Importing the US 1990 Census Data into a Relational DBMS

Technical Report

Peter McBrien
Dept. of Computing, Imperial College,
180 Queen's Gate, London SW7 2BZ
pjm@doc.ic.ac.uk

Tuesday 9th April 2002

1 Introduction

This report documents the relational database version of the 1990 US Census data [Cen90], as created by the Java program originally written as part of a student project at Imperial College [Gun01], and augmented by the author of this document.

The database is mainly intended to be used for project work in the **Dept. of Computing (DOC)**, Imperial College, that requires a medium sized 'real world' database to work on. However the Java program way prove useful to anyone who wishes to import the US Census data into a relational database.

2 Summary of Data

This section presents a summary of the information present in the relational database after importation using the Java program related to this report. It is largely a copy of the information presented on [Cen90], altered to reflect the manner in which Java program processes the information.

2.1 place

The Place file contains all places in the U.S. and outlying areas (excluding MCDs and CCDs) as of the 1990 Census. The fields are as follows:

- stateCode: state FIPS Code
- countyCode: county FIPS Code
- name Includes name of county. Counties don't have a type, but do have the word 'County' as part of the name.
- population
- housingUnits
- landArea: land area included in km^2.
- waterArea: water area included in km^2.

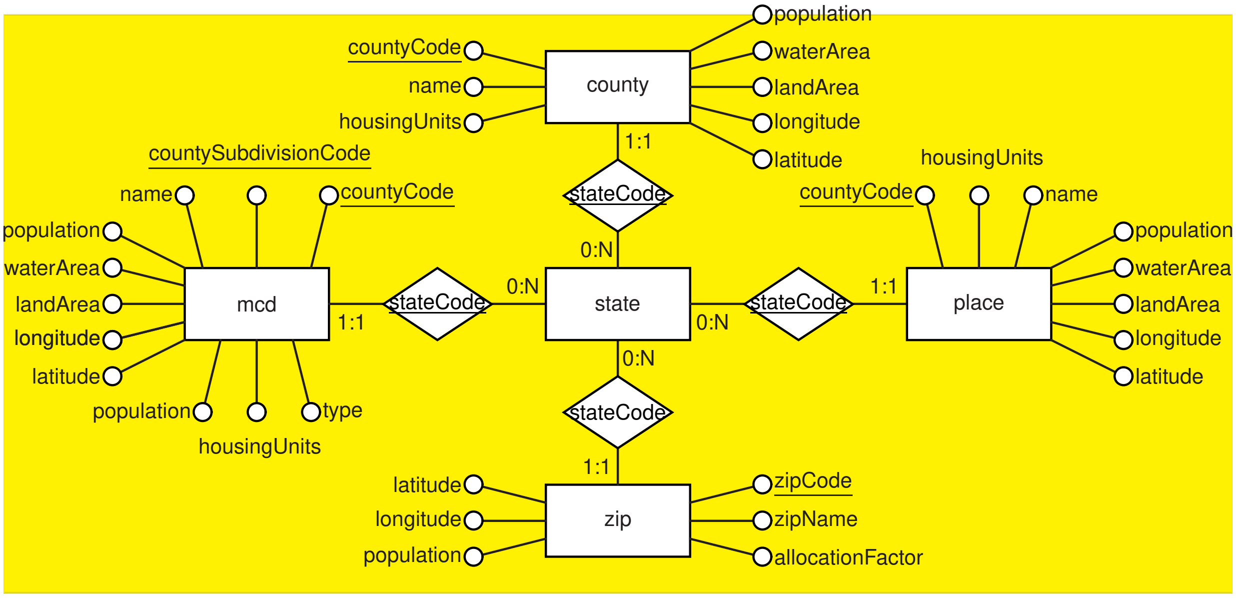

Figure 1: ER model for relational database schema

- latitude: in degrees, + or − denoting N or S latitude, respectively
- longitude: in degrees, + or − denoting E or W longitude, respectively

2.2 mcd

The MCDs file contains all MCDs and CCDs (county subdivisions record type 060) places in the U.S. as of the 1990 Census. The fields are as follows:

- stateCode: state FIPS Code
- countyCode: county FIPS Code
- countySubdivisionCode: county subdivision FIPS Code
- name
- type of place (either city, town, borough, village, . . .).
- population
- housingUnits
- landArea: land area included in km^2.
- waterArea: water area included in km^2.
- latitude: in degrees, + or − denoting N or S latitude, respectively
- longitude: in degrees, + or − denoting E or W longitude, respectively

2.3 county

The Counties file contains all counties in the U.S. as of the 1990 Census. The fields are as follows:

- stateCode: state FIPS Code
- countyCode: county FIPS Code

- name Includes name of county. Counties don't have a type, but do have the word 'County' as part of the name.
- population
- housingUnits
- landArea: land area included in km^2.
- waterArea: water area included in km^2.
- latitude: in degrees, + or − denoting N or S latitude, respectively
- longitude: in degrees, + or − denoting E or W longitude, respectively

2.4 zip

The 1990 Zip Code by State file. The fields are as follows:

- stateCode: state FIPS Code
- zipCode: 5-digit Zipcode
- zipName: Zipcode Name
- latitude: in degrees, + or − denoting N or S latitude, respectively
- longitude: in degrees, + or − denoting E or W longitude, respectively
- population
- allocationFactor: decimal portion of state within zipcode

2.5 state

Not avaialable from the USCensus web site is the list of FIPs statecodes, which is used to tie the four USCensus tables together.

- code: two digit FIPs code for the state
- abbr: two character FIPs abbreviation for the state
- name: name of state (which might be a name of a US protectorate)

3 Using the Imperial College version of the Database

A version of the database is maintained in a Sybase DBMS within DOC, for use only within the department. Any users of the database are expected not to make excessive use of the server.

To use the database, called `USCensus1990`, you should use the jConnect JDBC driver (available from the Sybase web site) to connect with the Sybase database, and use the following details in a `Connection` class:

- `url`: jdbc:sybase:Tds:uranium.doc.ic.ac.uk:4100/uscensus1990
- `username`: guest
- `password`: guestuser

References

[Cen90] U.S. Census. U.S. gazetteer place and zipcode files. Technical report, U.S. Census, 1990.

[Gun01] R. Gun. Implementing an RDBMS inside a jdbc driver. Technical report, Dept. of Computing, Imperial College London, 2001.