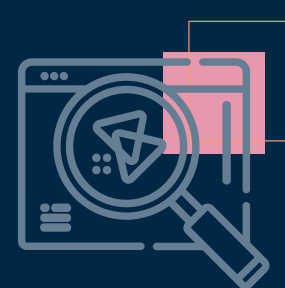


## NTU DSSI 109-1 資料科學與社會分析

公衛三 劉德駿  
社會三 王聖夫  
日文四 陳佳妤

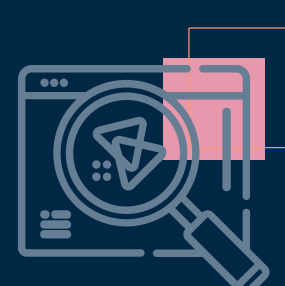
# 以氣象與站點資料 預測YouBike租借量



## Motivation & Goal

- 為什麼當我們有用車需求時，平常塞滿YouBike的車架總是無車可借呢？能不能租借到車，是可預測的嗎？
- 我們將探討如何透過建立模型，找出能解釋YouBike租借量的因子，進而做租借量的預測。

⇒ 藉由預測站點租借量的變化情況，改善使用者對YouBike的取用性（accessibility），進而增加資源配置的效率



## Data

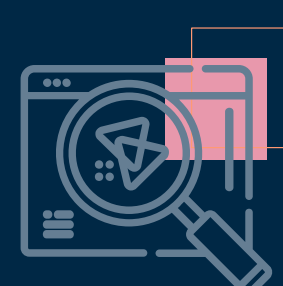
### Source

2018年1月至12月，松山與信義區的資料

- 臺北市自行車租借紀錄（臺北市資料大平臺）
- 天氣觀測資料（氣象局）
- 站點資訊（維基百科，公共自行車站點列表）

### Features

- 依變項：租借量、租借率（租借量／該站車柱總數）
- 自變項：月份、小時、氣壓、氣溫、濕度、降雨量、行政區、附近是否有捷運、附近是否有學校、是否為週末



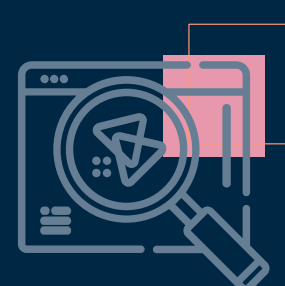
## Methods & Models

### Regression Model

- 預測：租借量
- 模型：Linear, SVM, Random Forest

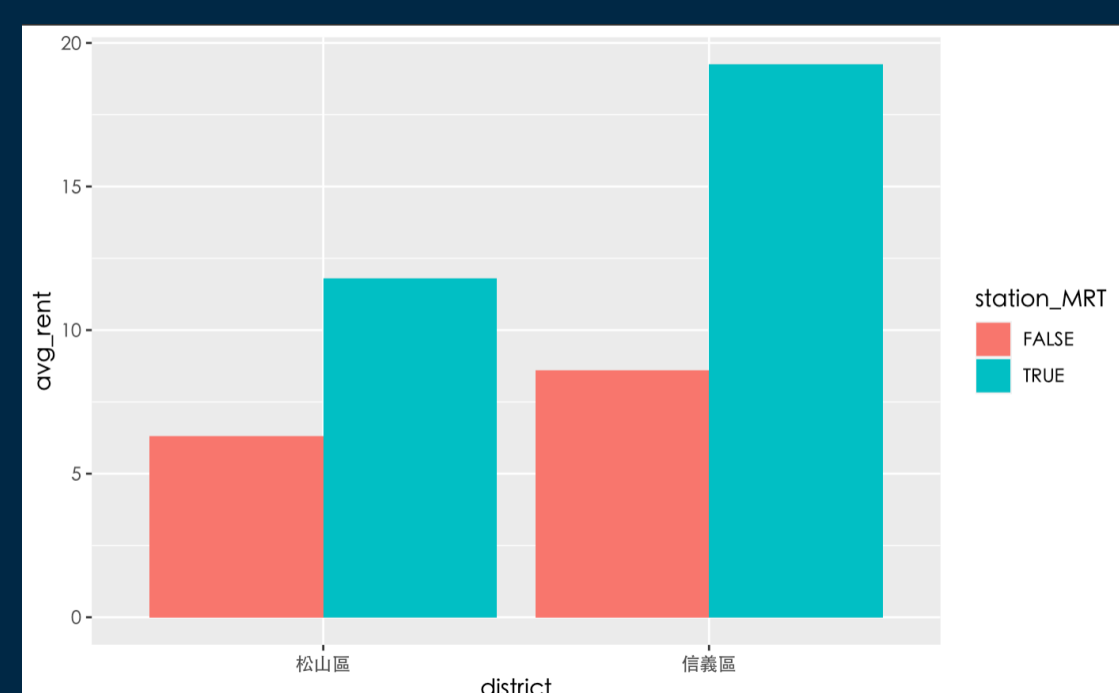
### Classification Model

- 預測：租借率
  - Busy:  $\text{rent\_rate} > 0.5$
  - Normal:  $0.5 \geq \text{rent\_rate} > 0.3$
  - Free:  $0.3 \geq \text{rent\_rate}$
- 模型：KNN, SVM, Random Forest

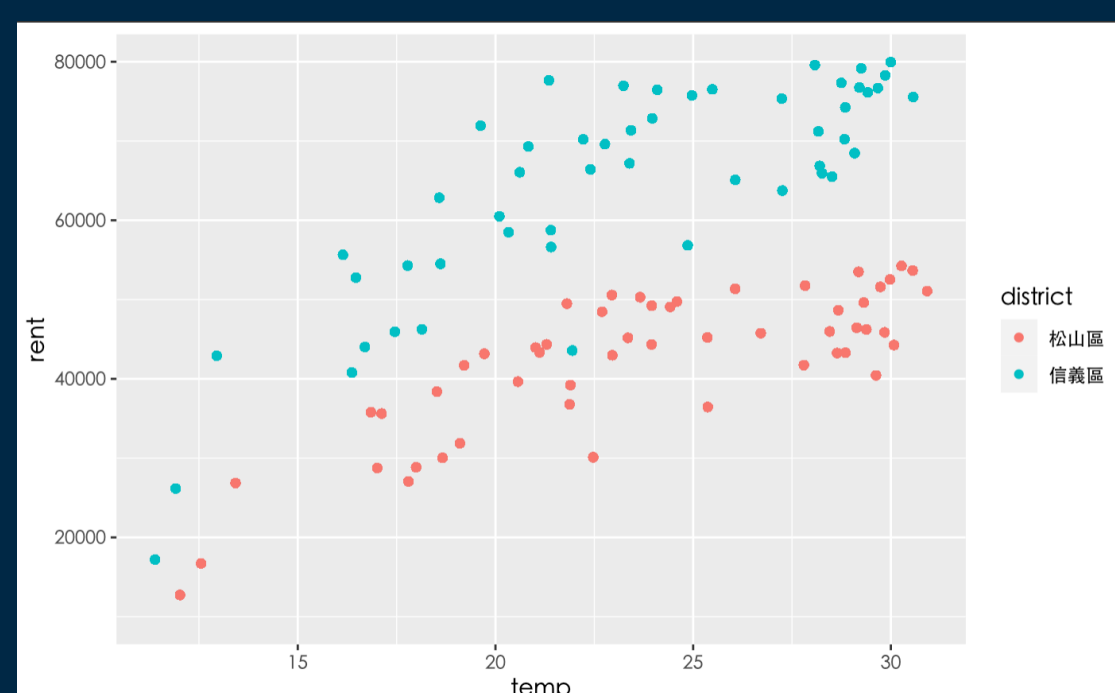


## Results

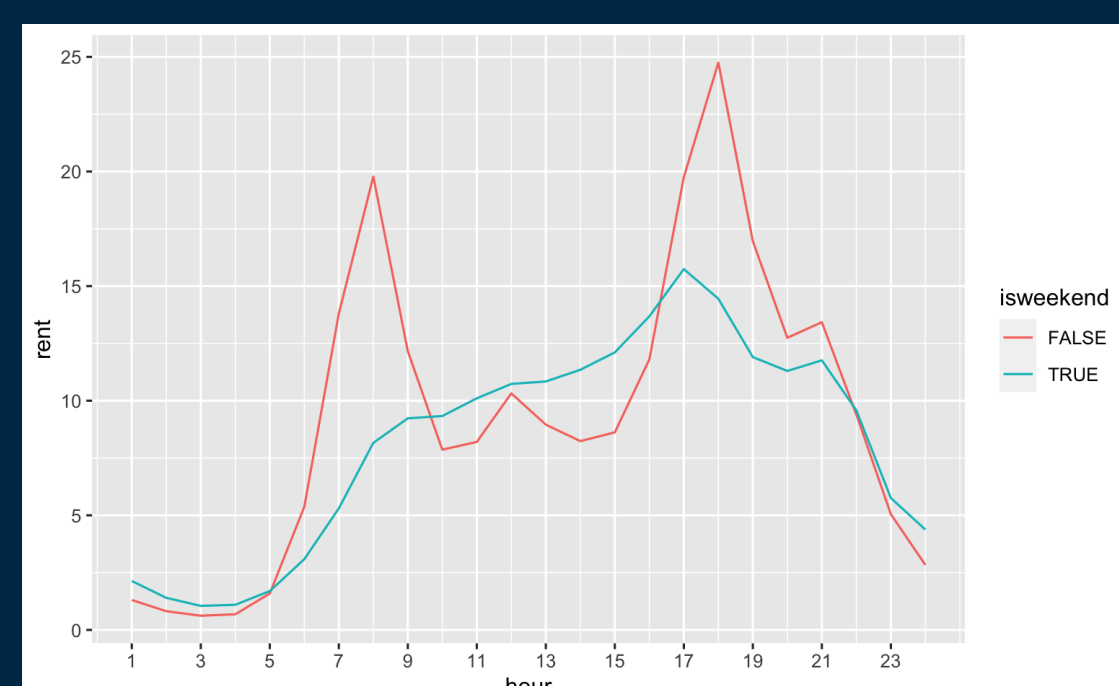
### Visualization



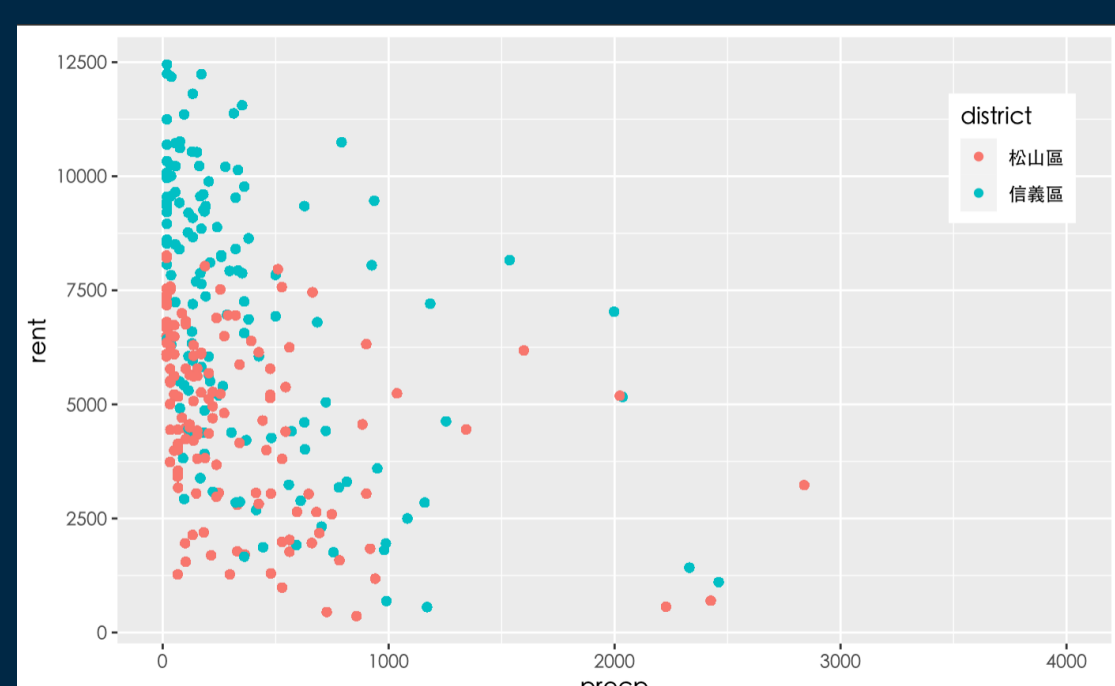
站點附近有捷運站，平均租借量較高



氣溫與租借量之散佈圖：兩者成正相關



一日時段的平均租借量：平日尖峰時刻較明顯

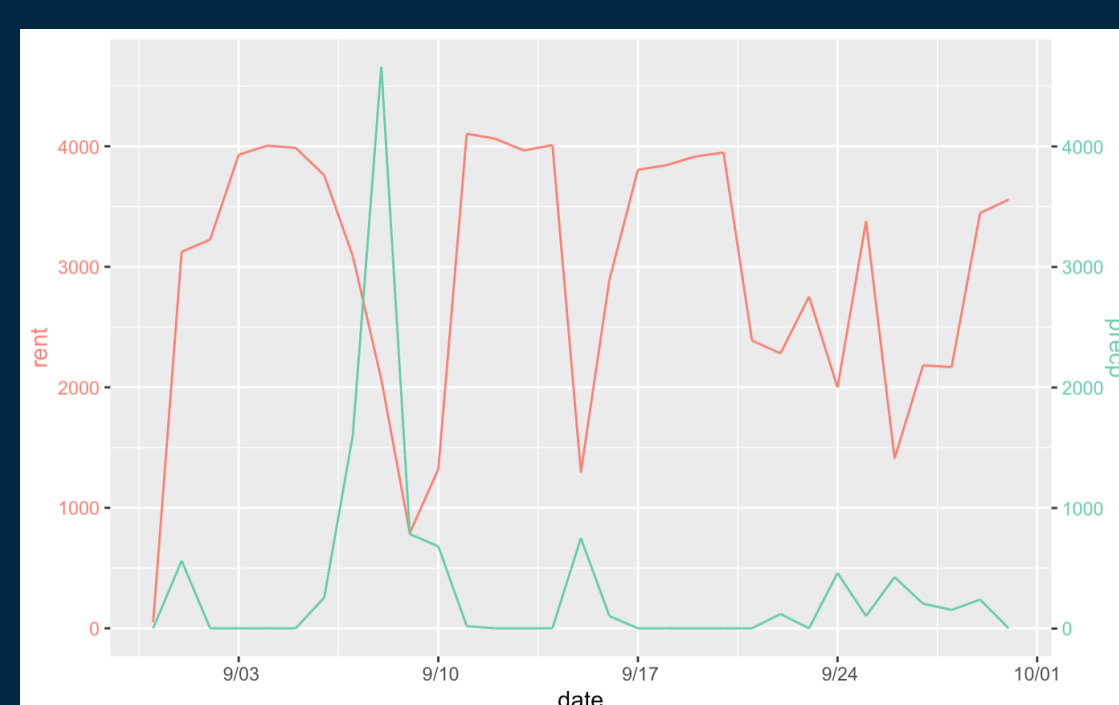


降雨量與租借量之散佈圖：兩者成負相關

### Model Evaluation

Regression	Linear	SVM	RF
RMSE	9.19	9.06	8.12

Classification	KNN	SVM	RF
Accuracy	75.6%	78.3%	79.4%



降雨量和租借量的折線圖

