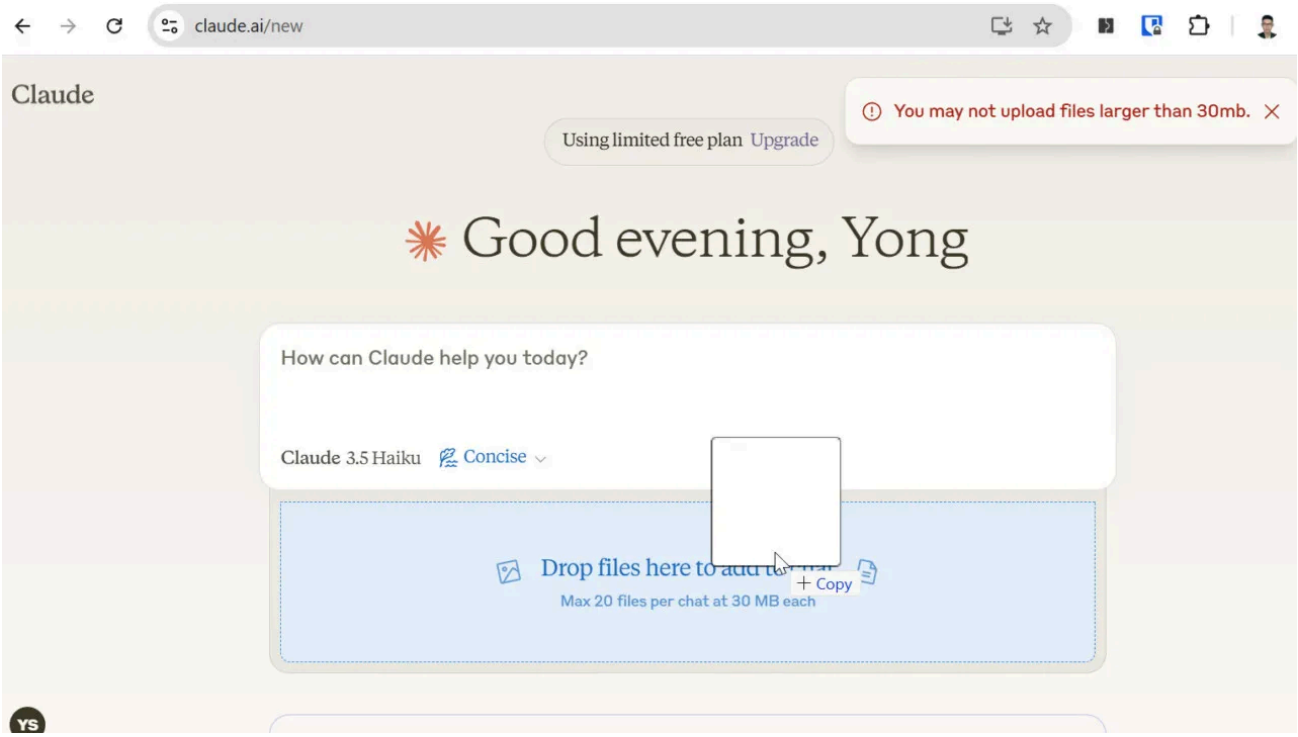
Claude’s Data Exploration with MCP Server

YONG SHENG TAN / FEBRUARY 8, 2025

Addressing Challenges for Data Exploration with LLM

Many people attempt to use LLMs for data exploration, but when uploading large CSV files, they often encounter quota errors due to context length limitations. LLMs have a fixed context window, restricting the amount of data they can process at once—for example, GPT-4 supports up to 32,000 tokens, while Claude 3.5 extends to 200,000 tokens. This constraint means that not all data can be analyzed simultaneously, potentially leading to incomplete insights.

For example, when I tried to upload a dataset around 174.6MB to Claude, an error will pop out and mention that we couldn’t upload files larger than 30MB.



When I face such error, previously, what would I do is that I will just upload a small sample of CSV files. But now, I found another way! 🤩

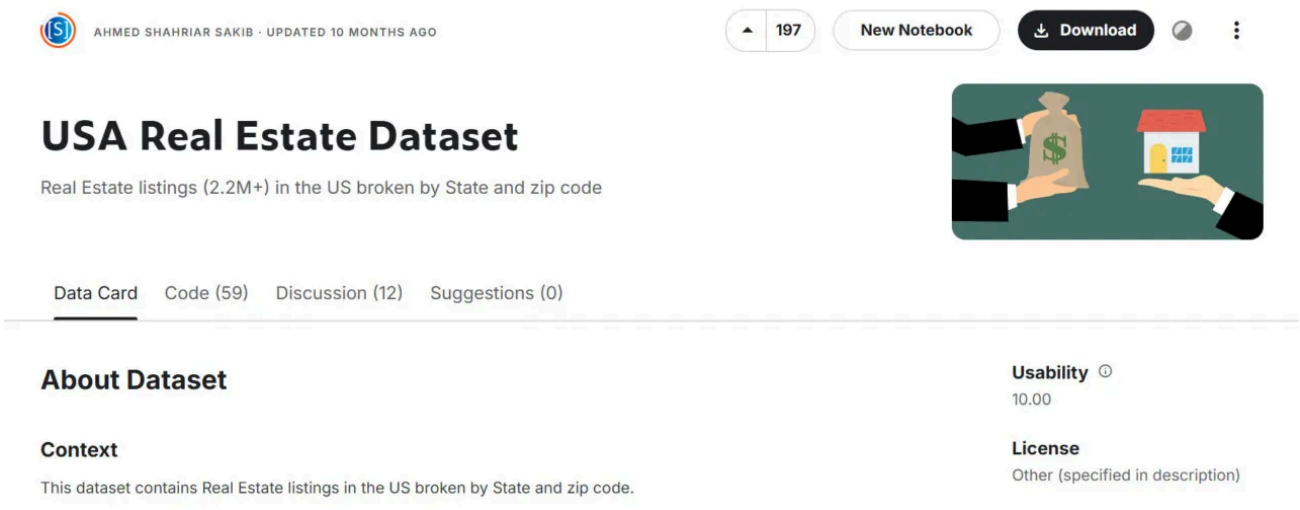
Table of Contents

Introduction to Dataset & Tools

Before introducing to the dataset and tools, I would like to clarify that I am setting this up in Windows 11.

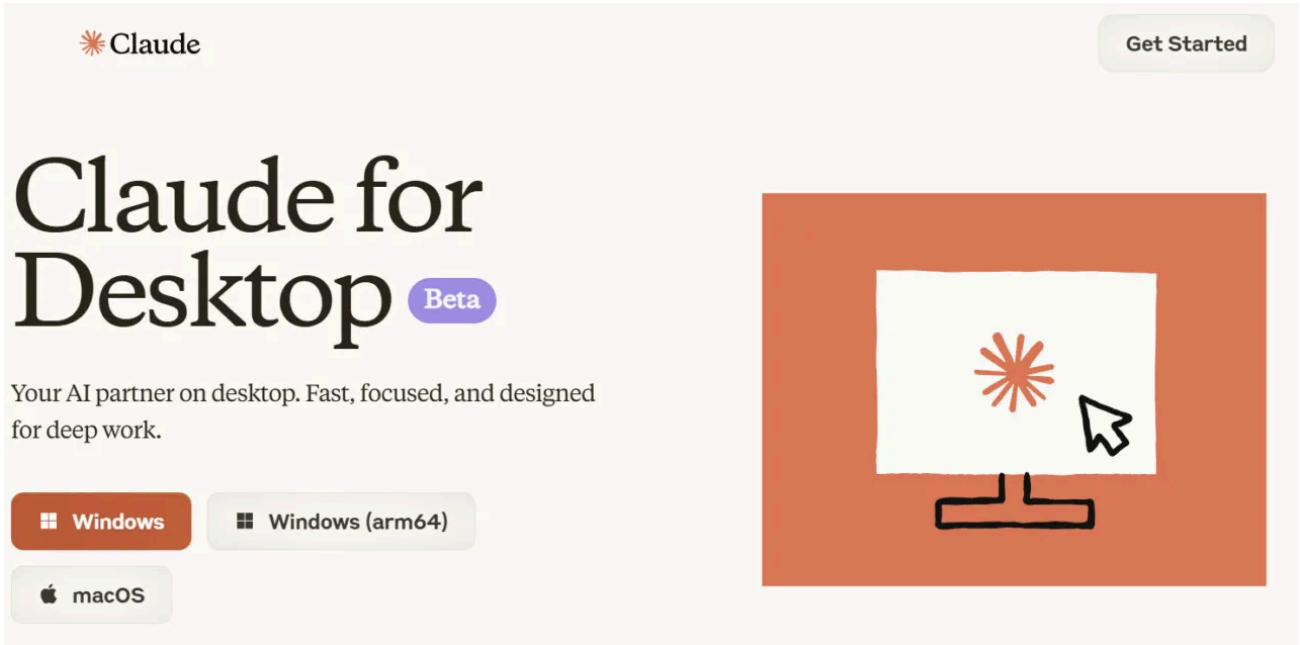
(i) Dataset – USA Real Estate Dataset

- **Source**: Kaggle
- **Size**: 2,226,382 entries (174.6 MB)



(ii) Claude Desktop

- **App**: https://claude.ai/download



(iii) MCP server

The Model Context Protocol, used by Claude LLM, alleviates the core challenge of building LLM-based applications by enabling seamless interaction with external tools and data.

For this blog, we're using **MCP server for data exploration**: https://github.com/reading-plus-ai/mcp-server-data-exploration. Appreciate the author's hard work on this.

Part 1: Setting up MCP Server for Data Exploration

Part 2: Setting up Claude Desktop

2.1

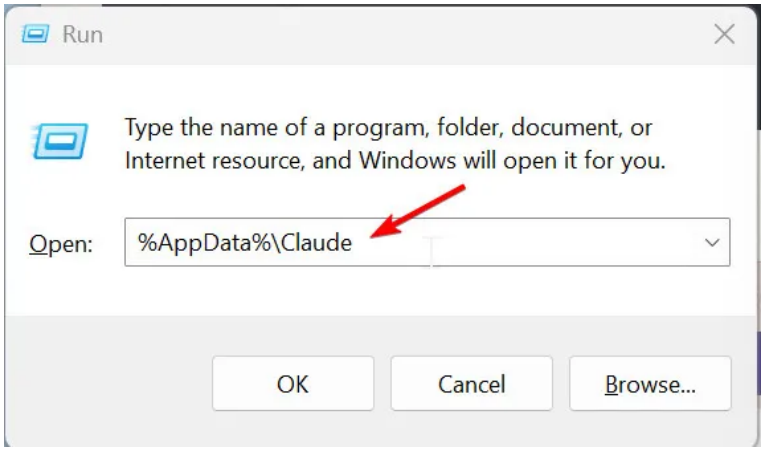
- git clone https://github.com/reading-plus-ai/mcp-server-data-exploration

2.2 Setting up Claude Desktop

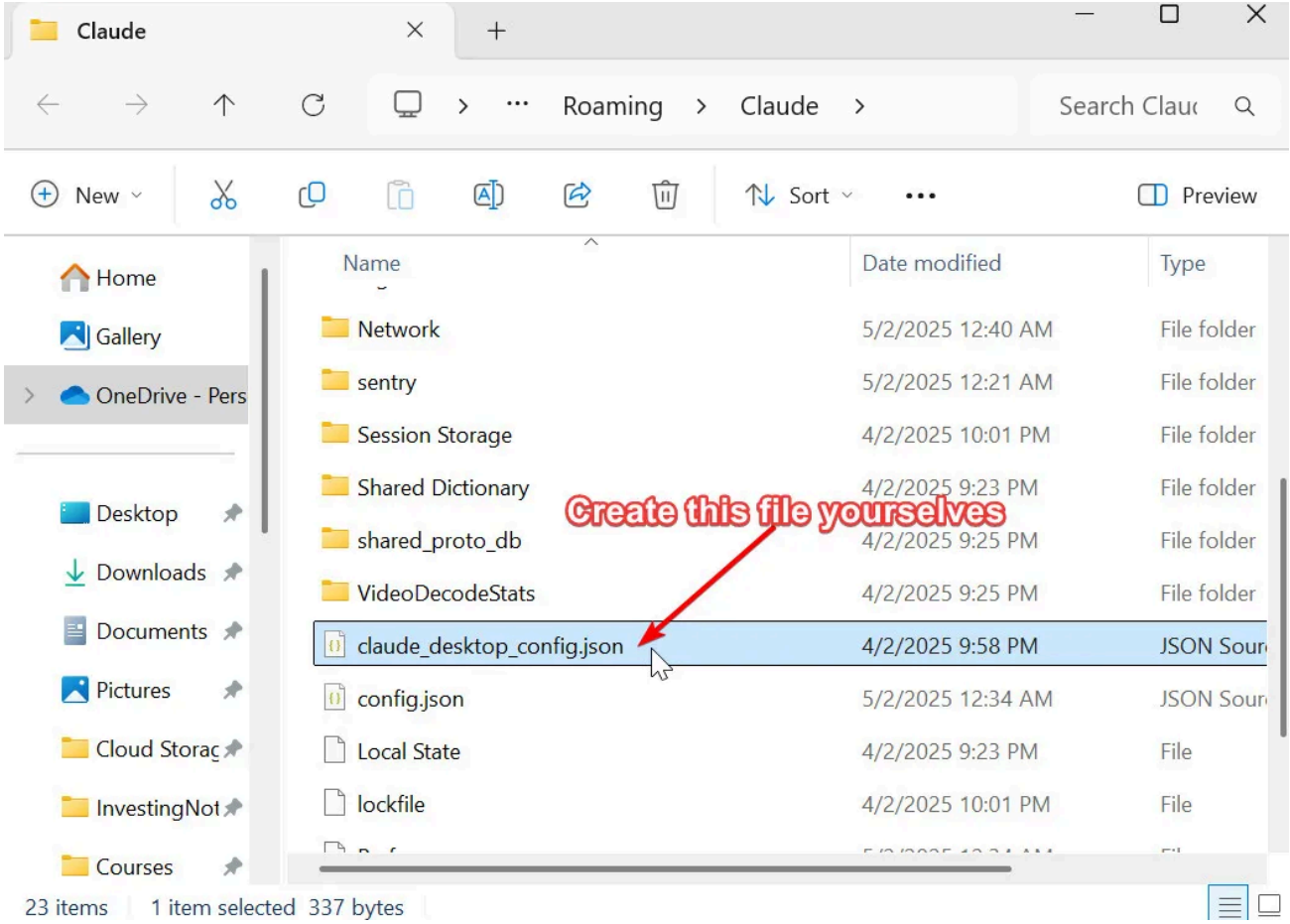
(i)

(ii) Create claude_desktop_config.json file

- Press `Ctrl + R`, then input `%AppData%\Claude` and hit Enter.



- Create `claude_desktop_config.json` file



- Then,

```
{
    "mcpServers": {
        "data_explore": {
            "command": "uv",
            "args": [
                "--directory",
                "C:\\Users\\tys\\Documents\\mcp_servers\\mcp-server-data-explo
                "run",
                "mcp-server-ds"
            ]
        }
    }
}
```

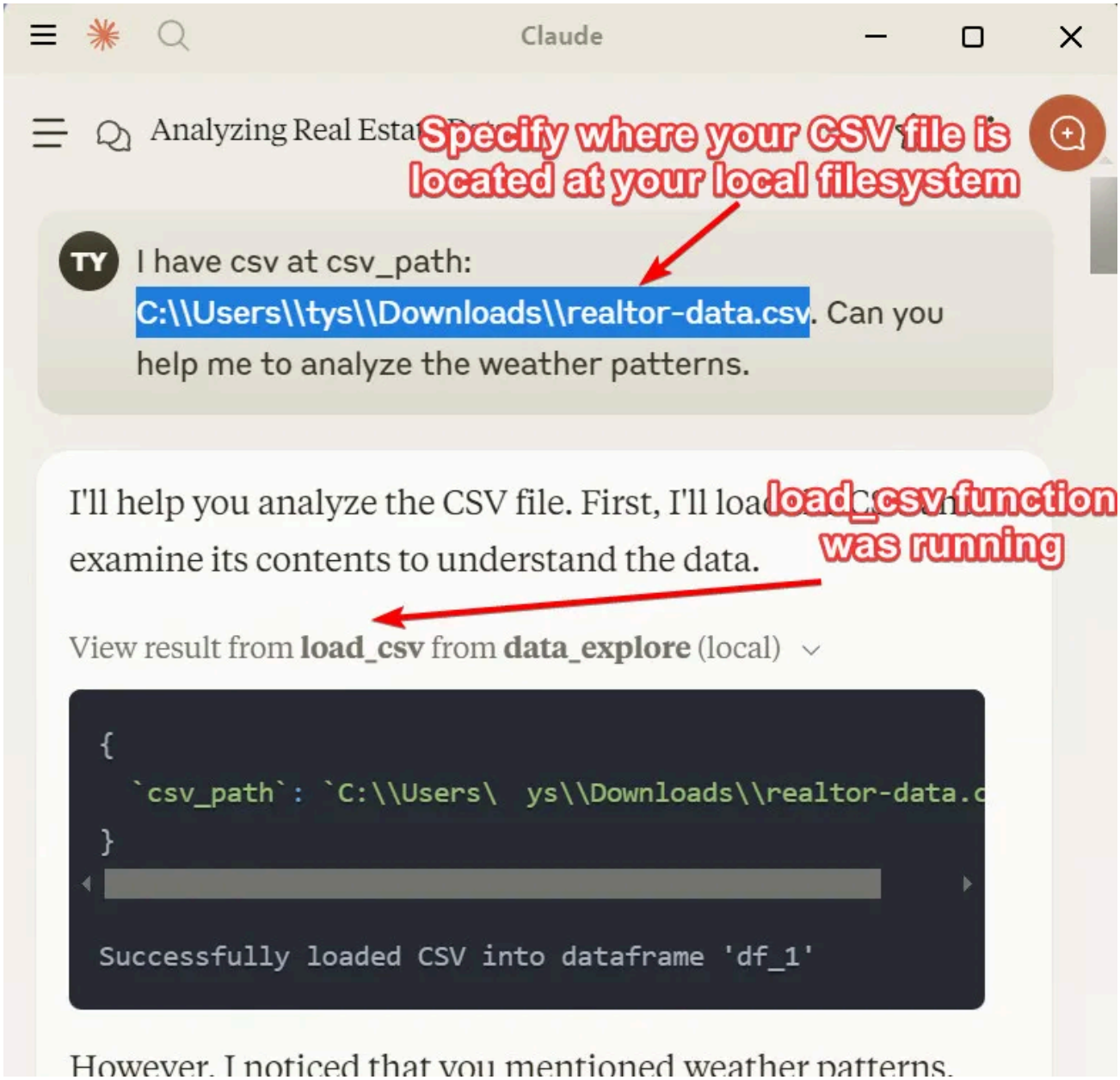
The claude_desktop_config.json file is where you'll configure your MCP servers (Windows). This file is typically located at:

- Windows: C:\Users\YourUsername\AppData\Roaming\Claude\claude_desktop_config.json

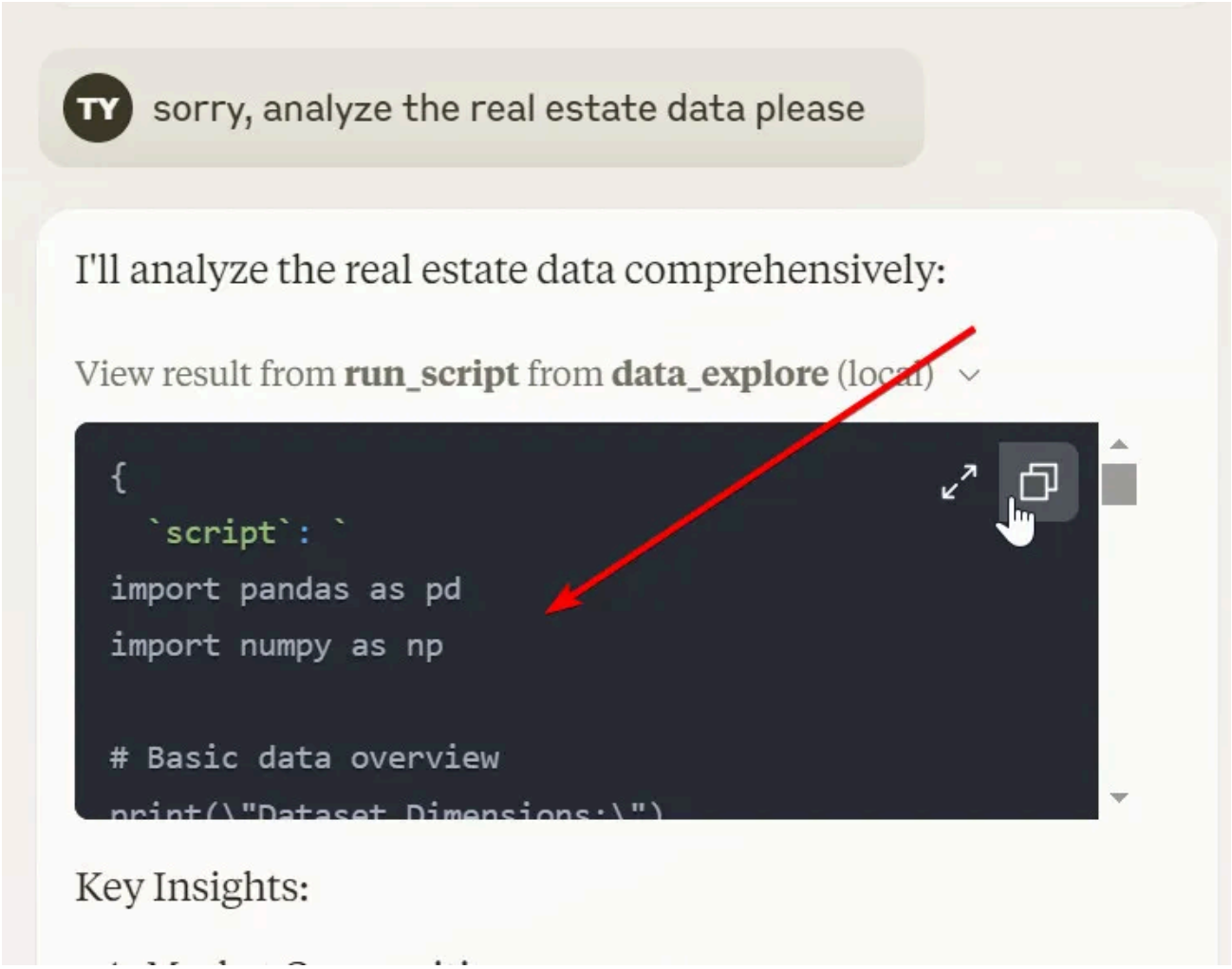
The server directories (where the MCP servers are installed) can be in any location; often in these two locations:

2.3 Try the results

(Note: try to change the prompt, the dataset is about US real estate, not weather patterns)

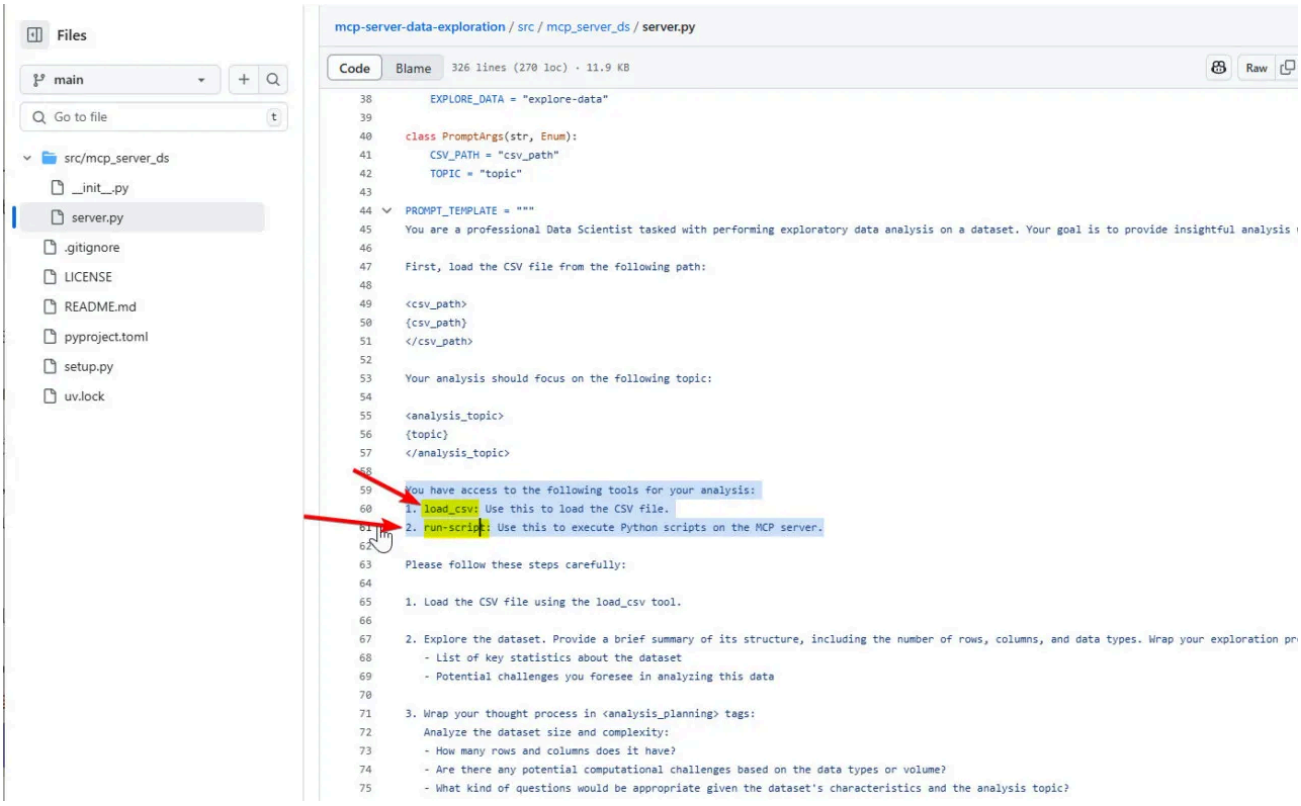


It use run_script module to run python script:



Part 3: Commentary on this approach

3.1 A Quick Look



3.2 Benefits

According to Dwarves Memo, below are a few benefits we could enjoy if using MCP Server to allow AI agents to connect with a variety of platforms and data sources using a standardized framework:

- **Standardized interactions**: MCP gives AI a consistent approach to access and use data across platforms, local resource and more.
- **Local-first architecture**: MCP prioritizes security, ensuring sensitive data is secured while providing frictionless processes.
- **Open-source innovation**: MCP creates a robust context which supports continuous enhancement and innovation.

Source: Dwarves Memo

3.3 Limitations

- If I add multiple MCP servers, when doing question answering, they might confuse which tools to use
- I am currently using Claude Free, so I will hit limits easily.
- Can only analyze a single CSV file. But, there are other MCP servers for database like for **MCP servers for SQLite** and **MCP servers for PostgreSQL** which I haven't tried it yet.

References

- **Data Exploration Server** https://github.com/reading-plus-ai/mcp-server-data-exploration
- **Model Context Protocol – Quickstart For Claude Desktop Users** https://modelcontextprotocol.io/quickstart/user
- **Model Context Protocol – Quickstart For Server Developers** https://modelcontextprotocol.io/quickstart/server
- **Use Claude Desktop and MCP Servers to Automate Your Desktop & Coding Workflow** https://medium.com/@mknebel/how-to-automate-

your-workflow-wtih-claude-desktop-and-mcp-servers-5072844b86d1

- **Exploring Model Context Protocol (MCP) with Claude Desktop: Simplifying AI Integration** https://blog.stackademic.com/exploring-model-context-protocol-mcp-with-claude-desktop-simplifying-ai-integration-e447087f95a1
- **How to Use MCP Tools on Claude Desktop App and Automate Your Daily Tasks** https://medium.com/@pedro.aquino.se/how-to-use-mcp-tools-on-claude-desktop-app-and-automate-your-daily-tasks-1c38e22bc4b0
- **How to use Anthropic MCP Server with open LLMs, OpenAI or Google Gemini** https://www.philschmid.de/mcp-example-llama
- **LLM Function-Calling vs. Model Context Protocol (MCP)** https://medium.com/@patc888/function-calling-vs-mcp-model-context-protocol-fc01e9c41eb4
- **Model Context Protocol vs Function Calling: What's the Big Difference?** https://www.reddit.com/r/ClaudeAI/comments/1h0w1z6/model_context_protocol_vs_function_calling_whats/
- **MCP enables Claude to Build, Run and Test Web Apps by Looking At Screenshots** https://wonderwhy-er.medium.com/mcp-enable-claude-to-build-run-and-test-web-apps-using-screenshots-3ae06aea6c4a
- **True Agentic RAG: How I Taught Claude to Talk to My PDFs using MCP (Model Context Protocol)** https://medium.com/@adkomyagin/true-agentic-rag-how-i-taught-claude-to-talk-to-my-pdfs-using-model-context-protocol-mcp-9b8671b00de1
- **Why Agentic RAG Is the Next Big Thing in AI** https://www.scalytics.io/blog/agentic-rag-next-big-thing-in-ai
- **Boost your Productivity with these Game-Changing GenerativeAI Tools!** 💡 https://www.linkedin.com/posts/riyazahd_ai-generativeai-rows-activity-7293470107320614913-EyTx

Search

Yong Sheng Tan

View Profile

Hi, I'm Yong Sheng. I am currently a marketing executive at ShareInvestor. I am good at utilizing tools to fulfil and achieve my work such as using ChatGPT for marketing copy, Mailchimp for email marketing, and some coding works (mostly in python) to automate some of task flows.

Leave a Reply

Logged in as Yong Sheng Tan. Edit your profile. Log out? Required fields are marked *

Add Comment *

Subscribe to our newsletter to receive the latest news and offers from us.

I accept the Privacy Policy

Post Comment

Subscribe to our newsletter to receive the latest news and offers from us.

Email

First Name

Last Name

Subscribe

Edit Form

Privacy Policy | Cookie Policy