WEB_LOGS 🔍

CUSTOMER_ID
ID of identified customer
DATE_KEY
Date of web event (YYYYMMDD format)
TIME
Time of web event (HH24MMSS format)
URL
URL of website visited by customer
REFERRER_SUBDOMAIN / _DOMAIN / _TLD
URL address of referring website
DEVICE_NAME
Event source device (Asus ME172V, LG-D390n…)
BROWSER_NAME
Used web browser
OS_NAME
Used operating system

CUSTOMER_BASE 👥

CUSTOMER_ID
Unique ID of customer
PRODUCT_ID
Unique ID of product (1 customer can have more products)
CUSTOMER_TYPE
Customer segment (PRIVATE or BUSINESS)
PRODUCT_GROUP
Product category (TV, Easy, Happy...)
PRODUCT_NAME
Detailed name of product (TV L, Easy Pecka, Happy XL…)
DEVICE_NAME
Name of used device on specific product (Nokia 3210…)
AGE
Age of customer (available only for PRIVATE segment)
GENDER
Gender of customer (available for PRIVATE segment)

VOICE_TRAFFIC 📞

PRODUCT_ID
ID of product on which the voice traffic is tracked
DIRECTION_TYPE
Direction of call from customer/product view (IN, OUT)
OTHER_PARTY_PRODUCT_ID
ID of product FROM/TO which the voice call is routed
DATE_KEY
Date of voice call (YYYYMMDD format)
TIME
Time of voice call (HH24MMSS format)
DURATION
Duration of voice call (in seconds)
OTHER_PARTY_COUNTRY
Country of origin of the other party product (SK, CZ, UK…)
GEO_LOCATION
GPS coordinates of BTS where the voice call is carried

IPTV_TRAFFIC 🖥️

PRODUCT_ID
ID of product on which the TV event is generated
DATE_KEY
Date of TV watching session (YYYYMMDD format)
TIME_KEY
Start time of TV watching session (HH24MMSS format)
DURATION
Duration of TV watching session (in seconds)
TV_SESSION_TYPE
Type of session (Life TV / Archive / Movie rental)
TV_SESSION_NAME
Name of TV channel / Show from TV archive / Movie rent

DATA_TRAFFIC 🌐

PRODUCT_ID
ID of product on which the data event is generated
DATE_KEY
Date of data event (YYYYMMDD format)
TIME_KEY
Start time of data event (HH24MMSS format)
DURATION
Duration of data event (in seconds)
NETWORK_TECHNOLOGY_TYPE
Technology of network used for data event (4G/3G/2G)
SESSION_END_TYPE
Reason for closing data session (TimeLimit, VolumeLimit…)
UPLOAD / DOWNLOAD_DATA_VOLUME
Volume of data uploaded/downloaded in bytes
GEO_LOCATION
GPS coordinates of BTS (Base Transceiver Station)

SMS_TRAFFIC ✉️

PRODUCT_ID
ID of product on which the SMS traffic is tracked
DIRECTION_TYPE
Direction of SMS from customer/product view (IN, OUT)
OTHER_PARTY_PRODUCT_ID
ID of product FROM/TO which the SMS is routed
DATE_KEY
Date of SMS (YYYYMMDD format)
TIME
Time of SMS (HH24MMSS format)
LENGTH
Length of SMS (in number of chars)
OTHER_PARTY_COUNTRY
Country of origin of the other party product (SK, CZ, UK...)

Big Data Hackathon
Data structure cheatsheet