

[Cloud AutoML Vision](#)

# REST Resource: projects.locations.datasets

## Resource: Dataset

A workspace for solving a single, particular machine learning (ML) problem. A workspace contains examples that may be annotated.

### JSON representation

## JSON representation

```
{
  "name": string,
  "displayName": string,
  "description": string,
  "exampleCount": number,
  "createTime": string,
  "etag": string,

  // Union field dataset_metadata can be only one of the following:
  "translationDatasetMetadata": {
    object (TranslationDatasetMetadata (https://cloud.google.com/vision/automl/docs/reference/
  },
  "imageClassificationDatasetMetadata": {
    object (ImageClassificationDatasetMetadata (https://cloud.google.com/vision/automl/docs/r
  },
  "textClassificationDatasetMetadata": {
    object (TextClassificationDatasetMetadata (https://cloud.google.com/vision/automl/docs/re
  },
  "imageObjectDetectionDatasetMetadata": {
    object (ImageObjectDetectionDatasetMetadata (https://cloud.google.com/vision/automl/docs/
  },
  "videoClassificationDatasetMetadata": {
    object (VideoClassificationDatasetMetadata (https://cloud.google.com/vision/automl/docs/r
  },
  "videoObjectTrackingDatasetMetadata": {
    object (VideoObjectTrackingDatasetMetadata (https://cloud.google.com/vision/automl/docs/r
  },
  "textExtractionDatasetMetadata": {
    object (TextExtractionDatasetMetadata (https://cloud.google.com/vision/automl/docs/referen
  },
  "textSentimentDatasetMetadata": {
    object (TextSentimentDatasetMetadata (https://cloud.google.com/vision/automl/docs/reference
  },
  "tablesDatasetMetadata": {
    object (TablesDatasetMetadata (https://cloud.google.com/vision/automl/docs/reference/rest/v1
  }
  // End of list of possible types for union field dataset_metadata.
}
```

## Fields

Fields	
<b>name</b>	<p><b>string</b></p> <p>Output only. The resource name of the dataset. Form: <b>projects/{project_id}/locations/{locationId}/datasets/{datasetId}</b></p>
<b>displayName</b>	<p><b>string</b></p> <p>Required. The name of the dataset to show in the interface. The name can be up to 32 characters long and can consist only of ASCII Latin letters A-Z and a-z, underscores (_), and ASCII digits 0-9.</p>
<b>description</b>	<p><b>string</b></p> <p>User-provided description of the dataset. The description can be up to 25000 characters long.</p>
<b>exampleCount</b>	<p><b>number</b></p> <p>Output only. The number of examples in the dataset.</p>
<b>createTime</b>	<p><b>string (Timestamp format)</b> (<a href="https://developers.google.com/protocol-buffers/docs/reference/google.protobuf#google.protobuf.Timestamp">https://developers.google.com/protocol-buffers/docs/reference/google.protobuf#google.protobuf.Timestamp</a>)</p> <p>Output only. Timestamp when this dataset was created.</p> <p>A timestamp in RFC3339 UTC "Zulu" format, accurate to nanoseconds. Example: "2014-10-02T15:01:23.045123456Z".</p>
<b>etag</b>	<p><b>string</b></p> <p>Used to perform consistent read-modify-write updates. If not set, a blind "overwrite" update happens.</p>

Union field **dataset\_metadata**. Required. The dataset metadata that is specific to the problem type. **dataset\_metadata** can be only one of the following:

## Fields

<b>translationDatasetMetadata</b>	<p>object (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TranslationDatasetMetadata">TranslationDatasetMetadata</a> (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TranslationDatasetMetadata">https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TranslationDatasetMetadata</a>) )</p> <p>Metadata for a dataset used for translation.</p>
<b>imageClassificationDatasetMetadata</b>	<p>object (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#ImageClassificationDatasetMetadata">ImageClassificationDatasetMetadata</a> (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#ImageClassificationDatasetMetadata">https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#ImageClassificationDatasetMetadata</a>) )</p> <p>Metadata for a dataset used for image classification.</p>
<b>textClassificationDatasetMetadata</b>	<p>object (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TextClassificationDatasetMetadata">TextClassificationDatasetMetadata</a> (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TextClassificationDatasetMetadata">https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TextClassificationDatasetMetadata</a>) )</p> <p>Metadata for a dataset used for text classification.</p>
<b>imageObjectDetectionDatasetMetadata</b>	<p>object (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#ImageObjectDetectionDatasetMetadata">ImageObjectDetectionDatasetMetadata</a> (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#ImageObjectDetectionDatasetMetadata">https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#ImageObjectDetectionDatasetMetadata</a>) )</p> <p>Metadata for a dataset used for image object detection.</p>
<b>videoClassificationDatasetMetadata</b>	<p>object (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#VideoClassificationDatasetMetadata">VideoClassificationDatasetMetadata</a> (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#VideoClassificationDatasetMetadata">https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#VideoClassificationDatasetMetadata</a>) )</p> <p>Metadata for a dataset used for video classification.</p>
<b>videoObjectTrackingDatasetMetadata</b>	<p>object (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#VideoObjectTrackingDatasetMetadata">VideoObjectTrackingDatasetMetadata</a> (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#VideoObjectTrackingDatasetMetadata">https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#VideoObjectTrackingDatasetMetadata</a>) )</p> <p>Metadata for a dataset used for video object tracking.</p>

## Fields

<b>textExtractionDatasetMetadata</b>	<p>object (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TextExtractionDatasetMetadata">TextExtractionDatasetMetadata</a> (https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TextExtractionDatasetMetadata))</p> <p>Metadata for a dataset used for text extraction.</p>
<b>textSentimentDatasetMetadata</b>	<p>object (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TextSentimentDatasetMetadata">TextSentimentDatasetMetadata</a> (https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TextSentimentDatasetMetadata))</p> <p>Metadata for a dataset used for text sentiment.</p>
<b>tablesDatasetMetadata</b>	<p>object (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TablesDatasetMetadata">TablesDatasetMetadata</a> (https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TablesDatasetMetadata))</p> <p>Metadata for a dataset used for Tables.</p>

## TranslationDatasetMetadata

Dataset metadata that is specific to translation.

### JSON representation

```
{
  "sourceLanguageCode": string,
  "targetLanguageCode": string
}
```

### Fields

<b>sourceLanguageCode</b>	<p><b>string</b></p> <p>Required. The BCP-47 language code of the source language.</p>
---------------------------	--

## Fields

<b>targetLanguageCode</b>	<b>string</b>
	Required. The BCP-47 language code of the target language.

## ImageClassificationDatasetMetadata

Dataset metadata that is specific to image classification.

### JSON representation

```
{
  "classificationType": enum (ClassificationType (https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#ClassificationType))
}
```

## Fields

<b>classificationType</b>	<b>enum (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#ClassificationType">ClassificationType</a> (https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#ClassificationType))</b>
	Required. Type of the classification problem.

## ClassificationType

Type of the classification problem.

### Enums

<b>CLASSIFICATION_TYPE_UNSPECIFIED</b>	An un-set value of this enum.
<b>MULTICLASS</b>	At most one label is allowed per example.
<b>MULTILABEL</b>	Multiple labels are allowed for one example.

## TextClassificationDatasetMetadata

Dataset metadata for classification.

### JSON representation

```
{  
  "classificationType": enum (ClassificationType (https://cloud.google.com/vision/automl/docs:  
})
```

### Fields

<b>classificationType</b>	enum ( <u>ClassificationType</u> ( <a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#ClassificationType">https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#ClassificationType</a> ) )  Required. Type of the classification problem.
---------------------------	---

## ImageObjectDetectionDatasetMetadata

Dataset metadata specific to image object detection.

## VideoClassificationDatasetMetadata

Dataset metadata specific to video classification. All Video Classification datasets are treated as multi label.

## VideoObjectTrackingDatasetMetadata

Dataset metadata specific to video object tracking.

## TextExtractionDatasetMetadata

Dataset metadata that is specific to text extraction

## TextSentimentDatasetMetadata

Dataset metadata for text sentiment.

### JSON representation

```
{  
  "sentimentMax": number  
}
```

### Fields

**sentimentMax**

**number**

Required. A sentiment is expressed as an integer ordinal, where higher value means a more positive sentiment. The range of sentiments that will be used is between 0 and sentimentMax (inclusive on both ends), and all the values in the range must be represented in the dataset before a model can be created. sentimentMax value must be between 1 and 10 (inclusive).

## TablesDatasetMetadata

Metadata for a dataset used for AutoML Tables.

### JSON representation



## JSON representation

```
{
  "primaryTableSpecId": string,
  "targetColumnSpecId": string,
  "weightColumnSpecId": string,
  "mlUseColumnSpecId": string,
  "targetColumnCorrelations": {
    string: {
      object(CorrelationStats (https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1
    ),
    ...
  },
  "statsUpdateTime": string
}
```

## Fields

<b>primaryTableSpecId</b>	<b>string</b>  Output only. The tableSpecId of the primary table of this dataset.
<b>targetColumnSpecId</b>	<b>string</b>  columnSpecId of the primary table's column that should be used as the training & prediction target. This column must be non-nullable and have one of following data types (otherwise model creation will error): <ul style="list-style-type: none"> <li>• CATEGORY</li> <li>• FLOAT64</li> </ul> <p>If the type is CATEGORY , only up to 100 unique values may exist in that column across all rows.</p> <p>NOTE: Updates of this field will instantly affect any other users concurrently working with the dataset.</p>

## Fields

<b>weightColumnSpecId</b>	<b>string</b>  columnSpecId of the primary table's column that should be used as the weight column, i.e. the higher the value the more important the row will be during model training. Required type: FLOAT64. Allowed values: 0 to 10000, inclusive on both ends; 0 means the row is ignored for training. If not set all rows are assumed to have equal weight of 1. NOTE: Updates of this field will instantly affect any other users concurrently working with the dataset.
<b>mlUseColumnSpecId</b>	<b>string</b>  columnSpecId of the primary table column which specifies a possible ML use of the row, i.e. the column will be used to split the rows into TRAIN, VALIDATE and TEST sets. Required type: STRING. This column, if set, must either have all of <b>TRAIN</b> , <b>VALIDATE</b> , <b>TEST</b> among its values, or only have <b>TEST</b> , <b>UNASSIGNED</b> values. In the latter case the rows with <b>UNASSIGNED</b> value will be assigned by AutoML. Note that if a given ml use distribution makes it impossible to create a "good" model, that call will error describing the issue. If both this columnSpecId and primary table's timeColumnSpecId are not set, then all rows are treated as <b>UNASSIGNED</b> . NOTE: Updates of this field will instantly affect any other users concurrently working with the dataset.

## Fields

<b>targetColumnCorrelations</b>	<p><b>map (key: string, value: object (<u>CorrelationStats</u> (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets.tableSpecs.columnSpecs#ColumnSpec.CorrelationStats">https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets.tableSpecs.columnSpecs#ColumnSpec.CorrelationStats</a>)))</b></p> <p>Output only. Correlations between</p> <p><b><u>TablesDatasetMetadata.target_column_spec_id</u></b> (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TablesDatasetMetadata.FIELDS.target_column_spec_id">https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TablesDatasetMetadata.FIELDS.target_column_spec_id</a>), and other columns of the</p> <p><b><u>TablesDatasetMetadatapprimary_table</u></b> (<a href="https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TablesDatasetMetadata.FIELDS.primary_table_spec_id">https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TablesDatasetMetadata.FIELDS.primary_table_spec_id</a>)</p> <p>. Only set if the target column is set. Mapping from other column spec id to its CorrelationStats with the target column. This field may be stale, see the statsUpdateTime field for for the timestamp at which these stats were last updated.</p> <p>An object containing a list of "key": value pairs. Example: { "name": "wrench", "mass": "1.3kg", "count": "3" }.</p>
<b>statsUpdateTime</b>	<p><b>string (<u>Timestamp</u> (<a href="https://developers.google.com/protocol-buffers/docs/reference/google.protobuf#google.protobuf.Timestamp">https://developers.google.com/protocol-buffers/docs/reference/google.protobuf#google.protobuf.Timestamp</a>) format)</b></p> <p>Output only. The most recent timestamp when targetColumnCorrelations field and all descendant ColumnSpec.data_stats and ColumnSpec.top_correlated_columns fields were last (re-)generated. Any changes that happened to the dataset afterwards are not reflected in these fields values. The regeneration happens in the background on a best effort basis.</p> <p>A timestamp in RFC3339 UTC "Zulu" format, accurate to nanoseconds. Example: "2014-10-02T15:01:23.045123456Z".</p>

Methods	
<p><b><u>create</u></b>            (https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets/create)</p>	Creates a dataset.
<p><b><u>delete</u></b>            (https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets/delete)</p>	Deletes a dataset and all of its contents.
<p><b><u>exportData</u></b>            (https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets/exportData)</p>	Exports dataset's data to the provided output location.
<p><b><u>get</u></b>            (https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets/get)</p>	Gets a dataset.
<p><b><u>getIamPolicy</u></b>            (https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets/getIamPolicy)</p>	Gets the access control policy for a resource.
<p><b><u>importData</u></b>            (https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets/importData)</p>	Imports data into a dataset.
<p><b><u>list</u></b>            (https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets/list)</p>	Lists datasets in a project.
<p><b><u>patch</u></b>            (https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets/patch)</p>	Updates a dataset.
<p><b><u>setIamPolicy</u></b>            (https://cloud.google.com/vision/automl/docs/reference/rest/v1beta1/projects.locations.datasets/setIamPolicy)</p>	Sets the access control policy on the specified resource.

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 October 9, 2019.