

Introduction

BusinessCounts is a geographic summary database of business establishments and employees for nearly ten million businesses and one hundred and thirty million employees. The database is available for all standard levels of geography including block group and is supplemented by occupational data from the Bureau of Labor Statistics and the County Business Patterns program. The primary variables available include:

- Total Establishments, Employees
- Size Establishments by size
- Occupation Employment by occupation
- Major Industry Establishments, Employees
- NAICS Establishments, Employees 3 and 4 digit
- Retail Sales estimates, employees, establishments

In addition to the current year estimates (January 2011), AGS provides a historical file in order to permit trending and tracking of changes over time. The historical file is five years prior to the current year release.

Methodology and Data Sources

The core source for the InfoGroup (formerly known as InfoUSA) Business Database that is built from a careful integration of commercial databases, compiled white and yellow page directory data, city directories, corporate annual reports, and securities filings. The BusinessCounts file is current to January 2011.

The database has been thoroughly cleansed for address consistency and geocoded. Virtually all records within the database are geocoded, although in some cases with less positional accuracy than others.

Occupation

The occupation tables were modified some years ago to include seventeen separate categories using improvements to the underlying industry-occupation tables released by the Bureau of Labor Statistics. The sales occupation (BCCYSALES) has been split into two categories: Sales Professionals (BCCYSALPR) and Sales Workers and Clerks (BCCYSALCL). The clerical group (BCCYCLER) has been split into Administrative Support Workers (BCCYCLERO) and Technical, Sales, and Administrative Field Operations (BCCYCLERF). The "Other" Services group has been broken into Site Based (BCCYSERV1) and Field Based (BCCYSERV2). Finally, the Production occupations group has been split into site and field based operations (BCCYPRODS, BCCYPRODF).

These changes are very useful to retailers and food services companies who are most interested in the daytime population, as many of these field operations staff who may be assigned to a particular location are in all likelihood absent on most workdays from that location.

Business Functional Classification

This set of 38 business types, based roughly on land use style, is adapted from one which was used several companies during the early 1990's and is, in our view, a very useful way of presenting business data to clients on the style of operation rather than the specific type of products produced. The variable ESCCYMFGSM, Manufacturing: Small Firms is a good indication of the presence of small industrial parks. A small manufacturer of plastics, for example, is likely to be different in its land use and employment profile than a large producer.

Administrative/Headquarters

We have added counts of employees and establishments for administrative or headquarters firms, as identified in the InfoGroup file.

SIC Classifications

The classification by two-digit SIC code has been slightly re-worked in order to merge certain small related categories while splitting other large categories. The result is a slightly shorter list of categories at this detail level. For example, the categories 01 (Crop production), 02 (Livestock production), 08 (Forestry), and 09 (Fishing) have been aggregated into a single category.

A significant number of establishments are not SIC coded within business list files, most often including those small firms for which neither private nor public records exist. Many of these uncoded firms are simply individual holding companies, DBA ("doing business as") names, and new firms that have not yet been well documented in multiple sources.

NAICS Classifications

A few years ago, a new classification system known as NAICS (North American Industrial Classification System) was introduced to replace the aging and outdated SIC (Standard Industrial Classification) system, which has been used for several decades (with some modifications over time).

The NAICS classification is intended to more accurately reflect the growing reliance on information-based companies which were not well classified under the SIC system, reflect the changes in retailing towards large, multi-faceted retailing (e.g. home stores, grocery and drug stores, etc.), and to eliminate business types which simply no longer exist. In most cases, the orientation is more towards the product/service being offered and less towards the style or raw materials of its creation. For example, the textile manufacturing SIC group has been split by whether the product is an "end-user" product or is simply a raw material in further manufacturing. The manufacture of automobile seat fabrics is now classified under the automobile manufacturing group rather than the textile manufacturing group, reflecting the ultimate use of the product rather than its composition.

NAICS classification data are presented at the 3 and 4 digit detail levels. A 2 digit summary can be

readily created using the formula features of Snap, should such be desirable. These are roughly comparable to the SIC “Major Industry” and 2 digit levels, although there are more categories under NAICS than under SIC.

Employees

The file includes both a size classification (e.g. 1 to 4 employees) and in a significant number of records, verified employee counts. When an actual employee count was given, this was used directly.

In order to estimate employees for those establishments either only a size class range is available, the latest County Business Patterns (CBP) database, published annually by the Census Bureau was thoroughly analyzed. For each four digit SIC code, the average number of employees per establishment of each size class was computed in order to provide a base estimate. These were further refined by using major industry average sizes by county, since much of the county level detail is suppressed within the CBP in order to avoid the possible disclosure of individual establishment employee, payroll, and sales volumes.

Once the initial estimates were applied, the results were evaluated on a county level basis in order to ensure consistency with county totals for each major SIC group, and nationally to ensure consistency with the detailed four-digit SIC level.

It should be noted that the employee size estimates for the Public Administration (SIC 91-98) major group are not particularly accurate. Employee estimates for individual government offices are simply not easily obtained and are generally afforded less attention by the major business list providers than private sector establishments. Further, neither the CBP nor the Economic Census databases cover this important sector of the economy. The total employee estimate is therefore rather low for this sector as a whole.

Retail Sales

Retail sales estimates were computed at the establishment level for those retail establishments which were not identified as administrative. Using the 2002 Economic Census (Retail Trade) as a benchmark, estimates by state and NAICS code were created which related retail sales to employee counts. These estimates (e.g. average sales per employee) were then converted to the SIC coding system as accurately as possible and updated to current figures obtained from the Monthly Retail Trade reports issued by the Census Bureau. Regional and store type differences were therefore accounted for in this procedure. Retail sales estimates are excluded for catalog and mail order firms, since these are not typically local in nature.

Comparability to Other Sources

Several additional sources of national and state level estimates from the BLS (Bureau of Labor Statistics) and the Census Bureau were used to verify summary counts in the final database. In general, the database agrees substantially with these estimates. The major sources of difference occur in several

areas. These areas of disagreement are noted below:

- Manufacturing employment in BusinessCounts is higher than corresponding statistics from the BLS. In large part, this reflects the use of the “primary” industry within BusinessCounts. Often, many manufacturing companies also have wholesale trade and finance divisions. The employment within these divisions is attributed to the main SIC category in this file, hence manufacturing estimates are higher than in BLS sources.
- Agricultural establishments, specifically farms, tend to be underreported in the database, so total agricultural establishment and employment counts are low relative to other sources. The so-called “primary” sector is not typically well represented in either the economic censuses or the annual County Business Patterns files, and is subsequently difficult to estimate with reliability.
- Service employment in the BusinessCounts file is higher than in equivalent BLS sources, primarily as a result of classification issues. In official BLS reports, educational institutions and employment is reported within the appropriate level of government (e.g. state versus local) whereas in BusinessCounts, these are reported in the educational services category.
- In addition, many public and quasi-public agencies are coded to the type of service they provide rather than as public sector establishments. Public sector estimates in BusinessCounts are therefore lower than published figures by an amount roughly equivalent to the over-count in services. In addition, the tendency within business list products such as InfoUSA is to put more emphasis on private sector establishments than on public sector establishments. Subsequently, in many cases not only is there no actual employee count, but often no size class information as well. Since the Census Bureau surveys of establishments typically exclude public sector establishments, and what statistics are available are typically only at a state level, the public sector employment estimates are substantially underestimated and should not be relied upon for many analytical applications.