

MINI PROYECTO TD 2025

Andrea Fu Castelló Sala
Sergio Del Carme Moreno
María Martínez Marí
Jorge Navarro Rodríguez
Iván Pérez Alonso
Edurne Serigó Troyano

ÍNDICE:

- INTRODUCCIÓN
- CARGA DE LIBRERÍAS
- IMPORTACIÓN DE DATOS
- MATERIAL Y MÉTODOS
- RESULTADOS DE LAS PREGUNTAS PLANTEADAS
- CONCLUSIÓN

INTRODUCCIÓN

- Contexto: mercadona ha implementado el ticket electrónico (los clientes pueden recibir el ticket en formato pdf por correo)
- Objetivo del proyecto: desarrollar un programa que analice los tickets del mercadona

CARGA DE LIBRERÍAS

- `library(pdftools)`
- `library(tidyverse)`
- `library(dplyr)`
- `library(stringr)`
- `library(ggplot2)`
- `library(knitr)`

MATERIAL Y MÉTODOS

- Fuente de datos: tickets del mercadona
- Lenguaje y herramientas usadas: lenguaje r y librería clave pdftools
- Proceso de lectura: se usa la función pdf_text()

IMPORTACIÓN DE DATOS

Pasos:

1. Definir la ruta del directorio
2. Identificar todos los archivos pdf del mercadona
3. Leer los archivos
4. Crear un data frame por archivo
5. Combinar todos los data frame en uno (df_lineas)
6. Crear una tabla para visualizar el data frame

1 MERCADONA, S.A. A-46103834
2 CTRA. GARRUCHA A VERA S/N
3 04630 GARRUCHA
4 TELÉFONO: 950133380
5 25/11/2023 09:09 OP: 78800
6 FACTURA SIMPLIFICADA: 2916-010-520925
7
8
9
10
11 Descripción P. Unit Importe
12 5 DONACIÓN 1,00 5,00
13 1 BARREÑO 2,20
14 3 PALO ANTIDESLIZANTE 1,90 5,70
15 1 MULTIUSOS 2,55
16 2 AGUA MINERAL 0,73 1,46
17 2 B.BASURA EXT.C.FÁCIL 1,60 3,20
18 1 ESCURRE FÁCIL 2,85
19 1 CUBO FREGAR C/RUEDAS 3,85
20 1 FRIEGASUELOS PINO 0,95
21 2 FIBRAS CRUZADAS 2,10 4,20
22 1 LOTE BAYETAS SUAVES 1.20

RESULTADOS

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

	producto	unidades_vendidas
1	LECHE DESNAT. CALCIO	24
2	QUESO LONCHAS CABRA	20
3	COPOS DE AVENA	16
4	BARRA CAMPESINA	14
5	GARBANZO M.COCIDO	14

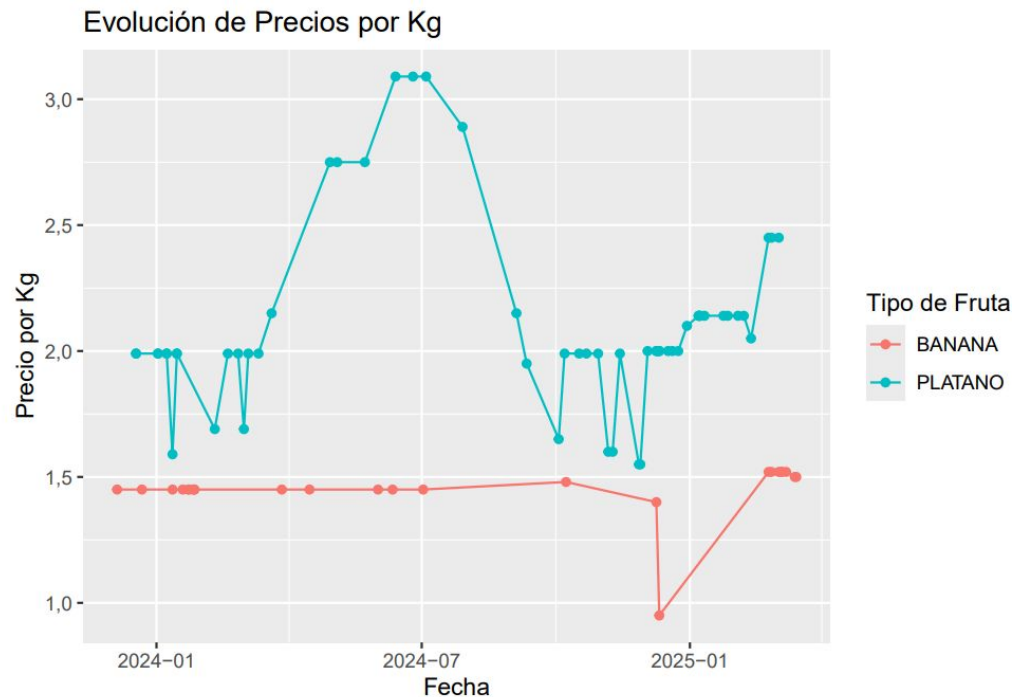
2. Si consideramos la categoría de FRUTAS Y VERDURAS. Cuáles son los 5 productos más vendidos? ¿Cuántos kilos se han vendido de cada uno de estos productos?

	producto	kilos_totales
1	NARANJA	82,06
2	PATATA	74,80
3	MANDARINA	31,19
4	CEBOLLA	22,00
5	PEPINO	19,62

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

	producto	kilos_totales
1	BACALADILLA	7,37
2	RODABALLO	3,67
3	SEPIA LONJA	3,58
4	DORADA	2,93
5	SEPIA FRESCA	2,64

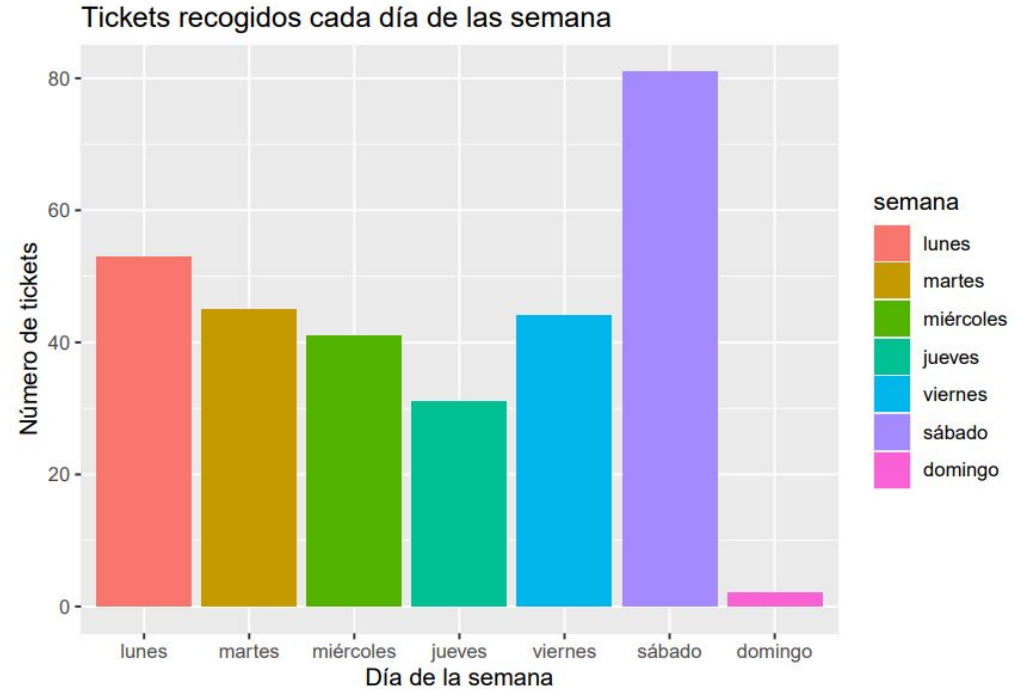
4. Muestra mediante un gráfico de líneas como ha variado el precio por kilo de las bananas y los plátanos en los tickets disponibles, a lo largo del tiempo.



5. ¿Cuál es la procedencia de los tickets? ¿Qué ciudad/ pueblo tiene un mayor número de tickets?

	procedencia	Freq
17	VALENCIA	130
1	ALBORAIA/ALBORAYA	50
7	BURJASSOT	26
10	MURO DE ALCOY	24
2	ALCOI/ALCOY	22
6	BUÑOL	12
5	BENIJOFAR	9
3	ALGINET	5
4	BENIFAIO	4
9	GARRUCHA	4
18	VERA	3
13	SAGUNT/SAGUNTO	2
8	GANDIA	1
11	ONTINYENT	1
12	REQUENA	1
14	SAN ANTONIO DE BENAGEBER	1
15	SANT JOSEP DE SA TALAIA	1
16	SANTA EULÀRIA DES RIU	1

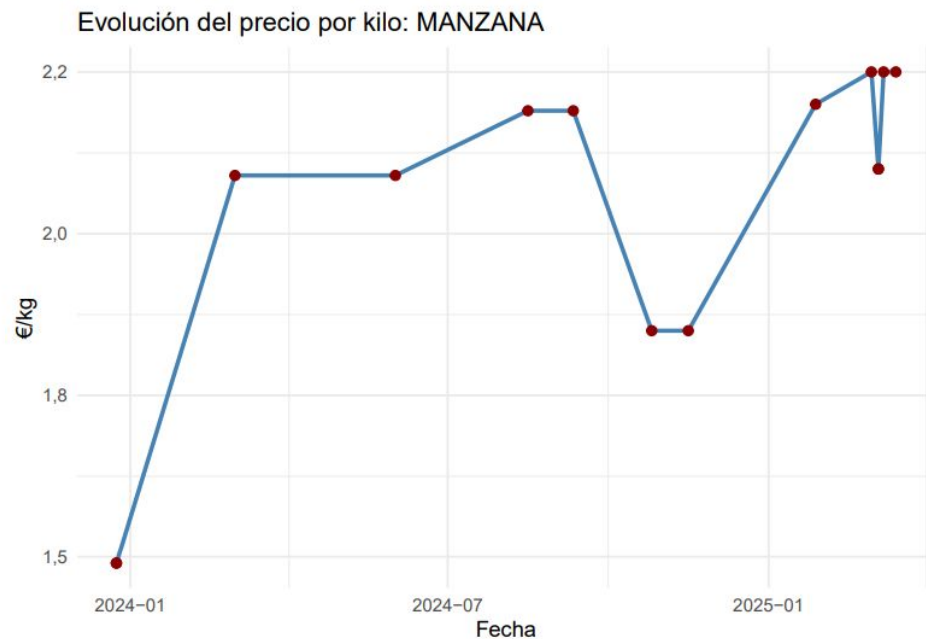
6. Muestra mediante un diagrama el número de tickets recogidos cada día de las semana. ¿Si tuvieses que cerrar un día entre semana qué día lo harías?



7. ¿Cuál es la media de ventas para cada día de la semana?

	▲ dia_semana	media_productos
1	lunes	17
2	martes	14
3	miércoles	16
4	jueves	11
5	viernes	16
6	sábado	19
7	domingo	14

8. ¿Cómo ha evolucionado el precio por unidad o por kilo de un producto específico a lo largo del tiempo?



9. 1. ¿Cuáles son los productos más vendidos en el conjunto de tickets disponibles?

producto <chr>	unida... <dbi>
LECHE DESNAT CALCIO	24
QUESO LONCHAS CAB...	20
COPOS DE AVENA	16
BARRA CAMPESINA	14
GARBANZO MCOCIDO	14
LECHE DESN SLACT	14
CHAPATA CRISTAL	12
TOMATE TRITURADO	12
BROTES TIERNOS MAXI	10
ISOTÓNICA LIMÓN	10

9. 2. ¿Cuáles son los productos menos vendidos en el conjunto de tickets disponibles?

producto <chr>	unida... <dbl>
A ALIÑADA PICADEDOS	1
ACEITE GIRASOL	1
ACEITUNA CHUPADED...	1
AJO Y PEREJIL	1
ALBONDIGAS UNID	1
ALCACHOFA PEQUEÑA	1
ALMENDRA LAMINADA	1
ALTRAMUCES	1
ANCHOA GPACK PLAS	1
ANGURIÑAS GR	1

10. ¿Cuál es el producto del que se adquieren más unidades en una sola compra?

Producto con el mayor número de unidades en un ticket

unidades	descripcion
----------	-------------

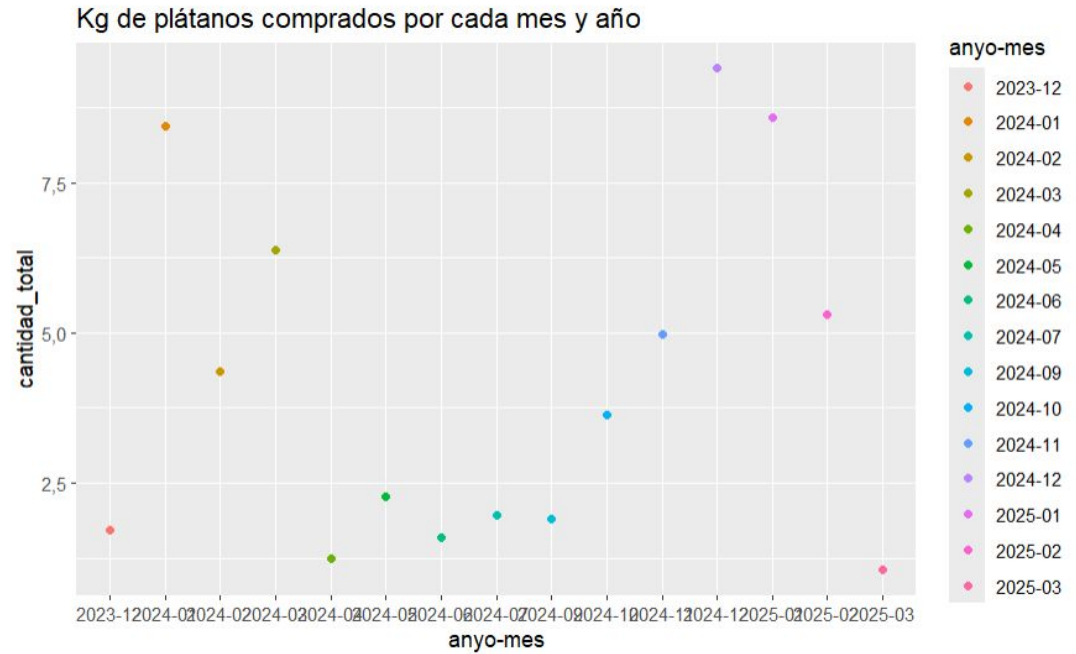
8	COLA SIN CAFEINA
---	------------------

11. ¿Cuáles son las combinaciones de productos más frecuentes?

Pares de productos diferentes que se compran más veces
juntos

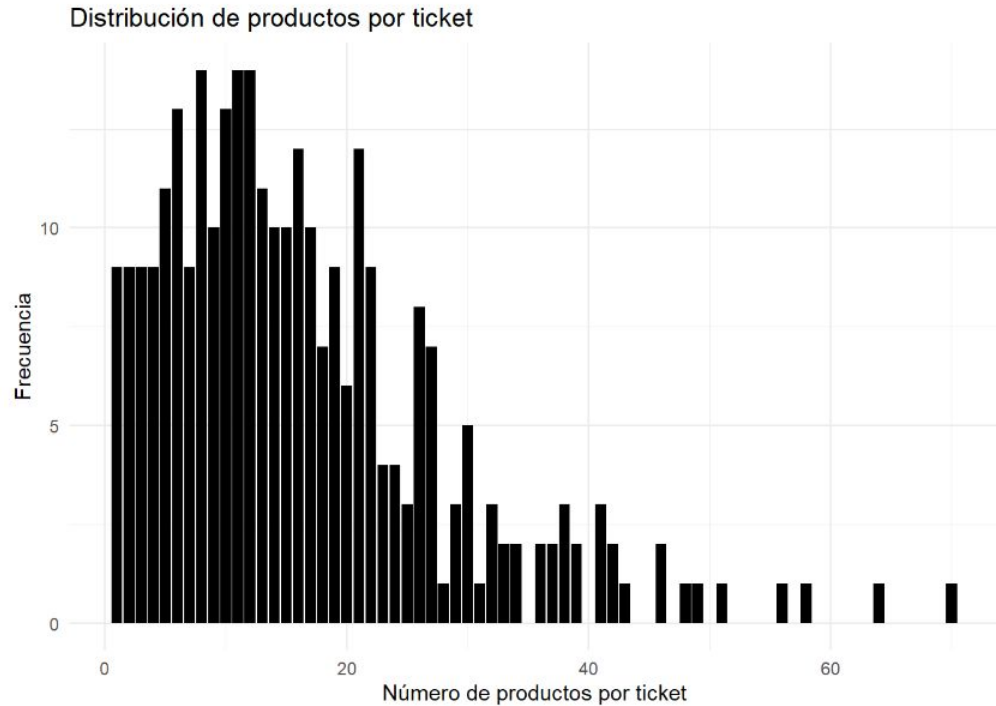
productos	n
BROTOS TIERNOS MAXI QUESO LONCHAS CABRA	16
LECHE DESNAT. CALCIO QUESO LONCHAS CABRA	14
PAN DE PUEBLO QUESO LONCHAS CABRA	14
PLATANO QUESO LONCHAS CABRA	14
C. POLLO 100% NAT QUESO LONCHAS CABRA	12
QUESO FRESCO CABRA QUESO LONCHAS CABRA	12

12. ¿Existe la estacionalidad en plátanos?



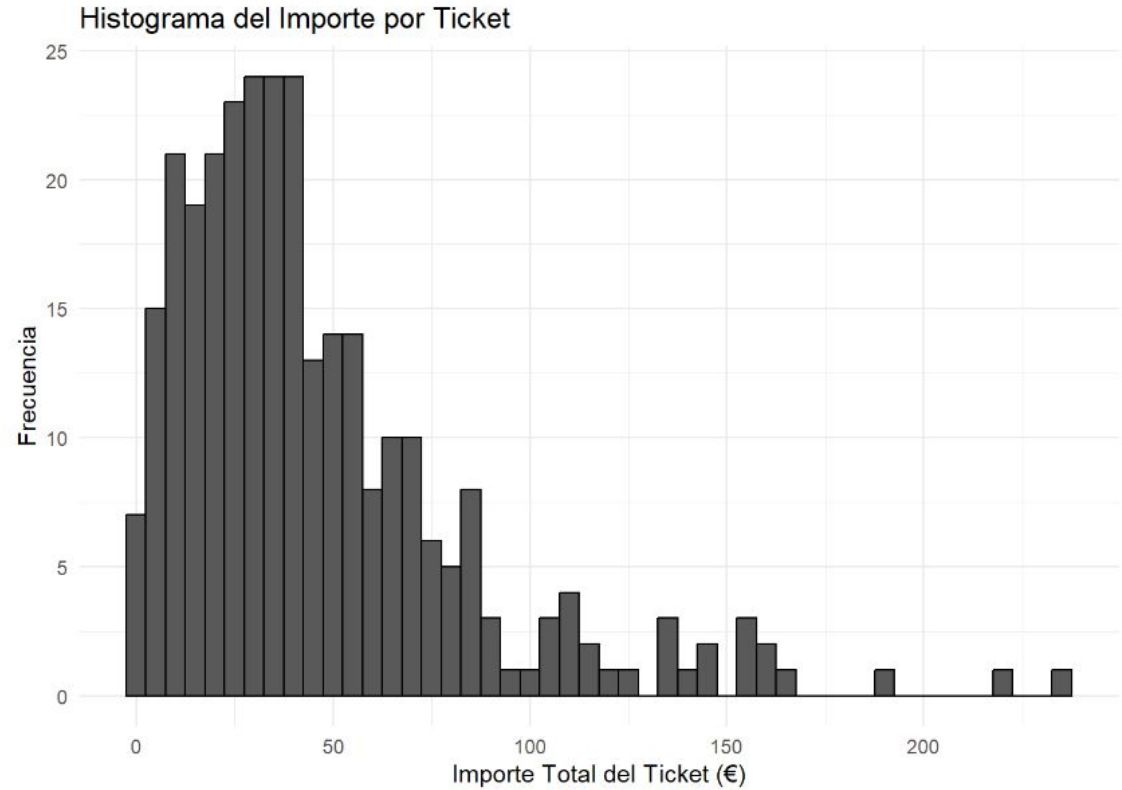
13. ¿Cuántos productos se compran por ticket, en promedio?

16



14. 1. ¿Cuál es el importe medio por ticket?

46€



14. 2.

¿Cuál es el ticket más caro registrado?

¿Y el ticket más barato registrado?

El más caro:

archivo <chr>	Metodo_Pago <chr>	Importe_Total <dbl>
20240323 Mercadona 234,20.pdf	Tarjeta	234

El más barato:

archivo <chr>	Metodo_Pago <chr>	Importe_Total <dbl>
20241227 Mercadona 0,43 €.pdf	Tarjeta	0,43

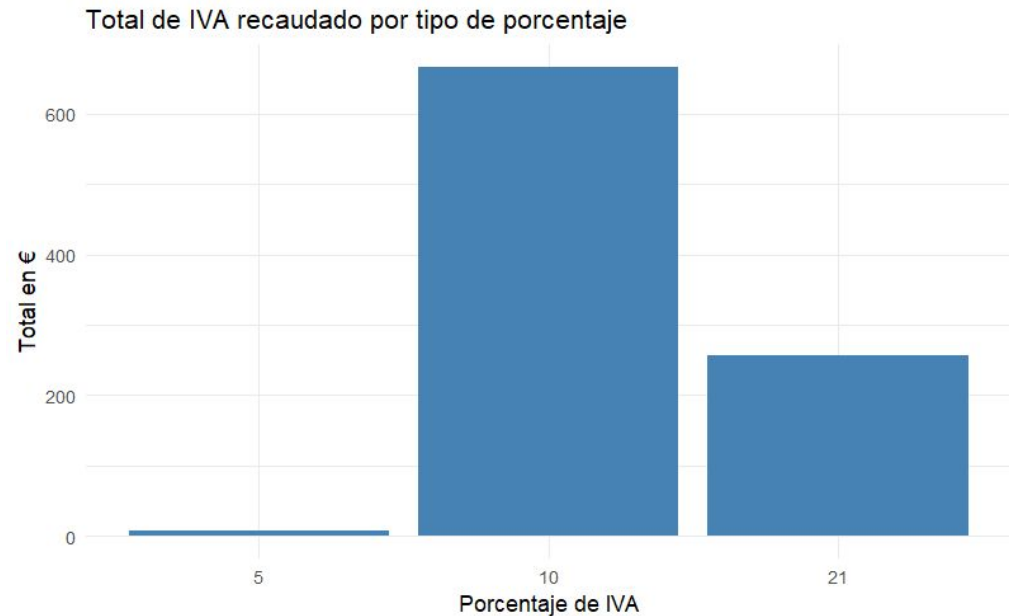
15. ¿Cuál es la cantidad total de dinero que se obtiene en impuestos cuando se venden alimentos con un 10% de IVA? ¿Y con un 21% y un 5%?

Total de IVA por
tipo de porcentaje
porcentaje Cuota

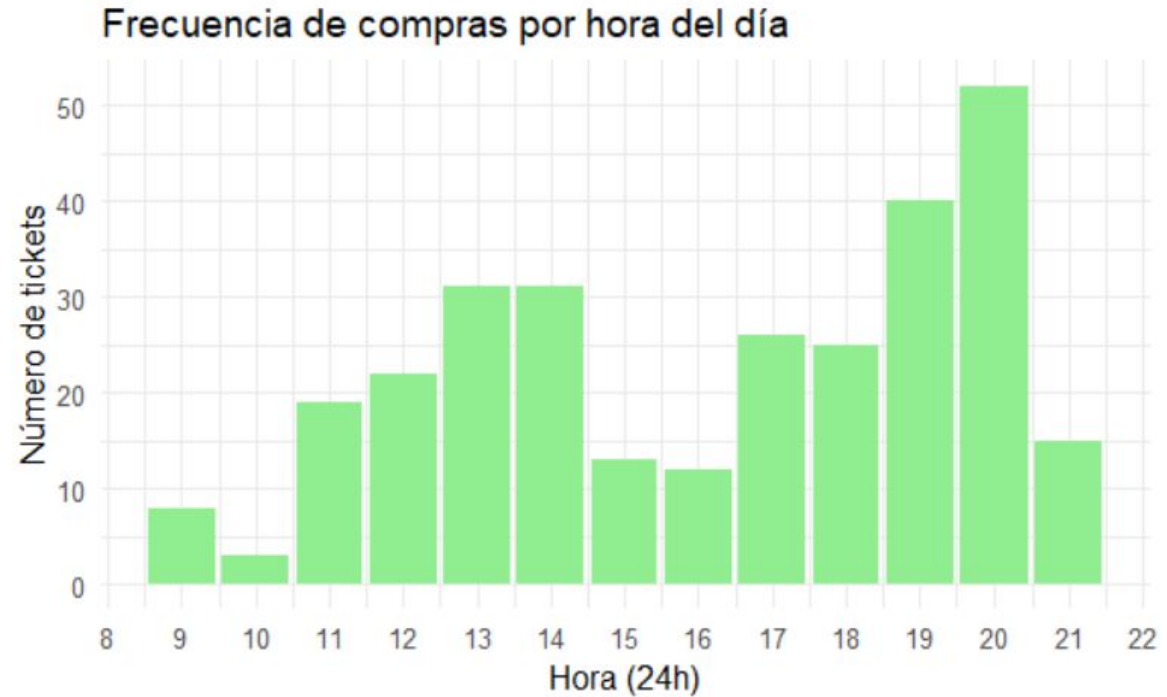
21 256.74

10 666.01

5 8.21

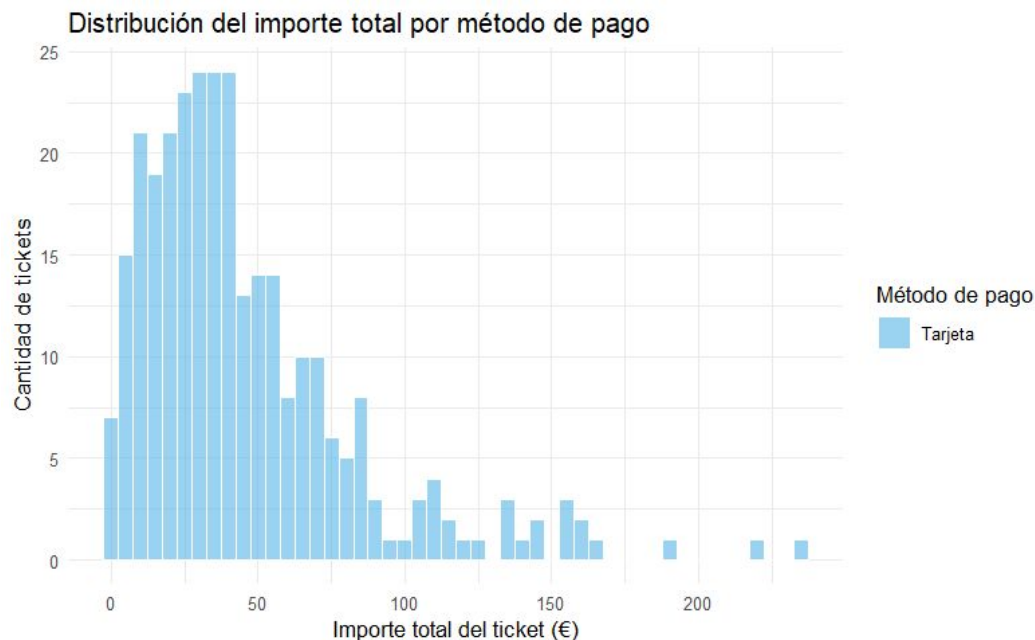


16. ¿A qué horas se suele ir más a comprar en los supermercados? ¿Cuáles son las que menos?



17. ¿Qué método de pago es más frecuente en los tickets: tarjeta o efectivo?
¿Cuánto se ha gastado en total con cada método de pago?

Importe total de cada método de pago	
Metodo_Pago	Total_Gastado
Tarjeta	13649.8



CONCLUSIÓN

- Objetivos cumplidos (automatizar lectura y análisis de tickets)
- Datos Analizados: productos, precios, fechas, cantidades, patrones de compra y combinaciones frecuentes.
- Que hemos conseguido: entender hábitos de consumo, apoyar decisiones de compra y mejorar la gestión del gasto familiar.