Where to find datasets

official documents/tutorial

official document is the most direct and right way to get the dataset/APIs/other information.

Stack Overflow

If you do not know where to go, try stack overflow

For stocking related data:

yahoo finance

how to collect yahoo finance data

yfinance offers a threaded and Pythonic way to download market data from Yahoo! finance.
yfinace
For example :

```python
import yfinance as yf
# valid intervals: 1m,2m,5m,15m,30m,60m,90m,1h,1d,5d,1wk,1mo,3mo
data = yf.download(tickers="^GSPC", interval="1d", start="1988-01-01", end="2023-12-31")
data.to_csv('./out.csv')
```

How to deal with data

package: pandas

pandas
Pandas is an extension of the Python language library for data analysis.

Pandas is an open source, BSD-licensed library that provides high-performance, easy-to-use

data structures and data analysis tools.

Pandas is a powerful tool set for analyzing structured data, based on Numpy (which provides high-performance matrix calculations).

Pandas can import data from various file formats such as CSV, JSON, SQL, and Microsoft Excel.

Pandas can perform computations on various data, such as merging, reshaping, and selecting, as well as data cleaning and data processing characteristics.

Pandas is widely used in various data analysis fields, such as academia, finance and statistics.

A official brief user guide of pandas

Using pandas to load csv data

The data structure in pandas

`Series` can be regarded as one dimension array, and `Datafram` can be regarded as two dimension array(matrix).

```
import pandas as pd
df = pd.read_csv('data/table.csv')
df.head() # The head(n) method is used to read the first n rows. By default, 5 rows are returned i
df.to_csv('data/new_table.csv') # write the dataframe to csv files
```

Detailed method can be shown in Series and Dataframe.

However, these documents are too detailed for beginners.
official user guide may helps, but they are still a little tedious.
A pretty good Chinese version tutorial.
A not bad tutorial.