1. Read doc

2. Try gpt
    a. Looks cool
    b. Run colab

3. Data exploration:
    1. Tricky part
        a. Amount is not an int
        b. Natural Language may not be accurate (names)

4. Research online
    a. SQL evaluation library
    b. https://github.com/whoiskatrin/sql-translator
    c. https://blog.futuresmart.ai/mastering-natural-language-to-sql-with-langchain-nl2sql#heading-incorporating-few-shot-examples-into-langchain
    d. https://platform.openai.com/docs/examples/default-sql-translate
    e. https://community.openai.com/t/converting-natural-language-to-sql-query/412565
        i. If you provide the database columns and a small description of the data that is stored in the columns plus a sample query or two at temperature 0, you should be able to get really good accuracy for SQL query generation. The trick lies in finding what exactly to put in the context and what to put into the prompt
    f. https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1184/reports/6907018.pdf

5. Approaches
    a. LangChain + Sqlite + SqlAlchemy + SqlAgent
    b. Raw openAI

6. Small twists
    a. Few-Shot Learning
        i. Add sample queries for clarity
    b. Prompt tuning
    c. Convert given csv into simpler structure for LLM understanding
    d. Validation for generated result
    e. RAG?

7. Main target
    a. Prompt
    b. Context