

# **DATA GENERATION GUIDELINES**

## *2009/2010 BLES Integrated Survey (BITS)*



*Republic of the Philippines*  
**DEPARTMENT OF LABOR AND EMPLOYMENT**  
**BUREAU OF LABOR AND EMPLOYMENT STATISTICS**  
Manila, Philippines

March 2011



*ISO 9001:2000 Certified*

**TABLE OF CONTENTS**

	<b>SECTION</b>	<b>PAGE</b>
	OBJECTIVE .....	1
A.	PREPARATORY ACTIVITIES.....	2
B.	RECONCILIATION OF STATUS CODES AND SELECTED DATA.....	2
C.	REQUIRED STATISTICAL TABLES.....	3
D.	FINAL ASSESSMENT OF ACCURACY IN DATA PROCESSING .....	5

## OBJECTIVE

These Guidelines are prepared to facilitate the generation of output tables needed in the preparation of the publication tables.

Preparatory activities should be done by concerned personnel before generating desired statistical tables. In addition, status codes need to be reconciled and necessary adjustments should be made before the generation of such tables.

From the 2009/2010 BLES Integrated Survey (BITS), statistics on the following variables will be generated:

- *Establishment Profile*
  - main economic activity;
  - major products/goods or services;
  - establishment characteristics as to ownership;
  - type of market for business process outsourcing;
  - unionism and membership, and existence and coverage of collective bargaining agreement/s; and
  - participation in global production network.
- *Employment*
  - total employment and its breakdown into working owners, unpaid workers and employees (managers/executives, supervisors/foremen, rank and file: regular and non-regular workers);
  - employment of specific group of workers;
  - jobs, work or service done within the premises of the establishment; and
  - jobs, work or service done outside the premises of the establishment.
- *Occupational Shortages and Surpluses*
  - job vacancies;
  - hard-to-fill and easy-to-fill occupations;
  - number of vacancies;
  - number of applicants;
  - length of recruitment period;
  - specialization/specific skills; and
  - reasons for hard-to-fill occupations.
- *Industrial Relations Practices*
  - Balancing Work and Family Responsibilities and Working Time Arrangements*
    - facilities available for child care, health and fitness, workplace accessibility and food service in the establishment;
    - programs implemented for child and family care, education, health and fitness, and financial and other support;
    - non-statutory leave arrangements;
    - flexible work arrangements;
    - workdays and work schedules;
    - usual operating hours of the establishment; and
    - arrangement for extra hours of work.
  - Occupational Safety and Health*

- preventive and control measures/activities against work safety and health hazards;
  - occupational safety and health policies/programs implemented in the establishment;
  - provisions/facilities/amenities available in the establishment;
  - OSH-related trainings/seminars provided for the employees;
  - agencies whose services/assistance were availed of in the conduct OSH trainings/seminars; and
  - health personnel who administers medical service to employees.
- *Occupational Injuries and Diseases*
    - incidence of occupational accidents;
    - cases of occupational injuries by incapacity for work (fatal, permanent, temporary);
    - cases of occupational injuries by classification (type of injury, part of body injured, cause of injury, agent of injury and major occupation group);
    - workdays lost of cases of occupational injuries by incapacity for work;
    - cases without lost workdays;
    - cases of occupational diseases;
    - incidence of commuting accidents including workers injured;
    - average employment; and
    - hours actually worked by all employed persons.

## **A. PREPARATORY ACTIVITIES**

1. All appropriate tabulations, cross-tabulations or disaggregation of the data categories should have been determined during the Pre-Field Operations stage of the survey, in order that,
  - a. tabulations of the same data variables/categories are maximized, hence processing time is minimized;
  - b. survey results that would be made accessible to the data users are comprehensive, thus, no need for further tabulations;
  - c. consistency checks are facilitated; and
  - d. reasonableness of the survey results at detailed levels are easily determined which would not be possible at aggregate levels. For one, errors in coding (industry) may be covered up when the data are tabulated at higher levels of disaggregation.
2. Based on the proposed data tabulations, the computer syntax should have been prepared prior to table generation.

## **B. RECONCILIATION OF STATUS CODES AND SELECTED DATA**

1. For DUP establishment, track its duplicate or the establishment to be retained.
  - a. The establishment with the lower EIN should be retained. If the establishment accomplished the questionnaire with the higher EIN, change it with the lower EIN; the DUP establishment should always be assigned with the higher EIN. Revise name, address, industry, geographic codes and status of the involved establishments accordingly in the status monitoring database.

Example:

Establishment Y with EIN 426 is DUP of Establishment X with EIN 678. Establishment X has responded.

The EIN of Establishment X should be replaced with 426 and its status code will be changed to RET or CET as applicable. The EIN of Establishment Y should be 678 and its status code becomes DUP of EIN 426.

- b. For establishments with more than one duplicate, the principle remains the same, the establishment to be retained should have the lower EIN and the duplicates should have the higher EIN. The status code of these establishments should be adjusted accordingly in the status monitoring database.
- c. For CET establishment, trace the EIN/s of the establishment/s that is/are included in the report of the CET questionnaire. The status code of these questionnaire/s should be CON with EIN \_\_\_\_ (EIN of the CET establishment).

In case the sample establishment/s of the CET establishment has also responded with status "RET" and data values have been encoded, change the status code of the establishment/s to CON with EIN \_\_\_\_ (EIN of the CET establishment) in the status monitoring database. Also, the data values that have been encoded should be deleted in the respondents' database, thus, retaining the responses of the CET establishment only.

3. Always check consistencies in the EIN, status code, industry code and employment for common retrieved BITS and OWS questionnaire. Print validation proof list to reconcile/ harmonize variables.
4. Ensure that the totals of retrieved/processed (RET, CET) and "spoilage" (REF, STR, TCL, CBL, PCL, OSE, OSP, DUP, CON, OTH) questionnaires in the database are the same as their corresponding total number of questionnaires recorded in *FM-BLES 03-3.19 Status Monitoring of Returned Questionnaires* (retrieved and "spoilage") and *FM-BLES 04-4.7 Monitoring of Data Processing Activities* (encoded questionnaires).

Confirm further that the sum of RFVs and unaccounted questionnaires is the difference of retrieved/processed and "spoilage" questionnaires from the sample size. Should there be any discrepancy, it would probably be due to the adjustments made in the status codes for DUP, CET and the CON questionnaires.

### C. REQUIRED STATISTICAL TABLES

1. After adjustments in status codes have been made, generate the *final survey status report (FM-BLES 03-3.17 Assessment on the Implementation of Field Operations of BLES Survey/s)*.
2. To aide in the preparation of the BUFs (Blowing-up Factors), generate the preliminary *Table A - Distribution of Establishments and Weighted Retrieval Rates by Industry Group and Employment Size*. Note that "transfers to and from" of establishments should have been taken into account in this distribution.
  - a. For each industry and employment size, ensure that:

$$N'_{hk,lm} > \text{eligible}_{hk,lm} > n'_{hk,lm}$$

where:

$N'_{hk,lm}$	estimated population in the initial stratum k and h and in the post-stratum l and m. It is estimated based on an eligibility ratio, i.e., $N'_{hk,lm} = N_{hk} \times \text{eligible}_{hk,lm}/n_{hk}$ where $N_{hk}$ is population count in the initial stratum k and h and $n_{hk}$ is sample count in the initial stratum k and h.
$\text{eligible}_{hk,lm}$	number of eligible samples (RET, CET, REF, STR, TCL, RFV, Unaccounted) in the initial stratum k and h and in the post-stratum l and m
$n'_{hk,lm}$	responding samples in the initial stratum k and h and in the post-stratum l and m

- b. Evaluate this distribution to determine the stratum (employment size) of the domain (industry) or cell (industry and employment size) that should be collapsed because of low response or non-response relative to the number of eligible samples in the stratum or cell. Make the necessary adjustments in the stratum or cell.
3. Generate the final *Table A - Distribution of Establishments and Weighted Retrieval Rates by Industry Group and Employment Size*. Note that adjustments for CET and CON establishments and collapsing of strata or cells undertaken should have been taken into account in this distribution table.
4. Compute the BUFs (ratio of the  $N'_{hk,lm}$  to  $n'_{hk,lm}$ ) per establishment record.
5. Generate the pre-determined *output tables*.

Evaluate numerical consistency of the data of a variable in a statistical table and across statistical tables where the same variable appears. In particular, check for consistency across totals of the same data variable that were disaggregated into different categories (industry and employment size).

6. Prepare *publication tables*
  - a. Since the presentation of the publication tables may be different from the output tables (e.g. percentages or averages), reference should be made to previously published survey results for comparability with updated survey data. Recent economic developments or issuances, e.g. wage orders between previous and current survey should also be considered to explain any variations in the data.
  - b. Part of the validation process also takes into account coherence checks with related survey data generated by BLES or by other establishment surveys (pay for normal regular working time per employee; total employment against similar data from the List of Establishments; trend in total employment against that of the employment index released by the National Statistical Coordination Board from the Quarterly Survey of Philippine Business and Industry, among others).

Reference to administrative statistics should also be made to determine coherence of survey data (injury statistics from BITS against those generated from the Work Accident and Illness Report of the Bureau of Working Conditions and employees compensation claims filed at the Social Security System, Government Service Insurance System and Employees' Compensation

Commission; BITS-based union and CBA statistics against data sourced from the Bureau of Labor Relations, among others).

Coherence does not necessarily mean full numerical consistency. “The coherence of statistical information reflects the degree to which it can be successfully brought together with other statistical information within a broad analytical framework and over time.”

#### **D. FINAL ASSESSMENT OF ACCURACY IN DATA PROCESSING**

Though controls are in place during data collection, and editing/validation and data encoding, there may have been lapses that were overlooked during these stages of survey operation. It is essential then that inaccuracies in data processing, to the extent possible, be finally detected and corrected during output table generation.

Should inconsistencies in output tables be noted, backtracking should start from the rejection list to validation prooflist then to the questionnaire itself.

To monitor possible inaccuracies at this point, refer to *FM-BLES 04-4.8 Monitoring of Accuracy in Data Processing*. This instrument, together with the other survey monitoring forms, will help survey managers to determine areas for improvement not only in the survey procedures but more importantly in enhancing the skills of data processors.