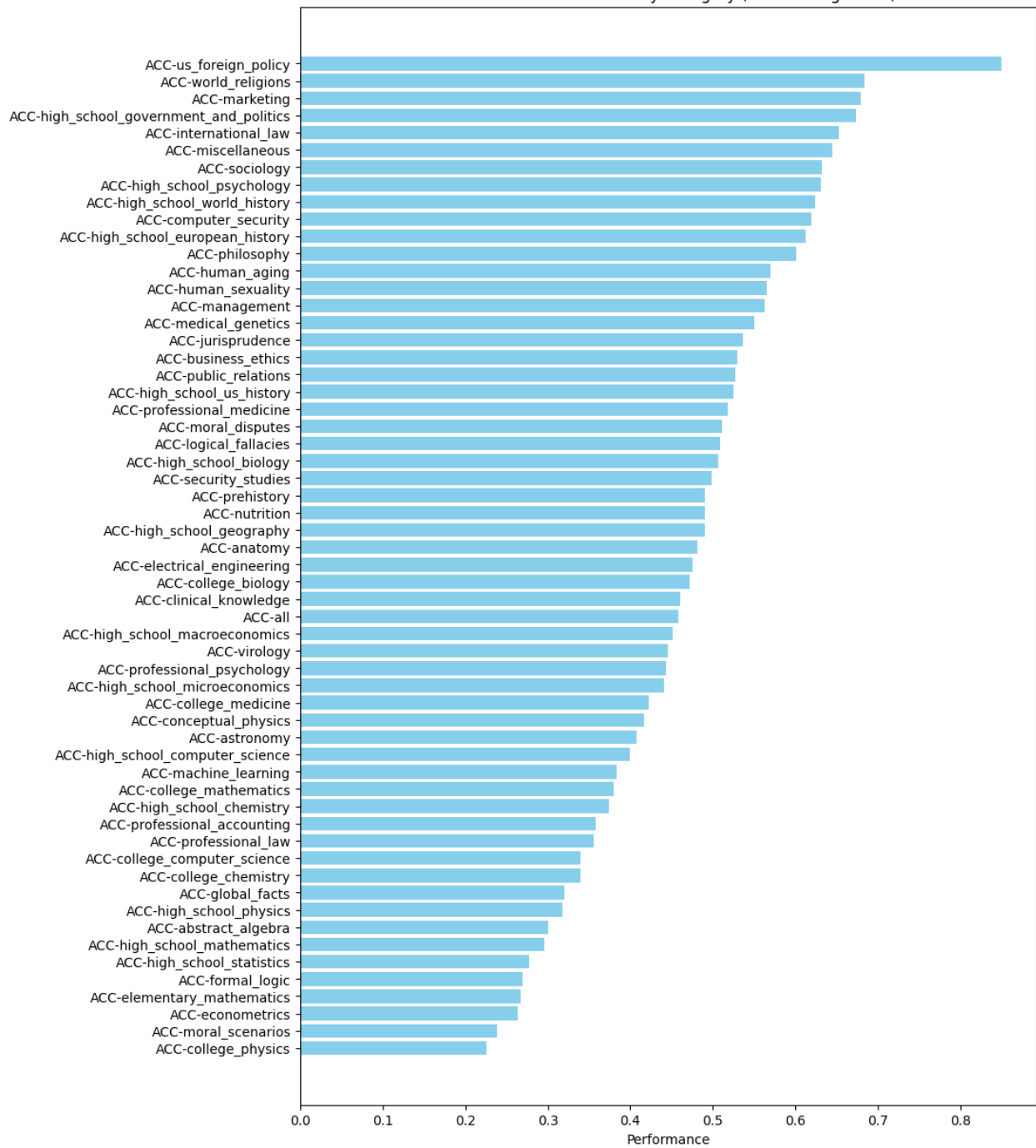
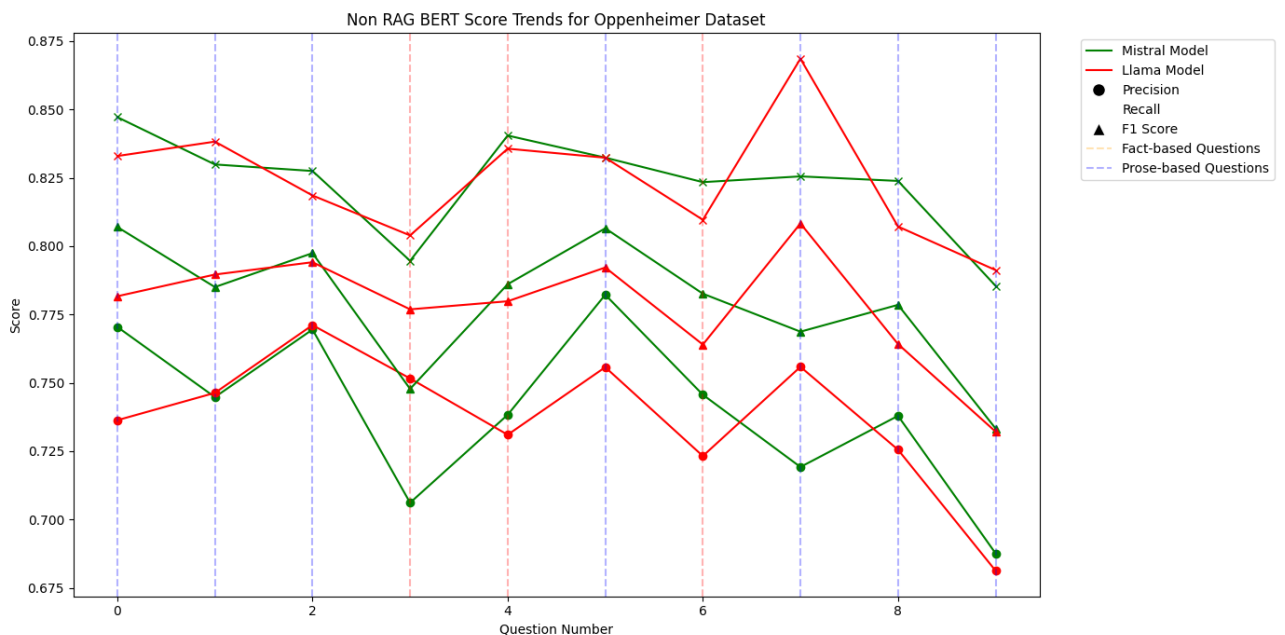


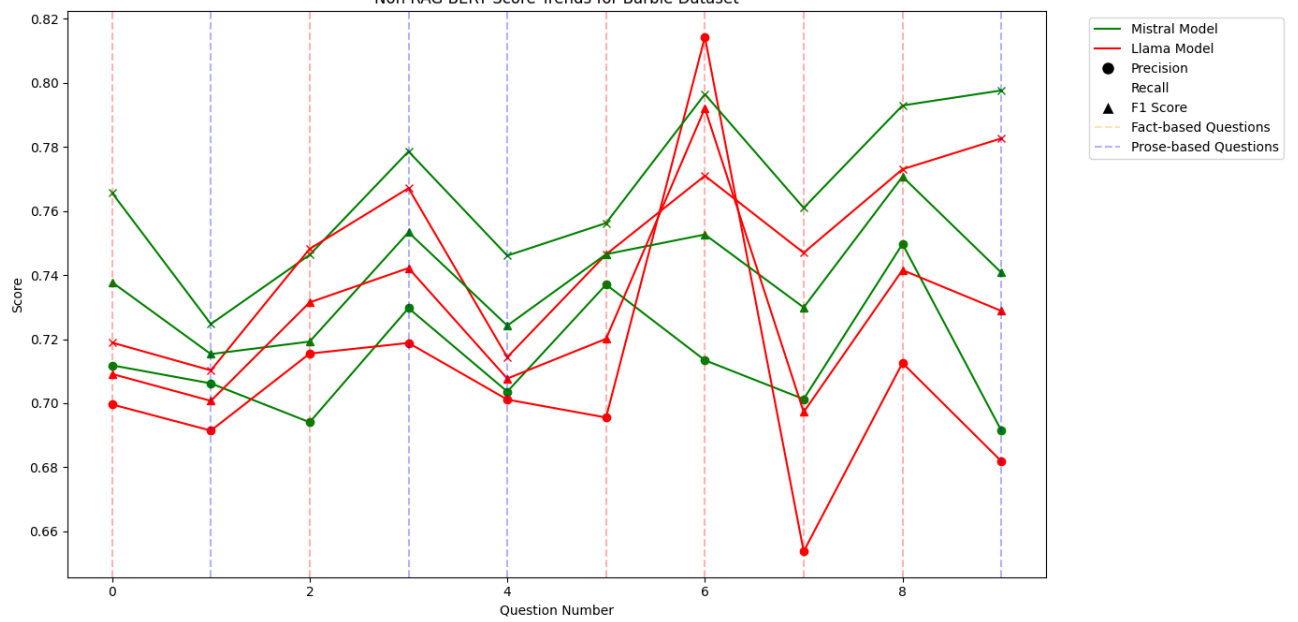
Llama-2-7B Performance by Category (Descending Order)



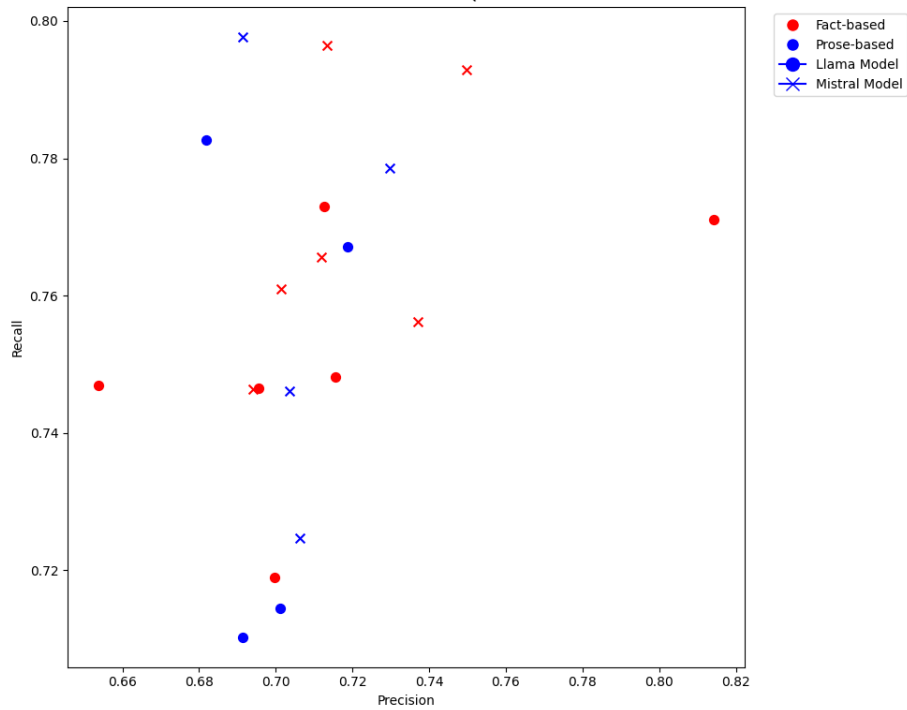
Non RAG BERT



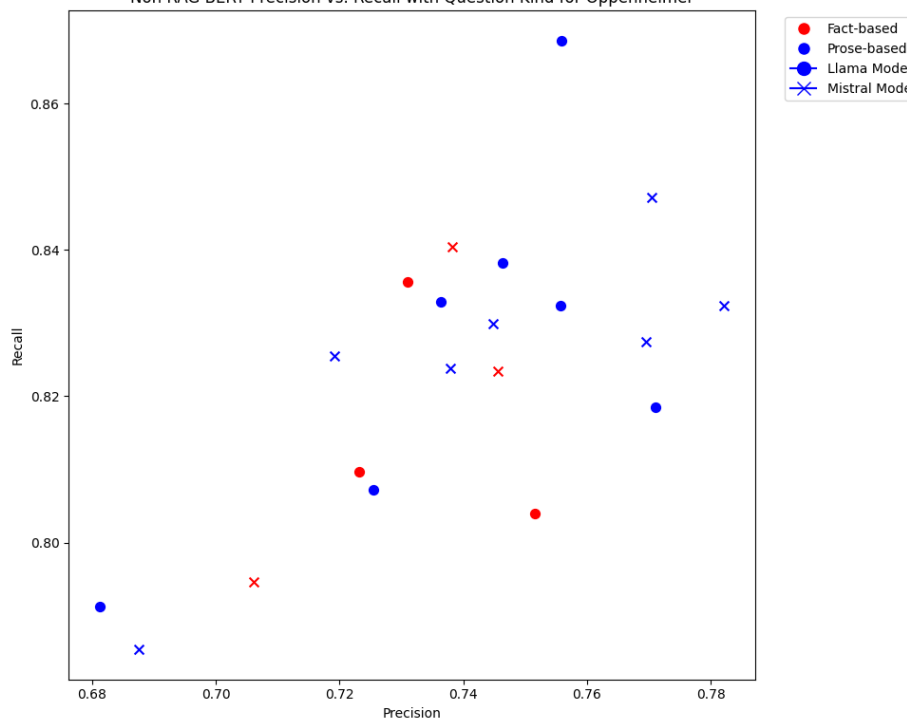
Non RAG BERT Score Trends for Barbie Dataset



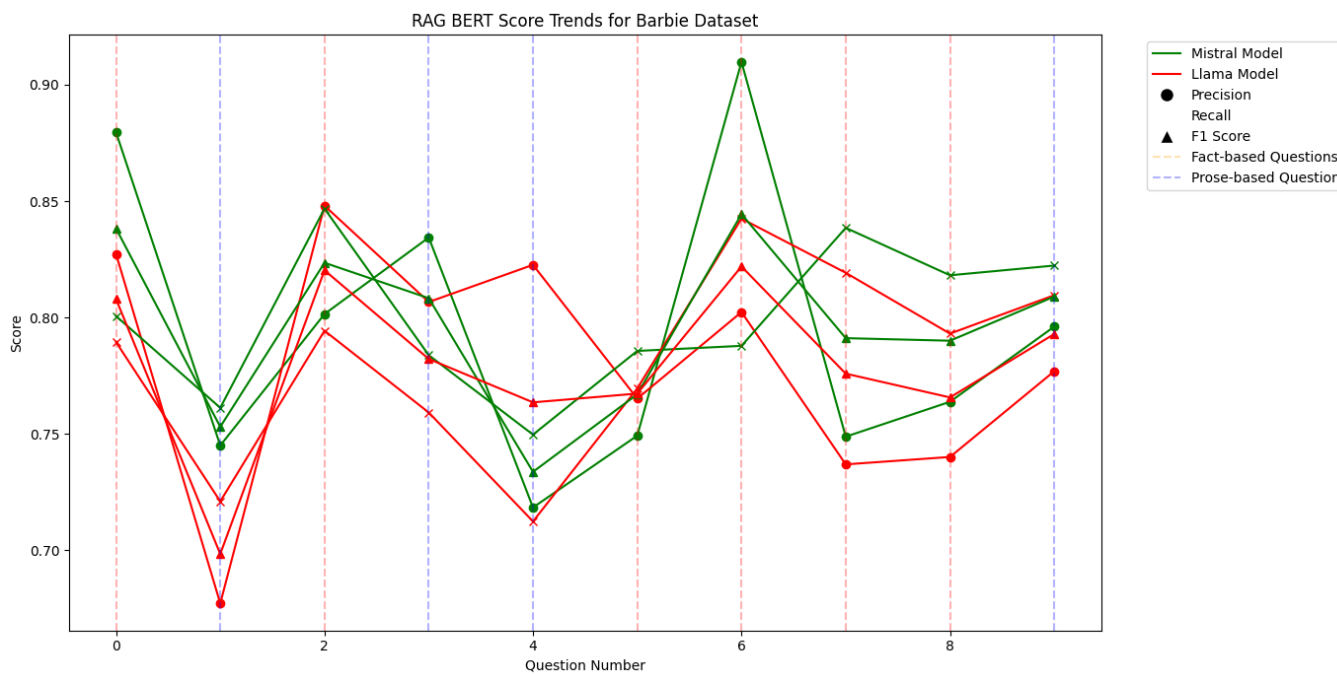
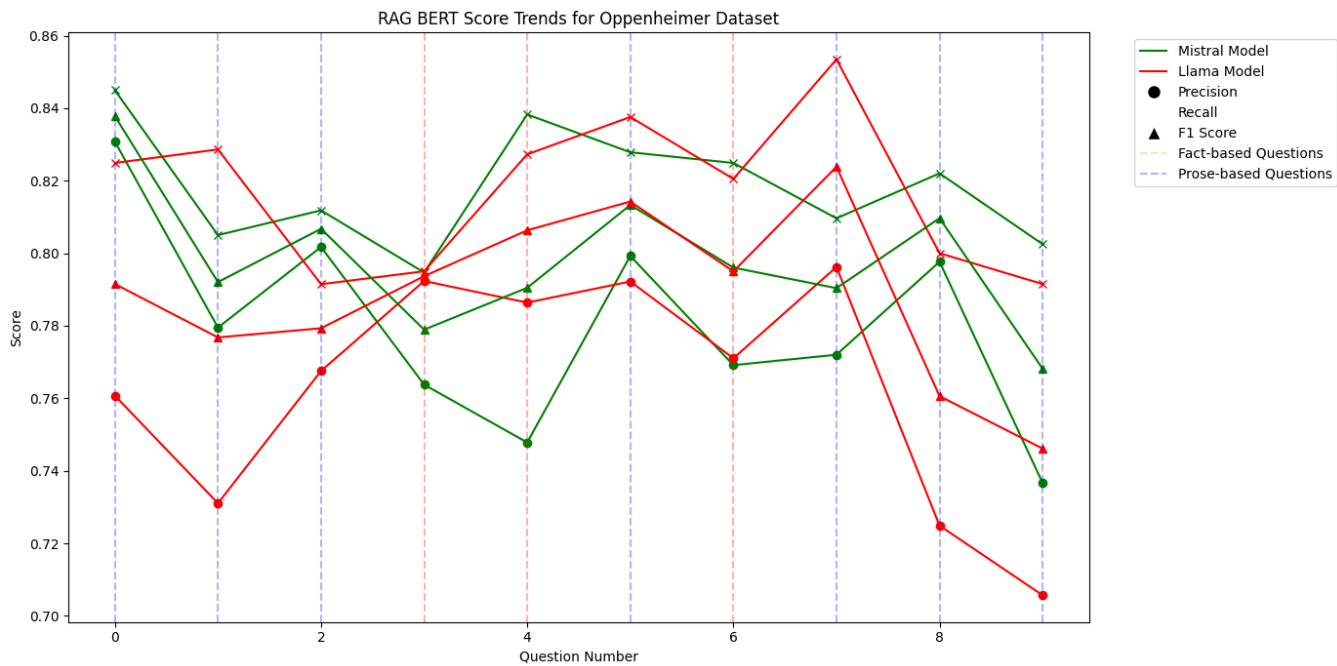
Non RAG BERT Precision vs. Recall with Question Kind for Barbie

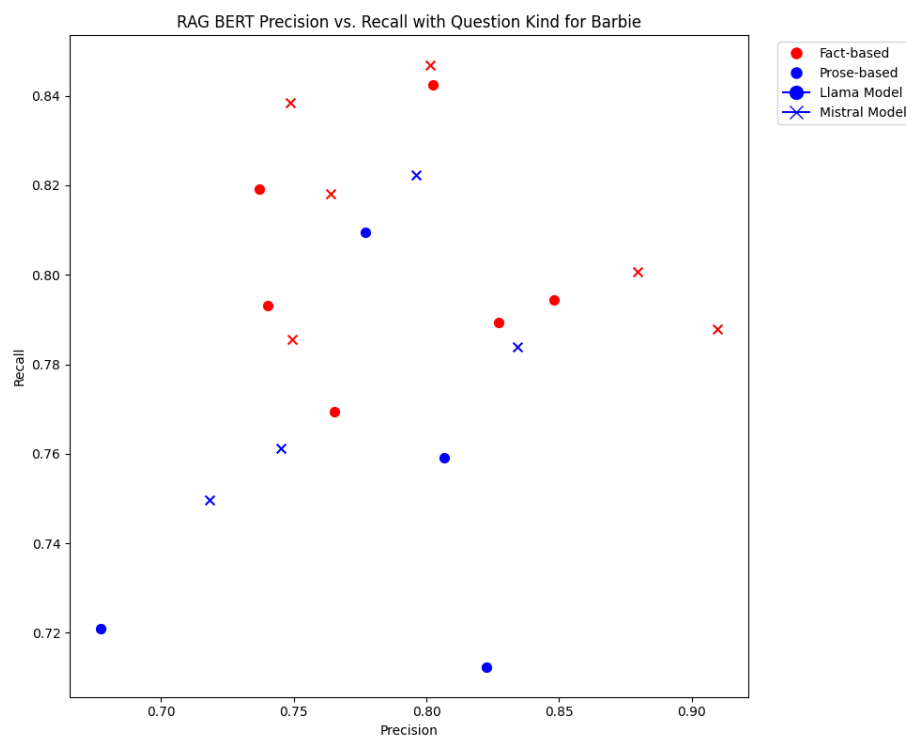
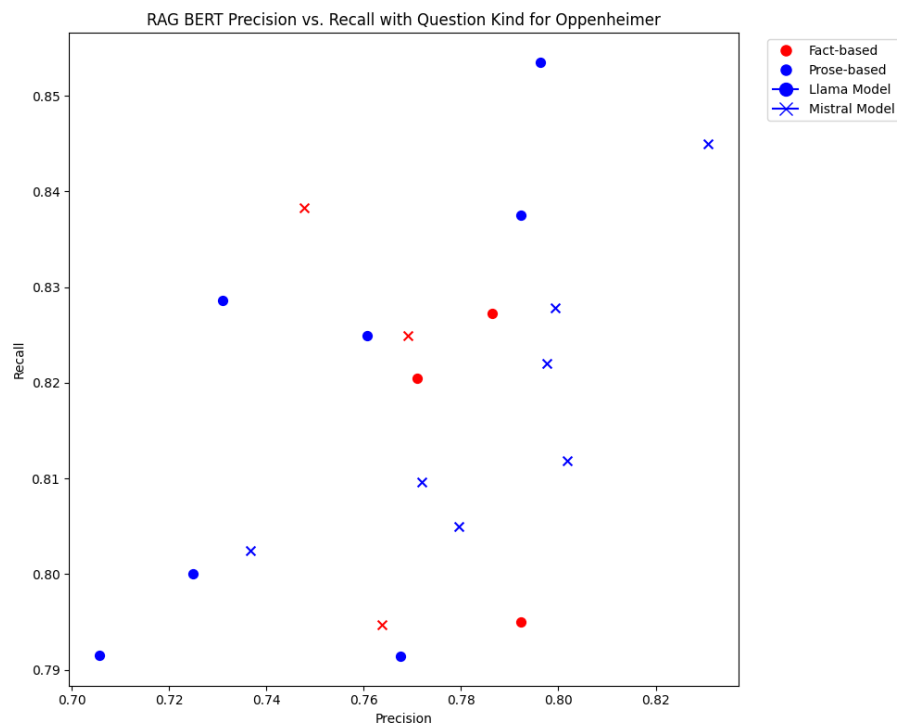


Non RAG BERT Precision vs. Recall with Question Kind for Oppenheimer



RAG BERT





Semantic Similarity:



