

Report CIDER

Date: 2023-02-07

Data: 'set_chlorbenzene-5.sdf', 'set_chlorbenzene.sdf', 'set_phenole.sdf'

Generated CIDER

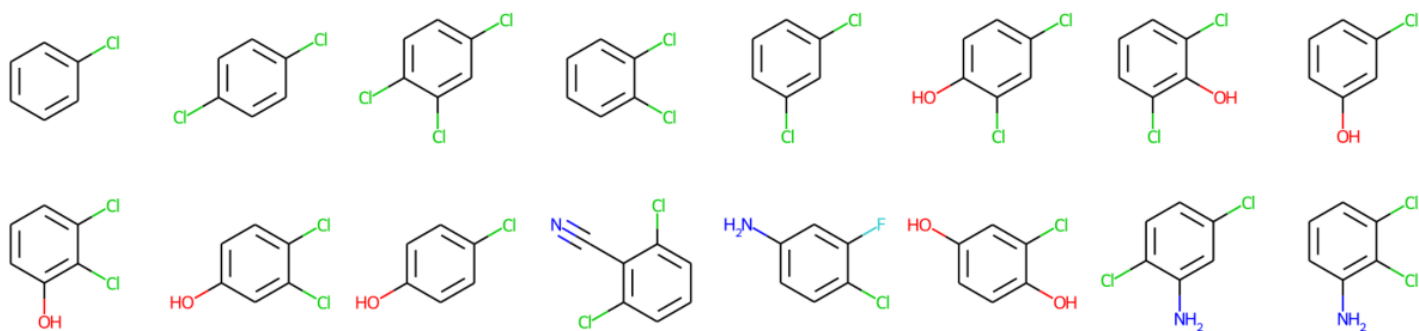
keys: 'SDMolSupplier_Object', 'number_of_molecules', 'coconut_id', 'identifier_list',
'number_of_duplicates', 'duplicates', 'number_of_shared_molecules',
'shared_molecules', 'Molecular Weight', 'Molecular Formula', 'Number of
H-Donors', 'Number of Rings', 'binned_Molecular Weight', 'binned_Number of
Rings', 'number_of_broken_Lipinski_Rules', 'Lipinski_Rule_of_5_summary',
'scaffold_list', 'scaffold_summary'

Generated CIDER

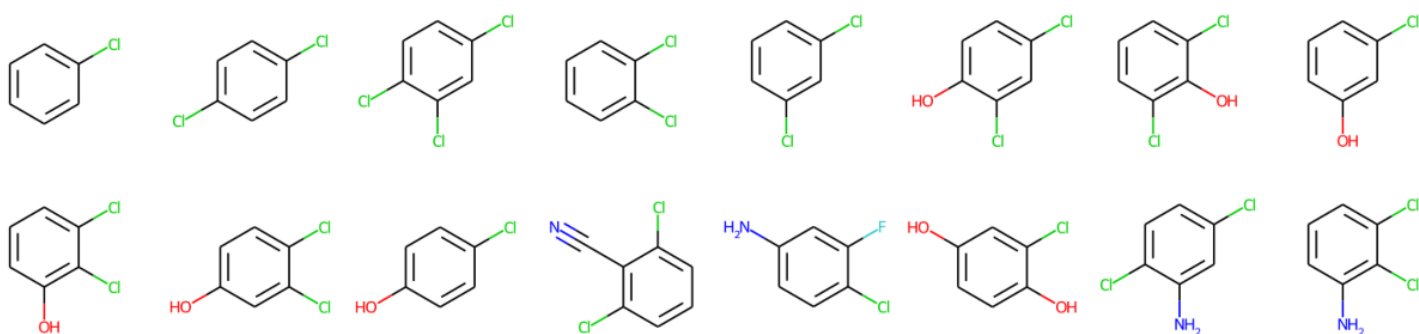
figures: 'mol_grid', 'intersection', 'distribution_of_Molecular Weight',
'distribution_of_Number of Rings', 'lipinski_plot', 'scaffold_grid',
'scaffold_grid_1', 'scaffold_grid_2', 'chemical_space', 'chemical_space_1',
'chemical_space_2'

Exemplary molecules from the datasets

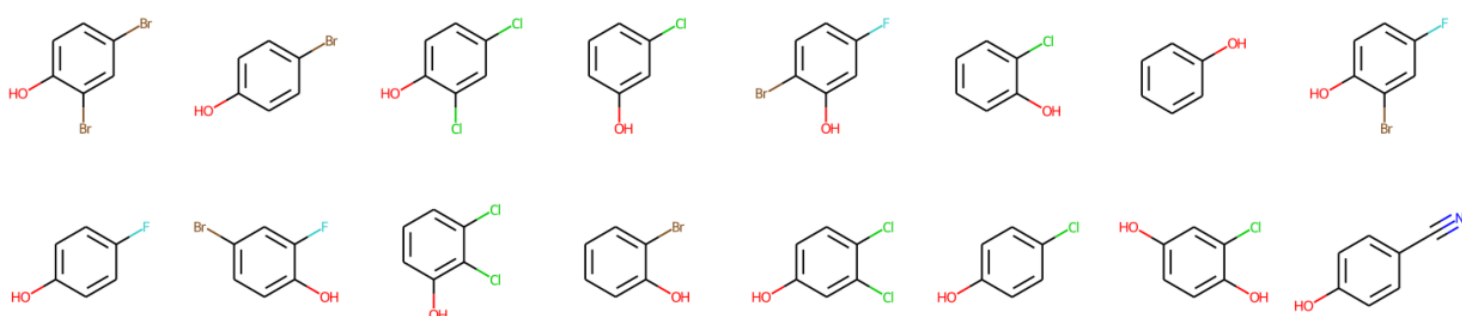
set_chlorbenzene-5.sdf



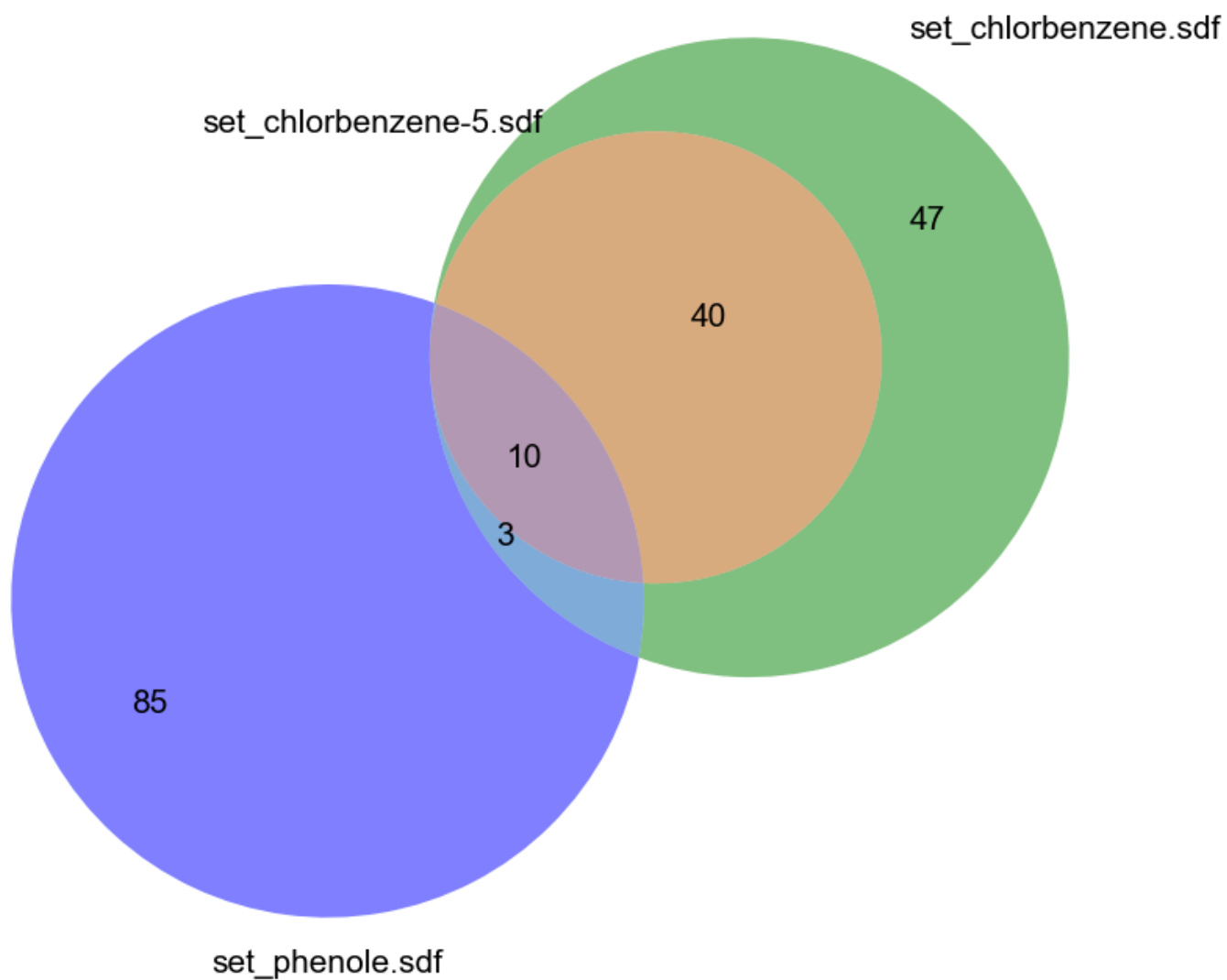
set_chlorbenzene.sdf



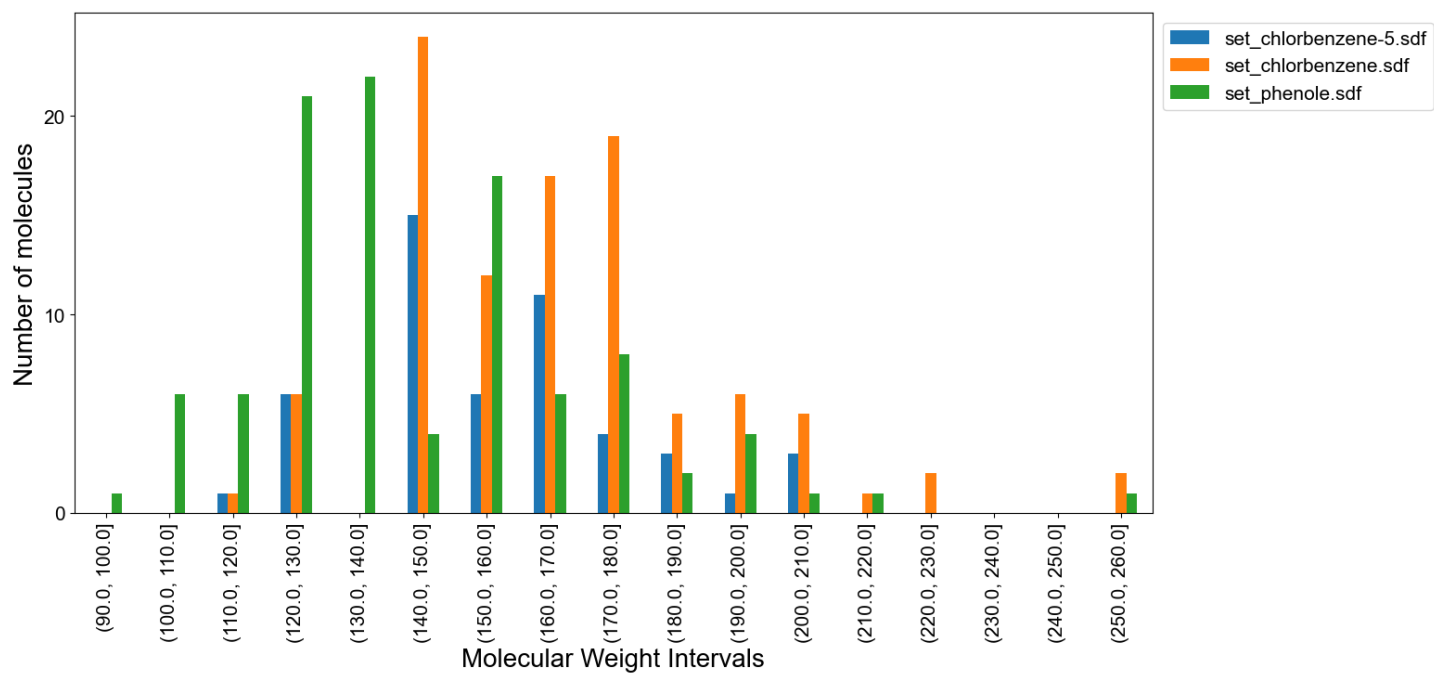
set_phenole.sdf



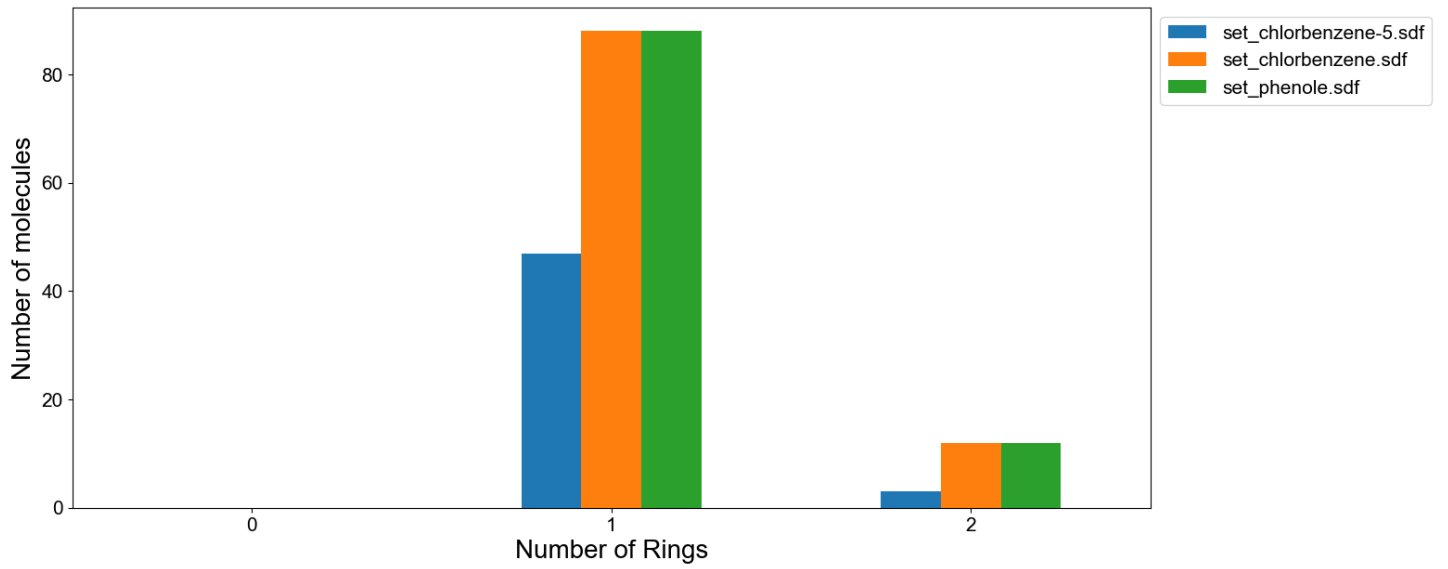
Intersection as Venn diagram



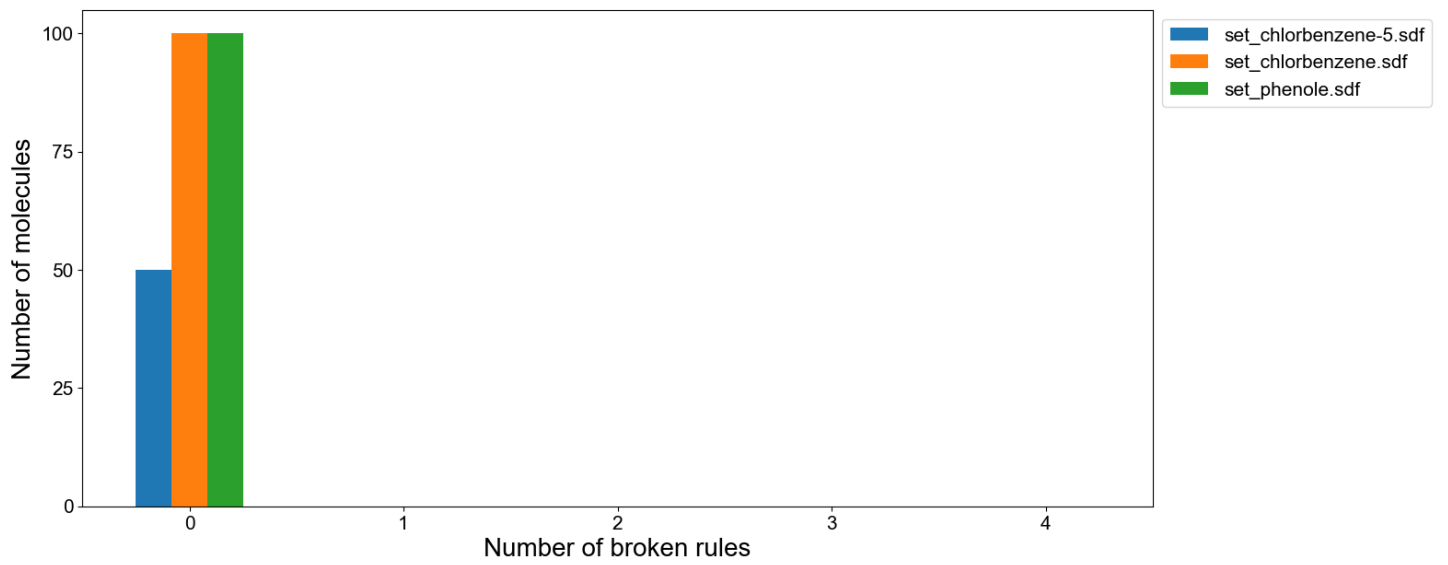
Distribution of Molecular Weight



Distribution of Number of Rings



Distribution of the number of broken Lipinski Rules

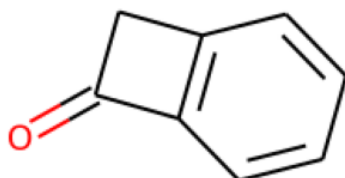


Most frequent scaffolds from the datasets

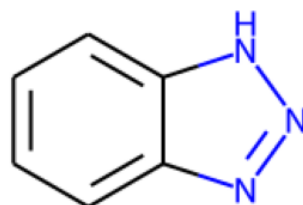
set_chlorbenzene-5.sdf



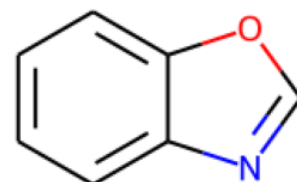
0.94



0.02



0.02

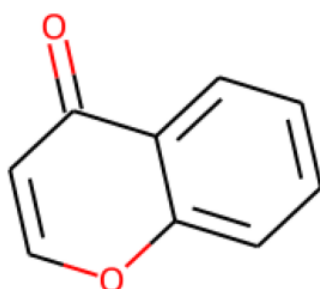


0.02

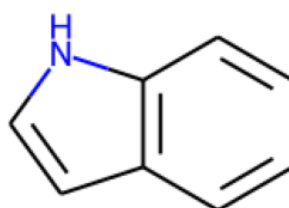
set_chlorbenzene.sdf



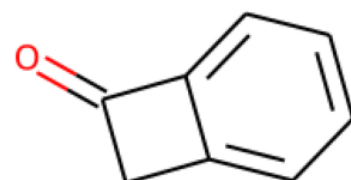
0.88



0.04



0.02

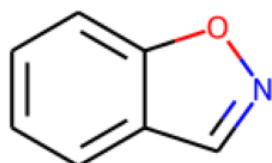


0.01

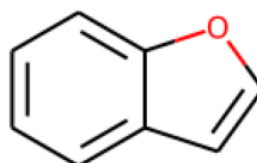
set_phenole.sdf



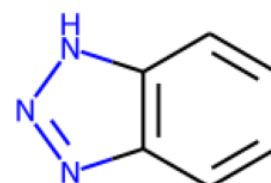
0.88



0.02



0.02



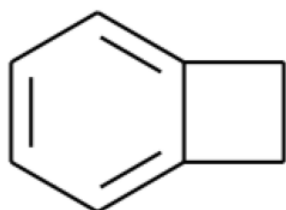
0.01

Most frequent scaffolds from the datasets

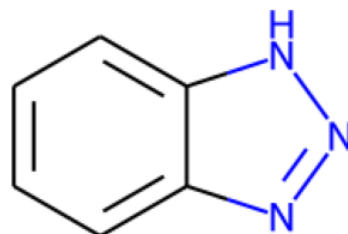
set_chlorbenzene-5.sdf



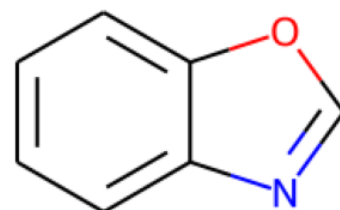
47



1



1

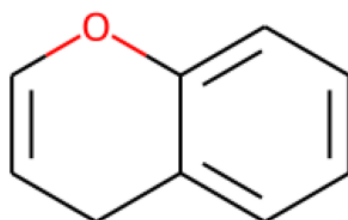


1

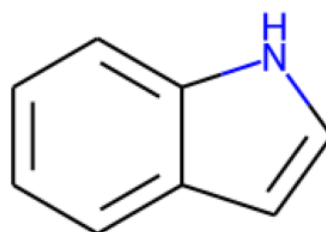
set_chlorbenzene.sdf



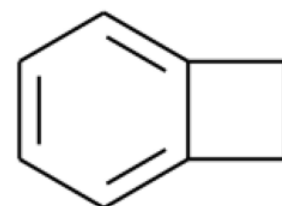
88



4



2

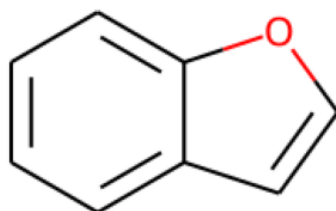


1

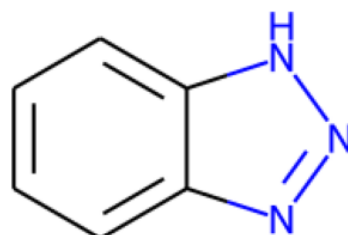
set_phenole.sdf



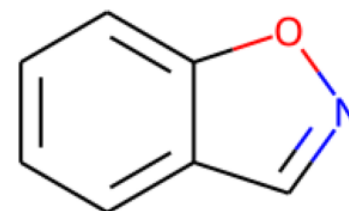
88



3



2



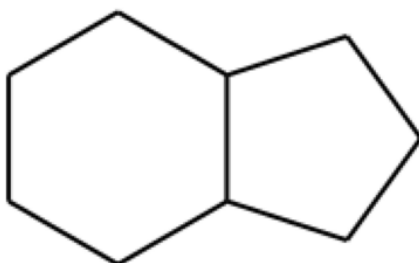
2

Most frequent scaffolds from the datasets

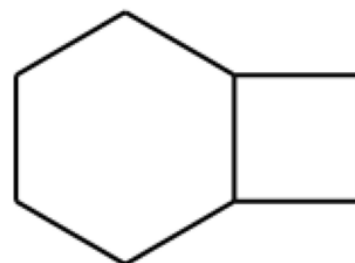
set_chlorbenzene-5.sdf



0.94



0.04

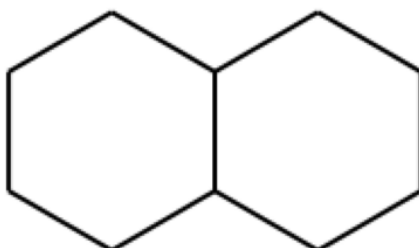


0.02

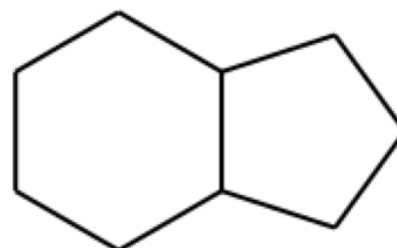
set_chlorbenzene.sdf



0.88



0.06

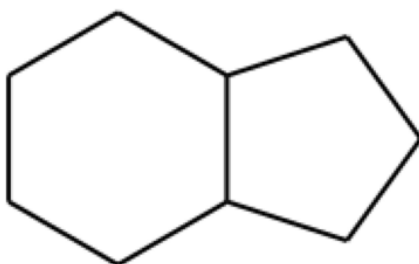


0.05

set_phenole.sdf



0.88



0.12

PCA plot

- set_chlorbenzene-5.sdf
- set_chlorbenzene.sdf
- set_phenole.sdf

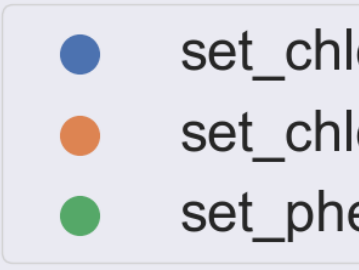
PC-2 (6%)

PC-1 (13%)



UMAP plot

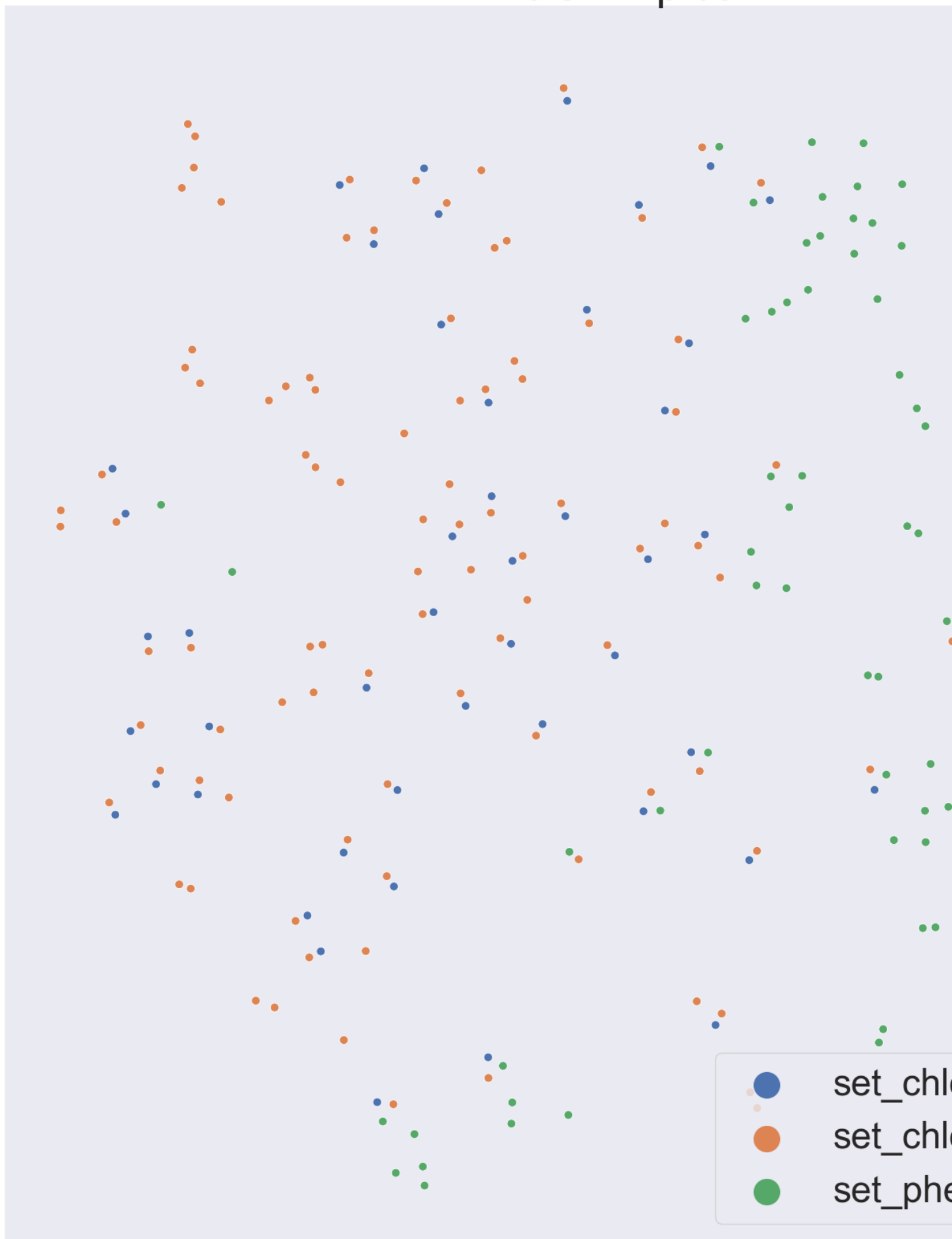
UMAP-2



UMAP-1

t-SNE plot

t-SNE-2



t-SNE-1