

Census Data Available at CMAP

- Xiaohong Zhang

Chicago Metropolitan Agency for Planning

Overview

- Why build our own collection
- What is available
- How do we collect the data and add values to it
- Data aggregation project
- How to access the data
- Question and answers

Why build our own collection?

- Historical data
- Large volume of data
- Block/Block Group level of data

Geography Filter Options

030 - Division (9)
040 - State (57)
050 - County (6,219)
060 - County Subdivision (44,378)
067 - Subminor Civil Division within
County within State (145)
070 - Place/Remainder (80,037)
080 - Census Tract within
Place/Remainder (288,208)
085 - Urban/Rural within Census
Tract within
Place/Remainder (191,731)
090 - Urban/Rural within Block
Group within Census Tract within
Place/Remainder (395,265)
140 - Census Tract (98,745)
144 - AIA/ANA/HHL within Census
Tract (2,746)
150 - Block Group within Census
Tract (223,351)
154 - AIA/ANA/HHL within Block
Group within Census Tract (2,091)

Why build our own collection?

- Historical data
- Large volume of data
- Block/Block Group level of data

Geography Filter Options

+ Geographic Type

- Summary Level

- 010 - United States (1)
- 020 - Region (4)
- 030 - Division (9)
- 040 - State (53)
- 050 - County (4,821)
- 060 - County Subdivision (41,526)
- 067 - Subminor Civil Division within County within State (145)
- 070 - Place/Remainder (69,892)
- 080 - Census Tract within Place/Remainder (210,809)
- 091 - Block Group within Census Tract within Place/Remainder (362,701)
- 100 - Block (143,337)
- 140 - Census Tract (78,824)
- 144 - AMANMNL within Census Tract (2,587)

Why build our own collection?

- Other value added services:
Data transformation
(DBF files and GIS files)

Data aggregation (Chicago Community Area (CCA) data)

Internal customized report
(Community profiles)

Load Geography List

Select ONE or MORE Standard Geographies

Municipality

County

Region

Select ONE Additional Geography (Optional)

Community Area

Census Tract

Block Group

Clear Form Selection

OK

Clear Values on Spreadsheet

Exit Tool

Note :

1. Clear Values on the spreadsheet before populate new data.
2. Given that some Census boundaries have changed since 2000, you might not be able to compare info from 2000 to 2010. For those that did not change, you will get a message telling you so. For those with boundary changes, you will not see information from Census 2000, or the changes from 2010 to 2000.
3. For CCAs, only data from Census 2010 has been populated. Other years of Data would be available in the future.

What is available at CMAP?

Data/Boundary Files

Census 1960 (Tract, Municipality)

Census 1970 (Tract)

Census 1980 (Tract)

Census 1990

- SF1 (County, Municipality, Tract, BG, Block)

- SF3 (County, Municipality, Tract, BG)

Census 2000

- SF1 (County, Municipality, Tract, Block, CCA)

- SF3 (County, Municipality, Tract, BG, CCA)

Census 2010

- County, Municipality, Tract, BG, Block, CCA

ACS 2005-2009 (County, Municipality, Tract, BG, CCA)

ACS 2006-2010 (County, Municipality, Tract, BG)

How do we collect data?

National Historical Geographical
Information System

Census FTP Site

Aggregate in-house



A screenshot of a directory listing from the U.S. Census Bureau website. The header shows 'U.S. Census Bureau' in a dark blue bar. Below it is a table with two columns: 'Name' and 'Last modified'. The table lists various folders representing different census data releases, including AHS, Econ2001 And Earlier, and multiple ACS (American Community Survey) releases from 2002 to 2008, with sub-folders for different time periods (e.g., 1yr, 3yr, 5yr).

<u>Name</u>	<u>Last modified</u>
 AHS/	07-Jul-2006 09:56
 Econ2001 And Earlier/	27-Apr-2011 08:40
 acs/	30-Dec-2010 10:05
 acs2002/	23-Sep-2008 08:25
 acs2003/	06-Oct-2004 07:07
 acs2004/	02-Feb-2006 10:40
 acs2005/	08-Feb-2011 09:06
 acs2005 2007 3yr/	09-Dec-2008 08:56
 acs2005 2009 5yr/	12-Dec-2010 19:28
 acs2006/	02-Sep-2010 11:55
 acs2006 2008 3yr/	27-Oct-2009 13:38
 acs2007 1yr/	26-Sep-2008 09:32
 acs2007 3yr/	09-Dec-2008 08:56
 acs2007 2009 3yr/	06-Jan-2011 14:51
 acs2008 1yr/	03-Nov-2009 16:42

Chicago Community Area data aggregation (Overview)

Project Rationale

- CMAP
- Accommodate forecasting criteria and methodology

Team

David Clark, Jon Hallas, Kermit Wies, Noel Peterson, Dave Morck and Xiaohong Zhang

Scope of Work

- Census 2010, most tables

Chicago Community Area data aggregation (Overview)

Major Steps

- Boundary change analysis
- Relationship analysis
- Tract issues
- Allocating ratios
- Data analysis
- Look-up tables
- Workflow
- Aggregation

Chicago Community Area data aggregation (Background)

Chicago Community Area

- Established in the 1920s by UC sociologists
- Aggregations of Census Tracts
- Drawn up with consideration given to:
 - a) settlement, growth and history
 - b) local identification
 - c) the local trade area
 - d) distribution of membership of local institutions
 - e) natural and artificial barriers

Chicago Community Area data aggregation (Background)

Chicago Community Areas boundary history

- Originally defined map had 75 community areas
- Two additions:
 - 1960 Census: O'Hare (#76)
 - 1980 Census: Edgewater (#77)

Chicago Community Area data aggregation (Background)

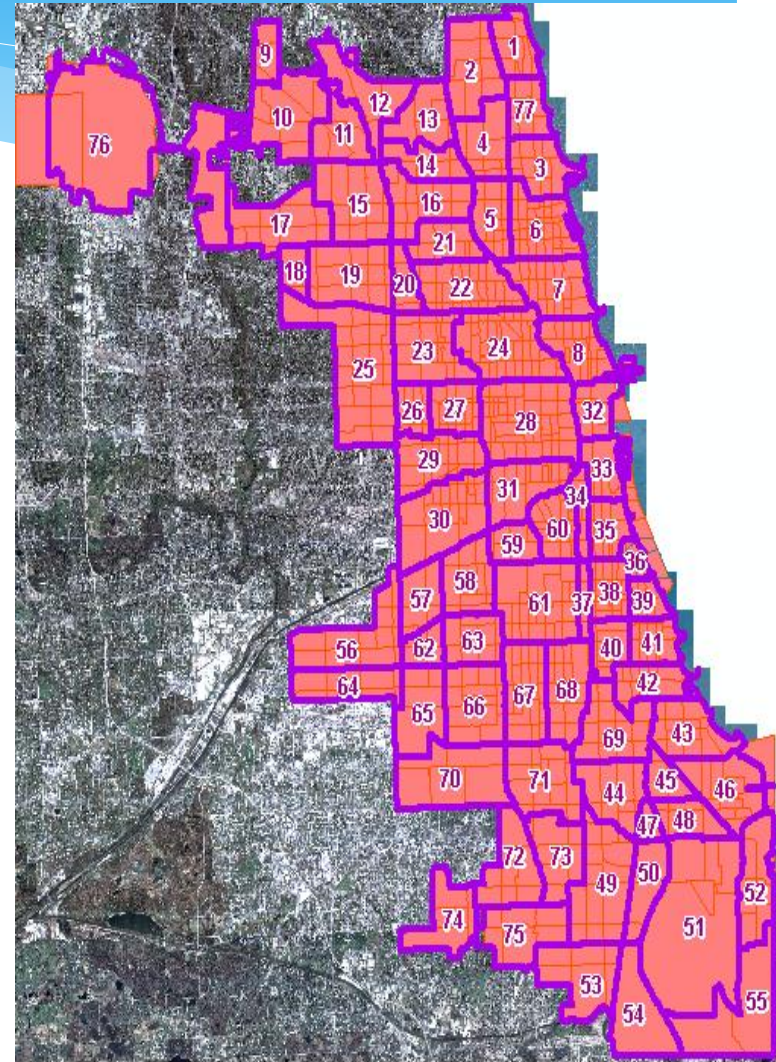
Census Boundary Update

- Program: Participant Statistical Areas Program (PSAP)
- Frequency: Once a decade
- Updates:
 - Revision
 - Merge
 - Split

Chicago Community Area data aggregation (Background)

CCA/Tract Boundary (2000)

- Aggregations of Census tracts
- Nested nicely within CCA boundaries
- Relationship:
First two digits of a 4-digit tract number correspond to the CCA #



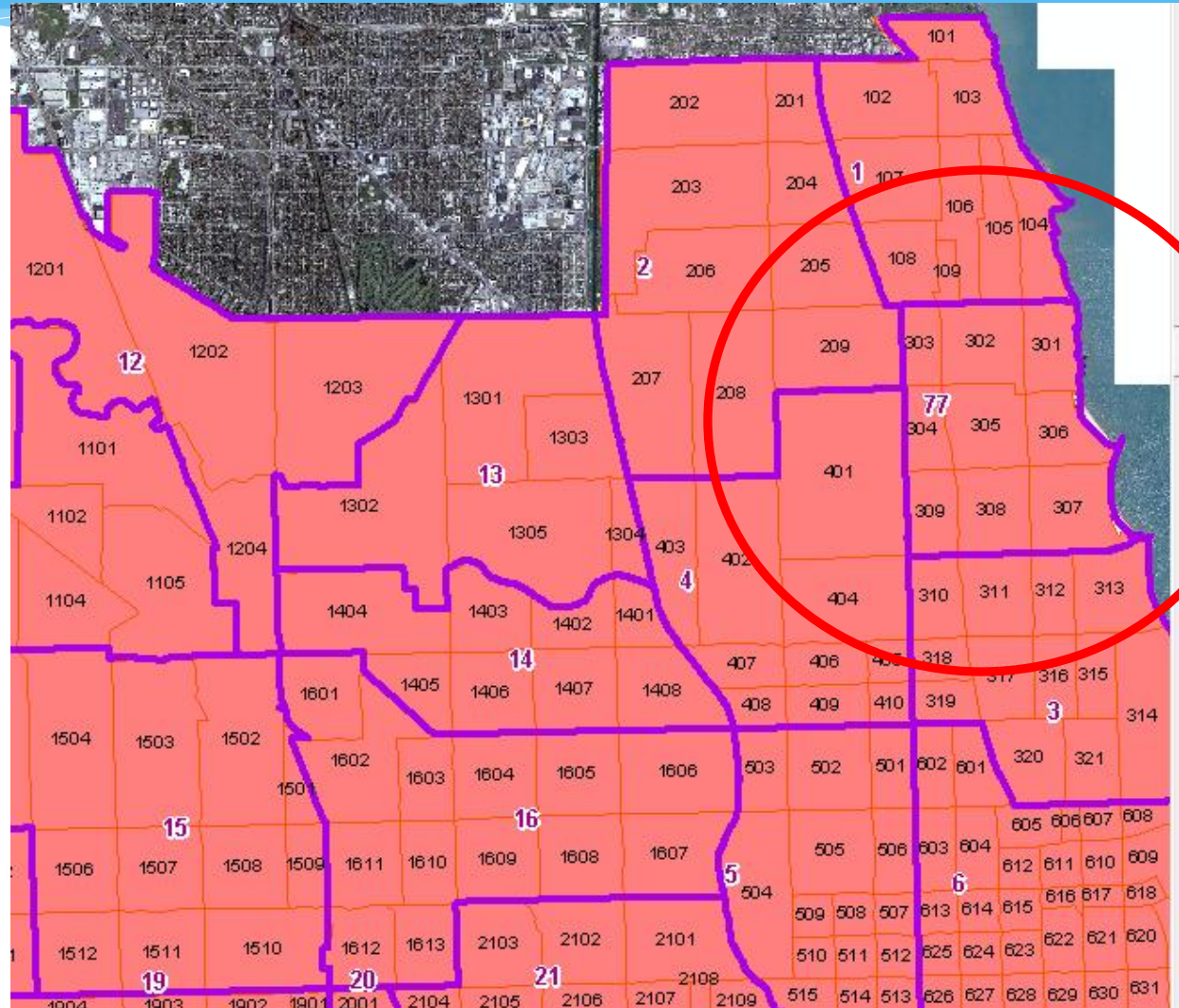
Tract_Short	CCAID	CCA_NAME
0101	1	Rogers Park
0102	1	Rogers Park
0103	1	Rogers Park
0104	1	Rogers Park
0105	1	Rogers Park
0106	1	Rogers Park
0107	1	Rogers Park
0108	1	Rogers Park
0109	1	Rogers Park
0201	2	West Ridge
0202	2	West Ridge
0203	2	West Ridge
0204	2	West Ridge
0205	2	West Ridge
0206	2	West Ridge
0207	2	West Ridge
0208	2	West Ridge
0209	2	West Ridge
0301	77	Edgewater
0302	77	Edgewater
0303	77	Edgewater
0304	77	Edgewater
0305	77	Edgewater
0306	77	Edgewater
0307	77	Edgewater
0308	77	Edgewater
0309	77	Edgewater

Tract_Short	CCAID	CCA_NAME
0310	3	Uptown
0311	3	Uptown
0312	3	Uptown
0313	3	Uptown
0314	3	Uptown
0315	3	Uptown
0316	3	Uptown
0317	3	Uptown
0318	3	Uptown
0319	3	Uptown
0320	3	Uptown
0321	3	Uptown
0401	4	Lincoln Square
0402	4	Lincoln Square
0403	4	Lincoln Square
0404	4	Lincoln Square
0405	4	Lincoln Square
0406	4	Lincoln Square
0407	4	Lincoln Square
0408	4	Lincoln Square
0409	4	Lincoln Square
0410	4	Lincoln Square
0501	5	North Center
0502	5	North Center
0503	5	North Center
0504	5	North Center
0505	5	North Center
0506	5	North Center
0507	5	North Center
0508	5	North Center
0509	5	North Center
0510	5	North Center
0511	5	North Center

Tract_Short	CCAID	CCA_NAME
0512	5	North Center
0513	5	North Center
0514	5	North Center
0515	5	North Center
0601	6	Lake View
0602	6	Lake View
0603	6	Lake View
0604	6	Lake View
0605	6	Lake View
0606	6	Lake View
0607	6	Lake View
0608	6	Lake View
0609	6	Lake View
0610	6	Lake View
0611	6	Lake View
0612	6	Lake View
0613	6	Lake View
0614	6	Lake View
0615	6	Lake View
0616	6	Lake View
0617	6	Lake View
0618	6	Lake View
0619	6	Lake View
0620	6	Lake View
0621	6	Lake View
0622	6	Lake View
0623	6	Lake View
0624	6	Lake View
0625	6	Lake View
0626	6	Lake View
0627	6	Lake View
0628	6	Lake View
0629	6	Lake View

Chicago Community Area data aggregation (Background)

CCA/Tract Boundary
(2000)



Census boundary change (Census 2000-2010)

Census 2010 Tracts for Chicago

Total: 809 tracts

No change: 565

Merge: 112

Split: 132

Census boundary change (Census 2000-2010)

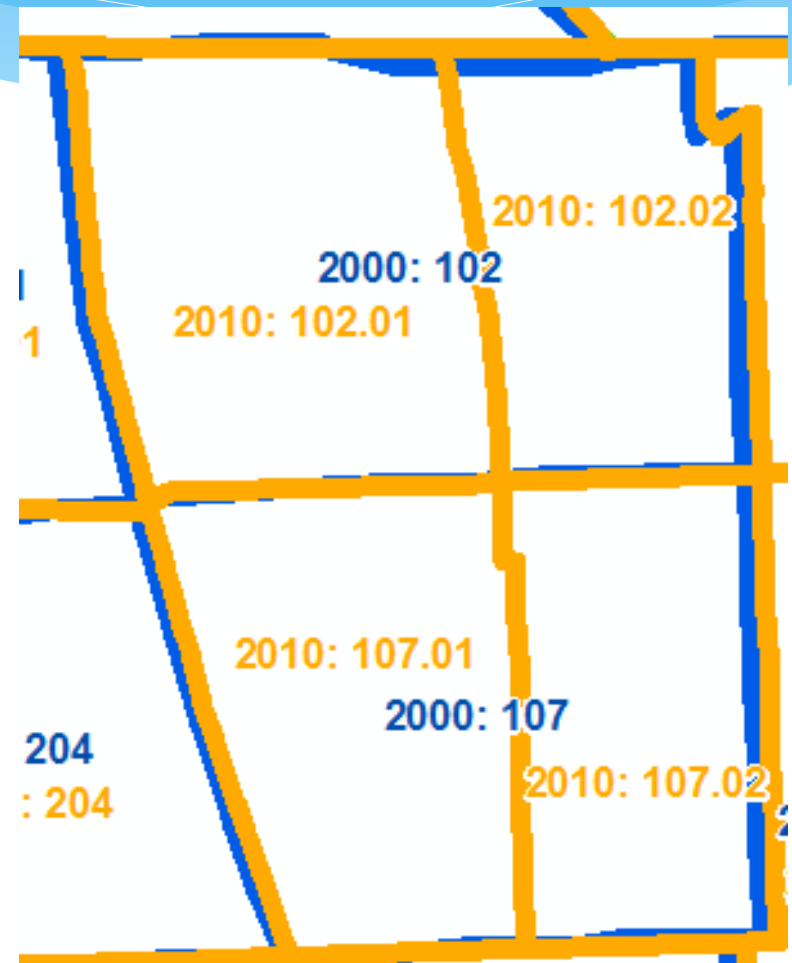
Split:

Census 2000: 102

Census 2010: 102.01 + 102.02

Census 2000: 107

Census 2010: 107.01 + 107.02

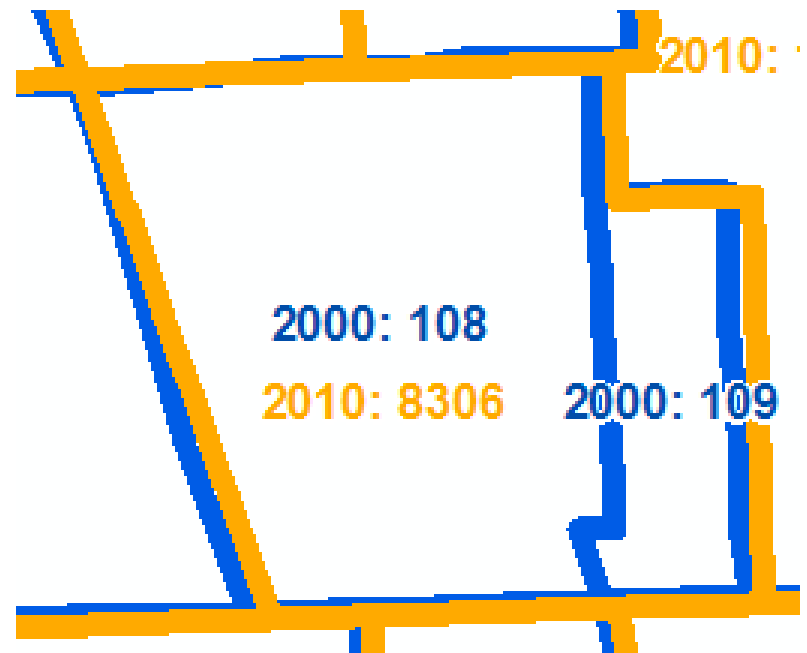


Census boundary change (Census 2000-2010)

Merge:

Census 2000: 108 + 109

Census 2010: 8306



Census boundary change (Census 2000-2010)

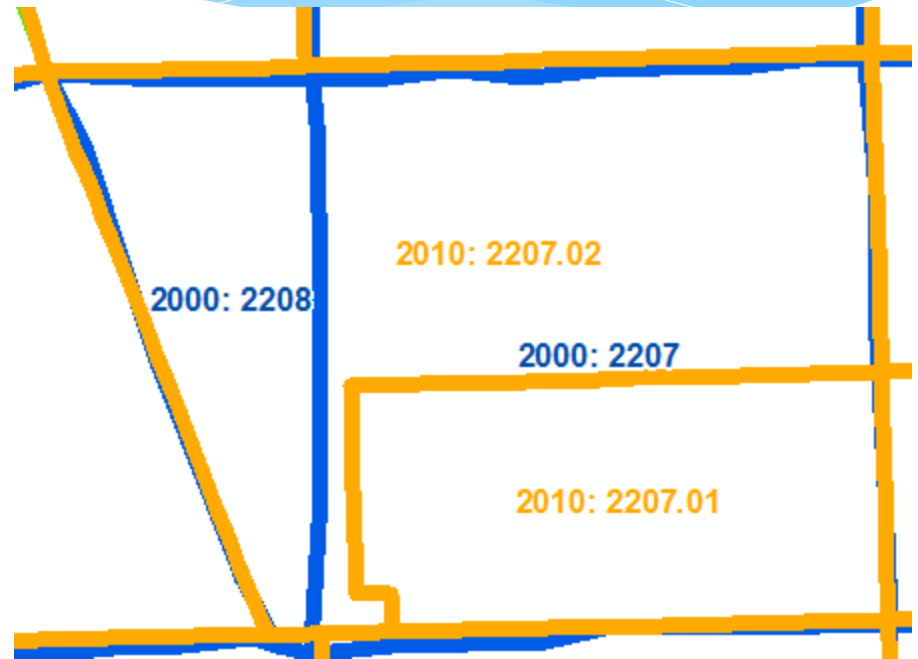
Split + Merge:

Census 2000: 2207, 2208

Census 2010:

Portion of 2207 → 2207.01

Portion of 2207 + 2208 → 2207.02

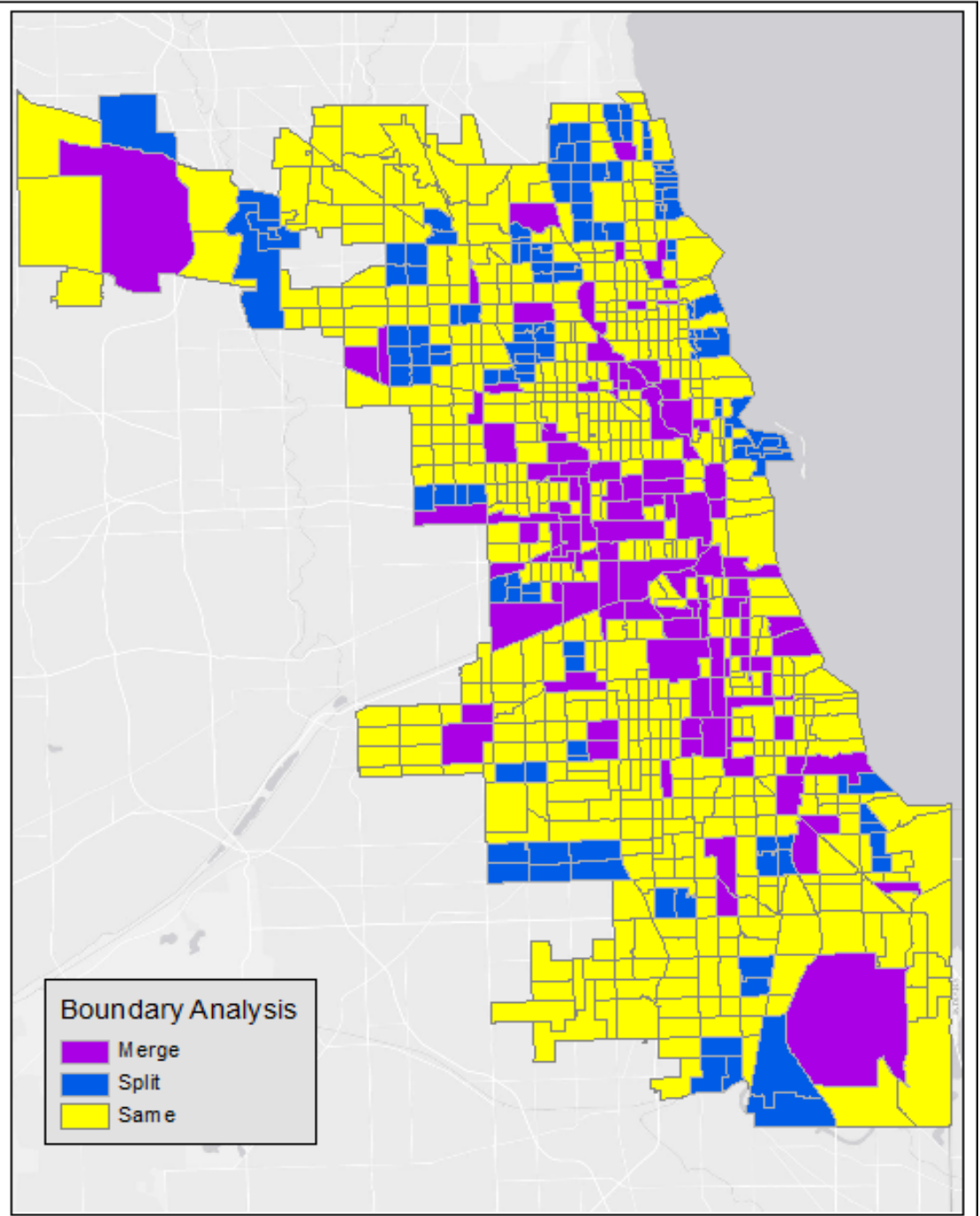


Boundary change (2000-2010)

Purple: merge

Blue: split

Yellow: no change



Census boundary change (Census 2000-2010)

Census Tract Relationship Files

- How 2000 census tracts relate to 2010 tracts
- One record per 2000/2010 census tract spatial set

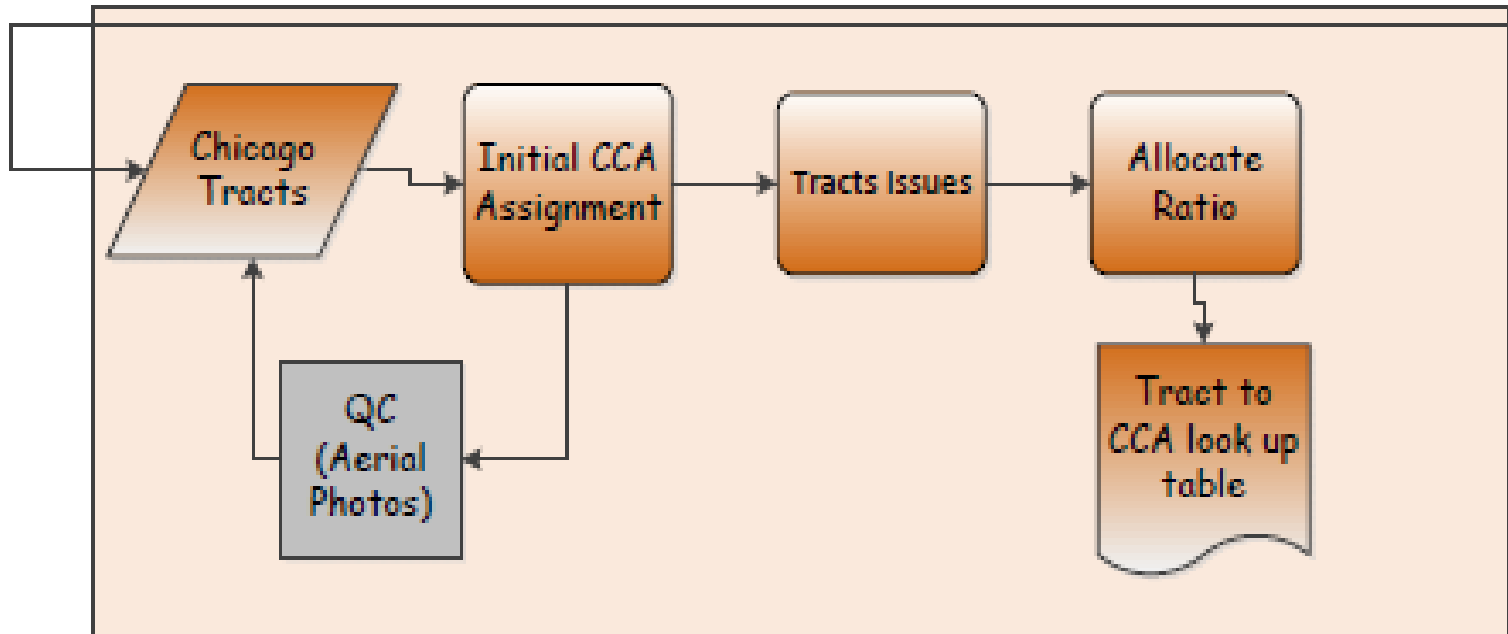
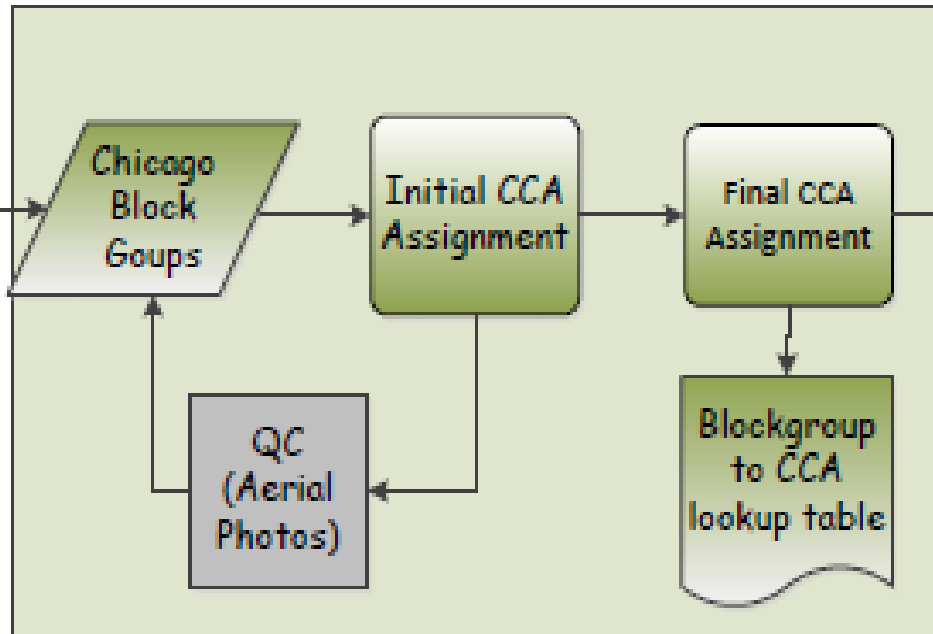
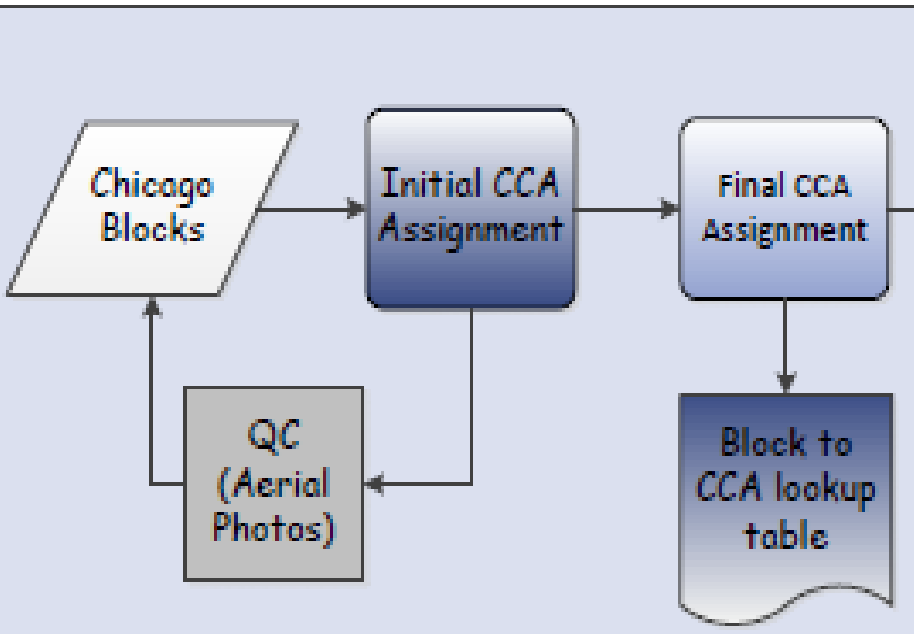
Three sets of files:

1. Census tract relationship file
2. Substantially changed 2000 Census Tract files
3. Substantially changed 2010 Census Tract files

CCA/Census boundary relationship (Census 2010)

Boundary Analysis

- Chicago Census Blocks
- Initial CCA assignment for Census Blocks
- QC and finalize assignment
- Initial BG/Tract relationship
- Identify tract issues
- Ratio and recommendations

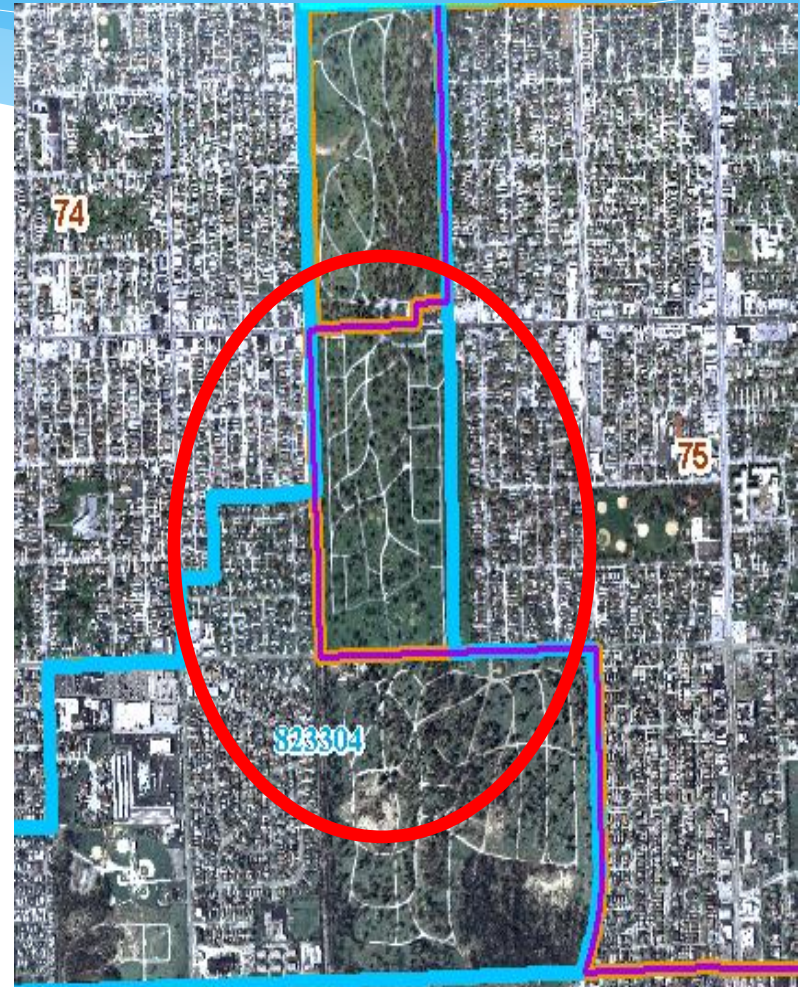


CCA/Census boundary relationship (Census 2010)

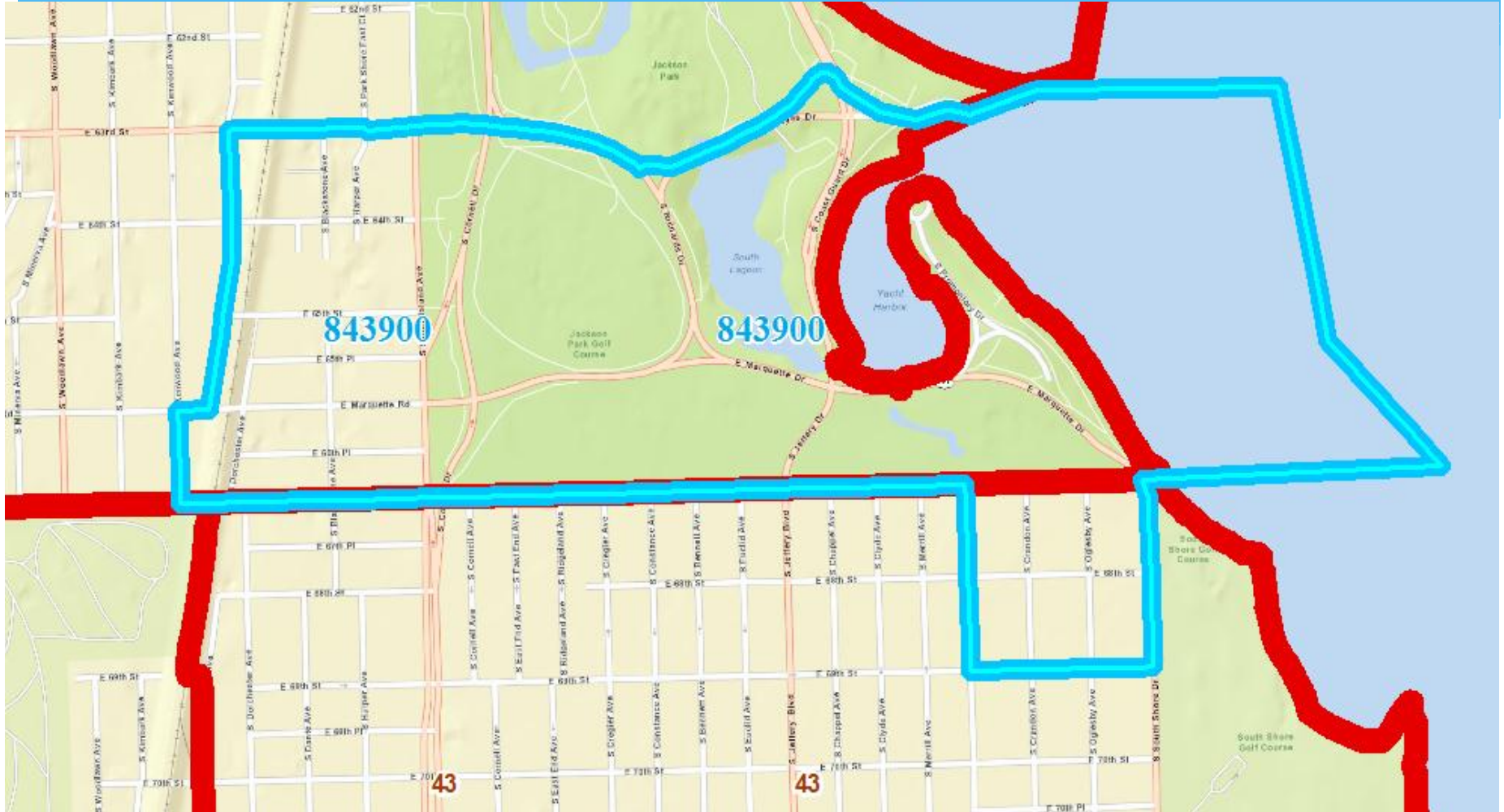
Problematic Tracts

Tract crosses the City of Chicago
Boundary:

- 1) All pop/HH **OUT**
- 2) Insignificant portion of
population **IN**
- 3) Unpopulated tract portion **OUT**
- 4) Significant population both
IN/OUT



CCA/Census boundary relationship (Census 2010)



Population/household allocation

GEOID10	CCA	Pop_Pct	HH_Pct	HU_Pct	PIH_Pct	NIGQ_Pct	IGQ_Pct
17031831000	22	0.624	0.604	0.602	0.624	0.000	0.000
17031831000	24	0.376	0.396	0.398	0.376	0.000	0.000
17031843900	42	0.365	0.305	0.329	0.360	1.000	0.000
17031843900	43	0.635	0.695	0.671	0.640	0.000	0.000
17031081403	8	1.000	1.000	1.000	1.000	0.000	0.000
17031840000	60	1.000	1.000	1.000	1.000	0.000	0.000
17031830900	22	1.000	1.000	1.000	1.000	1.000	0.000
17031020702	2	1.000	1.000	1.000	1.000	1.000	1.000
17031770902	76	0.056	0.077	0.081	0.056	0.000	0.000
17031823304	75	0.002	0.003	0.003	0.002	0.000	0.000
17031821402	53	0.003	0.004	0.004	0.003	0.000	0.000
17031810400	10	0.053	0.070	0.072	0.052	1.000	0.000
17031090200	9	0.987	0.984	0.983	0.987	1.000	0.000
17031090300	9	0.951	0.956	0.958	0.951	0.000	0.000

Chicago Community Area data aggregation (Data analysis)

Data Structure

Theme

Availabilities

Universes

SEGMENT	TABLE NUMBER	FIELD NAME
01	P1	TOTAL POPULATION [1]
01	P1	Universe: Total Population
01	P1	Total
02	P2	URBAN AND RURAL [6]
02	P2	Universe: Total Population
02	P2	Total:
02	P2	Urban:
02	P2	Inside urbanized areas
02	P2	Inside urban clusters
02	P2	Rural

Chicago Community Area data aggregation (Data analysis)

Statistics

Dataset	Segment	Tables			
		County Level	Tract Level	Block Level	Total
Census 2010	47	10	86	235	331

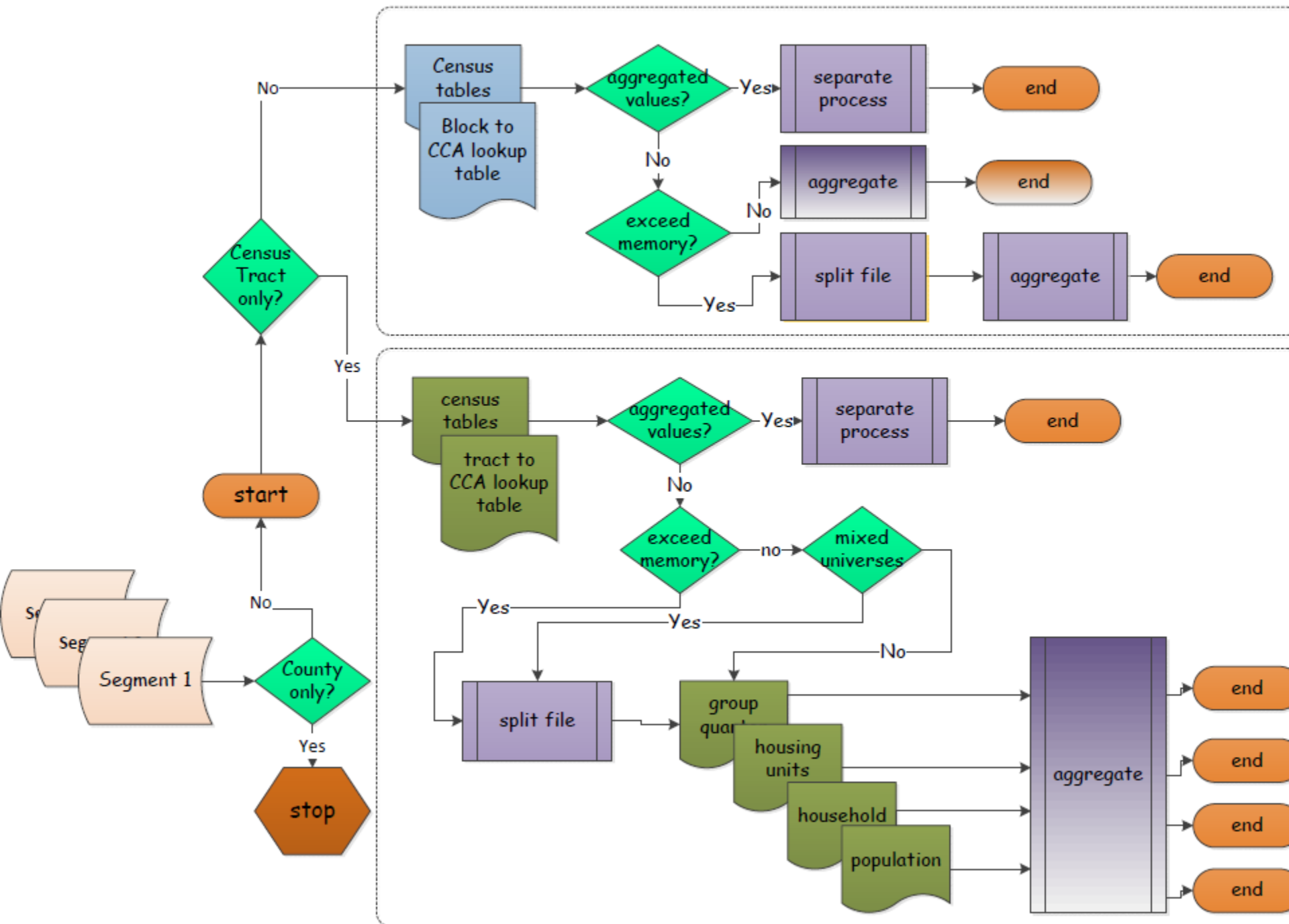
Lookup tables

Lookup Tables:

Block to CCA lookup table

Blockgroup to CCA lookup table

Tract to CCA lookup table



Data aggregation (Python)

```
import sqlite3
import csv
from dbfpy import dbf
import sys, string, os, time, difflib, math, shutil, commands, subprocess

def GetFileHeadings(InFile):
    dbfFile=dbf.Dbf(open(InFile,'r'))
    Headers_Cnt = range(len(dbfFile.fieldNames))
    Headers_Lst=dbfFile.fieldNames
    Headers_Tpl=tuple(Headers_Lst)
    return (Headers_Lst, Headers_Tpl, Headers_Cnt, dbfFile)

def CreateTable(Headers_A_Tpl,dbfFile_A, HeadersCnt_A, FileNmShrt):
    values_lst1=[]
    try:
        cursor.execute('drop table if exists Tab;')
        command11 = "create table Tab" + str(Headers_A_Tpl)
        for j in range(0,len(Headers_A_Tpl)):
            values_lst1.append("?")
        command12 = "insert into Tab values" + str(tuple(values_lst1))
        command13=command12.replace("'", '')
        cursor.execute(command11)
        for row_A in dbfFile_A:
            rows_A = []
            for num_A in HeadersCnt_A:
                rows_A.append(row_A[num_A])
            cursor.execute(command13, rows_A)
    except sqlite3.Error, message:
        print message
    except IndexError, message:
        print message
    return cursor

def CalculateTable(Headers_A_Lst, Headers_A_Tpl,dbfFile_A, HeadersCnt_A, FileNmShrt, i, IssueTracts_Lst, SplitTracts_Lst, UpdateDict_Pct,SegNum):
    try:
        values_lst1=[]
        values_lst2=[]
        values_lst3=[]
        Pop_Pct_Headers_Lst=[]
        default=0
        ##### Calculate the issues tracts #####
        Headers_A_Lst.remove('LOGRECNO')
        Headers_A_Lst.remove('GEOGNAME')
        Headers_A_Lst.remove('GEOGKEYX')
        Final_Headers_Lst=[]
        CSV_Headers_Lst=[]
        CSV_Headers_Tpl=()
        command50="DROP TABLE IF EXISTS Union_Tab_"+SegNum + ";"
        command61="CREATE TABLE Union_Tab_"+SegNum+" as SELECT * FROM TabA " # Pass on all data created in main function to Union_Tab table
        print "We are strating to process sgement : "+ SegNum
```

```

for tract_no in IssueTracts_Lst:
    Pop_Pct_Headers_Lst =[]
    print "This is the tract number we start to work with: " + tract_no
    print "***** Original Data for each issue tract just to compare the difference *****"
    #~~~~~ To see if the tracts are the two tracts that cross CCA boundaries ~~~~~
    print "len(tract_no) = " + str(len(tract_no) )
    if len(tract_no)>11 and (tract_no[:11] in SplitTracts_Lst):
        command11="SELECT * FROM TabA WHERE GEOGKEYX='" + tract_no[:11] + "'"
        cursor.execute(command11)
        rows0=cursor.fetchall()
        print rows0
        ratio = UpdateDict_Pct.get(tract_no,default)
        print "ratio that I am going to use: " + tract_no + ": " + str(ratio)
        ##### apply the ratio to the columns #####
        for Col in Headers_A_Lst:
            Pop_Pct_Headers_Lst.append(str(ratio) + "*" + Col + " as " + Col)

#print Pop_Pct_Headers_Lst
##### This is to do the calculation based on Population Ratio #####
command30="DROP TABLE IF EXISTS Tract_"+tract_no + ";"
cursor.executescript(command30)
command31="CREATE TABLE Tract_"+tract_no+" as SELECT LOGRECNO, "+str(Pop_Pct_Headers_Lst)
command32=command31.replace("'",'')
command33=command32.replace("[",'')
command34=command33.replace("]","")
command35=command34 + ", GEOGKEYX, GEOGNAME from TabA where GEOGKEYX='" + tract_no[:11] + "'"
cursor.execute(command35)
command40="UPDATE Tract_"+tract_no + " SET GEOGKEYX='"+tract_no + "'WHERE GEOGKEYX='"+ tract_no[:11] + "'"
cursor.execute(command40)

command41="SELECT * from Tract_"+ tract_no
cursor.execute(command41)
rows1=cursor.fetchall()
for row1 in rows1:
    print row1
print "Done with priting table Tract"
print "*****g"

print "This is to print out the last character from split tract : " + tract_no[12:]
if tract_no[12:]=='2' :
    ##### remove the last original tracts that have issues #####
    command51="DELETE FROM TabA WHERE GEOGKEYX='" + tract_no[:11] + "'"
    print" command51 : "+ command51
    cursor.execute(command51)
else:
    pass

```

What's next

What's next:

- * ACS 2006-2010 data
- * Improve the accuracy
- * Working on the median/average variables

Data access

- For external users:
 - * Data requests: Jon Hallas at Jhallas@cmap.illinois.gov
 - * Metropulse (www.metropulsechicago.org)
 - Incomplete coverage
 - Description might be a little bit different from AFF2
- For Internal users:
Data Depot (V:\Demographic_and_Forecast\Census)

Reference

National Historical Geographic Information System (NHGIS)

<https://www.nhgis.org/>

Tract to tract relationship tables

http://www.census.gov/geo/www/2010census/tract_rel/tract_rel.html

Census FTP site

<http://www2.census.gov>

Chicago Community Area History

<http://www.encyclopedia.chicagohistory.org/pages/319.html>

Thank You!

Questions?

Xiaohong Zhang

Data Librarian/Senior Analyst

Chicago Metropolitan Agency for Planning

Tel: 312-386-8719

Email: xzhang@cmaphillinois.gov