LLM-Judge Score vs TrueSkill ELO
(Dimension-wise Correlations: Pearson + Spearman)