

Python初級數據分析員證書
(六) 數據分析及可視化專案

14. 數據分析專案 Storytelling for Data Analysis

1

14. 數據分析專案 Storytelling for Data Analysis

Recap

In previous chapters we had learnt:

- Statistics
- Algebra
- Python
- Pandas, NumPy, SciPy
- Data Analysis
- Data Visualization, Matplotlib, Seaborn, Plotly
- Dashboard Visualization, Business Intelligence

Python初級數據分析員證書
(六) 數據分析及可視化專案

2

13.數據分析專案 Storytelling for Data Analysis

Chapter Summary

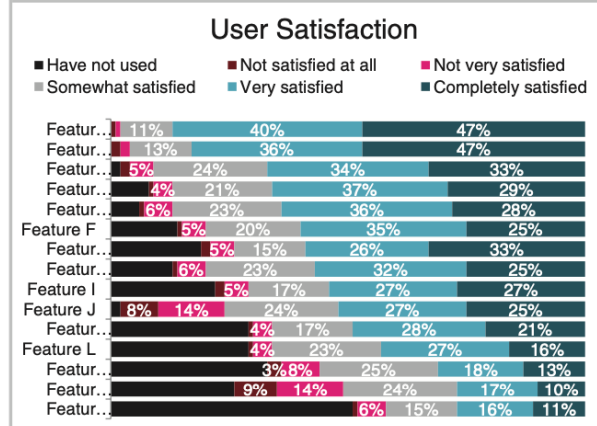
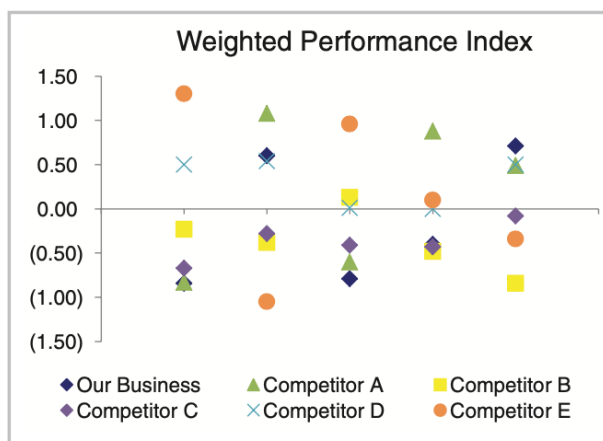
- 故事表達 – 內容
- 選擇有效的視覺效果
- 一切整潔
- 集中觀眾的注意力
- 明智地使用顏色
- 故事表達
- 個案研究: MacroMicro

Python初級數據分析員證書
(六) 數據分析 及可視化專案

3

14.數據分析專案 Storytelling for Data Analysis

Sample of ineffective charts



Python初級數據分析員證書
(六) 數據分析 及可視化專案

4

Storytelling - Context

在開始創建數據可視化或通信之前，應注意並花時間理解**上下文** 為了溝通的需要。

有兩種類型的分析：

- **探索** – 為同行業同行提供數據技術研究
- **說明** – 普通受眾使用的數據資訊易於理解在我們的課程中，我們主要關注解釋和交流。

Who, What, How

當涉及到解釋性分析時，有幾件事需要考慮並非常清楚：**who, what, and how**

Who

- 內部和外部利益相關者？
- 有人可能感興趣嗎？
- 對數據及相關領域的知識水準如何？

14. 數據分析專案 Storytelling for Data Analysis

Who, What, How

What

- 你需要你的聽眾知道什麼或做什麼？
- 他們在分析數據方面的需求是什麼？
- 目前的痛點是什麼？
- 他們對分析的期望是什麼？

Python初級數據分析員證書
(六) 數據分析 及可視化專案

7

14. 數據分析專案 Storytelling for Data Analysis

Who, What, How

How

- 你將如何與你的聽眾（演示文稿、文檔）進行交流？
- 應提供多詳細的資訊？
- 建議可能的後續步驟可能是讓對話繼續進行的好方法

Python初級數據分析員證書
(六) 數據分析 及可視化專案

8

Live presentation tips

- 寫下**要點**，用每張幻燈片來製作。
- **實踐** 你想對自己大聲說的話：這會點燃大腦的不同部分，說明你記住你的談話要點。它還迫使您闡明幻燈片之間的過渡，這些過渡有時會使演示者感到困惑。
- 給一個 **預備演示** 給朋友或同事。
- 錄製您的模擬演示和**回顧** 用於修改。



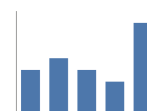
Choosing effective visual

91%

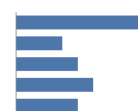
Simple text



Scatterplot



Vertical bar



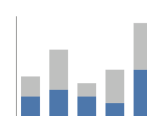
Horizontal bar

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

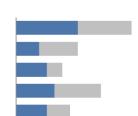
Table



Line



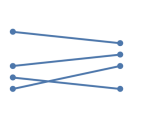
Stacked vertical bar



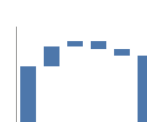
Stacked horizontal bar

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

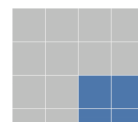
Heatmap



Slopegraph



Waterfall



Square area

Choosing effective visual - Simple text

當你只有一兩個數位要分享時, **簡單文本**可以是一種很好的溝通方式。您可以放大臨界數位以產生深遠影響。

20%

of children had a **traditional stay-at-home mom** in 2012, compared to 41% in 1970

選擇有效的視覺物件 - 表格

表單 非常適合這一點——與混合的觀眾交流，這些觀眾的成員都會尋找他們特定的興趣所在。

- 當觀眾閱讀表格時，您可能**失去了他們的耳朵和注意力**。
- 如果您必須向觀眾展示一**表單**，請提出一個問題，例如“您在表格中發現了什麼有趣的東西？”並給他們 20 秒。
- 當我面前有一**表單**時，我通常會有我的**伸出食指**：我正在跨行和向下列閱讀，或者我正在比較值。

Choosing effective visual – Heat Map

一種混合可以包含在表格中的細節的方法是利用視覺提示，方法是通過熱圖。

Table

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

Heatmap

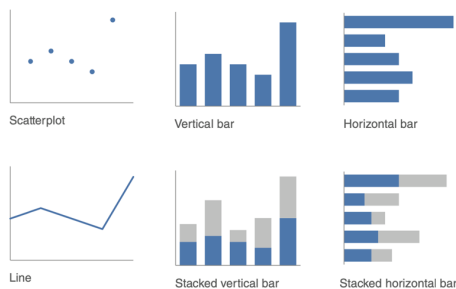
LOW-HIGH

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

Choosing effective visual – Graph

一個精心設計的圖表通常會獲得更多的資訊 比一張精心設計的表單還要快。

而表與我們的語言系統交互，圖形與我們的語言系統交互 視覺系統，這是更快地處理資訊。

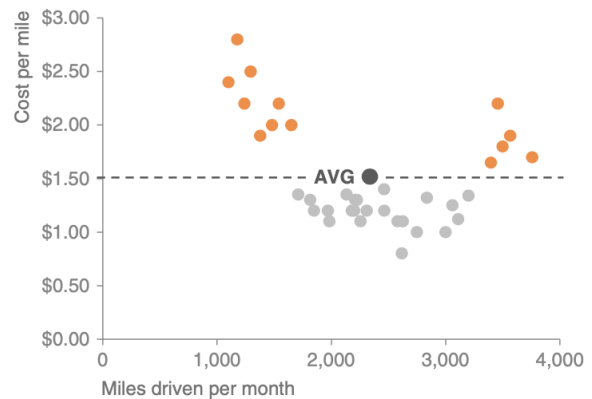


14. 數據分析專案 Storytelling for Data Analysis

選擇有效的視覺物件 - 散點圖

Scatterplots 可用於顯示兩件事之間的關係, 因為它們允許您在水平方向上同時對數據進行編碼 x -axis and 垂直 y -axis 以查看是否存在以及存在什麼關係。

Cost per mile by miles driven



Python初級數據分析員證書
(六) 數據分析 及可視化專案

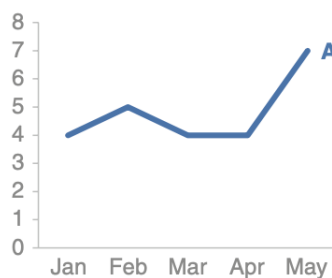
15

14. 數據分析專案 Storytelling for Data Analysis

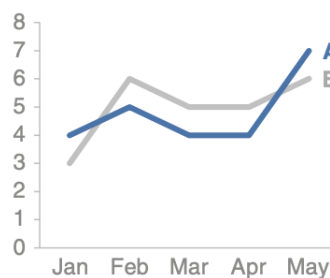
選擇有效的視覺效果 - 折線圖

折線圖最常用於繪製連續數據。由於這些點是通過線進行物理連接的，因此它意味著點之間的連接對於分類數據可能沒有意義

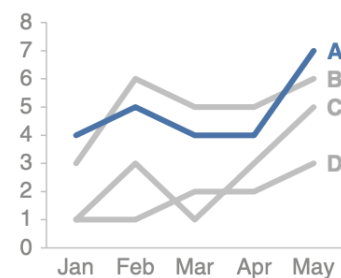
Single series



Two series



Multiple series

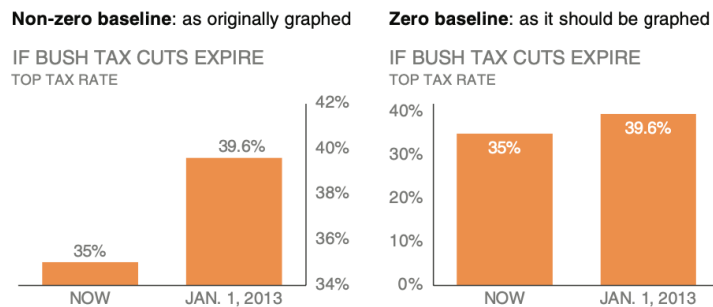


Python初級數據分析員證書
(六) 數據分析 及可視化專案

16

選擇有效的視覺物件 - 條形圖

Bar charts 很容易讓我們的眼睛閱讀。我們的眼晴比較了柱線的終點，因此很容易看到更**迅速** 哪個類別最大，哪個最小，以及類別之間的增量差異。

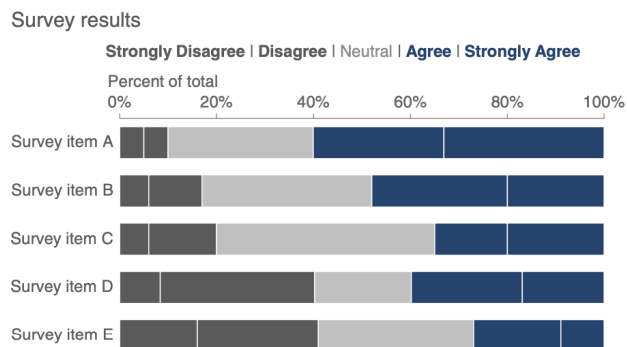


Python初級數據分析員證書
(六) 數據分析 及可視化專案

17

Stacked horizontal bar charts

Stacked horizontal bar charts 可用於顯示不同類別的總數，但也可用於顯示子元件部分。它們可以被結構化為顯示絕對值或總和100%。

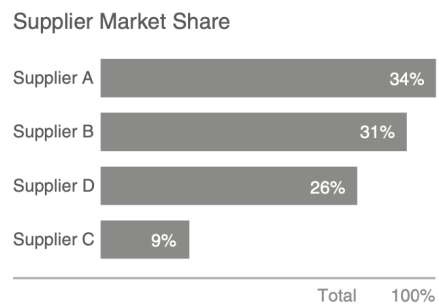
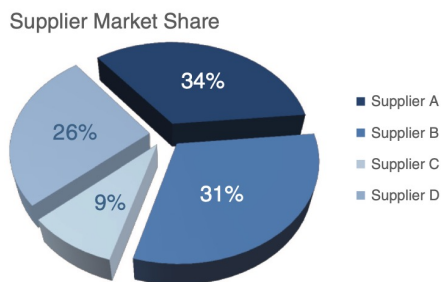


Python初級數據分析員證書
(六) 數據分析 及可視化專案

18

要避免的一些圖表

您還應該避免使用一些特定的圖形類型和元素：餅圖、圓環圖、3D 和輔助 y 軸。因為理解它們需要更多時間。

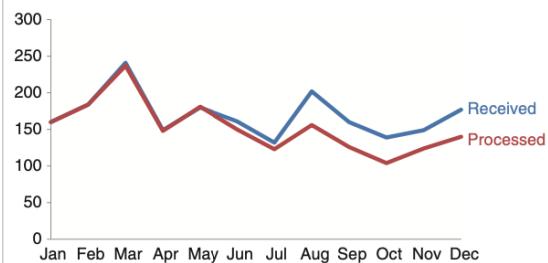
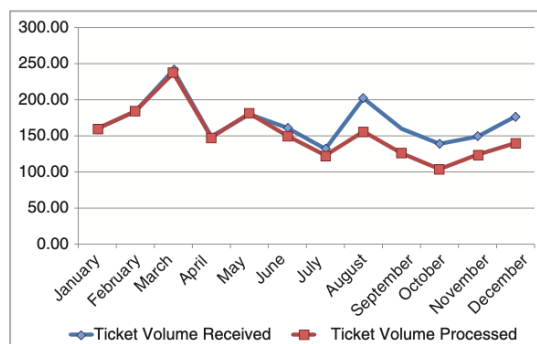


Python初級數據分析員證書
(六) 數據分析 及可視化專案

19

一切整潔

簡潔永遠是最好的。刪除不必要的詳細資訊和資訊，讓您的聽眾快速閱讀。沒有花哨的顏色和風格。

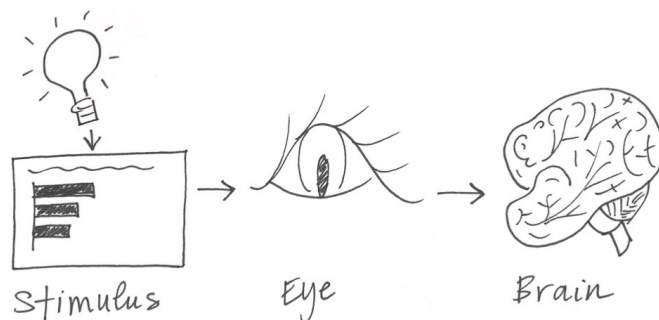


Python初級數據分析員證書
(六) 數據分析 及可視化專案

20

集中觀眾的注意力

簡化的圖片人們如何看待. 光從刺激物反射出來。這被我們的眼睛捕捉到了。我們不能用眼睛完全看清; 那裡會進行一些處理, 但大多數情況下是這樣發生在我們的大腦中。



Python初級數據分析員證書
(六) 數據分析 及可視化專案

21

集中觀眾的注意力

在大腦內部, 有三樣在我們設計視覺通信時必須理解的記憶類型: **iconic memory, short-term memory, and long-term memory.**

Iconic memory (標誌性的記憶體)

標誌性的記憶體速度超快。資訊會在你的標誌性記憶中停留幾分之一秒, 然後才會被轉發到你的短期記憶中。

Python初級數據分析員證書
(六) 數據分析 及可視化專案

22

Focus your audience's attention

Short-term memory (短期記憶)

人們可以在他們的短期記憶中保留大約四塊視覺資訊。如果我們創建一個包含十個不同數據系列的圖表，這些數據系列是十種不同的顏色，十種不同形狀的數據標記和旁邊的圖例，我們就會讓我們的觀眾非常努力地來回工作。

Focus your audience's attention

Long-term memory (長期記憶)

當某物離開短期記憶時，它要麼變成垃圾，很可能永遠丟失，要麼就是 **進入長期記憶**。

長期記憶是在一生中積累起來的，對於 **模式識別** 和一般 **認知處理**。

讓我們在下一張幻燈片中測試我們的大腦。在下一張幻燈片中，您將看到一大堆數位，系統要求您計算**3**的**總出現次數**，越快越好。並**計算你的時間**來完成這項任務。

14. 數據分析專案 Storytelling for Data Analysis

Focus your audience's attention

756395068473
658663037576
860372658602
846589107830

Python初級數據分析員證書
(六) 數據分析 及可視化專案

25

14. 數據分析專案 Storytelling for Data Analysis

Focus your audience's attention

正確答案是**六**. 沒有視覺提示可以說明您得出這個結論。

看看當我們對數位塊進行一次更改時會發生什麼。轉動**下一張幻燈片**並重複計算**3**.

Python初級數據分析員證書
(六) 數據分析 及可視化專案

26

14. 數據分析專案 Storytelling for Data Analysis

Focus your audience's attention

756**3**9506847**3**
 65866**3**0**3**7576
 860**3**72658602
 8465891078**3**0

Python初級數據分析員證書
 (六) 數據分析 及可視化專案

27

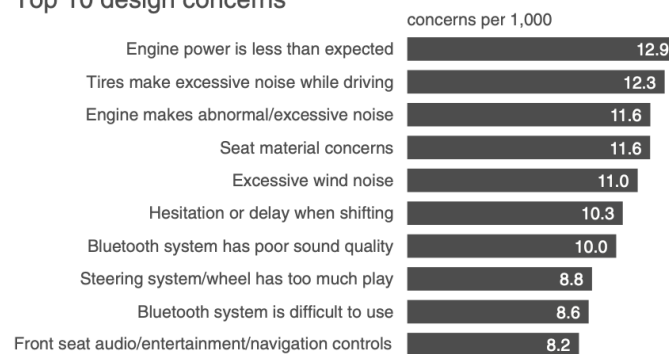
14. 數據分析專案 Storytelling for Data Analysis

集中觀眾的注意力

請注意，使用最後一張幻燈片進行相同的練習要容易得多、快速得多。

他之所以如此明顯，是因為在第二次反覆運算中，你的**標誌性的記憶**正被**槓桿**。讓我們看一個示例圖

Top 10 design concerns



Python初級數據分析員證書
 (六) 數據分析 及可視化專案

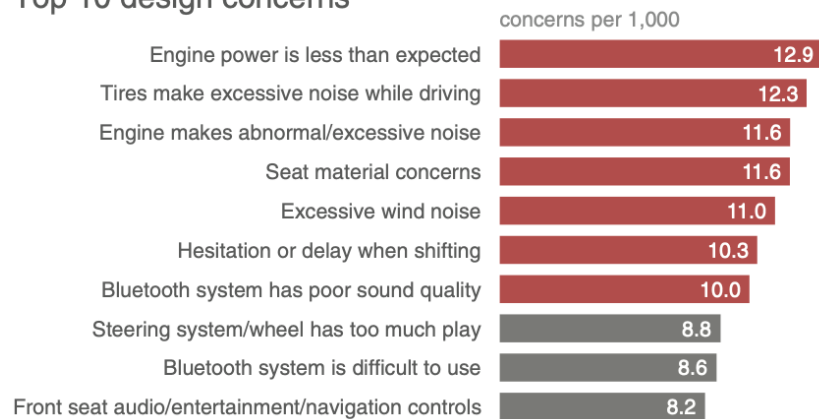
28

Focus your audience's attention

7 of the top 10 design concerns have 10 or more concerns per 1,000.

Discussion: is this an acceptable default rate?

Top 10 design concerns

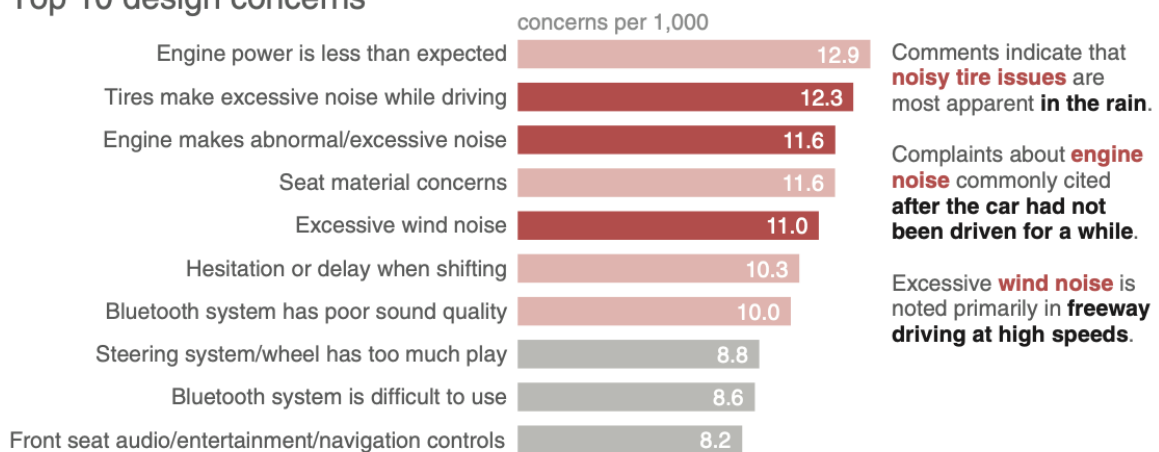


Python初級數據分析員證書
(六) 數據分析及可視化專案

29

Focus your audience's attention

Top 10 design concerns



Python初級數據分析員證書
(六) 數據分析及可視化專案

30

Use colour wisely

Rainbow colour may **not** help the audience understand the chart quickly.

Country Level Sales Rank Top 5 Drugs

Rainbow distribution in color indicates sales rank in given country from #1 (red) to #10 or higher (dark purple)

Country	A	B	C	D	E
AUS	1	2	3	6	7
BRA	1	3	4	5	6
CAN	2	3	6	12	8
CHI	1	2	8	4	7
FRA	3	2	4	8	10
GER	3	1	6	5	4
IND	4	1	8	10	5
ITA	2	4	10	9	8
MEX	1	5	4	6	3
RUS	4	3	7	9	12
SPA	2	3	4	5	11
TUR	7	2	3	4	8
UK	1	2	3	6	7
US	1	2	4	3	5

Top 5 drugs: country-level sales rank

RANK	1	2	3	4	5+
COUNTRY DRUG	A	B	C	D	E
Australia	1	2	3	6	7
Brazil	1	3	4	5	6
Canada	2	3	6	12	8
China	1	2	8	4	7
France	3	2	4	8	10
Germany	3	1	6	5	4
India	4	1	8	10	5
Italy	2	4	10	9	8
Mexico	1	5	4	6	3
Russia	4	3	7	9	12
Spain	2	3	4	5	11
Turkey	7	2	3	4	8
United Kingdom	1	2	3	6	7
United States	1	2	4	3	5

Python初級數據分析員證書
(六) 數據分析及可視化專案

31

Storytelling

- **查找你關心的主題.** 正是這種真正的關懷，而不是你的語言遊戲，這將是你風格中最引人注目和最誘人的元素。
- **不要漫無邊際**
- **保持簡單.** 偉大的大師們在他們的主題最深刻的時候寫出的句子幾乎像孩子一樣。“是還是不是？”莎士比亞的《哈姆雷特》問道。最長的單詞是三個字母。

Python初級數據分析員證書
(六) 數據分析及可視化專案

32

Storytelling

- **擁有要削減的膽量.** 如果一個句子，無論多麼出色，都沒有以某種新的和有用的方式闡明你的主題，那就把它劃掉。
- **聽起來像你自己.** 我自己發現我最信任自己的寫作，其他人似乎也最信任它，當我聽起來最像一個來自印第安那波利斯的人時，這就是我。

Storytelling

- **說出你想說的話.** 如果我打破了所有標點符號的規則，讓單詞意味著我想要它們的含義，然後把它們雜亂無章地串在一起，我就不會被理解。
- **關注讀者.** 我們的聽眾要求我們成為富有同情心和耐心的教師，永遠願意簡化和澄清。與觀眾互動。

最後的思考

數據 **可視化** 和 **溝通** 一般來說，數據位於科學與藝術的交匯點。這當然有一些科學依據：可以遵循的最佳實踐和準則。但也有藝術的成分。這就是這個地區如此有趣的原因之一！



結論

最後的提示：

- Tip #1: 學好你的工具
- Tip #2: 反覆運算並尋求反饋
- Tip #3: 花時間用數據講故事
- Tip #4: 通過良好的範例尋求靈感表達
- Tip #5: 開心地表達，找到你的風格

