Heat Map of All Scores

Method	Chinese (Faithfulness)	Chinese (Creativity)	Japanese (Faithfulness)	Japanese (Creativity)	Korean (Faithfulness)	Korean (Creativity)	Plausible Chinese (Faithfulness)	Plausible Chinese (Creativity)
Google	4.05	3.43	3.77	3.43	3.14	3.15	3.74	3.59
DeepL	3.77	3.46	3.41	3.40	3.13	3.38	3.45	3.66
Baseline	4.26	3.70	4.11	3.63	3.62	3.46	4.09	3.84
Diversity Explicit	4.22	3.86	4.06	3.78	3.62	3.68	4.08	3.95
Diversity Dialog	4.15	3.94	4.00	3.87	3.60	3.79	4.02	4.02
Zero-Shot Creatively	4.30	4.09	4.16	3.99	3.76	3.97	4.21	4.10
Context Explicit	4.29	3.73	4.11	3.62	3.61	3.48	4.12	3.86
Analogy Natural	3.95	4.06	3.70	4.03	3.55	3.97	3.79	4.06
Analogy Creative	3.19	4.25	3.07	4.29	2.80	4.23	3.10	4.22
Shuffle Order	4.29	3.74	4.16	3.65	3.65	3.52	4.14	3.85
Succinct	4.21	3.71	4.10	3.61	3.66	3.55	4.05	3.81
Two-Step	3.89	3.36	3.68	3.27	3.26	3.07	3.71	3.53
Discontinuous 1	4.25	3.74	4.09	3.63	3.56	3.59	4.05	3.84
Discontinuous 2	4.13	3.58	3.95	3.47	3.44	3.37	3.90	3.72
Few-Shot	4.34	3.86	4.21	3.79	3.70	3.72	4.16	3.93
Few-Shot Creatively	4.22	4.08	4.14	4.05	3.72	4.01	4.18	4.08