Enhanced Programme on Promoting Mathematical Modelling for Teachers and Students in Secondary Schools
Student Workshop 2025/26 (Senior)
推廣中學教師及學生數學建模計劃
學生工作坊 2025/26（高中）

Part II: Advanced Methods for Mathematical Modelling
第二部份：數學建模進階方法

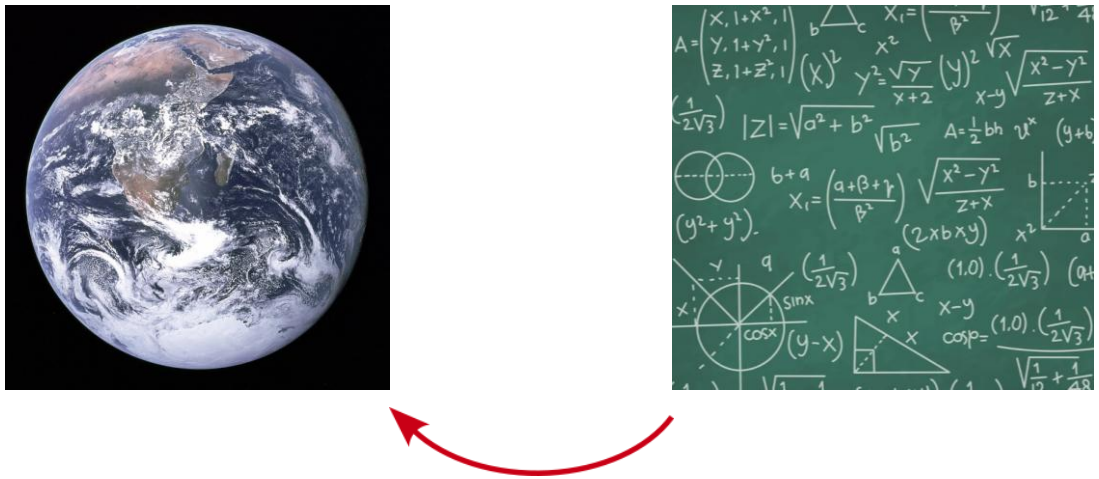
Prof. Benny Yiu-Chung Hon 韓耀宗教授

Department of Mathematics, The Chinese University of Hong Kong

香港中文大學數學系

Mathematical Modelling is … 數學建模是 ...

Understanding a real-world problem using mathematics
用數學了解現實生活問題



Mathematical Modelling Process
5 Steps of Mathematical Modelling
數學建模過程
數學建模 5 部曲

Real-world problem 現實世界問題

1. Analyzing the real-world Problem 分析現實世界問題
2. Formulating the Mathematical Model 建立數學模型

Mathematical model 數學模型

3. Solving the Mathematical Model 求解數學模型

Mathematical solution 數學模型的解

4. Interpreting the Mathematical Solution 解釋數學模型的解

Real-world Solution 現實世界的解

5. Evaluating the Mathematical Model 評估數學模型

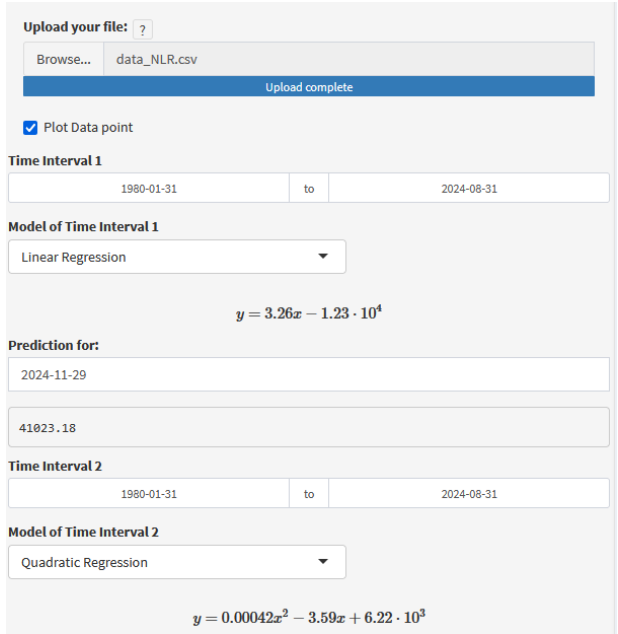
What Mathematical Concepts may be involved in Mathematical Modelling? 數學建模會用到甚麼數學概念？

- Short answer: Everything is possible!
  簡短答案：一切皆有可能！
  - Algebra 代數
  - Geometry 幾何
  - Probability and statistics 概率與統計
  - Calculus and optimization 微積分與最佳化
  - …

- However 然而:
  - The key is to have a mathematical mindset 關鍵在於擁有數學思維
  - Suitably transform the real-world problem into a math problem that we know how to solve 適當地將現實世界問題轉化為我們懂得求解的數學問題
  - Complicated model ≠ better model!
    複雜的模型 ≠ 更好的模型！

What Mathematical Concepts may be involved in Mathematical Modelling? 數學建模會用到甚麼數學概念？

- Algebra 代數:
  - Solving equations 解方程
  - Variations
    - Direct Variation 正變　　$y \propto x$, e.g. $y = 3x$
    - Inverse Variation 反變　　$y \propto \frac{1}{x}$, e.g. $y = \frac{2}{x}$
    - Joint Variation 聯變,　　e.g. $y = \frac{x}{z^2}$
    - Partial Variation 部分變, e.g. $y = k_1 x + k_2 z^3$



What Mathematical Concepts may be involved in Mathematical Modelling? 數學建模會用到甚麼數學概念？

- Modelling with functions 使用函數建模
  - Linear 線性函數

$$y = ax + b$$

  - Polynomial 多項式函數

$$y = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$$

  - Exponential 指數函數

$$y = ab^x$$

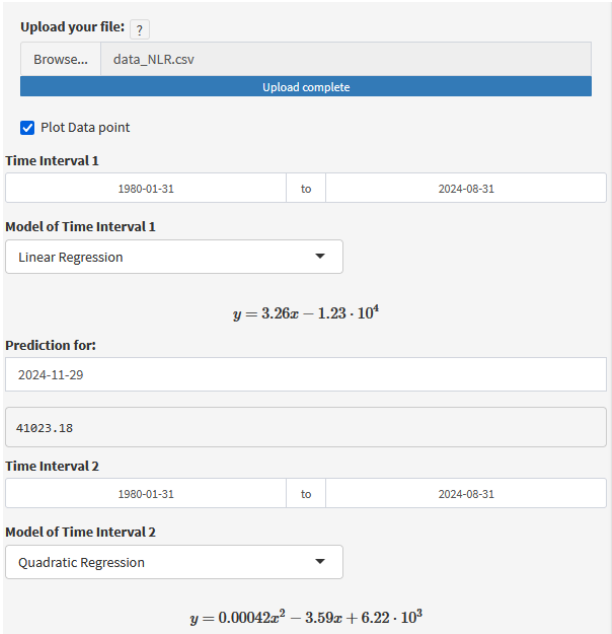
  - Logarithmic 對數函數

$$y = a + b \log x$$

  - Power model 冪模型

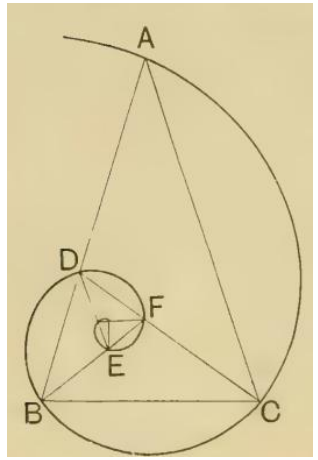
$$y = ax^b$$

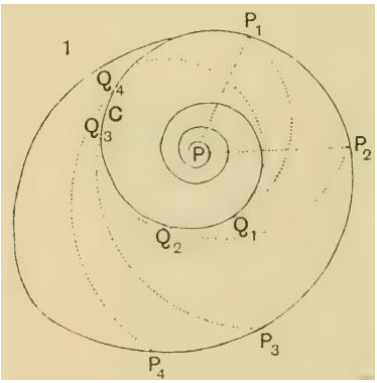
  - …



What Mathematical Concepts may be involved in Mathematical Modelling? 數學建模會用到甚麼數學概念？

- Geometry幾何:
  - Geometric measurements 幾何測量
    - Length 長度
    - Area 面積
    - Volume 體積
  - Deductive geometry 演繹幾何
  - Coordinate geometry 座標幾何
    - Equation of straight lines 直線方程
    - Distance formula 距離公式
  - Trigonometry 三角函數
    - $\sin x$, $\cos x$, …





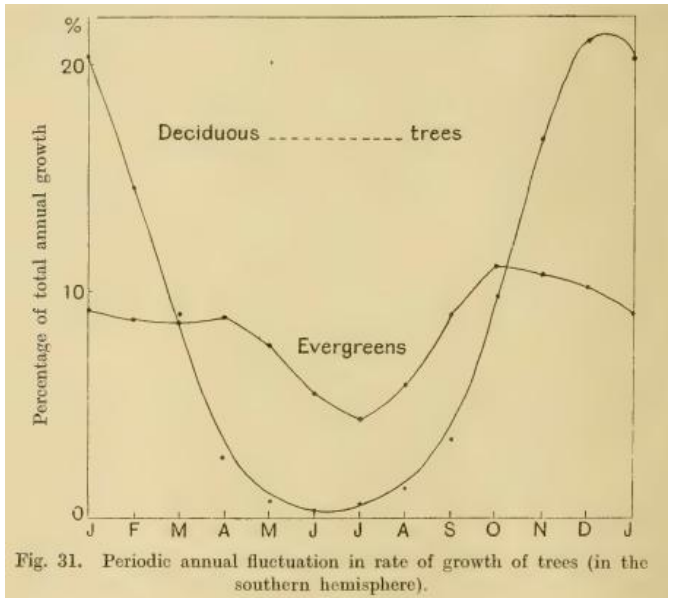
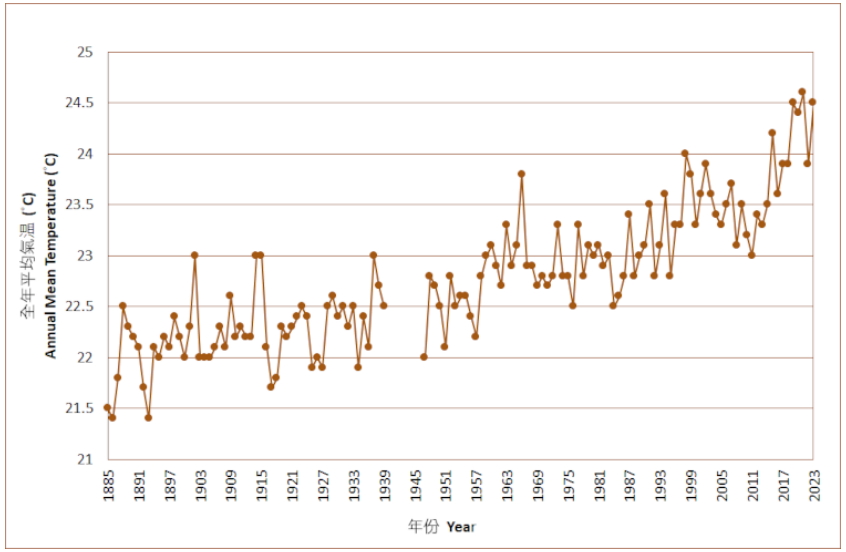


Fig. 31. Periodic annual fluctuation in rate of growth of trees (in the southern hemisphere).

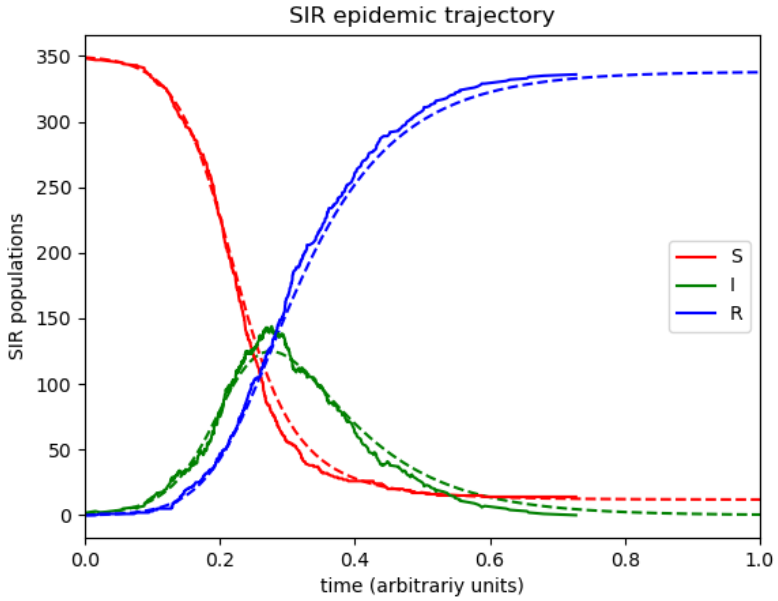
D'Arcy Thompson, *On Growth and Form* (1917)

What Mathematical Concepts may be involved in Mathematical Modelling? 數學建模會用到甚麼數學概念？

- Probability and Statistics 機率與統計:
  - Randomness 隨機性
  - Statistical analysis 統計分析



- More advanced techniques 更進階的技巧:
  - Calculus 微積分
  - Optimization 最優化
  - …



Susceptible
Infectious
Recovered/Removed

More Examples of Mathematical Modelling
更多數學建模的例子

Example: Estimating the Percentage of Electric Vehicles among All Cars in Hong Kong in 2030
估算在2030年香港電動車佔所有車輛的百分比

- Past MMCSS competition problem (2023/24 Junior)
  以往MMCSS比賽題目（2023/24年初中組）

Estimate the Percentage of Electric Vehicles among All Cars in Hong Kong in 2030

The Government of the Hong Kong Special Administrative Region has set an ambitious carbon intensity target of 65% to 70% by 2030 using 2005 as the base, which is equivalent to 26% to 36% absolute reduction and a reduction to 3.3-3.8 tonnes on a per capita basis. It is observed that Hong Kong's carbon emissions have shown a decreasing trend since 2014. In 2021, the Government announced Hong Kong's Climate Action Plan 2050, which outlined four major decarbonisation strategies, namely "net-zero electricity generation", "energy saving and green buildings", "green transport" and "waste reduction", that would lead Hong Kong towards the goal of carbon neutrality before 2050. It also sets out a more vigorous interim decarbonisation targets to reduce Hong Kong's carbon emissions by 50% before 2035 as compared to the 2005 level.

To strive for attaining the target of carbon neutrality in Hong Kong before 2050, promoting zero carbon emissions transport is one of the indispensable strategies.

Estimate the percentage of electric vehicles among all cars in Hong Kong in 2030.

State the data you have collected clearly. Your data must be accurate, with sources cited, and your argument must be logical and sound. State clearly the assumption(s) you need in your modelling process.

估算在 2030 年香港電動車佔所有車輛的百分比

香港特別行政區政府訂立了進取的碳強度目標，在 2030 年把碳強度由 2005 年的水平降低 65%至 70%，相當於 26%至 36%絕對減排量，以及人均排放量減至 3.3 至 3.8 公噸。香港的碳排放總量自 2014 年起呈現下降趨勢。政府在 2021 年發布了《香港氣候行動藍圖 2050》，以「淨零發電」、「節能綠建」、「綠色運輸」和「全民減廢」為四大減碳策略，帶領香港於 2050 年前邁向碳中和，並加強減碳中期目標，力爭在 2035 年前把香港的碳排放量從 2005 年的水平減半。

香港致力爭取 2050 年前實現碳中和的目標，推動零碳交通運輸是其中一項不可或缺的策略。

請你估計在 2030 年香港電動車佔所有車輛的百分比。

請列出你所收集的資料。資料要準確和寫出來源，論證要合乎邏輯。建模過程中所設立的所有假設均需清晰列出。

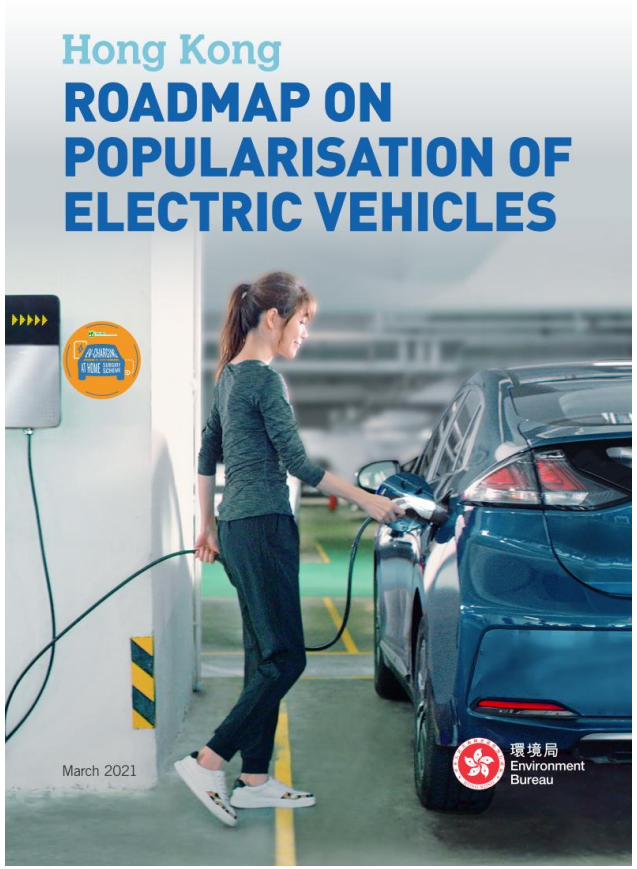
Example: Estimating the Percentage of Electric Vehicles among All Cars in Hong Kong in 2030
估算在2030年香港電動車佔所有車輛的百分比

- **1. Analyzing the real-world problem 分析現實世界問題:**

  - Hong Kong Climate Action Plan 2050
    香港氣候行動藍圖2050
    - Net-zero electricity generation 淨零發電
    - Energy saving and green buildings 節能綠建
    - Green transport 綠色運輸
    - Waste reduction 全民減廢

  - Need to promote zero carbon emissions transport
    需要推廣零碳排放交通

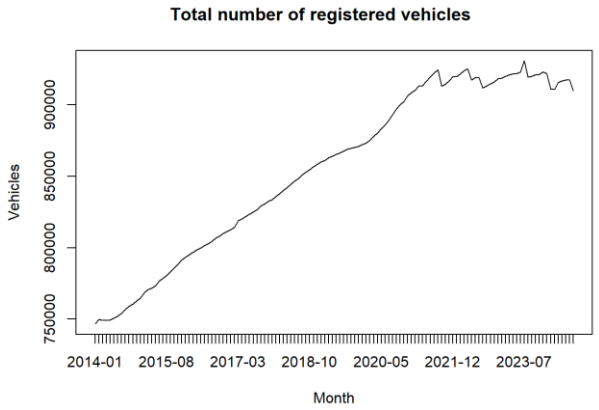
  - How can we analyze the percentage of electric vehicles among all cars in Hong Kong?
    如何分析香港所有汽車中電動車的比例？

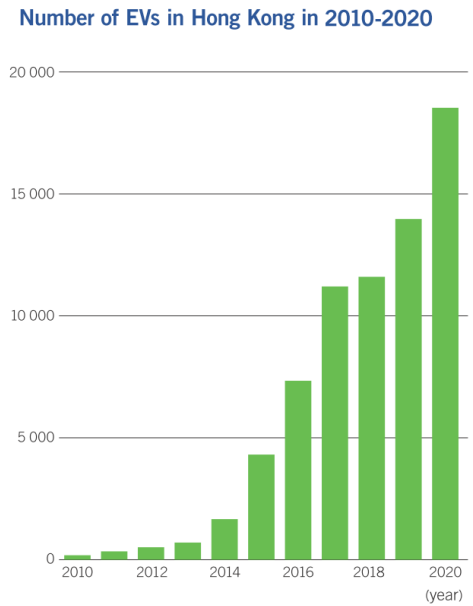


Source: Environmental Protection Department

Example: Estimating the Percentage of Electric Vehicles among All Cars in Hong Kong in 2030
估算在2030年香港電動車佔所有車輛的百分比

- **2. Formulating the mathematical model 建立數學模型:**
  - Factors 影響因素：
    - Number of electric vehicles 電動車數量
    - Number of total vehicles 車輛總數
  - Assumption 假設:
    - Current policies (first registration tax, electric vehicle subsidy scheme) 現行政策（首次登記稅、電動車補貼計畫）
    - Technological development 技術發展
  - Relevant data 相關數據：
    - Transport Department 運輸署
      https://www.td.gov.hk/en/transport_in_hong_kong/transport_figures/monthly_traffic_and_transport_digest/index.html
    - Environmental Protection Department 環保署
      https://www.epd.gov.hk/epd/english/environmentinhk/air/promotion_ev/promotion_ev.html
  - Develop a mathematical model 建立數學模型
    - Linear/nonlinear regression model in year 基於年份的線性/非線性迴歸模型
    - Regression model on various factors 基於各種因素的迴歸模型
    - Probabilistic model of switching to EV 電動車轉換機率模型



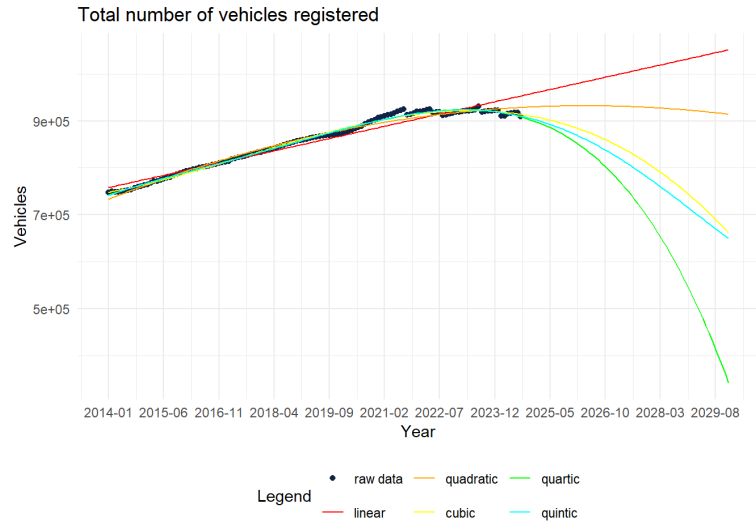


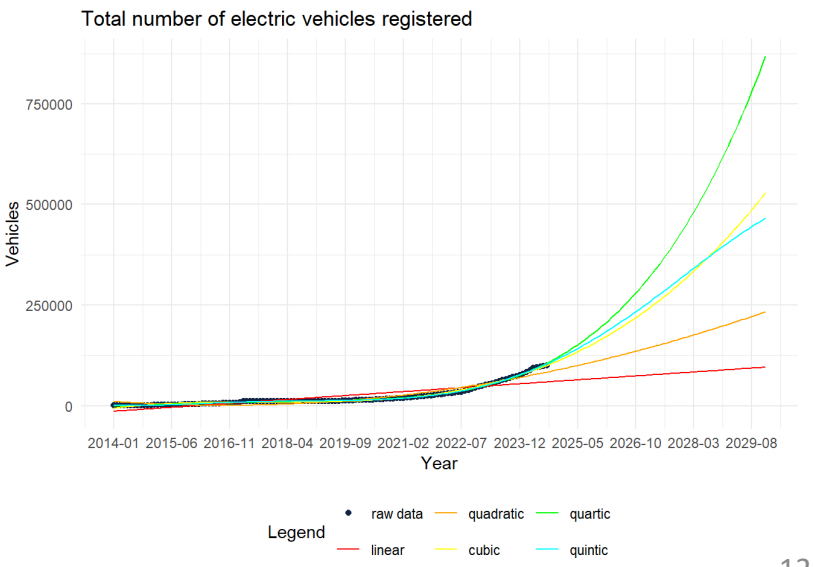
Data from Environmental Protection Department
https://www.epd.gov.hk/epd/english/top.html

Example: Estimating the Percentage of Electric Vehicles among All Cars in Hong Kong in 2030
估算在2030年香港電動車佔所有車輛的百分比

- **3. Solving the mathematical model:**
  求解數學模型：
  - Find the best-fit parameters using the given data and IT tools
    利用給定資料及 IT 工具找出最佳擬合參數

- **4. Interpreting the mathematical solution:**
  解釋數學模型的解：
  - Is the result increasing/decreasing? 增加還是減少？
    - Imply the trend of EV percentage 體現了電動車佔比的趨勢
  - Any sharp changes in the value? 數值是否存在劇烈變化？
    - Imply some rapid change in certain years
      顯示某些年份的快速變化
    - Align with technology/policy change?
      是否與技術/政策變化相符？



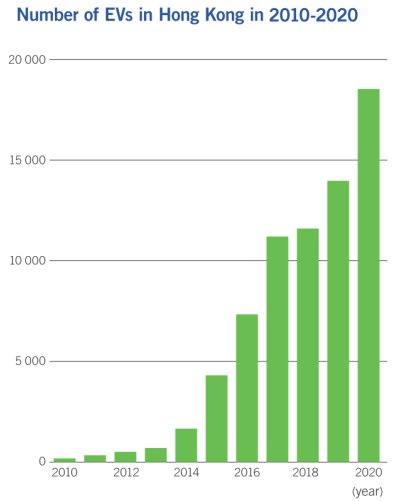


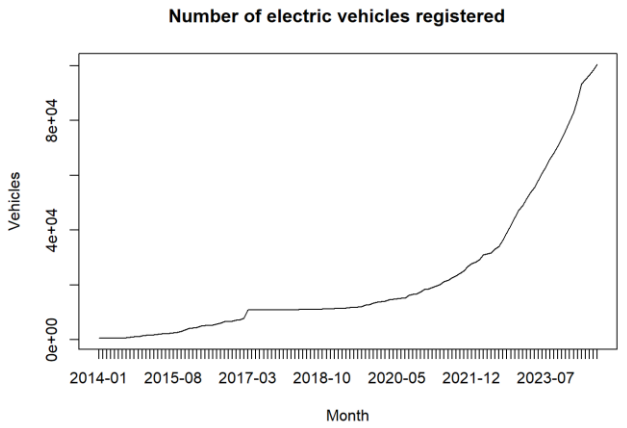
Example: Estimating the Percentage of Electric Vehicles among All Cars in Hong Kong in 2030
估算在2030年香港電動車佔所有車輛的百分比

- **5. Evaluating the mathematical model:**
  評估數學模型：

  - Test the model with other data
    使用其他數據測試模型
    - Data for some other years 使用其他年份的數據
    - Comparing with data for other economies with similar car targets
    - 與具有類似目標的其他經濟體的數據進行比較

  - Refining the model 改進模型:
    - Use different models for different factors (total number of cars, number of electric vehicles, electric vehicle price, etc.)
      針對不同因素（汽車總數、電動車數量、電動車價格等）使用不同的模型
    - Consider different year periods based on policy change 考慮基於政策變動的不同年份週期





Targets of zero emission private car sales of various economies

Year	Economy
2025	Norway[1]
2030	Denmark[2], Iceland[3], Ireland[4], Netherlands[5], Singapore[6], United Kingdom[7]
2035	United States (California)[8]
2040	Canada[9], France[10], Taiwan[11], Spain[12]

Data from Environmental Protection Department
https://www.epd.gov.hk/epd/english/top.html

Example: Estimating the Carbon Emissions of Hong Kong in 2030
估算香港在2030年的碳排放量

- Past MMCSS competition problem (2023/24 Senior)
  以往MMCSS比賽題目（2023/24年高中組）

Estimate the Carbon Emissions of Hong Kong in 2030

The Paris Agreement, an ambitious multilateral treaty agreed in December 2015, succeeds the Kyoto Protocol that expired in 2020. China formally signed it on Earth Day, 22 April 2016, and ratified it on 3 September 2016. The Paris Agreement came into force on 4 November 2016. Hong Kong plays a part to fulfill the obligations that China has under the Paris Agreement. As such, Hong Kong will need to review the climate change efforts every 5 years and align them with the requirements under the Paris Agreement.

Estimate the carbon emissions of Hong Kong in 2030.

State the data you have collected clearly. Your data must be accurate, with sources cited, and your argument must be logical and sound. State clearly the assumption(s) you need in your modelling process.

估算香港在 2030 年的碳排放量

《巴黎協定》在 2015 年 12 月通過，是一份目標進取的多邊協議，承接將在 2020 年屆滿的《京都議定書》。中國在 2016 年 4 月 22 日地球日正式簽署《巴黎協定》，並於 2016 年 9 月 3 日予以批准。《巴黎協定》已於 2016 年 11 月 4 日生效。中國須履行《巴黎協定》下的責任，而香港也擔當着一定的角色。因此，香港須每五年檢討應對氣候變化工作，以配合《巴黎協定》的要求。

請你估計香港在 2030 年的碳排放量。

請列出你所收集的資料。資料要準確和寫出來源，論證要合乎邏輯。建模過程中所設立的所有假設均需清晰列出。

Example: Estimating the Carbon Emissions of Hong Kong in 2030
估算香港在2030年的碳排放量

- **1. Analyzing the real-world problem:**
  分析現實世界問題

  - Understanding carbon emissions is crucial for creating effective environmental policies and meeting sustainability goals
    理解碳排放對於制定有效的環境政策和實現可持續發展目標至關重要

  - The Paris Agreement in December 2015
    2015年12月的巴黎協定

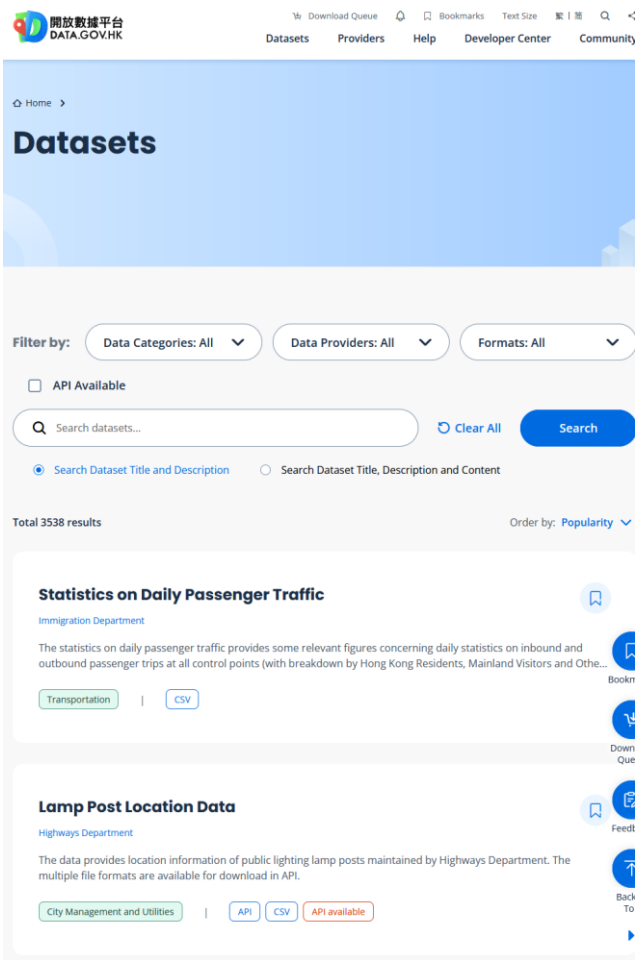
  - How can we estimate the carbon emissions of Hong Kong for reviewing climate change efforts?
    我們如何估計香港的碳排放，以檢討應對氣候變化工作？

Example: Estimating the Carbon Emissions of Hong Kong in 2030
估算香港在2030年的碳排放量

- **2. Formulating the mathematical model 建立數學模型:**
  - Assume that the carbon emission y is a function of the year x
    假設碳排放量 y 是年份 x 的函數
  - Assume the carbon emission depends certain factors, e.g.
    假設碳排放量取決於某些因素，例如：
    - Population change 人口變化
    - New technologies 新技術
    - International agreement (Paris Agreement) 國際協議（巴黎協定）

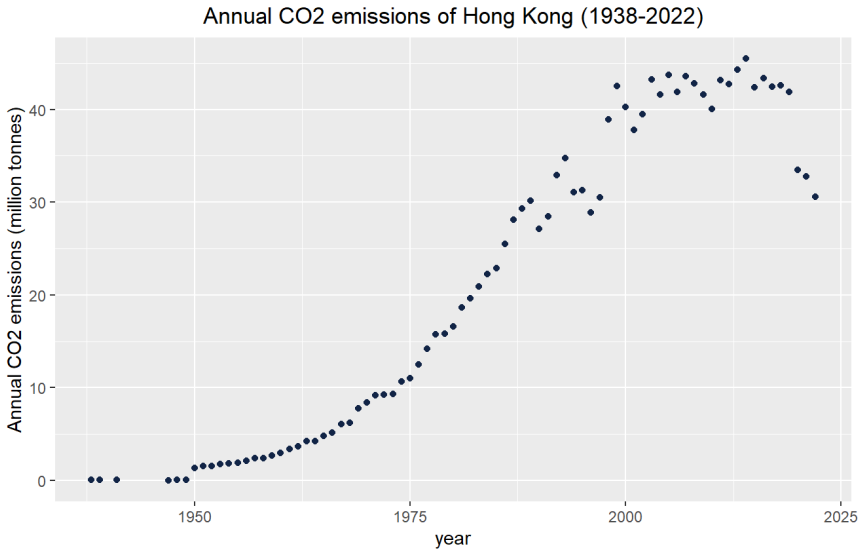
  - Relevant data相關數據：
    - DATA.GOV.HK香港政府開放數據平台 https://data.gov.hk/
    - Our World in Data https://ourworldindata.org/

  - Develop a mathematical model 建立數學模型
    - Linear/nonlinear regression model in year關於年份的線性/非線性迴歸模型
    - Regression model on various sources of carbon emission關於各種碳排放源的迴歸模型
    - Probabilistic model (e.g. on weather)機率模型（例如，關於天氣的機率模型）
    - …

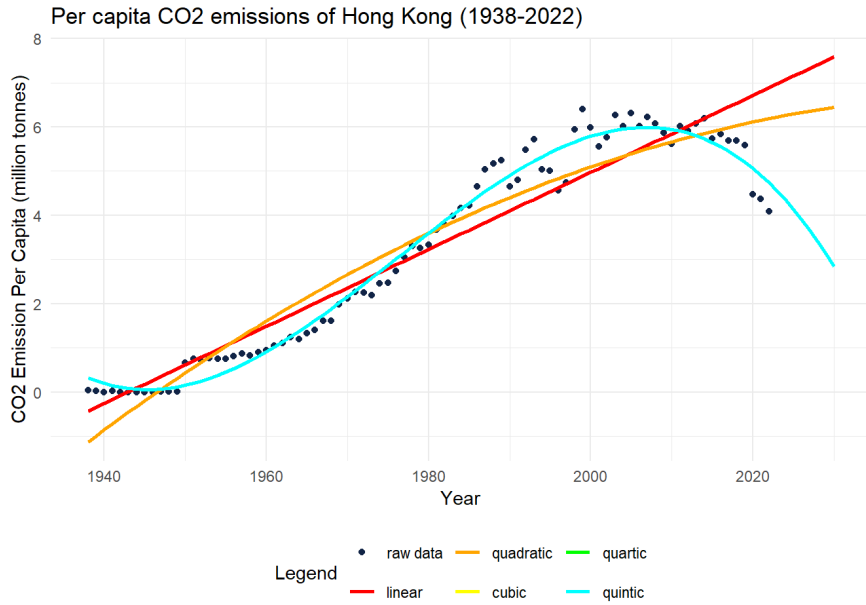


Example: Estimating the Carbon Emissions of Hong Kong in 2030
估算香港在2030年的碳排放量

- **3. Solving the mathematical model:**
  求解數學模型：
  - Find the best-fit parameters using the given data and IT tools
    利用給定資料及 IT 工具找出最佳擬合參數

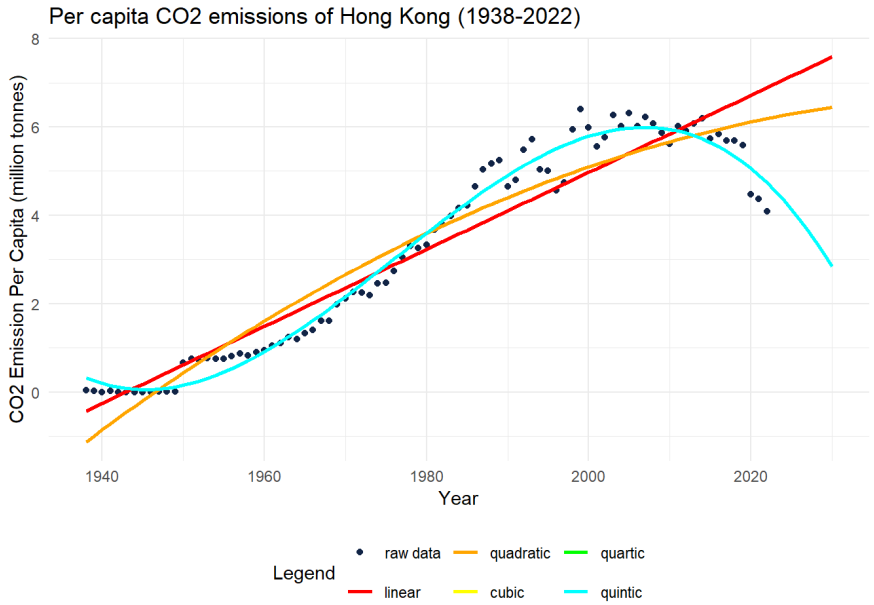


- **4. Interpreting the mathematical solution:**
  解釋數學模型的解：
  - Is the result increasing/decreasing?增加還是減少？
    - Imply the trend of carbon emission體現了碳排放的趨勢
  - Any sharp changes in the value? 數值是否存在劇烈變化？
    - Imply some rapid change in certain years
      顯示某些年份的快速變化
    - Align with technology/policy change?
      是否與技術/政策變化相符？



Example: Estimating the Carbon Emissions of Hong Kong in 2030
估算香港在2030年的碳排放量

- **5. Evaluating the mathematical model:**
  評估數學模型：



  - Test the model with other data 使用其他數據測試模型
    - Data for different years, e.g.
      使用不同年份的數據，例如：
      - Use 1990-2020 for model construction
        使用 1990-2020 年的數據建立模型
      - Test it using data in 2021-2024
        使用 2021-2024 年的數據測試模型
    - Comparing with data for other cities 與其他城市數據比較

  - Refining the model 改進模型：
    - Use other models 使用其他模型
    - Analyze the carbon emission by different sectors (e.g. household, transportation, industry, …) 分析不同類別（如家居、交通、工業等）的碳排放量
    - Consider different year periods using different models (e.g. before/after the Paris Agreement) 使用不同的模型考慮不同的年份時期（例如《巴黎協定》簽署前後）

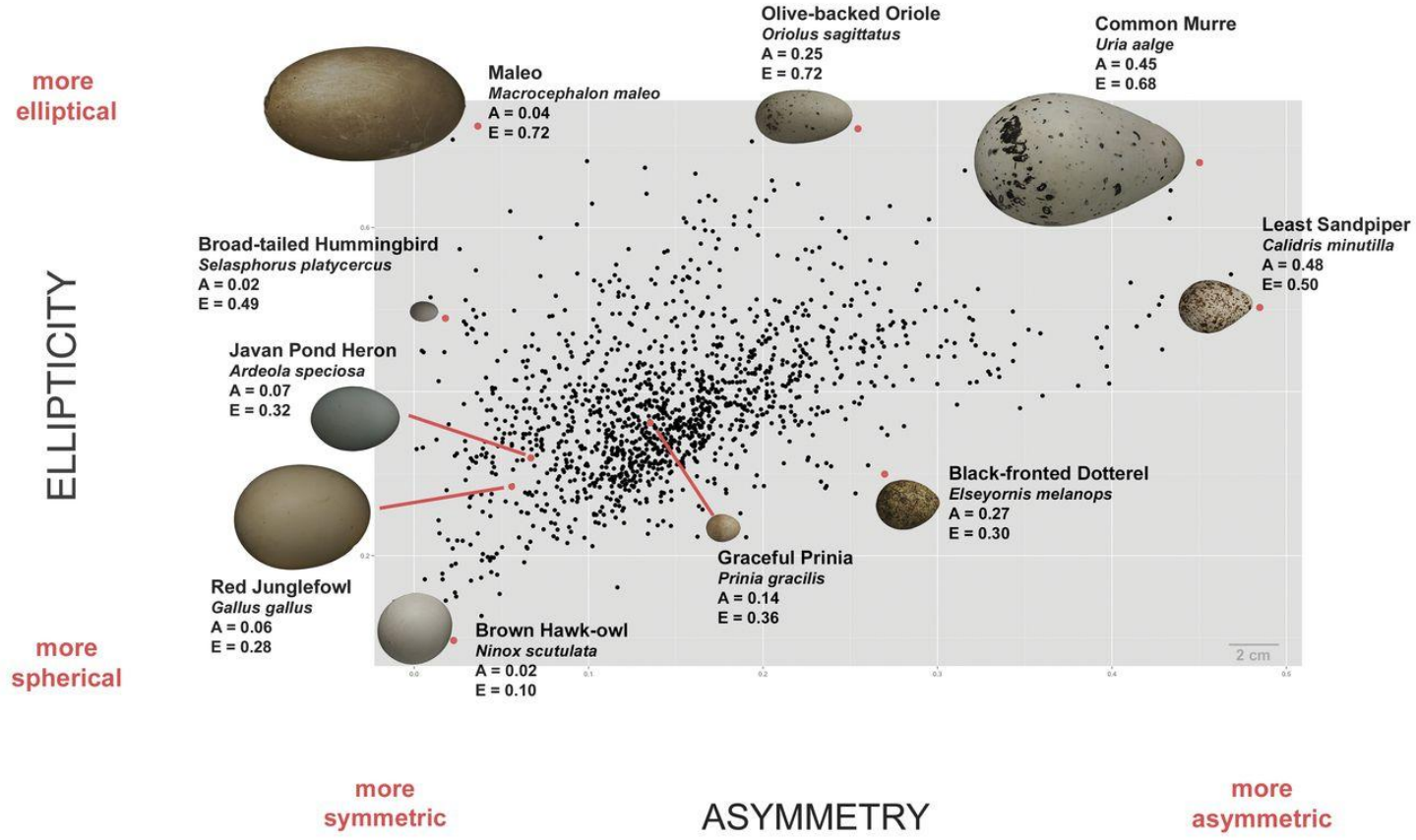
Example: Modelling the Shape of Eggs 蛋形狀的數學建模

- **Real-world problem 現實生活問題:**
  - There is a large variation in egg sizes and shapes in nature
    自然界中蛋的大小和形狀差異很大
  - Are the size and shape of eggs related to the features of the birds (body size, flying ability etc.)?
    蛋的大小和形狀是否與鳥類的特徵（體型、飛行能力等）有關？

Source: Stoddard et al., “Avian egg shape: Form, function, and evolution”, Science (2017)

Example: Modelling the Shape of Eggs 蛋形狀的數學建模

- **Mathematical problem 數學問題:**
  - How to represent the egg shapes? 如何表示蛋的形狀？
    - Ellipticity 橢圓度
    - Asymmetry 不對稱性
    - Area 面積
    - …
  - How to build a mathematical model to relate egg size /shape with some given measurements of the birds? 如何建立一個數學模型來將蛋的大小/形狀與鳥類的某些給定測量值聯繫起來？



Source: Stoddard et al., “Avian egg shape: Form, function, and evolution”, Science (2017)

Example: Modelling the Shape of Trees and their Roots
樹木和樹根的數學建模

- **Real-world problem 現實生活問題:**

  - We see different types of trees in everyday life
    我們在日常生活中看到不同類型的樹

  - Is there any relationship between the shape of the trees above ground and the shape of their roots underground?
    樹木地面上的形狀和地底根部的形狀有關係嗎？

Source: Carmona et al., "Fine-root traits in the global spectrum of plant form and function", Nature (2021)

Example: Modelling the Shape of Trees and their Roots
樹木和樹根的數學建模

- **Mathematical problem 數學問題:**

  - How to represent the above-ground and underground shapes mathematically?
    如何用數學的方式表示樹木地面上的形狀和地底根部的形狀？

    - Plant height 植物高度, leaf area樹葉面積, …
    - Root diameter 樹根直徑, root length 樹根長度, …

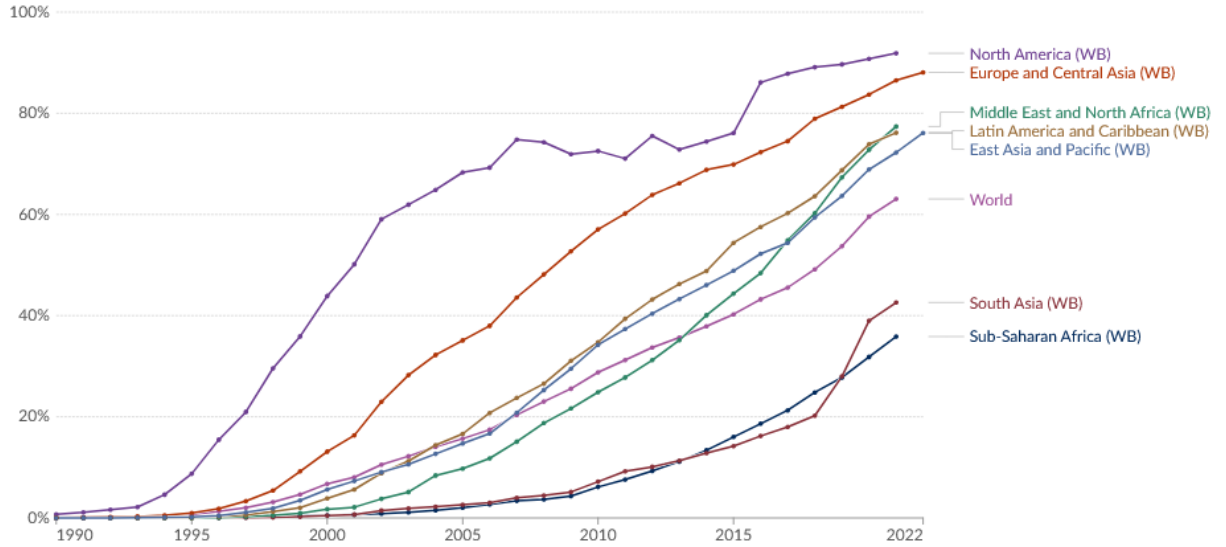
  - How to develop a mathematical model to study their relationship?
    如何建立數學模型來研究它們的關係？

Source: Carmona et al., "Fine-root traits in the global spectrum of plant form and function", Nature (2021)

Example: Modelling Global Internet Usage 全球網絡使用的數學建模

- **Real-world problem 現實生活問題:**
  - Internet access is highly related to education and economic development
    網路存取與教育和經濟發展高度相關
  - Understanding the trend of internet usage is important for policy making by governments and resource allocation by NGOs
    了解網路使用趨勢對於政府制定政策和非政府組織分配資源具有重要意義

- **Mathematical problem 數學問題:**
  - How to estimate the trend of internet access in different regions?
    如何評估不同地區的網路存取趨勢？
  - How to identify the relationship between internet access and other sociological factors?
    如何辨識網路存取與其他社會因素的關係？

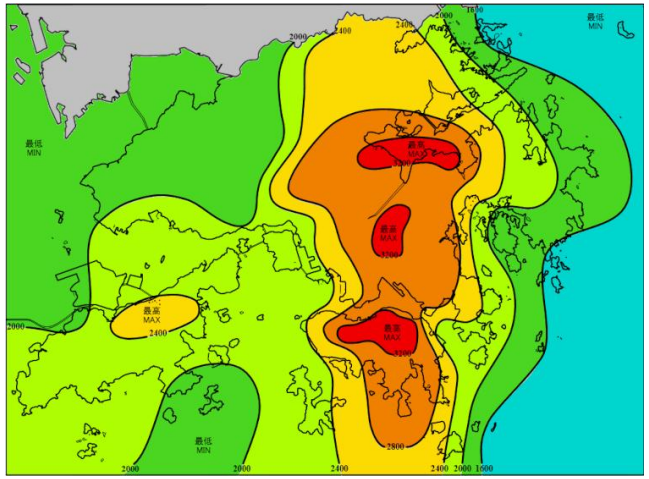


https://en.wikipedia.org/wiki/Global_Internet_usage

Example: Modelling Weather and Climate 天氣和氣候的數學建模

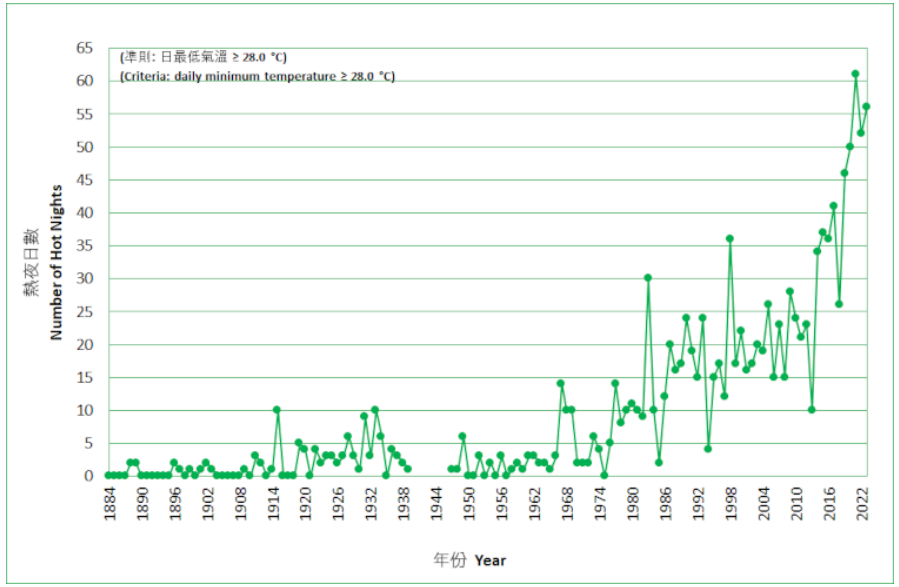
- **Real-world problem 現實生活問題:**
  - Weather prediction 天氣預報
  - Understanding climate change了解氣候變化
  - How can we predict weather or climate change?
    我們如何預測天氣或氣候變遷？

Annual rainfall distribution in 2023
2023年年降雨量分佈



- **Mathematical modelling 數學問題:**
  - Identify the relevant factors: urbanization, deforestation, energy consumption, …
    找出相關因素：城市化、森林砍伐、能源消耗…
  - Develop mathematical models for predicting weather or climate change
    建立數學模型來預測天氣或氣候變遷

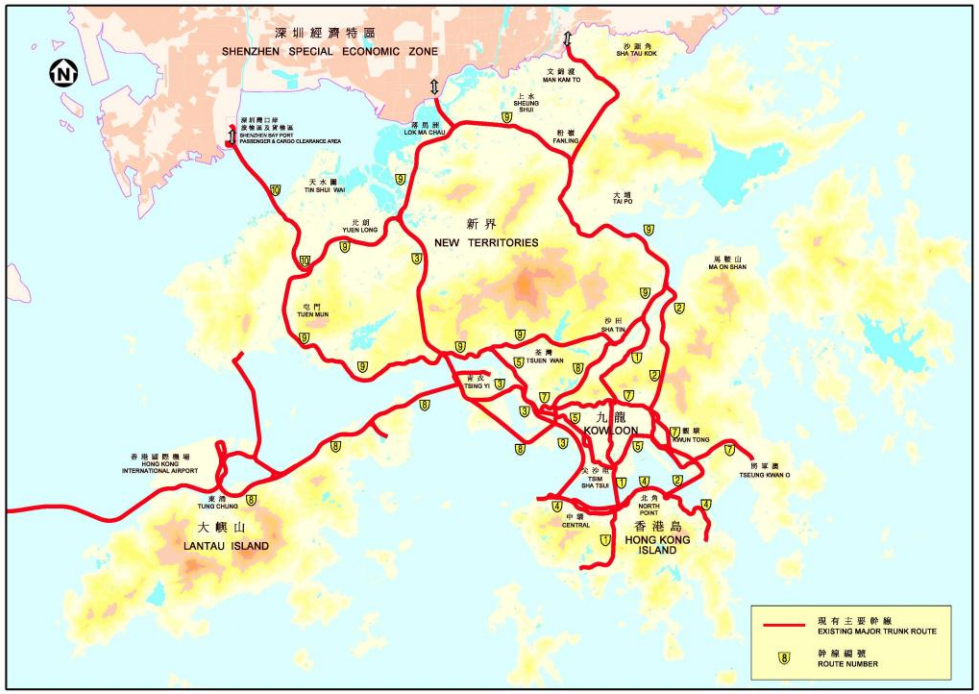
Hong Kong Observatory 香港天文台:
https://www.hko.gov.hk/en/index.html



Number of Hot Nights in Hong Kong 1884-2023
1884 年至 2023 年香港熱夜晚數

Example: Modelling Traffic Condition交通狀況的數學建模

- **Real-world problem 現實生活問題:**
  - Finding an optimal route 尋找最佳路線
  - Traffic control/management 交通管制/管理
  - Road network planning 道路網路規劃
  - How can we analyze traffic condition under different scenarios and study the future trend for better planning?
    我們如何分析不同情況下的交通狀況並研究未來趨勢，以便更好地進行規劃？



- **Mathematical modelling 數學問題:**
  - Identify key factors: population growth, number of private cars, use of public transportation, future urban planning,…
    找出相關因素：人口增長、私家車數量、公共運輸使用情況、未來城市規劃...
  - Develop mathematical models for predicting the trend/assessing the influence of different factors on traffic 建立交通狀況的數學模型

Transport Department 運輸署: https://www.td.gov.hk/en/home/index.html
Highways Department 路政署: https://www.hyd.gov.hk/en/home/index.html

How can we use AI to assist us with math modelling?

我們如何使用AI來協助我們進行數學建模？

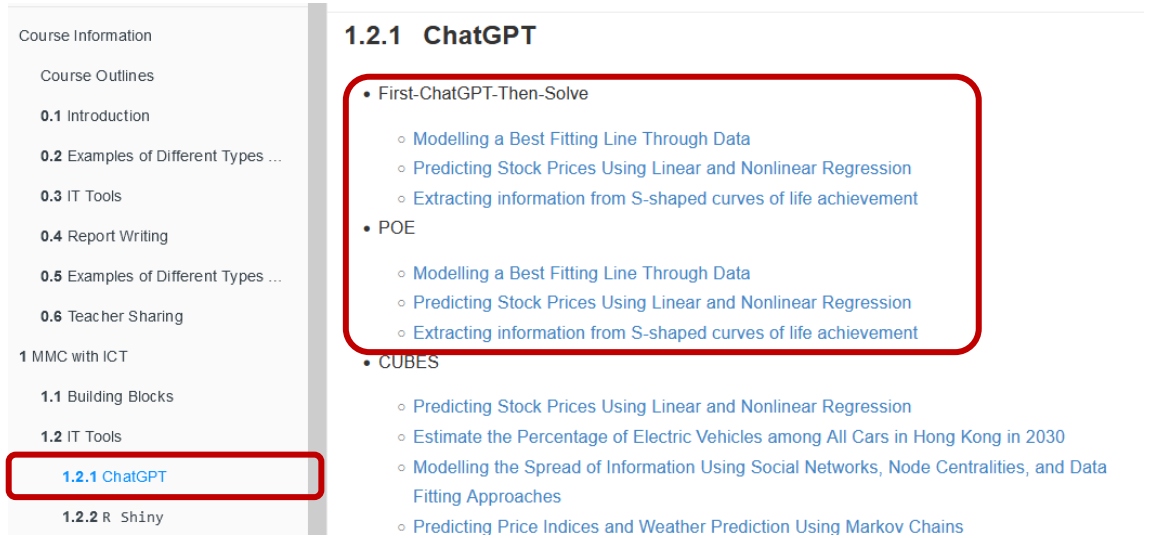
AI tools for math modelling 以 AI 工具輔助數學建模

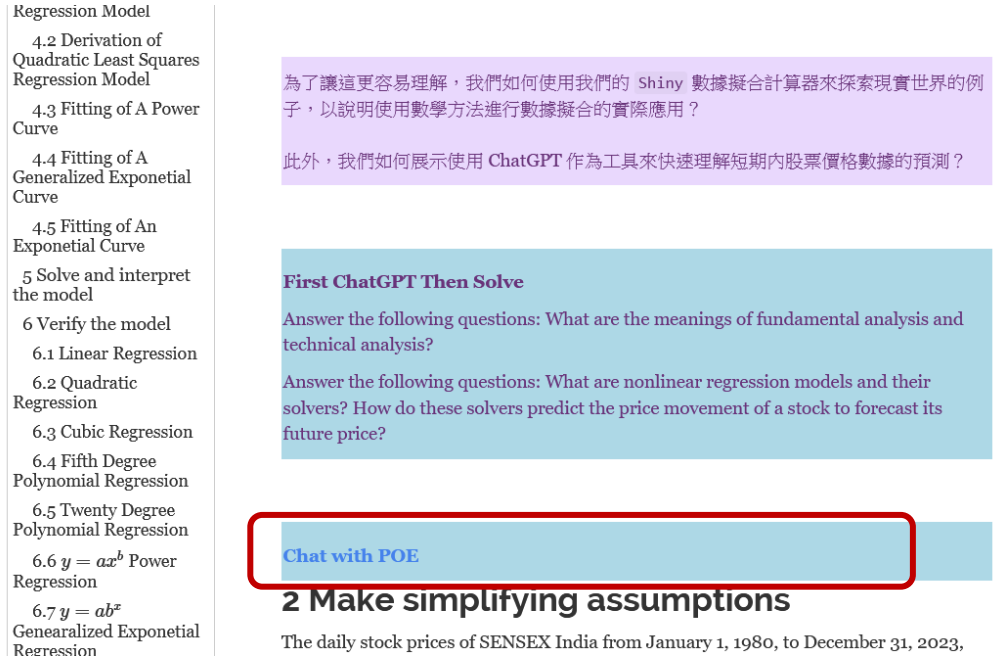
- **How can we use AI to assist us with math modelling?**
  我們如何使用AI來協助我們進行數學建模？

- *First-ChatGPT-Then-Solve (FCTS)* strategy for mathematical modelling

- Use AI-based tools to help us:
  - Understand problem background 了解問題背景
  - Identify relevant factors 找出相關因素
  - Locate datasets 尋找數據集
  - …
  - Fact-checking is important! 核實事實很重要！

- Freely available AI tools 免費AI工具
  - Poe https://poe.com/
  - Microsoft Copilot https://copilot.microsoft.com/

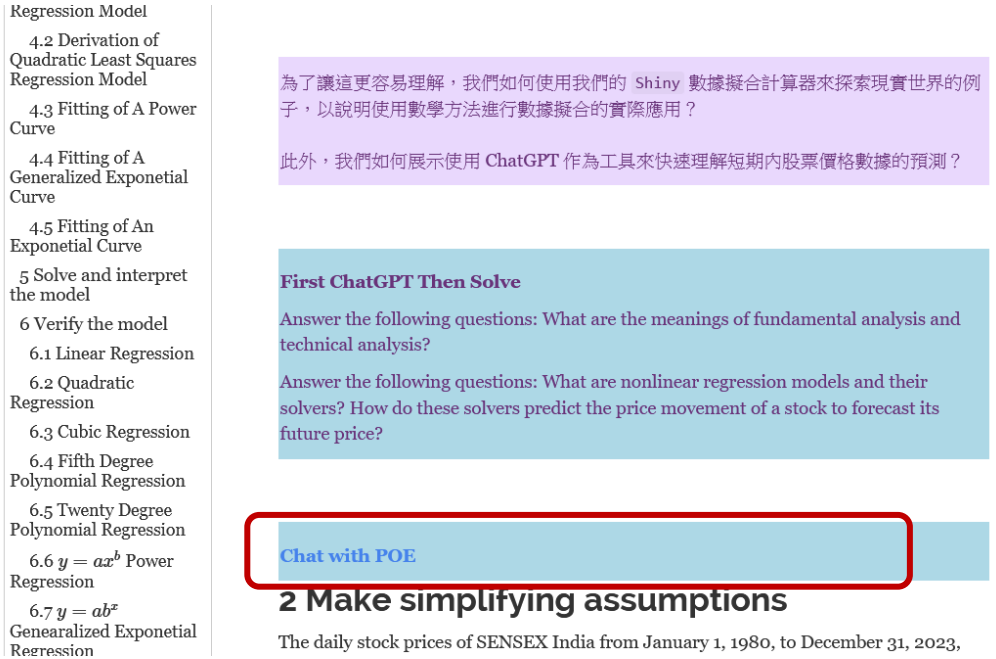




AI tools for math modelling 以 AI 工具輔助數學建模

- The art of asking questions 提問的藝術
  - Questions about the question 關於問題的疑問
  - Questions for clarification purposes 澄清問題
  - Questions that probe assumptions 探究假設的問題
  - Questions that probe for reasons and evidence 探究原因和證據的問題
  - Questions about viewpoints and perspectives 關於觀點和看法的問題
  - Questions that probe implications and consequences 探究意義和後果的問題

- What should be avoided when using AI?
  使用AI時應該避免什麼？
  - Outsourcing your thinking to AI
    把你的想法外判給人工智能
  - Trusting the answers by AI without checking
    不加檢查地相信人工智能的答案
  - Directly taking AI outputs as your answer
    直接將人工智能的輸出作為答案



Common free AI tools 常用免費 AI 工具

- **Poe** (https://poe.com/)
  - Multi-model platform with access to various engines 多模型平台，可存取各種引擎
- **Copilot** (https://copilot.microsoft.com/)
  - Integration with Microsoft functionalities 與 Microsoft 功能的整合
- **Perplexity** (https://www.perplexity.ai/)
  - source-backed answers with citations 基於來源的答案，附有引用
- **Grok** (https://grok.com/)
  - Transparent, shows AI reasoning steps and sources 透明，顯示 AI 推理步驟和來源
- **DeepSeek** (https://www.deepseek.com/)
  - Strong multilingual support (English and Chinese) 強大的多語言支持（英語和中文）

Effective AI prompt (prompt engineering) skill
有效的 AI 提示（提示工程）技能

- AI is very powerful if we use it smartly!
  如果我們聰明地使用 AI，它將非常強大！
  - Important to craft effective instructions 製作有效的指令很重要
  - Guide AI models to provide useful outputs 引導 AI 模型提供有用的輸出

- Be clear and specific 要清晰且具體
  - Avoid ambiguity避免歧義
  - Help AI understand your question and requirement well 幫助 AI 很好地理解您的問題和要求

- Provide context 提供內容
  - Give the AI all the background information it needs 給 AI 所有它需要的背景資訊
  - Relevant location, constraints, … 相關的位置、限制…

Effective AI prompt (prompt engineering) skill
有效的 AI 提示（提示工程）技能

- Iterate on prompts 反覆迭代提示詞
  - back-and-forth clarifications/explanations to refine the answers
    來回澄清/解釋以精煉答案

- Other useful tricks: 其他有用的技巧：
  - "Role prompting": Assign a role to AI so that it gives more tailored outputs
    「角色提示」：為 AI 分配角色，讓它提供更量身訂製的輸出
  - "Explain Like I'm 14": Ask the AI to simplify the explanations based on the prescribed level
    「請用14歲小孩都能聽懂的方式解釋」：要求 AI 根據指定的水平簡化解釋

- Fact-checking is important! 事實查核很重要！
  - Hallucination: AI may produce inaccurate or misleading information!
    幻覺：AI 可能產生不準確或誤導性的資訊！
  - Try multiple AI tools and compare the answers 試用多個 AI 工具並比較答案

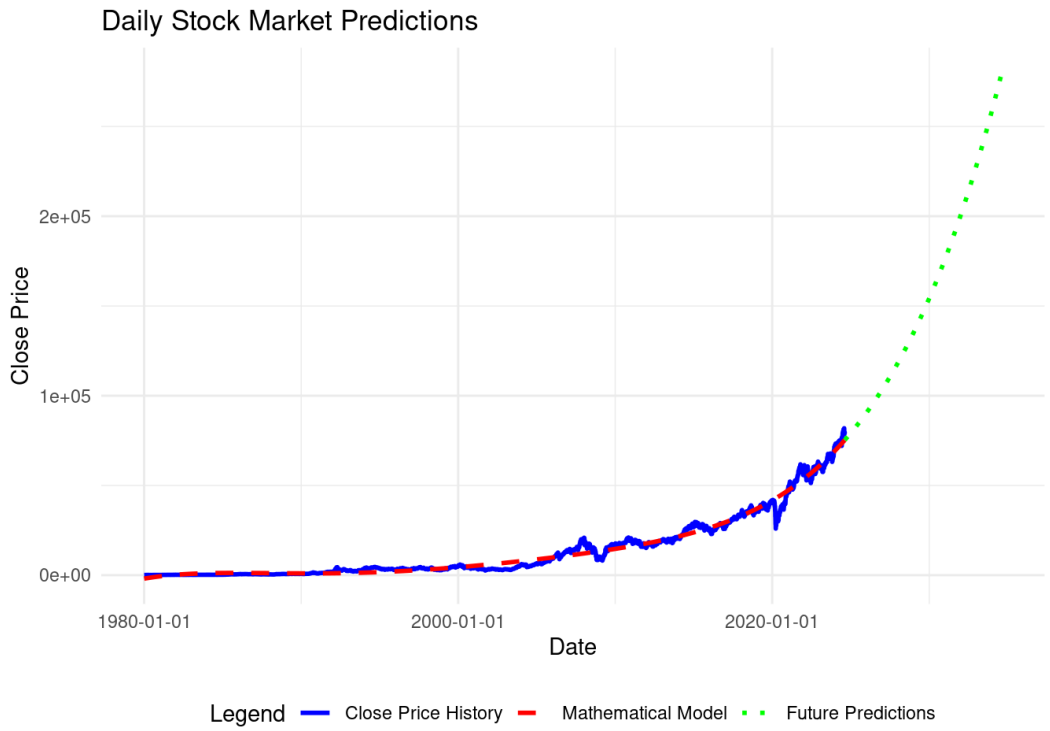
Some examples of mathematical modelling prompts
一些數學建模提示範例

- What exactly are we trying to predict or understand? 我們在試圖預測或理解什麼？
- List five factors that may be important 列出五個可能重要的因素
- What simplifications can we make? 我們可以做哪些簡化？
- Tell me three math concepts or equations that we may use
  告訴我三個我們可能使用的數學概念或方程式
- How could we check if our model is good? 我們如何檢查我們的模型是否良好？
- Tell me three aspects of this model that are oversimplified or might be wrong in real life.
  告訴我這個模型的在現實生活中被過度簡化或可能錯誤的三個方面。

Integrated Examples 綜合例子

- Example with more useful prompting strategies: Modelling for Car Price Analysis
  更多實用提示策略範例：汽車價格分析建模
  https://www.math.cuhk.edu.hk/~mathcal/MM2025/December25a/

- Predicting Stock Prices Using Linear and Nonlinear Regression
  使用線性和非線性迴歸預測股票價格
  http://mathcal.math.cuhk.edu.hk:7537/



Non-Linear Regression with R Shiny (for time data)
利用 R Shiny 進行非線性迴歸 (時序數據)

- https://www.math.cuhk.edu.hk/app/mathmodel/tool.html

- Non-linear regression R Shiny tool (for time data)
  - The x values (column 1) must be some dates 第1 欄必須是日期格式: YYYY-MM-DD, YYYY/MM/DD, DD-MM-YYYY, DD/MM/YYYY, YYYY-MM, YYYY/MM, MM-YYYY, MM/YYYY, or YYYY.
  - Can customize the time interval for data fitting 可以自訂資料擬合的時間段

- Many models available 多種模型可供選擇:
  - Linear model 線性模型
  - Quadratic model 二次模型
  - Cubic model 三次模型
  - Polynomial model 多項式模型
  - Power model 冪模型
  - Exponential model 指數模型
  - Logarithmic model 對數模型

Non-Linear Regression with R Shiny (for time data)
利用 R Shiny 進行非線性迴歸 (時序數據)

Upload your file: ?

Browse... data_NLR.csv

Upload complete

☑ Plot Data point

Time Interval 1

1990-01-31 to 2010-08-31

Model ?

Quadratic Regression

$$y = 0.000448x^2 - 4.82x + 1.5 \cdot 10^4$$

Residual sum of squares: ?

$$RSS = 1.05 \cdot 10^9$$

Prediction for:

2030-01-01

76066.12

Time Interval 2

2010-08-31 to 2024-08-31

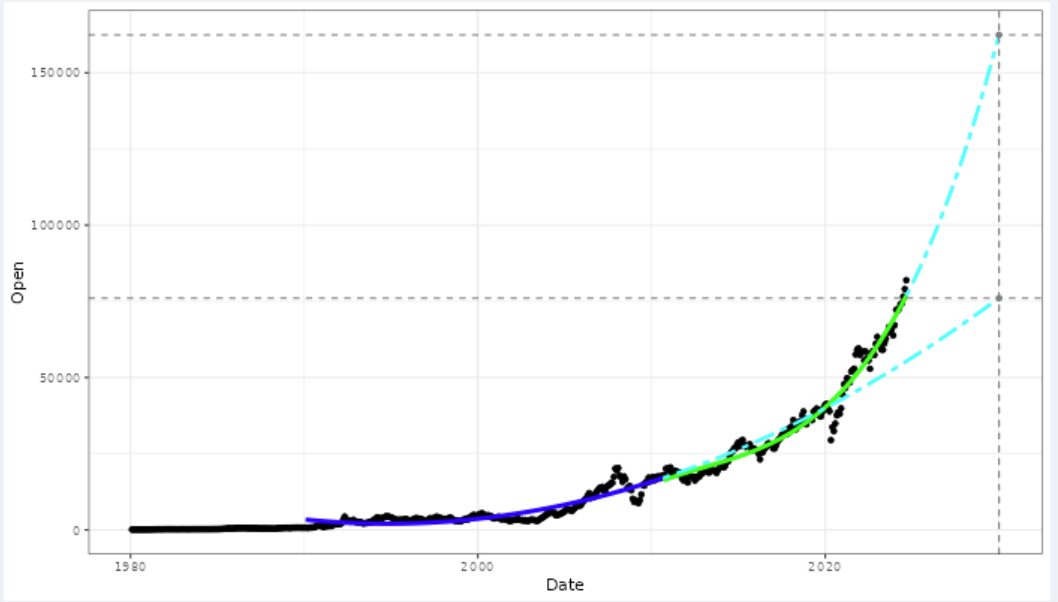
Model ?

Cubic Regression

$$y = 4.7 \cdot 10^{-7}x^3 - 0.017x^2 + 208x - 8.5 \cdot 10^5$$

Residual sum of squares: ?

$$RSS = 1.5 \cdot 10^9$$



Non-Linear Regression with R Shiny (for time data)
利用 R Shiny 進行非線性迴歸 (時序數據)

- **Exercise 練習**

  Non-Linear Regression R Shiny tool 非線性迴歸 R Shiny 工具
  https://www.math.cuhk.edu.hk/app/mathmodel/tool.html

  1. Consider the sample stock market data 股票市場資料 in the tool (also available as data_NLR.csv on our website)

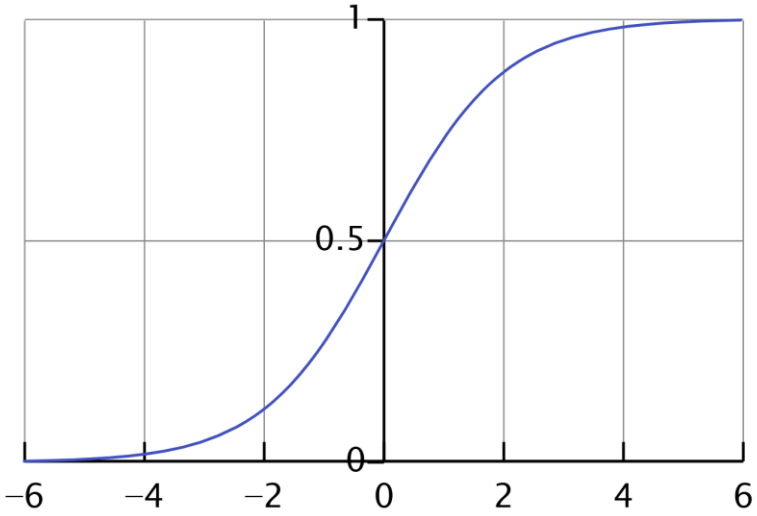
  2. Use the tool to try different models 使用工具嘗試不同的模型

  3. For different time periods, consider different models. What do you observe? 對於不同的時間段，考慮不同的模型。你觀察到什麼？

Advanced methods and tools for mathematical modelling

數學建模進階方法及工具

Modelling with Sigmoidal functions (S-shaped curves)
使用 S 形函數建模



- Sigmoidal function is a type of functions whose graph has a characteristic S-shaped curve.
  S 型函數是一種函數，因其函數圖像形狀像字母 S 得名。

- Features 特色:
  - Progress from small beginnings 從微小開始
  - Then accelerates 逐漸加速
  - Approaches a maximum over time 隨著時間的推移接近最大值

- More examples of S-shaped curves S 形函數的更多例子
  https://www.math.cuhk.edu.hk/~mathcal/MM/Sigmoidal.html

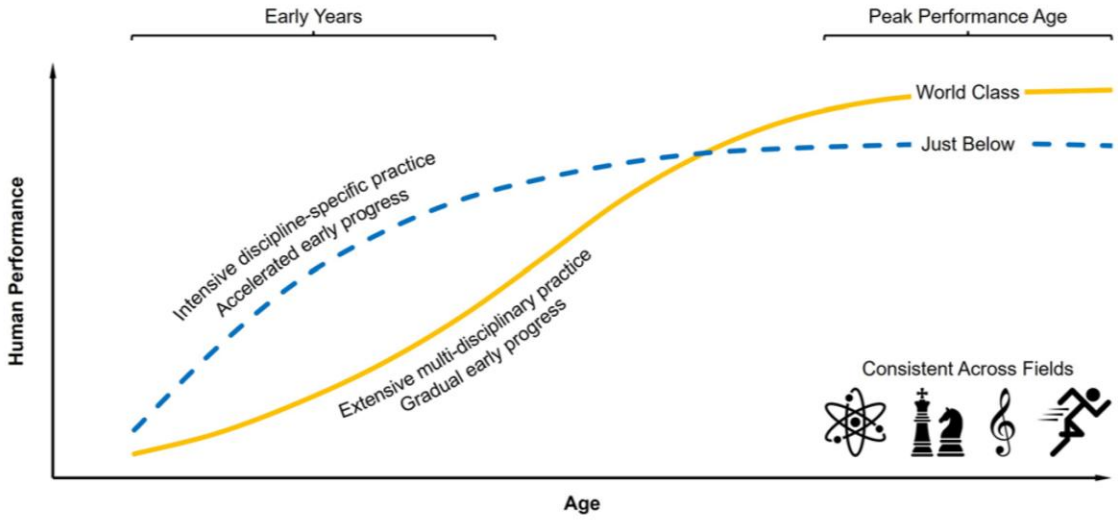
- Modelling life achievement using S-shaped curves透過 S 形函數進行人生成就的建模
  http://mathcal.math.cuhk.edu.hk:7562/

S-shaped curves of life achievement 人生成就的 S 形函數

- Modelling using S-shaped curves actually matches latest scientific findings!
  使用 S 形曲線的建模實際上符合最新的科學發現！

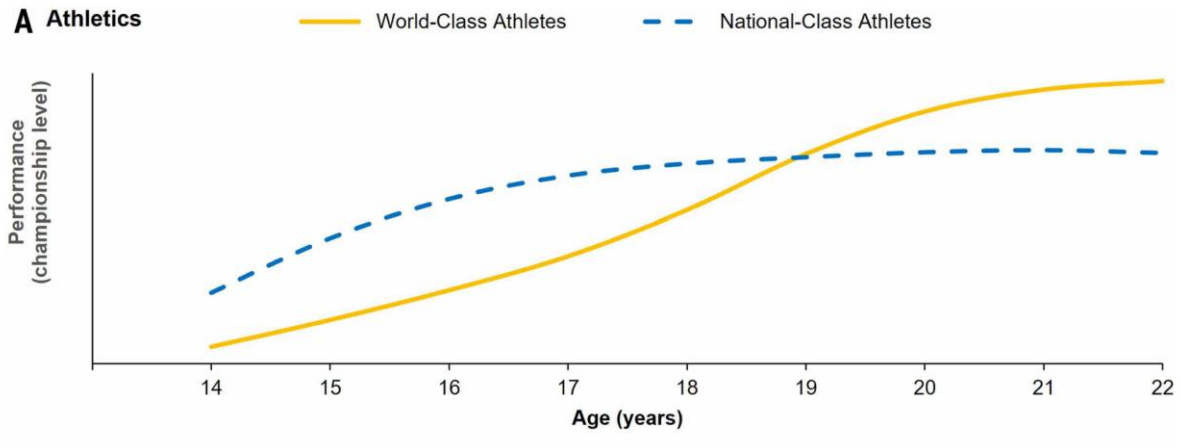
- *Recent discoveries on the acquisition of the highest levels of human performance*, **Science**, December 2025 *關於獲得人類最高水平表現的最新發現*，《科學》，2025 年 12 月 https://www.science.org/doi/10.1126/science.adt7790

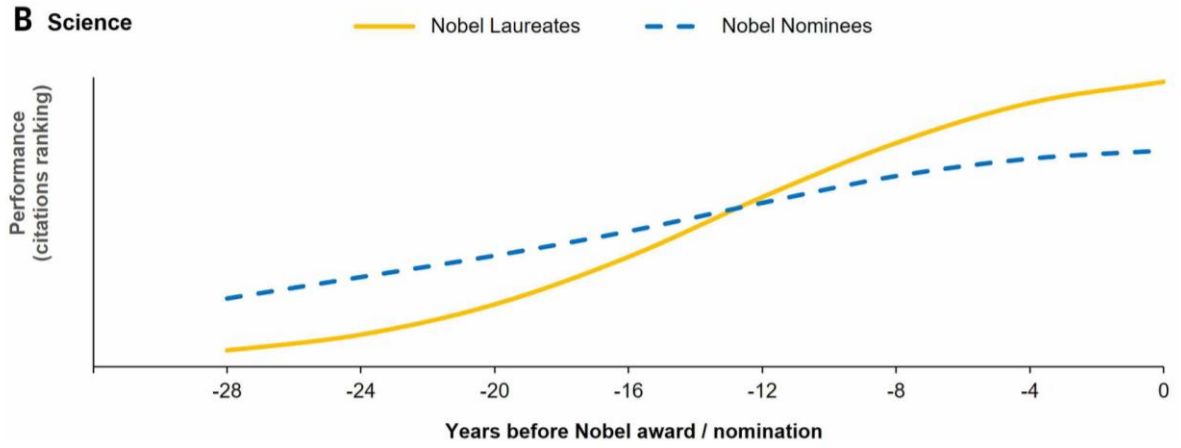
- Analyzed the development of more than 34,000 adult international top performers in different domains, including Nobel laureates, the most renowned classical music composers, Olympic champions, and the world's best chess players.
  綜合研究了 34,000 多名不同領域的國際頂尖成年人的發展情況，其中包括諾貝爾獎得主、最著名的古典音樂作曲家、奧運冠軍和世界上最好的國際象棋棋手。



S-shaped curves of life achievement 人生成就的 S 形函數

- Very similar developmental pattern of world-class performers across different domains 不同領域世界級人才的發展模式非常相似！
  - (A) Performance level of 508 world- class versus 420 national- class athletes from age 14 to 22 years.
    508 名世界級運動員與420名國家級運動員（年齡在14至22歲之間）的競賽表現比較。
  - (B) Performance level of 330 physics and chemistry Nobel laureates versus 1595 physics and chemistry nominees who have not earned the Nobel Prize. 330名諾貝爾物理學獎和化學獎得主與1595名未獲諾貝爾獎的物理和化學獎提名者的競技水平對比。





General Fitting XY R Shiny tool XY數據一般擬合 R Shiny 工具

- How can we fit datasets using S-shaped curves or other complicated functions?
  如何使用 S 形曲線或其他複雜函數擬合數據？

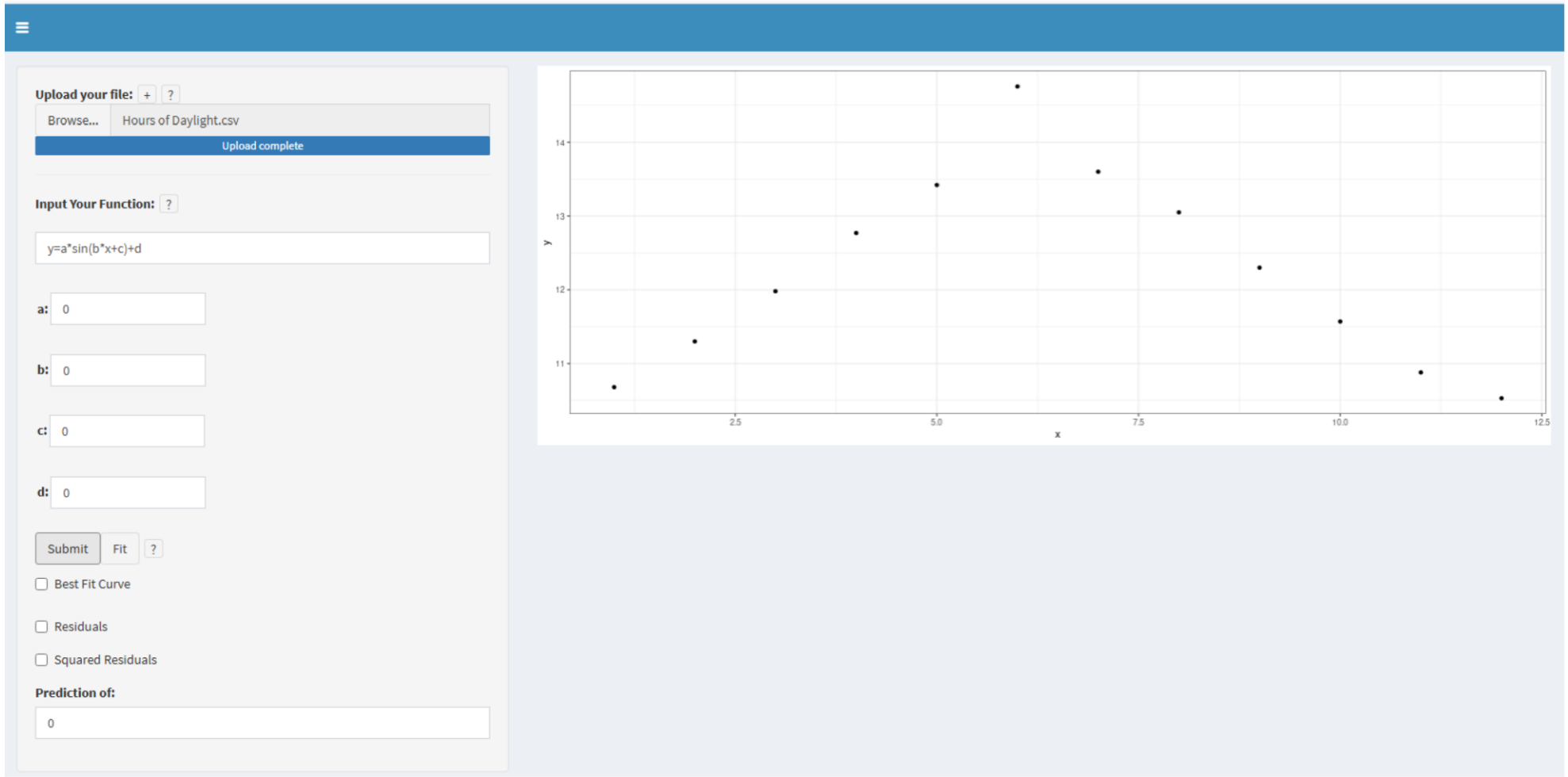
- **General Fitting XY R Shiny tool XY數據一般擬合 R Shiny 工具**
  https://mathmodelcuhk.shinyapps.io/general-fitting/

- Key functionalities:主要功能：
  - Easily fit different customized functions including 輕鬆適配各種自訂函數，包括：
    - Polynomial functions 多項式函數
    - Trigonometric functions 三角函數
    - Logarithmic functions 對數函數
    - Exponential functions 指數函數

  - Different real-life data can be used as input 可使用各種實際數據作為輸入

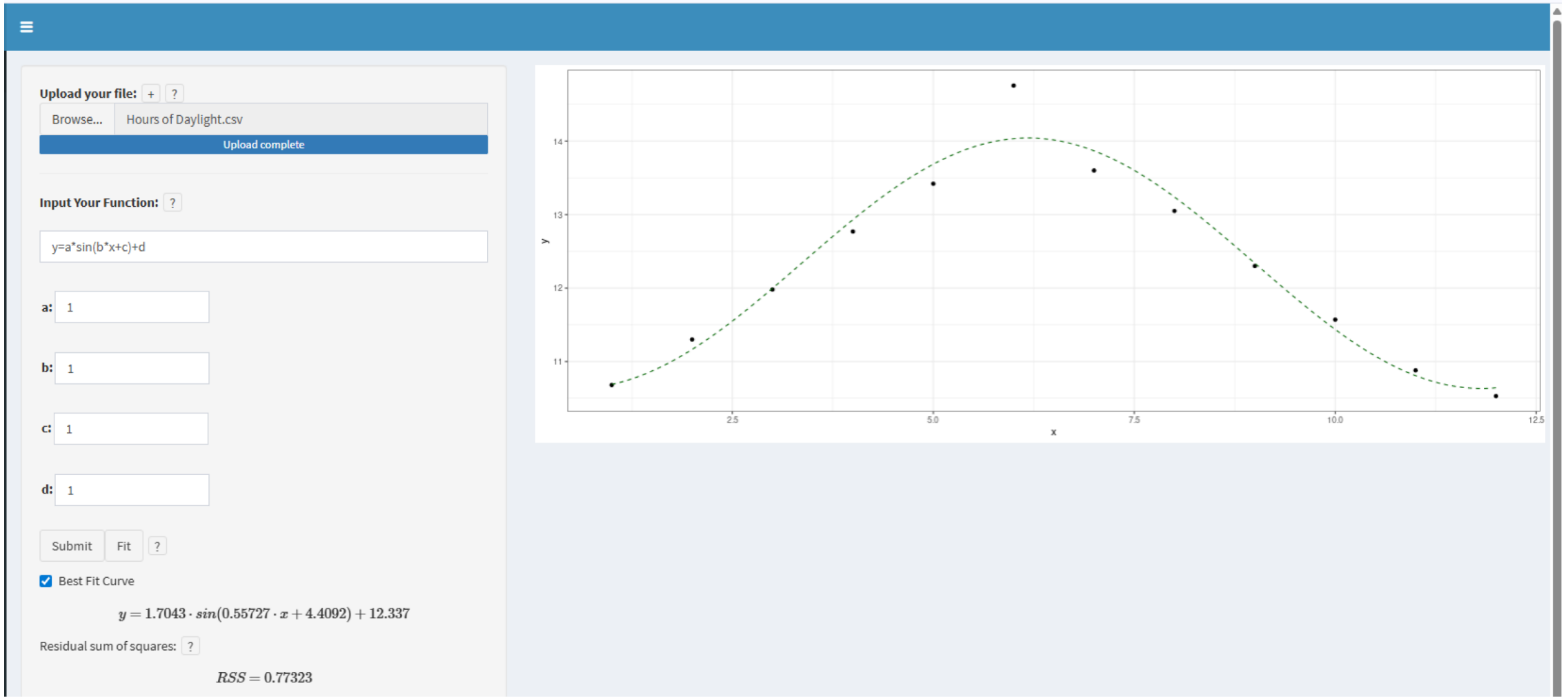
General Fitting XY R Shiny tool XY數據一般擬合 R Shiny 工具

- Loading dataset and prescribing the desired form of the function
  載入數據集並指定所需的函數形式



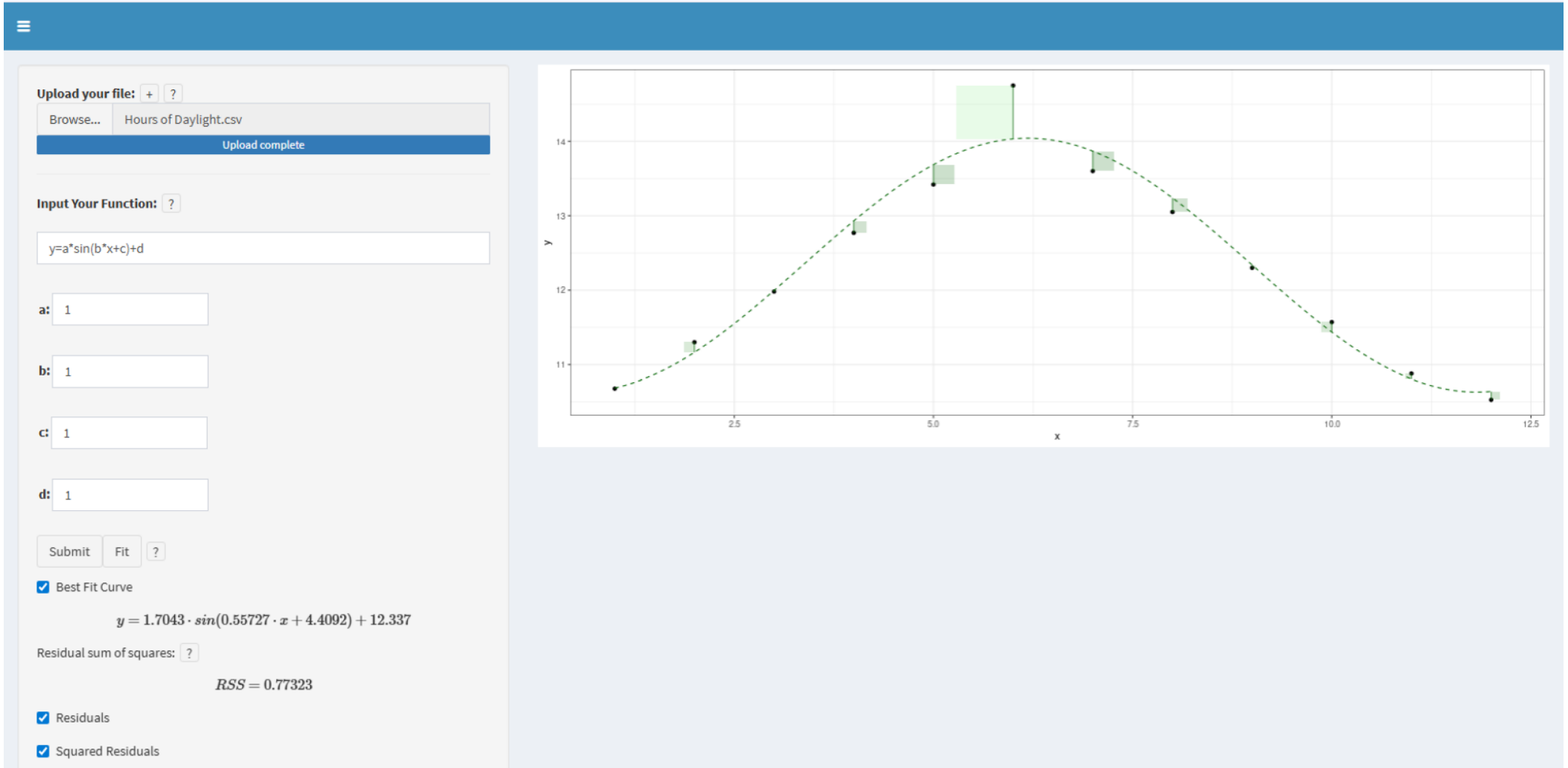
General Fitting XY R Shiny tool XY數據一般擬合 R Shiny 工具

- Obtain the best-fit model 得出最佳擬合模型



General Fitting XY R Shiny tool XY數據一般擬合 R Shiny 工具

- Evaluate residuals and errors 評估殘差和誤差



Exercise: the growth of common water hyacinth 水葫蘆的生長

• 高中数学建模活动设计 •

水葫芦的生长

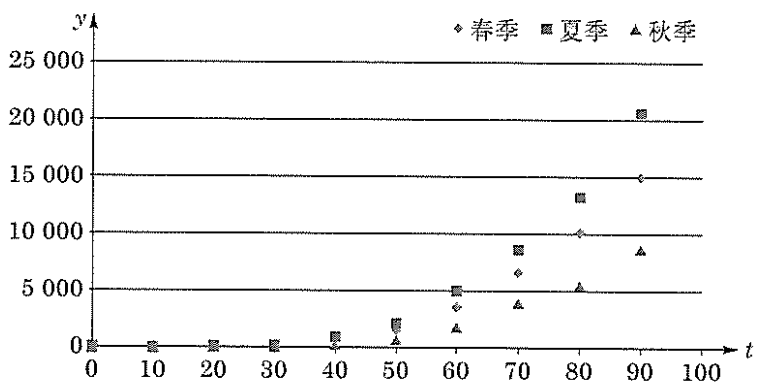
(选择性必修·第三册)

(教材、教学设计、学生活动手册)

表 4-1　水葫芦在春、夏、秋三个季节的植物量数据

调查相隔时段/天	春季	夏季	秋季
0	25.01	21.17	26.83
10	57.77	46.59	46.67
20	126.79	116.53	82.87
30	254.74	301.94	162.70
40	625.95	878.01	361.05
50	1 578.94	2 166.43	842.95
60	3 621.85	5 085.05	1 866.37
70	6 721.45	8 620.32	3 972.83
80	10 189.24	13 298.84	5 644.10
90	15 009.17	20 713.92	8 778.56

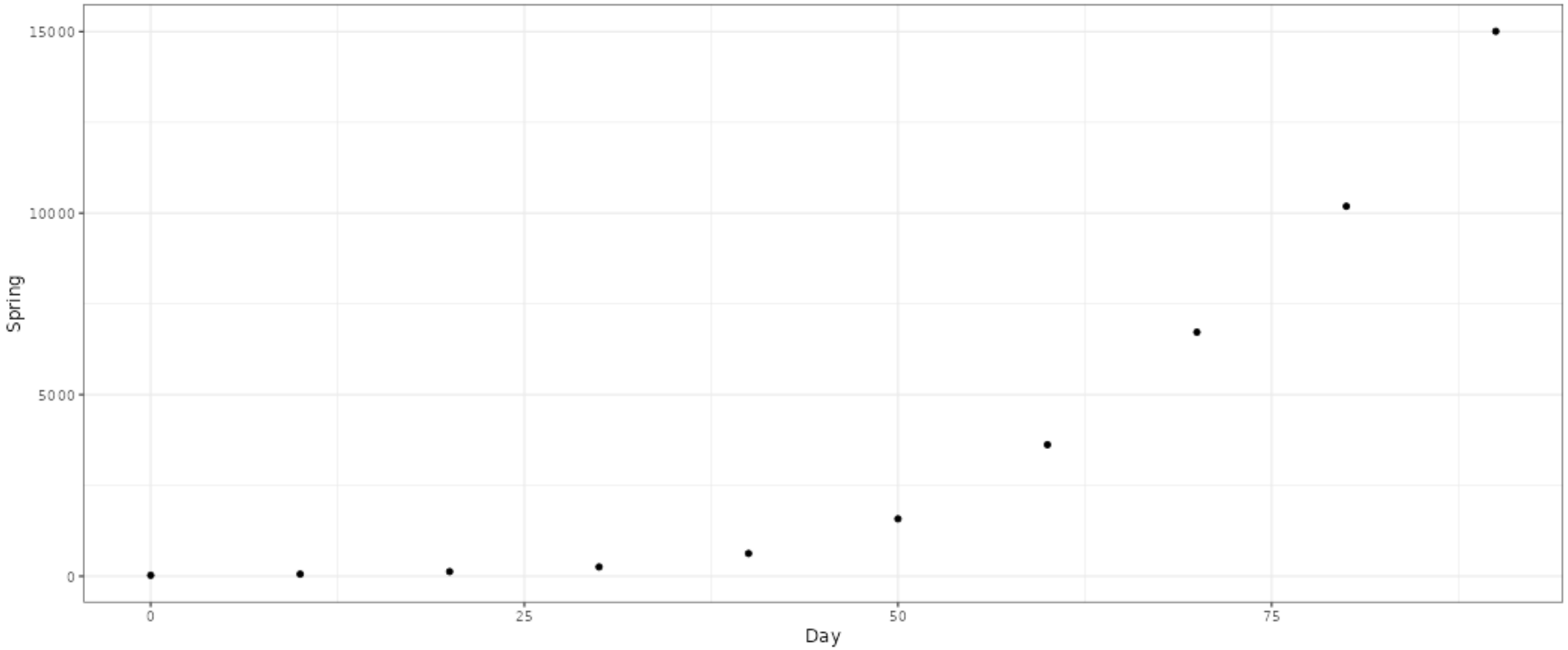
注：表中数据为 30 株水葫芦调查结果的平均值，其中调查时段 0 代表水葫芦接入时的初始值. 数据源于华南农业大学冯熠荣(2003)的《水葫芦种群生态控制的基础研究》.



Exercise: the growth of common water hyacinth 水葫蘆的生長

- https://www.math.cuhk.edu.hk/app/mathmodel/tool.html
  Use different tools to analyze the growth data in Spring
  使用不同工具來分析春季成長數據

Day	Spring
0	25.01
10	57.77
20	126.79
30	254.74
40	625.95
50	1578.94
60	3621.8
70	6721.45
80	10189.24
90	15009.17



Exercise: the growth of common water hyacinth 水葫蘆的生長

- **Let's try the polynomial model 試試多項式模型!**

Upload your file: ?

Browse... Ex1Growth-Spring.csv

Upload complete

☑ Best Fit Curve

Model ?

Quadratic Regression

$$y = 3.16x^2 - 133x + 799$$

Residual sum of squares: ?

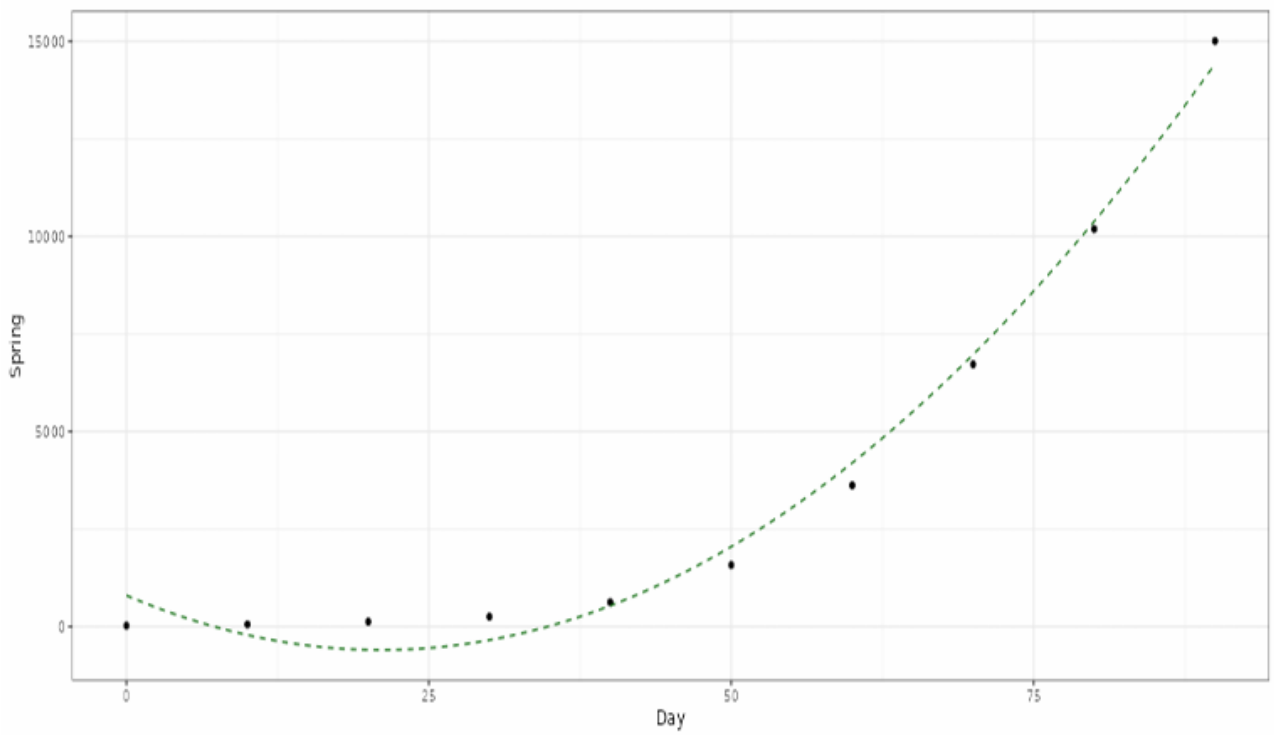
$$RSS = 2.56 \cdot 10^6$$

☐ Residuals

☐ Squared Residuals

Prediction of:

0

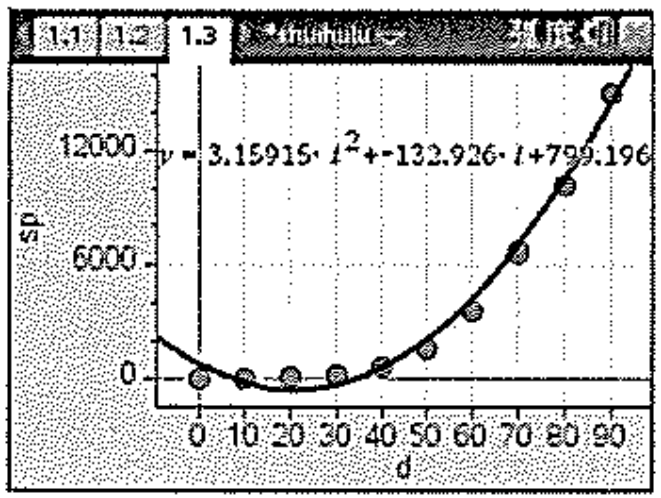


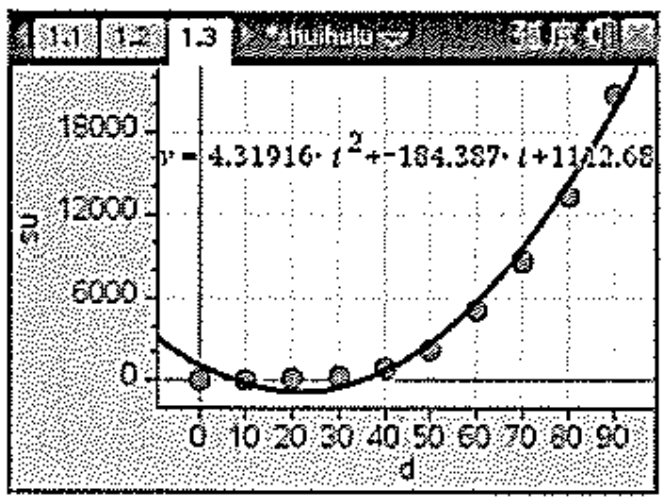
Exercise: the growth of common water hyacinth 水葫蘆的生長

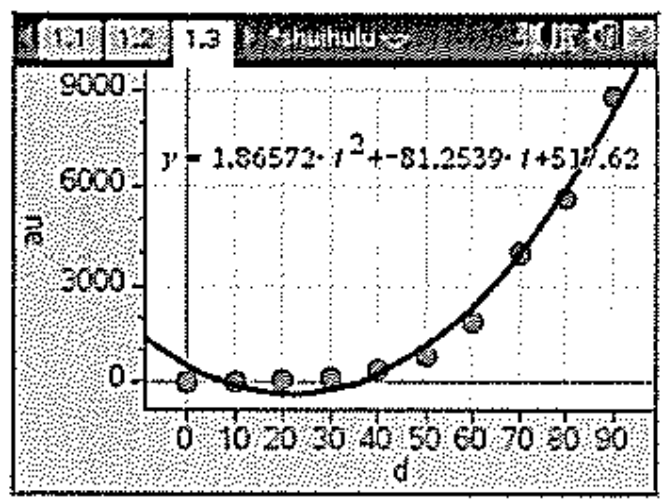
春季：$y=799.196-132.926t+3.159\,15t^2$，

夏季：$y=1\,112.68-184.387t+4.319\,16t^2$，

秋季：$y=517.62-81.253\,9t+1.865\,72t^2$，



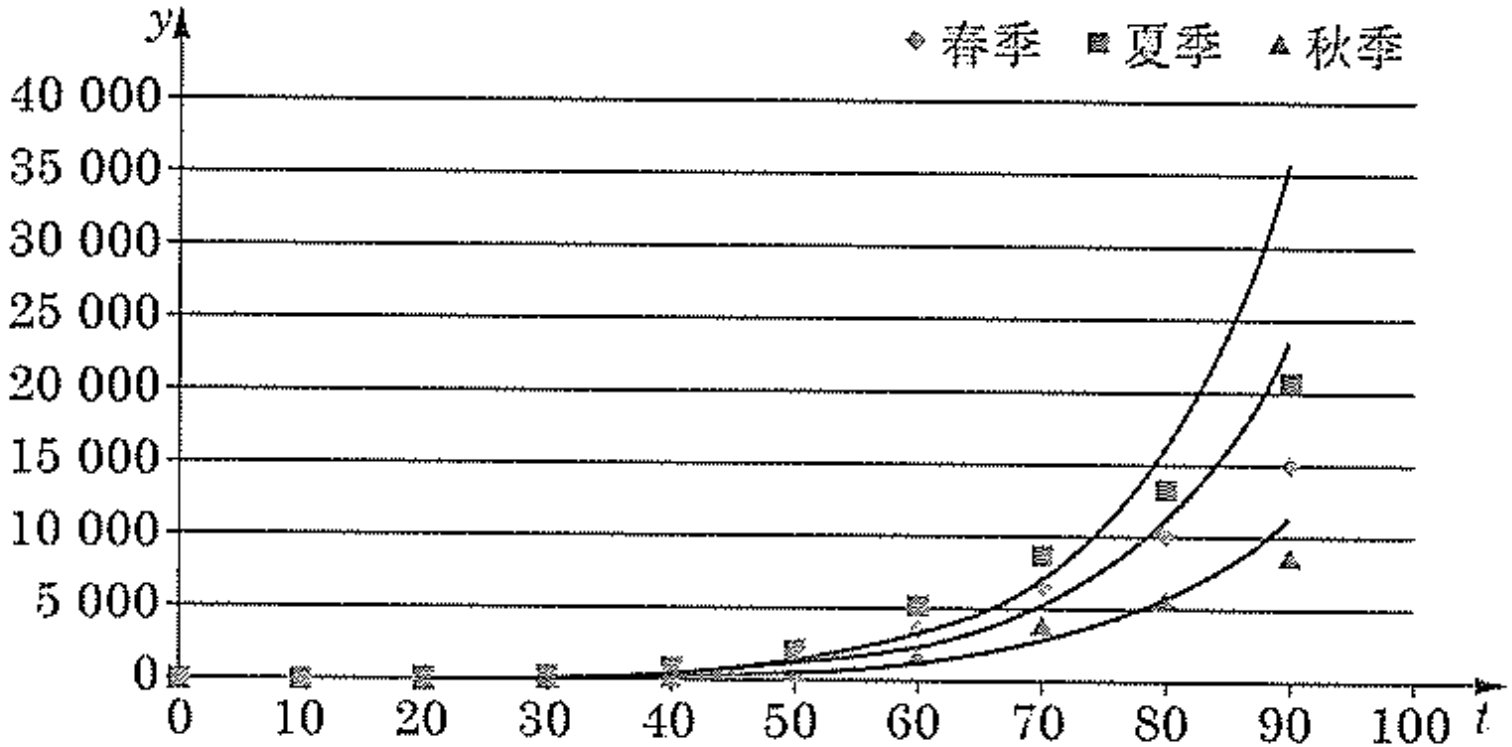




Exercise: the growth of common water hyacinth 水葫蘆的生長

- **Let's also try the exponential model 試試指數模型!**

线性回归模型：$\ln y=at+b$，其中 a 和 b 是待定系数. 记 $\ln y$ 为 s(即 $s=at+b$)，可得三个季节中 s 与 t 的线性拟合模型：春季为 $s=0.080\ 3t+3.27$，夏季为 $s=0.074\ 3t+3.39$，秋季为 $s=0.068\ 6t+3.19$. 将这些函数还原为指数函数，就可以分别得到：春季的生长模型为 $y=26.3\times1.08^t$，夏季的生长模型为 $y=29.7\times1.08^t$，秋季的生长模型为 $y=24.3\times1.07^t$



Exercise: the growth of common water hyacinth 水葫蘆的生長

Upload your file: ?

Browse... Ex1Growth-Spring.csv

Upload complete

☑ Best Fit Curve

Model ?

Exponential Regression

$$y = 29.7 \cdot 1.08^x$$

Residual sum of squares: ?

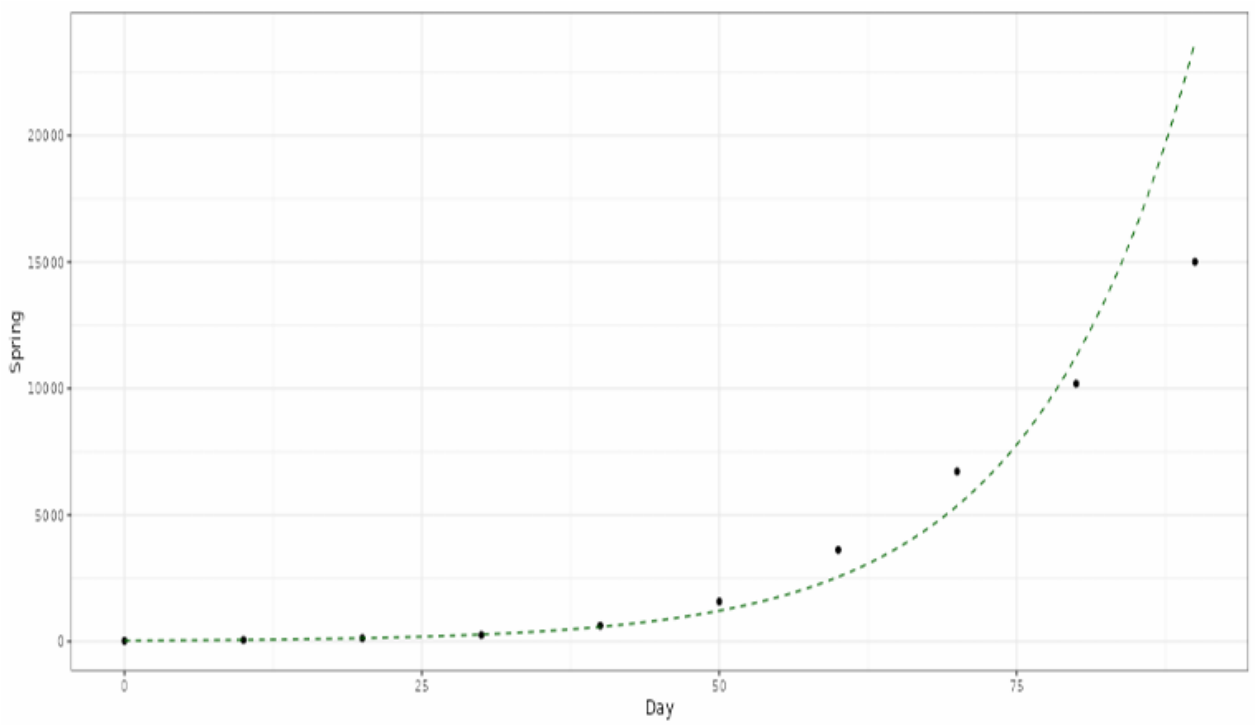
$$RSS = 7.98 \cdot 10^7$$

☐ Residuals

☐ Squared Residuals

Prediction of:

0



Exercise: the growth of common water hyacinth 水葫蘆的生長

- In fact, biological growth may be limited by different environmental factors and cannot be unlimited!
  事實上，生物生長可能受到各種環境因素的限制，未必能無限增長！

- Based on this phenomenon, Mathematical biologist P. F. Verhulst adjusted the exponential growth model in the 1840s and proposed the famous Logistic Function model (a type of S-shaped curve model):
  基於這一現象，數學生物學家 P. F. Verhulst 在 1840 年間調整了指數增長模型，並提出了著名的邏輯函數模型（一種 S 形曲線模型）：

$$y = \frac{A}{1 + e^{b-kx}}$$

春季：$y=\frac{23\,005.460}{1+\mathrm{e}^{6.402-0.078t}}$，$0\leqslant t\leqslant 90$；

夏季：$y=\frac{39\,605.959}{1+\mathrm{e}^{6.254-0.070t}}$，$0\leqslant t\leqslant 90$；

秋季：$y=\frac{14\,499.129}{1+\mathrm{e}^{6.446-0.076t}}$，$0\leqslant t\leqslant 90$.

- **Let's also try this model using the general fitting R Shiny tools!**
  我們不妨也用一般擬合 R Shiny 工具試試這個模型！

Other IT tools for math modelling 其他數學建模的IT工具

- There are many different freely available IT tools suitable for math modelling!
  有很多免費的IT工具適用於數學建模！

- **Microsoft Excel**
  - Statistics 統計, visualization 視覺化, …

- **GeoGebra**
  - Geometric modelling 幾何建模, functions 函數, …

- **Desmos**
  - Functions and graphs 函數和圖形, data fitting 數據擬合, …

- **R Shiny** (by CUHK Mathematics 香港中文大學數學系開發)
  - Regression (linear, nonlinear, multiple, general fitting, …)
    迴歸（線性、非線性、多元、一般擬合…）, Probability 概率, …

- **Other programming tools 其他程式設計工具 (Python, R, …)**
  - More functionalities and greater flexibilities 更多功能和更大靈活性

Microsoft Excel for math modelling 用於數學建模

- Many functionalities useful for math modelling 許多對數學建模有用的功能：
  - Statistics 統計分析
  - Regression analysis 迴歸分析
  - Data visualization (scatter plots, bar charts, …) 數據視覺化（散點圖、棒形圖等）

Paste | B I U | Clipboard | Font

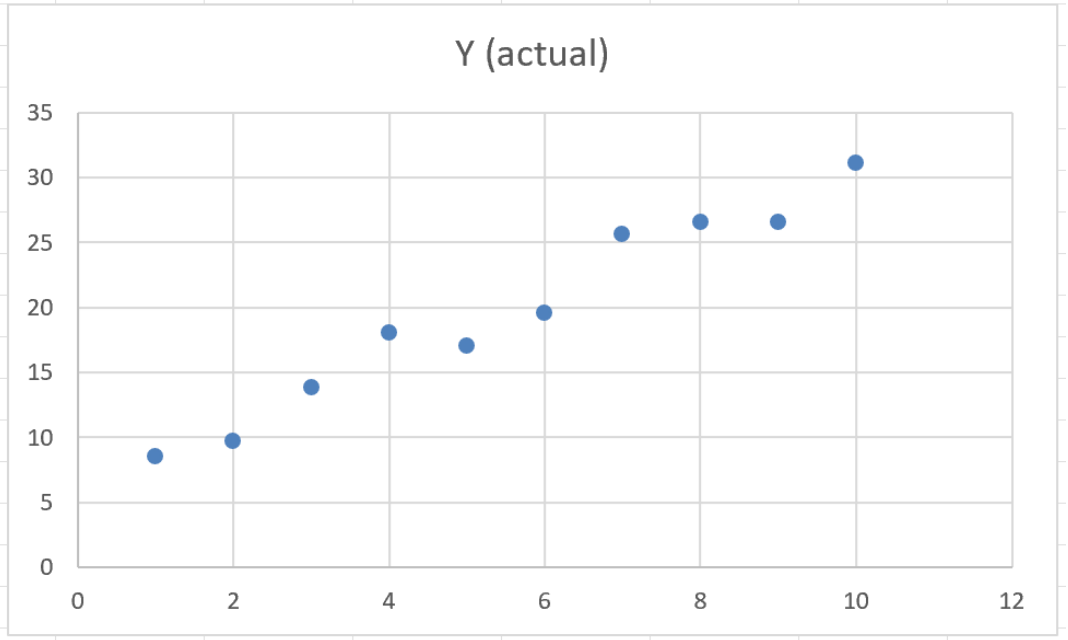
G15

	A	B	C	D
1	X	Y (actual)	Y (predicted)	Residuals
2	1	8.49	8.46	0.04
3	2	9.72	10.94	-1.22
4	3	13.8	13.43	0.36
5	4	18.05	15.92	2.13
6	5	17.03	18.4	-1.37
7	6	19.53	20.89	-1.36
8	7	25.66	23.38	2.28
9	8	26.53	25.86	0.67
10	9	26.56	28.35	-1.79
11	10	31.09	30.83	0.25
12				
13				

Clipboard | Font

I15

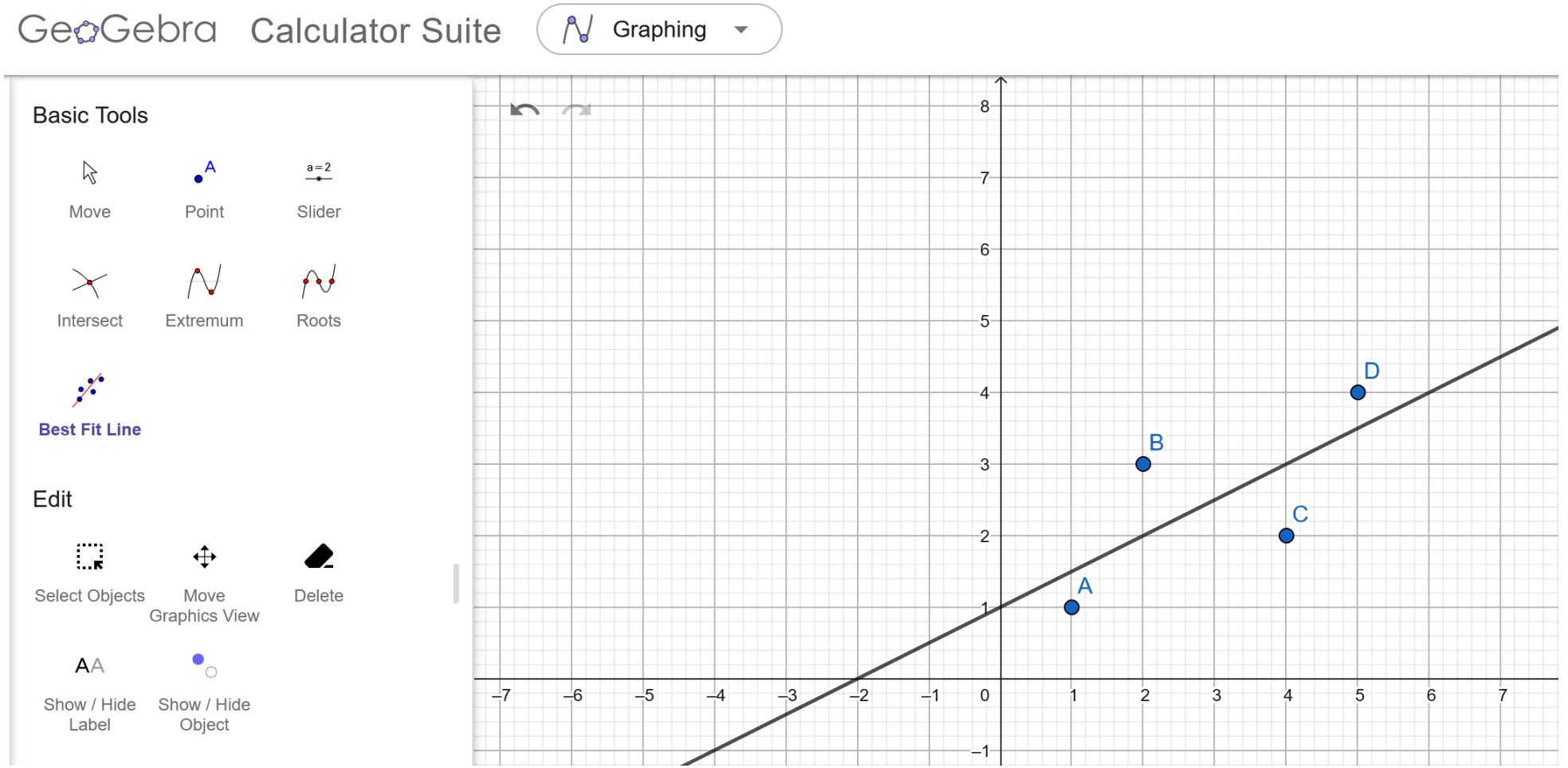
	A	B	C
1	Metric	Value	
2	Mean X	5.5	
3	Mean Y	19.64612	
4	Std X	3.02765	
5	Std Y	7.664861	
6	Correlation	0.982058	
7	Slope (b)	2.486199	
8	Intercept (a)	5.972026	
9			



GeoGebra for math modelling 用於數學建模

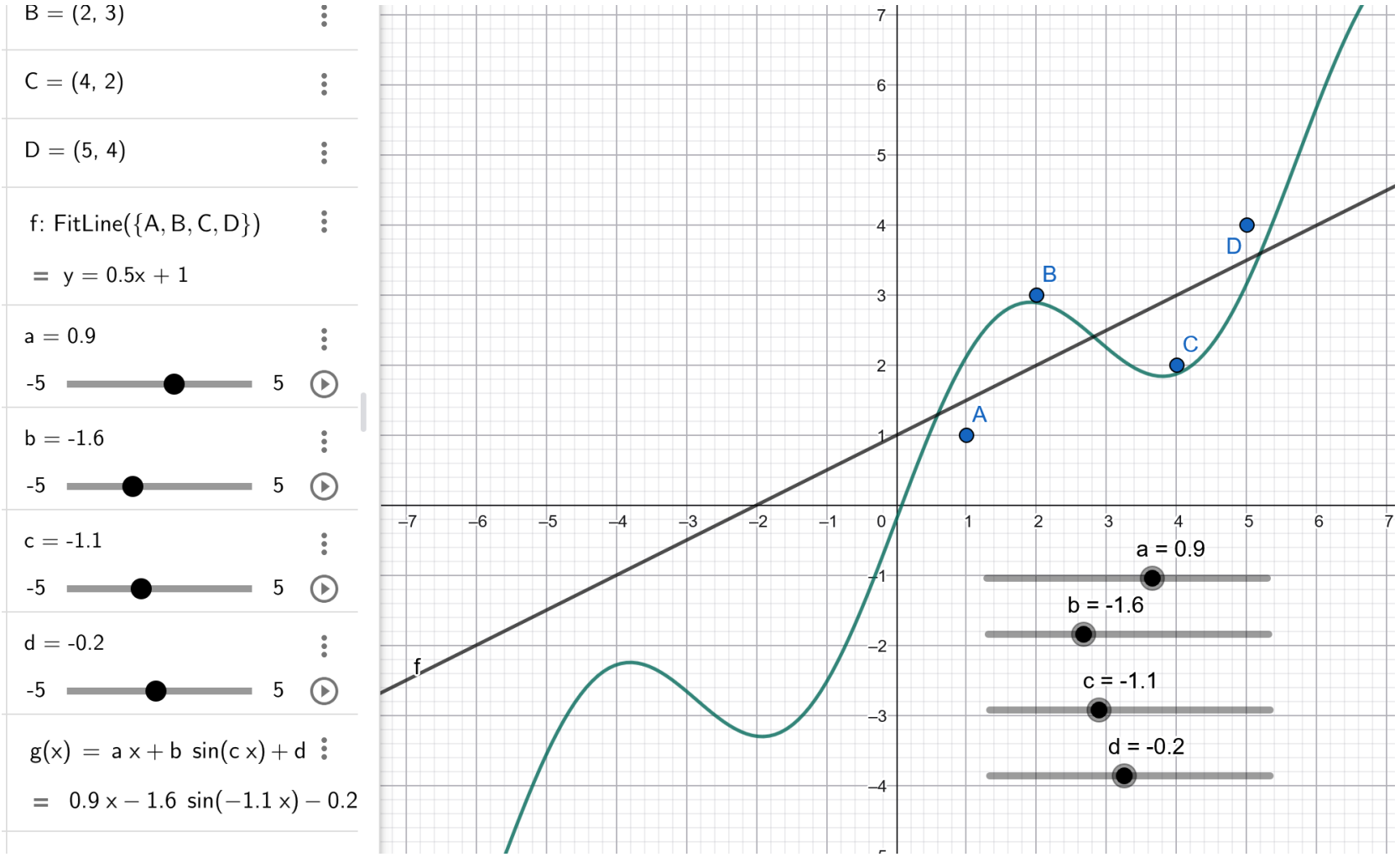
- **GeoGebra** (https://www.geogebra.org/)

- Adding data points and regression lines 新增數據點和迴歸線



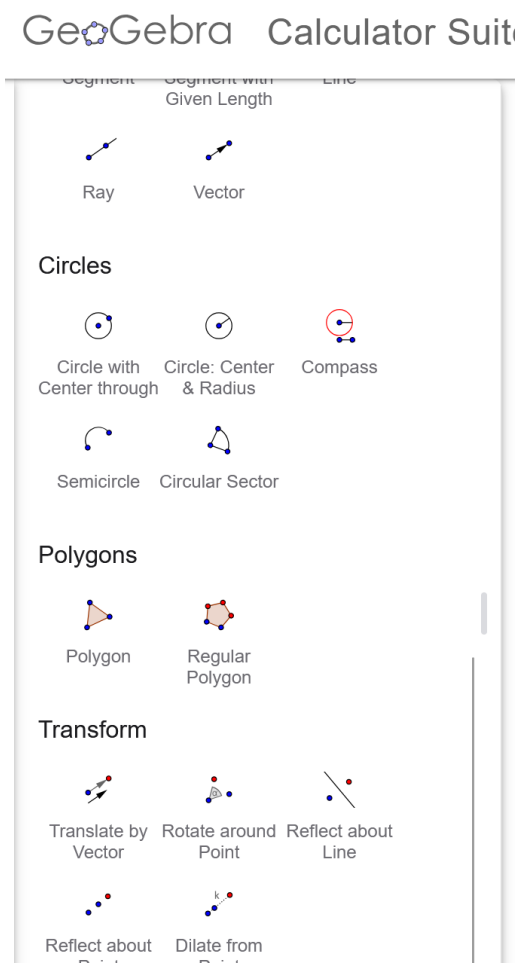
GeoGebra for math modelling 用於數學建模

- Modelling with functions interactively 使用函數進行互動式建模



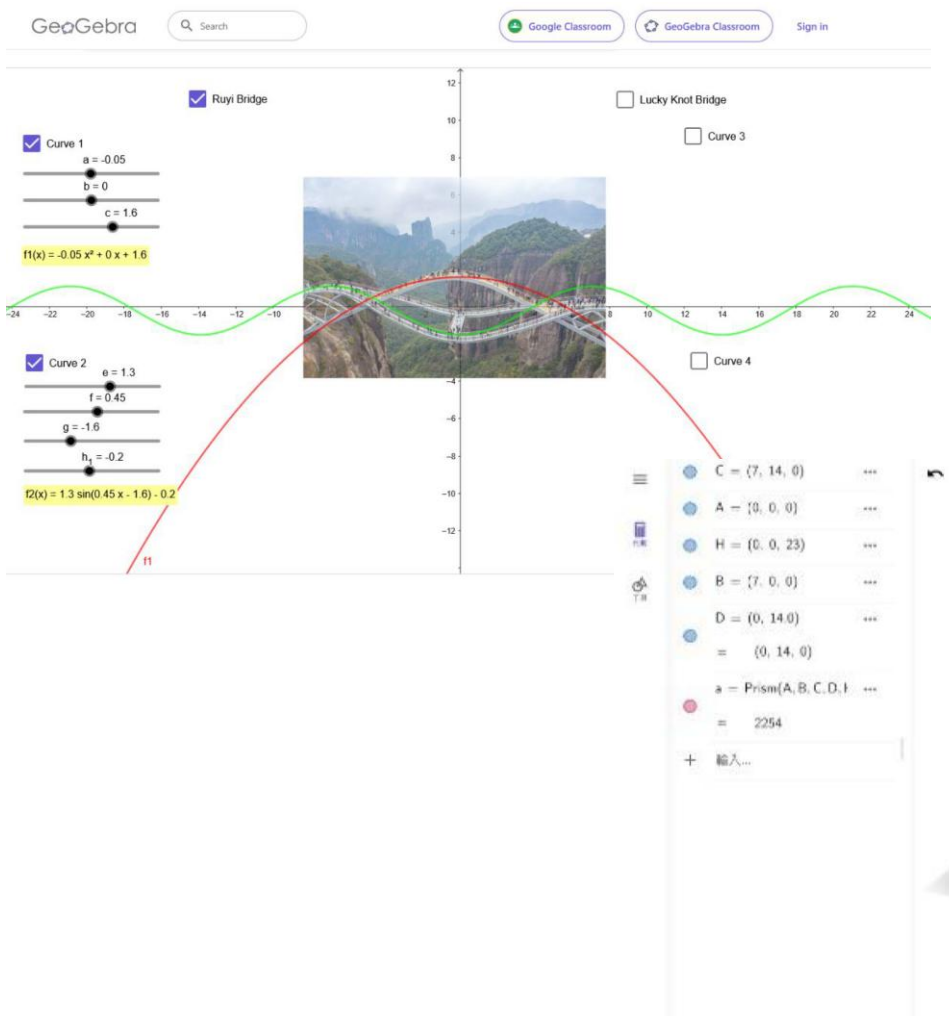
GeoGebra for math modelling 用於數學建模

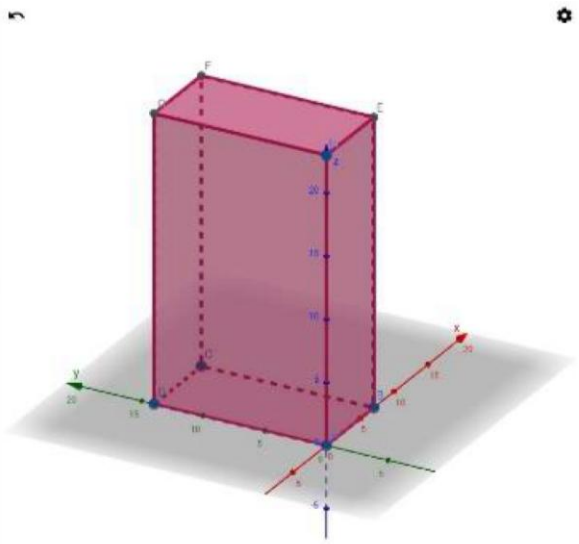
- Geometric modelling 幾何建模



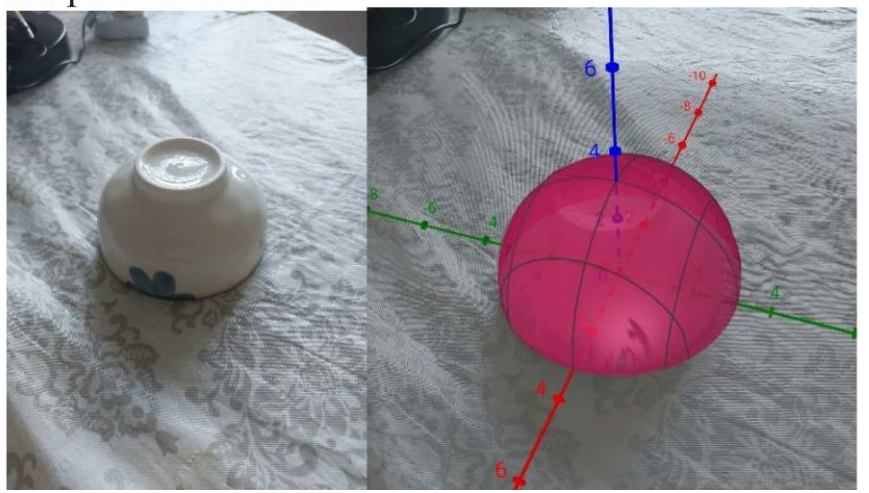
GeoGebra for math modelling 用於數學建模

- Other interactive functionalities for modelling
  其他用於建模的互動式功能

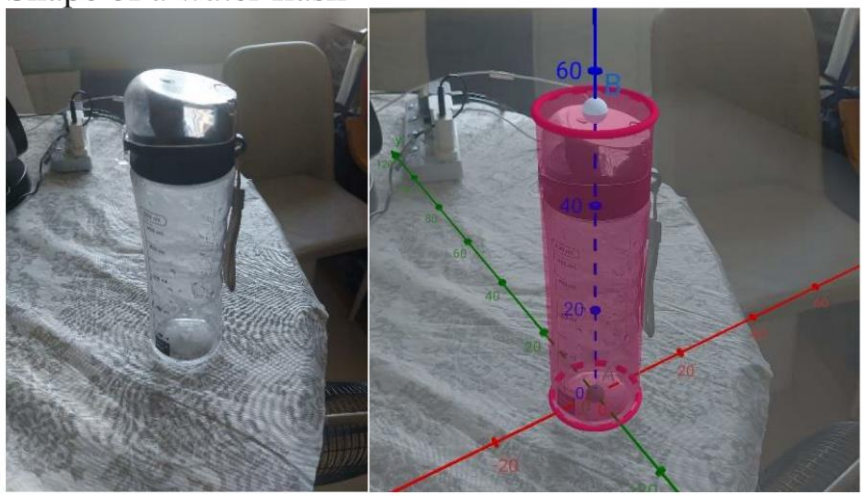




Shape of a rice bowl



Shape of a water flask



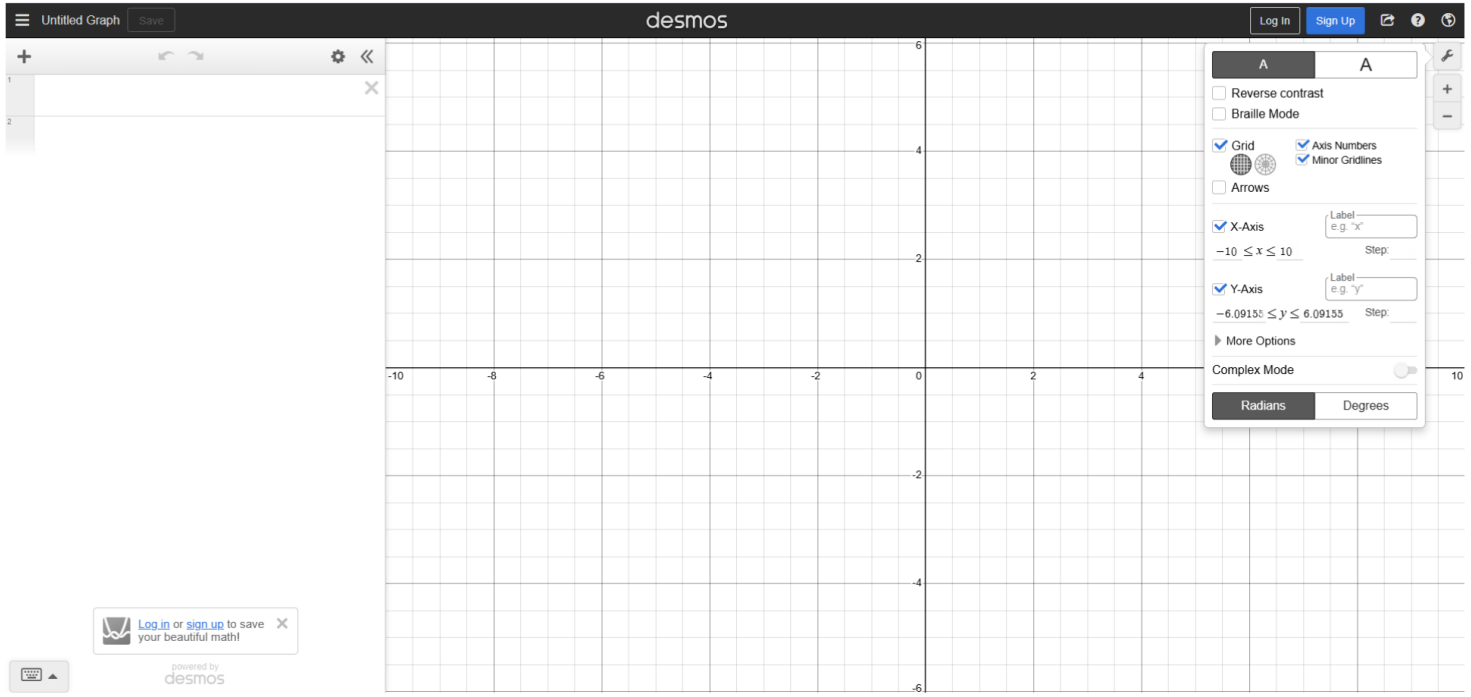
Desmos for math modelling 用於數學建模

- **Desmos** (https://www.desmos.com/calculator)

- Interface 介面:
  - Left: equation bars for inputting equations and expressions.
    左邊：用於輸入方程式和表達式的方程式輸入框。

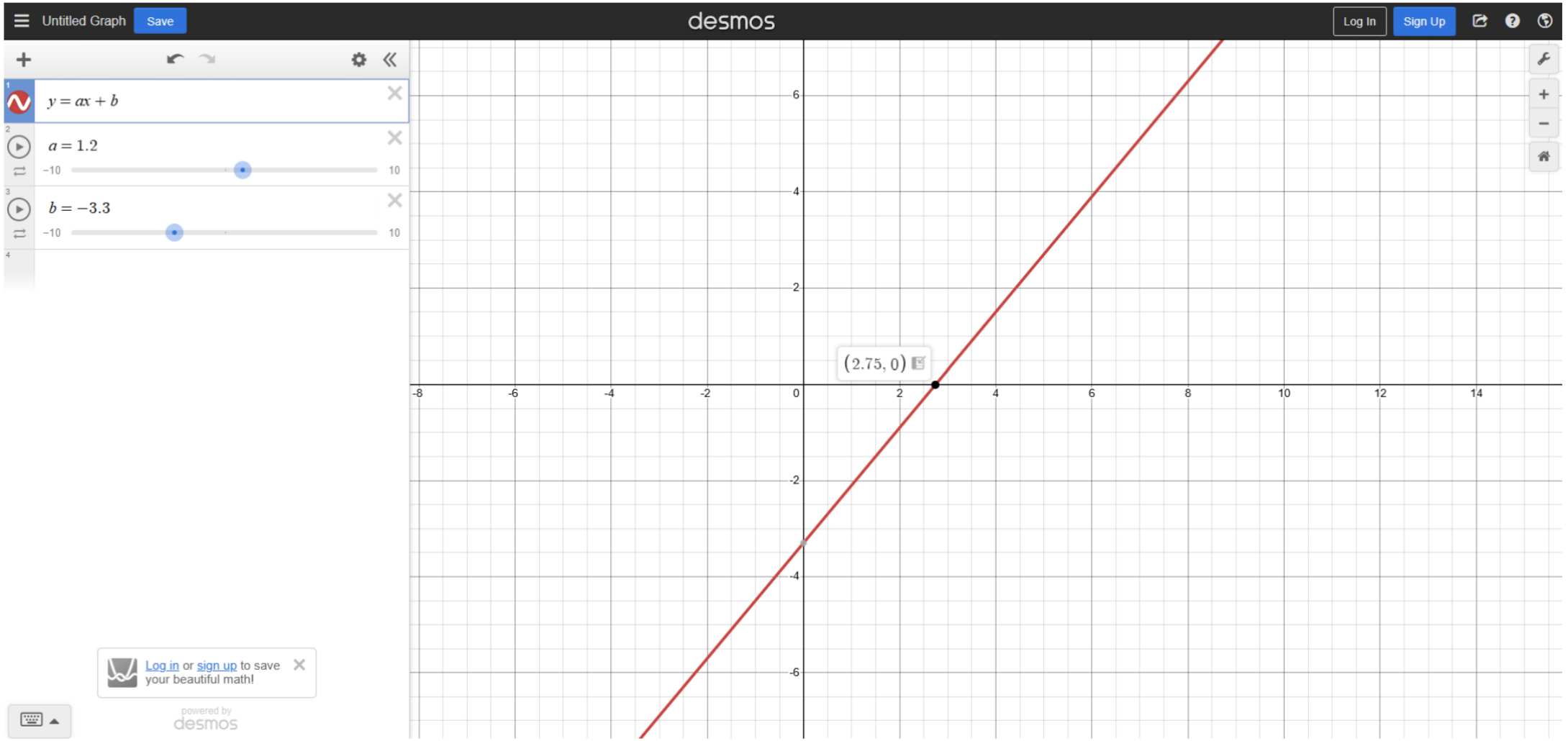
  - Middle: the graph window that visualizes the equation.
    中間：用於可視化方程式的圖形視窗。

  - Top-right-hand corner: the toolbox for different options of the grid.
    右上角：用於設定網格不同選項的工具箱。



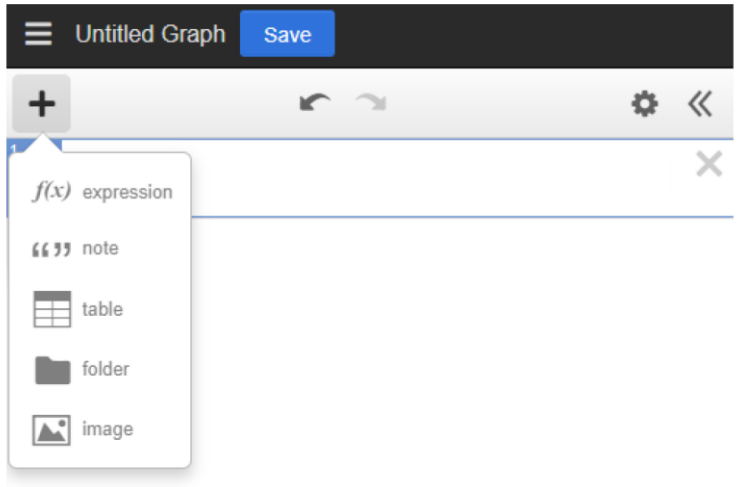
Desmos for math modelling 用於數學建模

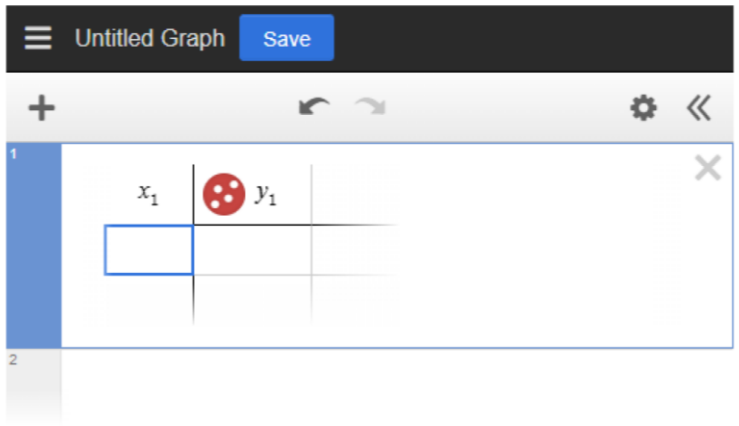
Adding curves and points 增加曲線和點

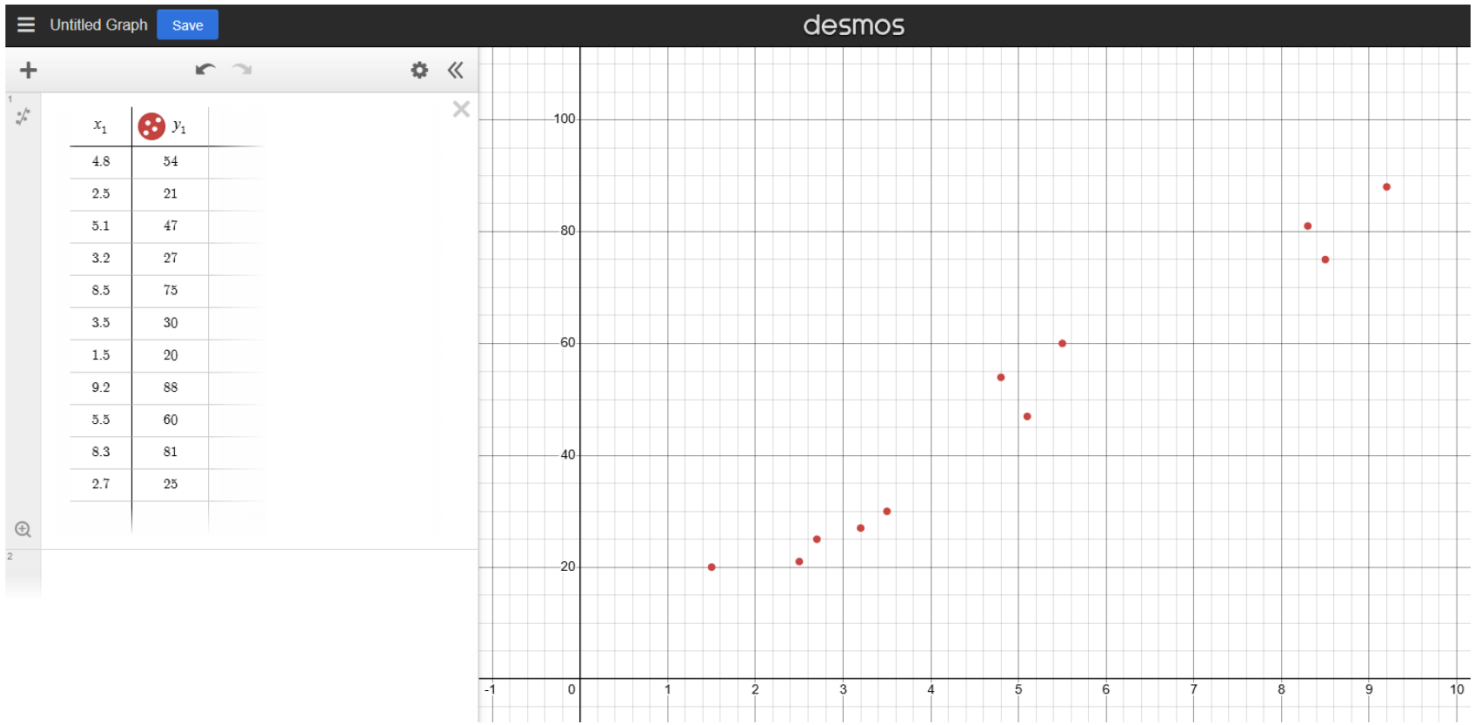


Desmos for math modelling 用於數學建模

Adding data points 新增數據點

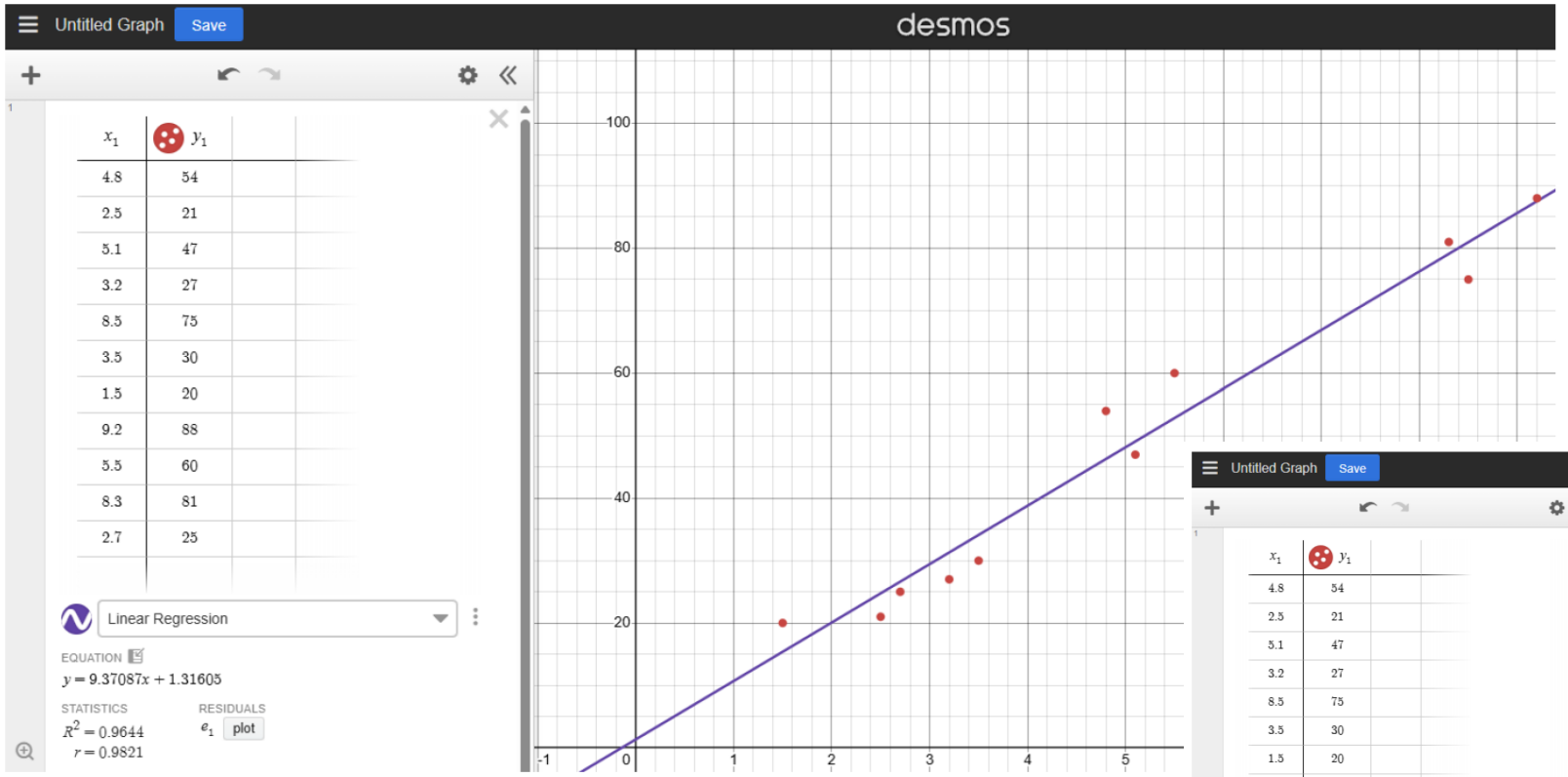


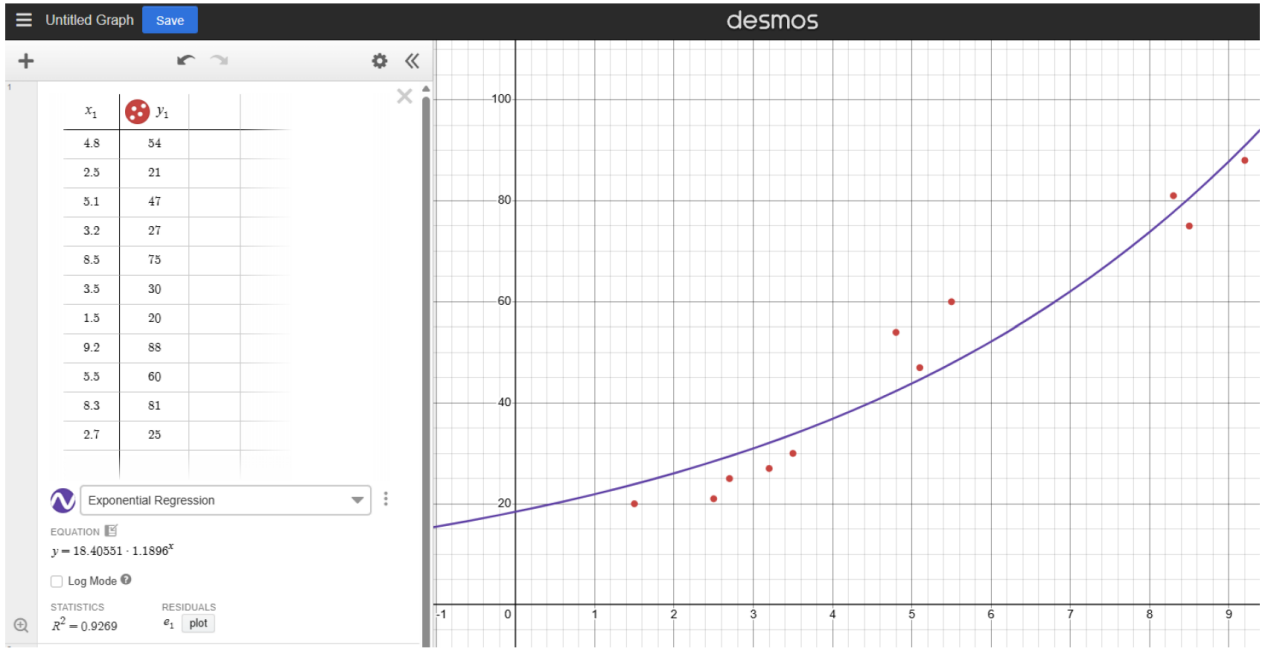




Desmos for math modelling 用於數學建模

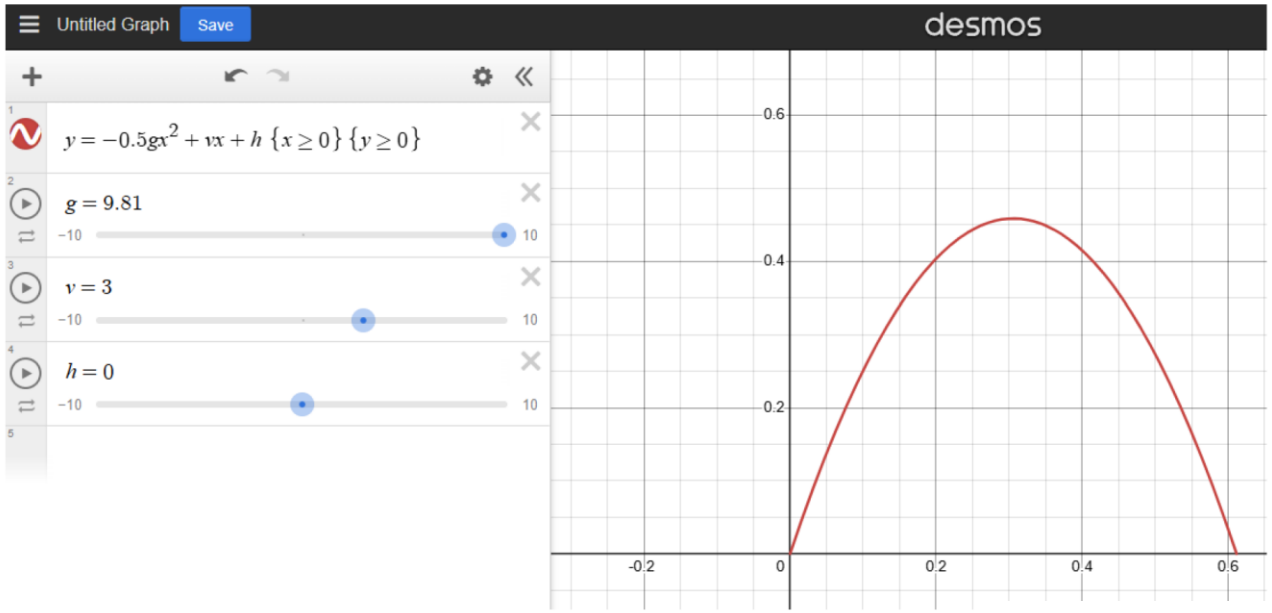
Choosing different regression functions 選擇不同的迴歸函數：



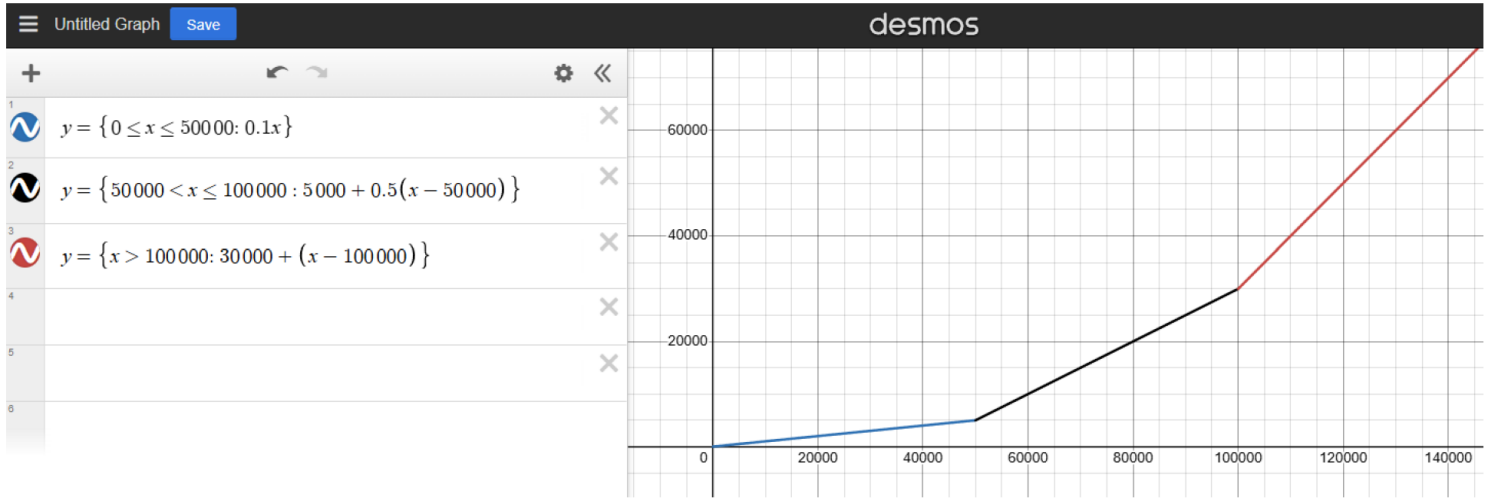


Desmos for math modelling 用於數學建模

Adding restrictions on functions 對函數添加限制



Piecewise functions 分段函數



Other programming tools for math modelling: Python
其他數學建模的程式設計工具：Python

- A very popular computer language for science and engineering
  一種非常流行的科學與工程電腦語言
- Highly customizable 高度可自訂
- Different packages and functionalities for math modelling 提供用於數學建模的各種軟體包和功能
  - Regression 迴歸分析
  - Probability 概率分析
  - Data input/output 數據輸入/輸出
  - Visualization 視覺化
  - Optimization 最佳化
- Freely accessible via online or offline compliers 可透過線上或離線編譯器免費使用
  - Google Colab https://colab.research.google.com/
  - Anaconda https://www.anaconda.com/

Python notebook examples (exercises) 範例（練習）

- **Basic operation of matrices 矩陣的基本運算**
  - https://colab.research.google.com/github/CUHKMathModel/Python/blob/main/Basic_Operation_of_Matrices.ipynb

- **Ordinary least squares 一般最小平方法**
  - https://colab.research.google.com/github/CUHKMathModel/Python/blob/main/Ordinary_Least_Square.ipynb

- **Basic statistics and statistical distributions 基礎統計和統計分佈**
  - https://colab.research.google.com/github/CUHKMathModel/Python/blob/main/Basic_Statistics_Tools.ipynb
  - https://colab.research.google.com/github/CUHKMathModel/Python/blob/main/Statistical_Distribution.ipynb

- **Dataset handling 數據集處理**
  - https://colab.research.google.com/github/CUHKMathModel/Python/blob/main/Introduction_to_Dataset_Handling_Using_Python.ipynb

Python notebook examples (solutions) 範例（答案）

- **Basic operation of matrices 矩陣的基本運算**
  - https://colab.research.google.com/github/CUHKMathModel/Python/blob/main/Basic_Operation_of_Matrices_(Solution).ipynb

- **Ordinary least squares 一般最小平方法**
  - https://colab.research.google.com/github/CUHKMathModel/Python/blob/main/Ordinary_Least_Square_(Solution).ipynb

- **Basic statistics and statistical distributions 基礎統計和統計分佈**
  - https://colab.research.google.com/github/CUHKMathModel/Python/blob/main/Basic_Statistics_Tools_(Solutions).ipynb
  - https://colab.research.google.com/github/CUHKMathModel/Python/blob/main/Statistical_Distribution_(Solution).ipynb

- **Dataset handling 數據集處理**
  - https://colab.research.google.com/github/CUHKMathModel/Python/blob/main/Introduction_to_Dataset_Handling_Using_Python_(Solution).ipynb

Other programming tools for math modelling: R
其他用於數學建模的程式設計工具：R

- **R** (https://www.r-project.org/)
- Useful for statistical analysis 用於統計分析
- A large variety of built-in functions available 提供豐富的內建函數
- Many examples in our **Math modelling e-book** contain detailed R codes:
  我們的**數學建模電子書**中包含許多範例的詳細 R 程式碼：
  - https://www.math.cuhk.edu.hk/~mathcal/MM/
  - Username: mathmodel
  - Password: mm@2024

How to Write a Mathematical Modelling Report?

如何撰寫數學建模報告？

How to Write a Mathematical Modelling Report? 如何撰寫數學建模報告？

- **Introduction 介紹**
  - Describe the real-world problem background 描述現實生活問題背景
  - Review the relevant prior works 以往的相關工作
  - Identify the current research gap 找出現有的研究空隙

- **Proposed mathematical models 提出的數學模型**
  - Introduce the mathematical tools involved 介紹所涉及的數學工具
  - Describe the proposed models in detail with reasoning and justification 詳細描述所提出的模型並給出推理和論證
  - Present the experimental results 展示實驗結果
  - Describe the subsequent model refinements with reasoning and justification 描述後續模型改進並給出理由和論證

- **Conclusion 結論**
  - Summarize the findings 總結研究結果
  - Discuss the advantages and limitations 討論優點和局限性
  - Discuss possible future directions 討論未來可能的方向

Example: Analyzing and Predicting Infectious Diseases in Hong Kong
例子：香港傳染病分析及預測

- Past MMCSS competition problem (2024/25 Junior) 以往MMCSS比賽題目（2024/25 初中組）

Analysing and Predicting Infectious Diseases in Hong Kong

Nowadays, there are many outbreaks of infectious diseases.

Established in 2004, the Centre for Health Protection (CHP) contributed to the development of the capacity of Hong Kong's public health system to deal with various important public health challenges. Through establishing a disease surveillance system, strengthening infection control, enhancing laboratory diagnostic capacity, conducting risk communication and health promotion, developing applied research and training programmes, and preparing emergency response plans, CHP seeks to prevent communicable and non-communicable diseases with the following commitment (the 3 'P's): "Protect the health of the community", "Promote healthy living" and "Partner with stakeholders".

Beside preventing communicable and non-communicable diseases, CHP also plays an important role in recording various notifiable infectious diseases, conducting risk communication and health promotion.

Consider the website of CHP (https://www.chp.gov.hk/en/static/24012.html), or other appropriate information, to complete the following tasks:

- Identify two diseases that you suggest to deserve special attention in Hong Kong in the coming future. Provide your justification.
- Among the two diseases identified above, select **ONE** to develop mathematical model(s) to predict the total number of cases in Hong Kong in 2026 and 2036 under different scenarios:
  (i) normal scenario;
  (ii) high-outbreak scenario; and
  (iii) low-outbreak scenario.

State the data you have collected clearly. Your data must be accurate, with sources cited, and your argument must be logical and sound. State clearly all the assumption(s) you need in your modelling process.

香港傳染病分析及預測

在現今社會，爆發傳染病的情況時有發生。

衛生防護中心（CHP）於 2004 年成立，旨在提升香港公共衞生系統，以應付公共衞生的挑戰。透過建立疾病監測網絡、加強感染控制、提升化驗診斷能力、進行風險通報及健康促進活動、發展應用研究及培訓計劃，以及擬定緊急應變計劃，致力預防傳染病和非傳染病，以實踐保障市民的健康、推廣健康生活及與相關各方建立伙伴關係的承諾。

除致力預防傳染病和非傳染病，衞生防護中心在記錄各類須呈報傳染病、進行風險通報及健康促進活動方面亦擔當重要角色。

請參考衞生防護中心網頁(https://www.chp.gov.hk/tc/static/24012.html)或其他合適的資料，以完成以下任務：

- 提出兩種你認為在不久將來需要在香港特別關注的傳染病，並提供理由。
- 對於這兩種傳染病的**其中一種**，建立數學模型以預測其於不同情況下在 2026 年和 2036 年的香港病例總數：
  (i) 一般情況；
  (ii) 高爆發情況；
  (iii) 低爆發情況。

請列出你所收集的資料。資料要準確並列出來源，論證亦要合乎邏輯。建模過程中所設立的所有假設均需清晰列出。

Example: Analyzing and Predicting Infectious Diseases in Hong Kong
例子：香港傳染病分析及預測

Techniques that may be useful:
可能有用的技巧：

- Regression迴歸分析
  - How to predict the trend? 如何預測趨勢？
  - What are the variables? 變數有哪些？
  - Polynomial/exponential/periodic/…
    多項式/指數/週期/…
  - Short term 短期
  - Long term 長期

- Probability and statistics 概率與統計
  - How to model high-outbreak and low-outbreak scenarios?
    如何模擬高爆發和低爆發情況？

香港傳染病分析及預測

在現今社會，爆發傳染病的情況時有發生。

衛生防護中心（CHP）於 2004 年成立，旨在提升香港公共衞生系統，以應付公共衞生的挑戰。透過建立疾病監測網絡、加強感染控制、提升化驗診斷能力、進行風險通報及健康促進活動、發展應用研究及培訓計劃，以及擬定緊急應變計劃，致力預防傳染病和非傳染病，以實踐保障市民的健康、推廣健康生活及與相關各方建立伙伴關係的承諾。

除致力預防傳染病和非傳染病，衞生防護中心在記錄各類須呈報傳染病、進行風險通報及健康促進活動方面亦擔當重要角色。

請參考衞生防護中心網頁(https://www.chp.gov.hk/tc/static/24012.html)或其他合適的資料，以完成以下任務：

- 提出兩種你認為在不久將來需要在香港特別關注的傳染病，並提供理由。
- 對於這兩種傳染病的**其中一種**，建立數學模型以預測其於不同情況下在 2026 年和 2036 年的香港病例總數：
  (i) 一般情況；
  (ii) 高爆發情況；
  (iii) 低爆發情況。

請列出你所收集的資料。資料要準確並列出來源，論證亦要合乎邏輯。建模過程中所設立的所有假設均需清晰列出。

Example: Analyzing Tourism in Hong Kong
例子：香港旅遊業分析

- Past MMCSS competition problem (2024/25 Senior) 以往MMCSS比賽題目（2024/25高中組）

Analysing Tourism in Hong Kong

Hong Kong is known as the "Pearl of the East" and is one of the most popular tourist destinations in the world. Understanding and analysing the trends of visitors to Hong Kong is of great help to the development of Hong Kong's tourism industry and even the overall economy of Hong Kong.

Complete the following tasks:

1. Design mathematical model(s) to predict the annual number of visitors to Hong Kong in the next five years (2026 to 2030).
2. How should Hong Kong allocate resources in different industries to promote our tourism industry? Please justify your suggestion(s) with mathematical model(s).

State the data you have collected clearly. Your data must be accurate, with sources cited, and your argument must be logical and sound. State clearly all the assumption(s) you need in your modelling process.

香港旅遊業分析

香港被譽為「東方之珠」，是全球最受歡迎的旅遊勝地之一。了解和分析訪港旅客的趨勢對香港旅遊業發展以至於香港整體經濟有極大幫助。

請完成以下任務：

1. 設計數學模型以預測未來五年（2026 年至 2030 年）的每年訪港旅客人次。
2. 香港應如何在不同行業投放資源以促進香港旅遊業？請以數學模型支持你的建議。

請列出你所收集的資料。資料要準確並列出來源，論證亦要合乎邏輯。建模過程中所設立的所有假設均需清晰列出。

Example: Analyzing Tourism in Hong Kong
例子：香港旅遊業分析

Analysis 分析:

- How to predict the trend?
  如何預測旅遊趨勢？
- Land/sea/air? 陸路/海路/空路？
- What are the variables affecting tourist number?
  哪些因素會影響遊客人數？
- What is the relationship between different industries and tourism?
  不同產業與旅遊業有何關聯？

香港旅遊業分析

香港被譽為「東方之珠」，是全球最受歡迎的旅遊勝地之一。了解和分析訪港旅客的趨勢對香港旅遊業發展以至於香港整體經濟有極大幫助。

請完成以下任務：

1. 設計數學模型以預測未來五年（2026 年至 2030 年）的每年訪港旅客人次。

2. 香港應如何在不同行業投放資源以促進香港旅遊業？請以數學模型支持你的建議。

請列出你所收集的資料。資料要準確並列出來源，論證亦要合乎邏輯。建模過程中所設立的所有假設均需清晰列出。

Sample report (1) – 1-page summary
參考作品 (1) – 1頁摘要

一、摘要

本文通過建立時間序列數學模型，預測香港未來五年（2026-2030 年）的訪客人數及通過多元線性回歸模型解釋香港在不同行業如何投放資源以促進旅遊業，最終優化香港資源配置以促進旅遊業發展。

針對問題一，本文收集歷史訪客人數[1]，以及相關外生變量數據[2][3][4][5]，分割了疫情前（2013-2019）、疫情中（2020-2022）、恢復期（2023-2025）三段歷史時期。考慮到數據有明顯的季節性趨勢與時間序列變化，若直接使用綫性建模，不考慮非綫性關係和外生變量，預測結果將不理想。故本文基於 2013-2024 年月度訪港遊客資料，採用 STL 分解法分析季節性模式，得出季節性強度 0.864，並構建季節性回歸積分滑動平均外生變量模[SARIMAX(2,1,1)(1,1,1,12)]模型進行預測。預測顯示香港旅遊業將於 2027 年完全恢覆至疫情前水準（6,300 萬人次/年），2030 年達 6,680 萬人次。進行滾動預測驗證後，模型驗證顯示平均絕對百分比誤差（MAPE）為 6.50%。

針對問題二，本文旨在構建量化模型，識別影響香港訪客人數的關鍵因素，並提出基於數據驅動的資源優化分配策略。收集 2015-2024 年香港旅遊業數據[6][7]（剔除疫情期間 2020-2022 年數據）後，採用多元線性回歸模型分析演唱會場次（文娛活動）、酒店房價（商業價值）和酒店入住率（服務業質量）等因素對訪客人數的影響，並通過彈性系數量化各因素的影響程度。彈性系數分析顯示商業價值（-0.87）、文娛活動（2.5）和服務業質量（0.79）是影響訪客人數的三大關鍵因素。優化模型解釋力達 99.0%（R^2=0.990），平均預測誤差為 3.59%。建議香港將 35%資源用於商業價值價格補貼（包括服務行業補貼 20%、旅遊套餐優惠 10%、交通費用補貼 5%），45%資源用於文化活動支持（包括演唱會場地補貼 25%、文化活動推廣 10%、煙花表演支持 10%），20%資源用於提升服務業質量，以最大化優化資源配置。

本研究通過數據分析和數學建模解決兩個關鍵問題：預測 2026-2030 年訪港旅客人次和制定最優資源投放策略。研究建議通過分段建模、實時數據更新和動態季節性調整等方法進一步提升模型精確度，並可將模型推廣至其他旅遊熱點地區及相關行業應用，為旅遊業復甦和長期發展提供科學決策支援。

Sample report (1) – Model assumption
參考作品 (1) – 模型假設

3. 1 基本假設

假設季節性模式在 2026 年至 2030 年將會維持，不會有大幅度改變（SARIMAX 模型）;

假設在經過新冠肺炎疫情之後，香港旅遊業仍然有增長的潛力;

假設香港旅遊業水準恢復至新冠肺炎疫情前水準仍然需要一段時間（SARIMAX 模型）;

假設在新冠肺炎疫情後全面開關訪港遊客快速回升，2026 年至 2030 年間將維持快速恢復的趨勢;

假設滾動驗證的 MAPE 值低於 15%時，模型具有可接受的預測能力;

假設異常值的影響可以通過對數轉換和季節性調整得到有效控制;

香港旅遊業的訪客人數受到多種因素的綜合影響，這些因素之間存在不同程度的相關性;

假設通過對歷史資料的分析，可以量化各因素對訪客人數的影響程度;

資源投入與旅遊業績效之間存在可測量的關係，可以通過模型優化資源配置;

假設多元線性回歸模型擾動項服從獨立的正態分佈。

3. 2 符號說明

表 3.2-1 符號說明

符號	含義	單位
y_t	表示實際遊客人數	百萬
T_t	表示旅客人數長期變化	百萬
S_t	表示旅客人數的週期性變化	百萬
R_t	表示無法被趨勢和季節性解釋的隨機波動	百萬
$y\hat{}_{t+1}$	表示下一個時間點的預測值	百萬
$y\hat{}_t$	表示當前時間點的預測值	百萬
α	表示平滑係數（$0<\alpha<1$）	-
N	表示樣本數量	-
Z	表示置信水準對應的 Z 值	-
σ	表示預測值的標準差	百萬
Y	表示訪客人數（應變量）	遊客
X_1	表示演唱會場次	場次
X_2	表示酒店入住率	百分比（%）
X_3	表示酒店房價	港幣/晚間
β_0	表示回歸方程常數項	-
β_1	表示演唱會場次的回歸係數	-
β_2	表示酒店入住率的回歸係數	-
β_3	表示酒店房價的回歸係數	-
E	彈性係數，表示自變量變動 1%導致因變量變動的百分比	-
E_1	表示演唱會場次的彈性係數	-
E_2	表示酒店入住率的彈性係數	-
E_3	表示酒店房價的彈性係數	-

注：未註明符號以出現處為準

Sample report (1) – Formulating and solving models
參考作品 (1) – 建構及求解模型

STL（Seasonal-Trend decomposition using Loess）分析是時間序列分解的核心步驟，用於識別數據中的季節性、趨勢和平穩成分。 我們將數據時間劃分爲疫情前（2013 年 1 月至 2019 年 5 月）和復蘇期（2023 年 3 月至 2025 年 1 月），對原始訪客數據通過公式（2）進行對數轉換：

$$\text{log\_arrivals} = \text{np.log1p(total arrivals)} \quad (2)$$

以減少異常值影響並使變化率更具可比性。 將數據進一步用 statsmodels 庫的 STL 方法設置關鍵參數為公式（3）和公式（4）：

$$\text{stl} = \text{STL(pre\_covid['log\_arrivals'], period=12)} \quad (3)$$

$$\text{result} = \text{stl.fit()} \quad (4)$$

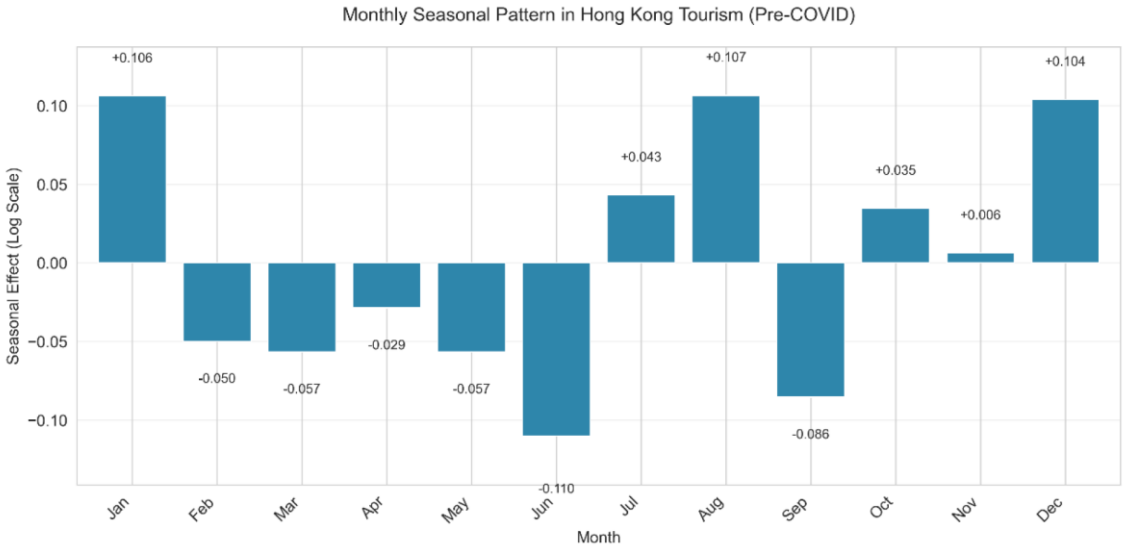


圖 4.1.2-1 疫情前季節性模式圖

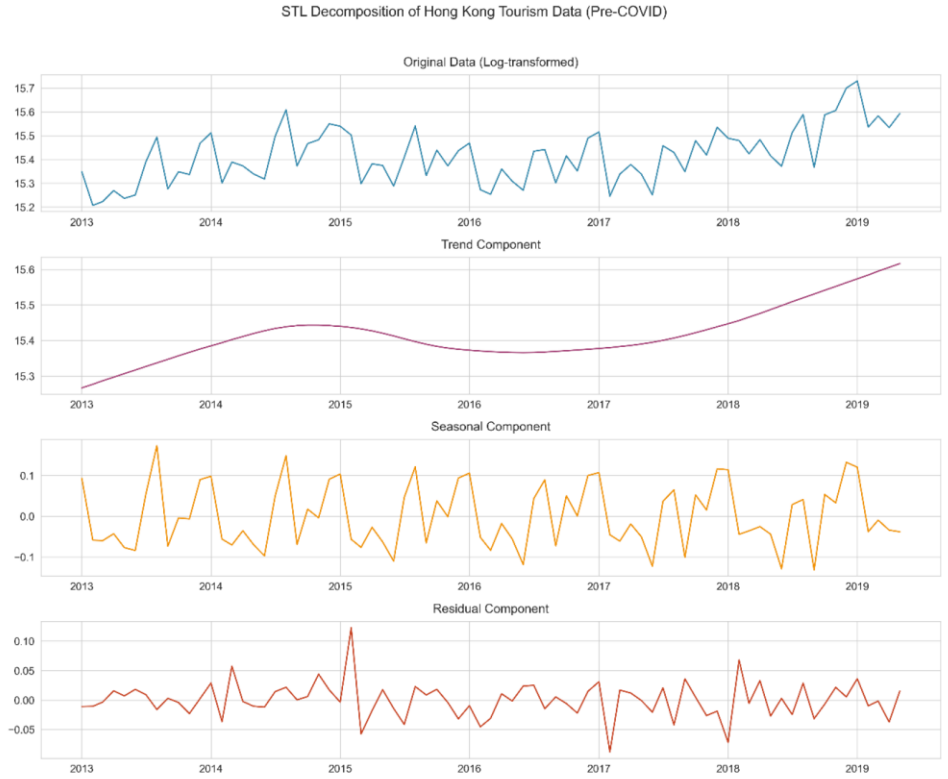
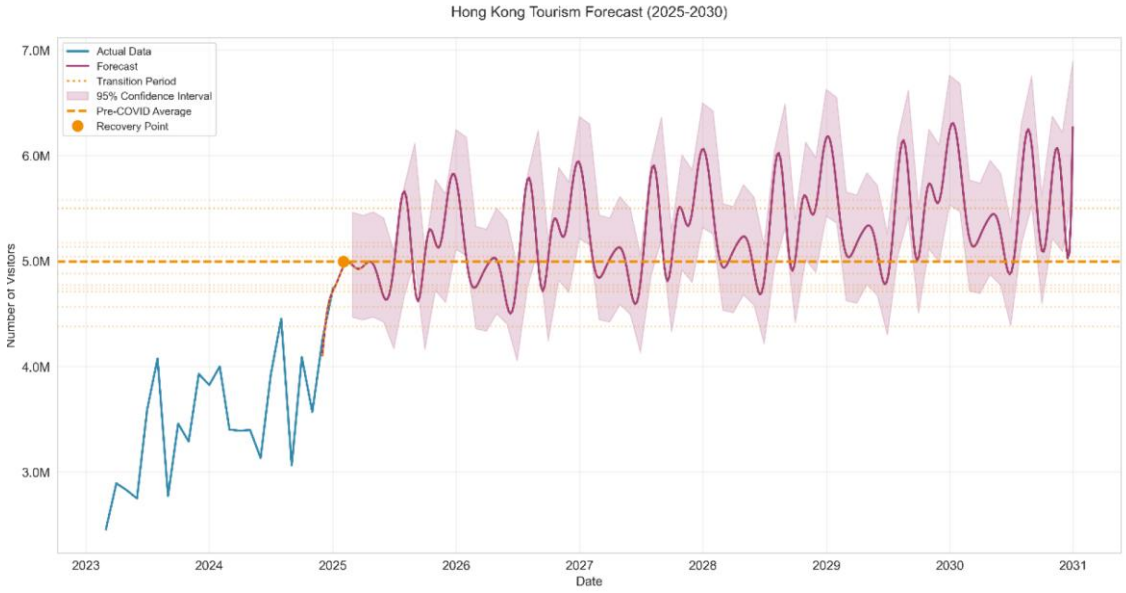
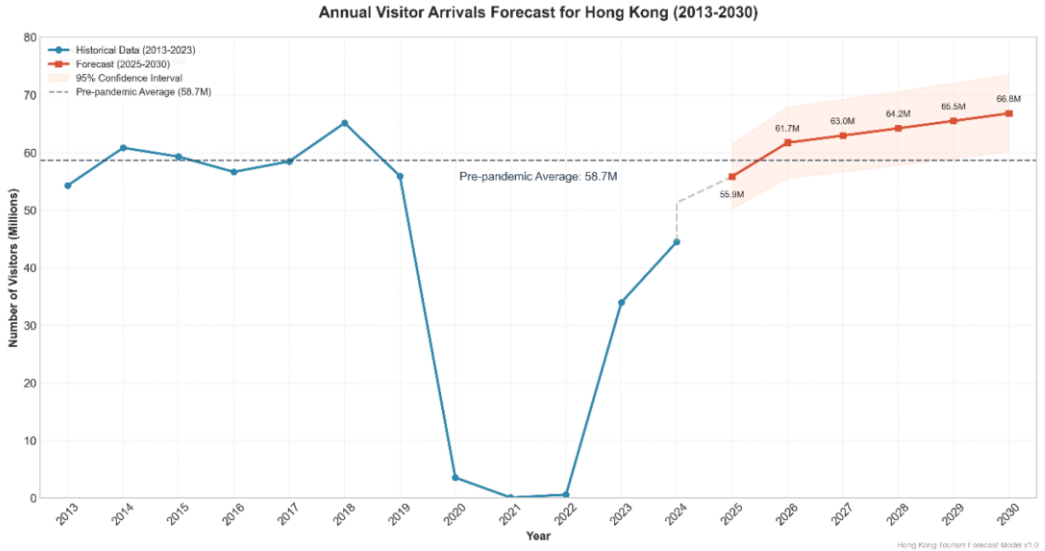


圖 4.1.2-2 疫情前 SLT 分解圖

發現數據季節性強度：0.864，趨勢性強度：0.864。 表明數據的季節性和趨勢性非常顯著，二者影響力相當。 同時呈現明顯的“雙峰”模式：夏季（7-8 月）和冬季（12-1 月）為旺季，且最大季節性波動達到 21.68 個百分點（8 月最高+10.65%，6 月最低-11.03%）。

Sample report (1) – Interpreting and validating models
參考作品 (1) – 解釋及驗證模型





在滾動預測驗證中，該模型的平均絕對百分比誤差（MAPE）為 6.50%，均方根誤差（RMSE）為 411,420.39（單位：百萬遊客），表明瞭該模型對歷史數據的擬合效果較好。

通過使用 STL 方法，我們可以清晰地分離出時間序列中的趨勢、季節性和其他波動，這使我們可以更好地理解遊客人數的變化規律，並爲未來的預測提供堅實的基礎。

在恢復期預測驗證中，MAPE 上升至 21.39%，RMSE 達到 982,670.61（單位：百萬遊客），表明瞭模型在預測疫情後恢復期的遊客人數時存在顯著的誤差。

模型主要依賴歷史數據，對 COVID-19 等突發事件的預測能力有限。

模型的季節性成分基於固定的週期假設，可能無法完全反應疫情後旅遊模式的變化

Sample report (1) – Formulating, solving, and interpreting models
參考作品 (1) – 建構、求解及解釋模型

4. 2 任務 2：如何在不同行業投放資源以促進香港旅遊業？

4. 2. 1 模型建構

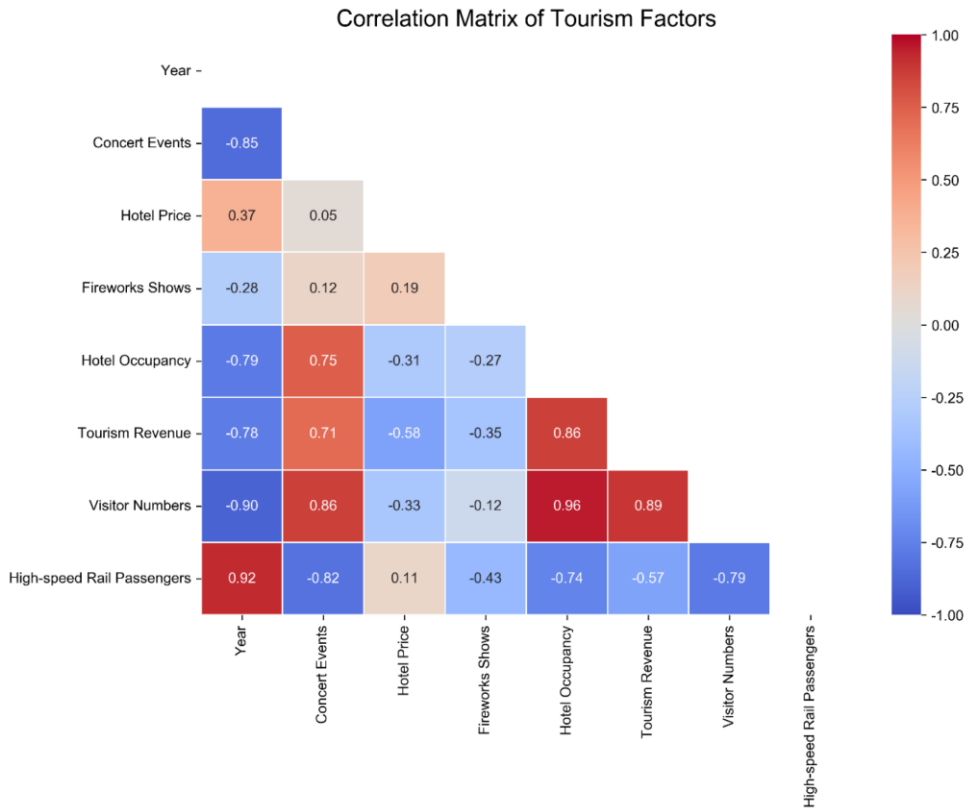
通過調查香港旅遊業年度數據及各因素與訪客人數的相關性我們得到表 4.2.1-1 表 4.2.1-2 和圖 4.2.1-1：

表 4.2.1-1 香港旅遊業年度數據概覽（2015-2024 年，剔除疫情期間）

年份	演唱會場次	酒店房價（港幣）	煙花表演（場）	酒店入住率（%）	旅遊業收入（億港幣）	訪客人數（人）
2016	116	1287.08	3	85.92	20.83	56,654,903
2017	99	1286.75	3	86.92	19.00	58,472,157
2018	122	1375.42	2	87.92	21.05	65,147,555
2019	93	1206.00	2	85.92	21.33	55,912,609
2023	84	1395.33	3	82.25	14.83	33,999,660
2024	72	1316.30	2	84.20	17.28	36,678,799

表 4.2.1-2 各因素與訪客人數的相關係數

因素	與訪客人數的相關係數	相關性強度
演唱會場次	0.86	極強正相關
酒店入住率	0.958	極強正相關
煙花表演	-0.124	極弱負正相關
旅遊業收入	0.893	極強正相關
酒店房價	-0.329	弱負相關
高鐵乘客量	-0.789	弱負相關



Sample report (1) – Formulating, solving, and interpreting models
參考作品 (1) – 建構、求解及解釋模型

可知旅遊業發展受到多種外部因素影響，對此問題我們打算考慮上圖 6-圖 7 中的所有變量，使用多元綫性回歸方法構建模型，其有公式（6）：

$$訪客人數 = \beta_0 + \beta_1 \times 演唱會場次 + \beta_2 \times 酒店房價 + \beta_3 \times 煙花表演 + \beta_4 \times 酒店入住率 + \beta_5 \times 高鐵乘客量 + \varepsilon \quad (6)$$

又基於相關性分析結果，我們選擇了相關性絕對值大於 0.6 的變量構建優化模型，其有公式（7）：

$$訪客人數 = \beta_0 + \beta_1 \times 演唱會場次 + \beta_2 \times 酒店入住率 + \beta_3 \times 酒店房價 + \varepsilon \quad (7)$$

該優化模型得到如下評估結果：

R^2: 0.990

調整後的 R^2： 0.981

F 統計量： 101.67

F 統計量的 p 值： 0.0016

均方誤差（MSE）： 3159332151780.70

平均絕對誤差（MAE）： 1389721.44

平均絕對百分比誤差（MAPE）： 3.59%

通過優化模型我們得到表 4.2.1-3：

表 4.1.2-3 優化模型的係數及統計顯著性

變數	係數	標準誤差	t 值	p 值	統計顯著性
常數項	45,239,351.71	1,030,000	44.084	0.000	極顯著
演唱會場次	13,126,584.99	1,310,000	10.038	0.002	極顯著
酒店入住率	4,752,518.81	1,160,000	4.111	0.026	顯著
酒店房價	-3,400,101.09	1,220,000	-2.782	0.069	邊緣顯著

表 4.2.1-4 彈性係數分析（各因素對訪客人數的影響程度）

因素	彈性係數	影響解釋
酒店房價減少	-0.79	入住率 1%，訪客人數增加約 2.5%
演唱會場次	2.5	演唱會場次增加 1%，訪客人數增加約 2.5%
酒店入住率	0.79	入住率提高 1%，訪客人數增加約 0.79%

這表明：

酒店房價對訪客人數有極強的負向影響：房價下降 1%，訪客人數增加約 100%

演唱會場次對訪客人數有較強的正向影響：演唱會場次增加 1%，訪客人數增加約 25%

酒店入住率對訪客人數有一定的正向影響：入住率提高 1%，訪客人數增加約 9%

Sample report (1) – Validating models
參考作品 (1) – 驗證模型

4. 2. 2 模型預測與驗證

通過將模型應用於歷史數據和未來預測，我們可以評估模型的預測能力。預測結果不僅能驗證模型的準確性，還能幫助我們理解模型捕捉到的關鍵趨勢和模式。 下模型在不同年份的預測表現如表 4.2.2-1 所示：

表 4.2.2-1 模型預測結果與實際值比較

年份	實際訪客人數	預測訪客人數	預測誤差	預測誤差百分比
2015	9,809,779	10,581,290	-771,510	-7.86%
2016	56,654,903	60,484,250	-3,829,346	-6.76%
2017	58,472,157	56,514,220	1,957,940	3.35%
2018	65,147,555	64,389,110	758,450	1.16%
2019	55,912,609	55,028,540	884,072	1.58%
2023	33,999,660	32,736,100	1,263,564	3.72%
2024	36,678,799	36,941,970	-263,169	-0.72%

從預測結果可分析出：

模型在 2015-2019 年的歷史數據上表現穩定，預測誤差百分比基本控制在±8%以內;

對 2023-2024 年的預測也顯示出較好的準確性，預測誤差百分比均在±4%以內

模型成功捕捉到了訪客人數的整體趨勢變化，包括 2019 年後的下降趨勢以及近期的逐步恢復

表明該模型不僅能夠解釋歷史數據中的變化規律，還具有較好的預測能力，可以為未來的政策制定和資源分配提供可靠的參考依據。

在對模型進行進一步評估後，我們得到表 4.2.2-2：

表 4.2.2-2 模型評估指標

評估指標	數值	解釋
R^2	0.990	模型解釋了 99.0%的訪客人數變異
調整後的 R^2	0.981	考慮變數數量後的解釋力
F 統計量	101.67	模型整體顯著性檢驗
F 統計量的 p 值	0.0016	模型整體極為顯著
均方誤差（MSE）	3,159,332,151,780.70	預測誤差的平方平均值
平均絕對誤差（MAE）	1,389,721.44	預測誤差的絕對值平均
平均絕對百分比誤（MAPE）	3.59%	預測誤差占實際值的百分比

該模型的平均絕對百分比誤差（MAPE）為 3.59%，表明模型具有較高的預測準確性。 所有年份的預測誤差百分比均在±8%範圍內，進一步證實了模型的可靠性。

Sample report (2) – 1-page summary
參考作品 (2) – 1頁摘要

Abstract

Hong Kong has long been regarded as one of the most vibrant cities in the world, with a fast-growing economy, a unique position due to its close ties to China, and various attractive sites. Known by many names—Pearl of the East, Shopping and Food Paradise, and East-meets-West Centre—Hong Kong is one of the top tourist destinations globally, attracting millions of visitors annually. As one of the four major pillars of Hong Kong's economy, tourism plays a critical role in further boosting the city's reputation and soft power internationally. Therefore, in this paper, we aim to investigate the trends of visitors to Hong Kong by developing a mathematical model to predict the number of visitors in the coming five years (2026-2030) and exploring ways to further elevate Hong Kong's position as a global tourist attraction.

In Task 1, we first identify the different factors affecting the number of tourists in Hong Kong and project individual factors to achieve the goal of predicting the future five-year trend. A total of eight significant factors are taken into account, such as the consumer price index, crime rate, and purchasing power parity. For each factor, we have made appropriate regression analyses and ultimately combined all factors to successfully predict the future number of tourists in Hong Kong over the next five years.

In Task 2, we have divided it into two subsections: allocating resources in the short term and long term. There are five main sectors in which the government can invest to boost tourism: shopping, hotels, dining, entertainment, and sightseeing. We analyzed the visiting trend of tourists from different origins throughout the year by separating them into two different groups and calculating their Popularity Index, then modeled their consumption patterns respectively using the Importance Indices for each sector. We found out found that shopping accounts for the largest share of Hong Kong's tourism. Therefore, it is recommended that the Hong Kong SAR government invest the most in the shopping industry. However, it is also important to consider the elasticity ratio in the long term; the more money invested, the smaller the increase in the effectiveness of the investment.

Sample report (2) – Problem introduction and analysis
參考作品 (2) – 題目簡介及分析

1.1 Background

Tourism is one of the traditional four pillars of Hong Kong's economy. In 2018, it accounted for 4.5% of Hong Kong's GDP and provided about 256 900 jobs, representing 7% of total employment in Hong Kong (HKTB, 2025). Up till mid-2019, the number of visitor arrivals has been rising consistently. However, as the global economy and tourism suffered from the COVID-19 cataclysm, the tourism industry in Hong Kong inevitably took a huge downturn. Fortunately, as what we can observe from the statistics, with the inherit advantage of Hong Kong as an international city, immense support from the Central Government and the unwavering determination of the Hong Kong people, one will not be accused to be over-optimistic for saying that tourism is on the right track of recovering in the year 2025. Even so, adequate government policy is still imperative for the visitor arrivals to bounce back to the pre-COVID era as soon as possible. Under such situation, mathematic modeling can prove useful by simulating real life conditions using numbers to help us with deeper understanding of the problem and help with decision makings.

1.2 Problem Restatement

The tasks 1 and 2 are restated as follows:

1. By creating a mathematical model, predict the trend of the annual number of tourists to Hong Kong in the coming five years, from 2026 to 2030.

2. Determine the method of allocating different resources to various sectors to maximize the effectiveness of Hong Kong's tourism promotion and provide suggestions to further boost Hong Kong's tourism given the scarcity of resources.

Definition of Important Terms

Terms	Definition
Tourists	Non-resident visitors who travel to a place, particularly Hong Kong, for leisure, business, or other purposes
Major events	Events that caused, or will cause, the number of tourists to significantly deviate from the predicted trend
Factors	Determinants of the number of tourists visiting Hong Kong
Industries	Sectors under the tourism industry which can affect the desire of tourists to visit Hong Kong
Resources	The amount of money or supply of other assets that can be drawn from the Hong Kong SAR government for investment, particularly, in the tourism industry

Sample report (2) – Model assumption
參考作品 (2) – 模型假設

Assumption	Justification
No major events e.g. wars, pandemics or destructive natural disasters that significantly affect the number of tourists will occur in the upcoming 5 years.	The probability of this happening is low, and even if it happens, its duration and effect can be hardly predictable.
The parameters that determine the number of tourists will follow the original trend and are also not altered by major events.	It would be difficult and impractical to predict the occurrence of these major events.
COVID-19, the pandemic, only have a significant impact on Hong Kong's tourism from 2020 to 2023.	While COVID-19 first appeared in late 2019, it was not widespread at that time. Policies restricting tourists were implemented in 2020, and these policies were relaxed and gradually phased out by 2023.
The cost of allocating resources in different categories is similar.	Most of the short-term policies involve promotional works or setting up minor events, whose costs will not differ significantly in different categories.

Definition of Variables

Variable	Definition
t	Time, in terms of years. For the n-th month of year q, the t value would be $q + \frac{n-1}{12}$, e.g. February 2008 is represented by $t = 2008\frac{1}{12}$
T	Number of tourists visiting Hong Kong in a particular month
C	The Composite Consumer Price Index (CPI) of a particular month
P	The Purchasing Power Parity (PPP) conversion rate of Hong Kong dollars of a particular month
E	The exchange rate of Chinese Yuan (CNY) to Hong Kong dollars (HKD) on the first day of a particular month
R	The crime rate of Hong Kong of a particular month, calculated by the number of crimes per 100,000 population
M	The rate of PM2.5 emissions of a particular month, in tonnes per year
V	The proportion of expected tourists in a given month that ended up not coming to Hong Kong due to a major event.
S	The seasonal variation of T in a particular month of the year for a particular group, calculated by dividing the average T for that group of that month by the average T for that group of all months over all years.
Φ	The popularity index of Hong Kong as a tourist destination in a month across different years
I	The importance index of a spending category when allocating resources
B	The budget of the government on the tourism sector each month
λ	The cost of investment in any one of the five industries: shopping, hotel bills, meal outside hotels, entertainment, and tours
ϵ	The elasticity ratio, defined to be $\frac{dT}{d\lambda}$, which represents the ratio of change in the number of tourists over the cost

Sample report (2) – Formulating and solving models
參考作品 (2) – 建構及求解模型

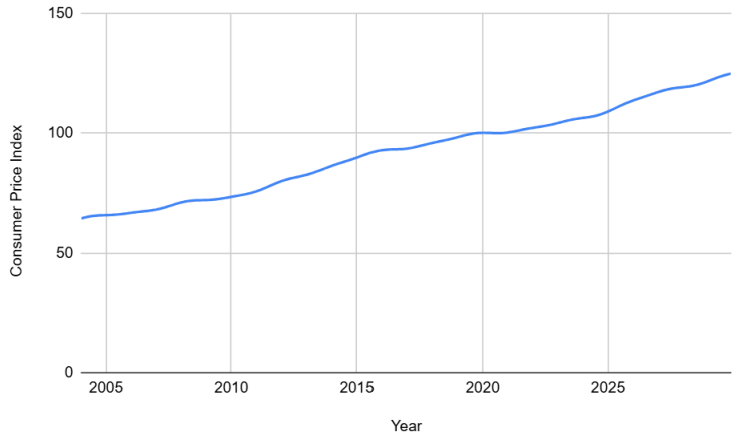
The number of tourists can be affected by many factors, which will be analyzed in this section. We will separate the tourists into two different groups based on their region of origin. The first group consists of tourists from Mainland China, Macao SAR, and Taiwan (Group 1), while the second group contains tourists from the rest of the world, such as Europe and Southeast Asia (Group 2).

We have outlined 7 important factors, listed as follows, and they will be addressed one by one:

1. The cost of travelling, measured by the **Composite Consumer Price Index** of Hong Kong (C). A higher value of C is expected to have a negative impact on the number of tourists.
2. The safety of travelling, measured by the **Crime Rate** of Hong Kong (R). A lower R will boost the number of tourists.
3. The competitiveness of Hong Kong, measured by the **Purchasing Power Parity**, quantified by the conversion rate of the Geary-Khamis Dollar to the Hong Kong Dollar (P). A larger value of P implies a lower competitiveness, hence harming the number of tourists.
4. The competitiveness of nearby cities, particularly those in Mainland China, measured by the **exchange rate of CNY to HKD** (E). A lower conversion rate means tourists tend to visit or stay in Mainland China due to its lower cost, thus having a negative impact on the number of tourists visiting Hong Kong. This effect is particularly significant for tourists originating in Mainland China.
5. The environmental quality, measured by the **PM2.5 emissions** (M). When M is lower, it means that the air quality is improved, and more tourists would tend to visit Hong Kong.
6. **Major events** affecting the number of tourists, which will be considered exceptional cases. They may either have a positive or negative effect on the number of tourists.
7. **Seasonal variations.**

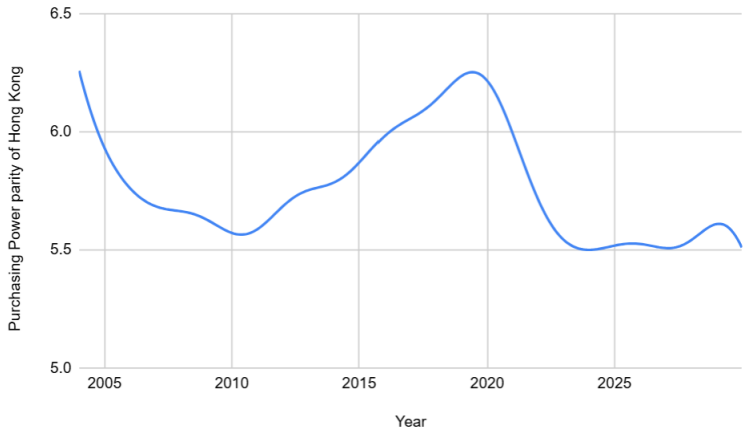
Composite Consumer Price Index

$$C(t) = mt + n + \sum_{i=1}^{4} a_i \cos(b_i t + c_i)$$



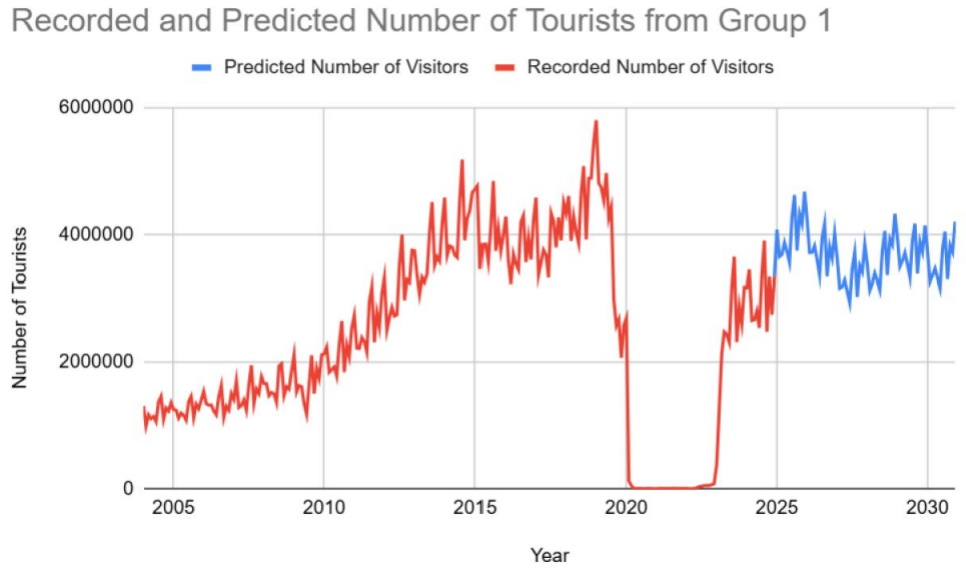
Purchasing Power Parity

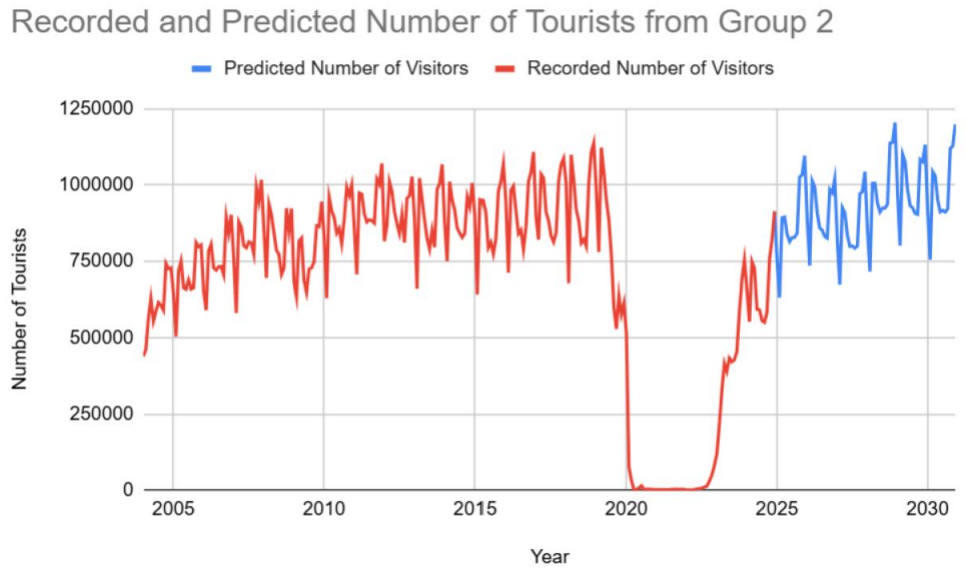
$$P(t) = p + \sum_{i=0}^{4} n_i t^i + \sum_{j=1}^{10} a_j \cos(b_j t + c_j)$$



Sample report (2) – Interpreting and validating models
參考作品 (2) – 解釋及驗證模型

Predicted Total Number of Tourists Visiting Hong Kong						
	Year					
Month	2025	2026	2027	2028	2029	2030
January	4072835	4244549	3625519	3522108	3949773	3732013
February	3649233	3711366	3161704	3128271	3494546	3263185
March	3683930	3716392	3186030	3234951	3583894	3355773
April	3870796	3817637	3285761	3389972	3712463	3462899
May	3728851	3598347	3115533	3268572	3529718	3310417
June	3556625	3358616	2934610	3133896	3328483	3160844
July	4270196	3934069	3459875	3735596	3907048	3764666
August	4620441	4168577	3714387	4053770	4169504	4044096
September	3743709	3339531	3028424	3359429	3382271	3295454
October	4321772	3803053	3508202	3925786	3880396	3827740
November	4166676	3625496	3408927	3834365	3726083	3720130
December	4672623	4018751	3841789	4324104	4140861	4199126



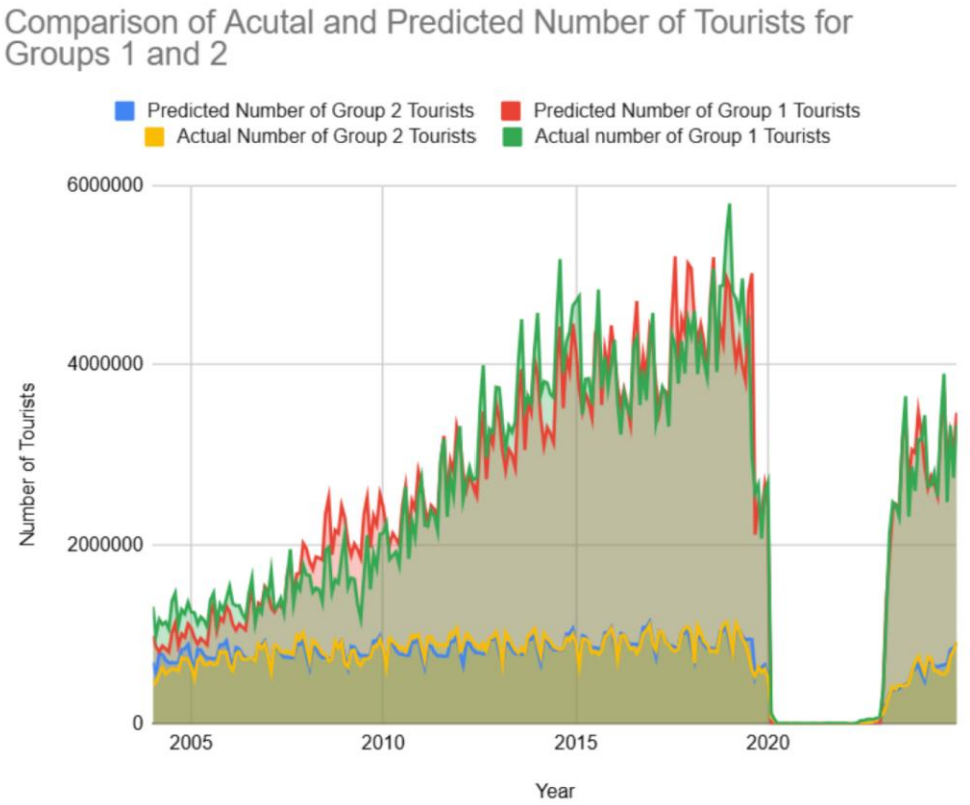


Sample report (2) – Interpreting and validating models
參考作品 (2) – 解釋及驗證模型

From this graph, we may conclude that tourist numbers from group 1, that is from mainland China, Taiwan or Macao SAR are unlikely to return to pre-COVID levels (2019) or grow further. A major contributing factor may be the expected decline in value of the Renmenbi, which incentivize tourists to take advantage of the lower touristic cost and stay in mainland China. On the other hand, tourist numbers from group 2 are expected to gradually return to pre-COVID levels.

Since a larger proportion of total tourist numbers in Hong Kong are contributed by Group 1, we propose that the total number of tourists may stagnate for the next 5 years. It may not be able to reach pre-COVID levels despite the easing off of the pandemic's effects.

Finally, we may evaluate the accuracy of the regression through this graph comparing T_{real} and T_{pred} from 2004 to 2024.



We see that T_{real} lines up well with T_{pred} for both groups. T_{pred} is particularly able to take into account the seasonal variations of T_{real} as well as the rate of recovery of T_{real} from the COVID-19 pandemic, which will have the greatest impact on tourist counts as Hong Kong continues to recover from it in the coming years. This suggests that the model will have a generally high accuracy for the prediction period.

Sample report (2) – Formulating and solving models
參考作品 (2) – 建構及求解模型

4 Task 2 (Part 1): Allocating Resources in the Short Term

4.1 Background

There are many factors to concern when making a resource allocation, especially on a complicated topic such as tourism. By examining the *Visitor Profile Report* and *Tourism Expenditure Associated to Inbound Tourism* published by HKTB over the past years, we observe that not only do visitors spend very differently on different industries, such consumption pattern varies among visitors from different origins and develop distinctively over the years. To ensure our method of allocation generates the maximal benefits, our model adopts a **divide and conquer** approach by separating the visitor demographic into two major separate groups based on their origins and allocate limited resources among the groups and within each group to each industry.

4.2 Short-term Popularity Fluctuations

To improve the allocation of resources in the short term, we must consider the number of tourists from a microscopic perspective i.e. the monthly trends. From Task 1, we noticed that the trends of Group 1 (Mainland China, Macau SAR and Taiwan) and Group 2 (other countries or regions) may differ. For example, by simply observing the number of visitors in each month from Group 1 and Group 2 respectively, we can see that Group 1 tourists tend to visit Hong Kong in August and January, whereas Group 2 visitors are more likely to visit Hong Kong in May/April and December. To further investigate this, we consider the Popularity Index (Φ) of each month, which can be calculated by the formula below:

$$\Phi(\tau) = \frac{10^{-4}}{w} \sum_{n=2004}^{2024} [T(n+\tau) \times \log_{10}(n-2003)] \tag{5}$$

Month	Φ_1	Φ_2
January	180.4	42.6
February	171.3	37.0
March	145.9	46.9
April	158.9	47.3
May	152.3	42.0
June	147.8	41.3
July	177.4	39.0
August	189.2	32.0
September	147.9	34.5
October	164.8	46.3
November	158.4	48.7
December	177.5	50.2

Sample report (2) – Formulating and solving models
參考作品 (2) – 建構及求解模型

4.3 Visitors Consumption Behavior Paterns

In order to do this, we take a look at the tourists' spending in different sectors over time over the past 10 years. Note that data from 2020 to 2022 are unavailable, which is likely due to the impact of the pandemic.

The table below shows the spending patterns of Group 1 tourists, demonstrated by their percentage of spending in each of the categories:

Spending Category	2015	2016	2017	2018	2019	2023	2024
Shopping, S (%)	68.8	65.0	60.2	60.9	57.1	48.8	43.7
Hotel Bills, H (%)	11.8	13.3	16.7	15.3	16.1	22.1	22.6
Meal Outside Hotels, M (%)	11.2	12.3	13.4	13.1	14.7	20.2	21.5
Entertainment, E (%)	2.5	2.8	3.3	3.5	3.7	3.2	5.7
Tours, T (%)	0.2	0.1	0.1	0.1	0.1	0.1	0.0
Others, O (%)	5.6	6.4	7.2	7.1	8.3	5.5	6.5

We can also make a similar table for Group 2 tourists:

Spending Category	2015	2016	2017	2018	2019	2023	2024
Shopping, S (%)	19.9	19.2	18.7	17.6	18.1	27.1	27.5
Hotel Bills, H (%)	45.7	44.8	43.7	44.8	43.4	39.5	37.1
Meal Outside Hotels, M (%)	18.6	19.4	20.0	29.3	20.0	10.9	21.5
Entertainment, E (%)	5.0	5.6	5.9	6.5	6.2	3.4	3.9
Tours, T (%)	0.8	0.8	1.0	0.7	0.6	0.5	0.7
Others, O (%)	10.0	10.2	10.6	11.1	11.7	8.7	9.3

we can calculate the Importance Index of a category, defined using the formula below:

$$I(cat) = \sum_{i=f(2015)}^{f(2023)} [(p_{cat,f(i+1)} + 3.5(p_{cat,f(i+1)} - p_{cat,f(i)}))(\log_{10}(i-2013))]$$

Category	I_1	I_2
Shopping, S	146.24	109.78
Hotel Bills, H	96.59	136.57
Meal Outside Hotels, M	91.28	87.99
Entertainment, E	24.75	15.08
Tours, T	-0.17	2.53
Others, O	27.00	36.39

Sample report (2) – Model refinement
參考作品 (2) – 改進模型

5 Task 2 (Part 2): Allocating Resources in the Long Term

5.1 Elasticity

We allocate Resources in the long term using the principle of the **Law of Diminishing Marginal Returns**, as the amount of resources provided to a certain investment, the increase in benefits gradually decreases. Similarly, the more money is invested to an industry to promote tourism, the increase in the effectiveness becomes less in the long run over time. It is therefore ineffective to pour all money in just one sector even if the sector is the most significant one.

In regard of this, we will introduce a new concept of Elasticity, ϵ. We define ϵ as the rate of change in the number of tourists over the cost, i.e. how sensitive the number of tourists is to changes in cost. There are five major sectors that the government can invest into for promoting tourism: Shopping (S), Hotel Bills (H), Meal Outside Hotels (M), Entertainment (E) and Tours (T). Let λ be the cost of each sector. The equation is as follows:

$$\epsilon = \frac{dT}{d\lambda} \tag{14}$$

Note that the ratio decreases over time, indicating the effect of cost changes on the number of tourists diminishes.

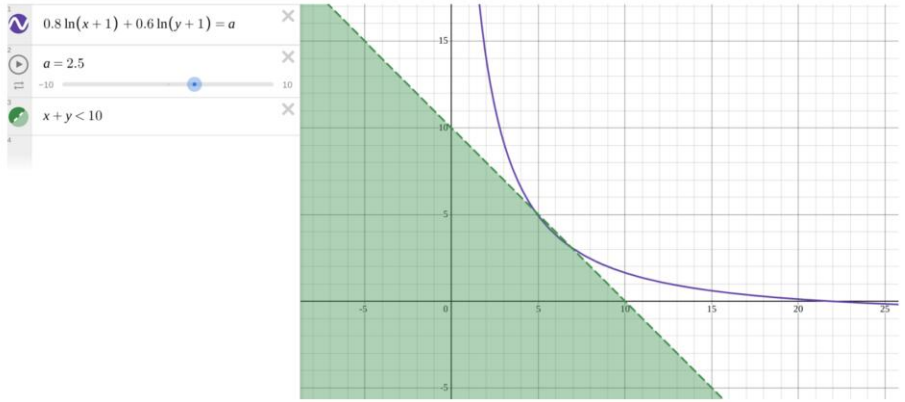
One way to model such effect is using this logarithmic equation:

$$T = 0.8\ln(\lambda + 1) \tag{15}$$

Differentiating the equation from both sides with respect to λ, we have,

$$\epsilon = \frac{dT}{d\lambda} = \frac{0.8}{\lambda + 1} \tag{16}$$

We can have a primary understanding of the government's strategy under such condition using a micro-scale experiment. Let's assume there are two sectors for the government to invest in. By using non-linear programming, we have come up with the following graph:



The green area represents the constraints, while a is T in arbitrary unit. x and y are resources allocated to each industry (also in arbitrary unit) with x being more effective than y. Despite so, it is not optimal to invest all the resources on x.

Example: Potential for agricultural development
例子： 農業發展潛力

- Past HSMMC competition problem (2024/25) (Part of the competition question)
  往屆 HSMMC 比賽題目（2024/25）（題目其中一部份）

> 請建立數學模型來選出一個中國在農業方面最具發展潛力的地區，並解釋你的觀點。
>
> Please develop mathematical model(s) to select a region in China with the greatest potential for agricultural development and explain your point of view.

- Importance of different factors不同因素的重要性
- Data collection 數據收集
- Quantify the development 量化發展

Example: MathWorks Math Modelling Challenge
例子：MathWorks 數學建模挑戰比賽

- https://m3challenge.siam.org

- A math modelling competition for high school students in USA/UK
  一個美國/英國高中生的數學建模競賽

- Past reports 歷屆作品：
  https://m3challenge.siam.org/resources/sample-problems/
  - Winning solutions for different years 歷屆得獎作品
  - Other sample papers for different years 歷屆其他參考作品
    - Average Sample Paper 平均水平作品
    - Above Average Sample Paper 高於平均水平作品
    - Excellent Paper優秀作品
  - Judge commentary for different years 歷年評審評語

More advanced-level mathematical modelling competition examples
更多進階數學建模比賽例子

- **The International Mathematical Modeling Challenge國際數學建模挑戰賽(IMMC)**
  https://immchallenge.org/Index.html

- **Mathematical Contest in Modeling/Interdisciplinary Contest in Modeling**
  美國數學建模競賽/交叉學科建模競賽 (MCM/ICM)
  https://www.comap.com/contests/mcm-icm

- **Contemporary Undergraduate Mathematical Contest in Modeling**
  全國大學生數學建模競賽 (CUMCM)
  https://en.mcm.edu.cn/

- **Sample report sharing 報告例子分享**
  https://www.math.cuhk.edu.hk/app/mathmodel/workshopmaterial2526/workshopmaterial.html
  - Username: mathmodel
  - Password: mm@2024

Discussion: CUMCM 2009 Problem B

2009 高教社杯全国大学生数学建模竞赛

联合赛区评奖结果

全国大学生数学建模竞赛组委会

2009 年 11 月 16 日

香港（本科组）

参赛学校	参赛队员	指导教师	获奖等级
香港城市大学	焦述铭、刘怡芳、杨天阳	Y. C. Hon	全国一等奖

澳门（本科组）

参赛学校	参赛队员	指导教师	获奖等级
澳门科技大学	周培媛、李洁明、郤天宇	梁勇等	全国二等奖

2009 高教社杯全国大学生数学建模竞赛题目

B 题　眼科病床的合理安排

医院就医排队是大家都非常熟悉的现象，它以这样或那样的形式出现在我们面前，例如，患者到门诊就诊、到收费处划价、到药房取药、到注射室打针、等待住院等，往往需要排队等待接受某种服务。

我们考虑某医院眼科病床的合理安排的数学建模问题。

该医院眼科门诊每天开放，住院部共有病床 79 张。该医院眼科手术主要分四大类：白内障、视网膜疾病、青光眼和外伤。附录中给出了 2008 年 7 月 13 日至 2008 年 9 月 11 日这段时间里各类病人的情况。

白内障手术较简单，而且没有急症。目前该院是每周一、三做白内障手术，此类病人的术前准备时间只需 1、2 天。做两只眼的病人比做一只眼的要多一些，大约占到 60%。如果要做双眼是周一先做一只，周三再做另一只。

外伤疾病通常属于急症，病床有空时立即安排住院，住院后第二天便会安排手术。

其他眼科疾病比较复杂，有各种不同情况，但大致住院以后 2-3 天内就可以接受手术，主要是术后的观察时间较长。这类疾病手术时间可根据需要安排，一般不安排在周一、周三。由于急症数量较少，建模时这些眼科疾病可不考虑急症。

该医院眼科手术条件比较充分，在考虑病床安排时可不考虑手术条件的限制，但考虑到手术医生的安排问题，通常情况下白内障手术与其他眼科手术（急症除外）不安排在同一天做。当前该住院部对全体非急症病人是按照 FCFS（First come, First serve）规则安排住院，但等待住院病人队列却越来越长，医院方面希望你们能通过数学建模来帮助解决该住院部的病床合理安排问题，以提高对医院资源的有效利用。

问题一：试分析确定合理的评价指标体系，用以评价该问题的病床安排模型的优劣。

问题二：试就该住院部当前的情况，建立合理的病床安排模型，以根据已知的第二天拟出院病人数来确定第二天应该安排哪些病人住院。并对你们的模型利用问题一中的指标体系作出评价。

Discussion: CUMCM 2009 Problem B

Contemporary Undergraduate Mathematical Contest in Modeling
全國大學生數學建模競賽 (CUMCM)

CUMCM-2009 Problem B: Assignment of Hospital Beds

The inpatient department of CREGOT Ophthalmology Hospital has seventy-nine hospital beds. Treatment and surgery services for cataract, retinal diseases, glaucoma and ocular trauma are offered. Related Information from 13 July 2008 to 11 September 2008 is given in the appendix.

The cataract surgery is relatively easy, and there is no emergency. According to current schedule, the cataract surgery service is on Monday and Wednesday. Surgery preparation needs one or two days. Patients who have two-eye cataract account for 60% of the whole number of cataract patients. The first eye surgery is on Monday, and the second is on Wednesday for them.

The ocular traumas patients usually need emergency treatment. They are hospitalized as soon as possible. The surgery will be on the following day when they are hospitalized.

For retinal diseases and glaucoma patients, usually, preparation needs two or three days before surgery when they are hospitalized, and their postoperative period will be longer. Generally, this type of surgery is not on Monday or Wednesday.

Considering the workload of oculists, CREGOT does not put the cataract surgery and other surgery (except for an emergency treatment) on the same day. The patients(except for emergency treatment) queue up for beds available under the FCFS(First Come, First Serve) principle, while the queue becomes longer and longer. CREGOT hope you can help him to improve the utilization of their resources with mathematical modeling.

Your tasks are as follows:

Task 1: To developed a reasonable evaluation index system to evaluate the merits and demerits of the beds assignment model.

Task 2: Build a model to determine which patients should be hospitalized based on the number of patients who will be discharged on the following day, and evaluate your model using the index system developed in Task 1.

Report sharing and analysis
報告分享與分析

Mathematical Contest in Modeling/Interdisciplinary Contest in Modeling 美國數學建模競賽/交叉學科建模競賽 (MCM/ICM) 2019

Figure 9: The Tourist Map of *Musée du Louvre* G Floor

2019
MCM/ICM
Summary Sheet

Summary

As the world's largest art museum and a landmark of Paris, *Musée du Louvre* attracts more than 10 millions people a year, which makes it necessary to construct efficient evacuation plans for different emergent circumstances. In this paper, two models are set up in order to simulate the evacuation situation from several connected rooms and calculate the time needed.

In the first model, rooms are simplified as small wall elements and exit elements. All the people inside the room are represented by agents. For each specific individual, he or she experiences both repulsive forces from the walls and other agents and attractive forces from the exits. All those forces together will change the agent's state of motion at any time. Meanwhile, the positions of agents also change ceaselessly, resulting in simultaneous change of forces. Such kind of dynamic system will eventually force all the people escape from the room. The model is implemented with MATLAB & Simulink. Some representative rooms of the museum are chosen as the test region and 500 people are placed inside the rooms with random initial positions. It takes 81 seconds for all people to evacuate from the test region, which coincides with the result of a former experiment conducted by Smith (1995).

After the construction of the first model, AnyLogic is used to simulate the evacuation situation for G floor of the museum. Referring to the guidance map given in the museum's official website, the floor plan was drawn. The tourists were represented with agents (circles) as before, while the diameters of the circles were calculated referring to proportional scale. The movement speed of agents was calculated with the same method. Each room could generate some visitors, and the generation rate is the same as that of the first model. After the agents get to their destination (stairs or exits in our case), they will simply disappear in the program.

Finally, we made some analysis on these two models, where one is microscopic while the other one is macroscopic. Robustness and rationality are enhanced after our modification for the first model. However, there still exist some remaining weakness in our model. In the end, we made some suggestions to the museum's personnel based on the simulation conduced by these two models.

Report sharing and analysis
報告分享與分析

Mathematical Contest in Modeling/Interdisciplinary Contest in Modeling 美國數學建模競賽/交叉學科建模競賽 (MCM/ICM) 2021

(a) New south wales ranchers look on as lightning strikes and a wildfire rages in 2019 (from Victoria file footage)

(b) Rare summer thunderstorm sparks new wildfires in

Problem Chosen	2021 MCM/ICM Summary Sheet	Team Control Number
B		2112262

Wildfire monitoring mechanism based on the optimization model

Summary

For the wildfire that occurred in Victoria, this article designed a complete monitoring plan for Country Fire Authority (CFA). From daily monitoring to large-scale fire handling, this article gave a specific drone monitoring plan.

First, this article analyzes the effective shooting range and communication range of the surveillance and situational awareness drone (SSA) and repeater drone (RR), and determines the optimal flight route to reach the economic optimal solution. Then, this article analyzes the topographic characteristics of Victoria (mountains and sea level). According to different terrain features, this article gives different drones flying heights and arrays.

Next, this article made an Optimal Control Model for the number of drones purchased, which is balanced the relationship between economy and reliability.Through the method of dynamic planning, we calculated the approximate plan of drone procurement quantity through the program. The two solutions deal with daily surveillance and information transmission when a wildfire occurs.

Then, we uses FDDI parameters to characterize the severity of the fire, and uses QR dynamic adjustment to meet different needs in different periods. At the same time, we innovatively applied the step-wise cluster analysis method to the prediction of surface temperature to establish a prediction model.

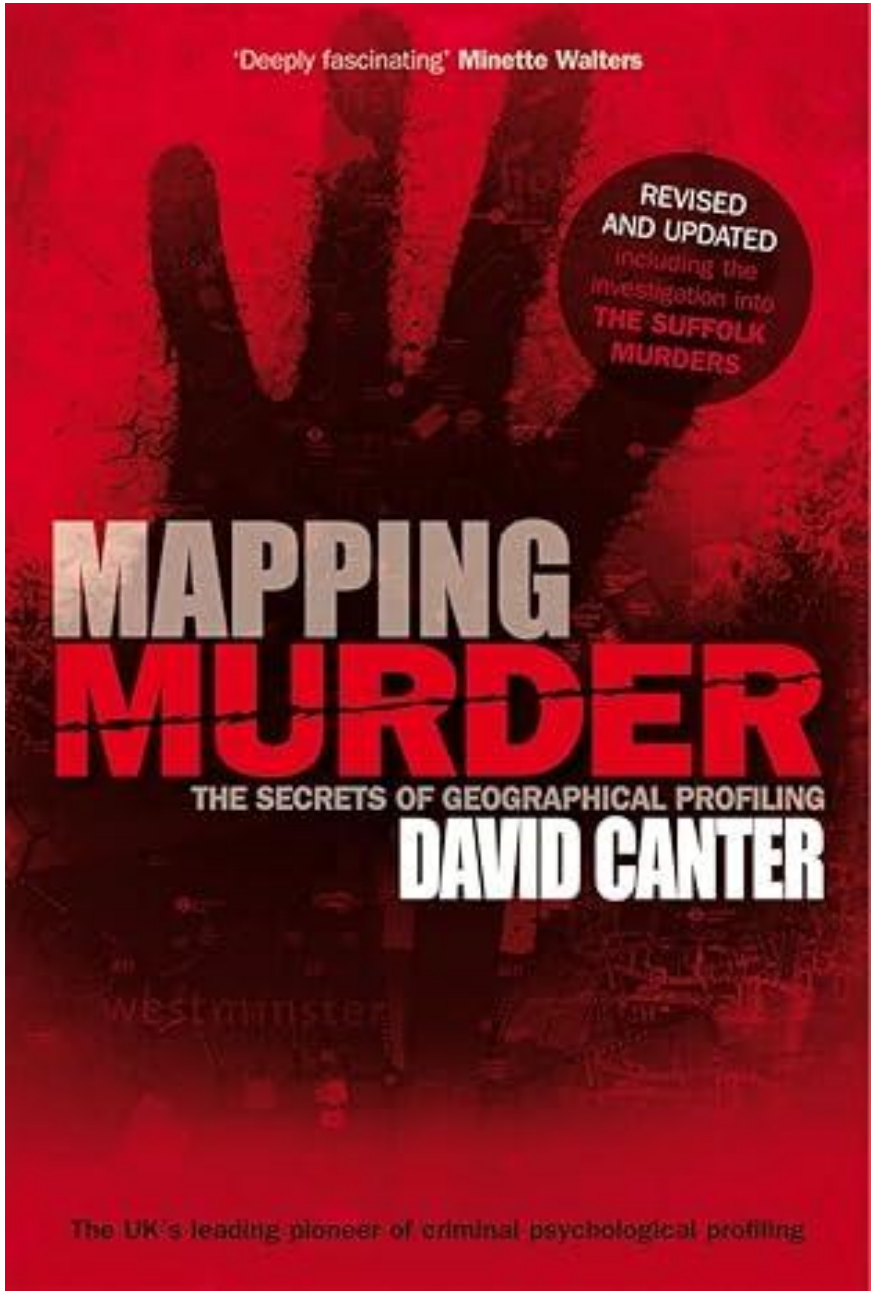
Finally, we optimized the drones distribution queue to adapt to different heights of terrain to ensure the effectiveness of monitoring and the safety of drones.And we gave the detailed budget of the entire system to help CFA lobby the Council to obtain the budget.

Keywords:

Optimal Control Model; Dynamic Adjustment Model; SCA Method;

Report sharing and analysis
報告分享與分析

Mathematical Contest in Modeling/Interdisciplinary Contest in Modeling 美國數學建模競賽/交叉學科建模競賽 (MCM/ICM) 2009



Advanced methods for mathematical modelling

數學建模進階方法

What other advanced mathematical methods may be useful for mathematical modelling?

還有哪些高階數學方法可能對數學建模有用？

- **Probability and statistics 概率與統計：**
  - Markov Chain 馬可夫鏈
  - Statistical modelling 統計建模

- **Optimization 最優化:**
  - Linear programming 線性規劃
  - Constrained and unconstrained optimization 約束優化與非約束優化

- **Other data analysis methods 其他數據分析方法：**
  - Dimensionality reduction降維
  - Clustering聚類
  - Classification分類

Some interesting mathematical model 一些有趣的數學建模

THE DAILY TELEGRAPH Friday, August 8, 2003

Eye-rolling + coldness + mockery = divorce

By Tom Peterkin

NEWLYWEDS can now discover if they are destined for the divorce courts by applying a mathematical model that can predict whether their marriage will succeed.

A mathematician has devised two formulae that he claims have a 94 per cent success rate when it comes to forecasting whether a couple are compatible. The suggestion that the secret of a happy marriage can be found within two complex lines of algebra was made by Prof James Murray of the University of Washington, Seattle.

The model was presented to an international conference for the first time yesterday when Prof Murray addressed academics at Mathematical Biology Conference at Dundee University.

The formulae were calculated during a 10-year study of 700 couples from King County, Seattle, conducted by Prof Murray and his colleague John Gottman, a psychologist.

The experiment, which began in the early 1990s, involved all the couples being observed during a 15-minute conversation when they were just married.

A contentious topic such as sex, child-rearing or money was chosen and the couple's ability to communicate was marked using a scale that gave positive points for good signals and negative points for bad signals.

For example, jokes, a positive tone of voice, smiles and affectionate gestures all resulted in positive scores. Bad signals such as rolling of the eyes, criticism, mocking and coldness led to a negative score.

"We used an accepted psychological scoring system to award them points, such as minus three for scorn and plus two for humour," said Prof Murray, the author of *Mathematics for Marriage*.

"Then we put their points on a graph and by converting them into algebraic terms were able to make our divorce predictions. We didn't tell the volunteers of course, as that could have defeated the object. And telling a couple their marriage is going to fail is not what they want to hear. But the success rate was amazing."

The results were fed into two equations – one for the husband and one for the wife.

The equations were used to calculate the compatibility of the couples by adding a series of other variables such as an "influence function" that differed for each couple.

The "influence function" measured how much someone's contribution to the conversation dictated the mood of his or her spouse.

The couples were tracked every two years and the model accurately predicted which marriages were doomed to failure in a country with a 50 per cent divorce rate.

Prof Murray, an exiled Scot who was been happily married to his wife Sheila for 40 years, said: "I was absolutely astonished. The key thing that comes out of it is that we have been able to calculate how people interact. For example, the wife might be a conflict avoider and the husband might be volatile. That marriage would not survive.

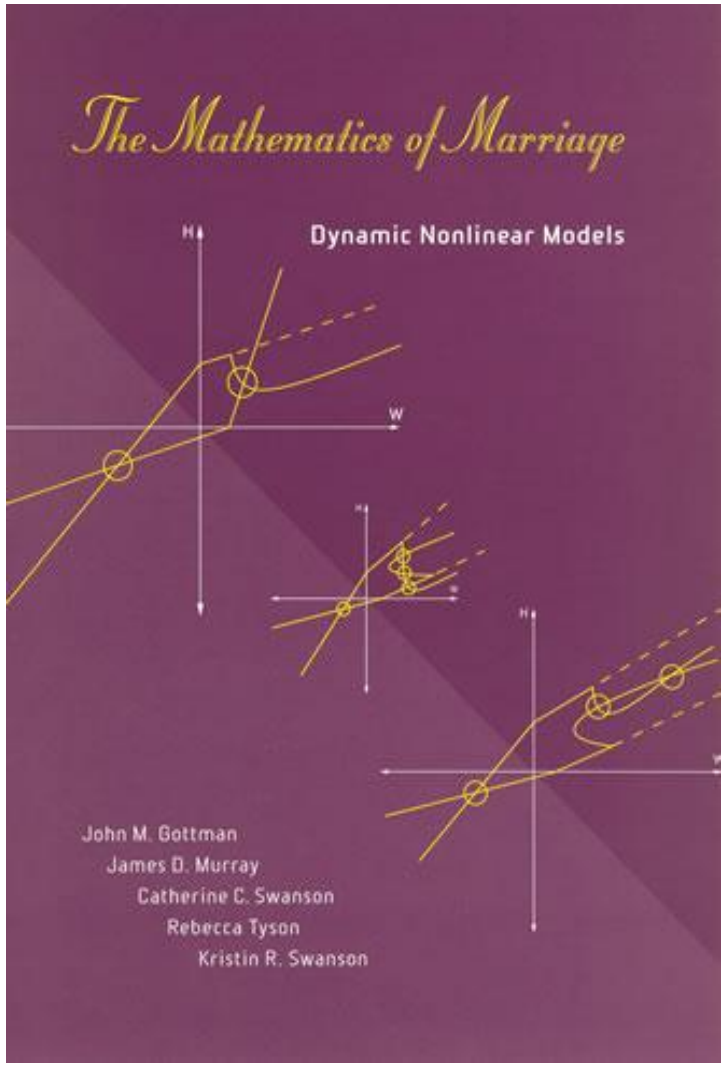
"But positive things can be taken from it. It points out why some people are having problems and can show what action has to be taken to save the marriage."

Editorial Comment: Page 25

反眼
+ 冷漠
+ 嘲笑
= 離婚

Photo: Jupiterimages, Brand X Pictures / Getty Images

Result? Dynamic Nonlinear Models (Chaos)
結果？動態非線性模型（混沌）



Prof. James Murray,
University of Washington (華盛頓大學)

10年對700對夫婦進行研究,
準確率94%

Remark 備註

- There is no perfect model, only better model
  沒有完美的模型，只有更好的模型

- Howard Emmons 霍華德·埃蒙斯:

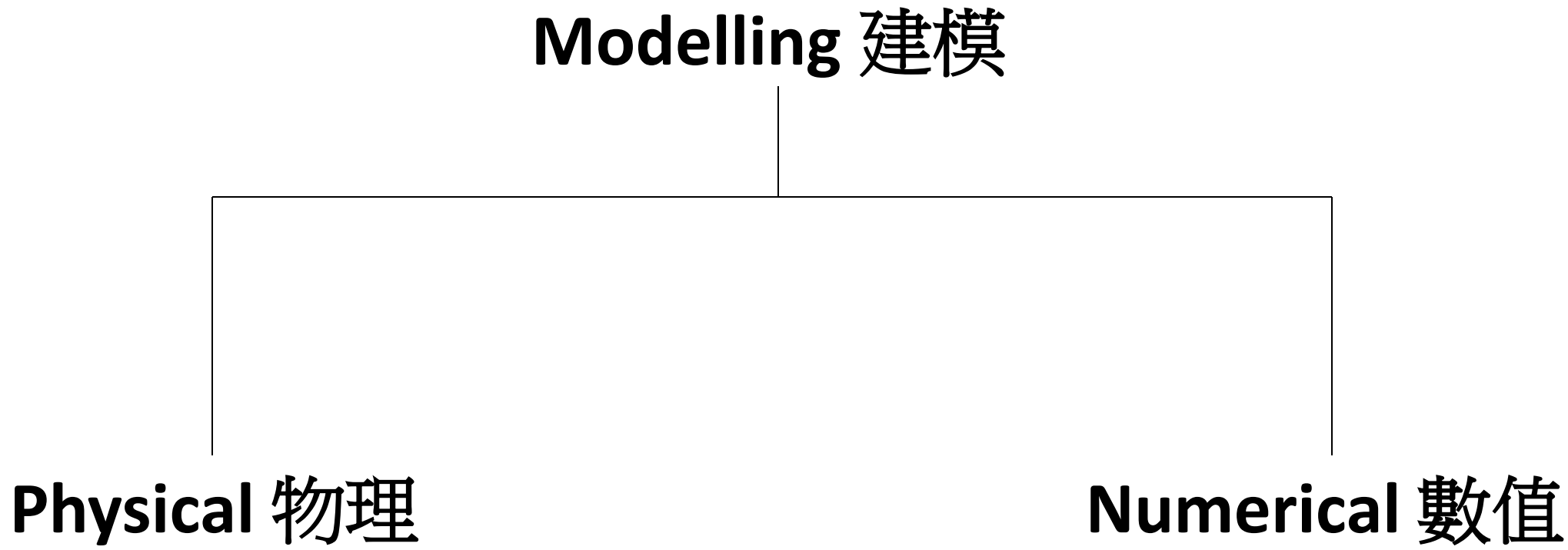
  “Mathematical modelling is not to produce the most comprehensive descriptive model

  but

  to produce the simplest possible model that incorporates the major features of the phenomenon of interest.”

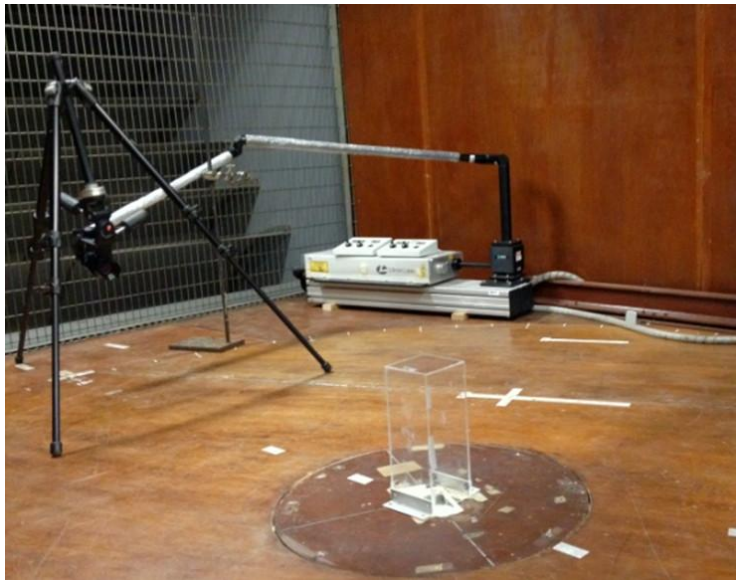
  「數學建模的目的不是建構最全面的描述模型，而是建構能夠涵蓋所關注現象主要特徵的最簡模型。」

Modelling of Real-Life Problems 建模現實生活中的問題



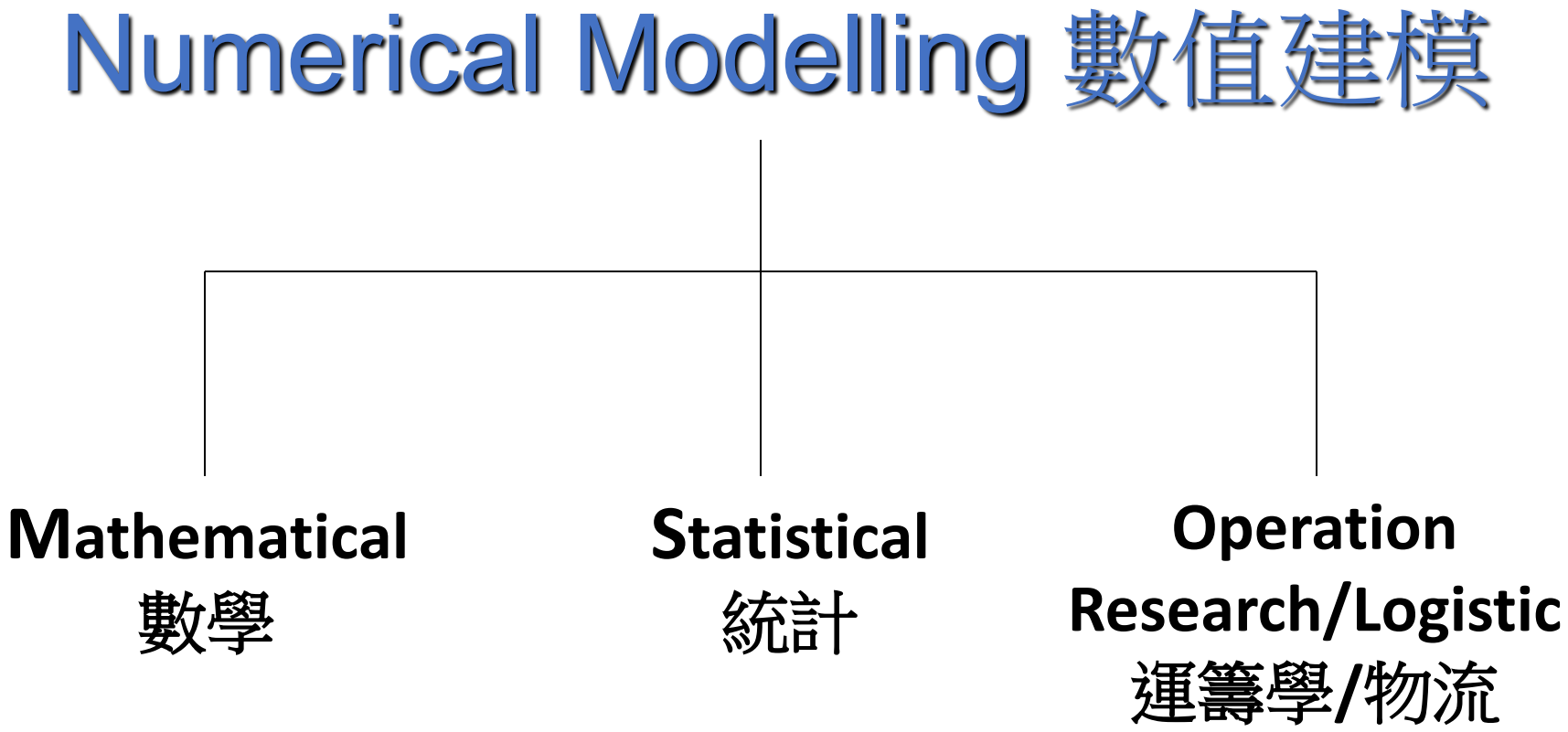
Physical Modelling 物理建模

- Wind/Water Tunnels 風洞/水洞
- Air Tunnel for airplanes 飛機風洞



Wind tunnel installed at the Department of Civil Engineering at HKU
香港大學土木工程系安裝的風洞

BEST and RELIABLE approximations but usually very **expensive**
最佳且可靠的近似值，但通常非常**昂貴**

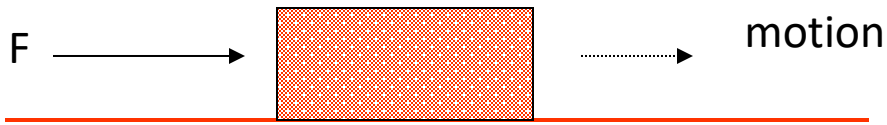

Numerical Modelling 數值建模
Mathematical
數學
Statistical
統計
Operation Research/Logistic
運籌學/物流

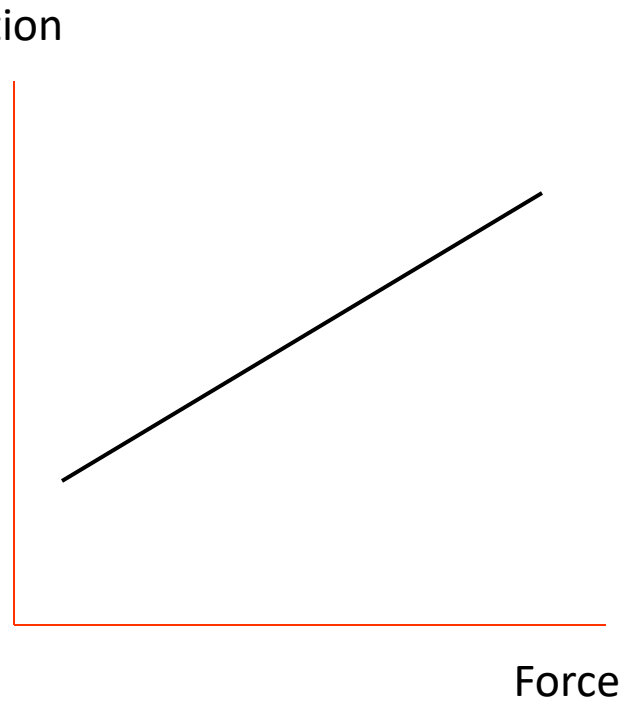
Mathematical Modelling 數學建模

Physical law 物理定律

Force = mass × acceleration

力 = 質量 × 加速度

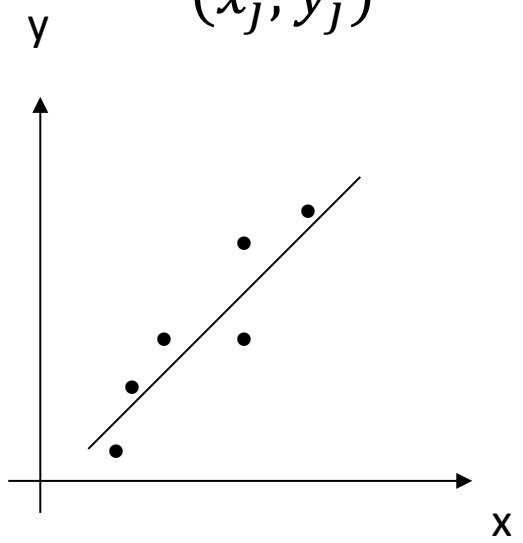




Statistical Modelling 統計建模

Regression Model 迴歸模型

Given data 給定數據 (x_j, y_j)



Find a function $f(x)$ so that $f(x_j)$ best fit y_j
求函數 $f(x)$，使得 $f(x_j)$最能擬合 y_j

Linear Regression Line 線性迴歸線

$$y = \alpha + \beta x$$

- Determine α and β by using method of least square. 使用最小平方法確定 α 和 β

Nonlinear Regression 非線性迴歸

$$\text{e.g. } f(x) = ae^{bx}$$

We have developed several R Shiny tools for regression
我們開發了多個 R Shiny 迴歸工具：
https://www.math.cuhk.edu.hk/app/mathmodel/tool.html

Statistical Modelling 統計建模

Simulation 模擬

Random Process 隨機過程

- Poisson Distribution泊松分佈

  $e^{-\lambda}\ \lambda^x/x!$

- Binomial Distribution 二項分佈

  ${}_nC_x\theta^x\,(1-\theta)^{n-x}$

- Normal Distribution 常態分佈

  $N(\mu,\sigma^2)$

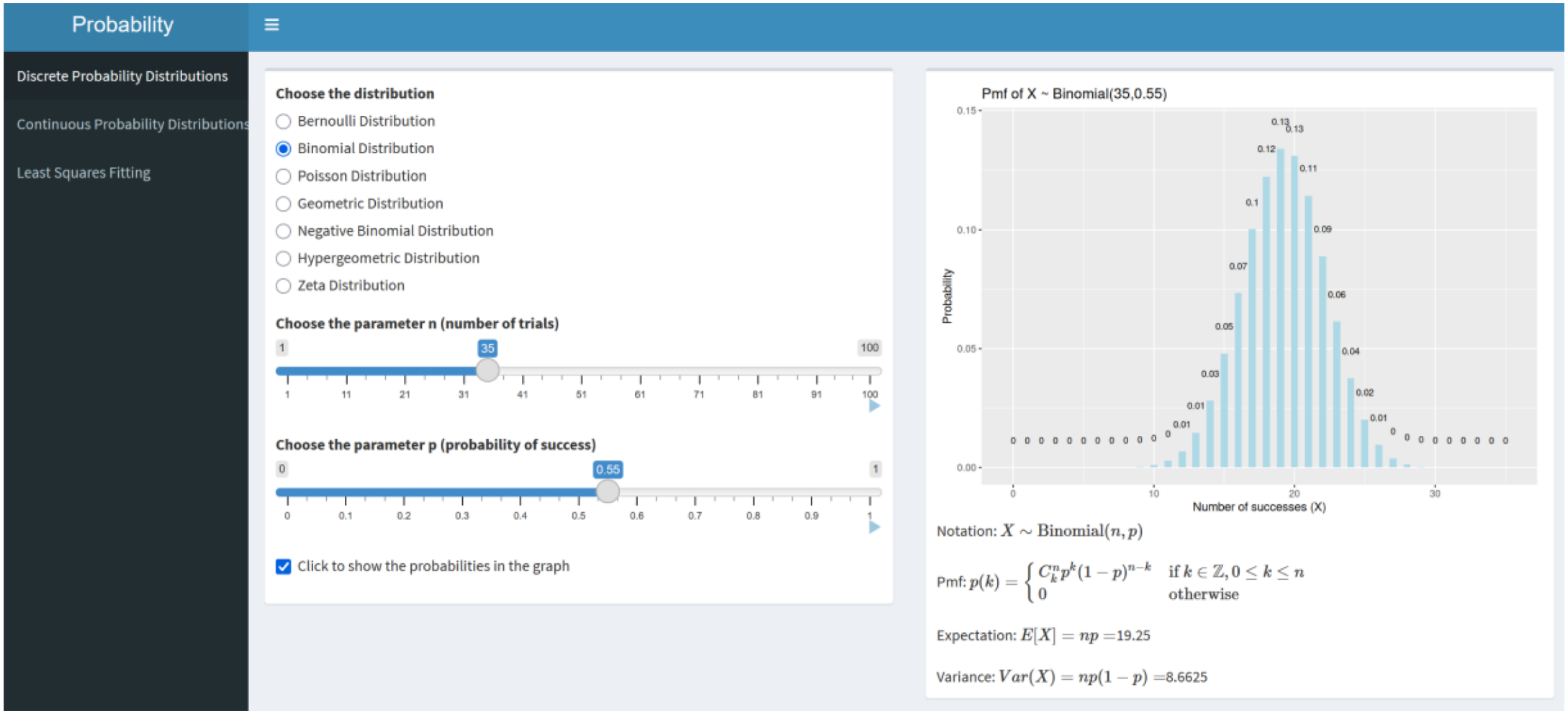
We have also developed a Probability Distribution R Shiny tool
我們亦建立了一個概率分佈 R Shiny工具 ：
https://www.math.cuhk.edu.hk/app/mathmodel/tool.html

Probability Distribution R Shiny tool 概率分佈 R Shiny工具

- https://mathmodelcuhk.shinyapps.io/probability-distribution/
- Key functionalities 主要功能：
  - visualize different discrete and continuous probability distributions
    可視化不同的離散和連續機率分佈
  - use slide bars to adjust the parameters interactively 使用滑桿互動式地調整參數
  - visualize the change in the distributions in real time 即時可視化分佈的變化
  - upload own data files and consider fitting the data using some probability distributions 上傳您自己的數據文件，並嘗試使用某些機率分佈來擬合數據

Probability Distribution R Shiny tool 概率分佈 R Shiny工具

- https://mathmodelcuhk.shinyapps.io/probability-distribution/



Statistical Modelling 統計建模

Statistical analysis 統計分析

- Parameter Identification 參數識別
- Group Behavior 群體趨勢: $N(\mu,\sigma^2)$

mean 均值 μ - central tendency 中心趨勢

variance方差 σ - central distribution 中心分佈

Statistical tools 統計工具：

- Comparison Test 比較檢定
- Hypothesis Test 假設檢定
- ANOVA 變異數分析
- Factor Analysis 因子分析
- Non-parametric Analysis 非參數分析
- Simple plotting tools 簡易繪圖工具

Operations Research / Logistic Modelling 運籌學 / 物流建模

- Optimization technique 優化技術
- Inventory control 庫存控制
- Queuing model 排隊模型
- Traffic model 流量模型

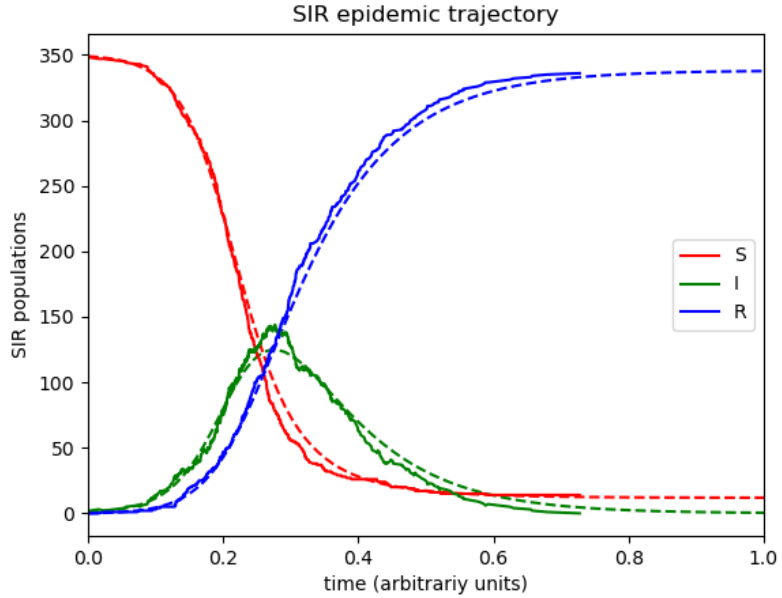
Can be achieved using packages/functions in Python, R etc.
可以使用 Python、R 等語言中的套件/函數來達到。

Recent Success in Mathematical Modelling?數學建模的近年成功案例?

Gifted Education Fund
資優教育基金 2019-004

Phase 3

Lesson 7-10 Individual Research (14 & 28 November 2020, 12 & 26 December 2020) Zoom at CityU and face-to-face consultation by appointment during 1-31 January 2021	
Topics:	• How to do research in mathematics? • Literature survey • Methodology in Mathematical Modelling • Numerical software Matlab for solving math models
List of Students' Proposals (a total of 8 students were admitted into Phase 3 on Individual Research	• Three-body movement problem in Astronomy • Cryptology for one-time password used in Internet • Mathematical and 3D Model for Covid-19 • Heat Conduction Model for Design of Clothes • Mathematical Model for Cooling of Liquids • Predicting the Maximum Growth Potential of Nations through Analysis of Current Energy and Resource Statistics • Queueing System for Minimizing Customers' Waiting Time • Mathematical Logic in Human's Behavior



Susceptible
Infectious
Recovered/Removed

$$\begin{cases} \frac{dS}{dt} = -\beta IS \\ \frac{dI}{dt} = \beta IS - \gamma I \\ \frac{dR}{dt} = \gamma I \end{cases}$$

Research Methodology: Data collection, analysis, and visualization
研究方法：資料收集、分析和視覺化

- **Data Sources 數據來源**
  - e.g. Centre for Health Protection 衛生防護中心

- **Tools for developing visualizations 視覺化工具**
  - e.g. R Shiny, Python, MATLAB, Tableau, Excel, …

- The graphical displays can help us identify trends patterns, outliers in the data and formulate hypotheses for further testing
  圖形化展示可以幫助我們識別數據中的趨勢、模式和異常值，並提出假設以進行進一步檢驗。

Big data visualization of multiple waves of COVID-19 infection in Hong Kong

International Conference on Public health and Well-being 2022
Effectiveness of public health interventions to control a pandemic, 18-19 November, 2022

Stephen CHONG[1], Jeffrey CHAN[2], Howard CHEUNG[3], Douglas NG[4], Benny HON[3*]

[1]The National Chi Nan University, Taiwan; [2]King George V School, Hong Kong;[3]The City University of Hong Kong, Hong Kong; [4]The University of Hong Kong, Hong Kong.

Corresponding author: Benny HON Benny.Hon@cityu.edu.hk

1. Background

The COVID-19 pandemic has caused significant disruptions to healthcare worldwide. Figure 1(A) shows the cumulative confirmed COVID-19 cases per million people and Figure 1(B) shows shows people's concern during the fifth wave of COVID-19 in Hong Kong. Data visualization can aid in persuading people to change their behavior. Once a virus has spread, public health officials must make critical judgments about how much information to communicate and when to dispense it. One of the most critical components of limiting an outbreak is encouraging people to adjust their behavior when it is not immediately evident that they should. Data visualization has become highly significant in communicating and convincing people. The study's aimed to visualize the COVID-19 pandemic's transmission in Hong Kong during the first five wave, as well as to improve overall preparedness and response for communicable diseases of importance to the public. The Hong Kong government has developed contingency plans and drills to test the ability of relevant government departments and organizations to deal with potential major outbreaks of infectious diseases and public health emergencies, as well as to raise community and healthcare personnel awareness and capability.

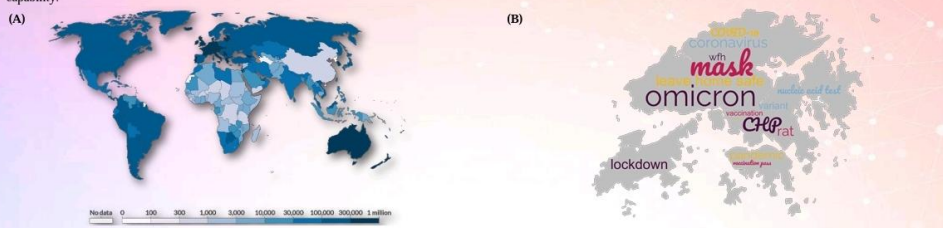


Figure 1. (A) Cumulative confirmed COVID-19 cases per million people between 22 January 2020 and 31 August 2022. (B) Word cloud shows people's concern during the fifth wave of COVID-19 in Hong Kong (source: Google analytics)

2. Methods

The daily number of confirmed COVID-19 cases in Hong Kong by the date of symptom onset was provided by the Centre for Health Protection (CHP). Data cleaning is the first step of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset. Different data visualization techniques such as line graph, bubble chart, heat map, and word clouds are used to unlock the benefits and make accurate decisions from the vast data.

Several software tools will be used for developing visualizations, such as MATLAB, Tableau and Excel. The graphical displays can help us identify trends, patterns and outliers in the data, and formulate hypotheses for further testing.

3. Results

The following figures show different visualization results and insights derived during the first five waves in Hong Kong.

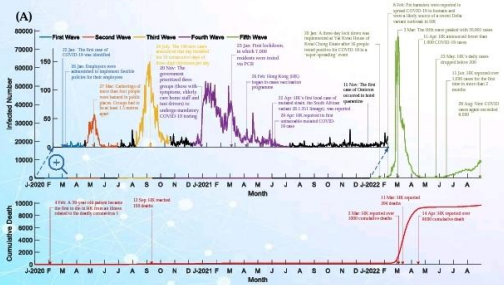


Figure 2. (A) Number of daily infected cases, cumulative death cases and highlights for the first five waves in Hong Kong. (B) Infographic: Current State of Hong Kong Economy – Five Key Points (source: https://www.amro-asia.org/infographic-current-state-of-hong-kong-economy-five-key-points/)

Figure 3. LeaveHomeSafe is a digital contact tracing app launched by the Hong Kong Government on 16 November 2020.

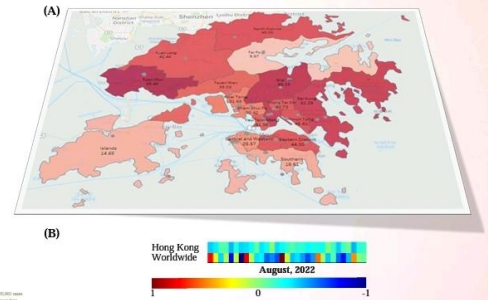


Figure 4. (A) Percentage changes of incidence of cases per 100,000 tested positive for COVID-19 by nucleic acid tests and rapid antigen tests by residential districts in July and August 2022, as well as the distribution of 43 public hospitals in Hong Kong. (B) Heat map shows the percentage of daily changes in COVID-19 Worldwide and in Hong Kong in August 2022.

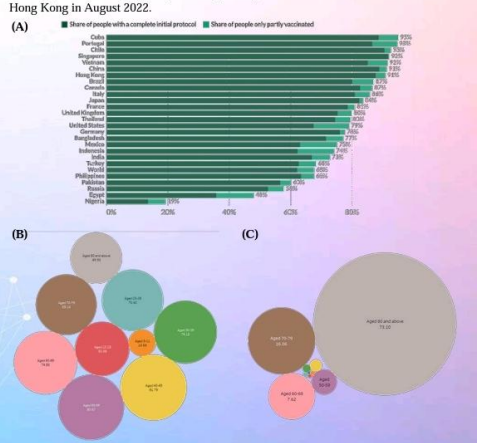


Figure 5. (A) Vaccination rate for different countries. (B) Vaccination rate for different age groups. (C) Death rate for COVID-19 patients who received one dose or less for different age groups.

COVID-19 pandemic data modelling: A case study of Hong Kong

International Conference on Public health and Well-being 2022
Effectiveness of public health interventions to control a pandemic, 18-19 November, 2022

Jeffrey CHAN[1], Stephen CHONG[2], Howard CHEUNG[3], Douglas NG[4], Benny HON[3*]

[1]King George V School, Hong Kong; [2]The National Chi Nan University, Taiwan; [3]The City University of Hong Kong, Hong Kong; [4]The University of Hong Kong, Hong Kong.

Corresponding author: Benny HON Benny.Hon@cityu.edu.hk

1. Background

Coronavirus, known as SARS-CoV-2 which causes the disease COVID-19, is a new strain of the coronavirus family. It is an RNA virus, which means it uses RNA as its genetic material. It is waterborne and partially airborne. SARS-CoV-2 includes flu-like symptoms and, in the extreme, MERS, pneumonia, loss of taste or smell, and even death. To limit the spread of this virus, how it spreads must be known; thus, mathematical models were developed, such as Susceptible-Infectious-Removed (SIR) and Susceptible-Exposed-Infectious-Removed (SEIR). We used the standard SIR model to help predict the infection number during the multiple waves of COVID-19 in Hong Kong. Given sufficient starting information and input on the duration of the outbreak, the model will assist in decision-making when preparing for a COVID-19 outbreak.

2. Methods

The SIR model [1, 2] is a general model for any infectious disease that measures the number of susceptible (S), infectious (I) and removed (R [initially always 0, as no individual has died or recovered from the disease yet] recovered or dead) individuals given the original number of susceptible and infected individuals over time (t). However, this mathematical model has some limitations.

Firstly, the mathematical model assumes that the COVID-19 epidemic is short; thus, the total population (susceptible + infectious + removed) is constant (no births or non-COVID-19 related deaths).

All models are wrong, but some are useful. **George E. P. Box**

Secondly, the mathematical model assumes that the constant rate of infection and the contacts (between the infected and susceptible individuals) is proportional and constant (β), so the more individuals are infected, the more likely a susceptible individual comes in contact with an infected individual. Thirdly, the mathematical model assumes a constant rate in which the removed individuals are increasing (γ) (death rate or recovery rate). Equations in the SIR model include:

$$\frac{dS}{dt} = -\beta SI, \quad (1)$$

$$\frac{dI}{dt} = \beta SI - \gamma I, \quad (2)$$

$$\frac{dR}{dt} = \gamma I. \quad (3)$$

The daily number of confirmed, death and recovery COVID-19 cases in Hong Kong were collected from the Centre for Health Protection (CHP), Hong Kong.

We first constructed a MATLAB (version 2022a; MathWorks, USA) function to represent the SIR model. The function included creating the basic model and passing in the parameters β, and γ, as well as the three populations at a given time. We then used an ODE solver in MATLAB, called `ode45`, to approximate the solutions to S, I and R. We computed the least squares error between our approximate solution and our gathered data at each time step. Finally, we used the MATLAB tool `fminsearch` to approximate the infection and recovery parameters.

3. Results

Figure 1 shows the susceptible, infectious and recovered populations for the first five waves in Hong Kong, **Table 1** shows the number of infected and death cases during multiple waves in Hong Kong.

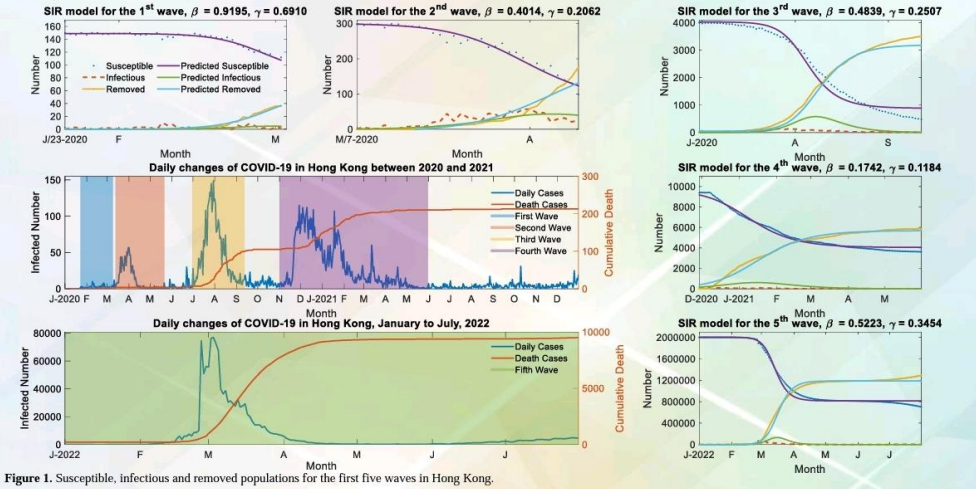


Figure 1. Susceptible, infectious and removed populations for the first five waves in Hong Kong.

Table 1. Number of infected and death cases in Hong Kong and its dominant variants

Wave	Variants	Days	Infected*	Death*	Basic reproduction ratio, β/γ
First	-	41	< 100	2	1.3307
Second	-	33	800+	2	1.9467
Third	Beta	75	3700+	93	1.9302
Fourth	Alpha, Beta, Delta, Gamma	184	5700+	102	1.4713
Fifth	Delta, Omicron	212+	1.34M+	9200+	1.5122

* Provisional figures
The data is up to 31/July/2022.

4. Discussions

We investigated the transmission dynamics and epidemiological features of COVID-19 cases in Hong Kong across epidemic waves. The estimated reproduction ratio R_0 ranged from 1.33 to 1.95. The first to fourth waves were primarily due to cases imported from Mainland China, overseas locations such as Europe and the United States, sea crew, aircrew members, domestic helpers and dancing groups. To combat the epidemic, the government implemented various measures, including work from home for civil servants, school suspension, closure of recreational facilities, reduction of importation pressure by closing some border control points and enhanced laboratory surveillance for early detection of cases. The SIR model showed that the transmission rate (β = 0.5223) for the Omicron variant was high during the fifth wave. However, the recovery rate is relatively high because of the introduction of the vaccination program in 2021. The majority of those who died had comorbidities and were unvaccinated.

5. Conclusions

COVID-19 is an exceptionally infectious disease that spreads rapidly. The SIR model can help predict the maximum number of people infected, recovered, and dead, given initial variables from current data on the infected people. However, to truly capture the dynamics of COVID-19, many modifications are needed. We would also have to consider the percentage of the populations that have received the COVID-19 vaccine, such as the Susceptible-Infectious-Removed-Vaccinated (SIRV) model. There is a significant latency period during which individuals have been infected but are not yet infectious for many vital infections. During this period, the individual is in compartment E (for exposure), and the SEIR model could be introduced in future studies.

References

[1] Kermack, W.O. and A.G. McKendrick, *A Contribution to the Mathematical Theory of Epidemics*. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, 1927. **115**(772): p. 700-721.

[2] Cooper, I., A. Mondal, and C.G. Antonopoulos, *A SIR model assumption for the spread of COVID-19 in different communities*. Chaos Solitons Fractals, 2020. **139**: p. 110057.

2 poster presentation at the 4th International Conference on Public Health and well-being 2022.

HONG KONG MEDICAL JOURNAL
LETTER TO THE EDITOR

Published in 2023. The Journal Impact Factor for the HKMJ is 3.125.

Suggestions to minimize the vaccination hesitancy of children in Hong Kong

Jeffrey Chan[1], Kei Shing Ng[2], Benny Yiu Chung Hon[3], Simon Ching Lam[4*]

[1]King George V School, Hong Kong
[2]Department of Diagnostic Radiology, Li Ka Shing Faculty of Medicine, The University of Hong Kong
[3]Department of Mathematics, The City University of Hong Kong
[4]School of Nursing, Tung Wah College, Hong Kong
[*] Corresponding author: simonlam@twc.edu.hk

To the Editor

COVID-19 vaccine hesitancy is high initially (1) and still prevalent in Hong Kong.

frontiers | Frontiers in Public Health

TYPE Brief Research Report
PUBLISHED xx xx 2022
DOI 10.3389/fpubh.2022.1053873

Check for updates

OPEN ACCESS

EDITED BY
Reza Lashgari,
Shahid Beheshti University, Iran

REVIEWED BY
Mohamed El-Kassas,
Helwan University, Egypt
Qi Ye,
South China Normal University, China
Saroj Chandra,
OP Jindal University, India

*CORRESPONDENCE
Simon Ching Lam
simonlam@twc.edu.hk
Benny Yiu Chung Hon
benny.hon@cityu.edu.hk

SPECIALTY SECTION
This article was submitted to Infectious Diseases: Epidemiology and Prevention, a section of the journal Frontiers in Public Health

RECEIVED 26 September 2022
ACCEPTED 21 November 2022
PUBLISHED xx xx 2022

CITATION
Ng KS, Hon JMH, Chong SCC, Cheung HHK, Chan J, Lam SC and Hon BYC (2022) Implementation of the compulsory universal testing scheme in Hong Kong: Mathematical simulations of a household-based pooling approach. *Front. Public Health* 10:1053873. doi: 10.3389/fpubh.2022.1053873

COPYRIGHT
© 2022 Ng, Hon, Chong, Cheung, Chan, Lam and Hon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Implementation of the compulsory universal testing scheme in Hong Kong: Mathematical simulations of a household-based pooling approach

Kei Shing Ng[1], Jeffrey Man Hin Hon[2,3], Stephen Chau Chun Chong[4], Howard Ho Kan Cheung[5], Jeffrey Chan[6], Simon Ching Lam[7]* and Benny Yiu Chung Hon[5,8]*

[1]Department of Diagnostic Radiology, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, Hong Kong SAR, China, [2]NVIDIA AI Technology Center (NVAITC), NVIDIA, Santa Clara, CA, United States, [3]Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong SAR, China, [4]Department of Applied Chemistry, National Chi Nan University, Puli, Taiwan, [5]Department of Mathematics, City University of Hong Kong, Kowloon Tong, Hong Kong SAR, China, [6]King George V School, Hong Kong, Hong Kong SAR, China, [7]School of Nursing, Tung Wah College, Hong Kong, Hong Kong SAR, China, [8]Department of Psychology, University of Science and Technology of China, Hefei, China

This study aims to propose a pooling approach to simulate the compulsory universal RT-PCR test in Hong Kong and explore the feasibility of implementing the pooling method on a household basis. The mathematical model is initially verified, and then the simulation is performed under different prevalence rates and pooled sizes. The simulated population is based in Hong Kong. The simulation included 10,000,000 swab samples, with a representative distribution of populations in Hong Kong. The samples were grouped into a batch size of 20. If the entire batch is positive, then the group is further divided into an identical group size of 10 for re-testing. Different combinations of mini-group sizes were also investigated. The proposed pooling method was extended to a household basis. A representative from each household is required to perform the RT-PCR test. Results of the simulation replications, indicate a significant reduction ($p < 0.001$) of 83.62, 64.18, and 48.46% in the testing volume for prevalence rate 1, 3, and 5%, respectively. Combined with the household-based pooling approach, the total number of RT-PCR is 437, 304, 956, 133, and 1,375,795 for prevalence rates 1, 3, and 5%, respectively. The household-based pooling strategy showed efficiency when the prevalence rates in the population were low. This pooling strategy can rapidly screen people in high-risk groups for COVID-19 infections and quarantine those who test positive, even when time and resources for testing are limited.

KEYWORDS
COVID-19, compulsory universal test, prevalence rate, sample pooling, mass screening, simulation

A brief research report was published at the journal ‘Frontiers in Public Health’, which has a high impact factor of 6.641. Frontiers is ranked the third most-cited publisher in USA.

一份簡短的研究報告發表在《公共衛生前沿》雜誌上，該雜誌的影響因子高達6.641。《前沿》是美國被引用次數排名第三的學術出版社。

Advanced Modelling Examples 進階建模例子

- See our **Math modelling e-book:** 見我們的**數學建模電子書：**
  - https://www.math.cuhk.edu.hk/~mathcal/MM/
  - Username: mathmodel
  - Password: mm@2024

- Predicting Stock Prices Using Linear and Nonlinear Regression
  使用線性和非線性迴歸預測股票價格
  http://mathcal.math.cuhk.edu.hk:7537/

- Estimate the Percentage of Electric Vehicles among All Cars in Hong Kong in 2030
  估算2030香港電動車佔總數的百分比
  http://mathcal.math.cuhk.edu.hk:7536/

- Extracting information from S-shaped curves of life achievement
  人生成就的 S 形曲線
  http://mathcal.math.cuhk.edu.hk:7562/

Advanced Modelling Examples 進階建模例子

- Modelling with Difference Equations 用差分方程建模
  http://mathcal.math.cuhk.edu.hk:7544/

- The Logistic or Inhibited Growth Model 邏輯/抑制生長模型
  http://mathcal.math.cuhk.edu.hk:7546/

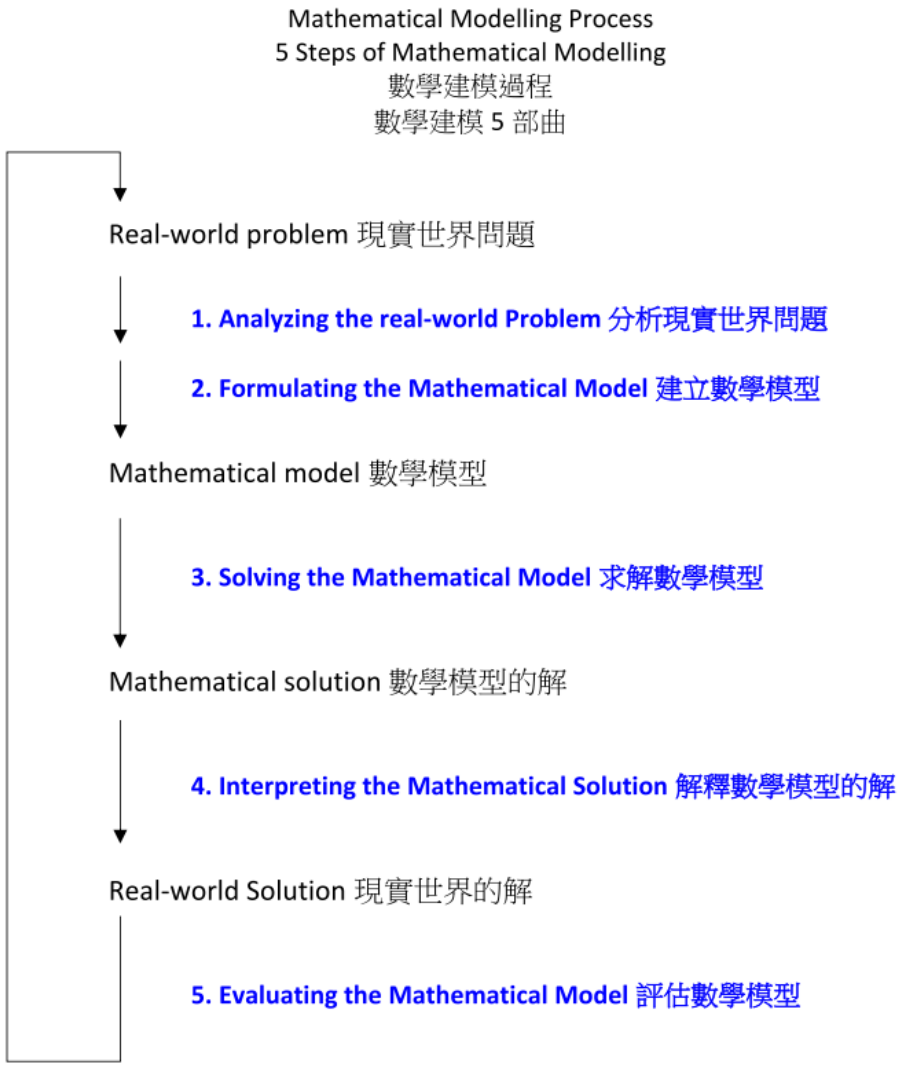
- Predicting Price Indices and Weather Prediction Using Markov Chains
  使用馬可夫鏈預測價格指數和天氣預報
  http://mathcal.math.cuhk.edu.hk:7539/

- Modelling the Spread of Information Using Social Networks, Node Centralities, and Data Fitting Approaches
  使用社交網路、節點中心性和資料擬合方法對資訊傳播進行建模
  http://mathcal.math.cuhk.edu.hk:7538/

Summary: Learning Resources for Math Modelling
數學建模的學習資源

- **Mathematical Modelling @ CUHK Mathematics**
  https://www.math.cuhk.edu.hk/app/mathmodel/

- **Mathematical Modelling e-book 數學建模電子書：**
  https://www.math.cuhk.edu.hk/~mathcal/MM/
  Username: mathmodel
  Password: mm@2024
  - Mathematical modelling concepts
  - A large variety of examples
  - IT tools freely available

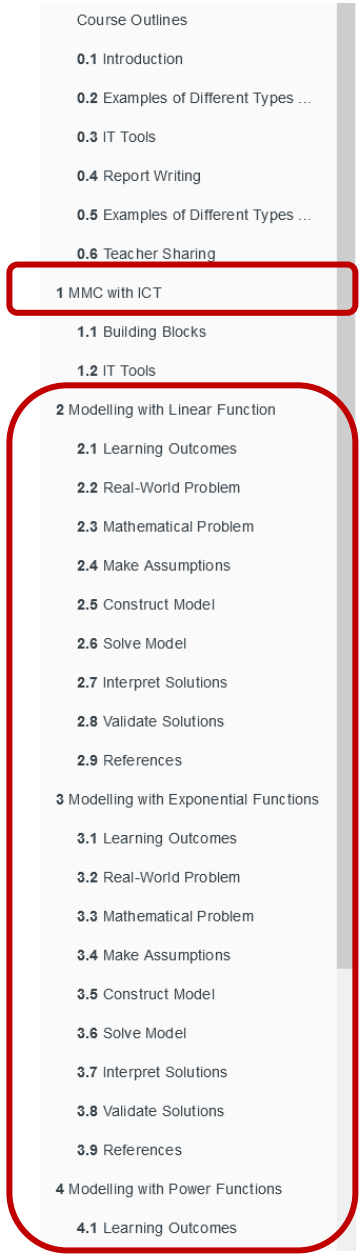


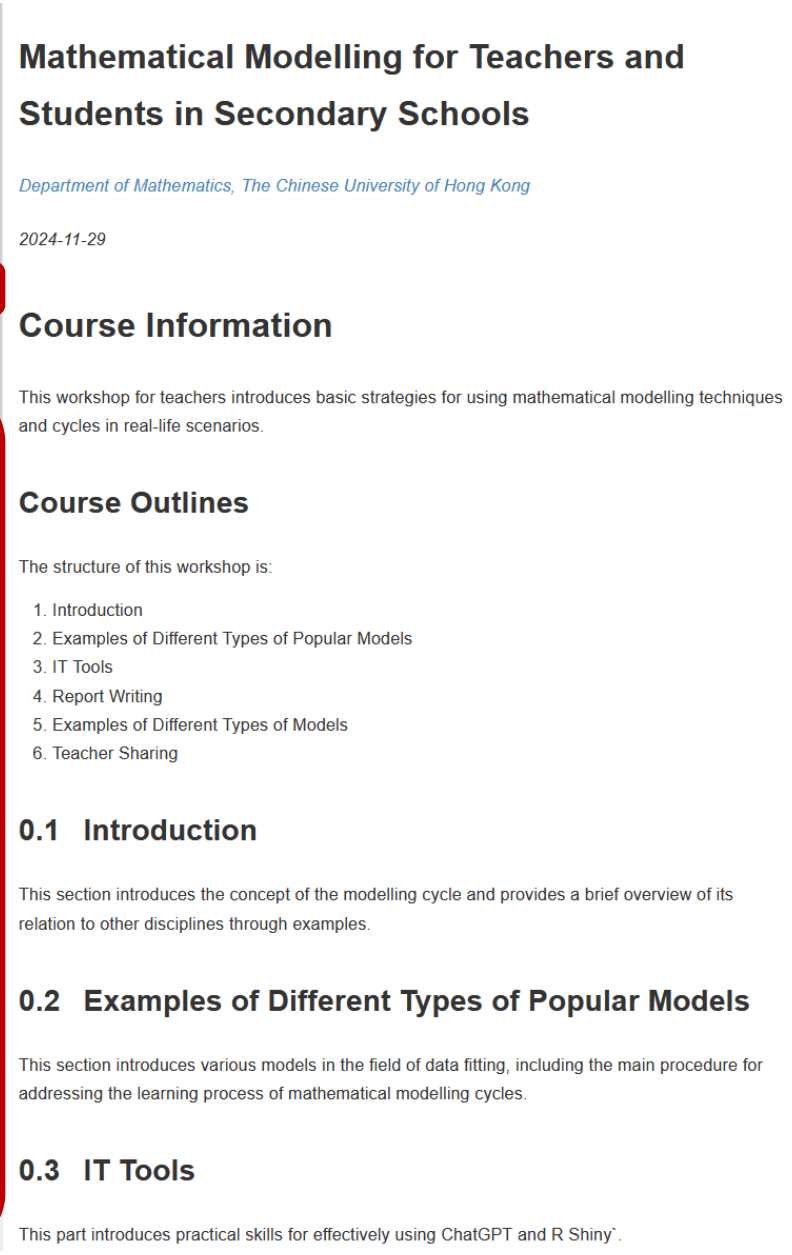
Summary: Math Modelling e-book 數學建模電子書

- **Mathematical modelling concepts**
  數學建模概念
  - Math modelling process 數學建模過程
  - Key steps關鍵步驟

- **Common models 常用模型：**
  - Linear functions線性函數
  - Exponential functions指數函數
  - Power functions冪函數
  - Trigonometric functions三角函數
  - Sigmoidal functions S形函數
  - All with detailed formulations, derivations, and examples 均提供詳細的公式、推導過程和範例

- **More advanced math tools 進階數學工具:**
  - Probability 概率, network model 網絡模型, …
  - Introduced via specific real-life problems (price prediction, social network etc.)
    透過具體問題（價格預測、社交網路等）介紹





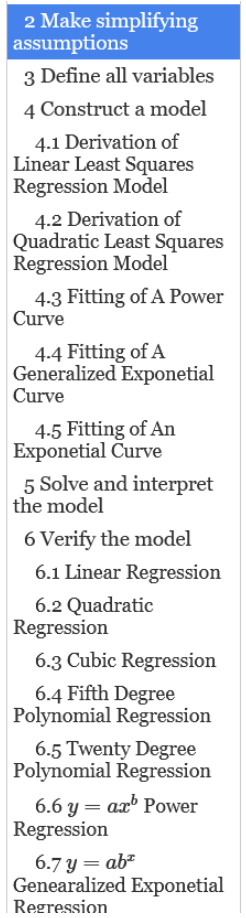
Summary: IT Tools for Math Modelling 數學建模 IT 工具

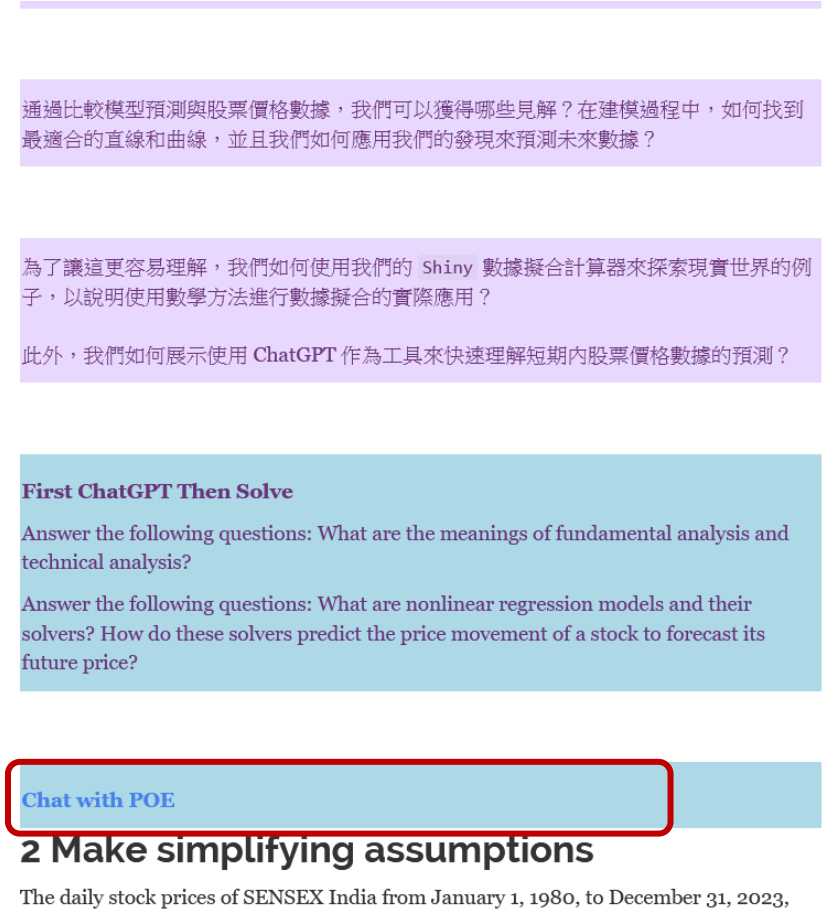
- *First-ChatGPT-Then-Solve*

- Use AI tools to help us:
  利用 AI 工具幫助我們：
  - Understand problem background 了解問題背景
  - Identify relevant factors 找出相關因素
  - Locate datasets 尋找數據集
  - …

- Fact-checking is important 核實事實很重要！

- Freely available AI tools 免費 AI 工具
  - Poe
  - Microsoft Copilot
  - DeepSeek
  - …





Summary: IT Tools for Math Modelling 數學建模 IT 工具

- **Computing and visualization tools 計算與視覺化工具 (R Shiny, GeoGebra, Desmos, Python, R, …)**

- R Shiny: a package for building interactive web apps based on the R programming language
  一個基於 R 程式語言的互動式網絡應用程式

- We have developed several R Shiny tools for math modelling我們開發了多個數學建模的R Shiny工具：
  https://www.math.cuhk.edu.hk/app/mathmodel/tool.html
  - *Find What Fits with R Shiny 使用 R Shiny 尋找最佳擬合線*
  - *Linear Regression with R Shiny 使用 R Shiny 進行線性迴歸*
  - *Nonlinear Regression with R Shiny 使用 R Shiny 進行非線性迴歸*
    - for XY data 適用於 XY 數據
    - for time series data 適用於時間序列數據
  - *General Fitting with R Shiny 使用 R Shiny 進行一般擬合*
  - … and more 以及更多！

- We have also provided some introductions to the functionalities of GeoGebra, Desmos, Python, and R on our website
  我們還在網站介紹了 GeoGebra、Desmos、Python 和 R 的功能。

Please complete the Course Evaluation:
請完成以下課程評估表：
https://forms.gle/kvEoHvziL7VoKYA38

Mathematical Modelling @ CUHK Mathematics:
https://www.math.cuhk.edu.hk/app/mathmodel

Contact:
mathmodel@math.cuhk.edu.hk

Thank you!