**Workshop Agenda**

**Morning Session (Presentations)**

**8:00 AM - 9:25 AM Pt I**

* **8:00 AM – 8:15 AM:** Workshop Introduction (Yoko)
* **8:15 AM – 8:55 AM:** Breaking Data Silos in the Water Industry: Centralized Data Storage, Accessibility, and Cybersecurity (Javad)
* **8:55 AM - 9:25 AM:** Ahead of the Curve: How to use Machine Learning to Predict Chemical Contaminant Removal in Potable Reuse Systems (Josh)

**9:25 AM - 9:40 AM: Break**

**9:40 AM - 12:00 PM Pt II**

* **9:40 AM - 10:05 AM:** Introduction to Data Analysis: A Systematic Approach to Water Industry Analytics (Farshid) (25 minutes)
* **10:05 AM - 10:30 AM:** Case Study (Evan) (25 minutes)
* **10:30 AM - 11:00 AM:** Lessons Learned from Applying Machine Learning Techniques for Research in a Utility (Andrew)
* **11:00 AM - 11:30 AM:** Leveraging AI for soft sensing water quality when there are sensor faults (Jeff)

**11:30 AM - 1:00 PM: Lunch Break**

**Afternoon Session (Hands-On Activities)**

*Moderated by Yoko, facilitated by Sierra, Javad, Josh, and Farshid*

**1:00 PM - 5:00 PM**

R/Python code examples from real-world case studies will be shared through hands-on programming exercises. Participants will run complete scripts via Google Colab with facilitator support. **Important:** Participants need a Gmail account for Google Colab access - please communicate this in advance. Bring your own laptop for the best learning experience. Use sticky notes on laptops to request help from facilitators moving around tables.

**Format:** One presenter per example with multiple facilitators answering questions. Participants will sit in groups by default. One-page front/back handouts with best practices for each example. Complete scripts provided to all participants.

* **1:00 PM - 1:25 PM:** Intro to Programming for Water Professionals (Sierra)   
  Topics to cover: how to use LLM to help with code (good starting point but need to QC), data security of LLMs, debugging, IDEs and languages, mention what Github is/does, resources like stackoverflow, libraries/packages, classes/types, focus on tabular data – dataframes, lists, numeric, string, etc., what is a function/arguments

**Example 1: Water Quality Data with Spatial Elements (Using Tacoma, WA local data)**

*Facilitated by Jeff Sparks*

* **1:25 PM - 3:05 PM:**
  + Reading spatial climate data
  + Spatial data wrangling techniques
  + Spatial visualization methods
  + Climate data analysis approaches
  + Best practices discussion and Q&A
* **3:05 PM - 3:15 PM:** Break

**Example 2: Time Series Data Analysis—ML** **Prediction for Ammonia Sensor**

*Facilitated by Evan Paul*

* **3:15 PM - 5:00 PM:**
  + How to read data
  + Rapid data visualization (multiple plot types)
  + Analysis: correlation of time series, stationarity testing
  + Best practices discussion and Q&A