
Trabajo Práctico 1

Black Friday

Un estudio de ventas a través del comportamiento del consumidor

Métodos Estadísticos Aplicados a los Negocios

Fecha de entrega: 15 de abril de 2019

1. INSTRUCCIONES

El trabajo puede ser grupal (con un máximo de 4 alumnos). Deberán entregar un informe final **a través del campus virtual** en el cual adjuntarán el Script en R para poder replicar sus resultados.

Tienen tiempo para entregar el TP hasta el lunes 15 de abril a las 23.55hs. Una entrega tardía implica desaprobación del trabajo práctico.

2. CONSIGNA

Utilice la base de datos `BlackFriday.csv`¹ que contiene una muestra de más de 500.000 transacciones realizadas en una tienda de retail en un Black Friday. La tienda quiere tener un mejor conocimiento de la compra de los clientes frente a diferentes productos. La base de datos incluye las siguientes variables:

¹Los datos provienen de una competencia organizada por Analytics Vidhya.

- Use_ID: código del cliente
- Product_ID: código del producto comprado
- Gender: M (Male), F (Female)
- Age: edad del cliente
- Occupation: código de la ocupación del cliente
- City_Category: código de la ciudad (A, B, C)
- Stay_In_Current_City_Years: cantidad de años que el cliente vive en dicha ciudad
- Marital Status: 1 (casado/a), el resto de los códigos indican otros estados civiles.
- Product_Category_1
- Product_Category_2
- Product_Category_3
- Purchase: monto de la compra en dólares

Se pide:

- (1) **(5 puntos)** Describa la base de datos y obtenga los principales estadísticos descriptivos. Comente.
- (2) **(5 puntos)** Construya un intervalo de confianza de la proporción de mujeres que realizan compras en el Black Friday.
- (3) **(10 puntos)** Estime intervalos de confianza del monto medio gastado por un individuo en el Black Friday, pero segmentando por edad. Intente expresarlo gráficamente. Comente. ¿Existen diferencias en el monto gastado de acuerdo a la edad?
- (4) **(10 puntos)** ¿Cuál es el producto más comprado? (Identifique a partir del ID). Pruebe si la proporción de mujeres que compran dicho producto es igual al 27%. Grafique la curva de potencia del test. Comente.
- (5) **(10 puntos)** ¿Hay evidencia suficiente para probar que un individuo gasta más de 9300 dólares en promedio? ¿Y más de 9400 dólares? Construya el estadístico y el p -valor en R. Concluya.
- (6) **(10 puntos)** El dueño de la tienda afirma que las mujeres gastan más dinero que los hombres. ¿Encuentra evidencia empírica que valide dicha afirmación?
- (7) **(10 puntos)** ¿Hay evidencia suficiente respecto de que las personas casadas gastan más dinero en el Black Friday? Justifique.
- (8) **(10 puntos)** Pruebe si existen diferencias significativas en el monto gastado entre las tres ciudades (A, B y C).
- (9) **(10 puntos)** ¿Hay evidencia suficiente respecto de que las mujeres casadas que tienen

entre 26-35 años gastan más dinero en un Black Friday que los hombres casados del mismo grupo etario?

- (10) **¡Vamos por el 10! (20 puntos)** Usando los datos de esta muestra diseñe una prueba para evaluar si los siguientes estimadores de la proporción poblacional (p) son insesgados, eficientes y consistentes. Justifique. (*Hint*: extraiga al azar muestras pequeñas y grandes a partir de alguna variable de la base de datos).

$$\text{Estimador 1: } \hat{p} = \frac{X}{n}$$

$$\text{Estimador 2: } \tilde{p} = \frac{X}{n-c}, \quad 0 < c < n$$

donde X denota el total de casos exitosos.