

PROYECTO TD 2025

Análisis de tickets de compra en supermercados Mercadona

Mar Alemany, Martín González, Mateo
Reina, Adrian Mena, Sina Taheri

INTRODUCCIÓN

- Digitalización de tickets en supermercados.
- Objetivo: Analizar datos de tickets de Mercadona usando R.
- Se busca entender hábitos de consumo y tendencias.



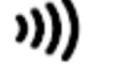
MERCADONA, S.A. A-46103834
C/ PINTOR PERIS ARAGÓ 34
46120 ALBORAIA/ALBORAYA
TELÉFONO: 961204862
22/11/2024 18:04 OP: 481195
FACTURA SIMPLIFICADA: 4523-015-509484



Descripción	P. Unit	Importe
1 COLA CAO		4,39
		TOTAL (€) 4,39
		TARJETA BANCARIA 4,39
IVA 10%	BASE IMPONIBLE (€) 3,99	CUOTA (€) 0,40
TOTAL	3,99	0,40

TARI. BANCARIA: **** * 6020
N.C: 036426237
AID: A000000041010

AUT: 045353
ARC: 00


MASTERCARD
Importe: 4,39 €


MASTERCARD


MERCADONA
800 500 220

SE ADMITEN DEVOLUCIONES CON TICKET

MATERIAL Y MÉTODOS

Librerías

Lorem ipsum dolor sit amet,
consectetur adipiscing elit.
Duis vulputate nulla at ante
rhoncus, vel efficitur felis
condimentum. Proin odio odio.

dataframes

- Dataframe general
- Dataframe productos

IMPORTACION Y CARGA DE DATOS

- 303 archivos PDF renombrados y cargados automáticamente
- Data frame a partir de los datos extraídos con variables:

Variable	Descripción
comercio	Nombre del comercio
empresa	Tipo y código de empresa
direccion	Dirección del comercio
cp	Código postal
telefono	Teléfono del comercio
fecha	Fecha de la compra (día-mes-año)
hora	Hora de la compra (horas y minutos)
op	Número del código de la operación
fs	Código de la factura simplificada
productos	Lista con los productos comprados
total	Dinero total de la compra
forma_pago	Forma de pago (tarjeta o en efectivo)
base_imp	Base Clasificación de Productos
cuota	Cuota del IVA

CLASIFICACIÓN DE PRODUCTOS

- 1 | Pescadería**
- 2 | Frutas y Verduras**
- 3 | Productos por unidad**

Table 2. Datos procesados para productos de pescadería

num_ticket	nombre	peso_kg	precio_kg	importe
168515	SARDINA GR 13/20	0.390	6.95	2.71
433003	BOQUERÓN MED 51/80	0.756	4.95	3.74
075372	LENGUADO FRESC EUROP	0.262	24.95	6.54
491735	CORVINA	1.560	9.95	15.52
491735	CIGALA PEQUEÑA REF	0.222	16.85	3.74
087622	LENGUADO FRESC EUROP	0.376	18.95	7.13
516718	LENGUADO FRESC EUROP	0.610	18.95	11.56
696429	BACALADILLA	1.590	3.95	6.28
682028	SARDINA MED 21/35	0.366	5.95	2.18
696429	BACALADILLA	1.590	3.95	6.28

Table 3. Descripción de variables del conjunto de frutas y verduras

num_ticket	nombre	peso_kg	precio_kg	importe
301952	MANZANA GOLDEN	1.694	1.49	2.52
168515	LANGOSTINO COCIDO	0.424	10.95	4.64
405598	PATATA	1.378	1.99	2.74
407209	PLATANO	1.202	2.75	3.31
070987	PATATA	1.930	1.99	3.84
070987	TOMATE ENSALADA	1.412	1.75	2.47
330410	PEPINO	0.652	1.45	0.95
330410	BANANA	0.936	1.45	1.36
330410	PARAGUAYO	0.594	3.79	2.25
330410	MANZANA GOLDEN	1.216	2.09	2.54

Tablas de productos

Table 4. Descripción de variables del conjunto de productos por unidad

num_ticket	nombre	cantidad	precio_unitario	importe
520925	DONACIÓN	5	1.00	5.00
520925	BARREÑO	1	2.20	2.20
520925	PALO ANTIDESLIZANTE	3	1.90	5.70
520925	MULTIUSOS	1	2.55	2.55
520925	AGUA MINERAL	2	0.73	1.46
520925	B.BASURA EXT.C.FÁCIL	2	1.60	3.20
520925	ESCURRE FÁCIL	1	2.85	2.85
520925	CUBO FREGAR C/RUEDAS	1	3.85	3.85
520925	FRIEGASUELOS PINO	1	0.95	0.95
520925	FIBRAS CRUZADAS	2	2.10	4.20

Data frame de productos

Table 5. Descripción de variables del conjunto de datos final

num_ticket	nombre	cantidad	precio	importe	tipo
007267	PERA CONFERENCIA	0.854	2.55	2.18	fruta_verdura
007267	MANZ. ROJA DULCE	1.474	2.90	4.27	fruta_verdura
007267	PIZZA BOLOÑESA	1.000	2.60	2.60	unidades
007267	MIX FRUTOS ROJOS	1.000	1.66	1.66	unidades
007267	PISTO DE VERDURAS	1.000	2.30	2.30	unidades
007267	FILETE DE TRUCHA	1.000	3.92	3.92	unidades
007267	PAN SEMILLAS	1.000	1.60	1.60	unidades
007267	SNACK PIPAS	1.000	1.40	1.40	unidades
007267	+PROT NATILLA VAINI	1.000	1.75	1.75	unidades
007267	SNACK CHOCOLATE	1.000	1.50	1.50	unidades

ANÁLISIS EXPLORATORIO

Análisis univariante

Variables categóricas
Variables de tipo numérico

El gasto total en todos los productos es: 14022.5 €

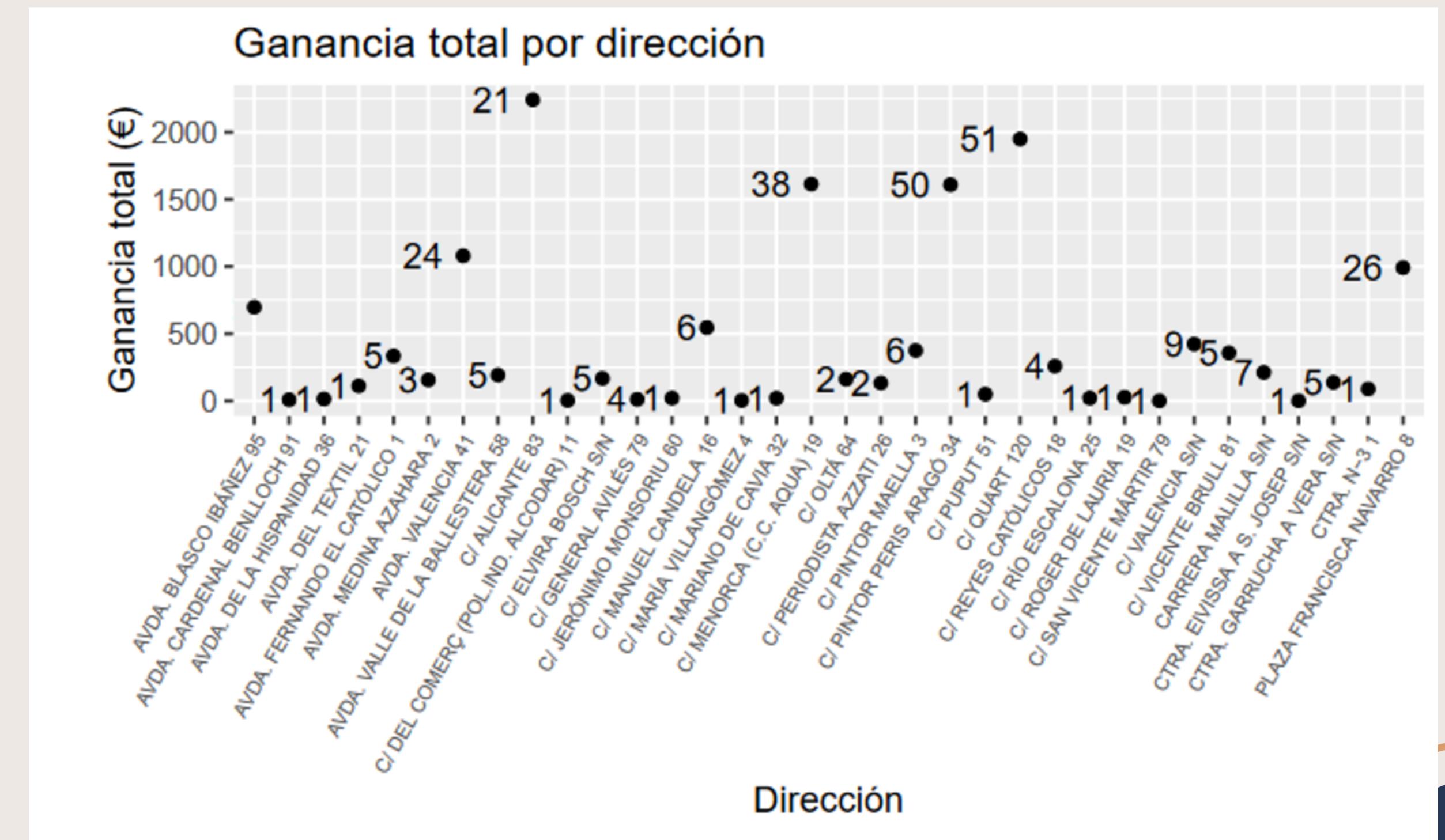
La media de productos comprados 138

por ticket es: 19.9 La media del importe total por ticket es: 48.35 €

ANÁLISIS EXPLORATORIO

Análisis multivariante

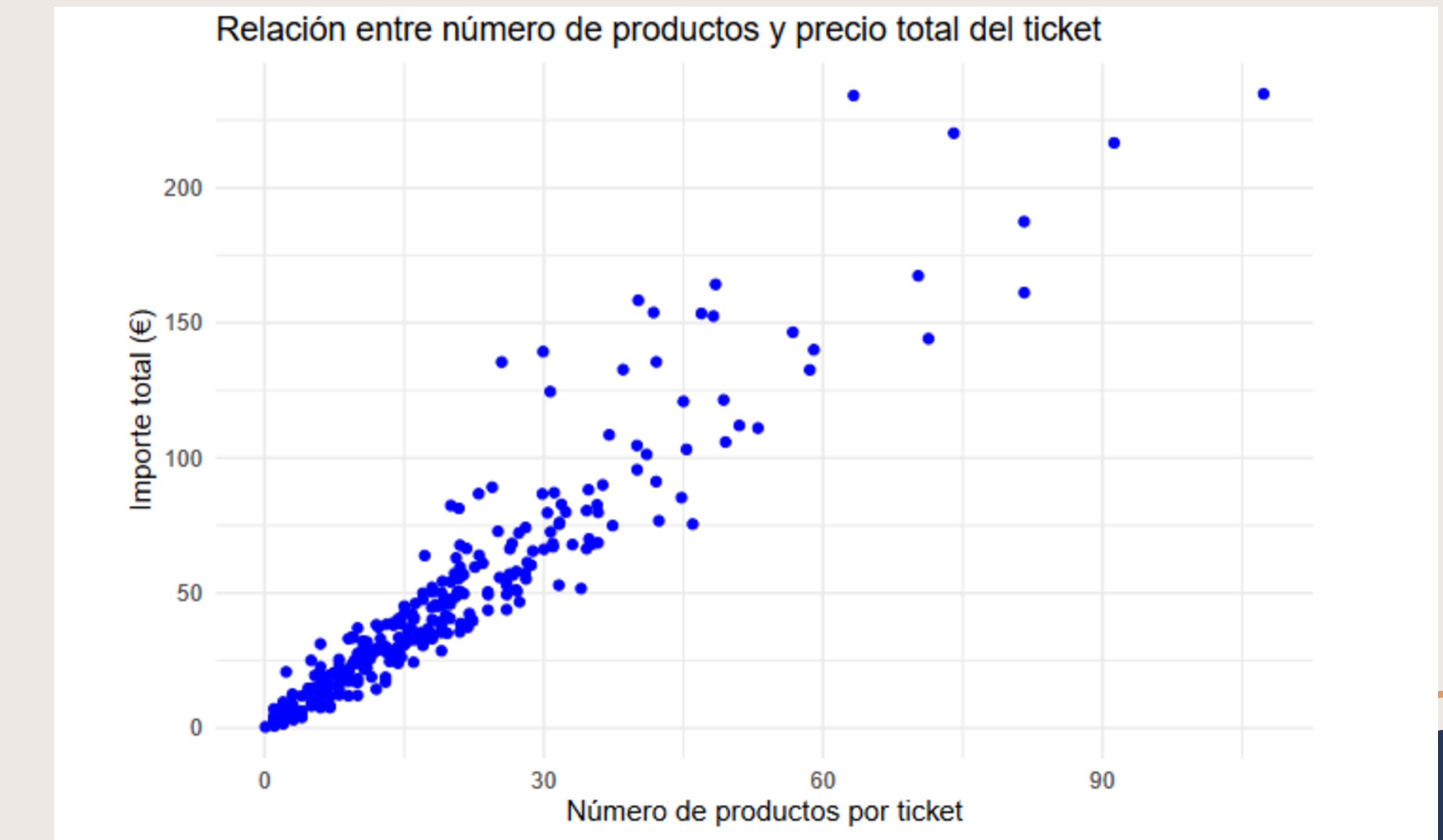
Relación entre la
Dirección y las
Ganancias



ANÁLISIS EXPLORATORIO

Análisis multivariante

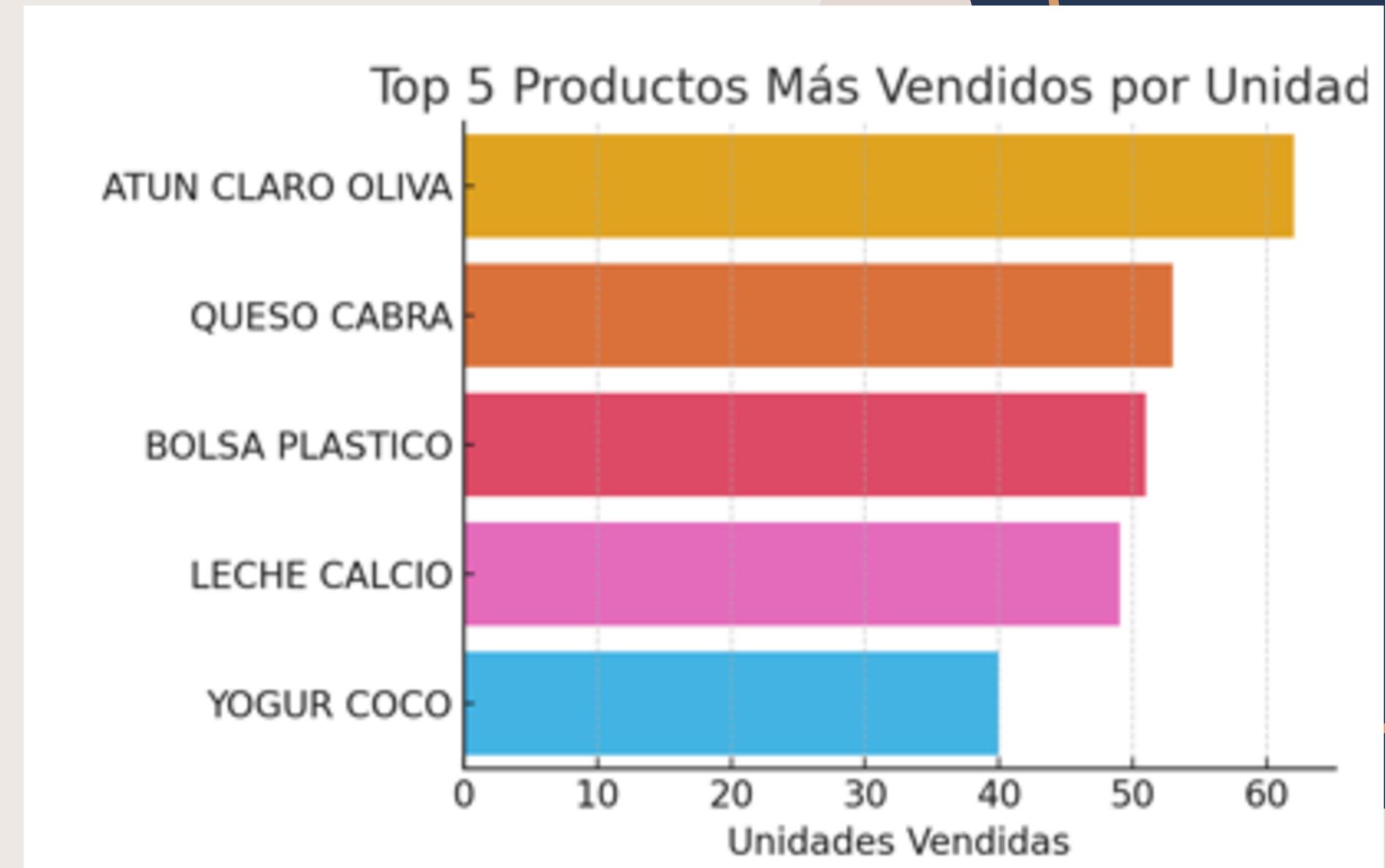
Relación entre el
Número de Productos
y el Precio Total por
Ticket



PREGUNTAS

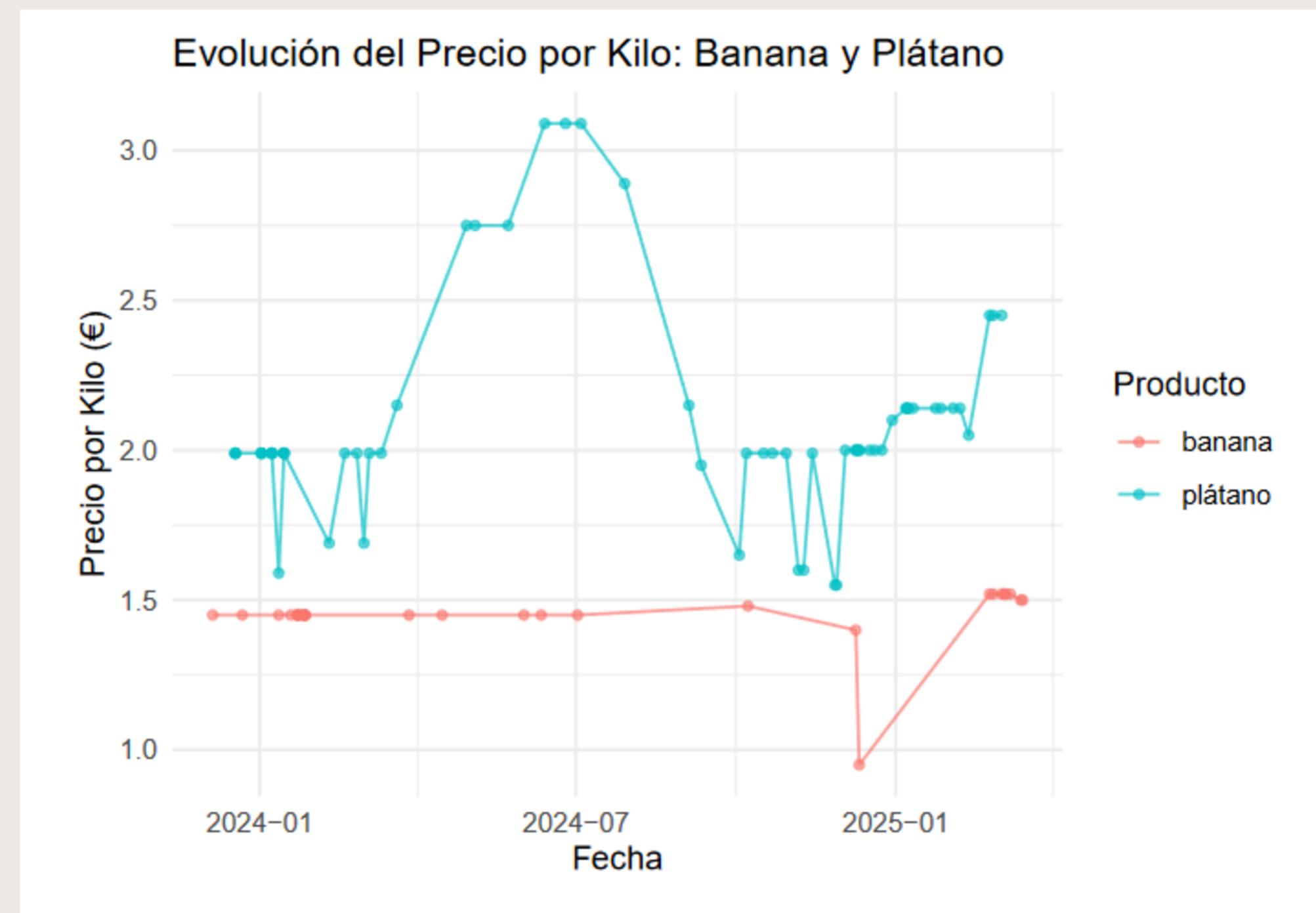
TOP PRODUCTOS POR UNIDADES

- Productos más vendidos:
 1. Atún Claro
 2. Queso de Cabra
 3. Bolsa Plástico
 4. Leche Calcio
 5. Yogur Coco

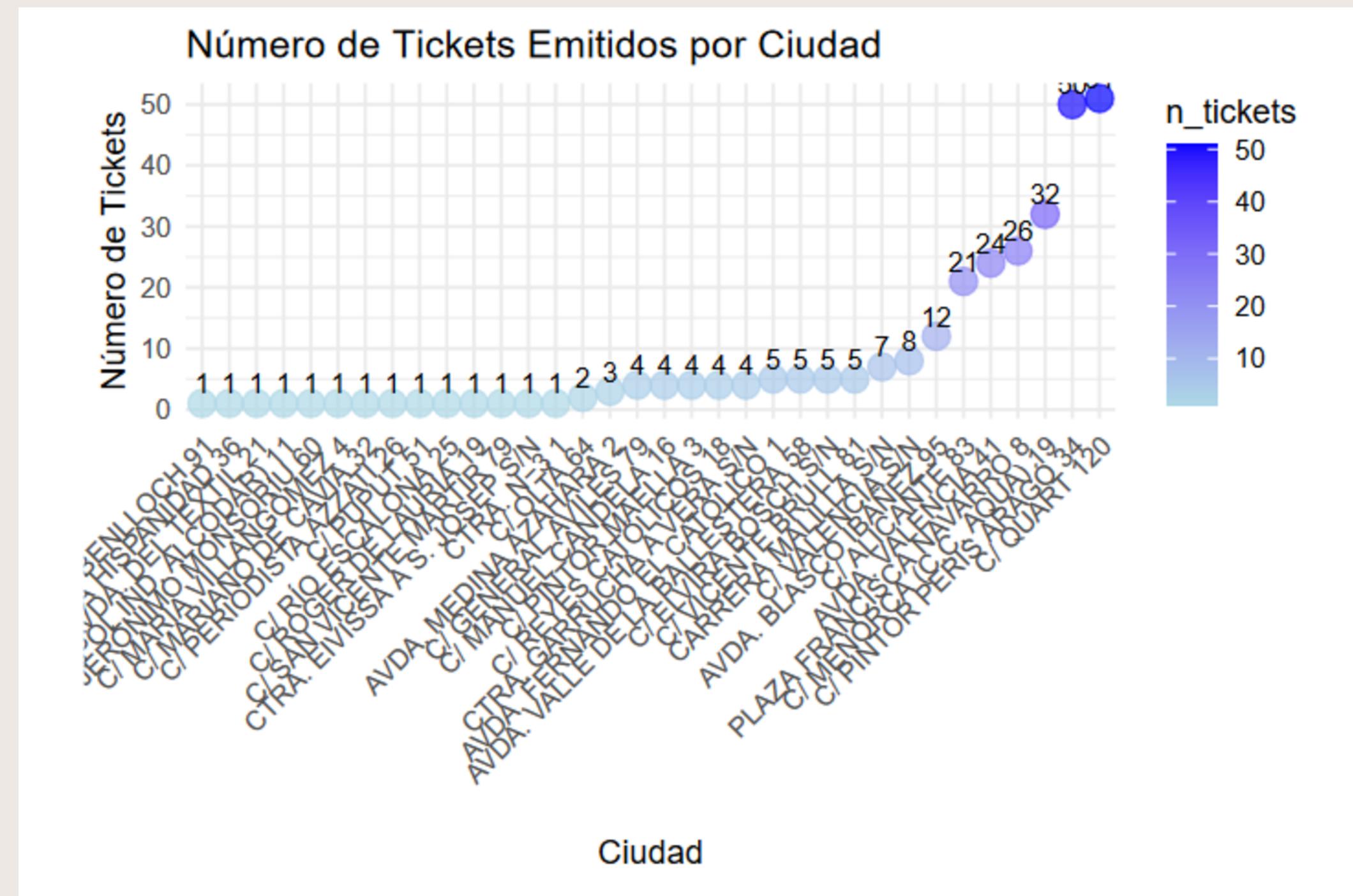


PRECIO BANANA VS PLÁTANO

- Evolución de precios por kilo a lo largo del tiempo.



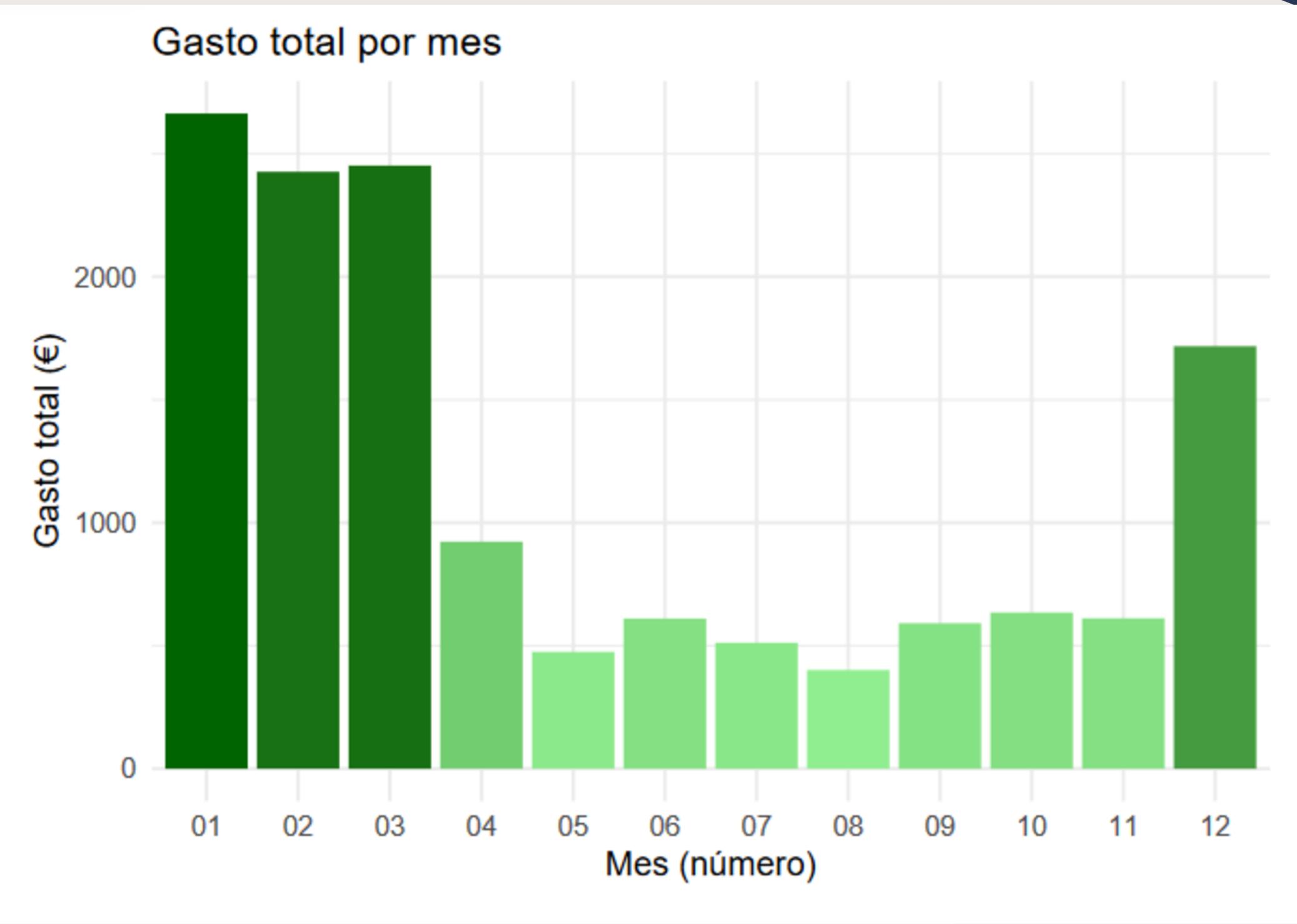
NÚMERO DE TICKETS EMITIDOS POR CIUDAD



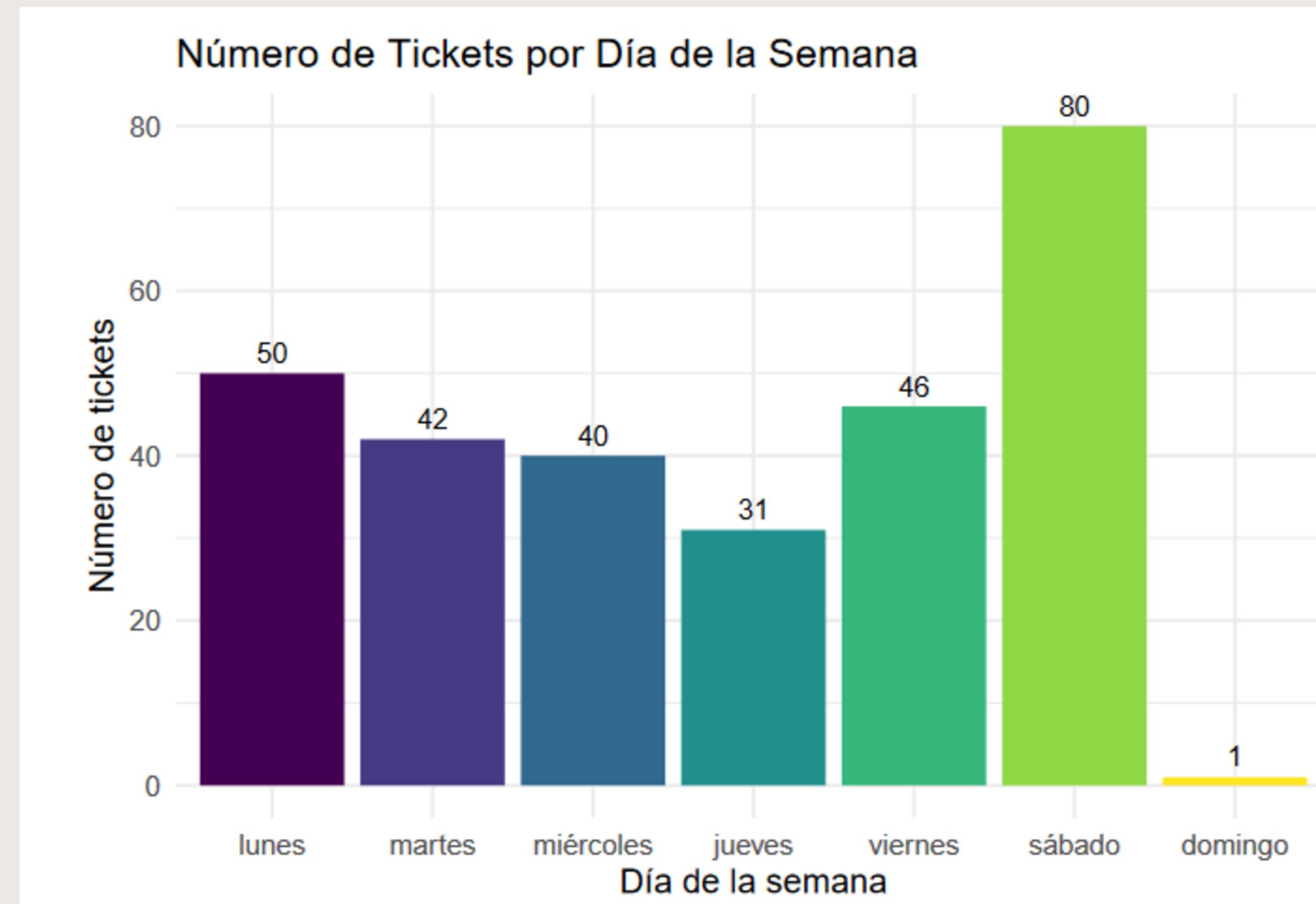
PRODUCTOS QUE SE SUELEN COMPRAR JUNTOS

item1	item2	frequency
leche desnat. calcio	queso lonchas cabra	44
pan de pueblo	queso lonchas cabra	43
brotes tiernos maxi	queso lonchas cabra	40
plátano	queso lonchas cabra	40
queso fresco cabra	queso lonchas cabra	40

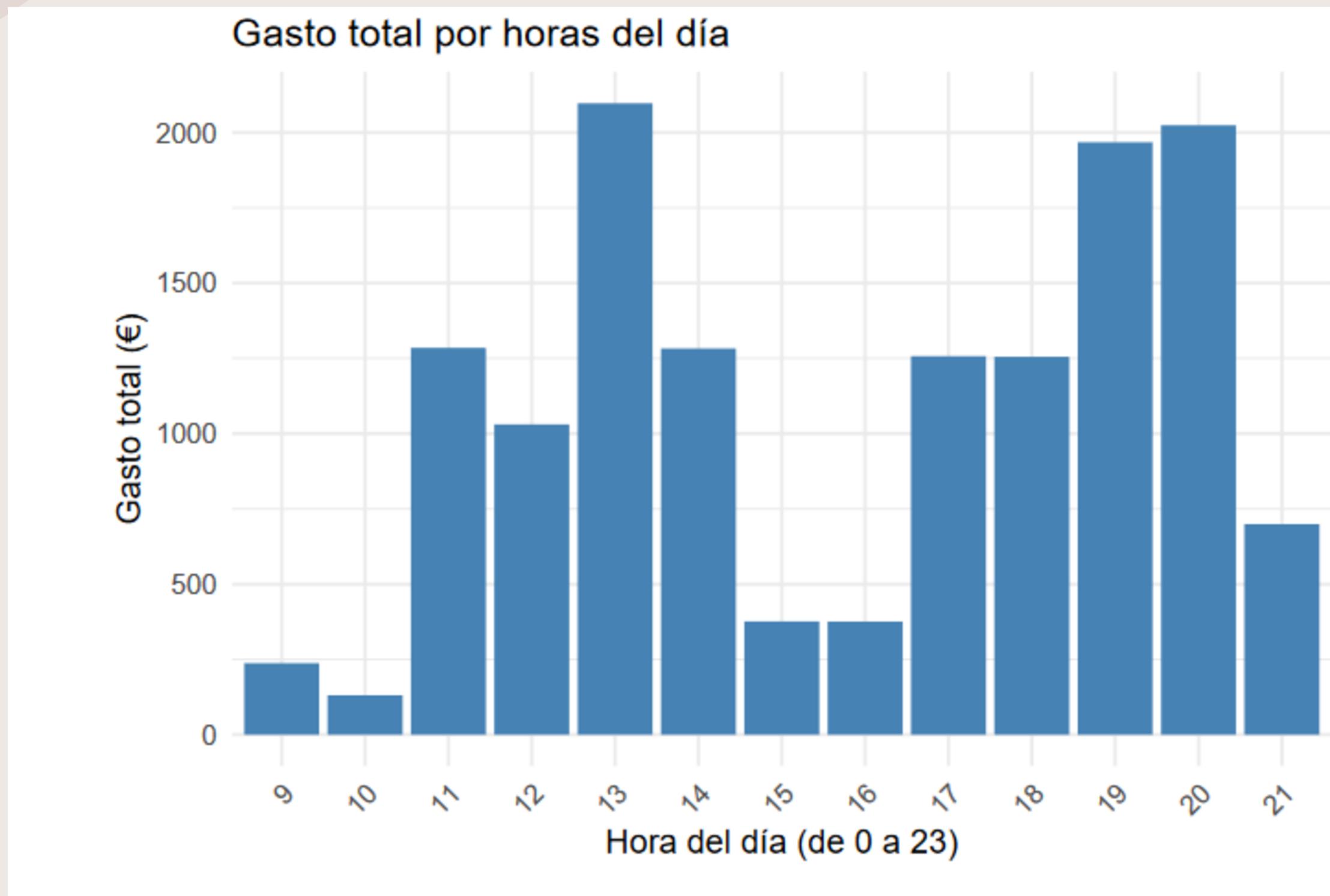
MES QUE MAS DINERO SE GASTA



TICKETS POR DÍA DE LA SEMANA



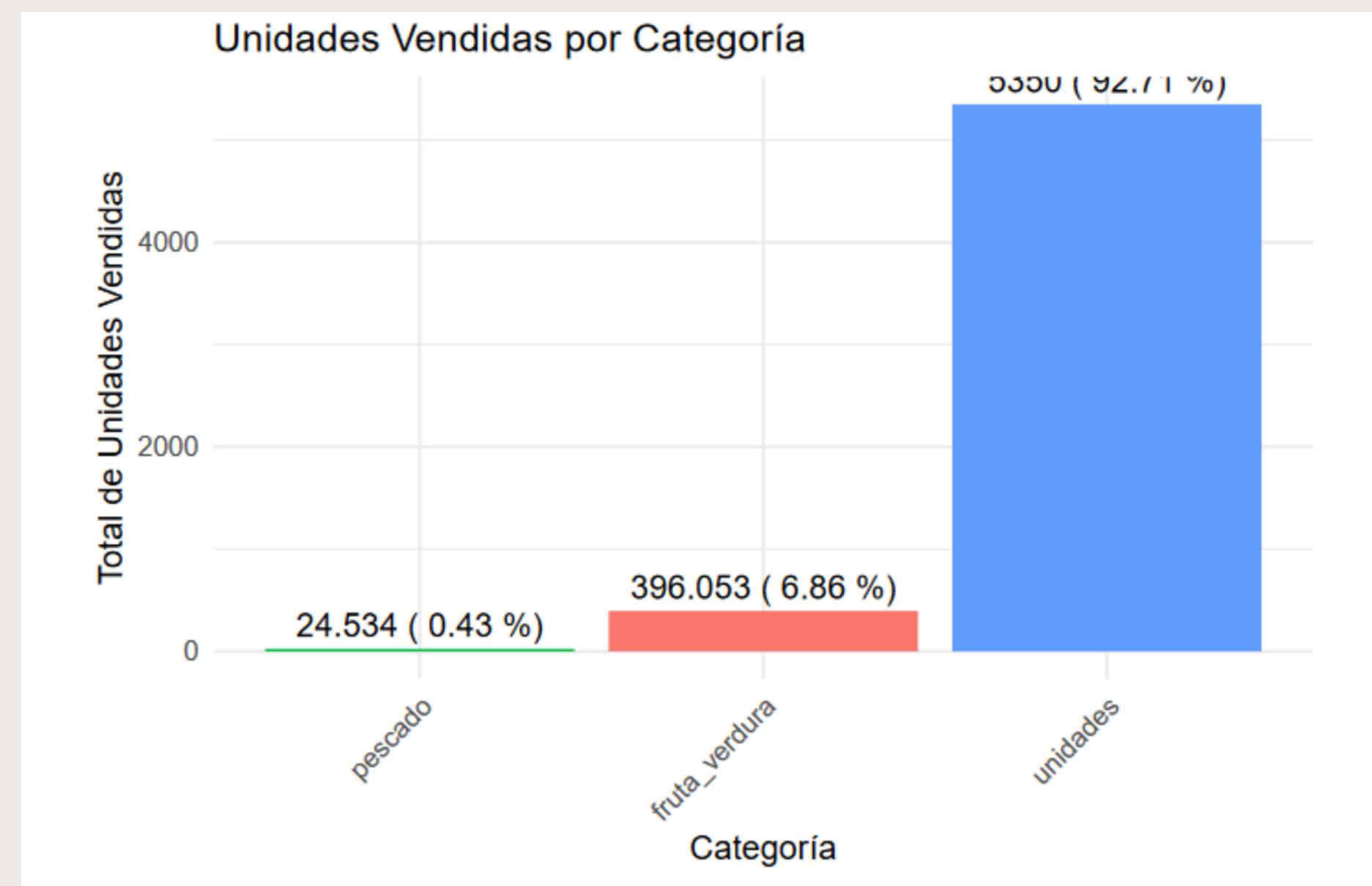
GASTO POR HORA DEL DÍA



DIFERENCIAS DE PRECIOS PARA EL MISMO PRODUCTO EN DIFERENTES TIENDAS

nombre	direccion	precio	n
+ proteína arandanos	AVDA. VALLE DE LA BALLESTERA 58	1.47	2
+ proteína arandanos	PLAZA FRANCISCA NAVARRO 8	1.55	5
+prot natilla choco	C/ ELVIRA BOSCH S/N	1.75	4
+prot natilla vaini	AVDA. BLASCO IBÁÑEZ 95	1.75	5
+prot natilla vaini	AVDA. VALENCIA 41	1.75	1
+prot pla-açaí	PLAZA FRANCISCA NAVARRO 8	1.15	2
+proteínas café	C/ ELVIRA BOSCH S/N	1.90	1
1/2 conejo troceado	AVDA. VALENCIA 41	5.13	1
1/2 conejo troceado	AVDA. VALENCIA 41	6.46	1
1/2 conejo troceado	C/ ALICANTE 83	4.73	1

CATEGORÍA DE PRODUCTOS CON MAYOR VENTA EN TÉRMINOS DE CANTIDAD



CONCLUSIONES

- identificado qué productos son los más y menos vendidos
- diferencias de precios del mismo producto entre tiendas
- permitido observar variaciones en el comportamiento de compra a lo largo de la semana, el mes y el día
- combinaciones frecuentes de productos