# Database Manual

# Outline

1. Database Manual
2. Organization
3. Table Definitions
4. Reports and Selectors
5. Macros
6. Archiving

# Organization and Table Definitions

The database is organized using the "Custom" method provided by Access, which allows for user defined categories. The categories are chosen to make it more efficient to find input data, queries, or macros than default organizations.

### Estimate\_Masters

These are the master tables for all of the estimates. Any data request or final table eventually links back to these tables. Internal to the database, these are the gold-standard for each geography.

20XXCTFEstimatefinal2 - This bizarrely named table holds the special district final estimates and gets created each year with a new year. This could probably be changed in future vintages because it's not necessary to rename with the year, but so far I haven't.

BGU\_Estimates\_2010 - While the District estimates are in the CTFEstimatesfinal2 table, this is where they get generated. Each BGU (Basic Geographic Unit) is in this table and the final estimates from this table are used to generate and populate the estimates tables in Oracle used to distribute CTF dollars. BGUs break the state up into unique geographies based on county, municipality, and special district boundaries. The table contains estimates from 2010 to the most recent vintage.

CityandCounty\_AnnualConstructionData\_2010 - This table is created using construction data input from the RCS and the Census Bureau. It is also the table where we set which housing data source to use for each municipality. This is the best table to use as a start for investigating issues with housing inputs, but doesn't hold any final estimate information. The table contains data from 2010 to the most recent vintage.

CityandCounty\_Estimate2010\_Master - This is likely the most useful of all the Masters because it contains the final estimates for all municipalities and counties, although it only shows housing unit method components (not births, deaths, and migration). This table contains out final estimates for all housing data and population data on municipalities.

County\_Estimates\_Master\_2010 - This is the gold standard table. It contains components and final estimates for all counties and the state. Since the counties are the gold standard for our estimates (everything must sum up to counties, and we generate the state by summing the counties), this table is incredibly important.

### County Estimates Queries

This section contains all the queries to generate the County\_Estimates\_Master\_2010 table. A run-down of these queries is available in the section on County Population estimates.

### Annual Construction Data Queries

This section contains all queries used to build the CityandCounty\_AnnualConstructionData\_2010 master table. These queries take construction data inputs and place them into the proper columns in the table, but they DO NOT change the source columns, that is done manually.

### City and County Estimates Queries

This section contains all queries to generate the CityandCounty\_Estimates2010\_Master table. More information on these queries is available in the section on City and County estimates.

### BGU Estimates Queries

This section contains queries to create the BGU\_Estimates\_2010 table. It gets larger every year since we have to add another full set of queries per year. More information is available in the section on BGU estimates.

### BGU New District Data

This section contains tables and queries to assist in appending data to the proper tables to add districts to the CTF process. You'll only need this once. See the BGU estimates sections for more information.

CTFDistrict\_NewBGUAppends - This table contains the 2010 BGU data for all new districts. This data is generated by the GIS person when we are notified of a new district.

CTFDistricts\_ExistingBGUUpdates - This data

### CTF Estimates Final Queries

This section contains queries to create the 20XXCTFEstimatesfinal2 table. More information is available in the section on CTF estimates.

### CTF IT Table Queries

This section contains the queries to generate the tables we send to Oracle for the CTF distribution. They are updated each year to link to the current vintage's final year then uploaded by Mark Krudwig. More information is available in the Section on CTF estimates.

### CTF Tables for IT

These tables are what we use for the CTF distribution upload.

DLG\_CTF\_BGU\_COUNTY\_TEMP - This is an Oracle 'holding' table containing the most recent (v2015 at writing) county estimates. This is used as a control for evaluating the BGU table.

DLG\_CTF\_BGU\_MUNI\_TEMP - This is an Oracle 'holding' table for the most recent municipal estimates (v2015 at writing). They are a control for evaluating the BGU table.

DLG\_CTF\_BGU\_SDIST\_TEMP - This is an Oracle 'holding' table for the most recent special district estimates (v2015 at writing). They are a control for evaluating the BGU table.

DLG\_CTF\_BGU\_TEMP - This is an Oracle 'holding' table for the most recent BGU estimates (v2015 at writing).

IT\_County - This is the Access table that we use to populate the Oracle 'holding table' for counties.

IT\_Muni - This is the Access table that we use to populate the Oracle 'holding table' for municipalities.

IT\_district - This is the Access table that we use to populate the Oracle 'holding table' for special districts.

IT\_BGU - This is the Access table that we use to populate the Oracle 'holding table' for BGUs.

### Macros

This is a section devoted to the various macros we use. These don't really need to be run, I use them for final runs and when I'm only debugging a single section or input data, but for everything except BGU I test queries by hand. I've removed many of the table delete commands, but if you're messing around, make sure to not use them. At least make sure to copy the database to the desktop.

ReviewCopyExportMacro - this macro is used to export copies of the data to the estimates\_review\_app directory at J:. I use this app to evaluate each place's estimate once I'm finalizing.

RunAllUpdaters - This one, as it's name suggests, runs the entire estimates process in one fell swoop. Don't run it if you don't mean it. I use it for final estimates or when we get input changes.

StandardizationPull - This pulls nicely formatted and named data from the database Masters. I don't currently use it, but keep it in case I ever want that data.

Step01\_CountyEstimatesMaster\_Updater - Runs all of the County Estimates Queries, must be run before other estimates queries.

Step02\_AnnualConstructionData\_Updater - Runs all queries to create the Annual Construction Data master we need to make City and County estimates. This one is my most used Macro since we're usually not changing queries, we're updating input data.

Step03\_CityandCountyEstimatesMaster\_Updater - Runs all of the City and County Queries.

Step04-09\_BGUEstimates20XX\_RunAll - These Macros run each individual year for the BGU model. We have to run it in year order because it requires the previous year to run to have inputs for the next year.

Step09\_BGUEstimatesOverlapped\_RunAll - This runs all of the overlapped queries to create the important BGU Overlapped table. This table has extra rows for overlapped areas to properly allocate the

Step10\_CTFEstimates\_Final\_RunAll - This marco runs all the final estimates table for the special districts, this has to be run last.

### Estimates Input Data

This section contains all of the input data we need to run the model. It's mostly linked Excel tables and Access tables. If you want to see everything that is linked externally, use the 'Linked Table Manager'. That will provide you file locations as well. Use this to find the files if you need them, to make sure we're linking properly.

BGU\_Estimates\_2010\_Overwrites: This table contains each BGU and any data that we need to force to be a certain number. This is often done as a result of challenges. Overwrites change the data in the model directly, although for BGU purposes, they don't necessarily replace the number directly.

CensusCityandCountyMaster\_2010: This is a linked table to a sheet that I add to each year featuring the latest Census data for subcounty areas. This table is where I pull in housing and population data from the Census for some of our reports (Table 2c in particular).

CensusCityandCountyMaster\_2010ft: This is a static version locked on v2013. I probably don't need it. I used it for evaluation I think.

CensusCityandCountyPlaceSums: Query that sums up parts to get place sums for evaluation purposes. Not technically an input, but it is based on inputs.

CensusHousingUnit\_Master\_2010: This table is used to populate the Annual Construction Master Census Permit columns. This table is more important than the population master from the Census.

CityandCounty\_BoundaryChange\_2010: This table contains housing and population changes for each place in the state due to annexations, deannexations, or other boundary changes. All annexations are entered here.

CityandCounty\_Estimates2010\_Overwrites: Similar to the BGU table, we keep any hard coded data for places in this table. These overwrites will populate directly in the Master with the exception of calculated values like household size and vacancy/occupancy measures.

CityandCountyGQbase\_2010: This table is the base for all of the group quarters estimates we publish. I change this table when we get updated to the 2010 data, but otherwise pretty rarely. It is essential though as we don't track totals, only change, so without it we can't generate GQ estimates.

DOLAMigrationAdjustments\_2010: This table is a bit of a relic from the days when we didn't trust the Census. We'd place migration data adjustments here. We used it that way for modeling our 2010 base from the April 1 data, but haven't used it since. So we need it, but we don't really edit it.

DPHEBirthData: This holds every birth data point we gather and is updated yearly.

DPHEDeathData: This holds all the death data, pre tabluated to the proper timeframe, and is updated yearly.

LocalBuildingPermitData\_2010: This is linked to the spreadsheet that contains the building permit data submitted through the residential construction survey.

LocalCertificateofOccupancyData\_2010: This is linked to the spreadsheet that contains the certificate of occupancy data submitted through the residential construction survey.

LocalDemolitionPermitDaata\_2010: This is linked to the spreadsheet that contains the demolition permit data submitted through the residential construction survey.

LocalMobileHomeData\_2010: This is linked to the spreadsheet that contains the mobile housing data submitted through the residential construction survey.

QCEWEmploymentData: This is a table we use to populate Table 1 for the counties during the Challenge period. We use the Q2 employment counts as symptomatic data in that table.

### Estimates Evaluation Queries

I don't really use these anymore, since I usually use my app, but these can be updated to look at percent and absolute change in different variables for evaluation purposes.

### Previous Vintage Tables for Evaluation

These are copies of the Masters from the previous vintage. I use these to populate the review app output mentioned in the Macro section. So we update these to get the revisions.

### Contact Tables

CityandCounty\_Contacts: I update this contact table at the same time as the MailChimp Contacts so that we attempt to keep it up to date. The only time I don't is when they ask for that.

District\_Contacts: This is a relic table we don't use anymore.

DLG\_LGCONTACT: This is a linked Oracle table with all local government contact information in it. It's used in the following queries along with other entity information tables.

LGISSurveyContactInfo\_CaC: I use this table to update the MailChimp and CityandCounty\_Contacts lists when something bounces. If this is the same contact, log into LGIS and move one contact up the hierarchy.

LGISSurveyContactInfoSD: This is the main list I use for CTF and RCS purposes at the Special District level. This list shows all district administrators for CTF districts.

SurveyCity&CountyContact: Don't use this, not sure how old it is.

### Entity Information

CENSUS\_COGEO: This is likely unnecessary now that we have the new database, but I used to use it to get different geography data. It's a linked Oracle table.

DLG\_CTF\_ENTITY: This is a linked Oracle table with data on CTF entities and whether or not they are eligible for distribution. I use this to help make the contact list.

DLG\_LG2CNTY: This is a linked ORacle table that crosswalks LGID numbers from LGIS to the corresponding county.

DLG\_LGBASIC: This is a linked Oracle table with a few basic data points on each local government that is used to link and filter queries.

DLG\_LGSTATUS: This linked Oracle table lets you know the status of an entity, it's been used in the past as a check against the City and County data, but it really wasn't needed due to the CTF table.

DLG\_V\_CTF\_CERT\_FOR\_POP\_CALC: This is a view of entities certified to be given funds. This is the table Mark uses when running his estimates. We are hoping to build a check system on our side that would help us eliminate failures when we push to Mark Krudwig. See the Email string in PDFs here: J:CTF Checks Round 1-3

regfips: This table is a crosswalk for places to planning and management regions.

### Review Reports

This section contains tables we use for CTF and RCS as well as queries I use for building the review app. The tables labeled DRAFT are turned into finals when I finish the year, I didn't in v2015 because I made a look-up based site in R, but I'm likely going back to this in v2016.

County\_estimates\_review\_current to muni\_estimates\_review\_previous: These tables are used to export to csv for the estimates review app. They are built from current and previous masters.

Table1\_Draft\_2015: This is a table we provide with Symptomatic Data. We'd moved to a lookup based format last year, but may be going back for simplicity's sake.

Table2\_Draft\_2015: This is the most detailed table we produce for each county and municipality. This is linked to a huge query to pull all this data in. It was previously Table2c, which is what Elizabeth calls it.

Table2\_rcs\_2015: This shows final data from the previous vintage and is used for the RCS.

Table2mc\_Draft\_2015: This is the same as Table2\_Draft\_2015, but has data for only multi-county places. This is needed due to the query structure we use.

Table2mc\_rcs\_2015: This is the same story as above.

Table3\_Draft\_2015: This table contains all the details of the population estimates for each place, but no symptomatic or input data.

Table3\_rcs\_2015: The same as the table above, but with the previous vintage data and used for the RCS.

Table4\_Draft\_2015: This table shows all the special district estimates. This is what I give to areas that want a table and that we post online.

Table4\_rcs\_2015: Same as above, but with previous data.

TableA\_rcs\_2015: This table contains the construction data we used in the previous vintage. We give this out for the RCS.

### Report Selectors

These are the queries that the reports in the above sections rely on. They get updated every year. Their names tell you where they link to, but you can use the reports to figure out which is which.

### Data Requests

These are tables I use to help respond to data requests. They are either frequently requested or annoying to pull, so if I do it once I can do it again.

CountyMuniPopHousing: This table contains total population and housing units from 2010 onward.

CountyNetMigration: This table contains net migration for the vintage period without GQ change.

countyplacehousingunits: This table has only total housing units for all county-place combinations.

DirectDistributionPull: This is the query that generates the table we give to Cynthia each August for the Direct Distribution.

LGID\_lookup: This is a quick table to help find LGID values for the areas we estimate for.

ParksandRecCTFPop-ParksandRecDistricts: We sometimes get asked about parks and rec districts (NOT including metro districts), it's happened twice, so I built these since it's a pain to pull.

TotalPopulationbyCounty: Exactly what it says.

TotalPopulationbyMuni: This has total populations for municipalities, it sums together the multi-county areas, so Aurora has one line.

VacancyRates: This is a query that shows every county-places vacancy rate. I use this for evaluation purposes mostly.

### Final Tables for the Annual Meeting

All of these queries are used to generate the data to update the Final Tables in the Tables directory. I typically use a vlookup to ensure we get the right data into the sheet.

CountyTable1: This one corresponds to the Table 1 we make for all of the counties, it also has the April 1 and 2010 Base numbers.

County\_FinalTable: This one just has county populations for each year, no bases.

CountyMasterwTableRegionNums: We used this to make a regional table, but I don't anymore.

CountyMuniHousing: This corresponds to the CountyMuniHousingXXXX table.

demog\_update\_table: This was used previous to update the codemog R package.

muniranked: This updates the muniranked table we produce for the annual meeting.

popbycountyandmuniXXXX: This updates the tables of the same name.

prregiontotals: This is used for region totals, we used to need this for the old online databases in Access.

RegionSumsnonPM: This was used for the different region definitions we displayed online, like Western Slope. It's not needed now since we use a view in Postgres to do that.

## Archiving the Database

Toward the end of March through April it is time to think about archiving the old Database. This process has varied over time, but I have a procedure that seems to work well for me. We used to archive the DB and ALL of the files associated, basically the whole directory. That was a bit bloaty and sort of overkill. I've outlined my process below, it's pretty simple.

#### Archival Process

1. Create a copy of the database on our local machine. I rename it with the vitange name on the end, so for vintage 2014 it's "Estimates 2010-20v2014.accdb".  
2. Convert all of the linked tables (including those from Oracle) to local tables. This makes the database self-contained and unable to be changed based on new data in the input files for the vintage. I highlight all of the linked tables in each section > right-click > then select "Convert to Local Table."  
3. Save the local database copy.  
4. I like to keep the previous vintage in the main Estimates directory as well as archived. So delete the previous vintage database already in the Estimates directory and then replace it with a copy of the one you just made. BEFORE you do that, double check that the one you're going to delete is archived.  
5. Right-click on the local copy you've been using and compress the database into a zip file. If you want to add other files to that, you can. I did in 2013, but I'm not sure what purpose I thought that would serve. Rename the zipped file "VintageXXXX" based on what vintage you're archiving.  
6. Place the new zip file into the Archive sub-directory of Estimates.   
7. You can delete the local copy.  
8. Send out an email to the staff to let them know what you've done and where to find the data. They'll appreciate it.  
9. Continue building new vintage in the DRAFT folder until you're ready then move it to the main directory.