SET Data Sanity Checks

2019-01-29

Read in processed data (all SETs in same CSV file) and perform some basic checks of the parameters.

library(knitr)  
library(readr)  
library(readxl)  
library(janitor)  
library(dplyr)  
library(tidyr)  
library(ggplot2)  
library(here)  
# library(flextable)  
source(here::here('R\_scripts', '000\_functions.R'))

Read in data. Can see how it’s formatted in the following output.

## Parsed with column specification:  
## cols(  
## reserve = col\_character(),  
## set\_id = col\_character(),  
## arm\_position = col\_character(),  
## date = col\_date(format = ""),  
## pin\_number = col\_character(),  
## flag = col\_double(),  
## pin\_height\_mm = col\_double()  
## )

# Introduce some wonkiness for testing purposes

When QC flag is [one of a subset of numbers; i’ll pick 2], make pin\_height NA to exclude that value from analysis. Generate another tabular summary.

**NOTE** At this point, I’ve inserted a few wonky values. A ‘2’ flag in the 2nd-4th rows of the data frame, and a negative pin reading as the first pin\_height. This is to test the script’s capabilities to detect these.

play with flextable. Just kidding; this won’t work without pandoc 2.0 or higher, and that’s not yet bundled with RStudio. Keeping the code here because I think that will be supported soon, and then we can auto-highlight some of the output.

Print out some information about each SET.

### CLMAJ-1

Summary of SET visits

|  |  |  |
| --- | --- | --- |
| times\_read | first\_reading | last\_reading |
| 19 | 2012-02-29 | 2016-11-21 |

Summary of reading dates that do not meet QC criteria

|  |  |  |  |
| --- | --- | --- | --- |
| date | pre\_n | post\_n | removed |
| 2012-02-29 | 36 | 33 | 3 |
| 2016-08-22 | 35 | 35 | 0 |

### CLMAJ-2

Summary of SET visits

|  |  |  |
| --- | --- | --- |
| times\_read | first\_reading | last\_reading |
| 19 | 2012-02-29 | 2016-11-21 |

[1] “All reading dates meet QC criteria (36 pins; no readings removed by QC codes)”

### CLMAJ-3

Summary of SET visits

|  |  |  |
| --- | --- | --- |
| times\_read | first\_reading | last\_reading |
| 19 | 2012-02-29 | 2016-11-21 |

[1] “All reading dates meet QC criteria (36 pins; no readings removed by QC codes)”

### JURO\_High-1

Summary of SET visits

|  |  |  |
| --- | --- | --- |
| times\_read | first\_reading | last\_reading |
| 19 | 2012-02-28 | 2016-11-22 |

[1] “All reading dates meet QC criteria (36 pins; no readings removed by QC codes)”

### JURO\_High-2

Summary of SET visits

|  |  |  |
| --- | --- | --- |
| times\_read | first\_reading | last\_reading |
| 19 | 2012-02-28 | 2016-11-22 |

Summary of reading dates that do not meet QC criteria

|  |  |  |  |
| --- | --- | --- | --- |
| date | pre\_n | post\_n | removed |
| 2013-07-01 | 35 | 35 | 0 |

### JURO\_High-3

Summary of SET visits

|  |  |  |
| --- | --- | --- |
| times\_read | first\_reading | last\_reading |
| 19 | 2012-02-28 | 2016-11-22 |

Summary of reading dates that do not meet QC criteria

|  |  |  |  |
| --- | --- | --- | --- |
| date | pre\_n | post\_n | removed |
| 2016-02-10 | 35 | 35 | 0 |

### JURO\_Low-1

Summary of SET visits

|  |  |  |
| --- | --- | --- |
| times\_read | first\_reading | last\_reading |
| 19 | 2012-03-02 | 2016-11-23 |

Summary of reading dates that do not meet QC criteria

|  |  |  |  |
| --- | --- | --- | --- |
| date | pre\_n | post\_n | removed |
| 2012-06-15 | 35 | 35 | 0 |
| 2014-05-13 | 35 | 35 | 0 |
| 2016-08-23 | 35 | 35 | 0 |

### JURO\_Low-2

Summary of SET visits

|  |  |  |
| --- | --- | --- |
| times\_read | first\_reading | last\_reading |
| 19 | 2012-03-02 | 2016-11-23 |

Summary of reading dates that do not meet QC criteria

|  |  |  |  |
| --- | --- | --- | --- |
| date | pre\_n | post\_n | removed |
| 2016-02-17 | 35 | 35 | 0 |

### JURO\_Low-3

Summary of SET visits

|  |  |  |
| --- | --- | --- |
| times\_read | first\_reading | last\_reading |
| 19 | 2012-03-02 | 2016-11-23 |

Summary of reading dates that do not meet QC criteria

|  |  |  |  |
| --- | --- | --- | --- |
| date | pre\_n | post\_n | removed |
| 2012-03-02 | 35 | 35 | 0 |
| 2012-09-14 | 35 | 35 | 0 |

### JURO\_Mid-1

Summary of SET visits

|  |  |  |
| --- | --- | --- |
| times\_read | first\_reading | last\_reading |
| 19 | 2012-02-28 | 2016-11-22 |

[1] “All reading dates meet QC criteria (36 pins; no readings removed by QC codes)”

### JURO\_Mid-2

Summary of SET visits

|  |  |  |
| --- | --- | --- |
| times\_read | first\_reading | last\_reading |
| 19 | 2012-02-28 | 2016-11-22 |

[1] “All reading dates meet QC criteria (36 pins; no readings removed by QC codes)”

### JURO\_Mid-3

Summary of SET visits

|  |  |  |
| --- | --- | --- |
| times\_read | first\_reading | last\_reading |
| 19 | 2012-02-28 | 2016-11-22 |

[1] “All reading dates meet QC criteria (36 pins; no readings removed by QC codes)”

### SPALT-1

Summary of SET visits

|  |  |  |
| --- | --- | --- |
| times\_read | first\_reading | last\_reading |
| 19 | 2012-03-02 | 2016-11-23 |

Summary of reading dates that do not meet QC criteria

|  |  |  |  |
| --- | --- | --- | --- |
| date | pre\_n | post\_n | removed |
| 2015-09-09 | 35 | 35 | 0 |
| 2016-02-17 | 35 | 35 | 0 |
| 2016-05-18 | 35 | 35 | 0 |

### SPALT-2

Summary of SET visits

|  |  |  |
| --- | --- | --- |
| times\_read | first\_reading | last\_reading |
| 19 | 2012-03-02 | 2016-11-23 |

Summary of reading dates that do not meet QC criteria

|  |  |  |  |
| --- | --- | --- | --- |
| date | pre\_n | post\_n | removed |
| 2012-11-29 | 35 | 35 | 0 |
| 2014-05-13 | 35 | 35 | 0 |
| 2014-08-29 | 35 | 35 | 0 |
| 2015-02-11 | 35 | 35 | 0 |
| 2015-09-09 | 35 | 35 | 0 |
| 2015-12-01 | 33 | 33 | 0 |
| 2016-02-17 | 35 | 35 | 0 |
| 2016-05-18 | 35 | 35 | 0 |

### SPALT-3

Summary of SET visits

|  |  |  |
| --- | --- | --- |
| times\_read | first\_reading | last\_reading |
| 19 | 2012-03-02 | 2016-11-23 |

[1] “All reading dates meet QC criteria (36 pins; no readings removed by QC codes)”

# Tests to perform

Find any counts that are not 36  
Especially find any that are > 36  
Find differences between pre- and post- QC point removal

# Tests and Checks

## Negative Pin Heights

## [1] "all pin heights are >0"