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

JSON representation

JSON representation

Fields

Fields	
name	<p>string</p> <p>Output only. The resource name of the dataset. Form: <code>projects/{project_id}/locations/{locationId}/datasets/{datasetId}</code></p>
displayName	<p>string</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 (<code>_</code>), and ASCII digits 0-9.</p>
description	<p>string</p> <p>User-provided description of the dataset. The description can be up to 25000 characters long.</p>
exampleCount	<p>integer</p> <p>Output only. The number of examples in the dataset.</p>
createTime	<p>string (Timestamp format) (https://developers.google.com/protocol-buffers/docs/reference/google.protobuf#google.protobuf.Timestamp)</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>
etag	<p>string</p> <p>Used to perform consistent read-modify-write updates. If not set, a blind "overwrite" update happens.</p>
<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:</p>	
translationDatasetMetadata	<p>object (TranslationDatasetMetadata) (/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TranslationDatasetMetadata)</p> <p>)</p> <p>Metadata for a dataset used for translation.</p>

Fields	
imageClassificationDatasetMetadata	<p>object (ImageClassificationDatasetMetadata (/automl/docs/reference/rest/v1beta1/projects.locations.datasets#ImageClassificationDatasetMetadata))</p> <p>Metadata for a dataset used for image classification.</p>
textClassificationDatasetMetadata	<p>object (TextClassificationDatasetMetadata (/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TextClassificationDatasetMetadata))</p> <p>Metadata for a dataset used for text classification.</p>
imageObjectDetectionDatasetMetadata	<p>object (ImageObjectDetectionDatasetMetadata (/automl/docs/reference/rest/v1beta1/projects.locations.datasets#ImageObjectDetectionDatasetMetadata))</p> <p>Metadata for a dataset used for image object detection.</p>
videoClassificationDatasetMetadata	<p>object (VideoClassificationDatasetMetadata (/automl/docs/reference/rest/v1beta1/projects.locations.datasets#VideoClassificationDatasetMetadata))</p> <p>Metadata for a dataset used for video classification.</p>
videoObjectTrackingDatasetMetadata	<p>object (VideoObjectTrackingDatasetMetadata (/automl/docs/reference/rest/v1beta1/projects.locations.datasets#VideoObjectTrackingDatasetMetadata))</p> <p>Metadata for a dataset used for video object tracking.</p>
textExtractionDatasetMetadata	<p>object (TextExtractionDatasetMetadata (/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TextExtractionDatasetMetadata))</p> <p>Metadata for a dataset used for text extraction.</p>

Fields

textSentimentDatasetMetadata	object (TextSentimentDatasetMetadata (/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TextSentimentDatasetMetadata)) <p>Metadata for a dataset used for text sentiment.</p>
tablesDatasetMetadata	object (TablesDatasetMetadata (/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TablesDatasetMetadata)) <p>Metadata for a dataset used for Tables.</p>

Dataset metadata that is specific to translation.

JSON representation

Fields

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

Dataset metadata that is specific to image classification.

JSON representation

Fields

classificationType	enum (<u>ClassificationType</u> (/automl/docs/reference/rest/v1beta1/ClassificationType)) Required. Type of the classification problem.
---------------------------	--

Dataset metadata for classification.

JSON representation

Fields

classificationType	enum (<u>ClassificationType</u> (/automl/docs/reference/rest/v1beta1/ClassificationType)) Required. Type of the classification problem.
---------------------------	--

Dataset metadata specific to image object detection.

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

Dataset metadata specific to video object tracking.

Dataset metadata that is specific to text extraction

Dataset metadata for text sentiment.

JSON representation

Fields

sentimentMax

integer

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).

Metadata for a dataset used for AutoML Tables.

JSON representation

Fields

primaryTableSpecId	string Output only. The tableSpecId of the primary table of this dataset.
targetColumnSpecId	string 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">• CATEGORY• FLOAT64 If the type is CATEGORY , only up to 100 unique values may exist in that column across all rows. NOTE: Updates of this field will instantly affect any other users concurrently working with the dataset.

Fields

weightColumnSpecId	<p>string</p> <p>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.</p>
mlUseColumnSpecId	<p>string</p> <p>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 TRAIN, VALIDATE, TEST among its values, or only have TEST, UNASSIGNED values. In the latter case the rows with UNASSIGNED 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 UNASSIGNED. NOTE: Updates of this field will instantly affect any other users concurrently working with the dataset.</p>
targetColumnCorrelations	<p>map (key: string, value: object (<u>CorrelationStats</u> (/automl/docs/reference/rest/v1beta1/projects.locations.datasets.tableSpecs.columnSpecs#ColumnSpec.CorrelationStats)))</p> <p>Output only. Correlations between</p> <p><u>TablesDatasetMetadata.target_column_spec_id</u> (/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TablesDatasetMetadata.FIELDS.target_column_spec_id) , and other columns of the</p> <p><u>TablesDatasetMetadataprimary_table</u> (/automl/docs/reference/rest/v1beta1/projects.locations.datasets#TablesDatasetMetadata.FIELDS.primary_table_spec_id) . 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>

Fields

statsUpdateTime	<p>string (Timestamp (https://developers.google.com/protocol-buffers/docs/reference/google.protobuf#google.protobuf.Timestamp) format)</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>
------------------------	--

Creates a dataset.

Deletes a dataset and all of its contents.

Exports dataset's data to the provided output location.

Gets a dataset.

Gets the access control policy for a resource.

Imports data into a dataset.

	Lists datasets in a project.
	Updates a dataset.
	Sets the access control policy on the specified resource.