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

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

Prof. CHOI Pui Tung Gary 蔡沛彤教授

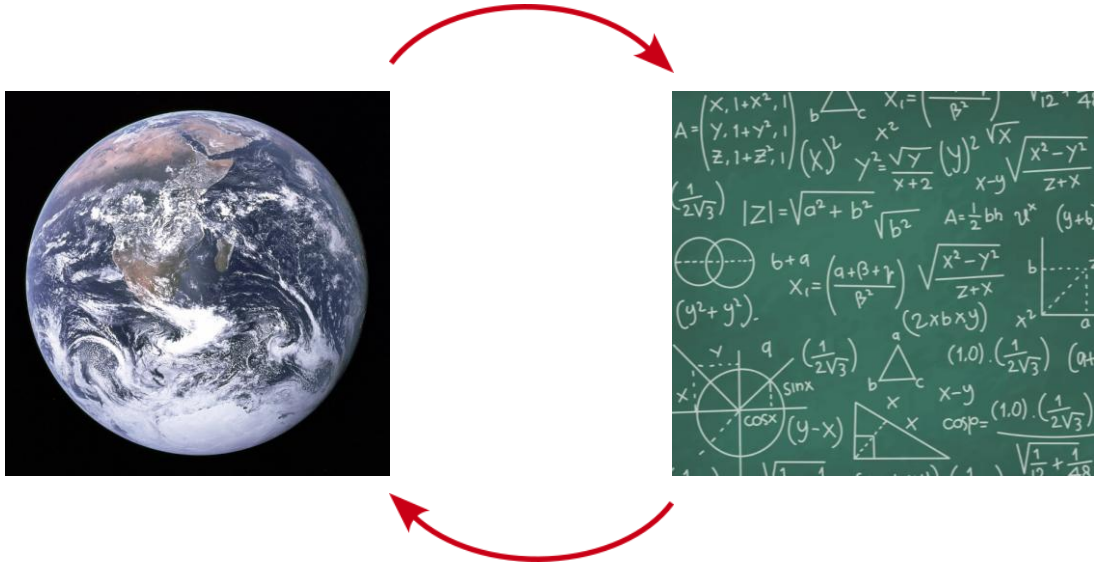
Dr. Jeff Chak-Fu Wong 黃澤富博士

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

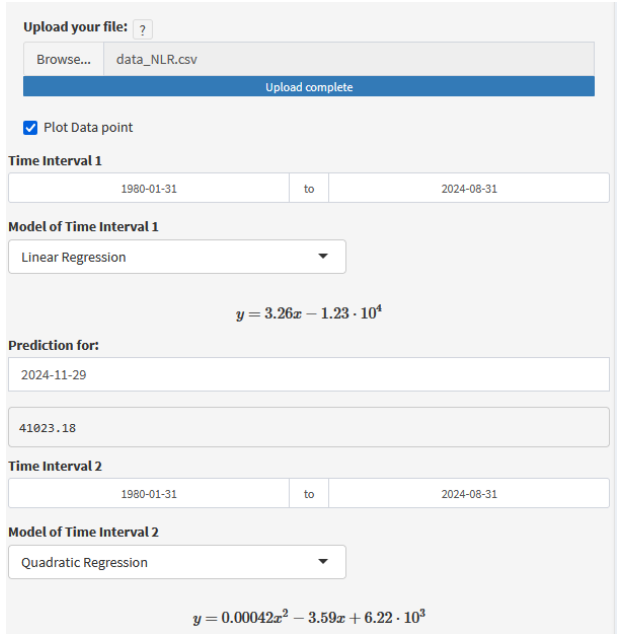
What Mathematical Concepts may be involved in Mathematical Modelling? 數學建模會用到甚麼數學概念？

- 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

$$y = ab^x$$

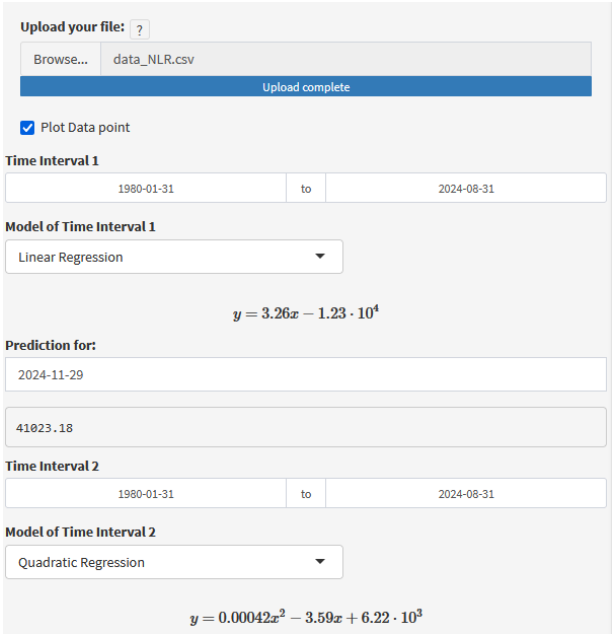
  - Logarithmic 對數函數

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

  - Power model 冪模型

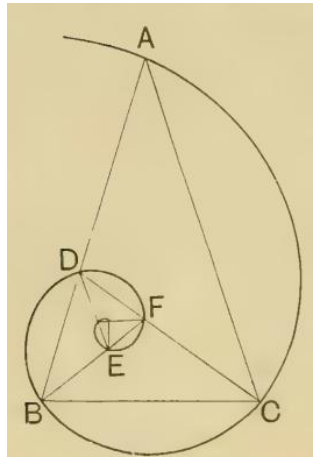
$$y = ax^b$$

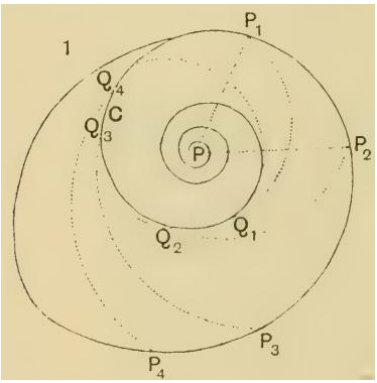
  - …



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
    - $\sin x$, $\cos x$, …





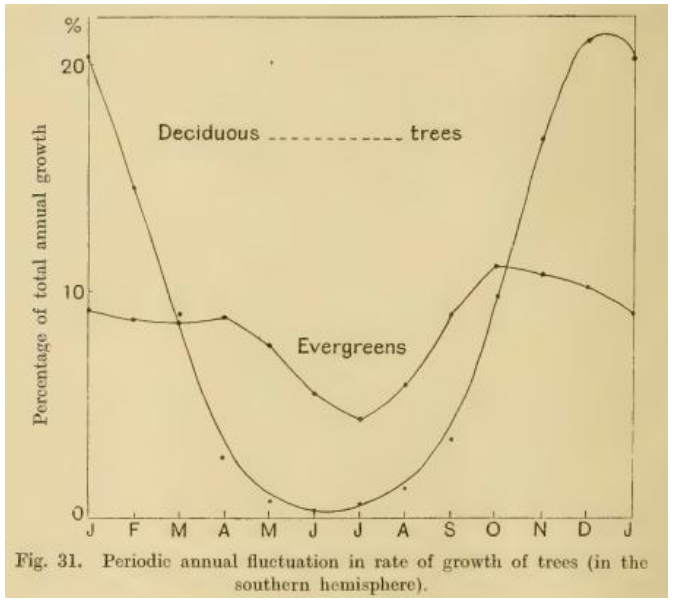
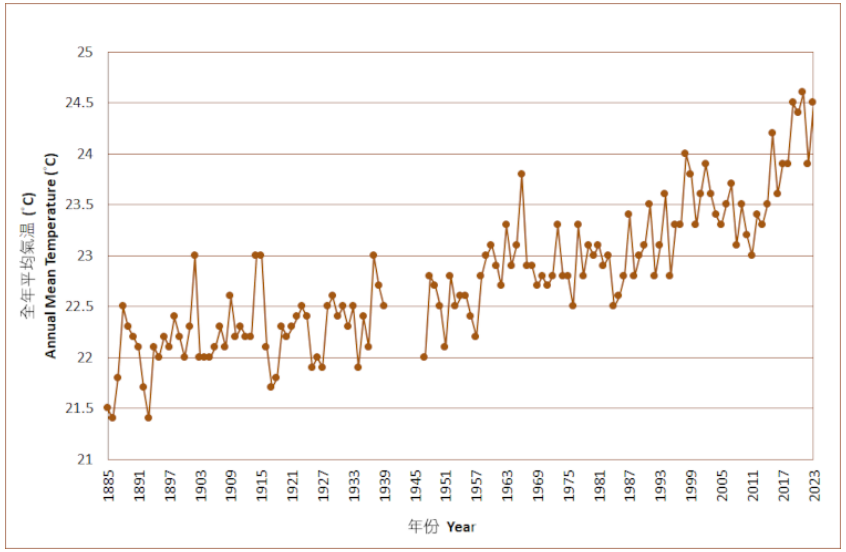


Fig. 31. Periodic annual fluctuation in rate of growth of trees (in the southern hemisphere).

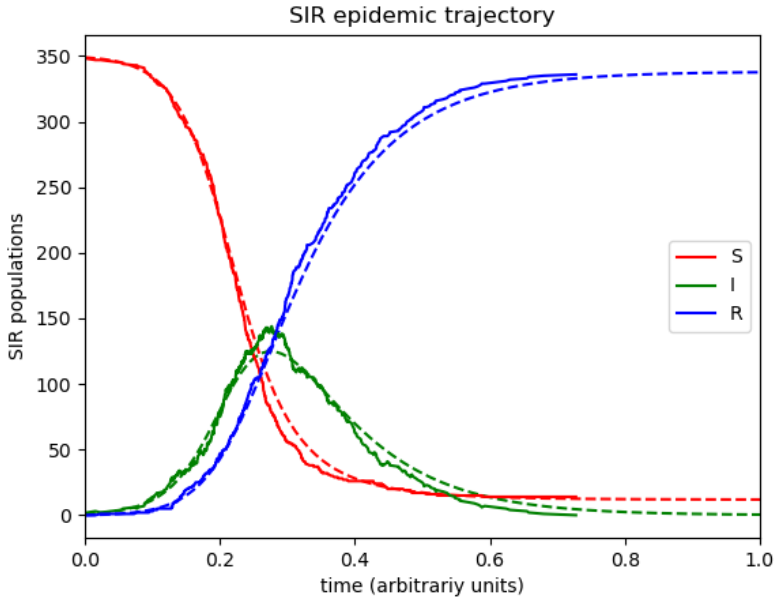
D'Arcy Thompson, *On Growth and Form* (1917)

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。

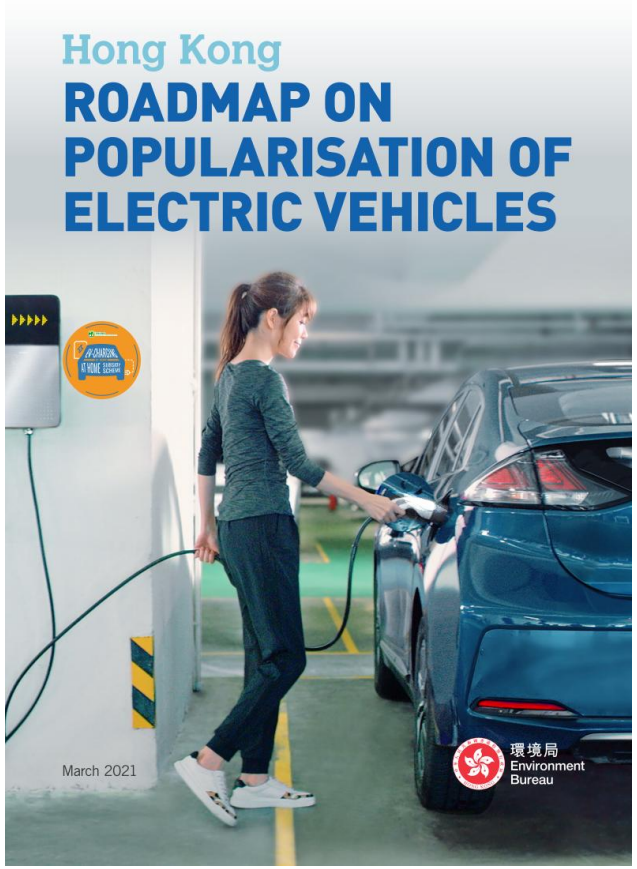
Example: Estimating the Percentage of Electric Vehicles among All Cars in Hong Kong in 2030
估算在2030年香港電動車佔所有車輛的百分比

- **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

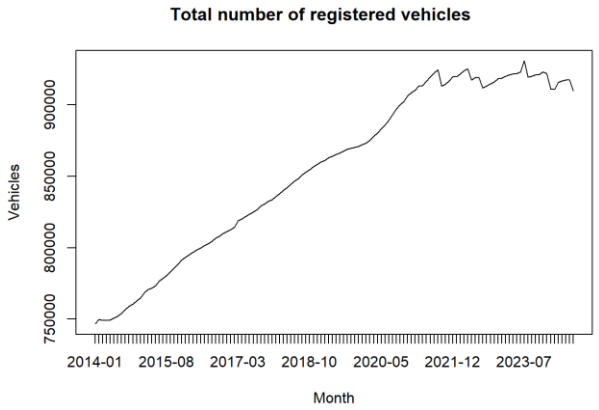
  - How can we analyze the percentage of electric vehicles among all cars in Hong Kong?
    如何分析香港所有汽車中電動車的比例？

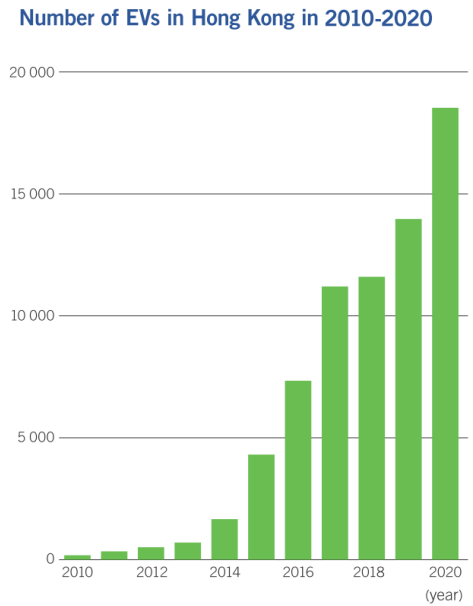


Source: Environmental Protection Department

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



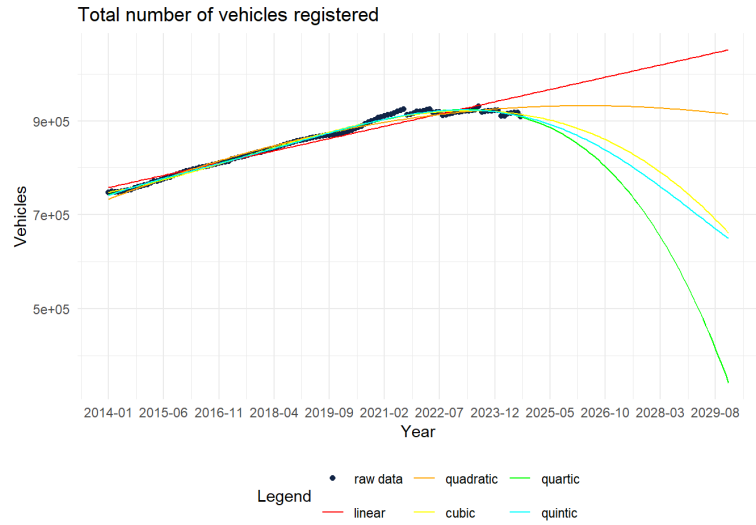


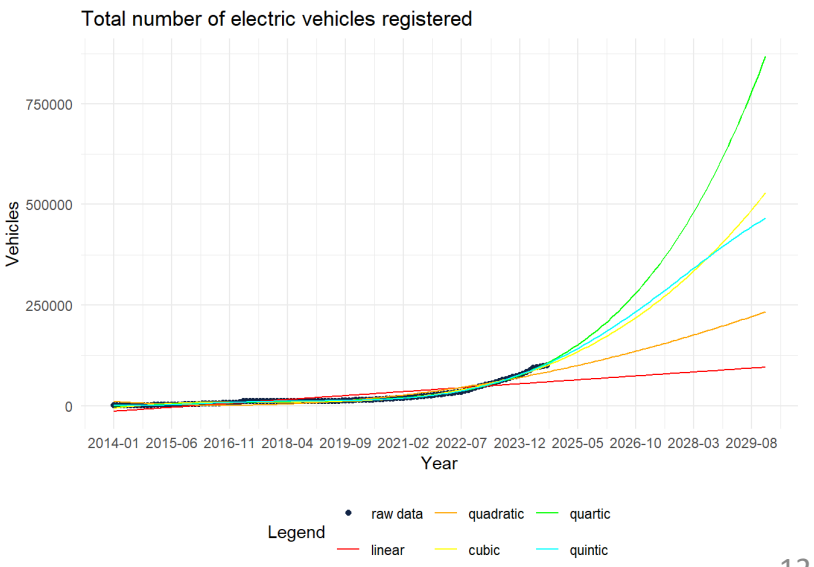
Data from Environmental Protection Department
https://www.epd.gov.hk/epd/english/top.html

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？



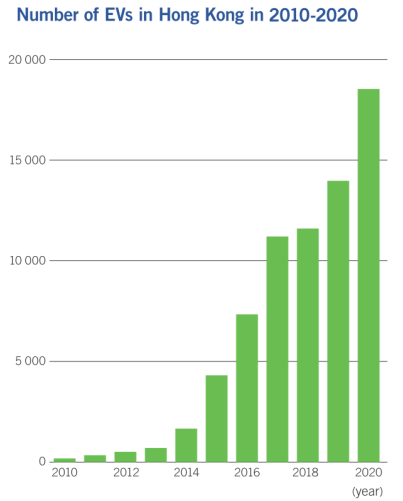


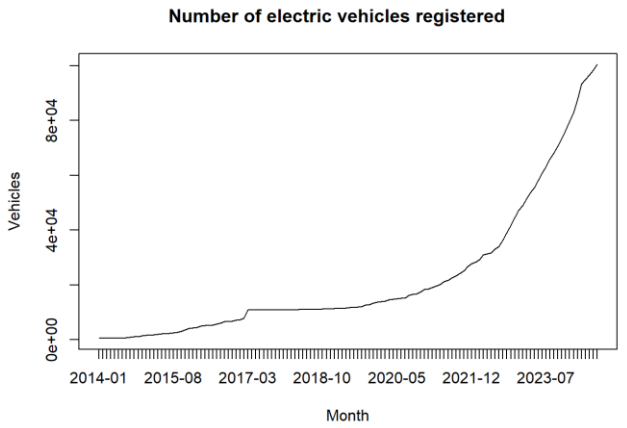
Example: Estimating the Percentage of Electric Vehicles among All Cars in Hong Kong in 2030
估算在2030年香港電動車佔所有車輛的百分比

- **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

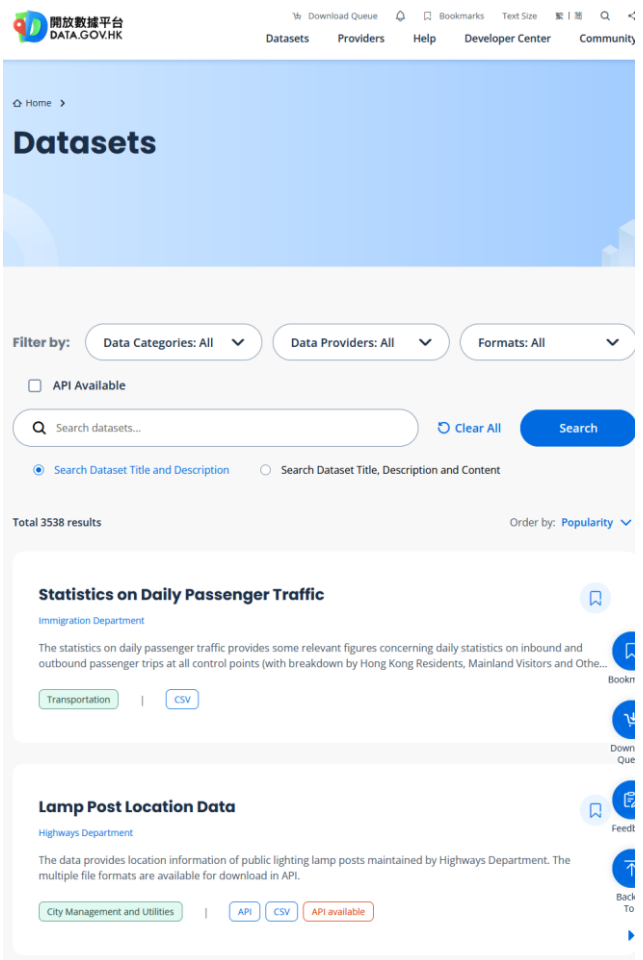
  - How can we estimate the carbon emissions of Hong Kong for reviewing climate change efforts?
    我們如何估計香港的碳排放，以檢討應對氣候變化工作？

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）

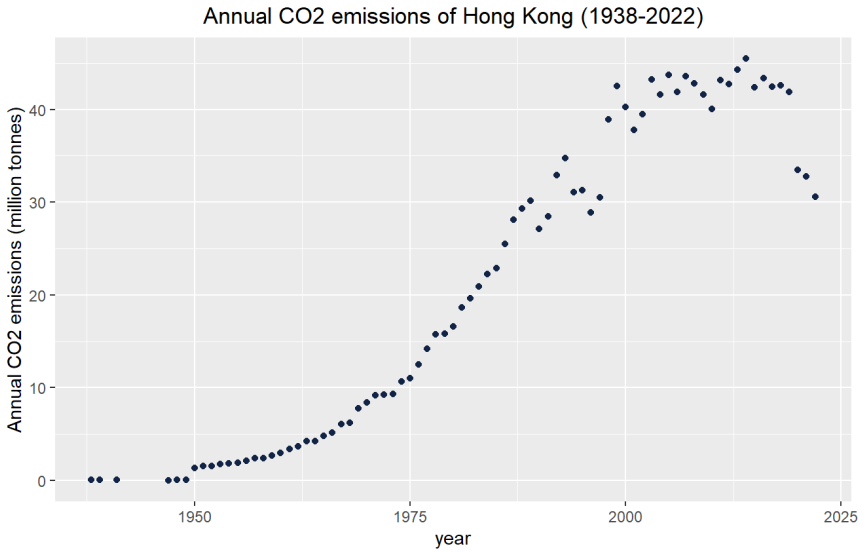
  - Relevant data相關數據：
    - DATA.GOV.HK香港政府開放數據平台 https://data.gov.hk/
    - Our World in Data https://ourworldindata.org/

  - 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）
    - …

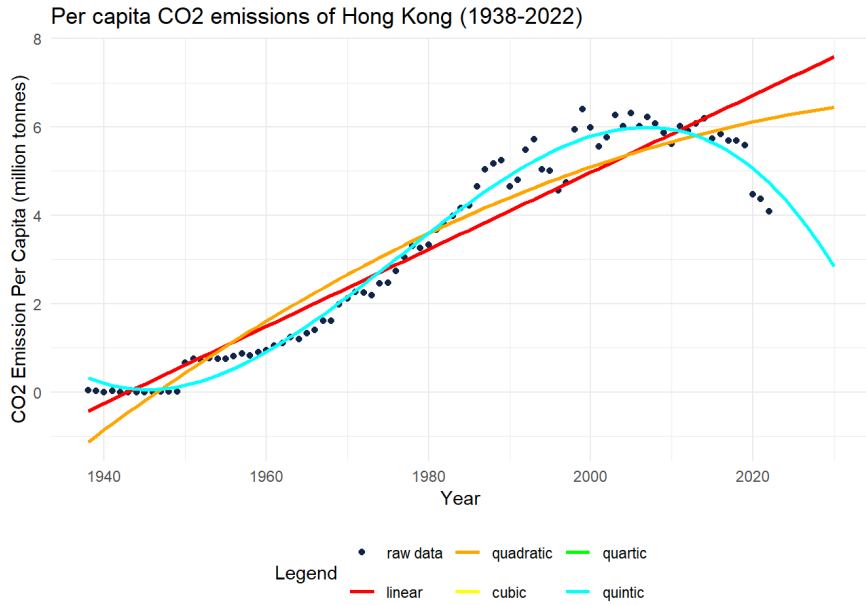


Example: Estimating the Carbon Emissions of Hong Kong in 2030
估算香港在2030年的碳排放量

- **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

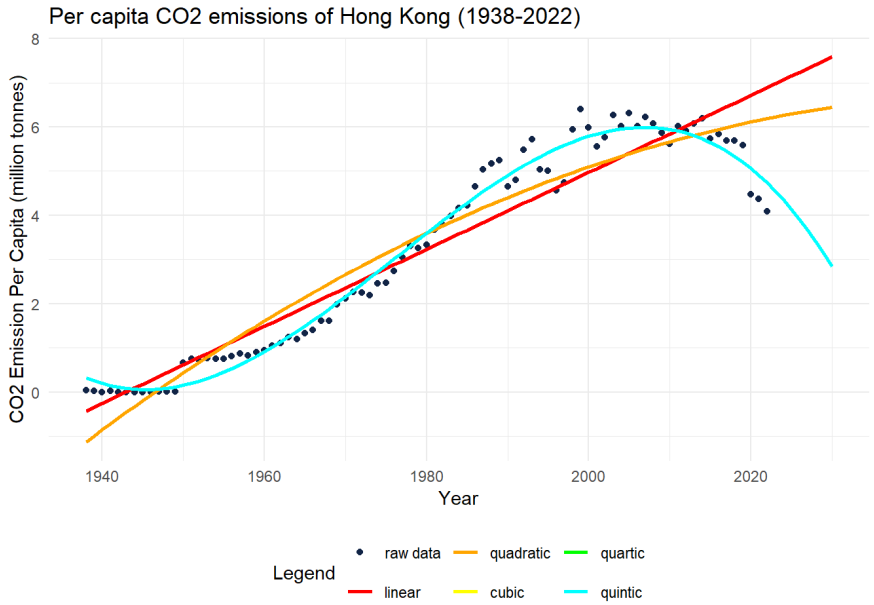


- **4. Interpreting the mathematical solution:**
  解釋數學模型的解：
  - Is the result increasing/decreasing?增加還是減少？
    - Imply the trend of carbon emission體現了碳排放的趨勢
  - Any sharp changes in the value? 數值是否存在劇烈變化？
    - Imply some rapid change in certain years
      顯示某些年份的快速變化
    - Align with technology/policy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

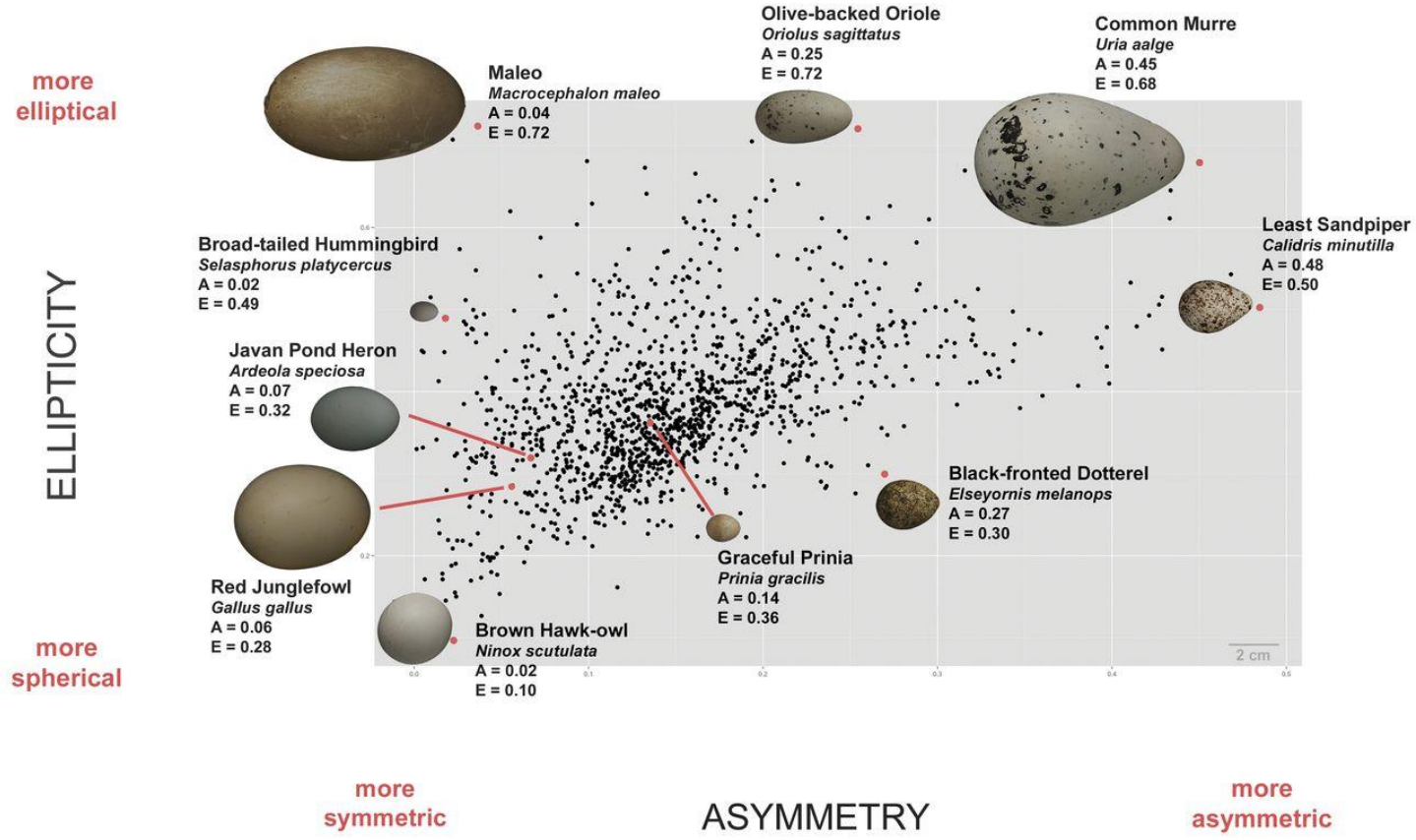
Example: Modelling the Shape of Eggs 蛋形狀的數學建模

- **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, …

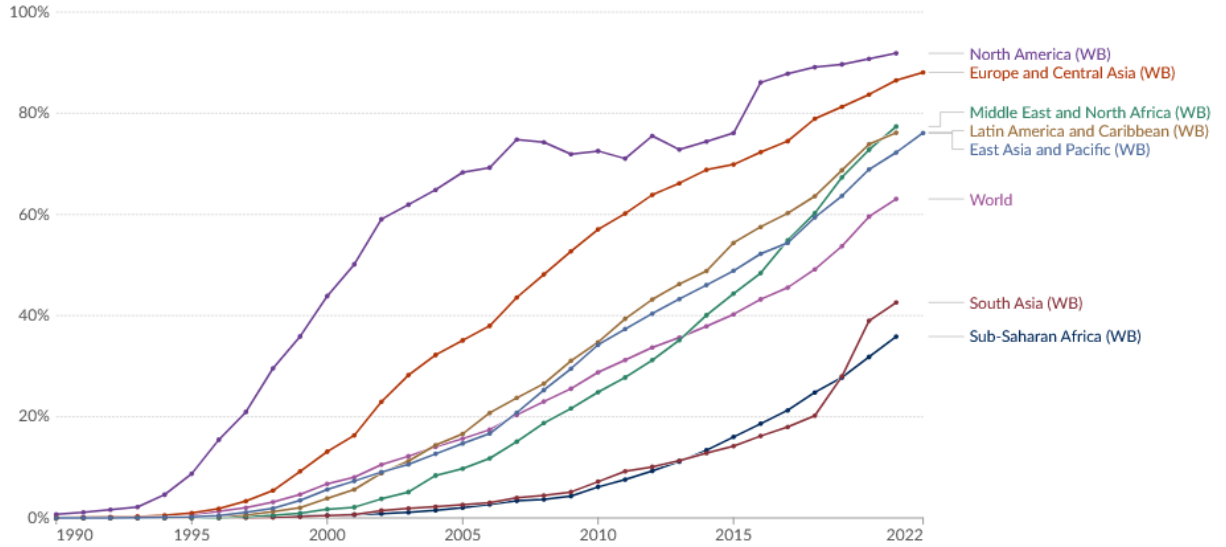
  - How to develop a mathematical model to study their relationship?
    如何建立數學模型來研究它們的關係？

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？

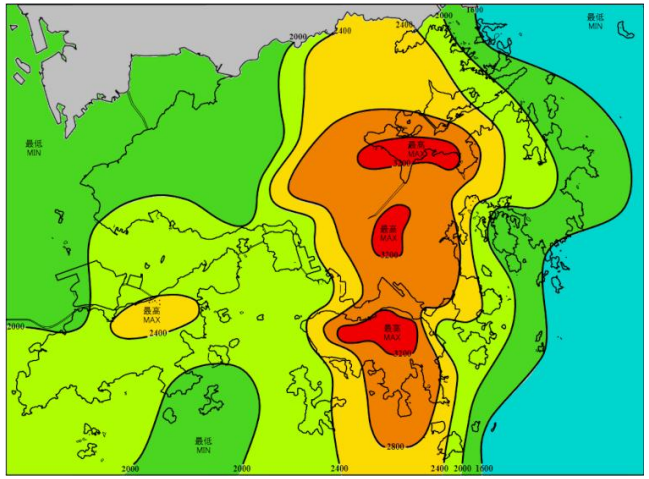


https://en.wikipedia.org/wiki/Global_Internet_usage

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

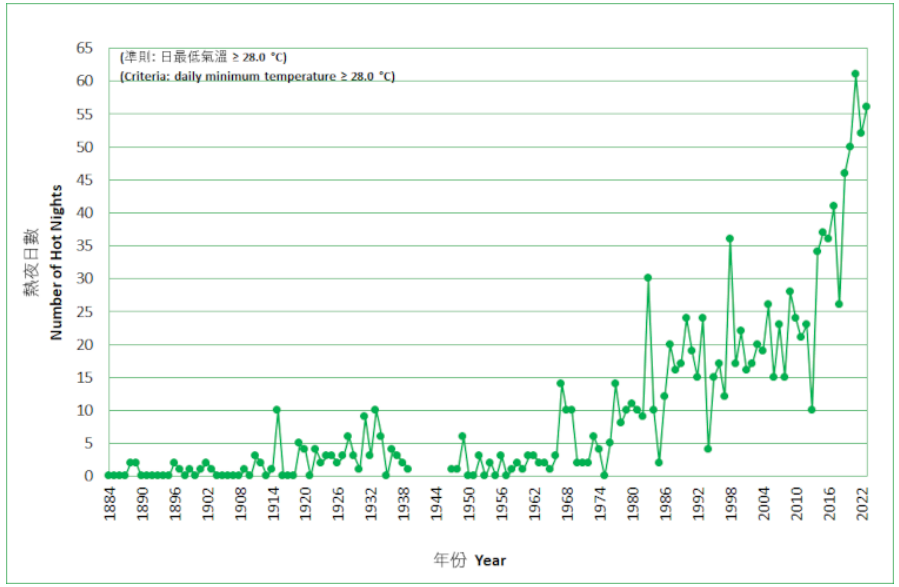
- **Real-world problem 現實生活問題:**
  - Weather prediction 天氣預報
  - Understanding climate change了解氣候變化
  - How can we predict weather or climate change?
    我們如何預測天氣或氣候變遷？

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

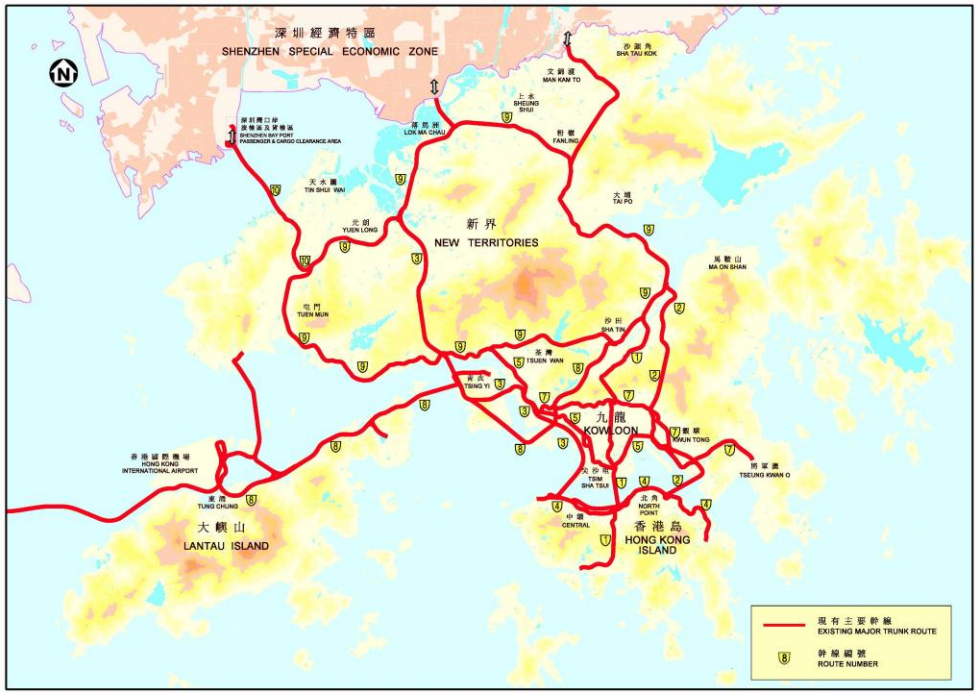
Hong Kong Observatory 香港天文台:
https://www.hko.gov.hk/en/index.html



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?

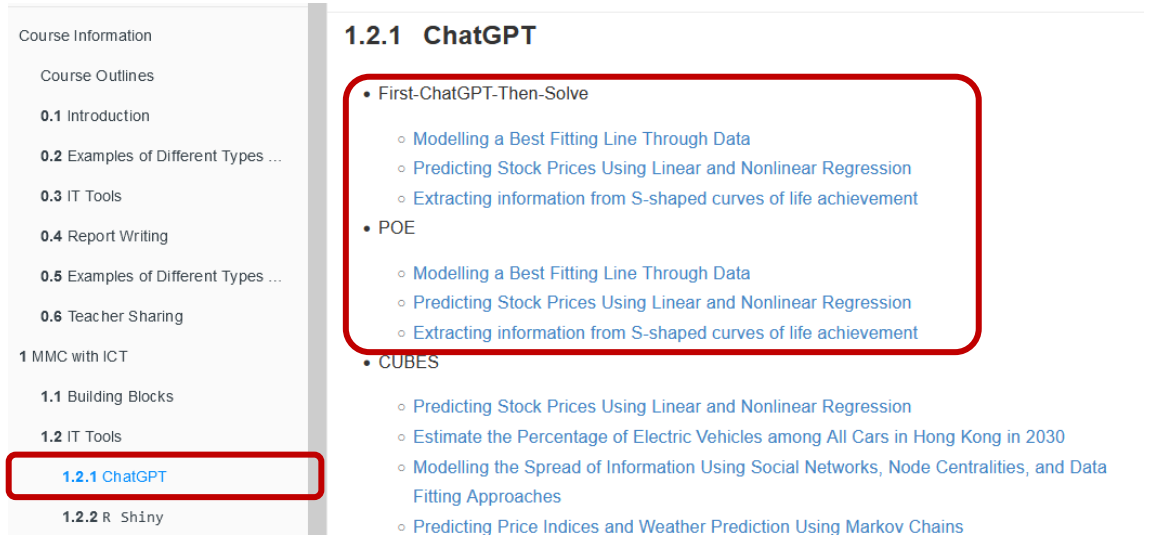
我們如何使用AI來協助我們進行數學建模？

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

- *First-ChatGPT-Then-Solve (FCTS)*
  strategy for mathematical modelling
  數學建模策略

- Use AI-based 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
  - …

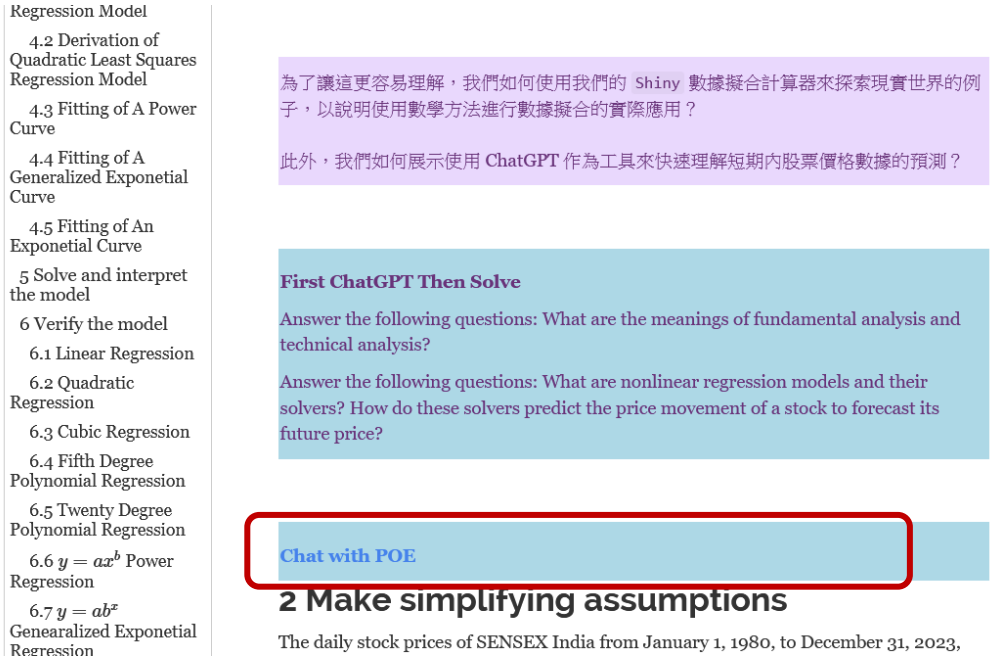
- Fact-checking is important!
  核實事實很重要！



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

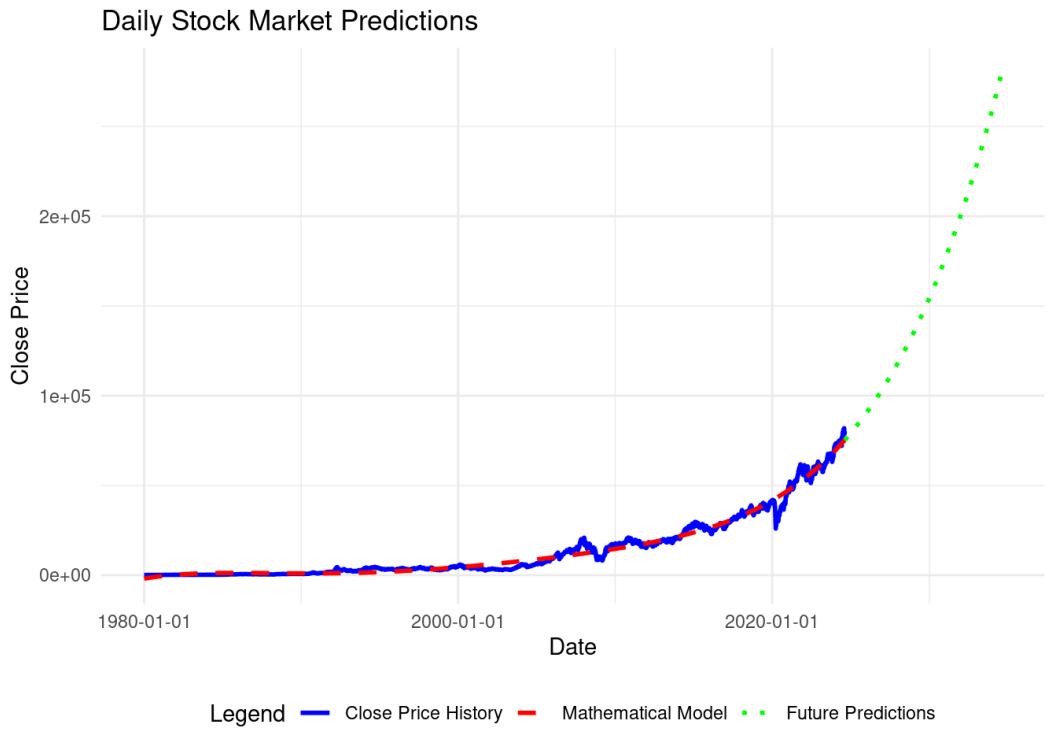
Some examples of mathematical modelling prompts
一些數學建模提示範例

- 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)**
  非線性迴歸 (時序數據) R Shiny 工具
  - 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

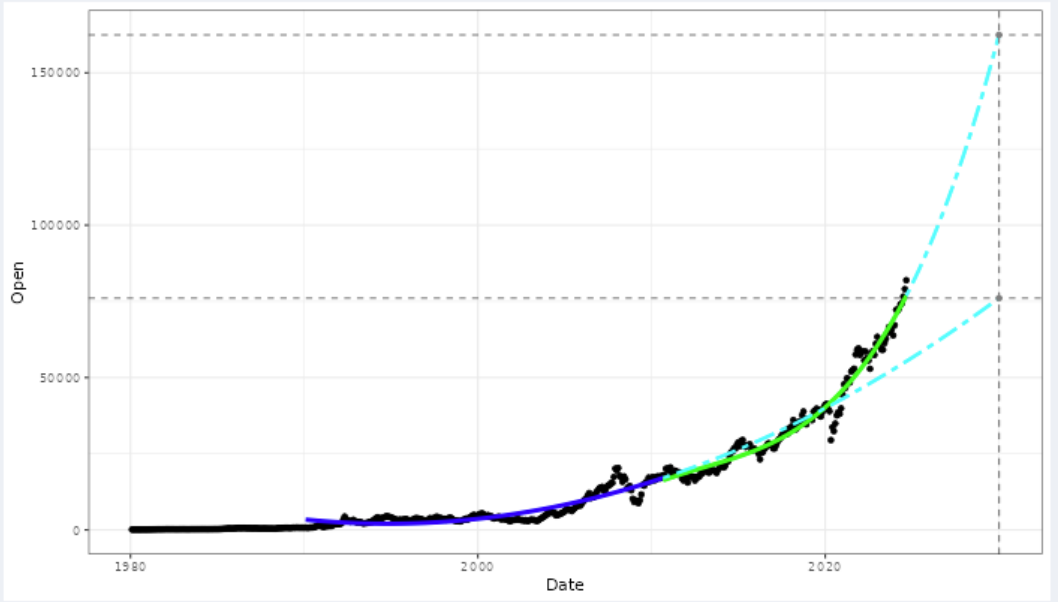
Model ?

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 in the tool (also available as data_NLR.csv on our website)
     考慮工具中的股票市場資料範例數據 (亦可見我們網頁上的data_NLR.csv)

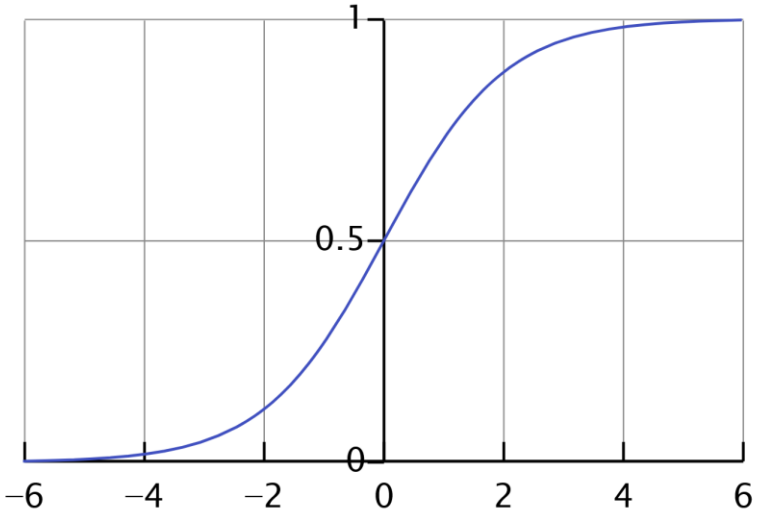
  2. Use the tool to try different models 使用工具嘗試不同的模型

  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

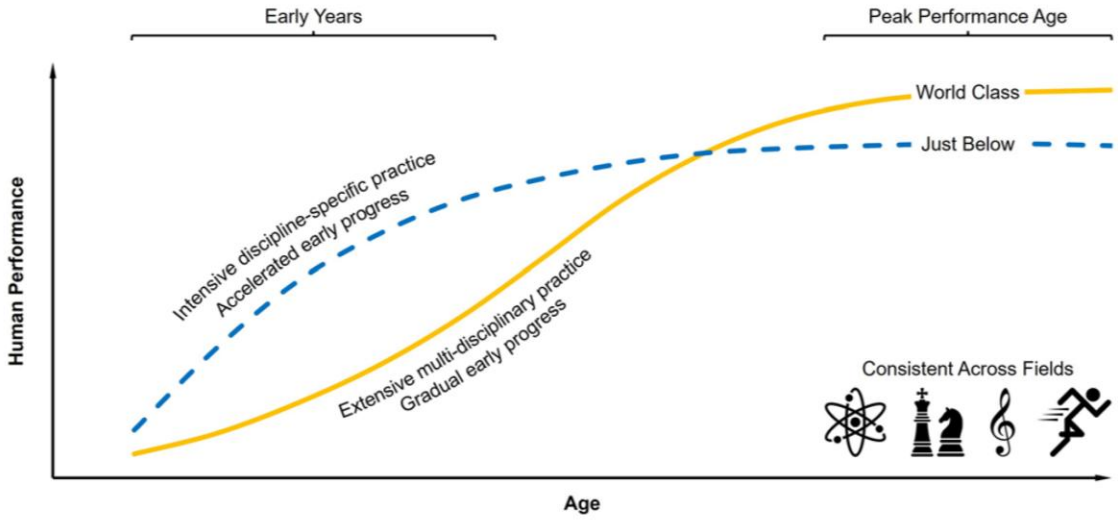
- Modelling life achievement using S-shaped curves透過 S 形函數進行人生成就的建模
  http://mathcal.math.cuhk.edu.hk:7562/

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！

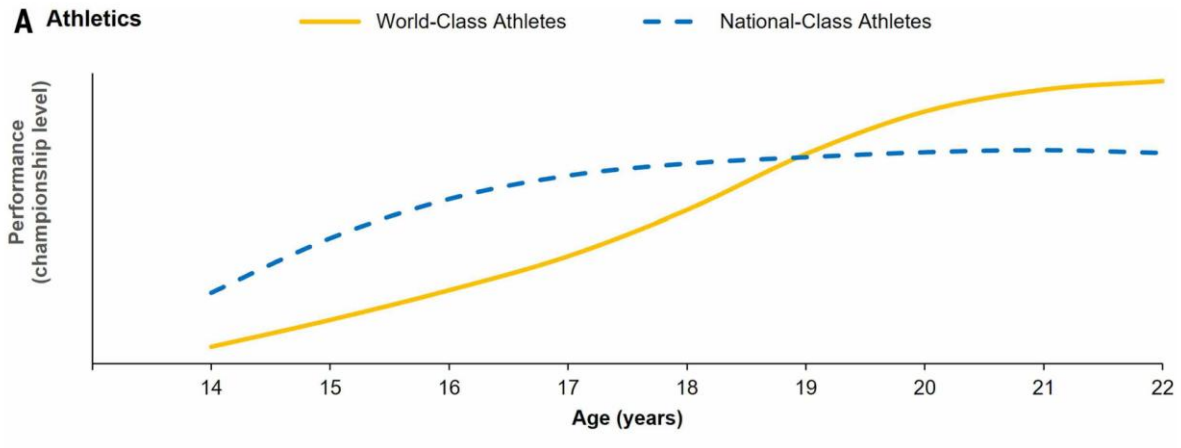
- *Recent discoveries on the acquisition of the highest levels of human performance*, **Science**, December 2025 *關於獲得人類最高水平表現的最新發現*，《科學》，2025 年 12 月 https://www.science.org/doi/10.1126/science.adt7790

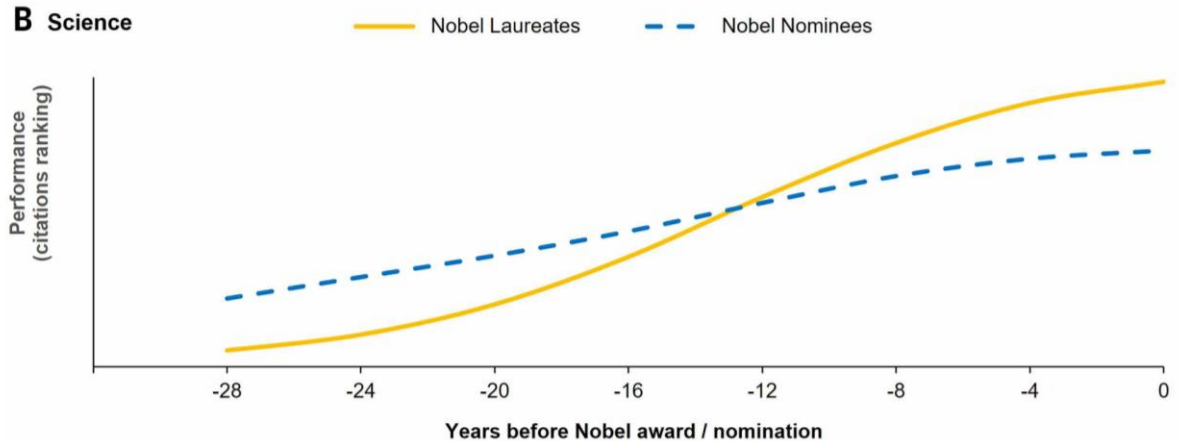
- Analyzed the development of more than 34,000 adult international top performers in different domains, including Nobel laureates, the most renowned classical music composers, Olympic champions, and the world's best chess players.
  綜合研究了 34,000 多名不同領域的國際頂尖成年人的發展情況，其中包括諾貝爾獎得主、最著名的古典音樂作曲家、奧運冠軍和世界上最好的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



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？

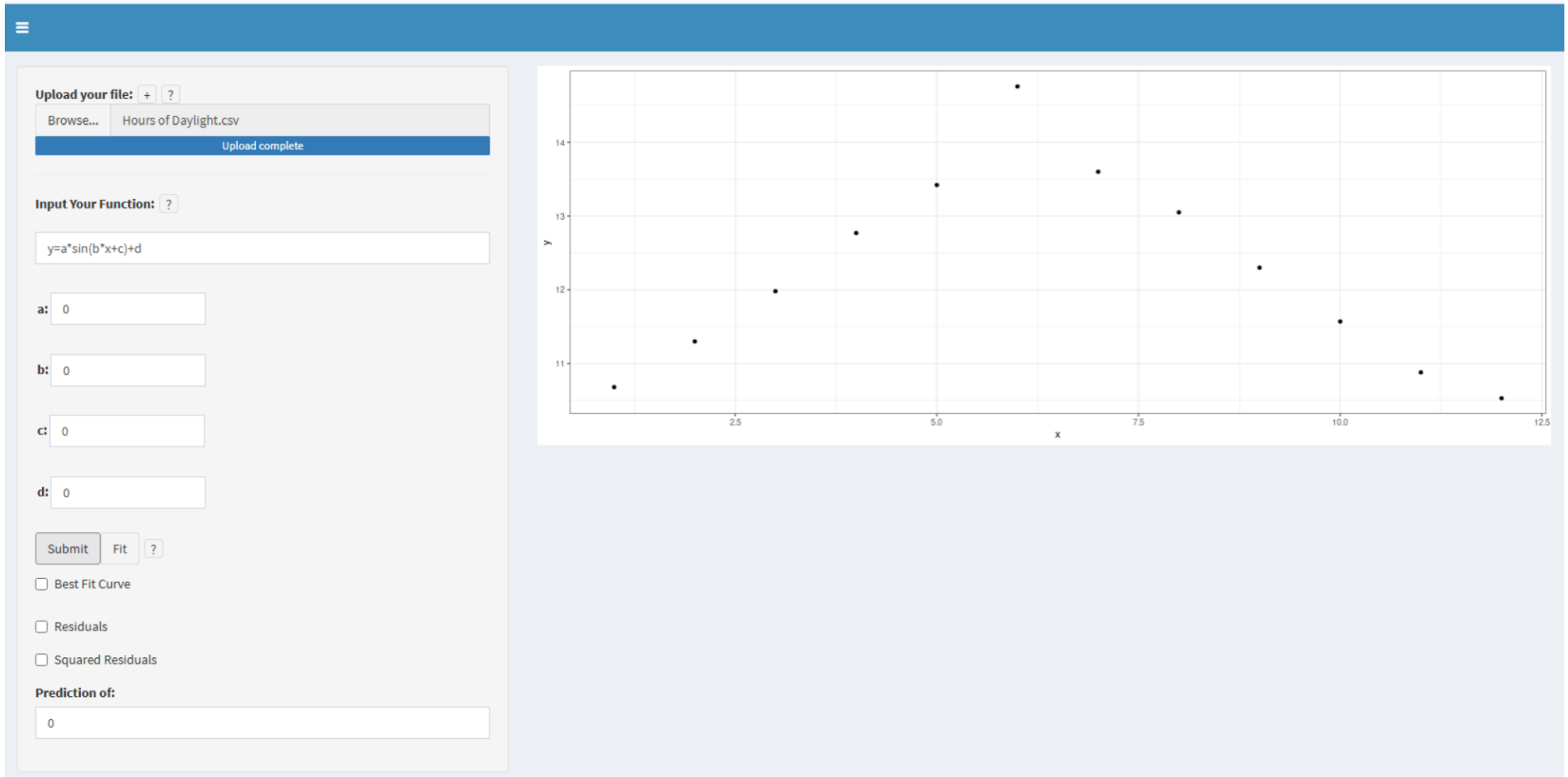
- **General Fitting XY R Shiny tool XY數據一般擬合 R Shiny 工具**
  https://mathmodelcuhk.shinyapps.io/general-fitting/

- 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

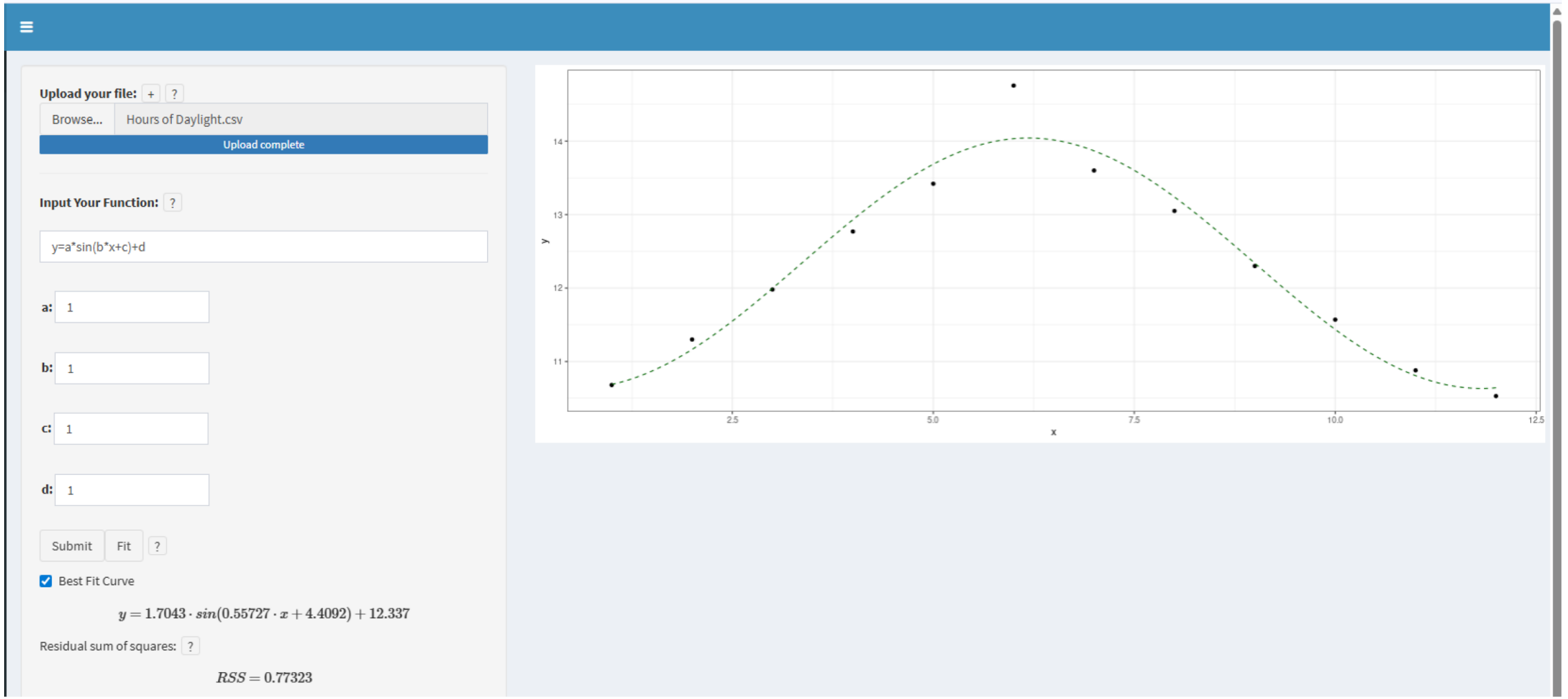
General Fitting XY R Shiny tool XY數據一般擬合 R Shiny 工具

- 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



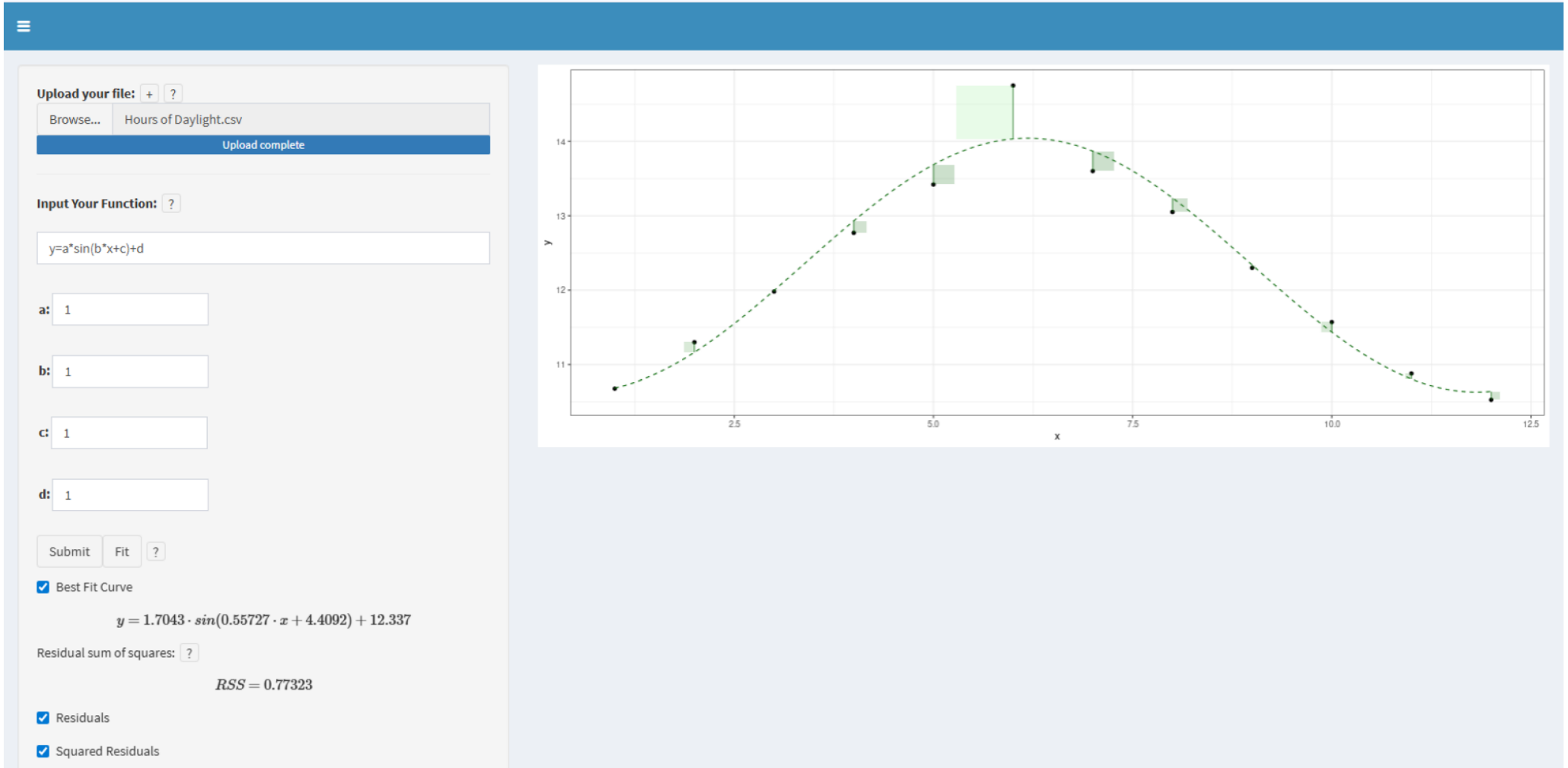
General Fitting XY R Shiny tool XY數據一般擬合 R Shiny 工具

- 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

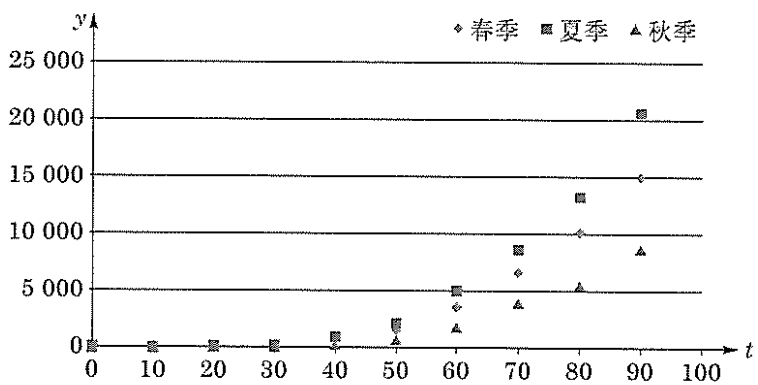
(选择性必修·第三册)

(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

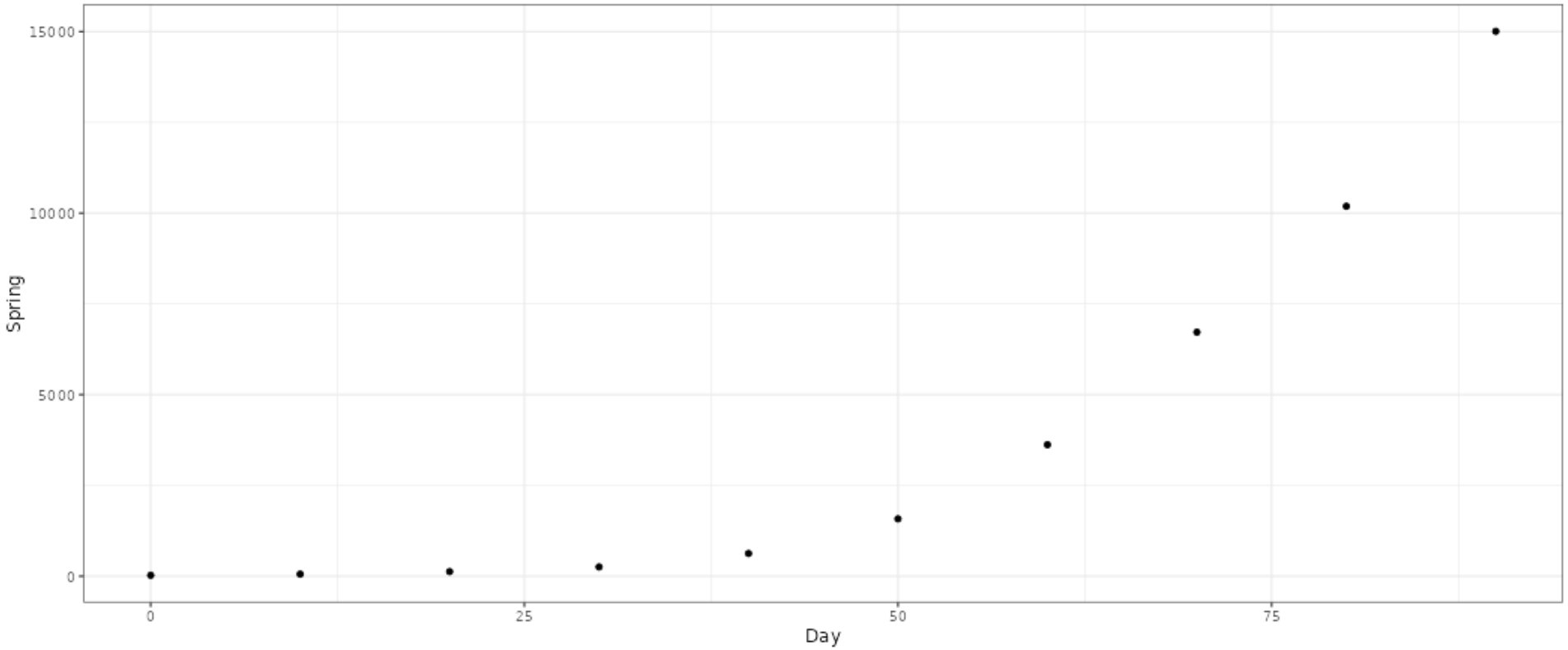
注：表中数据为 30 株水葫芦调查结果的平均值，其中调查时段 0 代表水葫芦接入时的初始值. 数据源于华南农业大学冯熠荣(2003)的《水葫芦种群生态控制的基础研究》.



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: ?

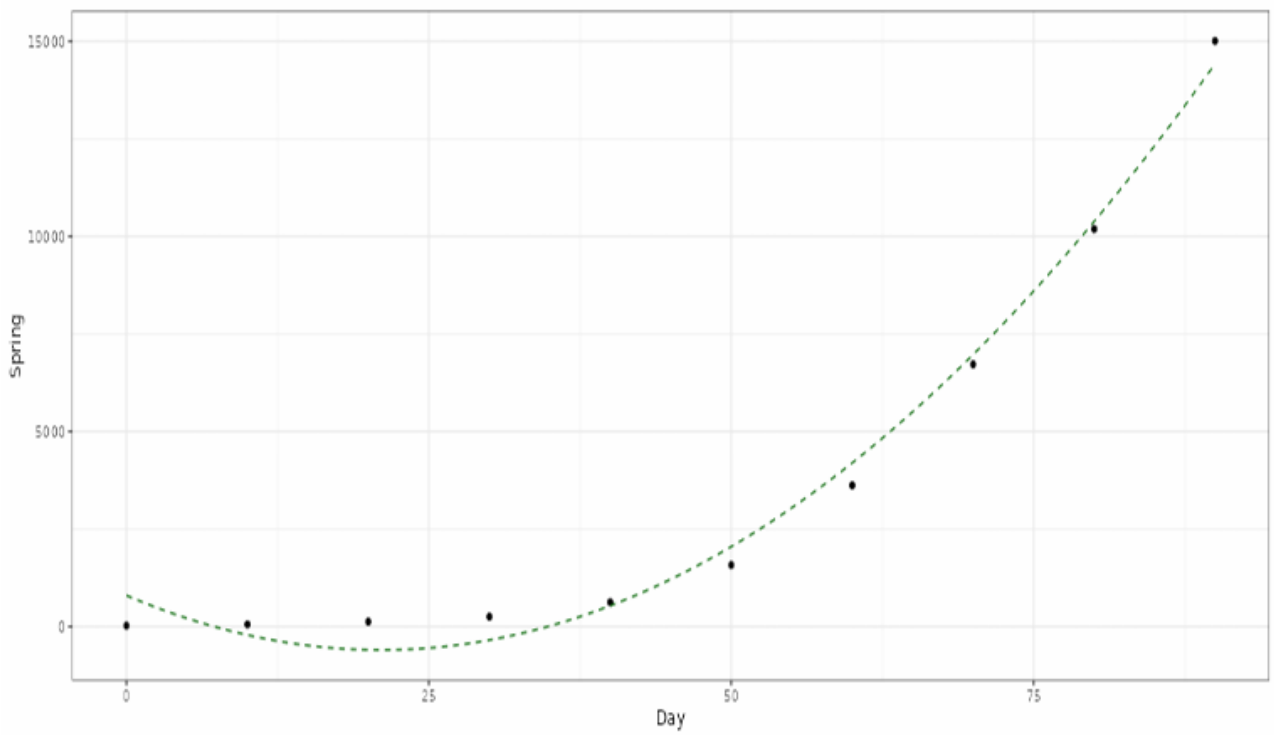
$$RSS = 2.56 \cdot 10^6$$

☐ Residuals

☐ Squared Residuals

Prediction of:

0

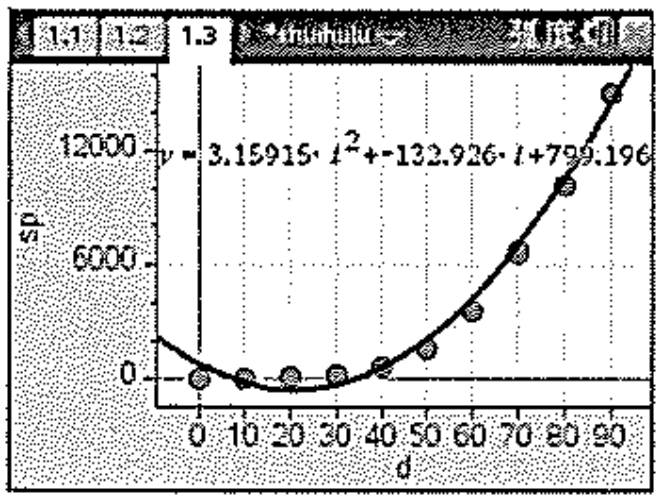


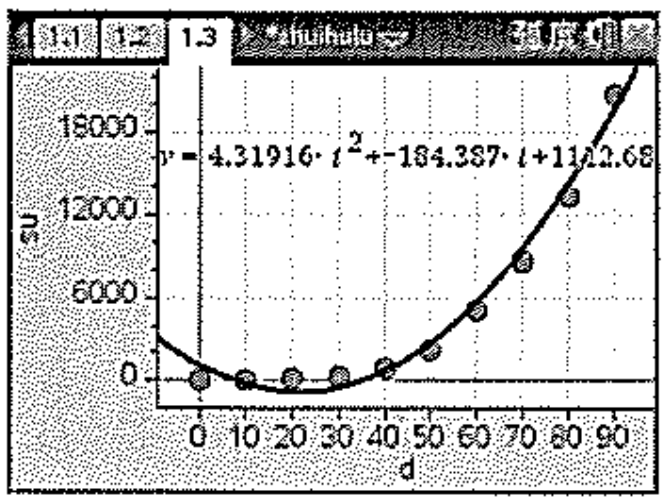
Exercise: the growth of common water hyacinth 水葫蘆的生長

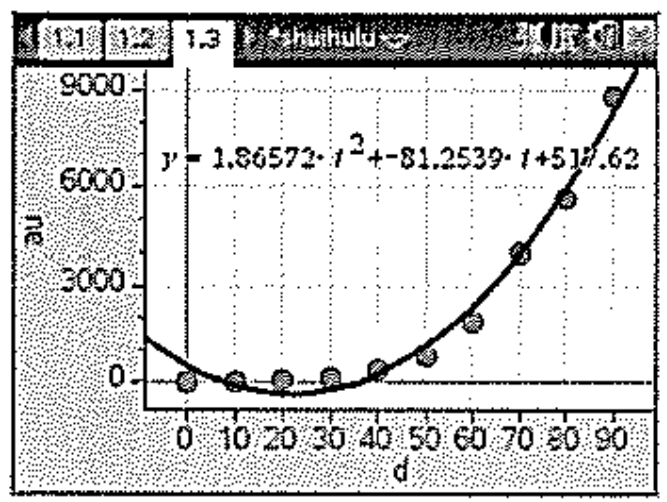
春季：$y=799.196-132.926t+3.159\,15t^2$，

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

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



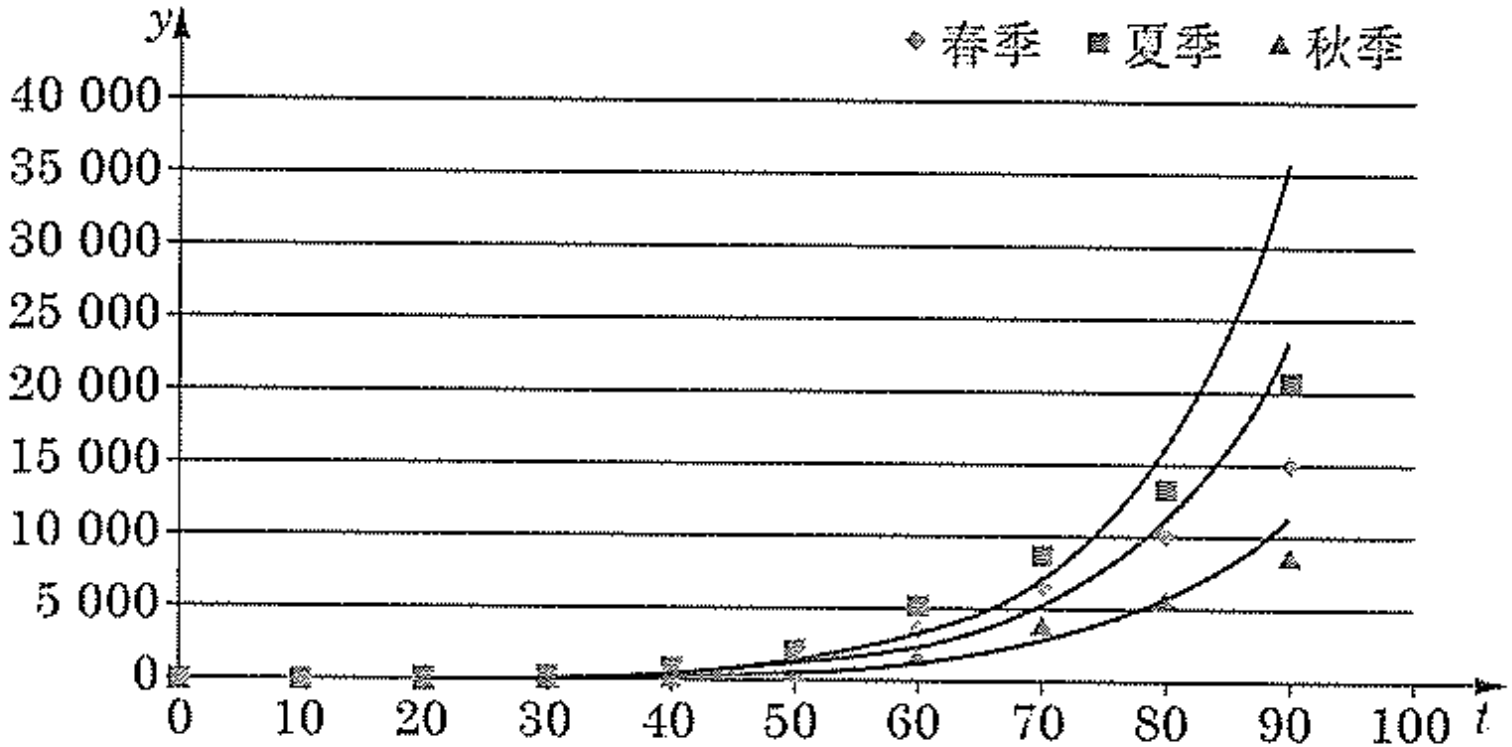




Exercise: the growth of common water hyacinth 水葫蘆的生長

- **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: ?

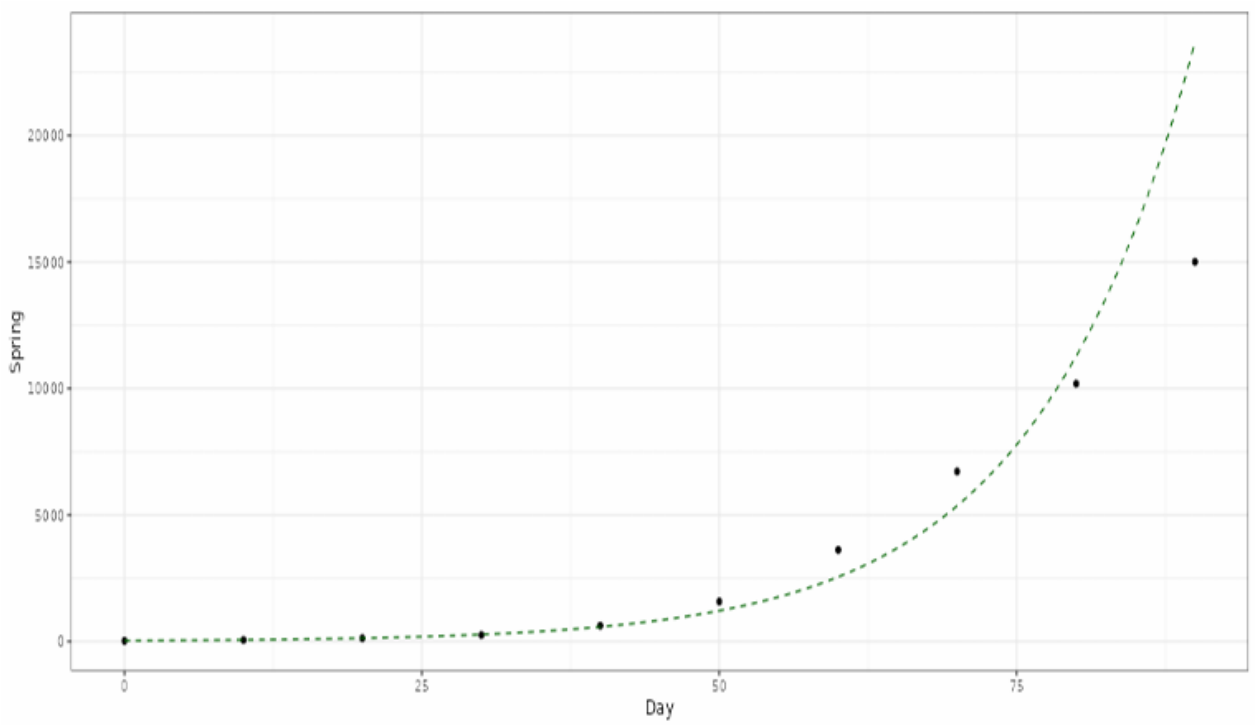
$$RSS = 7.98 \cdot 10^7$$

☐ 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！

Modelling Using Multiple Regression 多元迴歸

- What if there are multiple factors to be considered? 如果有多個因素需要考慮怎麼辦？
- How to determine whether a factor is important? 如何判斷一個因素是否重要？
- Example:
  How does advertisement in different media channels (TV, radio, newspaper) affect the product sales?
  不同媒體（電視、廣播、報紙）上的廣告如何影響產品銷售？

TV	radio	newspaper	sales
230.1	37.8	69.2	22.1
44.5	39.3	45.1	10.4
17.2	45.9	69.3	9.3
151.5	41.3	58.5	18.5
180.8	10.8	58.4	12.9
8.7	48.9	75.0	7.2

- **Multiple regression**: we can consider
  多元迴歸：我們可以考慮

$$y = f(x_1, x_2, \ldots, x_d) = a_0 + a_1 x_1 + a_2 x_2 + \cdots + a_d x_d$$

and minimize 並最小化

$$\text{RSS} = \sum_{i=1}^{n} (y_i - a_0 - a_1 x_{i,1} - \cdots - a_d x_{i,d})^2$$

Modelling Using Multiple Regression 多元迴歸

- Approaches 方法:
  - Consider each variable separately
    分別考慮每個變數

$$\text{sales} = \beta_0 + \beta_1 \text{TV}$$

  - Including the variables 包括所有變數

$$\text{sales} = \beta_0 + \beta_1 \text{TV} + \beta_2 \text{radio} + \beta_3 \text{newpaper}$$

TV	radio	newspaper	sales
230.1	37.8	69.2	22.1
44.5	39.3	45.1	10.4
17.2	45.9	69.3	9.3
151.5	41.3	58.5	18.5
180.8	10.8	58.4	12.9
8.7	48.9	75.0	7.2

  - Including all variables and their interaction terms 包括所有變數及它們互相的影響

$$\begin{aligned}\text{sales} = \beta_0 &+ \beta_1 \text{TV} + \beta_2 \text{radio} + \beta_3 \text{newpaper} \\ &+ \beta_4 (\text{TV} \cdot \text{radio}) + \beta_5 (\text{TV} \cdot \text{newpaper}) + \beta_6 (\text{radio} \cdot \text{newpaper})\end{aligned}$$

  - And many more possibilities 還有更多可能性

- Detailed description 詳細說明: http://mathcal.math.cuhk.edu.hk:7541/

Modelling Using Multiple Regression 多元迴歸

- **Interpretation of importance of different factors via multiple regression**:
  透過多元迴歸分析不同因素的重要性：

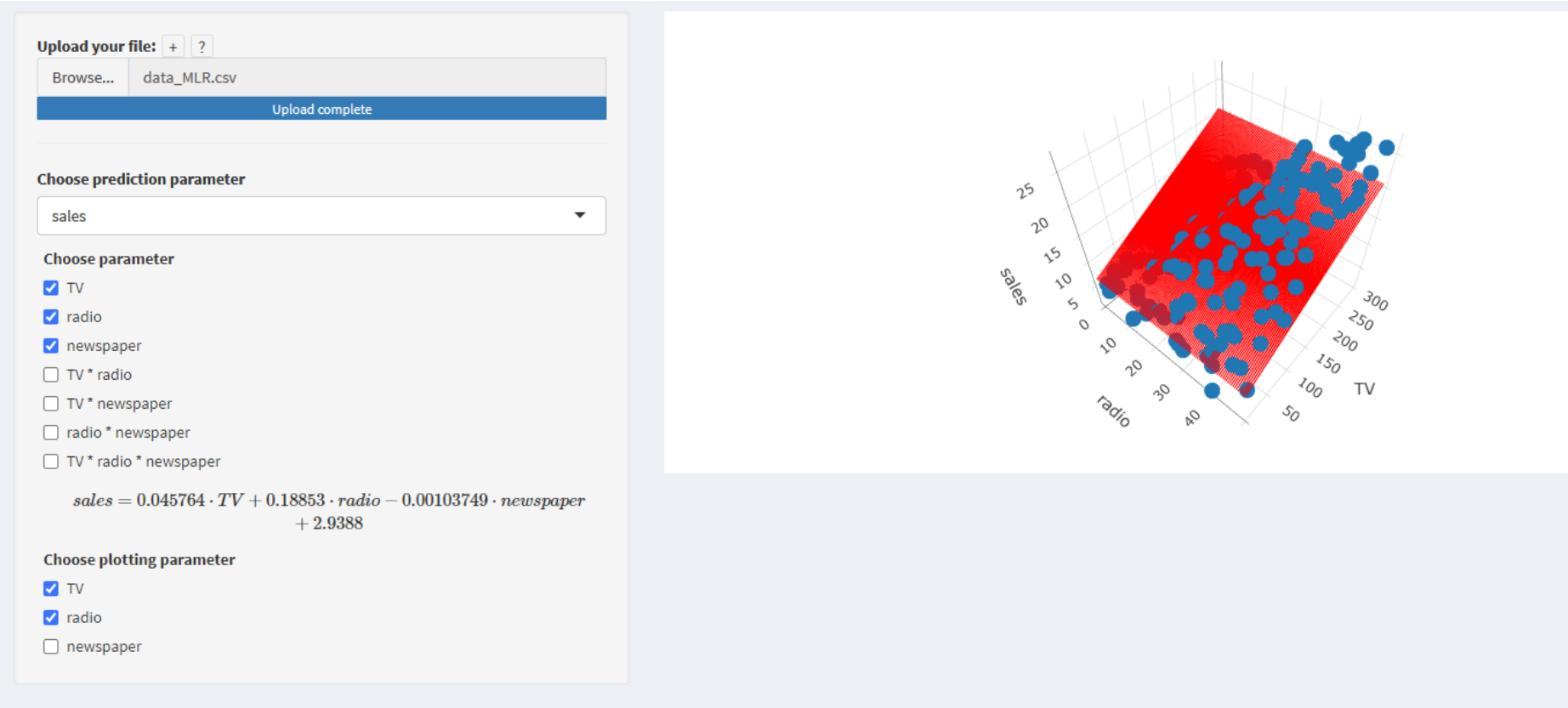
  - We first standardize all variables (i.e., transform the data values for each variable to have a mean of 0 and a standard deviation of 1) and then solve for the best-fit multiple regression model
    首先，我們將所有變數進行標準化處理（即將每個變數的資料值轉換為平均值為 0、標準差為 1 的平均值），然後求解最佳擬合的多元迴歸模型
  - The magnitude of the resulting coefficients can indicate the relative importance of the variables 所得系數的大小可以表示變數的相對重要性
  - Can be used for feature selection or model simplification
    可用於特徵選擇或模型簡化

- **Multi-Regression R Shiny Tool 多元迴歸 R Shiny 工具**
  https://mathmodelcuhk.shinyapps.io/multi-regression/

Multi-Regression R Shiny Tool 多元線性迴歸 R Shiny工具

- https://mathmodelcuhk.shinyapps.io/multi-regression/
  - Performing multiple linear regression 進行多元線性迴歸
  - Input data format: CSV, XLSX, or TXT file containing the data points (at least 2 columns of data) 輸入數據格式：包含數據點的 CSV、XLSX 或 TXT 檔案（最少 2 列資料）
  - Sample data file 參考數據檔案: [data_MLR.csv]

	A	B	C	D
1	TV	radio	newspape	sales
2	230.1	37.8	69.2	22.1
3	44.5	39.3	45.1	10.4
4	17.2	45.9	69.3	9.3
5	151.5	41.3	58.5	18.5
6	180.8	10.8	58.4	12.9
7	8.7	48.9	75	7.2
8	57.5	32.8	23.5	11.8
9	120.2	19.6	11.6	13.2
10	8.6	2.1	1	4.8
11	199.8	2.6	21.2	10.6
12	66.1	5.8	24.2	8.6
13	214.7	24	4	17.4
14	23.8	35.1	65.9	9.2
15	97.5	7.6	7.2	9.7
16	204.1	32.9	46	19
17	195.4	47.7	52.9	22.4
18	67.8	36.6	114	12.5



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

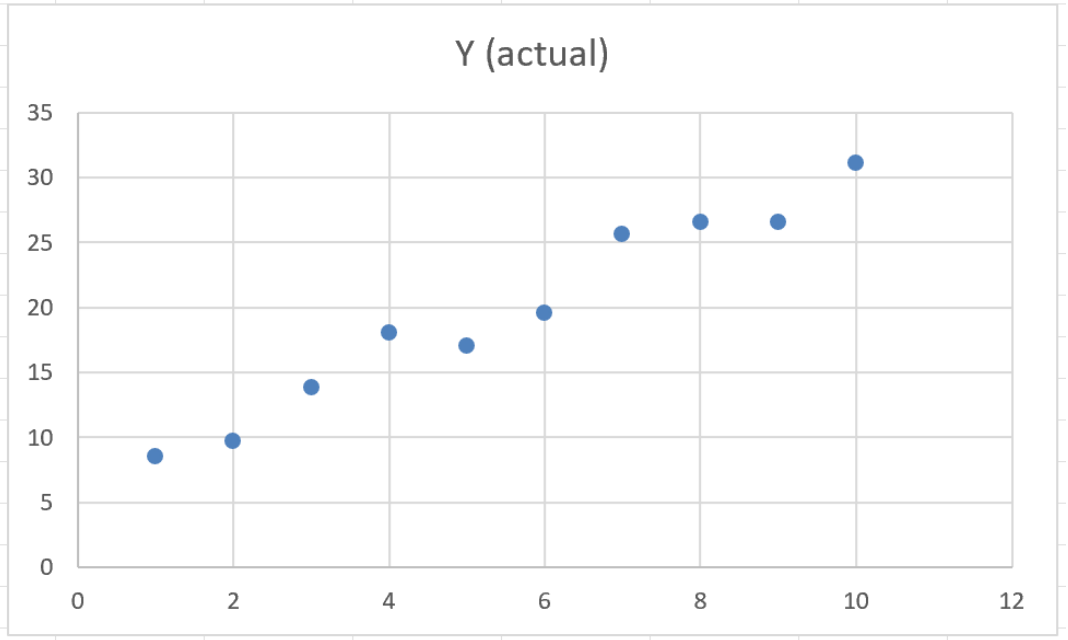
- Many functionalities useful for math modelling 許多對數學建模有用的功能：
  - Statistics 統計分析
  - Regression analysis 迴歸分析
  - Data visualization (scatter plots, bar charts, …) 數據視覺化（散點圖、棒形圖等）

Paste | B I U | Clipboard | Font | G15 | fx

	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 | fx

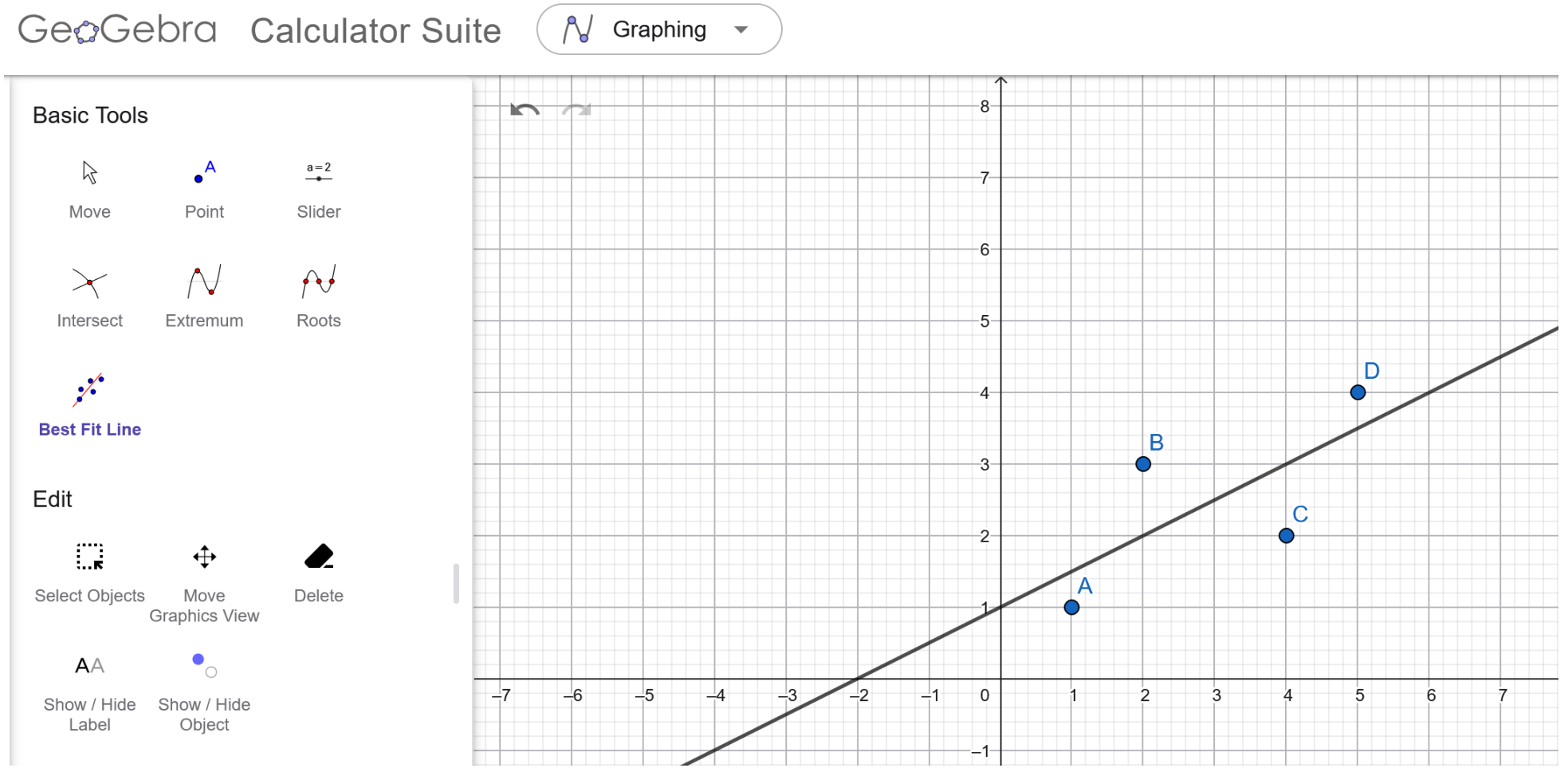
	A	B	C
1	Metric	Value	
2	Mean X	5.5	
3	Mean Y	19.64612	
4	Std X	3.02765	
5	Std Y	7.664861	
6	Correlation	0.982058	
7	Slope (b)	2.486199	
8	Intercept (a)	5.972026	
9			



GeoGebra for math modelling 用於數學建模

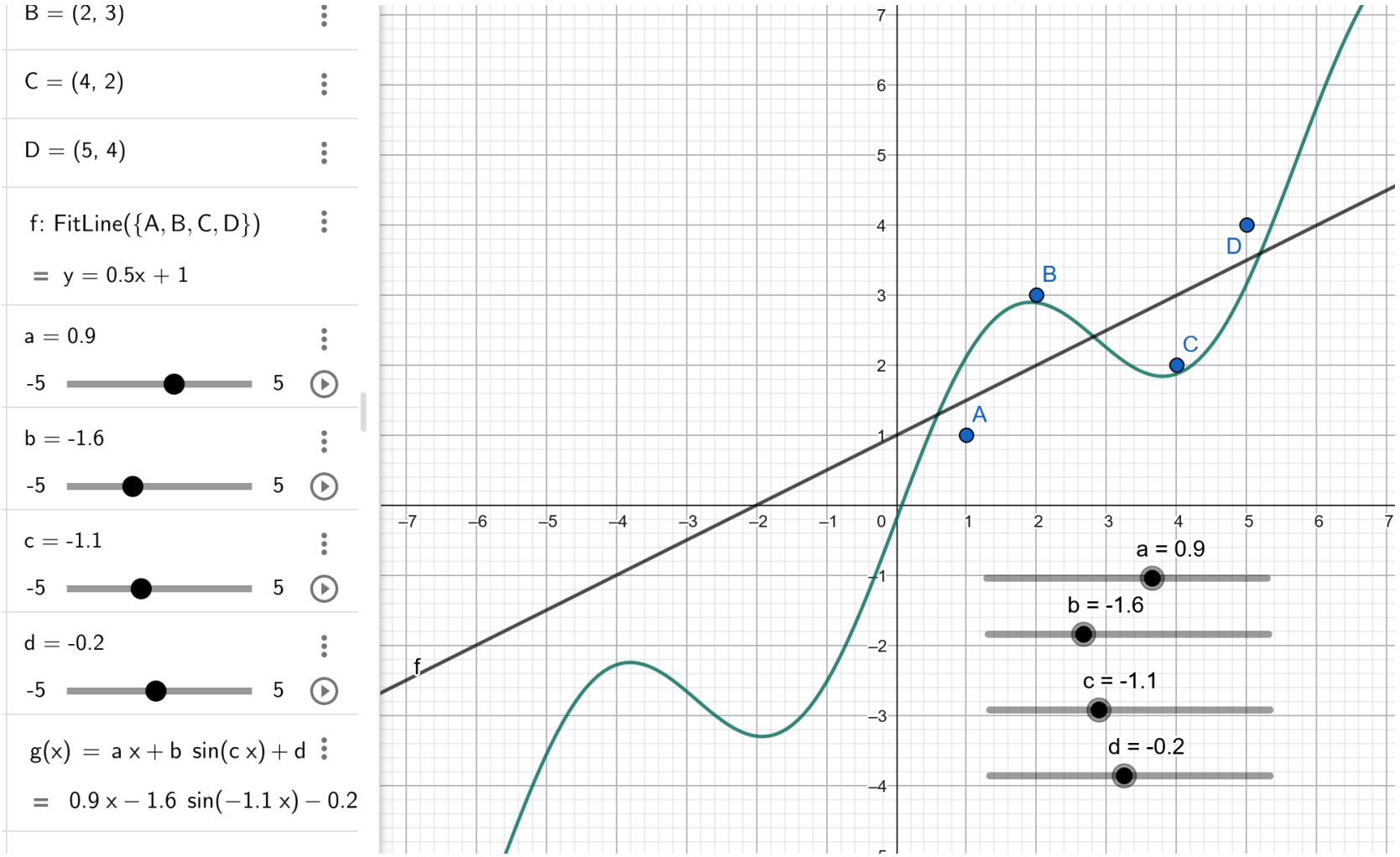
- **GeoGebra** (https://www.geogebra.org/)

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



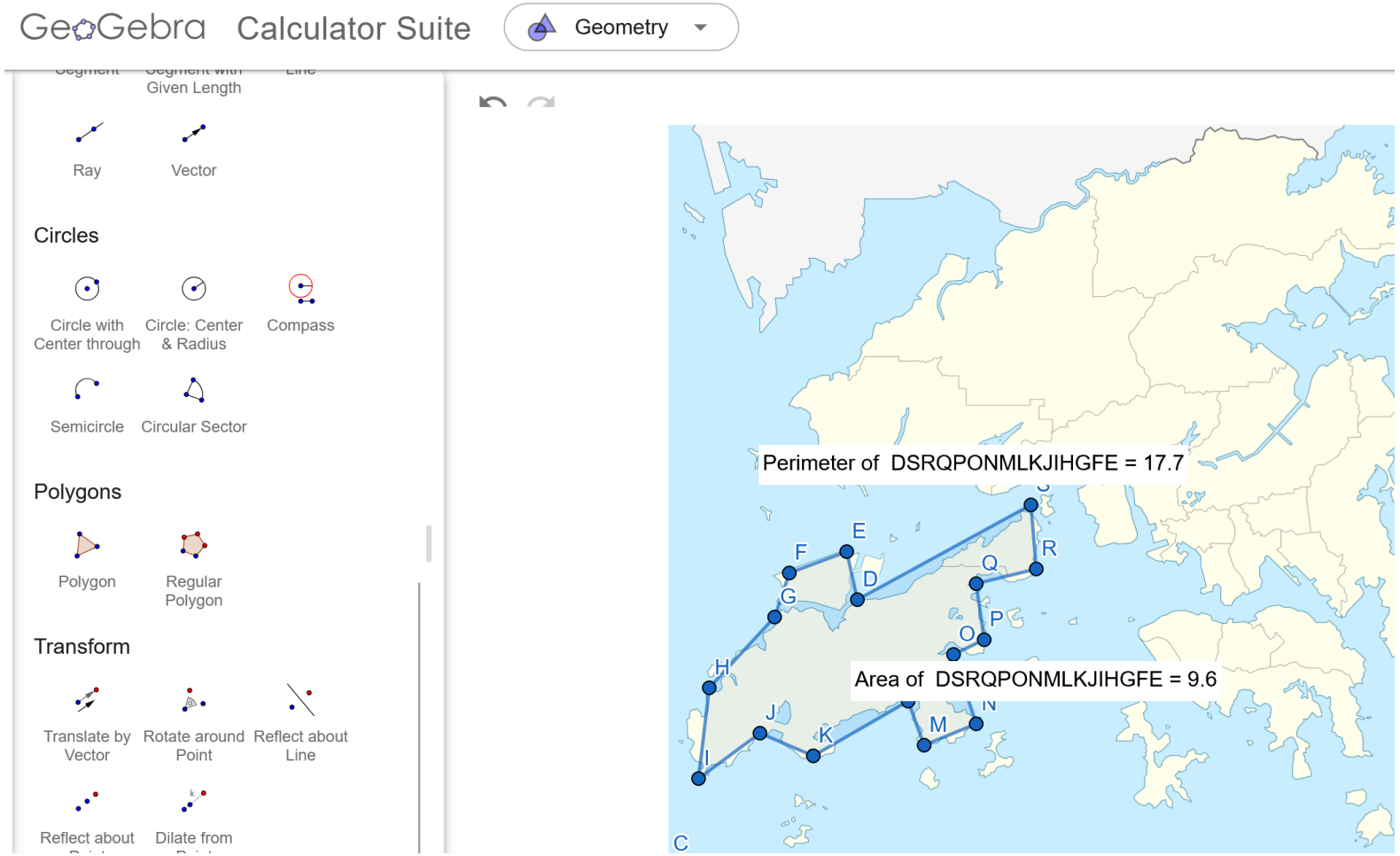
GeoGebra for math modelling 用於數學建模

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



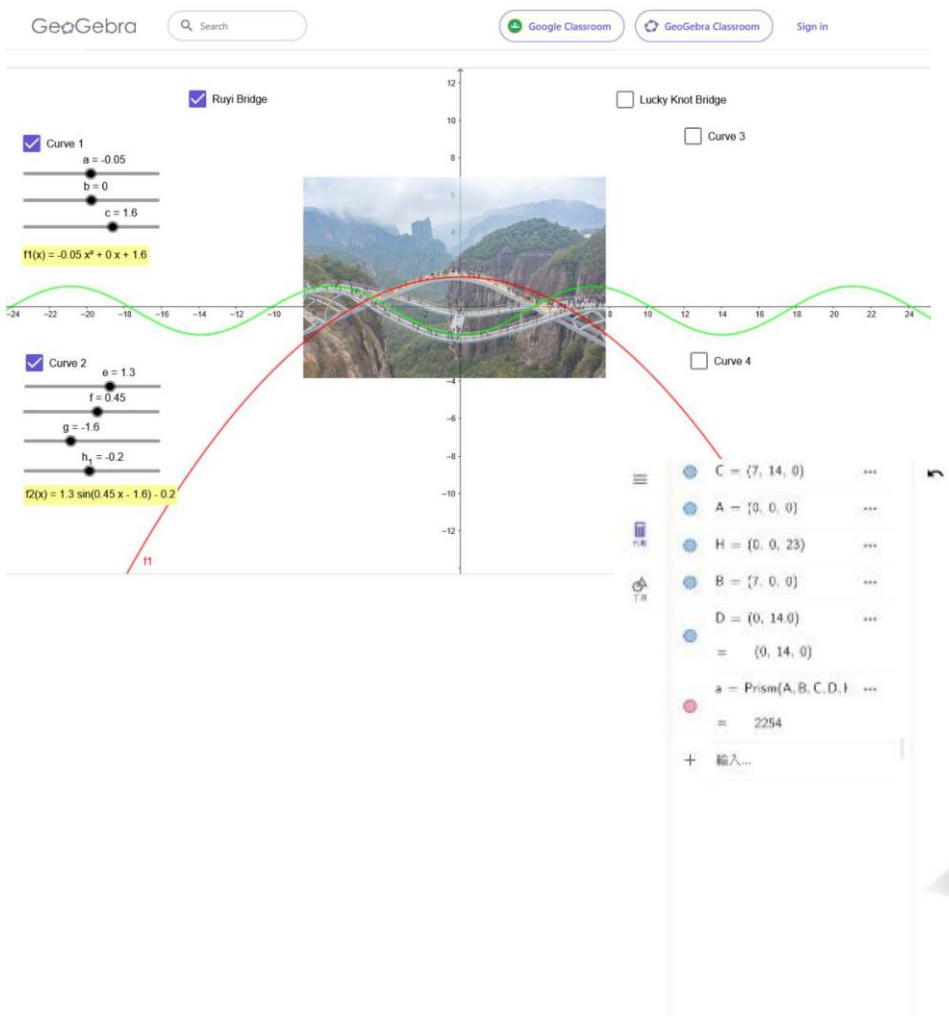
GeoGebra for math modelling 用於數學建模

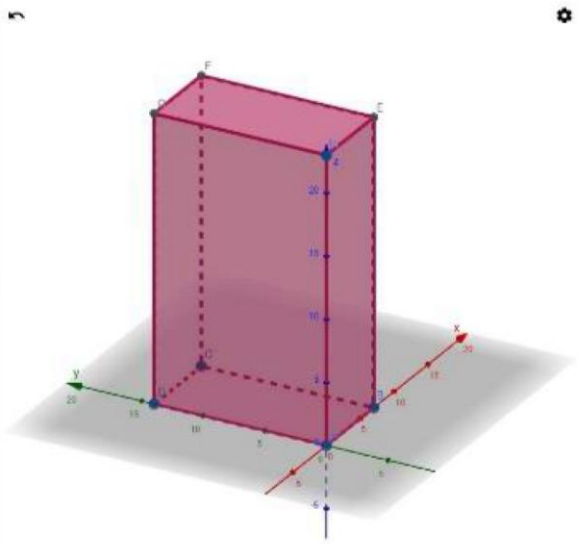
- Geometric modelling 幾何建模



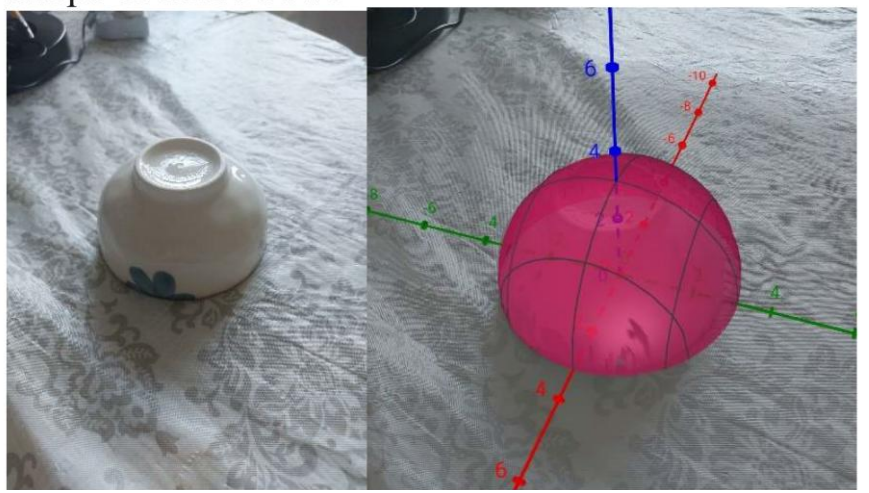
GeoGebra for math modelling 用於數學建模

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

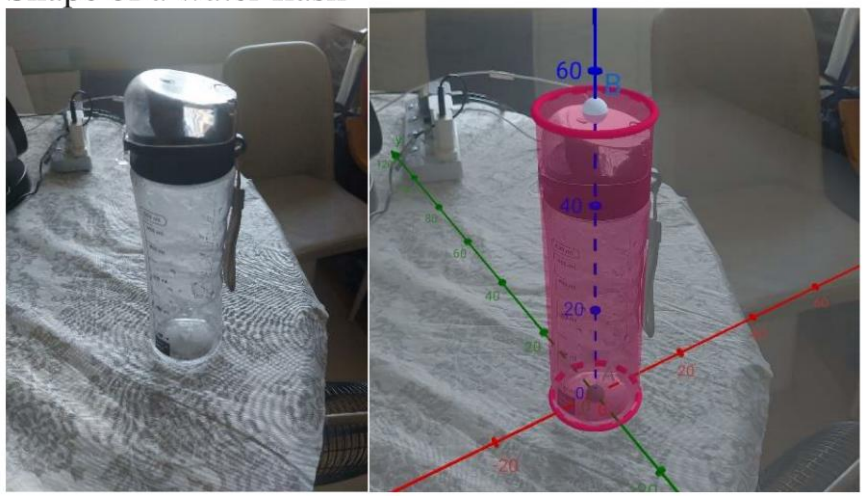




Shape of a rice bowl



Shape of a water flask



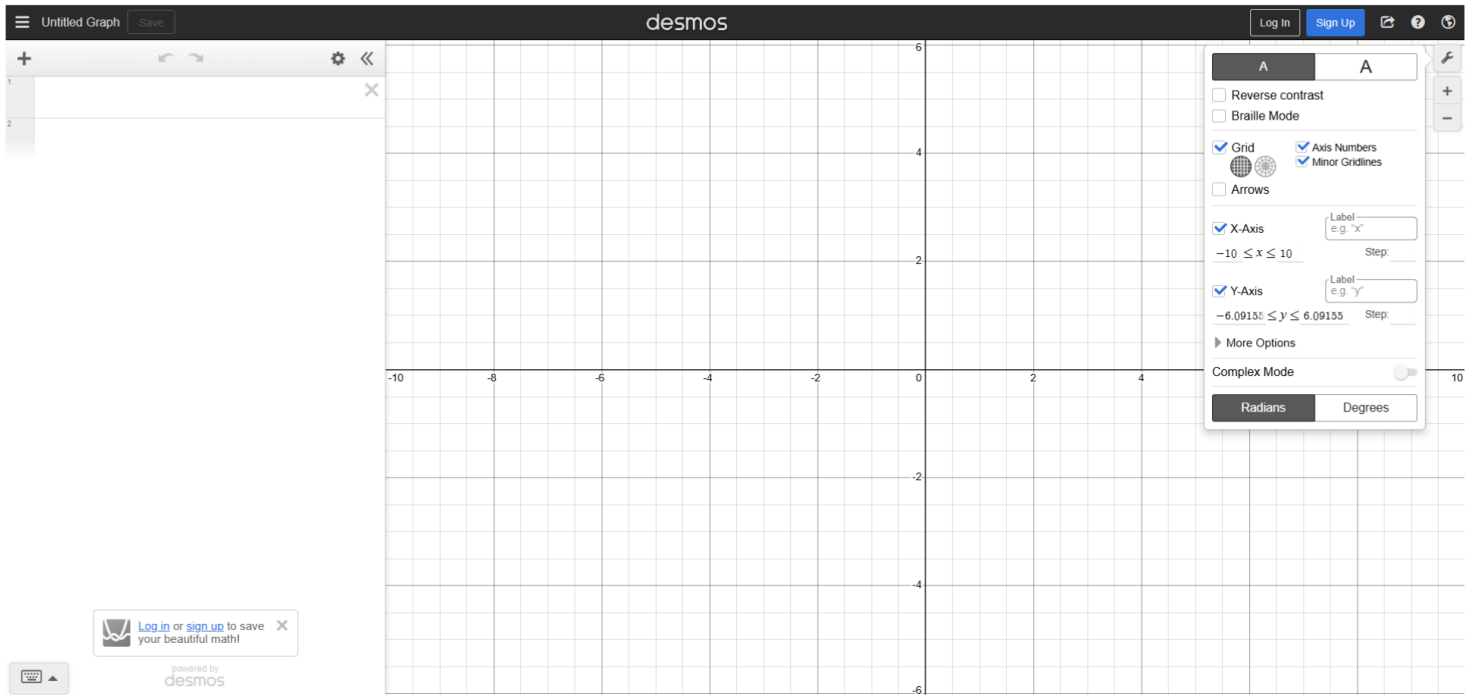
Desmos for math modelling 用於數學建模

- **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。

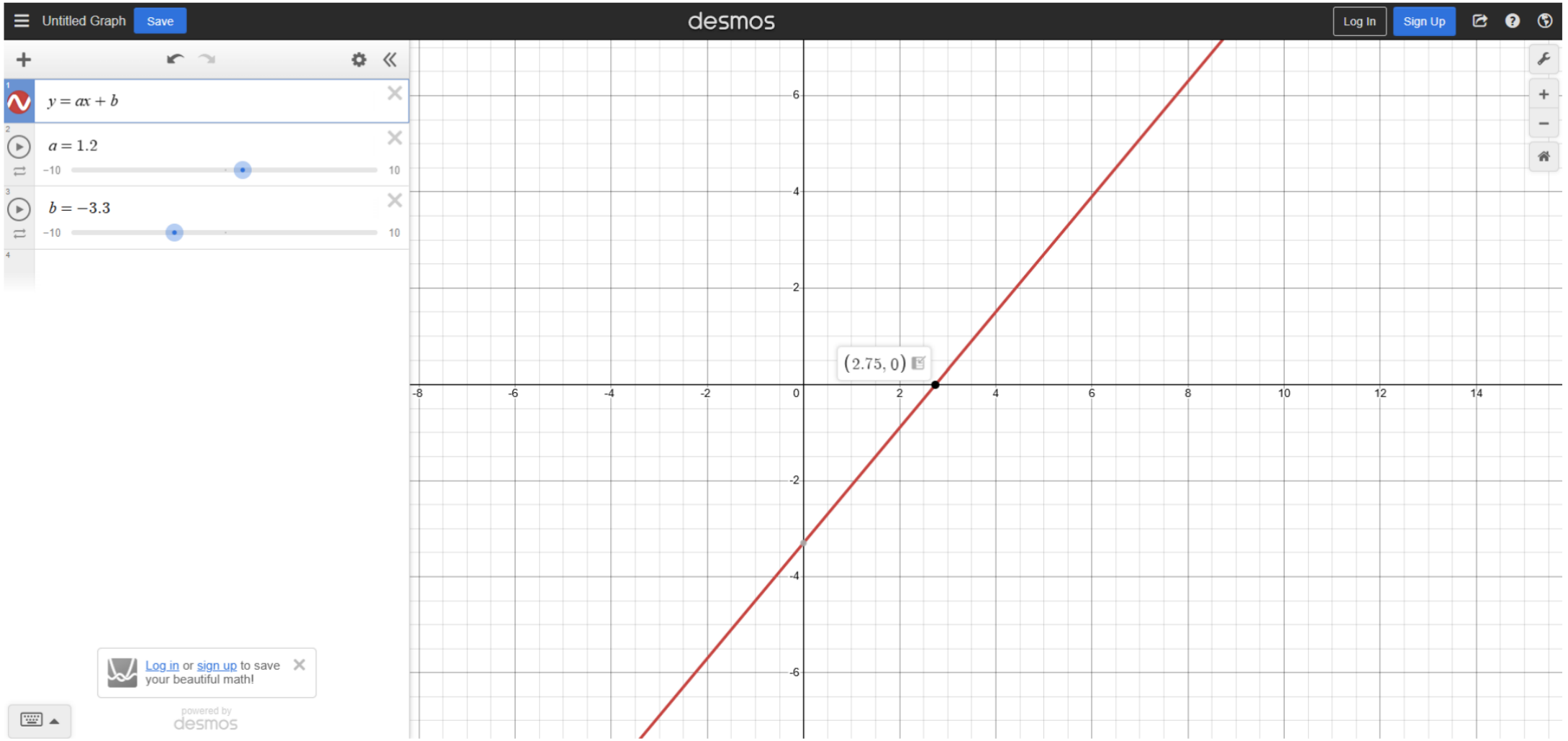
  - Middle: the graph window that visualizes the equation.
    中間：用於可視化方程式的圖形視窗。

  - 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。



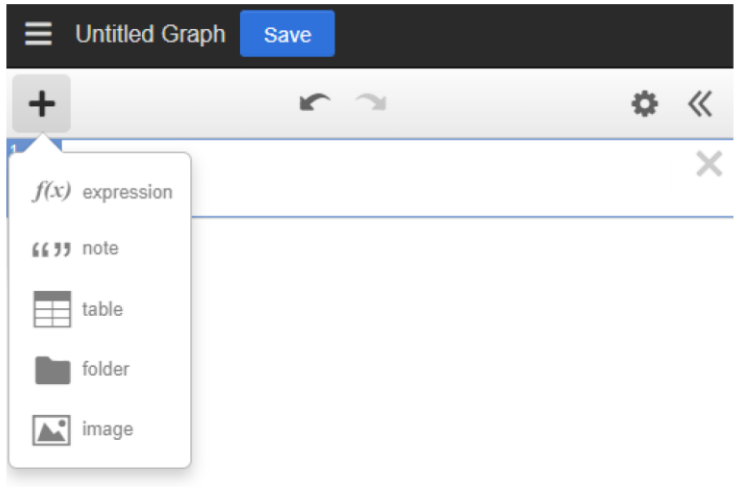
Desmos for math modelling 用於數學建模

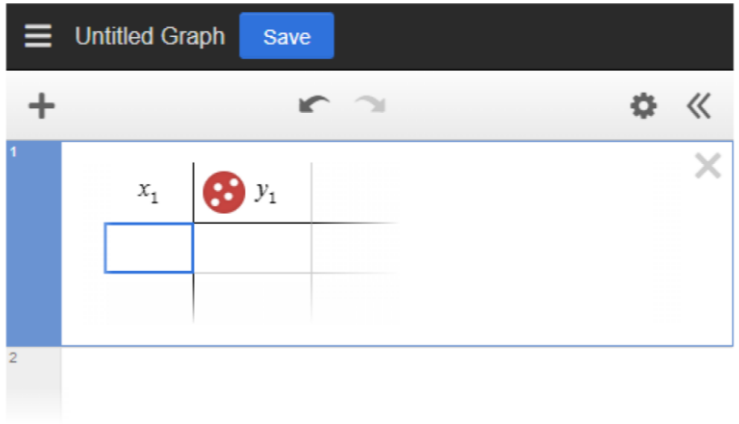
Adding curves and points 增加曲線和點

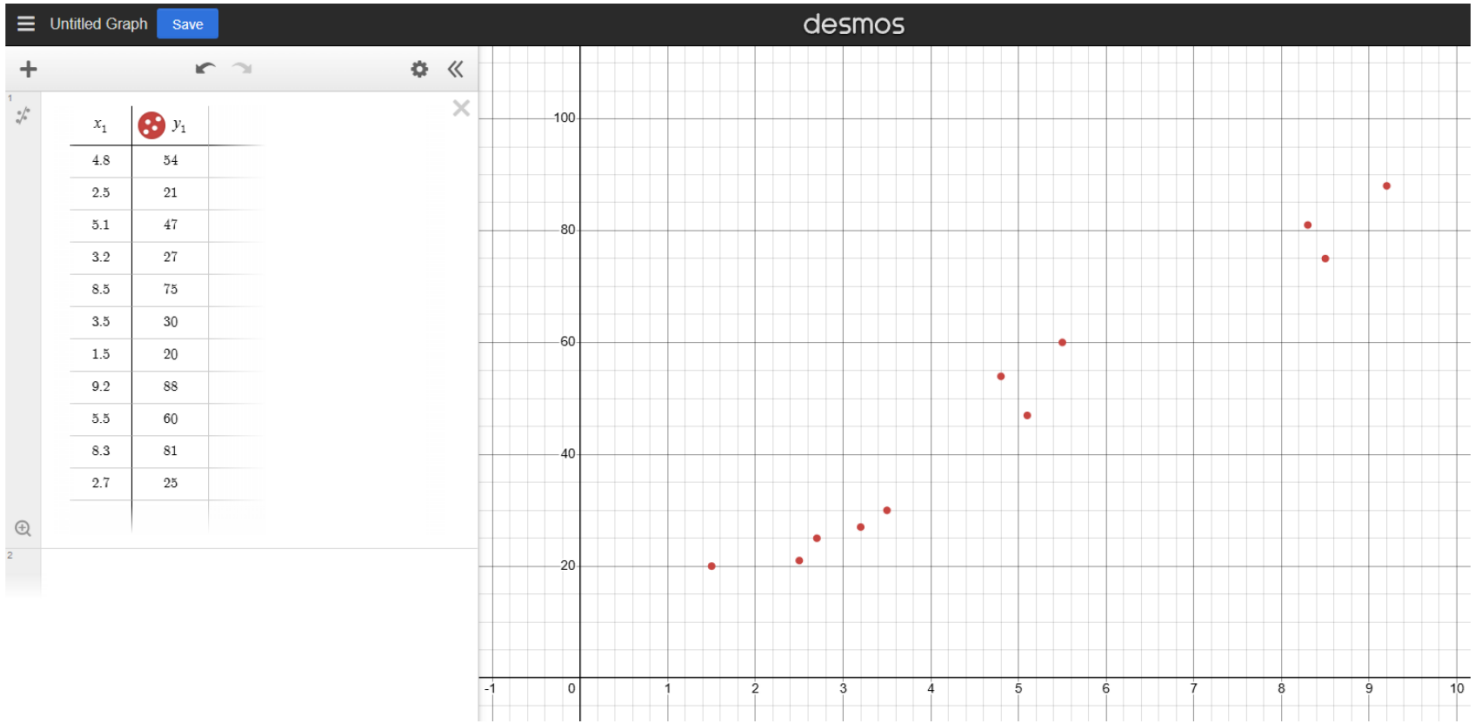


Desmos for math modelling 用於數學建模

Adding data points 新增數據點

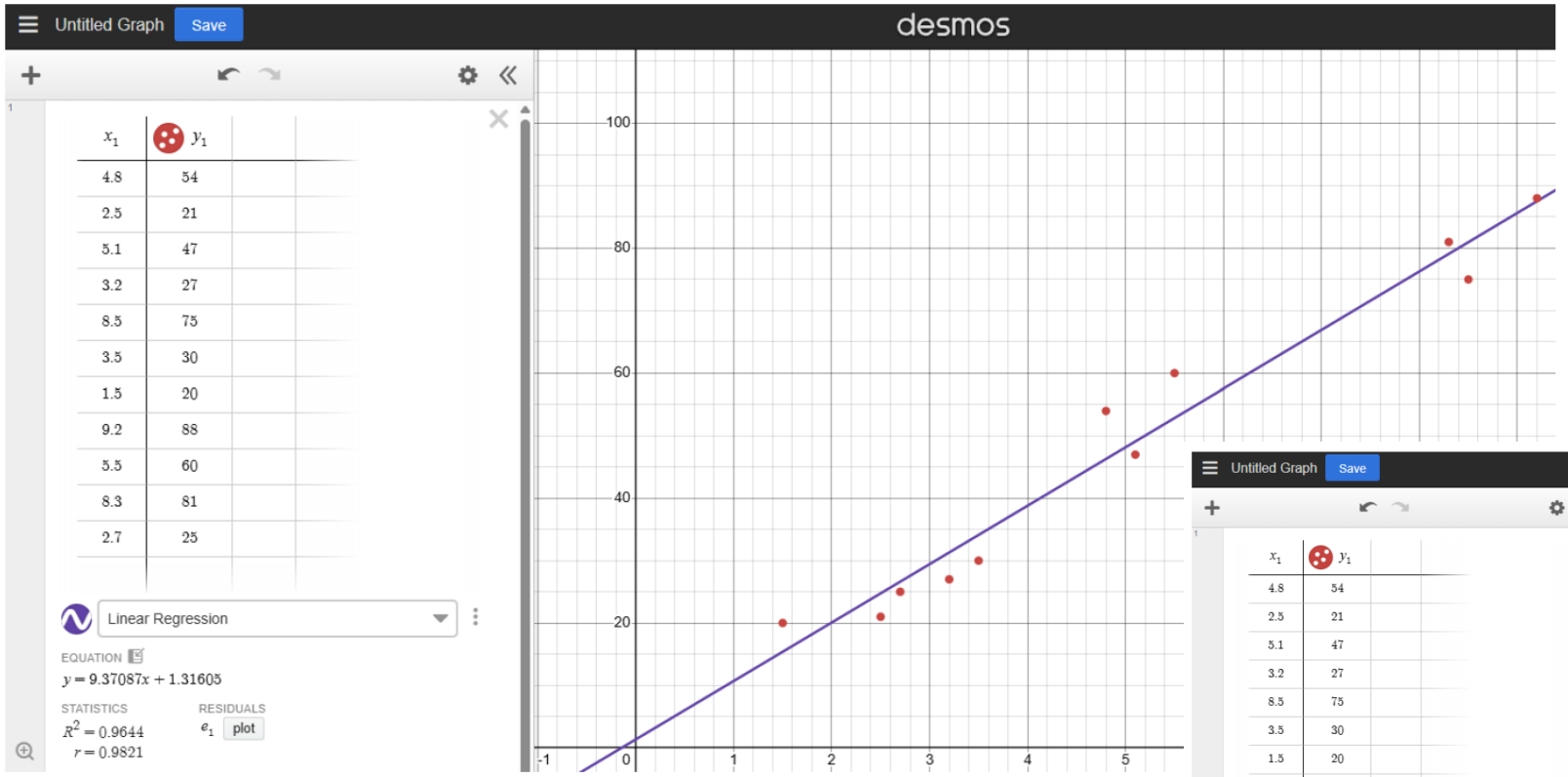


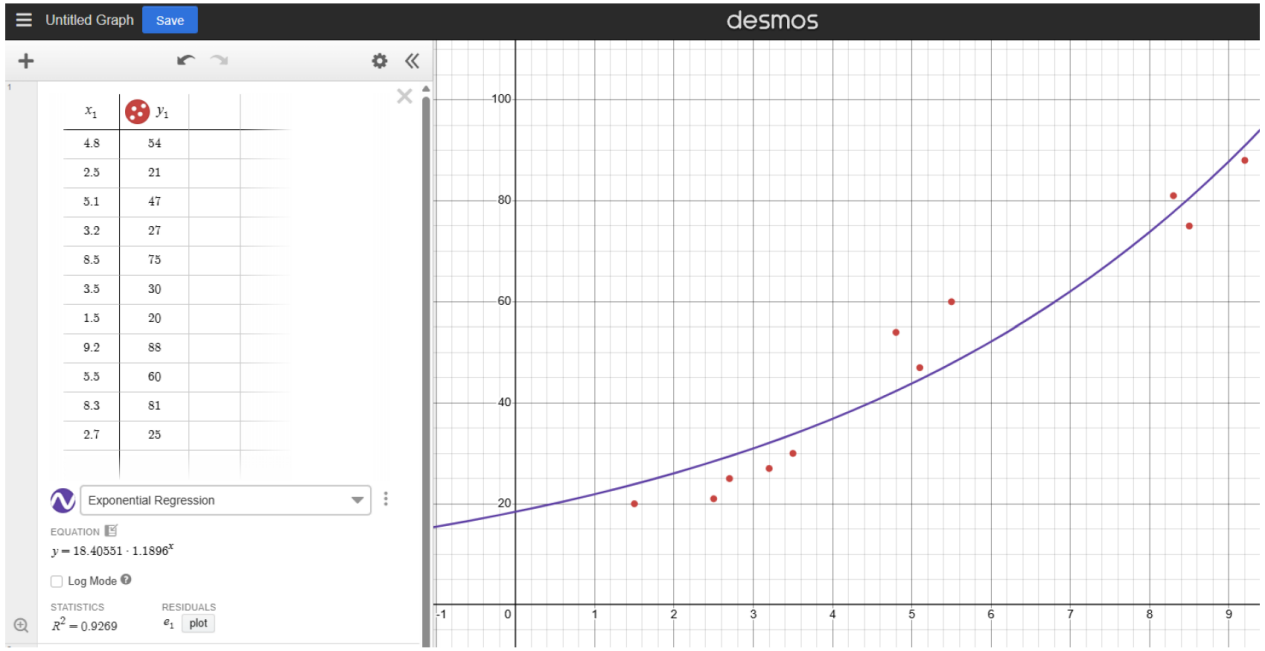




Desmos for math modelling 用於數學建模

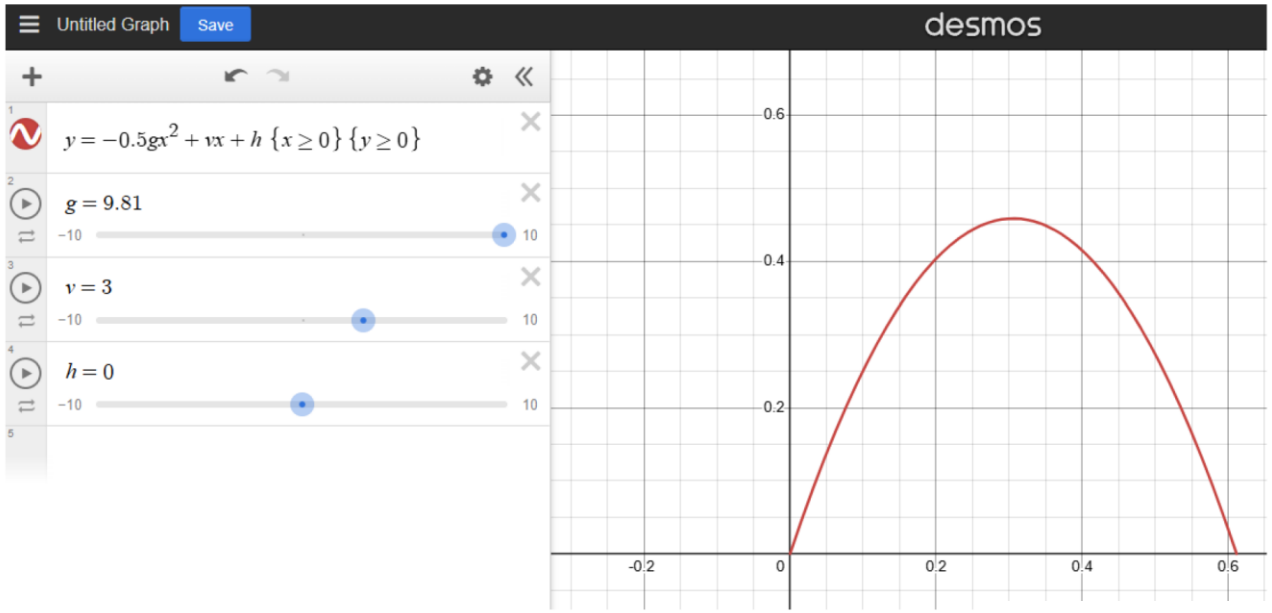
Choosing different regression functions 選擇不同的迴歸函數：



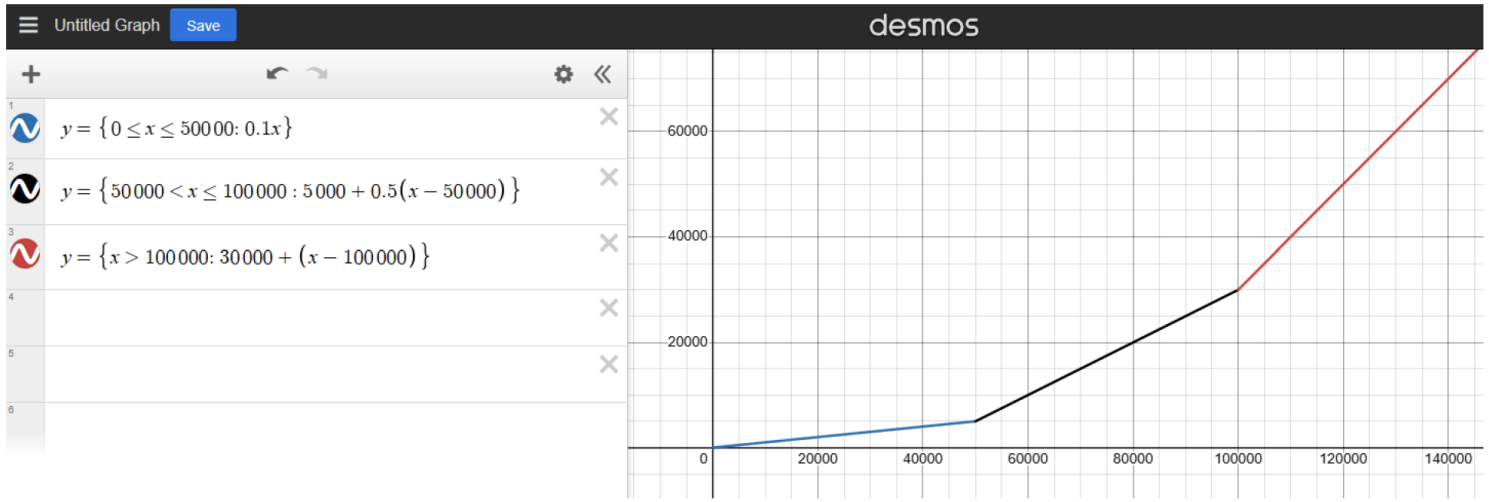


Desmos for math modelling 用於數學建模

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。

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

摘要

我方認爲社區型耐甲氧西林金黃葡萄球菌感染和猩紅熱是在不久將來在香港需要特別關注的傳染病，根據社區型耐甲氧西林金黃葡萄球菌感染和猩紅熱的研究價值比對後，我方選取社區型耐甲氧西林金黃葡萄球菌感染建立數學模型與預測其於不同情況下在 2026 年和 2036 年的香港病例總數，首先我方試用直綫方程後發現數據不貼合趨勢綫，其次我方試用四次方程後發現數據過度擬合趨勢綫，不符合常理，最後我方試用正弦方程後發現最適合我方進行預測，反思中發現其不符合常理，最終選擇將直綫方程與正弦方程相結合，發現其週期性及具上升趨勢，預測出社區型耐甲氧西林金黃葡萄球菌感染的香港病例總數在 2026 年時，一般情況的數目為 735、高爆發情況的數目為 1478、低爆發情況為 543；在 2036 年時，一般情況的數目為 1170、高爆發情況的數目為 1661、低爆發情況的數目為 725。

Sample report (1) – Problem introduction and analysis
參考作品 (1) – 題目簡介及分析

1.1 問題中的定義

傳染病

傳染病是指一些具傳播性的疾病。傳染病是由於病原體入侵人體繁殖或產生的毒素，破壞身體細胞及其功能所致，嚴重時會引致死亡。形成傳染病的傳播四項主要因素有：病原體、傳染源、傳播途徑 及宿主，稱之為傳染鏈。引致感染的病原體是一些很細小的微生物，如細菌、病毒、真菌（黴菌）和寄生蟲。(衛生防護中心 [1])

病原體

引致感染的病原體是一些很細小的微生物，如細菌、病毒、真菌（黴菌）和寄生蟲。(衛生防護中心 [1])

傳染源

任何讓病原體可以存活、寄居和繁殖的環境，如受感染的人類（例如病人、帶菌者和隱性感染患者）、禽畜、昆蟲和泥土。人類可透過接觸傳染源感染病原體。 (衛生防護中心 [1])

1.2 題目分析

爲了探究香港傳染病，我方通過衛生防護中心查看了 2007 年至 2024 年的傳染病感染人數，發現呈報的傳染病統計數字較爲突出的有六種：水痘、社區型耐甲氧西林金黃葡萄球菌感染、2019 冠狀病毒病、食物中毒、猩紅熱、結核病。(衛生防護中心 [3])

據我方調查發現，水痘、2019 冠狀病毒病、結核病均研發出了疫苗，接種率分別為 95%以上 （衛生防護中心 [4]），93.8%（衛生防護中 [5]），95%以上（衛生防護中心 [4]），減少了傳染率，而社區型耐甲氧西林金黃葡萄球菌感染、食物中毒、猩紅熱到目前爲止沒有研發出疫苗。不過食物中毒可以通過醫療乾預傳染率，具局部地區性未造成廣汎影響前控制，傳播途徑較爲可控。而社區型耐甲氧西林金黃葡萄球菌感染、猩紅熱的危害性較大，可能導致並發症，影響壽命和健康，所以我方認爲在不久將來香港特別關注的傳染病分別為社區型耐甲氧西林金黃葡萄球菌感染、猩紅熱。

Sample report (1) – Model assumption and data collection
參考作品 (1) – 模型假設及數據收集

2.1 假設

我們在計算數學模型前，先作以下假設。

1.沒有新的傳染病出現

我方假設 2025 年到 2036 年沒有新的和變異傳染病出現

2. 社區型耐甲氧西林金黃葡萄球菌(CA-MRSA)穩定不變

我方假設社區型耐甲氧西林金黃葡萄球菌(CA-MRSA)不會發生變異，遺傳物質 不發生改變

3.人口總數目不變

我方假設 2025 年到 2036 年人口總數目不變(7531800)(二零二四年年中人口數字)

4.年齡比例正常

我方假設香港年齡結構不變 ,出生數目(367000)(政府統計處[1]) 、死亡數目(514000)(政府統計處[1])保持不變以及香港人口年齡比例結構保持正常(1.5:6.5:2)(人口金字塔)

2.3 收集的數據

根據以上解決問題方向的規劃，我們在政府的衛生防護中心收集了由 2007 年至 2024 年(截至 12 月份)的感染社區型耐甲氧西林金黃葡萄球菌(CA-MRSA)的感染人數。

年份	2007	2008	2009	2010	2011	2012	2013	2014	2015
感染人數	173	282	368	495	624	815	990	997	1046

年份	2016	2017	2018	2019	2020	2021	2022	2023	2024
感染人數	1164	1258	1218	1236	813	582	416	469	567

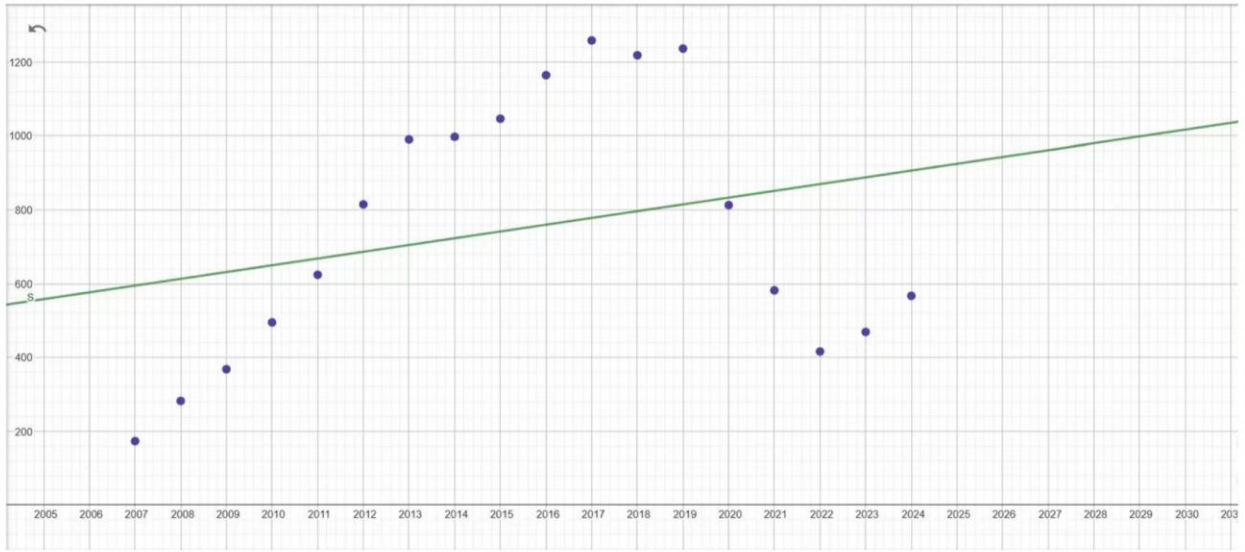
(衛生防護中心，2024)

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

3. 進行數學模型計算

3.1 香港的社區型耐甲氧西林金黃葡萄球菌感染者數目的數據的數學模型

根據以上香港的社區型耐甲氧西林金黃葡萄球菌感染者數目的數據，利用計算機套件的趨勢綫，我方得出以下直綫方程，四次方程及正弦的方程圖像：



以直綫方程模擬數據趨勢的方程為：$y = ax + b$

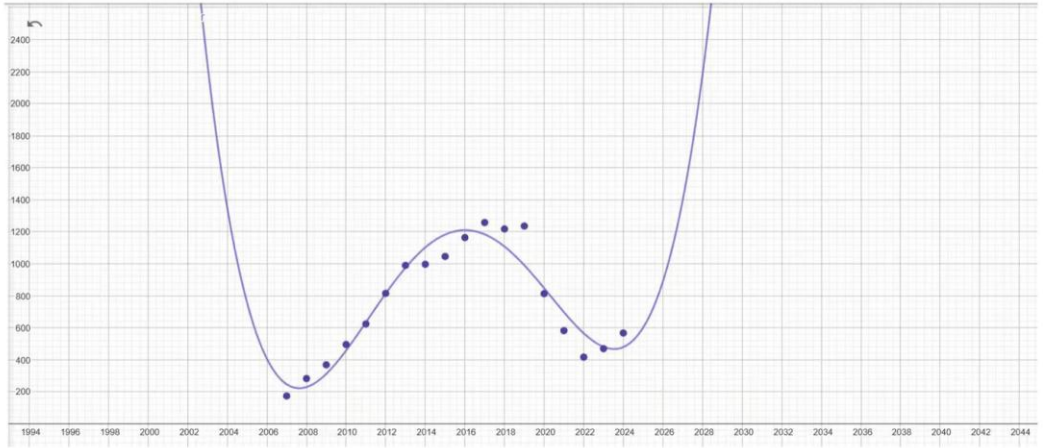
$a = 18.3415892672859$、$b = -36216.75094599243$

$\therefore y = 18.3416x - 36216.7509$

決定係數 $R^2 = 0.0765188093604$

因爲直綫方程未能擬合數據，散點與趨勢綫不貼合，與預期不理想化，決定係數太小，所以我方排除。

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



以四次方程模擬數據趨勢的方程為：$y = ax^4 + bx^3 + cx^2 + dx + e$

$a = 0.2156309759268$、$b = -1738.6105813401796$

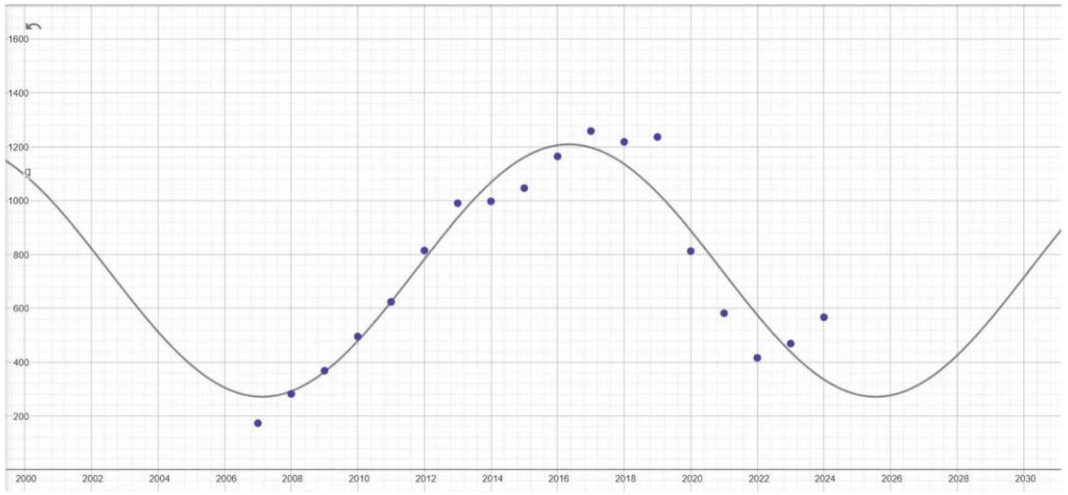
$c = 5256812.684050306$、$d = -7074117469.820244$

$e = 3559773246733.694$

$\therefore y = 0.2156x^4 - 1738.6106x^3 + 5256812.6841x^2 - 7074117469.8202x + 3559773246733.694$

決定係數 $R^2 = 0.9214340633808$

因爲四次方程過度擬合，與預期太過於理想化，2026 年之後不可能出現持續高速上升的趨勢，不符合現實，預測不准測，所以我方排除。



以正弦方程模擬數據趨勢為：$y = a \cdot \sin(bx + c) + d$

$a = 468.6964665768187$、$b = 0.340944285899$

$c = -1.0212951108444$、$d = 739.968794179509$

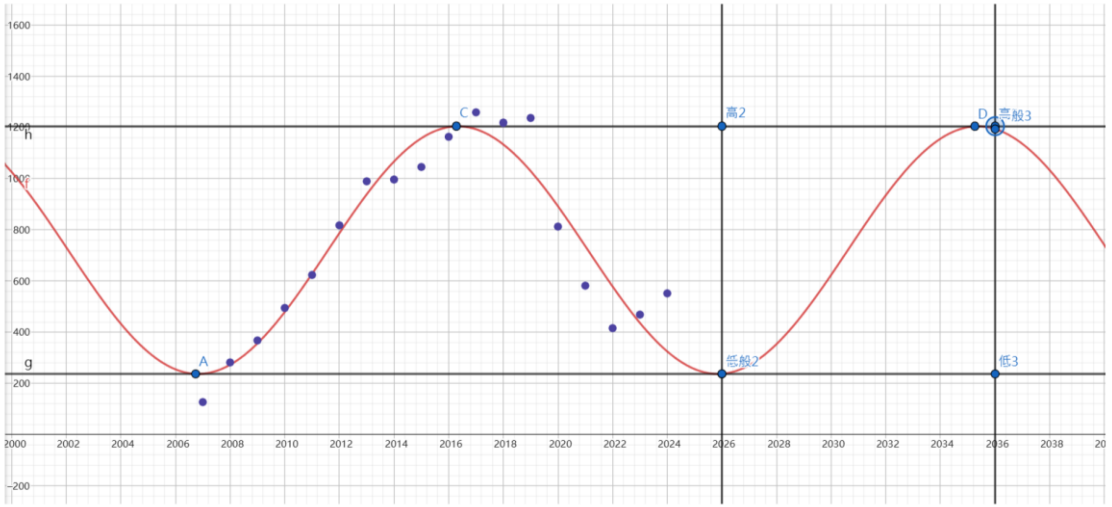
$\therefore y = 468.6965 \cdot \sin(0.3409x - 1.0213) + 739.9688$

決定係數 $R^2 = 0.9083044004034$

因爲正弦方程擬合數據的決定係數較接近於 1，且呈現出週期性的性質，所以我方採用正弦方程。

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

從以上方程圖像及決定係數比較，我方決定使用正弦方程的模擬數據趨勢方程來 2026 和 2036 年香港的社區型耐甲氧西林金黃葡萄球菌感染者數目。根據使用正弦方程模擬趨勢方程，我們可以知道：



在 2026 年時：

一般情況的數目:238 宗數；

高爆發情況的數目:1205 宗數;

低爆發情況:237 宗數。

在 2036 年時:

一般情況的數目:1194 宗數；

高爆發情況的數目:1205 宗數；

低爆發情況的數目:237 宗數。

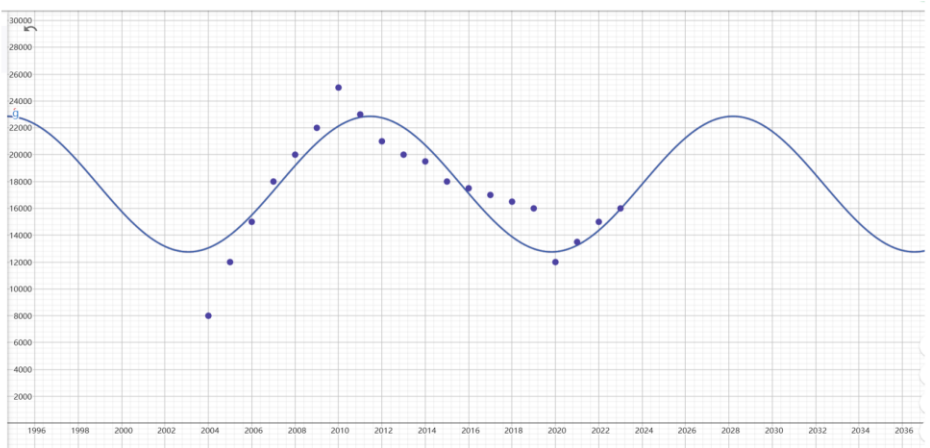
以下是基於美國公共衛生機構（如 CDC）和研究文獻中可獲取的社區型耐甲氧西林金黃色葡萄球菌（CA-MRSA）感染年度數據的整理。

美國社區型耐甲氧西林金黃葡萄球菌感染人數及年份

年份	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
感染人數	8000	12000	15000	18000	20000	22000	25000	23000	21000	20000

年份	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
感染人數	19500	18000	17500	17000	16500	16000	12000	13500	15000	16000

(U.S.CENTERS FOR DISEASECONTROL AND PREVENTION)



首先，美國的社區型耐甲氧西林金黃葡萄球菌感染者均主要集中在美國經濟發達的地區，例如紐約、洛杉磯、芝加哥等城市，與香港的環境相似，人口流動性大。其次，根據上圖示，美國的模擬數據趨勢同樣出現週期性性質，與我方的模擬數據趨勢相貼合。

Sample report (1) – Model refinement
參考作品 (1) – 改進模型

5. 評估模型--對模型的批評

在以上的數學模型中,我們發現 2026 年高或低爆發情況下的感染人數與 2036 年完全相同，不斷地循環一個週期，不符合常理。現實中一定會因某些因素而發生改變，會出現上升或下降的趨勢（高峰期或低峰期），而我方的數學模型中未考慮此問題，導致整個數學模型無法呈現一個不斷週期性循環的趨勢。因此以上的模型並不能完全準確地預測社區型耐甲氧西林金黃葡萄球菌感染並於不同情況下在 2026 年和 2036 年的香港病例總數。因為正弦方程的性質，所以在十年之間高爆發與低爆發不可能呈一樣數目出現，在現實中也不會發生這種情況，因此以上的數學模型的預期較不理想。

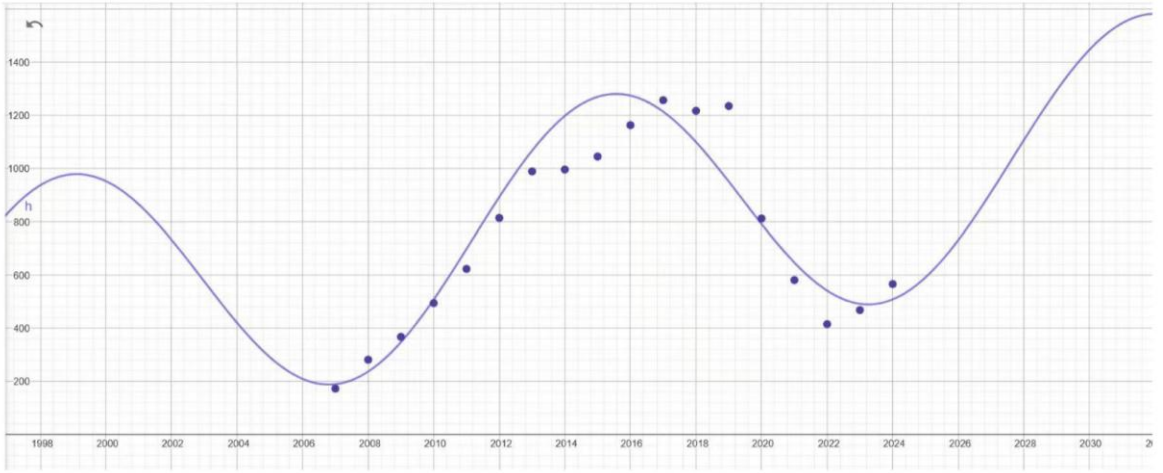
6.1 修訂數學模型

基於以上的反思，我方嘗試修訂以上的數學模型。將直綫方程模擬數據趨勢與正弦方程模擬數據趨勢相結合，使其出現不斷週期性循環上升的趨勢，讓數學模型更能準確預測社區型耐甲氧西林金黃葡萄球菌感染並於不同情況下在 2026 年和 2036 年的香港病例總數。

6.1.2 新的假設

7. 社區型耐甲氧西林金黃葡萄球菌感染具有週期性爆發的趨勢

8. 社區型耐甲氧西林金黃葡萄球菌感染有上升或下降的綫性趨勢



以正弦方程與直綫方程相結合的模擬數據趨勢為：$y = a \cdot \sin(bx + c) + ex + d$

$a = 468.6964665768187$、$b = 0.3818$

$c = -1.3212$、$d = -37629.968$

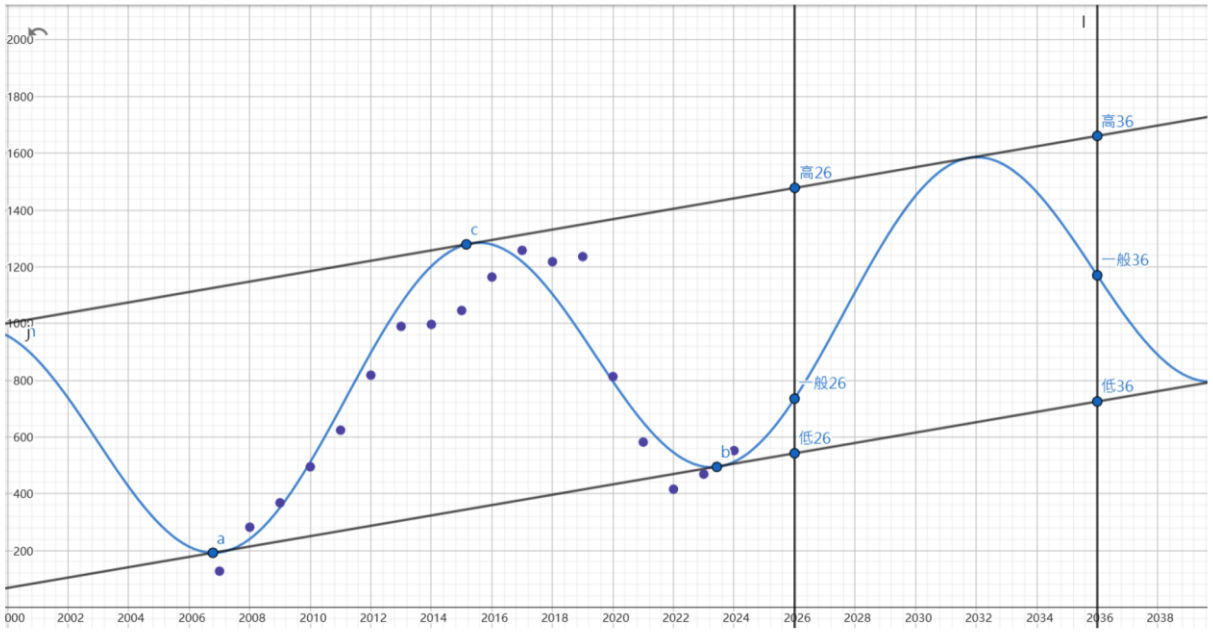
$e = 18.3415892672859$

$\therefore y = 468.6965 \cdot \sin(0.3818x - 1.3212) + 18.3416x - 37629.968$

決定係數 $R^2 = 0.980248034$

Sample report (1) – Model refinement
參考作品 (1) – 改進模型

6.2　修訂模型的預測



在 2026 年時：

一般情況的數目:735 宗數；

高爆發情況的數目:1478 宗數;

低爆發情況:543 宗數。

在 2036 年時:

一般情況的數目:1170 宗數；

高爆發情況的數目:1661 宗數；

低爆發情況的數目:725 宗數。

6.3 修訂模型的評估

根據正弦方程與直綫方程相結合的模擬數據趨勢圖像，隨機預測過去的五年，與現實的原本的數據做對比：

年份	現實數據	預測數據	誤差（宗數）	誤差百分比
2009	368	352	16	4.31%
2010	495	512	18	3.64%
2017	1258	1217	40	3.18%
2020	813	796	17	2.09%
2023	469	496	27	5.76%

我方預估在 2026 年時，一般情況的數目為 735、高爆發情況的數目為 1478、低爆發情況為 543；在 2036 年時，一般情況的數目為 1170、高爆發情況的數目為 1661、低爆發情況的數目為 725，發現現實數據與預測數據的誤差小，誤差百分比小，與現實數據相差不大，因此我方預測的社區型耐甲氧西林金黃葡萄球菌感染並於不同情況下在 2026 年和 2036 年的香港病例總數屬符合常理及預期。

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

Summary

The outbreak of infectious diseases has been serious for decades. In this paper, we will introduce 3 models. The first one is designed to calculate the risk levels of different diseases. The second one uses multiple regressions to predict the number of infected cases in different scenarios. The third model is used to predict the future level of outbreaks under different scenarios.

Our first model, the **Level of Infection Model** is used to predict the level of infection of different infectious diseases. We first deduce the future number of infections of each disease using multiple regressions and select the 5 highest infection as our target. We then search for different factors that affect the intensity of a disease and apply the data to our model, coming to the conclusion that Scarlet Fever and Food Poisoning has the highest level of infection, and will thus require the most attention.

For our second model, we made a model which is a formula that takes variables like: death rates, reproduction number, reinfection numbers, incubation period and different transmission ways of each respective disease into account. The formula will show that the larger the number we obtained through replacing the variables mentioned, the higher the level of infection, and thus requires a higher attention. Assisting in the accumulation of the found 5 notifiable diseases to find out the one disease to be researched into for the second task.

For our third model, three sets of **Gaussian Process Model** are used to predict the future trends (2026-2036) of Scarlet Fever to find out about the total number of its cases in Hong Kong in 2026 and 2036 under different scenarios: **normal scenario**, **high-outbreak scenario** and **low outbreak scenario**.

As the second and third model are both proposing solutions to the second question, we compare the result at the end to analyse the difference.

At the end of each model, we have listed out the strengths and weaknesses of the models we created and used to further reflect on our work for this topic.

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

1.1. Background

Infectious disease outbreaks have been a major problem in Hong Kong. From Tuberculosis to Dengue Fever, these infectious diseases have affected many people in Hong Kong since 1997.

To define infectious diseases, it refers to a disorder that is primarily caused by microscopic organisms like bacteria, viruses, fungi, or parasites. They are known to be fast-spreading and high-risk, causing outbreaks of infectious diseases to become a common phenomenon. Thus, focusing on the prediction of those outbreaks can greatly reduce the threat and danger of infectious diseases.

1.2. Problem Restatement

In this paper, we will solve 2 problems.

For the first problem, we are tasked to suggest and identify two diseases that deserve special attention in Hong Kong for the coming future. Thus, we will start off by collecting data of the number of notifiable infectious diseases each year (1997-2024), which is data provided by the CHP's website[1]. To tackle this problem, we have created a level of infection model which addresses which notifiable diseases have always had an ongoing and significant presence over the years. After identifying a few of those diseases, we will put them through a series of multi-regressions to predict their impact on Hong Kong in 2025. The two diseases with the largest number of predicted infectants will be chosen for this problem.

For the second problem, we are tasked to select one of the two found diseases to develop a mathematical model to predict the total number of cases in Hong Kong 2026 and 2036 under 3 different scenarios: a normal scenario, high-outbreak scenario and low-outbreak scenario. Thus, to tackle this problem, we have selected scarlet fever as our target and will put the data we got beforehand into a Gaussian Process Regression which will predict ranges of the number of infectants from 2026-2036.

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

2.1. Assumption and justification

Assumption

1. The R_0 (basic reproduction number) and the death rate of the diseases will not change.
2. The recurrent infections and the incubation period that happen on each person are the same.
3. There will not be new vaccines or new medicines in the future that will decrease the number of infections.
4. The diseases will not undergo genetic mutation in the future.

Justification

1. This is for the simplicity of our model as R_0 and death rate ranges widely.
2. Some people may encounter recurrent infections and a longer incubation period depending on their health. This is for easier calculation.
3. We are unable to predict when new inventions aimed at combating diseases will be released. This assumption simplifies our model.
4. There is no observable trend of the genetic mutation of infectious diseases. Also, genetic mutation of diseases may weaken or strengthen the infectious rate which is unpredictable.

2.2. Definition of Variables

Pn	Predicted Number of Infections in 2025
R_0	Basic reproduction number
Dr	Death Rate
Ri	Recurrent Infections
Ip	Incubation Period
Tw	Transmission ways

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

2.3. Model Description

Our model is designed to calculate the level of risks of different infectious diseases. The larger the number obtained through using this model, the higher the risk of the disease.

2.3.1. Model Overview

The following is the model we constructed:

$$f(Pn, R_0, Dr, Ri, Ip, Tw) = \frac{Pn \cdot R_0 \cdot (1-Dr) \cdot (1+Ri)}{Ip} \cdot Tw$$

The larger the number we obtained through replacing the variables in the above model, the higher the level of infection, and thus requires a higher attention.

In this model, Pn is the predicted number of infections. We use data from the recorded year to predict the infection number in 2025-01-01 using **multiple regression**.

R_0 is the basic reproduction number, which influences the spread of the infection directly. The higher this value, the more infections occur.

Dr is the death rate. $(1 - Dr)$ shows the proportion of individuals who survive in the infection. A higher death rate reduces the level of infection.

Ri represents the proportion of reinfections. $(1 + Ri)$ shows the increased level of infection due to reinfections.

Ip refers to the incubation period. It acts as a divisor as a longer incubation period may delay the symptoms onset and reduce the immediate impact of the infection.

Tw is the ways of transmission of diseases. We weight the different transmission ways according to their prominence . If the disease has more than one way of transmission, the weight will be added up before the multiplication.
The following is the weighting we calculated:

Respiratory droplet transmission	5
Transmission through physical contact	4
Aerosol	3
Fecal-oral route	1

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

2.3.2. Model Application

After calculating the predicted infections in 2025-01-01, we select the top 5 diseases that will have the most infections[*]. They are:

Tuberculosis

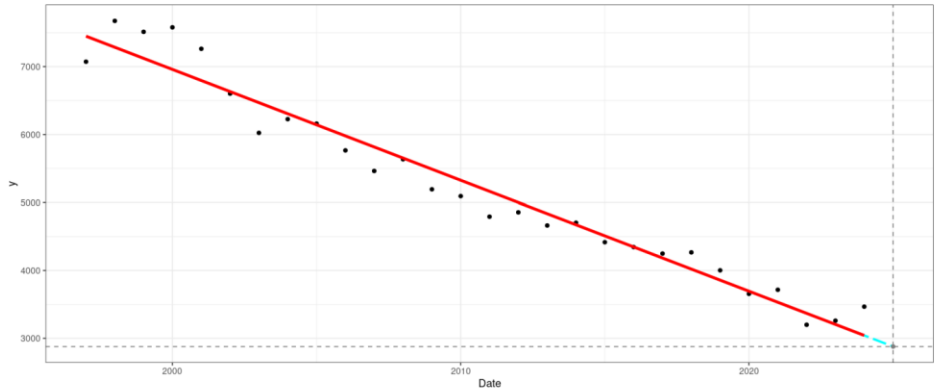


Fig. 3 *Linear Graph of no. of Tuberculosis cases*

$y = -0.446735081139859x + 7447.77017470293$

Predicted number: 2879.01

Result after applying the model: 174.78[2]

Community-associated methicillin-resistant Staphylococcus aureus infection (CAMRSAI) 社區型耐藥性金黃葡萄球菌

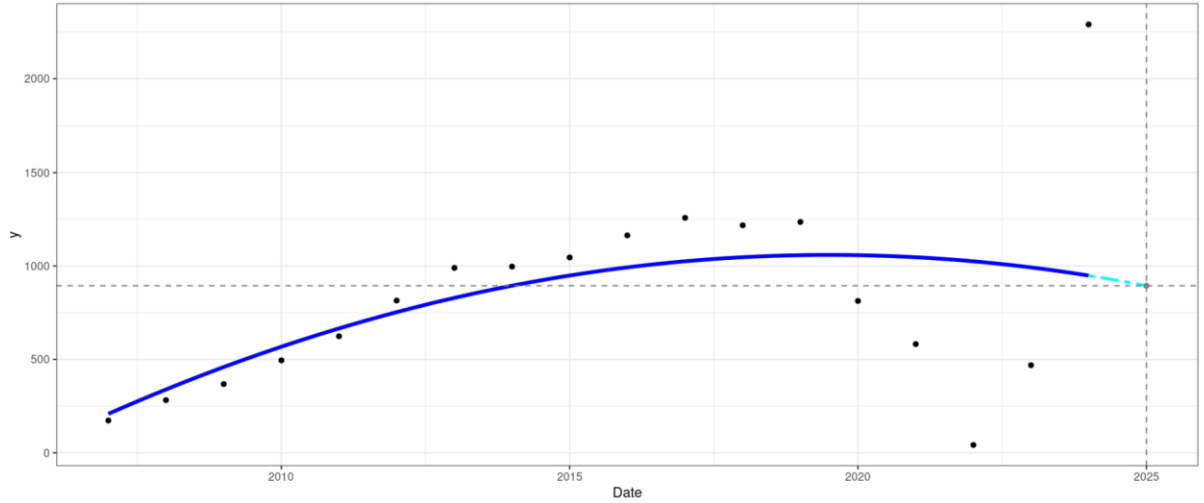


Fig. 5 *Quadratic Graph of no. of CAMRSAI cases*

$y = -4.08401169084952e-05x^2 + 0.372820681972283x + 208.482531481418$

Predicted number: 894.2347

Result after applying the model: 703.89[4]

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

Scarlet fever

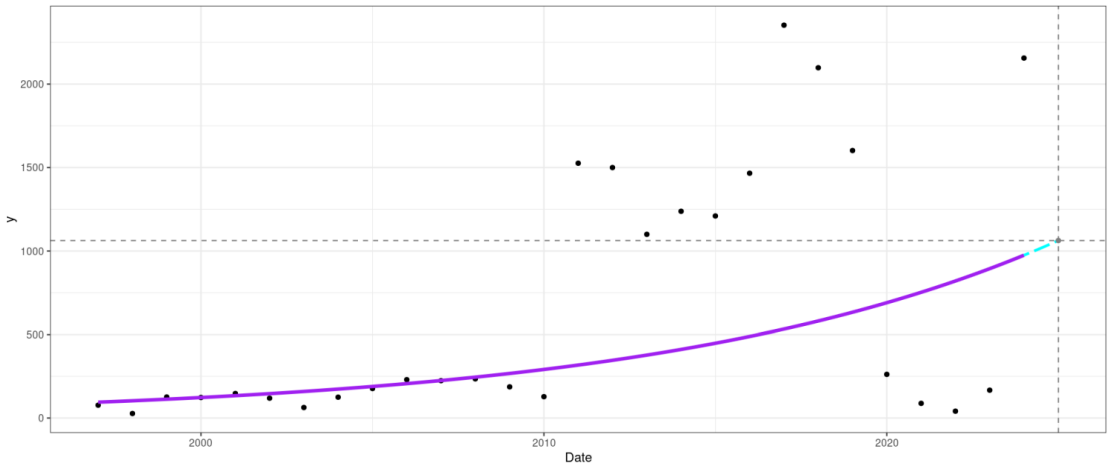


Fig. 4 *Exponential Graph of no. of Scarlet Fever cases*

$y = 94.9542577594466 \cdot 1.00023619423202^{x}$

Predicted number: 1062.807

Result after applying the model: 16926.93[3]

Viral hepatitis - E

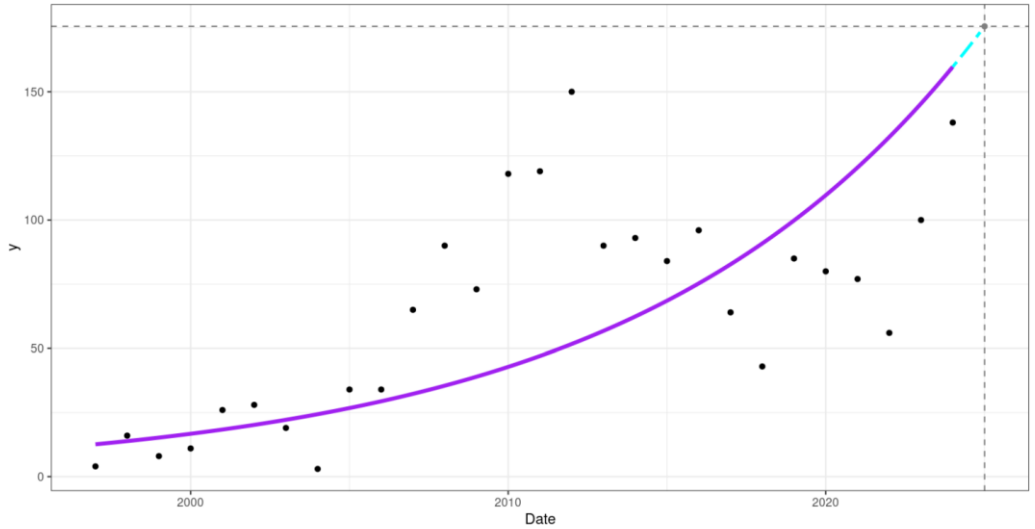


Fig. 7 *Exponential Graph of no. of Viral Hepatitis-E cases*

$y = 12.6185011506163 \cdot 1.00025744483431^{x}$

Predicted number: 175.5131

Result after applying the model: 11.16[6]

2.5 Strengths and Weaknesses

Strength

Comprehensive

In this model, most of the important factors of the spreading of infectious disease are considered. They include the predicted number of infections, basic reproduction number (R_0), death rate, recurrent infections, incubation period, and common transmission ways. Our model is comprehensive which makes the results more reliable and valid.

Weakness

Unpredictable change

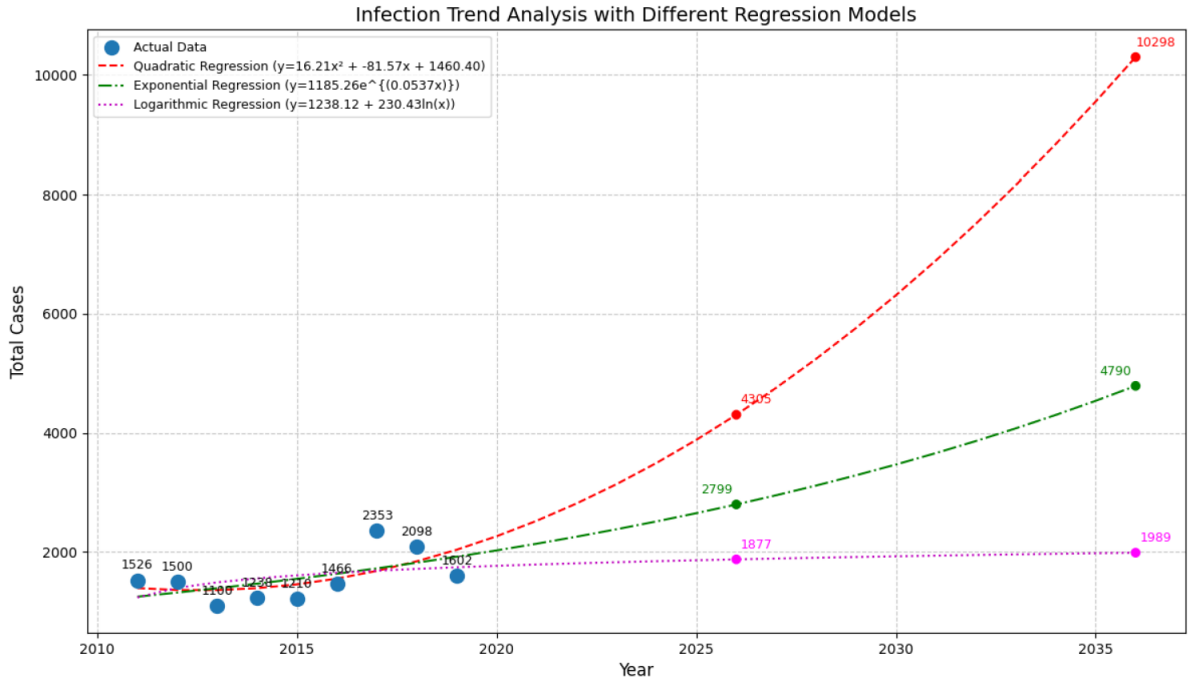
As we use data from previous years to predict future infection numbers, we cannot ensure that the number of infections in the future will change according to the trend we observe. There may be a sudden outbreak of a certain infectious disease due to factors such as genetic mutation.

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

3.2. Outbreak Scenario Model 1

3.2.1. Model Overview[**]

Quadratic regression, **exponential regression** and **logarithmic regression** are used to predict the number of cases in 2026 and 2036 in high-outbreak scenario, low-outbreak and normal scenarios respectively.



High-outbreak scenario: $y = 16.21x^2 - 81.57x + 1460.4$
In a high-outbreak scenario, the predicted number in 2026 will be 4305 and in 2036 will be 10298.

Low-outbreak scenario: $y = 1185.26e^{0.0537x}$
In a low-outbreak scenario, the predicted number in 2026 will be 2799 and in 2036 will be 4790.

Normal scenario: $y = 1238.12 + 230.43ln(x)$
In a normal scenario, the predicted number in 2026 will be 1877 and in 2036 will be

Scenario	2026 Predictions	2036 Predictions
High-outbreak scenario	4305	10298
Low-outbreak scenario	2799	4790
Normal scenario	1877	1989

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

3.3. Outbreak Scenario Model 2[7]

2. Core Modeling Technique: Gaussian Process (GP)

- **What is a GP?**
  - A flexible, probabilistic machine learning method that predicts future trends while quantifying uncertainty. It is like a "smart curve-fitting" tool that accounts for noise and patterns in data.
- **Key Idea:**
  - The GP assumes infections follow a pattern defined by:

$$Infections = Long\ Term\ trend + Seasonality + Noise$$

4. Scenario Configuration

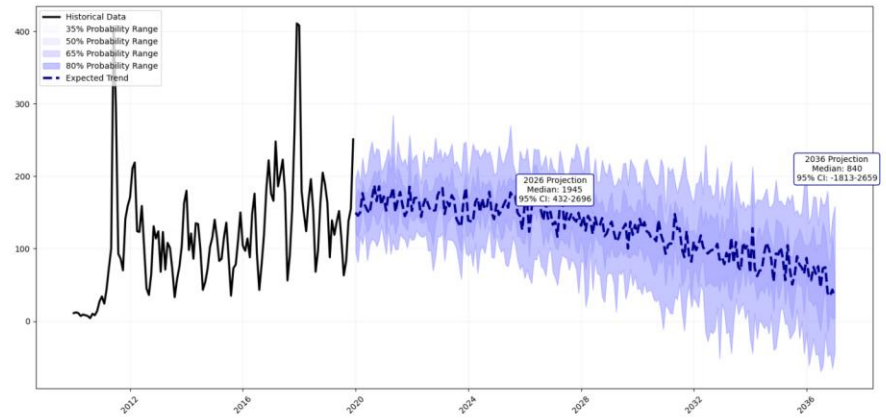
Each scenario defines how these kernels are combined:

Scenario	RBF Weight	Seasonality Weight	Trend Weight	Noise Level	Interpretation
Normal	Moderate	Disabled	Disabled	Low	Stable baseline transmission
High-outbreak scenario	Strong	Strong	Strong	High	Rapid spread seasonal spikes
Low-outbreak scenario	Weak	Weak	Minimal	Very Low	Contained transmission

Key Metrics:

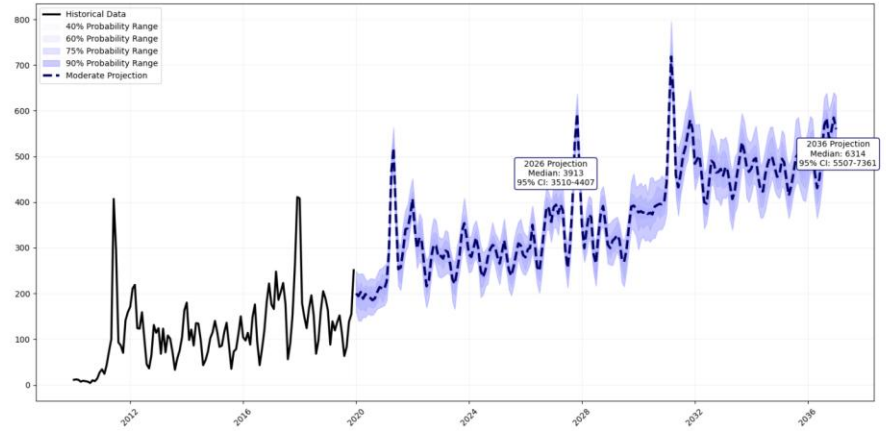
- **Median**: Most likely outcome.
- **95% Confidence Interval**: Range covering 95% of predictions.
- **Peak Infections**: Worst-case estimates (e.g., 99th percentile).

Normal Scenario[***]



Normal scenario: the trend is gradually decreasing. In 2026, the range of Scarlet Fever infection is 432 to 2696 people, with a median of 1945 people. In 2036, it is projected to have a range of -1813 to 2659 people, and the median is 840 people.

High-outbreak Scenario[****]



High-outbreak scenario: the trend is generally increasing. In 2026, the predicted range of infection is 3510 to 4407 people, while the median is 3913 people. In 2036, the range will be 5507 to 7361 people, with a median of 6314 people.

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。

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

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/

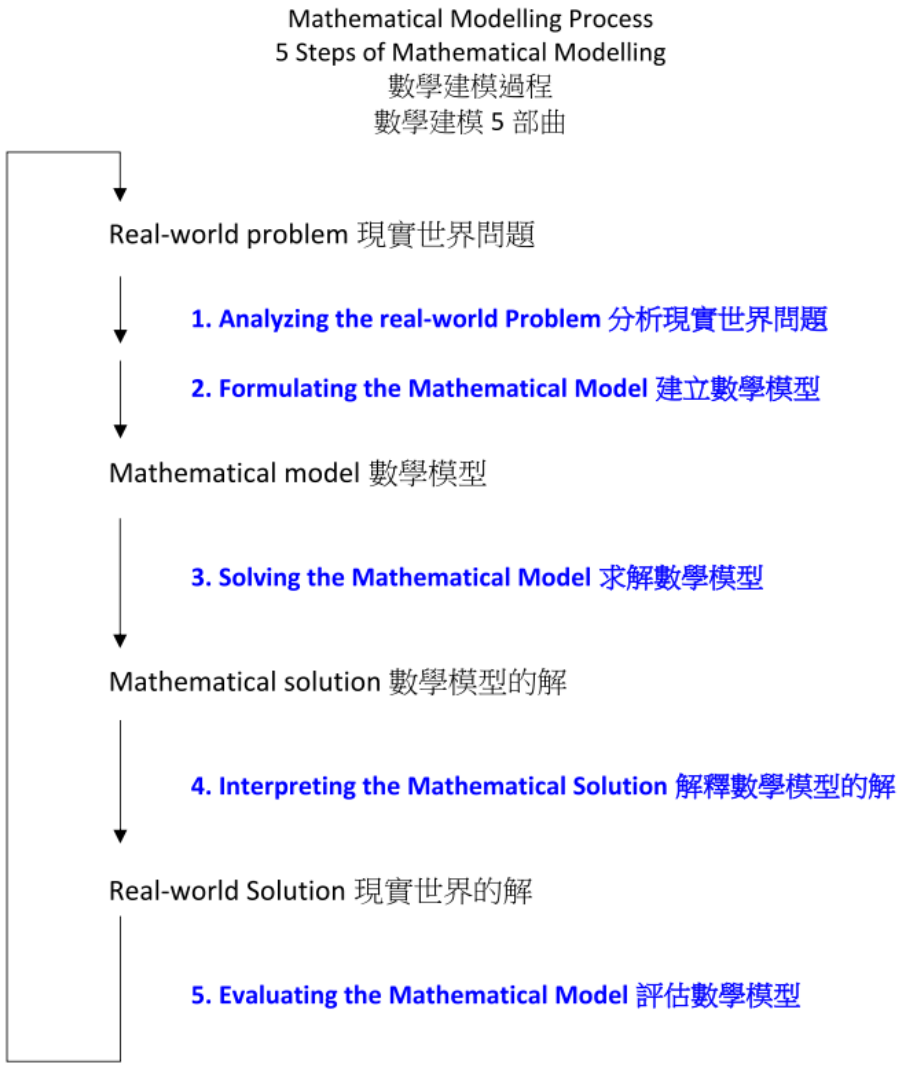
- Predicting Price Indices and Weather Prediction Using Markov Chains
  使用馬可夫鏈預測價格指數和天氣預報
  http://mathcal.math.cuhk.edu.hk:7539/

- 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

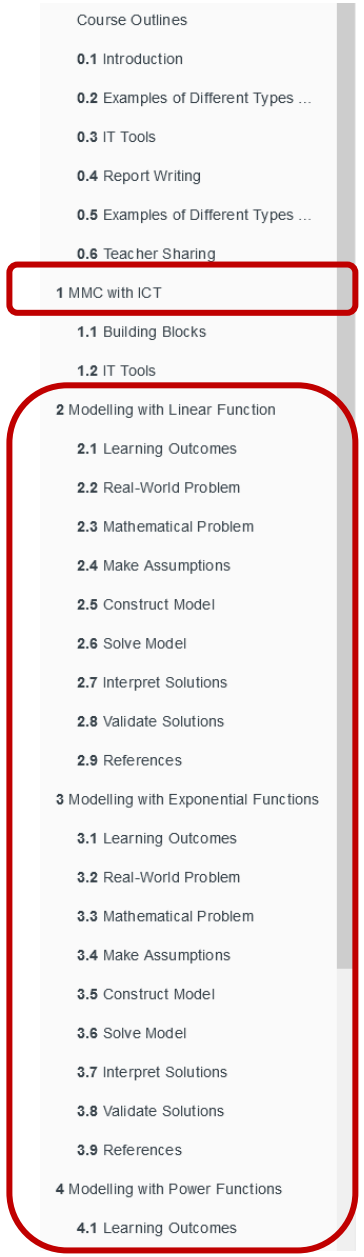


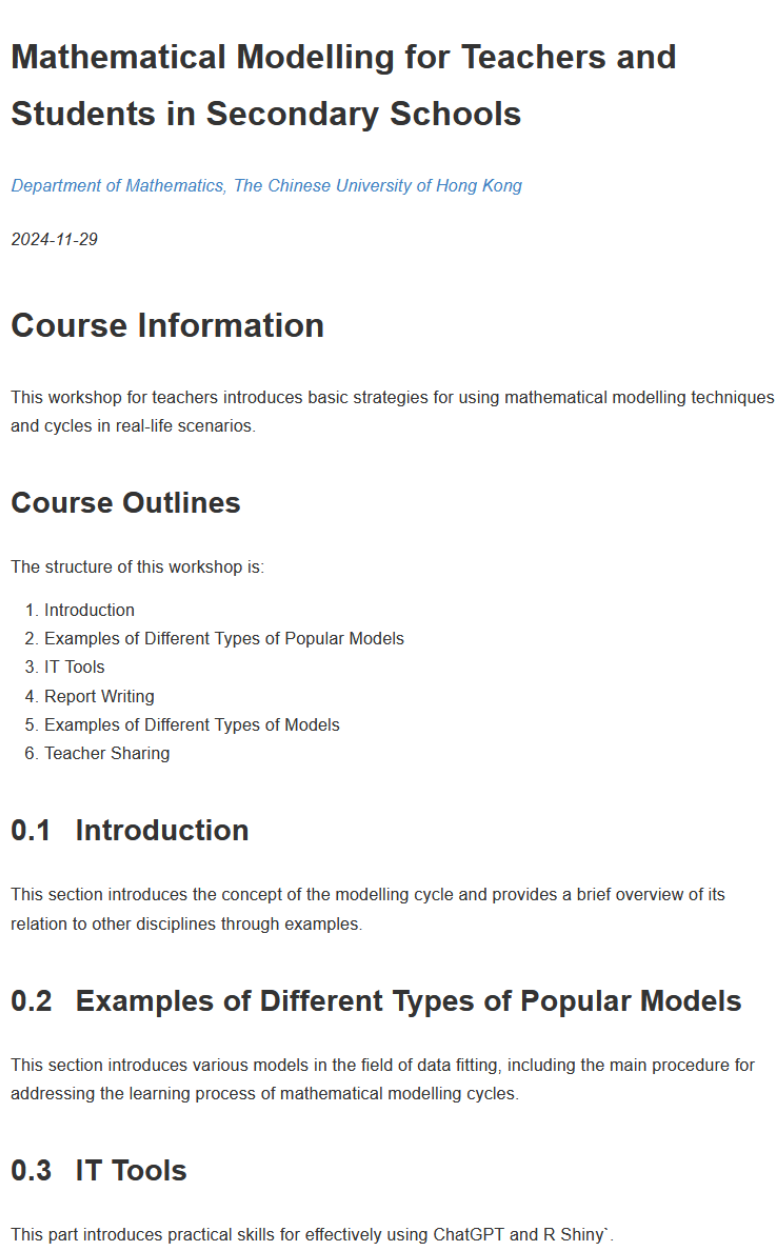
Summary: Math Modelling e-book 數學建模電子書

- **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





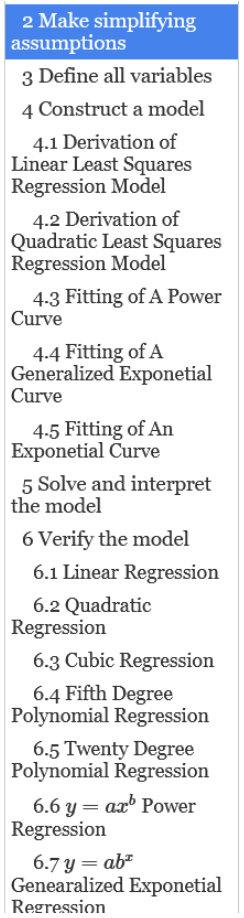
Summary: IT Tools for Math Modelling 數學建模 IT 工具

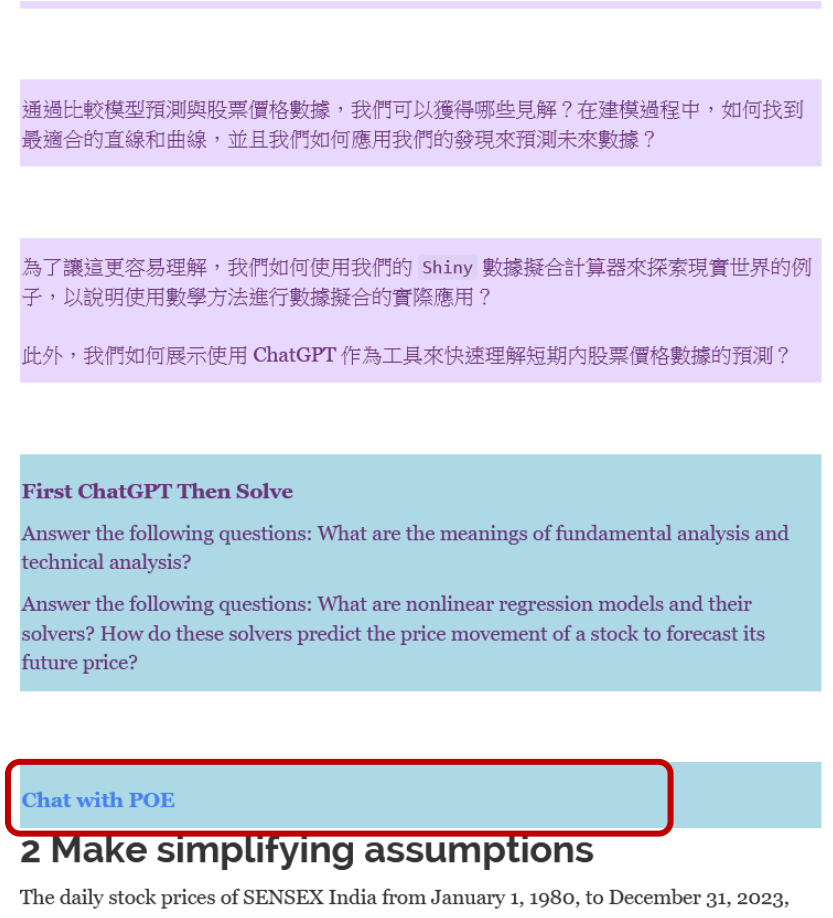
- *First-ChatGPT-Then-Solve*

- 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/ZPTbiEK5JP5REwvq5

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

Contact:
mathmodel@math.cuhk.edu.hk

Thank you!