



VD動態資訊和 路況即時路況動態 的資料關聯性驗證

第13組:
楊守荃、黃浚瑀、程品諭

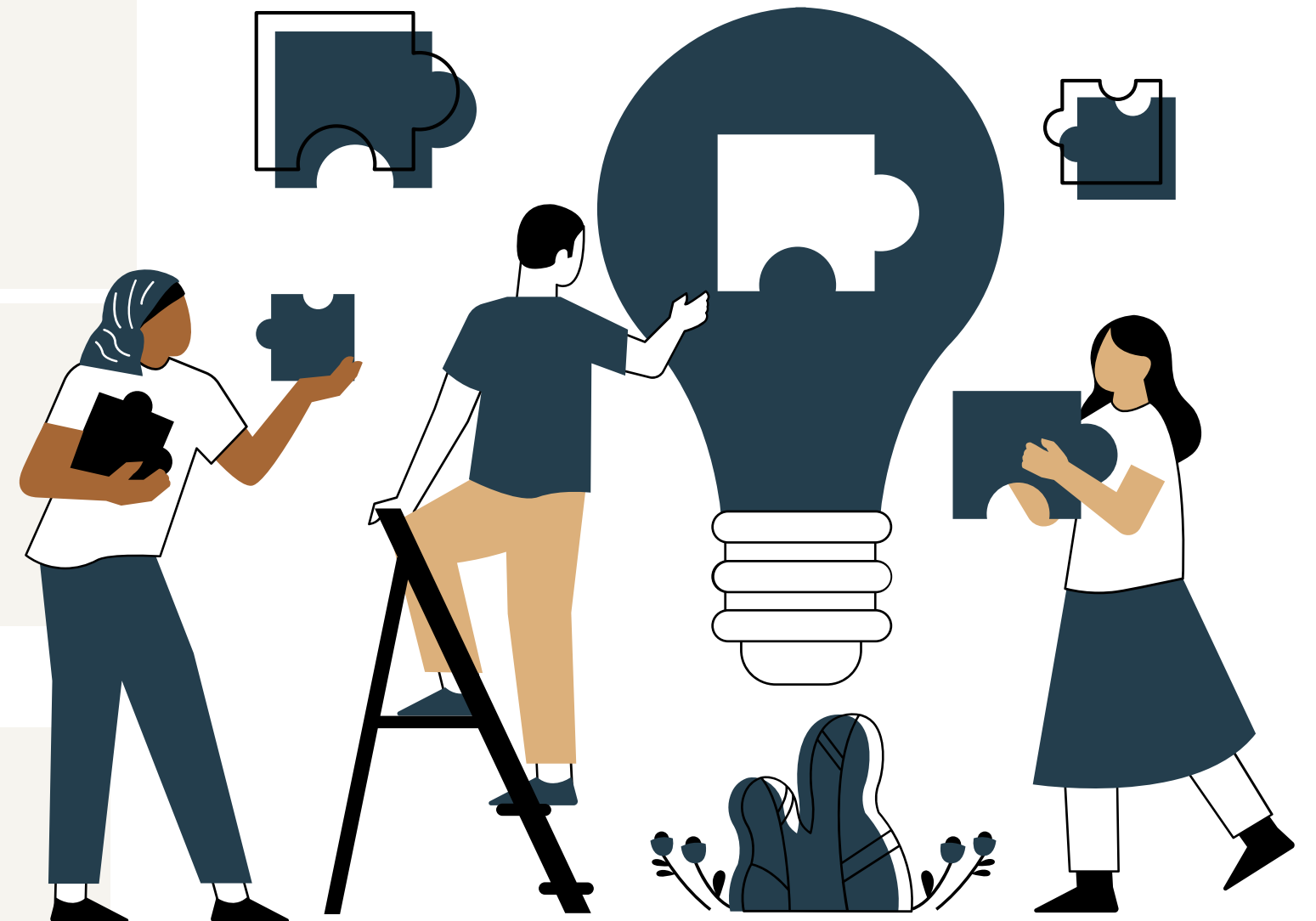
目錄

1 資料來源清洗與整理

2 資料來源視覺化及
程式碼介紹

3 結論與分析

4 額外功能或創意





資料來源

所選資料來源



1. 高公局歷史資料庫VD 動態資訊(v2.0)

網址：<https://tisvcloud.freeway.gov.tw/history/motc20/VD/>

VD 動態資訊 (v2.0)	1min	xml	./history/motc20/VD/YYYYMM DD/VDLive_HHmm.xml.gz	1分鐘定時更新及統整單一VD的動 態偵測值		
-------------------	------	-----	---	--------------------------	---	---

2. 高公局歷史資料庫路段即時路況動態資訊(v2.0)

網址：<https://tisvcloud.freeway.gov.tw/history/motc20/Section/>

路段即時路況動態 資訊(v2.0)	1min	xml	./history/motc20/Section/YYYY MMDD/LiveTraffic_HHmm.x ml.gz	提供機關發布路段或基礎路段的即 時路況動態資料		
----------------------	------	-----	---	----------------------------	---	---

所選空間範圍

1..Link ID:0000100108900J
Section ID:0056
國道一號北上竹北路段



2. Link ID:0000100102700F
Section ID:0024
國道一號北上三重路段

所選空間範圍

3. Link ID:0000100109860

Section ID:0060

國道一號北上寶山路段



LINK ID取得方式

使用交通資訊基礎路段編碼查詢系統
網址：<https://link.motc.gov.tw/>



點選所選路段，可得出路段ID



LINK ID取得方式

同時也可去VD資料庫查詢對應的LINK ID

```
37818 <VDID>VD-N3-N-14.610-N-LOOP</VDID>
37819 <LinkFlows>
37820   <LinkFlow>
37821     <LinkID>0000300101420F</LinkID>
37822     <Lanes>
37823       <Lane>
37824         <LaneID>0</LaneID>
37825         <LaneType>2</LaneType>
37826         <Speed>94</Speed>
```


SECTION ID取得方式

在

SectionLink_0000.xml.gz

2023-12-29 02:00

15K

檔案中以Link ID查詢對應的Section ID

```
<SectionLink>
  <SectionID>0304</SectionID>
  <LinkIDs>
    <LinkID>0000300131150D</LinkID>
    <LinkID>0000300131100D</LinkID>
    <LinkID>0000300131000D</LinkID>
    <LinkID>0000300130900D</LinkID>
    <LinkID>0000300130800D</LinkID>
    <LinkID>0000300130700D</LinkID>
    <LinkID>0000300130600D</LinkID>
    <LinkID>0000300130500D</LinkID>
    <LinkID>0000300130420D</LinkID>
    <LinkID>0000300130400Q</LinkID>
    <LinkID>0000300130300D</LinkID>
```

所選時間範圍

- 日期:2023/12/29
- 時間:17:00~20:59
- 時長:4小時(240分鐘)
- 含下午尖峰(17:00-19:00)

資料分析方式

由於VD資料和即時路況資料中的車道速率來源是同向鄰近車道，我們假設速率在兩資料來源中有很高的一致性。

資料分析方式

目標：

1. 求出該路段VD以及即時路況資料每分鐘的平均速率
2. 使用python將平均速率用折線圖畫出來
3. 使用excel 進行平均速率關係圖的迴歸分析
4. 比較兩組圖表，觀察其關聯性

程式碼解說

程式碼解說-爬蟲

```
# 偽裝成瀏覽器
headers = {
    "User-Agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/120.0.0.0 Safari/537.36"}
```

- 由於有些網站會防止被以爬蟲的方式取得資料，且高公局的網站中也有提到:若使用者擷取資料太頻繁，會中斷其連線。
- 為了防止這些發生，我們將headers更改為瀏覽器版本，就能避開這些風險，網站會認為這只是一個瀏覽器在下載而非使用者

程式碼解說-爬蟲

```
with open(f'./vd_compressed_data/{hour}{i}{j}.xml.gz', "wb") as file:  
    # 防止記憶體爆滿，改成以每102400 bytes的資料量讀進來  
    for chunk in response.iter_content(102400):  
        file.write(chunk)
```

- 由於不知道資料量有多大，因此我們以這種比較安全的方式將資料讀進來。
- 記憶體每讀102400 bytes就先寫進去，再換下一個102400 bytes，避免一次將所有東西讀入

程式碼解說-資料處理

我們寫了兩個函式，
以方便我們抓取我們
想要的資料

```
vd_speed_matrix1 = get_vdlive_speed('0000100108900J') # 竹北
vd_speed_matrix2 = get_vdlive_speed('0000100102700F') # 三重
vd_speed_matrix3 = get_vdlive_speed('0000100109860J') # 寶山

live_traffic_matrix1 = get_live_traffic_speed('0056') # 竹北
live_traffic_matrix2 = get_live_traffic_speed('0024') # 三重
live_traffic_matrix3 = get_live_traffic_speed('0060') # 寶山
```

程式碼解說-資料處理

```
for lane in lanes:
    speed_by_lane = int(
        lane.find("ns:Speed", namespaces=namespace).text)
    if speed_by_lane != 0:
        temp.append(speed_by_lane) # speed_matrix[車道]
for a in temp:
    sum = sum + a
avg = sum / len(temp)
speed_matrix.append(avg)
```


```
<Lanes>
  <Lane>
    <LaneID>0</LaneID>
    <LaneType>2</LaneType>
    <Speed>80</Speed>
```

由於一個lanes中有許多個lane，我們將所有lane中的平均速度抓下來後，再做一次平均

程式碼解說-繪圖

利用enumerate這個函數，幫我們分類我們想要畫的是哪張圖，並根據資料的結果來繪出六張折線圖

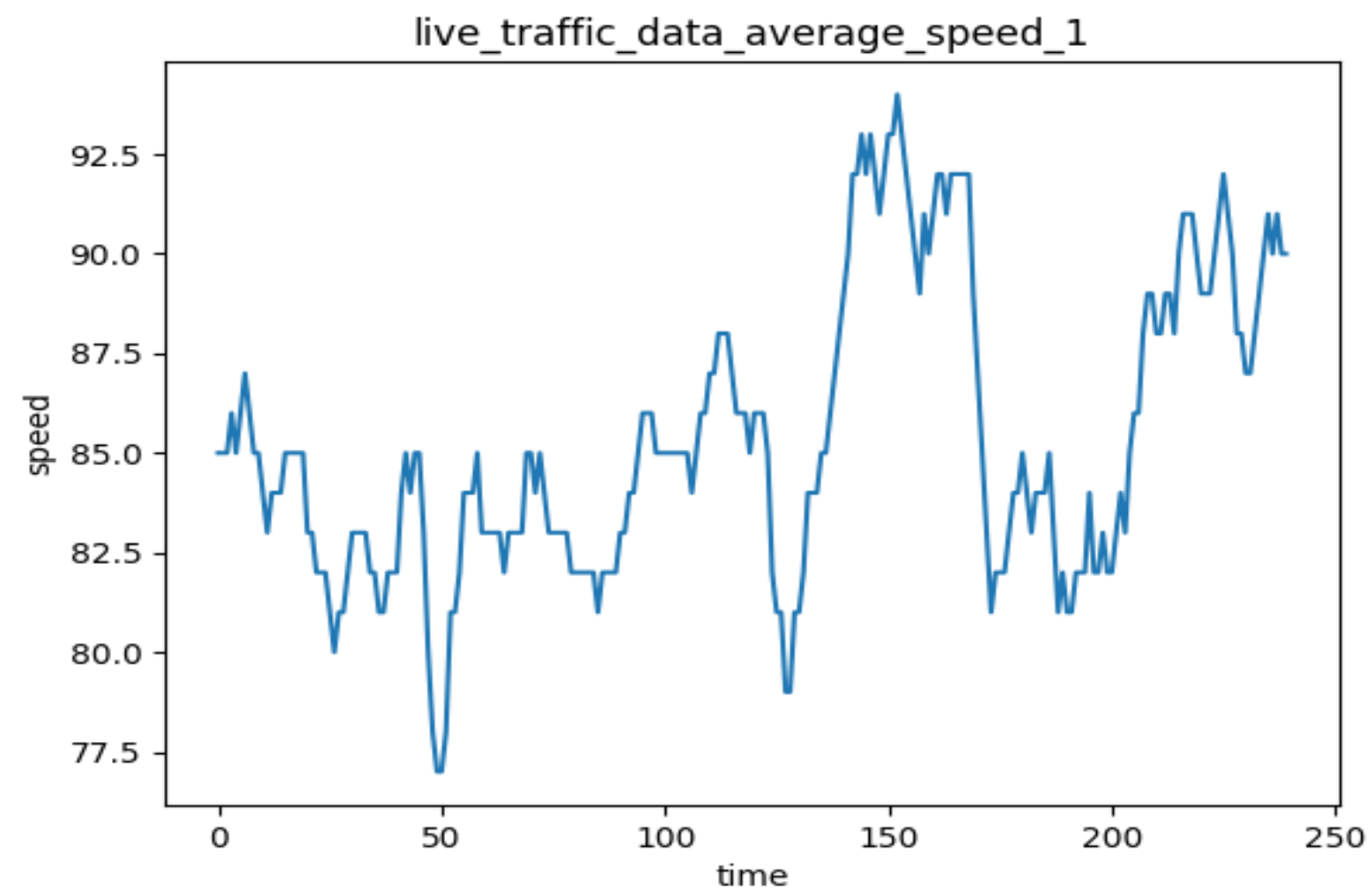
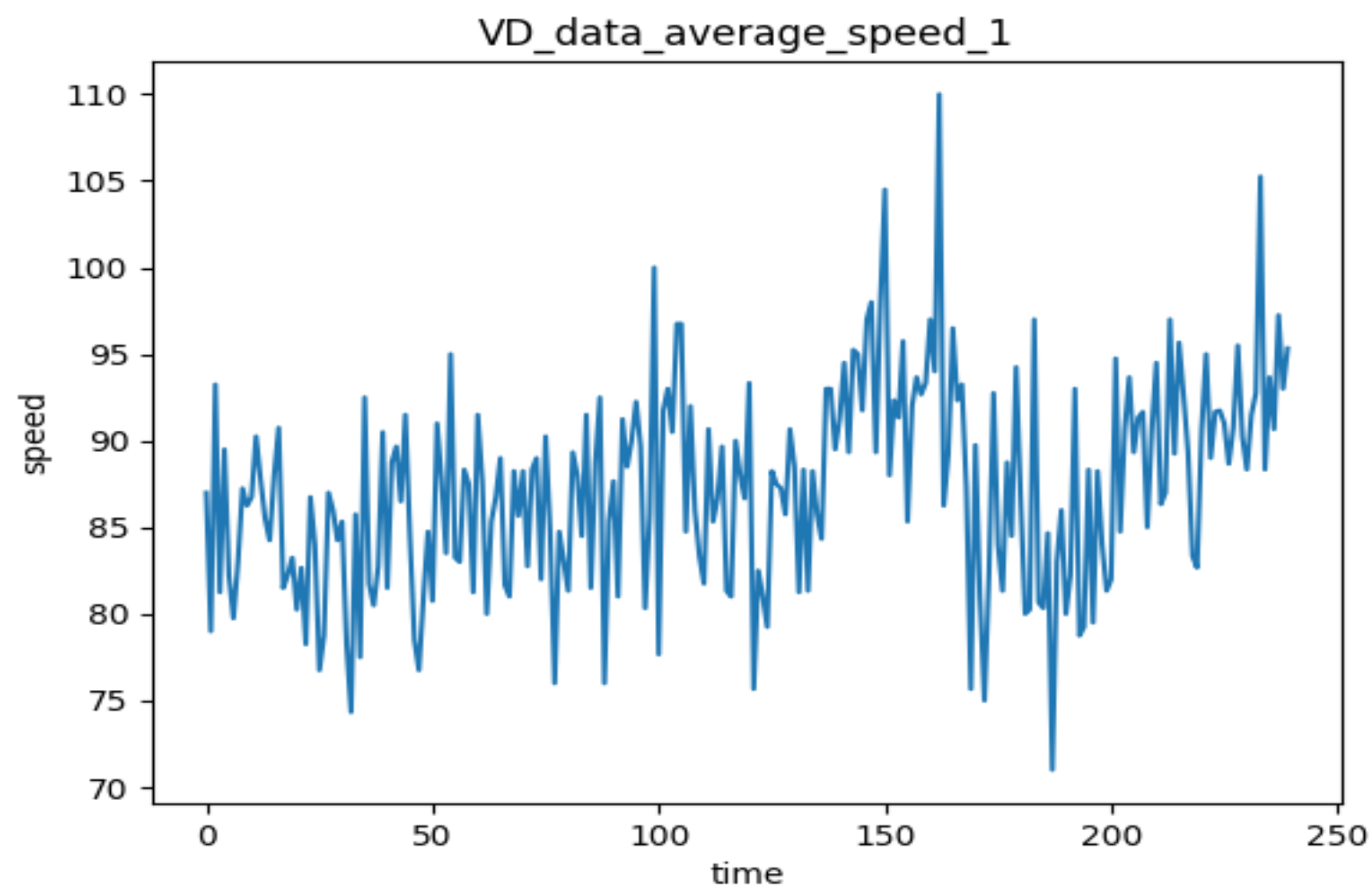
```
for i, matrix in enumerate([vd_speed_matrix1, vd_speed_matrix2, vd_speed_matrix3]):  
    plt.plot(range(len(matrix)), matrix)  
    plt.xlabel('time')  
    plt.ylabel('speed')  
    plt.title(f'VD_data_average_speed_{i+1}')  
    if i == 0:  
        district = '竹北'  
    elif i == 1:  
        district = '三重'  
    else:  
        district = '寶山'  
    plt.savefig(f'VD_speed_{district}.png')  
    plt.clf()
```



時間與速率關係圖

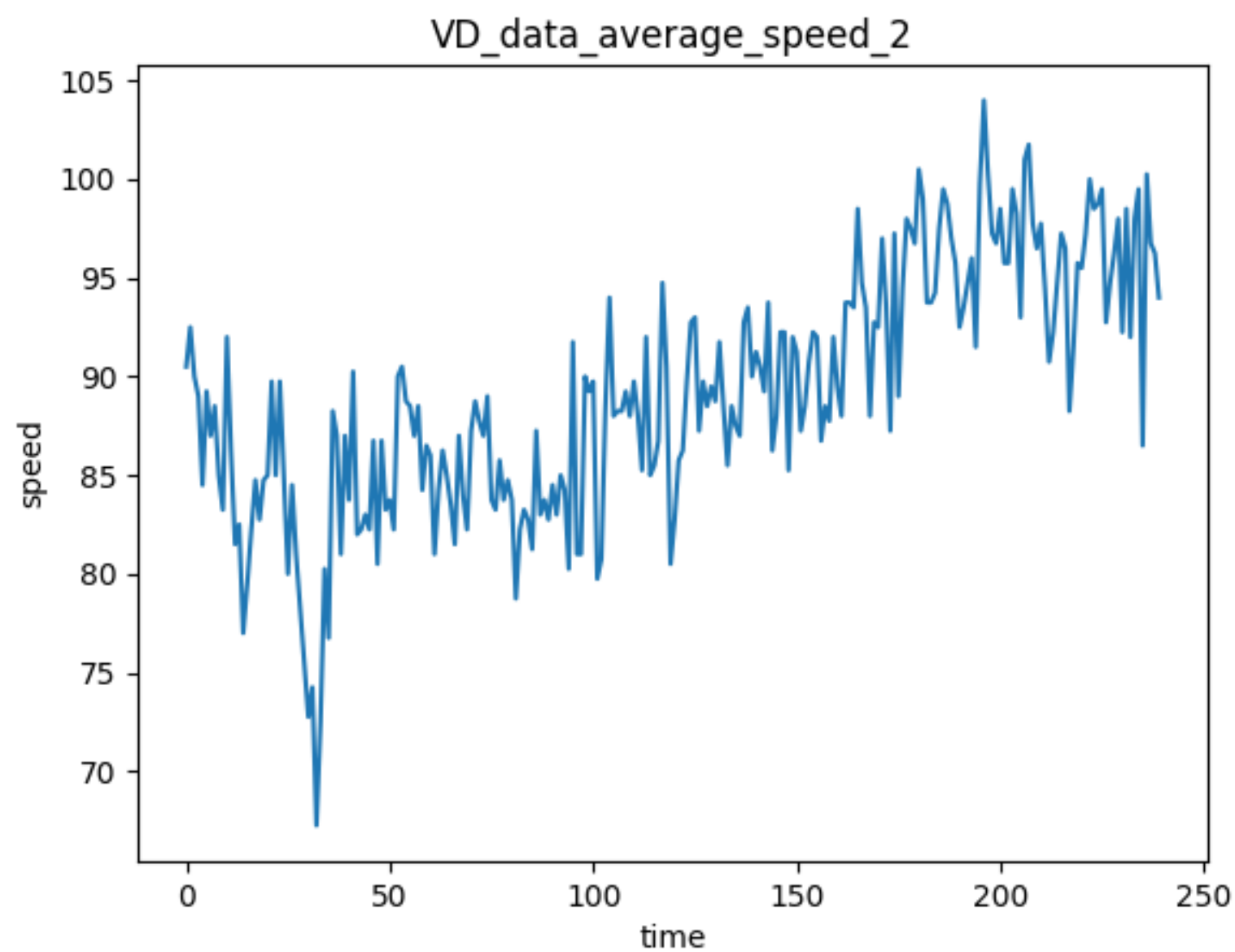
時間與速率關係圖

1. 國道一號北上竹北路段



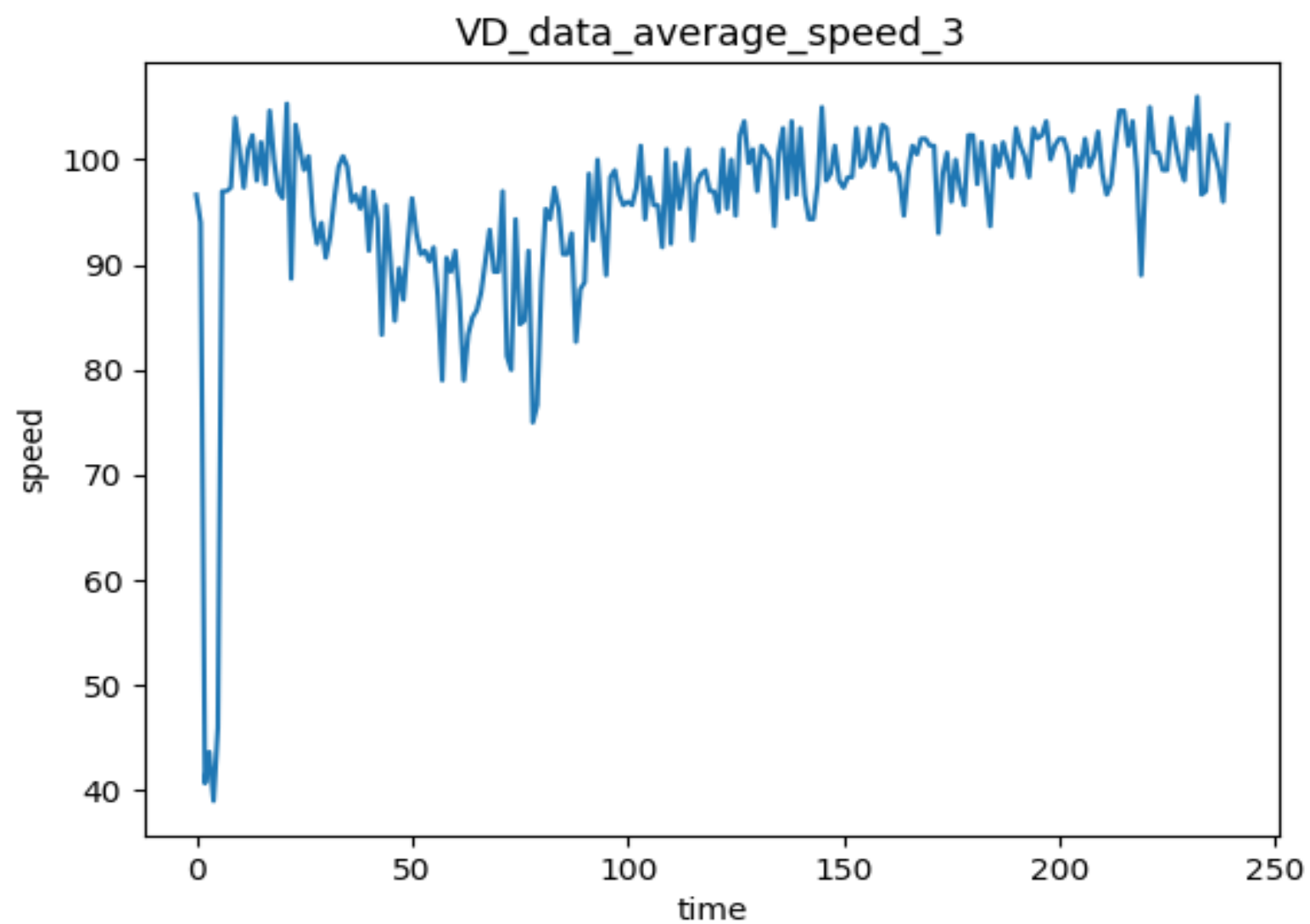
時間與速率關係圖

2. 國道一號北上三重路段




時間與速率關係圖

3. 國道一號北上寶山路段



時間與速率關係圖

以巨觀的角度而言，波形吻合度最佳的是竹北路段，再者是三重路段，最後則是寶山路段。
即時路況資料的速度相較VD資料而言偏低。



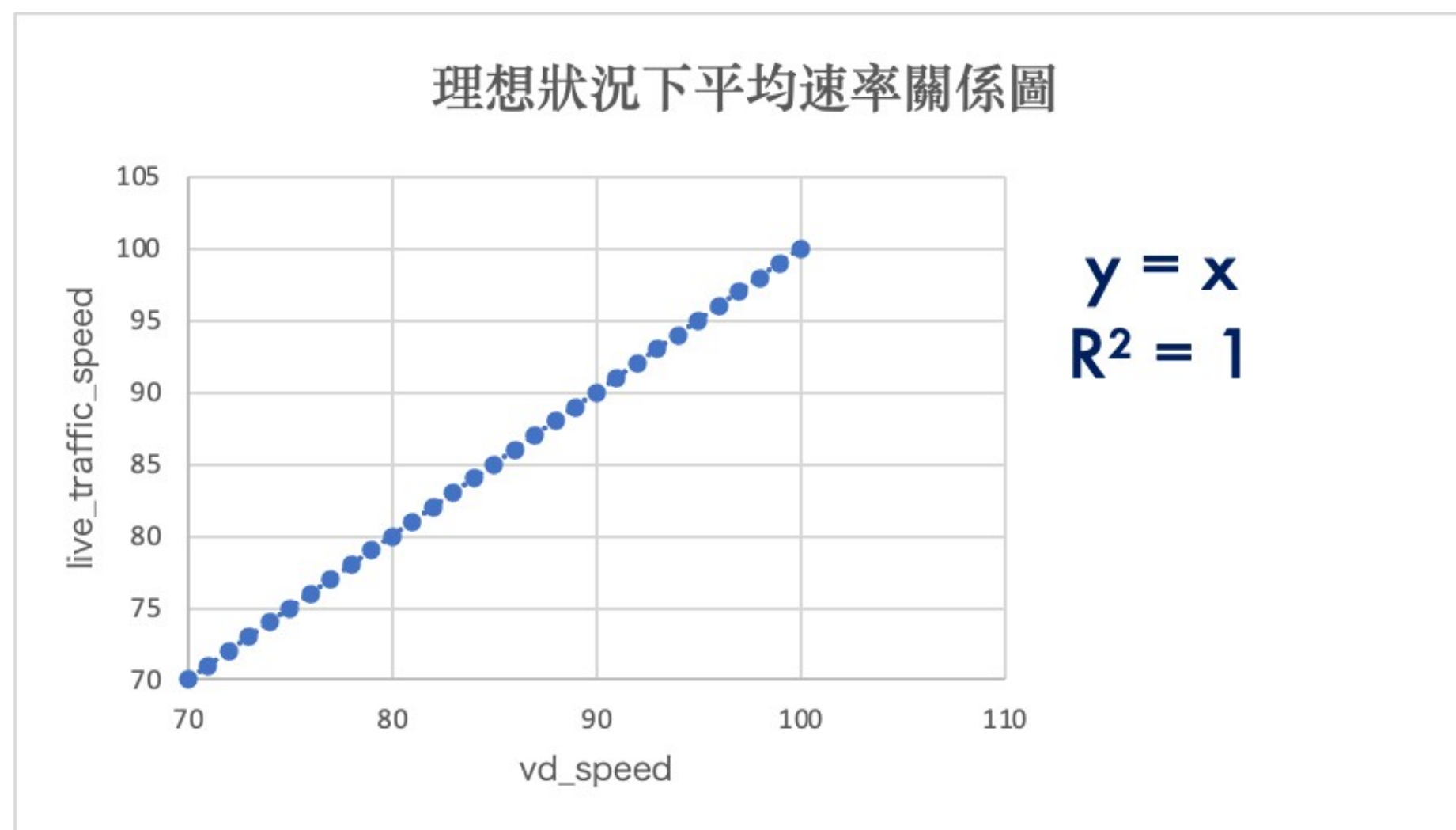
平均速率關係圖

平均速率關係圖

為了觀察同路段之VD資料及即時路況資料在速率上的相關性，我們用Excel繪製平均速率關係圖。

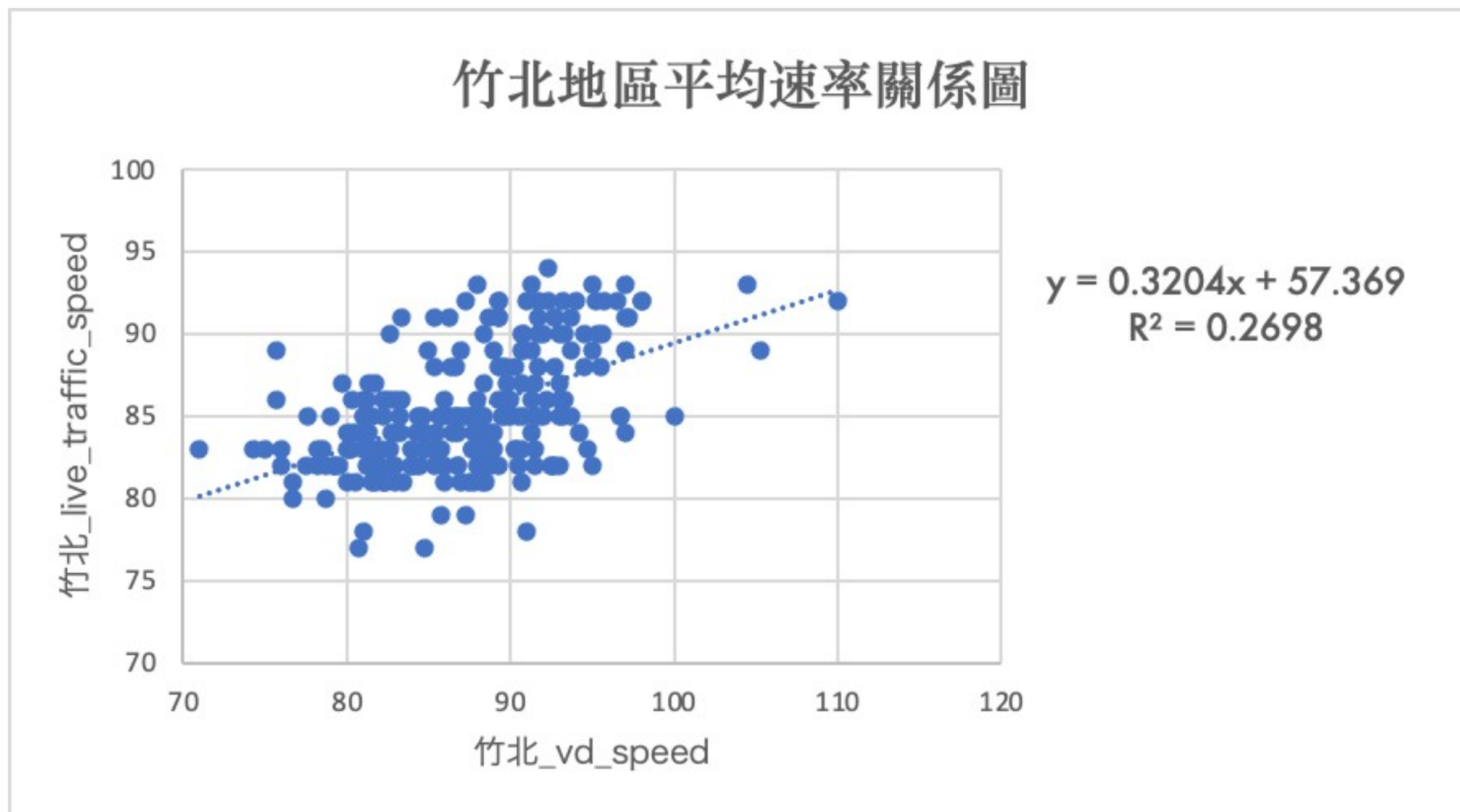
理想情況下的迴歸分析模型

- 斜率為1
- 截距為0
- 決定係數為1



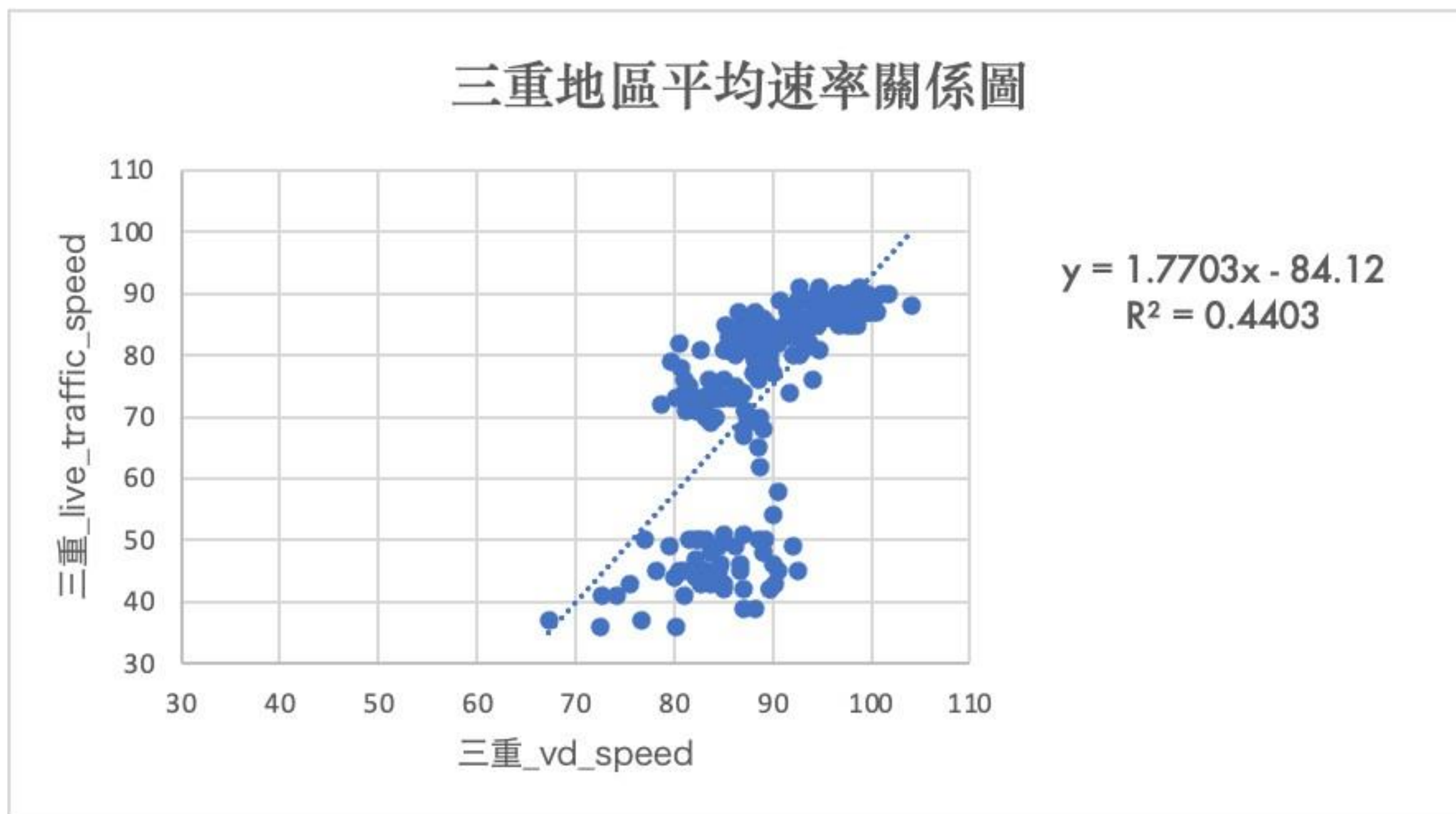
平均速率關係圖

1. 國道一號北上竹北路段



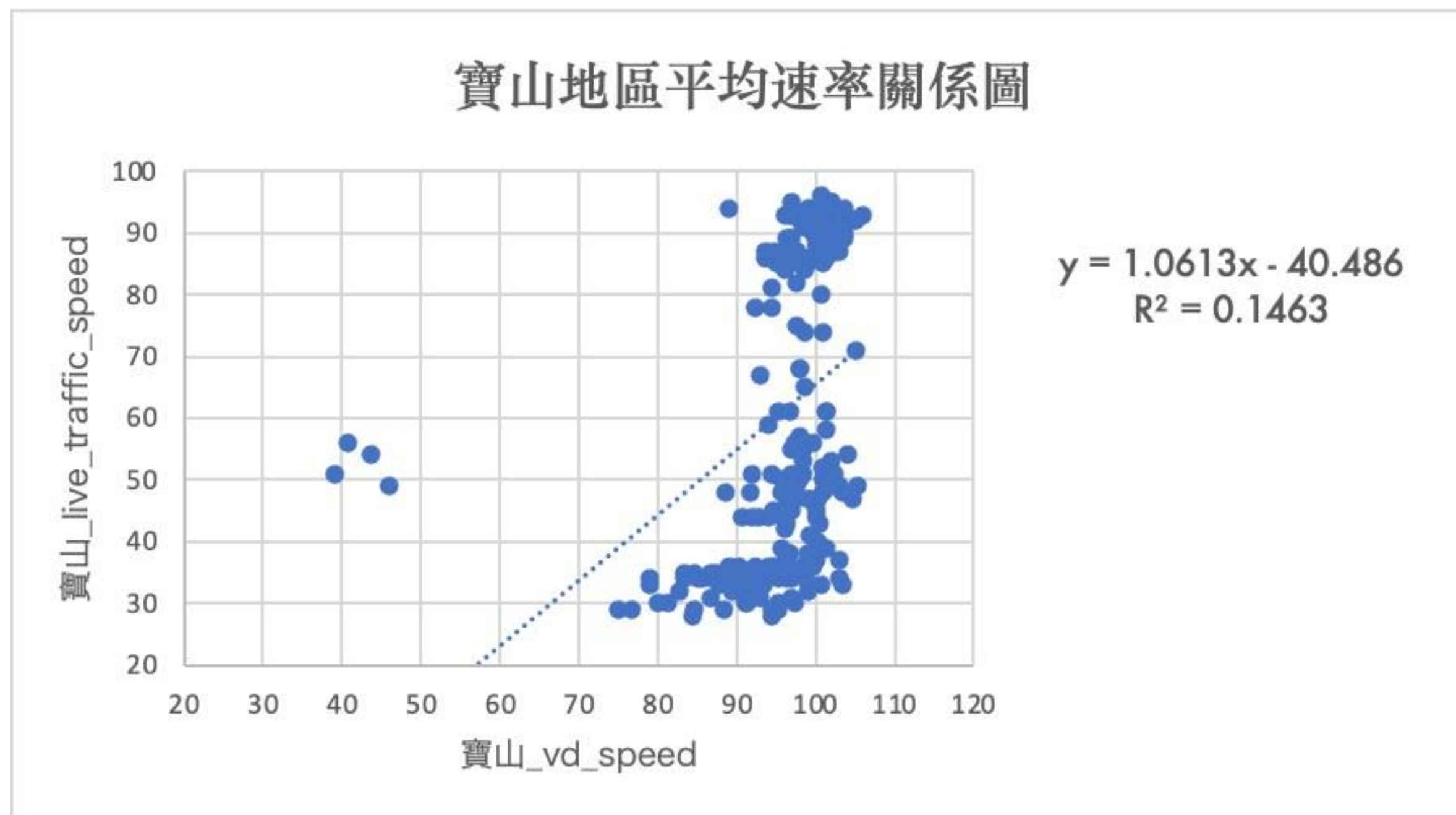
平均速率關係圖

2. 國道一號北上三重路段



平均速率關係圖

3. 國道一號北上寶山路段



平均速率關係圖

1. 斜率：

根據預期，兩資料來源中的平均速率應相近，繪製出的方程式斜率趨近於1，然竹北與三重路段之斜率偏差明顯。

	斜率	決定係數
國道一號北上 竹北路段	0.3204	0.2698
國道一號北上 三重路段	1.7703	0.4403
國道一號北上 寶山路段	1.0613	0.1463

平均速率關係圖

2. 決定係數：

由於三組資料的決定係數極低，表示回歸模型表現不佳，三組數據的 x , y 相關性不高

	斜率	決定係數
國道一號北上 竹北路段	0.3204	0.2698
國道一號北上 三重路段	1.7703	0.4403
國道一號北上 寶山路段	1.0613	0.1463

總結與分析

經過製圖分析，我們得出兩組資料來源在我們所選的時間範圍內並沒有一致性。

1. 儀器不同

VD資料使用紅外線偵測器測速，而即時路況資訊則是使用測速照相

2. 即時路況資料的範圍較大，影響因素較多

3. 可能內外車道受影響因素不同，速率會不相同

	斜率	決定係數
國道一號北上竹北路段	0.3204	0.2698
國道一號北上三重路段	1.7703	0.4403
國道一號北上寶山路段	1.0613	0.1463

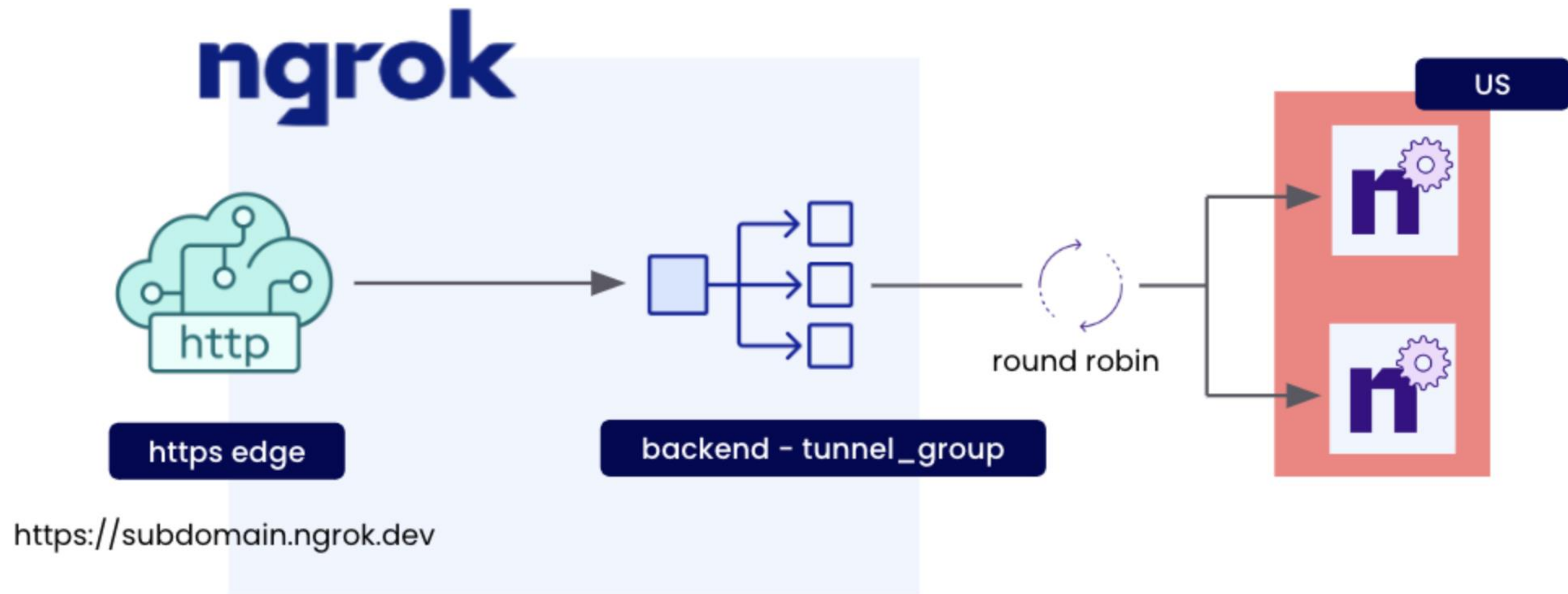


額外功能或創意

LINE Bot-交通辣妞

我們做了一個Line Bot-交通辣妞，可
先先選擇查詢VD資料或即時路況資料，
再輸入時間及路段Link ID，即可得到
2023/12/29該路段的路況資料。





- 可以將本機的IP埠號 (http://127.0.0.1:8000), 對應到一個隨機產生的且公開的HTTPS網址
- 外部使用者只要透過這個HTTPS網址, 就能夠存取本機所運行的服務

LINE Bot-交通辣妞

用網址建立伺服器後，將產生之網址串連 Line Developer，交通辣妞就動起來拉！



程式碼

```
from django.shortcuts import render
from django.http import HttpResponse, HttpResponseBadRequest, HttpResponseForbidden
from django.views.decorators.csrf import csrf_exempt
from django.conf import settings

from linebot import LineBotApi, WebhookParser
from linebot.exceptions import LineBotApiError
from linebot.models import TextMessage, TextSendMessage

import xml.etree.ElementTree as ET

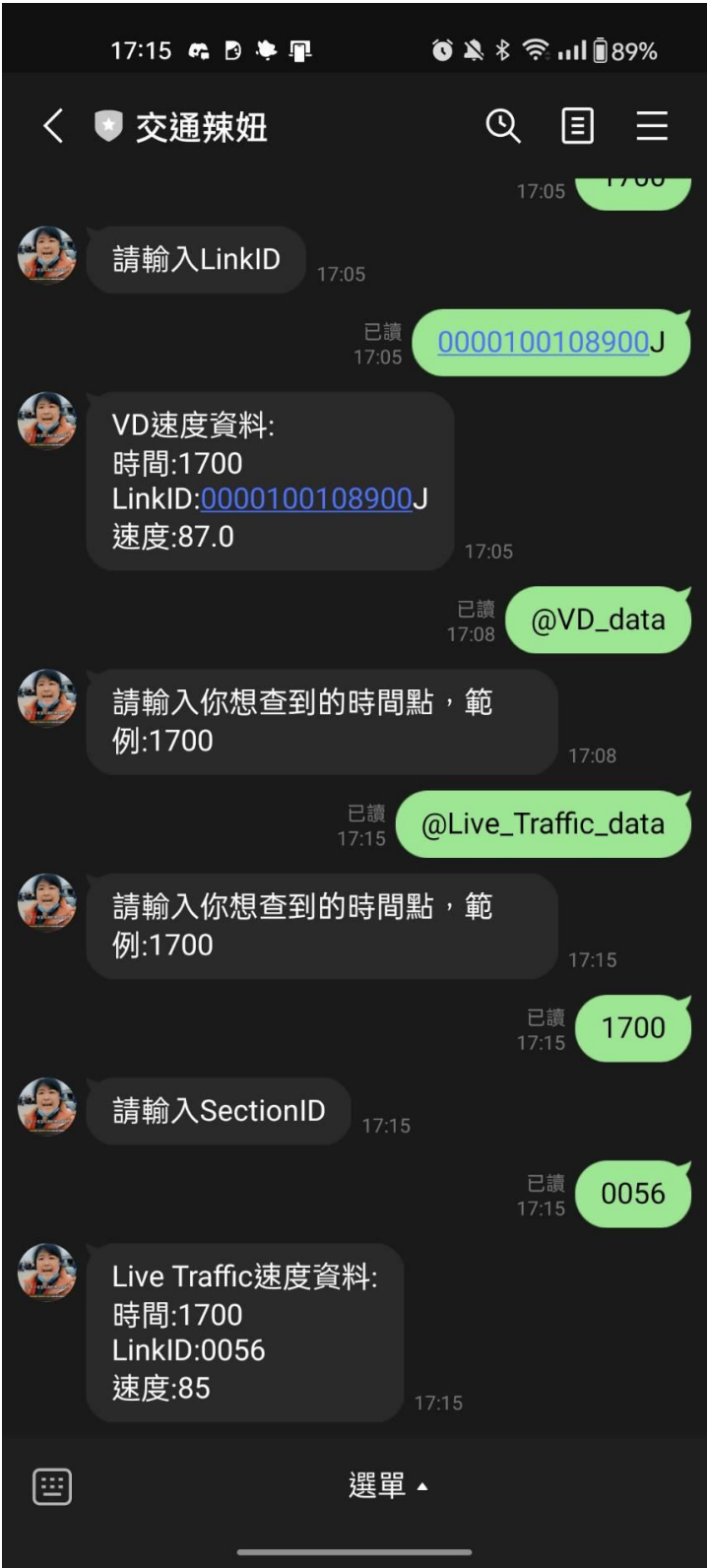
line_bot_api = LineBotApi(settings.LINE_BOT_API_KEY)
parser = WebhookParser(settings.LINE_BOT_SECRET)

def callback(request):
    @csrf_exempt
    def callback(request):
        if request.method == 'POST':
            signature = request.META['HTTP_X_LINE_SIGNATURE']
            body = request.body.decode('utf-8')

            try:
                events = parser.parse(body, signature) # 傳入的事件
            except InvalidSignatureError:
                return HttpResponseForbidden()
            except LineBotApiError:
                return HttpResponseBadRequest()

        for event in events:
            if isinstance(event, MessageEvent): # 如果有訊息事件
                if event.message.text == '@VD_data':
                    history.clear()
                else:
                    history.append(event.message.text)
                    if history[0] == 'VD':
                        speed = get_vd_live_speed(history[1], history[2])
                        if speed == False:
                            line_bot_api.reply_message(
                                event.reply_token,
                                TextSendMessage(text='格式輸入錯誤或找無資料，請重新輸入!'))
                        else:
                            line_bot_api.reply_message(
                                event.reply_token,
                                TextSendMessage(
                                    text=f'VD速度資料:\n時間:{history[1]}\nLinkID:{history[2]}\n速度:{speed}'))
                    elif history[0] == 'Live Traffic':
                        speed = get_live_traffic_speed(history[1], history[2])
                        if speed == False:
                            line_bot_api.reply_message(
                                event.reply_token,
                                TextSendMessage(text='格式輸入錯誤或找無資料，請重新輸入!'))
                        else:
                            line_bot_api.reply_message(
                                event.reply_token,
                                TextSendMessage(
                                    text=f'Live Traffic速度資料:\n時間:{history[1]}\nLinkID:{history[2]}\n速度:{speed}'))
                    history.clear()
            return HttpResponse()
```


實際對話情況



分工表

	楊守荃	黃浚瑀	程品諭
工作內容	簡報製作 資料分析 錄製影片	資料分析 資料處理 繪製圖表	資料分析 資料處理 繪製圖表
%數	33%	33%	33%



THANK YOU

