Google play app

資料探勘與分析

第29組

組長:

簡浚丞 404412131 61

組員:

黃子銓 404411687 38

吳芊儀 404411299 21

一、本文

我們使用的資料集是<https://www.kaggle.com/lava18/google-play-store-apps/downloads/google-play-store-apps.zip/6>

內容包含googleplaystore中所有的App資料。

而我們要分析的是Googleplaystore中所有app資料集

的category分類數量統計。



程式碼:



而步驟如下:

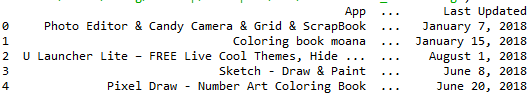
此開放資料總共有12個類別，取出我們想要的類別

googleplay = pd.read\_csv("google-play-store-apps/googleplaystore.csv",usecols = [0,1,2,3,4,5,6,7,8,9,10])

(這裡取了10類，不過我們只會用到Category)。

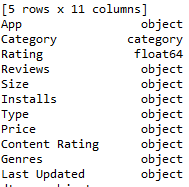
googleplay.head()

1.列出資料集的開頭



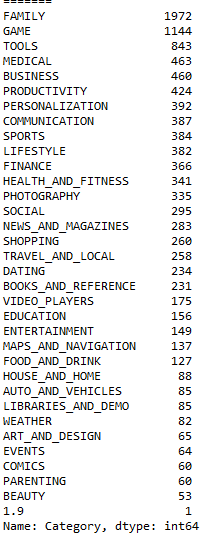
googleplay.dtypes

2.列出資料集的分類



googleplay.Category.value\_counts()

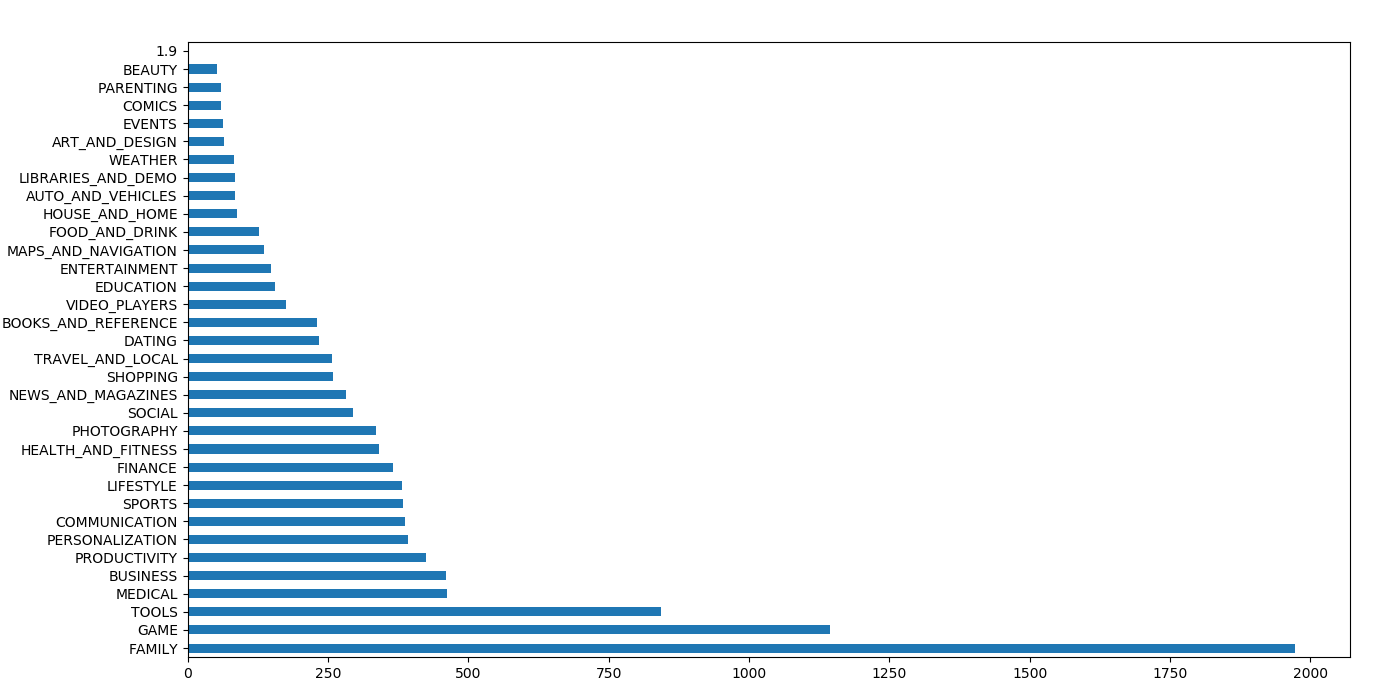
3.將Category中的資料分開計算並計算數量



4.繪製圖表

分類出Category各個項目的統計之後

將各個統計項目數量繪製成圖表



二、結論

Google app store裡面

最熱門的分類還是屬於家庭分類的app，

而我們的猜測是因為每個人都有家庭，

因此家庭類的app會是市場最大的一塊。

而其次當然是目前手遊當道的Game分類了，

畢竟家用主機市場逐漸在萎縮，而手遊正是新世代玩家的新紅海。

三、參考文獻<https://blog.csdn.net/NockinOnHeavensDoor/article/details/80565764>

<https://blog.csdn.net/lin00jian/article/details/51496893>

<https://blog.csdn.net/Jinlong_Xu/article/details/70175107>

<https://stackoverflow.com/questions/33483707/python-pandas-dataframe-head-displays-nothing>

<https://dataholic.wordpress.com/2016/10/18/%E7%AC%AC%E4%B8%80%E7%AF%87%E7%B6%B2%E8%AA%8C%E6%96%87%E7%AB%A0/>

<https://www.cnblogs.com/alanma/p/6877118.html?fbclid=IwAR20JbrwcU40rKWKWmAleIMZ5URTmYQYAMwxReu6IEwblaN9_avRMtXoDxk>

四、組員心得

吳芊儀:

時間過得很快，轉眼間這個學期也過了大半。這一學期，我修了資料探勘這門課，對Python有更深入的了解，整體來說我認為這堂課使我獲益良多。

老師在講課時會給我們看不同的程式範例，所以，在短時間內，我們會對資料探勘有簡單且整體的認識。

期末要繳交分組作業，實際從動手操作的過程當中，更加深入了解Python資料分析。做中學，先整理，後拆解。在完成報告的過程中，我得到許多更實際的收獲，是與直接從課堂上所得到的知識不同的。比如說要怎麼將主題分工、以及怎麼讓主題有個明確的輪廓，跟以往做個人報告時都不太相同。在分組準備的過程中花了很多時間，也做了很多討論與修改，雖然最後演示的結果不見得完美無缺，但是這對於準備踏入職場且尚缺乏社會歷練的我們，是一項很寶貴的經驗，也是一段很有趣的回憶。

對於自己在這堂課的表現，我深知自己還有很多需要加強改進的地方。也謝謝老師那麼有耐心，引導我們對資料探勘有更多的認識。

簡浚丞:

這堂課我因為老師的關係我做了這個報告，因而這樣可以跟組員增進感情，在各方面來說都是很好的活動，我們在找資料的時候我們非常的認真在網路上找了很多資料分析有關的資料，我們使用python 的

Jupyternotebook 實作，訓練我們使用不同的東西的重要性也讓我學會看別人的資料開始做學習，學習到了很多。

而且我們也熬夜趕出這個報告結合了大家的意見跟資料集大成而成的曠世巨作。

總而言之我資料分析出來的東西雖說沒有超級好，但我跟我的組員們絞盡腦汁想出來的，大家一起成長一起努力出來的結果，對於這次的報告大家都非常努力非常值得。整組的表現都在我的預期上。

黃子銓:

修到這門課的時候，才發現我已經快要畢業了，讓人覺得有些惋惜怎麼沒有早點接觸到這門課。

一直以來，我都專注於網頁的前端後端，在去年的專題結束後，已經一年沒有跨及python相關的技術。這次回來使用後才發現，python還是保持著他一直以來的特點:簡單，高效。可以將程式非常有效率地寫出來。再加上他非常廣泛的library造就了非常豐富多元的應用。

這次打code的經驗也是我第一次使用matlab繪圖，第一次覺得自己在一個資料及中分析出有意義的資料是一件多麼有趣的事情。當然，對於缺乏實作經驗的我們，其實還有非常多東西必須要努力去琢磨，這次是一個很好的經驗(還順便讓我發覺原來ipython和python的用法有點不一樣)。