**Challenge Overview:**

Your task is to write a Node.js script that processes and combines data from two sources to create a training dataset that can be used to train a digital assistant powered by LLaMA2 7b, the dataset should be formatted correctly for compatibility with FastChat, following the Vicuna training format for fine-tuning (<https://github.com/lm-sys/FastChat?tab=readme-ov-file#fine-tuning>). This assistant's goal is to provide users with the relevant product category based on given product data. The dataset you will work with includes:

* Trade-specific Information (JSON File): Contains details about various trades, including product IDs and other product features.
* Product Categorization (Excel Spreadsheet): Lists the same products by ID and names along with their relevant categorizations.

**Objectives:**

Data Extraction and Transformation:

Extract data from both the JSON file and the Excel spreadsheet.

Dataset Creation:

Create a final dataset that could be used to train a conversational agent. The dataset should include dialogues where the user presents the chatbot with product details, and the chatbot returns the relevant product category.

Please note that the user should not submit the request in json format, but provide a natural language description of the product based on the features included in the json (for example, they may provide a list of features to describe the product).

Documentation:

Document your code thoroughly.

Write a brief report explaining your approach, any challenges you encountered, and how you overcame them.

**Requirements:**

Use Node.js for scripting.

Ensure your script handles errors gracefully and logs meaningful messages that can help debug issues.

The final dataset should be clean and well-organized.

Evaluation Criteria:

Efficiency: How effectively and efficiently your script processes the data.

Creativity and Problem-Solving: Innovative approaches to handling any challenges encountered.

Documentation: Clarity and thoroughness of your report and code comments.

Include a README file with instructions on how to run your script and any dependencies required.

**Bonus points:**

The following are not mandatory, but their completion of one or more bonus points will be considered as a plus.

* Account for variability in the training data, the user does not always submit the request with the same format and using the same exact wording
* When applicable, make the product name dynamic so that it’s preceded by “autocallable” or “issuer callable” – you should be able to derive this from the json
* Add follow up questions to your training data: in cases where the user does not provide enough information for determining the product category, the chatbot should ask for further information. Once this has been provided, it can then return the product category
* A few products in the json file may not be categorized in the spreadsheet, can you derive their product category and include them in your dataset?