

DIY Search Tools

Practical Applications of  and 

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

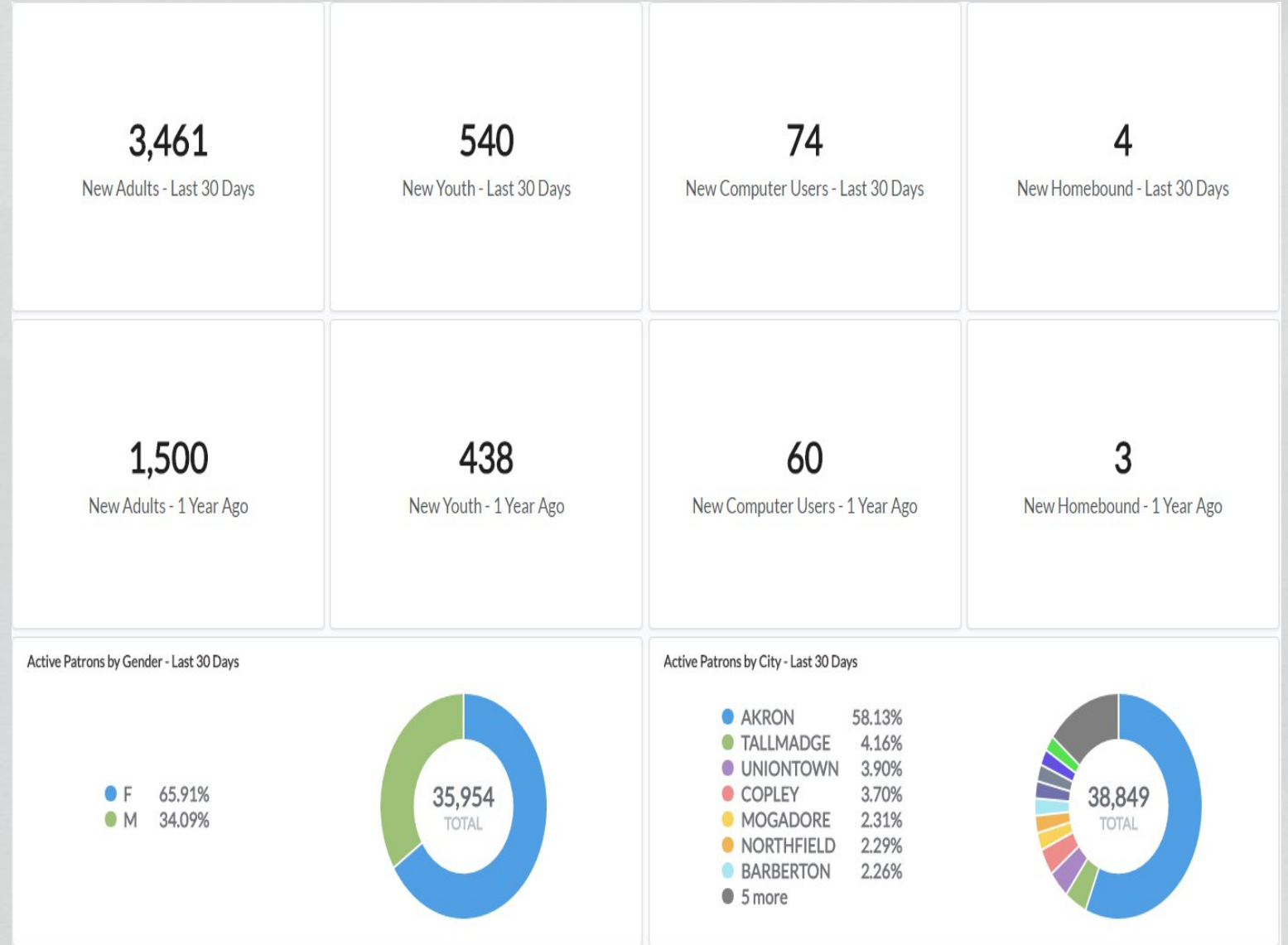
Metabase

Dashboards

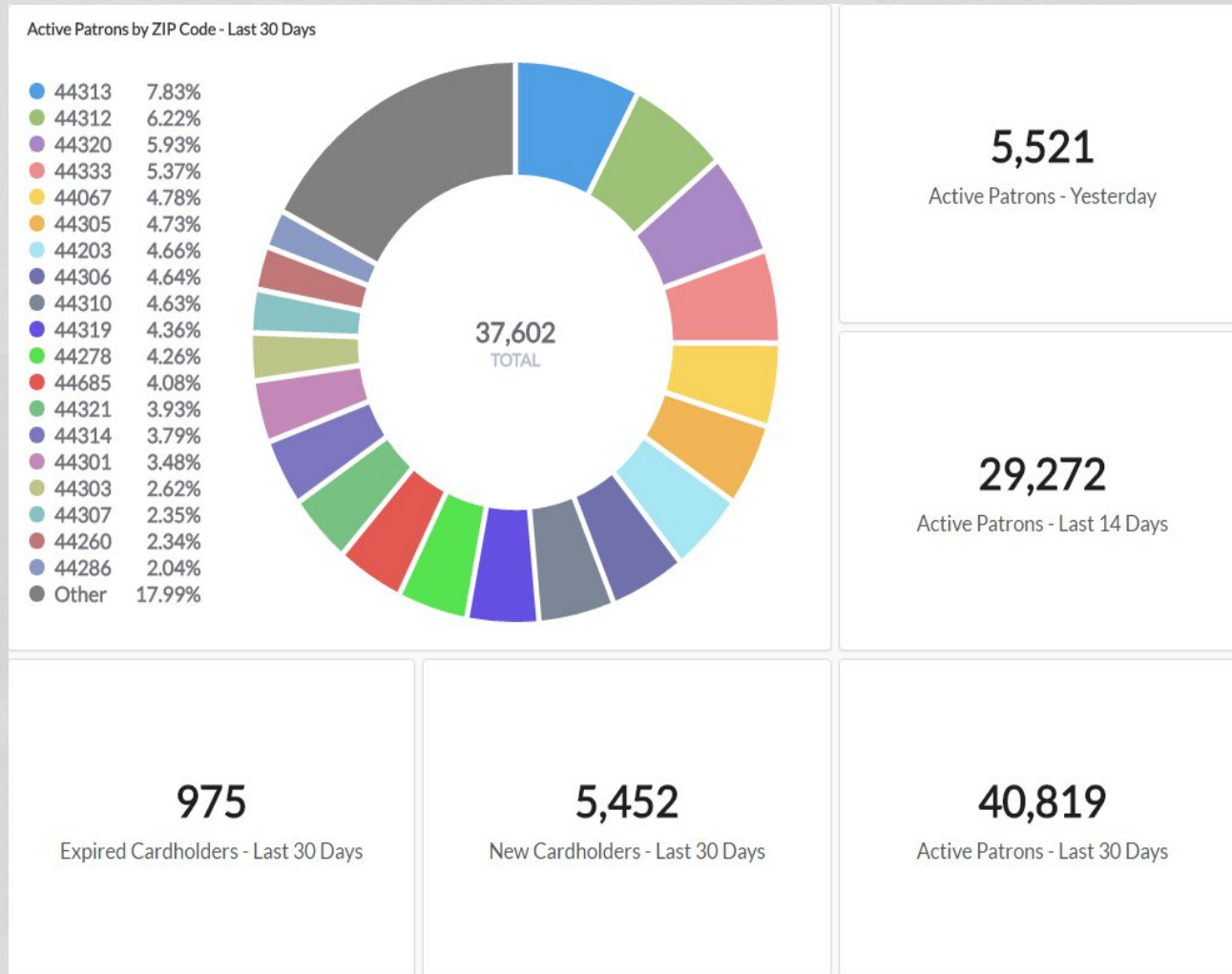
- Acquisitions / Cataloging metrics
- Branch Audience reports
- Cardholder metrics
- Circulation metrics
- Demographic metrics
- Teleforms metrics

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

[Basic Info](#) **Applications** [Permissions](#) [Workflows](#) [Options](#) [Settings](#) [Preferences](#)

Assigned applications (1) [remove all](#)

Sierra SQL Access

Available applications (2) [add all](#)

Admin App

Sierra Application

[Delete user](#) [SAVE](#) [Discard changes](#)

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/>

```
08-23 09:49:57 [1mINFO metabase.util[0m :: Loading Metabase...
08-23 09:50:01 [1mINFO util.encryption[0m :: DB details encryption is DISABLED for this Metabase instance. ?
08-23 09:50:13 [1mINFO metabase.core[0m :: Starting Metabase in STANDALONE mode
08-23 09:50:13 [1mINFO metabase.core[0m :: Launching Embedded Jetty Webserver with config:
{:port 3000}

08-23 09:50:13 [1mINFO metabase.core[0m :: Starting Metabase version v0.25.2 (35bce2d release-0.25.2) ...
08-23 09:50:13 [1mINFO metabase.core[0m :: System timezone is 'America/New_York' ...
08-23 09:50:13 [1mDEBUG metabase.driver[0m :: Registered driver [34m:bigquery[0m ?
08-23 09:50:13 [1mDEBUG metabase.driver[0m :: Registered driver [34m:crate[0m ?
08-23 09:50:13 [1mDEBUG metabase.driver[0m :: Registered driver [34m:druid[0m ?
08-23 09:50:13 [1mWARN metabase.driver[0m :: No -init-driver function found for 'metabase.driver.google'
08-23 09:50:13 [1mDEBUG metabase.driver[0m :: Registered driver [34m:googleanalytics[0m ?
08-23 09:50:13 [1mDEBUG metabase.driver[0m :: Registered driver [34m:h2[0m ?
08-23 09:50:14 [1mDEBUG metabase.driver[0m :: Registered driver [34m:mongo[0m ?
08-23 09:50:14 [1mDEBUG metabase.driver[0m :: Registered driver [34m:mysql[0m ?
08-23 09:50:14 [1mDEBUG metabase.driver[0m :: Registered driver [34m:postgres[0m ?
08-23 09:50:14 [1mDEBUG metabase.driver[0m :: Registered driver [34m:presto[0m ?
08-23 09:50:14 [1mDEBUG metabase.driver[0m :: Registered driver [34m:redshift[0m ?
08-23 09:50:14 [1mDEBUG metabase.driver[0m :: Registered driver [34m:sqlite[0m ?
08-23 09:50:14 [1mDEBUG metabase.driver[0m :: Registered driver [34m:sqlserver[0m ?
08-23 09:50:14 [1mINFO metabase.core[0m :: Setting up and migrating Metabase DB. Please sit tight, this may take a minute...
08-23 09:50:14 [1mINFO metabase.db[0m :: [36mVerifying h2 Database Connection ...[0m
08-23 09:50:14 [1mINFO metabase.db[0m :: Verify Database Connection ... ?
08-23 09:50:14 [1mINFO metabase.db[0m :: Running Database Migrations...
08-23 09:50:14 [1mINFO metabase.db[0m :: Setting up Liquibase...
08-23 09:50:14 [1mINFO metabase.db[0m :: Liquibase is ready.
08-23 09:50:14 [1mINFO metabase.db[0m :: Checking if Database has unrun migrations...
```


Metabase

Create Admin User

- In charge of adding new users

Connect Database

- Use the picture as a guide

Finish Setup

Database type:
PostgreSQL

Name:
Sierra DNA

Host:

Port:
1032

Database name:
iii

Database username:

Database password:

Additional JDBC connection string options:
sslmode=require

☒ Enable in-depth database analysis

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>

SMTP HOST
The address of the SMTP server that handles your emails.

SMTP PORT
The port your SMTP server uses for outgoing emails.

SMTP SECURITY
None SSL **TLS**

SMTP USERNAME

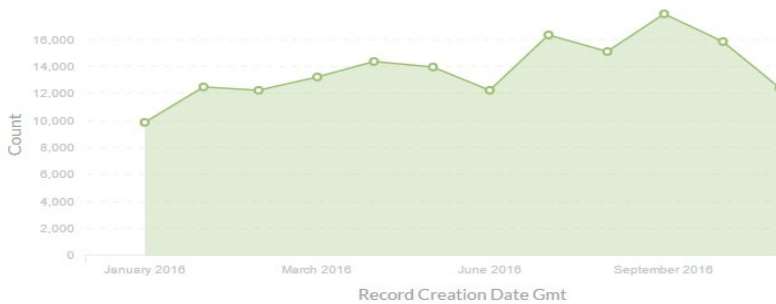
SMTP PASSWORD

FROM ADDRESS
Email address you want to use as the sender of Metabase.

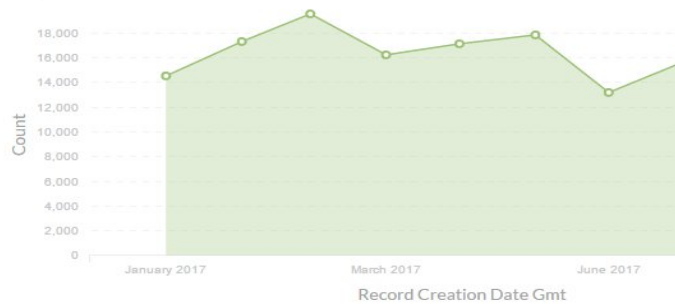
Visualizing Workflows

New Question

Copies Received for Last Year



Copies Received for 2017



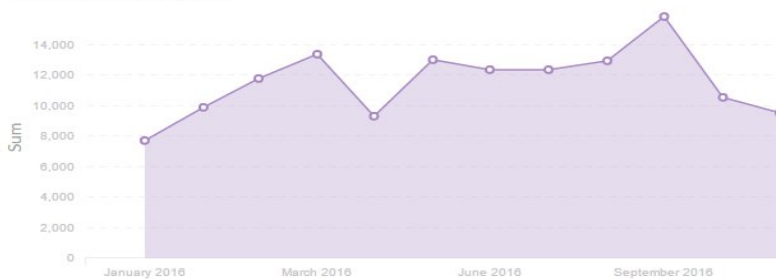
Copies Paid for Last Year



Copies Paid for 2017



Copies Ordered for Last Year



Copies Ordered for 2017



- **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

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
ela1f	28	34	36	39	27	47	32	46
ela1fc	15	18	8	-	-	-	5	-
ela1fi	3	3	9	5	5	6	4	7
ela1fl	6	4	7	5	11	18	6	10
ela1fm	7	6	11	7	13	6	8	12
ela1fp	-	1	1	-	-	-	-	1
ela1fr	35	34	39	36	36	35	34	6
ela1fs	5	5	3	3	2	7	3	2
ela1fw	-	1	1	-	1	-	1	-
ela1h	3	2	7	10	5	8	3	8
ela1n	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"


by Month

Raw SQL with Variable(s)


New Question

DATABASE

Sierra 

 More than x holds

```
1 SELECT
2   b.record_num           AS "bib_record",
3   date(b.cataloging_date_gmt) AS "bib_cataloged_dt",
4   b.title                 AS "bib_title",
5   count(h.id)             AS "hold_count"
6 FROM   sierra_view.bib_view b
7 JOIN   sierra_view.hold h ON h.record_id = b.id
8 GROUP BY 1,2,3
9 HAVING count(h.id) > {{more_than_x_holds}}
10 ORDER BY 4 desc
```

Variables 

SETTINGS


HELP

more_than_x_holds

Filter label


More than x holds

Variable type

Number 

Required? ☒

Default value

1 

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}}]]

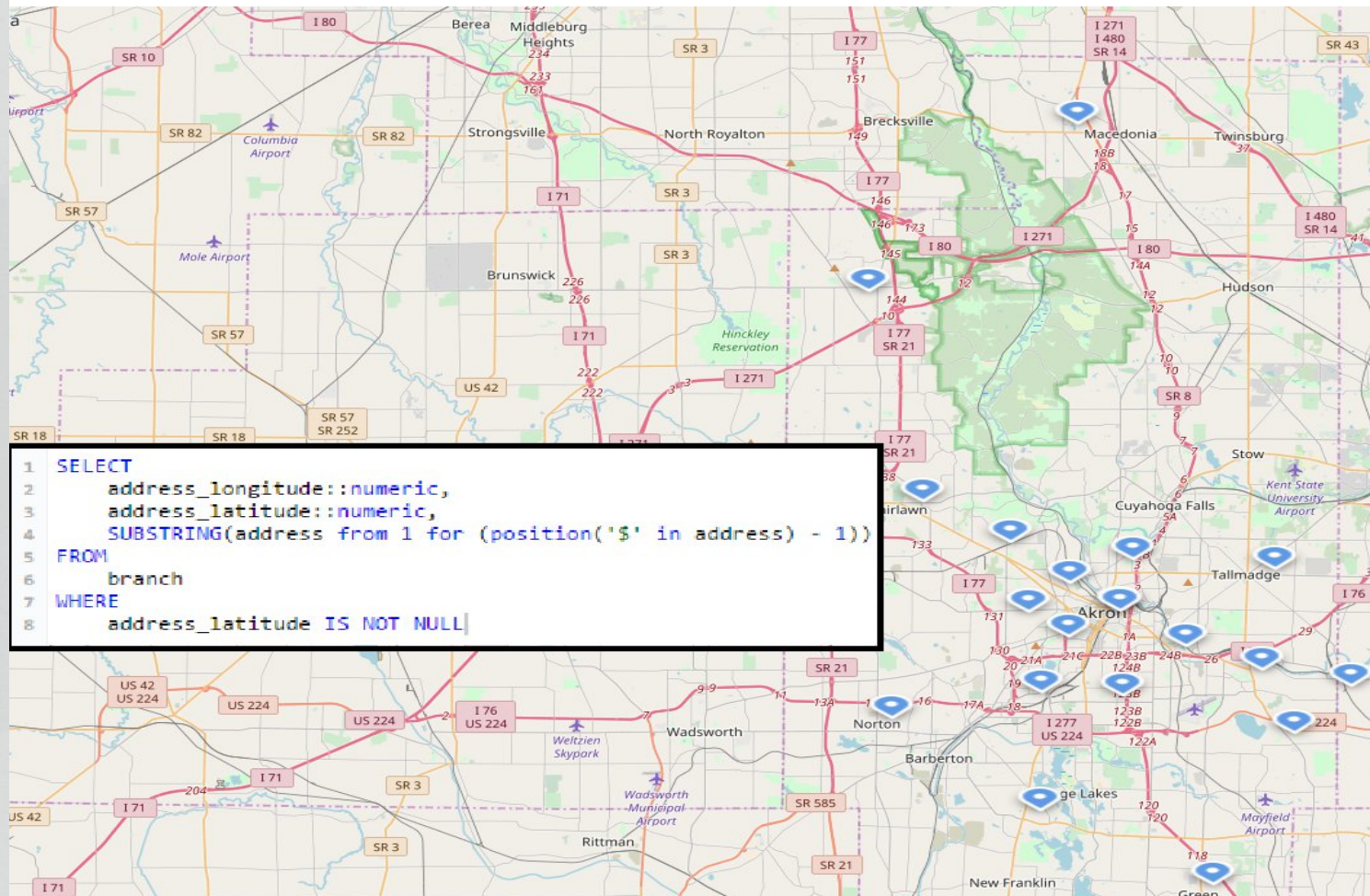
5. Share URL

Example

Maps

New Question

Branches



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

Popular Titles - Last 30 Days



Encore Catalog ▾

```
1 {  
2   "ids": "ga:",  
3   "start-date": "30daysAgo",  
4   "end-date": "yesterday",  
5   "metrics": "ga:uniqueEvents",  
6   "dimensions": "ga:eventLabel",  
7   "sort": "-ga:uniqueEvents",  
8   "filters": "ga:eventCategory==Engagement;ga:eventAction==Placed Hold",  
9   "include-empty-rows": true,  
10  "max-results": 15|  
11 }
```

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

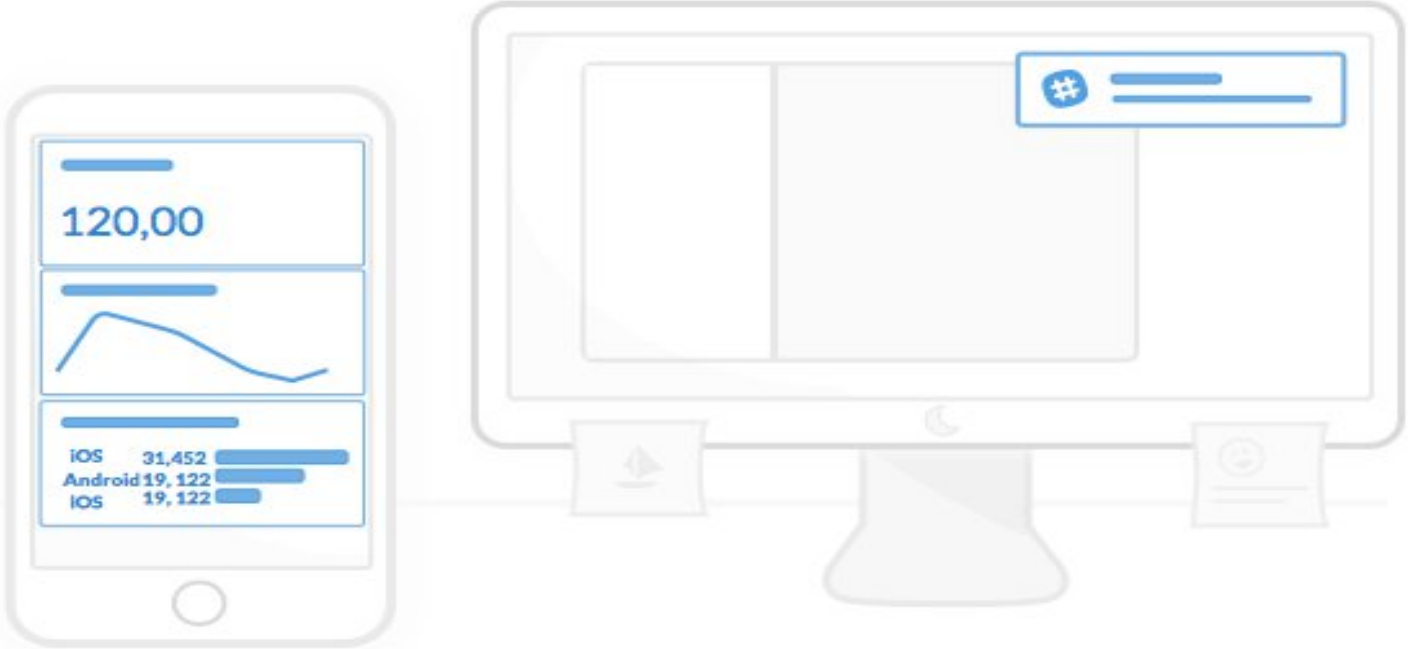
Misc Features

- 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

Help everyone on your team stay in sync with your data.



The illustration shows a mobile phone on the left and a desktop monitor on the right. The phone screen displays a dashboard with three sections: a top section with a blue bar and the number '120,00'; a middle section with a line graph showing a peak and then a decline; and a bottom section with a table of data. The table has three rows: 'IOS' with '31,452' and a long blue bar, 'Android' with '19,122' and a medium blue bar, and 'IOS' with '19,122' and a short blue bar. The desktop monitor shows a similar dashboard layout with a search bar at the top right containing a hashtag icon and a text input field. Below the monitor, there are two small icons: a sailboat and a person with a speech bubble.

Pulses let you send data from Metabase to email or Slack on the schedule of your choice.

Create a pulse

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

Help everyone on your team stay in sync with your data.



Pulses let you send data from Metabase to email or Slack on the schedule of your choice.

Create a pulse

Tips & Tricks

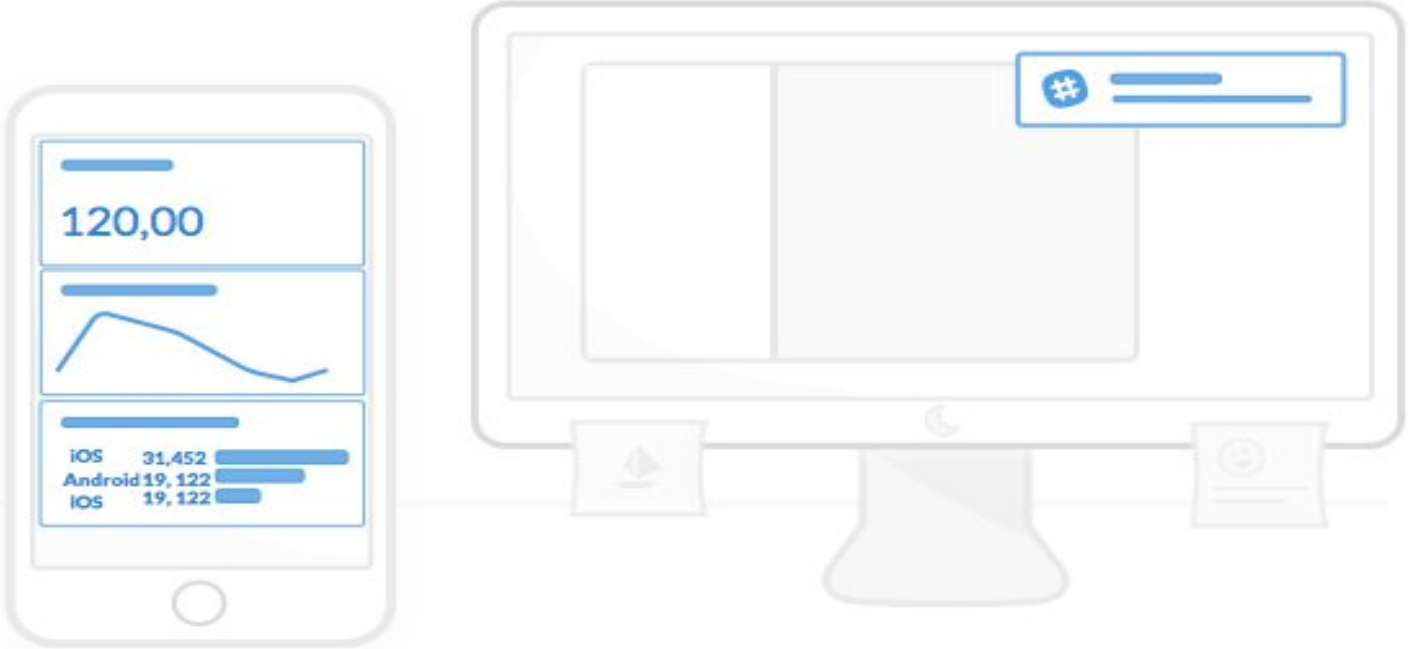
Library-Wide

- 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

Help everyone on your team stay in sync with your data.



Pulses let you send data from Metabase to email or Slack on the schedule of your choice.

Create a pulse

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 misspellings
[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

```
(k) KEEP Host: gmai;  
(1) Use FIRST Change: gmail.com  
Similarity Score = 4  
(2) Use SECOND Change: gmail.com  
Similarity Score = 4
```

Pref	Hostname
1	mail.flip-mail.com

Test
 DNS Record Published

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 misspellings 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...
*****
```