Caso práctico Big Data: Almacenamiento

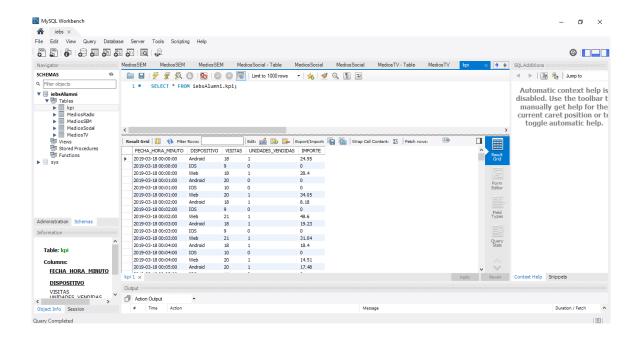
TIPOS DE DATOS:

- DATETIME: Campo que contiene información de la fecha, hora, minuto
- DATE: Campo que contiene la fecha
- YEAR: Campo que contiene el año
- TIME: Campo que contiene la hora, minuto, segundo
- VARCHAR: Campo que almacena datos alfanuméricos de longitud variable
- CHARACTER: Campo que almacena datos alfanuméricos de longitud fija
- INT: Campo que almacena números enteros
- FLOAT: Campo que almacena números con su parte decimal

FUENTES DE INFORMACIÓN:

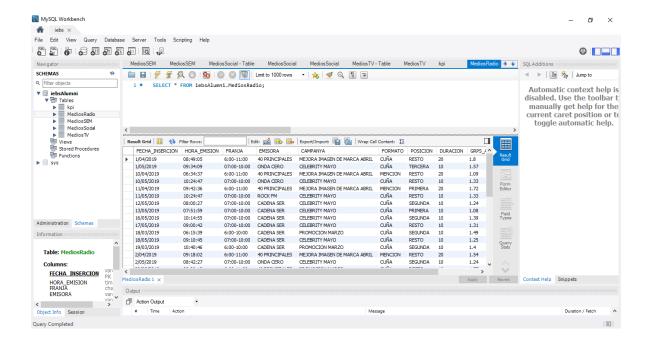
Las fuentes de datos siguientes son clasificadas como small data por la cantidad de datos almacenados y son de tipo relacionales, serán almacenadas en Cloud SQL, dado que es compatible con MySQL. Este sistema de almacenamiento presta las herramientas necesarias para el tipo de dato usado en las tablas.

TABLA	IEBS_KPIS	
VARIABLES	Tipo	Descripcion
FECHA_HORA_MINUTO	DATETIME	Campo que contiene la fecha, hora y minuto
DISPOSITIVO	VARCHAR	Dispositivo por el que se produce la visita, unidad vendida e importe
VISITAS	INT	Nº de visitas por unidad de minuto y dispositivo
UNIDADES_VENDIDAS	INT	Nº de unidades vendidas por unidad de minuto y dispositivo
IMPORTE	FLOAT	Suma del importe por unidad de minuto y dispositivo



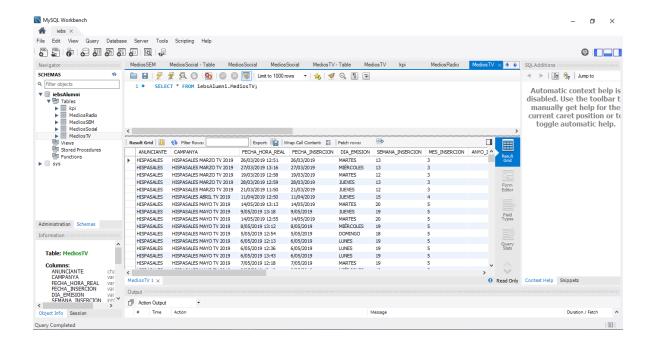
FICHERO IEBS008_PLANMEDIOSRD

VARIABLES	Tipo	Descripcion
FECHA_INSERCION	VARCHAR	Fecha en la que produce el anuncio en radio
HORA_EMISION	TIME	Hora del anuncio
FRANJA	CHARACTER	Franja del día en la que se produce
EMISORA	VARCHAR	Emisora de radio en la que emite
CAMPANYA	VARCHAR	Nombre de la campaña
FORMATO	VARCHAR	Formato del anuncio
POSICION	VARCHAR	Posición dentro del bloque de anuncios
DURACION	INT	Duración del anuncio
GRPS_AD16	FLOAT	GRPs para el target adultos mayores de 16
INVERSION	FLOAT	Coste del anuncio



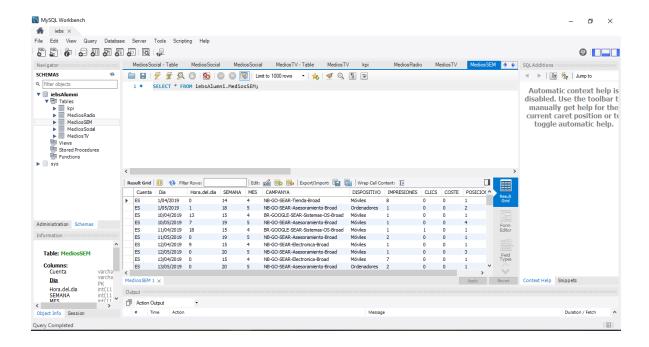
FICHERO IEBS008 PLANMEDIOSTV

VARIABLES	Tipo	Descripcion
ANUNCIANTE	CHARACTER	Nombre del anunciante
CAMPANYA	VARCHAR	Nombre de la campaña
FECHA_HORA_REAL	VARCHAR	Día, hora y minuto en la que se produce el anuncio
FECHA_INSERCION	VARCHAR	Día en el que se produce el spot
DIA_EMISION	VARCHAR	Literal del día de la semana en el se produce el anuncio
SEMANA_INSERCION	INT	Nº de semana en el se produce el anuncio
MES_INSERCION	INT	Mes en el se produce el anuncio
ANYO_INSERCION	YEAR	Año en el se produce el anuncio
SOPORTE_ID	INT	ID de la cadena de TV en la se produce el anuncio
FRANJA_ID	INT	ID de la franja horaria en la se produce el anuncio
PT_DT	CHARACTER	Horario de prime time (PT) o day time (DT)
BLOQUE	VARCHAR	Bloque en el que se emite el anuncio
CREATIVIDAD	VARCHAR	Creatividad asociada al anuncio
DURACION	INT	Duración en segundos del anuncio
POSICION_LIT	VARCHAR	Posición del anuncio en el bloque publicitario
POSICION_NUM	INT	Posición numérica del anuncio en el bloque publicitario
SPOTS_BLOQUE	INT	Nº de spots del bloque de anuncios
PROGRAMA	VARCHAR	Nombre del programa donde se emite el anuncio
GRPS_AD16	FLOAT	GRPs para el target de adultos mayores de 16 años
GRPS20_AD16	FLOAT	GRPs a 20" para el target de adultos mayores de 16 años
INVERSION	FLOAT	Coste del anuncio



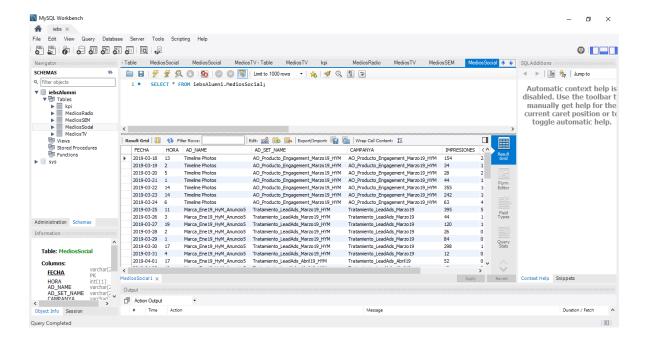
FICHERO IEBS008 PLANMEDIOSSEM

VARIABLES	Tipo	Descripcion
CUENTA	VARCHAR	Contiene el nombre de la cuenta
DÍA	VARCHAR	Fecha con el día
HORA DEL DÍA	INT	Hora del día
SEMANA	INT	Nº de semana del año
MES	INT	Mes del año
CAMPANYA	VARCHAR	Nombre de la campaña
DISPOSITIVO	VARCHAR	Dispositivo en el que se sirven las impresiones
IMPRESIONES	INT	Nº de impresiones
CLICS	INT	Nº de clics
COSTE	FLOAT	Coste por dia, hora, campaña y dispositivo
POSICION_MEDIA	FLOAT	Posición media de las impresiones servidas



FICHERO IEBS008_PLANMEDIOSSOCIAL

VARIABLES	Tipo	Descripcion
FECHA	VARCHAR	Fecha con el día
HORA	INT	Hora del día
AD_NAME	VARCHAR	Campo asociado a la campaña
AD_SET_NAME	VARCHAR	Campo asociado a la campaña
CAMPANYA	VARCHAR	Nombre de la campaña
IMPRESIONES	INT	Impresiones servidas
CLICS	INT	Clics conseguidos
LEADS	INT	Leads conseguidos
INVERSION	FLOAT	Coste por dia, hora, campaña y dispositivo



FUENTES DE DATOS:

Estas fuentes de datos tienen una cantidad de información superior a las mencionadas anteriormente, son consideradas big data y aunque puedan ser almacenadas de forma relacional requieren de una capacidad de almacenamiento superior, por lo tanto, se almacenarán en BigQuery.

