

Average F1 vs Average Time by Model and Prompt Version

