

UNIVERSIDADE DA CORUÑA

Facultade de Economía y Empresa

Trabajo de
fin de grado

Base de datos y minería
de datos en contornos
empresariales

Autora: Rosana María Blanco Amado

Tutores: Julio González Soto

Miguel Rodríguez Penabad

Grado en (Ciencias Empresariales)

Año 2014

Resumen

El objetivo de este trabajo consiste en demostrar la utilidad que supone tener la información de la empresa almacenada en una base de datos con el propósito de tomar buenas decisiones. En este caso, se crea una base de datos para una distribuidora de productos ecológicos asentada en la Comunidad Autónoma de Galicia, concretamente en la provincia de A Coruña. Esto se debe al inmenso crecimiento tanto del número de productores ecológicos como de las hectáreas dedicadas al cultivo de productos ecológicos en Galicia. La base de datos de esta empresa soporta todo tipo de información sobre su funcionamiento (nombre de productores ecológicos, de clientes, de los productos que se comercializan, fechas de compra y venta entre otros). Con ayuda de estos datos, la dirección o el departamento de gestión de la empresa toman decisiones con el fin de mejorar el funcionamiento de la empresa a la hora de distribuir sus productos.

Palabras clave: SGBD, Base de datos, diseño de base de datos, toma de decisiones, inteligencia de negocio, distribuidora de productos ecológicos, almacén de datos, data warehouse.

Número de palabras: 13.299

Abstract

The objective of this work is to demonstrate the usefulness of having the information stored in a company database in order to make good decisions. In this case, a database for organic products distributor based in the Autonomous Community of Galicia, specifically in the province of A Coruña is created. This is due to the immense growth in both the number of organic producers and area cultivated with organic products in Galicia. The database of this company supports all kinds of operating information (name of organic producers, customers, products sold, dates of purchase and sale among others). Using these data, the direction or management department business decision makers in order to improve the operation of the company in distributing their products.

Keywords: DBMS, database, database design, decision making, business intelligence, distributor of organic products, data warehouse, data warehouse.

Índice

Introducción	7
1. Planificación	9
2. Fundamentos Teóricos.....	14
2.1. La evolución de las bases de datos	14
2.2. Definición de Bases de Datos	15
2.3. Diseño de Bases de Datos.....	17
2.4. Modelo Entidad-Relación (ER)	18
2.4.1. Entidad, tipo de entidad y atributo.....	18
2.4.2. Identificadores	19
2.4.3. Relación y tipo de relación	20
2.4.4. Participación y Cardinalidad.....	22
2.4.5. Entidad Débil.....	23
2.5. Modelo Relacional.....	24
2.5.1. Relación.....	25
2.5.2. Restricciones del Modelo Relacional	26
2.5.3. Transformación del modelo ER al modelo relacional	27
2.6. Implementación física.....	32
2.6.1. Lenguaje SQL.....	32
3. Diseño de la base de datos de la distribuidora de productos ecológicos	34
3.1. Distribuidora de Productos ecológicos	34
3.2. Entidad Relación (Diagrama ER)	37
3.3. Diccionario de datos	39
3.4. Paso a relacional	44
3.5. Implementación con Microsoft Office Access	44
3.5.1. Tablas	45
3.5.2. Relaciones.....	48
3.5.3. Consultas	48
3.5.3.1. Estructura Básica de consultas SQL.....	49
4. Alimentación de la base de datos.....	50
5. Explotación de la información	52
6 Conclusiones.....	64
Bibliografía.....	65
Anexos	67

Índice de Gráficos

Gráfico 1. Duración estimada del trabajo	12
Gráfico 2. Ingresos por ventas según tipo de negocio desde 2012 hasta 2014.....	53
Gráfico 3. Importe por ventas/tipo de negocio según distribución geográfica	55
Gráfico 4.- Nº de Ventas por empleado desde 2012 hasta 2014.	57
Gráfico 5.- Clientes con mayor diferencia media en días en la venta.	58
Gráfico 6.- Cantidad de producto por Productor ecológico (diez mejores).....	59
Gráfico 7.- Cantidad de Productos por ayuntamiento (ambos diez mejores).....	60
Gráfico8.- Productores y diferencia media entre la Fecha real y la estimada en las Compras	61
Gráfico 9.- Mayores ingresos por producto (diez mejores).....	62
Gráfico 10.- Peores ingresos por productos (diez peores).....	63
Gráfico 11.- Productos (diez mejores) por Negocio	63
Gráfico 12.- Puntos de compra preferente de productos ecológicos	67
Gráfico 13.- Medios utilizados para dar a conocer los productos ecológicos	67
Gráfico14.- Evolución del número de productores ecológicos y de la superficie cultivada en Galicia 2001-2012.....	68

Índice de figuras

Figura 1. Relación Reflexiva.....	21
Figura 2. Relación Binaria.....	21
Figura 3. Relación Ternaria.....	21
Figura 4. Ejemplo representación entidad débil.....	23
Figura 5. Representación Clave Ajena.....	27
Figura 6. Relación N:M.....	29
Figura 7. Transformación al modelo relacional a partir de la relación N: M.....	29
Figura 8. Relación 1: N.....	30
Figura 9. Relación 1: N con entidad débil.....	31
Figura 10. Relación N: M : P.....	31
Figura 11. Transformación de una relación N: M : P al modelo relacional.....	32
Figura 12. Diagrama ER.....	38
Figura 13. Creación de la Tabla Cliente.....	46
Figura 14. Propiedades Tabla Cliente.....	46
Figura 15. Ejemplo de clave ajena en la Tabla Cliente.....	47
Figura 16. Ejemplo de filtro “Población”.....	47
Figura 17. Construcción de relaciones tabla Cliente y Población.....	48
Figura 18. Alimentación de la Tabla Productor Ecológico.....	50
Figura 19. Tabla Productor Ecológico.....	50
Figura 20. Tabla Producto.....	51
Figura 21. Tabla Compra.....	51
Figura 22. Transformación del diagrama ER al modelo relacional.....	69
Figura 23. Relaciones de las tablas de la base de datos.....	70

Índice de tablas

Tabla 1. Tabla de esfuerzo estimada	10
Tabla 2. Tabla esfuerzo real	13
Tabla 3. Representación de Entidad Fuerte	20
Tabla 4. Representación de atributos con diferente notación	20
Tabla 5. Representación de Dominios y Relación Empleado	25
Tabla 6. Definición de la entidad “Productores Ecológicos” y sus propiedades	39
Tabla 7. Definición de la entidad entidad “Población” y sus propiedades	39
Tabla 8. Definición de la entidad “Ayuntamiento” y sus propiedades.....	40
Tabla 9. Definición de la entidad “Compra” y sus propiedades	40
Tabla 10. Definición de la entidad “Línea_Compra” y sus propiedades	41
Tabla 11. Definición de la entidad “Productos Ecológicos” y sus propiedades.....	41
Tabla 12. Definición de la entidad “Línea_Venta” y sus propiedades	42
Tabla 13. Definición de la entidad “Venta” y sus propiedades	42
Tabla 14. Definición de la entidad “Empleado” y sus propiedades	43
Tabla 15. Definición de la entidad “Vehículo” y sus propiedades	43
Tabla 16. Definición de la entidad “Cliente” y sus propiedades	44
Tabla 17. Comparación(%)de ingresos de todos los negocios en relación al supermercado	54

Introducción

A lo largo de los años, las empresas se caracterizaron por el manejo de un volumen creciente de información lo que llevó a utilizar técnicas de almacenamiento de datos. Entre estas técnicas ocupan una posición muy importante las bases de datos tanto por la información almacenada como por el posible acceso rápido a ella (De Miguel y Piattini, 1997).

El interés que origina una base de datos es inmenso no sólo por la atención que alcanza en los planes de estudio sino también por la frecuente aparición en libros, proyectos..., lo que nos hace indicar que es una técnica universal (De Miguel y Piattini, 1997), he ahí la razón de llevar a cabo el presente trabajo.

Las bases de datos, como hemos mencionado, se utilizan para el almacenamiento de información y rápido acceso a los datos, y además como soporte para la toma de decisiones. Por ello una de sus principales ventajas es poder convertir los datos en información útil para la toma de decisiones, lo cual favorecerá a la dirección de la empresa en cuanto a su gestión empresarial (Marcano y Talavera, 2007). En este trabajo se crea una base de datos para una empresa: una distribuidora de productos ecológicos. La dirección o el departamento de gestión se encargarán de decidir que medidas tomar para mejorar su gestión empresarial. Para ello utilizarán los datos almacenados en la base de datos de la empresa.

Toda la información que se quiera o necesite almacenar de la distribuidora se representa en una base de datos. Algunos ejemplos de ello serían:

- ❖ Nombre y demás características de los productores ecológicos.
- ❖ Nombre y demás características de los clientes.
- ❖ Compras y ventas que se realicen en sus respectivas fechas.
- ❖ Detalle de los pedidos que recibe la empresa.
- ❖ El transporte de los alimentos ecológicos.

La razón que me ha llevado a elegir esta empresa es debida al interés por conocer la evolución de la producción de *productos ecológicos*¹ en Galicia. Aunque en este trabajo no se estudia su evolución, con ayuda de los datos extraídos del Ministerio de agricultura, alimentación y medio ambiente (2001-2012), más adelante en el trabajo, podremos comprobar que estamos ante una actividad creciente en el tiempo.

Son dos los objetivos que se pretenden alcanzar en el presente trabajo. Por un lado demostrar la importancia que tiene para una empresa tener toda la información almacenada en una base de datos para la gestión día a día de sus procesos. La información que se almacena debe ser precisa, completa y fiable. Por otro lado, que la información almacenada en la base datos facilite la toma de decisiones por parte de la dirección o departamento de gestión de la empresa.

Por último, este trabajo está estructurado de la siguiente manera: Primero mostraremos la planificación tanto estimada como real que conlleva la realización del trabajo. Después se mostrará una breve descripción de la evolución de las bases de datos en el tiempo. Luego explicaremos las diferentes fases a seguir para la creación de la base de datos. Finalmente se mostrará la utilidad del almacenamiento de datos para lograr el objetivo final que es la toma de decisiones usando un caso real desde el inicio (diseño de base de datos) hasta el análisis de datos que posibilita esa toma de decisiones.

¹ Se refiere a los alimentos extraídos de la agricultura o ganadería ecológica. Son productos que se obtienen sin el uso de sustancias químicas como pueden ser: pesticidas, fertilizantes, medicamentos...Se obtienen respetando el ritmo de crecimiento de plantas y animales. Además es un producto elaborado sin adición de sustancias artificiales: adictivos, colorantes, aromas.... (Ministerio de Medio Ambiente y Medio Rural y Marino, 2009).

1. Planificación

1.-Tareas y esfuerzo

Las tareas que se necesitan realizar para conseguir los objetivos del trabajo vienen dictadas por la metodología del diseño de Base de datos (apartado 2.3). De todas formas concretamente estas tareas que habrá que realizar son:

■ **Capacitación**

En esta tarea se trata de adquirir los conocimientos necesarios para elaborar el trabajo de fin de grado. Para ello se realizarán:

- **Curso Trabajo fin de grado.**
- **Revisión de los aspectos relacionados con el diseño de la Base de datos.**

■ **Diseño conceptual de la Base de datos**

Esta tarea consiste en modelar formalmente los requisitos de almacenamiento de la información y extraer la selección del dominio.

■ **Diseño Lógico**

Consiste en obtener las tablas que deberá contener la base de datos a partir de una serie de reglas de transformación descritas el apartado 2.5.3 del presente trabajo.

■ **Implementación de la base de datos**

Consiste en crear la base de datos a partir de las tablas en el SGBD elegido. Se ha elegido Access, como SGBD, por su fácil manejo y a la vez por su potencia para la gestión de base de datos de envergadura media.

■ **Alimentación de la base de datos.**

Este paso consiste en incorporar en cada una de las tablas de la base de datos toda la información que permita probar tanto la estructura de la base de datos como las consultas SQL y Data mining para la explotación de datos.

Los datos serán introducidos manualmente, introduciendo datos a la interface de Access, y también automáticamente creando baterías de datos.

■ Explotación de datos.

Esta tarea engloba realmente dos subtareas: Diseño de consultas SQL y Diseño Data Mining. Estas dos tareas consisten en:

- **Consultas SQL:** Se trata de diseñar las sentencias para recuperar la información relevante para el día a día de la empresa.
- **Data Mining:** Se trata de diseñar consultas relevantes a nivel gerencial.

■ Elaboración de la memoria.

Consiste en documentar las tareas realizadas en la memoria del trabajo, es decir, los objetivos, las tareas y el resultado que se obtendrá.

En la tabla 1 presentamos la estimación del tiempo en horas que se consideró para la realización de estas tareas, las cuales, como se puede comprobar, supondrán un total de 287 horas de esfuerzo.

Tabla 1.- Tabla de esfuerzo estimada

Id Tarea	Tarea	Horas
1	Capacitación	25
1.1	<i>Curso TFG</i>	<i>10</i>
1.2	<i>Revisión de los aspectos relacionales con el diseño de la Base de datos</i>	<i>15</i>
2	Diseño Conceptual de la base de datos	30
3	Diseño Lógico	30
4	Implementación	15
5	Alimentación de la Base de datos	27
6	Explotación de los datos	70
6.1	<i>Consultas SQL</i>	<i>35</i>
6.2	<i>Data Mining</i>	<i>35</i>
7	Elaboración de la Memoria	90
TOTAL HORAS		287

2.- Recursos

Para realizar el presente trabajo, son necesarios los siguientes recursos:

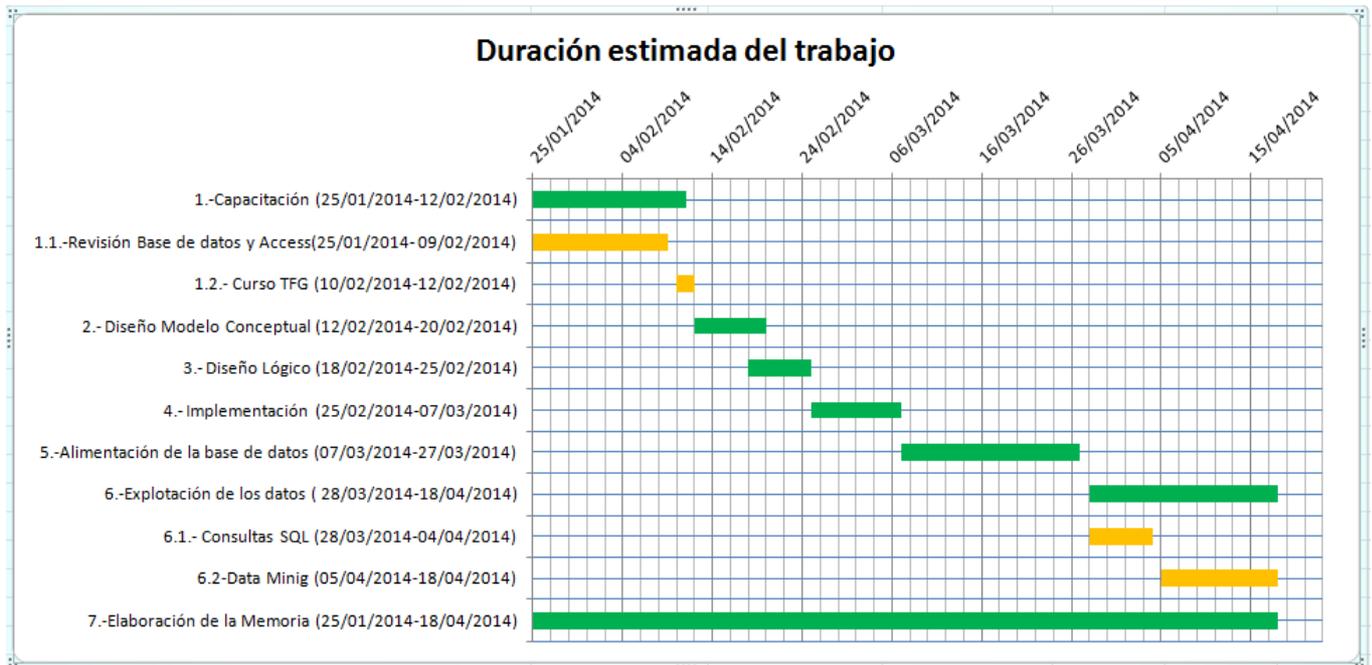
- **Recursos Humanos.** En este caso tenemos como único recurso la alumna que realiza el trabajo. El coste que supone tener a una persona graduada que se encargue de este trabajo es de 30 euros/hora. Por lo tanto, el coste estimado del proyecto en RR.HH. es de 9.090
- **Recursos materiales.** Para realizar este trabajo se necesitarán herramientas como un ordenador que tendrá instalado un sistema operativo, un navegador web (Google Chrome), un SGBD (Access), y principalmente Excel y Word. El coste de los recursos materiales suponen un total de 1.300 euros. Debido a que los materiales son amortizables en cuatro años si los prorrateamos según la duración del trabajo (en este caso cuatro meses), el coste alcanzará los 108,33 euros.

3.-Temporalización

Al presente trabajo le serán dedicadas 4 horas diarias (1 hora para la elaboración de la memoria y las 3 horas restantes se repartirán entre las tareas que haya que realizar cada uno de los días comprendidos entre la fecha inicial del trabajo (25 de enero de 2014) y la fecha final (18 de abril de 2014)).

A través de la Gráfica de Gantt podemos observar la duración estimada en días que conlleva realizar cada tarea. Entonces como conocemos la fecha de inicio del trabajo y tenemos estimado un número determinado de días para realizar cada tarea podemos calcular que la fecha final será el 18 de abril.

Gráfico 1.- Duración estimada del trabajo



En el gráfico, cada tarea aparece sombreada en color verde. Dos de ellas (Capacitación y Explotación de datos) aparecen desglosadas en 2 tareas de color amarillo cada una.

4.-Seguimiento

Se ha realizado un seguimiento contabilizando el número de horas reales que consumió cada fase. Como podemos observar en la tabla 2, tres son las tareas en las que se detectan desviaciones en las horas dedicadas a su elaboración. En dos de ellas (Diseño Lógico y Explotación de datos) no se necesitaron tantas horas como se había planificado en la tabla 1, pero en cambio en la tarea (Elaboración de la memoria) fue necesario incrementar el esfuerzo.

Estas desviaciones fueron producidas por no ajustar de forma exacta las horas que conlleva realizar cada tarea. Para intentar corregir las desviaciones y mantener la fecha final (18 de abril de 2014) fue necesario dedicar, en una de las tareas (Elaboración de memoria) más horas al día.

Tabla 2.- Tabla Esfuerzo Real

Id Tarea	Tarea	Horas
1	Capacitación	25
1.1	<i>Curso TFG</i>	10
1.2	<i>Revisión sobre base de datos y Access</i>	15
2	Diseño Conceptual de la base de datos	30
3	Diseño Lógico	28
4	Implementación	15
5	Alimentación de la Base de datos	27
6	Explotación de los datos	65
6.1	<i>Consultas SQL</i>	30
25 6.2	<i>Data Mining</i>	35
7	Elaboración de la Memoria	100
TOTAL HORAS		290

Finalmente este trabajo supone un total de 290 horas de esfuerzo. La duración inicial y final se mantuvo, pero no sucede lo mismo con el coste del proyecto el cual ascendió hasta 10.500€.

2. Fundamentos Teóricos

2.1. La evolución de las bases de datos

Las bases de datos se han convertido en una de las herramientas más utilizadas en la sociedad actual de la información. Se utilizaban para recuperar información o también como almacén de datos en todos los campos (científico, social, económico y político). En este apartado se presenta una breve panorámica histórica sobre las bases de datos.

Charles Bachman, en los años sesenta, desarrolla un sistema de base de datos que será conocido como “Almacén de datos Integrados (IDS)”. A raíz de este sistema tiene su origen el primer Sistema de Gestión de Base de Datos (SGBD). Este era un sistema bastante potente capaz de representar relaciones de datos de forma más compleja lo que llevó a que adquiriera la fuerza necesaria para convertirse en estándar. El SGBD contenía lo fundamental del modelo de datos en red, posteriormente sería estandarizado por el grupo “Conference on Data Systems Languages” (CODASYL), (Marqués, 2011).

No será hasta 1970 cuando Edgar Codd proponga el modelo de datos relacional basado en una teoría matemática (Codd, 1970). Pese a los esfuerzos ingentes en el desarrollo de prototipos que sigan este modelo, sería en 1979 cuando se creó un SGDB que se denominó relacional. Este carecía de características destacadas del modelo, como la integridad referencial que no será incluida hasta mucho después (Marqués, 2011).

En los años 80 se consolida el modelo relacional y se adopta el SQL (Structured Query Language), lenguaje estructurado de consultas, a finales de la década. El SQL se convierte en un lenguaje estándar para trabajar con base de datos relacionales, es decir, es soportado por la mayoría de los programas (Date y Ruíz Faudón, 2001). Aunque se conoce como lenguaje para consultas permite, además, definir la estructura de los datos, modificar los datos de la base de datos y especificar restricciones de seguridad.

En los años 90 se producen nuevos avances en muchas áreas de la base de datos. Los productos permiten definir nuevos tipos de datos, como imágenes, textos y también la realización de consultas complejas (consiste en utilizar varias tablas en una misma consulta). También se desarrollaron sistemas especializados para crear datawarehouses. El término datawarehouse fue impuesto por William H. Inmon y se traduce como almacén de datos. Una de sus principales características es que permite la toma de decisiones en cualquier área funcional basándose en información integrada y global del negocio (Inmon, 2002).

A pesar de los enormes avances de las bases de datos, era necesario que estuvieran orientadas a los objetos. Surgen por ello la base de datos relacional que introduce elementos de orientación a objetos por lo que seguirán teniendo una posición predominante en el mercado (Connolly y Begg, 2005).

Las bases de datos sustituyen a los ficheros en los Web sites, almacenando datos que se consultan desde un navegador y devolviendo las respuestas e integrándose con lenguajes de marca como HTML (Hyper Text Markup Language) o XML (Extensive Markup Language) para permitir su visualización por los navegadores (Connolly y Begg, 2005).

Las bases de datos siguen avanzando y hoy están presentes al hablar de librerías digitales, proyectos científicos, procesos de decisión etc.

2.2. Definición de Bases de Datos

Una base de datos consiste en una gran cantidad de datos relacionados entre sí y a la vez almacenados en un soporte informático que tienen un significado implícito para el usuario final al estar interesados en conocer su contenido (Elmasri y Navathe, 2007).

Un sistema de gestión de bases de datos (SGBD) es *“un sistema software que permite a los usuarios definir, crear, mantener y controlar el acceso a la base de datos.* (Connolly y Begg, 2005, p.15). El SGBD permite definir la base de datos utilizando un lenguaje de definición de datos mediante el cual los usuarios pueden especificar las estructuras, tipos de datos y todas aquellas restricciones que le sean aplicadas a los datos almacenados en la base.

El SGBD utiliza un lenguaje de manipulación de datos para crear y mantener la base de datos. Una vez que toda la información esté almacenada y descrita, este lenguaje proporciona un mecanismo general de consulta de datos (lenguaje de consulta). Este mecanismo solucionó el problema que existía anteriormente cuando se utilizaban bases de datos pero con archivos permanentes.

Por otro lado, como se ha mencionado en la definición del SGBD, se encarga de tener controlado el acceso a la base de datos. Esto es, proporciona garantía ante intentos de accesos no autorizados, garantiza la fiabilidad de la información almacenada y mantiene una coherencia con la información mediante un sistema de integridad con el que ampara un control de recuperación en caso de caídas del sistema (restaura la base de datos en un estado previo coherente) (Connolly y Begg, 2005).

La combinación del SGBD, la propia base de datos y los usuarios constituyen lo que se conoce como sistema de bases de datos.

Hoy en día prácticamente todas las empresas utilizan las bases de datos como herramienta esencial en su trabajo. Pues bien un sistema de bases de datos se puede construir para diferentes tipos de empresas con la finalidad de que aporte la información necesaria en el momento que se necesite. Por ejemplo en las universidades se utilizan con el objetivo de informar a los estudiantes sobre cursos, libros de texto que necesiten consultar..., respecto a líneas aéreas, la base de datos trata de que el cliente puede realizar las reservas que mejor le convenga teniendo a su disposición la información que necesiten, la banca etc., (Silberschatz *et al.*, 2006).

El propósito de un sistema de bases de datos surge a raíz de los sistemas convencionales que consistían en sistemas que almacenaban información de forma permanente en varios archivos como se mencionó anteriormente.

El objetivo de toda base de datos es que sea precisa, es decir, que contenga datos actualizados en el tiempo y fiable, que no contenga información errónea (Elmasri y Navathe, 2007).

2.3. Diseño de Bases de Datos

La tarea de crear una base de datos no es sencilla, esta consta de tres fases: conceptual, lógica y física. Es importante seguir el orden de las fases debido a su labor compleja pero también es importante que el diseñador de la base de datos comprenda las necesidades de los usuarios (Silberschatz *et al.*, 2006).

Por lo tanto es necesario entender y recopilar toda la información que se quiere almacenar en la base de datos y esta información la proporcionará el usuario del dominio². Para llevar a cabo esta primera tarea los diseñadores de la base de datos deberán interactuar con los expertos y los usuarios del dominio.

Una vez que el diseñador tenga claro la información con la que quiere trabajar, ya entraría en la primera fase. Esta fase se conoce como **diseño conceptual** que consiste en la creación de un esquema o modelo conceptual de la base de datos. Este esquema o modelo conceptual es independiente de los sistemas de gestión de base de datos o lenguajes de programación. Por ello los usuarios no técnicos deberán entender el esquema, el cuál no contendrá detalles sobre cómo se implementa la base de datos. Durante esta fase, los diseñadores suelen crear un diagrama que les sirva como ayuda para visualizar la base de datos. El diagrama permite identificar cada entidad, las relaciones entre ellas, cada atributo y las restricciones.

La siguiente fase se conoce como **diseño lógico** que consiste en transformar el esquema conceptual a un modelo de datos para un sistema de gestión de bases de datos determinado. Este modelo de datos consiste en un esquema relacional. Representa como se estructuran los datos más cerca de la implementación.

En la última fase, **diseño físico**, el diseñador tiene que materializar el esquema relacional, es decir, mantener una coherencia entre el anterior diseño de la base de datos, tablas y restricciones de integridad. Además deberá seleccionar métodos de acceso específico para los datos con la finalidad de conseguir buenas prestaciones y por otro lado diseñar las medidas de seguridad que requieran los datos (Connolly y Begg, 2005).

² Aquellas personas que proporcionan información para crear la base de datos.

2.4. Modelo Entidad-Relación (ER)

El modelo de entidad-relación (ER), fue creado por Peter Chen en 1976. Es uno de los diferentes modelos que se utilizan en la fase de diseño conceptual para la representación, relación e interacción de los diferentes datos. El modelo ER representa las entidades, las relaciones entre ellas, las propiedades de cada entidad (atributos) y las restricciones (Chen, 1976).

2.4.1. Entidad, tipo de entidad y atributo.

Una **entidad** es un concepto que puede ser concreto (Pedro, casa, gato...) o bien abstracto (trabajo, curso...) sobre la cual se recoge información. Es importante señalar que no admite duplicados, es decir, no pueden existir dos entidades iguales (Connolly y Begg, 2005). Un **tipo de entidad** es una abstracción que permite definir la colección de entidades que tienen los mismos atributos (propiedades), por ejemplo: persona, cuenta bancaria (Elmasri y Navathe, 2007). Cada tipo de entidad se simboliza con un rectángulo con el texto dentro (De Miguel Castaño y Piattini Velthuis, 1997).

Cada entidad tiene sus propiedades (información que la caracteriza y diferencia de las demás) estos son los **atributos** los cuales tienen diferente valor para cada entidad.

Existen distintos tipos de atributos:

- **Monovalorados /Multivalorados:** Cada atributo puede tener un valor único o más de uno. Por ejemplo si tenemos como atributo teléfono y contiene más de una línea de teléfono diremos que estamos ante un atributo multivalorado. Por lo contrario si sólo tienen un único valor estaremos ante atributos monovalorados.
- **Simple /Compuesto:** la diferencia radica en la división de subpartes. En este caso estaremos ante atributos compuestos. Un ejemplo de ello sería el atributo *dirección* el cual podemos dividirlo en *calle*, *número* y *población*. Mientras que un atributo simple sería por ejemplo DNI.
- **Almacenado/Derivado:** la diferencia está en la obtención de los datos, no es lo mismo insertar un dato que calcularlo. Un ejemplo de *dato derivado* sería la edad, si tenemos como *dato almacenado* la fecha de nacimiento de una persona podremos calcular la *edad*.

Los atributos tienen diferente valor para cada entidad, por lo tanto, el dominio contiene todos esos valores que puede tomar cada atributo. El **dominio** se define como “*un conjunto de valores homogéneos y atómicos*. Por homogéneo se entiende que todos los valores son del mismo tipo y por atómico, que cada valor de un dominio es indivisible” (De Miguel y Piattini, 1997, pp.131-133). El dominio es estático, es decir, los valores no varían en el transcurso del tiempo (en caso de variar serían dominios distintos), pero en cambio la relación si es dinámica.

Es importante que todos los dominios tengan un nombre por el que nos resulte más fácil referirnos a él y un tipo de datos.

Los dominios pueden definirse por intensión o por extensión. Por ejemplo las edades de personas son dominios que se definen por intensión como entero de longitud comprendido entre 18 y 67 (si hablamos de población activa); mientras que un dominio por extensión sería el nombre de las distintas nacionalidades.

Además de dominios y atributos simples también existen los compuestos. Se entiende por dominio compuesto el conjunto de dominios simples al que se pueden aplicar algunas restricciones de integridad. Por ejemplo tres dominios simples serían Día, Mes y Año, por lo que un dominio compuesto sería Fecha. Las restricciones de integridad, para este caso, tienen por finalidad el no permitir que aparezcan valores no válidos para la fecha (De Miguel y Piattini, 1997).

2.4.2. Identificadores

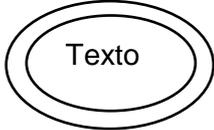
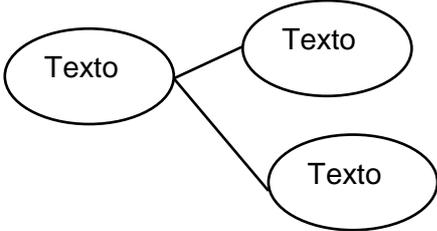
Una restricción importante para las entidades es la clave de los atributos. Normalmente un tipo de entidad tiene un atributo cuyos valores son diferentes para cada entidad individual de todo el conjunto de entidades. Este atributo se le conoce como **atributo clave** y sus valores sirven para identificar a cada entidad. La función del atributo clave o también conocida como restricción de unicidad es prohibir que dos entidades contengan el mismo valor para el atributo clave al mismo tiempo (Elmasri y Navathe, 2007). Para reconocer a un atributo como clave se subrayará el nombre de este dentro del óvalo.

En el apartado 2.4.1 mencionabamos el término tipos de entidad, por lo tanto, una vez explicado el término de identificación, podemos decir que existen dos tipos de entidad: fuerte y débil. Una entidad fuerte es aquella que se identifica por sí misma, es decir, contiene un atributo clave que sirve para identificar a cada entidad sin lugar a dudas; mientras que una entidad débil depende de la identificación (atributo clave) de una entidad fuerte. El tipo de entidad débil se explica más detalladamente en el apartado 2.4.5 debido a que es necesario mencionar antes el concepto de relación y tipos de relación.

Tabla 3. Representación de Entidad Fuerte.

Tipo de Entidad	Símbolo
Fuerte	

Tabla 4. Representación de atributos con diferente notación.

Atributo	Símbolo
Multivalorado	
Compuesto	
Derivado	

2.4.3. Relación y tipo de relación

Una relación consiste en asociar varias entidades. Un conjunto de relaciones es una relación matemática entre n entidades y cada tupla (fila) de la entidad es una relación. Se

representa mediante un rombo que estará unido a la/s entidad/es (Chen, 1976). En el rombo se introduce un texto que tendrá que ser una forma verbal que deberá tener concordancia a la hora de leer la relación entre entidades.

Un tipo de relación se conoce como la unión entre una o más tipos de entidad participantes (Connolly y Begg, 2005).

El número de entidades que se relacionan se conoce como **grado**. Existen varios tipos: Cuando mencionamos grado 1 nos estamos refiriendo a las relaciones reflexivas, grado 2 a las relaciones binarias, de grado 3 a las ternarias (De Miguel Castaño y Piattini Velthuis, 1997).

A continuación se muestran los diferentes tipos de relación según el grado:

- **Relaciones Reflexivas: consiste en relacionar una entidad consigo misma.**

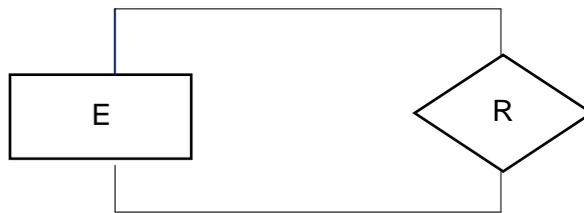


Figura 1. Relación Reflexiva

- **Relaciones binarias: consiste en unir dos entidades mediante una relación.**

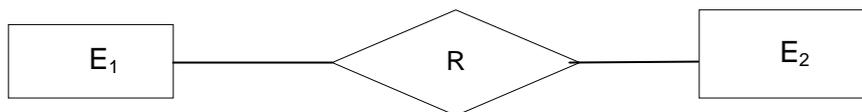


Figura 2. Relación binaria.

- **Relaciones ternarias: Son aquellas que unen tres entidades mediante una relación.**

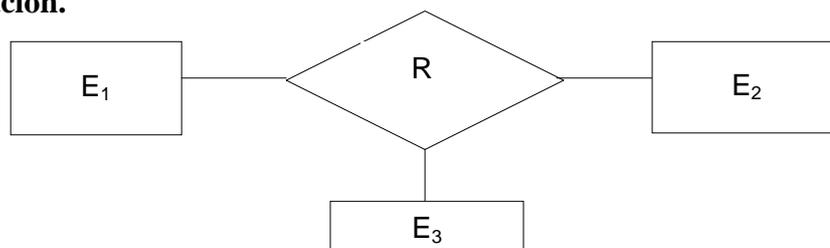


Figura 3. Relación Ternaria

2.4.4. Participación y Cardinalidad

Existen una serie de restricciones estructurales como son la cardinalidad y la participación las cuales hay que tener en cuenta a la hora de explicar el modelo entidad relación.

Por un lado está la **participación** que es el número mínimo de entidades que están asociadas por un conjunto de relaciones. Puede ser total o parcial, esto es, “*si cada entidad de E participa al menos en una relación de R se conoce como participación total; por el contrario “si sólo algunas entidades de E participan en relaciones de R, se dice que la participación del conjunto de entidades E en la relación R es parcial”* (Silberschatz et al. 2006, p.167).

Mientras que la **cardinalidad** es el número máximo de entidades asociadas mediante un conjunto de relaciones.

Por tanto la notación que se va usar para representar estas dos restricciones es (x, x) mínimos y máximos respectivamente (Elmasri y Navathe, 2007). Pueden ser de varios tipos:

- **Uno a uno (1:1):** Cada entidad A se puede asociar como máximo con una entidad de B y cada entidad B sólo se puede asociar como máximo con una entidad de A.
- **Uno a varios (1: N):** Cada entidad A se puede asociar con cuántas quiera entidades de B; en cambio la entidad B sólo se puede asociar como máximo a una entidad de A.
- **Varios a uno (N: 1):** Cada entidad A se puede asociar como máximo a una entidad de B, sin embargo la entidad de B se puede asociar con cuántas quiera entidades de A.
- **Varios a Varios (N: M):** Cada entidad de A se puede asociar con cuántas entidades de B quiera y cada entidad de B se puede asociar con cuántas entidades de A quiera también.

2.4.5. Entidad Débil

Hemos mencionado anteriormente que existen dos tipos de entidades: las fuertes que ya fueron explicadas anteriormente y las débiles.

Las entidades débiles tienen la particularidad de no tener suficientes atributos con los que pueda identificarse. Por ello necesita estar relacionada con una entidad fuerte. La clave de una entidad débil está compuesta por el atributo clave de la entidad fuerte más el atributo, que funcione como clave, de la entidad débil. Como podemos observar en la figura 4, este tipo de entidad se representa mediante un rectángulo de línea doble.

Cuando relacionamos un tipo de entidad débil y una fuerte, la relación tiene que representarse mediante un rombo de doble línea también.

Cuando queramos expresar la participación y la cardinalidad debemos de tener en cuenta que a la entidad débil le corresponde la notación: Uno a Uno, (1,1), es decir que cada entidad se puede relacionar como mínimo y como máximo con otra entidad (De Miguel Castaño y Piattini Velthuis, 1997).

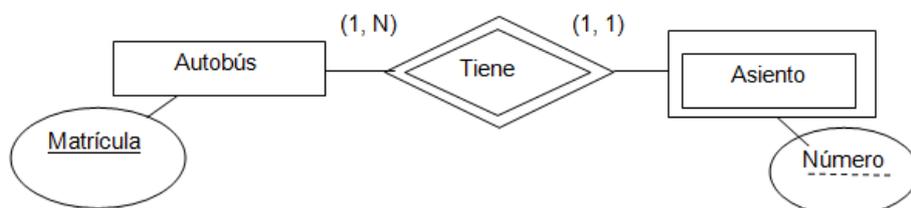


Figura 4.- Ejemplo representación entidad débil.

Como podemos observar en la figura 4, existe una relación entre dos tipos de entidades: *autobús* y *asiento*. La entidad *autobús* representa a una entidad fuerte y se identifica con el atributo matrícula, la cual es un buen identificador (clave) ya que no va a existir dos autobuses con el mismo número de matrícula y todos tienen una (no existen nullos).

Por otro lado tenemos la entidad débil, *asiento*, que necesita estar relacionada con la entidad *autobús* (entidad fuerte) ya que el atributo “número” por sí solo no especifica con claridad de qué asiento o asientos nos queremos referir, pero si lo asociamos con la clave de la entidad *autobús* podremos conocer el número del asiento al cual nos estamos refiriendo.

2.5. Modelo Relacional

El modelo relacional ha sido propuesto por E.F Codd en 1970 en su artículo titulado “*A relational model of data for large shared data banks*”. Además “*es un modelo aceptado generalmente como uno de los hitos en los sistemas de bases de datos*” (Connolly y Begg, 2005, p.64).

Este modelo relacional está basado en la teoría de las relaciones, en donde los datos se estructuran en forma de relaciones (tablas) siendo el objetivo principal del modelo mantener la independencia de la estructura lógica respecto al modo de almacenar y a otras características de tipo físico.

El trabajo presentado por Codd perseguía una serie de objetivos:

- **Independencia física:** la forma en que se almacenan los datos no debe influir en la manipulación lógica y por lo tanto los usuarios que tienen acceso a estos datos no han de modificar sus programas por cambios en el almacenamiento físico.
- **Independencia lógica:** realizar cambios como por ejemplo añadir, eliminar o modificar elementos de la base de datos no debe repercutir ni en los programas ni en los usuarios que acceden a los datos.
- **Flexibilidad:** Se trata de poder ofrecer a cada usuario los datos que necesiten de la forma más apropiada a la correspondiente aplicación.
- **Uniformidad:** el modelo relacional presenta un aspecto uniforme (tablas), y esto facilita al usuario a que pueda comprender y modificar la base de datos.
- **Sencillez:** de acuerdo a las características anteriores junto con los lenguajes de usuario que son sencillos, originan que este modelo relacional sea fácil de comprender y utilizar por parte del usuario.

Para lograr estos objetivos, Codd introduce el concepto de relación la cual será la estructura básica del modelo, es decir, que el modelo relacional representa la base de datos relacional como un conjunto de relaciones (Connolly y Begg, 2005).

2.5.1. Relación

Como elemento central del modelo relacional, **una relación** se define como un subconjunto del producto cartesiano de varios dominios donde cada elemento de la relación es una serie de n valores ordenados (Connolly y Begg, 2005).

Se representa mediante una tabla bidimensional compuesta por filas también conocido como tuplas (registros individuales) y columnas (atributos). Las relaciones (tablas) se utilizan para guardar información acerca de objetos, seres, lugares...todo aquello el cual se quiera representar en la base de datos.

Tabla 5.- Representación de Dominio y Relación Empleado

Dominios			
Dni	Nombre	Nacionalidad	Edad
Xxxxxx Extensión de 9 caracteres	Xxxxxx Extensión de n caracteres (por ejemplo 20)	Española Italiana Inglesa	Entero de longitud dos comprendido entre 18 y 67 (población activa)

Empleado				Atributo
Dni	Nombre	Nacionalidad	Edad	
45632569N	María	Española	18	T U P L A S Cardinalidad 3
79563258V	Pablo	Inglesa	52	
36952874L	Jesús	Italiana	45	

Grado 4

Fuente: Elaboración Propia a partir de (De Miguel y Piattini, 1997).

A modo de ejemplo en la tabla 5 se representa la relación Empleado, en donde aparece la estructura del modelo relacional. En ella podemos observar el **nombre** de la relación (Empleado); **los atributos** (Dni, Nombre, Nacionalidad y Edad); **los dominios** (de donde los atributos toman sus valores); **las tuplas** (contiene los valores que toma el Dni, el nombre, la nacionalidad y la edad para un empleado); **el grado** (indica el número de atributos) y **la cardinalidad** (indica el número de tuplas).

Las **relaciones** contienen una serie de propiedades:

- Cada relación contiene un nombre diferente.
- Cada celda de la relación contiene un único valor atómico.
- Cada atributo tiene un nombre diferenciado.
- Los valores de un atributo pertenecen al mismo dominio.
- Cada tupla es distinta (no hay tuplas duplicadas).
- El orden de los atributos y de las tuplas no tiene importancia. En el caso de las tuplas, en la práctica, el orden puede afectar a la eficiencia de acceso a las tuplas. (Connolly y Begg, 2005).

En el modelo relacional se utiliza el **esquema relacional** para representar cada entidad junto con sus respectivas propiedades: Entidad (atributo, atributo,..., atributo).

2.5.2. Restricciones del Modelo Relacional

El modelo relacional está sujeto a varias restricciones. Algunos autores (Cuadra *et al.*, 2013), (Elmasri y Navathe, 2007) explican las diferentes reglas que se utilizan para verificar la validación de los datos almacenados:

- **Restricciones de dominio:** Dentro de cada tupla, el valor de un atributo (A) debe ser un valor atómico (valor indivisible) del dominio (A).
- **Restricciones de clave:**
 - La restricción de *valores nulos*: todo atributo clave en una relación debe tomar un valor, nunca valores nulos.
 - *Superclave* se define como un atributo o conjunto de atributos que, de forma conjunta, identifican de forma unívoca a cada tupla. Sin embargo una superclave puede contener varios atributos adicionales que no sean necesarios para la identificación, por ello es preciso identificar las superclaves que sólo contengan el número mínimo de atributos necesarios para la identificación.

- *Clave candidata.* Un esquema de relación puede contener más de una clave. Cada una de ellas recibe el nombre de clave candidata. Una de las claves candidatas se designará como clave principal para identificar las tuplas de la relación.
- *Clave primaria:* es aquella clave candidata elegida como elemento principal de identificación de las entidades pertenecientes a un conjunto de entidades.
- Clave ajena (integridad referencial). Para explicar esta restricción nos basaremos en un ejemplo:

Como podemos observar en la figura 5, la entidad *Departamento* tiene dos atributos: Código (clave primaria) y Nombre. En cambio la entidad *Empleado* contiene los siguientes atributos Dni (clave primaria), Nombre y Código. Este último atributo (Código) es clave ajena ya que proviene de la entidad *Departamento*. En el esquema relacional la notación que se suele utilizar es una flecha, que desde la clave ajena se desplaza hasta la clave primaria.



Figura 5.- Representación Clave Ajena.

Las restricciones proporcionan las bases lógicas para asegurarse de que los valores de las bases de datos sean válidos puesto que uno de los principales propósitos del procesamiento de base de datos es proporcionar información fiable y precisa en todo momento para poder tomar decisiones.

2.5.3. Transformación del modelo ER al modelo relacional

El diseño de base de datos toma el modelo conceptual (ER) y lo transforma en el modelo lógico (modelo relacional).

Para transformar un modelo ER a un modelo relacional se utilizan los esquemas relacionales. Para llevar a cabo el proceso de transformación, como indican algunos autores (Cuadra *et al.*, 2013) se tienen en cuenta una serie de reglas³:

1. Toda entidad se transforma en una relación.
2. Las relaciones N: M se transforman en una relación.
3. Las relaciones 1: N dan lugar a una propagación de claves.
4. Transformación de entidades débiles.
5. Transformación de relaciones de grado n.
6. Transformación de relaciones 1:1.

1.-Toda entidad se transforma en una relación.

Cada entidad del modelo entidad relación se transforma en una relación cuya clave primaria será el identificador de la entidad.

Toda entidad, como hemos visto, tiene asociado una serie de atributos, pues cada uno de ellos se convertirá en atributos en cada relación creada a partir de las entidades, aunque hay que tener en cuenta una serie de restricciones:

- **Atributo monovalorado:** dan lugar a un atributo en la relación.
- **Atributos multivalorados:** dan lugar a una nueva relación cuya clave principal es la unión del propio atributo multievaluado, la clave primaria de la entidad a la cual pertenece ese atributo multievaluado y las claves ajenas.
- **Atributos compuestos:** son aquellos que contienen más de un atributo .En el modelo relacional sólo se cogerían los atributos simples que están dentro del compuesto.
- **Atributos derivados:** son aquellos cuyos valores se obtienen del resultado de algún cálculo. Como hemos mencionado anteriormente, existen varias formas de tratar con este tipo de atributos en el modelo relacional, en este caso no formarán parte de la relación.

³ Para explicar las reglas de transformación se utilizarán a modo de ejemplo las relaciones que aparecen en la página 38 del presente trabajo.

2.- Transformación de relaciones N: M

Las relaciones N: M dan lugar a una relación cuya clave será la unión de los identificadores principales de las entidades que enlaza la relación. Cada id será una clave ajena de la tabla de origen.

Para poder explicar este tipo de transformación hemos tenido que dividir una relación ternaria, ya que, en nuestro diagrama no existen relación N: M. La relación N: M que vamos a utilizar para llevar a cabo en la transformación al modelo relacional se muestra en la figura 6.

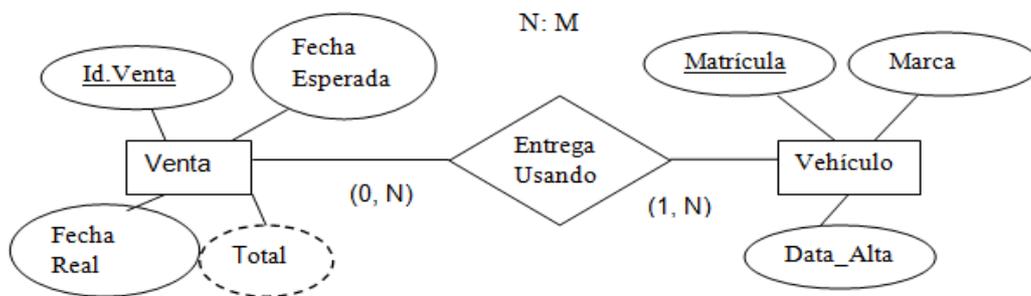


Figura 6.- Relación N: M

Para realizar la transformación al modelo relacional cuando estamos ante relaciones N: M, la solución es crear **un relación nueva**, en este caso “Entrega Usando”, que contenga los atributos clave de las entidades de las cuales provienen.

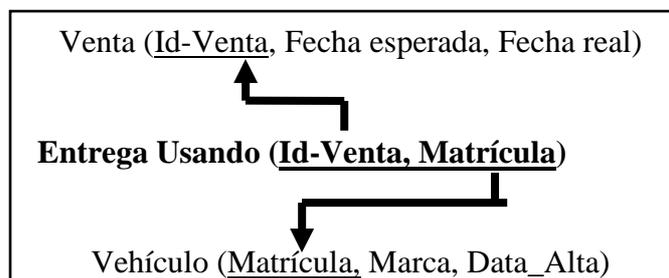


Figura 7.- Transformación al modelo relacional a partir de la relación N: M

3.- Transformación de relaciones 1: N

Para realizar transformaciones de relaciones 1: N, la solución que se adopta es la propagación de las claves.

A continuación podemos observar en la figura 8 un ejemplo de relación 1: N del modelo entidad relación.

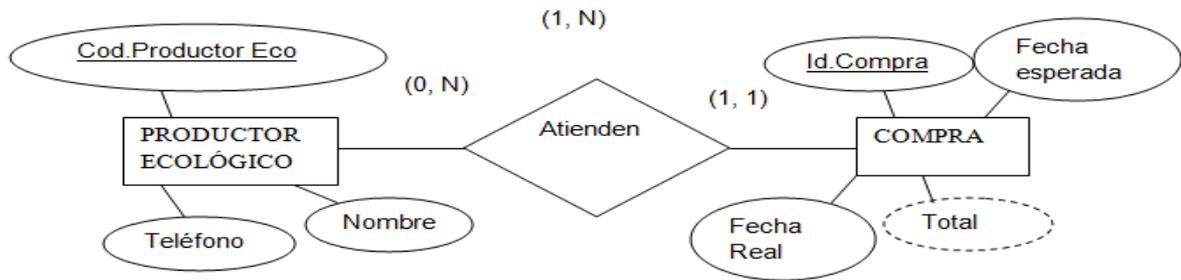


Figura 8.- Relación 1: N

Si llevamos a cabo la transformación a partir del modelo ER (figura 8) sería:

1.-Creamos dos relaciones a partir de las dos entidades existentes.

- PRODUCTOR_ ECOLÓGICO (Cod.ProductorEco, Nombre, Teléfono)
- COMPRA⁴ (Id. Compra, Fecha real, Fecha esperada)

2.-Se propaga la clave ajena (Id.Compra) desde la entidad Productor_Ecológico hasta la clave principal (Id.Compra) de la entidad Compra.

- PRODUCTOR_ ECOLÓGICO (NIF, Nombre, Web, Teléfono, Id. Compra)
- COMPRA (Id. Compra, Fecha real, Fecha esperada)



4.-Transformación de entidades débiles.

Una relación 1: N de dependencia de identificación (entidades débiles) da lugar a la propagación de la clave primaria desde la entidad débil hacia la entidad fuerte. En la figura 9 se muestra una relación entre Línea de Venta y Venta, con sus atributos.

⁴ No se incluye el atributo total, ya que es un atributo derivado (ver página 28).

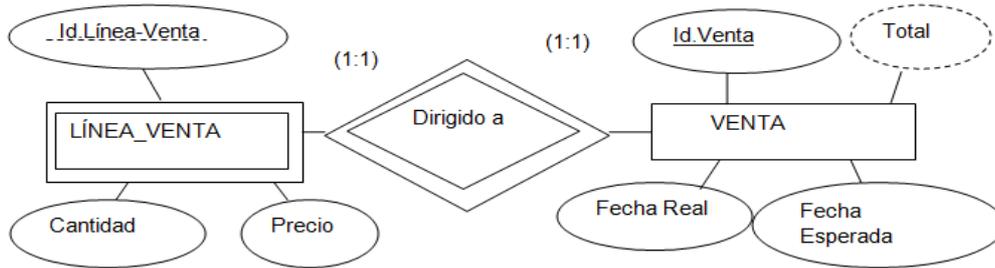


Figura 9. Relación 1: N con entidad débil.

Si llevamos a cabo la transformación a partir del modelo ER (figura 9) sería:

1.-Creamos dos relaciones a partir de las dos entidades existentes:

- LÍNEA_VENTA (Id.Venta, Id.Línea-Venta, Cantidad, Precio)→Entidad débil
- VENTA (Id.Venta, Fecha Real, Fecha Esperada)→Entidad fuerte

2.-Se propaga la clave principal desde la entidad Línea_Venta hasta la entidad Venta.

- LÍNEA_VENTA (Id.Venta, Id.Línea-Venta, Cantidad, Precio)→Entidad débil
- VENTA (Id.Venta, Fecha Real, Fecha Esperada)→Entidad fuerte



5.-Transformación relación grado n (en este caso grado 3).

Hasta ahora sólo hemos trabajado con relaciones binarias. En el esquema entidad relación podemos encontrar una relación ternaria, como podemos observar en la figura 10.

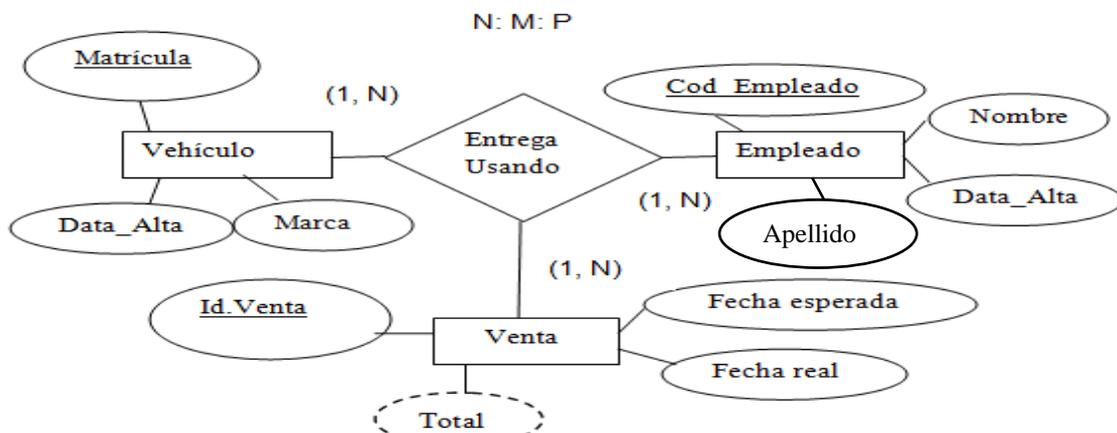


Figura 10.- Relación N: M: P

Generalmente la transformación de relaciones ternarias es similar a la relación binaria. En este caso, surge una nueva relación “Entrega Usando” cuya clave primaria es la unión de las claves principales de las entidades: **VEHÍCULO, EMPLEADO y VENTA.**

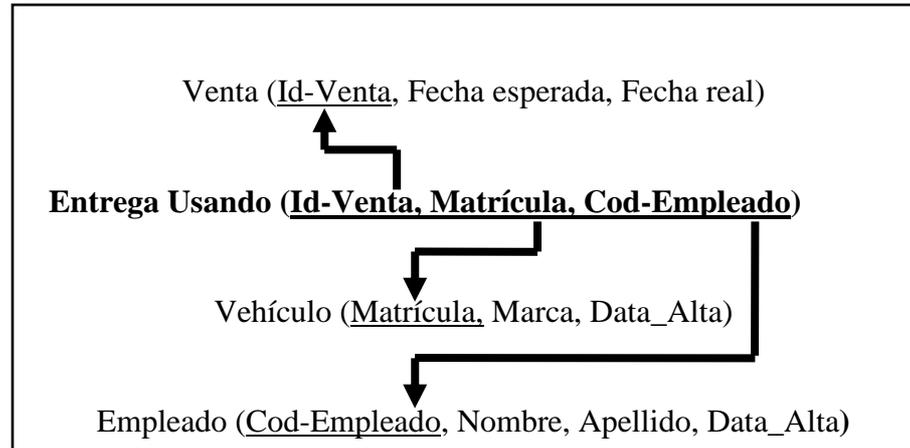


Figura 11.- Transformación de una relación N: M: P al modelo relacional

6.- Transformación de relación 1:1.

En este último caso la transformación al esquema relacional se hace de igual manera que la cardinalidad 1: N, con la particularidad de que el diseñador elige el lado en donde se prefiere colocar la clave ajena.

2.6. Implementación física

Cuando se procede a la creación de la base de datos es necesario escoger el Sistema de Gestión de Base de datos (SGBD) que vamos a utilizar.

Dependiendo de la selección del sistema, la creación de la base de datos puede realizarse de diversas formas. Todos ofrecen normalmente SQL y alguna alternativa propietaria normalmente más visual.

2.6.1. Lenguaje SQL

El lenguaje SQL es el lenguaje estándar ANSI e ISO que permite crear, consultar y modificar datos relacionales. Es un lenguaje que se parece al lenguaje natural, el inglés. Por ello se convierte en un lenguaje que permite acceder a todos los sistemas relacionales comerciales (Silberschatz *et al.*, 2006). La versión más reciente es SQL 2011.

SQL además de realizar consultas a la base de datos permite realizar otras funciones de definición, control y gestión de la base de datos. Las sentencias SQL (expresiones en inglés que permiten obtener datos de las tablas que deseamos) se pueden clasificar según la finalidad dando origen a tres “lenguajes”:

- DDL (Data Description Language: lenguaje de definición) permite definir, modificar o borrar las tablas en las que almacenan los datos y las relaciones entre estas.
- DCL (Data Control Language: lenguaje de control de datos) permite trabajar en un entorno multiusuario en el cual es importante la protección de datos, la seguridad de las tablas y el establecimiento de restricciones en el acceso.
- DML (Data Manipulation Language: lenguaje de manipulación de datos) permite recuperar los datos almacenados en la base de datos, actualizar la base de datos añadiendo nuevos datos o suprimiendo datos antiguos.

En el presente trabajo se va a utilizar Microsoft Office Access (programa de base de datos relacionales) que opera con SQL para trabajar con datos. Además se utilizará la alternativa visual para crear las tablas y las relaciones entre ellas.

3. Diseño de la base de datos de la distribuidora de productos ecológicos

3.1. Distribuidora de Productos ecológicos

La distribuidora de productos ecológicos suministra productos ecológicos a los clientes. La misión de la empresa es transportar estos productos, desde los productores ecológicos, hasta los clientes.

La empresa dispone de una nave amplia y acondicionada en Carballo (A Coruña). El motivo del porqué se implanta en esa zona es porque existen varios productores ecológicos a su alrededor (Ayuntamiento de Malpica, Coristanco, Cabana de bergantiños, Laxe...). Como es una zona periférica será más fácil y viable atender a los clientes.

Según estudios referidos a la percepción y consumo de alimentos ecológicos en Galicia, Craega⁵ (2010)⁶, los puntos de compra preferente son supermercados cerca de domicilios o trabajo, hipermercados, tiendas especializadas, entre otras (véase Anexo1).

Para llevar a cabo esta distribución, la empresa dispone de vehículos homologados para el transporte de estos alimentos.

La distribuidora pretende conseguir los siguientes objetivos: fomentar la agricultura y ganadería ecológica, concienciar a la gente del consumo de estos alimentos y potenciar la demanda.

⁵ Consello Regulador de Agricultura Ecolóxica de Galicia.

⁶ Según la última publicación de datos del Craega de 2013, los puntos de compra preferente no varían.

Para poder cumplir estos objetivos, la empresa tendrá que tener claro las estrategias a seguir. En este caso se inclina por dos de los cuatro instrumentos del marketing ecológico⁷: la distribución y la promoción. Antes de analizar cada uno de ellos, mencionar que lo llamativo de este tipo de marketing es que a la hora de llevar a cabo la comercialización de estos productos se tendrán en cuenta que son tan importantes las necesidades de los consumidores como los intereses presentes y futuros que quiera conseguir la empresa y siempre respetando el medio natural (Calomarde, 2000).

En cuanto a la distribución, esta empresa transporta los productos ecológicos en un radio no muy amplio, pero aún así pretende conseguir venderlos en un gran número de establecimientos de la zona para potenciar su consumo e incrementar la demanda.

La estrategia de distribución es muy importante porque la mayoría de los consumidores sólo compran productos ecológicos si están a su alcance. Para ello deberá asegurarse de:

- Tener controlados a todos los productores.
- Conocer que productos cultivan cada uno de ellos y en que fechas los tendrán preparados.
- Poder contar con un gran número de productores para ser capaz de abastecer a todos clientes en las fechas que ellos soliciten la mercancía.

Si todo eso funciona correctamente, cada vez que la empresa distribuidora reciba un pedido para una determinada fecha, sabrá con que productores contactar e irá en busca de la mercancía. Nos referimos a productores en plural porque puede darse el caso de que la cantidad que se solicite sea lo demasiado grande para que un solo productor pueda proporcionarlo. Por tanto la distribuidora tendrá que ir en busca de más cantidad de ese producto demandado a su segunda opción de productor. La mercancía se recogerá en cajas o bolsas dependerá de la cantidad que se requiera y se transportará al cliente que la solicitó.

⁷ Forma de concebir la relación de intercambio para conseguir que se satisfagan las partes que en ella intervienen, es decir, la sociedad y el medio natural (Calomarde, 2000).

Por otro lado para conocer cuáles son las mejores vías de comunicación de este tipo de alimentos nos centramos en la encuesta realiza por Craega en Galicia, en 2010⁸, (véase Anexo1). Podemos observar que la televisión, seguido de la presencia de estos alimentos en la mayoría de los establecimientos, por medio de revistas y el boca a boca son las principales vías.

Con la promoción lo que se pretende es fomentar la conciencia ecológica, es decir, impulsar a la gente a tener interés sobre los productos ecológicos. La empresa lo que pretende es conseguir que haya una mayor demanda y ser capaz de abastecer a un gran número de negocios.

Una vez que se conocen que medios repercuten más en los consumidores, la empresa podría realizar anuncios en TV, promocionando su nombre, dirección y forma de contactar con la empresa.

La segunda vía de comunicación es la presencia de estos alimentos en los diferentes negocios. La empresa cuando se disponga a repartir estos productos, la caja u otro material que recoge en cada productor podría llevar incorporado un logo de la distribuidora. Y también podría hacer lo mismo con el vehículo.

Otro buen recurso puede ser darse a conocer a través de revistas especializadas e incluso de los periódicos de la zona.

Una vez que se han descrito los objetivos y las estrategias, sería interesante destacar el porque interesa esta empresa. Pues bien es de interés conocer una actividad que se mantenga creciente y que pueda llegar a ser sostenible⁹ en el tiempo (con la difusión del producto se conseguirán más clientes potenciales que retroalimenten al negocio).

⁸ Según la última publicación de datos del Craega de 2013, los medios más influyentes en el conocimiento de productos ecológicos no varían.

⁹ Dado que en los productos ecológicos no se utilizan recursos químicos, el aumento del precio de fertilizantes, pesticidas etc, no le afecta tanto a este tipo de agricultura. Por otro lado como la venta de estos alimentos se realiza en un radio pequeño tampoco le afectaría tanto la subida del combustible.

Nos referimos a que nos adentramos en una actividad creciente porque como podemos observar en la Gráfica 14.- Evolución de superficie y nº de productores en Galicia 2001 - 2012 (véase Anexo2) tanto el número de hectáreas dedicadas al cultivo como el número de productores ha seguido un ritmo creciente a lo largo de esos años.

Según datos extraídos del Ministerio de Agricultura, alimentación y medio ambiente (2013) somos conocedores del descenso, a nivel nacional, de la superficie de cultivos ecológicos. Sufrió un descenso del 11,81% (*gran parte de este descenso se debe a una nueva metodología de cómputo de superficies, en la que ya no se tienen en cuenta determinados usos de la tierra*) En cambio, en Galicia se incrementa hasta un 9,31% la superficie de estos cultivos.

En 2012, Galicia asumía unas 15.091 (ha) dedicadas al cultivo de alimentos ecológicos. Y en cuanto al número de productores desde 2001 hasta 2012 sufrió un crecimiento moderado llegando a sumar en el último año 461 productores.

3.2. Entidad Relación (Diagrama ER)

El diagrama de la distribuidora de productos ecológicos, como podemos observar en la figura 12 de la siguiente página, nos permite visualizar los distintos conceptos explicados en el apartado 2.4 (modelo entidad relación) como son: las entidades, sus respectivos atributos, las relaciones existentes y las restricciones.

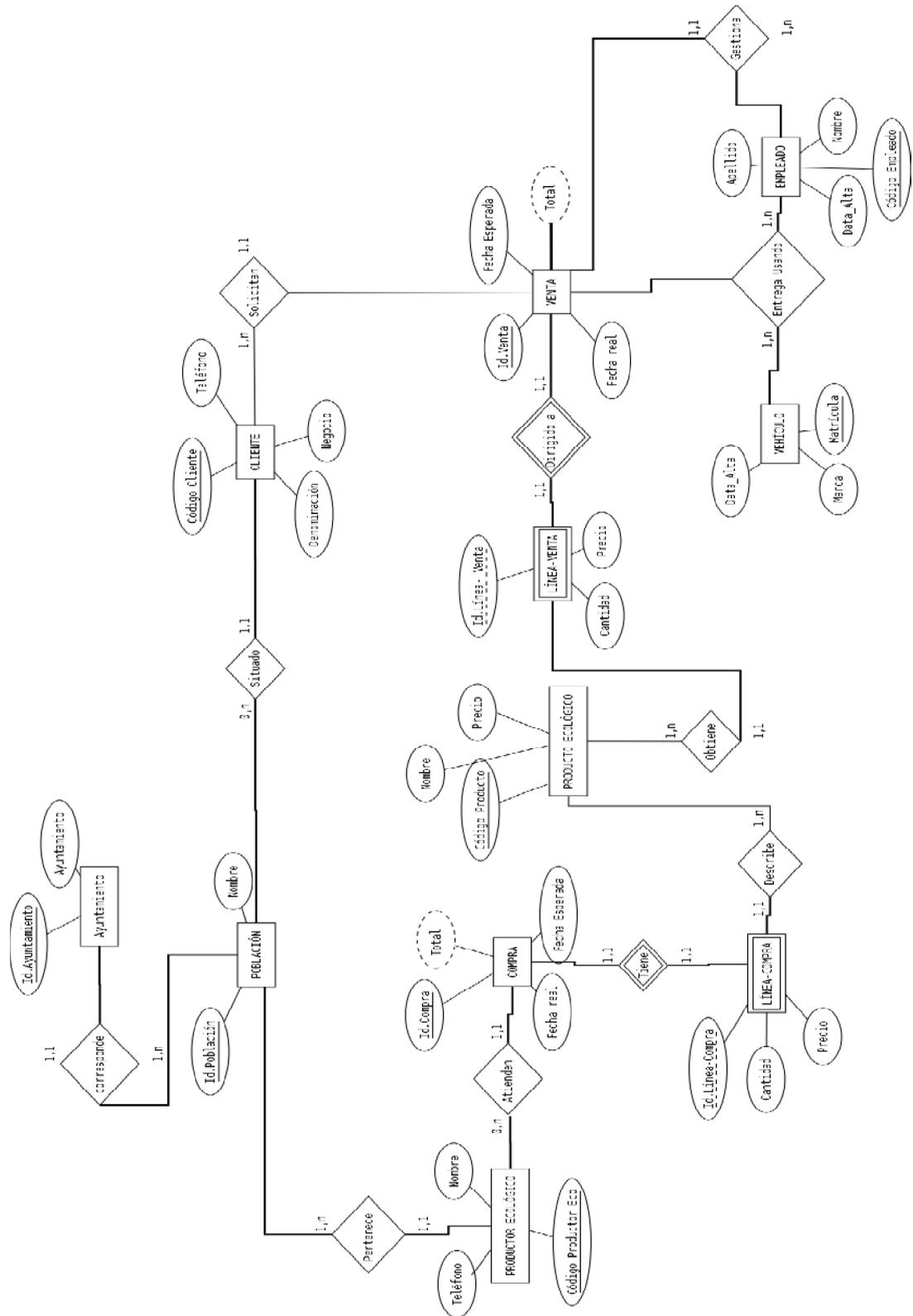


Figura 12. Diagrama ER

3.3. Diccionario de datos

Con el diccionario de datos lo que se pretende es explicar el significado de cada una de las entidades junto con los atributos que forman parte del diagrama de la distribuidora.

Tabla 6. Definición de la entidad Productores ecológicos y sus propiedades.

ENTIDAD: “Productores Ecológicos”: Agricultores o Ganaderos dedicados al cultivo de productos ecológicos.		
CARACTERÍSTICAS : Entidad Fuerte SE IDENTIFICAN POR : Cod_Productor ATRIBUTOS		
NOMBRE	DESCRIPCIÓN	CARACTERÍSTICAS
Cod_Productor	Código alfanumérico que permite identificar a los productores.	Monovalorado Simple Almacenado
Nombre	Nombre de los agricultores y ganaderos que cultivan productos ecológicos.	
Teléfono	Conjunto de caracteres de tamaño 9.	

Tabla 7. Definición de la entidad Población y sus propiedades

ENTIDAD: “Población”: Parroquias en donde están situados los productores ecológicos y los clientes.		
CARACTERÍSTICAS : Entidad Fuerte IDENTIFICADOR: Id_Población ATRIBUTOS		
NOMBRE	DESCRIPCIÓN	CARACTERÍSTICAS
Id_Población	Código alfanumérico que permite identificar a cada parroquia.	Monovalorado Simple Almacenado
Nombre	Nombre de las distintas parroquias.	

Tabla 8. Definición de la entidad Ayuntamiento y sus propiedades.

ENTIDAD: “Ayuntamiento”: se refiere a los distintos Ayuntamientos en donde opera la distribuidora.		
CARACTERÍSTICAS : Entidad Fuerte SE IDENTIFICAN POR : Id_Ayuntamiento		
ATRIBUTOS		
NOMBRE	DESCRIPCIÓN	CARACTERÍSTICAS
Id_Ayuntamiento	Código alfanumérico que permite identificar algunos de los ayuntamientos de la provincia de A Coruña.	Monovalorado Simple Almacenado
Ayuntamiento	Nombres de los ayuntamientos.	

Tabla 9. Definición de la entidad Compra y sus propiedades.

ENTIDAD: “Compra”: Indica las distintas adquisiciones de productos ecológicos.		
CARACTERÍSTICAS : Entidad Fuerte SE IDENTIFICAN POR: Id_Compra		
ATRIBUTOS		
NOMBRE	DESCRIPCIÓN	CARACTERÍSTICAS
Id_Compra	Código alfanumérico que permite identificar cada una de las compras.	Monovalorado Simple Almacenado
Fecha Esperada	Fecha en la cual se prevee adquirir la mercancía a los productores.	
Fecha Real	Fecha exacta en la cual se realiza la compra a los productores.	
Total	Sumatorio de todas las compras realizadas.	Monovalorado, Simple, Derivado.

Tabla 10. Definición de la entidad Línea_Compra y sus propiedades.

ENTIDAD: “Línea_Compra” detalle de las compras que realiza la distribuidora.		
CARACTERÍSTICAS : Entidad Débil SE IDENTIFICAN POR: ID_Compra y Id_Línea_Compra		
ATRIBUTOS		
NOMBRE	DESCRIPCIÓN	CARACTERÍSTICAS
Id_Compra	Código alfanumérico que permite identificar la compra	Monovalorado Simple Almacenado
ID_ Línea _Compra	Código numérico	
Cantidad	Número de productos que se adquieren.	
Precio	Valor monetario en Euros de cada producto ecológico.	

Tabla 11. Definición de la entidad Productos y sus propiedades.

ENTIDAD: “Productos”: productos naturales que se obtienen sin la utilización de residuos químicos (Ministerio de Medio Ambiente y Medio Rural y Marino, 2009).		
CARACTERÍSTICAS : Entidad Fuerte SE IDENTIFICAN POR: Cod_Producto		
ATRIBUTOS		
NOMBRE	DESCRIPCIÓN	CARACTERÍSTICAS
Cod_Producto:	Código alfanumérico que identifica a los productos ecológicos.	Monovalorado Simple Almacenado
Nombre	Nombre de los productos ecológicos que se comercializan.	
Precio	Valor monetario en Euros de cada producto ecológico.	

Tabla 12. Definición de la entidad Línea_Venta y sus propiedades.

ENTIDAD: “Línea_Venta” detalle de las ventas que realiza la distribuidora.		
CARACTERÍSTICAS : Entidad Débil SE IDENTIFICAN POR: Id_Venta y Id.Línea_Venta		
ATRIBUTOS		
NOMBRE	DESCRIPCIÓN	CARACTERÍSTICAS
Id_Venta	Código alfanumérico que permite identificar la venta.	Monvalorado Simple Almacenado
Id_Línea_Venta	Código numérico	
Cantidad	Número de productos ecológicos que se venden.	
Precio	Valor monetario por el cual se venden los productos	

Tabla 13. Definición de la entidad Venta y sus propiedades.

ENTIDAD: “Venta” Indica las distintas ventas realizadas por la distribuidora a sus clientes.		
CARACTERÍSTICAS : Entidad Fuerte SE IDENTIFICAN POR: Id_Venta		
ATRIBUTOS		
NOMBRE	DESCRIPCIÓN	CARACTERÍSTICAS
Id_Venta	Código alfanumérico que identifica el número de cada venta.	Monovalorado Simple Almacenado
Fecha prevista	Fecha prevista para la venta	
Fecha Real	Fecha real de la venta	
Total	Sumatorio de las ventas	Monovalorado, Simple y Derivado

Tabla 14. Definición de la entidad Empleado y sus propiedades.

ENTIDAD: “Empleado” (entidad fuerte): Persona que se encarga del transporte de los alimentos desde los productores hasta los clientes.		
CARACTERÍSTICAS : Entidad Fuerte SE IDENTIFICAN POR: Cod_Empleado ATRIBUTOS		
NOMBRE	DESCRIPCIÓN	CARACTERÍSTICAS
Cod_Empleado	Código numérico que identifica a los empleados de la empresa.	Monoevaluado Simple Almacenado
Nombre	Nombre del empleado	
Apellidos	Apellidos del empleado	
Data_Alta	Día de alta de los empleados	

Tabla 15. Definición de la entidad Vehículo y sus propiedades.

ENTIDAD: “Vehículo” medio de transporte con el cual se traslada la mercancía desde el productor hasta el cliente.		
CARACTERÍSTICAS : Entidad Fuerte SE IDENTIFICAN POR: Matrícula ATRIBUTOS		
NOMBRE	DESCRIPCIÓN	CARACTERÍSTICAS
Matrícula	Identifica el vehículo combinando letras y números.	Monoevaluado Simple Almacenado
Marca	Nombre que se utiliza para la identificación del vehículo	
Data_Alta	Día de alta de Vehículo	

Tabla 16. Definición de la entidad Cliente y sus propiedades.

ENTIDAD: “Clientes” empresas que realizan pedidos a la distribuidora para adquirir productos ecológicos.		
CARACTERÍSTICAS : Entidad Fuerte SE IDENTIFICAN POR: Cod_Cliente		
ATRIBUTOS		
NOMBRE	DESCRIPCIÓN	CARACTERÍSTICAS
Cod_Cliente	Código numérico que identifica a cada uno de los clientes.	Monovalorado Simple Almacenado
Denominación	Nombre de la empresa (Cliente)	
Negocio	Tipo de establecimiento	
Teléfono	Conjunto de caracteres de tamaño 9.	

3.4. Paso a relacional

El esquema relacional (véase Anexo3) surge a raíz de la transformación del modelo ER. Como se mencionó anteriormente toda entidad da lugar a una relación, las claves primarias están subrayadas y las ajenas están identificadas mediante flechas. Por último resaltar que en la entidad compra y venta aparece un atributo derivado “TOTAL” el cual no se introduce en la relación de Compra ni de Venta, por ser atributo derivado, como se explicó en el apartado 2.5.3. También decir que como única relación N: M: P tenemos “ENTREGA USANDO” para la cual se crea una nueva relación independiente.

3.5. Implementación con Microsoft Office Access

Microsoft Access es uno de los sistemas de gestión de base de datos (SGBD) más utilizado debido a su fácil manejo y a la vez por su potencia para la gestión de base de datos de envergadura media. . Sus principales funciones son almacenar, mantener, recuperar e interpretar los datos, además de mantenerlos actualizados pudiendo compartirlos con otros

(Andersen, 2008). Por otro lado, proporciona una serie de herramientas necesarias para llevar a cabo la gestión completa de los datos:

- **Tablas:** Soporte básico de la información de una base de datos y una de las primeras tareas en el desarrollo de la base de datos es la construcción de tablas.
- **Consultas :** La información se deriva del contenido de una o varias tablas tras “preguntar”o “consultar” en Access.
- **Formularios:** La función de los formularios es facilitar la visión, introducción y modificación de datos que están almacenados en las tablas.
- **Informes:** El objetivo es mostrar los datos almacenados en tablas de forma elegante. Normalmente, se realizan informes sobre aquella información que necesita ser impresa (Casas Luengo, 2007).

3.5.1. Tablas

Es el elemento básico de la base de datos. Esta organizada en filas (denominadas registros) y columnas (denominadas campos). Para iniciar la construcción de una tabla es necesario iniciar la base de datos con la que vamos a trabajar, dirigirse al comando Tabla y en la pestaña “crear” tenemos la opción de diseño de tabla.

Cuando procedemos a la construcción de la tabla debemos tener claros los datos que temos que almacenar. Como se puede observar en el Anexo 4 de presente trabajo cada relación tiene un nombre y unos respectivos atributos pues esos serán los datos de la tabla. Por nombre llevará el de la relación y como datos internos los atributos.

En el presente trabajo, la base de datos de la distribuidora de productos ecológicos tiene un total de 12 tablas:

- | | |
|-------------------|------------------------|
| 1. Cliente | 7. Población |
| 2. Compra | 8. Producto |
| 3. Empleado | 9. Productor Ecológico |
| 4. Entrega_Usando | 10. Vehículo |
| 5. Línea_Compra | 11. Venta |
| 6. Línea_Venta | 12. Ayuntamiento |

Utilizaremos la tabla “Cliente” para explicar su construcción.

Nombre del campo	Tipo de datos
Cod_Cliente	Número
Nombre	Texto
Negocio	Texto
Teléfono	Texto
Id_Población	Texto

Figura 13. Creación de la Tabla Cliente

En la columna “ Nombre del Campo” se escriben todos los atributos de la relación, en este caso , de la relación “Cliente” y al lado en la columna“ Tipo de datos”¹⁰ especificamos cada atributo.

De todos los campos de la tabla uno de ellos por lo menos tiene una relevancia significativa es por ello que debemos identificarlo. En el ejemplo anterior observamos como Cod_Cliente es el campo que identifica la tabla para ello lleva asignado a su lado izquierdo una llave (clave principal). En la parte de debajo de la tabla aparecen las propiedades de los campos:

General	Búsqueda
Tamaño del campo	Entero largo
Formato	
Lugares decimales	Automático
Máscara de entrada	
Título	
Valor predeterminado	
Regla de validación	
Texto de validación	
Requerido	Sí
Indexado	Sí (Sin duplicados)
Etiquetas inteligentes	
Alineación del texto	General

Figura 14. Propiedades Tabla Cliente

¹⁰ Pueden ser Número, Texto, Moneda, Hipervínculo, Datos adjuntos entre otros.

En el ejemplo anterior observamos las diferentes propiedades asignadas al campo Cod_Cliente. En la pestaña “GENERAL” resaltamos el tamaño del campo que en este caso el tipo de dato al ser un número existen diferentes opciones y para mayor facilidad a la hora de incorporar los números le hemos asignado “Entero Largo”.

Otro dato a resaltar son las propiedades requerido e indexado, pues cuando estemos asignando las propiedades de un atributo clave (como es el caso) a las dos propiedades le asignaremos sí y además en indexado “sí sin duplicado” (la clave primaria no admite duplicados, como así se mencionó anteriormente).

Por otro lado, cuando un campo precise de validación de otra tabla o necesitemos especificar su valor directamente, podemos especificar esas características del campo “BÚSQUEDA”. Siguiendo con el ejemplo anterior, el campo Id_Población (tabla Cliente) lo tenemos que asignar a la tabla Población.

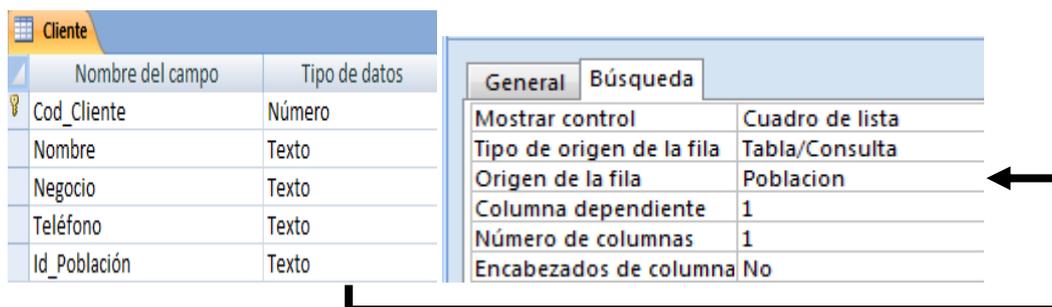


Figura 15. Ejemplo de Clave ajena en la Tabla Cliente

De esta forma una vez cubiertos los datos de la tabla Población, cuando pasemos a cubrir la tabla cliente , en el campo Población ya no tendremos que escribir pues aparecerán diferentes opciones (las incorporadas en la tabla Población) en una pestaña que funciona como filtro.

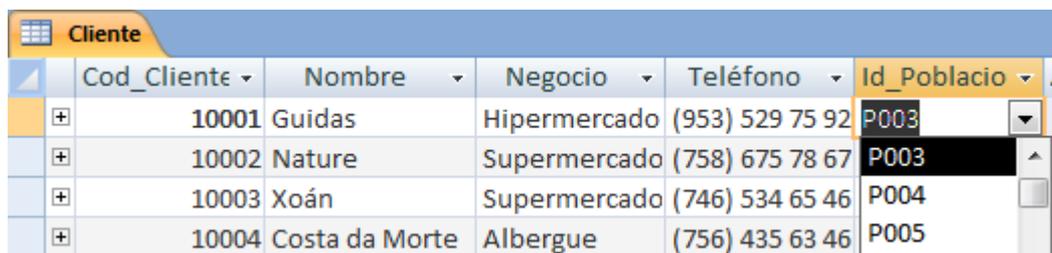


Figura 16. Ejemplo de filtro “Población”

3.5.2. Relaciones

Para definir una relación entre dos tablas es preciso indentificar que campos tienen en común las tablas. Ambos campos son claves principales en sus respectivas tablas, pero si nos encontramos ante una relación uno a varios, el campo de la tabla principal se le denomina tabla primaria mientras que el campo de la tabla secundaria se le denomina clave ajena como hemos visto anteriormente.

Las relaciones tienen como misión la recuperación rápida de datos al estar la tabla secundaria indexada a la tabla principal (Andersen, 2007). Además es preciso que las relaciones se realicen una vez construídas las tablas cuando aún no tiene introducido ningún dato.

A continuación se muestra un ejemplo del cómo crear relaciones entre dos tablas. En este caso entre Cliente y Población.

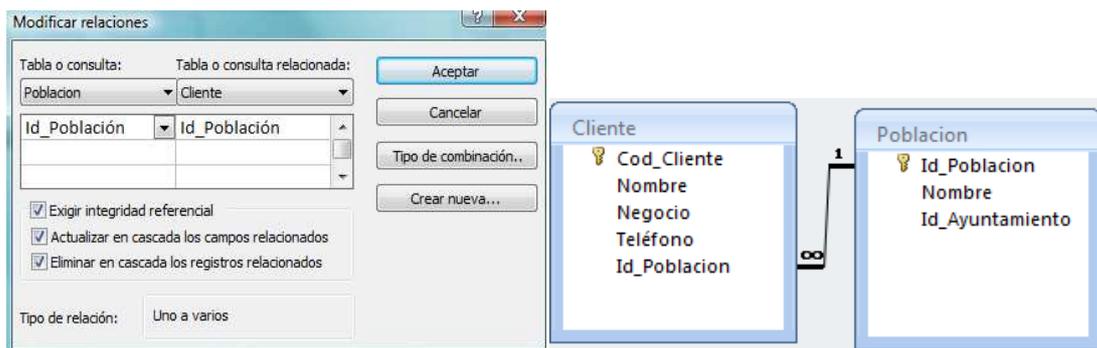


Figura 17. Construcción de relaciones tabla Cliente y Población.

En el Anexo 4 se muestran las relaciones de la base de datos creadas para la distribuidora.

3.5.3. Consultas

Las consultas en Access pueden ser de cuatro categorías: de selección, de propósito especial, de acción y específicas de lenguajes de consulta estructurado (SQL) (Andersen, 2007). En este trabajo fin de grado se escogió la opción “consultas SQL específicas” las cuales sólo son accesibles a través del lenguaje SQL. Todas las opciones tienen lenguaje SQL la diferencia es que la cuarta opción (la elegida) utiliza un lenguaje de programación.

Antes de adentrarnos a la realización de consultas, es preciso explicar la estructura básica de consultas a través de SQL.

3.5.3.1. Estructura Básica de consultas SQL

Las sentencias de SQL comienzan con una forma verbal y se expresan en mayúsculas, seguida del resto de cláusulas que pueden ser obligatorias u opcionales (Andersen, 2007) .A continuación se muestran algunas de las cláusulas que más se utilizan:

- **SELECT** (obligatoria)

Incluye los datos que se solicitan en la consulta, normalmente una o varias expresiones. Alternativamente un * indica todas las columnas de las tablas especificadas. Si no quisiéramos que aparecieran filas repetidas tendríamos que poner DISTINCT.

- **FROM** (obligatoria)

Determina la tabla o tablas de las cuales queremos extraer los datos.

- **WHERE** (opcional)

Indica un predicado que expresa la condición que debe cumplir cada fila que interviene en la consulta. Así la consulta se limita a las filas que cumplen la condición.

- **ORDER BY** (opcional)

Permite determinar el criterio de ordenación de las filas de la tabla resultado. Sin ella se obtendrán las mismas filas, pero puede ser en órdenes distintos, depende de la estrategia seguida polo SGBD para extraer los datos.

- **GROUP BY** (opcional)

Contiene funciones de agregado y muestra los campos que no se resumen en la cláusula SELECT.

- **HAVING** (opcional)

Contiene funciones de agregado, pero a diferencia del anterior, especifica las condiciones que se aplican a los campos que se resumen en la instrucción SELECT.

4. Alimentación de la base de datos

El proceso de alimentar una base de datos consiste en incorporar datos en cada una de las tablas creadas con el fin de poder realizar consultas que sean útiles para tomar decisiones. Los datos que se refieren a nombres de personas, denominaciones sociales, matrículas de vehículos o bien códigos no son reales debido a la ley de protección de datos. Por otro lado, páginas web como Granxa Familiar nos ha permitido conocer que productos ecológicos se cultivan.

Para iniciar la alimentación de la base de datos nos dirigimos a una de las tablas, por ejemplo, Productor Ecológico y comenzamos a incorporar datos (véase figura 18):

Cod_Produc	Nombre	Teléfono	Id_Poblacio
PR01	O Campo	(964) 287 65 82	P001
*			P002
			P003

Figura 18 .Alimentación de la Tabla Productor Ecológico

A continuación podemos observar algunas de las tablas con su información almacenada:

Cod_ProductorEco	Nombre	Teléfono	Id_Població
PR01	O Campo	(964) 287 65 82	P029
PR02	Rellán	(754) 627 36 57	P020
PR03	Mato	(746) 523 52 37	P019
PR04	Constante	(765) 823 68 56	P018
PR05	Nicolás	(765) 728 65 72	P017
PR06	Julián	(927) 598 27 95	P016

Figura 19.- Tabla Productor Ecológico

Producto			
	Cod_Produc	Precio	Nombre
+	A01	5,00 €	Alubia (1kg)
+	A02	3,00 €	Aguacate (1kg)
+	A03	3,50 €	Arroz largo Integral (1kg)
+	A04	1,42 €	Acelga (1kg)
+	A05	33,00 €	Albariño Corisca Eco (botella)
+	B01	13,50 €	Berengena (1kg)
+	B02	2,50 €	Brócoli (1kg)

Figura 20 . Tabla Producto

Compra				
	Id_Compra	Fecha esperada	Fecha Real	Cod_ProductorEco
+	C01	17-jun-11	23-jun-11	PR06
+	C02	17-jun-11	23-jun-11	PR13
+	C03	17-jun-11	23-jun-11	PR01
+	C04	06-jul-11	12-jul-11	PR15
+	C05	08-jul-11	14-jul-11	PR10
+	C06	13-jul-11	19-jul-11	PR02
+	C07	21-jul-11	27-jul-11	PR47
+	C08	04-ago-11	10-ago-11	PR40

Figura 21 . Tabla Compra

5. Explotación de la información

En este apartado se refleja el objetivo del trabajo, el análisis de la información para que sirva de soporte en la toma de decisiones. Para ello la dirección o el departamento de gestión de la empresa realizan consultas, sobre los datos almacenados en la base de datos para su posterior análisis.

Para realizar las consultas se tienen en cuenta tres análisis: Ventas, Compras y Productos. Dentro de estos análisis, la empresa realizará varias consultas sobre aquella información que considere relevante y que le permita tomar decisiones.

Por lo tanto, las consultas llevadas a cabo por la dirección de la empresa o departamento de gestión con posterior toma de decisiones son las siguientes:

ANÁLISIS DE VENTAS

En este tipo de análisis obtenemos información genérica acerca de las ventas en los periodos 2012-2014. Para ello se ha realizado, en Access, una consulta en SQL incluyendo todas aquellas tablas de las cuales queremos extraer información. Esto es:

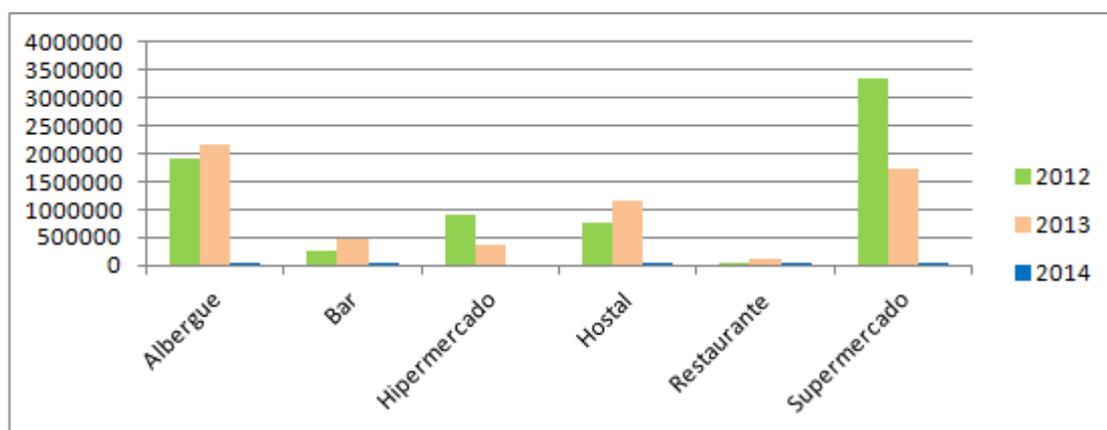
```
SELECT      Poblacion.Nombre,      Cliente.Negocio,      Venta.[Fecha      real],
[Linea_Venta]![Cantidad]*[Linea_Venta]![Precio]      AS      Importe,      Cliente.Nombre,
Ayuntamiento.Ayuntamiento, Empleado.Nombre, [Venta]![Fecha real]-[Venta]![Fecha esperada] AS
Dif_Días
FROM ((Ayuntamiento INNER JOIN Poblacion ON Ayuntamiento.Id_Ayuntamiento =
Poblacion.Id_Ayuntamiento) INNER JOIN (Empleado INNER JOIN (Cliente INNER JOIN Venta ON
Cliente.Cod_Cliente = Venta.Cod_Cliente) ON Empleado.Cod_Empleado = Venta.Cod_Empleado) ON
Poblacion.Id_Poblacion = Cliente.Id_Poblacion) INNER JOIN Linea_Venta ON Venta.Id_Venta =
Linea_Venta.Id_Venta;
```

Una vez extraída la información necesaria se procede a la creación de tablas y gráficos dinámicos en Excel. En este caso se crearon los siguientes:

1. Importe de ventas según el tipo de negocio entre los periodos 2012 y 2014.

El objetivo es analizar qué tipo de negocio ha generado mayores ingresos durante estos tres años. Es importante remarcar qué en el periodo 2014 sólo existen datos almacenados en la base de datos hasta el 5 de abril de 2014.

Gráfico 2.- Ingresos por ventas según tipo de negocio desde 2012 hasta 2014.



La empresa, durante los periodos analizados, atiende a un total de 139 ventas y las ventas recogidas del año 2014 son hasta abril, como antes se ha mencionado, por eso es el periodo que menos ingresos de ventas refleja.

Como podemos observar, dos son los negocios que mayores ingresos han generado a la distribuidora. Por un lado están las ventas al supermercado, en 2012, las cuales han superado los 300.000 euros aunque en el 2013 los ingresos bajaron más o menos hasta la mitad.

Por otro lado están las ventas al albergue que también refleja cifras significativas en cuanto a los ingresos. Podemos observar que durante el año 2012 roza los 200.000 euros y en el 2013 los supera.

Otra forma de comprobarlo sería a través de la siguiente tabla dinámica en donde podemos comparar en tanto por cien, los ingresos que supone cada negocio en relación al supermercado:

Tabla 17.- Comparación (%) de ingresos de todos los negocios en relación al supermercado.

Importes Negocios	Años			Total general
	2012	2013	2014	
Albergue	57%	124%	44%	80%
Bar	8%	27%	69%	15%
Hipermercado	27%	21%	0%	25%
Hostal	23%	67%	21%	38%
Restaurante	1%	7%	6%	3%
Supermercado	100%	100%	100%	100%
Total general				

En esta tabla podemos observar también varios datos significativos. Por un lado los ingresos por ventas al albergue en el año 2013 llegan a los del supermercado. En el 2014, el Bar consigue altos ingresos (a pesar de sólo tener datos sobre las ventas hasta el 5 de abril). Por último destacar que el restaurante es el que menos ingresos aporta pero se observa que poco a poco van ascendiendo.

La razón del porqué estos dos negocios (supermercado y albergue) aportan mayores ingresos puede deberse; por un lado, el supermercado es un tipo de negocio que está presente en la mayoría de las áreas donde suministra nuestra empresa, por su tamaño necesita más cantidad de alimentos que los otros negocios y como pudimos ver en la encuesta realizada por Craega es uno de los negocios preferentes para el consumo de estos alimentos.

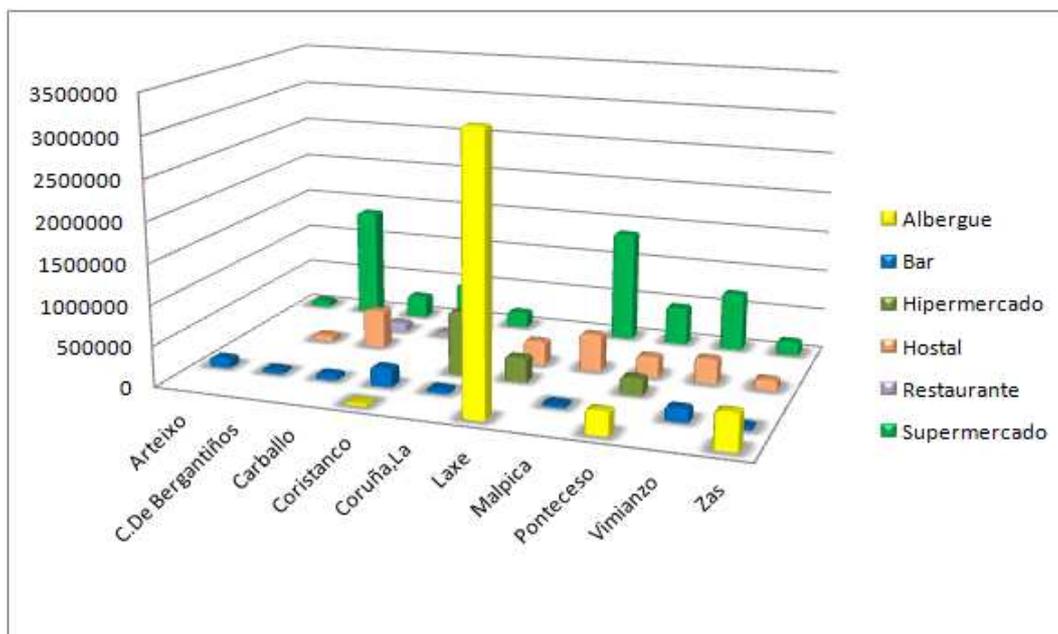
En cuanto al albergue, en la zona donde opera la empresa existen varios albergues con bastante éxito de ahí a su incremento de ventas durante 2012 y 2013.

Debido a la existencia de competidores en la zona, la distribuidora no consiguió vender a todos los negocios de forma igualitaria. Pese a esto la empresa pretende ofrecer descuentos según la cantidad de compra, para conseguir, por un lado competir con los productos convencionales y por otro lograr que los clientes confíen en este tipo de productos. En cambio en el restaurante, una buena estrategia podría ser incrementar el precio de algunos productos ecológicos para que fueran considerados como productos de alta calidad con el propósito de que negocios como los restaurantes se interesen por estos productos.

2. Importe de ventas por tipo de negocio y distribución geográfica.

Con el siguiente gráfico se pretende analizar en qué lugares (ayuntamientos) se generan más ingresos y además conocer qué tipo de negocio es el que genera tales ingresos.

Gráfico 3.- Importe por ventas/tipo de negocio según distribución geográfica.



Como hemos mencionado antes, el supermercado es el negocio que más ingresos aporta a nuestra empresa y en este gráfico podemos observar en color verde que está presente en todos los ayuntamientos¹¹ (excepto en Laxe). Por otro lado el “albergue”, en color amarillo, decir que está presente en cuatro ayuntamientos (principalmente Laxe, seguido de Zas, Ponteceso y Coristanco).

El ayuntamiento con mayor éxito es Laxe principalmente con el “Albergue”. Este ayuntamiento es reconocido como zona turística de ahí a que sea uno de los negocios que más productos necesitan y más ingresos generen.

Otro dato que llama la atención son los ingresos obtenidos en Carballo. Al principio del trabajo hemos mencionado que la empresa tiene una nave asentada en esa zona pero en el gráfico 3 podemos observar que los mayores ingresos no provienen de ese ayuntamiento.

¹¹ Nos referimos a los ayuntamientos que están registrados en la base de datos de nuestra empresa.

La empresa deberá darse más a conocer en los siguientes años (potenciar más sus estrategias de comunicación: anuncios publicitarios, revistas ecológicas) ya que, se demuestra la dificultad que tiene la distribuidora para adentrarse a suministrar en las ciudades (véase en el gráfico el caso de A Coruña, Arteixo e incluso ciudades pequeñas como Carballo).

3. Número de ventas por empleado y año.

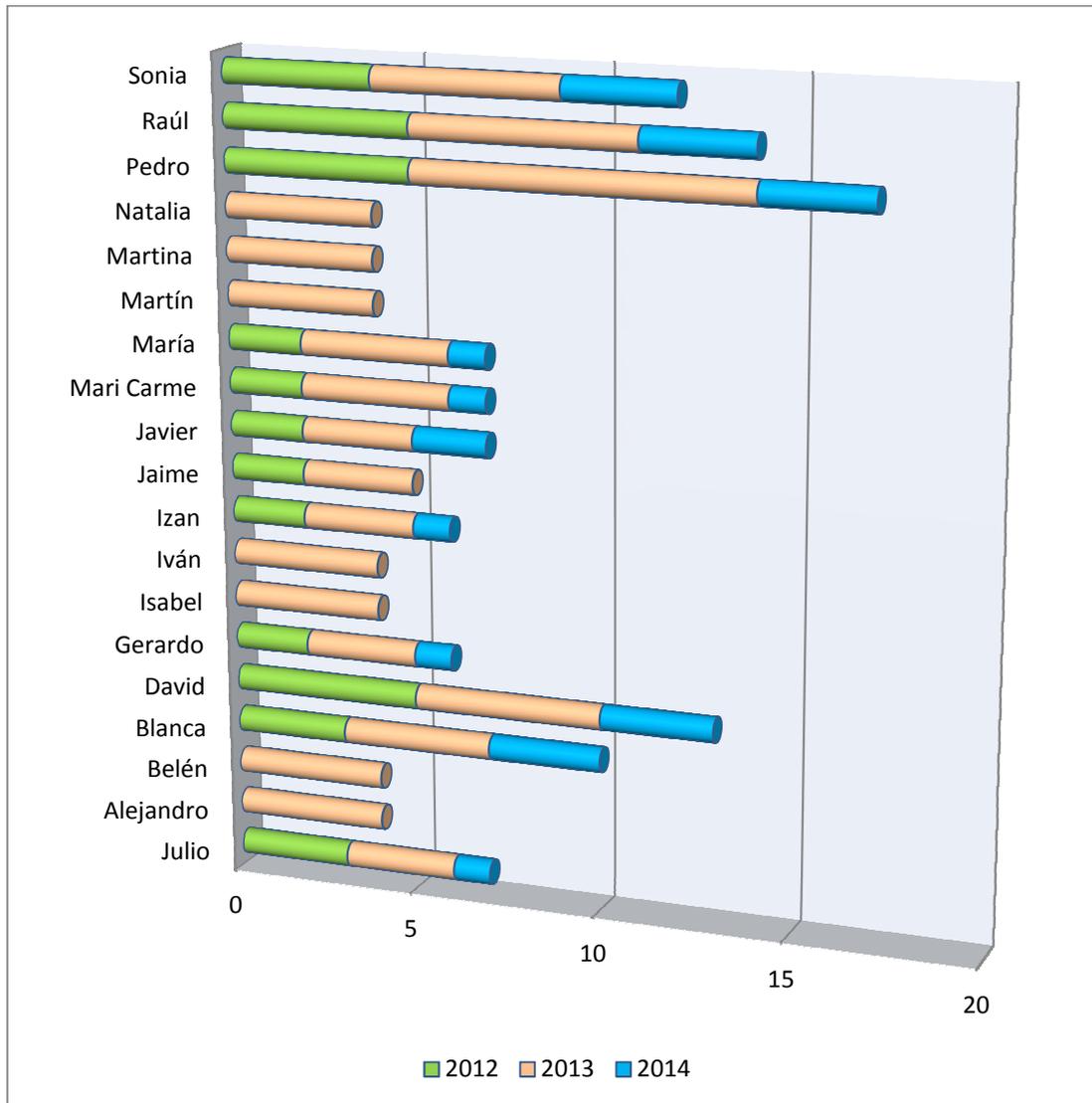
Este tipo de análisis nos permite comparar la eficiencia de los empleados. En el gráfico 4 se puede observar las entregas realizadas por cada empleado de la empresa en cada periodo.

La empresa tiene contratado un total de 20 empleados (12 fueron contratados en 2012 y 8 en el 2013). De todos los empleados cuatro superan las 10 ventas (Raúl, Sonia, Pedro y David) y en concreto Pedro supera las 15 ventas. Estos cuatro empleados se reconocen como empleados más eficientes sobre el resto de empleados contratados en su mismo año. Esto se debe a que son empleados con la formación y experiencia necesaria para desarrollar su trabajo.

Por último en el 2013 es cuando más se refleja la eficiencia de los trabajadores ya que en ese periodo se registraron el mayor número de entregas. En este dato puede influir la incorporación de 8 nuevos trabajadores en ese mismo año.

La empresa tendrá que valorar, si, en los próximos años, siguen siendo eficientes los mismos empleados, o bien consiguen que todos lo sean por igual. Por ello la empresa deberá dar formación a todos sus empleados con el objetivo de conseguir que todos sean eficientes por igual.

Gráfico 4.- N° de Ventas por empleado desde 2012 hasta 2014.

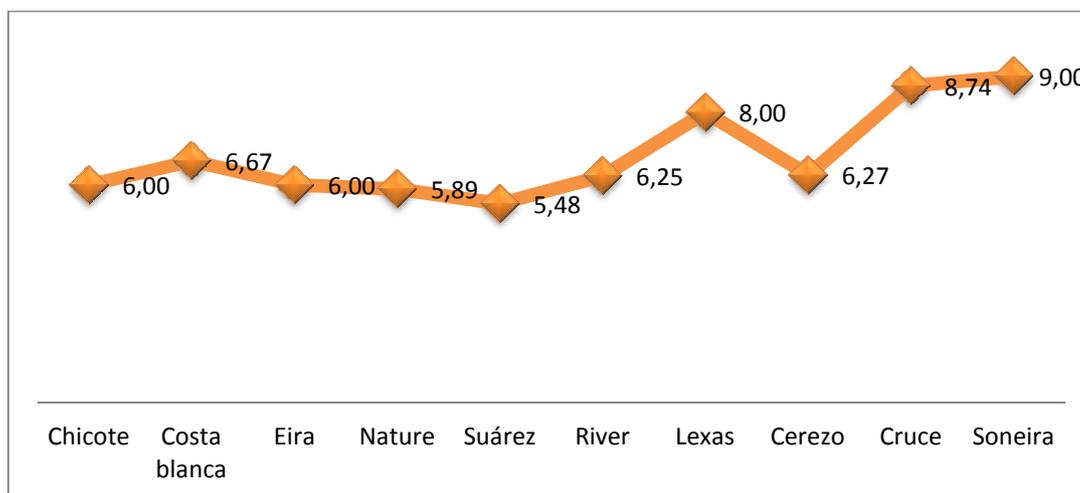


4.- Diferencia media entre la Fecha Real y Estimada en la Venta.

En el siguiente gráfico podemos observar a diez clientes quienes han sufrido más las demoras de la mercancía. La empresa poco puede hacer al respecto ya que es difícil que nunca exista ninguna diferencia de días cuando se suministra mercancía. Una de las razones que puede desviar la fecha real de la venta con la estimada es la diferencia de días que puedan existir en las fechas de compra a los productores.

Aunque la distribuidora deberá asegurarse de que no sean siempre los mismos clientes quienes reciben sus pedidos más tarde de lo previsto, en este caso se podría aplicar un descuento en el precio del producto.

Gráfico 5.- Clientes con mayor diferencia media en días en la venta.



ANÁLISIS DE COMPRAS

En este tipo de análisis obtenemos información genérica sobre las compras en los periodos 2011-2014. Para ello se ha realizado, en Access, una consulta en SQL incluyendo todas aquellas tablas de las cuales queremos extraer información. Esto es:

```
SELECT Productor_Ecologico.Nombre, Compra.[Fecha real], Linea_Compra.Cantidad,
Linea_Compra.Precio, [Cantidad]*[Precio] AS Importe, Poblacion.Nombre, [Compra]![Fecha real]-
[Compra]![Fecha esperada] AS Diferencia, Ayuntamiento.Ayuntamiento
FROM ((Ayuntamiento INNER JOIN Poblacion ON Ayuntamiento.Id_Ayuntamiento =
Poblacion.Id_Ayuntamiento) INNER JOIN Productor_Ecologico ON Poblacion.Id_Poblacion =
Productor_Ecologico.Id_Poblacion) INNER JOIN (Compra INNER JOIN Linea_Compra ON Compra.Id_Compra =
Linea_Compra.Id_Compra) ON Productor_Ecologico.Cod_ProductorEco = Compra.Cod_ProductorEco;
```

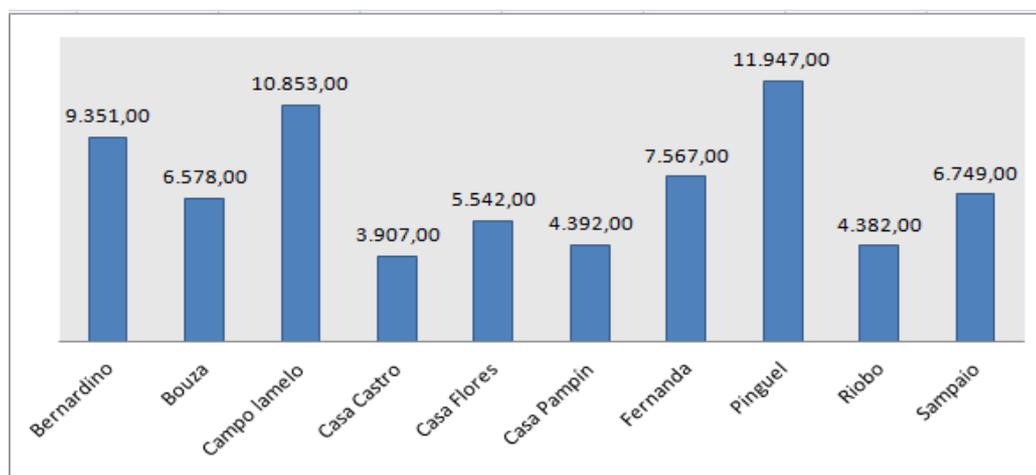
Una vez extraída la información necesaria la empresa procede a la creación de tablas y gráficos dinámicos en Excel. En este caso se crearon los siguientes:

1.-Compras realizadas (cantidad de productos) a cada Productor Ecológico.

Con este tipo de análisis la distribuidora podrá conocer el nombre de los diez productores que más cantidad de productos ecológicos suministran a la empresa.

Cada productor ecológico suministra más de un producto, por ello la empresa tendrá que decidir si le conviene prescindir de alguno ellos o no. Una de las estrategias que podrá llevar a cabo será comprobar si al año siguiente, siguen siendo los mismos productores quienes suministran mayores cantidades de productos.

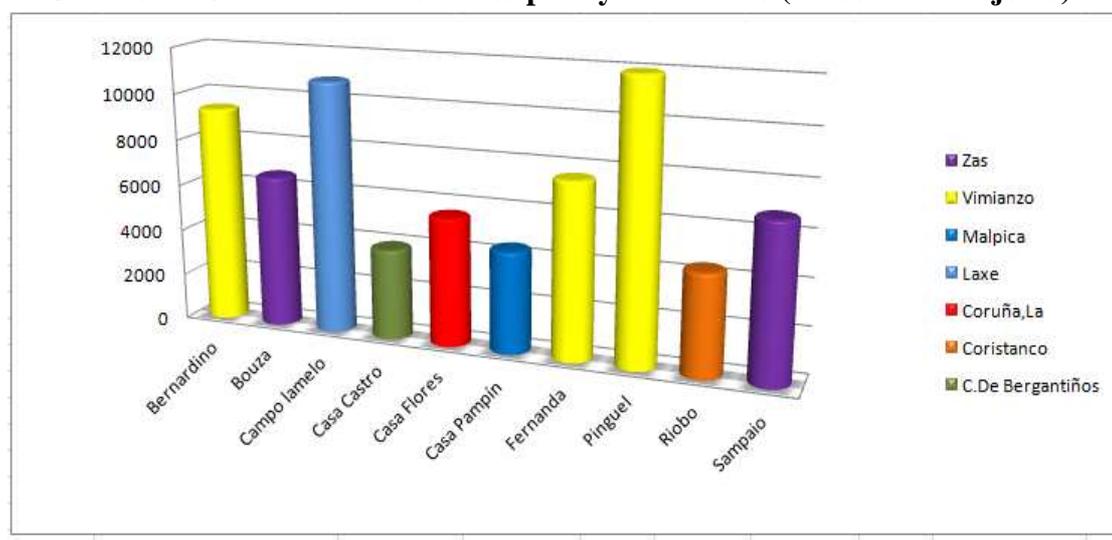
Gráfico 6.- Cantidad de producto por Productor ecológico (diez mejores).



2.- Distribución geográfica por cantidad de producto adquirido.

Con este análisis la empresa conoce el nombre del ayuntamiento en donde residen los productores que más cantidad de productos han suministrado a la distribuidora. Por orden serían: Vimianzo con tres productores los cuales superan las 6.000 cantidades, luego estaría Laxe en donde reside un único productor que aporta alrededor de 10.000 cantidades (como pudimos ver en el gráfico 3, en este ayuntamiento, debido al albergue es en donde se registran más ingresos por ventas); Zas con dos productores los cuales rozan las 6.000 entres otros.

Gráfico 7.- Cantidad de Productos por ayuntamiento (ambos diez mejores).

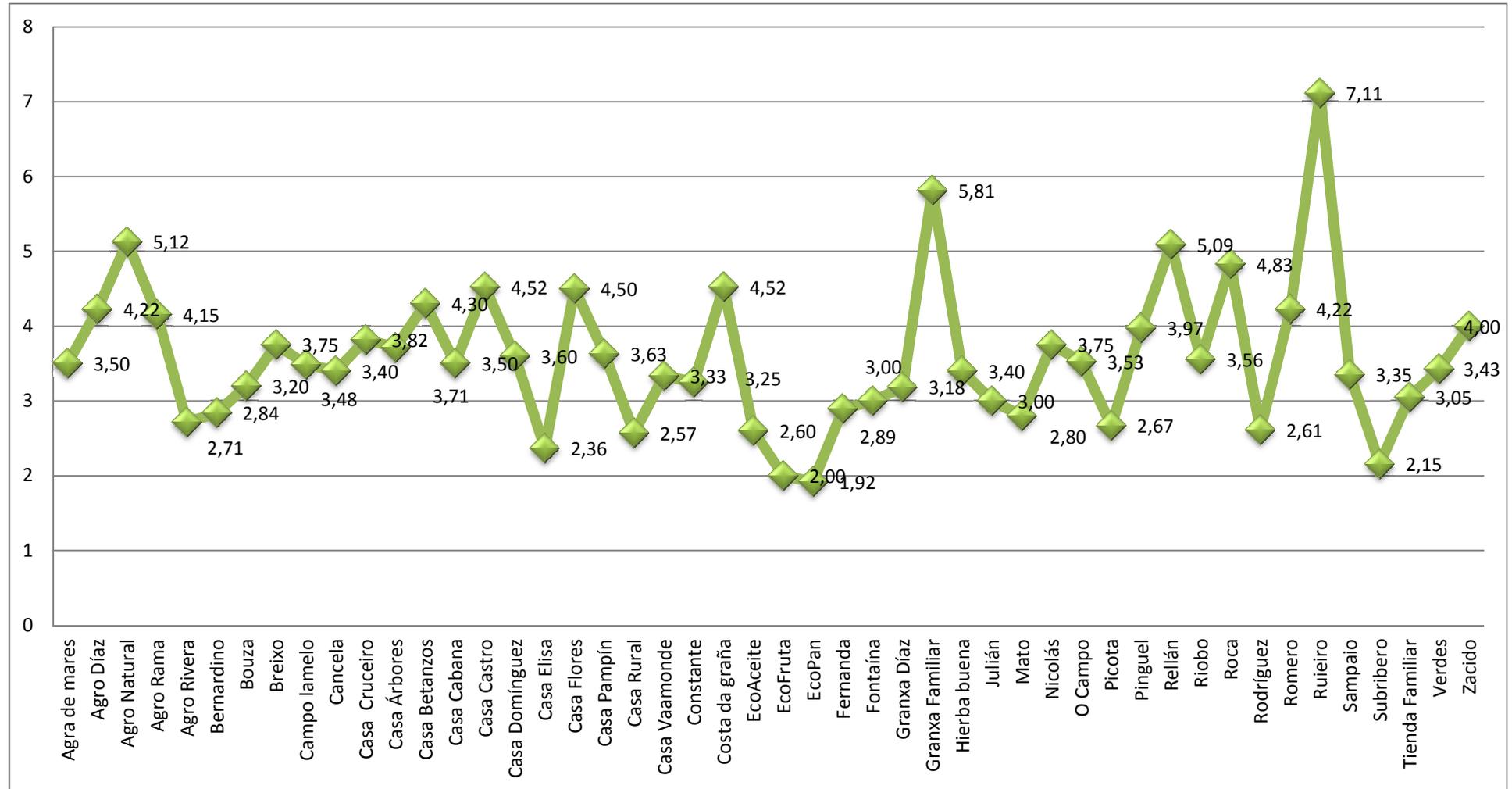


Como hemos mencionado antes, cada productor ecológico suministra más de un producto, por ello la empresa deberá decidir si le conviene prescindir de alguno de los productores. Por ejemplo la empresa tiene registrado en la base de datos a dos productores: Granxa Familiar y Granxa Díaz residentes en Muxía y Cee respectivamente, en cambio no se tiene registrado ningún cliente en esos dos ayuntamientos. Por tanto si tenemos en cuenta los gastos de desplazamientos lo mejor para la empresa es prescindir de ellos y centrarse en aquellos que están más cercanos geográficamente y que más productos le suministran.

3.- Diferencia media entre la Fecha real y Estimada en la Compra.

Como podemos comprobar también existes diferencias entre fecha real de compra y fecha estimada. Esto puede deberse a que un productor estime tener una cantidad de productos en una fecha determinada, pero la obtención de los alimentos puede retrasarse o aligerarse. Como se ha mencionado antes, esta disparidad entre los días es uno de los hechos que provoca que haya diferencias entre las fechas de venta.

Gráfico 8.- Productores y diferencia media entre la Fecha real y la estimada en las Compras.



ANÁLISIS DE PRODUCTOS

En este tipo de análisis obtenemos información sobre los productos ecológicos. Para ello se ha realizado, en Access, una consulta en SQL:

```
SELECT Producto.Nombre, [Linea_Venta]![Cantidad]*[Linea_Venta]![Precio]
AS Importe, Linea_Venta.Cantidad, Venta.[Fecha real], Cliente.Negocio,
Cliente.Nombre, Poblacion.Nombre, Ayuntamiento.Ayuntamiento
FROM ((Ayuntamiento INNER JOIN Poblacion ON
Ayuntamiento.Id_Ayuntamiento = Poblacion.Id_Ayuntamiento) INNER JOIN (Cliente
INNER JOIN Venta ON Cliente.Cod_Cliente = Venta.Cod_Cliente) ON
Poblacion.Id_Poblacion = Cliente.Id_Poblacion) INNER JOIN (Producto INNER JOIN
Linea_Venta ON Producto.Cod_Producto = Linea_Venta.Cod_Producto) ON
```

Una vez extraída la información necesaria la empresa procede a la creación de tablas y gráficos dinámicos en Excel. En este caso se crearon los siguientes:

1.- Productos según ingresos.

Con este análisis la empresa conocerá que productos han supuesto mejores y peores ingresos a la empresa. A continuación, en el gráfico 9 y 10, se observan que en ambos aparecen productos como: verduras, frutas, hortalizas...Pese a estos datos, la empresa no puede decantarse por comprar unos u otros sólo se ciñe a los pedidos de los clientes.

Gráfico 9.- Mayores ingresos por producto (diez mejores).

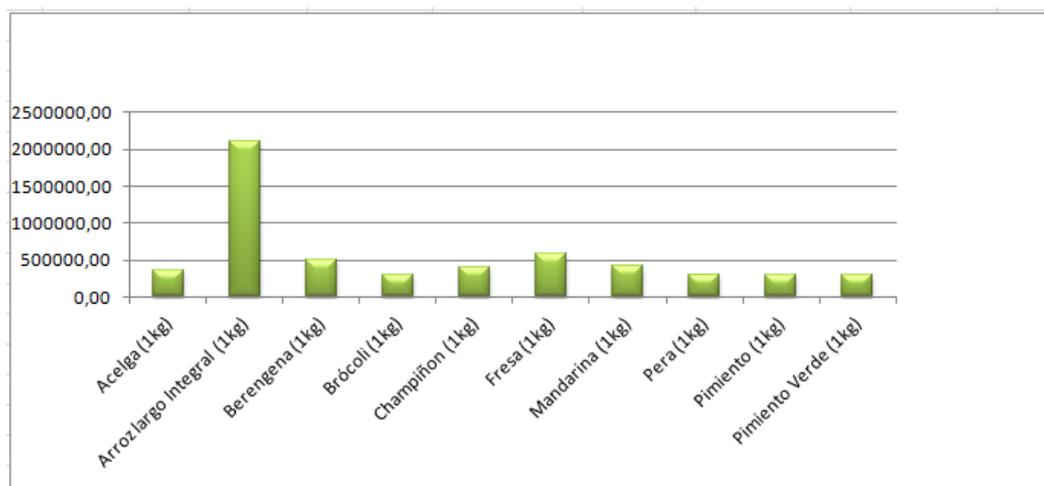
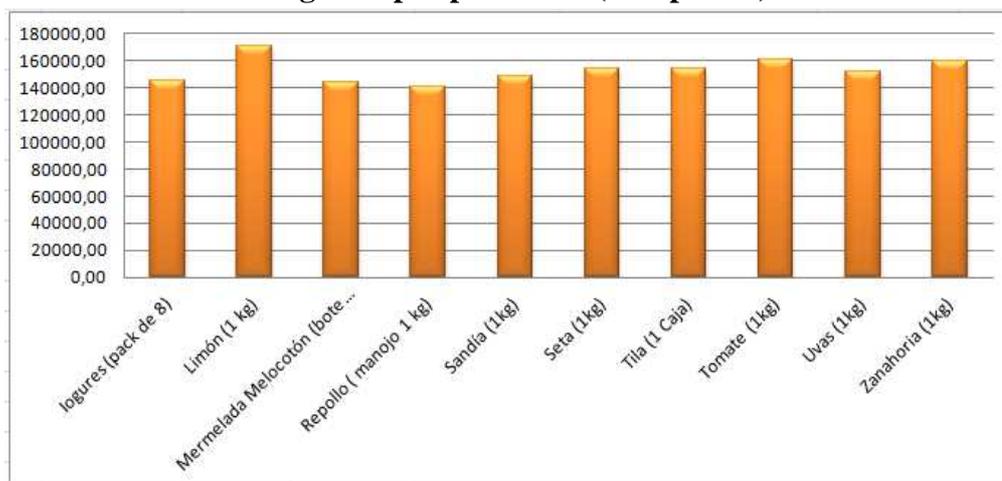


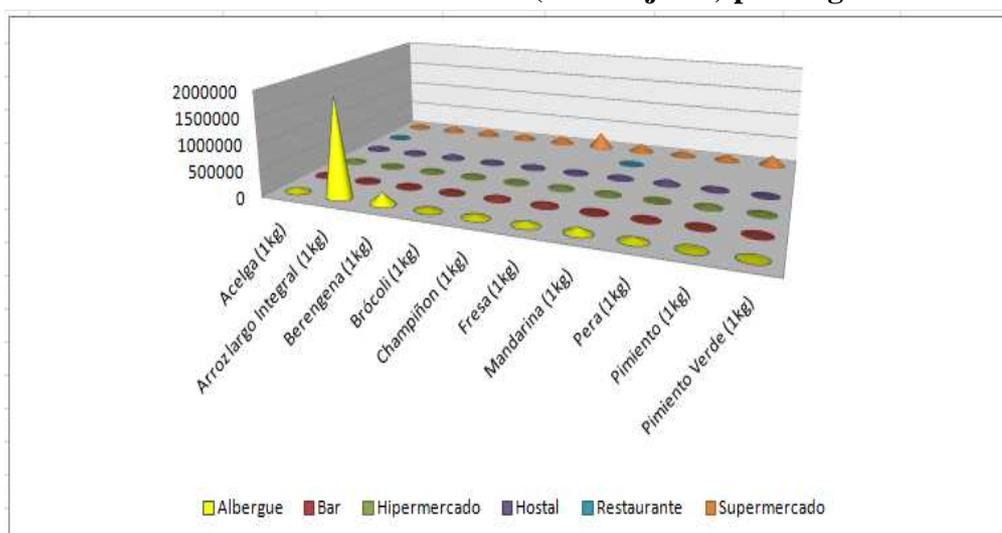
Gráfico 10.- Peores ingresos por productos (diez peores).



2.- Productos (diez mejores) por Negocio.

Con este tipo de análisis la empresa conocerá el tipo de negocio que reclama los productos que generan mayores ingresos a la distribuidora.

Gráfico 11.- Productos (diez mejores) por Negocio.



Como podemos comprobar tanto el albergue como el supermercado contienen a estos diez productos. Entre ellos podemos resaltar el arroz con un gran éxito en el albergue, y del resto con gran presencia en el supermercado. La distribuidora, una vez que es conocedora de los datos, podrá asegurarse de tener grandes cantidades de esos diez productos más demandados para atender a cada negocio.

6 Conclusiones

Una base de datos se construye sobre una arquitectura empresarial que muestra las relaciones entre los elementos que la integran, en esta empresa: clientes, productores ecológicos, departamentos (compras y ventas) y vehículos. Esta relación se muestra específica en cada empresa. Como lo mencionábamos al principio del trabajo, el objetivo es mostrar esta realidad para una empresa que se dedica a distribuir productos ecológicos. Una vez esquematizada la distribuidora, con la información que disponemos, hemos conseguido analizarla con el objetivo último de servir de apoyo en la toma de decisiones.

En este trabajo se pone de manifiesto la importancia del análisis de la información y como se pueden analizar realidades que no son evidentes.

Bibliografía

- Andersen, V. (2007). *Access 2007: Manual de referencia*. Madrid: McGraw Hill.
- Andersen, V. (2007). *Microsoft Access 2007 paso a paso*. Madrid: McGraw Hill.
- Calomarde, J. (2000). *Marketing Ecológico*. España: Pirámide-Esic.
- Casas Luengo, J. (2007). *Access 2007*. Madrid: Anaya Multimedia
- Connolly, T. y Begg, C. (2005). *Sistemas de base de datos 4/e* (4º ed.). Madrid: Pearson Addison-Wesley.
- Cuadra, D., Castro, E., Iglesias, A.Mª., Martínez, P., Calle, F.J., Harith, C.deP., Moreno, L., García, S., Martínez, J.L., Rivero, J., Segura, I. (2013). *Desarrollo de base de datos: casos prácticos desde el análisis a la implementación* (2º ed.). Madrid: RA-MA
- Date, C.J. y Ruíz Faudón, S.L.M. (2001). *Introducción a los sistemas de bases de datos* (7º ed.). México: Pearson Educación.
- De Miguel Castaño, A y Piattini Velthuis, M. G. (1997). *Fundamentos y modelos de base de datos*. Madrid: RA-MA.
- Elmasri, R. & B. Navathe, S. (2007). *Fundamentos de sistemas de bases de datos* (5º ed.). Madrid: Addison-Wesley.
- Inmon, W.H. (2002). *Building the Data Warehouse Evolution* (3º ed.). United States of America: MacAllister Publishing Services, LLC.
- Marqués, M. (2011). *Base de datos*. Publicaciones de la Universitat Jaume I. Servei de Comunicació i Publicacions.

Silberschatz, A., Korth, H. F. y Sudarshan, S. (2006). *Fundamentos de bases de datos* (5ª ed.). Madrid: McGraw-Hill.

Codd, E. F. (1970). A relational model of data for large shared data banks. *Communications of the ACM*, 13(6), 377-387. doi:10.1145/362384.362685

Pin-Shan Chen, P. (1976). The Entity-Relationship Model-Toward a Unified View of Data. *ACM Transactions on Database Systems*, 1, 9-36.

Marcano Aular, Y.J. y Talavera Pereira, R. (2007). Minería de Datos como soporte a la toma de decisiones. *Revistas Científicas y Humanísticas*, 23, nº52,104-118. Recuperado de: <http://revistas.luz.edu.ve/index.php/op/article/view/2028/1956>

Consello Regulador de Agricultura Ecolóxica Galega (2013). Percepción e consumo de alimentos ecolóxicos en Galicia. Recuperado de:

<http://www.craega.es/index.php/agricultura-ecoloxica-ga/estudios-ga>

Ministerio de Agricultura, alimentación y medio ambiente. (2001-2012) .Estadísticas de agricultura ecológica 2001-2012. Recuperado de:

<http://www.magrama.gob.es/es/alimentacion/temas/la-agricultura-ecologica/documentos-de-interes/>

Ministerio de Agricultura, alimentación y medio ambiente. (2012). El número de establecimientos dedicados a actividades industriales ecológicas aumenta un 11,63 por ciento. Recuperado de:

<http://www.magrama.gob.es/es/prensa/noticias/el-n%C3%BAmero-de-establecimientos-dedicados-a-actividades-industriales-ecol%C3%B3gicas-aumenta-un-11,63-por-ciento-/tcm7-295116-16>

Ministerio de medio ambiente, y medio rural y marino (2009). Marketing y elementos ecológicos: *Manual de aplicación a la venta detallista*. Recuperado de: http://www.magrama.gob.es/es/alimentacion/temas/la-agricultura-ecologica/marketing_tcm7-8096.pdf

Anexos

Anexo 1

Gráfico 12.- Puntos de compra preferente de productos ecológicos.



Fuente: Craega 2010

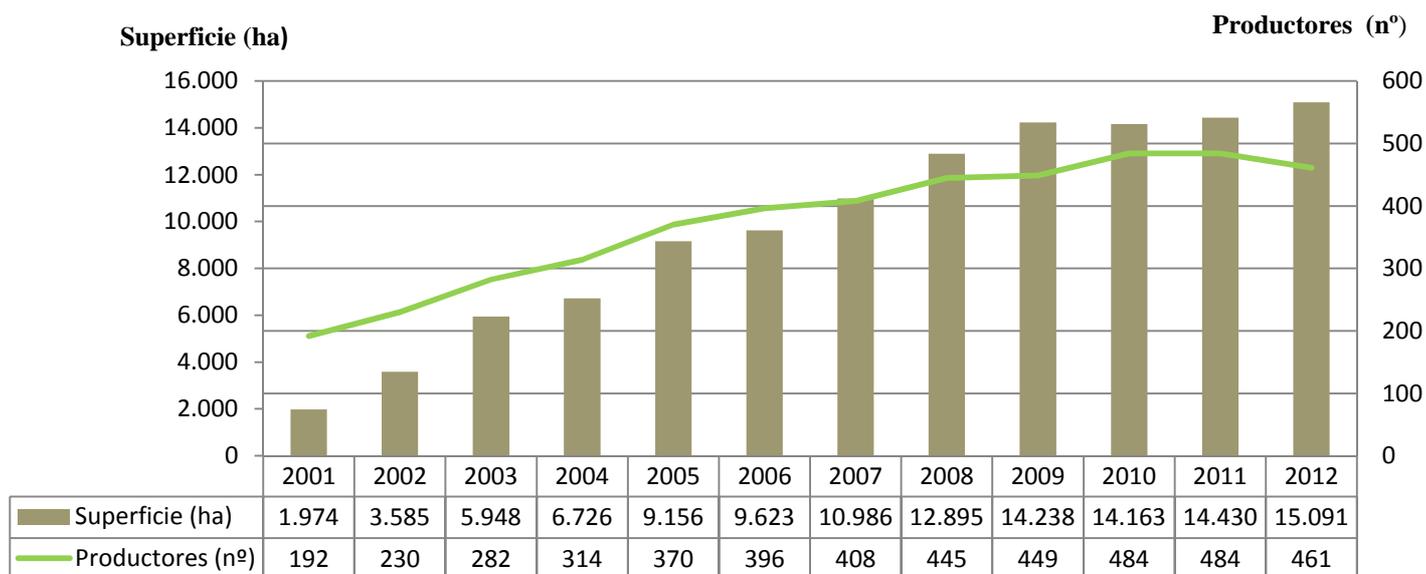
Gráfico 13.- Medios utilizados para dar a conocer los productos ecológicos.



Fuente: Craega 2010

Anexo 2

Gráfico 14.- Evolución del número de productores ecológicos y de la superficie cultivada en Galicia 2001-2012.



Fuente: Elaboración propia a partir de datos extraídos del Ministerio de Agricultura, alimentación y Medio ambiente (2001-2012).

Anexo 3

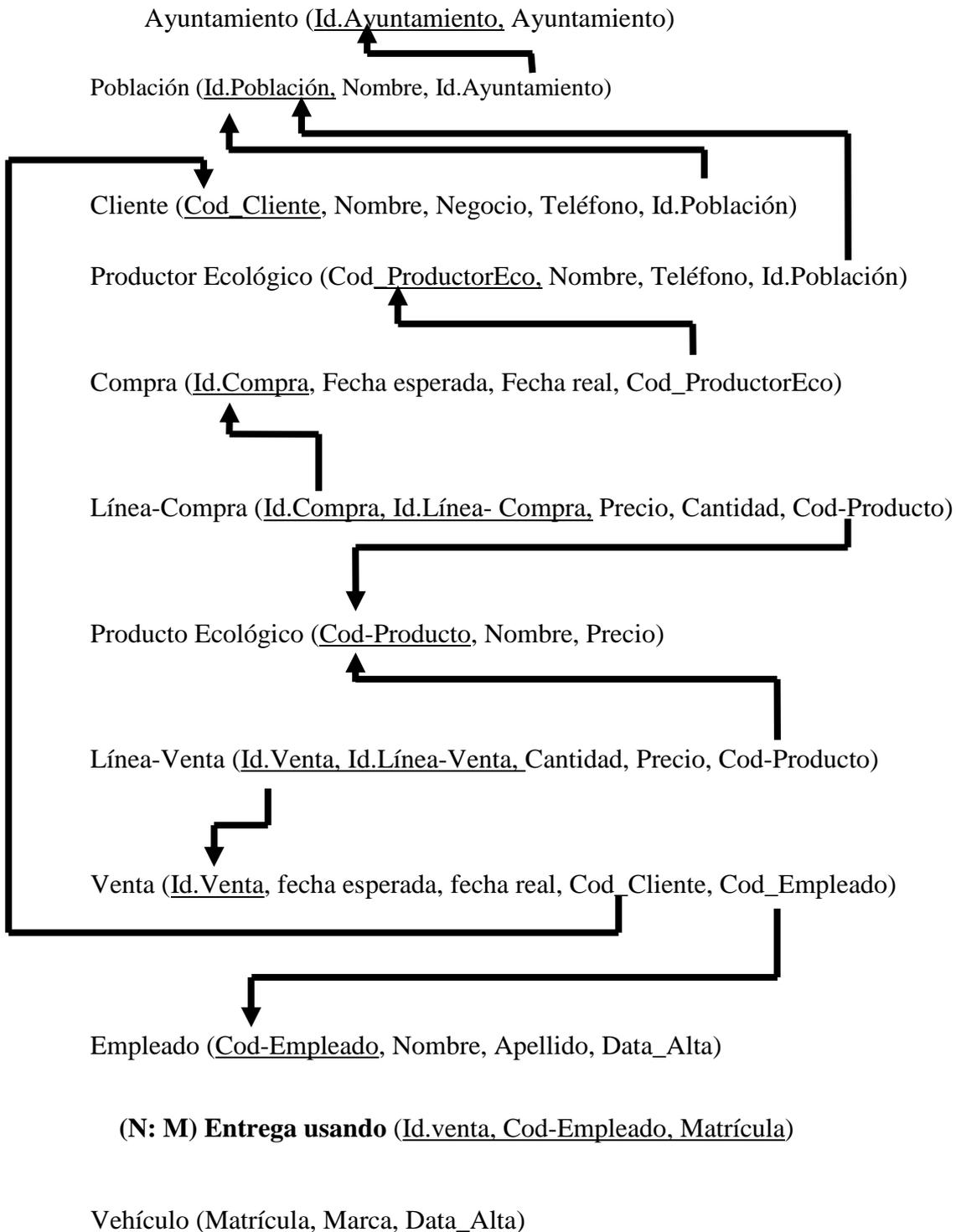


Figura 22. Transformación del diagrama ER al modelo relacional.

Anexo 4

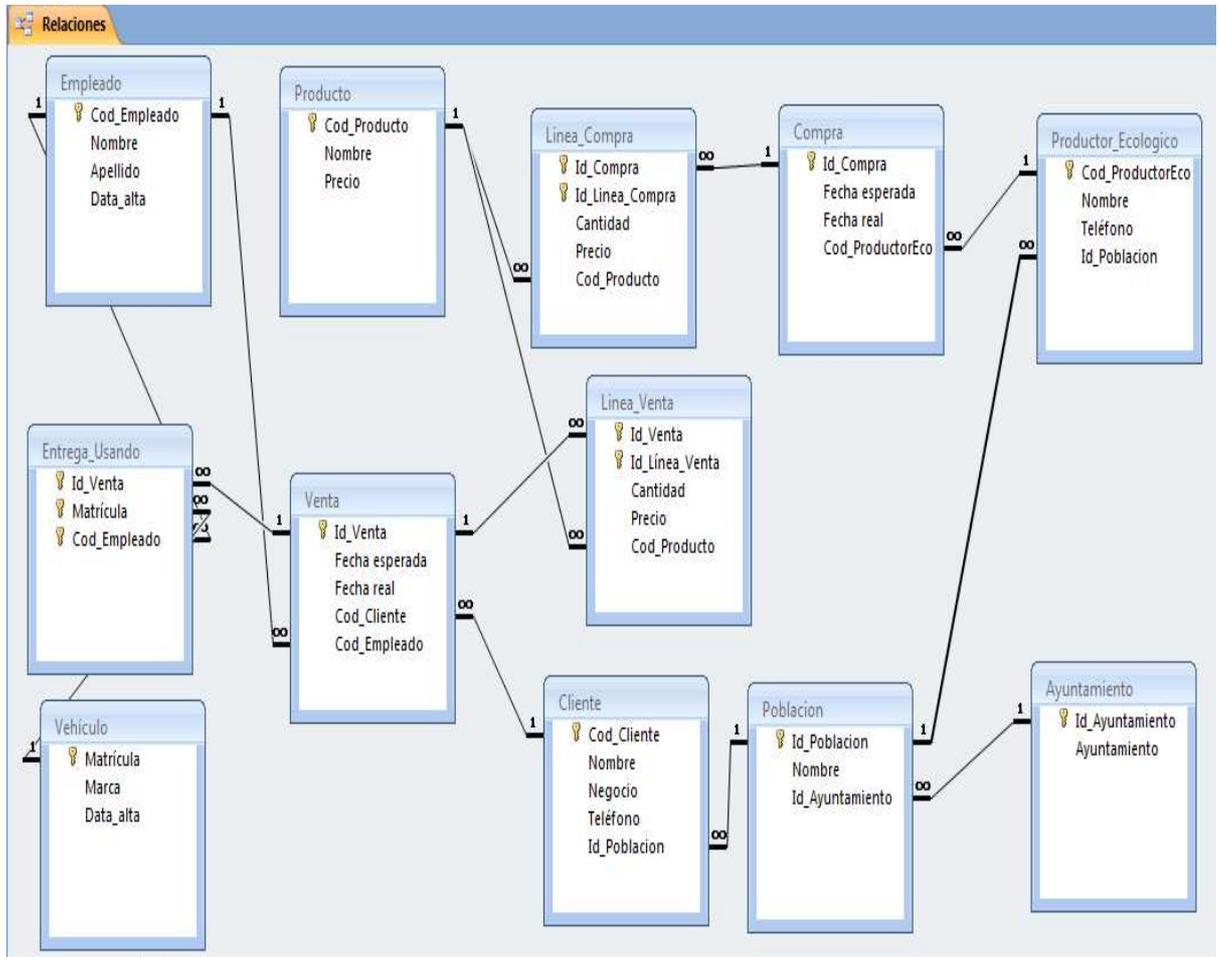


Figura 23.- Relaciones de las tablas de la base de datos

