Datos sociodemográficos, renta, gastos de los hogares

DESCRIPCIÓN DEL SERVICIO

25/9/20

info@unica360.com
www.unica360.com

CONTENIDO:

El presente documento el dataset de datos sociodemográficos, renta y gastos a nivel de grid 100m, desarrollado por unica360.

índice

1. ANTECEDENTES Y OBJETIVOS 2
2. DATOS DE MERCADO POTENCIAL, APLICACIONES 2
3. DATOS SOCIODEMOGRÁFICOS, RENTA, GASTOS DE LOS HOGARES 3
3.1. Sociodemografía 3
3.2. Renta media de los hogares. 3
3.3. Gastos de los hogares por categoría 4
4. GRANULARIDAD, CUADRÍCULA DE 100M 7

1. ANTECEDENTES Y OBJETIVOS

En unica360 desarrollamos proyectos de análisis de clientes y geomarketing. En el curso de estos trabajos, hemos ido generando una serie de sets de datos de mercado potencial, con componente espacial, de gran utilidad para:

- Conocer el entorno de un establecimiento comercial
- Estimar la demanda en función del tipo de entorno
- Enriquecer datos de clientes: en función de dónde vives, podemos deducir, hasta cierto punto, cómo eres

En los últimos años, nos llega una creciente demanda de este tipo de datos, ligada al crecimiento del análisis de datos de negocio, Big data, Business Analyitcs… Ante esta necesidad, hemos modelado sets de datos estandarizados para servirlos de maneras diversas.

2. DATOS DE MERCADO POTENCIAL, APLICACIONES

Algunos casos en que este tipo de datos es aplicable, de manera muy resumida:

- **Retail**, tiendas o servicios físicas: conocimiento del entorno de cada establecimiento, segmentación de red comercial, selección de ubicaciones óptimas en expansión, benchmark con tiendas de la competencia.
- **e-commerce**: enriquecimiento de datos de clientes, como renta, hábitat, oferta comercial en su zona), probabilidad de tener niños, vivienda unifamiliar aislada… (afectan a la elección de canal online).
- **Gran Consumo**: enriquecimiento de potenciales clientes detallistas, estimación de su potencial de demanda, adecuación del surtido ideal.
- **Business analytics**: enriquecimiento de datos de clientes, establecimientos, rutas, centros logísticos… en entornos de análisis de datos, Big data.
- **portales e inmobiliarias online**: enriquecimiento de datos de inmuebles, búsqueda por tipo de zona, generación de descriptivos a partir de la zona, cualificación de leads, descriptivos automáticos y SEO optimizado.
- **Smart city y transporte**: ubicación y dimensionamiento de equipamientos, simulación de desplazamientos y demanda de transporte, diseño de ciudad distribuida.
- **publicidad exterior**: cualificación de cada uno de los soportes, para adecuar oferta al público objetivo.
- **investigación de mercados**: cualificación de datos de miembros de paneles, diseño de muestras de nicho, segmentadas.
- **Servicios a domicilio** de todo tipo, enriqueciendo los clientes por sus sociodemográficos, tipo de hábitat -urbano vs rural vs diseminado-, tipo de vivienda -unifamiliar, tamaño…-

En cualquier otro sector -banca, seguros, servicios personales…-se obtienen igualmente grandes ventajas competitivas enriqueciendo datos sobre puntos de venta y clientes.

3. DATOS SOCIODEMOGRÁFICOS, RENTA, GASTOS DE LOS HOGARES

La API sirve, actualmente, los siguientes tipos de datos:

3.1. Sociodemografía. Tabla básica referida a la población residente, recogiendo las siguientes dimensiones:

- Población por sexo y edad, hogares, viviendas totales
- Renta media de los hogares
- Tasa de inmigrantes por tipos
- Nivel de estudios

3.2. Renta media de los hogares. Estima los ingresos medios anuales de las personas, los declarantes tributarios y los hogares, en € brutos y netos (antes de impuestos y cotizaciones por parte del trabajador o después de la liquidación de éstos).

Informa también del origen de las rentas -trabajo, pensión, prestación por desempleo, otras prestaciones, capital/rentas-.

Se modeliza a partir de diversas fuentes de datos de origen público, difundidos por el INE, institutos de estadística autonómicos y open data varios. Las principales fuentes son:

1. INE, estadística experimental, Atlas de distribución de la Rnta de los Hogares (ADRH) por sección censal
2. AEAT, estadística de declarantes por código postal, municipio
3. Precio medio de vivienda por sección censal
4. Modelización sobre INE, Encuesta de Presupuestos Familiares (EPF), Encuesta de Condiciones de Vida (ECV)
5. Banco de España, Encuesta Financiera de las Familias (EFF)

Como se ha mencionado, se completa con otras fuentes a nivel de comunidad autónoma, para mejorar los datos públicos en tres aspectos:

- Estimación precisa de la renta en áreas no informadas, por tener Hacienda propia (País Vasco, Navarra), por tamaño de la zona (no se informan secciones y códigos postales con poca población).
- Desagregación espacial hasta el nivel de grid 100m o portal, gracias a datos de características de viviendas y edificios y precios medios de venta, alquiler.

- Corrección de áreas que en alguna de las operaciones descritas son "recortadas" por considerarse valores extremos, cuando en realidad reflejen la asimétrica distribución de los ingresos.

El modelo incluye el origen de la renta -trabajo, prestaciones varias, rentas…- como aportación a la renta, para la zona tributaria común (excluidos País Vasco y Navarra).

3.3. Gastos de los hogares por categoría

Gasto medio por hogar en los 12 grupos COICOP, modelizado a partir de la *Encuesta de Presupuestos Familiares*, censo, padrón..

Los niveles de gasto estimados para cada sección censal presentan una variabilidad limitada –tantos valores como nodos en el árbol de decisión- marcada por dos limitaciones de los datos de origen:

- Tamaño y representatividad de la muestra de la EPF, nodos muy pequeños no se comportan de manera robusta en set de validación
- Limitaciones de cruces posibles en datos de censos y padrones de INE

Por ello, se ha optado siempre por la robustez del modelo, aun cuando la variabilidad sea limitada y –probablemente- la precisión también.

Los 12 grandes grupos son:

01 Alimentos y bebidas no alcohólicas

02 Bebidas alcohólicas y tabaco

03 Vestido y calzado

04 Vivienda, agua, electricidad, gas y otros combustibles

05 Muebles, artículos del hogar y artículos para el mantenimiento corriente del hogar

06 Sanidad

07 Transporte

08 Comunicaciones

09 Ocio y cultura

10 Enseñanza

11 Restaurantes y hoteles

12 Otros bienes y servicios

La tabla siguiente muestra un Informe de gasto por grupo en el área de influencia de un establecimiento. El porcentaje se compara con la media española.

EJEMPLO DE GASTO MEDIO Y TOTAL EN UN AREA DE INFLUENCIA FICTICIA

Grupo	Descripción	Gasto total	gasto / hogar	pct	Total España
01	Alimentos y bebidas no alcohólicas	15,495,702 €	4,624 €	12.5%	14.4%
011	Productos alimenticios[1]	14,411,003 €	4,301 €	11.6%	13.5%
012	Bebidas no alcohólicas	1,084,699 €	324 €	0.9%	0.9%
02	Bebidas alcohólicas, tabaco y narcóticos	2,055,128 €	613 €	1.7%	2.1%
03	Artículos de vestir y calzado	6,557,215 €	1,957 €	5.3%	5.6%
04	Vivienda, agua, electricidad, gas y otros combustibles	38,905,871 €	11,610 €	31.3%	30.0%
05	mobiliario, equipamiento del hogar y gastos corrientes de conservación de vivienda	6,138,289 €	1,832 €	4.9%	4.9%
06	Salud	3,871,843 €	1,155 €	3.1%	3.2%
07	Transportes	13,375,848 €	3,992 €	10.8%	12.4%
08	Comunicaciones	3,284,159 €	980 €	2.6%	3.1%
09	Ocio, espectáculos y cultura	8,675,851 €	2,589 €	7.0%	6.6%
10	Enseñanza	2,811,084 €	839 €	2.3%	1.0%
11	Hoteles, cafés y restaurantes	12,754,843 €	3,806 €	10.3%	9.1%
12	Otros bienes y servicios	10,210,854 €	3,047 €	8.2%	7.6%
TOTAL		**124,136,687 €**	**37,045 €**	**100.0%**	

(1)Se excluyen los productos vendidos por hoteles, restaurantes, cafés, bares, máquinas automáticas... para ser consumidos normalmente en el lugar de compra como tales) que van al código 09342.tampoco los alimentos para animales de compañía (vendidos específicamente(grupo 11) y los platos preparados para llevar, elaborados en tratorías o similares, aunque sean consumidos en el hogar (código 11113). No se incluyen tampoco los alimentos para animales de compañía (vendidos específicamente como tales).

Imagen 1. Ejemplo de gasto de los hogares en torno a un establecimiento minorista

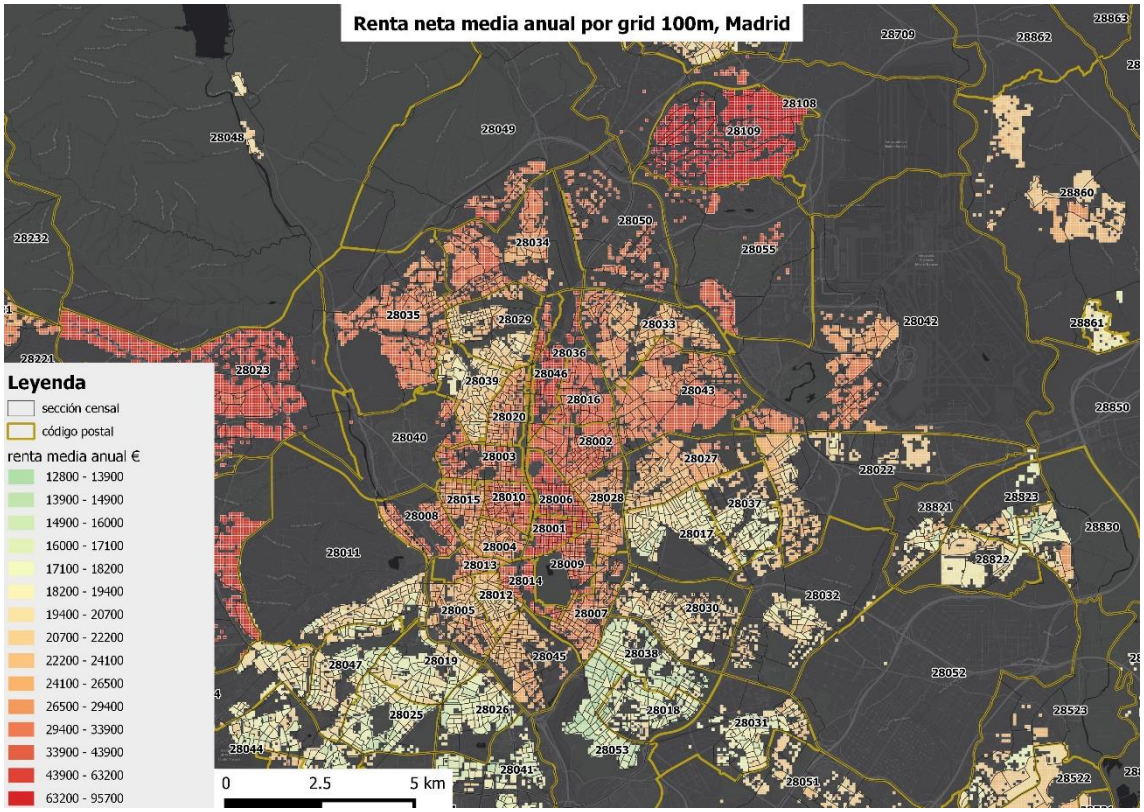

Imagen 2. Renta neta por titular media, grid 100m, Madrid

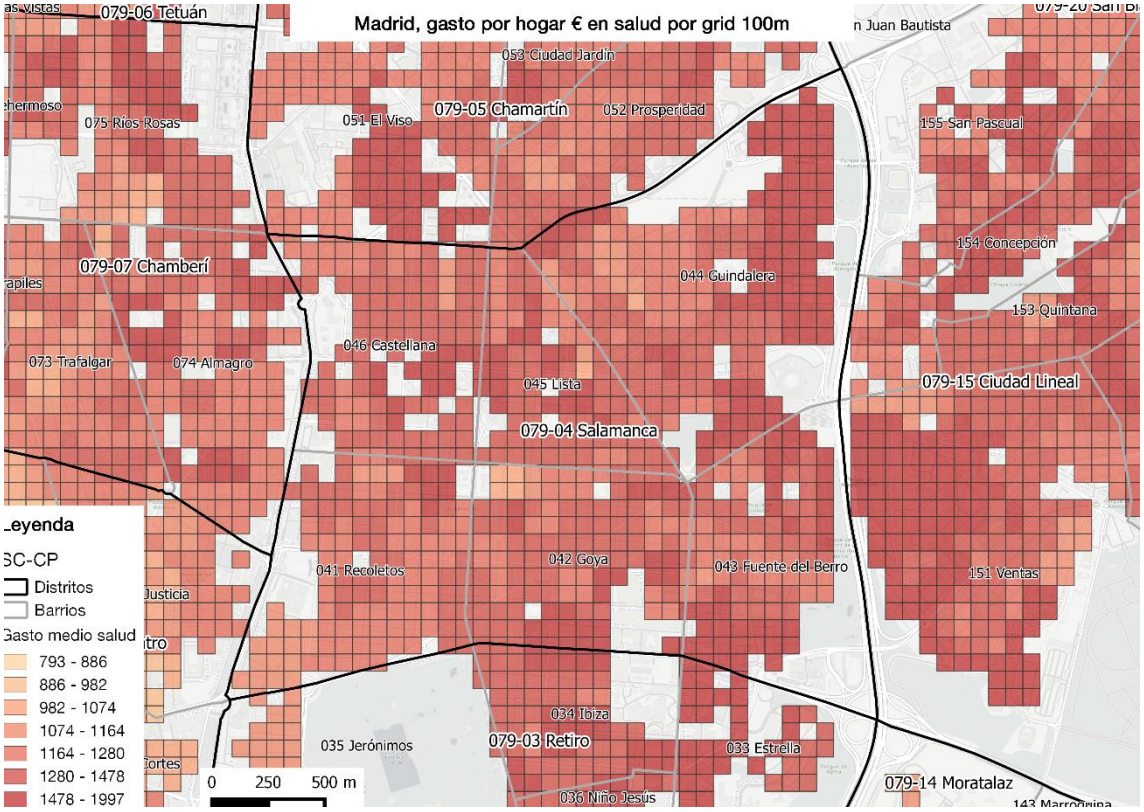

Imagen 3. Grid 100m, gasto medio por hogar en salud, Distrito de Salamanca

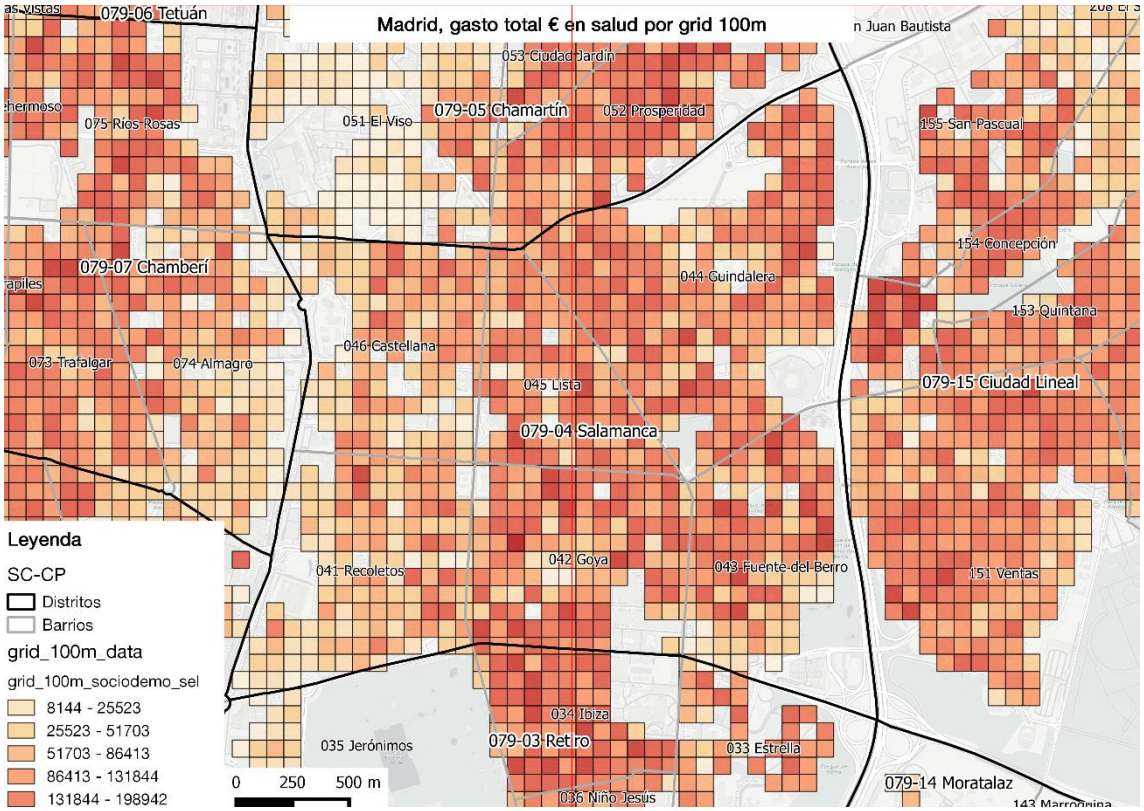

Imagen 4. Grid 100m, gasto total de los hogares en salud, Distrito de Salamanca

4. GRANULARIDAD, CUADRÍCULA DE 100M

Los datos pueden ser devueltos a diferente granularidad territorial:

- Municipio
- Código postal
- Sección censal
- Cuadrícula (grid) de 100m

Este último nivel representa la granularidad más fina. Presenta varias ventajas:

- Independencia de procesos cartográficos simplificando la asignación de celdas a puntos con coordenadas y evitando actualizaciones de capas
- Estabilidad geométrica necesaria en cálculos de densidades
- Al referirse el centro de cada cuadrado a centenas de metros en el sistema de coordenadas proyectadas EPSG:3857, es sencillo convertir los resultados en capas geográficas, sin necesidad de la cartografía (grid) de referencia
- Finalmente, este sistema de referencia de coordenadas es universal, usado por servicios de mapas web como Google maps.