

EDUCACIÓN Y ANÁLISIS GEOESPACIAL

HERRAMIENTAS PARA LA PLANIFICACIÓN

Education and geospatial analysis

Tools for planning

- *Diego Quartulli*

UNAJ - Instituto de Ciencias Sociales y Administración, UNAJ - Dirección de Gestión Académica, DGCyE – Subsecretaría de Planificación

- *Manuel Riveiro*

UNAJ - Instituto de Ciencias de la Salud

RESUMEN

En este documento se exploran diferentes técnicas que hacen uso de datos espaciales que se consideran útiles para la planificación de, en términos más generales, diferentes sistemas sociales, y en especial del sistema educativo. De este modo, el objetivo de este trabajo es promover el uso de información georreferenciada para la toma de decisiones a diferentes niveles de los sistemas educativos. Para ello se revisan algunas herramientas comúnmente utilizadas en los análisis espaciales y se propone su aplicación al análisis del sistema educativo. Esto permite visibilizar una serie de relaciones entre el espacio y el sistema educativo que, en el mejor de los casos, sólo se intuyen y, en el peor de los casos, se desconocen.

A lo largo de este trabajo se detallan la construcción de indicadores analíticos (como la distancia lineal, de ruteo y sus derivados), el uso de límites políticos - administrativos (partidos, fracciones y radios censales), el empleo de técnicas con base en la geometría (buffers, polígonos envolventes convexos y cóncavos, centroides y diagramas de Voronoi) y en el análisis de patrones de puntos (densidades).

Algunas de las técnicas propuestas pueden servir para la construcción de medidas tanto normativas como descriptivas. Se afirma que estos últimos pueden ser útiles tanto para la exploración de los propios datos como para analizar el grado de cumplimiento de los primeros. Por otro lado, algunas técnicas pueden servir de insumo para la toma de decisiones a nivel local mientras que otras se pueden aprovechar más a nivel central.

ABSTRACT

This document explores different techniques that make use of spatial data that are considered useful for the planning of, in more general terms, different social systems, and especially the educational system. Thus, the objective of this work is to promote the use of georeferenced information for decision making at different levels of educational systems. To this end, some tools commonly used in spatial analyzes are reviewed and their application to the analysis of the educational system is proposed. This makes visible a series of relationships between space and the educational system that, in the best of cases, are only intuited and, in the worst of cases, are unknown.

Throughout this work, we detail the construction of analytical indicators (such as linear distance, routing and their derivatives), the use of political-administrative limits (parties, fractions and census radio), the use of techniques based on geometry (buffers, convex and concave envelope polygons, centroids and Voronoi diagrams) and in the analysis of point patterns (densities).

Some of the proposed techniques can be used to construct both normative and descriptive measures. It is stated that the latter can be useful both for exploring the data itself and for analyzing the degree of compliance with the former. On the other hand, some techniques can serve as input for decision-making at the local level while others can be used more at the central level.

PALABRAS CLAVES: Análisis Espacial - Planificación - Educación

KEYWORDS: Spatial Analysis - Planification – Education

3. INTRODUCCIÓN

“Todo está relacionado con todo lo demás, pero las cosas cercanas están más relacionadas que las cosas distantes”
(Waldo Tobler 1970, 285)

Este trabajo parte del supuesto que diferentes cuestiones espaciales afectan el funcionamiento del sistema educativo (Brock, 2016). Por esta razón se considera pertinente su utilización tanto en la planificación como en la toma de decisiones. En particular, dada las nuevas posibilidades que ofrece la posibilidad de georreferenciar grandes cantidades de datos nominales, busca promover el análisis territorial en mayor detalle.

En este sentido, a lo largo de este trabajo se analizarán una serie de técnicas unidas por el denominador común de aportar algún grado de utilidad cuando la cuestión espacial importa. Hasta acá llega la similaridad y comienza la disparidad. Las herramientas pueden aportar métricas cuantitativas (p.e. distancias), cómo también representaciones visuales (p.e. mapas)

y estas pueden usarse tanto para crear y resumir información espacial como generar interpretaciones que ayuden a la toma de decisiones en diferentes niveles.

Por esta razón, antes de entrar en la especificidad de cada técnica quizá sea pertinente aclarar lo siguiente. Existen varios usos para cada una de ellas y no todas sirven para el mismo uso. Lo anterior complica un poco la redacción/lectura del trabajo porque no se ofrece una taxonomía, una tipología o una simple clasificación de las técnicas que las ordene/clasifique. A cambio, siguiendo una bibliografía clásica, sí se ofrece un perfil de cada técnica, basada en comentarios sobre cada una de las siguientes dimensiones¹:

- Su utilidad para el análisis y diseño de la política pública (p.e. para planeamiento y evaluación)
- Su utilidad para la ilustración y/o empoderamiento (p.e. en el contexto de datos abiertos)
- Su utilidad para diferentes niveles como el centro (p.e. un ministerio) y lo local (p.e. un colegio)

Su uso para objetivos normativos y/o descriptivos (p.e. cuantos estudiantes cumplen tal o cual ley referida a cuestiones espaciales o de tiempo)

La estructura del trabajo será la siguiente. Primero se contextualizará teóricamente la problemática de la cuestión espacial en la educación ([Sección 3.1](#)). Luego se comentará someramente el origen de los datos con los que se trabajará para ejemplificar las técnicas ([Sección 4](#)) para posteriormente comenzar con la caracterización específica de cada técnica ([Sección 5](#)).

El orden en la presentación de estas últimas será el siguiente:

1. Técnicas espaciales **analíticas** ([Sección 5.1](#))
2. Técnicas espaciales basadas en **límites políticos administrativos** ([Sección 5.1.1](#))
3. Técnicas basadas en la construcción de **buffers** ([Sección 5.2](#))
4. Técnicas basadas en **polígonos de Voronoi** ([Sección 5.3](#))
5. Técnicas basadas en **polígonos envolventes** ([Sección 5.4](#))
 - 5.1 Construcción de **polígonos convexos** ([Sección 5.4.1](#))
 - 5.2 Construcción de **polígonos cóncavos** ([Sección 5.4.2](#))
6. Técnicas basadas en análisis de **centroides** ([Sección 5.5](#))
7. Técnicas basadas en análisis de **densidades** ([Sección 5.6](#))

¹ Las dimensiones se basan en una adaptación de las propuestas por Kenneth Land [Land (1983)](Land y Ferriss 2007, 522-23).

Finalmente se mostrará, a modo de recomendación, una aplicación práctica que incluya algunas de las técnicas vistas ([Sección 6.1](#)) y, finalmente, se pasará a las conclusiones del trabajo ([Sección 7](#)).

Finalmente se hará una conclusión en donde se realizará una síntesis que, más que repetir las conclusiones individuales de cada técnica, hará énfasis en la complementariedad y sinergia de los diferentes instrumentos analizados.

3.1 MARCO TEÓRICO

Las ciencias siempre han tenido una relación bifronte con las cuestiones espaciales y las ciencias sociales no son la excepción (Sack, 1980, p. 3). Especialmente en el pasado, para algunos el espacio sólo se ligaba a ciertas disciplinas que implicaban cuestiones formales aplicadas, cercanas a la geometría y la topología, como pueden considerarse los casos de la agrimensura o la geodésica. Para otros, sólo se ligaba a las ciencias naturales en donde se entendía al espacio como algo concreto caracterizado por variables que son el resultados de operaciones empíricas de medición como el largo o el ancho sobre cosas concretas o, como se decía en tiempos de Descartes, sobre la *res-extensa*. Una expansión de esta última acepción es la que estuvo detrás del surgimiento del concepto de “región” dentro de la geografía en forma complementaria al surgimiento de los Estados-Naciones y su creciente necesidad de planificar el desarrollo de estos (Llanos Hernández, 2010).

Con el paso del tiempo, especialmente dentro de la geografía humana y con una vinculación cada vez más transparente con las ciencias sociales, el concepto de “región” fue cediendo frente a la mayor difusión del concepto de “territorio” que fue incluyendo dentro de sus significados no sólo a las propias personas sino a las relaciones que se dan entre ellas y como estas modifican el propio territorio (Capel, 2016).

Si bien los párrafos anteriores pueden considerarse excesivamente generales sirven para mostrar algunas de las tensiones (presentes en este trabajo a través de las diferentes técnicas) que se pueden observar en la aplicación del análisis espacial dentro del mundo educativo. En este sentido, por un lado aquí se hará un uso moderado pero explícito de cuestiones geométricas, topológicas, estadísticas y teoría de redes que guardan una fuerte relación con disciplinas formales. Como se verá, la mayor explotación de este tipo de técnicas es permitida por la georreferenciación de los establecimientos educativos y, principalmente, la de los propios estudiantes. Por otro lado, las tensiones internas provenientes del concepto de territorio también están presentes a lo largo de las diferentes técnicas. Algunas de los usos de las técnicas parecerán más útiles para analizar el efecto local de las personas en los establecimientos educativos y, de manera derivada, de estos sobre el territorio del cual forman parte. Otros usos, a veces a través de una misma técnica, tenderán a indagar el efecto

inverso, esto es, el efecto que ejerce el territorio, como una gran variable contextual, en el propio funcionamiento de los establecimientos educativos y, de manera derivada, sobre el comportamiento de las personas que forman parte de ellos. En efecto, la propia idea intuitiva de “zona de influencia” posee un carácter bifronte. Se la puede utilizar tanto para analizar la influencia de las personas y el establecimiento sobre el territorio como la influencia del mismo territorio sobre aquellos.

La incorporación de la cuestión de las relaciones sociales en el significado del concepto de “terreno” hace visible otra tensión. Es la que existe entre las ganancias que trae una estrategia que maximice las ventajas de lo local frente a los costos de oportunidad de no favorecer la diversidad. Nos explicaremos en los siguientes párrafos.

Maximizar las ventajas de lo local, por ejemplo con políticas que induzcan a que los estudiantes se seleccionen en función de la distancia a los establecimientos educativos, permite la instauración de un sistema objetivo (Roth 2016, cap. 9), beneficios en salud pública sea tanto vía aumento de viajes peatonales (McDonald, 2008) y/o reducción de accidentes de tránsito vehiculares cuando se acompaña de caminos o senderos seguros (Dimaggio y Li, 2013) (Larouche et al. 2018). Por otro lado, también es plausible que este tipo de estrategias, vía el mecanismo social de clausura, produzcan un aumento del capital social tanto a nivel (de cada uno) de los miembros como de las propias comunidades de los establecimientos educativos (Coleman, 1988) (Coleman, 1990) (Burt, 2005). En efecto, este último razonamiento ha sido un supuesto clásico en la ubicación espacial de la escuela en la planificación urbana (Perry, 1914).

A nivel del establecimiento en general, pero más en particular en la composición de cada aula, el principal costo de oportunidad de estas estrategias es la (potencial) escasa diversidad que, por diseño, no permite que se produzca el beneficio de la interacción e intercambio entre diferentes (Burt, 2005). Estos costos parecen incrementarse en contextos de segregación residencial en donde, *ceteris paribus*, la estrategias de maximizar lo local favorecen una reproducción de la ya (pre)existente desigualdad residencial (Saporito y Van Riper, 2016). Una manera de reducir este costo de oportunidad es la creación de sistemas que, más que crear áreas que minimicen la distancia al establecimiento, prioricen la creación de áreas que maximicen la desigualdad social dentro de un determinado límite de distancia (Saporito, 2017) (Bryant, 2018) (Monarrez, 2019).

Como puede intuirse las tensiones anteriores suelen ser una fuente de discusión tanto política como académica constante entre las intersecciones de la educación y los análisis espaciales, el análisis de redes sociales y la bibliografía acerca del capital social. Aquí sólo los nombramos para advertir de que no se nos escapa que la aplicación de diferentes técnicas espaciales en la esfera de la educación no se trata simplemente de un problema técnico acerca de cómo resumir o visualizar datos.

Una característica en común de todas las medidas que se presentan a continuación es que, a tono con el epígrafe del comienzo de esta sección, en este trabajo se asume que los objetos que interactúan (sean estos el territorio, los establecimientos o las personas que forman parte de cualquiera de ambos) se ven menos influenciados por la distancia a medida que aumenta el valor de la misma. Más allá del parámetro que mejor describa esta relación (que puede ser diferente para cada nivel) usualmente se asume, y aquí poniendo el eje causal en las decisiones de las personas, que (la mayoría de) los agentes tienen en cuenta opciones educativas cercanas y “podan” de su árbol de decisión las opciones lejanas (Kjellström y Regnér, 1999).

4. METODOLOGÍA

Muchos de los datos que se utilizaron en este trabajo son datos oficiales disponibles de forma pública. Sin embargo, otros, de forma razonable por cuestiones de confidencialidad, son privados.

Ejemplos de los primeros son las ubicaciones de escuelas provistas por el equipo del Mapa Educativo de la Dirección de Información y Estadística de la Provincia de Buenos Aires y los polígonos de los departamentos y provincias provistos por el Instituto Geográfico Nacional. En esta misma categoría entran los polígonos de los radios censales provistos por el INDEC. Los polígonos de las fracciones censales son de construcción propia en base a la cartografía de los radios que forman parte de las mismas.

Ejemplos de los segundos, pueden considerarse los datos nominales (no públicos) de la distribución de los estudiantes tienen origen en la Dirección General de Cultura y Educación de la Provincia de Buenos Aires. El uso de este tipo de datos nominales es posible debido a los avances en la carga de los propios sistemas administrativos, más precisamente, del sistema Mis Alumnos. Debido a que se trata de datos confidenciales, la información fue anonimizada para su uso, imposibilitando así vincularla a datos de identificación individual, resguardando así la privacidad de las personas. Como se verá a lo largo de este trabajo, la utilización de datos nominales para la toma de decisiones, aún cuando su finalidad original tenga un objetivo principalmente administrativo, aumenta el tipo y calidad de las inferencias posibles. Al trabajar sobre datos anonimizados, y agregando la información individual a conjuntos vinculados a los establecimientos educativos y/o alguna otra unidad espacial, permite utilizar los datos para una gestión más inteligente sin representar ningún riesgo de privacidad de las personas.

5. RESULTADOS

En el análisis de datos educativos desde una lógica o perspectiva espacial, suele predominar el análisis a partir de zonas delimitadas por identidades político- administrativas (provincia, distrito, localidad, etc.) o el mapeo de información a partir de “puntos” para su posterior análisis e interpretación. La pregunta disparadora de este trabajo es: ¿cómo identificamos una zona de influencia de una escuela?²

En otras palabras, cómo pasamos de un “punto” a una “área” entendiendo, como dice el epígrafe, que las cercanías son (más) relevantes que las lejanías a la hora de entender el funcionamiento de los procesos sociales. En esta búsqueda repasamos algunos usos comunes y sistematizamos algunas técnicas cuya aplicación puede nutrir el análisis del sistema educativo. Un análisis que busca ofrecer herramientas que permitan captar aspectos más complejos y lograr así una mejor representación del proceso real que estudia.

Como se detalló en la introducción a continuación se describirán una serie de técnicas que incorporan la dimensión espacial al análisis educativo. Si bien muchas de las técnicas pueden aplicarse en diferentes niveles (local, provincial), a la hora de la ejemplificación de cada una de estas en este documento, se ha preferido el nivel local. La razón es que este nivel permite la captación de cierto nivel de detalle que sería inviable (en una publicación de este tipo con imágenes estáticas sin posibilidad de ajustar el zoom ni la cantidad de capas) si se hubiera optado por el nivel provincial. La estrategia anterior se complementó con una comparación sistemática entre dos establecimientos del área céntrica del partido de La Plata (uno de nivel primario y otro de nivel secundario) y otros dos establecimientos de áreas más menos urbanas.

A modo de agrupar y distinguir algunas de las características de las técnicas se dirá lo siguiente. Las técnicas acá denominadas *analíticas* son aquellas que, a pesar que para su construcción necesiten operaciones espaciales, mucha de su utilidad surge de su posterior independencia de su visualización en un mapa. En cambio, en el caso de los polígonos políticos – administrativos la existencia de esos polígonos es previa o ajena al propio sistema educativo pero su uso permite la relación visual y espacial con o bien otros datos sociales (como los censales) o bien con niveles de decisión políticos (como el partido).

En cambio, en las secciones siguientes (§4.3-4.7) se pasará a analizar una serie de técnicas que poseen en común la siguiente característica: Visualizan supuestos normativos o aspectos descriptivos del propio sistema educativo (sea de cada establecimiento o bien de todo el sistema educativo) y con ella producen un (nuevo) producto georreferenciable. En general,

² Sin embargo, como se indicará en varias notas al pie de este documento, las mismas técnicas se podría utilizar para otros objetivos que también implican alguna relación entre la educación y el análisis geo-espacial.

todas las técnicas aquí utilizadas producen como producto final un dato vectorial que luego se puede visualizar en un mapa.

En este sentido, primero analizaremos algunas aplicaciones para la construcción de polígonos como los Buffers (§4.3) y los polígonos de Voronoi (§4.4). Los primeros los utilizaremos aquí para visualizar principalmente criterios normativos y los segundos para visualizar el establecimientos más cercano a cada punto de la provincia. Luego, especialmente en el dominio del análisis de la matrícula de cada establecimiento, se analizará, para la captación de la forma o silueta de la nube de puntos de los estudiantes, la construcción de polígonos envolventes (§4.5). Luego, para la captación de la tendencia espacial central, se analizará la construcción de centroides (§4.6). Finalmente, para la captación de diferentes zonas de densidad, se realizará un análisis de densidades (§4.7).

5.1 TÉCNICAS ESPACIALES ANALÍTICAS

Existen técnicas que pueden ser consideradas espaciales pero que, sin embargo, sus resultados pueden tener una interpretación clara aún si aquellos no se encuentran posicionados o georreferenciados (y por lo tanto no pueden visualizarse en un mapa o insertarse como un nodo de una red). A este tipo de técnicas espaciales las llamaremos aquí *analíticas*.

Lo anterior, al menos cuando se complementa con otras técnicas que veremos más adelante, más que una desventaja puede considerarse una virtud ya que si bien se necesita de herramientas específicas de análisis espacial (p.e. GIS) para su **producción**, se puede prescindir de ellos para (muchos de) sus **usos** posteriores.

El ejemplo típico de este tipo de técnicas es la **distancia** (y muchos de sus derivados posteriores). La distancia, como lo sugiere la primera ley de la geografía de Tobler, puede considerarse como un indicador clave de los análisis espaciales en las ciencias sociales (Logan 2012, pag. 4-8).

Si bien para su producción o cálculo se necesita de datos georreferenciados su producto es un número con una unidad de medida asociada. Si pensamos que esta medida se puede calcular a un nivel desagregado de cada estudiante las opciones de síntesis posteriores son muchas. En este sentido, podremos por ejemplo producir una medida de distancia para cada estudiante:

- con su establecimiento educativo,
- con el establecimiento educativo más cercano (así como el con el segundo, tercero, etc.)
- con el compañero de curso más cercano (así como con el segundo, tercero, etc.).

De manera derivada cada medida anterior se puede agrupar en categorías de distancias y/o agregar y sintetizar mediante funciones estadísticas de tendencia central (media, mediana, etc.) así como con funciones de dispersión (desvíos, varianza, etc.) a diferentes niveles como el curso, el establecimiento, el distrito y la propia provincia. Inversamente luego se puede ordenar a cada estudiante, a cada establecimiento y a cada distrito en algún n -til (decil, quintil, etc.) con respecto a sus pares.

Más en profundidad, la misma idea de distancia también se puede complejizar en varias dimensiones. Por un lado, en vez de calcularse la distancia euclidiana (o alguna otra) sobre coordenadas proyectadas también se puede calcular sobre coordenadas geodésicas. Esta técnica, que arroja un resultado más preciso a costa de una mayor carga computacional del cálculo, no suele tener diferencias prácticas para distancias cortas como la de un estudiante a un establecimiento educativo.

Por otro lado, también se pueden agregar constricciones reales al cálculo como pueden considerarse la familia de las distancias de ruteo. Estas implican que la distancia se calcule siempre sobre alguna “ruta” considerada viable como caminos, puentes, etc. Este resultado será siempre igual o mayor al simple cálculo en “línea recta”. Un punto importante de este tipo de cálculo es que posteriormente permite transformaciones de la distancia en tiempo y se puede realizar un cálculo diferente para cada tipo de transporte (a pie, transporte público, auto, etc.) como los que ofrecen los servicios de GPS de los celulares modernos.

Desde un punto de vista conceptual la principal diferencia de la distancia de ruteo frente a otros tipos de distancias quizá sea su fuerte anclaje en la teoría abstracta de grafos. Esta es una característica compartida con otras estrategia de análisis de las ciencias sociales como la teoría de redes sociales. En efecto, una misma medida abstracta de la teoría de grafos es interpretada de manera diferente para la distancia de ruteo que para la teoría de las redes sociales.

Para mostrar esa particularidad quizá el mejor ejemplo sea la comparación del cálculo de las distancias de ruteo con el cálculo de otra distancia mucha más conocida como es la distancia euclideana. Aunque los usos de esta última sean múltiples, sus fundamentos provienen más de la geometría que de la teoría de grafos. Cuando se la suele aplicar al análisis espacial geográfico el cálculo de la distancia euclideana también se apoya en algún sistema de representación que “traduzca” la ubicación real de algún objeto a algún sistema de coordenadas proyectado. Esto último, que se suele encargar la geodesia como disciplina, constituye un sistema euclideo de 2 dimensiones y es sobre estos valores que se realiza el respectivo cálculo.³

³ Para aplicaciones simples de la vida cotidiana donde es posible trazar una línea recta entre 2 puntos (por ejemplo al medir una mesa con una regla) el cálculo es simplemente la diferencia absoluta entre los 2 puntos a medir. En este caso se dice que se trata de un espacio euclideo de una dimensión.

En cambio, en la familia de las distancias de ruteo el fundamento conceptual que se encuentra por detrás, más que provenir de la geometría o de la topología (de la que también utiliza muchos conceptos) proviene principalmente de la teoría de grafos. Especialmente aquí nos van a interesar el análisis de distancias de ruteo sobre mapas callejeros en donde las esquinas se interpretan como nodos y las cuerdas como aristas.⁴

Esto es posible tanto por la digitalización creciente de los mapas (especialmente los de acceso abierto como OpenStreetMap) como el mayor acceso a diferentes librerías (de distintos lenguajes de programación) que se basan o aplican algoritmos de la teoría de redes por detrás (Dijkstra, A-Star, etc.). En el caso particular de este trabajo, el grafo se construyó con la información vial de OpenStreetMap y en cuanto a las librerías utilizadas la mayoría fueron del lenguaje R aunque algunas de ellas ejecutarán por detrás diferentes librerías en C por cuestiones de eficiencia.

Dadas que las posibilidades de cálculos son muchas aquí se mostraran, a modo de ejemplo, sólo unas pocas.⁵ Por otro lado, ya que estas técnicas no implican su visualización a través de mapas, se pueden hacer análisis a altos niveles políticos-administrativos sin que haya problemas de resolución. En efecto, en esta ocasión la ejemplificación se mostrarán a nivel de toda la provincia para así mostrar que estas técnicas son escalables ya que aquí se utilizan datos de casi 20 mil establecimientos educativos y casi 3 millones de estudiantes.

Tabla 1: Media de medidas según tipo de ruteo a pie y según Sector, Ámbito y Región. Valores en Km

Característica	Estatal	Privado	Global	Estatal	Privado	Global
Total	1,7	1,9	1,8	1,3	1,5	1,4
Ámbito						
Rural Agrupado	6,8	16	7,3	3,7	11	4,2
Rural Disperso	14	9,9	13	8,5	6,8	8,4
Urbano	1,5	1,9	1,6	1,2	1,5	1,3
Nivel						
Nivel Inicial	1,4	1,7	1,5	1,0	1,4	1,1
Nivel Primario	1,5	1,8	1,6	1,1	1,5	1,2
Nivel Secundario	2,1	2,1	2,1	1,6	1,6	1,6

De manera derivada las técnicas anteriores también permiten la construcción de indicadores que indiquen si el estudiante cumple la respectiva normativa de distancia de su nivel educativo. *A posteriori* esto permite construir porcentajes para cada establecimiento, distrito, región y la propia provincia. Como veremos más adelante, algunos de estos cálculos también

⁴ Existen aplicaciones de las distancias de ruteo que no tienen de insumo un mapa carretero o vial. Piensese en su uso para cálculos sobre las cuencas (o redes) de los ríos en donde se intenta analizar las relaciones entre los ríos principales y sus afluentes y las propias relaciones entre diferentes cuencas (Eros y Lowe, 2019).

⁵ En (Quartulli y Aispuro, 2021) se realiza una muestra más exhaustiva de los posibles análisis que se pueden realizar con datos de distancias. En ese trabajo sólo se refieren a distancias lineales o euclidianas pero los mismos se pueden replicar a las distancias de ruteo.

se pueden visualizar en un mapa (Buffers, Voronoi) pero, nuevamente, los indicadores analíticos tienen la virtud de poder independizarse de los mapas e integrarse fácilmente al flujo de trabajo de equipos más familiarizados con hojas de cálculo (excel, drive, etc.) y, consiguientemente, integrarse más fácilmente con el resto de los datos educativos existentes.

Por último, las técnicas anteriores permiten que cada establecimiento, donde el propio manejo de sistemas GIS o de lenguajes de programación se encuentra mucho menos difundido, reutilice esos datos para alguna estadística interna con una simple hoja de cálculo sin depender de sistemas GIS. En cuanto a la publicación de este tipo de datos, siempre que se agreguen de alguna manera, podrían ser utilizados sin aparentes problemas legales.

A continuación, a modo de ejemplo que mantendremos a lo largo de todo el texto, hemos seleccionado 2 establecimientos públicos que pertenecen al partido de La Plata. Uno, de nivel primario y que siempre se ubicará a la izquierda y otro, de nivel secundario y que siempre se ubicará a la derecha. Indirectamente las medidas anteriores también sirven para controlar la calidad del dato que sirve como insumo para las técnicas posteriores. Por ejemplo, sirven para excluir excluyendo del análisis a ciertos valores extremos considerados improbables.

5.1.1 Límites políticos - administrativos

El sistema educativo puede considerarse como un sistema social artificial ya que es diseñado por la propia sociedad (Bunge, 1995). En este diseño existen divisiones **políticas** (región, distrito, etc.) sobre las cuales se toman decisiones. Análogamente puede decirse que el sistema estadístico construye divisiones **administrativas** (partido, radio, etc.) que, al menos en sus niveles más agregados, suelen coincidir con los primeros. El punto importante aquí es que sobre estas divisiones administrativas se publican datos que pueden servir para la evaluación o la toma de decisiones educativas. Cualquiera de los motivos anteriores hace que sea pertinente el análisis de la conveniencia de la producción de datos espaciales educativos a estos niveles. Ambos tipos de datos se suelen representar dentro del análisis espacial como polígonos vectoriales, esto es, como una figura que representa los límites de nivel respectivo.

En Argentina el mínimo denominador en común de ambos tipos de límites (del político y del administrativo) es a nivel de partido o departamento. En este sentido, los análisis a nivel de departamento representan el nivel más bajo en donde el nivel político coincide con la existencia de datos censales. Como veremos en seguida, este resulta un nivel muy agregado para hacer análisis para cada establecimiento aunque nada impide su utilización para otros objetivos. Por ejemplo, pueden usarse las fracciones censales para caracterizar variables contextuales del territorio de un establecimiento educativo (Born, 2019, pag. 36). Si bien se podría descender hasta el nivel del radio de cada establecimiento usualmente se prefiere la fracción, especialmente en zonas urbanas, por considerarlo:

- a) un estimador más robusto y menos volátil que el radio específico del establecimiento y
- b) por entender que las áreas normativas de distancia y tiempo de cada establecimiento generalmente se acercan más a los tamaños de las fracciones que a los de los radios.

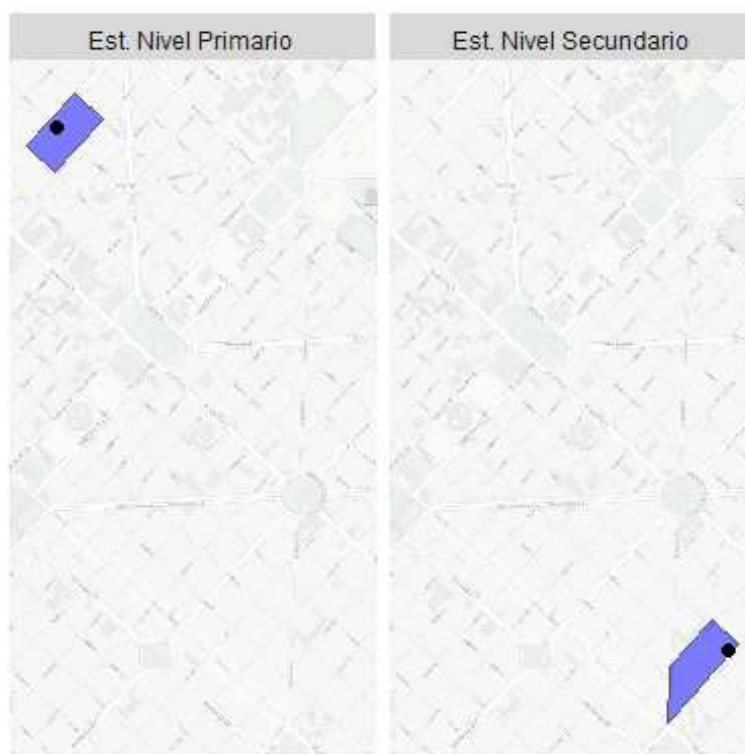
Sin embargo, es posible usar la propia información de los radios censales para mitigar los problemas anteriores. En efecto, es posible utilizar una combinación del tamaño de la matrícula con la densidad poblacional de cada uno los radios que conforman la fracción (Born y Scasso 2019, pag. 58). Si se prefiere un indicador contextual que no incluya información de la matrícula (por considerarse algo volátil en el tiempo y algo ajena al control de la planificación estatal) puede usarse como alternativa la información de los radios que se intersecan con el respectivo buffer normativo (v) y ponderar a cada uno de ellos en función del tamaño del área intersecada y la distancia del centroide de cada radio al establecimiento. Nuevamente, el supuesto que se encuentra por detrás de esta medida es que los radios más cercanos influyen (o son influidos) más que los radios más lejanos aún cuando todos formen parte de una misma fracción.

Como puede observarse en Figura 2 y en Figura 1 los radios y las fracciones censales tienen la característica de poseer una superficie variable. En general, cuanto más densa poblacionalmente es determinada área más pequeñas serán sus fracciones y radios. En este caso, como el establecimiento de la izquierda se encuentra en un área más densa que el establecimiento de la derecha tanto su radio (Figura 2) como su fracción (Figura 1) poseen una superficie menor al establecimiento de la derecha.

Figura 1: Fracciones censales



Figura 2: Radios censales



El ejemplo anterior suele usarse como un intento por imputar características sociales contextuales (a través de datos censales) a todo el establecimiento. Este punto es importante. Más que entender esta característica como un defecto puede entenderse como una característica buscada intencionalmente, en el sentido que el desiderata del indicador es la construcción de un indicador del establecimiento que relacione su ubicación con el contexto social del territorio y no, por ejemplo, con el contexto social de la matrícula (Born y Scasso 2019, pag. 6).

El uso de las divisiones políticas puede ser interesante para visualizar criterios presupuestarios que se deban realizar desde el nivel provincial para algunos de los niveles inferiores, especialmente a nivel de los departamentos ya que estos representan a la institución política de los municipios. Sin embargo, esta mínimo denominador común todavía parece extenso para otros objetivos. En especial, no parecen pertinentes para delimitar algo que intuitivamente se puede denominar “la zona de influencia de un establecimiento” sea tanto en su interpretación como efecto del territorio sobre el establecimiento o a la inversa. Esto se debe a la propia escala geográfica de las regiones, los distritos o los partidos es muy

grande y suele haber varios establecimientos por partido.⁶ Lo anterior impide la realización de discriminaciones finas a nivel local ya que, especialmente en zonas urbanas, la gran mayoría de los establecimientos tendría otros establecimientos en esa misma unidad política/administrativa.

Esto último problema se reduciría en el caso del uso de divisiones administrativas censales como las fracciones o los radios dado que razonablemente se puede esperar que la “zona de influencia” de un establecimiento se expanda a través de varios radios censales dado que usualmente existen más radios censales que establecimientos educativos (Figura 2).

5.1.2 Datos políticos administrativos y datos matrícula nominal

En la sección anterior se mostraron algunas potencialidades y limitaciones de la combinación de datos de los establecimientos con datos espaciales (en especial censales) del territorio. En esta sección el objetivo será la vinculación de los datos de los establecimientos con datos espaciales de la propia matrícula. Esto es algo relativamente nuevo ya que depende de la posibilidad de la georeferenciación de los propios estudiantes algo que, aparte de ser información no disponible en forma pública, es algo relativamente difícil de realizar.

El caso particular de la producción de los datos espaciales educativos en combinación con datos censales puede ser útiles en cuestiones de construcción de variables contextuales de los establecimientos en función de las características sociales de la matrícula (Born y Scasso 2019, pag. 6). Específicamente pueden tener una importancia crucial como una segunda mejor respuesta en cuestiones de evaluación de la calidad educativa.⁷

Como veremos a continuación, muchas de las estrategias usuales cuando se trabaja con este tipo de técnica es el empleo de una colorpleta en donde el color mediante el cual se rellenan los polígonos se encuentra en función de algún valor de una variable considerada pertinente.

En la Figura 3 puede observarse una colorpleta que indica, para ambos establecimientos, el porcentaje de estudiantes que proviene de cada departamento. En este caso en particular no se aprecian diferencias sustantivas entre ambos establecimientos. Como enseguida se podrá apreciar esto se debe a un punto ciego de la propia técnica más que a la extrema similitud de

⁶ Este supuesto puede relativizarse en el nivel universitario ya que en algunas áreas sólo existe una sola institución por departamento o incluso por región, entendiéndose por esta como una región educativa conformada por varios departamentos.

⁷ La racionalidad de la cláusula de segunda mejor respuesta es que siempre, desde un punto de vista metodológico, debe preferirse (si existieran) datos propios del estudiante como cuestiones de su personalidad, sus relaciones sociales, su hogar y su vivienda. En este sentido, en situaciones en donde no exista este tipo de datos, se puede imputar a cada estudiante (y de manera derivada a cada establecimiento) algunas características consideradas pertinentes de su radio censal.

los datos. En efecto, en Figura 4 y en Figura 5 se observa que los datos de ambos establecimientos efectivamente poseen distribuciones diferentes.

Figura 3: Porcentaje de estudiantes según departamento



Figura 4: Porcentaje de estudiantes según fracción censal

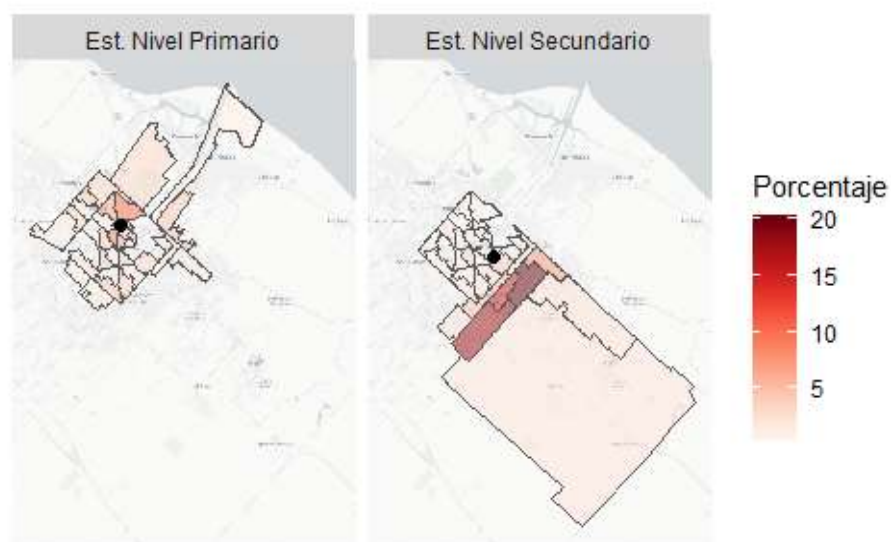
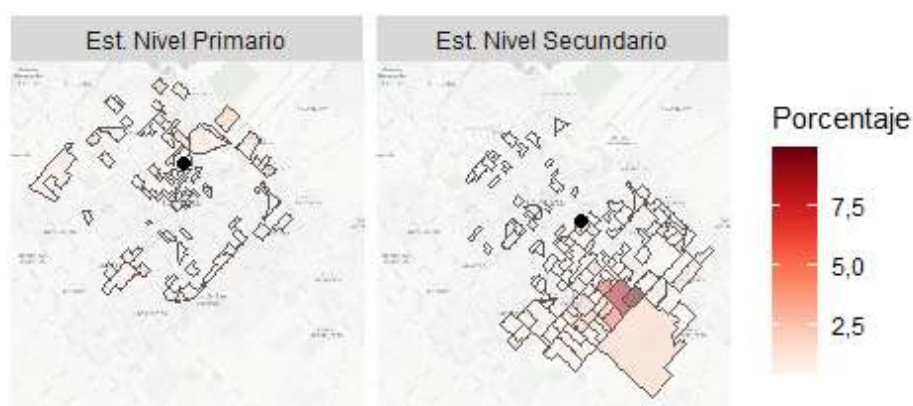


Figura 5: Porcentaje de estudiantes según radio censal



Sin embargo, si bien los datos a nivel de fracción y radio son más desagregados y permiten visibilizar diferencias entre ambos establecimientos, la irregularidad en las formas y en la superficie de los polígonos hace difícil una lectura fluida. No sólo eso. La irregularidad de los polígonos sobre los cuales se muestran los datos hace que, aún para un mismo establecimiento, se tratara de poblaciones de estudiantes diferentes cuando se los observa a la luz de los departamentos, las fracciones y los radios. Se recuerda que esto no es así, ya que los datos que se encuentran por detrás de los gráficos de cada establecimiento son siempre los datos de las y los estudiantes de cada establecimiento. Más allá de las deficiencias de su visualización, los datos de las y los estudiantes se podrían utilizar para la construcción de indicadores espaciales analíticos cuando se tenga el objetivo de construir variables contextuales de los establecimientos con datos de la matrícula.

Antes de pasar a la siguiente sección es pertinente el siguiente comentario. Un punto metodológico importante es que el área de cada polígono es diferente y esto puede llevar tanto a errores metodológicos como a interpretaciones pocos robustas ya que estas son excesivamente dependientes de los cortes espaciales utilizados. Para evitar errores metodológicos una recomendación usual es llenar los colores con valores normalizados más que con valores absolutos (por ejemplo, dividiendo por el área de cada polígono). El problema de la dependencia de las interpretaciones en función de los cortes espaciales utilizados es

parte de un problema compartido con otras técnicas de análisis, y se lo suele denominar problema de la unidad de área modificable (MAUP).⁸

Un punto a favor de estas estrategias es que, al tratarse de información agregada, casi siempre preservan la confidencialidad y por lo tanto es compatible con una política de datos abiertos. En especial la vinculación de datos políticos y censales a nivel de la mínima unidad geográfica en común (usualmente a nivel de departamento) puede ser una buena candidata a publicaciones seriales abiertas.

5.2 BUFFERS

Los buffers⁹ son polígonos que representan el área entre una distancia dada (1Km, 5Km, etc.) y una figura geométrica (punto, línea, polígono, etc.). Independientemente de cuál sea el insumo (un punto para un colegio, una línea para un camino o un polígono para un departamento) el resultado siempre es un **polígono**. Dado que, para fines de este trabajo, los establecimientos se representan mediante un punto, el polígono resultante siempre es un **círculo** con un radio establecido (1Km, 5Km, etc.). Si los establecimientos incluyeran las dimensiones del predio entonces se usaría un polígono para representar a esos predios y el respectivo buffer sería tan irregular como son las formas del respectivo predio.

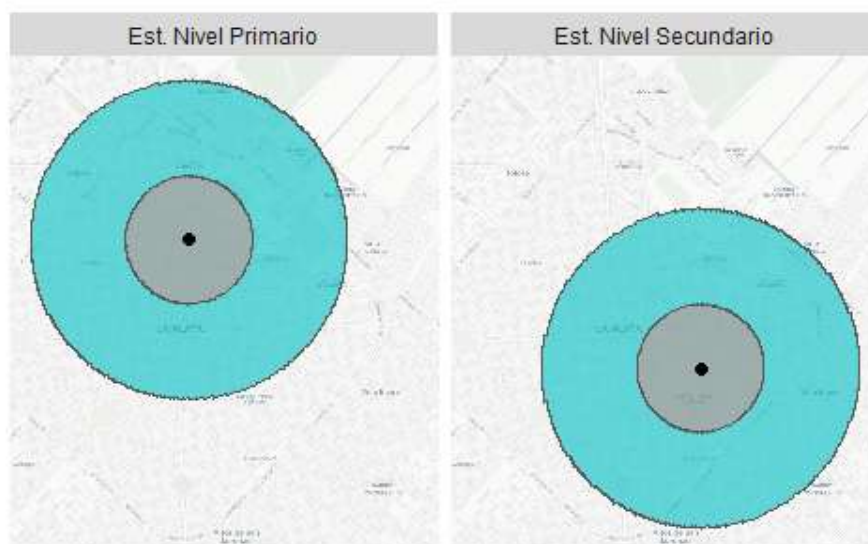
Es posible la construcción de un buffer específico para cada establecimiento en función de ciertos criterios. La dimensión del buffer puede estar determinada por criterios analíticos (ej. distancias promedio, tamaño de la matrícula, etc.) o criterios legales/normativos para cada nivel educativo (definidos por ejemplo por el Reglamento Escolar). En efecto, una de sus mayores virtudes es la fácil visualización espacial de la extensión de los criterios normativos. Esto puede observarse en [Figura 6](#) en donde se estipulan dos buffers para cada establecimiento. Uno de 1 km y otro de 2,5 km de un radio.¹⁰

⁸ Más adelante veremos como la propuesta de los análisis basados en densidades de puntos (Sección XX) mitigan estos problemas aunque con un mayor costo de producción. Sobre la problemática específica MAUP puede consultarse (Arbia, 1989).

⁹ Dada su difusión dentro del léxico espacial se ha preferido la introducción del anglicismo *buffer*. Las traducciones disponibles en español (amortiguación, regulación, etc.) usualmente parecen tener significados diferentes al específico del mundo del análisis espacial.

¹⁰ Estrictamente la reglamentación sobre el nivel secundario habla sobre tiempo de viaje y no sobre distancia. Como una aproximación, y asumiendo una conversión de tiempo en distancia, la misma se visualiza a través de un buffer. Si se quiere aumentar la complejidad se pueden calcular isolíneas de tiempo de viaje para cada medio de transporte para cada establecimiento. Esta es una técnica derivada de las distancias de ruteo y en ese caso cada establecimiento tendría un polígono variable envolvente con esos límites, pero no un buffer.

Figura 6: Buffer normativo



Ante la ausencia de información nominal georreferenciada, los buffers pueden ser utilizados para inferir, en base a estimaciones teóricas, variables contextuales acerca de las características sociales de la matrícula de los establecimientos. Por ejemplo, se podría estimar la influencia del territorio sobre el establecimiento en función del tamaño del establecimiento y/o la densidad poblacional de su entorno y luego, en función de los radios (o fracciones) censales que se encuentran en ese buffer, inferir características contextuales sociales del establecimiento presente en los propios datos censales (Born y Scasso 2019, pag. 58).

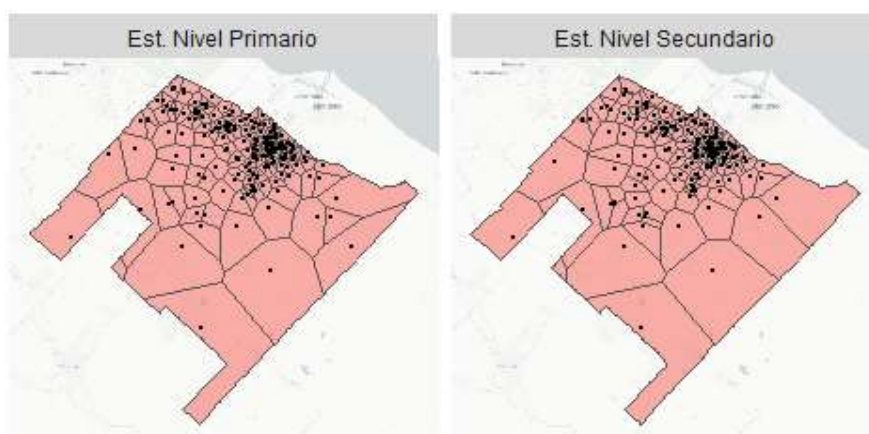
Si se posee datos georreferenciados de cada estudiante, los buffers también pueden servir para averiguar el grado de cumplimiento de la normativa. En efecto, se pueden realizar intersecciones espaciales para averiguar, por ejemplo, cuantos estudiantes (y de qué lugar) caen dentro/fuera del círculo conformado por el buffer.

Por lo anterior, la utilidad de los buffers puede ser tanto a nivel central como local. Es especial, estos son útiles para la construcción espacial de criterios normativos y su posterior intersección con otros datos sociales georreferenciados. Por otro lado, en todos sus usos parece tratarse de una técnica compatible con políticas de datos abiertos dado que no parece comprometer la confidencialidad de los datos de los estudiantes.

5.3 POLÍGONOS DE VORONOI

Dado un conjunto de objetos (p.e. establecimientos educativos) en un plano (p.e una provincia) un diagrama de Voronoi es una partición (o teselación) de ese plano en tantos polígonos como objetos. Su característica distintiva es que todos los puntos interiores de cada polígono (posibles ubicaciones de estudiantes) comparten la propiedad de tener el mismo objeto (establecimiento educativo) más cercano. En esta partición ningún polígono se superpone con otro y tampoco quedan espacios del plano sin ser ocupado por algún polígono. El resultado visual de este tipo de particiones puede observarse en la Figura 7. En este caso, para que se aprecie su potencial uso a un nivel superior al del propio establecimiento se ha realizado esta partición para el partido de La Plata tanto para el nivel primario (izquierda) como para el secundario (derecha).

Figura 7: Diagramas de Voronoi según nivel educativo



Como puede observarse en Figura 7 dentro de cada polígono hay un punto que representa a un establecimiento educativo y cada punto se encuentra bordeado por un polígono diferente. También puede verse como, al sur de ambas imágenes, dada la mayor lejanía entre los establecimientos educativos, los polígonos poseen un área mayor. Dado que la oferta educativa de nivel primario es mayor a la oferta de nivel secundario los polígonos de esa zona sur también son más extensos en el nivel secundario en comparación con el nivel primario.

El hecho que muchos establecimientos secundarios se encuentran en un predio contiguo al de un establecimiento de nivel primario facilita la comparación visual entre ambos niveles.

Por ejemplo, el punto más “al oeste” de la figura de la izquierda no tiene su contraparte en la figura de la derecha. Esto permite intuir, si existiera un establecimiento secundario allí en la figura derecha, como cambiaría los diagramas de Voronoi y, de este modo indirecto, intuir el modo de construcción de los mismos. Un punto a favor de esta técnica es que sólo presupone información georreferenciada de los establecimientos y no de los estudiantes. Otro punto a favor de esta técnica es que su resultado final genera un mapa muy intuitivo.¹¹

Como muchas de las otras técnicas de este documento permite que se puedan agregar algunas restricciones, como características naturales (ríos, picos montañosos, etc.) o convenciones (límites administrativos, rutas, etc.), y que estas sirvan como divisiones explícitas dentro de los polígonos y que estas no queden dentro de algún polígono.¹²

Los posibles usos de esta técnica son varios. Algunos son útiles a nivel central y otros a nivel local. En especial la técnica parece idónea como insumo para la construcción de zonas de cobertura para cada establecimiento que ayuden a ordenar espacialmente la demanda efectiva a la oferta existente. A pesar de ser una técnica que está diseñada para ser utilizada como más de un establecimiento como insumo puede también tener una utilidad local a nivel de cada establecimiento. Como se verá en la sección XX, la creación del respectivo polígono de Voronoi de un establecimiento puede ayudar, de modo visible en un mapa, a entender cuantos y de dónde son los estudiantes que llegan a ese establecimiento cuando ese establecimiento no es su establecimiento más cercano o, expresado de manera alternativa, estimar cuantos y de dónde son aquellos estudiantes de su matrícula que tienen a ese establecimiento como su establecimiento más cercano.

Otro uso posible es la comparación de las áreas de los polígonos resultantes de la partición con algún indicador de densidad de la demanda (ver sección XX) y los tamaños de los establecimientos para corregir desbalances en la infraestructura existente.

Por último, al igual que sucedía con los usos de los buffers, algunos de los beneficios de este tipo de técnicas puede también conseguirse en base a técnicas analíticas. En efecto, si uno sabe la distancia de cada estudiante a cada establecimiento puede estimar o bien cuál es el establecimiento más cercano o bien ordenar a los 3 o 5 más cercanos, aunque ninguna de las operaciones anteriores se visualice en un mapa como sí lo hace un diagrama de Voronoi.

¹¹ En efecto, es una técnica que, con muchas variantes, se comenzó a usar antes del surgimiento y difusión de las modernas técnicas de análisis espacial. En Estados Unidos sus inicios se deben a los esfuerzos por implementar diseños compatibles con la legislación “Brown vs el Consejo de Educación” (1954) y, más en general, con la problemática de la segregación educativa (Goldring y Swain, 2014).

¹² En relación con esto la isla Martín García es parte del departamento de La Plata por lo que, estrictamente, ese departamento se representa con un multipolígono más que con un polígono. Gracias a la flexibilidad de esta técnica, se puede excluir a la isla entendiendo al Río de la Plata como una característica natural que oficia de límite natural.

5.4 POLÍGONOS ENVOLVENTES

En los usos de polígonos puede importar tanto la visualización y cómputo del perímetro como su área. En el primer caso lo que importa es el límite entre el exterior y el interior del polígono y en el segundo caso lo que importa es sólo su interior. En las técnicas que se analizarán en esta sección, especialmente en su construcción más que en su uso posterior, es clave la idea que aquí denominaremos de “envolvente”. En este sentido, estas técnicas comparten el espíritu de poder captar la envoltura (o la silueta o la sombra) más que la forma precisa, de una nube de puntos. De este modo, el resultado usualmente se trata de un polígono irregular que logra envolver a la nube de puntos formado por las ubicaciones de los estudiantes de cada establecimiento.

Una característica de esta técnica es que, por diseño, es influenciada principalmente por los valores o puntos más extremos de la nube de puntos (más adelante se especificará más sobre esta afirmación) y no distingue zonas de diferentes densidades de puntos en su interior (ver sección XXX). El grado de injerencia de los valores extremos se puede matizar, tanto para los polígonos convexos como para los cóncavos, filtrando previamente a los valores muy extremos sea a través de la estrategia de excluir del análisis a algún ntil de la distribución de distancia explícito. Otra estrategia, que en este caso tiene que ver con la diferencia específica de los polígonos cóncavos frente a los convexos, es la modificación de los criterios que tiene que cumplir cada punto para formar parte (o no) del perímetro del polígono.

Parece oportuno destacar que la idea anterior de puntos extremos (p.e. aquellos estudiantes más alejados en términos de alguna distancia) ilumina una característica de esta técnica pero también también oscurece esta otra: que serían los puntos extremos de, para ponerlos en términos simples, diferentes puntos cardinales. Justamente este último detalle es lo que marca la diferencia entre una técnica analítica y otra espacial ya que en la primera existe un valor de de una unidad de medida como el metro, el kilómetro, etc. y en la segunda se pueden calcular los primeros pero sin perder sus coordenadas espaciales. Puesto en un ejemplo concreto, supongamos que un establecimiento educativo tiene sólo 10 estudiantes, 7 al sur, 1 al norte, 1 al este y 1 al oeste lo que implica una nube de puntos más densa hacia la parte sur del establecimiento con unos pocos ejemplos en los otros tres puntos cardinales. Suponiendo que el establecimiento se encuentre justo en el cruce entre dos calles, los que viven en el sur se encuentran a 5 y 4 km de distancia mientras que los que viven en los otros puntos cardinales se encuentran a 1 Km de distancia con respecto al establecimiento. En esta situación, la técnica de los polígonos envolventes siempre incluirá a los casos del norte, este y oeste distorsionando en cierto sentido la nube de puntos original. La estrategia complementaria de hacer un ranking de distancias y luego, por ejemplo, de retirar de la confección del polígono envolvente a aquellos casos que se encuentran a más del 95 ntil sólo tendría efecto en el límite del “lado sur” del polígono ya que ahí se encuentran los casos más lejanos.

El ejemplo anterior sirve para mostrar que existe otra estrategia para mejorar la utilidad de los polígonos envolventes convexos (aunque también se pueda aplicar a los cóncavos) que consiste en calcular, más que a los outliers en términos de distancia al establecimiento, a los outliers en términos de distancia a los otros miembros de la matrícula. Si se prefiere se los podía denominar como outliers espaciales. Esta estrategia es más difícil de implementar por que implica:

- 1) Primero calcular la distancia entre todos los estudiantes de ese establecimiento,
- 2) Luego quedarse con el valor menor para cada estudiante (que sería a que distancia se encuentra el compañero de matrícula más cercano) y,
- 3) Finalmente hacer un ranking de todos los estudiantes y, recién en este paso, excluir del análisis a aquellos estudiantes que se encuentren más lejos de otros estudiantes.

Como veremos a continuación, en algunas situaciones la captación de la envoltura dice bastante sobre la forma de la nube de puntos pero otras veces esto no sucede. En general, la técnica basada en la construcción de polígonos convexos funciona mejor con distribuciones de puntos que se presuman “compactas” y la técnica de polígonos cóncavos, al ser más general, también funciona aceptablemente para captar distribuciones de puntos que presentan fuertes cavidades o agujeros.

En cualquier caso, en la sección XX veremos cómo estas técnicas se pueden complementar con otras para lograr el objetivo de transmitir distintos aspectos de la nube de puntos de los estudiantes. Cualquiera de esta familia de técnicas tiene a favor que, como todo polígono, preserva la confidencialidad del dato aunque, estrictamente, los (pocos) puntos extremos que se seleccionan para su construcción son puntos reales de algún estudiante que se visualizan en el mapa.

5.4.1 *Polígonos envolventes convexos*

La construcción de un polígono envolvente convexo implica la construcción de un polígono que envuelva a los puntos más extremos con la condición que su perímetro sea el más corto posible y que el mismo no posea ningún ángulo interior mayor a 180° . Esto suele producir, en comparación a un polígono cóncavo, una figura más simple (en el sentido que posee menos lados) aunque esto puede representar una figura bastante diferente a la distribución subyacente de la nube de puntos, especialmente en casos de distribuciones poco compactas. Expresado de manera alternativa, las formas de los polígonos convexos suelen ser algo toscas y estas formas no suelen ser buenas representaciones de nube puntos irregulares.

Figura 8: Polígonos convexos



Como puede observarse en la Figura 8 tanto en la figura de la izquierda como en la de derecha el polígono resultante se forma por unos pocos puntos extremos y se cumple la condición que ningún ángulo interior de los vértices del perímetro posee más de 180° . Para visualizar el efecto de la poda de los estudiantes más lejanos en los polígonos convexos de cada establecimiento, se han construido polígonos con el 100% de los estudiantes (polígono más grande) y otros que sólo incluyen al 95% de los estudiantes más cercanos y otro con el 90% más cercano (polígono más chico).

5.4.2 Polígonos envolventes cóncavos

Esta técnica es similar a la envolvente convexa pero ahora se permiten cavidades o agujeros o huecos que copien más fielmente la forma de la distribución subyacente de la silueta de la nube de puntos.¹³ La inclusión de estas cavidades es lo que justamente produce un polígono cóncavo. El costo es una figura, en comparación al respectivo polígono convexo, más irregular

¹³ Para las personas que no vienen de las matemáticas quizá la siguiente regla memotécnica sirva para distinguir la diferencia entre los diferentes tipos de polígonos envolventes. Como lo sugiere su etimología, la palabra “cóncavo”, de origen de latina, es una unión de “con” y “cavus” en donde esta última quiere decir “agujero”, “hueco” o, de manera más transparente para el español “cavidad”, por lo que el significado de “cóncavo” se puede asociar a algo “con-cavidades”. En cambio, la etimología de “convexo” no parece ser tan iluminadora sobre su significado actual, por lo que alcanza con saber que serían aquellos que no tienen cavidades.

(con más lados), con un área menor y de mayor perímetro que la anterior. Otro costo de esta representación más fiel de la nube de puntos, es que, al aumentar su irregularidad, esperablemente se torna más difícil de clasificar su forma final. Siguiendo el famoso proverbio inglés “no puedes tener la torta y también comerla a la vez”¹⁴ es claro que cada investigador debe tener en claro que prefiere en cada situación teniendo como extremos la propia nube de puntos con toda su riqueza (pero muchas prohibitiva por cuestiones de confidencialidad) y las diferentes opciones de síntesis (como los polígonos envolventes) más o menos simples de interpretar pero menos realista en cuanto a su representación de la realidad.

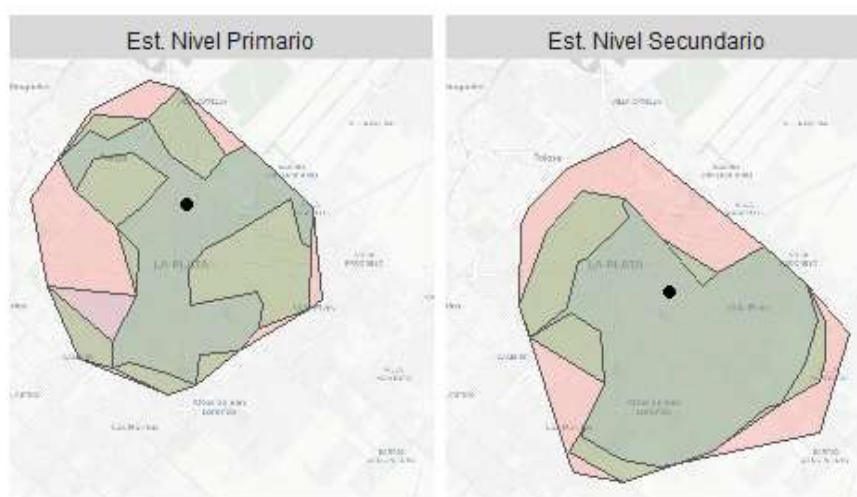
El cálculo de los polígonos cóncavos contiene un parametro *alpha* que permite decidir hasta donde la envolvente crea cavidades que van a buscar, aparte de los puntos que incluye el polígono convexo, a otros puntos cercanos a los extremos. Este parámetro, que no existe en el cálculo del polígono convexo, y que se puede configurar entre un rango de $-\infty$ y $+\infty$, es lo que permite conceptualizar a los polígonos envolventes convexos como un caso especial de los polígonos cóncavos, a saber, aquellos polígonos cóncavos en donde su parámetro *alpha* es igual a 0.¹⁵

Aquí hemos asumido que estas técnicas nos importan para captar la envolvente de la nube de puntos. En sentido estricto, y dependiendo de la forma de los puntos y de la configuración de los parámetros, estas técnicas pueden también producir polígonos con “huecos” como el de un anillo o, de forma más extrema, un multipolígono que contenga varias “islas”. En estos últimos casos se habla de “forma α ” (*α shape*) que, a su turno, generalizan la noción de “envolvente α ” (*α hull*).

¹⁴ El proverbio original en cuestión es “you can’t have your cake and eat it”.

¹⁵ El parámetro *alpha* tiene más difusión en la literatura de la geometría computacional. La bibliografía canónica al respecto es (Edelsbrunner 2011) (Edelsbrunner, Kirkpatrick, y Seidel 1983) (Edelsbrunner et al. 1995). En campos más cercanos a la programación y la optimización de algoritmos existe una mayor variedad de términos y posibles configuraciones. En efecto, existen diferentes algoritmos para diferentes lenguajes de programación. Algunos pueden utilizar otros términos (como el parámetro “convexity”) y otros valores de referencia en vez del rango de $-\infty$ $+\infty$. Algunos algoritmos también presentan la opción de un parámetro que permita una configuración de un umbral de distancia a partir del cual se permite que entre otro punto en la construcción del perímetro del polígono cóncavo.

Figura 9: Polígonos convexos



Como puede observarse en la Figura 9 se han construido tres polígonos cóncavos por cada establecimiento. En la Figura 8 las diferencias entre aquellos polígonos se debían a que se incluía diferentes poblaciones como insumo en su cálculo (100%, 95% y 90%). En cambio en la Figura 9 siempre se ha trabajado con el 100% de la población y las diferencias en los tres polígonos de cada establecimiento difieren en los parámetros en que se han configurado su concavidad. Por esta razón, si se compara el polígono más grande de cada establecimiento de la Figura 8 con los polígonos más grande de cada establecimiento de la Figura 9 se vera que estos últimos son levementes menores en su superficie y esta diferencia aumenta a medida que se aumenta el grado de concavidad o el parámetro *alpha*. Esto es así porque a mayor concavidad los polígonos incluyen varios puntos más en su construcción, poseen una superficie menor y presentan una forma que, si bien se acerca más a la silueta de la nube de puntos subyacente, es difícil de describir dada su pronunciada irregularidad. En general, cuanto menor sea el umbral de la distancia entre los puntos y cuanto mayor sea el grado de concavidad permitida más irregular será el polígono resultante.¹⁶

¹⁶ Existen más de un algoritmo disponible para este fin en diferentes lenguajes. En un primer momento (2020) utilizó el algoritmo concaveman (existente en varios lenguajes) que tiene sus bases en (Park y Oh, 2012). En el léxico de la propia librería se puede configurar un parámetro del grado de concavidad (a menor valor mayor concavidad) y un umbral mínimo de distancia entre los puntos. Posteriormente (2024) la importante librería GEOS (también existente en diferentes lenguajes) incorporó el algoritmo Graham's scan (Graham, 1972) para el cálculo de este tipo de polígonos. Para fijar las ideas, las imágenes de este documento, que fueron realizadas

5.5 CENTROIDES

Cómo se anticipó anteriormente, las técnicas de polígonos envolventes (sea en su versión cóncava o en su versión convexa) son intentos por captar la forma de la nube de puntos de los estudiantes en donde las versiones convexas son más simples pero menos realistas y lo inverso sucede con las diferentes versiones de los polígonos cóncavos. En otras palabras, son intentos por prestar atención a los bordes o extremos de la nube de puntos. Por el contrario, la técnica de los centroides permite concentrarse en su tendencia media. Esto último, más que una debilidad, es una oportunidad para ofrecer, gracias a su complementación, una mejor visualización de la complejidad de los datos.

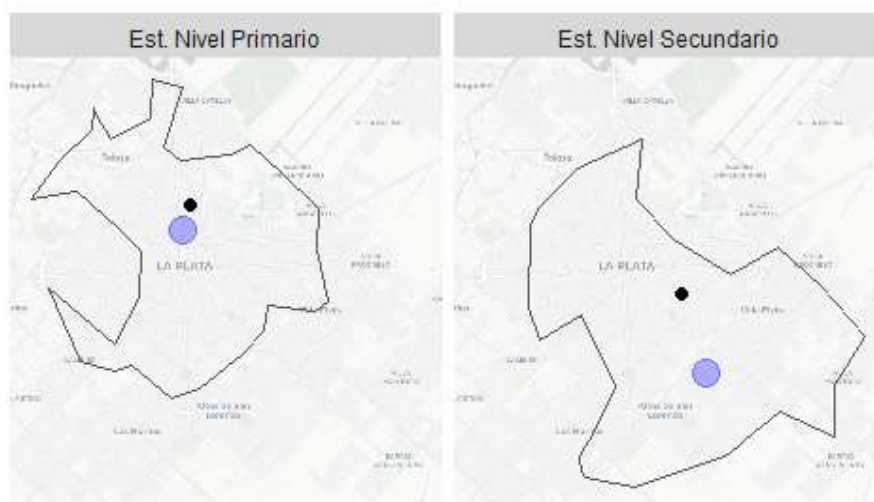
El centroide puede considerarse como el punto que minimiza la suma de las distancias (euclidianas) cuadradas entre ese punto y el resto de los puntos que conforman la figura. En este trabajo principalmente nos interesará el centroide formado por la nube de puntos (en vez de, por ejemplo, algún polígono envolvente) de las ubicaciones de los estudiantes. Esta técnica, como la media aritmética de un conjunto de números, también se ve afectada por los valores extremos, pero a diferencia de los polígonos utilizados para la construcción de las envolventes (Ver sección XX) **todos** los valores intermedios también cuentan por lo que, de alguna manera, existe un mayor contrapeso del resto de los valores.¹⁷

En este sentido, como lo sugiere la Figura 10 el centroide de una nube de puntos puede considerarse como una aceptable estimador de la tendencia central en el sentido que en la posición final del centroide intervienen como insumo todos los puntos y no sólo aquellos con posiciones extremas.

con la librería `sf` del programa R están configuradas con un valor en su concavidad de 0.8, 0.6 y 0.4 respectivamente, siendo 1 un polígono convexo y 0 un polígono con una máxima concavidad.

¹⁷ Si se calcula el centroide de un polígono envolvente (sea este cóncavo o convexo) el cálculo produce un punto “como si” todos los estudiantes se distribuyeran de forma pareja a lo largo del área del polígono. Dado que los polígonos envolventes, por su propia construcción, (sobre)estiman los casos con valores extremos de la nube de puntos, el cálculo del centroide sobre estas figuras produciría una doble (sobre)estimación de los valores extremos. Por esta razón, en presencia de los datos individuales de cada estudiante, se prefiere la construcción de los centroides directamente sobre la nube de puntos de los propios estudiantes.

Figura 10: Mapa centroide



En la Figura 10 para una mejor contextualización se ha dejado la capa del polígono cóncavo (concavidad = 0,4) junto con el recientemente calculado centroide. Como puede observarse, en en la figura de la izquierda el centroide se encuentra más cerca que el respectivo establecimiento educativo que en la figura de la derecha.

La utilidad en solitario de esta técnica remite a que produce una sola posición como resumen de una nube de puntos. Dado que de modo explícito o implícito cada miembro de cada comunidad educativa suele tener algún supuesto sobre este punto, la intuitiva visualización facilita el proceso de comprobación sobre la adecuación de los mismos. En efecto, la distancia (como medida analítica) del centroide al establecimiento puede ser diferente a la media de la distancia (euclídeana o de ruteo) de los estudiantes. Esto último, que en principio puede resultar contraintuitivo, es otras de las diferencias entre las técnicas analíticas y las espaciales.

La diferencia con la media de las distancias se vuelve más transparente cuando se observan los insumos de cada técnica. En en caso del centroide no sólo se le puede calcular su distancia con respecto a el establecimiento sino que cada insumo del mismo también posee una posición (coordenadas) en el plano y, por lo tanto, se puede introducir en el cálculo y visualizar en un mapa. Para fijar las ideas, si un establecimiento posee sólo 2 estudiantes, ambos a 1 km de distancia, pero uno en una dirección y otro en la contraria (supongamos norte-sur) la media de las distancias será 1 km y, sin embargo, el centroide coincidirá con el punto del propio establecimiento, por lo que su distancia al establecimiento será de 0 km. En este

sentido, su uso a nivel local de cada establecimiento puede ser pertinente para la “toma de conciencia” de aspectos que, por su propia naturaleza, son poco intuitivos.

Por último, a pesar de tratarse de un punto y de calcularse sobre la nube de puntos de los estudiantes, es una medida que, al igual que el resto de las técnicas aquí presentadas, preserva la confidencialidad del dato por tratarse de una agregación. Por estas razones, parece compatible con una política de datos abiertos. Como muchas de las técnicas de este trabajo, su utilidad aumenta si se lo utiliza como complemento de algún otro indicador que capte la heterogeneidad de los puntos. Esto último es lo que se hará en la próxima sección.

5.6 ANÁLISIS DE DENSIDADES

Si los polígonos envolventes ayudan a captar la silueta de la nube de puntos y el centroide su tendencia central espacial, la técnica que veremos en esta sección puede decirse que ayuda a cuantificar el grado de heterogeneidad espacial de esos puntos. De manera intuitiva la técnica de cálculo que se encuentra *por detrás*, sirve para discriminar zonas de la nube de puntos con diferentes densidades de puntos y la técnica de visualización que se encuentra *por delante* remite a la noción de isobanda (e isolínea) usual en la cartografía. Por lo anterior, en esta sección se intentará explicar tanto la técnica de cálculo como la técnica de visualización.

Esta técnica de cálculo puede considerar algo más compleja que las anteriores. Sin embargo, o gracias precisamente a ello, también permite cumplir muchos de los desideratas presentes en la construcción intuitiva de una zona de influencia. Si bien se la puede complementar con varias de las técnicas anteriores (y eso se hará en la sección XX) posiblemente sea la técnica que mejor se desempeñe de forma aislada para los fines buscados en este documento.

En este último sentido, cuando se analizó en la [Sección 5.1.2](#) se había anticipado que las técnicas basadas en análisis de densidades lograban mitigar el problema de la unidad de área modificable (MAUP). La razón es que se construye una función de densidad **continua** que intenta estimar la intensidad de los puntos en función de la distribución de los puntos **discretos** observados.¹⁸

Expresado de otro modo, en vez de calcular la densidad en polígonos predeterminadas (p.e. aquellos con orígenes políticos o administrativos), ahora estas áreas primero se calculan en función de las diferencia o similitud de la densidad de los propios puntos observados y,

¹⁸ En efecto, es una técnica usual en el campo académico denominado “Análisis espacial de patrones de puntos”. Fuentes de consulta de esta disciplina pueden considerarse (Baddeley, Rubak, y Turner 2016) [Gelfand et al. (2010)](Møller y Waagepetersen 2004). Para una exposición más general acerca de las funciones de densidades puede consultarse (Silverman 1998) (Chacón y Duong 2018). En este caso particular hemos utilizado la técnica de estimación de densidades basados en kernels gaussianos en vez de cuadrantes (Gimond 2020, cap. 11).

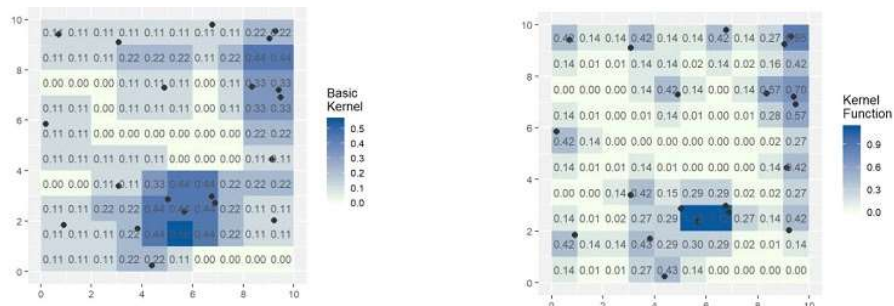
posteriormente se visualizan como isobandas cartográficas en un plano. Así, la técnica puede verse tanto como una técnica de estimación (de una variable latente continua) como una técnica de visualización que sirve tanto para suavizar los patrones encontrados en los valores observados como para visualizarlos de una manera sintética. Más adelante veremos que aquí importan las dos características.

Dentro de la función para realizar la estimación del cálculo de la densidad existen dos parámetros principales: el **ancho de banda** (*bandwidth*) y el **tipo de núcleo** (*kernel*). El primero define hasta que distancia se va a considerar a otro punto como “vecino” y el segundo determina la manera en que se “integran” a los vecinos encontrados. En este último sentido, se obtendrá un resultado si todos los vecinos encontrados pesan lo mismo (uniforme) o si los vecinos más cercanos influyen más que los vecinos más lejanos.

Por otro lado, cuanto más pequeña sea la distancia para ser considerado como “vecino”, más irregular (o menos suavizadas) serán las **isobandas** resultantes. En otras palabras, cuanto menor sea la distancia para considerar a otros puntos como un vecino la técnica más replicará la irregular distribución observada de puntos porque el resumen o síntesis ofrecida será de pocos puntos. En el otro extremo, cuanto más grande sea la distancia para considerar a otro punto como vecino, más suavizadas serán las isobandas construidas porque justamente en esa síntesis entran más puntos. Al igual que sucedía con los polígonos envolventes cada investigador deberá encontrar su respectiva configuración óptima entre ambos extremos.

La combinación de determinado valor del ancho de banda y el tipo de núcleo determinará, en conjunción con la propia distribución observada de puntos, los valores de la función de densidad. Este proceso puede observarse en la Figura 11 en donde, en base a un plano de 10 x 10 celdas se crean unas 100 posiciones para las cuales se calcula, para cada una, un valor específico de densidad.

Figura 11: Imágenes extraídas de Gimond Manuel (2020) CC 4.0.



En esta figura tenemos la visualización de los resultados de dos funciones de densidad que tienen:

a) igual insumo, esto es, una misma distribución de puntos a analizar y,

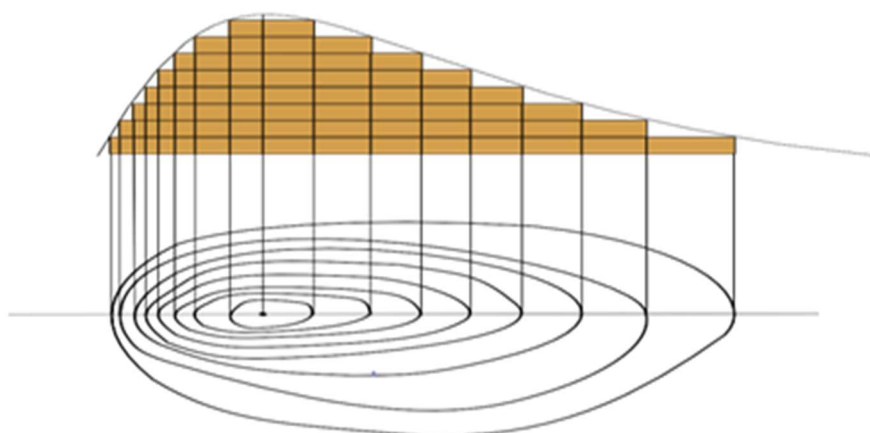
- b) igual configuración del *bandwidth* o ancho de banda y,
- c) diferente configuración del *kernel* o núcleo.

Para estos ejemplos podemos suponer que el ancho de banda siempre es 1 por lo que para calcular la densidad de puntos de cada celda se integran los valores de los puntos adyacentes de la celda vecina. En efecto, lo que cambia entre la imagen de la izquierda y la de la derecha es el modo en que se integran los puntos considerados como vecinos detectados por el ancho de banda. En la imagen de la izquierda todos los puntos detectados como vecinos suman igual ya que se utiliza un kernel uniforme o básico. En la imagen de la derecha los puntos pesan de forma inversamente proporcional a la distancia ya que se utiliza un kernel gaussiano. Si uno se imagina los valores de las densidades como una carta cartográfica se observa que en las zonas con los colores más altos de la imagen de la izquierda (que serían los colores más oscuros de la imagen) posee subidas más suavizadas con mesetas y sierras, mientras que esas mismas zonas en la figura de la derecha se observan montañas no sólo más altas sino con una mayor pendiente.

Un punto de la idea de “vecino” es que, si bien se trata de un parámetro libre desde el punto de vista estadístico, este se puede llenar con contenido sociológico y/o normativo. En este sentido, en el contexto de la planificación educativa por ejemplo, se podría pensar que un vecino en nivel inicial sea alguien a menos de 500 metros o menos de 1000 metros y que en nivel secundario sea alguien a menos de 1000 metros o 2500 metros.

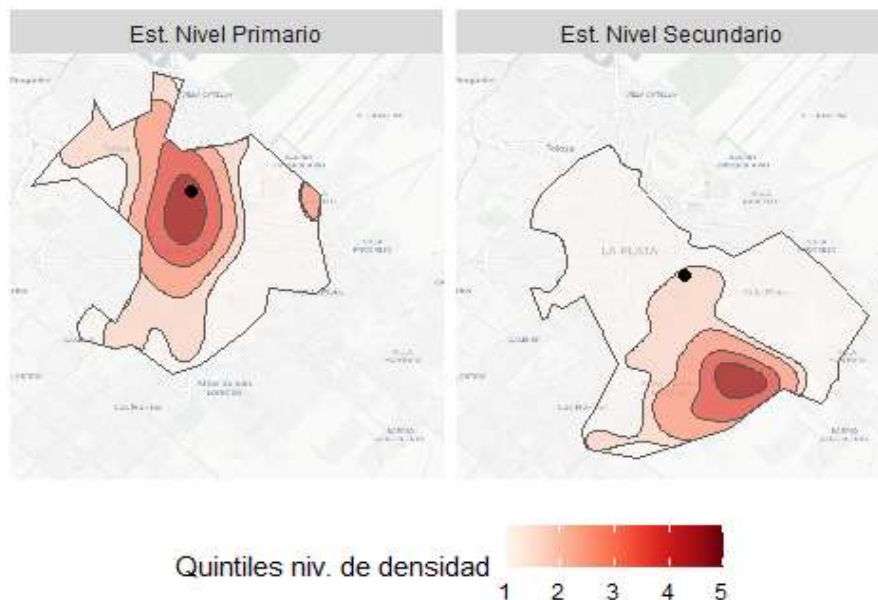
Unos párrafos atrás, al pasar, se introdujo de contrabando la noción de **isobanda** y luego se usó la metáfora de una carta cartográfica, por lo que aquí aprovecharemos para decir algunas palabras al respecto. La idea de isobanda y su aplicación a la construcción de las cartas cartográficas permite traducir en polígonos vectoriales los diferentes colores de las celdas de la Figura 11 o, más precisamente, los diferentes valores de la función de densidad. Para dilucidar esta cuestión introduciremos el concepto hermano de **isolínea**. Este es un recurso muy usado en cartografía para visualizar “altitudes” o “elevaciones” (3D) en un plano (2D). En principio, no importa si se trata de elevaciones reales como en los mapas cartográficos de relieve o de elevaciones analíticas productos de histogramas, cantidad de tiempo o funciones de densidad como en el caso de la Figura 11. El proceso, como lo sugiere la Figura 12, consiste en que cada isolínea en un plano representa zonas de igual “altitud”, permitiendo de este modo construir, de modo derivado, polígonos de contorno (isobandas) con los intervalos entre las isolíneas que pueden interpretarse como áreas de similar (no iguales) “altitud”. Cuantas más isolíneas se calculen más intervalos se crearán y, por lo tanto, más isobandas (y con una menor heterogeneidad interna) se podrán construir.

Figura 12: Imagen extraída de Wikipedia (artículo “contour line”) CC 2.5



Ahora pasaremos a aplicar la técnica a los mismos establecimientos de los ejemplos anteriores. Eso precisamente es lo que se realiza en la Figura 13.

Figura 13: Mapas densidades



Como puede observarse aquí se están utilizando como marco para aplicar el análisis de densidades los polígonos envolventes cóncavos que se han descrito anteriormente ([Sección 5.4.2](#)) aunque de manera similar se podría haber usado el polígono envolvente convexo. Por otro lado, ahora aparecen una serie de polígonos en su interior los cuales, cada uno, representa zonas de similares valores de la función de densidad. Esto implica que los

polígonos con los colores más oscuros representan zonas en donde se encuentran muchos puntos (estudiantes) cercanos y las zonas más claras representan zonas en donde estos están más espaciados. En la Figura 13 los valores de los niveles de densidad están visualizados como quintiles por lo que el 20% de las celdas con mayores niveles de densidad forman un misma isobanda (polígono o multipolígono) y así sucesivamente con el resto de la distribución. Esta manera de visualizar los datos no tiene una escala absoluta sino que se trata de un valor relativo a cada distribución de puntos de cada establecimiento y para fines prácticos se consideró aceptable una partición de 5 isobandas, aunque, nada impide utilizar más isobandas.

En comparación con las técnicas que provienen de la idea de los polígonos envolventes ([Sección 5.4](#)), los análisis de densidades rescatan información de todo el espectro de puntos y no justamente de aquellos que tienen alguna relación con los “bordes”. Por otro lado, si bien se acercan a la distribución real de la nube de puntos se alejan en 2 direcciones interesantes. Primero, al suavizar la tendencia (y de este modo distorsionar los datos de los puntos) la técnica es compatible con un formato de datos abiertos. Segundo, la producción de isobandas produce síntesis intuitivas aun cuando estas se construyan sobre una distribución extremadamente irregular de puntos.

La idea de densidad es potente y puede servir para otras cuestiones de análisis espacial y educación. Aquí se lo presentó en el contexto de la producción de mapas para visualizar la distribución de la población de estudiantes de cada establecimiento pero su utilidad no termina allí.¹⁹

Por último, es una técnica que realiza los cálculos sobre datos desagregados, lo que si bien es una limitación dado la mayor difusión de datos agregados, también es una oportunidad para aprovechar de forma flexible la información a nivel de cada punto de los estudiantes sin caer en agregaciones o síntesis groseras al tiempo que preserva la confidencialidad de los datos.

6. RECOMENDACIONES O DISCUSIÓN

Antes de pasar a una aplicación práctica ([Sección 6.1](#)) que muestre una combinación específica de las técnicas descriptas anteriormente es pertinente una breve recapitulación que condense las principales características de cada una de ellas. En la tabla Tabla 2 que se

¹⁹ Por ejemplo, se puede utilizar como parte de un estimador geográfico desagregado de demanda potencial de estudiantes. En efecto, la técnica es aplicable teniendo como covariables datos demográficos y tasas de deserción. La técnica también es compatible con la exclusión de zonas que por alguna información secundaria se considere que allí la estimación no es correcta sea tanto por accidentes geográficos (lagunas, montañas, etc.) como sociales (basurales, límites provinciales, etc.).

presenta más abajo se condensan las principales características de cada una de las técnicas vistas y se destacan algunos posibles usos de la misma.

Tabla 2: Características principales de las técnicas vistas

Técnica	Nivel	Posibles usos	Observaciones
Analíticas (Distancias lineales y de ruteo)	Todos	Gráficos exploratorios de distancias Control de calidad de otros datos espaciales Construcción de tablas resumen a diferentes niveles y con diferentes cruces de variables	Una vez producida, su utilidad no depende de un sistema GIS Información compatible con el conocimiento estadístico disponible (no depende ni de la geometría ni de la teoría de redes)
Límites Políticos Administrativos	Provincial Regional Partido	Visualizar información educativa en polígonos en donde o bien a) exista capacidad de decisión y/o b) exista información social como la censal	Existencia previa de los polígonos utilizados. No se trata de polígonos que surjan de información del propio sistema educativo. Problemas metodológicos por las diferentes áreas de cada polígono.
Límites Políticos Administrativos	Local	Primera opción para imputar características contextuales del territorio a cada establecimiento (fracción o radios) Segunda mejor opción para imputar características contextuales de la matrícula (radios)	No es necesario su visualización en mapas. Su finalidad principal es generar atributos “sociales” a cada establecimiento y, de manera derivada, a su matrícula.
Buffer	Local	Construcción de zonas normativas como las líneas de contorno de un polígono circular.	Alcanza con información georreferenciada de los establecimientos.
Buffer	Centro	Construcción de un polígono (o un multipolígono) que represente todas las zonas en donde se cumplen los criterios normativos.	Alcanza con información georreferenciada de los establecimientos.
Voronoi	Todos	Construcción de un polígono que indica los límites de hasta que punto del plano un establecimiento es el más cercano.	Alcanza con información georreferenciada de los establecimientos
Polígonos envolventes convexos	Local	Construcción de un polígono convexo que forma una envolvente que copia la forma exterior de la nube de puntos. Posee mayor superficie, menos lados y su figura es más regular (y se suele alejar más de la nube de puntos) que su hermano cóncavo.	Utiliza información georreferenciada de los estudiantes
Polígonos envolventes cóncavos	Local	Construcción de un polígono cóncavo que forma una envolvente que copia la forma exterior de la nube de puntos. Posee menor superficie, más lados y su figura es menos regular (y se suele acercar más a la nube de puntos) que su hermano convexo.	Utiliza información georreferenciada de los estudiantes

Centroides	Local	Medida de tendencia central espacial de la nube de puntos de los estudiantes de cada establecimiento	Utiliza información georreferenciada de los estudiantes
Centroides	Radio censal	El Punto medio de los polígonos de los radios puede servir para imputar distancia desde cada uno de ellos al establecimiento educativo y esto incluirlo como insumo en una variable contextual del territorio del establecimiento	Utiliza sólo información georreferenciada de los establecimientos en el caso de variable contextual sobre el territorio.
Densidades	Local	Creación de polígonos para la discriminación de diferentes zonas de densidad de la nube de puntos de los estudiantes.	Utiliza información georreferenciada de los estudiantes
Densidades	Central	Estimación espacial de la demanda teórica en función de la demanda efectiva más un factor de corrección contextual a nivel de radio censal.	Utiliza información georreferenciada de los estudiantes

En función de lo anterior a continuación se realiza un pequeño resumen en función de diferentes objetivos. En este sentido:

- Cuando el objetivo sea la construcción de zonas de influencia de cada establecimiento, los análisis basados en técnicas de densidad parecen ser una técnica idónea. Estos permiten la construcción de diferentes zonas de densidad dentro de un polígono envolvente conformados por la nube de puntos de los estudiantes (sea este en versión cóncava o convexa). Lo anterior se mejora si se complementa con a) el centroide de la nube de puntos de las ubicaciones de los estudiantes, b) un buffer normativo según el nivel del establecimiento y c) el respectivo polígono de Voronoi del establecimiento que indique los límites hasta donde ese establecimiento es el establecimiento más cercano para ese nivel.
- Cuando el objetivo sea la asignación de cada punto del espacio provincial (o distrital o departamental) al establecimiento educativo más cercano, la técnica de los diagramas de Voronoi parece la más idónea. Esta técnica permite su construcción en función de cada nivel (inicial, primario, etc.) o alguna otra clasificación del establecimiento que se considere (p.e. gestión estatal o privada).
- Cuando el objetivo sea la visualización de todas las zonas en donde no se está cumpliendo los criterios normativos de distancia (o tiempo de viaje), se puede calcular la diferencia lógica entre a) el plano formado por el polígono de la provincia (o del distrito o partido) y b) la unión lógica espacial de los buffers (normativos) de los diferentes establecimientos de aquel plano.
- Más en general, las posibilidades que surgen de la combinación de los datos georreferenciados de los establecimientos y los estudiantes parece conformar un área novedosa, sub-explotada y con un futuro promisorio para la planificación y evaluación de las políticas educativas.

- El uso de las distancias de ruteo, si bien plantea una complejidad mayor que el cálculo de la una distancia lineal, tiene como recompensa que puede utilizarse tanto para una mejor adaptación del sistema social a la distribución actual de los establecimientos educativos y su demanda vía:
 - a) la planificación del transporte público (p.e. modificando ramales de colectivos),
 - b) el diseño de senderos seguros escolares y la optimización de la distribución de las fuerzas de seguridad y
 - c) estimar el impacto de la reducción de la distancia si se modifica la distribución de edificios (vía construcción/alquiler de nuevos edificios), la propia infraestructura (p.e. construcción de puentes) y/o la propia demanda de estudiantes (vía políticas que incorporen la distancia o cercanía como criterio de admisión).

Como se puede percibir, ninguna de las herramientas presentadas es de aplicación universal, sino que cada una se adapta mejor a distintos objetivos. También es posible combinar aspectos de una metodología y otra para adaptarlas a las necesidades de cada proyecto.

6.1 APLICACIÓN PRÁCTICA

En más de una sección de este documento se comentó que las técnicas descriptas anteriormente aumentan su valor cuando se las complementa con otra. Dado que cada una de ellas tiene sus fortalezas y debilidades existen algunas combinaciones que, a pesar de su eclecticismo, logran una sinergia positiva que ninguna técnica puede lograr en solitario. En esta sección se intentará mostrar un pequeño ejemplo práctico que muestre esta potencialidad. Siguiendo el mismo hilo conductor del resto del documento, las aplicaciones se efectuarán sobre los mismos establecimientos sobre las cuales se fueron aplicando las técnicas en forma individual. Aquí, a diferencia de la sección anterior, se opta por la estrategia de hacer 2 mapas por cada establecimiento, cada uno con distinto nivel de zoom ya que en cada nivel se visibiliza mejor diferentes aspectos de la técnica.²⁰

En esta ocasión se ha preferido combinar, visualmente, las siguientes técnicas:

- El punto del establecimiento
- El buffer normativo del establecimiento
- El respectivo polígono de Voronoi del establecimiento

²⁰ Cabe destacar que este tipo de estrategias tienen su origen en la limitación de un documento escrito con un procesador de texto en donde el autor, como los antiguos cartógrafos de oficio, tiene que trabajar con una bala de plata en donde se debe decidir, de una vez y para siempre, el nivel de zoom así como la cantidad de capas y leyendas. La tecnología utilizada para producir los análisis y visualizaciones de este documento (lenguaje R) permite realizar mapas interactivos en archivos html que se pueden ver desde un celular y donde el problema anterior es casi inexistente ya que el lector puede ajustar el zoom a los niveles deseados así como seleccionar las capas o layer que prefiere.

- El polígono envolvente cóncavo de la matrícula del establecimiento
- El centroide de la matrícula del establecimiento
- El análisis de densidades de la matrícula del establecimiento

Figura 14: Aplicación práctica. Menor zoom.

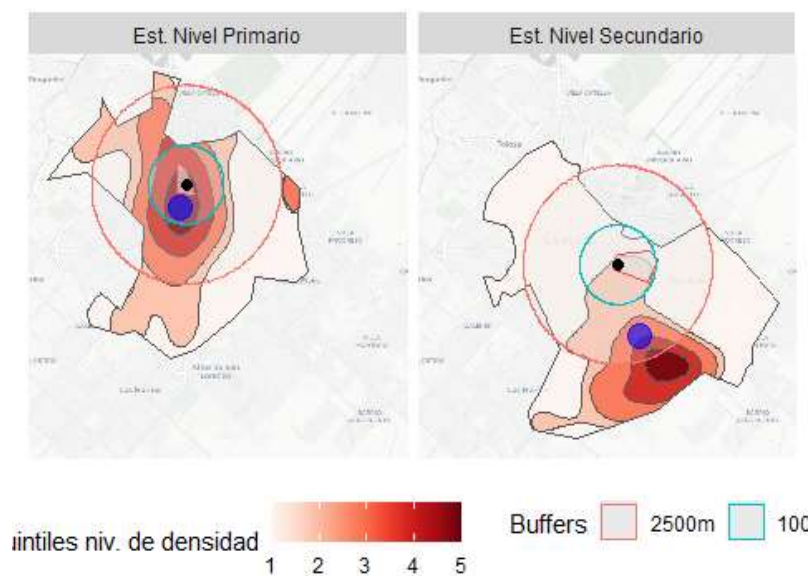
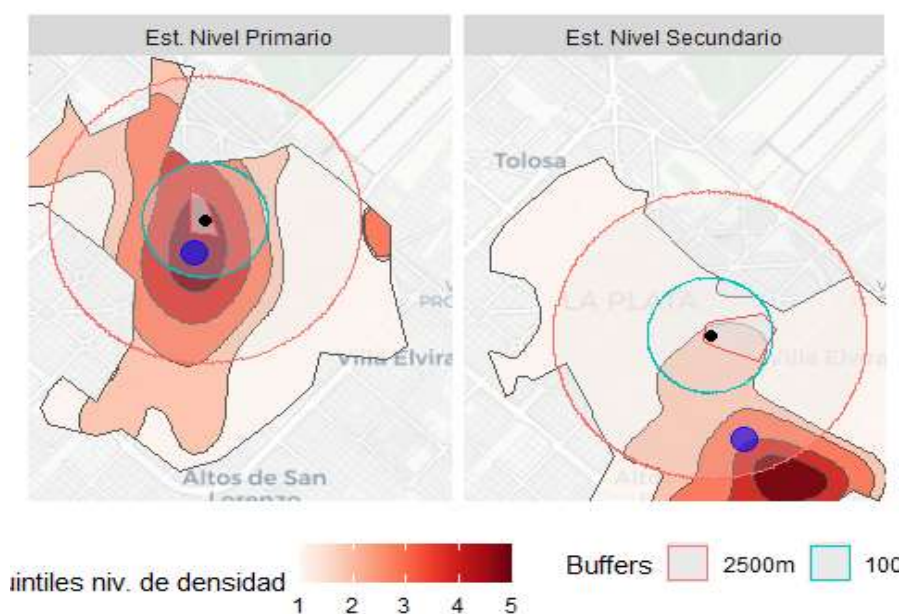


Figura 15: Aplicación práctica. Mayor zoom.



Como se observa en la Figura 14 y en la Figura 15 para cada establecimiento se tiene una imagen que incluye varias de las técnicas que fueron analizadas anteriormente. En ambos establecimientos, en parte por ubicarse en áreas urbanas donde existe una variada oferta de establecimientos, los respectivos polígonos de Voronoi (el polígono irregular que bordea el punto negro de cada establecimiento) poseen un área muy pequeña con relación al área del polígono envolvente cóncavo. Esto ya da una idea que, al menos para ejemplo usado, pocos estudiantes concurren a su establecimiento más cercano.

En el establecimiento del nivel primario se ve cómo el centroide se encuentra mucho más cerca que el respectivo centroide del establecimiento de nivel secundario. Por otro lado se puede apreciar, en ambos casos, como la ubicación del centroide se vuelve razonable dado los respectivos análisis de densidades. En el caso de la imagen de la izquierda el centroide se encuentra relativamente cerca del establecimiento porque la distribución de las densidades es bastante más concéntrica que el establecimiento de nivel secundario. En este último caso, como se observa en el análisis de densidades la mayor cantidad de casos proviene del sur por lo que el centroide también se inclina hacia el sur.

En resumen, la combinación de las diferentes técnicas aquí descritas parece ofrecer una sinergia en tanto (algunos de) los resultados se hacen más razonables en función de los resultados de (algunas de) las otras técnicas. Por otro lado, se pueden observar intuitivamente el grado de cumplimiento de la normativa vigente y cuantos de los estudiantes que forman parte de la matrícula tienen a ese establecimiento como el establecimiento más cercano.

Finalmente, mucho de lo anterior se puede complementar con tablas y gráficos que incorporen datos que sean productos de lo que aquí se denominó “técnicas espaciales analíticas”. En este sentido los informes sobre distancia parecen ser los que primero vienen a la cabeza, pero también se pueden construir indicadores analíticos que indiquen el porcentaje de estudiantes que cumplen la normativa de distancia o tiempo y que porcentaje tiene a ese establecimiento como el establecimiento más cercano.

7. CONCLUSIONES

A lo largo del trabajo se han repasado diferentes alternativas para la realización de análisis espaciales dentro del campo de la planificación y gestión educativa. Algunos de las técnicas propuestas hacen uso de información exclusivas de los establecimientos, otras de información exclusiva de los estudiantes, otras aprovechan la información provenientes de otras fuentes (p.e. censos) y, finalmente, otras sacan ventaja de la sinergia resultante de alguna combinación las fuentes anteriores.

En este punto es útil destacar la novedad y potencialidad que posee la georreferenciación de los estudiantes para muchas de estas técnicas de análisis. Si bien la georreferenciación, esto es, la técnica de relacionar información con un punto o zona geográfica, es una práctica antigua, tuvo un relajamiento fuerte a mediados de 1970 con el surgimiento de los modernos sistemas GIS y la posibilidad de hacer mapas con diferentes capas o layers (Hill 2006, cap. 1). El cambio anterior aumentó la utilidad de la georreferenciación pero no su facilidad de construcción. Esto lo hicieron en forma posterior, tanto la difusión de internet como la posterior difusión de los teléfonos celulares ya que ahora cualquier casi cualquier persona podía producir información georreferenciada. Esto posibilitó lo que a veces se denomina el surgimiento de la ciencia ciudadana (Haklay, Mazumdar, y Wardlaw 2018). Esta se caracteriza por la participación activa o pasiva de ciudadanos no científicos en proyectos científicos y, a modo de ejemplo, cualquier proyecto científico que utilice como insumo datos de OpenStreetMap (OSM) estaría haciendo uso de la ciencia ciudadana. En este trabajo en particular, la información almacenada en la base de datos de OSM puede ser explotada mediante diferentes librerías de manera bastante automatizada lo que permite que a grandes cantidades de direcciones de personas particulares se les pueda asignar una coordenada geográfica específica.

Si bien este último tipo de dato no se puede publicar por cuestiones de confidencialidad de la información, tampoco se trata de un dato en que su importancia remite a aquella característica que no lo hace publicable. Más bien, como se ha visto a lo largo de este documento, existen una variedad de técnicas que, usando esta información como insumo, al tiempo que preservan el anonimato del dato producen síntesis útiles para la toma de decisiones. En este sentido es pertinente aclarar que, especialmente cuando se combinan apropiadamente ([Sección 6.1](#)), el uso de diferentes técnicas puede dar una idea cercana a la distribución real de la nube de puntos de los estudiantes a pesar de nunca mostrarse esta explícitamente. No sólo eso. La utilización de técnicas que captan visualmente de modo sintético la forma o silueta ([Sección 5.4](#)), el centro espacial ([Sección 5.5](#)) y diferentes zonas de densidad ([Sección 5.6](#)) de una nube de puntos puede ser de más ayuda que la visualización de la propia (e irregular) nube de puntos ya que la misma, en la gran mayoría de las situaciones, es difícil de resumir o sintetizar.

Robert K. Merton afirmó hace más de medio siglo que “la teoría sólida sólo prospera sobre una dieta rica en hechos pertinentes, y procedimientos recién inventados ayudan a suministrar los ingredientes de esa dieta” (Merton 2002, pag. 189). En este sentido, la posibilidad de contar con datos georreferenciados de los estudiantes, aceptables tanto en su calidad como en su representatividad, puede considerarse un novedoso tipo de ingrediente para la dieta de algunas teorías educativas, así como para la planificación y evaluación de las políticas públicas tanto a nivel central como local.

Desde un punto de vista académico, la unión de la georreferenciación a nivel de estudiante con la georreferenciación de los establecimientos puede ser el primer paso para una fertilización cruzada entre diferentes campos disciplinarios de las ciencias sociales y los análisis espaciales. Desde un punto de vista más cercano a la planificación y la gestión el cruzamiento anterior aporta herramientas analíticas y visuales para la creación de nuevos indicadores y una planificación más inteligente de las políticas educativas y una evaluación más precisa de ellas. Desde un costado más instrumental, la economía de escala de un organismo no local (p.e. un organismo provincial) para sistematizar y georeferenciar las diferentes informaciones locales no parece comparable a la propia capacidad de cada establecimiento para sistematizar su propia información local. Este hecho crea una ventana de oportunidad para que desde el centro se puedan generar informes automáticos interactivos e individualizados para cada uno de los establecimientos que difícilmente pueda ser producido por cada uno de los establecimientos que produce cada información local.

REFERENCIAS BIBLIOGRÁFICAS

- Baddeley, Adrian, Ege Rubak, y Rolf Turner. 2016. *Spatial Point Patterns*. Boca Raton: CRC Press.
- Born, Diego. 2019. «Índice de contexto social de la educación. ICSE.»
- Born, Diego, y Martín Scasso. 2019. «Uso y producción de índices para la medición de temas sociales y educativos».
- Brock, Colin. 2016. *Scale, space and location in the study of education*. Blommsbury Academic.
- Bryant, Bryant. 2018. «Income segregation across schools and the shapes of school attendance zones». Tesis doctoral, Williamsburg.
- Bunge, Mario. 1995. *Sistemas sociales y filosofía*. Buenos Aires: Sudamericana.
- Burt, Ronald. 2005. *Brokerage & Closure*. Great Clarendon Street: Oxford University Press.
- Capel, Horacio. 2016. «Las ciencias sociales y el estudio del territorio». *Biblio 3W. Revista bibliográfica de geografía y ciencias sociales* XXI (1149).
- Chacón, José, y Tarn Duong. 2018. *Multivariate kernel smoothing and its applications*. Boca Raton: Taylor & Francis.
- Coleman, James. 1988. «Social capital in the creation of human capital». *American Journal of Sociology* 94: 95-120.
- . 1990. *Foundations of social theory*. Cambridge: Harvard University Press.

- Dimaggio, Charles, y Guohua Li. 2013. «Effectiveness of a safe routes to school program in preventing school aged pedestrian injury». *Pediatrics* 131 (2): 290-96.
- Edelsbrunner, Herbert. 2011. «Alpha Shapes - A survey». En. Springer.
- Edelsbrunner, Herbert, Michael Facello, Ping Fu, Nataraj Akkiraju, Carlos Varela, y Ernst Mücke. 1995. «Alpha shapes: Definition and Software». *International Computational Geometry Software Workshop*, 1-4.
- Edelsbrunner, Herbert, David Kirkpatrick, y Raimund Seidel. 1983. «On the shape of a set of points in the plane». *Transactions on information theory* 29 (4): 551-59.
- Eros, Tibor, y Winsor Lowe. 2019. «The Landscape Ecology of Rivers: from Patch-Based to Spatial Network Analyses». *Current Landscape Ecology Reports* 4: 103-12.
- Gelfand, Alan, Peter Diggle, Montserrat Fuentes, y Peter Guttorp, eds. 2010. *Handbook of spatial statistics*. Boca Raton: Chapman & Hall /CRC.
- Gimond, Manuel. 2020. «Intro to GIS and Spatial Analysis». <https://mgimond.github.io/Spatial/index.html>.
- Goldring, Ellen, y Walker Swain. 2014. «The school attendance and residential location balancing act. Community, choice, diversity and achievement». En, editado por Gregory Ingram y Daphne Kenyon, 92-116. Cambridge: Lincoln Institute.
- Graham, Ronald. 1972. «An efficient algorithm for determining the convex hull of a finite planar set». *Information Processing Letters* 1: 132-33.
- Haklay, Muki, Suvodeep Mazumdar, y Jessica Wardlaw. 2018. «Citizen Science for observing and understanding the earth». En, editado por Pierre Mathieu y Christoph Aubrecht, 69-88. Cham: Springer.
- Hill, Linda. 2006. *Georeferencing. The geographic associations of information*. Cambridge: MIT Press.
- Kjellström, Christian, y Håkan Regnér. 1999. «The Effects of Geographical Distance on the Decision to Enrol in University Education». *Scandinavian Journal of Educational Research* 43 (4): 335-48. <https://doi.org/10.1080/0031383990430401>.
- Knuth, Donald. 1984. «Literate Programming». *The Computer Journal* 27: 97-111.
- Land, Kenneth C. 1983. «Social indicators». *Annual Review of Sociology* 9 (1): 1-26. <https://doi.org/10.1146/annurev.so.09.080183.000245>.
- Land, Kenneth C., y Abbott Ferriss. 2007. «The sociology of social indicators». En, editado por Clifton Bryant y Dennis Peck, 518-26. California: Thousand Oaks. SAGE Publications.

- Larouche, Richard, George Mammen, David Rowe, y Guy Faulkner. 2018. «Effectiveness of active school transport interventions: A systematic review and update». *BMC Public Health* 2016 (18): 1-18.
- Llanos Hernández, Luis. 2010. «El concepto del territorio y la investigación en las ciencias sociales». *Agricultura, sociedad y desarrollo* 7 (3): 207-20.
- Logan, John R. 2012. «Making a Place for Space: Spatial Thinking in Social Science». *Annual review of sociology* 38 (agosto). <https://doi.org/10.1146/annurev-soc-071811-145531>.
- McDonald, Noreen. 2008. «Children's mode choice for the school trip: The role of distance and school location in walking to school». *Transportation* 35: 23-35.
- Merton, Robert K. 2002. *Teoría y estructura sociales*. Mexico: Fondo de Cultura Económica.
- Møller, Jesper, y Rasmus Waagepetersen. 2004. *Statistical inference and simulation for spatial point processes*. Boca Raton: Chapman & Hall /CRC.
- Monarrez, Tomas. 2019. «Rising to the Challenge: Engaging diverse perspectives on issues and evidence». En. Denver.
- Park, Jin Seo, y Se Jong Oh. 2012. «A new concave hull algorithm and concaveness measure for n-dimensional datasets». *Journal of information science and engineering* 28: 587-600.
- Perry, Clarence. 1914. «The school as a factor in neighborhood development». En. Memphis: Department of recreation Russell Sage Foundation.
- Quartulli, Diego, y Fermín Aispuro. 2021. «Análisis de la distancia de las escuelas a los hogares de los estudiantes de la Provincia de Buenos Aires».
- Roth, Alvin. 2016. *Who get what and why*. London: William Collins.
- Sack, Robert. 1980. *Conceptions of space in social thought*. London: The Macmillan Press LTD.
- Saporito, Salvatore. 2017. «Irregularly shaped school attendance zones and racial integration». *Social Science Research* 64: 299-318.
- Saporito, Salvatore, y David Van Riper. 2016. «Do irregularly shaped school attendance zones contribute to racial segregation or integration?» *Soc. Curr.* 3 (1): 64-83.
- Silverman, Bernard. 1998. *Density estimation for statistics and data analysis*. Boca Raton: Chapman & Hall /CRC.
- Tobler, Waldo. 1970. «A computer movie simulation urban growth in the Detroit region». *Economic Geography* 46: 234-40.

