MIP Data Repository Queries

Dashboard - Queries

The Media Impact Project Dashboard has a number of weekly scheduled automated background tasks that run queries on the sessionized data within Google Big Query.

Prepare Queries

- **Donors**
  - **Query:**
    - **SELECT date, IF(database_donors IS NULL AND ea = 'Intent' AND ec = 'Support', cid, NULL) AS donor_cid FROM (SELECT a.donor_cid AS database_donors, TIMESTAMP(b.year + '-' + b.month + '-' + b.day) AS date, b.ec AS ec, b.ea AS ea, b.cid AS cid FROM [116430105.donors] a FULL OUTER JOIN EACH (SELECT hits.eventInfo.eventCategory AS ec, hits.eventInfo.eventAction AS ea, fullVisitorId AS cid, /*Date*/ REGEXP_EXTRACT(date, '(^[0-9]{{4}})') as year, REGEXP_EXTRACT(date, '^[0-9]{{4}}([0-9]{{2}})') as month, REGEXP_EXTRACT(date, '([0-9]{{2}}$)') as day FROM (TABLE_DATE_RANGE([116430105.ga_sessions_], TIMESTAMP('{min_date}'), TIMESTAMP('{max_date}'))) WHERE fullVisitorId IS NOT NULL AND hits.eventInfo.eventAction = 'Intent' AND hits.eventInfo.eventCategory = 'Support') b ON b.cid = a.donor_cid) WHERE IF(database_donors IS NULL AND ea = 'Intent' AND ec = 'Support', cid, NULL) IS NOT NULL GROUP BY date, donor_cid ORDER BY date desc**

- **Email Subscribers**
  - **Query:**

- SELECT date, cid AS cid_subscriber, IF(database_subscribers IS NULL AND mc_eid IS NOT NULL, mc_eid, NULL) AS mc_eid FROM (SELECT a.mc_eid AS database_subscribers, b.cid AS cid, TIMESTAMP(b.year + '-' + b.month + '-' + b.day) AS date, b.mc_eid AS mc_eid FROM [116430105.email_subscribers_mailchimp] a FULL OUTER JOIN EACH FLATTEN((SELECT fullVisitorId AS cid, REGEXP_EXTRACT(hits.page.pagePath, r'mc_eid=(.+)') AS mc_eid, /*Date*/ REGEXP_EXTRACT(date, '(^[0-9]{{4}})') as year, REGEXP_EXTRACT(date, '^[0-9]{{4}}([0-9]{{2}})') as month, REGEXP_EXTRACT(date, '([0-9]{{2}}$)') as day FROM (TABLE_DATE_RANGE([116430105.ga_sessions_], TIMESTAMP('{min_date}'), TIMESTAMP('{max_date}'))) WHERE REGEXP_EXTRACT(hits.page.pagePath, r'mc_eid=(.+)') IS NOT NULL),mc_eid) b ON b.mc_eid = a.mc_eid) WHERE IF(database_subscribers IS NULL AND mc_eid IS NOT NULL, mc_eid, NULL) IS NOT NULL GROUP BY date, cid_subscriber, mc_eid ORDER BY date desc

- **Unduplicated Known**
  - **Query:**
    - SELECT date, IF(new_subscriber_only = existing_subscriber_only, NULL, new_subscriber_only) AS subscriber_only, IF(new_donor_only = existing_donor_only, NULL, new_donor_only) AS donor_only, IF(new_member_only = existing_member_only, NULL, new_member_only) AS member_only, IF(new_member_donor = existing_member_donor, NULL, new_member_donor) AS member_donor_cd20, IF(new_donor_subscriber = existing_donor_subscriber, NULL, new_donor_subscriber) AS donor_subscriber_mceid, IF(new_member_subscriber = existing_member_subscriber, NULL, new_member_subscriber) AS member_subscriber_cd20, IF(new_member_donor_subscriber = existing_member_donor_subscriber, NULL, new_member_donor_subscriber) AS member_donor_subscriber_cd20 FROM (SELECT date, subscriber_only AS existing_subscriber_only, donor_only AS existing_donor_only, member_only AS existing_member_only, member_donor_cd20 AS

existing_member_donor, donor_subscriber_mceid AS existing_donor_subscriber, member_subscriber_cd20 AS existing_member_subscriber, member_donor_subscriber_cd20 AS existing_member_donor_subscriber, new_subscriber_only, new_donor_only, new_member_only, new_member_donor, new_donor_subscriber, new_member_subscriber, new_member_donor_subscriber FROM [116430105.unduplicated_known_people], (SELECT TIMESTAMP(year + '-' + month + '-' + day) AS date, IF(NOT ea = 'Intent' AND NOT ec = 'Support' AND cd19 IS NULL AND cd20 IS NOT NULL AND mc_eid IS NULL, cd20, NULL) AS new_member_only, IF(NOT ea = 'Intent' AND NOT ec = 'Support' AND cd19 IS NULL AND cd20 IS NULL AND mc_eid IS NOT NULL, mc_eid, NULL) AS new_subscriber_only, IF(ea = 'Intent' AND ec = 'Support' AND cd19 IS NULL AND cd20 IS NULL AND mc_eid IS NULL, cid, NULL) AS new_donor_only, IF(ea = 'Intent' AND ec = 'Support' AND cd19 IS NULL AND cd20 IS NOT NULL AND mc_eid IS NULL, cd20, NULL) AS new_member_donor, IF(ea = 'Intent' AND ec = 'Support' AND cd19 IS NULL AND cd20 IS NULL AND mc_eid IS NOT NULL, mc_eid, NULL) AS new_donor_subscriber, IF(NOT ea = 'Intent' AND NOT ec = 'Support' AND cd19 IS NULL AND cd20 IS NOT NULL AND mc_eid IS NOT NULL, cd20, NULL) AS new_member_subscriber, IF(ea = 'Intent' AND ec = 'Support' AND cd19 IS NULL AND cd20 IS NOT NULL AND mc_eid IS NOT NULL, cd20, NULL) AS new_member_donor_subscriber FROM (SELECT fullVisitorId AS cid, hits.eventInfo.eventCategory AS ec, hits.eventInfo.eventAction AS ea, IF(hits.customDimensions.index = 19, hits.customDimensions.value, NULL) AS cd19, IF(hits.customDimensions.index = 20, hits.customDimensions.value, NULL) AS cd20, REGEXP_EXTRACT(hits.page.pagePath, r'mc_eid=(.+)') AS mc_eid, /*Date*/ REGEXP_EXTRACT(date, '(^[0-9]{{4}})') as year, REGEXP_EXTRACT(date, '^[0-9]{{4}}([0-9]{{2}})') as month, REGEXP_EXTRACT(date, '([0-9]{{2}}$)') as day FROM TABLE_DATE_RANGE([116430105.ga_sessions_], TIMESTAMP('{min_date}'), TIMESTAMP('{max_date}'))) WHERE IF(NOT ea = 'Intent' AND NOT ec = 'Support' AND cd19 IS NULL AND cd20 IS NOT NULL AND mc_eid IS NULL, cd20, NULL) IS NOT NULL OR IF(NOT ea = 'Intent' AND NOT ec = 'Support' AND cd19 IS NULL AND cd20 IS NULL AND mc_eid IS NOT NULL, mc_eid, NULL) IS NOT NULL OR IF(ea = 'Intent' AND ec = 'Support' AND cd19 IS NULL AND cd20 IS NULL AND

mc_eid IS NULL, cid, NULL) IS NOT NULL OR IF(ea = 'Intent' AND ec = 'Support' AND cd19 IS NULL AND cd20 IS NOT NULL AND mc_eid IS NULL, cd20, NULL) IS NOT NULL OR IF(ea = 'Intent' AND ec = 'Support' AND cd19 IS NULL AND cd20 IS NULL AND mc_eid IS NOT NULL, mc_eid, NULL) IS NOT NULL OR IF(NOT ea = 'Intent' AND NOT ec = 'Support' AND cd19 IS NULL AND cd20 IS NOT NULL AND mc_eid IS NOT NULL, cd20, NULL) IS NOT NULL OR IF(ea = 'Intent' AND ec = 'Support' AND cd19 IS NULL AND cd20 IS NOT NULL AND mc_eid IS NOT NULL, cd20, NULL) IS NOT NULL)) WHERE IF(new_subscriber_only = existing_subscriber_only, NULL, new_subscriber_only) IS NOT NULL OR IF(new_donor_only = existing_donor_only, NULL, new_donor_only) IS NOT NULL OR IF(new_member_only = existing_member_only, NULL, new_member_only) IS NOT NULL OR IF(new_member_donor = existing_member_donor, NULL, new_member_donor) IS NOT NULL OR IF(new_donor_subscriber = existing_donor_subscriber, NULL, new_donor_subscriber) IS NOT NULL OR IF(new_member_subscriber = existing_member_subscriber, NULL, new_member_subscriber) IS NOT NULL OR IF(new_member_donor_subscriber = existing_member_donor_subscriber, NULL, new_member_donor_subscriber) IS NOT NULL GROUP BY date, subscriber_only, donor_only, member_only, member_donor_cd20, donor_subscriber_mceid, member_subscriber_cd20, member_donor_subscriber_cd20, ORDER BY date desc

- **Members**
  - **Query:**
    - SELECT date, cid AS cid_member, IF(database_members IS NULL AND cd20 IS NOT NULL, cd20, NULL) AS cd20 FROM (SELECT a.cd20 AS database_members, TIMESTAMP(b.year + '-' + b.month + '-' + b.day) AS date, b.cd20 AS cd20, b.cid AS cid FROM [116430105.members] a FULL OUTER JOIN EACH FLATTEN(FLATTEN((SELECT fullVisitorId AS cid, IF(hits.customDimensions.index = 20, hits.customDimensions.value, NULL) AS cd20, /*Date*/ REGEXP_EXTRACT(date, '(^[0-9]{{4}})') as year, REGEXP_EXTRACT(date, '^[0-9]{{4}}([0-9]{{2}})') as month, REGEXP_EXTRACT(date, '([0-9]{{2}}$)') as day FROM (TABLE_DATE_RANGE([116430105.ga_sessions_],

TIMESTAMP('{min_date}'), TIMESTAMP('{max_date}'))) WHERE IF(hits.customDimensions.index = 20, hits.customDimensions.value, NULL) IS NOT NULL),cd20), cd20) b ON b.cd20 = a.cd20) WHERE IF(database_members IS NULL AND cd20 IS NOT NULL, cd20, NULL) IS NOT NULL GROUP BY date, cid_member, cd20 ORDER BY date desc

Data Users Query

- SELECT
---MEMBERS---
EXACT_COUNT_DISTINCT(IF (cd20 IS NOT NULL, cd20, NULL)) AS TotalMembersThisWeek,
EXACT_COUNT_DISTINCT(database_members) +
  EXACT_COUNT_DISTINCT(IF (cd20 IS NOT NULL, cd20, NULL)) AS
KPI_TotalMembersKnownToMIP,
 ---EMAIL NEWSLETTER SUBSCRIBERS---
EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL, mc_eid, NULL)) AS
CameToSiteThroughEmail,
EXACT_COUNT_DISTINCT(database_subscribers) +
  EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL, mc_eid, NULL)) +
  EXACT_COUNT_DISTINCT(IF(cd19 IS NOT NULL, cd19, NULL)) AS
KPI_TotalEmailSubscribersKnownToMIP,
 ((EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL, mc_eid, NULL))*100)/
  (EXACT_COUNT_DISTINCT(database_subscribers) +
    EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL, mc_eid, NULL)) +
    EXACT_COUNT_DISTINCT(IF(cd19 IS NOT NULL, cd19, NULL)))) AS
KPI_PercentKnownSubsWhoCame,
EXACT_COUNT_DISTINCT(IF(cd19 IS NOT NULL, cd19, NULL)) AS KPI_NewEmailSubscribers,
 ---DONORS---
EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
  AND ec = 'Support', cid, NULL)) AS TotalDonorsThisWeek,
EXACT_COUNT_DISTINCT(database_donors) +
  EXACT_COUNT_DISTINCT(IF(cd19 IS NOT NULL, cid, NULL)) AS KPI_TotalDonorsKnownToMIP,
 ---TOTAL KNOWN USERS---
EXACT_COUNT_DISTINCT(IF (cd20 IS NOT NULL, cd20, NULL)) +
  EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL, mc_eid, NULL)) +
  EXACT_COUNT_DISTINCT(IF(ea = 'Intent'

```
      AND ec = 'Support', cid, NULL )) AS
Duplicated_MembersPlusCameThroughEmailPlusDonors,
  EXACT_COUNT_DISTINCT(IF(NOT ea = 'Intent'
    AND NOT ec = 'Support'
    AND cd19 IS NULL
    AND cd20 IS NOT NULL
    AND mc_eid IS NULL, cd20, NULL)) + -- member only
   EXACT_COUNT_DISTINCT(IF(NOT ea = 'Intent'
    AND NOT ec = 'Support'
    AND cd19 IS NULL
    AND cd20 IS NULL
    AND mc_eid IS NOT NULL, mc_eid, NULL)) + -- mc_eid subs only
   EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
    AND ec = 'Support'
    AND cd19 IS NULL
    AND cd20 IS NULL
    AND mc_eid IS NULL, cid, NULL)) + -- donors only
   EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
    AND ec = 'Support'
    AND cd19 IS NULL
    AND cd20 IS NOT NULL
    AND mc_eid IS NULL, cd20, NULL)) + -- member and donor
   EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
    AND ec = 'Support'
    AND cd19 IS NULL
    AND cd20 IS NULL
    AND mc_eid IS NOT NULL, mc_eid, NULL)) + -- donor and subscriber
   EXACT_COUNT_DISTINCT(IF(NOT ea = 'Intent'
    AND NOT ec = 'Support'
    AND cd19 IS NULL
    AND cd20 IS NOT NULL
    AND mc_eid IS NOT NULL, cd20, NULL)) + -- member and subscriber
   EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
    AND ec = 'Support'
    AND cd19 IS NULL
    AND cd20 IS NOT NULL -- member and subscriber and donor
    AND mc_eid IS NOT NULL, cd20, NULL)) AS Unduplicated_TotalUsersKPI,
 EXACT_COUNT_DISTINCT(database_members) +
```

```
EXACT_COUNT_DISTINCT(IF(cd20 IS NOT NULL, cd20, NULL)) +
EXACT_COUNT_DISTINCT(database_subscribers) +
EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL, mc_eid, NULL )) +
EXACT_COUNT_DISTINCT(IF(cd19 IS NOT NULL, cd19, NULL )) +
EXACT_COUNT_DISTINCT(database_donors) +
EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
  AND ec = 'Support', cid, NULL )) AS
Duplicated_Database_MembersPlusCameThroughEmailPlusDonors,
 EXACT_COUNT_DISTINCT(database_member_only) +
   EXACT_COUNT_DISTINCT(IF(NOT ea = 'Intent'
   AND NOT ec = 'Support'
   AND cd19 IS NULL
   AND cd20 IS NOT NULL
   AND mc_eid IS NULL, cd20, NULL)) + -- member only
   EXACT_COUNT_DISTINCT(database_subscriber_only) +
   EXACT_COUNT_DISTINCT(IF(NOT ea = 'Intent'
   AND NOT ec = 'Support'
   AND cd19 IS NULL
   AND cd20 IS NULL
   AND mc_eid IS NOT NULL, mc_eid, NULL)) + -- mc_eid subs only
   EXACT_COUNT_DISTINCT(database_donor_only) +
   EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
   AND ec = 'Support'
   AND cd19 IS NULL
   AND cd20 IS NULL
   AND mc_eid IS NULL, cid, NULL)) + -- donors only
   EXACT_COUNT_DISTINCT(database_member_donor) +
   EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
   AND ec = 'Support'
   AND cd19 IS NULL
   AND cd20 IS NOT NULL
   AND mc_eid IS NULL, cd20, NULL)) + -- member and donor
   EXACT_COUNT_DISTINCT(database_donor_subscriber) +
   EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
   AND ec = 'Support'
   AND cd19 IS NULL
   AND cd20 IS NULL
   AND mc_eid IS NOT NULL, mc_eid, NULL)) + -- donor and subscriber
```

```sql
    EXACT_COUNT_DISTINCT(database_member_subscriber) +
    EXACT_COUNT_DISTINCT(IF(NOT ea = 'Intent'
     AND NOT ec = 'Support'
     AND cd19 IS NULL
     AND cd20 IS NOT NULL
     AND mc_eid IS NOT NULL, cd20, NULL)) + -- member and subscriber
    EXACT_COUNT_DISTINCT(database_member_donor_subscriber) +
    EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
     AND ec = 'Support'
     AND cd19 IS NULL
     AND cd20 IS NOT NULL -- member and donor and subscriber
     AND mc_eid IS NOT NULL, cd20, NULL)) AS Unduplicated_Database_TotalUsersKPI
FROM
  (SELECT
    database_members, database_subscribers, database_donors,
    database_member_only, database_subscriber_only, database_donor_only,
    database_member_donor, database_donor_subscriber, database_member_subscriber,
    database_member_donor_subscriber,
    mc_eid_new AS mc_eid,
    cd19_new AS cd19,
    cd20_new AS cd20,
    cid, ec, ea
   FROM
     /*Only change ending timestamp*/
     (SELECT cd20 AS database_members
     FROM [116430105.members]
     WHERE date BETWEEN TIMESTAMP('2015-06-29') AND TIMESTAMP('{prev_max_date}')),
     (SELECT mc_eid AS database_subscribers
     FROM [116430105.email_subscribers_mailchimp]
     WHERE date BETWEEN TIMESTAMP('2015-06-29') AND TIMESTAMP('{prev_max_date}')),
     (SELECT donor_cid AS database_donors
     FROM [116430105.donors]
     WHERE date BETWEEN TIMESTAMP('2015-06-29') AND TIMESTAMP('{prev_max_date}')),
     (SELECT
       member_only AS database_member_only,
       subscriber_only AS database_subscriber_only,
       donor_only AS database_donor_only,
       member_donor_cd20 AS database_member_donor,
```

```
    donor_subscriber_mceid AS database_donor_subscriber,
    member_subscriber_cd20 AS database_member_subscriber,
    member_donor_subscriber_cd20 AS database_member_donor_subscriber
  FROM [116430105.unduplicated_known_people]
  WHERE date BETWEEN TIMESTAMP('2015-06-29') AND TIMESTAMP('{prev_max_date}')),
  (SELECT
    fullVisitorId AS cid,
    hits.eventInfo.eventCategory AS ec,
    hits.eventInfo.eventAction AS ea,
    IF(hits.customDimensions.index = 19, hits.customDimensions.value, NULL) AS cd19_new,
    IF(hits.customDimensions.index = 20, hits.customDimensions.value, NULL) AS cd20_new,
    REGEXP_EXTRACT(hits.page.pagePath, r'mc_eid=(.+)') AS mc_eid_new,
    /*Change both timestamps*/
    FROM (TABLE_DATE_RANGE([116430105.ga_sessions_], TIMESTAMP('{min_date}'),
TIMESTAMP('{max_date}') )) ) )
```

Data Stories Query

```
SELECT
  Combo_URL,
  Page_Title AS Article,
  SUM(IF(Hit_Type='PAGE', 1, NULL)) AS Pageviews,
  ------ SCROLL DEPTH ------
  SUM(IF(Event_Category='Scrolling'
    AND Event_Action='start', 1, NULL)) AS Scroll_Start,
  SUM(IF(Event_Category='Scrolling'
    AND Event_Action='25%', 1, NULL)) AS Scroll_25,
  SUM(IF(Event_Category='Scrolling'
    AND Event_Action='50%', 1, NULL)) AS Scroll_50,
  SUM(IF(Event_Category='Scrolling'
    AND Event_Action='75%', 1, NULL)) AS Scroll_75,
  SUM(IF(Event_Category='Scrolling'
    AND Event_Action='100%', 1, NULL)) AS Scroll_100,
  SUM(IF(Event_Category='Scrolling'
    AND Event_Action='supplemental', 1, NULL)) AS Scroll_Supplemental,
  SUM(IF(Event_Category='Scrolling'
    AND Event_Action='end', 1, NULL)) AS Scroll_End,
```

```
----- TIME ON ARTICLE -----
SUM(IF(Event_Category='Time'
  AND Event_Action='15s', 1, NULL)) AS Time_15,
SUM(IF(Event_Category='Time'
  AND Event_Action='30s', 1, NULL)) AS Time_30,
SUM(IF(Event_Category='Time'
  AND Event_Action='45s', 1, NULL)) AS Time_45,
SUM(IF(Event_Category='Time'
  AND Event_Action='60s', 1, NULL)) AS Time_60,
SUM(IF(Event_Category='Time'
  AND Event_Action='75s', 1, NULL)) AS Time_75,
SUM(IF(Event_Category='Time'
  AND Event_Action='90s', 1, NULL)) AS Time_90,
------ SOCIAL ACTIONS ------
SUM(IF(Social_Action='comment-submit', 1, NULL)) AS Comments,
SUM(IF(Social_Action='republish', 1, NULL)) AS Republish,
SUM(IF(Social_Action='email', 1, NULL)) AS Emails,
SUM(IF(Social_Action='tweet' AND Social_Network='Twitter', 1, NULL)) AS Tweets,
SUM(IF(Social_Action='recommend' AND Social_Network='Facebook', 1, NULL)) AS
Facebook_Recommendations,
SUM(IF(Event_Category='Related Link' AND (Event_Label CONTAINS "tribpedia"), 1, NULL)) AS
Tribpedia_Related_Clicks,
SUM(IF(Event_Category='Related Link', 1, NULL)) AS Related_Clicks
FROM (
SELECT
  hits.type AS Hit_Type,
  hits.page.pagePath AS Page_Path,
  REGEXP_EXTRACT(hits.page.pagePath, r'(\/\d{{4}}\/\d{{2}}\/\d{{2}}\/[^\/]+)') AS Combo_URL,
  IF(hits.page.pageTitle CONTAINS "Comments on" or hits.page.pageTitle CONTAINS
"Comentarios sobre ",
    REGEXP_EXTRACT(hits.page.pageTitle,r'["“"](.*)["”"]'),
    REGEXP_EXTRACT(hits.page.pageTitle,r'(.*) \| The Texas Tribune')) AS Page_Title,
  hits.eventInfo.eventCategory AS Event_Category,
  hits.eventInfo.eventAction AS Event_Action,
  hits.social.socialInteractionNetwork AS Social_Network,
  hits.social.socialInteractionAction AS Social_Action,
  hits.eventInfo.eventLabel AS Event_Label
FROM (
```

```
    TABLE_DATE_RANGE([116430105.ga_sessions_], TIMESTAMP('{min_date}'),
TIMESTAMP('{max_date}') )
 )
  HAVING
    REGEXP_MATCH(Page_Path, r'/\d{{4}}/\d{{2}}/\d{{2}}/')=true)
GROUP BY
  Combo_URL, Article
ORDER BY
  Pageviews DESC
```

Data Quality Queries

Table 1 Query
```
SELECT
  --------- IDENTIFIED USERS ---------
  EXACT_COUNT_DISTINCT(cid) AS MIP_Users
FROM (
  SELECT
    fullVisitorId AS cid,
    FROM ( TABLE_DATE_RANGE([116430105.ga_sessions_], TIMESTAMP('{min_date}'),
TIMESTAMP('{max_date}') )) )
```

Table 2 Query
```
SELECT
  --Already in database and came to site this week (didn't subscribe)--
  EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL
    AND database_subscribers IS NOT NULL
    AND cd19 IS NULL, cid, NULL)) AS I_inDatabaseCameToSite,
  EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL
    AND database_subscribers IS NOT NULL
    AND cd19 IS NULL, mc_eid, NULL)) AS K_inDatabaseCameToSite,
  --Not in database and came to site this week (didn't subscribe)--
  EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL
    AND database_subscribers IS NULL
    AND cd19 IS NULL, cid, NULL)) AS I_notInDatabaseCameToSite,
  EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL
    AND database_subscribers IS NULL
```

```
    AND cd19 IS NULL, mc_eid, NULL)) AS K_notInDatabaseCameToSite,
  --New subscribers who clicked on email this week--
  EXACT_COUNT_DISTINCT(IF(IF(mc_eid IS NOT NULL, cid, NULL) =
    IF(cd19 IS NOT NULL AND ec = 'Email Subscribe', cid, NULL), cid, NULL)) AS
I_newSubscriberCameThroughEmail,
  EXACT_COUNT_DISTINCT(IF(IF(mc_eid IS NOT NULL, cid, NULL) =
    IF(cd19 IS NOT NULL AND ec = 'Email Subscribe', cid, NULL), mc_eid, NULL)) AS
K_newSubscriberCameThroughEmail,
  --Total email visitors this week--
  EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL, cid, NULL)) AS I_SubscribersThisWeek,
  EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL, mc_eid, NULL)) AS K_SubscribersThisWeek,
  --New subscribers--
  EXACT_COUNT_DISTINCT(IF(cd19 IS NOT NULL
    AND ec = 'Email Subscribe', cid, NULL)) AS I_NewSubscribers,
  EXACT_COUNT_DISTINCT(IF(cd19 IS NOT NULL
    AND ec = 'Email Subscribe', cd19, NULL)) AS K_NewSubscribers,
  --Total subscribers in database now--
  EXACT_COUNT_DISTINCT(database_cid) +
    EXACT_COUNT_DISTINCT(IF(cd19 IS NOT NULL
    AND ec = 'Email Subscribe', cid, NULL)) +
    EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL
    AND database_subscribers IS NULL
    AND cd19 IS NULL, cid, NULL)) AS I_TotalDatabaseSubscribers,
  EXACT_COUNT_DISTINCT(database_subscribers) +
    EXACT_COUNT_DISTINCT(IF(cd19 IS NOT NULL
    AND ec = 'Email Subscribe', cd19, NULL)) +
    EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL
    AND database_subscribers IS NULL
    AND cd19 IS NULL, mc_eid, NULL)) AS K_TotalDatabaseSubscribers,
  --Percent of database subscribers who clicked on email this week--
  (EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL, mc_eid, NULL)) * 100)/
    (EXACT_COUNT_DISTINCT(database_subscribers) +
      EXACT_COUNT_DISTINCT(IF(cd19 IS NOT NULL
    AND ec = 'Email Subscribe', cd19, NULL)) +
      EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL
      AND database_subscribers IS NULL
      AND cd19 IS NULL, mc_eid, NULL))) AS K_PercentDatabaseSubscribersWhoCame,
  --Email newsletter clicks--
```

```
   NULL AS EmailNewsletterClicks
FROM
 (SELECT
   a.mc_eid AS database_subscribers, a.cid_subscriber AS database_cid,
   b.mc_eid AS mc_eid, b.ec AS ec, b.ea AS ea, b.cd19 AS cd19, b.cid AS cid
  FROM
  /*everything in database from beginning (6/29/15) until the start of this week*/
  (SELECT mc_eid, cid_subscriber
   FROM [116430105.email_subscribers_mailchimp]
   WHERE date BETWEEN TIMESTAMP('2015-06-29') AND TIMESTAMP('{prev_max_date}') )a
   FULL OUTER JOIN EACH
   /*this week*/
   FLATTEN((SELECT
      fullVisitorId AS cid,
      REGEXP_EXTRACT(hits.page.pagePath, r'mc_eid=(.+)') AS mc_eid,
      hits.eventInfo.eventCategory AS ec,
      hits.eventInfo.eventAction AS ea,
      IF(hits.customDimensions.index = 19, hits.customDimensions.value, NULL) AS cd19
    FROM (TABLE_DATE_RANGE([116430105.ga_sessions_], TIMESTAMP('{min_date}'),
TIMESTAMP('{max_date}') ))
    WHERE REGEXP_EXTRACT(hits.page.pagePath, r'mc_eid=(.+)') IS NOT NULL
      OR IF(hits.customDimensions.index = 19, hits.customDimensions.value, NULL) IS NOT
NULL),mc_eid) b
   ON b.mc_eid = a.mc_eid)
```

Table 3 Query

```
SELECT
 --Database donors who visited again--
 EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
  AND ec = 'Support'
  AND database_donors IS NOT NULL, cid, NULL)) AS I_databaseDonorsWhoVisited,
 NULL AS K_databaseDonorsWhoVisited,
 --New donors not in database--
 EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
  AND ec = 'Support'
  AND database_donors IS NULL, cid, NULL)) AS I_donatedOnSiteForFirstTime,
 NULL AS K_donatedOnSiteForFirstTime,
```

```
  --Total donors on site this week--
  EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
    AND ec = 'Support', cid, NULL)) AS I_totalDonorsOnSiteThisWeek,
  NULL AS K_totalDonorsOnSiteThisWeek,
  --Total donors in database--
  EXACT_COUNT_DISTINCT(database_donors) +
    EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
    AND ec = 'Support'
    AND database_donors IS NULL, cid, NULL)) AS I_totalDonorsInDatabase,
  NULL AS K_totalDonorsInDatabase,
  --Percent database donors on site this week--
  NULL AS K_percentDatabaseDonorsWhoCame
FROM
  (SELECT
    a.Donors AS database_donors,
    b.ec AS ec, b.ea AS ea, b.cid AS cid
   FROM
    (SELECT donor_cid AS Donors
     FROM [116430105.donors]
     WHERE date BETWEEN TIMESTAMP('2015-06-29') AND TIMESTAMP('{prev_max_date}') )a
     FULL OUTER JOIN EACH
     (SELECT
       fullVisitorId AS cid,
       hits.eventInfo.eventCategory AS ec,
       hits.eventInfo.eventAction AS ea,
      FROM (TABLE_DATE_RANGE([116430105.ga_sessions_], TIMESTAMP('{min_date}'),
TIMESTAMP('{max_date}') ))
      WHERE hits.eventInfo.eventAction = 'Intent'
       AND hits.eventInfo.eventCategory = 'Support') b
     ON b.cid = a.Donors)
```

Table 4 Query
```
SELECT
  ---Total known individuals---
 EXACT_COUNT_DISTINCT(IF(NOT ea = 'Intent'
    AND NOT ec = 'Support'
    AND cd19 IS NULL
    AND cd20 IS NOT NULL
```

```
 AND mc_eid IS NULL, cd20, NULL)) + -- member only
EXACT_COUNT_DISTINCT(IF(NOT ea = 'Intent'
 AND NOT ec = 'Support'
 AND cd19 IS NULL
 AND cd20 IS NULL
 AND mc_eid IS NOT NULL, mc_eid, NULL)) + -- mc_eid subs only
EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
 AND ec = 'Support'
 AND cd19 IS NULL
 AND cd20 IS NULL
 AND mc_eid IS NULL, cid, NULL)) + -- donors only
EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
 AND ec = 'Support'
 AND cd19 IS NULL
 AND cd20 IS NOT NULL
 AND mc_eid IS NULL, cd20, NULL)) + -- member and donor
EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
 AND ec = 'Support'
 AND cd19 IS NULL
 AND cd20 IS NULL
 AND mc_eid IS NOT NULL, mc_eid, NULL)) + -- donor and subscriber
EXACT_COUNT_DISTINCT(IF(NOT ea = 'Intent'
 AND NOT ec = 'Support'
 AND cd19 IS NULL
 AND cd20 IS NOT NULL
 AND mc_eid IS NOT NULL, cd20, NULL)) + -- member and subscriber
EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
 AND ec = 'Support'
 AND cd19 IS NULL
 AND cd20 IS NOT NULL -- member and subscriber and donor
 AND mc_eid IS NOT NULL, cd20, NULL)) AS K_individualsWhoCameThisWeek,
EXACT_COUNT_DISTINCT(database_member_only) +
 EXACT_COUNT_DISTINCT(IF(NOT ea = 'Intent'
 AND NOT ec = 'Support'
 AND cd19 IS NULL
 AND cd20 IS NOT NULL
 AND mc_eid IS NULL, cd20, NULL)) + -- member only
 EXACT_COUNT_DISTINCT(database_subscriber_only) +
```

```
    EXACT_COUNT_DISTINCT(IF(NOT ea = 'Intent'
     AND NOT ec = 'Support'
     AND cd19 IS NULL
     AND cd20 IS NULL
     AND mc_eid IS NOT NULL, mc_eid, NULL)) + -- mc_eid subs only
    EXACT_COUNT_DISTINCT(database_donor_only) +
    EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
     AND ec = 'Support'
     AND cd19 IS NULL
     AND cd20 IS NULL
     AND mc_eid IS NULL, cid, NULL)) + -- donors only
    EXACT_COUNT_DISTINCT(database_member_donor) +
    EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
     AND ec = 'Support'
     AND cd19 IS NULL
     AND cd20 IS NOT NULL
     AND mc_eid IS NULL, cd20, NULL)) + -- member and donor
    EXACT_COUNT_DISTINCT(database_donor_subscriber) +
    EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
     AND ec = 'Support'
     AND cd19 IS NULL
     AND cd20 IS NULL
     AND mc_eid IS NOT NULL, mc_eid, NULL)) + -- donor and subscriber
    EXACT_COUNT_DISTINCT(database_member_subscriber) +
    EXACT_COUNT_DISTINCT(IF(NOT ea = 'Intent'
     AND NOT ec = 'Support'
     AND cd19 IS NULL
     AND cd20 IS NOT NULL
     AND mc_eid IS NOT NULL, cd20, NULL)) + -- member and subscriber
    EXACT_COUNT_DISTINCT(database_member_donor_subscriber) +
    EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
     AND ec = 'Support'
     AND cd19 IS NULL
     AND cd20 IS NOT NULL -- member and donor and subscriber
     AND mc_eid IS NOT NULL, cd20, NULL)) AS K_individualsInDatabase,
  ((EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL
     AND cd20 IS NULL
     AND NOT ea = 'Intent'
```

```sql
  AND NOT ec = 'Support', mc_eid, NULL )) + -- email clicker only
NULL +                         -- donors only
EXACT_COUNT_DISTINCT(IF(mc_eid IS NULL
    AND cd20 IS NOT NULL
    AND NOT ea = 'Intent'
    AND NOT ec = 'Support', cd20, NULL )) + -- member only
EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL
    AND cd20 IS NULL
    AND ea = 'Intent'
    AND ec = 'Support', mc_eid, NULL )) + -- email clicker and donor
EXACT_COUNT_DISTINCT(IF(mc_eid IS NULL
    AND cd20 IS NOT NULL
    AND ea = 'Intent'
    AND ec = 'Support', cd20, NULL )) + -- member and donor
EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL
    AND cd20 IS NOT NULL
    AND NOT ea = 'Intent'
    AND NOT ec = 'Support', cd20, NULL )) + -- member and email clicker
EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL   -- member and email clicker and donor
    AND cd20 IS NOT NULL
    AND ea = 'Intent'
    AND ec = 'Support', cd20, NULL )))*100) /
 (EXACT_COUNT_DISTINCT(database_subscriber_only) +
EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL
 AND cd20 IS NULL
 AND NOT ea = 'Intent'
 AND NOT ec = 'Support', mc_eid, NULL )) + -- email clicker only
EXACT_COUNT_DISTINCT(database_donor_only) +
NULL +                      -- donors only
EXACT_COUNT_DISTINCT(database_member_only) +
EXACT_COUNT_DISTINCT(IF(mc_eid IS NULL
    AND cd20 IS NOT NULL
    AND NOT ea = 'Intent'
    AND NOT ec = 'Support', cd20, NULL )) + -- member only
EXACT_COUNT_DISTINCT(database_donor_subscriber) +
EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL
    AND cd20 IS NULL
    AND ea = 'Intent'
```

```
                AND ec = 'Support', mc_eid, NULL )) + -- email clicker and donor
          EXACT_COUNT_DISTINCT(database_member_donor) +
          EXACT_COUNT_DISTINCT(IF(mc_eid IS NULL
                AND cd20 IS NOT NULL
                AND ea = 'Intent'
                AND ec = 'Support', cd20, NULL )) + -- member and donor
          EXACT_COUNT_DISTINCT(database_member_subscriber) +
          EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL
                AND cd20 IS NOT NULL
                AND NOT ea = 'Intent'
                AND NOT ec = 'Support', cd20, NULL )) + -- member and email clicker
          EXACT_COUNT_DISTINCT(database_member_donor_subscriber) +
          EXACT_COUNT_DISTINCT(IF(mc_eid IS NOT NULL  -- member and email clicker and donor
                AND cd20 IS NOT NULL
                AND ea = 'Intent'
                AND ec = 'Support', cd20, NULL ))) AS K_percentDatabaseIndividualsWhoCame
  FROM
    (SELECT
      database_member_only, database_subscriber_only,
      database_donor_only, database_member_donor,
      database_donor_subscriber, database_member_subscriber,
      database_member_donor_subscriber, cid,
      mc_eid_new AS mc_eid, ec, ea, cd19, cd20
     FROM
       /*Only change ending timestamp*/
       (SELECT member_only AS database_member_only,
       subscriber_only AS database_subscriber_only,
       donor_only AS database_donor_only,
       member_donor_cd20 AS database_member_donor,
       donor_subscriber_mceid AS database_donor_subscriber,
       member_subscriber_cd20 AS database_member_subscriber,
       member_donor_subscriber_cd20 AS database_member_donor_subscriber
       FROM [116430105.unduplicated_known_people]
       WHERE date BETWEEN TIMESTAMP('2015-06-29') AND TIMESTAMP('{prev_max_date}')),
       (SELECT
         fullVisitorId AS cid,
         REGEXP_EXTRACT(hits.page.pagePath, r'mc_eid=(.+)') AS mc_eid_new,
         hits.eventInfo.eventCategory AS ec,
```

```
    hits.eventInfo.eventAction AS ea,
    IF(hits.customDimensions.index = 19, hits.customDimensions.value, NULL) AS cd19,
    IF(hits.customDimensions.index = 20, hits.customDimensions.value, NULL) AS cd20
    /*Change both timestamps*/
    FROM (TABLE_DATE_RANGE([116430105.ga_sessions_], TIMESTAMP('{min_date}'),
TIMESTAMP('{max_date}') )) ) )
```

Table5 Query
```
SELECT
  EXACT_COUNT_DISTINCT(IF(ea = 'Intent'
    AND ec = 'Support', cid, NULL )) AS Total_Identified_Donors_This_Week
FROM
  (SELECT
    hits.eventInfo.eventCategory AS ec,
    hits.eventInfo.eventAction AS ea,
    fullVisitorId AS cid
   FROM (TABLE_DATE_RANGE([116430105.ga_sessions_], TIMESTAMP('{min_date}'),
TIMESTAMP('{max_date}') )) )
```

Table 6 Query
```
SELECT
  --Database members who visited again--
  EXACT_COUNT_DISTINCT(IF(cd20 IS NOT NULL
    AND database_members IS NOT NULL, cid, NULL)) AS I_databaseMembersWhoVisited,
  EXACT_COUNT_DISTINCT(IF(cd20 IS NOT NULL
    AND database_members IS NOT NULL, cd20, NULL)) AS K_databaseMembersWhoVisited,
  --New members not in database--
  EXACT_COUNT_DISTINCT(IF(cd20 IS NOT NULL
    AND database_members IS NULL, cid, NULL)) AS I_loggedInOnSiteForFirstTime,
  EXACT_COUNT_DISTINCT(IF(cd20 IS NOT NULL
    AND database_members IS NULL, cd20, NULL)) AS K_loggedInOnSiteForFirstTime,
  --Total members on site this week--
  EXACT_COUNT_DISTINCT(IF(cd20 IS NOT NULL, cid, NULL)) AS
I_totalMembersOnSiteThisWeek,
  EXACT_COUNT_DISTINCT(IF(cd20 IS NOT NULL, cd20, NULL)) AS
K_totalMembersOnSiteThisWeek,
  --Total members in database--
```

```
     EXACT_COUNT_DISTINCT(database_members) +
      EXACT_COUNT_DISTINCT(IF(cd20 IS NOT NULL
      AND database_members IS NULL, cid, NULL)) AS I_totalMembersInDatabase,
     EXACT_COUNT_DISTINCT(database_members) +
      EXACT_COUNT_DISTINCT(IF(cd20 IS NOT NULL
      AND database_members IS NULL, cd20, NULL)) AS K_totalMembersInDatabase,
     --Percent database members on site this week--
     (EXACT_COUNT_DISTINCT(IF(cd20 IS NOT NULL, cd20, NULL))*100)/
      (EXACT_COUNT_DISTINCT(database_members) +
      EXACT_COUNT_DISTINCT(IF(cd20 IS NOT NULL
      AND database_members IS NULL, cd20, NULL))) AS K_percentDatabaseMembersWhoCame
FROM
  (SELECT
     a.cd20 AS database_members, a.cid_member AS database_cid,
     b.cd20 AS cd20, b.cid AS cid
    FROM
    (SELECT cd20, cid_member
    FROM [116430105.members]
    WHERE date BETWEEN TIMESTAMP('2015-06-29') AND TIMESTAMP('{prev_max_date}') )a
    FULL OUTER JOIN EACH
     FLATTEN(FLATTEN((SELECT
        fullVisitorId AS cid,
        IF(hits.customDimensions.index = 20, hits.customDimensions.value, NULL) AS cd20
      FROM (TABLE_DATE_RANGE([116430105.ga_sessions_], TIMESTAMP('{min_date}'),
TIMESTAMP('{max_date}') ))
      WHERE IF(hits.customDimensions.index = 20, hits.customDimensions.value, NULL) IS NOT
NULL), cd20), cd20) b
     ON b.cd20 = a.cd20)
```

OTHER QUERIES

User Interactions Quarterly
```
SELECT
  Page_Title,
  COMBO_URL,
  SUM(IF(Hit_Type='PAGE', 1, NULL)) AS Pageviews,
  ------ SCROLL DEPTH ------
  SUM(IF(Event_Category='Scrolling'
```

```sql
      AND Event_Action='start', 1, NULL)) AS Scroll_Start,
  SUM(IF(Event_Category='Scrolling'
    AND Event_Action='25%', 1, NULL)) AS Scroll_25,
  SUM(IF(Event_Category='Scrolling'
    AND Event_Action='50%', 1, NULL)) AS Scroll_50,
  SUM(IF(Event_Category='Scrolling'
    AND Event_Action='75%', 1, NULL)) AS Scroll_75,
  SUM(IF(Event_Category='Scrolling'
    AND Event_Action='100%', 1, NULL)) AS Scroll_100,
  SUM(IF(Event_Category='Scrolling'
    AND Event_Action='supplemental', 1, NULL)) AS Scroll_Supplemental,
  SUM(IF(Event_Category='Scrolling'
    AND Event_Action='end', 1, NULL)) AS Scroll_End,
  ----- TIME ON ARTICLE -----
  SUM(IF(Event_Category='Time'
    AND Event_Action='15s', 1, NULL)) AS Time_15,
  SUM(IF(Event_Category='Time'
    AND Event_Action='30s', 1, NULL)) AS Time_30,
  SUM(IF(Event_Category='Time'
    AND Event_Action='45s', 1, NULL)) AS Time_45,
  SUM(IF(Event_Category='Time'
    AND Event_Action='60s', 1, NULL)) AS Time_60,
  SUM(IF(Event_Category='Time'
    AND Event_Action='75s', 1, NULL)) AS Time_75,
  SUM(IF(Event_Category='Time'
    AND Event_Action='90s', 1, NULL)) AS Time_90,
  ------ SOCIAL ACTIONS ------
  SUM(IF(Social_Action='comment-submit', 1, NULL)) AS Comments,
  SUM(IF(Social_Action='republish', 1, NULL)) AS Republish,
  SUM(IF(Social_Action='email', 1, NULL)) AS Emails,
  SUM(IF(Social_Action='tweet'AND Social_Network='Twitter', 1, NULL)) AS Tweets,
  SUM(IF(Social_Action='recommend' AND Social_Network='Facebook', 1, NULL)) AS
Facebook_Recommendations,
  SUM(IF(Event_Category='Related Link', 1, NULL)) AS Related_Clicks,
  SUM(IF(Event_Category='Related Link' AND (Event_Label CONTAINS "tribpedia"), 1, NULL)) as
TribpediaRelatedLinks
FROM (
  SELECT
```

```sql
    hits.type AS Hit_Type,
    IF(hits.page.pageTitle CONTAINS "Comments on" or hits.page.pageTitle CONTAINS
"Comentarios sobre",
      REGEXP_EXTRACT(hits.page.pageTitle,r'["""](.*)["""]'),
      REGEXP_EXTRACT(hits.page.pageTitle,r'(.*) \| The Texas Tribune')) AS Page_Title,
    hits.page.pagePath AS Page_Path,
    REGEXP_EXTRACT(hits.page.pagePath, r'(\/\d{4}\/\d{2}\/\d{2}\/[^\/]+)') AS COMBO_URL,
    /*Date*/
    REGEXP_EXTRACT(hits.page.pagePath, r'/(\d{4})/\d{2}/\d{2}/') as yyyy,
    REGEXP_EXTRACT(hits.page.pagePath, r'/\d{4}/(\d{2})/\d{2}/') as mm,
    REGEXP_EXTRACT(hits.page.pagePath, r'/\d{4}/\d{2}/(\d{2})/') as dd,
    hits.eventInfo.eventCategory AS Event_Category,
    hits.eventInfo.eventAction AS Event_Action,
    hits.social.socialInteractionNetwork AS Social_Network,
    hits.social.socialInteractionAction AS Social_Action,
    hits.eventInfo.eventLabel AS Event_Label
    FROM (
    TABLE_DATE_RANGE([116430105.ga_sessions_], TIMESTAMP('2016-04-15'),
TIMESTAMP('2016-07-23')))
  HAVING
    REGEXP_MATCH(Page_Path, r'/\d{4}/\d{2}/\d{2}/')=true
    )
  WHERE TIMESTAMP(yyyy + '-' + mm + '-' + dd) BETWEEN TIMESTAMP('2016-04-01') AND
TIMESTAMP('2016-06-30')
GROUP BY
  COMBO_URL, Page_Title
ORDER BY
  Pageviews DESC
```