

\*\*\* Files needed for exercise: CA\_BG\_2010\_pop.shp, CA\_tract\_2010.shp, CA\_tract\_pharm\_OD\_5closest.shp, and LA\_greater\_500Cities\_tract\_HBP.shp

**Goals:** The goal of this exercise is to explore how to analyze patterns in your data. You will first create population weighted centroids (PwC) for census tracts in California to conceptualize where your population lives; and you are going to use recently generated population centers to inform the hotspot analysis on high blood pressure prevalence in the Greater Los Angeles Area.

**Skills:** After completing this exercise, you will be able to generate population weighted centroids at different scales. You will learn how to use hotspot analysis tool provided by ArcGIS, and properly interpret the hotspot map.

#### Preparing Workspace

- 1. In the Windows Explorer Start Menu, open ArcMap.
- In the folder tree on the left side, browse to New Maps > My Templates and select a Blank Map. Click OK.

Q ArcMap - Getting Started		22				
Open existing map or make new ma	ap using a template					
Existing Maps  Recent  Browse for more  New Maps  My Templates  Templates  Standard Page Sizes  Architectural Pag  Sizes  Standard Page Sizes  North American (  Industry  USA  World  Browse for more   III	My Templates	•				
C:\Users\sitkonic\AppData\Roamir	ng\ESRI\Desktop10.1\ArcMap\Templates\Normal.mxt					
Default geodatabase for this map:  What is this?    C: \Users \sitkonic\Documents\ArcGIS\Default.gdb						
Do not show this dialog in the	future.	OK Cancel				



- 3. Click the Add Data icon to add new data to the map.
- 4. Click on the **Connect to Folder** button 🖴.
- 5. Browse to the data folder for this exercise (*Analyzing\_patterns\_Exercise\_data*) and click **OK**. You now have a permanent connection to that folder making it quicker to find and add data.
- 6. From the folder, hold Ctrl and click the shapefile *CA\_tract\_2010.shp* and

*CA\_BG\_2010\_pop.shp*. These files are projected census boundaries for California. And the block group file includes population counts.

Add Data	10. 80. 3			-				x
Look in: LA_greater CA_tract_pl CA_tract_20 CA_BG_201	Analyzing_patterns_Exercise_d; ▼ _500Cities_tract_HBP.shp harm_OD_5closest.shp D10.shp 0_pop.shp	仓		•	<b>2</b>		Ŭ	
Name: Show of type:	Datasets, Layers and Results			•		A	idd ncel	

- 7. Once you've selected the files, click Add and the data will appear in the Table of Contents as a layer.
- Now you have your census boundaries added into your workspace as shapefile layers.
  Because of the complexity of CA's boundaries, uncheck the display box before both of layers to disable the display for better performance.





#### **Generating Population weighted Centroid for Census Tracts**

- In order to generate the population weighted centroid for census tracts, we need sub-units of tracts with total population counts. Right click on CA\_BG\_2010\_pop to Open Attribute Table. We have total population but no tract identifier to categorize on. We are going to use Spatial Join tool to resolve that.
- 2. Click on the Search icon to call out the search tool tip.



 You want to convert all shapes into points in order to simplify the spatial relationship for spatial join. Type feature to point in the search box, and select Feature To Point (Data Management) tool in the search results.



 Select CA\_BG\_2010\_pop in the pop up window and name the output correspondingly. Please make sure to check the **Inside** option to keep the topological relation to the census tracts.



✓ Feature To Point	
Input Features	^
CA_BG_2010_pop	- 🖻
Output Feature Class	
R:\Projects\CDC\Thematic\Workspace\CA_BG_2010_pop_pt.shp	
✓ Inside (optional)	-
OK Cancel Environments	Show Help >>

5. Use the Search tool again as in step 2 to find Spatial Join Geoprocessing tool under Analysis toolbox. Select the newly generated CA\_BG\_2010\_pop\_pt as the input dataset (because we need to attach tract ID to block groups). Use CA\_tract\_2010 as Join feature. Join Operation: you want to JOIN\_ONE\_TO\_ONE. Keep the default: keep all target features checked, and Match option: INTERSECT. Make sure you save it in your Exercise\_Data folder as a shapefile.

CA_BG_2010_pop_pt	<b>–</b>	
oin Features		_
CA_tract_2010	-	E
utput Feature Class		
R:\Projects\CDC\Thematic\Workspace\CA_BG_pop_pt_tractID.shp		P
oin Operation (optional)		_
JOIN_ONE_TO_ONE		•
Keep All Target Features (optional)		
eld Map of Join Features (optional)		_
E COUNTYER (Text)	<u> </u>	
TRACTCE (Text)		_
BIKGPDCE (Text)		×
ECOID (Text)		
NAMELSAD (Text)	E	
MTFCC (Text)		
EUNCSTAT (Text)		
ALAND (Double)		
AWATER (Double)		
🗄 TotalPop (Double)		
TATEFP_1 (Text)		
COUNTYFP_1 (Text)		
TRACTCE_1 (Text)	*	
latch Option (optional)		
INTERSECT		•
earch Radius (optional)		



 Now that you have census block groups' centroid, population, and census tract ID, you can start to generate the population weighted centroid. Use the Search tool to find the Mean Center tool under Spatial Statistics category and click it.

ALL	Maps	<u>Data</u>	<u>Tools</u>	Images	
Mea	Mean center <u>Inv Extent</u> • <u>Search returned 4 items</u> • Mean Center (Spatial Statistics) (Tool)	9			
<u>Any E</u>	ixtent 🔻				
Searc	h returne	d 4 items	+		Sort By 🔻
S M Id to	ean Ce entifies olboxes	nter (S the geo system	patial S graphic toolbox	tatistics) (Tool) center (or the cen es\spatial statistics	nter of concentratio s tools.tbx\measurin

7. Choose the block group points with population and tract ID as input feature data. Set Weight Field to be TotalPop and Case Field to be GEOID\_1 (GEOID from census tracts). Don't forget to save your output file to an accessible location with a meaningful name. Then click OK.

Input Feature Class		
CA_BG_pop_pt_tractID		- 🖻
Output Feature Class		
:\Users\rl53\Documents\ArcGIS\Default.gdb\CA_1	tract_Pwc_2010	
Weight Field (optional)		
TotalPop		•
Case Field (optional)		
GEOID_1		•
Dimension Field (optional)		
		-

- You will see the warning message display in the result window. It is because we didn't exclude the census tracts with zero population. So there is no population center generated for those census tracts. That won't affect our analysis.
- 9. Remove all dataset except CA\_tract\_PwC\_2010 to make your workspace clean. You are going to use the newly generated population weighted centroids to run hotspot analysis.



#### Define your Study Scale for Hotspot analysis

- Next, you will try to use your newly generated PwC data to define the study scale for hotspot analysis. As mentioned in the presentation, we are trying to use the average driving distance to the closest 5 pharmacies as the proximity of the community size. We can start to use the point level pharmacy data in module 1 and OD cost matrix in module 2 to generate the OD cost lines. In this exercise we provided the file CA\_tract\_pharm\_OD\_5closest.shp in the folder. Add CA\_tract\_pharm\_OD\_5closest.shp to the workspace.
- 2. With the help of **Statistics** tool, you can calculate the average driving distance to the nearest 5 pharmacies.
- Open Attribute Table of CA\_tract\_pharm\_OD\_5closest.shp. Right click on the field Total\_Kilo. This field contains the driving distance from PwC to pharmacies in kilometer. Select Statistics in the dropdown menu.



4. Find **Mean** for the Total\_Kilo in the pop up. It reads 3.529 km.





#### Hotspot Analysis on Greater Los-Angeles area

- Browse to the data folder and add LA\_greater\_500Cities\_tract\_HBP.shp to your workspace. This data comes from CDC's 500 cities project. It includes both model smoothed crude high blood pressure prevalence and high blood press medical adherence at census tract level.
- 2. Use **Search** tool to find **Hot Spot Analysis (Getis-Ord G\*)** in spatial Statistics tool set; and click on it.

ALL	Maps	<u>Data</u>	<u>Tools</u>	Images
hots	pot analy	ysis		9
Any E	extent 🔻			
Searc	h returned	d 2 items	•	Sort By 🔻
<u>ج</u> ال	ot Spot iven a se	Analys t of wei	<b>is</b> (Geti ighted fe	is-Ord Gi*) (Spatial Statistics) (Tool) aatures, identifies statistically significa
to	olboxes\	system	toolbox	es\spatial statistics tools.tbx\mapping
ن الآ Gi to	ptimized ven incid olboxes\	l Hot S dent poi system	pot <b>Ana</b> nts or w toolbox	alysis (Spatial Statistics) (Tool) reighted features (points or polygons), res\spatial statistics tools.tbx\mapping

3. Before you run the Hot Spot Analysis tool, you want to find out the projection information of

the input dataset so you can put in the number for **Distance Band** parameter.

Seperal Source Selection Displa	v Symbology Fields	Definition Query	Labele	Joine & Relates	Time	HTML Popu
Jelection Displa	y Symbology Tields	Deminition Query		Juins & Melates	Time	ттысторо
Extent						
Тор	: -407658.861532 m					
Left: 122739.221969 m		Right: 181095.0	19589 m			
Bottom	: -477509.426364 m					
Data Source						
Projected Coordinate System:	NAD_1983_California	Albers		<b>*</b>		
Projection:	Albers					
False_Lasting:	-400000000					
Central Meridian:	-120.00000000					
Standard Parallel 1:	34.00000000			=		
Standard Parallel 2:	40.50000000					
Latitude Of Origin:	0.00000000					
Linear Unit:	Meter					
				-		
•	III			P.		
		Se	t Data Sour	rce .		
			t Data 300			



Right click on *LA\_greater\_500Cities\_tract\_HBP* and click **Properties**. You can find coordinate system information under **Source** tab. Your data has been projected to right system and the measuring unit in your coordinate system is **Meter**. Since the number you get from last step is 3.529 km, your input for **Distance Band** should be 3529 meters.

4. Go back to Hot Spot Analysis, select LA\_greater\_500Cities\_tract\_HBP as Input Feature Class; BPHIGH\_Cru (smoothed high blood pressure crude rate) as Input Field; specify the output folder and file name; choose FIXED\_DISTANCE\_BAND for Conceptualization of Spatial Relationships; leave Distance Method as default (EUCLIDEAN\_DISTANCE); and put in 3529 meters in Distance Band. Finally, don't forget to check Apply False Discovery Rate to minimize type I error.

💐 Hot Spot Analysis (Getis-Ord Gi*)		J
Input Feature Class	A	
LA_greater_500Cities_tract_HBP	- 🖻	
Input Field		I
BPHIGH_Cru	-	
Output Feature Class		
R:\Projects\CDC\Thematic\Workspace\LA_greater_HBP_hotspot.shp	e 1	h
Conceptualization of Spatial Relationships		
FIXED_DISTANCE_BAND	-	
Distance Method		1
EUCLIDEAN_DISTANCE	-	I
Standardization		I
NONE		
Distance Band or Threshold Distance (optional)		1
	3529	I
Self Potential Field (optional)		1
	-	I
Weights Matrix File (optional)		1
		I
Apply False Discovery Rate (FDR) Correction (optional)		I
	-	
OK Cancel Environments	Show Help >>	

5. Click **OK** and you have done your first Hot Spot analysis. Try to symbolize the BPHIGH\_Cru field in *LA\_greater\_500Cities\_tract\_HBP*. Please think about the following questions:



- Any similarity of two maps?
- o Any difference of two maps?
- How is the hotspot map going to help you in your working environment?
- Feel free to run multiple test by using different number in Distance Band. What do you observed? How is the observation going to impact the usefulness of the map?