

Análisis del conversion rate

Se entiende como **conversion rate** el ratio de transacciones realizadas por sesión en un periodo de tiempo x. Esta métrica es una métrica muy utilizada en los e-commerce para hacer una valoración del funcionamiento del negocio. Sobre el valor de esta métrica se toman decisiones de diseño, de optimización y a veces de inversión. El objetivo principal es tener una CR tan alto como sea posible y a su vez procurar que no haya caídas.

Cuando se produce una caída de CR puede ser producida por diferentes factores: menos usuarios que acuden a nuestro site, creación de campañas sobre audiencias con baja propensión de compra, aumento del precio, etc..

El caso que nos ocupa se focaliza sobre las audiencias. Se requiere contestar a la siguiente pregunta: ¿qué tipo de impacto tiene el aumento o reducción de un tipo de audiencia sobre el CR?.

En este test adjuntamos un dataset con datos históricos a nivel diario desde enero del 2019. A continuación detallamos los metadatos del mismo:

Variable	Tipo	Descripción
<i>date</i>	String	Fecha en formato YYYYMMDD
<i>channelGrouping</i>	String	Agrupación de tráfico por canal de entrada
<i>userAgeBracket</i>	String	Rango de edad
<i>userType</i>	String	Tipología de usuario por recurrencia: New Visitor / Returning Visitor
<i>sessions</i>	int	Número de sesiones
<i>transactions</i>	int	Número de transacciones

Así pues viendo la cabecera de los datos tenemos la siguiente información:

```
date,channelGrouping,userAgeBracket,userType,sessions,transactions
20190101,(Other),25-34,New Visitor,17,0
20190101,(Other),25-34,Returning Visitor,30,2
20190101,(Other),35-44,New Visitor,11,0
20190101,(Other),35-44,Returning Visitor,71,2
20190101,(Other),45-54,Returning Visitor,11,0
20190101,(Other),55-64,Returning Visitor,10,0
20190101,Direct - Non Paid,18-24,New Visitor,48,0
20190101,Direct - Non Paid,18-24,Returning Visitor,90,2
20190101,Direct - Non Paid,25-34,New Visitor,288,3
20190101,Direct - Non Paid,25-34,Returning Visitor,651,14
20190101,Direct - Non Paid,35-44,New Visitor,310,2
20190101,Direct - Non Paid,35-44,Returning Visitor,881,13
20190101,Direct - Non Paid,45-54,New Visitor,182,2
20190101,Direct - Non Paid,45-54,Returning Visitor,562,5
20190101,Direct - Non Paid,55-64,New Visitor,110,0
```

Esto se lee de la siguiente manera:

fila 1: 20190101,(Other),25-34,New Visitor,17,0

La fila 1 nos indica que en el día 1 de enero del 2019 hubo 17 sesiones y 0 transacciones de gente entre 25-34 años que entraron por el canal Other y era su primera visita al web.

fila 15: 20190101,Direct - Non Paid,55-64,New Visitor,110,0

La fila 15 nos indica que en el día 1 de enero del 2019 hubo 110 sesiones y 0 transacciones de gente entre 55-64 años que entraron por el canal Direct - Non Paid y era su primera visita al web.

Se requiere de **una solución que nos ayude a explicar cómo los cambios de sesiones en función de la audiencia afectan al conversion rate del día**, es decir, poder crear unos insights que de información relevante como: por cada mil sesiones que se pierde del grupo de Returning Visitors entre 45-54 años que entran por newsletter se pierde x décimas de CR o bien por cada mil sesiones recurrentes de 45-54 años se pierden y décimas de CR.

Para facilitar la comprensión de la solución, en el caso de canal, se requiere que se agregue la información por el primer ítem que aparece a la izquierda en la descripción, de manera que consideremos sólo el primer nivel de descripción: (Other), Direct, Email, Social....