



# Universidad Ricardo Palma

RECTORADO

PROGRAMA DE ESPECIALIZACIÓN EN CIENCIA DE DATOS

*Formamos seres humanos para una cultura de paz*

## PROGRAMA DE ESPECIALIZACIÓN DATA SCIENCE NIVEL I



R + Python

MÓDULO VIII



python™

**Modelos Semi - Supervisados :**  
**RFM – Analisis Cesta Canasta**



## A nuestro recordado Maestro

**Dr. Erwin Kraenau Espinal, Presidente de la Comisión de Creación de la Maestría en Ciencia de los Datos**



**PROGRAMA DE ESPECIALIZACIÓN EN DATA SCIENCE NIVEL I**

« Para poder **seguir** a veces hay que **empezar de nuevo**»



# Agenda

- Segmentación de Clientes.
- Parámetros de la Segmentación de Clientes
- Life Time Value.
- RFM.
- Ventajas Segmentación RFM.
- Análisis de Asociaciones :  
Canasta de Mercado





# SEGMENTACIÓN DE CLIENTES



# SEGMENTACIÓN DE CLIENTES

- Es el proceso de dividir **clientes** en grupos basados en características comunes para que las compañías puedan mercadear cada grupo efectiva y apropiadamente.
- Los grupos o segmentos deben ser homogéneos intragrupos y heterogéneos intergrupos.



# SEGMENTACIÓN DE CLIENTES

## Demográfica

Edad, género, nivel de estudios, nivel de ingresos, etc.

## Geográfica

Lugar de residencia

## Psicográfica

Estilo de vida, valores, personalidad, intereses

## Sociocultural

Ciclo de vida familiar, clase social, valores culturales



# Parámetros de Segmentación

- **Objetivos de rentabilidad por cliente:** o cómo optimizar el rendimiento de los clientes que ya tenemos reduciendo variables de alto impacto como por ejemplo la tasa de abandono.
- **Definición del formato del negocio:** dependiendo de a qué segmento de clientes nos dirijamos primordialmente, nuestro negocio presentará un formato u otro.
- **Optimización de la oferta:** si sabemos cuál es nuestro segmento de cliente prioritario y qué valora, podremos realizar cambios sustanciales en nuestra oferta de producto para responder a las necesidades de ese segmento. Es una decisión estratégica importante.
- **Lanzamiento de nuevos productos o servicios:** el conocimiento estratégico de nuestros clientes nos permitirá prever si nos resultará rentable el lanzamiento de determinados productos o de nuevos servicios integrados en nuestra oferta existente.





# LIFE TIME VALUE



**PROGRAMA DE ESPECIALIZACIÓN EN DATA SCIENCE NIVEL I**

# Lifetime Value

- El *Customer Lifetime Value*(CLV), o Valor vitalicio del cliente(VVC) es una de las métricas comúnmente asociadas con el CRM(*Customer Relationship Management*).
- Se trata de algo aparentemente tan simple como estimar el valor que representa para la empresa la relación con un cliente, a lo largo de la vida de dicha relación.



# Lifetime Value

- En términos sencillos se puede decir que el lifetime value es el valor presente de la ganancia a futuro que un cliente generaría para nuestra organización.



# Lifetime Value

- **Margen (m):** El margen es la ganancia anual que genera un cliente menos los costos de operación.
- **Retención de clientes (r):** Si asumimos que un cliente provee un flujo de ganancia constante para la organización, se podría decir que es leal a nuestros productos y/o servicios.
- **Costo de capital (c):** El costo de capital depende de la estructura financiera de una compañía.

$$\text{Customer Lifetime Value} = m (r / (1 + c - r))$$



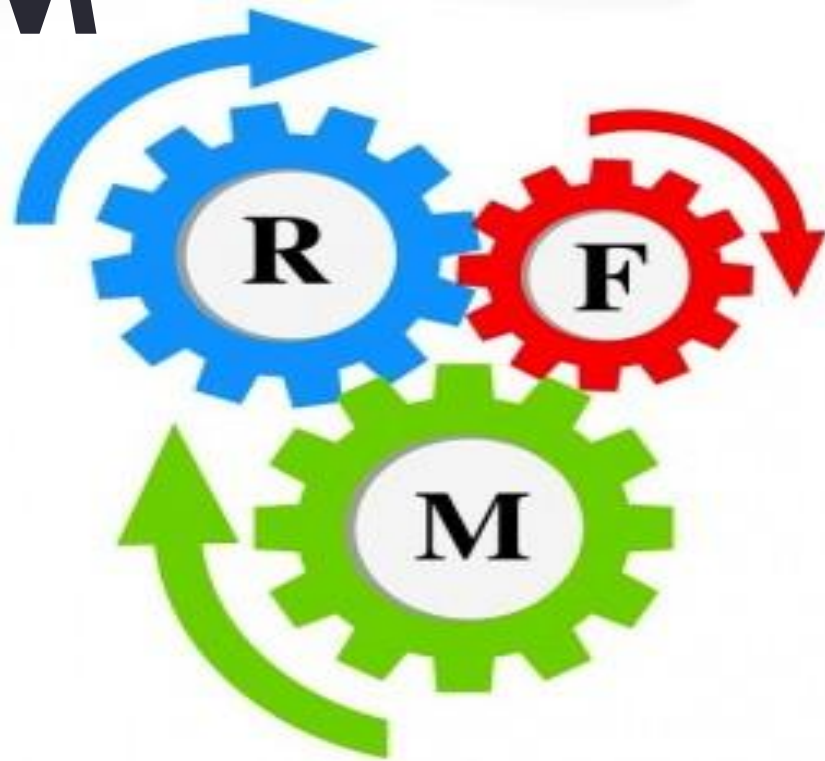


# Lifetime Value : Importancia

- En primer lugar, el CLV tiene una utilidad directa y evidente: resulta indispensable conocerlo antes y después de aplicar cualquier estrategia de CRM, para ser capaz de evaluar y monitorizar los resultados de la misma. Es decir conocer el tiempo de duración de la relación con el cliente y cuanto genera esto para la empresa.



# Segmentación RFM



PROGRAMA DE ESPECIALIZACIÓN EN DATA SCIENCE NIVEL I

# Segmentación rfm

- El análisis RFM (por Recency, Frequency, Monetary) es una técnica relativamente reciente de mercadeo usada para determinar cuantitativamente cuales clientes son los que invierten más en una empresa.
- Esto se consigue examinando lo que el cliente ha comprado utilizando tres factores: (R) Recientes compras, (F) Frecuencia de compra y (M) Monto de la compra en términos monetarios.



Recency -  
How recent was the last purchase?



Frequency -  
How many purchases have been made?



Monetary Value -  
What is the value of the purchases?

copyright © Jigsaw Academy Education Pvt. Ltd.

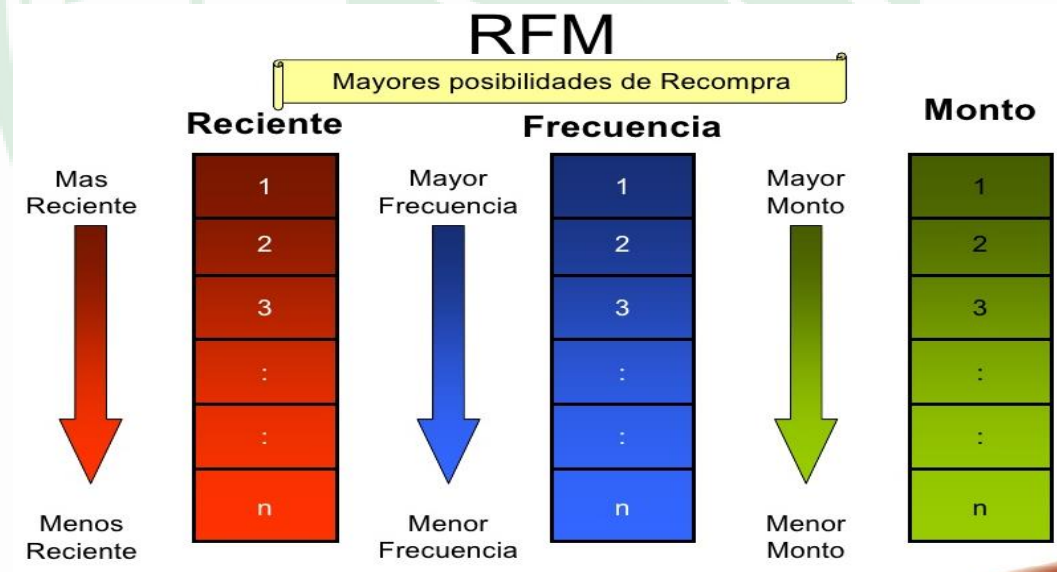
# Segmentación rfm

- El mejor predictor del futuro es el comportamiento pasado.

**Recency** – tiempo desde última compra

**Frequency** – tiempo entre compras

**Monetary Value** – monto de compras





# Rfm : Funcionamiento

- **Como Funciona:**
- Recency: Con la última información de compra de los clientes:
- Se divide la base de datos en 5 segmentos iguales (quintiles).
- -Número 5 ( Para el más antiguo)
- -Número 1 (Es el más reciente )



# Ventajas : Segmentación RFM

- Para la gestión de clientes, el análisis RFM balancea las tres dimensiones características mas potentes.
- RFM es de fácil comprensión.
- RFM puede aumentar la performance de campañas o promociones de marketing directo al apuntar a clientes del segmento más alto de todas las dimensiones.
- Se Puede identificar y ver la evolución en el tiempo de mis clientes.
- Permite focalizar las medidas de fidelización de clientes.



# Análisis de Canasta de mercado



# Análisis de Canasta (Market Basket Analysis)

Los hábitos de compra de los clientes pueden ser representados a través de asociaciones o correlaciones entre los diferentes productos que compran en sus “canastas”.



Cliente 1:  
Arroz, puré,  
bebida



Cliente 2:  
Arroz, helado,  
pan



Cliente 1:  
Arroz, bebida,  
cerveza





# Asociación de Productos



+



=



PROGRAMA DE ESPECIALIZACIÓN EN DATA SCIENCE NIVEL I

# Análisis de Canasta: Indicadores

- Las relaciones entre productos se miden por:

*Importancia Relativa (support):* indica el porcentaje de transacciones que llevan el antecedente y el consecuente, con respecto al total de transacciones analizadas

Ej:

{arroz, puré, cerveza}, {arroz, helado, pan}, {arroz, bebida, cerveza}

*Importancia Relativa(arroz → cerveza) =  $2/3 = 66\%$*

Este indicador señala que tan frecuente es la relación entre productos con respecto al universo de boletas. Pero no indica si existe vínculo entre ambos productos.



# Análisis de Canasta: Indicadores

*Confiabilidad (confidence): indica el porcentaje de transacciones que llevan el antecedente y el consecuente juntos, con respecto al total de transacciones que llevan el antecedente.*

*Ej:*

*{arroz, puré, cerveza}, {arroz, helado, pan}, {arroz, bebida, cerveza}*

*Confianza (cerveza → puré) =  $1/2 = 50\%$*

*Esta relación señala el vínculo o sinergia entre ambos productos (probabilidad condicional).*

*Pero, qué pasa en el siguiente caso:*

*Confianza(helado → arroz) =  $1/1 = 100\%$*

*¿Quiere decir que hay una fuerte relación entre estos productos?*



# Análisis de Canasta: Indicadores

*Ganancia (gain or Improvement):* puntaje que representa el aumento en la probabilidad de selección del consecuente, al ser comprado en conjunto con el antecedente

$$\text{Ganancia } (A \rightarrow B) = \text{Confianza}(A \rightarrow B) / \text{Importancia Relativa}(B)$$

*Ej:*

{arroz, puré, cerveza}, {arroz, helado, pan}, {arroz, bebida, cerveza}

$$\text{Ganancia}(\text{helado} \rightarrow \text{arroz}) = 1/1 = 1$$

$$\text{Ganancia}(\text{bebida} \rightarrow \text{puré}) = 0,5/0,3 = 1,5$$

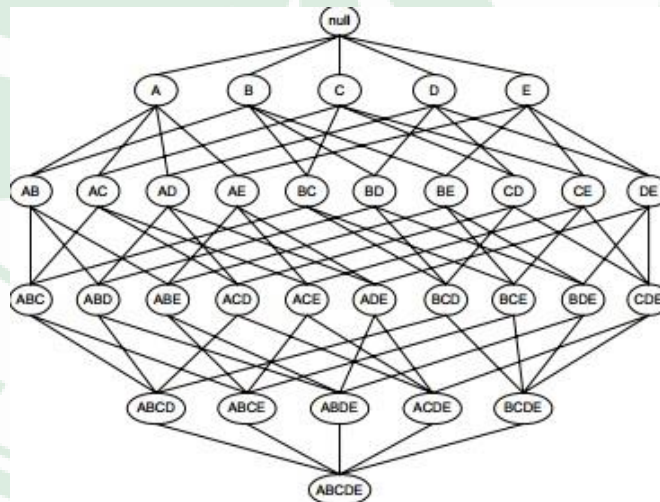
Con estos indicadores podemos entender las relaciones entre clientes





# Algoritmo APRIORI

- Algoritmo de Búsqueda de Asociaciones de Productos (Agrawal, 1994).
- Dado valores de confianza e importancias relativas, busca las relaciones entre productos que cumplan esos requisitos.
- Tiempos de solución razonables
- Encuentra sólo las relaciones más fuertes entre productos y no las de un producto en particular.



# Uso y aplicación : market basket analysis

- Las relaciones entre productos permiten apoyar decisiones como:
  - Armado de Pack's (Consulta de Canasta)
    - Entre 2 productos de alta rotación de distintas categorías pero del mismo proveedor + un producto de baja rotación y alto margen.
    - Entre 2 productos de alta rotación de categorías y proveedores distintos + un producto de baja rotación y alto margen.
  - *Descuentos: Políticas agresivas para productos que se venden juntos, con alta confiabilidad.*
  - Distribución en sala de Categorías – Diseño de Layout
  - Relaciones entre proveedores y sus productos
  - Relaciones entre líneas y sublíneas de productos



# Beneficios : market basket analysis

- Identificación de preferencias y gustos de los clientes
  - Reforzar esas preferencias
  - Incentivar a otros clientes a tomar estas preferencias
- Conocimiento encontrado:
  - Identificación de la “canasta típica” de los clientes de la cadena.
  - Los clientes compran habitualmente productos sustitutos por ejemplo:
    - Omo y Drive
    - 1+1 y yogurt batido
  - Se encuentran relaciones entre productos de distintas categorías por ejemplo:
    - Detergente con Yogurt
    - Detergente con Leche
    - Yogurt con Mayonesa



# Beneficios : market basquet analysis

- El sistema de Análisis de Canasta permitirá:
  - Conocer y entender el comportamiento de compra de los clientes
  - En base a las preferencias, ofrecer ofertas y promociones de acuerdo a las asociaciones
    - Campañas más exitosas
    - Campañas menos costosas
  - Anticipar cambios en las preferencias de los clientes y actuar oportunamente
  - Aumento en las ventas de la cadena
  - Clientes más satisfechos y leales





# ¡Gracias!



**Comunidad Data Science Perú**



**Comunidad Data Science Perú**

**PROGRAMA DE ESPECIALIZACIÓN EN DATA SCIENCE NIVEL I**