

INSTITUTO NACIONAL DE ESTADÍSTICA

EN

**Estadística Piloto sobre
Movilidad a partir del
posicionamiento de teléfonos
móviles (Censos de Población y
Viviendas 2021).**

Subdirección General de
Estadísticas Sociodemográficas

Octubre de 2019

Índice

1	Introducción	3
2	Objetivos	5
3	Ámbito de la investigación	5
3.1.	ÁMBITO POBLACIONAL	5
3.2.	ÁMBITO GEOGRÁFICO	5
3.3.	ÁMBITO TEMPORAL	5
4	Características de la investigación	6
5	Recogida de datos	8
6	Tratamiento de la información	8
7	Difusión de los resultados	9
8	Calendario	9

ÁNEXO.

Ejemplo: división de la provincia de Cádiz en celdas de movilidad

Estadística Piloto sobre Movilidad a partir del posicionamiento de teléfonos móviles (Censo de Población y Viviendas 2021).

1 Introducción

Los Censos de Población y Viviendas, que se realizan cada 10 años, constituyen una operación estadística de mucha tradición y de mucha envergadura. Hasta 2001 se realizaron mediante el envío de cuestionarios a todos los hogares residentes en España. En 2011 el censo de población y viviendas se realizó con una muestra de aproximadamente el 10% de la población. El censo de 2021, que se realizará con fecha de referencia 1 de enero estará, por primera vez, basado en el uso de registros. España se unirá al reducido número de países a nivel mundial que está en condiciones de acometer el censo con este tipo de metodología.

El censo de 2021 será un reto metodológico importante. El uso de registros de calidad como los que actualmente existen en España permitirá la difusión de información con una periodicidad inferior a los 10 años y aumentar la calidad de la información.

Pero aunque casi todas las variables que se investigan en el censo pueden obtenerse a través de registros administrativos, existe una serie de ellas que tienen una gran demanda por parte de los usuarios y que no están disponibles a través de estas fuentes administrativas. La continuidad de las series censales de esas variables solo puede quedar garantizada mediante la realización de operaciones específicas.

El censo de 2011, como ya ocurriera en ediciones anteriores, introducía varias preguntas que permitían medir tanto la movilidad cotidiana como una cierta aproximación muy específica a la medición de los movimientos turísticos (el estudio de lo que se conoce como “población vinculada”¹).

La información relacionada con las preguntas de movilidad fue de la más demandada del Censo 2011. En él se preguntó a las personas por la ubicación de su lugar de trabajo/estudio (pregunta 19) y también por si pasaba 14 noches o más (por motivos laborales o de veraneo) en un municipio diferente del de residencia (pregunta 4). Con estas dos preguntas se calculó la denominada población vinculada, incluida por primera vez en el Censo 2001, que permite disponer de una mejor estimación de la carga poblacional real que soporta un municipio.

Adicionalmente se incluyeron otras tres preguntas relacionadas con el número de viajes diarios, el medio de transporte utilizado y el tiempo transcurrido para llegar al lugar de trabajo/estudio.

Conseguir toda esta información a partir de fuentes administrativas no resulta posible. Pero existe una nueva y prometedora fuente que podría dar respuesta a gran parte de la información proporcionada en 2011: el uso de la **información espacio-temporal de los teléfonos móviles**.

¹ Para más detalles ver https://www.ine.es/censos2011/censos2011_meto_pobla_vinculada.pdf

El estudio que aquí se propone es un análisis de idoneidad de la información estadística obtenida a partir del posicionamiento de los teléfonos móviles. El objetivo es que esta información sustituya a la que se ofrece tradicionalmente en los Censos de Población y Viviendas.

En concreto se trata de construir, idealmente con información de los tres principales operadores de telefonía móvil (OTM), tres matrices de movilidad que se detallarán más adelante.

Para el estudio piloto se han escogido aquellas compañías de telefonía móvil que se considera a priori que pueden tener suficiente nivel de representatividad en todo el territorio español.

Según los datos oficiales publicados por la Comisión Nacional de los Mercados y a Competencia (julio de 2019) el número de teléfonos móviles por operador es el siguiente:

<u>Operador</u>	<u>Terminales</u>	<u>cuota mercado (%del total)</u>
Movistar:	16.270.132	30,3%
Orange:	13.705.972	25,5%
Vodafone:	12.293.129	22,9%
Grupo Masmóvil:	7.018.220	13,1%
Operadores móviles virtuales:	4.471.348	8,3%
Total:	53.758.801	100%

En concreto las tres primeras compañías tienen una cuota de telefonía móvil que ampliamente supera el 20% y puede considerarse como aceptable para los objetivos que se persiguen en este proyecto.

Parece lógico suponer que aquellos OTM con cuotas superiores al 20% de mercado van a poder ofrecer datos para todas las celdas geográficas determinadas por el INE (unas 3.200, tal como se precisa más adelante). Según los datos por CCAA publicados por la CNMC², las tres primeras compañías se distribuyen de forma bastante homogénea en el territorio. En el peor de los casos, estas compañías ofrecen cuotas de mercado en torno al 17% (caso de Orange en el País Vasco). No ocurre así con la cuarta compañía en cuanto a cuota de mercado, el Grupo Masmóvil, que apenas alcanza el 5% en algunas CCAA, siempre según esta fuente y que podría arrojar información no significativa en un número elevado de las áreas objeto de estudio.

² <http://data.cnmc.es/datagraph/jsp/graph/grafico-cuota-mercado.jsp> actualizada a datos de 2017

2 Objetivos

El principal objetivo de esta estadística piloto es evaluar la idoneidad de la información sobre el posicionamiento de los teléfonos móviles para su uso como alternativa al método actual, basada en los censos de población, con el fin de proporcionar matrices de movilidad.

Para ello se construirán tres matrices de movilidad, que se publicarán en la página web del INE dentro del apartado de estadística experimental para así obtener retroalimentación de los usuarios e incluir en el futuro esta fuente de datos dentro del plan de utilización regular de las estadísticas demográficas.

Además, a partir de la información de movilidad cotidiana recogida en los censos de 2011 se puede delimitar una nueva unidad de tipo geográfico, que va camino de tener carácter oficial, el **área urbana funcional** (AUF)³. La información recogida en esta investigación permite analizar la posible delimitación de las AUF a partir de esta nueva fuente.

Aunque la definición es compleja, esencialmente una AUF es una ciudad grande (normalmente mayor de 100.000 habitantes) y un conjunto de municipios que la rodean y en los cuales más del 15% de los residentes se desplaza cotidianamente a ese municipio principal. Ese concepto de construir áreas urbanas a partir de la movilidad cotidiana, medida de una u otra forma y usando distintos umbrales, es el que está presente en este estudio.

Toda la información se construirá con la suma de resultados de los tres operadores, de forma que será imposible determinar dato alguno de ningún operador en concreto.

3 Ámbito de la investigación

3.1 ÁMBITO POBLACIONAL

El ámbito poblacional está constituido por los teléfonos móviles de población residente en España de los tres operadores antes citados. Se excluyen teléfonos extranjeros en roaming. La información solicitada, además, se refiere solo a teléfonos móviles, no a todos los dispositivos susceptibles de tener una tarjeta SIM.

3.2 ÁMBITO GEOGRÁFICO

El ámbito territorial es todo el territorio nacional. El territorio se divide en aproximadamente 3.200 “celdas de movilidad” y se ofrece información para cada una de ellas.

3.3 ÁMBITO TEMPORAL

Hay varios ámbitos según la matriz concreta tal como se detalla en el apartado posterior.

³https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Functional_urban_area

El período de recogida previsto en las matrices 1 y 2 es de cuatro días laborables consecutivos. La elección de estas fechas es poco relevante. Se trata de escoger una semana “normal” en cuanto a movilidad, es decir, con actividad educativa y laboral normal, sin festivos. No se puede precisar todavía la semana concreta, que depende de la fecha de inicio del estudio. Tentativamente se ha propuesto una semana de noviembre de 2019.

Para la matriz 3 se toman cuatro días concretos, por definir, pero tentativamente se han fijado los siguientes: 20 de julio, 15 de agosto, 24 de noviembre y 25 de diciembre, todos ellos de 2019. Se pretende con ellos tener dos días de verano –un momento de máxima población estacional como puede ser el 15 de agosto y otro no tan intenso como un sábado de julio–, el día de navidad, y un fin de semana “valle” que coincidiría con la semana de los días laborables elegidos anteriormente, para tener una imagen casi completa de una semana “normal”.

4 Características de la investigación

Se plantea construir tres matrices que se describen a continuación. Se trata de tres tablas concretas, construidas como suma de los datos de los distintos operadores.

La unidad mínima de observación geográfica sería la “**celda de movilidad INE**” (en el anexo figura una provincia de ejemplo). Para la determinación de las celdas se considera la población empadronada en ellas, de forma que cada celda debe superar el umbral de 5.000 empadronados (algunas pocas celdas quedan muy ligeramente por debajo de ese umbral). Con esto se pretende garantizar que el número de usuarios de cualquier OTM en cada celda sea suficiente para el análisis de movilidad entre celdas. Las celdas se construirían así:

- Dentro de cada provincia, los municipios menores de 5.000 habitantes, se agrupan geográficamente para alcanzar al menos ese umbral de población.
- Cada uno de los municipios de entre 5.000 y 50.000 habitantes constituye una celda.
- Los municipios de más de 50.000 habitantes se pueden desagregar en distritos o barrios, siempre con un umbral superior a esos 5.000 habitantes (se estima que eso supondría unas 1.150 celdas para toda España). En el caso más extremo, Madrid, se divide el municipio en los 128 barrios que la componen dado que todos superan el umbral de 5.000 habitantes contemplado.

Con una primera estimación, el número total de celdas rondaría las 3.220 con un mínimo 5.000 habitantes salvo mínimas excepciones. El tamaño medio de cada celda es de unas 15.000 personas, lo que significa un número medio estimado de 13.000 teléfonos sumando los tres operadores.

Delimitadas estas celdas, los tres productos a construir serían:

TABLA Nº 1: MATRIZ DE MOVILIDAD COTIDIANA

Esta es la matriz principal del proyecto, pues es la que permite delimitar las AUFs. Se trata de una matriz de origen-destino formada por N (~ 3.220) celdas x N celdas. En la casilla “fila A-columna B” de ella figuraría el número de terminales que, de media, se desplazan cotidianamente de la celda de residencia (A) a la celda de destino cotidiano (B).

Para la construcción de esta matriz, en primer lugar se debe determinar, la celda de residencia de cada uno de los **teléfonos móviles de cada OTM**, mediante este método:

Se toman como mínimo 4 días laborables consecutivos de noviembre de 2019 (pueden ser los días lunes 18 a jueves 21). La celda de residencia de cada teléfono móvil es aquella en donde el teléfono móvil se encuentra durante más tiempo entre las 00:01 y las 06:00 de los cuatro días observados.

Se está estudiando con los OTM alguna otra forma de asignar la celda de residencia basada en un número distinto de días u horarios.

Asignada para cada teléfono móvil la celda de residencia, se intenta determinar la celda de destino cotidiano. De nuevo, este destino recurrente se construirá tomando como referencia solo los mismos días elegidos anteriormente.

El método es parecido al de asignación de celda de residencia pero no dará siempre un valor: se deberán tomar todas las celdas (incluida la de residencia) en las que se encuentra el terminal durante el horario 9:00 a 18:00 siempre y cuando se encuentre a ese terminal en esa celda al menos durante cuatro horas al día y al menos en dos días de los cuatro observados. Si hay más de una celda, se tomará aquella más frecuente (en la que está más tiempo), que no sea la de residencia. Si solo la celda de residencia cumple esta condición, entonces la celda de residencia será también la celda de destino.

Tal y como se ha dicho anteriormente, este método no siempre va a poder obtener una celda de destino, pues en algunos casos no se podrá determinar que exista una celda a la que se acude de forma recurrente. Por tanto no es necesario asignar celda de destino a todos los teléfonos móviles, pero sin embargo sí que es necesario que todos los teléfonos móviles tengan asociada una celda de origen.

TABLA N° 2. MATRIZ DE POBLACIÓN DE DÍA Y DE NOCHE

Para esta tabla no se usa ninguna información de origen ni de destino. Se trata simplemente de conocer cuántos teléfonos móviles hay en cada celda de movilidad cotidiana, en M momentos distintos del día. Es por tanto una tabla de tamaño ($N \times M$). Permite tener información sobre población flotante. Se tomarían estos M puntos de corte en el día a acordar con los OTM pero el mínimo es de 6 momentos: 06:00, 10:00, 14:00, 18:00, 22:00, 02:00, para tipificar las zonas según las horas en las que tienen actividad.

Es importante tener en cuenta que la tabla n° 2 se construye tomando la media de los mismos cuatro días de noviembre de 2019 antes mencionados.

Cada teléfono móvil se debe contabilizar en una única celda en cada momento. Para ello, cuando no pueda determinarse en qué celda está en el instante dado y se tenga información de un instante anterior o posterior, se le asignará la celda en la que estuviera en el instante más cercano en el tiempo del instante elegido. Así si un teléfono está localizado en la celda A a las 17:35 y en la celda B a las 18:10 se considera que a las 18:00 está en la celda B.

TABLA Nº 3. MATRIZ DE POBLACIÓN ESTACIONAL

Se trata de construir una matriz $N \times N \times 4$ (o cuatro matrices $N \times N$). Cada una de las cuatro corresponde a un día concreto. En principio se consideran los días 20 de julio, 15 de agosto, 24 de noviembre y 25 de diciembre, todos ellos de 2019.

Para cada día elegido, en la casilla fila A-columna B figuraría el número de personas que, residiendo en la celda A, pernoctan en la celda B durante ese día, según las siguientes especificaciones:

Para el día elegido, se determina para cada terminal la celda donde pernocta (donde ese teléfono móvil se encuentra más frecuentemente durante las horas 02:00 a 06:00 en esa fecha).

Este método debe poder proporcionar una celda de pernoctación a todos los teléfonos presentes en territorio español en esas fechas.

Para cada uno de los teléfonos se debe encontrar su celda de residencia, bien usando el método descrito para la tabla nº 1, bien tomando como referencia fechas anteriores a la fecha objeto de estudio en cada caso. Es decir, para determinar la celda de residencia necesaria para elaborar la submatriz del 25 de diciembre se puede tomar la celda de pernoctación más frecuente de la semana anterior (días 16 a 22 de diciembre) o los siete días anteriores (18 a 24 de diciembre) o el mes anterior, o de los días de noviembre mencionados en la tabla nº 1. El método elegido se debe acordar entre el INE y los OTM.

5 Recogida de datos

El INE recibirá las matrices ya agregadas elaboradas por cada operador. Por tanto el INE no va a disponer en ningún momento de información individual de ningún tipo sino solo de tres matrices de resultados finales elaboradas por cada uno de los OTM.

6 Tratamiento de la información

Las nueve matrices de resultados recibidas (tres por cada OTM) se integran en tres sumando los datos de los OTM.

Siempre que el número de observaciones sea inferior a 10 en una celda, para un operador dado, el resultado vendrá censurado en origen. De esta forma no se podrán ofrecer datos de movilidad origen-destino para celdas en las cuales al menos uno de los tres OTM cuente menos de ese umbral.

Estas limitaciones son irrelevantes de cara al proyecto que pretende solo conocer movilidad cotidiana de cierta entidad. No considerar flujos menores de 50 teléfonos para celdas de 5.000 teléfonos significa despreciar flujos inferiores al 1% de la población en el peor de los casos.

Los OTM también proporcionarán información sobre aquellas celdas en la que la cobertura de telefonía móvil pueda arrojar resultados menos fiables, para su consideración en las tablas de resultados

7 Difusión de resultados

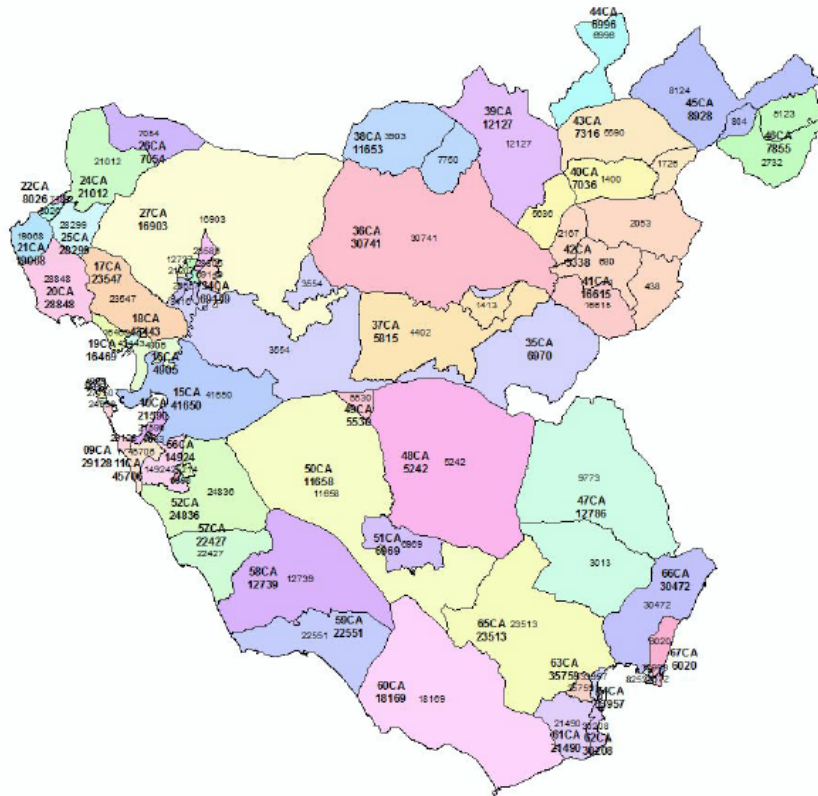
Se publicarán las tres matrices completas en forma numérica y también en forma de mapas. En cuanto a la información numérica, lo relevante del proyecto es conocer porcentajes de personas que se mueven entre áreas, por lo que en principio se optará por ofrecer información en forma de porcentajes, no de números absolutos, que no reflejarían a la población en su conjunto sino al universo de teléfonos de los tres OTM

8 Calendario

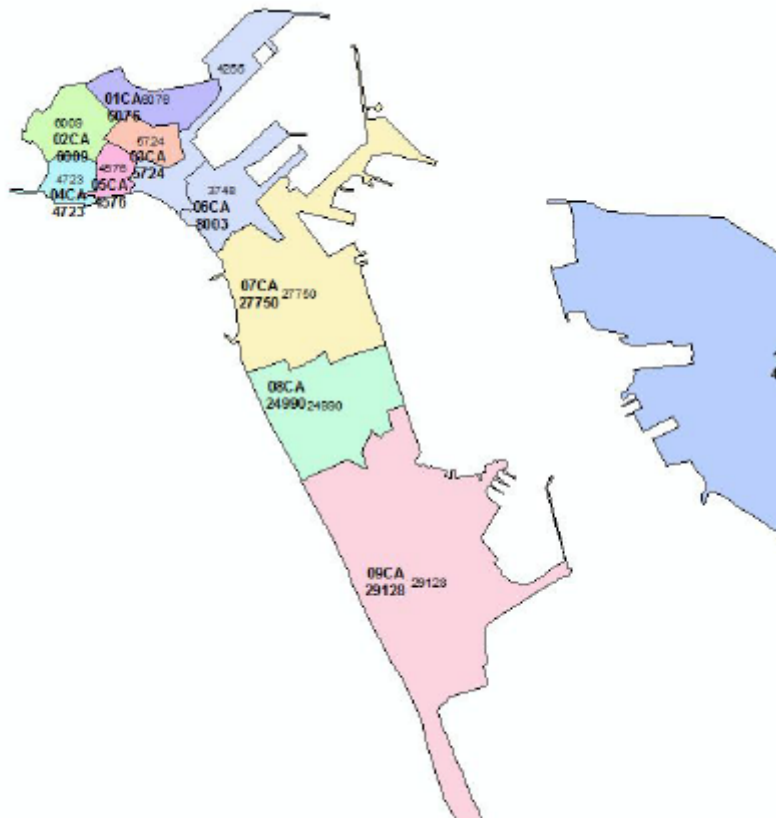
El calendario de las principales fases de la investigación es el siguiente:

- Elaboración de datos por parte de los OTM: de diciembre 2019 a marzo 2020
- Integración de la información y elaboración de los productos de difusión: abril 2020
- Publicación: mayo 2020

Anexo. Ejemplo: división de la provincia de Cádiz en celdas de movilidad



Detalle del Municipio de Cádiz:



Celdas de movilidad. Provincia de Cádiz (total: 87)

CUMUN	NMUN	CUDIS	POB_DIST18	NPRO	CPRO	POB_MUN18	ID_GRUPO	POB_GRUPO
11012	Cádiz	1101205	4576	Cádiz	11	116979	05CA	4576
11012	Cádiz	1101204	4723	Cádiz	11	116979	04CA	4723
11020	Jerez de la Frontera	1102001	4875	Cádiz	11	212879	32CA	4875
11027	Puerto de Santa María, El	1102704	4905	Cádiz	11	88364	16CA	4905
11001	Alcalá de los Gazules		0	Cádiz	11	5242	48CA	5242
11009	Benaocaz		0	Cádiz	11	680	42CA	5338
11011	Bosque, El		0	Cádiz	11	2167	42CA	5338
11019	Grazalema		0	Cádiz	11	2053	42CA	5338
11040	Villaluenga del Rosario		0	Cádiz	11	438	42CA	5338
11025	Paterna de Rivera		0	Cádiz	11	5530	49CA	5530
11012	Cádiz	1101202	5724	Cádiz	11	116979	03CA	5724
11003	Algar		0	Cádiz	11	1413	37CA	5815
11902	San José del Valle		0	Cádiz	11	4402	37CA	5815
11012	Cádiz	1101203	6009	Cádiz	11	116979	02CA	6009
11022	Línea de la Concepción, La	1102207	6020	Cádiz	11	62940	67CA	6020
11012	Cádiz	1101201	6076	Cádiz	11	116979	01CA	6076
11901	Benalup-Casas Viejas		0	Cádiz	11	6969	51CA	6969
11020	Jerez de la Frontera	1102008	3416	Cádiz	11	212879	35CA	6970
11020	Jerez de la Frontera	1102010	3554	Cádiz	11	212879	35CA	6970
11029	Puerto Serrano		0	Cádiz	11	6996	44CA	6996
11026	Prado del Rey		0	Cádiz	11	5636	40CA	7036
11042	Zahara		0	Cádiz	11	1400	40CA	7036
11037	Trebujena		0	Cádiz	11	7054	26CA	7054
11022	Línea de la Concepción, La	1102204	7079	Cádiz	11	62940	71CA	7079
11022	Línea de la Concepción, La	1102206	7167	Cádiz	11	62940	68CA	7167
11005	Algodonales		0	Cádiz	11	5590	43CA	7316
11018	Gastor, El		0	Cádiz	11	1726	43CA	7316
11022	Línea de la Concepción, La	1102205	7650	Cádiz	11	62940	70CA	7650
11031	San Fernando	1103104	7850	Cádiz	11	95174	13CA	7850
11002	Alcalá del Valle		0	Cádiz	11	5123	46CA	7855
11034	Setenil de las Bodegas		0	Cádiz	11	2732	46CA	7855
11012	Cádiz	1101206	4255	Cádiz	11	116979	06CA	8003
11012	Cádiz	1101207	3748	Cádiz	11	116979	06CA	8003
11032	Sanlúcar de Barrameda	1103202	8026	Cádiz	11	68037	22CA	8026
11022	Línea de la Concepción, La	1102201	8253	Cádiz	11	62940	72CA	8253
11024	Olvera		0	Cádiz	11	8124	45CA	8928
11036	Torre Alháuquime		0	Cádiz	11	804	45CA	8928
11015	Chiclana de la Frontera	1101502	9199	Cádiz	11	83831	55CA	9199
11031	San Fernando	1103102	960	Cádiz	11	95174	14CA	9303
11031	San Fernando	1103103	8343	Cádiz	11	95174	14CA	9303
11015	Chiclana de la Frontera	1101503	9658	Cádiz	11	83831	54CA	9658
11022	Línea de la Concepción, La	1102202	9672	Cádiz	11	62940	73CA	9672
11032	Sanlúcar de Barrameda	1103203	7497	Cádiz	11	68037	23CA	10700
11032	Sanlúcar de Barrameda	1103204	3203	Cádiz	11	68037	23CA	10700
11031	San Fernando	1103105	10725	Cádiz	11	95174	12CA	10725
11010	Bornos		0	Cádiz	11	7750	38CA	11653
11017	Espera		0	Cádiz	11	3903	38CA	11653
11023	Medina-Sidonia		0	Cádiz	11	11658	50CA	11658
11041	Villamartín		0	Cádiz	11	12127	39CA	12127
11020	Jerez de la Frontera	1102005	12727	Cádiz	11	212879	29CA	12727
11039	Vejer de la Frontera		0	Cádiz	11	12739	58CA	12739
11013	Castellar de la Frontera		0	Cádiz	11	3013	47CA	12786
11021	Jimena de la Frontera		0	Cádiz	11	9773	47CA	12786
11015	Chiclana de la Frontera	1101501	14924	Cádiz	11	83831	56CA	14924
11027	Puerto de Santa María, El	1102702	16469	Cádiz	11	88364	19CA	16469
11038	Utrique		0	Cádiz	11	16615	41CA	16615
11020	Jerez de la Frontera	1102009	16903	Cádiz	11	212879	27CA	16903
11022	Línea de la Concepción, La	1102203	17099	Cádiz	11	62940	69CA	17099
11035	Tarifa		0	Cádiz	11	18169	60CA	18169
11016	Chipiona		0	Cádiz	11	19068	21CA	19068
11032	Sanlúcar de Barrameda	1103205	21012	Cádiz	11	68037	24CA	21012
11020	Jerez de la Frontera	1102003	21207	Cádiz	11	212879	30CA	21207
11004	Algeciras	1100403	21490	Cádiz	11	121414	61CA	21490
11031	San Fernando	1103101	21590	Cádiz	11	95174	10CA	21590
11014	Conil de la Frontera		0	Cádiz	11	22427	57CA	22427
11007	Barbate		0	Cádiz	11	22551	59CA	22551
11008	Barrios, Los		0	Cádiz	11	23513	65CA	23513
11027	Puerto de Santa María, El	1102703	23547	Cádiz	11	88364	17CA	23547
11020	Jerez de la Frontera	1102004	23586	Cádiz	11	212879	28CA	23586
11015	Chiclana de la Frontera	1101505	24836	Cádiz	11	83831	52CA	24836
11012	Cádiz	1101209	24990	Cádiz	11	116979	08CA	24990
11015	Chiclana de la Frontera	1101504	25214	Cádiz	11	83831	53CA	25214
11012	Cádiz	1101208	27750	Cádiz	11	116979	07CA	27750
11032	Sanlúcar de Barrameda	1103201	28299	Cádiz	11	68037	25CA	28299
11020	Jerez de la Frontera	1102007	28657	Cádiz	11	212879	31CA	28657
11020	Jerez de la Frontera	1102002	28805	Cádiz	11	212879	33CA	28805
11030	Rota		0	Cádiz	11	28848	20CA	28848
11012	Cádiz	1101210	29128	Cádiz	11	116979	09CA	29128
11004	Algeciras	1100402	30208	Cádiz	11	121414	62CA	30208
11033	San Roque		0	Cádiz	11	30472	66CA	30472
11006	Arcos de la Frontera		0	Cádiz	11	30741	36CA	30741
11004	Algeciras	1100401	33957	Cádiz	11	121414	64CA	33957
11004	Algeciras	1100404	35759	Cádiz	11	121414	63CA	35759
11028	Puerto Real		0	Cádiz	11	41650	15CA	41650
11027	Puerto de Santa María, El	1102701	43443	Cádiz	11	88364	18CA	43443
11031	San Fernando	1103106	45706	Cádiz	11	95174	11CA	45706
11020	Jerez de la Frontera	1102006	69149	Cádiz	11	212879	34CA	69149