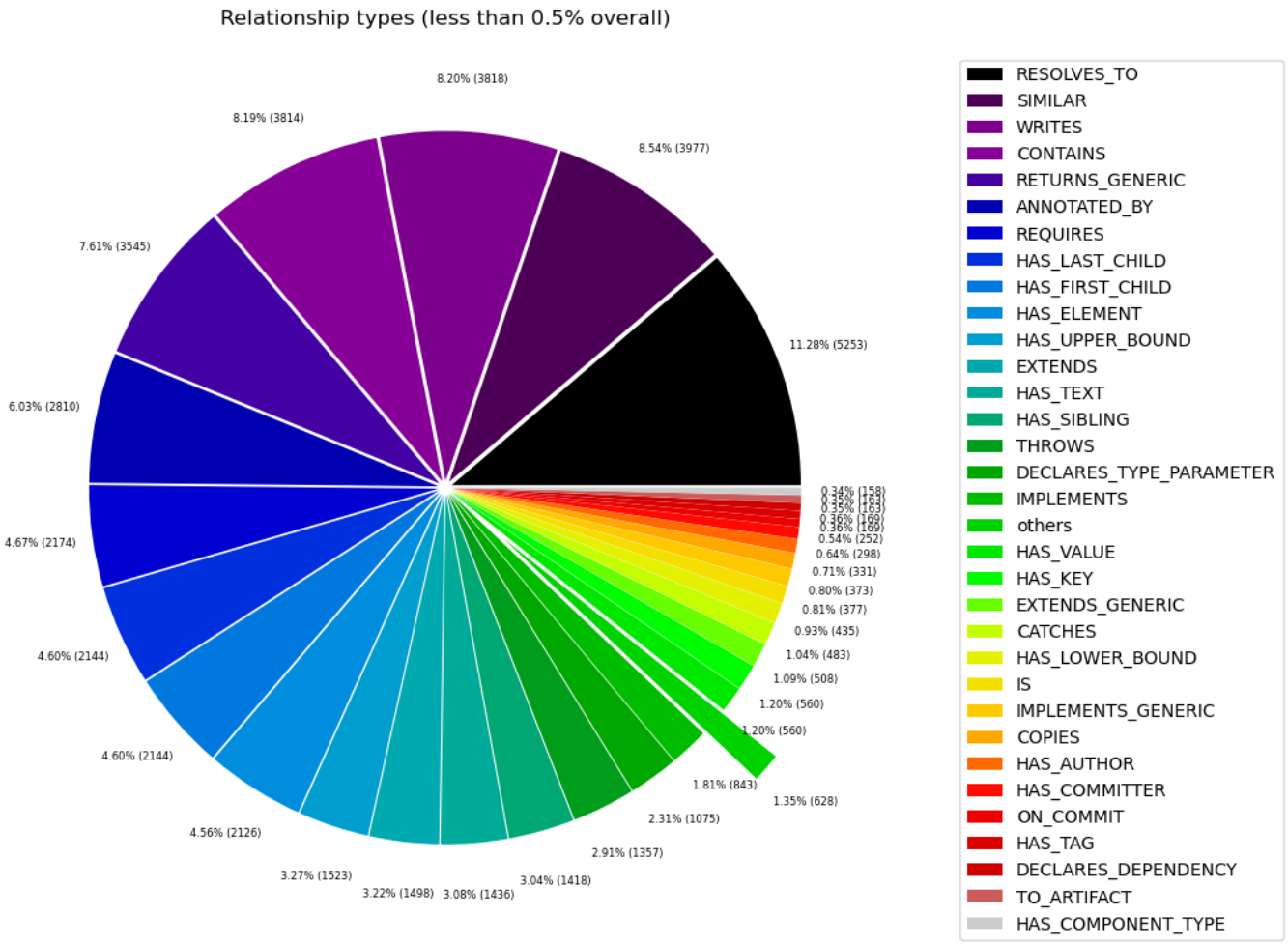
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	HAS_PROPERTY	1	0.000092
1	THROWS_GENERIC	5	0.000462
2	EXCLUDES	5	0.000462
3	DESCRIBES	9	0.000831
4	HAS_ROOT_ELEMENT	13	0.001201
5	HAS_GOAL	16	0.001478
6	HAS_EXECUTION	16	0.001478
7	OF_NAMESPACE	18	0.001663
8	HAS_ATTRIBUTE	18	0.001663
9	INCLUDES_CONCEPT	19	0.001755
10	USES_PLUGIN	21	0.001940
11	IS_ARTIFACT	21	0.001940
12	HAS_CONFIGURATION	21	0.001940
13	REQUIRES_TYPE_PARAMETER	21	0.001940
14	REQUIRES_CONCEPT	28	0.002586
15	HAS_DEFAULT	36	0.003325
16	DECLARES_NAMESPACE	36	0.003325
17	HAS_BRANCH	38	0.003510
18	HAS_HEAD	39	0.003602
19	CONTAINS_VALUE	121	0.011176
20	COPY_OF	126	0.011638
21	HAS_COMPONENT_TYPE	158	0.014594
22	DECLARES_DEPENDENCY	163	0.015056
23	TO_ARTIFACT	163	0.015056
24	ON_COMMIT	169	0.015610
25	HAS_TAG	169	0.015610
26	HAS_COMMITTER	252	0.023276
27	HAS_AUTHOR	298	0.027525
28	COPIES	331	0.030573
29	IMPLEMENTS_GENERIC	373	0.034453

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	numberOfNodesWithS
0	[Git, Change]	MODIFIES	[File, Git]	259035	259035	
1	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	259035	12706	
2	[Git, Change]	UPDATES	[File, Git]	170412	259035	
3	[Git, Change]	CREATES	[File, Git]	61193	259035	
4	[Git, Change]	DELETES	[File, Git]	37976	259035	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21955	13192	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	15378	12706	
7	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	12706	1	
8	[Author, Git, Person]	COMMITTED	[Git, Commit]	12706	298	
9	[Committer, Git, Person]	COMMITTED	[Git, Commit]	12706	252	
10	[Repository, File, Git]	HAS_FILE	[File, Git]	10810	1	
11	[Git, Change]	RENAMES	[File, Git]	10546	259035	
12	[Java, ByteCode, Member, Method]	HAS	[Java, ByteCode, Parameter]	8361	13192	
13	[Java, ByteCode, Member, Method]	READS	[Java, ByteCode, Member, Field]	8192	13192	
14	[File, Git]	HAS_NEW_NAME	[File, Git]	6237	10810	
15	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6079	13029	
16	[Type, File, Java, ByteCode, Class]	DECLARES	[Java, ByteCode, Member, Method]	5013	760	
17	[Type, File, Java, ByteCode, Class]	DEPENDS_ON	[Type, File, Java, ByteCode, JavaType]	3993	760	
18	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3427	7118	
19	[Java, ByteCode, ParameterizedType, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3053	7176	
20	[Java, ByteCode, ParameterizedType, Bound]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, WildcardType]	2908	7176	
21	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, ParameterizedType, Bound]	2657	13029	
22	[Java, ByteCode, ParameterizedType, Bound]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, TypeVariable]	2454	7176	
23	[Java, ByteCode, Member, Constructor, Method]	WRITES	[Java, ByteCode, Member, Field]	2438	2068	
24	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalAnnotatio...	2341	2822	
25	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
26	[Java, ByteCode, Member, Constructor, Method]	INVOKES	[Java, ByteCode, Member, Constructor, Method]	2123	2068	
27	[Type, File, Java, ByteCode, Class]	DECLARES	[Java, ByteCode, Member, Field]	2091	760	
28	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2078	13192	
29	[Type, File, Java, ByteCode, Class, GenericDec...	DECLARES	[Java, ByteCode, Member, Method]	2070	229	

Graph Density

```
total_number_of_nodes (vertices): 348652
total_number_of_relationships (edges): 1082640
-> total directed graph density: 8.906375323844265e-06
-> total directed graph density in percent: 0.0008906375323844265
```