

DATAVISION



UNIVERSITYHACK 2018®
DATATHON

UniversityHack 2018 - DATATHON
Reto Wefferent Card Analytics

Martín García Ripoll Muñoz
Lucas Gil Melby

ÍNDICE

Introducción.....	3
Fuentes de Datos.....	4
Funcionalidades.....	5
Análisis General.....	5
Optimización.....	9
Nuevo Negocio.....	12
Rutas Comunes.....	13
¿Dónde Comprar?.....	15
Posibles Mejoras.....	16

A fin de facilitar la lectura de este documento nos hemos limitado al uso del género masculino, por ej. "ciudadanos" o "usuarios". Por supuesto, toda la información aquí descrita, está destinada por igual a mujeres y a hombres. Por ello, las referencias al género deben considerarse neutrales.

INTRODUCCIÓN

Desde el equipo de DATAVISION estamos convencidos de la importancia del crecimiento empresarial a pequeña y gran escala, y pensamos también que este crecimiento, beneficioso no solo para las empresas sino para la sociedad en su totalidad pues crean riqueza y empleo a nivel local y nacional, debe ir de la mano de métodos científicos que no releguen las decisiones empresariales a la suerte o a métodos tradicionales. Es por esto que creemos que el recabado, tratamiento y análisis de datos masivos puede ayudar enormemente a este objetivo que tanto beneficia a la sociedad.

DATAVISION ha querido centrarse en crear una herramienta para el análisis y visualización de datos enfocado a simplificar la toma de decisiones comerciales basándose en datos objetivos, especialmente para un público menos familiarizado con el análisis de datos estadísticos.

En cuanto a la ayuda de toma de decisiones comerciales DATAVISION cuenta con dos apartados principales, '*NUEVO NEGOCIO*', enfocado a aquellas personas o entidades interesadas en abrir un negocio o un nuevo establecimiento y '*OPTIMIZACIÓN*', orientado a ayudar aplicar nuevas y/o mejores estrategias de *marketing* en establecimientos ya existentes para aumentar los ingresos y mejorar los servicios a los clientes.

Además de estos dos apartados de orientación, el apartado '*ANÁLISIS GENERAL*' ofrece una serie de estadísticas, gráficos y mapas que aportan una visión general de la situación comercial de un municipio o provincia, Murcia en este caso. Esta información es de gran utilidad para cualquier empresario interesado en conocer la situación comercial de su sector o sectores relacionados, para las instituciones públicas interesadas en tener un mejor conocimiento de su sectores comerciales y para el público general interesado en conocer los comercios de su localidad o provincia y que esté interesado en abrir su propio negocio.

DATAVISION también aprovecha los datos del Grupo Cajamar en conjunción con datos abiertos para crear el apartado '*RUTAS COMUNES*' que proporciona información de utilidad social a las instituciones públicas interesadas en mejorar las infraestructuras de transporte para sus ciudadanos.

Finalmente, el apartado '*¿DÓNDE COMPRAR?*' Ayuda a los usuarios a determinar dónde hacer la compra de la forma más óptima.

La versión online desplegada en servidor del panel de control de DATAVISION puede ser accedida desde datavision.mainware.es.

FUENTES DE DATOS

Para este proyecto los datos más explotados han sido los aportados por el Grupo Cajamar, no obstante, otros datos han sido utilizados para enriquecer las funcionalidades de DATAVISION. A continuación se detallan los datos abiertos utilizados:

- **AEMET:** De la Agencia Española de Meteorología han sido empleados los datos meteorológicos de temperatura media y precipitaciones diarias.
- **DICES:** El Directorio Cartográfico de España y América ha sido empleado para obtener los datos relativos a las coordenadas de cada código postal.
- **idealista:** La información de idealista ha sido de utilidad para extraer el precio medio del metro cuadrado de alquiler y venta de locales en cada código postal.
- **Turismo de Murcia:** La localización de los aparcamientos públicos han sido obtenidas de esta fuente de datos pública.
- **INE:** Del Instituto Nacional de Estadística se han obtenido los datos de afluencia de turistas nacionales y extranjeros.
- **MUTRANS:** La localización de las paradas de autobús y tranvía han sido obtenidas de este portal.
- **Íñigo Flores:** El repositorio público "*ds-codigos-postales*" que ofrece Íñigo Flores ha sido de utilidad para encontrar las coordenadas de los polígonos que conforman los perímetros de los códigos postales.
- **Datos Abiertos de la Región de Murcia:** De este portal se ha obtenido la información referente a los centros comerciales de la Región de Murcia. También han sido obtenidos de este portal los datos de localización de restaurantes en la ciudad de Murcia.

En algunos casos, los datos han sido obtenidos mediante una API REST de fácil acceso, en otros los datos estaban disponibles para descarga en su totalidad, etcétera. No obstante, en cualquier caso, debido a la falta de estandarización, para optimizar la carga de datos y para hacer un tratamiento mejor de los datos, todos los datos de las distintas plataformas han sido descargados, post-procesados e introducidos en una base de datos común desde donde se realizan los cálculos necesarios y desde donde son accedidos por el panel de DATAVISION.

Gracias a este método, el acceso y sobre todo la aplicación de cálculos sobre millones de datos se realiza a gran velocidad en un servidor remoto y sin depender de la potencia del ordenador utilizado para acceder a DATAVISION ni de la conexión a internet excepto para descargar los resultados finales de los cálculos.

ANÁLISIS GENERAL

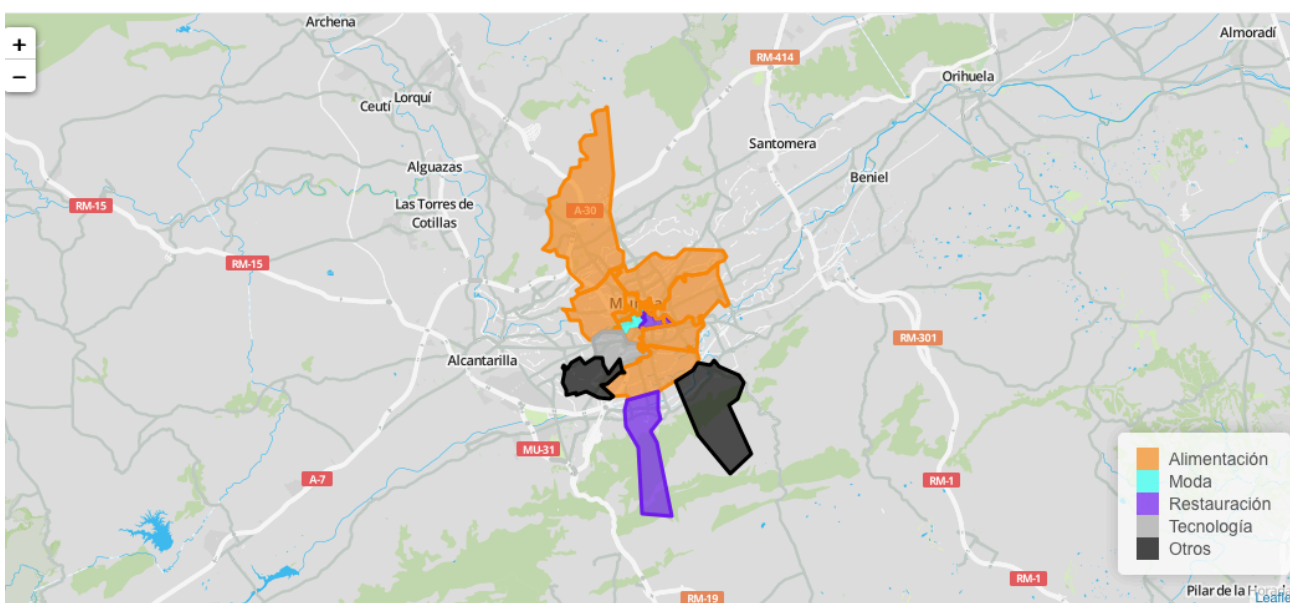
El objetivo principal de este apartado está alineado con el objetivo general de DATAVISION, acercar los datos y su representación visual al público para poder extraer conclusiones útiles para la gestión de negocios actuales, determinar el tipo y la localización de nuevos negocios o simplemente para conocer mejor el estado del sector terciario de su localidad o provincia.

En esta página se incluye un mapa para ayudar a visualizar el sector comercial dominante y el segundo más dominante en cada uno de los 16 códigos postales de la ciudad de Murcia. Gracias a esta información el público puede saber qué tipos de negocios acumulan la mayor parte de los ingresos anuales, y las instituciones públicas pueden hacerse una idea general de la composición de sus comercios y así mejorar sus políticas de ayudas, de recaudación de impuestos, impulsar sectores minoritarios y fomentar la distribución equitativa de los establecimientos de cada sector a lo largo de las áreas metropolitanas, zonas residenciales y pueblos para atender las necesidades de los ciudadanos.

SECTOR DOMINANTE

El siguiente mapa muestra el sector comercial dominante y segundo más dominante (en facturación total) en cada uno de los 16 códigos postales de la ciudad de Murcia.

Sector dominante ☒ Segundo sector dominante

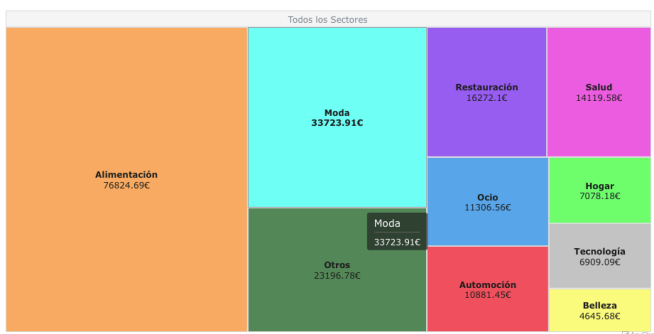


ANÁLISIS GENERAL

Conocer el origen de los ingresos de un sector es de vital importancia para conocer mejor este sector. La distribución por sexos, edad o renta es muy importante a la hora de entender la dinámica de ciertos sectores y por ende adaptar las estrategias comerciales y adaptar las políticas públicas para incentivar el consumo en ciertos sectores. Para ayudar a conocer este tipo de datos demográficos se muestra el apartado Distribución Demográfica por Sectores.

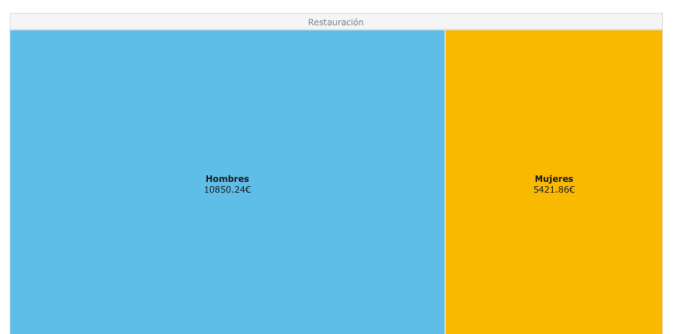
DISTRIBUCIÓN DEMOGRÁFICA POR SECTORES

Se muestra la distribución de ingresos por grupos de sexo, edad y poder adquisitivo por cada sector.



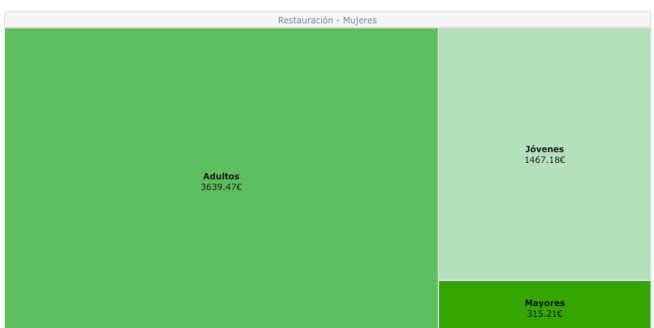
DISTRIBUCIÓN DEMOGRÁFICA POR SECTORES

Se muestra la distribución de ingresos por grupos de sexo, edad y poder adquisitivo por cada sector.



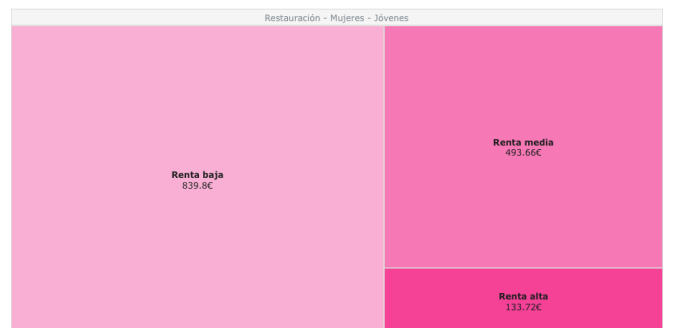
DISTRIBUCIÓN DEMOGRÁFICA POR SECTORES

Se muestra la distribución de ingresos por grupos de sexo, edad y poder adquisitivo por cada sector.



DISTRIBUCIÓN DEMOGRÁFICA POR SECTORES

Se muestra la distribución de ingresos por grupos de sexo, edad y poder adquisitivo por cada sector.



ANÁLISIS GENERAL

Seguidamente, también se le ofrece la posibilidad al usuario de ver los detalles demográficos del público responsable de la media de ingresos a lo largo de los dos años de los que tenemos datos así como la afluencia de turistas españoles y extranjeros. Ver en qué momento del año se deben intensificar estos esfuerzos de captación y fidelización para turistas y para cada uno de los sexos, edades y poderes adquisitivos es de vital importancia para maximizar los ingresos a lo largo del año.

Para ayudar a visualizar estos datos se muestra una gráfica que muestra el promedio de ingresos mensuales en los sectores seleccionados, qué porcentaje corresponde a mujeres y hombres, cuál a jóvenes, adultos y mayores y cuál a un público de poder adquisitivo bajo, medio y alto y los datos de afluencia de turistas nacionales y extranjeros.

TURISMO Y DISTRIBUCIÓN DEMOGRÁFICA POR MESES

Se muestra la distribución del gasto por meses en cada sector comercial en función de rasgos demográficos de sexo, edad y poder adquisitivo y la afluencia de turistas nacionales y extranjeros.



ANÁLISIS GENERAL

Finalmente, se ofrece una visión de la influencia de los factores meteorológicos en los ingresos de cada sector. Para ello, el usuario selecciona una semana de las 104 comprendidas entre octubre de 2015 y septiembre de 2017 y se muestra el porcentaje de variación de los ingresos de cada sector en cada día de la semana respecto a la media de cada uno de los días de la semana de los dos años de datos.

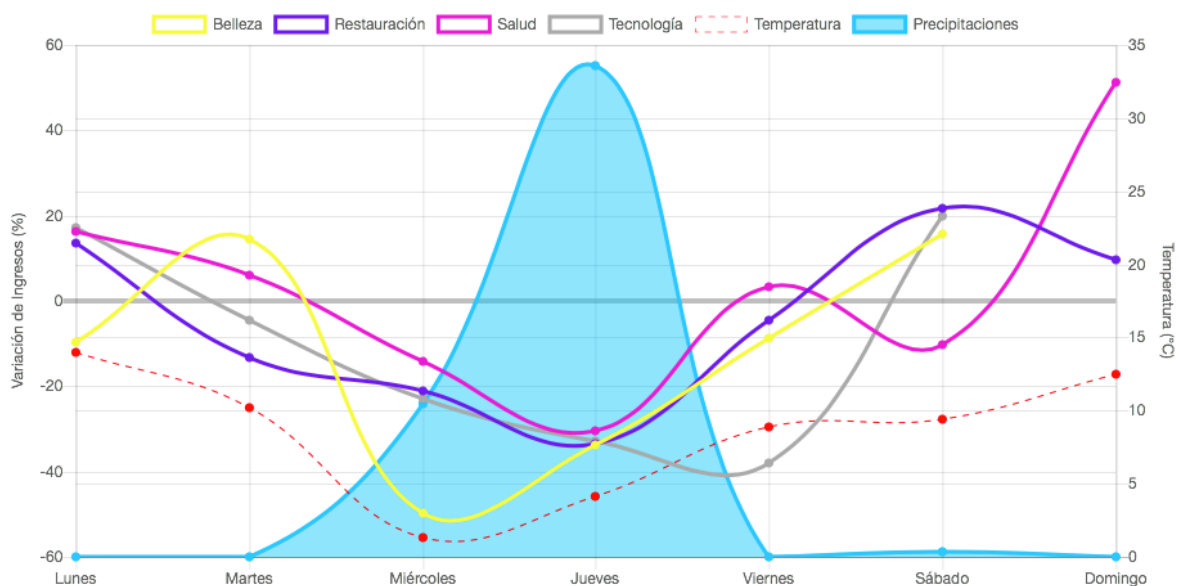
INFLUENCIA DE LA METEOROLOGÍA

En esta sección se muestra la influencia que tiene sobre las ventas de una semana las condiciones climatológicas. En ella se presenta la variación en % (positiva o negativa) de los ingresos de cada día en una determinada semana con respecto a la media de todas las semanas entre octubre de 2015 y septiembre de 2017 así como datos meteorológicos como la precipitación y la temperatura media de cada día.

* Algunos sectores no tienen actividad comercial los domingos, por tanto solo se muestran datos hasta el sábado para estos sectores.

SELECCIONE UNA SEMANA:

2017-01-16



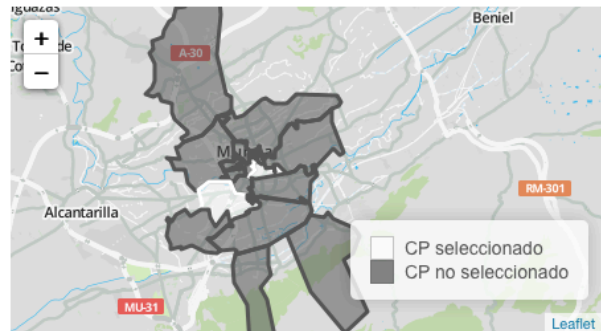
Gracias a este gráfico es fácil ver el efecto que tiene la temperatura media del día y las precipitaciones en los ingresos. En el ejemplo que se presenta se muestra claramente que la bajada de temperaturas y las precipitaciones han tenido un efecto importante en todos los sectores, en especial en belleza, donde los ingresos del miércoles de la semana del 16 de enero de 2017 cayeron sobre un 50% respecto a todos los otros miércoles, etc.

OPTIMIZACIÓN

Para los interesados en mejorar los resultados de sus negocios actuales, DATAVISION ofrece la página 'OPTIMIZACIÓN', que orienta al usuario a la hora de tomar decisiones comerciales que aumenten los ingresos de su negocio. Para ello, el usuario selecciona el/los sector(es) y el/los código(s) postal(es) en los que está interesado.

Seleccione uno o más sectores y códigos postales.

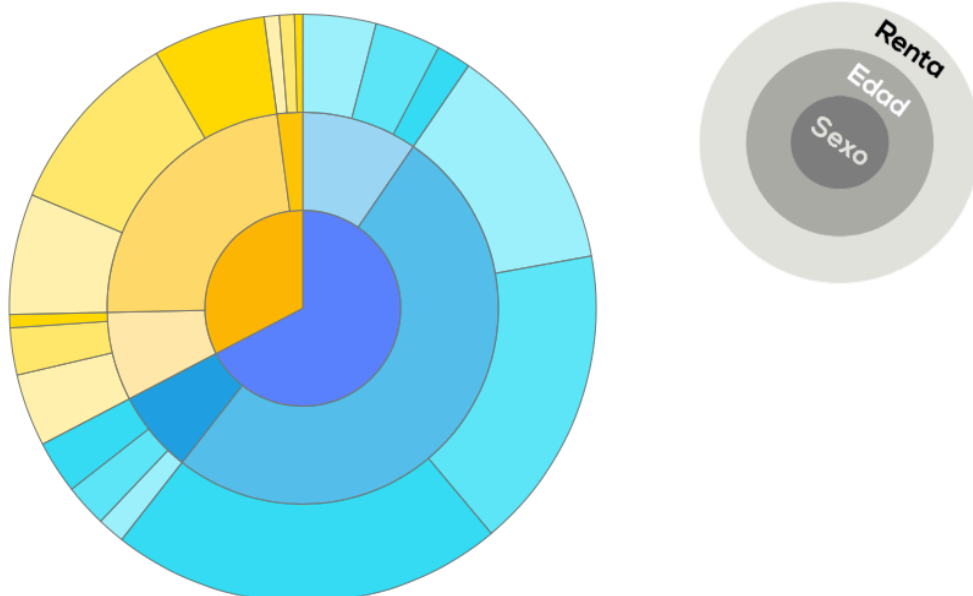
- | | |
|--|---|
| <input type="checkbox"/> Alimentación | <input type="checkbox"/> Automoción |
| <input type="checkbox"/> Belleza | <input checked="" type="checkbox"/> Hogar |
| <input type="checkbox"/> Moda | <input checked="" type="checkbox"/> Ocio |
| <input checked="" type="checkbox"/> Restauración | <input type="checkbox"/> Salud |
| <input type="checkbox"/> Tecnología | <input type="checkbox"/> Otros |



Seguidamente, teniendo en cuenta los parámetros seleccionados se muestra un gráfico de distribución demográfica de los ingresos. Como se expone en el apartado anterior, conocer los detalles demográficos de los clientes de un negocio es básico para adaptar las estrategias de *marketing*, el catálogo de productos y servicios a ofrecer, etc.

DISTRIBUCIÓN DEMOGRÁFICA DE INGRESOS

A continuación se muestra la distribución demográfica del gasto de los consumidores en función de su sexo, edad, renta y combinaciones de las mismas.

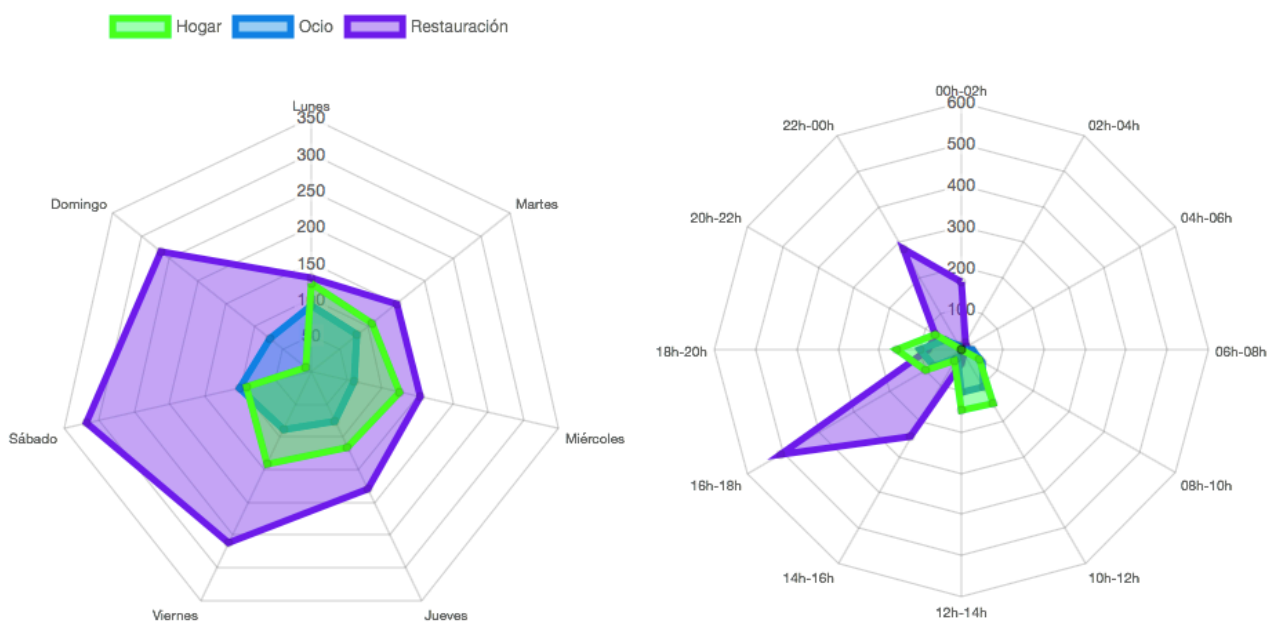


OPTIMIZACIÓN

A continuación, al usuario se le muestra una distribución temporal de la media de los ingresos obtenidos por los establecimientos de los sectores y códigos postales seleccionados. A la izquierda se muestra la variación de los ingresos en función del día de la semana, gracias a esta información se puede determinar qué día de la semana son mayores los ingresos y qué días menores, y tomar decisiones comerciales en consecuencia, como cerrar el establecimiento ciertos días para reducir costes, aumentar las promociones otros días, etc. A la derecha se muestra la distribución de ingresos por franjas horarias de dos horas, de lo que se pueden extraer conclusiones como cambiar la hora de la comida de los empleados, cambiar los horarios de apertura y cierre del establecimiento, la introducción de prácticas de promoción como "Happy Hour", etc.

DISTRIBUCIÓN TEMPORAL DE INGRESOS

Los siguientes gráficos muestran el promedio de ingresos a lo largo del tiempo. Quedan reflejados los ingresos a lo largo del día a la derecha y de la semana a la izquierda.



Muchas veces los esfuerzos de promoción y federación fracasan por ser aplicados en momentos poco adecuados, y la clave del éxito reside en aplicar las promociones en los momentos en los que esa promoción es necesaria y no en momentos donde ya hay clientes de sobra. Para todo esto, DATAVISION ofrece una herramienta de gran utilidad.

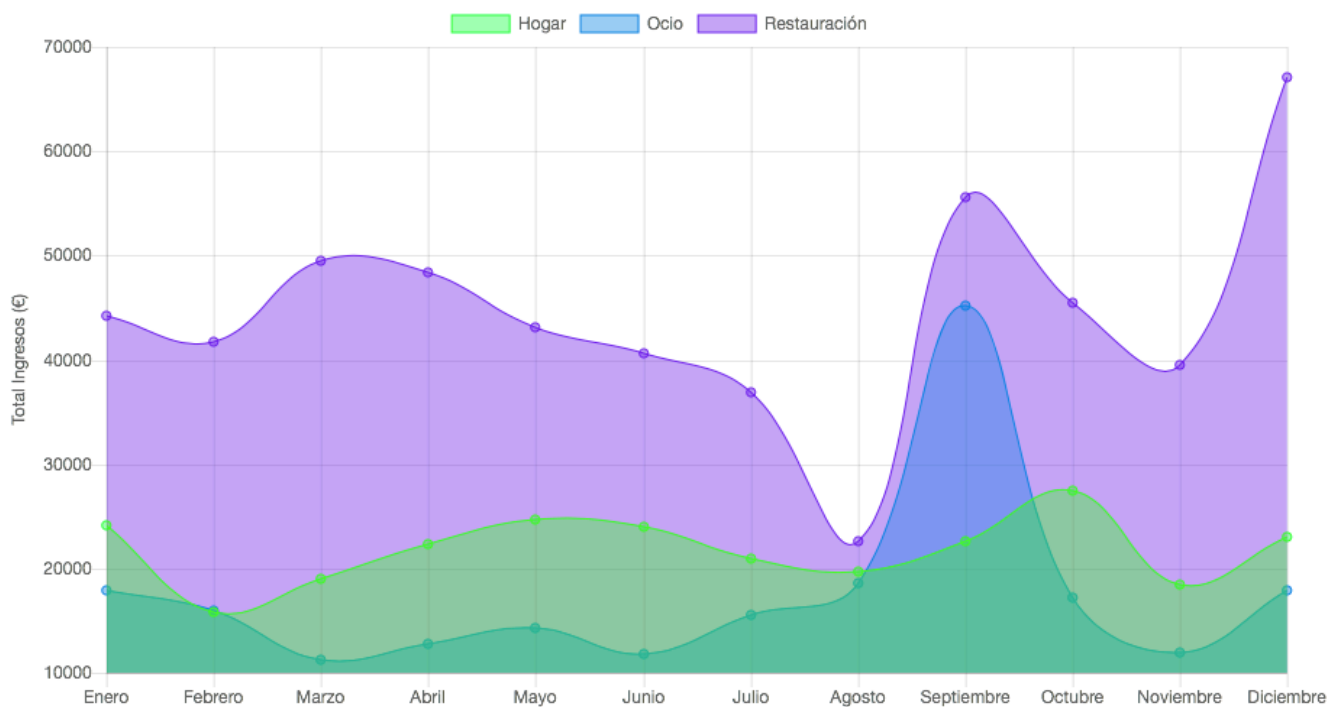
OPTIMIZACIÓN

Finalmente, para la sección de optimización se muestran datos relativos a los ingresos obtenidos en los sectores y códigos postales seleccionados en cada mes del año.

Algunos sectores aumentan o disminuyen sus ingresos en ciertas épocas del año y otros sectores en otros momentos del año. Además, ciertos sectores se ven más afectados por estos efectos que otros, teniendo que tener en cuenta esta temporalidad en sus estrategias de negocio. Negocios que vean sus ingresos muy reducidos en verano (como el sector moda) o en navidades (como el sector hogar) pueden tener esto en cuenta para aumentar sus esfuerzos de promoción o conceder vacaciones a sus empleados en estas fechas. Asimismo, los negocios que tengan un aumento considerable de ingresos en verano (como el sector ocio) o en navidades (como en restauración) pueden contemplar la opción de incrementar su plantilla o incrementar su inventario en estas fechas.

DISTRIBUCIÓN ANUAL DE INGRESOS

Se muestra el promedio de ingresos obtenido mensualmente por cada sector comercial.

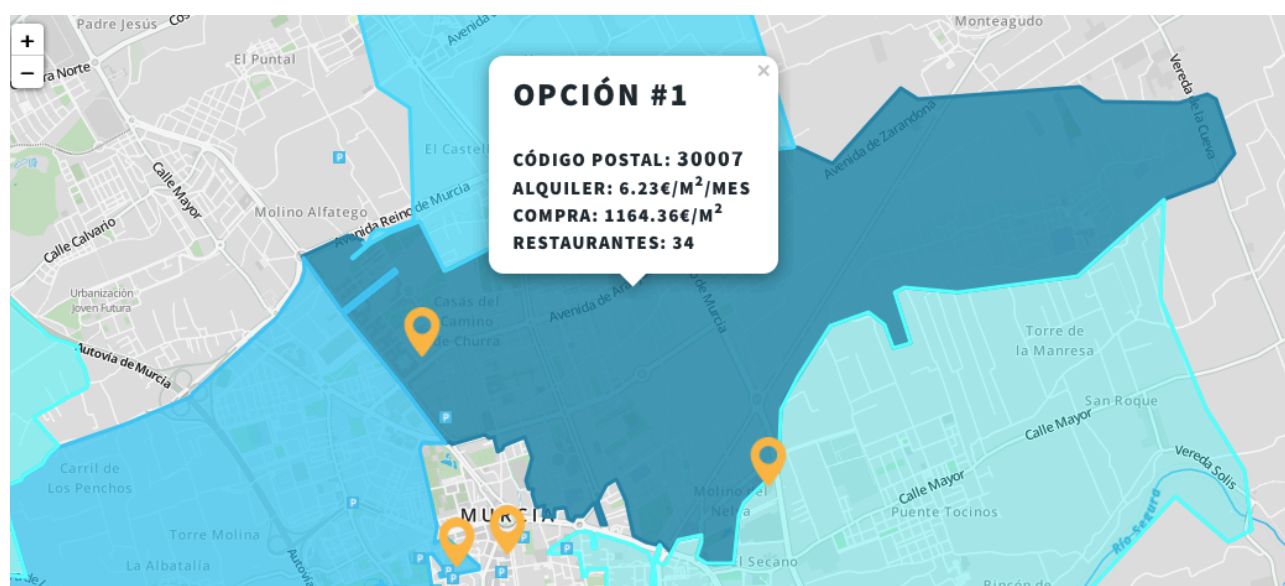


NUEVO NEGOCIO

En este apartado se pretende atender las necesidades de aquellos usuarios que pretenden abrir un nuevo establecimiento, sea su primer negocio, o sea con la idea de expandir su negocio con otro establecimiento.

Para ello, indicando el sector del nuevo establecimiento y a qué público irá dirigido en función de su poder adquisitivo, DATAVISION le mostrará los códigos postales más indicados para abrir el nuevo establecimiento.

El sistema se basa en una serie de parámetros para determinar cuáles son los códigos postales más indicados. La heurística seguida es la de determinar en qué códigos postales vive más gente del poder adquisitivo seleccionado que acostumbran a desplazarse más a otros códigos postales para hacer compras en el sector seleccionado.



El mapa, además de mostrar los códigos postales más indicados para abrir un establecimiento de las características seleccionadas también muestra el precio medio del metro cuadrado para el alquiler y la venta de locales y el número de restaurantes en el código postal seleccionado así como la situación de los centros comerciales y aparcamientos públicos. Información que puede ser útil a la hora de determinar dónde abrir un nuevo establecimiento. La cercanía a centros comerciales y aparcamientos y coste que supondrá alquilar o comprar un local pueden hacer que la decisión acerca de dónde establecerse sea más fácil, fundamentada y correcta.

RUTAS COMUNES

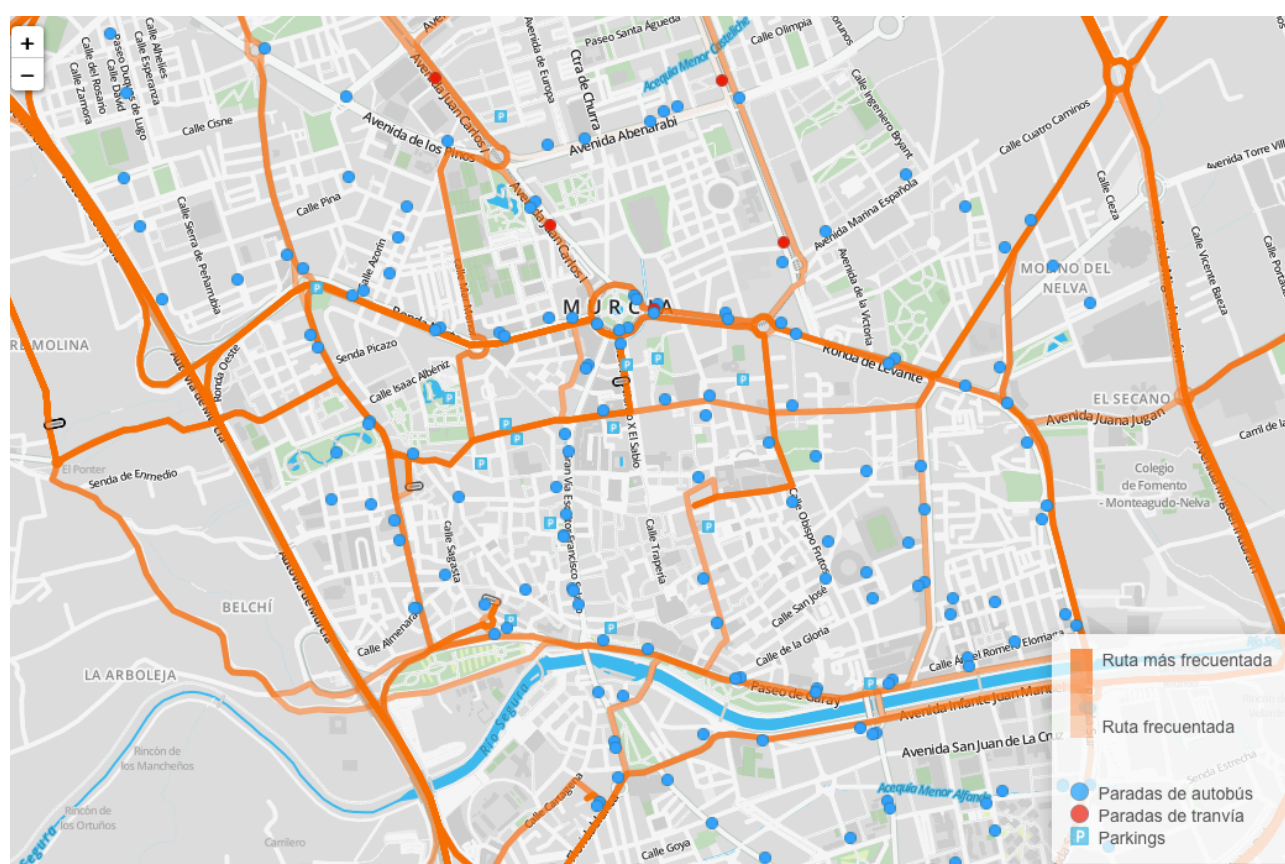
También haciendo uso de los datos aportados por el Grupo Cajamar, DATAVISION pretende ser una herramienta de utilidad social. Para ello, el apartado "RUTAS COMUNES" representa en un mapa las rutas más comunes.

La heurística seguida es la de seleccionar las N rutas más frecuentadas por los consumidores a la hora de desplazarse desde el código postal de su residencia habitual hasta el código postal donde realiza compras dentro de la franja horaria seleccionada por el usuario, siempre y cuando esta ruta supere los D kilómetros. Donde N y D son parámetros configurables por el usuario. De esta forma, con el parámetro D , el usuario puede obviar rutas de pequeño recorrido intraurbano y centrarse en las de mayor recorrido, y con el parámetro N puede filtrar para ver solo un pequeño número de rutas que representan las más concurridas.



RUTAS COMUNES

Además de las rutas más comunes, el mapa también muestra los aparcamientos públicos, las paradas de autobús y de tranvía en la ciudad de Murcia.

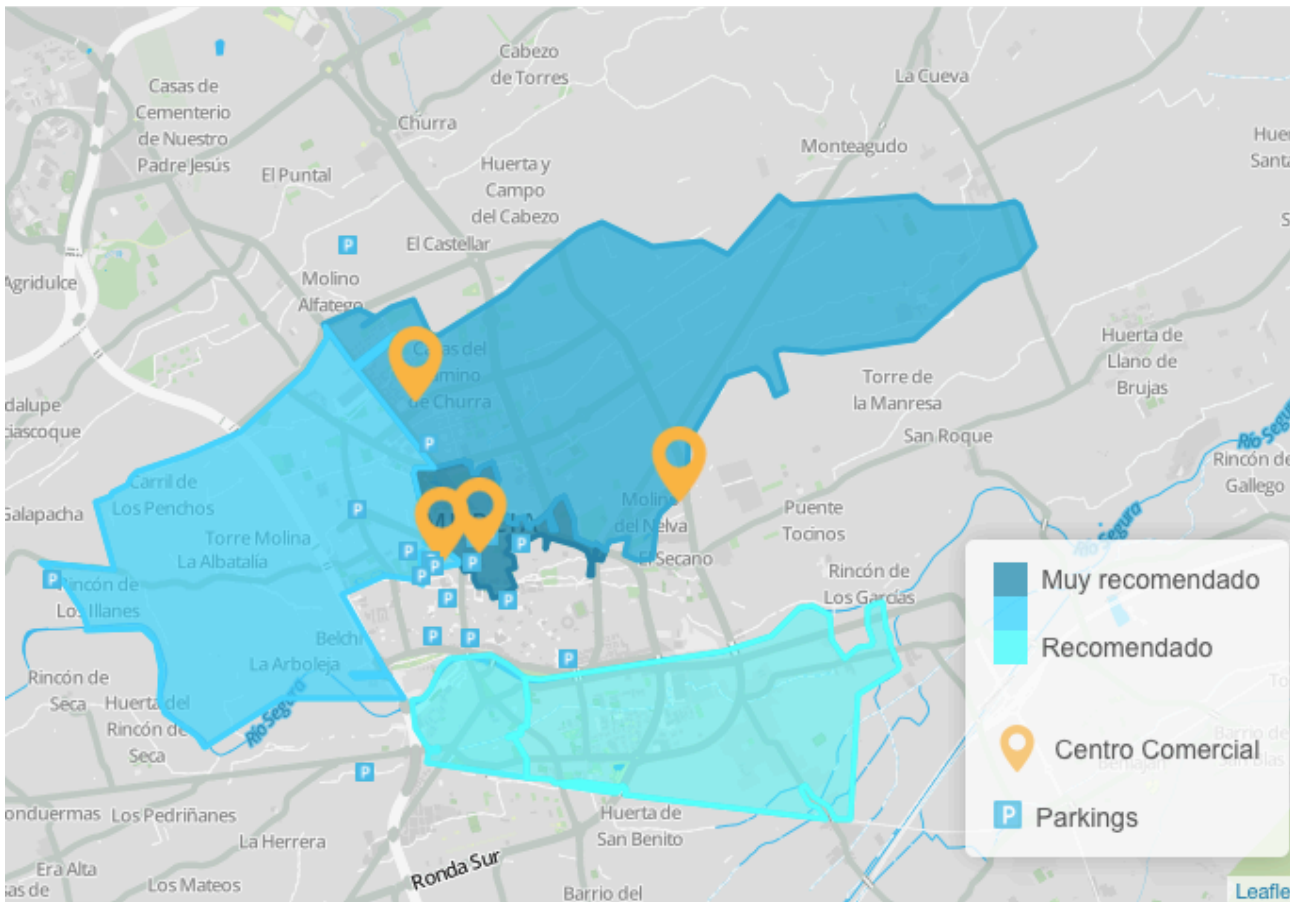


Toda esta información puede ser de gran utilidad en muchos sentidos, no solo beneficia al sector privado que puede determinar dónde establecerse si considera que el paso de clientes por su localización es importante, como es el caso de las gasolineras, panaderías, etcétera, o para empresas de transporte que pueden utilizar esta información para determinar si reforzar estas rutas si transportan personas o evitarlas si se dedican al transporte de mercancías.

Asimismo, las instituciones públicas pueden ser las más beneficiadas de esta información, pues pueden mejorar las infraestructuras de comunicación vial en las carreteras más concurridas así como reforzar el transporte público en las zonas más transitadas ayudando así a los consumidores a ahorrar tiempo y dinero, y beneficiando a la totalidad del municipio reduciendo el tráfico y la contaminación acústica y aérea.

¿DÓNDE COMPRAR?

Esta sección está especialmente pensada para aquellos usuarios que quieren optimizar el uso de su tiempo en el día a día a la hora de hacer la compra. Esta página determina en qué código postal es más recomendable realizar la compra para poder adquirir los productos de todos los sectores seleccionado en un solo código postal.



Gracias a esta página un poco de optimización basada en evidencia puede ser aplicada en el día a día de los usuarios que busquen el código postal ideal que les ofrezca todos los servicios que buscan.

POSIBLES MEJORAS

Desde el equipo de DATAVISION creemos que el proyecto descrito es de gran utilidad para el público general, tanto empresas privadas, como a individuales pensando en crear un nuevo negocio pasando por las instituciones públicas. No obstante, también somos conscientes de que como en cualquier proyecto, existen posibles mejoras y ampliaciones. A continuación se describen posibles mejoras o ampliaciones que planteamos para DATAVISION:

- En cuanto a ampliaciones, obviamente cualquier ampliación de funcionalidades enfocadas a ayudar en la toma de decisiones y en la visualización de información siempre es bienvenida.
- También es una ampliación de interés la creación de nuevos gráficos que representen más datos de una forma nueva.
- En esta versión de DATAVISION, se ha incluido información del transporte público de la ciudad de Murcia, se plantea la adición de más información similar como puede ser la relativa al transporte público de Lorca u otras ciudades siempre y cuando estos datos se complementen con una ampliación de los datos bancarios a toda la Región de Murcia.
- También creemos que sería buena idea añadir más datos meteorológicos como viento, temperaturas máximas y mínimas y añadir nuevas representaciones de estos datos que ayuden a ser relacionados con datos comerciales, de transporte, etc.
- Por otro lado, consideramos que la mejora de las heurísticas utilizadas para la recomendación de códigos postales para establecer nuevos negocios puede ser mejorada y ampliada analizando los datos de una forma distinta y añadiendo nuevos factores a la ecuación.
- En esta línea también, mejorar y ampliar la heurística que determina las rutas a reforzar con transporte público e infraestructuras viales.
- Finalmente, las heurísticas empleadas para determinar el código postal más recomendado para hacer la compra también puede ser mejorado con más datos.

Estas son solo algunas de las posibles mejoras que consideramos para DATAVISION. La posibilidad de crear todo tipo de nuevas funcionalidades y características siempre está abierta.