

Basic outline to guide discussion and work on automating code for Washtenaw County Health Department

1. Data Source Documentation

- a. Review documentation put together by WCHD
- b. **Discussion #1:** Are there opportunities to reduce human effort in data exchange? (For example: Direct Database connection, SFTP or other shared folder with partners?)
- c. **Discussion #2:** What are some checks to implement inside the data?
 - i. General: Naming convention checks, dates not in the future, etc.
 - ii. Specific: Look out for specific values? Values inside/outside expected ranges?
 - iii. What is an “auto-stop”? What is a “warning” we need to know about but information can go through?
 - iv. What are the best ways to notify relevant individuals about successful/unsuccessful runs?
 1. Email with attachment?
 2. File generation, someone checks?
 3. Teams/Slack notifications?

2. Code - Re-Writing for Automation

- a. **Discussion #1:** Think about the desired end product. (regardless if you'd like to pursue any of the “More Advanced” options, the “Basic” ones need to have been addressed- you'll still need the code)
 - i. Basic
 1. .PNG image files? What size should output charts be, and where do we want those files output to? Do we want to record each version, or overwrite previous versions?
 2. Data files for other use? Where do we want those files output to? .csv files or excel files? Do we want to record each version, or overwrite previous versions?
 - ii. More Advanced (skipping the details on these options for now - that can be a later “workshop”)
 1. Fill in powerpoint slides
 2. Rmarkdown reports
 3. Rshiny Dashboard
 - a. Frame the question(s)
- b. Inside the code
 - i. Implementation of data checks
 1. What are some ways to know if something has gone wrong? Can we articulate those in such a way to test for them?
 - ii. Data alerts, warnings, or other notes to mark progress through code
 - iii. Path names for automation
 - iv. Functions for repeated tasks & single points of edits

1. Store these in such a way that they can be used for other projects
- c. Documenting the code
 - i. Comments/etc. inside the code
 - ii. Flow charts/other write-ups
- d. Testing the code
 - i. Repeat old projects
 - ii. Run new projects side-by-side prior to full implementation

3. Automation

- a. Windows vs. Mac
- b. **Discussion #1:** Where does the automation live? (On someone's work machine, on a designated "job" machine?)
- c. **Discussion #2:** How often does the job need to run?

4. System Management

- a. **Discussion #1:** Who is responsible for checking for errors? Who is responsible for fixing errors?
- b. **Discussion #2:** How will you monitor for changing data structures?
- c. **Discussion #3:** How will you monitor changes in the code/notify stakeholders about changes in any calculations?