Universidad Central de Venezuela

Facultad de Ciencias

Maestría en Modelos Aleatorios

Tópicos Especiales I

**Proyecto**

Se tiene un conjunto de datos de una compañía que realiza ventas por internet. Para ello esta compañía utiliza técnicas publicitarias en línea. Dicha compañía desea incrementar sus beneficios, es decir maximizar sus ventas y minimizar los costos de publicidad. Los datos asociados a la publicidad se encuentran en la hoja de excel anexa, en ella encontraran el registro durante un año de las siguientes variables. Ad.group, MaxCpc, Clicks, Impresiones, CTR, Avg.Cpc, Avg.Posicion, Qscore, Cost.conv, Costo, Ventas, Conv.rate.

* **Ad.group** = grupos de campañas publicitarias. Estos pueden aparecer repetidos.
* **MaxCPC** = representa lo que el dueño de la compañía está dispuesto a pagar por la publicidad.
* **Clicks** = número de veces que un usuario de internet hace click sobre la publicidad.
* **Impresiones** = total de veces que la publicidad fue impresa en internet.
* **CTR** = al cociente entre clicks e impresiones.
* **Avg.Cpc** = lo que termina pagando por la publicación de la publicidad.
* **Avg.Posicion** = Posición promedio en la que sale el anuncio publicitario (primera página de búsqueda, al principio de la pagina, al final de la pagina, etc.).
* **Costo** = el gasto de publicidad durante el año.
* **Ventas** = total de ventas, esto es cuando luego de hacer click en la publicidad el usuario decide comprar el articulo.
* **Conv.rate** = es la tasa de ventas, es decir el cociente de ventas entre el total de clicks.

Con la finalidad de responder a las inquietudes del cliente, deben realizar lo siguiente:

* Aplicar las técnicas de estadística descriptiva que considere necesarias para caracterizar los datos y dar una primera caracterización de los mismos. ¿existe correlación entre las variables?,
* Estudiar la distribución de cada una de las variables
* Clasifique, si es posible,  los grupos. Aplique las técnicas vistas en el curso (Arboles, cluster, bosques, etc.) o cualquier otra que considere apropiada.
* Estudiar el comportamiento del costo frente a las otras variables.
* Estudiar los casos en los cuales las ventas son mayores a cero e iguales a cero y combine con los casos en los cuales el  costo=0 y  costo>0 ¿existe algún patrón de comportamiento que permita caracterizar estos grupos? ¿Qué haría usted para que aquellos grupos que no generan ventas se comporten como los que generan ventas? ¿qué haría usted para que los grupos que generan costos y no generan ventas empiecen a generar ventas?

Siéntase en libertad de aplicar cualquier otra técnica, estudio o análisis exploratorio que considere pertinente.

Deben entregar un informe donde queden clara las técnicas utilizadas y los resultados obtenidos. Deben entregar el script debidamente identificado y donde quede claro que se hace en cada paso. Todos los gráficos deben llevar título y epígrafe que explique bien lo que representan.

**Fecha de entrega:** Viernes 19 de mayo de 2017.