

# 英雄聯盟電競資料視覺化：數據解析與比賽趨勢分析

## 期中規劃報告

### 目錄

- 摘要
- 第1章：簡介
  - 1-1：專題研究動機
  - 1-2：專題目標
  - 1-3：所遇問題及解決方法
  - 1-4：專題進度表
  - 1-5：專題分工
- 第2章：相關研究
  - 2-1：電子競技產業概況
  - 2-2：League of Legends遊戲機制
  - 2-3：資料視覺化在電競分析的應用
- 第3章：系統設計與架構
  - 3-1：資料收集方法
  - 3-2：資料處理流程
  - 3-3：視覺化工具選擇
- 第4章：資料分析面向
  - 4-1：英雄選用統計
  - 4-2：比賽表現分析
  - 4-3：選手與戰隊分析
  - 4-4：版本更新影響分析
  - 4-5：觀眾與市場分析
- 第5章：預期成果
- 第6章：參考資料

### 摘要

本專題旨在透過對League of Legends ( 英雄聯盟 ) 電競賽事的資料視覺化分析，探索遊戲策略、選手表現與觀眾互動之間的關聯。英雄聯盟作為全球最受歡迎的電子競技遊戲之一，每年產生海量比賽資料，蘊含著豐富的分析價值。本研究將收集2022-2024年主要賽區 ( LPL、LCK、LEC、LCS ) 及國際賽事的資

料，運用Python進行處理與視覺化，創建互動式儀表板展示英雄選擇趨勢、比賽節奏變化、選手表現指標等多維度分析，為電競愛好者、分析師和產業相關者提供數據支持的洞察。

## 第1章：簡介

### 1-1：專題研究動機

電子競技（eSports）產業近年來經歷了爆發性增長，其中League of Legends（英雄聯盟）作為最具影響力的電競遊戲之一，吸引了全球數億玩家和觀眾。這款遊戲的職業比賽體系完善，每年產生大量結構化資料，包括選手統計、英雄表現、比賽策略等。

然而，這些豐富的資料往往以分散且難以理解的形式存在，缺乏系統性的視覺化呈現。對於普通觀眾、遊戲玩家甚至專業分析師而言，從原始數據中提取有價值的洞察是極具挑戰性的。

此外，作為資訊管理系學生，我對於如何運用資料視覺化技術解析複雜系統極具興趣。英雄聯盟作為一個涉及策略、技術與心理因素的複雜競技系統，提供了理想的研究對象。透過視覺化分析這些資料，不僅能夠發現遊戲內部的規律與趨勢，也能練習和應用資料收集、清理、分析與視覺呈現的全流程技能。

### 1-2：專題目標

本專題擬定以下具體目標：

- 建立資料收集管道：**利用Riot Games API、專業數據網站（如Oracle's Elixir、Games of Legends）等資源，建立穩定的資料收集流程，獲取2022-2024年英雄聯盟主要賽區的職業比賽資料。
- 開發互動式視覺化儀表板：**使用Python生態系（如Pandas、Matplotlib、Plotly等）開發一系列互動式視覺化圖表，清晰呈現多維度分析結果。
- 多角度分析電競生態：**從英雄選用率、選手表現、比賽策略、版本更新影響等多個角度，揭示英雄聯盟電競生態的發展規律與趨勢。
- 特定場景案例分析：**選取2-3個典型賽事（如世界賽、季中邀請賽），進行深度案例分析，展示數據視覺化在理解高水平比賽中的應用價值。
- 創建網頁展示平台：**利用基本的網頁技術（HTML、CSS、JavaScript），構建簡潔的網頁界面，整合所有視覺化成果。

### 1-3：所遇問題及解決方法

在專題初期規劃階段，已預見以下可能的挑戰及對應解決方案：

- 資料獲取限制：**
  - 問題：**Riot API有使用頻率限制，且部分歷史資料可能不完整。
  - 解決方法：**結合使用官方API與第三方數據網站（如Oracle's Elixir提供的開放資料集），同時實施資料緩存機制減少API呼叫。
- 大量非結構化資料處理：**
  - 問題：**賽事評論、選手訪談等質性資料難以系統化分析。

- **解決方法：**使用YouTube Data API提取影片資訊，結合簡單的自然語言處理工具分析評論趨勢，輔助量化資料的解讀。

3. 視覺化交互性限制：

- **問題：**靜態報告難以呈現互動式視覺化的完整功能。
- **解決方法：**在報告中展示關鍵靜態圖表，同時開發基於HTML的簡易網頁，實現完整互動功能，並提供網站連結或視頻演示。

4. 多維資料的整合挑戰：

- **問題：**需整合英雄數據、選手表現、比賽環境等多維度因素。
- **解決方法：**設計多層次視覺化架構，從概覽到細節逐層展開，使用儀表板方式組織不同維度的視覺化。

1-4：專題進度表

週次	日期	工作內容
第1週	4/15~4/21	確定研究範圍，初步資料來源調研
第2週	4/22~4/28	資料收集腳本開發，API調用測試
第3週	4/29~5/5	建立資料清理與轉換流程
第4週	5/6~5/12	英雄選用分析模組開發
第5週	5/13~5/19	比賽表現與選手分析模組開發
第6週	5/20~5/26	版本更新影響與市場分析模組開發
第7週	5/27~6/2	網頁界面設計與整合
第8週	6/3~6/9	系統測試與完善，報告撰寫
第9週	6/10	期末專題發表準備

1-5：專題分工

本專題由個人獨立完成，總體工作分為以下幾個模組：

1. 資料收集與處理模組
2. 英雄選用分析模組
3. 比賽與選手分析模組
4. 版本與市場分析模組
5. 網頁展示平台開發
6. 報告撰寫與發表準備

第2章：相關研究

2-1：電子競技產業概況

電子競技 ( eSports ) 已發展成為一個價值數十億美元的全球產業。根據Newzoo的報告，2023年全球電競市場規模達到13.8億美元，預計到2026年將增長至18.7億美元。League of Legends作為最具影響力的電競遊戲之一，其職業聯賽體系包括中國LPL、韓國LCK、歐洲LEC、北美LCS等區域賽事，以及世界賽 ( World Championship )、季中邀請賽 ( Mid-Season Invitational ) 等國際賽事。

英雄聯盟電競產業形成了完整的生態系統，包括職業戰隊、賽事組織、遊戲開發商、贊助商、媒體平台與觀眾群體。2022年世界賽總決賽的全球觀看人數超過5000萬，展現了這款遊戲持續的全球影響力。

## 2-2：League of Legends遊戲機制

League of Legends是一款5v5團隊競技遊戲，每位玩家控制一名「英雄」，與隊友合作摧毀敵方基地。遊戲中有超過160名英雄可供選擇，每名英雄擁有獨特的技能組與特性。比賽通常分為三個階段：

1. **選角階段**：雙方隊伍輪流禁用 ( Ban ) 與選擇 ( Pick ) 英雄，這一階段體現了戰隊的策略思考與準備。
2. **對線期**：比賽初期，選手專注於積累個人資源與優勢。
3. **團隊戰階段**：中後期，隊伍圍繞地圖資源與目標展開團隊戰鬥，最終推進敵方基地。

遊戲公司Riot Games每隔約兩週發布一次平衡性更新，調整英雄屬性與遊戲機制，這些變動直接影響比賽策略與英雄選擇偏好。

## 2-3：資料視覺化在電競分析的應用

資料視覺化在電競分析中扮演著越來越重要的角色。目前的相關應用主要包括：

1. **賽前分析**：視覺化展示雙方戰隊歷史對戰數據、選手專精英雄與表現，輔助預測比賽走向。
2. **即時比賽視覺化**：如金錢差距圖、經濟變化曲線等，幫助觀眾理解比賽局勢發展。
3. **賽後分析**：深入分析比賽數據，評估戰術有效性，總結選手表現。
4. **長期趨勢分析**：研究英雄選用率變化、戰術演變等宏觀趨勢。

然而，目前電競領域的資料視覺化仍存在若干限制：大多數分析停留在描述性層面，缺乏深度洞察；視覺化形式較為單一，互動性不足；跨維度分析不足，難以全面呈現電競生態的複雜性。本專題旨在嘗試突破這些限制，探索更全面、深入的電競資料視覺化應用。

# 第3章：系統設計與架構

## 3-1：資料收集方法

本專題將採用多源資料收集策略，主要包括：

1. **Riot Games API**：使用官方API獲取比賽基礎數據，主要針對高水平排位賽的信息。將使用Python的requests庫開發API調用模組，實現自動化資料擷取。
2. **專業數據網站**：
  - **Oracle's Elixir**：提供詳細的職業比賽統計數據，包括選手KDA、經濟指標、視野控制等。
  - **Games of Legends**：提供比賽選角禁用記錄與基本比賽記錄。

- **LoLEsports**：官方賽事資訊平台，提供賽程與基本統計。

### 3. 社群媒體資料：

- **YouTube**：收集主要賽事直播回放的觀看數據與評論。
- **Twitter/微博**：通過關鍵詞搜索收集粉絲討論與熱門話題。

資料將重點關注以下幾項核心內容：

- 比賽匹配數據（勝負、時長、擊殺等）
- 英雄選用統計（選/禁率、勝率、登場次數）
- 選手績效數據（KDA、經濟、地圖控制）
- 版本更新信息（英雄調整、遊戲機制變化）
- 觀眾互動數據（觀看人數、評論熱度）

## 3-2：資料處理流程

資料處理將遵循以下流程：

1. **資料提取**：使用Python腳本自動從上述各來源提取原始資料，建立本地數據庫。
2. **資料清理**：
  - 處理缺失值與異常值
  - 標準化不同來源的數據格式
  - 驗證數據一致性
  - 合併重複記錄
3. **資料轉換**：
  - 計算派生指標（如平均KDA、經濟轉換效率等）
  - 時間序列重組（按版本、賽季、年份等分組）
  - 特徵工程（構建新的分析維度）
4. **資料儲存**：
  - 使用SQLite本地數據庫存儲結構化數據
  - 使用CSV文件存儲中間處理結果
  - 使用JSON格式存儲用於視覺化的最終數據

## 3-3：視覺化工具選擇

考慮到專題目標與技術可行性，本項目將主要使用以下工具：

1. **數據處理**：
  - **Pandas**：用於數據清洗、轉換與分析
  - **NumPy**：用於數值計算

- **SQLite**：用於資料儲存與查詢
2. 靜態視覺化：
    - **Matplotlib**：創建基礎統計圖表
    - **Seaborn**：創建較為美觀的統計圖表
  3. 互動式視覺化：
    - **Plotly**：創建互動式圖表
    - **Dash**（可選）：構建簡單的分析儀表板
  4. 網頁展示：
    - **HTML/CSS/JavaScript**：構建基本網頁框架
    - **Chart.js**：輕量級前端圖表庫

這些工具組合將支持開發以下視覺化類型：

- 時間序列分析（折線圖、面積圖）
- 分布分析（直方圖、盒型圖、小提琴圖）
- 關聯分析（散點圖、熱力圖、相關矩陣）
- 比例分析（餅圖、堆疊柱狀圖）
- 網絡分析（英雄關聯網絡）
- 地理分析（全球賽區表現對比）

## 第4章：資料分析面向

### 4-1：英雄選用統計

英雄選用分析是理解遊戲元數據（Meta）的關鍵，將聚焦於以下方面：

1. 選用率與禁用率趨勢：
  - 開發交互式時間序列圖表，展示主要英雄的選用率與禁用率變化
  - 分賽區對比不同地區的英雄偏好差異
  - 視覺化版本更新對英雄選用的影響
2. 位置與角色分析：
  - 創建英雄池多樣性指數，評估各位置的英雄選擇靈活度
  - 分析特定位置（如中路、打野）的英雄生態變化
  - 視覺化展示英雄類型（如刺客、法師、坦克）的比例變化
3. 英雄協同與克制關係：
  - 開發英雄協同網絡圖，展示哪些英雄經常一起出現
  - 分析特定英雄的最佳搭配與最強克制

- 創建交互式熱力圖，直觀展示英雄間的勝率對比

## 4-2：比賽表現分析

比賽表現分析將著眼於遊戲內動態與節奏變化：

### 1. 比賽時長與節奏：

- 開發比賽時長分布視覺化，分析不同賽區和比賽階段的節奏差異
- 創建時間-事件視覺化，展示關鍵目標（首塔、峽谷先鋒、大龍）的控制時間分布
- 分析版本更新對比賽節奏的影響

### 2. 團隊資源分配：

- 視覺化不同位置的資源分配策略（如經濟分布、傷害佔比）
- 分析成功戰隊的資源分配模式
- 創建經濟領先-勝率關係圖，分析經濟優勢與勝率的相關性

### 3. 戰術模式識別：

- 開發戰術分類視覺化，區分推塔流、團戰流、分推流等不同戰術取向
- 分析不同戰術的成功率與流行度變化
- 創建互動式案例分析，深入剖析典型比賽中的戰術執行

## 4-3：選手與戰隊分析

選手與戰隊分析將聚焦於職業電競的人員表現：

### 1. 選手績效評估：

- 開發多維績效雷達圖，全面評估選手在各方面的能力
- 分析選手的英雄池與專精度
- 創建選手-英雄表現熱力圖，識別獨特的「招牌英雄」

### 2. 戰隊風格與演變：

- 視覺化不同戰隊的遊戲風格特徵（如團戰頻率、經濟效率）
- 分析戰隊在不同版本與賽季的表现變化
- 創建戰隊間的風格相似性聚類分析

### 3. 勝負因素分析：

- 開發決定比賽勝負的關鍵指標視覺化
- 分析不同階段（如選角階段、前期、中期、後期）對勝率的影響權重
- 創建交互式「假設情境」分析工具，模擬關鍵指標變化對勝率的影響

## 4-4：版本更新影響分析

版本更新是影響遊戲生態的關鍵因素，本面向將重點分析：

### 1. 平衡性調整影響：

- 視覺化英雄調整前後的選用率與勝率變化
- 分析遊戲機制變更（如目標屬性調整）對比賽的影響
- 創建版本更新影響力指數，量化評估不同更新的影響程度

### 2. 適應性與創新：

- 分析戰隊和選手對新版本的適應速度
- 視覺化展示新策略的出現與傳播過程
- 對比不同賽區對版本變化的反應差異

### 3. 元數據穩定性：

- 開發元數據穩定性視覺化，評估遊戲生態的穩定與變化周期
- 分析版本更新頻率與幅度對職業比賽的影響
- 創建「版本友好度」指標，識別哪些英雄受版本影響較小

## 4-5：觀眾與市場分析

電競的核心在於其觀賞性與市場影響力，此面向將分析：

### 1. 觀眾參與度：

- 視覺化不同賽事與比賽的觀看人數分布
- 分析觀眾偏好與比賽特徵（如對戰隊伍、時長、激烈程度）的關聯
- 創建社交媒體互動熱度地圖，展示比賽的社交影響力

### 2. 市場趨勢：

- 分析電競市場與遊戲內容更新的關聯
- 視覺化不同賽區的市場增長趨勢
- 創建市場關注度與遊戲元素（如特定英雄、戰隊）的關聯分析

### 3. 觀眾體驗優化：

- 基於數據分析，提出提升觀賞體驗的視覺化建議
- 分析哪些類型的比賽最受歡迎，並探討原因
- 創建觀眾參與預測模型，評估不同因素對觀眾參與度的影響

## 第5章：預期成果

本專題預期將產出以下成果：

### 1. 互動式資料視覺化網頁：

- 包含15-20個互動式圖表，涵蓋上述五個分析面向
- 提供簡潔的導航與篩選功能，允許使用者自定義視角
- 適配桌面與移動設備的響應式設計



## 2. 分析報告與案例研究：

- 基於數據視覺化的綜合分析報告
- 2-3個深度案例研究，展示如何運用視覺化理解特定比賽或趨勢
- 提供對未來電競發展的數據支持的預測與建議

## 3. 技術實現文檔：

- 詳細的資料收集與處理流程文檔
- 視覺化實現的技術說明與代碼示例
- 系統架構與功能模組說明

## 4. 實用工具與資源：

- 開發的資料收集與處理腳本
- 整理後的結構化數據集
- 可復用的視覺化模組與模板

這些成果不僅將滿足學術要求，也將為電競愛好者、分析師和產業從業者提供實用的參考工具與資源。

## 第6章：參考資料

1. Newzoo. (2023). Global Esports & Live Streaming Market Report.
2. Riot Games. (2023). League of Legends Professional Game Data Guidelines.
3. Oracle's Elixir. (2023). Professional LoL Statistics. <https://oracleselixir.com/>
4. Games of Legends. (2023). LoL Esports Statistics. <https://gol.gg/>
5. Kaytoue, M., Silva, A., Cerf, L., Meira, W., & Raïssi, C. (2012). Watch me playing, I am a professional: a first study on video game live streaming. In Proceedings of the 21st international conference on World Wide Web.
6. Yang, X., Zhao, Z., & Liu, S. (2021). LoLVis: Visual Analysis of League of Legends Game Data. IEEE Conference on Visual Analytics Science and Technology.
7. Pobiedina, N., Neidhardt, J., Calatrava Moreno, M. D. C., & Werthner, H. (2013). Ranking factors of team success. In Proceedings of the 22nd International Conference on World Wide Web.
8. Agarwala, A., & Pearce, M. (2014). Learning Dota 2 team compositions. Technical report, Stanford University.
9. Kahn, A. S., Shen, C., Lu, L., Ratan, R. A., Coary, S., Hou, J., ... & Williams, D. (2015). The Trojan Player Typology: A cross-genre, cross-cultural, behaviorally validated scale of video game play motivations. Computers in Human Behavior, 49, 354-361.
10. Schubert, M., Drachen, A., & Mahlmann, T. (2016). Esports Analytics Through Encounter Detection. In Proceedings of the MIT Sloan Sports Analytics Conference.