

Portland State University

**PDXScholar**

---

Engineering and Technology Management  
Student Projects

Engineering and Technology Management

---

Winter 2018

# Analysis of Residential and Auto Break-in Records in Taipei City

Afnan Althoupety  
*Portland State University*

Aishwarya Joy  
*Portland State University*

Juchun Cheng  
*Portland State University*

Priyanka Patil  
*Portland State University*

Tejas Deshpande  
*Portland State University*

Follow this and additional works at: [https://pdxscholar.library.pdx.edu/etm\\_studentprojects](https://pdxscholar.library.pdx.edu/etm_studentprojects)



Part of the [Applied Statistics Commons](#), and the [Social Control, Law, Crime, and Deviance Commons](#)

**Let us know how access to this document benefits you.**

---

## Citation Details

Althoupety, Afnan; Joy, Aishwarya; Cheng, Juchun; Patil, Priyanka; and Deshpande, Tejas, "Analysis of Residential and Auto Break-in Records in Taipei City" (2018). *Engineering and Technology Management Student Projects*. 1943.

[https://pdxscholar.library.pdx.edu/etm\\_studentprojects/1943](https://pdxscholar.library.pdx.edu/etm_studentprojects/1943)

This Project is brought to you for free and open access. It has been accepted for inclusion in Engineering and Technology Management Student Projects by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: [pdxscholar@pdx.edu](mailto:pdxscholar@pdx.edu).



## **Analysis of Residential and Auto Break-in Records in Taipei City**



**Course:** ETM 538

**Term:** Winter 2018

**Instructors:** Daniel Sagalowicz, Mike Freiling

**Date:** 03/09/2018

**Authors:** Afnan Althoupety , Aishwarya Joy , Juchun Cheng , Priyanka Patil , Tejas Deshpande

# 1. Introduction

Taipei City is the capital of Taiwan. It has population of 2.7 million living in the city area of 271 km<sup>2</sup> (104 mi<sup>2</sup>). There are totally 12 administrative districts in this city. To maintain the safety of the city, Taipei City Police Bureau has arranged regular patrol routes with focus on the high-risk area where residential and auto break-in occurs. Due to limited police resource, resident neighborhood also organized volunteered patrol teams to enhance the security in residential area. Based on past file record history, the Bureau would like to understand the high-risk districts and time schedule to improve their patrol schedule and route accordingly. The goal of this project is to help the bureau to identify the higher risk area and schedule for such criminal incidents so that to develop better prevention plan with its limited resources.

## 2. Project Objective

Residential break-in and burglaries have been most frequent crimes in Taiwan nowadays and they are closely related to geographical locations. Citizens have been relied on the police patrol to decrease the chance of these crimes. However, the police resources have not been increased enough to catch up the increased density of residents in Taipei city. To enhance the patrol route, community volunteer patrol teams have been organized to make up the police resources. However, the police bureau had only limited utilization of their past criminal record history to predict for the high-risk time and location of burglary crimes.

The objective of this project is to analyze Taipei City Police file records for house and auto break-in, burglary records from 2015 to 2017 to identify the high-risk situation which includes the district and the time. The total number of files records in the data set is 1988 entries. Various analysis methodologies, including association and 1R, would be tried to develop the best prediction and suggestion.

### 3. Background Information

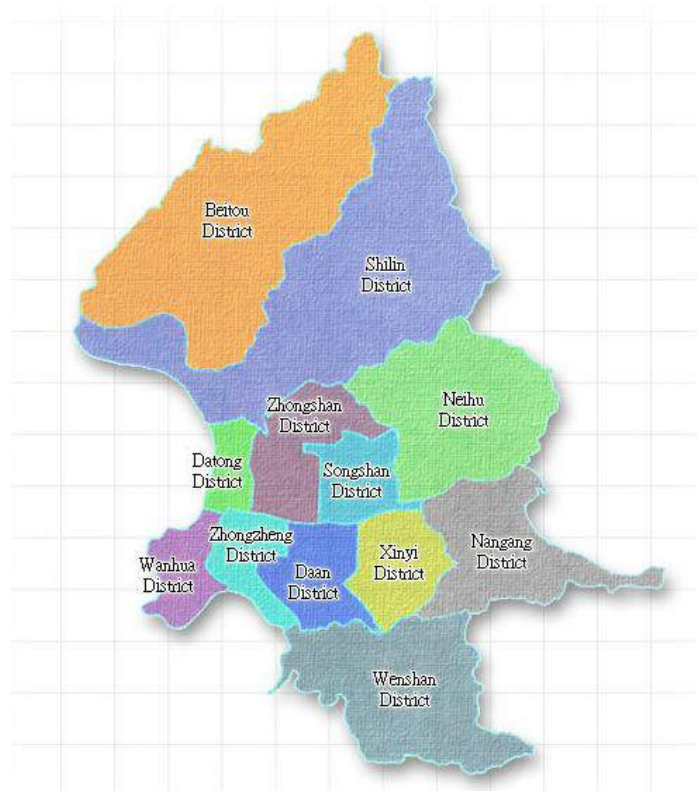


Figure 1: Taipei City Administrative District Map [1]

As shown in **Figure 1**, there are totally 12 administrative districts in the Taipei City, including Songshan, Xinyi, Daan, Zhongshan, Zhongzheng, Datong, Wanhua, Wenshan, Nangang, Neihu, Shilin, and Beitou. **Figure 2** below shows the area size and population of each district. Due to the varied topography, economic development and time of development, the population is unevenly distributed. Daan, Songshan and Datong districts are the most populated.

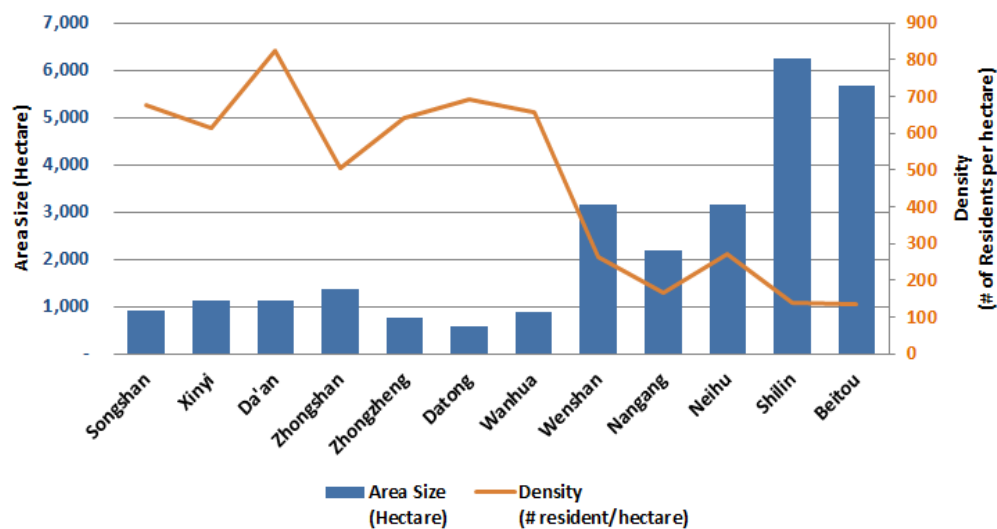


Figure 2: Area Size and Population for Each District [2]

On the other hand, the growth of household number goes beyond the increased number of policemen. As we can see from **Figure 3**, the actual number of policemen in Taipei City has not increased much when compared with the increased number of households. In the past 25 years, the households per police station and per policeman has increased by around 30%. This had dramatically increased the burden of policemen's workload. At the same time, it has increased the risk for household's safety.

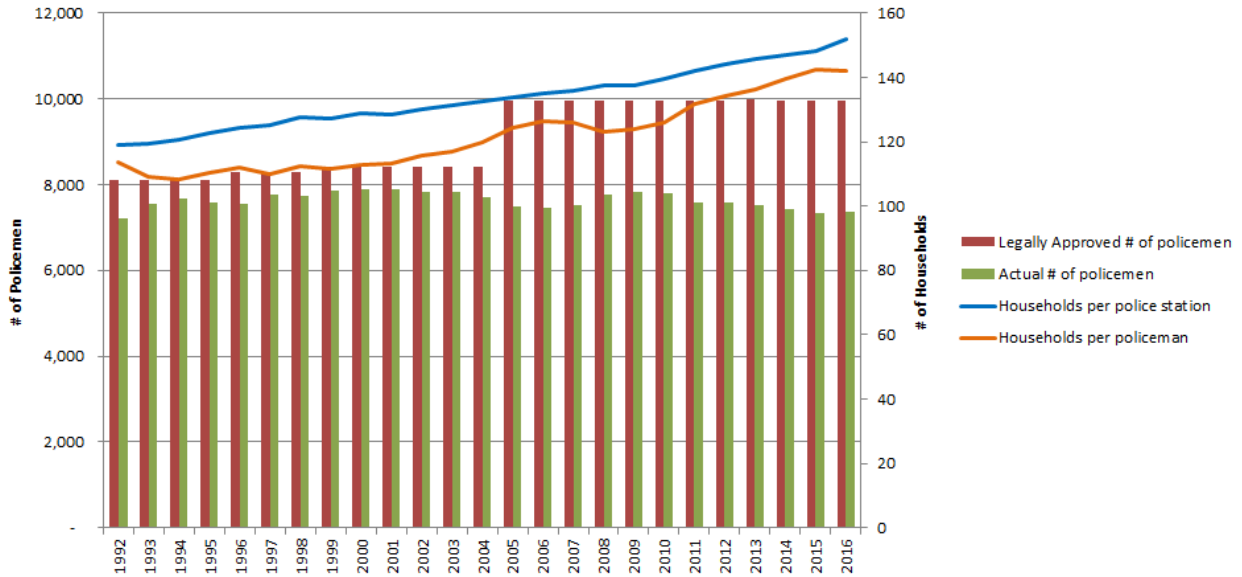


Figure 3: Taipei City Policemen Resource Situation [2]

While comparing among 12 districts, we can see the workload is very uneven among 12 districts for the workload of each policeman. **Figure 4** below shows the actual number of policemen of each district and the number of household for each policeman and in that district in the end of 2016. Neihu district has almost doubled number of households for each policeman when compared with Zhongzheng district. Zhongzheng District is the home of most of the national government buildings of Taiwan. It includes the Presidential Office and various government ministries. It is the reason for its high number of police resources. Da'an is the home of three major national universities and has numerous shopping areas. Da'an also offers some of Taipei's most expensive residential real estate. This is the reason for its higher number of police resources. However, Beitou district is the second highest on the chart for policeman workload while Beitou has the lowest density (# of residents per hectare) from **Figure 2**.

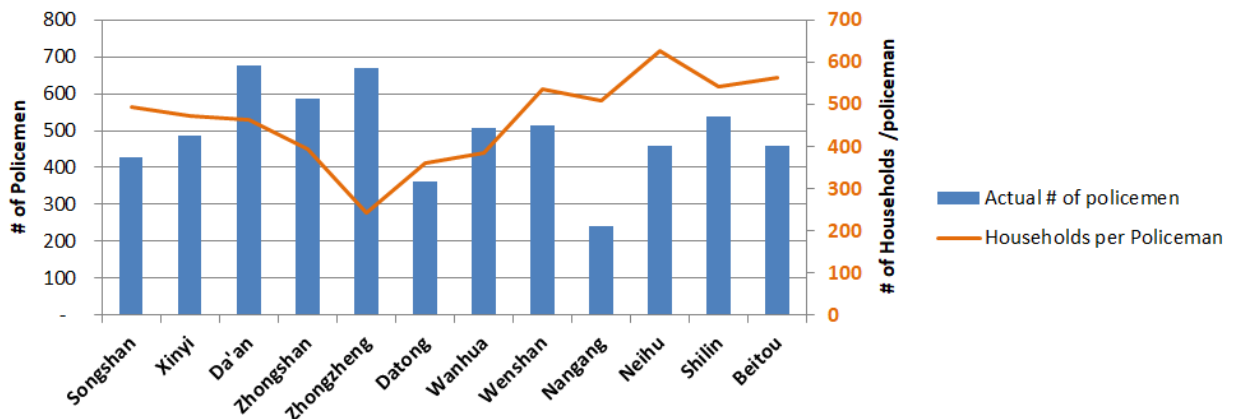


Figure 4: Policeman Resource among 12 Districts in the end of 2016 [2]

From above analysis, it is noticed that the distribution of policemen resources is very uneven among 12 districts. To best utilize the limited resource for the criminal prevention plan would be an urgent task for the Taipei City Police Bureau.

## 4. Dataset and Variable Description

Taipei City government has been promoting open-data policy for Open Government campaign. Single-entry integrated website and add-on value application development program had been arranged to allow the citizens to access more transparent data directly. On <http://data.taipei> website, various of data regarding citizen's daily life has been published to allow more variety application development by third parties from academic and private vendors. We were able to acquire the residential and auto crime history records from this website for our project analysis. Table 1 below show the dataset and its variable description.

Field	Type	Description
Index	Integer	File record indexing number, starting from 1,2, 3,... so on
Type	Text	Indicate the record type: Housebreak-in or Auto break-in/stolen
Year	Integer	Indicate the case year as 2015, 2016 or 2017
Date	Date	Indicate the case date of break-in as mmdd format
Time	Text	Indicate the case time in hour range, e.g. 07~09, 10~12, 19~21, 22~24
Location District	Text	Indicate the district name where the case happened

Table 1: Police Records

## 5. Initial Data Analysis

As we wanted to analyze the chances of a break-in in a district on certain day at certain time, we explored and analyzed each predictor associated with total number of break-ins including for both House and Auto before building any algorithm on the data. There are few variables which can be used to predict the chances of House or Auto break-ins district, hence we analyzed them individually first and then in combination with each other to understand the correlation between them.

## 5.1 Total Number of Break-ins for all districts per year

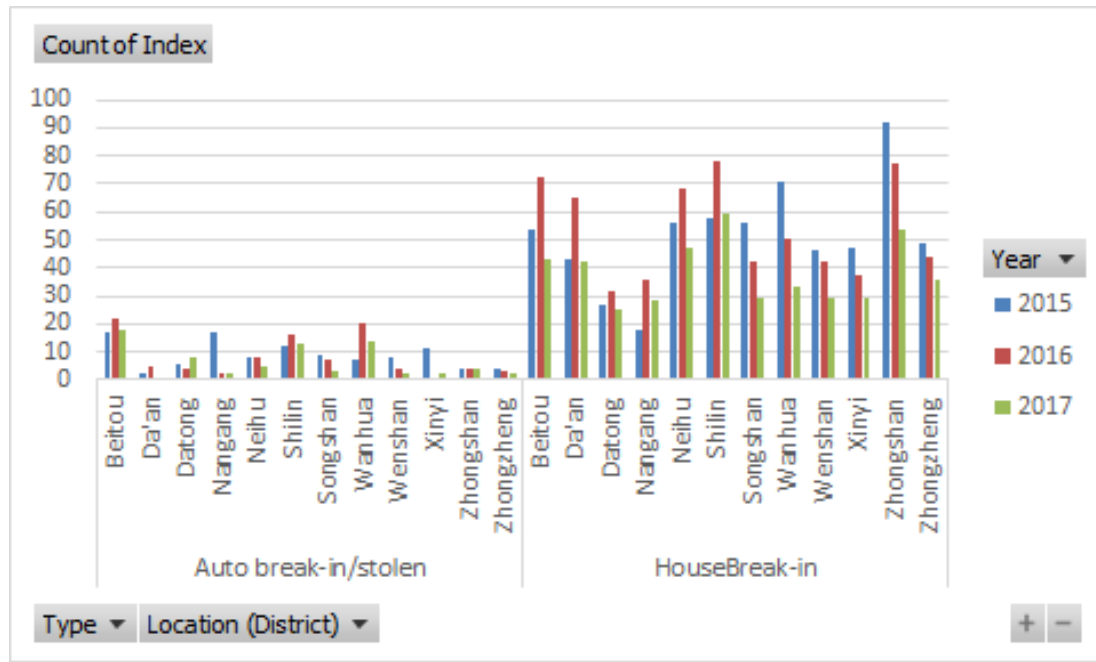


Figure 5: Break-Ins per year for each district

The graph shows that highest number of house break-in were recorded in the year 2016, whereas number of auto break-ins are pretty much constant for all three years.

## 5.2 Total Number of Break-ins per District

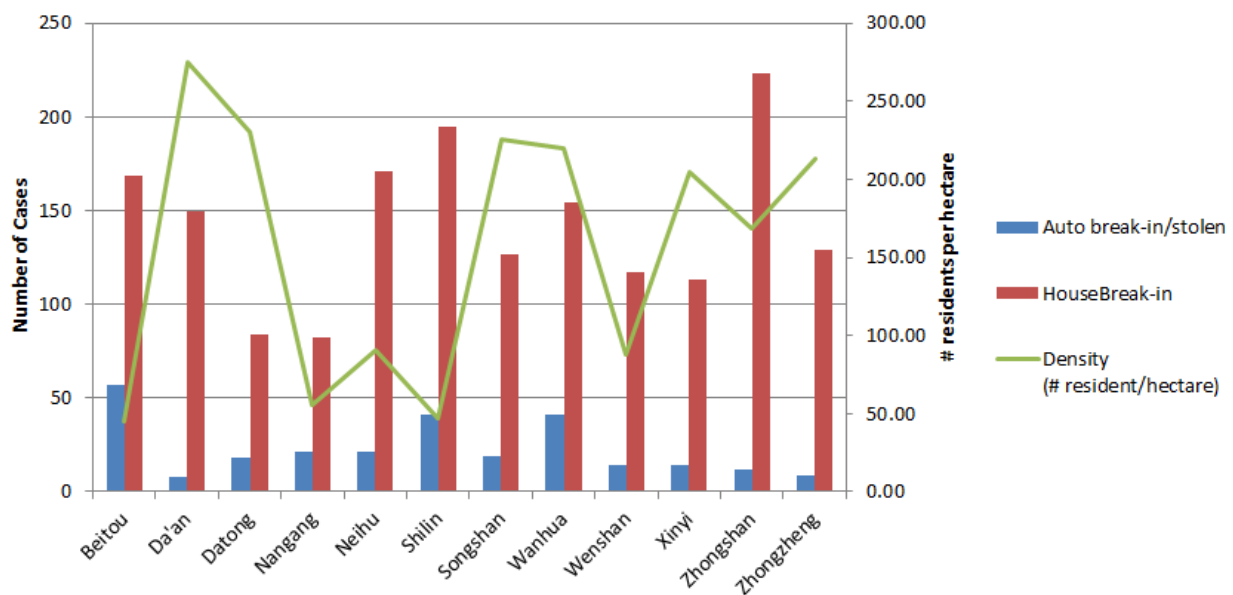


Figure 6: Break-Ins per district



The graph above is used to analyze the most unsafe district in the Taipei city which has maximum number of Auto and House Break-ins. As we can see, the highest number of house break-ins are in the district Zhongshan and other districts have quite constant count of house break-ins, where Datong & Nangang being the lowest in house and auto break-ins. In case of only auto break-ins, Beitou has the highest number of auto break-in counts whereas remaining districts have constant auto-break ins.

### 5.3 The day of the week & number of break-ins

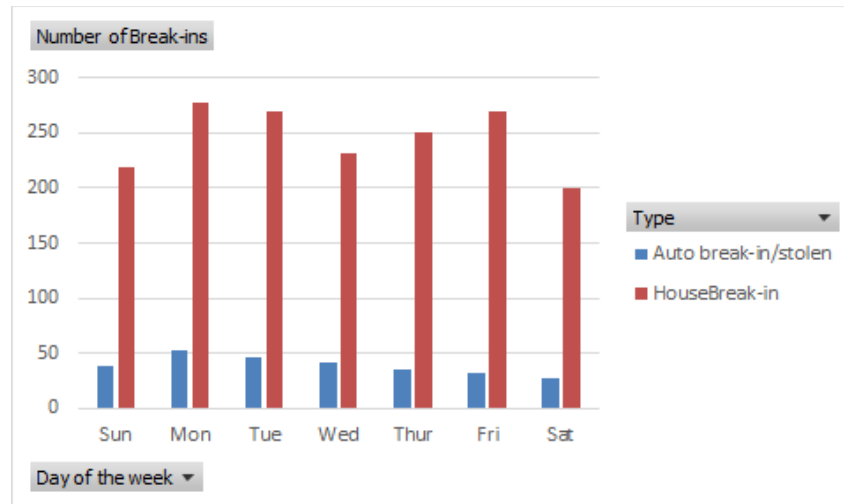


Figure 7: Break-Ins per day of week

As it can be seen from the graph above, weekdays have the highest number of house break-ins whereas auto break-in count remain constant over the whole week except for Tuesday.

### 5.4 Time of the day & number of Break-ins

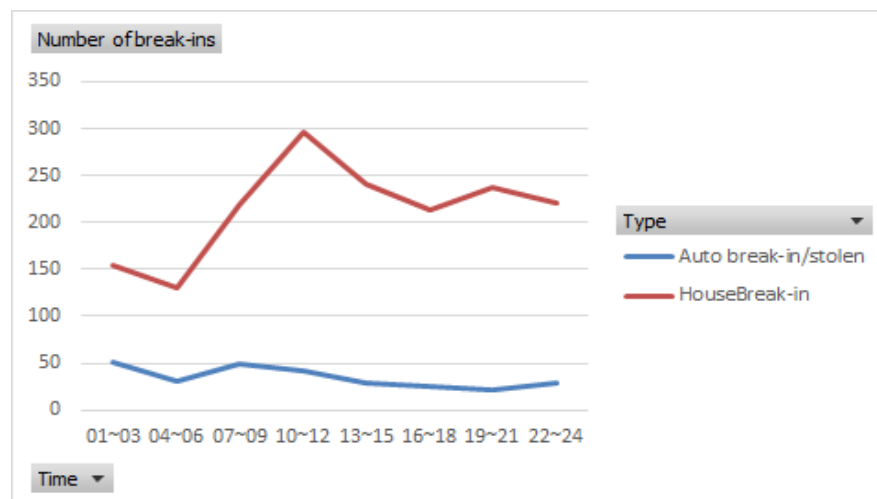


Figure 8: Break-Ins per Time of day

The plot shows that maximum number of house break-ins occur in between 10 am to 3 pm, it is evident that people leave house during this period for work which ultimately creates the risk of House Break-in. The maximum number of auto break-ins happen during the night between 01 am to 03 am as well as early in



the morning at 07 am- 09am. The probability for Auto Break-in during the night time increases as most of the vehicles residential area are parked at their respective place with owners sleeping at that time.

To help Police station improve their Patrol routing and frequency, we considered most vulnerable hours of House break-ins in detail. Considering that 10-12 and 1-3 pm show the highest count of ‘House break-ins’, we further analyzed, which locations get affected the most during these 2 times. (As auto break-in seems to be constant during the whole day, we only considered House Break-ins).

#### 5.4.1 During 13:00-15:00

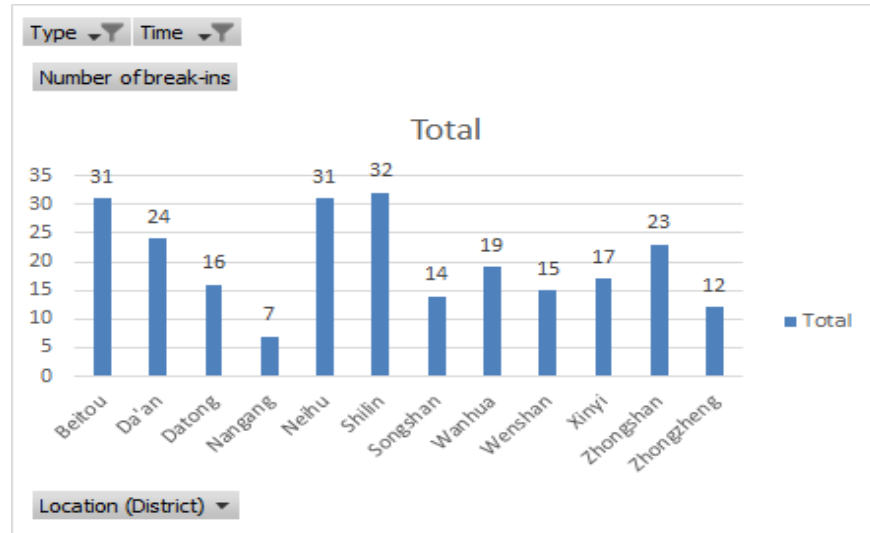


Figure 9: House Break-Ins during 13:00-15:00

As it can be seen from the graph, Beitou, Neihu and Shilin have maximum number of house break-ins during the time 13:00-15:00.

#### 5.4.2 During 10:00-12:00

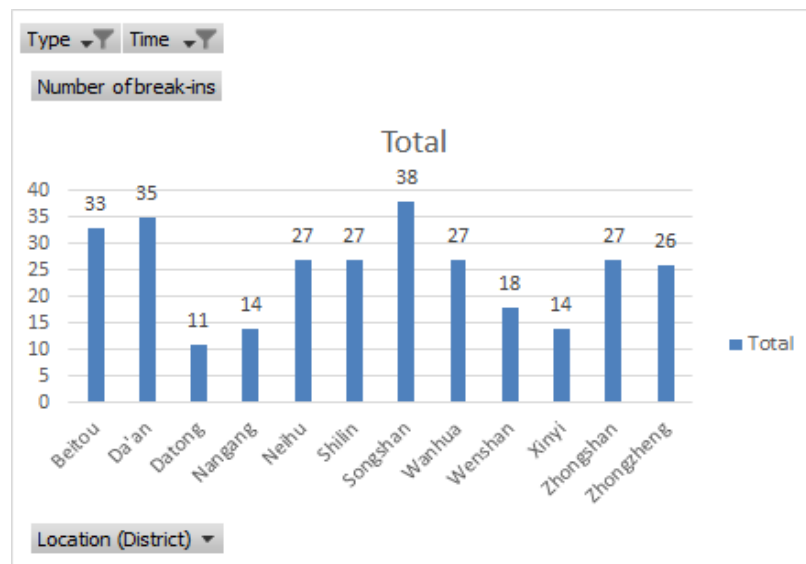


Figure 10: House Break-Ins during 10:00-12:00

It can be analyzed from the graph that Beitou, Daan and Songshan have highest number of house break-in during 10:00-12:00. Considering both the graphs and analysis together, Beitou seems to be vulnerable between 10 am to 3pm.

## 5.5 Number of Break-ins in each Season

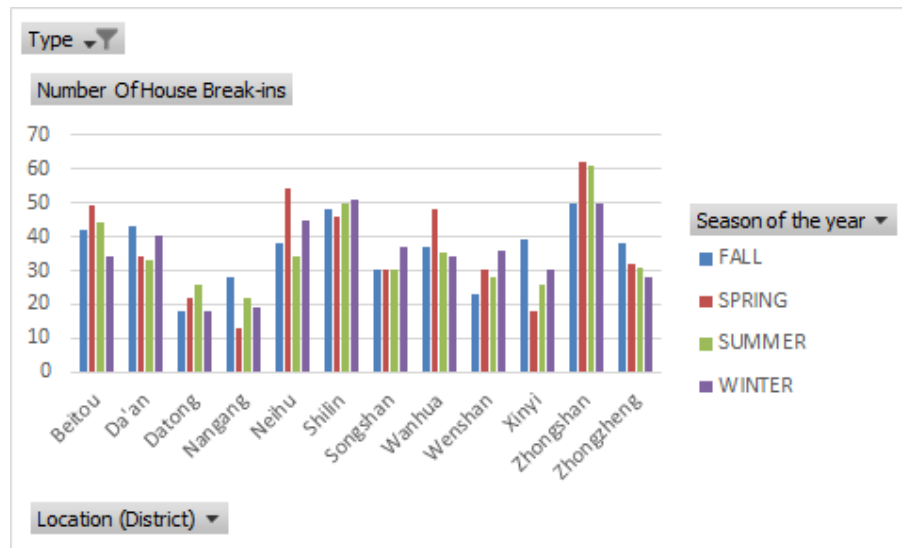


Figure 11: House Break-Ins during each Season by district

The figure above shows Spring and Summer have more number of House break-ins in almost all districts, whereas districts Shilin which has almost equal number of break-ins throughout the year.

**Note:** please refer to section 6.2 for how we bucket months into seasons.

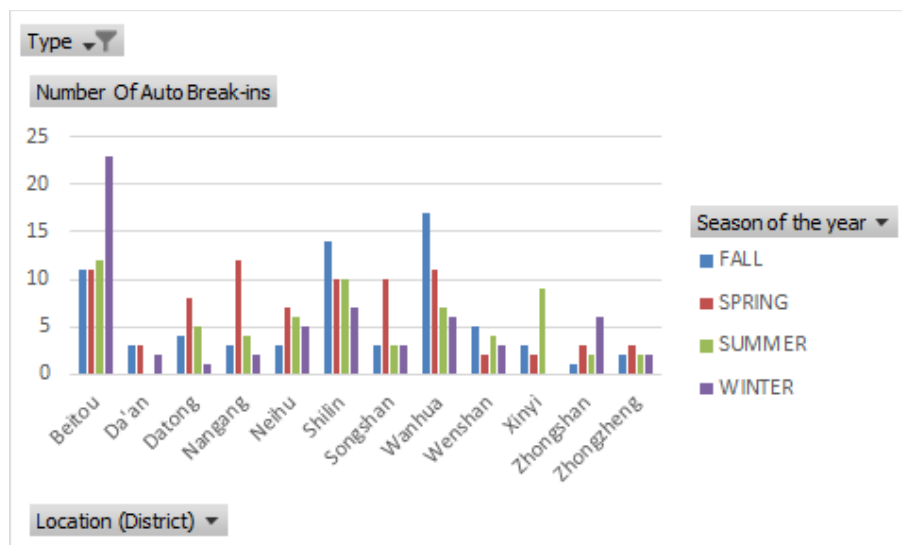


Figure 12; Auto Break-Ins during each Season by district

## 5.6 Number of Holidays in a year & Break-ins

To analyze the number of Break-ins in a particular year/season, we explored the data in correlation with Holidays involved in that year and season. According to the data available regarding National Holidays in Taiwan in year 2015, 2016 and 2017, we got following numbers.

Sum of Total number of holidays Column Labels					
Row Labels	Fall	Spring	Summer	Winter	Grand Total
2015	6	13	6	4	29
2016	7	16	4	6	33
2017	7	8	4	5	24
<b>Grand Total</b>	<b>20</b>	<b>37</b>	<b>14</b>	<b>15</b>	<b>86</b>

**Table 2: Count of holidays in each season in years 2015-2017**

To analyze whether having Holiday would affect the number of Break-in or not, we calculated percentage of Break-ins on holidays for House break-ins and Auto break-ins in each district in particular season.

Count of Index Column Labels													
Row Labels	FALL			SPRING			SUMMER			WINTER			Grand Total
	2015	2016	2017	2015	2016	2017	2015	2016	2017	2015	2016	2017	
Beitou	6.666667	10.52632	12.5	14.28571	20	7.692308	0	15.78947	7.142857	0	0	12.5	9.467456
Da'an	8.333333	0	0	12.5	28.57143	0	0	5.882353	0	0	5.882353	0	5.333333
Datong	0	0	0	50	0	0	0	0	0	0	0	0	2.380952
Nangang	0	0	12.5	33.33333	16.66667	0	0	0	0	0	0	14.28571	4.878049
Neihu	12.5	5	0	4.347826	9.090909	15	0	8.333333	0	0	4	0	5.263158
Shilin	0	5.263158	0	18.18182	16.66667	0	8.333333	5	0	11.11111	9.52381	25	7.692308
Songshan	0	0	0	11.11111	9.090909	0	11.11111	0	0	0	6.666667	16.66667	4.724409
Wanhua	11.11111	0	10	13.63636	43.75	20	0	6.25	0	10	0	0	11.68831
Wenshan	0	11.11111	0	21.42857	0	11.11111	8.333333	9.090909	0	16.66667	6.666667	11.11111	9.401709
Xinyi	6.666667	11.11111	0	50	33.33333	0	22.22222	0	7.692308	13.33333	11.11111	33.33333	15.04425
Zhongshan	4	10	0	0	12.5	0	7.407407	0	11.11111	8.695652	17.64706	40	8.520179
Zhongzheng	4.761905	0	11.11111	0	0	0	12.5	11.11111	7.142857	0	15.38462	0	5.426357
<b>Grand Total</b>													

**Table 3: Percentage House break-ins on Holidays each year in all the districts**

Table 3 shows percentage of House Break-ins in Holidays are not significantly high, except in SPRING 2015, district Nangang had 33.33% of House break-ins on Holidays whereas Xinyi has 33.33% break-ins on Holidays in WINTER 2017. All other numbers, such as Da'an in SPRING 2016, Xinyi in SUMMER 2015 are similar considering percentage break-ins on Holidays.

Count of Index Column Labels													
Row Labels	FALL			SPRING			SUMMER			WINTER			Grand Total
	2015	2016	2017	2015	2016	2017	2015	2016	2017	2015	2016	2017	
Beitou	0	20	50	0	33.33333	16.66667	0	0	0	0	0	0	7.017544
Da'an	0	0 N/A	0	0	0	100 N/A	N/A	N/A	N/A	0 N/A	0 N/A	0 N/A	12.5
Datong	0 N/A	0	0	0	0	0	0 N/A	0 N/A	0 N/A	0 N/A	0 N/A	0 N/A	0
Nangang	0 N/A	0	0	18.18182	0 N/A	0 N/A	0 N/A	0	0	0	0 N/A	0 N/A	9.52381
Neihu	N/A	0	0	0	0 N/A	0	0	0	0	0	0	0	0
Shilin	0	0	0	33.33333	0	0	0	0	14.28571	0	0	0	7.317073
Songshan	N/A	0	0	16.66667	0	0	0	0 N/A	0 N/A	50	0 N/A	0	10.52632
Wanhua	33.33333	0	0	0	0	0	0	0	0	0	0	0	2.439024
Wenshan	0	0 N/A	0	0 N/A	0	0	0	0 N/A	0 N/A	0 N/A	0	0	0
Xinyi	50 N/A	100	0 N/A	0 N/A	N/A	0	0	0	0 N/A	N/A	N/A	0	14.28571
Zhongshan	0 N/A	N/A	N/A	0	0	0	0	0 N/A	0	0	50	0	8.333333
Zhongzheng	0 N/A	0	0	0	0 N/A	0	0 N/A	N/A	N/A	0	0	0	0
<b>Grand Total</b>	<b>10</b>	<b>3.225806</b>	<b>11.11111</b>	<b>13.15789</b>	<b>4.166667</b>	<b>10</b>	<b>0</b>	<b>0</b>	<b>4.761905</b>	<b>4.761905</b>	<b>4.166667</b>	<b>0</b>	<b>5.818182</b>

**Table 4: Percentage Auto break-ins on Holidays each year in all the districts**

We can see from the table above that, Da'an has 100% of Auto break-ins on Holidays. But, that seems to be overfitting the data. Other than Da'an, districts Beitou has 33.33% of Auto break-ins on holidays in Spring 2016 and 50% in Fall 2017. Similarly, Wanhua and Xinyi has 33.33% and 50% break-ins on Holidays in Fall 2015. In case of Auto break-ins, Winter and Spring seem to have higher chances for almost all districts except Xinyi which has no Auto break-in in Winter and possesses higher chance of break-ins in Summer.

## 6. Methodology/Approach

### 6.1 Data Mining

Knowledge discovery in data is defined as the extraction of implicit patterns or information by leveraging the massive amount of data. Indeed, data mining is the core step in knowledge discovery process and its domain is effective in many applications including business, science, and health. Moreover, as a growing discipline, data mining can be applied using different state-of-the-art algorithms including Neural Network, Naïve Bays, Regression, K-nearest Neighbors, and Decision Tree. This project used data mining and algorithms majorly to find patterns in different variables and provide probabilities for a definite outcome. Considering the nature of the data, we came up with multiple algorithms which can be used for data mining. Amongst them, Association and 1R seemed to be workable on this data.

### 6.2 Buckets Used

As the project aims to find the vulnerable districts on certain day at certain time, we analyzed it further to check whether it has any relation with particular season and holidays in it. Hence, we divided the data using 4 buckets for 4 seasons depending on the weather conditions in Taipei city as following:

Season	Months
Winter	Nov, Dec, Jan
Spring	Feb, Mar, Apr
Summer	May, June, July
Fall	Sept, Oct, Nov

Table 5: Buckets created

## 7. Algorithms Used

### 7.1 1R Algorithm:

This algorithm can be simply defined as constructing one rule for each predictor in the data, and then selecting the rule with the minimum error. To illustrate, 1R algorithm generates the rule for each predictor value by counting the frequency of this value against the target (confidence) and assigning this value to the target based on the most frequent. After that, the total error of each predictor is calculated and finally we

make the decision by choosing the rule with the smallest error. We built 1R algorithm on 3 main predictors i.e. Season, Time and Day of the week to analyze which predictor has highest significance on number of Break-ins in certain district. The results and probabilities associated with the outcome showed the relationships between vulnerability of the district and respective predictor. As there are two types of break-ins, House break-in and Auto Break-in, with house break-ins occurring exponentially more often, algorithms on both of them was built separately.

### 7.1.1 1R algorithm using Season of the year for all districts

#### For House Break-in:

House Break-ins						Error Rate					
Row Labels	FALL	SPRING	SUMMER	WINTER	Grand Total	FALL	SPRING	SUMMER	WINTER	Min. Error rate	Total Error
Beitou	42	49	44	34	169	0.75147929	0.710059	0.739645	0.798817	0.710059172	120 SPRING
Da'an	43	34	33	40	150	0.713333333	0.773333	0.78	0.733333	0.713333333	107 FALL
Datong	18	22	26	18	84	0.785714286	0.738095	0.690476	0.785714	0.69047619	58 SUMMER
Nangang	28	13	22	19	82	0.658536585	0.841463	0.731707	0.768293	0.658536585	54 FALL
Neihu	38	54	34	45	171	0.777777778	0.684211	0.80117	0.736842	0.684210526	117 SPRING
Shilin	48	46	50	51	195	0.753846154	0.764103	0.74359	0.738462	0.738461538	144 WINTER
Songshan	30	30	30	37	127	0.763779528	0.76378	0.76378	0.708661	0.708661417	90 WINTER
Wanhua	37	48	35	34	154	0.75974026	0.688312	0.772727	0.779221	0.688311688	106 SPRING
Wenshan	23	30	28	36	117	0.803418803	0.74359	0.760684	0.692308	0.692307692	81 WINTER
Xinyi	39	18	26	30	113	0.654867257	0.840708	0.769912	0.734513	0.654867257	74 FALL
Zhongshan	50	62	61	50	223	0.775784753	0.721973	0.726457	0.775785	0.721973094	161 SPRING
Zhongzheng	38	32	31	28	129	0.705426357	0.751938	0.75969	0.782946	0.705426357	91 FALL
<b>Grand Total</b>	<b>434</b>	<b>438</b>	<b>420</b>	<b>422</b>	<b>1714</b>					Sum	1203
										Er. Rate	0.701867

Table 6:1R rule for house break-ins by seasons

- Error rate associated with 1R using Season = **70.1%**
- District Beitou, Neihu, Wanhua and Zhongshan have highest number of House Break-ins in Spring, with the error rate of 71%, 68%, 68% and 72%, whereas Zhongshan has almost equal number of House Break-ins in Summer as well with the error rate of 72%.
- Number of House Break-ins are maximum for Districts Da'an, Nangang, Xinyi, and Zhongzheng in Fall with the error rate 71%, 65%, 65% and 70% respectively.
- Similarly, Districts Shilin, Songshan and Wenshan have higher number of House Break-ins in winter with the error rate of 73%, 70% and 69%.
- District Datong has number of House break-ins maximum in Summer with the error rate of 69%.

#### For Auto Break-in:

Auto						Error Rate					
Row Labels	FALL	SPRING	SUMMER	WINTER	Grand Total	FALL	SPRING	SUMMER	WINTER	Min. Error rate	Total Error
Beitou	11	11	12	23	57	0.807017544	0.807018	0.789474	0.596491	0.596491228	34 WINTER
Da'an	3	3		2	8	0.625	0.625		0.75	0.625	5 FALL
Datong	4	8	5	1	18	0.777777778	0.555556	0.722222	0.944444	0.555555556	10 SPRING
Nangang	3	12	4	2	21	0.857142857	0.428571	0.809524	0.904762	0.428571429	9 SPRING
Neihu	3	7	6	5	21	0.857142857	0.666667	0.714286	0.761905	0.666666667	14 SPRING
Shilin	14	10	10	7	41	0.658536585	0.756098	0.756098	0.829268	0.658536585	27 FALL
Songshan	3	10	3	3	19	0.842105263	0.473684	0.842105	0.842105	0.473684211	9 SPRING
Wanhua	17	11	7	6	41	0.585365854	0.731707	0.829268	0.853659	0.585365854	24 FALL
Wenshan	5	2	4	3	14	0.642857143	0.857143	0.714286	0.785714	0.642857143	9 FALL
Xinyi	3	2	9		14	0.785714286	0.857143	0.357143	1	0.357142857	5 SUMMER
Zhongshan	1	3	2	6	12	0.916666667	0.75	0.833333	0.5	0.5	6 WINTER
Zhongzheng	2	3	2	2	9	0.777777778	0.666667	0.777778	0.777778	0.666666667	6 SPRING
<b>Grand Total</b>	<b>69</b>	<b>82</b>	<b>64</b>	<b>60</b>	<b>275</b>					Sum	158
										Er. Rate	0.574545

Table 7:1R rule for auto break-ins by seasons



- Error rate associated with 1R using Season= **57.45%**
- District Datong, nangang, Neihu, Songshan and Zhongzheng have maximum number of Auto break-ins in Spring with the error rate of 55%, 42%, 67%, 47% and 66% respectively.
- Districts Da'an, Shilin, Wanhua, and Wenshan have number of Auto break-ins maximum in Fall with the error rate 62%, 65%, 58% and 64% respectively.

### 7.1.2 1R algorithm using Day of the week for all districts

#### For House Break-in:

House Break-ins								Error Rate											
Row Labels	1	2	3	4	5	6	7	Grand Total	1	2	3	4	5	6	7	Min. Error Rate	Total Error		
Beitou	24	27	26	24	27	22	19	169	0.857988	0.840237	0.846154	0.857988	0.840237	0.869822	0.887574	0.840236686	142	2 Monday	
Da'an	21	18	26	17	25	28	15	150	0.86	0.88	0.826667	0.886667	0.833333	0.813333	0.9	0.813333333	122	6 Friday	
Datong	8	6	14	14	19	11	12	84	0.904762	0.928571	0.833333	0.833333	0.77381	0.869048	0.857143	0.773809524	65	5 Thursday	
Nangang	9	10	21	11	12	11	8	82	0.890244	0.878049	0.743902	0.865854	0.853659	0.865854	0.902439	0.743902439	61	3 Tuesday	
Neihu	19	36	26	20	25	29	16	171	0.888889	0.789474	0.847953	0.883041	0.853801	0.830409	0.906433	0.789473684	135	2 Monday	
Shilin	24	33	32	25	34	23	24	195	0.876923	0.830769	0.835897	0.871795	0.825641	0.882051	0.876923	0.825641026	161	5 Thursday	
Songshan	21	27	13	13	12	26	15	127	0.834646	0.787402	0.897638	0.897638	0.905512	0.795276	0.88189	0.787401575	100	2 Monday	
Wanhua	16	26	28	22	15	31	16	154	0.895104	0.831169	0.818182	0.857143	0.902597	0.798701	0.896104	0.798701299	123	6 Friday	
Wenshan	14	16	17	13	23	16	18	117	0.880342	0.863248	0.854701	0.888889	0.803419	0.861248	0.846154	0.803418803	94	5 Thursday	
Xinyi	23	19	14	12	20	12	13	113	0.79646	0.831858	0.876106	0.893805	0.823009	0.893805	0.884956	0.796460177	90	1 Sunday	
Zhongshan	29	38	29	38	21	40	28	223	0.869955	0.829596	0.869955	0.829596	0.90583	0.820628	0.874439	0.820627803	183	6 Friday	
Zhongzheng	10	21	23	22	17	20	16	129	0.922481	0.837209	0.821705	0.829457	0.868217	0.844961	0.875969	0.821705426	106	3 Tuesday	
Grand Total	218	277	269	231	250	269	200	1714								Sum	1382		
																Er. Rate	0.806301		

Table 8:1R rule for home break-ins by Day of the week

- Error rate associated with 1R using Day of the week= **80.6%**
- District Nangang has maximum number of Auto break-ins on Tuesday with the lowest error rate of 74.3%.
- Districts Beitou, Neihu, Songshan have number of House Break-ins maximum on Monday with the error rate 84%, 78.9%, 78.7% respectively.
- Zhongzheng is vulnerable district on all 3 days Monday, Tuesday and Wednesday almost equally, making it most unsafe district with respect to House Break-ins.

#### For Auto Break-in:

Auto Break-ins								Error Rate												
Row Labels	1	2	3	4	5	6	7	Grand Total	1	2	3	4	5	6	7	Min. Error Rate	Total Error			
Beitou	12	7	14	6	5	7	6	57	0.789474	0.877193	0.754386	0.894737	0.912281	0.877193	0.894737	0.754385965	43	3 Tuesday		
Da'an	1	2	2	1	1	1	1	8	0.875	0.75	1	0.75	0.875	0.875	0.875	0.75	6	2 Monday		
Datong			6	3	4	2	1	2	18	1	0.666667	0.833333	0.777778	0.888889	0.944444	0.888889	0.666666667	12	2 Monday	
Nangang	2	3	6	4	2	3	1	21	0.904762	0.857143	0.714286	0.809524	0.904762	0.857143	0.952381	0.714285714	15	3 Tuesday		
Neihu	3	3	2	3	5	2	3	21	0.857143	0.857143	0.904762	0.857143	0.761905	0.904762	0.857143	0.761904762	16	5 Thursday		
Shilin	3	9	3	9	7	3	7	41	0.926829	0.780488	0.926829	0.780488	0.829268	0.926829	0.829268	0.780487805	32	2 Monday		
Songshan	4	3	4	2	3	3		19	0.789474	0.842105	0.789474	0.894737	0.842105	0.842105	1	0.789473684	15	1 Sunday		
Wanhua	4	10	2	8	7	6	4	41	0.902439	0.756098	0.95122	0.804878	0.829268	0.853659	0.902439	0.756097561	31	2 Monday		
Wenshan	4	2	5	1	1	1	1	14	0.714286	0.857143	0.642857	0.928571	1	0.928571	0.928571	0.642857143	9	3 Tuesday		
Xinyi	2	2	3	1	2	2	2	14	0.857143	0.857143	0.785714	0.928571	0.857143	0.857143	0.857143	0.785714286	11	3 Tuesday		
Zhongshan	3	3	1	1	3	1		12	0.75	0.75	1	0.916667	0.916667	0.75	0.916667	0.75	9	1 Sunday		
Zhongzheng	1	3	4	1	1	1		9	0.888889	0.666667	0.555556	0.888889	1	1	1	0.555555556	5	3 Tuesday		
Grand Total	39	53	46	42	35	32	28	275								Sum	204			
																Er. Rate	0.741818			

Table 9:1R rule for auto break-ins by Day of the week

- Error rate associated with 1R using Day of the week = **74.18%**
- In case of Auto break-ins, Tuesday has maximum number of break-ins in the districts, Beitou, Nangang, Wenshan, Xinyi, and Zhongzheng with the error rate of 75%, 71%, 64.2%, 78.5%, 55.5%.
- Secondly, Monday has second largest count in number of Auto break-ins in districts, Da'an, Datong, Shilin and Wanhua with error rate 75%, 66%, 78%, 75%.

### 7.1.3 1R algorithm using Time of the day for all districts

#### For House Break-in:

House		Error Rate																		
Row Labels	01-03	04-06	07-09	10-12	13-15	16-18	19-21	22-24	Grand Total	01-03	04-06	07-09	10-12	13-15	16-18	19-21	22-24	Min. Error	Total Error	
Beitou	19	16	14	33	31	27	14	15	169	0.887574	0.905325	0.91716	0.804734	0.816568	0.840237	0.91716	0.911243	0.804734	136	10-12
Da'an	8	7	21	35	24	22	15	18	150	0.946667	0.953333	0.86	0.766667	0.84	0.853333	0.9	0.88	0.766667	115	10-12
Datong	12	7	15	11	16	10	7	6	84	0.857143	0.916667	0.821429	0.869048	0.809524	0.880952	0.916667	0.928571	0.809524	68	13-15
Nangang	7	5	16	14	7	10	13	10	82	0.914634	0.939024	0.804878	0.829268	0.914634	0.878049	0.841463	0.878049	0.804878	66	07-09
Neihu	10	20	22	27	31	21	22	18	171	0.94152	0.883041	0.871345	0.842105	0.818713	0.872193	0.871345	0.894737	0.818713	140	13-15
Shilin	23	13	19	27	32	28	33	20	195	0.882051	0.933333	0.902564	0.861538	0.835897	0.85641	0.830769	0.897436	0.830769	162	19-21
Songshan	8	5	17	38	14	12	22	11	127	0.937008	0.96063	0.866142	0.700787	0.889764	0.905512	0.826772	0.913386	0.700787	89	10-12
Wanhua	12	13	22	27	19	22	26	13	154	0.922078	0.915584	0.857143	0.824675	0.876623	0.857143	0.831169	0.915584	0.824675	127	10-12
Wenshan	11	9	16	18	15	15	13	20	117	0.905983	0.923077	0.863248	0.846154	0.871795	0.871795	0.888889	0.82906	0.82906	97	22-24
Xinyi	8	11	12	14	17	11	17	23	113	0.929204	0.902655	0.893805	0.876106	0.849558	0.902655	0.849558	0.79646	0.79646	90	22-24
Zhongshan	25	17	28	27	23	19	36	48	223	0.887892	0.923767	0.874439	0.878924	0.896861	0.914798	0.838565	0.784753	0.784753	175	22-24
Zhongzheng	12	8	16	26	12	17	19	19	129	0.906977	0.937984	0.875969	0.79845	0.906977	0.868217	0.852713	0.852713	0.79845	103	10-12
Grand Total	155	131	218	297	241	214	237	221	1714										Sum	1368
																			Er Rate	0.798133

Table 10:1R rule for auto break-ins by Time of Day

- Error rate associated with 1R using time of the day = **79.8%**
- In case of House break-ins, time period between 10-12 has maximum number of break-ins in Districts Beitou, Da'an. Songshan, Wanhua, and Zhongzheng with error rate of 80.4%, 76.6%, 70%, 82% and 79.8% respectively.
- Secondly, time period 22-24 has higher number of House Break-ins in districts Xinyi, Zhongshan, Wenshan with the error rate of 79%, 78% and 82%.
- Nangang is almost equally vulnerable at time 7-9 and 10-12 with the error rate 80% and 82%.
- District Neihu and Datong are vulnerable at 13-15 with the error rate of 80% and 81%.

#### For Auto Break-in:

Auto	Error Rate																											
Row Labels	01-03									04-06	07-09	10-12	13-15	16-18	19-21	22-24	Grand Total	01-03	04-06	07-09	10-12	13-15	16-18	19-21	22-24	Min. Error	Total Error	
Beitou		6	11	9	10	7	4	5	5	57	0.894737	0.807018	0.842105	0.824561	0.877193	0.929825	0.912281	0.912281	0.807018								46	04-06
Da'an		2		1	2	1			2	8	0.75	1	0.875	0.75	0.875	1	1	0.75	0.75								6	01-03
Datong		3	2	8	1	2	1	1		18	0.833333	0.888889	0.555556	0.944444	0.888889	0.944444	0.944444	1	0.555556								10	07-09
Nangang		4	2	6	1	3	1		4	21	0.809524	0.904762	0.714286	0.952381	0.857143	0.952381	1	0.809524	0.714286								15	07-09
Neihu		5	3	6		2	3	1	1	21	0.761905	0.857143	0.714286	1	0.904762	0.857143	0.952381	0.952381	0.714286								15	07-09
Shilin		9	3	7	8	4	5	2	3	41	0.780488	0.926829	0.829268	0.804878	0.902439	0.878049	0.95122	0.926829	0.780488								32	01-03
Songshan		2	3	3	2	1	3	3	2	19	0.894737	0.842105	0.842105	0.894737	0.947368	0.842105	0.842105	0.894737	0.842105								16	04-06
Wanhua		9	3	3	11	6	3	2	4	41	0.780488	0.926829	0.926829	0.731707	0.853659	0.926829	0.95122	0.902439	0.731707								30	10-12
Wenshan		4	1	3	2	1	3			14	0.714286	0.928571	0.785714	0.857143	1	0.928571	0.785714	1	0.714286								10	01-03
Xinyi		2		1	2	2	1	1	5	14	0.857143	1	0.928571	0.857143	0.857143	0.928571	0.928571	0.642857	0.642857								9	22-24
Zhongshan		3	1	1	1	1	1	2	2	12	0.75	0.916667	0.916667	0.916667	0.916667	0.916667	0.833333	0.833333	0.75								9	01-03
Zhongzheng		2	1	1	1		2	2		9	0.777778	0.888889	0.888889	0.888889	1	0.777778	0.777778	1	0.777778								7	01-03
Grand Total		51	30	49	41	29	25	22	28	275										Sum							205	
																				Er_Rate							0.745455	

Table 11:1R rule for auto break-ins by Time of Day

- Error rate associated with 1R using time of the day = **74.55%**
- Time period of 7-9 has highest number of Auto break-ins in District Datong, Nangang and Neihu with the error rate of 55%, 71.42% and 71.4%.
- Secondly, time period of 1-3 is vulnerable in case of districts Da'an, Shilin, Zhongshan and Zhongzheng with the error rate of 75%, 78%, 75%, and 77.7% respectively.
- Time 4-6 also represents risk of Auto break-ins for districts Beitou, and Songshan with the error rate 80% and 84%.
- District Xinyi represents minimum error rate 64.2% at 22-24 of Auto break-ins.

### 7.1.4 1R algorithm using Holidays

#### For House Break-in:

Error rate associated with 1R using Holiday Data = **7.7%**



House Break-ins				Error Rate					
District	Non Holidays		HolidGrand	Non Holidays	Holidays	Min. Err Rate	Total Error	Outcome	
Beitou	153	16	169	0.094675	0.905325	0.094675	16	Non Holidays	
Da'an	142	8	150	0.053333	0.946667	0.053333	8	Non Holidays	
Datong	82	2	84	0.023810	0.976190	0.023810	2	Non Holidays	
Nangang	78	4	82	0.048780	0.951220	0.048780	4	Non Holidays	
Neihu	162	9	171	0.052632	0.947368	0.052632	9	Non Holidays	
Shilin	180	15	195	0.076923	0.923077	0.076923	15	Non Holidays	
Songshan	121	6	127	0.047244	0.952756	0.047244	6	Non Holidays	
Wanhua	136	18	154	0.116883	0.883117	0.116883	18	Non Holidays	
Wenshan	106	11	117	0.094017	0.905983	0.094017	11	Non Holidays	
Xinyi	96	17	113	0.150442	0.849558	0.150442	17	Non Holidays	
Zhongshan	204	19	223	0.085202	0.914798	0.085202	19	Non Holidays	
Zhongzheng	122	7	129	0.054264	0.945736	0.054264	7	Non Holidays	
<b>Grand Total</b>	<b>1582</b>	<b>132</b>	<b>1714</b>			<b>Total Er</b>	<b>132</b>		
						<b>Min. Err Rate</b>	<b>0.0770128</b>		

Table 12:1R rule for house break-ins by Holiday

### For Auto Break-in:

Error rate associated with 1R using Holiday Data = **5.8%**

Auto break-in				Error Rate					
District	Non Holidays		HolidGrand	Non Holidays	Holidays	Min. Err Rate	Total Error	Outcome	
Beitou	53	4	57	0.070175	0.929825	0.070175	4	Non Holidays	
Da'an	7	1	8	0.125000	0.875000	0.125000	1	Non Holidays	
Datong	18		18	0.000000	1.000000	0.000000	0	Non Holidays	
Nangang	19	2	21	0.095238	0.904762	0.095238	2	Non Holidays	
Neihu	21		21	0.000000	1.000000	0.000000	0	Non Holidays	
Shilin	38	3	41	0.073171	0.926829	0.073171	3	Non Holidays	
Songshan	17	2	19	0.105263	0.894737	0.105263	2	Non Holidays	
Wanhua	40	1	41	0.024390	0.975610	0.024390	1	Non Holidays	
Wenshan	14		14	0.000000	1.000000	0.000000	0	Non Holidays	
Xinyi	12	2	14	0.142857	0.857143	0.142857	2	Non Holidays	
Zhongshan	11	1	12	0.083333	0.916667	0.083333	1	Non Holidays	
Zhongzheng	9		9	0.000000	1.000000	0.000000	0	Non Holidays	
<b>Grand Total</b>	<b>259</b>	<b>16</b>	<b>275</b>			<b>Total Er</b>	<b>16</b>		
						<b>Min. Err Rate</b>	<b>0.0581818</b>		

Table 13:1R rule for auto break-ins by Holiday

- 1R algorithm using Holiday/Non-Holiday attribute, results/outcome for all districts is Non-Holidays.
- As number of days which are not Holidays will be more than number of Holidays in any year, the number of Break-ins will be higher for Non-Holidays than those on Holidays, ultimately providing lower error rate for Non-Holidays in 1R analysis.
- Hence, using Holidays/ Non-Holidays as attribute to build 1R algorithm might give biased findings and misdirect the conclusion. Thus, even though this 1R algorithm has the minimum error rate amongst all, it's **not considered** for comparison with other three 1R algorithms

### The best 1R Rule:

Amongst all 1R algorithms built above, algorithm with Seasons as attribute is best as it has lowest error rate of 70.18% for House Break-ins, and 57% error rate for Auto break-ins.

## 7.2 Association Rule

This algorithm is a rule-based machine learning method. This algorithm has two phases:

- Finding the antecedents (if), an item that found in the data.
  - Finding the consequent (then), an item that found in the data in a combination with the antecedent.
- We can determine the association rules of if/then pattern by calculating the support, the count of the value or item in the data, and the confidence, the frequency of this if/then found to be true.

For this data, we used Association algorithms on 3 relevant variables step by step. Additionally, algorithms are built separately for House Break-ins and Auto break-ins to avoid biased results.

### Antecedents:

1. District
2. Season
3. Time of the day

**Consequent:** House Break-ins

**Support:** Number of House break-ins satisfying all antecedents.

**Error rate:** 1- Confidence = 1- (count satisfying all antecedents / total number of count for that instance)  
OR 1-(support count/ total count for that instance)

**Confidence:** Number of times consequent is true (House Break-ins or Auto break-ins).

**Minimum Error:** Minimum Error rate among all cases for those antecedents.

**Building a rule:** The rule is chosen based on minimum rate amongst all cases.

Here, to build the rule and for the sake of estimation, we used error rate (nothing but 1-confidence) and used minimum rate to find the best case matching all antecedents.

### 7.2.1 Association using District and Season:

#### For House Break-ins

House Break-ins					Error Rate							
District	FALL	SPRING	SUMMER	WINTER	Grand Total	FALL	SPRING	SUMMER	WINTER	Minimum	Total Error	
Beitou		42	49	44	34	169	0.751479	0.710059	0.739645	0.798817	0.710059	120 SPRING
Da'an		43	34	33	40	150	0.713333	0.773333	0.78	0.733333	0.713333	107 FALL
Datong		18	22	26	18	84	0.785714	0.738095	0.690476	0.785714	0.690476	58 SUMMER
Nangang		28	13	22	19	82	0.658537	0.841463	0.731707	0.768293	0.658537	54 FALL
Neihu		38	54	34	45	171	0.777778	0.684211	0.80117	0.736842	0.684211	117 SPRING
Shilin		48	46	50	51	195	0.753846	0.764103	0.74359	0.738462	0.738462	144 WINTER
Songshan		30	30	30	37	127	0.76378	0.76378	0.76378	0.708661	0.708661	90 WINTER
Wanhua		37	48	35	34	154	0.75974	0.688312	0.772727	0.779221	0.688312	106 SPRING
Wenshan		23	30	28	36	117	0.803419	0.74359	0.760684	0.692308	0.692308	81 WINTER
Xinyi		39	18	26	30	113	0.654867	0.840708	0.769912	0.734513	0.654867	74 FALL
Zhongshan		50	62	61	50	223	0.775785	0.721973	0.726457	0.775785	0.721973	161 SPRING
Zhongzheng		38	32	31	28	129	0.705426	0.751938	0.75969	0.782946	0.705426	91 FALL
Grand Total		434	438	420	422	1714						

Table 14: Association using District and Season for House Break-in

The table above represents the association of Antecedents District and Season with the consequent of Maximum Number of House break-ins season according to the error rate estimated. E.g. if district is Xinyi then maximum number of House break-ins are predicted in Fall with the error rate = 65.48%. In case of Districts, Beitou, Neihu, Wanhua and Zhongshan, season SPRING has maximum number of House break-ins with the error rate = 71%, 68.4%, 68.8% and 72%.

## For Auto Break-ins

Auto Break-ins										Error rate				Min Error	Total Error
District	FALL	SPRING	SUMMER	WINTER	Grand Total	FALL	SPRING	SUMMER	WINTER						
Beitou		11	11	12	23	57	0.807018	0.807018	0.789474	0.596491	0.596491			34	WINTER
Da'an		3	3		2	8	0.625	0.625	1	0.75	0.625			5	FALL
Datong		4	8	5	1	18	0.777778	0.555556	0.722222	0.944444	0.555556			10	SPRING
Nangang		3	12	4	2	21	0.857143	0.428571	0.809524	0.904762	0.428571			9	SPRING
Neihu		3	7	6	5	21	0.857143	0.666667	0.714286	0.761905	0.666667			14	SPRING
Shilin		14	10	10	7	41	0.658537	0.756098	0.756098	0.829268	0.658537			27	FALL
Songshan		3	10	3	3	19	0.842105	0.473684	0.842105	0.842105	0.473684			9	SPRING
Wanhua		17	11	7	6	41	0.585366	0.731707	0.829268	0.853659	0.585366			24	FALL
Wenshan		5	2	4	3	14	0.642857	0.857143	0.714286	0.785714	0.642857			9	FALL
Xinyi		3	2	9		14	0.785714	0.857143	0.357143	1	0.357143			5	SUMMER
Zhongshan		1	3	2	6	12	0.916667	0.75	0.833333	0.5	0.5			6	WINTER
Zhongzheng		2	3	2	2	9	0.777778	0.666667	0.777778	0.777778	0.666667			6	SPRING
Grand Total		69	82	64	60	275									

Table 15: Association using District and Season for Auto break-in

In SPRING season Districts Datong, Nangang, Neihu, Songshan and Zhongzheng have highest number of Auto break-ins with the error rate 55.5%, 42.8%, 66.7%, 47.3%, 66.67% respectively. Districts Da'an, Shilin, Wanhua, Wenshan have highest chance of Auto break-ins in FALL season. District Beitou and Zhongshan have higher chance of Auto break-ins in WINTER season with error rate of 59% and 50%. District Xinyi has higher chance of Auto break-ins in SUMMER with minimum error rate of 35.7%.

## 7.2.2. Association using District, season and Time of Day

### For House Break-ins

After associating District with Season, we joined 3rd attribute of Time in association with last two. The table below shows the association using antecedents District, Season and Time for the consequent of House or Auto break-ins with minimum error rate.

House Break-in									Error Rate										
Row Labels	01-03	04-06	07-09	10-12	13-15	16-18	19-21	22-24	Grand Total	01-03	04-06	07-09	10-12	13-15	16-18	19-21	22-24	Min. Error Rate	Total Error
FALL	47	35	53	75	66	48	62	48	434										
Beitou	7	2	2	7	11	8	4	1	42	0.833333	0.952381	0.952381	0.833333	0.738095	0.809524	0.904762	0.97619	0.738095238	31 13-15
Da'an	4	1	4	15	8	2	5	4	43	0.906977	0.976744	0.906977	0.651163	0.813953	0.953488	0.883721	0.906977	0.651162791	28 10-12
Datong	6	1	3	1	1	4	1	1	18	0.666667	0.944444	0.833333	0.944444	0.944444	0.777778	0.944444	0.944444	0.666666667	12 01-03
Nangang	2	2	6	4	2	5	4	3	28	0.928571	0.928571	0.785714	0.857143	0.928571	0.821429	0.857143	0.892857	0.785714286	22 07-09
Neihu	5	6	3	5	5	5	4	5	38	0.868421	0.842105	0.921053	0.868421	0.868421	0.868421	0.824737	0.868421	0.842105263	32 04-06
Shilin	6	8	1	8	8	5	6	5	48	0.875	0.833333	0.979167	0.833333	0.833333	0.895833	0.875	0.875	0.833333333	40 04-06
Songshan		2	7	5	4	3	7	2	30	1	0.933333	0.766667	0.833333	0.866667	0	0.766667	0.933333	0.766666667	23 07-09
Wanhua	1	3	5	9	6	3	8	2	37	0.972973	0.918919	0.964865	0.756757	0.837838	0.918919	0.783284	0.945946	0.756756757	28 10-12
Wenshan	3	1	6	1	3	3	1	5	23	0.869565	0.956522	0.73913	0.956522	0.869565	0.869565	0.956522	0.782609	0.739130435	17 07-09
Xinyi	6	4	8	6	6	2	2	5	39	0.846154	0.897436	0.794872	0.846154	0.846154	0.948718	0.948718	0.871795	0.794871795	31 07-09
Zhongshan	4	2	3	9	9	3	13	7	50	0.92	0.96	0.94	0.82	0.82	0.94	0.74	0.86	0.74	37 19-21
Zhongzheng	3	3	5	5	3	5	7	7	38	0.921053	0.921053	0.868421	0.868421	0.921053	0.868421	0.815789	0.815789	0.815789474	31 19-21
SPRING	44	30	58	73	48	57	72	56	438										
Beitou	4	7	7	8	6	7	4	6	49	0.918367	0.857143	0.857143	0.836735	0.877551	0.857143	0.918367	0.877551	0.836734694	41 10-12
Da'an	1	2	4	6	3	7	6	5	34	0.970588	0.941176	0.882353	0.821529	0.911765	0.794118	0.823529	0.852941	0.794117647	27 16-18
Datong	2	2	3	3	5	3	2	2	22	0.909091	0.909091	0.863636	0.861636	0.772727	0.863636	0.909091	0.909091	0.772727273	17 13-15
Nangang	2	1	2	2			1	5	13	0.846154	0.923077	0.846154	0.846154	1	0.923077	0.615385	1	0.615384615	8 19-21
Neihu	3	3	7	9	10	9	7	6	54	0.944444	0.944444	0.87037	0.833333	0.814815	0.833333	0.87037	0.888889	0.814814815	44 13-15
Shilin	8	1	4	8	4	5	10	6	46	0.825087	0.978261	0.913043	0.826087	0.913043	0.891304	0.782609	0.869565	0.782609696	36 19-21
Songshan	3		5	9	2	2	5	4	30	0.9	1	0.833333	0.7	0.933333	0.933333	0.833333	0.866667	0.7	21 10-12
Wanhua	4	1	4	8	5	8	12	6	48	0.916667	0.979167	0.916667	0.833333	0.895833	0.833333	0.75	0.875	0.75	36 19-21
Wenshan	5	3	4	5	3	5	4	1	30	0.833333	0.9	0.866667	0.833333	0.9	0.833333	0.866667	0.966667	0.833333333	25 01-03
Xinyi		1		2	2	3	6	4	18	1	0.944444	1	0.888889	0.888889	0.833333	0.666667	0.777778	0.666666667	12 19-21
Zhongshan	8	4	13	6	6	6	7	12	62	0.870968	0.935484	0.790323	0.903226	0.903226	0.903226	0.887097	0.806452	0.790322581	49 07-09
Zhongzheng	4	5	5	7	2	1	4	4	32	0.875	0.84375	0.84375	0.78125	0.9375	0.96875	0.875	0.875	0.78125	25 10-12
SUMMER	41	33	55	66	70	48	51	56	420										
Beitou	5	4	3	7	9	5	5	6	44	0.880364	0.909091	0.931818	0.840909	0.795455	0.886364	0.886364	0.863636	0.795454545	35 13-15
Da'an	3	2	5	4	7	6	2	4	33	0.909091	0.939394	0.848485	0.878788	0.787879	0.818182	0.939394	0.878788	0.787878788	26 13-15
Datong	3	3	5	3	7	2	1	2	26	0.884615	0.884615	0.807692	0.846154	0.730769	0.923077	0.961538	0.923077	0.730769231	19 13-15
Nangang	3	2	5	4	2	3	3		22	0.863636	0.909091	0.772727	0.818182	0.909091	0.863636	0.863636	1	0.772727273	17 07-09
Neihu	1	5	5	6	8	2	5	2	34	0.970588	0.852941	0.852941	0.823529	0.764706	0.941176	0.852941	0.941176	0.764705882	26 13-15
Shilin	6	3	10	5	12	6	7	1	50	0.88	0.94	0.8	0.9	0.76	0.88	0.86	0.98	0.76	38 13-15
Songshan	3	3	3	11	5			1	30	0.9	0.9	0.9	0.633333	0.833333	1	0.866667	0.966667	0.633333333	19 10-12
Wanhua	2	2	4	6	5	7	4	5	35	0.942857	0.942857	0.885714	0.828571	0.857143	0.8	0.885714	0.857143	0.8	28 16-18
Wenshan		2	4	4	4	3	4	7	28	1	0.928571	0.857143	0.857143	0.857143	0.892857	0.857143	0.75	0.75	21 22-24
Xinyi	2	2	2	3	4	3	3	7	26	0.923077	0.923077	0.923077	0.846154	0.846154	0.846154	0.730769	0.730769231	19 22-24	
Zhongshan	10	5	5	6	5	6	8	16	61	0.836066	0.918033	0.918033	0.901639	0.918033	0.901639	0.868852	0.737705	0.737704918	45 22-24
Zhongzheng	3		4	7	2	5	5	5	31	0.903226	1	0.870968	0.774194	0.935484	0.83871	0.83871	0.83871	0.774193548	24 10-12



WINTER	23	33	52	83	57	61	52	61	422										
Beltou	3	3	2	11	5	7	1	2	34	0.911765	0.911765	0.941176	0.675471	0.852941	0.794118	0.970588	0.941176	0.676470588	23 10-12
Da'an		2	8	10	6	7	2	5	40	1	0.95	0.8	0.75	0.85	0.825	0.95	0.875	0.75	30 10-12
Datong	1	1	4	4	3	1	1	1	18	0.944444	0.944444	0.777778	0.777778	0.833333	0.944444	0.833333	0.944444	0.777777778	14 07-09
Nangang				4	3	1	1	1	19	1	1	0.842105	0.789474	0.842105	0.947368	0.631579		0.631578947	12 22-24
Nelhu	1	6	7	7	8	5	6	5	45	0.977778	0.866667	0.844444	0.844444	0.852222	0.888889	0.866667	0.888889	0.822222222	37 13-15
Shilin			4	6	8	12	10	7	51	0.941176	0.908392	0.921569	0.882353	0.842105	0.764706	0.803922	0.862745	0.764705882	39 15-18
Songshan	2		2	13	3	7	6	4	37	0.945946	1	0.945946	0.648469	0.918919	0.810811	0.837838	0.891892	0.648648649	24 10-12
Wanhua	5	7	9	4		4	2		34	0.852941	0.794118	0.735294	0.882353	0.911765	0.882353	0.941176	1	0.735294118	25 07-09
Wenshan	3	3	2	8	5	4	4	7	36	0.916667	0.916667	0.944444	0.777778	0.861111	0.888889	0.888889	0.805556	0.777777778	28 10-12
Xinyi			3	2	3	5	3	6	30	1	0.866667	0.933333	0.9	0.833333	0.9	0.8	0.766667	0.766666667	23 02-24
Zhongshan	3	6	7	6	3	4	8	13	50	0.94	0.88	0.86	0.88	0.94	0.92	0.84	0.74	0.74	37 22-24
Zhongzheng	2		2	7	5	6	3	3	28	0.928571	1	0.928571	0.9	0.821429	0.785714	0.892857	0.892857	0.75	21 10-24

**Table 16: Association using District, Season and Time of Day for House break-in**

Using above association algorithm, we can give following examples,

- If season = Fall and time = 10-12 then Da'an and Wanhua have higher chances of House break-ins with minimum error rate of 65.1% and 75.6% respectively.
- If season = SPRING and time = 19-21 then district Nangang has higher chance of House break-ins with minimum error rate = 61.5%
- Season = SPRING, and time = 10-12, then district Songshan, Zhongzheng and Beitou have higher chances of House Break-ins with error rate = 70%, 78% and 83% respectively.
- For season = SUMMER, and time = 10-12 then district Wanhua has higher chance of House Break-ins with error rate = 63.3%.
- Season = WINTER and time = 22-24 the district Neihs has higher chance of House break-ins with the error rate = 63.15%.

### For Auto Break-ins

Auto Break-in									Error Rate									Min. Error Rate	Total Error Time affected
Row Labels	01'03	04'06	07'09	10'12	13'15	16'18	19'21	22'24	Grand Total	01'03	04'06	07'09	10'12	13'15	16'18	19'21	22'24		
FALL	19	7	8	15	6	2	6	6	89										
Beitou	4	2		3	1			1	11	0.616364	0.818182	1	0.727273	0.909091	1	1	0.909091	0.6163636	7'01'03
Da'an	1								3	0.666667	1	1	0.666667	1	1	1	0.666667	0.6666666	2'01'03
Datong	2		2	1					5	0.33	1	0.5	1	1	1	1	0.666667	0.6666666	2'01'03
Nangang	1				1			1	3	1	1	0.666667	1	0.666667	1	1	0.666667	0.6666666	2'07'09
Neihu			1						3	0.666667	1	0.666667	1	0.666667	1	1	1	0.6666667	2'01'03
Shilin	3	2	2	4		1	1	1	14	0.785714	0.857143	0.857143	0.714286	1	0.928571	0.928571	0.928571	0.714285714	10'10'12
Songshan	1				1	1			3	0.666667	1	1	0.666667	0.666667	1	1	1	0.6666667	2'01'03
Wanhua	5	2		5	2	1	2	3	17	0.705882	0.882353	1	0.705882	0.882353	0.941176	0.882353	1	0.705882353	12'01'03
Wenshan	2		1				2	5	0.6	1	0.8	1	1	1	1	0.6	1	0.6	3'01'03
Xinyi				1			1	1	3	1	1	1	0.666667	1	1	0.666667	0.666667	0.6666666	2'10'12
Zhongshan								1	1	1	1	1	1	1	1	1	0	0	2'22'24
Zhongzheng		1	1						2	1	0.5	0.5	1	1	1	1	1	0.5	1'04'06
SPRING	18	8	13	11	11	10	4	7	82										
Beitou	1			1	3	2	2	2	11	0.909091	1	1	0.909091	0.727273	0.818182	0.818182	1	0.727272727	8'13'15
Da'an	3			1	1				3	0.666667	1	1	0.666667	0.666667	1	1	1	0.6666666	2'01'03
Datong	1	2	3	1	1				8	0.75	0.75	0.625	0.875	0.875	1	1	1	0.625	5'07'09
Nangang	4	2	2	1	1			2	12	0.666667	0.833333	0.833333	0.316667	0.916667	1	1	0.833333	0.6666666	8'01'03
Neihu	3		3			1			7	0.571429	1	0.571429	1	1	0.857143	1	1	0.571428571	4'03'03
Shilin		1	2	2	1		1	1	10	0.7	1	0.9	0.8	0.8	0.9	1	0.9	0.7	7'01'03
Songshan	1	2	2			3	1		10	0.9	0.8	0.8	0.9	1	0.7	0.9	1	0.7	7'16'18
Wanhua	2	1	1	3	2	1		1	11	0.818182	0.909091	0.909091	0.727273	0.818182	0.909091	1	0.909091	0.727272727	8'10'12
Wenshan		1		1					2	1	0.5	1	0.5	1	1	1	0.5	1'08'06	
Xinyi						1			3	1	1	1	1	0.5	1	1	0.5	0.5	1'13'15
Zhongshan			1			1	1	1	3	1	1	0.666667	1	1	0.666667	0.666667	1	0.6666666	2'07'09
Zhongzheng	2						1		3	0.333333	1	1	1	1	0.666667	1	1	0.333333333	1'01'03
SUMMER	8	4	14	7	7	8	10	6	64										
Beitou	1	2	4	1	2	2	2	2	12	0.916667	0.833333	0.666667	0.916667	0.833333	1	0.833333	1	0.6666666	8'07'09
Datong			2		1	1	1		5	1	1	0.6	1	0.8	0.8	1	0.6	3'07'09	
Nangang								1	4	1	1	0.5	1	1	0.75	1	0.75	0.5	2'01'09
Neihu	1	2				2			6	0.833333	0.666667	1	1	1	0.666667	0.833333	1	0.6666667	4'04'06
Shilin	1		2	1	2	2	1	1	10	0.9	1	0.8	0.9	0.8	0.8	0.9	0.9	0.8	8'07'09
Songshan							2		3	1	1	0.666667	1	1	1	0.333333	1	0.333333333	1'09'21
Wanhua	2		1	2	1				7	0.714286	1	0.857143	0.714286	0.857143	1	1	0.857143	0.714285714	5'01'03
Wenshan		1				1	1	1	4	0.75	1	0.75	1	1	0.75	1	0.75	3'01'03	
Xinyi	2		1	1	1	1		3	9	0.777778	1	0.888889	0.888889	0.888889	0.888889	1	0.666667	0.6666667	6'22'24
Zhongshan							1		2	1	1	1	1	1	1	0.5	1	0.5	1'10'12
Zhongzheng				1			1		2	1	1	1	0.5	1	1	0.5	1	0.5	1'10'12
WINTER	6	11	14	8	5	5	2	9	60										
Beitou		7	5	5	1	2	1	2	23	1	0.695652	0.782609	0.782609	0.956522	0.913043	0.956522	0.913043	0.695652174	16'04'06
Da'an	1								2	1	1	0.5	1	1	1	1	0.5	0.5	1'07'09
Datong			1						1	1	1	0	1	1	1	1	1	0	0'07'09
Nangang			1		1				2	1	1	0.5	1	0.5	1	1	1	0.5	1'07'09
Neihu		1	2		1			1	5	1	0.8	0.6	1	0.8	1	1	0.8	0.6	3'07'09
Shilin	2	1	2	1		1			7	0.714286	0.857143	0.714286	0.857143	1	0.857143	1	1	0.714285714	5'01'03
Songshan		1							3	1	1	0.666667	1	1	1	1	0.333333	0.333333333	1'22'24
Wanhua				1	1	1	1	2	6	1	1	0.833333	0.833333	0.833333	0.833333	1	0.666667	0.6666666	4'22'24
Wenshan			1	1	1				3	0.666667	1	0.666667	0.666667	1	1	1	0.6666667	0.6666666	2'01'03
Zhongshan	3	1				1		1	6	0.5	0.833333	2	1	0.833333	1	1	0.833333	0.5	3'02'03
Zhongzheng							1	1	2	1	1	1	1	1	0.5	0.5	1	0.5	1'16'18

**Table 17: Association using District, Season and Time of Day for Auto break-in**

- If season = Fall, and time = 1-3, then districts Beitou, Da'an, Datong, Neihu, Songshan, Wanhua, Wenshan and Xinyi have the higher chance of Auto break-ins with minimum error rates amongst all.

- Datong has the minimum error rate of 50% having higher vulnerability associated with Auto break-in at 1-3.
- If season = SPRING, and time= 1-3 then district Zhongzheng has higher chance of Auto break-ins with minimum error rate of 33.33%.
- If season = SUMMER, and time = 19-21 then Songshan is highly vulnerable with the error rate = 33.3%.
- Districts Beitou, Datong, and Nangang have higher chance of Auto break-ins in SUMMER at 7-9am.

## 8. Results & Analysis

The below table shows the summary result for 1R rules with its corresponding error rate:

House Break-in				Auto Break-in			
District	Seasons(71.1%)	Day(80.63%)	Time(79.81%)	District	Seasons (57.45%)	Day(74.18%)	Time(74.55%)
Beitou	SPRING	Monday	10~12	Beitou	WINTER	Tuesday	04~06
Da'an	FALL	Friday	10~12	Da'an	FALL	Monday	01~03
Datong	SUMMER	Thursday	13~15	Datong	SPRING	Monday	07~09
Nangang	FALL	Tuesday	07~09	Nangang	SPRING	Tuesday	07~09
Neihu	SPRING	Monday	13~15	Neihu	SPRING	Thursday	07~09
Shilin	WINTER	Thursday	19~21	Shilin	FALL	Monday	01~03
Songshan	WINTER	Monday	10~12	Songshan	SPRING	Sunday	04~06
Wanhua	SPRING	Friday	10~12	Wanhua	FALL	Monday	10~12
Wenshan	WINTER	Thursday	22~24	Wenshan	FALL	Tuesday	01~03
Xinyi	FALL	Sunday	22~24	Xinyi	SUMMER	Tuesday	22~24
Zhongshan	SPRING	Friday	22~24	Zhongshan	WINTER	Sunday	01~03
Zhongzheng	FALL	Tuesday	10~12	Zhongzheng	SPRING	Tuesday	01~03

**Table 17: 1R Analysis Result Summary**

According to the findings of **1R algorithms**, it can be said that,

- Beitou needs more number of resources in Spring, as it has higher chances of House break-ins in Spring. Similarly, it is found that Monday is the most vulnerable day of the week in which 10-12 represents the most vulnerable time which needs more inspection/ patrolling in Beitou district. In addition, it needs attention in Winter, as well as Tuesday in a week at 4 to 6 am considering higher number of Auto break-ins.
- District Da'an needs more resources in Fall as well as on Friday and Monday at the time between 10 am to 12pm and 1 to 3 am considering House and Auto Break-ins.
- District Datong is vulnerable to House Break-ins on Thursday, specifically at 1 to 3 pm, additionally more number of House break-ins are there in Summer. Whereas, considering the auto break-ins the district, patrolling resources need to be managed more in Spring as well as on Monday at 7 to 9 am.
- District Nangang has higher chances of House Break-ins during Tuesday, specifically at 7 to 9am, additionally more number of House break-ins are there in Fall, which may need to have more patrolling resources in Fall. In addition, it needs attention in Spring, as well as on Tuesday in a week at 7 to 9 am considering higher number of Auto break-ins.
- District Neihu is vulnerable during Monday, specifically at 1 to 3 pm, additionally more number of House break-ins are there in Spring making it considered for more number of policemen/resources in the district for patrolling. In addition, it needs attention on Thursday in a week at 7 to 9 am considering higher number of Auto break-ins.

- District Shilin has higher chances of House Break-ins during Thursday, specifically at 7 pm to 9 pm, additionally more number of House break-ins are there in Winter, which may need more policemen/ resources in Winter. Additionally, it also requires more number of resources in Fall as well as on Monday at 1 to 3 am to control Auto break-ins in the area.
- District Songshan has higher chances of House Break-ins during Monday, specifically at 10 am to 12 pm, whereas more number of House break-ins are there in Winter, which may need more policemen/ resources in Winter. In addition, it needs attention in Spring, as well as on Sunday in a week at 4 to 6 am considering higher number of Auto break-ins.
- District Wanhua is vulnerable during Friday, specifically at 10 am to 12 pm, additionally more number of House break-ins are there in Spring making it considered for more number of policemen/resources in the district for patrolling. Additionally, it also requires more number of resources in Fall as well as on Monday at 10 to 12 pm to control Auto break-ins in the area.
- District Wenshan has higher chances of House Break-ins during Thursday, specifically at 10 pm to 12 am, whereas more number of House break-ins are there in Winter, which may need more policemen/ resources in Winter. In addition, it needs attention in Fall, as well as on Tuesday in a week at 1 to 3 am considering higher number of Auto break-ins.
- Similarly, district Xinyi needs more resources in Fall, as well as on Sunday at the time between 10 pm to 12 am considering number of House break-ins the areas, whereas to control Auto break-ins, district needs more number of resources in Summer as well as on Tuesday at 10 pm to 12 am.
- Zhongshan needs more number of resources in Spring, as it has higher chances of House break-ins in Spring. Similarly, It is found that Friday is the most vulnerable day of the week in which 10 pm to 12 am represents the most vulnerable time which needs more inspection/ patrolling in Zhongshan district. Additionally, it also requires more number of resources in Winter as well as on Sunday at 1 to 3 am to control Auto break-ins in the area.
- Similarly, district Zhongzheng needs more resources in Fall, as well as on Tuesday at the time between 10 am to 12 pm, considering the risk of House break-ins the area. Additionally, it also requires more number of resources in Spring as well as on Tuesday at 1 to 3 am to control Auto break-ins in the area.

According to the findings of **Association Rules**, there are few results which can be considered to manage the policemen resources district in particular season at certain time.

- In Fall, districts Beitou at 1-3 pm, Da'an at 10 am-12 pm, Wenshan at 7-9 am will require more attention to prevent House break-ins. Additionally, in Fall season, Districts Beitou, Datong and Wenshan requires more resources at 1 to 3 am in the morning to prevent Auto break-ins.
- Spring will require more attention on districts Songshan at 10-12 pm, Nangang at 7 pm-9 pm, and Zhongzheng at 10 am-12 pm, as they represent minimum error rate amongst all other districts for number of House Break-ins in the area. Additionally, considering number of Auto break-ins all the districts, in Spring season, Districts Zhongzheng, Neihu, Nangang, Da'an require more resources at 1 to 3 am in the morning.
- Season Summer will require more attention on districts Songshan at 10-12 pm, Xinyi and Zhongshan at 10 pm-12am, and Nangang at 1-3 pm, as they represent minimum error rate amongst all other districts for higher number of House Break-ins in the area. Additionally, considering number of Auto break-ins all the districts, in Summer season, Districts Songshan will require more

resources at 7 - 9 pm, whereas Beitou, Datong and Nangang require more resources at 7 to 9 am in the morning. In addition, Districts Zhongshan and Zhengzhong are equally vulnerable at 10- 12 pm in case of Auto Break-ins.

- Season Winter will require more attention on districts Beitou, Da'an, Songshan, Weshan, Zhongzheng at 10 am-12 pm, whereas Nangang at 10pm- 12 am, as they represent minimum error rate amongst all other districts for number of House Break-ins in the area. Additionally, considering number of Auto break-ins all the districts, in Winter season, Districts Da'an, Datong, Nangang and Neihu requires more resources at 7 to 9 am in the morning whereas Songshan requires attention at 10pm – 12am.

## 9. Conclusion & Recommendations

After analyzing results of both 1R and Association, we tried to find the most unsafe times for all districts in case of both House break-ins and Auto break-ins, so that policemen and patrolling resources can be rearranged/increased for districts in Taipei city depending on the recommendations provided.

- Considering the Figure 6 the density of population in each district and number of House Break-ins in that district over the period of 3 years (from 2015 to 2017), districts Zhongzheng, Zhongshan, Songshan and Da'an are **highly unsafe** in case of both House and Auto break-ins.
- Even though Districts Beitou and Shilin seem to have much lower population density due to its remote location and more hills, they have significantly larger number of Auto and House break-ins. The larger patrol route network would require more resources than other districts.
- Overall, Spring and Summer are seen to be having higher chances of Break-ins for both House and Auto Break-ins. Additionally, in case of House Break-ins, time between 10am to 12pm and 10pm to 12 am are seen to be affecting most of the districts. Whereas, 7am to 9am and 1 to 3 am are the most unsafe times with high probability of Auto break-ins in many districts.
- The time between 4 to 6 pm and 7 to 9 pm possess less chance of Auto break-ins, whereas times 4 to 6 am and 16 to 18 pm possess less chance of House break-ins according to the findings of both the algorithms. Hence, policemen or other resources used for these times can be rearranged and utilized for the times which are highly vulnerable.

Following are the findings in case of each district individually, as policemen and resources are generally arranged for each district separately in Taipei city.

- Beitou has maximum number of House Break-ins in Spring and Auto break-ins in Winter. It is recommended that resources should be added for this district in these seasons, specifically at 10pm-12am for House Break-ins and 4-6pm, for controlling Auto break-ins.
- Similarly, Da'an is vulnerable majorly in Fall in case of both kinds of break-ins. It's recommended that policemen and resources should be more in Fall to control Break-ins, specially between 10 am to 12 pm and 1 to 3 am.
- Datong has maximum number of House Break-ins in Summer and Auto break-ins in Spring. It is recommended that patrolling & resources should be increased for this district in these seasons, specifically at 1-3 pm for House Break-ins and 7-9 am for controlling Auto break-ins.
- Similarly, Nangang is one of the most vulnerable district in Fall in case of House break-ins and in Spring in case of Auto Break-ins. So, it's recommended that policemen and resources should be more in during these seasons to control Break-ins, specially between 7 am to 9 pm and 1 to 3 am.



- Similarly, Neihu is vulnerable majorly in Spring to both kind of break-ins. So, it's recommended that policemen and resources should be more in Spring to control Break-ins, specially between 1 to 3 pm and 1 to 3 am.
- Shilin has maximum number of House Break-ins in Winter and Auto break-ins in Fall. It is recommended that resources should be increased in these seasons, specifically at 4-6 pm for House Break-ins and 10pm-12 am for controlling Auto break-ins.
- Similarly, Songshan is seen to be unsafe majorly in Winter in case of House Break-ins and in Spring for Auto break-ins. So, it's recommended that policemen and resources should be more in Winter & Spring to control Break-ins, especially between 10 pm to 12am and 4 to 6 pm respectively.
- Wanhua is seen to be vulnerable to House Break-ins in Spring and for Auto break-ins in Fall. Additionally, there is need of more resources for patrolling between the time 19- 21 pm and 1-3 am respectively.
- Wenshan is estimated to be unsafe in Winter in case of House Break-ins and for Auto break-ins it is vulnerable in Fall. Hence, recommendation is to increase number of policemen and resources in the area for patrolling, especially during the time 10 am to 12 pm and 1 -3 am in the morning. This district has more population of retired people.
- In case of Xinyi district, maximum number of House Break-ins are seen in Fall where maximum number of Auto break-ins are seen in Summer. Hence, it is recommended that number of policemen for patrolling should be increased in these seasons for the time between 7 to 9 am and 10pm to 12 am to control Break-ins.
- Zhongshan is estimated to be unsafe in Spring in case of House Break-ins and for Auto break-ins it is vulnerable in Winter. Hence, recommendation is to increase number of policemen and resources in the area for patrolling, especially during the time 7 to 9 am and 1 -3 am.
- Similarly, Zhongzheng is seen to be unsafe majorly in Fall in case of House Break-ins and in Spring for Auto break-ins. So, it's recommended that policemen and resources should be arranged in such a way that patrolling rounds are more in Fall & Spring to control Break-ins, especially between 7 to 9 pm and 1 to 3 am respectively.

## 10. Scope for Future Research

1. The type crimes can be increased to include gang related violence, retail break ins and other crimes that can be mitigated by immediate police presence.
2. The recommended patrol routes can be better improved with additional data related to income group of victims, age of home/car owner, size of house, model of car. and information on sophistication of criminal i.e. if it was by an amateur or a professional.
3. If we get specific details regarding current route map and number of resources available to the police in each district, then we can make an optimized route map that will help them patrol the target areas better and thus ultimately make it more difficult for the crime to occur.

## 11. Appendices

(Attached separately)

**A:** APPENDIX A\_Break-ins Records\_Final Data analysis

**B:** APPENDIX B\_1R

**C:** APPENDIX\_Association rule

## 12. References

[1] Taipei City Government, website:

[https://english.gov.taipei/Content\\_List.aspx?n=02D12F5BE6C0FC93](https://english.gov.taipei/Content_List.aspx?n=02D12F5BE6C0FC93)

[2] Taipei City Police Department Statistic Table Yearly Report, website:

<https://police.gov.taipei/cp.aspx?n=41ACB3325FE2ACE3#7>