PTAGIS Workshop: In-stream PIT-array Data Wrangling

# Full group presentation (40 minutes)

## Title:

PITcleanr: An R Package to Wrangle Messy PIT Tag Data for Analysis

## Abstract:

TBD

## Outline:

1. Background to In-stream Arrays
   1. History and Purpose
   2. Problems with PIT-array interrogation data
   3. How/what data is summarized, analyzed, etc.
2. Introduce PITcleanR
   1. How did it originate?
      1. To deal with the various issues with large, messy PIT-tag datasets,
      2. Initially developed and specialized for DABOM,
      3. But has evolved as a more general tool
      4. What can it do for you?
         1. Data queries
         2. Truncate and compress observations
         3. Order observations following a stream path or network
         4. Prep data for future analysis; CJS models, DABOM models, arrival/travel time
   2. What PITcleanR isn’t?
      1. Doesn’t answer questions.
      2. Doesn’t make bad data good.
      3. It is not the end.
3. Review Package
   1. Github Website (<https://github.com/KevinSee/PITcleanr>)
   2. Package Website (<https://kevinsee.github.io/PITcleanr/>)
   3. Vignettes (<https://kevinsee.github.io/PITcleanr/articles/>)
   4. Tutorials, Examples, Applications (maybe we develop more vignettes before workshop?)
4. Summary
   1. Our hope is to continue to develop PITcleanr to make it useful for a broader audience (w/ funding)
   2. Resources
      1. How to contact developers???
      2. How to submit bugs/issues???
   3. Present Afternoon Worksop. What will we do?

* Outline
  + Benefits / challenges of PIT tag data
  + PITcleanr
    - Purpose
    - Functionality
    - Package resource
  + Summary
* PIT tag data
  + Good things, bad things
  + Example of how much data can be generated
* PITcleanr package
  + Overview
* Example dataset
  + Show data – complete tag history
  + Read in from PTAGIS
    - Mention can read in from other sources
  + Compress
    - Need configuration
  + Direction of movement
    - Parent-child table
      * What does that mean
      * How to build one
        + From scratch
        + Using PITcleanr

Sites\_sf

Query flowlines

Build Parent-Child

* + Filter detections
    - By hand
      * Use min / max dates?
    - filterDetections() for one-way movements
  + Construct capture histories
* Other functionality
  + QA/QC for orphan or disowned tags
  + Estimate efficiency
  + Query PTAGIS for:
    - Interrogation / MRR site metadata
    - MRR data file
  + Other queries being worked on (require API key from PTAGIS):
    - PIT tag metadata
    - Test tag detections at single site for a calendar year
    - Complete tag history for one tag (don’t use for lots of tags)
* Package resources
  + Package website
  + Various vignettes
* Summary

# Break-out Workgroups (2 hours)

## Requirements:

Laptop computer with R (>version 4.3.0), tidyverse, and PITcleanR (>version 2.0.1) loaded.

install.packages(“devtools”)

remotes::install\_github(“KevinSee/PITcleanr”, build\_vignettes = TRUE)

browseVignettes(package = “PITcleanr”)

Own dataset?

## Pre-Workshop (15-30 minutes before start time):

Assist attendees with downloading R, tidyverse and PITcleanR

## Outline:

1. Introductions and housekeeping
2. Navigating the website
3. Downloading from GitHub
4. Getting package help and vignettes.
5. Walk through vignettes:
   1. Queries?
   2. Building node network
      1. Arrival and travel times
      2. Plotting nodes and map?
   3. Basic figures and tables
   4. Capture-history matrices
      1. Detection efficiencies
      2. CJS
6. Examples or Applications?
7. Work on your own (maybe 30-45 mins?)
8. Wrap-up
   1. Next steps
   2. Contributing ideas