1/31/2022

Keith Dorais

**Fka SV Test Script Notes**

1. **AOI Cleanup**

* Filter out all invalid data and replace with -1 if not already done so
* Use OR statement to get a series of all valid AOI data points in Boolean
* Replace all FALSE points with -1 since these are invalid

1. **Validity Check**

* Create an AND condition to find all valid data points
* Replace all FALSE points with AOI value of -1
  + Add Boolean value to end of data frame

1. **Durations**

* Find points where each condition is true
* Sum duration at each point and save to a variable
* Add durations for each condition into a new DataFrame

1. **Data Quality Check**

* Compare duration with difference in timestamp from current point and next point
  + (Duration == Timestamp – Timestamp-1)
* Create boolean series with False for matching data (no difference) and True for wrong data (duration difference)
* Use loc to mark invalid data

1. **Reduction – NOT FINISHED**

* Check if Data-Validity == False
* Merge consecutive False data points
  + Use shift to compare index # to check if points are consecutive
  + This part is what has me stuck, I can’t figure out how to merge the consecutive cells
  + Drop\_duplicates doesn’t work, all duplicate cells are lost
  + Not sure how to proceed
* Set duration according to new Timestamp (Same as last step)

1. **Recalculated Durations**

* Run same code as in step 3 with updated duration data

1. **Validity**

* Use normalized value\_counts function to find % of Valid & Invalid data points
* Find total duration of Valid data points
* Divide sum of Valid data duration by total duration
* Add Valid data points % and Duration % to a new DataFrame and export for review

1. **Export**

* Export final modified data log for reference