Distr.
GENERAL

CES/SEM.47/26
29 January 2002

ENGLISH ONLY

STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

COMMISSION OF THE EUROPEAN COMMUNITIES

CONFERENCE OF EUROPEAN STATISTICIANS

EUROSTAT

Joint UNECE/Eurostat Seminar on Integrated Statistical Information Systems and Related Matters (ISIS 2002)
(17-19 April 2002, Geneva, Switzerland)

Topic IV: Ways of making statistical information systems more responsive to users

FULFILLING A STATISTICAL VISION AND STRATEGIC ALLIANCE THROUGH AN INTEGRATED AND DIMENSIONAL DATA SYSTEM

Contributed paper

Submitted by the National Agricultural Statistics Service, Dept. of Agriculture, United States [1]

1. The Strategic Plan for the National Agricultural Statistics Service (NASS) provides the blueprint for its statistical vision and the critical success factors that must be addressed to achieve this vision. One of the keys to achieving its statistical vision is to make much more effective use of underutilized previously reported survey and census data in NASS. Therefore, NASS developed an integrated, easy-to-use, and high performance data system, which will be discussed in detail in this paper. The NASS Dimensional Data System is one of the largest IBM Red Brick databases in the world and contains well over one billion records or individual survey responses. In addition, real-time data from ongoing surveys are continuously loaded into the data system to provide "one-stop data shopping" of current and previous data for input into systems and applications.

2. The NASS Dimensional Data System was designed, developed, and implemented by statisticians in the program divisions with resistance from management in the information technology division. Therefore, the central role of the data system in achieving NASS's statistical vision was not fully embraced throughout the organization. NASS used the services of a senior consultant to work with employees in all divisions to help develop a strategic alliance to pursue NASS's statistical vision. The process used by the consultant will also be discussed in this paper.

Keywords: Data Warehouse, Dimensional Design, Data Access, Data Integration

---

[1] Prepared by Jack Nealon (jnealon@nass.usda.gov) and Jim Cotter.

I. INTRODUCTION

3. The National Agricultural Statistics Service (NASS) administers the United States Department of Agriculture's program for collecting and publishing national, state, and county agricultural statistics. Thousands of farmers, ranchers, agri-businesses, and others voluntarily respond to nationwide surveys conducted by NASS to obtain relevant data on crops, livestock, prices, labor, chemical usage, cost of production, and other agricultural items. Data from these surveys are the primary input to generate agricultural statistics for about 120 crops and 45 livestock items that are published in over 425 national reports each year. In 1997, NASS began conducting the Census of Agriculture, which was previously the responsibility of the Bureau of Census in the U.S. Department of Commerce. The Census of Agriculture provides the only source of comprehensive county-level data which are valuable for measuring agricultural conditions at the local level.

4. U. S. agriculture is a very dynamic industry. For example, the number of hog operations declined dramatically from 247,090 in 1991 to 81,130 in 2001 while total hog inventory increased almost two percent. The trend in agriculture is toward fewer, larger, and more vertically integrated farms. To achieve desired levels of precision from surveys, these fewer, but larger, operations will continue to experience respondent burden from more and more surveys. In addition, there are ever increasing demands for more information on the farm sector. As government policy has been shifting from a planned to an open market economy, there has been more volatility in production and more need for survey information. New kinds of information are desired in areas such as pest management, farm safety, and animal health. The trend toward more individualized farm products due to genetic engineering will also likely create the need for more information. The demand, therefore, for more agricultural statistics is not expected to ease, adding to the challenge of maintaining high response rates.

5. External data users have become more sophisticated. They increasingly are scrutinizing, analyzing, and critiquing published agricultural statistics. They have rising expectations concerning the accuracy of agricultural statistics. Therefore, to meet these expectations, NASS must continue to ensure the accuracy of published statistics.

6. There continues to be a call from the public and private sectors for government to do more with less**.** To obtain more accurate estimates using the same or a smaller budget, NASS is constantly challenged to look for more efficient solutions when developing survey and census procedures and systems.

7. In summary, based on these trends – demand for more information, demand for more accuracy, and demand for more efficiency -- the following three factors are critical to the success of NASS's future agricultural statistics program: (1) maintaining high survey cooperation from data suppliers, (2) maintaining a high level of accuracy in our data and resulting statistics for data users, and (3) creating opportunities for employees to be even more efficient in conducting surveys and censuses.

II. STATISTICAL VISION

8. In the Agency's Strategic Plan, NASS outlined strategies to address these three critical success factors. To maintain **respondent cooperation**, strategies included: (a) investigating alternative estimation approaches to reduce data collection requirements, (b) using data from other sources to reduce the data needed from respondents, (c) developing individualized questionnaires to only ask items needed from a particular operation, and (d) using previously reported survey data to reduce the length of interviews. Strategies for maintaining the **accuracy** of data and the resulting statistics included: (a) investigating new sampling and estimation methodologies to improve the efficiency of designs, (b) integrating previously reported data into more data collection instruments to resolve survey-to-survey data

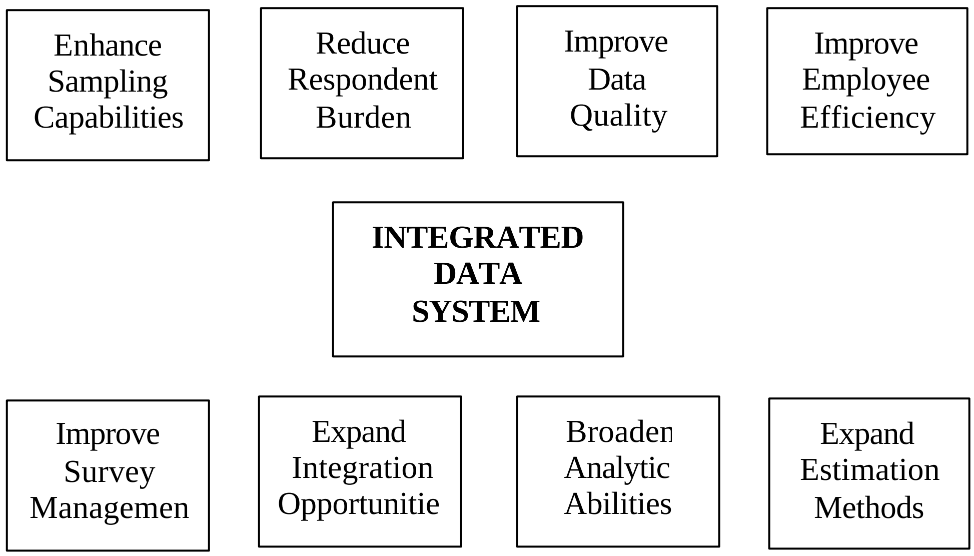
inconsistencies, (c) providing NASS employees easy access to current and historical data to improve the data review and analysis processes, and (d) developing improved procedures for adjusting or imputing for non-response. Finally, to be more **efficient** in carrying out the Agency's mission, strategies included: (a) improving data management to provide efficient and effective movement of data within NASS, (b) adopting technological advances in data collection and data processing to improve the quality and timeliness of agricultural statistics, and (c) integrating survey and census concepts, data, methodology, and processing systems [6].

9. The key to pursuing these strategies to meet NASS's statistical vision rests not with the discovery of new statistical techniques, but on making much more effective use of underutilized previously reported survey and census data in the organization. Most organizations have no shortage of data. However, the data are often not consistently defined and not readily available since the data reside on multiple files, databases, and platforms, and are managed by multiple individuals. For a statistical organization to prosper in the 21st century, it is imperative that its critical data be integrated and easily accessible by its employees and the systems that process the data. This is no longer a luxury, but a business need. Without an integrated data infrastructure, a statistical organization will not be effective in realizing a shared, strategic statistical vision. Realizing this, NASS set forth to develop a data system that would integrate previous survey and census data and be easily accessible by all employees.

10. As depicted in Figure 1, NASS expected the integrated data system to satisfy eight critical components in its statistical vision. These were:

a) Sampling capabilities will be enhanced by having multi-year, multi-survey, and multi-variable data readily available for sampling. Integrating data across states will facilitate national and regional sampling.
b) Respondents' previous data can be used to simplify the response process for some survey questions by providing this data during an interview or, if the data is still current, by using it in an ongoing survey in lieu of re-asking these questions. This would help reduce respondent burden.
c) Data quality can be improved by expanding the use of previous data in the data editing, data imputation, and data analysis phases to minimize survey-to-survey data inconsistencies.
d) Improve the efficiency of employees by providing one-stop shopping of data and metadata.
e) Tracking an operator's response history will facilitate improved data collection strategies, thus improving response rates. Historical data may point to the mode of collection preferred, types of surveys refused, and other information that will help employees make sound survey management plans.
f) Survey and census procedures and processing systems can be integrated and standardized using the integrated data system as the source of data.
g) The analytic abilities of employees will be broadened by making reporter-level data easily accessible across states, years, surveys, commodities, and respondents.
h) Valuable estimation procedures, such as survey-to-survey ratios for matched sample units, not frequently used currently due to lack of data integration, can be used more since previous survey and census data will be readily available.

Figure 1: NASS's Eight Expectations For The Integrated Data System



III. DIMENSIONAL DATA SYSTEM

11. To make this vision a reality, NASS developed an integrated, easy-to-use, and high performance data system, which is one of the largest IBM Red Brick databases in the world. The Data Warehouse contains well over one billion records or individual survey responses, collected from over 200 surveys from 1997 to present, and the 1997 Census of Agriculture. In addition, real-time data from ongoing surveys are continuously loaded into an Operational Data Store that uses the same metadata tables available to the Data Warehouse. Employees have direct access to current and previous data without requiring intervention or assistance from computer specialists or power users. This marriage of current and previous survey and census data from farmers and ranchers in an integrated and easy-to-use system provides "one-stop data shopping" for employees to develop canned or ad-hoc applications and systems to improve the quality of surveys and censuses.

12. There are seven features that characterize the NASS integrated data system which will now be discussed.

13. **Easy To Track Detailed Data History:** By 2003, NASS's Data Warehouse will contain reporter-level data from the 1997 and 2002 Censuses of Agriculture and many surveys from 1997 through 2002. Despite its large size (over 500 gigabytes), the database must still allow employees easy and fast access to data across surveys, censuses, years, states, respondents, and other pertinent factors.

14. In nearly all implementations of tracking data history easily, a Data Warehouse benefits more from the use of a database management system than file-based storage. Since NASS generates thousands of data files each year, an integrated database approach was better than a file-based approach to readily track previous survey and census data.

15. The types of databases most suited for data warehousing are "relational" and "multidimensional." A relational database is designed around tables (rows and columns) while a multidimensional database is designed around multidimensional arrays. Due to its design, a multidimensional database supports a significantly smaller volume of data than a relational database. With the need to track large-volume,

detailed data, NASS selected a **relational database** approach since relational databases scale well beyond NASS's estimated 500-gigabyte data system [8].

16. **Easy To Understand:** The traditional design used for relational databases is called entity-relation or ER modeling. ER modeling is a logical design technique for transaction processing that seeks to remove redundancy in data so that transactions to update, delete, or add database records can be performed extremely fast. The practice of removing data redundancy is called normalizing the data. One can think of an ER design as modeling the "data rules" since its goal is to eliminate data redundancy to maximize the speed of data updates. ER modeling creates many, many tables in most situations since the data are normalized for transaction efficiency. In many cases, end users cannot understand, remember, or navigate an ER design since tables are linked together by a bewildering spider web of joins. Unfortunately, the success of making transaction processing incredibly efficient resulted in databases that are also difficult to query [4].

17. Database designers, realizing the difficulty of presenting a complex schema or model to end users, searched for simpler designs, which led to ***dimensional modeling*** [5]. A dimensional model contains the same information as an ER model, but packages it differently with the goals of providing user understandability and query performance. In a dimensional design, data are denormalized (opposite of ER design) so there are considerably fewer tables. Denormalization encourages data redundancy, such as having a single table with state, district, and county columns rather than having three separate tables. This design is easier to understand and navigate and provides faster data access when querying. Actually, dimensional modeling is the only viable technique for databases designed for ad-hoc query processing. One can think of a dimensional design as modeling the "business rules" since information is combined in tables to make it easy for users to navigate the tables to create queries. Figure 2 summarizes the key differences between ER and dimensional database designs.

Figure 2: Entity-Relation (ER) Design Versus Dimensional Design

ER Design	**Dimensional Design**
Data Normalized	**Data Denormalized**
Many Tables	**Considerably Fewer Tables**
Fast Updating	**Fast Querying**
Usually Pre-defined Queries	**Ad-hoc Or Pre-defined Queries**
Models Data Rules	**Models Business Rules**

18. Every dimensional model is composed of a central table, called a fact or data table, with a multi-part key, and a set of smaller tables called dimension tables that contain the metadata or descriptive, textual information. Each dimension table has a single-part key that corresponds exactly to one of the multi-part keys in the data table. The charm of this design is that it is recognizable and understood by end users and is easy to navigate since a few dimension or metadata tables are logically joined by a central data table.

19. How do NASS's survey and census data fit into a dimensional design? NASS conducts hundreds of surveys annually and an Agricultural Census every five years. About 10,000 unique survey questions are obtained each year, which cover 120 crop and 45 livestock items. Surveys and survey questions can be added, dropped, or changed over time. Traditionally, data have been stored in thousands of data files each year in an "n by p" structure (n rows by p columns). It is worth noting that an ER design for NASS's survey and census data would require over one hundred tables, which would be challenging for users to navigate. Also, ad-hoc query speeds would often be unsatisfactory due to the need to perform many unanticipated table joins when executing queries from many tables.

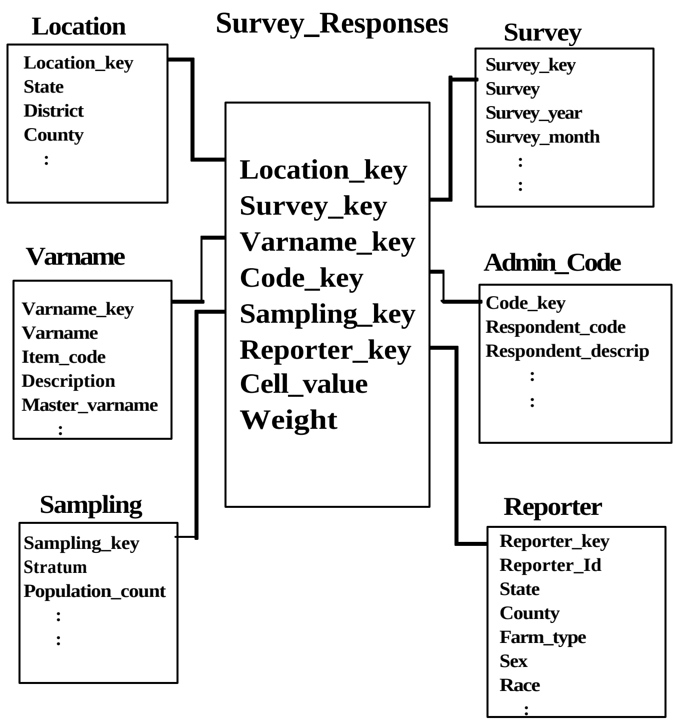
20. NASS decided on the number of tables and the columns in each table for the dimensional design by interviewing employees from all NASS divisions and by building a prototype design for these end users to critique. Several modifications were then made to the dimensional design based on suggestions from users.

21. The two goals of NASS's dimensional design were ***simplicity and speed*** so that: (a) it would be easy for users to understand and navigate the tables to create queries, and (b) queries would be executed in a very timely manner by the database. This led to an effort to have a dimensional design that contained as few tables as practical. There are common dimension tables in many Data Warehouse applications. One can envision a *time* dimension in NASS's design to track survey and/or census data across years, a *location* dimension (state, district, county) to track data across or within states, a *variable name* dimension to tract data for a questionnaire item, and a *reporter* dimension to track data on a particular respondent. Four other dimensions were initially included in the dimensional design to capture analytic needs in NASS, namely, a *survey* dimension containing a list of surveys and censuses, a *commodity* dimension containing a list of commodities, such as cattle, corn, and cotton, a *sampling* dimension containing sampling metadata, and an *administrative codes* dimension containing pertinent questionnaire codes, such as response and respondent codes. Some organizations also create a separate demographic table with information on customers, such as age, race, and sex. NASS decided to include demographic information in the reporter table for simplicity and speed.

22. Based on feedback received from users, time and survey information that were in two dimension tables in the prototype were combined into a single table to simplify the development of queries. Likewise, commodity (corn, soybeans, etc.) and variable name information (corn acres, corn production, etc.) were in separate dimension tables for the prototype, but were later combined to facilitate the navigation of tables and creation of queries. The resulting dimensional design that satisfies NASS's diverse survey and census program is shown in Figure 3 [7].

23. Notice in Figure 3 that a user must navigate only a maximum of seven tables to access any data they need. This design is not the traditional n rows by p columns (n by p) design, but an np by 1 design where each survey response is a row in the survey responses table. The central data table contains the cell value or response and the corresponding weight to adjust or expand the data to an aggregate level, such as state or stratum. The six dimension tables contain the metadata or survey specifications relevant to each survey response. A row in the survey table corresponds to a unique survey or census. A row in the location table corresponds to a unique state and county, and so forth. These six dimension or metadata tables allow for ad-hoc and pre-defined slicing and dicing of the data by any combination of analytic attributes, such as state, survey, commodity, questionnaire item, demographic characteristics, sampling stratum, and respondent. Clearly, NASS's dimensional database design is easy to understand.

Figure 3: NASS's Dimensional Data Warehouse Design



24. **Easy To Use:** The most important aspects of a data system to users are the portions that rise above the surface: the schema, the metadata, and the desktop access tool(s) [3]. The simplicity of the schema has already been described so the metadata and data access tool(s) will be discussed next.

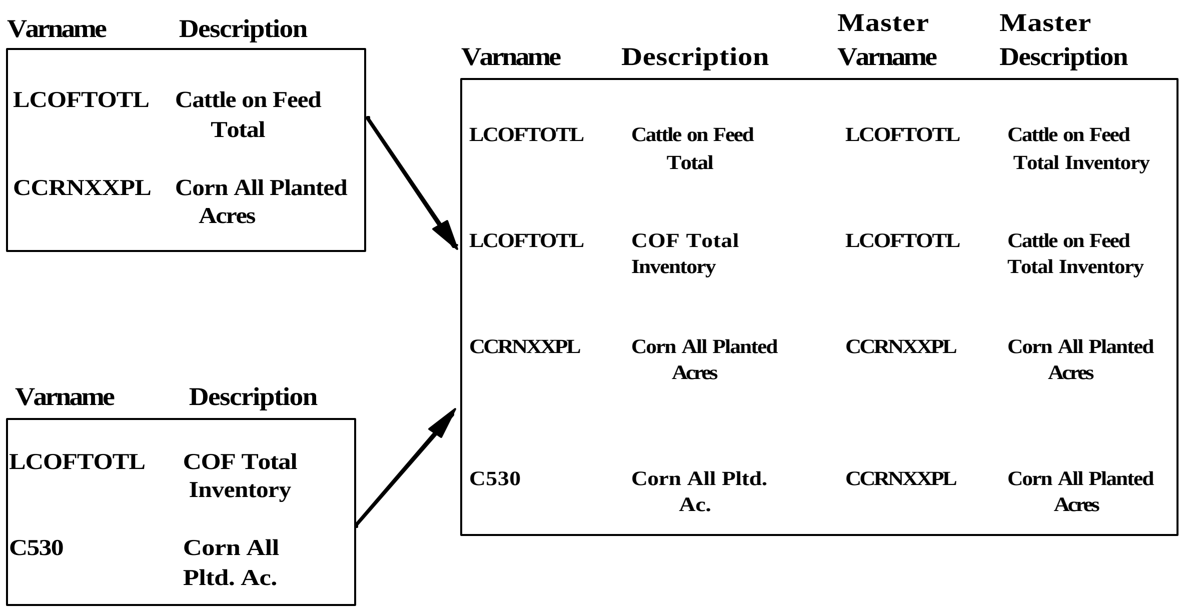
25. NASS's metadata for the Data Warehouse was determined by statisticians throughout the Agency. It is very descriptive and self-explanatory so that users can easily browse the dimension tables to set up a query. The metadata are expressed in business terms that are familiar to employees. A full description for each column in each table is available on-line.

26. The development of the Data Warehouse created an excellent opportunity to standardize metadata. For example, the Master Varname and Master Varname Description columns were added to the Varname table to ensure common definitions across surveys and states. As illustrated in Figure 4, the Varname or variable name on the input data file for corn planted acres is CCRNXXPL for one survey and C530 for another survey. Different metadata from different surveys are preserved, but, simultaneously, common definitions have been created to facilitate analysis across surveys, years, and states. In Figure 4, the Master Varname for corn planted acres is CCRNXXPL. Users, therefore, have a choice when developing a query to use either descriptions from the source system or the standardized descriptions [10].

27. NASS's employees actively participated in the evaluation of the primary data access software to ensure the selected product was easy to use. NASS purchased a drag and drop/point and click query tool called BrioQuery from Brio Technology. This tool supports multidimensional analysis and charting, has

pivoting or column rotation capability for analysis, does calculations, exports results to multiple formats, accesses multiple databases, and can produce standard reports and executive information systems. This easy-to-use tool is so popular with employees that they now also use it with other databases in the Agency. BrioQuery can satisfy most of the basic statistical needs of NASS employees. NASS also uses SAS software, which satisfies additional needs from power users, such as advanced statistical programs, complex statistical analysis, and data mining.

Figure 4: Metadata Standardization



28. **Easy To Administer:** When visiting organizations who had developed or attempted to develop a Data Warehouse System, an issue that often surfaced was the "complexity" of some database management systems and data access tools. Depending on the software, an organization might have from one to three database administrators. NASS did not have the luxury of having several full-time database administrators devoted to its data system. NASS researched and selected a primary data access tool (BrioQuery) and a database management system (IBM Red Brick) that required modest administration. NASS currently has well over one billion records in the Data Warehouse and 700 users throughout the Agency, but only one database administrator. This is testimony to the ease of administration of the Data Warehouse.

29. **Fast Data Loading:** It is important that a database system load data in a timely manner so users have access to the data without delay. The hardware for the Data Warehouse is an IBM/RS6000 M80 four-processor UNIX machine with eight gigabytes of random access memory or RAM. The IBM Red Brick database loads up to 40 million rows of data per hour using this machine, which includes building indexes and conducting referential integrity or edit checks. This translates to loading all the edited and imputed data files for all states for a major national NASS survey in about 10 minutes.

30. **Fast Data Access:** The best illustration of query speed in the Data Warehouse (with over one billion records from over 200 different surveys) is to provide examples. After pressing the "process" button in BrioQuery for each of the following five, simple queries, each query was sent to the Data Warehouse on the IBM/RS6000 M80 in Missouri, and the results returned to a desktop computer in Washington, D.C. in less than five seconds for each query. This very impressive query speed was unheard of in NASS prior to the implementation of this data system.

a) ***One Reporter*** - all data for one reporter in Iowa's March 1999 Hog Survey (27 records returned)
b) ***One Stratum*** - total hogs and pigs for each reporter in sampling stratum 98 for two Hog Surveys in North Carolina (91 records returned)
c) ***One County*** - non-response-adjusted total hogs and pigs (Cell value multiplied by Weight in Figure 3) for Carroll County in Illinois' 1997 Census of Agriculture (1 record returned)
d) ***One State*** - all reporters with cattle in Montana's 1998 July Cattle Survey (280 records returned)
e) ***One Survey*** - all reporters with hogs in the 1997 U.S. Census Not-on-the-Mail List Survey (205 records returned)

31. **Current and Previous Data Integrated:** Real-time data from ongoing surveys are continuously loaded into an Operational Data Store that uses the same six metadata or dimension tables available to the Data Warehouse. The extraction, transformation, and loading system detect, processes, and loads new data files within an average of six minutes after the data arrive on the server. The versioning facility in the IBM Red Brick database system allows users to constantly query the database even while the database is being loaded, with no degradation in query performance. Therefore, "one-stop data shopping" of current (Operational Data Store) and previous (Data Warehouse) data is provided to employees in an integrated environment. Enhancements in survey sampling, survey management, data editing, non-response imputation, data analysis, modeling, and estimation, which are called for in NASS's statistical vision, can now be achieved using the Dimensional Data System as the foundation for all census and survey methodology and processing systems.

IV. STRATEGIC ALLIANCE

32. The NASS Data System was designed, developed, and implemented by statisticians in the program divisions. Input from end users throughout NASS was actively sought and incorporated into the design. However, the system was developed and implemented with resistance and negligible assistance from management in the information technology (IT) division who did not understand the strategic value of data or its role in satisfying NASS's statistical vision [9]. Even today, the responsibility for the Data Warehouse resides with statisticians in a program division. As the Data Warehouse has grown over the years, opportunities have emerged to leverage existing partnerships and build a strategic alliance across all divisions, including the IT division.

33. NASS is a functionally-structured organization that depends upon cross-functional communication and collaboration to be efficient, effective, and successful. Managers realize that they need to collaborate. Still, attempts to collaborate sometimes fail, primarily due to a lack of embracing shared goals and benefits. To build a strategic alliance across all divisions requires a shift in culture for some managers from knowledge hoarding to knowledge sharing, from controlling to partnering, and from independence to interdependence. This shift can occur if there is a shared vision to rally around and a willingness to change the culture.

34. NASS is currently undertaking a major reengineering and integration effort in preparation for the 2002 Census of Agriculture. This effort is called PRISM (Project to Reengineer and Integrate Statistical Methods), which embodies NASS's statistical vision. All divisions are involved in this effort and embrace the shared statistical vision. However, the central role of the Dimensional Data Warehouse System in achieving this vision had not been fully embraced. Anyone who has actually implemented a Data Warehouse knows that much of their energy was likely spent dealing with political or organizational issues, which frequently sidetrack or derail Data Warehouse initiatives [1]. Some of NASS's senior management realized that political or organizational issues concerning the integrated data system were impeding progress in achieving the PRISM goals. Therefore, NASS used the services of a senior consultant who was familiar with NASS's Data Warehouse and trained in facilitation and building

alliances. The senior consultant worked with employees from all divisions involved in PRISM to remove barriers to communication and collaboration, help articulate the shared statistical vision, and develop a strategic alliance across divisions to pursue the PRISM vision [2]. The process or steps used by the consultant to help build the strategic alliance using the Dimensional Data System as the focal point for integration will now be outlined.

35. *Step 1:* The consultant met with NASS's senior executives to discuss any concerns or questions they had about the planned processing architecture for the PRISM initiative. This meeting provided an opportunity for the consultant to build some trust and credibility with the senior executives. Fortunately, he was already respected by a couple senior executives due to his earlier contributions assisting NASS with the design and implementation of the Data Warehouse.

36. *Step 2:* The consultant then met with all the major players in each division who were working on the PRISM processing system. The main purpose of this two-day meeting was to listen to their concerns. He encouraged open and honest information sharing and promised to meet later with the senior executives to share their concerns and recommendations. Issues discussed included ways to improve communications, the PRISM management structure, concerns about unclear or changing direction, performance expectations for the processing system, and the role of the Dimensional Data System.

37. *Step 3:* The consultant met with the senior executives again to share the concerns from employees and propose several recommendations, such as appointing a single program manager, dedicating more full-time employees to the PRISM effort, and better leveraging current technology, such as the Data Warehouse. The senior executives embraced and implemented the recommendations. Their support also helped establish the credibility of the consultant with the PRISM processing team as he began to help build an alliance across divisions.

38. *Step 4:* The consultant returned monthly to meet with the PRISM processing team to continue the process of encouraging communications, knowledge sharing, and collaboration. He was able to ignore political or organizational issues that were clouding decision-making and encourage communications across functional areas by giving everyone an opportunity to be heard and contribute. NASS is not optimally designed to handle initiatives of this magnitude due to its functional structure. Therefore, the consultant emphasized a program or process orientation rather than a functional view and initially focused the discussions on developing a shared or common direction and focus. He also facilitated discussions on overcoming some barriers to success, ways to proceed efficiently and effectively, clarifying expectations for the processing system, and utilizing the Dimensional Data System to solve several of the processing challenges. Through these open discussions, a strategic alliance across divisions was forming.

39. *Step 5:* The consultant returned several more times to work closely with the PRISM program manager on the project plan, time lines, and resource needs. He also facilitated additional discussion sessions with team members on complex parts of the PRISM system. The discussions resulted in agreement on critical issues, such as the central role of the Dimensional Data System.

40. *Step 6:* The consultant met again with NASS's senior executives to update them on the progress that had occurred and discuss some issues of concern. Action was then taken to address the concerns.

41. *Step 7:* The consultant has continued to meet with the program manager and team members, when requested. Through his facilitation, an effective strategic alliance has been developed across divisions with agreement that future census and survey processing systems need to be fully integrated with the Dimensional Data System.

V. NEAR FUTURE

42. NASS has only “scratched the surface” in realizing the vast potential of the Dimensional Data System. To date, usage has focused on ad-hoc data access and data analysis. In the future, NASS’s census and survey processing systems will be integrated with the Dimensional Data System. By 2003, the system will contain detailed, reporter-level data for many surveys from 1997 through 2002 and for the 1997 and 2002 Censuses of Agriculture. In addition, aggregate data, such as survey estimates and official statistics, will soon be integrated into the Dimensional Data System for analytic purposes. Survey estimates include not only expanded totals and ratios, but also valuable analytic information, such as the number of usable reports, number of positive reports, and the coefficient of variation, that are available in the summary or tabulation system. This integration will provide “one-stop shopping” of reporter-level data, survey estimates, official statistics, and the corresponding metadata. Therefore, an evolution of the Dimensional Data Warehouse is underway from “the final resting place for data” to an “integrated data environment serving as the source of current and previous data during a survey or census.” Systems are being integrated in a hub and spoke fashion with the Dimensional Data System serving as the hub. The dimensional design has not only been used to model the business rules, but is now being used to model the systems architecture of the future, thereby maximizing the role of the Dimensional Data System.

43. NASS must continue to strengthen its analytic capabilities to meet the rising expectations for accurate statistics. Ad-hoc data analysis, which has traditionally been conducted primarily by power users in the Agency, should be performed by all statisticians in the future. This expectation can be achieved through the Dimensional Data System.

44. In closing, NASS has entered a new era as an agricultural statistics organization where current and historical data are easily accessible by employees to provide maximum benefits to the agricultural statistics program. Many opportunities for improving the agricultural statistics program that once were a dream of some senior managers in NASS are now readily achievable. It is now in the hands of employees throughout NASS to make NASS’s statistical vision a reality by making greater use of the Agency’s Dimensional Data System.

VI. REFERENCES

[1] Demarest, Marc (1997), *The Politics of Data Warehousing*.
[2] Hocutt, Robert J. (December 2001), *PRISM Correspondence to Jack Nealon*.
[3] Inmon, W.H., Welch, J.D., and Glassey, Katherine L. (1997), *Managing the Data Warehouse*, Wiley.
[4] Kimball, Ralph (August 1997), *A Dimensional Modeling Manifesto*, DBMS Magazine: Tools & Strategies For IS Professionals.
[5] Kimball, R., Reeves, L., Ross, M., and Thornthwaite, W.(1998), *The Data Warehouse Lifecycle Toolkit*; Wiley.
[6] *NASS Action Plan 1998* (updates and complements previous NASS Strategic Plan) (1998); National Agricultural Statistics Service (NASS); U.S. Department of Agriculture (USDA).
[7] Nealon, Jack (2000), *Improving Our Agricultural Statistics Program Through Easy And Fast Access By Employees To Previous Survey And Census Data*, NASS/USDA.
[8] Simon, Alan R.(1997), *Data Warehousing For Dummies*, IDG Books Worldwide, Inc.
[9] Vogel, Fred A. (2001), *The Data Warehouse: A Modern System For Managing Data*, NASS/USDA, Prepared for the 2001 Conference on Agricultural and Environmental Statistical Applications in Rome (CAESAR).
[10] Yost, Mickey (1999), *Using A Dimensional Data Warehouse To Standardize Survey And Census Metadata*, NASS/USDA, Prepared for the 1999 Federal Committee On Statistical Methodology Conference.