# Classifying Images of Clouds in the Cloud with AutoML Vision

# Overview

In this lab, you upload images to Cloud Storage and use them to train a custom model to recognize different types of clouds (cumulus, cumulonimbus, etc.).

**What you learn**

In this lab, you learn how to perform the following tasks:

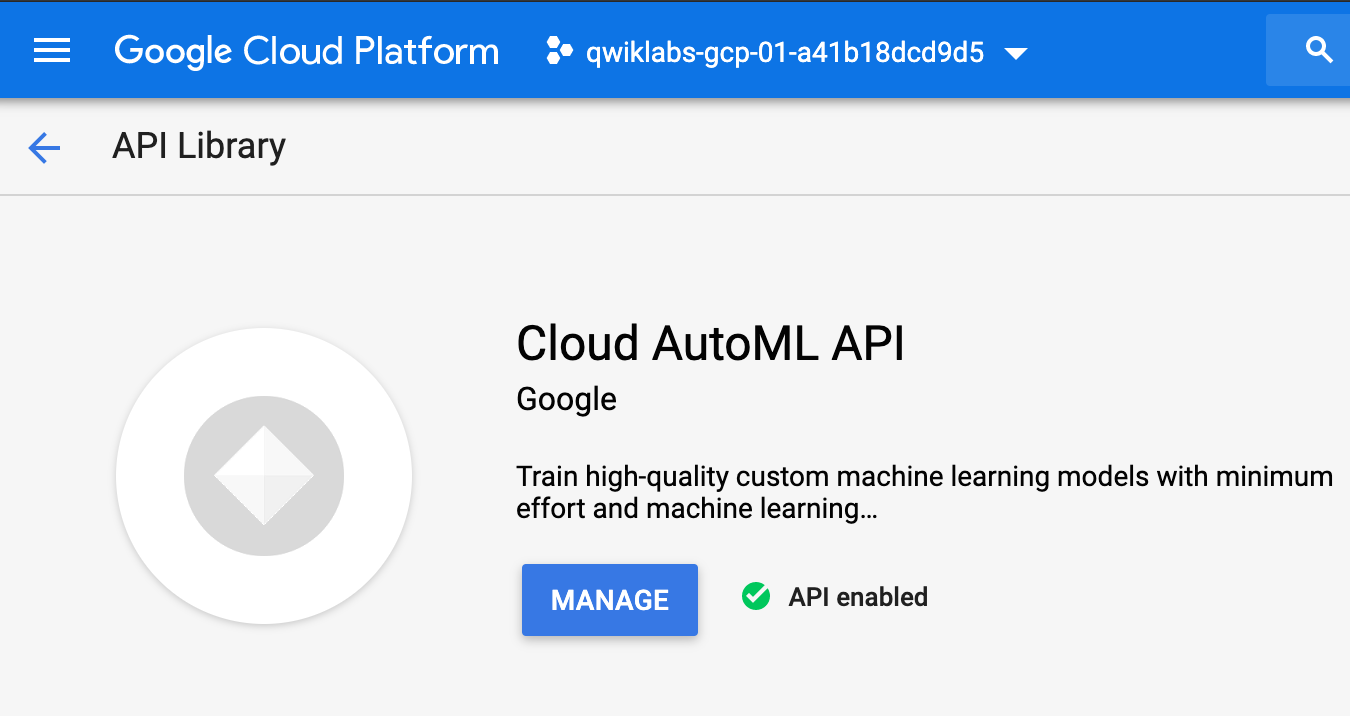
* Upload a labeled dataset to Cloud Storage and connect it to AutoML Vision with a CSV label file
* Train a model with AutoML Vision and evaluate its accuracy
* Generate predictions on your trained model

Task 1. Set up AutoML Vision

AutoML Vision provides an interface for all the steps in training an image classification model and generating predictions on it. Start by enabling the AutoML API.

Open the navigation menu and and select **APIs & Services** > **Library**. In the search bar type in "Cloud AutoML API". Click on the **Cloud AutoML API** result and then click **Enable**.

This may take a minute. You should now be on the following page (ensure that the **Manage** button appears and **API enabled** is also displayed):



### **Create a Cloud Storage bucket for you**

### **Create a Cloud Storage bucket for your training data**

In Cloud Shell, paste the below command to make a new bucket to hold your training. We use the magic variable $DEVSHELL\_PROJECT\_ID, which knows your current project, and simply add -vcm to the end.

Run the below command in Cloud Shell:

gsutil mb -p $DEVSHELL\_PROJECT\_ID \

-c regional \

-l us-central1 \

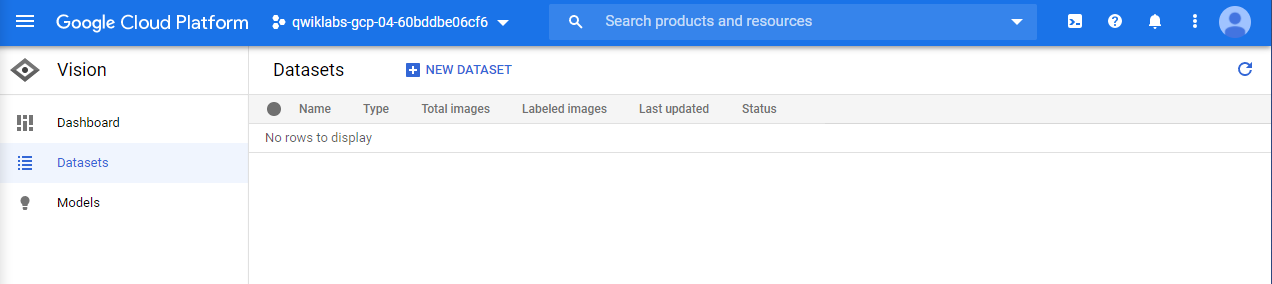
gs://$DEVSHELL\_PROJECT\_ID-vcm/

content\_copy

Leave your Cloud Shell window open for additional steps to follow.

Open a new browser tab and navigate to the [AutoML UI](https://console.cloud.google.com/vision/datasets" \t "_blank).

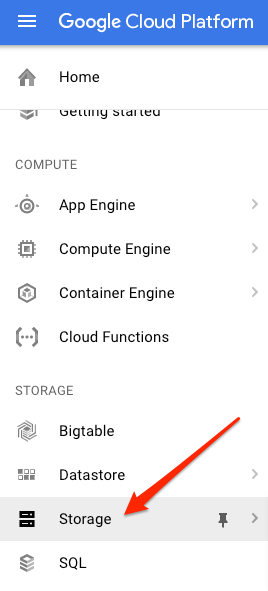
You will be taken to the AutoML Vision datasets page once the APIs are verified.



## Task 2. Upload training images to Cloud Storage

In order to train a model to classify images of clouds, you need to provide labeled training data so the model can develop an understanding of the image features associated with different types of clouds. In this example, your model will learn to classify three different types of clouds: cirrus, cumulus, and cumulonimbus. To use AutoML Vision you need to put your training images in Cloud Storage.

In the Cloud Console, open the **Navigation menu** and select **Storage** > **Browser**:



Once there, you should see the bucket from the last step.

The training images are publicly available in a Cloud Storage bucket. Use the gsutil command-line utility for Cloud Storage to copy the training images into your bucket:

gsutil -m cp -r gs://automl-codelab-clouds/\* gs://$DEVSHELL\_PROJECT\_ID-vcm/

content\_copy

Once copying is complete you can view the three types of clouds you have images for:

gsutil ls gs://$DEVSHELL\_PROJECT\_ID-vcm/

content\_copy

Which types of clouds do you have training data (images) for?

checkCumulus

checkCumulonimbus



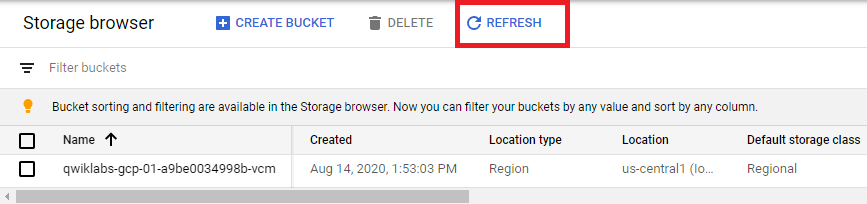
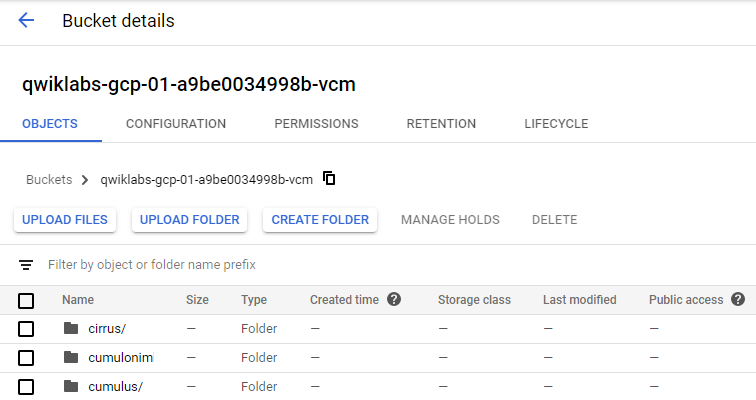
Stratus

checkCirrus

Submit

### **Optional: View the images using the Cloud Storage Console UI**

When the images finish copying, click the **Refresh** button at the top of the Cloud Storage browser. Then click on your bucket name. You should see 3 folders of photos for each of the 3 different cloud types to be classified:

If you click on the individual image files in each folder, and then click once more when you see the URL, you can see the photos you'll be using to train your model for each type of cloud.

## Task 3. Create an AutoML Vision training dataset

Now that your training data is in Cloud Storage, you need a way for AutoML Vision to access it. You'll create a CSV file where each row contains a URL to a training image and the associated label for that image. This CSV file has been created for you; you just need to update it with your bucket name.

Run the following commands which:

* Copy the template file to your Cloud Shell instance
* Update the CSV with the files in your project
* Upload this file to your Cloud Storage bucket
* Show the bucket to confirm the data.csv file is present

gsutil cp gs://automl-codelab-metadata/data.csv .

head --lines=10 data.csv

sed -i -e "s/placeholder/$DEVSHELL\_PROJECT\_ID-vcm/g" ./data.csv

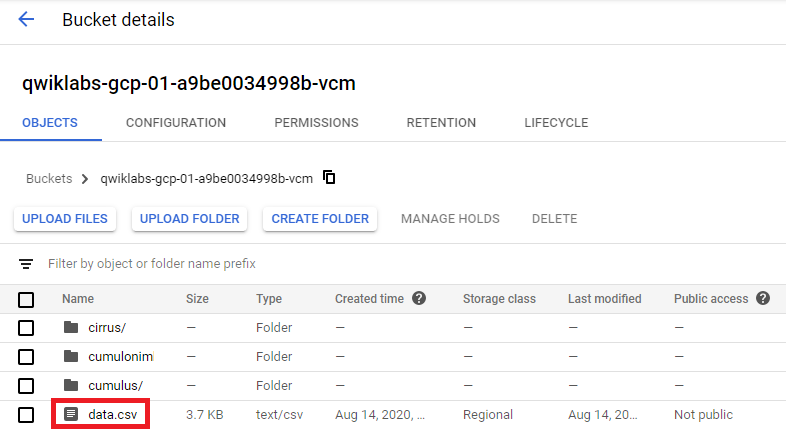
head --lines=10 data.csv

gsutil cp ./data.csv gs://$DEVSHELL\_PROJECT\_ID-vcm/

gsutil ls gs://$DEVSHELL\_PROJECT\_ID-vcm/

content\_copy

Confirm that you see the data.csv file in your bucket.



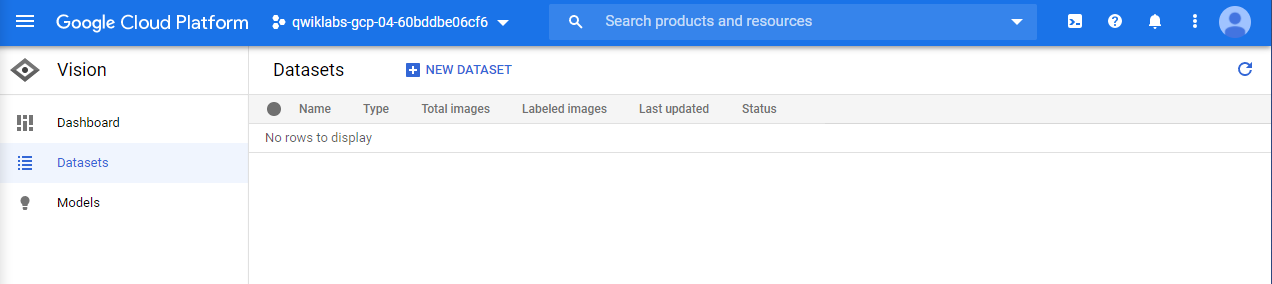
View all the folders and files in your bucket you can add a wildcard to gsutil ls like so:

gsutil ls gs://$DEVSHELL\_PROJECT\_ID-vcm/\*

content\_copy

Highlight and copy the location of your data file to your clipboard which will look similar to: gs://qwiklabs-gcp-your-project-id-will-be-here-vcm/data.csv

Navigate back to the [AutoML Vision](https://console.cloud.google.com/vision/datasets" \t "_blank) datasets page.

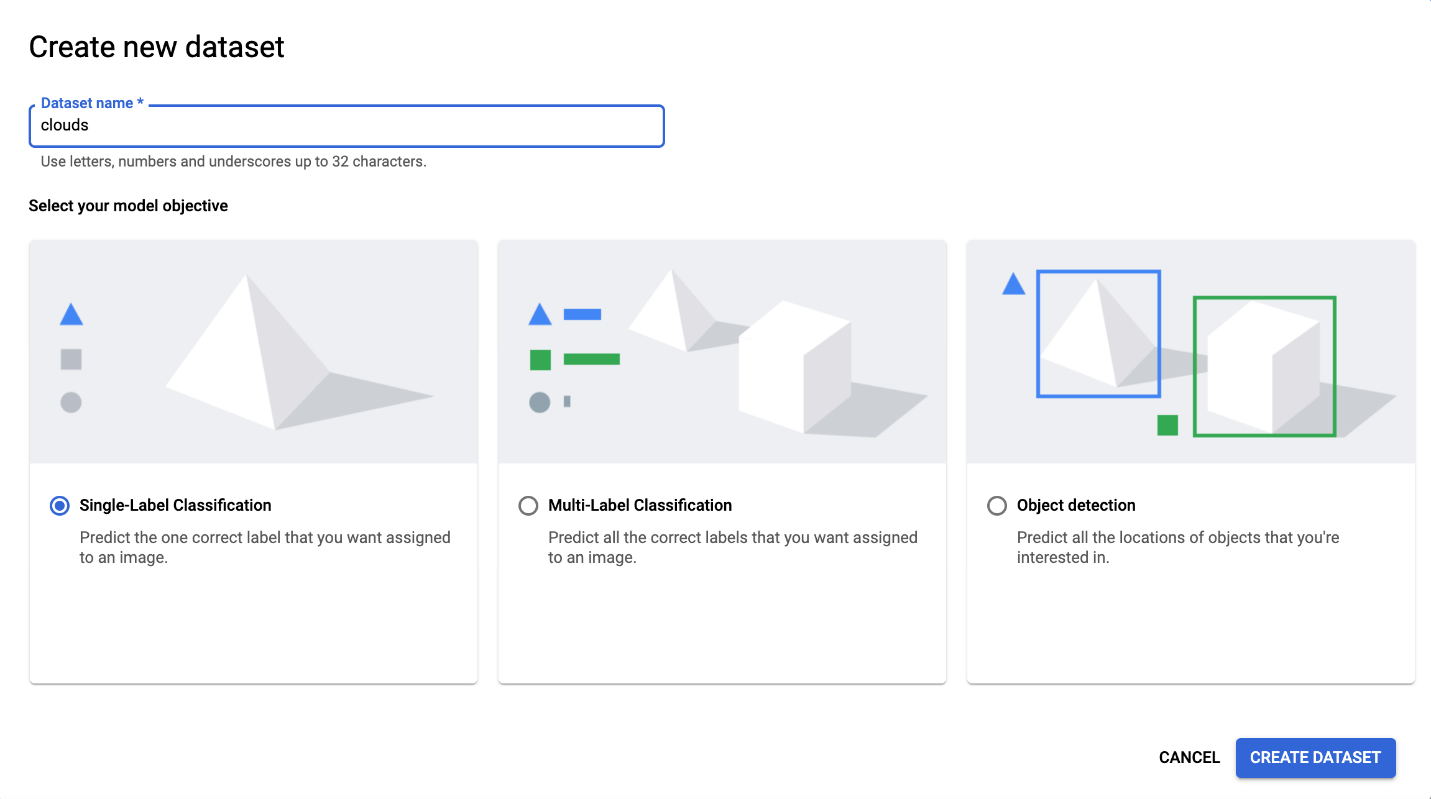


At the top of the Cloud Console, click **+ New dataset**.

Type **clouds** for the dataset name.

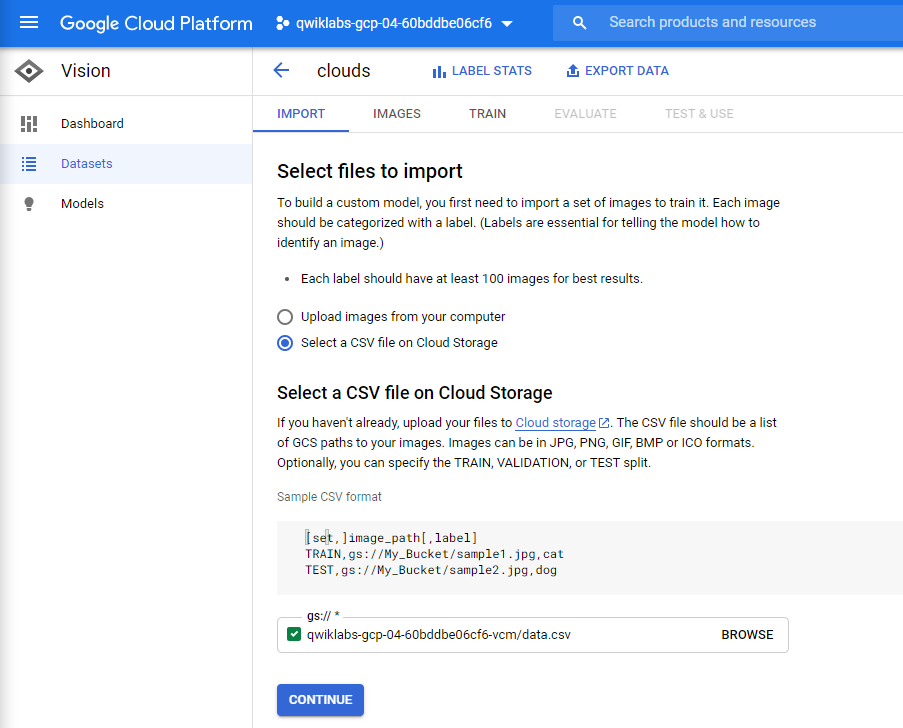
Leave **Single-label Classification** checked.

Click **Create dataset** to continue.



On the next screen you will choose the location of your training images (the ones you uploaded in the previous step).

Choose **Select a CSV file on Cloud Storage** and add the file name to the URL for the file that is in your clipboard from the previous step. You may also use the browse function to find the csv file. Once you see the white in green checkbox you may select **Continue** to proceed.

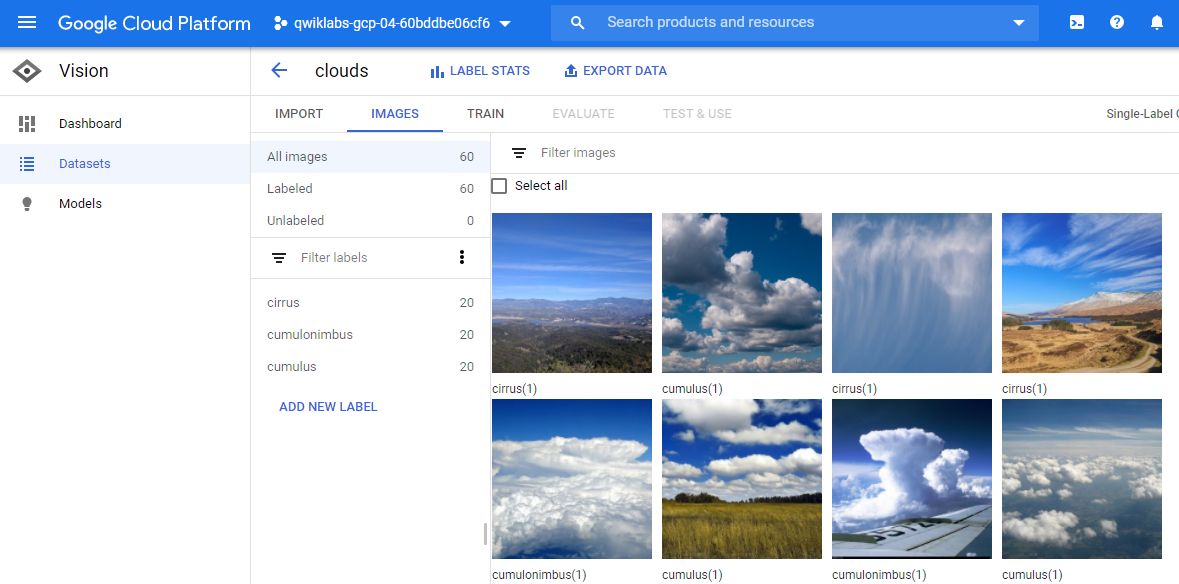


Once the import has completed click the **Images** tab to see the images in your dataset.

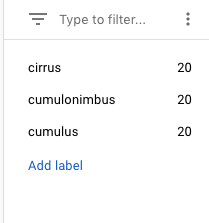
It will take 8 to 12 minutes while the image metadata is processed ("Running: Importing Images" will appear on the screen). Once complete, the images will appear by category.

## Task 4. Inspect the images

Next, proceed with a brief examination of the images.

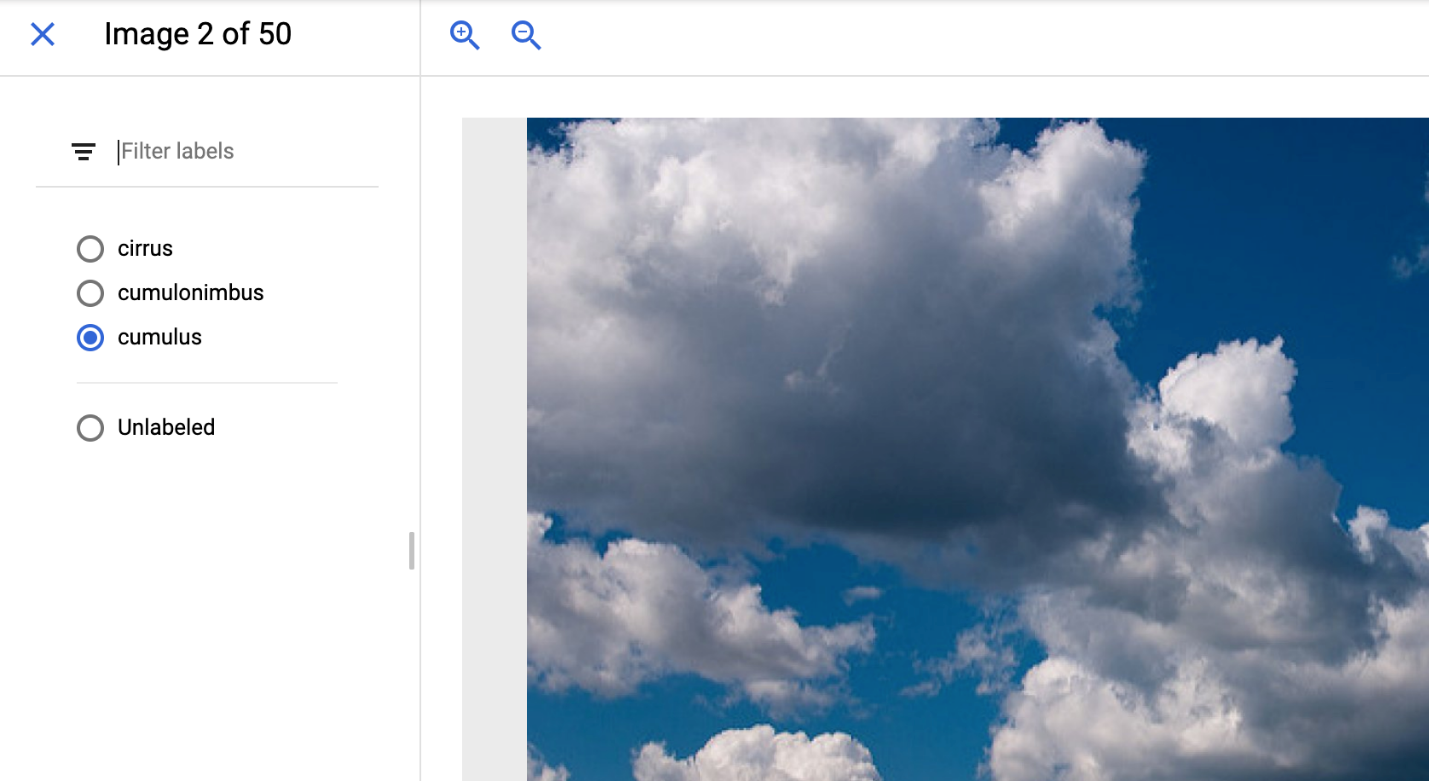


Try filtering by different labels in the left menu (i.e. click cumulus) to review the training images:

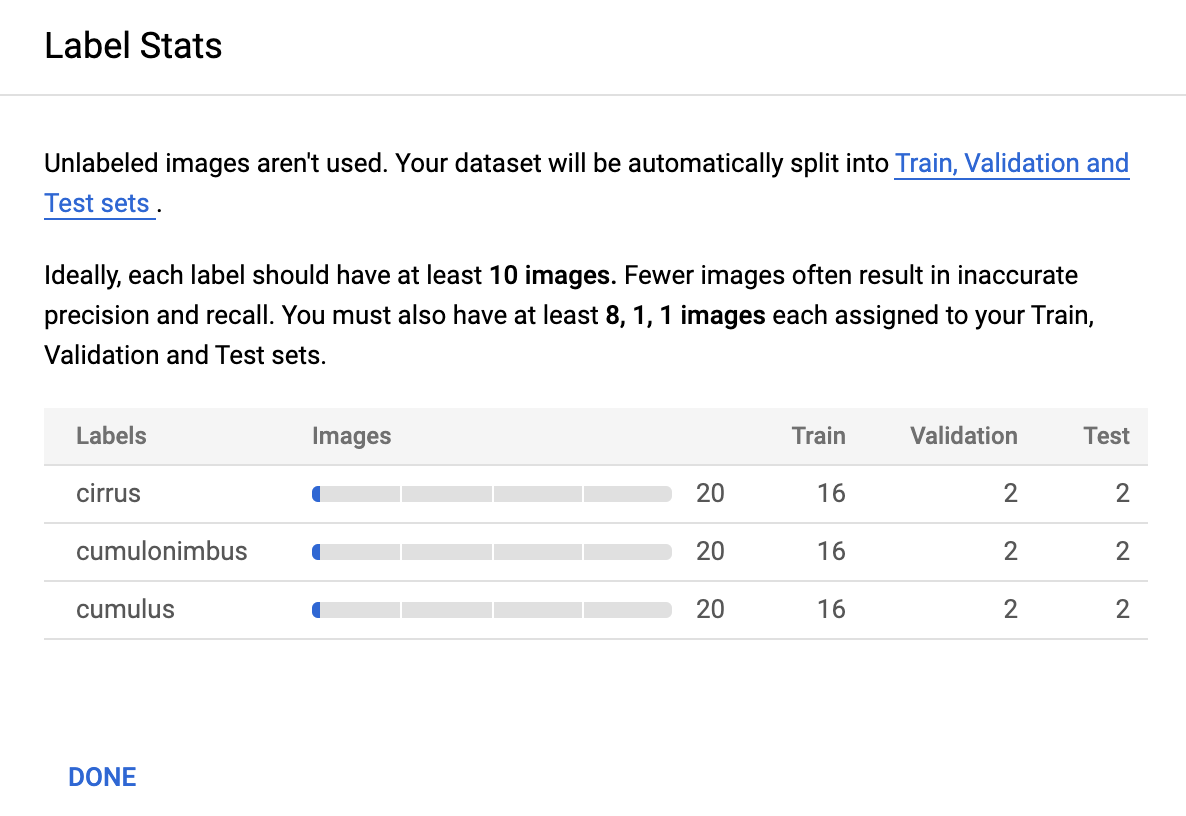


**Note:**If you were building a production model, you'd want at least 100 images per label to ensure high accuracy. This is just a demo so only 20 images of each type were used so the model could train quickly.

If any images are labeled incorrectly you can click on them to switch the label or delete the image from your training set:



To see a summary of how many images you have for each label, click on **Label stats**. You should see the following pop-out box show up on the right side of your browser. Press **Done** after reviewing the list.

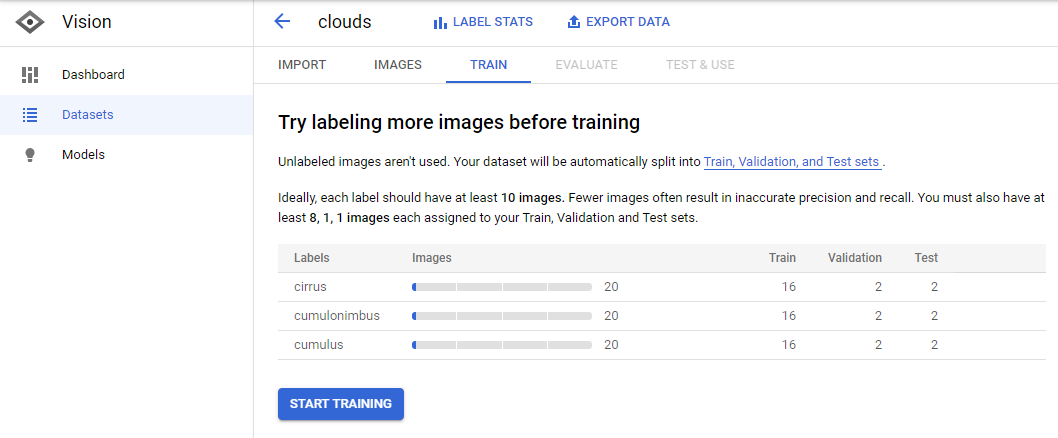


**Note:**If you are working with a dataset that isn't already labeled, AutoML Vision provides an in-house human labeling service.

## Task 5. Train your model

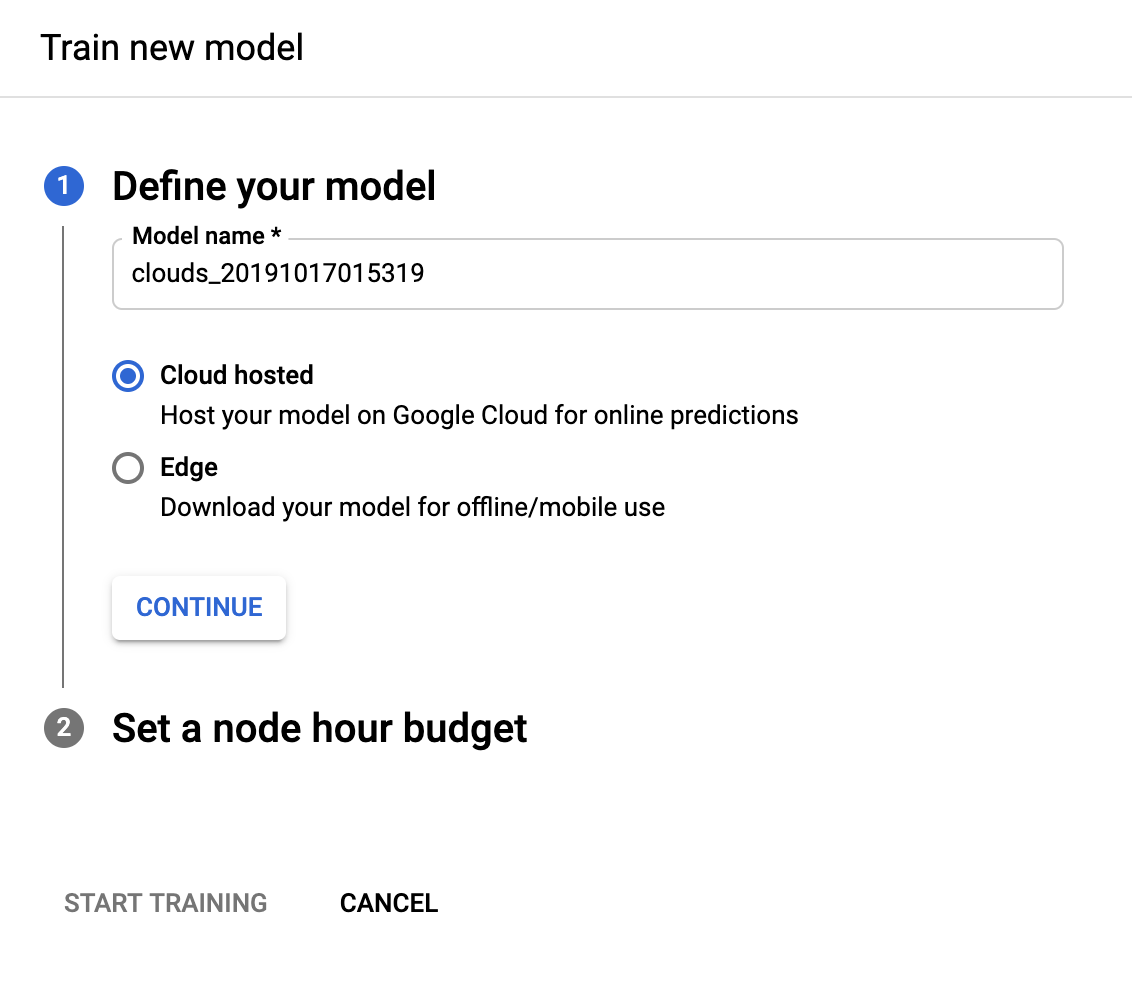
You're ready to start training your model! AutoML Vision handles this for you automatically, without requiring you to write any of the model code.

To train your clouds model, go to the **Train** tab and click **Start training**.



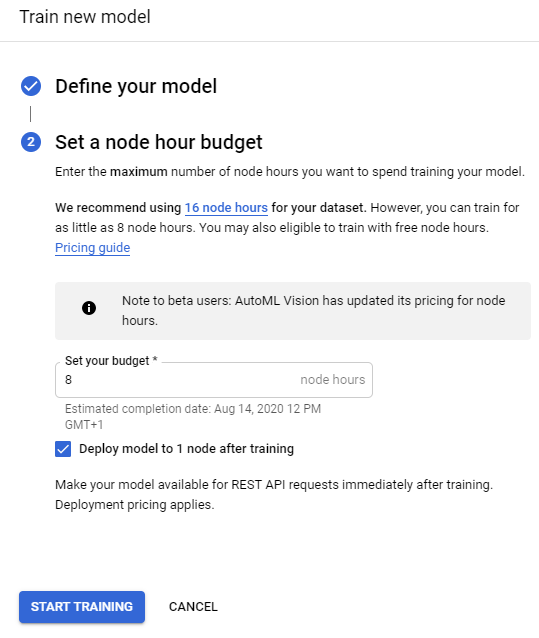
Enter a name for your model, or use the default auto-generated name.

Leave **Cloud hosted** selected and click **Continue**.



For the next step, type the value "**8**" into the **Set your budget \*** box and check **Deploy model to 1 node after training**. This process (auto-deploy) will make your model immediately available for predictions after testing is complete.

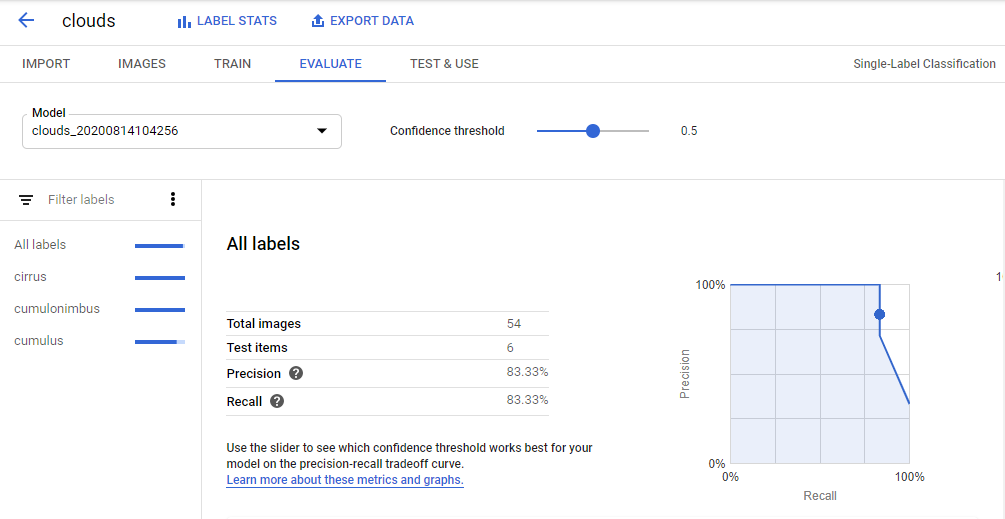
Click **Start training**.



**Note:** Training this custom model can be expected to take over an hour to complete (**55 to 90 minutes** on average). The total training time includes node training time as well as infrastructure set up and tear down. To get full credit for the lab you do not need to wait for training to complete and can simply review at the below screenshots from Evalution and Prediction.

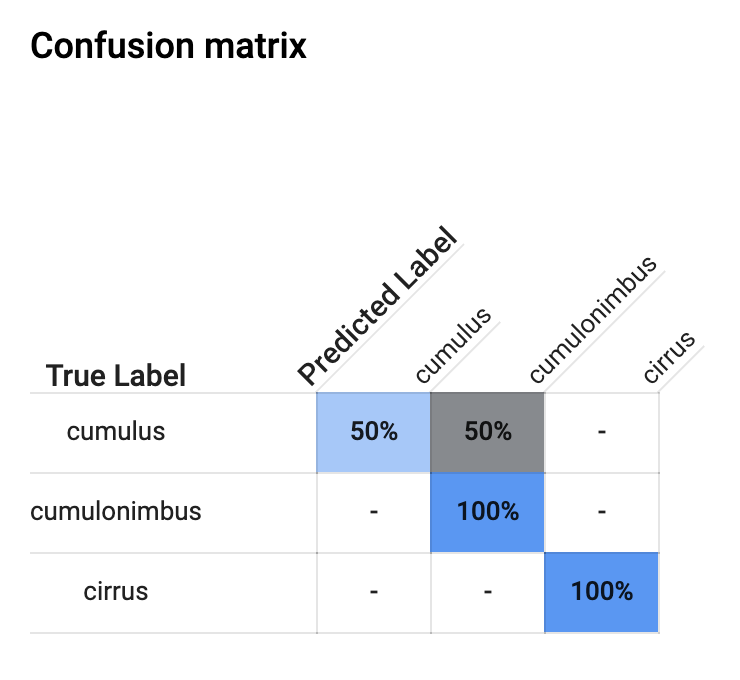
## Task 6. Evaluate your model

After training is complete, select the **Evaluate** tab. Here you'll see information about Precision and Recall of the model. It should resemble the following:



You can also adjust the **Confidence threshold** slider to see its impact.

Finally, scroll down to take a look at the **Confusion matrix**.



This tab provides some common machine learning metrics to evaluate your model accuracy and see where you can improve your training data. Since the focus for this lab was not on accuracy, move on to the next section about predictions section. Feel free to browse the accuracy metrics on your own.

## Task 7. Generate predictions

Now it's time for the most important part: generating predictions on your trained model using data it hasn't seen before.

There are a few ways to generate predictions. In this lab you use the UI to upload images. You'll see how your model does classifying these two images (the first is a cirrus cloud, the second is a cumulonimbus).

First, download these images to your local machine by right-clicking on each of them (**Note**: You may want to assign a simple name like 'Image1' and 'Image2' to assist with uploading later):

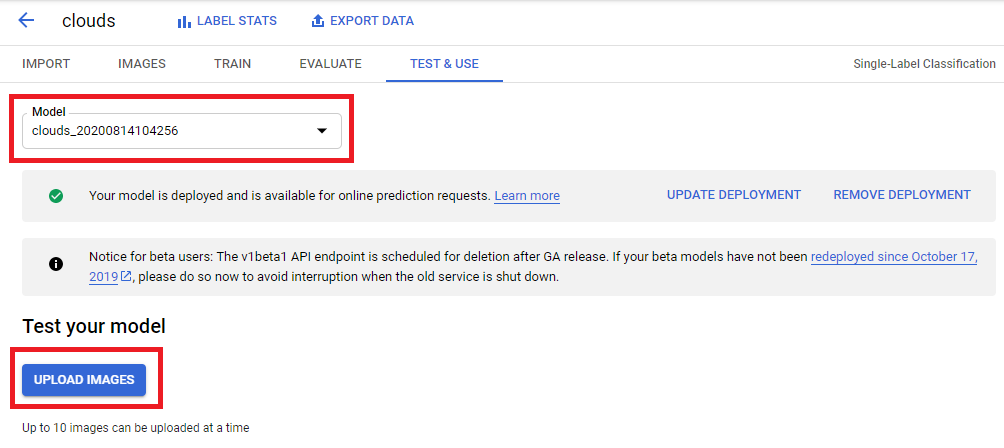




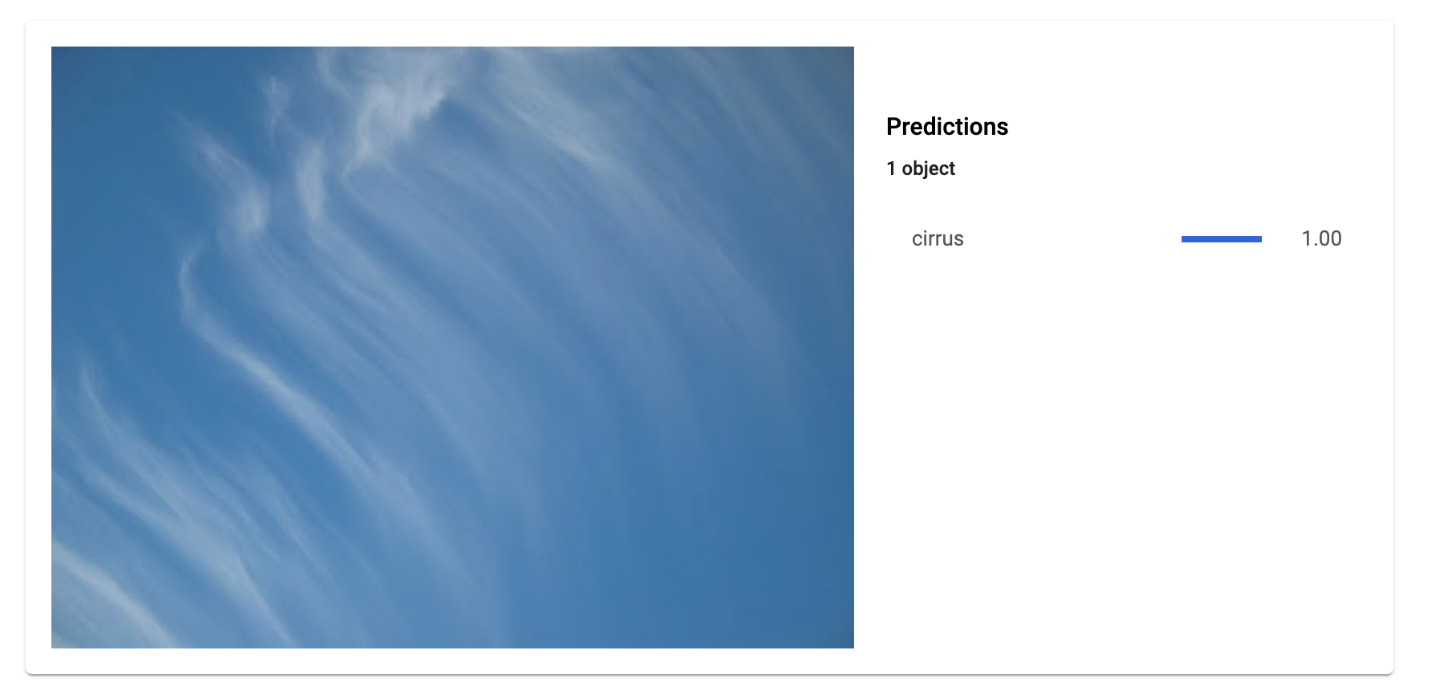
Navigate to the **Test & Use** tab in the AutoML UI:

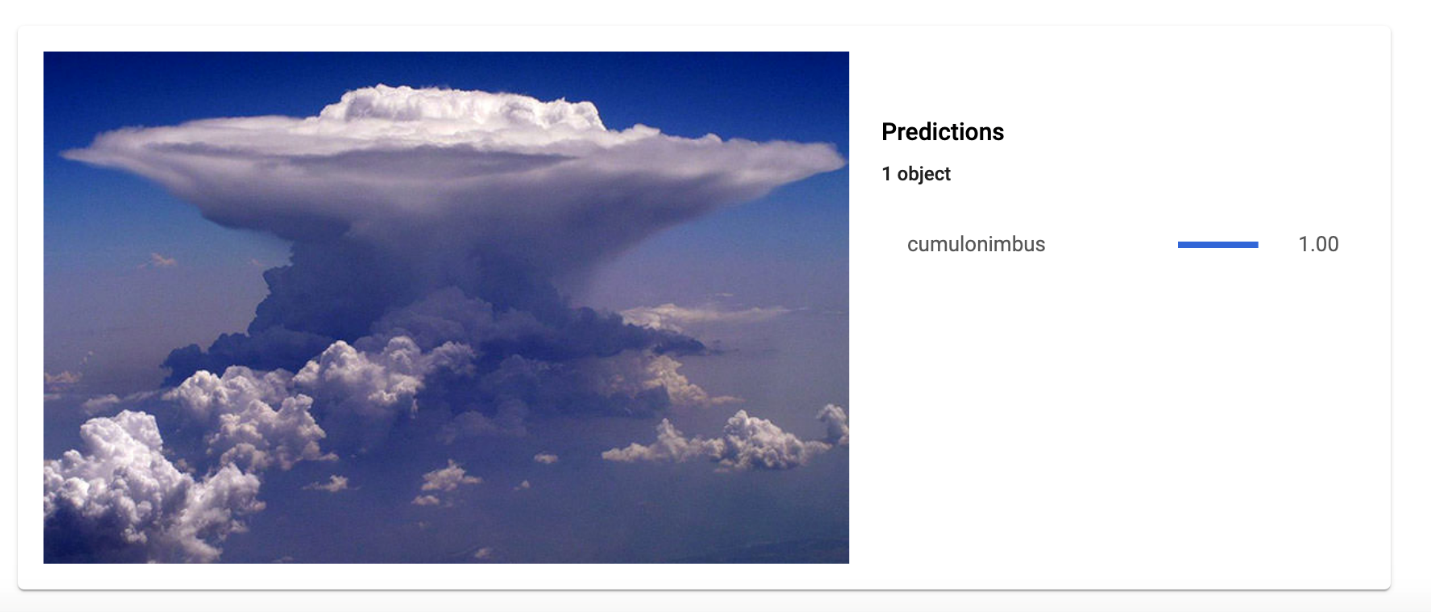
On this page you will see that the model you just trained and deployed is listed in the **Model** pick list.

Click **Upload images** and upload the cloud sample images you just saved to your local disk (you can select both images at the same time).



When the prediction request completes you should see something like the following:





Excellent - the model classified each type of cloud correctly!

Congratulations!

You've learned how to train your own custom machine learning model and generate predictions on it through the web UI. Now you've got what it takes to train a model on your own image dataset.

#### **What was covered**

* Uploading training images to Cloud Storage and creating a CSV for AutoML Vision to find these images.
* Reviewing labels and training a model in the AutoML Vision UI.
* Generating predictions on new cloud images.

END YOUR LAB