

VISUALIZING DATA: SIERRA SQL API AND PLOTLY

2018 SCIUG Conference at Chapman University
(October 23)

Presented by Seong Heon Lee, Systems & Technology Librarian
Hugh & Hazel Darling Law Library, Chapman University

SUGGESTED PROPOSAL TOPICS

- ILL Customer Support / new functionality
- Weeding
- More Polaris presentations
- MySQL workshop / SQL report-writing
- 3rd-party tools w/ ILL products
- Best practices
- User experiences (migration)
- ERMS
- Encore Duet
- Sierra APIs
- Stats / create lists / other ways of reporting
- Linked Data
- Electronic ordering
- Access/ User Services issues
- Interlibrary Loan topics (ex. ILL basics/update)
- Technology / marketing ideas and the library

FINAL PRODUCT PREVIEW

Circulation Transaction Report

WHY (1)

Sierra has Web Management Reports

- How many of you are using WMR?
- What is your user experience?
- Decision Center

WHY (2)?

Who uses what?

- Compare resource usages by libraries
- Compare by the following criteria:
 - locations
 - patron types
 - item types
 - transaction times
 - etc...

WHY (3)?

Why not?

- Plenty of data visualization tools (Open Source)
- Open Library Stacks (Direct SQL & REST API)
- Customize a data report as you wish

**EXPLORING DATA
IS
FUN!**

CONTENTS

1. Plotly (Online & Offline)
2. Sierra SQL API
3. Examples (Sierra & Plotly)
4. Web App (Sierra & Plotly)

WHAT IS PLOTLY?

- A data analytics and visualization tool
- Charts & Dashboard
- Online and Offline
- Open Source
- Support multiple languages (Python, JavaScript, R, MATLAB, etc)
- D3.js(SVG) and stack.gl (WebGL) for web graphics
- Great API Documentation & Examples

PLOTLY ONLINE

- **Plotly examples**
- **Plotly Create page**
 1. Import data: File Upload, URL Link, SQL DB Connector
 2. Select a chart type (bar/pie/scatter) and attributes (X, Y axes)
 3. Adjust chart styles
 4. multiple traces in scatter plot (**data**).

PLOTLY OFFLINE (1)

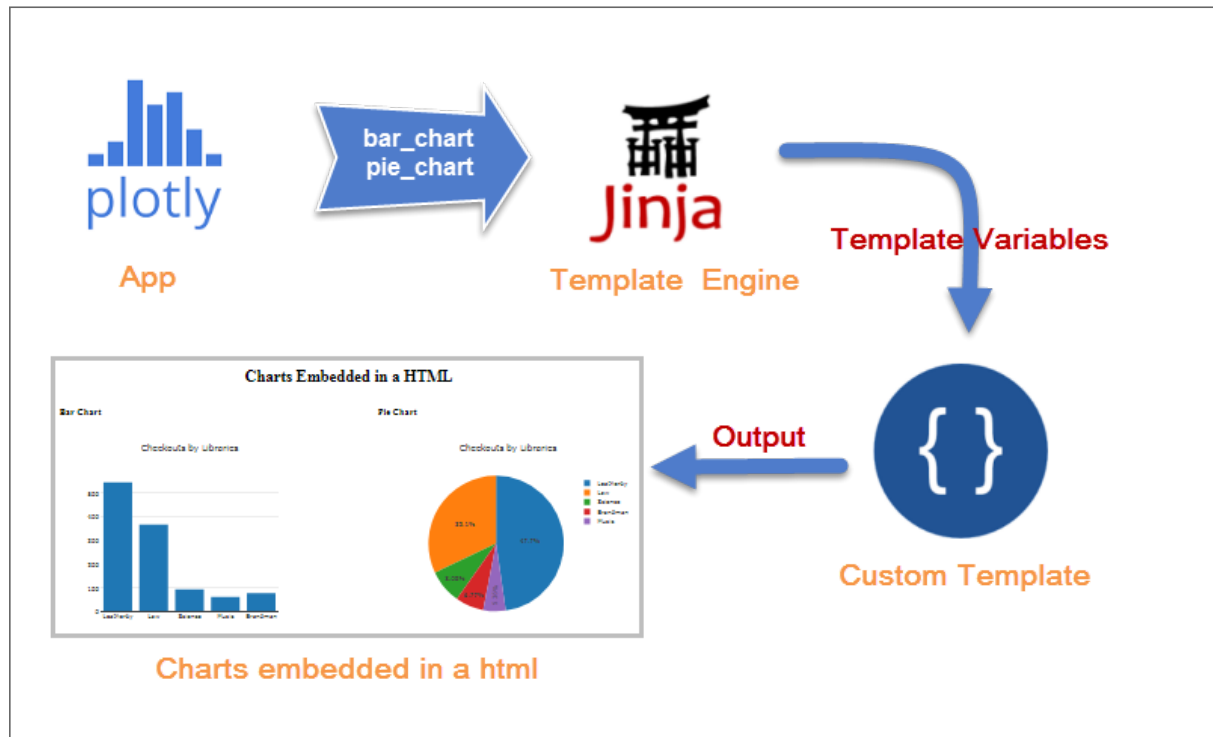
- **Python Library**
- **Plotly Offline Setup & Syntax**
 1. import plotly library and chart objects
 2. Use offline plot function
 3. feed data
 4. decorate layout
 5. pass optional arguments (filename, show_link)
 6. Interactive display modeBar

PLOTLY OFFLINE (2)

How to embed charts in a html?

- Follow the same rules
- Use div for output_type, instead of filename
- Hold a chart in a variable
- Pass the variable to a template (Jinja2 engine)
- A template file will receive multiple chart variables

PLOTLY OFFLINE (3)



PLOTLY: SUMMARY

- Online & Offline
- Various types of charts
- Data, Layout, Other attributes
- Interactive
- Offline
 - `plotly.offline.plot()` function
 - `output_type='div'`
 - Template engine (Jinja2) and a template html

SIERRA SQL API

- PostgreSQL relational database
- Sierra_view schema & 349 tables (read only)
- User permission to "Sierra SQL Access" required
- Access to:
 - bibliographic data (bib, item, holding)
 - transactional data (circ_trans, fine, patron)
 - system parameters (location, custom codes, loanrule, properties)
- By default, 5 concurrent connections per user

PGADMIN, DESTKOP SQL CLIENT

- Version 4
- **Connect** to database
- Run SQL **queries**
- Show **results** in tables
- **Creating/testing** SQL queries

SETUP PGADMIN

Server 192.168.20.220

Properties | SSL | SSH Tunnel | Advanced

Name: Sierra SQL Access

Host: 192.168.20.220

Port: 1032

Service:

Maintenance DB: iii

Username: shl

Password:

Store password: ☐

Colour:

Group: Servers

Help OK Cancel

KNOW SIERRA DATABASE

- **SierraDNA**
- Learn the structure of Sierra database
- Find a table in category and search
- Detailed view of a table
- ERD (Entity Relationship Diagram) view
- Example: Transactions > Circulation > checkout

FIND CHECKOUTS

(PGADMIN DEMO)

```
-- Search for checkedout
SELECT pv.home_library_code, pv.ptype_code, c.checkout_gmt, c.renewal_count, pv.barcode
FROM sierra_view.checkout AS c
JOIN sierra_view.patron_view AS pv ON pv.id = c.patron_record_id
WHERE NOW() - c.checkout_gmt <= interval '3 hours'
AND pv.home_library_code = 'lhome'
```

PLOTLY & SIERRA

Transactions by Patron Home Library

- Reuse an offline script
- Use Sierra data, not a dummy one
- New things:
 - Connection & query: psycpg2
 - Transform Sierra data to plotly data

SIERRA SQL OUTPUT

Output pane						
Data Output						
	patron_home_library_code text	ptype_code text	time text	op_code character	itype_code_num integer	item_location_code text
266	lhome	20	2017/09/17	r	117	lill
267	lhome	20	2017/09/17	r	117	lill
268	lhome	20	2017/09/17	r	117	lill
269	lhome	20	2017/09/17	r	117	lill
270	lhome	20	2017/09/17	r	117	lill
271	lhome	20	2017/09/17	r	117	lill
272	chome	0	2017/09/17	i	16	cresm
273	lhome	20	2017/09/17	r	117	lill
274	chome	0	2017/09/17	o	16	cresm
275	chome	4	2017/09/17	i	16	cresm
276	chome	0	2017/09/17	i	16	cresm
277	chome	0	2017/09/17	o	16	cresm
278	chome	0	2017/09/17	r	16	cresm
279	chome	28	2017/09/17	i	16	cresm
280	chome	0	2017/09/17	o	16	cresm
281	chome	0	2017/09/17	r	16	cresm
282	chome	0	2017/09/17	i	16	cresm
283	chome	0	2017/09/17	o	1	cresh

PLOTLY X Y AXES

patron_home_library_code
lhome
lhome
lhome
lhome
lhome
lhome
lhome
chome
lhome
chome
chome
chome
chome
chome
chome
chome
chome
chome
chome
chome
chome

[['cbran', 'chome', 'lhome'], [32, 5429, 880]]

X axis

Y axis

```
plotly.offline.plot({
  "data": [Bar(x=home_libraries[0], y=home_libraries[1])],
  "layout": Layout(title="Home Libraries")
, output_type='div')
```

EMBED SIERRA CHARTS

Can we embed multiple charts in one HTML page?

- Use the same SQL output from circ_trans
- Run data transformations on each column
- Pltoly option: output_type='div'
- Pass chart outputs to the template engine
- Multiple charts will be embedded as instructed in a template file

INTERACTIVE PROGRAM

Can we specify days and transaction types?

- Python `sys.argv` - read user inputs from terminal
- User Inputs:
 - days
 - transaction types(o i f r)
 - output filename

```
python sierra_chart_embed.py 60 oi sierra_interactive.html
```


SIERRA SQL API & PLOTLY: SUMMARY

- Plotly Python Offline
- SierraDNA to learn Sierra database
- PGAdmin to run test queries
- Python database adapter (psycopg2)
- Transform data for Plotly
- Template engine Jinja2 to embed multiple charts
- Interactive program with user inputs

MORE IDEAS...

- Create the reports of special purposes
- Automatically run and deliver reports (scheduler, clone)
- Explore more data from Sierra (collections, circulation, acquisition)
- Collect non-Sierra data (counter, space usages, research/reference)
- Data from other APIs (Googlesheet, LibGuides)
- Plotly's different charts (map, bubble, heatmap)

WHAT IS NEXT?

[LIBRARY DATA GROUP]

SEONG HEON LEE, SELEE@CHAPMAN.EDU

THANK YOU.