Todo’s left on hitmarker API Project:

* Cron “Code” to our backup folder “Cron Jobs”
* snapshot VM process
* Append missing data from May 6 and 7
* De Dupe May 5
  + Save code for dedupe
* Migrate and rename VM
  + Update documentation if we do that
* Document
  + how to access the backup JSON file on the VM
  + how to manually execute .py scrip to append ia\_summary
* Test process for when we add new search from voxjar build to Py to Gcloud to PowerBI
* Validate we are getting all calls out (GUI shows more results per day than API)

**Voxjar to Google Cloud BigQuery API Process Overview:**

**Note: this is for HIT MARKERS denoting which searches did/did not match calls. This does not pull any transcripts or tags.**

For the Oakstreet Health Voxjar instance, we are using a voxjar created, Dougherty-Martinsen edited, python script, ‘voxjarapipull.py’.

In this file, row#362 denotes the time frame we are pulling data from. We configure this to be (0,1) pulling a single day of data for yesterday.

We pull with 1-day lag time giving voxjar 24 hours to process audio files into text transcripts.

Which Searches that come from voxjar are denoted in Lines #448 🡪

“search name”,

BEST PRACTICE in VOXJAR for making searches

* When using voxjar GUI/Front end do not use “Exp” searches unless one is specifically working on the API
* Production searches set to last 7 days
* named:
  + ‘exp7”searchname”’ all lowercase
  + Names don’t contain spaces. If it’s absolutely necessary to breakup text one has to, use “\_“ for spaces
  + exp denoting excellent search for production
* Do not apply any filters (other than 7days) to exp searches

**Cron Jobs:**

**Best Practice: snapshot every 90days for BC**

**Specific Details:**

* Google project: dm-bv-data-storage-4-6-2020
  + Storage bucket
    - voxjar\_daily\_py\_script
    - Daily JSON run: Buckets/voxjar\_daily\_py\_script/daily\_voxjar\_run.json
  + BigQ dataset:
    - vjar\_oh\_prod
      * oh = 1st and last letter of client name **O**ak Street Healt**H**
  + BigQ dataset tables
    - ia\_summary
      * Interaction Analytics Summary
      * production table for BI tools
      * Schema is located
  + SQL Queries for table maintenance
    - Delete column
    - List of dupes
    - Union\_Query
    - DeDupe