

9. 簡介及比較數據可視化工具

章節內容

- 簡介可視化工具
- Matplotlib
- Seaborn
- Plotly
- Superset
- Tableau
- PowerBI



可視化的四個重點

- 為什麼選擇可視化？
- 可視化什麼？
- 誰是觀眾？
- 如何可視化？

溝通的形式

1. 口頭交流

- 敏事
- 談話

2. 書面交流

- 紙質檔
- 電子文件
- 社交媒體

3. 物理通信

- 身體姿勢、眼神交流、面部表情
- 情感、語氣、語調、音調

4. 視覺傳達

- 圖表，插圖，繪圖，標誌
- 電子資源

文字與圖表的對比

主頁 > 教育與就業 > **失業率跌至3.5%**



失業率跌至3.5%

2023年1月19日

政府統計處公布，去年10月至12月經季節性調整的失業率為3.5%，較9月至11月的3.7%下跌0.2個百分點；就業不足率下跌0.1個百分點至1.5%。

期內總就業人數增加約8,300人至3,665,300人，總勞動人口則減少約4,400人至3,791,300人。

失業人數減少約12,700人至126,000人，就業不足人數也減少約4,400人至55,300人。

按行業分析，與消費及旅遊相關行業合計的失業率下跌0.3個百分點至4.7%。大多數其他行業，尤其是建造業、運輸業和教育業的失業率也下跌。

勞工及福利局局長孫玉菡表示，勞工市場短期內應會繼續改善。隨着政府取消大部分社交距離措施，以及香港與內地逐步通關，本地經濟活動料更顯著恢復，為勞工市場提供進一步支持，但金融狀況收緊可能抵銷部分正面影響。



Hong Kong SAR (China): Unemployment Rate		
Mnemonic	LBR.IHKG	
Unit	% 3-mo. MA, SA	
Adjustments	Seasonally Adjusted	
	Monthly	
Data	Feb 2023	3.3
	Jan 2023	3.4

Series Information	
Source	Census and Statistics Department Hong Kong Special Administrative Region (HKSARG)
Release	Employment Situation - Monthly
Frequency	Monthly
Start Date	10/31/1981
End Date	2/28/2023



長江和記實業有限公司 CK HUTCHISON HOLDINGS LIMITED

(Incorporated in the Cayman Islands with limited liability)
Stock code: 1

Chairman's Statement

Following the strong economic rebound in the second half of 2020, global recovery continued during 2021 helped by the accelerated vaccination rollout programs and easing of lock-downs and economic reopenings in many parts of the world. However, the global recovery was not a smooth journey in 2021 as new COVID-19 variants continued to emerge during the year. Different policy responses in various markets led to operational and cost challenges exacerbated by pandemic-related supply-demand mismatches. Nevertheless, the Group was able to respond nimbly to changing economic and business environments in various sectors and geographies and is able to report pleasing results for the full year.

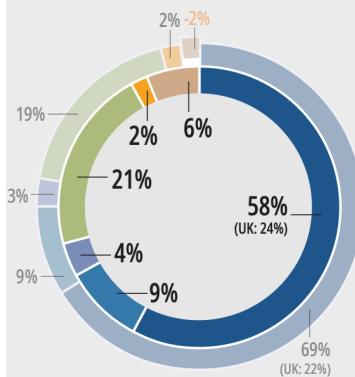
The Group reported EBITDA and EBIT growth of 15% and 20% respectively in reported currency compared to last year was primarily driven by improvements in the Ports and Retail divisions together with positive underlying results of Cenovus Energy⁽¹⁾ as opposed to significant losses reported by Husky Energy in 2020. The results also benefitted from favourable currency translation impacts, as well as the net impact of one-off items in 2021 as compared to those in 2020. These improvements were partly offset by lower contributions from the Telecommunications division as the operating environment remains challenging, particularly in Italy. In local currencies, the Group's reported EBITDA and EBIT grew 10% and 16% respectively from last year.

Analyses by Core Business Segments

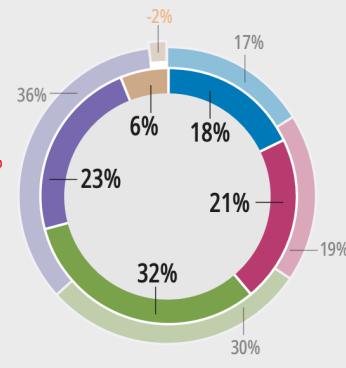
	Pre-IFRS 16 ⁽¹⁾ 2021 HK\$ million		Pre-IFRS 16 ⁽¹⁾⁽²⁾ 2020 HK\$ million		Change %	Local currencies change %
Revenue⁽³⁾						
Ports and Related Services ⁽³⁾	42,285	9%	32,865	8%	29%	26%
Retail	173,601	39%	159,619	40%	9%	6%
Infrastructure	56,100	13%	52,792	13%	6%	-
CK Hutchison Group Telecom	92,575	21%	90,663	22%	2%	-3%
Hutchison Asia Telecommunications	8,786	2%	9,147	2%	-4%	-5%
Finance & Investments and Others	72,036	16%	58,760	15%	23%	20%
Total Revenue	445,383	100%	403,846	100%	10%	7%
EBITDA⁽³⁾						
Ports and Related Services ⁽³⁾	15,157	13%	10,914	12%	39%	35%
Retail	16,034	14%	14,397	15%	11%	9%
Infrastructure	29,636	27%	29,066	30%	2%	-5%
CK Hutchison Group Telecom	43,052	39%	48,540	50%	-11%	-14%
Hutchison Asia Telecommunications	2,036	2%	2,034	2%	-	-1%
Finance & Investments and Others	5,312	5%	(8,007)	-9%	166%	165%
Total EBITDA	111,227	100%	96,944	100%	15%	10%

2021 Total EBIT⁽²⁾⁽³⁾

Reported: HK\$64,744 million
Underlying: HK\$59,965 million



Europe	HK\$34,646 million (Reported: HK\$44,433 million)
Mainland China	6% 18% 26% 50%
Mainland China	8% 24% 33% 35%
Mainland China	63% 30% 8% -1%
Hong Kong	37% 24% 20% 19%
Asia, Australia & Others ⁽¹⁾	HK\$12,507 million
Mainland China	29% 16% 45% 10%
Mainland China	100%
Canada	HK\$1,118 million



- Reported ● Underlying
- Europe
- Mainland China
- Hong Kong
- Asia, Australia & Others⁽¹⁾
- Canada
- Finance & Investments and Others

- Note 1: Includes Panama, Mexico and the Middle East
 Note 2: Prepared under Pre-IFRS 16 basis which is set out in note 1 on page 5
 Note 3: The outer pie chart represents EBITDA and EBIT %-mix on a reported basis. The inner pie chart represents underlying EBITDA and EBIT %-mix, which excludes the gain on disposal of tower assets completed in 2021 of HK\$25.3 billion, non-cash impairment of goodwill of the Group's Italian telecommunication business of HK\$(15.5) billion, impairment charge of Cenovus of HK\$(1.5) billion and a non-cash foreign exchange reserve loss of HK\$(3.5) billion following the Cenovus-Husky merger
 Note 4: Represents EBITDA and EBIT %-mix for Europe on a reported basis
 Note 5: Represents EBITDA and EBIT %-mix for Europe on an underlying basis

- Reported ● Underlying
- Ports & Related Services
- Retail
- Infrastructure
- Telecommunications
- Finance & Investments and Others

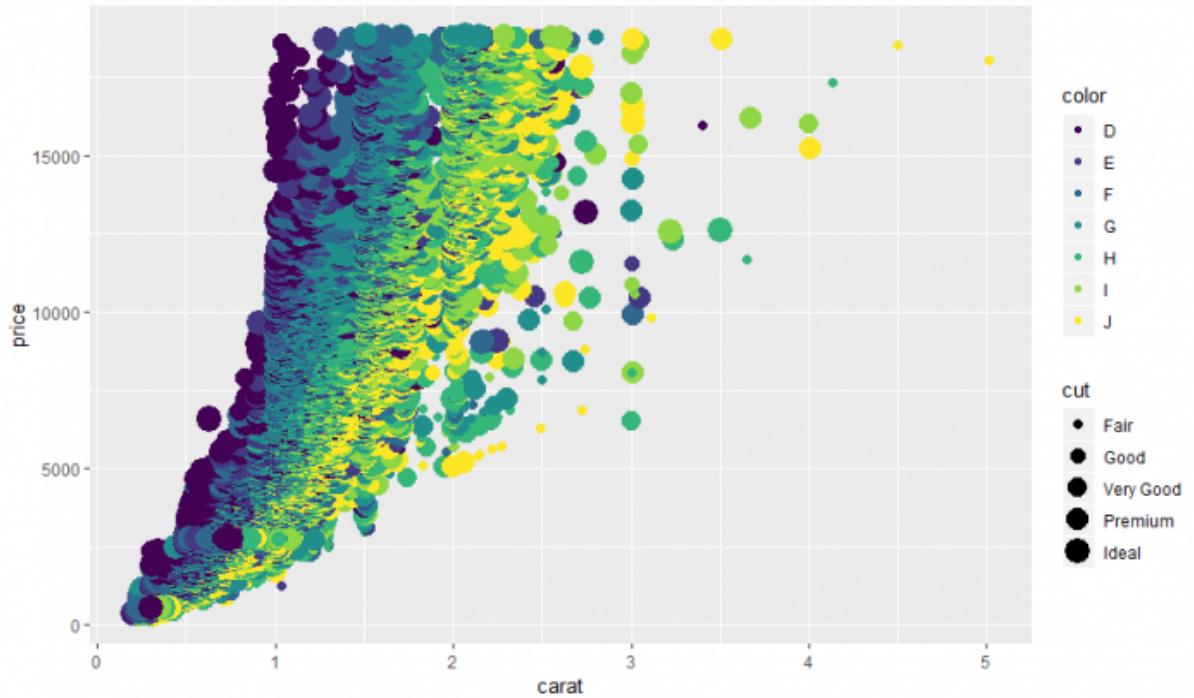
數據分析用途的視像圖表

在數據分析和數據科學領域，數據量是巨大的。我們必須以人性化的方式展示結果。

Univ	Course Name	MS	DMIN	OMA	PROG	CAP	STOR	DMGT	SEC	BIG	PREP	GOV	DEC	COMM	VIS	ETH	CASE	EVAL
CC	Data Mining	x	x	x	x		x				x							
	Introduction to Data Science	x	x	x						x								
	Dataset Organization and Management						x	x	x		x							
	Data Science Capstone		x						x									x
NKU	Data Visualization														x			
	Data Mining	x	x		x											x		
	Big Data						x	x		x								
	Intro to Data Science								x									
	Data Analytics			x										x	x			
	Data Science Capstone			x						x		x	x			x	x	
OSU	Intro to Scientific Visualization															x		
	Introduction to Data Mining	x	x	x						x								
UR	Introduction to Data Mining	x	x						x		x			x			x	
USF	Data Visualization				x	x				x			x			x		
	Data Mining	x	x	x	x						x	x						
	Mathematical Modeling			x														x

Table 10: Mapping of competencies to courses in data science programs

Scatterplot example on diamonds dataset

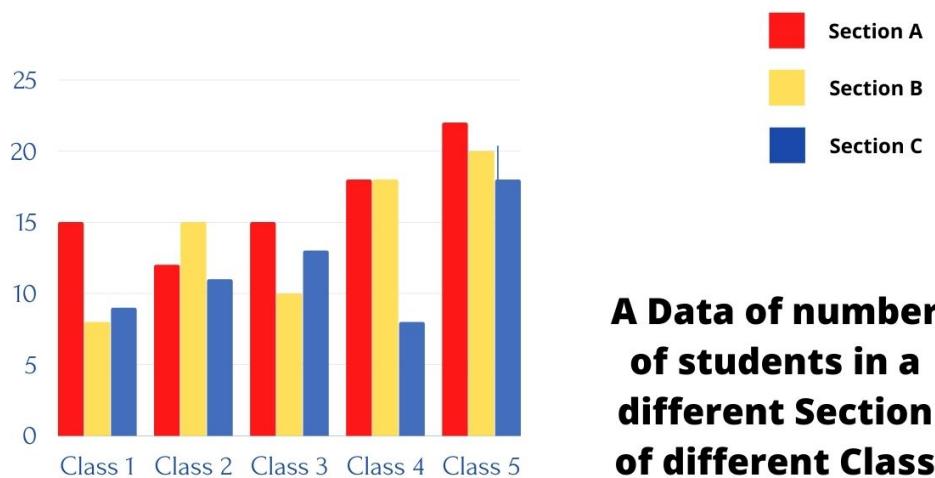


有效數據可視化的關鍵原則

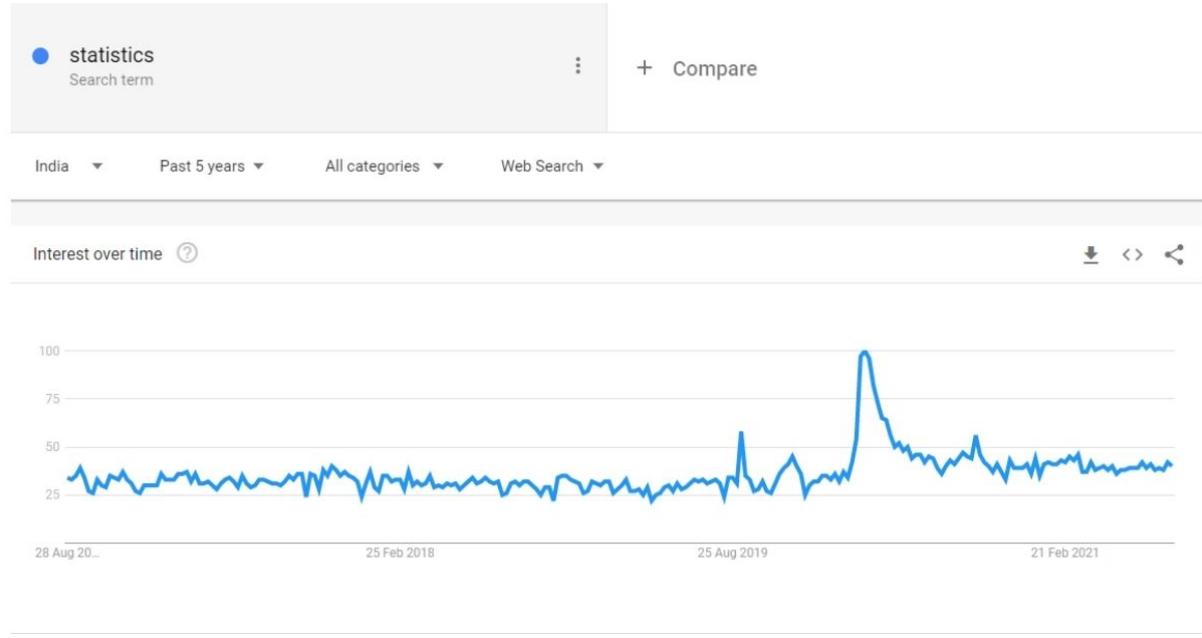
1. 確定最佳視覺物件 **Determine the best visual**
2. 平衡設計 **Balance the design**
3. 聚焦關鍵領域 **Focus on the key areas**
4. 保持簡單 **Keep it simple**
5. 使用模式 **Use patterns**
6. 比較方面 **Compare aspects**

基本圖表種類

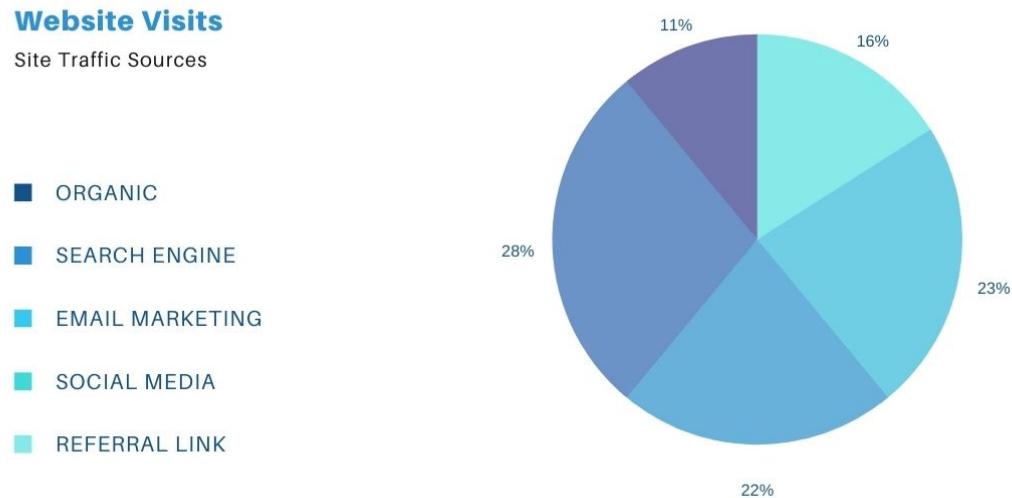
條形圖 **Bar Chart** 用於快速比較不同類別的數據、強調差異、披露歷史高點和低點以及指示趨勢和異常值。



折線圖 Line Chart 用於可視化數據趨勢，通常隨時間變化（例如五年內的股價變化或當月的網站頁面流覽量）

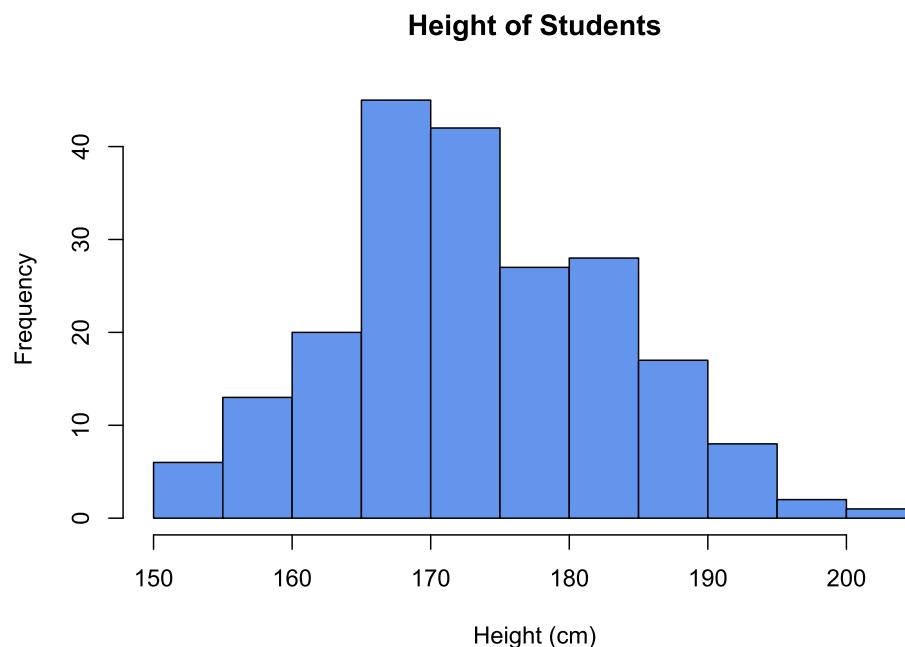


餅形圖 Pie Chart 是圓形統計圖形，它被分成不同的切片，以數位比例顯示各數據。

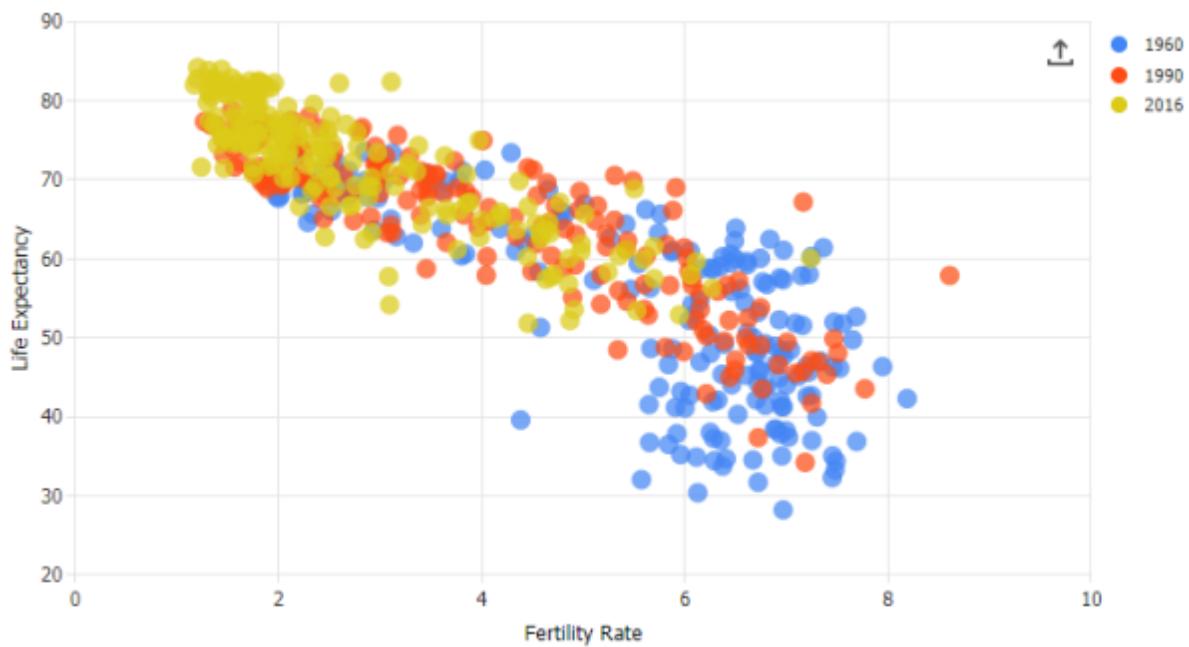


直方分佈圖 Histogram 用於匯總離散或連續的數據，並按區間刻度進行測量。

它經常用於以方便的格式可視化數據分佈的重要特徵。Y-軸可以是頻率、出現次數、或其對數。



分散點圖 Scatter plots 有助於檢查兩個或多個變數之間的關係，揭示一個變數是否是另一個變數的良好預測變數，或者它們是否傾向於獨立變化。



數據分析之可視化工具

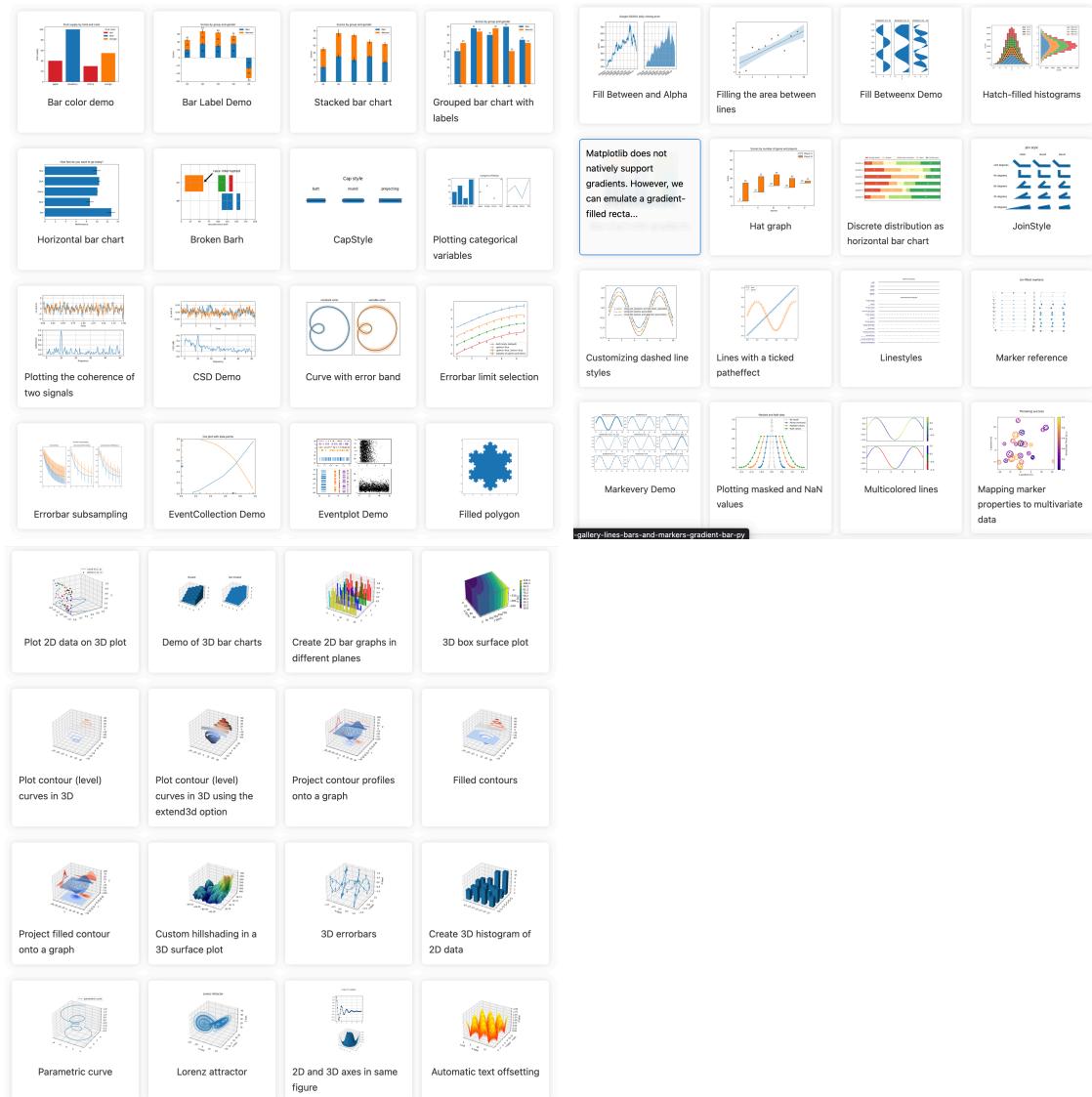


Matplotlib 是 John Hunter (1968-2012) 和該專案的眾多貢獻者努力開發的結果。

Matplotlib 是一個為其用戶維護並由其用戶維護的社區專案。

Matplotlib 是數據分析和數據科學行業中最常用的工具。

Matplotlib 是一個全面的可視化工具，可與 iPython 配合使用。



Matplotlib Bar Chart 條形圖

```
import matplotlib.pyplot as plt

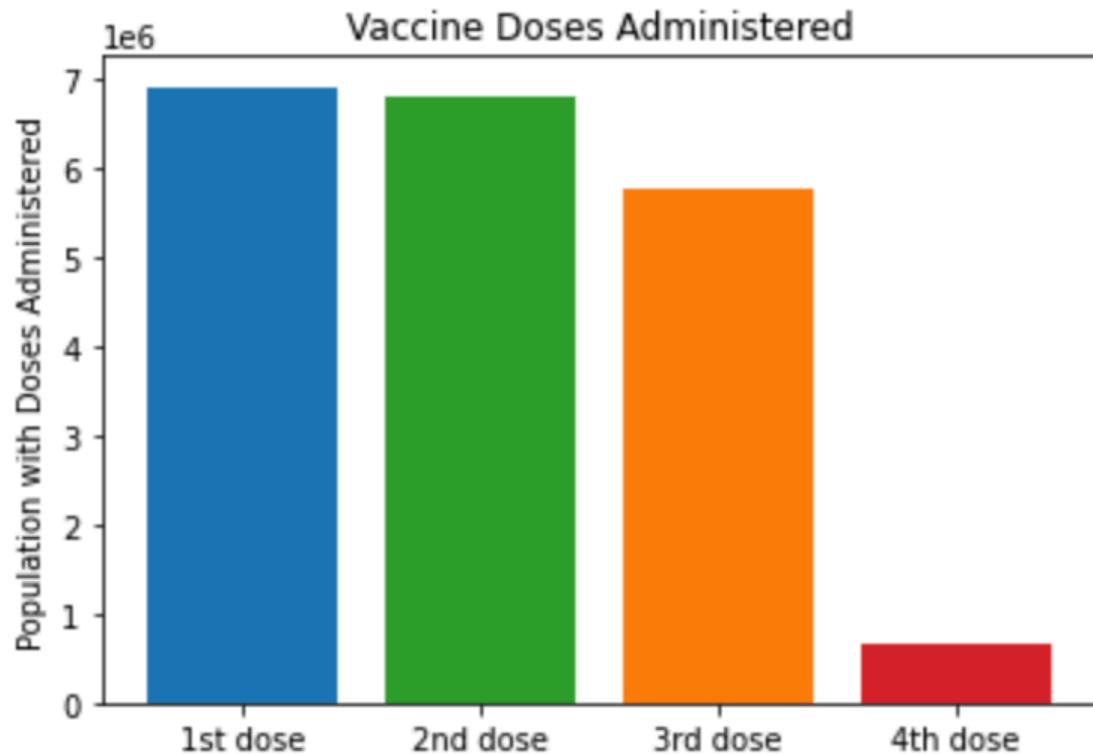
fig, ax = plt.subplots()

vaccine = ['1st dose', '2nd dose', '3rd dose', '4th dose']
counts = [6904243, 6779247, 5764953, 675118]
bar_colors = ['tab:blue', 'tab:green', 'tab:orange', 'tab:red']

ax.bar(vaccine, counts, label=bar_labels, color=bar_colors)

ax.set_ylabel('Population with Doses Administered')
ax.set_title('Vaccine Doses Administered')

plt.show()
```

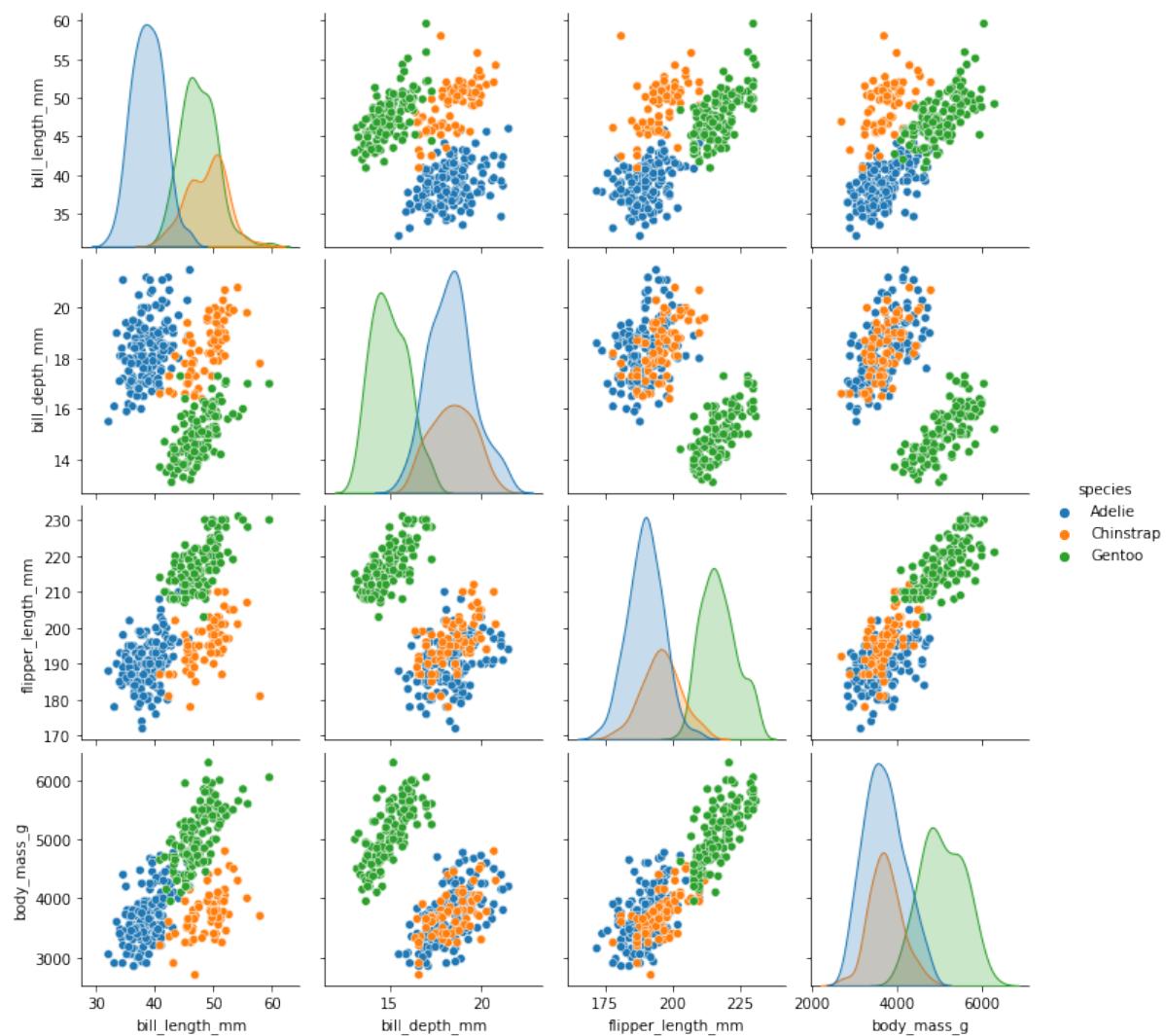


Seaborn 是一個高級介面，建基於 **Matplotlib** 繪製統計圖形。它旨在簡化可視化成複雜數據集的核心部分的編程。

您可以將 **Seaborn** 視為 **Matplotlib** 的外掛程式，以簡化編碼和優良設計。

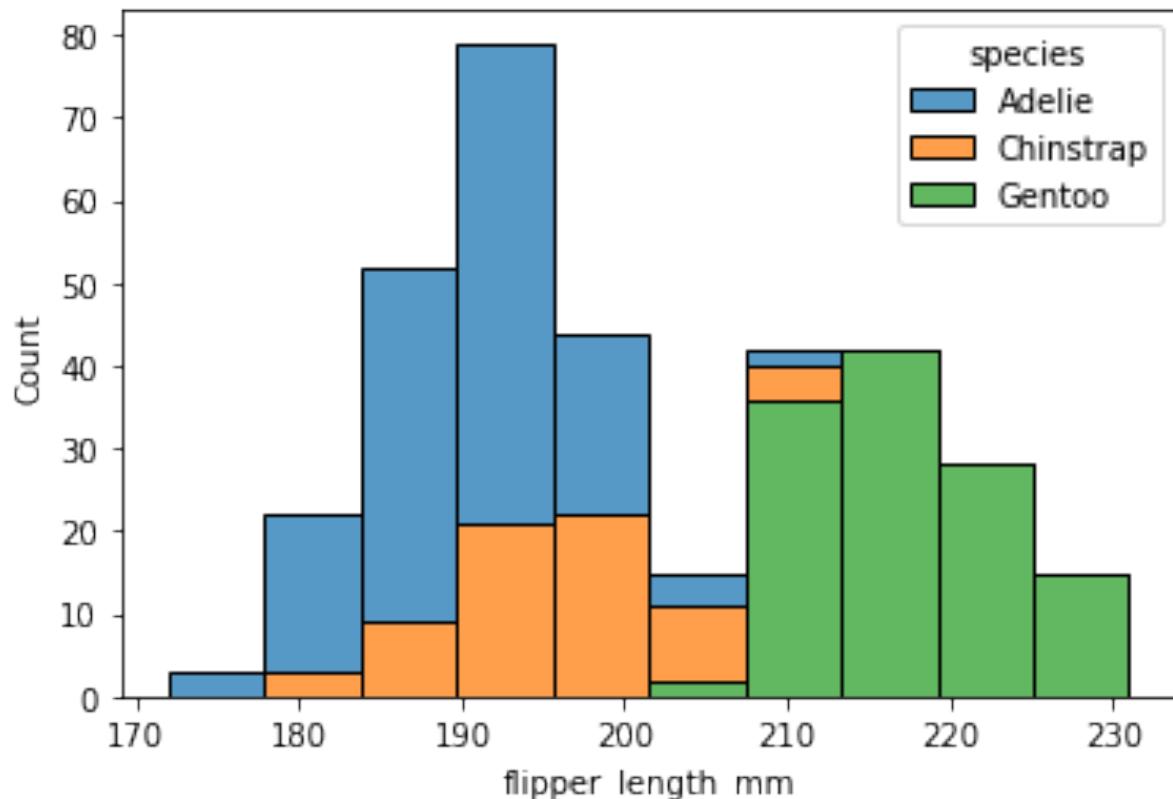
Seaborn 在 **Matplotlib** 起源的簡單代碼庫中設計得更好，這意味著您必須在使用前安裝兩者。

```
1 sns.pairplot(penguins, hue="species")
```



Seaborn 的簡化程式碼

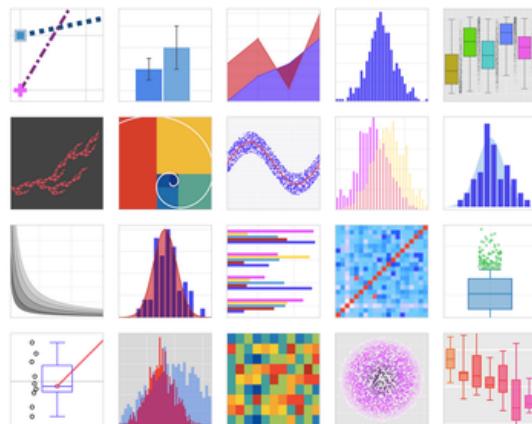
```
1 sns.histplot(data=penguins, x="flipper_length_mm", hue="species", multiple="stack")  
<AxesSubplot:xlabel='flipper_length_mm', ylabel='Count'>
```





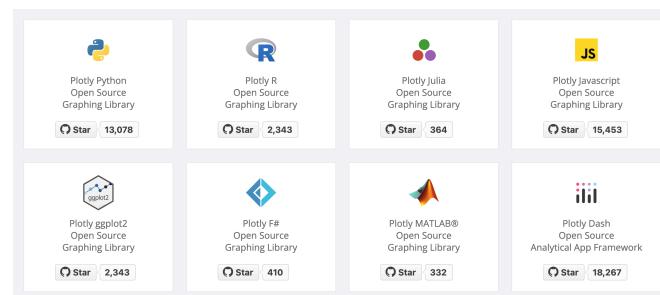
Plotly 建立於 2012 年，比 Matplotlib 相對新。它集中開發在線/離線數據分析和可視化工具。Plotly 為個人和開發人員提供在線圖形、分析和統計工具，以及能用於 Python、R、MATLAB、Perl、Julia、Arduino 和 REST 的科學圖形庫。

由於其簡單的編碼功能和互動式精心設計的圖形可視化，用戶群體正在快速增長！



plotly | Graphing Libraries

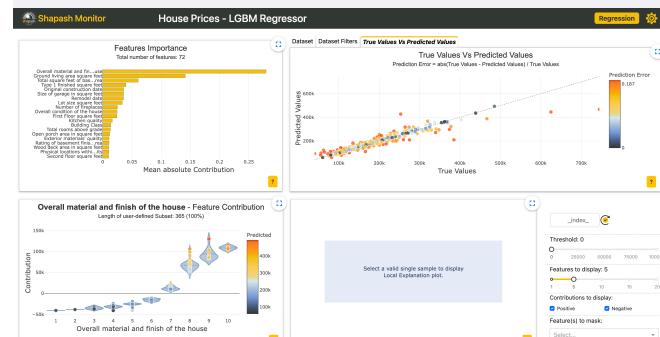
Plotly's Python graphing library 製作互動式、出版品質的圖表。在本章中，我們將重點介紹此模組的數據分析。



plotly | Dash

Dash 是用於使用 Python、R、Julia 和 F#（實驗性）快速構建數據應用的原始低代碼框架。

Dash 基於 Plotly.js 和 React.js 編寫，非常適合構建和部署具有自定義使用者介面的數據應用程式。它特別適合任何處理數據的人。



plotly 最方便的好處之一是它關於懸停的互動式數據。大大增加了展示性。

```
1 fig = px.scatter(df, x="sepal_width", y="sepal_length", color="species",
2                   size='petal_length', hover_data=['petal_width'])
3 fig.show()
```

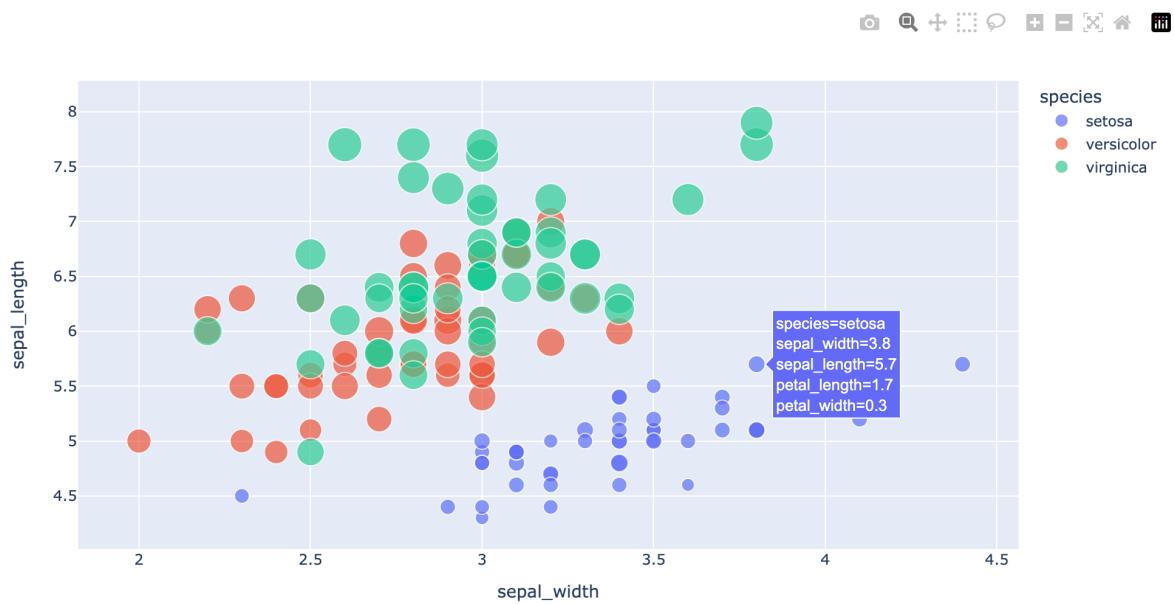




Tableau 是用於數據分析和商業智慧的領先數據可視化工具。

它也是一個儀錶板環境可視化工具。

Tableau 特色

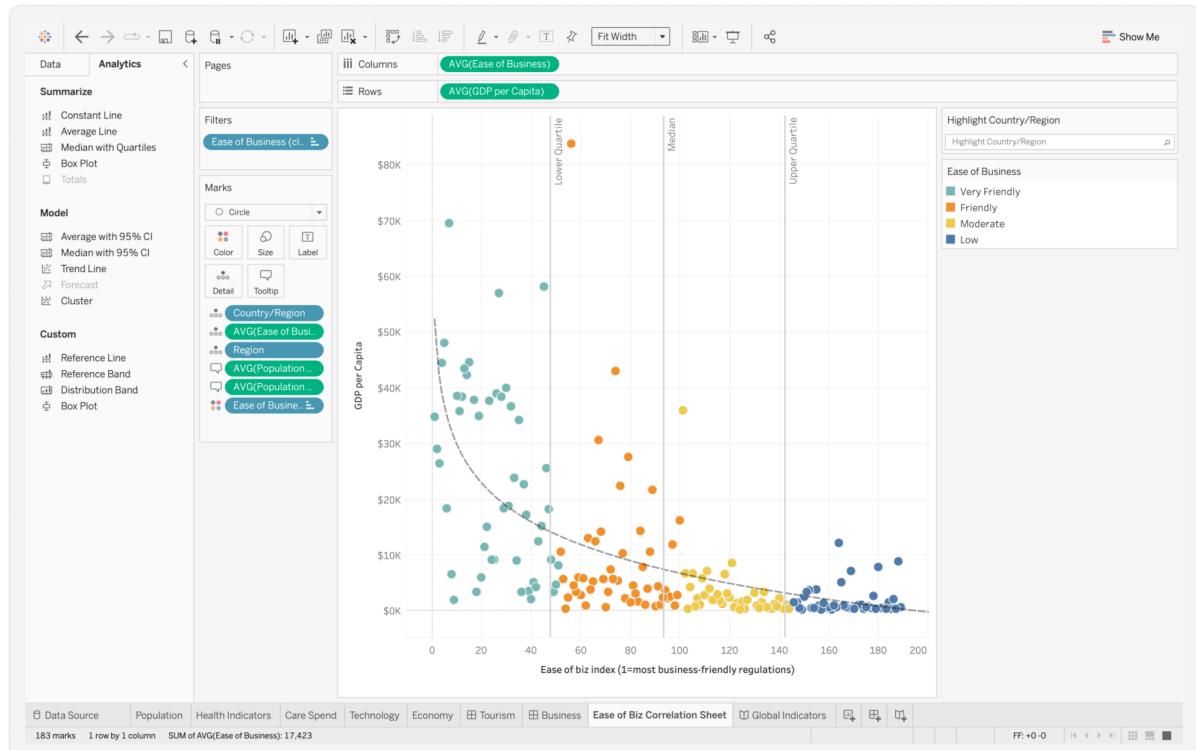
1. Tableau 支援強大的數據發現和探索功能，使用戶能夠在幾秒鐘內回答重要問題
2. 不需要先驗程式設計知識；沒有相關經驗的使用者可以立即開始使用 Tableau 創建可視化效果
3. 它可以連接到其他 Business Intelligence BI 工具不支援的多個資料來源。Tableau 使用戶能夠通過聯接和混合不同的數據集來創建報表

Tableau Server 支援集中位置來管理組織內所有已發佈的數據源

Tableau 通過完全集成的數據管理和治理、可視化分析和數據故事講述以及協作，推動更好的業務成果，所有這些都是 Salesforce 行業領先的內置功能。



Tableau Desktop 交付以訪問、視覺化和分析您的資料。通過直觀的拖放介面，您可以發現隱藏的見解，從而更快地做出有影響力的業務決策。

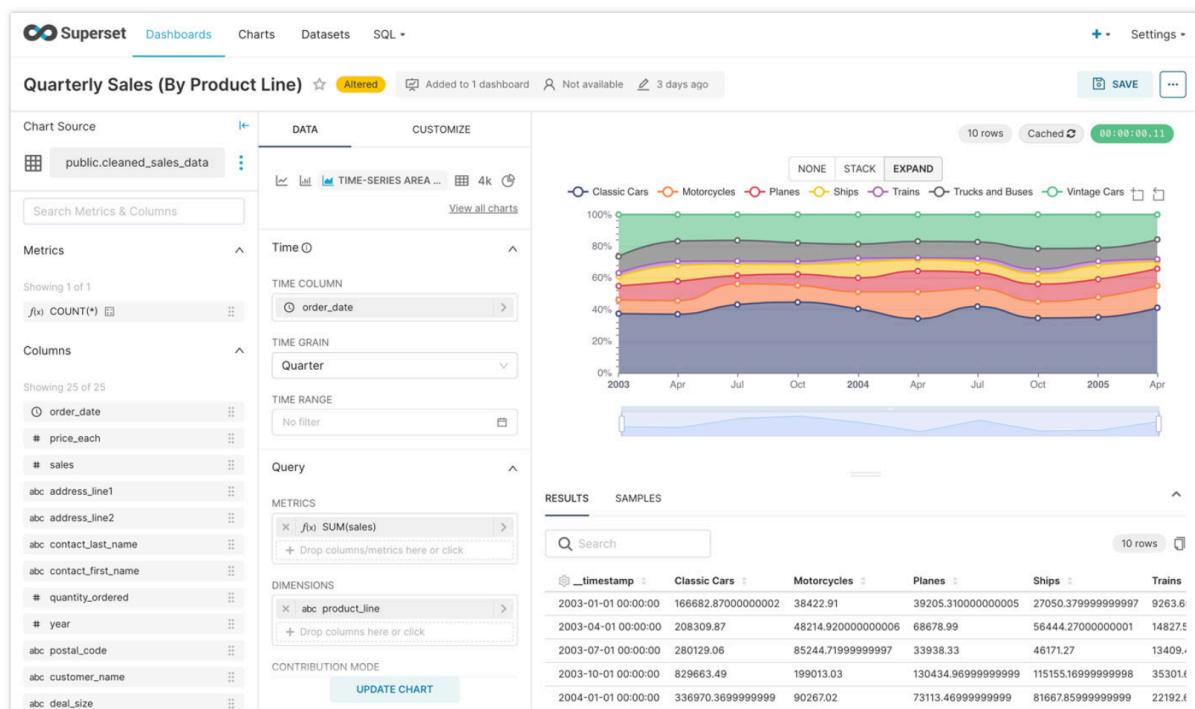




Superset 快速、輕量級、直觀，並載入了各種選項，使所有技能的使用者都可以輕鬆流覽和可視化他們的數據，從簡單的折線圖到高度詳細的地理空間圖表。

特色：

1. 儀錶板創建
2. 企業身份驗證（[OpenID](#)、[LDAP](#)、[OAuth](#) 等）
3. 與 [Apache ECharts](#) 集成
4. 輕量級語義層
5. 可視化外掛程式支援
6. 與大多數[使用 SQL](#) 的數據源相容

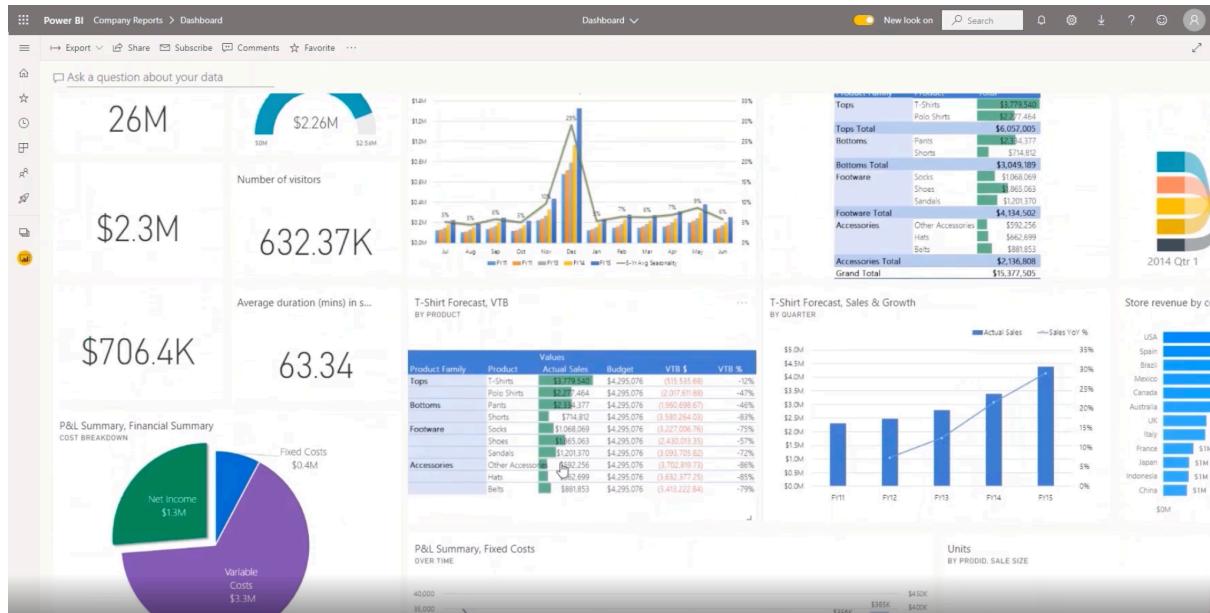


- 使用數據可視化數位探索數據。
- 通過互動式儀錶板查看數據
- 使用 **SQL Lab** 編寫查詢以瀏覽數據

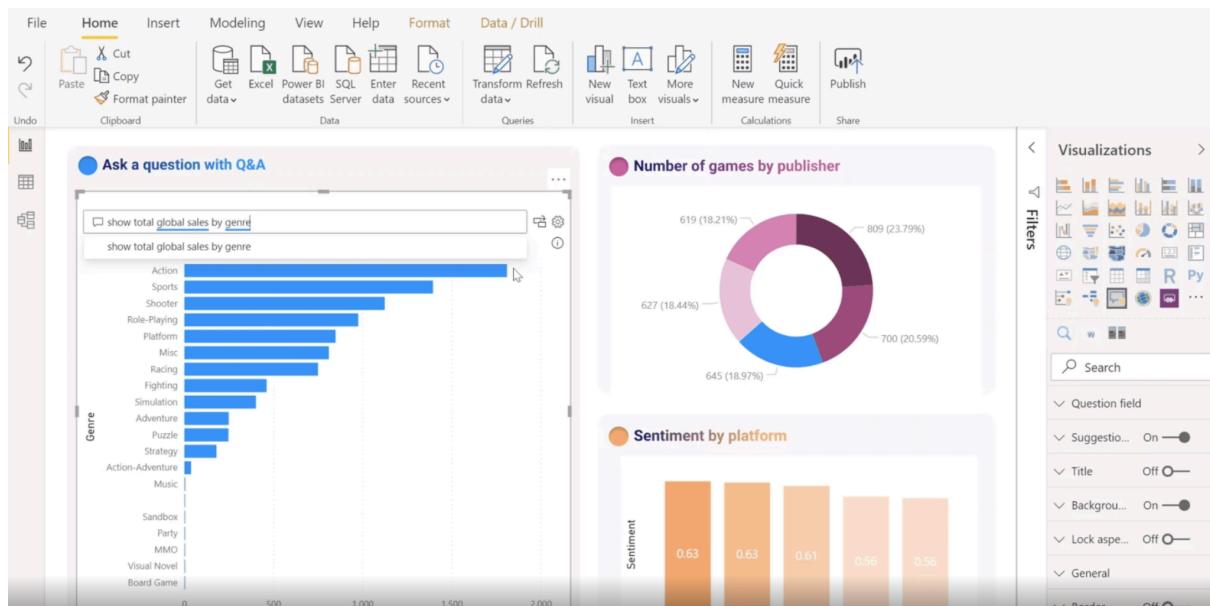


Power BI 由 Microsoft 創建。Power BI 是一個統一、可縮放的自助服務和企業商業智慧（BI）平臺。連接並視覺效果注入應用。

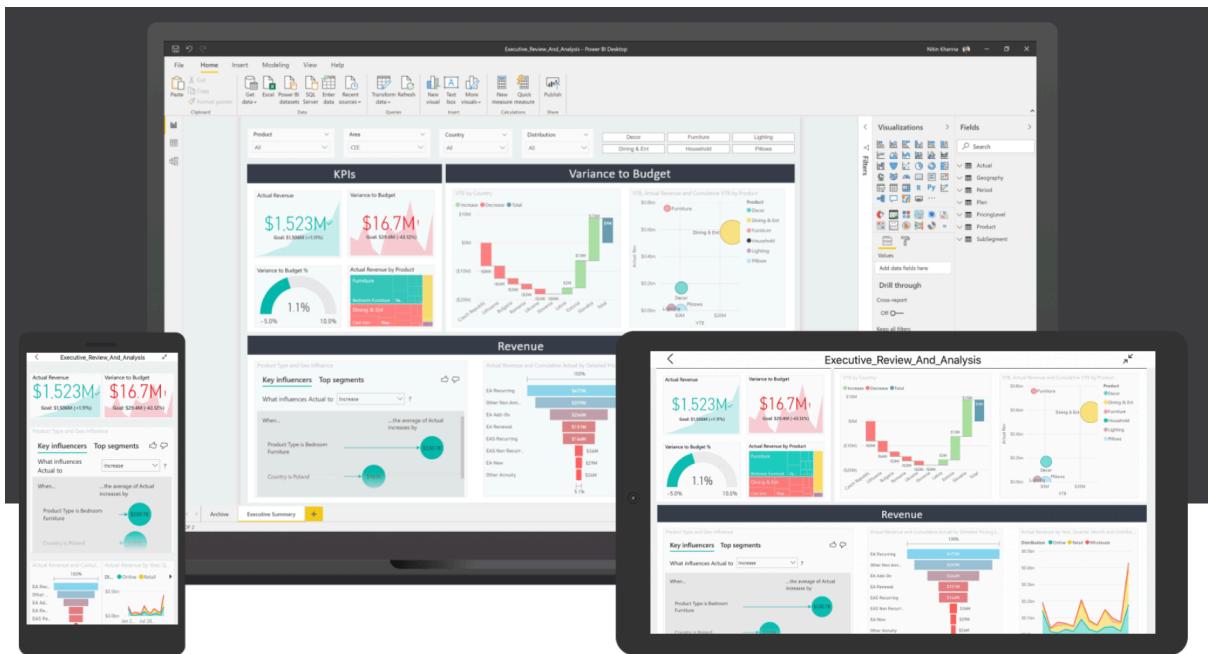
輕鬆連接、建模和可視化數據，創建具有 KPI 和品牌個人化的令人難忘的報告。甚至獲得 AI 驅動的業務問題的快速答案。



在相同的數據上輕鬆協作，協作處理報表，並在流行的 Microsoft Office 應用程式（如 Microsoft Teams 和 Excel）之間共用見解，以快速做出數據驅動的決策，從而推動戰略行動。



Power BI: MS Office 般的介面



Python 相容
matplotlib



plotly | Graphing Libraries

plotly | Dash

獨立運作



Apache **Superset**™



參考資料：

- <https://matplotlib.org/>
- <https://seaborn.pydata.org/>
- <https://plotly.com/graphing-libraries/>
- <https://dash.plotly.com/>
- <https://www.tableau.com/>
- <https://superset.apache.org/>
- <https://powerbi.microsoft.com/>