

### **III. The Reduction of Duplication of Statistical Survey Objects with the Statistical Frame of Establishments and Enterprises**

**Bunzo HIRAI**

Statistical Standards Department

Statistics Bureau

Ministry of Public Management, Home Affairs,

Posts and Telecommunications

JAPAN

In spite of measures to reduce reporting burden of respondents of statistical surveys, the continuation of the decreasing tendency of cooperation of respondents is one of the greatest problems ministries of the Government of Japan face as survey conductors. In order to tackle with this problem the Government of Japan tries to introduce the system in which duplication of statistical survey objects be reduced by registering the data of the population from which the sample is selected for surveys, managing the data unitarily, checking the survey history records of the sampled objects in the database of enterprises and establishments.

This article shows the framework of the system and the accomplishments to date.

#### 1. Historical Background

##### (1) Development of the Statistical Frame for Enterprise and Establishments

The idea to use the database of population information for the reduction of duplication of survey objects goes back to the New Strategies for Government Statistical Services for the Coming Decade<sup>1</sup>. In this recommendation, compilation of a Statistical Frame for Enterprise and Establishments was proposed. The Frame was to be basically composed of information of the list basically made by the Enterprise and Establishment Census and to be converted to the database. It was to be maintained and renewed with the results of censuses such as Commerce Census and Manufacturing Census. The main object was to provide population information for statistical surveys conducted by ministries, but the reduction of reporting burden was also aimed by avoiding the convergence of the same entity to be the objects of a number of surveys through the strict thorough management of samples with the common database of population information.

Responding to this recommendation the Management and Coordination Agency (MCA)

---

<sup>1</sup> The New Strategies for Government Statistical Services for the Coming Decade is the comprehensive policy vision on official statistics in the Government of Japan submitted by the Statistical Council in March 1995. The Statistical Council is an advisory body of the Statistical Standards Department, Statistical Bureau, Ministry of Public Management, Home Affairs, Posts and Telecommunications (MPHPT) (Management and Coordination Agency up to January 2001), composed of statistical experts.

transformed the existing Establishment Census to the Establishment and Enterprise Census and the first Census was conducted in 1996<sup>2</sup>. Then MCA developed the Establishment and Enterprise Information Database System (EEDIS) in 1998, but the EEDIS has stored only the information from the results of the Establishment and Enterprise Census and has had no historical information as to whether a particular establishment or enterprise had been selected as an object of past surveys, thus it cannot be used as a system for avoiding the convergence of the same entity to be the objects of different surveys, but be used only for providing the list of all the establishments and enterprises as the basic sample frame for surveys.

(2) The Measures to Reduce Reporting Burden of Statistical Surveys in the Administrative Reform Related to the Central Government Reorganization

In December 1997, Administrative Reform Council, which was chaired by the then Prime Minister Ryutaro Hashimoto and deliberated the government policy to reorganize the central ministries and agencies, recommended in its final report that the statistical organization should still be decentralized and that the organization which would be in charge of coordination of government statistical activities (to be organized under the proposed Ministry of Public Management, Home Affairs, Posts and Telecommunications (MPHPT)) perform leadership to correct duplication of statistical administration. This policy was legally authorized by the enactment of the Basic Law to Reorganize Central Government in 1998.

Based on that policy the Government considered the measures to correct duplications of the objects of statistical surveys as one of the sub-policies of the promotion of streamline and efficiency of administrative organization. As a result, in April 1999, the Cabinet decided "the Basic Plan concerning the Promotion of Streamline and Efficiency of Administrative Organization". In the Plan, MPHPT was to establish a statistical frame of establishments and enterprises, including the historical information of the number of times that each establishment or enterprise has been surveyed, using the results of statistical surveys and the administrative records available of each office or ministry<sup>3</sup>, and was to maintain unitary population information. Each office or ministry was to take measures to correct the sample using the frame when sampling them in order to avoid such situation as a survey object is surveyed more frequently than the frequency fixed in advance in a certain period.

In order to make up concrete measures to establish a statistical frame including survey

---

<sup>2</sup> MCA was merged into the Ministry of Public Management, Home Affairs, Posts and Telecommunications (MPHPT) with the Ministry of Home Affairs and the Ministry of Posts and Telecommunications in January 2001, at the time of drastic central government reorganization.

<sup>3</sup> Before the central government reorganization in 2001 we used "ministries and agencies (each ministry or agency)" when we mention central government organization in general, because there were a number of agencies which have the same status as ministries. But at the reorganization most of such agencies were consolidated to ministries, and the Cabinet Office was newly established. Thus now we use "office and ministries (each office or ministry)" for that purpose.

history data and to reduce duplication with the frame, office and ministries organize an inter-ministerial association. The association reached agreement on the details of the development of the new system in December 2000. MPHPT started to develop another new database in order for each office or ministry to check too much concentration of surveys to the same objects on the database. Finally in June 2002 the government started the operation of the new Statistical Frame of Establishments and Enterprises (SFE), and in July it also started activity for the reduction of duplication of the objects of statistical surveys.

## 2. The Main Features of the new SFE

### (1) Roles of the SFE

The roles of the new database can be summarised in the following three points.

a) Providing the latest list of information related to all establishments and enterprises available as the basic frame for statistical surveys on establishments and enterprises

b) Registering and providing the historical information of survey objects (establishments and enterprises) on all censuses and surveys, i.e. the information whether a specific survey object was previously requested to report for such censuses and surveys, as information to avoid any unnecessary or undesirable concentration of surveys to the same survey objects in subsequent surveys

### (2) Outline of the SFE

#### a) System composition

The SFE was developed as a client/server system using the ORACLE DBMS. By connecting the SFE server with the PCs located at government ministries (the clients), clients can perform on-line-based work to avoid duplications. The SFE is composed of the population information table group, the historical information table group and the control information table group. Tables are linked through the uniform establishment codes so that the information among the different tables can be interconnected. The following figure shows the conceptual framework of the SFE.

#### b) Contents of the database

##### [Population information table group]

Population information is compiled from the results of the Establishment and Enterprise Census conducted by MPHPT, the Commerce Census and the Manufacturing Census by the Ministry of Economy, Trade and Industry, and the Financial Statements of Corporations by Industry by the Ministry of Finance, as the basic (updating) sources. In addition to these sources, data prepared by a private company (Teikoku Databank) are used to complement the information on enterprises.

About 6.6 million private establishments have been recorded in the SFE, while public establishments have not yet been recorded. The number of enterprises (companies) is approximately 1.7 million.

The items recorded in the SFE are as follows:

- Basic items ... area, name, address and telephone number of the establishment or enterprise
- Attributive items ... number of persons engaged, kind of business activities, legal organizational status, etc.
- Items on enterprises ... capital, number of branch offices, status of the parent company, subsidiary companies, and associate companies, etc.

##### [Historical information table group]

The historical information, including the number of times that each establishment has been surveyed and the survey names since January 2001, is registered concerning all the censuses and surveys conducted by the government. The number of censuses and surveys from which the historical information will be collected and stored is around 230 as of June 2002, which have been conducted since January 2002, covering 13.0 million establishments in total, while the surveys for which duplication should be corrected number 140 covering 3.0 million establishments.

The items recorded in the SFE are as follows:

- Items to be recorded in the survey history (for establishments) ... flag indicating an object of a census or a sample survey, flag indicating a survey candidate, flag indicating an object being the object of the reduction of duplication, etc.
- Items on the survey attributes (for censuses or surveys) ... name of the office or ministry, name of the census or the survey, year and month census or survey was conducted, etc.

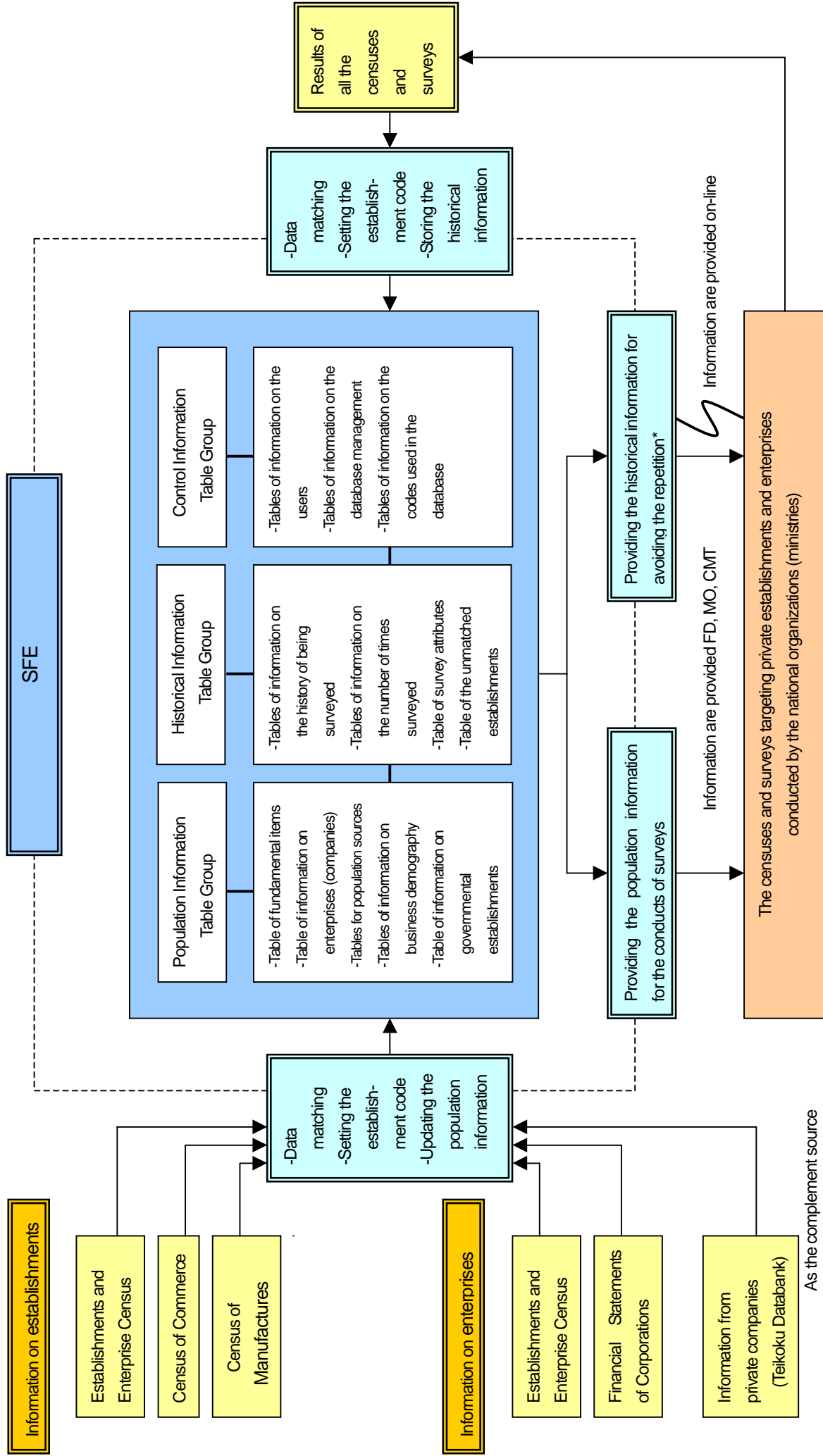
c) Users and uses

For the time being, only the office and ministries of the national government are allowed to use the SFE. The database provides the users with information that can be utilized in the following ways: to elucidate the objects of censuses or surveys, to check whether there is duplication of survey objects more frequently than the frequency set in advance, and to analyze the population structure of establishments and enterprises.

d) Access to the information

In principle, the SFE provides the office and ministries on-line with the information needed to avoid duplications, while also providing information for carrying out censuses and surveys on establishments and enterprises using off-line media such as FD, MO, and CMT, taking account of the large volume of the data needed.

# Conceptual Framework of the Statistical Frame of Establishments and Enterprises



\* Avoiding the repetition is to avoid the situation that the same establishment or enterprise is selected as samples doubly or more in a short period.

### 3. The System to Reduce Duplication of Statistical Survey Objects

#### (1) Purpose

The purpose of the system is to avoid excess concentration of the statistical surveys to the same survey objects in statistical surveys on establishments and enterprises.

#### (2) Content

In the system, the standard maximum frequency in which each establishment or enterprise is surveyed is set, and when the frequency in a certain period (one year) before the survey seems to exceed the maximum, the alternative object which has been surveyed less than the maximum are to be selected.

#### (3) Covered Statistical Survey and Covered Establishments and Enterprises

This system covers all statistical censuses and surveys of the central government the objects of which are private establishments and enterprises (including public corporations).

The system covers establishments and enterprises covered in the previous paragraph except:

- a) the objects of censuses
- b) the objects of surveys using cluster sampling in which complete enumeration is adapted for each cluster extracted
- c) the objects which are decided at the time of survey
- d) the objects of surveys using stratified sampling in which complete enumeration is adapted for one or some of strata.

#### (4) The Maximum Frequency

The maximum frequency is preliminary set based on the results of the storage of survey history information in the SFE. It is currently set at the frequency of the object at 1% below the top in relation to the number of registered survey records (5% for the branch establishments of an enterprise with less than 30 % employees). The maximum frequency for establishments which do not coincide with population information is set as the same as the maximum of all the categories of attributes of survey objects (i.e. 70).

Attributes of Survey Objects	Maximum Frequency
Branch Establishments of an Enterprise	
Less than 30 employees	16
30 or more and less than 100 employees	34
100 or more employees	60
Establishments not belonging to any Enterprises	
Less than 30 employees	16
30 or more and less than 100 employees	34
100 or more employees	40
Headquarters of an Enterprise or Establishments without branches	
Less than 100 million yen capital	30
100 million yen or more and less than 1 billion yen capital	50
1 billion yen or more capital	70
Establishments which Do Not Coincide with Population Information	70

(5) Reduction Process

a) Before conducting a statistical survey, each office or ministry submits a list of survey object candidates made on a magnetic media in a uniform format to MPHPT.

b) MPHPT registers the list on the SFE. Then MPHPT calculates the frequency at which each establishment or enterprise on the SFE has been surveyed based on the registered history information. MPHPT returns the list to the office or ministry after adding a flag which indicates that the object be corrected to the establishment or enterprise whose frequency of surveys exceeds the maximum frequency based on the calculation results.

c) Each office or ministry, when noticed the necessity of correction, selects establishments or enterprises which has been surveyed less than the maximum in place of those with correction flags.

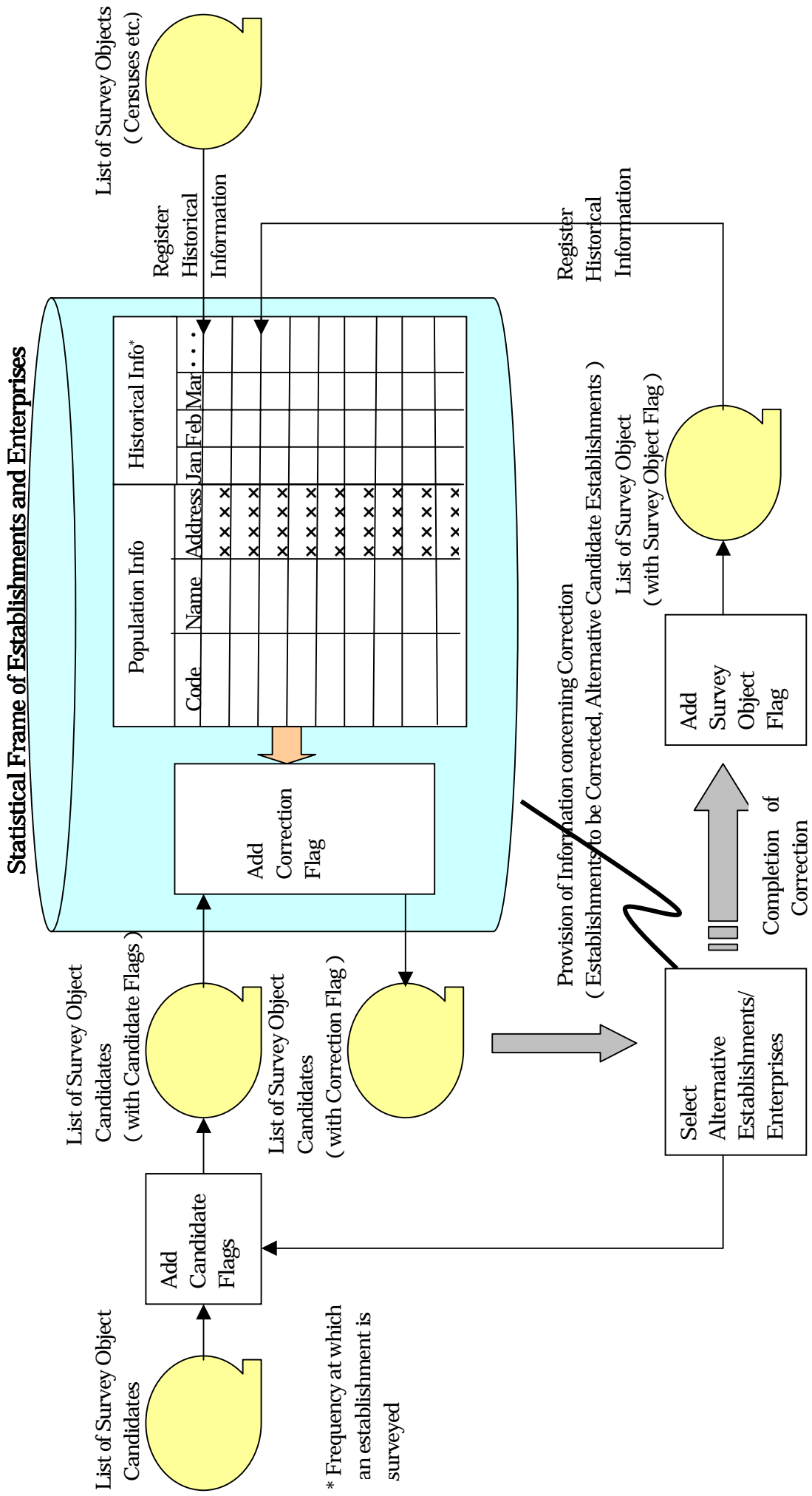
MPHPT provides information on whether a establishment or enterprise is the object of correction via Kasumigaseki WAN (wide area network of the central government offices).

d) After each office or ministry corrects duplication, it must report the results and submit the list of survey objects with survey object flag to MPHPT.

e) MPHPT registers historical information of each establishment or enterprise to the SFE based on the list of survey objects submitted by each office or ministry.



# Framework of the System to Reduce Duplication of Statistical Survey Objects



#### 4. The Results of Duplication Correction

From July, when the duplication correction measures through this new system started, to September 2002, 11 surveys used this system based on the reception register, and 3 were corrected based on the correction results reports. In these three surveys 16 objects had correction flags and 11 of them were actually corrected. 5 objects were not corrected because alternative objects cannot be selected in sampling.

**The Reduction of Duplication of Statistical Survey Objects**

As of September 1, 2002

	Total	2002		
		June	July	August
<b>Number of Statistical Surveys</b> (Based on Reception Register) Correction without Correction Flag (Reporting is not required) with Correction Flag (Reporting is required)	11 5 6	3 2 1	6 2 4	2 1 1
(Based on Correction Results Reports) Reception of Correction Results Reports	3	0	2	1
<b>Number of Establishments/Enterprises</b> (Based on Correction Results Reports) Number of Survey Objects without Correction Flag with Correction Flag	39,331 39,315 16	0 0 0	24,832 24,825 7	14,499 14,490 9
Corrected	11	0	2	9
Not Corrected (Reason)	5	0	5	0
① Population is too small.	0	0	0	0
② The said establishments are indispensable for accuracy.	5	0	5	0
③ The said establishments cannot be replaced to other geographical areas.	0	0	0	0
④ The said establishments cannot be replaced to the same classification of industry or size.	0	0	0	0
⑤ Other reasons.	0	0	0	0

**The Registration of Historical Information**

As of September 1, 2002

	Total	2002		
		June	July	August
(Based on Reception Register) Number of Registered Statistical Surveys Number of Registered Establishments/Enterprises	20 126,467	3 7,313	9 47,532	8 71,622

## 5. Future Issues

MPHPT, which is in charge of comprehensive coordination in the decentralized statistical system, promotes simple and efficient operation in the registration to the SFE in cooperation with other office and ministries,. Each office or ministry is expected to submit the list of survey object candidates with enough time in advance in order to allow MPHPT to smoothly process the measures needed for the correction of the duplication of survey objects.

The SFE is originally intended for the correction of the duplication of survey objects, not for the sampling population as the EEDIS, thus so far the development of the system for sampling population on the SFE is still on the way. MPHPT plans to develop the system, then the sampling and the correction would be connected, and the process would be simplified.