

# QGIS

Parte del material obtenido de [geodacenter.github.io/workbook/1\\_datascience/lab1.html](https://geodacenter.github.io/workbook/1_datascience/lab1.html)

# GIS

- Un GIS es un conjunto de programas integrado, capaz de almacenar, editar, analizar, compartir y mostrar información geográfica. QGIS es un GIS libre y gratuito.
- Permite leer capas vectoriales y rasters.

# ¿Qué vamos a hacer?

- Descargar datos de portal de datos abierto. En este caso del portal Sócrata, datos abiertos de la Ciudad de Chicago
- Crear una tabla de datos a partir de un archivo con formato csv
- Manipular formatos de las variables, incluyendo formatos de fecha/hora
- Seleccionar las observaciones usando la herramienta de selección
- Crear una nueva capa de datos a partir de observaciones seleccionadas
- Convertir una tabla con coordenadas de puntos en una capa de puntos
- Selección espacial
- Agregación espacial
- Uniendo las tablas
- Cálculos de variables en una tabla
- Mapeo básico de coropletas

# Mapas coropléticos

- Mapas de coropletas o temáticos
- Visualización de una distribución espacial:
  - contrapartida de un histograma.
  - aproximación discreta de una distribución.
  - todas las observaciones con igual intervalo de valor se visualizan con el mismo color.

Mapa

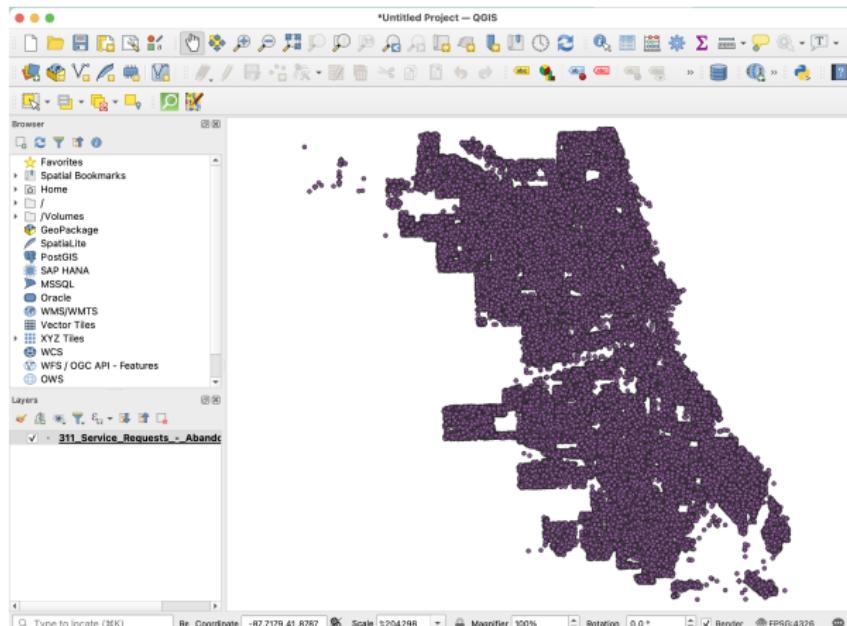
# Descargar datos de Socrata

- <https://data.cityofchicago.org/>
- Service Requests
- <https://data.cityofchicago.org/Service-Requests/311-Service-Requests-Abandoned-Vehicles/3c9v-pnva>
- Download button → CSV

The screenshot shows the Chicago Data Portal interface. At the top, there's a navigation bar with links for 'Browse', 'Tutorial', 'Feedback', social media icons (Facebook, Twitter, YouTube), a search icon, and a 'Sign In' button. Below the navigation is the dataset title '311 Service Requests - Abandoned Vehicles'. A 'Service Requests' button is visible. The main content area contains a brief description of the dataset: 'All open abandoned vehicle complaints made to 311 and all requests completed since January 1, 2011. A vehicle can be classified as abandoned if it meets one or more of the following criteria: All open abandoned vehicle complaints made to 311 and all requests completed since January 1, 2011.' A 'More' link is present. To the right, there's a sidebar with download options ('View Data', 'Visualize', 'Export', 'API', '...', 'Download 311 Service Requests - Abandoned Vehicles', 'CSV for Excel'), additional formats ('CSV for Excel (Europe)', 'TSV for Excel', 'RDF', 'XML', 'RSS'), and a link to 'Additional Formats'. At the bottom left, there's a section titled 'About this Dataset' with tables for 'Updated' (June 15, 2020) and 'Metadata' (Data Owner: Streets & Sanit, Time Period: All open requests and responses completed since January 1, 2011, Frequency: Data is updated daily, Last Updated Date via Automated Load).

# Agregar datos

- Layer → Add Layer → Add Delimited Text Layer:  
311\_Service\_Requests\_-\_Abandoned\_Vehicles.csv
- Point Coordinates → X: Longitude, Y: Latitude
- El color varía



## Manipular formatos de las variables - queremos sólo Sept. 2016

- Botón derecho sobre layer → Export → Save Feature As... →  
Hacer click en [...] → Chicago\_full.sph (ESRI Shapefile)
- Quedarse con: Creation Date, Address, ZipCode, X, Y, Ward, PoliceD, Comm, Latitude and Longitude, Location
- Remove Layer (botón derecho) 311\_Service...
- Ver tipo de variables
  - Botón derecho → Properties
  - Fields → Querríamos que "Creation Date" sea *Date*, no *String*

# Manipular formatos de las variables - queremos sólo Sept. 2016

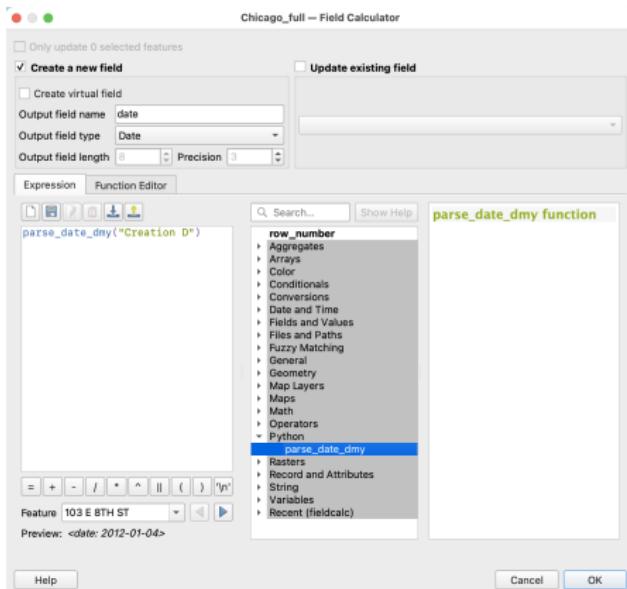
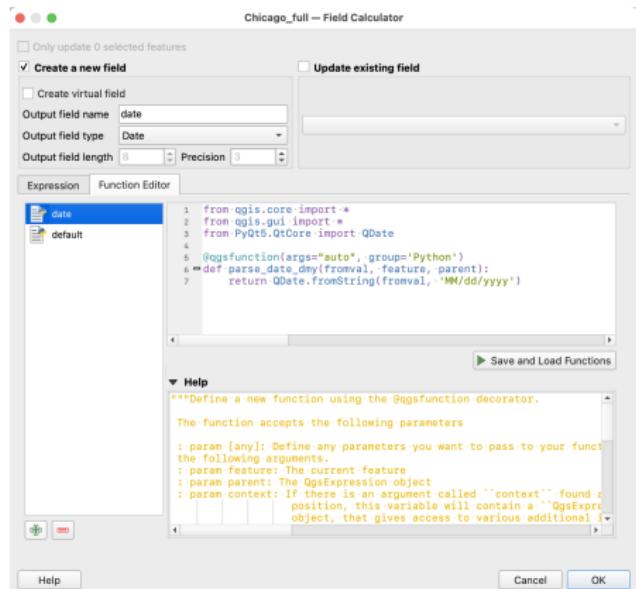
- Queremos cambiar variable a formato *Date*
  - Botón derecho → Toggle Editing
  - Open Field Calculator 
  - Function Editor → Click en el signo + en verde → Poner nombre que quieran.
  - Agregar este texto:

```
from qgis.core import *
from qgis.gui import *
from PyQt5.QtCore import QDate

@qgsfunction(args="auto", group='Python')
def parse_date_dmy(fromval, feature, parent):
    return QDate.fromString(fromval, 'MM/dd/yyyy')
```

- Ir a Expression tab → Python → parse\_date\_dmy
- Output Field Name: Date
- Output Field Type: Date
- parse\_date\_dmy( "Creation D" )

# Frame Title



## Filter

- Open Attribute Table (botón derecho o ícono)
  - Ícono filtro 
  - date → Between (inclusive) → 2016-09-01 to 2016-09-30 →  
Select features (abajo a la derecha) → Selected: 2637
  - Ícono "Inverse Selection" () → 258849
  - Ícono "Delete Selected Features" ()

# Borrar variables (field)

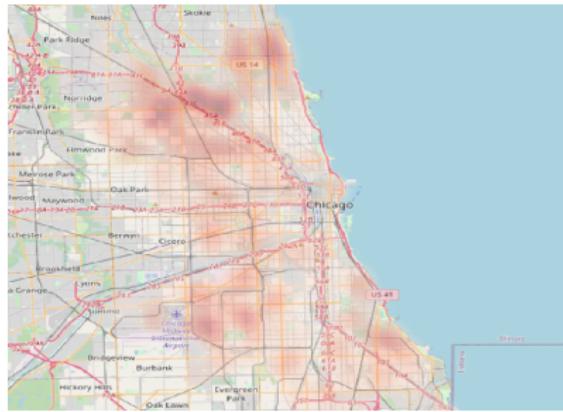
- Ícono “Delete Field” (
- Location
- Export → Save Feature As... → Hacer click en [...] →  
**Chicago\_sept16.sph** (ESRI Shapefile)
- Remove Layer **Chicago\_full**

# Amendment

- Buscar en tabla los que tienen Comm=0 (haciendo click en Comm, ordena la tabla por esa variable)
- Seleccionar una fila
- Click en el ícono “Pan map to the selected rows” ( ) y buscar en el mapa el punto que se selecciona
- Ir a mapa y usar la herramienta de Select ( ) para seleccionar un área chica de puntos alrededor de los seleccionados ( $\approx 10$ )
- Move Selected to Top
- Ver número de comunidad de las de alrededor
- Corregir tabla (si no está activado, “Toggle Editing” )
- Botón Derecho → Save Layer Edits

# Point Map to Heat map

- Plugins → Manage and install plugins → QuickOSM y QuickMapServices
  - Web → Quick Map Services → OSM Standard
  - Botón Derecho sobre Layer → Properties → Symbology
  - Heat Map ([Help](#))
  - Color Ramp: Reds
  - Layer Randering → Opacity: 35 %
  - Ojo! Esta es una variable *spatially extensive*



# Cuidado!

- La cantidad de autos abandonados es una variable **spatially extensive**: mayor para áreas más grandes, sólo porque el área es más grande → evitar graficar estas variables!
- Transformarlas en variables **spatially intensive**

# Objetivo: Análisis por Community Area → *spatial aggregation*

- Sin el identificador, en GIS necesitaríamos hacer una operación *point in polygon*
- Pero no tenemos los polígonos!
- Queremos contar la cantidad de autos abandonados por Community Area
  - Field Calculator
  - Enumerar filas → \$rownumber → Output field name: *id*
  - Contar distintos por Community →  
`count_distinct("id", "Community")` → Output field name: *count*
  - Nos falta la variable de cantidad de habitantes.

Objetivo: *Choropleth Map* con la cantidad de vehículos abandonados cada 100,000 hab.c

- Idealmente conseguir *boundary file* y Población en mismo shape file
  - [Chicago Community Area Population 2000 and 2010](#)
  - Add Layer → Add Vector Layer (**chicago\_commpop.shp**)
  - Opción 1:
    - Properties → Joints (barra a la izquierda)
    - Join layer: Chicago\_Sept2016, Join field: Community, Target field: NID
  - Opción 2:
    - Vector → Data Management Tools → Join Attributes by Location
    - Ver Attributes Table de la Joined Layer

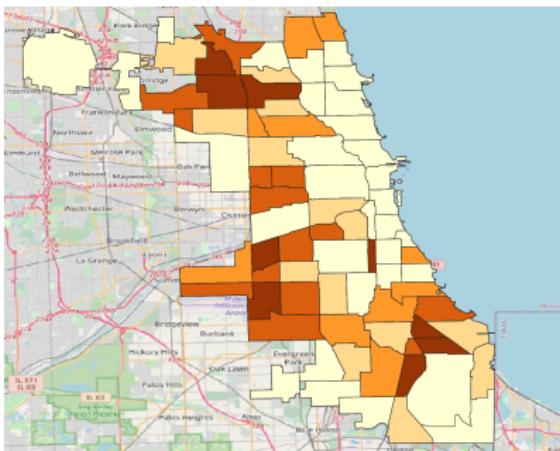
# Objetivo: *Choropleth Map* con la cantidad de vehículos abandonados cada 100,000 hab.

- Calcular vehículos abandonados cada 100,000 habitantes
  - Field Calculator
  - Output field name: av100000
  - count/POP2010\*100000
- Quedarnos con un feature por community
  - Export (así guardamos la versión anterior)
  - Enumerar por comunidad
    - Opción 1: Processing → ToolBox (✿)
    - Add autoincremental tool → Start values at 1, Group Values by "Community"
    - Opción 2:

```
array\_find(array\_agg($id, "community"), $id)+1
```
- Eliminar los que no tienen 1 en la nueva variable (tienen que quedar 77 observaciones)

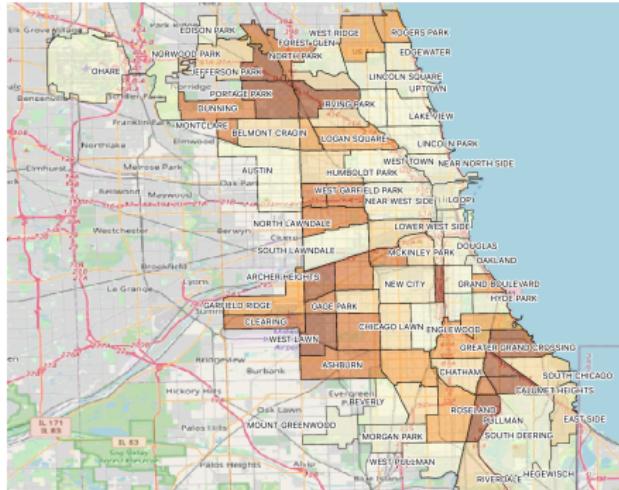
## Objetivo: *Choropleth Map*

- Properties → Symbology → Graduated → Value: av100000
  - Color ramp: YOrBr
  - Equal count (Quantile) - 5



## Objetivo: *Choropleth Map*

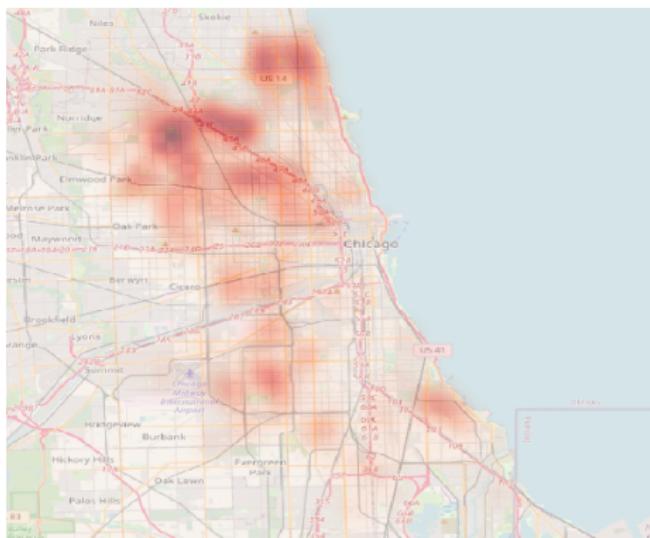
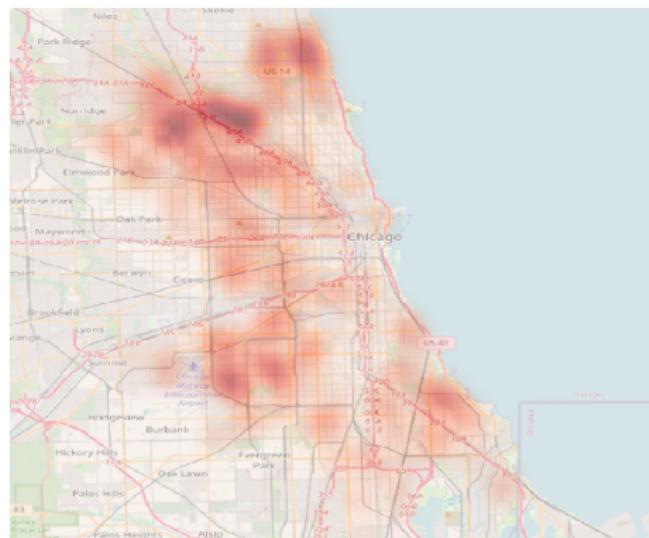
- Más transparencia para ver mapa de fondo
    - Hace click en cada color → Edit Symbol → Simple Fill → Fill Style → Dense 4
  - Agregar leyendas
    - Properties → Labels → Single Labels → community
    - Buffer 



# Heat Map con weight

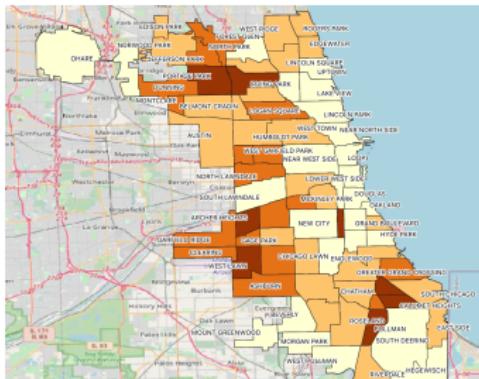
- Add Delimited Text Layer **commpopulation.csv**
- Layer → Layer Properties (o botón derecho del mouse sobre layer → Joints (sobre barra a la izquierda)
- Joint Layer: **commpopulation.csv**, Join Field: NID, Target Field: Community
- Symbology: Heatmap, Weight points by:  
*commpopulation\_POP2010*

# Heat Map con weight



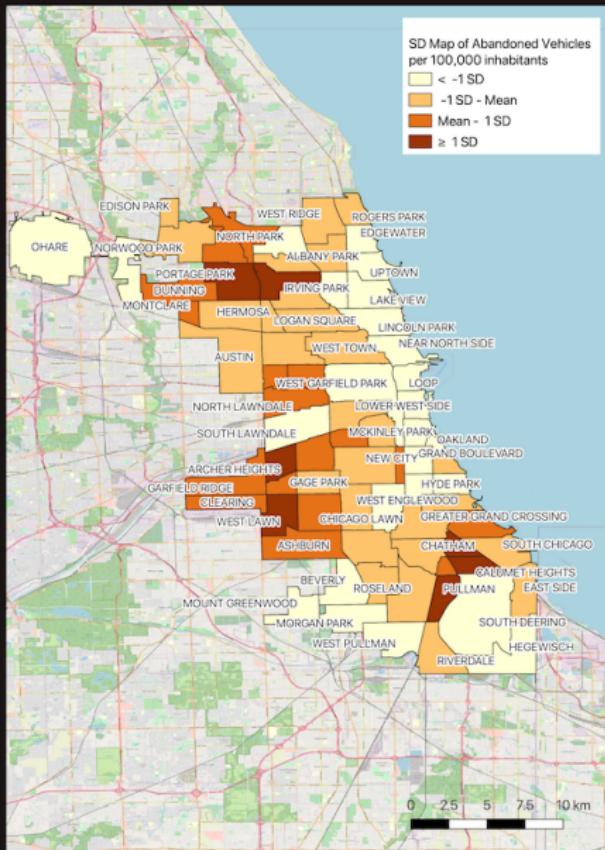
# Mapa de desviación estándar

- Los intervalos corresponden a una desviación extra.
- Los outliers son identificados como los que superan 1 desviaciones respecto a la media.
- Las categorías son 4 (o de a un o medio SD) predefinidas:
  - ( $<\text{media}-1\text{SD}$ )
  - ( $\text{media}-1\text{SD}, \text{media}$ )
  - ( $\text{media}, \text{media}+1\text{SD}$ )
  - ( $>\text{media}+1\text{SD}$ )



# Exportar mapa

- Project → New Print Layout 
- Botón Derecho → Page properties → Item Properties → Orientation: Portrait
- Add Item → Add Map ()
- Scroll para fit o a mano (Scale)
- Deshabilitar Select/Move Item ()
- Botón Derecho → Page properties → Background: Black
- Add scalebar ()
- Add legend ()
- Export as Image or PDF

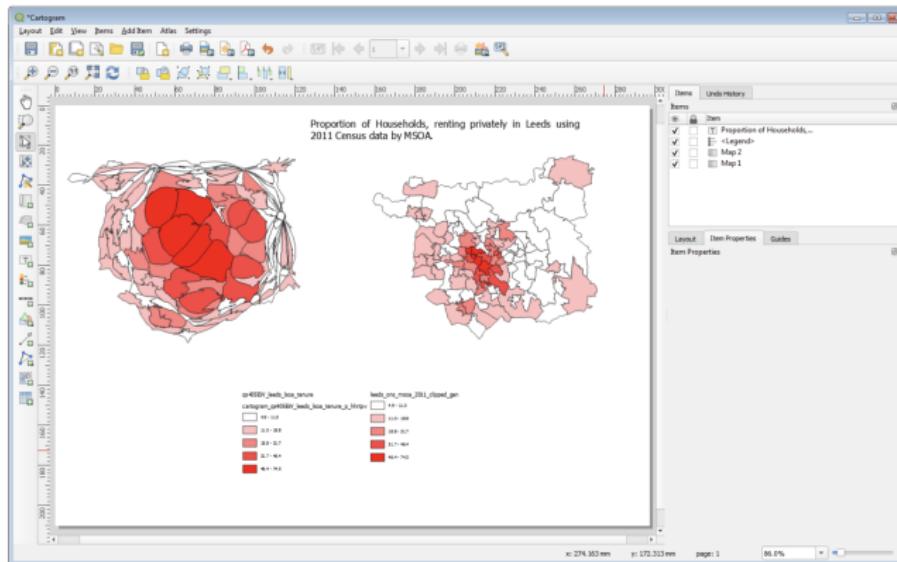


# Otros mapas

- Natural Breaks
  - Agrupa minimizando la varianza dentro de cada clase.
  - Diferente número de polígonos en la clase.
- Pretty Intervals
  - Sólo redondea
- Equal Intervals
  - Las clases tienen la misma amplitud de valores.
  - Diferente número de polígonos en la clase.
- Cartograma
  - Regiones en tamaño proporcional de la variable de interés.
  - Evita efectos engañosos por tamaño de área.
  - Utiliza formas de áreas transformadas respetando la topología.
  - Variable mapeada según tamaño y color.

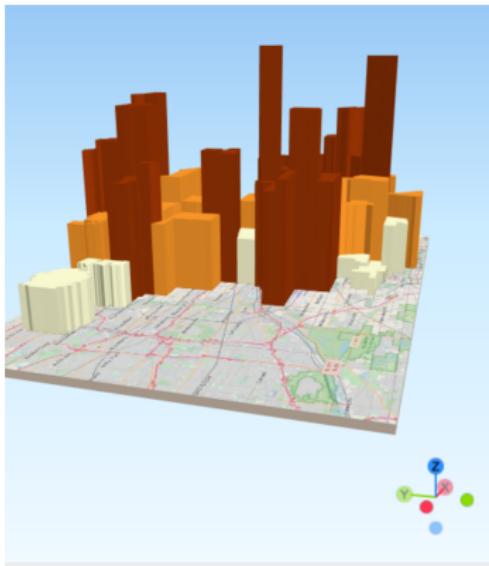
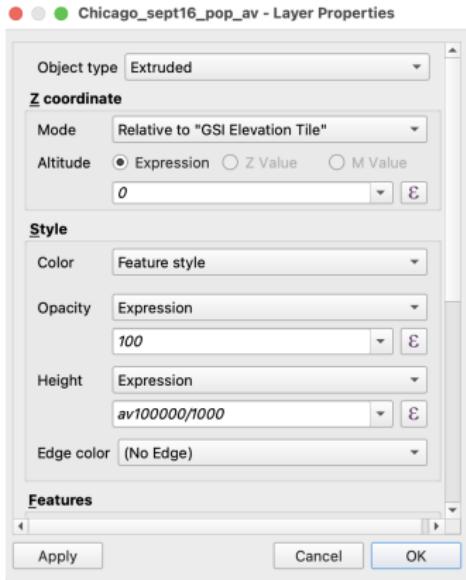
# Cartograma

- Plugins → Manage and Install Plugins → cartogram3
  - Vector → Cartogram
  - Example



# Mapa en 3D

- Plugins → Manage and Install Plugins → qgis2threejs
- Hacer mapa de cuantiles (3)
- Web → qgis2threejs Exporter
- Botón derecho del mouse sobre layer → Properties



# De mapa a gráfico

- Si queremos volver al Point Vector y eliminemos lo que no es septiembre 2016 **con Python**
  - Add Layer → Delimited Text Layer
  - Export
  - Year: Field Calculator → right( "Creation D" , 4)
  - Month: Field Calculator → substr( "Creation D" , 1 , 2)

```
layers = QgsProject.instance().mapLayersByName('test')
layer = layers[0]
caps = layer.dataProvider().capabilities()
feats = layer.getFeatures()
dfeats = []

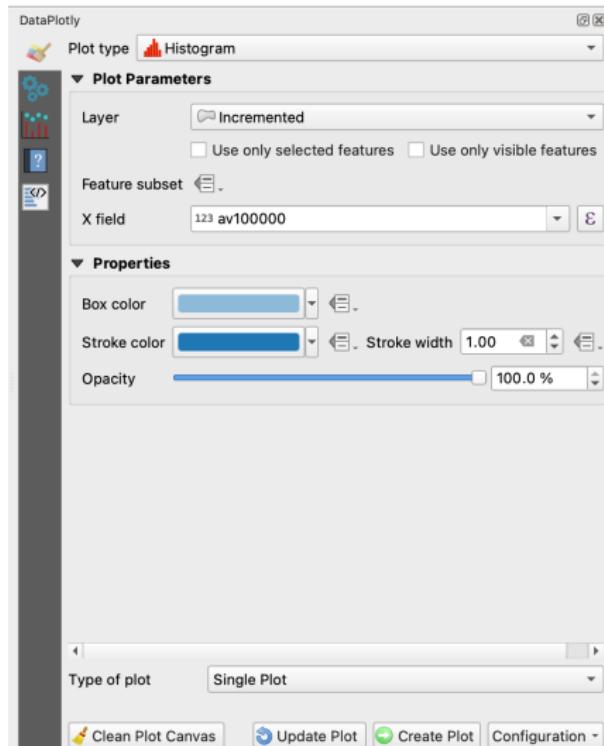
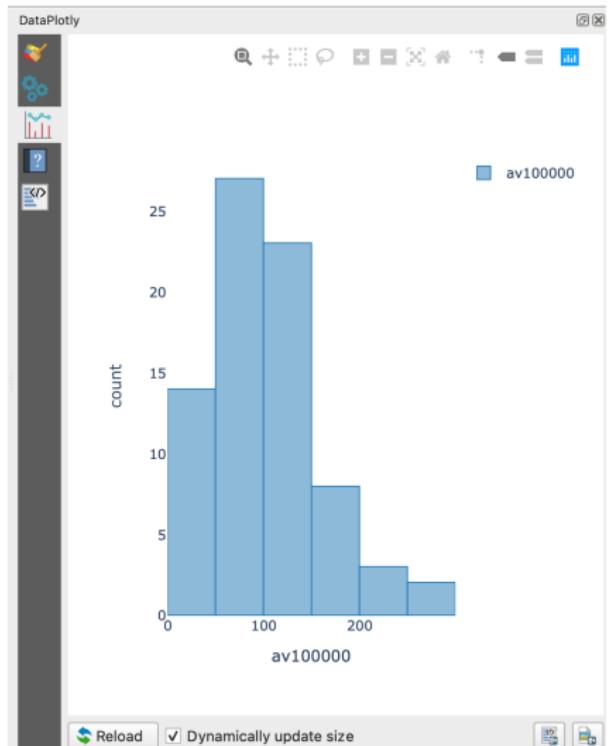
if caps & QgsVectorDataProvider.DeleteFeatures:
    for feat in feats:
        if feat['Year']!=2016:
            dfeats.append(feat.id())
    res= layer.dataProvider().deleteFeatures(dfeats)
    layer.triggerRepaint()
```

- Repetir para Month y eliminar lo que no cumpla
- [Source](#)

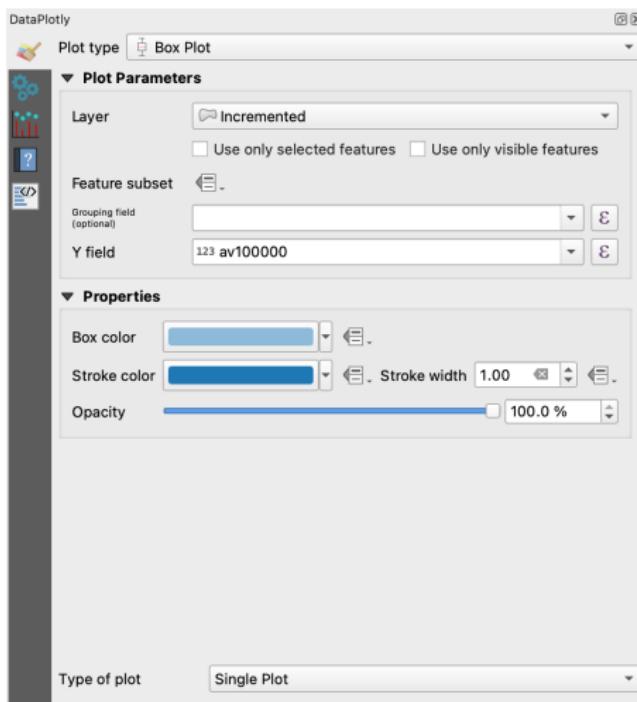
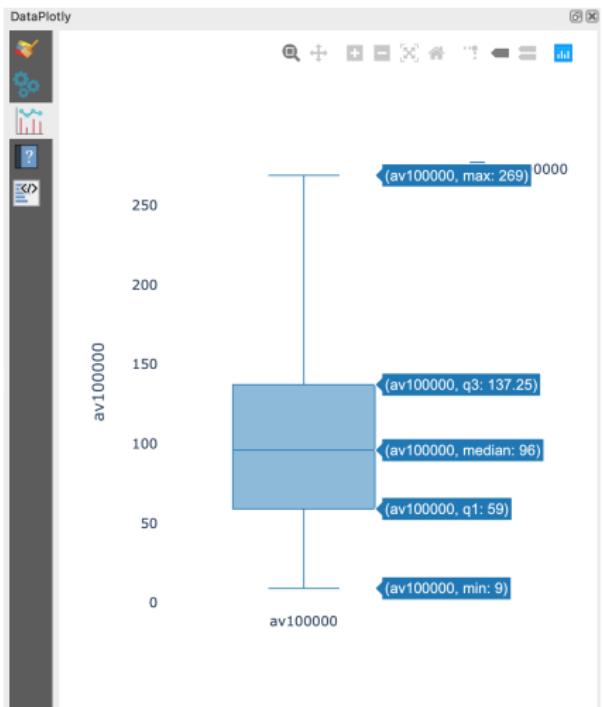
## De mapa a gráfico

- Para hacer los gráficos necesitamos la base con 77 observaciones, es decir, a nivel de community.
  - Plugins → Manage and Install Plugins → [PlotPy](#)

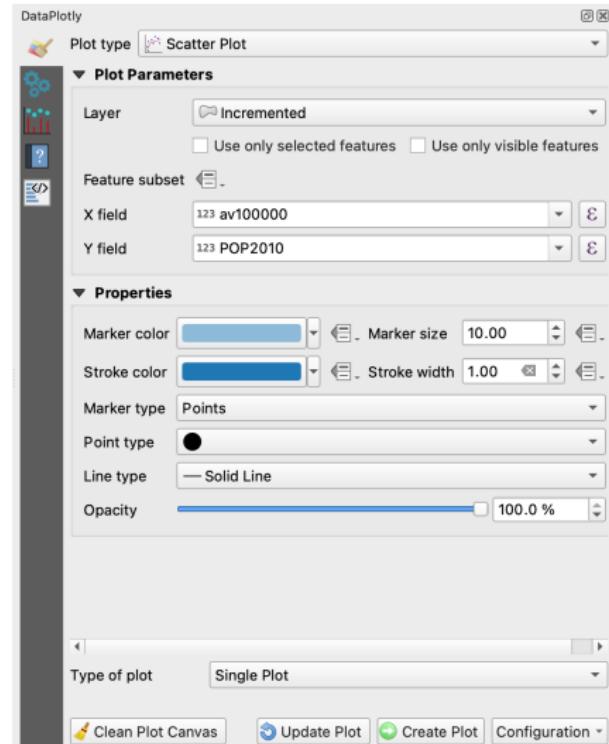
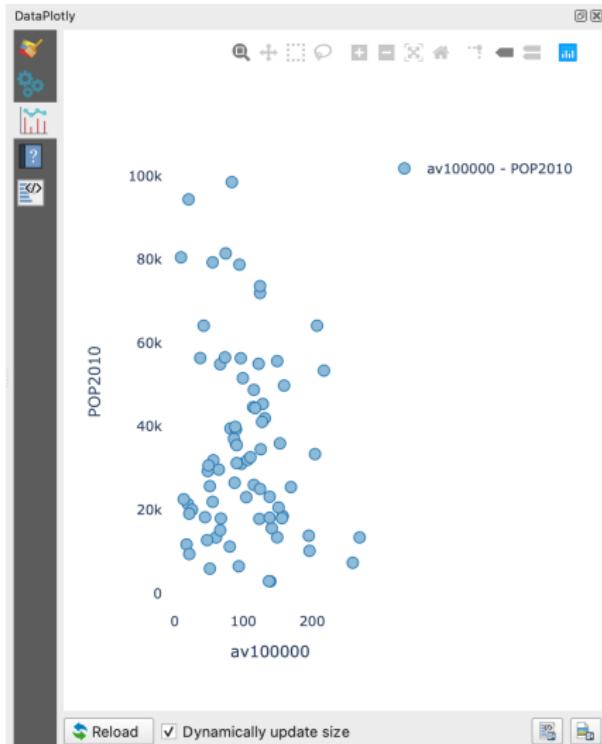
# Histograma



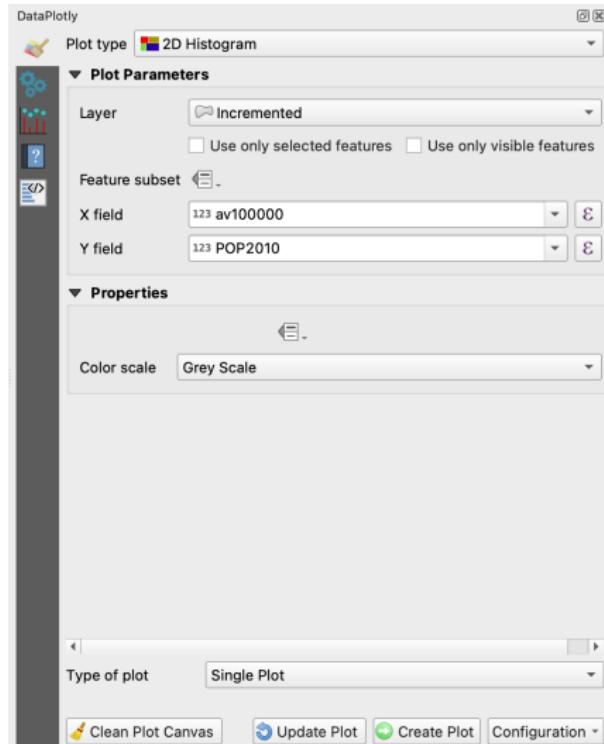
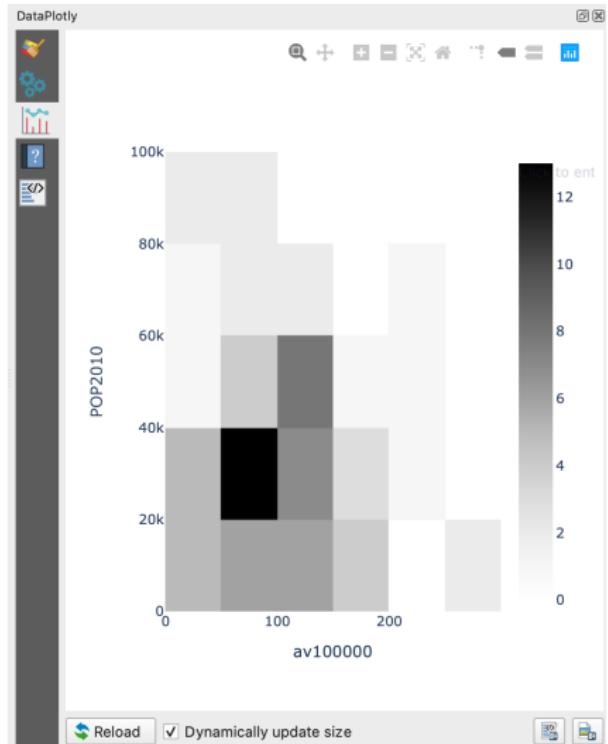
# Box Plot



# Scatter



## Histograma 2D

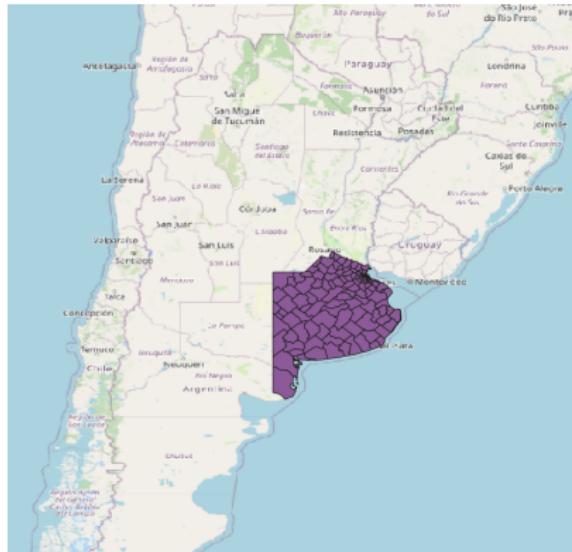


# Buenos Aires - INDEC

- Ir a  
[www.indec.gob.ar/indec/web/Institucional-Indec-Codgeo](http://www.indec.gob.ar/indec/web/Institucional-Indec-Codgeo)
- Seleccionar: *Descargue el archivo: Total país por departamento.*  
*(archivo rectificado 8/02/2017)*
- Descargar **Codgeo\_Pais\_x\_dpto\_con\_datos.zip**
- Layer → Add Layer → Add Vector Layer → Seleccionar el archivo **pxdptodos.shp**
- Me quiero quedar con Buenos Aires y CABA:
  - Layer → Toggle Editing (ó botón derecho del mouse sobre la layer)
  - Layer → Open Attribute Table (ó botón derecho del mouse sobre la layer)
  - Filtro → Starts with → link 0 → Select Features → Invert Selection → Delete
  - Layer → Save as... → Hacer click en los [...] de File Name → Guardar como Shapefile **partidos\_bsas.shp**

# Agregar mapa detrás

- Plugins → Manage and install plugins → QuickOSM y QuickMapServices
- Web → Quick Map Services → OSM Standard



# Buenos Aires - INDEC

- Ir al INDEC [https://www.indec.gob.ar/indec/web/  
Institucional-Indec-BasesDeDatos-6](https://www.indec.gob.ar/indec/web/Institucional-Indec-BasesDeDatos-6)
- Censo 2010: Base de datos REDATAM del cuestionario ampliado
- Resultados Básicos → Población → Condición de actividad
- Resultados Básicos → Hogares → Al menos un indicador NBI
- Corte de Área: Departamento
- Ejecutar → Abajo de todo aparece "Descargar en formato Excel"
- Botón derecho del mouse sobre link → Save link As...

# Buenos Aires - INDEC

**Censo 2010**  
Instituto Nacional de Estadística y Censos de la República Argentina

**indec**

**Cuestionario ampliado**

**Base de dato REDATAM**

**Censo 2010**  
Instituto Nacional de Estadística y Censos de la República Argentina

**indec**

**Cuestionario ampliado**

**Base de dato REDATAM**

**Población**

**Definir Parámetros**

Título de la Tabla

Completo el nivel  
Utiliza computadora  
Condición de actividad  
Cobertura de salud  
Dificultad o limitación permanente para oír  
Dificultad o limitación permanente para caminar o subir  
Dificultad o limitación permanente para agarrar objetos  
Dificultad o limitación permanente para entender y/o aprender  
Recibe jubilación o pensión

Selección Variable(s)

Corte de Área

Área Geográfica    
Definición del Universo

Tipo de salida

**Ejecutar**

WebMaster [agenciamaster@indec.mecion.gob.ar](mailto:agenciamaster@indec.mecion.gob.ar)  
 © 2002-2014 Consultas de información estadística: [cens@indec.mecion.gob.ar](mailto:cens@indec.mecion.gob.ar)  
 Comentarios sobre nuestro sitio: [indecweb@indec.mecion.gob.ar](mailto:indecweb@indec.mecion.gob.ar)  
 INDEC - Av. Julio A. Roca 609, PB, C1057ABB Ciudad Autónoma de Buenos Aires, Argentina. Teléfono: 54 11 4349-9235  
 Powered by **REDATAM**

**Hogares**

**Definir Parámetros**

Título de la Tabla

Total de habitaciones o piezas  
Habitaciones  
Computadora  
Teléfono celular  
Teléfono de línea  
Régimen de tenencia  
Haciembamiento  
Al menos un indicador NBI  
Total de Personas en el Hogar  
Indicadores de población de los hogares

Corte de Área

Área Geográfica    
Definición del Universo

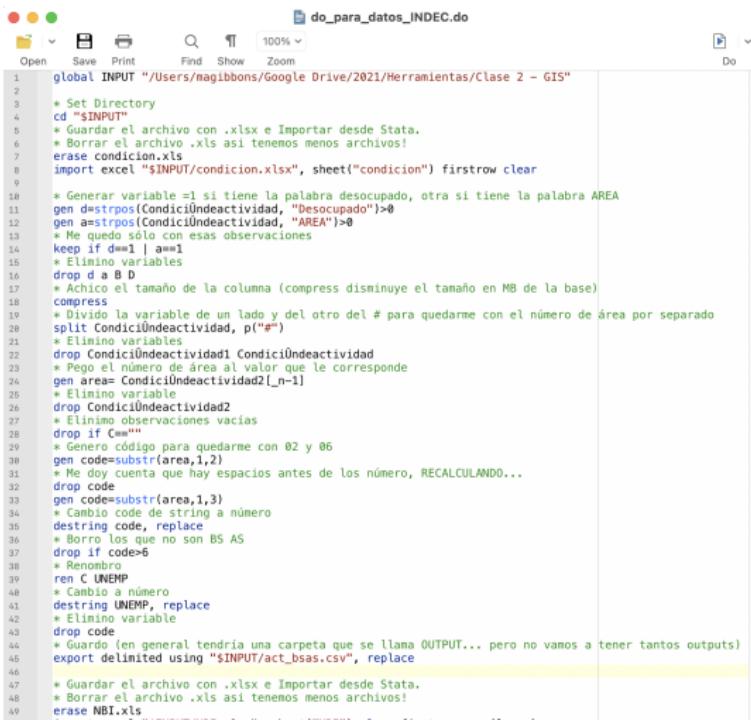
Tipo de salida

**Ejecutar**

WebMaster [agenciamaster@indec.mecion.gob.ar](mailto:agenciamaster@indec.mecion.gob.ar)  
 © 2002-2014 Consultas de información estadística: [cens@indec.mecion.gob.ar](mailto:cens@indec.mecion.gob.ar)  
 Comentarios sobre nuestro sitio: [indecweb@indec.mecion.gob.ar](mailto:indecweb@indec.mecion.gob.ar)  
 INDEC - Av. Julio A. Roca 609, PB, C1057ABB Ciudad Autónoma de Buenos Aires, Argentina. Teléfono: 54 11 4349-9235  
 Powered by **REDATAM**

# Buenos Aires - INDEC

- Ir a Editor de Datos para armar la base
- Ejemplo Stata



The screenshot shows a Stata do-file editor window titled "do\_para\_datos\_INDEC.do". The code in the editor is as follows:

```
1 global INPUT "/Users/magibbons/Google Drive/2021/Herramientas/Clase 2 - GIS"
2
3 * Set Directory
4 cd "$INPUT"
5 * Guardar el archivo con .xlsx e Importar desde Stata.
6 * Borrar el archivo .xls asi tenemos menos archivos!
7 erase condicion.xls
8 import excel "$INPUT/condicion.xlsx", sheet("condicion") firstrow clear
9
10 * Generar variable =1 si tiene la palabra desocupado, otra si tiene la palabra AREA
11 gen destrpos(CondiciÙndeactividad, "Desocupado")>0
12 gen aestrpos(CondiciÙndeactividad, "AREA")>0
13 * Me quedo sólo con esas observaciones
14 keep if dwi==1 | awi
15 * Elimino variables
16 drop d a B D
17 * Achico el tamaño de la columna (compress disminuye el tamaño en MB de la base)
18 compress
19 * Divide la variable de un lado y del otro del # para quedarme con el número de área por separada
20 split CondiciÙndeactividad, p("#")
21 * Elimino variables
22 drop CondiciÙndeactividad1 CondiciÙndeactividad
23 * Pego el número de área al valor que le corresponde
24 gen area= CondiciÙndeactividad2[_n-1]
25 * Elimino variable
26 drop CondiciÙndeactividad2
27 * Elimino observaciones vacías
28 drop if C==""
29 * Genero código para quedarme con 02 y 06
30 gen code=substr(area,1,2)
31 * Me doy cuenta que hay espacios antes de los número, RECALCULANDO...
32 drop code
33 gen code=substr(area,1,3)
34 * Cambio code de string a número
35 destring code, replace
36 * Borro los que no son BS AS
37 drop if code<>8
38 * Renombro
39 ren C LINEMP
40 * Cambio a número
41 destring LINEMP, replace
42 * Elimino variable
43 drop code
44 * Guardo (en general tendría una carpeta que se llama OUTPUT... pero no vamos a tener tantos outputs)
45 export delimited using "$INPUT/act_bsas.csv", replace
46
47 * Guardar el archivo con .xlsx e Importar desde Stata.
48 * Borrar el archivo .xls asi tenemos menos archivos!
49 erase NBI.xls
```

# Merge datos (sólo datos, nada geográfico)

- Layer → Add Layer → Add Delimited Text Layer → nbi\_bsas.csv
- Layer → Add Layer → Add Delimited Text Layer → act\_bsas.csv
- Layer → Layer Properties (o botón derecho del mouse sobre **partidos\_bsas.shp**) → Joints (sobre barra a la izquierda)
- Join layer: act\_bsas, Join field: area, Target field: link
- Desclickear "Cache join layer in memory"
- Chequear que quedó bien!
- Botón derecho sobre **partidos\_bsas.shp** → Export → Save Features As... → Hacer click en los [...] → **bsas\_nbi\_unemp.shp**