

IOT 物聯網使用記錄分析： 偵測異常的使用狀態

姚璇亨 Henry Yao

2021.11.13

偵測異常的使用狀態

偵測異常的使用狀態

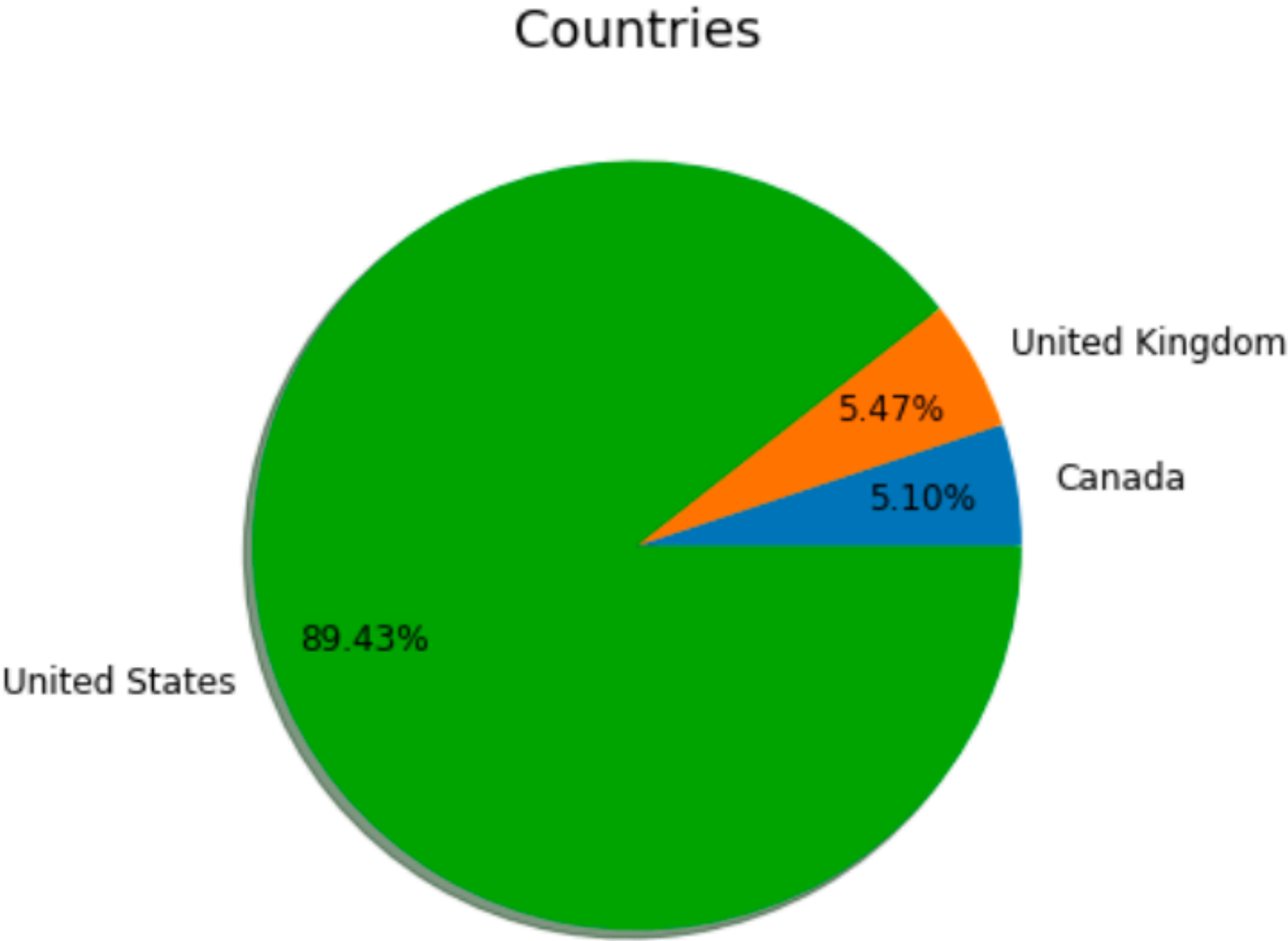
1. 資料概況

原始資料筆數	8,347,801
資料清理後筆數	6,813,301
起始日期	台灣時間 西元 2019 年 8 月 31 日
結束日期	台灣時間 西元 2019 年 9 月 29 日
資料天數	30 天
每日平均資料筆數	227,043.37

偵測異常的使用狀態

2. 國家分佈

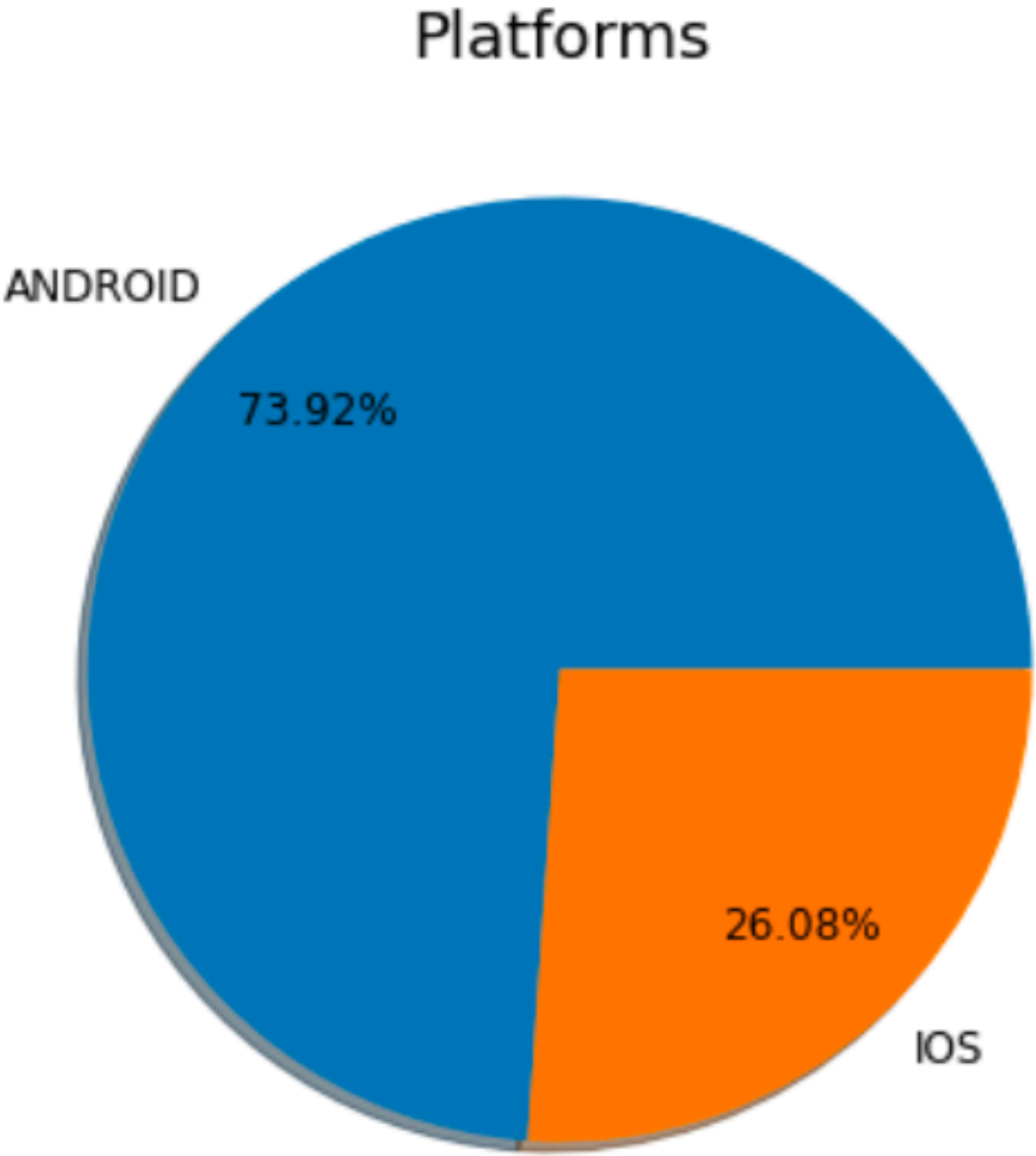
Canada	347518
United Kingdom	372597
United States	6091186



偵測異常的使用狀態

3. 平台分佈

ANDROID	5034890
IOS	1776411



偵測異常的使用狀態

4. 關鍵指標定義：「期間內 live 連線的成功率」

考量數據紀錄錯誤者（first_video_frame > 10000）仍有建立影像連線，故將連線「成功」定義如下：

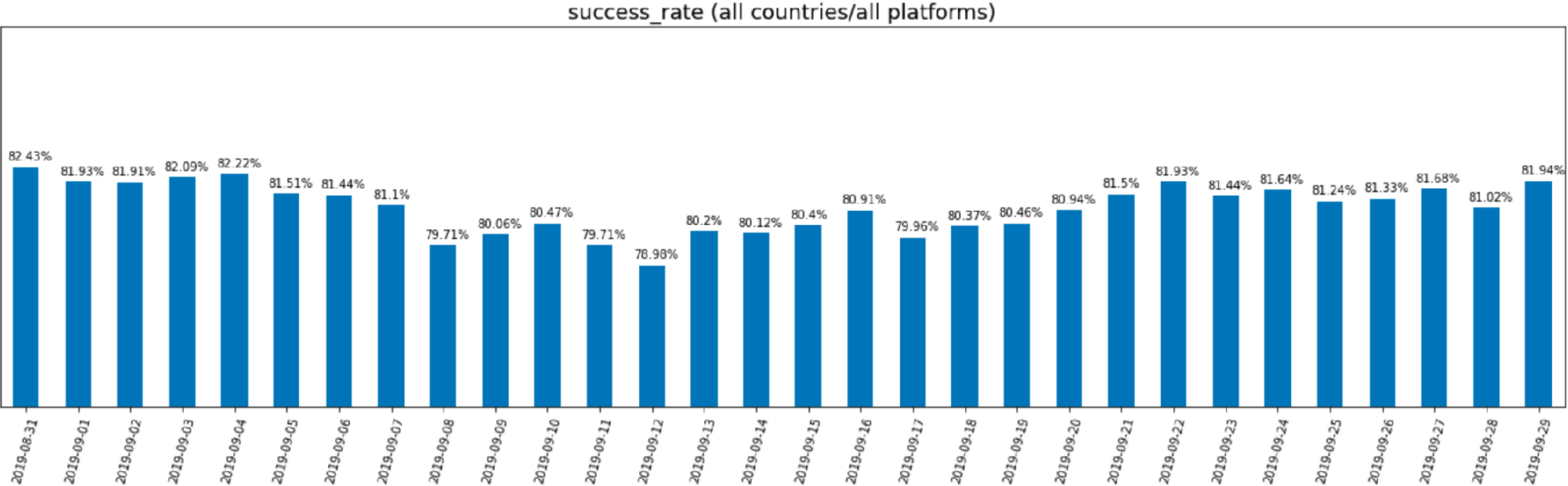
first_video_frame 不是 -1

期間內 live 成功率，則以以下定義來表示：

期間內 live 成功率 = 期間內成功 live 影像連線總數 ÷ 期間內 live 總數 × 100%

偵測異常的使用狀態

5. 每天 live 連線的成功率分佈

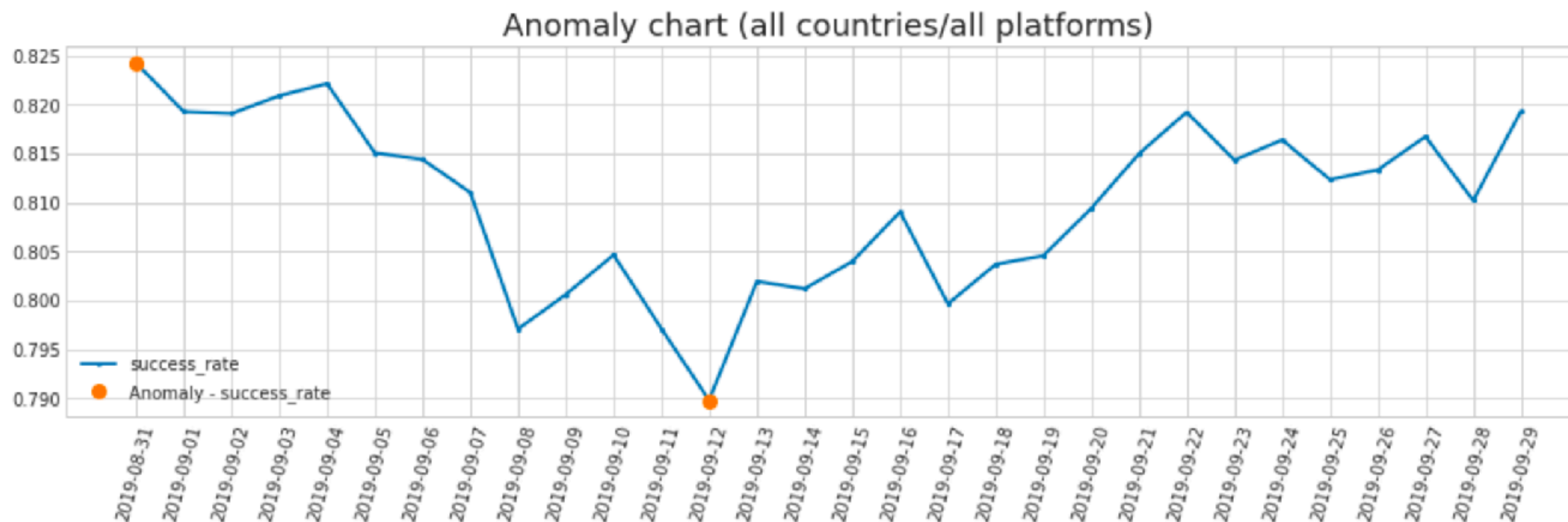


偵測異常的使用狀態

6. 尋找趨勢異常

採用時間序列異常偵測法的 **Quantile Anomaly Detection**（分位數異常檢測）

- 將正常上下限分別訂為 99%、1% 百分位數以偵測範圍外的離群值
- 發現 **2019-08-31** 的成功率異常地高、**2019-09-12** 的成功率異常地低
- 與前一日比較，成功率變化幅度各為 **0.005%**、**-0.007%**

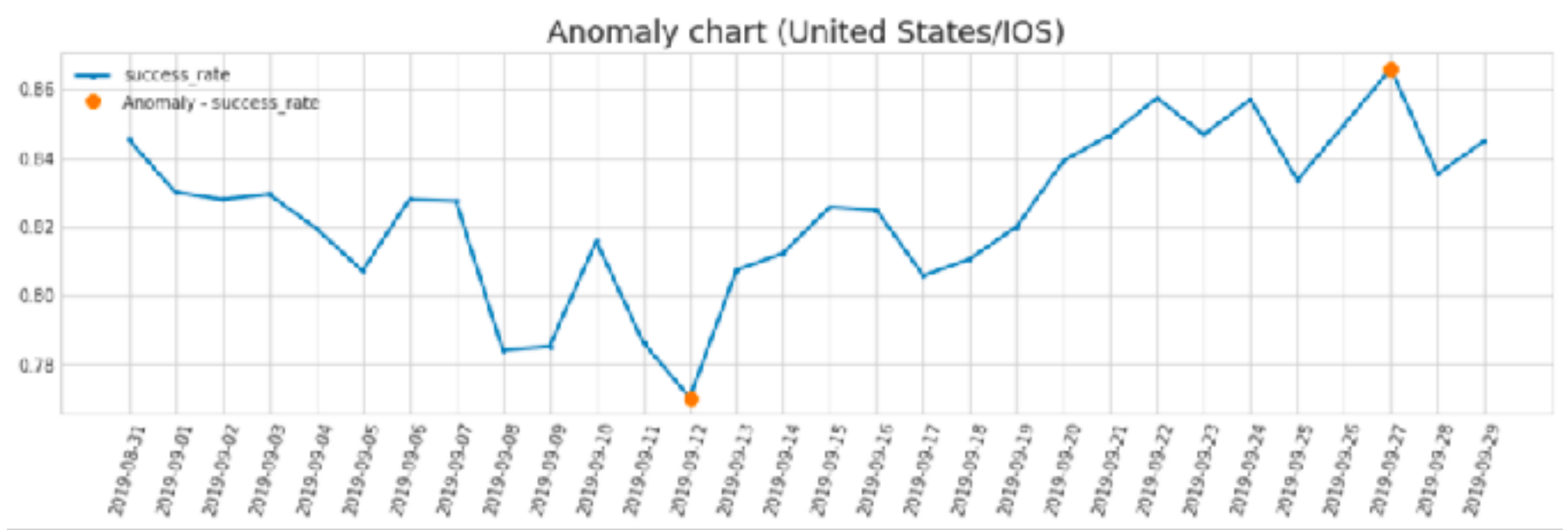
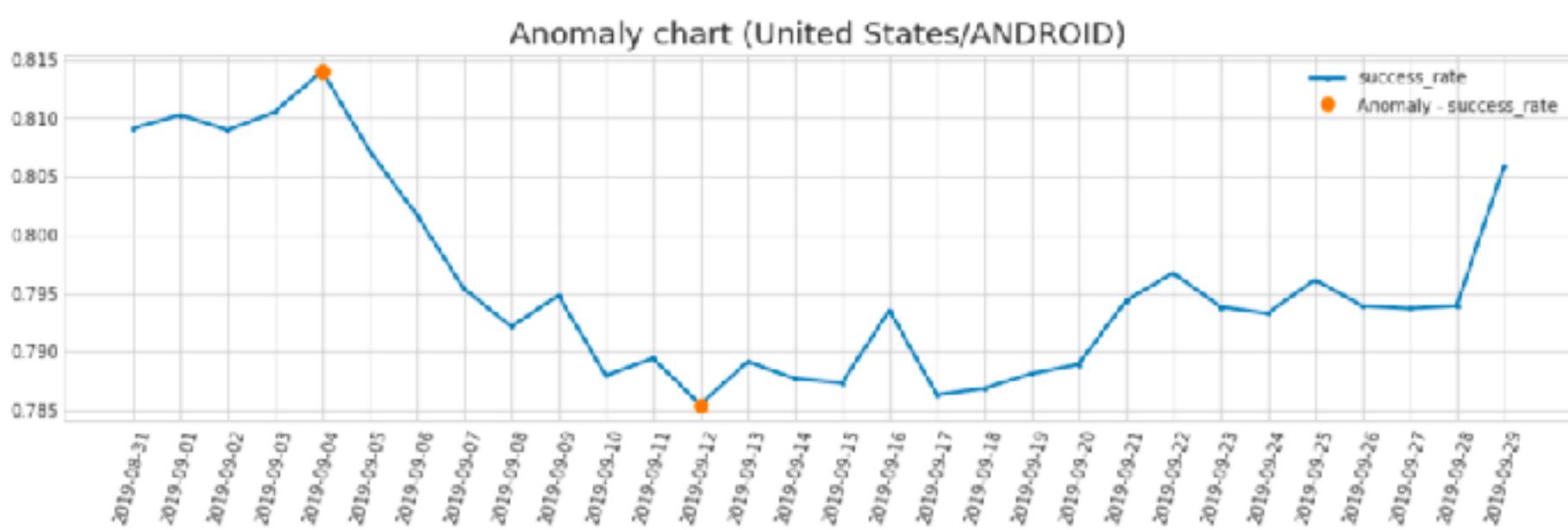
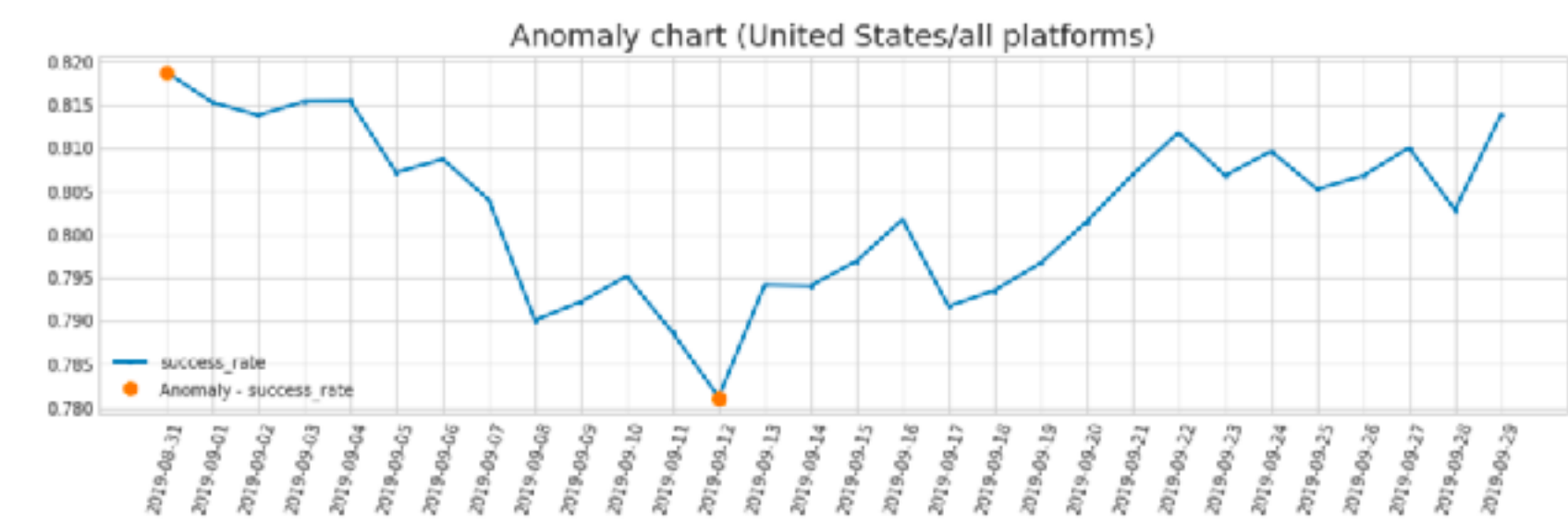


偵測異常的使用狀態

7. 檢查不同「國家」、「平臺」的「每天 live 影像連線的成功率」

將資料透過個國家、平台維度分割後，針對每個組合再度分別進行異常偵測

- 成功率異常地低的 2019-09-12 日期，在 United States 的每個平台組合皆被偵測出來，其他國家的組合皆無
- 仍懷疑異常確在，且發生原因可能與國家有關、與平台無關



偵測異常的使用狀態

8. 尋找其他可能乘載本次趨勢變化要素的欄位資訊

分別檢查「candidate_type」、「network_type」、「xmpp_type」、「camera_pipeline」四個欄位

- 篩選出日期為 2019-09-12、國家為 United States 的資料
- 計算以上欄位各個值在當天所有連線數中所佔的比例
- 接著再與前一天（2019-09-11）的資料做比較，計算出變化幅度
- 初步檢視，發現 candidate_type 的 local/local 有最高的減少幅度，降低了 4.26%

event_date	candidate_type	
2019-09-12	local/local	-0.0426

- 初步假設異常原因：local/local 這個理論上最穩定連線方式的減少，造成當天 live video 連線成功率偏低
- 待透過觀察更長期間資料以及統計檢定來進一步驗證假設