

# USING R TO TRACK CHANGES IN ASSESSMENTS

June 23, 2023

Presenter: Mikayla Baker  
Environmental Program Manager





**Jason Jones**, Developer/Assessment Lead



**Glen Holroyd**, Architect, Code QC, IT Rep.

# Assessments From Months to Minutes

Old Way



External  
Data

1. Gather  
(3 months)



AZDEQ  
Database

2. Format  
(6 months)



3. Assess  
(4 months)



ID	Name	Water Type	Size	Units	Status
AZ10010101-000A_01	NEW RIVER - NHR - S - MG - MG - 10010101	STREAM	231	Miles	Active
AZ10010101-000A_02	Cox Gulch, Fish Headwaters to THREE R CANYON	STREAM	2.3	Miles	Active
AZ10010101-000A_03	Lake Cochise Head Fork, Non-Headwaters to GOVERNMENT SPG	STREAM	9.075	Miles	Active
AZ10010101-000A_04	Reguere Lake	LAKE	1002	Acres	Active
AZ10010101-000A_05	FISH CREEK/LOWER SALT RIVER BASIN, FISH - S - MG - LR - 10010101A	STREAM	136	Miles	Active

4. Send to ATAINS

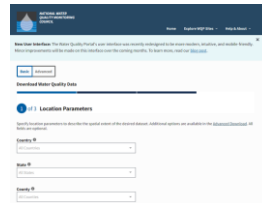
Time = 13 months

New Way



Various  
Databases

1. Gather  
(7 minutes)



Portal

2. Format  
(0 minutes)



3. Assess  
(5 minutes)

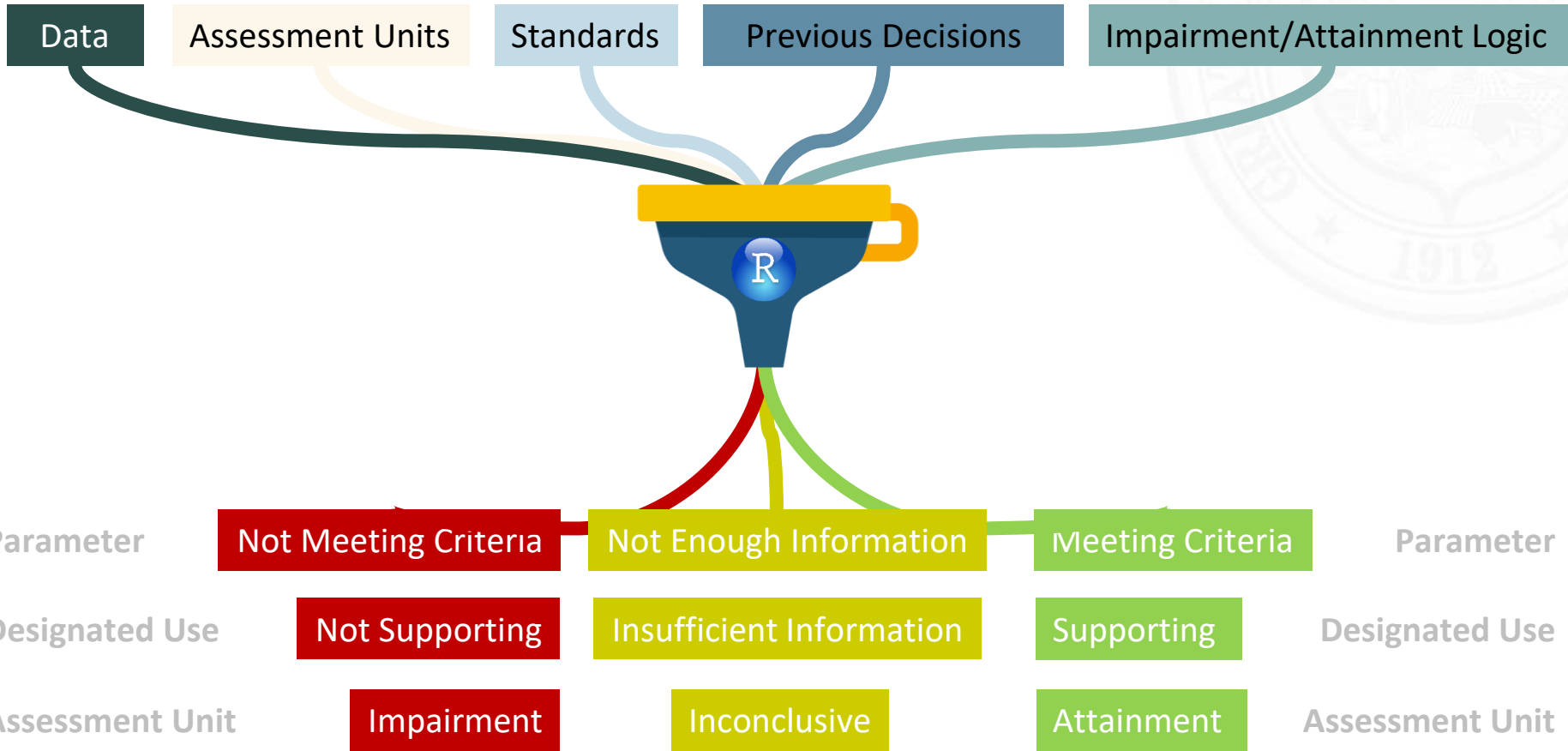
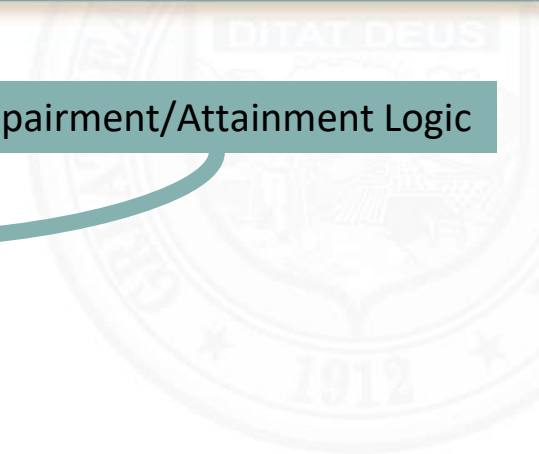


ID	Name	Water Type	Size	Units	Status
AZ10010101-000A_01	NEW RIVER - NHR - S - MG - MG - 10010101	STREAM	231	Miles	Active
AZ10010101-000A_02	Cox Gulch, Fish Headwaters to THREE R CANYON	STREAM	2.3	Miles	Active
AZ10010101-000A_03	Lake Cochise Head Fork, Non-Headwaters to GOVERNMENT SPG	STREAM	9.075	Miles	Active
AZ10010101-000A_04	Reguere Lake	LAKE	1002	Acres	Active
AZ10010101-000A_05	FISH CREEK/LOWER SALT RIVER BASIN, FISH - S - MG - LR - 10010101A	STREAM	136	Miles	Active

4. Send to ATAINS

Time = 12 minutes!

# It is Just Math



**ADEQ Shiny**

Arizona's 2018 Water Quality Assessment Dashboard


**Assessment Results Table:**


Parameter	Assessment Unit	Not Meeting Criteria	Not Enough Information	Meeting Criteria
Population	Population	100	100	100
Use	Use	100	100	100
Waterbody	Waterbody	100	100	100

*Shiny*

# Option 1 - Use ATTAINS

**ATTAINS** - Arizona (21ARIZ)

 (Mr Jason Jones)

 Help



Home



Assessment Units



Assessments



Actions



Reports



Priorities




Surveys



Administration

## Cycle Summary

**Report  
Type\***

Cycle Summary 

Select Report Type

Cycle Summary

**Cycle Comparison**

Assessment Units

Action Summary

**Report\***

**Cycle\***

2024 - Organization Draft (It 

**Run Report**

# Option 1 - Use ATTAINS

## Cycle Comparison

[+ New Report](#)

Assessments

Parameters

Columns

Download

Assessment Unit ID	Assessment Unit Name	Previous Assessment Unit Name	EPA IR Category	Previous EPA IR Category	EPA IR Category Summary
AZ15020001-013A_00	Little Colorado West Fork, from Headwaters to GOVERNMENT SPG	Little Colorado West Fork, from Headwaters to GOVERNMENT SPG	2	2	--
AZ15050301-560_00	Cox Gulch, from Headwaters to THREE R CANYON	Cox Gulch, from Headwaters to THREE R CANYON	4A	4A	--
AZ15070102-006A_00	New River, headwaters - Interstate 17	New River, headwaters - Interstate 17	2	2	--
AZL15060106A-1290_00	Saguaro Lake	Saguaro Lake	2	3	CHANGED
AZ15030104-015_00	COLORADO RIVER	COLORADO RIVER	1	2	CHANGED
AZ15060106A-583_00	Fish Creek, from Headwaters to SALT RIVER	Fish Creek, from Headwaters to SALT RIVER	3	3	--

What observations are in **Table 1** are not in **Table 2**?

## Table 1

a
b
c

## Table 2

a
c
d

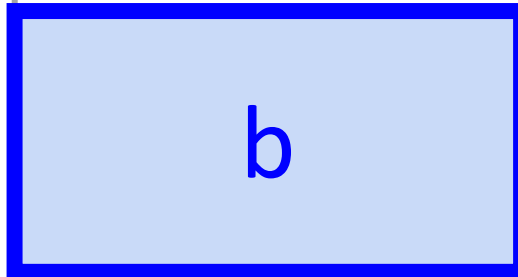
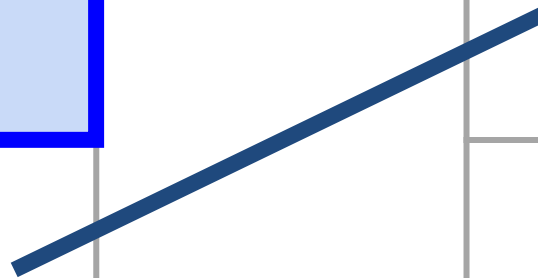
What observations are in **Table 1** are not in **Table 2**?

## Table 1

a
b
c

## Table 2

a
c
d





# Option 2 - Use R

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins Project: (None)

Assessment Table Comparision.R

```
5 library(tidyverse)
6
7 # 1. A working directory is where on your computer R puts things.
8 # 2. Set by navigating to the folder you want to save things by clicking
9 #   the ... in the files pane. Navigate to folder where the files live on y
10 #   It is good practice to put each project you work on in a folder.
11 # 3. Click 'More' in the file pane and 'set as working directory'.
12 # 4. Copy this code into your script (don't copy the '>').
13
14 # CHANGE THE CODE BELOW AND SET YOUR WORKING DIRECTORY
15 setwd("~/R Presnetation/antijoin example-20230516T191222Z-001/antijoin examp
16
17
18 #### QUESTION 1 - NEW DELISTS/IMPAIRMENTS ####
19
20 # Import Impairments
21 j.imp_2026_par <- read_csv("j.imp_2026_par.csv")
22 j.imp_2024_par <- read_csv("j.imp_2024_par.csv")
23
```

Environment History Connections Tutorial

Import Dataset 209 MiB

R Global Environment

Data

j.imp_2024_par	293 obs. of 2 variables
j.imp_2026_par	290 obs. of 2 variables

Files Plots Packages Help Viewer Presentation

New Folder New Blank File Delete Rename More

Home > R Presnetation > antijoin example-20230516T191222Z-001

Name	Size	Modified
..		
antijoin example		

22:49 # QUESTION 1 - NEW DELISTS/IMPAIRMENTS R Script

# Option 2 - Use R

Assessment Table Comparison.R x j.imp\_2024\_par x j.imp\_2026\_par x

Filter

WBID	CharacteristicName
1 14070006-001	SELENIUM
2 14070006-1130	MERCURY IN FISH TISSUE
3 14070007-123	ESCHERICHIA COLI
4 14070007-123	SELENIUM
5 14070007-123	SUSPENDED SEDIMENT CONCENTRATION (SSC)
6 15010001-003	SELENIUM
7 15010001-005	SELENIUM
8 15010001-006	SELENIUM
9 15010001-008	SELENIUM
10 15010001-010	SELENIUM
11 15010001-011	SELENIUM
12 15010001-022	SELENIUM
13 15010002-001	SELENIUM
14 15010002-003	SELENIUM
15 15010002-003	SUSPENDED SEDIMENT CONCENTRATION (SSC)
16 15010002-004	SELENIUM
17 15010002-007	SELENIUM
18 15010002-009	SELENIUM
19 15010002-012	SELENIUM
20 15010002-013	SELENIUM
21 15010002-020B	SELENIUM
22 15010002-318	SELENIUM
23 15010002-871	SELENIUM
24 15010002-871	SELENIUM

Showing 1 to 25 of 293 entries, 2 total columns

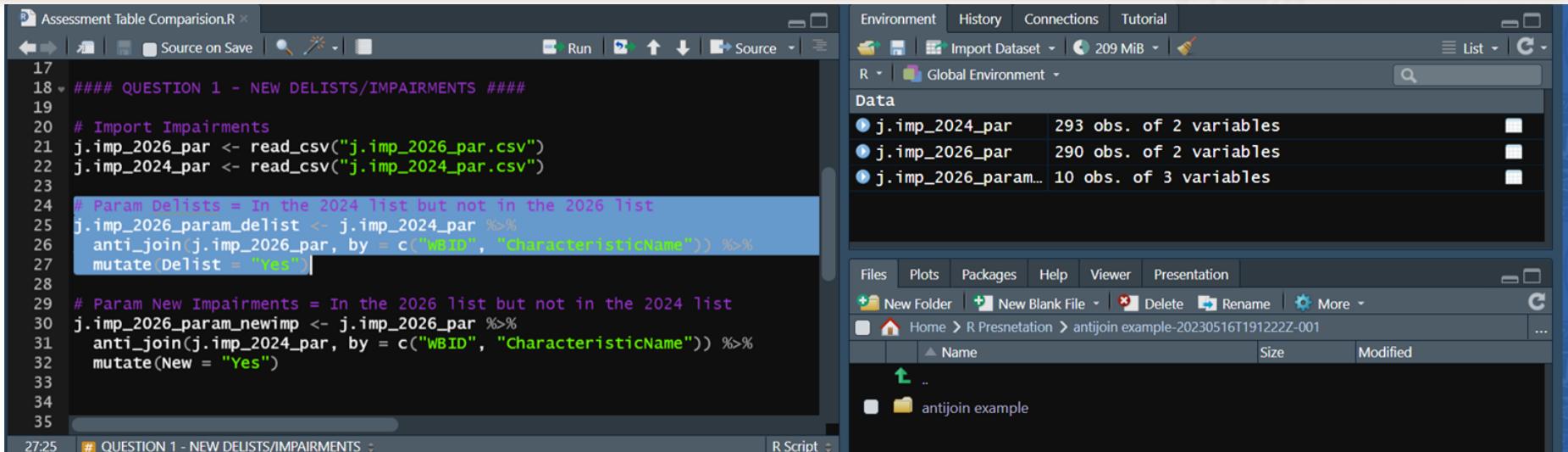
Assessment Table Comparison.R x j.imp\_2024\_par x j.imp\_2026\_par x

Filter

WBID	CharacteristicName
1 14070006-001	SELENIUM
2 14070006-1130	MERCURY IN FISH TISSUE
3 14070007-123	ESCHERICHIA COLI
4 14070007-123	SELENIUM
5 14070007-123	SUSPENDED SEDIMENT CONCENTRATION (SSC)
6 15010001-003	SELENIUM
7 15010001-005	SELENIUM
8 15010001-006	SELENIUM
9 15010001-008	SELENIUM
10 15010001-010	SELENIUM
11 15010001-011	SELENIUM
12 15010001-022	SELENIUM
13 15010002-001	SELENIUM
14 15010002-003	SELENIUM
15 15010002-003	SUSPENDED SEDIMENT CONCENTRATION (SSC)
16 15010002-004	SELENIUM
17 15010002-007	SELENIUM
18 15010002-009	SELENIUM
19 15010002-012	SELENIUM
20 15010002-013	SELENIUM
21 15010002-020B	SELENIUM
22 15010002-318	SELENIUM
23 15010002-871	SELENIUM
24 15010002-871	SELENIUM

Showing 1 to 25 of 290 entries, 2 total columns

# Option 2 - Use R

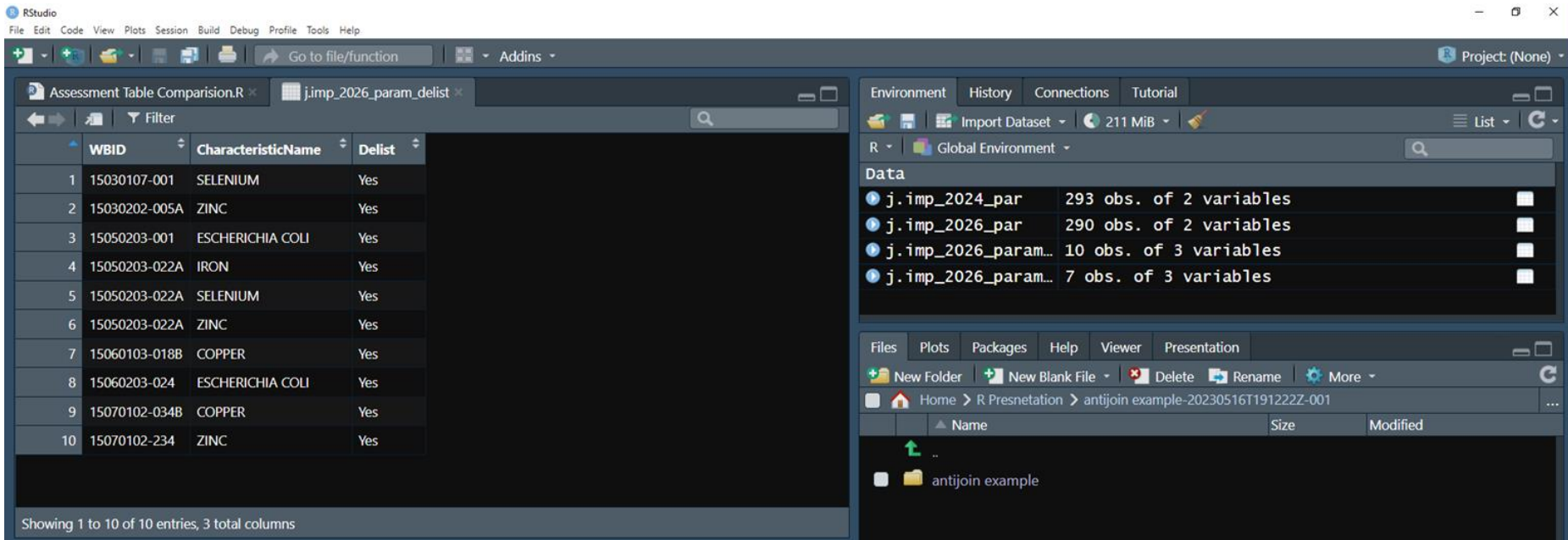


```
17
18 ### QUESTION 1 - NEW DELISTS/IMPAIRMENTS ###
19
20 # Import Impairments
21 j.imp_2026_par <- read_csv("j.imp_2026_par.csv")
22 j.imp_2024_par <- read_csv("j.imp_2024_par.csv")
23
24 # Param Delists = In the 2024 list but not in the 2026 list
25 j.imp_2026_param_delist <- j.imp_2024_par %>%
26   anti_join(j.imp_2026_par, by = c("WBID", "CharacteristicName")) %>%
27   mutate(Delist = "Yes")
28
29 # Param New Impairments = In the 2026 list but not in the 2024 list
30 j.imp_2026_param_newimp <- j.imp_2026_par %>%
31   anti_join(j.imp_2024_par, by = c("WBID", "CharacteristicName")) %>%
32   mutate(New = "Yes")
33
34
35
```

Environment: Global Environment (209 MiB)

Data:

- j.imp\_2024\_par: 293 obs. of 2 variables
- j.imp\_2026\_par: 290 obs. of 2 variables
- j.imp\_2026\_param...: 10 obs. of 3 variables



	WBID	CharacteristicName	Delist
1	15030107-001	SELENIUM	Yes
2	15030202-005A	ZINC	Yes
3	15050203-001	ESCHERICHIA COLI	Yes
4	15050203-022A	IRON	Yes
5	15050203-022A	SELENIUM	Yes
6	15050203-022A	ZINC	Yes
7	15060103-018B	COPPER	Yes
8	15060203-024	ESCHERICHIA COLI	Yes
9	15070102-034B	COPPER	Yes
10	15070102-234	ZINC	Yes

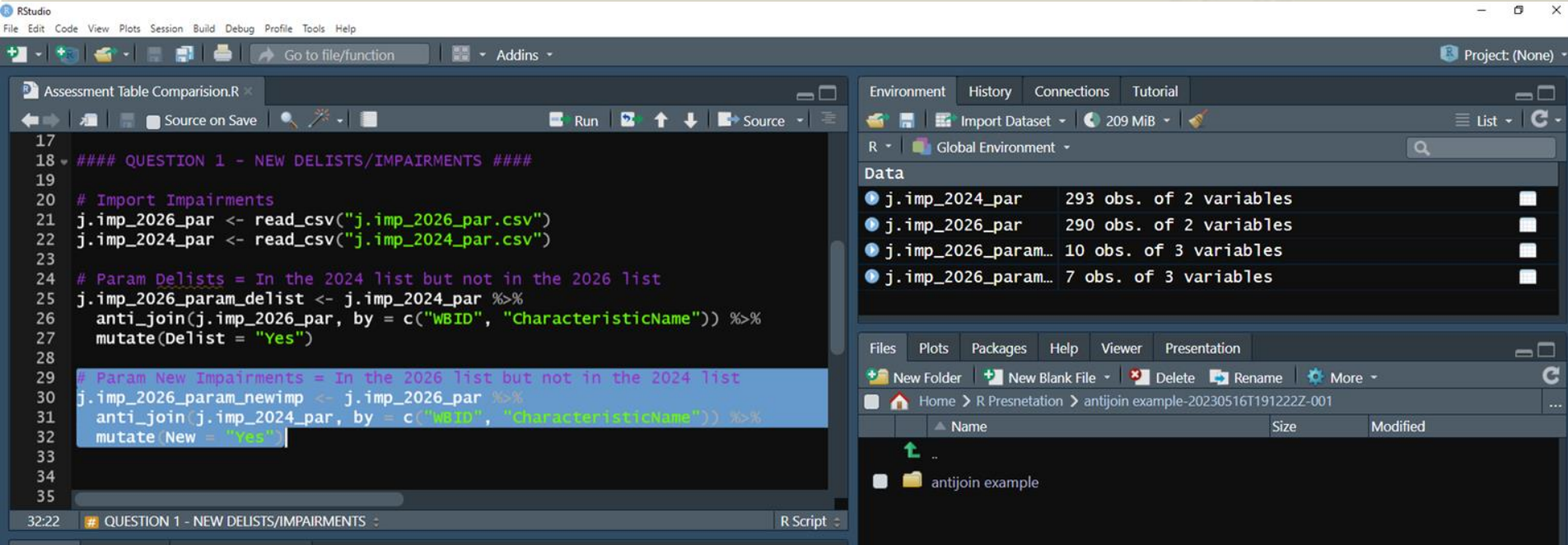
Showing 1 to 10 of 10 entries, 3 total columns

Environment: Global Environment (211 MiB)

Data:

- j.imp\_2024\_par: 293 obs. of 2 variables
- j.imp\_2026\_par: 290 obs. of 2 variables
- j.imp\_2026\_param...: 10 obs. of 3 variables
- j.imp\_2026\_param...: 7 obs. of 3 variables

# Option 2 - Use R

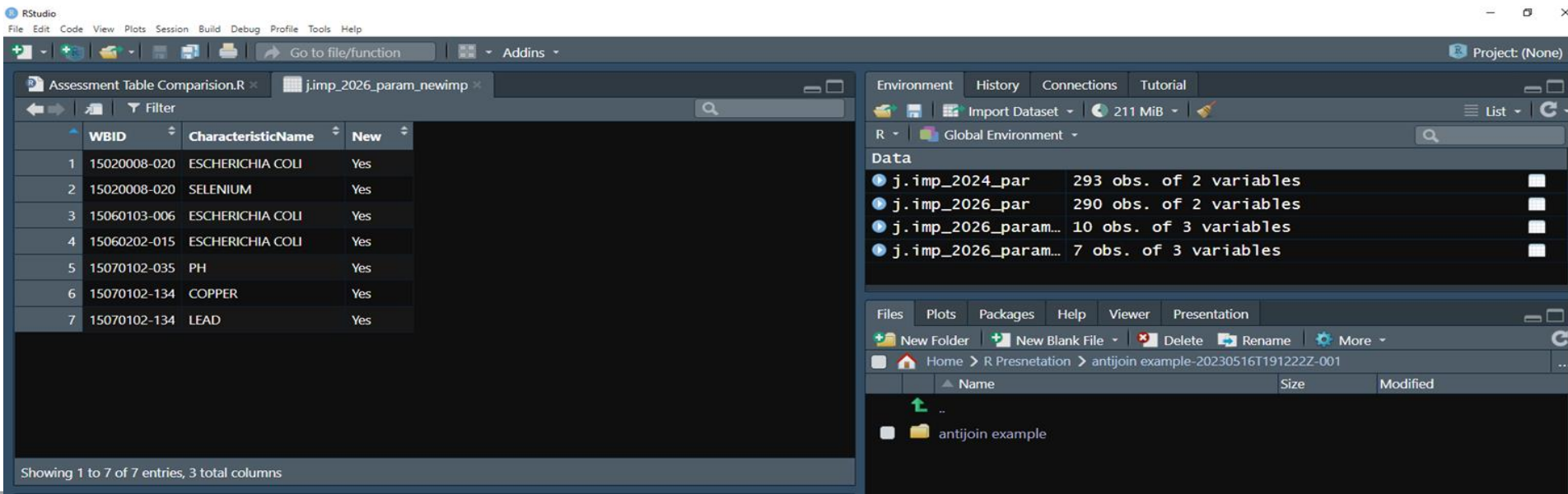


RStudio interface showing the R script editor with the following code:

```
17
18 ##### QUESTION 1 - NEW DELISTS/IMPAIRMENTS #####
19
20 # Import Impairments
21 j.imp_2026_par <- read_csv("j.imp_2026_par.csv")
22 j.imp_2024_par <- read_csv("j.imp_2024_par.csv")
23
24 # Param Delists = In the 2024 list but not in the 2026 list
25 j.imp_2026_param_delist <- j.imp_2024_par %>%
26   anti_join(j.imp_2026_par, by = c("WBID", "CharacteristicName")) %>%
27   mutate(Delist = "Yes")
28
29 # Param New Impairments = In the 2026 list but not in the 2024 list
30 j.imp_2026_param_newimp <- j.imp_2026_par %>%
31   anti_join(j.imp_2024_par, by = c("WBID", "CharacteristicName")) %>%
32   mutate(New = "Yes")
33
34
35
```

The Environment pane shows the following data objects:

Object	Observations	Variables
j.imp_2024_par	293	2
j.imp_2026_par	290	2
j.imp_2026_param...	10	3
j.imp_2026_param...	7	3



RStudio interface showing the Data Viewer pane with the following table:

	WBID	CharacteristicName	New
1	15020008-020	ESCHERICHIA COLI	Yes
2	15020008-020	SELENIUM	Yes
3	15060103-006	ESCHERICHIA COLI	Yes
4	15060202-015	ESCHERICHIA COLI	Yes
5	15070102-035	PH	Yes
6	15070102-134	COPPER	Yes
7	15070102-134	LEAD	Yes

Showing 1 to 7 of 7 entries, 3 total columns



**Questions?**

**Jason Jones**

**[jdj@azdeq.gov](mailto:jdj@azdeq.gov)**

**Mikayla Baker**

**[baker.mikayla@azdeq.gov](mailto:baker.mikayla@azdeq.gov)**