

BASE DE DATOS 2

Dra. Rosanna Costaguta (Prof. Responsable)

Lic. Nevelin Salazar

LSI - UNSE - 2019

CONTENIDO PROGRAMÁTICO

UNIDAD 1: Descubrimiento de conocimientos

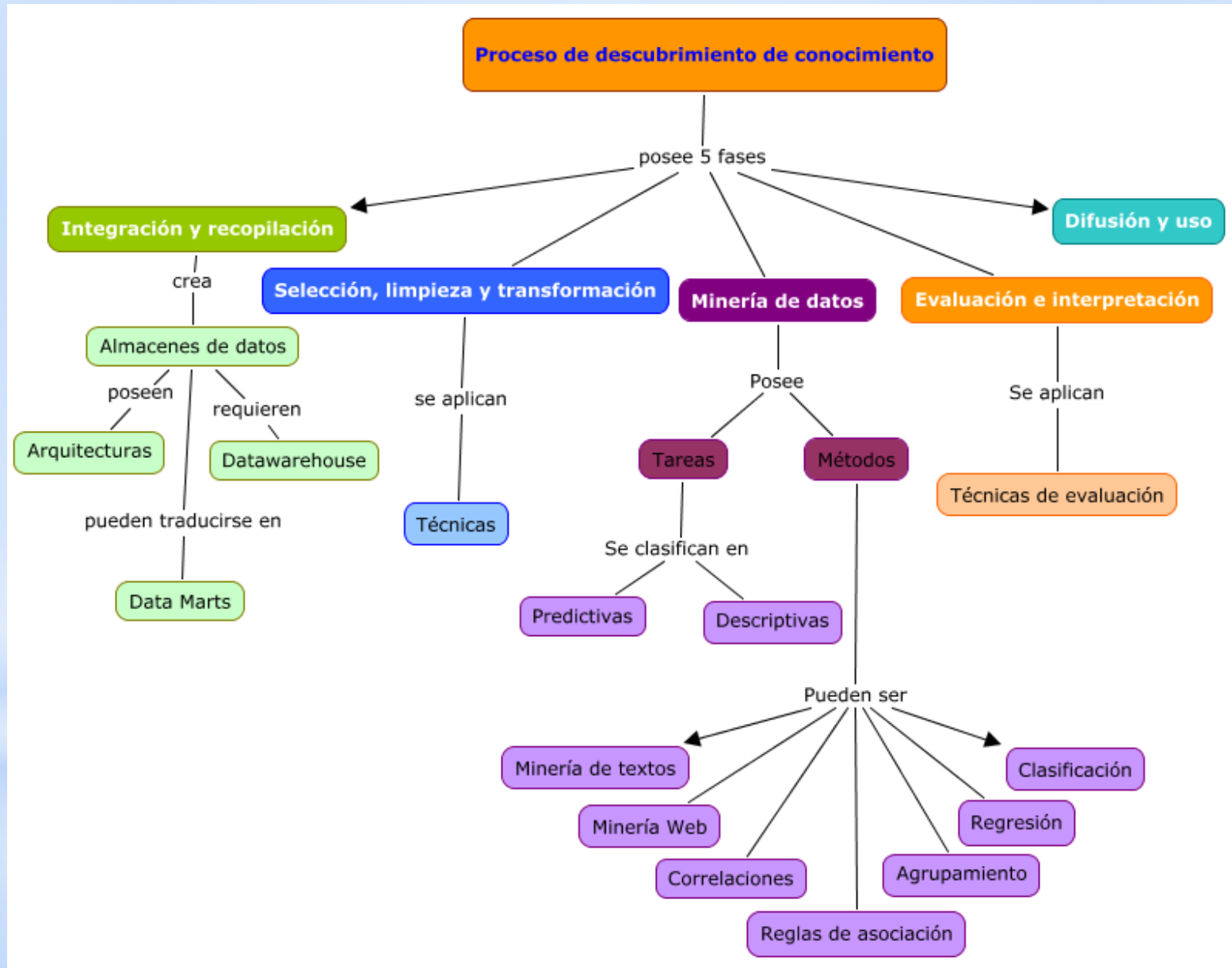


UNIDAD 2: Minería de datos



UNIDAD 3: Gestión de datos masivos

ARTICULACIÓN TEMÁTICA



ACTIVIDADES A DESARROLLAR

- ✓ **Clases**
- ✓ **Actividades colaborativas (Foros, Wikis, etc.)**
- ✓ **Actividades de investigación**
- ✓ **Taller de programación**

EVALUACIONES PREVISTAS

- ✓ **Parcial 1: 3/10**
- ✓ **Recuperatorio 1: 10/10**

- ✓ **Parcial 2: 7/11**
- ✓ **Recuperatorio 2: /1411**

- ✓ **Defensa taller: 5/11**

Para PROMOCIONAR la Asignatura:

- Aprobar los parciales con 7 puntos mínimo.
- Aprobar taller de programación con 7 puntos mínimo.
- Aprobar actividades colaborativas con 7 puntos mínimo.
- Presentar todas las actividades solicitadas, en tiempo y forma.

Para REGULARIZAR la Asignatura:

- Aprobar parciales o sus recuperatorios.
- Aprobar el taller de programación.
- Aprobar las actividades colaborativas.

ACTUALIDAD

Incremento en el volumen y
en la variedad de datos
almacenados digitalmente...




¿Aspecto positivo?: memoria organizacional y
fuente de conocimiento...

¿Aspecto negativo?: forma de interpretar...

CAMPAÑAS DE PREVENCIÓN

PARA EVITAR DENGUE, ZIKA Y CHIKUNGUNYA

ELIMINÁ LOS RECIPIENTES EN DESUSO QUE ACUMULAN AGUA.



CAMPAÑA DE PREVENCIÓN EN LA DETECCIÓN DEL CÁNCER DE CUELLO UTERINO

CUIDA TU CUERPO NO TIENE REEMPLAZO

La campaña se desarrollará en la Ciudad de San Miguel de Tucumán desde el 27/06 al 01/07 de 2016. Los puestos de atención se concentrarán en la Delegación de OSFATUN y en el Centro Médico 26 de Noviembre. Está destinada a todas las afiliadas de la Obra Social OSFATUN.

¡Participa de los controles, son gratuitos. Comunícate con tu Delegación y solicita el turno!

Delegación Tucumán Centro: (0381) 420-3744

Centro Médico 26 de Noviembre: (0381) 436-4163

CAMPAÑA NACIONAL DE VACUNACIÓN

SARAMPIÓN Y RUBÉOLA

del 1º de octubre al 30 de noviembre

El sarampión y la rubéola son enfermedades virales muy contagiosas que se transmiten por vía aérea.

CONTAGIO

Toser Estornudar

Aunque en la Argentina ya no circulan estos virus, es indispensable que todos los niños estén vacunados.

Cuando algún portador del virus tose o estornuda, las gotitas que contienen el virus se esparcen por el aire y se depositan sobre las superficies cercanas. Cualquier persona puede contagiarse el virus inhalando estas gotitas o tocándolas y luego llevando las manos a su rostro, boca, ojos u oídos.

SARAMPIÓN

Provoca fiebre alta, secreción nasal, conjuntivitis, tos y erupciones en la cara y el cuello. Puede llegar a ser mortal. Los casos más graves se dan en los menores de 5 años.

RUBÉOLA

Genera erupciones, fiebre, dolor de garganta, conjuntivitis, e inflamación en los ganglios. Si una mujer embarazada la contrae, su bebé puede tener graves complicaciones.

Niños y niñas de 13 meses a 4 años

Deben recibir una dosis **obligatoria** de la vacuna.

Es gratuita

No es necesario tener una orden médica.

En todos los hospitales

Y centros de vacunación de todo el país.

Línea gratuita para más información, que atiende de lunes a viernes de 9 a 21 hs. Sábados, domingos y feriados de 9 a 18 hs. **0800-222-1002**

FUENTE: Ministerio de Salud. Presidencia de la Nación. www.argentina.gob.ar/salud

+ 65 AÑOS PUÉRPERAS EMBARAZADAS PERSONAL DE SALUD NIÑOS + 6 MESES - 2 AÑOS FACTORES DE RIESGO

CAMPAÑA 2019

VACUNACIÓN ANTIGRI PAL

DESDE EL 15 DE ABRIL

EN TODOS LOS CENTROS DE ATENCIÓN PRIMARIA DE LA SALUD (CAPS) DEL DISTRITO DE ALTE BROWN DE LUNES A VIERNES DE 8:00 A 18:00 HS.

CAMPAÑA DE INMUNIZACIÓN GRATUITA A:

SIN ORDEN MÉDICA

- > NIÑOS DE 6 MESES A DOS AÑOS DE EDAD,
- > EMBARAZADAS Y PUÉRPERAS,
- > PERSONAS MAYORES DE 65 AÑOS,
- > Y PERSONAL DE SALUD.

CON ORDEN MÉDICA

- > PERSONAS ENTRE LOS 2 AÑOS Y HASTA LOS 64 AÑOS CON FACTORES DE RIESGO.

CAMPAÑA CÁNCER DE PULMÓN

LALCEC

Turnos:
a partir
del 29 de Mayo
En la Sede del LALCEC
de Lunes a Viernes
de 9 a 11

MANTENETE ALEJADO DEL TABACO; PREVENÍ EL CÁNCER

La mejor manera de reducir el riesgo de cáncer de pulmón es no fumar y evitar el humo de las personas que fuman.

Un estudio a tiempo puede salvarte la vida

LALCEC Rojás
DIRECCION APS MUNICIPIO DE ROJAS
CPA de Rojás

Dra. Rosanna Costaguta

¿CÓMO SE DISEÑAN ESTAS CAMPAÑAS





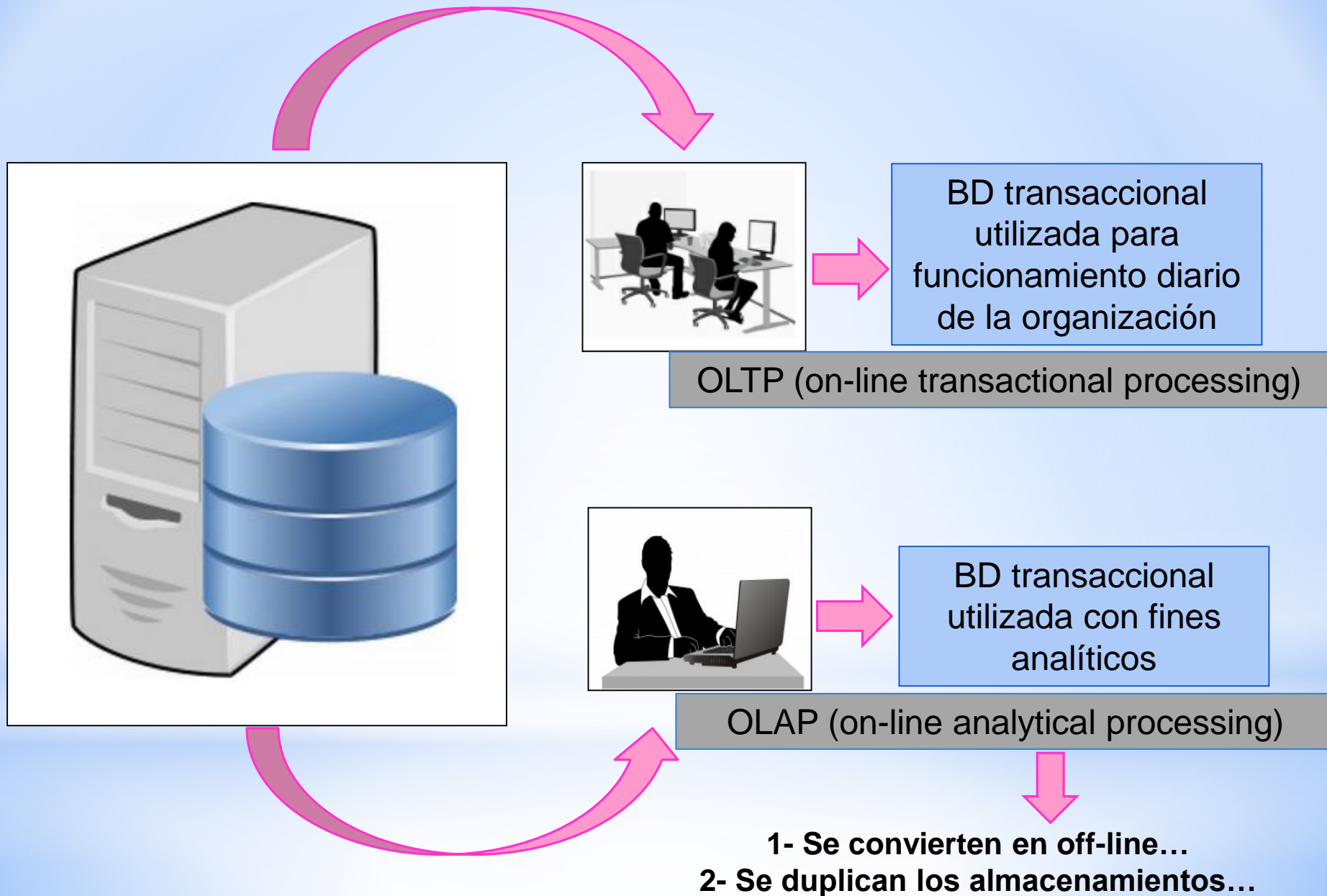
Decisiones

PROCEDER:

- Caro
- Lento
- Subjetivo



**Enorme cantidad de datos y
variabilidad en sus tipos →
Capacidades humanas de
análisis e interpretación no
alcanzan**



SITUACIÓN HIPOTÉTICA 2

Hospitales y UPAs de la ciudad X



Embarazadas atendidas



*Evolución o tendencia
en bebés prematuros...*

OLAP



“10 % de nacimientos
prematurados”

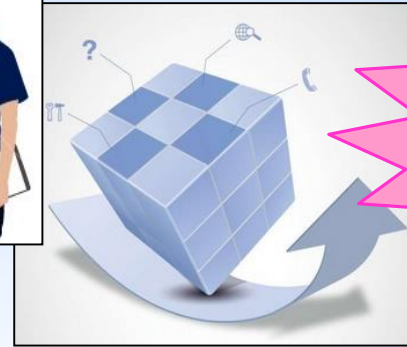
**minería
de
datos**



Patrones o reglas...

Historias clínicas + hábitos + datos socioeconómicos + ...

Dra. Rosanna Costaguta



**minería
de
datos**

Patrones o reglas...

Historias clínicas + hábitos + datos socioeconómicos + ...



**PREDECIR SI UN EMBARAZO TERMINARÁ EN
UN BEBÉ PREMATURO**



**Toma de
decisiones**

- Definición de Políticas de prevención adecuadas
- Definición de Políticas de mitigación adecuadas

A large, multi-pointed pink starburst shape with a thin pink outline, centered in the upper half of the slide. The text 'MINERÍA DE DATOS' is written in white, bold, uppercase letters inside the starburst.

MINERÍA DE DATOS

Proceso de extraer conocimiento útil y comprensible, previamente desconocido, desde grandes volúmenes de datos almacenados en diferentes formatos.

(Witten y Frank, 2000)

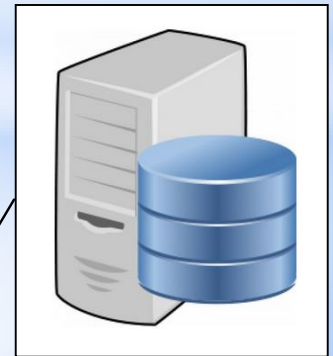
EJEMPLOS DE APLICACIÓN DE MD



CASO 1: Análisis de créditos bancarios

Se aplican técnicas de MD sobre datos históricos de clientes y préstamos para encontrar **reglas** que permitan evaluar si un nuevo crédito será o no devuelto.

*Nro-crédito/plazo-meses/monto\$/ID-Cliente/
ing-mensual/casa-propia/estado-civil/cant-hijos/
vehículo/empleado/estudio-máximo/cuotas-morosas/
devolvió-crédito*





CASO 1: Análisis de créditos bancarios

REGLAS/PATRONES

IF *cuotas-morosas* $< > 0$ AND *monto*\$ > 7.000 AND
casa-propia = NO AND *ing-mensual* < 10.000 AND
estado-civil = divorciado THEN *devolvió-crédito* = NO

IF *cuotas-morosas* = 0 AND *plazo-meses* = 24 AND
ing-mensual > 10.000 AND *estado-civil* = casado
THEN *devolvió-crédito* = SI

¿Qué puede hacer el banco con este conocimiento? Puede clasificar cada nueva solicitud y rechazar aquellas que reúnan las características de clientes que no devuelven el crédito.

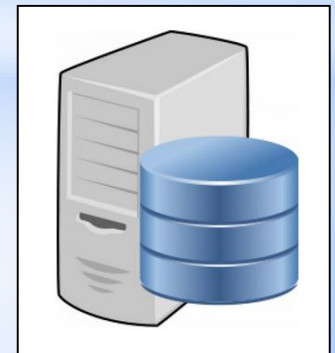
EJEMPLOS DE APLICACIÓN DE MD



CASO 2: Análisis de la canasta o cesta de compras

Se aplican técnicas de MD sobre datos almacenados de clientes, productos y ventas para encontrar **reglas** que describan el comportamiento de compra de los clientes.

*Nro-ticket/ID-Cliente/leche/pañales/lechuga/asado/
coca-cola/sal/carbón/cerveza/ketchup/huevos/
maní/.../desodorante/dentífrico/...*





CASO 2: Análisis de la cesta de compras

REGLAS/PATRONES

IF *leche* = SI AND *pañales* = SI AND *desodorante* = SI
THEN *ketchup* = SI

IF *asado* = SI AND *carbón* = SI THEN *coca-cola* = SI

IF *pañales* = SI THEN *coca-cola* = SI AND *cerveza* = SI

¿Qué puede hacer el supermercado con este conocimiento?

Puede reordenar sus góndolas poniendo los productos juntos, puede lanzar campañas de ofertas para incentivar el consumo de determinados productos, etc.

EJEMPLOS DE APLICACIÓN DE MD



CASO 3: Análisis de las ventas de un producto

Se aplican técnicas de MD sobre datos almacenados de productos, ventas y stock para encontrar **reglas** que permitan identificar la cantidad ideal de productos para los distintos meses del año.

ID-Producto/denominación/sept2017/oct2017/nov2017/dic2017/ene2018/feb2018/mar2018/abril2018/mayo2018/jun2018/jul2018/ago2018





CASO 3: Análisis de las ventas de un producto

REGLAS/PATRONES

IF *denominación* = TV led 55` AND *oct2017* > 150
AND... AND *ago2018* > 280
THEN *sept2018* > 150 AND *sept2018* < 200

IF *denominación* = cafetera express AND *nov2017* < 5
AND... AND *ago2018* < 5
THEN *sept2018* < 5

¿Qué puede hacer la empresa con este conocimiento?
Puede generar sus órdenes de reposición de productos sin
generar “sobre-stock” ni sufrir venta insatisfecha.

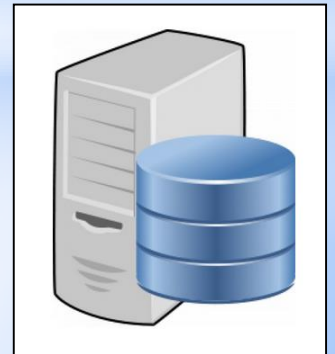
EJEMPLOS DE APLICACIÓN DE MD



CASO 4: Análisis de grupos humanos

Se aplican técnicas de MD sobre datos históricos de empleados para categorizarlos en diferentes **grupos** que compartan determinadas particularidades.

ID-Empleado/Ape/Nom/estado-civil/título-máx/estado-civil/cant-hijos/sexo/antigüedad/sueldo-promedio/cant-inasistencias/...



Dra. Rosanna Costaguta



CASO 4: Análisis de grupos humanos

REGLAS/PATRONES

Grupo 1: *estado-civil* = soltero AND *cant-hijos* = 0
AND *sexo* = masculino AND *antigüedad* = 3
AND *cant-inasistencias* = 35

Grupo 2: *estado-civil* = casado AND *cant-hijos* = 2
AND *sexo* = masculino AND *sueldo-promedio* = 10.000
AND *título-máx* = secundario

Grupo 3: *estado-civil* = casado AND *cant-hijos* = 3
AND *sexo* = femenino AND *cant-inasistencias* = 5

¿Qué puede hacer la empresa con este conocimiento?
Puede establecer políticas de contratación, políticas de capacitación, etc.

¿SOBRE QUÉ TIPOS DE DATOS SE PUEDE APLICAR MD?



DATOS ESTRUCTURADOS PROVENIENTES DE BD RELACIONALES



BD relacional = colección de tablas

Tabla = conjunto de tuplas (filas) y atributos (columnas)

Tupla = Instancia

Los datos de una BD relacional responden a una estructura y cumplen restricciones de integridad...

TIPOS DE DATOS ÚTILES PARA MD



- *Numéricos* (valores enteros o reales).

Ejemplos: ing-mensual = 10.165,80

cant-hijos = 2

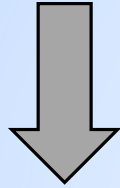
- *Catégoricos o nominales* (valores pertenecientes a categorías predefinidas).

Ejemplos: sexo = M o sexo = F

casa-propia = sí o casa-propia = no

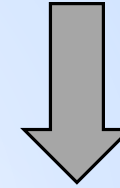
estado-civil = 1 .. estado-civil = 5

¿QUÉ TIPOS DE MODELOS SE USAN EN MD?



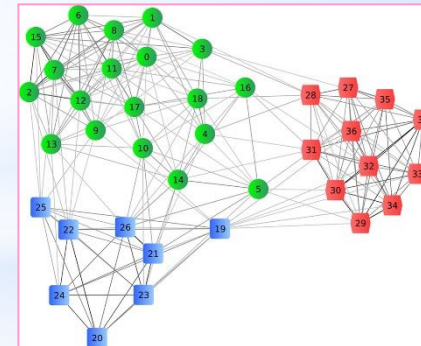
Modelos predictivos:

Tratan de establecer valores futuros o desconocidos para una variable en base a variables conocidas.



Modelos descriptivos:

Identifican patrones que resumen o explican los datos analizados.

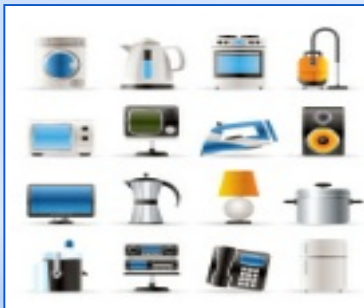




CASO 1: Análisis de créditos bancarios
MODELO PREDICTIVO



CASO 2: Análisis de la cesta de compras
MODELO DESCRIPTIVO



CASO 3: Análisis de las ventas de un producto
MODELO PREDICTIVO



CASO 4: Análisis de grupos humanos
MODELO DESCRIPTIVO

ACTIVIDAD

- 1) Reúnase con dos compañeros.
- 2) Lean el artículo que les fue asignado.
- 3) Respondan:
 - Contexto en el que se aplicó MD
 - Objetivo perseguido
 - Datos utilizados
 - Resultados alcanzados
 - Tipo de modelo aplicado
- 4) Presenten las respuestas a toda el aula.