

**Data Base Project**  
**Dr. Robert Fountain**  
**fountain@RegionalEconomics.org**

**Notice to Users:** This project contains not only addresses and citations for frequently used data, but is annotated to show processes for accessing and interpreting specific data.

The file is updated frequently, but probably not as often as data providers change their websites, so if you find an obsolete link or description, please send a note to the email address above. Please also let us know about additional data sources in the regional economics sphere which you can share with us.

## **Data Sources and Notes**

**Higher Education** Enrollments, graduations, degrees granted, by type of higher education institution, annually, most by county or region.

Source: California Postsecondary Education Commission

<http://www.cpec.ca.gov/OnLineData/FindRpt.asp>

High School graduates

Higher education enrollment by segment (i.e., CSUS, UC, Community colleges, Independent Institutions, etc.)

Statewide Degrees granted (by segment, but not by county or region)

Subsets by gender and ethnicity.

Data files: Data by Topic/Education/ in csv format

Analysis file: Comparisons of Higher Education Data.xls

**Employment US** Non Agricultural Wage & Salary Monthly (SIC industry data also available from this data source.)

Bureau of Labor Statistics Selective Access

File Name: US Employ Non-farm W&S 1985-2000 mthly.txt

File Name: US Computer and Office Equipment 1990-2000.txt

File Name: US Computer and Data Processing Services 1990-2000.txt

Link to data source: <http://146.142.4.24/labjava/outside.jsp?survey=ee>

Select variables, choose **not seasonally adjusted**. Use the reformat option to add format instructions (table, text) and time interval. Download as text, convert to excel. Note that rows contain a 13<sup>th</sup> column for annual, which has to be manually separated from monthly data for computations. This is done in our USA MONTHLY W&S EMPLOY worksheet, in the FORMAT CONVERSION page.

**Housing Permits** US MSA's, single and multi family, monthly

US Census Bureau

Folder Name: Housing/Permits

Links <http://www.census.gov/const/www/permitsindex.html>

New search entry

## **Data Base Project**

### **Dr. Robert Fountain**

**fountain@RegionalEconomics.org**

<http://www.census.gov/const/www/C40/table3.html#monthly> Old direct table access

The data are in Table 3 which includes all US MSA's, Table 2, data for States, and Table 1, US. The tables are in the form: TB3YU1299 for table 3, year to date, units, 12<sup>th</sup> month, 1999. (Use the year-to-date 12<sup>th</sup> month to get annual totals). Other tables are for the current month only (T3U1299) or the permit values (T3V1299).

Be careful of VALUE vs. NUMBER OF UNITS which are different tables.

Note that the tables contain CMSA and PMSA values. The PMSA values are listed just under the CMSA entry, not by individual alphabetical order (i.e., under San Francisco-Oakland-San Jose CMSA one finds all the individual PMSA's.)

Beware that in converting the text file to Excel (using fixed column width, as the txt files are not delimited) the column titles get garbled. Compare to the text file to make corrections.

There is no "multifamily" entry, but rather individual entries for 2 units, etc. The easiest way to compute all multifamily is to subtract single family from total. USA values are in a separate table, Table 1a, use largest sample category.

**Housing Affordability**, US MSA's, quarterly  
National Association of Home Builders.

Folder Name: /Housing/Affordability

Link: <http://www.nahb.net/>

(Note: data restricted to NAHB members; our access courtesy of John Orr, BIASC.

### **Interest Rates**

Source: Federal Reserve Economic Data site

<http://www.stls.frb.org/fred/data/irates.html>

Data for many types of interest rates, bond yields, and mortgage rates at monthly intervals for US. Some series date back to 1919, mostly from 1970's. Data is in text form.

The FHA 30 year mortgage rates series was formerly widely used but has been discontinued in favor of the Conventional Mortgages series.

### **Population and Projections**

**Population for California Counties 1990 – 2040**

Compiled in data file

Compiled from source <http://www.dof.ca.gov/newdr/california.txt>

See column id codes in <http://www.dof.ca.gov/HTML/DEMOGRAP/Race.htm>

**Data Base Project**  
**Dr. Robert Fountain**  
**fountain@RegionalEconomics.org**  
**Source: California Department of Finance**

**ES-202 Covered Wage & Salary Employment**, also called the Employment Security Series. Provides detailed sic detail on employees, firms, earnings. Covers all employees in every industry, the most detailed data available. Availability of data is approximately 2 years old. Download htm format directly into excel.

<http://www.calmis.cahwnet.gov/file/es202/cew-select.htm>

## Data Base Project

Dr. Robert Fountain

fountain@RegionalEconomics.org

### Some computation methods and tips using Excel

1. When creating a research product which uses data from multiple source files, the following process is recommended:
  - a) Keep the original downloaded files in the DATA directory, with filenames annotated by download date, even when obsolete and even when newer downloads contain the entire historical record. Beware of “revisions and benchmarking” in which historical data are suddenly changed without notice by data providers.
  - b) To create a research product, make a new Excel spreadsheet, titled to reflect the research product, which contains as separate worksheets all of the original data files to be used, without any modification. Then create additional “computations” worksheets in the Excel file to extract and manipulate only the relevant parts of the data files (see technique using (=address) rather than copy paste). The transfer of relevant data from the original data files to the computation worksheet is therefore conducted within the single Excel spreadsheet (but across worksheets) which keeps everything in one file.
  - c) To update such a file, pasting new original data sheets onto the data worksheets makes updating easy, and even automatic in some cases.

### 2. Prepare worksheets to avoid errors.

When moving data from one worksheet to another, the following steps avoid confusion and mistakes.

- a) Move the source and destination worksheet tabs at the bottom of the Excel screen so the ones you are working on are adjacent.
- b) On the source worksheet, set up the screen with appropriate zoom magnification, freeze panes, and use the fill colors to mark entry ranges (like the date columns) before transferring any data. A few minutes setting up the worksheets before transferring data can save days of tracking down errors which occur when one gets disoriented in large spreadsheets.

3. **Data transfer method.** When copying anything within a spreadsheet or worksheet, do not use copy-paste but rather the linked address method =(address). This allows a trail by which to verify the data and update it very simply. Learn how to use the “FILL HANDLE” with this feature.

When copying between different spreadsheets, especially if they are in different folders, it is better to use the copy-paste because if you ever move any of the files linked by =address the destination file will not be able to find the source file.

**Data Base Project**  
**Dr. Robert Fountain**  
**fountain@RegionalEconomics.org**

4. **Value tables rather than relative address tables.** Before sorting or graphing, copy the final computed data you are working on to a new table using (copy-paste-special-values only). Graph and Sort cannot work on relative value address entries like =sum( ) or (=D12), and may appear to have worked correctly but give outrageous results.
  
5. **Data File organization.** There may be three files created in the steps between raw data and finished research product (a graph or table ready to insert into a report).
  - a. Downloaded raw data files in the DATA directory, in native formats, with origin file names modified to indicate the download date.
  - b. A topic spreadsheet in the DATA directory, which may have many data files inserted into it as well as computations, selections, or reformats of the source data. The topic spreadsheets may become large over time, especially if they contain many geographies.
  - c. A PRODUCT file, in the projects directory, which has some computations or data tables copied from the topic spreadsheet but also has graphs or tables created from the computations.

## Data Citations

- Worksheets and data products must contain the information below, preferably posted in a “readme” worksheet or in the upper left the computations page.
- - Variable name(s)
  - Source description and provider
  - Source File name in the Data directory
  - Geography (i.e. County, city, MSA, etc.)
  - Time intervals reflected in the product
  - Frequency (monthly, quarterly, annually, etc.)
  - Notations, assumptions, instructions. Seasonally adjusted, annual average of latest 12 months, 1999 benchmark (earlier data has different values), etc.
  
- Leave a complete trail. Remember that researchers other than yourself may be working with this file and need to know what the data are. Research integrity can be seriously compromised if others are not able to verify your data and results.