



TAREA 2: A PRIORI

En esta tarea a partir de los datos proporcionados por el INE sobre graduandos 2023, se aplicó la misma regla dada en clase $\text{support}=0.2$, $\text{confidence} = 0.5$.

Primero se exploró las 130 reglas que aprobaron dicho umbral de manera general para esto se obtuvieron los siguientes resultados interesantes:

- Para todo el set de datos se tienen mas Licenciaturas en nivel educativo.
- Para todo el set de datos se tienen más mujeres en graduandos en el 2023.
- Más graduandos en el sector privado.

Como segundo punto se quiso explorar el filtro por departamento de Guatemala y resto del país esto únicamente para ingenierías. Los resultados interesantes son los siguientes:

a) Ingenierías en Guatemala (capital)

- $\{\text{Sector}=\text{Público}\} \Rightarrow \{\text{Departamento}=\text{Guatemala}\}$
- $\{\} \Rightarrow \{\text{Pueblo_Pertenencia}=\text{Ladino}\}.$
- $\{\} \Rightarrow \{\text{Sexo}=\text{Hombre}\}$

Lectura: en Ingenierías (capital) se mantiene el problema de “pueblo” ignorado en el privado (muy fuerte). Además, cuando el “pueblo” está ignorado predominan hombres. Entre mujeres en Ingeniería-Guatemala, “Ladino” está sobrerrepresentado.

Considerar que hay fuerte relación entre sector público y el departamento de Guatemala en ingenierías, aunque predomina la cantidad en universidades privadas según la relación general

La mayoría de estudiantes de ingeniería son latinos y hombres esto para el departamento de Guatemala.

b) Ingenierías en Resto del país

- $\text{Pueblo}=\text{Ignorado} \Rightarrow \text{Sector}=\text{Privado}: \text{conf}\approx 0.990, \text{lift}\approx 1.192.$
- $\text{Sector}=\text{Privado} \Rightarrow \text{Pueblo}=\text{Ignorado}: \text{conf}\approx 0.734, \text{lift}\approx 1.192.$
- $\text{Edad}\geq 30 \Rightarrow \text{Pueblo}=\text{Ignorado}: \text{conf}\approx 0.704, \text{lift}\approx 1.144.$

Lectura: el patrón de privado \leftrightarrow “pueblo” ignorado vuelve a aparecer (muy fuerte). Además, en el resto del país, mayores de 30 muestran mayor probabilidad de tener el “pueblo” no consignado.