

Segments

Date published: 2020-10-30

Date modified: 2024-05-30

CLOUDERA

Legal Notice

© Cloudera Inc. 2024. All rights reserved.

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Unless otherwise noted, scripts and sample code are licensed under the Apache License, Version 2.0.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 (“ASLv2”), the Affero General Public License version 3 (AGPLv3), or other license terms. Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks in the United States and other countries. All other trademarks are the property of their respective owners.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

Contents

Creating segments.....	4
Cloning segments.....	8
Editing segments.....	13
Creating entity segments.....	16
Deleting segments.....	21
Creating segments from filter definitions.....	23
Using segments in visuals.....	34

Creating segments

CDP Data Visualization lets you easily create a new segment definition.

Procedure

1. On the main navigation bar, click DATA.

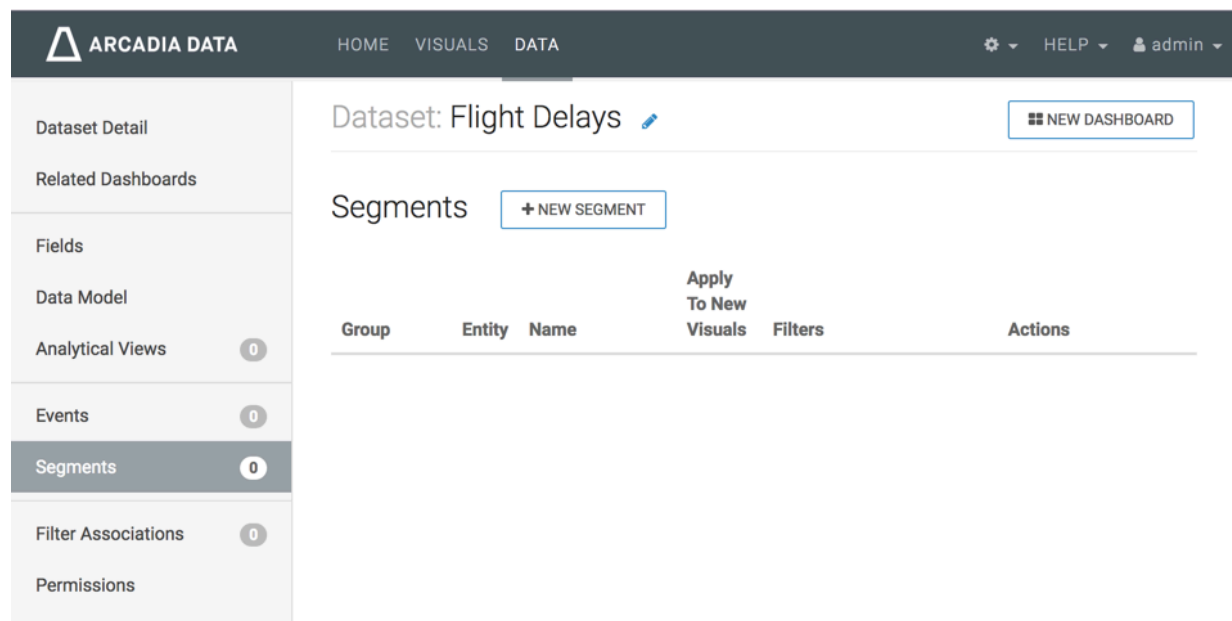
The Data view appears.

2. Open the Datasets tab.
3. Find the dataset in the list of datasets, either by scrolling or by using search, and click it.

We are using the dataset Flight Delays based on data previously imported into Data Visualization from a data file . We are also using additional data files.

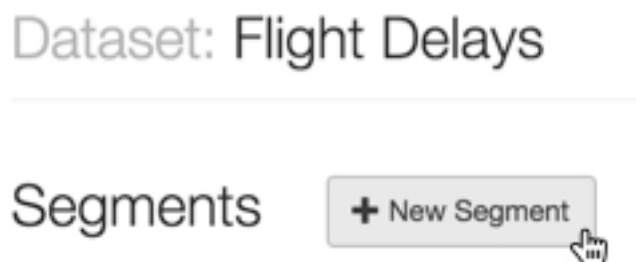
Dataset side navigation appears, open at Dataset Detail view.

4. In the side navigation menu, click Segments.



The Segments view appears.

5. In the Segments interface, click New Segment.



The Add Segment modal window appears.

6. Make the following entries:

- Under Segment Name, enter Delayed Flights.
- Under Segment Group, enter Delay Status.
- Under Filters, click the text box to open the Filter Expression modal window.

Add Segment ×

Segment Name

Enter segment name


Segment Group

Enter a group that links segments of the same type together

☐ **Apply as a filter to new visuals**

Filters

You can enter multiple filters and they will automatically have an "AND" logic between them

Click to update in SQL expression editor 

ADD FILTER

Entity to extract (only necessary for behavior based segments)

Click to update in SQL expression editor

ADD ENTITY

CANCEL

SAVE

7. In the Filter Expression modal window, enter the following expression to show flights that are delayed:

```
[airtime] IS NOT NULL AND [depdelay] > 0
```

- Click Validate Expression to verify correctness.

Filter Expression

Enter an expression that can serve as a WHERE clause. For example: event_type="login"

[airtime] IS NOT NULL AND [depdelay] > 0

VALIDATE EXPRESSION

☒ Autocomplete on

All Functions

abs
acos
add_months
adddate
AND
appx_median
ascii
asin
atan

All Fields & Tables

actualelapse...
airlineid
airtime
T/F arrdel15
arrdelay
arrdelayminut...
arrivaldelaygr...
arrtime
arrtimeblk

CANCEL

APPLY

- Click Apply to save the filter and return to the Add Segment modal window.

The new filter appears under Filters.

10. Click Save.

Add Segment

Segment Name

Delayed Flights

Segment Group

Delay Status

☐ Apply as a filter to new visuals

Filters

You can enter multiple filters and they will automatically have an "AND" logic between them

[airtime] IS NOT NULL AND [depdelay] > 0

ADD FILTER

Entity to extract (only necessary for behavior based segments)

Click to update in SQL expression editor

ADD ENTITY

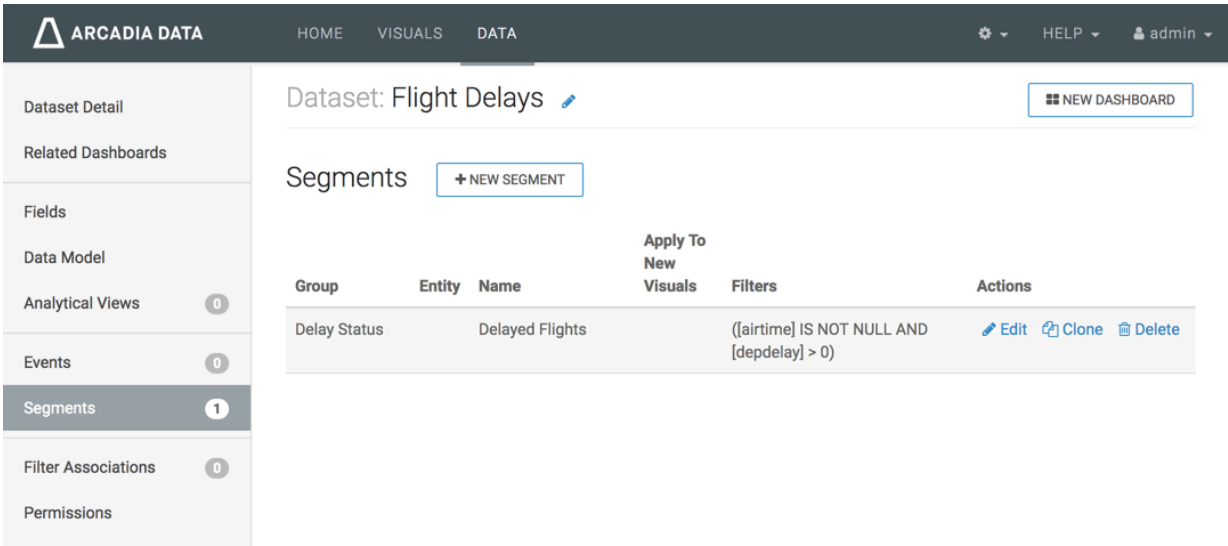
CANCEL

SAVE

The new segment definition appears in the Segments interface.



Note: The actions available for this segment: Edit, Clone, and Delete.



To view how segments work in a visual, see [Using Segments in Visuals](#).

Related Information
[Using segments in visuals](#)

Cloning segments

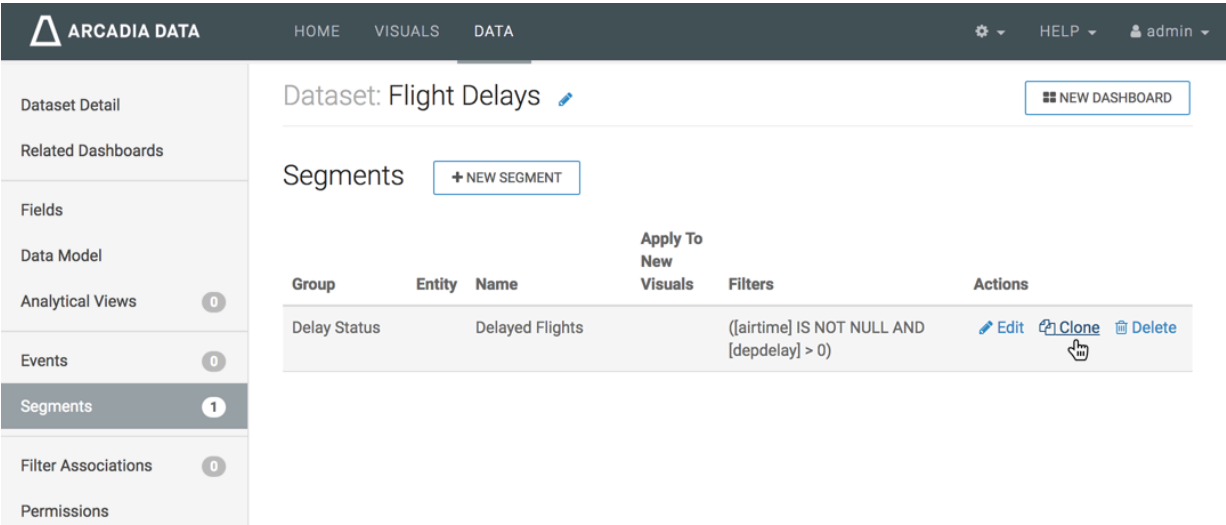
CDP Data Visualization lets you clone an existing segment definition.

Procedure

1. On the main navigation bar, click DATA.
2. Select Flight Delays dataset.
Dataset side navigation appears, open at Dataset Detail view.
3. In the side navigation menu, click Segments.

- 4. In the Segments list, find the segment to clone, and click Clone.

In this example, we are cloning Delayed Flights to create a new segment On-Time Departure.



A pre-populated Add Segment modal window appears named Copy of Delayed Flights.

5. Edit the Add Segment modal window by changing its entries.
 - Change Segment Name, to On-Time Departure.
 - Leave Segment Group as is.
 - Under Filters, click the text box to open the Filter Expression modal window.

Add Segment ✕

Segment Name

On-Time Departure

Segment Group

Delay Status

☒ **Apply as a filter to new visuals**

Filters

You can enter multiple filters and they will automatically have an "AND" logic between them

[airtime] IS NOT NULL AND [depdelay] > 0 ✕

Click to update in SQL expression editor ✕

ADD FILTER

Entity to extract (only necessary for behavior based segments)

Click to update in SQL expression editor

ADD ENTITY

CANCEL

SAVE

6. Replace the filter expression. In our example, we are adding the following expression:

```
[airtime] IS NOT NULL AND [depdelay] < = 0
```

7. Click Validate Expression to verify correctness.

Filter Expression

×

Enter an expression that can serve as a WHERE clause. For example: event_type="login"

[airtime] IS NOT NULL AND [depdelay] < = 0

VALIDATE EXPRESSION

☒ Autocomplete on

All Functions

abs
acos
add_months
adddate
AND
appx_median
ascii
asin
atan

All Fields & Tables

actualelapse...
airlineid
airtime
TIF arrdel15
arrdelay
arrdelayminut...
arrivaldelaygr...
arrtime
arrtimeblk

CANCEL

APPLY

8. Click Apply to save the filter and return to the Add Segment modal window.
9. Under Filters, notice the new filter expression.

10. Click Save.

Add Segment

×

Segment Name

On-Time Departure

Segment Group

Delay Status

☒ Apply as a filter to new visuals

Filters

You can enter multiple filters and they will automatically have an "AND" logic between them

[airtime] IS NOT NULL AND [depdelay] < = 0

Click to update in SQL expression editor

ADD FILTER

Entity to extract (only necessary for behavior based segments)

Click to update in SQL expression editor

ADD ENTITY

CANCEL

SAVE

11. The new segment definition appears in the Segments interface.

Note the similarity in the definition of the two filters, Delayed Flights and On-Time Departure.

The screenshot shows the Arcadia Data interface. The top navigation bar includes 'HOME', 'VISUALS', and 'DATA'. The left sidebar contains a navigation menu with items: Dataset Detail, Related Dashboards, Fields, Data Model, Analytical Views (0), Events (0), Segments (2), Filter Associations (0), and Permissions. The main content area is titled 'Dataset: Flight Delays' and features a '+ NEW SEGMENT' button. Below this is a table of segments.

Group	Entity	Name	Apply To New Visuals	Filters	Actions
Delay Status		Delayed Flights		([airtime] IS NOT NULL AND [depdelay] > 0)	Edit Clone Delete
Delay Status		On-Time Departure		([airtime] IS NOT NULL AND [depdelay] <= 0)	Edit Clone Delete

Editing segments

CDP Data Visualization lets you easily edit an existing segment definition.

Procedure

1. On the main navigation bar, click DATA.
2. Select Flight Delays dataset.
Dataset side navigation appears, open at Dataset Detail view.
3. In the side navigation menu, click Segments.

- 4. In the Segments list, find the segment to edit, and click Edit.

In our example, let's edit the segment Delayed Flights to change the segment name from Delayed Flights to Delayed Departure.

ARCADIA DATA

HOMEVISUALSDATA

HELPadmin

Dataset Detail

Related Dashboards

Fields

Data Model

Analytical Views0

Events0

Segments2

Filter Associations0

Permissions

Dataset: Flight Delays

NEW DASHBOARD

SegmentsNEW SEGMENT


Group	Entity	Name	Apply To New Visuals	Filters	Actions
Delay Status		Delayed Flights		([airtime] IS NOT NULL AND [depdelay] > 0)	<a>Edit <a>Clone <a>Delete
Delay Status		On-Time Departure		([airtime] IS NOT NULL AND [depdelay] <= 0)	<a>Edit <a>Clone <a>Delete

A pre-populated Edit Segment modal window appears.

5. Change the entries in Edit Segment modal window.
 - Change Segment Name from Delayed Flights to Delayed Departure.
 - Leave all other fields 'as is'.

Add Segment ✕

Segment Name

Delayed Departure 


Segment Group


Delay Status

☒ **Apply as a filter to new visuals**

Filters

You can enter multiple filters and they will automatically have an "AND" logic between them

[airtime] IS NOT NULL AND [depdelay] > 0 

Click to update in SQL expression editor 

ADD FILTER

Entity to extract (only necessary for behavior based segments)

Click to update in SQL expression editor

ADD ENTITY

CANCEL

SAVE

6. Click Save.

The renamed/edited segment appears in the Segments interface.

The screenshot shows the Arcadia Data interface for the 'Flight Delays' dataset. The left sidebar contains navigation links: Dataset Detail, Related Dashboards, Fields, Data Model, Analytical Views (0), Events (0), Segments (2), Filter Associations (0), and Permissions. The main content area is titled 'Dataset: Flight Delays' and includes a '+ NEW SEGMENT' button. Below this is a 'Segments' section with a table listing existing segments.

Group	Entity	Name	Apply To New Visuals	Filters	Actions
Delay Status		Delayed Departure		([airtime] IS NOT NULL AND [depdelay] > 0)	Edit Clone Delete
Delay Status		On-Time Departure		([airtime] IS NOT NULL AND [depdelay] <= 0)	Edit Clone Delete

Creating entity segments

Entity segments are a more advanced use of segments, where the extraction of an entity enables cohort analysis.

Procedure

1. On the main navigation bar, click Data.
2. Select Flight Delays dataset.
Dataset side navigation appears, open at Dataset Detail view.
3. In the side navigation menu, click Segments.

- 4. In the Segments list, find the segment to clone, and click sClone.

To make our task simpler, we are first cloning the Delayed Departure segment that was created in Cloning Segments .

ARCADIA DATA

HOMEVISUALSDATA

⚙️

HELP

admin

Dataset Detail

Related Dashboards

Fields

Data Model

Analytical Views0

Events0

Segments2

Filter Associations0

Permissions

Dataset: Flight Delays

NEW DASHBOARD

Segments

NEW SEGMENT

Group	Entity	Name	Apply To New Visuals	Filters	Actions
Delay Status		Delayed Departure		([airtime] IS NOT NULL AND [depdelay] > 0)	<div>EditCloneDelete</div>
Delay Status		On-Time Departure		([airtime] IS NOT NULL AND [depdelay] <= 0)	<div>EditCloneDelete</div>

A pre-populated Add Segment modal window appears.

5. Edit the Add Segment modal window by changing its entries.
 - Change Segment Name to Delayed Airlines.
 - Leave other fields 'as is'.
 - Under Entity to extract, click the text box to open the Entity Expression modal window.

Add Segment

Segment Name

Delayed Airlines

Segment Group

Delay Status

☐ Apply as a filter to new visuals

Filters

You can enter multiple filters and they will automatically have an "AND" logic between them

[airtime] IS NOT NULL AND [depdelay] > 0

Click to update in SQL expression editor

ADD FILTER

Entity to extract (only necessary for behavior based segments)

Click to update in SQL expression editor

ADD ENTITY

CANCEL

SAVE

6. In the Entity Expression modal window, enter the following expression:

```
[airlineid]
```

7. Click Validate Expression to verify correctness.

Entity Expression ✕

[airlineid]

VALIDATE EXPRESSION

☒ Autocomplete on

All Functions

abs
acos
add_months
adddate
AND
appx_median
ascii
asin
atan

All Fields & Tables

A abbreviation
actualelapse...
airlineid
A airport_codes...
A airport_codes...
A airport_codes...
A airport_codes...
1.2 airport_lat_lo...
A airport_lat_lo

CANCEL

APPLY

8. Click Apply to save the expression and return to the Add Segment modal window.
9. In the Add Segment modal window, notice the new expression in the Entity to extract field.

10. Click Save.

Add Segment

Segment Name

Delayed Airlines

Segment Group

Delay Status

☐ Apply as a filter to new visuals

Filters

You can enter multiple filters and they will automatically have an "AND" logic between them

[airtime] IS NOT NULL AND [depdelay] > 0

Click to update in SQL expression editor

ADD FILTER

Entity to extract (only necessary for behavior based segments)

[airlineid]

ADD ENTITY

CANCEL

SAVE

The new segment definition appears in the Segments interface. The segment Delayed Airlines has an entry in the Entity column, airlineid.

ARCADIA DATA HOME VISUALS DATA ⚙️ HELP admin

Dataset: Flight Delays NEW DASHBOARD

Segments + NEW SEGMENT

Group	Entity	Name	Apply To New Visuals	Filters	Actions
Delay Status		Delayed Departure		((airtime) IS NOT NULL AND [depdelay] > 0)	Edit Clone Delete
Delay Status		On-Time Departure		((airtime) IS NOT NULL AND [depdelay] <= 0)	Edit Clone Delete
Delay Status	[airlineid]	Delayed Airlines		((airtime) IS NOT NULL AND [depdelay] > 0)	Edit Clone Delete

To view how segments work in a visual, see [Using Segments in Visuals](#).

Related Information

[Cloning segments](#)

[Using segments in visuals](#)

Deleting segments

CDP Data Visualization lets you easily delete dataset segment definition.

Procedure

1. On the main navigation bar, click DATA.
2. Select Flight Delays dataset.
Dataset side navigation appears, open at Dataset Detail view.
3. In the side navigation menu, click Segments.

4. In the Segments list, find the segment to delete, and click Delete.
- In this example, we are deleting Short Flights.

ARCADIA DATA

HOMEVISUALSDATA

HELPadmin

Dataset Detail

Related Dashboards

Fields

Data Model

Analytical Views0

Events0

Segments3

Filter Associations0

Permissions

Dataset: Flight Delays

NEW DASHBOARD

SegmentsNEW SEGMENT

Group	Entity	Name	Apply To New Visuals	Filters	Actions
Delay Status		Delayed Flights		([airtime] IS NOT NULL AND [depdelay] > 0)	EditCloneDelete
Delay Status		On-Time Departure		([airtime] IS NOT NULL AND [depdelay] <= 0)	EditCloneDelete
Flight Time		Short Flights		([airtime] < 60)	EditCloneDelete

5. Click Delete in the confirmation modal window.

Are you sure you want to delete selected segment?

Cancel

Delete

The segment Short Flights no longer appears in the Segments interface.

ARCADIA DATA

HOMEVISUALSDATA

HELPadmin

Dataset Detail

Related Dashboards

Fields

Data Model

Analytical Views0

Events0

Segments2

Filter Associations0

Permissions

Dataset: Flight Delays

NEW DASHBOARD

SegmentsNEW SEGMENT

Group	Entity	Name	Apply To New Visuals	Filters	Actions
Delay Status		Delayed Flights		([airtime] IS NOT NULL AND [depdelay] > 0)	EditCloneDelete
Delay Status		On-Time Departure		([airtime] IS NOT NULL AND [depdelay] <= 0)	EditCloneDelete

Creating segments from filter definitions

Besides creating segments in the dataset detail view, Cloudera Data Visualization also lets you create segments from a filter definition of a visual.

About this task

See the following topics to create a new segment from a filter in a visual, then edit the filter to add another segment, and finally update an existing segment.

- [Creating the first segment](#)
- [Adding the second segment](#)
- [Viewing the defined segments](#)
- [Adding values to an existing segment](#)

Procedure

Creating the first segment

Follow these steps to create a new data segment, Mountain from a filter definition:

1. Start a new visual based on dataset Flight Delays, based on data previously imported into Data Visualization from a data file. See, [Creating a visual](#).
2. Add `originstate` to the Filters shelf.

The screenshot shows the Cloudera Data Visualization interface. On the left, a table titled "Flight Delays" displays data for the year 2014, quarter 1, month 1, with columns for dayofmonth and dayofweek. The table has 5 columns: year, quarter, month, dayofmonth, and dayofweek. The data rows show values for days 28, 29, 30, and 31 of the month.

On the right, the "VISUALS" and "DATA" panels are visible. The "VISUALS" panel shows a "Table" view selected. The "DATA" panel shows the "Flight Delays" dataset with "Sample Mode" set to "OFF". The "Dimensions" shelf contains fields: origin, origincityname, originstate, originstatename, and dest. The "Measures" shelf contains fields: flights_2014, Record Count, year, quarter, month, dayofmonth, and dayofweek. A green arrow points to the "originstate" field in the "Filters" shelf, indicating it has been added as a filter.

3. In the Filter for originstate window modal, select CO and WY, and click Apply.

Filter for originstate ✕

Values Pattern Set values

☐ AZ

☐ CA

☒ CO

☐ CT

☐ WI

☐ WV

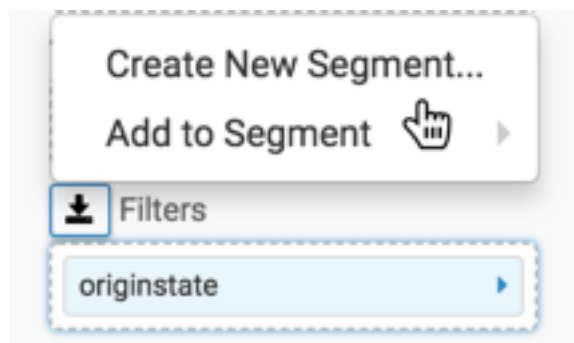
☒ WY

CANCEL APPLY

4. Click the Save Down icon to the left of the Filters shelf.



5. From the menu, select Create New Segment.



6. In the Add Segment window modal, specify the following:

- Under Segment Name, enter Mountain.
- Under Segment Group, enter Geographic Region.
- Under Filters, notice the new filter expression, [originstate] in ('CO', 'WY').
- Leave Entity to Extract field empty.

7. Click Save.

Add Segment

×

Segment Name

Mountain

Segment Group

Geographic Region

☐ Apply as a filter to new visuals

Filters

You can enter multiple filters and they will automatically have an "AND" logic between them

[originstate] in ('CO','WY')


ADD FILTER

Entity to extract (only necessary for behavior based segments)

Click to update in SQL expression editor

ADD ENTITY

CANCEL

SAVE 

The new segment definition appears in Segments section on the right navigation menu.

Filter for arc_segment

Values

Delay Status

☐ is Delayed Airline

☐ is Delayed Departure

☐ is On-Time Departure

Geographic Region

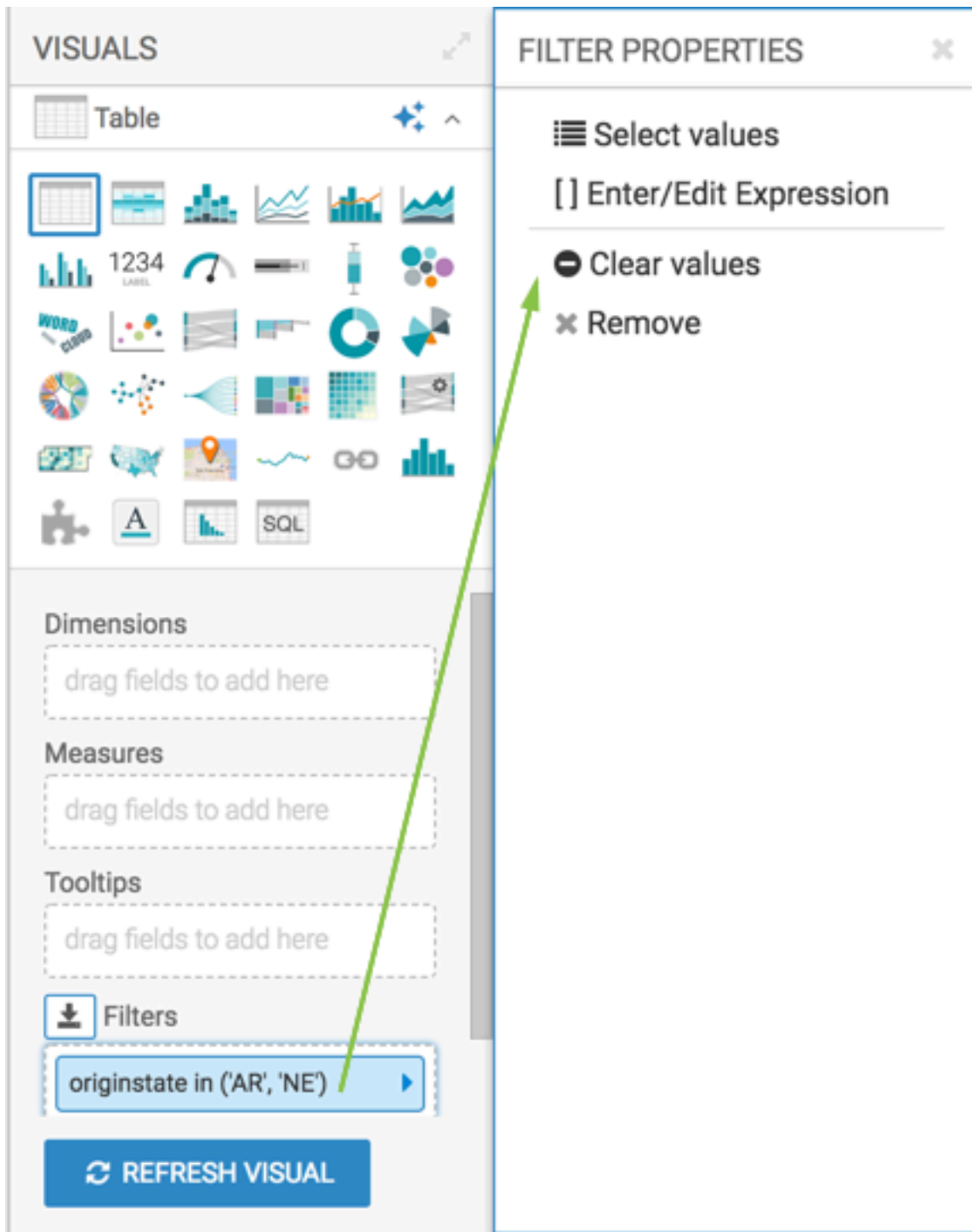
☐ is Mountain

Adding the second segment

Let's edit the filter defined in the first segment and create a second segment.

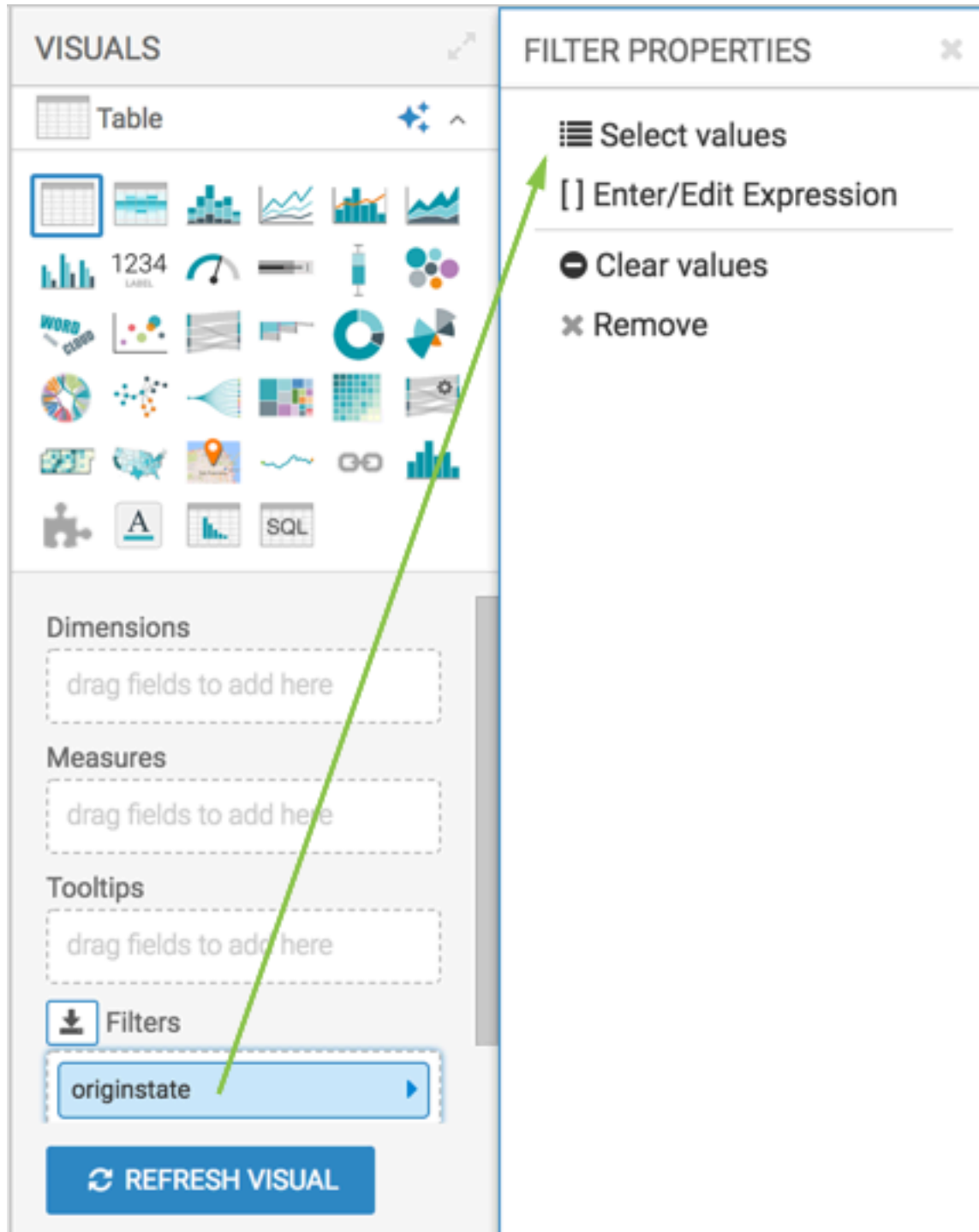
8. On the Filters shelf, at the right of the `originstate` filter, click the Down icon to open the Filter Properties menu.

9. From the Filter Properties menu, select Clear values.



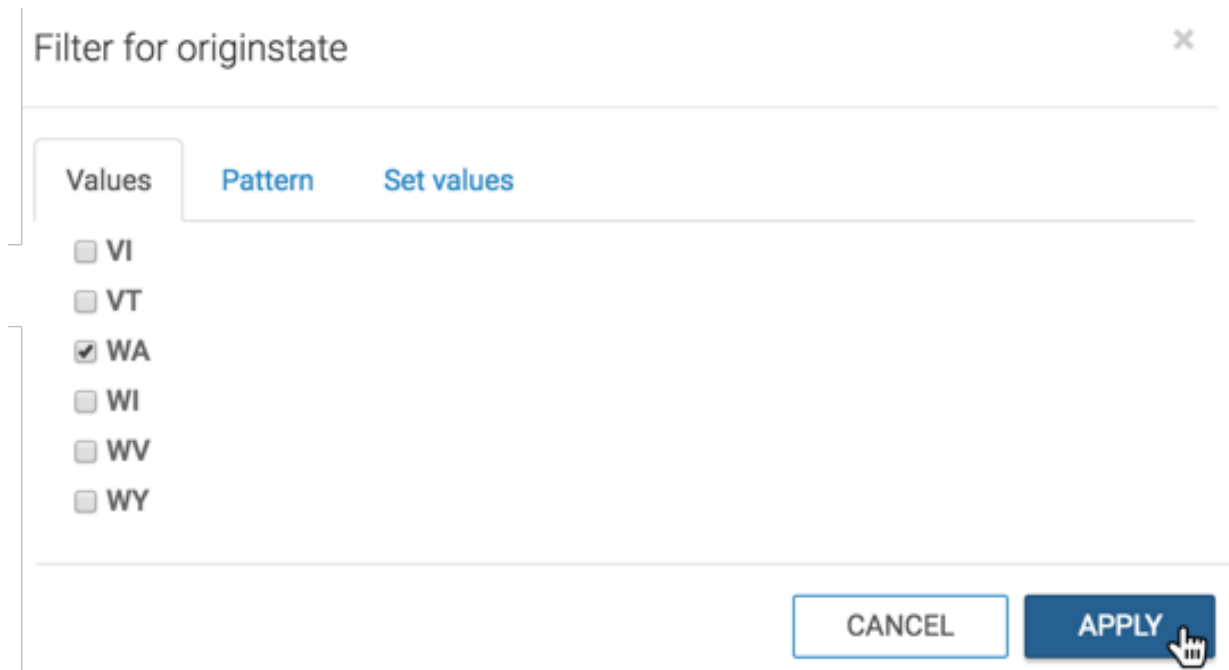
10. Click the Down icon to the right of the originstate filter to open the Filter Properties menu.

11. From the Filter Properties menu, select Select values.



12. In the Filter for originstate window modal, select WA.

13. Click Apply.



14. To define the next segment, select Create New Segment from the Save as Segment menu.



15. In the Add Segment modal window, specify the following:

- Under Segment Name, enter North West.
- Under Segment Group, enter Geographic Region that we used earlier.
- Under Filters, notice the new filter expression, [originate] in ('WA').
- Leave Entity to Extract field empty.

16. Click Save.

Add Segment ×

Segment Name

North West

Segment Group

Geographic Region

☐ Apply as a filter to new visuals

Filters

You can enter multiple filters and they will automatically have an "AND" logic between them

[originstate] in ('WA')

Click to update in SQL expression editor


ADD FILTER

Entity to extract (only necessary for behavior based segments)

Click to update in SQL expression editor

ADD ENTITY

CANCEL

SAVE 

Viewing the defined segments

17. In the side navigation bar of the visual interface, click Segments.

Notice that the two segments (and the segment group) we just defined appear in the Segments menu.

Filter for arc_segment

×

Values

Delay Status

☐ is Delayed Airline

☐ is Delayed Departure

☐ is On-Time Departure

Geographic Region

☐ is Mountain

☐ is North West

CANCEL

APPLY

Adding values to an existing segment

Let's add a new state, OR, to our existing segment North West.

18. Click inside the Filter shelf to open the Field Properties menu.

19. Click Clear values.

20. Click Select values.

21. Select OR in the Filter for originstate window modal.

22. Click Apply.

Filter for originstate

×

Values

Pattern

Set values

☐ OK

☒ OR

☐ PA

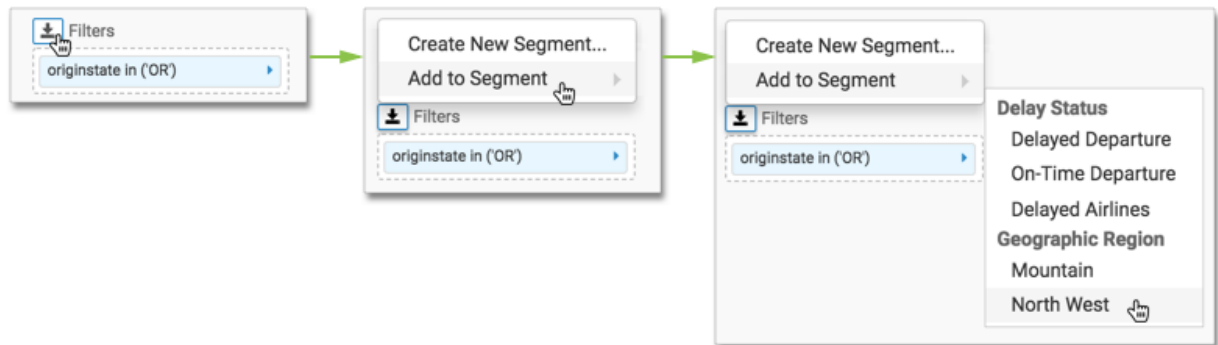
☐ PR

CANCEL

APPLY

23. To add the new selection to an existing segment, do the following:

- Click Filters.
- In the menu, select Add to Segment from the Save as Segment menu.
- In the secondary menu, under Geographic Regions, select North West.



Notice that when the Edit Segment modal window appears, it shows the existing configuration for the segment North West.

However, its Filters specification now includes Oregon: [originstate] in ('WA', 'OR').

24. Click Save to finish adding the specification to the segment.

Edit Segment

Segment Name

North West

Segment Group

Geographic Region

☐ Apply as a filter to new visuals

Filters

You can enter multiple filters and they will automatically have an "AND" logic between them

[originstate] in ('WA', 'OR')

Click to update in SQL expression editor

ADD FILTER

Entity to extract (only necessary for behavior based segments)

Click to update in SQL expression editor

ADD ENTITY

CANCEL

SAVE

What to do next

To view how segments work in a visual, see [Using Segments in Visuals](#)

Related Information

[Using segments in visuals](#)

[Creating a visual](#)

Using segments in visuals

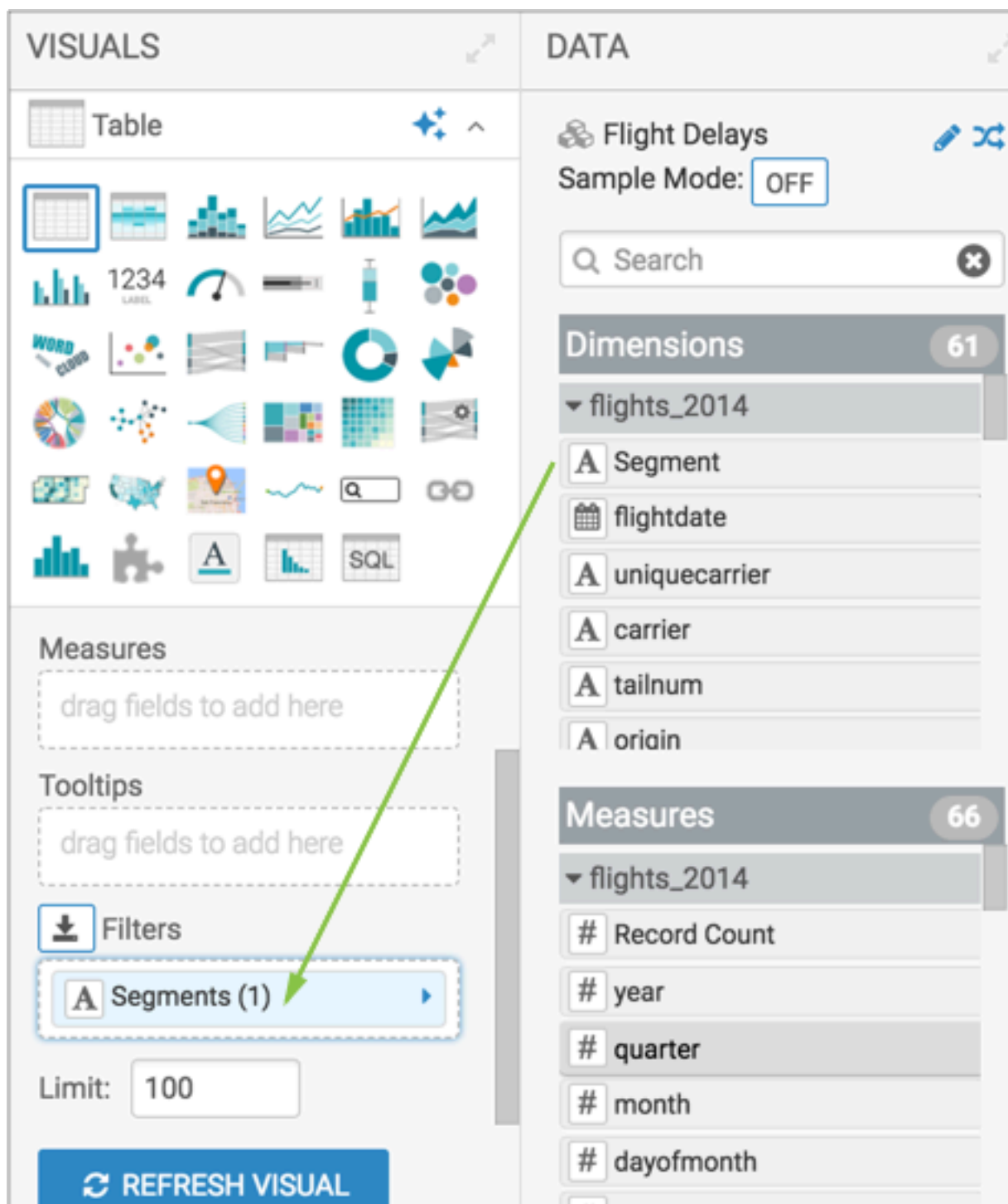
This article demonstrates how segments work in a visual.

About this task

After creating segments in the Dataset Segments user interface. See, [Creating Segments](#). Or creating segments from filter definitions of a visual, you can view this segmented data in a visual. See [Creating Segments from Filters](#).

Procedure

1. Click Data in the side navigation bar of the visual interface.



The segments (and the segment group) that you defined earlier appear in the Segments menu.

Filter for arc_segment

Values

Delay Status

☐ is Delayed Airline

☐ is Delayed Departure

☐ is On-Time Departure

Geographic Region

☐ is Mountain

☒ is North West


CANCEL


APPLY

2. Select one of the segments, and click Refresh Visual.

We selected North West.

The visual only shows the rows that match the segment criteria. In our case, the states WA and OR.

Flight Data 

enter subtitle... 

origin	origincityname	originstate	originstatefips
PDX	Portland, OR	OR	41
SEA	Seattle, WA	WA	53
BLI	Bellingham, WA	WA	53
SEA	Seattle, WA	WA	53
PDX	Portland, OR	OR	41
SEA	Seattle, WA	WA	53
SEA	Seattle, WA	WA	53

Related Information

[Creating segments](#)

[Creating segments from filter definitions](#)