An R-based Web Application to Search, Analyze and Display Water Quality Data in Oregon State, USA

Peter Bryant

May 4th, 2016
National Water Quality Monitoring Council
10th National Monitoring Conference
Tampa, Florida
Web Application Goals

Provide an interactive means for users to:

- Query data from multiple databases
- Evaluate status and trend at individual sampling stations
- Display the results

Non-R users are able to generate information on status and trends
Water Quality Data Start Here…
...And Go Here

• DEQ databases – Current and legacy databases for grab and continuous data

• NWQMC Water Quality Portal for grab data
  – EPA STORET
  – USGS NWIS

• USGS National Water Information System for continuous data
Then What?

Use the data to:

• Determine status of water quality

• Look for trends in water quality
But How?

That’s a lot of places to look and potentially a lot of data to analyze

Luckily there’s
What is R?

- Open source statistical programming language
- It’s free (r-project.org)
- Has a large user base
- Lots of documentation and learning resources
How Does R Work?

• Use a text editor to write out the analysis

• Pass the analysis code to the R software to actually run the analysis

• R users have built packages to do many different types of analysis

• A package is a collection of functions to perform an analysis
The Shiny Package

• The Shiny package is built by RStudio

  – Provides R functions that translate R code to an interactive javascript web application

  – Builds a user interface to edit settings in the analysis

  – Puts all the code to run an analysis behind the user interface
While I see this
You see this

Oregon Water Quality Status and Trend Beta Version 2.0

Select Plan Area
Burnt River

Select Parameters to Query
- Temperature
- pH
- Bacteria

Select the Start and End Dates
2006-04-05 to 2016-04-05

Select Database(s) to Query:
- Water Quality Portal
- DEQ

Run Query
Submit

You just submitted Burnt River Plan Area Query for Temperature, pH, Bacteria from 2006-04-05 to 2016-04-05

<table>
<thead>
<tr>
<th>Analyte</th>
<th>Stations</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. Coli</td>
<td>16</td>
<td>502</td>
</tr>
<tr>
<td>Fecal Coliform</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>pH</td>
<td>23</td>
<td>7376</td>
</tr>
<tr>
<td>Temperature</td>
<td>43</td>
<td>94621</td>
</tr>
</tbody>
</table>

Download the data
View map
Application Structure
Search For Data

Query Databases

• RODBC package for DEQ data
  – Provides connection from R to SQL Server databases

• dataRetrieval package for Water Quality Portal data
  – A collection of functions to get data from USGS and EPA National Databases
  – Uses web services to communicate with the databases
## Search For Data

**Oregon Water Quality Status and Trend Beta Version 2.0**

### Select Plan Area
- Burnt River

### Select Parameters to Query
- Temperature
- pH
- Bacteria

### Select the Start and End Dates
- 2006-04-05 to 2016-04-05

### Select Database(s) to Query:
- Water Quality Portal
- DEQ

### Run Query
- Submit

You just submitted Burnt River Plan Area Query for Temperature, pH, Bacteria from 2006-04-05 to 2016-04-05.

### Analyte Results Table

<table>
<thead>
<tr>
<th>Analyte</th>
<th>Stations</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. Coli</td>
<td>16</td>
<td>502</td>
</tr>
<tr>
<td>Fecal Coliform</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>pH</td>
<td>23</td>
<td>7376</td>
</tr>
<tr>
<td>Temperature</td>
<td>43</td>
<td>94621</td>
</tr>
</tbody>
</table>

- Download the data
- View map
Search For Data
Analyze

Format

• Each database has different field names

• Built function to combine results from each database into a single table
Analyze

Look for trend

• Uses Seasonal Kendall trend analysis from wq package

• Apply to each station for each parameter
Analyze

Compare to Water Quality Standard

- Many of Oregon’s standards are geo-specific (e.g. Temperature and pH)
- Data are gathered from many different sources
- Station location information may not be accurate
- User is able to specify geo-specific criteria to apply to the station for analysis
Analyze

Compare to Water Quality Standard

• Temperature standard is based on 7 Day Average Daily Maximum

• E. Coli and Enterococcus are based on 30-day Geometric Mean
Display

• Provide tabular summaries of formatted data

• Create interactive charts using ggplot2 package

• Provide zooming to specific time periods in the charts

• Provide download of chart to insert into report

• Construct responsive table of exceedances based on changes to selected water quality standard
## Oregon Water Quality Status and Trend Beta Version 2.0

<table>
<thead>
<tr>
<th>Client</th>
<th>Analyte</th>
<th>Station_ID</th>
<th>Station_Description</th>
<th>SampleType</th>
<th>Result</th>
<th>MRL</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E. Coli</td>
<td>11494</td>
<td>Burnt River at Snake River Road (Huntington)</td>
<td>Grab Sample::GS</td>
<td>22.0</td>
<td>1.0</td>
<td>MPN/100 mL</td>
</tr>
<tr>
<td></td>
<td>pH</td>
<td>11494</td>
<td>Burnt River at Snake River Road (Huntington)</td>
<td>Grab Sample::GS</td>
<td>8.5</td>
<td>0.1</td>
<td>pH Units</td>
</tr>
<tr>
<td></td>
<td>Temperature</td>
<td>11494</td>
<td>Burnt River at Snake River Road (Huntington)</td>
<td>Grab Sample::GS</td>
<td>8.3</td>
<td>0.0</td>
<td>°C</td>
</tr>
</tbody>
</table>
Display
Display
Display
Burnt River at Clarks Creek bridge, ID = 34256

p value = 1, Not Significant, slope = 37.17, n = 89

- Exceeds Single Sample
- Geometric Mean WQS
- Meets Single Sample
- Exceeds Geometric Mean
- Meets Geometric Mean
- Single Sample WQS

E Coli

Date

May
Jun
Jul
Aug
Sep

Save plot
Display
Summary

This R based web application is built using the Shiny package in R Studio and:

• Uses database queries to search for data

• R code to facilitate streamlined analysis

• Includes interactive tables and charts
If you want to see this
Go To This Project’s Repository

https://github.com/petertbryant/StatusAndTrends
Thanks to

- Oregon DEQ Laboratory Water Quality Monitoring Group and Steve Hanson, Wade Peerman, Michael Tichenor and Lori Pillsbury

- John Paul Schmit with the National Park Service whose Shiny app provided inspiration and whose Github repository helped get me started
Questions?

Contact info:

bryant.peter@deq.state.or.us
• Github is a website where you share your R code with others online

• Github uses the version control software git

• Version control software keeps track of the changes you make to your code