Polishing Your Analysis Tools Until They Are Shiny

Andrew Dau
United States Department of Agriculture – National Agricultural Statistics Service
The findings and conclusions in this preliminary presentation have not been formally disseminated by the U. S. Department of Agriculture and should not be construed to represent any agency determination or policy.
National Agricultural Statistics Service (NASS)

- NASS is committed to providing timely, accurate, and useful statistics in service to U.S. agriculture.

- Over 400 reports annually.
  - Census of Agriculture every 5 years.

- Majority of reports driven by surveys
Traditional NASS Survey Process

Data Collection

Editing of Data

Summarization

Estimation Process

Interactive Data Analysis System (IDAS)

Analysis of Data
What is IDAS?

• Analysis tool built using SAS/AF

• First tool deployed in 1994
  – Used by analysts in headquarters and our twelve regional offices – guided analysis
  – Written and maintained by statisticians in Methodology Division

• Motivation to change...
  – SAS no longer releasing any AF updates or enhancements
  – Architecture updates
  – Enhance our analysis capabilities
What is R Shiny?

• Shiny helps you make interactive web applications for visualizing data. Bring R data analysis to life.

• Shiny combines the computational power of R with the interactivity of the modern web.
R Shiny

HTML

JavaScript

jQuery1.com

DB
R Shiny App Basics

- server.R
  - Defines functionality of app

- ui.R
  - Defines app layout
Why R Shiny?

- Shiny **User Showcase** looks promising
- Some familiarity with R
  - More familiarity coming?

<table>
<thead>
<tr>
<th>Use</th>
<th>Java</th>
<th>JMP</th>
<th>Python</th>
<th>R</th>
<th>SAS</th>
<th>Tableau</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rarely</td>
<td>122</td>
<td>128</td>
<td>90</td>
<td>74</td>
<td>101</td>
<td>103</td>
</tr>
<tr>
<td>Monthly</td>
<td>6</td>
<td>4</td>
<td>21</td>
<td>22</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>Weekly</td>
<td>2</td>
<td>4</td>
<td>10</td>
<td>14</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Daily</td>
<td>6</td>
<td>0</td>
<td>27</td>
<td>29</td>
<td>16</td>
<td>14</td>
</tr>
</tbody>
</table>

(2018, Pierson - ASA Survey of Recent Graduates)
Where should you start?
Proof of Concept

• Assess the viability of a software product solving a business need
  – Does it have functionality to replace old application?
  – Is it secure?
  – Is it scalable?
  – etc.
Proof of Concept Part 1

Scatter Plots
DemoScatterPlot

The diagram shows a scatter plot with the states IA, IN, KS, and KY on the x-axis and current_value_num on the y-axis. The data points are color-coded as follows:

- **Operator** in blue
- **Spouse** in green
- **Partner** in purple

The state KS has a data point labeled (KS, 1209) for the Operator category.
DemoScatterPlot
DemoScatterPlot

![Scatter plot with data points indicating current value numbers for different states (KS, KY) with legend for Operator, Spouse, and Partner.](scatter_plot.png)
DemoScatterPlot

Variable Name
Fake VAR 1

Marker Color
- Mode
- Respondent
- Reviewed Status
- Comments
DemoScatterPlot

Variable Name
Fake VAR 2

Marker Color
- Mode
- Respondent
- Reviewed Status
- Comments

State
IA  IL  IN  KS  KY

Fake VAR 2
0  500  1000  1500  2000
DemoScatterPlot

Variable Name
- Fake VAR 2

Marker Color
- Mode
- Respondent
- Reviewed Status
- Comments
### DemoScatterPlot

<table>
<thead>
<tr>
<th>id</th>
<th>varname</th>
<th>current_value_num</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>217 DUM1</td>
<td>1707</td>
</tr>
</tbody>
</table>

Showing 1 to 1 of 1 entries

**Variable Name**
- Fake VAR 1

**Marker Color**
- Mode
- Respondent
- Reviewed Status
- Comments

**Click Event**
- Record
- Listing-Above
- Listing-Below
- Listing-All
- Listing-State
### Demo Scatter Plot

#### Scatter Plot

<table>
<thead>
<tr>
<th>id</th>
<th>varname</th>
<th>current_value_num</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>209 DUM1</td>
<td>1368</td>
</tr>
<tr>
<td>2</td>
<td>214 DUM1</td>
<td>1520</td>
</tr>
<tr>
<td>3</td>
<td>207 DUM1</td>
<td>1646</td>
</tr>
<tr>
<td>4</td>
<td>217 DUM1</td>
<td>1707</td>
</tr>
</tbody>
</table>

Showing 1 to 4 of 4 entries

#### Variable Name

- Fake VAR 1

#### Marker Color

- Mode
- Respondent
- Reviewed Status
- Comments

#### Click Event

- Record
- Listing-Above
- Listing-Below
- Listing-All
- Listing-State
25 Records Displayed
Useful Packages

- ggplot2
- plotly
- shinydashboard
- shinydashboardPlus
- data.table
- esquiesse
- leaflet
Conclusion

• R shiny has shown necessary capability...so far.
  – Need further development
  – Need to test deployment strategy
  – Need to stress test
  – etc.

• Lots of questions to be answered...
Contact Info

Andrew Dau

andrew.dau@nass.usda.gov
Selected References

www.google.com
www.rstudio.com
https://divadnojnarg.github.io/blog/awesomedashboards/
https://cran.r-project.org
https://github.com/dreamRs/esquisse
