

Task 8

Tema: Market Basket Analysis (Panamericana)

Profesor: Eduard F. Martínez González

Octubre 2025

Introducción

El presente taller tiene como propósito aplicar técnicas de **análisis de canasta de mercado (Market Basket Analysis)** a partir de un conjunto de transacciones realizadas en la librería Panamericana.

Este tipo de análisis busca descubrir **patrones de compra conjunta y asociaciones entre productos** utilizando el algoritmo **Apriori**. A partir de las transacciones observadas, el objetivo será identificar reglas del tipo:

$$\{\textit{Producto A}\} \Rightarrow \{\textit{Producto B}\}$$

que permitan interpretar cuáles artículos tienden a adquirirse de manera conjunta y con qué frecuencia o probabilidad ocurre dicha relación.

El resultado de este análisis es fundamental para el diseño de estrategias comerciales, promociones cruzadas y distribución de productos dentro de un punto de venta.

Instrucciones

- El acceso al repositorio del taller está disponible en el siguiente enlace:
- La entrega debe incluir:
 - Un archivo en formato `.R` con el desarrollo del taller.
 - Un archivo en formato `PDF` con las respuestas e interpretaciones solicitadas.
- El script en R debe contener:
 - Nombres de los integrantes y versión de R utilizada.
 - Carga de librerías necesarias (`dplyr`, `arules`, `arulesViz`, etc.).
 - Código organizado y documentado con comentarios: `## Punto 1`, `## Punto 2`, etc.

```

## Nombre(s) de Autor(es)
## R version 4.5.0

rm(list=ls()) # limpiar entorno

## cargar librerias
require(dplyr)
require(arules)
require(arulesViz)

# cargar datos
df <- read.csv("data_panamericana.csv")

# vista previa
head(df)

# crear lista de productos por transaccion
lista <- strsplit(df$productos, ",")

# convertir al formato transaccional
transa <- as(lista, "transactions")

# verificar estructura
summary(transa)

```

Taller: Análisis de Canasta de Mercado – Panamericana

1. (15 pts) Revise la base de datos y explique brevemente la estructura del conjunto **transa**: número de transacciones, número de productos distintos y densidad promedio.
2. (20 pts) Aplique el algoritmo **Apriori** con los parámetros base: soporte mínimo = 0.1, confianza mínima = 0.4, y longitud mínima = 2. Reporte el número total de reglas generadas.
3. (20 pts) Inspeccione las primeras 10 reglas obtenidas con la función **inspect()** e interprete dos de ellas en lenguaje natural, explicando su significado comercial.
4. (15 pts) Identifique y elimine reglas redundantes utilizando **is.redundant()** y conserve únicamente las reglas informativas. Explique brevemente el criterio de eliminación.
5. (15 pts) Visualice las reglas finales mediante un gráfico de dispersión (**plot()**) con el soporte en el eje X, la confianza en el eje Y y el sombreado dado por el **lift**. Describa lo que representa este gráfico.
6. (15 pts) Interprete los hallazgos más relevantes:
 - ¿Qué productos aparecen con mayor frecuencia en las asociaciones?

- ¿Qué combinaciones podrían ser útiles para promociones o estrategias de venta cruzada?

Total: 100 pts