

# Developing model for effective implement of Disposal Cup Deposit system

Myunghee Shin

September 26, 2020

## 1. Introduction

Have you seen a pile of plastic cups in the street? In Seoul, Korea, I can easily see lots of disposable cups thrown up in bus stops and streets. After seeing garbage dumped on the streets, I felt that it is a serious problem for our environment. This problem is not only seriously damaging the city's aesthetics, but also increasing the burden of public management such as disposable cup processing costs and urban beautification costs. Therefore, we need to reuse and recycle the items including disposable cups for our future generation.

### 1.1 Background

Fortunately, the Korean government is going to implement a new law for preventing our environment. In order to solve the above problems, the government made Article 15-2 of the 「Act on the Promotion of Saving and Recycling of Resources」 which stipulates the 'Disposable Cup Deposit System' and promoted resource circulation of empty containers and disposable cups. Plan to be implemented, but no specific implementation plan has been established.

### 1.2 Problem

As I mentioned above, there is no specific implementation plan about this law yet. The Korean government will need an astronomical amount of money to implement this law to all of the cafes. So, I try to suggest enough number of cafe and unmanned recovery machine

for 'disposable cup deposit system' by region in Seoul Gang-nam. It could be really useful for Korean local government to manage this law implement. By this suggestion, I hope I can help preventing our environment effectively.

### **1.3 Purpose of analysis**

As a factor analysis for improving the effectiveness of the 'Disposable Cup Deposit System', which is scheduled to be implemented in June 2022, the best area to increase the recovery rate is selected, and the number of recovery systems required is calculated to present the policy direction.

## **2. Data acquisition and cleaning**

### **2.1 Data sources**

I would like to analyze the data and recommend Policy suggestions about Disposal cup Deposit system. I will use the data about cafe sales figures, the number and location of cafes, the number and location of offices, the number and location of cultural facilities, the number and location of bus stops, the number and location of subway stations in seoul and so on. All of these datas could be found on this website. (<https://data.seoul.go.kr/dataList/datasetList.do>)

So, I try to find out the degree of necessity of Disposal cup Deposit system by region of seoul. And I will find out the number of cafes or unmanned recovery machine for this system by region in seoul Gangnamgu. With this calculating model, We can find out a useful information in other region in Korea.

### **2.2 Data cleaning and Feature selection**

- 1) Sales of coffee shops/cafes by autonomous district in Seoul
  - Creation of derived variables related to monthly average total sales
  - 'Monthly average total sales (10,000 won)' = (first half sales + second half sales) / 2 \* 'number of cafes')
  - Extract required column

- Based on the data of 'Old' and 'Monthly average total sales (10,000 won)', create a 'Monthly average total sales (10,000 won)' column
- 2) Information on permits for restaurants in Seoul
  - Extract only data (rows) related to coffee shops and elements (columns) to use
  - Excluding the data for which the business status is closed
- 3) Seoul City Building Information
  - Extract required column
  - 'GIS building integrated identification number', 'Main use name' column extraction
- 4) Seoul city bus stop location information data
  - Distinguished aggregation
  - 'Number of bus stops' by grouping and counting the number
  - Buffer setting
- 5) Seoul Transportation Corporation address status data by station
  - Distinguished aggregation
  - 'Number of subway stations' is generated by counting the number by grouping 'gu'
  - Buffer setting

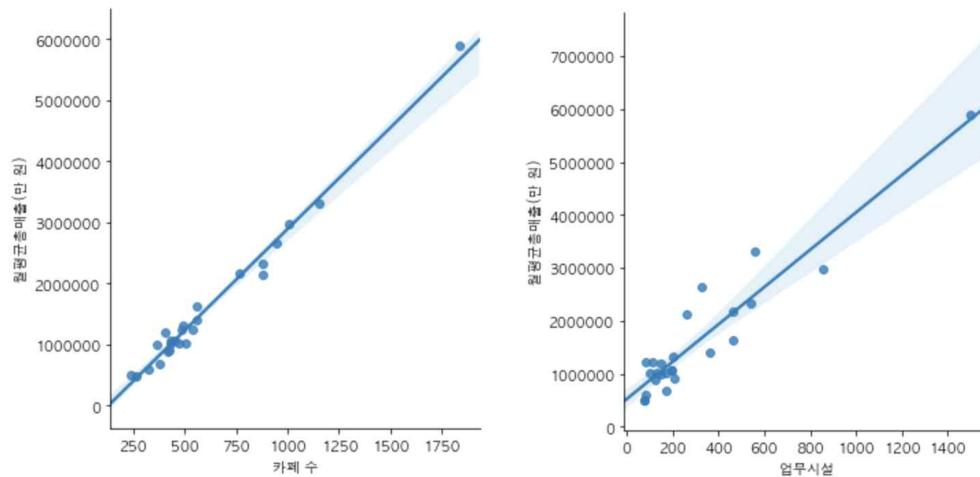
### 3. Exploratory Data Analysis

I analyzed significant factors affecting the use of disposable cups and user accessibility between autonomous districts. Because there is no data about disposable cup trash amount and recovery rate. So I tried to find out another factor to prevent the use of disposable cups. Therefore, I assumed that cafe sales figures(Y) would be related with disposable cup amount generated. And I did a correlation analysis about several factors related with cafe sales figures. So I tried to find necessity need this system by region in Seoul.

Furthermore, I assume that People would do participate in this law if the place of cafe and unmanned recovery machine is easy to access. So, I also tried to find a 'composite indicator' mixed with the

recovery accessibility and necessity need of this system by region in Seoul.

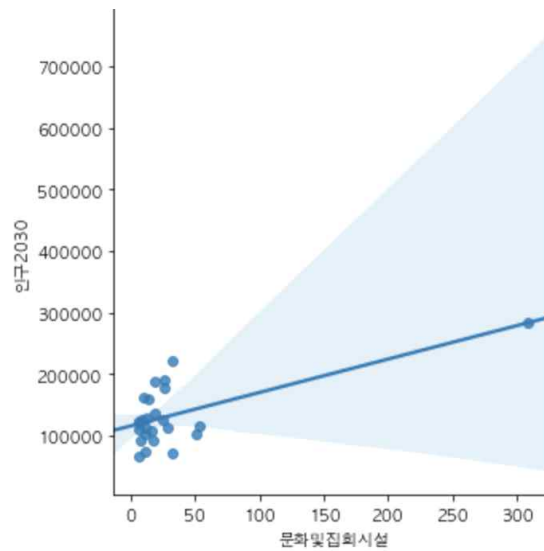
### 3.1 Relationship between cafe sales figures and several variables



[cafe sales figures(y)-number of cafe(x)] [cafe sales figures-number of office]

- Analysis of correlation between 'average monthly sales (10,000 won)' and 'number of cafes'
  - \* The correlation coefficient was 0.99 and the p-value was  $7.2e-22$ , which rejected the null hypothesis and adopted the alternative hypothesis. In other words, there is a correlation between average monthly sales and the number of cafes.
  - \* As the sales of cafes are high in regions with a large number of cafes, it can be said to be a factor that increases the amount of disposable cups generated.
- Correlation analysis between 'average monthly sales (10,000 won)' and the number of office buildings
  - \* The correlation coefficient was 0.86 and the p-value was  $1.0e-07$ , rejecting the null hypothesis and adopting the alternative hypothesis. In other words, there is a correlation between average monthly sales and business facilities.
  - \* It can be said that it is a factor that increases the amount of disposable cups because the sales of cafes are high in the business facility areas where office workers with high frequency of visiting cafes reside.

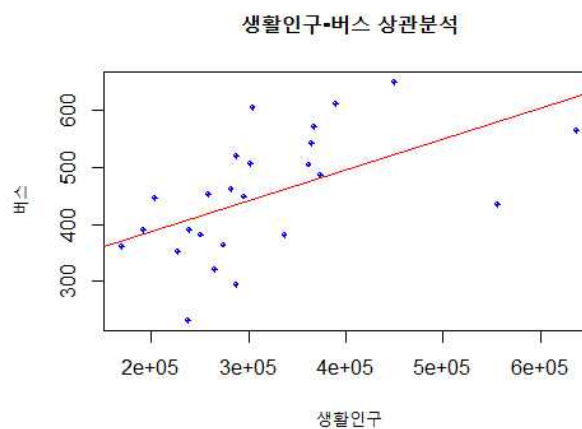
### 3.2 Relationship between number of age 2030 and cultural facilities



[number of age 2030(Y) - number of cultural facilities(X)]

- Correlation analysis of 'Population 2030' and number of cultural facilities
  - \* The correlation coefficient is 0.66 and the p-value is 0.00045, rejecting the null hypothesis and adopting the alternative hypothesis. In other words, it is judged that there is a correlation between the number of people in their 20s and 30s and culture and assembly facilities.
  - \* The number of cultural facilities can be said to be a factor that increases the amount of disposable cups, as 20-30 households who take out drinks in cafes often use cultural facilities.

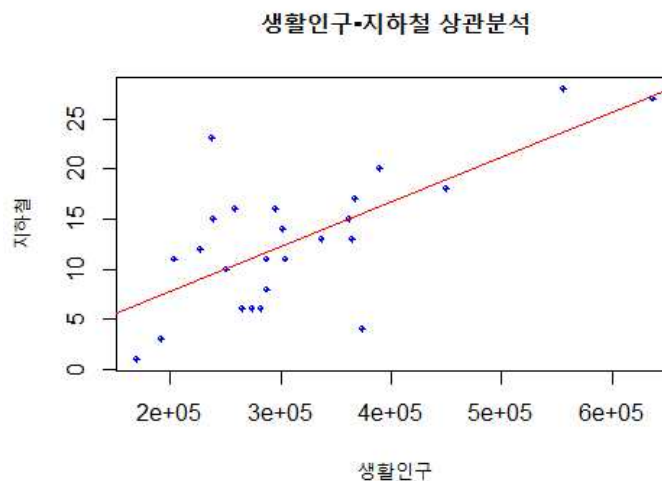
### 3.3 Relationship between living population and number of bus stops



[number of bus stops(Y) - living population(X)]

- Correlation analysis between living population and number of bus stops
  - \* The correlation coefficient was 0.5501 and the p-value was 0.004395, rejecting the null hypothesis and adopting the alternative hypothesis. In other words, there is a correlation between the living population and the number of bus stops.
  - \* Since there is a significant relationship between the living population and the number of bus stops, it can be said that it is a factor that improves accessibility of bus stops.

### 3.4 Relationship between living population and number of subway stations



[number of subway station(Y) - living population(X)]

- Correlation analysis of living population and number of subway stations
  - \* The correlation coefficient was 0.6944 and the p-value was 0.0001178, rejecting the null hypothesis and adopting the alternative hypothesis. In other words, there is a correlation between the living population and the number of subway stations.
  - \* Since there is a significant relationship between the living population and the number of subway stations, it can be said

that subway stations improve accessibility.

## 4. Predictive Modeling

### 4.1 Expenditure of Indexes

#### 1) Occurrence point regression analysis

- ☐ Substitute the amount generated by sales
- ☐ Perform regression analysis
  - Independent variables: number of cafes, number of business facilities, number of cultural facilities, number of short-term foreigners
  - Dependent variable: Cafe sales
- ☐ Derive a regression equation through R
  - $(\text{Cafe sales}) = (2.441\text{e}+03) * \text{number of cafes} + (9.540\text{e} + 02) * \text{number of business facilities} + (6.153\text{e} + 02) * \text{number of cultural facilities} + (3.286\text{e} +00) * \text{short-term foreigners} + (-2.253\text{e}+05)$
  - Significant probability (p-value)  $< 2.2\text{e}-16$ , which is significant

#### 2) Weight setting using mathematical method

- ☐ Mathematical method: Set the regression coefficient of regression analysis as weight
- ☐ Regression coefficient: number of cafes 2441, business facilities 954, cultural facilities 615, foreign short-term 3
- ☐ Short-term foreigner variable is removed and weight is set
- ☐ Number of cafes: Number of business facilities: Number of cultural facilities = 8: 3: 2 = 24: 9: 6
- ☐ Based on the regression model, based on the regression model, 8:3: Set the weight to 2

#### 3) User accessibility viewpoint weight is set by referring to literature and press releases.

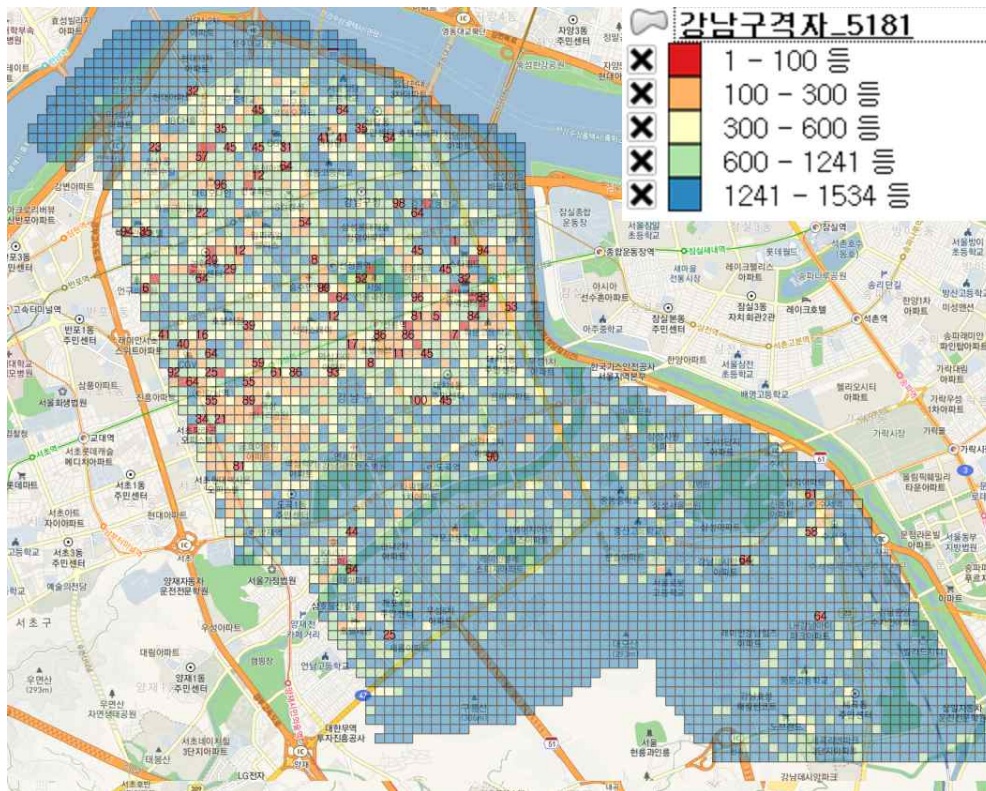
- ☐ Replace accessibility with living population
- ☐ Derivation of regression equation
  - Independent variable: number of bus stops, number of subway stations

- Dependent variable: living population
- ☐ Weight is arbitrarily set to 1:2
  - Bus stop: Subway = 1: 2 = 13: 26
- 4) Occurrence point and collection point are summed by 3: 1 weight
  - ☐ Derivation of regression equation
    - $Y \sim \text{number of cafes} + \text{business facilities} + \text{cultural facilities} + \text{bus stop} + \text{subway}$
  - ☐ Total result
    - Number of cafes: Business facilities: Cultural facilities: Bus stops: Subway = 72: 27: 18: 13: 26
  - ☐ Adjust the sum of weights to 1
    - Coffee: Business: Culture: Bus: Subway = 0.46: 0.17: 0.12: 0.08: 0.17

## 4.2 Visualization from QGIS

- ☐ Gangnam-gu grid unit visualization
  - Going further from the administrative building standard, we visualized the modeling in consideration of the amount of disposable cups and user accessibility in the entire Gangnam-gu in a grid method. As a result, it was confirmed that the 1st, 2nd, and 3rd ranks were evenly distributed in centers such as Yeoksam-dong, Samseong-dong, and Daechi-dong, whereas the outer regions of Gangnam-gu (Segok-dong, Gaepo-dong, Ilwon-dong, etc.) were distributed in the 4th and 5th ranks. Can
- ☐ The basic assumption of this study is verified that the area with a large amount of disposable cups will recover well.



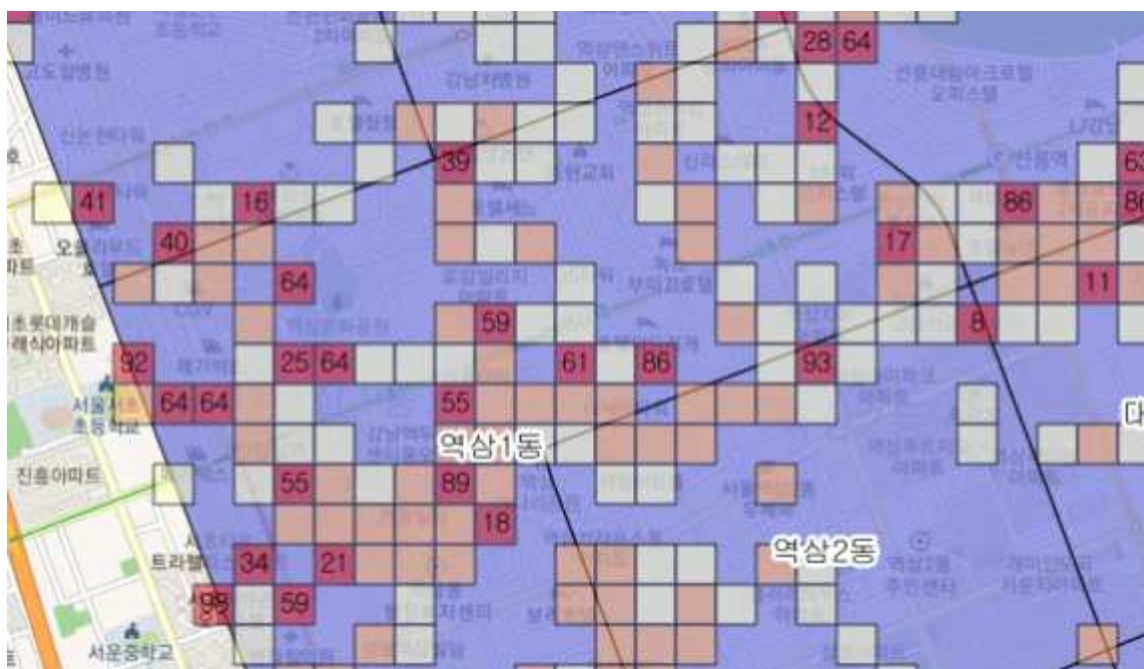


[Final visualization model (grid) considering both the amount of occurrence and user accessibility]

## 5. Conclusions

- Finding the location and required number of disposable cup collection systems (cafes and unmanned collection machines)
  - Visualize the best ranking (1-3) visualization model of the collection system. In the order of the highest modeling index, red grids from 1 to 100, orange grids from 101 to 300, and yellow grids from 301 to 600 were expressed. Therefore, it is judged that the red, orange, and yellow disposable cups are effectively collected in the order of red, orange, and yellow, and three, two, and one disposable cup collection systems for each color should be installed.
  - When the weight per grid is set to 3:2:1, the total required number of disposable cup collection systems in Gangnam-gu is 1,023 units.

※ The total number of cafes in Gangnam-gu is 1,800



[Collection system optimal ranking (1-3 ranking) visualization model]

- If the number of color grids is summed for each administrative building area, the results as shown in table below can be derived.
- With a distance of 105m per grid, at least one disposable cup collection system is required per grid of 3rd or higher.

Rank	dong	1 grid (red)	2 grid (orange)	3 grid (yellow)	total number
1	역삼1동	24	60	51	243
2	삼성1동	13	21	23	104
3	논현2동	6	17	26	78
4	대치2동	6	15	26	74
5	삼성2동	10	12	19	73
6	논현1동	10	5	25	65
7	압구정동	6	12	23	65

8	역삼2동	2	17	22	62
9	청담동	6	6	23	53
10	신사동	3	11	18	49
11	대치4동	6	7	9	41
12	도곡2동	3	5	9	28
13	도곡1동	0	3	12	18
14	수서동	2	3	4	16
15	개포4동	1	2	5	12
16	일원본동	1	1	3	8
17	개포2동	0	3	2	8
18	세곡동	1	0	4	7
19	일원2동	0	1	2	4
20	대치1동	0	0	2	2
21	일원1동	0	0	2	2
22	개포1동	0	0	1	1

- In Yeoksam 1-dong, there are 24 areas with priority 1, 60 areas with priority 2, and 51 areas with 3rd priority. Applying the previously set 3:2:1 weight, it is necessary to install a total of 72 disposable cup recovery systems in 24 first-order areas. In addition, 120 recovery systems should be installed in the second priority area and 51 recovery systems in the third priority area. Through this, the insight that Yeoksam 1-dong needs to install and manage a total of 243 disposable cup collection systems.
- Based on the above data, the government can calculate the number of management of the dong-specific or distinct

disposable cup collection system, and prepare and operate a more efficient management plan.

## 6. Future directions

- ※ In this analysis, Gangnam-gu is selected as an example and results are derived. By applying other autonomous district data based on the derived model, it is possible to analyze and predict the optimal area for each distinction and related information.

Assistance in implementing the one-time goods deposit system for each autonomous district

- ☐ This analysis can be used to implement and effectively manage the disposable deposit system in each autonomous district.
  - Data-based decision-making of each autonomous district can be supported through the model for selecting the best area for implementation stores, in which regions to increase the number of stores to be implemented, and in which locations to select and intensively manage them.
  - Through systematic management, urban aesthetics can be improved and public management burdens (environment beautification costs and disposable cup incineration costs, etc.) can be reduced.
- ☐ Based on the final model, the location of the disposable cup unmanned recoverer can be selected in the future.
  - Consider installing an unmanned collection device if there is no or insufficient cafe store at the location.