

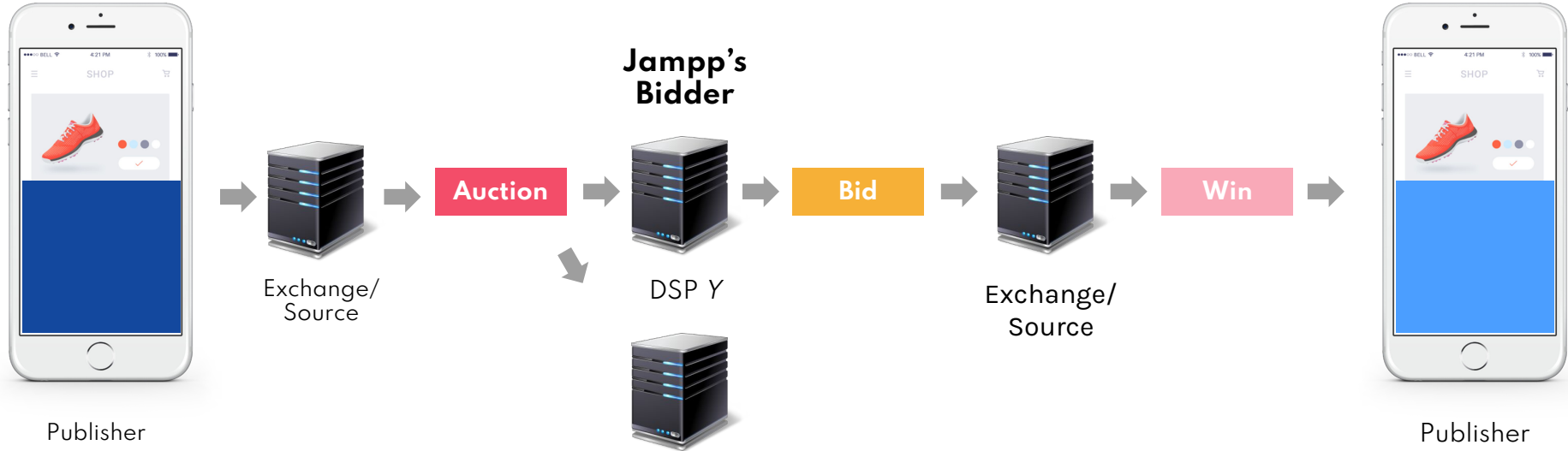


Jampp <> ECI

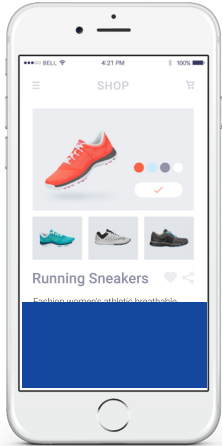
Inferencia Causal

Jorge Brea
Data Scientist

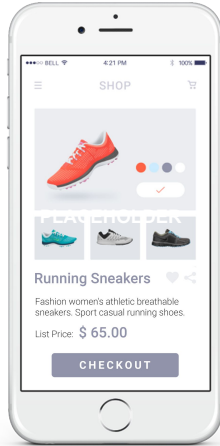
Real Time Bidding (RTB)



Paradoja de Simpson



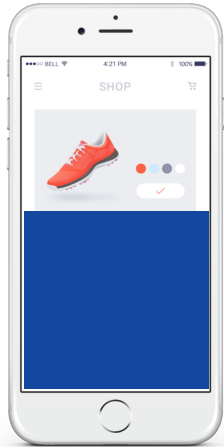
Banner



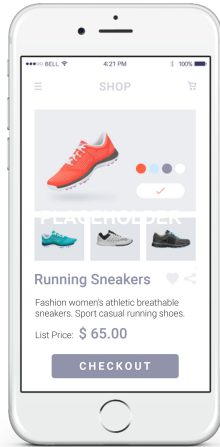
Interstitial

¿Cuál tipo de ad es más 
probable que resulte en un click?

Paradoja de Simpson



Banner



Interstitial

Ad Type	$E[\text{Click} \text{AdType}]$
Interstitial (I)	12% 130/1100
Banner (B)	18% 550/3000

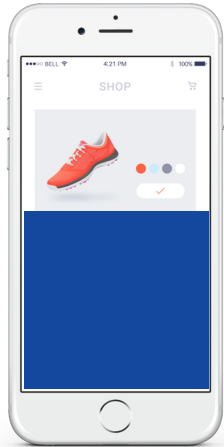
¿Cuál tipo de ad es más
probable que resulte en un click?



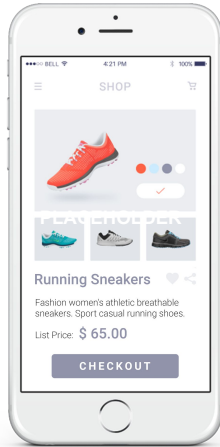
Paradoja de Simpson



By source



Banner



Interstitial

Ad Type	$E[\text{Click} \text{AdType}, \text{Source 1}]$	$E[\text{Click} \text{AdType}, \text{Source 2}]$	$E[\text{Click} \text{AdType}]$
Interstitial (I)	10% 10/1000	30% 30/100	12% 130/1100
Banner (B)	5% 50/1000	25% 500/2000	18% 550/3000

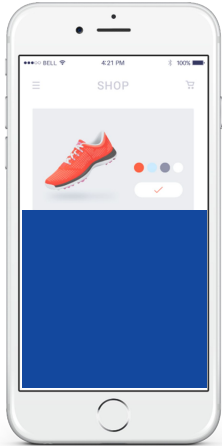
¿Cuál tipo de ad es más
probable que resulte en un click?



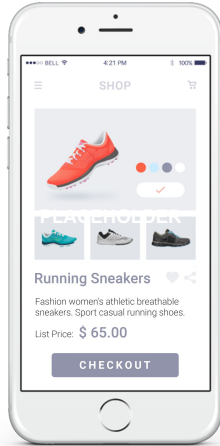
Paradoja de Simpson



By source



Banner



Interstitial

Ad Type	$E[\text{Click} \text{AdType}, \text{Source 1}]$	$E[\text{Click} \text{AdType}, \text{Source 2}]$	$E[\text{Click} \text{AdType}]$
Interstitial (I)	10% 10/1000	30% 30/100	12% 130/1100
Banner (B)	5% 50/1000	25% 500/2000	18% 550/3000

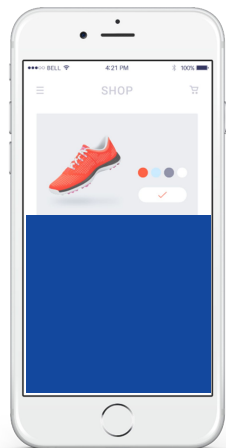


¿Deberíamos entonces usar Interstitial?

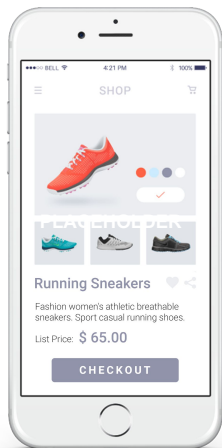
Paradoja de Simpson



By source

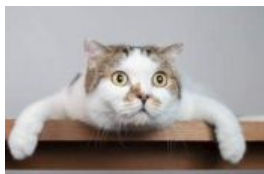


Banner



Interstitial

Ad Type	$E[\text{Click} \text{AdType}, \text{Source 1}]$	$E[\text{Click} \text{AdType}, \text{Source 2}]$	$E[\text{Click} \text{AdType}]$
Interstitial (I)	10% 10/1000	30% 30/100	12% 130/1100
Banner (B)	5% 50/1000	25% 500/2000	18% 550/3000



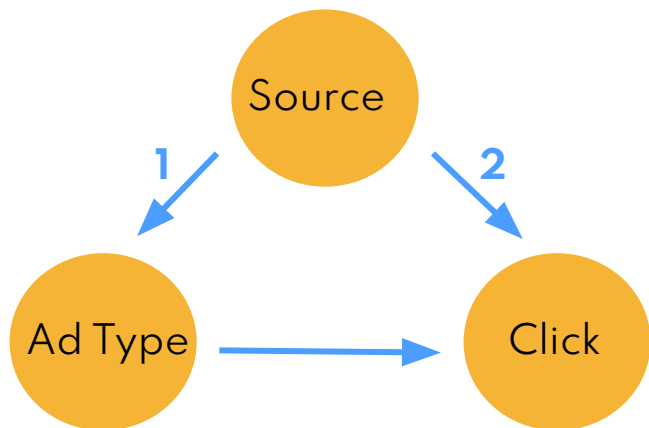
¿Deberíamos entonces usar Interstitial?

¡Depende! 🤔

Escenario 1



Ad presentado depende del source.



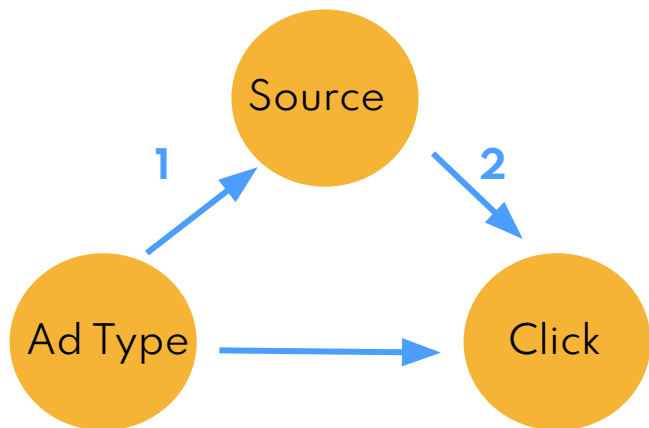
Ad Type	$E[\text{Click} \text{AdType}, \text{Source 1}]$	$E[\text{Click} \text{AdType}, \text{Source 2}]$	$E[\text{Click} \text{AdType}]$
Interstitial (I)	10% 10/1000	30% 30/100	12% 130/1100
Banner (B)	5% 50/1000	25% 500/2000	18% 550/3000

1. Source 1 tiene **más interstitial que banner** ads.
2. Source 1 es **menos confiable**.

Deberíamos condicionar al source para evaluar el efecto del ad en la respuesta de le usuarie.

Escenario 2

Ad type elige el source.



Ad Type	$E[\text{Click} \text{AdType}, \text{Source 1}]$	$E[\text{Click} \text{AdType}, \text{Source 2}]$	$E[\text{Click} \text{AdType}]$
Interstitial (I)	10% 10/1000	30% 30/100	12% 130/1100
Banner (B)	5% 50/1000	25% 500/2000	18% 550/3000

1. Source es El Ad Type elegido.
2. Source 1 es **menos confiable**.

Deberíamos tomar $E[\text{Click}|\text{Ad Type}]$

Correlación no implica causalidad

Ad type elige el source.

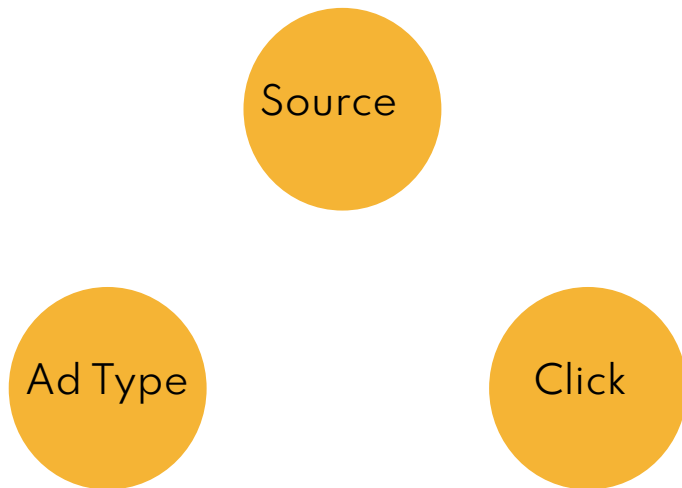


Ad Type

Click

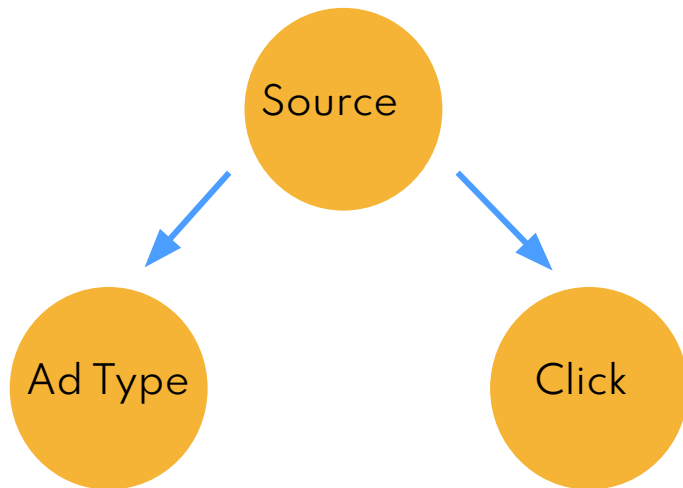
Correlación no implica causalidad

Ad type elige el source.



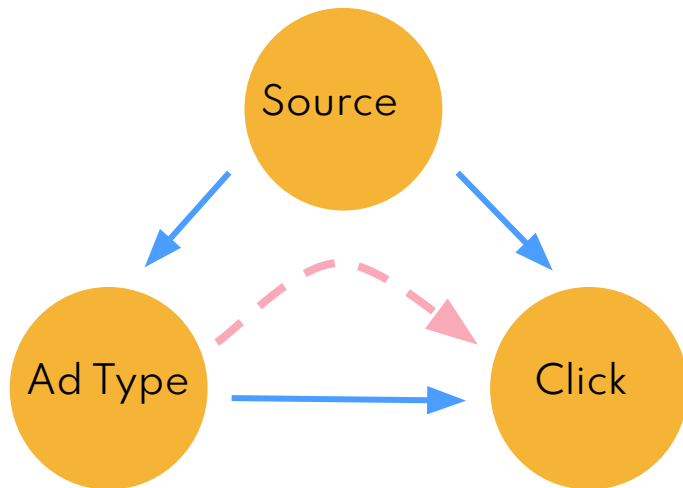
Correlación no implica causalidad

Ad type elige el source.



Correlación no implica causalidad

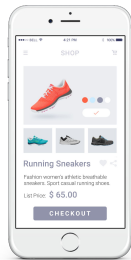
Ad type elige el source.



¿Qué implica causalidad?

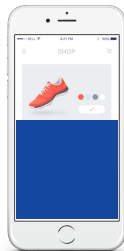


$do(\text{Ad Type} = \text{Interstitial}) \longrightarrow$



$$Y_i |_{\lambda_0(\text{AdType}=I)} \hat{=} Y_i(\text{Interstitial})$$

$do(\text{Ad Type} = \text{Banner}) \longrightarrow$



$$Y_i |_{\lambda_0(\text{AdType}=B)} \hat{=} Y_i(\text{Banner})$$

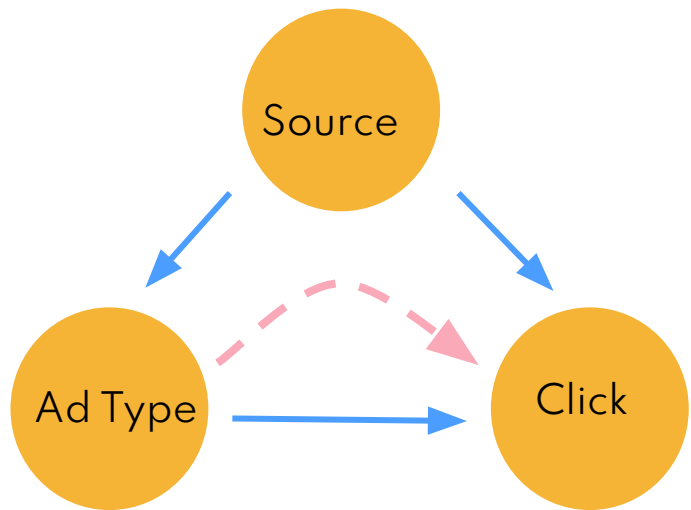
$$Y_i(I) - Y_i(B)$$

El efecto causal de mostrar un ad type....

Average treatment effect (ATE)



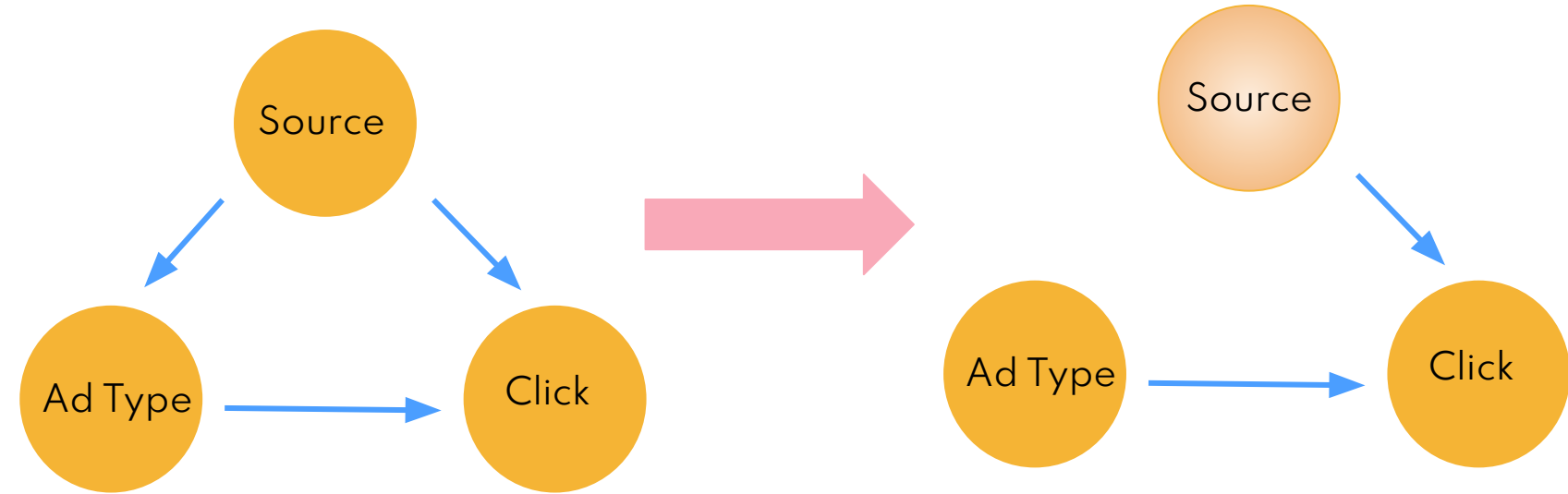
Variable de confusión



Correlación ≠ Causalidad

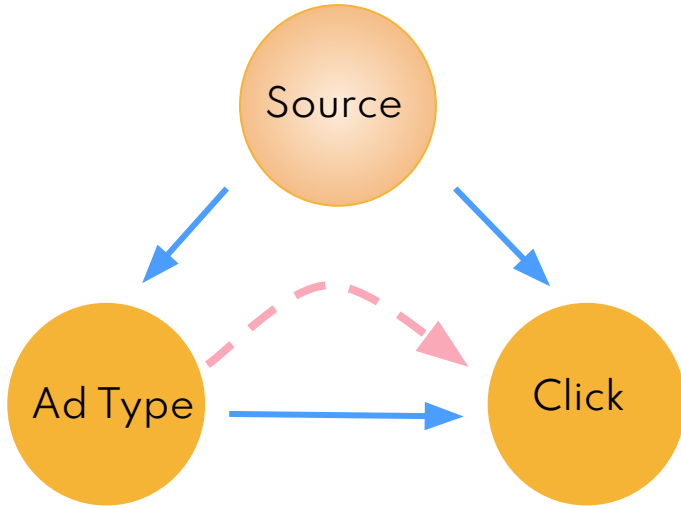
$$E[Y_i(I) - Y_i(B)] = E[Y_i(I)] - E[Y_i(B)] \\ \neq E[Y_i | I] - E[Y_i | B]$$

Randomized Control Trials

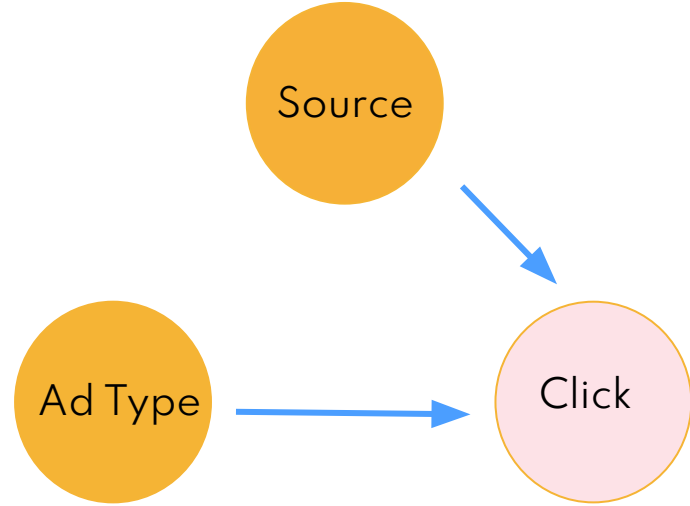


1. Se selecciona el ad type **independientemente** del source.
2. Se repite el 50% para el otro.

Observational Study



Lo que tenemos



Lo que sería ideal

$$E[\text{Click} \mid do(\text{Ad Type}), \text{Source}] = E[\text{Click} \mid \text{Ad Type}, \text{Source}]$$

Volviendo al principio...



Ad Type	$E[\text{Click} \text{AdType}, \text{Source 1}]$	$E[\text{Click} \text{AdType}, \text{Source 2}]$	$E[\text{Click} \text{AdType}]$
Interstitial (I)	10% 10/1000	30% 30/1000	12% 130/1100
Banner (B)	5% 50/1000	25% 500/2000	18% 550/3000

$$E[\text{Click} | do(\text{Ad Type})] = \sum_{\text{Source}} E[\text{Click} | do(\text{Ad Type})] P(\text{Source})$$

Volviendo al principio...



Ad Type	$E[\text{Click} \text{Ad Type, Source 1}]$	$E[\text{Click} \text{Ad Type, Source 2}]$	$E[\text{Click} \text{Ad Type}]$
Interstitial (I)	10% 10/1000	30% 30/1000	12% 130/1100
Banner (B)	5% 50/1000	25% 500/2000	18% 550/3000

$$E[\text{Click} | do(\text{Ad Type})] = \sum_{\text{Source}} E[\text{Click} | do(\text{Ad Type})] P(\text{Source})$$

$$P(S_1) = \frac{20}{31}$$


$$P(S_2) = \frac{11}{31}$$

$$E[\text{Click} | do(I)] = 0.17$$

$$E[\text{Click} | do(B)] = 0.12$$

Volviendo al principio...



Ad Type	$E[\text{Click} \text{AdType}, \text{Source 1}]$	$E[\text{Click} \text{AdType}, \text{Source 2}]$	$E[\text{Click} \text{AdType}]$	Causality
Interstitial (I)	10% 10/1000	30% 30/1000	12% 130/1100	17% 
Banner (B)	5% 50/1000	25% 500/2000	18% 550/3000	12%

$$E[\text{Click} | do(\text{Ad Type})] = \sum_{\text{Source}} E[\text{Click} | do(\text{Ad Type})] P(\text{Source})$$

$$P(S_1) = \frac{20}{31}$$

$$E[\text{Click} | do(I)] = 0.17$$

$$P(S_2) = \frac{11}{31}$$

$$E[\text{Click} | do(B)] = 0.12$$

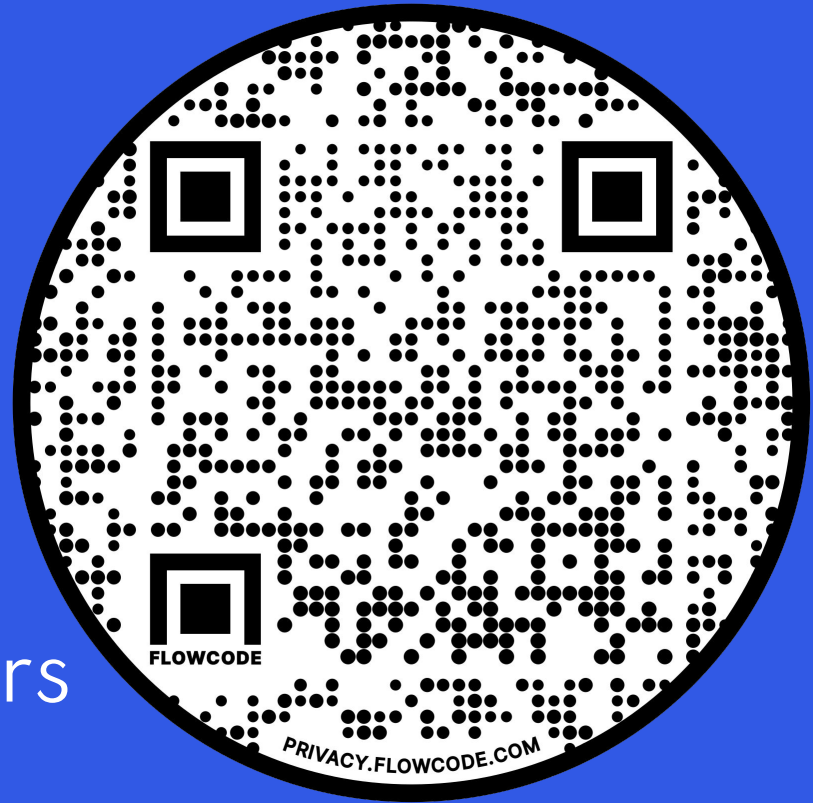
¿Preguntas?

We are hiring!

www.jampp.com/jobs



Data Analysts
ML Engineers
Performance
Engineers
Data Engineers
Cloud Engineers
Full Stack Engineers
....& more!





¡Gracias!