

datasets_default_prompt_mixed_aso_test

	geoguessr	tourist	aerial
geoguessr	1.0	0.999930238880864	0.0
tourist	0.0	1.0	0.0
aerial	0.9949143829158908	0.9957416865472509	1.0

datasets_default_prompt_mixed_ttest

	geoguessr	tourist	aerial
geoguessr	1.0	3.384149415528034e-08	2.3704622822061674e-25
tourist	3.384149415528034e-08	1.0	1.9566483775816043e-25
aerial	2.3704622822061674e-25	1.9566483775816043e-25	1.0

datasets_default_prompt_mixed_mixed_stats

Model	Median	Mean	Standard Deviation
geoguessr	0.3691250351428327	0.3674616316889966	0.019392575233285445
tourist	0.4368640379509945	0.44123647858754744	0.04239671883765639
aerial	0.07692307692307693	0.07129786879786879	0.04731126461165153

datasets_default_prompt_mixed_mixed

geoguessr	tourist	aerial
0.35052456723203357	0.4154503105590063	0.125
0.33261627906976743	0.42287687931735063	0.0
0.387124788715976	0.45396888711607497	0.14285714285714285
0.3798613930290553	0.39794050343249426	0.07692307692307693
0.3607508142391863	0.5241676459067763	0.038461538461538464
0.3824833064701819	0.42764260071798965	0.09090909090909091
0.39265209070347384	0.46843944099378887	0.07692307692307693
0.3948935800034797	0.5176886792452831	0.07692307692307693
0.35588787172343206	0.4182289589150379	0.08333333333333333
0.36006066734074826	0.36920622257678504	0.0
0.36511627906976746	0.47690217391304346	0.0
0.3811522464999122	0.4600790513833992	0.17857142857142855
0.39161736277898196	0.4810131631431991	0.11904761904761904
0.3536061230497498	0.4848823294774631	0.041666666666666664

0.38290697674418606	0.43373590982286636	0.05
0.36949249627677855	0.4399921660791226	0.045454545454545456
0.3712831192872845	0.3877163360067539	0.045454545454545456
0.32431163954943676	0.46251468860164513	0.08333333333333333
0.3441334579876125	0.38679398618447786	0.03571428571428571
0.36875757400888687	0.3954896383583909	0.11538461538461539