VLM Performance by Region
(Mean ± Std across 23 models)