



### **Scenario**

You use the Facets view when you want to see a list of keywords that are associated with a given facet. You can select any facet that you or your administrator has created as well as the Part of Speech (POS) facet and Phrase Constituent facet are defined and populated automatically by Content Analytics.

In this scenario, you are researching the popularity of various models of cars to see which are the most popular. You begin with the Facets view because you want to see the frequencies and correlation values for various models of cars.

## Explore Frequency and Correlation Values

The screenshot displays the IBM Content Analytics web interface. At the top, the header includes the IBM logo, the text "IBM Content Analytics", and navigation links for "Collection: NH... (change)", "Logged in as: Not ...", "Preferences", "My Profile", "Help", and "About". Below the header is a navigation bar with several tabs: "Documents", "Facets", "Time Series", "Deviations", "Trends", "Facet Pairs", "Connections", "Dashboard", and "Reports". The "Facets" tab is highlighted with a red rectangular box. A yellow callout box with a pointer to the "Facets" tab contains the text: "First, click the **Facets** tab to display the view." Below the navigation bar, the main content area shows "575366/575366 results" and a "Facet Navigation" section. This section includes a "Filter:" input field with a "Clear" button and a list of facets: "Part of Speech", "Phrase Constituent", "Named entity", "Troubles", "Category for Auto", "State", "City", "Vehicle/Equipment Corp", "Vehicle/Equipment Make", "Model", and "Model Year". To the right of the facets, there is a "keywords" dropdown menu and a "Filter:" input field. The main content area displays the message: "There are no returned documents that have keywords belonging to the selected facet. Try another facet."

## Explore Frequency and Correlation Values

The screenshot displays the IBM Content Analytics interface. At the top, the header shows 'IBM Content Analytics' and 'Collection: NH... (change)'. Below the header, there is a navigation bar with tabs for 'Documents', 'Facets', 'Time Series', 'Deviations', 'Trends', 'Facet Pairs', 'Connections', 'Dashboard', and 'Re'. The main content area shows '575366/575366 results matched'. On the left, there is a 'Facet Navigation' panel with a 'Filter:' input field and a 'Clear' button. Below the filter, there is a list of facets: 'Part of Speech', 'Phrase Constituent', 'Named entity', 'Troubles', 'Category for Auto', 'State', 'City', 'Vehicle/Equipment Corp', 'Vehicle/Equipment Make', 'Model', and 'Model Year'. In the center, there is a 'Show:' dropdown menu with 'Keywords' selected and circled in red. A yellow callout box points to the dropdown menu with the text: 'In the Show drop-down list, you can specify whether to show keywords or subfacets in your search results. For our purposes we will select **Keywords**.' Below the dropdown, there is a message: 'There are no returned documents matching the selected facet. Try another facet.'

## Explore Frequency and Correlation Values

IBM Content Analytics Collection: NH... (change) Logged in as: Not ... | Preferences | My Profile | Help | About

Documents Facets Time Series Deviations Trends Facet Pairs Connections Dashboard Re

575366/575366 results matched

Facet Navigation Default order

Filter:  Clear

Part of Speech <sup>2</sup>

Phrase Constituent <sup>2</sup>

Named entity <sup>2</sup>

Troubles

Category for Auto

State

City

Vehicle/Equipment Corp

Vehicle/Equipment Make

**Model**

Model Year

Show: Keywords Filter:

<input type="checkbox"/>	Keywords	Frequency	1	Correlation
<input type="checkbox"/>	EXPLORER	16254	1.0	
<input type="checkbox"/>	TAURUS	14357	1.0	
<input type="checkbox"/>	F150	10281	1.0	
<input type="checkbox"/>	GRAND CHEROKEE	9883	1.0	
<input type="checkbox"/>	WINDSTAR	9836	1.0	
<input type="checkbox"/>		9648	1.0	
<input type="checkbox"/>		8360	1.0	
<input type="checkbox"/>		7611	1.0	
<input type="checkbox"/>		7176	1.0	
<input type="checkbox"/>	FOCUS	7129	1.0	
<input type="checkbox"/>	CHILD SAFETY SEAT	6818	1.0	
<input type="checkbox"/>	SILVERADO	6733	1.0	

In the Facet Navigation panel, click **Model** to display all of the keywords that are associated with the facet.

## Explore Frequency and Correlation Values

IBM Content Analytics Collection: NH... (change) Logged in as: Not ... | Preferences | My Profile | Help | About

Documents Facets Time Series Deviations Trends Facet Pairs Connections Dashboard Re

575366/575366 results matched

Facet Navigation Default order

Filter: [ ] Clear

- Part of Speech
- Phrase Constituent
- Named entity
- Troubles
- Category for Auto
- State
- City
- Vehicle/Equipment Corp
- Vehicle/Equipment Make
- Model**
- Model Year
- Component Description

Show: Keywords Filter: [ ]

Keywords	Frequency	Correlation
<input type="checkbox"/> EXPLORER	16254	1.0
<input type="checkbox"/>		1.0
<input type="checkbox"/>		1.0
<input type="checkbox"/>		1.0
<input type="checkbox"/>		1.0
<input type="checkbox"/>		1.0
<input type="checkbox"/>		1.0
<input type="checkbox"/>		1.0
<input type="checkbox"/>		1.0
<input type="checkbox"/>		1.0
<input type="checkbox"/>		1.0
<input type="checkbox"/>		1.0
<input type="checkbox"/>		1.0
<input type="checkbox"/>		1.0
<input type="checkbox"/>		1.0
<input type="checkbox"/>		1.0
<input type="checkbox"/> SILVERADO	6733	1.0

When your query is set to the initial wild card expression \*\*, all correlation values are set to 1.0. The reason is because all of the documents in the collection are being compared to themselves, thus resulting in a correlation value of 1.0.

After you add a keyword to your query or enter a search expression, the correlation values are recalculated and reflect the degree of uniqueness of this keyword to the other documents in the corpus that match the query.

## Explore Frequency and Correlation Values

The screenshot displays the IBM Content Analytics interface. At the top, it shows the IBM logo, 'IBM Content Analytics', and user information: 'Collection: NH... (change)', 'Logged in as: Not ...', and links for 'Preferences', 'My Profile', 'Help', and 'About'. Below this is a navigation bar with tabs for 'Documents', 'Facets', 'Time Series', 'Deviations', 'Trends', 'Facet Pairs', 'Connections', 'Dashboard', and 'Reports'. A status bar indicates '575366/575366 results matched'. The main area is divided into a 'Facet Navigation' sidebar on the left and a main content area. The sidebar includes a 'Filter:' input field with a 'Clear' button and a list of facet categories: 'Part of Speech', 'Phrase Constituent', 'Named entity', 'Troubles', 'Category for Auto', 'State', 'City', 'Vehicle/Equipment Corp', 'Vehicle/Equipment Make', 'Model', and 'Model Year'. The 'Model' facet is selected. The main content area shows a table with columns for 'Keywords', 'Frequency', and 'Correlation'. The table lists several car models with their respective frequencies and correlation values. A callout box is overlaid on the table, providing instructions on how to narrow search results by selecting keywords.

Keywords	Frequency	Correlation
<input checked="" type="checkbox"/> EXPLORER	16254	1.0
<input type="checkbox"/> TAURUS		
<input type="checkbox"/> F150		
<input checked="" type="checkbox"/> GRAND CHEROKEE		
<input type="checkbox"/> WINDSTAR		
<input type="checkbox"/> CARAVAN		
<input type="checkbox"/> ACCORD		
<input checked="" type="checkbox"/> BLAZER		
<input type="checkbox"/> SUBURBAN		
<input type="checkbox"/> FOCUS		
<input type="checkbox"/> CHILD SAFETY SEAT	6076	1.0
<input type="checkbox"/> SILVERADO	6733	1.0

For your research, you decide that you want to narrow your search to the Ford Explorer, Jeep Grand Cherokee, and Chevy Blazer keywords, so you select the checkboxes next to each keyword.

## Explore Frequency and Correlation Values

The screenshot displays the IBM Content Analytics interface. At the top, the header shows 'IBM Content Analytics', 'Collection: NH... (change)', and 'Logged in as: Not ...'. Below the header, there are navigation tabs for 'Documents', 'Facets', 'Time Series', 'Deviations', 'Trends', and 'Reports'. The main content area shows '575366/575366 results matched'. On the left, there is a 'Facet Navigation' panel with a 'Filter:' input field and a 'Clear' button. The facets listed include 'Part of Speech', 'Phrase Constituent', 'Named entity', 'Troubles', 'Category for Auto', 'State', 'City', 'Vehicle/Equipment Corp', 'Vehicle/Equipment Make', 'Model', and 'Model Year'. The 'Model' facet is currently selected. The main results table has columns for 'Keywords', 'Frequency', 'Add to search with Boolean OR.', and 'Correlation'. The table lists various car models with their respective frequencies and correlation values. A tooltip points to the 'Add to search with Boolean OR.' button, stating: 'You have three options for adding a keyword with a Boolean to your search. First, you can click **Add to search with Boolean OR.**'

<input type="checkbox"/>	Keywords	Frequency	Add to search with Boolean OR.	Correlation
<input checked="" type="checkbox"/>	EXPLORER	16254	1.0	1.0
<input type="checkbox"/>	TAURUS	14357	1.0	1.0
<input type="checkbox"/>	F150	10281	1.0	1.0
<input checked="" type="checkbox"/>	GRAND CHEROKEE	9883	1.0	1.0
<input type="checkbox"/>	WINDSTAR	9836	1.0	1.0
<input type="checkbox"/>	CARAVAN	9648	1.0	1.0
<input type="checkbox"/>	ACCORD	8360	1.0	1.0
<input checked="" type="checkbox"/>	BLAZER	7611	1.0	1.0
<input type="checkbox"/>	SUBURBAN	7176	1.0	1.0
<input type="checkbox"/>	FOCUS	7129	1.0	1.0
<input type="checkbox"/>	CHILD SAFETY SEAT	6818	1.0	1.0
<input type="checkbox"/>	SILVERADO	6733	1.0	1.0



## Explore Frequency and Correlation Values

The screenshot displays the IBM Content Analytics interface. At the top, the header includes the IBM logo, 'IBM Content Analytics', and user information: 'Collection: NH... (change)', 'Logged in as: Not ...', and links for 'Preferences', 'My Profile', 'Help', and 'About'. Below the header is a navigation bar with tabs for 'Documents', 'Facets', 'Time Series', 'Deviations', 'Trends', 'Facet Pairs', 'Connections', 'Dashboard', and 'Reports'. The main content area shows '575366/575366 results matched'. On the left, a 'Facet Navigation' panel is visible with a 'Filter:' input and a 'Clear' button. The main table displays search results with columns for 'Keywords', 'Frequency', 'Correlation', and 'Add to search with Boolean AND NOT'. The 'Grand Cherokee' row is highlighted in blue. A yellow callout box points to the 'Add to search with Boolean AND NOT' button, which is circled in red.

<input type="checkbox"/>	Keywords	F	Add to search with Boolean AND NOT.	Correlation
<input checked="" type="checkbox"/>	EXPLORER	16254	<input type="checkbox"/>	1.0
<input type="checkbox"/>	TAURUS	14357	<input type="checkbox"/>	1.0
<input type="checkbox"/>	F150	10281	<input type="checkbox"/>	1.0
<input checked="" type="checkbox"/>	GRAND CHEROKEE	9883	<input type="checkbox"/>	1.0
<input type="checkbox"/>	WINDSTAR	9836	<input type="checkbox"/>	1.0
<input type="checkbox"/>	CARAVAN	9648	<input type="checkbox"/>	1.0
<input type="checkbox"/>	ACCORD	8360	<input type="checkbox"/>	1.0
<input checked="" type="checkbox"/>	BLAZER	7611	<input type="checkbox"/>	1.0
<input type="checkbox"/>	SUBURBAN	7176	<input type="checkbox"/>	1.0
<input type="checkbox"/>	FOCUS	7129	<input type="checkbox"/>	1.0
<input type="checkbox"/>	CHILD SAFETY SEAT	6818	<input type="checkbox"/>	1.0
<input type="checkbox"/>	SILVERADO	6733	<input type="checkbox"/>	1.0

## Explore Frequency and Correlation Values

IBM Content Analytics Collection: NH... (change) Logged in as: Not ... | Preferences | My Profile | Help | About

Documents Facets Time Series Deviations Tre Connections Dashboard Reports

575366/575366 results matched

Facet Navigation Default order

Filter:

- Part of Speech <sup>2</sup>
- Phrase Constituent <sup>2</sup>
- Named entity <sup>2</sup>
- Troubles
- Category for Auto
  - State
  - City
  - Vehicle/Equipment Corp
  - Vehicle/Equipment Make
  - Model**
  - Model Year

Show: Keywords

<input type="checkbox"/>	Keywords		Add to search with Boolean AND.	Correlation
<input checked="" type="checkbox"/>	EXPLORER	16254		1.0
<input type="checkbox"/>	TAURUS	14357		1.0
<input type="checkbox"/>	F150	10281		1.0
<input checked="" type="checkbox"/>	GRAND CHEROKEE	9883		1.0
<input type="checkbox"/>	WINDSTAR	9836		1.0
<input type="checkbox"/>	CARAVAN	9648		1.0
<input type="checkbox"/>	ACCORD	8360		1.0
<input checked="" type="checkbox"/>	BLAZER	7611		1.0
<input type="checkbox"/>	SUBURBAN	7176		1.0
<input type="checkbox"/>	FOCUS	7129		1.0
<input type="checkbox"/>	CHILD SAFETY SEAT	6818		1.0
<input type="checkbox"/>	SILVERADO	6733		1.0

But for our purposes we will use the third option, **Add to search with Boolean AND.**

## Explore Frequency and Correlation Values

The screenshot displays the IBM Content Analytics interface. At the top, the header shows 'IBM Content Analytics' and 'Collection: NH... (change)'. The user is logged in as 'Not ...'. The main navigation bar includes 'Documents', 'Facets', 'Time Series', 'Deviations', 'Trends', 'Facet Pairs', 'Connections', 'Dashboard', and 'Reports'. Below this, a status bar indicates '575366/575366 results matched'. A 'Facet Navigation' panel on the left lists various filters, with 'Model' selected. The main content area shows a table of car models with columns for selection, model name, frequency, and correlation. A tooltip points to a 'Show query input area' button, stating: 'Alternatively, you could click **Show query input area** and create your own query. Building queries is explained in a separate tutorial.'

	Model	Frequency	Correlation
<input type="checkbox"/>	TAURUS	14357	1.0
<input type="checkbox"/>	F150	10281	1.0
<input checked="" type="checkbox"/>	GRAND CHEROKEE	9883	1.0
<input type="checkbox"/>	WINDSTAR	9836	1.0
<input type="checkbox"/>	CARAVAN	9648	1.0
<input type="checkbox"/>	ACCORD	8360	1.0
<input checked="" type="checkbox"/>	BLAZER	7611	1.0
<input type="checkbox"/>	SUBURBAN	7176	1.0
<input type="checkbox"/>	FOCUS	7129	1.0
<input type="checkbox"/>	CHILD SAFETY SEAT	6818	1.0
<input type="checkbox"/>	SILVERADO	6733	1.0

## Explore Frequency and Correlation Values

The screenshot shows the IBM Content Analytics interface. At the top, it displays 'IBM Content Analytics' and 'Collection: NH... (change)'. The user is logged in as 'Not ...'. The interface includes a navigation bar with options like Documents, Facets, Time Series, Deviations, Trends, Facet Pairs, Connections, Dashboard, and Reports. Below this, it shows '33748/575366 results matched'. The main area is divided into a 'Facet Navigation' panel on the left and a table of results on the right. The 'Facet Navigation' panel has a 'Filter' input and a 'Clear' button, with a list of facets including 'Part of Speech', 'Phrase Constituent', 'Named entity', 'Troubles', 'Category for Auto', 'State', 'City', 'Vehicle/Equipment Corp', 'Vehicle/Equipment Make', 'Model', and 'Model Year'. The 'Model' facet is selected. The table on the right has columns for 'Keywords', 'Frequency', and 'Correlation'. A tooltip is overlaid on the table, stating: 'Your results are displayed. Note that the correlation values now differ.'

Keywords	Frequency	Correlation
EXPLORER	16254	16.8
GRAND CHEROKEE	9883	16.7
		16.6
		4.2
		5.9
		3.6
		8.4
		9.5
		0.1
		0.1
		0.4
ROLLING BOARD KIT (27		1.9

## Explore Frequency and Correlation Values

The screenshot displays the IBM Content Analytics interface. At the top, the header shows 'IBM Content Analytics' and 'Collection: NH... (change)'. The user is logged in as 'Not ...'. The main navigation bar includes 'Documents', 'Facets', 'Time Series', 'Deviations', 'Trends', 'Facet Pairs', 'Connections', 'Dashboard', and 'Reports'. Below this, a status bar indicates '33748/575366 results matched'. The 'Facet Navigation' panel on the left shows various filters, with 'Model' selected. The main content area displays a table of search results for the 'Model' facet, sorted by frequency in descending order. A callout box points to the 'Facet Pairs' icon in the navigation bar, explaining that documents are sorted by frequency by default and that users can change this by clicking 'Specify your preferences for viewing search results'.

Keywords	Frequency	Correlation
EXPLORER	16254	16.6
GRAND CHEROKEE	9883	16.6
BLAZER	7611	16.6
FIRESTONE	1233	4.2
WILDERNESS	660	5.9
WILDERNESS AT	283	3.6
ATX	278	8.4
RADIAL ATX	108	9.5
UNKNOWN	77	0.1
WRANGLER	46	0.1
GOODYEAR	44	0.4
RUNNING BOARD KIT	27	1.9

## Explore Frequency and Correlation Values

The screenshot shows the 'Search and Result Preferences' dialog box in IBM Content Analytics. The 'Facets' tab is selected. The 'Default sort order' dropdown menu is highlighted with a red circle and set to 'Frequency'. A blue informational box explains that selecting correlation as the default sort order would result in facets chosen by frequency, and that using the default of 100 keywords would only show the 100 most frequent keywords. The 'Number of results per page' is set to 100, and 'Show keywords or subfacets by default' is set to 'Keywords'. 'Save' and 'Cancel' buttons are at the bottom right.

IBM Content Analytics Collection: NH... (change) Logged in as: Not ... | Preferences | My Profile | Help | About

Search and Result Preferences

Search Results Result Columns Facets Time Series Deviations Trends Facet Pairs C

Number of results per page: 100

Default sort order: Frequency

Show keywords or subfacets by default: Keywords

However, even if you select correlation as the default sort order, the list of facets that are displayed are chosen by frequency. Therefore, if you use the default of 100 keywords, you get only the 100 most frequent keywords.

Save Cancel

## Explore Frequency and Correlation Values

The screenshot displays the IBM Content Analytics interface. At the top, the header shows 'IBM Content Analytics' and 'Collection: NH... (change)'. The user is logged in as 'Not ...'. The main navigation bar includes 'Documents', 'Facets', 'Time Series', 'Deviations', 'Trends', 'Facet Pairs', 'Connections', 'Dashboard', and 'Reports'. Below this, a status bar indicates '33748/575366 results matched'. The 'Facet Navigation' section on the left shows a list of facets, with 'Model' selected. A callout box is overlaid on the table, stating: 'To use the Facets view effectively, it is important to understand the difference between frequency and correlation values.' The table itself has columns for 'Keywords', 'Frequency', and 'Correlation'. The 'Frequency' column includes a bar chart and a numerical value. The 'Correlation' column includes a bar chart and a numerical value. The table is sorted by 'Keywords'.

Keywords	Frequency	Correlation
EXPLORER	16254	16.8
EE	9883	16.7
	7611	16.6
	1233	4.2
	660	5.9
	283	3.6
	278	8.4
	108	9.5
	77	0.1
	46	0.1
GOODYEAR	44	0.4
RUNNING BOARD KIT	27	1.9

## Explore Frequency and Correlation Values

The screenshot displays the IBM Content Analytics interface. At the top, the header shows 'IBM Content Analytics' and 'Collection: NH... (change)'. The user is logged in as 'Not ...'. The main navigation bar includes 'Documents', 'Facets', 'Time Series', 'Deviations', 'Trends', 'Facet Pairs', 'Connections', 'Dashboard', and 'Reports'. Below this, a status bar indicates '33748/575366 results matched'. The 'Facet Navigation' section on the left lists various facets, with 'Model' selected. The main content area shows a table with columns for 'Keywords', 'Frequency', and 'Correlation'. A red box highlights the first few rows of the table, and a yellow callout box explains the frequency column.

Keywords	Frequency	Correlation
	16254	16.8
	9883	16.7
	7611	16.6
	1233	4.2
	660	5.9
<input type="checkbox"/> WILDERNESS AT	283	3.6
<input type="checkbox"/> ATX	278	3.4
<input type="checkbox"/> RADIAL ATX	108	9.5
<input type="checkbox"/> UNKNOWN	77	0.1
<input type="checkbox"/> WRANGLER	46	0.1
<input type="checkbox"/> GOODYEAR	44	0.4
<input type="checkbox"/> RUNNING BOARD KIT	27	1.9

The frequency value for a keyword indicates the number of documents found that contain the given keyword.



## Explore Frequency and Correlation Values

The screenshot displays the IBM Content Analytics interface. At the top, the header shows 'IBM Content Analytics' and 'Collection: NH... (change)'. The user is logged in as 'Not ...'. The main navigation bar includes 'Documents', 'Facets', 'Time Series', 'Deviations', 'Trends', 'Facet Pairs', 'Connections', 'Dashboard', and 'Reports'. Below this, it indicates '33748/575366 results matched'. The 'Facet Navigation' panel on the left shows various filters, with 'Model' selected. The main table displays the following data:

Keywords	Frequency	Correlation
EXPLORER	16.8	16.8
GRAND CH	16.7	16.7
BLAZER	16.6	16.6
FIRESTONE	4.2	4.2
WILDERNE	5.9	5.9
WILDERNES	3.6	3.6
ATX	278	8.4
RADIAL ATX	108	9.5
UNKNOWN	77	0.1
WRANGLER	46	0.1
GOODYEAR	44	0.4
RUNNING BOARD KIT	27	1.9

A tooltip points to the 'Correlation' column, stating: 'The correlation value for a keyword is a measure of the uniqueness of a high frequency as compared to all of the documents that match the current query.'

## Explore Frequency and Correlation Values

IBM Content Analytics Collection: NH... (change) Logged in as: Not ... | Preferences | My Profile | Help | About

Documents Facets Time Series Deviations Trends Facet Pairs Connections Dashboard Reports

33748/575366 results matched

Facet Navigation Default order Show: Keywords Filter:

Filter:

Part of Speech <sup>?</sup>  
Phrase Constituent <sup>?</sup>  
Named entity <sup>?</sup>  
Troubles  
Category for Auto  
State  
City  
Vehicle/Equipment Corp  
Vehicle/Equipment Make  
Model  
Model Year

Keywords	Frequency	Correlation
	16254	16.8
	9883	16.7
	7611	16.6
	1233	4.2
	660	5.9
	283	3.6
	278	8.4
	108	9.5
<input type="checkbox"/> WRANGLER	77	0.1
<input type="checkbox"/> GOODYEAR	46	0.1
<input type="checkbox"/> RUNNING BOARD KIT	44	0.4
	27	1.9

What does this mean in terms of your data? It means that the frequency value, while useful, might not be as revealing as the correlation value.

## Explore Frequency and Correlation Values

IBM Content Analytics

Collection: NH... (change) Logged in as: Not ... | Preferences | My Profile | Help | About

Documents Facets Time Series Deviations Trends Facet Pairs Connections Dashboard Reports

33748/575366 results matched

Facet Navigation Default order

Filter:

For example, in your results, the Grand Cherokee has a much higher correlation value than the Wrangler. This might mean that the Grand Cherokee is the more popular Jeep model. In any case, it indicates that there is something unusual happening in the data that requires more research.

Keywords	Frequency	Correlation
EXPLORER	16254	16.8
GRAND CHEROKEE	9883	16.7
BLAZER	7611	16.6
FIRESTONE	1233	4.2
WILDERNESS	660	5.9
WILDERNESS AT	283	3.6
ATX	278	8.4
RADIAL ATX	108	9.5
UNKNOWN	77	0.1
WRANGLER	46	0.1
GOODYEAR	44	0.4
RUNNING BOARD KIT	27	1.9

## Explore Frequency and Correlation Values

The screenshot displays the IBM Content Analytics interface. At the top, the header shows 'IBM Content Analytics' and 'Collection: NH... (change)'. The user is logged in as 'Not ...'. The main navigation bar includes 'Documents', 'Facets', 'Time Series', 'Deviations', 'Trends', 'Facet Pairs', 'Connections', 'Dashboard', and 'Reports'. Below this, a status bar indicates '33748/575366 results matched'. The 'Facet Navigation' section on the left lists various filters, with 'Model' selected. A tooltip over the 'Filter' section states: 'You can also sort and filter the information in the view.' The main content area shows a table with columns for 'Keywords', 'Frequency', and 'Correlation'. The table lists various keywords and their corresponding frequency and correlation values.

<input type="checkbox"/>	Keywords	Frequency	1	Correlation
<input type="checkbox"/>	EXPLORER	16254	16.8	
<input type="checkbox"/>	GRAND CHEROKEE	9883	16.7	
<input type="checkbox"/>	BLAZER	7611	16.6	
<input type="checkbox"/>	FIRESTONE	1233	4.2	
<input type="checkbox"/>	WILDERNESS	660	5.9	
<input type="checkbox"/>	WILDERNESS AT	283	3.6	
<input type="checkbox"/>	ATX	278	8.4	
<input type="checkbox"/>	RADIAL ATX	108	9.5	
<input type="checkbox"/>	UNKNOWN	77	0.1	
<input type="checkbox"/>	WRANGLER	46	0.1	
<input type="checkbox"/>	GOODYEAR	44	0.4	
<input type="checkbox"/>	RUNNING BOARD KIT	27	1.9	

## Explore Frequency and Correlation Values

The screenshot displays the IBM Content Analytics interface. At the top, the header shows 'IBM Content Analytics', 'Collection: NH... (change)', and 'Logged in as: Not ...'. Below the header is a navigation bar with tabs for Documents, Facets, Time Series, Deviations, Trends, Facet Pairs, Connections, Dashboard, and Reports. The main content area shows '33748/575366 results matched'. On the left, there is a 'Facet Navigation' panel with a 'Filter:' input field and a 'Clear' button. Below this, several facet categories are listed, including 'Part of Speech', 'Phrase Constituent', 'Named entity', 'Troubles', 'Category for Auto', 'State', 'City', 'Vehicle/Equipment Corp', 'Vehicle/Equipment Make', 'Model', and 'Model Year'. The 'Model' facet is currently selected. The main table displays the following data:

<input type="checkbox"/>	Keyword			Correlation	
<input type="checkbox"/>	EXPLORER			16.8	<div style="width: 100%;"></div>
<input type="checkbox"/>	GRAND CHERO			16.7	<div style="width: 100%;"></div>
<input type="checkbox"/>	BLAZER			16.6	<div style="width: 100%;"></div>
<input type="checkbox"/>	FIRESTONE			4.2	<div style="width: 20%;"></div>
<input type="checkbox"/>	WILDERNESS	660		5.9	<div style="width: 40%;"></div>
<input type="checkbox"/>	WILDERNESS AT	283		3.6	<div style="width: 25%;"></div>
<input type="checkbox"/>	ATX	278		8.4	<div style="width: 60%;"></div>
<input type="checkbox"/>	RADIAL ATX	108		9.5	<div style="width: 70%;"></div>
<input type="checkbox"/>	UNKNOWN	77		0.1	<div style="width: 5%;"></div>
<input type="checkbox"/>	WRANGLER	46		0.1	<div style="width: 5%;"></div>
<input type="checkbox"/>	GOODYEAR	44		0.4	<div style="width: 10%;"></div>
<input type="checkbox"/>	RUNNING BOARD KIT	27		1.9	<div style="width: 15%;"></div>

A yellow tooltip box is overlaid on the table, containing the text: 'You can click any of the headers to sort the results by that header. For example, you can click Correlation to sort the data by correlation value.'

## Explore Frequency and Correlation Values

The screenshot displays the IBM Content Analytics interface. At the top, the header shows "IBM Content Analytics" and "Collection: NH... (change)". The user is logged in as "Not ...". The main navigation bar includes "Documents", "Facets", "Time Series", "Deviations", "Trends", "Facet Pairs", "Connections", "Dashboard", and "Reports". Below this, it indicates "33748/575366 results matched".

The "Facet Navigation" panel on the left shows a filter for "Model" selected. The main table displays a list of keywords with their corresponding correlation values. The table is sorted in descending order of correlation, as indicated by a downward arrow in the "Correlation" header. A tooltip explains: "The data is now sorted in descending order by correlation value. Note the arrow in the header that indicates the sort type is descending. Click the heading again, and the column is sorted in ascending order." The "Correlation" header is highlighted with a red box and a "1" with a downward arrow.

Keywords	Frequency	Correlation
EXPLORER	270	16.8
GRAND	1660	16.7
BLAZER	1660	16.6
RADIAL	1660	9.5
ATX	270	8.4
WILDERNESS	660	5.9
FIRESTONE	1233	4.2
TERRAIN	5	3.6
FORTERA HL (P)	4	3.6
WILDERNESS AT	283	3.6
ATX II	9	3.4
P23575R15	0	2.4

## Explore Frequency and Correlation Values

The screenshot displays the IBM Content Analytics interface. At the top, the header shows "IBM Content Analytics" and "Collection: NH... (change)". The user is logged in as "Not ...". The main navigation bar includes tabs for Documents, Facets, Time Series, Deviations, Trends, Facet Pairs, Connections, Dashboard, and Reports. Below this, a status bar indicates "33748/575366 results matched".

The interface is divided into several sections:

- Facet Navigation:** A sidebar on the left with a "Filter:" input field and a "Clear" button. Below it, a list of facets is shown, including "Part of Speech", "Phrase Constituent", "Named entity", "Troubles", "Category for Auto", "State", "City", "Vehicle/Equipment Corp", "Vehicle/Equipment Make", "Model" (highlighted), and "Model Year".
- Show:** A dropdown menu set to "Keywords".
- Filter:** A text input field, highlighted with a red box in the image.
- Table:** A table with columns for "Keywords", "Frequency", and "Correlation". The table lists various keywords and their corresponding frequency and correlation values.

A callout box points to the "Filter:" field with the following text:

You can also use the Filter field to show only the results that you want to look at.

For example, there are different types of Wranglers, so maybe you want to look at the correlation and frequency values for all of the Wranglers as you research the question of popularity.

Keywords	Frequency	Correlation
EXPLORER	16254	
GRAND CHEROKEE	9883	
BLAZER	7611	
RADIAL ATX	108	
ATX	278	
WILDERNESS	660	
FIRESTONE	1233	
TERRAIN	5	
FORTERA HL (P)	4	3.6
WILDERNESS AT	283	3.6
ATX II	9	3.4
P23575R15	0	2.4

## Explore Frequency and Correlation Values

The screenshot displays the IBM Content Analytics interface. At the top, the header includes the IBM logo, 'IBM Content Analytics', and navigation links for 'Collection: NH...', 'Logged in as: Not ...', 'Preferences', 'My Profile', 'Help', and 'About'. Below the header, there are tabs for 'Documents', 'Facets', and 'T'. A yellow callout box points to a 'Filter:' input field containing the text 'wr', with a red box around it. The callout text reads: 'In the Filter field, type a few letters from the keyword Wrangler. The view dynamically changes to show only the results that contain those letters.' Below the filter field, a table lists search results with columns for 'Keywords', 'Frequency', and 'Correlation'. The 'Model' facet is selected in the left sidebar.

33748/575366 results matched

Facet Navigation Default order

Filter:  Clear

<input type="checkbox"/>	Keywords	Frequency	Correlation	
<input type="checkbox"/>	WRANGLER AT	11	0.8	■
<input type="checkbox"/>	WRANGLER ST	9	0.7	■
<input type="checkbox"/>	WRANGLER RTS	12	0.7	■
<input type="checkbox"/>	WRANGLER SRA	4	0.2	■
<input type="checkbox"/>	WRANGLER	46	0.1	■
<input type="checkbox"/>	WRANGLER HP	2	0.1	■
<input type="checkbox"/>	WRANGLER HT	3	0.1	■

Facet Pairs Connections Dashboard Reports

Part of Speech <sup>2</sup>

Phrase Constituent <sup>2</sup>

Named entity <sup>2</sup>

Troubles

Category for Auto

State

City

Vehicle/Equipment Corp

Vehicle/Equipment Make

Model

Model Year



## Explore Frequency and Correlation Values

The screenshot displays the IBM Content Analytics interface. At the top, the header shows "IBM Content Analytics" and "Collection: NH... (change)". The user is logged in as "Not ...". Navigation tabs include Documents, Facets, Time Series, Deviations, Trends, Facet Pairs, Connections, Dashboard, and Reports. The main area shows "33748/575366 results matched". A "Facet Navigation" panel on the left lists various facets, with "Model" selected. A table displays the following data:

Keywords	Frequency	Correlation
EXPLORER	16254	16.8
GRAND CHEROKEE	9883	16.7
BLAZER	7611	16.6
RADIAL ATX	108	9.5
ATX	278	8.4
WILDERNESS	660	5.9
FIRESTONE	1233	4.2
TERRAIN	5	3.6
FORTERA HL (P)	4	3.6
WILDERNESS AT	283	3.6
ATX II	9	3.4
P23575R15	9	2.4

A callout box with a yellow background and a speech bubble icon contains the text: "To remove the filter, just clear the Filter field." This callout points to a "Filter:" input field, which is currently empty and highlighted with a red border.

## Explore Frequency and Correlation Values

The screenshot displays the IBM Content Analytics interface. At the top, the navigation bar includes the IBM logo, 'IBM Content Analytics', and user information: 'Collection: NH... (change)', 'Logged in as: Not ...', and links for 'Preferences', 'My Profile', 'Help', and 'About'. Below this is a secondary navigation bar with tabs for 'Documents', 'Facets', 'Time Series', 'Deviations', 'Trends', 'Facet Pairs', 'Connections', 'Dashboard', and 'Reports'. The 'Documents' tab is highlighted with a red box.

A yellow callout box points to the 'Documents' tab with the following text: 'When you determine which keywords you want to research further, you can click the **Documents** tab to look at the actual content of the documents that contain the keywords of interest. In this way, you can begin to narrow down which keywords are most meaningful to your research.'

The main content area shows a table of keywords. The table has columns for 'Keywords', 'Frequency', and 'Correlation'. The 'Frequency' column includes a numerical value and a small blue bar chart. The 'Correlation' column includes a numerical value and a green bar chart. The table is sorted by correlation in descending order.

<input type="checkbox"/>	Keywords	Frequency	Correlation
<input type="checkbox"/>	EXPLORER	16254	16.8
<input type="checkbox"/>	GRAND CHEROKEE	9883	16.7
<input type="checkbox"/>	BLAZER	7611	16.6
<input type="checkbox"/>	RADIAL ATX	108	9.5
<input type="checkbox"/>	ATX	278	8.4
<input type="checkbox"/>	WILDERNESS	660	5.9
<input type="checkbox"/>	FIRESTONE	1233	4.2
<input type="checkbox"/>	TERRAIN	5	3.6
<input type="checkbox"/>	FORTERA HL (P)	4	3.6
<input type="checkbox"/>	WILDERNESS AT	283	3.6
<input type="checkbox"/>	ATX II	9	3.4
<input type="checkbox"/>	P23575R15	9	2.4

## Explore Frequency and Correlation Values

IBM Content Analytics

Collection: NH... (change) Logged in as: Not ... | Preferences | My Profile | Help | About

Documents | Facets | Time Series | Deviations | Trends | Facet Pairs | Connections | Dashboard | Reports

Results 1-10 of 3000  
(33748/575366 results matched)

Results per page: 10

Flags	Actions	Source	Date	Title	Thumbnail
<input type="checkbox"/>		ica	12/17/99	196117	
VEHICLE DID NOT GO INTO REVERSE. DEALER HAS REPAIRED THE VEHICLE. *AK					
<input type="checkbox"/>				15436	
WELD ON CONTR... N FRAME TO BREAK. *SKD					
<input type="checkbox"/>				73623	
EA94-038, BRAKE... EN APPLYING THE BRAKES, VEHICLE... OSITION , SPEED 10MPH. *AK					
<input type="checkbox"/>				262535	
SOMETIMES WIN... SOME STAYS... SOMETIMES THEY WILL... CONSUMER IS HAVING THE... SAME PROBLEM AS MENTIONED IN DECAL 98V150001. CONSUMER STATES VEHICLE WILL NOT PASS...					

Facet Navigation Default order

Filter:  Clear

- Part of Speech
- Phrase Constituent
- Named entity
- Troubles
- Category for Auto
- State
- City
- Vehicle/Equipment Corp
- Vehicle/Equipment Make
- Model

This concludes the tutorial on exploring frequency and correlation values.