



# TRABAJO PRÁCTICO FINAL

Minería de datos aplicada al  
análisis de navegación en un  
e-commerce

Lopez - Saldaña

# Introducción a los datos

- Tienda online de ropa para embarazada
- Sede en Polonia
- 5 meses del año 2008
- Clientes de varios países
- CSV + TXT explicativo



# Estructura del Dataset

- 165.474 observaciones
- 14 variables
- Cada fila representa un clic
- Cada sesión con múltiples clics

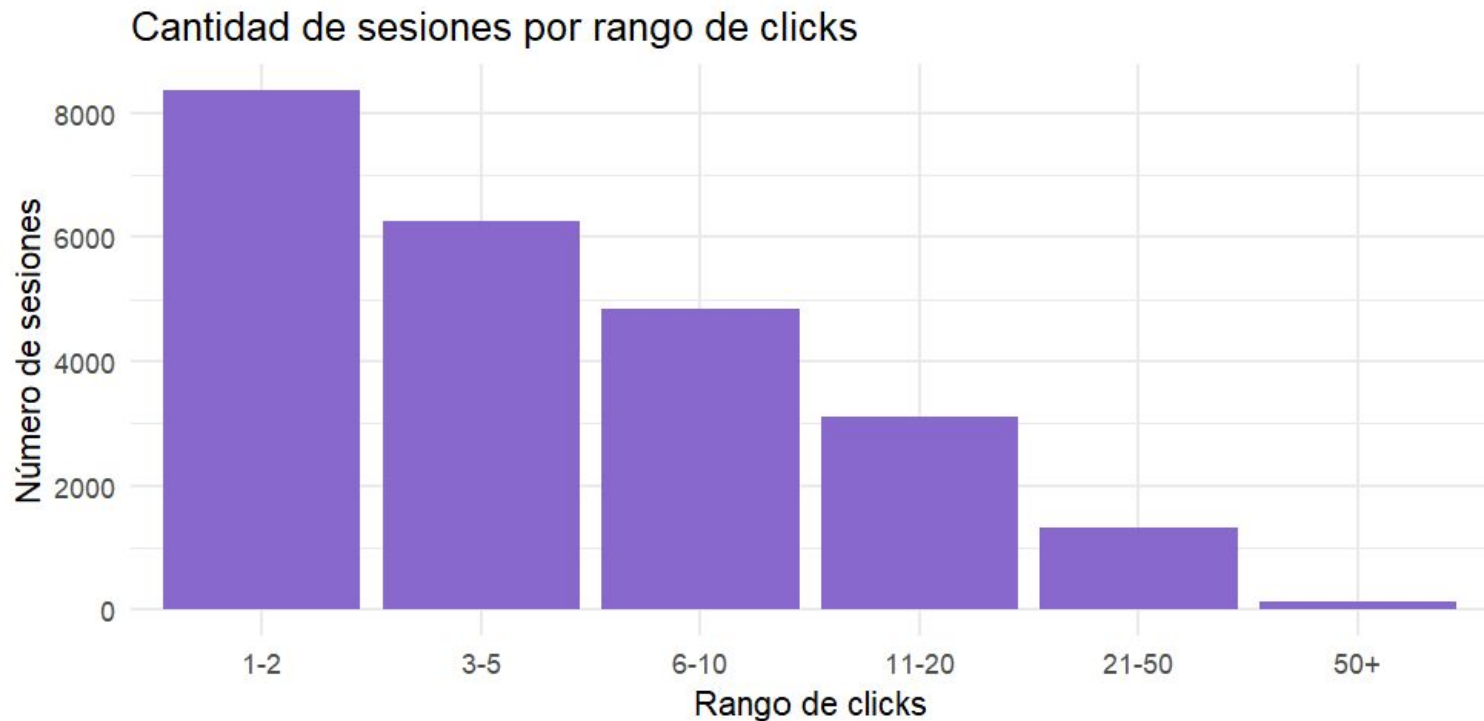


# Exploración del Dataset

- Exploración inicial (head, str, summary)
- Sin valores nulos
- Variables numéricas y categóricas
- Precios entre 18 y 82 USD
- Hasta 195 clics por sesión

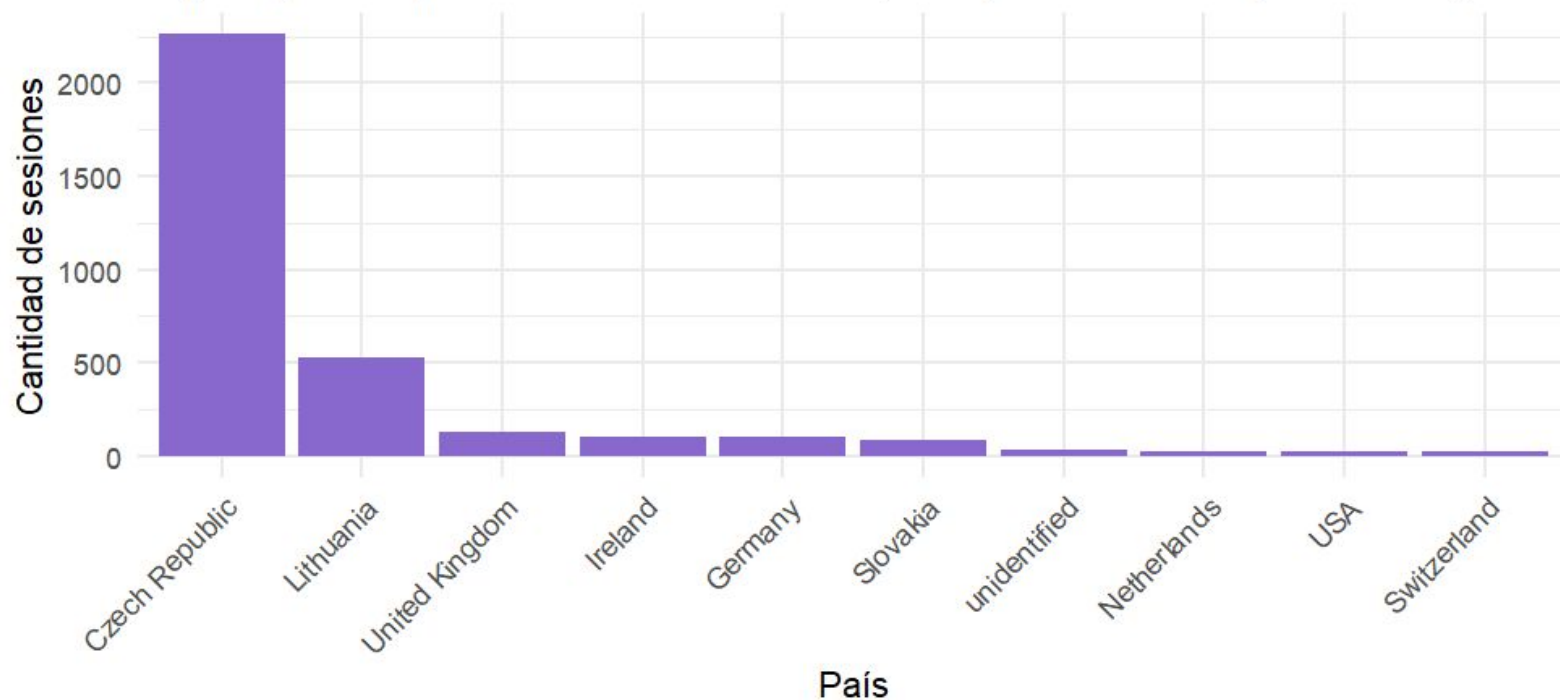


# Clicks por sesión

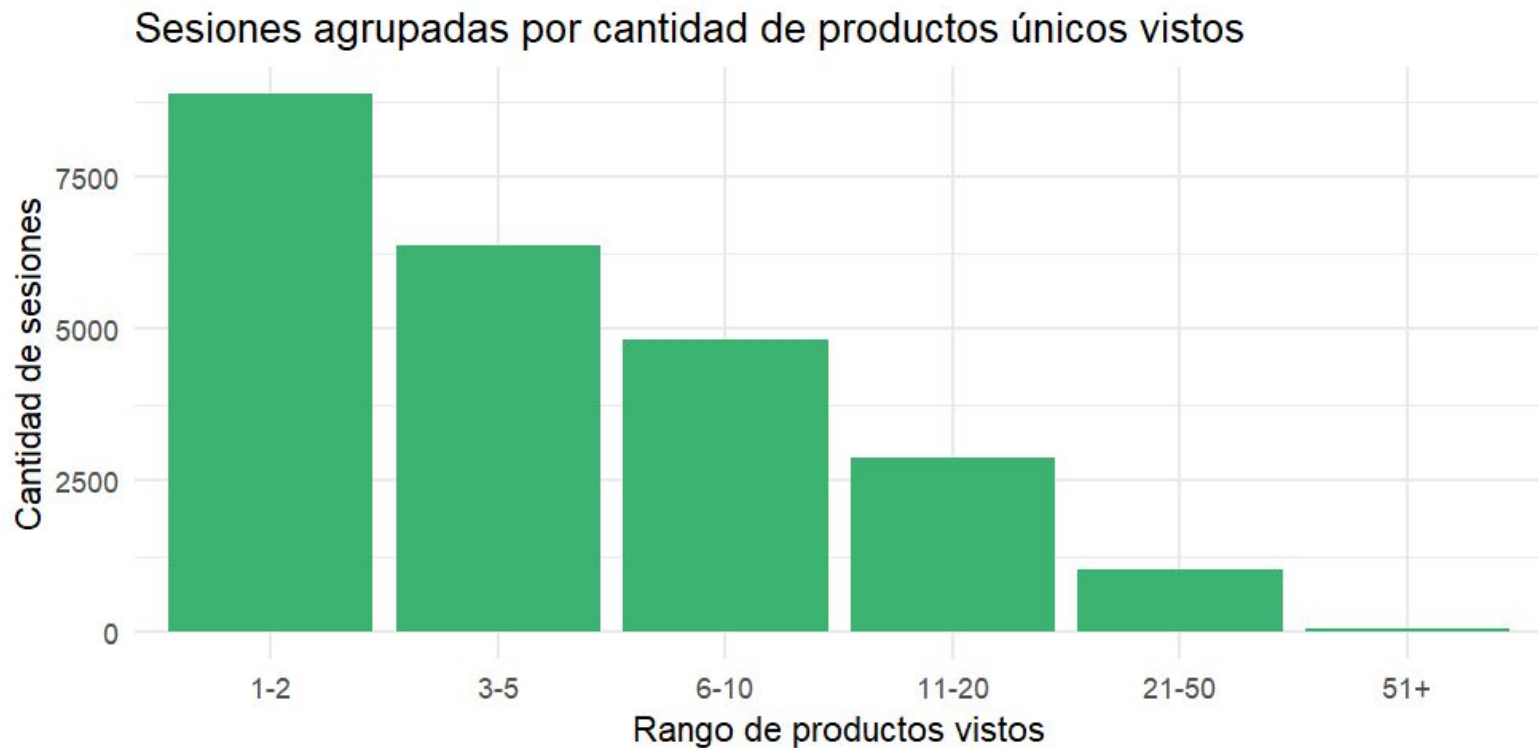


# Sesiones por país

Top 10 países por número de sesiones (excluyendo Polonia y dominios)



# Productos vistos por sesión



# Categorías de producto por sesión

Treemap de categorías de productos más vistas por sesión

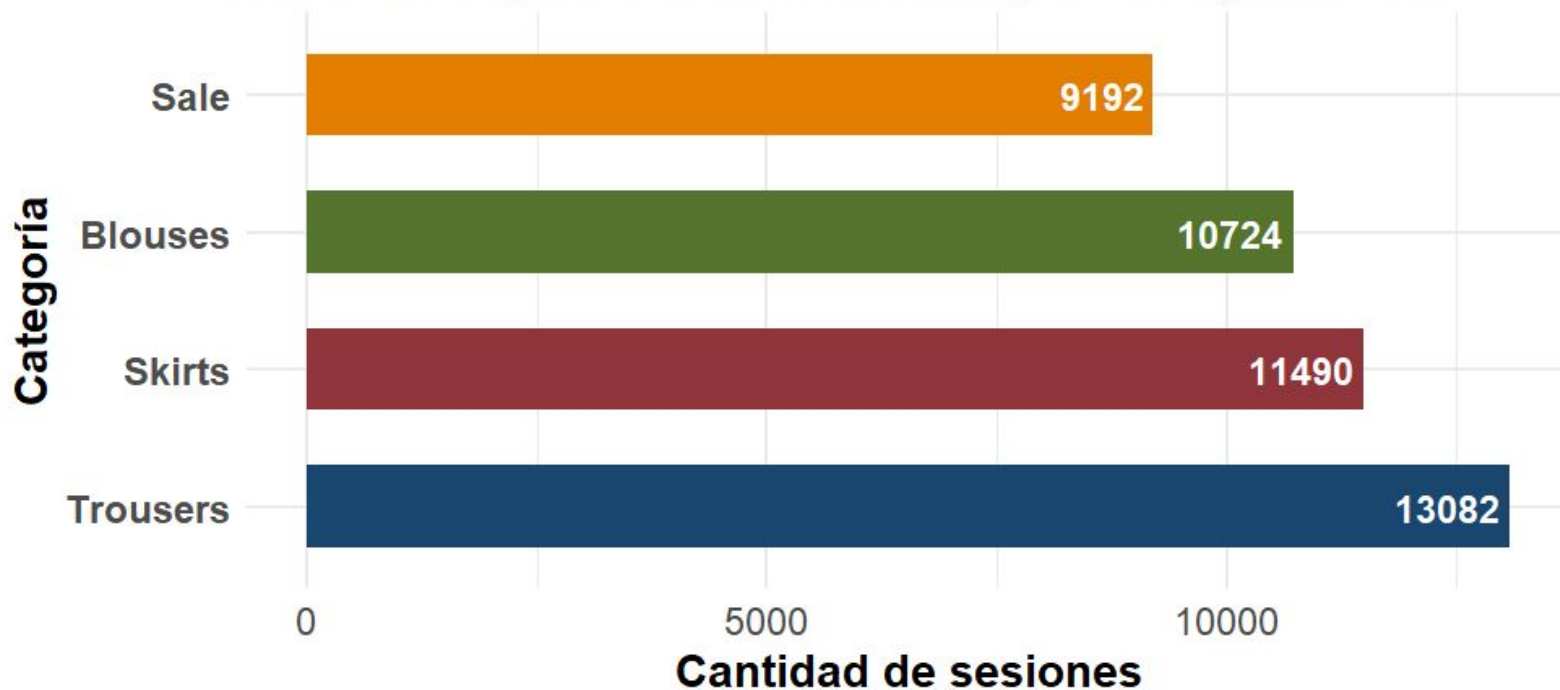


Cantidad de categorías 1 2 3 4



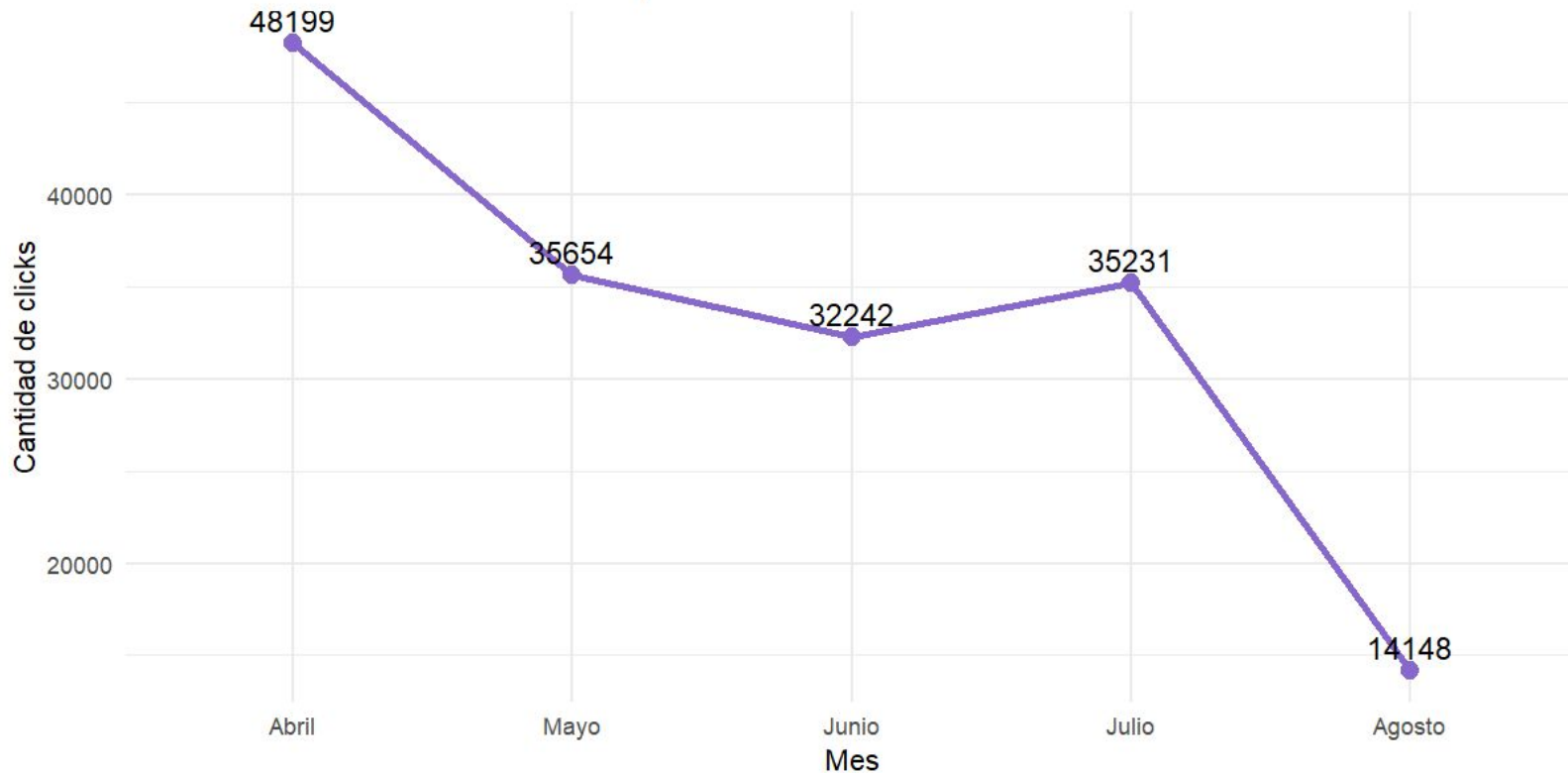
# Categorías de productos más vistas

Sesiones que vieron cada categoría de producto



# Evolución de clicks a lo largo de los meses

Evolución de la cantidad de clicks por sesión



# **Análisis de patrones de compras**

## **General:**

- 24026 sesiones de navegación
- 217 modelos de ropa distintos
- Algoritmo Apriori - soporte mínimo 2%
- A2, A5, A3, A11 y A1 productos centrales

# Análisis de patrones de compras

Polonia:

- Fuertes asociaciones entre C12 - C17 y C56 - C57
- C17 punto central
- 16 reglas de asociación



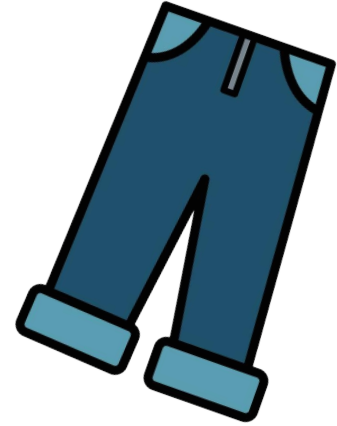
# Análisis de patrones de compras

## República Checa:

- Regla que más se destaca: C57 - C56
- Fuertes asociaciones entre C50 - C49 y C29 - C40
- 7 reglas de asociación



# Análisis de secuencias frecuentes



- Algoritmo cSPADE
- Secuencias frecuentes:

$A2 \rightarrow A5$  ,  $A1 \rightarrow A2$ ,  $A2 \rightarrow A3$  y  $A2 \rightarrow A11$  (Pantalones)

$B10 \rightarrow B13$  (Polleras)



# Conclusión

- Sesiones breves, pocos clics por usuario
- Navegan dentro de una sola categoría
- Diferencias claras entre países analizados



**¡MUCHAS  
GRACIAS!**