Grandes de Bases de Datos

Minería de datos

Minería de Datos

- Google procesa 20 PB por día
  - … en el 2008
- INTERNET ARCHIVE WayBackMachine posee 3 PB + 100 TB/mensuales
  - … en el 2009
- facebook tiene 2.5 PB de datos de usuarios + 15 TB/diarios
  - … en el 2009

Minería de Datos

- ebay concentra 6.5 PB de datos de usuarios + 50 TB/diarios
  - … en el 2009
- CERN GCH - CERN genera 15 PB anuales
  - cuando funciona…

Minería de Datos

- Retos principales:
  - ¿Qué hacer con tantos datos?
  - ¿Cuándo son útiles?
  - ¿Cómo los almacenamos?
  - ¿Cómo se encuentran disponibles?

¿Qué es Minería de Datos?

¿Qué es la minería de datos?

"***Proceso no trivial de identificación de patrones valiosos, novedosos, potencialmente útiles y entendibles a partir de grandes volúmenes de datos***"

U. Fayyad, et al. 's definición de KDD en KDD96

Por partes!!

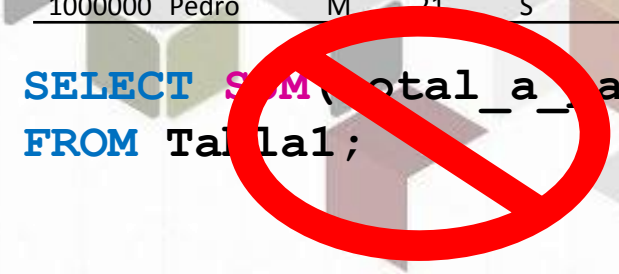
- ¿Grandes volúmenes de datos?
  - El analizar volúmenes pequeños no requiere de la minería de datos.

Id Cliente	Nombre	Genero	Edad	Examen Prostata	Total a pagar
1	Ana	F	23	N	100,000
2	Luis	M	22	S	120,000
3	Hugo	M	21	S	105,000
4	Mary	F	20	N	100,000
5	Sonia	F	22	N	120,000
6	Paco	M	23	S	102,000
7	Toño	M	22	S	99,000
8	Susan	F	21	N	115,000
9	Karla	F	20	N	108,000
10	Pedro	M	21	S	109,000

Por partes!!

- ¿Proceso no trivial?
  - La minería de datos no son simples consultas de SQL

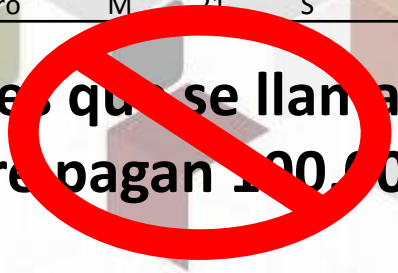
Id Cliente	Nombre	Genero	Edad	Examen Prostata	Total a pagar
1	Ana	F	23	N	100,000
2	Luis	M	22	S	120,000
3	Hugo	M	21	S	105,000
4	Mary	F	20	N	100,000
…					
…					
…					
999998	Susan	F	21	N	115,000
999999	Karla	F	20	N	108,000
1000000	Pedro	M	21	S	109,000



Por partes!!

- ¿Patrones valiosos?
  - Podemos encontrar patrones incorrectos y estos no tienen ningún valor

Id Cliente	Nombre	Genero	Edad	Examen Prostata	Total a pagar
1	Ana	F	22	N	100,000
2	Luis	M	22	S	120,000
3	Hugo	M	21	S	105,000
4	Mary	F	20	N	100,000
…					
500000	Ana	F	23	N	100,000
…					
999998	Susan	F	21	N	115,000
999999	Ana	F	19	N	100,000
1000000	Pedro	M	21	S	109,000

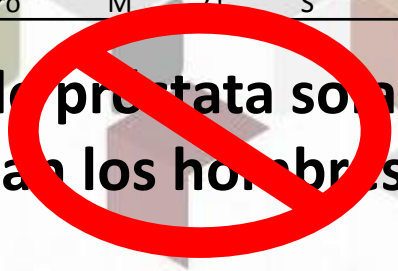


“Las mujeres que se llaman ‘Ana’ siempre pagan 100,000”

Por partes!!

- ¿ Patrones novedosos?
  - El desgastarnos en conocimiento conocido es un desperdicio

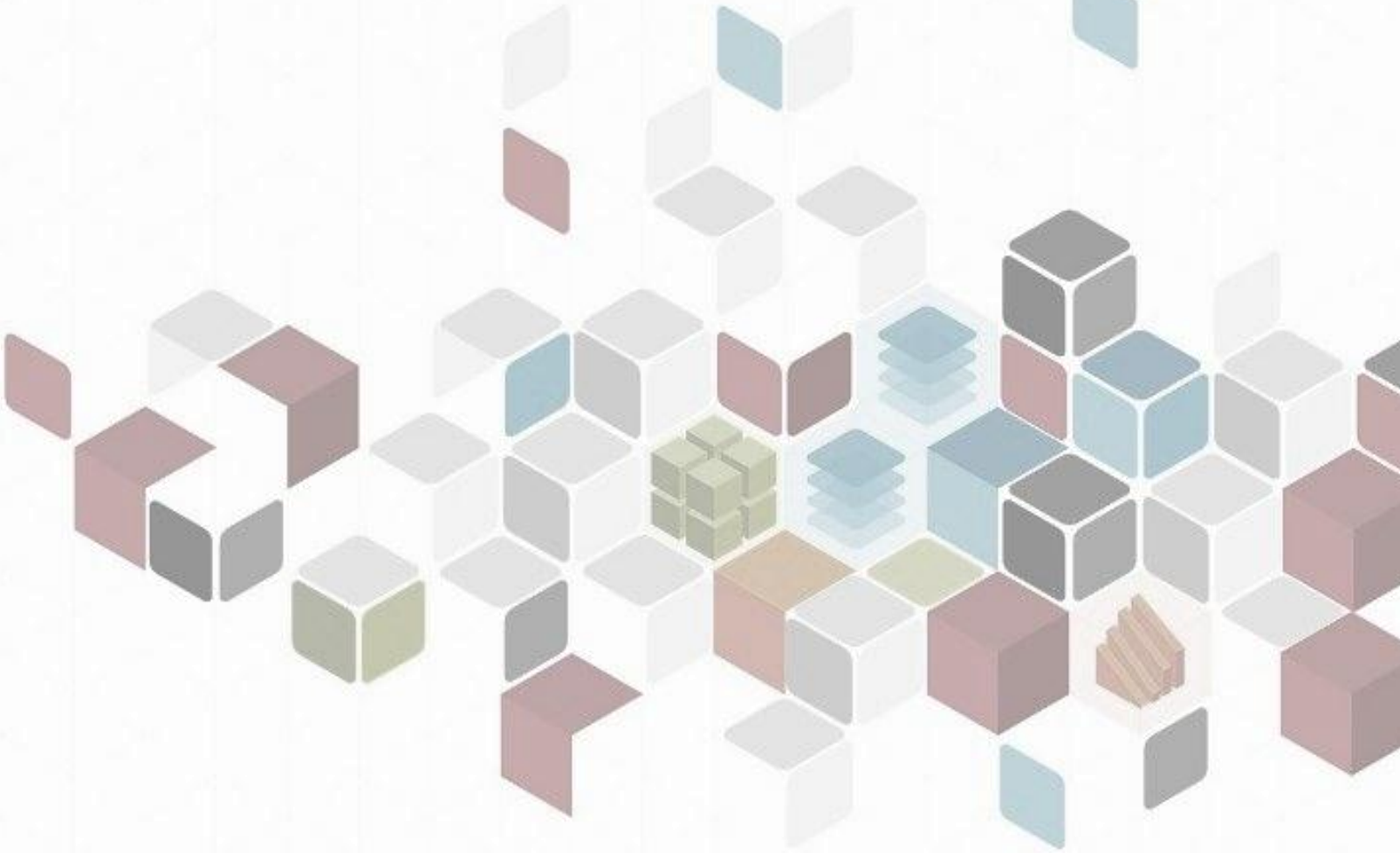
Id Cliente	Nombre	Genero	Edad	Examen Prostata	Total a pagar
1	Ana	F	23	N	100,000
2	Luis	M	22	S	120,000
3	Hugo	M	21	S	105,000
4	Mary	F	20	N	100,000
…					
…					
…					
999998	Susan	F	21	N	115,000
999999	Karla	F	20	N	108,000
1000000	Pedro	M	21	S	109,000



“El examen de prostata solamente lo realizan los hombres”

Por partes!!

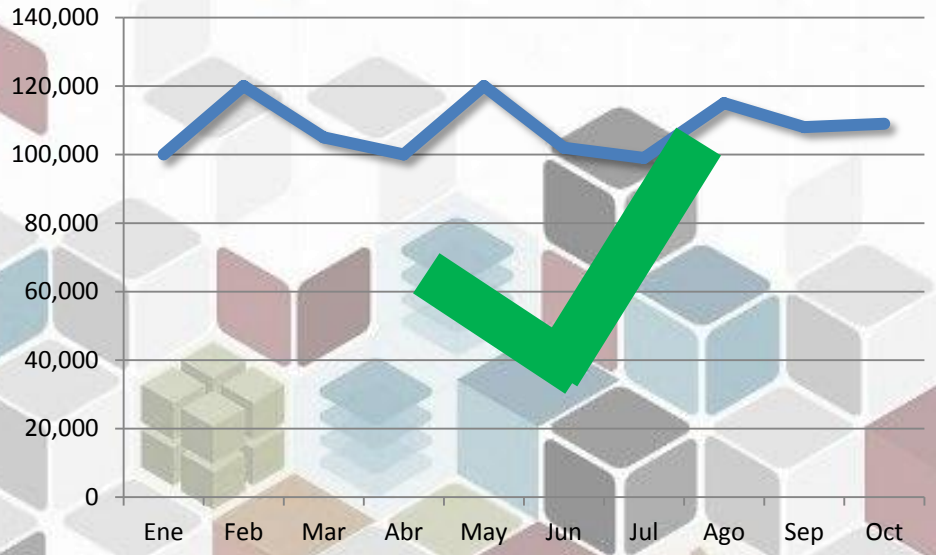
- ¿Potencialmente útiles?
  - Los patrones encontrados se utilizarán para la toma de decisiones



Por partes!!

- ¿Potencialmente entendibles?
  - Los patrones serán presentados y serán visualizados por tomadores de decisiones

Id Cliente	Nombre	Genero	Edad	Examen Prostata	Total a pagar
1	Ana	F	23	N	100,000
2	Luis		22		120,000
3	Hugo		21	S	105,000
4	Mary	F	20	N	100,000
…					
…					
…					
999998	Susan	F	21	N	115,000
999999	Karla	F	20	N	108,000
1000000	Pedro	M	21	S	109,000



Objetivo

- General
  - Permitir a una empresa, el mejorar sus _____ a través de un mejor entendimiento de su _____.
  - Potencial para un mayor numero de ventajas competitivas.

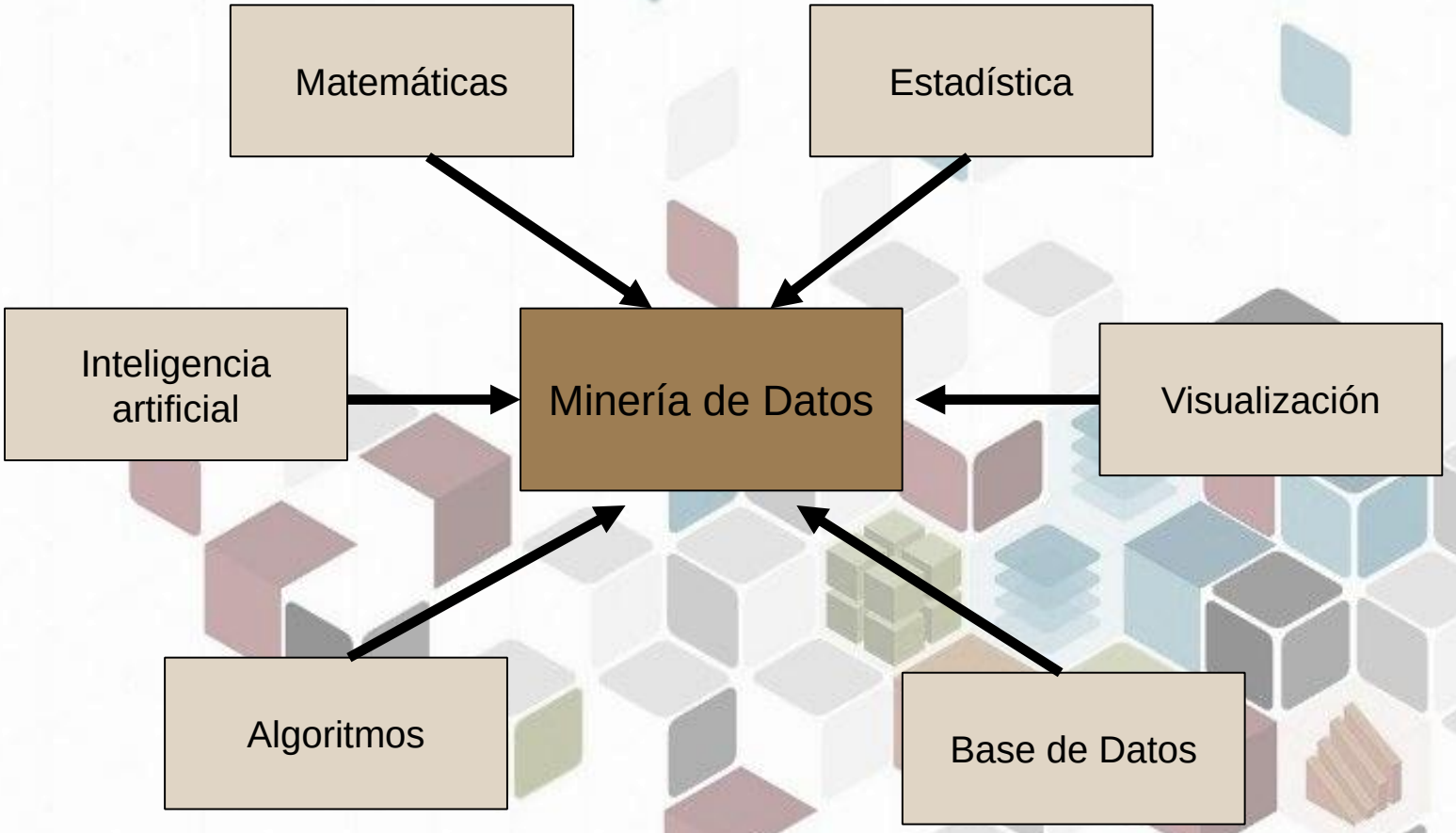
¿Sinónimos?

- Descubrimiento de conocimiento en bases de datos (KDD)
- Análisis exploratorio de datos (*exploratory data analysis*)
- Estadística aplicada (*applied statistics*)
- Aprendizaje de máquina (*machine learning*).
- Extracción de conocimiento (*knowledge extraction*).
- Análisis de datos/patrones (*data/pattern analysis*).
- Inteligencia de negocios (*business intelligence*).

Fundamentos de la MD

- Utilizar datos “crudos” para inferir relaciones importantes de negocio.
- A pesar del valor de la minería de datos, existe gran confusión sobre lo que es realmente,… no es magia
- Es una colección de técnicas para analizar grandes cantidades de datos.
- No existe un única aproximación de minería de datos, sino un conjunto de técnicas que pueden ser utilizadas tanto individual como colectivamente.

Principales disciplinas involucradas



En general

- Tecnologías de Sistemas de Bases de Datos y ***Datawarehouse***.
- Estadística.
- Inteligencia artificial.
- Cómputo de alto rendimiento.
- Reconocimiento de patrones
- Redes Neuronales.
- Visualización de datos.
- Recuperación de Información (***Information retrieva***l)
- Procesamiento de imágenes y señales.
- Análisis de datos espacio-temporal.

Minería de datos vs. Estadística

Minería de Datos

- Cantidades de datos:
  - +1,000,000,000 filas
  - 3,000 columnas
- Datos que suceden y se captan.
- ¿Muestrear?
  - No tiene caso si tenemos tanto poder de cómputo y con mucha capacidad.
- Precio:
  - Elevado

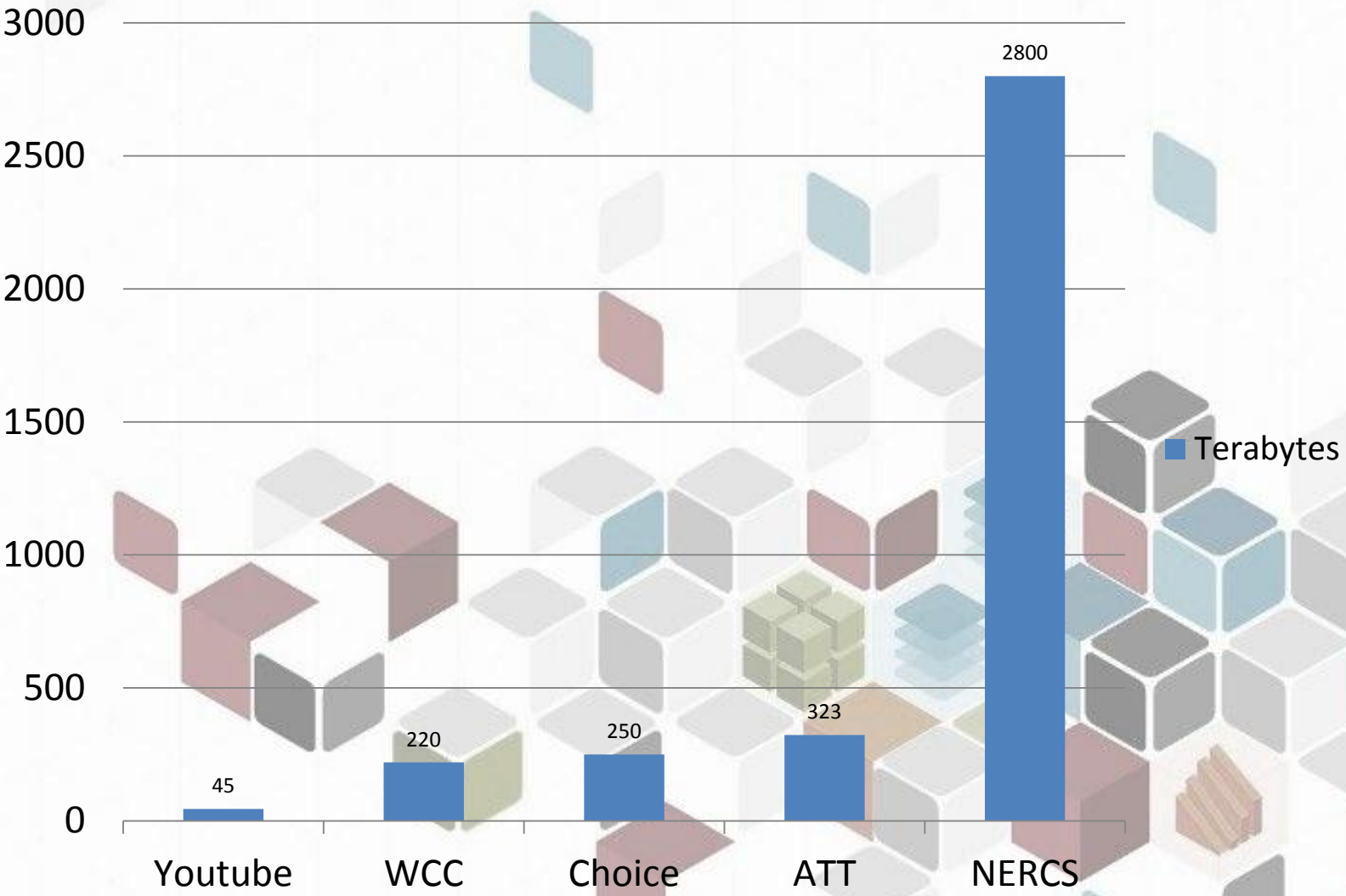
Estadística

- Cantidad de datos:
  - 10,000 filas
  - 20 columnas
- Datos obtenidos sistemáticamente.
- Muestrear
  - Se obtienen hasta estimaciones de errores!!
- Precio:
  - Bajo

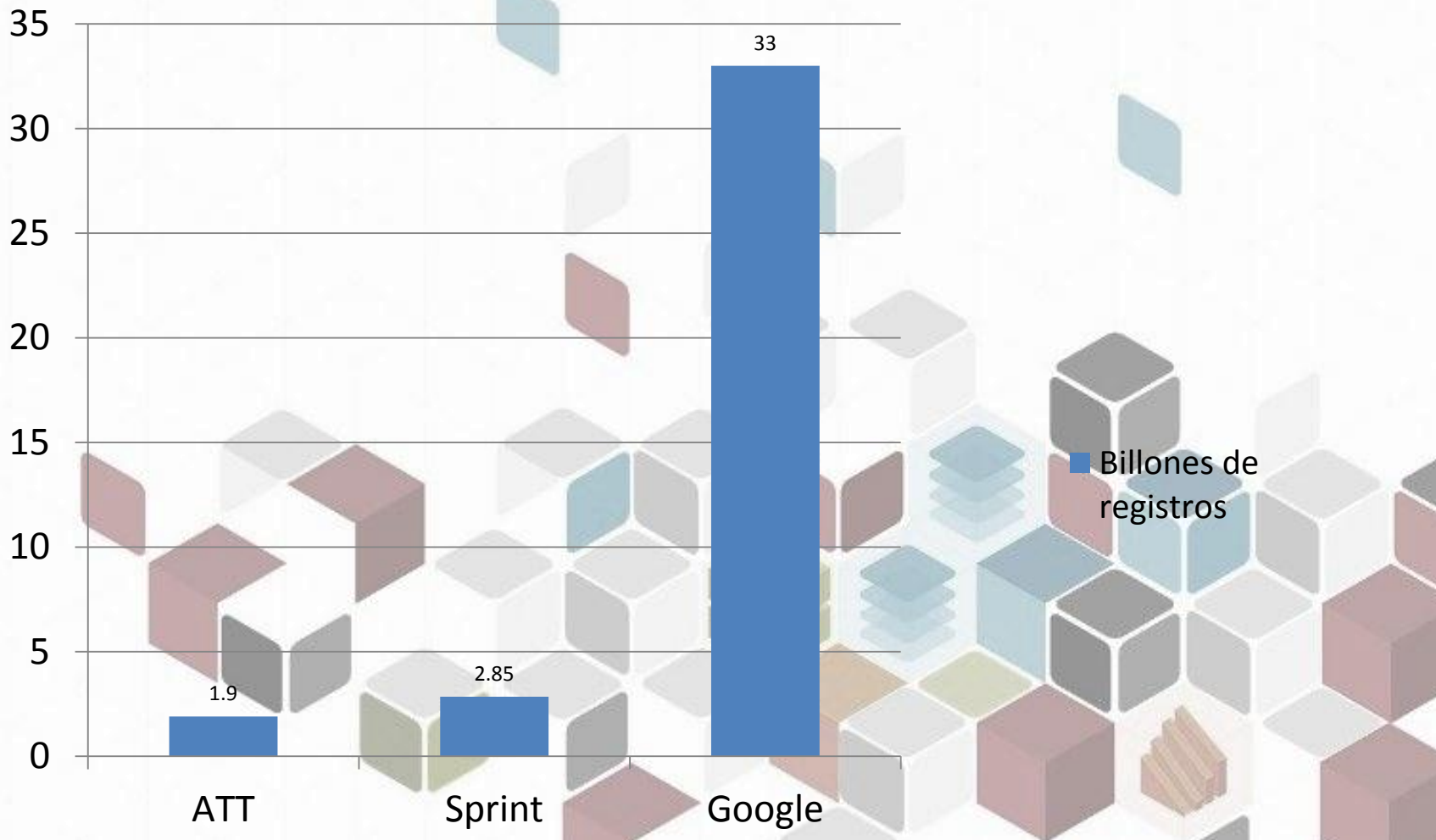
Minería de Datos – ¿Por que?

1. Se producen datos
2. Se almacenan datos
3. El poder de computo es costeable
4. Las presiones competitivas aumentan
5. Sistemas de Minado de Datos están disponibles

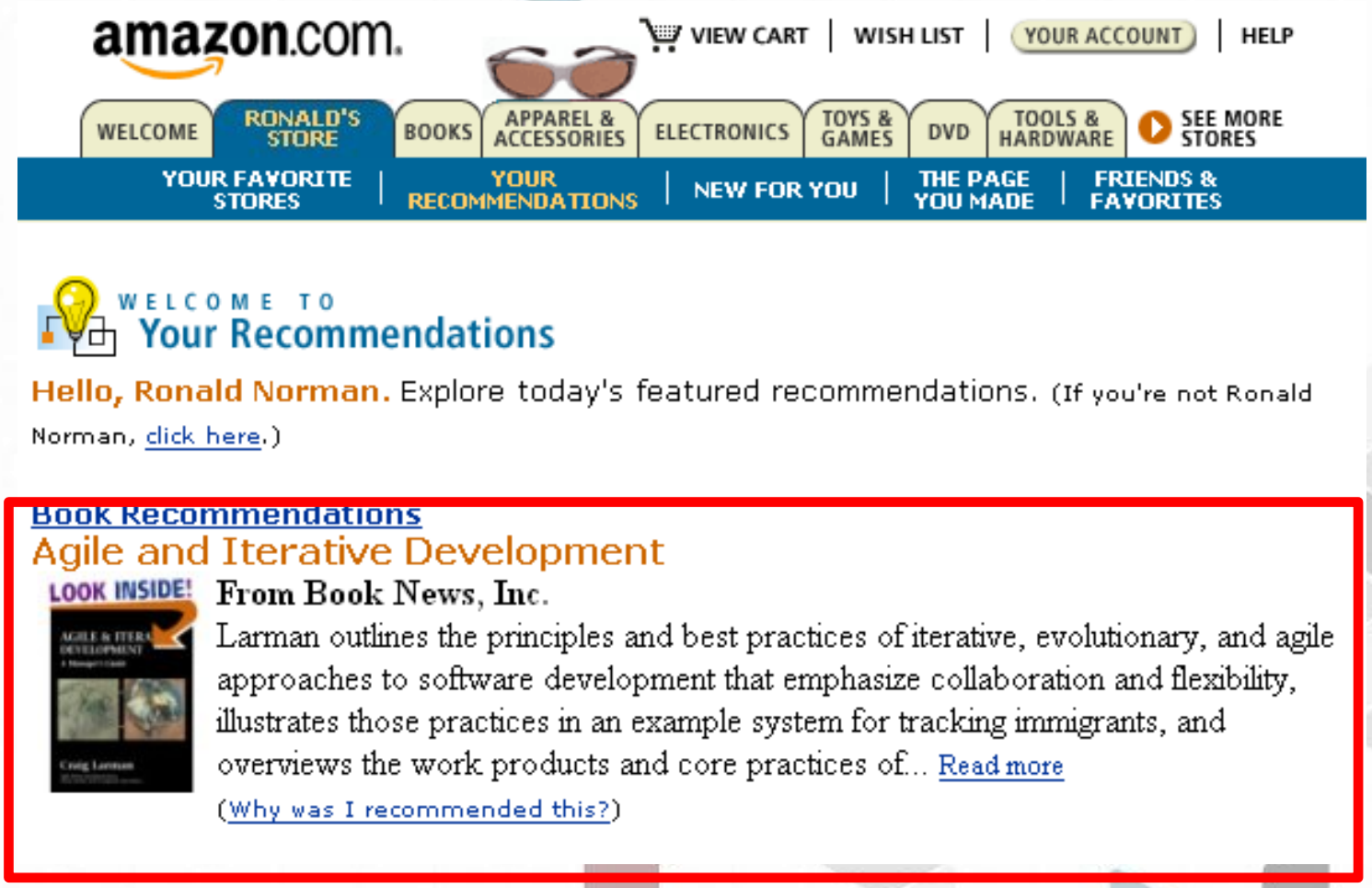
Crecimiento de GBD



Crecimiento de GBD



Customer Relationship Management (CRM)



Customer Relationship Management (CRM)

Una organización debe ser capaz de:

1. **Notar** – ¿Cómo hacen las cosas sus clientes?
2. **Recordar** – ¿Qué hace cada cliente?
3. **Aprender** – Acciones anteriores implican comportamiento
4. **Actuar** – sobre lo que se aprendió, para mejorar la relación con el cliente

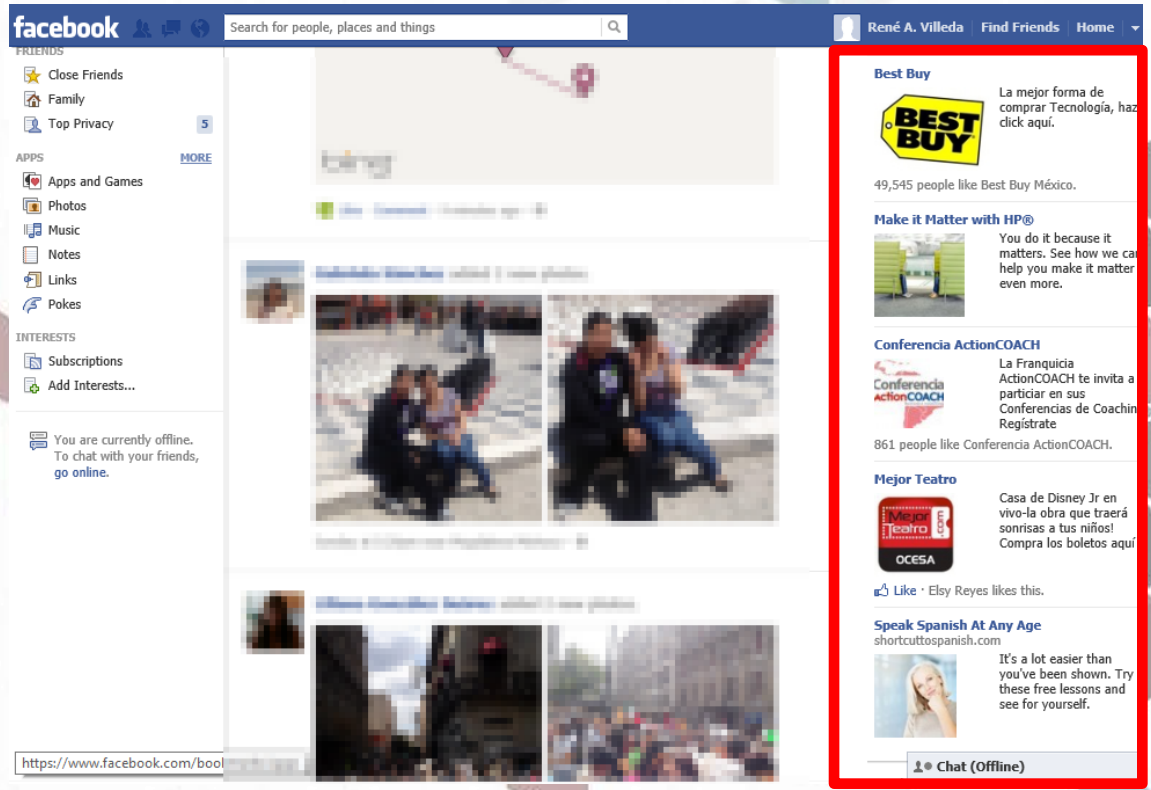
Basados en datos de las transacciones

Basados en datos de las transacciones

	Date	Time	Rate	Minutes	Origination+		Phone number	Destination		Usage type	Call type	Airtime charges
1	05/10	12:12P	P	2	LA Mesa	CA	[redacted]	Voice Mail	CL	AR		Included
2	05/10	01:54P	P	1	LA Mesa	CA		San Diego	CA	MN		.00
3	05/10	01:55P	P	1	LA Mesa	CA		San Diego	CA	MN		.00
4	05/10	02:26P	P	8	Calexico	CA		Incoming	CL	MN		.00
5	05/10	02:59P	P	2	Calexico	CA		Mobile	CL	MN		.00
6	05/10	03:19P	P	1	Calexico	CA	[redacted]	Mobile	CL	MN		.00
7	05/10	04:07P	P	30	LA Mesa	CA		Incoming	CL	A		Included
8	05/11	11:06A	P	3	LA Mesa	CA		Incoming	CL	MN		.00
9	05/11	11:15A	P	1	San Diego	CA		LA Mesa	CA	A		Included
10	05/11	02:26P	P	1	Encinitas	CA		Voice Mail	CL	AR		Included
11	05/11	02:27P	P	2	Encinitas	CA	[redacted]	Chulavista	CA	A		Included
12	05/11	02:47P	P	3	San Diego	CA		Incoming	CL	MN		.00
13	05/11	08:31P	P	4	LA Mesa	CA		Rnchpnsqts	CA	A		Included
14	05/12	11:17A	P	8	LA Mesa	CA		Incoming	CL	MN		.00
15	05/12	11:33A	P	2	LA Mesa	CA		Mobile	CL	MN		.00

Identificar y recordar las relaciones es lo importante

- ***Comscore*** identifica que un usuario promedio pasa 408 minutos al mes en ***Facebook***



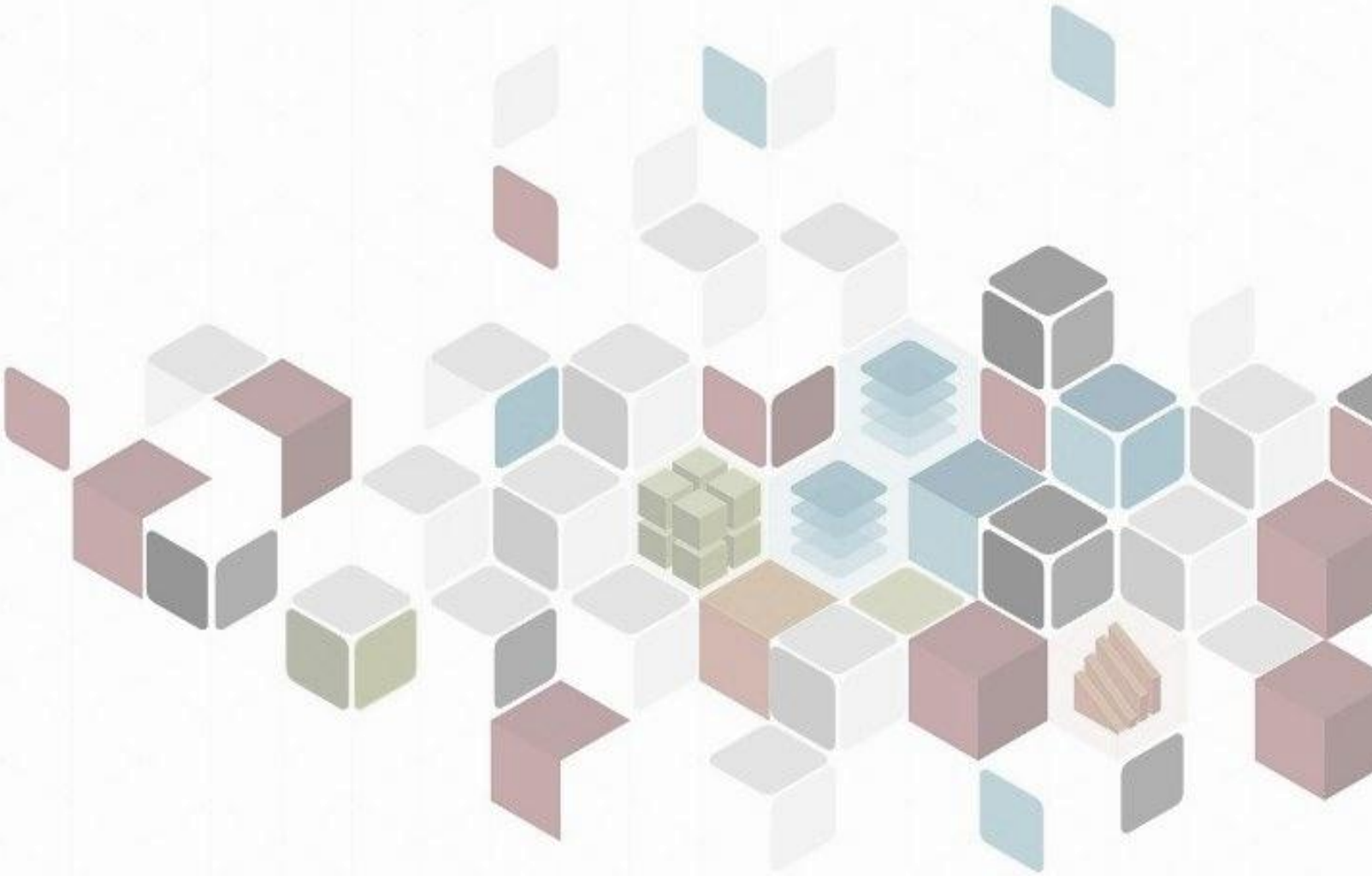
¿Qué motiva la Minería de datos?

- **Ricos en datos pero pobres en información**
  - “Nos estamos ahogando en datos pero estamos sedientos de conocimiento”
- Necesidad de transformar datos en información y conocimiento útil.
  - Análisis de mercados.
  - Detección de fraudes.
  - Retener clientes.
  - Control de producción.
  - Investigación científica.
  - Etc.

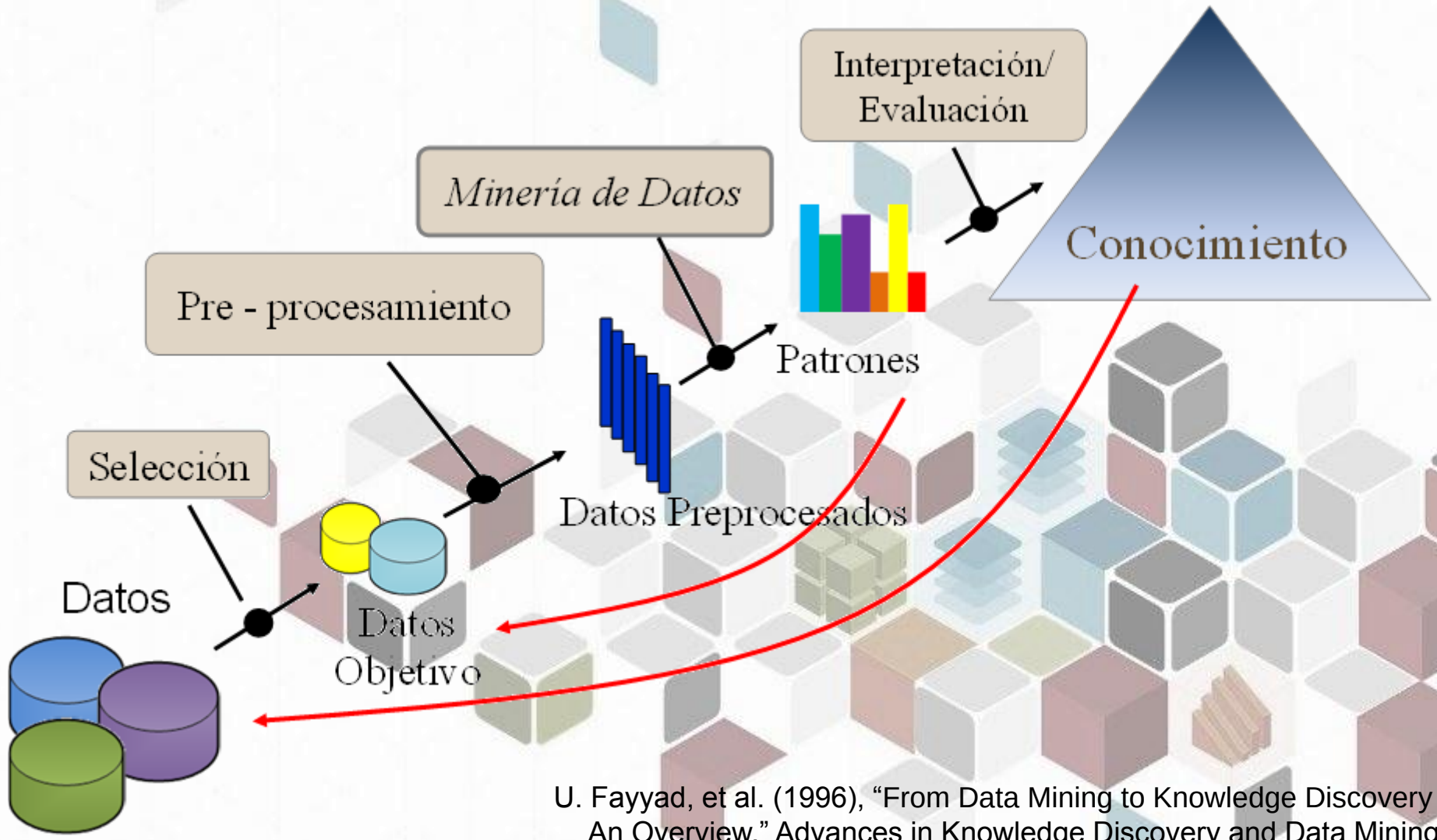
Minería de datos

- Abundancia de datos + necesidad de herramientas para el análisis de datos.
- Las habilidades humanas no pueden.
- Las decisiones importantes no están basadas en los datos almacenados en los repositorios.
- Escasez de herramientas para extraer el conocimiento contenido en las Bases de Datos.
- Alternativa: Sistemas expertos.
  - Consumen mucho tiempo.
  - Costosos.

Proceso de descubrimiento de conocimiento

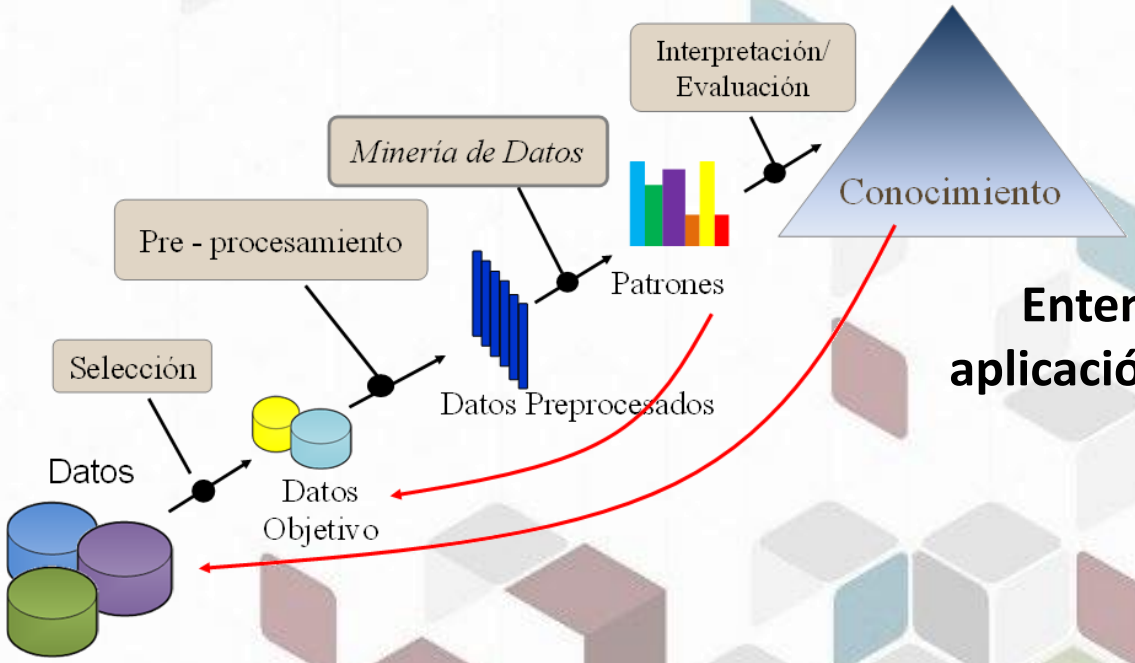


Proceso de descubrimiento de conocimiento



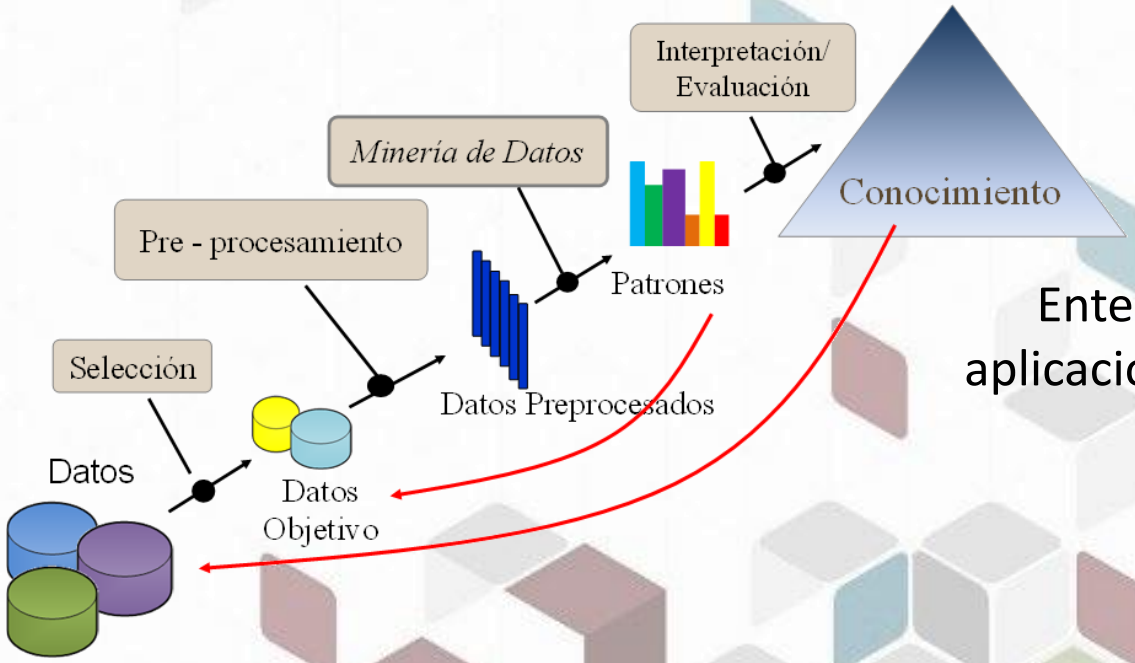
U. Fayyad, et al. (1996), “From Data Mining to Knowledge Discovery : An Overview,” Advances in Knowledge Discovery and Data Mining, U. Fayyad et al. (Eds.), AAAI/MIT Press

Proceso de descubrimiento de conocimiento



Entendimiento del dominio de la aplicación (saber que conocimiento se ha descubierto antes)

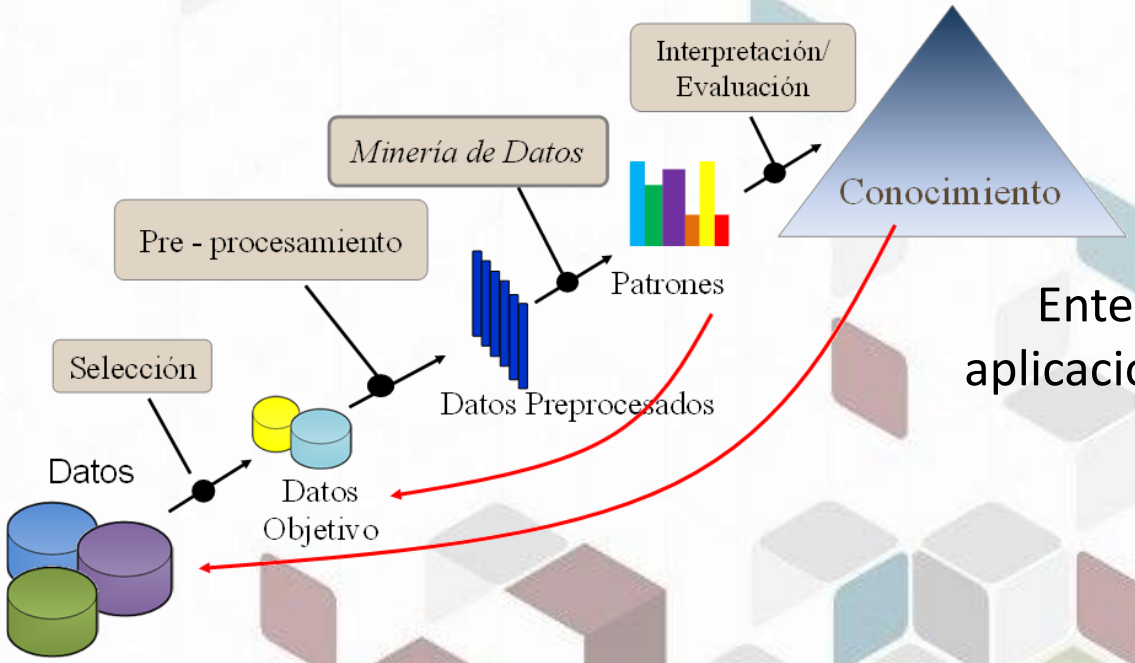
Proceso de descubrimiento de conocimiento



Entendimiento del dominio de la aplicación (saber que conocimiento se ha descubierto antes)

Integración de los datos

Proceso de descubrimiento de conocimiento

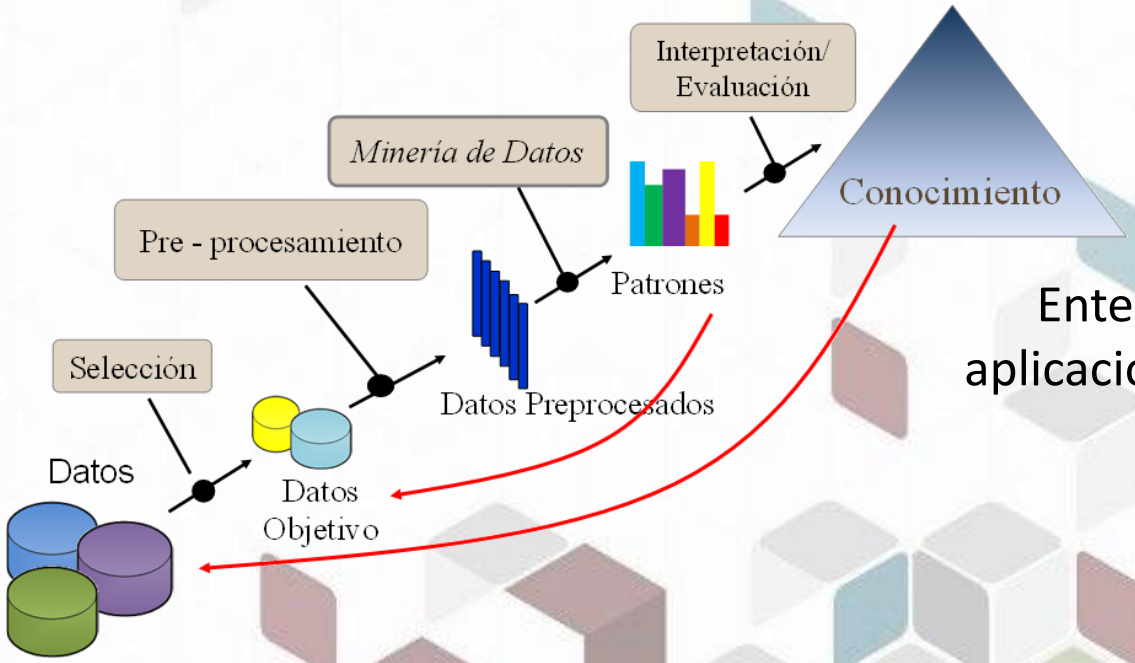


Entendimiento del dominio de la aplicación (saber que conocimiento se ha descubierto antes)

Integración de los datos

Selección de los datos

Proceso de descubrimiento de conocimiento



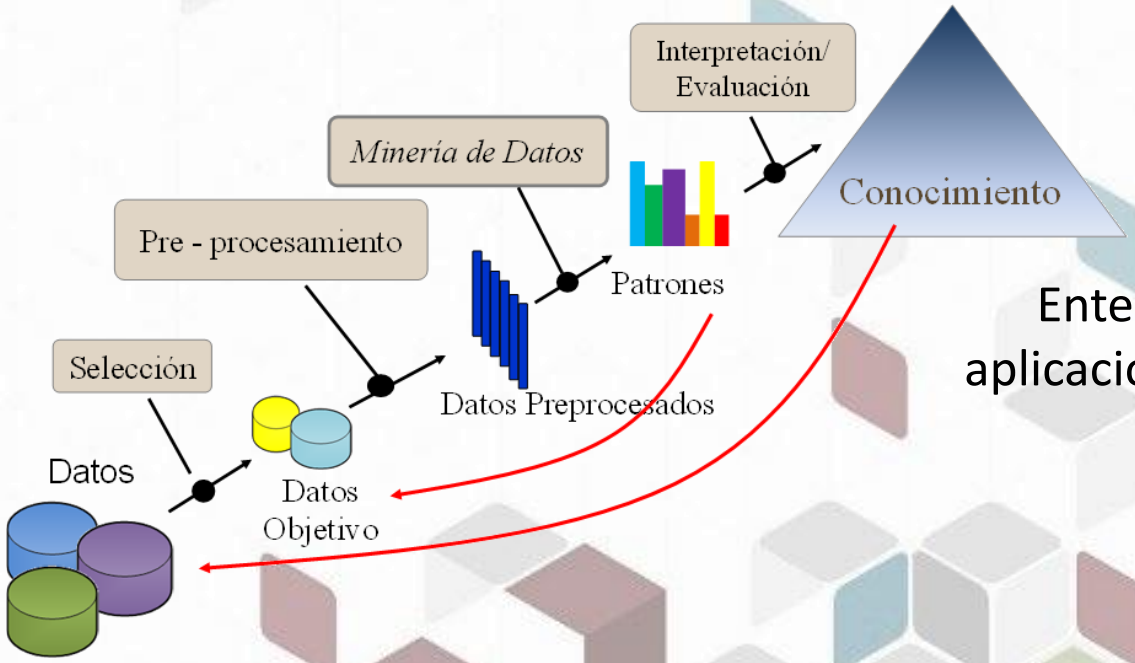
Entendimiento del dominio de la aplicación (saber que conocimiento se ha descubierto antes)

Integración de los datos

Selección de los datos

Limpieza de datos y pre – procesamiento

Proceso de descubrimiento de conocimiento



Entendimiento del dominio de la aplicación (saber que conocimiento se ha descubierto antes)

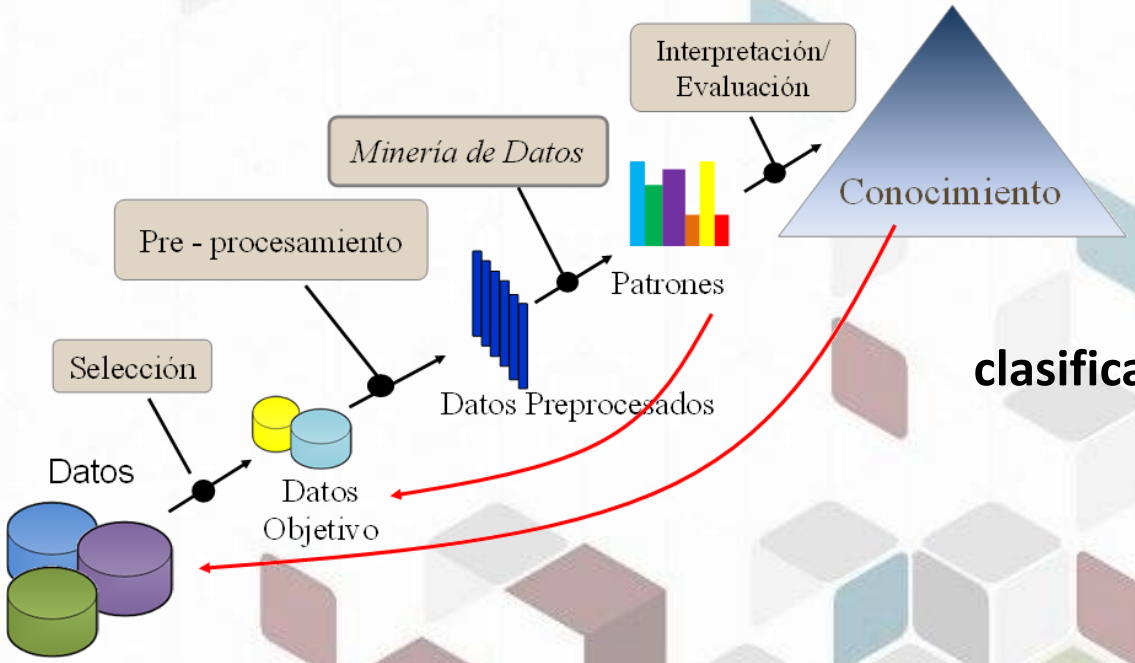
Integración de los datos

Selección de los datos

Limpieza de datos y pre – procesamiento

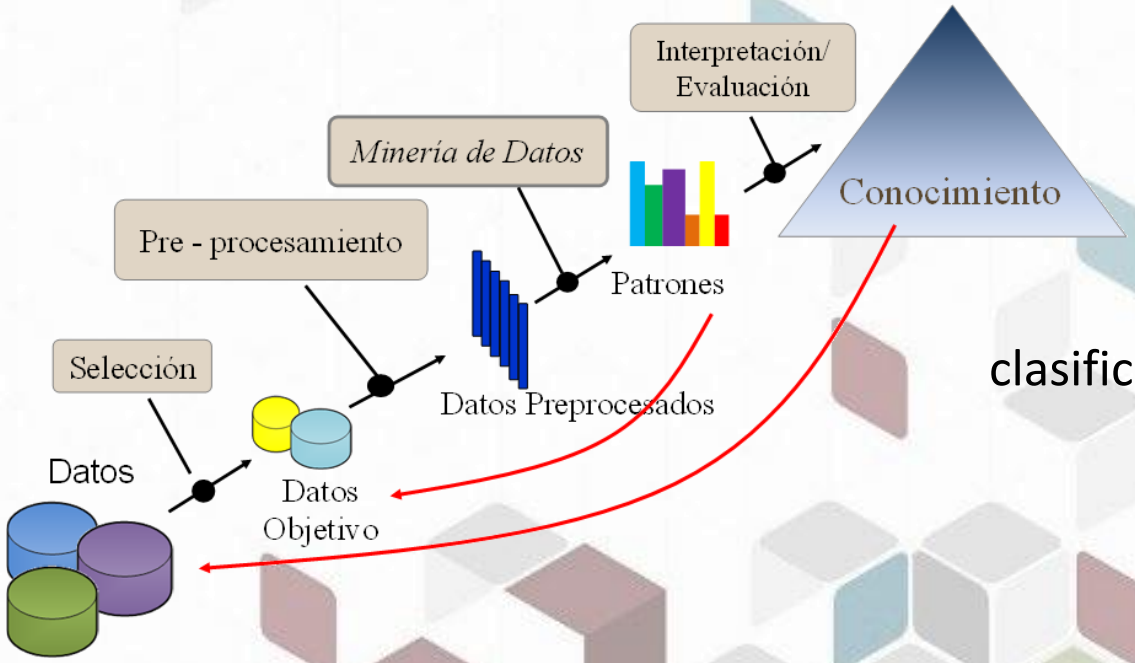
Reducción y Transformación de los datos

Proceso de descubrimiento de conocimiento



Minería de datos (resumen, clasificación, regresión, asociación, agrupación)

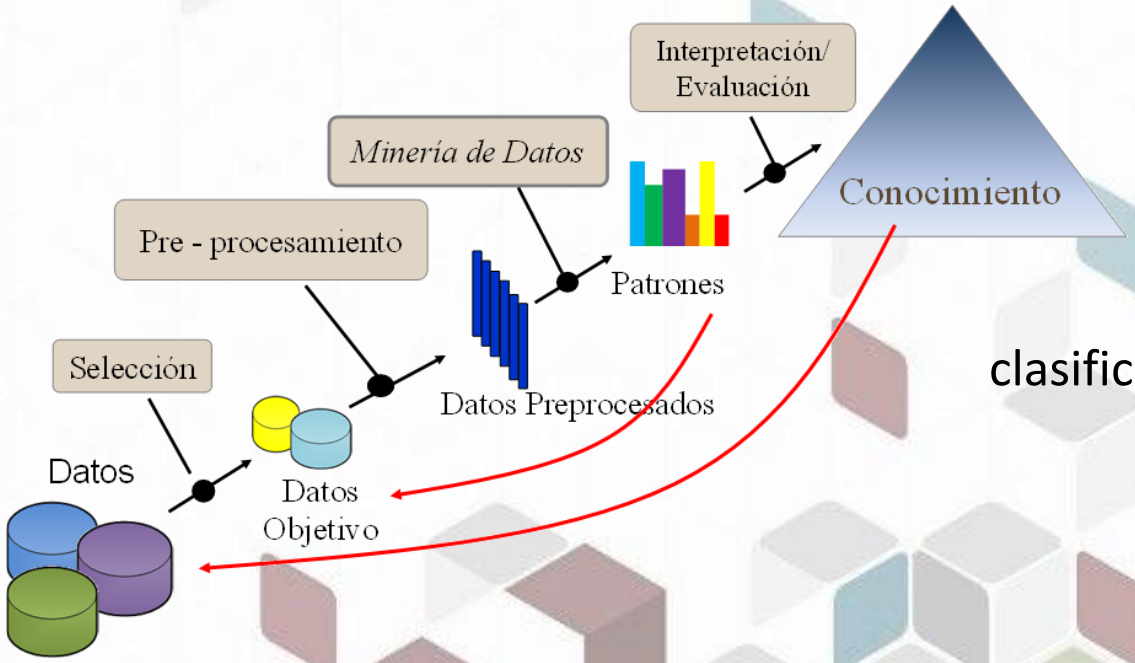
Proceso de descubrimiento de conocimiento



Minería de datos (resumen, clasificación, regresión, asociación, agrupación)

Evaluación de los patrones

Proceso de descubrimiento de conocimiento



Minería de datos (resumen, clasificación, regresión, asociación, agrupación)

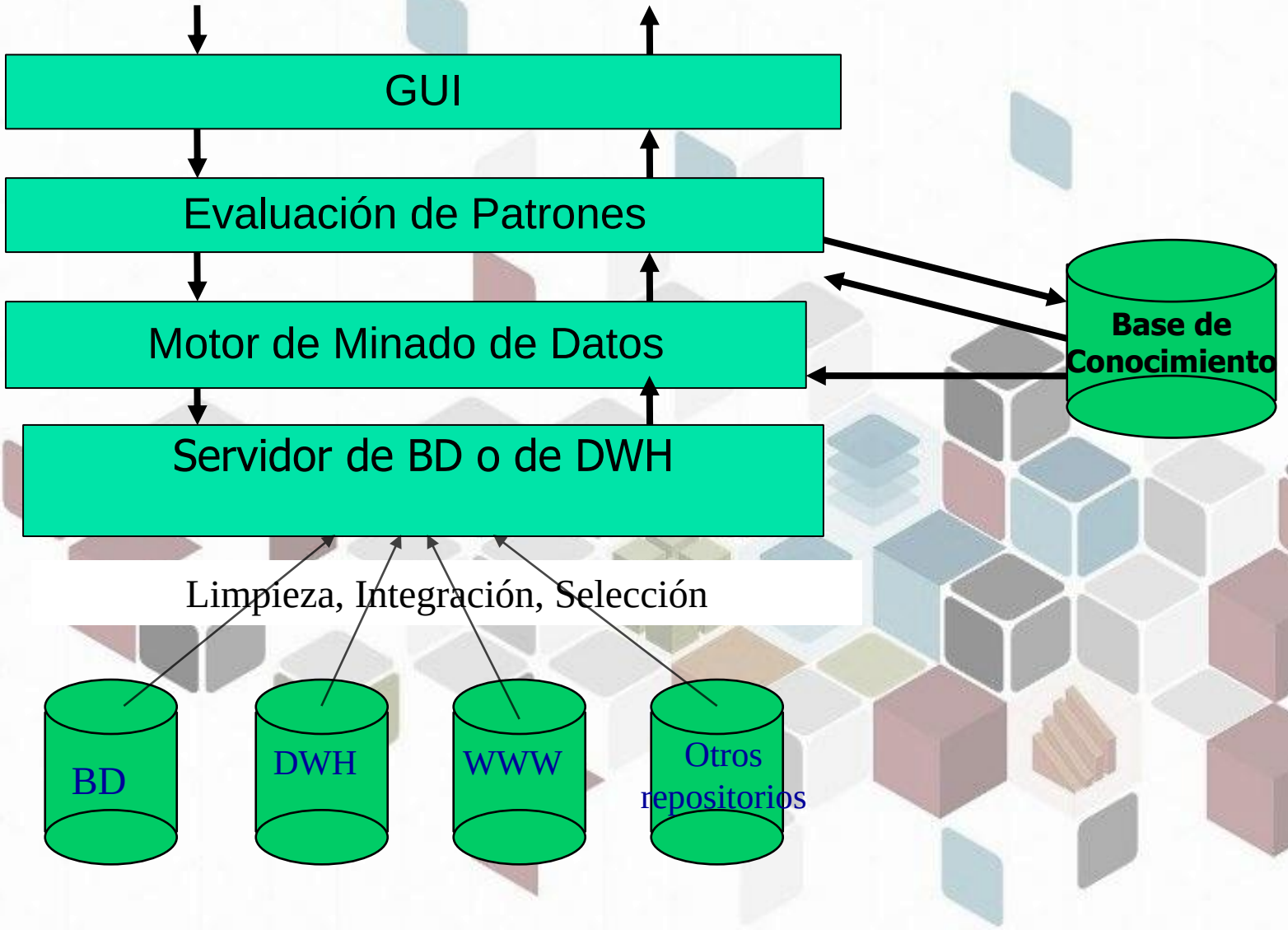
Evaluación de los patrones

Presentar el conocimiento

Arquitectura general de un sistema de minería de datos

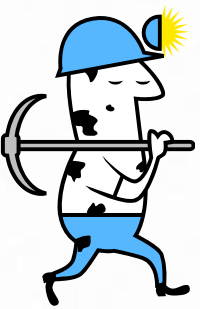
- Repositorios.
- Servidor de Base de Datos o Data Warehouse.
- Base de conocimiento.
- *Data mining engine*.
- Modulo de evaluación de patrones.
- Interfaz con el usuario.

Arquitectura general de un sistema de minería de datos

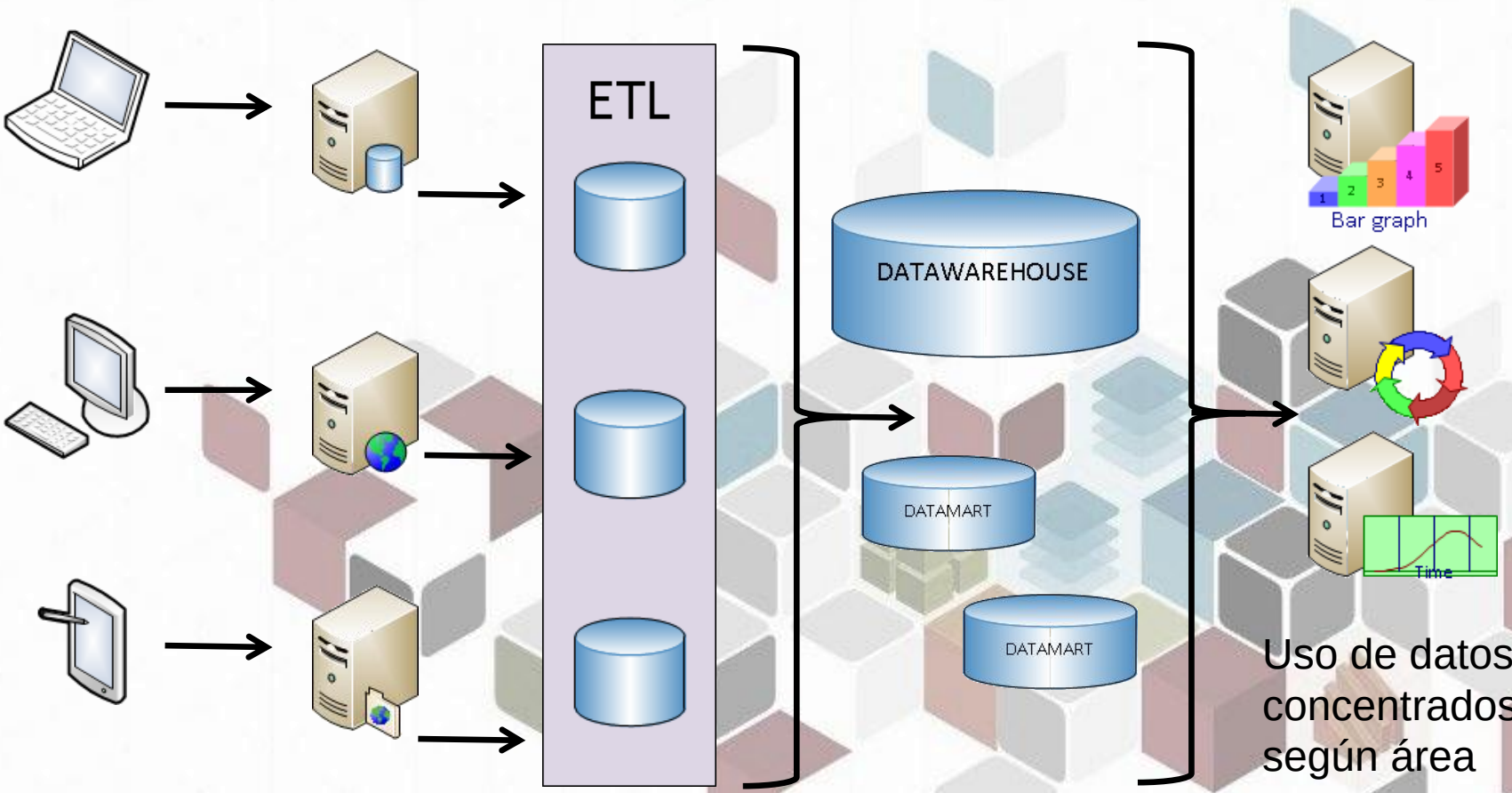


Minería de datos: ¿En donde?

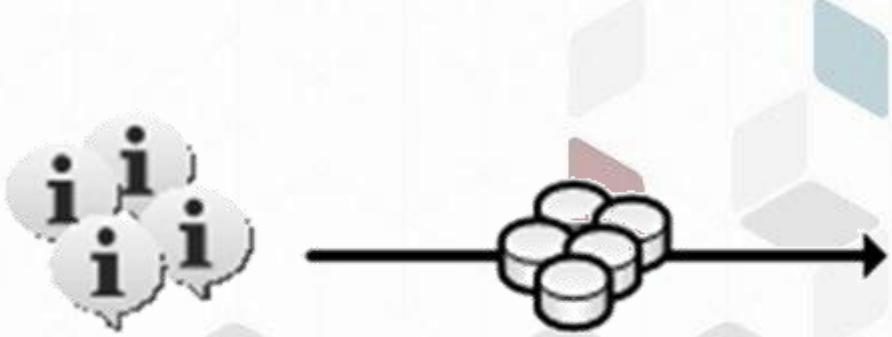
- Archivos planos
- Bases de datos relacionales.
- *Data Warehouse.*
- Bases de datos transaccionales.
- Sistemas avanzados de bases de datos y aplicaciones avanzadas de bases de datos:
  - Bases de datos Objeto-Relacionales.
  - Bases de datos temporales y de series de tiempo.
  - BD espaciales y espacio-temporales.
  - Bases de Datos de texto y bases de datos multimedia.
  - Bases de Datos Heterogéneas y Legadas.
  - *Data Streams*.
  - WWW.



Almacenes de datos



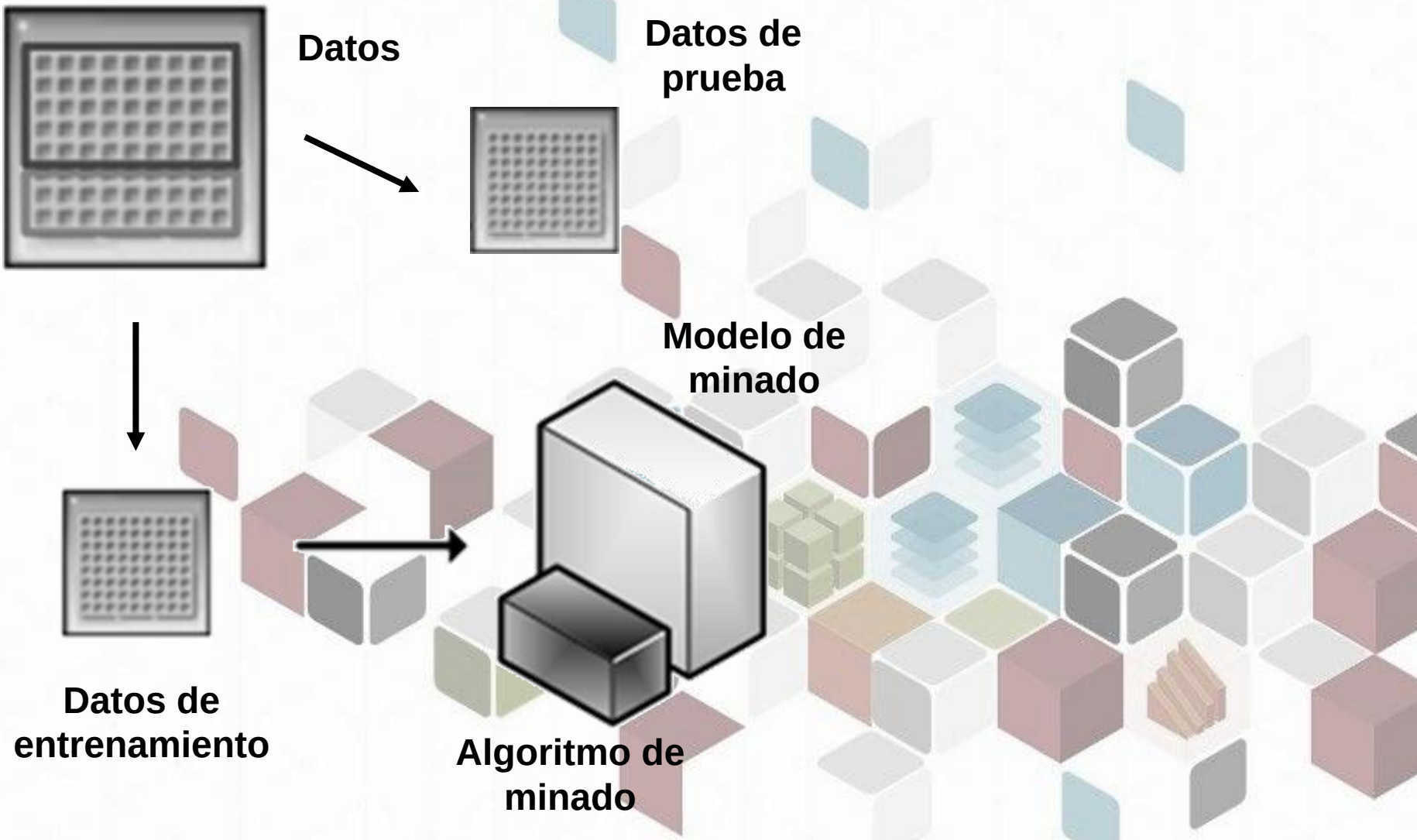
Modelos de Minado



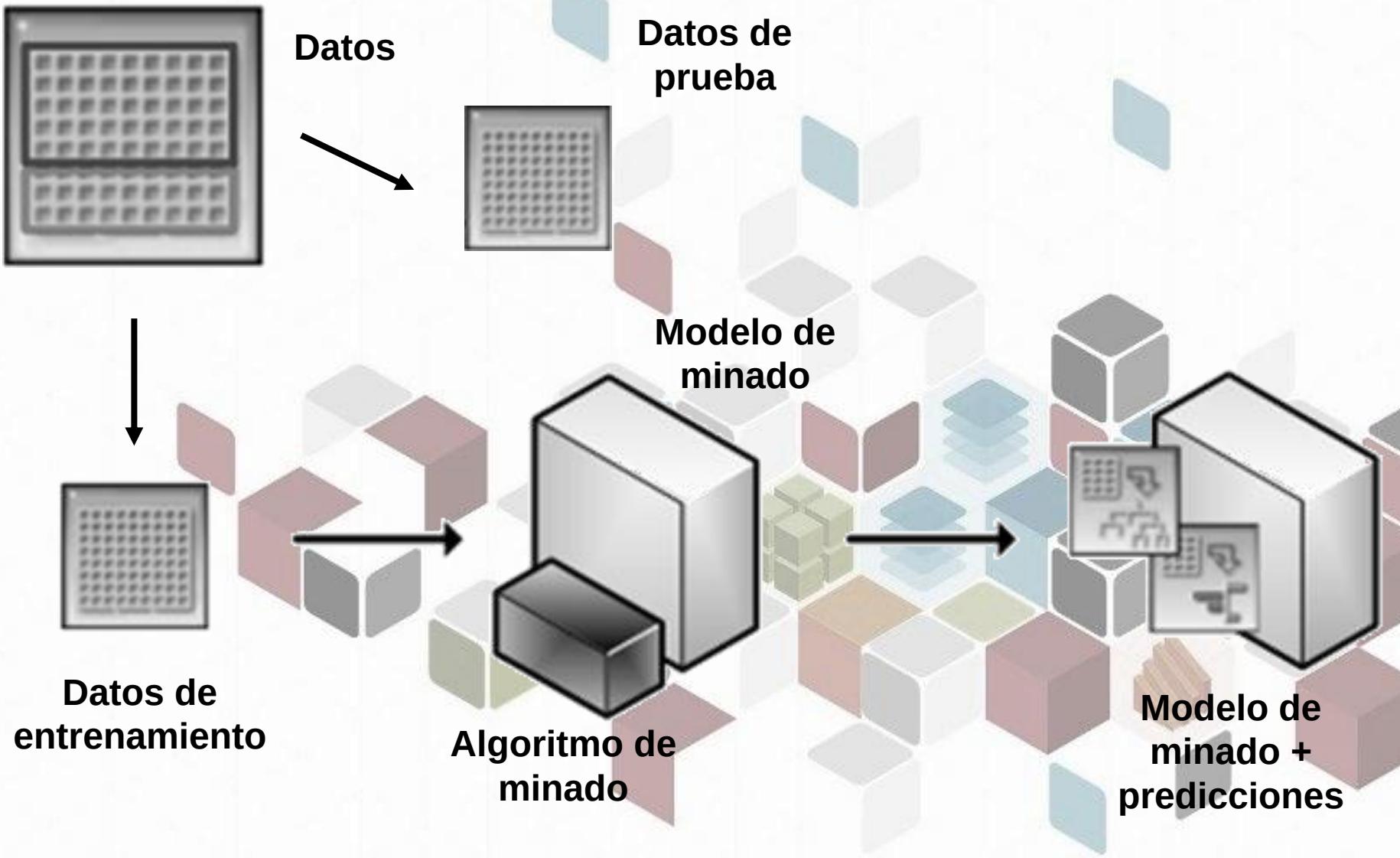
Conocimiento

Modelo de minado

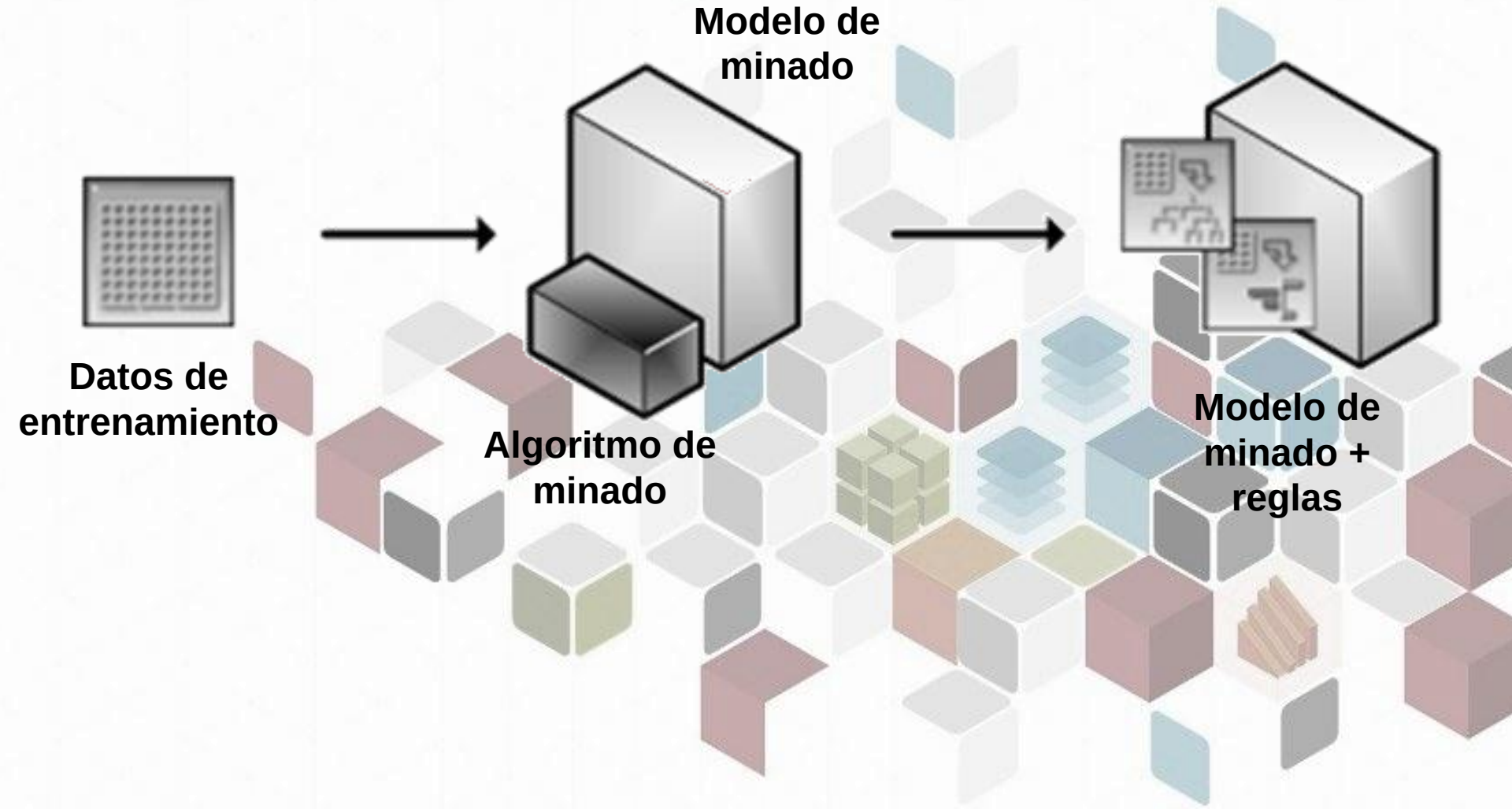
Modelos de Minado



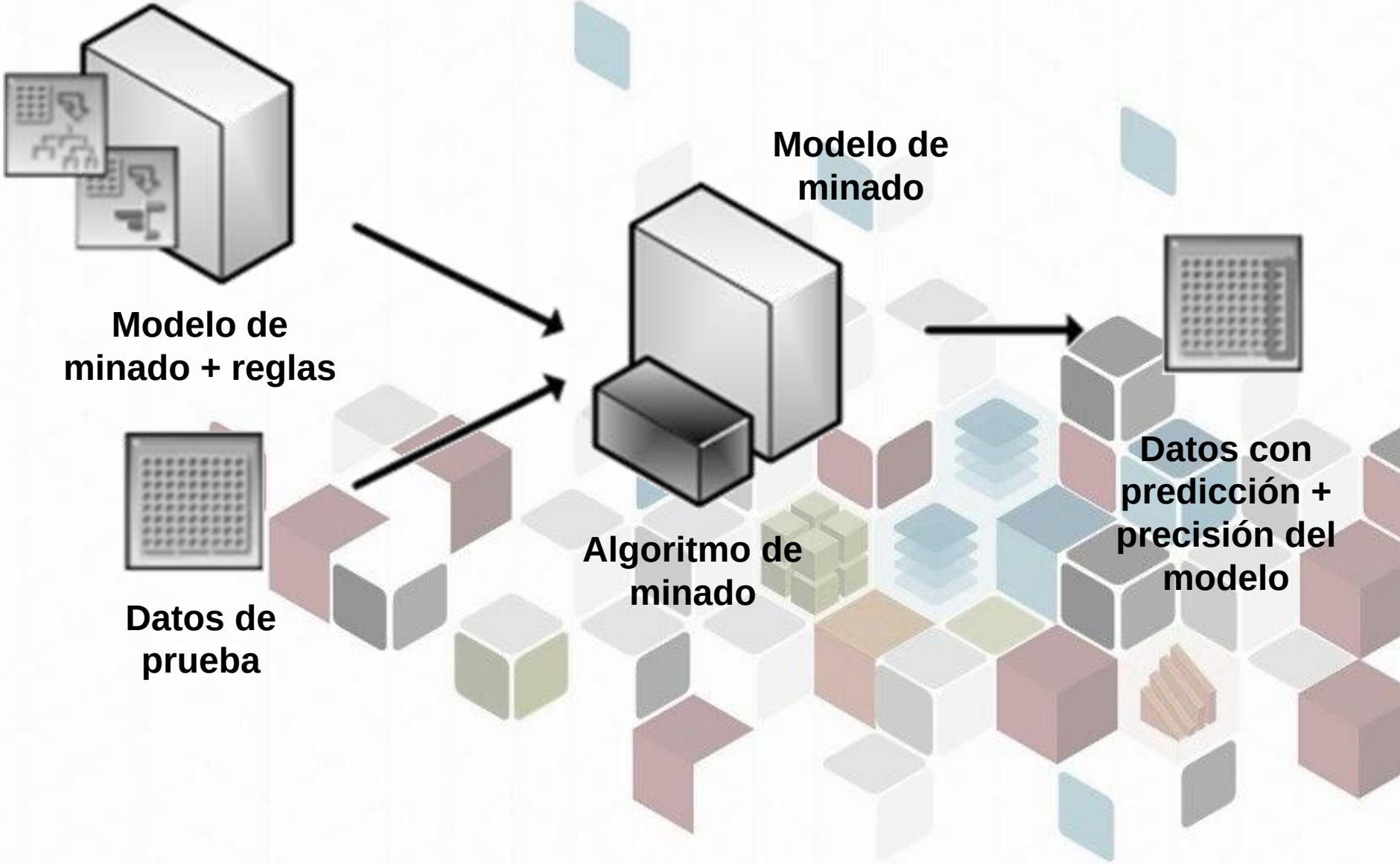
Modelos de Minado



Modelos de Minado



Modelos de Minado



Mecanismos de aprendizaje

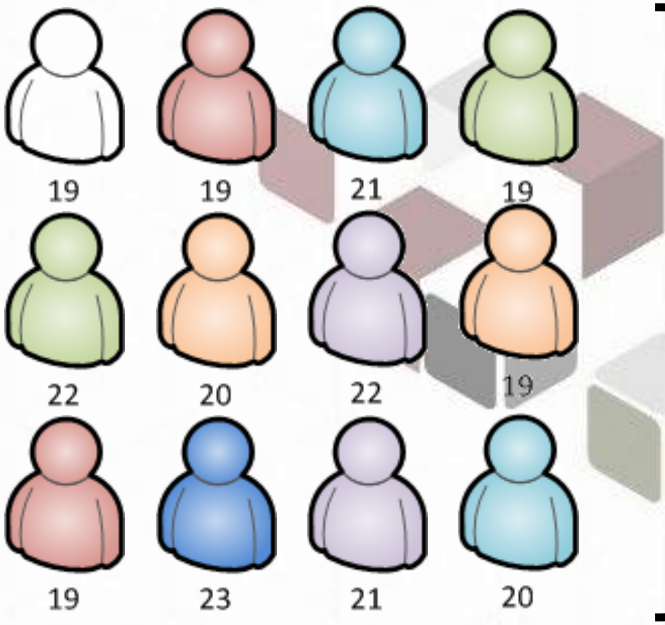
- **Aprendizaje supervisado**
  - Se asocia un elemento representado por un vector $x = (x_1, \ldots, x_n)$ a una de las r_k clases de la variable C. Dichas etiquetas ya están determinadas con anterioridad.
- **Aprendizaje no supervisado**
  - Las etiquetas o clases existentes se desconocen así como también la pertenencia de cada elemento a la clase. Los algoritmos se encargan de determinar ambos problemas.
- **Aprendizaje semi supervisado**
  - Es una combinación de los dos anteriores.

Categorización de las tareas de minería de datos

- **Tareas Descriptivas**
  - Caracterizar, describir las propiedades generales de los datos
- **Tareas Predictivas**
  - Realizar inferencias en los datos para hacer predicciones.

¿Qué tipo de patrones pueden minarse?

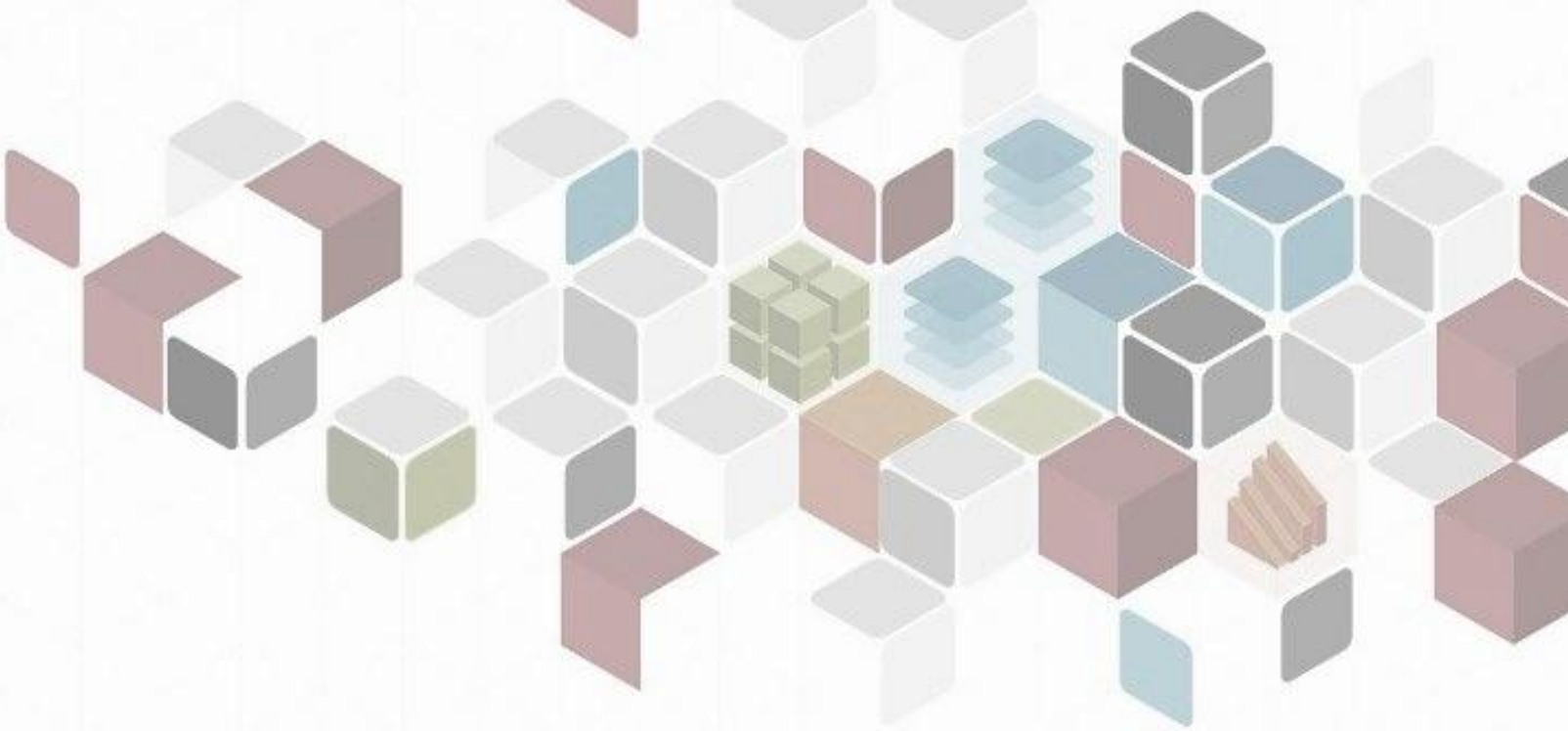
- **Caracterización**
  - Resumir las características generales de una clase objetivo



Los clientes son personas entre 19 y 20 años

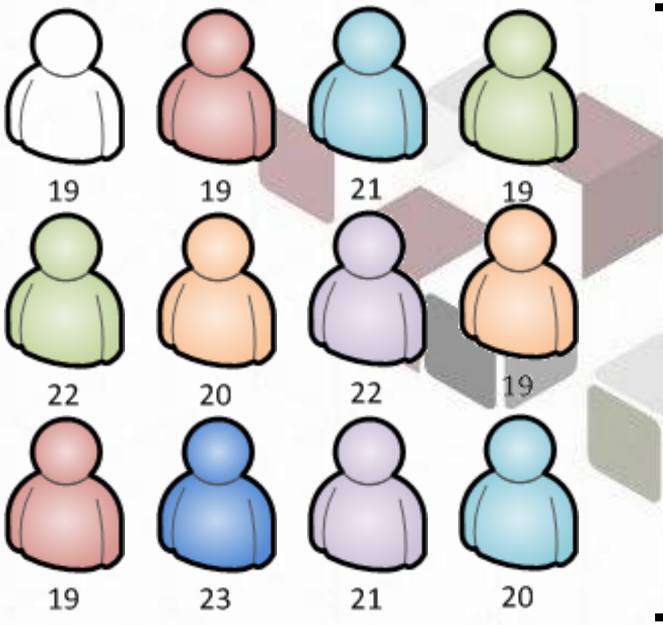
¿Qué tipo de patrones pueden minarse?

- **Discriminar**
  - Comparar las características de una clase objetivo con la(s) clase(s) con las que contrasta.



¿Qué tipo de patrones pueden minarse?

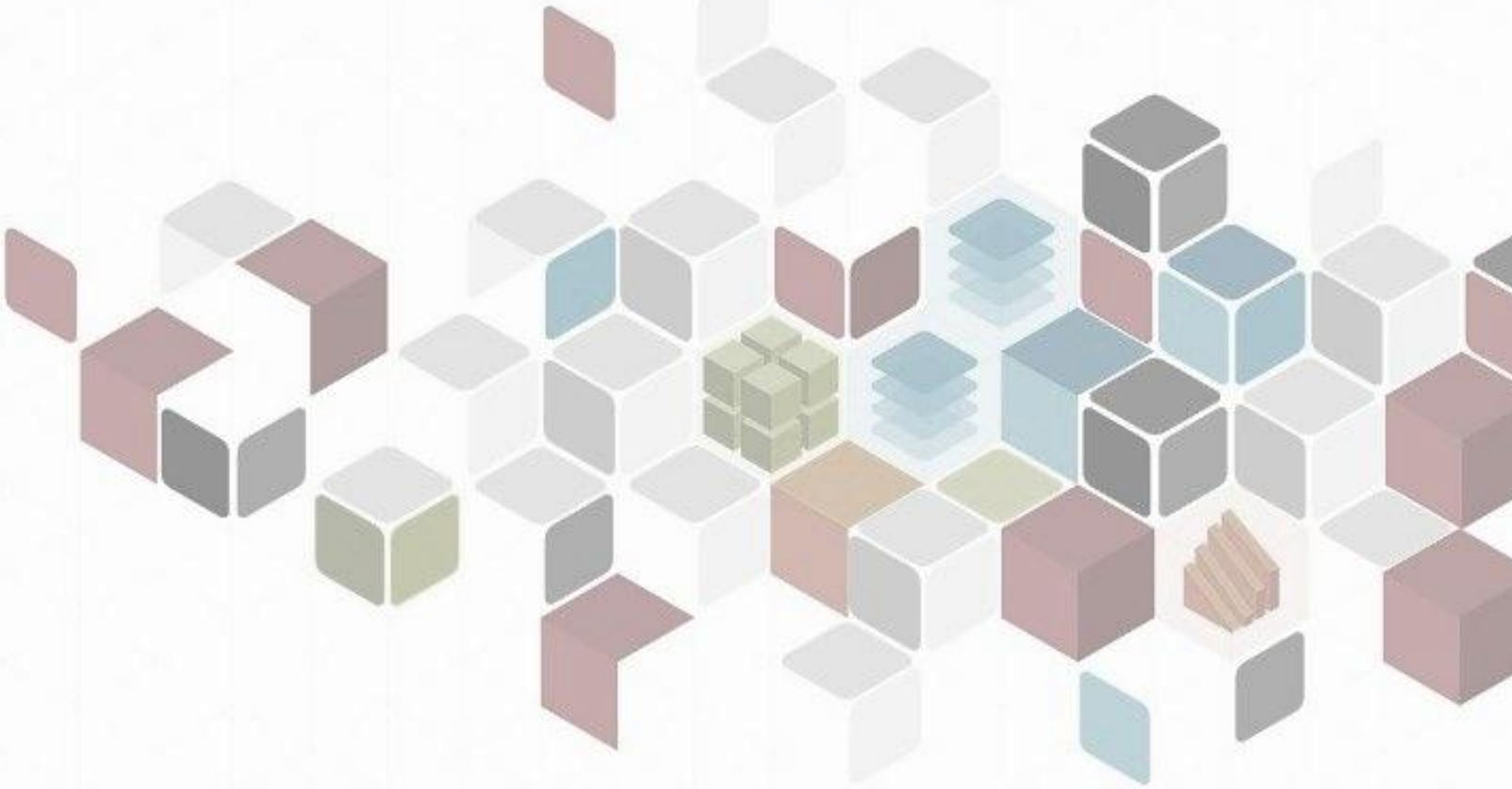
- **Discriminar**
  - Comparar las características de una clase objetivo con la(s) clase(s) con las que contrasta.



Los clientes que les gusta el color rojo, tienen 19 años y son los más jóvenes que los de otros colores

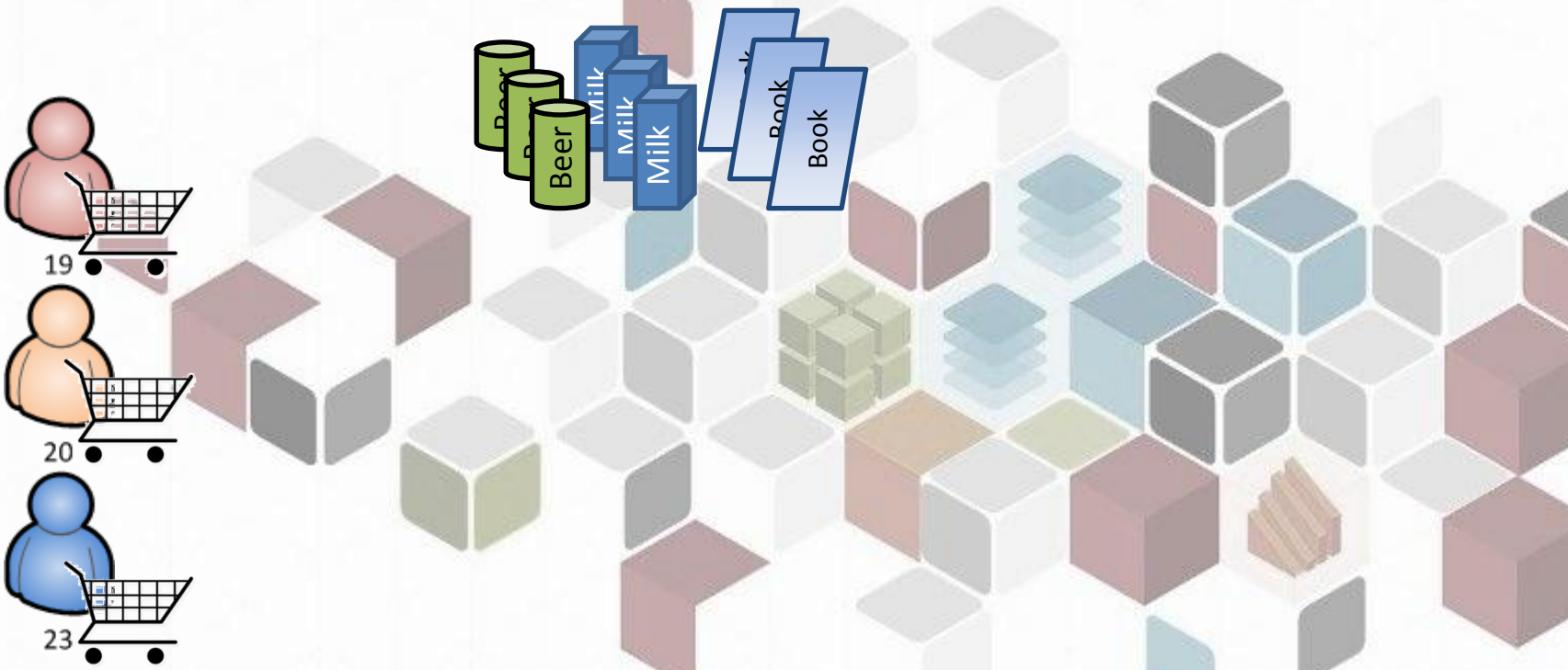
¿Qué tipo de patrones pueden minarse?

- **Asociaciones**
  - Descubrir objetos que ocurren juntos en forma frecuente.



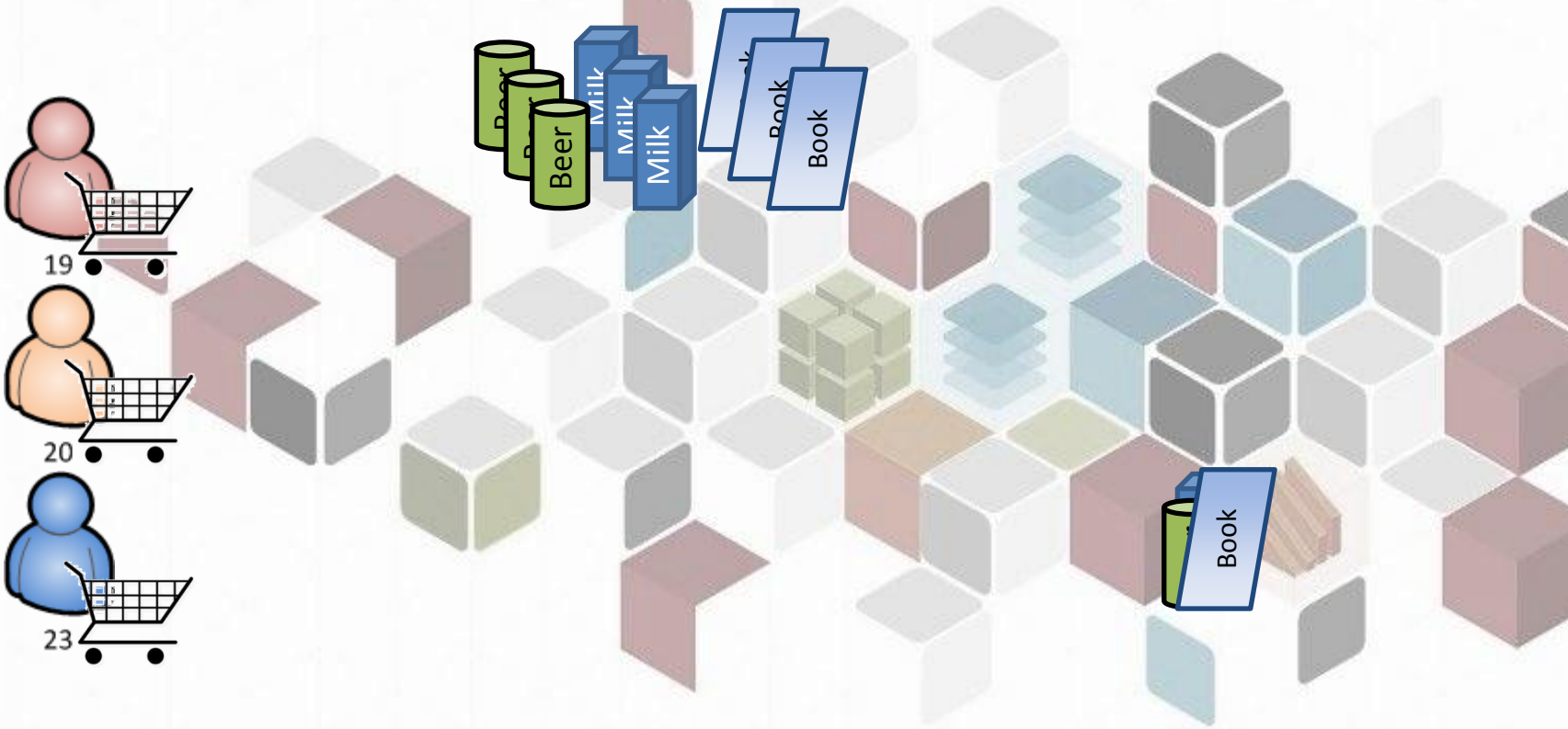
¿Qué tipo de patrones pueden minarse?

- **Asociaciones**
  - Descubrir objetos que ocurren juntos en forma frecuente.



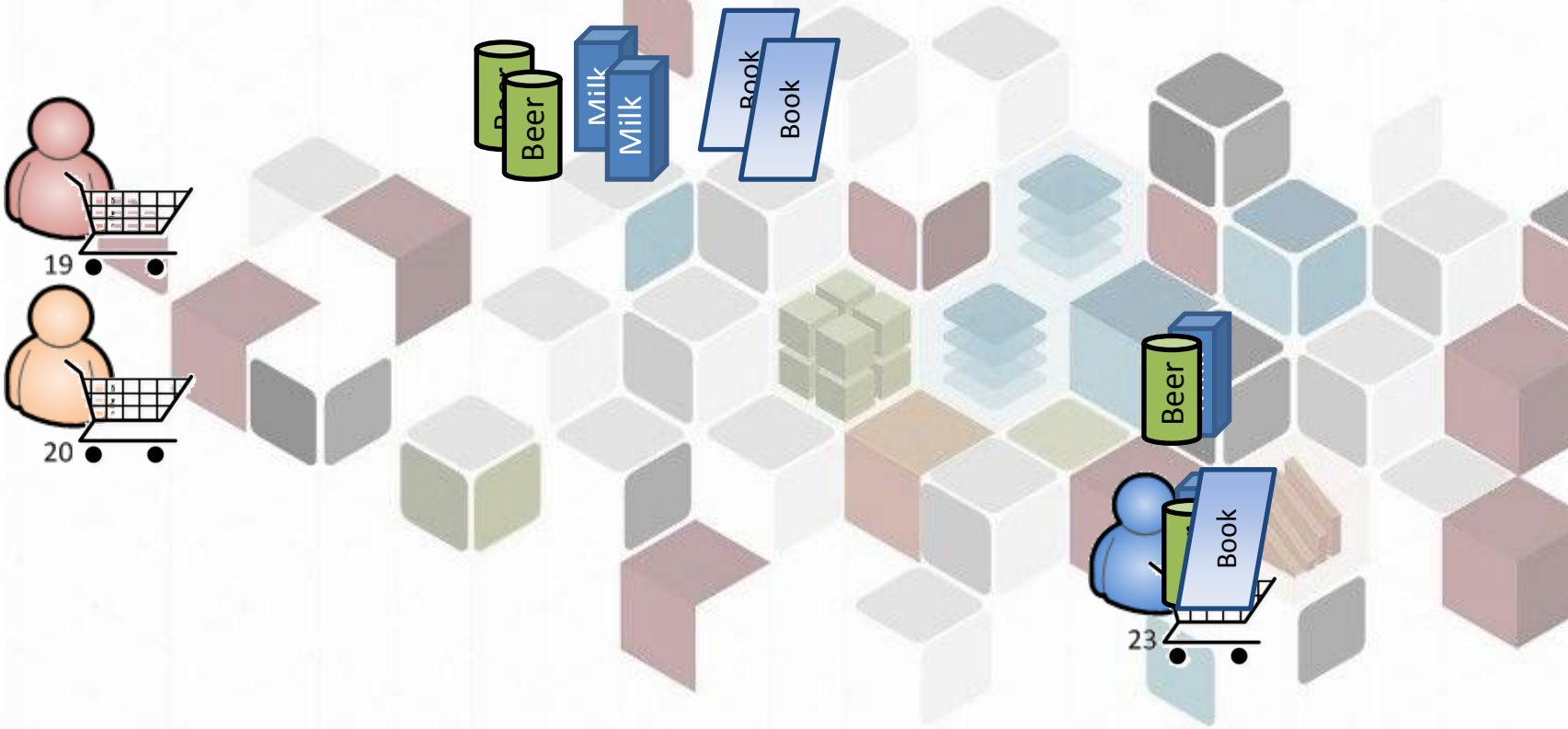
¿Qué tipo de patrones pueden minarse?

- **Asociaciones**
  - Descubrir objetos que ocurren juntos en forma frecuente.



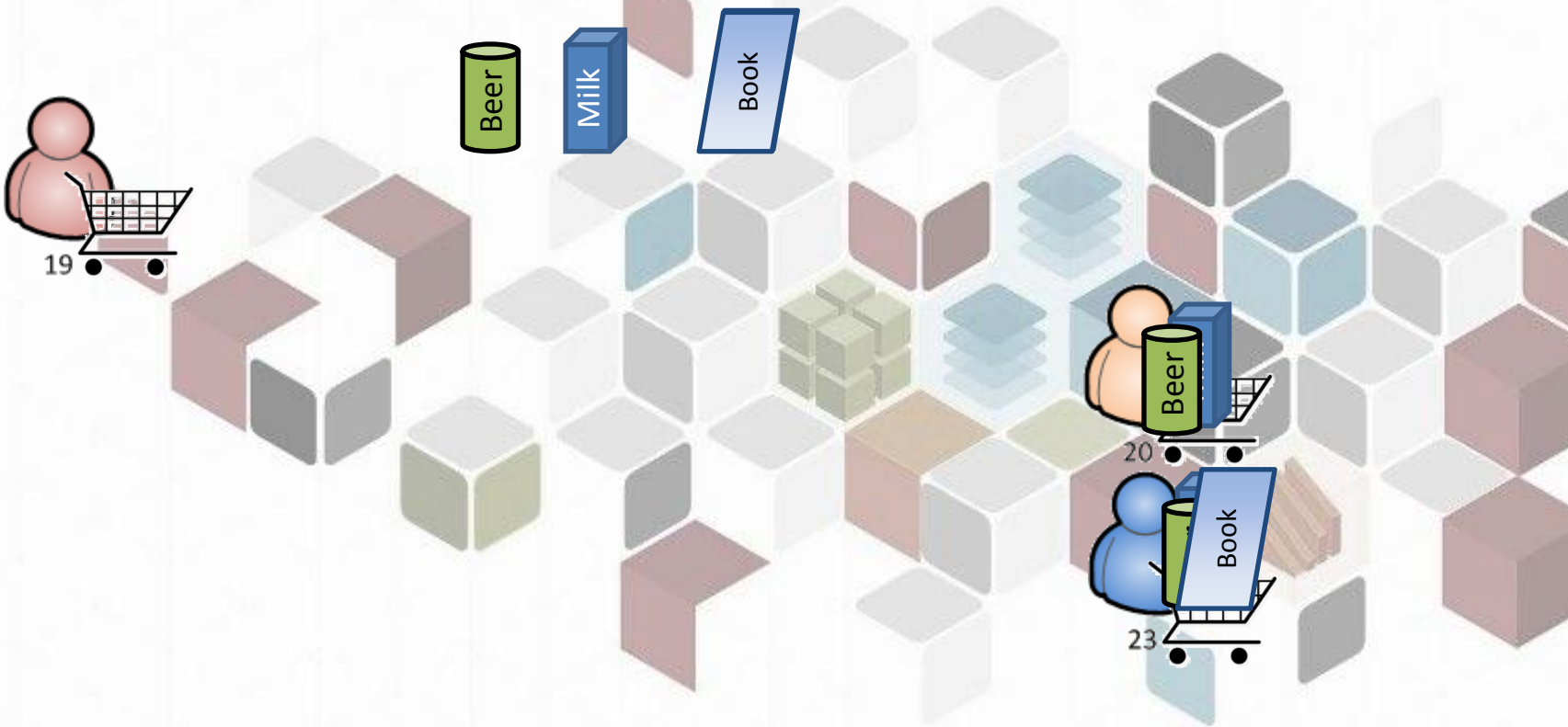
¿Qué tipo de patrones pueden minarse?

- **Asociaciones**
  - Descubrir objetos que ocurren juntos en forma frecuente.



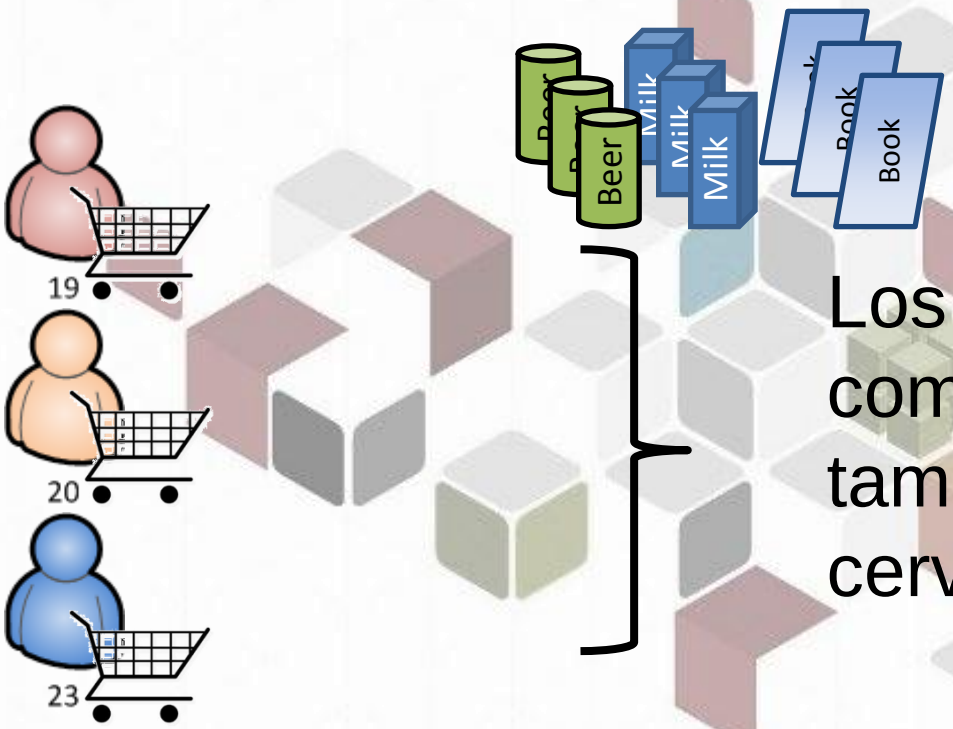
¿Qué tipo de patrones pueden minarse?

- **Asociaciones**
  - Descubrir objetos que ocurren juntos en forma frecuente.



¿Qué tipo de patrones pueden minarse?

- **Asociaciones**
  - Descubrir objetos que ocurren juntos en forma frecuente.

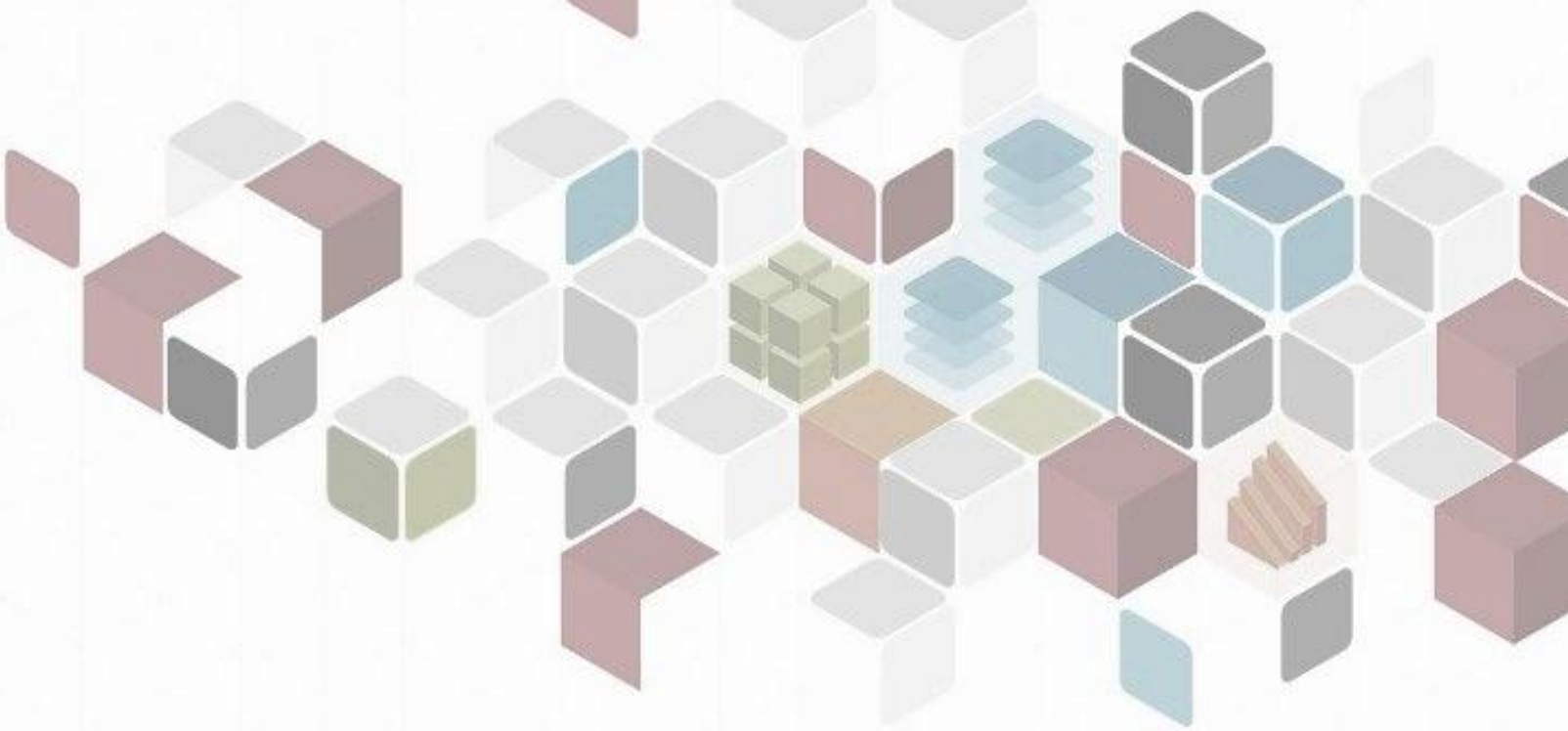


¿Qué tipo de patrones pueden minarse?

- **Asociaciones**
  - Descubrir objetos que ocurren juntos en forma frecuente.
  - Reglas de asociación “A priori”
  - Redes neuronales

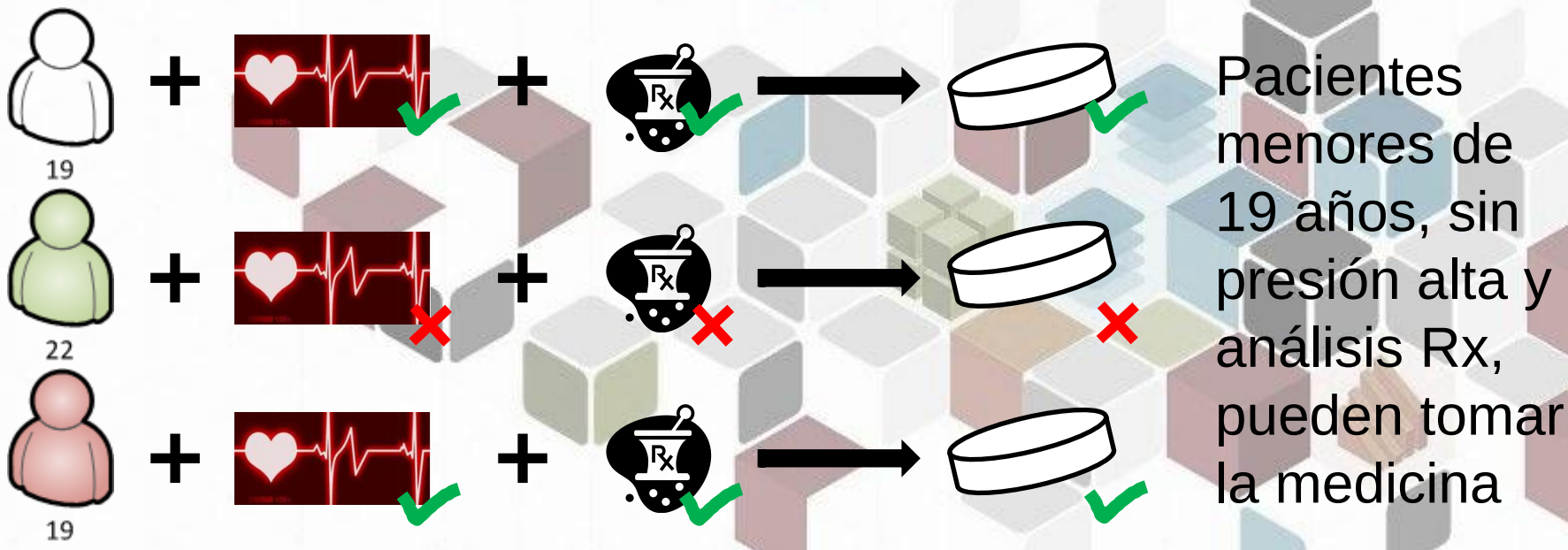
¿Qué tipo de patrones pueden minarse?

- **Clasificación (aprendizaje supervisado)**
  - Encontrar un modelo para predecir la clasificación de datos no observados (conociendo las etiquetas).



¿Qué tipo de patrones pueden minarse?

- **Clasificación (aprendizaje supervisado)**
  - Encontrar un modelo para predecir la clasificación de datos no observados (conociendo las etiquetas).

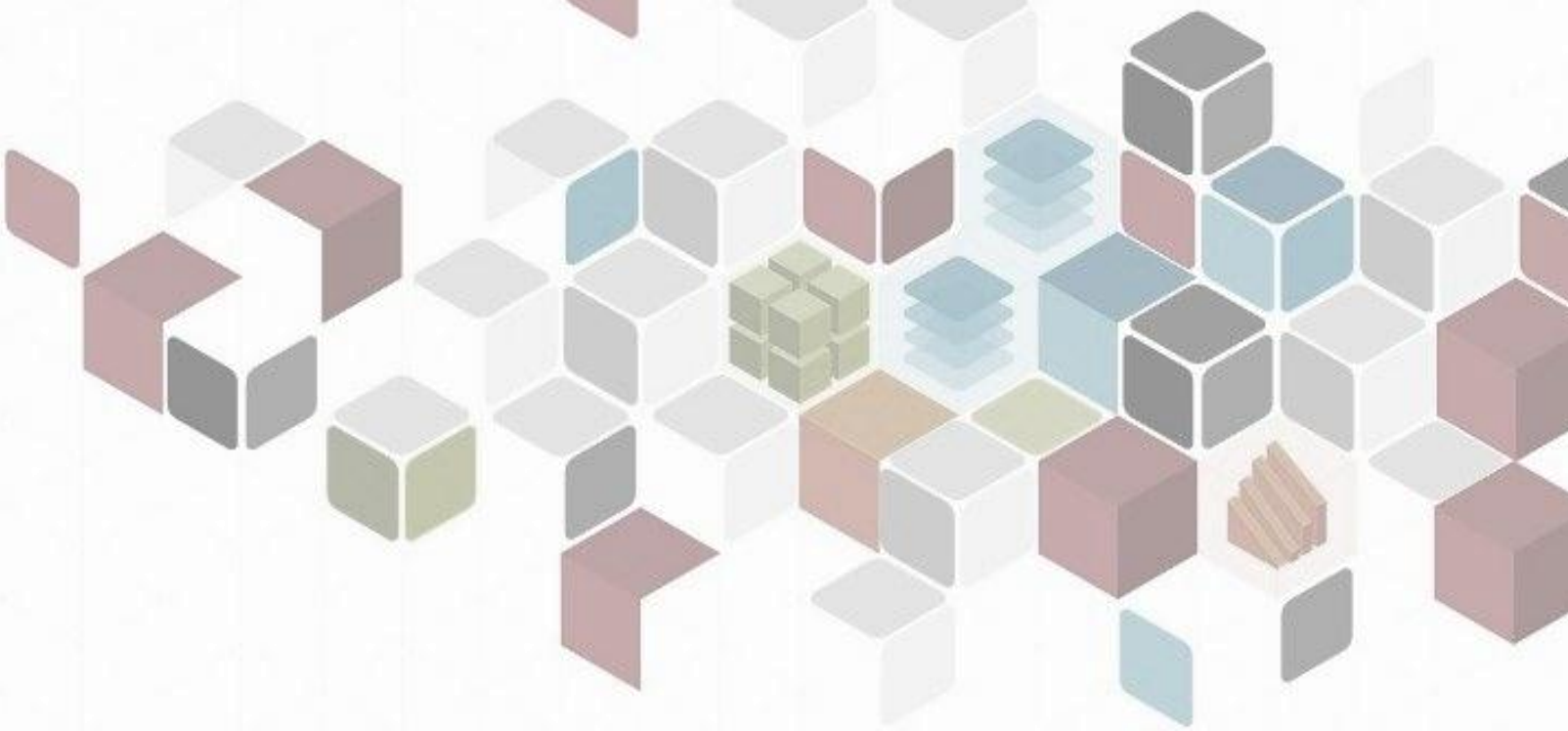


¿Qué tipo de patrones pueden minarse?

- **Clasificación (aprendizaje supervisado)**
  - Encontrar un modelo para predecir la clasificación de datos no observados (conociendo las etiquetas).
  - Algoritmos de clasificación
  - Árboles de clasificación
  - Redes neuronales
  - K-vecinos más próximos
  - Clasificador ingenuo de Bayes.

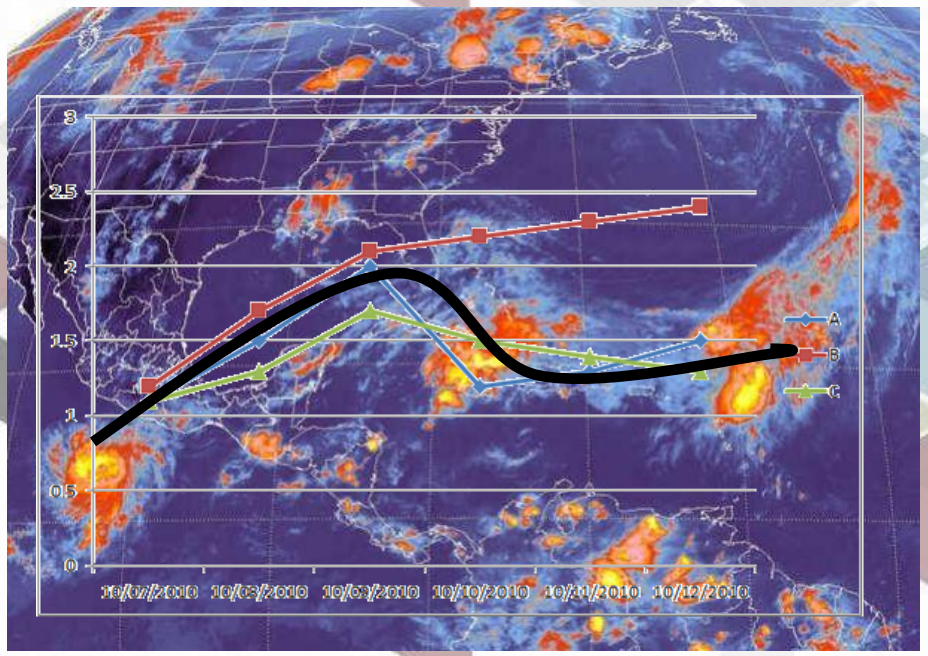
¿Qué tipo de patrones pueden minarse?

- **Regresión**
  - Encontrar un modelo para predecir los valores numéricos de datos no observados.



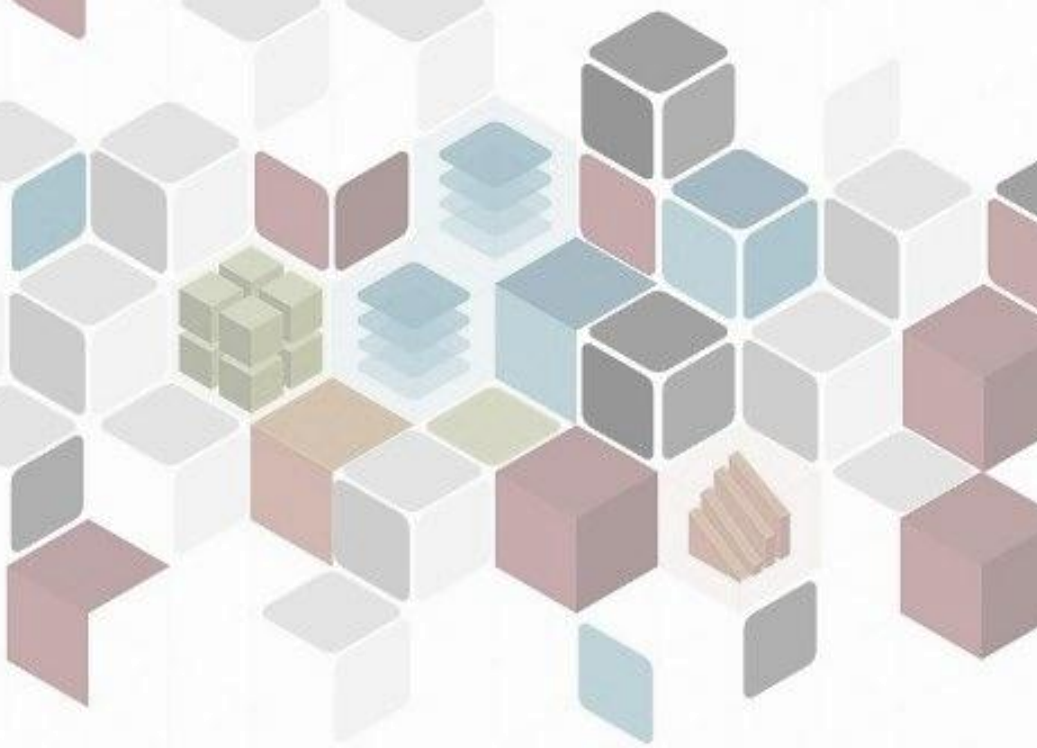
¿Qué tipo de patrones pueden minarse?

- **Regresión**
  - Encontrar un modelo para predecir los valores numéricos de datos no observados.



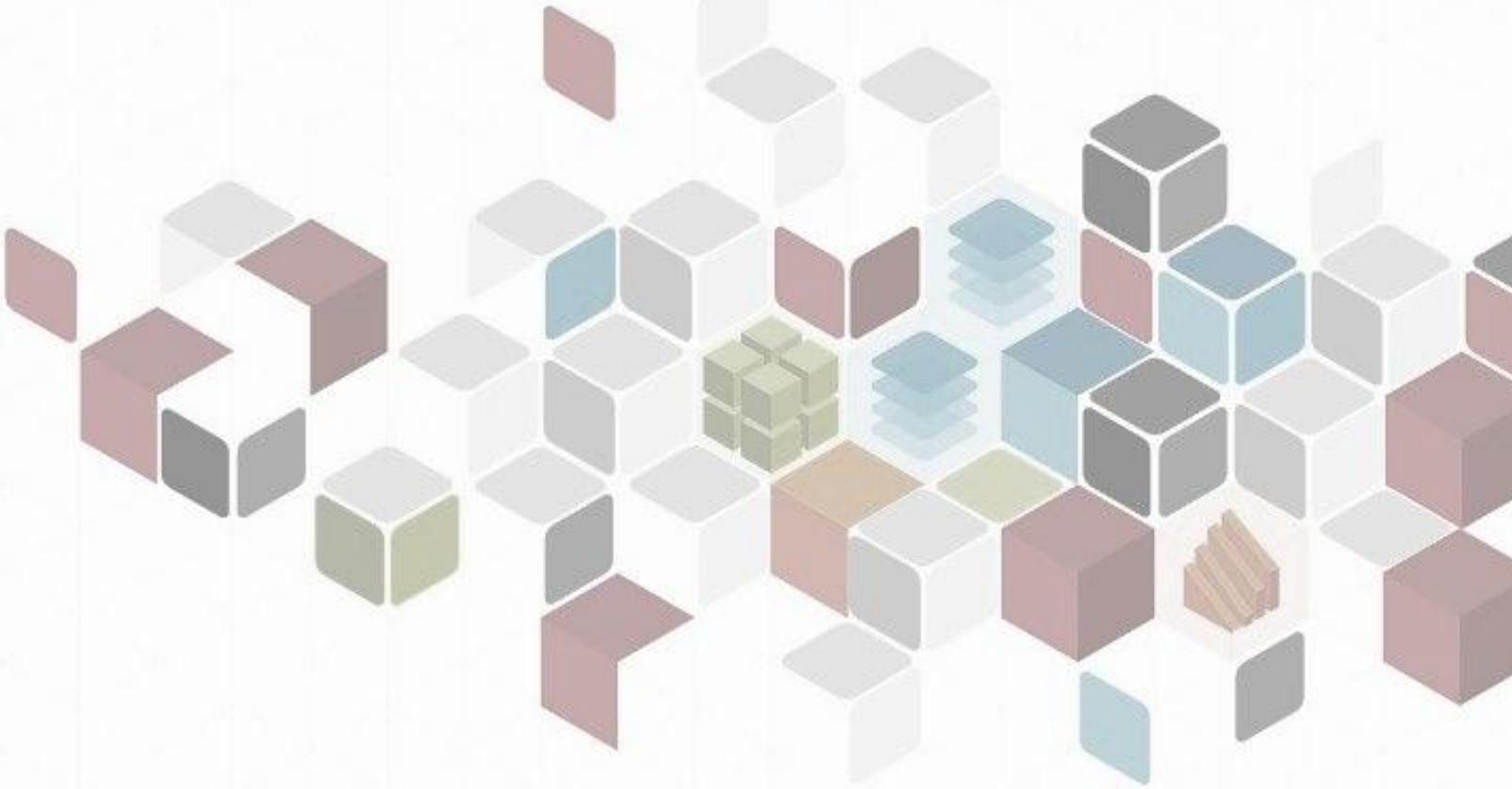
¿Qué tipo de patrones pueden minarse?

- **Regresión**
  - Encontrar un modelo para predecir los valores numéricos de datos no observados.
  - Regresión lineal
  - Regresión logística
  - Árboles de regresión
  - Redes neuronales.



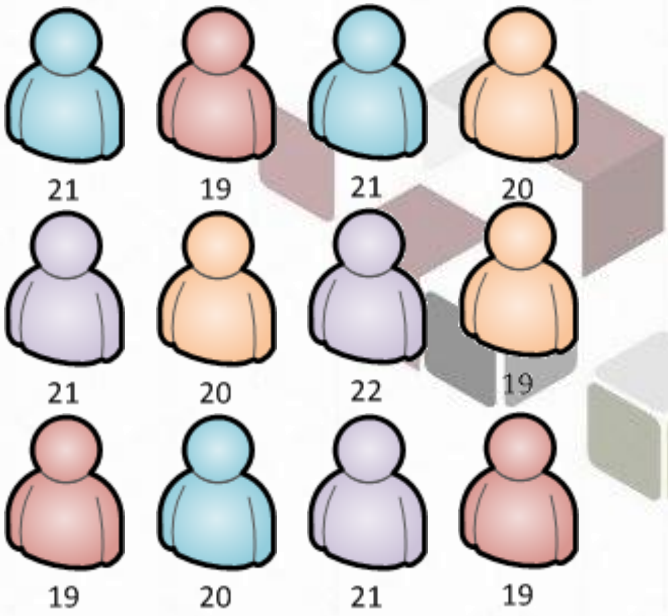
¿Qué tipo de patrones pueden minarse?

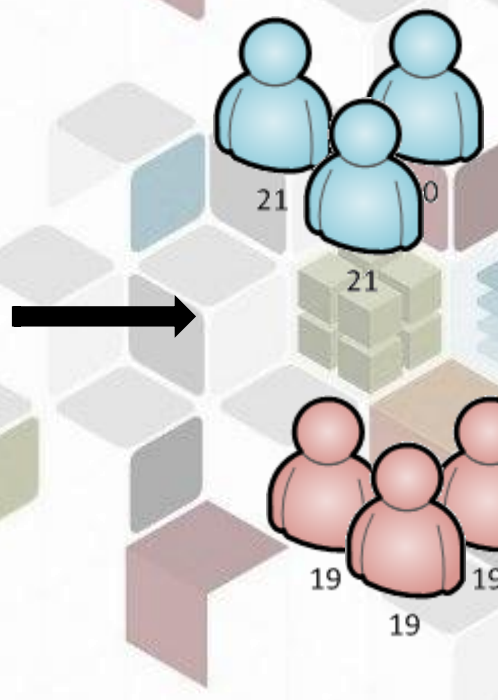
- **Clustering (aprendizaje no supervisado)**
  - Agrupar datos en clases sin conocer su clasificación previa.



¿Qué tipo de patrones pueden minarse?

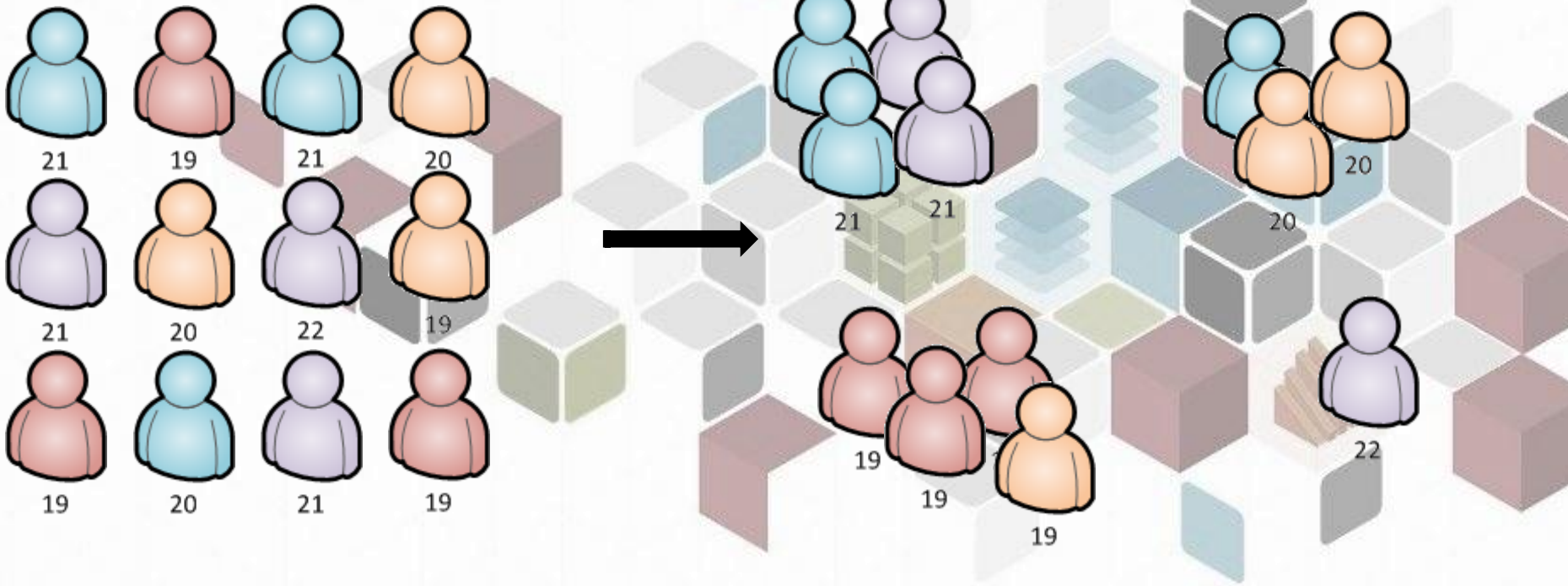
- **Clustering (aprendizaje no supervisado)**
  - Agrupar datos en clases sin conocer su clasificación previa.





¿Qué tipo de patrones pueden minarse?

- **Clustering (aprendizaje no supervisado)**
  - Agrupar datos en clases sin conocer su clasificación previa.

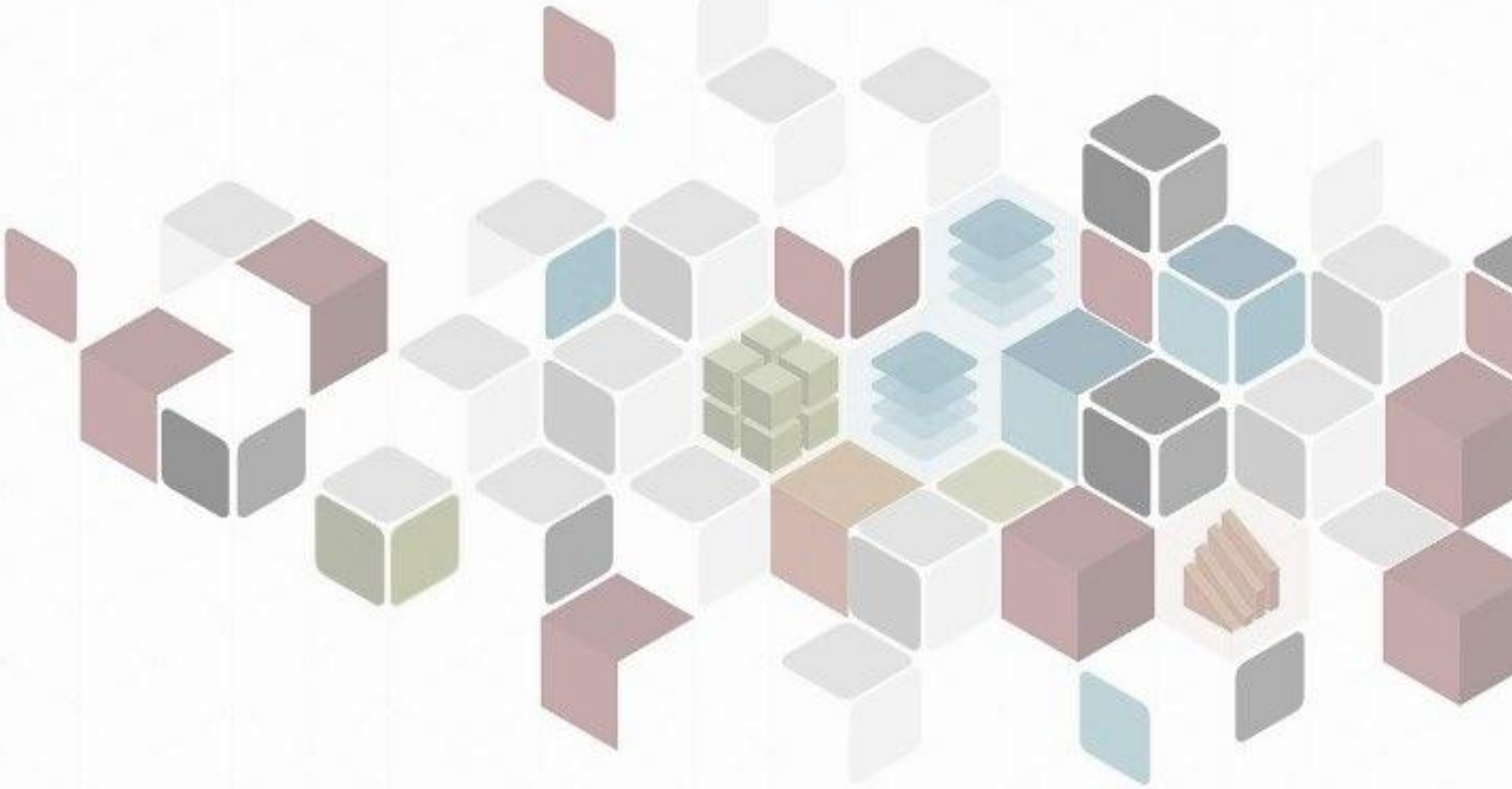


¿Qué tipo de patrones pueden minarse?

- **Clustering (aprendizaje no supervisado)**
  - Agrupar datos en clases sin conocer su clasificación previa.
  - Algoritmo de K-medias
  - Algoritmo de C-medias difusas
  - Agrupamiento jerárquico
  - Mezcla de Gaussianas

¿Qué tipo de patrones pueden minarse?

- **Análisis de datos anormales (Outlier analysis)**
  - Identificar y explicar las excepciones

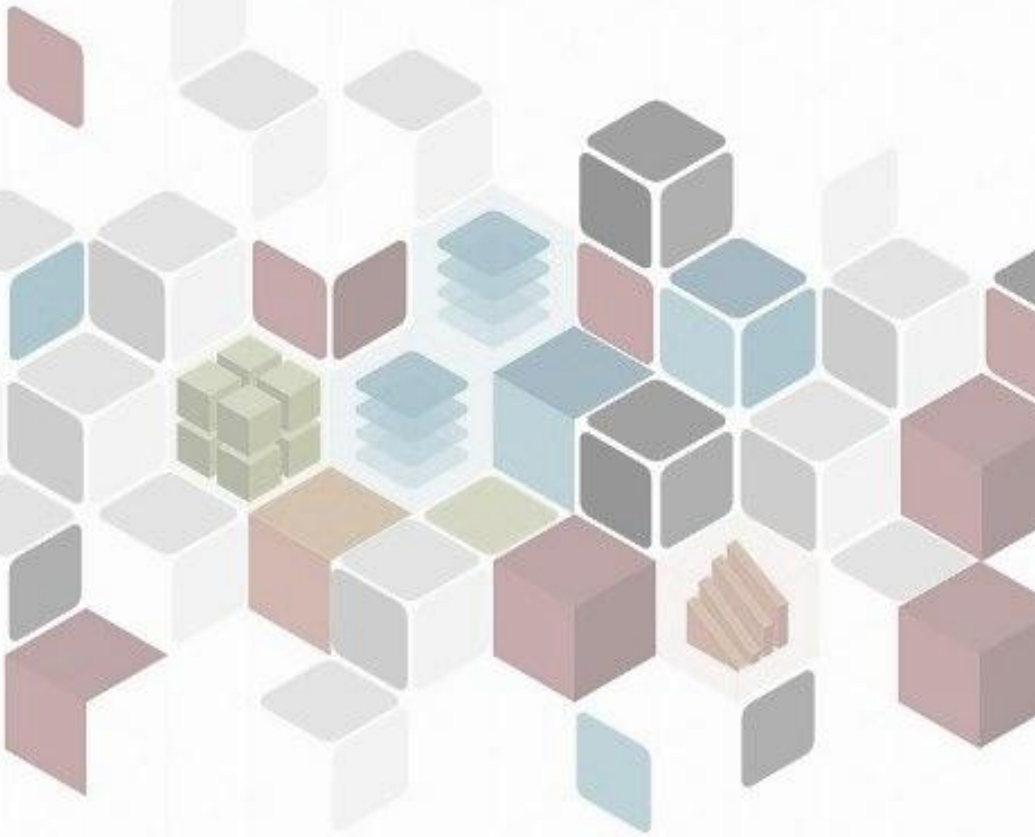


¿Qué tipo de patrones pueden minarse?

- **Análisis de datos anormales (Outlier analysis)**
  - Identificar y explicar las excepciones

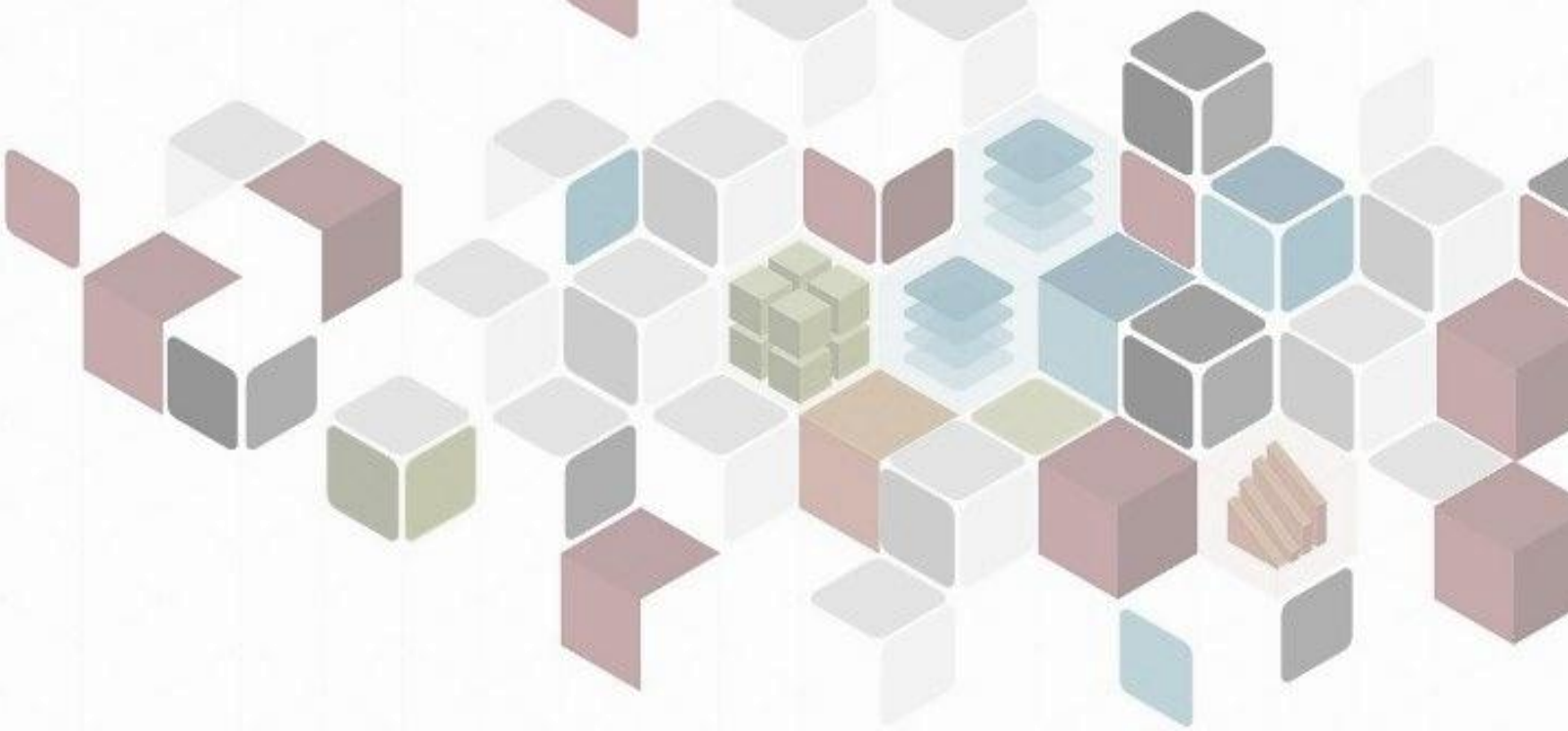
¿Qué tipo de patrones pueden minarse?

- **Análisis de datos anormales (Outlier analysis)**
  - Identificar y explicar las excepciones
  - Árboles de decisión
  - Redes neuronales
  - Algoritmos genéticos.



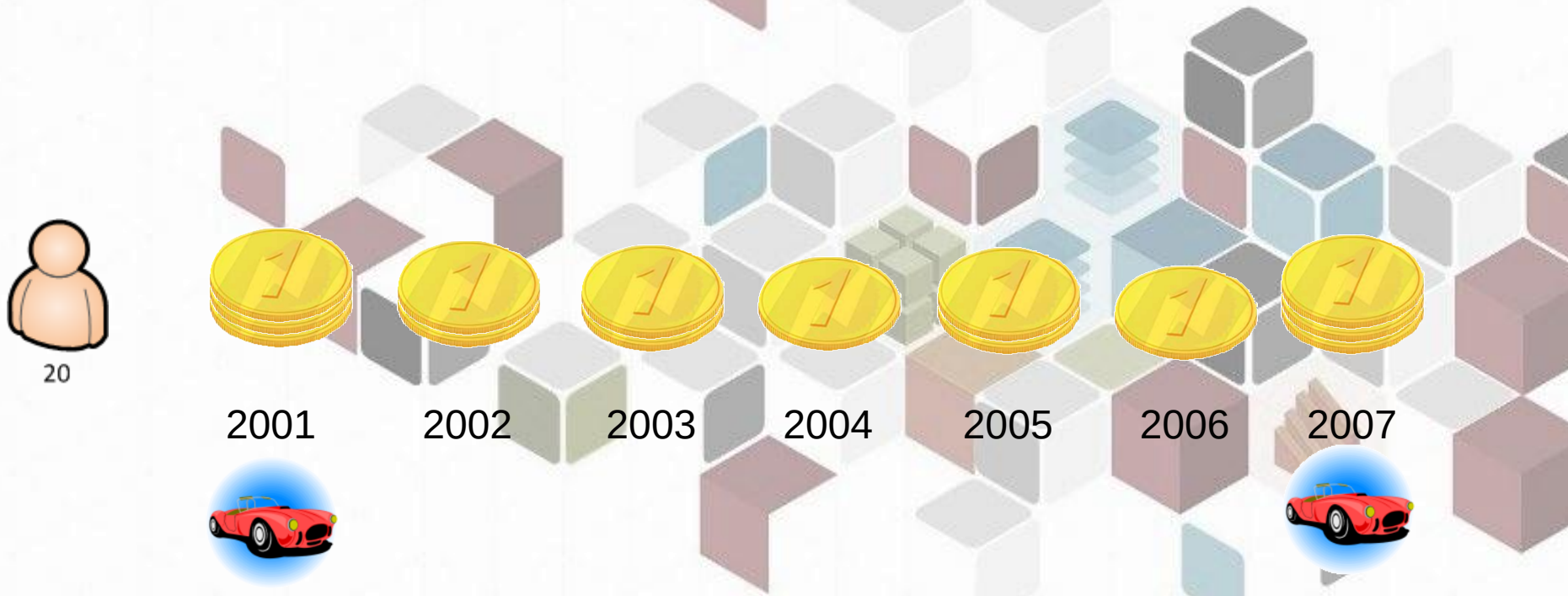
¿Qué tipo de patrones pueden minarse?

- **Tendencias y análisis evolutivo**
  - Tendencias de objetos que varían su comportamiento a través del tiempo o modelos regulares



¿Qué tipo de patrones pueden minarse?

- **Tendencias y análisis evolutivo**
  - Tendencias de objetos que varían su comportamiento a través del tiempo o modelos regulares



¿Qué tipo de patrones pueden minarse?

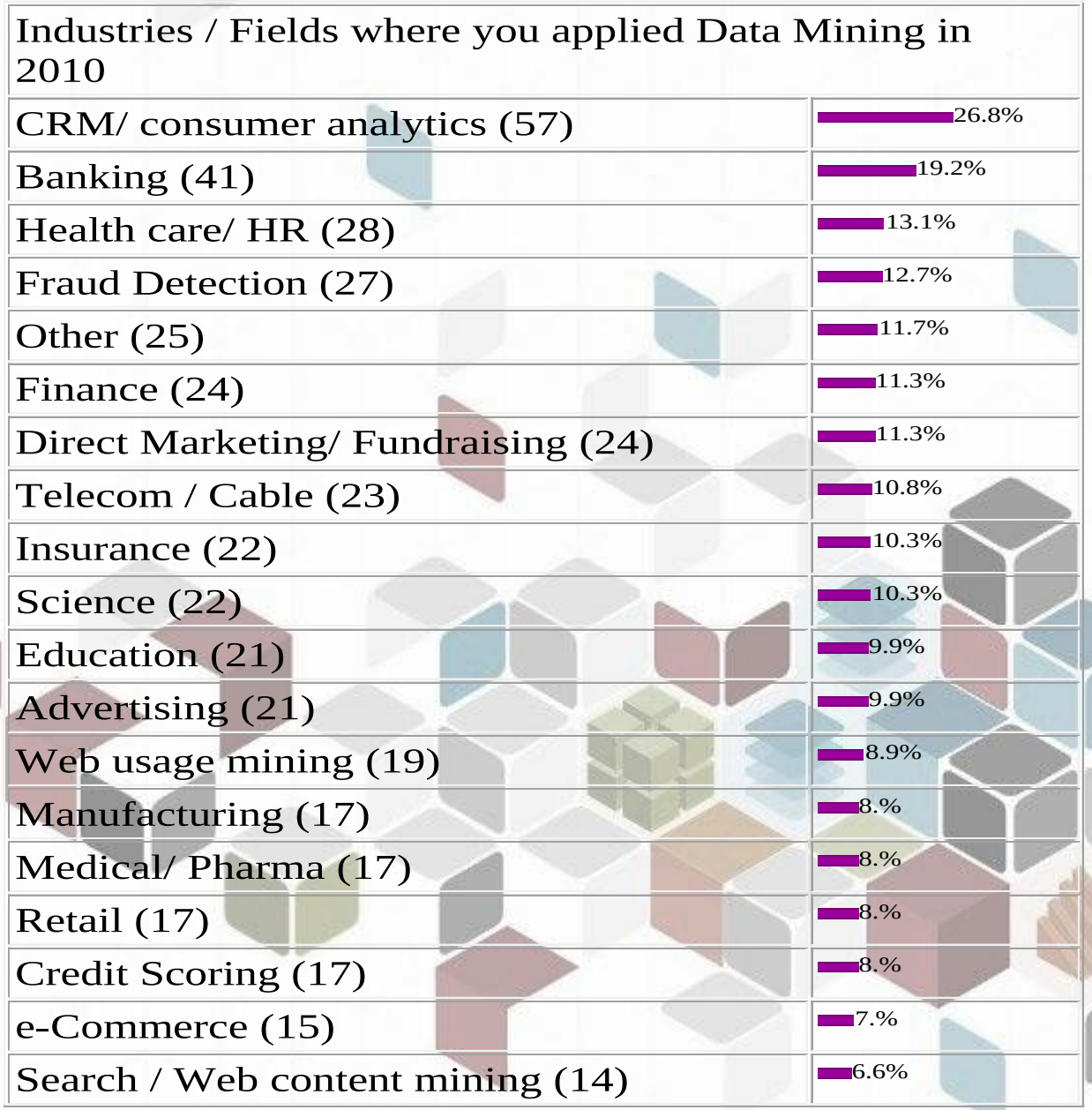
- **Tendencias y análisis evolutivo**
  - Tendencias de objetos que varían su comportamiento a través del tiempo o modelos regulares
  - AutoRegressive Integrated Moving Average

¿Cómo medir la utilidad de los resultados?

- **Medidas objetivas**
  - Basadas en estadísticas y en la estructura de patrones.
- **Medidas subjetivas**
  - Basadas en la expectativa del usuario.



Minería de datos en acción



Complicaciones en minería de datos

- Volumen de datos
  - Se requieren algoritmos con un desempeño razonable
- Resultados interesantes
  - Cómo garantizamos que se obtendrán resultados “interesantes”
- Se requieren habilidades para el proceso
  - Cómo seleccionar la herramienta adecuada o los datos preparados
- Calidad de los datos
  - Cómo interpretamos los resultados a la luz de datos de baja calidad
- Heterogeneidad de los datos
  - Cómo combinamos los datos de múltiples fuentes
- Privacidad
  - Cómo garantizamos la privacidad de los datos individuales

Ética en la minería de datos

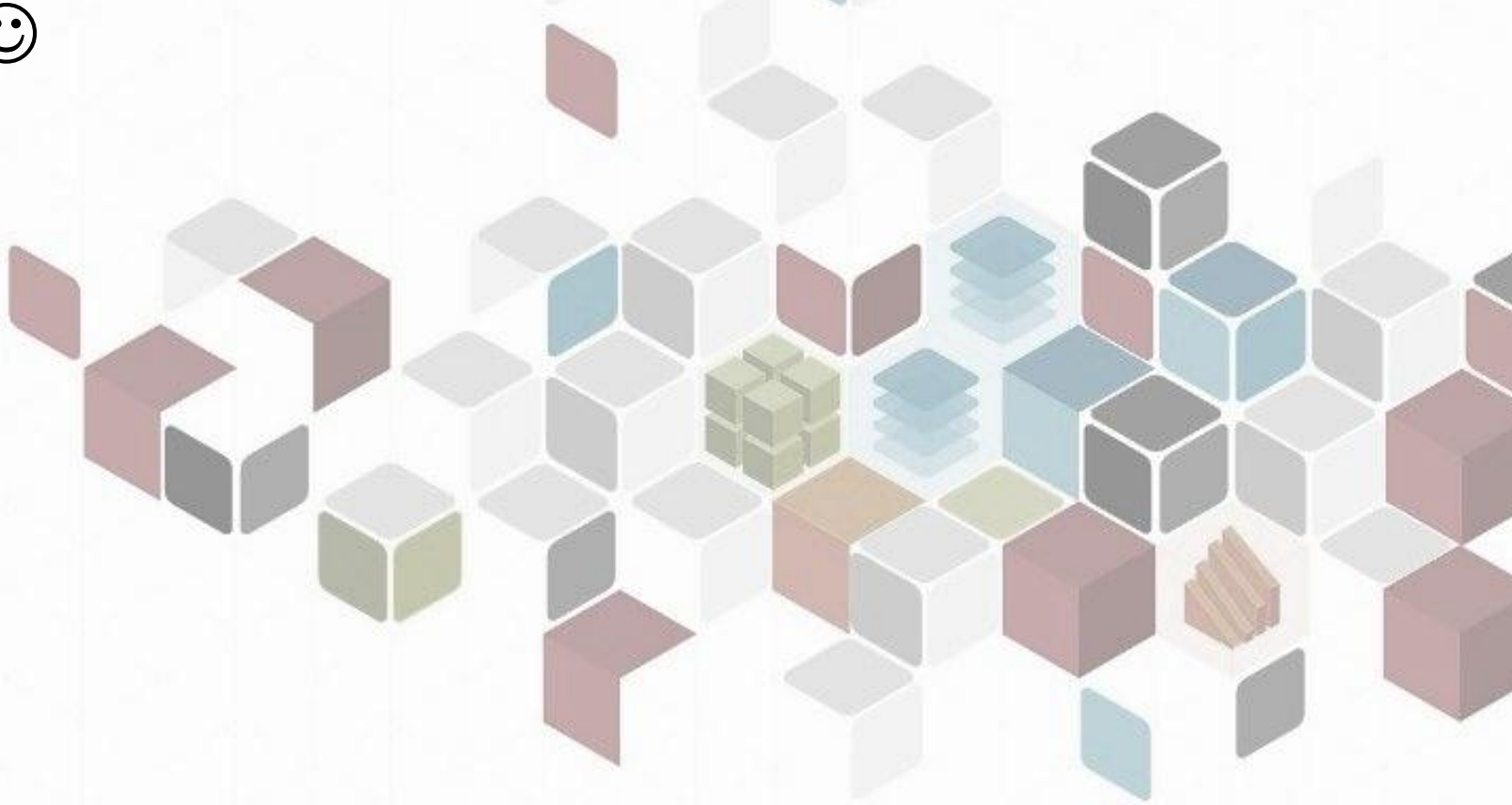
- Colección
- Propósito
- Uso
- Privacidad, seguridad

Ética en la minería de datos

- Proteger los datos privados (ingresos, médicos, creencias – políticas, religiosas, historia laboral, etc)
- Experiencia en el descubrimiento de errores en BD
- Experiencia en trabajo supervisando BD – placas, teléfonos, impuestos. (actividades de empleados: búsqueda de datos personales, familiares, de amigos, de vecinos; para vender información, para cambiar información – multas, impuestos)

Complejidad

Cuando el análisis de algoritmos toma sentido… más ☺



Complejidad

Descripción	Datos Tamaño		Métrica	Dispositivo
Pequeño	10^2	25600	KB	Documento
Chico	10^4	2560000	MB	Memoria USB
Mediano	10^6	256000000	MB	Memoria USB
Grande	10^8	25600000000	GB	Disco duro
Gigante	10^{10}	2560000000000	TB	Arreglos discos duros
Masivo	10^{12}	256000000000000	TB	BD Paralelas / Clusters
Supermasivo	10^{15}	25600000000000000	PB	BD Distribuidas

Taxonomía Huber-Wegman de tamaños de conjuntos de datos (1995)

Complejidad

Orden	Problema
$O(n^{1/2})$	Graficar una gráfica no densa
$O(n)$	
$O(n \log(n))$	Calcular transformadas de Fourier
$O(n^c)$	Resolver una regresión lineal múltiple
$O(n^2)$	Resolver la mayoría de los algoritmos de agrupación
$O(a^n)$	Detectar valores anormales multivariados.

Complejidad

Numero de operaciones de diversos algoritmos
con complejidades distintas para varios tamaños de datos

n	$n^{1/2}$	n	$n \log(n)$	$n^{3/2}$	n^2
Pequeño	10	10^2	$2x10^2$	10^3	10^4
Chico	10^2	10^4	$4x10^4$	10^6	10^8
Mediano	10^3	10^6	$6x10^6$	10^9	10^{12}
Grande	10^4	10^8	$8x10^8$	10^{12}	10^{16}
Gigante	10^5	10^{10}	10^{11}	10^{15}	10^{20}

Complejidad

Teraflop Grand Challenge Computer
1000 gigaflop

n	$n^{1/2}$	*n*	*n log(n)*	$n^{3/2}$	n^2
Pequeño	10^{-11} segundos	10^{-10} segundos	$2x10^{-10}$ segundos	10^{-9} segundos	10^{-8} segundos
Chico	10^{-10} segundos	10^{-8} segundos	$4x10^{-8}$ segundos	10^{-6} segundos	10^{-4} segundos
Mediano	10^{-9} segundos	10^{-6} segundos	$6x10^{-6}$ segundos	.001 segundos	1 segundos
Grande	10^{-8} segundos	10^{-4} segundos	$8x10^{-4}$ segundos	1 segundos	2.8 horas
Gigante	10^{-7} segundos	.01 segundos	.1 segundos	16.7 minutos	3.2 años

Complejidad

Asumiendo que una maquina ejecuta una operación de punto flotante por lectura de un dato, se tiene lo siguiente:

1KB toman en ser analizados .00000001 segundos con una maquina de 1Teraflop
10KB toman en ser analizados .00000001 segundos con una maquina de 1Teraflop
100KB toman en ser analizados .0000001 segundos con una maquina de 1Teraflop
1 MB toman en ser analizados .000001 segundos con una maquina de 1Teraflop
10 MB toman en ser analizados .000001 segundos con una maquina de 1Teraflop
100 MB toman en ser analizados .00001 segundos con una maquina de 1Teraflop
1 GB toman en ser analizados .0001 segundos con una maquina de 1Teraflop
10 GB toman en ser analizados .001 segundos con una maquina de 1Teraflop
100 GB toman en ser analizados .01 segundos con una maquina de 1Teraflop
1 TB toman en ser analizados 1 segundo con una maquina de 1Teraflop

Esto asumiendo un comportamiento lineal del algoritmo que analiza los datos y un escenario ideal (lo cual es poco factible, pero para efectos prácticos, lo asumimos).

Complejidad

Así que lo siguiente es, llenar la siguiente tabla:

Complejidad/Datos	$n^{½}$	n	n log(n)	n^2	n^3
Pequeño					
Chico					
Mediano					
Grande					
Masivo					
Súper masivo					

Esto asumiendo la nueva tabla de tamaños de datos para cada clasificación, así como el desempeño computacional de una computadora que posee un Intel Core i7 980 XE a 109 GFLOPS

La paradoja de Rhine: un buen ejemplo de cómo no llevar a cabo investigación científica

- Joseph Rhine, parapsicólogo en los 50's lanzó la hipótesis de que ciertas personas tenían percepción extrasensorial (PES)
- Llevó a cabo un "experimento" en el que a las personas se les preguntaba que adivinaran 10 cartas ocultas – rojas y azules.
- Descubrió que casi 1 en 1000 tenían PES – eran capaces de adivinar las 10 cartas.
  - ¿Cuál es la prob. de que una persona adivine 10 cartas?

La paradoja de Rhine: un buen ejemplo de cómo no llevar a cabo investigación científica

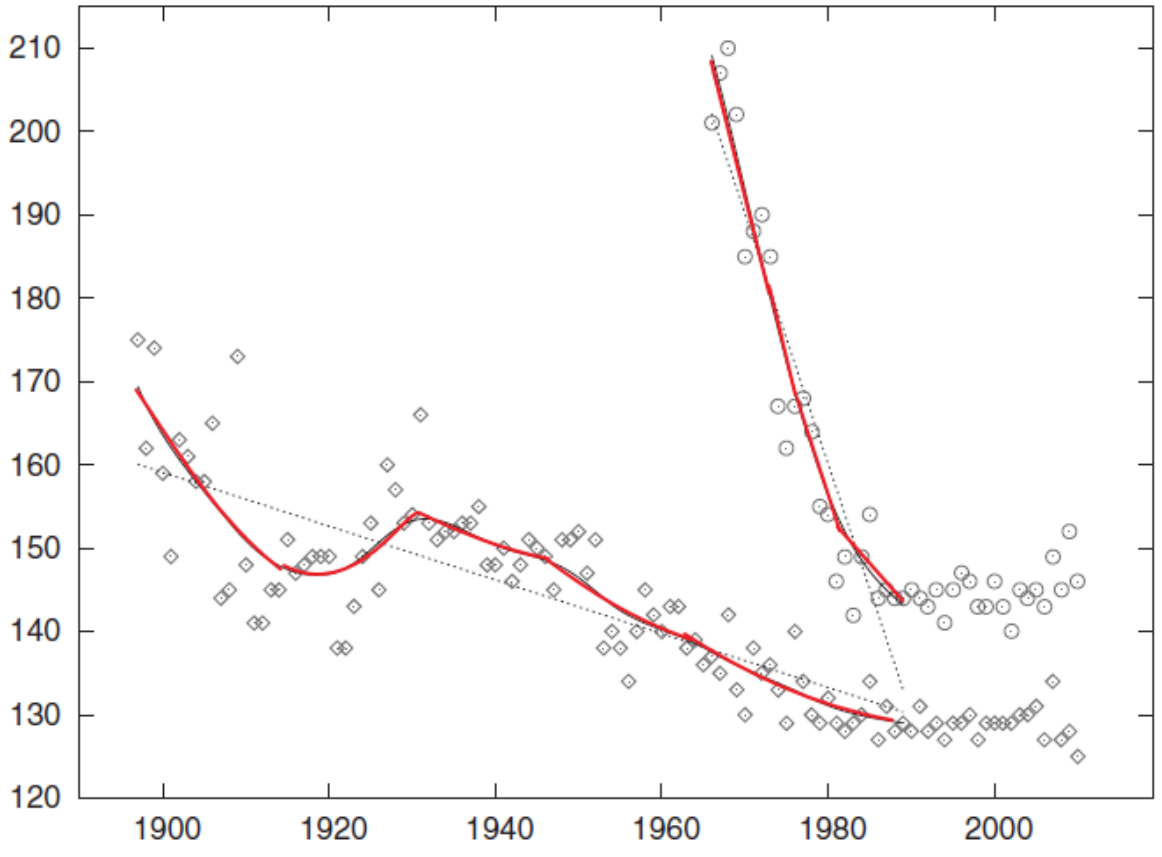
- Entonces le dijo a estas personas que tenían PES y las llamó para realizar otro experimento del mismo tipo: descubrió que casi todos habían perdido su PES.
- ¿Qué concluyó?

La paradoja de Rhine: un buen ejemplo de cómo no llevar a cabo investigación científica

- Entonces le dijo a estas personas que tenían PES y las llamó para realizar otro experimento del mismo tipo: descubrió que casi todos habían perdido su PES.
- ¿Qué concluyó?
  - Concluyó que no debes mencionarle a la gente que tiene percepción extrasensorial pues esto causa que la pierdan….

Contexto

- Observen la siguiente tendencia
  - ¿Qué indica?
  - ¿Qué predicen?



Hallazgos

- En cuanto a patrones:
  - **Excelente**: hallazgos nuevos y útiles
  - **Mejor**: hallazgos nuevos pero no tan útiles
  - **Bueno**: ningún hallazgo
  - **Malo**: falsos negativos
    - Determinar que no existe algo que si existe
  - **Peor**: falsos positivos
    - Determinar que existe algo que no existe