

Cálculo de métricas

DATASET: LogsDataiku.csv

Este *dataset* almacena información de acceso al sitio web de Dataiku, durante 2 meses. Contiene información sobre cada vista de página individual en el sitio web www.Dataiku.com

Estructura del dataset

El *dataset*, está en un fichero csv con 18 campos:

- *server_ts*: fecha de conexión al servidor
- *client_ts*: fecha de conexión del usuario
- *client_addr*: dirección IP del usuario
- *visitor_id*: id asignado al usuario (~_utma)
- *session_id*: id de la sesión/visita (~_utmb)
- *location*: url de la página a la que se accede
- *referer*: url de la página desde la que se accedió a location. Si está vacío se considera una búsqueda directa
- *user_agent*: navegador desde el que se conectó el usuario
- *type*: tipo de acceso
- *visitor_param*
- *session_param*:
- *event_param*:
- *br_width*: ancho del navegador del usuario
- *br_height*: alto del navegador del usuario
- *sc_width*: ancho de la pantalla del usuario (resolución)
- *sc_height*: alto de la pantalla del usuario (resolución)
- *br_lang*: idioma del navegador del usuario
- *tz_off*: time zona. El número indica la diferencia en minutos con la hora GMT . -60 indica zona horaria GMT+1 y 60 indica zona horaria GMT-1

Ejercicio

A partir del *dataset* anterior, calcula las siguientes métricas para el periodo de tiempo registrado

- N.º de visitas → Número de visitas (sesiones que ha tenido el sitio web)
- N.º de visitantes únicos → Número de usuarios diferentes que han visitado el sitio web
- N.º medio de páginas/visitas → Para cada visita (sesión) cuántas páginas se han visitado. Media para todas las visitas
- Tasa de rebote → Número de visitas (sesiones) que solo han accedido a una página
- Tasa de salida para cada página → % de veces que cada página ha sido una página de salida
- Tráfico directo → Número de visitas (sesiones) que provienen de escribir la url directamente
- Tráfico de búsqueda → Número de visitas (sesiones) que provienen de pinchar en una búsqueda
- Tráfico referido → Número de visitas (sesiones) que provienen de pinchar en un enlace de otra página

Qué hay que entregar

1. Documento en el que se explique cómo se han calculado las diferentes métricas y los resultados obtenidos para cada una de las métricas
2. En caso de realizar una implementación en algún lenguaje, código para probarlo.