github: <https://github.com/graduatecrisis/CSX-Data-Science/tree/master/final>

11/29

* 決定主題：steam遊戲重分類並標籤
  + 價格標籤：找出適當的價格區間
    - 定價
    - 資料來源：steamspy.com
  + 評論標籤：
    - 同時考慮評論分數與評論數量
    - 專業評論與玩家評論分開看
    - 資料來源：
  + 遊戲類型標籤：
    - 利用發行前新聞稿、遊戲大綱內容進行重點擷取
    - 需要再思考標籤的可能內容、資料來源
* 先利用2012-2017的資料做標籤並分組，最後用2018的資料做預測組
* 先找2012-2017的list https://steamspy.com
* 找出價格分布的模式：彭
* 找出可能的應用程式或奇怪的東西
  + 12, 13, 14：洪
  + 15, 16, 17：黃