Penn Treebank Validation (Lower) & Test Correlation Matrix

	WMD	Perplexity	SkipThought	Transformer	BLEU-2	BLEU-3	InferSent	BLEU-4	Av-Cosine
WMD	1	−0.21	0.25	0.54	−0.4	−0.43	0.09	−0.11	0.25
Perplexity	0.2	1	−0.07	−0.74	−0.81	−0.78	−0.78	−0.87	−0.67
SkipThought	0.16	−0.38	1	0.85	0.03	0.34	0.23	0.1	0.49
Transformer	0.2	−0.59	0.16	1	0.82	0.74	0.21	0.4	0.36
BLEU-2	0.06	−0.96	0.16	0.35	1	0.77	0.81	0.93	0.53
BLEU-3	0.65	−0.81	0.24	0.49	0.94	1	0.72	0.93	0.55
InferSent	−0.58	−0.83	0.54	0.46	0.53	0.65	1	0.84	0.64
BLEU-4	0.4	−0.97	0.35	0.58	0.43	0.46	0.83	1	0.69
Av-Cosine	0.38	−0.74	0.87	0.96	0.18	0.33	0.65	0.86	1