

[Cloud AutoML Vision Object Detection](#)

# Managing datasets

A project can have multiple datasets, each used to train a separate model. You can get a [list of the available datasets](#) (#list-datasets), [get a specific dataset](#) (#get-dataset), [export a dataset](#) (#export-dataset) and can [delete a dataset](#) (#delete-dataset) you no longer need.

Use labels with user-defined metadata to organize your datasets. See [Using labels to organize Google Cloud Platform resources](#) (<https://goo.gl/xmQnxf>) for more information on labels as well as examples of labels.

## Listing datasets

A project can include numerous datasets. This section describes how to retrieve a list of the available datasets for a project.

WEB UI
REST & CMD LINE
MORE ▾

To see a list of the available datasets using the [Cloud AutoML Vision Object Detection UI](#) (<https://console.cloud.google.com/vision>), click the **Datasets** link at the top of the left navigation menu.

	Name	Type	Total Images	Labeled Images	Last updated	Status
<ul style="list-style-type: none"> <li>VISION</li> <li>DASHBOARD</li> <li style="background-color: #e0e0e0;">DATASETS</li> <li>MODELS</li> </ul>	untitled_1563557657669 <small>IOD6620964353350828032</small>	Object detection	0	0	Jul 19, 2019, 10:38:06 AM	Running: Importing images
	docs_july1_fishfooding <small>IOD8163482410097311744</small>	Object detection	225	225	Jul 18, 2019, 1:55:01 PM	Success: Training model

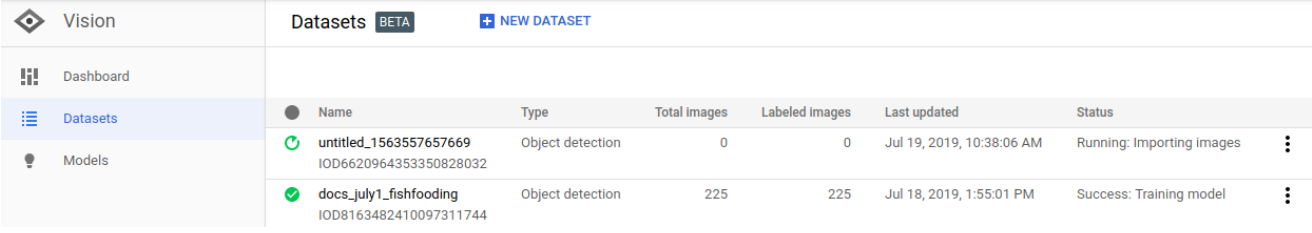
To see the datasets for a different project, select the project from the drop-down list on the left side of the title bar.

## Get a dataset

You can also get a specific dataset using a dataset ID.

WEB UI
REST & CMD LINE
MORE ▾

To see a list of the available datasets using the [Cloud AutoML Vision Object Detection UI](https://console.cloud.google.com/vision) (<https://console.cloud.google.com/vision>), click the **Datasets** link at the top of the left navigation menu.



Name	Type	Total Images	Labeled Images	Last updated	Status
untitled_1563557657669 IOD6620964353350828032	Object detection	0	0	Jul 19, 2019, 10:38:06 AM	Running: Importing images
docs_july1_fishfooding IOD8163482410097311744	Object detection	225	225	Jul 18, 2019, 1:55:01 PM	Success: Training model

To see the datasets for a different project, select the project from the drop-down list on the left side of the title bar.

Access a specific dataset by selecting its name from the list.

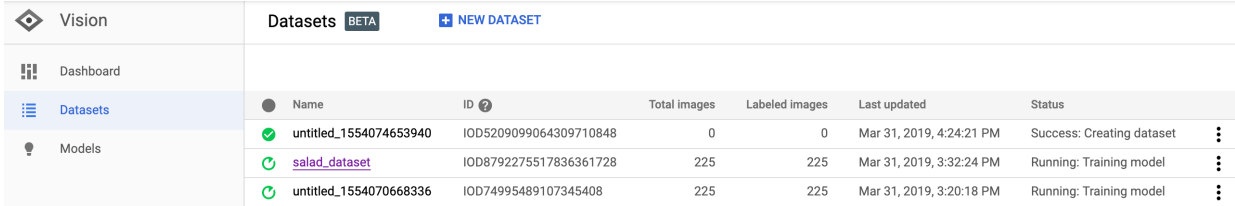
## Export a dataset

You can export a CSV file with all a dataset's information to a Google Cloud Storage bucket. This is particularly helpful for the cases where you added, deleted, or modified (<https://cloud.google.com/vision/automl/object-detection/docs/label>) training image annotations in the UI.

WEB UI
REST & CMD LINE
MORE ▾

To export a non-empty dataset, complete the following steps:

- Select the non-empty dataset from the **Datasets** page.



Name	ID	Total images	Labeled images	Last updated	Status
untitled_1554074653940	IOD5209099064309710848	0	0	Mar 31, 2019, 4:24:21 PM	Success: Creating dataset
salad_dataset	IOD8792275517836361728	225	225	Mar 31, 2019, 3:32:24 PM	Running: Training model
untitled_1554070668336	IOD74995489107345408	225	225	Mar 31, 2019, 3:20:18 PM	Running: Training model

Selecting the non-empty dataset will take you to the **Dataset details** page.

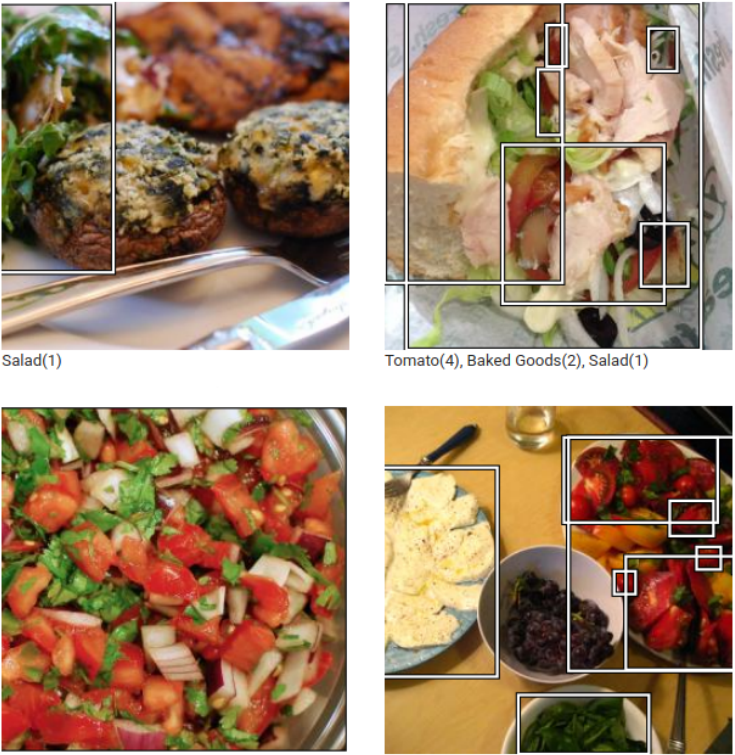
← salad\_dataset BETA || LABEL STATS EXPORT DATA

IMPORT **IMAGES** TRAIN EVALUATE TEST & USE

Filter images

All images	225
Labeled	225
Unlabeled	0
Filter labels	⋮
Baked Goods	44
Cheese	89
Salad	203
Seafood	48
Tomato	162

[ADD NEW LABEL](#)



Salad(1)

Tomato(4), Baked Goods(2), Salad(1)

Salad(1), Tomato(1)

Tomato(6), Salad(2), Cheese(1)

2. Select the **Export data** option at the top of the Dataset details page.

The screenshot shows the Google Cloud Vision console interface for a dataset named 'salad\_dataset'. The left sidebar contains navigation options: Vision, Dashboard, Datasets, and Models. The main content area has tabs for IMPORT, IMAGES, TRAIN, EVALUATE, and TEST & USE. The 'EXPORT DATA' button is highlighted with a red arrow. Below the tabs, there is a table of dataset statistics and a list of labels with their counts. The 'Filter images' and 'Filter labels' options are also visible.

Category	Count
All images	225
Labeled	225
Unlabeled	0
Filter labels	
Baked Goods	44
Cheese	89
Salad	203
Seafood	48
Tomato	162

ADD NEW LABEL

This opens a window where you can choose a Google Cloud Storage bucket location, or create a new bucket and designate it as the location to store the CSV file.

**Export data**

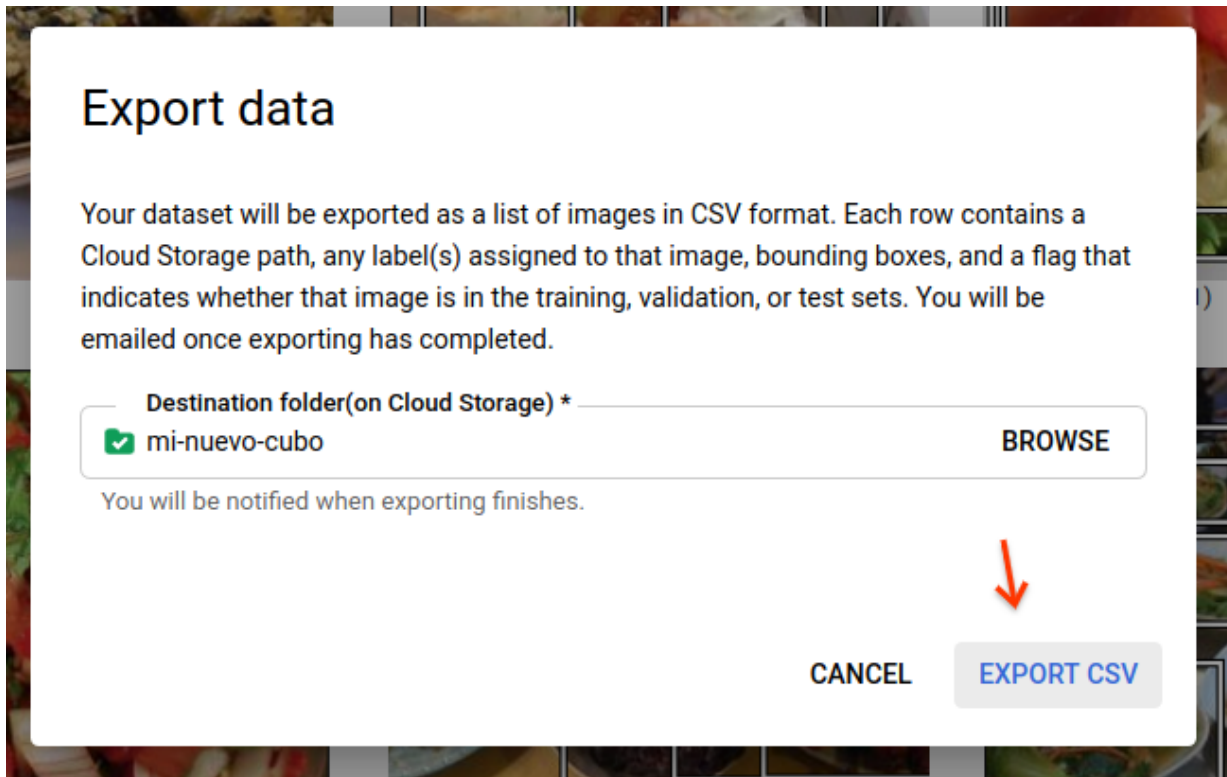
Your dataset will be exported as a list of images in CSV format. Each row contains a Cloud Storage path, any label(s) assigned to that image, bounding boxes, and a flag that indicates whether that image is in the training, validation, or test sets. You will be emailed once exporting has completed.

Destination folder(on Cloud Storage) \* **BROWSE**

Destination folder is required.

**CANCEL** **EXPORT CSV**

3. Select **Export CSV** after you have selected a new or existing Google Cloud Storage bucket location.



4. You will receive an email when the data export process has completed.

AutoML Vision finished exporting data from dataset "salad\_dataset" Inbox x



**AutoML Vision** <noreply-automl-vision@google.com> •

to me ▾

Hello AutoML Vision Customer,

AutoML Vision finished exporting data from dataset "salad\_dataset".

Additional Details:

Resource Name:

projects/5[redacted]76/locations/us-central1/datasets/IOD8[redacted]728

Operation State: Succeeded

To continue your progress, go back to your dataset using

[https://pantheon.corp.google.com/storage/browser/mi-nuevo-cubo?project=\[redacted\]-object-detection](https://pantheon.corp.google.com/storage/browser/mi-nuevo-cubo?project=[redacted]-object-detection)

Sincerely,

The Google Cloud AI Team

## Exported CSV format

The exported CSV file contains the same formatting as the [training data import CSV](https://cloud.google.com/vision/automl/object-detection/docs/csv-format) (<https://cloud.google.com/vision/automl/object-detection/docs/csv-format>):

```
set,path,label,x_min,y_min,x_max,y_min,x_max,y_max,x_min,y_max
```

This CSV file is saved in a created export folder that is distinguished by a unique timestamp. Below are some sample lines from an exported CSV file:

```
/export_data-salad_dataset-2019-05-29T18:12:18.750Z/image_object_detection_1.csv
```

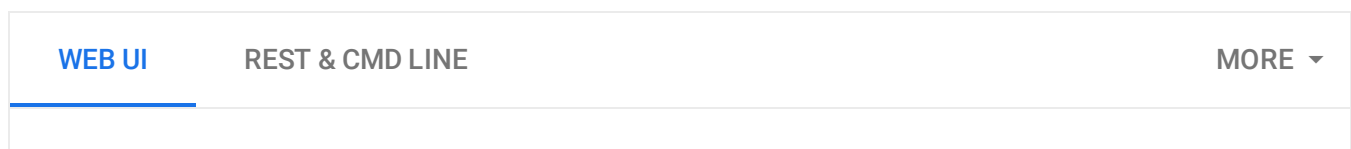
```
TRAIN,gs://my-storage-bucket/img/img009.jpg,Cheese,0.643239,0.362779,0.662498,0.362779,0.662498,0.362779,0.662498
TRAIN,gs://my-storage-bucket/img/img009.jpg,Salad,0.205697,0.255249,0.459074,0.255249,0.459074,0.255249,0.459074
TEST,gs://my-storage-bucket/img/img118.jpg,Cheese,0.320334,0.501238,0.726751,0.501238,0.726751,0.501238,0.726751
TEST,gs://my-storage-bucket/img/img118.jpg,Salad,0.0,0.037361,1.0,0.037361,1.0,0.037361,1.0,0.037361
TEST,gs://my-storage-bucket/img/img118.jpg,Cheese,0.358745,0.29076,0.740381,0.29076,0.740381,0.29076,0.740381
TRAIN,gs://my-storage-bucket/img/img375.jpg,Tomato,0.027274,0.41247,0.43122,0.41247,0.43122,0.41247,0.43122,0.41247
VALIDATION,gs://my-storage-bucket/img/img852.jpg,Tomato,0.716958,0.178534,0.805999,0.178534,0.805999,0.178534,0.805999,0.178534
VALIDATION,gs://my-storage-bucket/img/img852.jpg,Tomato,0.858044,0.297255,0.950847,0.297255,0.950847,0.297255,0.950847,0.297255
VALIDATION,gs://my-storage-bucket/img/img852.jpg,Tomato,0.199644,0.624155,0.321919,0.624155,0.321919,0.624155,0.321919,0.624155
VALIDATION,gs://my-storage-bucket/img/img852.jpg,Cheese,0.399672,0.277189,0.600955,0.277189,0.600955,0.277189,0.600955,0.277189
```

Note that there is still only one representation of a bounding box and its label per line. This information tells you the following:

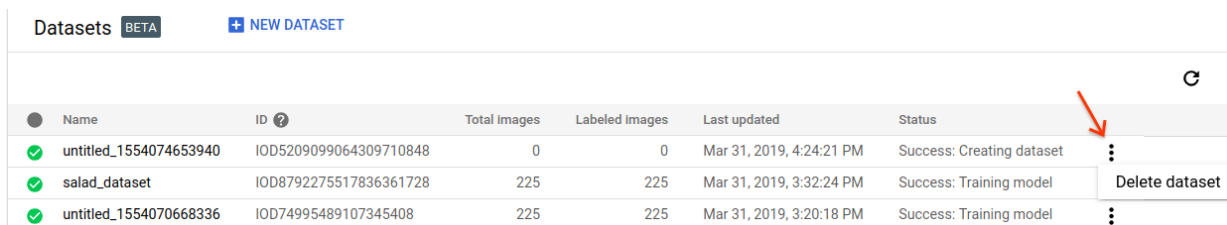
- `img009.jpg` - is in the TRAIN set and contains two bounding boxes with Cheese and Salad labels
- `img118.jpg` - is in the TEST set and has three bounding boxes with Cheese, Salad, and Cheese labels
- `img375.jpg` - is in the TRAIN set and has one bounding box with a Tomato label
- `img852.jpg` is in the VALIDATION set and has 4 distinct bounding boxes with Tomato, Tomato, Tomato, and Cheese labels

## Deleting a dataset

You can delete a dataset resource using the dataset's ID.



1. In the [Cloud AutoML Vision Object Detection UI](https://console.cloud.google.com/vision) (<https://console.cloud.google.com/vision>), click the **Datasets** link at the top of the left navigation menu to display the list of available datasets.



Datasets BETA + NEW DATASET

Name	ID	Total images	Labeled images	Last updated	Status	
untitled_1554074653940	IOD5209099064309710848	0	0	Mar 31, 2019, 4:24:21 PM	Success: Creating dataset	⋮
salad_dataset	IOD8792275517836361728	225	225	Mar 31, 2019, 3:32:24 PM	Success: Training model	Delete dataset
untitled_1554070668336	IOD74995489107345408	225	225	Mar 31, 2019, 3:20:18 PM	Success: Training model	⋮

2. Click the three-dot menu at the far right of the row you want to delete and select **Delete dataset**.
3. Click **Delete** in the confirmation dialog box.

Except as otherwise noted, the content of this page is licensed under the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) (<https://creativecommons.org/licenses/by/4.0/>), and code samples are licensed under the [Apache 2.0 License](https://www.apache.org/licenses/LICENSE-2.0) (<https://www.apache.org/licenses/LICENSE-2.0>). For details, see our [Site Policies](https://developers.google.com/terms/site-policies) (<https://developers.google.com/terms/site-policies>). Java is a registered trademark of Oracle and/or its affiliates.

Last updated January 22, 2020.