

Quick Start Guide to Facebook Co-Location Data for Level 3 Administrative/Geographical Units

Based on 03/13/2020 Release

Shenyue Jia, Ph.D., Chapman University, Orange, California USA
sjia@chapman.edu

Created on 03/14/2020

Revised on 03/15/2020

0. LEVEL 3 ADMINISTRATIVE/GEOGRAPHICAL UNIT OF A COUNTRY

	Meaning	Examples
Level 1*	Country/Region	
Level 2	Largest admin/geo unit within a country	State (US), province (China), prefecture (Japan), etc.
Level 3	Largest admin/geo unit within a level 2 unit	County (US), city (China), city (Japan), etc.

* Some geographical data providers such as [GADM](#) refer country as level 0, state as level 1, and county as level 2.

Some resources for polygon type level 3 admin/geo unit

1. [GADM](#) provides geospatial boundaries for almost every country in the world in multiple formats, including the format that can be directly read by the sp package of R. One thing to note for GADM data is that it defines Country as Level 0 unit. In this case, the level 3 admin/geo unit used in FB co-location data is level 2 data from GADM.
2. [Esri Living Atlas](#) provides geospatial boundaries for almost every country with a better quality and extra demographic information. Sometimes it even provides level 4 unit boundaries.
3. [DIVA GIS](#) groups geospatial data by country, including geographical boundaries and else.
4. [Humanitarian Data Exchange](#) also provides abundant administrative boundaries and other relevant data, especially the High Resolution Population Density by Facebook and CIESIN of Columbia University.

1. VARIABLES AVAILABLE FROM FACEBOOK CO-LOCATION DATA

	Variable	Indication
Auxiliary Variables	polygon1_ID	ID of the level 3 “From” polygon
	polygon1_name	Name of the level 3 “From” polygon
	lon_1	Longitude of he level 3 “From” polygon
	lat_1	Latitude of the level 3 “From” polygon
	name_stack_1	Name of the level 3 “From” polygon with the level 2 polygon it belongs to
	polygon1_ID	ID of the level 3 “From” polygon
	polygon1_name	Name of the level 3 “From” polygon
	lon_1	Longitude of he level 3 “From” polygon
	lat_1	Latitude of the level 3 “From” polygon
	name_stack_1	Name of the level 3 “From” polygon with the level 2 polygon it belongs to
Thematic Variables	ds	Time stamp
	country	Country
	fb_population_1	The number of completed trajectories from the level 3 admin/geo unit where the link STARTS
	fb_population_2	The number of completed trajectories from the level 3 admin/geo unit where the link ENDS
	link_value	Probability of colocation between polygon 1 (from) and polygon 2 (to): (# observed colocation events in a week) / (# possible colocation events in a week)

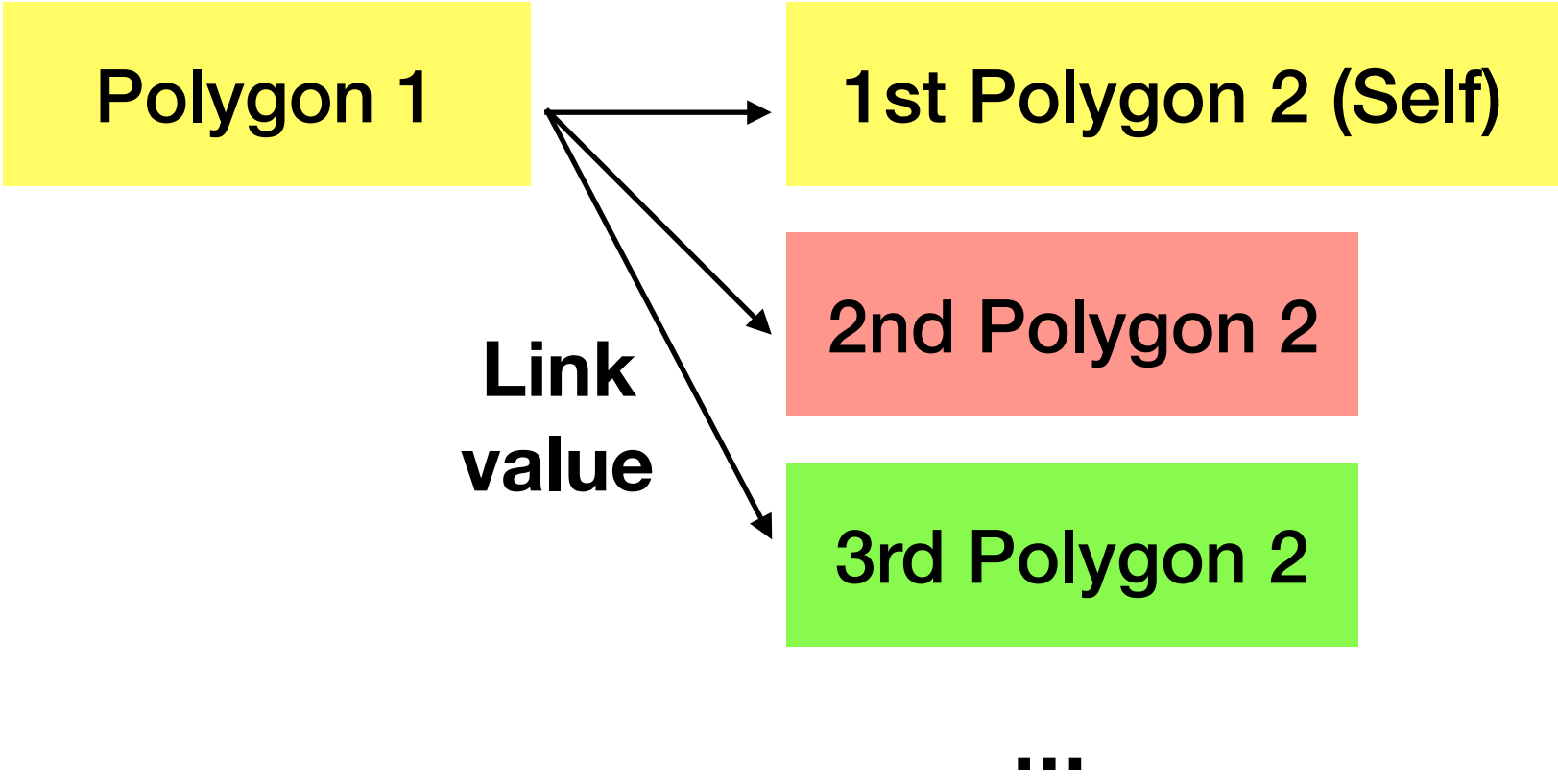
2. UNDERSTAND THE CO-LOCATION DATA TABLE STRUCTURE

Co-Location Map Example

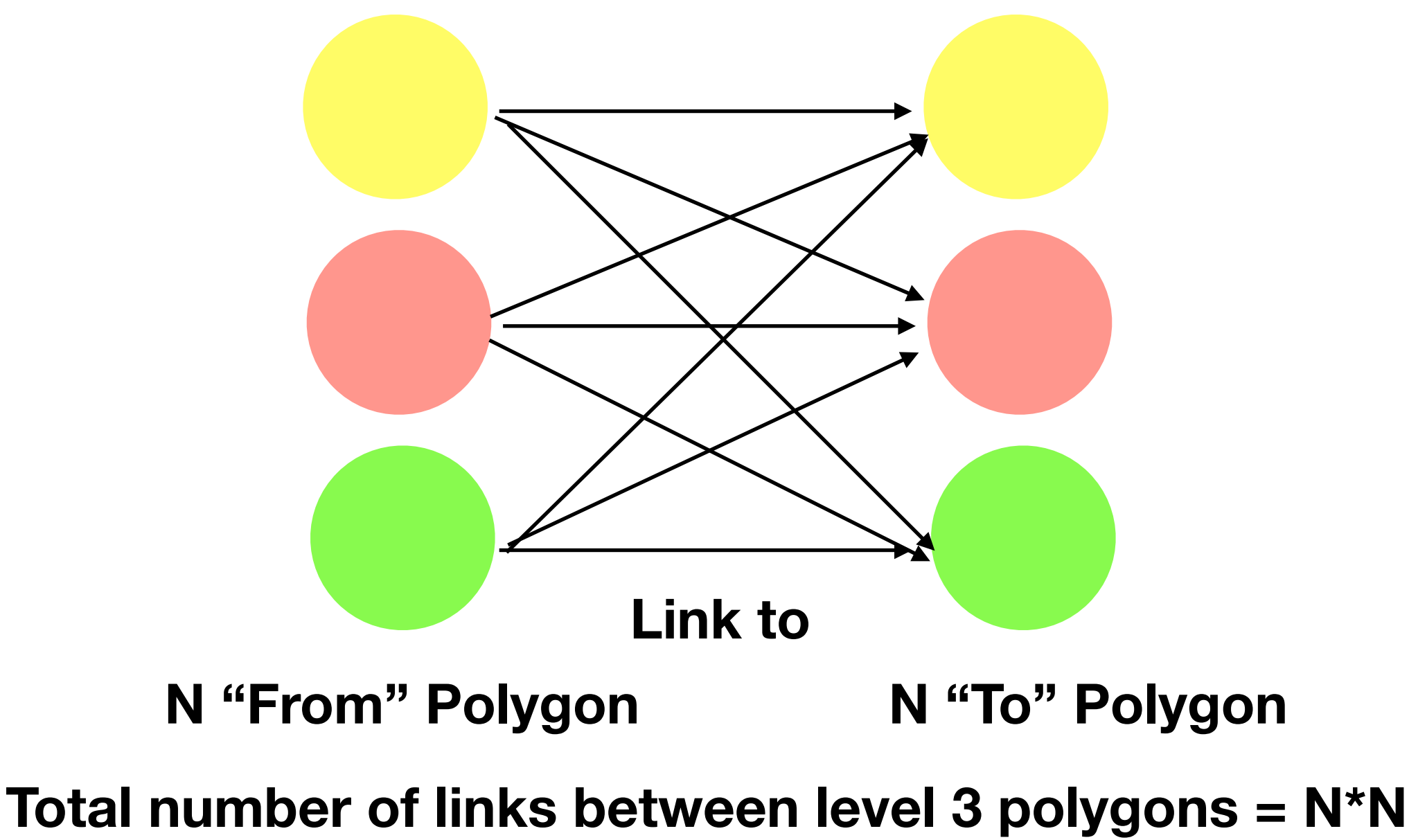
polygon1_id	polygon1_name	lon_1	lat_1	name_stack_1	fb_population_1	polygon2_id	polygon2_name	lon_2	lat_2	name_stack_2	fb_population_2	link_value	country	ds
881048	A Coruña	-8.4641	43.12576	A Coruña	11164	881048	A Coruña	-8.46417	43.1257	A Coruña	11164	0.00163054	ESP	2020-03
881048	A Coruña	-8.4641	43.12576	A Coruña	11164	881035	Albacete	-1.98053	38.8251	Castille-La Mancha /	5004	1.52917056	ESP	2020-03
881048	A Coruña	-8.4641	43.12576	A Coruña	11164	881036	Alicante	-0.56854	38.4785	Community of Valenc	29483	3.03642701	ESP	2020-03
881035	Albacete	-1.9805	38.82511	Castille-La Manc	5004	881048	A Coruña	-8.46417	43.1257	A Coruña	11164	0.00163054	ESP	2020-03
881035	Albacete	-1.9805	38.82511	Castille-La Manc	5004	881035	Albacete	-1.98053	38.8251	Castille-La Mancha /	5004	1.52917056	ESP	2020-03
881035	Albacete	-1.9805	38.82511	Castille-La Manc	5004	881036	Alicante	-0.56854	38.4785	Community of Valenc	29483	3.03642701	ESP	2020-03
881036	Alicante	-0.5685	38.47859	Community of Va	29483	881048	A Coruña	-8.46417	43.1257	A Coruña	11164	0.00163054	ESP	2020-03
881036	Alicante	-0.5685	38.47859	Community of Va	29483	881035	Albacete	-1.98053	38.8251	Castille-La Mancha /	5004	1.52917056	ESP	2020-03
881036	Alicante	-0.5685	38.47859	Community of Va	29483	881036	Alicante	-0.56854	38.4785	Community of Valenc	29483	3.03642701	ESP	2020-03

Information of Polygon 1 (From Polygon)

Information of Polygon 2 (To Polygon)



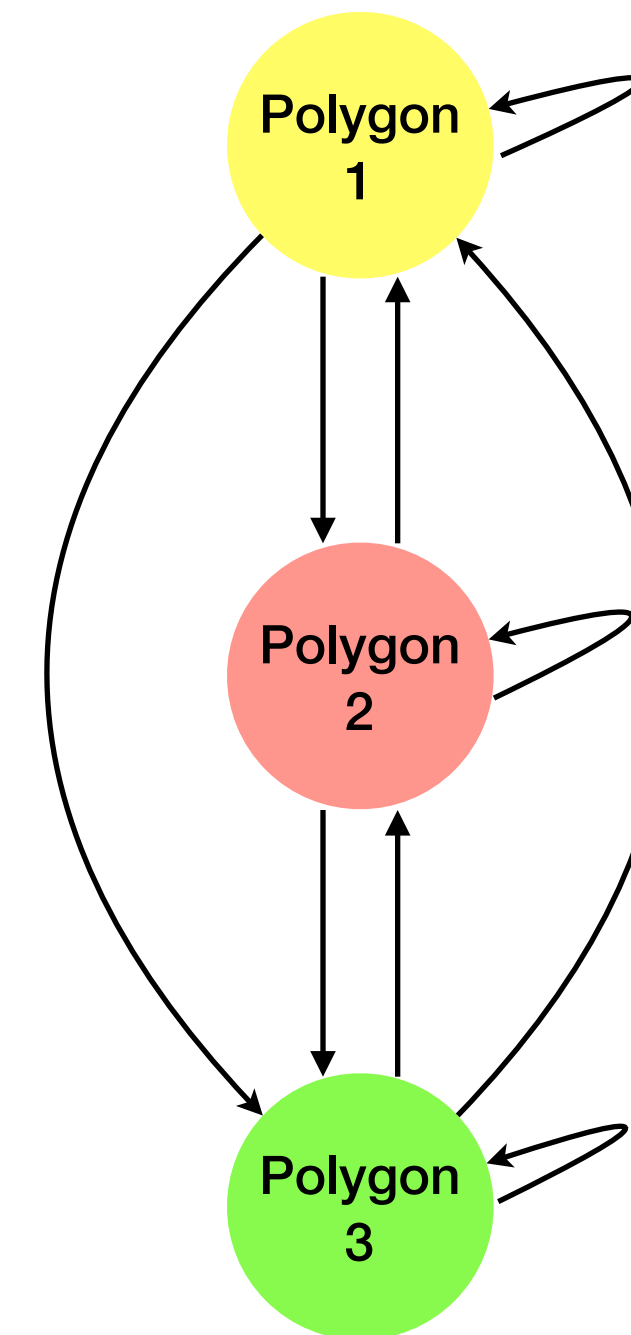
One “From” polygon has *n* “To” polygon, *n* equals to the total number of level 3 admin/geo unit polygons available for the COUNTRY, including the “From” polygon itself.



3. UNDERSTAND THE LINK_VALUE FIELD

- $\text{link_value} = \text{Number of observed colocation events in a week} / (\text{Number of possible colocation events in a week})$.
- If polygon 1 and 2 are different,
 - $\text{link_value} = (\text{Number of five minute colocations in a week for polygon 1 and polygon 2}) / (\text{population from polygon 1} * \text{population from polygon 2} * \text{number of five minute bins in a week})$.
- If polygon 1 is the same as polygon 2,
 - $\text{link_value} = (\text{Number of five minute colocations in a week within polygon 1}) / ((\text{population from polygon 1}) * (\text{population from polygon}_1 - 1) / 2 * \text{number of five minute bins in a week})$.

- The denominator is the maximum number of colocations observed in the week. The link values are hence always **SYMMETRIC** as well.



4. MAP FACEBOOK POPULATION DATA FOR EACH LEVEL 3 ADMIN/GEO UNIT

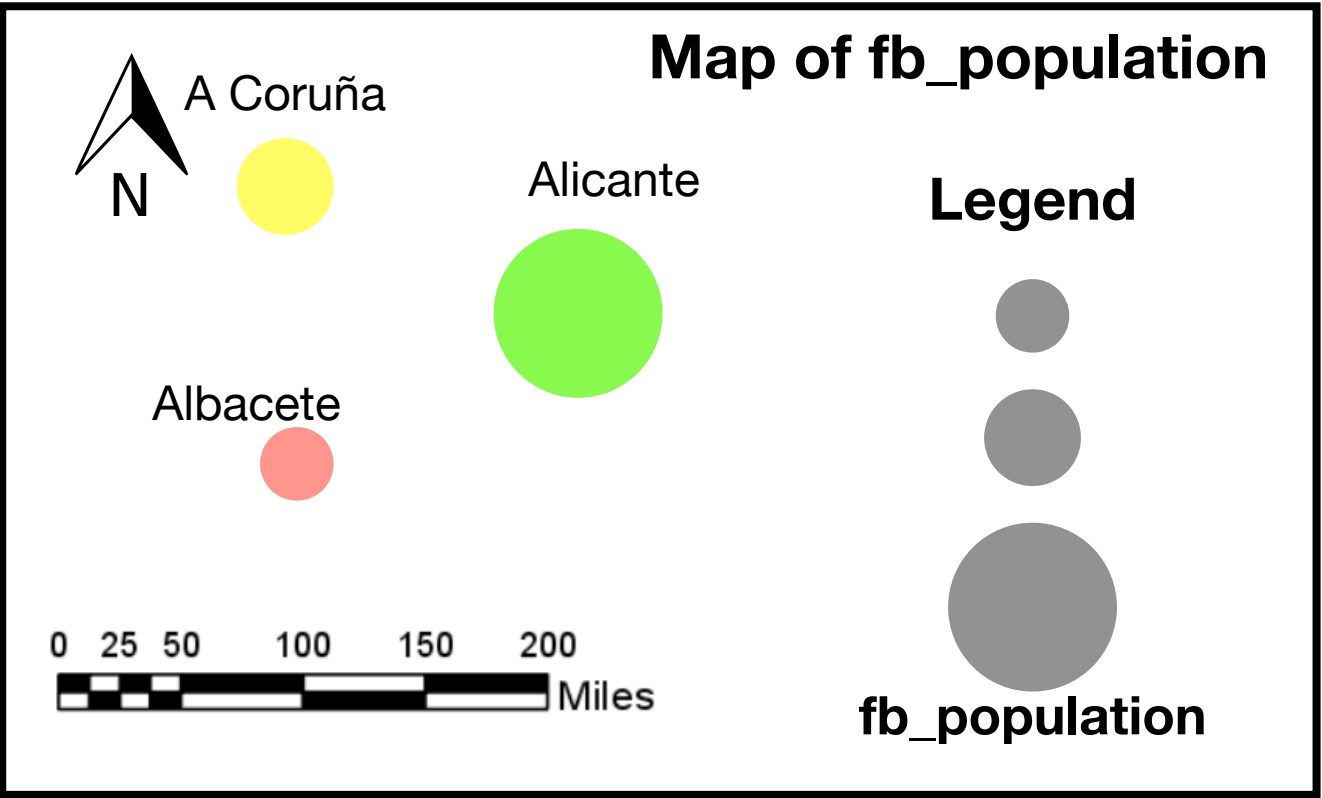
Co-Location Map Example

polygon1_id	polygon1_name	lon_1	lat_1	name_stack_1	fb_population_1	polygon2_id	polygon2_name	lon_2	lat_2	name_stack_2	fb_population_2	link_value	country	ds
881048	A Coruña	-8.4641	43.1257	A Coruña	11164	881048	A Coruña	-8.464172	43.1257	A Coruña	11164	0.00163054	ESP	2020-03
881048	A Coruña	-8.4641	43.1257	A Coruña	11164	881035	Albacete	-1.980530	38.8251	Castille-La Mancha /	5004	1.52917056	ESP	2020-03
881048	A Coruña	-8.4641	43.1257	A Coruña	11164	881036	Alicante	-0.568540	38.4785	Community of Valenc	29483	3.03642701	ESP	2020-03
881035	Albacete	-1.9805	38.8251	Castille-La Manc	5004	881048	A Coruña	-8.464172	43.1257	A Coruña	11164	0.00163054	ESP	2020-03
881035	Albacete	-1.9805	38.8251	Castille-La Manc	5004	881035	Albacete	-1.980530	38.8251	Castille-La Mancha /	5004	1.52917056	ESP	2020-03
881035	Albacete	-1.9805	38.8251	Castille-La Manc	5004	881036	Alicante	-0.568540	38.4785	Community of Valenc	29483	3.03642701	ESP	2020-03
881036	Alicante	-0.5685	38.4785	Community of Va	29483	881048	A Coruña	-8.464172	43.1257	A Coruña	11164	0.00163054	ESP	2020-03
881036	Alicante	-0.5685	38.4785	Community of Va	29483	881035	Albacete	-1.980530	38.8251	Castille-La Mancha /	5004	1.52917056	ESP	2020-03
881036	Alicante	-0.5685	38.4785	Community of Va	29483	881036	Alicante	-0.568540	38.4785	Community of Valenc	29483	3.03642701	ESP	2020-03



Polygon Information Example

polygon_id	polygon_name	lon	lat	name_stack	fb_population
881048	A Coruña	-8.464172	43.1257	A Coruña	11164
881035	Albacete	-1.980530	38.8251	Castille-La Mancha /	5004
881036	Alicante	-0.568540	38.4785	Community of Valenc	29483



- The information of each “From” polygon is repeated for N times in the table.
- Remember to obtain **UNIQUE** value of each polygon information column to reduce the table to N rows to map Facebook Population of each polygon

5. MAP LINK VALUE BETWEEN LEVEL 3 ADMIN/GEO UNIT

Co-Location Map Example

polygon1_id	polygon1_name	lon_1	lat_1	name_stack_1	fb_population_1	polygon2_id	polygon2_name	lon_2	lat_2	name_stack_2	fb_population_2	link_value	country	ds
881048	A Coruña	-8.4641	43.1257	A Coruña	11164	881048	A Coruña	-8.464172	43.1257	A Coruña	11164	0.00163054	ESP	2020-03
881048	A Coruña	-8.4641	43.1257	A Coruña	11164	881035	Albacete	-1.980530	38.8251	Castille-La Mancha /	5004	1.52917056	ESP	2020-03
881048	A Coruña	-8.4641	43.1257	A Coruña	11164	881036	Alicante	-0.568540	38.4785	Community of Valenc	29483	3.03642701	ESP	2020-03
881035	Albacete	-1.9805	38.8251	Castille-La Manc	5004	881048	A Coruña	-8.464172	43.1257	A Coruña	11164	1.52917056	ESP	2020-03
881035	Albacete	-1.9805	38.8251	Castille-La Manc	5004	881035	Albacete	-1.980530	38.8251	Castille-La Mancha /	5004	0.00548153	ESP	2020-03
881035	Albacete	-1.9805	38.8251	Castille-La Manc	5004	881036	Alicante	-0.568540	38.4785	Community of Valenc	29483	1.18484643	ESP	2020-03
881036	Alicante	-0.5685	38.4785	Community of Va	29483	881048	A Coruña	-8.464172	43.1257	A Coruña	11164	3.03642701	ESP	2020-03
881036	Alicante	-0.5685	38.4785	Community of Va	29483	881035	Albacete	-1.980530	38.8251	Castille-La Mancha /	5004	1.18484643	ESP	2020-03
881036	Alicante	-0.5685	38.4785	Community of Va	29483	881036	Alicante	-0.568540	38.4785	Community of Valenc	29483	0.00129893	ESP	2020-03

- Assign a unique **ID of links** (row in the table)
- Use lon_1, lat_1, lon_2, lat_2 **AND the ID** of link to create line features/objects (start and end point coordinates) by
 - XY To Line (ArcGIS Pro/ArcMap)
 - Function developed using sp and maptools library (R Language)
 - Python GDAL/OGR
 - ...
- Add link_value from the table to the newly created line features/objects with the shared ID of link
 - Or simply add this while creating the link if your tool permits you to do so

