

Universidad de Santiago de Compostela  
Facultad de Ciencias Políticas y Sociales  
Grado en Ciencia Política y de la Administración



## TRABAJO FIN DE GRADO

*Matrices de vecindad, análisis geoespacial y puticlubes.*

*Matrices de vecindade, análise xeoespacial e puticlubes.*

*Weights matrices, spatial analysis and brothels.*

Alumno: Francisco Filgueira Ferreiro

Director: Celestino García Arias

Santiago de Compostela

Julio 2020

Este trabajo se realizó con LaTeX y RStudio en su totalidad<sup>1</sup>, funcionando ambos mediante lenguaje de programación y comandos. Para los menos familiarizados con la programación, tanto LaTeX como R pertenecen al conjunto de programas llamados de *software libre y de código abierto*. Éstos permiten un fácil acceso [gratuito] al tratamiento de datos, así como la modificación de su código, lo que da la posibilidad de mejorarlos y distribuirlos sin coste para los interesados; además, son herramientas increíblemente potentes, que dejan atrás a los programas con licencia y de pago que se utilizan en esta facultad <sup>2</sup> —y en la mayoría—.

Desde aquí se quiere animar a la implementación de temario que utilice estas herramientas, aunque sea de manera introductoria. Sí es cierto que el manejo de comandos puede presentar más complejidad que el uso de programas mediante interfaz gráfico, pero desde luego esta complejidad no es tal para que el alumnado de universidad se vea superado por ello. Además, permiten la manipulación precisa y concreta de todos los pasos en el manejo de datos, aumentando el conocimiento y la calidad de los análisis. Programas como RStudio en un entorno Linux no solo permiten el manejo en profundidad de los datos, sino que también son capaces de integrar texto mediante código con softwares como LaTeX o Markdown que son realmente útiles y adecuadas para la escritura de texto e investigación reproducible. Por otra parte, disponen de millones de opciones [expandibles mediante paquetes] para la edición de texto, gráficas, tablas, etc. La mayor parte de las pruebas estadísticas y de las gráficas [por ejemplo los mapas] realizadas en este trabajo no se podrían lograr mediante Word y SPSS. Si bien estos últimos sirven para la edición de texto y manejo estadístico *básico* y de uso general, se ven limitados a medida que aumenta la complejidad [o la especificidad] de las técnicas de manejo y análisis de datos. En cambio R —entre otros tantos— corrige estos defectos, convirtiéndose así en una herramienta más versátil y dinámica en comparación con los programas de licencia anteriormente mencionados.

---

<sup>1</sup>Todos los elementos gráficos [mapas, tablas, etc.], cálculos y análisis de datos se hicieron para este trabajo dentro de el entorno de RStudio

<sup>2</sup>SPSS y el paquete Office de Windows.

## **Agradecimientos**

A mis padres, por su inestimable labor en el trabajo de campo.

A mi tutor, por su ayuda y consejos.

# Índice

Introducción . . . . .	4
<b>I Matriz de vecindad de orden variable</b>	<b>6</b>
¿Qué es una matriz de vecindad? . . . . .	6
Notación utilizada . . . . .	7
Definición formal de las matrices . . . . .	8
Ejemplo gráfico . . . . .	8
Restricciones . . . . .	10
<b>II Análisis geoespacial de prostíbulos</b>	<b>12</b>
¿Por qué análisis geoespacial? . . . . .	12
<b>Revisión bibliográfica</b>	<b>12</b>
<b>Variables e indicadores</b>	<b>16</b>
<b>Fuentes de datos y selección de casos</b>	<b>18</b>
<b>Técnicas de análisis</b>	<b>19</b>
Regresión espacial . . . . .	21
Modelos espaciales . . . . .	21
<b>Hipótesis</b>	<b>22</b>
<b>Análisis de datos</b>	<b>23</b>
Análisis exploratorio . . . . .	23
Mapas . . . . .	24
Test . . . . .	28
Elección del modelo mediante AIC . . . . .	29
Estimación del modelo y resultados . . . . .	33
Conclusiones . . . . .	39
<b>Referencias bibliográfica</b>	<b>41</b>

## Introducción

Este trabajo se asienta sobre tres elementos principales: por un lado la matriz de vecindad, algo básico en modelos de regresión espacial, pero cuyas formas más simples son generadores de problemas por imprecisión o por su mala adaptación a las particularidades de cada territorio; de segundo tenemos el análisis geoespacial, que no es tratado como un tema en sí, sino como una herramienta para alcanzar el fin aquí propuesto; y por último los prostíbulos, el elemento central del trabajo, y más concretamente, su distribución en el espacio.

Lo que se hará a continuación es, explicar un tipo de matriz de vecindad ideada en y para este trabajo [aunque seguramente ya ideada anteriormente por alguien] que, mediante el uso de herramientas de análisis espacial, ayudará a determinar qué elementos influyen en la distribución de los prostíbulos a lo largo y ancho de la provincia.

¿Por qué la elección de estos temas? Principalmente por curiosidad. Es muy común que yendo en coche por las carreteras de la provincia de Pontevedra se encuentren edificios con neones que indican —disimuladamente— la presencia de un prostíbulo. También es común que los *vecinos* conozcan su existencia y su ubicación, y a pesar de ser algo que juega en contra de la legislación, no se esconden. Además, si bien están presentes en zonas de tráfico, también lo están en zonas alejadas de núcleos de población y de menor tránsito, al menos aparentemente, y dónde —en principio— no cabría esperar su presencia. Entonces las preguntas que me planteo resolver son: ¿Por qué se ubican como lo hacen los prostíbulos? ¿Se distribuyen aleatoriamente? ¿Existen elementos que condicionen su distribución espacial? Y si es así, ¿cuáles son?

Las respuestas a estas preguntas serían interesantes y útiles —por ejemplo— para las entidades locales en las que se encuentran. Si la ley es igual en todo el territorio de un país, ¿por qué tenemos actividades ilegales situadas en puntos concretos del espacio? Es posible que por simple azar, pero también es posible que no sea así.

Si la regulación de este tipo de negocios es algo —al parecer— complicado, conociendo los condicionantes que intervienen en su distribución podríamos tener herramientas más potentes para una regulación más eficiente y eficaz. Y dejando de lado la importancia *política* de la prostitución, tenemos un problema de tipo social; estamos ante un *fenómeno* que genera sufrimiento en las profesionales que se dedican a esto, con una [re]presión legal y social —aparentemente— desigual en según qué zonas, además del sometimiento al estigma y rechazo de una gran parte de la población. Un ejemplo de esta desigualdad son las numerosas las noticias de cierres de burdeles en las ciudades, pero muchos de los prostíbulos de carretera continúan su actividad hoy día tras una *trayectoria profesional* de más de diez años de antigüedad, y en muchos casos

asociados al tráfico de drogas y de seres humanos. A simple vista también parece que existe más volatilidad en el negocio en entornos urbanos, en contraste con los rurales. En estos últimos es común que las casas de prostitutas sostengan de mejor manera su actividad en el tiempo, en muchos casos formando parte de la 'normalidad' del pueblo, por lo que parece que existe un importante componente territorial.

Por lo tanto me propongo, mediante el uso de técnicas de estadística espacial, señalar algunos de los elementos que condicionan la distribución de la prostitución en la provincia de Pontevedra, y realizar un estudio *piloto* en el problema de la prostitución desde un enfoque espacial que pueda servir de inspiración o apoyo en futuros trabajos.

## Parte I

# Matriz de vecindad de orden variable

## ¿Qué es una matriz de vecindad?

Una matriz de vecindad, de ponderaciones o pesos, es uno de los elementos principales en modelos econométricos para datos georreferenciados (Herrera, Mur, Ruiz, 2011). Es utilizada para reflejar relaciones entre elementos en el espacio, generalmente de vecindad y, en los casos más sencillos, esto se hace mediante 0 y 1 para no existencia de vecindad y sí existencia respectivamente. Cada fila  $i$  corresponde a un objeto<sup>3</sup> espacial, y cada columna  $j$  corresponde al mismo objeto espacial; es decir, los mismos objetos se disponen en el mismo orden a lo largo de las filas y las columnas para así relacionarlos en cada celda  $M_{ij}$ . A cada par de vecinos le corresponderá un número en cada celda de la matriz.

Puesto que  $i = j$ , las matrices de ponderación de vecinos son cuadradas, ya que las relaciones se computan entre cada objeto  $1, 2, \dots, ij$ , y pudiendo reflejarse en ellas relaciones simétricas o asimétricas.

Una de las matrices de vecindad más utilizada en análisis geoespacial es la basada en un modelo *queen* de contigüidad. En ella, se consideran como vecinos del polígono  $i$  aquellos polígonos  $j$  que lindan con su borde; de otro modo,  $i$  y  $j$  son vecinos si la intersección de sus límites no es igual a 0. En este caso, estamos ante uno de los modelos más simples y más utilizados de matrices de vecindad por su facilidad de cálculo, y sobre todo, por su simple definición de *vecindad*.

También existen matrices de orden superior, en las que se consideran vecinos de  $i$  a aquellos polígonos que también lo son de  $j$  y no de  $i$  de manera directa, aquí denominados  $N_n$ ; de forma que,  $N_{(2)}$  será vecino de  $i$  si también lo es de  $j$  aunque los bordes de  $N_{(2)}$  y  $i$  no se toquen, siguiendo la definición anterior y para vecinos de orden 2.

Pero estos modelos no tienen en cuenta las características del territorio; al aplicarlos se consideran vecinos a *todos* los polígonos que cumplan estos requisitos, sin importar las características de éstos. Por eso se propone la utilización de una matriz de orden superior y variable para una región dada. De esta manera, a parte de contar con las ventajas de la contigüidad *queen*, también se considerarán como vecinas aquellas áreas que, aunque estén lejos en el espacio, sus

---

<sup>3</sup>Polígonos, vectores, puntos, áreas, etc.

características permitan considerarse vecinos de otras áreas, bien por su importancia geoespacial o por su influencia. Así, se pueden considerar vecinos polígonos que no lindan directamente con un determinado polígono, pero que la influencia de aquellos sobre éstos sea suficiente para considerarse como relacionados<sup>4</sup> espacialmente. La diferencia está en considerar vecinos a aquellos polígonos de orden superior en función de una *variable* que marque la importancia del polígono y evitando considerar como vecinos a *todos* los polígonos —independientemente de sus características— que cumplan con las definiciones por contigüidad anteriormente mencionadas.

## Notación utilizada

Antes de empezar con la expresión formal de la matriz, se explicará la notación utilizada.  $n$  se utiliza para expresar el *orden* de una matriz o polígono. Una matriz de orden 2 sería aquella en la que:  $n = 2$ .

$d$  hace referencia a la *distancia* entre dos puntos o polígonos. Ésta guarda relación directa con el orden [se explica esto más adelante]. Dos polígonos con intersección  $\neq \emptyset$  de sus bordes tienen distancia 0; polígonos separados por otro *intermedio* están a distancia 1, y así sucesivamente.

$N$  —neighbor— se emplea para referirse a *vecinos*. Para vecinos de orden 2, por ejemplo, tenemos:  $N_2$ . Para cualquier orden, tendremos:  $N_n$ . Del mismo modo, para un vecino de  $i$  tendremos:  $N_{(i)}$ . Y para un vecino de orden 2 de  $i$  tendremos:  $N_{(i2)}$ .

$P$  es utilizado para representar a un *polígono*, siendo estos las unidades básicas que conforman el mapa. La distinción entre *vecino* y *polígono* es útil, también, para referirse a un área cuando aún no sabemos si es vecino o no. Como en el caso anterior, para un polígono con relación de orden 2 respecto de  $i$  tenemos:  $P_{(i2)}$ .

$bnd$  —boundary— se utiliza como *límite* o *frontera*, indistintamente. Como en los casos anteriores, para el *límite* del polígono  $i$  tendremos:  $bnd_{(i)}$ .

$a$  se utiliza como una *variable* cualquiera, a elegir. Ejemplos de ésta serían la población, la renta, densidad poblacional, etc. Del modo anterior, para hacer referencia a una variable de  $i$  —por ejemplo—, tenemos:  $a_{(i)}$ .

---

<sup>4</sup>Generalmente en una relación asimétrica, en la cuál  $N_{(2)}$  es vecino de  $i$  pero  $i$  no es vecino de  $N_{(2)}$ .

## Definición formal de las matrices

Según la notación anterior, la definición de una matriz tipo queen de orden 1 sería:

**Matriz contigüidad queen** de orden 1.

$$(1)(bnd_{(i)} \cap bnd_{(j)} \neq \emptyset) \implies (j) \in N_{(i)}$$

$$(2)(bnd_{(i)} \cap bnd_{(j)} = \emptyset) \implies (j) \notin N_{(i)}$$

Se escribe 1 en la matriz si la intersección de  $i$  y  $j$  es distinta de cero, es decir, si intersecan sus fronteras en algún punto; por el contrario, si ésta es igual a cero, se reflejará con un 0.

La matriz de orden superior basada en contigüidad queen ideada para este trabajo sería:

**Matriz orden superior variable**

$$(3)(a_{P_i} \geq x) \text{ y } (1 \leq d_{(P_i-P_n)} \leq n-1) \implies P_i \in N_{(P_n)}$$

$$(4)(a_{P_i} < x) \implies (1) \text{ o } (2)$$

Lo que se ha hecho es añadir una condición a las definiciones de vecindad según un modelo queen simple. En (3) tenemos que, si una variable  $[a]$  a elegir de un polígono en concreto  $i$   $[P_i]$  iguala o supera un determinado umbral  $[x]$  y la distancia entre este polígono  $i$  y uno de orden  $n$  es mayor o igual a 1 y menor que el orden máximo permitido menos 1  $[n-1]$ , entonces el polígono  $i$  pertenece al conjunto de vecinos del polígono de orden  $n$ . Lo que dice (4) es que, si la variable de un polígono  $i$  no supera este umbral, entonces las reglas de decisión de vecindad/no vecindad son las propias de un modelo queen de orden 1, que vienen explicadas por (1) y (2).

## Ejemplo gráfico

Para ver mejor qué hace esta matriz, la explicaremos con ayuda de la figura 1; primero utilizando un modelo de contigüidad *queen* y después uno de *orden variable*.



Figura 1: Elaboración propia

El polígono de referencia será  $i$ ,  $j$  su vecino por contigüidad y  $P$  es vecino de  $j$  pero no de  $i$ , utilizando la contigüidad *queen*. Si se trabaja con un modelo de orden superior *queen simple*, como por ejemplo  $n = 2$  tenemos que  $P$  y  $i$  son vecinos. En definitiva, con matrices de orden 2, se considera vecino de  $i$  todo vecino de  $j$  —independientemente de sus características—, siempre que de primeras no linde con  $i$  [ya que de esta forma sería  $j$ ].

En el caso de modelos de orden variable, habría que considerar las características de un polígono  $P$  de orden  $n$ . Partiendo de que, polígonos de tamaño más grande *podrían* tener más influencia en sus cercanos, sería aceptable decir que para la *figura 1* el polígono  $P$  —por su tamaño— ejerce influencia en  $i$ , pero no al revés; considerando de esta forma a  $P$  vecino de orden 2 de  $i$  pero no al revés.

Esto pasaría al establecer un umbral  $x$  para una variable  $a$  de un polígono  $n$ . Si estamos interesados en mostrar la importancia del tamaño del polígono, podríamos escoger como  $a$  la variable  $m^2$ ; situando un umbral —por ejemplo— en  $200 m^2$ , consideraríamos vecino, en la figura 1 a  $P$  si su área supera este umbral.

Expresado de otra manera tendríamos:

*queen*  $j \in N_{(i)}$  y  $i \in N_{(j)}$

*orden variable* ( $P$  y  $j$ )  $\in N_{(i)}$  pero  $i \notin N_{(P)}$  aunque  $i \in N_{(j)}$ .

Las matrices correspondientes serían:

$$\text{Matriz orden variable} = \begin{bmatrix} & \mathbf{p} & \mathbf{j} & \mathbf{i} \\ \mathbf{p} & 0 & 1 & 0 \\ \mathbf{j} & 1 & 0 & 1 \\ \mathbf{i} & 1 & 1 & 0 \end{bmatrix} \quad \text{Matriz queen} = \begin{bmatrix} & \mathbf{p} & \mathbf{j} & \mathbf{i} \\ \mathbf{p} & 0 & 1 & 0 \\ \mathbf{j} & 1 & 0 & 1 \\ \mathbf{i} & 0 & 1 & 0 \end{bmatrix}$$

En la primera matriz, las columnas reflejan la condición de vecinos de las filas, y de este modo tendría igualmente sentido trabajar con una matriz estandarizada por fila. Existe asimetría en la primera, ya que la vecindad puede considerarse asimétrica en caso de existir relaciones unidireccionales; es decir, de  $j$  hacia  $i$ , y no al revés.

## Restricciones

La primera limitación que existe viene dada por la relación entre distancia  $d$  y orden  $n$ . Ésta viene definida por:

$$d = n - 1.$$

Un polígono  $j$  anexo a uno  $i$  presentan distancia 0 y orden 1; si anexamos otro polígono  $P$  a  $j$  sin que toque  $i$ , entonces tenemos una distancia de 1 y de orden 2 entre  $i$  y  $P$ : respectivamente tendríamos  $n = 1, d = 0$ ; y  $n = 2, d = 1$ .

Para el caso de la figura 1, la distancia entre  $i$  y  $j$  es de 0, y de  $i$  y  $P$  es de 1; respectivamente, el orden de la relación es 1 para el primer caso y 2 para el segundo.

Una restricción para la construcción de esta matriz aparece en la relación orden-distancia; la distancia entre dos polígonos, no puede superar al orden  $-1$ , tal que:

$$\underline{d \leq n - 1}$$

Si la distancia de  $P$  supera a  $n - 1$ , entonces éste escapa del alcance de su orden, por lo que de ninguna manera sería considerado como vecino  $N$ . Las consideraciones sobre vecindad serán hechas siempre dentro del rango marcado por el orden. Por ejemplo, para un polígono  $n = 3$ , su distancia no podrá ser  $> 2$ , puesto que si fuera así significaría que existen 3 o más polígonos en el medio, escapando al límite marcado por su orden.

Por otra parte, dados tres polígonos  $P$ ,  $j$  y  $i$  deben guardar relación de proximidad entre ellos.  $P$  debe lindar con  $j$ , y  $j$  lindar con  $i$ , pero  $P$  y  $i$  no deben hacerlo entre sí. Formalmente tenemos:

$$[bnd_{P_n} \cap bnd_{(P_{n-1})} \neq \emptyset] \wedge [bnd_{P_{n-1}, \dots, n=2} \cap bnd_{(P_{n-2}, \dots, n=1)} \neq \emptyset] \wedge [bnd_{(P_N)} \cap bnd_{(i)} = \emptyset]$$

La intersección de frontera del polígono de orden  $n$  y la de su adyacente  $n - 1$  debe ser  $\neq 0$ , al igual que la del polígono  $n - 1$  y  $n - 2$  [en el caso de darse] y así sucesivamente hasta llegar a  $i$ ; esto es,  $P_n$  y  $P_{n-1}$  deben ser vecinos entre sí, continuando la cadena de enlace hasta  $j$  y  $i$ . Pero como se ha dicho, los polígonos no pueden intersecar fronteras con otros a una distancia mayor a 0, ya que la vecindad en órdenes superiores siempre se producirá mediante la cadena de adyacencia.

Por otra parte, esta relación estará presente en todos los polígonos contiguos, desde  $n$  hasta  $n - p$ , llegando hasta  $n = 1$ , es decir,  $j$ .

Por lo tanto:  $[bnd_{n,n-1,n-p,\dots,n=2} \cap bnd_{(i)} = \emptyset]$ , es decir, las fronteras de los polígonos de ordenes superiores a 1 que continúan esta cadena  $[n, n - 1, n - 2, \dots]$  y las del polígono de referencia  $i$  deben de ser cero; no se deben tocar de manera directa.

En las matrices de orden superior, los polígonos considerados como vecinos están definidos a partir de un polígono de orden  $n$ , por lo que según lo explicado, tenemos que:  $i \in N_n$  y  $N_j$ , pero  $P_n \notin N_i$ , aunque  $j \in N_i$ .

Explicado todo esto pasamos con el análisis de la geografía de la prostitución en la provincia de Pontevedra.

## Parte II

# Análisis geoespacial de prostíbulos

## ¿Por qué análisis geoespacial?

El componente espacial de los fenómenos es algo que adquiere especial relevancia en según qué casos. Por ejemplo, es innegable que la propagación de una enfermedad tiene un comportamiento determinado a lo largo del espacio, ya que la proximidad juega un papel fundamental en el aumento de la probabilidad de contagio. En la distribución de especies de árboles, plantas o animales también es importante tener en cuenta los factores geográficos que puedan influir, puesto que de no hacer esto estaríamos ignorando un elemento clave en nuestro análisis.

En ciencias sociales esto cobra especial importancia a la hora de estudiar ciertos fenómenos. Se sabe que la renta, el voto, industria, etc. se sitúan en determinadas regiones y se agrupan en zonas de características similares, siendo habitual su distribución en *cluster* o en grupos. Es normal que elementos próximos en el espacio posean características similares y presenten relaciones debido a su proximidad. Esto nos lleva a la primera ley de la geografía de Tobler, la cual dice que *todas las cosas se relacionan entre sí, pero las cosas más próximas tienen una mayor relación que las distantes*.

## Estudios previos

Desde el siglo XIX hasta casi mediados del siglo XX, los lugares de concentración de prostíbulos y burdeles estaban inmersos en la vida cotidiana de la ciudad. Era común la agrupación de estos negocios en sectores concretos, generalmente en zonas céntricas (Bertozzi, 2005, p. 6). Éstos se situaban en calles, barrios o áreas de las urbes dedicadas al *vicio* y al ocio masculino en general, pero estaban lejos de esconderse. Entre los negocios habituales en estas zonas destacaban las licorerías, casinos y teatros eróticos<sup>5</sup>, entre los que se encontraban los prostíbulos.

Entre las actividades de ocio de la época destacaban por excelencia el consumo de alcohol y el sexo, que se podían satisfacer fácilmente en los *distritos de luz roja*. Éstos eran frecuentados por hombres, aportando cantidades de dinero suficientes como para sostener de buena forma la actividad. Además, era en las ciudades donde el poder adquisitivo era más elevado, y por lo

---

<sup>5</sup>Generalmente en capitales de país o grandes ciudades.

tanto, más fructíferos este tipo de negocios.

Por otra parte, el rechazo moral de la prostitución no era el mismo que es en la actualidad. Anteriormente, ésta era considerada —incluso— como una forma de prevenir la infidelidad, pues el sexo con estas *mujeres perdidas* no estaba considerado como motivo de ruptura matrimonial. Comúnmente las mujeres [especialmente en la clase media] consideraban como normal el hecho de que sus maridos visitasen prostitutas y mantuviesen sexo con ellas fuera del matrimonio (Bertozzi, 2005, p. 10).

Fue a partir de cuartos del siglo XX cuando esta situación empezó a cambiar con el reforzamiento de la moral católica y el endurecimiento de las leyes, provocando que los prostíbulos tuvieran que reducir su visibilidad y ejercer sus actividades en zonas con una permisividad más alta. La estigmatización moral fue uno de los grandes motores de la descentralización de la prostitución en los territorios, acompañada de una mayor presión policial y con la transformación de estos códigos morales en leyes. Es evidente que las leyes se elaboran para controlar las acciones de la gente, y que éstas tienen un carácter universal, o al menos nacional. Pero en realidad, la interpretación legal y su promulgación se sitúan de determinada manera en el espacio, a través de códigos morales y normas que varían de un lugar a otro (Pue 1990). Por otra parte, la aplicación de la ley por parte de la policía no es en todos casos rigurosa. En muchas ocasiones, la policía es más permisiva ante la situación de algunas de las prostitutas, ya que en realidad no desean arrestarlas (Hubbard, 1999, p. 114). Además, como también dice este autor, las relaciones entre las prostitutas y la policía suelen ser cordiales, hasta el punto ocasionar molestias en los vecinos por la indulgencia del trato.

Como también añade Sara Bertozzi (2005, p. 19) en su estudio, la estigmatización social y la represión policial están estrechamente relacionadas, ya que los gobiernos municipales —incluso los estatales— intervienen en el negocio del sexo —en parte— como respuesta a la oposición pública generada. En relación a esto tenemos una forma indirecta de vigilancia policial, puesto que en muchos casos no son los propios cuerpos de seguridad los únicos responsables del cumplimiento de la legislación; es frecuente que las denuncias provengan de vecinos molestos o afectados por tales actividades.

En la actualidad, los estudios sobre la distribución geográfica de la prostitución indican que los elementos que la condicionan son puramente de carácter mercantil, buscando la maximización de sus beneficios —mediante un flujo constante de clientela— y reduciendo los riesgos asociados —mediante el alejamiento de núcleos urbanos—. Además, estudios empíricos sugieren que la

prostitución está más aceptada en ciertos sitios, generalmente de bajo nivel socioeconómico (Helderop, Huff, Morstatter, Grubestic y Wallace, 2019, p. 946), lo que genera problemas de clase y estigmas sociales (Hubbard 1997; Larsen, 1992), pero que permite encontrar a los propietarios de prostíbulos un lugar idóneo para su actividad. De manera sencilla, lo que está aceptado en un sitio no lo está necesariamente en otro, y como sugiere Chris Philo (1991), esto guarda relación con las diferencias entre las personas y los lugares en términos de clase social, ideología política, grupo étnico, creencias religiosas y demás.

Pero todas estas características no significan un descenso en la oferta de prostitución o un debilitamiento de la actividad; lo único que ha variado es su distribución geográfica. Es más, el gran cambio que ha sufrido la prostitución en el siglo XX ha sido pasar de concentrarse en *distritos de luz roja* a concentrarse en áreas marginales y empobrecidas, alejadas de la sociedad dominante (Hubbard, 1999). Es entonces cuando el estigma en contra de la prostitución aumenta, ya que se relaciona con diversos tipos de criminalidad. De hecho, a menudo la prostitución se considera una causa de marginalidad y conflictividad, y no un síntoma propio de zonas desestructuradas según Hubbard.

Por un lado, tenemos los riesgos a los que están expuestos los clientes o las zonas cercanas a estos negocios; y por otro tenemos los problemas a los que se enfrentan las propias prostitutas. De entre los primeros cabe destacar el incremento de la criminalidad y el tráfico de drogas, actividades generalmente ligadas a la prostitución. Como dice Sara Bertozzi, la prostitución a menudo ha estado asociada con el incremento de la criminalidad, y con la proliferación del tráfico de drogas en los lugares en las que ésta se encuentra (2005). También destaca el riesgo de contagio de enfermedades de transmisión sexual, de gran peligro para la salud pública (Voloshin, Derevitskiy, Mukhina y Karbovskii, 2016, p. 346). Y como dicen Helderop et al., el consumo de alcohol está asociado a la práctica de sexo sin protección por dinero. Por lo tanto, mediante la elaboración de sistemas predictivos que modelen esta distribución, se pueden reubicar recursos destinados a la seguridad sexual de las trabajadoras y personas en general (Voloshin et al. 2016). Pero los problemas más importantes son sufridos por las mujeres que trabajan en el sexo. El primero es el tráfico de seres humanos <sup>6</sup>, que afecta directamente a muchas mujeres en esta profesión; también, la violencia física a las que se enfrentan en ocasiones las prostitutas atañe a problemas de salud mental, siendo los más comunes la depresión y estrés postraumático (Helde-

---

<sup>6</sup>A lo largo del trabajo de campo, al localizar los prostíbulos y buscar información en Google, muchas eran las noticias relacionadas con el tráfico de mujeres en condiciones de esclavitud.

rop et al.).

La distribución espacial de la prostitución se ha convertido en un tema importante de clarificar. Determinar los factores que influyen en su localización es de gran importancia de cara a la regulación e intervención en la actividad. En la actualidad muchos son los gobiernos que están interesados en el cese de la prostitución y de las actividades delictivas asociadas, como la explotación y el tráfico de seres humanos. Pero el sexo, como casi cualquier actividad del ser humano, se distribuye de ciertas maneras en el espacio; por poner un ejemplo, es común que la búsqueda de la satisfacción del deseo sexual se dé en clubes nocturnos, pero sería ignorada —como poco— si se hiciera en el supermercado (Hubbard, 1999, p. 37). Esto condiciona espacialmente la distribución de la actividad, de igual modo que ocurre con los burdeles.

Pero el carácter de la prostitución ha cambiado en los últimos diez años con el aumento de la importancia de internet. Muchas prostitutas utilizan esta herramienta para promocionar y ejercer su profesión, llegando así a un número más amplio de potenciales clientes a un coste más bajo. Esto supone un cambio en la distribución de la prostitución, ya que probablemente estemos ante una vuelta a la concentración en los centros de las ciudades, pero más *escondida* que anteriormente. Es un hecho que muchas prostitutas ofrecen sus servicios por internet (Hubbard, 1999, p. 124), trayendo consigo la posibilidad de cambiar de ubicación con mucha más facilidad que hasta ahora. Además, internet hace mucho más difícil detectar y comprobar las actividades ilegales; en muchos países la policía no lleva acciones en contra de los anuncios de burdeles en internet como también sugiere Philip Hubbard.

Comentados los problemas ligados a la prostitución, ¿por qué es relevante la determinación de los factores que influyen en la distribución geográfica? Pues porque de distribuirse ésta en el espacio de una determinada manera, tendríamos puntos de mayor exposición a los riesgos anteriormente mencionados, algo de gran interés para los gobiernos locales. También es posible generar modelos explicativos [y predictivos] que controlen, eliminen o alerten de este riesgo, y trayendo consigo la posibilidad de legislar de manera más concreta para eliminarlos.

Por lo tanto, no se debe ignorar el aspecto geográfico, puesto que es evidente que su distribución no se da de forma aleatoria. Como afirma Hubbard(1999, p. 33), estudios recientes en estadística espacial han empezado a demostrar que el espacio está inevitablemente sexualizado de muchas y complejas formas, trasladando los problemas relacionados con el sexo a la agenda geográfica, ya que la prostitución usualmente está concentrada en determinados puntos del espacio (Aalbers, Deinema, 2012, p. 129).

Según esta revisión los entornos de prostitución se tienden a encontrar en zonas marginales alejadas de los núcleos de población [gozando así de mayor anonimato], pero lo suficientemente cerca de éstos para disponer de la clientela necesaria. Además estos lugares presentan un estatus socioeconómico menor con tasas de delincuencia más elevadas, y con niveles de vida más bajos. Generalmente son suburbios caracterizados por precios más bajos en relación a áreas metropolitanas.

Esto los hace lugares idóneos para la localización de la prostitución. En cambio en la provincia de Pontevedra no existen entornos como éstos que destaquen de manera significativa respecto de otros lugares [característicos de grandes ciudades], como típicamente se han descrito en la literatura. Sin embargo, existen zonas que comparten gran parte de los rasgos descritos; éstas son las *áreas industriales*.

Las zonas calificadas como de uso industrial son entornos caracterizados por estar bastante alejadas de los núcleos de población, pero lo suficientemente cerca como para que no suponga un problema el desplazamiento. También se caracterizan por la afluencia de personas [mayormente hombres] pertenecientes o ajenas a trabajos de la industria, lo que permite el anonimato necesario para evitar el rechazo social. Entre estas personas encontramos a los trabajadores del propio polígono y a los que se desplazan hasta él por diversas cuestiones. Además estas zonas presentan tasas altas de población masculina que pasan un buen número de horas al día lejos de sus hogares, y con el poder adquisitivo suficiente para frecuentar locales de prostitución.

Por último, estos entornos presentan precios del suelo más baratos y con mayores facilidades para la construcción de locales o casas.

Todo esto lleva a suponer que es posible que exista algún tipo de correlación entre la distribución geográfica de los polígonos industriales y la de los prostíbulos; no porque exista relación de dependencia entre ambos, sino porque son las características del terreno industrial lo que —se supone— buscan los burdeles. Pero hay que remarcar que no se considera que los polígonos industriales sean la causa de la geografía de la prostitución, aunque es posible que exista relación.

## **Prostíbulos, variables e indicadores**

Como lo que aquí interesa es la dimensión espacial de la prostitución se trabajará con dos variables dependientes de carácter territorial:

- (1) **la distancia de los municipios al prostíbulo más cercano**

- (2) **ausencia/presencia de prostíbulos por municipio.**

La variable (1) da cuenta de 'irregularidades' en la distribución de la prostitución, ya que ante la presencia de elementos que impidan o dificulten el negocio de la prostitución, (1) aumenta, y en caso contrario disminuye. Además, esta variable permite captar los efectos directos e indirectos, estos últimos mediante la matriz de vecindad. Los municipios con actividad de prostitución han de presentar valores más bajos de (1) como efectos directos, y sus vecinos, valores más bajos como efectos indirectos. (1) permite también generar más valores al ser continua, en contraste con (2). Como indicador de (1) se considera el número en kilómetros de separación [dado un plano euclídeo de dos dimensiones] entre el centroide del polígono  $n$  y el punto [en representación de un prostíbulo] más cercano.

La variable (2) registra la presencia o ausencia de prostíbulos en el municipio. Los valores se obtienen mediante un indicador binario que mide esta presencia/ausencia, tal que; si existen uno o más puntos dentro del polígono, entonces el valor es 1; si no, 0.

Pudo haberse utilizado una variable que diese cuenta del número de prostíbulos por ayuntamiento, sin embargo tendríamos poca variación de datos; muchos de valor 0, lo que dificultaría la interpretación de mapas, así como el análisis estadístico.

Se consideran variables explicativas:

- La **renta media** por municipio, en base al promedio de renta anual bruta ingresada por persona en euros para el año 2018 para cada uno de los 62 municipios de Pontevedra .
- La **población**, que mide el número de habitantes por ayuntamiento para el año 2018.
- El **área industrial**. Ésta mide la cantidad de superficie total ocupada por la industria en kilómetros cuadrados.
- La **distancia a la carretera más cercana**, para el conjunto de carreteras nacionales y los centroides de cada polígono [del mismo modo que (1)]. Se indica —también— en kilómetros.

Respecto a las carreteras y según la literatura, éstas condicionan en gran medida la localización de la prostitución, especialmente de la prostitución 'de carretera'[la analizada en este trabajo]. Sin embargo, aquí solamente se tendrán en cuenta las carreteras nacionales por dos factores;

el tráfico y las salidas. Las carreteras nacionales son las que más tráfico tienen seguidas de autopistas y autovías, pero éstas últimas tienen muchas más limitaciones a la hora de desviarse o detenerse. Tampoco se espera que las carreteras de nivel inferior tengan un tráfico suficiente como para ejercer influencia, por lo que se descartan a pesar de que tengan también mucha facilidad de paro.

### **Fuentes de datos y casos**

El tratamiento estadístico y la generación de los mapas se hará con R a partir de los datos geográficos en formato vectorial [.shp], el cual permite calcular áreas y distancias. Éstos se obtuvieron de la base de datos online del Servizo Galego de Saúde.

Por otra parte, los datos para las variables sociodemográficas [renta, población, etc.] se obtuvieron de la base de datos del portal digital del Instituto Galego de Estatística.

Los emplazamientos de los prostíbulos se consiguieron mediante un extenso trabajo de campo. Dada la condición de estos negocios como ilegales es difícil el acceso a las ubicaciones de cada establecimiento. Parte de ellos se obtuvieron buscando en motores de búsqueda comunes [Google y GoogleMaps]; en listados de negocios como Páxinas Galegas y Páginas Amarillas principalmente; en preguntas a personas relacionadas con los distintos ayuntamientos; y por último, buscando noticias en medios digitales, como periódicos y blogs. Posteriormente se contrasta la noticia con su ubicación en longitud y latitud con alguna base de datos cartográficos. Las más comunes y usadas fueron GoogleMaps y OpenStreetMap. Pero el grueso de prostíbulos se localizaron tras un largo trabajo de campo recorriendo las carreteras de la provincia, lo que proporcionó un listado de 29 prostíbulos con su ubicación geográfica exacta en longitud y latitud. Posteriormente se convirtieron —mediante un conversor de coordenadas— a un sistema de coordenadas en dos dimensiones, el sistema universal transversal de Mercator.

Por las características del fenómeno [y los recursos disponibles] no fue posible realizar algún tipo de muestreo probabilístico con garantías de representación estadística, de modo que el muestreo es de tipo no probabilístico.

Los casos seleccionados son el conjunto de los concellos que conforman la provincia de Pontevedra, con un total de 62. Se trabaja con un listado completo de los ayuntamientos por la facilidad y conveniencia de trabajar con la población. El conjunto de bases de datos utilizan

el ayuntamiento como unidad de agregación mínima, siendo casi imposible encontrar listas de datos completos y fiables para unidades territoriales más pequeñas. Además, trabajar con unidades más pequeñas generaría problemas al relacionar variables, puesto que es probable que las relaciones que se esperan encontrar no se den en niveles de agregación inferiores por la excesiva fragmentación del terreno. Realizar el trabajo con una muestra completa de concellos es tarea sencilla y adecuada en este caso. Tenemos un total de 62 subdivisiones, con un valor para cada variable correspondiente con cada una de estas.

## Técnicas de análisis

Como las variables tienen un componente espacial importante, la primera tarea es realizar pruebas de autocorrelación espacial con el fin de determinar estadísticamente si éstas se distribuyen de manera aleatoria o no. La autocorrelación espacial es algo muy común en geografía; se da cuando los objetos son similares a otros objetos cercanos. Para el caso de las variables, decimos que existe autocorrelación espacial cuando los valores de una variable se correlacionan con los valores promedio de sus vecinos. En caso de distribución geográfica aleatoria, la autocorrelación espacial sería muy baja o —idealmente— inexistente; si la distribución no fuese aleatoria cabría esperar la existencia de relaciones entre los valores de los objetos y los valores de los objetos vecinos, tanto de signo positivo como negativo.

Las principales pruebas para determinar correlación en el espacio son la I de Moran junto con diagramas de dispersión con ponderación espacial<sup>7</sup> para variables de razón, y técnica de *Join Count* para variables cualitativas. La primera es similar al coeficiente de correlación de Pearson, variando entre -1 y 1; en caso de un coeficiente de -1 tendríamos una correlación espacial perfecta negativa, de tal manera que valores altos en un punto del espacio se rodearán de valores bajos en sus cercanías; del modo contrario, con un coeficiente de 1, los valores altos de un punto del espacio estarán rodeados de valores altos en el conjunto de vecinos; un coeficiente de Moran de 0 indica ausencia de relación espacial. La segunda se utiliza en variables binarias para la determinación de relaciones espaciales entre elementos que puedan tomar solamente dos valores, generalmente ausencia y presencia del fenómeno. Mediante el análisis por pares de todos los objetos se comparan los casos coincidentes y no coincidentes con los que cabría esperar dada una distribución aleatoria, y mediante simulación se repite este proceso  $n$  veces para completar el test. Como resultado obtenemos evidencia estadística sobre la aleatoriedad o no aleatoriedad

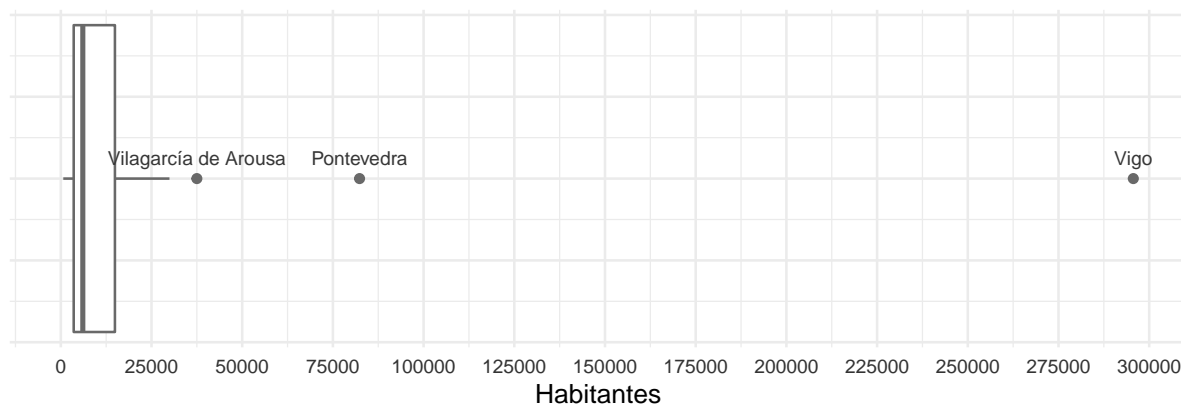
---

<sup>7</sup>Moran Scatter Plot

del fenómeno.

La matriz utilizada será del tipo explicado en la primera parte del trabajo; una matriz de orden superior variable que tendrá un umbral de vecindad en 25.000 habitantes. Se pensó así para dar mayor *peso* a los ayuntamientos de Vigo, Pontevedra y Vilagarcía de Arousa, ya que éstos no afectan únicamente a sus vecinos de primer orden, sino que afectan —como poco— a los de segundo por su gran población. Esto permite estimar las relaciones entre las variables de mejor manera por tener en cuenta las grandes diferencias en población, uno de los factores que más influye en su entorno. Para una idea general de las grandes diferencias en población tenemos el siguiente gráfico de caja:

Figura 2  
Dispersión de la población de los municipios en Pontevedra



Nota: Elaboración propia

La diferencia entre Vigo, Pontevedra y Vilagarcía con el resto de municipios es considerable; por ello la influencia en su entorno ha de ser mayor en comparación con los más pequeños. En cambio, esta influencia se considera residual —o inexistente— en municipios situados debajo de este umbral.

Por último se realizarán tests de calidad para determinar qué modelo es el que mejor se acerca al modelo *real* de generación de datos mediante el criterio de información de Akaike [AIC]. Este estadístico mide la calidad del ajuste entre el modelo a estimar y el modelo *real* que genera los datos, y nos permite la comparación de la 'calidad' entre distintos modelos de igual forma. Sin entrar en detalles, diremos que cuánto más bajo sea el AIC de un modelo, mejor.

## Regresión espacial

La regresión espacial emplea la matriz de vecindad para controlar los efectos de la autocorrelación en el espacio. Por ejemplo, la riqueza de un municipio influye en la riqueza de los de su entorno, y a su vez este entorno influye en el propio municipio. Estos modelos permiten *controlar* y medir los efectos del entorno a través de cualquiera de sus variables, llamados en ocasiones *efectos indirectos*. De utilizar una estimación mediante MCO para un caso como el anterior estaríamos ignorando elementos de difusión espacial, lo que nos llevaría a estimadores sesgados e inconsistentes (Darmofal, 2015).

Los siguientes modelos consideran este fenómeno tanto en la variable dependiente como en el resto de sus elementos [variables independientes o residuos]. Si se asume, por ejemplo, que los residuos se distribuyen normalmente pero en realidad están autocorrelacionados en el espacio, tendríamos los problemas de estimación anteriormente mencionados. Por esto es especialmente útil la matriz de vecindad junto con modelos espaciales, ya que permite la ponderación de los valores de la variable cercanos en el espacio, y de este modo cuantificar el efecto que ejerce el entorno. En muchos de los procesos sociales el entorno más cercano influye considerablemente en lo particular, y a su vez lo particular influye en su entorno. La regresión espacial por un lado controla esta influencia, y por otro estima sus efectos, lo que nos permite generar modelos más ajustados a la realidad y determinar cuánta influencia ejerce lo que nos rodea.

## Modelos

En este trabajo se tendrán en cuenta los siguientes modelos, explicados a partir de un modelo de mínimos cuadrados ordinarios para facilitar las explicaciones para los menos familiarizados:

- Siendo el modelo lineal por MCO:  $y = \beta_0 + \beta_1x + \beta_2x \dots \beta_nx + \varepsilon$ , se utilizarán los siguientes modelos de tipo espacial.
- Retardo espacial  $LM_{lag}$ , tal que:  $y = \rho W y + \beta_nx + \varepsilon$ , en el cual  $W$  corresponde a una matriz de vecindad y  $\rho$  su coeficiente. Este modelo es utilizado para controlar la autocorrelación en el espacio de la variable dependiente, incluida en el modelo como variable independiente multiplicada por la matriz de vecindad. Se utiliza cuando la variable dependiente influye en sí misma a través de puntos cercanos en el espacio.
- Modelo de retardo en los residuos  $LM_{err}$ :  $y = \beta_nx + \varepsilon$ , donde  $\varepsilon = \lambda W \varepsilon + \zeta$ , siendo  $\lambda$  el

parámetro de la correlación espacial en los residuos. Se utiliza cuando los valores de los residuos del modelo se correlacionan entre sí en puntos cercanos del espacio. A través de la utilización de la matriz, es posible controlar la relación espacial de los residuos.

- Modelo error de Durbin  $LM_D$ :  $y = \beta_n x + \rho Wx + \varepsilon$ , dónde —igual que en el anterior modelo—  $\varepsilon = \lambda W\varepsilon + \zeta$  y se incorpora ponderada una de las variables independientes. Es necesario utilizarlo cuando las variables independientes y los residuos del modelo presentan autocorrelación en el espacio.
- El último modelo espacial es el SARMA<sup>8</sup> que combina algunos de los elementos de los modelos anteriores. En particular, el que se espera usar es:  $y = \beta_n x + \varphi W y + \xi W x + \lambda W \varepsilon + \zeta$ . De forma simplificada, se combinan las ponderaciones de vecindad de la variable dependiente, independientes y los residuos, siendo  $\beta$  los efectos de las variables independientes,  $\varphi$  los efectos del proceso autorregresivo de la variable dependiente,  $\xi$  los efectos del proceso autorregresivo de la variables independientes y  $\lambda$  los efectos del proceso autorregresivo de los residuos.
- Para el conjunto de hipótesis de variable independiente categórica (2) se utilizará una regresión logística de un modelo generalizado lineal (GLM), combinándolo —si es necesario— con las formas de regresión anteriormente descritas. La interpretación de la probabilidad de ocurrencia será:  $y = \frac{\exp(\beta_0 + \beta_1 x \dots \beta_n x)}{1 + \exp(\beta_0 + \beta_1 x \dots \beta_n x)}$

## Hipótesis

Las hipótesis se muestran agrupadas en función de las variables dependientes (1) y (2) para cada municipio.

---

<sup>8</sup>Spatial Autorregressive Moving Average Model

( $H1_{(1)}$ ). A mayor área industrial, menor distancia al prostíbulo más cercano.

( $H2_{(1)}$ ). A mayor población, mayor distancia al prostíbulo más cercano.

( $H3_{(1)}$ ). A menor distancia a carreteras, menor será la distancia al prostíbulo más cercano.

( $H4_{(1)}$ ). A mayor renta media por municipio, mayor será la distancia al prostíbulo más cercano.

( $H1_{(2)}$ ). A mayor área industrial, mayor probabilidad de presencia de prostíbulos.

( $H2_{(2)}$ ). A mayor población, mayor probabilidad de presencia de prostíbulos.

( $H3_{(2)}$ ). A menor distancia a carreteras, mayor probabilidad de presencia de prostíbulos.

( $H4_{(2)}$ ). A mayor renta media por ayuntamiento, menor probabilidad de presencia de prostíbulos

## Exploración de datos

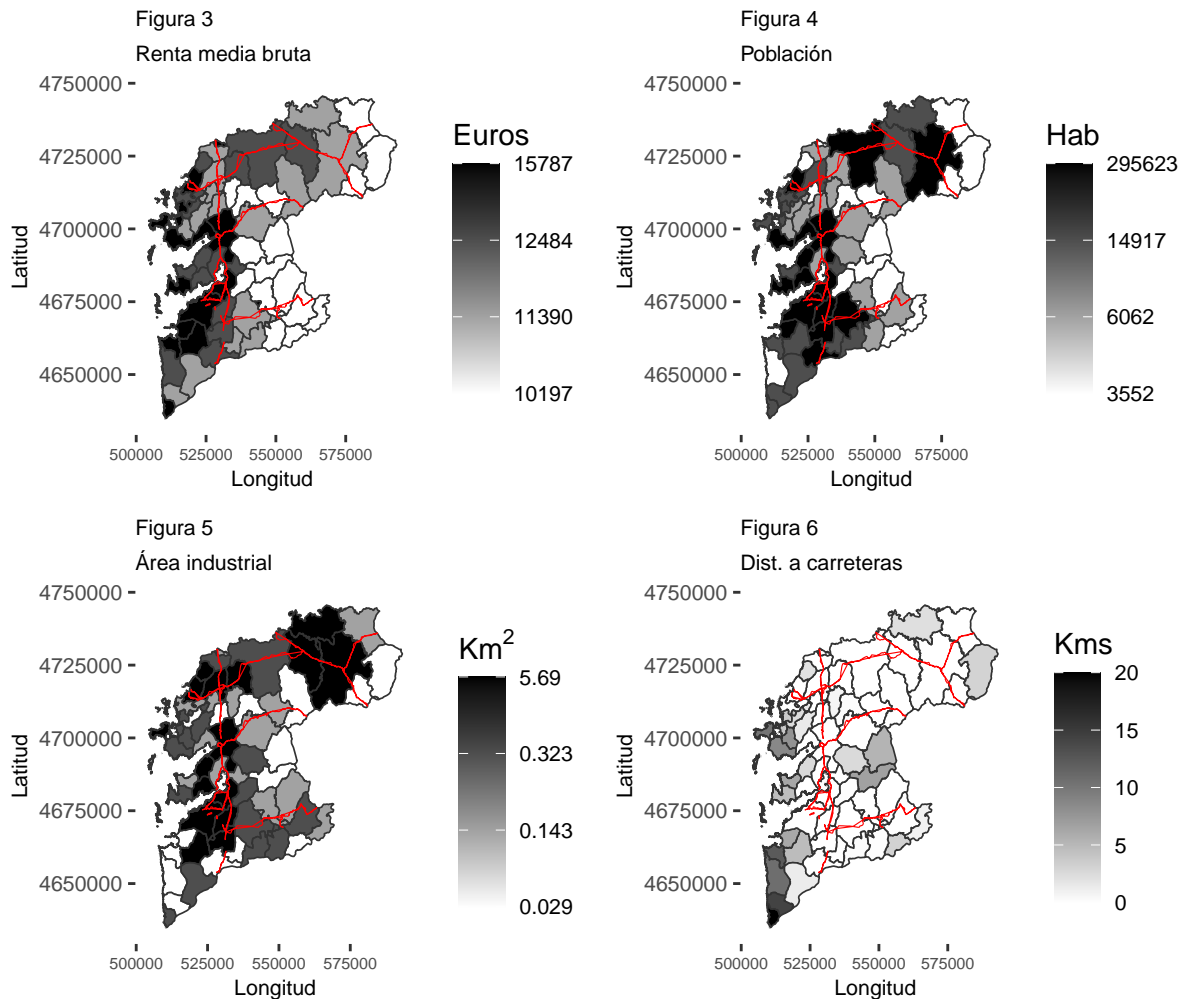
En estadística espacial, la exploración de los datos es algo que se da de manera *natural* en el análisis de datos (Haining, 2003), y se necesita para detectar —si hay— errores en los datos. Además, si bien quien escribe el trabajo conoce en profundidad las características de los procesos de los que habla, es probable que quién lo lea no lo haga y agradezca la inclusión de algún mapa para hacerse una idea.

El siguiente mapeo utiliza el sistema de coordenadas Universal Transversal de Mercator [UTM], basado en la proyección geográfica transversa de Mercator. En un principio se utilizó un sistema de proyección GPS, con longitudes y latitudes en grados decimales, pero el cálculo de las distancias era mucho más complejo e impreciso, ya que con este método no solo se tienen en cuenta longitud y latitud, sino que se incorpora un tercer factor [haciendo del mapa un objeto tridimensional] con un eje de altura. El sistema UTM es adecuado para calcular áreas y distancias, tanto por la sencillez de cálculo como de interpretación de mapas.

## Mapeo

### Variables independientes

De esta forma se disponen los siguientes mapas seguidos de una —muy— breve descripción, con un valor para cada polígono [concello] y un total de 62 unidades:



El conjunto de figuras sugiere cierto patrón similar de distribución. Hay una clara zona de influencia en el eje atlántico que modifica la distribución de las variables. Además los valores parecen estar concentrados en cúmulos, posiblemente por efectos de autocorrelación espacial. Las regiones tienden a parecerse entre ellas, tanto en valores altos como en valores bajos de las variables. Pero aunque las variables se asemejen entre sí en su distribución, existen muchas diferencias entre ellas como para tratarlas de manera conjunta. En la siguiente tabla lo vemos más claro:

### Descriptivos de las variables

Variable	N	Media	Desviación	Mediana	Min	Max
Área industrial	62	0,43	0,986	0,14	0	6
Dist. carreteras	62	2,47	4.147	0,11	0	20
Población	62	15.401	38.166	6062	678	295.623
Renta	62	11.334	1.934	11390	7.429	15.787

Cuadro 1: Elaboración propia

La renta parece ser la variable que menos dispersión tiene. Con un coeficiente de variación del 17% se sitúa bastante lejos de las dispersiones del resto de variables. Recordamos que se trata de una variable agregada, que mide la renta media de los municipios, por lo que tampoco cabe esperar gran dispersión. La renta y la mediana coinciden casi totalmente, lo que es propio de una distribución simétrica. Aun así existen diferencias considerables entre los municipios de mayor y menor renta, con un valor en diferencia de 8300 euros.

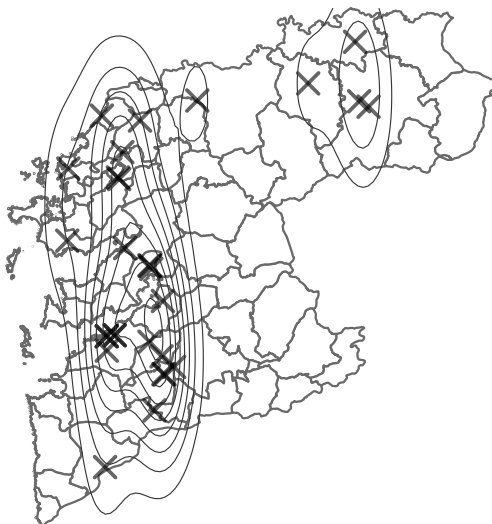
La población, el área industrial y la distancia a las carreteras son variables de asimetría positiva; sus medianas son más bajas que sus medias, indicando mayor concentración de casos en valores más bajos. Sin embargo todas tienen una alta dispersión como resultado de las grandes diferencias entre municipios, pero como indican los mapas, estas diferencias no parecen estar distribuidas de manera homogénea por todo el territorio, sino que se concentran en determinadas zonas de la provincia.

### **Variables dependientes**

En esta parte se representarán mediante mapas y puntos la distribución espacial de los prostíbulos y las variables dependientes operacionalizadas de diferente manera.

Comenzaremos por localizar los burdeles:

Figura 7  
Distribución de prostíbulos con curvas de densidad



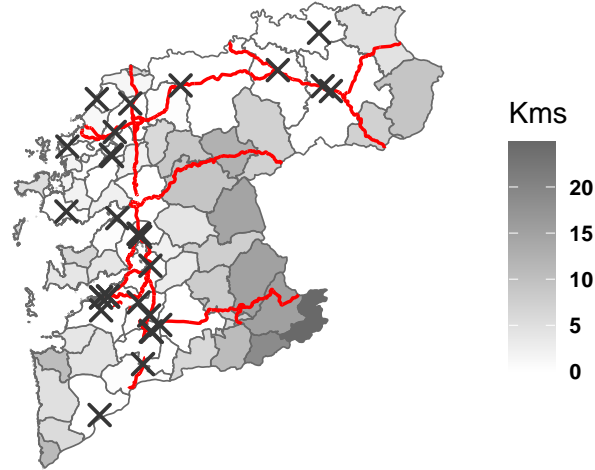
Nota: Elaboración propia

Las curvas de color oscuro se corresponden con la densidad de la prostitución. La mayor parte de locales están ubicados en la zona occidental de la provincia. Existe un pequeño grupo concentrado en la región noroeste, en las proximidades de Lalín.

Aunque no se pueda extraer —aún— conclusiones, podemos percibir una distribución no homogénea en el territorio, que sugiere la existencia de elementos que la condicionen. Sin embargo, uno de los problemas de la operacionalización de esta variable en número de prostíbulos por ayuntamiento —como ya se comentó— está en la marcada agrupación de los datos reflejada en el mapa y en la existencia de muchos polígonos de valor 0. Para evitar esto se trabaja con la variable (1), que se trata de la distancia de cada polígono al prostíbulo más cercano. Así conseguimos la variación en cada zona a excepción de los que tienen prostíbulos, que serán en este caso los que tengan valor 0.

Representamos (1) en el siguiente mapa:

Figura 8  
Distancia en km al punto de carretera más cercano

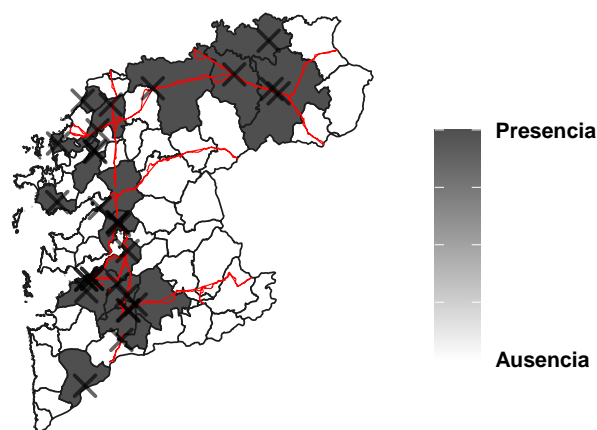


Nota: Elaboración propia

Los polígonos más oscuros se corresponden con distancias más altas tal y como se dijo. La distancia máxima es de aproximadamente 25 kilómetros, y se corresponde con Crecente, pero como se ve a simple vista, los valores disminuyen a medida que nos acercamos al este y al norte. Esto sugiere presencia de autocorrelación espacial.

La variable binaria (2) genera el siguiente mapa:

Figura 9  
Existencia de uno o más prostíbulos por municipio y carreteras nac



Nota: Elaboración propia

A excepción de dos ayuntamientos, todos los que tienen algún prostíbulo limitan al menos con otro ayuntamiento que también tiene, lo que indica agrupación de burdeles. Pero de nada sirve

la simple observación de los mapas a fin de sacar conclusiones, por lo que determinaremos si las variables están significativamente agrupadas o diseminadas en el espacio mediante las pruebas correspondientes.

## Test

En la siguiente tabla se muestran los estadísticos de Moran para cada variable independiente:

I de Moran para v. independientes		
Variable	Observado	Valor p.
Área industrial	0,185	0,00***
Dist. carreteras	0,597	0,00***
Población	0.036	0,20
Renta	0,497	0,00***

*Significación:* 0 '\*\*\*' 0,001 '\*\*' 0,05 '\*' 0,1 '.'

Cuadro 2: Elaboración propia

Todos presentan autocorelación espacial en mayor o menor medida, a excepción de la población con valor  $p$  no significativo.

Para las variables dependientes utilizaremos la I de Moran y el estadístico *Join count* para datos categóricos. La I de Moran para (1) es:

I de Moran para la v. dependiente (1)		
	Observado	Valor p.
Dist. prostíbulos	0,565	0,00***

*Significación:* 0 '\*\*\*' 0,001 '\*\*' 0,05 '\*' 0,1 '.'

Cuadro 3: Elaboración propia

(1) se relaciona consigo misma a través de los valores de su entorno, pues existe una considerable correlación espacial positiva de 0.56.

La siguiente prueba es el Join Count para variables binarias:

Join Count para presencia en v. dependiente (2)

Coincidentes	Esperados	Varianza
17	11,37	8,57
Estadístico	1.92	
Valor p.	0.027**	

*Significación:* 0 '\*\*\*\*' 0,001 '\*\*' 0,05 '\*' 0,1 '.'

Cuadro 4: Elaboración propia

Interpretando los resultados para los polígonos de color negro tenemos que, dada una distribución aleatoria se esperaría que el número de casos coincidentes fuera cercano a 11; sin embargo el número obtenido es de 17, con un valor p 0,027 y rechazando así la hipótesis nula de no autocorrelación.

Para el caso de los polígonos de color blanco [para determinar si se autocorelacionan las zonas sin prostíbulos] tenemos valores no significativos, mostrados en la siguiente tabla:

Join Count para ausencia en la v. dependiente (2)

Coincidentes	Esperados	Varianza
59	70,3	24,5
Estadístico	-2.28	
Valor p.	0.988	

*Significación:* 0 '\*\*\*\*' 0,001 '\*\*' 0,05 '\*' 0,1 '.'

Cuadro 5: Elaboración propia

No existe relación entre ayuntamientos sin prostitución y sus vecinos para una matriz de vecindad como la anterior.

## Elección del modelo más ajustado mediante AIC

Todas las variables menos una se relacionan consigo mismas a través de los valores de sus respectivos vecindarios, por lo que la utilización de modelos *simples* que contemplen la autocorrelación únicamente en los residuos de los modelos, en las variables dependientes o en las

variables independientes, seguramente dé lugar a modelos poco ajustados. Utilizando modelos *mixtos* que contemplen estas relaciones [modelos Durbin, SAC<sup>9</sup>, SMA<sup>10</sup> o SARMA] podremos aumentar la calidad de nuestras estimaciones.

Pero antes de estimar el AIC se deben hacer contrastes de hipótesis para determinar si realmente se debe utilizar modelos de tipo espacial, ya que este estadístico no contrasta ninguna hipótesis. El más conocido y simple es el test del Multiplicador de Lagrange<sup>11</sup> para contrastar la dependencia espacial tanto en los residuos como en las variables o ambas, que se realiza a continuación para cada conjunto de hipótesis:

Test multiplicador de Lagrange para (1)

Tipo	Estadístico	Valor p.
LMerr	16,52	0.0000480326***
LMlag	25,61	0.0000004173***
RLMerr	0.61	0.431647
RLMlag	9.71	0.001836**
SARMA	26.23	0.0000020135***

*Significación:* 0 '\*\*\*' 0,001 '\*\*' 0,05 '\*' 0,1 '.'

Cuadro 6: Elaboración propia

Los contrastes para dependencia en los errores y en las variables son significativos para el primer conjunto, no así sus correspondientes robustos. Estos últimos indican un mejor ajuste si se utilizasen modelos que tuviesen en cuenta únicamente la dependencia espacial en las variables y no únicamente en sus errores. Según el test, parece también acertado utilizar modelos SARMA, ya que contemplan ambas situaciones. Frente a la significación de varios tipos de modelos, la elección se hará atendiendo al AIC más bajo.

Para el modelo logístico del segundo conjunto de hipótesis tenemos:

<sup>9</sup>Spatial Autoregressive Combined model del paquete 'spatialreg'.

<sup>10</sup>Spatial Moving Average model, más simple que los modelos SARMA, también dentro de 'spatialreg'.

<sup>11</sup>Lagrange Multiplier Test

Test multiplicador de Lagrange para (2)

Tipo	Estadístico	Valor p.
LMerr	3,36	0.067 .
LMlag	-1,40	1
RLMerr	4,67	0.03*
RLMlag	-0,085	1
SARMA	3,27	0,19

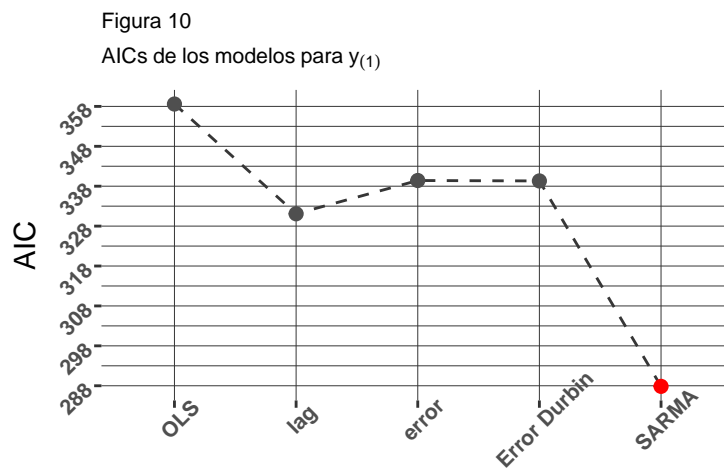
Significación: 0 '\*\*\*\*', 0,001 '\*\*', 0,05 '\*' 0,1 '.'

Cuadro 7: Elaboración propia

En este caso también parece adecuada la consideración de los efectos espaciales en el modelo. En concreto la prueba sugiere un modelo que contemple la autocorrelación únicamente en los residuos.

Pero el test omite el tipo de error de Durbin. Hasta aquí solo se quiere justificar el uso de modelos espaciales.

Una vez hecho esto, se calculan los AICs correspondientes para los modelos descritos en secciones anteriores, ahora sí, para la elección del modelo:

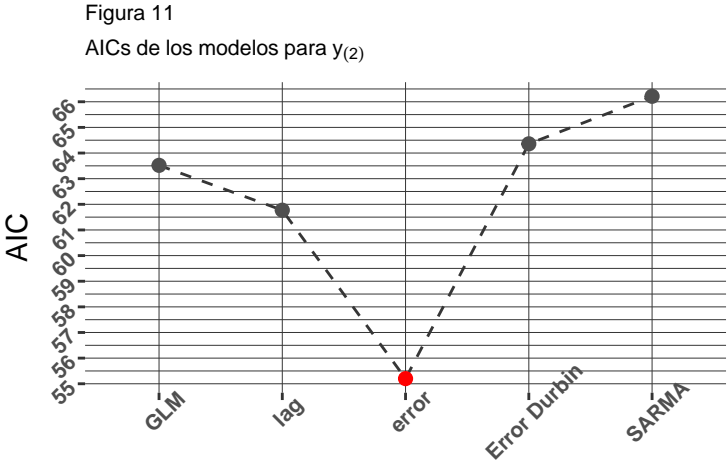


Nota: Elaboración propia

Para el caso de (1), el mejor modelo es el SARMA. Recordemos que, en términos generales, menor el AIC mejor el ajuste del modelo. El AIC permite comparar la calidad de modelos de distinto tipo, tanto anidados como no anidados, estimando la *información perdida* en el proceso

de modelado [respecto del proceso real de generación de los datos]. Valores más bajos indican menos pérdida de información.

Para (2) tenemos lo siguiente:



Nota: Elaboración propia

Para este caso y, al cambiar de método de estimación a una regresión logística con una variable dependiente categórica, parece ser que el modelo mejor ajustado es el que contempla los efectos espaciales únicamente en los residuos.

## Estimación del modelo

Teniendo en cuenta las características de las variables se harán dos modelos regresión para cada conjunto de hipótesis; un modelo SARMA para (1) y un modelo logístico con el control espacial de sus residuos:

(1)

$$y = \beta_n x_n + \varphi W y + \xi W x_n + \lambda W \varepsilon + \zeta$$
$$d.prostibulos \sim M * d.prostibulos + M * industria + M * d.carreteras + M * poblacion + M * renta + industria + d.carrreteras + poblacion + renta + M * residuos$$

(2)

$$p_y = \frac{\exp(\beta_0 + \beta_1 x \dots \beta_n x)}{1 + \exp(\beta_0 + \beta_1 x \dots \beta_n x)}$$
$$bin.prostibulos \sim industria + d.carreteras + poblacion + renta + M * residuos$$

En ambos casos  $M$  es la misma matriz de ponderaciones de vecinos.

### Estimación de los modelos de regresión

	<i>Modelo:</i>	
	Distancia prostíbulo (1)	Presencia/ausencia (2)
	<i>SARMA</i>	<i>Logit</i>
industria	-1.391** (0.541)	4.042** (1.679)
d.carreteras	0.534*** (0.134)	-0.401* (0.229)
poblacion	0.0001*** (0.00002)	-0.00001 (0.0001)
renta	-0.0004* (0.0003)	0.001 (0.0003)
$M_{ij}$ * d.prostibulos	1.143*** (0.119)	
$M_{ij}$ * industria	-4.519*** (1.383)	
$M_{ij}$ * d.carreteras	-0.677*** (0.212)	
$M_{ij}$ * poblacion	0.0002*** (0.00004)	
$M_{ij}$ * renta	-0.0002 (0.0002)	
$M_{ij}$ * residuos	-1.476*** (0.255)	-1.270*** (0.472)
Observations	62	62
Adjusted R <sup>2</sup>	0.828	
Log Likelihood		-21.600
Residual Std. Error	2.240 (df = 51)	
F Statistic	30.449*** (df = 10; 51)	
<i>Significación:</i>	0 '***' 0,001 '**' 0,05 '*' 0,1 '.'	

Cuadro 8: Elaboración propia

Primero vamos con los efectos directos del modelo (1) en la descripción de los resultados de los modelos. Todos los coeficientes de efecto directo son significativos a excepción de la renta. Tenemos que: a mayor área industrial por ayuntamiento, menor distancia existe entre éste y el prostíbulo más cercano [(1)]; la distancia a la carretera más cercana está relacionada con signo positivo con (1); según aumenta la población aumenta (1), dándose una relación similar a la planteada en la hipótesis  $H2_{(1)}$ ; la renta también parece tener un efecto distinto al planteado en las hipótesis, ya que —curiosamente— a mayor renta mayor proximidad de prostíbulos, sin embargo sin significación estadística al 95 %.

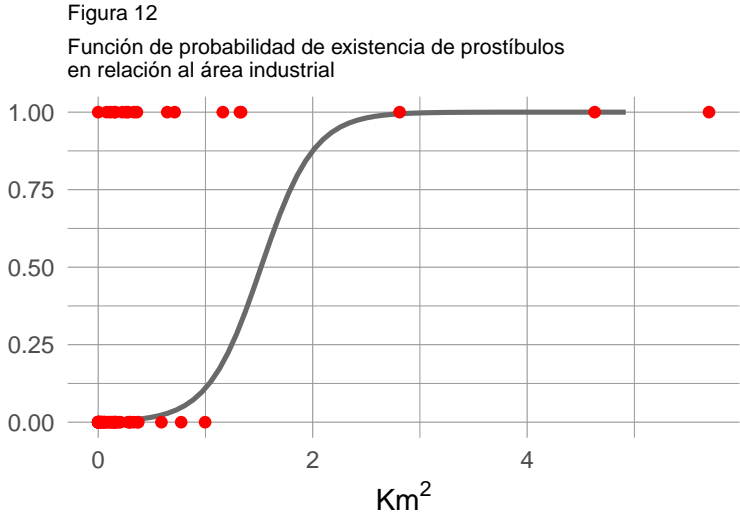
Existen grandes diferencias entre los valores de los estimadores fruto de las distintas unidades de medidas de las variables independientes y de sus rangos, aunque una interpretación estricta de los coeficientes no tiene sentido. De hacerlo así tendríamos que: al aumentar en  $1km^2$  el área industrial de un ayuntamiento, la distancia al prostíbulo más cercano se reduce alrededor de  $1,39kms$ ; al aumentar en  $1km$  la distancia a la carretera más cercana, (1) aumenta en  $0,5kms$ ; al aumentar en una unidad la población, (1) aumenta en  $0,000052kms$  o, lo que es lo mismo,  $5,2cms$ ; por último, al aumentar la renta media en 1 la distancia al prostíbulo más cercano disminuye en  $44cms$ . Pero como el modelo no se ajusta de manera perfecta, una interpretación como la anterior serviría de bien poco.

Como es de esperar la variable dependiente se relaciona consigo misma en el espacio, esto es, la distancia al prostíbulo del municipio desciende conforme desciende el promedio de distancias de los vecinos; la industria tiene un efecto más fuerte de manera indirecta de igual signo. Parece que la influencia de la industria *vecina* ejerce más influencia en el propio ayuntamiento. Esto puede deberse a que la matriz de vecindad pondera los efectos de los polígonos vecinos, y cada polígono tiene un promedio de 5 vecinos aproximadamente, lo que *agrand*a los efectos; la influencia de la renta no es significativa, pero la relación de efectos directos e indirectos podría sugerir que la distancia a los prostíbulos es menor en sitios de renta más alta; los efectos indirectos de la población son de igual signo que los indirectos, pero de mayor magnitud. Si el promedio de población de los vecinos de un municipio aumenta, aumentará la distancia del propio municipio al prostíbulo más cercano; y por último tenemos la curiosa relación inversa entre efectos indirectos y directos de la distancia a las carreteras. Mientras que como efecto directo tenemos que si aumenta la distancia a las carreteras también aumenta la distancia al prostíbulo, como efecto indirecto tenemos que si aumenta el promedio de la distancia de los vecinos a las carreteras, disminuye la distancia a prostíbulos del municipio. Esto posiblemente

esté explicado por la naturaleza de la geografía de las distribuciones de las variables. Mientras los prostíbulos se dan de manera agrupada en el espacio, las carreteras ocupan un área mayor en el mapa. Esto aumenta la probabilidad de que, en conjunto, los ayuntamientos se sitúen más cerca de una carretera que de un prostíbulo. Pero cuando trabajamos con un ayuntamiento singular, esta relación no se encuentra, ya que solamente se centra en un único polígono.

La matriz de vecindad multiplicada por los residuos solamente se utiliza como control de la autocorrelación espacial que éstos presentan, por lo que su interpretación no tiene demasiado interés en este trabajo.

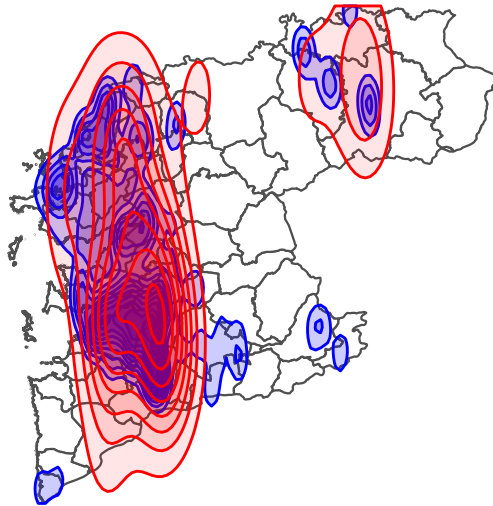
El segundo modelo (2) trabaja con menos variables. La interpretación de los residuos se hace de la misma manera que en el caso anterior; como forma de control. Las variables tienen significación más baja, pareciendo que lo único relevante en este modelo es el área industrial. Pero la interpretación de los coeficientes de esta regresión logística es un poco más compleja. El signo es igual que en el anterior modelo; el área industrial aumenta la probabilidad —significativamente— de existencia de prostíbulos en el municipio. Los coeficientes expresan como logaritmo de las razones de la probabilidad de que ocurra *y* al aumentar, en este caso, el área industrial de un municipio. De manera simplificada, al aumentar el área industrial, la probabilidad de presencia de prostíbulos es cercana a una razón de 4. Para conocer las probabilidades predichas de ocurrencia hay que *transformar* los coeficientes mediante lo explicado en el apartado de técnicas de análisis. Si lo hacemos con la variable del área industrial [la única significativa en el modelo (2)] tenemos la siguiente gráfica de distribución de probabilidad para el modelo:



Nota: Elaboración propia

La probabilidad de existencia de prostíbulos aumenta considerablemente a partir de  $1km^2$ . Los puntos y la línea gris es la función de probabilidad, mientras que los puntos en rojo son los casos existentes de prostíbulos en relación al terreno industrial. A partir del kilómetro cuadrado no hay locales de prostitución en esta provincia, y como indica esta última regresión, el área industrial es la única que muestra significación estadística [una vez se controlan los residuos]. La gran influencia del suelo industrial en la prostitución surge de sus —casi— idénticas distribuciones. Esto se muestra a continuación:

Figura 13  
Curvas de densidad de prostíbulos [rojo] y  
área industrial [azul]



Nota: Elaboración propia

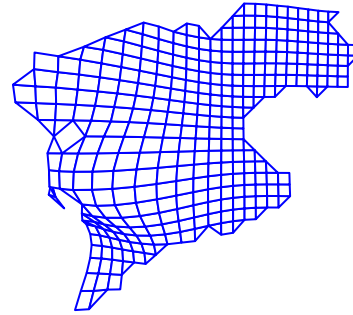
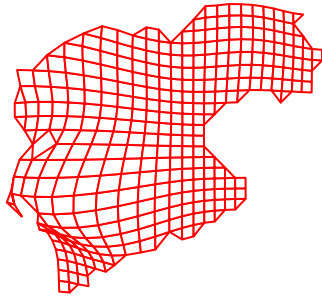
Las curvas rojas corresponden con la densidad de la industria; las azules con la de los prostíbulos. A excepción de algunas zonas industriales al sur y al sureste, ambas distribuciones son especialmente parecidas.

Desde otra perspectiva podemos dibujar un mapa en tres dimensiones de la provincia; los ejes  $x$  y  $y$  son los mismos que en el caso anterior, pero se añade un tercero como densidad:

Mapa 3D de densidades

**Dens. prost. Pontevedra 90°**

**Dens. indust. Pontevedra 90°**



**Dens. prost. Pontevedra 0°**

**Dens. indust. Pontevedra 0°**

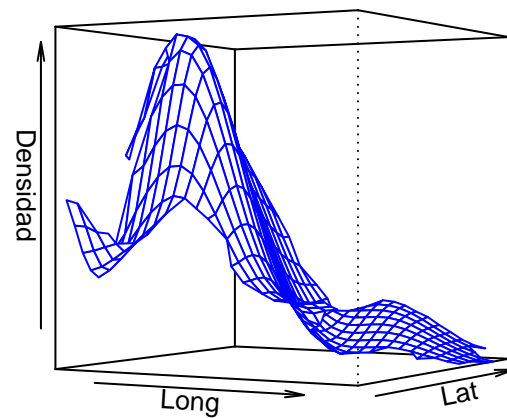
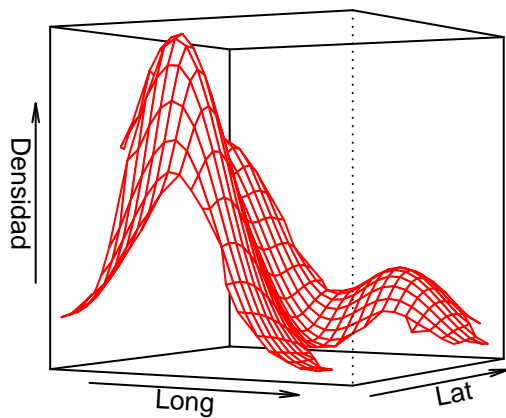


Figura 14: Elaboración propia

Los mapas se corresponden entre sí en columnas, variando únicamente el ángulo de visión. Tal y como sugiere la figura 13, ambas distribuciones se asemejan especialmente; la alta significación de los coeficientes del modelo para el área industrial posiblemente se deba a sus muy parecidas distribuciones espaciales.

De acuerdo con una explicación en términos *sustantivos*, las relaciones entre variables encontradas, algunas aún distintas de las planteadas en las hipótesis, son fácilmente entendibles. En primer lugar tenemos la curiosa relación entre zonas industriales y prostitución. Ésta podría

venir explicada por la alta tasa de población masculina dentro de un entorno que favorece la anonimidad. No es difícil imaginarse a un grupo de trabajadores en busca de *ocio* después de una jornada de trabajo. Las facilidades de edificación y de apertura de negocios también son elementos que favorecen el negocio; estos terrenos suelen estar lejos de núcleos urbanos, y por lo tanto tienen precios más bajos. También existe una amplia oferta de locales *reconvertibles* en prostíbulos, y además, están lejos de ocasionar posibles problemas a vecinos, ya que en estas zonas la densidad de población y de viviendas es considerablemente más baja.

La distancia a las carreteras también parece ser un elemento condicionante. Esto poco desarrollo necesita; es evidente que en zonas de poco tránsito o lejos de las principales vías de comunicación no habrá cerca prostíbulos. Como se explicó, las normas que rigen hoy en día la prostitución de carretera son en exclusivas del mercado, y si no hay tráfico de clientes, simplemente no habrá negocios.

La interpretación de los efectos poblacionales puede ser más problemática. No parece tener demasiado sentido que, por el simple hecho de haber mucha población en un ayuntamiento, aumente significativamente la distancia a los prostíbulos, ya que gente es justamente lo que se busca en este negocio. Lo probable es que, en ayuntamientos de gran población ésta se dé de manera concentrada en ciudades. La *ciudad* genera muchos problemas para los prostíbulos, sobre todo los de carretera. Por una parte tenemos una vigilancia continua, tanto por parte de la gente como de la policía. Por otro lado, el anonimato en zonas densamente pobladas es mucho menor, y supondría un buen freno para clientes y 'emprendedores' del negocio.

La renta no parece tener ningún efecto en ninguno de los modelos, por lo que —en principio— tratar de explicar sus efectos parece no servir de mucho.

## Conclusiones

Según lo visto, la distribución de prostitución en Pontevedra parece estar lejos de la ausencia de condicionantes que le afecten. Las causas de la estrecha relación entre la prostitución y las zonas industriales sería algo interesante de determinar y, aunque no parezca algo extremadamente raro, sí es algo, cuanto menos, curioso. Posiblemente existan nuevos elementos que condicionan la distribución de la prostitución en comparación a otros lugares o tiempos, o sencillamente estos condicionantes sean ya conocidos y afecten de distinta manera en esta región.

En cualquier caso no se debería ignorar el carácter espacial del fenómeno. Como ocurre muchas

veces en sociología, la cercanía provoca que las relaciones entre los actores sean más intensas. Éste es uno de esos casos, pues las regiones más cercanas muestran mayor similitud de características, y a medida que nos alejamos se debilita la fuerza de las relaciones. Si se hubiese hecho este trabajo mediante técnicas *clásicas* de estimación tendríamos aproximaciones sesgadas e inconsistentes que ignoran una parte fundamental. Un ejemplo de esto lo tendríamos en las explicaciones típicas de condicionantes del voto [riqueza, estatus social, estudios del padre/madre, etc]. Aunque todas estas variables influyan de manera significativa, también es evidente que el contexto también lo hace. El cambio de país, de ciudad o hasta de barrio pueden provocar cambios en la ideología de las personas, y seguramente también en la tendencia de voto. Por esto tenemos concentración de zonas de determinado color político en ciertas regiones que contrastan notablemente con otras, como es el caso de las zonas rurales y urbanas y sus *concretas* distribuciones geográficas.

La estadística espacial puede utilizarse para conocer mejor los efectos de la cercanía en una sociedad y corregir de manera localizada parte de sus problemas. Estos resultados podrían utilizarse por parte de los municipios en la regulación de la prostitución, o al menos, podrían preguntarse porqué se distribuyen así. Si las zonas industriales tienen una mayor permisividad, tanto social como policial, podrían tomarse medidas que *dificultaran* la consolidación de negocios de este tipo, moviéndose hacia otras zonas en las que exista mayor vigilancia.

Respecto al resto de elementos que parecen intervenir de manera significativa, no parece que la explicación pueda ahondar demasiado en el fenómeno; en muchos de los casos —seguro— se nos ocurrían más de un motivo para que se den de tal forma, como por ejemplo en el caso de la renta o la distancia a carreteras. Sí cabe destacar la mayor importancia de los condicionantes económicos frente a los sociales, descritos estos últimos ampliamente en la literatura. Como negocio, la prostitución busca maximizar los beneficios, y en la actualidad el estigma social posiblemente suponga un menor freno respecto al pasado. Puede que sea por eso por lo que no encontramos prostíbulos en zonas de bajo nivel socioeconómico pero sí los encontramos cerca de vías de comunicación y de tráfico habitual.

Por último, se quiere insistir en que no se señalan las causas de la distribución geográfica de la prostitución, sino que se señalan algunos de los elementos que intervienen en el proceso de generación de los datos. La explicación teórica sobre lo que pasa en estos sitios y en municipios de determinadas características puede que tenga mucha relevancia social, pues muchos de los elementos intervinientes son de este tipo; sin embargo, eso es tarea para otro trabajo.

## Referencias

- Bertozzi, Sarah. (2005). *Vicious Geography: The Spatial Organization of Prostitution in Twentieth Century Philadelphia*. Recuperado de: <https://researchgate.net>
- Darmofal, D. (2015). *Spatial Analysis for the Social Sciences (Analytical Methods for Social Research)*. Cambridge: Cambridge University Press. DOI:10.1017/CBO9781139051293
- Haining, R. (2003). *Spatial Data Analysis: Theory and Practice*. Cambridge: Cambridge University Press. DOI:10.1017/CBO9780511754944
- Helderop, E., Huff, J., Morstatter, F., Grubestic, A., & Wallace, D. (2019). *Hidden in Plain Sight: A Machine Learning Approach for Detecting Prostitution Activity in Phoenix, Arizona*. *Applied Spatial Analysis and Policy*, 12(4), 941-963. DOI: 10.1007/s12061-018-9279-1
- Herrera Gomez, Marcos & Mur, Jesus & Ruiz Marín, Manuel. (2011). *¿Cuál matriz de pesos espaciales? Un enfoque sobre selección de modelos*. Recuperado de: <https://www.researchgate.net>
- Hubbard, Philip. (1997). *Red-light districts and toleration zones: geographies of female street prostitution in England and Wales*. *The Royal Geographical Society. Area*, 29(2), 129–140. DOI: 10.1111/j.1475-4762.1997.tb00015.x
- Hubbard, Philip. (1999). *Sex and the city: geographies of prostitution in the urban West*. Recuperado de: <https://www.academia.edu>
- Inga Santiviáñez, Rosa María. (2015). *El criterio de información de Akaike en el análisis de datos categorizados*. (Tesis doctoral). Universidad Complutense de Madrid, Madrid. Recuperado de: <https://eprints.ucm.es>
- Larsen, E. N. (1992). *The politics of prostitution control: interest group politics in four Canadian cities*. *International Journal of Urban and Regional Research*, 16,169–189. DOI: 10.1111/j.1468-2427.1992.tb00167.x
- Manuel B. Aalbers & Michaël Deinema. (2012). *Placing prostitution: The spatial-sexual order of Amsterdam and its growth coalition*. Recuperado de: <https://researchgate.net>
- Philo, Chris. (1991). *New Words, New Worlds: Reconceptualising Social and Cultural Geography*. Recuperado de: <https://www.raco.cat>

- Pue, W. Wesley. (1990). *Wrestling with law: (geographical) specificity vs. (legal) abstraction*. *Urban Geography*, 11:6, 566-585, DOI: 10.2747/0272-3638.11.6.566
- Voloshin, Daniil & Derevitskii, Ivan & Mukhina, Ksenia & Karbovskii, Vladislav. (2016). *Identifying Venues for Female Commercial Sex Work using Spatial Analysis of Geocoded Advertisements*. Recuperado de: <https://researchgate.net>

## Anexo

Se adjunta un código qr y un enlace del mismo sitio web para la consulta de datos de este trabajo, con la opción de descargar el script de R y la matriz de datos. Además, se elaboró un mapa interactivo con las variables aquí utilizadas que puede servir de ayuda para una exploración más *intuitiva*.



<https://rpubs.com/frife/tfgusc>