DIY Search Tools

Practical Applications of sierra DNA and sierra API

Cameron Schrode – ILS Administrator

cschrode@akronlibrary.org

Akron-Summit County Public Library

Introduction

- **Sierra DNA *(SQL)***
  - Fetch information from database
  - Cannot edit or add records
  - Useful for Statistics and Reports
  - Today's Example: Business Intelligence Dashboard

- **Sierra API *(Programmatic)***
  - Fetch, Edit, and Add Records
  - Useful for Scheduled Tasks, Rapid Update, Complex I/O
  - Today's Example: Email Address Repair

Business Intelligence

- **What is business intelligence?**
  1. Metrics and Benchmarks
  2. Measure Input / Output of Processes
  3. Reporting
  4. Collaborative
  5. Knowledge Management
  6. (optional) Real-Time Data

- **Metabase**
  - Alternatives
    - Tableau
    - Power BI for Office 365

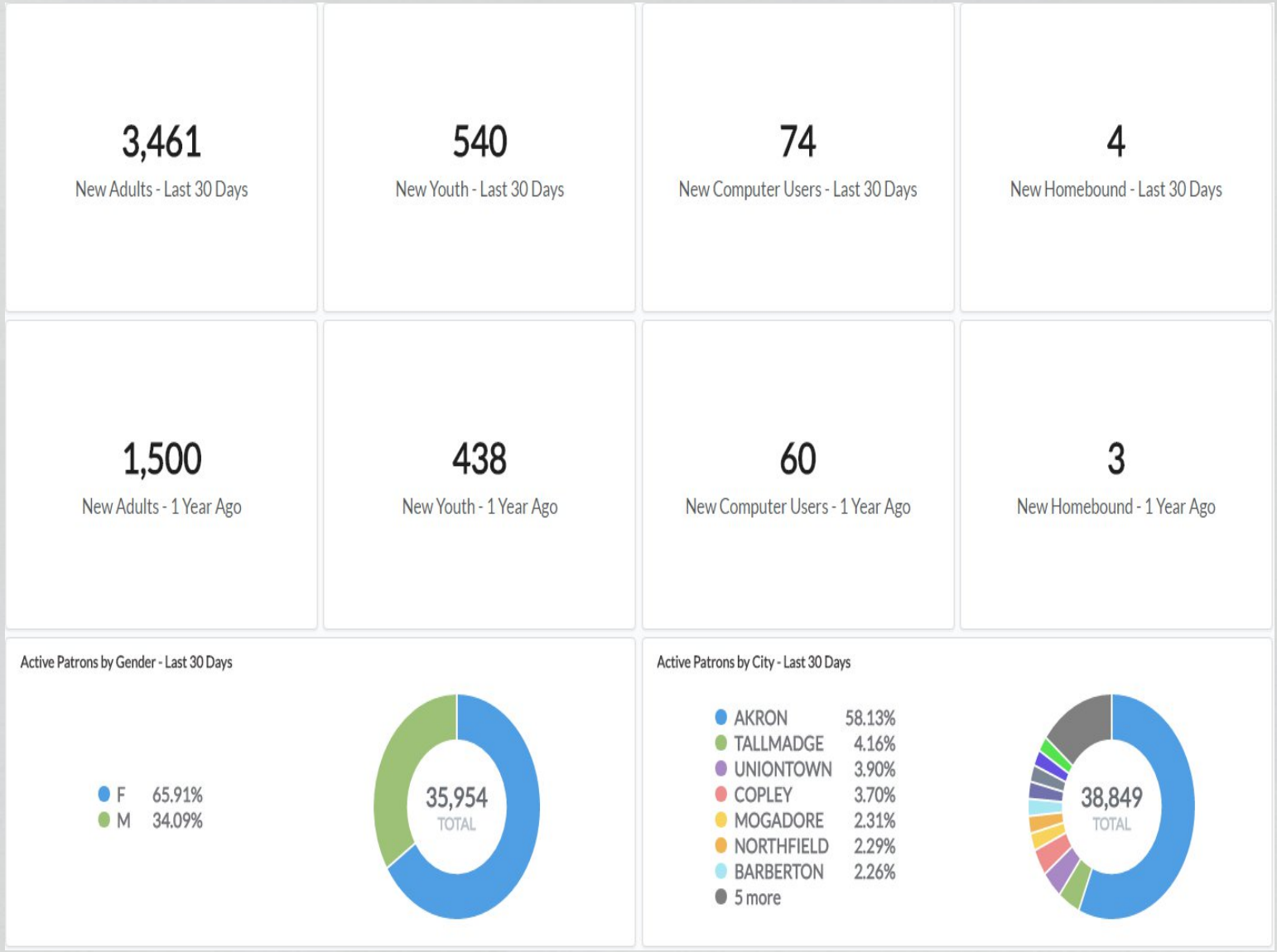
Metabase

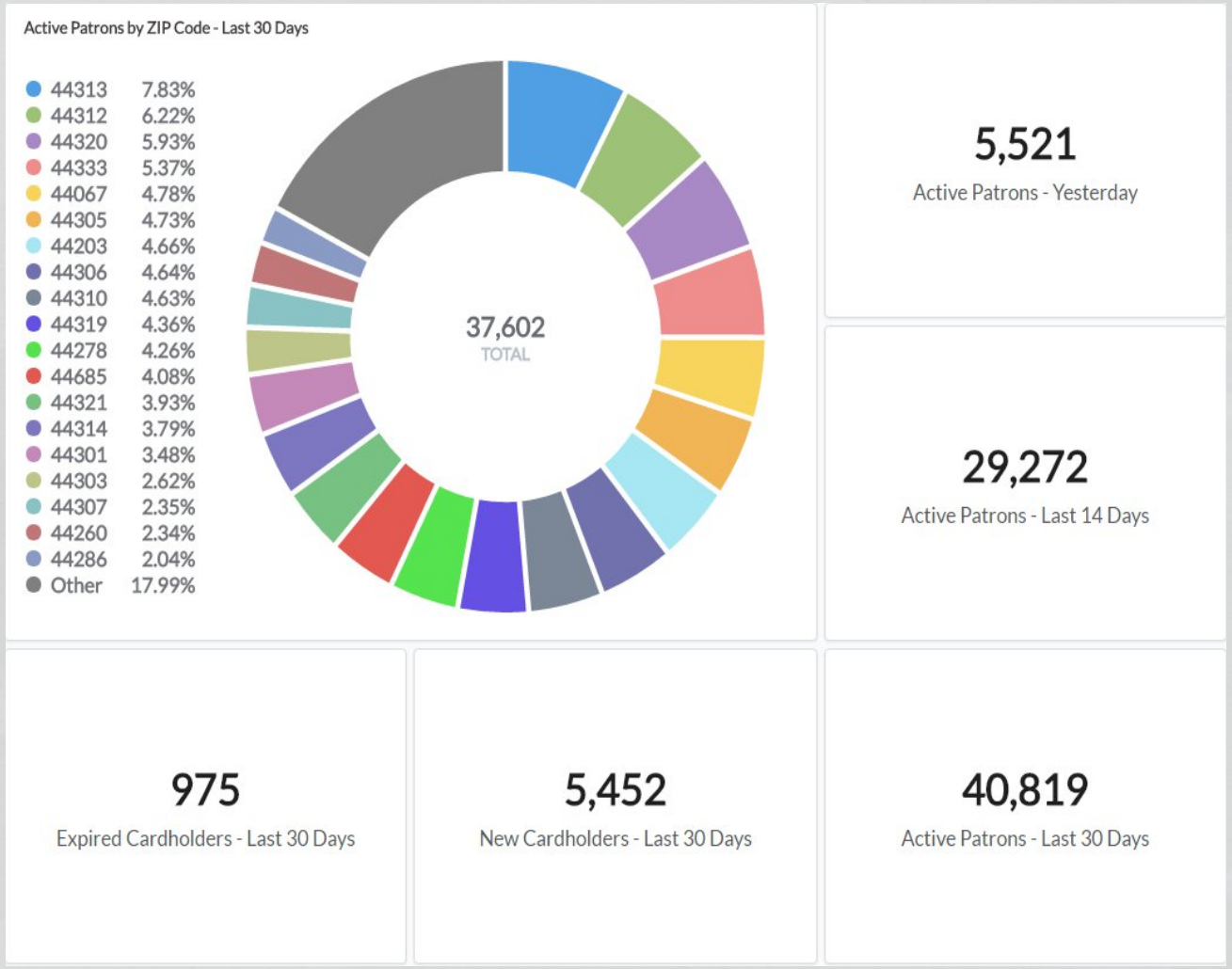
Dashboards

- Acquisitions / Cataloging metrics
- Branch Audience reports
- Cardholder metrics
- Circulation metrics
- Demographic metrics
- Teleforms metrics

Components

- Question => Answer
- Collection
  - Information Architecture
- Dashboard
  - Visual Representation
- Pulse
  - Reporting and Distribution





Metabase

1. **Requires Java 7+**
   http://java.com/
2. **Requires Metabase.jar**
   http://metabase.com/

Metabase uses H2 by default

- PostgreSQL for backup capability

Gmail for "Pulse" Emails

- Archival
- Year-End Statistics

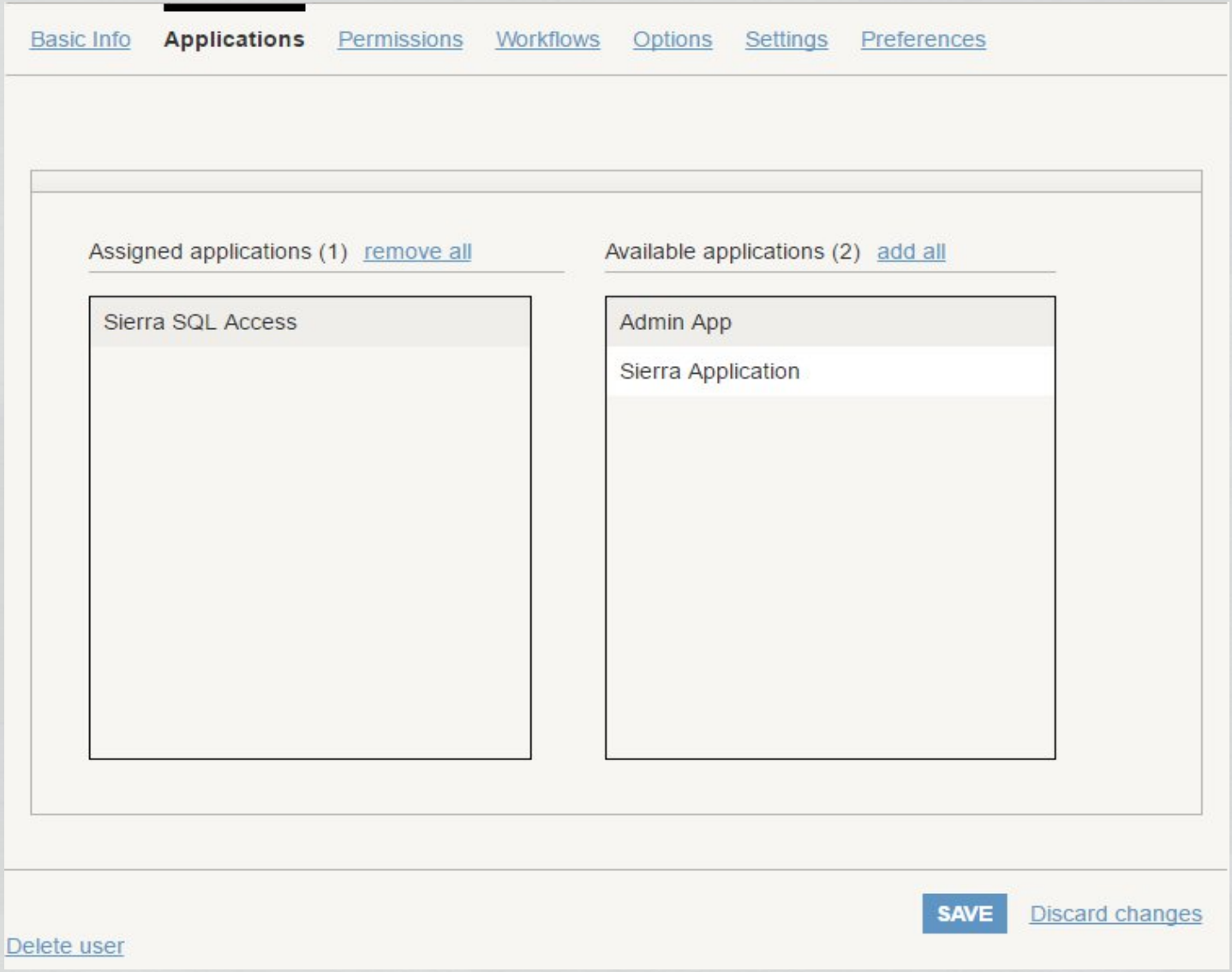
Initial Questions to Ask

1. **What is my Sierra Database server?**
   - Primary data source

2. **Am I running a recent version (2.0+) of Sierra?**
   - Sierra DNA was not publicly released until then

3. **Is my Sierra Desktop App bundled with Java 6 or Java 8?**
   - May be able to run Metabase ***right now*** if you've upgraded to use Java 8
   - Requires a more recent version of Sierra (~2.4+)

4. **Are we also using Teleforms or Google Analytics?**
   - Secondary data source(s)
   - *Deprecated:* Teleforms impending replacement by i-Tiva



Start with Sierra

1. Create a new Sierra user
   - Service account, only for Metabase
2. Assign application
   - Sierra SQL Access
3. Save user

Remember username and password

- Password manager, etc.

Metabase

Download

http://www.metabase.com/start/jar.html

Create Shortcut

If using Sierra's Java Runtime

C:\Sierra Desktop App\jre\1.8.##_##\bin\

Then

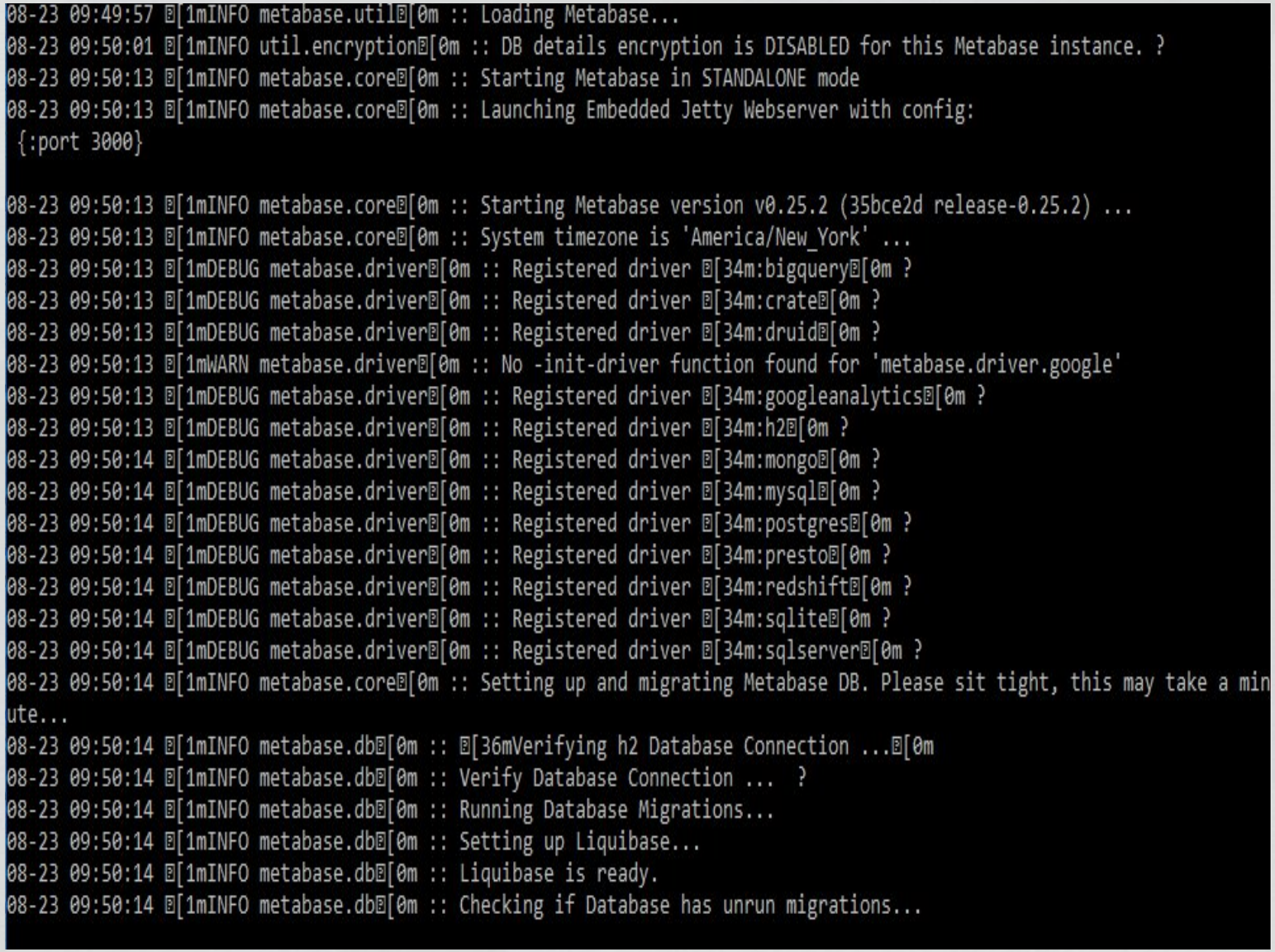
java.exe -jar "C:\path\to\metabase.jar"

Run Shortcut

Should see something like this picture:

Then open web browser to

http://localhost:3000/



Metabase

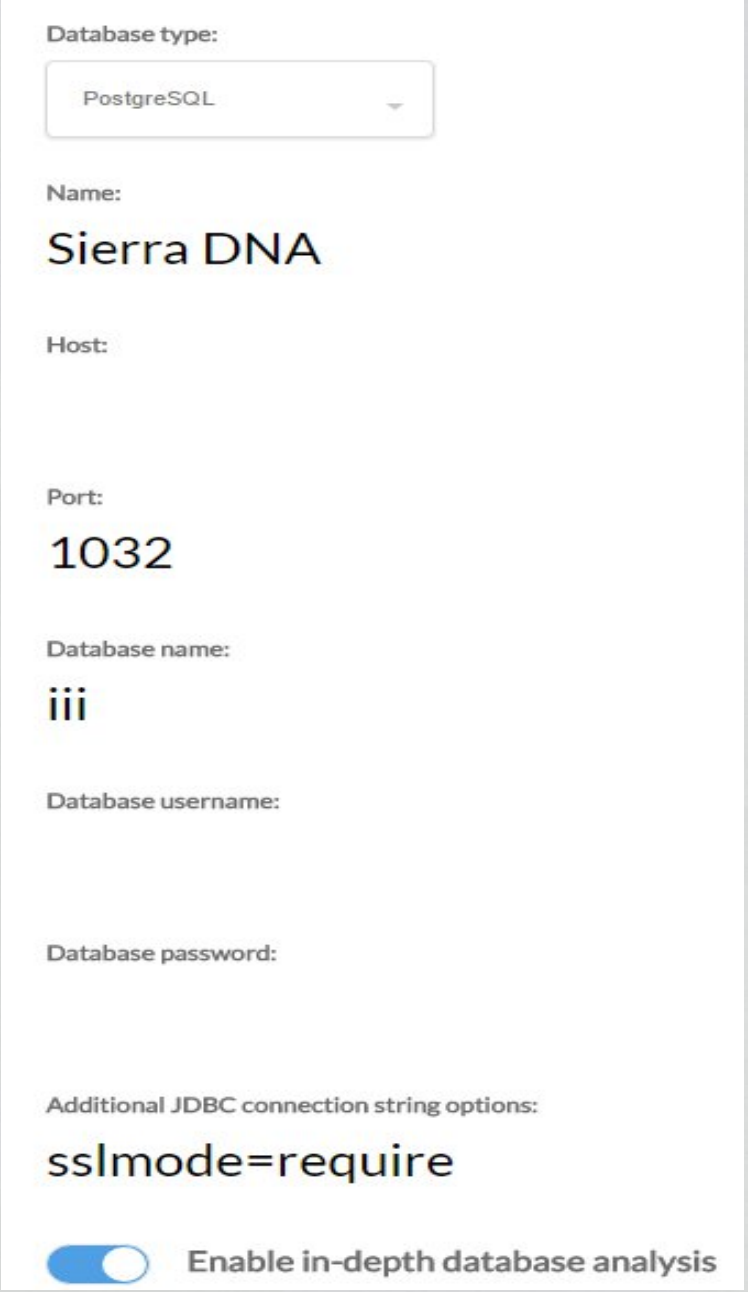
Create Admin User

- In charge of adding new users

Connect Database

- Use the picture as a guide

Finish Setup



Email Setup

Create Gmail Account

- Separate account for storage purposes

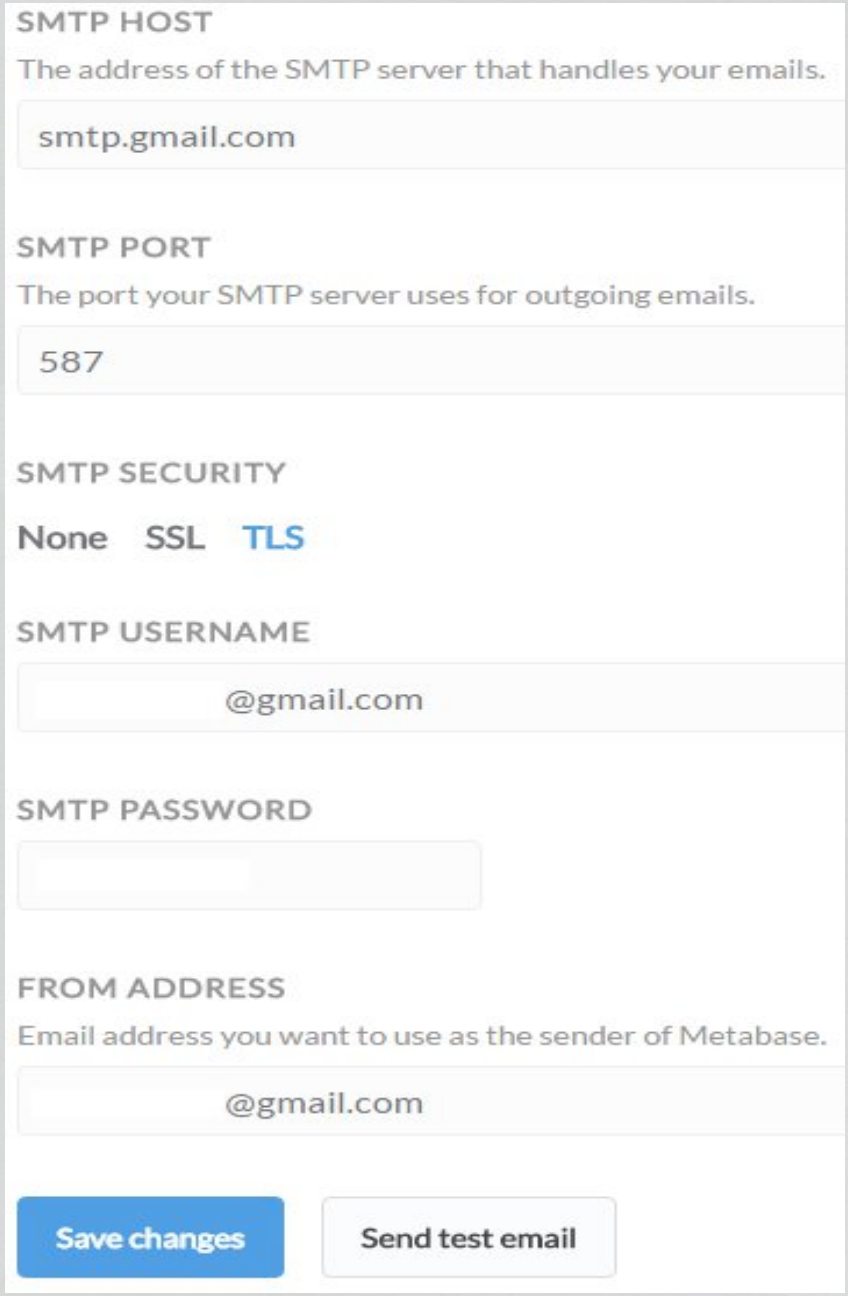
Admin Panel in Metabase

- Email settings, fill in the blanks

Test Email & Save

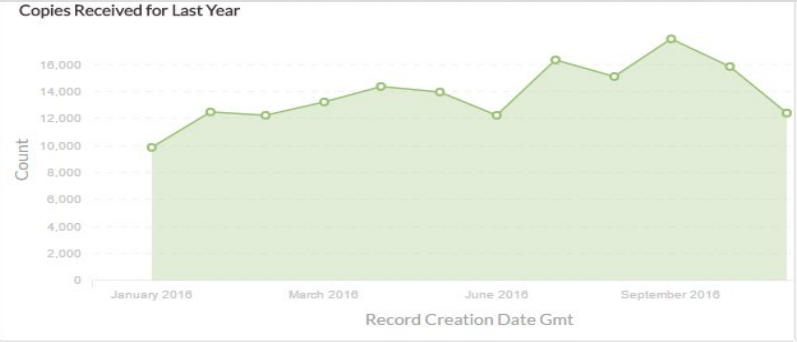
- If it fails, ensure info is correct

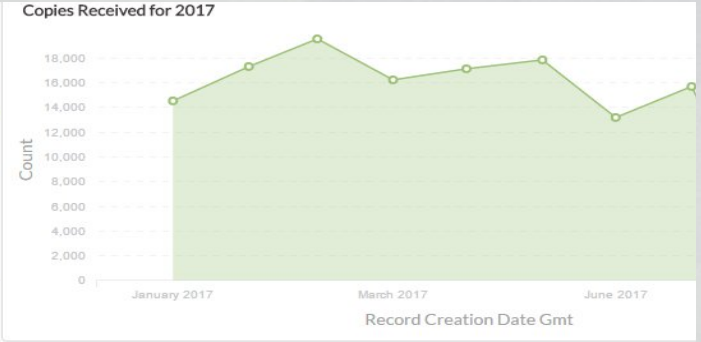
- If it still fails, try this link and enable:
https://myaccount.google.com/lesssecureapps

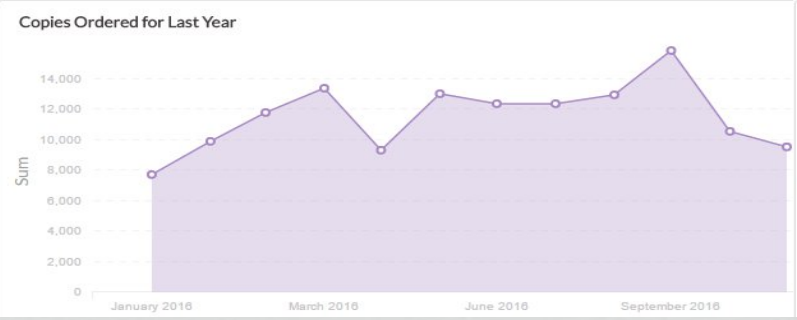


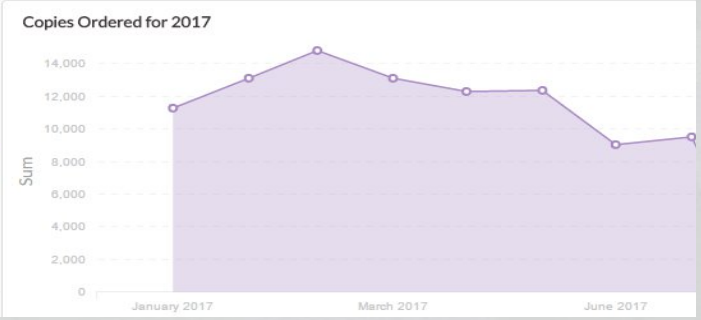
Visualizing Workflows

New Question









- **Collection Development**
  - Copies Ordered, Paid, and Received
  - Hold Ratios & Biggest Hold List by Branch
  - Allocation by Item Location
- **Technical Services**
  - Items Cataloged
  - Bib Records Cataloged
  - Purchases by Branch & Agency
- **Marketing / Search / Discovery**
  - Holds Placed via WebPAC / Encore
  - # of Logins by Segment
  - # of Holds Placed by Segment
  - # of Searches by Segment
- **Circulation**
  - Local & Interlibrary Loan Checkouts
  - INN-Reach Checkouts
  - # of Items Due Today
  - Renewals

New Question

DATA: Order Record Received

FILTERED BY: Received Date Gmt — This Year

VIEW: Sum of Copy Num

GROUPED BY: Received Date Gmt: Month and Location Code

Copies Received by Location Code in 2017

Location Code	January 2017	February 2017	March 2017	April 2017	May 2017	June 2017	July 2017	August 2017
elaf	28	34	36	39	27	47	32	46
elafc	15	18	8	-	-	-	5	-
elafi	3	3	9	5	5	6	4	7
elafl	6	4	7	5	11	18	6	10
elafm	7	6	11	7	13	6	8	12
elafp	-	1	1	-	-	-	-	1
elafr	35	34	39	36	36	35	34	6
elafs	5	5	3	3	2	7	3	2
elafw	-	1	1	-	1	-	1	-
elah	3	2	7	10	5	8	3	8
elan	50	65	71	50	69	55	38	31

Rows 1-11 of 1048

Pivot Chart

1. **Data**

   Sierra DNA

   ⇨ Sierra View

   ⇨ Order Record Received

2. **Filtered By**

   - Received Date Gmt (click "Previous")
     - Click "Current"
     - Set to "Year"

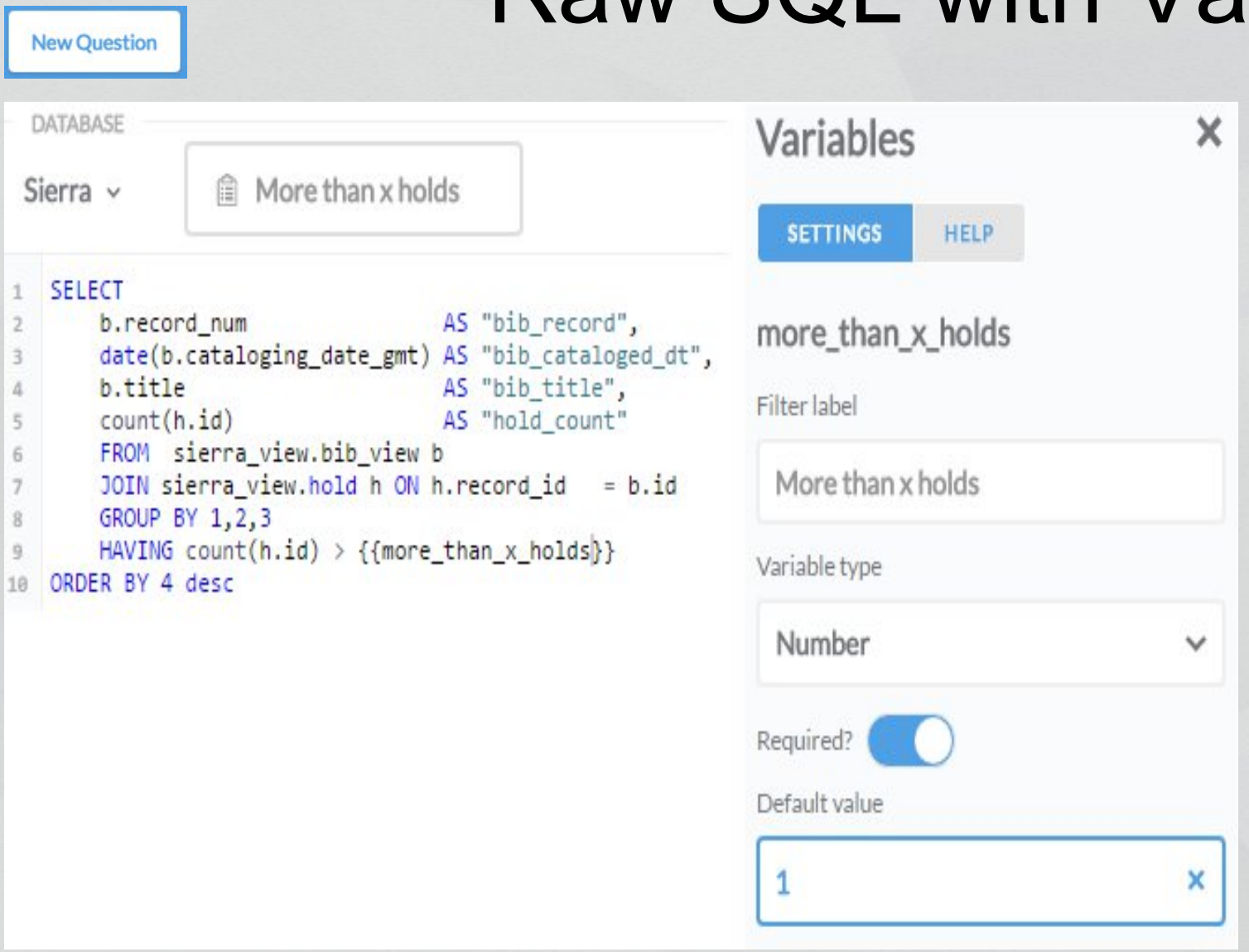
3. **View**

   - Sum of...
     - Copy Num

4. **Grouped By**

   - Received Date Gmt (click "by Day")
     - Set to "by Month"

Raw SQL with Variable(s)



1. **Open Editor**

2. **Write (or Paste) SQL Query**

3. **Substitute Variables: {{variable_name}}**
   - Valid Types:
     - Number
     - Text
     - Date
     - Field Filter

4. **Optional Criteria: [[criteria]]**

   Example

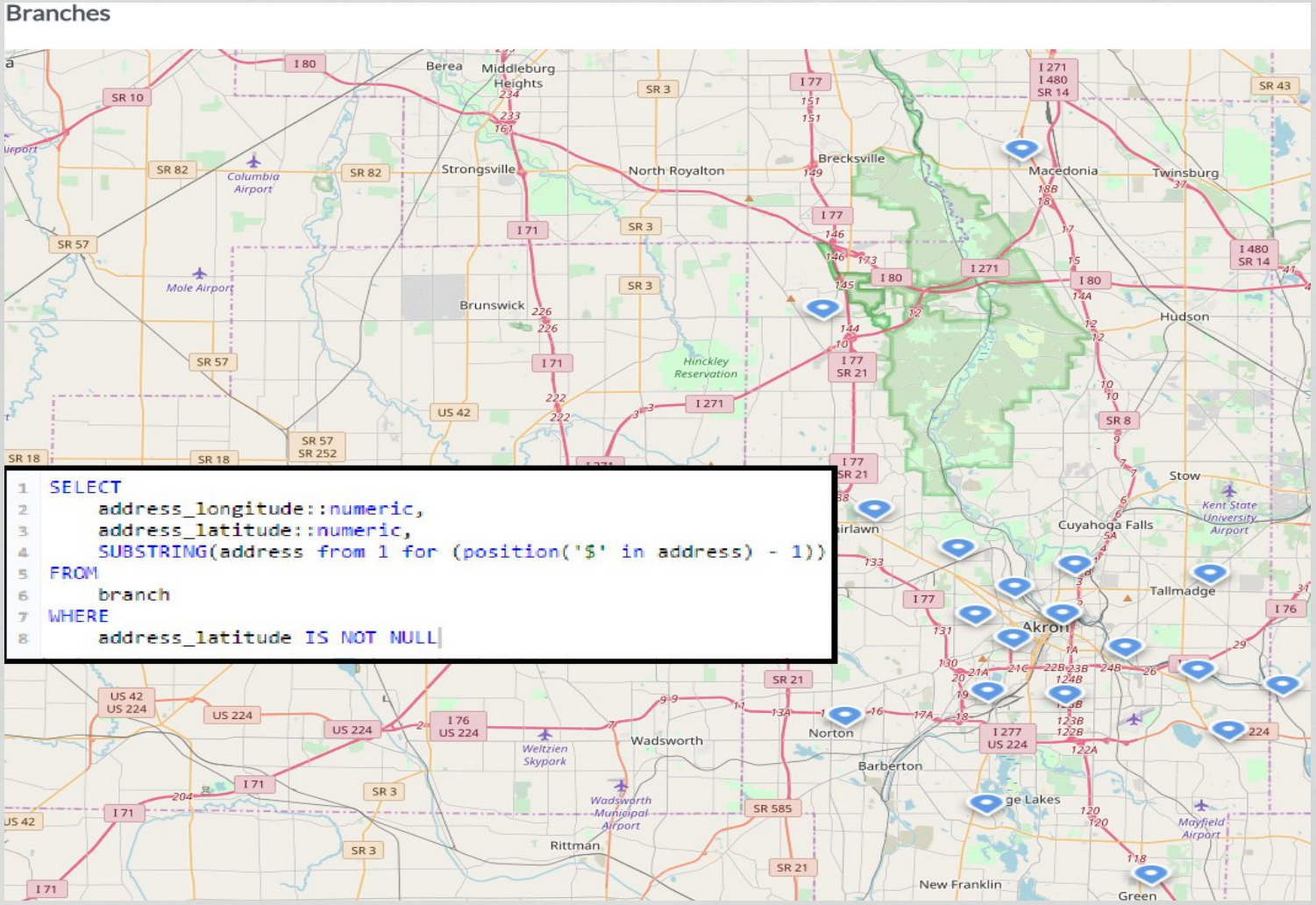
   [[HAVING count(h.id) > {{more_than_x_holds}}]]

5. **Share URL**

   Example

Maps

New Question



1. **Data**
   - Sierra DNA
     - ⇨ Sierra View
       - ⇨ Zip Info (Patrons)
       - ⇨ Branch (Branch)

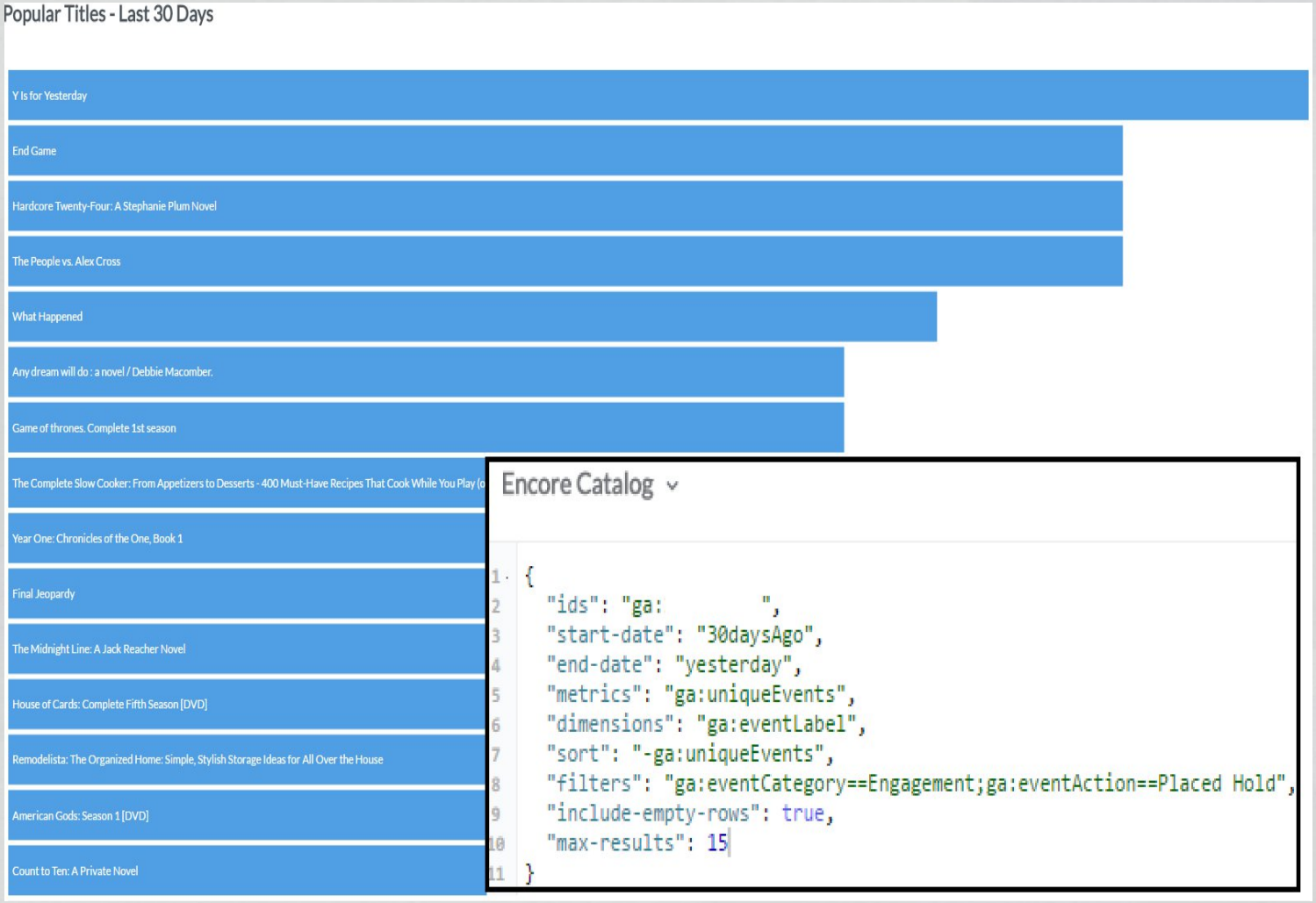
2. **Filtered By**
   - Address Latitude
     - Is Not Empty

3. **View**
   - Address Longitude (as numeric)
   - Address Latitude (as numeric)

4. **Notes**
   - Two Types: Regional Map & Pin Map
   - Pin Map requires coordinates
   - Regional Map requires a State or Province

Google Analytics

New Question



1. **Data**

   Google Analytics

   ⇨ Encore Catalog

2. **Filtered By**
   - Event Category
     - Engagement
   - Event Action
     - Placed Hold
   - Date
     - 30 Days Ago - Yesterday

3. **View**
   - Top 15 Records by Unique Events
   - Event Label (Record Title)
   - Unique Events

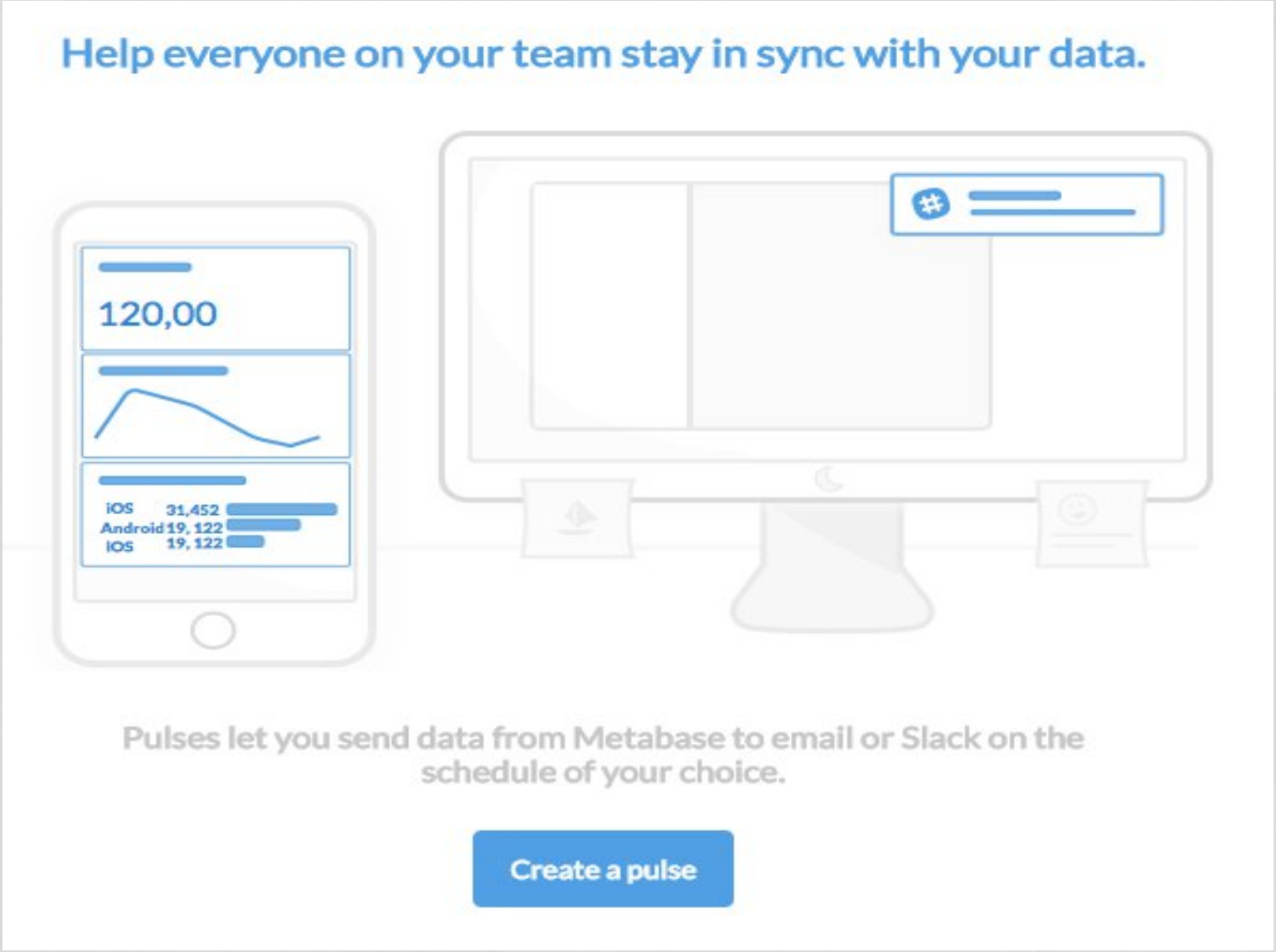
Metabase

Misc Features

- LDAP / Windows Single Sign-On
- Dashboard-Level Filters
- Graph Customizations
- Built-in Knowledgebase
- Connect to Ticket Systems, etc.
- AI Chat Bot using Slack

Issues We've Had

- Pulses (email) can time-out
- Public dashboards prompt login
- Unintuitive Permissions
- Google Analytics missing fields
- LDAP Group Mapping



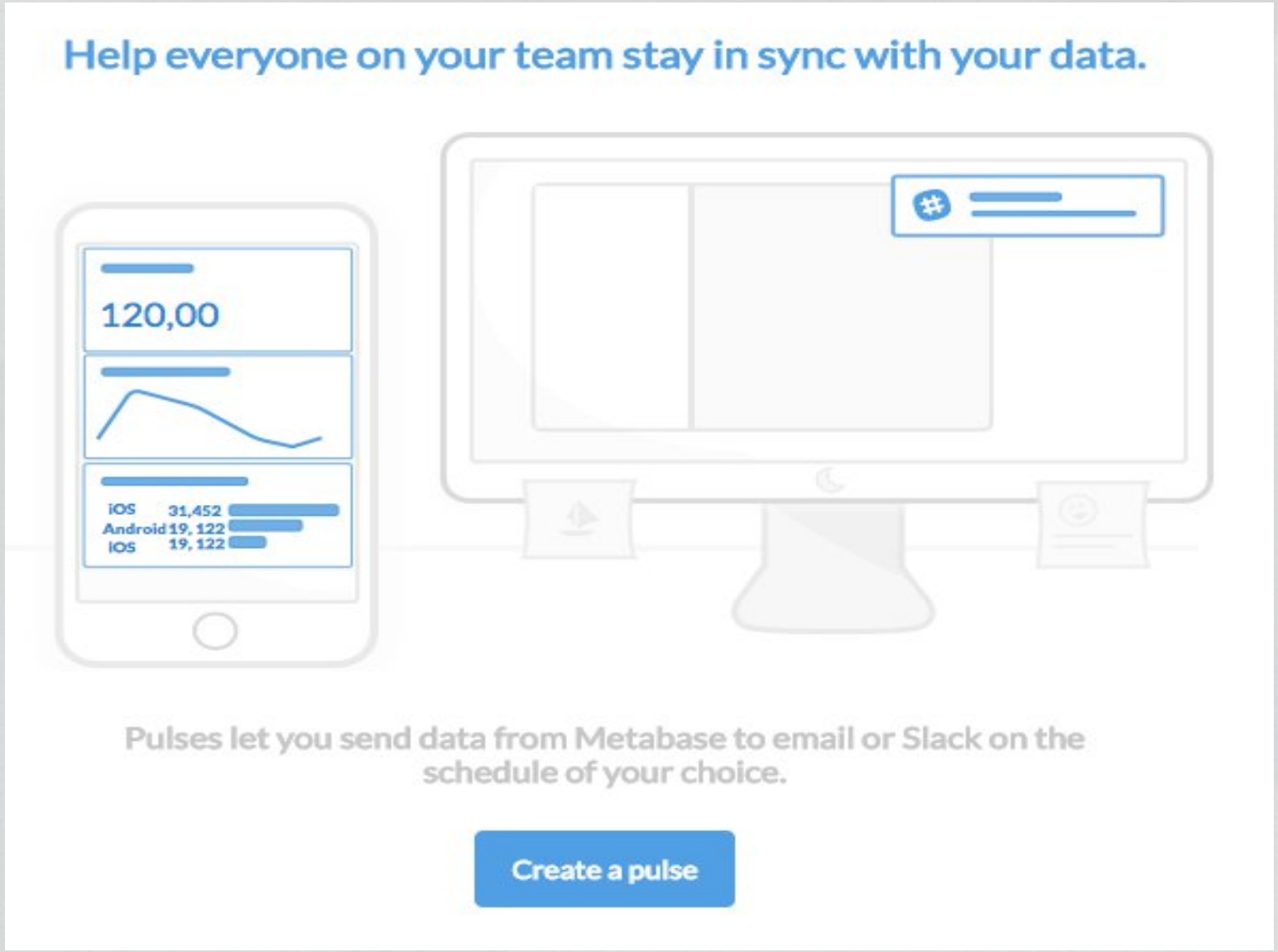
Why Use It?

Strategic

- Historical Archival and Trends
- Bottlenecks and Waste
- Impact Study
- Business Case Development
- More Accountability

Tactical

- INN-Reach vs Local Workload
- Registration Drives
- Agency / Branch Specific Details
- Programs based on Audience
- Floating Collections Evaluation



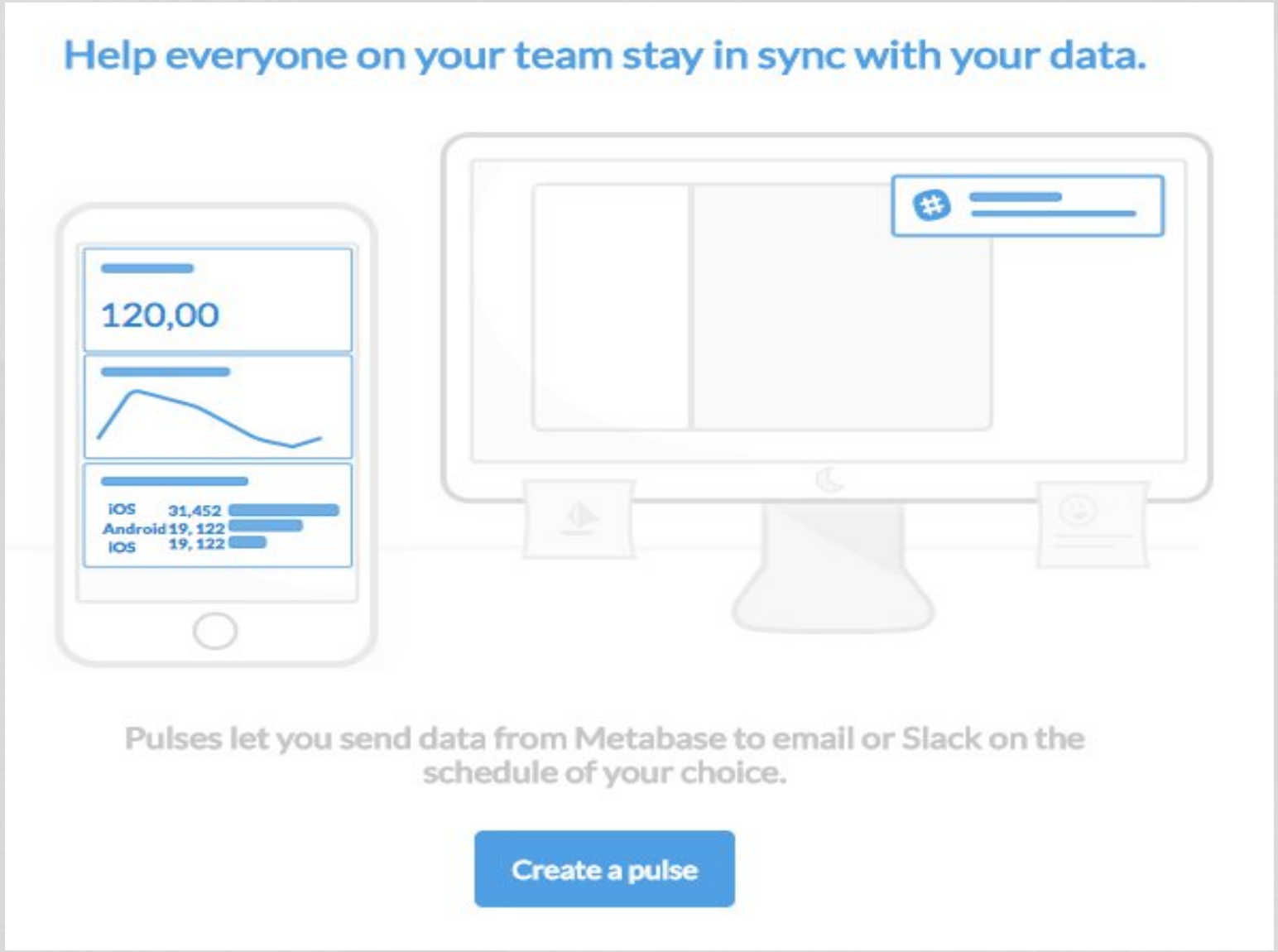
Tips & Tricks

Library-Wide

- Enable Sharing & Embedding
  - Use with Intranets
- Use LDAP if in production
  - Works for Linux servers too!
- PostgreSQL if in production
  - Back up those queries!

User

- Finished queries as starting point for new queries
- Overwriting is non-destructive
  - Built-in Revision History
- Theme your dashboards
- Use Data Reference, especially its "Getting Started" page



Online Resources

Metabase

- http://metabase.com/docs/latest
- http://0.0.0.0:3000/reference/guide (change 0.0.0.0 to host, click "Getting Started")

SQL

- http://w3resource.com/PostgreSQL/tutorial.php

Google Analytics

- https://developers.google.com/analytics/devguides/reporting/core/v3/common-queries
- https://console.developers.google.com/

Sierra API

What can it do for me?

- Repetitive tasks
- Scheduled tasks

Heads Up!

- Relies on framework I developed & released for last year's talk for integration
- You can find that framework at http://bit.ly/ascpl-2016-iug

Use Case

- Proactively repair "bounce" emails, improve eNewsletter outreach, repair email notices

Concepts

Can I Use This Too?

- Not meant for production use - yet!
- If you're interested in contributing, whether testing or bugfixing, feel free to get in touch

Technical Goals

- Go behind-the-scenes on what a Sierra API application looks like
- Maintain open access while making quality control less error-prone
- Share quality control datasets between libraries - easy and privacy-protected

Initial Questions to Ask

What version of Sierra are we on?

- Sierra API is available for Sierra ~2.4+

Do we have a capable server?

- Requirements:
  - PHP & php-zip, php-pgsql, php-curl, php-xml modules
  - Composer (http://getcomposer.org/)

Can we write our own?

- Sure! If you use PHP, our framework is easy to plug-in queries and logic
- If you're using a different language, use the process breakdown on the "Process" slide

Process / Algorithm

- **Add domain to "known good" list if MX record exists, do not prompt for change**

- **If no MX record detected:**
  1. Auto-correct mistakes that hamper analysis (e.g. comma to period and 2 to @)
  2. Run similarity checks against "known good" list
  3. Auto-correct to an exact match, if the host is already in "known good"

- **If similarity check doesn't match any host exactly:**
  1. Return top two matches, along with "other" to enter your best guess & "keep" to skip record

- **If email was (auto-)corrected:**
  1. Add user ID and corrected email address to "changed patrons" list

- **Finally:**
  1. Cycle through "changed patrons" list
  2. PUT patron/[id] - no brackets - with the 'emails' field populated by list (array) of corrected email address(es)
  3. (Archive &) Clear "changed patrons" list once all patrons are finished

Working Repairs

- test@example,ocm, test@exa...
- test@example,ocm
- test2example.com
- test2#example.com
- test example.com
- test@example .com

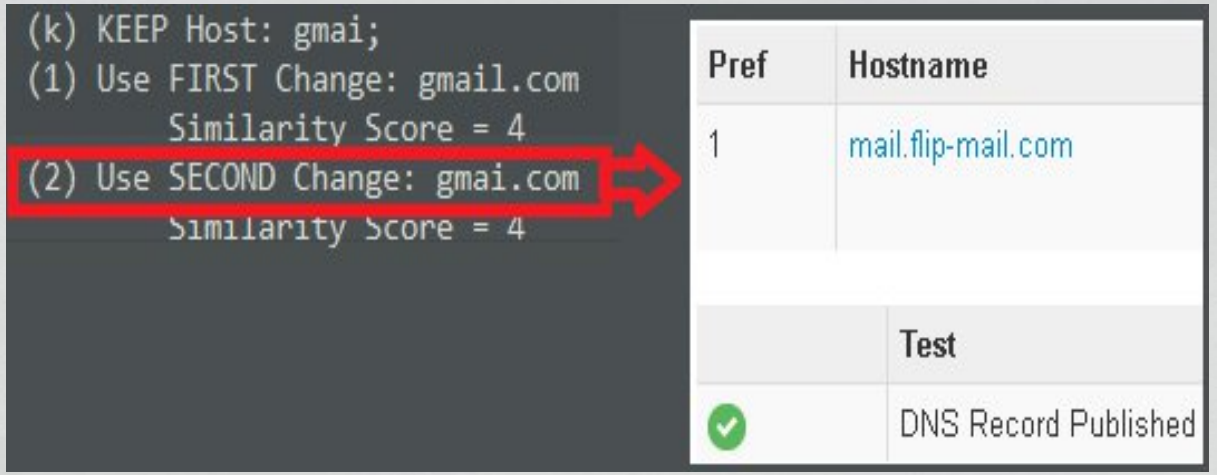
Output

- Pending Changes
- Common misspellings
  - Quality assurance dataset
  - Long-term memory
- Known hosts
  - Quality assurance dataset
  - Short-term memory

```
[1506053530] [DEBUG] Splitting multiple emails
[1506053530] [DEBUG] Finished splitting up multiple emails
[1506053530] [DEBUG] AutoCorrect: @ sign
[1506053530] [DEBUG] Autocorrecting: [/gmail.com] to [@gmail.com] in following address...
[1506053530] [DEBUG] *********@gmail.com
[1506053530] [DEBUG] Autocorrecting: [2gmail.com] to [@gmail.com] in following address...
[1506053530] [DEBUG] ********@gmail.com
[1506053530] [DEBUG] Autocorrecting: [2yahoo.com] to [@yahoo.com] in following address...
[1506053530] [DEBUG] ***********@yahoo.com
[1506053530] [DEBUG] Autocorrecting: [#starkstate.net] to [@starkstate.net] in following address...
[1506053530] [DEBUG] **********@starkstate.net
[1506053530] [DEBUG] Finished AutoCorrect: @ sign
[1506053530] [DEBUG] Autocorrecting: [ZOOM.INTERNET.NET] to [zoominternet.net] in following email:
[1506053530] [DEBUG] **********@ZOOM.INTERNET.NET
[1506053530] [DEBUG] Autocorrecting: [gmail.cm] to [gmail.com] in following email:
[1506053530] [DEBUG] **********@gmail.cm
[1506053530] [DEBUG] Autocorrecting: [GMAIL.CM] to [gmail.com] in following email:
[1506053530] [DEBUG] **********@GMAIL.CM
[1506053530] [DEBUG] Finished AutoCorrect: Previously learned mispellings
[1506053530] [DEBUG] Normalize Domain Names
[1506053530] [DEBUG] Finished Normalizing Domain Names
[1506053530] [DEBUG] Calculating number of new domain names
[1506053530] [DEBUG] Finished Calculating number of new domain names
[1506053530] [DEBUG] Running DNS Validation
[1506053535] [DEBUG] Finished DNS Validation
[1506053535] [DEBUG] AutoCorrect: @ sign
[1506053535] [DEBUG] Finished AutoCorrect: @ sign
[1506053535] [DEBUG] Filtering emails with no @
[1506053535] [DEBUG] Finished Filtering emails with no @
[1506053535] [DEBUG] Calculating number of new domain names
[1506053535] [DEBUG] Finished Calculating number of new domain names
[1506053535] [DEBUG] AutoCorrect: Remove illegal characters
[1506053535] [DEBUG] Finished AutoCorrect: Remove illegal characters
```

Quirks

- Not all MX records are good
- Popular hosts learn fastest
- Internal DNS maybe filtered
- Big dataset is more thorough
- Defunct companies



Learning API

- Obtain "API Key" from Admin Corner
- Use config.json template
- Use developers.iii.com to find API endpoint
- Fill in config.json key and secret
- Use the sandbox (Swagger) to figure out fields and values for desired outcome
- Try to replicate GET queries in scripts
  - *Remember* GET queries do not modify records
  - Everything else will modify records

```
//--- Use access token to query API ---//
$response = $client->query(
    $access->token(),
    'GET',
    'bibs',
    array(
        'query' => array(
            'deleted' => 'false',
            'suppressed' => 'false',
            'createdDate' => '[2013-07-30T19:20:28Z,]',
            'fields' => 'title,author,publishYear'
        )
    )
    /** For POST / PUT
    array(
        'json' => array(
            'deleted' => 'false',
            'suppressed' => 'false',
            'createdDate' => '[2013-07-30T19:20:28Z,]',
            'fields' => 'title,author,publishYear'
        )
    )*/
);
```

Online Resources

Sierra API

- https://developer.iii.com/
  - Requires registration
  - Fantastic for understanding these modules

- http://bit.ly/ascpl-2016-iug
  - Configurable swiss-army tool to get running with SQL and API quickly (if using PHP)

Proof-of-Concept Interface

```
[1506053535] [DEBUG] AutoCorrect: Remove illegal characters
[1506053535] [DEBUG] Finished AutoCorrect: Remove illegal characters
**********************************************************************
Supervision ready! Please press enter to continue...
**********************************************************************
**********************************************************************
Reminder:
**********************************************************************
Selections you make are not permanent. pending-changes.json will be generated
and may be reviewed (along with mispellings and known-hosts history)
before making your selections permanent.


**********************************************************************
INSTRUCTIONS:
**********************************************************************
The following emails require training to fix. You will be prompted until
one of these options are selected. You can quit at any time with CTRL+C


(k) KEEP Host
        -- Ignore the host and DO NOT attempt to learn it
(1) Use FIRST Change / (2) Use SECOND Change
        -- Displays similar from previously learned hosts (see known-hosts.json)
        -- This will remember the typo and fix it in the future too!
(o) OTHER
        -- Use this option on hosts that require research to find the correct
           spelling for the host.
        -- Alternatively, if the domain is trustworthy, you can re-type carefully
           here.
        -- This will remember the typo and fix it in the future too!
**********************************************************************
Supervision ready! Please press enter to continue...
**********************************************************************
```