To achieve our project goals, the first and most fundamental step is to gather our data, which includes layoff data and stock data.

Between them, stock data is the easier one to get, for there is already a well-established Python open-source package that allows us to download and process stock price data from Yahoo Finance. What we should input is just the company list whose data you want to get. Based on that, what we should do first is to explore the other part, the layoff data, to see which companies' data we can obtain. From there, we will determine our company list.

We have a web: [*https://layoffs.fyi/*](https://layoffs.fyi/), from which we get the layoff data. 图形用户界面, 应用程序

描述已自动生成

However, we encountered obstacles when trying to obtain data from the website through web scraping or API access. We initially attempted to use the Airtable website's API (as the layoffs table is hosted on the Airtable platform) to retrieve the data. Unfortunately, the creator of the layoffs table has restricted this method.

图形用户界面, 应用程序, Word

描述已自动生成

Next, we attempted to use the web scraping techniques we had learned to obtain the data. Below are the methods we used and the results we achieved.

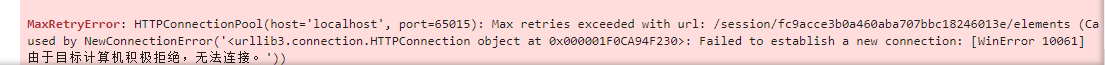


文本

低可信度描述已自动生成

手机屏幕截图

描述已自动生成



The Chinese sentence in the image means:

The connection was actively refused by the target machine.

We refers to the ChatGPT for the reason of this error, and it replies:  
图形用户界面, 文本, 应用程序

描述已自动生成

Therefore, we decided to use a different approach to obtain this data.

ChatGPT has developed powerful text recognition capabilities. So, we took screenshots of the data we needed from the layoffs website and used GPT to convert the text into a table.

Finally, we successfully obtained the layoff data. After cleaning and processing, we also generated the Company list we needed.

图形用户界面, 表格, Excel

描述已自动生成

图片包含 图形用户界面

描述已自动生成

Once we have the company list. The stock data is on its way.

The **“yfinance”** package is a Python library that provides a simple interface for retrieving financial data from Yahoo Finance. It is widely used for accessing historical stock prices, real-time market data, dividends, splits, and other financial metrics.

However, retrieving data from Yahoo Finance requires the company's ticker rather than its name. Therefore, we created a dictionary to add this information.

文本

描述已自动生成

Then, using the ticker\_list we just obtained, we retrieved another major piece of data we needed: the stock data.

图片包含 表格

描述已自动生成

Next comes the data cleaning process, and we obtained the results as shown in the figure.

文本

描述已自动生成