


Project Details

3



To create a new project, go to **File** → **New Project** on the menu bar at the top, or by clicking the **New Project** icon  at the top left of the page

Project Details

4

- ▶ Choose a name for your project – (Example: *Raleigh_04282015*)
- ▶ Select a project file location on your computer
- ▶ Enter a project description (Optional)
- ▶ Select all counties in a state for which you want to generate synthetic population
 - ▶ If you want to select multiple counties, hold the **Ctrl** button while selecting. All selected counties must be within a single state
 - ▶ To select all counties in a state, double click on the name of the state

5

Project Details

Project Setup Wizard

Step 1: Region

a. Enter project name
Raleigh, NC 2015

b. Select a project file location
C:\Test

c. Enter project description (Optional)

d. Select one or more counties

State/Country
☐ New Hampshire
☐ New Jersey
☐ New Mexico
☐ New York
☒ North Carolina
☐ North Dakota

Note: Counties cannot be chosen across multiple states.

< Back **Next >** Cancel

6

Project Details

Project Setup Wizard

Step 2: Geographic Resolution

a. Choose the geographic resolution at which you want to synthesize the population

County
 County
 Census Tract
 Census Blockgroup
 Traffic Analysis Zone (TAZ)

b. Do you want to include geographic correspondence, sample data, and marginal totals must be provided for the chosen resolution? ☒ No

Note: If **No** is chosen, MABLE/Geocorr2K: Geographic Correspondence Engine will be used.

c. User provided

Select the Geographic Correspondence file

< Back **Next >** Cancel

- Select the geographical resolution at which you want to generate the synthetic population
- Using data from census, synthetic population can be generated at county, census tract and blockgroup levels
- For generating data at the TAZ level, a file that provides correspondence between the geography and PUMA (from where the population will be sampled) should be provided by the user

Project Details

7

Project Setup Wizard

Step 2: Geographic Resolution

a. Choose the geographic resolution at which you want to synthesize the population
Traffic Analysis Zone (TAZ) [v]

Note: If **TAZ** is chosen, all information including geographic correspondence, sample data, and marginal totals must be provided.

b. Will you provide Geographic Correspondence between the chosen geography and PUMA boundaries?
☒ Yes ☐ No

Note: If **No** is chosen, MABLE/Geocon2K: Geographic Correspondence Engine will be used.

c. User provided
 Select the Geographic Correspondence file
 [v]
 Browse to select file...

< Back Next > Cancel

- If TAZ is selected as the geographical resolution, **Step 2b** defaults to **Yes**. Alternately, if the user wants to provide data at any other geographical resolution, the **Yes** button can be clicked manually
- Click on the dropdown, navigate and select the pertinent file

Project Details

8

Project Setup Wizard

Step 2: Geographic Resolution

a. Choose the geographic resolution at which you want to synthesize the population
Traffic Analysis Zone (TAZ) [v]

Note: If **TAZ** is chosen, all information including geographic correspondence, sample data, and marginal totals must be provided.

b. Will you provide Geographic Correspondence between the chosen geography and PUMA boundaries?
☒ Yes ☐ No

Note: If **No** is chosen, MABLE/Geocon2K: Geographic Correspondence Engine will be used.

c. User provided
 Select the Geographic Correspondence file
 C:/Users/vparikapati3/Downloads/Downloads/Raleigh_Data/Raleigh_Data/geog_corr_rev_08-01-2013.csv [v]

< Back Next > Cancel

Project Details

9

Project Setup Wizard

Step 3: Population Sample

a. Will you provide sample data?

☒ Yes ☐ No

Note: If **No** is chosen, select the US Census data source to use for population synthesis. Note that .

b. Choose the Census data source you want PopGen to use.

c. User provided

Select the household sample file

Browse to select file...

Select the person sample file

Browse to select file...

Note: Groupquarter data is optional, but if the person control totals include residents of groupquarters, then provide groupquarter information as well to generate a representative synthetic population.

< Back Next > Cancel

Select the appropriate files for sample data

Project Details

10

Project Setup Wizard

Step 3: Population Sample

a. Will you provide sample data?

☒ Yes ☐ No

Note: If **No** is chosen, select the US Census data source to use for population synthesis. Note that .

b. Choose the Census data source you want PopGen to use.

c. User provided

Select the household sample file

C:/Users/vgarikapatt3/Downloads/Downloads/Raleigh_Data/Raleigh_Data/HM4_sample_06-04-2013.csv

Select the groupquarter sample file

Select the person sample file

C:/Users/vgarikapatt3/Downloads/Downloads/Raleigh_Data/Raleigh_Data/persons_sample_08-01-2013.csv

Note: Groupquarter data is optional, but if the person control totals include residents of groupquarters, then provide groupquarter information as well to generate a representative synthetic population.

< Back **Next >** Cancel

Project Details

11

Project Setup Wizard

Step 4: Marginal Totals

a. Will you provide the marginal totals for population characteristics of interest?

☒ Yes ☐ No

Note: If **No** is chosen, US Census Summary Files (SF) for year 2000 will be used.

b. Choose the Census data source you want PopGen to use.

c. User provided

Select the household marginal total file

C:/Users/vperkapets/Downloads/Downloads/Raleigh_Data/Raleigh_Data/hh_marginals_rev_08-01-2013.csv

Select the groupquarter marginal total file

C:/Users/vperkapets/Downloads/Downloads/Raleigh_Data/Raleigh_Data/per_marginals_rev_08-01-2013.csv

Select the person marginal total file

C:/Users/vperkapets/Downloads/Downloads/Raleigh_Data/Raleigh_Data/per_marginals_rev_08-01-2013.csv

Note: Groupquarter data is optional; but if the person marginal totals include residents of groupquarters, then provide groupquarter information as well to generate a representative synthetic population.

< Back **Next >** Cancel

Select the appropriate files for marginal data

Project Details

12

Project Setup Wizard

Step 5: MySQL Connection Settings

a. Hostname 127.0.0.1

b. Username root

c. Password 1234

d. Port 3306

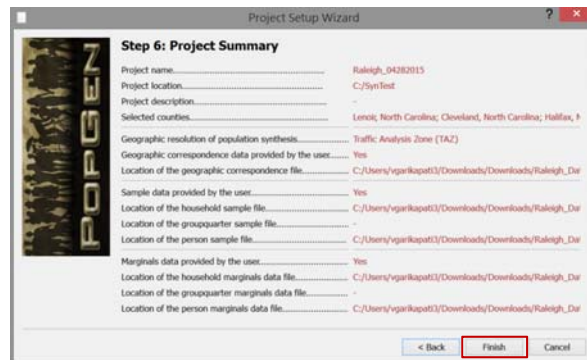
Note: A MySQL connection needs to be established before proceeding.

< Back **Next >** Cancel

- The hostname should be set to 127.0.0.1
- The default password is 1234. Do Not change this
- The default port is '3306'

Project Details

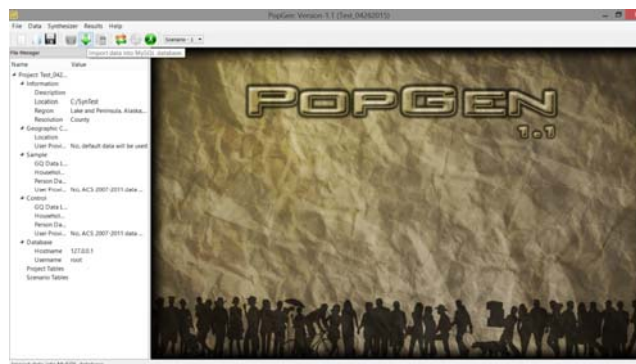
13



- A window showing all the previous steps with the selected options is displayed
- Review and ensure that everything is correct
- Click **Finish** to complete the project setup process

Data Import

14



Go to **Data** → **Import** at the top of the screen to start importing data or click the **Data Import** icon in the toolbar

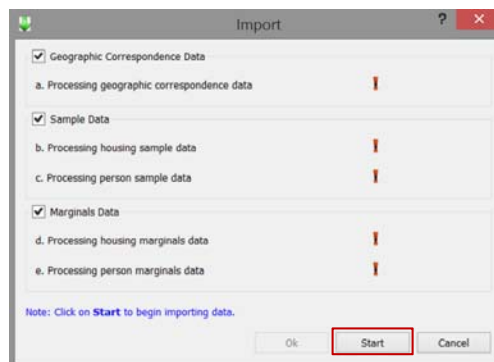
15

Data Import

- ▶ Click **Start** to begin the import process
- ▶ This step downloads the user provided data files, parses them, and then uploads them into the MySQL database as tables
- ▶ While importing, you will be prompted to overwrite any files that already exist

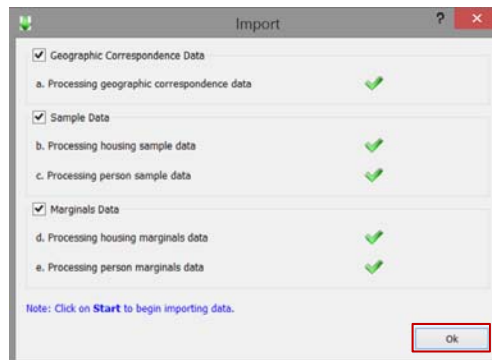
16

Data Import



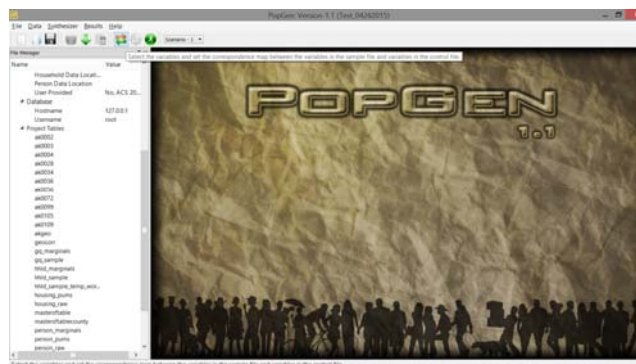
Data Import


17



Data Correspondence

18



Go to **Synthesizer** → **Set Corresponding Variables** in the menu bar at the top or click the  icon

19

Data Correspondence

- ▶ Select **Yes** for "Do you wish to match distributions of both Persons and Household attributes of interest?"
 - Note:** If you select **No**, the synthesizer will not control for person attributes of interest
- ▶ Under the '*Household Variables*' tab, **No** will be selected by default
 - ❑ If you select **Yes**, the program will try to mitigate person total inconsistencies by the modifying household marginal distributions

20

Data Correspondence

Corresponding Sample Categories with Marginal Variables

a. Do you wish to match distributions of both Persons and Household attributes of interest?

☒ Yes ☐ No

b. Set the correspondence between the sample variable categories and the columns in the Marginals table

Household Variables Person Variables Groupquarters Variables

Do you wish to modify the household size marginal distribution?

☐ Yes ☒ No

Select the household size variable name Enter the average value for the last household size category Select the person variable to obtain the person total

hhldtype agep

21

Data Correspondence

- ▶ To create correspondences between the variables, select the appropriate variables from the '*Sample Variable*' column and click **Select**
- ▶ The variables now appear in the '*Selected Variable*' column and are ready to be assigned correspondences
- ▶ This process is meant to establish correspondences between variables and their categories in the sample files and variables and their categories in the marginal files

22

Data Correspondence

Corresponding Sample Categories with Marginal Variables

a. Do you wish to match distributions of both Persons and Household attributes of interest?

☒ Yes ☐ No

b. Set the correspondence between the sample variable categories and the columns in the Marginals table

Household Variables ☐ Person Variables ☒

Do you wish to modify the household size marginal distribution?

☐ Yes ☒ No

Select the household size variable name. Enter the average value for the last household size category. Select the person variable to obtain the person total

pp1rhh: pp1rhh:

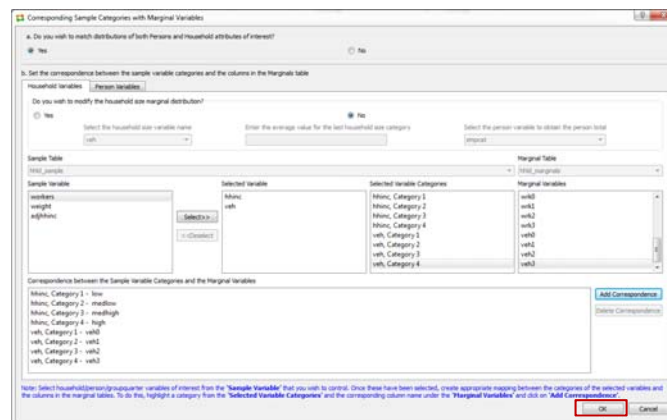
Sample Table	Marginal Table																								
<table border="1"> <thead> <tr> <th>Sample Variable</th> <th>Selected Variable</th> </tr> </thead> <tbody> <tr> <td>veh</td> <td>pp1rhh</td> </tr> <tr> <td>hhinc</td> <td></td> </tr> <tr> <td>numundTB</td> <td></td> </tr> <tr> <td>workers</td> <td></td> </tr> <tr> <td>numover65</td> <td></td> </tr> </tbody> </table>	Sample Variable	Selected Variable	veh	pp1rhh	hhinc		numundTB		workers		numover65		<table border="1"> <thead> <tr> <th>Selected Variable Categories</th> <th>Marginal Variables</th> </tr> </thead> <tbody> <tr> <td>pp1rhh, Category 1</td> <td>state</td> </tr> <tr> <td>pp1rhh, Category 2</td> <td>county</td> </tr> <tr> <td>pp1rhh, Category 3</td> <td>tract</td> </tr> <tr> <td>pp1rhh, Category 4</td> <td>log</td> </tr> <tr> <td>pp1rhh, Category 5</td> <td>hh</td> </tr> </tbody> </table>	Selected Variable Categories	Marginal Variables	pp1rhh, Category 1	state	pp1rhh, Category 2	county	pp1rhh, Category 3	tract	pp1rhh, Category 4	log	pp1rhh, Category 5	hh
Sample Variable	Selected Variable																								
veh	pp1rhh																								
hhinc																									
numundTB																									
workers																									
numover65																									
Selected Variable Categories	Marginal Variables																								
pp1rhh, Category 1	state																								
pp1rhh, Category 2	county																								
pp1rhh, Category 3	tract																								
pp1rhh, Category 4	log																								
pp1rhh, Category 5	hh																								

Correspondence between the Sample Variable Categories and the Marginal Variables

Note: Select household(person)/groupquarter variables of interest from the "Sample Variable" that you wish to control. Once these have been selected, create appropriate mapping between the categories of the selected variables and the columns in the marginal tables. To do this, highlight a category from the "Selected Variable Categories" and the corresponding column name under the "Marginal Variables" and click on "Add Correspondence".

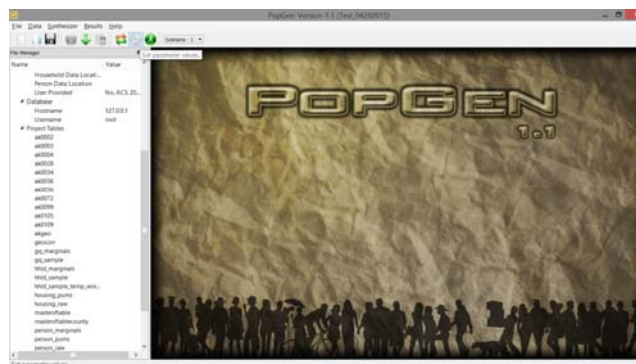
25


Data Correspondence



26

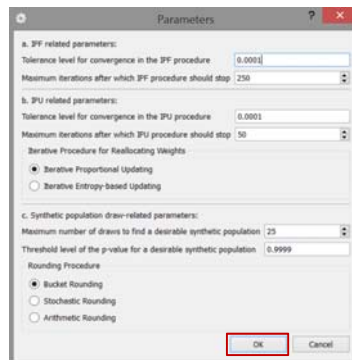
Parameters



Go to **Synthesizer** → **Parameters/Settings** in the menu bar at the top or click the  icon

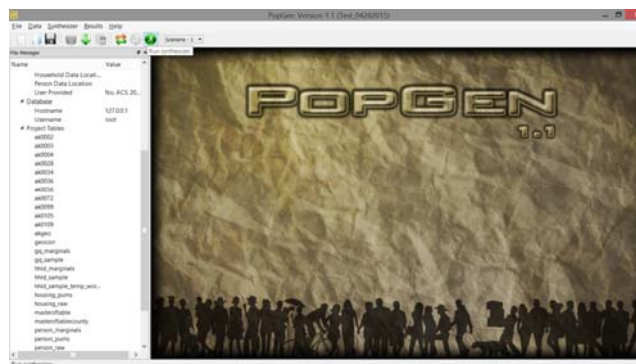
Parameters


27



Running the Synthesizer

28



Go to **Synthesizer** → **Run** in the menu bar at the top or click the  icon in the toolbar

29

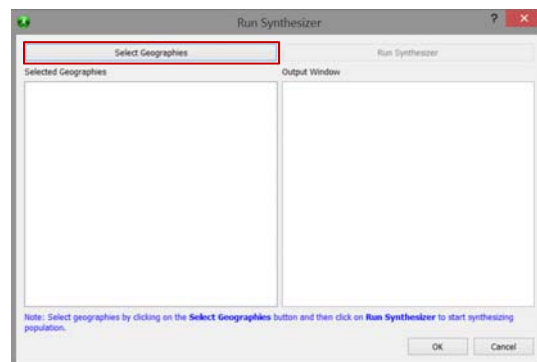
Running the Synthesizer

- ▶ Once the synthesizer window opens, click **Select Geographies**
- ▶ A new window opens with a list of all census block groups (or individual geographies corresponding to the chosen geographic resolution)
- ▶ Select all the geographies you intend to run by holding the **Ctrl** button or select all geographies using the **Ctrl + A**
- ▶ Click **OK** to continue

Note: You can choose as few or as many geographies as you like.

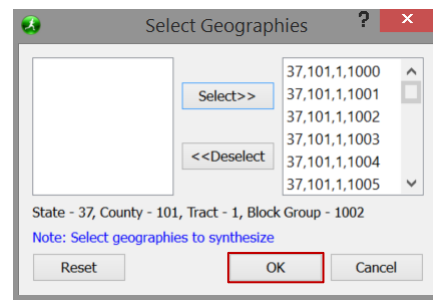
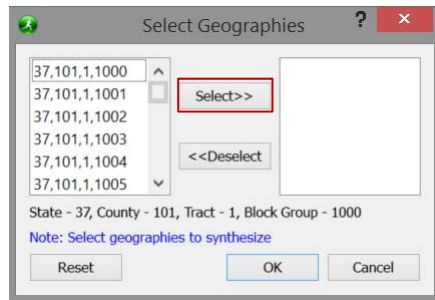
30

Running the Synthesizer



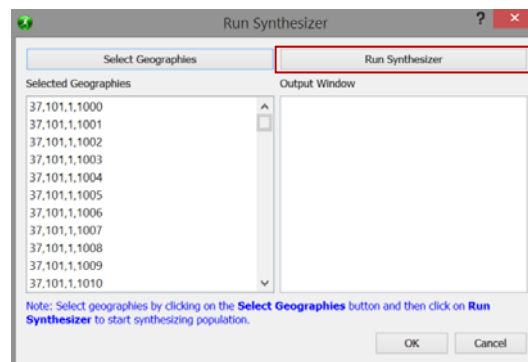
31

Running the Synthesizer



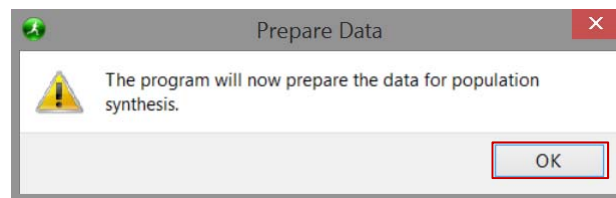
32

Running the Synthesizer



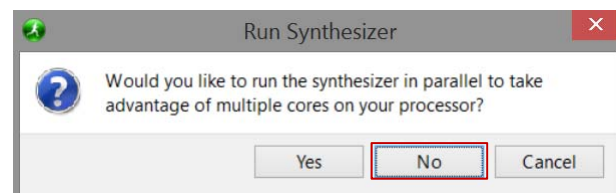
Running the Synthesizer

33



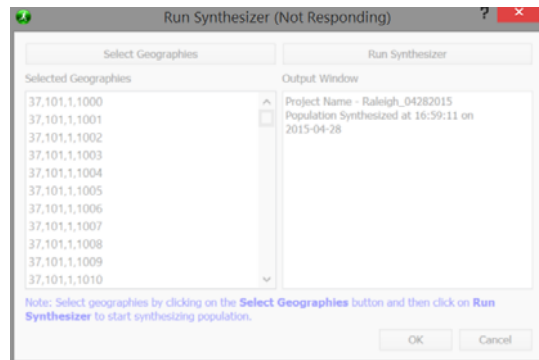
Running the Synthesizer

34



35

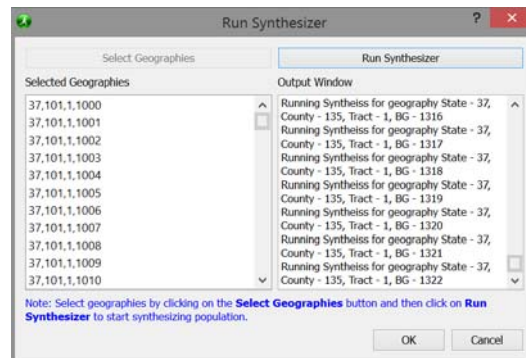
Running the Synthesizer



The synthesizer window appears frozen when the run is in progress. This is *OK!*

36

Running the Synthesizer



Once the run is completed, the 'Output Window' on the right hand side of the 'Run Synthesizer Window' will display each geography that was processed

37

