USING R TO TRACK CHANGES IN ASSESSMENTS

June 23, 2023

Presentor: Mikayla Baker
Environmental Program Manager
Resources

Jason Jones, Developer/Assessment Lead

Assessments From Months to Minutes

**Old Way**

1. Gather (3 months)
2. Format (6 months)
3. Assess (4 months)
4. Send to ATTAINS

**Time = 13 months**

**New Way**

1. Gather (7 minutes)
2. Format (0 minutes)
3. Assess (5 minutes)
4. Send to ATTAINS

**Time = 12 minutes!**
It is Just Math

Data | Assessment Units | Standards | Previous Decisions | Impairment/Attainment Logic

Parameter

Designated Use

Assessment Unit

Not Meeting Criteria | Meeting Criteria

Not Enough Information | Insufficient Information

Not Supporting | Supporting

Impairment | Attainment

Inconclusive

Insufficient Information

Not Enough Information

Meeting Criteria

Not Meeting Criteria
Option 1 - Use ATTAINS

ATTAINS - Arizona (21ARIZ)

Home  Assessment Units  Assessments  Actions  Reports  Priorities  Surveys  Administration

Cycle Summary

Report Type*
- Cycle Summary
- Select Report Type
- Cycle Summary

Report*
- Cycle Comparison
- Assessment Units
- Action Summary

Cycle*
- 2024 - Organization Draft (IR)

Run Report
## Cycle Comparison

<table>
<thead>
<tr>
<th>Assessment Unit ID</th>
<th>Assessment Unit Name</th>
<th>Previous Assessment Unit Name</th>
<th>EPA IR Category</th>
<th>Previous EPA IR Category</th>
<th>EPA IR Category Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>AZ15020001-013A_00</td>
<td>Little Colorado West Fork, from Headwaters to GOVERNMENT SPG</td>
<td>Little Colorado West Fork, from Headwaters to GOVERNMENT SPG</td>
<td>2</td>
<td>2</td>
<td>--</td>
</tr>
<tr>
<td>AZ15050301-550_00</td>
<td>Cox Gulch, from Headwaters to THREE R CANYON</td>
<td>Cox Gulch, from Headwaters to THREE R CANYON</td>
<td>4A</td>
<td>4A</td>
<td>--</td>
</tr>
<tr>
<td>AZ15070102-006A_00</td>
<td>New River, headwaters - Interstate 17</td>
<td>New River, headwaters - Interstate 17</td>
<td>2</td>
<td>2</td>
<td>--</td>
</tr>
<tr>
<td>AZL15060106A-1290_00</td>
<td>Saguaro Lake</td>
<td>Saguaro Lake</td>
<td>2</td>
<td>3</td>
<td>CHANGED</td>
</tr>
<tr>
<td>AZ15030104-015_00</td>
<td>COLORADO RIVER</td>
<td>COLORADO RIVER</td>
<td>1</td>
<td>2</td>
<td>CHANGED</td>
</tr>
<tr>
<td>AZ15060106A-583_00</td>
<td>Fish Creek, from Headwaters to SALT RIVER</td>
<td>Fish Creek, from Headwaters to SALT RIVER</td>
<td>3</td>
<td>3</td>
<td>--</td>
</tr>
</tbody>
</table>
What observations are in **Table 1** are not in **Table 2**?

**Table 1**

- a
- b
- c

**Table 2**

- a
- c
- d
What observations are in **Table 1** are not in **Table 2**?
Option 2 - Use R

```r
library(tidyverse)

# 1. A working directory is where on your computer R puts things.
# 2. Set by navigating to the folder you want to save things by clicking
#    the ... in the files pane. Navigate to folder where the files live on y
#    It is good practice to put each project you work on in a folder.
# 3. Click 'More' in the file pane and 'set as working directory'.
# 4. Copy this code into your script (don't copy the '>').

# CHANGE THE CODE BELOW AND SET YOUR WORKING DIRECTORY
setwd("~/R Presentation/antijoin example-20230516T191222-001/antijoin example")

## QUESTION 1 - NEW DELISTS/IMPAIRMENTS ##

# Import Impairments
j.imp_2026_par <- read_csv("j.imp_2026_par.csv")
j.imp_2024_par <- read_csv("j.imp_2024_par.csv")

```

Data

- j.imp_2026_par: 293 obs. of 2 variables
- j.imp_2024_par: 290 obs. of 2 variables
### Option 2 - Use R

<table>
<thead>
<tr>
<th>WBID</th>
<th>CharacteristicName</th>
</tr>
</thead>
<tbody>
<tr>
<td>14070006-001</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>14070006-1130</td>
<td>MERCURY IN FISH TISSUE</td>
</tr>
<tr>
<td>14070007-123</td>
<td>ESCHERICHIA COLI</td>
</tr>
<tr>
<td>14070007-123</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>14070007-123</td>
<td>SUSPENDED SEDIMENT CONCENTRATION (SSC)</td>
</tr>
<tr>
<td>15010001-003</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010001-005</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010001-006</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010001-008</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010001-010</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010001-011</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010001-012</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010001-013</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-001</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-003</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-004</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-007</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-009</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-011</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-012</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-013</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-014</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-016</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-017</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-018</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-019</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-020</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-021</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-022</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-023</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-024</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-025</td>
<td>SELENIUM</td>
</tr>
<tr>
<td>15010002-026</td>
<td>SELENIUM</td>
</tr>
</tbody>
</table>

Showing 1 to 25 of 293 entries, 2 total columns.
Option 2 - Use R
Option 2 - Use R

```r
# QUESTION 1 - NEW DELISTS/IMPAIRMENTS

# Import Impairments
j.imp_2024_par <- read_csv("j.imp_2024_par.csv")
j.imp_2024_par <- read_csv("j.imp_2024_par.csv")

# Param Delists = In the 2024 list but not in the 2026 list
j.imp_2024_param_delist <- anti_join(j.imp_2024_par, by = c("WBID", "CharacteristicName"))
mutable(delist = "Yes")

# Param New Impairments = In the 2026 list but not in the 2024 list
j.imp_2026_param_newimp <- anti_join(j.imp_2024_par, by = c("WBID", "CharacteristicName"))
mutable(newimp = "Yes")
```

---

```
---

# QUESTION 1 - NEW DELISTS/IMPAIRMENTS

j.imp_2024_par <- read_csv("j.imp_2024_par.csv")
j.imp_2024_par <- read_csv("j.imp_2024_par.csv")

j.imp_2026_param_delist <- anti_join(j.imp_2024_par, by = c("WBID", "CharacteristicName"))
mutable(delist = "Yes")

j.imp_2026_param_newimp <- anti_join(j.imp_2024_par, by = c("WBID", "CharacteristicName"))
mutable(newimp = "Yes")
```
Questions?

Jason Jones
jdj@azdeq.gov

Mikayla Baker
baker.mikayla@azdeq.gov