# DART multi-day experiment analysis workflow

This document guides users through basic analysis of multi-day 2-photon imaging experiments as used for DART projects. The analysis pipeline assumes that two days (or two sessions on the same day) of data from the same imaging field of view have been collected. One will be used as the reference, and the other dataset will be matched to this reference.

This sequence of MATLAB scripts will first segment and extract timecourses from each day’s data, then the data from cell sidentifiable across days will be matched. Key metrics will be extracted from cells that were matched and were active on at least one day. These metrics include the trial-averaged timecourse and mean response at each cell’s preferred direction, the normalized difference of this trial-averaged response, and noise correlations between cells and the population. Additionally, response data will be separated by locomotion and pupil dilation as measures of behavioral state.

The analysis pipeline uses three configuration files that work together:

1. An Instructions file specifies which session to analyze, points to the appropriate dataset file and defines which day (DART or control) was used as reference for cell matching. Most of the time this instruction file should be the only one you need to edit.
2. A datasheet file that contains metadata for all experimental sessions in an expt array. Each array element stores one session's parameters (mouse ID, date, imaging settings, drug condition, etc.). Importantly, the multiday\_matchdays field links paired sessions (e.g., control/DART).
3. An RC that defines directory paths based on hostname and username. This returns a struct rc with fields like rc.analysis and rc.data. This approach Centralizes path management across different computers/users.

All the scripts listed below are in “ImagingCode-Glifeld-Hull/celine/DART pipeline”. A number of functions are called that are in “ImagingCode-Glifeld-Hull/celine”.

## Steps and considerations

**Before beginning analysis:**

* Add your paths to behavConstsDART RC file.
* Create a datasheet using step0\_exampleDataSheet.m.
* Make an instruction file based on instr\_example.m

**Analysis pipeline**

1. **Segment each day’s data using step1\_DART\_segmenting.m**
   1. Enter the datasheet and experiment ID and then run this script all at once (rather than in chunks).
   2. Consider modifying what images are used for segmentation and how many times each image appears to optimize segmentation for different fluorophores or expression patterns.
   3. Key output: registration shifts, masks and timecourses for each individual day.
2. **Match the two days using step2\_DART\_matching.m**
   1. If only a green (GCaMP) channel was collected, use step2\_DART\_matching\_greenOnly.m
   2. Enter the datasheet and experiment ID and then run this script all at once (rather than in chunks).
   3. Consider which day is best to use as the reference vs. the match day for a given project.
   4. Key outputs: alignment shifts to match two FOVs, timecourses and input structure as cell arrays where each day is a cell (for all cells, and subset to matched cells), indices for which cells were red.
3. **Get timeseries of pupil size using step3\_DART\_eyeTracker.m**
   1. Enter the datasheet and experiment ID and then run this script all at once (rather than in chunks).
   2. The pupil radian range can be modified if pupil detection fails.
   3. Key output: timecourses of pupil diameter in radians for each day
4. **Extract key metrics using step4\_DART\_dataExtraction.m**
   1. Run this script all at once (rather than in chunks); it will ask for the datasheet and experiment ID through the command line.
   2. Consider updating if other outputs are desired.
   3. Current version focuses on cells that are active on at least one day. These are referred to as “keep” cells.
   4. Key outputs: trial-averaged timecourses and mean responses for keep cells at their preferred direction and separated by locomotion and pupil size, noise correlations, normalized difference. Outputs are cell arrays where each day is a cell.
   5. Outputs with “tc” in the name are timecourses; output with “response” in the name are averaged over the response window.
5. **Visualize results using step5\_DART\_multiDayVisualizations.m**
   1. Enter the datasheet and experiment ID then run this script one chunk at a time, so that you can see the output as you go. You may not need want all chunks, depending on what is relevant to your experiment.
   2. This script has basic visualizations, but few statistical comparisons. As such it is suitable for preliminary exploration of data to guide subsequent statistical testing.
   3. Key outputs: plots of grand average timecourses for HTP+/- (that is, red and green) neurons on the control and DART days at each contrast and size for various behavioral states, contrast response and size response function plots for HTP+/- neurons on the control and DART days, plots of fraction HTP+ cells suppressed and facilitated at each contrast and size, timecourses split by noise correlation.