

Importacion

Morales Martorell, Josep Vicent, Rocio Bono Moreno, Laura Horjales Rivas, Arnau Hernández Lucas y I

2025-04-03

Contents

| | | |
|-----|-----------------------------|---|
| 0.1 | 1.Librerías | 1 |
| 0.2 | 2. Carga Pdfs | 3 |
| 0.3 | 3. Transformación | 3 |
| 0.4 | 4 PREGUNTAS | 4 |

0.1 1.Librerías

Conjunto de librerías a utilizar:

```
knitr::opts_chunk$set(echo = FALSE, warning = FALSE, message = FALSE)
library(pdftools)
```

```
## Warning: package 'pdftools' was built under R version 4.4.3
```

```
## Using poppler version 23.08.0
```

```
library(tidyr)
library(dplyr)
```

```
##
## Adjuntando el paquete: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##   filter, lag
```

```
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
library(ggplot2)
library(stringr)
library(lubridate)
```

```
## Warning: package 'lubridate' was built under R version 4.4.3
```

```
##
```

```
## Adjuntando el paquete: 'lubridate'
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      date, intersect, setdiff, union
```

```
library(knitr)
```

```
library(kableExtra)
```

```
## Warning: package 'kableExtra' was built under R version 4.4.3
```

```
##
```

```
## Adjuntando el paquete: 'kableExtra'
```

```
## The following object is masked from 'package:dplyr':
```

```
##
```

```
##      group_rows
```

```
# Configurar nombres en español para gráficos
```

```
options(OutDec = ",", digits = 2)
```

Mini Proyecto de Análisis Exploratorio de Datos:

Grupo: L

Participantes: Josep Vicent Morales Martorell, Rocio Bono Moreno, Laura Horjales Rivas, Arnau Hernández Lucas y Pau García Pérez.

INTRODUCCIÓN:

CONTEXTO DEL PROYECTO

En la actualidad donde la información ha pasado a ser uno de los elementos más preciados, el análisis de datos juega un papel muy importante tanto en la toma de decisiones estratégicas como operativas. La gran cantidad de datos ha impulsado la necesidad de convertir estos mismos en conocimientos aplicables, lo que convierte el análisis exploratorio en una herramienta esencial en múltiples sectores.

Como futuros científicos de datos, nuestro trabajo consistirá en extraer la información más relevante de una gran cantidad de datos. En este proyecto, los tickets representaran nuestra fuente de datos, donde mediante estos, podremos descubrir: preferencias, hábitos y comportamientos de los consumidores. A través de ellos, también podemos interpretar patrones ocultos, prever necesidades futuras y generar estrategias que optimicen tanto la experiencia del cliente como la gestión empresarial.

Analizando los registros de venta recogidos en los tickets, es posible obtener conclusiones relevantes para áreas como el control de inventario, el diseño de promociones o la personalización de servicios. Esto no solo mejora la eficiencia interna, sino que permite responder de forma ágil y efectiva a las demandas del mercado.

NUESTRO PROYECTO

Este proyecto nace con la necesidad de convertir datos sin procesar - provenientes de tickets de Mercadona- en información valiosa mediante técnicas de limpieza, transformación y análisis exploratorio. A través de una serie de scripts en R, hemos diseñado un modelo de trabajo que permite capturar los aspectos más relevantes del consumo cotidiano, con el objetivo de ofrecer herramientas útiles para la toma de decisiones comerciales.

La metodología utilizada combina la programación en R con recursos de visualización y manipulación de datos. Nuestro trabajo no se limita en organizar los datos, sino que busca identificar correlaciones, tendencias y patrones de comportamiento. A través de este proceso, contribuimos al entendimiento de dinámicas reales del mercado y fomentamos un entorno donde el dato procesado se convierte en una ventaja competitiva.

RECURSOS UTILIZADOS Y ESTRUCTURA DE TRABAJO

Para llevar a cabo este análisis, hemos empleado librerías fundamentales de R como readr, dplyr, stringr, tibble y ggplot2, entre otras. Estas herramientas nos han permitido estructurar los datos de manera eficiente, realizar filtrados específicos, transformar cadenas de texto y crear visualizaciones útiles.

El proceso inicial incluyó la conversión de los archivos PDF a texto plano mediante el uso de las funciones de la librería pdftools. Luego, se procedió a extraer la información de interés, organizándola en dataframes preparados para su análisis posterior. Una parte esencial fue la limpieza de los datos: corregir errores, homogeneizar formatos y validar los campos esenciales para obtener un dataset sólido y confiable.

VARIABLES Y ESTRUCTURA DE LOS DATOS

En esta primera fase, trabajamos con un conjunto de diferentes variables, que abarcan tanto información general del ticket (fecha, hora, importe total, tienda, caja, número de ticket) como detalles específicos de los productos adquiridos (nombre del producto, cantidad, precio unitario, tipo de producto, peso o unidad, entre otros).

0.2 2. Carga Pdfs

Los tickets que queremos analizar se encuentran en formato .pdf, y se encuentran en la carpeta data, por tanto en este bloque importamos todos los .pdf de la carpeta.

0.3 3. Transformación

En este apartado realizamos el tratamiento de los datos para dividirlos y reorganizarlos para posteriormente poder consultarlos de manera más fácil y eficiente.

Como hemos observado, en la carpeta data se incluyen tickets que no son de Mercadona o directamente no son tickets por lo que antes de proceder con el transformado, vamos a eliminar estos tickets erróneos.

0.3.1 3.1 Tabla Precios

En la Figura se puede apreciar qué estructura sigue la tabla de manera genérica.

Como las comidas que van por peso como verduras frutas o pescados, se organizan de otra manera en los tickets, mostrándose en una línea la cantidad, y el nombre del elemento, y en la de abajo el peso del elemento en Kg, seguido del precio en Kg/€ y finalmente, a la derecha, el precio pagado. Se ha creado esta función para poder implementarlo en nuestra tabla.

Extraemos por completo los datos referentes a los elementos comprados precios y unidades, y se transforma en una tabla esto se realiza para todos los distintos tickets. Extrae tabla (en pruebas)

Extracción de tabla del pdf.

Transformación datos en tabla

Estos datos presentan algunos problemas, entre ellos, los decimales están separados por “,” en vez de por “.”, y a su vez, las tablas no están llenas de datos, ya que algunos datos del Importe salen desplazados a la columna de Precio debido a que esta está vacía cuando la cantidad de unidades es 1, ya que es el mismo valor.

Por tanto, en el siguiente bloque arreglaremos estos 2 problemas.

| | Descripción | P. Unit | Importe |
|---|----------------------|---------|---------|
| 1 | CAFE MOLIDO NATURAL | | 3,20 |
| 1 | LAVAVAJILLAS ULTRA | | 1,90 |
| 1 | CEBOLLA 2 KG | | 2,79 |
| 1 | CACAHUETE SIN SAL | | 1,60 |
| 1 | NUEZ CASCARA NATURAL | | 2,90 |
| 1 | 40 B. CONG. MEDIANAS | | 0,85 |
| 1 | BOLSAS.G CONGELADO | | 1,10 |
| 1 | PISTACHO TOST 0% SAL | | 3,35 |
| 1 | AVELLANA TOSTADA | | 3,10 |
| 2 | PAÑUELO CAJA | 1,30 | 2,60 |
| 1 | NARANJA 5 KG. | | 5,95 |

Figure 1: Figura 1: Vista de productos

Ahora para una consulta más fácil, unimos todos los tickets en una misma tabla y usamos la columna index (índice) para poder diferenciarlos entre ellos con mayor facilidad.

0.3.2 3.2 Datos Generales

Aquí vamos a tratar de recoger y colocar en una tabla la información general de los tickets como dirección, teléfono... entre otros.

Extracción bloque cabecera:

En este bloque organizamos toda la información de la cabecera de los tiquets en un único dataframe

Con este bloque extraemos de cada pdf el precio total pagado en euros y lo añadimos al dataframe anterior de datos generales.

0.3.3 3.3 Tabla IVA

En este punto extraemos las tablas de los distintos IVA aplicados en los tickets.

Con este bloque extraemos la tabla entera desde el ticket colocándolas en una lista que referencia el ticket.

En este bloque, colocamos todas las tablas en un único data frame con una variable de índice para poder saber a que tiquet hacen referencia y los enlaza con los otros data frames que hemos creado anteriormente.

Guardamos todos los data frames en carpeta data

0.4 4 PREGUNTAS

En la siguietne parte vamos a resolver las siguientes preguntas propuestas por el profesor:

- ¿Cuáles son los 5 productos, de los vendidos por unidades, con más ventas ?- ¿Cuántas unidades de cada uno se han vendido ?

| Descripcion | Total_Unidades_Vendidas |
|---------------------|-------------------------|
| ATUN CLARO OLIVA | 62 |
| BOLSA PLASTICO | 50 |
| QUESO LONCHAS CABRA | 43 |
| YOGUR COCO | 40 |
| PAN SEMILLAS | 38 |

Como podemos observar los 5 productos más vendidos por unidades son: el atún claro oliva(62 unidades), queso lonchas cabra(53 unidades), bolsa plástico (51 unidades), leche desnatada de calcio (49 unidades) y por último yogur coco (40 unidades), esto muestra una gran tendencia hacia los lácteos ya que 3 de estos 5 productos son lácteos.

- Si consideramos la categoría de FRUTAS Y VERDURAS. Cuáles son los 5 productos más vendidos ?
¿Cuántos kilos se han vendido de cada uno de estos productos ?

| Descripcion | total_kilos_vendidos |
|----------------------|----------------------|
| PLATANO | 58 |
| BANANA | 26 |
| SANDIA BAJA SEMILLAS | 23 |
| PEPINO | 20 |
| MELON PIEL SAPO | 16 |

Como se observa en esta tabla los 5 productos más vendidos de frutas o verduras son: el plátano (vendiendo un total de 62.868 kg), la banana (vendiendo 28.140 kg), la sandia baja semillas (22.843 kg), pepino (19.624 kg) y el calabacín verde (17.754 kg).

- Si consideramos la categoría de PESCADO. Cuáles son los 5 productos más vendidos ? ¿Cuántos kilos se han vendido de cada uno de estos productos ?

| Descripcion | Peso_Total | Cantidad |
|----------------------|------------|----------|
| LENGUADO FRESC EUROP | 1,2 | 3 |
| SEPIA FRESCA | 1,6 | 3 |
| SEPIA LONJA | 2,0 | 3 |

- Muestra mediante un gráfico de líneas como ha variado el precio por kilo de las bananas y los plátanos en los tickets disponibles, a lo largo del tiempo.

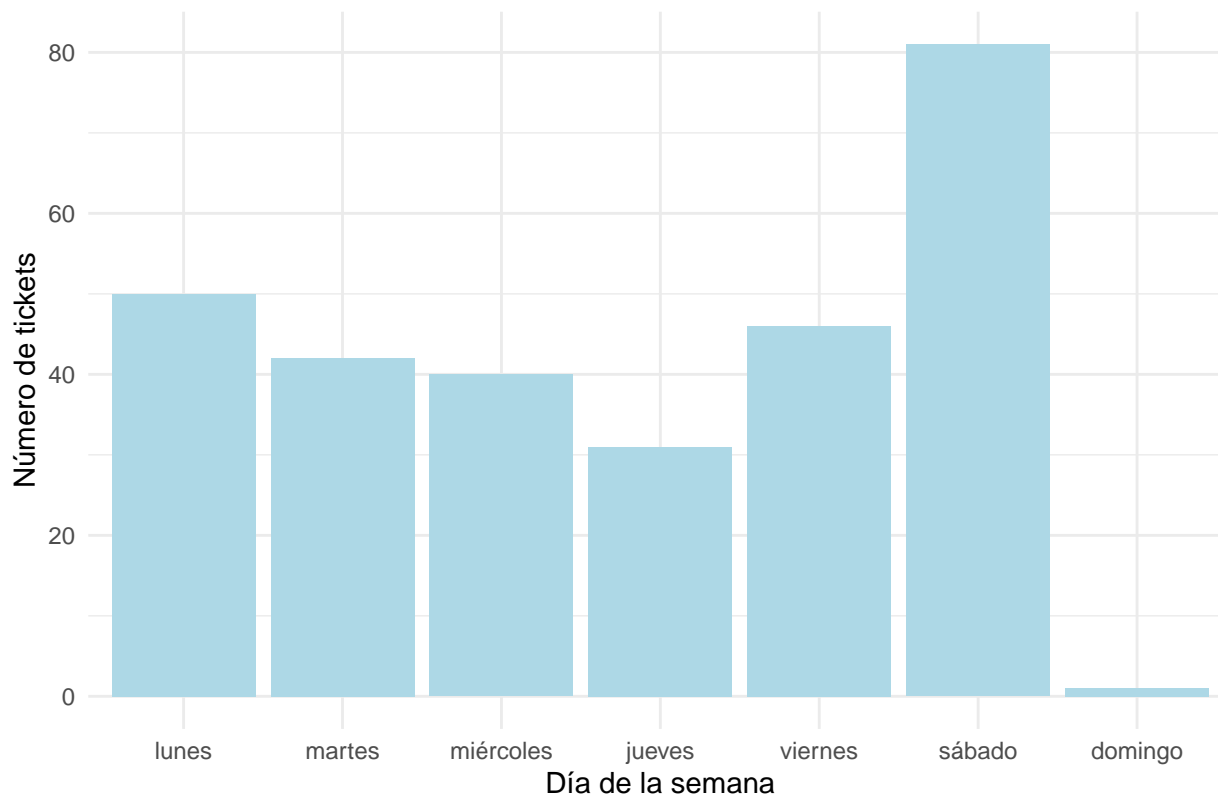
Para ello primero sacaremos todas las veces que se ha comprado bananas y platanos(como necesitamos que este ordenado por la fecha haremos un inner_join)

| Ciudad | Compras |
|--------------------------|---------|
| VALENCIA | 120 |
| ALBORAIA/ALBORAYA | 50 |
| BURJASSOT | 27 |
| MURO DE ALCOY | 24 |
| ALCOI/ALCOY | 22 |
| BUÑOL | 12 |
| BENIJOFAR | 8 |
| ALGINET | 5 |
| MONCADA | 5 |
| BENIFAIO | 4 |
| GARRUCHA | 4 |
| VERA | 3 |
| GANDIA | 1 |
| ONTINYENT | 1 |
| REQUENA | 1 |
| SAGUNT/SAGUNTO | 1 |
| SAN ANTONIO DE BENAGEBER | 1 |
| SANT JOSEP DE SA TALAIA | 1 |
| SANTA EULÀRIA DES RIU | 1 |

Como era previsible al ser los tickets de los alumnos, era mucha más probable que estos fueran de Valencia o de alrededores, por lo que hay 120 de Valencia y a continuación Alboraya y Burjassot.

- Muestra mediante un diagrama el número de tickets recogidos cada día de la semana. ¿Si tuvieses que cerrar un día entre semana qué día lo harías ?

Número de tickets por día de la semana



| | index | Direccion | Ciudad | CP | Telefono | Fecha | Hora | OP | Num_Tienda |
|------|-------|-----------------|-----------|-------|-----------|------------|-------|--------|------------|
| TP.7 | 8 | C/ VALENCIA S/N | BENIJOFAR | 03178 | 966713008 | 2023-12-24 | 13:41 | 211165 | 3095 |

Según el diagrama el día entre semana con menos tickets y por lo tanto ventas, es el jueves. Por lo que lo más inteligente sería cerrar este día, ya que es cuando menos gente va a Mercadona.

También se responderan a las siguientes preguntas propuestas por nuestro equipo:

- ¿Cuál es la hora más habitual para realizar la compra? ¿Este horario varía entre los días laborales y los fines de semana?

| Tipo_dia | Hora_hora | N_compras |
|---------------|-----------|-----------|
| Fin de semana | 13 | 20 |
| Laboral | 20 | 48 |

Como podemos comprobar el fin de semana la hora con más compras (21) es a la 13 a diferencia que los días laborales que es a las 20 con 48 tickets, esto puede darse, debido a que entre semana tenemos obligaciones y extraescolares que atender y por eso las 20 es más frecuente mientras que los fines de semana aprovechamos las mañanas para poder comprar y tener unas tardes más relajadas.

- ¿Existe alguna diferencia clara en el perfil de compra entre los días de semana y los fines de semana?

| Tipo_dia | Descripcion | Frecuencia |
|----------|---------------------|------------|
| Laboral | PLATANO | 50 |
| Laboral | QUESO LONCHAS CABRA | 32 |
| Laboral | BOLSA PLASTICO | 31 |
| Laboral | FILETE PECHUGA | 25 |
| Laboral | UVA BLANCA S/SEM | 24 |

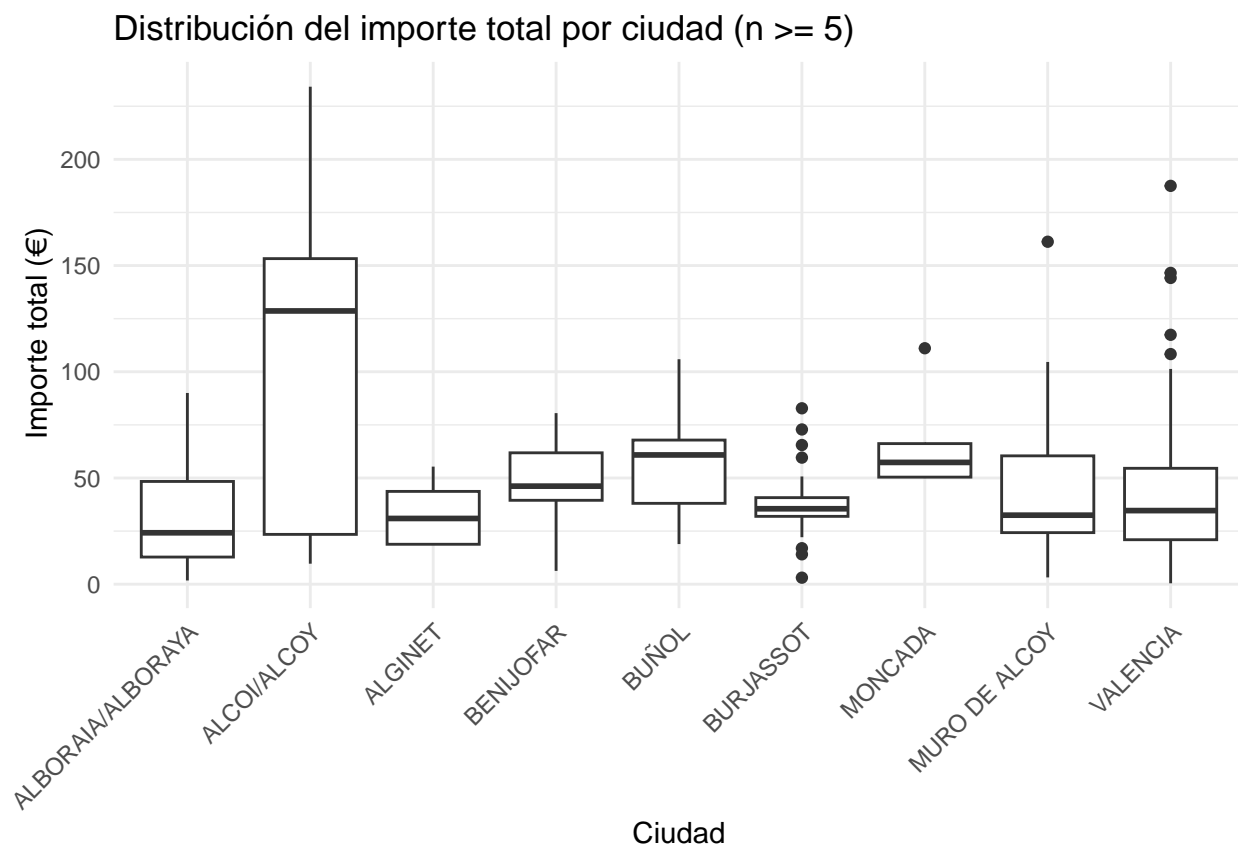
| Tipo_dia | Descripcion | Frecuencia |
|---------------|----------------------|------------|
| Fin de semana | PEPINO | 18 |
| Fin de semana | ALMENDRA NATURAL | 16 |
| Fin de semana | CACAHUETE CHOCOLATE | 15 |
| Fin de semana | KIWI VERDE | 15 |
| Fin de semana | Z.NARANJA 330ML REF. | 14 |

Durante los fines de semana, los productos más comprados tienden a ser snacks, frutos secos y frutas como el pepino, kiwi verde y el zumo de naranja. En cambio, los días laborales, los productos más vendidos son elementos básicos como plátanos, queso en lonchas o carne, la gran compra de bolsas nos indica que entre semana se realizan compras más grande, por eso, su gran cantidad de venta, mientras que los fines de semana es una compra más orientada al consumo inmediato o pequeños caprichos.

- ¿El precio total de la compra varían según la ciudad o zona?

| Ciudad | n_tickets | media_total | mediana | sd_total | minimo | maximo |
|--------------------------|-----------|-------------|---------|----------|--------|--------|
| ONTINYENT | 1 | 112,02 | 112,02 | NA | 112,02 | 112,02 |
| ALCOI/ALCOY | 22 | 102,55 | 128,64 | 71,48 | 9,66 | 234,20 |
| REQUENA | 1 | 89,18 | 89,18 | NA | 89,18 | 89,18 |
| MONCADA | 5 | 66,99 | 57,33 | 25,51 | 49,94 | 111,08 |
| SAGUNT/SAGUNTO | 1 | 66,30 | 66,30 | NA | 66,30 | 66,30 |
| BENIFAIO | 4 | 64,69 | 62,35 | 48,34 | 12,56 | 121,49 |
| BUÑOL | 12 | 58,12 | 60,84 | 28,52 | 18,90 | 105,91 |
| VERA | 3 | 52,07 | 52,12 | 0,86 | 51,19 | 52,91 |
| SAN ANTONIO DE BENAGEBER | 1 | 49,74 | 49,74 | NA | 49,74 | 49,74 |
| BENIJOFAR | 8 | 48,12 | 46,16 | 22,48 | 6,27 | 80,53 |
| MURO DE ALCOY | 24 | 45,02 | 32,49 | 36,08 | 3,20 | 161,22 |
| VALENCIA | 120 | 41,19 | 34,66 | 31,07 | 0,43 | 187,52 |
| BURJASSOT | 27 | 38,16 | 35,52 | 17,58 | 3,07 | 82,81 |
| ALGINET | 5 | 33,48 | 30,94 | 16,01 | 18,64 | 55,35 |
| ALBORAIA/ALBORAYA | 50 | 32,20 | 24,20 | 23,87 | 1,75 | 90,03 |
| GARRUCHA | 4 | 24,89 | 24,86 | 11,75 | 12,10 | 37,76 |
| GANDIA | 1 | 3,90 | 3,90 | NA | 3,90 | 3,90 |
| SANTA EULÀRIA DES RIU | 1 | 3,08 | 3,08 | NA | 3,08 | 3,08 |
| SANT JOSEP DE SA TALAIA | 1 | 0,77 | 0,77 | NA | 0,77 | 0,77 |

Para verlo de manera más visual, podemos hacer un diagrama de boxplot, donde la línea central refleja la mediana y los cuartiles la caja



Como podemos ver en el boxplot Alcoy es la ciudad con más tickets y donde más se gasta en promedio, ciudades como valencia y burjassot tienen una media bastante estable en todas las compras aunque encontramos casos donde hay compras de más de 100€. La elevada media de Alcoy puede deberse a que la mayoría de los

tickets son de compras muy altas lo que también podría estar relacionada con los precios en ese Mercadona, mientras en los otros municipios la media se mantiene bastante constante y dentro de rangos más bajos.

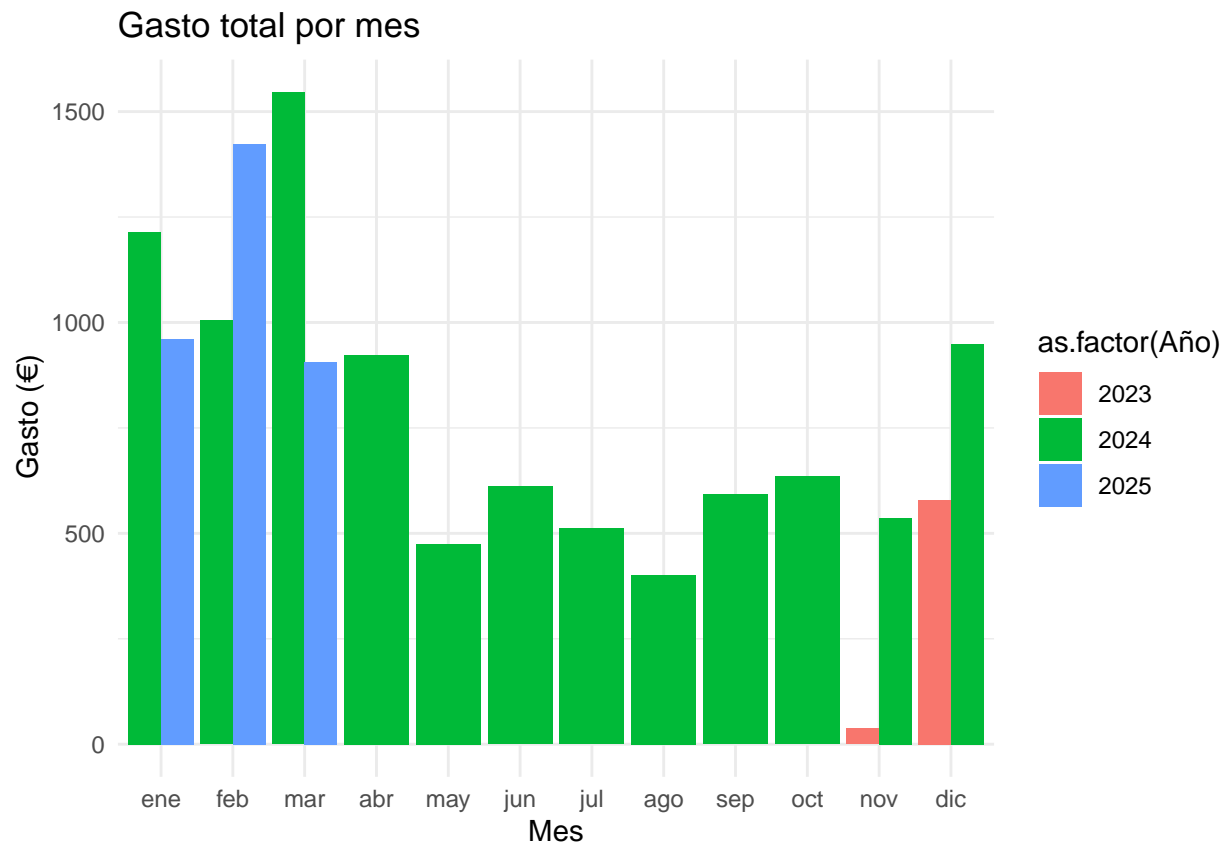
- ¿Cuánto dinero de media se gasta cada cliente en una compra?

| n_compras | gasto_medio | gasto_mediana | sd_gasto |
|-----------|-------------|---------------|----------|
| 291 | 46 | 37 | 38 |

En nuestro estudio tenemos 291 tickets, el gasto medio de cada cliente es de 45,8€ por compra, mientras que la mediana nos indica que el 50% de las compras cuestan menos de 37,05. La desviación de 38 nos indica que el gasto por compra varía bastante de un cliente a otro, es decir hay compras pequeñas de 20€y otras muy grandes de más de 100€.

- ¿Cuál es el mes o período en el que más gastos se realizan? ¿Durante las vacaciones de navidad?

| Año | Mes | Gasto_Total |
|------|-----|-------------|
| 2024 | mar | 1546 |
| 2025 | feb | 1423 |
| 2024 | ene | 1214 |
| 2024 | feb | 1004 |
| 2025 | ene | 961 |
| 2024 | dic | 949 |



Si nos fijamos en los tickets del año 2023 (color rosa) el mes con más gastos es diciembre. Fijandonos en el año 2024 (verde), el período con más gastos es en enero (siendo el único año con tickets todos los meses). En el año 2025 (azul) es febrero el mes con más gastos. Por lo que el período con más gastos no es en la época de Navidad.

- ¿Qué productos suelen comprarse juntos?

| Descripcion.x | Descripcion.y | Frecuencia |
|----------------------|----------------------|------------|
| ALMENDRA NATURAL | CACAHUETE CHOCOLATE | 16 |
| ALMENDRA NATURAL | PEPINO | 16 |
| CACAHUETE CHOCOLATE | ALMENDRA NATURAL | 16 |
| CACAHUETE CHOCOLATE | PEPINO | 16 |
| PEPINO | ALMENDRA NATURAL | 16 |
| PEPINO | CACAHUETE CHOCOLATE | 16 |
| ALMENDRA NATURAL | BARRITA MUESLI CHOCO | 14 |
| ATUN CLARO OLIVA | PAN SEMILLAS | 14 |
| BARRITA MUESLI CHOCO | ALMENDRA NATURAL | 14 |
| BARRITA MUESLI CHOCO | CACAHUETE CHOCOLATE | 14 |

Se observa que “almendra natural” se suele comprar con “cacahuete chocolate”, “pepino” o “barrita muesli choco” y el “pan semillas” con “atun claro oliva”.

- ¿Cuál es el producto más caro registrado en los tickets?

| | index | Cantidad | Descripcion | Precio | Importe | Peso | Tipo |
|------|-------|----------|------------------|--------|---------|------|------|
| 2421 | 52 | 1 | ALISTADO MEDIANO | 31 | 31 | 1 | |

El producto más caro registrado es el “alariado mediano”, que tiene un precio de 31.33 euros.

- ¿Hay una cantidad media de productos por tickets?

[1] 19

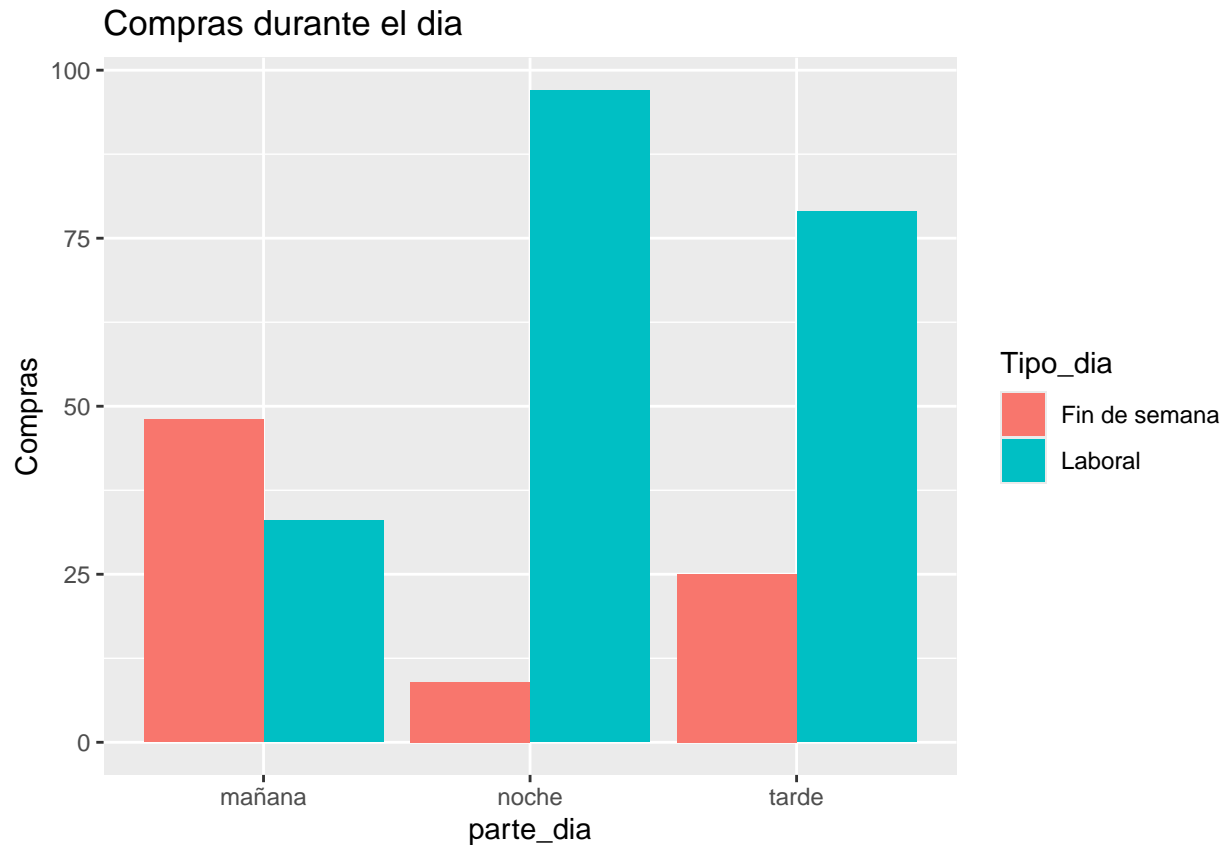
La cantidad media de productos por ticket es de 18.80198 productos.

- ¿Cuando se hacen más compras en las diferentes partes del día: mañana, tarde y noche?

Para esta pregunta separaremos el día en mañana (9:00 - 13:59), tarde (14:00 - 18:59) y noche (19:00 - 21:45). Para ello crearemos una nueva columna.

| parte_dia | Compras |
|-----------|---------|
| noche | 106 |
| tarde | 104 |
| mañana | 81 |

Podemos observar que durante la noche y la tarde se hacen muchas más compras que en comparación por la mañana, ya que las compras entre semana se hacen más usualmente por la tarde, ya que por la mañana las personas normalmente trabajan o estudian. Por esto queremos ir más allá y vamos a separarlo entre días de la semana y días del fin de semana



Se confirma lo que pensabamos se va más por la mañana en el fin de semana en cambio durante los dias laborales por la tarde y la noche se va mucho más a comprar

- ¿Cuánto supone económicamente de media el IVA en las compras

```
«««< HEAD
```

```
'''
## [1] 3,4
'''
```

===== - ¿Que productos son los más comprados durante las diferentes estaciones, verano (junio - agosto), invierno(diciembre - marzo)... ?

Para ello haremos los mismo que hemos hecho previamente pero esta vez dividiendo por estaciones: verano (junio - agosto), invierno(diciembre - febrero), otoño (septiembre - noviembre) y primavera (marzo - mayo)

| estacion | Descripcion | Frecuencia |
|----------|---------------------|------------|
| invierno | PLATANO | 30 |
| invierno | PAN SEMILLAS | 17 |
| invierno | KIWI VERDE | 15 |
| invierno | QUESO LONCHAS CABRA | 15 |
| invierno | FILETE PECHUGA | 14 |

| estacion | Descripcion | Frecuencia |
|-----------|------------------|------------|
| primavera | TOMATE TRITURADO | 12 |
| primavera | UVA BLANCA S/SEM | 9 |
| primavera | BANANA | 8 |
| primavera | COPOS DE AVENA | 8 |
| primavera | NARANJA 5 KG. | 8 |

| estacion | Descripcion | Frecuencia |
|----------|---------------------|------------|
| verano | QUESO LONCHAS CABRA | 10 |
| verano | PAN SEMILLAS | 9 |
| verano | CANÓNIGOS | 8 |
| verano | FILETE PECHUGA | 8 |
| verano | FRESA | 8 |

| estacion | Descripcion | Frecuencia |
|----------|---------------------|------------|
| otoño | PLATANO | 12 |
| otoño | QUESO LONCHAS CABRA | 10 |
| otoño | 12 HUEVOS GRANDES-L | 9 |
| otoño | FILETE PECHUGA | 9 |
| otoño | ACT 0% NAT ED 8 | 8 |

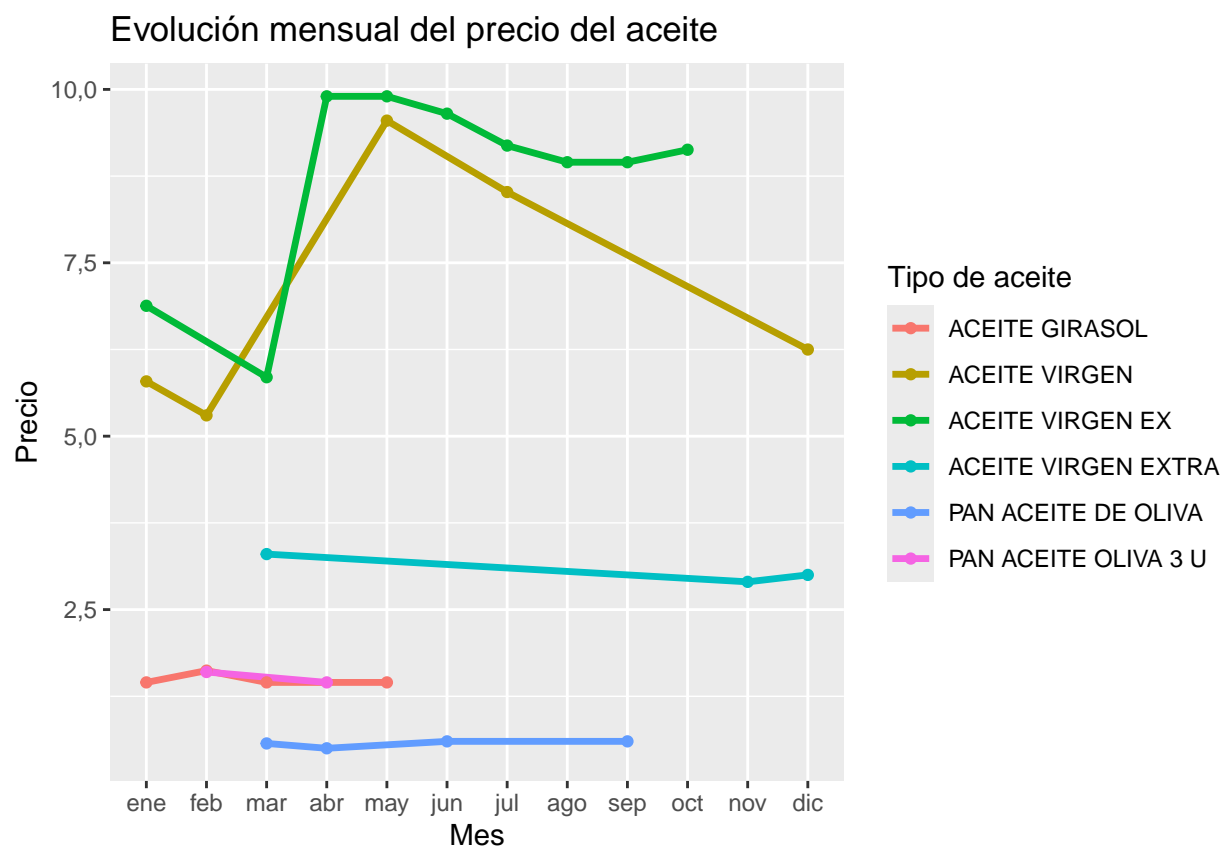
Una de las cosas destacables ya ha sido comentado previamente en una de las preguntas, ya que en invierno y en otoño se puede ver que el platano es el producto más comprado cuando luego en la tabla de verano y primavera no aparece y en su defecto aparece la Banana. Esto es debido a que durante las estaciones de primavera y verano el precio del platano es mucho más superior al estar fuera de temporada. Igualmente con otras frutas que salen en el top únicamente en una estación o en dos, podríamos atribuir este hecho a que cada fruta tiene una temporada.

- ¿Hay relación con precio medio de los productos y la cantidad de productos de la compra?

[1] 0,4

Nos sale una correlación de 0.4 (si utilizamos el metodo de spearman), lo que quiere decir que si que hay una pequeña relación. Una teoria posible puede ser que cuando más grande es la compra compramos productos más caros como carnes o pescados ya que suele ser la “compra de la semana”, en cambio si son compras más pequeñas la media de los productos es más baja ya que suelen ser snaks o productos más baratos para el consumo en corto plazo. »»»> 8f57bc26aec6548212ade53ee2de8a931251e5dc

- ¿Como ha variado el precio del aceite, y relacionados, a lo largo del tiempo, haz una gráfica para apreciarlo mejor?



Según nuestros datos, como se ve en la gráfica, se nota una gran subida de precio en el aceite virgen y

- ¿En que se ha gastado(€) más en pescado o en otros productos que van por peso? ¿Es equivalente al número de veces que se han comprado unidades de estos?

En la tabla 1 se puede apreciar el gasto total en euros, y la cantidad vendidad de ese tipo de producto.

Table 1: Venta Pescado Fruta y Verdura

| Tipo | Gastado | Compras |
|-----------------|---------|---------|
| Fruta o Verdura | 916 | 374 |
| Pescado | 243 | 27 |

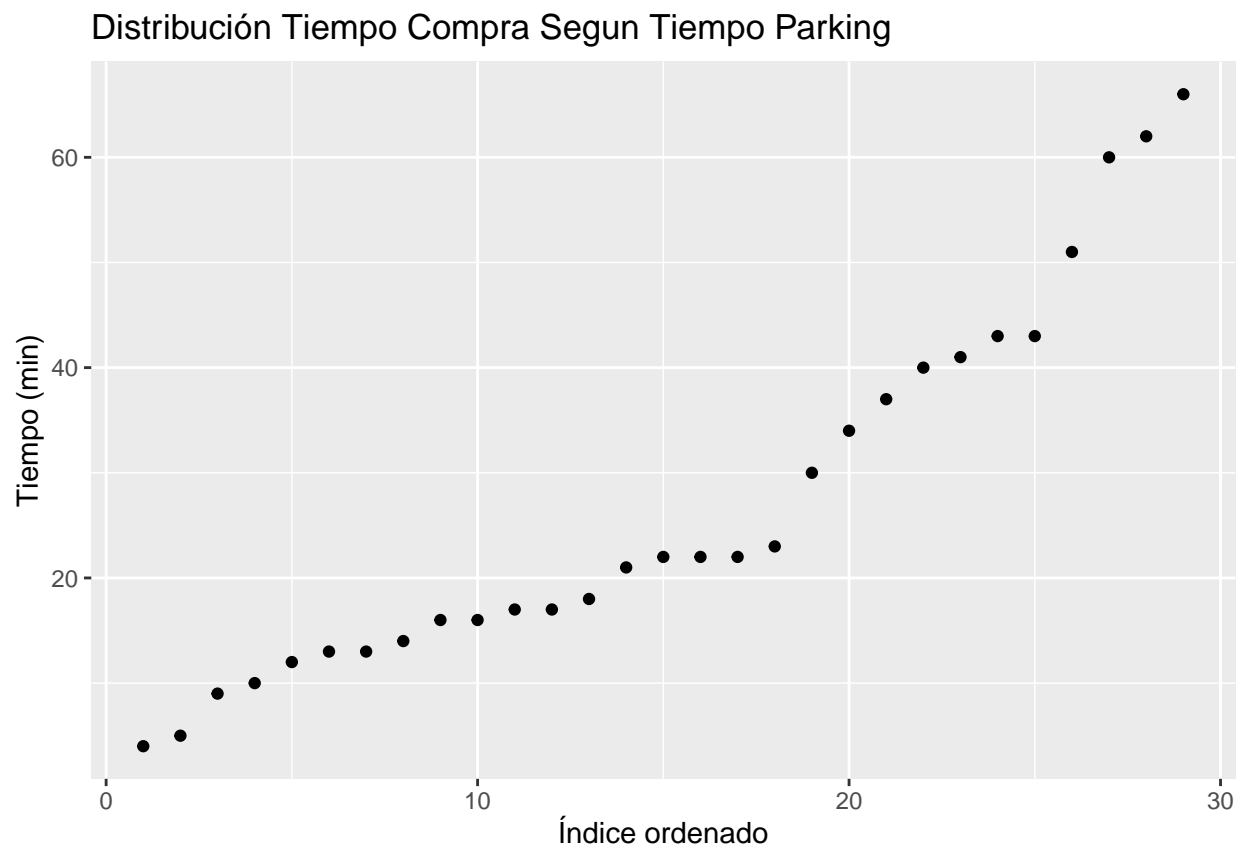
- ¿Hay una relación entre la cantidad de productos que se compra y si ha utilizado el parking o no?

| usado_parking | mediana_prod | total_prod |
|---------------|--------------|------------|
| FALSE | 13 | 262 |
| TRUE | 17 | 29 |

Como se ve en el dataframe, cuando no se utiliza el parking, el número de productos es más bajo (mediana

Por lo que podemos concluir, que cuando se utiliza el parking es más probable que se compren más produc

- ¿Cual es el tiempo medio que tardan en hacer la compra las personas que han utilizado el parking?



El tiempo medio que tardan en hacer la compra las personas que utilizan el parking es de 28 minutos y m