



**Universidad
Europea**

UNIVERSIDAD EUROPEA DE MADRID

ESCUELA DE ARQUITECTURA, INGENIERÍA Y DISEÑO

GRADO EN INGENIERÍA INFORMÁTICA

PROYECTO FIN DE GRADO

**COMPARADOR ECONÓMICO DE
PRODUCTOS EN SUPERMERCADOS CON
LA AYUDA DE UN CHATBOT**

SOFÍA CORRAL CABALLERO

Dirigido por

ÓSCAR CABANILLAS NUÑEZ

CURSO 2023-2024

TÍTULO: COMPARADOR ECONÓMICO DE PRODUCTOS EN SUPERMERCADOS CON LA AYUDA DE
UN CHATBOT

AUTOR: SOFÍA CORRAL CABALLERO

TITULACIÓN: GRADO EN INGENIERÍA INFORMÁTICA

DIRECTOR/ES DEL PROYECTO: ÓSCAR CABANILLAS NUÑEZ

FECHA: Junio de 2024

1. RESUMEN

Con la subida de los precios en los supermercados en estos últimos años, los españoles comparan mucho más el coste de cada artículo, pero para ello deben ir a diferentes establecimientos o buscar en cada una de sus páginas webs por separado.

De ahí surge SUPER-ahorro, un proyecto que pretende resolver el problema con la creación de una aplicación capaz de mostrar al usuario el coste de cada producto obtenido mediante técnicas de web scraping ejecutadas semanalmente para obtener dicha información de diferentes supermercados, y, así, mostrar una comparativa para que el cliente pueda decidir el establecimiento más económico haciendo uso de un único servicio. Además, se pretende ayudar a conocer la evolución en el tiempo del precio de cada artículo, su evolución con respecto a la semana anterior y su predicción de los precios a futuro. Con la finalidad de mejorar la experiencia del usuario, se implementa un chatbot para asistir en el uso de la aplicación y dar respuesta a diversas preguntas relacionadas.

Palabras clave: Comparador de precios de productos en supermercados, Chatbot, Web scraping, Aplicación, ParseHub, Power BI

2. ABSTRACT

With the rise of prices in the supermarkets in recent years, the Spanish people compare much more the cost of each item, but to do so they must go to different stores or search in each of their web pages individually.

SUPER-ahorro, a project that aims to solve the problem with the creation of an application capable of showing the user the cost of each product obtained through web scraping techniques executed weekly to get this information from different supermarkets, and thus show a comparison so that the customer can decide the cheapest store using a single service. In addition, it is intended to help to know the evolution over time of the price of each item, its evolution with respect to the previous week and its prediction of future prices. In order to improve the user experience, a chatbot is implemented to assist in the use of the application and provide answers to various related questions.

Keywords: Supermarket product price comparator, Chatbot, Web scraping, Application, ParseHub, Power BI

AGRADECIMIENTOS

A todos los que me han apoyado durante estos años, familiares, amigos y profesores que han hecho más ameno y llevadero este camino. En especial a mis padres que, con mucho esfuerzo, me han dado la oportunidad de tener el futuro que desde pequeña quería.

Cita - frase célebre / Dedicatoria

“La satisfacción radica en el esfuerzo, no en el logro.

El esfuerzo total es la victoria total”

3. TABLA RESUMEN

	DATOS
Nombre y apellidos:	Sofía Corral Caballero
Título del proyecto:	Comparador económico de productos en supermercados con la ayuda de un chatbot
Directores del proyecto:	Óscar Cabanillas Núñez
El proyecto ha implementado un producto: (esta entrada se puede marcar junto a la siguiente)	SI
El proyecto ha consistido en el desarrollo de una investigación o innovación: (esta entrada se puede marcar junto a la anterior)	SI (Innovación)
Objetivo general del proyecto:	Crear un servicio capaz de comparar los precios de productos entre diferentes supermercados, comprobando su evolución y aportando valor al usuario mediante un chatbot de ayuda.

Índice

RESUMEN	3
ABSTRACT	4
TABLA RESUMEN	7
Capítulo 1. ANTECEDENTES / ESTADO DEL ARTE	15
1.1 Introducción.....	15
1.2 Estado del arte	15
1.3 Contexto y justificación	21
1.4 Planteamiento del problema	22
Capítulo 2. OBJETIVOS.....	24
2.1 Objetivos generales	24
2.2 Objetivos específicos	24
2.3 Beneficios del proyecto	24
Capítulo 3. DESARROLLO DEL PROYECTO.....	25
3.1 Planificación del proyecto.....	25
3.2 Descripción de la solución, metodologías y herramientas empleadas	28
3.2.1 Descripción de la solución.....	28
3.2.1.1 Obtención de los datos.....	29
3.2.1.2 Limpieza del dato.....	44
3.2.1.3 Modelo de datos.....	57
3.2.1.4 Calidad del dato	58
3.2.1.5 Aplicación.....	62
3.2.2 Metodología	76
3.2.3 Herramientas empleadas	78
3.3 Recursos requeridos	78
3.4 Presupuesto	79
3.5 Viabilidad	80
3.6 Resultados del proyecto	81
Capítulo 4. DISCUSIÓN	83
Capítulo 5. CONCLUSIONES.....	84
5.1 Conclusiones del trabajo.....	84
5.2 Conclusiones personales.....	84
Capítulo 6. FUTURAS LÍNEAS DE TRABAJO	85

Capítulo 7. REFERENCIAS.....	86
Capítulo 8. ANEXOS	89

Índice de Figuras

Figura 1. Porcentaje variación del precio en marzo de 2022 y 2023 (Statista, 2023)

Figura 2. Ranking de las ciudades donde más se ahorra en la compra (Los Supermercados Más Baratos de 2023, 2023)

Figura 3. Ranking de las ciudades donde menos se ahorra en la compra (Los Supermercados Más Baratos de 2023, 2023)

Figura 4. Mapa con los supermercados más elegidos por CCAA (Minutos, 2022)

Figura 5. Mapa de los supermercados con más establecimientos por CCAA (Minutos, 2022)

Figura 6. Ejemplo del primer nivel del conjunto de datos de Mercadona

Figura 7. Ejemplo del segundo nivel del conjunto de datos de Mercadona

Figura 8. Ejemplo del fichero de Mercadona final resultante con los campos extraídos y los añadidos posteriormente

Figura 9. Ejemplo del primer nivel del conjunto de datos de Dia

Figura 10. Ejemplo del Segundo nivel del conjunto de datos de Dia

Figura 11. Ejemplo del tercer nivel del conjunto de datos de Dia

Figura 12. Ejemplo del fichero de Dia final resultante con los campos extraídos y los añadidos posteriormente

Figura 13. Imagen de las funciones/pasos a los que llama el pipeline main para obtener el conjunto de datos de Mercadona

Figura 14. Imagen de ejemplo que muestra cómo se seleccionan en ParseHub los datos que se quieren extraer de Mercadona

Figura 15. Imagen del proceso que se ejecuta para cada una de las categorías a las que llama el fichero main para extraer los datos de Mercadona

Figura 16. Ejemplo del formato e información que se recoge sobre los datos de Mercadona

Figura 17. Imagen de las funciones/pasos a los que llama el pipeline main para obtener el conjunto de datos de Dia

Figura 18. Imagen de ejemplo que muestra cómo se seleccionan el primer nivel en ParseHub los datos que se quieren extraer de Dia

Figura 19. Imagen de ejemplo que muestra cómo se seleccionan el segundo nivel en ParseHub los datos que se quieren extraer de Dia

Figura 20. Imagen de ejemplo que muestra cómo se seleccionan el tercer nivel en ParseHub los datos que se quieren extraer de Dia

Figura 21. Ejemplo del resultado obtenido tras cargar el fichero con los datos de Mercadona en Power BI

Figura 22. Ejemplo del resultado obtenido tras promover los encabezados del conjunto de datos de Mercadona

Figura 23. Ejemplo del resultado obtenido tras crear la columna de Supermercado en el conjunto de datos de Mercadona

Figura 24. Ejemplo del resultado obtenido tras crear la columna de Fecha en el conjunto de datos de Mercadona

Figura 25. Ejemplo del resultado obtenido tras configurar el tipo de cada dato correctamente en el conjunto de datos de Mercadona

Figura 26. Ejemplo del resultado obtenido tras dividir la columna 'name' en el conjunto de datos de Mercadona

Figura 27. Ejemplo del resultado obtenido tras dividir la columna 'precio' en el conjunto de datos de Mercadona

Figura 28. Ejemplo del resultado obtenido tras modificar el nombre de las columnas en el conjunto de datos de Mercadona

Figura 29. Ejemplo del resultado obtenido tras fraccionar la columna 'precio' en el conjunto de datos de Mercadona

Figura 30. Ejemplo del resultado obtenido tras rellenar hacia abajo los datos del campo 'categoría' en el conjunto de datos de Mercadona

Figura 31. Ejemplo del resultado obtenido tras crear una nueva columna ID en el conjunto de datos de Mercadona

Figura 32. Ejemplo del resultado obtenido tras cargar el fichero con los datos de Dia en Power BI

Figura 33. Ejemplo del resultado obtenido tras dividir la columna por delimitador en el conjunto de datos de Dia

Figura 34. Ejemplo del resultado obtenido tras promover los encabezados del conjunto de datos de Dia

Figura 35. Ejemplo del resultado obtenido tras crear la columna de "Supermercado" en el conjunto de datos de Dia

Figura 36. Ejemplo del resultado obtenido tras crear la columna de "Fecha" en el conjunto de datos de Dia

Figura 37. Ejemplo del resultado obtenido tras configurar el tipo de cada dato correctamente en el conjunto de datos de Dia

Figura 38. Ejemplo del resultado obtenido tras modificar el nombre de las columnas en el conjunto de datos de Dia

Figura 39. Ejemplo del resultado obtenido tras rellenar hacia abajo los datos del campo 'categoría', 'url' y 'nombre' en el conjunto de datos de Dia

Figura 40. Ejemplo del resultado obtenido tras agrupar las filas según el nombre del producto en el conjunto de datos de Dia

Figura 41. Ejemplo del resultado obtenido tras la creación de una columna nueva 'ListaTemporal' en el conjunto de datos de Dia

Figura 42. Ejemplo del resultado obtenido tras expandir la columna compuesta por una lista en el conjunto de datos de Dia

Figura 43. Ejemplo del resultado obtenido tras la creación de una columna nueva 'ListaUnida' en el conjunto de datos de Dia

Figura 44. Ejemplo del resultado obtenido tras la expansión de la columna 'ListaUnida' en el conjunto de datos de Dia

Figura 45. Ejemplo del resultado obtenido tras la división de la columna generada con el precio en el conjunto de datos de Dia

Figura 46. Ejemplo del resultado obtenido tras la eliminación de los paréntesis en el conjunto de datos de Dia

Figura 47. Ejemplo del resultado obtenido tras dividir la columna precio en el conjunto de datos de Dia

Figura 48. Esquema del modelo de datos implementado y su cardinalidad

Figura 49. Página creada en Power BI con el conjunto entero de datos de Dia

Figura 50. Página creada en Power BI con el conjunto entero de datos de Mercadona

Figura 51. Página creada en Power BI para comprobar la calidad del dato de ambos conjuntos de datos

Figura 52. Página creada en Power BI que muestra la evolución del precio por producto de Mercadona

Figura 53. Página creada en Power BI que muestra la evolución del precio por producto de Dia

Figura 54. Prototipo de la pestaña de "Carrito" para la aplicación

Figura 55. Prototipo de la pestaña de "Ahorro" para la aplicación

Figura 56. Prototipo de la pestaña de "Análisis" para la aplicación

Figura 57. Prototipo de la pestaña del chatbot para la aplicación

Figura 58. Imagen del visual del filtro de productos que muestra la opción de buscador implementada

Figura 59. Imagen del visual del filtro de productos que muestra la opción seleccionar múltiples productos

Figura 60. Imagen de la matriz que compara los precios de ambos supermercados

Figura 61. Gráfica lineal que muestra la evolución en el tiempo del precio de los productos para ambos supermercados

Figura 62. KPIs que muestran el total del precio de los productos seleccionados para ambos supermercados

Figura 63. KPIs que muestran el precio total de los productos seleccionados y su comparativa con el precio de la semana anterior para ambos supermercados

Figura 64. KPI que indica el número total de productos añadidos a la cesta de la compra

Figura 65. KPIs que muestran el número total de productos seleccionados en la lista de la compra y la suma de los mismos por supermercado en función del precio

Figura 66. Gráfica de líneas evolución de los precios de Mercadona con proyección a futuro

Figura 67. Gráfica de líneas evolución de los precios de Mercadona con proyección a futuro

Figura 68. Listado de tablas usadas para implementar en el chatbot

Figura 69. Ejemplo de sinónimos añadidos en el conjunto de datos del chatbot

Figura 70. Preguntas recomendadas al usuario para hacer en el chatbot

Figura 71. Resultado final del chatbot

Figura 72. Logo de la aplicación y del proyecto SUPER-ahorro

Figura 73. Diseño base implementado para la aplicación

Figura 74. Botones añadidos en la aplicación para interactuar entre páginas

Figura 75. Botón con acceso directo al chatbot implementado

Figura 76. Aplicación subida al servicio de Power BI

Figura 77. Esquema de la metodología implementada en el proyecto

Índice de Tablas

Tabla 1. Listado de ciudades y número de habitantes (Wikipedia, 2023)

Tabla 2. Planificación del desarrollo del proyecto por actividades y semanas

Tabla 3. Nombre del campo, descripción, tipo y ejemplo de las variables obtenidas para los datos de Mercadona

Tabla 4. Nombre del campo, descripción, tipo y ejemplo de los campos añadidos a la extracción de Mercadona

Tabla 5. Nombre del campo, descripción, tipo y ejemplo de las variables obtenidas para los datos de Dia

Tabla 6. Nombre del campo, descripción, tipo y ejemplo de los campos añadidos a la extracción de Dia

Tabla 7. Listado de las diferentes categorías de las que se han extraído datos de Mercadona junto con el url y el código identificador de la misma

Tabla 8. Listado de las diferentes categorías de las que se han extraído datos de Dia junto con el url de la misma

Tabla 9. Resumen de los entregables llevados a cabo en el proyecto usando metodología ágil

Tabla 10. Presupuesto total del proyecto desglosado

Capítulo 1. ANTECEDENTES / ESTADO DEL ARTE

1.1 Introducción

Con el incremento de los precios en la cesta de la compra en los últimos años, los españoles analizan con más detenimiento los productos que consumen y la evolución de los precios en el tiempo. Para ello, se pretende ayudar al usuario con la creación de un servicio capaz de comparar el precio de los productos entre diferentes supermercados y así lograr el supermercado más económico, facilitando a su vez la creación de una lista de la compra. También, se pretende aportar valor adicional mediante el uso de un chatbot de asistencia al usuario.

1.2 Estado del arte

Con la segunda mayor subida de los precios en la cesta de la compra en los últimos 35 años, los consumidores dedican más tiempo para contrastar el precio de los productos, incluso comparando entre diferentes supermercados, demostrando un menor arraigo a “su” supermercado de confianza e interponiendo en numerosas ocasiones el coste del producto en su bolsillo.

En los últimos dos años (2022-2023) los precios se han incrementado en un 30,8% llegando a afectar al 90% de los productos (Pelayo, 2023).

Además, se debe tener en cuenta que algunos de los artículos que han tenido un mayor porcentaje de crecimiento en el precio del producto son el azúcar, el aceite, los huevos o la leche, considerados productos básicos.

La siguiente gráfica muestra la variación porcentual interanual que han sufrido los precios de los alimentos en marzo de 2023 y el mismo mes de 2022.

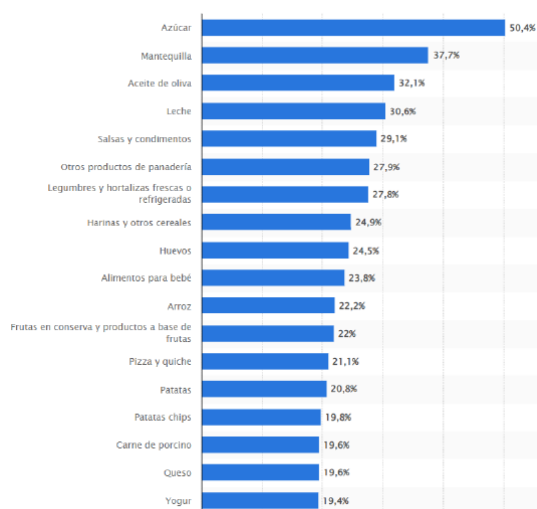


Figura 1. Porcentaje variación del precio en marzo de 2022 y 2023 (Statista, 2023)

Como solución al incremento de los precios, más del 50% de los artículos que las familias españolas consumen son de marca blanca debido a su menor coste en referencia a otro de marca original (Vargas, 2023).

Según la OCU, las principales subidas se produjeron en supermercados como Carrefour (45%), Alcampo (43%), Mercadona (38%) o Día (32%). Además, el estudio se muestra preocupado al comparar dichas subidas con el aproximadamente 6% que subieron los salarios en el mismo período (Agencias, 2024).

Teniendo en cuenta que en el 2022 y, según cálculos realizados por la OCU, los hogares tendrán un sobrecoste de 830€ más en alimentación que el año anterior los consumidores se han visto con la necesidad de cambiar sus hábitos dejando de consumir algunos productos básicos (Los Precios de Alimentos Siguen Al Alza, 2022).

Con la rebaja del IVA que llegó en 2023, las frutas, verduras, el pan, los huevos, las legumbres y el arroz disminuyeron su IVA pasando del 4% al 0% y, la pasta y el aceite bajo de un 10% al 5% se pudo ver una pequeña rebaja en los precios que no fue suficiente y en poco tiempo volvieron a desbordarse los precios (Los Supermercados Más Baratos de 2023, 2023).

De media, el coste anual de la cesta de la compra de una familia tipo se sitúa en 6.109€. Pero existe una diferencia aproximada de 1.056€ entre comprar los productos en la tienda más cara o hacerlo en otra más barata, es decir, dependiendo del supermercado elegido para comprar cada uno de los artículos seleccionados es posible llegar a ahorrarse aproximadamente hasta un 17,5% del presupuesto.

Dependiendo en la ciudad en la que se esté o si se decide hacer la compra por internet se puede llegar a ahorrar desde 430€/año como máximo en el caso de Cuenca hasta 3.928€/año como máximo en el caso de Madrid. La siguiente imagen nos indica las ciudades en las que más se puede ahorrar.

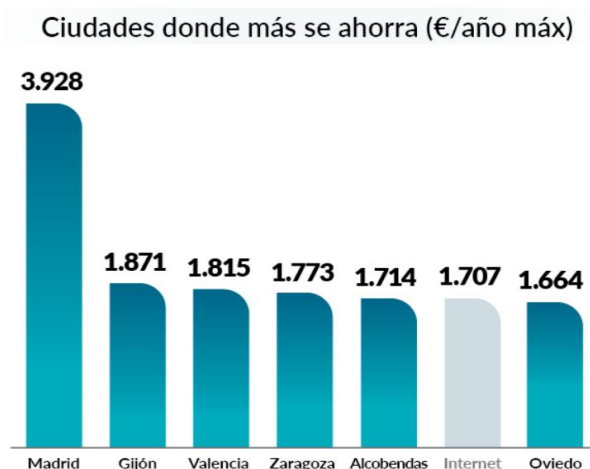


Figura 2. Ranking de las ciudades donde más se ahorra en la compra (Los Supermercados Más Baratos de 2023, 2023)

Como se puede observar Madrid está en la cabeza con bastante diferencia del resto de ciudades y, a su vez, se puede apreciar como aparece la opción de internet dentro de las seis primeras. Por otro lado, cerrando el ranking se encuentran las siguientes ciudades:

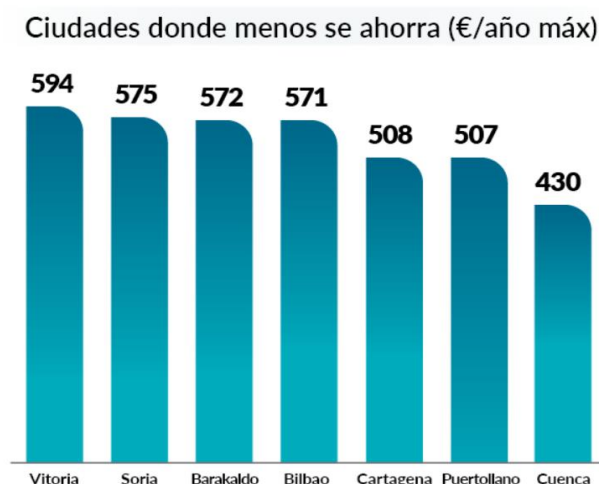


Figura 3. Ranking de las ciudades donde menos se ahorra en la compra (Los Supermercados Más Baratos de 2023, 2023)

Como se puede advertir son ciudades más pequeñas con un menor número de habitantes y, por consiguiente, de supermercados y marcas sobre las que elegir. En la siguiente tabla se puede ver en detalle los habitantes que hay en cada una de las ciudades mostradas en los gráficos anteriores a día 1 de enero de 2023 y, aunque existe alguna excepción cuantos menos habitantes tiene la ciudad menos dinero se ahorra por año (Los Supermercados Más Baratos de 2023, 2023).

Ciudad	Nº de habitantes
Madrid	3.332.035
Gijón	268.313
Valencia	807.693
Zaragoza	682.513
Alcobendas	119.416
Oviedo	217.584
Vitoria	255.886

Soria	40.096
Baracaldo	101.229
Bilbao	346.096
Cartagena	218.050
Puertollano	45.127
Cuenca	53.512

Tabla 1. Listado de ciudades y número de habitantes (Wikipedia, 2023)

Además, El Español informa de que existe un supermercado por cada 1.886 habitantes lo que explica la lógica comentada con anterioridad entre el ahorro y el número de habitantes que hay en cada ciudad (Tobar, 2023).

Tras una encuesta realizada por Data Centric en 2022 y, según publica el periódico 20 minutos, Mercadona es el supermercado favorito en gran parte de las comunidades autónomas a excepción de Euskadi y Melilla. Este dato es curioso ya que no es la cadena con más supermercados en el territorio, pero aún así los españoles la siguen eligiendo. En el siguiente mapa se puede observar visualmente los supermercados elegidos por los españoles para hacer la compra (Minutos, 2022).



Figura 4. Mapa con los supermercados más elegidos por CCAA (Minutos, 2022)

Por otro lado y, si se compara por comunidad autónoma los supermercados con más establecimientos, se denota como el grupo Dia se coloca en primera posición, pero en este caso, el mapa está más dividido. Los resultados obtenidos fueron:



Figura 5. Mapa de los supermercados con más establecimientos por CCAA (Minutos, 2022)

Tanto el grupo Dia como los supermercados Coviran se encuentran dentro del TOP 3 nacional de número de establecimientos porque son las dos grandes empresas que más presencia tienen en municipios inferiores a los 10.000 habitantes (Tobar, 2023).

Otras aplicaciones similares

Existen diferentes aplicaciones análogas encargadas de ayudar a los consumidores a realizar un listado con los productos que necesitan adquirir en su próxima visita al supermercado.

Las más comunes son el asistente de Google o la propia aplicación del asistente virtual de Amazon. En el caso de Google, se puede apreciar como parte negativa que no es capaz de integrarse en otros servicios y es necesario acceder a la web para ver y gestionar la lista.

Bring o Listonic son otras aplicaciones que ayudan al consumidor a realizar la lista de la compra, la última incluso desde un reloj inteligente. Ambas tienen como única funcionalidad agregar los productos indicados por el cliente a una lista en la aplicación (Fernández, 2019).

Este trabajo final de grado, además de dar soporte en la elaboración de la lista de la compra según se vaya requiriendo de un nuevo producto, va a efectuar una comparativa entre diferentes supermercados aconsejando el establecimiento donde comprar cada uno de los productos de la lista a un menor coste. Asimismo, mostrará un análisis con el histórico de precios

de los productos y una predicción a futuro de la evolución de los mismos para que, de una manera visual e intuitiva, se puedan tomar conclusiones útiles.

Otra diferencia notable entre la propuesta para este proyecto y las aplicaciones mencionadas anteriormente es que se implementará un chatbot que mejorará la experiencia del usuario de cara a solucionar dudas al consumidor y ejecutar aquellas actividades para las que el propio chatbot haya sido entrenado.

1.3 Contexto y justificación

¿Existe alguna forma de ahorrar en la cesta de la compra? En la actualidad se puede observar como los precios de los productos en los supermercados siguen al alza provocando mayores dificultades a los españoles para llegar a final de mes. Este incremento ha hecho que se fijen más en los precios, se compare entre diferentes supermercados e incluso se haga la compra en varios de ellos en función del coste.

Con la creación de esta aplicación se pretende ayudar a los clientes de esta a comparar sus precios y determinar qué productos comprar y en qué supermercado. De esta forma, el usuario puede plantear la solución de dos formas diferentes teniendo en cuenta no solo el importe de la compra sino el tiempo o el carburante. La primera opción es elegir un solo supermercado para comprar el total de los productos que se necesitan comparando el precio total de la compra. La segunda opción que se puede plantear es comprar en diferentes establecimientos en función al coste menor del producto individualmente, en esta segunda opción el cliente debe tener en cuenta aspectos secundarios no menos importantes.

Por otro lado, esta aplicación ayuda a estimar el valor de su lista de la compra para priorizar los ahorros y reservarlo mensualmente.

Como valor añadido, se le quiere dotar al usuario con información histórica en la que se puede controlar perfectamente el sentido que llevan los precios de sus productos y detectar si alguno de ellos ha modificado su valor. De esta manera, se pretende ayudar dando a conocer si aquellos productos que son secundarios o no tengo existe la necesidad de ser comprado de continuo han subido o bajado y, por consiguiente, es buen momento hacerse con ellos.

Finalmente, es muy común que se haga una lista de la compra antes de ir al supermercado para que no se olvide ningún artículo y, el último servicio ofrecido ayuda a tratar este punto con la posibilidad de ir añadiendo los productos necesarios.

Alguno de los casos de uso que se pueden ver solucionados con la creación de la aplicación serían los siguientes:

- **Quiero comprar una serie de productos y no sé a qué supermercado ir:** Con esta aplicación, se añaden los productos a la cesta de la compra e inmediatamente saldrá el coste total de los mismos comparándolo con cada uno de los supermercados.

- **Quiero comprar una serie de productos y no sé si comprarlos en el mismo o distinto supermercado:** Otra de las casuísticas puede venir dada porque de los productos que he añadido, unos de ellos son más baratos en el supermercado A y los otros en el B, ¿Los compro todos en el mismo supermercado?, ¿me ahorro lo suficiente si me reparto entre supermercados?. Este producto ayuda a detectar el número de artículos que están más baratos para cada uno de los establecimientos y ser capaces de ver si es rentable o no dividir la compra entre supermercados o hacerla al completo en uno pese a no tener todos los productos al mejor precio teniendo en cuenta el incremento de tiempo o coste que esto supone.
- **Quiero hacer la lista de la compra:** Me gustaría que cuando vaya a comprar no se olvide nada. Para este caso de hacer una lista de la compra existen múltiples opciones, pero ninguna de ellas da el valor añadido de comprar los precios a posterior.
- **Me estoy quedando sin un producto, pero puedo esperar a comprarlo la semana que viene, ¿qué hago?:** Con las gráficas de predicción a futuro de los precios se puede observar si el producto está bajando o subiendo su valor y poder tomar la decisión de comprarlo esta semana o mejor esperar a la siguiente.

Como valor añadido, el producto final pretende dar la mayor información posible sin necesidad de interactuar durante mucho tiempo por la aplicación, es decir, con solo acceder a ella y seleccionar los productos que deseas comprar o analizar se puede contemplar el resto de información que se te pretende facilitar.

A modo de conclusión, los españoles no tienen acceso directo a comparar los productos y obtener información rápida y sencilla. Por ello, parte la idea de crear una aplicación capaz de mostrar las conclusiones que el usuario necesita en el momento y en una sola aplicación sin necesidad de perder el tiempo accediendo a las páginas oficiales de cada uno de los supermercados de manera individual.

1.4 Planteamiento del problema

Tras conocer la necesidad que hay de ayudar a los Españoles en la compra, una tarea común y que se realiza en su mayoría semanalmente, existe la idea de crear una aplicación que, en poco tiempo, sea capaz de dar información valiosa al consumidor para ahorrar en la cesta de la compra.

Con la ayuda de obtener toda la información de diferentes supermercados en una misma aplicación, se puede reducir el tiempo que se dedica a comprar productos y acceder a cada una de las páginas de los supermercados individualmente.

Para cumplir con el objetivo de ayudar al usuario a resolver los diferentes casos de uso que se han planteado, se ha creado una aplicación compuesta por diferentes páginas. La primera de ellas, está destinada a añadir los productos que se pretenden comparar, comprar o analizar y que servirá a su vez como listado de la compra.

También, se ofrece otra vista con los precios para cada uno de los productos en los diferentes supermercados dando información a su vez del supermercado en el que está más barato o el total del coste del conjunto total de la cesta de la compra. También, se pretende ayudar al usuario en la toma de decisiones mostrando una tercera página con una serie de visuales que informan de la evolución de los precios en el tiempo, el coste total, su evolución con respecto a la semana pasada, la cantidad de productos totales que se tiene en la cesta y el número de los mismos que son más baratos en cada uno de ellos y, por último, una estimación a futuro de la evolución del precio en ambos supermercados.

Con toda esta información se pretende aportar valor ayudando a facilitar y disminuir el coste de la cesta de la compra a la vez que ahorrar tiempo.

Capítulo 2. OBJETIVOS

2.1 Objetivos generales

El objetivo principal es crear un servicio final capaz de comparar el precio de los productos en diferentes supermercados, analizar su evolución en el tiempo y su predicción a futuro almacenando un histórico de los datos recopilados, empleando la metodología scrum para su desarrollo y cuya finalidad principal es hacer que los clientes ahorren en su lista de la compra en supermercados.

2.2 Objetivos específicos

Siguiendo las directrices mencionadas en el objetivo general, los objetivos específicos localizados son los siguientes:

- Crear un proceso automático de obtención de los datos.
- Tratar, limpiar y realizar un análisis exploratorio de los datos recopilados.
- Implementar un modelo de datos capaz de cumplimentar todos los requisitos teniendo en cuenta las relaciones entre diferentes tablas y los identificadores únicos de las mismas.
- Ser capaz de obtener conclusiones y verificar la veracidad de los datos mediante el uso de herramientas de visualización.
- Implementar un producto final para los usuarios que represente todo el trabajo realizado.
- Emplear metodologías ágiles para el desarrollo del trabajo al completo.
- Aplicar todos los conceptos adquiridos a lo largo de la carrera en el desarrollo del proyecto.

2.3 Beneficios del proyecto

Teniendo en cuenta las subidas de precios de los últimos meses en la lista de la compra, con la creación de esta aplicación, los usuarios de esta serán capaces de, en primer lugar, ahorrar en la cesta comparando el precio de diferentes supermercados y adquiriendo cada producto al mejor precio y, secundariamente pero no menos importante, facilitar una lista de todos los productos que necesitamos compra para que, de este modo, los usuarios no se olviden de dichos artículos cuando acudan al supermercado.

Capítulo 3. DESARROLLO DEL PROYECTO

3.1 Planificación del proyecto

Durante el desarrollo del proyecto, se ha seguido una metodología scrum compuesta de ocho sprints diferentes y cuya duración es de tres semanas cada uno. Las diferentes actividades que se han llevado a cabo durante la implementación del producto se pueden ver representadas en los siguientes bloques:

Bloque 1: Actividades previas al desarrollo del proyecto

- Contextualizar el problema que se quiere resolver y estudiar las posibles soluciones que se pueden aplicar teniendo en cuenta los proyectos similares ya existentes.
- Crear un anteproyecto indicando los antecedentes, problemas que se pretenden resolver, metodología prevista y su planificación.
- Definir conceptos necesarios previos a la implementación del desarrollo como son el alcance, los requisitos funcionales y no funcionales, un logotipo o el nombre final del producto.
- Preparación de las herramientas necesarias previstas para implementar el proyecto como pueden ser la descarga de nuevas librerías o softwares y creación de cuentas gratuitas para algunos softwares.
- Formarse mediante el uso de vídeos, cursos o artículos sobre las tareas que se van a llevar a cabo en el desarrollo del producto.

Bloque 2: Desarrollo del proyecto

1. Obtención de los datos:
 - Estudiar las diferentes opciones que hay de obtención de los datos y elegir la correcta en base a su facilidad para obtener los datos, las variables que conforman el conjunto o el formato final del conjunto entre otras.
 - Determinar el formato correcto de obtención de los datos (extensión del fichero, campos, categorías...), para que el resto de los procesos sean más sencillos de implementar.
 - Implementar un proceso de extracción de la información para el primer supermercado elegido (Mercadona) teniendo en cuenta que este desarrollo tiene que ejecutarse semanalmente y debe funcionar de la misma forma. Este proceso dará como resultado final un fichero con todo el detalle necesario (productos, categoría y precios principalmente).
 - Realizar un proceso similar al anterior de extracción de la información para el segundo supermercado elegido (Dia) teniendo en cuenta que este desarrollo tiene que ejecutarse semanalmente y debe funcionar de la misma forma. Además, se debe intentar que el resultado final tenga un formato similar al obtenido para el primer supermercado.

2. Estudio y análisis de la calidad del dato:
 - Limpiar el conjunto de datos obtenido con los datos de ambos supermercados hasta obtener un mismo fichero de igual formato para ambos supermercados.
 - Crear un proceso de unificación de los ficheros obtenidos semanalmente para un mismo supermercado.
 - Elaborar un fichero que relacione los artículos de ambos supermercados para posteriormente poder comparar su precio.
 - Analizar y comprobar la calidad del dato con la creación de tablas, visuales y medidas en Power BI.
3. Chatbot:
 - Realizar un listado de posibles dudas o gráficas que el usuario pueda demandar mediante el uso del chatbot.
 - Entrenar el chatbot con posibles sinónimos y preguntas que el usuario pueda realizar.

Bloque 3: Creación del producto final para el usuario

- Creación de un prototipo con las ideas que se implementarán posteriormente para el producto final y que englobe todas las pestañas que va a tener la aplicación, el estilo de las mismas y su interacción.
- Cargar los diferentes conjuntos de datos obtenidos, unificados y limpiados previamente. Incluyendo también la tabla que relaciona los IDs de ambos supermercados.
- Crear un modelo de datos que relacione las diferentes tablas de hechos y dimensiones necesarias. Para ello, será necesario la generación de nuevas tablas como pueden ser la tabla calendario o tablas intermedias en las relaciones.
- Elaboración de gráficas, KPIs, filtros o tablas necesarias para su representación en la aplicación y que se han determinado previamente en el mock-up.
- Implementar la interacción entre los diferentes gráficos (teniendo en cuenta los productos señalados) para las diferentes páginas de la aplicación.
- Implementar una interfaz de usuario clara, con colores acordes al logotipo y marca creados.
- Desplegar o publicar el producto final para ser consumida por todos los públicos.

Bloque 4: Pruebas

- Se realizan pruebas del proceso de obtención de los datos cada semana verificando que la información que se obtiene es la que se muestra en la fuente origen.
- Se realizan las pruebas pertinentes para comprobar que la aplicación final funciona correctamente, es decir, se verificará que el dato es correcto, que la iteración entre las diferentes páginas es la adecuada o que los filtros se aplican bien en cada uno de los visuales. De esta forma, se pretende garantizar una buena experiencia de usuario.

A continuación, se verá representado de forma esquemática en una tabla la planificación llevada a cabo por semanas, teniendo en cuenta los bloques y actividades mencionados anteriormente.

El tiempo empleado para el desarrollo total del proyecto ha sido de un total de 450 horas (cuatro meses y medio).

	Enero	Febrero	Marzo	Abril	Mayo	Junio
ACTIVIDADES PREVIAS AL DESARROLLO DEL PROYECTO						
Redactar la idea del proyecto						
Contextualizar la idea						
Antecedentes y estado del arte						
Definir el alcance del proyecto						
Definir los requisitos funcionales						
Definir los requisitos no funcionales						
Definir el objetivo principal						
Definir los objetivos secundarios						
Crear un nombre y logotipo de marca						
Planificar el desarrollo del proyecto						
Instalar y comprobar el funcionamiento del entorno						
Formación en herramientas nuevas						
DESARROLLO DEL PROYECTO						
<i>Obtención de los datos</i>						
Localizar los orígenes de los datos						
Determinar un formato para el fichero						
Creación de un proceso para el primer supermercado						
Creación de un proceso para el segundo supermercado						
<i>Estudio y análisis de la calidad del dato</i>						
Limpiar el conjunto de datos						
Implementar un proceso de unión de la información						
Crear fichero lds						
Analizar la calidad del dato						

Chatbot						
Creación de un listado con posibles preguntas						
Entrenar el chatbot						
CREACIÓN DEL PRODUCTO FINAL PARA EL USUARIO						
Creación del prototipo						
Cargar los ficheros de datos						
Implementar el modelo de datos						
Crear las métricas necesarias						
Crear los KPIs						
Crear las gráficas y tablas						
Implementar la interacción entre páginas						
Implementar la interfaz						
Publicación del producto						
PRUEBAS						
Pruebas del proceso de obtención de los datos						
Pruebas de la aplicación						
DOCUMENTACIÓN						
Creación de la memoria del trabajo						

Tabla 2. Planificación del desarrollo del proyecto por actividades y semanas

3.2 Descripción de la solución, metodologías y herramientas empleadas

3.2.1 Descripción de la solución

El desarrollo del proyecto ha estado dividido en diferentes bloques. El primero de ellos está relacionado con la obtención de los diferentes conjuntos de datos, y cuya finalidad es obtener diferentes ficheros con los precios de cada uno de los supermercados a comparar.

El segundo bloque engloba todos los pasos llevados a cabo para la limpieza del dato tanto del fichero obtenido con los datos de precios de Mercadona, como la información recogida del supermercado Dia.

El tercer apartado engloba todo el proceso de construcción del modelo de datos, creando tablas de dimensiones y relacionándose estas últimas con diferentes tablas de hechos. Además, ha sido necesaria la creación de nuevas tablas para relacionar los productos de los dos supermercados sobre los que hemos cogido la información.

El cuarto bloque se encarga de comprobar la calidad del dato. Una vez obtenido, limpiado y estructurado el dato, este bloque se encarga de estudiar y analizar la veracidad del dato creando métricas y comparando la información.

Finalmente, el último apartado se compone del desarrollo de un servicio final sobre el que los usuarios puedan emplear la información y comparar el precio de diferentes productos en los supermercados para así tomar una decisión.

Dentro de cada bloque se han realizado diversas tareas enfocadas al cumplimiento de los objetivos propuestos. A continuación, se explicará más en detalle cada uno de los bloques mencionados:

3.2.1.1 Obtención de los datos

Para este desarrollo ha sido necesario partir de dos orígenes diferentes. Ambas son las páginas oficiales de cada uno de los supermercados. Para los dos casos se ha procedido a hacer web scraping de los datos usando ParseHub una herramienta dedicada a extraer información de cualquier web y que ofrece parte de sus apartados gratuitos (Capterra, 2022).

En ambos casos, se ha intentado obtener la mayor información posible de los datos para poder aportar más valor en la aplicación y, también, se ha procurado obtener un formato similar en ambos casos.

En el caso de Mercadona, el fichero resultante está en formato JSON y, tras el proceso de extracción de datos podemos obtener información de las siguientes variables:

Campo	Descripción	Tipo	Ejemplo
Categoría	Grupo por el que se engloban el resto de los productos. Puede tomar valores como aceite, leche o azúcar y, posteriormente, en el nombre del producto entrar en un mayor detalle	texto	Gaseosa
Nombre	Nombre completo del producto	texto	Gaseosa Hacendado grande
Tamaño	Unidades en la que se mide el producto, pueden ser kilos, litros, unidades...	texto	Botella 1,5 L

Precio inicio	Valor por el que se vende el producto en cuestión y la cantidad/tamaño indicado	numérico	0,50 €
Precio oferta	En caso de que el producto tenga alguna oferta, este campo vendrá relleno con el precio del producto	numérico	0,45 €
Unidad	Cantidad por la que se notifica el precio del producto. Puede tomar el valor de unidades, pack...	texto	ud

Tabla 3. Nombre del campo, descripción, tipo y ejemplo de las variables obtenidas para los datos de Mercadona

Estos datos vienen dados en dos niveles diferentes, el primero de ellos nos indica todas las categorías que hay y, dentro de cada una de ellas, se puede ver el resto de campos indicados anteriormente:

```
> "Agua_sin_gas": [ ...  
  ],  
> "Agua_con_gas": [ ...  
  ],  
> "Gaseosa": [ ...  
  ],  
> "cocacola_clasica": [ ...  
  ],  
> "Cocacola_zero": [ ...  
  ],  
> "Cocacola_sin_cafeina": [ ...  
  ],  
  "Fanta_de_limon": [
```

Figura 6. Ejemplo del primer nivel del conjunto de datos de Mercadona

```
"Agua_sin_gas":[
{
"name":"Agua mineral grande Fuente Dehesa\nGarrafa 6 L\n1,26 € /ud."
},
{
"name":"Agua mineral grande Fuente Dehesa\n6 botellas x 1,5 L\n2,34 € /pack"
},
{
"name":"Agua mineral grande Fuente Dehesa\nBotella 1,5 L\n0,39 € /ud."
},
{
"name":"Agua mineral mediana Fuente Dehesa\n6 botellas x 1 L\n3,00 € /pack"
},
{
"name":"Agua mineral mediana Fuente Dehesa\nBotella 1 L\n0,50 € /ud."
},
]
```

Figura 7. Ejemplo del segundo nivel del conjunto de datos de Mercadona

Adicionalmente a la información que se extrae directamente de la página web, se añaden los siguientes campos:

Campo	Descripción	Tipo	Ejemplo
Supermercado	Indica el nombre del establecimiento sobre el que se han recogido los datos. En este caso, todas las filas tienen el mismo valor	texto	Mercadona
Fecha	Muestra la fecha del día en el que se recopiló la información. En este caso en concreto como el proceso se ha realizado todos los lunes, las fechas coinciden con dicho día de la semana	Fecha	01/04/2024 0:00

Tabla 4. Nombre del campo, descripción, tipo y ejemplo de los campos añadidos a la extracción de Mercadona

El fichero en este caso tiene la siguiente apariencia:

1	Categoría	Nombre	Cantidad	Precio Inicio[€]	Precio Of[€]	Cantidad	Supermercado	Fecha
2	Aceite_de_girasol_semillas_y_maiz	Aceite de coco virgen Hacendado	Bote 450 ml	4.4		ud.	Mercadona	01/04/2024 0:00
3	Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0-2º Hacendado	Botella 1 L	1.45		ud.	Mercadona	01/04/2024 0:00
4	Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0-2º Hacendado	Garrafa 5 L	6.55	6.75	ud.	Mercadona	01/04/2024 0:00
5	Aceite_de_oliva	Aceite de oliva 0-4º Hacendado	Botella 1 L	9.25		ud.	Mercadona	01/04/2024 0:00
6	Aceite_de_oliva	Aceite de oliva 1º Hacendado	Botella 1 L	9.25		ud.	Mercadona	01/04/2024 0:00
7	Aceite_de_oliva	Aceite de oliva intenso Hacendado	Garrafa 3 L	27.6		ud.	Mercadona	01/04/2024 0:00
8	Aceite_de_oliva	Aceite de oliva suave Hacendado	Garrafa 3 L	27.6		ud.	Mercadona	01/04/2024 0:00
9	Aceite_de_oliva	Aceite de oliva virgen extra Hacendado	Botella 1 L	9.9		ud.	Mercadona	01/04/2024 0:00
10	Aceite_de_oliva	Aceite de oliva virgen extra Hacendado	Garrafa 3 L	29.55		ud.	Mercadona	01/04/2024 0:00
11	Aceite_de_oliva	Aceite de oliva virgen extra Hacendado	Spray 200 ml	3.3		ud.	Mercadona	01/04/2024 0:00
12	Aceite_de_oliva	Aceite de oliva virgen extra Hacendado Gran Selección	Botella 750 ml	8.1		ud.	Mercadona	01/04/2024 0:00
13	Aceite_de_oliva	Aceite de oliva virgen extra Picual Casa Juncal	Botella 500 ml	7.85		ud.	Mercadona	01/04/2024 0:00
14	Aceite_de_oliva	Aceite de oliva virgen Hacendado	Botella 1 L	9.55		ud.	Mercadona	01/04/2024 0:00
15	Aceite_de_oliva	Aceite de oliva virgen Hacendado	Garrafa 3 L	28.5		ud.	Mercadona	01/04/2024 0:00
16	Agua_con_gas	Agua de soda con sifón La Casa	Botella 1,5 L	1.6		ud.	Mercadona	01/04/2024 0:00
17	Agua_con_gas	Agua mineral con gas grande Aguadoy	6 botellas x 1,5 L	2.46		pack	Mercadona	01/04/2024 0:00
18	Agua_con_gas	Agua mineral con gas grande Aguadoy	Botella 1,5 L	0.41		ud.	Mercadona	01/04/2024 0:00
19	Agua_con_gas	Agua mineral con gas grande Vichy Catalan	Botella 1 L	1.49		ud.	Mercadona	01/04/2024 0:00
20	Agua_con_gas	Agua mineral con gas pequeña Aguadoy	6 botellas x 500 ml	1.62		pack	Mercadona	01/04/2024 0:00
21	Agua_con_gas	Agua mineral con gas pequeña Aguadoy	Botella 500 ml	0.27		ud.	Mercadona	01/04/2024 0:00
22	Agua_con_gas	Agua mineral con gas pequeña Vichy Catalan	6 botellas x 300 ml	5.89		pack	Mercadona	01/04/2024 0:00

Figura 8. Ejemplo del fichero de Mercadona final resultante con los campos extraídos y los añadidos posteriormente

En el caso del supermercado Dia, aunque los campos sean diferentes, se ha pretendido recoger como mínimo la misma información que para Mercadona, pero se le ha añadido algún campo adicional que, en este caso, sí ofrecía la página web.

El fichero de salida, al igual que en el caso anterior, es un JSON compuesto en primera instancia por los siguientes campos:

Campo	Descripción	Tipo	Ejemplo
Categoría	Grupo por el que se engloban el resto de los productos. Puede tomar valores como aceite, leche o azúcar y, posteriormente, en el nombre del producto entrar en un mayor detalle	texto	Aceites
Nombre del producto	Nombre completo del producto	texto	Aceite de oliva virgen extra La Almazara del Olivar garrafa 3 l
url	Link para acceder a ver el producto en la página oficial de Dia	texto	Aceite de oliva virgen extra La Almazara del Olivar de Dia garrafa 3 l - Supermercados DIA

Precio inicio	Valor por el que se vende el producto en cuestión y la cantidad/tamaño indicado	numérico	29,55€
Precio por unidad	Valor del producto por unidad, es decir por litro, kilo...	numérico	9,85 €/LITRO

Tabla 5. Nombre del campo, descripción, tipo y ejemplo de las variables obtenidas para los datos de Dia

El formato del fichero está compuesto por tres niveles diferentes:

- **Nivel categoría:** En él se puede observar únicamente las distintas categorías en las que se divide el conjunto de datos.

```
> "Aceites":[ ...
],
> "Vinagre_y_alineos":[ ...
],
> "Sal_y_especies":[ ...
],
> "Agua":[ ...
],
> "Cocacola":[ ...
],
> "Fanta_de_naranja":[ ...
],
> "Fanta_de_limon":[ ...
],
> "Arroz":[ ...
```

Figura 9. Ejemplo del primer nivel del conjunto de datos de Dia

- **Nivel producto:** Dentro de cada categoría existen diversos productos con su nombre, url y precio.

```
"Aceites":[
{
"name":"Aceite de oliva virgen extra La Almazara del Olivar garrafa 3 l",
"url":"https://www.dia.es/aceites-salsas-y-especias/aceites/p/266129",
"Precio":[
{
"name":"29,55 €"
},
{
"name":"(9,85 €/LITRO)"
}
]
},
]
```

Figura 10. Ejemplo del Segundo nivel del conjunto de datos de Dia

- **Nivel precio:** Dentro del apartado de precio, existen dos valores diferentes. El primero de ellos hace referencia al precio del conjunto entero mientras que el segundo hace referencia al precio si se comprara por unidad, es decir, un litro o un kilogramo.

```
"Precio":[
{
  "name":"29,55 €"
},
{
  "name":"(9,85 €/LITRO)"
}
]
```

Figura 11. Ejemplo del tercer nivel del conjunto de datos de Dia

De la misma forma que en el apartado anterior, se le añadieron columnas adicionales para facilitar nueva información útil al conjunto de datos. Estos valores son los siguientes:

Campo	Descripción	Tipo	Ejemplo
Supermercado	Indica el nombre del establecimiento sobre el que se han recogido los datos. En este caso, todas las filas tienen el mismo valor	texto	Dia
Fecha	Muestra la fecha del día en el que se recopiló la información. En este caso en concreto como el proceso se ha realizado todos los lunes, las fechas coinciden con esos días	Fecha	01/04/2024 0:00

Tabla 6. Nombre del campo, descripción, tipo y ejemplo de los campos añadidos a la extracción de Dia

El fichero resultante, por tanto, es el siguiente:

Categoría	Nombre	url	Precio Inicio(€)	Precio/unidad (€)	Unidad	Supermercado	Fecha
Aceites	Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/107364	46 €	9,20 €	LITRO	Dia	01/04/2024
Aceites	Aceite de oliva suave Carbonell botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100	13,89 €	13,89 €	LITRO	Dia	01/04/2024
Aceites	Aceite de oliva suave La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/104	9,25 €	9,25 €	LITRO	Dia	01/04/2024
Aceites	Aceite de oliva suave La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/49895	46 €	9,20 €	LITRO	Dia	01/04/2024
Aceites	Aceite de oliva suave La Española botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/1348	13,45 €	9,42 €	LITRO	Dia	01/04/2024
Aceites	Aceite de oliva virgen extra Carbonell botella 750 ml	https://www.dia.es/aceites-salsas-y-especias/aceites/p/12013	10,89 €	14,52 €	LITRO	Dia	01/04/2024
Aceites	Aceite de oliva virgen extra La Almazara del Olivar garrafa 3 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/266129	29,55 €	9,85 €	LITRO	Dia	01/04/2024
Aceites	Aceite de oliva virgen La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/216284	9,55 €	9,55 €	LITRO	Dia	01/04/2024
Aceites	Aceite de orujo de oliva La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/12668	5,24 €	5,24 €	LITRO	Dia	01/04/2024
Aceites	Aceite refinado de girasol Diadol botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/101	1,49 €	1,49 €	LITRO	Dia	01/04/2024
Agua	Agua mineral con gas Dia botella 1,5 l	https://www.dia.es/agua-refrescos-y-zumos/agua/p/99285	0,41 €	0,27 €	LITRO	Dia	01/04/2024
Agua	Agua mineral natural Aquarel botella 1,5 l	https://www.dia.es/agua-refrescos-y-zumos/agua/p/38301	0,59 €	0,39 €	LITRO	Dia	01/04/2024
Agua	Agua mineral natural Bezoya botella 6 x 1,5 l	https://www.dia.es/agua-refrescos-y-zumos/agua/p/4430296	4,44 €	0,49 €	LITRO	Dia	01/04/2024
Agua	Agua mineral natural Dia botella 1,5 l	https://www.dia.es/agua-refrescos-y-zumos/agua/p/26995	0,25 €	0,17 €	LITRO	Dia	01/04/2024
Agua	Agua mineral natural Dia botella 33 cl	https://www.dia.es/agua-refrescos-y-zumos/agua/p/30967	0,34 €	0,73 €	LITRO	Dia	01/04/2024
Agua	Agua mineral natural Dia botella 50 cl	https://www.dia.es/agua-refrescos-y-zumos/agua/p/37776	0,22 €	0,44 €	LITRO	Dia	01/04/2024
Agua	Agua mineral natural Dia garrafa 5 l	https://www.dia.es/agua-refrescos-y-zumos/agua/p/26996	0,65 €	0,13 €	LITRO	Dia	01/04/2024
Agua	Agua mineral natural Font Vella botella 1,5 l	https://www.dia.es/agua-refrescos-y-zumos/agua/p/85	0,73 €	0,49 €	LITRO	Dia	01/04/2024
Agua	Agua mineral natural Font Vella garrafa 2,5 l	https://www.dia.es/agua-refrescos-y-zumos/agua/p/120320	2,29 €	0,37 €	LITRO	Dia	01/04/2024
Agua	Agua mineral natural Solán de Cabras botella 1,5 l	https://www.dia.es/agua-refrescos-y-zumos/agua/p/36684	0,89 €	0,59 €	LITRO	Dia	01/04/2024
Ajos_cebollas_y_puerros	Ajo morado malla 400 g	https://www.dia.es/verduras/ajos-cebollas-y-puerros/p/171396	2,19 €	5,48 €	KILO	Dia	01/04/2024
Ajos_cebollas_y_puerros	Ajos malla 250 g	https://www.dia.es/verduras/ajos-cebollas-y-puerros/p/17653	1,59 €	6,36 €	KILO	Dia	01/04/2024
Ajos_cebollas_y_puerros	Cebolla dulce granel 1 kg aprox.	https://www.dia.es/verduras/ajos-cebollas-y-puerros/p/45862	2,49 €	2,49 €	KILO	Dia	01/04/2024
Ajos_cebollas_y_puerros	Cebolla dulce malla 1 kg	https://www.dia.es/verduras/ajos-cebollas-y-puerros/p/132920	2,49 €	2,49 €	KILO	Dia	01/04/2024

Figura 12. Ejemplo del fichero de Dia final resultante con los campos extraídos y los añadidos posteriormente

ParseHub:

En ambos casos, el proceso de extracción del conjunto de datos se ha hecho usando la herramienta ParseHub y se ha creado un pipeline denominado main que llama al resto de procesos, pero aunque el proceso es similar, tiene algunas discrepancias.

- **Extracción datos Mercadona:** Los pasos llevados a cabo para el desarrollo del proceso para obtener un único fichero JSON son los siguientes:
 - o Generación de un pipeline denominado ‘main’ que engloba el resto de los procesos. Este desarrollo se compone de la misma función repetida tantas veces como categorías de producto se extraen. Dicha función lo único que hace es buscar, ir y ejecutar el pipeline indicado.



Figura 13. Imagen de las funciones/pasos a los que llama el pipeline main para obtener el conjunto de datos de Mercadona

- Creación de un proceso para cada categoría de productos, esto es diferente para cada categoría ya que depende de los subapartados de los que esté compuesto cada grupo. El siguiente ejemplo engloba el desarrollo al completo de extracción de los productos reconocidos dentro del grupo 'Aceite', este consta de 4 subapartados (Aceite de oliva, aceite de girasol, vinagre y otros aledaños y sal y bicarbonato). Para cada uno de ellos se pretende recoger el máximo de información posible.



Figura 14. Imagen de ejemplo que muestra cómo se seleccionan en ParseHub los datos que se quieren extraer de Mercadona

Adicionalmente, y, para personalizar y obtener un resultado, se añade en la función el nombre de cada subapartado. De este modo, se tiene configurado al completo el pipeline.



Figura 15. Imagen del proceso que se ejecuta para cada una de las categorías a las que llama el fichero main para extraer los datos de Mercadona

Para el resto de las categorías el proceso es similar, la única diferencia entre cada uno de los desarrollos viene dada por el nombre que se debe indicar en la función `relative` y el link vinculado a cada pipeline que, en este caso, varía únicamente en el ID final.

- Antes de procesar el fichero, es necesario configurar los urls que se deben cargar en cada categoría. Para ello, se observó cómo el link que accede a la página oficial de Mercadona tiene una parte igual y, los últimos dígitos son los que varían. Estos valores se introducen en cada uno de los pipelines.

Categoría	URL	Valor
Aceite, vinagre y sal	https://tienda.mercadona.es/categories/112	112
Agua	https://tienda.mercadona.es/categories/156	156
Cocacola	https://tienda.mercadona.es/categories/158	158
Fanta naranja/limon	https://tienda.mercadona.es/categories/159	159
Arroz	https://tienda.mercadona.es/categories/118	118
Legumbres	https://tienda.mercadona.es/categories/121	121
Pasta y fideos	https://tienda.mercadona.es/categories/120	120
Azucar	https://tienda.mercadona.es/categories/89	89
Cereales	https://tienda.mercadona.es/categories/78	78
Galletas	https://tienda.mercadona.es/categories/80	80
Fruta	https://tienda.mercadona.es/categories/27	27
Verdura	https://tienda.mercadona.es/categories/29	29
Huevos	https://tienda.mercadona.es/categories/77	77
Leche	https://tienda.mercadona.es/categories/72	72

Tabla 7. Listado de las diferentes categorías de las que se han extraído datos de Mercadona junto con el url y el código identificador de la misma

- El fichero de salida tenía el siguiente formato:

```
CSV/Excel  JSON  CSV/Excel Wide (beta)
{
  "Aceite_de_oliva": [
    {
      "name": "Aceite de oliva 0,4° Hacendado\nBotella 1 L\n8,95 € /ud."
    },
    {
      "name": "Aceite de oliva virgen extra Hacendado\nGarrafa 3 L\n29,55 € /ud."
    },
    {
      "name": "Aceite de oliva virgen extra Hacendado\nBotella 1 L\n9,90 € /ud."
    },
  ],
}
```

Figura 16. Ejemplo del formato e información que se recoge sobre los datos de Mercadona

- **Extracción datos Dia:** Del mismo modo que en el apartado anterior, el pipeline main llama al resto de procesos que se han creado, uno por cada categoría. A diferencia que, con Mercadona, el formato de la página web está dividido únicamente por categorías, es decir, tras entrar en cada una de ellas no se localizan subapartados



Figura 17. Imagen de las funciones/pasos a los que llama el pipeline main para obtener el conjunto de datos de Dia

Dentro de cada template se llama a todos los productos que hay y se recoge el máximo de información posible. En este caso, el formato de extracción varía ya que está dividido en más pasos y se obtiene también el url del producto.

En primer lugar, se recoge el link relacionándolo con la imagen ya que esta contiene dicha información.



Figura 18. Imagen de ejemplo que muestra cómo se seleccionan el primer nivel en ParseHub los datos que se quieren extraer de Dia

Secundariamente, se recoge el nombre del producto:



Figura 19. Imagen de ejemplo que muestra cómo se seleccionan el segundo nivel en ParseHub los datos que se quieren extraer de Dia

Finalmente, se recogen los precios para mostrarlos en el JSON final como otro apartado compuesto por el precio de venta y el coste por unidad.



Figura 20. Imagen de ejemplo que muestra cómo se seleccionan el tercer nivel en ParseHub los datos que se quieren extraer de Dia

Este proceso es similar para todos los templates a diferencia de cambiar el nombre del 'select' y los urls para cada una de las categorías. Los links en este caso varían no solo en el número sino que aparece también el nombre de la categoría. Los links para cada categoría son:

Categoría	URL
Aceite	https://www.dia.es/aceites-salsas-y-especias/aceites/c/L2046
Vinagre	https://www.dia.es/aceites-salsas-y-especias/vinagres-y-alinos/c/L2047
Sal	https://www.dia.es/aceites-salsas-y-especias/sal-y-especias/c/L2048
Agua	https://www.dia.es/agua-refrescos-y-zumos/agua/c/L2107
Cocacola	https://www.dia.es/agua-refrescos-y-zumos/cola/c/L2108
Fanta Naranja	https://www.dia.es/agua-refrescos-y-zumos/naranja/c/L2212
Fanta Limón	https://www.dia.es/agua-refrescos-y-zumos/limon-lima-limon/c/L2109
Arroz	https://www.dia.es/arroz-pastas-y-legumbres/arroz/c/L2042
Pastas y fideos	https://www.dia.es/arroz-pastas-y-legumbres/pastas/c/L2044

Garbanzos	https://www.dia.es/arroz-pastas-y-legumbres/garbanzos/c/L2191
Alubias	https://www.dia.es/arroz-pastas-y-legumbres/alubias/c/L2178
Lentejas	https://www.dia.es/arroz-pastas-y-legumbres/lentejas/c/L2193
Azucar	https://www.dia.es/azucar-chocolates-y-caramelos/azucar-y-edulcorantes/c/L2060
Cereales	https://www.dia.es/galletas-bollos-y-cereales/cereales/c/L2068
Galletas	https://www.dia.es/galletas-bollos-y-cereales/galletas/c/L2065
Manzana	https://www.dia.es/frutas/manzanas/c/L2032
Platanos	https://www.dia.es/frutas/platanos/c/L2033
Peras	https://www.dia.es/frutas/peras/c/L2034
Naranjas y Mandarinas	https://www.dia.es/frutas/naranjas-y-mandarinas/c/L2196
Uvas	https://www.dia.es/frutas/uvas/c/L2035
Limonos y pomelos	https://www.dia.es/frutas/limones-y-pomelos/c/L2037
Frutas del bosque	https://www.dia.es/frutas/frutas-del-bosque/c/L2038
Frutas tropicales	https://www.dia.es/frutas/frutas-tropicales/c/L2039
Ajos, Cebollas y puerros	https://www.dia.es/verduras/ajos-cebollas-y-puerros/c/L2022
Tomates, pimientos y pepinos	https://www.dia.es/verduras/tomates-pimientos-y-pepinos/c/L2023
Calabacín, calabaza y berenjena	https://www.dia.es/verduras/calabacin-calabaza-y-berenjena/c/L2181
Judías, brocoli y coliflores	https://www.dia.es/verduras/judias-brocolis-y-coliflores/c/L2024
Lechuga, escarolas y endivias	https://www.dia.es/verduras/lechuga-escarolas-y-endivias/c/L2027
Patatas y zanahorias	https://www.dia.es/verduras/patatas-y-zanahorias/c/L2028
Setas y champiñones	https://www.dia.es/verduras/setas-y-champinones/c/L2029
Otras verduras	https://www.dia.es/verduras/otras-verduras/c/L2031
Huevos	https://www.dia.es/leche-huevos-y-

	mantequilla/huevos/c/L2055
Leche	https://www.dia.es/leche-huevos-y-mantequilla/leche/c/L2051

Tabla 8. Listado de las diferentes categorías de las que se han extraído datos de Dia junto con el url de la misma

3.2.1.2 Limpieza del dato

Dentro de este apartado se debe diferenciar el proceso que se ha llevado a cabo para cada uno de los conjuntos de datos ya que tanto el formato como los campos varían.

- **Limpieza del conjunto de datos Mercadona:** En primera instancia y, tras cargar el conjunto de datos en Power BI se podía observar como era necesaria la implementación de cambios. Partiendo de la siguiente información:

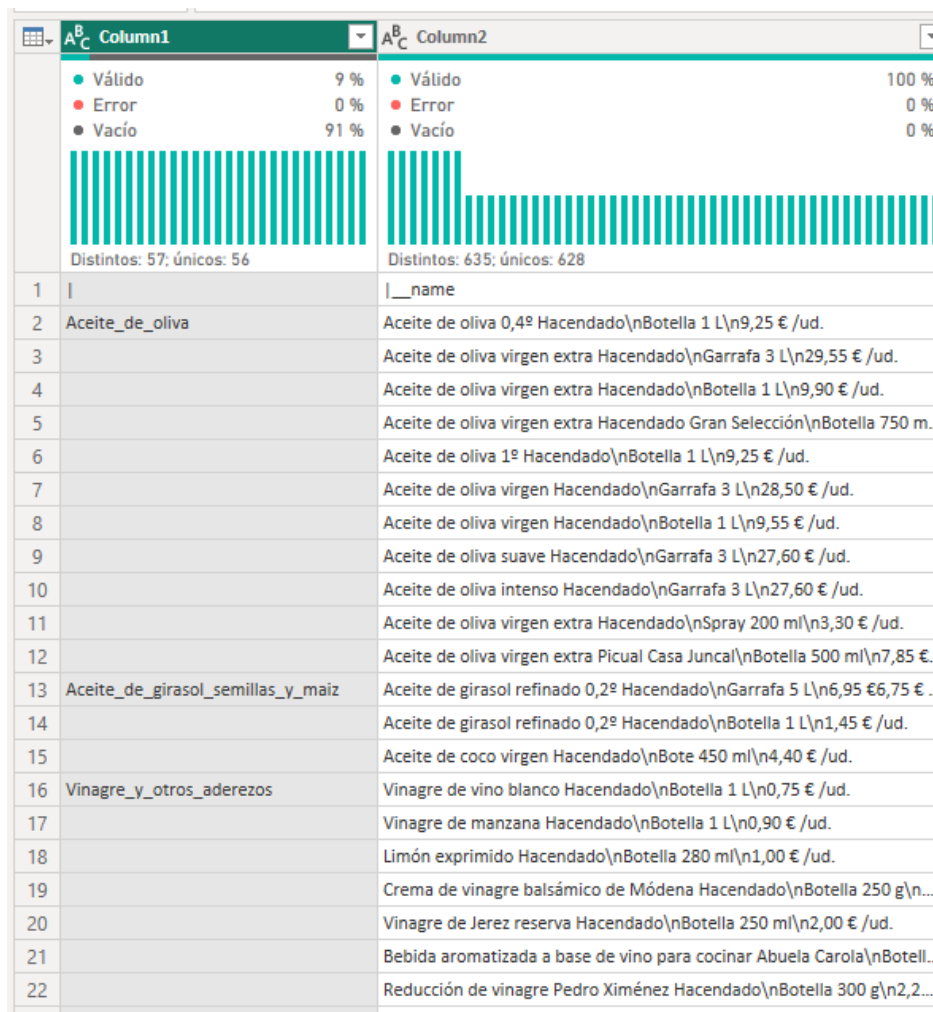


Figura 21. Ejemplo del resultado obtenido tras cargar el fichero con los datos de Mercadona en Power BI

Los pasos llevados a cabo para su limpieza fueron:

- **Encabezados promovidos:** Se indicó la primera fila como cabecera del conjunto de datos.

	A ^B C I	A ^B C I__name
1	Aceite_de_oliva	Aceite de oliva 0,4º Hacendado\nBotella 1 L\n9,25 € /ud.
2		Aceite de oliva virgen extra Hacendado\nGarrafa 3 L\n29,55 € /ud.
3		Aceite de oliva virgen extra Hacendado\nBotella 1 L\n9,90 € /ud.
4		Aceite de oliva virgen extra Hacendado Gran Selección\nBotella 750 m...
5		Aceite de oliva 1º Hacendado\nBotella 1 L\n9,25 € /ud.
6		Aceite de oliva virgen Hacendado\nGarrafa 3 L\n28,50 € /ud.
7		Aceite de oliva virgen Hacendado\nBotella 1 L\n9,55 € /ud.
8		Aceite de oliva suave Hacendado\nGarrafa 3 L\n27,60 € /ud.
9		Aceite de oliva intenso Hacendado\nGarrafa 3 L\n27,60 € /ud.
10		Aceite de oliva virgen extra Hacendado\nSpray 200 ml\n3,30 € /ud.
11		Aceite de oliva virgen extra Picual Casa Juncal\nBotella 500 ml\n7,85 €...
12	Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0,2º Hacendado\nGarrafa 5 L\n6,95 €6,75 € ...
13		Aceite de girasol refinado 0,2º Hacendado\nBotella 1 L\n1,45 € /ud.
14		Aceite de coco virgen Hacendado\nBote 450 ml\n4,40 € /ud.
15	Vinagre_y_otros_aderezos	Vinagre de vino blanco Hacendado\nBotella 1 L\n0,75 € /ud.
16		Vinagre de manzana Hacendado\nBotella 1 L\n0,90 € /ud.

Figura 22. Ejemplo del resultado obtenido tras promover los encabezados del conjunto de datos de Mercadona

- **Se agregó la columna Supermercado:** Al ser esta columna un valor fijo, la forma de incluirlo fue creando una nueva columna personalizada.

	A ^B C I	A ^B C I__name	ABC 123 Supermercado
1	Aceite_de_oliva	Aceite de oliva 0,4º Hacendado\nBotella 1 L\n9,25 € /ud.	Mercadona
2		Aceite de oliva virgen extra Hacendado\nGarrafa 3 L\n29,55 € /ud.	Mercadona
3		Aceite de oliva virgen extra Hacendado\nBotella 1 L\n9,90 € /ud.	Mercadona
4		Aceite de oliva virgen extra Hacendado Gran Selección\nBotella 750 m...	Mercadona
5		Aceite de oliva 1º Hacendado\nBotella 1 L\n9,25 € /ud.	Mercadona
6		Aceite de oliva virgen Hacendado\nGarrafa 3 L\n28,50 € /ud.	Mercadona
7		Aceite de oliva virgen Hacendado\nBotella 1 L\n9,55 € /ud.	Mercadona
8		Aceite de oliva suave Hacendado\nGarrafa 3 L\n27,60 € /ud.	Mercadona
9		Aceite de oliva intenso Hacendado\nGarrafa 3 L\n27,60 € /ud.	Mercadona
10		Aceite de oliva virgen extra Hacendado\nSpray 200 ml\n3,30 € /ud.	Mercadona
11		Aceite de oliva virgen extra Picual Casa Juncal\nBotella 500 ml\n7,85 €...	Mercadona
12	Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0,2º Hacendado\nGarrafa 5 L\n6,95 €6,75 € ...	Mercadona
13		Aceite de girasol refinado 0,2º Hacendado\nBotella 1 L\n1,45 € /ud.	Mercadona
14		Aceite de coco virgen Hacendado\nBote 450 ml\n4,40 € /ud.	Mercadona
15	Vinagre_y_otros_aderezos	Vinagre de vino blanco Hacendado\nBotella 1 L\n0,75 € /ud.	Mercadona
16		Vinagre de manzana Hacendado\nBotella 1 L\n0,90 € /ud.	Mercadona

Figura 23. Ejemplo del resultado obtenido tras crear la columna de Supermercado en el conjunto de datos de Mercadona

- **Se agregó la columna fecha:** En este caso la columna fecha se incluyó añadiendo del mismo modo una columna personalizada.

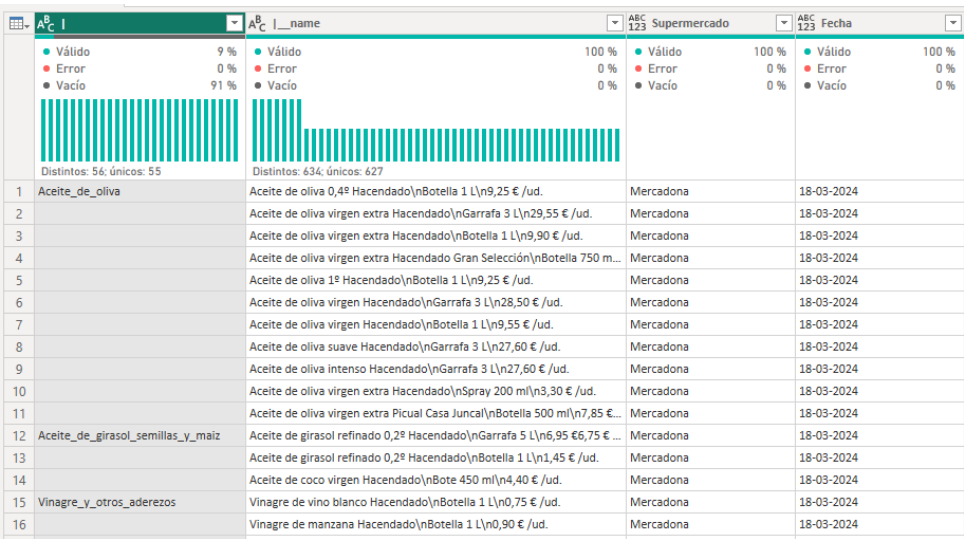


Figura 24. Ejemplo del resultado obtenido tras crear la columna de Fecha en el conjunto de datos de Mercadona

- **Se configuró el tipo de dato correctamente:** Principalmente se indicó que la nueva columna de fecha era de tipo 'date'.

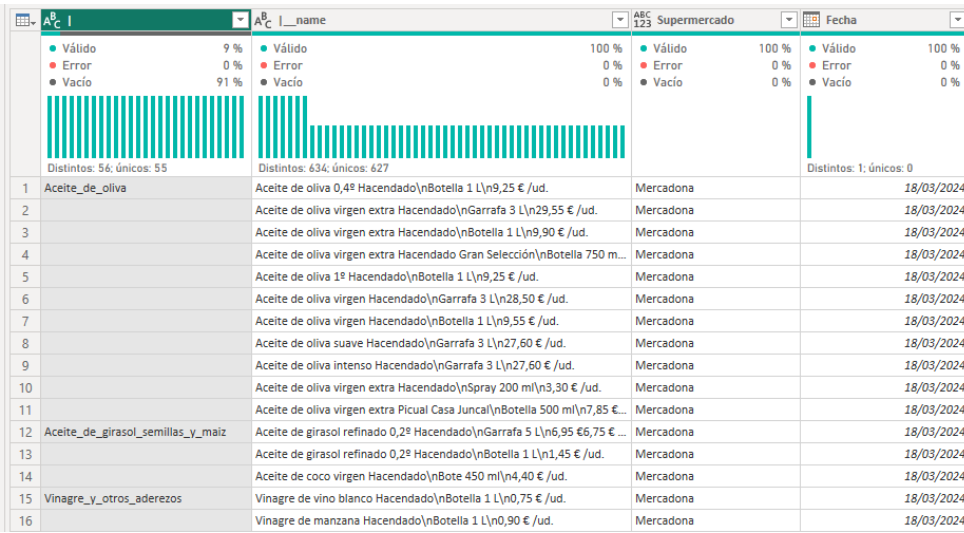


Figura 25. Ejemplo del resultado obtenido tras configurar el tipo de cada dato correctamente en el conjunto de datos de Mercadona

- **Se divide la columna 'name' por delimitador:** Dentro de la columna 'name' se podía encontrar separado por '\n' diferentes conceptos. Por lo tanto, se procedió a dividir en columnas separadas.

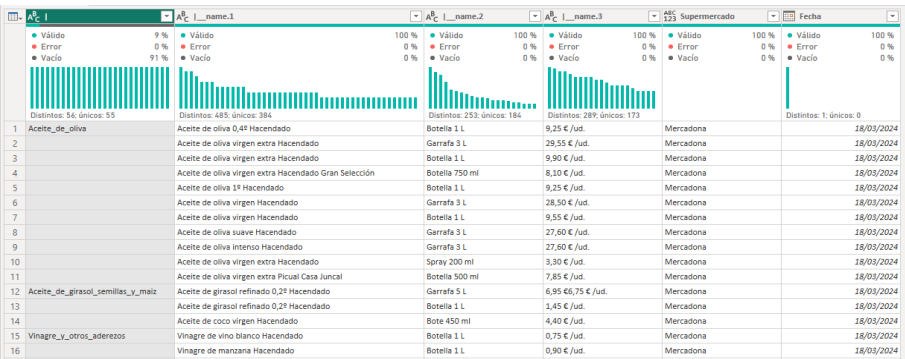


Figura 26. Ejemplo del resultado obtenido tras dividir la columna ‘name’ en el conjunto de datos de Mercadona

- **Se dividió la columna ‘precio’ por delimitador:** Dentro de las diferentes columnas generadas tras el apartado anterior, se decidió que era conveniente dividir la columna que aportaba el coste y la unidad en dos

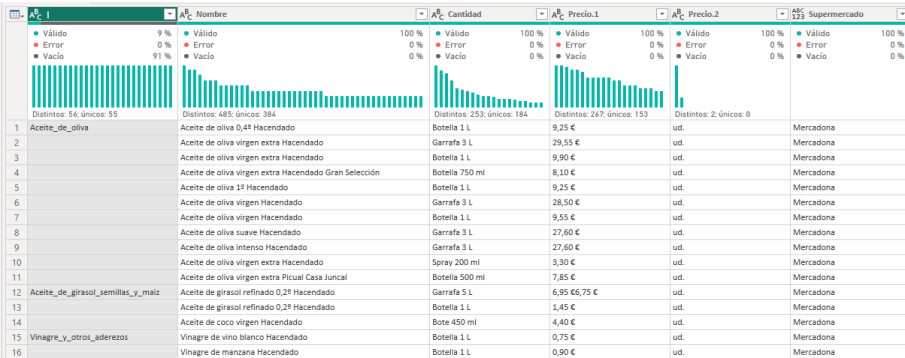


Figura 27. Ejemplo del resultado obtenido tras dividir la columna ‘precio’ en el conjunto de datos de Mercadona

- **Se cambió el nombre de las columnas:** Tras la separación de las columnas y, tras tener la estructura final del fichero, se procedió a poner los nombres definitivos de cada una de las columnas.

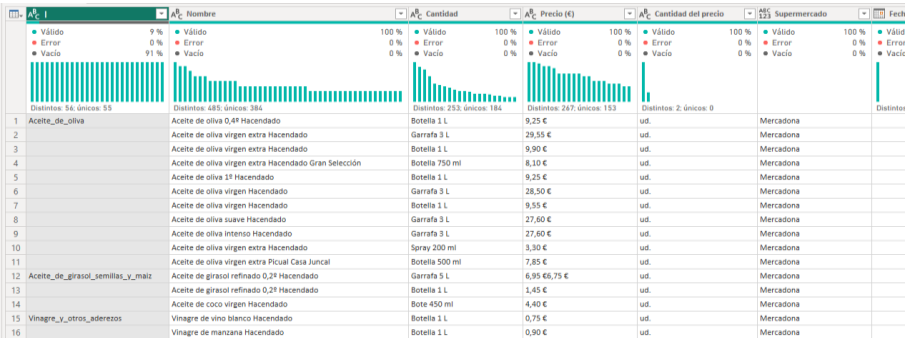


Figura 28. Ejemplo del resultado obtenido tras modificar el nombre de las columnas en el conjunto de datos de Mercadona

- **Se fraccionó una columna por delimitador:** De nuevo la columna del precio se duplicó en precio inicial y precio de la oferta.

A ₁ Categoría	A ₂ Nombre	A ₃ Cantidad	A ₄ Precio (€),1	A ₅ Precio (€),2	A ₆ Precio (€),3	A ₇ Cantidad d
1	Aceite de oliva	Botella 1 L	9,25			null ud.
2	Aceite de oliva virgen extra Hacendado	Garrafa 3 L	29,55			null ud.
3	Aceite de oliva virgen extra Hacendado	Botella 1 L	9,90			null ud.
4	Aceite de oliva virgen extra Hacendado Gran Selección	Botella 750 ml	8,10			null ud.
5	Aceite de oliva 1 ^{er} Hacendado	Botella 1 L	9,25			null ud.
6	Aceite de oliva virgen Hacendado	Garrafa 3 L	28,50			null ud.
7	Aceite de oliva virgen Hacendado	Botella 1 L	9,55			null ud.
8	Aceite de oliva suave Hacendado	Garrafa 3 L	27,60			null ud.
9	Aceite de oliva intenso Hacendado	Garrafa 3 L	27,60			null ud.
10	Aceite de oliva virgen extra Hacendado	Spray 200 ml	3,30			null ud.
11	Aceite de oliva virgen extra Pical Casa Juncal	Botella 500 ml	7,85			null ud.
12	Aceite de girasol semillas y maíz	Garrafa 5 L	6,95	6,75		null ud.
13	Aceite de girasol refinado 0,2 ^{er} Hacendado	Botella 1 L	1,45			null ud.
14	Aceite de coco virgen Hacendado	Bote 450 ml	4,40			null ud.
15	Vinagre de vino blanco Hacendado	Botella 1 L	0,75			null ud.
16	Vinagre de manzana Hacendado	Botella 1 L	0,90			null ud.

Figura 29. Ejemplo del resultado obtenido tras fraccionar la columna 'precio' en el conjunto de datos de Mercadona

- **Se rellenó hacia abajo el valor de la columna categoría:** Previamente fue necesario cambiar todos los espacios en blanco por valores nulos para, posteriormente, rellenarlos con el valor que se encuentra encima del espacio.

A ₁ Categoría	A ₂ Nombre	A ₃ Cantidad	1,2 Precio Inicio(€)	1,2 Precio Oferta(€)	A ₆ Cantidad del precio	1,1 Supermer
1	Aceite de oliva	Botella 1 L	9,25		null ud.	Mercadona
2	Aceite de oliva	Garrafa 3 L	29,55		null ud.	Mercadona
3	Aceite de oliva	Botella 1 L	9,90		null ud.	Mercadona
4	Aceite de oliva	Botella 750 ml	8,10		null ud.	Mercadona
5	Aceite de oliva	Botella 1 L	9,25		null ud.	Mercadona
6	Aceite de oliva	Garrafa 3 L	28,50		null ud.	Mercadona
7	Aceite de oliva	Botella 1 L	9,55		null ud.	Mercadona
8	Aceite de oliva	Garrafa 3 L	27,60		null ud.	Mercadona
9	Aceite de oliva	Garrafa 3 L	27,60		null ud.	Mercadona
10	Aceite de oliva	Spray 200 ml	3,30		null ud.	Mercadona
11	Aceite de oliva	Botella 500 ml	7,85		null ud.	Mercadona
12	Aceite de girasol semillas y maíz	Garrafa 5 L	6,95	6,75	null ud.	Mercadona
13	Aceite de girasol refinado 0,2 ^{er} Hacendado	Botella 1 L	1,45		null ud.	Mercadona
14	Aceite de coco virgen Hacendado	Bote 450 ml	4,40		null ud.	Mercadona
15	Vinagre de vino blanco Hacendado	Botella 1 L	0,75		null ud.	Mercadona
16	Vinagre de manzana Hacendado	Botella 1 L	0,90		null ud.	Mercadona

Figura 30. Ejemplo del resultado obtenido tras rellenar hacia abajo los datos del campo 'categoría' en el conjunto de datos de Mercadona

- **Se creó un ID:** En este caso fue necesaria la unión de la columna nombre y cantidad del producto.

1	1,2 Precio Inicio(€)	1,2 Precio Oferta(€)	A ₃ Cantidad del precio	1,1 Supermercado	Fecha	A ₇ ID
1	9,25		null ud.	Mercadona	18/03/2024	Aceite de oliva 0-4 ^{er} Hacendado-Botella 1 L
2	29,55		null ud.	Mercadona	18/03/2024	Aceite de oliva virgen extra Hacendado-Garrafa 3 L
3	9,90		null ud.	Mercadona	18/03/2024	Aceite de oliva virgen extra Hacendado-Botella 1 L
4	8,10		null ud.	Mercadona	18/03/2024	Aceite de oliva virgen extra Hacendado Gran Selección-Botella 750 ml
5	9,25		null ud.	Mercadona	18/03/2024	Aceite de oliva 1 ^{er} Hacendado-Botella 1 L
6	28,50		null ud.	Mercadona	18/03/2024	Aceite de oliva virgen Hacendado-Garrafa 3 L
7	9,55		null ud.	Mercadona	18/03/2024	Aceite de oliva virgen Hacendado-Botella 1 L
8	27,60		null ud.	Mercadona	18/03/2024	Aceite de oliva suave Hacendado-Garrafa 3 L
9	27,60		null ud.	Mercadona	18/03/2024	Aceite de oliva intenso Hacendado-Garrafa 3 L
10	3,30		null ud.	Mercadona	18/03/2024	Aceite de oliva virgen extra Hacendado-Spray 200 ml
11	7,85		null ud.	Mercadona	18/03/2024	Aceite de oliva virgen extra Pical Casa Juncal-Botella 500 ml
12	6,95	6,75	null ud.	Mercadona	18/03/2024	Aceite de girasol refinado 0-2 ^{er} Hacendado-Botella 5 L
13	1,45		null ud.	Mercadona	18/03/2024	Aceite de girasol refinado 0-2 ^{er} Hacendado-Botella 1 L
14	4,40		null ud.	Mercadona	18/03/2024	Aceite de coco virgen Hacendado-Bote 450 ml
15	0,75		null ud.	Mercadona	18/03/2024	Vinagre de vino blanco Hacendado-Botella 1 L
16	0,90		null ud.	Mercadona	18/03/2024	Vinagre de manzana Hacendado-Botella 1 L

Figura 31. Ejemplo del resultado obtenido tras crear una nueva columna ID en el conjunto de datos de Mercadona

- **Limpieza del conjunto de datos Dia:** Tras cargar la información recogida en Power BI se observa cómo era necesaria la implementación de cambios. Partiendo de la siguiente información:

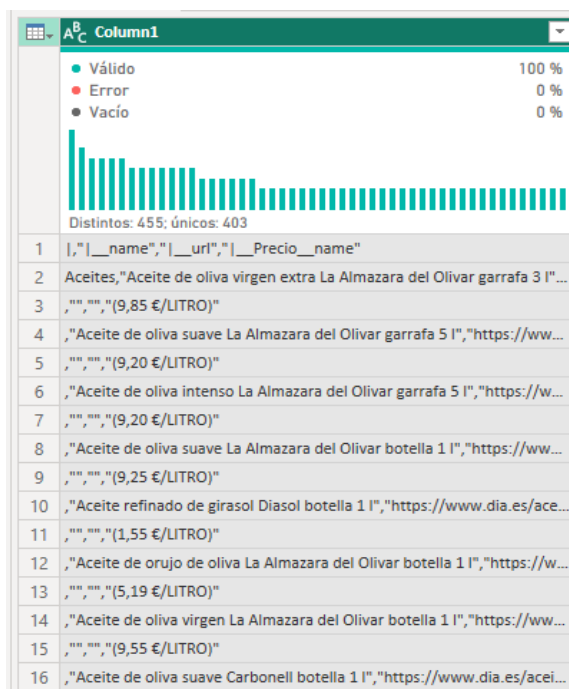


Figura 32. Ejemplo del resultado obtenido tras cargar el fichero con los datos de Dia en Power BI

Los pasos implementados para su limpieza fueron:

- **Se divide la columna por delimitador:** Teniendo en cuenta que toda la información se encontraba en la misma columna y esta venía separada por una coma, se procede a dividir en las diversas columnas.

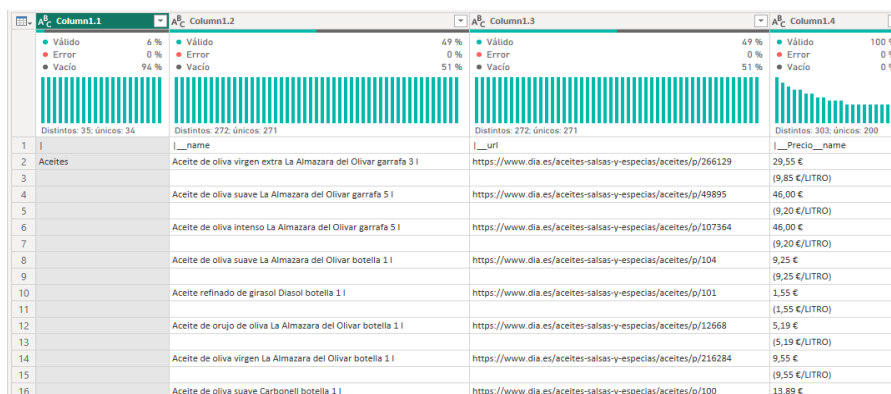


Figura 33. Ejemplo del resultado obtenido tras dividir la columna por delimitador en el conjunto de datos de Dia

- **Encabezados promovidos:** Se indicó que el valor de la primera fila debe ser la cabecera del conjunto de datos.

	A_c, I	A_c, I_name	A_c, I_url	A_c, I_Precio_name
1	Aceites	Aceite de oliva virgen extra La Almazara del Olivar garrafa 3 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/266129	29,55 €
2				(9,85 €/LITRO)
3		Aceite de oliva suave La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/49895	46,00 €
4				(9,20 €/LITRO)
5		Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/107364	46,00 €
6				(9,20 €/LITRO)
7		Aceite de oliva suave La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/104	9,25 €
8				(9,25 €/LITRO)
9		Aceite refinado de girasol Diasol botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/101	1,55 €
10				(1,55 €/LITRO)
11		Aceite de orujo de oliva La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/12668	5,19 €
12				(5,19 €/LITRO)
13		Aceite de oliva virgen La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/216284	9,55 €
14				(9,55 €/LITRO)
15		Aceite de oliva suave Carbonell botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100	13,89 €
16				(13,89 €/LITRO)

Figura 34. Ejemplo del resultado obtenido tras promover los encabezados del conjunto de datos de Dia

- **Se agrega la columna Supermercado:** Al ser esta columna un valor fijo, la forma de incluirlo fue creando una nueva columna personalizada.

	A_c, I	A_c, I_name	A_c, I_url	A_c, I_Precio_name	Supermercado
1	Aceites	Aceite de oliva virgen extra La Almazara del Olivar garrafa 3 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/266129	29,55 €	Dia
2				(9,85 €/LITRO)	Dia
3		Aceite de oliva suave La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/49895	46,00 €	Dia
4				(9,20 €/LITRO)	Dia
5		Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/107364	46,00 €	Dia
6				(9,20 €/LITRO)	Dia
7		Aceite de oliva suave La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/104	9,25 €	Dia
8				(9,25 €/LITRO)	Dia
9		Aceite refinado de girasol Diasol botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/101	1,55 €	Dia
10				(1,55 €/LITRO)	Dia
11		Aceite de orujo de oliva La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/12668	5,19 €	Dia
12				(5,19 €/LITRO)	Dia
13		Aceite de oliva virgen La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/216284	9,55 €	Dia
14				(9,55 €/LITRO)	Dia
15		Aceite de oliva suave Carbonell botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100	13,89 €	Dia
16				(13,89 €/LITRO)	Dia

Figura 35. Ejemplo del resultado obtenido tras crear la columna de “Supermercado” en el conjunto de datos de Dia

- **Se agrega la columna fecha:** En este caso la columna fecha se incluye, del mismo modo que el supermercado, mediante una columna personalizada.

	A_c, I	A_c, I_name	A_c, I_url	A_c, I_Precio_name	Supermercado	Fecha
1	Aceites	Aceite de oliva virgen extra La Almazara del Olivar garrafa 3 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/266129	29,55 €	Dia	18-03-2024
2				(9,85 €/LITRO)	Dia	18-03-2024
3		Aceite de oliva suave La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/49895	46,00 €	Dia	18-03-2024
4				(9,20 €/LITRO)	Dia	18-03-2024
5		Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/107364	46,00 €	Dia	18-03-2024
6				(9,20 €/LITRO)	Dia	18-03-2024
7		Aceite de oliva suave La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/104	9,25 €	Dia	18-03-2024
8				(9,25 €/LITRO)	Dia	18-03-2024
9		Aceite refinado de girasol Diasol botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/101	1,55 €	Dia	18-03-2024
10				(1,55 €/LITRO)	Dia	18-03-2024
11		Aceite de orujo de oliva La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/12668	5,19 €	Dia	18-03-2024
12				(5,19 €/LITRO)	Dia	18-03-2024
13		Aceite de oliva virgen La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/216284	9,55 €	Dia	18-03-2024
14				(9,55 €/LITRO)	Dia	18-03-2024
15		Aceite de oliva suave Carbonell botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100	13,89 €	Dia	18-03-2024
16				(13,89 €/LITRO)	Dia	18-03-2024

Figura 36. Ejemplo del resultado obtenido tras crear la columna de “Fecha” en el conjunto de datos de Dia

- **Se cambia el tipo de los campos:** Principalmente se indica que el campo de fecha creado anteriormente debe ser de tipo date.

A_Cat	I_name	A_Cat	I_url	A_Cat	I_Precio_name	ABC	Supermercado	Fecha
49 % Válido 0 % Error 51 % Vacio	49 % Válido 0 % Error 51 % Vacio	49 % Válido 0 % Error 51 % Vacio	100 % Válido 0 % Error 0 % Vacio	100 % Válido 0 % Error 0 % Vacio	100 % Válido 0 % Error 0 % Vacio	100 % Válido 0 % Error 0 % Vacio	100 % Válido 0 % Error 0 % Vacio	100 % Válido 0 % Error 0 % Vacio
Distintos: 271; únicos: 270	Distintos: 271; únicos: 270	Distintos: 302; únicos: 199	Distintos: 1; únicos: 0					
	Aceite de oliva virgen extra La Almazara del Olivar garrafa 3 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/266129	29,55 €	Dia	18/03/2024			
	Aceite de oliva suave La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/49895	(9,85 €/LITRO) 46,00 €	Dia	18/03/2024			
	Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/107364	(9,20 €/LITRO) 46,00 €	Dia	18/03/2024			
	Aceite de oliva suave La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/104	(9,25 €/LITRO) 9,25 €	Dia	18/03/2024			
	Aceite refinado de girasol Diasol botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/101	(1,55 €/LITRO) 1,55 €	Dia	18/03/2024			
	Aceite de orujo de oliva La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/12668	(5,19 €/LITRO) 5,19 €	Dia	18/03/2024			
	Aceite de oliva virgen La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/216284	(9,55 €/LITRO) 9,55 €	Dia	18/03/2024			
	Aceite de oliva suave Carbonell botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100	(13,89 €/LITRO) 13,89 €	Dia	18/03/2024			

Figura 37. Ejemplo del resultado obtenido tras configurar el tipo de cada dato correctamente en el conjunto de datos de Dia

- **Se cambian los nombres de las columnas:** Sustituyendo nombres que ayuden a identificar el contenido que se va a mostrar en cada una de las columnas.

A_Cat	Categoría	A_Cat	Nombre	A_Cat	url	A_Cat	I_Precio_name	ABC	Supermercado
6 % Válido 0 % Error 94 % Vacio	49 % Válido 0 % Error 51 % Vacio	49 % Válido 0 % Error 51 % Vacio	100 % Válido 0 % Error 0 % Vacio	100 % Válido 0 % Error 0 % Vacio	100 % Válido 0 % Error 0 % Vacio	100 % Válido 0 % Error 0 % Vacio	100 % Válido 0 % Error 0 % Vacio	100 % Válido 0 % Error 0 % Vacio	100 % Válido 0 % Error 0 % Vacio
Distintos: 34; únicos: 33	Distintos: 271; únicos: 270	Distintos: 271; únicos: 270	Distintos: 302; únicos: 199						
1	Aceites	Aceite de oliva virgen extra La Almazara del Olivar garrafa 3 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/266129	29,55 €	Dia				
2		Aceite de oliva suave La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/49895	(9,85 €/LITRO) 46,00 €	Dia				
3		Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/107364	(9,20 €/LITRO) 46,00 €	Dia				
4		Aceite de oliva suave La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/104	(9,25 €/LITRO) 9,25 €	Dia				
5		Aceite refinado de girasol Diasol botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/101	(1,55 €/LITRO) 1,55 €	Dia				
6		Aceite de orujo de oliva La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/12668	(5,19 €/LITRO) 5,19 €	Dia				
7		Aceite de oliva virgen La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/216284	(9,55 €/LITRO) 9,55 €	Dia				
8		Aceite de oliva suave Carbonell botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100	(13,89 €/LITRO) 13,89 €	Dia				
9									
10									
11									
12									
13									
14									
15									
16									

Figura 38. Ejemplo del resultado obtenido tras modificar el nombre de las columnas en el conjunto de datos de Dia

- **Se rellena hacia abajo el campo categoría, url y nombre:** Como se puede observar en las imágenes, para las tres columnas había espacios en blanco por lo que se procede, en primer lugar, a poner el valor como nulo y, secundariamente, a rellenar el campo con el valor que se encontraba en la fila anterior.

A _C Categoría	A _C Nombre	A _C url	A _C Precio_name	Supermercado
<div> <div>Válido 100 %</div> <div>Error 0 %</div> <div>Vacío 0 %</div> </div> <div>Distintos: 33; únicos: 0</div>	<div> <div>Válido 100 %</div> <div>Error 0 %</div> <div>Vacío 0 %</div> </div> <div>Distintos: 270; únicos: 0</div>	<div> <div>Válido 100 %</div> <div>Error 0 %</div> <div>Vacío 0 %</div> </div> <div>Distintos: 270; únicos: 0</div>	<div> <div>Válido 100 %</div> <div>Error 0 %</div> <div>Vacío 0 %</div> </div> <div>Distintos: 302; únicos: 199</div>	<div> <div>Válido 100 %</div> <div>Error 0 %</div> <div>Vacío 0 %</div> </div> <div>Distintos: 465; únicos: 133</div>
1 Aceites	Aceite de oliva virgen extra La Almazara del Olivar garrafa 3 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/266129	29,55 €	Dia
2 Aceites	Aceite de oliva virgen extra La Almazara del Olivar garrafa 3 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/266129	(9,85 €/LITRO)	Dia
3 Aceites	Aceite de oliva suave La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/49895	46,00 €	Dia
4 Aceites	Aceite de oliva suave La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/49895	(9,20 €/LITRO)	Dia
5 Aceites	Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/107364	46,00 €	Dia
6 Aceites	Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/107364	(9,20 €/LITRO)	Dia
7 Aceites	Aceite de oliva suave La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/104	9,25 €	Dia
8 Aceites	Aceite de oliva suave La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/104	(9,25 €/LITRO)	Dia
9 Aceites	Aceite refinado de girasol Diasol botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/101	1,55 €	Dia
10 Aceites	Aceite refinado de girasol Diasol botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/101	(1,55 €/LITRO)	Dia
11 Aceites	Aceite de orujo de oliva La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/12668	5,19 €	Dia
12 Aceites	Aceite de orujo de oliva La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/12668	(5,19 €/LITRO)	Dia
13 Aceites	Aceite de oliva virgen La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/216284	9,55 €	Dia
14 Aceites	Aceite de oliva virgen La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/216284	(9,55 €/LITRO)	Dia
15 Aceites	Aceite de oliva suave Carbonell botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100	13,89 €	Dia
16 Aceites	Aceite de oliva suave Carbonell botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100	(13,89 €/LITRO)	Dia

Figura 39. Ejemplo del resultado obtenido tras rellenar hacia abajo los datos del campo 'categoría', 'url' y 'nombre' en el conjunto de datos de Dia

- **Se agrupan las filas con el mismo nombre de documento:** Hasta el momento para cada producto había dos filas, una con el precio del producto y otra con el valor por unidad. Para solventar el problema se unifican ambas filas en una sola, durante el proceso se realizan los siguientes pasos:
 - Agrupación de las filas por el campo nombre del producto.

A _C Nombre	Recuento
<div> <div>Válido 100 %</div> <div>Error 0 %</div> <div>Vacío 0 %</div> </div> <div>Distintos: 270; únicos: 270</div>	<div> <div>Válido 100 %</div> <div>Error 0 %</div> <div>Vacío 0 %</div> </div>
1 Aceite de oliva virgen extra La Almazara del Olivar garrafa 3 l	Table
2 Aceite de oliva suave La Almazara del Olivar garrafa 5 l	Table
3 Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	Table
4 Aceite de oliva suave La Almazara del Olivar botella 1 l	Table
5 Aceite refinado de girasol Diasol botella 1 l	Table
6 Aceite de orujo de oliva La Almazara del Olivar botella 1 l	Table
7 Aceite de oliva virgen La Almazara del Olivar botella 1 l	Table
8 Aceite de oliva suave Carbonell botella 1 l	Table
9 Aceite de oliva suave La española botella 1 l	Table
10 Aceite de oliva virgen extra Carbonell botella 750 ml	Table

Figura 40. Ejemplo del resultado obtenido tras agrupar las filas según el nombre del producto en el conjunto de datos de Dia

- Se crea una columna nueva personalizada, esta era proveniente de una lista que contenía los valores de los precios.

ABC Nombre		Recuento	ABC 123 ListaTemporal
<div><div></div><div>Válido100 %</div><div>Error0 %</div><div>Vacío0 %</div></div> <div><div></div><div>Distintos: 270; únicos: 270</div></div>		<div><div></div><div>Válido100 %</div><div>Error0 %</div><div>Vacío0 %</div></div>	<div><div></div><div>Válido100 %</div><div>Error0 %</div><div>Vacío0 %</div></div>
1	Aceite de oliva virgen extra La Almazara del Olivar garrafa 3 l	Table	List
2	Aceite de oliva suave La Almazara del Olivar garrafa 5 l	Table	List
3	Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	Table	List
4	Aceite de oliva suave La Almazara del Olivar botella 1 l	Table	List
5	Aceite refinado de girasol Diasol botella 1 l	Table	List
6	Aceite de orujo de oliva La Almazara del Olivar botella 1 l	Table	List
7	Aceite de oliva virgen La Almazara del Olivar botella 1 l	Table	List
8	Aceite de oliva suave Carbonell botella 1 l	Table	List
9	Aceite de oliva suave La española botella 1 l	Table	List
10	Aceite de oliva virgen extra Carbonell botella 750 ml	Table	List

Figura 41. Ejemplo del resultado obtenido tras la creación de una columna nueva ‘ListaTemporal’ en el conjunto de datos de Dia

- Se expande la lista.

ABC Nombre		Recuento	ABC 123 ListaTemporal
<div><div></div><div>Válido100 %</div><div>Error0 %</div><div>Vacío0 %</div></div> <div><div></div><div>Distintos: 270; únicos: 270</div></div>		<div><div></div><div>Válido100 %</div><div>Error0 %</div><div>Vacío0 %</div></div>	<div><div></div><div>Válido100 %</div><div>Error0 %</div><div>Vacío0 %</div></div> <div><div></div><div>Distintos: 207; únicos: 167</div></div>
1	Aceite de oliva virgen extra La Almazara del Olivar garrafa 3 l	Table	29,55 €;(9,85 €/LITRO)
2	Aceite de oliva suave La Almazara del Olivar garrafa 5 l	Table	46,00 €;(9,20 €/LITRO)
3	Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	Table	46,00 €;(9,20 €/LITRO)
4	Aceite de oliva suave La Almazara del Olivar botella 1 l	Table	9,25 €;(9,25 €/LITRO)
5	Aceite refinado de girasol Diasol botella 1 l	Table	1,55 €;(1,55 €/LITRO)
6	Aceite de orujo de oliva La Almazara del Olivar botella 1 l	Table	5,19 €;(5,19 €/LITRO)
7	Aceite de oliva virgen La Almazara del Olivar botella 1 l	Table	9,55 €;(9,55 €/LITRO)
8	Aceite de oliva suave Carbonell botella 1 l	Table	13,89 €;(13,89 €/LITRO)
9	Aceite de oliva suave La española botella 1 l	Table	13,45 €;(13,45 €/LITRO)
10	Aceite de oliva virgen extra Carbonell botella 750 ml	Table	10,89 €;(14,52 €/LITRO)

Figura 42. Ejemplo del resultado obtenido tras expandir la columna compuesta por una lista en el conjunto de datos de Dia

- Se crea una nueva columna con el resto de campos.

A ^B _C Nombre	Recuento	A ^B _C ListaTemporal	A ^B _C ListaUnida
<div> <div>Válido 100 %</div> <div>Error 0 %</div> <div>Vacío 0 %</div> </div> <div>Distintos: 270; únicos: 270</div>	<div> <div>Válido 100 %</div> <div>Error 0 %</div> <div>Vacío 0 %</div> </div>	<div> <div>Válido 100 %</div> <div>Error 0 %</div> <div>Vacío 0 %</div> </div> <div>Distintos: 207; únicos: 167</div>	
1 Aceite de oliva virgen extra La Almazara del Olivar garrafa 3 l	Table	29,55 €;(9,85 €/LITRO)	Table
2 Aceite de oliva suave La Almazara del Olivar garrafa 5 l	Table	46,00 €;(9,20 €/LITRO)	Table
3 Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	Table	46,00 €;(9,20 €/LITRO)	Table
4 Aceite de oliva suave La Almazara del Olivar botella 1 l	Table	9,25 €;(9,25 €/LITRO)	Table
5 Aceite refinado de girasol Diasol botella 1 l	Table	1,55 €;(1,55 €/LITRO)	Table
6 Aceite de orujo de oliva La Almazara del Olivar botella 1 l	Table	5,19 €;(5,19 €/LITRO)	Table
7 Aceite de oliva virgen La Almazara del Olivar botella 1 l	Table	9,55 €;(9,55 €/LITRO)	Table
8 Aceite de oliva suave Carbonell botella 1 l	Table	13,89 €;(13,89 €/LITRO)	Table
9 Aceite de oliva suave La española botella 1 l	Table	13,45 €;(13,45 €/LITRO)	Table
10 Aceite de oliva virgen extra Carbonell botella 750 ml	Table	10,89 €;(14,52 €/LITRO)	Table

Figura 43. Ejemplo del resultado obtenido tras la creación de una columna nueva 'ListaUnida' en el conjunto de datos de Dia

- Se expanden todos los valores.

A ^B _C Nombre	Recuento	A ^B _C ListaTemporal	A ^B _C ListaUnida.Categoría	A ^B _C ListaUnida.url
<div> <div>Válido 100 %</div> <div>Error 0 %</div> <div>Vacío 0 %</div> </div> <div>Distintos: 270; únicos: 270</div>	<div> <div>Válido 100 %</div> <div>Error 0 %</div> <div>Vacío 0 %</div> </div>	<div> <div>Válido 100 %</div> <div>Error 0 %</div> <div>Vacío 0 %</div> </div> <div>Distintos: 207; únicos: 167</div>		
1 Aceite de oliva virgen extra La Almazara del Olivar garrafa 3 l	Table	29,55 €;(9,85 €/LITRO)	Aceites	https://www.dia.es/aceites-salsas-y-especias/aceites/p/266129
2 Aceite de oliva suave La Almazara del Olivar garrafa 5 l	Table	46,00 €;(9,20 €/LITRO)	Aceites	https://www.dia.es/aceites-salsas-y-especias/aceites/p/49895
3 Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	Table	46,00 €;(9,20 €/LITRO)	Aceites	https://www.dia.es/aceites-salsas-y-especias/aceites/p/107364
4 Aceite de oliva suave La Almazara del Olivar botella 1 l	Table	9,25 €;(9,25 €/LITRO)	Aceites	https://www.dia.es/aceites-salsas-y-especias/aceites/p/104
5 Aceite refinado de girasol Diasol botella 1 l	Table	1,55 €;(1,55 €/LITRO)	Aceites	https://www.dia.es/aceites-salsas-y-especias/aceites/p/101
6 Aceite de orujo de oliva La Almazara del Olivar botella 1 l	Table	5,19 €;(5,19 €/LITRO)	Aceites	https://www.dia.es/aceites-salsas-y-especias/aceites/p/12668
7 Aceite de oliva virgen La Almazara del Olivar botella 1 l	Table	9,55 €;(9,55 €/LITRO)	Aceites	https://www.dia.es/aceites-salsas-y-especias/aceites/p/216284
8 Aceite de oliva suave Carbonell botella 1 l	Table	13,89 €;(13,89 €/LITRO)	Aceites	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100
9 Aceite de oliva suave La española botella 1 l	Table	13,45 €;(13,45 €/LITRO)	Aceites	https://www.dia.es/aceites-salsas-y-especias/aceites/p/1548
10 Aceite de oliva virgen extra Carbonell botella 750 ml	Table	10,89 €;(14,52 €/LITRO)	Aceites	https://www.dia.es/aceites-salsas-y-especias/aceites/p/18013

Figura 44. Ejemplo del resultado obtenido tras la expansión de la columna 'ListaUnida' en el conjunto de datos de Dia

- Se divide la columna generada con el precio: Tras la agrupación esta columna mostraba ambos precios y, mediante los siguientes pasos se procede a dividir:
 - Se elimina de la segunda columna el valor del producto ya que aparece duplicado.

Sofía Corral Caballero

ABC 123	ListaUnida. __Precio__name	A ^B _C ListaTemporal
<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div></div>
<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div></div>
		<div><div></div><div></div><div></div></div>
		Distintos: 179; únicos: 129
29,55 €		(9,85 €/LITRO)
46,00 €		(9,20 €/LITRO)
46,00 €		(9,20 €/LITRO)
9,25 €		(9,25 €/LITRO)
1,55 €		(1,55 €/LITRO)
5,19 €		(5,19 €/LITRO)
9,55 €		(9,55 €/LITRO)
13,89 €		(13,89 €/LITRO)
13,45 €		(13,45 €/LITRO)

Figura 45. Ejemplo del resultado obtenido tras la división de la columna generada con el precio en el conjunto de datos de Dia

- Se eliminan los paréntesis.

ABC 123	ListaUnida. __Precio__name	A ^B _C ListaTemporal
<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div></div>
<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div></div>
		<div><div></div><div></div><div></div></div>
		Distintos: 178; únicos: 128
29,55 €		9,85 €/LITRO
46,00 €		9,20 €/LITRO
46,00 €		9,20 €/LITRO
9,25 €		9,25 €/LITRO
1,55 €		1,55 €/LITRO
5,19 €		5,19 €/LITRO
9,55 €		9,55 €/LITRO
13,89 €		13,89 €/LITRO
13,45 €		13,45 €/LITRO

Figura 46. Ejemplo del resultado obtenido tras la eliminación de los paréntesis en el conjunto de datos de Dia

- Se divide la columna en dos con el uso del delimitador, creando una columna con el precio por unidad y otra con la unidad en la que se mide.

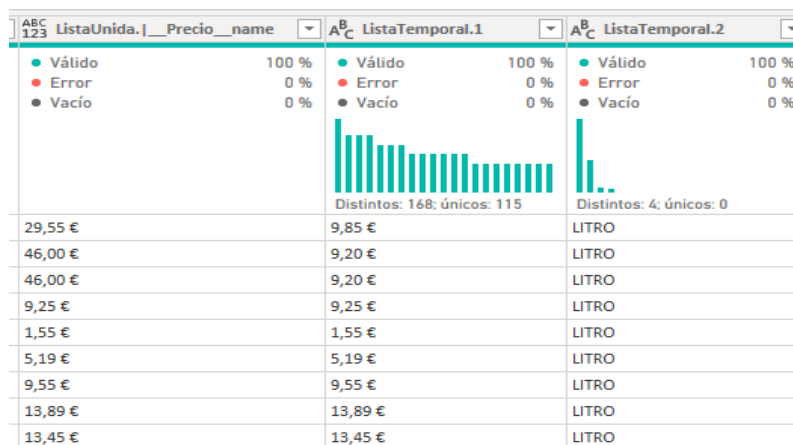


Figura 47. Ejemplo del resultado obtenido tras dividir la columna precio en el conjunto de datos de Dia

Con todos los cambios implementados en ambos orígenes, se obtienen dos conjuntos de datos limpios y preparados para ser consumidos en la aplicación final para el usuario.

3.2.1.3 Modelo de datos

Tras obtener y limpiar la información que se necesita para la creación de la aplicación, es importante que se dedique un tiempo a la creación de un modelo de datos correcto ya que ayuda a entender con mayor facilidad los datos y poder tomar decisiones.

Dentro del modelo se pueden encontrar diferentes tipos de tablas (Dimensiones y hechos), en este caso se pueden encontrar las siguientes:

- **Tablas de hechos:** Son aquellas que contienen los campos claves o la información que queremos analizar (Nieto, 2018). Ellas son:
 - Tabla de hechos de Mercadona.
 - Tabla de hechos de Dia.
- **Tablas de dimensiones:** Son aquellas que aportan información adicional a las tablas de hechos y que necesitan ir unidas de estas para tener sentido. En este caso en concreto son las siguientes:
 - Tabla de dimensiones de fecha.
 - Tabla de dimensiones con los IDs.

Adicionalmente, un modelo de datos se caracteriza por las relaciones entre tablas, ya que existen diferentes opciones. En este caso, únicamente se ha usado la relación 1...n. Por ejemplo, para un valor de fecha puede existir uno o más productos en la tabla de Mercadona.

El modelo de datos implementado en el informe de Power BI y usado para el producto final es el siguiente:

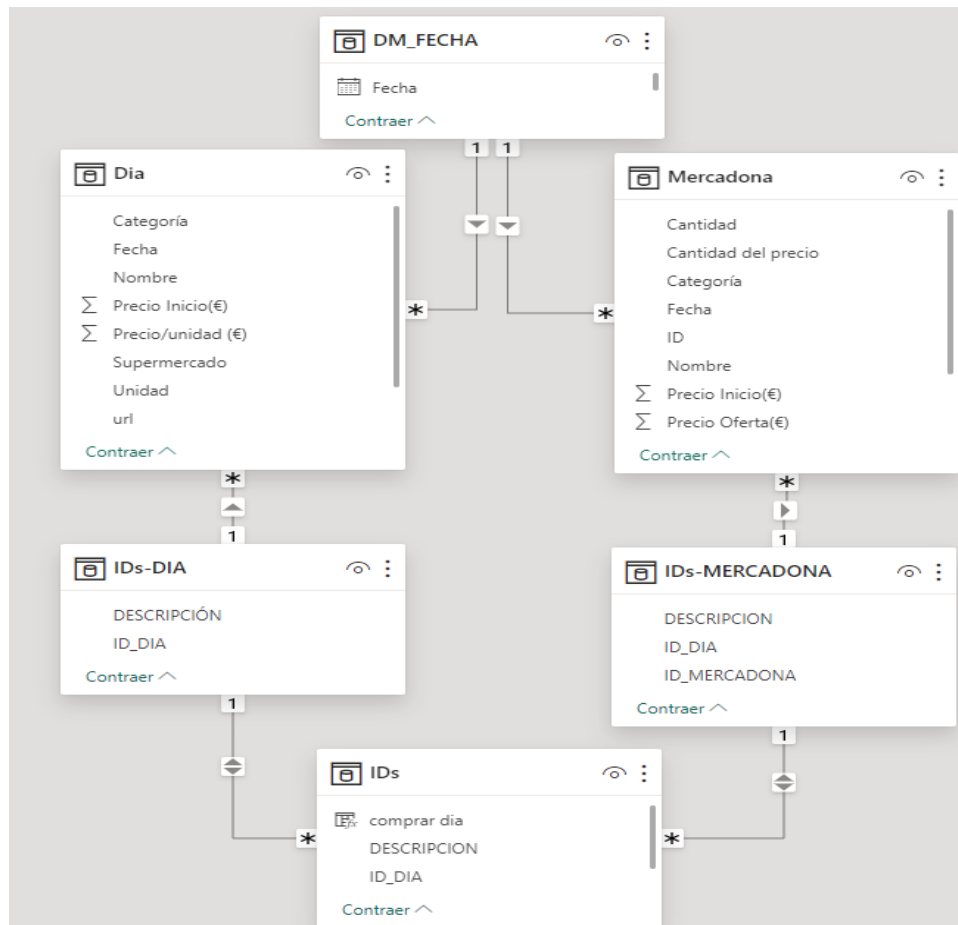


Figura 48. Esquema del modelo de datos implementado y su cardinalidad

3.2.1.4 Calidad del dato

Una vez obtenidos los datos, limpiados y creado un modelo de datos único, se ha procedido a verificar su veracidad mediante diferentes visualizaciones, métricas o KPIs en Power BI. Para ello, se ha comprobado lo siguiente:

- **El nombre de cada columna corresponde con su dato:** En este caso, lo que se quería comprobar es que los datos se vieran correctos para cada una de las columnas que componen el conjunto de datos, es decir, se ha creado una tabla con todos los datos y revisado sus valores. Adicionalmente, se han incluido en la misma visualización una serie de filtros como son el nombre del producto, la categoría y la fecha con el objetivo de verificar la calidad del dato en el histórico al completo. Para dicha comprobación se han creado dos páginas distintas por supermercado:

DIA		Fecha	Categoría	Nombre					
		Todas	Todas	Todas					
Categoría	Nombre	url	Precio Inicio(€)	Precio/unidad (€)	Unidad	Supermercado	Fecha		
Aceites	Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/107364	42.50 €	9.10 €	LITRO	Dia	29/04/2024		
Aceites	Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/107364	45.50 €	9.10 €	LITRO	Dia	15/04/2024		
Aceites	Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/107364	45.50 €	9.10 €	LITRO	Dia	22/04/2024		
Aceites	Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/107364	46.00 €	9.20 €	LITRO	Dia	11/03/2024		
Aceites	Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/107364	46.00 €	9.20 €	LITRO	Dia	18/03/2024		
Aceites	Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/107364	46.00 €	9.20 €	LITRO	Dia	25/03/2024		
Aceites	Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/107364	46.00 €	9.20 €	LITRO	Dia	01/04/2024		
Aceites	Aceite de oliva intenso La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/107364	46.00 €	9.20 €	LITRO	Dia	08/04/2024		
Aceites	Aceite de oliva suave Carbonell botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100	13.89 €	13.89 €	LITRO	Dia	06/05/2024		
Aceites	Aceite de oliva suave Carbonell botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100	13.89 €	13.89 €	LITRO	Dia	11/03/2024		
Aceites	Aceite de oliva suave Carbonell botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100	13.89 €	13.89 €	LITRO	Dia	18/03/2024		
Aceites	Aceite de oliva suave Carbonell botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100	13.89 €	13.89 €	LITRO	Dia	25/03/2024		
Aceites	Aceite de oliva suave Carbonell botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100	13.89 €	13.89 €	LITRO	Dia	01/04/2024		
Aceites	Aceite de oliva suave Carbonell botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100	13.89 €	13.89 €	LITRO	Dia	08/04/2024		
Aceites	Aceite de oliva suave Carbonell botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100	13.89 €	13.89 €	LITRO	Dia	15/04/2024		
Aceites	Aceite de oliva suave Carbonell botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100	13.89 €	13.89 €	LITRO	Dia	22/04/2024		
Aceites	Aceite de oliva suave Carbonell botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100	13.89 €	13.89 €	LITRO	Dia	06/05/2024		
Aceites	Aceite de oliva suave Carbonell botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/100	18.89 €	13.89 €	LITRO	Dia	29/04/2024		
Aceites	Aceite de oliva suave La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/104	9.25 €	9.25 €	LITRO	Dia	29/04/2024		
Aceites	Aceite de oliva suave La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/104	9.25 €	9.25 €	LITRO	Dia	11/03/2024		
Aceites	Aceite de oliva suave La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/104	9.25 €	9.25 €	LITRO	Dia	18/03/2024		
Aceites	Aceite de oliva suave La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/104	9.25 €	9.25 €	LITRO	Dia	25/03/2024		
Aceites	Aceite de oliva suave La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/104	9.25 €	9.25 €	LITRO	Dia	01/04/2024		
Aceites	Aceite de oliva suave La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/104	9.25 €	9.25 €	LITRO	Dia	08/04/2024		
Aceites	Aceite de oliva suave La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/104	9.25 €	9.25 €	LITRO	Dia	15/04/2024		
Aceites	Aceite de oliva suave La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/104	9.25 €	9.25 €	LITRO	Dia	22/04/2024		
Aceites	Aceite de oliva suave La Almazara del Olivar botella 1 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/104	9.25 €	9.25 €	LITRO	Dia	06/05/2024		
Aceites	Aceite de oliva suave La Almazara del Olivar garrafa 5 l	https://www.dia.es/aceites-salsas-y-especias/aceites/p/49895	42.50 €	9.10 €	LITRO	Dia	29/04/2024		

Figura 49. Página creada en Power BI con el conjunto entero de datos de Dia

MERCADONA		Fecha	Categoría	Nombre					
		Todas	Todas	Todas					
Categoría	Nombre	Cantidad	Precio Inicio(€)	Precio Oferta(€)	Cantidad del precio	Supermercado	Fecha		
Aceite_de_girasol_semillas_y_maiz	Aceite de coco virgen Hacendado	Bote 450 ml	4.40 €	ud.		Mercadona	11/03/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de coco virgen Hacendado	Bote 450 ml	4.40 €	ud.		Mercadona	18/03/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de coco virgen Hacendado	Bote 450 ml	4.40 €	ud.		Mercadona	25/03/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de coco virgen Hacendado	Bote 450 ml	4.40 €	ud.		Mercadona	01/04/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de coco virgen Hacendado	Bote 450 ml	4.40 €	ud.		Mercadona	08/04/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de coco virgen Hacendado	Bote 450 ml	4.40 €	ud.		Mercadona	15/04/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de coco virgen Hacendado	Bote 450 ml	4.40 €	ud.		Mercadona	22/04/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de coco virgen Hacendado	Bote 450 ml	4.40 €	ud.		Mercadona	29/04/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de coco virgen Hacendado	Bote 450 ml	4.40 €	ud.		Mercadona	06/05/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0.2º Hacendado	Botella 1 L	1.45 €	ud.		Mercadona	11/03/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0.2º Hacendado	Botella 1 L	1.45 €	ud.		Mercadona	25/03/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0.2º Hacendado	Garrafa 5 L	6.95 €	6.75 ud.		Mercadona	11/03/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0.2º Hacendado	Garrafa 5 L	6.95 €	6.75 ud.		Mercadona	25/03/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0.2º Hacendado	Botella 1 L	1.45 €	ud.		Mercadona	18/03/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0.2º Hacendado	Botella 1 L	1.45 €	ud.		Mercadona	01/04/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0.2º Hacendado	Botella 1 L	1.45 €	ud.		Mercadona	08/04/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0.2º Hacendado	Botella 1 L	1.45 €	ud.		Mercadona	15/04/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0.2º Hacendado	Botella 1 L	1.45 €	ud.		Mercadona	22/04/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0.2º Hacendado	Botella 1 L	1.45 €	ud.		Mercadona	29/04/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0.2º Hacendado	Botella 1 L	1.45 €	ud.		Mercadona	06/05/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0.2º Hacendado	Garrafa 5 L	6.95 €	6.75 ud.		Mercadona	18/03/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0.2º Hacendado	Garrafa 5 L	6.95 €	6.75 ud.		Mercadona	01/04/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0.2º Hacendado	Garrafa 5 L	6.75 €	ud.		Mercadona	08/04/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0.2º Hacendado	Garrafa 5 L	6.75 €	ud.		Mercadona	15/04/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0.2º Hacendado	Garrafa 5 L	6.75 €	ud.		Mercadona	22/04/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0.2º Hacendado	Garrafa 5 L	6.75 €	ud.		Mercadona	29/04/2024		
Aceite_de_girasol_semillas_y_maiz	Aceite de girasol refinado 0.2º Hacendado	Garrafa 5 L	6.75 €	ud.		Mercadona	06/05/2024		

Figura 50. Página creada en Power BI con el conjunto entero de datos de Mercadona

Esta visualización ayuda a poner todas las fechas de los datos con el mismo formato, que los precios tengan dos decimales... Además, se ha comprobado que la relación entre la categoría y el producto tenía sentido o que, en el caso de los datos de Dia, el link que nos llevaba a su página y su producto funcionaba correctamente.

- **Recuentos de diferentes columnas:** Otra manera de comprobar la veracidad del conjunto de datos es realizar diferentes recuentos distintos de valores y de esta forma

se comprueba que el dato tiene sentido. Los cálculos que se han realizado para ambos supermercados han sido los siguientes:

- **Recuento distinto del nombre del producto:** Teniendo en cuenta que en Mercadona se han registrado más productos que en Dia, usando el filtro de fechas, se pretende analizar la diferencia que hay entre la semana pasada y la actual ya que, aunque cada semana al hacer web scraping a los datos su número podría variar, este diferencia debería ser mínima.
 - **Recuento distinto de fechas:** Con dicha información, se pretende poder revisar si el proceso de almacenamiento de los datos semanalmente ha ido correctamente.
 - **Recuento distinto de categorías:** Del mismo modo que el recuento distinto del producto, la finalidad de este KPI es la misma, verificar que se han almacenado todos los datos correctamente, pero en este caso debería aparecer el mismo valor para cada día.
- **Comprobación de los precios en los productos:** Otro punto importante es comprobar si los precios están correctamente almacenados y, para ello, se ha creado una matriz que muestra la suma de precios por categorías de productos o por nombre para los últimos 4 días almacenados.
- La finalidad principal es comprobar si se han almacenado correctamente los datos teniendo en cuenta que, aunque el precio puede variar, no habrá mucha diferencia en siete días.

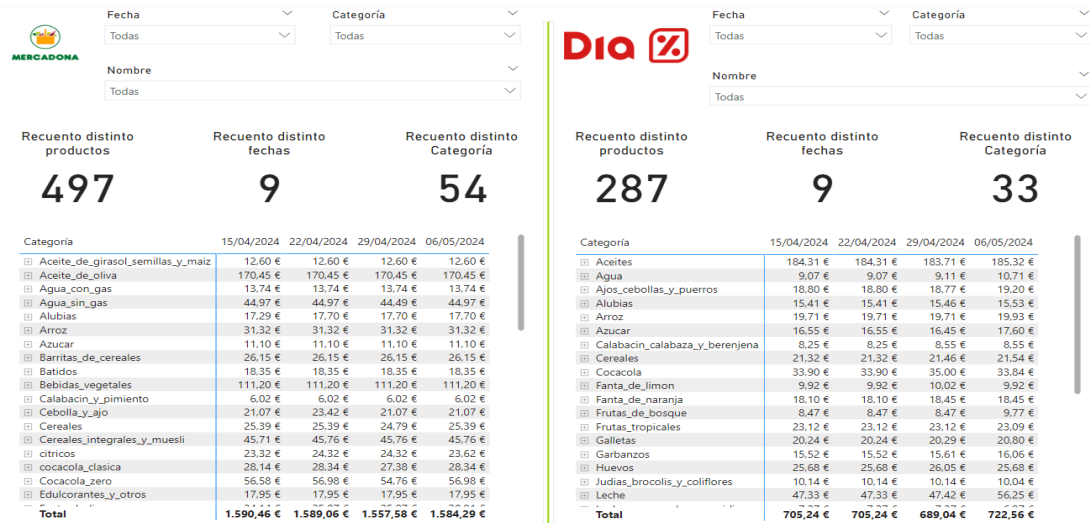


Figura 51. Página creada en Power BI para comprobar la calidad del dato de ambos conjuntos de datos

- **Evolución:** Como valor adicional, se han creado dos páginas nuevas en el informe de Power BI una para cada supermercado con un filtro por nombre y en la que se muestra la evolución de precios.

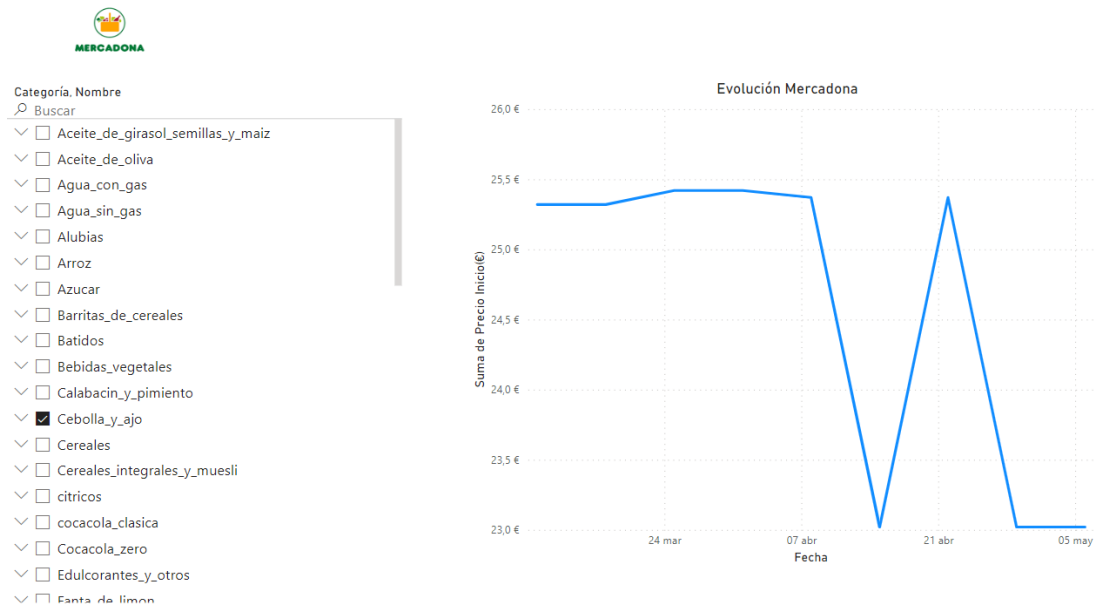


Figura 52. Página creada en Power BI que muestra la evolución del precio por producto de Mercadona

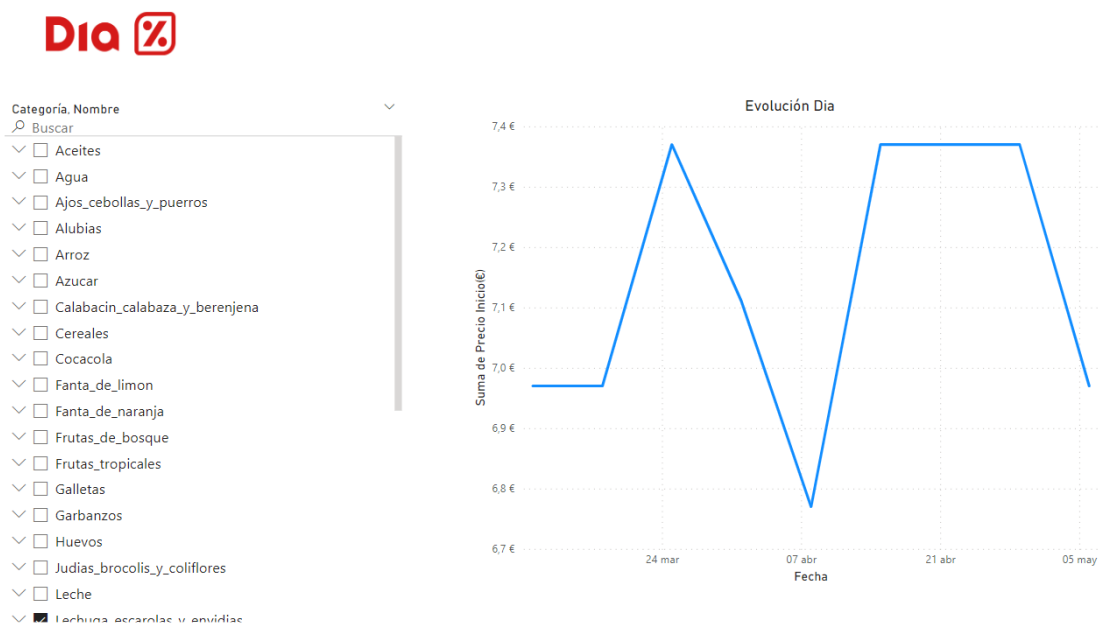


Figura 53. Página creada en Power BI que muestra la evolución del precio por producto de Dia

3.2.1.5 Aplicación

Tras realizar todo el proceso de recopilación de los datos, limpieza de los mismos y creación de un modelo de datos se pretende crear un servicio enfocado al usuario final. En este caso en concreto y, teniendo en cuenta el uso actual de los dispositivos móviles, se ha decidido crear una aplicación ya que de este modo el usuario va a poder acceder a ella desde cualquier lugar, es decir, podrá añadir productos a su cesta de la compra desde casa o desde cualquier lugar y también, podrá ver su lista de productos cuando esté en el supermercado.

En primer lugar y una vez determinado el servicio, se requiere conocer la información a representar en ella, que se quiere enseñar al usuario y que información puede resultar útil y, por último, ayudar a obtener conclusiones y poder tomar decisiones. Tras realizar un brainstorming de ideas, las siguientes opciones de información son las consideradas más apropiadas para enseñar al cliente:

- **Elaboración propia de una lista de la compra:** Con este apartado, se pretende solicitar al usuario los productos que desea solicitar y, por consiguiente, los que desea comparar. De este modo, puede obtener por un lado una comparación de precios y, por el otro, su lista de la compra para usar cuando vaya a los supermercados.
- **Ayudar en la recomendación del mejor supermercado sobre el que comprar cada producto:** De esta forma se le presentará al usuario una matriz comparativa con los precios de ambos supermercados que le ayudará a ver cual de ellos es más barato y su diferencia.
- **Información adicional sobre la evolución de los precios:** Esto ayudará al usuario a la toma de decisiones a futuro, ya que podrá conocer la evolución de sus productos y compararlo con semanas anteriores.
- **Asistir al usuario ante cualquier duda:** Con este punto, se pretende asistir al usuario dando respuesta a sus preguntas mediante un chatbot.

En segundo lugar, se debe estudiar la mejor manera de mostrar al usuario la información, ya que debe ser fácil de entender y estar bien estructurada para facilitar su experiencia. Para ello, se ha realizado un prototipo compuesto de tres pantallas diferentes sobre las que el usuario puede navegar con la ayuda de unos botones. Los prototipos son los siguientes:

- **Pantalla Carrito:** Dentro de esta pantalla, la idea es mostrar al usuario el listado de todos los productos para que pueda ir añadiendo a la lista todos los que pretenda comprar. Además, se puede acceder al resto de pantallas con el uso de botones.



Figura 54. Prototipo de la pestaña de "Carrito" para la aplicación

- **Pantalla Ahorro:** En esta pantalla se pretende mostrar al usuario una tabla comparativa de los precios de ambos supermercados con los productos previamente seleccionados en la pantalla anterior. De igual modo, podrá moverse hasta cualquier otra pantalla sin problema con la ayuda de botones.



Figura 55. Prototipo de la pestaña de "Ahorro" para la aplicación

- **Pantalla Análisis:** Para que el usuario sea capaz de analizar en detalle la evolución de los precios y tomar sus propias conclusiones. Pretende mostrar una serie de gráficos, KPIs o medidas que ayuden a ello.

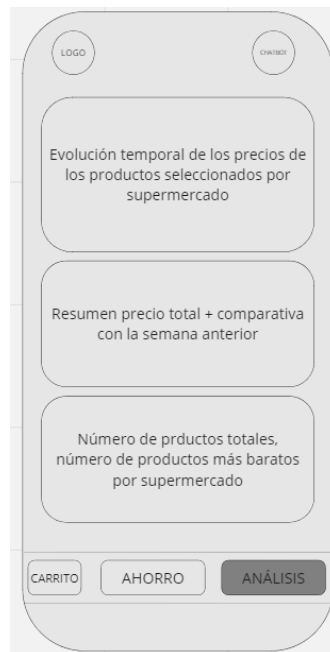


Figura 56. Prototipo de la pestaña de "Análisis" para la aplicación

- **Pantalla Chatbot:** A esta pantalla se podrá acceder desde cualquier otra pantalla y en su interior habrá un chatbot con una barra para introducir la pregunta en cuestión, una serie de preguntas propuestas y un espacio para su respuesta.



Figura 57. Prototipo de la pestaña del chatbot para la aplicación

Con el prototipo e ideas diseñados, se procede a implementarlo. Para ello, se ha usado Power BI ya que es una herramienta usada previamente en el estudio de la calidad del dato y, para uso personal no requiere licencia.

Para la creación de la aplicación ha sido necesario cargar el conjunto de datos obtenido de los diferentes supermercados, así como la tabla que relaciona los nombres de cada uno de ellos y su modelo de datos.

El desarrollo está dividido en tres grandes grupos. El primero de ellos es el relacionado con los visuales, los KPIs o las medidas necesarias a calcular para su visualización, por otro lado, se encuentra tanto la implementación como el entrenamiento de un chatbot y, el tercer grupo engloba la parte visual y experiencia del usuario en la cual se debe tener en cuenta no solo los colores o colocación de cada visual sino que también se deben implementar los botones correspondientes para moverse por las diferentes páginas.

Dentro del primer grupo se han realizado los siguientes visuales:

- **Filtro del nombre de los productos:** Para la primera pantalla (pantalla denominada 'Carrito' en el prototipo) se ha creado un visual con el nombre de todos los productos que se tienen almacenados en la base de datos. Este visual es un filtro al cual se le han implementado los siguientes cambios:
 - Se ha añadido la opción de buscador para facilitar al usuario su experiencia, de este modo podrá introducir el nombre del producto a elegir y seleccionar aquel que desee.

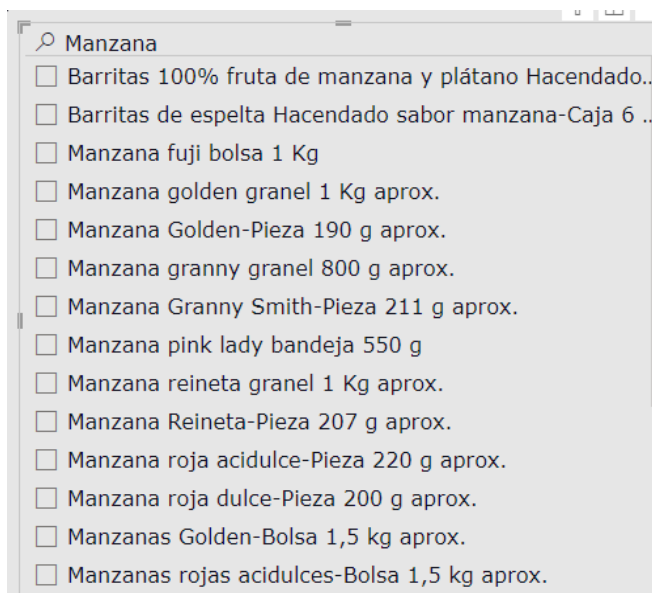


Figura 58. Imagen del visual del filtro de productos que muestra la opción de buscador implementada

- Se ha incluido la opción de selección múltiple para que se puede elegir más de un producto a la vez.

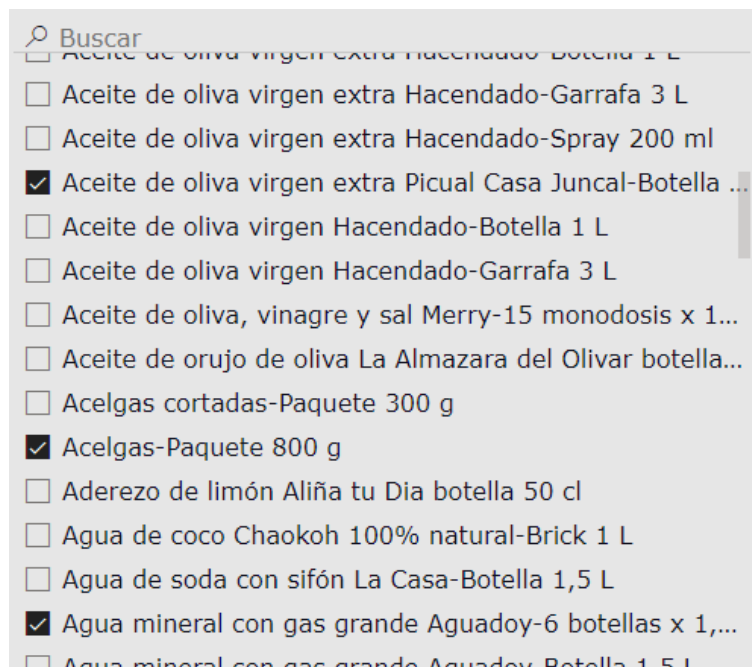


Figura 59. Imagen del visual del filtro de productos que muestra la opción seleccionar múltiples productos

- Se ha habilitado la opción de seleccionar todos los productos con la finalidad de comparar la evolución genérica de los productos.
 - Se ha creado un filtro no visible en el resto de pantallas sincronizados entre sí con este primero para que las demás pantallas muestren información correspondiente únicamente a los productos seleccionados.
- **Matriz comparativa de precios:** La segunda pantalla tiene como finalidad mostrar una comparativa de precios entre ambos supermercados para los productos seleccionados. Para ello se ha implementado una matriz que muestra en las filas los nombres de cada producto y, como columnas los precios en cada supermercado (Dia y Mercadona). Adicionalmente a esto, se ha añadido la suma total de los productos y se ha filtrado el visual para que muestre únicamente el precio del último día registrado, es decir, el precio más actual que se tiene almacenado en la base de datos.

DESCRIPCION	DIA	MERCADONA
Aceite de oliva virgen extra Picual Casa Juncal-Botella 500 ml		7,85 €
Acelgas-Paquete 800 g	1,50 €	2,26 €
Agua mineral con gas grande Aguadoy-6 botellas x 1,5 L		2,46 €
Agua mineral grande Aguadoy-Garrafa 6 L	0,65 €	0,78 €
Cebollas dulces-Malla 1 kg	2,49 €	2,44 €
Champiñón Portobello-Bandeja 300 g aprox.	1,69 €	2,03 €
Galletas tostadas Tosta Rica-Caja 3 paquetes (570 g)	2,99 €	3,00 €
Limonas-Malla 1 kg	1,89 €	1,79 €
Vinagre balsámico de Módena Hacendado-Botella 500 ml	2,39 €	2,50 €
	13,60 €	25,11 €

Figura 60. Imagen de la matriz que compara los precios de ambos supermercados

- **Evolución de los precios de ambos supermercados:** Se ha creado una gráfica de líneas que muestra la evolución en el tiempo de los precios de ambos supermercados. Para ello se pueden observar dos líneas diferentes, la verde correspondiente con los precios de Mercadona y la roja relacionada con los precios de Dia. Adicionalmente se ha implementado una línea de tendencia entre ambos supermercados

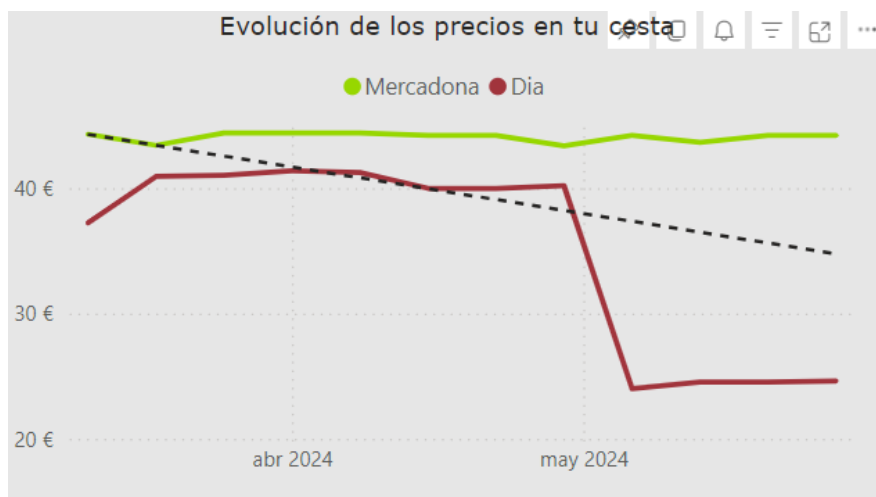


Figura 61. Gráfica lineal que muestra la evolución en el tiempo del precio de los productos para ambos supermercados

- **KPI precio total de los productos para cada supermercado:** Este dato es la suma de todos los valores que se encuentran en la columna de precios, es decir, teniendo en cuenta que se aplica el filtro de los productos marcados en la primera página, este KPI muestra el coste total de los productos seleccionados para cada supermercado.

Precio Total	
	25,11 €
	13,60 €

Figura 62. KPIs que muestran el total del precio de los productos seleccionados para ambos supermercados

- **Comparativa costes vs semana anterior:** En este caso, se realiza una comparativa de costes de esta semana con los mismos productos y su precio en la semana anterior. Para esto, ha sido necesario la creación de dos columnas binarias previas, una indicando la última fecha registrada en la base de datos y otra con la penúltima. La medida final calcula representa el porcentaje y lo hace sumando el precio de la semana pasada entre el precio actual del producto tal y como se muestra a continuación:

$$\text{mercadonaincremento} = 1 - \left(\frac{\text{CALCULATE(SUM(Mercadona[Precio Inicio(€)]), DM_FECHA[penultimaFecha] = "Si")}}{\text{CALCULATE(SUM(Mercadona[Precio Inicio(€)]), DM_FECHA[ultimaFecha] = "Si")}} \right)$$

Como valor adicional, se ha incluido una condición que, si el precio ha disminuido o se ha mantenido igual, se marcará de color verde y si ha aumentado se marcará de color rojo.

	Precio Total	vs Semana Anterior
	25,11 €	0,00 %
	13,60 €	-38,38 %

Figura 63. KPIs que muestran el precio total de los productos seleccionados y su comparativa con el precio de la semana anterior para ambos supermercados

- **Número de productos totales:** Este dato nos dice el número total de productos que tenemos añadidos a la cesta de la compra y se calcula haciendo un recuento distinto del campo 'descripción'.



Figura 64. KPI que indica el número total de productos añadidos a la cesta de la compra

- **Número de productos por supermercado:** En este caso se hace un recuento del mismo modo que en el apartado anterior, pero se añade la condición de que solo se suma al KPI del supermercado más barato, es decir, añadirá el producto al recuento únicamente del supermercado más barato y, por lo tanto, la suma del KPI para Mercadona más la de Dia debe dar la suma total calculada en el apartado anterior. Para la creación de esta métrica ha sido necesario crear una columna calculada que nos indique en qué supermercado está el producto más barato, este cálculo es el siguiente:

comprar dia = IF(CALCULATE(sum(Dia[Precio Inicio(€)]), DM_FECHA[ultimaFecha]= "Si") > CALCULATE(SUM(Mercadona[Precio Inicio(€)]), DM_FECHA[ultimaFecha] = "Si"), "No", "Si")

El resultado de ambos KPIs (Mercadona y Dia) junto con el sumatorio total tienen la siguiente forma en la aplicación:

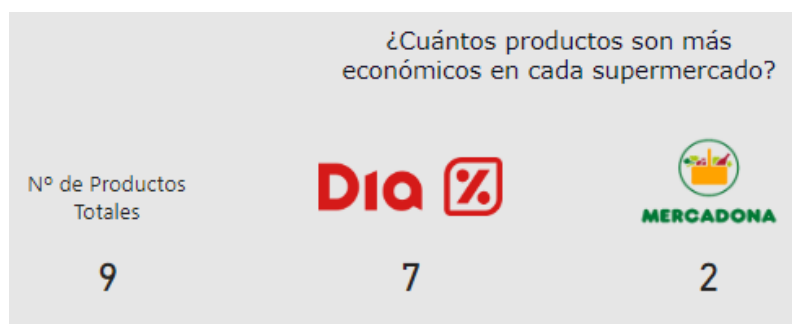


Figura 65. KPIs que muestran el número total de productos seleccionados en la lista de la compra y la suma de los mismos por supermercado en función del precio

- **Predicción a futuro de los precios de Mercadona:** Como valor añadido se ha implementado una gráfica de líneas que muestra por un lado el histórico de datos y precios que se tiene almacenado y, por otro lado, una predicción a futuro de los precios para las próximas 5 semanas y con un 75% de intervalo de confianza.

En este caso, se elimina gran parte de las tablas y nos quedamos únicamente con la de los supermercados y la clave de dimensiones que relaciona ambos establecimientos.

DM_FECHA	<input type="checkbox"/> Inclusión en Preguntas y respuestas
Dia	<input checked="" type="checkbox"/> Inclusión en Preguntas y respuestas
IDs	<input checked="" type="checkbox"/> Inclusión en Preguntas y respuestas
IDs-DIA	<input type="checkbox"/> Inclusión en Preguntas y respuestas
IDs-MERCADONA	<input type="checkbox"/> Inclusión en Preguntas y respuestas
Medidas	<input type="checkbox"/> Inclusión en Preguntas y respuestas
Mercadona	<input checked="" type="checkbox"/> Inclusión en Preguntas y respuestas

Figura 68. Listado de tablas usadas para implementar en el chatbot

- **Sinónimos:** Para un mejor acceso al chatbot se ha necesitado incluir sinónimos de los nombres de columnas para que tenga un mayor contexto y, sea indiferente preguntar por el 'precio' de un producto o por su 'coste'.

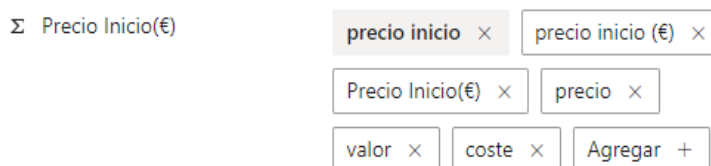


Figura 69. Ejemplo de sinónimos añadidos en el conjunto de datos del chatbot

- **Sugerir preguntas:** Finalmente, se han incluido algunas series de preguntas de ejemplo para que el usuario pueda empezar a 'interactuar' con el chatbot. Estas preguntas también han servido para comprobar el funcionamiento del chatbot y la información que muestra como resultado final.

Reordenación de las preguntas sugeridas

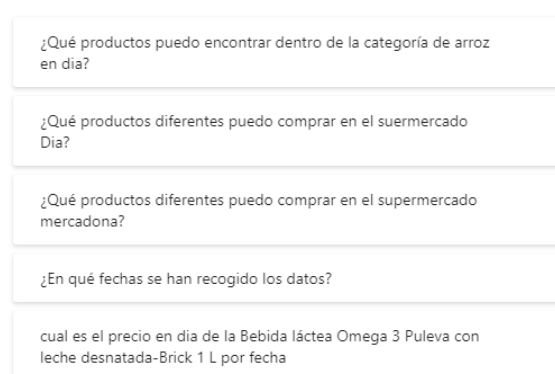


Figura 70. Preguntas recomendadas al usuario para hacer en al chatbot

Con todos estos pasos, se ha creado el chatbot como servicio final:

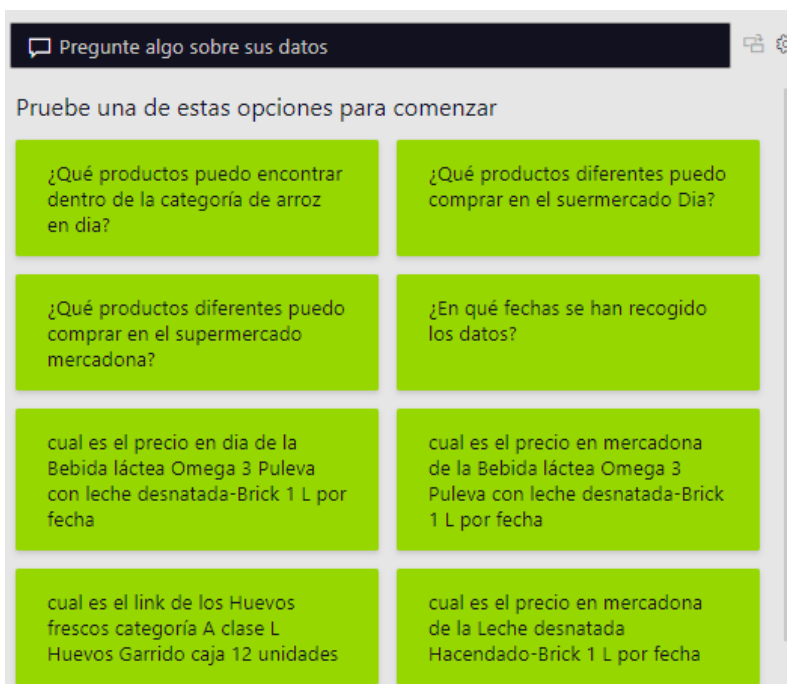


Figura 71 .Resultado final del chatbot

Dentro del tercer grupo, el encargado de mejorar la experiencia del usuario se han realizado los siguientes pasos:

- **Creación de una marca y logo:** Durante el desarrollo del proyecto se determina la necesidad de crear una marca y, por consiguiente, un logo para la aplicación. De ahí surge la idea de crear **SUPER ahorro** un nombre sencillo que pretende dar a conocer al usuario la idea de ahorrar 'mucho' con la de incluir la palabra 'supermercado' acortada. Para el logo se usó el apoyo de la IA a través del asistente Copilot Designer y la tecnología de DALL-E 3. El logo es el siguiente:



Figura 72. Logo de la aplicación y del proyecto SUPER-ahorro

- **Diseño:** Tras seleccionar como colores principales para la aplicación un color casi negro y un gris claro, se procedió a implementar el formato de la aplicación creando el fondo, el lugar para los botones e incluso los recuadros para colocar los visuales.

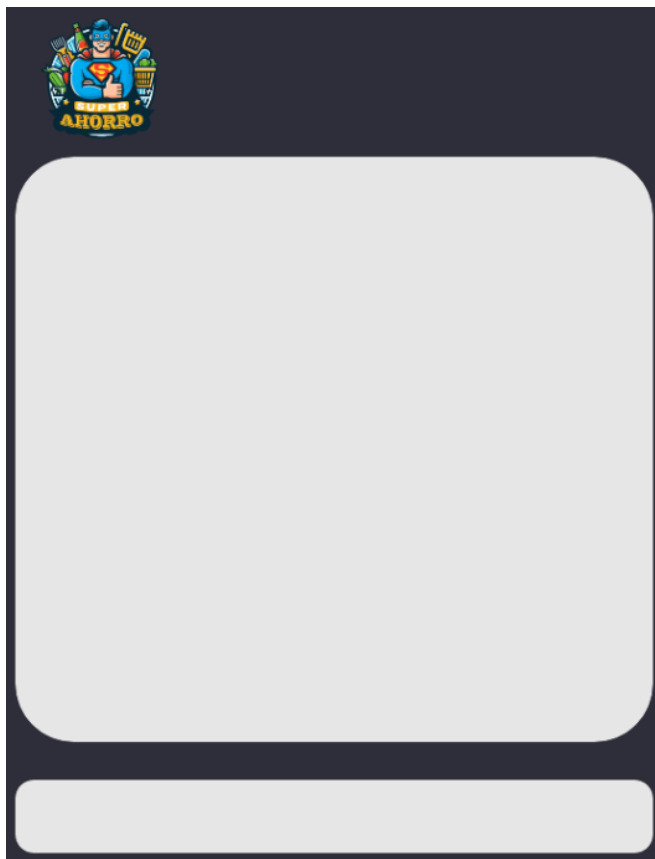


Figura 73. Diseño base implementado para la aplicación

- **Interacción entre páginas:** Dentro de cada página, en la parte inferior se ha creado un espacio con diferentes botones para que el usuario se pueda mover entre páginas. A cada botón se le ha añadido una imagen que simbolice lo que se va a ver en su interior (un carro de la compra para la página del carrito, una hucha con una moneda para la página del ahorro y una gráfica para la página del análisis). Estos botones están conectados entre sí.

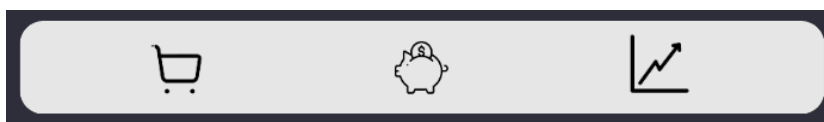


Figura 74. Botones añadidos en la aplicación para interactuar entre páginas

- **Chatbot:** Una vez creado, implementado y entrenado el chatbot se decide utilizar en la parte superior de todas las páginas un acceso al mismo tras hacer click en un botón. Este botón está configurado similar a los anteriores. Adicionalmente se ha tenido que modificar el color del diseño para que siga la misma línea de tonos que la usada para toda la aplicación y al logo.



Figura 75. Botón con acceso directo al chatbot implementado

Por último, este producto final podrá ser consumido desde el servicio de Power BI. Para ello ha sido necesario publicar el informe en el área de trabajo, crear una licencia PRO gratis temporal y, dar acceso a todos aquellos usuarios con el correo terminado en '@live.uem.es'

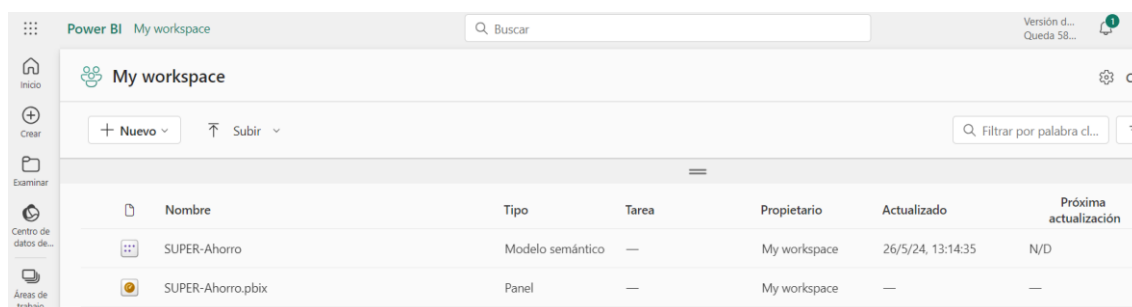


Figura 76. Aplicación subida al servicio de Power BI

Adicionalmente, se puede acceder al servicio de Power BI desde la aplicación de dispositivo móvil e incluso generar un acceso directo a la misma.

3.2.2 Metodología

Durante el desarrollo del proyecto se ha llevado a cabo una metodología Scrum compuesta por una serie de sprints, cada uno de ellos con una duración aproximada de 3 semanas. Al finalizar cada sprint se ha enviado un email al tutor con el estado del proyecto y los avances nuevos que se han desarrollado.

Los diferentes hitos que se han obtenido al finalizar cada uno de los sprints son los siguientes:

Sprint 1	Documentación, análisis y entrega del anteproyecto
Sprint 2	Documentación, requisitos funcionales y no funcionales

	Entorno preparado para el desarrollo
Sprint 3	Fichero de ejemplo del conjunto de datos recopilado para el primer supermercado
	Fichero de ejemplo del conjunto de datos recopilados para el segundo supermercado
	Limpieza de los datos obtenidos por el primer supermercado
Sprint 4	Limpieza de los datos obtenidos por el segundo supermercado
	Mock-up/prototipo de la aplicación
Sprint 5	Primeros análisis y comprobaciones de la calidad del dato
	Creación de Excel relacionando los productos de ambos supermercados
Sprint 6	Creación del modelo de datos y sus relaciones entre diferentes tablas
	Creación de la primera y segunda página del producto final
Sprint 7	Creación de la tercera página del producto final
	Implementación del chatbot
Sprint 8	Entrega de la memoria del TFG

Tabla 9. Resumen de los entregables llevados a cabo en el proyecto usando metodología ágil

Al finalizar cada sprint, se realiza un análisis y se determinan los futuros pasos a realizar. En paralelo, cada semana se estudiaban las tareas realizadas, las futuras y algún posible bloqueo.

Como reflexión final, cabe destacar que, con este formato de metodología se ha podido avanzar todas las semanas en mayor o menor medida y, al implementar una metodología ágil hemos podido observar cómo avanza el servicio final en cada entrega, es decir, en cada sprint.

El proceso llevado a cabo durante el proyecto de una manera visual tiene el siguiente aspecto:

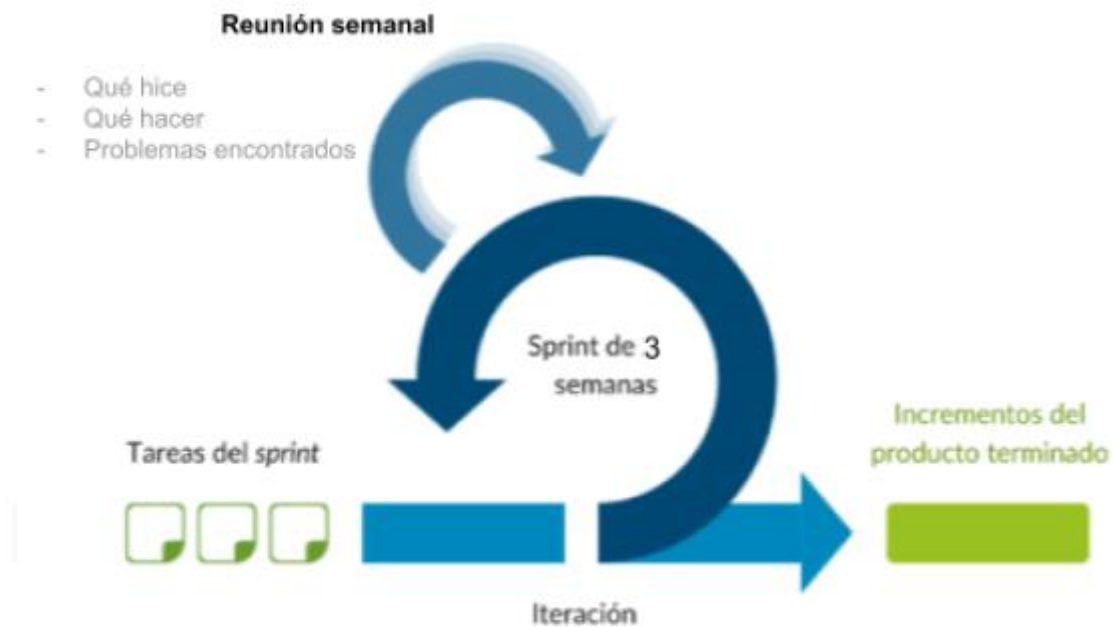


Figura 77. Esquema de la metodología implementada en el proyecto

3.2.3 Herramientas empleadas

Las tecnologías empleadas durante el proyecto son las siguientes:

- Lenguajes de programación
 - o JSON: Formato inicial en el que se obtenían los datos semanalmente
 - o M: Lenguaje con el que se realiza el ETL de los datos
- IDEs
 - o Visual Studio Code
 - o PaseHub
- Control de versiones
 - o Git
- Gestor de paquetes
 - o Pip
- Diseño
 - o Miro
- Visualización de datos
 - o Power BI

3.3 Recursos requeridos

Los recursos requeridos para llevar a cabo este proyecto son principalmente computacionales. Para ello, es necesario:

- **Ordenador:** Con un sistema operativo actualizado de Windows. En este ordenador se ha instalado todo el Software necesario para el desarrollo del proyecto mencionado en el apartado anterior, programas y librerías. También ha sido necesaria la creación de una cuenta de ParseHub para la obtención de los datos.
- **Dispositivo móvil:** Para testear la aplicación generada como producto final para los usuarios con acceso a internet. Las pruebas se han realizado en diferentes teléfonos móviles para verificar su funcionamiento independientemente del sistema operativo del dispositivo.

3.4 Presupuesto

Dentro del presupuesto se ha tenido en cuenta tanto las horas empleadas en el personal encargado de gestionar y desarrollar el proyecto como son el tutor de este y la persona encargada de la implementación del producto. El total de horas empleadas por el personal llega a un total de 470 horas, de las cuales 450 horas son de desarrollo (49 €/hora) y las 20 horas restantes han sido dedicadas por el tutor para la gestión (59 €/hora). Por consiguiente, el coste total del personal asciende a 23.230€.

En cuanto al resto de recursos necesarios, como son el ordenador, monitor, dispositivos telefónicos o material de papelería empleados para el desarrollo del proyecto no han supuesto ningún coste, pero si se hiciera el presupuesto sin tener en cuenta los recursos ya adquiridos, la inversión prevista sería de 1.651,94€. Este valor podría aumentar teniendo en cuenta que el producto implementado tiene algún proceso manual actualmente para evitar costes adicionales. No obstante, en la siguiente tabla se puede observar la información más detallada de cada uno de los recursos:

Tipo de coste	Valor	Comentarios
Horas de trabajo en el proyecto		
Horas personales de trabajo	450 horas	Cada hora tiene un coste de 49€
Horas del tutor	20 horas	Cada hora tiene un coste de 59€
Equipo técnico utilizado		
Ordenador portátil – HP 15s-fq5059ns	649 €	Recurso ya adquirido, precio actual en el mercado.

Monitor LG	151,05 €	Recurso ya adquirido, precio actual en el mercado.
Dispositivo telefónico – Xiaomi Redmi 13C	132,89 €	Recurso ya adquirido, precio actual en el mercado.
Dispositivo telefónico – iPhone 14	699 €	Recurso ya adquirido, precio actual en el mercado.
Software utilizado		
Visual Studio Code	0 €	No requiere pago por licencia
Power BI	0 €	No requiere pago por licencia
GitHub	0 €	No requiere pago por licencia
Librerías	0 €	No requiere pago por licencia
Estudios e informes		
Estudios e informes	0 €	Todos los recursos utilizados (informes de consulta o revistas de investigación) eran gratuitos
Materiales empleados		
Material empleado	20€	Gasto en papelería (bolígrafos, folios e impresión de la memoria)

Tabla 10. Presupuesto total del proyecto desglosado

3.5 Viabilidad

Teniendo en cuenta el coste total de la creación del producto y que hasta el momento los únicos gastos son los del personal porque el material ya estaba adquirido previamente, se puede afirmar que la viabilidad es positiva ya que el resultado final es un producto que puede ayudar al usuario y, mediante diferentes enfoques se puede sacar tanto económicamente como a nivel información valores positivos.

Los diferentes enfoques en los que se puede generar valor con dicha aplicación son:

- **Anuncios:** Un enfoque muy común y utilizado enfocado en la amortización del importe invertido durante el desarrollo del proyecto con la incorporación de anuncios publicitarios enfocados en el tema del producto.

- **Supermercados:** En la actualidad, la aplicación está compuesta de únicamente dos supermercados (Mercadona y Día), a futuro se puede valorar la opción de ir incorporando nuevos supermercados para que los usuarios puedan tenerlo en cuenta de cara a su compra y, valorar la opción de cobrar a los supermercados.
- **Datos:** Durante la navegación del usuario por la aplicación, este va a ir aportando información que es aconsejable almacenar en una tabla de histórico, principalmente el listado de todos los productos seleccionados en cada inicio de sesión por cada usuario, esto puede ayudar a conocer los hábitos de consumo de cada persona o aquellos productos más recurrentes por los españoles. Con este detalle, se pueden tomar decisiones como la elección de los anuncios e incluso personalizarlos en función a los gustos.

A futuro y, según vaya ampliándose el uso de la aplicación será necesario migrar los procesos a una herramienta en la nube como puede ser Azure, Amazon o Google. Este cambio supondrá un incremento en el gasto fijo mensual de costes a tener en cuenta.

Otro punto que puede aportar viabilidad a la aplicación es, a medida que vayan aumentando los usuarios, valorar la opción de crear un nuevo apartado de usuarios premium con información adicional y valiosa de los supermercados como pueden ser las ofertas de la semana o consejos en función a sus compras semanales y que, en este caso, los usuarios tengan un coste mensual por suscripción.

Por último, se deberá valorar la inversión en un proyecto de captación de nuevos usuarios y fidelización de los ya existentes. Esta propuesta requiere una gran inversión y se deberá analizar más en detalle para revisar su viabilidad real.

A modo de conclusión, actualmente la aplicación es viable por el bajo coste de inversión que ha tenido y se pretende recuperar a medio plazo, pero hay que tener en cuenta que a futuro y, para que el producto continúe creciendo, es necesario aumentar la inversión.

3.6 Resultados del proyecto

Como resultado final al desarrollo, se ha logrado crear una aplicación con información útil para los usuarios sobre los precios de los productos en diferentes supermercados incluyendo en la misma gráficas temporales.

Durante el desarrollo del proyecto se ha tenido que modificar la planificación planteada en un primer momento debido al cambio de fechas en las entregas del proyecto y, también, por un cambio de enfoque de la aplicación ya que en una primera instancia se pretendía crear la aplicación de un modo diferente al final y, el tiempo estimado para dedicar a cada una de las tareas se vio modificado.

Pese a los imprevistos encontrados principalmente con la disminución de tiempo y, teniendo en cuenta la limitación de tiempo dedicada al desarrollo al compaginarlo con el trabajo se ha conseguido llegar a la entrega prevista con todos los apartados planificados y cumpliendo con el alcance.

Durante todo el proceso, se han realizado diferentes pruebas para validar el producto final. Estas pruebas han estado enfocadas principalmente en la aplicación, comprobando que tanto el registro como el inicio de sesión funcionan correctamente así como que los datos que se muestran al usuario son correctos.

Las diferentes pruebas que se han llevado a cabo para comprobar que el producto final funciona correctamente y muestra bien los datos son:

- Comprobar la conexión con diferentes usuarios de la universidad europea.
- Comprobar que si no tienes usuario con permisos no puedes acceder al sistema.
- Comprobar que se pueden seleccionar y borrar los diferentes productos en la primera página de la aplicación.
- Comprobar la relación entre las tres pestañas diferentes sobre las que se compone la aplicación.
- Verificar que los valores filtrados en la primera página se filtran en las demás.
- Verificar que las cifras de precios son las correctas y que muestra el campo adecuado.
- Comprobar que se puede ver la aplicación desde un dispositivo móvil.
- Comprobar que se puede acceder desde diferentes dispositivos.
- Verificar que el formato de los visuales se adecúa al de cada dispositivo móvil.
- Comprobar la velocidad de respuesta.

Todas estas pruebas se han realizado en diferentes dispositivos electrónicos para verificar que funciona correctamente en diferentes sistemas operativos.

Capítulo 4. DISCUSIÓN

La creación de este proyecto viene dada con la idea de ayudar a los usuarios de la aplicación a reducir el tiempo dedicado a disminuir el coste de la cesta de la compra. Para ello, se propone la idea de comparar el coste de los diferentes productos que se pueden encontrar en distintos supermercados en una única aplicación, pero la planificación y el alcance presentados en primera instancia ha surgido una serie de modificaciones.

Con el cambio en las fechas de entrega del proyecto hubo que modificar el enfoque de la aplicación, cambiando la tecnología que se tenía planificada utilizar para su implementación y, por consiguiente, el enfoque del chatbot que estaba planificado también. En primera instancia, se pretendía crear la aplicación con angular, una herramienta menos usada, menos conocida y más compleja para mi, pero la limitación de tiempo hizo que se cambiara a Power BI tanto la creación del producto final como la creación, implementación y entreno del chatbot.

Además, el tiempo dedicado en la obtención, almacenamiento y limpieza del conjunto de datos fue superior al estimado en un primer momento, pero teniendo en cuenta que es muy importante cerciorarnos de que los datos que se muestran en la aplicación son correctos se hizo necesario aumentar el tiempo destinado para este tema.

Por otro lado, con la limitación de utilizar versiones de prueba o gratuitas hace que algunos de los proyectos implementados dejen de ser automáticos como el caso de la obtención de los datos semanalmente o la publicación de la aplicación y el acceso a todo el público.

En relación a la metodología utilizada, se propuso una metodología ágil con entregas y reuniones tras la finalización de cada sprint, pero, aunque se ha cumplido con las entregas, las revisiones y reuniones que hacía con carácter semanal o al finalizar cada sprint eran individuales.

Capítulo 5. CONCLUSIONES

5.1 Conclusiones del trabajo

Con las subidas de precios de los productos en los supermercados, la gente ha comenzado a buscar soluciones para intentar ahorrar lo máximo posible. La creación de esta aplicación pretende agrupar toda la información necesaria para facilitar la toma de decisiones en un solo servicio.

La idea principal de SUPER-ahorro, la aplicación creada como producto final, es encontrar el supermercado más económico para los artículos añadidos en la lista de la compra del usuario y, a su vez, facilitar esta información en un único producto en lugar de tenerlo que hacer buscando individualmente en la web oficial de cada establecimiento. También, se pretende mostrar la evolución de sus precios y una estimación a futuro del sentido del coste de los productos.

Adicionalmente, durante la realización del trabajo han surgido diferentes problemas que han hecho cambiar las fechas del proyecto como son el tiempo de entrega o la dificultad de algunas tecnologías, pero todos ellos han sido solventados con un cambio de planificación e incluso con el cambio de ciertas tecnologías no utilizadas anteriormente por otras más conocidas.

5.2 Conclusiones personales

La realización de este proyecto ha sido un reto satisfactorio pues es un trabajo que engloba muchos puntos diferentes aprendidos a lo largo de la carrera. El resultado final es positivo ya que he sido capaz de completar las diferentes actividades que comprenden el alcance del proyecto.

Hasta la actualidad, en la mayoría de los trabajos realizados es inevitable centrarse más en una tarea, pero cuando una persona se debe enfrentar sola a pensar una idea de proyecto, sus posibles soluciones, solventar los errores y cumplir con los tiempos de planificación es cuando se ve la complejidad. Haber sido capaz de llegar a obtener un producto final cumpliendo los objetivos previstos es el momento más gratificante de la carrera.

En la actualidad, dedico parte de mi tiempo a comparar el precio de los productos de manera manual, entrando a cada una de las páginas webs de los distintos supermercados, buscando y comparando artículos para tratar de ahorrar.

El TFG es el trabajo más importante de la carrera y mi idea principal era realizarlo sobre algo que pueda ayudarme a mí y, por supuesto, al resto de usuarios que consuman la aplicación a solventar los problemas rudimentarios de comparación de artículos uno a uno, con la consiguiente pérdida de tiempo personal.

Capítulo 6. FUTURAS LÍNEAS DE TRABAJO

Existen diferentes líneas de trabajo en las que se puede ampliar el desarrollo del proyecto y que principalmente por tiempo no se ha podido implementar. Algunas de las ideas son:

- **Ampliar el conjunto de datos:** En la actualidad, únicamente se recoge información de dos supermercados (Mercadona y Dia) con una periodicidad mensual (todos los lunes). Sería muy positivo ampliar el conjunto de datos en primer lugar para más supermercados y, secundariamente, con una periodicidad mayor a la actual.
- **Chatbot:** Actualmente, el funcionamiento del chatbot está limitado a los sinónimos que se han incluido manualmente de los campos y a escribir el nombre de los productos literalmente igual que se encuentran en el modelo de datos. Para mejorar la experiencia del usuario sería aconsejable dotarlo de una más amplia alimentación de palabras y situaciones.
- **Implementar procesamiento del lenguaje natural (PLN):** El nombre de los diferentes productos en la actualidad esté almacenado al completo y con las mismas palabras que se encuentran en la página web. A futuro y, enfocado en la aplicación, sería aconsejable usar el procesamiento del lenguaje para que no haya problema si el usuario escribe palabras similares a las usadas en la descripción original (León, 2020).

Capítulo 7. REFERENCIAS

- Pelayo, F. (2023, 28 septiembre). La cesta de la compra pega un subidón de un 30,8% en los últimos dos años que ha afectado a 9 de cada 10 de los productos del Súper. [www.20minutos.es- Últimas Noticias](https://www.20minutos.es/noticia/5176655/0/-cesta-compra-peg-un-subidon-un-30-8-los-ultimos-dos-anos-que-ha-afectado-9-cada-10-los-productos-super/). <https://www.20minutos.es/noticia/5176655/0/-cesta-compra-peg-un-subidon-un-30-8-los-ultimos-dos-anos-que-ha-afectado-9-cada-10-los-productos-super/>
- Statista. (2023, 30 mayo). Ranking de los alimentos con mayor subida de los precios en España en 2023. <https://es.statista.com/estadisticas/1379186/ranking-de-los-alimentos-con-mayor-subida-de-los-precios-en-espana/>
- Vargas, R. L. (2023, 20 junio). Los españoles ya compran más del 50% de los productos del supermercado de marca blanca. La Razón. https://www.larazon.es/economia/espanoles-compran-mas-50-productos-supermercado-marca-blanca_202306206490b0fee9f5530001c5b045.html
- Agencias. (2024, 4 abril). Los precios de supermercados han subido “de media” un 38% en tres años, según la OCU. La Vanguardia. <https://www.lavanguardia.com/vida/20240404/9587664/precios-supermercados-han-subido-media-38-tres-anos-ocu-pmv.html>
- Los precios de alimentos siguen al alza. (2022, 14 diciembre). [www.ocu.org](https://www.ocu.org/consumo-familia/supermercados/noticias/subida-precios-alimentacion). <https://www.ocu.org/consumo-familia/supermercados/noticias/subida-precios-alimentacion>
- Los supermercados más baratos de 2023. (2023, 28 septiembre). [www.ocu.org](https://www.ocu.org/consumo-familia/supermercados/noticias/supermercados-mas-baratos-2023). <https://www.ocu.org/consumo-familia/supermercados/noticias/supermercados-mas-baratos-2023>
- Wikipedia, (2023). Anexo: Municipios de España por población. https://es.wikipedia.org/wiki/Anexo:Municipios_de_Espa%C3%B1a_por_poblaci%C3%B3n
- Tobar, S. (2023, 3 julio). Un ‘súper’ por cada 1.886 habitantes: radiografía de las cadenas que resisten en la España vaciada. El Español. https://www.elespanol.com/invertia/empresas/distribucion/20230703/super-habitantes-radiografia-cadenas-resisten-espana-vaciada/775422704_0.html
- Minutos. (2022, 12 noviembre). El mapa de los supermercados en cada comunidad: cuál es el preferido en tu zona y de cuál hay más. [www.20minutos.es - Últimas Noticias](https://www.20minutos.es/noticia/5075924/0/mapa-supermercados-comunidades-autonomas-preferido-hay-mas/). <https://www.20minutos.es/noticia/5075924/0/mapa-supermercados-comunidades-autonomas-preferido-hay-mas/>
- Fernández, Y. (2019, 25 mayo). 17 aplicaciones para organizar y gestionar la lista de la compra. Xataka. <https://www.xataka.com/basics/17-aplicaciones-para-organizar-gestionar-lista-compra>
- Capterra, (2022). Software parsehub. <https://www.capterra.es/software/145965/parsehub>

- Nieto, C. P. (2018, 19 diciembre). Qué son las tablas de hechos y de dimensión. Carlos Pesquera Nieto. <https://carlospesquera.com/que-son-las-tablas-de-hechos-y-de-dimension/>
- León, E. (2020, 16 diciembre). Procesamiento del lenguaje natural (PLN) con Python - BAOSS. BAOSS. <https://www.baoss.es/procesamiento-del-lenguaje-natural-pln-con-python/>

Enlaces genéricos necesarios para el desarrollo:

- **Extracción conjunto de datos**
 - Parsehub (2020). Help Parsehub. <https://help.parsehub.com/hc/en-us/articles/217753378-Select>
 - Mercadona compra online. (s. f.). <https://tienda.mercadona.es/categories>
 - Supermercado online | ¡Recibe tu compra hoy mismo! | Día. (s. f.). <https://www.dia.es/>
- **Limpieza de los datos**
 - Rellenar una fórmula hacia abajo en celdas adyacentes - Soporte técnico de Microsoft. (s. f.). <https://support.microsoft.com/es-es/office/rellenar-una-f%C3%B3rmula-hacia-abajo-en-celdas-adyacentes-041edfe2-05bc-40e6-b933-ef48c3f308c6#:~:text=Haga%20clic%20en%20Inicio%20%3E%20Rellenar,%2C%20Derecha%2C%20Arriba%20o%20Izquierda.>
 - Dhania. (2021, 28 noviembre). APPEND y MERGE: ¿Qué opción para qué uso? Biist Pro. https://biist.pro/___trashed-2
 - Aprende Power BI con Salvador Ramos. (2022, 8 diciembre). Deshabilitar carga en consultas en Power Query [Vídeo]. YouTube. <https://www.youtube.com/watch?v=tpTVRbEu87o>
 - De un Autodidacta, B. (2022, 4 enero). 4 formas de crear una columna condicional en Power BI (DAX y Power Query). Medium. <https://bitacordeunautodidacta.medium.com/4-formas-de-crear-una-columna-condicional-en-power-bi-dax-y-power-query-6fddbb1f5706>
- **Aplicación**
 - Lizana, M. (2023, 5 septiembre). Análisis predictivo en Power BI. Forecast. - Mar Lizana - Medium. Medium. <https://medium.com/@akanemar/an%C3%A1lisis-predictivo-en-power-bi-forecast-3f37dbc47021>
- **Chatbot**
 - maggiesMSFT. (2023, 11 octubre). Usar Preguntas y respuestas de Power BI para explorar y crear objetos visuales en los informes - Power BI. Microsoft Learn. <https://learn.microsoft.com/es-es/power-bi/natural-language/power-bi-tutorial-q-and-a>

- maggiesMSFT. (2024, 11 enero). Enseñanza de Preguntas y respuestas para comprender las preguntas y los términos de Preguntas y respuestas de Power BI - Power BI. Microsoft Learn. <https://learn.microsoft.com/es-es/power-bi/natural-language/q-and-a-tooling-teach-q-and-a>
- Pivotalstats. (2023, 13 febrero). AI Visualization - Implement CHATBOT quickly using Q&A visual | Power BI [Vídeo]. YouTube. <https://www.youtube.com/watch?v=c7icpgEEkqE>

Capítulo 8. ANEXOS

Anexo 1. Código fuente

El código que se ha ido generando durante el desarrollo del proyecto junto con el conjunto de datos o la aplicación está disponible en un repositorio público de GitHub el cual se puede acceder con el siguiente enlace: <https://github.com/sofiaCorral/TFG-INFO>. Dentro del mismo se pueden encontrar diferentes carpetas que dividen el contenido en bloques. Estos son:

- **Scraping:** En su interior se puede encontrar una carpeta denominada ParseHub donde se localizan los desarrollos llevados a cabo para la obtención de los datos, otra carpeta con los datos obtenidos en primera instancia en formato JSON (Datos limpios JSON), otra carpeta con esos datos una vez convertidos en CSV (Datos CSV), otra con los datos limpios tras el proceso realizado en Power BI (datos limpios CSV) y un fichero con la relación entre los productos de ambos supermercados.
- **Visualización:** En este lugar se encuentra el fichero de Power BI encargado de hacer la limpieza del conjunto de datos y estudiar la calidad del dato obtenido.
- **Aplicación:** Por un lado, existe una carpeta denominada logos con las imágenes usadas en la aplicación y, por otro lado el fichero de Power BI con los visuales de la aplicación, es decir, el que se ha subido al servicio para que los usuarios utilicen.
- **Documentación:** Se adjuntan todos los documentos generados durante el desarrollo del proyecto.

