

El Guachinche

LUIS ARDÉVOL MESA, CARLOS MARTÍNEZ GARCÍA, MIGUEL MATO MARTÍNEZ

CONTENTS

1	Introducción	1
2	Diseño y construcción del almacén de datos	2
A	Diseño del almacén de datos	2
B	Creación de los cubos de datos	8
3	Extracción, transformación y carga de datos	10
4	Análisis de datos con MDX y ROLAP	10
A	Consultas MDX	10
B	Consultas ROLAP	10
5	Cuadros de mando e informes	10

1. INTRODUCCIÓN

Los *guachinches* son establecimientos (en la propia vivienda del propietario) que surgen en la zona del norte del Tenerife como solución para vender el excedente de vino de cada cosecha. El vino se acompaña típicamente con platos tradicionales de la cocina canaria. Actualmente, según el Decreto 83/2013¹, hay una regulación muy estricta amparando a estos locales y al viticultor; muchos establecimientos comúnmente denominados “guachinches”, entran dentro de la categoría *restaurante*, al salirse de lo establecido en este decreto.

El Guachinche nace de una idea sencilla pero atractiva: traer el espíritu de los tradicionales guachinches canarios al resto de España, pero con un toque especial. En nuestra cocina, las recetas isleñas comparten protagonismo con una cuidada selección de arepas para todos los públicos, todo esto acompañado de los mejores vinos de las Islas. Es ese equilibrio entre tradición y adaptación lo que nos hace únicos: cada local incorpora platos de la región, porque creemos que la gastronomía local es parte fundamental de nuestra historia.

No obstante, quien entra a El Guachinche no solo viene a comer. Para dar una experiencia lo más cercana posible y crear una atmósfera atractiva y familiar para los comensales, nuestros locales presentan una ambientación tradicional y hogareña. La experiencia no sería completa de no ser por nuestro personal, siempre cercano.

Actualmente, El Guachinche cuenta con 13 locales, repartidos en Canarias (7), Galicia (2), Andalucía (2) y Comunidad Valenciana (2). El crecimiento en los últimos dos años ha sido exponencial, por lo que desde la dirección de la empresa nos vemos en la necesidad de implementar un modelo de inteligencia de negocio que

¹<https://www.gobiernodecanarias.org/boc/2013/153/001.html>

nos permita tomar decisiones basadas en datos, priorizando siempre al cliente.

Los objetivos principales perseguidos por la cadena son los siguientes:

- Fidelización de clientes. En esta línea, sería conveniente que la economía de cada local no dependiera de la nueva clientela, ya que la fuente de ingresos sería menos estable.
- Mejorar la experiencia general del cliente en nuestros locales.
- Adaptar la oferta gastronómica a las preferencias de cada región.
- Ampliar beneficios, especialmente en las regiones menos rentables actualmente.
- Estudio de la rentabilidad del servicio a domicilio.

El servicio a domicilio corre a cargo de dos empresas: Uber Eats y Glovo. Conocida la comisión de estos servicios, cada local descuenta esta comisión del total y almacena el beneficio real de cada pedido a domicilio. Para cumplir con los objetivos, es importante conocer:

- La cantidad de nuevos clientes en comparación a los habituales.
- La satisfacción de los clientes con los servicios prestados en nuestros establecimientos.
- El rendimiento económico de cada uno de los locales, así como de la totalidad de la cadena. Esto incluye:
 - Ventas de cada producto ofertado.
 - Gastos e ingresos de cada local. En estos últimos sería necesario distinguir entre ingresos presenciales e ingresos a domicilio.

A continuación, se describen los procesos de diseño y construcción del almacén de datos para gestionar la información necesaria. Posteriormente se comentarán los procesos de extracción, transformación y carga que permiten la creación y actualización del almacén de datos. Se sigue con algunas consultas de interés para los objetivos descritos anteriormente y, para terminar, se incluyen cuadros de mando e informes adecuados al caso de estudio.

2. DISEÑO Y CONSTRUCCIÓN DEL ALMACÉN DE DATOS

Nuestros locales llevan un registro cuidadoso de los aspectos más importantes de la actividad diaria, como son las ventas, los gastos o la opinion de los clientes. Cada local genera reportes mensuales que recogen estos datos y son enviados a la central para su análisis. Estos reportes son los que nos permitirán obtener la información necesaria para la toma de decisiones.

A. Diseño del almacén de datos

Para realizar el diseño del almacén de datos, es necesario conocer primero la estructura de los datos recogidos en los reportes de cada local. Concretamente, cada mes se reciben dos archivos de datos por cada local de nuestra cadena. El primero de ellos contiene toda la información relacionada con la economía del restaurante: desglose de gastos en cuatro categorías (personal, suministros, alquiler y otros), ingresos por ventas (diferenciando entre los servicios presencial y domicilio) y clientela. Esto será beneficio para varios de los objetivos de la cadena. Siendo un poco más específicos, se proporcionan los siguientes datos:

- Un identificador entero para cada local, generado por la central en orden de apertura.
- La fecha de envío del reporte (establecida para el primer día del mes).
- Los gastos en alquiler, personal, suministros (proveedores), y extras (limpieza, mantenimiento, etc.), todo en euros.

- Los ingresos, tanto por vía presencial como a domicilio, en euros. Como ya se ha mencionado, estos ingresos ya tienen descontada la comisión de Uber Eats o Glovo.
- Los clientes totales por ambos canales, así como el número de clientes nuevos.
- Los platos ofertados en el local, cada uno con los siguientes datos:
 - Un identificador entero para cada plato, generado por la central.
 - El nombre, precio y número de ventas del plato.

Estos datos se recogen en un archivo XML como el que se muestra en la figura 1.

```
<datos>
  <finanzas id="1" emision="2021-02-01">
    <alquiler>1300.00</alquiler>
    <personal>12000.00</personal>
    <proveedores>3254.40</proveedores>
    <extra>1045.11</extra>
    <ingresos_presencial>25911.13</ingresos_presencial>
    <ingresos_domicilio>1741.15</ingresos_domicilio>
    <numero_clientes_presencial>1798</numero_clientes_presencial>
    <nuevos_clientes_presencial>129</nuevos_clientes_presencial>
    <numero_clientes_domicilio>188</numero_clientes_domicilio>
    <nuevos_clientes_domicilio>73</nuevos_clientes_domicilio>
    <platos>
      <plato id="1" nombre="Almogrote Gomero" precio="5.00" ventas="277" />
      <plato id="2" nombre="Papas arrugadas con mojo" precio="6.50" ventas="337" />
      <plato id="3" nombre="Queso asado con mojo" precio="7.00" ventas="19" />
      <plato id="4" nombre="Escaldon" precio="7.00" ventas="131" />
      <plato id="5" nombre="Ropa vieja" precio="9.00" ventas="206" />
      <plato id="6" nombre="Costilla con papas y pina" precio="10.00" ventas="160" />
      <plato id="7" nombre="Carne fiesta" precio="9.50" ventas="228" />
      <plato id="8" nombre="Quesillo" precio="4.50" ventas="388" />
      <plato id="9" nombre="Bienmesabe" precio="4.50" ventas="470" />
      <plato id="10" nombre="Arepas reina pepiada" precio="4.00" ventas="504" />
      <plato id="11" nombre="Arepas pabellon" precio="4.00" ventas="274" />
      <plato id="12" nombre="Arepas full equipo" precio="4.00" ventas="183" />
      <plato id="13" nombre="Arepas veganas" precio="4.00" ventas="71" />
      <plato id="14" nombre="Arepas blancas" precio="4.00" ventas="503" />
      <plato id="15" nombre="Vino tinto canario" precio="5.00" ventas="93" />
      <plato id="16" nombre="Vino blanco canario" precio="5.00" ventas="256" />
      <plato id="17" nombre="Agua" precio="1.50" ventas="688" />
      <plato id="18" nombre="Cola" precio="2.00" ventas="159" />
      <plato id="19" nombre="Limon" precio="2.00" ventas="248" />
      <plato id="20" nombre="Naranja" precio="2.00" ventas="926" />
      <plato id="21" nombre="Aquarius" precio="2.00" ventas="547" />
      <plato id="22" nombre="Nestea mangopina" precio="2.00" ventas="435" />
    </platos>
  </finanzas>
</datos>
```

Fig. 1. Ejemplo de archivo XML con los datos económicos mensuales de un local.

Los identificadores de los locales son los siguientes:

- 1: La Laguna, Tenerife.
- 2: Hermigua, La Gomera.
- 3: Cotillo, Fuerteventura.
- 4: Las Palmas de Gran Canaria, Gran Canaria.
- 5: Tazacorte, La Palma.
- 6: Valverde, El Hierro.
- 7: Arrecife, Lanzarote.
- 8: Granada, Granada.
- 9: Sevilla, Sevilla.
- 10: Santiago de Compostela, A Coruña.
- 11: Vigo, Pontevedra.
- 12: Alicante, Comunidad Valenciana.
- 13: Valencia, Comunidad Valenciana.

Los identificadores de los productos son los siguientes:

- 1: Almogrote Gomero.
- 2: Papas arrugadas con mojo.
- 3: Queso asado con mojo.
- 4: Escaldón.
- 5: Ropa vieja.
- 6: Costilla con papas y piña.
- 7: Carne fiesta.
- 8: Quesillo.
- 9: Bienmesabe.
- 10: Arepa reina pepiada.
- 11: Arepa pabellón.
- 12: Arepa full equipo.
- 13: Arepa vegana.
- 14: Arepa blanca.
- 15: Vino tinto canario.
- 16: Vino blanco canario.
- 17: Agua.
- 18: Refresco de cola.
- 19: Refresco de limón.
- 20: Refresco de naranja.
- 21: Aquarius.
- 22: Nestea mango-piña.
- 23: Salmorejo.
- 24: Pescaito frito.
- 25: Gambitas de Huelva.
- 26: Pestiños.
- 27: Vino tinto andaluz.
- 28: Vino blanco andaluz.
- 29: Pulpo a feira.
- 30: Empanada (porción).
- 31: Lacon con grelos.
- 32: Tarta de Santiago.
- 33: Vino tinto gallego.
- 34: Vino blanco gallego.
- 35: Paella.
- 36: Arroz negro.
- 37: Esgarraet.
- 38: Fartons.
- 39: Vino tinto valenciano.
- 40: Vino blanco valenciano.

Como bien describe el objetivo del negocio, cada región tiene sus platos típicos. Los 22 primeros productos son comunes a todos los locales de la cadena, mientras que a partir de ahí, cada región incorpora distintos productos, como es el caso del *salmorejo* en Andalucía, el *lacçon con grelos* en Galicia o el *esgarraet* en la Comunidad Valenciana. Así mismo, cada local ofrece vinos de proximidad.

Para cumplir los objetivos relacionados con la satisfacción y fidelización de clientes, cada local proporciona un segundo archivo, en este caso en formato CSV, con valoraciones de los clientes: en caso de ser clientes presenciales, se valora el ambiente del local, el personal y la calidad de la comida, mientras que los clientes a domicilio simplemente hacen llegar una valoración del servicio general. Así, cada fila del archivo da el identificador del local, la fecha de la valoración, y las votaciones correspondientes (3 en el caso de clientes presenciales, una en el caso de clientes a domicilio). Un ejemplo de este archivo se muestra en la figura 2.

```

restaurante,fecha,valoracion_ambiente,valoracion_personal,valoracion_comida
1,2021-02-01,3.4,4.3,4.7
1,2021-02-01,3.8,3.3,4.1
1,2021-02-01,4.6,3.8,4.7
1,2021-02-01,3.7,3.5,4.6
1,2021-02-01,4.8,4.6,3.4
1,2021-02-01,3.8,4.4,4.3
1,2021-02-01,4.7,4.1,4.5
1,2021-02-01,4.8,4.9,4.4
1,2021-02-01,4.4,3.5,3.8
1,2021-02-01,3.7,5.0,4.5
1,2021-02-01,4.5,4.0,4.6
1,2021-02-01,4.3,4.0,4.2
1,2021-02-01,4.4,4.6,4.8
1,2021-02-01,4.3,4.3,4.0
1,2021-02-01,3.3,4.1,4.7

```

Fig. 2. Ejemplo de archivo CSV con las valoraciones de los clientes de un local.

Se dará a los datos almacenados de cada uno de los archivos anteriores la máxima granularidad posible, es decir, mensual. Las dimensiones seleccionadas son las siguientes:

- **Tiempo:** la dimensión temporal de los datos se representa mediante dos niveles: año y mes.
- **Restaurante:** la dimensión restaurante actúa como dimensión geográfica. Se tendrán en cuenta todos los locales de la cadena y se mostrarán dos niveles, uno para el país y otro para la ciudad.
- **Producto:** la dimensión producto recoge todos los platos ofertados en los locales de la cadena. Esta es una dimensión plana que muestra los productos y el precio asociado a cada uno de ellos.

Los hechos seleccionados para el almacén de datos son los siguientes:

- **Finanzas:** se almacenan los costes e ingresos de los locales, junto con los datos de los clientes para cada combinación de las siguientes dimensiones: tiempo (mes de envío de los reportes) y restaurante (mediante el identificador de cada local).
- **Producto:** se almacenan los hechos relacionados con las ventas de los productos ofertados en los locales, para cada combinación de las siguientes dimensiones: tiempo (mes de envío de los reportes), restaurante (mediante el identificador de cada local) y producto (mediante el identificador de cada plato). Se tiene el número total de ventas mensuales de un producto, así como un cálculo de los ingresos generados por ese producto, a partir de sus ventas y su precio.
- **Feedback:** se almacenan las valoraciones de los clientes para cada combinación de las siguientes dimensiones: tiempo (mes de envío de los reportes) y restaurante (mediante el identificador de cada local). Se almacena la media de las valoraciones de los clientes para cada uno de los aspectos recogidos en los reportes.

Por tanto, el modelo conceptual del almacén de datos se muestra en la figura 3, siendo un esquema con tres estrellas y tres dimensiones.

Los detalles de la implementación del almacén de datos se describen en la siguiente tabla:

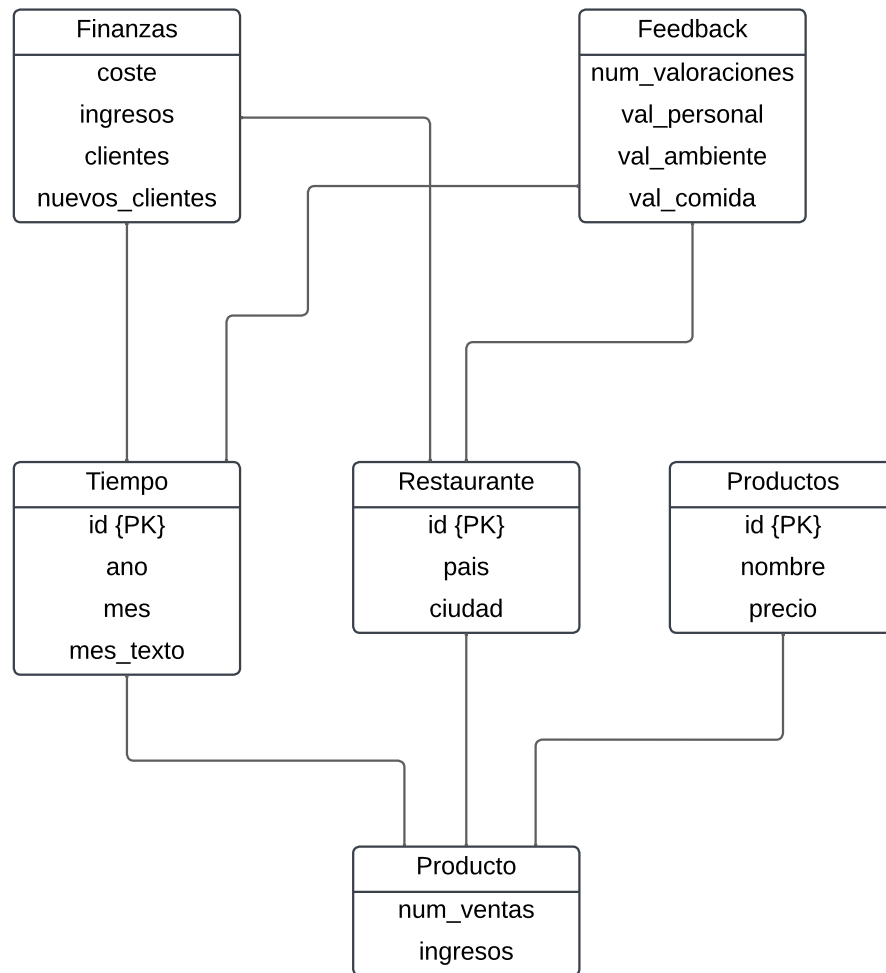


Fig. 3. Esquema conceptual del almacén de datos. **Hacer el bueno**

Dimensiones		
-------------	--	--

Tiempo		
Atributo	Tipo	Descripción
id (pk)	Integer	Identificador autoincremental generado por el almacén cada vez que se inserta un nuevo mes.
ano	Integer	Número correspondiente al año.
mes	Integer	Número correspondiente al mes dentro del año.
mes_texto	String	Texto con el nombre del mes.

Restaurante		
Atributo	Tipo	Descripción
id (pk)	Integer	Identificador de cada local, generado por la central.

Hechos		
Finanzas		
Atributo	Tipo	Descripción
alquiler	Numeric	Gasto en alquiler de un local, en euros.
personal	Numeric	Gasto en personal de un local, en euros.
proveedores	Numeric	Gasto en suministros de un local, en euros.
extra	Numeric	Gasto en extras de un local, en euros.
ingresos_presencial	Numeric	Ingresos por ventas presenciales de un local, en euros.
ingresos_domicilio	Numeric	Ingresos por ventas a domicilio de un local, en euros.
numero_clientes_presencial	Integer	Número total de clientes presenciales de un local.
numero_clientes_domicilio	Integer	Número total de clientes a domicilio de un local.
nuevos_clientes_presencial	Integer	Número de clientes presenciales nuevos de un local.
nuevos_clientes_domicilio	Integer	Número de clientes nuevos a domicilio de un local.
Producto		
Atributo	Tipo	Descripción
ventas	Integer	Número total de ventas de un producto en un local.
Feedback		
Atributo	Tipo	Descripción
valoracion_ambiente	Numeric	Valoración promedio (entre 0 y 5 con un decimal) del ambiente de un local, según sus clientes.
valoracion_personal	Numeric	Valoración promedio (entre 0 y 5 con un decimal) del personal de un local, según sus clientes.
valoracion_comida	Numeric	Valoración promedio (entre 0 y 5 con un decimal) de la calidad de la comida de un local, según sus clientes.

Table 1. Borrar esto: Básicamente en cada tabla de estas habrá una fila por cada restaurante y por cada mes y año.

B. Creación de los cubos de datos

Utilizando como base las estructuras de datos relacionales descritas en el apartado anterior, se crea un esquema multidimensional con tres cubos y tres dimensiones compartidas entre ellos. Estas se definen a nivel de esquema y posteriormente se reutilizan a nivel de cubo:

- Dimensión: **Tiempo**. Dimensión de tipo temporal.
 - Jerarquía: *jerarquiaTiempo*. Definida por el atributo *id*.
 - * Nivel: *ano*. Nivel de tipo *TimeYears* definido por el atributo *ano*.
 - * Nivel: *mes*. Nivel de tipo *TimeMonths* definido por el atributo *mes*. Se utiliza el atributo *mes_texto* para nombrar a los elementos de este nivel.
 - * Tabla: tiempo
- Dimensión: **Restaurante**. Dimensión de tipo estándar.
 - Jerarquía: *jerarquiaRestaurantes*. Definida por el atributo *id*.
 - * Nivel: *pais*. Nivel de tipo regular definido por el atributo *pais*.
 - * Nivel: *ciudad*. Nivel de tipo regular definido por el atributo *ciudad*.
 - * Tabla: restaurante
- Dimensión: **Producto**. Dimensión de tipo estándar.
 - Jerarquía: *jerarquiaProductos*. Definida por el atributo *id*.
 - * Nivel: *nombre*. Nivel de tipo regular definido por el atributo *nombre*.
 - * Tabla: producto
- Cubo: **Finanzas**.
 - Tabla: *finanzas*
 - Dimensiones
 - * Dimensión usada: **Tiempo**. Se referencia a la dimensión *Tiempo*. Se utiliza el atributo *fecha* como clave foránea.
 - * Dimensión usada: **Restaurante**. Se referencia a la dimensión *Restaurante*. Se utiliza el atributo *restaurante* como clave foránea.
 - Medidas **MIRAR BIEN PORQUE NO SE AGREGAN ?**
 - * Medida: **Alquiler**. ??
 - * Medida: **Personal**. ??
 - * Medida: **Proveedores**. ??
 - * Medida: **Extra**. ??
 - * Medida: **Ingresos presencial**. ??
 - * Medida: **Ingresos domicilio**. ??
 - * Medida: **Número clientes presencial**. ??
 - * Medida: **Número clientes domicilio**. ??
 - * Medida: **Nuevos clientes presencial**. ??
 - * Medida: **Nuevos clientes domicilio**. ??
 - * Medida calculada: **Gastos totales**. Se genera un nuevo valor para la dimensión *Measures* usando la expresión MDX: “[Measures].[Alquiler] + [Measures].[Personal] + [Measures].[Proveedores] + [Measures].[Extra]”

- * Medida calculada: **Ingresos totales**. Se genera un nuevo valor para la dimensión *Measures* usando la expresión MDX: “[Measures].[Ingresos presencial] + [Measures].[Ingresos domicilio]”
- * Medida calculada: **Beneficio**. Se genera un nuevo valor para la dimensión *Measures* usando la expresión MDX: “[Measures].[Ingresos totales] - [Measures].[Gastos totales]”
- * Medida calculada: **Número total clientes**. Se genera un nuevo valor para la dimensión *Measures* usando la expresión MDX: “[Measures].[Número clientes presencial] + [Measures].[Número clientes domicilio]”
- * Medida calculada: **Número total nuevos clientes**. Se genera un nuevo valor para la dimensión *Measures* usando la expresión MDX: “[Measures].[Nuevos clientes] + [Measures].[Nuevos clientes domicilio]”
- * Medida calculada: **Beneficio por cliente**. Se genera un nuevo valor para la dimensión *Measures* usando la expresión MDX: “[Measures].[Beneficio] / [Measures].[Número total clientes]”

- Cubo: **Producto**.

- Tabla: *producto*
- Dimensiones
 - * Dimensión usada: **Tiempo**. Se referencia a la dimensión *Tiempo*. Se utiliza el atributo *fecha* como clave foránea.
 - * Dimensión usada: **Restaurante**. Se referencia a la dimensión *Restaurante*. Se utiliza el atributo *restaurante* como clave foránea.
 - * Dimensión usada: **Producto**. Se referencia a la dimensión *Producto*. Se utiliza el atributo *producto* como clave foránea.
- Medidas **como trato a precio? lo meto en la dimension producto?**
 - * Medida: **Ventas**. ??
 - * Medida calculada: **Ingresos por producto**. Se genera un nuevo valor para la dimensión *Measures* usando la expresión MDX: “[Measures].[Ventas] * [Producto].[precio]”

- Cubo: **Feedback**.

- Tabla: *feedback*
- Dimensiones
 - * Dimensión usada: **Tiempo**. Se referencia a la dimensión *Tiempo*. Se utiliza el atributo *fecha* como clave foránea.
 - * Dimensión usada: **Restaurante**. Se referencia a la dimensión *Restaurante*. Se utiliza el atributo *restaurante* como clave foránea.
- Medidas
 - * Medida: **Valoración ambiente**. Se agrega el atributo *valoracion_ambiente* usando la funcion AVG.
 - * Medida: **Valoración personal**. Se agrega el atributo *valoracion_personal* usando la funcion AVG.
 - * Medida: **Valoración comida**. Se agrega el atributo *valoracion_comida* usando la funcion AVG.
 - * Medida calculada: **Valoración media restaurante**. Se genera un nuevo valor para la dimensión *Measures* usando la expresión MDX: “([Measures].[Valoración ambiente] + [Measures].[Valoración personal] + [Measures].[Valoración comida]) / 3”

3. EXTRACCIÓN, TRANSFORMACIÓN Y CARGA DE DATOS

4. ANÁLISIS DE DATOS CON MDX Y ROLAP

A. Consultas MDX

B. Consultas ROLAP

5. CUADROS DE MANDO E INFORMES