---Para obtener los contactos que no cuentan con correo electronico, se hará la siguiente consulta. Una vez hecha la consulta, lo que haremos es eliminarlos de la base para evitar un sesgo.

SELECT

EMPRESA,

NOMBRE1,

ALCALDIA\_MUNICIPIO,

ESTADO,

ETIQUETA,

REVISTA\_IMPRESA,

GIROPRINCIPAL,

CORREO

FROM `proyecto-final-db-381602.BasePublicidad.contactos`

WHERE CORREO IS NULL

---Ademas queremos asegurarnos de que de los contactos que no tienen correo, existe la posibilidad de que se les haga llegar la revista impresa.

AND REVISTA\_IMPRESA LIKE "%IMPRESA%";

--- Para los fines de este estudio, necesitamos eliminar los contactos sin correo, el motivo es lo mencionado anteriormente

DELETE `proyecto-final-db-381602.BasePublicidad.contactos`

WHERE CORREO IS NULL;

--- Una vez eliminados los contactos sin correo, tenemos que saber también que contactos se encuentran duplicados y elimnarlos también.

SELECT

  COUNT(CORREO),

  CORREO

FROM

  proyecto-final-db-381602.BasePublicidad.contactos

  GROUP BY CORREO

HAVING COUNT (CORREO)>1;

---Ahora, para poder eliminar estos contactos y quedarnos con nuestros valores unicos, tenemos que asignar primero un valor de ID y que sea unico.

with contactosfinal as (SELECT

EMPRESA,

TITULO,

NOMBRE1,

NOMBRE2,

APELLIDO1,

APELLIDO2,

PUESTO,

CALLE\_Y\_NUMERO

INTERIOR,

COLONIA,

CP,

ALCALDIA\_MUNICIPIO,

ESTADO,

CORREO,

ETIQUETA,

REVISTA\_IMPRESA,

GIROPRINCIPAL,

FECHACREACION,

ROW\_NUMBER() OVER (PARTITION BY CORREO) AS ID

FROM

  proyecto-final-db-381602.BasePublicidad.contactos)

  SELECT \*

FROM contactosfinal

WHERE ID = 1;

---Creamos una nueva tabla con los datos que obtuvimos en la consukta anterior

CREATE TABLE CONTACTOSFINAL (

  EMPRESA STRING (100),

  NOMBRE1 STRING (100),

  NOMBRE2 STRING (100),

  APELLIDO1 STRING (100),

  APELLIDO2 STRING (100),

  PUESTO STRING (100),

  CALLE\_Y\_NUMERO STRING (100),

  INTERIOR STRING (100),

  COLONIA STRING (100),

  CP STRING (100),

  ALCALDIA\_MUNICIPIO STRING (100),

  ESTADO STRING (100),

  CORREO STRING (100),

  ETIQUETA STRING (100),

  REVISTA\_IMPRESA STRING (100),

  GIROPRINCIPAL STRING (100),

  FECHACREACION DATETIME);

ALTER TABLE Customers

DROP COLUMN Email;

INSERT INTO proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL(EMPRESA,

  TITULO,

  NOMBRE1,

  NOMBRE2,

  APELLIDO1,

  APELLIDO2,

  PUESTO,

  CALLE\_Y\_NUMERO,

  INTERIOR,

  COLONIA,

  CP,

  ALCALDIA\_MUNICIPIO,

  ESTADO,

  CORREO,

  ETIQUETA,

  REVISTA\_IMPRESA,

  GIROPRINCIPAL,

  FECHACREACION)

with contactosfinal2 as (SELECT

EMPRESA,

TITULO,

NOMBRE1,

NOMBRE2,

APELLIDO1,

APELLIDO2,

PUESTO,

CALLE\_Y\_NUMERO,

INTERIOR,

COLONIA,

CP,

ALCALDIA\_MUNICIPIO,

ESTADO,

CORREO,

ETIQUETA,

REVISTA\_IMPRESA,

GIROPRINCIPAL,

FECHACREACION,

ROW\_NUMBER() OVER (PARTITION BY CORREO) AS ID

FROM

  proyecto-final-db-381602.BasePublicidad.contactos)

SELECT EMPRESA,

TITULO,

NOMBRE1,

NOMBRE2,

APELLIDO1,

APELLIDO2,

PUESTO,

CALLE\_Y\_NUMERO,

INTERIOR,

COLONIA,

CP,

ALCALDIA\_MUNICIPIO,

ESTADO,

CORREO,

ETIQUETA,

REVISTA\_IMPRESA,

GIROPRINCIPAL,

FROM contactosfinal2

WHERE ID = 1;

---Primero veremos cuales son los valores unicos para saber cuales son los que se tienen que modificar para unificar la base

SELECT

DISTINCT ESTADO

FROM proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL;

---Ahora procederemos a unificar datos y eliminar datos de Estados que no sean de México

UPDATE proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL

SET ESTADO = 'Guanajuato'

WHERE ESTADO = ' Guanajuato';

UPDATE proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL

SET ESTADO = 'Baja California'

WHERE ESTADO = 'Baja Califnornia';

UPDATE proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL

SET ESTADO = 'Coahuila'

WHERE ESTADO = 'Coahuila ';

UPDATE proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL

SET ESTADO = 'Edo. de Mexico'

WHERE ESTADO = 'Edo. De México';

UPDATE proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL

SET ESTADO = 'Edo. de Mexico'

WHERE ESTADO = 'EDOMEX';

UPDATE proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL

SET ESTADO = 'Edo. de Mexico'

WHERE ESTADO = 'Estado de México';

UPDATE proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL

SET ESTADO = 'Edo. de Mexico'

WHERE ESTADO = 'México';

---Ahora eliminaremos los datos que no nos sirven

DELETE `proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL`

WHERE ESTADO = ' Turkey'

DELETE `proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL`

WHERE ESTADO = 'Arizona'

DELETE `proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL`

WHERE ESTADO = 'California'

DELETE `proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL`

WHERE ESTADO = 'Canada'

DELETE `proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL`

WHERE ESTADO = 'China'

DELETE `proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL`

WHERE ESTADO = 'España'

DELETE `proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL`

WHERE ESTADO = 'EUA'

DELETE `proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL`

WHERE ESTADO = 'FL'

DELETE `proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL`

WHERE ESTADO = 'Francia'

DELETE `proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL`

WHERE ESTADO = 'La Aurora'

DELETE `proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL`

WHERE ESTADO = 'MA'

DELETE `proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL`

WHERE ESTADO = 'Miami'

DELETE `proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL`

WHERE ESTADO = 'OH'

DELETE `proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL`

WHERE ESTADO = 'Oregon'

DELETE `proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL`

WHERE ESTADO = 'Perú'

DELETE `proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL`

WHERE ESTADO = 'Texas';

SELECT

DISTINCT ESTADO

FROM proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL;

DELETE `proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL`

WHERE ESTADO = 'ESTADO';

---Tampoco nos funcionan los que son Nulos

DELETE `proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL`

WHERE ESTADO IS NULL;

---Para comprobar que si tenemos unicamente 32 valores unicos (32 por los estados de la Republica Mexicana)

SELECT

DISTINCT ESTADO

FROM proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL;

---Vemos nuestra tabla para saber cuantos valores nos quedaron despues de limpiarla

SELECT \*

FROM proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL;

---Procedemos a mostrar cuanta audiencia se tiene por Estado de la Republica Méxicana.

SELECT

ESTADO,

  COUNT (ESTADO) AS VECESREPETIDAS

FROM proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL

GROUP BY ESTADO

ORDER BY VECESREPETIDAS ASC;

---Ahora contaremos cuantos contactos tienen asignada una etiqueta y cuales son las etiquetas

SELECT

DISTINCT ETIQUETA

FROM proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL;

UPDATE proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL

SET ETIQUETA = 'E&C'

WHERE ETIQUETA = 'Energia electrica ';

UPDATE proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL

SET ETIQUETA = 'NINGUNA'

WHERE ETIQUETA IS NULL;

---Procedemos a hacer el conteo de cuantos contactos tenemos por ETIQUETA

SELECT

ETIQUETA,

  COUNT (ETIQUETA) AS CONTACTOSASIGNADOS

FROM proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL

GROUP BY ETIQUETA

ORDER BY CONTACTOSASIGNADOS ASC;

---Ahora, queremos saber cuales son los sectores que se encuentran en nuestra base

SELECT

DISTINCT GIROPRINCIPAL

FROM proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL

ORDER BY GIROPRINCIPAL ASC;

---Tenemos que volver a unificar los datos porque algunos se encuentran mal escritos (Energía Renovable-Energías Renovables, Oil&Gas-Oil & Gas, Tansporte- Transporte y Logística)

UPDATE proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL

SET GIROPRINCIPAL = 'Energía Renovable'

WHERE GIROPRINCIPAL = 'Energías Renovables';

UPDATE proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL

SET GIROPRINCIPAL = 'Oil & Gas'

WHERE GIROPRINCIPAL = 'Oil&Gas';

UPDATE proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL

SET GIROPRINCIPAL = 'Transporte y Logística'

WHERE GIROPRINCIPAL = 'Transporte';

UPDATE proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL

SET GIROPRINCIPAL = 'SIN ASIGNACION'

WHERE GIROPRINCIPAL IS NULL;

SELECT

DISTINCT GIROPRINCIPAL

FROM proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL

ORDER BY GIROPRINCIPAL ASC;

---Procedemos a hacer el conteo de cuantos contactos tenemos por Giro Principal

SELECT

GIROPRINCIPAL,

  COUNT (GIROPRINCIPAL) AS TOTALCONTACTOS

FROM proyecto-final-db-381602.BasePublicidad.CONTACTOSFINAL

GROUP BY GIROPRINCIPAL

ORDER BY TOTALCONTACTOS ASC;