

**10<sup>th</sup> National Convention on Statistics (NCS)**  
EDSA Shangri-La Hotel  
October 1-2, 2007

**Ensuring Fitness-to-Use of Administrative-Based Statistics  
at the Department of Labor and Employment (DOLE)**

by

Editha B. Rivera

For additional information, please contact:

Author's name : Editha B. Rivera  
Designation : Chief, Labor and Employment Officer, Technical Services  
Division  
Affiliation : Bureau of Labor and Employment Statistics  
Address : 3/F DOLE Bldg. Intramuros, Manila  
Tel. no. : (0632) 527-3000 loc 316  
E-mail : [ebrivera49@yahoo.com](mailto:ebrivera49@yahoo.com)

# Ensuring Fitness to Use of Administrative-based Statistics At the Department of Labor and Employment (DOLE)<sup>1</sup>

by

Editha B. Rivera<sup>2</sup>

## ABSTRACT

The paper shares with readers the experiences of the Bureau of Labor and Employment Statistics (BLES) in its attempt to ensure fitness to use of administrative-based statistics generated as by-product of DOLE's regulatory functions. The Project, undertaken in partnership with the Statistical Research and Training Center (SRTC) aims to develop a data quality assessment framework (DQAF) both for survey data and administrative-based statistics. This paper focuses on the latter.

Some 23 data systems in 15 bureaus and attached agencies and 12 regional offices were assessed, using a structured questionnaire, along the six data quality dimensions adopted by Statistics Canada which are **relevance, timeliness, accuracy, interpretability, accessibility** and **coherence**. Assessment was done at two levels: one, at the data capture level, and the other, at the data processing or consolidation level. Results of the data gathering have been validated with the bureaus and attached agencies. To institutionalize the system in the DOLE, a Policy Guidelines shall be issued identifying among others, data quality focal persons in the various offices of DOLE who shall ensure that data quality are observed in the collection and processing of data/information from administrative records and reports. The training of focal persons is an important component of the system.

## 1. Background of the Study/Project

As a public statistical agency, the Bureau of Labor and Employment Statistics (BLES) is expected to provide statistical information and services that will assist the government, including its stakeholders, in policy and decision making processes in the area of labor and employment. It is therefore the responsibility of BLES to ensure that the data it produces are of good quality, i.e., fit to use, by its intended users. To achieve this, the organization has established a Quality Management System that is ISO 9001:2000 compliant.

As the statistical arm of the Department of Labor and Employment (DOLE), the BLES is likewise expected to make available in the Department, statistics for planning, policy review and development as well as for program development. The DOLE as the executive agency for labor and employment produces a large amount of labor statistics as by-product of its administrative functions which the BLES regularly publishes. In the pursuance of its

---

<sup>1</sup> Paper presented at the 10<sup>th</sup> National Convention on Statistics on October 1-2, 2007 at the EDSA Shangri-la Hotel, Mandaluyong City. The author gratefully acknowledges the comments and support given by Ma. Criselda R. Sy, Director IV, Bureau of Labor and Employment Statistics (BLES), DOLE.

<sup>2</sup> Chief Labor and Employment Officer, Technical Services Division, Bureau of Labor and Employment Statistics, Department of Labor and Employment and DQAS Project Coordinator for Administrative-based Statistics.

mandate to develop and prescribe statistical standards in the Department, the Bureau assists in the development of performance indicators, including the formulation of reporting forms and its attendant manuals, as well as in prototyping data processing system.

In the course of formulating the BLES Strategic Plan for 2005-2010, top management had in several forums, aired its concern over the quality of data being generated by agencies under the administrative supervision of DOLE. The low compliance to mandatory reports, the timeliness of data, the layered reporting flow from the operational units to consolidation by staff bureaus, and at times inconsistent statistics from as many sources from within the Department have not gone unnoticed by top management.

In response to this observation, the BLES decided to adopt as one of its strategic objectives for 2005-2010, the development of a data quality assessment framework (DQAF) not only for its survey data, as initially planned, but also for the use of offices within the DOLE to ensure the quality of statistics generated from administrative records or reports. Thus, the development of a Data Quality Assessment System (DQAS) was prioritized as the flagship project for 2006. The DQAS project was undertaken in partnership with the Statistical Research and Training Center (SRTC).

This paper shares with readers the experiences of BLES in the development of the data quality assessment framework for administrative-based statistics. Earlier review of literatures done by the author revealed that DQAF finds applicability in survey data (Rivera, 2006 p. 3-4). However, there are indications that the Australia Bureau of Statistics (ABS) (Yu, 2003) and the Office of National Statistics (ONS) of the United Kingdom (Birch, 2005) have started works on developing DQAF for administrative-based statistics.

## **2. Objectives and Expected Outputs of the Study/Project**

**General:** To develop a Data Quality Assessment Framework (DQAF) for BLES surveys and labor statistics generated from administrative reports/records, and subsequently establish a culture of data quality in the DOLE

**Specific:**

1. To develop parameters of data quality for BLES surveys and DOLE administrative-based statistics;

2. To determine key data systems maintained by DOLE agencies that will be subjected to quality assessment;
3. To formulate guidelines that will ensure data quality from BLES surveys as well as from administrative records/reports;
4. To identify appropriate institutional arrangements within DOLE for managing the implementation of the Data Quality Assessment System (DQAS); and
5. To advocate data quality assessment framework (DQAF) in the Philippine Statistical System (PSS) but foremost in the major statistics-producing agencies.

**Expected Outputs:**

- 📄 Data Quality Assessment Framework
- 📄 Data Quality Standards and Guidelines
- 📄 DQAS Institutional Infrastructure
- 📄 Administrative Issuances Implementing the DQAF in DOLE

**3. Research Methodology/Project Strategies**

Because almost all agencies in the DOLE produce statistics as by-product of their administrative functions, it became necessary that at least one data system per agency be covered by the Project. Thus, with the help of representatives from member agencies of the DOLE Inter-Agency Committee on Statistical Matters (IAC on StatMatters), an inventory of data system per agency was done. From this, one data system was chosen by the agency head to be enrolled for coverage in the DQAS Project.

It is important to note at this point the support and cooperation of the various agency heads in all stages of the Project, as they were consulted and advised on the Project status either through the DOLE IAC on StatMatters or through briefings/orientations conducted.

The Project Consultant recommended that the Data Quality Assessment Framework (DQAF) of Statistics Canada (StatCan) be adapted not only because of its straightforwardness but also because DQAFs adopted by other national statistical organizations seem to emanate from StatCan's framework. In this connection, data elements

for each data quality dimension were formulated. These data elements would be tested on the enrolled data systems.

Thus, a questionnaire was designed and tested in two (2) bureaus, three (3) attached agencies and two (2) regional offices of the DOLE.

On the basis of the findings of the pilot test, the questionnaire was finalized. Personal interviews using the 10-page guide questionnaire were conducted in five (5) bureaus, nine (9) attached agencies, one (1) service office and 12 regional offices covering 23 data systems. Before interviews with responsible staff, a briefing/orientation on DQAF would be presented to top management of the agency in order to solicit their support and cooperation.

The results of the data-gathering phase were validated by concerned agencies. Further, it became the basis in formulating the quality guidelines in the collection, processing and dissemination of administrative-based statistics in the DOLE. The draft guidelines that include quality measures have been presented to the DOLE IAC on StatMatters, and in a workshop participated in by representatives from the BSP, BAS, NCSB and DENR wherein comments were solicited from them.

Likewise, as part of the advocacy component of the project, the draft framework was presented at the NSCB Executive Board meeting to solicit comments and drum up support for the adoption of data quality in the PSS. Meanwhile, presentation of the framework and the draft guidelines are being scheduled in the months of October and November for the provincial staff in all regional offices of the DOLE.

#### **4. Statistics Canada' Data Quality Assessment Framework (DQAF)**

Statistics Canada defines the quality of information in terms of its fitness for use. It is a multidimensional concept embracing both the relevance of information to users' needs, and the characteristics of information such as relevance, timeliness, accuracy, interpretability, accessibility, and coherence that affect how it can be used (Statistics Canada' Quality Assurance Framework 2002, p.1). These different dimensions are defined as follows:

**Relevance:** The *relevance* of statistical information reflects the degree to which it meets the real needs of clients. It is concerned with whether the available information sheds light on the issues that are important to users. Assessing relevance is a subjective matter dependent upon the varying needs of users.

**Timeliness:** The *timeliness* of statistical information refers to the delay between the reference point (and the end of the reference period) to which the information pertains, and the date on which the information becomes available. It is typically involved in a trade-off against *accuracy*. The *timeliness* of information will influence its relevance.

**Accuracy:** The *accuracy* of statistical information is the degree to which the information correctly describes the phenomena it was designed to measure. It is usually characterized in terms of error in statistical estimates and is traditionally decomposed into bias (systematic error) and variance (random error) components. It may also be described in terms of major sources of errors that potentially cause inaccuracy (e.g., coverage, sampling, nonresponse, response).

**Interpretability:** The *interpretability* of statistical information reflects the availability of the supplementary information and metadata necessary to interpret and utilize it appropriately. This information normally covers the underlying concepts, variables and classifications used, the methodology of data collection and processing, and indications of the accuracy of the statistical information.

**Accessibility:** The *accessibility* of statistical information refers to the ease with which it can be obtained from the source agency. This includes the ease with which users can learn of the existence of the information, locate it, and import it into their working environment. It also refers to the suitability of the form or medium through which the information can be accessed. The cost of the information may also be an aspect of accessibility to some users.

**Coherence:** The *coherence* of statistical information reflects the degree to which it can be successfully brought together with other statistical information within a broad analytic framework and over time. This includes coherence between different data items pertaining to the same point in time, coherence between the same data items for different points in time, and international coherence. It does not necessarily imply full numerical consistency.

## 5. Findings of the Data Gathering Phase

### RELEVANCE

- Except for one (1) agency, the rest release their data to both internal (within DOLE) and external clients.
- Feedback mechanisms in place are mostly for internal users but some conduct client satisfaction surveys, users' forum, press conferences/briefings, etc. for external clients. However, some have no system of recording the feedback. Feedback received is mostly on completeness of information, data desegregations and timeliness.
- Many were able to identify potential users of data, i.e., they know who among non-users yet can find usefulness of their statistics not only for academic purposes but even for public policy.

### ACCURACY

- Since most data systems function as part of program monitoring, the concept of client universe (introduced in the questionnaire as equivalent to target universe in order to determine coverage rate) was confused with expected coverage of the program. They took it to mean as target clients, which is difficult to determine. Also, most indicators were considered performance indicators.
- Not all units are covered by the collection procedures, mostly caused by the absence of a list of their supposedly target clients as in the case of the inspection programs and family welfare programs, and non-reporting, in cases where clients are required to submit or register with the DOLE. On the other hand, over-reporting was noted in some data systems that would require review and improvement in the existing reporting system.

- Validation of data from original reporting units is in place and documentation is available especially those data systems included in the Statistical and Performance Reporting System (SPRS). Validation is in most cases, in the form of checking for numerical consistency and completeness of entries. Further, if the bureau that consolidates the data from the reporting units finds inconsistencies or inaccuracies in reports submitted, it returns the reports to the originating unit. However, in some attached agencies and bureaus, validation is not being done at all.

### **TIMELINESS**

- Reports from regional offices have specified time for submission to head offices. Head offices however are forced to release preliminary data due to incomplete reports. Some reasons given for delayed submission of reports were lack of manpower and the adjustments due to the Rationalization Program.
- The usual practice is to submit preliminary reports first to DOLE top officials before releasing to external clients.
- One data system has a practice of considering the data as always preliminary because cases can be re-docketed even after a long period of time.
- There is no revision policy uniformly adopted in the DOLE. Hence, agencies have the discretion to revise their preliminary releases once the data or figures are validated by reporting units.

### **ACCESSIBILITY**

- Statistics are readily available (except in very few agencies where clearance is sought from heads of office in the case of external users) to both internal and external users. They are mostly in printed copies, others in soft copy and at no cost, and available at agency webpage.
- In the presentation of the survey findings, agencies realized the need to adopt a release policy in the DOLE. The appreciation came out as inconsistent data were noted due to varying cut off dates of data releases. In general, this practice leads to confusion on the accuracy of the final figures that are released for consumption of external clients.

### **INTERPRETABILITY**

- There is a Manual of Instructions for data systems covered by the SPRS, mostly in printed form and few in soft copies. It contains concepts and definition, methodology as well as discussions on data limitations. Only timeliness among the data quality dimensions is discussed in the Manual. However, whether it is readily available cannot be said in most offices.
- Statistical reports do not contain discussions on concepts and definitions as well as limitations of the data.
- Some offices entertain queries about the data (usually additional information or unpublished ones) but few have formal system of recording and addressing them.

### **COHERENCE**

- Almost all agencies said that they are not aware of similar data sets even if there are existing ones as in the case of Labor Force Survey data on mode of job search which would be checked for coherence with data on job applicants registering with the Public Employment Service Offices (PESOs). Similar data sets that were used to check include the Commission for Filipinos Overseas (CFO) data for the Overseas Workers Welfare Administration (OWWA) data system as well as the Government Service Insurance System (GSIS) and the Social Security System (SSS) data for compensation cases data.

## **6. Proposed Quality Indicators and Measures for Administrative-Based Statistics**

A data quality assessment framework consists of dimensions with corresponding elements. The framework was tested during the data gathering phase as to its applicability for administrative –based statistics. Another important dimension of such framework is the measurement of quality. Quality measurement is concerned with providing users with information to ascertain whether the data is of sufficient quality and thus fit for use. **Quality measures** (QM) are defined here as direct (often numerical) measures of a particular aspect of quality. These numbers are interpreted against accepted standards often determined from a consensus and depending on particular usage of data. However, in many instances, many quality measures are difficult and/or costly to measure. In some instances, some cannot be measured at all. In such situations, one uses **quality indicators** (QI) instead. Quality indicators consist of information that is a by-product of the statistical process. This may be either present or absent in a particular component of quality. They do not provide

direct measurements of quality but can provide an insight into the efforts and activities exerted to ensure that data would conform to acceptable standards of quality. The Project thus proposes the following quality measures and/or indicators for each element. These quality measures/indicators were initially drafted by the Project Consultant and later validated in a workshop participated in by agency representatives.

Quality Element	Quality Measures/Indicators
<b>RELEVANCE</b>	
R1. Bases (purposes) for the collection of data and indicators collected	R1.1 Enumerate the legal/administrative bases.
	R1.2 Major Indicators
	R1.3 Definition of key concepts
	R1.4 Consultations with users of the data in the determination of key indicators
R2. Identification of Users of the data set	R2.1 Identification of external users
	R2.2 Identification of internal users
	R2.3 Other potential users (aside from those identified in R2.1 and R2.2)
	R2.4 Identification of probable reasons why users identified in have not been using the data.
R3. Feedback Mechanism from External Users	R3.1 Presence of a Feedback Mechanism from external users
	R3.2 From of Feedback Mechanism ( Client Satisfaction Surveys, Public/Users Forum, FGD, others)
	R3.3 Type of Feedback Received (Adequacy in terms of indicators, level of disaggregation, and completeness of information; timeliness)
	R3.4 Presence of a formal system of recording feedback
R4. Feedback Mechanism from Internal Users	R4.1 Presence of a Feedback Mechanism from Internal users
	R4.2 From of Feedback Mechanism ( memo/letters, marginal or personal notes, telephone calls, e-mail, personal meetings)
	R4.3 Type of Feedback Received (Adequacy in terms of indicators, level of disaggregation, and completeness of information; timeliness)
	R4.4 Presence of a formal system of recording feedback
R5. Actions taken on feedback/queries received	R5.1 Type and nature of action taken
	R5.2 Response time
	R5.3 Queries received that cannot be addressed
	R5.4 Satisfaction of user on course of action taken
<b>Quality Element</b>	<b>Quality Measures/Indicators</b>

R6. Incorporation in different stages/ phases of the planning process.	R6.1 Discussed or a major agenda during the agency's planning workshop.
	R6.2 Discussed or a major agenda during DOLE-wide planning workshops.
	R6.3 Involvement of major players (e.g. regional offices, planning office)
<b>ACCURACY</b>	
A1. Coverage	A1.1 Assessment if all cases/events were actually covered
A2. Data System Design	A2.1 Mode of Data Collection
	A2.2 Identification of reporting units
A3. Nonresponse	A3.1 Nonresponse rates
	A3.2 Identification of primary cases of nonresponse (types)
	A3.3 Reasons for nonresponse
	A3.4 Characteristics of nonrespondents vs. respondents
	A3.5 Efforts taken to minimize nonresponse
	A3.6 Methods used in compensating for nonresponse (weight adjustments, imputations) and bases for choice of method
A4. Measurement	A4.1 Validation procedures for basic forms
	A4.2 Validation procedures for summaries / consolidated reports
	A4.3 Validity of items included in all forms
	A4.4 Efforts undertaken to minimize potential measurement errors.
A5. Processing	A5.1 Description of current system of processing
	A5.2 Types of processing errors and their causes
	A5.3 Processing error rates
	A5.4 Number of stages of editing and editing time per stage.
	A5.5 Description of procedures of correcting entries during editing
	A5.6 Efforts undertaken to minimize processing errors
A6. Respondent Burden	A6.1 Number of times an establishment has been covered.
	A6.2 Number of callbacks made
	A6.3 Efforts taken to minimize respondent burden
A7. Evaluation/ Review Process	A7.1 Schedule for review process of the data system design.
	A7.2 Documentation and dissemination of the results of the review process (form and type)
	A7.3 Parties involved in the review process (external reviewer)
<b>Quality Element</b>	<b>Quality Measures/Indicators</b>
<b>TIMELINESS</b>	
T1. Determination of the gap between end of reference	T1.1 Bases for the determination of schedule of release of final data (including parties responsible for the setting)

period and release schedule	T1.2 Description of the role of users in the determination of release schedule
	T1.3 Analysis of implication of gap between end of reference period and schedule of release
T2. Release Schedules (Calendar)	T2.1 Schedule for submission of basic forms
	T2.3 Schedule for validation of submitted basic forms
	T2.4 Schedule for completion of consolidated reports
	T2.5 Description of Preliminary report policy
	T2.6 Dissemination of release schedules (release calendar)
T3. Submission Performance	T3.1 Description of submission policy
	T3.2 Submission rates for basic forms
	T3.3 Reasons for delays in submission
	T3.4 Description of efforts to minimize late submissions
<b>ACCESSIBILITY</b>	
Ac1. Release Policy	Ac1.1 Mechanism for data release
	Ac1.2 Stages
	Ac1.3 Enforcement mechanism including sanctions
	Ac1.4 Disclosure policy (level of details of information to be released)
	Ac1.5 Responsibilities
Ac2. Availability of statistical data	Ac2.1 Access by internal users
	Ac2.2 Access by external users
Ac3. Dissemination Format	Ac3.1 Printed, Electronic, Website
Ac4. Cost of Access	Ac4.1 Description and bases in the determination of rates (cost of access)
Ac5. Form and Mode of Access	Ac5.1 Evaluation of the ease of access
Ac6. Form of Assistance Given	Ac6.1 Describe mechanism of providing assistance in accessing data set.
<b>INTERPRETABILITY</b>	
I1. Documentation of concepts and data collection process	I1.1 Presence of manuals containing concepts and description of data collection process
	I1.2 Availability and form of manuals
I2. Characteristics of publications / statistical reports	I2.1 Form of the report
	I2.2 Report content (definition of key concepts, methodology, limitations)
<b>Quality Element</b>	<b>Quality Measures/Indicators</b>
I3. Interaction and feedback from data analysts	I3.1 Feedback regarding content of reports from data analyst/users.
	I3.2 Feedback regarding form of the report
	I3.3 Actions taken in addressing feedbacks
<b>COHERENCE</b>	
C1. Numerical consistency from	C1.1 Reasons for multiple releases

multiple releases	C1.2 Analysis of the Differences in the figures
	C1.3 Actions taken to ensure numerical consistency
C2. Coherence with other data sets	C2.1 Enumerate similar or related data sets
	C2.2 Description of comparisons made and reasons identified for non-numerical consistency
	C2.3 Actions taken to achieve coherence (not necessarily numerical consistency)
C3. Interaction with data analysts	C3.1 Feedback given affecting coherence
	C3.2 Actions taken to ensure coherence

## 7. Recommendations

The framework proposed here may evolve over time as some dimensions, elements or quality measure/indicators may be changed or added. In the meantime, it would be interesting to test the framework initially in the data systems enrolled in the project. Thus, an administrative issuance for its implementation in the Department is urgently needed. This should be accompanied by a Policy Manual that describes the institutional mechanism for its monitoring (which means identifying focal persons and the monitoring scheme), as well as the training plan for the focal persons. Action Plans may be drawn by the various agencies in improving the quality of their statistical outputs. An audit scheme must also be put in place to ensure the continuing adherence of all DOLE agencies in the data quality system.

In view of the appreciation for the need of a release policy in the Department as a result of the discussions with agency heads, it is imperative that the DOLE adopts one. This shall include identifying the responsible person/unit to release the data, the type and level of desegregation, the timing of release and administrative sanctions for infractions committed. Such policy must be widely circulated and fully understood by all employees. In this regard, an intensive advocacy campaign must be implemented to inculcate the culture of data quality in all DOLE offices.

## 8. Future Activities of the Project

Project DQAS is now at Phase 2. Future activities involve the preparation of the Plan of Action by the different agencies in DOLE to improve the data quality of systems that they have enrolled in the Project, the development of a training module for the focal persons, and advocacy campaign. The Plan of Action may include identifying quality measures and/or quality indicators that are currently practice in the Philippine Statistical System and/or

internationally acceptable practices such as on dissemination policies. It is also envisioned to facilitate monitoring progress in improvement of data quality in the various data system of the Department.

In October and November 2007, the BLES shall undertake advocacy program in DOLE regional offices for its various projects and programs. Part of these is the DQAF in which the draft guidelines will be presented to staff of the various offices of the DOLE in the region.

## References

- Birch, J., (2005) *Quality Measurement and Reporting at the Office for National Statistics (UK)*. Available from [www.dtistats.net/bsug/2005-01/qualitymeasurement.pdf](http://www.dtistats.net/bsug/2005-01/qualitymeasurement.pdf)
- BLES-SRTC DQAS Project Documents including draft report of the Project Consultant Dr. Arturo Pacificador Jr. Manila, 2006-07.
- International Labor Office: *Labour statistics based on administrative records: Guidelines on compilation and presentation* (Bangkok, 1997).
- Office for National Statistics (2004). *Guidelines for measuring Statistical quality* version 1.1. (Online). UK. Available from <http://www.statistics.gov.uk>
- Rivera, E.B., (2006) *Are Labor Statistics From Administrative Records Fit To Use? Development of a Data Quality Assessment Framework*. Paper written with Research Advisor Dr. Arturo Pacificador as a requirement for the UNSIAP-SRTC Fourth Regional Course Research-Based Training Programme, November 2- December 9, 2005.
- Statistics Canada (2002). *Statistics Canada's Quality Assurance Framework – 2002*. Statistics Canada Catalogue no. 12-586-XIE.
- Statistics Canada (2003). *Statistics Canada's Quality Guidelines – Fourth Edition-October 2003*. Statistics Canada Catalogue no. 12-536-XIE.
- Yu, F. (2003). *Making Data Quality Visible in Practice*. **Proceedings of Statistics Canada Symposium 2003 Challenges in Survey Taking for the Next Decade**. Statistics Canada Catalogue no.11-522-XIE.