

Rank	Goal	Treatment.Dataset	Median Reduction	Spread Reduction	Reduction Quartiles 50%	Avg Estimate (Before →After)	Estimate Stdev (Before →After)
1	score	Baseline.flight::1	58.45%	48.85%		26.0→12.9	24.0→15.1
1	score	Baseline.osp::1	49.58%	42.18%		26.2→12.6	23.8→12.3
1	score	Baseline.ground::1	47.22%	73.77%		23.1→15.2	24.4→19.1
1	score	Baseline.osp2::1	32.91%	39.60%		24.8→17.5	24.0→19.8
1	defects	Baseline.flight::1	65.78%	28.34%		4273.2→2096.8	5281.1→3483.6
1	defects	Baseline.osp::1	60.91%	45.30%		5037.7→2790.4	6763.0→3459.9
1	defects	Baseline.ground::1	58.46%	42.99%		5451.7→3172.2	8602.5→4329.2
2	defects	Baseline.osp2::2	25.92%	60.29%		5231.8→4343.4	7344.5→5311.4
1	effort	Baseline.ground::1	77.09%	47.29%		675.7→375.0	1185.4→943.3
1	effort	Baseline.osp::1	75.00%	34.79%		586.3→227.0	1007.1→336.5
1	effort	Baseline.flight::1	66.57%	43.74%		511.5→187.3	991.6→407.1
2	effort	Baseline.osp2::2	35.48%	45.19%		593.2→312.1	1000.0→404.0
1	months	Baseline.flight::1	36.02%	13.12%		17.2→12.5	7.6→5.5
1	months	Baseline.osp::1	35.49%	22.28%		18.5→13.6	8.6→5.3
1	months	Baseline.ground::1	33.86%	21.97%		18.9→14.7	10.5→7.6
1	months	Baseline.osp2::1	22.57%	31.89%		18.4→14.6	9.2→5.6