

PROBLEMA FINAL

Programación en R

Analizamos datos de marketing en R

Mediante el lenguaje de programación R, vamos a desarrollar un análisis de datos sobre el histórico de ventas de una compañía y su inversión publicitaria para distintos medios (TV, SEM y redes sociales).

Objetivos

La finalidad de esta actividad es poner en práctica todo el conocimiento que has ido desarrollando a lo largo de la asignatura a través de un caso de uso 100% real.

Descripción de la actividad

Para esta actividad, vas a trabajar con una serie de ficheros:

- **PF_sales.csv**: contiene las ventas mensuales de la compañía entre el 01/11/2015 y el 01/11/2019.
- **PF_tv.csv**: contiene la inversión mensual de la compañía en publicidad televisiva entre el 01/01/2016 y el 01/11/2019.
- **PF_sem.csv**: contiene a nivel semanal tanto la inversión de la compañía en SEM como los clics generados entre el 23/11/2015 y el 24/06/2019.
- **PF_rrss.csv**: contiene a nivel semanal tanto la inversión de la compañía en redes sociales como las impresiones conseguidas entre el 23/11/2015 y el 24/06/2019.

La resolución de la actividad pasa por generar un proyecto de RStudio y su correspondiente código R para, a través de él, llevar a cabo el análisis de los datos anteriores.

Tendrás que respetar los buenos hábitos de programación vistos durante la asignatura (la estructura de directorio, la estructura de código a través de funciones, el nombrado de variables...) y utilizar el control de versiones Git.

El desarrollo de la actividad se va a dividir en tres partes:

1. Adquisición y preparación del dato.
2. Desarrollo de análisis.
3. Generación de reporting R Markdown.

Veamos el detalle de cada una:

1. Adquisición y preparación del dato

Debes desarrollar una o más funciones que:

- Realicen la lectura de las 4 fuentes de datos asignando para cada una de ellas un dataframe.
- Imputen la media en los NAs de la variable inversión en TV.
- Unan los cuatro dataframes en un mismo dataframe, usando granularidad mensual y quedándose únicamente con el periodo de fechas coincidente.

2. Desarrollo de análisis

Debes desarrollar al menos tres funciones (una por cada elemento de la siguiente lista) que permitan:

- Construir una tabla que muestre la inversión de cada medio en cada año (descartando 2015 y 2019, dado que no están completos).
- Construir una tabla que muestre la venta media de cada mes, descartando también 2015 y 2019. Es decir, deberíamos ser capaces de ver cuál es la venta media en enero, febrero, marzo, etc. para entender qué estacionalidad tiene la serie.
- Construir una tabla que muestre la correlación de cada variable de medios con la venta para entender cuál de ellas está teniendo mayor impacto.

3. Generación de reporting R Markdown

Debes construir un documento R Markdown (exportable a HTML) a modo de reporte que incluya el código R generado, así como sus resultados (tanto la adquisición y preparación del dato como el desarrollo de análisis).

Para ello, ten en cuenta que:

- Se ha de explicar el racional de los pasos seguidos para facilitar la comprensión del documento, así como reforzar o destacar las conclusiones observadas en el dato.
- Los chunks de código R se pueden mostrar u ocultar en la salida en función de tu interés.
- Debes utilizar elementos que mejoren la estética del documento (encabezados, listados, índice de contenido, temas CSS...).

Formato de entrega

Debes comprimir en un ZIP la carpeta del proyecto y subirlo al campus. Dicha carpeta debe contener todo lo necesario para poder replicar tu ejecución.

Criterios de calificación

1. Adquisición y preparación del dato: 4 puntos.
2. Desarrollo de análisis: 3 puntos.
3. Generación de reporting R Markdown: 3 puntos.

Nota al alumnado: *aquellos que consigáis resolver la mayoría de los ejercicios de forma correcta podréis optar al sobresaliente. No obstante, podréis superar la asignatura si conseguís una resolución aproximada, es decir, las resoluciones parciales son también válidas. Y recordad: en el último webinar, el profesor os orientará para la realización del ejercicio y resolverá vuestras dudas.*