



EGADE Business School
Tecnológico de Monterrey

Aplicaciones de analítica de datos a los negocios II

PROF: JUAN C. BUSTAMANTE

JUCBUSTAM@TEC.MX

Normas para la conexión síncrona:

1. La clases tiene un back-up garantizado (Grabación disponible en la nube de Zoom).
2. Ingresar a la clase con la cámara del equipo de computo encendida.
3. La **cámara deberá permanecer encendida a lo largo de la clase.**
4. Al ingresar a la clase deben silenciar el micrófono del equipo de computo.
5. Levantar la mano es una opción cuando se quiere preguntar algo durante la sesión de clase, pero les recomiendo que mejor hagamos uso intensivo del chat del canal general para hacer preguntas.
6. En caso de necesitar hacer una pregunta, puede interrumpir la clase sin problema, activando el micrófono de vuestro equipo de computo, luego de la pregunta desactíVELO nuevamente.
7. Para una buena clase online es indispensable **el debate, así que foméntelo!!!.**
8. Toda la información se gestiona en CANVAS LMS.



Cronograma de trabajo:

Sesiones	Contenidos	Actividad		Fecha
1	Información general del curso	Utility of classification algorithms		Martes 18/04
2	Algoritmo de regresión logística	Ejecutar script	Solución caso: Retention modelling at Scholastic Travel Company (A) and (B)	Martes 25/04
3	Algoritmo Naïve Bayes	Ejecutar script		Martes 02/05
4	Algoritmo k-nearest-neighbors (KNN)	Ejecutar script		Martes 09/05
5	Algoritmo Support vector machine	Ejecutar script		Martes 16/05
6	Algoritmo Decision Trees	Ejecutar script		Martes 23/05
7	Algoritmo Random Forest	Ejecutar script		Martes 30/05
8	Modelo RFM	Ejecutar script	Solución caso: CD Now	Martes 06/06
9	Modelo valor de vida del cliente (I)	Ejecutar script		Martes 13/06
10	Modelo valor de vida del cliente (II)	Ejecutar script		Martes 20/06
11	Análisis de series de tiempo	Ejecutar script		Martes 27/06
12	Proyecto final	Presentación en equipos		Martes 04/07
	Evaluación final			

1

< **Modelo RFM**

RECENCY
FREQUENCY
MONETARY

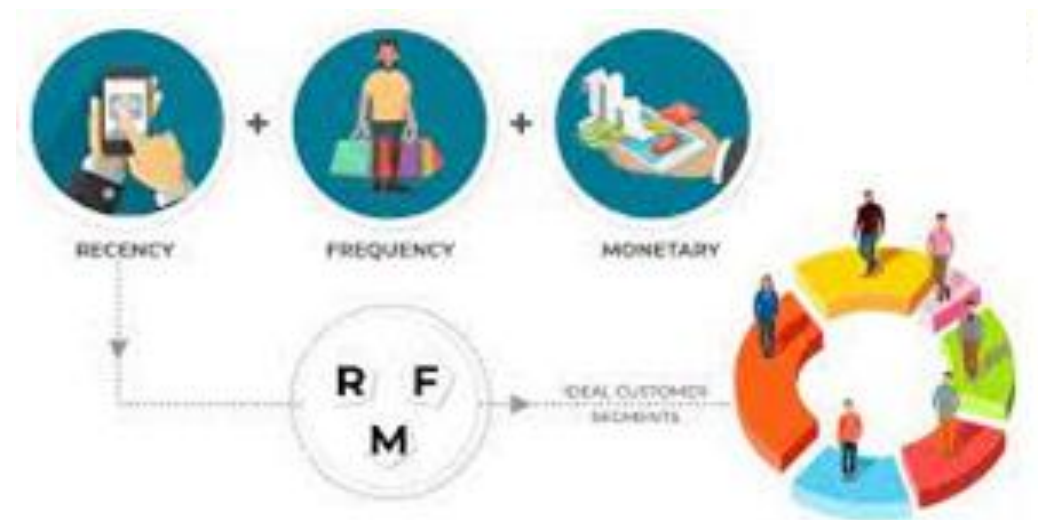


Modelo RFM?



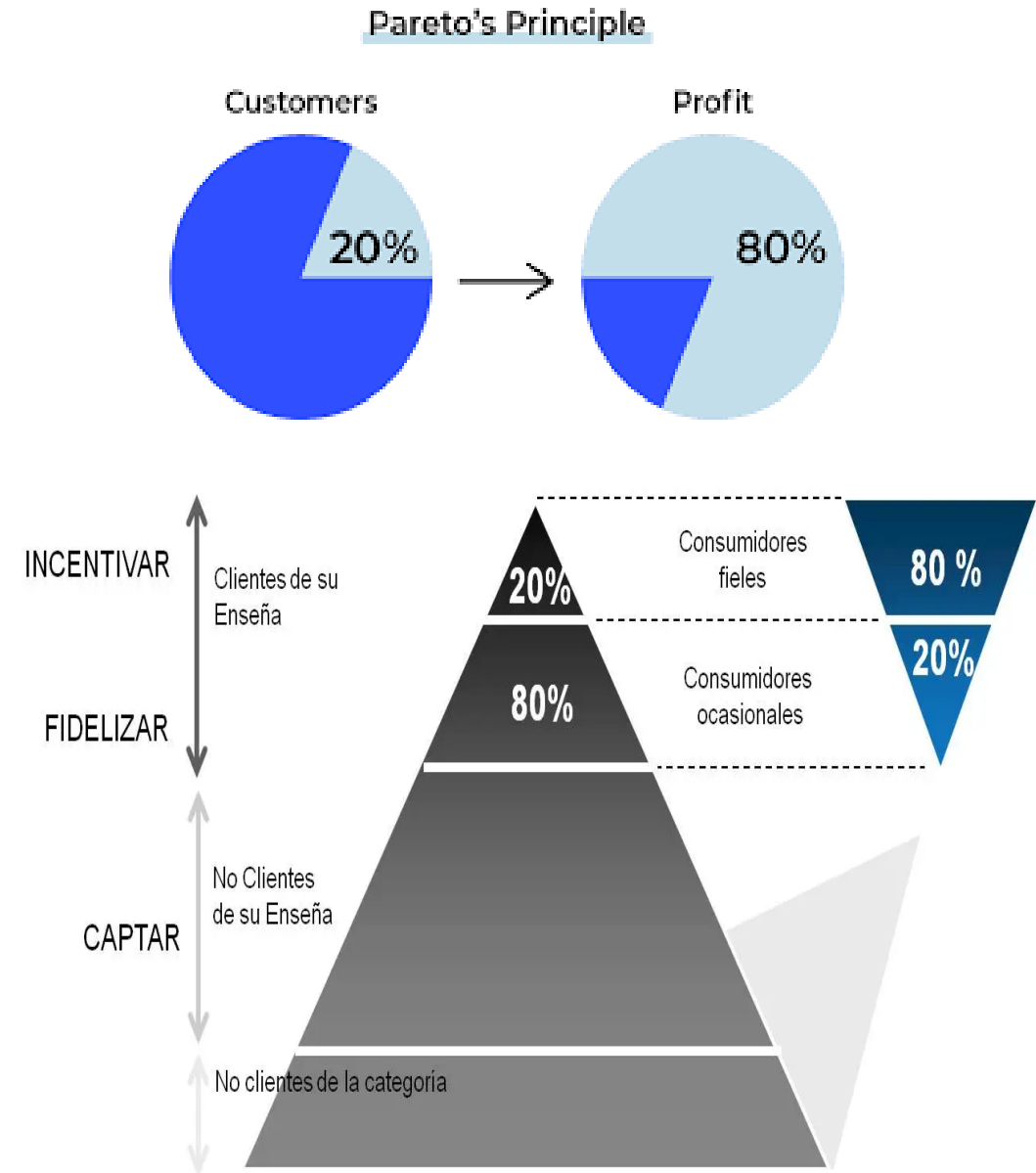
Este modelo contempla lo Reciente, Frecuente y el Valor Monetario para cada cliente a partir del cual determinamos el comportamiento o evolución de compra.

El modelo RFM sigue la premisa de que “los más propensos a comprar son aquellos que han comprado más recientemente, con más frecuencia y gastan más dinero”..



Modelo RFM

- Se basa en la conocida “Ley de Pareto” o del “80/20” enunciada por el economista italiano Vilfredo Pareto, en el siglo XIX.
- El RFM se aplica sobre esta “Ley de Pareto” y se refiere a que “el 80% de las compras las realizan el 20% de los clientes”.



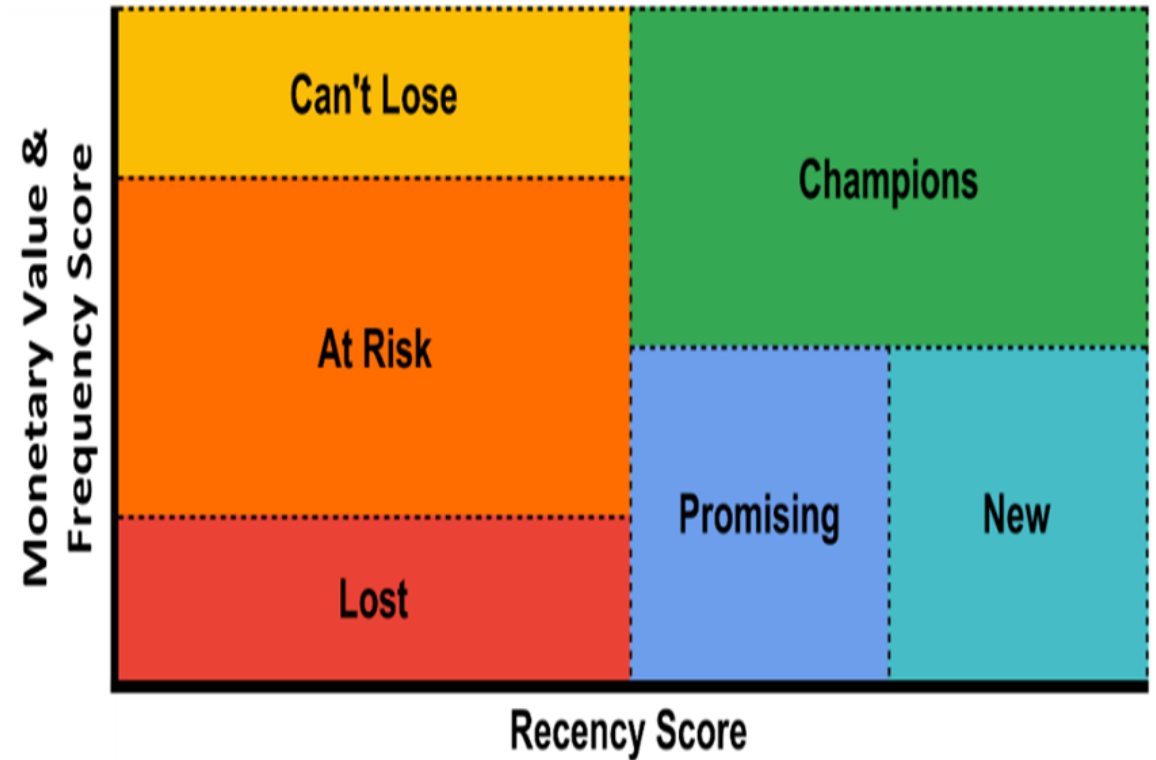
Lógica del modelo RFM

- Con Reciente (Recency) medimos los días que han pasado desde hoy (o cualquier fecha a futuro) hasta la fecha en que realizó su última compra.
- Con la Frecuencia (Frequency) medimos el número de compras que ha hecho cada cliente en total.
- Y el Valor Monetario (Monetary) es la suma total de cantidad de dinero que el cliente lleva gastado en sus compras

Segmentación con RFM



- Con el RFM podemos construir escalas basadas en estas tres variables, dando a cada cliente un valor según el percentil en el que se encuentre.
- Lo más común es escalar por cuartiles ó quintiles, es decir, tanto a Reciente, como a la Frecuencia, como al Valor Monetario les asignamos un valor del 1 al 4 ó 5, siendo 1 la peor puntuación y 4 ó 5 la mejor, como una escala de “Likert”.
- Por tanto, para cada una de las tres variables los clientes cuya última compra sea más reciente, más frecuente y cuya suma de ingresos sea la más elevada tendrán una puntuación de (R=4 ó 5) (F=4 ó 5) (M=4 ó 5).
- Con estas puntuaciones se dividen a los clientes en 25 ó 20%, tratándose de quintiles.



Segmentación con RFM

Segmentos RFM (Hasta $4 \star 4 \star 4 = 64$)

4	Más reciente	4	Más frecuente	4	Valor alto	25%
3		3		3		25%
2		2		2		25%
1	Más antiguo	1	Menos frecuente	1	Valor bajo	25%

Ejemplo: “4 3 1”

Importe bajo
Frecuencia alta
Transacción más frecuente

PREMISAS DEL MODELO RFM

- Los clientes que compran de forma reciente son más favorables a comprar que aquellos que no lo han hecho últimamente.
- Los clientes que compran más frecuentemente están más dispuestos a comprar nuevamente que aquellos que han hecho una o dos compras.
- Los clientes que gastan más, están más dispuestos a comprar nuevamente.
- Los clientes más valiosos son aquellos que pueden llegar hacerlo aún más.

RFM

VENTAJAS

- Modelo muy intuitivo
- Facilita la segmentación de clientes con lógica financiera
- Contextualiza la aportación de cliente en una relación.

DESVENTAJAS

- No sirve para detectar el cambio migratorio de cada cliente de un valor a otro (de un cuartil a otro cuartil) cada vez que se actualice el análisis.
- Además, tampoco detalla cuándo un cliente es nuevo (de hace una semana, o de hace un mes). Se podría interpretar que si la Frecuencia es baja y la Recencia es alta podría ser nuevo, pero lo usual es que los datos se mezclan con clientes con fecha de registro muy antigua que han empezado a comprar este año por primera vez. Realmente no es nuevo en la base de datos pero sí en comprar.

Bibliotecas



Manipulación y análisis de datos



Creación de vectores y matrices
Colección de funciones matemáticas



Generación de gráficos



Biblioteca de visualización basada en Matplotlib



Biblioteca de aprendizaje automático



Dataset Retail
e-commerce