Spearman Correlation Heatmap

	BERTScore	ROUGE	METEOR	BLEU	OE-Eval
QwenVL2.5-72B	61.16	67.04	69.00	59.32	68.31
Gemini-2.5-Flash	60.34	66.90	67.12	58.96	67.13
GPT-4o	63.06	70.01	67.34	60.57	69.85
InternVL3-78B	59.79	63.29	64.55	57.14	63.99
Overall	61.09	66.81	66.75	59.00	67.32

Metric