Big\_Data\_Application Project

Data Analysis of food services in highway rest area

*김지수* 1876092 , 김진유 1785054

key words: rest area, sales, variation of food provided

1. Introduction

This project provides a thorough analysis of food services provided in Korea’s highway rest-stop. The data used in this project includes the official data inquired from Korea Expressway Corporation ([http://data.ex.co.kr/portal/fdwn/ vi](http://data.ex.co.kr/portal/fdwn/%20vi)ew?type=ETC&num=R3&requestfrom=dataset#) and data collected from actual surveys.

The goal of the analysis is to provide a statistically meaningful result concerning the distribution of rest-area and variations of food preferences depending on consumer’s age.

2. Tables, Figures and Equations

2.1 Tables and Figures

Table 1 Area.

|  |  |  |
| --- | --- | --- |
| variable | type | annotation |
| id | integer | PRIMARY |
| Area | varchar |  |

**Table 1** contains 2 variables id and area. ‘Id’ indicates the primary key of area where the rest-stops are located. The area is divided into 9 parts grouped by states and variable ‘Area’ is the name of each grouped states.

Table 2 CustomerInfo.

|  |  |  |
| --- | --- | --- |
| variable | type | annotation |
| id | integer | PRIMARY |
| age | integer |  |
| Store\_type | integer | FOREIGN |
| sex | integer |  |
| Customer\_name | varchar |  |

**Table 2** contains 5 variables collected from real-life survey. It includes the age, name, sex and the category of food they mostly enjoy at rest\_stops.

Table 3 RestAreaInfo.

|  |  |  |
| --- | --- | --- |
| variable | type | annotation |
| id | integer | PRIMARY |
| name | integer |  |
| area | integer | FOREIGN |

**Table 3** contains 3 variables of rest-area existing in korea

Table 4 StoreInfo.

|  |  |  |
| --- | --- | --- |
| variable | type | annotation |
| id | integer | PRIMARY |
| Store\_name | varchar |  |
| Store\_type | integer | FOREIGN |

**Table 4** contains 3 variables of stores launched in rest area. It is the parent table of Rest\_Area\_Info

Table 5 StoreType

|  |  |  |
| --- | --- | --- |
| variable | Type | annotation |

|  |  |  |
| --- | --- | --- |
| id | integer | PRIMARY |
| Store\_type | integer |  |

Table 6 wholestore

|  |  |  |
| --- | --- | --- |
| variable | type | annotation |
| id | integer | PRIMARY |
| Ra\_id | integer | FOREIGN |
| Store\_id | integer | FOREIGN |
| Area\_id | integer | FOREIGN |

**Table 6** are mostly composed of FOREIGN KEYS from other tables. It acts as a sum of rest of the tables.

3. Related Work

The initial idea of the project was originated from an article introducing top 10 rest-area delicacy around the country.[1] We found out that despite the frequent use of rest-areas there weren’t enough data analyzed. So, we used the official data and borrowed the idea to order the rest-area by sales. Also, by modifying the top5 sales item of each rest-area we integrated sales data with customer’s age and their preferences together. The difference be

4. System Design and Requirements

4.1 Problem and Scope

The end-user of the project are people who need the rest-area information such as pre-founder of rest-areas or a business manager of rest-areas. They can use the distribution of rest-areas across the country and each of their sales and the popular stores of each rest-areas in business planning.

4.2 Component Design

4.3 Constraints and Non-Functional Requirements

* The system should provide clear and understandable data.
* Every user who needs the information should be able to browse the website freely.
* Only administrator can update, insert and delete items from the database.
* Login sessions should provide the authority of editing
* For high maintainability, php scripts should be analyzable and it must include sufficient amount of annotations.
* Data transaction should be precise and should catch exceptional case in Time behavior under 0.01s.

- The average response time is under 0.01 second since the functions are quite simple and concisely implemented.

- The current size of the database is total 192KB (total 6 tables all varying between 20~60KB) and can vary as more rest-areas open and close.

4.4 Abstract Design

- Code Base : PHP scripts, HTML

- Storage Type : SQL

- Authentication: PHP login Session

- separation between public and restricted areas

: Public area : All of the analyzed data

Restricted area : editing data

5. Advanced SQL Implementation

5.1 Pivot Table implementing GROUP\_BY and CASE

Pivot Table is a table-valued expression. It turns the unique values in one column into multiple columns in the output and performs aggregations on any remaining column values. Columns were selected from store\_type table, each column divided one by one with CASE function. Row were selected from area table and each items were calculated using SUM function.

5.2 Roll-up

ROLLUP is an extension of the GROUP BY clause. The ROLLUP option allows you to include extra rows that represent the total value(or whatever value wanted), which are commonly referred to as super-aggregate rows, along with the grand total row. In this project ROLLUP was used to represent sales of each store. Instead of plainly printing sales for each store, ROLLUP was use to represent the rest-area each store is included along with the sum of store sales which indicates sales of rest-area.

5.3 Ranking[2]

RANK() function is a window function that assigns a rank to each row within a partition of a result set. In this case, rank was implemented and ordered in a descending manner.

5.4 Dense Rank

DENSE\_RANK() is a window function that assigns a rank to each row within a partition of a result set. and returns consecutive rank values

5.5 CASE WHEN THEN[3]

WHEN statement specifies the condition to be tested. The THEN statement specifies the action if the WHEN condition returns TRUE. Case was used in this project to divide the sales value into parts.

6. User Scenario

(FYI. Stores differ from Rest-area. Stores indicate the shops that are launched in rest-area.)

6.1 Actor : Business Founder

Business Founder (henceforth founder) accesses to index.html, the main webpage of the system. Founder may choose from 4 different contents in the index.html.

1. rest\_area\_list.php

server sends a list of existing rest-areas with there are, name, and top5 stores launched with each of their sales ranked.

1. whole\_store\_list.php

server sends the list of top 5 stores from rest-area. Founder may choose to browse the list of stores according to the range of sales founder chooses to browse.

Founder may choose to browse a pivot table showing what kind of stores and how many there are in each area sector.

1. customer\_ranking.php

Server sends the list of food type and the average age of the people who prefer certain food\_types

Founder may choose to browse the top5 preferences for each age range. Founder can choose the age to browse starting from age 0-10 to 80+.

1. customer\_input.php

Server sends a form for user to fill out. It collects age, name , sex and the type of food he/she prefers.

6.2 Actor : Administrator

Administrator follows all the user cases of Business Founder. Additional to scenarios above, Administrator have access to edit function.

1. Login.php

Administrator must be authenticated before accessing edit.php. Administrator inputs admin-password to php session and if it matches, accessor is considered as administrator and automatically moves to the edit.php

1. Edit.php

Administrator can choose from 3 option.

Administrator (henceforth admin) may edit customer information. Server sends the result of customer survey in a form of data table. Admin choose what to delete or update from the list provided by the server.

Admin may insert an item to rest-area list. Admin types in the name of rest-area and chooses the area from the options given. Server receives the input and reflects it to the database instantly.

Admin may edit items by updating or deleting from the rest-area-list.

1. Logout.php

Once all the editing job is done admin must logout from the php session and terminate it safely.

6.3 Actor : Server System

Server accesses database creates table and insert items. For each requests from user(Administrator / Business Founder) Server executes SQL queries and returns the result.

7. Conclusion

This project was first devised to simply provide information to people who use rest-area in normal life. However, our team thought that sales data was too good to be wasted. So, we expanded the target user to business managers . By doing so, we were able to create statistically meaningful data such as sales according to types of stores or preferences and customer experiences.

We think that nowadays the market lacks online services related to rest-area. For further development we look forward to collect more customer surveys and increase the accuracy by actively advertising. Also, we will place importance on services for normal customers. The ultimate goal is for this website to come in handy and provide guidelines to the recommended food for each rest-areas.

References

[1]https://www.korea.kr/news/cultureColumnView.do?newsId=148840288

[2]<https://www.sqlservertutorial.net/sql-server-window-functions/>

[3] https://www.sqlshack.com/understanding-sql-server-case-statement/