

A stylized sunburst graphic in shades of purple and blue, located in the top-left corner of the slide.

# Métodos descriptivos

Máster en Ciencia de Datos

**UCO**  
ONLINE

Four horizontal bars of varying lengths in yellow and red, located at the bottom of the slide.

# Aplicaciones de métodos descriptivos (Parte 2)

*Minería de patrones poco frecuentes*

# Minería de patrones poco frecuentes

- Definición formal

- Patrones (Itemsets)

- Representa algún tipo de homogeneidad y regularidad en los datos. Conjuntos de ítems que aparecían frecuentemente en los datos, denotando regularidad.
    - En ocasiones, lo que se está buscando no tiene que ser frecuente:
      - Enfermedad rara
      - Algún tipo de anomalía o mal funcionamiento
      - Etc

- Regla de asociación

- Relación entre *itemsets* de un patrón P de la forma  $X \rightarrow Y$
    - Si el antecedente X de la regla se satisface, entonces es bastante probable que el consecuente Y de la regla también se satisfaga. **Esto es lo más importante ahora!**

**MEDIDA DE CONFIANZA COMO BASE**

$$I = \{i_1, \dots, i_n\} \in \Omega$$

$$P \subseteq I$$

$$X \subset P \subseteq I$$

$$Y = P \setminus X$$

# Minería de patrones poco frecuentes

- Espacio de búsqueda
  - Patrones o Itemsets: existen  $(2^k - 1)$  patrones diferentes
  - Reglas de asociación clásicas: un conjunto de datos con  $k$  ítems diferentes tiene un total de  $3^k - 2^{k+1} + 1$  reglas diferentes

| Nivel de estudios | Estado civil | Sexo   | Superó la prueba |
|-------------------|--------------|--------|------------------|
| Primaria          | Soltero      | Hombre | Sí               |
| Primaria          | Soltera      | Mujer  | No               |
| Universitarios    | Casado       | Hombre | Sí               |
| Universitarios    | Casada       | Mujer  | Sí               |
| Primaria          | Divorciado   | Hombre | No               |
| Secundaria        | Casada       | Mujer  | Sí               |

# Minería de patrones poco frecuentes

- Reglas poco frecuentes

- Si primaria y soltero ENTONCES hombre

soporte = 1/6 confianza = 0.5

- Si divorciado ENTONCES primaria

soporte = 1/6 confianza = 1.0

- Si primaria ENTONCES divorciado

soporte = 1/6 confianza = 0.33

| Nivel de estudios | Estado civil | Sexo   | Superó la prueba |
|-------------------|--------------|--------|------------------|
| Primaria          | Soltero      | Hombre | Sí               |
| Primaria          | Soltera      | Mujer  | No               |
| Universitarios    | Casado       | Hombre | Sí               |
| Universitarios    | Casada       | Mujer  | Sí               |
| Primaria          | Divorciado   | Hombre | No               |
| Secundaria        | Casada       | Mujer  | Sí               |

# *Minería de patrones poco frecuentes*

- ¿Aplicar algoritmos clásicos?
  - Apriori
  - FP-Growth
  - LCM
  - Etc
- Problema:
  - Enorme espacio de soluciones
  - Post procesado muy costoso
- Soluciones:
  - Apriori Inverse
  - Apriori Rare
  - Etc

A stylized sunburst graphic in shades of purple and blue, located in the top-left corner of the slide. It features a semi-circle on the left with several rays extending outwards to the right.

¡Gracias!

**UCO**  
ONLINE

A horizontal bar at the bottom of the slide consisting of three equal-width rectangles of yellow, red, and blue, representing the colors of the Spanish flag.