

**CHAT APT**

**BSID-3 吳家可歸**

# System Architecture

**FRONT END  
FLUTTER**

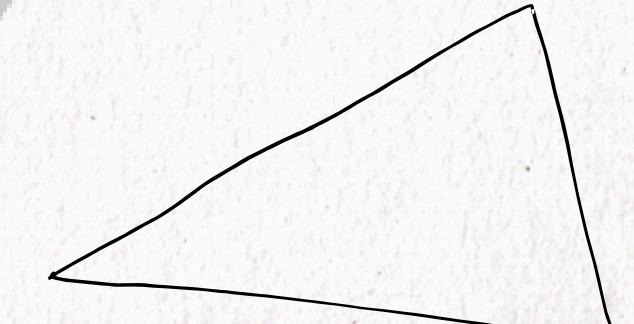
Message, Login credentials

Respond, Token, Permissions

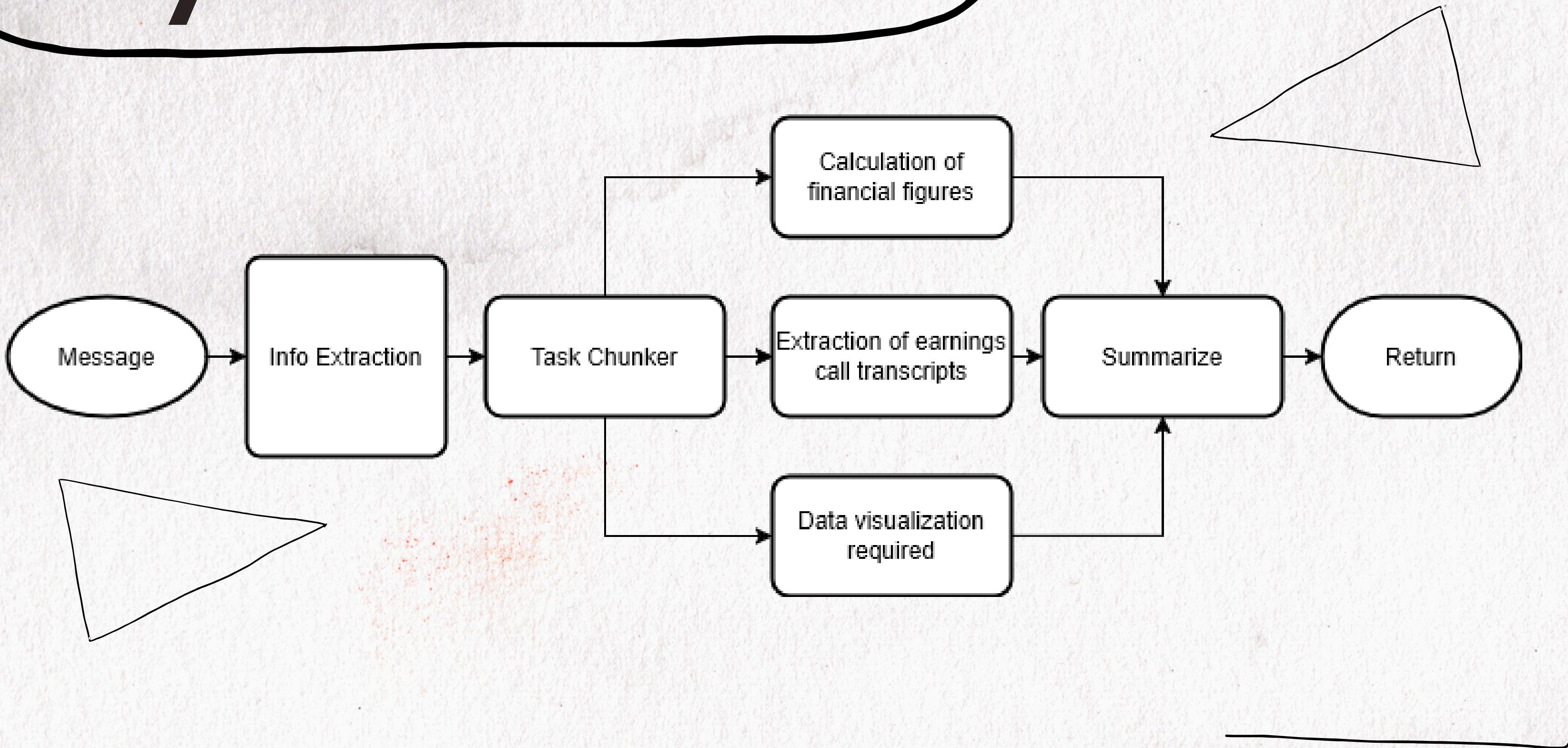
**BACK END  
FASTAPI**

Message Displaying

Login Authentication

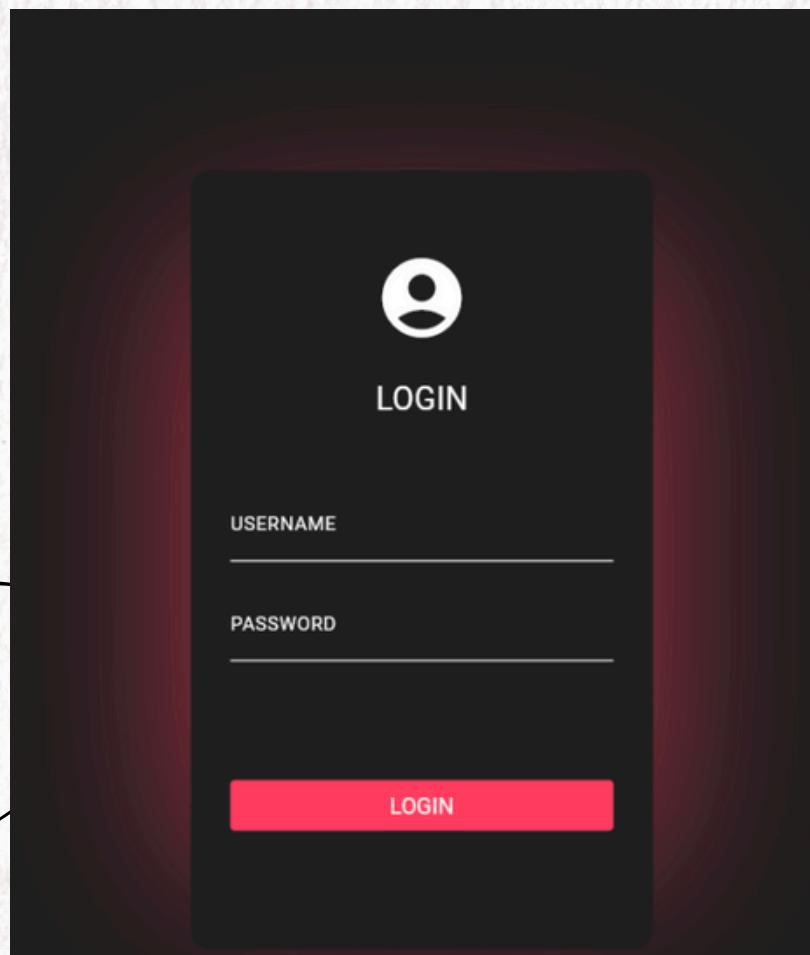


# System Flow



# FastAPI

我們使用 fastapi 進行 login 以及提供前端 request  
登入的部分使用 token/permission/md5



## Financial Data Analysis Backend

[/openapi.json](#)

### default

**POST** /login Login

**POST** /query Process Query

**POST** /query2 Process Query2

# Info Extraction

- 把所有特徵換成標準特徵（強制轉成英文問題，使用歷年以及 USD）
- 根據使用者 permission 篩選可使用資料
- 根據可使用資料，若問句中含有不可使用資料，更改使用者輸入（未授權部分拿掉）

# Task Chunker

根據使用者的問題，我們區分了三種問題

- 計算財務數值相關（用 LLM 生成程式碼運行）
- 撈取法說會稿，並且分析資料來源
- 資料須視覺化（用 LLM 生成程式碼運行）

並且，這三種問題是可重疊的，會區分該問題的關鍵字

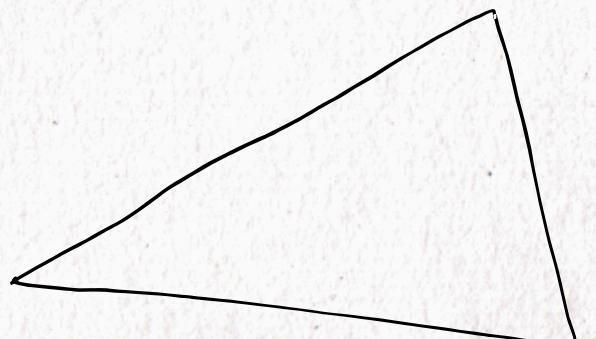
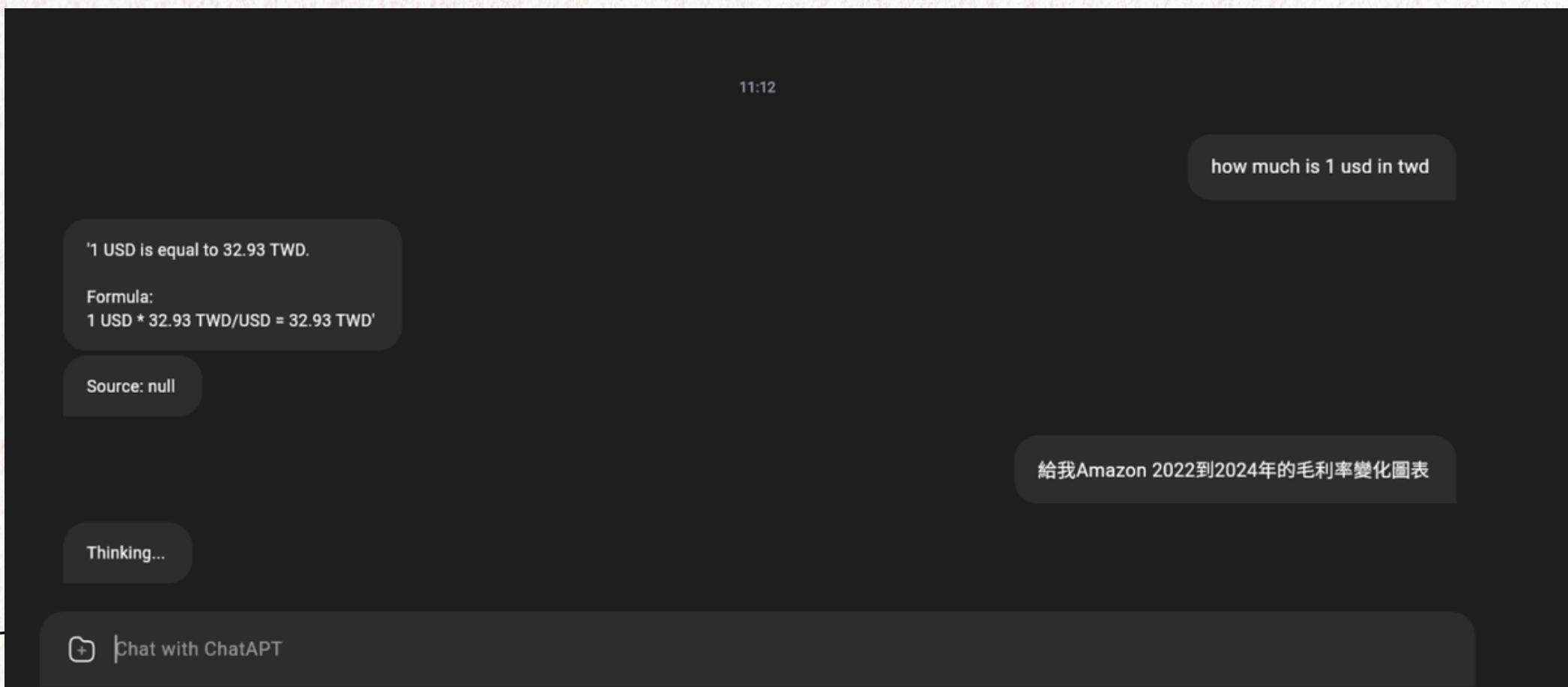
若無這三種，則直接用 LLM 回答使用者問題（例如：你好）

# Hallucination

- 把資料統一化，使得處理上較為輕鬆
- 因為使用生成程式碼來產生數值，所以資料較為準確。（不是由語言直接運算）
- 並沒有使用 vector search，不會對完整資料截斷喪失完整性
- 在使用資料時，會讓 LLM 生成對應的 source
- 把未授權的公司或不存在的資訊，進行問題改寫，使模型不會去生成不存在內容。

# Summarization

- 根據先前所有任務的 Chunker 結果，我們會給 LLM 進行彙整
- 並且根據先前使用者使用的語言，把彙整結果改成該語言



# Pulldown Menu

另外一個介面，我們提供使用者選擇三個參數（公司 年分 季度），和先前的不一樣的是，我們這邊 LLM 選擇陳述內容，以及是否需要資料視覺化

The screenshot shows a dark-themed web application interface. At the top, there is a navigation bar with a menu icon and the title "ChatAPT". Below the title, there are four dropdown menus: "Amazon" (selected), "at quarter", "One", and "in 2022". To the right of these dropdowns is a "Search" button. Below the search bar, the word "Description" is centered. A large text box contains the generated text:

```
'The data provided indicates a situation where an attempt was made to analyze or report on certain business aspects such as financial data trends, product strategy, market competition, or future outlook; however, the document was found empty. This situation highlights the importance of ensuring data integrity and availability for comprehensive analysis and reporting in business contexts.\n\nGiven the absence of data, visual representations can still be generated to illustrate the situation. For example, a pie chart can be created with a single slice labeled "none," signifying the lack of data. This can be accomplished using the `matplotlib` library in Python. Here is a conceptual flow:\n\n1. Define a list containing only one value representing 100% of the distribution. Assign the label "none" to this slice.\n2. Use `matplotlib` to generate and save the pie chart as an image file in the `./img` directory.\n\nAnother visualization could be a timeline (Gantt chart) to represent the expected data input schedule:\n\n1. Use `matplotlib` to create a Gantt chart that shows tasks or data collection events along a timeline. This helps visualize their scheduling and duration.\n2. Arrange tasks sequentially to avoid overlaps that could result in void outcomes.\n3. Save the resulting chart as an image in the `./img` directory.\n\nFor scenarios anticipating missing data, a bar graph can be developed:\n\n1. Define standard categories such as financials, products, and market metrics. These categories could potentially be where data might be missing.\n2. Plot a bar for each category, possibly using random or placeholder data to visualize the missing areas.\n3. Use `matplotlib` to create the chart and save it as an image in the `./img` directory.\n\nThese steps outline how visualizations can be crafted using Python's `matplotlib` library, even in the absence of substantial data, to aid in depicting missing information and planning future data-related tasks or reports.'
```

# DEMO1

≡ ChatAPT

Source: num



給我Amazon 2022到2024年的毛利率變化圖表

要計算2022年至2024年的毛利率變化，必須擁有這些年度的收入、銷售成本（COGS）以及任何相關的價格變化數據。提供的數據僅包括美元和新台幣的匯率。若沒有關於2022年和2024年的收入和銷售成本的具體財務數據，就無法準確計算毛利率變化。對於涉及不同貨幣操作的計算，匯率可能是必要的，但它們本身並未提供足夠的毛利率計算信息。

毛利率公式：

$$\text{毛利率} = (\text{收入} - \text{銷售成本}) / \text{收入} * 100$$

視覺化：

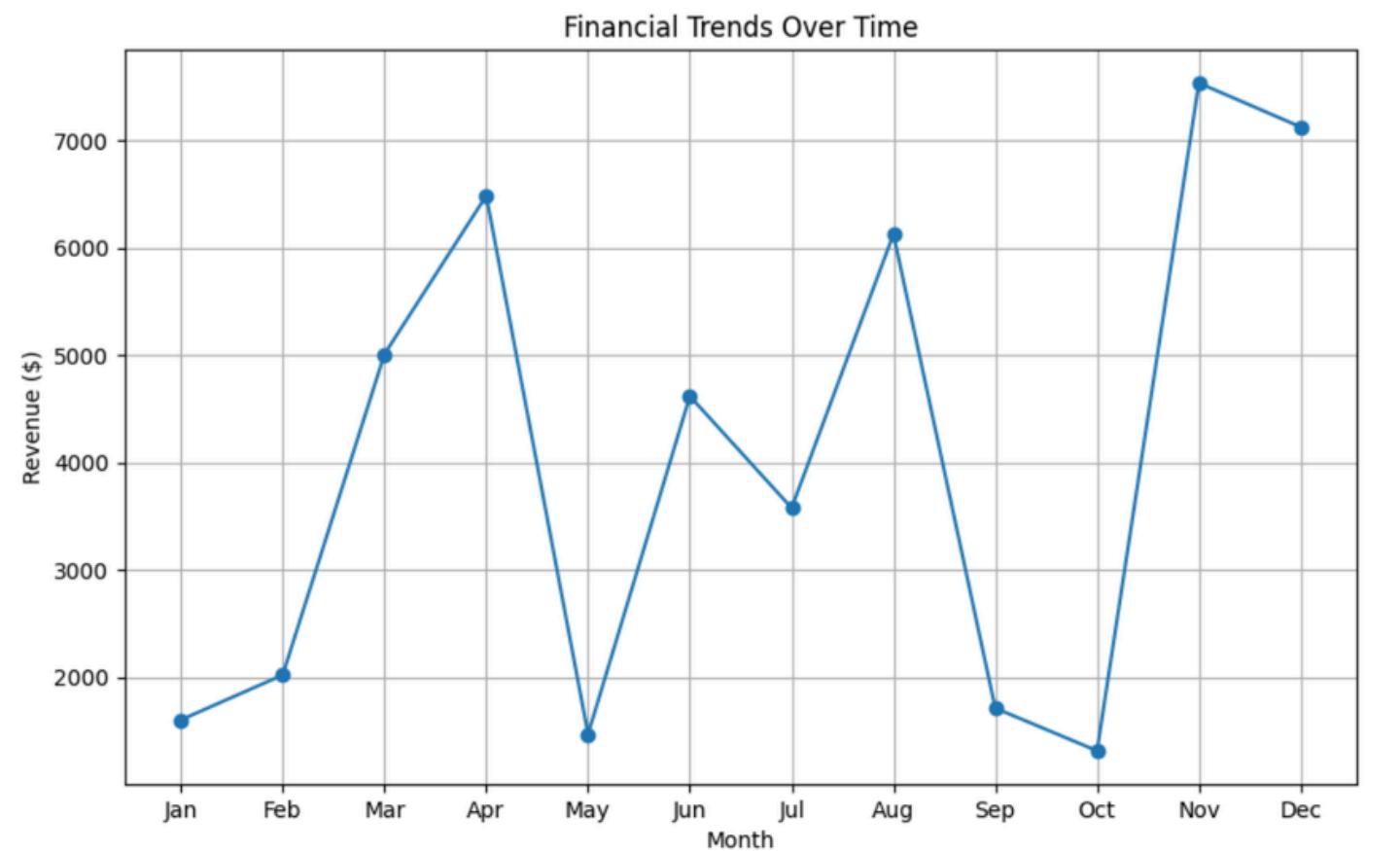
要生成2022年至2024年毛利率變化的折線圖，可以使用Python中的matplotlib庫。您需要定義年份和相應的毛利率值，然後使用折線圖繪製這些值。最後，將圖表保存為名為'gross\_margin\_changes.png'的圖像。



Chat with ChatAPT

# DEMO2

## Charts



Amazon at quarter One in 2022 Search

ChatAPT Amazon at quarter One in 2022 Search

Description

The data provided indicates a situation where an analysis was made to analyze or report on certain business aspects such as financial data trends, product strategy, market competition, or future outlook. However, the document used for analysis highlights the importance of ensuring data integrity and accuracy. It is crucial to verify the source of data and ensure it is up-to-date. This document also emphasizes the need for a clear understanding of the context in which the data is presented. For example, a pie chart can be created with a single data label 'Total' if the intent is to represent the whole 'Total' for this slice. Instead, it's recommended to generate and use the chart as a single data label 'Total', signifying the lack of detail. This can be accomplished using the 'Total' label in the chart's title or subtitle. Another important aspect is the use of descriptive labels for each data series. For instance, 'Revenue' and 'Profitability' are good labels for a line chart. These labels help measure their scheduling and duration. Average sales quarterly to avoid overlaps that could result in misleading data. The chart also includes a legend to identify the data series. These steps outline how modifications can be made using Python's Matplotlib library to ensure the accuracy of analytical data and its usefulness in planning future data-related tasks or reports.

Charts

Financial Trends Over Time

This line chart shows revenue trends from 2021 Q1 to 2021 Q4. The Y-axis represents Revenue (\$) ranging from 100 to 300. The X-axis represents the quarters from 2021 Q1 to 2021 Q4. The data shows a steady upward trend with a slight dip in Q2 and Q3.

Quarter	Revenue (\$)
2021 Q1	150
2021 Q2	180
2021 Q3	160
2021 Q4	200

Product Sales Performance

This bar chart compares sales volume for five products. The Y-axis represents Sales Volume ranging from 0 to 1000. The X-axis lists products A through E. Product C has the highest sales volume, followed by Product D and Product B.

Product	Sales Volume
Product A	700
Product B	700
Product C	1000
Product D	800
Product E	800

Trends in Revenue and Profitability Over Time

This line chart tracks two metrics over time: Revenue (blue line) and Profitability (orange line). The Y-axis represents both metrics in dollars, ranging from 100 to 300. The X-axis represents the quarters from 2021 Q1 to 2021 Q4. Both metrics show a consistent upward trend over the period.

Quarter	Revenue (\$)	Profitability (\$)
2021 Q1	150	100
2021 Q2	180	120
2021 Q3	160	110
2021 Q4	200	130

Comparison of Product Performance Metrics

This bar chart compares four products across three performance metrics. The Y-axis represents the metric value ranging from 0 to 500. The X-axis lists products A through D. Product C consistently outperforms the others across all metrics.

Metric	Product A	Product B	Product C	Product D
Revenue	450	400	500	480
Profitability	350	300	420	380
Growth Rate	250	200	300	280

Market Share Distribution Among Competitors

This pie chart illustrates the market share distribution among four competitors. The total market share is 100%. Competitor A holds the largest share at 35%, followed by Competitor C at 30%, Competitor B at 20%, and Competitor D at 15%.

Competitor	Market Share (%)
Competitor A	35%
Competitor B	20%
Competitor C	30%
Competitor D	15%

Scatter Plot of Predicted Market Opportunities vs Associated Risks

This scatter plot displays the relationship between predicted market opportunities (X-axis) and associated risks (Y-axis). The X-axis ranges from 0 to 100, and the Y-axis ranges from 0 to 100. The data points show a positive correlation, indicating that higher predicted market opportunities are generally associated with higher risks.

Predicted Market Opportunities	Associated Risks
10	10
20	20
30	30
40	40
50	50
60	60
70	70
80	80
90	90
100	100

***Thank you***