

Graph Mining HW1

梁謙行 110061644

許睿哲 110061592

程式及說明

程式的部分以及如何使用的 Readme 將會放在以下 Github 連結，使用前請 clone repo。

https://github.com/enochliang/Graph_Mining_Course_HW1.git

分工

梁謙行 -- 主題發想、取得地震資料、撰寫繪圖程式、撰寫數據輸出結果報告

許睿哲 -- 主題發想、使用遊戲資料取得 API 撰寫程式碼取得 dataset、解釋 dataset 資料含意

構思過程

一開始我們想到在台灣規模極大的地震發生次數較少，而小地震幾乎是家常便飯，所以我們猜測或許地震的規模對上發生次數可以符合 power law，印出數據之後發現並沒有，因為在地震規模較小的部分，有著規模越小次數越小的趨勢，所以不符合。

後來我們找到了一個遊戲，在此遊戲中發現，世界第 1000 名與第 500 名對戰的勝率只有些微差異，世界第 10 與第 50 的勝率卻幾乎不會輸。因此推測出技術與人數可能符合 power law。

取得 dataset 後發現，技術與其人數應該是屬於常態分布，只是也有 long tail。但觀測其他遊戲數據，意外發現遊戲次數與其對應的人數符合 power law。

遊戲統計數據意義解釋

APM: Attacks Per Minute，每分鐘的攻擊行數。舉例，消一行的攻擊為 0，消兩行的攻擊為 1，消三行的攻擊為 2，T 轉與消四行攻擊為 4，且隨 combo 次數越多攻擊越高。

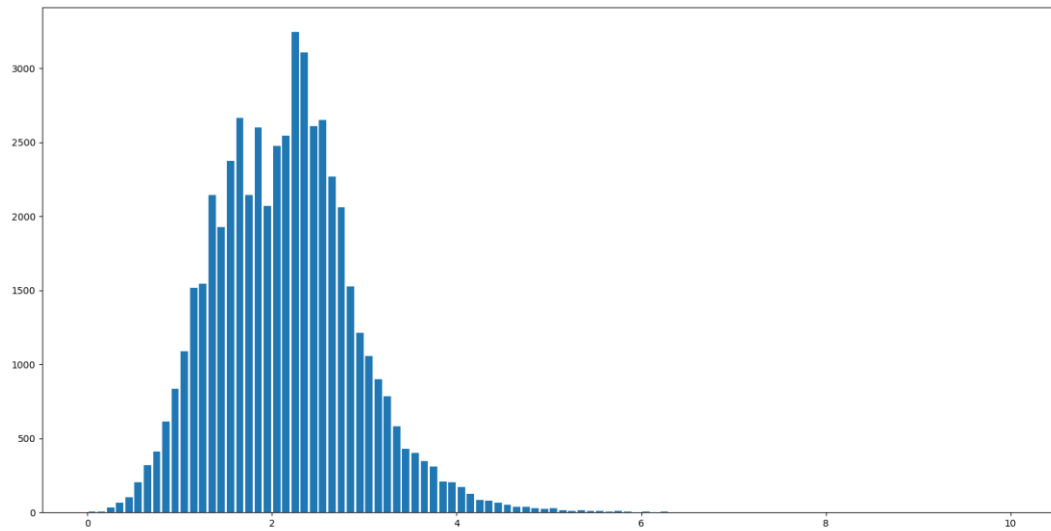
VS: Versus Score，每場遊戲中技術的指標，考慮到對敵人的攻擊(lines sent)與消除敵人的攻擊(garbage cleared)。

$$VS = \left[\frac{LinesSent + GarbageCleared}{pieces} \right] * pps * 100$$

pieces: 遊戲中的總方塊數

pps: pieces per second，每秒鐘的方塊數。

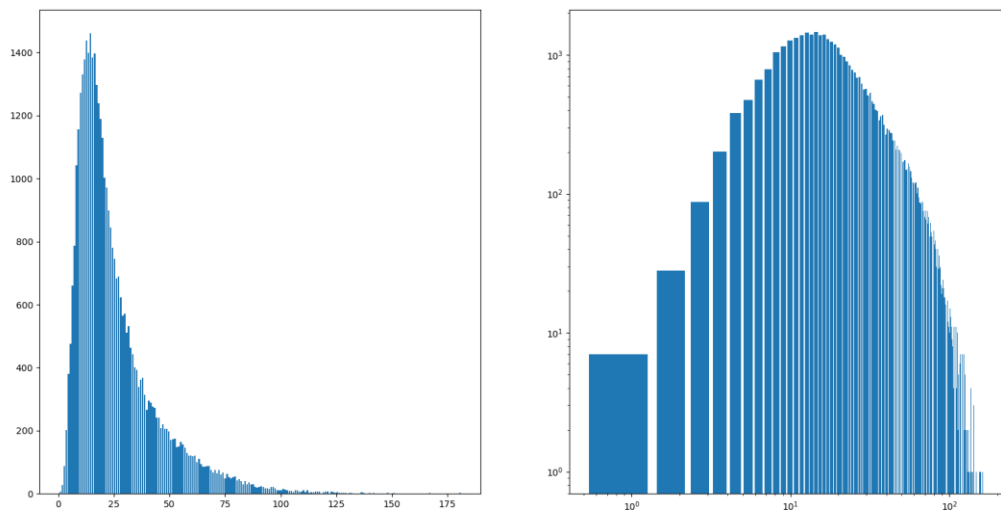
數據圖繪製



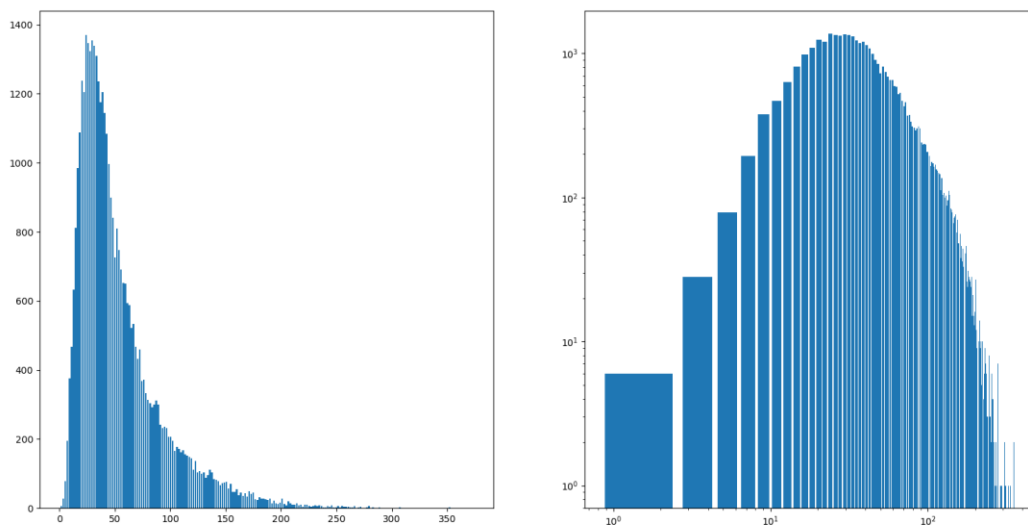
次數 v.s. 地震規模

印出數據後我們發現雖然我們原本的設想地震規模越大發生次數越低，地震規模越小規模應該較多，但實際上在規模及小的部分反而有檢哨次數的跡象，所以地震不符合 power law。原因或許是因為地震具有能量累積後爆發的性質，所以大部分的地震仍有一定的規模，並不是規模越小發生次數越頻繁。

以下為遊戲數據(左半部為 linear scale · 右半部為 log-log scale)

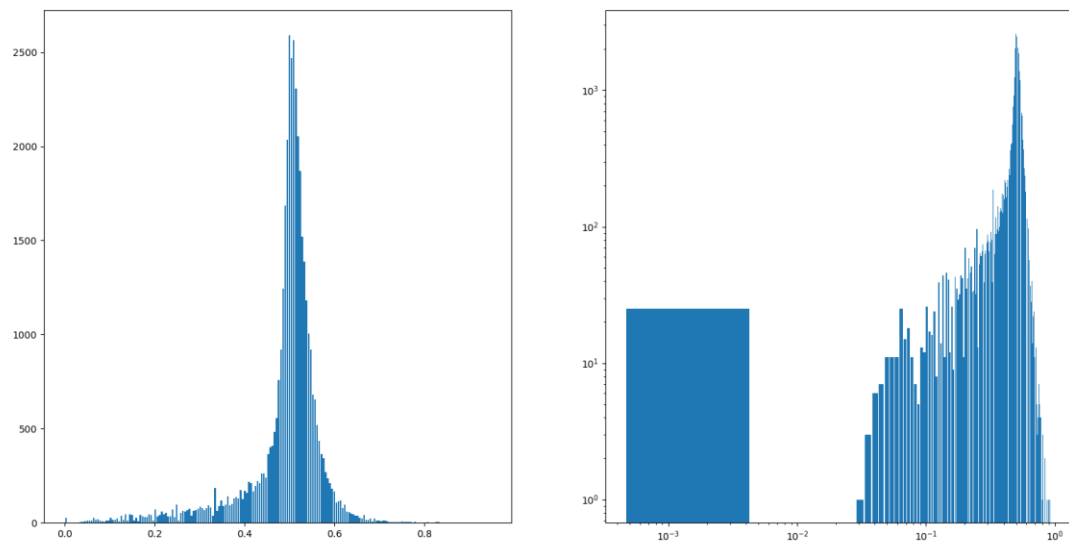


人數 v.s. APM



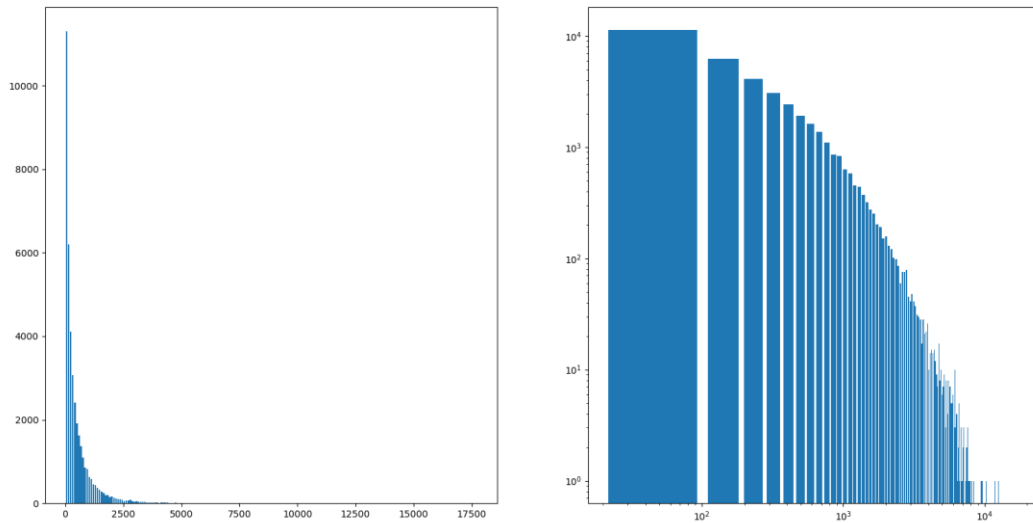
人數 v.s. VS_score

這兩筆數據的型態比較像，他們都具有 long tail 但是在左半邊人數的數值出現下降的現象，所以這兩組數據不符合 power law。

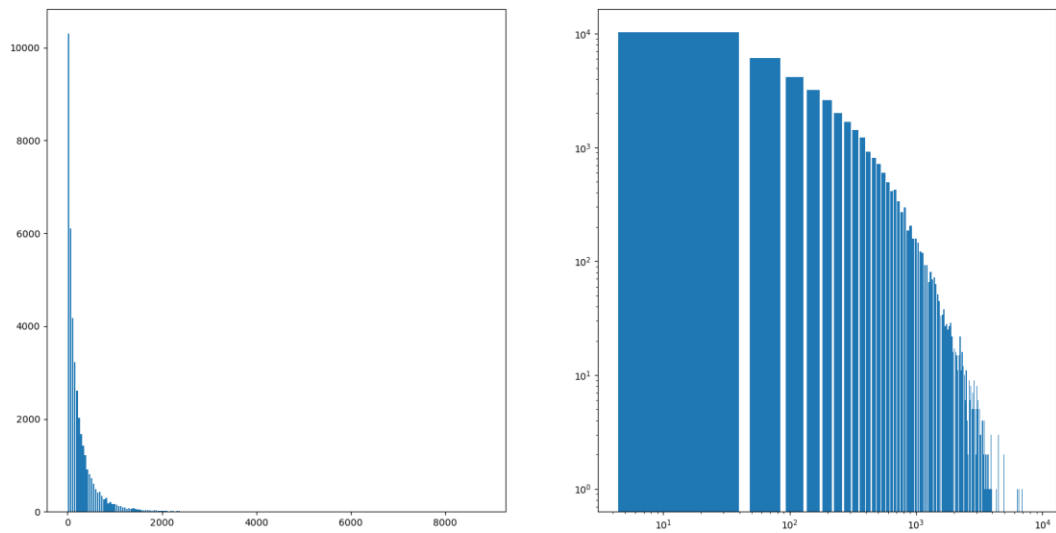


人數 v.s. winRate

勝率的數據明顯不符合 power law，反而比較像常態分布。



人數 v.s. gamesPlayed



人數 v.s. gamesWon

經過思考我們認為勝場的分布應該是受到遊玩場數的分布影響，因為勝場數是包含在遊玩場數中，而這兩組數據都同時具有 power law 的特徵，有 long tail 且在 loglog scale 的圖中，大致上有向右均勻下降的趨勢，雖然以 10 為底還是有一點弧度。

推論

在其他遊戲方面可能也是這樣，遊玩次數越多，其對應的人數也越少。

參考連結

<https://github.com/hickatheworld/tetrio.js.git>