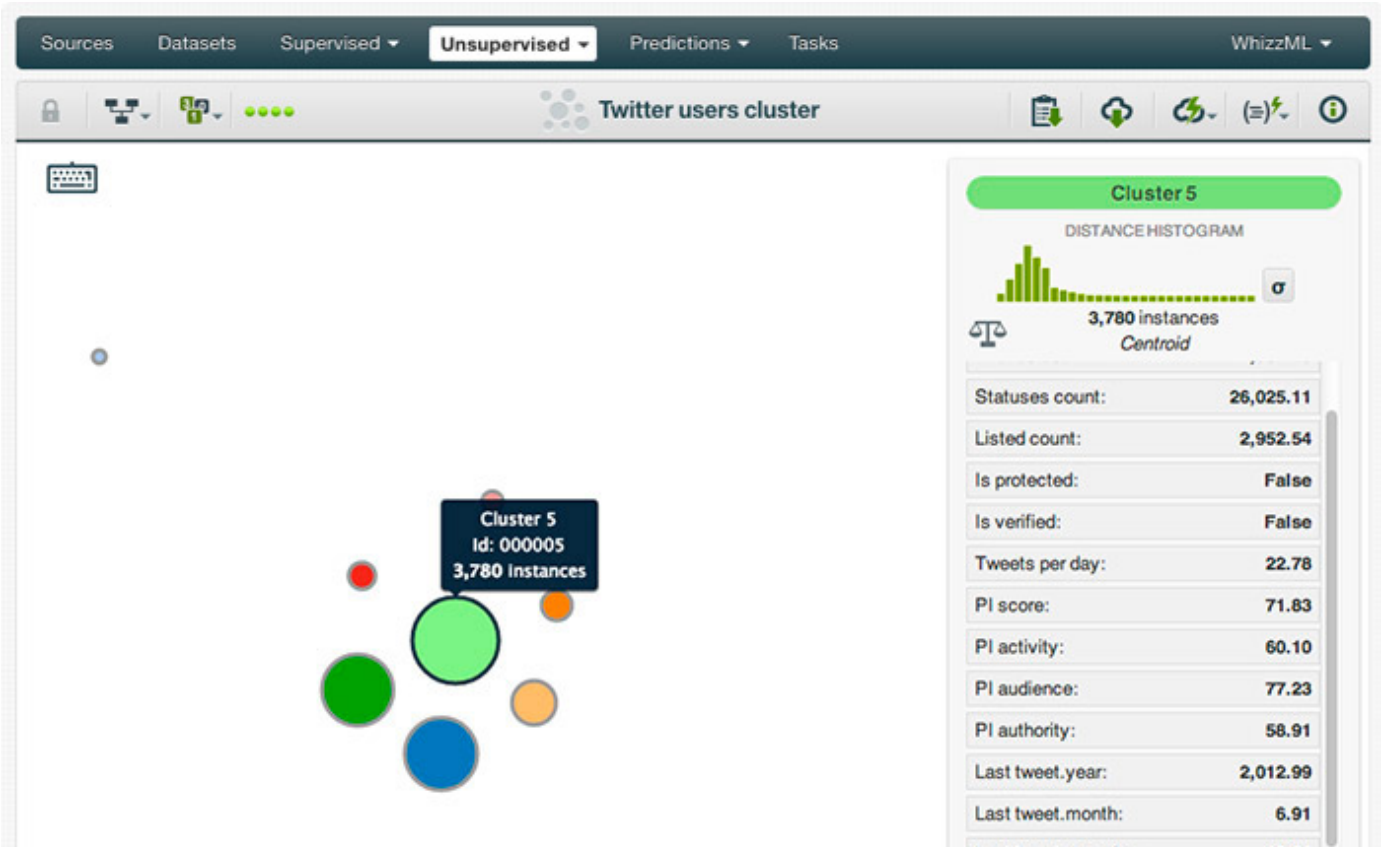
Clusters

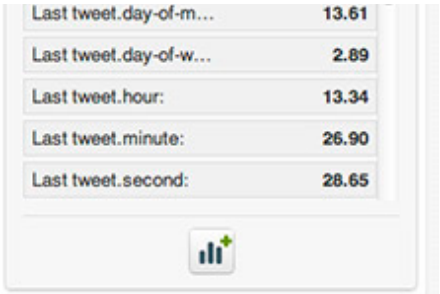
A **cluster** is a set of groups (i.e., clusters) of instances of a **dataset** that have been automatically classified together according to a distance measure computed using the fields of the dataset. Clusters can handle numeric, categorical, text and items fields as inputs:

- Numeric fields: the Eucledian distance is computed between the instances numeric values.
- Categorical fields: a common way to handle categorical data is to take each category as a new field and assign 0 or 1 depending on the category. So a field with 20 categories will become 20 separate binary fields. BigML uses a technique called k-prototypes which modifies the distance function to operate as though the categories were transformed to binary values.
- Text and item fields: each instance is assigned a vector of terms and then cosine similarity is computed to determine closeness between instances.

Each cluster group is represented by a **centroid** or center that is computed using the mean for each numeric field and the mode for each categorical field. For text and items fields each centroid contains the terms or items which minimize the average cosine distance between the centroid and the points in its neighborhood.

To create a **cluster**, you can select an arbitrary number of clusters (i.e., **k**) and also select an arbitrary subset of fields from your **dataset** as **input_fields**. You can use scales to select how each field influences the distance measure used to group instances together.





BigML.io allows you to create, retrieve, update, and delete your **cluster**. You can also list all of your **clusters**.

Jump to:

- Cluster Base URL
- Creating a Cluster
- Cluster Arguments
- Retrieving a Cluster
- Cluster Properties
- PMML
- Filtering and Paginating Fields from a Cluster
- Create a Dataset Using a Cluster and a Centroid
- Create a Model Using a Cluster and a Centroid
- Updating a Cluster
- Deleting a Cluster
- Listing Clusters

Cluster Base URL

You can use the following base URL to create, retrieve, update, and delete **clusters**.

```bash
https://bigml.io/andromeda/cluster
```

All requests to manage your **clusters** must use HTTPS and be authenticated using your **username** and **API key** to verify your identity. See this section for more details.

Creating a Cluster

To create a new **cluster**, you need to POST to the **cluster** base URL an object containing at least the **dataset/id** that you want to use to create the **cluster**. The **content-type** must always be **"application/json"**.

```
POST /cluster?$BIGML_AUTH HTTP/1.1
Host: bigml.io
Content-Type: application/json
```

You can easily create a new **cluster** using **curl** as follows. All you need is a valid **dataset/id** and your authentication variable set up as shown above.

curl

```
curl "https://bigml.io/andromeda/cluster?$BIGML_AUTH" \
    -X POST \
    -H 'content-type: application/json' \
    -d '{"dataset": "dataset/603e20a91f386f43db000004"}'
```

BigML.io will return a newly created **cluster** document, if the request succeeded.

json


```
{
    "balance_fields": true,
    "category": 0,
    "cluster_datasets": {},
    "cluster_models": {},
    "cluster_seed": "2c249dda00fbf54ab4cdd850532a584f286af5b6",
    "clusters": {},
    "code": 201,
    "columns": 5,
    "configuration": null,
    "configuration_status": false,
    "created": "2021-03-15T08:48:56.041192",
    "creator": "alfred",
    "critical_value": null,
```


Cluster Arguments

In addition to the **dataset**, you can also POST the following arguments.

Argument	Type	Description
all_fields_preferred optional	Boolean, default is **false**	Setting this parameter to **true** will set the preferred flag of all fields at once, instead of doing it one by one in a **fields** map. **Example**: true
balance_fields optional	Boolean, default is **true**	When this parameter is enabled, all the numeric fields will be scaled so that their standard deviations are 1. This makes each field have roughly equivalent influence. **Example**: true
category optional	Integer, default is **the category of the dataset**	The category that best describes the **cluster**. See the category codes for the complete list of categories. **Example**: 1
cluster_seed	String	A string to generate deterministic clusters. **Example**: "My Seed"
critical_value optional	Integer, default is **5**	The clustering algorithm G-means is parameter free except for one, the **critical_value** parameter. G-means iteratively takes existing clusters and tests whether the cluster's neighborhood appears Gaussian. If it doesn't the cluster is split into two. The **critical_value** sets how strict the test is when deciding whether data looks Gaussian. The default is to 5, which seems to work well in most cases. A range of 1 - 10 is acceptable. A **critical_value** of 1 means data must look very Gaussian to pass the test, and can lead to more clusters being detected. Higher **critical_value** will tend to find fewer clusters. **Example**: 3
dataset	String	A valid **dataset/id**. **Example**: dataset/603e20a91f386f43db000004
		A list of dataset ids or objects to be used to build the new **cluster**. See the Section on Multi-Datasets and Section on Resources Accepting Multi-Datasets Input for more details. **Example**:

Argument	Type	Description

datasets
optional
	Array	[{
		"id": "dataset/603e20a91f386f43db000004",
		"sample_rate": 0.5,
		"out_of_bag": true
		},
		{
		"id": "dataset/603e20a91f386f43db000005",
		"sample_rate": 0.8,
		"replacement": true
		}]

deep optional	Boolean, default is **false**	Clone the dataset used to build the **cluster** while cloning the original . Must be used along with the **origin** or **shared_hash** option. **Example**: true
default_numeric_value optional	String	It accepts any of the following strings to substitute missing numeric values across all the numeric fields in the dataset: **mean**, **median**, **minimum**, **maximum**, **zero**. **Example**: "median"
description optional	String	A description of the **cluster** up to 8192 characters long. **Example**: "This is a description of my new cluster"
excluded_fields optional	Array, default is **[], an empty list. None of the fields in the dataset is excluded**.	Specifies the fields that won't be included in the **cluster**. **Example**: ["000000", "000002"]
field_scales optional	Object, default is **{}, an empty dictionary. That is, no special**	With this argument you can pick your own scaling for each field. If a field isn't included in **field_scales**, BigML will treat the scale as 1 (no scale change). If both **balance_fields** and **field_scales** are present, then **balance_fields** will be applied first. This will make it easy for you do things like balancing age and salary, but then request that age be twice as important. **Example**:

Argument	Type	Description

	scaling is used.	{
		"000001": 4,
		"000003": 2
		}

fields
optional
	Object, default is **{}, an empty dictionary. That is, no names or preferred statuses are changed**.	This can be used to change the names of the fields in the cluster with respect to the original names in the dataset or to tell **BigML** that certain fields should be preferred. An entry keyed with the field id generated in the **source** for each field that you want the name updated.
		Example:
		{
		"000001": {"name": "length_1"},
		"000003": {"name": "length_2"},
		}

fields_maps
optional
	Object	A dictionary keyed by **dataset/id** and **object** values. Each entry maps fields in the first dataset to fields in the dataset referenced by the key.
		Example:
		{
		"dataset/603e20a91f386f43db000004": {
		"000000":"000023",
		"000001":"000024",
		"000002":"00003a"},
		"dataset/603e20a91f386f43db000005": {
		"000000":"000023",
		"000001":"000004",
		"000002":"00000f"
		}

include_extracted_features optional	Boolean or Array of IDs	Extracted image features to use as model inputs. Available options are: • **true**: include all extracted features, unless explicitly excluded (this is the default for all non-deepnet models.) • **false**: don't include any extracted features (this is the default for deepnet models, if not given.) • **list of ids**: a explicit list of field ids, corresponding to extracted fields to add to the default set

Argument	Type	Description
		Example: true
input_fields optional	Array, default is **[]. All the fields in the dataset**.	Specifies the fields to be considered to create the **cluster**. **Example**: ["000001", "000003"]
k optional	Integer, default is **null to use g-means cluster**	The number of clusters. Must be null or a number greater than or equal to 1 and less than or equal to 300. **Example**: 3
model_clusters optional	Boolean default is **false**	Whether a model for every cluster will be generated or not. Each model predicts whether or not an instance is part of its respective cluster. **Example**: true
name optional	String default is **dataset's name**	The name you want to give to the new **cluster**. **Example**: "my new cluster"
origin optional	String	The **cluster/id** of the cluster to be cloned. The price of the cluster must be 0 to be cloned via API. Set **deep** to **true** to clone the dataset used to build the cluster too. Note that the dataset can be cloned only if it is already in the public gallery and free. If multiple datasets have been used to create the cluster, only the first dataset will be cloned. **Example**: "cluster/604f1f781f386f4240000000"
out_of_bag optional	Boolean, default is **false**	Setting this parameter to **true** will return a sequence of the out-of-bag instances instead of the sampled instances. See the Section on Sampling for more details. **Example**: true
project optional	String	The **project/id** you want the **cluster** to belong to. **Example**: "project/603de73d1f386f7360000000"

Argument	Type	Description
range optional	Array, default is **[1, max rows in the dataset]**	The range of successive instances to build the **cluster**. **Example**: [1, 150]
replacement optional	Boolean, default is **false**	Whether sampling should be performed with or without replacement. See the Section on Sampling for more details. **Example**: true
sample_rate optional	Float, default is **1.0**	A real number between 0 and 1 specifying the sample rate. See the Section on Sampling for more details. **Example**: 0.5
seed optional	String	A string to be hashed to generate deterministic samples. See the Section on Sampling for more details. **Example**: "MySample"
shared_hash	String	The **shared hash** of the shared **cluster** to be cloned. Set **deep** to **true** to clone the dataset used to build the cluster too. Note that the dataset can be cloned only if it is already shared and set clonable. If multiple datasets have been used to create the cluster, only the first dataset will be cloned. **Example**: "kpY46mNuNVReITw0Z1mAqoQ9ySW"
summary_fields optional	Array, default is **[]**	Specifies the ids for fields which will be included when generating the per cluster summaries/datasets, but will not be used for clustering. The **summary_fields** must be a strict subset of the input_fields, where the latter is adjusted before passing it to the model creation algorithm by setting it to all non-preferred fields if not provided explicitly, adding to it explicit **summary_fields**, and subtracting explicit **excluded_fields**. You can use either field identifiers or field names. **Example**: ["000004"]

Argument	Type	Description
tags optional	Array of Strings	A list of strings that help classify and index your **cluster**. **Example**: ["best customers", "2021"]

webhook
optional
	Object	A webhook url and an optional secret phrase. See the Section on Webhooks for more details.
		Example:
		{
		"url": "http://myhost/path/to/webhook",
		"secret": "mysecret"
		}

weight_field optional	String	Any numeric field with no negative or missing values is valid as a weight field. Each instance will be weighted individually according to the weight field's value. For more information, see the Weight Field section. **Example**: "000005"

You can also use **curl** to customize a new **cluster**. For example, to create a new **cluster** named "my cluster", with only certain rows, and with only three fields:

curl

```
curl "https://bigml.io/andromeda/cluster?$BIGML_AUTH" \
     -X POST \
     -H 'content-type: application/json' \
     -d '{"dataset": "dataset/603e20a91f386f43db000004",
          "input_fields": ["000001", "000003"],
          "name": "my cluster",
          "range": [25, 125]
         }'
```

If you do not specify a name, **BigML.io** will assign to the new **cluster** the **dataset**'s name. If you do not specify a **range** of instances, **BigML.io** will use all the instances in the **dataset**. If you do not specify any input **fields**, **BigML.io** will include all the input fields in the **dataset**.

Retrieving a Cluster

Each **cluster** has a unique identifier in the form **"cluster/id"** where id is a string of 24 alpha-numeric

characters that you can use to retrieve the **cluster**.

To retrieve a **cluster** with **curl**:

```curl
curl "https://bigml.io/andromeda/cluster/604f1f781f386f4240000000?$BIGML_AUTH"
```

You can also use your browser to visualize the **cluster** using the full **BigML.io** URL or pasting the **cluster/id** into the BigML labs dashboard.

Cluster Properties

Once a **cluster** has been successfully created it will have the following properties.

Property	Type	Description
balance_fields filterable, sortable	Boolean	Whether all the numeric fields have been scaled so that their standard deviations are 1.
category filterable, sortable, updatable	Integer	One of the categories in the table of categories that help classify this resource according to the domain of application.
cluster_datasets	Object	A dictionary that maps cluster ids to dataset resources offering per field distribution summaries for each cluster. Each dataset resource can be serialized on-demand using the neighborhood of the cluster.
cluster_seed	String	With no seed, the cluster locations can vary from run to run. With a seed, the clusters are deterministic.
		All the information that you need to recreate or use the cluster on your own. It includes: • **clusters**: a list of centroids with a cluster object for each centroid. • **fields**: a dictionary with an entry per field in the dataset used to build the cluster. Fields are paginated according to the **field_meta** attribute. Each entry includes the column number in original

Property	Type	Description
clusters	Object	source, the name of the field, the type of the field, and the **summary**. See this Section for more details. • **global**: the global **center** of the full data (equivalent to the cluster for a **k** of 1). It also contains the **distance** distribution of the entire dataset to that center. • **total_ss**: the sum of squares for the global cluster's distances. • **within_ss**: the total within-cluster sum of squares for each cluster. • **between_ss**: total_ss - within_ss • **ratio_ss**: between_ss / total_ss. A large **ratio_ss** (close to 1) indicates a better clustering, while a smaller value (closer to 0) indicates a weaker cluster. Generally the more clusters, the larger ratio_ss will be, but with diminishing returns.
code	Integer	HTTP status code. This will be 201 upon successful creation of the **cluster** and 200 afterwards. Make sure that you check the code that comes with the status attribute to make sure that the **cluster** creation has been completed without errors.
columns filterable, sortable	Integer	The number of fields in the **cluster**.
composites filterable, sortable	Array of Strings	The list of composite ids that reference this **cluster**.
created filterable, sortable	ISO-8601 Datetime	This is the date and time in which the **cluster** was created with microsecond precision. It follows this pattern yyyy-MM-ddThh:mm:ss.SSSSSS. All times are provided in Coordinated Universal Time (UTC).
creator	String	The user that created the **cluster**.
dataset filterable, sortable	String	The **dataset/id** that was used to build the **cluster**.

Property	Type	Description
dataset_field_types	Object	A dictionary that informs about the number of fields of each type in the dataset used to create the cluster. It has an entry per each field type (**categorical**, **datetime**, **numeric**, **text**, **image**, **path** and **regions**), an entry for **preferred** fields, and an entry for the **total** number of fields.
dataset_status filterable, sortable	Boolean	Whether the **dataset** is still available or has been deleted.
datasets	Array	A list of dataset ids or objects used to build the **cluster**.
default_numeric_value	String	Any of the following strings to substitute missing numeric values across all the numeric fields in the dataset: **mean**, **median**, **minimum**, **maximum**, **zero**.
description updatable	String	A text describing the **cluster**. It can contain restricted markdown to decorate the text.
excluded_fields	Array	The list of **fields**'s ids that were excluded to build the **cluster**.
execution_id filterable, sortable	String	The **execution/id** that created the **cluster**.
execution_status filterable, sortable	Boolean	Whether the **execution** is still available or has been deleted.
field_scales	Object	The specific scales used for each field, if any.
fields_meta	Object	A dictionary with meta information about the fields dictionary. It specifies the **total** number of fields, the current **offset**, and **limit**, and the number of fields (**count**) returned.
input_fields	Array	The list of **input fields**' ids used the models of the **cluster**.
k filterable, sortable	Integer	The number of clusters.

Property	Type	Description
locale	String	The **dataset**'s locale.
max_columns filterable, sortable	Integer	The total number of fields in the **dataset** used to build the **cluster**.
max_rows filterable, sortable	Integer	The maximum number of instances in the **dataset** that can be used to build the **cluster**.
model_clusters filterable, sortable	Boolean	Whether a model for each cluster was created or not.
name filterable, sortable, updatable	String	The name of the **cluster** as your provided or based on the name of the **dataset** by default.
name_options filterable, sortable	String	Information about the **cluster**.
number_of_batchcentroids filterable, sortable	Integer	The current number of **batch centroids** that use this **cluster**.
number_of_centroids filterable, sortable	Integer	The current number of **centroids** that use this **cluster**.
number_of_public_centroids filterable, sortable	Integer	The current number of **public centroids** that use this **cluster**.
ordering filterable, sortable	Integer	The order used to chose instances from the dataset to build **cluster**. There are three different types: • **0** Deterministic • **1** Linear • **2** Random
origin filterable, sortable	String	The **cluster/id** of the original **cluster**.

Property	Type	Description
out_of_bag filterable, sortable	Boolean	Whether the out-of-bag instances were used to create the **cluster** instead of the sampled instances.
price filterable, sortable, updatable	Float	The price other users must pay to clone your **cluster**.
private filterable, sortable, updatable	Boolean	Whether the **cluster** is public or not.
project filterable, sortable, updatable	String	The **project/id** the resource belongs to.
range	Array	The **range** of instances used to build the **cluster**.
replacement filterable, sortable	Boolean	Whether the instances sampled to build the **cluster** were selected using replacement or not.
resource	String	The **cluster/id**.
rows filterable, sortable	Integer	The total number of instances used to build the **cluster**.
sample_rate filterable, sortable	Float	The sample rate used to select instances from the **dataset** to build the **cluster**.
scales	Object	A dictionary that represents the combination of user requested **field_scales** and **balance_fields**.
seed filterable, sortable	String	The string that was used to generate the sample.
shared filterable, sortable, updatable	Boolean	Whether the **cluster** is shared using a private link or not.
shared_clonable filterable, sortable, updatable	Boolean	Whether the shared **cluster** can be cloned or not.
shared_hash	String	The hash that gives access to this **cluster** if it has been shared using a private link.

Property	Type	Description
sharing_key	String	The alternative key that gives read access to this **cluster**.
size filterable, sortable	Integer	The number of bytes of the **dataset** that were used to create this **cluster**.
source filterable, sortable	String	The **source/id** that was used to build the **dataset**.
source_status filterable, sortable	Boolean	Whether the **source** is still available or has been deleted.
status	Object	A description of the status of the **cluster**. It includes a code, a message, and some extra information. See the table below.
subscription filterable, sortable	Boolean	Whether the **cluster** was created using a subscription plan or not.
summary_fields	Array	The list of field's ids that are included when generating the cluster's summaries but were not used for clustering.
tags filterable, updatable	Array of Strings	A list of user tags that can help classify and index this resource.
updated filterable, sortable	ISO-8601 Datetime	This is the date and time in which the **cluster** was updated with microsecond precision. It follows this pattern yyyy-MM-ddThh:mm:ss.SSSSSS. All times are provided in Coordinated Universal Time (UTC).
webhook	Object	A webhook url and an optional secret phrase. See the Section on Webhooks for more details.
weight_field optional	String	Any numeric field with no negative or missing values is valid as a weight field. Each instance will be weighted individually according to the weight field's value. For more information, see the Weight Field section. **Example**: "000001"

Property	Type	Description
white_box filterable, sortable	Boolean	Whether the **cluster** is publicly shared as a white-box.

Cluster Object

Property	Type	Description
center	Object	A dictionary with the mean for numeric fields, the mode for categorical fields and the terms and items that minimize the average cosine distance for text and items fields.
count	Integer	The count gives the size of that neighborhood.
distance	Object	A dictionary that gives a numeric summary capturing the distribution of distances from the cluster's center to each of the points that fall into its neighborhood.
id	String	The id of the cluster.
name	String	The name of the cluster.

Cluster Status

Creating a **cluster** is a process that can take just a few seconds or a few days depending on the size of the **dataset** used as input and on the workload of **BigML**'s systems. The **cluster** goes through a number of states until its fully completed. Through the status field in the **cluster** you can determine when the cluster has been fully processed and ready to be used to create predictions. These are the properties that **cluster**'s **status** has:

Property	Type	Description
code	Integer	A status code that reflects the status of the **resource** creation. It can be any of those that are explained here.
elapsed	Integer	Number of milliseconds that **BigML.io** took to process the **resource**.
message	String	A human readable message explaining the status.

Property	Type	Description
progress	Float, between 0 and 1	How far **BigML.io** has progressed building the **resource**.

Once **cluster** has been successfully created, it will look like:

```json
{
    "balance_fields": true,
    "category": 0,
    "cluster_datasets": {
        "000000": "",
        "000001": ""
    },
    "cluster_models": {},
    "cluster_seed": "2c249dda00fbf54ab4cdd850532a584f286af5b6",
    "clusters": {
        "between_ss": 10.44563,
        "clusters": [
            {
                "center": {
```

PMML

The default cluster output format is JSON. However, the **pmml** parameter allows to include a PMML version of the cluster. The cluster will include a XML document that fullfils PMML v4.1. For example:

```curl
curl "https://bigml.io/andromeda/cluster/604f1f781f386f4240000000?$BIGML_AUTH&pmml=yes"
```

will include the PMML version of the cluster together with the JSON representation. While:

```curl
curl "https://bigml.io/andromeda/cluster/604f1f781f386f4240000000?$BIGML_AUTH&pmml=only"
```

will include the PMML version of the cluster but not all the usual JSON fields. Some fields will be incomplete or not even be returned.

Filtering and Paginating Fields from a Cluster

A **cluster** might be composed of hundreds or even thousands of fields. Thus when retrieving a **cluster**, it's possible to specify that only a subset of fields be retrieved, by using any combination of the following parameters in the query string (unrecognized parameters are ignored):

Parameter	Type	Description
fields optional	Comma-separated list	A comma-separated list of **field** IDs to retrieve. **Example**: "fields=000000,000002"
full optional	Boolean	If false, no information about fields is returned. **Example**: "full=false"
iprefix optional	String	A case-insensitive string to retrieve fields whose name start with the given prefix; It is possible to specify more than one iprefix by repeating the parameter, in which case the union of the results is returned. **Example**: "iprefix=INCOME"
limit optional	Integer	Maximum number of **fields** that you will get in the **fields** field. **Example**: "limit=100"
offset optional	Integer	How far off from the first **field** in your **dataset** is the first **field** in the **fields** field. **Example**: "offset=100"
order_by optional	String	Sorting criteria; possible values are **"count"**, **"max"**, **"min"**, **"name"**, and **"type"**, and their negated values (**"-count"**, **"-name"**, etc.) to specify a descending order. **Example**: "order_by=name"
prefix optional	String	A case-sensitive string to retrieve fields whose name start with the given prefix; It is possible to specify more than one prefix by repeating the parameter, in which case the union of the results is returned. **Example**: "prefix=income"

Since **fields** is a map and therefore not ordered, the returned fields contain an additional key, **order**, whose integer (increasing) value gives you their ordering. In all other respects, the source is the same as the one you would get without any filtering parameter above.

The **fields_meta** field can help you paginate fields. Its structure is as follows:

Property	Type	Description
count optional	Integer	Specifies the current number of fields in the resource.
limit optional	Integer	The maximum number of fields that will be returned in the resource.
offset optional	Integer	The current offset in the pagination of fields.
total optional	Integer	The total number of fields in the resource.

Note that paginating fields might only be worth if you are going to deal with really wide (i.e., more than 200 fields).

Create a Dataset Using a Cluster and a Centroid

Each centroid has associated a pre-computed dataset that has been created using all the instances in the neighborhood. You can create a new dataset using the corresponding **cluster/id** and **centroid** id as follows:

```curl
curl "https://bigml.io/andromeda/dataset?$BIGML_AUTH" \
     -X POST \
     -H 'content-type: application/json' \
     -d '{"cluster": "cluster/604f1f781f386f4240000000",
         "centroid": "000000"}'
```

Create a Model Using a Cluster and a Centroid

If you created a cluster setting the **model_clusters** option to **true**, then each centroid has associated a pre-computed model that has been created using all the instances of the dataset. Each model separates between those instances that belong to the centroid neighborhood and those that belong to other neighborhoods. You can create a new model using the corresponding **cluster/id** and **centroid** id as follows:

```curl
curl "https://bigml.io/andromeda/dataset?$BIGML_AUTH" \
     -X POST \
     -H 'content-type: application/json' \
     -d '{"cluster": "cluster/604f1f781f386f4240000000",
         "centroid": "000000"}'
```

Updating a Cluster

To update a **cluster**, you need to PUT an object containing the fields that you want to update to the **cluster**'s base URL. The content-type must always be: **"application/json"**. If the request succeeds, **BigML.io** will return with an HTTP 202 response with the updated **cluster**.

For example, to update **cluster** with a new name you can use **curl** like this:

```curl
curl "https://bigml.io/andromeda/cluster/604f1f781f386f4240000000?$BIGML_AUTH" \
     -X PUT \
     -H 'content-type: application/json' \
     -d '{"name": "a new name"}'
```

If you want to update a **cluster** with a new label and description for a specific field you can use **curl** like this:

```curl
curl "https://bigml.io/andromeda/cluster/604f1f781f386f4240000000?$BIGML_AUTH" \
     -X PUT \
     -H 'content-type: application/json' \
     -d '{"fields": {
            "000000": {
                "label": "a longer name",
                "description": "an even longer description"
            }
          }}'
```

See this section for more details.

Deleting a Cluster

To delete a **cluster**, you need to issue a HTTP DELETE request to the **cluster/id** to be deleted.

Using **curl** you can do something like this to delete a **cluster**:

```curl
curl -X DELETE "https://bigml.io/andromeda/cluster/604f1f781f386f4240000000?$BIGML_AUTH"
```

If the request succeeds you will not see anything on the command line unless you executed the command in verbose mode. Successful DELETEs will return **"204 no content"** responses with no body.

Once you delete a **cluster**, it is permanently deleted. That is, a delete request cannot be undone. If you try to delete a **cluster** a second time, or a **cluster** that does not exist, you will receive a **"404 not found"** response.

However, if you try to delete a **cluster** that is being used at the moment, then **BigML.io** will not accept the request and will respond with a **"400 bad request"** response.

See this section for more details.

Listing Clusters

To list all the **clusters**, you can use the **cluster** base URL. By default, only the 20 most recent **clusters** will be returned. You can see below how to change this number using the **limit** parameter.

You can get your list of **clusters** using **curl**.

```curl
curl "https://bigml.io/andromeda/cluster?$BIGML_AUTH"
```

See this section for more details. You can also paginate, filter, and order your **clusters**.