C. Edward Porter

Spring 2016

INFM 600

**Data Manipulation and Database Construction**

To combine the two datasets, a little manual manipulation was required. Both datasets contained a similar column “Agency”, in DC Purchase Orders - 2013 (DC Purchase Orders, 2013b) and “Agency\_Full\_Name”, in DC Agency Performance Data – 2013 (DC Agency Performance Data, 2013b). However, the entries within these rows were not identical.

First, we copied the unique values out of each dataset by importing both datasets into Excel via the “Get External Data” command and filtering out duplicates. Then, we matched the unique values from one to the other. Finally, we manually replaced all the “Agency\_Full\_Name” values in the Performance Data dataset via find/replace in the csv file. These matched agencies are contained in Table 1.

Both data sets also needed a primary key row.  Both had an “OBJECTID” column; however, the DC Agency Performance Data (2013) row was not unique, as the numbering appeared to restart multiple times. These were replaced with sequential values in Excel.

With both datasets prepared, each was run through a CSV to SQL conversion process by using this site:<http://www.convertcsv.com/csv-to-sql.htm>. These SQL statements were then run on a locally hosted instance of MariaDB, thus creating the DCDATADB.

Finally, to formally link the two tables, we created a third table—the “agencies” table. This table contained all the unique agency names extract earlier, along with their agency acronym, if applicable, and a unique agency\_ID functioning as a primary key. This table was also constructed via creation as a CSV file in Excel, followed by CSV to SQL, and finally running the SQL query on the locally hosted MariaDB instance.

With this in place, a new column was introduced to both the purchase\_orders table and the performance\_metrics table in DCDATA called agency\_ID. These were made to be foreign\_keys with the agency\_ID from the agencies table. Finally, we ran the following SQL queries for the two tables to fill in this agency\_ID column:

UPDATE performance\_metrics SET performance\_metrics.agency\_ID=

  (SELECT agency\_ID FROM agencies where performance\_metrics.Agency\_Full\_Name=agencies.Agency\_Full\_Name)

UPDATE purchase\_orders SET purchase\_orders.agency\_ID=

  (SELECT agency\_ID FROM agencies where purchase\_orders.Agency=agencies.Agency\_Full\_Name)

With the database constructed, we performed a data dump to produce the SQL file hosted here.

Word Count: 350

|  |  |
| --- | --- |
| **Agencies from Purchase Order** | **Agencies from Performance Metrics** |
| ALCOHOLIC BEVERAGE REGULATION ADMINISTRATION | Alcoholic Beverage Regulation Administration |
| BD OF REAL PROPERTY ASSESSMENT & APPEALS |  |
| BOARD OF ELECTIONS | Elections, Board of |
| CHILD AND FAMILY SERVICES | Child and Family Services Agency |
| CITY ADMINISTRATOR / DEPUTY MAYOR | City Administrator, Office of the |
| COMM OF JUDICIAL DISABILITIES & TENURE | Judicial Disabilities and Tenure, Commission on |
| COMMISSION ON ARTS & HUMANITIES | Arts and Humanities, Commission on the |
| COUNCIL OF THE DISTRICT OF COLUMBIA |  |
| CRIMINAL JUSTICE COORDINATING COUNCIL | Criminal Justice Coordinating Council |
| D.C HEALTH BENEFIT EXCHANGE AUTHORITY |  |
| D.C. DEPARTMENT OF HUMAN RESOURCES | Human Resources, Department of |
| D.C. SENTENCING & CRIM. CODE REV. COMM. | Sentencing and Criminal Code Revision Commission, District of Columbia |
| DC BD OF ETHICS AND GOVT ACCOUNTABILITY | District of Columbia Board of Ethics and Government Accountability |
| DC NATIONAL GUARD | National Guard, District of Columbia |
| DC PUBLIC LIBRARY | Public Library, District of Columbia |
| DC STATE BOARD OF EDUCATION |  |
| DEPART OF INSURANCE,SECURITIES & BANKING | Insurance,Securities and Banking, Department of |
| DEPART OF YOUTH REHABILITATION SERVICES | Youth Rehabilitation Services, Department of |
| DEPARTMENT OF CORRECTIONS | Corrections, Department of |
| DEPARTMENT OF EMPLOYMENT SERVICES | Employment Services, Department of |
| DEPARTMENT OF FORENSICS SCIENCES |  |
| DEPARTMENT OF GENERAL SERVICES | Department of General Services |
| DEPARTMENT OF HEALTH | Health, Department of |
| DEPARTMENT OF HEALTH CARE FINANCE | Health Care Finance, Department of |
| DEPARTMENT OF HUMAN SERVICES | Human Services, Department of |
| DEPARTMENT OF MENTAL HEALTH |  |
| DEPARTMENT OF MOTOR VEHICLES | Motor Vehicles, Department of |
| DEPARTMENT OF PARKS AND RECREATION | Parks and Recreation, Department of |
| DEPARTMENT OF PUBLIC WORKS | Public Works, Department of |
| DEPARTMENT OF TRANSPORTATION | Transportation, District Department of |
| DEPARTMENT ON DISABILITY SERVICES | Disability Services, Department of |
| DEPT OF SMALL & LOCAL BUSINESS DEVELOPMT | Small and Local Business Development, Department of |
| DEPT. OF CONSUMER AND REGULATORY AFFAIRS | Consumer and Regulatory Affairs, Department of |
| DEPT. OF HOUSING AND COMM. DEVELOPMENT | Housing and Community Development, Department of |
| DEPUTY MAYOR FOR ECONOMIC DEVELOPMENT | Deputy Mayor for Planning and Economic Development, Office of the |
| DEPUTY MAYOR FOR EDUCATION | Deputy Mayor for Education, Office of the |
| DEPUTY MAYOR FOR HEALTH & HUMAN SERVICES | Deputy Mayor for Health and Human Svcs, Office of |
| DEPUTY MAYOR FOR PUBLIC SAFETY & JUSTICE | Deputy Mayor of Public Safety and Justice |
| DISTRICT DEPARTMENT OF THE ENVIRONMENT | Environment, District Department of the |
| DISTRICT OF COLUMBIA PUBLIC SCHOOLS | Public Schools, District of Columbia |
| EMPLOYEESCOMPENSATION FUND |  |
| EQUIPMENT LEASE - CAPITAL |  |
| FIRE AND EMERGENCY MEDICAL SERVICES | Fire and Emergency Medical Services |
| FORENSIC LABORATORY TECHNICIAN TRAIN PRG |  |
| HOMELAND SECURITY GRANTS |  |
| HOMELAND SECURITY/EMERGENCY MANAGEMENT | Homeland Security and Emergency Management Agency |
| INAUGURAL EXPENSES |  |
| JUDICIAL NOMINATION COMMISSION | Judicial Nomination Commission |
| LOTTERY & CHARITABLE GAMES CONTROL BOARD |  |
| MEDICAL LIABILITY CAPTIVE INS AGENCY |  |
| METROPOLITAN POLICE DEPARTMENT | Metropolitan Police Department |
| OFFICE FOR NON-PUBLIC TUITION |  |
| OFFICE OF ADMINISTRATIVE HEARINGS | Administrative Hearings, Office of |
| OFFICE OF CABLE TV | Cable Television, Office of |
| OFFICE OF CAMPAIGN FINANCE | Campaign Finance, Office of |
| OFFICE OF CHIEF FINANCIAL OFFICER |  |
| OFFICE OF CHIEF TECHNOLOGY OFFICER | Chief Technology Officer, Office of the |
| OFFICE OF CONTRACTING AND PROCUREMENT | Contracting and Procurement, Office of |
| OFFICE OF DISABILITY RIGHTS | Disability Rights, Office of |
| OFFICE OF EMPLOYEE APPEALS | Employee Appeals, Office of |
| OFFICE OF FINANCE & RESOURCE MGMT |  |
| OFFICE OF HUMAN RIGHTS | Human Rights, Office of |
| OFFICE OF LATINO AFFAIRS | Latino Affairs, Office on |
| OFFICE OF MOTION PICTURES & TELEVISION | Motion Pictures and Television, Office of |
| OFFICE OF MUNICIPAL PLANNING | Planning, Office of |
| OFFICE OF PEOPLES COUNSEL | Peoples Counsel, Office of |
| OFFICE OF POLICE COMPLAINTS | Police Complaints, Office of |
| OFFICE OF RISK MANAGEMENT | Risk Management, Office of |
| OFFICE OF TENANT ADVOCATE | Tenant Advocate, Office of the |
| OFFICE OF THE ATTORNEY GENERAL | Attorney General, Office of the |
| OFFICE OF THE CHIEF MEDICAL EXAMINER | Chief Medical Examiner, Office of the |
| OFFICE OF THE D.C. AUDITOR | Auditor, Office of the District of Columbia |
| OFFICE OF THE INSPECTOR GENERAL | Inspector General, Office of the |
| OFFICE OF THE MAYOR | Mayor, Office of the |
| OFFICE OF THE SECRETARY | Secretary, Office of the |
| OFFICE OF UNIFIED COMMUNICATIONS | Unified Communications, Office of |
| OFFICE OF ZONING | Zoning, Office of |
| OFFICE ON AGING | Aging, District of Columbia Office on |
| OFFICE ON ASIAN/PACIFIC AFFAIRS | Asian and Pacific Islander Affairs, Office of |
| PUBLIC EMPLOYEE RELATIONS BOARD | Public Employee Relations Board |
| PUBLIC SERVICE COMMISSION | Public Service Commission |
| SPECIAL EDUCATION TRANSPORTATION |  |
| STATE SUPERINTENDENT OF EDUCATION (OSSE) | Superintendent of Education, Office of the State |
| TAXI CAB COMMISSION | Taxicab Commission, District of Columbia |
|  | Contract Appeals Board |
|  | Behavioral Health, Department of |
|  | Veteran Affairs, Office of |
|  | DC Public Charter School Board |