cell specimen table
table/pandas dataframe

Table of cell information where non-cell segmented rois have been filtered out (only contains roi_valid = True entries)

column name	data type	description
cell_specimen_id [index]	int 64	unified id of segmented cell across experiments (assigned after cell matching)
cell_roi_id	int	experiment specific id of segmented roi (assigned before cell matching)
height	int 64	height of ROI in pixels
mask_image_plane	int 64	which image plane an ROI resides on in data processing file
max_correction_down	float 64	max motion correction in down direction in pixels
max_correction_left	float 64	max motion correction in left direction in pixels
max_correction_right	float 64	max motion correction in right direction in pixels
max_correction_up	float 64	max motion correction in up direction in pixels
roi_mask	array of bool	an image array that displays the location of the roi mask in the field of view
valid_roi	bool	indicates if cell classification found the ROI to be a cell or not (True = cell, False = not cell).
width	int 64	width of ROI in pixels
x	float 64	x position of ROI in image plane in pixels (top left corner)
y	float 64	y position of ROI in image plane in pixels (top left corner)

dff traces
table/pandas dataframe

Baseline normalized fluorescence traces for all valid cells in an ophys experiment, organized by cell_specimen_id. For more information on dF/F calculation see technical whitepaper dF/F calculation section. Sampling rate can be found in metadata 'ophys_frame_rate'.

column name	data type	description
cell_specimen_id [index]	int 64	unified id of segmented cell across experiments (assigned after cell matching)
cell_roi_id	int 64	experiment specific id of segmented roi (assigned before cell matching)
dff	array of float 64	fractional fluorescence values relative to baseline (baseline is rolling 10 min median value)

events
table/pandas dataframe

A dataframe containing spiking events in traces derived from the two photon movies, organized by cell specimen id. For more information on event detection processing please see the event detection portion of the white paper.

column name	data type	description
cell_specimen_id	int 64	[index] unified id of segmented cell across experiments (assigned after cell matching)
cell_roi_id	int 64	experiment specific id of segmented roi (assigned before cell matching)
events	array of float	event trace where events correspond to the rise time of a calcium transient in the dF/F trace, with a magnitude roughly proportional the magnitude of the increase in dF/F
filtered_events	array of float	events array with a 1d causal half-gaussian filter used for smoothing. Uses a halfnorm distribution as weights to the filter
lambda	float 64	regularization value selected to make the minimum event size be close to N * noise_std
noise_std	float 64	estimated noise standard deviation for the events trace

ophys timestamps
numpy array

Timestamps associated with 2-photon frames captured by the microscope. Sampled at different rates depending upon the microscope. Sampling rate can be found in metadata 'ophys_frame_rate'.

max_projection
image

2D maximum intensity projection image of the microscope field of view.

average_projection
image

2D average intensity projection image of the microscope field of view.

segmenation_mask_image
image

Binary mask image showing all segmented ROIs in the imaging plane

motion correction
table/pandas dataframe

x and y offsets (in pixels) applied during motion correction

column name	data type	description
x	int	frame shift on x axis
y	int	frame shift on y axis

raw running speed
table/pandas dataframe

Unfiltered running speed and timestamps sampled at 60Hz.

column name	data type	description
speed	float 64	speed in cm/sec
timestamps	float 64	time in seconds

stimulus templates
table/pandas dataframe

A table containing the stimulus images for a session. Indices are image names, 'warped' and 'unwarped' columns contain image arrays.

column name	data type	description
image_name [index]	string	name of stimulus presented
unwarped	array of int	image array of unwarped stimulus image
warped	array of int	image array of warped stimulus image

stimulus presentations
table/pandas dataframe

Table whose rows are stimulus presentations and whose columns are presentation characteristics. Table contains every image presented , including those associated with trials that were aborted, as well as stimuli that were omitted.

column name	data type	description
stimulus_presentations_id	int 64	[index] identifier for each unique stimulus presentation
duration	float 64	duration of image flash in seconds (stop time - start time), NaN if omitted
end_frame	float 64	image presentation end frame
image_index	int 64	image index (0-7) for each image, corresponding to each image name in the session
image_name	string	name of stimulus presented; 'omitted' if no stimulus was presented
image_set	string	Filename of image set used
index	int 64	An index assigned to each stimulus presentation
omitted	bool	True if no image was shown for this stimulus presentation
start_frame	int 64	image presentation start frame; can be used to index into stimulus_timestamps
start_time	float 64	image presentation start time in seconds
stop_time	float 64	image presentation end time in seconds

licks
table/pandas dataframe

lick timestamps (in seconds) and frames, sampled at 60Hz.

column name	data type	description
timestamps	float 64	time of each lick in seconds
frame	int	Stimulus frame of lick

running speed
table/pandas dataframe

Running speed and timestamps sampled at 60hz.

column name	data type	description
speed	float 64	speed in cm/sec
timestamps	float 64	time in seconds

stimulus timestamps
numpy array

Timestamps associated with the stimulus presentations on the monitor. Stimuli are presented at 60Hz. An array of float 64.

rewards
table/pandas dataframe

A dataframe containing timestamps of delivered rewards. Timestamps are sampled at 60 Hz.

column name	data type	description
timestamps	float 64	time of each reward in seconds
volume	float 64	volume of water reward in ml.
autorewarded	bool	True if free reward was delivered for that trial. Occurs during the first 5 trials.

reward rate
numpy array

Reward rate is calculated of a 25 trial rolling window and provides a measure of the rewards earned per unit time (in units of rewards/min)

task parameters
dictionary

Parameters used to define the task runtime behavior

key	Value dtype	value
blank_duration_sec	list of float	duration in seconds of inter stimulus interval.
stimulus_duration_sec	float	duration in seconds of each stimulus image presentation
omitted_flash_fraction	float	probability that a stimulus image presentations is omitted. Change stimuli, and the stimulus immediately preceding the change, are never omitted.
response_window_sec	list of float	the range of the period following an image change, in seconds, where mouse response influences trial outcome. The first value represents response window start. The second value represents response window end. The values represent time before display lag is accounted for and applied.
reward_volume	float	volume of earned water reward in ml (0.007)
auto_reward_voume	float	volume of auto rewards in ml (0.005ml)
session_type	string	visual stimulus type displayed during behavior session
stimulus	string	stimulus type
stimulus_distribution	string	distribution for drawing change times. Either 'exponential' or 'geometric'.
task	string	type of visual stimulus task, typically 'change detection'
n_stimulus_frames	int	total number of visual stimulus frames presented during a behavior session

trials

table/pandas dataframe

trial and behavioral response data

column name	data type	description
trials_id [index]	int 64	Unique identifier for each trial (including go, catch, and aborted trials)
lick_times	int_64	Timestamps of licks that occurred during the trial
reward_time	NaN or float	Time the reward is delivered following a correct response, or on autorewarded trials.
reward_volume	float	volume of reward in ml. (0.005 for auto reward and 0.007 for earned reward)
hit	bool	[behavior response type] on a go trial, when mouse correctly licks within reward window
miss	bool	[behavior response type] on a go trial, mouse does not lick at all, or licks after reward window
correct_reject	bool	[behavior response type] on a catch trial, mouse does not lick at all or licks after reward window
false_alarm	bool	[behavior response type] on a catch trial, mouse incorrectly licks within reward window
aborted	bool	[behavior response type] True if mouse licks before the scheduled change time .
stimulus_change	bool	True if an image change occurs during the trial (if the trial was both a 'go' trial and not aborted)
go	bool	[trial type] True if there was a change in stimulus identity on this trial
catch	bool	[trial type] True if there was not a change in stimulus identity on this trial
auto_rewarded	bool	True if free reward was delivered for that trial. Occurs during the first 5 trials of a session.
start_time	float 64	start time of the trial in seconds. Change times are drawn at the start of a trial.
stop_time	float 64	end time of the trial in seconds. Trial ends ~3 seconds after change occurs.
trial_length	float 64	duration of trial, in seconds (stop time - start time)
response_time	float 64	time of first lick in trial in seconds, and NaN if trial aborted
change_frame	float 64	frame of stimulus change
change_time	float 64	time of stimulus change in seconds
response_latency	float 64	time of first lick relative to the change time; NaN if trial aborted. (response_time - change_time)
initial_image_name	string	name of stimulus presented at start of trial
change_image_name	string	name of stimulus that is changed to at the change time, on go trials

performance metrics

table/pandas dataframe

Table with session scale summary behavior response performance metrics

column name	data type	description
trial_count	int	The length of the trial dataframe (including all go, catch and aborted trials)
go_trial_count	int 64	number of go trial types
hit_trial_count	int 64	number of catch trial types
miss_trial_count	int 64	number of trials with a hit behavior response
false_alarm_trial_count	int 64	number of trials with a miss behavior response
correct_reject_trial_count	int 64	number of trials with a false alarm behavior response
auto_rewareded_trial_count	int 64	number of trials with a correct reject behavior response
rewarded_trial_count	int 64	number of trials where the mouse received an auto reward of water.
total_reward_count	int	number of trials where the mouse was eligible to receive a water reward (go trials) and did receive an earned water reward
total_reward_volume	float 64	volume of all water rewards received (earned and auto rewarded)
maximum_reward_rate	float 64	The peak of the rolling reward rate (rewards/minute)
engaged_trial_count	int 64	number trials where the mouse was engaged (reward rate>2 rewards/min)
mean_hit_rate	float 64	The mean of the rolling hit_rate
mean_hit_rate_uncorrected	float 64	The mean of the rolling hit_rate_raw
mean_hit_rate_engaged	float 64	The mean of the rolling hit_rate, excluding epochs when the rolling reward rate was below 2 rewards/min
mean_false_alarm_rate	float 64	The mean of the rolling false_alarm_rate, excluding epochs when the rolling reward rate was below 2 rewards/min
mean_false_alarm_rate_uncorrected	float 64	The mean of the rolling false_alarm_rate_raw
mean_false_alarm_rate_engaged	float 64	The mean of the rolling false_alarm_rate, excluding epochs when the rolling reward rate was below 2 rewards/min
mean_dprime	float 64	The mean of the rolling d_prime
mean_dprime_engaged	float 64	The mean of the rolling d_prime, excluding epochs when the rolling reward rate was below 2 rewards/min
max_dprime	float 64	The peak of the rolling d_prime
max_dprime_engaged	float 64	The peak of the rolling d_prime, excluding epochs when the rolling reward rate was below 2 rewards/min

metadata

dictionary

dataset specific metadata. keys with * are only found in behavior ophys experiment datasets and not behavior ophys session datasets

key	Value dtype	value
age_in_days	int	age of mouse in days
behavior_session_id	int	unique identifier for a behavior session
behavior_session_uuid	uuid	unique identifier for a behavior session
cre_line	string	cre driver line for a transgenic mouse
date_of_acquisition	date time object	date and time of experiment acquisition, yyyy-mm-dd hh:mm:ss.
driver_line	list of string	all driver lines for transgenic mouse
* emission_lambda	float	wavelength (in nm) the 2-photon microscope is tuned to based on mouse reporter line
equipment_name	string	identifier for equipment data was collected on (microscope or behavior training box)
* excitation_lambda	float 64	wavelength (in nm) of the 2-photon laser
* experiment_container_id	int	unique identifier for a container
* field_of_view_height	int	field of view height in pixels
*field_of_view_width	int	field of view width in pixels
full_genotype	string	full genotype of transgenic mouse
* imaging_depth	int	depth in microns from the cortical surface, where the data was collected
* imaging_plane_group	none or int	a grouping number assigned to grouped planes (experiments) for multiscope sessions
* imaging_plane_group_count	int	total number of imaging groups for multiscope session
indicator	string	transgenic mouse indicator
mouse_id	int	unique identifier for a mouse
* ophys_experiment_id	int	unique identifier for an ophys experiment
* ophys_frame_rate	float 64	2-photon imaging frame rate (in Hz)
* ophys_session_id	int	unique identifier for an ophys session
reporter_line	string	reporter line for transgenic mouse
session_type	string	visual stimulus type displayed during behavior session
stimulus_frame_rate	float	rate (in Hz) that the visual stimulus was displayed
* targeted_structure	string	cortical area targeted by the experiment field of view

eye tracking

table/pandas dataframe

Ellipse fit parameters for the eye, pupil and corneal reflection (cr). Fits are derived from tracking points from a DeepLabCut model applied to video (collected at 30hz) frames of a subject's right eye.

column name	data type	description
frame [index]	int 64	frame index of eye tracking video
timestamps	float 64	time in seconds
likely_blink	float 64	True if frame has outlier ellipse fits, which is often caused by blinking / squinting of the eye.
eye_area	float 64	area of eye ellipse in pixels squared, after setting likely blinks to NaN
eye_area_raw	float 64	pupil area with no outliers/likely blinks removed.
eye_center_x	float 64	center of eye ellipse on x axis in pixels
eye_center_y	float 64	center of eye ellipse on y axis in pixels
eye_height	float 64	eye height (minor axis of the eye ellipse) in pixels
eye_width	float 64	eye width (major axis of the eye ellipse) in pixels
eye_phi	float 64	eye rotation angle in radians
pupil_area	float 64	area of pupil ellipse in pixels squared, after setting likely blinks to NaN
pupil_area_raw	float 64	pupil area with no outliers/likely blinks removed.
pupil_center_x	float 64	center of pupil ellipse on x axis in pixels
pupil_center_y	float 64	center of pupil ellipse on y axis in pixels
pupil_height	float 64	pupil height (minor axis of the pupil ellipse) in pixels
pupil_width	float 64	pupil width (major axis of the pupil ellipse) in pixels
pupil_phi	float 64	pupil rotation angle in radians
cr_area	float 64	area of corneal reflection ellipse in pixels squared

Loading Cache Tables

```
from allensdk.brain_observatory.behavior.behavior_project_cache import VisualBehaviorOphysProjectCache
```

To load the Visual Behavior cache on AWS:
```
cache = VisualBehaviorOphysProjectCache.from_local_cache(cache_dir=data_storage_directory, use_static_cache=True)
Where data_storage_directory = "/data/visual-behavior-ophys_data"
```

To load the Visual Behavior cache on a local drive:
```
cache = VisualBehaviorOphysProjectCache.from_s3_cache(cache_dir=cache_dir)
Where cache_dir is a folder on your local drive
```

To load the cache tables:
```
ophys_experiment_table = cache.get_ophys_experiment_table()
ophys_session_table    = cache.get_ophys_session_table()
behavior_session_table = cache.get_behavior_session_table()
ophys_cell_table       = cache.get_ophys_cell_table()
```

Cache Summary Tables

table/pandas dataframe

The cache contains 4 tables which provide a summary level view of all the currently released datasets: ophys_experiment_table, ophys_session_table, behavior_session_table and, ophys_cells_table. These tables contain roughly the same information but are organized differently. Each table is indexed by its corresponding id type

column name	data type	description	ophys_experiment_table	ophys_session_table	behavior_session_table	ophys_cell_table
age_in_days	int	age of mouse in days	x	x	x	
behavior_session_id	int	unique identifier for a behavior session	x	x	x	
cell_roi_id	int	experiment specific id of segmented roi				x
cell_specimen_id	int 64	unified id of segmented cell matched across experiments				x
cre_line	string	cre driver line for a transgenic mouse	x	x	x	
date_of_acquisition	datetime	date and time of data acquisition, yyyy-mm-dd hh:mm:ss.	x	x	x	
driver_line	list	all driver lines for transgenic mouse	x	x	x	
equipment_name	string	identifier for equipment data was collected on	x	x	x	
experience_level	string	'Familiar': image set mouse was trained on (not novel) 'Novel 1': first exposure to novel image set 'Novel >1': novel image set but not first exposure	x			
file_id	int	lookup id to retrieve NWB file from S3 or the local cache.	x		x	
full_genotype	string	full genotype of transgenic mouse	x	x	x	
image_set	string	image set shown for a particular session	x			
imaging_depth	int	depth of the imaging plane in microns from the cortical surface	x			
indicator	string	transgenic mouse indicator	x	x	x	
mouse_id	int	unique identifier for a mouse	x	x	x	
ophys_container_id	int	unique identifier for an ophys container (grouping of ophys experiments by FOV)	x	x	x	
ophys_experiment_id	int	unique identifier for an ophys experiment	x	x	x	x
ophys_session_id	int	unique identifier for a ophys session	x	x	x	
passive	bool	True if the session was passive (no lickspout), False if the session was active (lickspout with water rewards)	x			
prior_exposures_to_image_set	float 64	Number of prior sessions (during training or ophys) where the mouse was exposed to the image set shown in the current session. Starts at 0 for first exposure		x	x	
prior_exposures_to_omissions	int 64	Number of prior sessions where the mouse was exposed to omissions. Starts at 0 for first exposure.		x	x	
prior_exposures_to_session_type	int 64	Number of previous sessions where the mouse was exposed to the current session type. Starts at 0 for first exposure	x	x	x	
project_code	string	Dataset variant the mouse belongs to (scientifica vs. mesoscope, trained on stim A vs. B)	x	x	x	
reporter_line	string	Reporter line for transgenic mouse	x	x	x	
session_number	float 64	For ophys (non-training) sessions only, session number is the number associated with the session type. (Eg. OPHYS_2_images_A_passive has session number 2)	x	x	x	
session_type	string	Visual stimulus type displayed during behavior session	x	x	x	
sex	string	Sex of the mouse	x	x	x	
targeted_structure	string	Visual area targeted by the experiment field of view	x	x	x	

Loading Dataset Objects

To load a behavior session from the cache:
```
behavior_session = cache.get_behavior_session(behavior_session_id)
```

To load an ophys_experiment from the cache:
```
ophys_experiment = cache.get_behavior_ophys_experiment(ophys_experiment_id)
```

Note: Attributes associated with both **ophys_experiment** datasets and **behavior_session** datasets (such as running speed) will be shown as **dataset.attribute** in the tables in this cheat sheet, while attributes associated only with the **ophys_experiment** datasets (such as cell fluorescence traces) will use **ophys_experiment.attribute**

Dataset attributes

color coded by data type

attributes and methods associated with a visual behavior dataset object.
* are only available for a behavior ophys experiment dataset object and not the behavior session dataset object.

attribute	type	description	attribute loading
* average projection	image	2D image of the microscope field of view, averaged across the experiment	dataset.average_projection
behavior session id	int	unique identifier for the behavior session	dataset.behavior_session_id
* cell specimen table	table	a dataframe with cell ROI information, including ROI masks	dataset.cell_specimen_table
* corrected fluorescence traces	table	traces of change in fluorescence	dataset.corrected_fluorescence_traces
* dff traces	table	traces of dff (change in fluorescence / fluorescence)	dataset.dff_traces
* events	table	a dataframe containing deconvolved events	dataset.events
* eye tracking	table	a table containing pupil area, pupil diameter, eye position, times of likely blinks, and other data derived from eye tracking movies	dataset.eye_tracking
licks	table	a dataframe containing lick times	dataset.licks
* max projection	image	2D maximum intensity projection image of the 2-photon movie	dataset.max_projection
metadata	dictionary	session specific metadata including genotype, session type, etc.	dataset.metadata
* motion correction	table	x and y offsets/shifts applied during motion correction	dataset.motion_correction
* ophy experiment id	int	unique identifier for the ophys experiment	dataset.ophys_experiment_id
* ophys session id	int	unique identifier for the ophys session	dataset.ophys_session_id
* ophys timestamps	array	timestamps associated with frames captured by the microscope	dataset.ophys_timestamps
performance metrics	table	a dataframe containing summary behavior performance metrics	dataset.get_performance_metrics()
raw running speed	table	a table containing timestamps and unfiltered running speed	dataset.raw_running_speed
rewards	table	timestamps of delivered rewards	dataset.rewards
reward rate	array	array of calculated reward rates over the course of the session	dataset.get_reward_rate()
rolling performance	table	a dataframe containing behavior performance metrics computed on a rolling basis across trials	dataset.get_rolling_performance_df()
running speed	table	a table containing timestamps and filtered running speed	dataset.running_speed
* segmentation mask image	image	a 2D binary image of all ROI masks	dataset.segmentation_mask_image
stimulus presentations	table	Start and end time of each stimulus presentation, info about which image was presented & whether the stimulus was omitted	dataset.stimulus_presentations
stimulus templates	table	a table containing arrays of presented stimuli	dataset.stimulus_templates
stimulus timestamps	array	timestamps of stimulus presentations on the monitor	dataset.stimulus_timestamps
task parameters	dictionary	parameters used to define task runtime behavior	dataset.task_parameters
trials	table	behavior trial data including change time and trial type	dataset.trials