

Clase 1 - Repaso de elementos de análisis territorial y ArcMap

Martín Montané

Primer semestre 2019

¡Bienvenidos a instrumentos de análisis territorial 2!

Antes que nada quería darles la bienvenida al segundo curso de análisis espacial. En otro documento van a ver un poco más técnicamente qué es lo que vamos a hacer. Sin embargo, en términos prácticos el objetivo de esta segunda parte es realmente simple: **darles las herramientas para poder realizar un análisis de datos espaciales desde el principio hasta el final**. Subrayo esta última parte porque *el principio* incluye la búsqueda y transformación de los datos desde su hábitat natural hasta un formato en el cuál puedan analizarlo espacialmente. También incluye no solo el análisis propiamente dicho, sino también la visualización y la comunicación de los resultados. Con este objetivo en mente es que vamos a tener una breve **Introducción a R**, además de usar ArcMap y GeoDa, programas en los cuales ya tienen experiencia.

El objetivo de la primera clase - que probablemente se extienda dos clases - es ni más ni menos que refrescar el uso de ArcMap y algunos elementos relacionados con los datos espaciales. Para esto, vamos a replicar uno de los ejercicios que hicieron el año pasado, pero con algunas modificaciones para que, junto con el repaso, empecemos a buscar los datos allí donde están y conozcamos algunas funciones que todavía no conocen.

El problema

Se han presentado quejas sobre la falta de Centros de Salud (conocidos como CeSac) del Gobierno de la Ciudad de Buenos Aires en algunas zonas del territorio. A pesar de conocer la residencia de los ciudadanos que presentaron las quejas, el GCBA quiere diagnosticar correctamente si es necesario construir uno y, de ser así, conocer en qué manzana de la ciudad debería hacerse. Eso es exactamente lo que vamos a intentar responder.

La metodología

Lo que vamos a hacer es generar un Índice de Demanda de Salud (IDS) para cada manzana de nuestra área de análisis. Este indicador será un promedio ponderado de cuatro indicadores:

1) *La facilidad de acceso (medido como la distancia a la avenida más cercana)* 2) *La falta de oferta hospitalaria (medido como qué tan lejos está de los actuales servicios de internación hospitalaria)* 3) *La cantidad de personas con Necesidades Básicas Insatisfechas (NBI) a nivel de radio censal* 4) *La densidad poblacional a nivel de radio censal*.

Los ponderadores van a poder ser elegidos por ustedes, pero primero debemos generar estas métricas. Luego, para hacerlo más realista, incorporaremos una restricción relevante: solo recomendaremos construir un CESAC si no se encuentra actualmente cubierto por otro ya existente!

Consiguiendo los datos

Para este curso van a poder descargar siempre los datos desde links que yo les voy a ir pasando en estos documentos. Sin embargo, siempre que sea posible voy a insistirles para que vayamos hasta las fuentes a descargarlos. En este caso vamos a descargar desde internet algunos de los datos (que en otras clases vamos a aprender cómo generar) y otros vamos a ir a buscarlos “donde viven”.

Desde el portal de datos de la Ciudad de Buenos Aires vamos a descargar algunos datasets que nos serán de mucha importancia¹:

- Límites de los barrios de la ciudad
- Calles
- Centros de salud

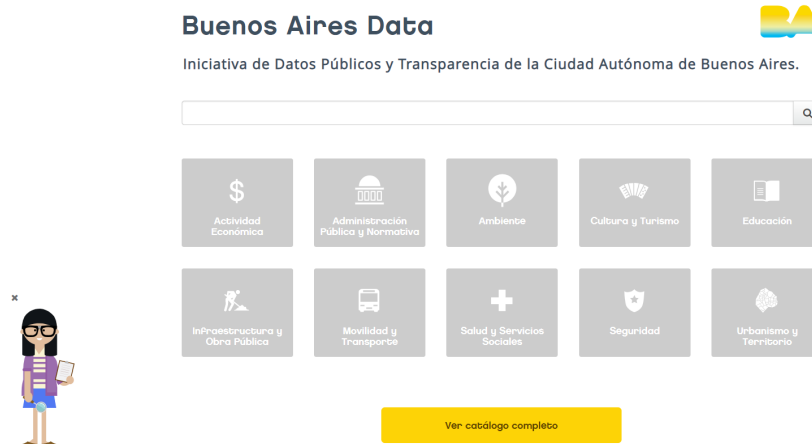


Figura 1: Buenos Aires Data. Accesible via <https://data.buenosaires.gob.ar/>

Para eso, simplemente escribiremos “barrios” en la barra de búsqueda (Ver figura 1) y descargaremos el ESRI Shapefile - mejor conocido como “Shapefile”. Lo mismo haremos con “Centros de salud” y “calles”. Como ya saben, los Shapefiles, al estar compuestos de más de un archivo, suelen comprimirse. La página del Gobierno permite descargarlo desde un .rar

Por el momento, los datos censales provenientes del 2010 van a descargarlos desde el siguiente link: <https://github.com/datalab-UTDT/GIS2/blob/master/Clase%20Repaso/cabaCENSO2010.rar>. Esto será temporal, y en clases posteriores vamos a aprender cómo conseguir los datos y generar algunas de sus variables. Adicionalmente van a descargar los datos de sobre las manzanas de la Ciudad (<https://github.com/datalab-UTDT/GIS2/raw/master/Clase%20Repaso/Manzanas.zip>) e información sobre la capacidad de internación de los hospitales de la ciudad (<https://github.com/datalab-UTDT/GIS2/raw/master/Clase%20Repaso/HospitalesInternacion.rar>).


Cargando los datos

Para comenzar con nuestro proyecto en ArcMap lo que debemos hacer es descomprimir todos los archivos en una carpeta, deberían tener seis Shapefiles, contando el que descargaron desde el repositorio del curso con los datos del censo 2010. Solo para ser muy claro, estos seis shapefiles deberían ser:

- Barrios de la ciudad
- Calles
- Centros de salud
- Hospitales de internación
- Datos a nivel de radio censal
- Manzanas de la Ciudad de Buenos Aires.

Los primeros datos espaciales que vamos a agregar son los datos del Censo 2010. Como ya saben, obedecen a los polígonos del censo del 2010 y contienen diversa información habitacional y sociodemográfica sobre los

¹El repositorio de datos abiertos de la Ciudad de Buenos Aires es un gran espacio para conseguir nuevos datos. Les recomiendo que siempre que tengan que realizar algún análisis sobre la ciudad lo tengan en cuenta.

habitantes de la Ciudad de Buenos Aires. Para agregar a estos datos, solo tienen que ir al botón de agregar datos , ir hasta la carpeta donde contamos con todos los shapefiles y hacer click en el correspondiente a los datos del censo de CABA.

¿En qué sistema de coordenadas están los datos del Censo? ¿Se encuentran proyectados? Como ustedes saben, un conjunto de **layers** pueden guardarse en **Data Frames**, que a su vez tienen un sistema de coordenadas determinado. Los datos sobre este sistema pueden ser consultados desde **Properties**. **¿Cuál es el sistema? ¿Es el mismo que el de los polígonos del Censo?**

Luego, vamos a agregar los polígonos de los Barrios, que los vamos a utilizar más que nada para poder vincular a los radios censales con los barrios y comunas, información que no está en la *Attribute Table* de los datos del Censo ¿Recuerdan cómo hacerlo?. Simplemente hagan click derecho desde el Table Of Contents (TOC) en el shapefile del Censo y abran la tabla de atributos. Van a ver que no hay datos para esto. Sin embargo, al tratarse de datos espaciales podemos establecer relaciones entre los dos datasets con los que contamos en base a su ubicación espacial usando el *Spatial Join*. Pero antes debemos trabajar con una de las propiedades más importantes de los datos espaciales: su sistema de coordenadas.

Corrección de proyecciones

Cuando agreguen los datos de los Barrios van a recibir una alerta sumamente informativa de ArcMap, donde dice que los datos espaciales que estamos por agregar no comparten el mismo sistema de coordenadas que el Data Frame al cual lo estamos asignando. Esto es importante, pero por el momento vamos a apretar “Close” y ver qué sucede si no hacemos ninguna transformación entre los sistemas de coordenadas.

Ahora coloquen los datos de barrios por encima de los de radios censales, agreguen algo de transparencia (**Properties -> Display -> Transparent**), y desde **Symbolology** aumenten un poco el tamaño de los bordes y elijan un color que haga que se vea el contraste. ¿Qué sucede en los límites de los barrios? Si bien tienen la misma forma que algunos de los límites de los radios censales, estos se encuentran desplazados. Este es solo uno de los efectos que pueden suceder cuando las proyecciones no son compatibles. Para solucionar este problema debemos “traducir” uno u otro sistema de coordenadas hacia el otro. Por suerte, ArcMap hace este trabajo un poco más fácil para las proyecciones que ya conoce, aunque no siempre tenemos tanta suerte.

Podríamos trabajar con cualquiera de las dos proyecciones, pero debido a que la mayoría de los datos espaciales que vamos a seguir usando se encuentran en la misma proyección que el dataset de Barrios, vamos a convertir a los polígonos de la layer del censo al de barrios. Desde el **ArcToolbox** vamos a **Data Management Tools** y luego a **Projections and Transformations** y finalmente a **Project**. Deben elegir como Input dataset al que desean transformar y como output la proyección a la que quieren convertirlo (Figura 2). Deberían ver algo como en la figura siguiente y luego don OK: ahora las líneas deberían estar alineadas.

Antes de cargar el resto de los archivos, cambien el sistema de coordenadas del Data Frame para que nos avise cada vez que se cargue algún archivo que no tiene el mismo sistema. Para esto, vayan a las propiedades del Data Frame y en la pestaña **Coordinate System** simplemente eligen el coordinate system que se encuentra dentro de la carpeta **Layers**. Presionen OK y ahora el Data Frame ya tiene el mismo sistema de coordenadas que las dos capas.

Cargando el resto de los archivos

Ahora ya estamos en condiciones de agregar el resto de los archivos (son 4). El primero que vamos a agregar es el de Manzanas de la capital federal. Ya saben como agregar capas desde el botón **Add Data**. Una vez que eligieron el de Manzanas les va a decir que no corresponde a la proyección del Data Frame. En realidad sí coincide, solo que ArcMap no reconoce al *geoides International 1909* Hayford como igual al *International 1924* (no se preocupen por este tecnicismo, lo vamos a ver un poco más profundamente en otras clases). Para que no continúe con este error de aquí en adelante, lo que vamos a hacer es ir a **Data Management**

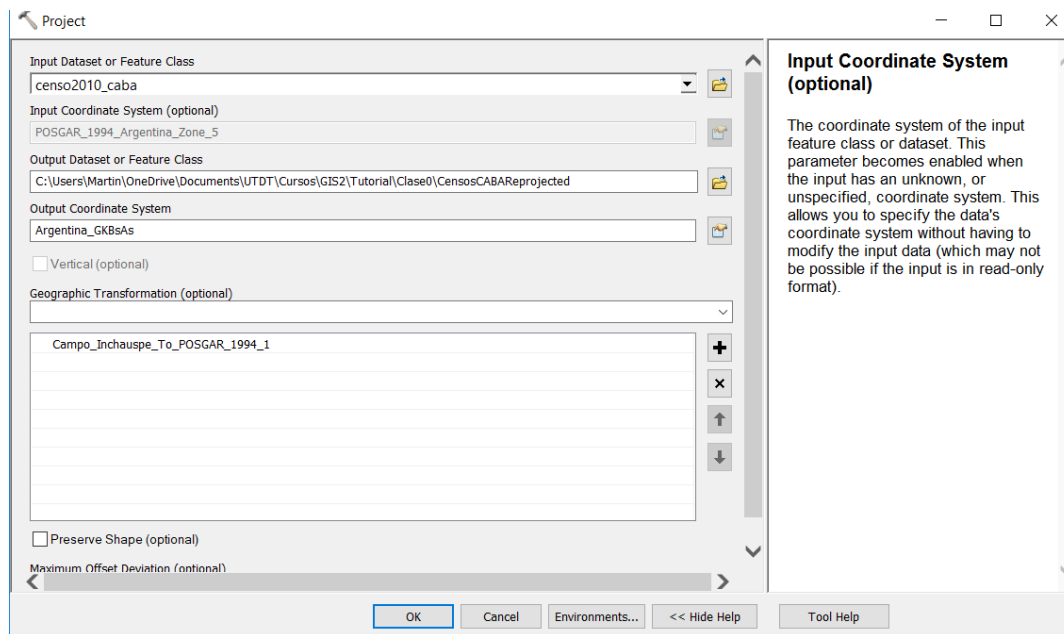


Figura 2: Menú de la herramienta Project de ArcMap. En el input dataset o feature class deben elegir la capa a modificar. En la segunda capa, la dirección y el nombre donde estará guardada la capa. En la tercera, el nombre del sistema de coordenadas de destino. Luego, se puede elegir la transformación que tenga ArcMap.

Tools -> Projections and Transformations -> Define Projection. Lo que hace esta herramienta es sobrescribir la información sobre el sistema de coordenadas de los objetos. No realiza una transformación y hay que estar siempre atentos a esto, pero nosotros sabemos que son exactamente el mismo sistema, solo que no reconoce a los *geoides* como idénticos. Elegimos el mismo sistema de coordenadas con el que estamos trabajando y damos OK.

Ya casi tenemos a todas las capas con el mismo sistema de coordenadas. Solo nos queda realizar lo mismo que hicimos con las manzanas en los hospitales de internación que, si bien ArcMap lo reconoce como el mismo sistema de coordenadas, tiene un nombre distinto y nos puede llegar a traer problemas. Una vez que hicieron la definición de la proyección, desde las proyecciones del Data Frame deberían ver solo una para las seis capas: “Argentina_GKBSAs”.

Sobre cómo almacenar la información del proyecto: existe una tentación a no guardar lo que se está haciendo, o hacerlo solo como un .mxd (un proyecto de ArcMap) para luego volver a utilizarlo más adelante. Sin embargo, es importante comprender que existen limitaciones en esta estrategia, más que nada es muy poco práctico para su divulgación. Lo que vamos a hacer es guardar en una **geodatabase** todas las layers que ya tenemos convertidas al mismo CRS para futuro uso en caso de ser necesario. Para esto primero les sugiero que cambien los nombres por default de las capas por unos que sean más descriptivos. Por ejemplo, en lugar de ‘centros-de-salud-nivel-1’ podría llamarse CESAC. Una vez que hayan hecho esto, vayan a **Catalog** y creen una nueva **Geodatabase** con el nombre que ustedes quieran. Hagan click derecho en ella una vez que esté creada, vayan a **Import** y luego **Feature Class (multiple)**. Una vez allí, seleccionen las 6 layers, arrástenlas a la ventana y den OK. Una vez que lo hayan hecho, desde **Catalog** van a poder administrar a esta base de datos. Van a ver que tiene 6 capas con los nombres que ustedes eligieron.

Procesamiento de datos

Una de los principales componentes de cualquier proyecto de datos es su procesamiento desde su forma inicial hasta que nos sirva como insumo para los análisis, en este caso la generación de un índice que nos indique la necesidad de la construcción de un nuevo CESAC. En este trabajo debemos hacer transformaciones para todas las capas. Comencemos con la capa de Calles

De calles a avenidas

La capa, cuyo nombre default es “callejero”, cuenta con información sobre todas las calles de la ciudad. Sin embargo, no todas las calles tienen la misma capacidad de mover pasajeros: las avenidas permiten una mayor cantidad de vehículos que puedan trasladar a las personas. Queríamos contar solo con las Avenidas para nuestro trabajo y vamos a hacerlo mediante la herramienta **Select by attributes (Selection -> Select by Attributes)**.

Como todas las herramientas en ArcMap, toma elementos y devuelve algo basado en los “parámetros”. En este caso, debemos elegir en qué capa hay que buscar los elementos, un método y luego una consulta SQL (Figura 3).

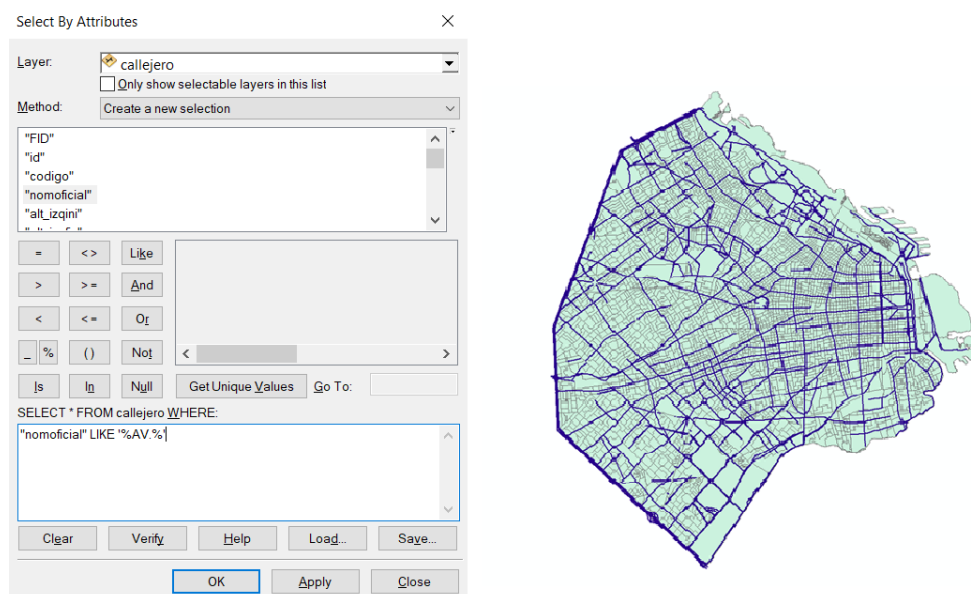


Figura 3: Menú de la herramienta Select by attributes. Permite seleccionar casos según una búsqueda SQL en la attribute table de una determinada capa.

La capa que tienen que elegir es la que tiene los datos de las calles. El método es “Create a new selection”, ya que queremos quedarnos con los datos de avenidas y descartar el resto. Para la construcción de la consulta SQL, la variable “nomoficial” tiene el nombre completo de la calle o avenida (pueden ver esto desde la **Attribute Table** asociada a los shapefiles de calles). Lo que vamos a hacer es quedarnos solo con los que tengan la palabra avenida en el nombre oficial. Si revisan algunos de los casos van a ver que la abreviación de avenida es AV., por lo que debemos seleccionar todas las filas que contengan dentro de todo su texto el *substring* “AV.”. El operador clave es LIKE, que busca en la columna (que se indica a la izquierda) lo que esté a la derecha. Además de la palabra que buscamos encontrar, agregamos dos signos de porcentaje, uno adelante y otro atrás, que nos dice que “AV.” puede estar precedido o sucedido por cualquier otro carácter.

Una vez que tenemos hecha la selección debemos crear una nueva capa solo con estas features. Esto es realmente simple: hacen click derecho sobre la capa de calles, van a **Selection** y desde allí a **Create Layer from Selected Features**. Llamen “Avenidas” a la nueva capa y guardenla en la geodatabase que crearon

anteriormente. Si quieren, pueden eliminar del proyecto la capa de Calles ya que no la usaremos de acá en adelante.

Asignando los barrios y comunas a los datos censales

En estas primeras clases vamos a trabajar primero con las Comunas 8 y 9, para luego expandir el análisis hacia otras comunas. Aun teniendo otra forma de segmentar el espacio para trabajar con las Comunas 8 y 9, vamos a hacer el trabajo de asignar barrios y comunas a los datos de los radios censales, ya que puede ser de utilidad en otros proyectos del curso. De cualquier manera, es una buena excusa para repasar algunas de las funciones de ArcMap

La herramienta clave es **Spatial Join**, herramienta que deben haber usado intesamente durante la materia del año pasado. A diferencia de un join en base a si dos columnas comparten el mismo valor, en spatial join se asignan los valores según dos entidades compartan cierto atributo espacial (por ejemplo, que se crucen en al menos un punto).

Para usar la herramienta tenemos que ir a **ArcToolbox** → **Analysis Tools** → **Spatial Join**. En este caso, nuestras target features son los polígonos del censo, a los que queremos asignarle información que está contenida en el dataset de barrios. Existen dos opciones de join (que entran en efecto cuando existen más de un match entre nuestras input feature y las output feature). Dejemoslo en el default. También usen “Intersect”, el default, como el método para determinar que dos objetos se encuentran en un mismo espacio geográfico.

Observen si el matcheo de los barrios fue perfecto. Si no lo fue, ¿Por qué les parece que esto fue así? Les doy una pista: hagan zoom en los bordes de los barrios y los polígonos de los censos ¿Hay un matcheo perfecto? Finalmente, realicen el mismo spatial join pero ahora usando el método **HAVE THEIR CENTER IN** ¿Mejoró el matcheo entre barrios y polígonos? ¿Hay algún polígono mal clasificado? ¿Por qué creen que esto sucede? Trabajaremos con este último *spatial join*

Hagan zoom (1:30000) en el límite entre los barrios de Palermo y Belgrano. Verán que uno de ellos se clasificó como de Palermo, aunque en realidad es de Belgrano. Esto obedece a que el centroide de este polígono, debido a su forma, se ubica dentro del polígono de Palermo. Esta es una buena oportunidad para recordar el uso del **editor**. Si no lo tienen activado, solo tienen que hacer click secundario donde está la barra de tareas y seleccionarlo (Ver Figura 4)

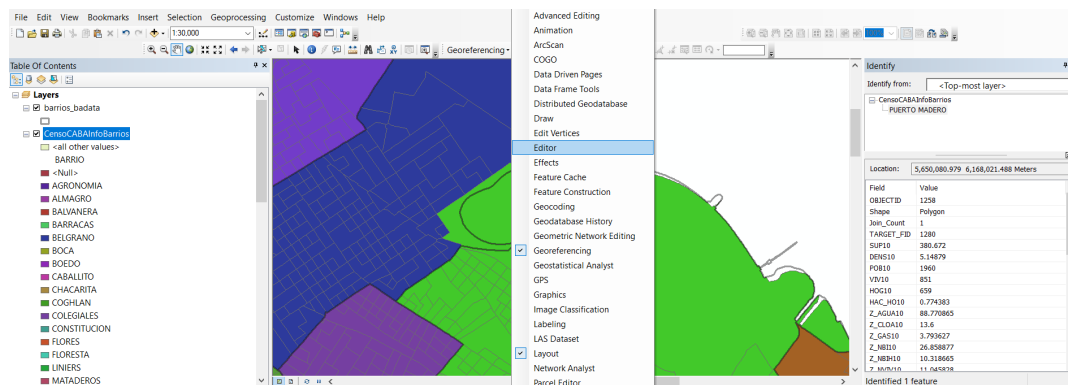


Figura 4: Haciendo click derecho en la barra de herramientas pueden elegir qué barra agregar a la ventana de trabajo. En este caso solo tienen que agregar el editor.

Una vez que pueden acceder al editor simplemente despliegan el menú **Editor** y hacen click en **Start Editing**. Solo pueden trabajar de a una capa a la vez, así que eligen la de los radios censales con datos derivados del spatial join, en mi caso se llama “CensoCABAInfoBarrios”. Hagan click en OK. Seleccionen el polígono, abran la **Attribute Table** correspondiente a la capa con la que están trabajando y finalmente reemplazan

PALERMO por BELGRANO. Como siempre, luego de haber hecho los cambios correspondientes deben seleccionar **STOP EDITING** para salir del modo edición.

Van a observar que existen algunos polígonos que no tienen ningún barrio asignado. En ArcMap estos valores no conocidos se representan como . Vamos a trabajar sobre estos datos utilizando **Select by attributes**. Tienen que elegir tanto la capa sobre la cual van a trabajar y la consulta, tal como aparece en la Figura 5. Presten atención a que cuando se quiere seleccionar el valor *Null* no hay que usar el signo igual, sino el operador IS y luego NULL.

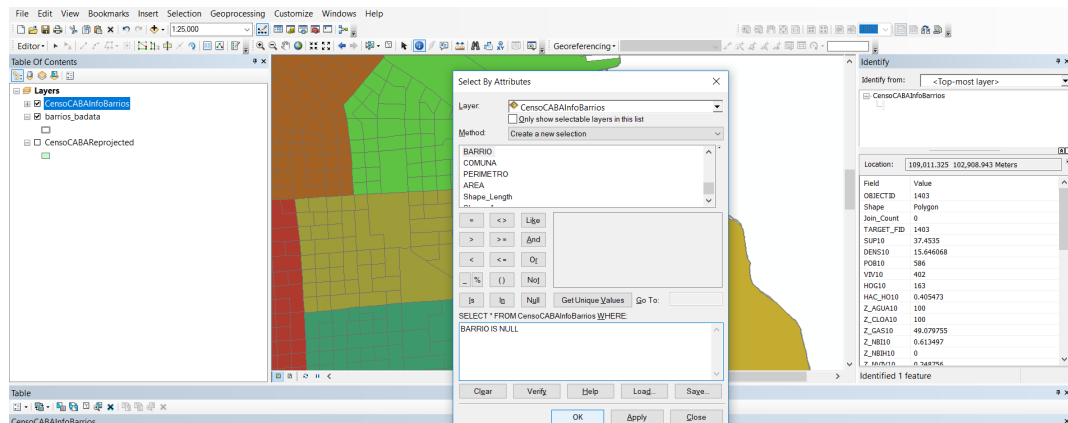


Figura 5: Select By Attributes. En esta ventana pueden elegir las feature en base al valor que tengan en su tabla de atributos. Para encontrar los valores Null no hay que usar el operador =, sino 'IS'

Una vez que apreten OK van a tener seleccionados algunos polígonos. Sin embargo, para visualizar mejor donde están los radios censales que no tienen datos, van a crear una layer que contenga solo estos polígonos. Para esto tienen que hacer click secundario en la capa donde tienen hecha la selección, ir a *Selection* y luego *Create Layer from Selected Features*. Desactiven todas las capas menos la última que crearon ¿Qué patrón observan en los polígonos que no tienen asignado ningún dato en el barrio?

Como ven, existen dos zonas de polígonos sin barrios asignados: Puerto Madero y la zona del Riachuelo: ¿Tienen alguna idea de por qué pasó esto?. Lo que vamos a hacer es editar el valor tanto del Barrio como de la Comuna para los de puerto madero (no se olviden de editar la capa que contiene a todos los polígonos y que el numero de comuna de Puerto Madero es la 1) y luego vamos a descartar a los que todavía no tienen barrios asignados (los que están en la zona del Riachuelo). Pista: para seleccionar a todos los que NO son Null pueden usar el operador NOT luego de IS. Una vez que hayan seleccionado a los polígonos con datos válidos, creen un nueva capa con esta selección.

Eliminen todas las capas de radios censales de la ciudad que no usaremos y guarden en la geodatabase los nuevos poligonos, limpios y con nueva información

Incorporando información a las manzanas

Ya estamos en condiciones de agregarle información a nuestras manzanas para poder generar el índice que nos dirá dónde hay una necesidad de construcción de los CESAC. Comenzaremos con la estimación de la distancia a la avenida más cercana con la herramienta **Near**, ubicada en **Analysis Tools**. Solo debemos agregar en INPUT FEATURE a los objetos a los que queremos agregarle la distancia, en este caso las manzanas, y en NEAR FEATURES a aquellos objetos con los que queremos medir la distancia más cercana (las avenidas). Hagan click en OK ¿Funcionó? Puede que sí o que no, dependiendo de la computadora (o al menos esa es mi experiencia) ArcMap suele tener algunos problemas cuando se buscan hacer operaciones sobre una gran cantidad de datos, luego veremos que otros programas (como R o Python) tiene menos límites en este sentido. Los que tuvieron algún problema trabajen de ahora en más solo con los datos de las comunas del sur (8 y 9).

¿Qué hacer si falló Near?

Primero vamos a generar una capa que contenga a las dos comunas para luego usar este “molde” para cortar al resto de las capas. Solo tenemos que usar la herramienta de geoprocessing **DISSOLVE (Geoprocessing -> Dissolve)**. Esta herramienta elimina los límites internos de los polígonos de acuerdo a una variable, en nuestro caso la variable COMUNA (Figura 6). Luego debemos seleccionar las comunas con las que vamos a trabajar mediante **Select by Attribute**. Creen una nueva layer en base a estos polígonos.

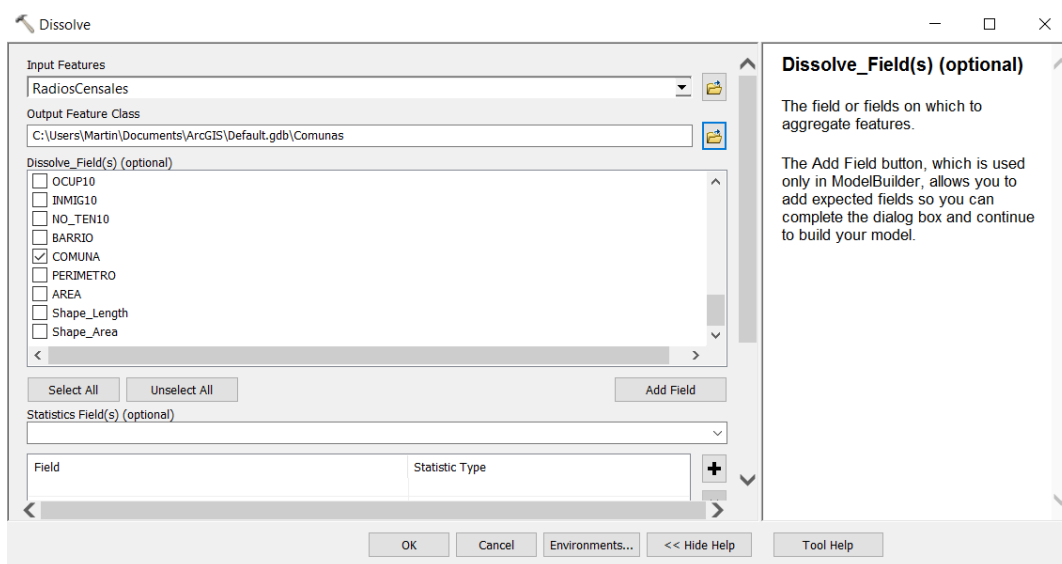


Figura 6: Cómo elegir las avenidas en base a un matcheo parcial de una cadena de texto.

Finalmente, con estos polígonos usen **CLIP (Geoprocessing -> Clip)** para avenidas y manzanas. A las nuevas capas les agregué “clip” al final, pero pueden nombrarlas como más fácil les resulte. Ahora sí podemos usar la herramienta **NEAR** y no va a fallar. Veamos qué es lo que nos devuelve.

Independientemente de si tuvieron o no que hacer el análisis sobre un subconjunto de datos, ahora en la capa de las manzanas deberían tener dos nuevas variables: “NEAR_FID” y “NEAR_DIST”, que respectivamente nos indican el ID de la feature más cercana y la distancia correspondiente (en metros). Como vamos a hacer otra aplicación de **NEAR**, que reemplazará el cálculo que hicimos recién, creamos una nueva variable y le asignamos el valor de **NEAR_DIST**. Ahora sí apliquen nuevamente **NEAR** para los datos de los CESAC ya existentes. Guarden la variable de distancia en una nueva columna (yo la llamé **CESAC_DIST**). Si bien la distancia con los CESAC no va a ser una variable que forme parte del trabajo en esta instancia

Nos falta la información sobre la densidad poblacional y hogares con NBI, que está en los polígonos a nivel Censal. Un simple **spatial join** entre los polígonos de las manzanas y los de los radios censales alcanzará. Elegí solo algunos campos para que se agregaran a los datos de las manzanas, ya que hay muchos más en los radios censales y no los usaremos. Como método de “asignación” de polígonos recomiendo **HAVE_THEIR_CENTER_IN** (Figura 7).

Agregando información sobre la cobertura actual de oferta de internación

La distancia con los hospitales con capacidad de internación es un poco más compleja, ya que además de la distancia física queremos que incluya alguna medida de la capacidad de internación que tienen. Para esto vamos a tener que hacer un **Kernel Density** (hay otras formas de hacerlo, pero esta es adecuada) que genere un espacio continuo a partir de los puntos, de tal manera que los puntos más cercanos con mayor cantidad de camas de internación tengan un valor más alto que los otros.

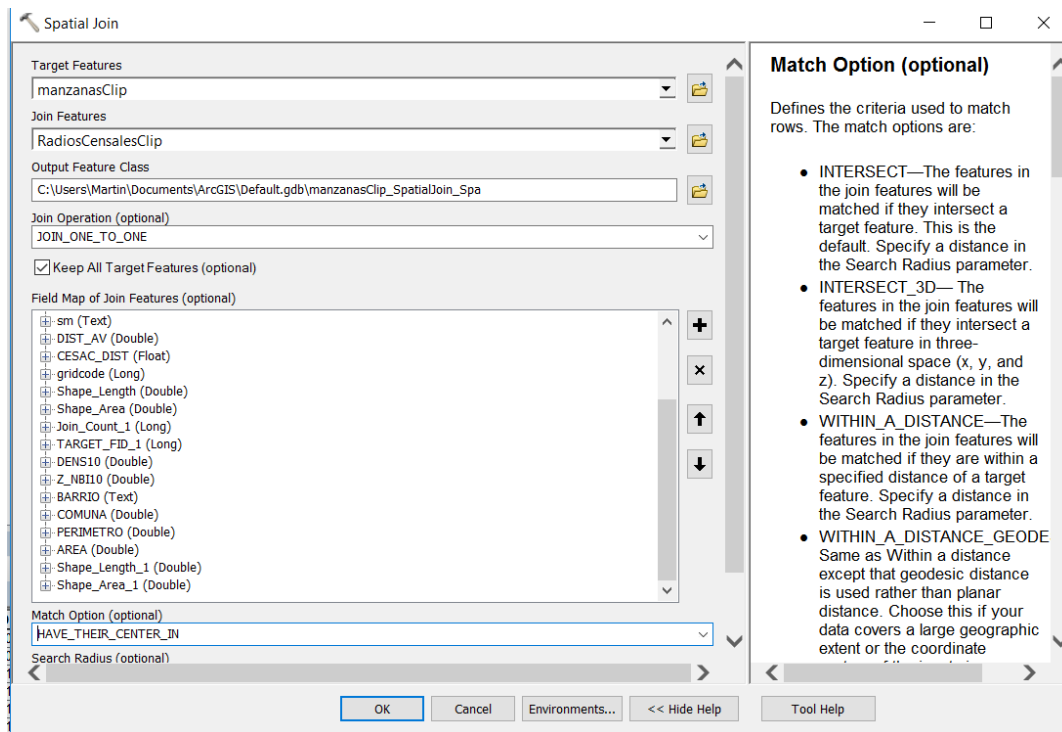


Figura 7: Menú de la herramienta spatial join

Para esto, vamos a ir a **Kernel Density (Spatial Analyst Tools -> Density -> Kernel Density)**. En la Figura 8 les muestro los parámetros que yo usé, pero pueden elegir los que crean necesarios. Por otro lado, tengan en cuenta que deberían elegir el espacio donde se va a realizar el análisis, porque sino ArcMap por alguna razón toma como límite a los hospitales más lejanos y no los computa. Para esto, deben ir a **Environments...** y en **Processing Extent** elijan **same as display**, aunque antes de esto procuren estar seleccionando en la ventana todo el área de que estamos trabajando.

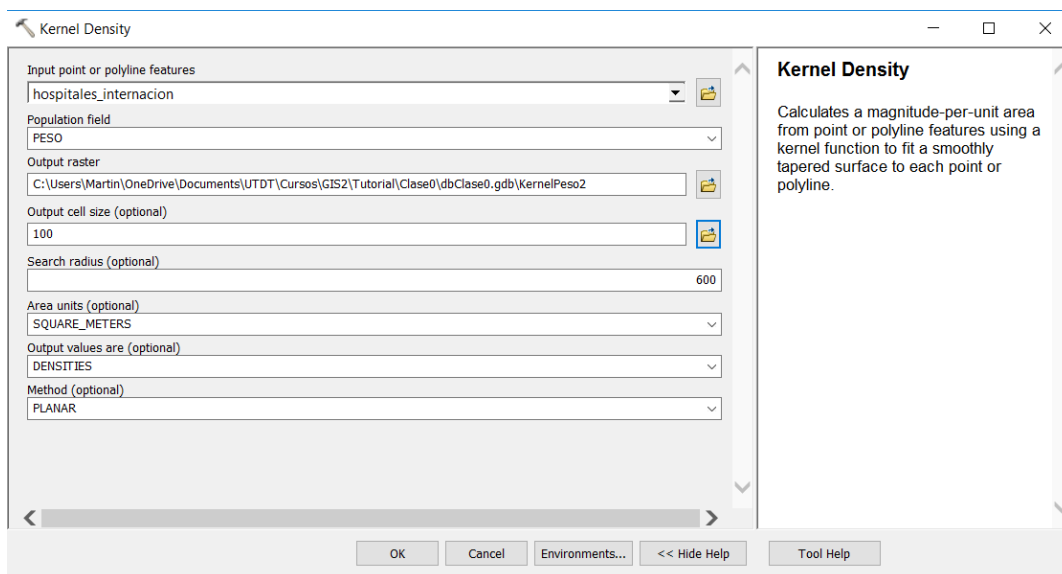


Figura 8: Menú para calcular un Kernel Density.

Una vez que contamos con nuestra capa que contiene una medida continua de la disponibilidad de camas de internación debemos reclasificarlo para que cuente con una tabla de atributos.. Ya que estamos, vamos a aprovechar esta oportunidad para reclasificar a los valores de 1 a 5 en base a quintiles (¿Saben lo que son? Simplemente se ordenan los casos de menor a mayor y se divide en conjuntos de igual tamaño, desde los más pequeños hasta los más grandes). Para hacer esto tenemos que ir a **Reclassify (Spatial Analyst Tools -> Reclass -> Reclassification)**.

Allí eligen al objeto que contiene el kernel, van a **Classify** y segmentan por *quantiles* y 5 *clases* (de esta manera serán quintiles). Dan OK y luego **Reverse new values**, ya que en nuestro índice ponderado queremos que este valor sea más alto cuanto menos cobertura haya por parte de los hospitales actuales. Lo que tenemos ahora es un **raster** con los valores reclasificados, pero necesitamos asignar estos valores a nuestras manzanas, que son polígonos. Si intentan hacer un spatial join, que es lo que haremos más tarde, verán que ArcMap no los dejará hacerlo. Para eso, necesitamos convertir nuestros rasters en polígonos.

La conversión de rasters a polígonos (y muchas otras) se hacen vía las herramientas que están localizadas en **Conversion Tools**. En este caso, hay que usar **Raster to Polygon**, dentro del menú “From Raster”. Tienen que elgir como INPUT RASTER al nuevo raster reclasificado. Luego, deben hacer un simple spatial join entre las manzanas y nuestra nueva capa de polígonos - recomiendo usar HAVE_THEIR_CENTER_IN.

¡Todavía no terminamos de incorporar la información de los polígonos! Es muy probable que varios de ellos no tengan información sobre la disponibilidad de oferta de salud y por lo tanto tengan un valor **Null**. Para generar nuestro indicador de “demanda insatisfecha” de un centro de salud debemos contar con valores numéricos en todas las variables. Tener un valor Null en está variable en nuestro caso equival a tener el valor más alto de todos (5) debido a que no hay oferta actual de servicios de internación. Reemplazaremos con “5” cuando el valor sea Null y el resto de los valores los vamos a dejar tal como están.

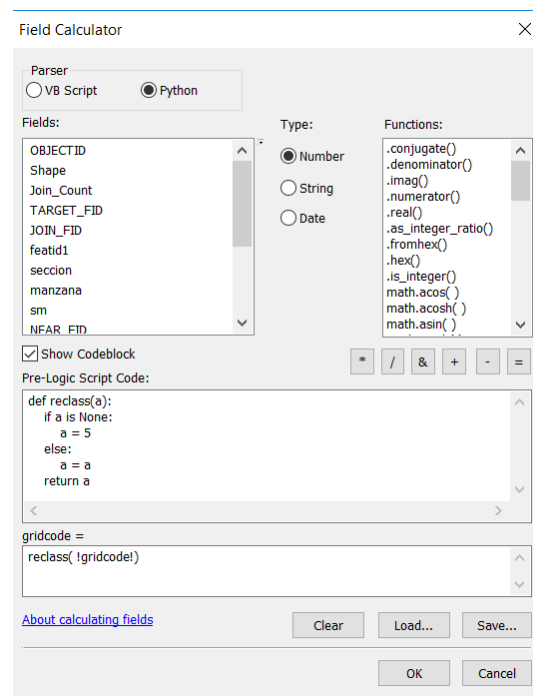


Figura 9: Reemplazando los valores de Null por un valor que pueda ser computable.

Esto lo podemos hacer de al menos dos maneras. En este caso usaremos Python desde el Field Calculator, tal como se muestra en la Figura 9. A grandes rasgos, lo que hace es definir una **función** en Python (ya verán como se hace esto en R), y luego aplican esta función a cada uno de los valores.

Ahora sí, finalmente tenemos la información codificada sobre la oferta hospitalaria a nivel de manzana!

Último paso del procesamiento de datos: quintiles.

ArcMap brinda la posibilidad de generar los quintiles o las agrupaciones de números continuos de una manera muy simple para los rasters, pero hay que escribir funciones más complejas en Python para hacer lo mismo en polígonos. En esta materia y en las que siguen de la especialización en datos van a ver R y no Python, por lo que no vamos a introducirnos de lleno en este lenguaje. Más adelante van a poder hacer estos calculos de una manera mucho más simple en R, y luego lo profundizarán en materias posteriores.

En este caso vamos a calcular los quintiles de una manera creativa, minimizando el código en Python. Comiencen con la distancia a la avenida más cercana. Selecciones está variable desde **Symbology** y eligan 5 grupos según Quintiles. ArcMap no tiene la función para guardar estos valores en la tabla de atributos, pero ustedes pueden ejecutar un simple código de Python en el que solo tienen que modificar los puntos de corte. Por ejemplo, en este caso la Figura les muestra los cortes que sugiere la simbología y cómo se pueden codificar estos en el código de Python. En Pre-Logic Script Code hay que definir la función, que en este caso se llama **quantile**. Lo que hace esta función es tomar los valores de una variable y devolver un número según si el valor es menor o igual a los quiebres que nosotros le indicamos. Es una forma manual de generar los quintiles.

Recomiendo fuertemente crear una nueva variable que contenga estos quintiles. Luego, en la caja inferior de **Field Calculator** tienen que llamar a la función y, entre paréntesis y signos de exclamación, la variable sobre la cual quieren generar los cuartiles (Ver figura 10). Todo esto parece difícil - y realmente lo es a esta altura - lo harán sin mayores problemas en R más adelante. Repitan este proceso para la **DENSIDAD** y **NBI**.

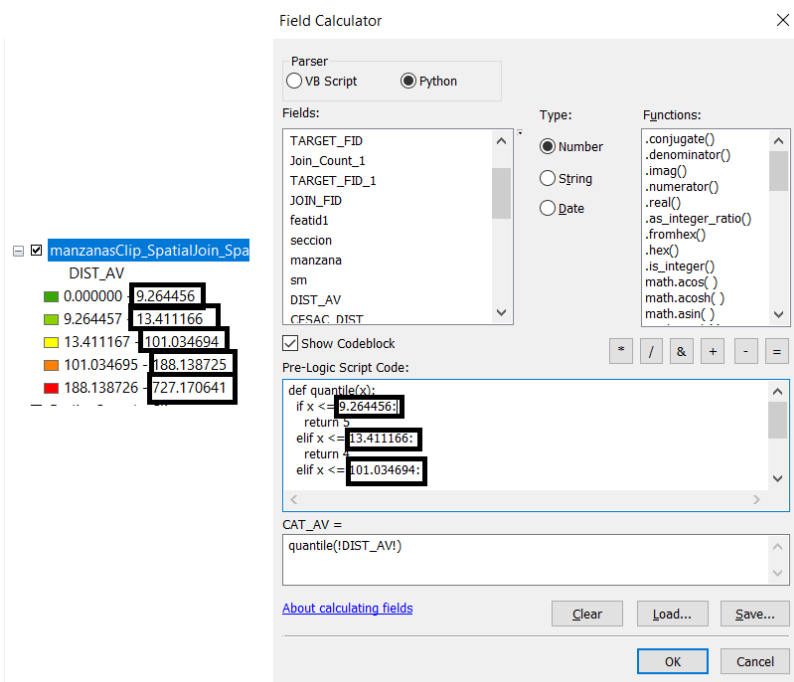


Figura 10: Cómo discretizar una variable continua cuando trabajamos con feature data.

```
def quantile(x):  
    if x <= 9.2644561:  
        return 5  
    elif x <= 13.411166:  
        return 4  
    elif x <= 101.034694:  
        return 3
```

```
elif x <= 188.138725:
    return 2
else:
    return 1
```

Generando el Índice de Demanda de Salud

Ya contamos con cuatro variables con las que vamos a construir nuestro Índice de Demanda de Salud de servicios de salud. Hay que crear una variable, que yo llamé **CESAC_INDEX**, y luego en **Field Calculator** podemos armar de manera sencilla un indicador que pondere cada valor para generar el índice. En mi caso particular, creé un índice que pondera un 10% el valor de distancia con la avenida, un 50% a la cantidad de NBI, un 10% la densidad poblacional y un 30% según que tan atendida por los hospitales por internación está la zona. El código en VB (Visual Basic) es la siguiente:

$$0.1 * [\text{CAT_AV}] + 0.5 * [\text{CAT_NBI}] + 0.1 * [\text{CAT_DENS}] + 0.3 * [\text{gridcode}]$$

Por supuesto, no olviden cambiar los nombres de las variables de acuerdo a los suyos y pueden hacer todos los cambios que quieran en los pesos, recuerden que siempre sumen 1. Vamos a hacer el análisis sobre las comunas del sur **así que si no tuvieron que hacerlo antes, ahora seleccionen las comunas del sur**. Pero lo vamos a hacer de una manera un poco distinta: un **spatial join** entre las manzanas de las ciudad y los barrios. Una vez que tengan los datos de comunas y barrios en la capa de manzanas **guardenla en la geodatabase**, ya que nos va a servir para más adelante.

Ahora quédense solo con las manzanas de los barrios del sur de la ciudad y mapeen el índice desde Symbology: ¿Dónde detectan los mayores más altos de demanda insatisfecha? Agreguen los datos sobre los CESAC actuales: ¿Creen que los CESAC están cubriendo la demanda insatisfecha que estimamos sin tenerlos en cuenta?

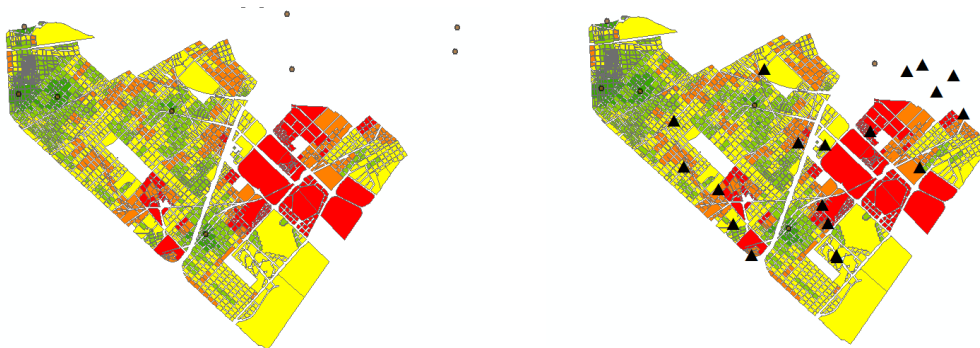


Figura 11: Mapa de quintiles del Índice de Demanda de Salud. Los puntos son hospitales y los triángulos son hospitales.

En mi caso la respuesta a esta última pregunta es que sí: que los CESAC están muy bien ubicados allí donde nuestro indicador dice que hay una demanda insatisfecha. Vamos a suponer entonces que las manzanas que están a un radio de 400m desde algún CESAC es un área que se encuentra bien servida, por lo que reemplazaremos a esas variables por el valor 1, el más bajo que puede tomar en nuestro indicador. Este código en Python es un poco más complejo, pero no se preocupen: ya van a aprender cómo hacer estos cálculos de una manera muy simple en R. El código que deben poner en el **Field Calculator** es el siguiente, tal como hicieron en otras ocasiones (Figura 12)

Ahora sí: ¿En qué manzana o manzanas propondrían crear los nuevos CESAC? Prueben distintos criterios de generación del índice ¿Siguen eligiendo las mismas manzanas? Una vez que estén cómodos con el criterio que

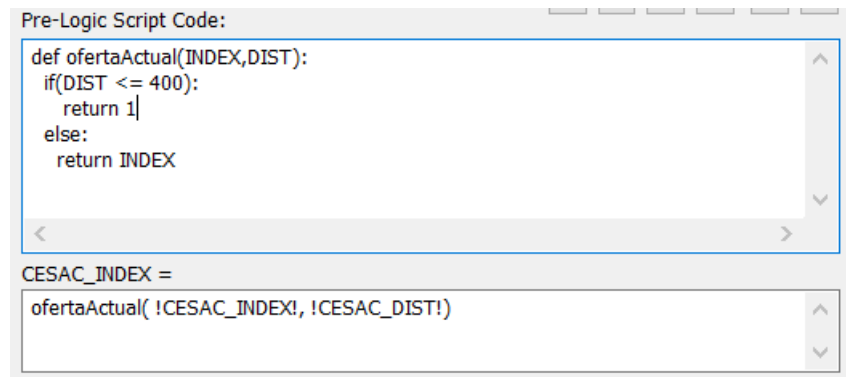


Figura 12: Detectando si una manzana tiene un CESAC en un radio de 400 metros o no.

eligieron, guarden el proyecto como “Clase Repaso”. Lo usaremos más adelante para comunicar los resultados y replicar lo hecho en otras comunas.