Programming Language Instruction in Libraries

Lorin Bruckner (UNC)  Joel Herndon (Duke)
Amanda Henley (UNC)  Jeremiah Trinidad-Christensen (Columbia)
Programming Language Instruction in Libraries

Research Data Services, Columbia University Libraries

Jeremiah Trinidad-Christensen
jt2118@columbia.edu
Map Club!

- Introduced Fall 2016
- Rapid prototyping
- Tackle something new each time
- Loosely structured
- Mostly JavaScript
Tools introduced

- Tangram
- Leaflet
- D3
- geojson.io
- Cartagen
Open Labs for R & Python

Why?

- Frustration from students & staff
  - Matching question to the right person
- Community building
- Needed alternative to workshop model
Open Lab format

What?

- Semi-structured format
  - 10-30 minute talk
- Modular, not built from previous session
- Intern led, part of the Digital Center Internship Program
- Starter kits for new users

```
In [147]: # load in flights dataset
flights = seaborn.load_dataset("flights")

In [148]: # print first five rows
flights.head()

Out[148]:

<table>
<thead>
<tr>
<th>year</th>
<th>month</th>
<th>passengers</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>January</td>
<td>112</td>
</tr>
<tr>
<td>1</td>
<td>February</td>
<td>118</td>
</tr>
<tr>
<td>2</td>
<td>March</td>
<td>132</td>
</tr>
<tr>
<td>3</td>
<td>April</td>
<td>129</td>
</tr>
<tr>
<td>4</td>
<td>May</td>
<td>121</td>
</tr>
</tbody>
</table>

In [149]: # print last five rows
flights.tail()

Out[149]:

<table>
<thead>
<tr>
<th>year</th>
<th>month</th>
<th>passengers</th>
</tr>
</thead>
<tbody>
<tr>
<td>139</td>
<td>August</td>
<td>606</td>
</tr>
<tr>
<td>140</td>
<td>September</td>
<td>508</td>
</tr>
<tr>
<td>141</td>
<td>October</td>
<td>461</td>
</tr>
<tr>
<td>142</td>
<td>November</td>
<td>390</td>
</tr>
<tr>
<td>143</td>
<td>December</td>
<td>432</td>
</tr>
</tbody>
</table>
```

```
In [150]: flights = flights.pivot("month", "year", "passengers")
# draw border
x = seaborn.heatmap(flights, linewidths=0.3)
```
Different needs

- R taught in different schools & departments
- Tied to curriculum
- Python not supported in the same way, but a strong demand for learning
- More structured, sessions tied more closely together
Takeaways

- Community building is tough!
- Takes time onboarding interns
- Need better communication with teaching faculty
- Assessment
Programming Pathways
University of North Carolina at Chapel Hill

Lorin Bruckner
lorin.bruckner@unc.edu

Amanda Henley
amanda.henley@unc.edu
Different Instruction Models

- Traditional standalone short courses
- Bootcamp/intensive style “Data Wrangling” event
- Columbia University Library’s Open Labs
Why Open Labs?

- Learning R takes time and practice
- Inspired by Columbia’s success with the model
- Previous success expanding GIS learning communities in the past
- Collaboration with the Odum Institute
Our Implementation

- Brief instruction followed by lab time
- Exercises are provided
- Lessons available online using R Markdown
  ropenlabs.web.unc.edu
- Pilot: Badging

Exercises

Note: If you are completing exercises to obtain a badge, please show your work instead of providing just the answers.

1. What are the different types of cancellation policies offered at Airbnb?
2. What is the most common and least common cancellation policy?
3. What percent of Airbnb property types are houses? Hint: To find what percent of \( X \) is \( Y \), use this formula: \( \frac{Y \times 100}{X} \)
4. How many listings get a review rating below 50?
5. What is the maximum number of people that can be accommodated at each property type? Hint: For `aggregate()`, we can look at other statistics than just the mean. Another option is "max" which gives us the highest value at each level.
What Worked

- Huge initial attendance
- Transitioned to a dedicated core group
- Diverse group of participants, from undergrads to faculty
- Corpus of instruction material, data
Assessment

- Feedback surveys
- Attendance data
- Voting
- (Future) focus groups

“Congratulations! The papers have been finally selected. Our team is looking forward to working with the authors. If you have any questions or need any assistance, please feel free to reach out.”

“I love these labs…”

“Thanks guys, this was helpful. I have zero knowledge in R but now, I am getting excited about R…”
What We’re Still Figuring Out

- Encouraging collaboration
- Helping latecomers catch up
- Reducing prep time
- Managing beginners’ sophisticated goals
- Teaching trial and error
Ideas/Next Steps

- Adjust format to encourage collaboration
- Alternative types of exercises
- Outsource materials
- Stick to the basics
- Increase outreach to faculty
New Frontiers in Services

Embracing code and building community @ Duke Libraries

Joel Herndon
joel.herndon@duke.edu
Twitter: @jherndon
Data and Visualization Services

http://library.duke.edu/data
askdata@duke.edu
SUPPORT AREAS

- Data Sources
- Data Cleaning
- Mapping and GIS
- Data Management
- Data Analysis
- Data Visualization
DVS Workshops
Spring 2018
January to April 2018
Early Instruction

- High demand
- Varying expectations
- Focused on syntax / data structures
Reboot
Interactive

- Less syntax
- More hands-on
- Intuitive examples
Reproducible

- Embed data management
- Illustrate best practices
Community Focus

DVS Workshops
Spring 2018

R Workshop Series
Intro to R
Data Management, Git, & RStudio
Visualization in R with ggplot2
Mapping with R
R Markdown
Shiny

rfun.library.duke.edu
Lessons Learned
Iterate

Created by Weltenraser from Noun Project
Simplify

Created by Chris Homan from Noun Project
Questions?

Joel Herndon
Email: joel.herndon@duke.edu
Twitter: @jherndon

Resources:
RFUN: rfun.library.duke.edu
DVS Workshops: library.duke.edu/data/workshops/past-workshops