Accuracy by Question Type (TinyLlama vs Your Model)