

| | | | |
|---|------|--|--------------|
| Project to Establish Standard Analysis Models for Public Big Data | 2016 | ▶ Establishing 6 standard analysis models, such as civil complaint analyses and analyses of CCTV blind spots | 11.8 |
| | 2017 | ▶ Establishing 10 standard analysis models, such as job matching model analyses | 13.7 |
| | 2018 | ▶ Establishing 10 standard analysis models, such as analyses to decide the location for electric vehicle charging stations (ongoing) | 18.0 |
| Total | | | 107.7 |

Source: Materials submitted by the MOIS

The ‘New Project on Analysis of Public Big Data’ (hereinafter referred to as the ‘New Analysis Project’) is for developing new big data analysis models in order to resolve pending issues of organizations (refer to the following case study): There is a review and confirmation process for data types and methods for analysis, followed by a verification process to review the effectiveness of the analysis results. The New Analysis Project has the characteristics like that of pilot projects. On the other hand, the Project to Establish Standard Analysis Models for Public Big Data (hereinafter referred to as ‘Project to Establish Standard Analysis Models’) selects models among ones that have been proven effective through the New Analysis Project, which can be commonly utilized by numerous organizations, including local governments. These models are spread and distributed accordingly.

[Case study]

- ▶ **(Analysis of civil complaints)** Comprehensively analyzing data related to Pohang City in Gyeongsangnam-do (province) regarding the civil complaint bulletin board, news article about Pohang City, and the meeting minutes of Pohang City Council. By visualizing top keywords by channel, area, and month, the analysis can be utilized to review civil complaints that are not favorable for the city and decide on the order of priority to proceed with civil complaints.
- ▶ **(Analysis of transportation)** Comprehensively analyzing data related to Jeonju City in Jeollabuk-do (province) regarding the usage logs of public transportation cards, bus lane information, and operation information of buses with travel routes, location information of bus stops, and floating population information provided by public telecommunications service providers. The analysis can be utilized to adjust bus operation intervals and develop blind spots for public transportation (new routes).
- ▶ **(Decision of locations for electric vehicle charging stations)** Comprehensively analyzing data related to Daegu Metropolitan City regarding location of mass gathering facilities, addresses of e-vehicle owners, locations of the existing charging stations, and traffic flow and floating population information. The analysis can be utilized to decide on the locations of new electric vehicle charging stations.

Also, the MOIS has been utilizing (analyzing) the Standard Analysis Models onto the Combined Evaluation of Local Government (CELG) in order to boost public big data analysis and its use since 2017. Accordingly, all local governments are recommended to utilize the Standard Analysis Models for policy establishment and decision-making in administration.

2. Insufficient Measures to Utilize Private Data for Standard Analysis Models

Through the Project to Establish Standard Analysis Models for Public Big Data in 2016 and 2017, the MOIS developed a total of 16 Standard Analysis Models (refer to Attached Table: “Establishment History of Standard Analysis Models in 2016 and 2017”). Among these models, 8 models (50%) including the analysis of CCTV blind spots, tourism, and public transportation are using private data, such as floating population data¹ (by telecommunication companies), credit card sales data (by credit card companies), and credit information (by credit information companies) for their big data analysis (Refer to Table 2).

[Table 2] History of Standard Analysis Models that require utilization of private data

| Establishment year | Model name | Description | Name of private data required |
|---------------------------|--|---|--|
| 2016 | Analysis of CCTV blind spots | ▶ Analyzing CCTV blind spots and prioritizing orders to install new cameras | Floating population data |
| | Analysis of tourism | ▶ Analyzing effects of local festivals (tourism and sales increases) | Floating population data, credit card sales data |
| | Analysis of public transportation | ▶ Analyzing public transportation blind spots and efficiency of flexible car allocation | Floating population data |
| | Analysis on efficiency of labor monitoring | ▶ Analyzing the order of priority at workplaces under labor monitoring | Corporate credit rating data |

¹The data is a statistical estimation of the floating population by the grid (50cmx50cm) based on cellphone locations.

| | | | |
|------|--|---|--|
| 2017 | Analysis of taxpayers with overdue local tax | ▶ Analyzing the order of priority for efficient imposition of overdue local taxes | Credit information of taxpayers with overdue taxes |
| | Analysis of job matching | ▶ Analyzing and developing customized jobs for jobseekers and job recommendations | Corporate credit rating data |
| | Analysis of golden hour for ER patients ① | ▶ Analyzing blind spots of 119 emergency centers | Transportation volume data |
| | Analysis of golden hour for ER patients ② | ▶ Analyzing the optimal measures to allocate and operate ambulances | Transportation volume data |

Source: Rearranged materials submitted by the MOIS

The 2nd Basic Plan to Provide and Promote the Use of Public Data (December 2016) and the Implementation Plan to Provide and Promote the Use of Public Data (2017) (hereinafter referred to as the ‘Implementation Plan of Public Data’) stipulates that it is the responsibility of the MOIS to secure government-wide joint measures to use (e.g. en bloc agreement, joint purchase) private data that are commonly required for administrative services for the general public, as well as to resolve pending issues of various government offices in finance, telecommunications, and sales information.

The final report of the Project to Establish Standard Analysis Models in 2016 and 2017 recommends that measures are needed to conveniently secure private data at lower prices, as private data is utilized to establish the Standard Analysis Models for various topics such as CCTV blind spots, tourism, and public transportation.

Also, numerous reports related to the status of big data implementation by local governments, including ‘Boosting Measures to Establish Big Data-based National Strategies,’² ‘Research on Effective Utilization Measures of Private Data for Big Data Analysis,’³ and ‘Survey on the Progress of Big Data Work Implementation of Local

²December 2017, Korea Institute of Public Administration (KIPA)

³November 2017, Korea Local Information Research & Development Institute (KLID)

Governments⁴ suggest that the level of securing and utilizing private data in the public sector is very low. To improve the situation, a rational distribution system needs to be established. To effectively and efficiently use private data, certain measures including joint purchases are needed.

Therefore, the MOIS needs to secure measures (such as individual or joint purchases of private data) so that local governments can establish their policies and make administrative decisions by fully utilizing the Standard Analysis Models.

But, the MOIS did not reflect the joint use measures for private data (including en bloc agreements and joint purchases) into its implementation plan of public data in 2018 or its budget plan (draft) of 2019. As of November 2018, during the audit period, the ministry had not prepared any necessary measures.

Accordingly, the Board of Audit and Inspection of Korea (hereinafter referred to as 'BAI') conducted a full inspection between October 29 and November 30, 2018, on 327⁵ cases of local government data analyses (until October 2018) by using the Standard Analysis Models of the MOIS to see whether such analyses used private data. As can be seen in Table 3 , among 100 cases that required using private data for analysis (such as CCTV blind spots and effects of tourism), it was found that only 27 cases used private data (by individually purchasing the data), while 73 cases did not.

⁴Conducting the survey among a total of 243 local governments including 17 metropolitan and provincial governments and 226 city and county governments regarding big data work (October 2017, Korea Local Information Research & Development)

⁵Based on the analysis report registered on the common ground system of big data of the MOIS between December 30, 2016, and October 31, 2018

[Table 3] Analysis cases of the Standard Analysis Models of local governments

(Unit: number of local governments)

| Model name | Name of private data | No. of local governments that utilize the data | No. of local governments that do not utilize the data |
|---|--|---|--|
| Analysis of CCTV blind spots | Floating population data | 14 | 61 |
| Analysis of tourism (Analyzing effects of local festivals) | Floating population data Credit card sales data | 10 | 5 |
| Analysis of public transportation | Floating population data | 3 | 7 |
| Total | | 27 | 73 |

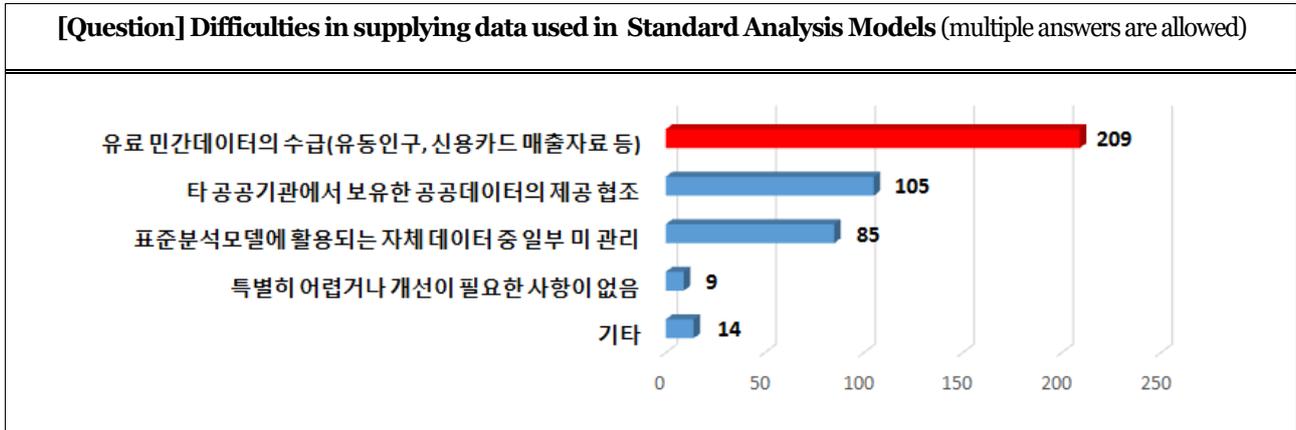
Source: Rearranged materials submitted by the MOIS

To analyze causes and find issues to be improved, the BAI carried out a survey among 245⁶ local governments nationwide regarding the difficulties of data supply when using the Standard Analysis Models. The results are presented in Graph 1: 85% (209 local governments) experienced difficulties in ‘supplying paid private data’ (such as floating population data and credit card sales data) when utilizing the Standard Analysis Models. Graph 2 indicates that 97% (237 local governments) want the MOIS to purchase private data in the form of a lump sum or joint purchase and provide such service to local governments, as local governments have difficulties making individual contacts for data purchases and lack budget and expertise.

⁶According to the current Local Autonomy Act and the Special Act on the Establishment of Jeju Special Self-Governing Province and the Development of Free International City, there are a total of 243 local governments, including 17 metropolitan and provincial governments and 226 city and county governments while not including Jeju City and Seoguipo City (they are designated as administrative cities rather than local governments). They are included in this survey, however, as the cities are carrying out big data work. Thus, a total of 245 organizations are the survey respondents.

[Graph 1] Survey results of difficulties in supplying data for use in Standard Analysis Models (1)

(Unit: No. of cases)



Source: status survey by the BAI

Key for Graph 1:

유료 민간데이터의 수급 (유동인구, 신용카드 매출자료 등): Supply of paid private data (e.g. floating population data, credit card sales data)

타 공공기관에서 보유한 공공데이터의 제공협조: Support to provide public data owned by other public organizations

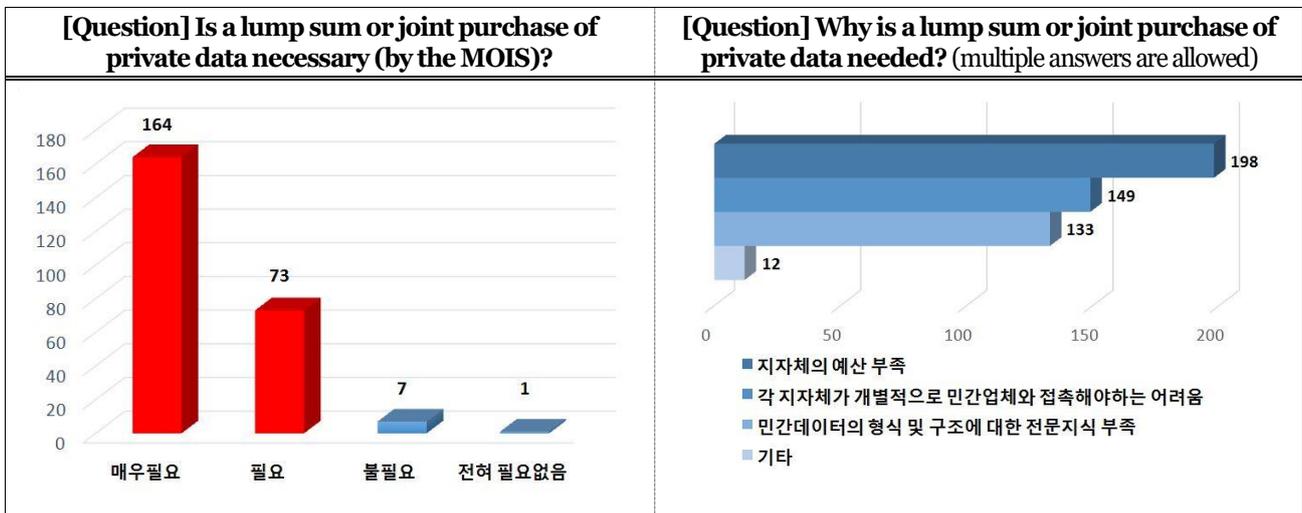
표준분석모델에 활용되는 자체 데이터 중 일부 미 관리: No management for some of the local government's own data, used for the Standard Analysis Models

특별히 어렵거나 개선이 필요한 사항이 없음: No specific issues or items to be improved

기타: Others

[Graph 2] Survey results of difficulties in supplying data for use in Standard Analysis Models (2)

(Unit: No. of cases)



Source: status survey by the BAI

Key for Graph 2:

매우 필요: Extremely necessary

필요: Necessary

불필요: Unnecessary

전혀 필요없음: Not necessary at all

지자체의 예산 부족: Lack of local government budget

각 지자체가 개별적으로 민간업체와 접촉해야하는 어려움: Difficulties of local governments not being able to individually contact private entities

The BAI also requested Corporation A (hereinafter referred to as ‘A’) and Corporation B (hereinafter referred to as ‘B’) to provide their opinions regarding the price policy of the floating population and data measures of lump sum or joint purchases of data. These are the companies that provided and sold the most floating population data as private data during the audit period, which is used the most for the Standard Analysis Models⁷. The results can be seen in Table 4: Both A and B agreed that a lump sum or nationwide joint purchase can lower the unit price with the advantage of unifying purchase channels compared to individual purchases made by each local government. This confirms that a lump sum or joint purchase can be an opportunity to provide the floating population data to local governments.

[Table 4] Opinions of telecommunications service provider on lump sum or joint purchase of floating population data

| Organization | Detailed opinions |
|--------------|--|
| A | ▶ Reducing unnecessary negotiation processes and costs between data providers and local governments by providing data through one unified channel in bulk. Able to provide data at lower costs compared to existing price policy |
| B | ▶ Able to provide data at lower costs compared to individual purchases made by each local government (if the sales channel can be unified and data providers deal with only one partner (MOIS) in the form of a lump sum (joint) purchase) |

Source: Rearranged materials submitted by Organization A and B

The MOIS expects to see a cost reduction as well as a boosting effect for analyzing and utilizing public data with the support of a smooth supply of private data compared to

⁷Among 16 Standard Analysis Models established by 2017, 8 models require analysis. 3 out of 8 used the floating population data.

individual purchases if the ministry implements a lump sum or joint purchase of private data to use in Standard Analysis Models.

3. Necessity of System Development Measures for Major Standard Analysis Models

The MOIS is currently distributing the ‘Guidelines for the Standard Analysis Models’ by model to help local governments use data and perform their own analysis. The guidelines are based on 16 analysis models that were developed through the Project to Establish Standard Analysis Models in 2016 and 2017.

According to the guidelines, the personnel of each local government are expected to proceed with the overall big data analysis process by installing programs such as the spatial information analysis program, database management program, and big data analysis program on their PCs, and editing and processing the collected data by using Program A. This work by using the PC, however, is manual.

This manual style of work can prove problematic as the guidelines can vary from 111 to 294 pages (maximum) when dealing with major⁸ Standard Analysis Models, which many local governments utilize, including in the analysis of civil complaints, CCTV blind spots, and tourism. Also, it is not easy for some civil servants, even in computing departments, to learn how to use these required programs, including spatial information and big data analysis programs.

⁸Top 3 models (until October 31, 2018) in the order of how much the models (registered onto the common ground system of big data) are utilized for analysis cases of local governments

The 1st advisory committee on public big data⁹ in 2016 (February 4, 2016) expressed its opinion that there should be a data standard system equipped with a common ground system¹⁰ to promote the Standard Analysis Models of public big data. The final report of the Project to Establish Standard Analysis Models in 2016 and 2017 also recommends that there should be measures to develop the Standard Analysis Models into a system and to secure data supply for stable analysis.

Therefore, the MOIS needs to develop major Standard Analysis Models (that are widely used for the common ground system for big data) into an automatic system, while securing measures to provide bulk data used for Standard Analysis Models after processing it in the designated form.

However, the MOIS recommends that local governments place individual contracts for individual big data analysis based on the ‘fee estimation guidelines for public big data analysis projects,’ as it is difficult for local governments to analyze and process the data on their own. This raises concerns for waste in budget and administrative power of each local government by placing same or similar contracts for projects. As of November 2018, there have not been measures to process and provide bulk data before utilizing it to develop the Standard Analysis Models into a system.

Taking this phenomenon into account, the BAI conducted another survey to review the utilization rate of the Standard Analysis Models (MOIS): The survey consists of 327 analysis cases (up to October 2018) of local governments using 6

⁹Advisory committee in big data under the Open Data Strategy Council, an intergovernmental control tower regarding public data

¹⁰Established and operated as intergovernmental common ground system (named ‘*Hyeon*’ meaning ‘insight’ in Korean) for big data analysis by the National Information Resources Service of the MOIS since 2012 to provide general big data analysis functions such as keyword analysis

Standard Analysis Models¹¹ established in 2016. The results are shown in Table 5: 4 models, except for civil complaint analysis and CCTV blind spots analysis, were hardly ever used (less than 10% utilization rate).

[Table 5] List of local governments' use of Standard Analysis Models (established in 2016)

(Unit: No. of local governments, %)

| Model name | No. of local governments that utilize the models (Utilization rate ^{Note}) | Model name | No. of local governments that utilize the models (Utilization rate) |
|-----------------------------------|--|--|--|
| Analysis of civil complaints | 171 (70) | Analysis of CCTV blind spots | 75 (31) |
| Analysis of tourism | 15 (6) | Analysis of maintenance fees for housing complexes | 14 (5) |
| Analysis of public transportation | 13 (5) | Analysis of Efficiency of labor monitoring | - (0) |

Note: The utilization rate is calculated by the number of local governments that are registered as analysis cases on the common ground system for big data of the MOIS compared to the total number of local governments (243).

Source: Materials submitted by the MOIS

To do cause analysis, there was a survey among 245 local governments during the audit period: 72% (176 local governments) responded that it is hard for them to conduct an analysis without any help, due to the difficulty of the Guidelines for the Standard Analysis Models. Of the 245 local governments, 44% (108 local governments) expressed that it is difficult to utilize open public data (owned by other organizations) as is and stated that it was inconvenient to reprocess the data on their own.

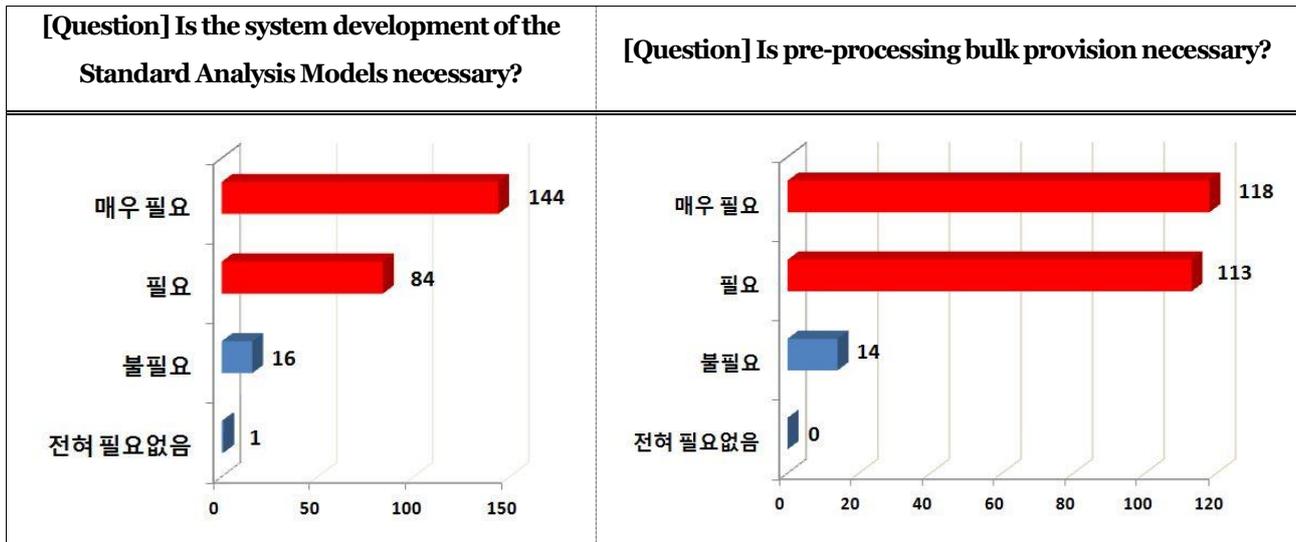
The necessity of system development for Standard Analysis Models and the necessity of pre-processing and bulk provision to use the data is shown in Graph 3: 93% (228 local governments) and 94% (231 local governments) think it is 'extremely necessary' or 'necessary'. This suggests that local governments want measures to

¹¹Excluding 10 Standard Analysis Models established in 2017 as it has been less than 1 year since the establishment

automate analysis and data processing of the Standard Analysis Models due to numerous difficulties, such as manual analysis.

[Table 3] Survey results regarding system development of the Standard Analysis Models and data collection & processing

(Unit: No. of local governments)



Source: status survey by the BAI

Key for Graph 3:

매우 필요: Extremely necessary

필요: Necessary

불필요: Unnecessary

전혀 필요없음: Not necessary at all

The survey results suggest that the analysis, data editing and processing of the Standard Analysis Models are being manually performed. It is too complicated for local governments to conduct their own analysis as the relevant guidelines are too difficult, and as such, the trend of insufficiently using the Standard Analysis Models will likely continue.

4. Insufficient Management of Performance Such as the Assessment on the Utilization of Public Big Data Analysis Models

According to the ‘Promotion Plan for Boosting the Utilization of Public Big Data’ (February 2016), it was the MOIS’ responsibility to secure an assessment system by the 1st half of 2016 and to carry out the performance management to boost the usage of public big data.

Then, the MOIS implemented the Project to Establish Standard Analysis Models in 2016 of which the results can be found in Table 6: Performance goals and index of the Standard Analysis Models were secured to see whether the analysis results of such models are, in fact, being used.

[Table 6] List of performance goals and index of Standard Analysis Models

(Unit: %)

| No. | Model name | Performance index | Performance goals | | |
|-----|--|--|---|------|------|
| | | | 2017 | 2018 | 2019 |
| 1 | Analysis of civil complaints | ▪ Designating personnel to civil complaints; rate of reducing delays in processing civil complaints | Based on the rate of reducing delays in processing civil complaints | | |
| 2 | Analysis of public transportation | ▪ Reviewing whether to use analysis results (e.g. blind spots for public transportation) to reflect onto public transportation policies | Use | Use | Use |
| 3 | Analysis of tourism | ▪ Rate of increase in domestic and international tourists compared to that of 2016 | 1.5 | 3.0 | 4.5 |
| 4 | Analysis of CCTV blind spots | ▪ Rate of applying analysis results to actual installation of CCTVs | 10 | 20 | 30 |
| 5 | Analysis of maintenance fees for housing complexes | ▪ Rate of improvement in finding apartment complexes suspected of inappropriately executing maintenance fees ▪ Rate of improvement in finding companies suspected of inappropriate construction bidding | 5 | 10 | 15 |
| 6 | Analysis of labor monitoring efficiency | ▪ Utilization rate of analysis results to the actual designation of workplaces under labor monitoring | 10 | 20 | 30 |

Source: Materials submitted by the MOIS

The MOIS reported plans to implement of performance assessments on the utilization of public big data analysis projects at the 3rd Open Data Strategy Council in 2016 and at the 4th one in 2017. The performance assessment goes in the order of ① preparation of performance assessment, ② self-assessment of organizations (e.g.

local governments), ③ comprehensive assessment by the organization in charge, (MOIS) and ④ assessment implementation and post-management.

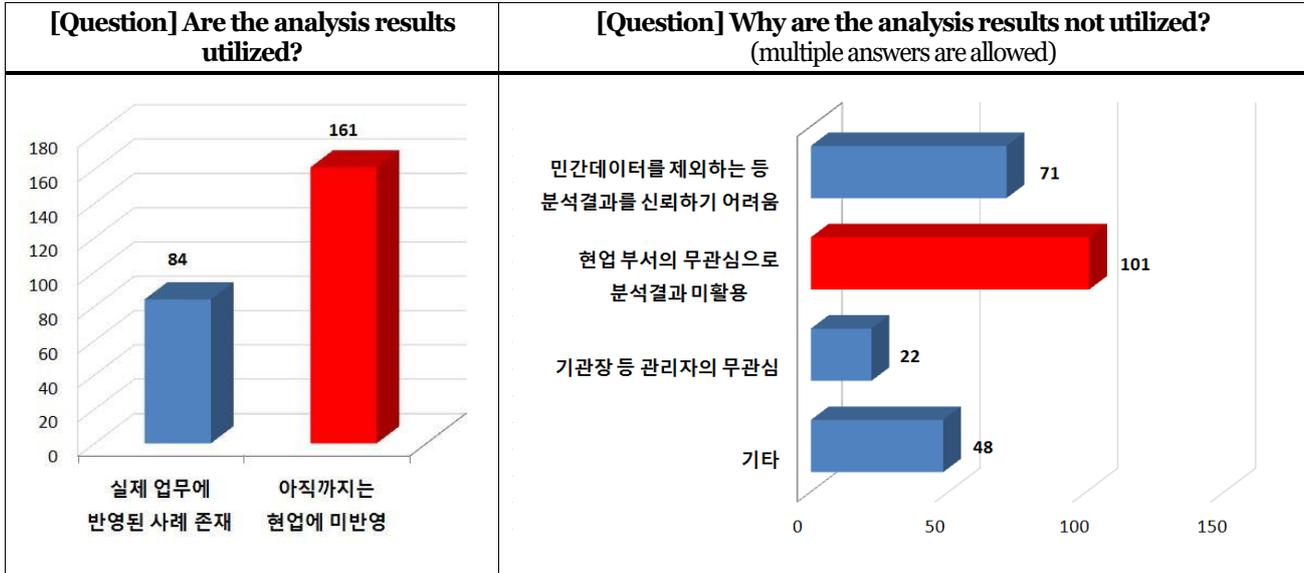
Accordingly, the MOIS collected the results of the self-assessments by organizations (such as local governments), on their utilization of the Standard Analysis Models. While carrying out the comprehensive assessment, the MOIS needed to identify low utilization cases and suggest improvements for continuous performance management.

The MOIS collected the results of self-assessments from organizations in 2016, but did not carry out the comprehensive assessment, assessment implementation, and post-management. In 2017 and 2018, there was no performance assessment at all.

During the audit period, the BAI researched to see whether 245 local governments nationwide adopted the analysis results of utilizing the Standard Analysis Models. The results, as indicated in Graph 4, show that 67% (161 local governments) have not yet reflected the analysis results to their work. The biggest reason (41%; 101 governments) the working level departments neglected the analysis results in utilizing them for work was due to indifference, and the actual rate of using the analysis results tends to be fairly low.

[Graph 4] Status survey results regarding the utilization of Standard Analysis Models

(Unit: No. of local governments)



Source: status survey by the BAI

Key for Graph 4:

실제 업무에 반영된 사례 존재: Cases of actual utilization

아직까지는 현업에 미반영: Not yet using analysis

민간데이터를 제외하는 등 분석결과를 신뢰하기 어려움: Hard to trust the analysis results (e.g. excluding private data)

현업 부서의 무관심으로 분석결과 미활용: Not using analysis results due to indifference

기관장 등 관리자의 무관심: Indifference of managers (e.g. president of organizations)

기타: Others

The survey results from items 1-3 indicate concern that if the utilization rate remains low, there might be potential failure in vitalizing the data-based administration, which would have been used to establish objective and scientific policies by applying the big data analysis results.

Opinions from Relevant Organizations and Review Results

The MOIS proposed the idea of a collaboration between the Ministry of Economy and Finance and local governments regarding lump sum or joint purchases of private data to secure fee estimation standards of private data and relevant budget measures through demand surveys and research of public organizations.

Regarding the system development of major Standard Analysis Models, the ministry suggested securing measures of pre-processing and bulk provision of public data that will be jointly used by comprehensively reviewing future circumstantial changes. Additionally, by developing a pilot system for 1-2 Standard Analysis Models used most frequently by local governments, any further developments needed could be identified. However, there are not enough legal grounds to carry out the continued monitoring of continuous performance assessments on public big data analyses. Further, there may be complaints regarding assessing the local governments by rank, and as such, the MOIS responded that it would prepare for relevant guidelines so that organizations using such guidelines can autonomously manage their own performance.

However, it is hard to accept the ministry's response if the following 2 reasons are taken into consideration: First, the Project to Establish Standard Analysis Models is an item for performance management according to Article 68 of the Electronic Government Act and Article 84 of the Enforcement Decree of the Act. Second, the performance of utilizing Standard Analysis Models is already reflected onto the joint assessment index of local governments.

Recommendation to the MOIS

The Board of Audit and Inspection of Korea would like to present the **Minister of the Interior and Safety** the following recommendation: The MOIS needs to secure measures to jointly use data, including bulk or joint purchases of private data that can be utilized for the Standard Analysis Models so as to encourage the use of Standard Analysis Models of public big data for actual administrative work. Also, it is necessary to develop

the system for the Standard Analysis Models that have higher utilization rates by local governments (e.g. analysis of CCTV blind spots) and secure measures to process and provide such data in bulk. The BAI also recommends the MOIS to prepare measures to systematically manage performance to continuously improve and supplement the weak points of the Standards Analysis Models. **(Notice)**

[Attached Table]

Establishment History of Standard Analysis Models (2016 and 2017)

| No. | Establishment year | Model name | Description |
|-----|--------------------|--|--|
| 1 | 2016 | Analysis of civil complaints | ▪ Analyzing complaints; steps of complaint proceedings with civil complaints |
| 2 | | Analysis of public transportation | ▪ Analyzing public transportation blind spots and efficiency of flexible car allocations |
| 3 | | Analysis of tourism | ▪ Analyzing effects of local festivals |
| 4 | | Analysis of CCTV blind spots | ▪ Analyzing CCTV blind spots and order of priority in installing new cameras |
| 5 | | Analysis of maintenance fees for housing complexes | ▪ Analyzing maintenance fees for housing complexes, bidding and unlawful relations between contractors and sub-contractors |
| 6 | | Analysis of labor monitoring efficiency | ▪ Analyzing order of priority of workplaces under labor monitoring |
| 7 | 2017 | Analysis of taxpayers with overdue local taxes | ▪ Calculating and analyzing grade possible to pay for overdue taxes |
| 8 | | Analysis of donation resources ① | ▪ Developing optimal model to allocate donation resources (physical resources) |
| 9 | | Analysis of donation resources ② | ▪ Developing optimal model to match with volunteer work (human resources) |
| 10 | | Analysis of water supply leakages | ▪ Analyzing water supply leakages and predicting leakage risk levels |
| 11 | | Analysis of golden hour for ER patients ① | ▪ Analyzing blind spots of 119 emergency centers |
| 12 | | Analysis of golden hour for ER patients ② | ▪ Analyzing optimal measures to allocate and operate ambulances |
| 13 | | Analysis of job matching | ▪ Analyzing and developing model providing customized jobs and job recommendations for jobseekers |
| 14 | | Analysis of road management ① | ▪ Analyzing risk areas with GIS-based potholes |
| 15 | | Analysis of road management ② | ▪ Predicting road damage risk levels and finding priority sections for re-pavement |