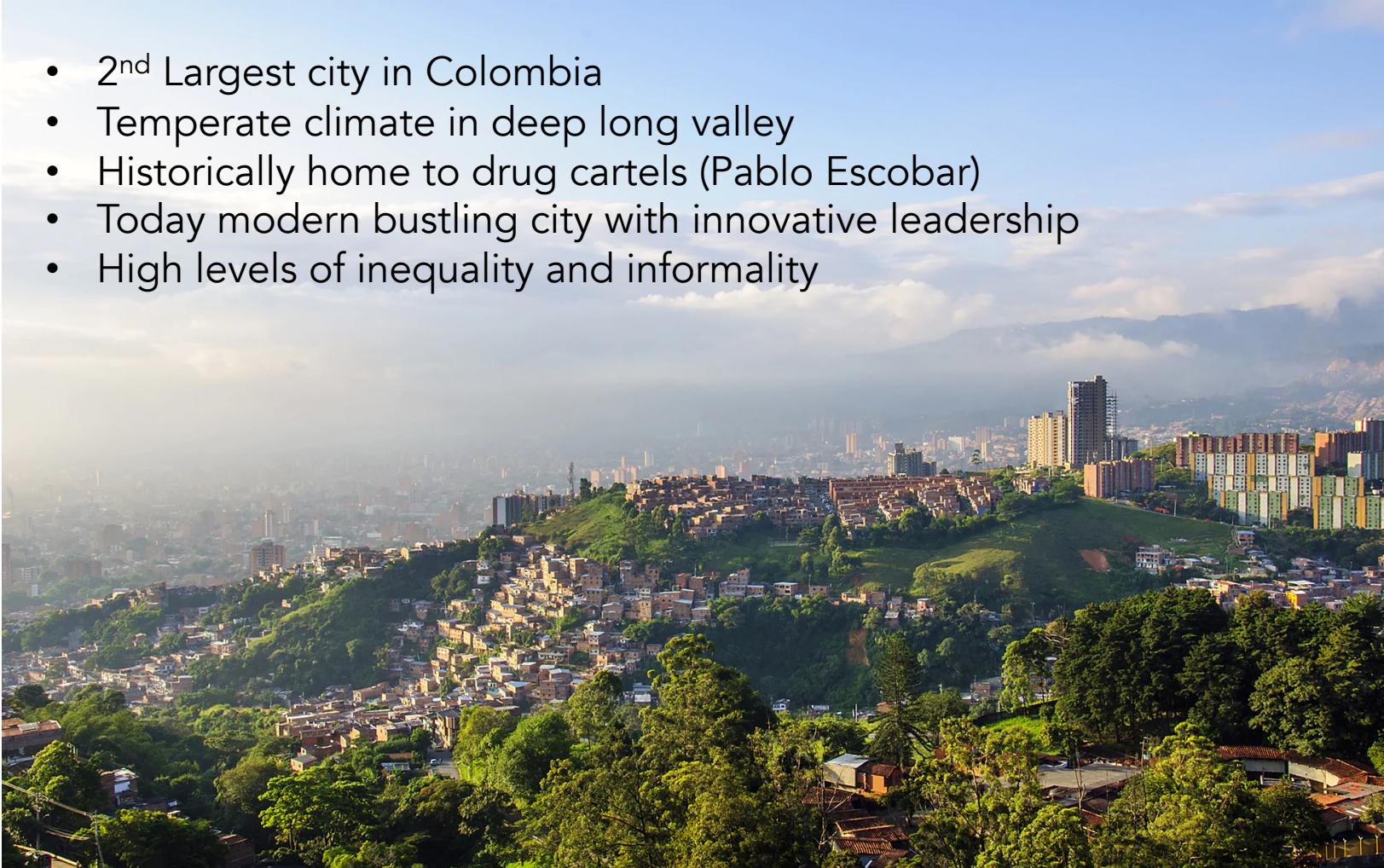


Medellin, Colombia

- 2nd Largest city in Colombia
- Temperate climate in deep long valley
- Historically home to drug cartels (Pablo Escobar)
- Today modern bustling city with innovative leadership
- High levels of inequality and informality



Medellín, Colombia



- Google street view/locations of informal firms in Medellin: <https://royalsocietypublishing.org/doi/10.1098/rsos.211841>
- Bogota GPS mobile phone data applied to study amenities and mobility: <https://journals.sagepub.com/doi/10.1177/23998083221141439>
- Original formality paper (formal firms are complex, predicting growth in formality rate across cities): <https://dash.harvard.edu/handle/1/37366600>
- Spin-off looking at commuting times to pool skills to generate formal employment in cities: <https://link.springer.com/article/10.1007/s41109-019-0235-z>
- CDR mobile phone data applied to study lockdown rules in Colombia: <https://royalsocietypublishing.org/doi/full/10.1098/rsif.2020.1035?af=R>
- World Bank paper (predicting employment at high spatial resolution in any city): <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/660611647960970611/where-are-all-the-jobs-a-machine-learning-approach-for-high-resolution-urban-employment-prediction-in-developing-countries>
- Inter-city migration in Colombia (using administrative data on formal employment): <https://www.sciencedirect.com/science/article/pii/S0198971522000187?dgcid=coauthor>

- The government of Colombia has just announced funding for a new metro station in Medellin, Colombia.
 - Medellin is the country's second largest city, composed of nearly 4 million people across 10 municipalities and 66 communes (administrative region).
 - You are a development consultancy and your task is to advise which comuna the new stop should be built in.
 - Comuna leaders are already squabbling: they all have arguments as to why their comuna would benefit the most from the new metro-stop.
 - You should give a data-driven assessment of which of the 66 communes should be awarded the metro-stop (and why), and how it should be connected to the existing infrastructure.
-
- You can assume that building a station anywhere costs the same, but the cost of a rail-line is approximately proportional to the length of the line.
 - Your answer should provide basic information regarding the location of the stop and how the metro stop is connected to the current network (what line it is on, and which stop or stops it is adjacent to).
 - It is acceptable to suggest building a metro-stop in a comuna that already has a stop, if your analysis supports this decision.
-
- Your group will have 8 mins to present your recommendation to the Colombian National Planning Department on Friday.



SITVA

Sistema Integrado de
Transporte del Valle del Aburrá

MAPA METRO

LÍNEAS DEL SISTEMA

METRO

Línea A Niquía - La Estrella - Niquía
Línea B San Antonio - San Javier - San Antonio

CABLE

Línea J San Javier - La Aurora - San Javier
Línea K Acevedo - Santo Domingo - Acevedo
Línea L Santo Domingo - Arví - Santo Domingo

BÚS

Línea 1 U. de M. - Av. del Ferrocarril - Parque Aranjuez - Av. del Ferrocarril - U. de M.
Línea 2 U. de M. - Av. Oriental - Parque Aranjuez - Av. Oriental - U. de M.

CONVENCIONES

- Estación
- Paradero
- Transferencia directa
- Transferencia peatonal
- Ruta integrada
- Parqueaderos para bicicletas
- PAC Punto de Atención al Cliente
- Todas las estaciones son accesibles
- En construcción
- Estación fuera de servicio

HORARIOS DE OPERACIÓN

- Lunes a sábado
4:30 a.m. - 11:00 p.m.
Domingos a festivos
5:00 a.m. - 10:00 p.m.
- Línea L (Cable turístico)
9:00 a.m. - 6:00 p.m.
No presta servicio el primer día hábil de la semana)

MAYORES INFORMES

HOLA METRO 444 95 98



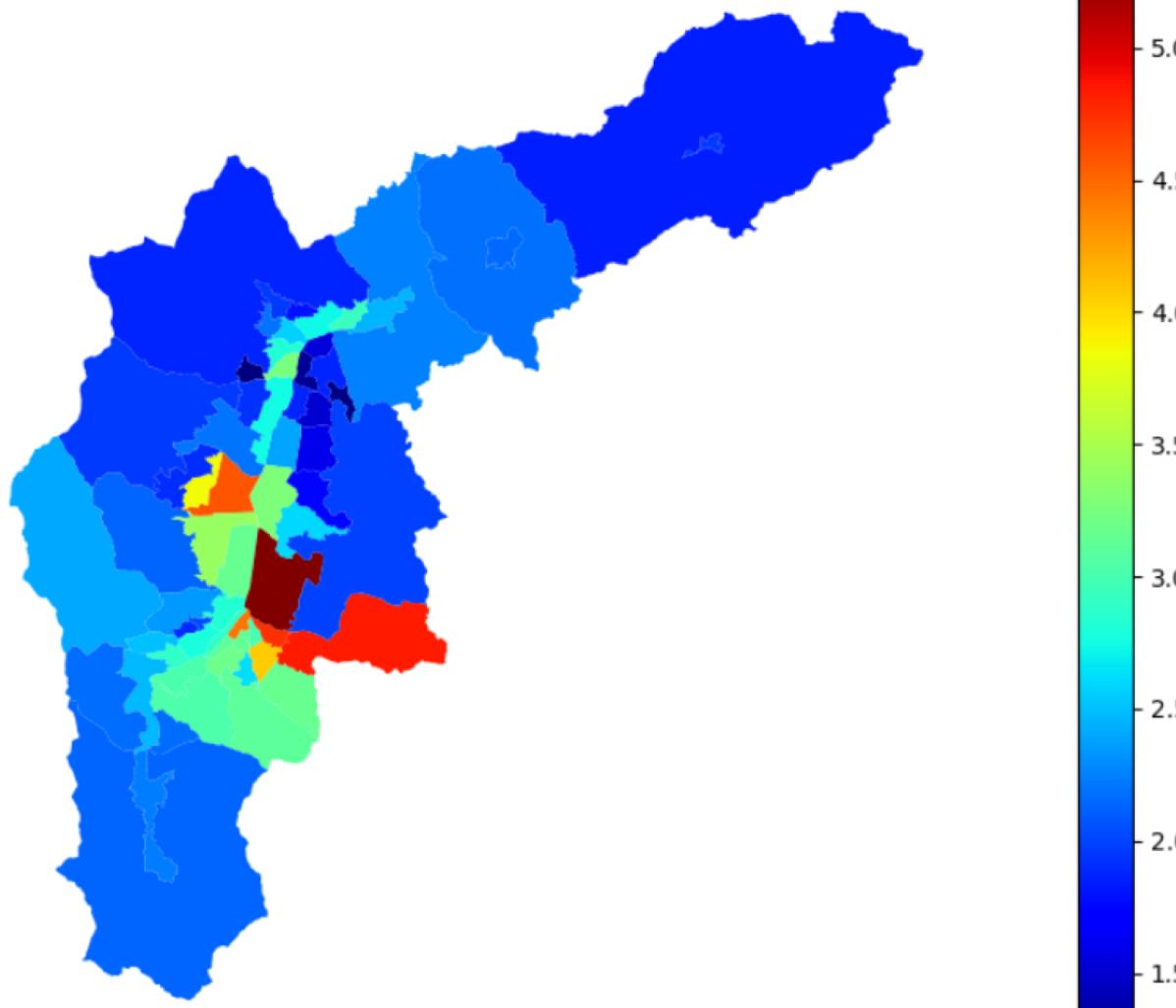
• This article is more than **6 years old**

'If I'm stratum 3, that's who I am': inside Bogotá's social stratification system

Every district in Colombia's capital is rated 1 to 6 for affluence, and its services subsidised accordingly. But is a laudable idea creating division and stigma?



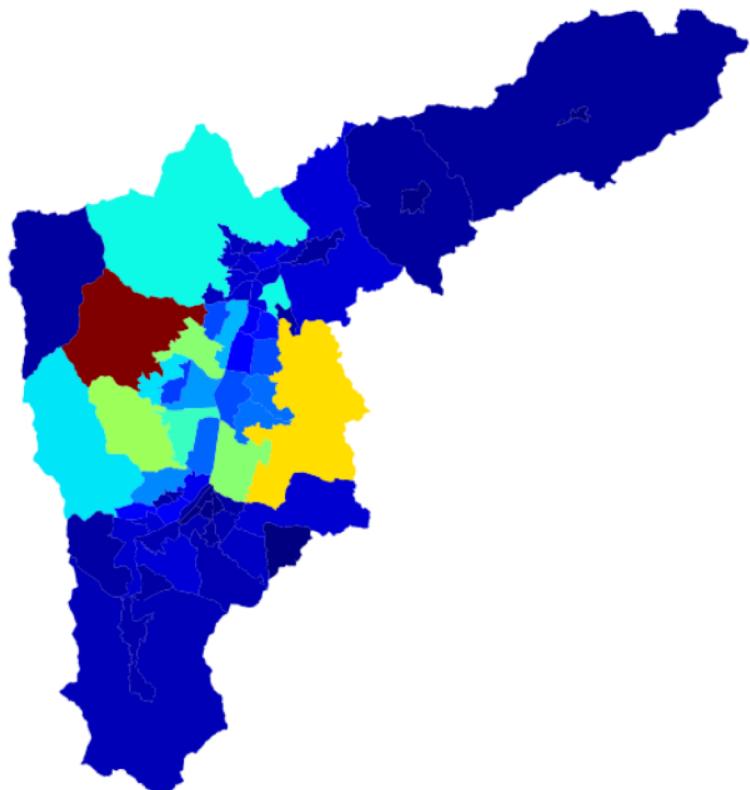
The Bogotá *barrio* of El Oasis, ... residents of poorer neighbourhoods pay less for water, telephone bills and rubbish collection than those in wealthy areas. Photograph: Alamy



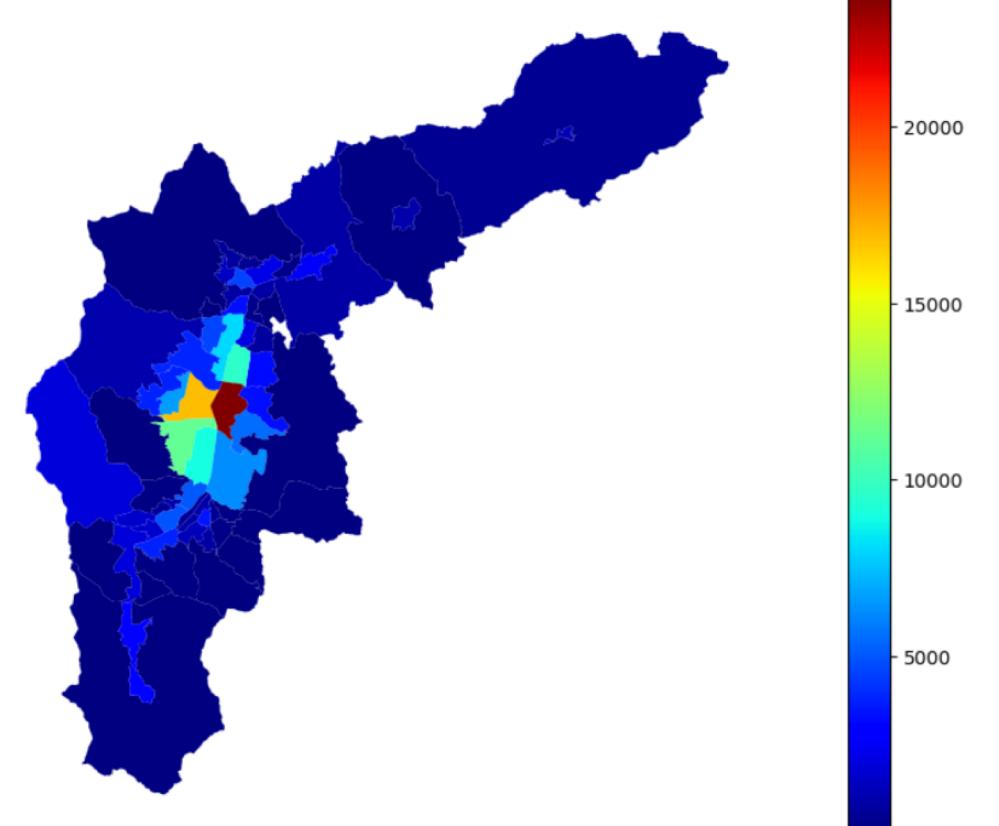
Ella Jessel in Bogotá

Thu 9 Nov 2017 07.30 GMT

Population



Visible commercial firms



Straulino, D., Saldarriaga, J. C., Gómez, J. A., Duque, J. C., & O'Clery, N. (2022). Uncovering commercial activity in informal cities. *Royal Society Open Science*, 9(11), 211841.

Data Sources:

- The Medellin metro system comprises nearly 100 stops across subway, tram, bus rapid transit (BRT), and cable cars (tramway). Daily ridership averages 1,000,000+. The current network is given in “**medellin_metro.csv**” each row represents a station, complete with its latitude and longitude coordinates, the comuna it is in, and what line it is on (as well as the type of line). Transfer stations are indicated. Successive stops along the same line are successive in the file. For instance, Sabaneto and La Estrella are adjacent on Line 1, but La Estrella (Line A) and San Antonio (Line B) are not.

Data Sources:

- The Medellin metro system comprises nearly 100 stops across subway, tram, bus rapid transit (BRT), and cable cars (tramway). Daily ridership averages 1,000,000+. The current network is given in “**medellin_metro.csv**” each row represents a station, complete with its latitude and longitude coordinates, the comuna it is in, and what line it is on (as well as the type of line). Transfer stations are indicated. Successive stops along the same line are successive in the file. For instance, Sabaneto and La Estrella are adjacent on Line 1, but La Estrella (Line A) and San Antonio (Line B) are not.
- The file “**comunas_commuters.csv**” contains the relative number of trips between each comuna and has been constructed from mobile phone records (tower pings). For each row, there are three columns - the first gives the origin (an index), the second gives the destination (another index), and the final column gives the relative number of trips. The file “**comunas.csv**” matches the index with the corresponding comunas.

Data Sources:

- The Medellin metro system comprises nearly 100 stops across subway, tram, bus rapid transit (BRT), and cable cars (tramway). Daily ridership averages 1,000,000+. The current network is given in "**medellin_metro.csv**" each row represents a station, complete with its latitude and longitude coordinates, the comuna it is in, and what line it is on (as well as the type of line). Transfer stations are indicated. Successive stops along the same line are successive in the file. For instance, Sabaneto and La Estrella are adjacent on Line 1, but La Estrella (Line A) and San Antonio (Line B) are not.
- The file "**comunas_commutes.csv**" contains the relative number of trips between each comuna and has been constructed from mobile phone records (tower pings). For each row, there are three columns - the first gives the origin (an index), the second gives the destination (another index), and the final column gives the relative number of trips. The file "comunas.csv" matches the index with the corresponding comunas.
- "**comunas_gmaps.csv**" contains the driving times, transit times and walking times (in seconds) between the different comunas which come from Google Maps. Edge (i,j) is how long it takes to commute from comuna i to j (at 9am in the morning during the week).

Data Sources:

- The Medellin metro system comprises nearly 100 stops across subway, tram, bus rapid transit (BRT), and cable cars (tramway). Daily ridership averages 1,000,000+. The current network is given in “**medellin_metro.csv**” each row represents a station, complete with its latitude and longitude coordinates, the comuna it is in, and what line it is on (as well as the type of line). Transfer stations are indicated. Successive stops along the same line are successive in the file. For instance, Sabaneto and La Estrella are adjacent on Line 1, but La Estrella (Line A) and San Antonio (Line B) are not.
- The file “**comunas_commutes.csv**” contains the relative number of trips between each comuna and has been constructed from mobile phone records (tower pings). For each row, there are three columns - the first gives the origin (an index), the second gives the destination (another index), and the final column gives the relative number of trips. The file “**comunas.csv**” matches the index with the corresponding comunas.
- “**comunas_gmaps.csv**” contains the driving times, transit times and walking times (in seconds) between the different comunas which come from Google Maps. Edge (i,j) is how long it takes to commute from comuna i to j (at 9am in the morning during the week).
- “**comunas_pop.csv**” gives the average population in 2020 within each comuna.

Data Sources:

- The Medellin metro system comprises nearly 100 stops across subway, tram, bus rapid transit (BRT), and cable cars (tramway). Daily ridership averages 1,000,000+. The current network is given in “**medellin_metro.csv**” each row represents a station, complete with its latitude and longitude coordinates, the comuna it is in, and what line it is on (as well as the type of line). Transfer stations are indicated. Successive stops along the same line are successive in the file. For instance, Sabaneto and La Estrella are adjacent on Line 1, but La Estrella (Line A) and San Antonio (Line B) are not.
- The file “**comunas_commutes.csv**” contains the relative number of trips between each comuna and has been constructed from mobile phone records (tower pings). For each row, there are three columns - the first gives the origin (an index), the second gives the destination (another index), and the final column gives the relative number of trips. The file “**comunas.csv**” matches the index with the corresponding comunas.
- “**comunas_gmaps.csv**” contains the driving times, transit times and walking times (in seconds) between the different comunas which come from Google Maps. Edge (i,j) is how long it takes to commute from comuna i to j (at 9am in the morning during the week).
- “**comunas_pop.csv**” gives the average population in 2020 within each comuna.
- “**comunas_firms.csv**” gives the predicted number of commercial firms that are visible in Google Street Imagery (both formal and informal) in each comuna.

Data Sources:

- The Medellin metro system comprises nearly 100 stops across subway, tram, bus rapid transit (BRT), and cable cars (tramway). Daily ridership averages 1,000,000+. The current network is given in “**medellin_metro.csv**” each row represents a station, complete with its latitude and longitude coordinates, the comuna it is in, and what line it is on (as well as the type of line). Transfer stations are indicated. Successive stops along the same line are successive in the file. For instance, Sabaneto and La Estrella are adjacent on Line 1, but La Estrella (Line A) and San Antonio (Line B) are not.
- The file “**comunas_commutes.csv**” contains the relative number of trips between each comuna and has been constructed from mobile phone records (tower pings). For each row, there are three columns - the first gives the origin (an index), the second gives the destination (another index), and the final column gives the relative number of trips. The file “**comunas.csv**” matches the index with the corresponding comunas.
- “**comunas_gmaps.csv**” contains the driving times, transit times and walking times (in seconds) between the different comunas which come from Google Maps. Edge (i,j) is how long it takes to commute from comuna i to j (at 9am in the morning during the week).
- “**comunas_pop.csv**” gives the average population in 2020 within each comuna.
- “**comunas_firms.csv**” gives the predicted number of commercial firms that are visible in Google Street Imagery (both formal and informal) in each comuna.
- “**comunas_stratum.csv**” contains the average stratum of the neighbourhoods in each comuna. Stratum is a socio-economic indicator that is based on the quality of housing and infrastructure. The highest socio-economic scale is 6 (wealthy areas), and the lowest 1 (poor areas). See:
<https://www.theguardian.com/cities/2017/nov/09/bogota-colombia-social-stratification-system>

Data Sources:

- The Medellin metro system comprises nearly 100 stops across subway, tram, bus rapid transit (BRT), and cable cars (tramway). Daily ridership averages 1,000,000+. The current network is given in “**medellin_metro.csv**” each row represents a station, complete with its latitude and longitude coordinates, the comuna it is in, and what line it is on (as well as the type of line). Transfer stations are indicated. Successive stops along the same line are successive in the file. For instance, Sabaneto and La Estrella are adjacent on Line 1, but La Estrella (Line A) and San Antonio (Line B) are not.
- The file “**comunas_commutes.csv**” contains the relative number of trips between each comuna and has been constructed from mobile phone records (tower pings). For each row, there are three columns - the first gives the origin (an index), the second gives the destination (another index), and the final column gives the relative number of trips. The file “comunas.csv” matches the index with the corresponding comunas.
- “**comunas_gmaps.csv**” contains the driving times, transit times and walking times (in seconds) between the different comunas which come from Google Maps. Edge (i,j) is how long it takes to commute from comuna i to j (at 9am in the morning during the week).
- “**comunas_pop.csv**” gives the average population in 2020 within each comuna.
- “**comunas_firms.csv**” gives the predicted number of commercial firms that are visible in Google Street Imagery (both formal and informal) in each comuna.
- “**comunas_stratum.csv**” contains the average stratum of the neighbourhoods in each comuna. Stratum is a socio-economic indicator that is based on the quality of housing and infrastructure. The highest socio-economic scale is 6 (wealthy areas), and the lowest 1 (poor areas). See: <https://www.theguardian.com/cities/2017/nov/09/bogota-colombia-social-stratification-system>
- We have also included the file “**comunas_boundaries.zip**” which contains the shapefiles of the various comunas. This file labels comunas in Macrozona numbers, and it can be linked to the comuna indexes of the rest of the data in the “comuna.csv” file.