**🚀 Local Multi-LLM Testing & Performance Tracker**

What does this do?

* **Trigger & Model Setup:**  
  The workflow kicks off when a chat message is received. It then reaches out to your local LM Studio server to fetch a list of available language models (make sure to update the base URL to match your LM Studio server’s IP as noted in the sticky notes).
* **Processing Chat Messages:**  
  When you send a chat message, the workflow automatically adds a system prompt that instructs the models to keep their responses concise and easy enough for a 5th grader to understand. It then dynamically sends your prompt to one or more of the local models for testing.
* **Timing & Analysis:**  
  It captures the start and end times of the response to calculate how long each model takes to reply. After receiving a response, a code node analyzes the text by calculating:
  + **Word and sentence counts**
  + **Average word and sentence lengths**
  + **A readability score** (using the Flesch-Kincaid formula) to see how easy the response is to read
* **Logging Results:**  
  All of the data—including the original prompt, the model used, the response details, the timing information, and the analysis metrics—is automatically logged into a Google Sheet. This lets you review and compare the performance of different models over time.
* **Built-In Guidance & Tips:**  
  Throughout the workflow, there are several sticky notes offering helpful setup instructions (like installing LM Studio, configuring model parameters, and even best practices such as deleting previous chats to avoid overlapping responses).