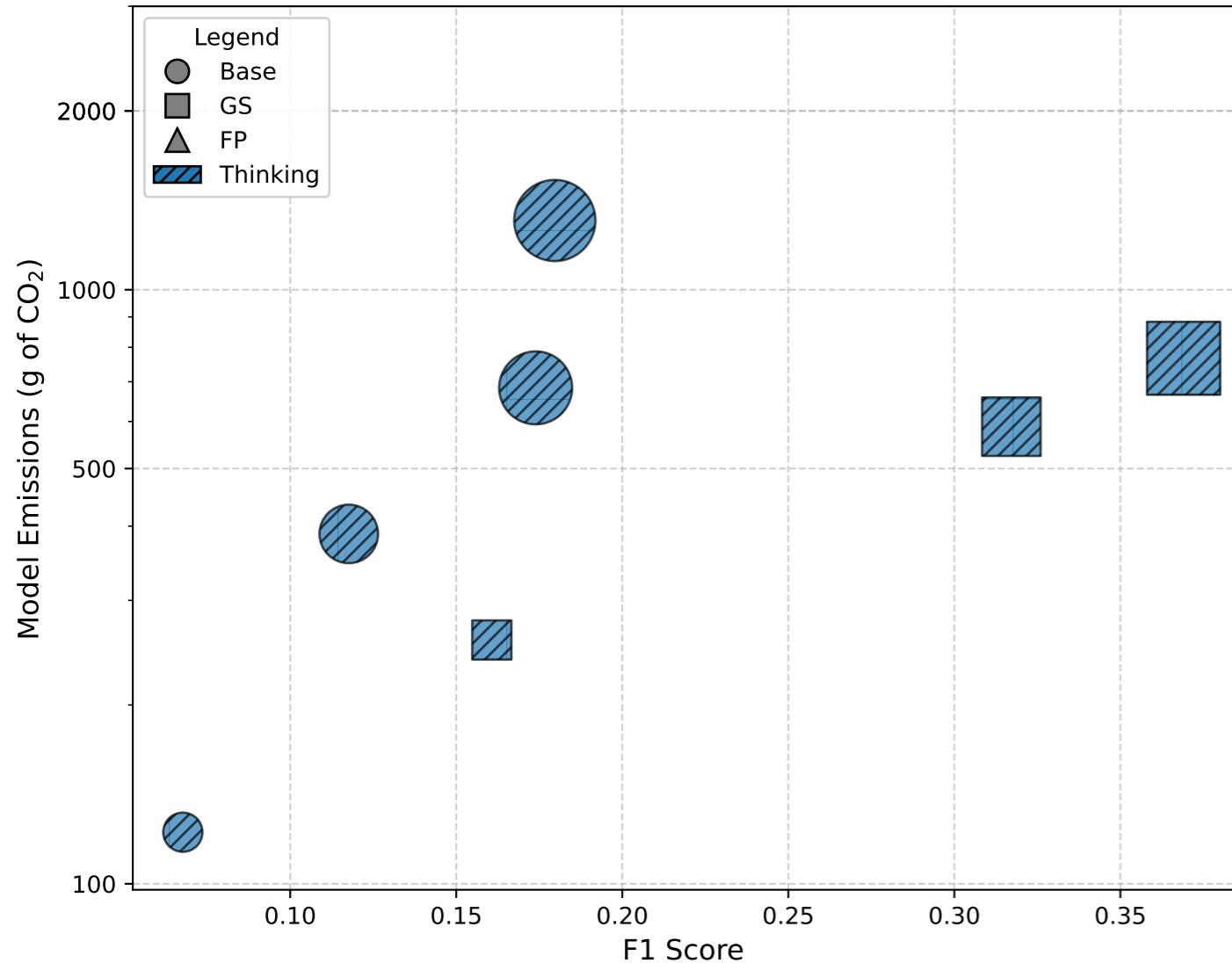


DeepSeek-r1 F1 Score vs Total Emissions

HotpotQA



Natural Questions

