商業分析HW1

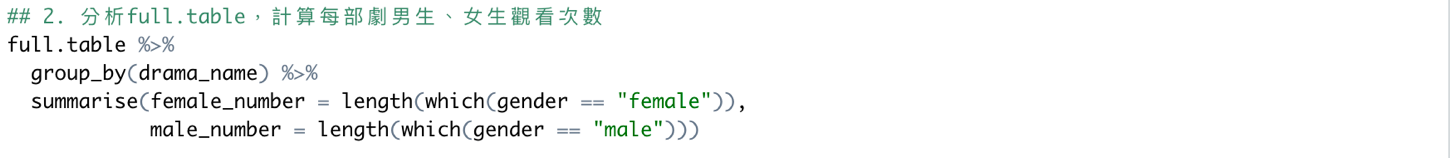
105305072 企管四 許惠甄

一、

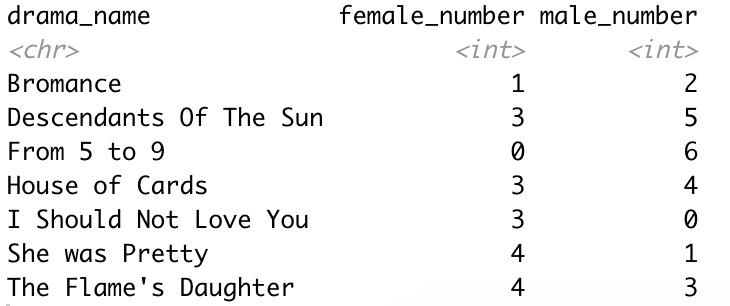
1. 將 watch.table 與其他兩個報表合併為full.table。



2. 分析full.table，計算每部劇男生、女生觀看次數 (group\_by(), summarize(), length())



呈現結果如下：

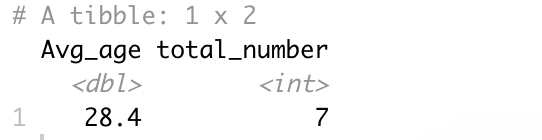


從表格中可看出《From 5 to 9》的男性觀看次數最多，而《She was Pretty》與《The Flame’s Daughter》是女性觀看次數較多的兩部劇。

3. 找出用Android系統的，針對這類客戶進行分析（例如平均年紀、性別、來自哪裡）。

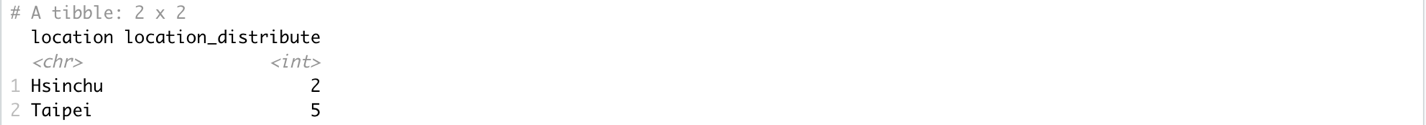


呈現結果如下：



從表格中可看出使用Android系統的平均年齡為28.4歲，佔原先39筆資料中的7筆。





從以上表格中可看出使用Android系統的共有兩位男性，一共觀看《From 5 to 9》5次及《House of Cards》2次，居住地區為新竹及台北。

4. 針對台北男性這類客戶進行分析。

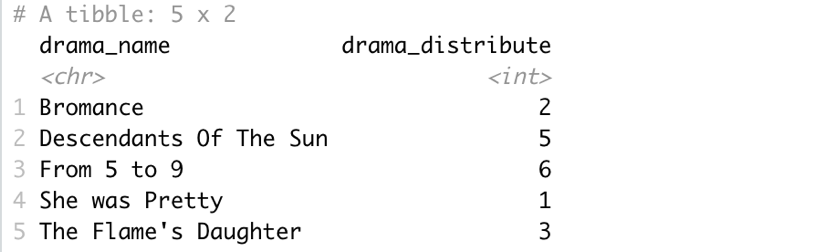
呈現結果如下：



從表格中可看出台北男性觀看者的平均年齡為26.1歲，佔原先39筆資料中的17筆。



呈現結果如下：

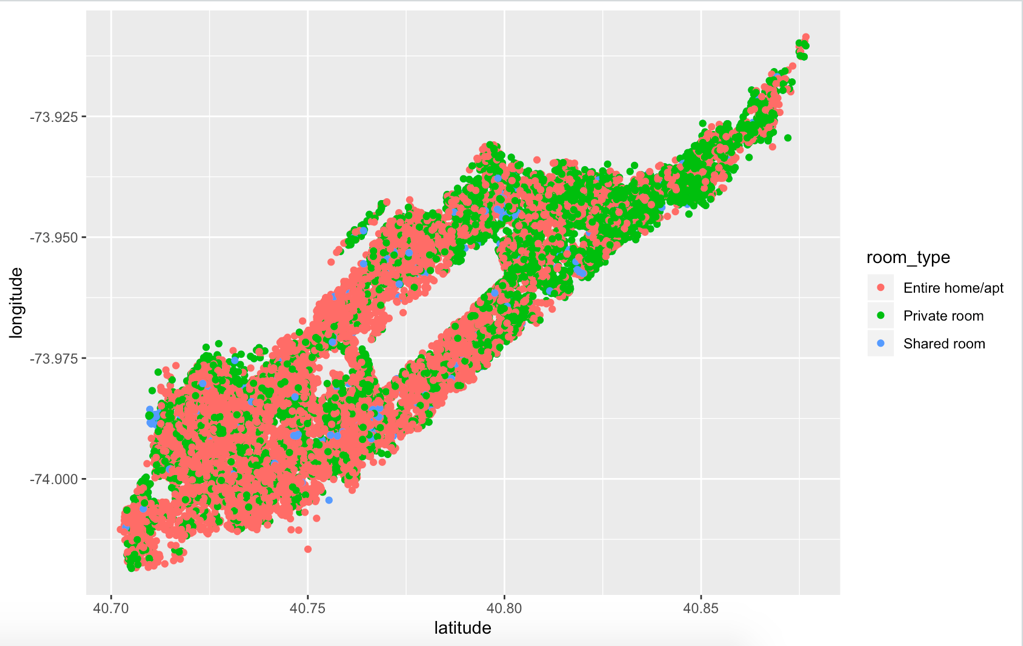


從表格中可看出台北男性觀看次數最多的影集是《From 5 to 9》，次數最少的是《She was Pretty》。

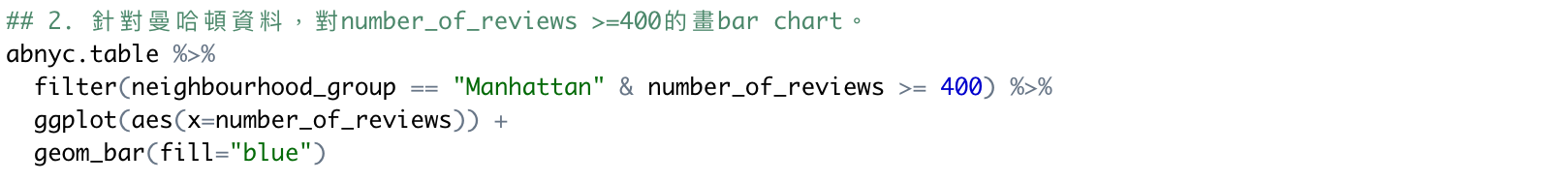
二、

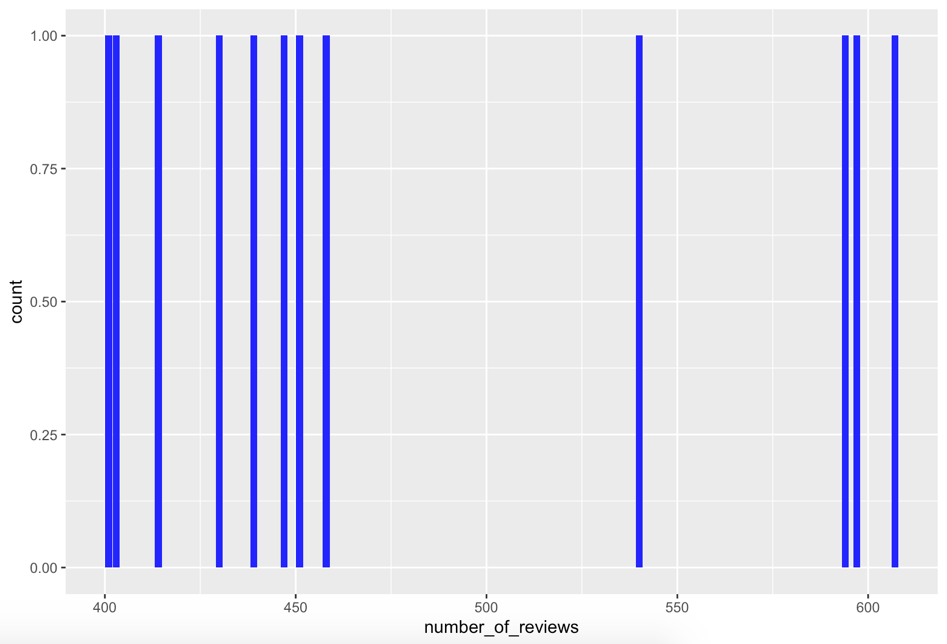
1. 找出 neighbourhood\_group == "Manhattan"的資料，利用ggplot畫經緯度的scatter plot。





2. 針對曼哈頓資料，對number\_of\_reviews >=400的畫bar chart。

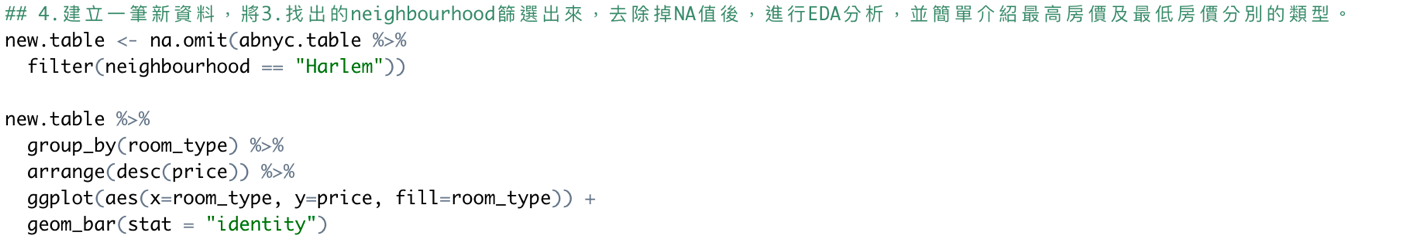


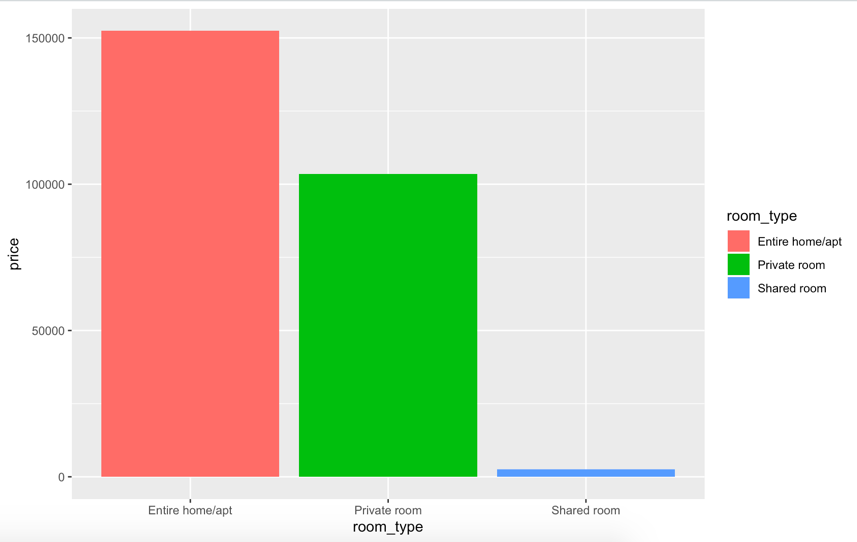


3. 針對曼哈頓資料，number\_of\_reviews >=400的中，哪個neighbourhood擁有最多number\_of\_reviews。



結果顯示「Harlem」區域的房子有最高的Review次數。

4. 建立一筆新資料，將3.找出的neighbourhood篩選出來，去除掉NA值後，進行EDA分析，並簡單介紹最高房價及最低房價分別的類型。



結果顯示最高房價是「Entire home/apt」類型，而最低房價是「Share room」類型。