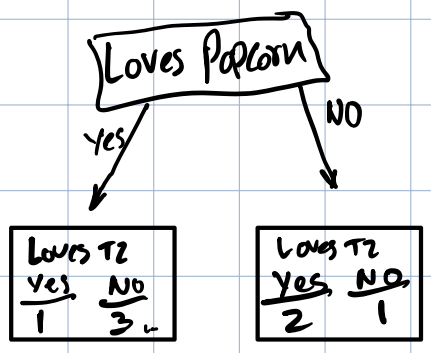
$$C\,Gini = \left(\frac{\#\ de\ si}{Total\ rama}\right)^2 - \left(\frac{\#\ de\ no}{Total\ rama}\right)^2$$

Cuando hay distinto número de personas en cada rama

$$C\,Gini_{WA} = \left(\frac{\#\ of\ People\ in\ left\ leaf}{\#\ of\ People\ both\ leaves}\right)^2 * C_{left} + \left(\frac{\#\ of\ people\ in\ right\ leaf}{\#\ people\ both\ (leaves)}\right)^2 * C_{right}$$

Loves Popcorn	Loves Soda	Age	Loves TRolls
yes	yes	7	NO
yes	NO	12	NO
NO	yes	18	yes
.	.	.	.
.	.		.
.	.		
.	.	.	



$$C_{G_{\text{Loves popcorn yes}}} = 1 - \left(\frac{1}{4}\right)^2 - \left(\frac{3}{1+3}\right)^2 \approx 0{,}375$$

$$C_{G_{\text{Loves popcorn NO}}} = 1 - \left(\frac{2}{3}\right)^2 - \left(\frac{1}{3}\right)^2 \approx 0.44$$

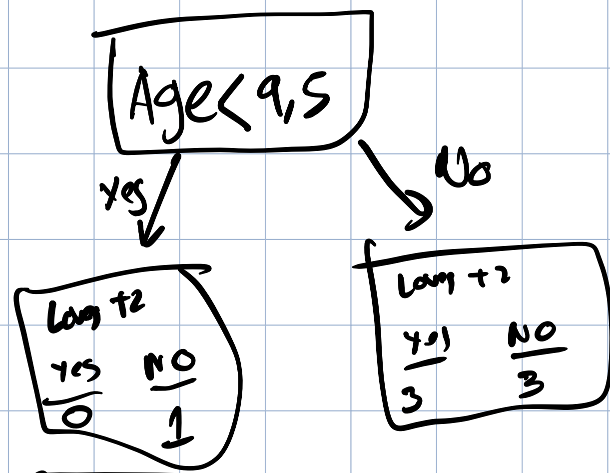
$$C_{G_{\text{Loves popcorn Ponderado}}} = \left(\frac{4}{7}\right) * 0{,}375 + \left(\frac{3}{7}\right) * 0{,}44$$

C_G Loves popo p... $= 0,405$ → Mejor C_G (más puro) → Nodo raíz.

Age	Cous Troll
7	NO
12	NO
18	yes
35	yes
38	yes
50	yes
83	NO

9.5, 15, 26.5, 36.5, 44, 66.5

→ 0,343



$C_{G_p} = 0,429$

Otras formas de cuantificar impureza →

- * Gini → Sencilla
- * Entropía
- * CHAID