From R Code to Data Driven Decisions

Jason Jones, Zac White, Valeria Bocanegra, Mackenzie Moore and Matthew Robinson

February 27, 2024

ADEQ Arizona Department of Environmental Quality

Clean Air, Safe Water, Healthy Land for Everyone

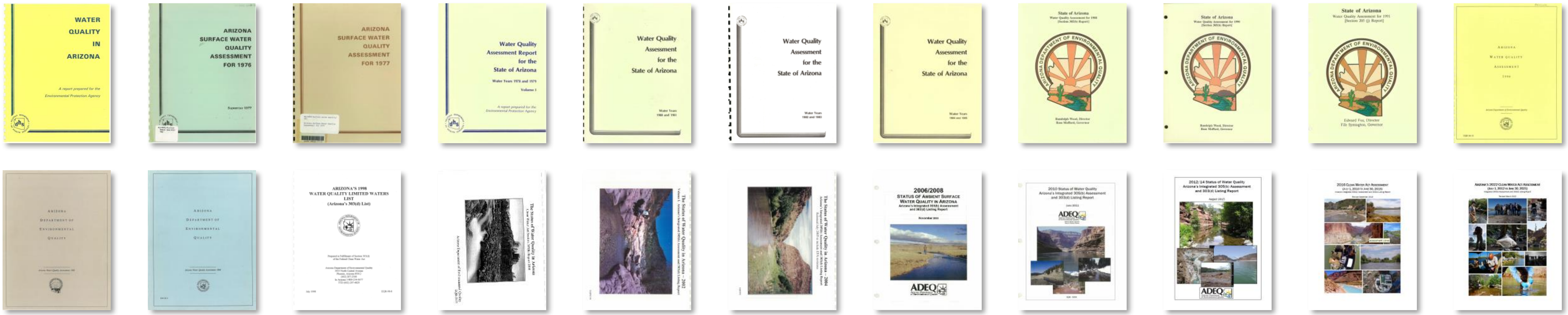
What is R ?

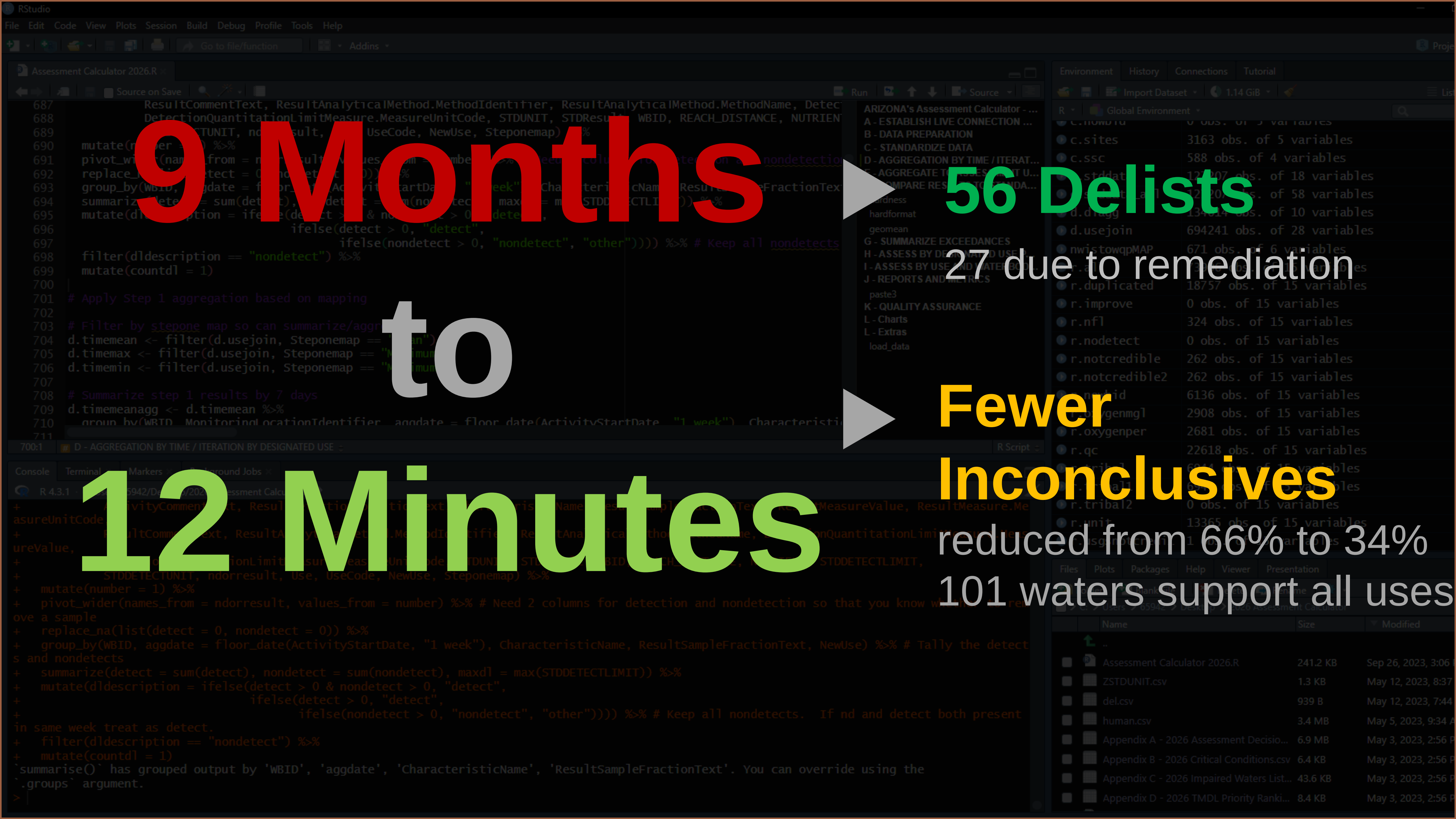
R is an advanced set of tools for data analysis.

Clean Water Act Assessment

ADEQ
Arizona Department of Environmental Quality

- The **Assessment** takes **raw water quality data** and determines if **standards are being met**
- Running the assessment took us over **9 months** (that is with a tool to help us do the math!)
- In 2018, the tool we had stopped working




9 Months
to
12 Minutes
56 Delists
27 due to remediation
Fewer Inconclusives
reduced from 66% to 34%
101 waters support all uses

ADEQ Arizona Department of Environmental Quality Arizona's Water Quality Assessment Dashboard - 2026 Cycle

Overview | Metrics | Impairments | Decisions | Datagaps | Help | Source Code

Welcome

DISCLAIMER - THIS DASHBOARD DISPLAYS ASSESSMENT INFORMATION. THIS DATA IS PROVISIONAL UNTIL EPA APPROVES THE OFFICIAL ASSESSMENT. OFFICIAL ASSESSMENTS ARE GENERALLY APPROVED EVERY EVEN YEAR AND ARE DUE TO EPA BY APRIL 1ST.

The 2026 assessment cycle includes data from 7/1/2019 to 6/30/2024. New data is added weekly.

Assumptions and detailed instructions for using the dashboard can be found by clicking the 'Help' page or by clicking HERE.

Status: **Not approved by EPA; Assessment Window Open**

Last Day Assessment Tool Ran: **2023-12-12**

Number of Records (formatted and aggregated): **118369**

Number of Provisional Impaired Waters (All): **186** (Click the Impairments page for more information)

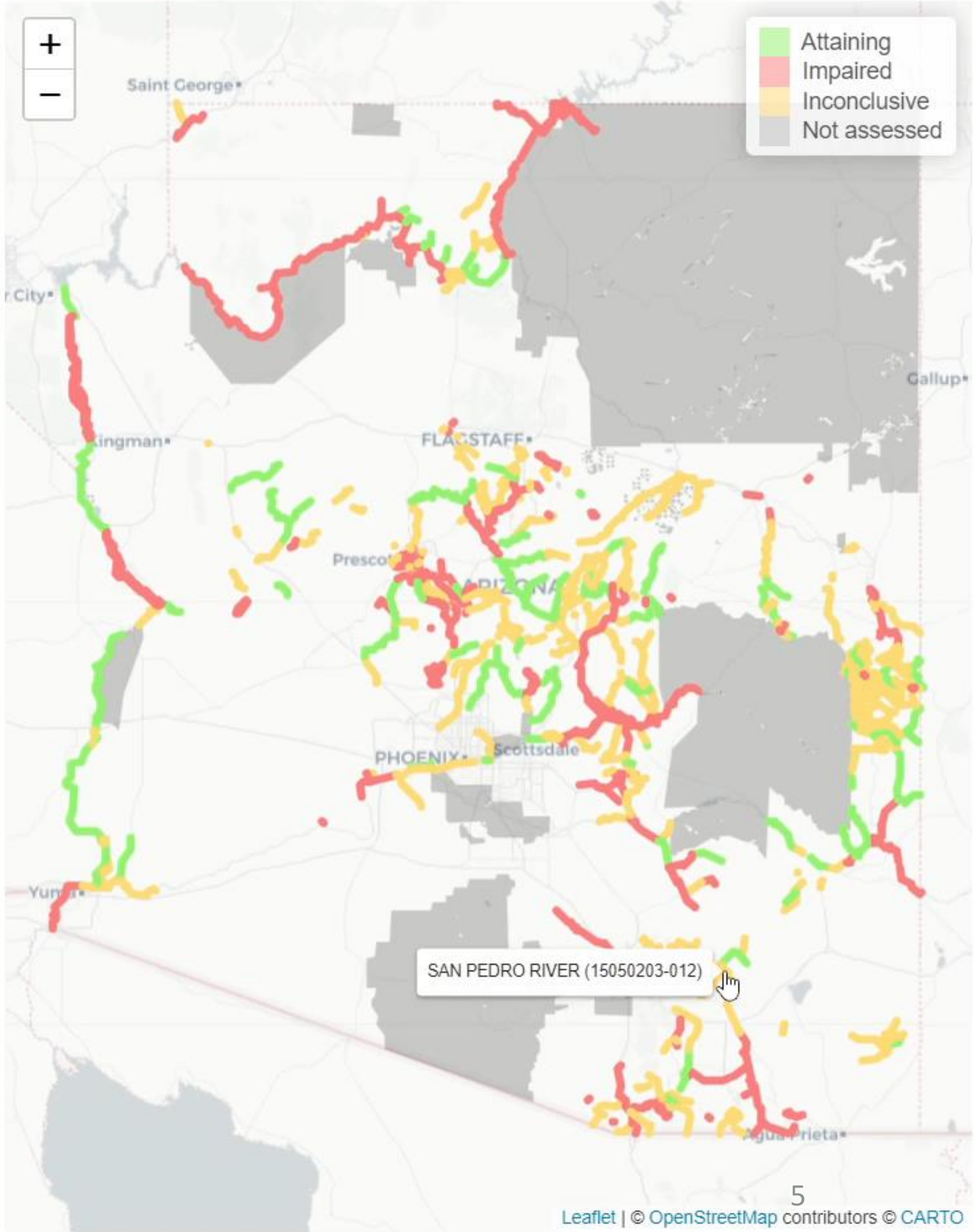
Number of Exceedances: **1691**. (Click the Decisions page for more information)

Number of Inconclusives (Use): **731**. (See below and/or click the Datagap page for more information)

Provisional Decisions at Each Assessment Level

Level	Attaining-Supporting-Meeting	Impaired-Not Supporting-Not Meeting	Inconclusive
Parameter	18549	346	6128
Use	1139	259	731
Waterbody	101	186	263

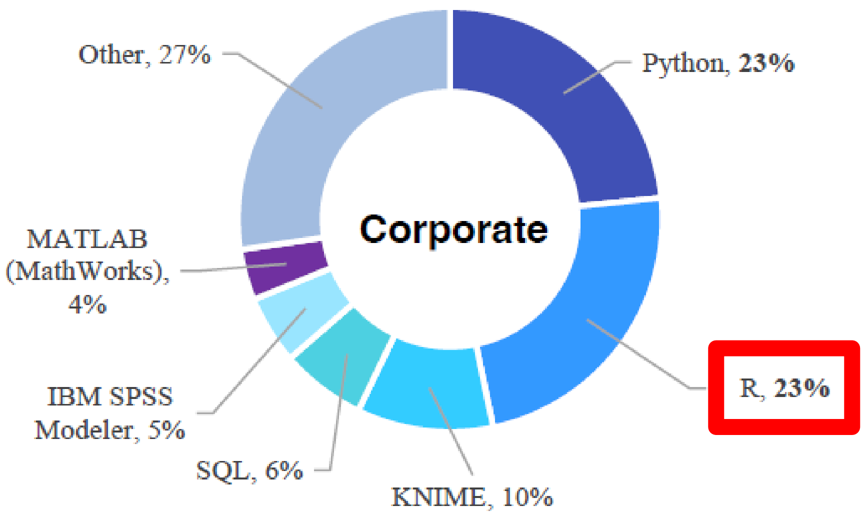
Map of Provisional Assessment Waterbody Decisions

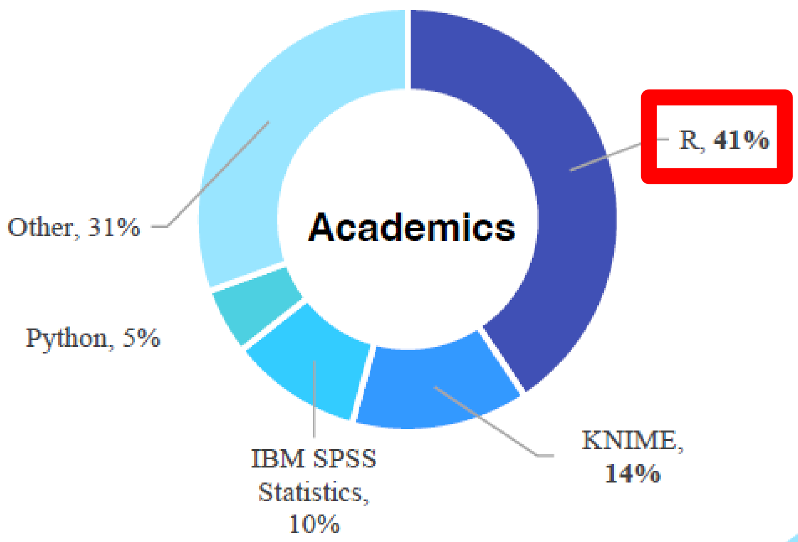


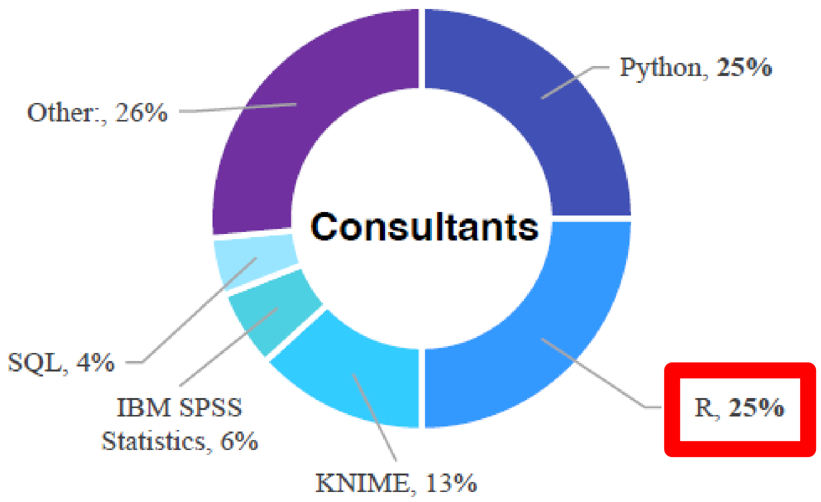
Why R ?

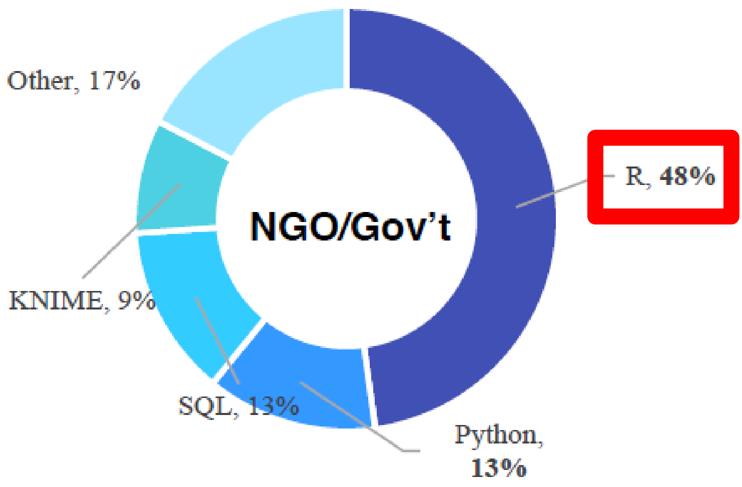
- Free
- Sharable
- Reproducible
- Large Community

Survey Says!









Source: 2020 Survey from www.RexerAnalytics.com

Training Program

ADEQ
Arizona Department of Environmental Quality

Data Science

datacamp

\+

Project

=

Results

Environmental Scientist
Water Quality Division
Surface Water Quality Improvement Section
Sampling and Source Identification Unit

Zac White

ADEQ
Arizona Department of Environmental Quality

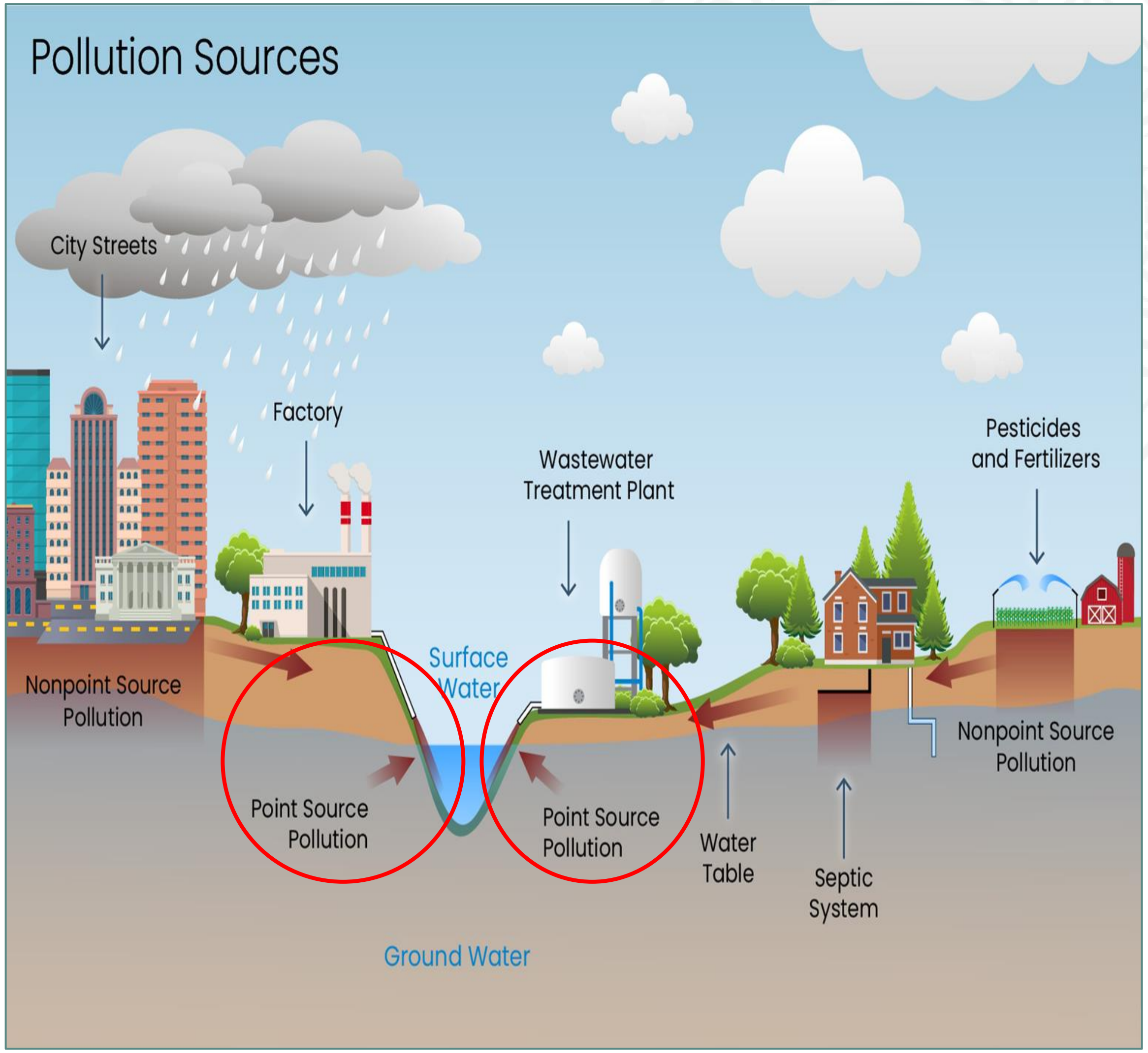
Question

What % of all pollutant discharges identified in Arizona Total Maximum Daily Load (TMDL) reports are Point Source (PS) vs Nonpoint Sources (NPS)?

Exploring relationships between pollutant discharges

$$TMDL = \sum WLA + \sum LA + MOS$$

- TMDL = Allowable Total Maximum Daily Load
- **WLA** = Point Source load allocation
- **LA** = Nonpoint Source load allocation
- MOS = Margin of Safety



https://www.bcwater.org/threats/

Challenges faced during the project:

- New R user
- New dataset / data exploration

- Trial and error in analysis

- Intervals of hitting walls

Exploring relationships between pollutant discharges

Evolution of project phases:

1. Imported datasets from ADEQ and EPA
1. Data quality control
1. Joined data into one dataframe
1. Filtered NA and duplicate data
1. Standardized units
1. Totaled LA and WLA data
1. Visualized results w/ graphs

Exploring relationships between pollutant discharges

ADEQ Arizona Department of Environmental Quality

-Specialized data science package

- Comprehensive vignette explaining commands used in package

The tidyverse

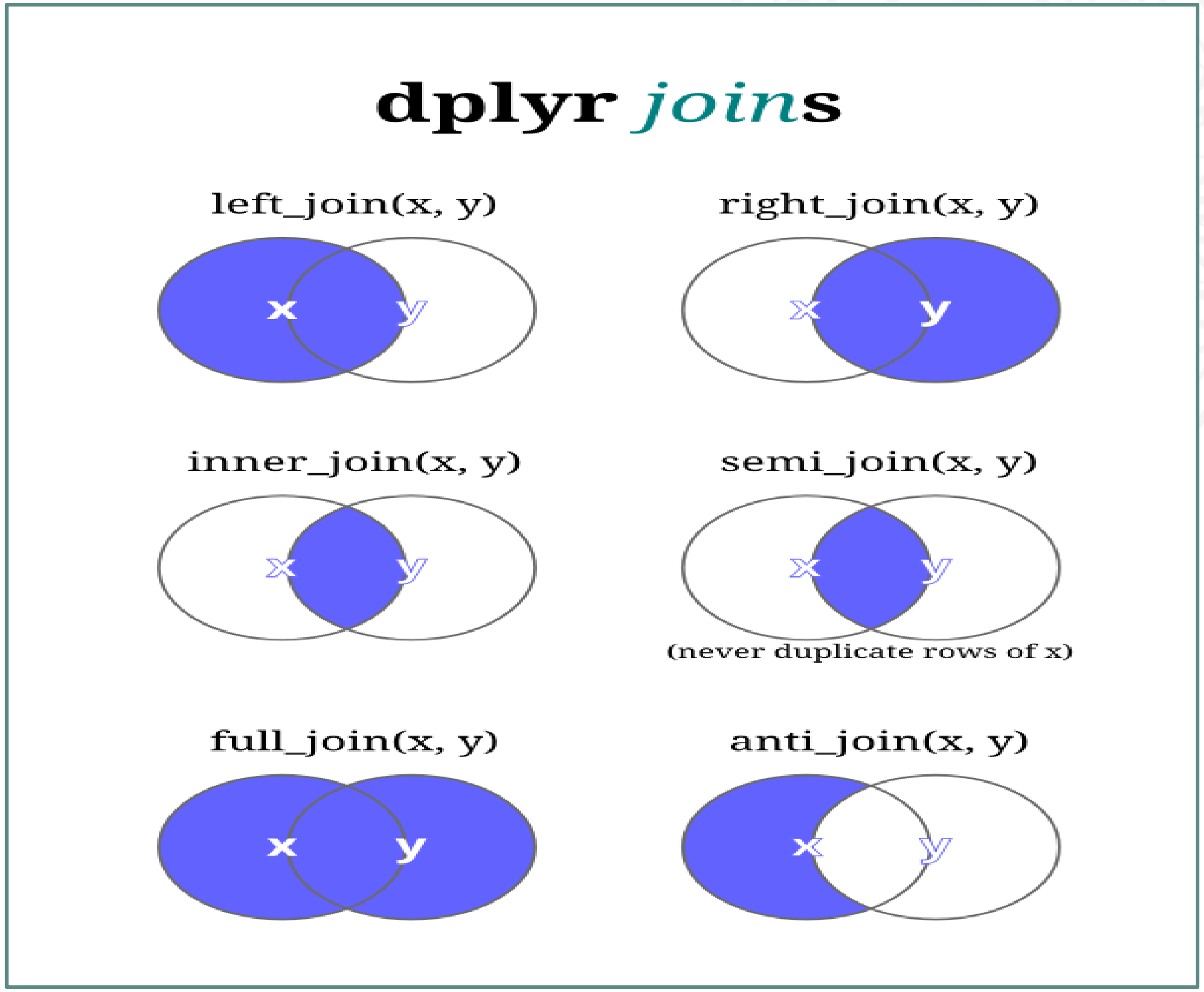
Components

The tidyverse is a collection of R packages that share common philosophies and are designed to work together. This site is a work-in-progress guide to the tidyverse and its packages.

https://hbctraining.github.io/Intro-to-R/lessons/08_intro_tidyverse.html

Exploring relationships between pollutant discharges

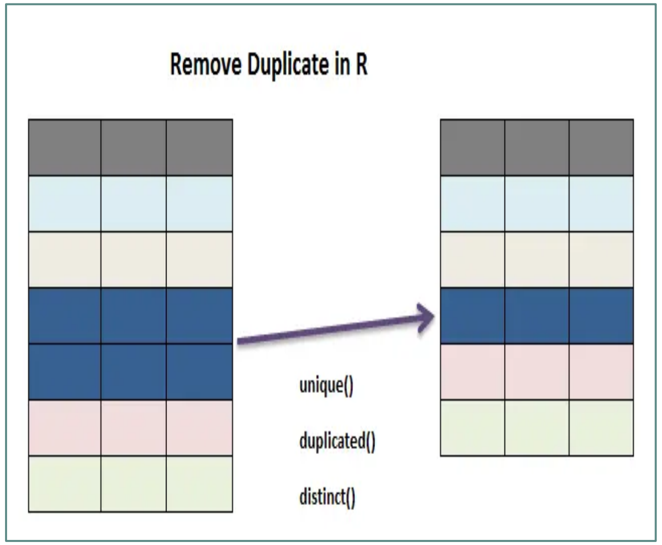
- Compare internal ADEQ database to EPA database
- table joins (anti / inner)

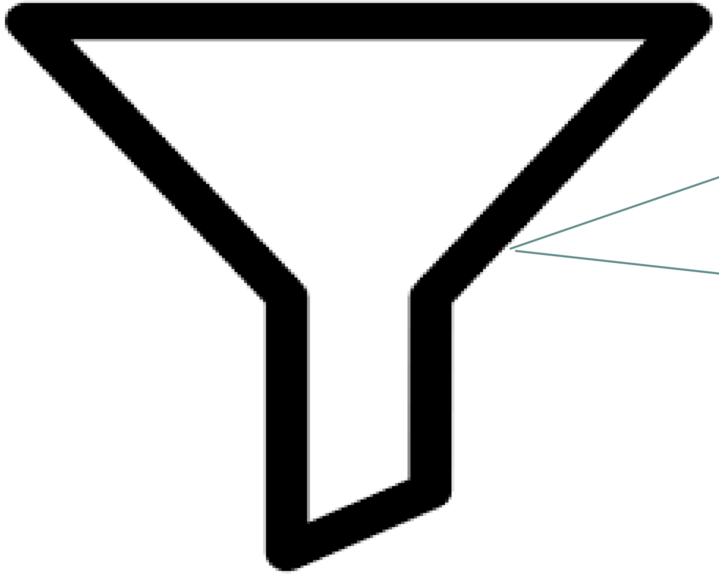


Exploring relationships between pollutant discharges

ADEQ Arizona Department of Environmental Quality

26,000 rows raw data





filter( )

distinct( )

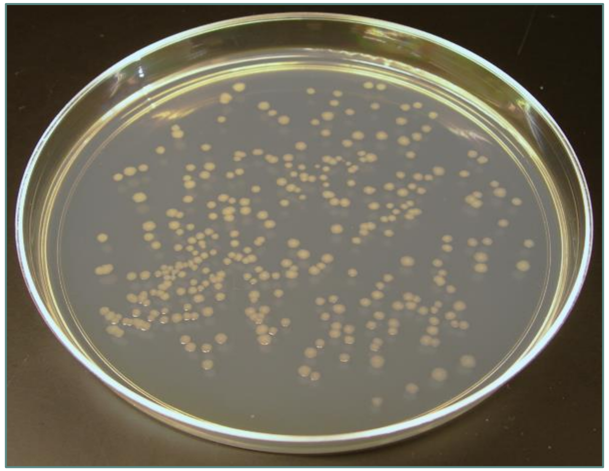
2,500 rows of usable data

dplyr

Exploring relationships between pollutant discharges

Standardize results w/ *units* package:

- E. coli CFUs → G-CFUs
- "X"-grams –>Kg



Totals for WLA and LA :

-sum( ) WLA and LA columns

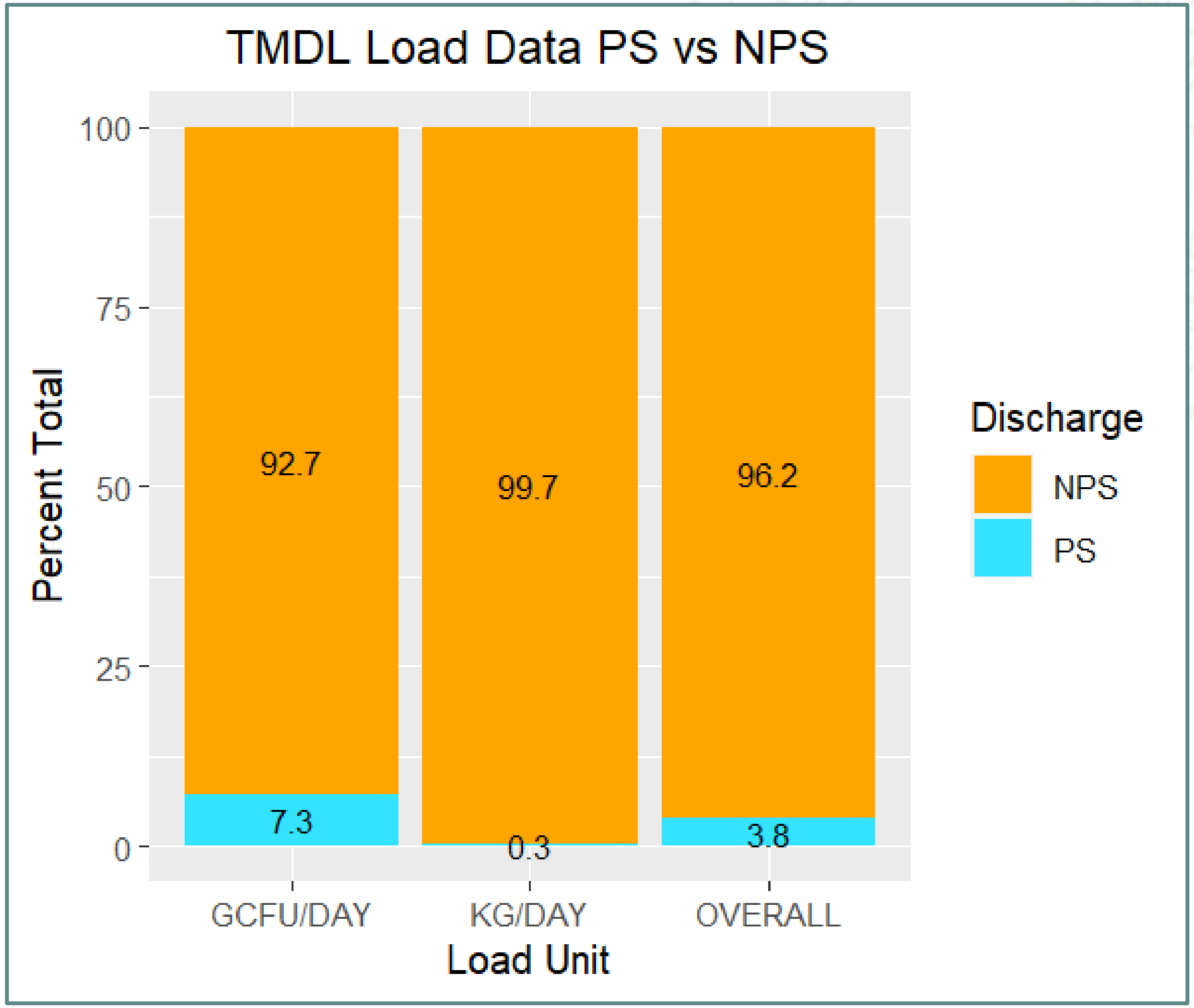
*% WLA = WLA/ LA *100*

% LA = 100 - %WLA

Exploring relationships between pollutant discharges

Visualize results

- flexible graphing tool
- easy to use, hard to master



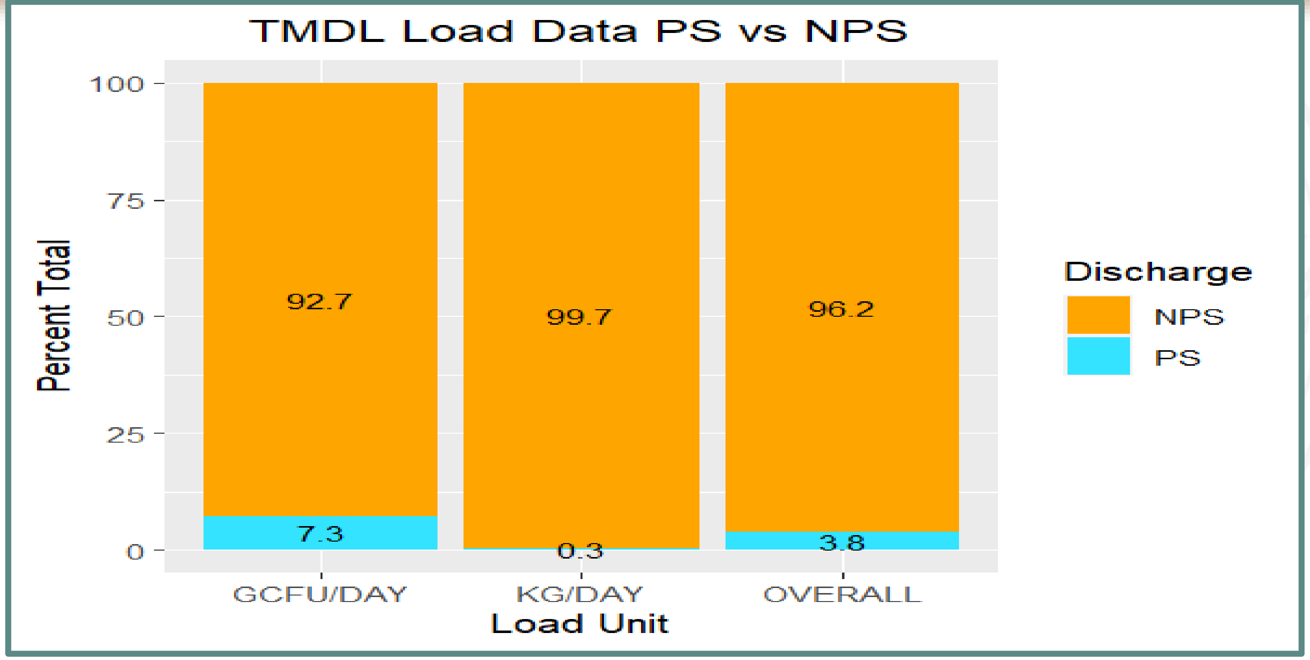
Exploring relationships between pollutant discharges

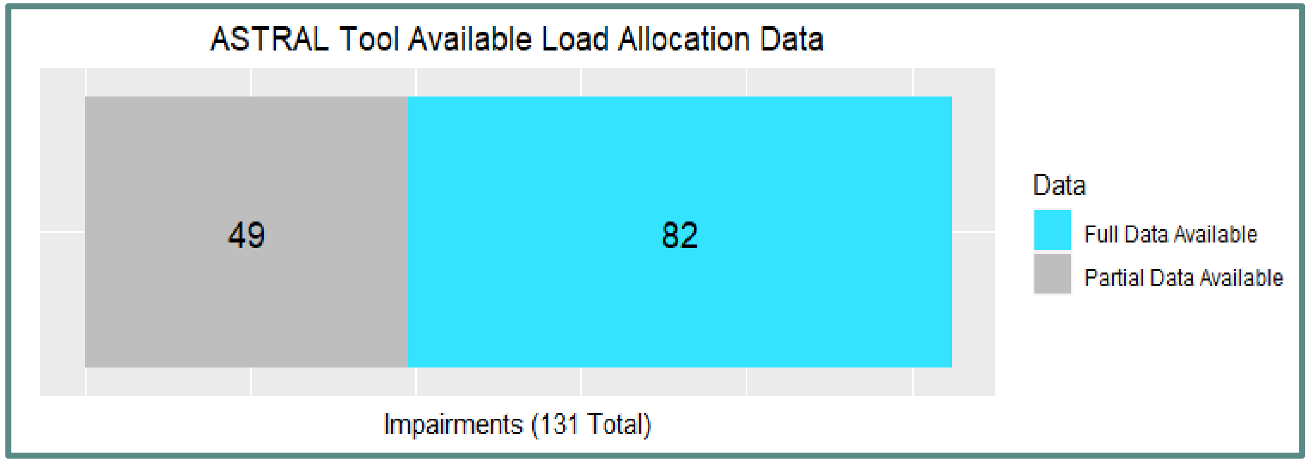
Question

What % of all TMDL loads are Point Source vs Nonpoint Sources?

Results

- **96 % Nonpoint Source loading**
- Data driven decisions
  - Permits and NPS teams
  - address data gaps





Environmental Scientist
Water Quality Division
Surface Water Quality Improvement Section
Sampling and Source Identification Unit

Valeria Bocanegra

Arizona Department of Environmental Quality

Statistical Analysis of Copper Content in Water Bodies

Questions

1. Is there a difference between storm samples and baseflow samples?
    a. To answer this question I used the Wilcoxon Rank sum test to test for significant difference.
2. Does sampling during stormflow conditions increase the chances of an exceedance?
    a. To answer this question I used the Chi-squared test of independence.

R Packages Leveraged for Analysis

ADEQ
Arizona Department
of Environmental Quality

dataRetrieval

- Collection of functions for pulling data from USGS and EPA water quality and hydrology data from the Water Quality Portal.

dplyr

- Package that provides a consistent set of verbs for data manipulation.

ggplot2

- A package that allows you to create graphics based on the data provided and aesthetics you choose

Retrieve Water Quality Data

```
#User Inputs
start.date <- "2010-01-01"
end.date <- "2023-12-01"
a.wqp <- readWQPdata(statecode = "Arizona", characteristicName = "Copper",
startDate = start.date, endDate = end.date)
a.wqp_sel <- a.wqp %>%
  select("MonitoringLocationIdentifier", "ActivityCommentText",
         "ResultMeasure.MeasureUnitCode", "ResultMeasureValue",
         "ResultSampleFractionText", "CharacteristicName",
         "ProjectIdentifier", "ActivityStartDate",
         "ActivityStartTime.Time", "ActivityMediaName",
         "ActivityIdentifier", "ActivityTypeCode", "ResultCommentText",
         "DetectionQuantitationLimitMeasure.MeasureValue",
         "DetectionQuantitationLimitMeasure.MeasureUnitCode") %>%
  mutate(ResultMeasureValue = as.numeric(ResultMeasureValue)) %>%
  filter(CharacteristicName == "Copper", ActivityMediaName == "Water") %>%
  filter(ResultMeasure.MeasureUnitCode != "mg/kg") %>%
  #filter(ActivityCommentText != "NA") %>%
  mutate(`ResultMeasure.MeasureUnitCode` = ifelse(
    `ResultMeasure.MeasureUnitCode` == "ppb", "ug/L", `ResultMeasure
.MeasureUnitCode`)) %>%
  mutate(`ResultMeasure.MeasureUnitCode` = ifelse(
    `ResultMeasure.MeasureUnitCode` == "ug/l", "ug/L", `ResultMeasure
.MeasureUnitCode`)) %>%
  filter(ActivityTypeCode == "Sample-Routine")
```

- dataRetrieval
  - readWQPdata
- dplyr
  - select()
  - mutate()
  - filter()

Wilcoxon Rank Sum Test for Significant Difference

ADEQ
Arizona Department of Environmental Quality

```r
c.diss.stat <- c.cu.diss %>%
  mutate(logresult = log(ResultMeasureValue))
# Wilcoxon test to measure statistical difference between two groups
wilcox.test(logresult ~ Storm, data = c.diss.stat, conf.int = TRUE)
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  logresult by Storm
## W = 273779, p-value < 0.00000000000000022
## alternative hypothesis: true location shift is not equal to 0
## 95 percent confidence interval:
##  -2.259985 -1.972432
## sample estimates:
## difference in location
##             -2.120206
```

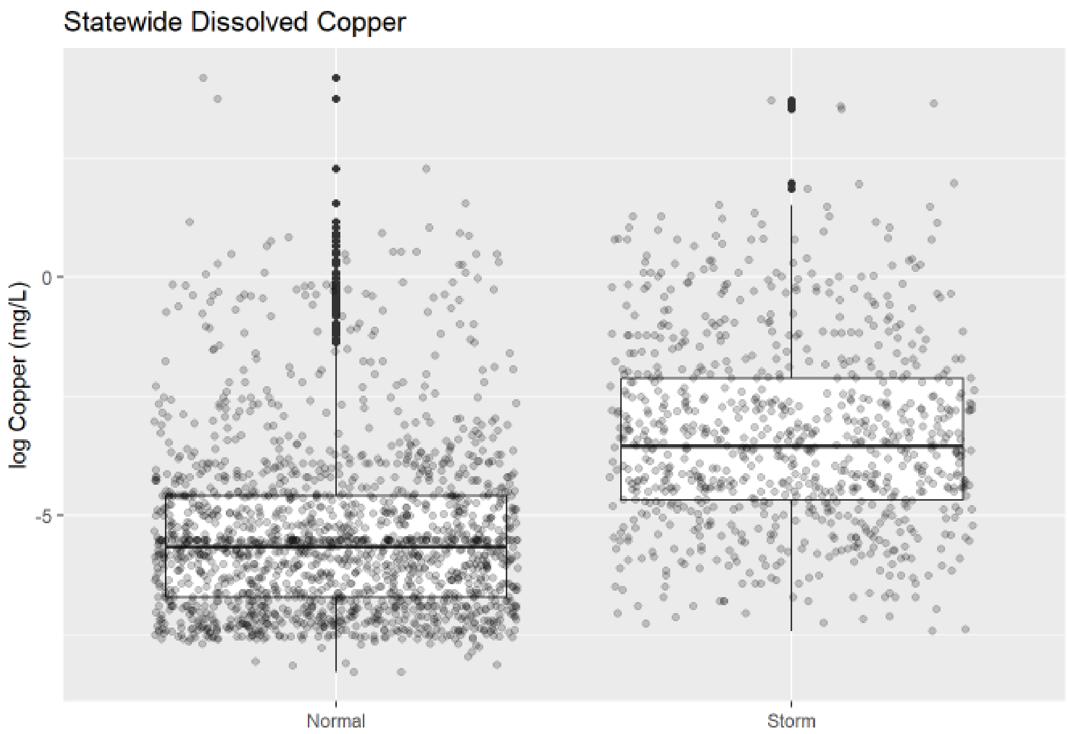
```r
c.total.stat <- c.cu.total %>%
  mutate(logresult = log(ResultMeasureValue))
# Wilcoxon test to measure statistical difference between two groups
wilcox.test(logresult ~ Storm, data = c.total.stat, conf.int = TRUE)
```

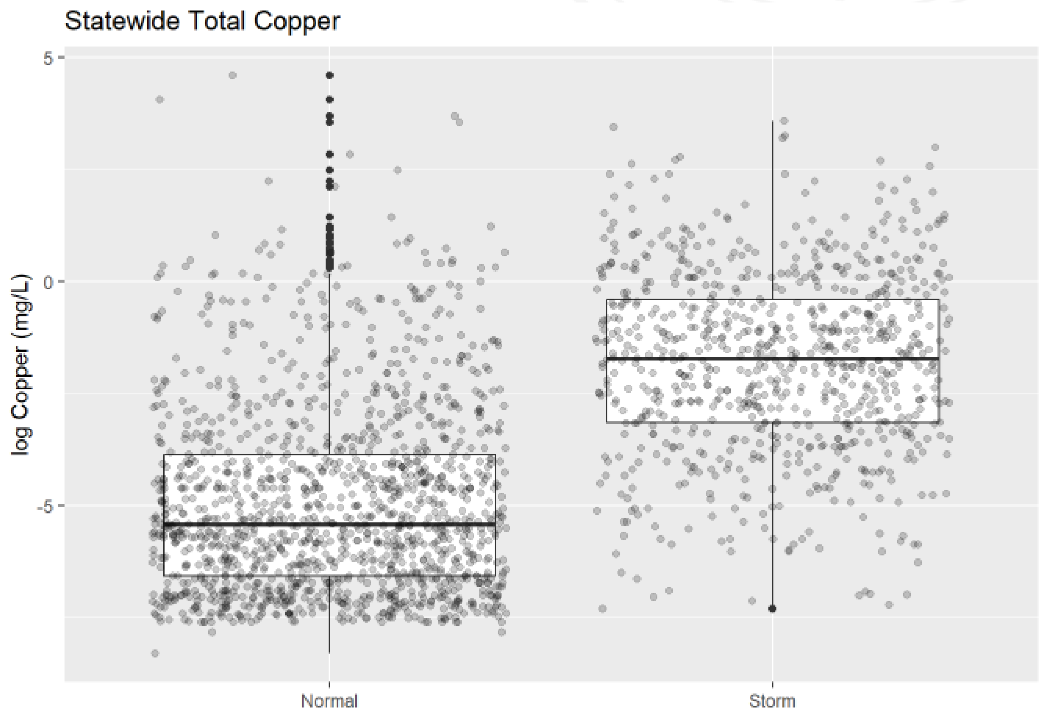
```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  logresult by Storm
## W = 154327, p-value < 0.00000000000000022
## alternative hypothesis: true location shift is not equal to 0
## 95 percent confidence interval:
##  -3.560668 -3.202771
## sample estimates:
## difference in location
##             -3.383441
```

- p-values < 0.05
- "true location shift is not equal to 0. "
- Sample sets are significantly different.

Visualizing Results

```r
# Dissolved Wilcoxon
c.cu.diss %>%
  ggplot(aes(x = Storm, y = log(ResultMeasureValue))) +
  geom_boxplot() +
  geom_jitter(alpha = 0.2) +
  labs(x = "", y = "log Copper (mg/L)", title = "Statewide Dissolved Copper")
```





Chi-squared test of independence

Table 3.0 Statewide Dissolved Copper Exceedance v. Flow

	Normal	Storm
No	1494	391
Yes	196	450

```
# Statewide Copper
# Dissolved
chisq.test(e.a.con.diss)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  e.a.con.diss
## X-squared = 516.68, df = 1, p-value < 0.00000000000000022
```

Categorical variables: Exceedance and Flow

- p-values < 0.05
- Variables are dependent.

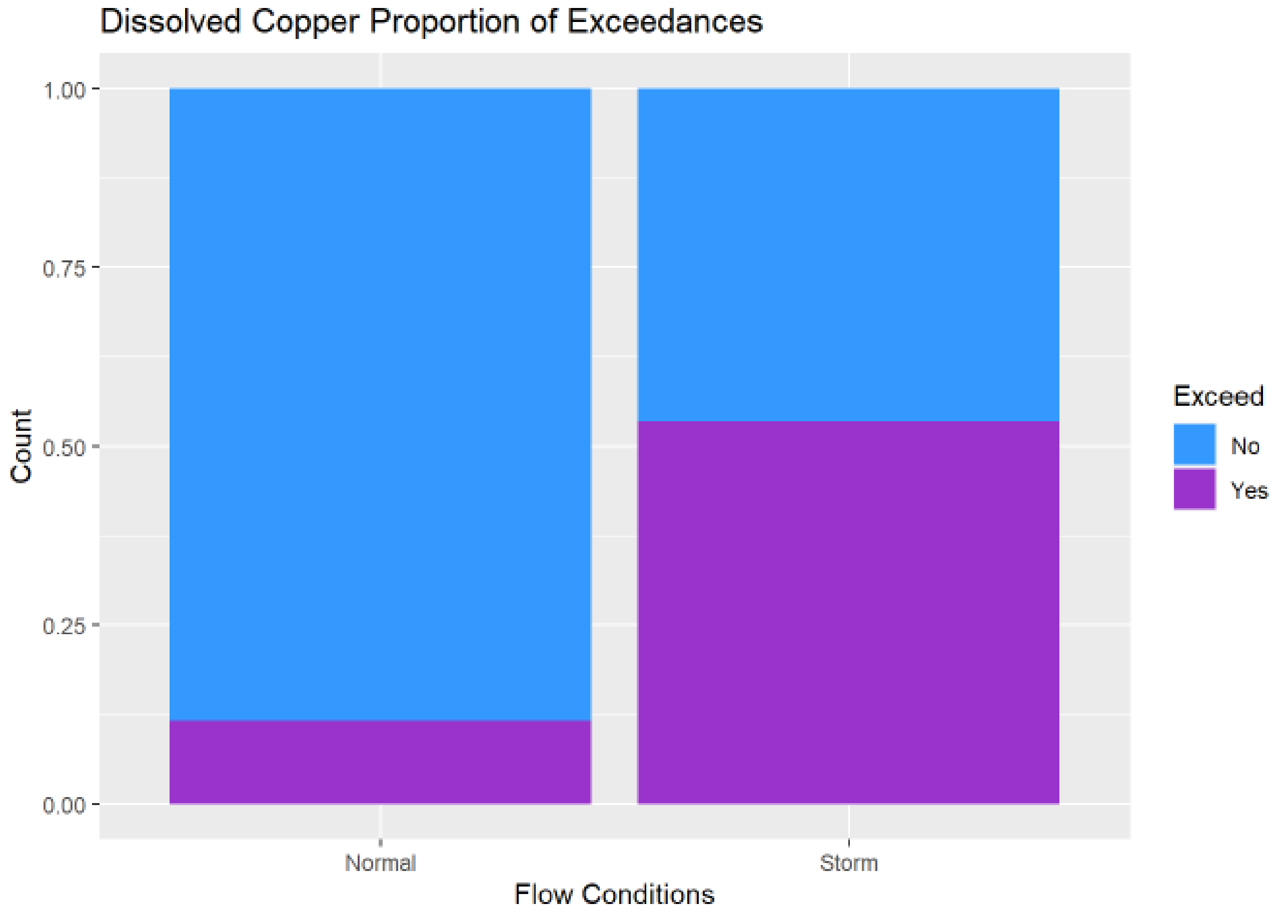
Table 4.0 Statewide Total Copper Exceedance v. Flow

	Normal	Storm
No	1338	676
Yes	28	114

```
# Statewide Copper
# Total
chisq.test(e.a.con.total)
```

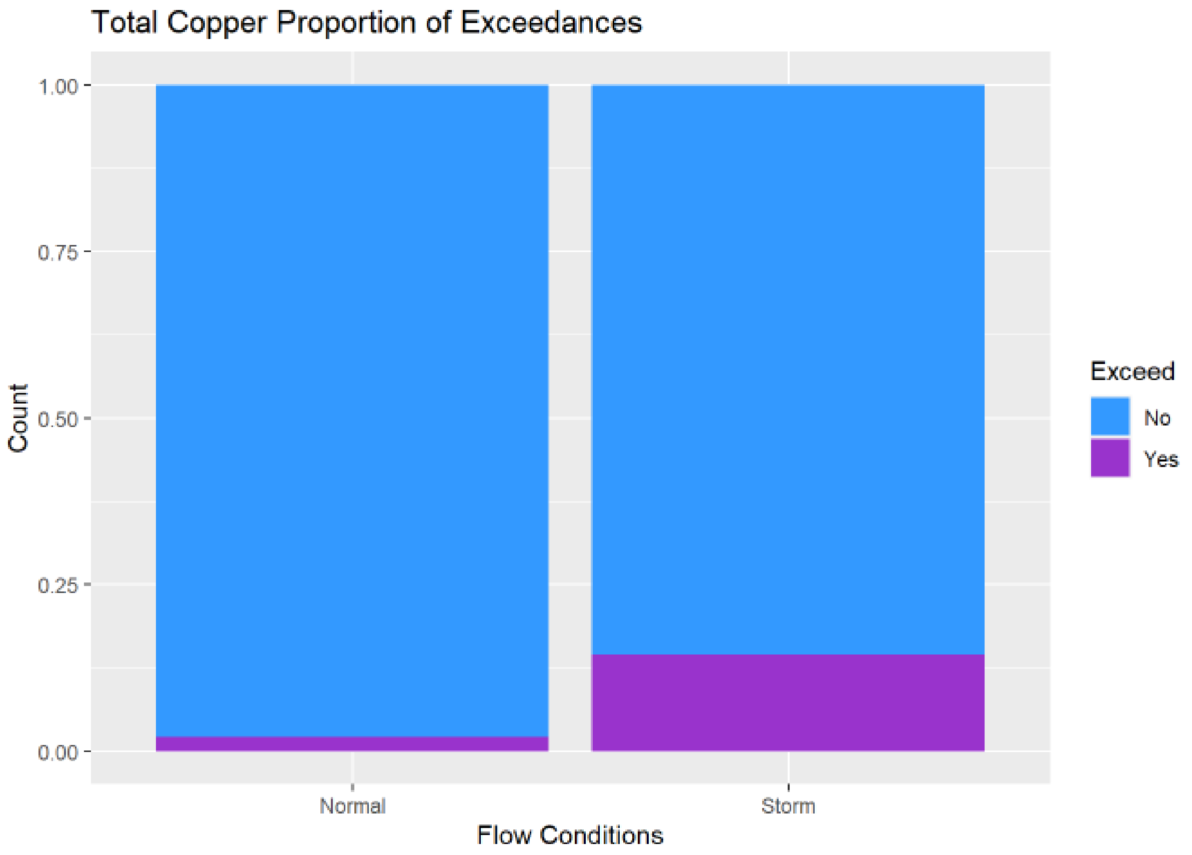
```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  e.a.con.total
## X-squared = 122.69, df = 1, p-value < 0.00000000000000022
```

Analyze Results



> 50% storm samples exceed the standard

More storm samples exceed than base flow samples.



Outcomes

Analysis

1. Significant difference in copper content between storm and baseflow samples.
2. Exceedances are not independent of flow.
3. Storm events introduce bias to the data and are not representative of normal conditions.

Environmental Scientist
Water Quality Division
Surface Water Quality Improvement Section
Standards and Assessment Unit

Mackenzie Moore

ADEQ
Arizona Department of Environmental Quality

Measuring EJ Considerations in TMDLs

Question

- Was there a bias in how ADEQ has historically chosen which impairments to complete Total Maximum Daily Load (TMDL) reports for?

Analysis

- Environmental Justice Parameters: Percent Low Income & Percent Minority
- Assessed Categories: Impaired with TMDL, Impaired without TMDL, Assessed – Not Impaired (Attaining or Inconclusive)

How We Leveraged R for the TMDL & EJ Analysis

tidyverse	ggplot	leaflet
Match EJ parameters from distinct HUC12s to ADEQ’s assessment units	Create boxplots for analysis of EJ parameters to waters with completed TMDLs	Map the assessed waterbodies with percent low income and percent minority data shown spatially

Comparing Assessment Decisions to EJ Data

ADEQ
Arizona Department of Environmental Quality

WBID	WaterbodyName	WaterbodyDecision	F__Low_Inc	F__Minorit	Median_Low_Inc	Median_Minorit	max_Low_Inc	max_Minorit
14070006-001	COLORADO RIVER	Impaired	40.69	69.14	41.775	50.670	42.86	69.14
15010001-003	COLORADO RIVER	Impaired	0.00	0.00	100.000	100.000	100.00	100.00
15010001-005	COLORADO RIVER	Impaired	0.00	0.00	50.000	50.000	100.00	100.00
15010001-006	COLORADO RIVER	Impaired	100.00	100.00	100.000	100.000	100.00	100.00
15010001-008	COLORADO RIVER	Impaired	62.50	100.00	64.585	100.000	100.00	100.00
15010001-010	COLORADO RIVER	Impaired	50.00	100.00	58.335	100.000	66.67	100.00
15010001-011	COLORADO RIVER	Impaired	33.33	33.33	50.000	66.665	66.67	100.00
15010001-022	COLORADO RIVER	Impaired	66.67	100.00	42.860	45.710	66.67	100.00
15010002-001	COLORADO RIVER	Impaired	0.00	0.00	0.000	0.000	0.00	0.00
15010002-004	COLORADO RIVER	Impaired	0.00	0.00	0.000	0.000	0.00	0.00
15010002-007	COLORADO RIVER	Impaired	0.00	0.00	0.000	0.000	0.00	0.00
15010002-009	COLORADO RIVER	Impaired	0.00	0.00	0.000	0.000	20.00	20.00
15010002-012	COLORADO RIVER	Impaired	0.00	0.00	0.000	0.000	0.00	0.00
15010002-013	COLORADO RIVER	Impaired	0.00	0.00	0.000	0.000	25.00	40.00
15010002-020B	HERMIT CREEK	Impaired	0.00	100.00	12.500	70.000	25.00	100.00
15010002-318	SPRING CANYON CREEK	Impaired	100.00	0.00	50.000	0.000	100.00	0.00
15010002-871	ROYAL ARCH CREEK	Impaired	0.00	0.00	0.000	0.000	0.00	0.00

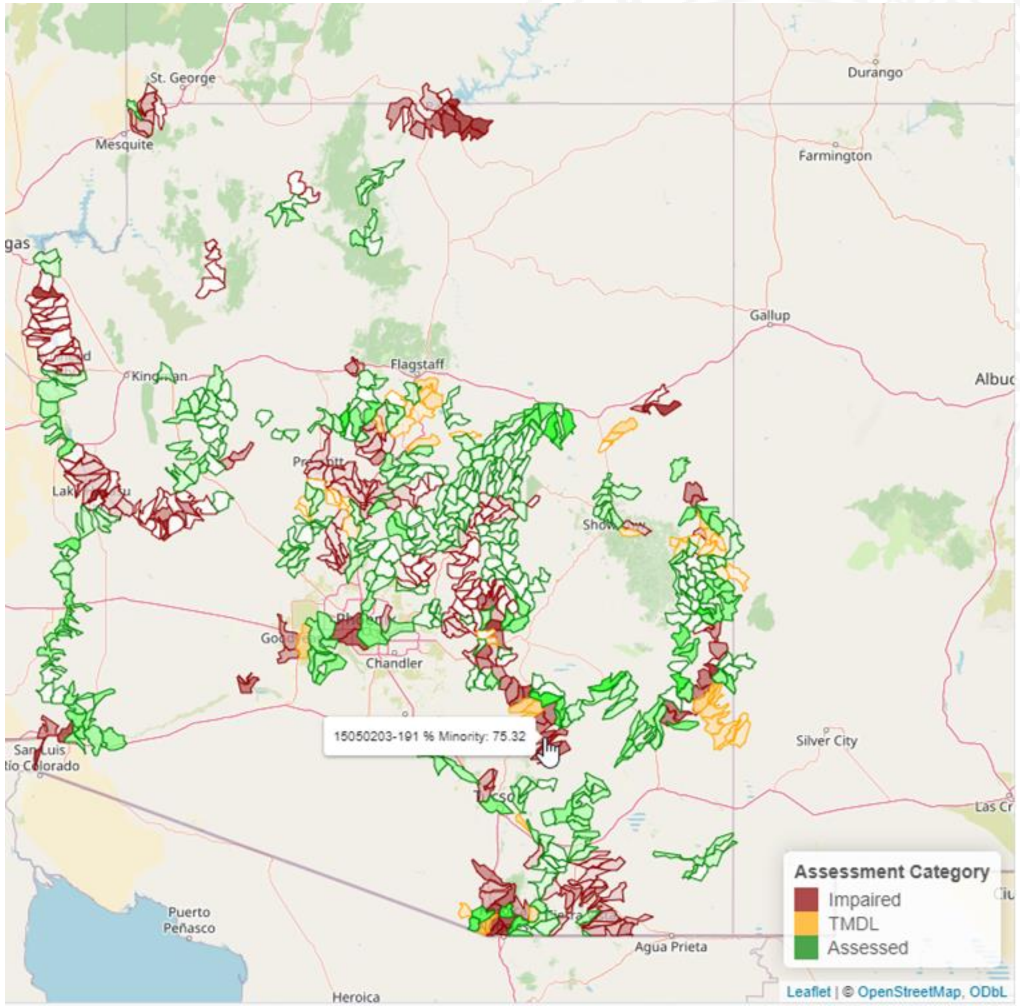
But where are these reaches?

There must be a better way to visualize this data…

Mapping EJ Parameters and Assessments in R

```
leaflet() %>%
  addTiles() %>%
  addPolygons(
    data = ejmap_imp,
    weight=1,
    opacity = 1,
    col = 'darkred',
    fillOpacity = 0.7,
    label = ~label_Income,
    fillColor = ~pal1(PCT_LOW_INCOME)) %>%
  addPolygons(
    data = ejmap_tmdl,
    weight=1,
    opacity = 1,
    col = 'orange',
    fillOpacity = 0.7,
    label = ~label_Income,
    fillColor = ~pal2(PCT_LOW_INCOME)) %>%
  addPolygons(
    data = ejmap_assessed,
    weight=1,
    opacity = 1,
    col = 'green',
    fillOpacity = 0.7,
    label = ~label_Income,
    fillColor = ~pal3(PCT_LOW_INCOME)) %>%
  addLegend(position = "bottomright",
            colors = (c("darkred", "orange", "green")),
            labels = c("Impaired", "TMDL", "Assessed"),
            title = "Assessment Category",
            opacity = 0.7)
# Define colors or each layer.  Each layer is under 'addPolygons()'.
pal4 <- colorNumeric(
  palette = colorRampPalette(c('white', 'darkred'))(length(ejmap_imp$PCT_MINORITY)),
  domain = ejmap_imp$PCT_MINORITY)

pal5 <- colorNumeric(
  palette = colorRampPalette(c('white', 'orange'))(length(ejmap_imp$PCT_MINORITY)),
  domain = ejmap_imp$PCT_MINORITY)
```



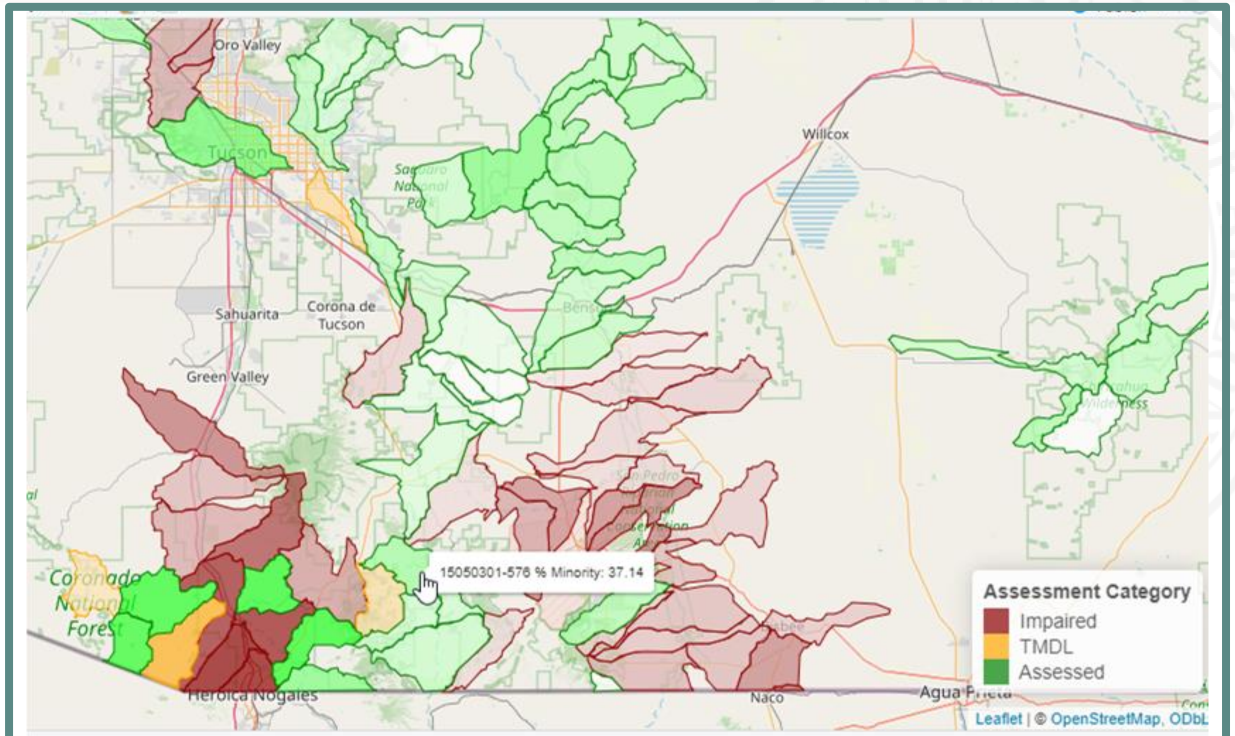
Leaflet - Coding a Custom Map

Colors representing each assessment decision:

```
# Define colors or each layer.  Each layer is under 'addPolygons()'.
pal1 <- colorNumeric(
  palette = colorRampPalette(c('white', 'darkred'))(length(ejmap_imp$PCT_LOW_INCOME)),
  domain = ejmap_imp$PCT_LOW_INCOME)

pal2 <- colorNumeric(
  palette = colorRampPalette(c('white', 'orange'))(length(ejmap_imp$PCT_LOW_INCOME)),
  domain = ejmap_imp$PCT_LOW_INCOME)

pal3 <- colorNumeric(
  palette = colorRampPalette(c('white', 'green'))(length(ejmap_imp$PCT_LOW_INCOME)),
  domain = ejmap_imp$PCT_LOW_INCOME)
```



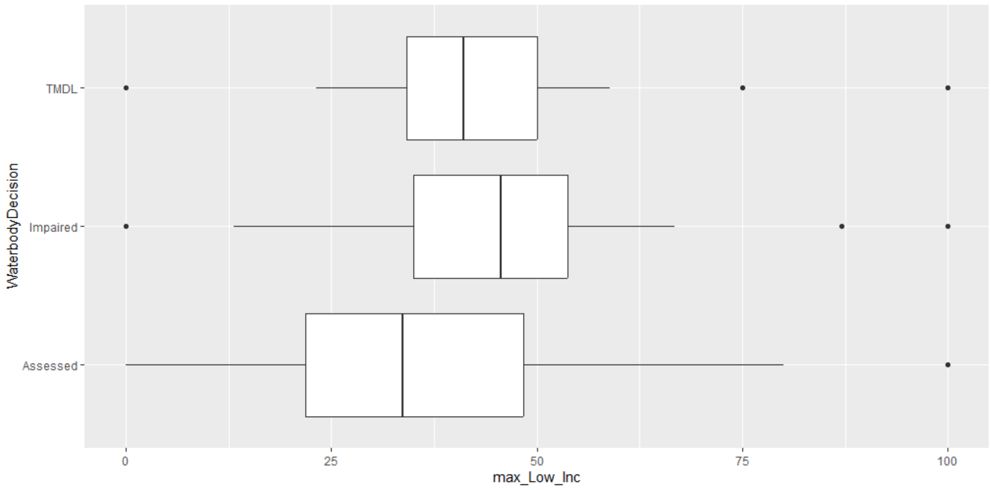
Intensity of color based on the EJ parameter percentage:

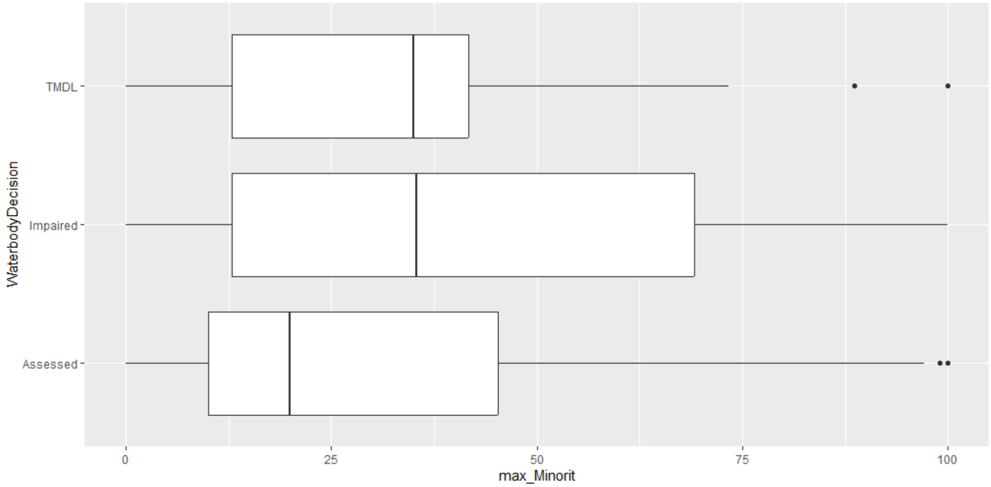
```
# This defines some 'bins' and colors.
bins <- c(0, 50, 100)
pal <- colorBin("YlOrRd", domain = ejmap$PCT_LOW_INCOME, bins = bins)
```

The leaflet map allows viewers to see each assessed waterbody's location, the assessment decision, and the EJ parameters all at once

Outcomes

- Determined that **there was not a bias** for or against communities with low income or high percent minority in selecting which impaired waters to complete TMDLS for
- Created a product to **easily visualize** where impairments are in relation to EJ populations
- Environmental Justice is included in TMDL Prioritization Tool as a weighted category for **intentional consideration**



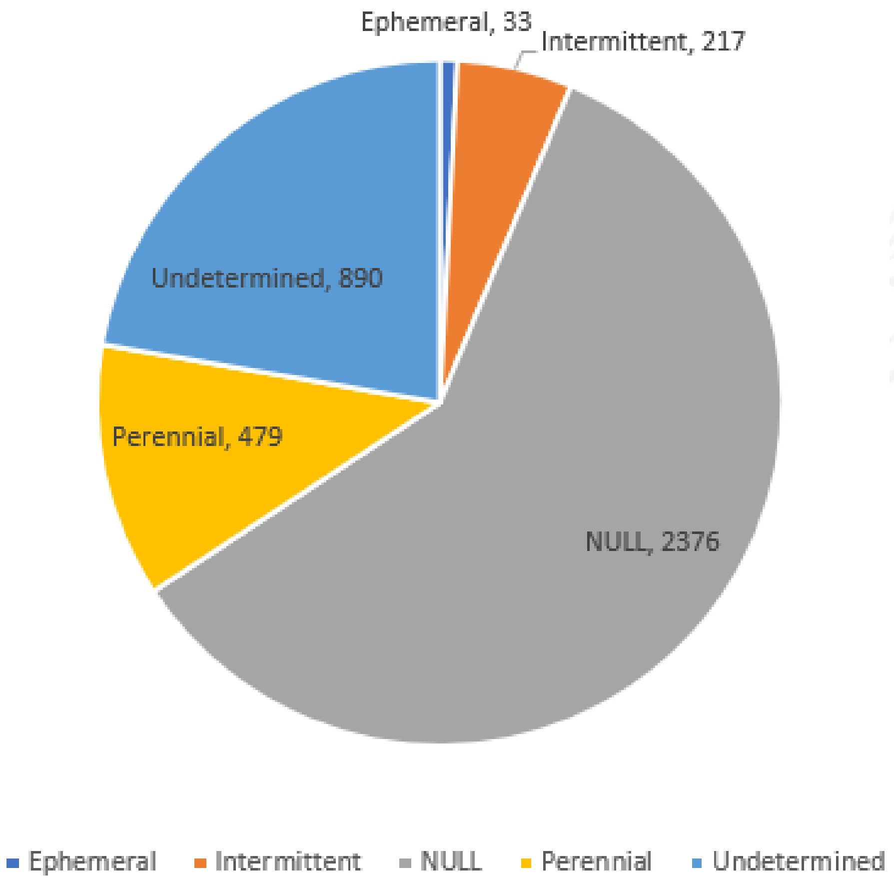


Environmental Scientist
Water Quality Division
Surface Water Quality Improvement Section
Standards and Assessment Unit

Matthew Robinson

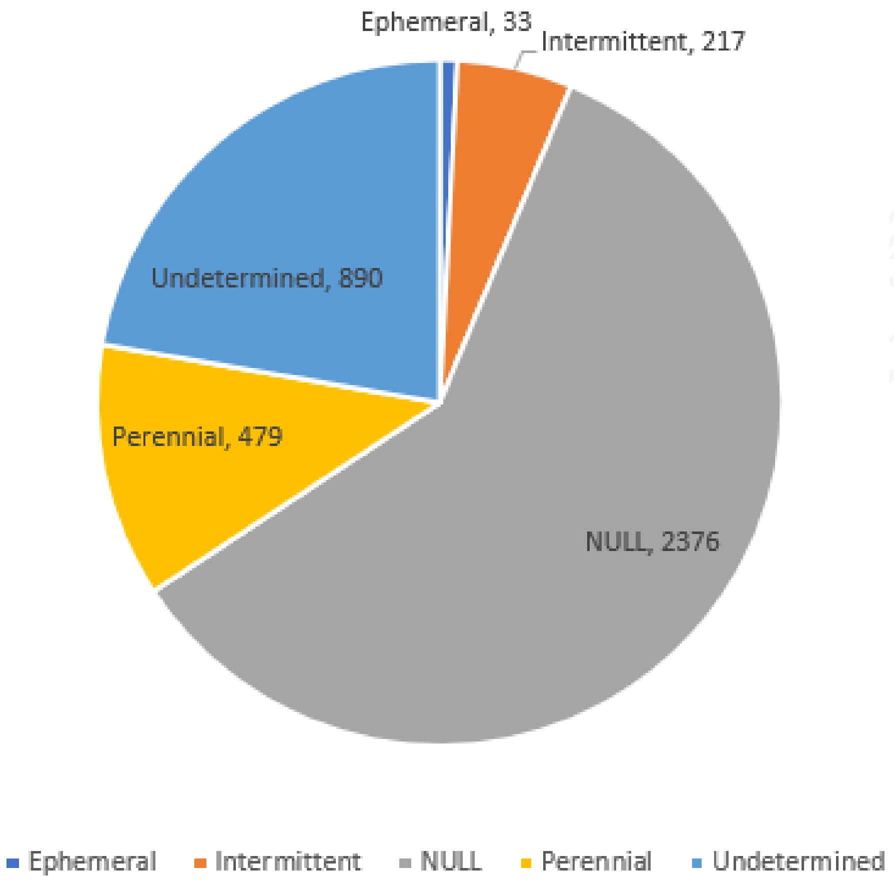
Assigning flow regimes

- Standards and Assessment Unit – Environmental Scientist
- My role: Classify the flow regimes of Arizona's water bodies
- Knowing a flow regime is essential to making water quality standards and regulatory decisions
- However, 82% of streams have no assigned flow regime



Assigning flow regimes

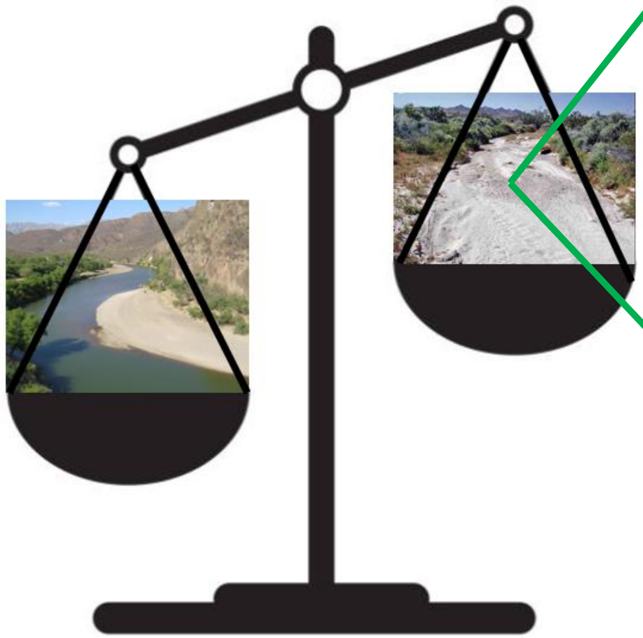
How can we use R to assign flow regimes?

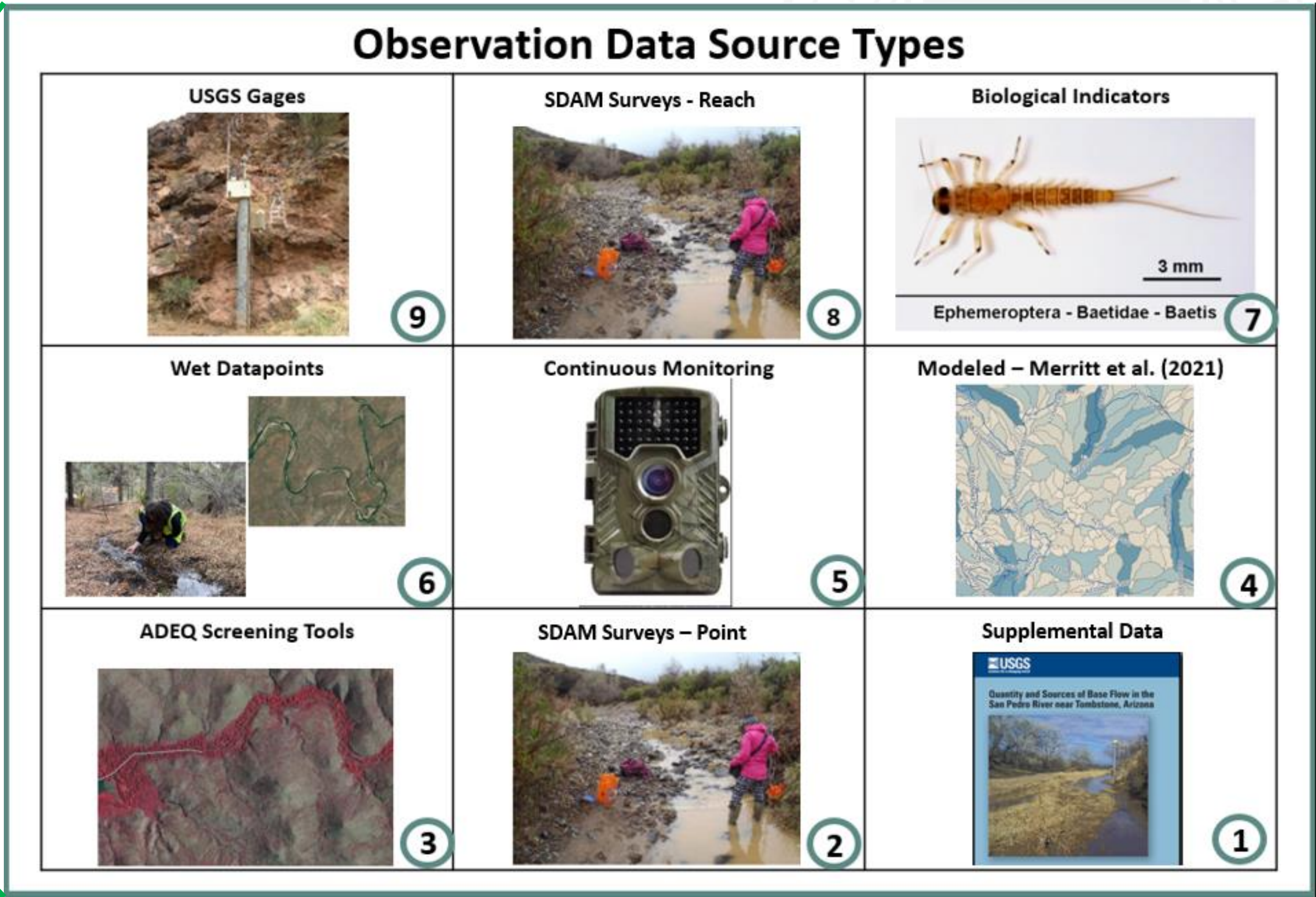


Assigning flow regimes

ADEQ
Arizona Department
of Environmental Quality

Weight of
Evidence Approach

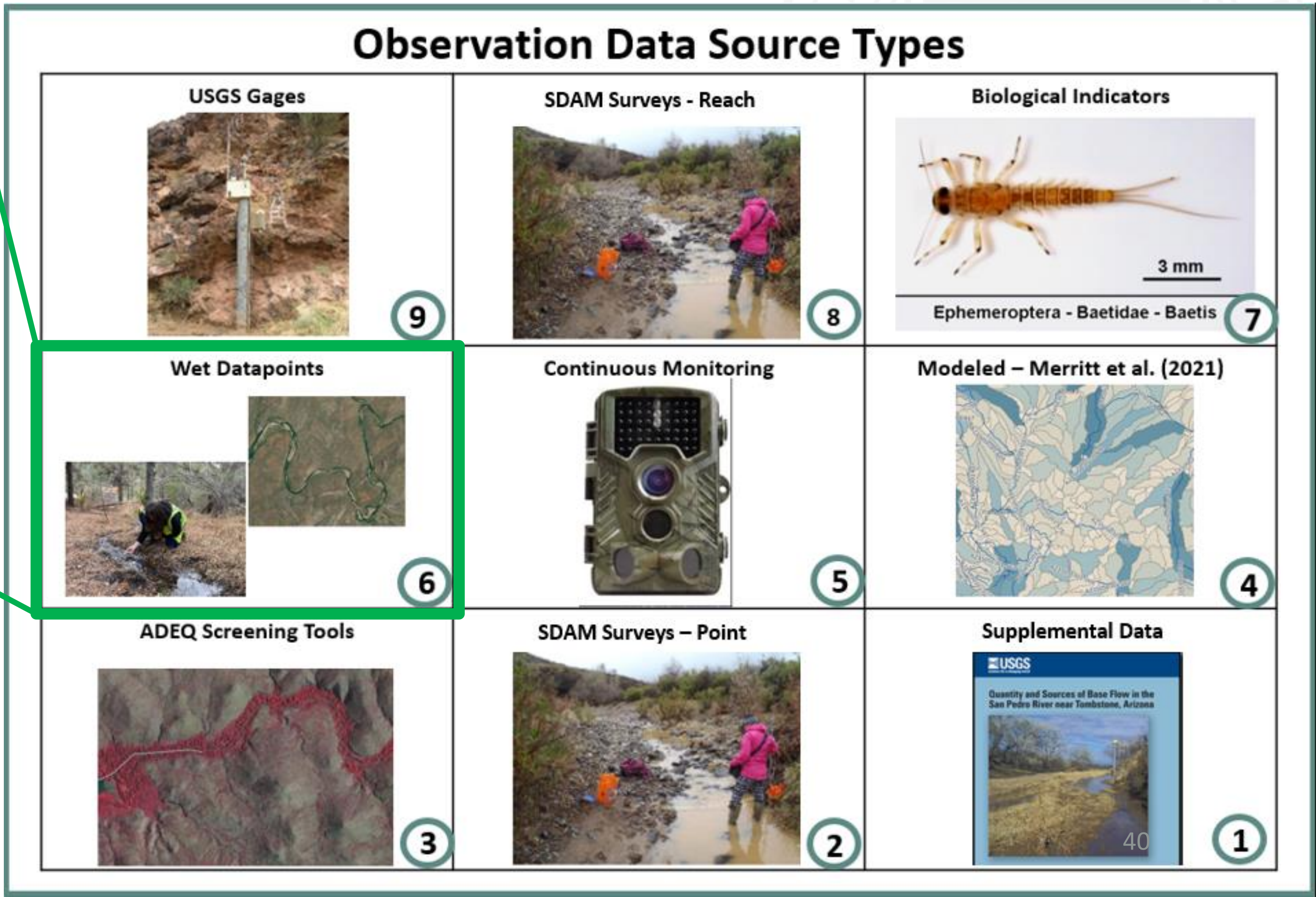




Assigning flow regimes

What does this have to do with R?

Wet Data Points – Level 6
Has THOUSANDS of records!



What are Wet Data Points?

Individual observations of flow on a single day

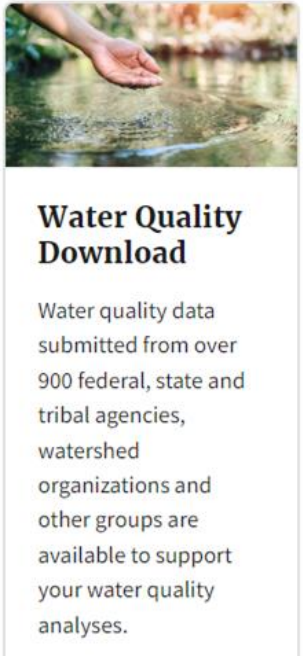
Wet Data Points

Ephemeral: “A surface water or portion of surface water that flows or pools **only in direct response to precipitation.”**

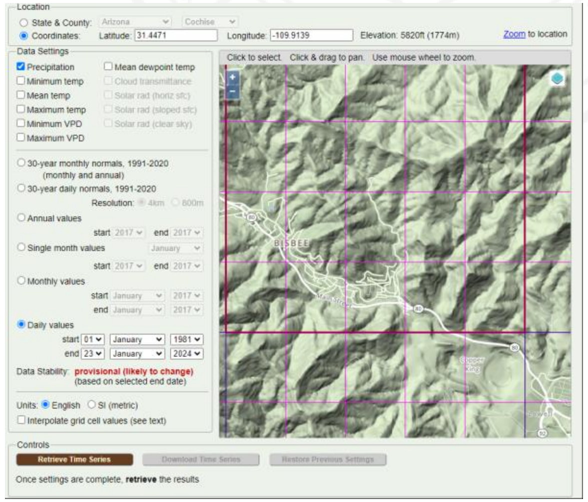
Flow observations must be paired with precipitation data to make a determination

0.4 inches of rain within a 48-hour period to be the precipitation magnitude that results in streamflow

Use wet data points that are NOT stormflow events to summarize flow over seasons and years: “At Least Intermittent”



+



coordinates from mid-reach

Assigning flow regimes

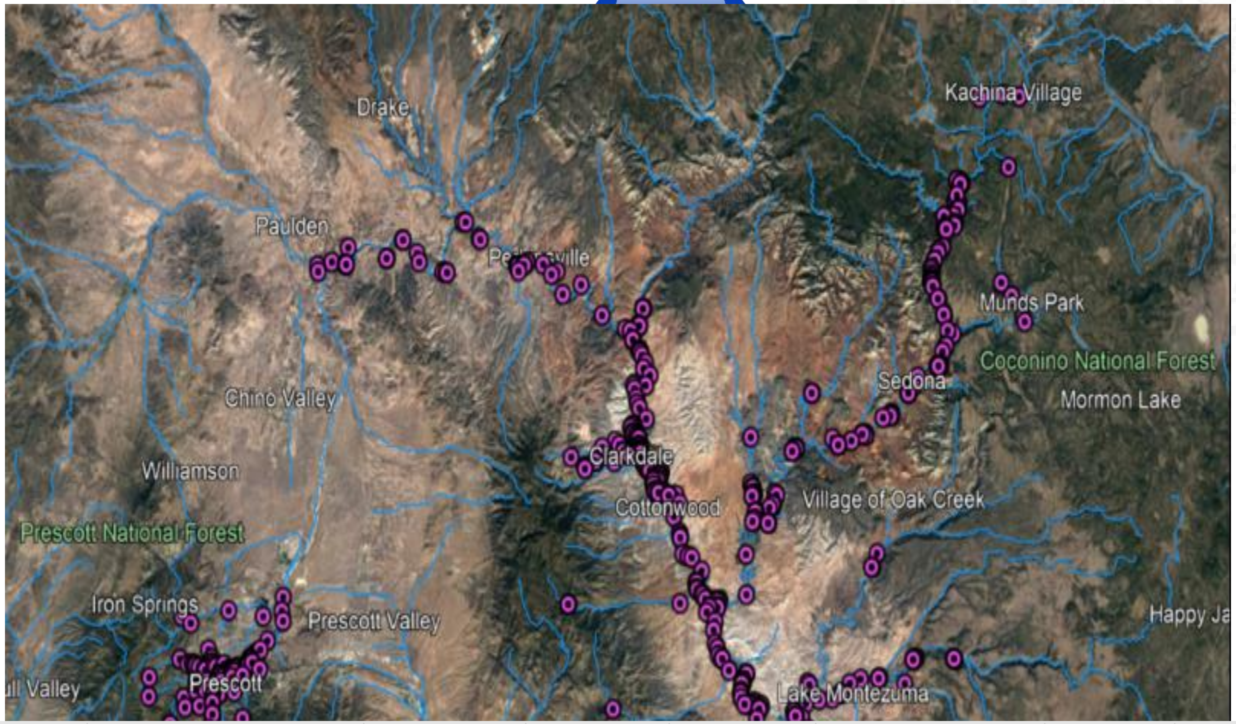
ADEQ Arizona Department of Environmental Quality

Wet Data Points R Process

1) Download + Clean WQX data

United States Environmental Protection Agency

2) Pull in precipitation data
3) Combine (loop)
4) Summarize
5) Assign a flow regime



	A	B	C	D	E	F	G	H
1	OrganizationIdentifier	OrganizationFormalName	ActivityIdentifier	ActivityTypeCode	ActivityMediaName	ActivityMediaSubdivisionName	ActivityStartDate	ActivityStartTime/Time
2	USGS-AZ	USGS Arizona Water Science Center	nwisaz.01.98101697	Sample-Routine	Water	Surface Water	6/4/1981	9:47:00
3	USGS-AZ	USGS Arizona Water Science Center	nwisaz.01.98101697	Sample-Routine	Water	Surface Water	6/4/1981	9:47:00
4	USGS-AZ	USGS Arizona Water Science Center	nwisaz.01.98101697	Sample-Routine	Water	Surface Water	6/4/1981	9:47:00
5	USGS-AZ	USGS Arizona Water Science Center	nwisaz.01.98101697	Sample-Routine	Water	Surface Water	6/4/1981	9:47:00
6	USGS-AZ	USGS Arizona Water Science Center	nwisaz.01.98100930	Sample-Routine	Water	Surface Water	7/30/1981	1:00:00

Assigning flow regimes

ADEQ Arizona Department of Environmental Quality

1) "Clean" WQX data

```
## Call Library
library(dplyr)
# Setup some Waterbody specific objects
wbid = "15070102-033A"
## Set the waterbody name
waterbody.name = "LY
# Read in the Master
all.az.site.data = r                                         station.csv",header=TRUE)

# Read in the HUC-le
wqp = read.csv(file                                          hem.csv"),header=TRUE)
## Prepare wqp datas
### reduce # of fiel
### reformat the dat
wqp = wqp %>%
  select(Organizatio                                         dentifier) %>%
  mutate(ActivitySta
 # create a characte                                          keep
 monitoring.types.ke                                         nnial", "Channelized Stream"
                                                             hemeral")

 ## Filter Monitorin
 wqp.joined = wqp.jo
   filter(Monitoring
 wqp.reduc = wqp.joi
   distinct(WBID, Wa                                         ityStartDate, season)

 # Set up the Season
 wqp.reduc = wqp.red
  mutate(quarter = q
  mutate(season = case_when(quarter == 1 ~ "winter",
                            quarter == 2 ~ "spring",
                            quarter == 3 ~ "summer",
                            quarter == 4 ~ "fall")) %>%
```

Only uses 1 package!

Set up some waterbody specific information…

Read in the WQX data…

Select() and mutate() remove all unwanted columns and changes the date format…

Filter() removes all types of locations we are not interested in

Distinct() gives us all unique sample dates per site
& mutate creates a "season" field

Assigning flow regimes

ADEQ
Arizona Department of Environmental Quality

2) Pull in precipitation data

WBID	Waterbody_Name	type.wet.points	sample_number	ActivityStartDate	season	Precipitation
15070103-239	French Gulch	AZDEQ_SW	AZDEQ_SW-859	10/11/2001	Summer	
15070103-239	French Gulch	AZDEQ_SW	AZDEQ_SW-859	1/29/2001	Winter	
15070103-239	French Gulch	AZDEQ_SW	AZDEQ_SW-859	2/23/2001	Winter	
15070103-239	French Gulch	AZDEQ_SW	AZDEQ_SW-859	4/24/2001	Spring	
15070103-239	French Gulch	AZDEQ_SW	AZDEQ_SW-859	3/29/2001	Winter	

PRISM CLIMATE GROUP

	A	B
1	Date	precipitation (inches)
2	10/7/2001	0.11
3	10/8/2001	0.03
4	10/9/2001	0.05
5	10/10/2001	0
6	10/11/2001	0
7	10/12/2001	0
8	10/13/2001	0

We don't just need 1 date, we need the antecedent 48-hours of precipitation…

How do we do this?

R!

Assigning flow regimes

ADEQ Arizona Department of Environmental Quality

EPA United States Environmental Protection Agency

3) Combine (using a loop!)

ActivityStartDate	season	Precipitation
10/11/2001	Summer	0.05
1/29/2001	Winter	0.00
2/23/2001	Winter	1.54
4/24/2001	Spring	0.00
3/29/2001	Winter	0.15

```
### Loop ###

for(i in 1:nrow(wdp

# define row in pr    ata wh    date is
date_1 = which(pris    te ==     s[i])   # pull the    st    ery date

# define the 3 date    pull    ip from (goes bac    days    original date)
dates_48 = c(date_1    te_1 -    date_1 - 2)

# Grab the precip c    bas    the rows of the    tes w    st defined
precip_tot = c(prism    t (i    s)`[dates_48])

# sum the 3 precip v    f    e 3 dates
precip_48_total = sum    ci    t)

# take the sum we just    ted and add it to th    precip c    n in the wdp dataset
wdp$precip[i] = precip_48_total
}
```

PRISM CLIMATE GROUP

	A	B
1	Date	precipitation (inches)
2	10/7/2001	0.11
3	10/8/2001	0.03
4	10/9/2001	0.05
5	10/10/2001	0
6	10/11/2001	0
7	10/12/2001	0
8	10/13/2001	0

= 0.05 inches = "precip_48_total"

Assigning flow regimes

Arizona Department of Environmental Quality

ActivityStartDate	season	Precipitation
10/11/2001	Summer	0.00
1/29/2001	Winter	0.00
3/20/2002	Winter	0.00
2/27/2003	Winter	0.00
5/15/2003	Spring	0.00
11/12/2003	Summer	0.00
12/26/2003	Fall	0.00
3/13/2004	Winter	0.00
4/3/2004	Spring	0.00
2/23/2004	Winter	0.00
5/5/2005	Spring	0.00
2/24/2005	Winter	0.00
2/23/2004	Winter	0.00

4) Summarize

Filter() retains only precip less than 0.4 inches

```
wet = wdp %>%
  filter(precip < 0.4)
```

Summarize the flow observations by season

```
obs_per_season = wet %>%
  group_by(WBID) %>%
  count(season) %>%
```

Count() the number of seasons

num_obs_in_fall	num_obs_in_spring	num_obs_in_summer	num_obs_in_winter
1	3	2	7

Use “ifelse” statements to answer how often flow is observed

```
# Run Logic
wet.summary = wet.2 %>%
  mutate(ali.1yr.1season = ifelse(num_years_in_fall >= 1 | num_years_in_winter >= 1
                        | num_years_in_summer >= 1,  "YES", "NA")) %>%
```

years1.3season	years2.3season	years3.3season	years4.3season
YES	YES	YES	NO

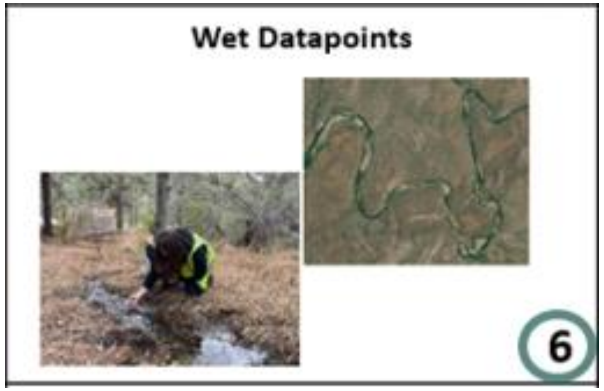
Assigning flow regimes

5) Assign Flow Regime!

Flow Regime Result: **At Least Intermittent**

Analysis: based on wet data point observations, there was at least 1 season with greater than 3 years of record after excluding all storm influenced flow events (precipitation threshold of 0.4inches of rain within 48-hours)

Recap:



= EPA + PRISM CLIMATE GROUP + R

Questions?

Jason Jones

Jones.Jason@azdeq.gov

ADEQ
Arizona Department of Environmental Quality

Clean Air, Safe Water,
Healthy Land for Everyone