

Application of Geographic Information of Statistics

Introduction

Geographic Information System (GIS) associates spatial data with related attribute data as well as process, query and analyze both at the same time. Through associating spatial data with attribute data, GIS provides functions such as querying, browsing, calculating statistics and analyzing maps. The GIS of Statistics visualizes spatial data distribution, pattern of clusters and statistics via maps to display implicit tendency and extraordinary data. It is beneficial to support various administrative decisions. Because the way that displays social-economic attribute data spatially facilitates decision makers to grasp more diverse information, it has great potential for many applications.

This article introduces the GIS of Statistics of the Ministry of the Interior (<http://moistgis.moi.gov.tw/moi92-1/login.cfm>) and demonstrates how to apply the system further to support administrative and business decisions.

The Development of GIS of Statistics of the Ministry of the Interior

With various functions of query, locating and statistic-calculation, the GIS of Statistics integrates map-display and query functions into an identical interface. Users can calculate statistics based on the buffer zone of an address, the buffer of a segment of road, the district or the basic statistic unit, as well as define a thematic platform to analyze certain project flexibly. It makes a breakthrough that calculates statistics based on not only traditional administrative districts but also flexible spatial units which are user-defined.

Considering issues of personal privacy and network security, the GIS of Statistics provides two interfaces with different authorities: one is for administrative users of government sectors and the other is for general users of the public. The former has more authorities of functions than the latter.

About the spatial data, some areas display the 1:1000 base maps while others display the 1:25000 topographic maps, it depends on whether there are address and base maps in the area. The areas with the 1:1000 base maps and address datasets include Taipei City, Kaohsiung City, Keelung City, Taichung City, Chiayi City and Tainan City, and the areas without the 1:1000 base maps and address datasets include Taipei County, Taoyuan County, Taichung County, Tainan County & Kaohsiung County. Consequently, in the former areas, the system provides more detailed analyzing functions, which are showed in Figure 1, than those in the latter ones.

Figure 1: The framework of system functions for administrative users (with the 1:1000 base maps and address datasets)

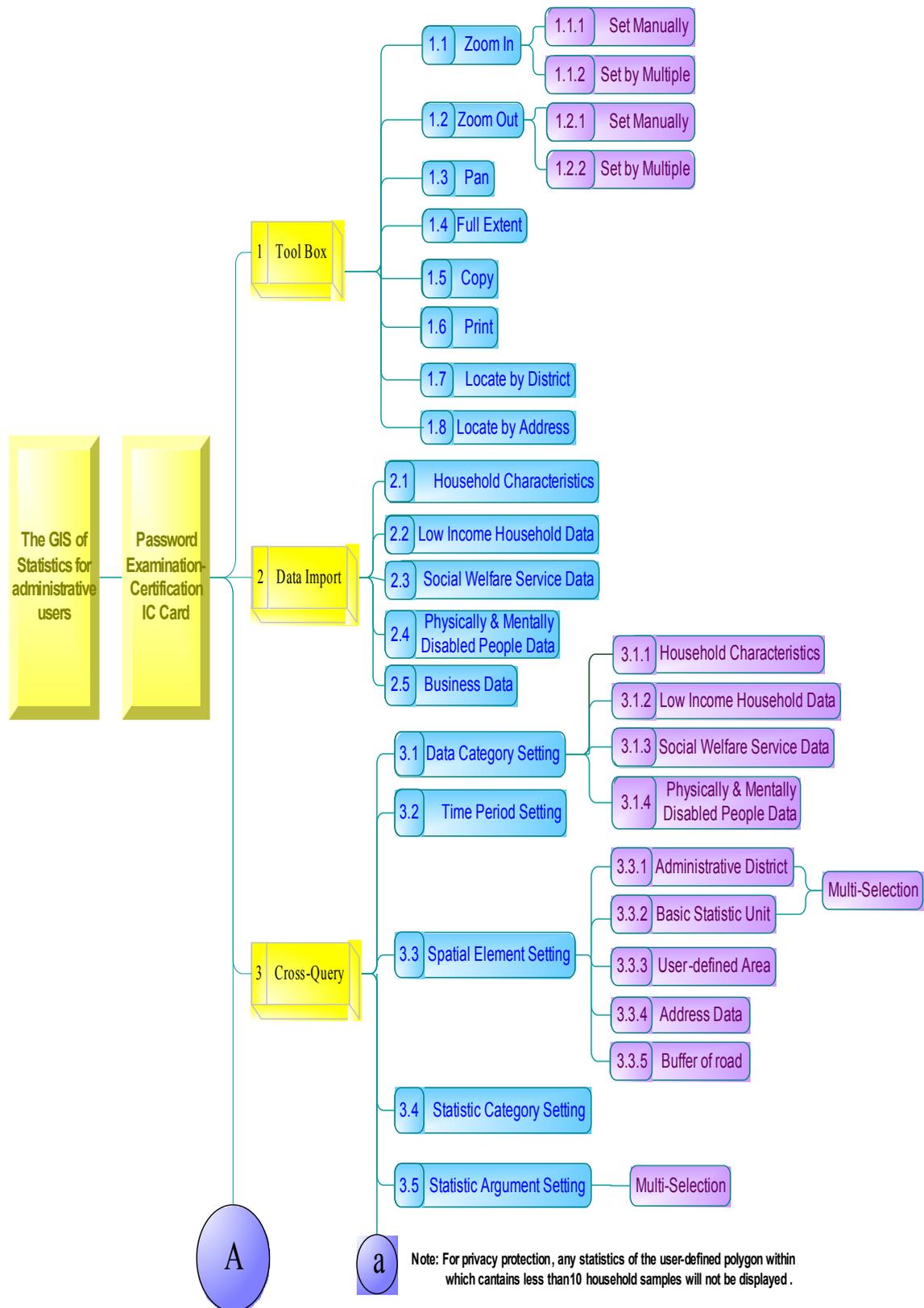
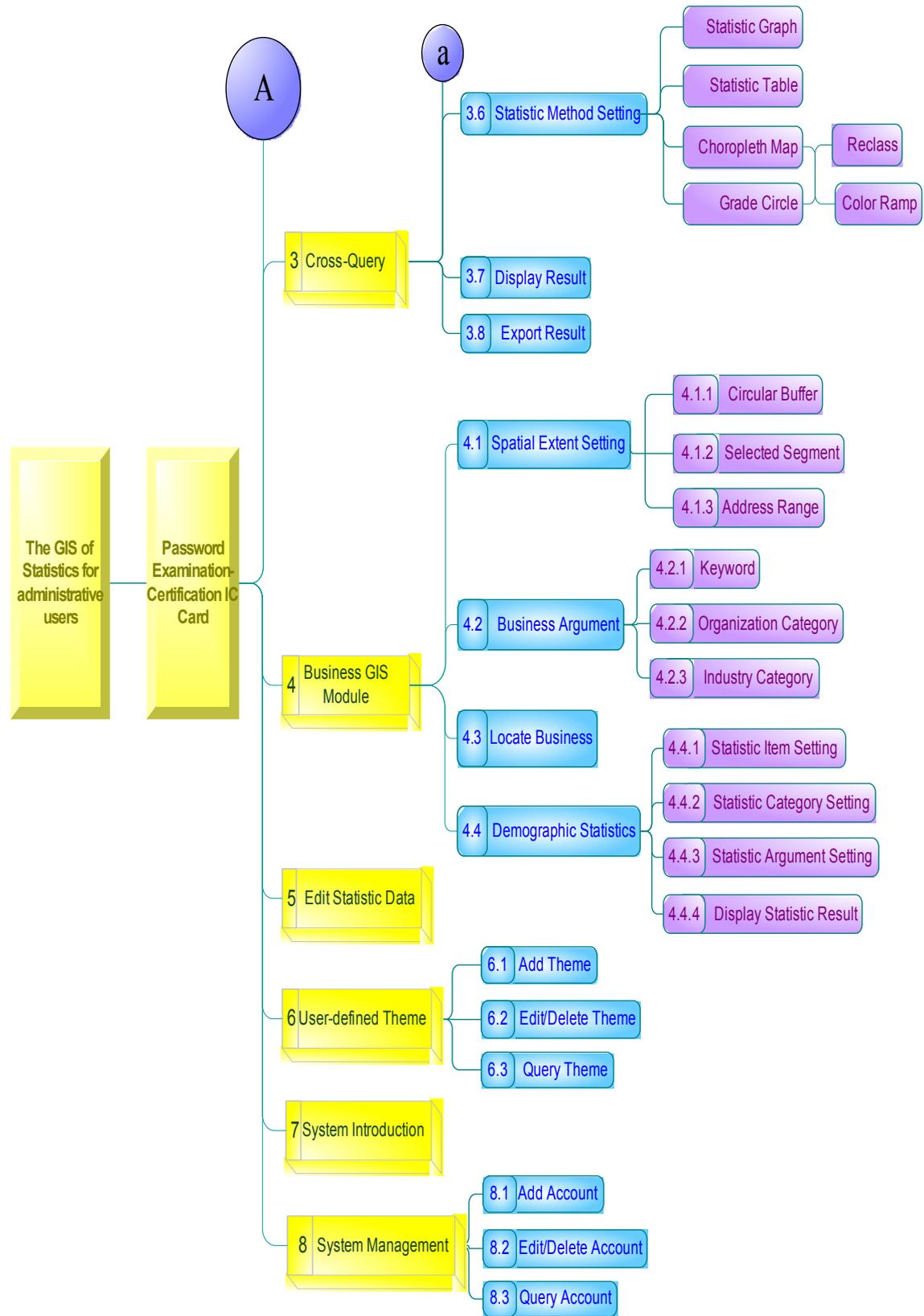


Figure 1(continued): The framework of system functions for administrative users (with the 1:1000 base maps and address datasets)



The Statistic categories and data items of the social-economic attribute database

are listed as Table 1. Each item of industrial, corporate and population attribute data can be cross-queried in an integrated way.

Table 1: Statistic categories and data items

Statistic Categories	Statistic Items & Statistic Arguments
1. Household characteristics	Gender, Age, Aborigine, Household, Marital status, Educational attainment
2. Characteristics of heads of low income household	Characteristics, Type, Illegal construction, Property right.
3. Low income household population characteristics	Appellation, Educational attainment, Counseled or not, Marital status, Health condition, Type of disability, Degree of disability, Type of profession, Insured or not.
4. Physically & mentally disabled people	Age, Gender, Type of disability, Degree of disability, Cause of disability.
5. Social welfare service institution	Type of institution, Address, Phone number, Main services.
6. business data	business name, address, type of organization, large/medium/small size industry type, main business items.

Case Demonstrations

Up to now, the GIS of Statistics has become increasingly complete and efficiently. The following will demonstrate several case applications in Taipei City to show the application potential of this system.

◆ Case 1: Market area of Carrefour

Carrefour is the second large retail corporation, which is an example of cooperation between French and Taiwan local business. To understand consumers' demands, the company spends 40 million dollars a year to study the market. Every time when Carrefour enters a new community, it always deployed a specialist to conduct a market survey first, the first step of the standard operation procedure to build a tremendous network of logistics.

According to the classical retailing theory, the location-selection of a shopping mall has to take the followings into consideration.

1. Estimate the potential of consumption in the market area. Centered in the site of the shopping mall, various market areas are defined by 5, 10 and 15

minutes' walk as well as 5, 10 and 15 minutes' driving trip. If there exists physical divides such a rail way or a block with competitors, the boundaries of market areas must be modified. After the zoning of the market areas, then the company analyzes the attributes of the population living in these regions. It includes calculating demographic indexes such as population density, age distribution, educational degree, occupational type, disposable income per capita and so on. The analysis that Carrefour conducts is more detailed. According to the distances and the disposable income index of these market areas, Carrefour discriminates critical market areas from ordinary market areas further.

2. Study the traffic and competitive situations around these market areas. The more convenient the traffic is, which means more bus stops or broader roads, in the neighborhood, the bigger the radius of the potential sales areas is.

The GIS of Statistics associates business information with demographic attributes, such as gender, marital status, educational degree, age, aborigine and household characteristics, and facilitates users to analyze market potentials based on various spatial elements, such as administrative districts, basic statistic units, addresses, user-defined polygons and user-selected streets. With the combination of geographic information (shown as Figure 2) and business data (shown as Figure 3), it enriches the business intelligence of analyzing sales performance, market potential and market trends. In the future, through the system, business like Carrefour can locate their outlets as well as analyze the demographic attributes within buffers of user-defined distances, and then incorporate the module of querying business data to find out market competitors as well as to formulate responsive strategies.

Figure 2: The module of querying demographic data

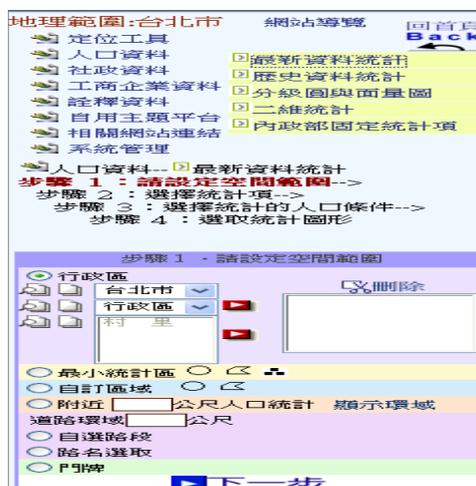
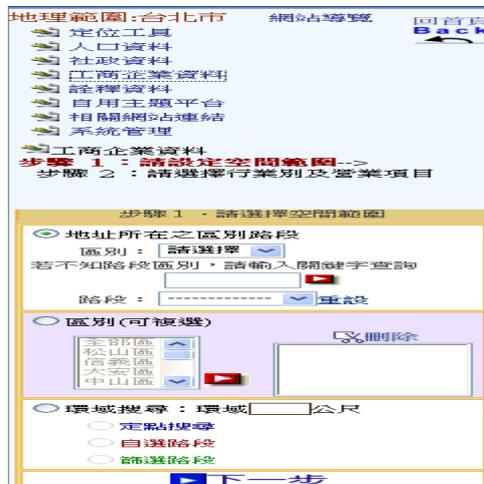


Figure 3: The module of querying business data



- ◆ Case 2: Distribution of 7-11 convenient stores (application of user-defined thematic platform)

The “user-defined thematic platform” module lets users define a specific theme to search related data and import these data into the system to analyze. It requires address datasets to locate features on the map.

According to the official website of 7-11, there are 617 stores in Taipei at the end of October in 1994. If we list all the addresses and names of every store in a EXCEL table and import the “add theme” function of this module, we can display the spatial distribution of the stores. Then, we use “theme vs. population query” or “theme vs. theme query” function, to conduct which kind of cross-analysis, joined with demographic data or with other thematic data (shown as Figure 4).

Entering the “business data” interface, and incorporating thematic data with industrial corporate data, we can search related traditional grocery stores. For example, there exists the Ton-Son Food Store, the Chen-Shin Store, Na-Bi-Me Food Store, the Chen-Yi Food Store, and Tan-Yu Corporation near the Shin-wan Outlet of 7-11 (shown as Figure 5). It illustrates how competitive the market is in the region and that the system can help us understand business operation in the neighborhood.

Figure 4: The spatial distribution and cross-analysis incorporating user-defined thematic data and demographic data

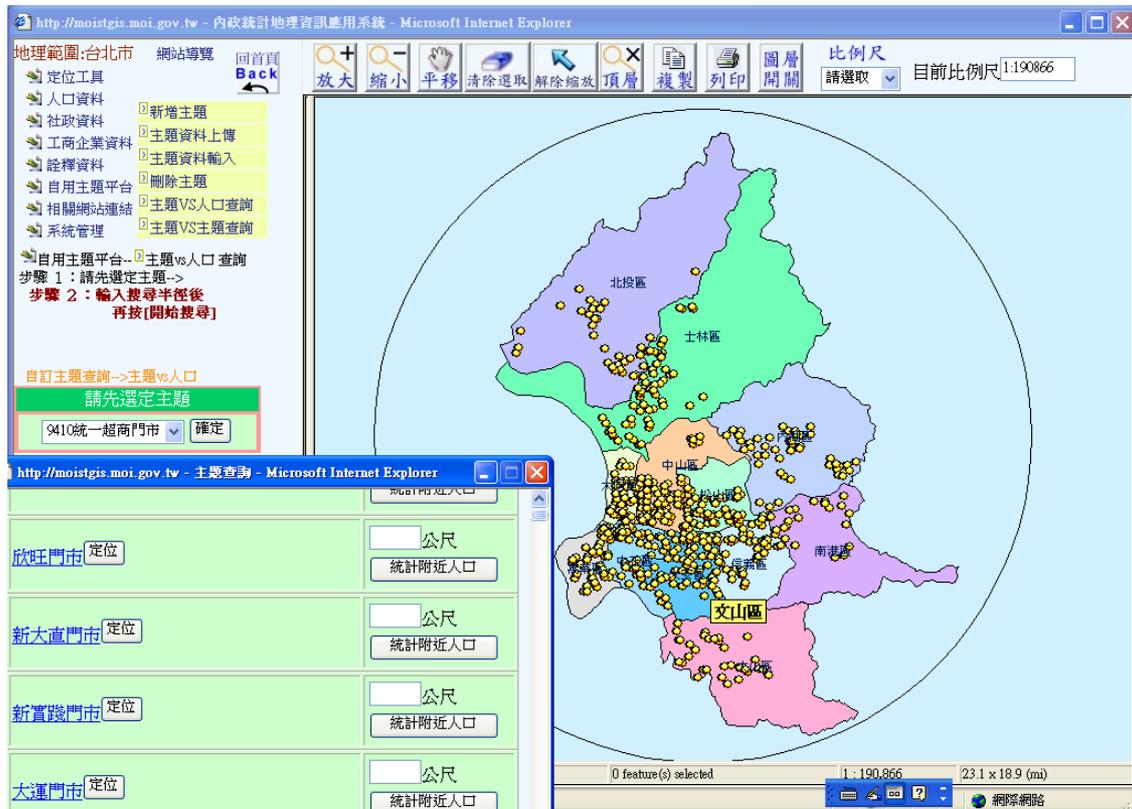
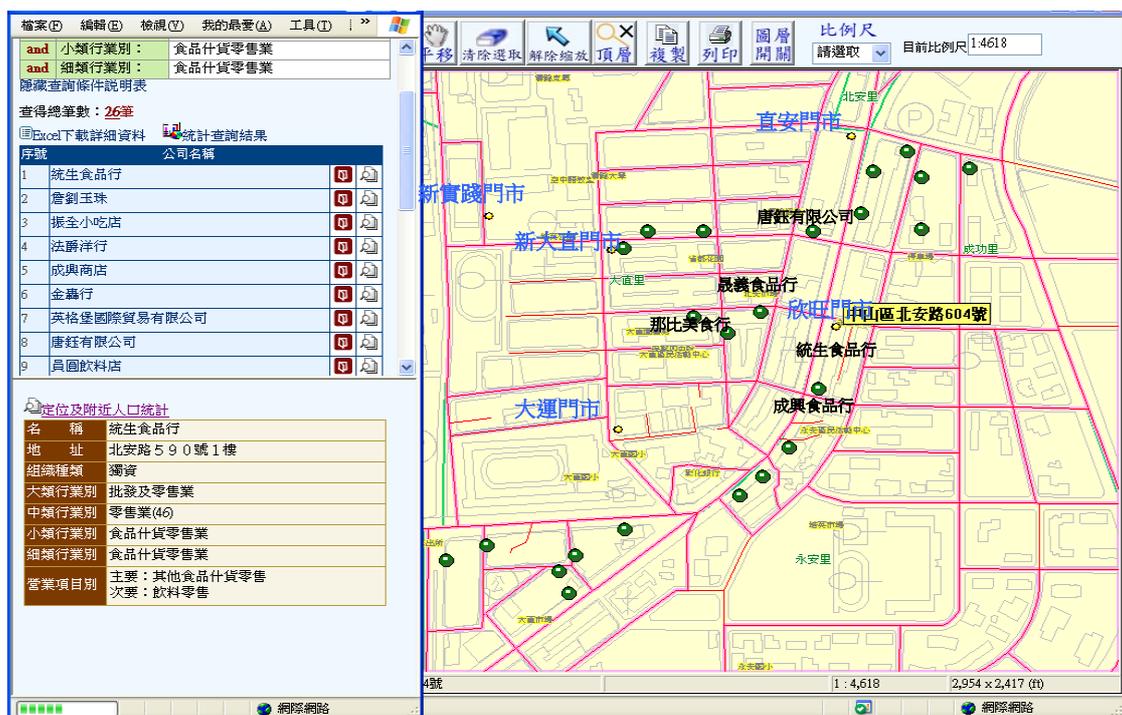


Figure 5: The locations of Shin-wan Outlet and other outlets or competitive retail stores



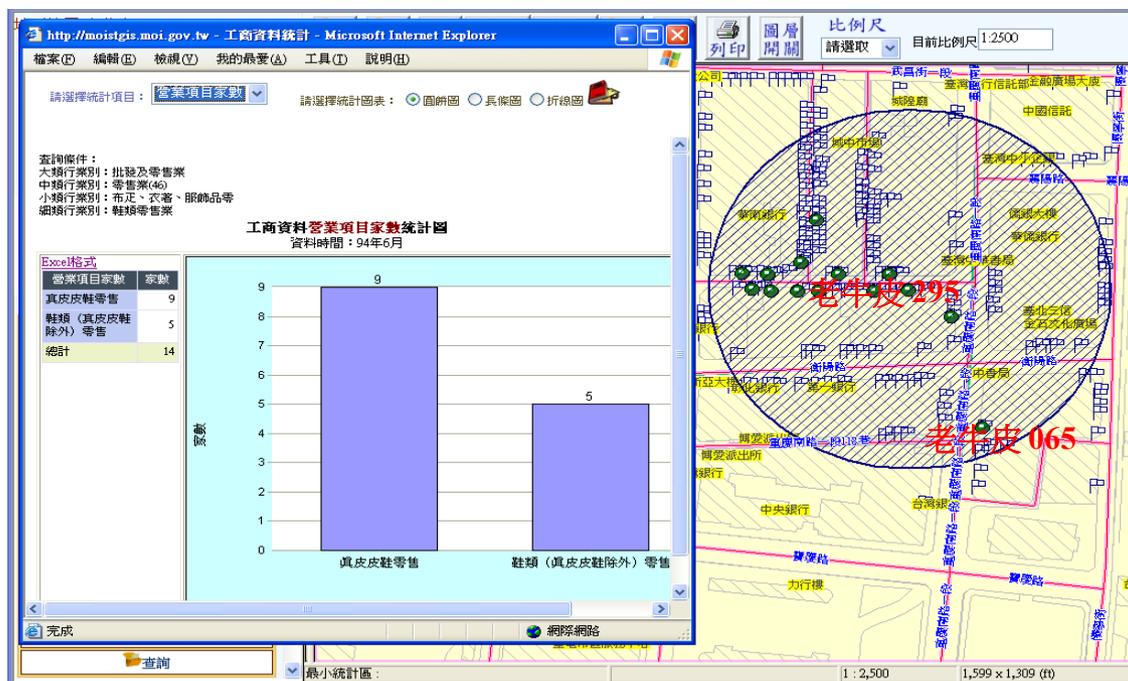
◆ Case 3: Query of business data

The system has incorporated and transferred business registration data into the module of “business geographic information application” to provide integrated application which can enhance the efficiency of regional planning, business administration, location selection of stores and academic researches.

For example, Yuanlin St. in Taipei City has long been famous as the “shoes street”. Entering the interface of “business GIS module”, we can figure out that there are 10 vendors of shoes along the street, which is not long. We can then not only calculate the statistics and download related attributes of these 10 companies, but also calculate the demographic statistics of the neighborhood.

In order to check whether there is any other competitor located nearby, we can run the function of “search within the buffer”. Figure 6 illustrates that the 120-meter buffer zone centered by the La New’s 295-th branch store exists other 14 shoes stores. Additionally, the La New’s 295-th branch store locates only 136 meters away from the 065-th branch store along the street. It might imply there is strong market demand in the region.

Figure 6: The locations of the competitors of the La New’s 295-th branch store with its 120-meter buffer zone



◆ Case 4: Query of social welfare service institution data

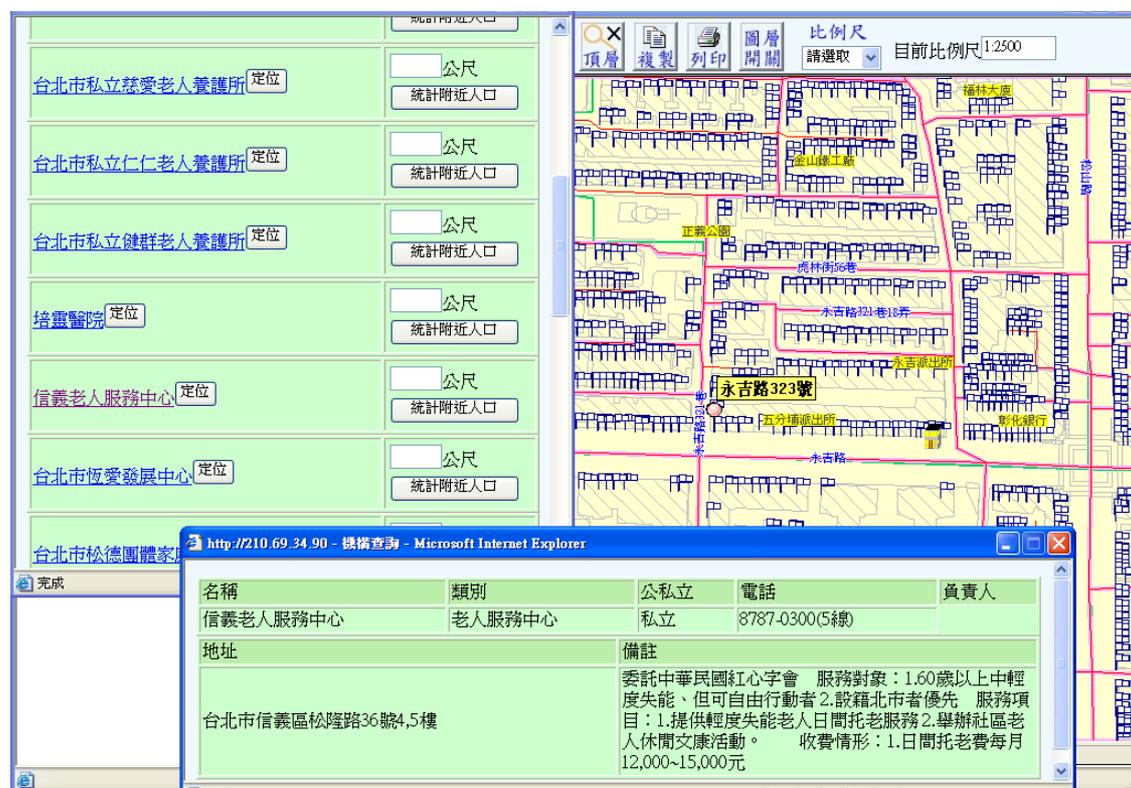
Let’s imagine a case that Mr. X is unable to take care of his slightly disoriented

father during the daytime so he wants to find any nursing facilities for old people in his neighborhood. By locating his home in the system, he can find that there are 17 social welfare service institutions within the 1,000 meters buffer area (shown as Figure 7) and glance over the services of the institutions.

After the procedures of query described above, Mr. X will know the profile of the Shin-Yi Elders' Care Center and realize he can reach it from his home along the Yong-Ji Road in the distance of 906 meters. If by foot, assuming the velocity is about 45 meters per minute, he needs to spend 20 minutes to get there.

Otherwise, the owner of the Shin-Yi Elders' Care Center can use the system to know that there are 35,727 people who are more than 60 years old within the 1500-meter buffer region centered by the care center. If assuming the velocity is about 75 meters per minute by foot, it costs about 20 minutes if these people want to come for services. Therefore, the care center can conduct a survey for their need. Besides, there are 78 similar institutions within the 1500-meter buffer region centered by the care center, hence, the care center can understand its competitors and formulate strategies for market segmentation.

Figure 7: Searching social welfare service institutions in the neighborhood



◆ Case 5: Two-dimensional statistic calculation with multi-argument setting

This function facilitates users to select two demographic items, such as gender, marital status, educational degree, age, aborigine and household characteristics as well as at most 4 statistic arguments to run cross-analysis.

For instance, take three attributes including divorce status, age and educational degree into consideration, there are most divorced population in the Chung-Shan District of Taipei City and the choropleth map of divorced people who are high-than-college educated in each Li zone of the Chung-Shan District is shown as Figure 8. If considering people who are more than 30 years old, the proportions of the divorced population with different educational degree category shows that the senior-high-school educated or junior-high-school educated population have high divorce ratio in Taipei, Taichung, Chiayi, Tainan and Kaohsiung (shown as Table 2), and it means high-than-college educated population are not the relatively high-divorce-ratio group.

Figure 8: The choropleth map of divorced population who are high-than-college educated in each Li zone of the Chung-Shan District

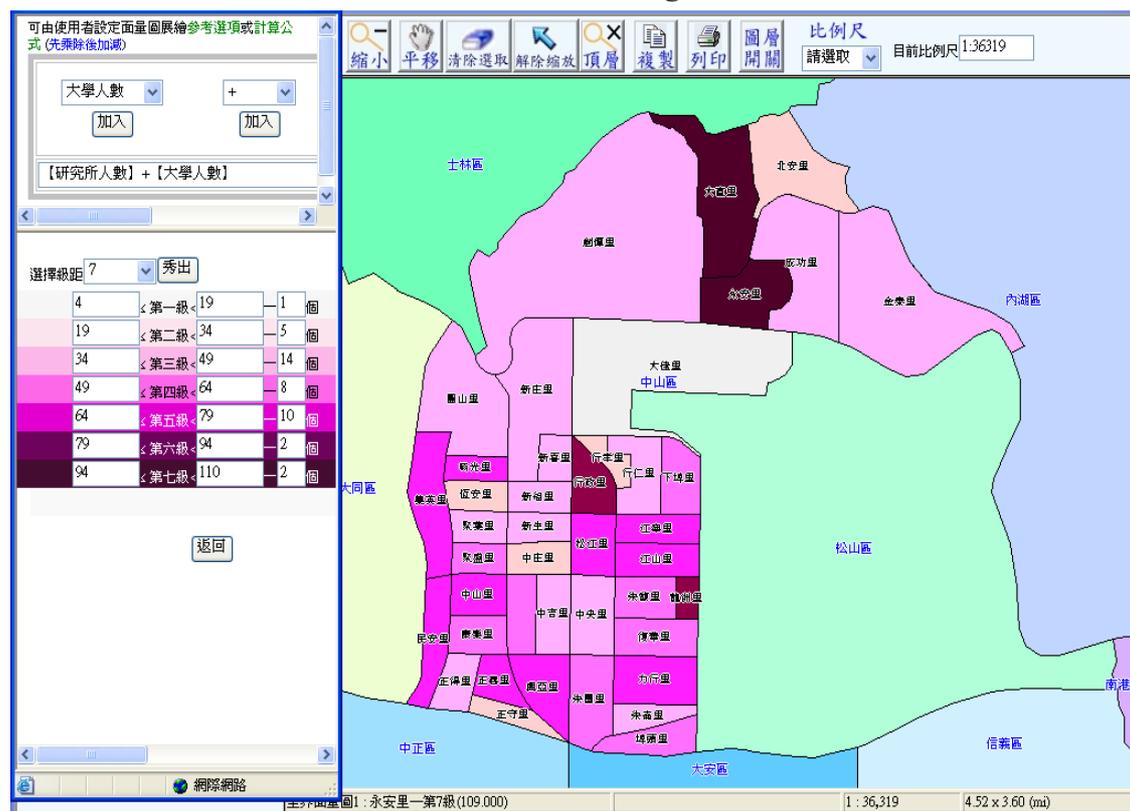


Table 2: The divorce ratio of people who are more than 30 years old

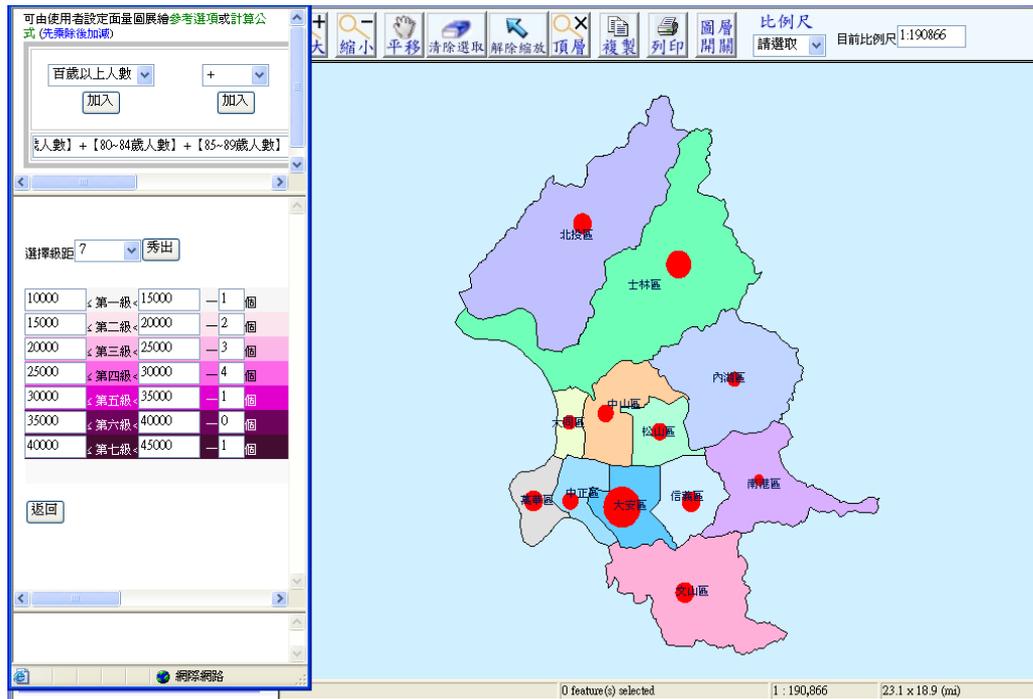
End of September, 2004		Unit : %			
Educational attainment	Taipei City	Taichung City	Chiayi City	Tainan City	Kaohsiung City
Total	7.55	8.93	7.10	7.91	8.77
Research Study	3.58	3.42	2.60	2.88	3.01
University/college	5.59	5.51	3.86	4.76	5.15
Specialist college	7.24	7.54	5.66	6.76	7.21
Senior high school	10.17	11.73	9.52	10.07	11.14
Junior high school	9.89	12.22	10.37	10.83	12.08
Elementary school	6.65	7.32	5.84	6.89	7.57
Self-studies	4.59	3.56	2.24	3.98	4.07
Illiterate	3.40	3.46	1.74	2.69	3.78

◆ Case 6: Grade circle maps and choropleth maps

Grade circle maps display the proportional variation with the area size of circle symbols scattered in different regions. Choropleth maps display the quantitative variation with a gradually changing color scheme on polygons, which represent different geographic regions not individual site points. Because psychologists suggest that people can only distinguish 7 different items simultaneously, the system provides 1 to at most 7 grades of classification options which can be defined by users.

With the trends of extension of human life and decreasing of birth rate, the aging of population is more prevalent. By using the query functions of “demographic data” and “latest statistics”, we can see there are more than 290,000 people who are more than 65 years old in Taipei City, which accounts for 11.1% of total population in Taipei City. In the grade circle map displaying the distribution of the old population in Taipei City, different district shows different size of circles. The bigger the circle is, the more the old people there are. The system shows that the number of old population in Da-an District is the greatest while the number of old population in Nan-kang District is the least (shown as Figure 9). Conclusively, we can grasp the demand for the service of the old simply from the visualization functions, “grade circle maps and choropleth maps”, of the system.

Figure 9: The grade circle map displaying the distribution of the old population in Taipei City



◆ Case 7: Query of Physically and mentally disabled people

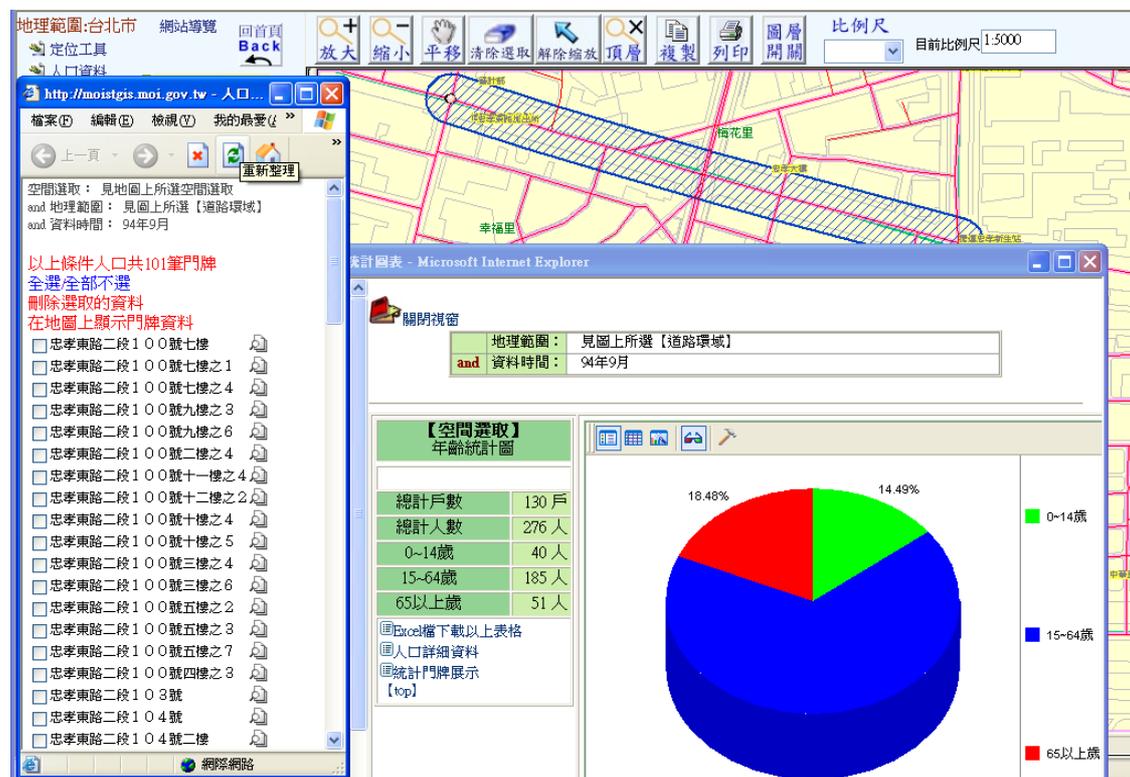
Physically and Mentally Disabled Citizens Protection Act defined disabled people as those who are limited or can not fully participate in activities of societal nature or production nature, either due to physiological or psychological reasons. The disabled can be divided into 16 types and 4 levels. According to the survey of the health care authorities of the central government, the needs of the disabled diversify from medical care, educational service, working situation to living service. Through the system, we can understand the reason, level and type of the disabled as well as their demographic attributes, such as gender, age and so on, and conduct the cross-analyses to set related policies.

Entering the functions of the disabled in the “social service data” category, defining the interested region and setting statistic item and arguments, we can get statistical tables and graphs (shown as Figure 10) to show cross-analyses.

- ◆ Case 8: Population statistic-calculation based on the buffer zone of given road segment selected by road name or user-defined

The statistics along the specific road segment often apply to traffic accident statistics, traffic flow statistics, police territory management, rent statistics of store, garbage can statistics, MRT targeted customers statistics, and so on, and can generate more benefits if incorporated with demographic data. For example, if we want to know the population of the 30-meter buffer of a certain MRT route segment, we can select the interested segment and input the range of the buffer, and then perform the calculation of demographic statistics. The example of Sect. 2 of Chung-shao East Road is shown as Figure 12.

Figure 12: The demographic statistics within the 30-meter buffer zone of Sect. 2 of Chung-shao East Road



- ◆ Case 9: Application of basic statistical units

The blocks in the city are the most integrated elements with physical boundaries and are suitable to be regarded as basic geographic units. The basic statistical units in the project are the close areas which are surrounded by geographic features, streets and other physical boundaries. With the basic statistical units, we can display social-economic attributes without violating personal privacy. The major framework of basic statistical units is street blocks. In the system, we can select multiple basic statistical

units, such as the illustrated in Figure 13. The educational levels of the population in the basic statistical unit covering the Shin-wan outlet of 7-11 are shown in Figure 14.

Figure 13: The function of multi-selecting basic statistical units

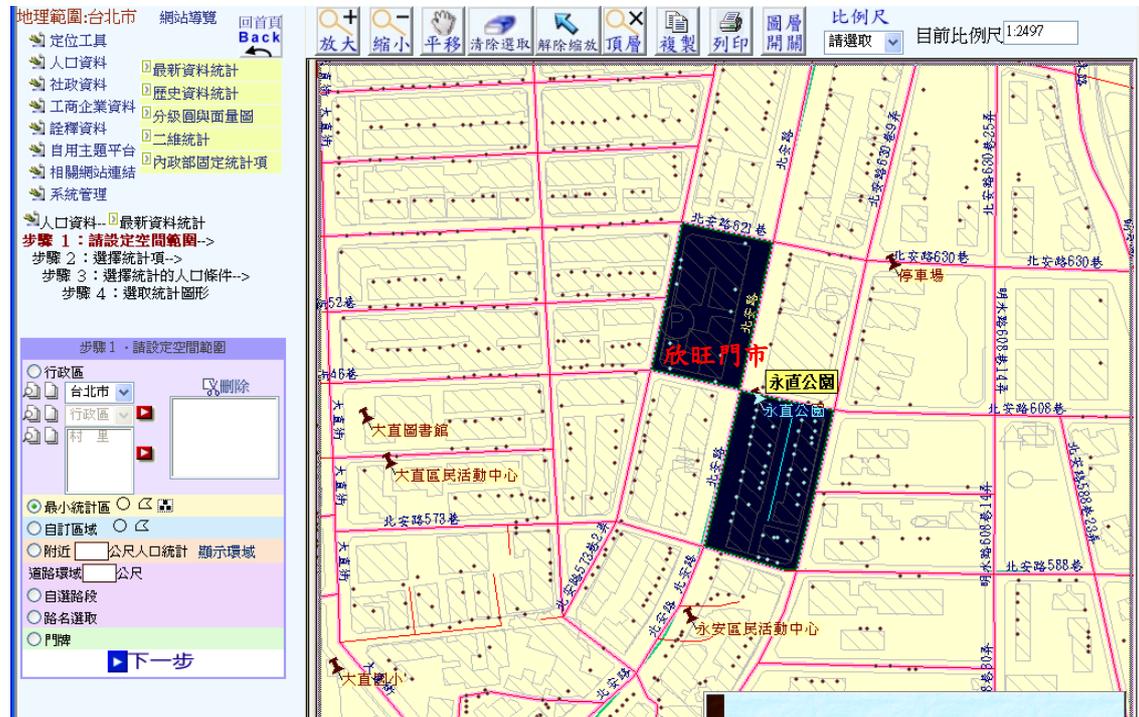
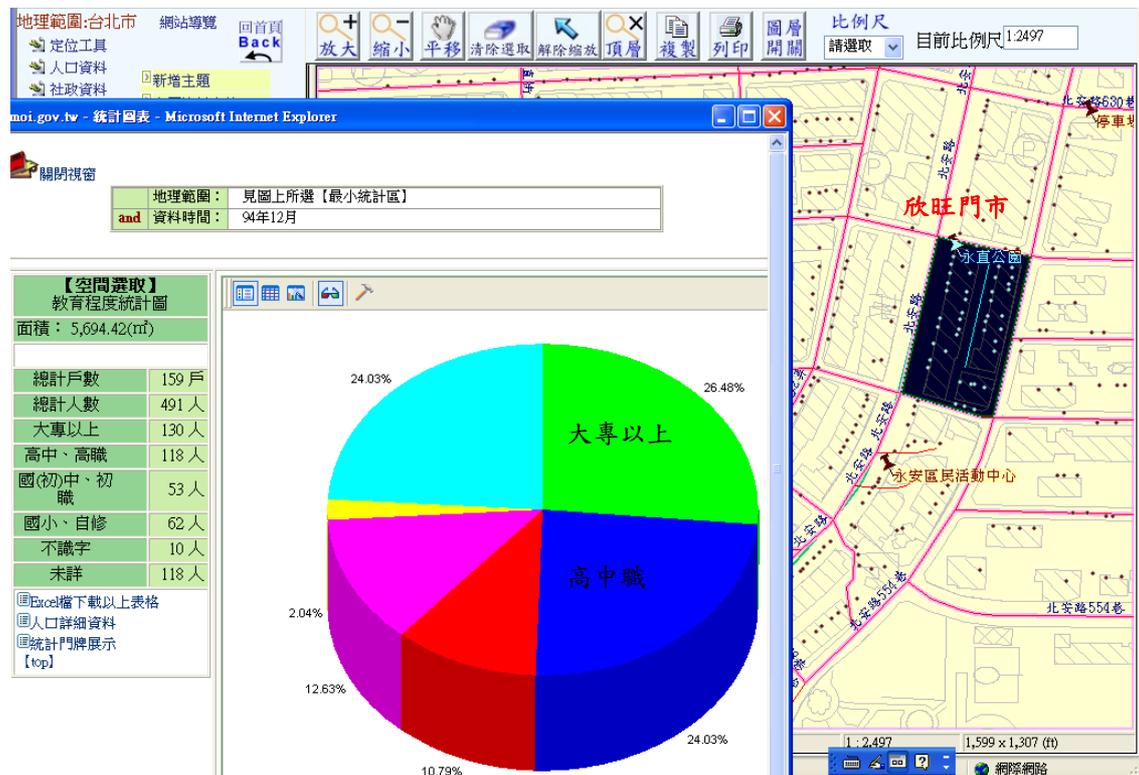


Figure 14: The educational levels of the population in the basic statistical unit covering the Shin-wan outlet of 7-11



Conclusion

The technologies and applications of GIS are interdisciplinary and can be extended to government agencies, academic institutions, private corporations and the public. In order to provide the government agency more comprehensive information to support decision-making, the GIS of Statistics creates basic statistic units based on address datasets, displays more detailed spatial distribution of population and business data, as well as improves the precision of the administrative decision reference. It includes six kinds of different population data which covers various sub-categories of attributes, and then users can display information with multi-dimensional query.

The system provides different interfaces for different type of users, administrative users of government agencies and general users of the public to take both confidentiality and accessibility into consideration. Additionally, it is convenient and costless because users can browse and display data via the internet browser without any extra costs of GIS software.

Accordingly, the maintenance of the digital maps and attribute data is crucial to the benefit of the system. For the sustainable development of the system and expansion of the application efficiency, the authorities should continue the tasks of data update. The foundation of public affairs data should be build based on a complete structure, as well as by establishing digital address datasets which is compatible with the Household Registration and Conscription Information System.

To sum up, applying the GIS of Statistics is helpful to support various administrative and business decision-making, such as defining the market areas, formulating marketing strategies, planning regional development scheme and so on. In the future, this project will coordinate with the National Geographic Information System (NGIS) Plan to expand the system to a national scale and to incorporate with data of medical services, urban planning and elementary educational services. It is predictable that the GIS of Statistics will play an increasing important role in various applications of decision-making support.

Reference

1. Chou Tien-yin, 2002, The theories and practices of Geographic Information System, Scholars Books Co., Ltd.
2. Department of Statistics, Ministry of the Interior, 2005, The proposal of the GIS of Statistics development project.
3. The official website of Carrefour (<http://www.carrefour.com.tw>)
4. The official website of 7-11 (<http://www.7-11.com.tw>)