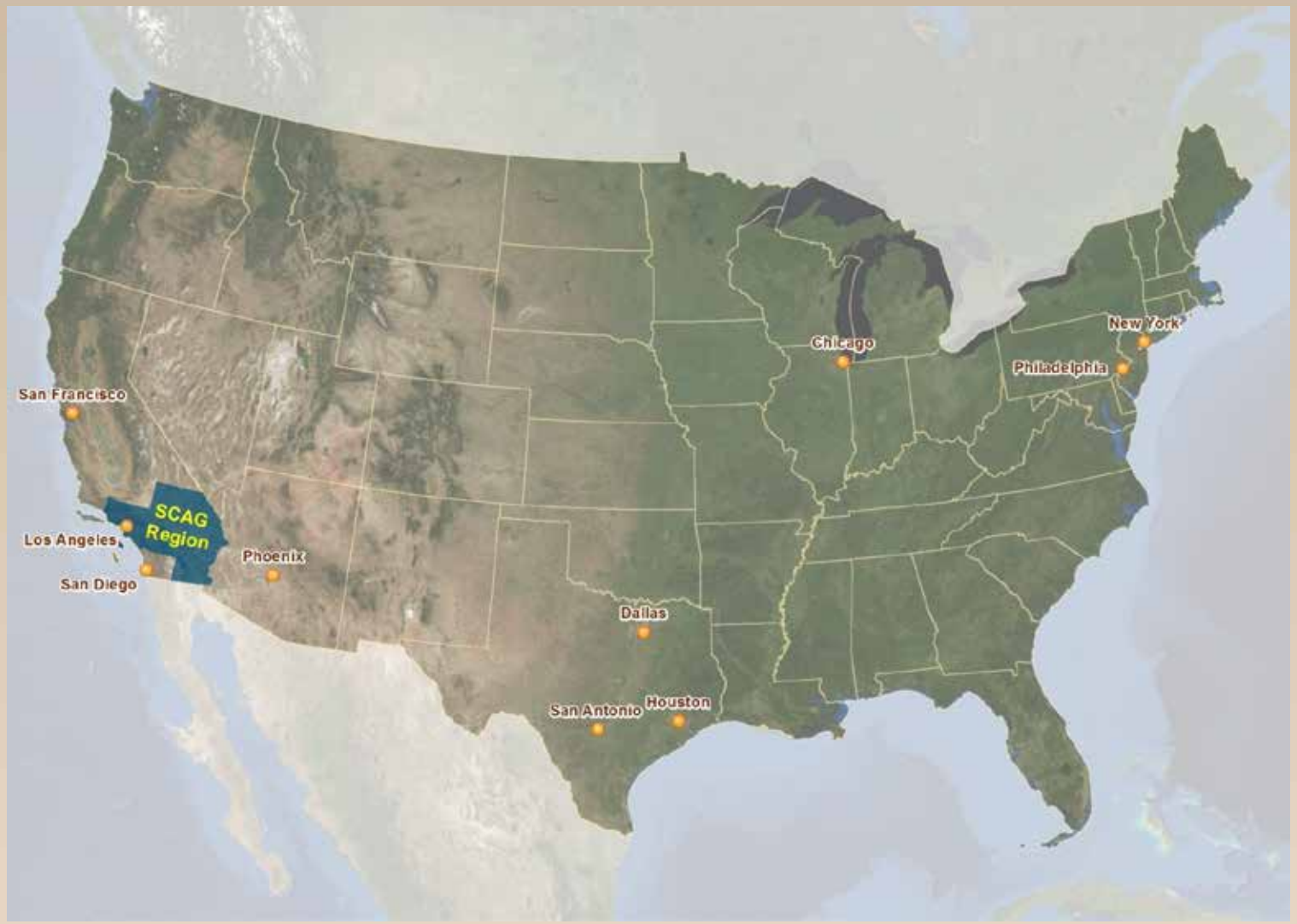
Python Scripting for Regional Land Use Data Management and QC Workflow

2015 ESRI User Conference | July 21, 2015 | San Diego, CA

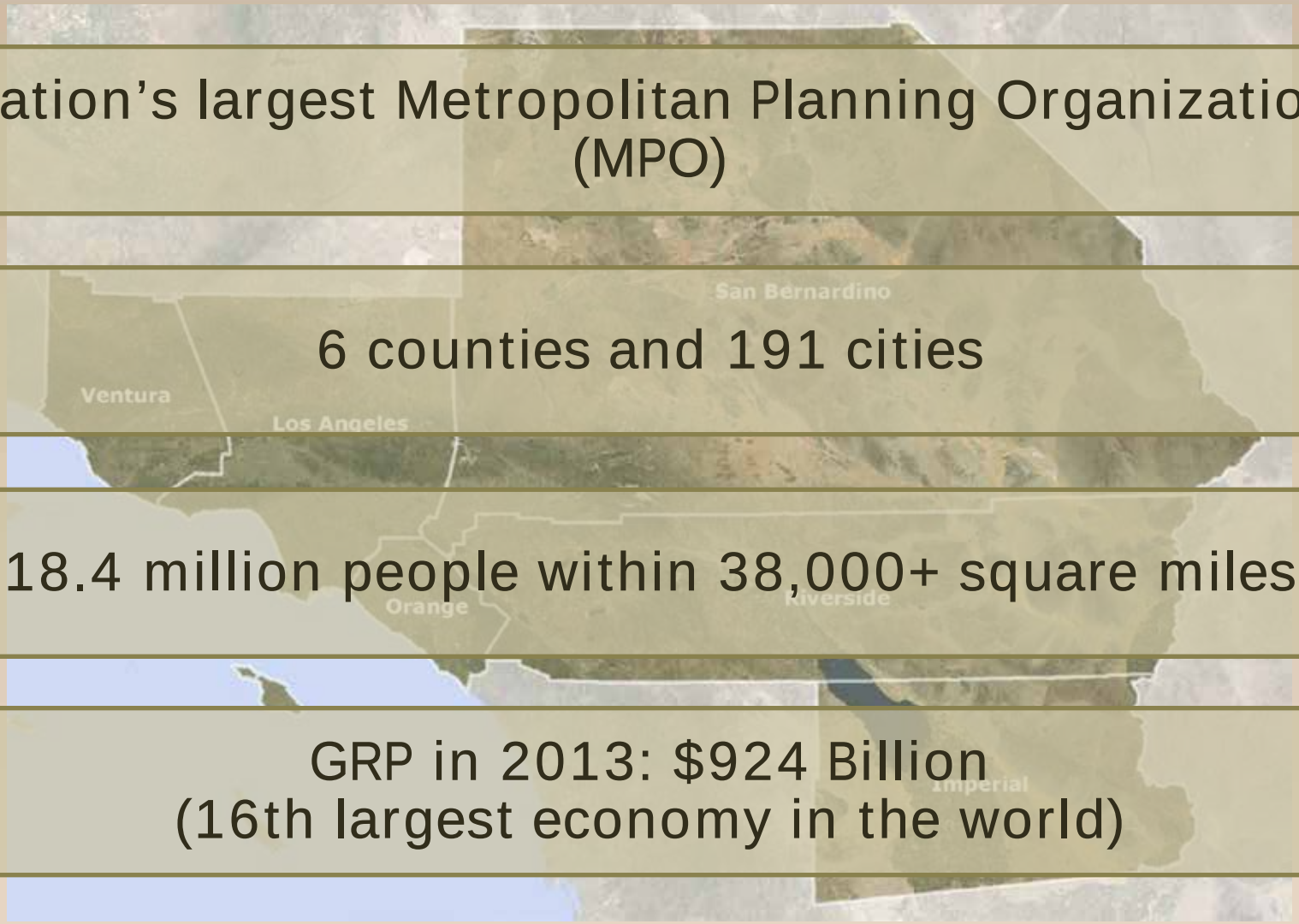
Jung Seo, Frank Wen, Simon Choi and Tom Vo

Research & Analysis
Southern California Association of Governments

Southern California Association of Governments (SCAG)



Southern California Association of Governments (SCAG)



Nation’s largest Metropolitan Planning Organization (MPO)

6 counties and 191 cities

18.4 million people within 38,000+ square miles

GRP in 2013: $924 Billion
(16th largest economy in the world)

Overview

- Background
- Objectives
- Methodology
- Conclusions

BACKGROUND

2016 RTP/SCS and Senate Bill 375

- 2016-2040 Regional Transportation Plan / Sustainable Communities Strategy (RTP/SCS)
  - A long-range transportation plan
- SB375 – California's Climate Protection Act
  - Integration of transportation, land use, housing and environmental planning to meet the regional GHG emission reduction targets

Bottom-Up Local Input Process

- Bottom-up local input process
  - Participation and cooperation of all 197 local government partners
- Review by local jurisdictions on SCAG's land use and resource areas datasets
  - SCAG Data/Map Book
  - One-on-one meeting
- Collect data changes, answer questions, provide technical guidance

Regional Land Use Database

- Development of regional land use database in preparation for the 2016 RTP/SCS
- Updated and reviewed thru the bottom-up local input process
- Base data for integrated growth forecast, scenario planning model, planning and policy analysis, etc.

Regional Land Use Database (Dataset Overview)

- Land use types
  - General plan land use and zoning (GPZN)
  - Existing land use (LU)
  - Specific plan land use (SP)
- Geographic level
  - Parcel data at city and county level
- Standardized land use code
  - 40 general plan land use classification
  - 110 existing land use classification

Table 2:
2012 SCAG Existing Land Use Codes - Legend

Legend	Land Use Description
Single Family Residential	1110 Single Family Residential 1111 High-Density Single Family Residential 1112 Low-Density Single Family Residential
Multi-Family Residential	1120 Multi-Family Residential 1121 Mixed Multi-Family Residential 1122 Duplexes, Triplexes and 2- or 3-Unit Condominiums and Townhouses 1123 Low-Rise Apartments, Condominiums, and Townhouses 1124 Medium-Rise Apartments and Condominiums 1125 High-Rise Apartments and Condominiums
Mobile Homes and Trailer Parks	1130 Mobile Homes and Trailer Parks 1131 Trailer Parks and Mobile Home Courts, High-Density 1132 Mobile Home Courts and Subdivisions, Low-Density
Mixed Residential	1140 Mixed Residential 1100 Residential
Rural Residential	1150 Rural Residential
General Office	1210 General Office Use 1211 Low- and Medium-Rise Major Office Use 1212 High-Rise Major Office Use 1213 Skyscrapers
Commercial and Services	1200 Commercial and Services 1220 Retail Stores and Commercial Services 1221 Regional Shopping Center 1222 Retail Centers (Non-Strip With Contiguous Interconnected Off-Street Parking) 1223 Retail Strip Development 1230 Other Commercial 1231 Commercial Storage 1232 Commercial Recreation 1233 Hotels and Motels
Facilities	1240 Public Facilities 1241 Government Offices 1242 Police and Sheriff Stations 1243 Fire Stations 1244 Major Medical Health Care Facilities 1245 Religious Facilities 1246 Other Public Facilities 1247 Public Parking Facilities 1250 Special Use Facilities 1251 Correctional Facilities 1252 Special Care Facilities 1253 Other Special Use Facilities
Education	1260 Educational Institutions 1261 Pre-Schools/Day Care Centers 1262 Elementary Schools 1263 Junior or Intermediate High Schools 1264 Senior High Schools 1265 Colleges and Universities 1266 Trade Schools and Professional Training Facilities
Military Installations	1270 Military Installations 1271 Base (Built-up Area) 1272 Vacant Area 1273 Air Field 1274 Former Base (Built-up Area) 1275 Former Base Vacant Area 1276 Former Base Air Field
Industrial	1300 Industrial 1310 Light Industrial 1311 Manufacturing, Assembly, and Industrial Services 1312 Motion Picture and Television Studio Lots 1313 Packing Houses and Grain Elevators 1314 Research and Development 1320 Heavy Industrial 1321 Manufacturing 1322 Petroleum Refining and Processing 1323 Open Storage 1324 Major Metal Processing 1325 Chemical Processing 1330 Extraction 1331 Mineral Extraction - Other Than Oil and Gas 1332 Mineral Extraction - Oil and Gas 1340 Wholesaling and Warehousing
Transportation, Communications, and Utilities	1400 Transportation, Communications, and Utilities 1410 Transportation 1411 Airports 1412 Railroads 1413 Freeways and Major Roads 1414 Park-and-Ride Lots 1415 Bus Terminals and Yards 1416 Truck Terminals 1417 Harbor Facilities 1418 Navigation Aids 1420 Communication Facilities 1430 Utility Facilities 1431 Electrical Power Facilities 1432 Solid Waste Disposal Facilities 1433 Liquid Waste Disposal Facilities 1434 Water Storage Facilities 1435 Natural Gas and Petroleum Facilities 1436 Water Transfer Facilities 1437 Improved Flood Waterways and Structures 1438 Mixed Utilities 1440 Maintenance Yards 1441 Bus Yards 1442 Rail Yards 1450 Mixed Transportation 1460 Mixed Transportation and Utility
Mixed Commercial and Industrial	1500 Mixed Commercial and Industrial
Mixed Residential and Commercial	1600 Mixed Residential and Commercial
Open Space and Recreation	1800 Open Space and Recreation 1810 Golf Courses 1820 Local Parks and Recreation 1830 Regional Parks and Recreation 1840 Cemeteries 1850 Wildlife Preserves and Sanctuaries 1860 Specimen Gardens and Arboreta 1870 Beach Parks 1880 Other Open Space and Recreation
Agriculture	2000 Agriculture 2100 Cropland and Improved Pasture Land 2110 Irrigated Cropland and Improved Pasture Land 2120 Non-Irrigated Cropland and Improved Pasture Land 2200 Orchards and Vineyards 2300 Nurseries 2400 Dairy, Intensive Livestock, and Associated Facilities 2500 Poultry Operations 2600 Other Agriculture 2700 Horse Ranches
Vacant	3000 Vacant 3100 Vacant Undifferentiated 3200 Abandoned Orchards and Vineyards 3300 Vacant With Limited Improvements 3400 Beaches (Vacant) 1900 Urban Vacant
Water	4000 Water 4100 Water, Undifferentiated 4200 Harbor Water Facilities 4300 Marina Water Facilities 4400 Water Within a Military Installation 4500 Area of Inundation (High Water)
Under Construction	1700 Under Construction

Regional Land Use Database (Attribute Information)

- General plan and zoning (GPZN)
  - City's GP designations
  - City's zoning codes
  - SCAG's standardized GP code
  - Residential density (average, min/max)
  - Year of adoption/amendment
- Existing land use (LU)
  - SCAG's standardized LU code
- SCAGUID12, APN, county, city, acreage, etc.

GPZN Attribute Table

Table

GeneralPlan_poly_IM_2012

APN	FIPS	X_CENTER	Y_CENTER	Shape_Leng	Shape_Area	CITY	COUNTY	DENSITY	LOW	HIGH	YEAR_ADOPT	ZONE_CODE	CITY_GP_CO	SCAG_GP_CO	NOTES	ACRES
037-140-011	06025	638511.3907	3654162.91866	867.94793	38760.070408	Brawley	Imperial	0	0	0	9/1/2008	P-F	Public Facilities	1240		9.577822
037-140-017	06025	638595.999944	3654473.75361	1812.635074	118042.283834	Brawley	Imperial	0	0	0	9/1/2008	P-F	Public Facilities	1240		29.168884
037-160-045	06025	638041.062908	3652403.85166	285.334168	5133.691849	Brawley	Imperial	0	0	0	9/1/2008	M-1	Industrial	1300		1.26854
037-160-059	06025	637921.155732	3652554.56655	165.468035	997.060951	Brawley	Imperial	0	0	0	9/1/2008	M-1	Industrial	1300		0.246379
037-160-061	06025	637989.556365	3652537.32488	129.962566	687.650531	Brawley	Imperial	0	0	0	9/1/2008	M-1	Industrial	1300		0.169922
037-160-062	06025	638004.2586	3652522.24847	57.056496	132.771006	Brawley	Imperial	0	0	0	9/1/2008	M-1	Industrial	1300		0.032808
037-160-063	06025	637937.948128	3652381.99157	109.25227	382.189054	Brawley	Imperial	0	0	0	9/1/2008	M-1	Industrial	1300		0.094435
037-160-064	06025	638021.856001	3652459.28663	578.685552	14021.979706	Brawley	Imperial	0	0	0	9/1/2008	M-1	Industrial	1300		3.46489
037-160-065	06025	637957.38317	3652440.92265	131.069994	575.675568	Brawley	Imperial	0	0	0	9/1/2008	M-1	Industrial	1300		0.142253
037-160-066	06025	637901.711136	3652460.24545	598.816962	22186.776359	Brawley	Imperial	0	0	0	9/1/2008	M-1	Industrial	1300		5.482461
037-160-067	06025	637837.984264	3652373.12774	101.655653	523.578865	Brawley	Imperial	0	0	0	9/1/2008	M-1	Industrial	1300		0.129379
037-160-068	06025	637732.136451	3652585.937	439.827889	10386.666588	Brawley	Imperial	0	13	17	9/1/2008	R-3	Medium Density Residential	1120		2.566601
037-160-069	06025	637960.201947	3652609.10518	1208.85293	44769.09197	Brawley	Imperial	0	13	17	9/1/2008	R-3	Medium Density Residential	1120		11.062684
037-160-070	06025	637733.987673	3652444.80238	440.450656	12527.401026	Brawley	Imperial	0	13	17	9/1/2008	R-3	Medium Density Residential	1120		3.095588
037-160-071	06025	637833.540848	3652498.00581	216.087925	1451.313029	Brawley	Imperial	0	13	17	9/1/2008	R-3	Medium Density Residential	1120		0.358627
040-130-008	06025	636784.741195	3647114.55746	953.938872	53790.613387	Brawley	Imperial	0	0	0	9/1/2008	M-1	Industrial	1300		13.29195
040-130-009	06025	636809.938224	3647320.64962	526.898606	16889.522459	Brawley	Imperial	0	0	0	9/1/2008	M-1	Industrial	1300		4.173492
040-130-011	06025	636763.787475	3646941.74893	657.526047	20325.357819	Brawley	Imperial	0	0	0	9/1/2008	M-1	Industrial	1300		5.022505
046-050-028	06025	635605.571326	3650917.67922	215.433814	1949.176497	Brawley	Imperial	0	5.5	9	9/1/2008	R-E	Low Density Residential	1110		0.481652
046-050-029	06025	635639.581791	3650878.58522	349.806291	6670.000252	Brawley	Imperial	0	5.5	9	9/1/2008	R-E	Low Density Residential	1110		1.648193
046-050-030	06025	635625.890428	3650787.54154	402.263718	8855.781884	Brawley	Imperial	0	5.5	9	9/1/2008	R-E	Low Density Residential	1110		2.188311
046-050-034	06025	635299.792737	3650890.03907	959.768199	24728.807589	Brawley	Imperial	0	0	0	9/1/2008	A-1	Open Space	1800		6.110621

1 (0 out of 85929 Selected)

LU Attribute Table

Table

landuse_poly_IM_2012

FID	Shape	SCAGUID12	APN	FIPS	X_CENTER	Y_CENTER	Shape_Leng	Shape_Area	CITY	COUNTY	LU12	ACRES
0	Polygon	0250037925	037-140-011	06025	638511.3907	3654162.91866	867.94793	38760.070408	Brawley	Imperial	1433	9.577822
1	Polygon	0250037930	037-140-017	06025	638595.999944	3654473.75361	1812.635074	118042.283834	Brawley	Imperial	1433	29.168884
2	Polygon	0250037971	037-160-045	06025	638041.062908	3652403.85166	285.334168	5133.691849	Brawley	Imperial	1310	1.26854
3	Polygon	0250037980	037-160-059	06025	637921.155732	3652554.56655	165.468035	997.060951	Brawley	Imperial	1310	0.246379
4	Polygon	0250037981	037-160-061	06025	637989.556365	3652537.32488	129.962566	687.650531	Brawley	Imperial	1311	0.169922
5	Polygon	0250037982	037-160-062	06025	638004.2586	3652522.24847	57.056496	132.771006	Brawley	Imperial	1311	0.032808
6	Polygon	0250037983	037-160-063	06025	637937.948128	3652381.99157	109.25227	382.189054	Brawley	Imperial	1311	0.094435
7	Polygon	0250037984	037-160-064	06025	638021.856001	3652459.28663	578.685552	14021.979706	Brawley	Imperial	3100	3.46489
8	Polygon	0250037985	037-160-065	06025	637957.38317	3652440.92265	131.069994	575.675568	Brawley	Imperial	3100	0.142253
9	Polygon	0250037986	037-160-066	06025	637901.711136	3652460.24545	598.816962	22186.776359	Brawley	Imperial	1311	5.482461
10	Polygon	0250037987	037-160-067	06025	637837.984264	3652373.12774	101.655653	523.578865	Brawley	Imperial	1450	0.129379
11	Polygon	0250037988	037-160-068	06025	637732.136451	3652585.937	439.827889	10386.666588	Brawley	Imperial	3100	2.566601
12	Polygon	0250037989	037-160-069	06025	637960.201947	3652609.10518	1208.85293	44769.09197	Brawley	Imperial	3100	11.062684
13	Polygon	0250037990	037-160-070	06025	637733.987673	3652444.80238	440.450656	12527.401026	Brawley	Imperial	3100	3.095588
14	Polygon	0250037991	037-160-071	06025	637833.540848	3652498.00581	216.087925	1451.313029	Brawley	Imperial	1300	0.358627
15	Polygon	0250039055	040-130-008	06025	636784.741195	3647114.55746	953.938872	53790.613387	Brawley	Imperial	1300	13.29195
16	Polygon	0250039056	040-130-009	06025	636809.938224	3647320.64962	526.898606	16889.522459	Brawley	Imperial	1300	4.173492
17	Polygon	0250039058	040-130-011	06025	636763.787475	3646941.74893	657.526047	20325.357819	Brawley	Imperial	1300	5.022505
18	Polygon	0250045973	046-050-028	06025	635605.571326	3650917.67922	215.433814	1949.176497	Brawley	Imperial	1800	0.481652
19	Polygon	0250045974	046-050-029	06025	635639.581791	3650878.58522	349.806291	6670.000252	Brawley	Imperial	1110	1.648193
20	Polygon	0250045975	046-050-030	06025	635625.890428	3650787.54154	402.263718	8855.781884	Brawley	Imperial	3100	2.188311
21	Polygon	0250045978	046-050-034	06025	635299.792737	3650890.03907	959.768199	24728.807589	Brawley	Imperial	1800	6.110621
22	Polygon	0250045979	046-050-035	06025	635205.242367	3650747.81545	239.232786	2440.783877	Brawley	Imperial	1820	0.603131

1 (0 out of 85929 Selected)

Regional Land Use Database (Dataset Size)

County	Parcel No.	GPZN Size	LU Size
Imperial	85,929	62 MB	33 MB
Los Angeles	2,092,552	1,620 MB	965 MB
Orange	661,051	682 MB	471 MB
Riverside	810,948	749 MB	490 MB
San Bernardino	804,529	615 MB	358 MB
Ventura	252,602	238 MB	158 MB
SCAG Region	**4,707,611**	**3.92 GB**	**2.41 GB**

OBJECTIVES

Objectives

- Development of an effective workflow for huge regional land use database
  - To develop an efficient regional land use data management and QC process
  - To develop a standardized and reliable workflow

METHODOLOGY

Base Data Development Process

- Sources for 2012 land use base data
  - SCAG's 2008 land use data
  - DMP LPS property data
  - DMP new construction data
- Data processing & standardization
  - Property data processing thru Statistical Analysis Software (SAS) by county
  - Geoprocessing thru ArcGIS applications and Python scripting

Base Data Development Process (Data Standardization)

- Data Standardization
  - 2008 land use codes è new 2012 codes
  - Attribute field properties
    - Field value type and lengths
  - Disaggregate to city-level for data entry and update process
  - Merge city-level datasets to county-level

Sample Python Scripts (Data Standardization)

- *SelectLayerByAttribute_management*

```
# Make a layer from the feature class
arcpy.MakeFeatureLayer_management(fc, "test")

# Select features by 'CITY' name
whereClause = '"' + cityField + '"= ' + "'" + cityValue + "'"
arcpy.SelectLayerByAttribute_management("test","NEW_SELECTION",whereClause)

# Write the selected features to a new featureclass
newFC = "P:/=general_plan_2012/shapes/=updates/County/GPZN_County_Feb2015/
=breakdown/" + cnty + "/" + cityName + "_GPZN.shp"
arcpy.CopyFeatures_management("test", newFC)
```

- *Merge_management, AddField_management*

```
# Merge city-level shapefiles to county-level shapefiles
fcList = arcpy.ListFeatureClasses("*.shp", "")
fcMerged = "P:/=general_plan_2012/shapes/=updates/County/GPZN_County_Feb2015/GeneralPlan
_poly_" + cnty + "_2012.shp"
arcpy.Merge_management(fcList, fcMerged)

# Add a field to include parcel acreage and calculate acreage
arcpy.AddField_management (fcMerged, "ACRES", "DOUBLE", 15, 10, "", "ACRES", "NULLABLE")
arcpy.CalculateField_management(fcMerged, "ACRES", "!shape.area@acres!", "PYTHON_9.3")
```

Data Entry and Update Process

- Manual update on city-level data by staff
  - Inputs received from jurisdictions
  - Edits on field values and parcel shape
- Merge city-level data to county-level data
- County-level update, e.g. open space
- *The Problem* – Potential human mistakes
  - Incorrect attribute field value – land use, city name, density info, etc.
  - Incorrect parcel shape & location

Data Review Process

- QC process for:
  - Data standardization
  - City-level datasets
  - County-level datasets
- Types of QC
  - Attribute field information
  - Feature comparison
  - Spatial match
- Development of Python-based workflow

Data Review Process (City-Level Datasets)

- QC for City-Level Datasets
  - Geographic comparison
    - Parcel location vs. city boundary
  - Feature comparison
    - Geometry, feature count, attribute field count, etc.
  - Attribute field information
    - Field value accuracy – city name, residential density, null values, etc.

Sample Python Scripts (Feature Comparison)

- *FeatureCompare_Management*

```
# Set variables for feature comparison
baseFeature = "P:/=general_plan_2012/shapes/=updates/City/" + cnty + "/" + cityName + "_GPZN.shp"
testFeature = "P:/=existing_landuse 2012/shapes/City/=updates/" + cnty + "/" + cityName + "_LU.shp"
sortField = "SCAGUID12"
xyTolerance = "1 METERS"
compareType = "GEOMETRY_ONLY"
continueCompare = "CONTINUE_COMPARE"

# Compare feature and print the results
compareResult = arcpy.FeatureCompare_management(baseFeature, testFeature, sortField, compareType, "",
xyTolerance, "", "",     "", "", continueCompare)
print compareResult
print arcpy.GetMessages()
```

- *output*

```
===== Begin of Feature Comparison for IM =====
Brawley (Total Features: 7319)

Executing: FeatureCompare P:/=general_plan_2012/shapes/=updates/City/IM/Brawley_GPZN.shp
"P:/=existing_landuse   2012/shapes/City/=updates/IM/Brawley_LU.shp" SCAGUID12 GEOMETRY_ONLY # "1 Meters"
0.001 0.001 # # CONTINUE_COMPARE #
Start Time: Mon Sep 08 17:08:44 2014
Table: Tables have different number of fields (Base: 24, Test: 13).
Table: Table row counts are the same.
SpatialReference: Spatial references are the same.
FeatureClass: Geometries are the same.
Succeeded at Mon Sep 08 17:08:45 2014 (Elapsed Time: 1.00 seconds)
```

Data Review Process (County-Level Datasets)

- QC for County-Level Datasets
  - Feature comparison
    - Feature count
    - New SCAGUID12 for subdivided parcel
  - Attribute field information
    - Field value accuracy – incorrect land use codes, SCAGUID12, etc.

Sample Python Scripts (Attribute Fields)

- *SearchCursor, getValue*

```
cursor = arcpy.SearchCursor(fc)
for row in cursor:
    citynameRow = row.getValue(cityField)
    citygpRow = row.getValue(citygpField)
    scaggpRow = row.getValue(scaggpField)
    zoneRow = row.getValue(zoneField)

    if citynameRow != cityName:
        citynameNull = citynameNull + 1
    if citygpRow == " ":
        citygpNull = citygpNull + 1
    if scaggpRow == " ":
        scaggpNull = scaggpNull + 1
    if zoneRow == " ":
        zoneNull = zoneNull + 1
    parcelNum = parcelNum + 1
```

```
cursor = arcpy.SearchCursor(fc)
for row in cursor:
    cityRow = row.getValue(cityField)
    gpRow = row.getValue(gpField)
    uidRow = row.getValue(uidField)

    if cityValue == cityRow:
        if gpRow not in correctGP:
            incorrectGP = incorrectGP + 1
            parcelNum = parcelNum + 1
            print cityValue + " (" + uidRow
            + "): " + gpRow
        else:
            parcelNum = parcelNum + 1
```

- *CalculateField_management*

```
if luRow == "1113":
    luRow_before = luRow
    arcpy.CalculateField_management (fc, luField, "1150")
```

CONCLUSIONS

Benefits of Python Scripting for Data Management and QC Workflow

- Time and labor efficiency in managing and reviewing numerous and sizable datasets
- High data consistency and reliability
  - Consistent chain of data management and review process
  - Conformity with standardized data format
- Effective file and directory operation
  - *os.makedirs, shutil.copy, etc.*

Thank you!

Jung H. Seo

Southern California Association of Governments
seo@scag.ca.gov

SOUTHERN CALIFORNIA ASSOCIATION of GOVERNMENTS
50 YEARS