# commit size

# into valid\_commit\_size

# legacy, allow large results

Select

commit

, count(distinct difference\_old\_path) as files

, count(distinct if((LENGTH(REGEXP\_REPLACE(lower(difference\_old\_path), 'test', '#'))

- LENGTH(REGEXP\_REPLACE(lower(difference\_old\_path), 'test', ''))) > 0, null, difference\_old\_path) ) as non\_test\_files

From

[citric-trees-187317:repo\_attributes.valid\_commits]

Group by

commit

# Create refactor size \_diff

Select

From

[citric-trees-187317:repo\_attributes.refactor\_stats\_2018] as r

Join

[citric-trees-187317:repo\_attributes.refactor\_stats\_2018\_commit\_size] as cs

# files per commit distribution

Select

Files

, count(\*) as commits

From

[citric-trees-187317:repo\_attributes.valid\_commit\_size]

Group by

Files

Order by

Files

# non test files files per commit distribution

Select

non\_test\_files

, count(\*) as commits

From

[citric-trees-187317:repo\_attributes.valid\_commit\_size]

Group by

non\_test\_files

Order by

non\_test\_files