Stratified split ensures all 7 domains
are represented in the test set

Training
(24 problems)
68.6%

Test
(11 problems)
31.4%

Individual Math Domains (Stratification Layers)

Note: Random seed = 42 (reproducible split)

0%   20%   40%   60%   80%   100%