

Application of DEGURBA to the Globe

Exercise notes

Disaggregation of statistical data by degree of urbanisation class

Last Edited: April 2022

Pietro Florio

European Commission – Joint Research Centre (JRC)
Disaster Risk Management Unit – JRC E.1
Via E. Fermi, 2749 - 21027 Ispra (VA) – Italy
e-mail: pietro.florio@ec.europa.eu

Michele Melchiorri

European Commission – Joint Research Centre (JRC)
Disaster Risk Management Unit – JRC E.1
Via E. Fermi, 2749 - 21027 Ispra (VA) – Italy
e-mail: michele.melchiorri@ec.europa.eu

This note contains the summary to break down statistics according to the Degree of Urbanisation. The classification of territorial units is achieved with the Degree of Urbanisation Territorial Units Classifier tool (GHS-DU-TUC) via automatic geospatial and geo-statistical processing. The Degree of Urbanisation classifies human settlements based on population size, population density and grid contiguity. Results are inspected in spreadsheet and GIS software. The Degree of Urbanisation training has been designed in the Global Human Settlement Layer framework, which serves the global commitment of the European Union and other five Intergovernmental Organisations to develop a harmonised people-based global definition of cities and settlements. The Degree of Urbanisation has been endorsed by the 51st Session of the United Nations Statistical Commission as method to delineate cities, urban and rural areas for international statistical comparison.

1 Exercise summary

The exercise demonstrates the disaggregation of statistics relative to given territorial units collected in a table, by degree of urbanisation class. This activity is particularly useful in the view of Sustainable Development Goals progress monitoring and reporting.

In the first phase, the Degree of Urbanisation class is joined as attribute to the statistics table, with the help of a code that univocally identifies each territorial unit. To succeed in the join operation, the territorial units listed in the statistics table should be indexed with the same univocal code that identifies them in the shape file used as input in the GHS-DU-TUC Tool. Subsequently, the statistics table with the appended DEGURBA class is exported to a spreadsheet.

In the second phase, MS Excel is used to produce a Pivot Table that operates on the values of the statistics table by Degree of Urbanisation class. In this case, we sum the number of residential households in the census having or not having access to electricity, by Degree of Urbanisation class.

The exercise is carried out with a spreadsheet software and a GIS software. The exercise is demonstrated with MS Excel and Qgis but operations can be replicated with other software as well. The exercise is guided, the duration is approximately 2h. Remember to save from time to time your progress. To perform the exercises your computer should have installed: a) Matlab Runtime R2020b (9.9)¹; b) ArcMAP (or Qgis²); c) MS Excel (or any other spreadsheet software).

¹ It can be downloaded from <https://it.mathworks.com/products/compiler/matlab-runtime.html>

² It can be downloaded from <https://qgis.org/en/site/forusers/download.html>

2 Exercise 1: Disaggregation of statistical data by degree of urbanisation class

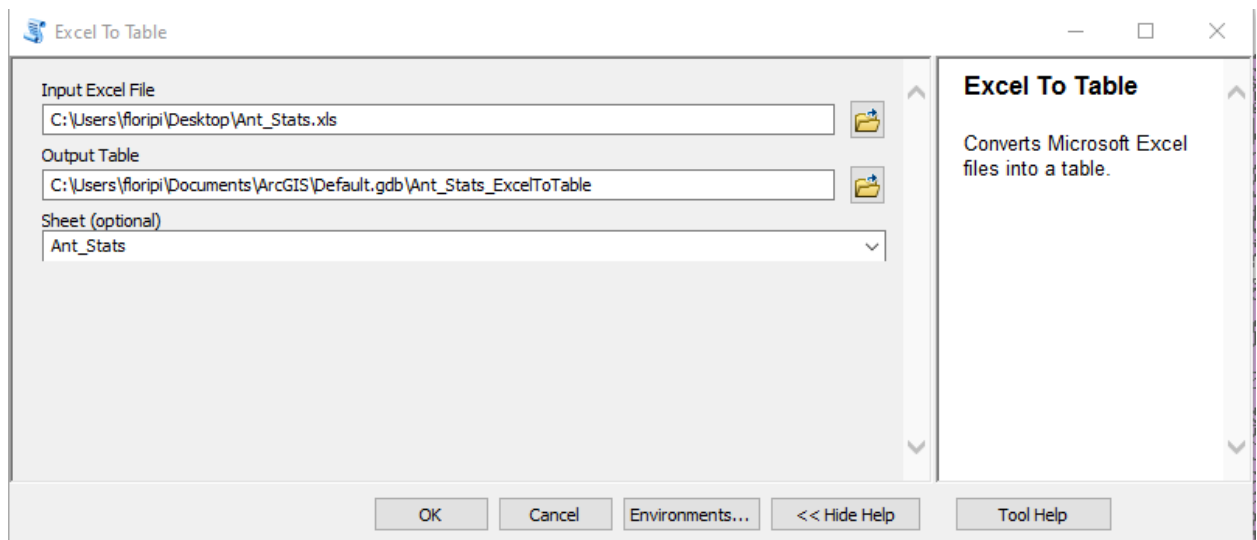
2.1 Input data

- 1) The exercise requires:
 - a) A statistics table listing territorial units in Colombia along with some attributes, which include the number of households having access to electricity and the number of those not having access to electricity: `Ant_Stats.xls`
 - b) Classified territorial units as output from the DU-TUC Tool `classified_polygons.shp`
- 2) The required input data are provided in the `SDG_training_exercise.zip` file
- 3) Expand the archive `SDG_training_exercise.zip` file in a local folder of your preference. Remember the path where you extracted the archive and use it throughout the exercise whenever you find path as `<localpath>`

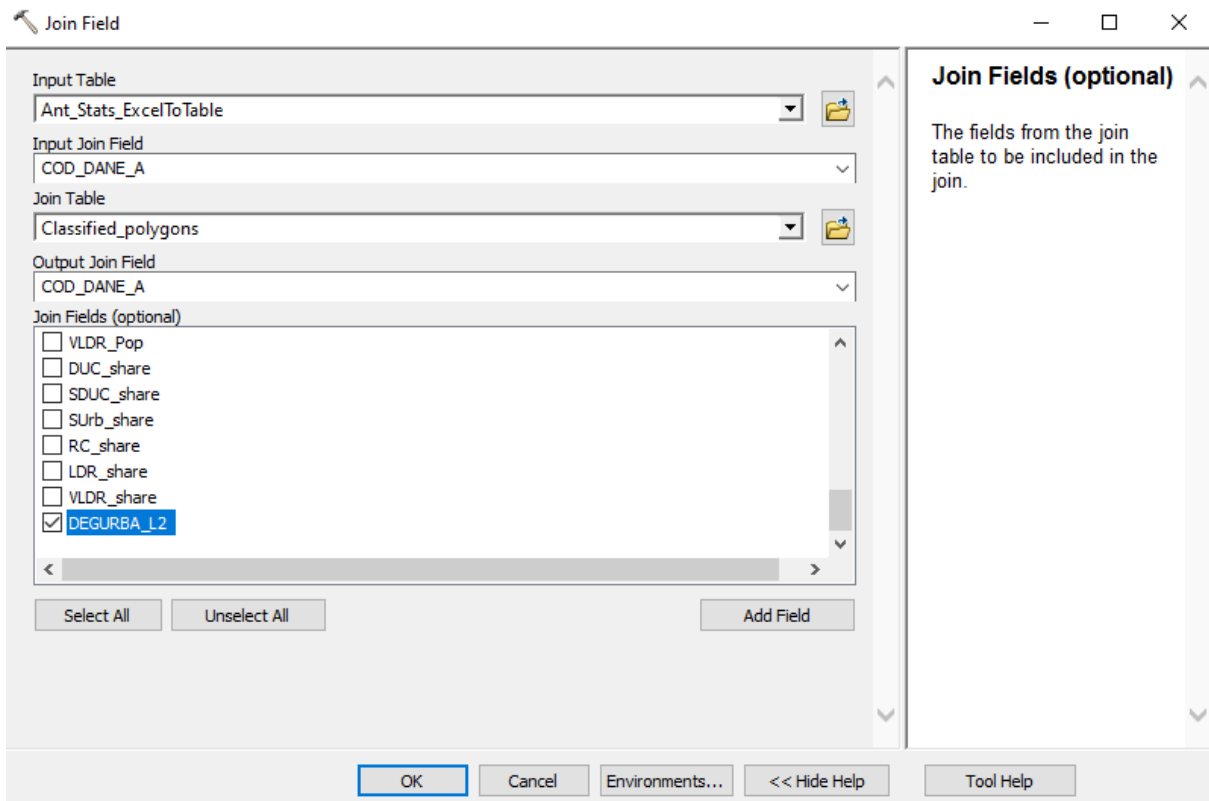
The file required by this exercise are located in the folder: `<localpath>\TESTSDG\input`

2.2 Practical steps

- 1) Open File Explorer
 - a) Browse to the folder `<localpath>\TESTSDG\`
 - b) Create the subfolder `Results`
- 2) Launch QGIS (or ArcMap)
 - a) Load the spreadsheet `Ant_Stats.xls` by dragging it to the canvas in QGIS.
 - a. In ArcMap, you need to convert it to a table readable by the software first. To do so, open the ArcToolbox in ArcMap and select `Conversion tools > Excel > Excel to Table`

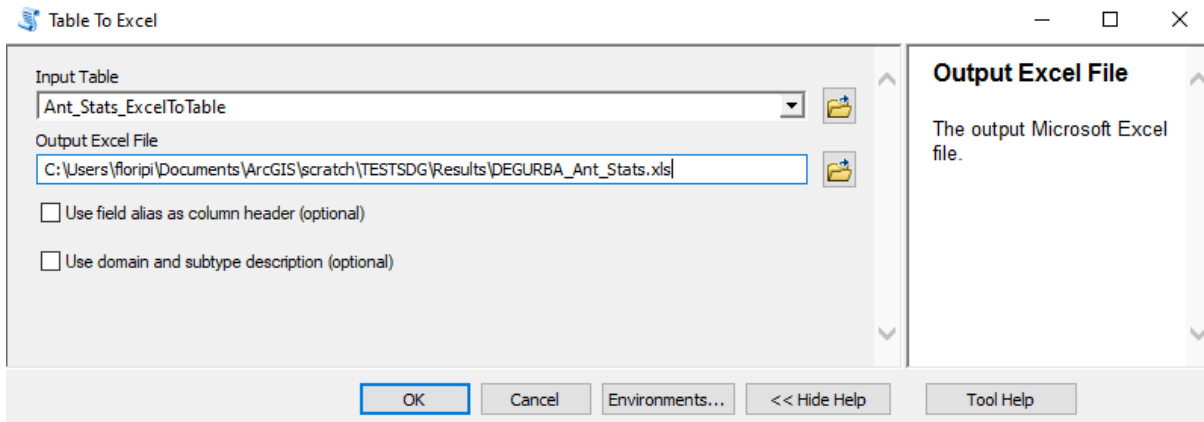



- b) Load the GHS-DU-TUC output `classified_polygons.shp` by dragging it to the canvas
 - a. In QGIS, you can set the correct symbology colors by right clicking on the `classified_polygons.shp`, selecting Properties > Symbology and Categorized from the dropdown menu, then clicking on the Style button on the bottom of the interface and Load style, then clicking on the three dots [...] and pointing to `localpath>\TESTSDG\input\symbology\GHS-DUTUC-L2.sld`
 - b. In ArcMap, you can do this by right clicking on the `classified_polygons.shp` then selecting Properties > Symbology > Categories > Unique Values, then Import, and navigate to files located in the folder `<localpath>\TESTSDG\input\symbology\GHS-DUTUC-L2.lyr`
- c) Join `classified_polygons.shp` to `Ant_Stats.xls` or `Ant_Stats_ExcelToTable` via Vector General > Join attributes by field value
 - a. In ArcMap, select ArcToolbox and then Data Management Tools > Joins > Join Field.
- d) Select `Ant_Stats` or `Ant_Stats_ExcelToTable` as Input table and `Classified_polygons` as Join table, `COD_DANE_A` as table field to base on, `DEGURBA_L1` and `DEGURBA_L2` as fields to copy.

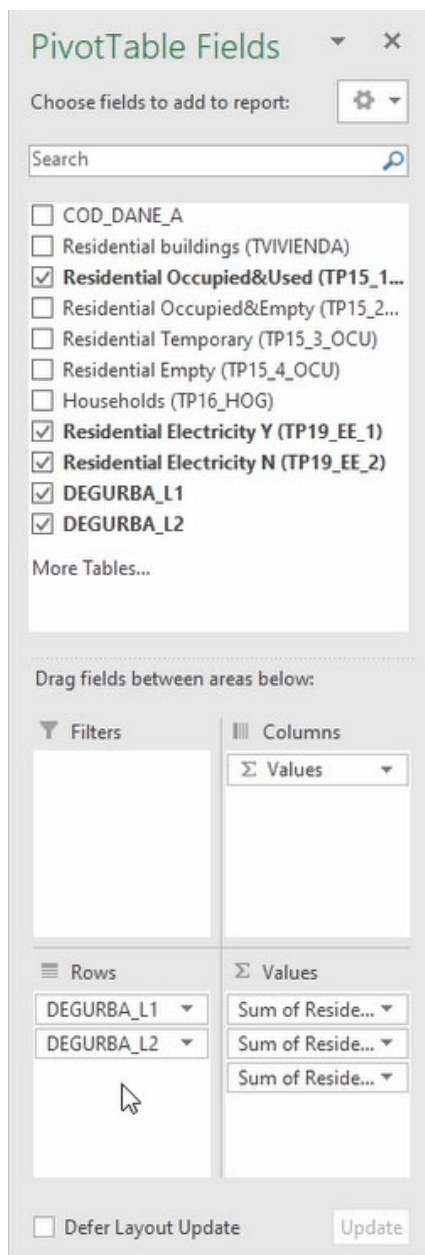


- e) Click Ok or Run to run the join operation

- f) Verify that the table Ant_Stats or Ant_Stats_ExcelToTable has 51235 features after the join operation
- g) In QGIS, Right click on the “Joined layer” that appears on the layer list and select Export > Save Features As...
 - a. In ArcMap, open the ArcToolbox and select Conversion tools > Excel > Table to Excel
- h) In QGIS, save the file as MS Office Open XML spreadsheet [XLSX] to the folder <localpath>\TESTSDG\Results\DEGURBA_Ant_Stats.xlsx
 - a. In ArcMap, save it as MS Excel 1993-2007 XLS to the folder <localpath>\TESTSDG\Results\DEGURBA_Ant_Stats.xls



- 3) Launch MS Excel
 - a) Load the result
`<localpath>\TESTSDG\results\DEGURBA_Ant_Stats.xls(x)`
 - b) Select all cells by clicking on the select all button  (or via the keyboard shortcut Ctrl + A)
 - c) Click Insert > Pivot Table, select you want to place it to a New worksheet and click OK
 - d) Drag the following fields to the Values box and set the aggregation function to Sum in the Value Field Settings accessible by clicking on the field item
 - i) Residential Occupied&Used
 - ii) Residential Electricity Y
 - iii) Residential Electricity N
 - e) Drag the DEGURBA_L1 and DEGURBA_L2 fields to the Rows box



- f) Copy the cells of the Pivot table and paste them to a new worksheet named “SDG stats” as formatted values.



	A	B	C	D
1				
2				
3	Row Labels	Sum of Residential Occupied&Used (TP15_1_OCU)	Sum of Residential Electricity Y (TP19_EE_1)	Sum of Residential Electricity N (TP19_EE_2)
4	=1	335029	320512	14517
5	11	29920	25537	4383
6	12	167216	160672	6544
7	13	137893	134303	3590
8	=2	406769	402433	4336
9	21	94132	92939	1193
10	22	34293	33879	414
11	23	278344	275615	2729
12	=3	1291716	1285682	6034
13	30	1291716	1285682	6034
14	= (blank)			
15	(blank)			
16	Grand Total	2033514	2008627	24887

The row numbers correspond to the settlement classes, according to the encoding of the Degree of Urbanisation standard

1 st Level DoU		2 nd Level DoU	
Urban Area	<div>UC</div> <div>3</div> <div>Urban Centre High-density cluster (City)</div>	<div>UC</div> <div>30</div> <div></div>	
	<div>UCI</div> <div>2</div> <div>Urban Cluster Moderate-density cluster (Town & semi-dense area)</div>	<div>DUC</div> <div>23</div> <div>Dense Urban Cluster (Dense Town)</div>	
		<div>SDUC</div> <div>22</div> <div>Semi-Dense Urban Cluster (Semi-dense Town)</div>	
		<div>SBRS</div> <div>21</div> <div>Suburban / Peri-urban grid cell (Suburban / Peri-urban area)</div>	
	<div>RUR</div> <div>1</div> <div>Rural grid cell Low-density cluster (Rural area)</div>	<div>RC</div> <div>13</div> <div>Rural Cluster (Village)</div>	
		<div>LDR</div> <div>12</div> <div>Low density rural grid cell (Dispersed rural density area)</div>	
		<div>VLDR</div> <div>11</div> <div>Very low density rural grid cell (Mostly uninhabited area)</div>	
<div>W</div> <div>10</div> <div>Water</div>			

- g) Copy the cells of the Pivot table from the worksheet named “SDG stats” and paste them on the same worksheet below
- h) Compute the relative frequencies of Residential Electricity Y and Residential Electricity N with respect to the total residential households

	A	B	C	D
1	Row Labels	Sum of Residential Occupied&Used (TP15_1_OCU)	Sum of Residential Electricity Y (TP19_EE_1)	Sum of Residential Electricity N (TP19_EE_2)
2	1	335029	320512	14517
3	11	29920	25537	4383
4	12	167216	160672	6544
5	13	137893	134303	3590
6	2	406769	402433	4336
7	21	94132	92939	1193
8	22	34293	33879	414
9	23	278344	275615	2729
10	3	1291716	1285682	6034
11	30	1291716	1285682	6034
12	(blank)			
13	(blank)			
14	Grand Total	2033514	2008627	24887
15				
16				
17				
18				
19	Row Labels	Sum of Residential Occupied&Used (TP15_1_OCU)	Sum of Residential Electricity Y (TP19_EE_1)	Sum of Residential Electricity N (TP19_EE_2)
20	1	=B2/\$B2	96%	4%
21	11	100%	85%	15%
22	12	100%	96%	4%
23	13	100%	97%	3%
24	2	100%	99%	1%
25	21	100%	99%	1%
26	22	100%	99%	1%
27	23	100%	99%	1%
28	3	100%	100%	0%
29	30	100%	100%	0%
30	(blank)			
31	(blank)			

- i) Plot the percentage breakdowns of Residential Electricity Y and Residential Electricity N on a bar plot or format the cells as data bars, to visualise the results

18				
19	Row Labels	Sum of Residential Occupied&Used (TP15_1_OCU)	Sum of Residential Electricity Y (TP19_EE_1)	Sum of Residential Electricity N (TP19_EE_2)
20	1	100%	96%	4%
21	11	100%	85%	15%
22	12	100%	96%	4%
23	13	100%	97%	3%
24	2	100%	99%	1%
25	21	100%	99%	1%
26	22	100%	99%	1%
27	23	100%	99%	1%
28	3	100%	100%	0%
29	30	100%	100%	0%
30	(blank)			
31	(blank)			
32				

END