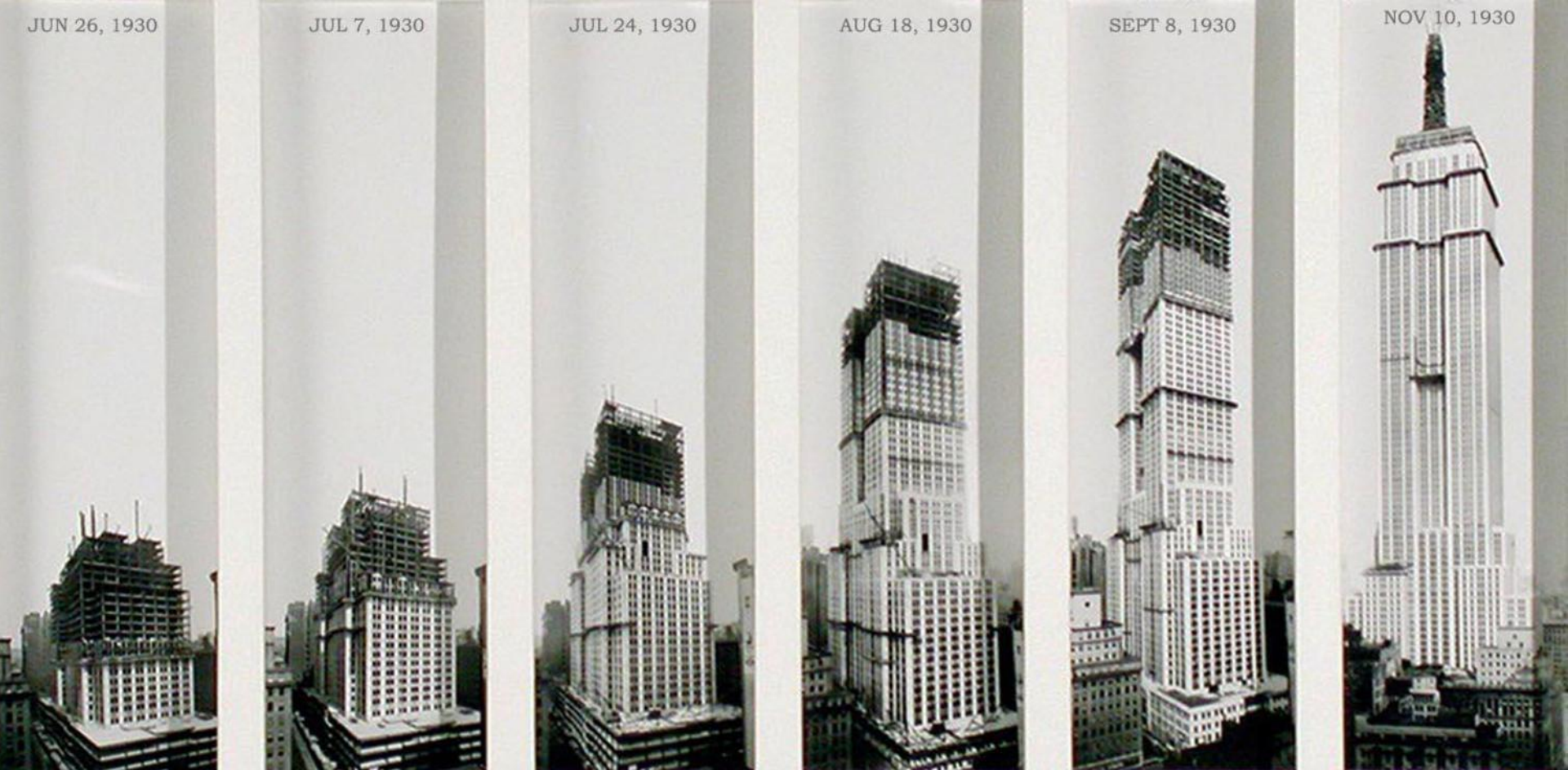




A Demonstration of a Free
Tool that Takes Data from
the Water Quality Portal
and Generates Files you
can Load to ATTAINS in Less
than 15 Minutes

May 29, 2019
Jason D. Jones



The Empire State Building was built in **410** days
(1/22/1930 to 3/8/1931)

Arizona CWA Assessment finished in **817** days
(7/1/2015 to 9/25/2017)

We use for

1. Real Time Assessments
2. Sending official results to ATTAINS

What is ?

R is an advanced set of tools
for data analysis.



Imports Data from Portal Automatic assignment of Assessment Unit

Formats Data (used to be 80% of the work!)

Segregate Data (Improvements, Credibility)

Removes results w/ detection limits greater than the standard

Aggregate by Space and Time

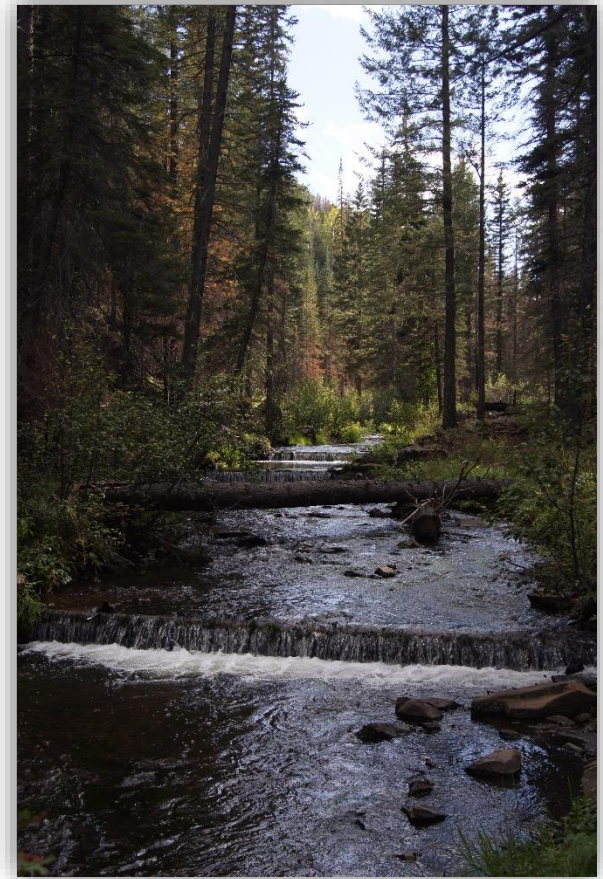
Handles Hard Standards (Ammonia, Site Specific, statistically based)

Real-Time Assessment with Human Override of Automated Decision

Assess at the Parameter, Use and Assessment Unit Levels

Export to ATTAINS

- **If-then** analysis to see how **changes in new standards** will impact impairments
- **Tailored Sample Plans.** Shows exactly what is needed to make a decision. **We could reduce our FY20 lab budget by \$180,000!**
- Able to see **improvements in real time** (delists for copper and beryllium for Boulder Creek)



Need 2 arsenic, 1 E. coli
and 1 total boron samples



Three tracks

1



Automated Decision.
Solely based on data.

2



Provisional Decision.
Human can override computer.

3



Official Decision.
Send to EPA's ATTAINS.

Move from Automated to Provisional

WBID	WATERBODY_DESC	NewUse	CharacteristicName	ResultSampleFractionText	Automated assess	No	Yes	samplecount	provassess	provdate
14070006-001	COLORADO RIVER	AWCChronic	SELENIUM	Total	Not meeting criteria	53	6	59	Not meeting criteria	12/27/2018
15010001-175B	UNKAR CREEK	AWWChronic	SELENIUM	Total	Not meeting criteria	1	2	3	Not enough information	4/11/2019
15010004-001B	HAVASU CREEK	AWWChronic	SELENIUM	Total	Not meeting criteria	3	4	7	Not meeting criteria	4/11/2019
15020002-016	LITTLE COLORADO RIVER	AWCAcute	SUSPENDED SEDIMENT CONCENTRATION (SSC)	Suspended	Not meeting criteria	0	3	3	Not enough information	12/27/2018
15020005-009	SILVER CREEK	FBC	ESCHERICHIA COLI	Total	Not meeting criteria	0	5	5	Not meeting criteria	12/28/2018
15020005-1380	SHOW LOW LAKE	AWCChronic	IRON	Dissolved	Not meeting criteria	0	2	2	NA	NA
15020015-206	NEWMAN CANYON	AWCAcute	SUSPENDED SEDIMENT CONCENTRATION (SSC)	Suspended	Not meeting criteria	0	2	2	Not enough information	3/13/2019

Automated – Live Data

Provisional

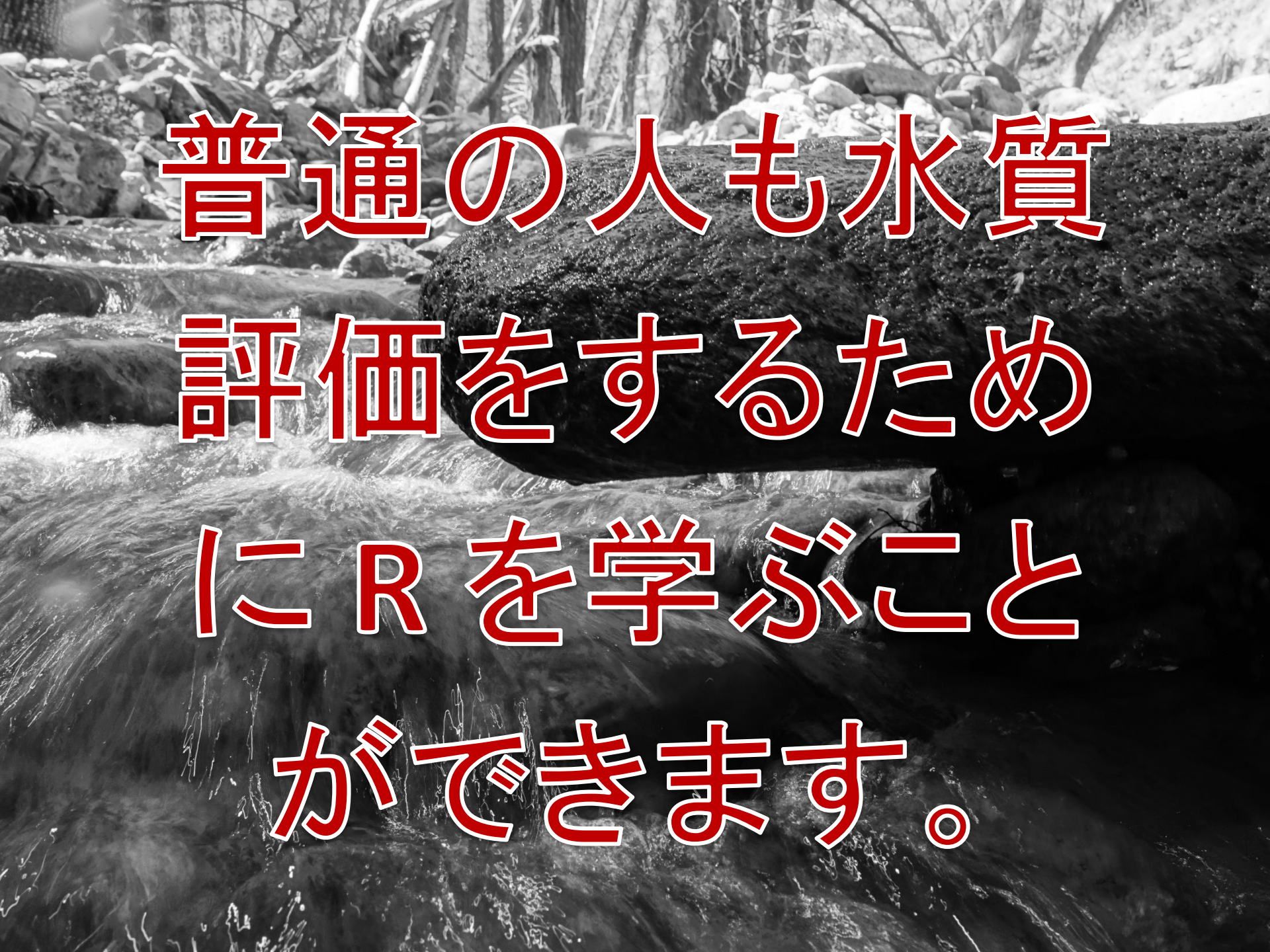
Highlights Possible New **Impairments** and New **Delists**

Generate **ATTAINS** files based on **provisional** data already reviewed

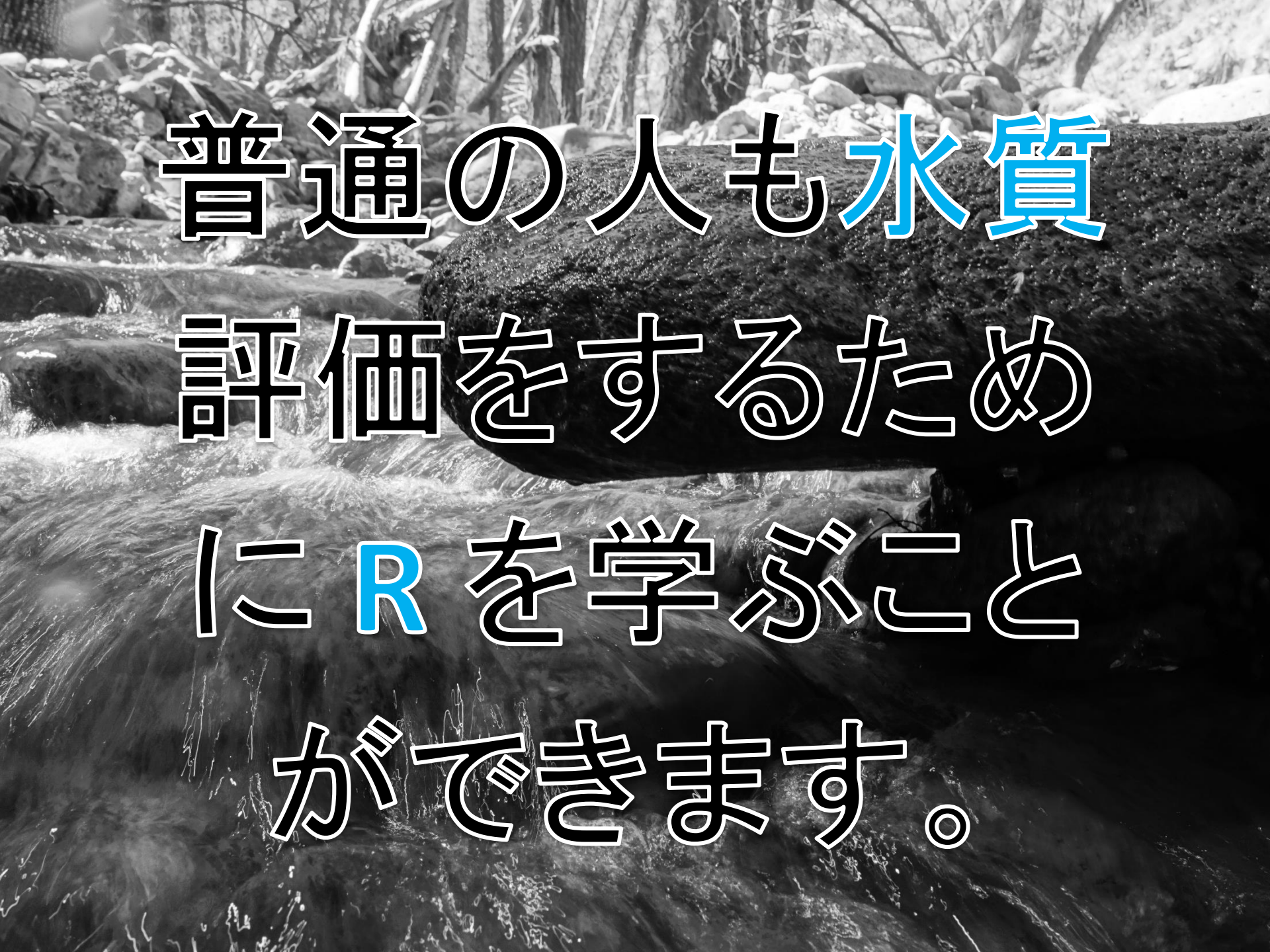
1. assessment.csv
2. associated-actions.csv
3. parameter.csv
4. sources.csv
5. uses.csv

Output = 2020 Arizona Assessment Calculator Parameter File

A	B	C	D	E
ASSESSMENT_UNIT_ID	PARAM_NAME	PARAM_USE_NAME	PARAM_STATUS_NAME	PARAM_ATTAINMENT_CODE
15020010-006	SELENIUM	Full Body Contact	Meeting Criteria	Meeting criteria
15020010-006	SELENIUM	Fish Consumption	Meeting Criteria	Meeting criteria
15020010-006	ZINC	Agricultural Irrigation	Meeting Criteria	Meeting criteria
15020010-006	ZINC	Agricultural Livestock Watering	Meeting Criteria	Meeting criteria
15020010-006	ZINC	Aquatic and Wildlife (Coldwater Fishery)	Meeting Criteria	Meeting criteria
15020010-006	ZINC	Full Body Contact	Meeting Criteria	Meeting criteria
15020010-006	ZINC	Fish Consumption	Meeting Criteria	Meeting criteria
15020010-0180	AMMONIA-NITROGEN	Aquatic and Wildlife (Coldwater Fishery)	Cause	Not meeting criteria
15020010-0180	ANTIMONY	Aquatic and Wildlife (Coldwater Fishery)	Insufficient Information	Not enough information



普通の人も水質
評価をするため
にRを学ぶこと
ができます。

A black and white photograph of a river flowing over rocks, with the text overlaid. The text is in a stylized font with white outlines and blue highlights for the words '水質' (water quality) and 'R'.

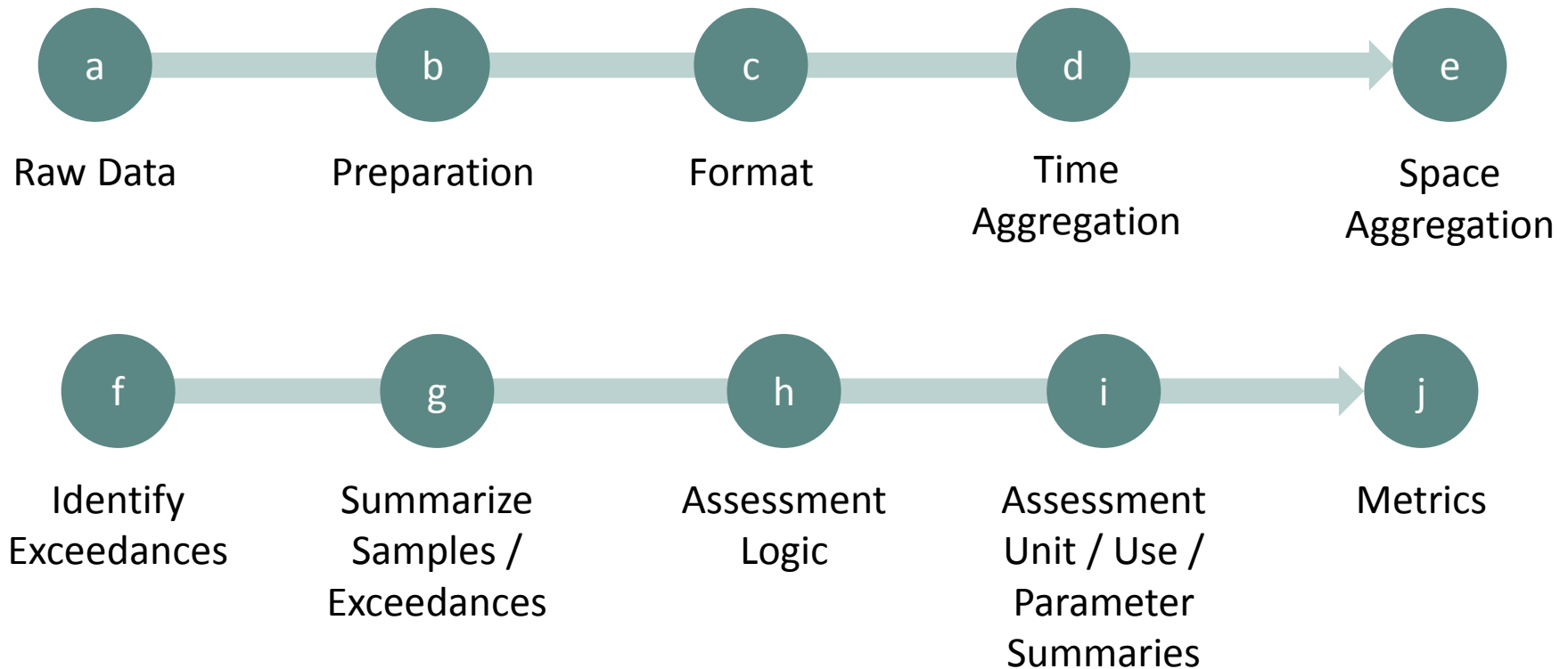
普通の人も水質
評価をするため
に R を学ぶこと
ができます。

Road Map

```
#### E - AGGREGATE TO ASSESSMENT UNIT / WBID ####

# Set up two choices. Almost every parameter = maximum. DO is a minimum. pH is a Max and Min
e.spaceaggregate <- mutate(d.timeaggfinal, Steptwomap = ifelse(CharacteristicName == "DISSOLVED OXYGEN (DO)", "Minimum", "Maximum"))










# Add in exceptions for pH and oxygen saturation
e.spaceaggregate[grepl("HIONMAX", e.spaceaggregate$CharacteristicName), "Steptwomap"] <- "Maximum"
e.spaceaggregate[grepl("HIONMIN", e.spaceaggregate$CharacteristicName), "Steptwomap"] <- "Minimum"
e.spaceaggregate[grepl("DISSOLVED OXYGEN SATURATION", e.spaceaggregate$CharacteristicName), "Steptwomap"] <- "Minimum"
```



Folder Structure – Inputs and Outputs

Working Directory

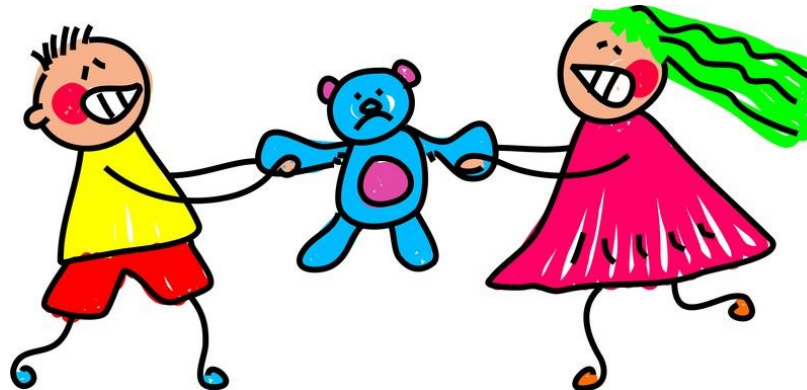
← → ↕ ↑ This PC > 65942 (M:) > Real Time Impairment

Name	Date modified	Type
 Real Time Assessment.R The code	5/7/2019 1:46 PM	R File
 .Rhistory	5/6/2019 2:54 PM	RHISTORY File
 human.csv The main output	5/3/2019 3:47 PM	Microsoft Excel C...
 human research.docx	5/3/2019 3:05 PM	Microsoft Word D...
 schema.ini	4/25/2019 10:51 AM	Configuration sett
 inputs Inputs (like standards)	5/7/2019 11:37 AM	File folder
 humancopies	5/3/2019 10:34 AM	File folder
 metrics	4/26/2019 11:56 AM	File folder
 buckets	3/4/2019 3:12 PM	File folder

A black and white photograph of a piece of driftwood with seaweed on a rocky shore. The driftwood is a long, thin, and gnarled piece of wood, possibly a branch or a piece of a tree trunk, that has been weathered and bleached by the sun. It is covered in various types of seaweed, including long, thin, and feathery fronds, as well as some larger, more rounded, and bulbous structures. The driftwood is resting on a dark, wet, and rocky surface, which appears to be a shoreline or a beach. The background is a blurred view of the ocean with waves breaking on the shore. The overall mood is somber and contemplative.

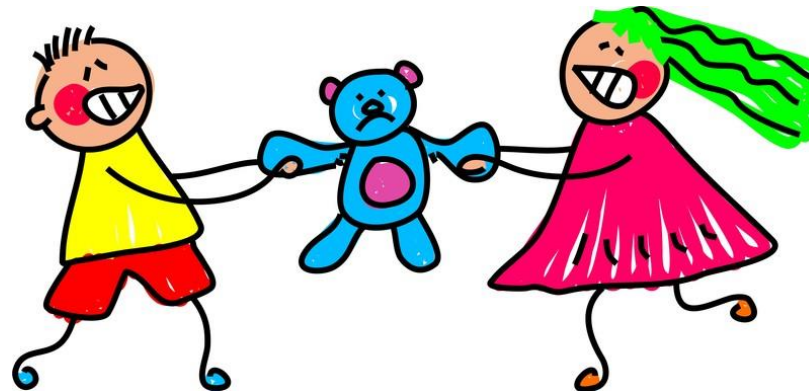
Now What?

1. Already have a methodology...why change?
2. Tool not sharable (proprietary code/software)
3. The open source learning curve
4. Special snowflakes of the Clean Water Act
5. We don't use common infrastructure for ~~data~~, assessment units, standards



Top 5 Reasons We Don't Share Methodology

1. Already have a methodology...why change?
2. Tool not sharable (proprietary code/software)
3. The open source learning curve
4. Special snowflakes of the Clean Water Act
5. We don't use common infrastructure for data, assessment units, standards



Biology Degree

www.datacamp.com



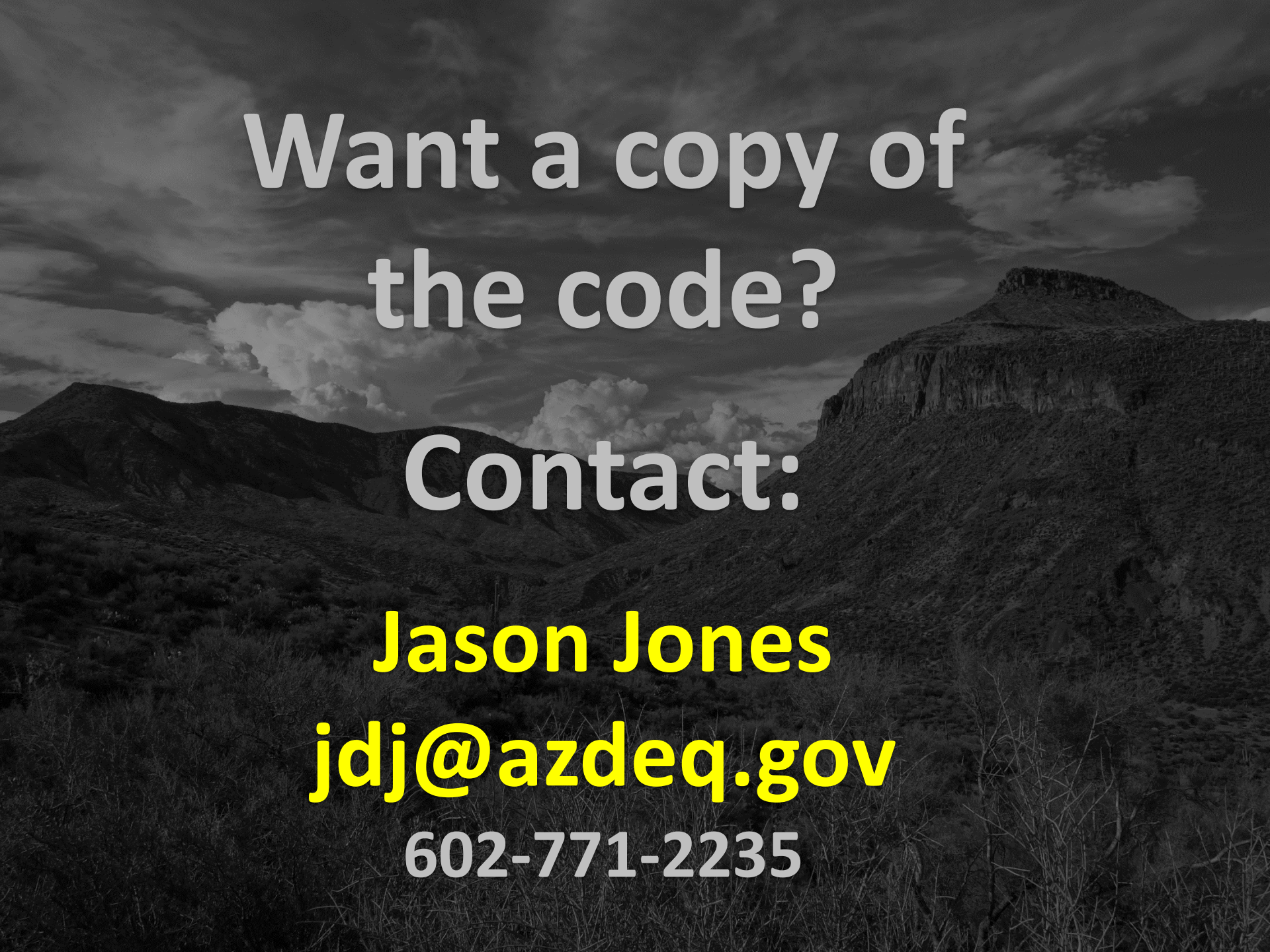
Cost = **\$300** for a Year all Content

Time = **20 hours**

Recommended Courses to do Basic Assessments

- 4 hours basic R
- 4 hours intermediate R
- 8 hours for dplyr (data manipulation)
- 4 hours R Studio



A black and white photograph of a desert landscape. In the foreground, there are dark, scrubby bushes. In the middle ground, there are rolling hills and mountains. On the right side, a prominent mesa with a flat top is visible. The sky is filled with dramatic, layered clouds. The overall tone is somber and rugged.

**Want a copy of
the code?**

Contact:

Jason Jones

jdj@azdeq.gov

602-771-2235