

Test perplexity grouped by model

