

**IBM® Content Analytics**

**Explore How Frequencies Deviate from the Expected Average**

**Copyright © 2010 International Business Machines, Inc.**

## Explore How Frequencies Deviate from the Expected Average

---

### **The Deviations view**

The Deviations view shows how the frequency of keywords deviates from the expected average for a given time period. Unlike the Trends view, the Deviations view does not focus on previous time periods, but instead focuses on the expected average and takes into account all the averages of the other frequency counts for the given time period.

The Deviations view is useful for identifying patterns that occur cyclically and alerts you when those cyclic patterns have an unexpected change. You can use this view to show seasonal patterns in your data that occur on a monthly, weekly, or daily basis.

## Explore How Frequencies Deviate from the Expected Average

---

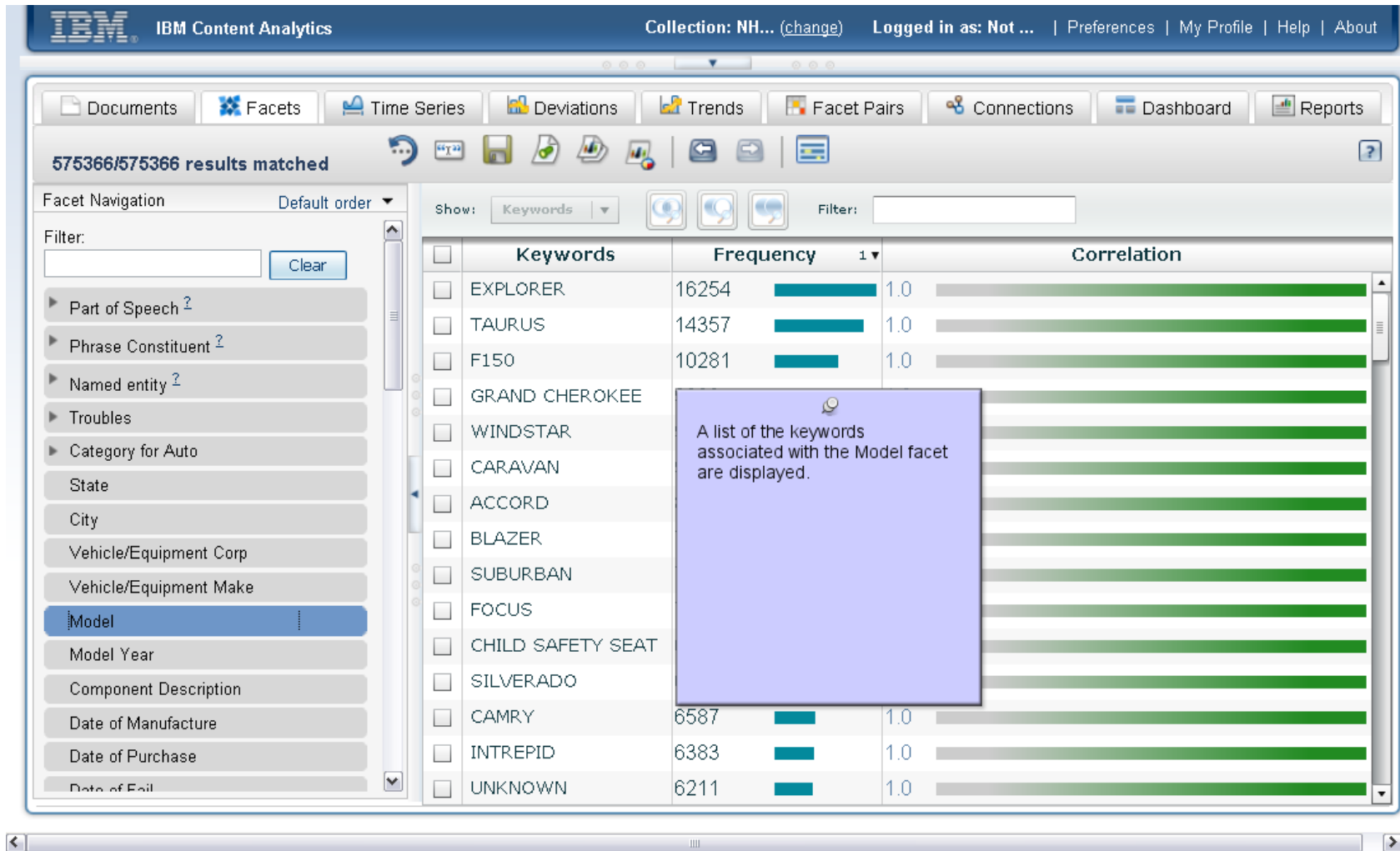
### **Scenario**

To explore deviations, you must select a facet. In this scenario, you are researching engine fires in Ford Explorers.

## Explore How Frequencies Deviate from the Expected Average

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 tabs for "Documents", "Facets", "Time Series", "Deviations", "Trends", "Facet Pairs", "Connections", "Dashboard", and "Reports". The main content area shows "575366/575366 results matched" and a "Facet Navigation" pane on the left. The "Facet Navigation" pane has a "Default order" dropdown and a "Filter:" input field with a "Clear" button. A list of facets is shown, with "Model" highlighted by a red border. A yellow callout box points to the "Model" facet with the text: "To add your first facet to the search, in the Facet Navigation pane, click **Model**." The main search results area shows "Show: Keywords" and a "Filter:" input field. Below this, a message states: "There are no returned documents that have keywords belonging to the selected facet. Try another facet."

## Explore How Frequencies Deviate from the Expected Average



## Explore How Frequencies Deviate from the Expected Average

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 are navigation 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, there is a 'Facet Navigation' panel with a 'Filter:' section and a list of facets including Part o, Phras, Named entity 2, Troubles, Category for Auto, State, City, Vehicle/Equipment Corp, Vehicle/Equipment Make, Model (highlighted), Model Year, Component Description, Date of Manufacture, Date of Purchase, and Date of Exit. A yellow callout box with a red border points to the 'EXPLORER' keyword in the table, containing the text: 'Click the checkbox next to the keyword EXPLORER.' The table has columns for 'Keywords', 'Frequency', and 'Correlation'. The 'EXPLORER' row is highlighted in blue, and its checkbox is checked. The table data is as follows:

Keywords	Frequency	Correlation
<input checked="" 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"/> CARAVAN	9648	1.0
<input type="checkbox"/> ACCORD	8360	1.0
<input 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
<input type="checkbox"/> CAMRY	6587	1.0
<input type="checkbox"/> INTREPID	6383	1.0
<input type="checkbox"/> UNKNOWN	6211	1.0

## Explore How Frequencies Deviate from the Expected Average

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, there are navigation tabs for 'Documents', 'Facets', 'Time S...', 'Trends', 'Facet Pairs', 'Connections', 'Dashboard', and 'Reports'. A status bar indicates '575366/575366 results matched'. On the left, a 'Facet Navigation' panel shows a 'Filter:' input field 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' (highlighted), 'Model Year', 'Component Description', 'Date of Manufacture', 'Date of Purchase', and 'Date of Exit'. The main content area features a table with columns for 'Keywords', 'Add to search with Boolean AND.', and 'Correlation'. A callout box points to the 'Add to search with Boolean AND.' icon, containing the text: 'To narrow your search to the Explorer model, click Add to search with Boolean AND.' The table lists various car models with their respective counts and correlation values.

<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 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 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
<input type="checkbox"/>	CAMRY	6587	1.0
<input type="checkbox"/>	INTREPID	6383	1.0
<input type="checkbox"/>	UNKNOWN	6211	1.0

## Explore How Frequencies Deviate from the Expected Average

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'. The main area shows '16254/575366 results matched'. On the left is a 'Facet Navigation' panel with a 'Filter' input and a 'Clear' button. The 'Category for Auto' facet is highlighted with a red box. A yellow callout box points to this facet with the text: 'Too add a second facet to the search, in the Facet Navigation panel, expand the **Category for Auto** section.' The main table displays the following data:

Keywords	Frequency	Correlation
EXPLORER	16254	35.0
	1225	8.7
	653	12.1
	274	17.2
	273	7.3
	106	19.3
UNKNOWN	52	0.2
WRANGLER	30	0.2
GOODYEAR	23	0.4
AIR BAG SWITCH	20	1.8
WILDERNESS HT	17	1.0
BRIDGESTONE	12	0.3
WRANGLER RTS	11	1.3
MICHELIN	11	0.2
ATX II	9	7.1



## Explore How Frequencies Deviate from the Expected Average

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 are navigation tabs for Documents, Facets, Time Series, Deviations, Trends, Facet Pairs, Connections, Dashboard, and Reports. The main content area shows '16254/575366 results matched'. On the left, the 'Facet Navigation' pane is visible, with a 'Filter' input field and a 'Clear' button. The 'Incident' subfacet is highlighted with a red box. A yellow callout box with the text 'Expand the Incident subfacet.' points to this subfacet. The main table displays the following data:

Keywords	Frequency	Correlation
EXPLORER	16254	35.0
FIRESTONE	1225	8.7
WILDERNESS	653	12.1
ATX	274	17.2
WILDERNESS AT	273	7.3
RADIAL ATX	106	19.3
UNKNOWN	52	0.2
WRANGLER	30	0.2
GOODYEAR	23	0.4
HTCH	20	1.8
HT	17	1.0
BRIDGESTONE	12	0.3
WRANGLER RTS	11	1.3
MICHELIN	11	0.2
ATX II	9	7.1

## Explore How Frequencies Deviate from the Expected Average

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

16254/575366 results matched

Facet Navigation Default order

- Part of Speech <sup>?</sup>
- Phrase Constituent <sup>?</sup>
- Named entity <sup>?</sup>
- Troubles
- Category for Auto
  - Authorities
  - Company
  - Component
  - Customer
  - Environment
  - Gas
  - Incident
    - Accident
    - Blowout
    - Burn**
    - Deficiency
    - Electricity

Show: Keywords Filter:

<input type="checkbox"/>	Keywords	Frequency	1	Correlation
<input type="checkbox"/>	fire	544	0.8	
<input type="checkbox"/>	smoke	261	0.9	
<input type="checkbox"/>	flame	95	0.7	
<input type="checkbox"/>	fire hazard	12	0.2	

Scroll down and click the **Burn** subfacet.

## Explore How Frequencies Deviate from the Expected Average

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 are navigation tabs for 'Documents', 'Facets', 'Time Series', 'Deviations', 'Trends', 'Facet Pairs', 'Connections', 'Dashboard', and 'Reports'. The main content area shows '16254/575366 results matched'. On the left, a 'Facet Navigation' sidebar lists various categories, with 'Burn' selected under the 'Incident' category. The main table displays a list of keywords with their frequency and correlation. A tooltip is visible over the table, stating: 'A list of the keywords associated with the Burn subfacet is displayed.'

Keywords	Frequency	Correlation
<input type="checkbox"/> fire		High
<input type="checkbox"/> smoke		High
<input type="checkbox"/> flame		Medium
<input type="checkbox"/> fire hazard		Low

## Explore How Frequencies Deviate from the Expected Average

The screenshot shows the IBM Content Analytics interface. The top navigation bar includes the IBM logo, 'IBM Content Analytics', and user information: 'Collection: NH... (change)', 'Logged in as: Not ...', '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 main content area displays a table of keywords. A callout box with a yellow background and black text says 'Click the top-most checkbox to select all of the keywords.' The table has columns for 'Keywords', 'Frequency', and 'Correlation'. The 'Frequency' column includes a bar chart and a numerical value. The 'Correlation' column includes a horizontal bar chart. The 'Incident' category is expanded in the left sidebar, with 'Burn' selected.

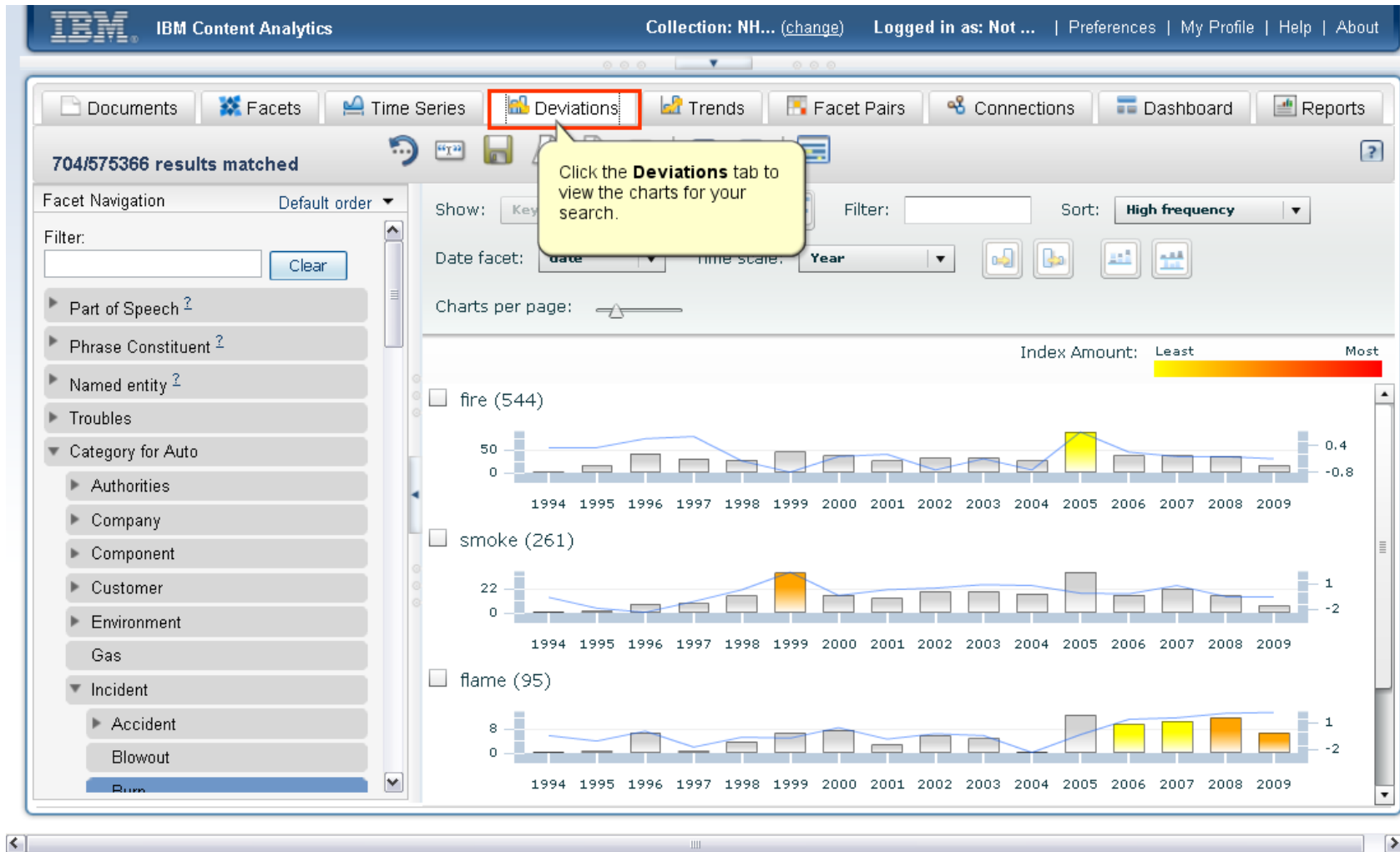
<input checked="" type="checkbox"/>	Keywords	Frequency	Correlation
<input checked="" type="checkbox"/>	fire	544	0.8
<input checked="" type="checkbox"/>	smoke	261	0.9
<input checked="" type="checkbox"/>	flame	95	0.7
<input checked="" type="checkbox"/>	fire hazard	12	0.2

## Explore How Frequencies Deviate from the Expected Average

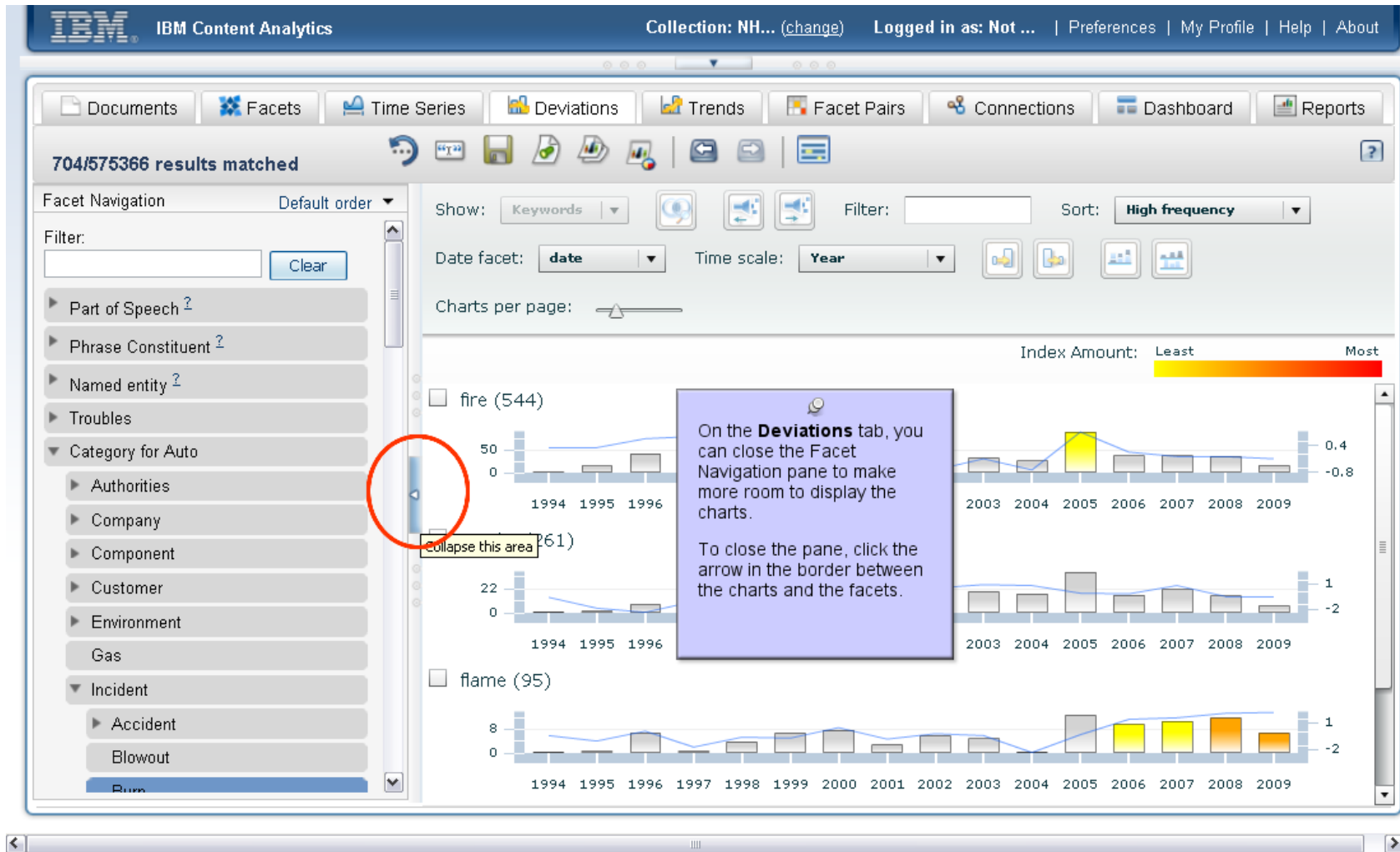
The screenshot displays the IBM Content Analytics interface. At the top, the header includes the IBM logo, 'IBM Content Analytics', and user information: 'Collection: NH... (change)', 'Logged in as: Not ...', and links for 'Preferences', 'My Profile', 'Help', and 'About'. Below the header, there are navigation tabs for 'Documents', 'Facets', 'Trends', 'Facet Pairs', 'Connections', 'Dashboard', and 'Reports'. The main content area shows '16254/575366 results matched' and a 'Facet Navigation' sidebar on the left. The sidebar lists various categories, with 'Incident' expanded to show sub-categories like 'Accident', 'Blowout', 'Burn', 'Deficiency', and 'Electricity'. The 'Burn' category is selected. A callout box points to a button with a magnifying glass icon, containing the text: 'Click Add to search with Boolean AND to add the keywords to your search.' The main table displays search results with columns for 'Keywords', 'Add to search with Boolean AND.', and 'Correlation'. The table contains four rows of data:

<input checked="" type="checkbox"/>	Keywords	Add to search with Boolean AND.	Correlation
<input checked="" type="checkbox"/>	fire	544	0.8
<input checked="" type="checkbox"/>	smoke	261	0.9
<input checked="" type="checkbox"/>	flame	95	0.7
<input checked="" type="checkbox"/>	fire hazard	12	0.2

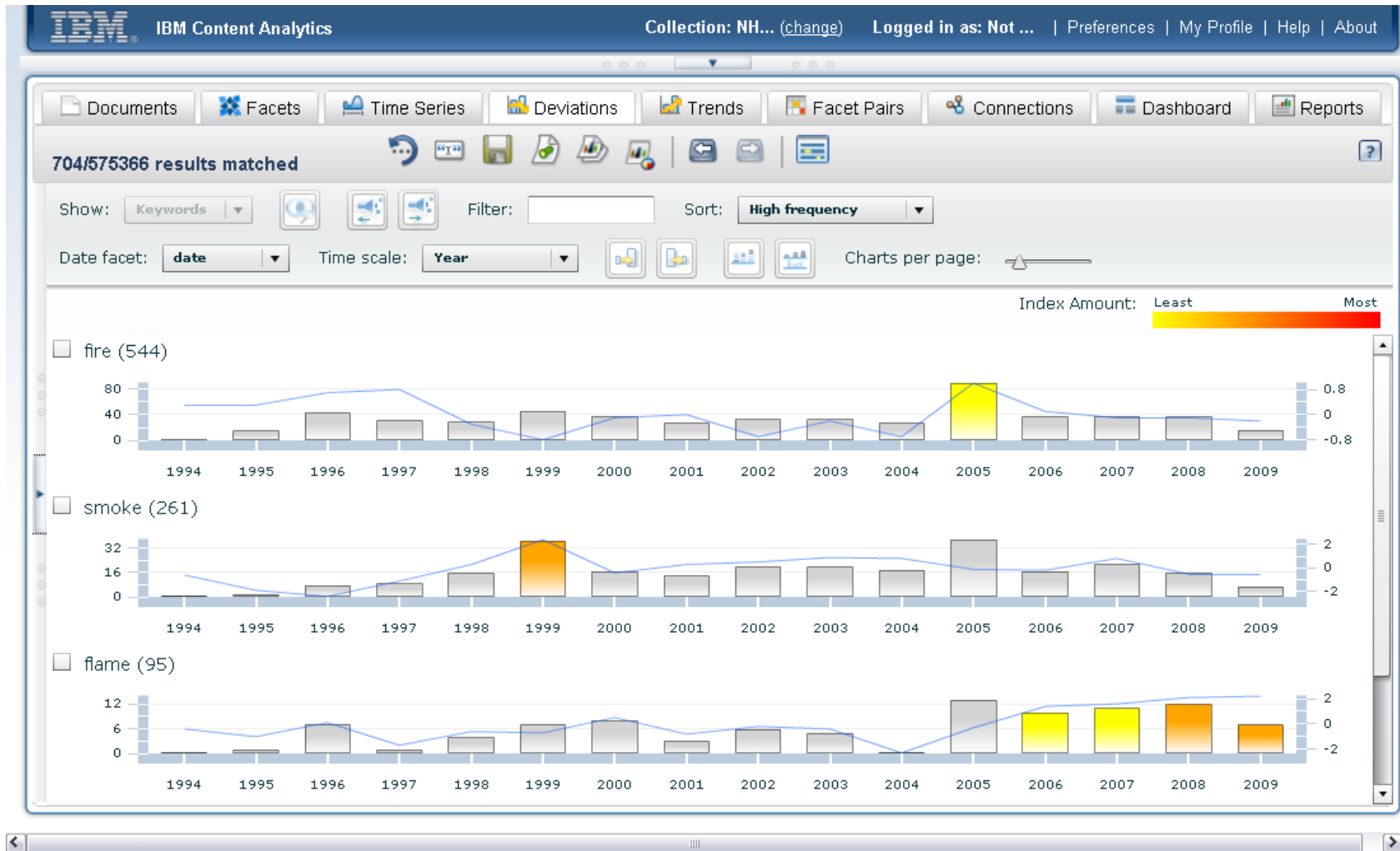
## Explore How Frequencies Deviate from the Expected Average



## Explore How Frequencies Deviate from the Expected Average

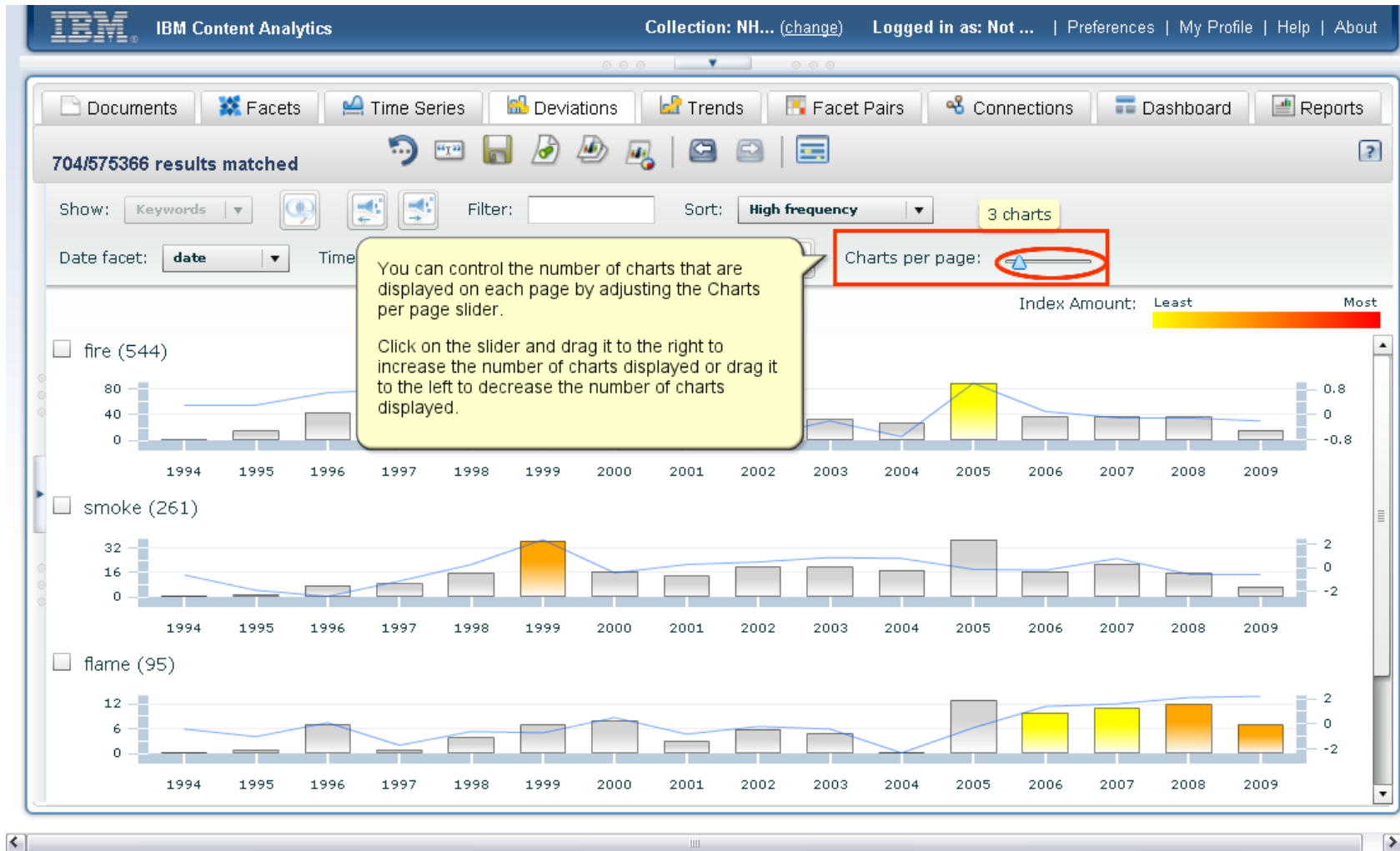


## Explore How Frequencies Deviate from the Expected Average

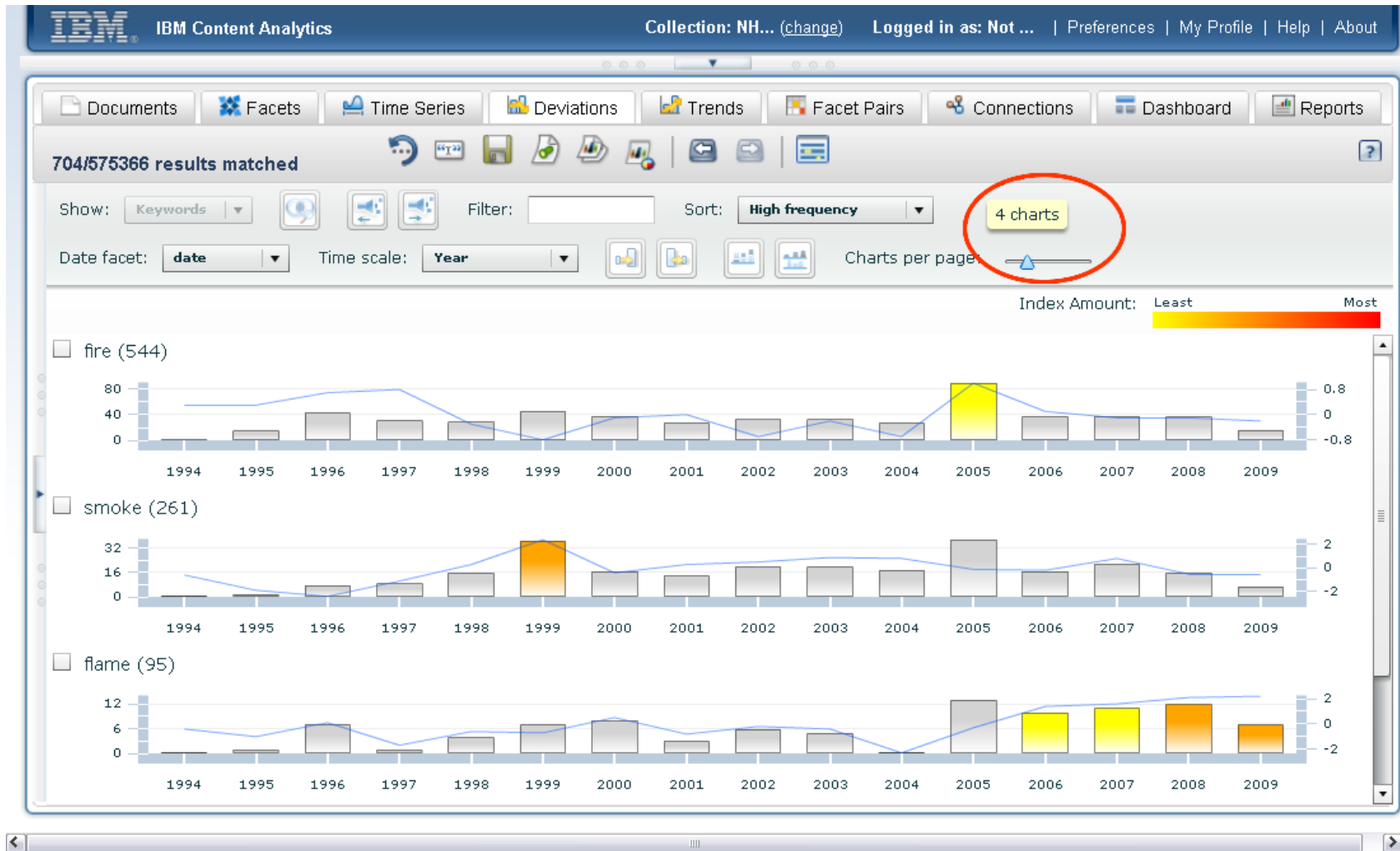




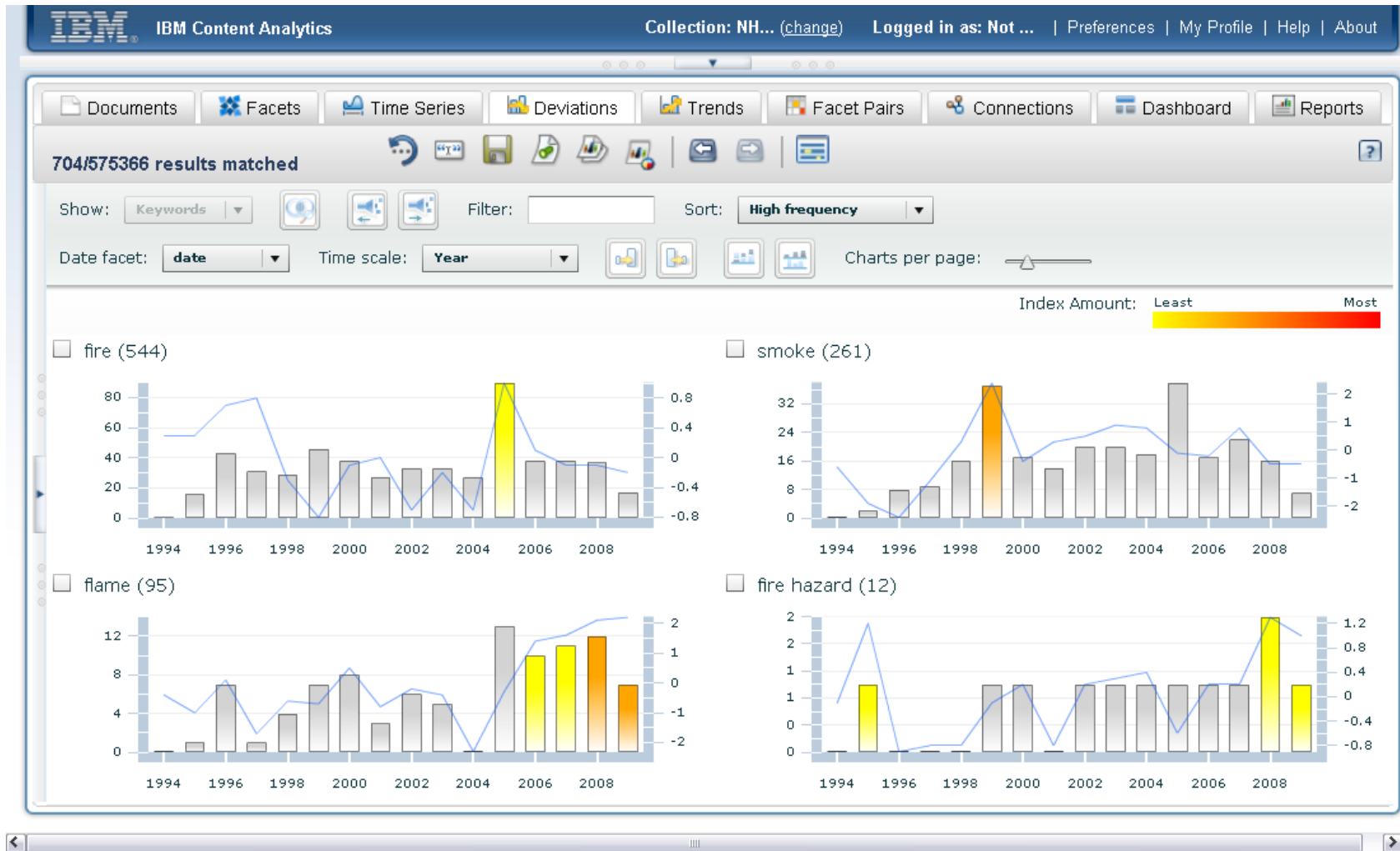
## Explore How Frequencies Deviate from the Expected Average



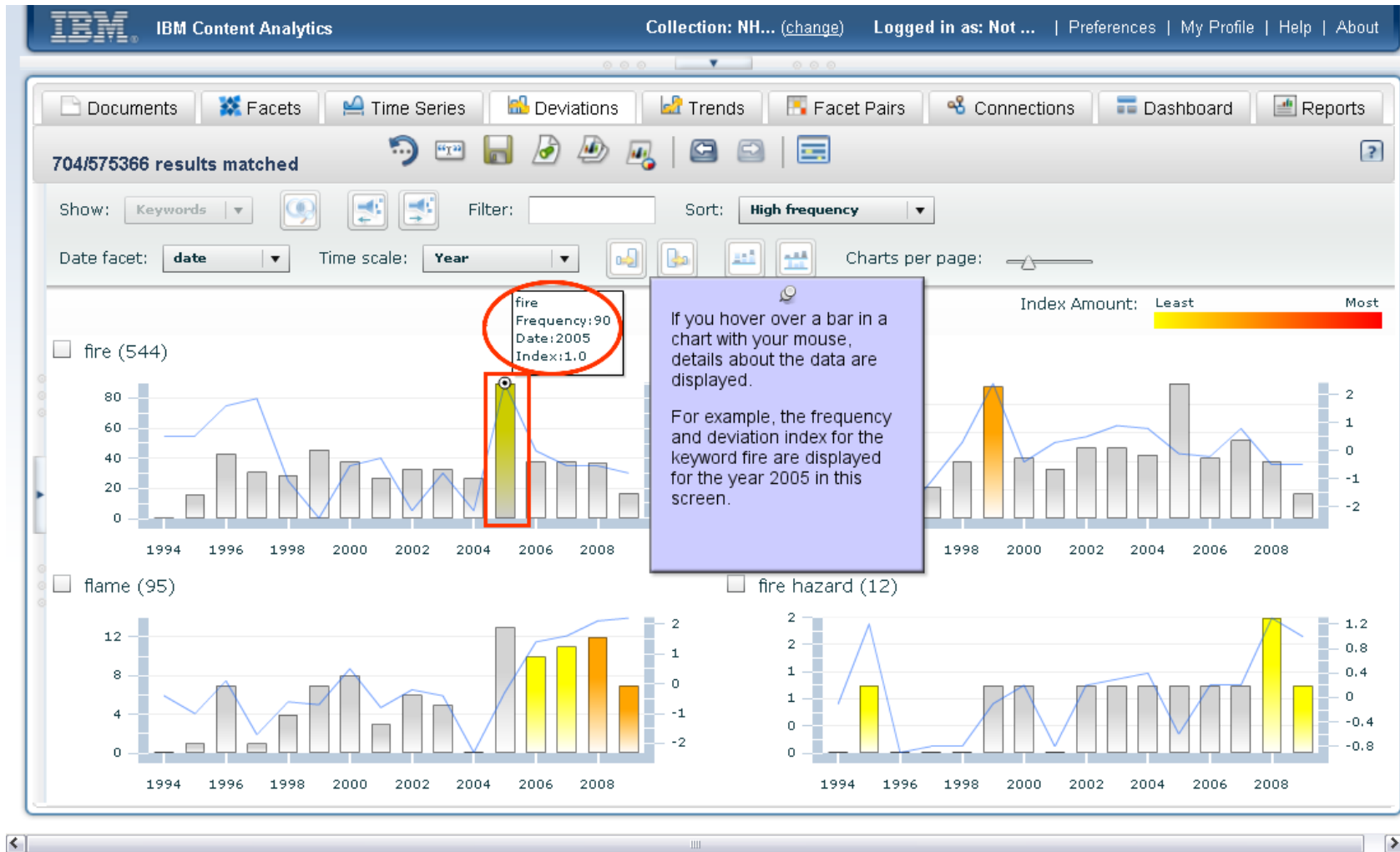
## Explore How Frequencies Deviate from the Expected Average



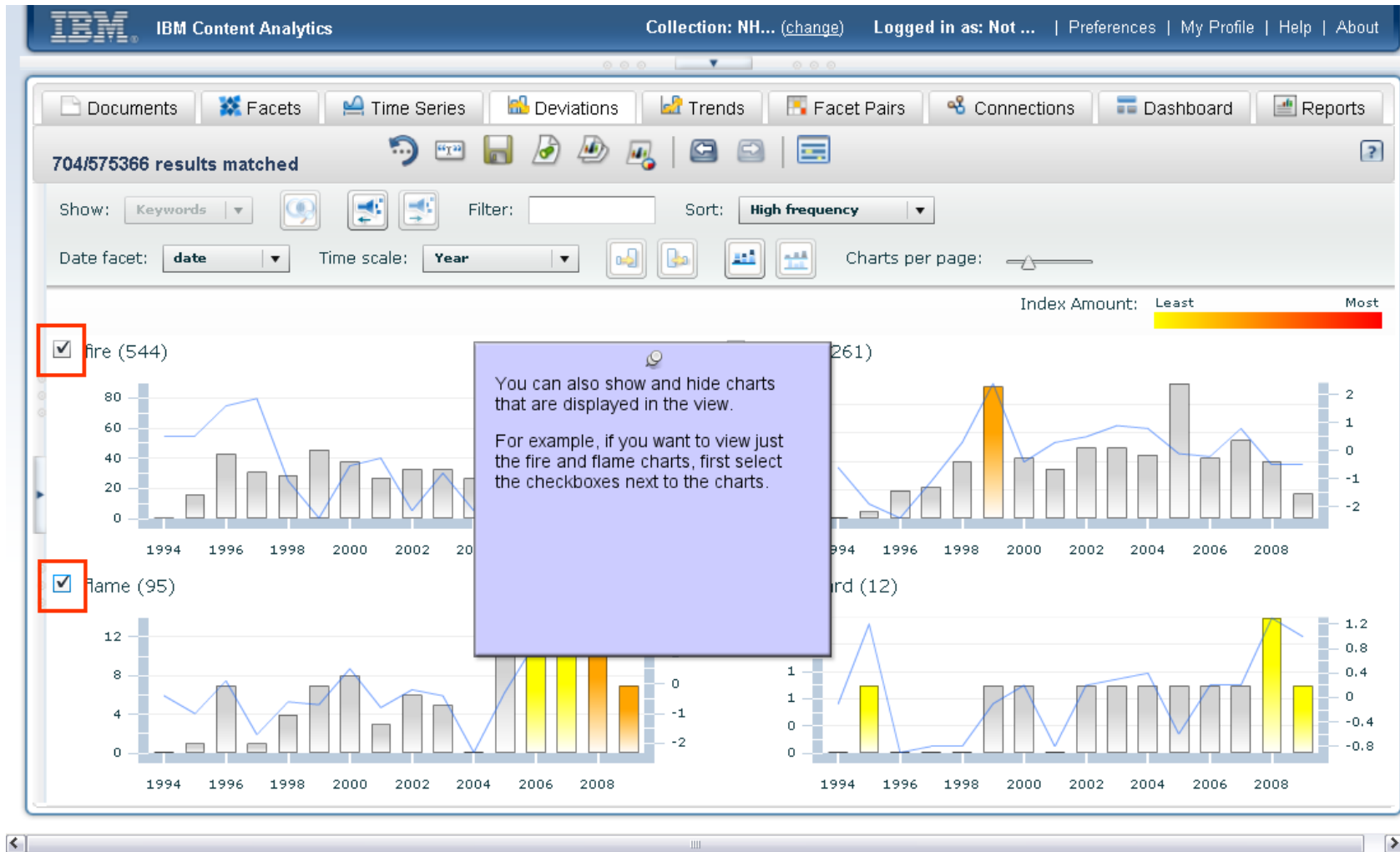
## Explore How Frequencies Deviate from the Expected Average



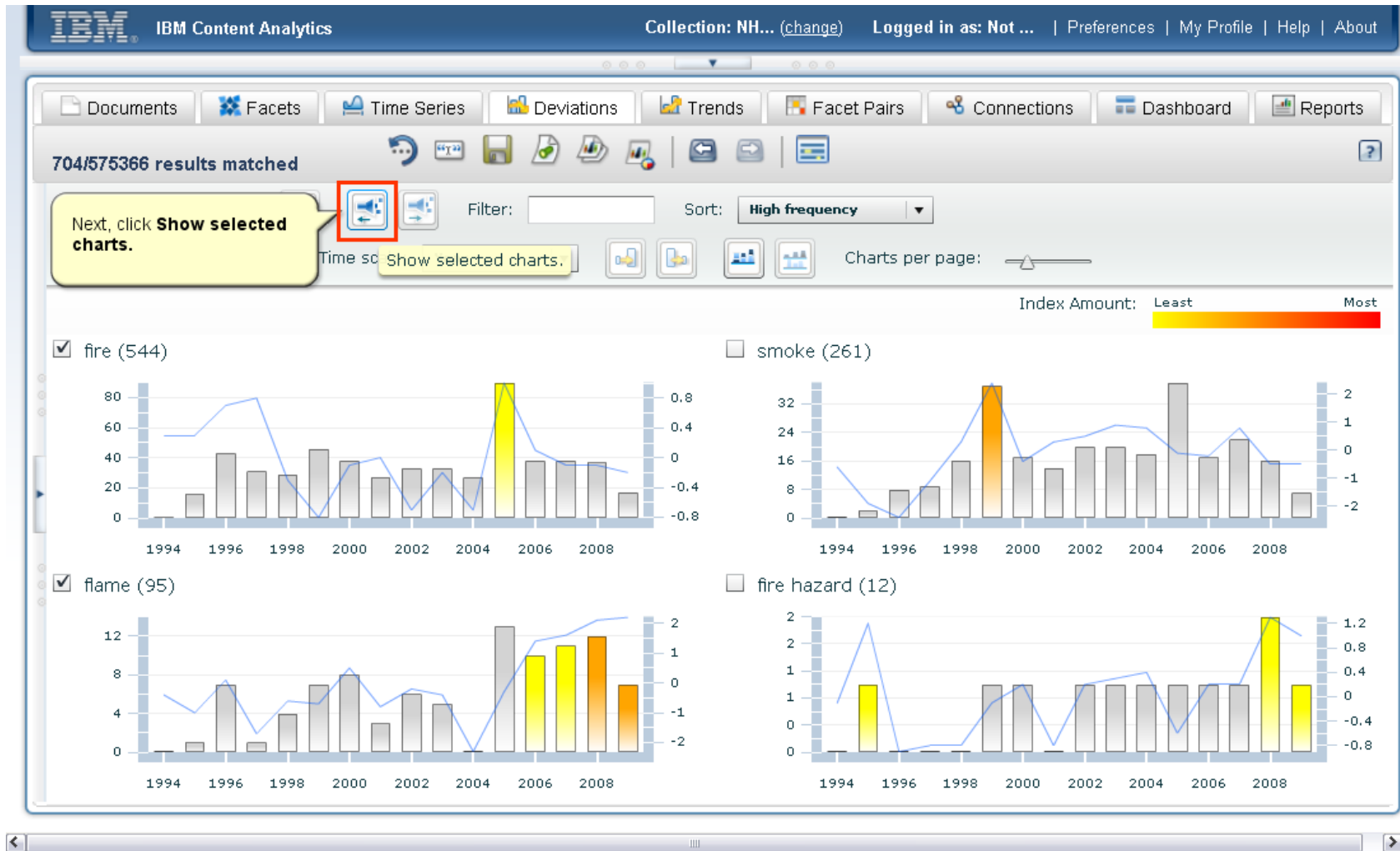
## Explore How Frequencies Deviate from the Expected Average



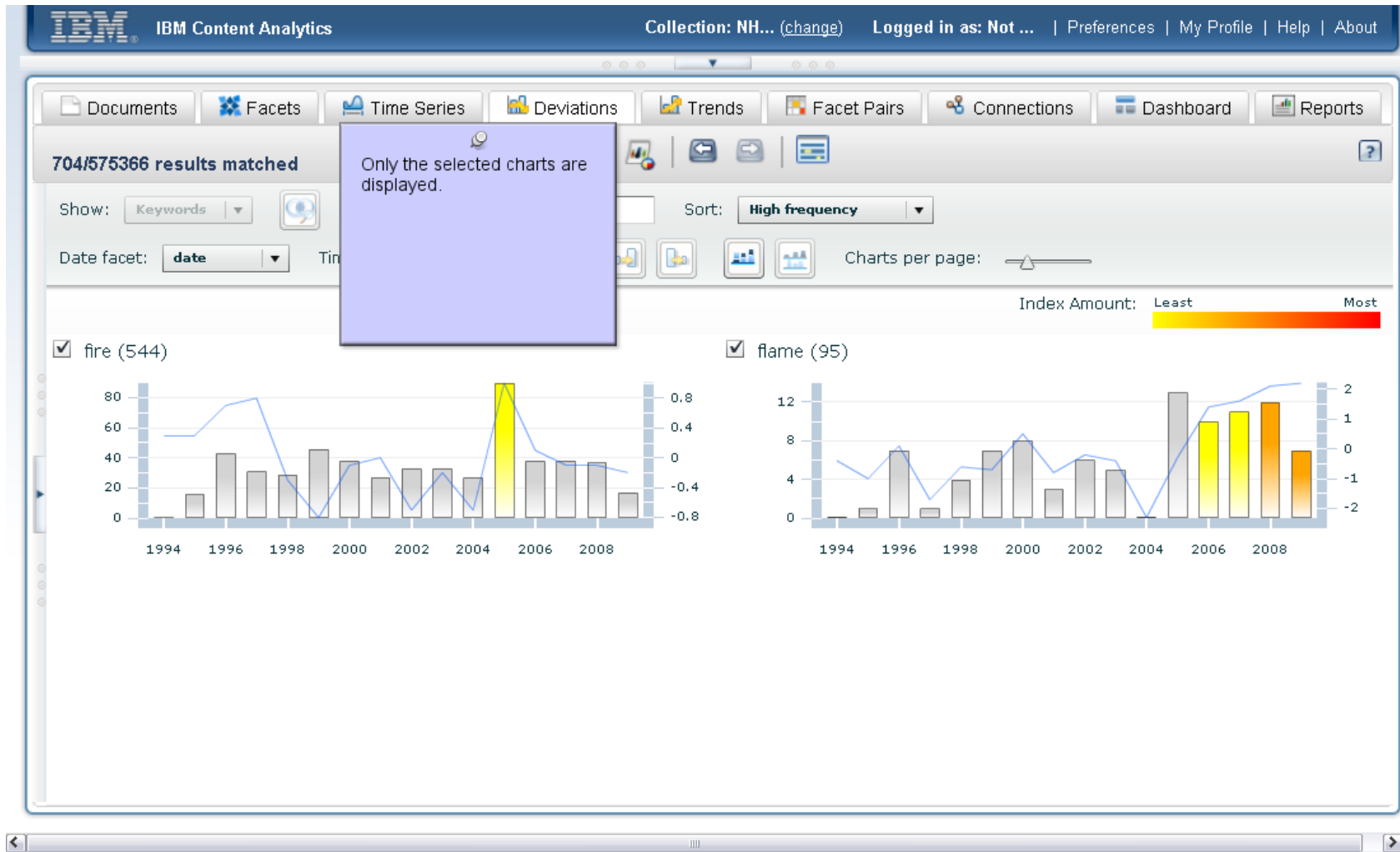
## Explore How Frequencies Deviate from the Expected Average



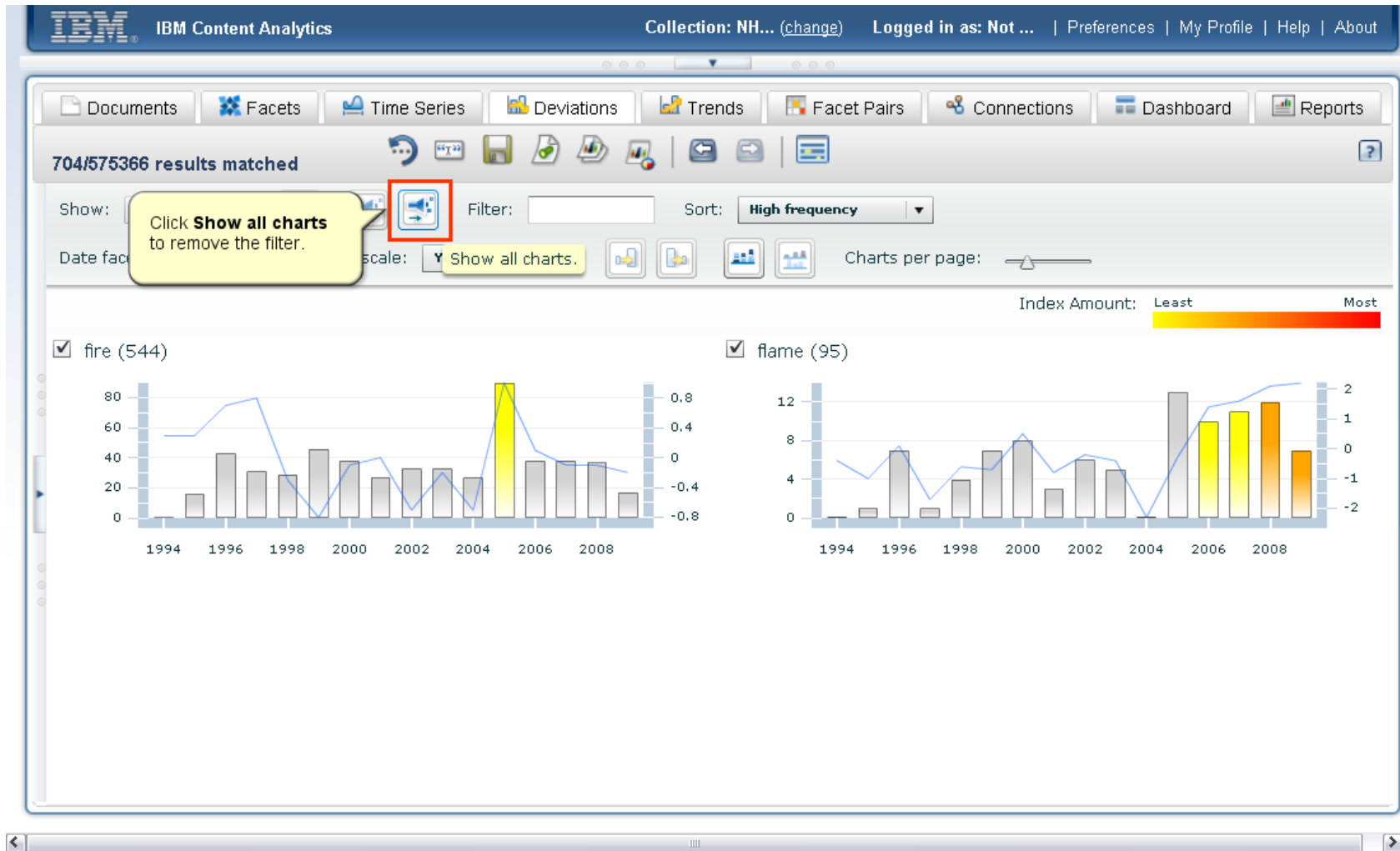
## Explore How Frequencies Deviate from the Expected Average



# Explore How Frequencies Deviate from the Expected Average

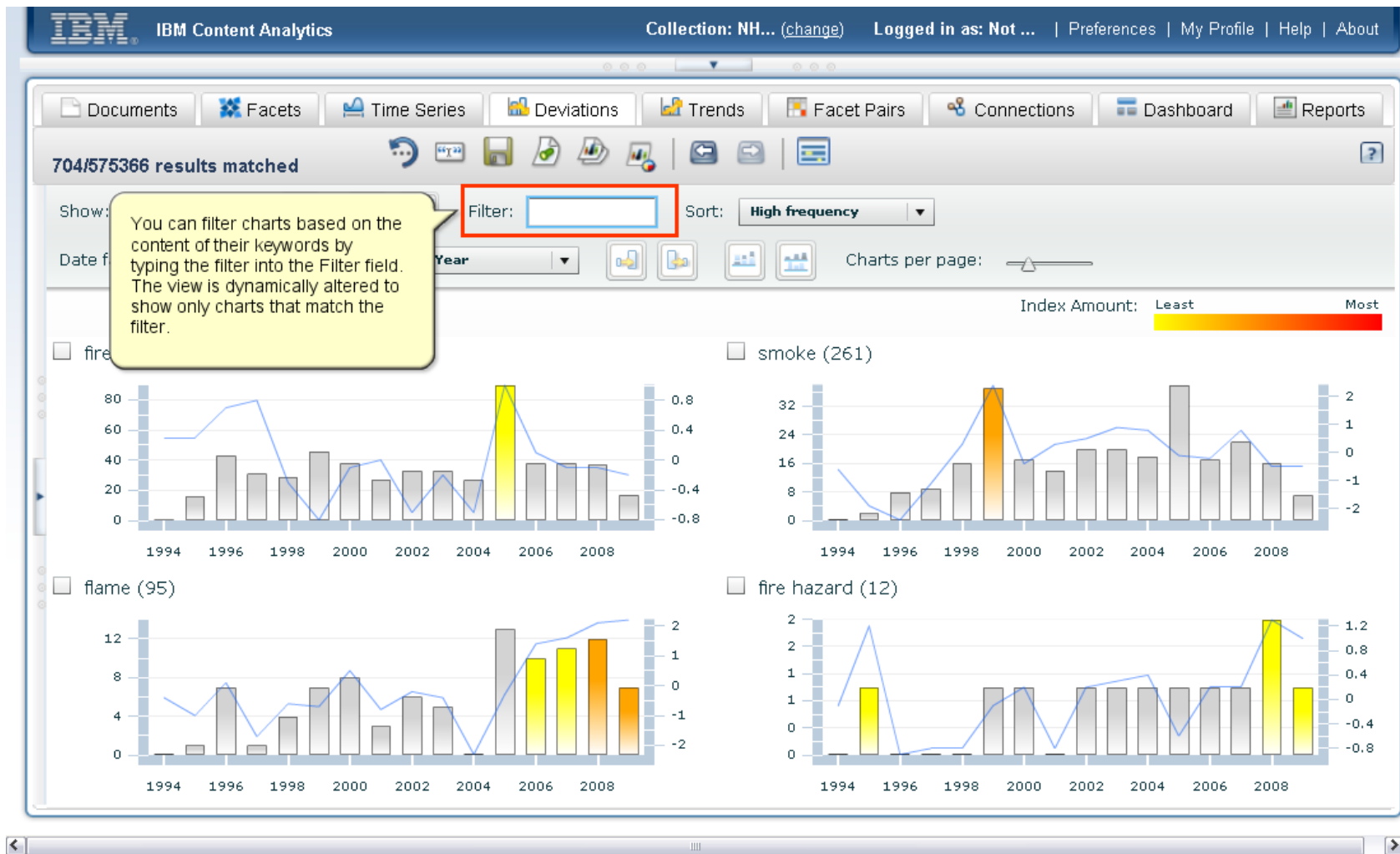


## Explore How Frequencies Deviate from the Expected Average

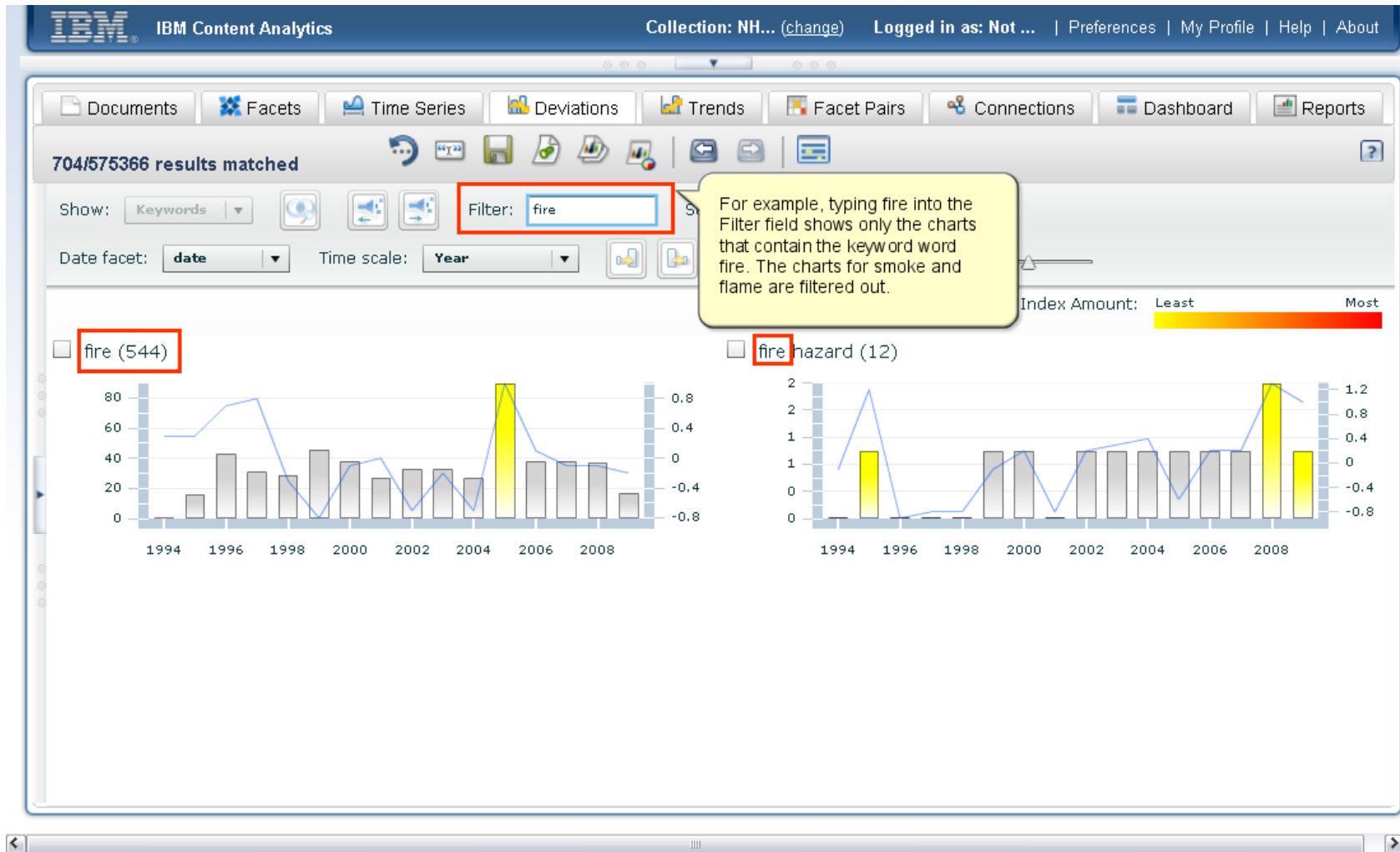




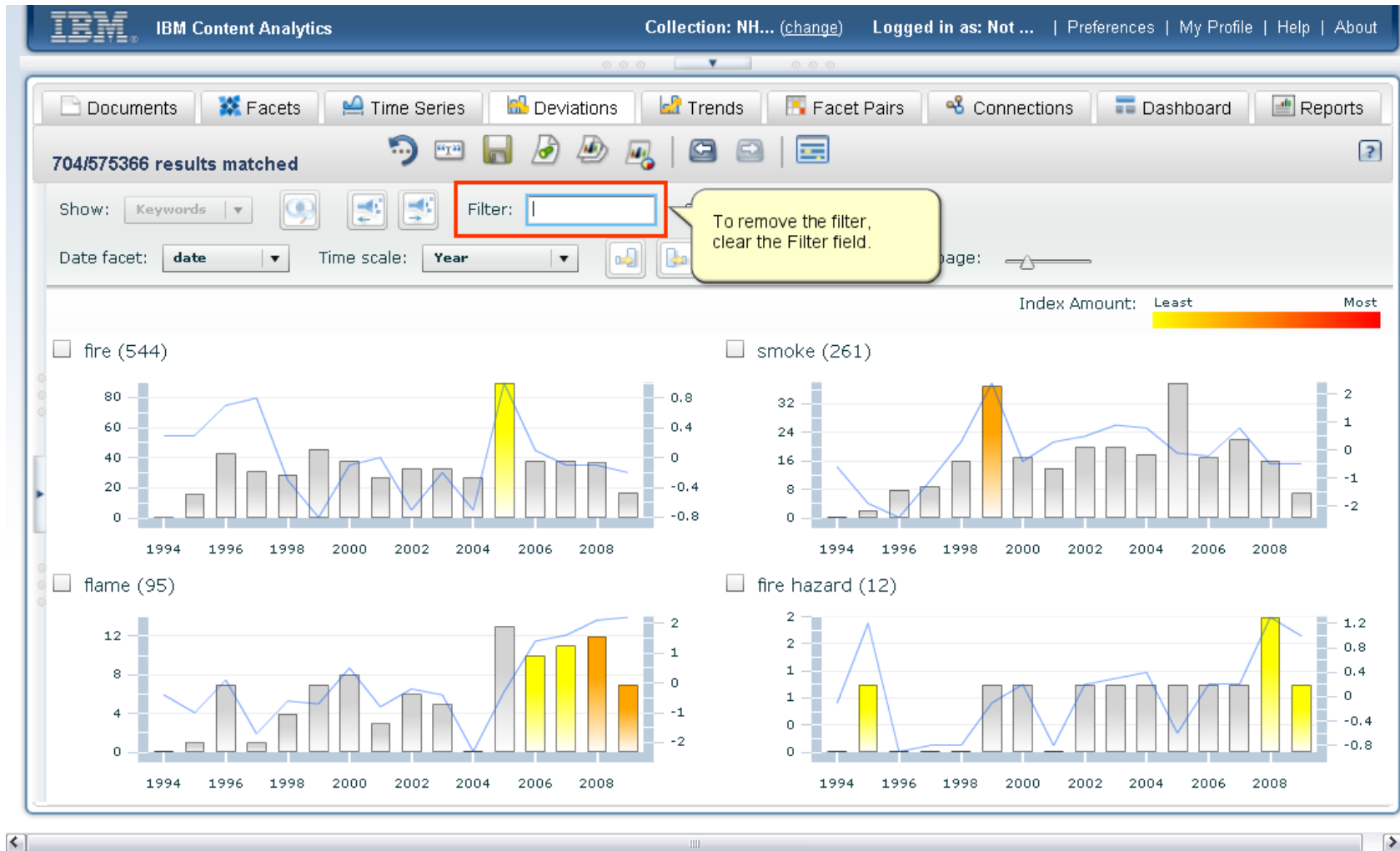
## Explore How Frequencies Deviate from the Expected Average



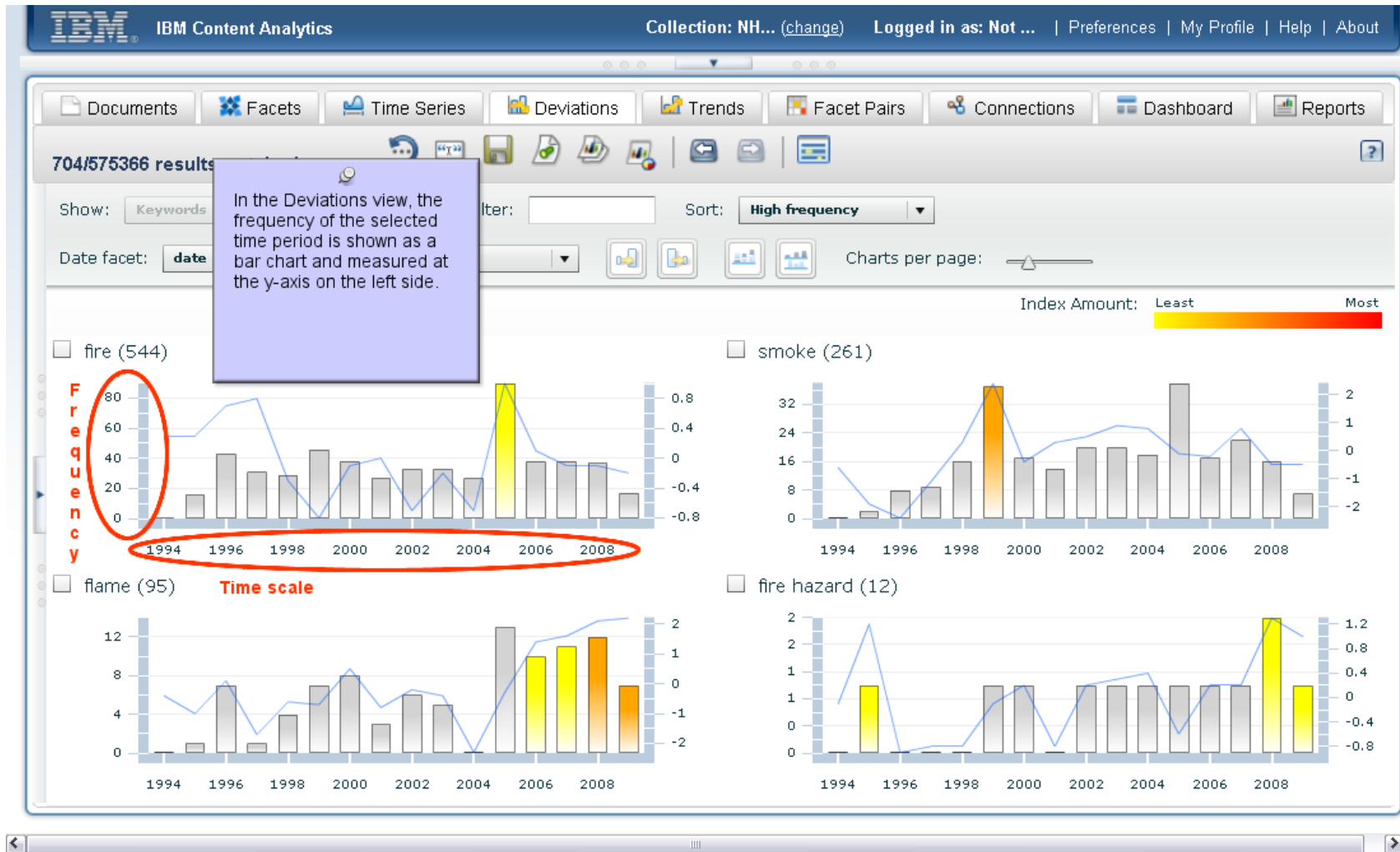
## Explore How Frequencies Deviate from the Expected Average



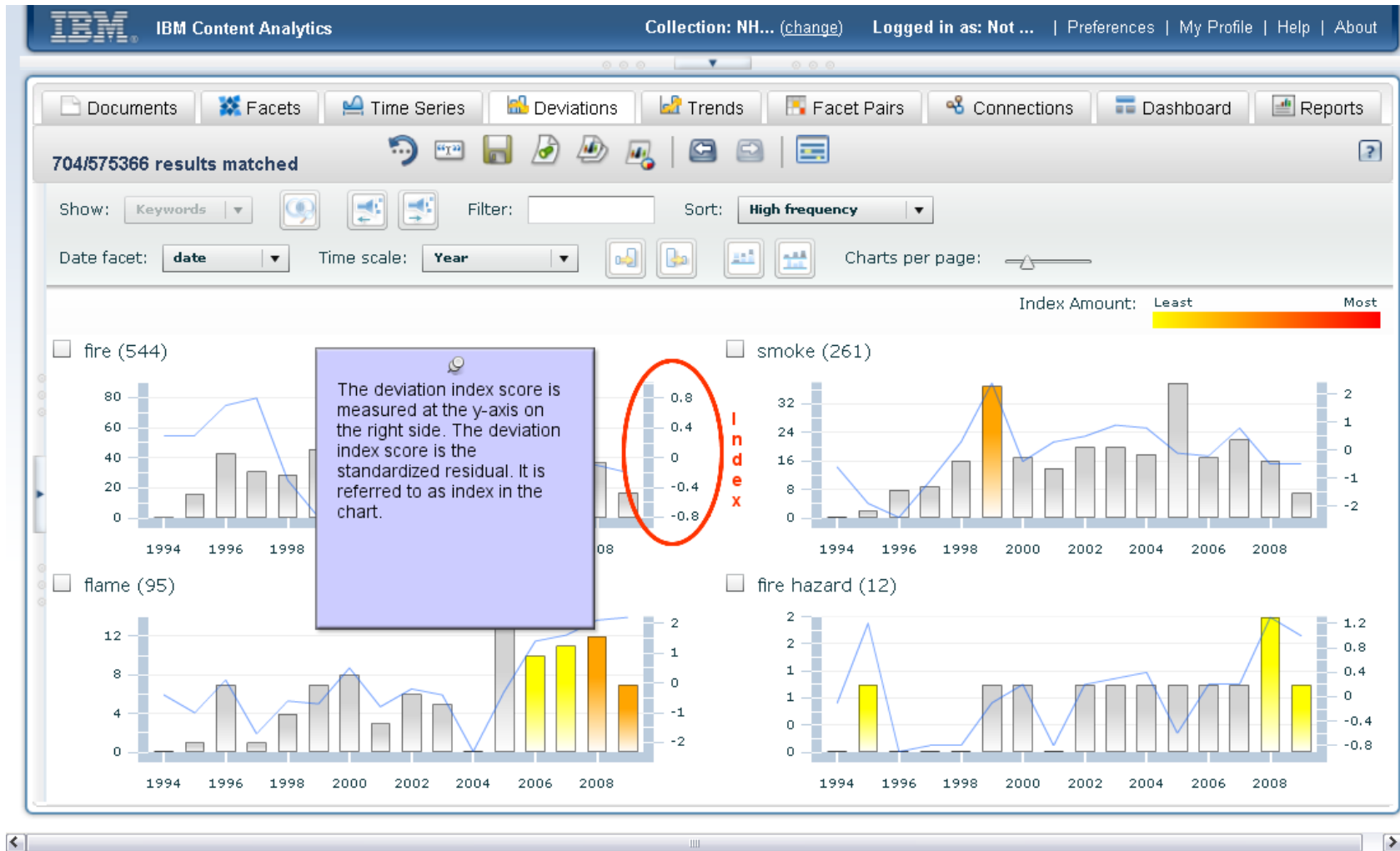
## Explore How Frequencies Deviate from the Expected Average



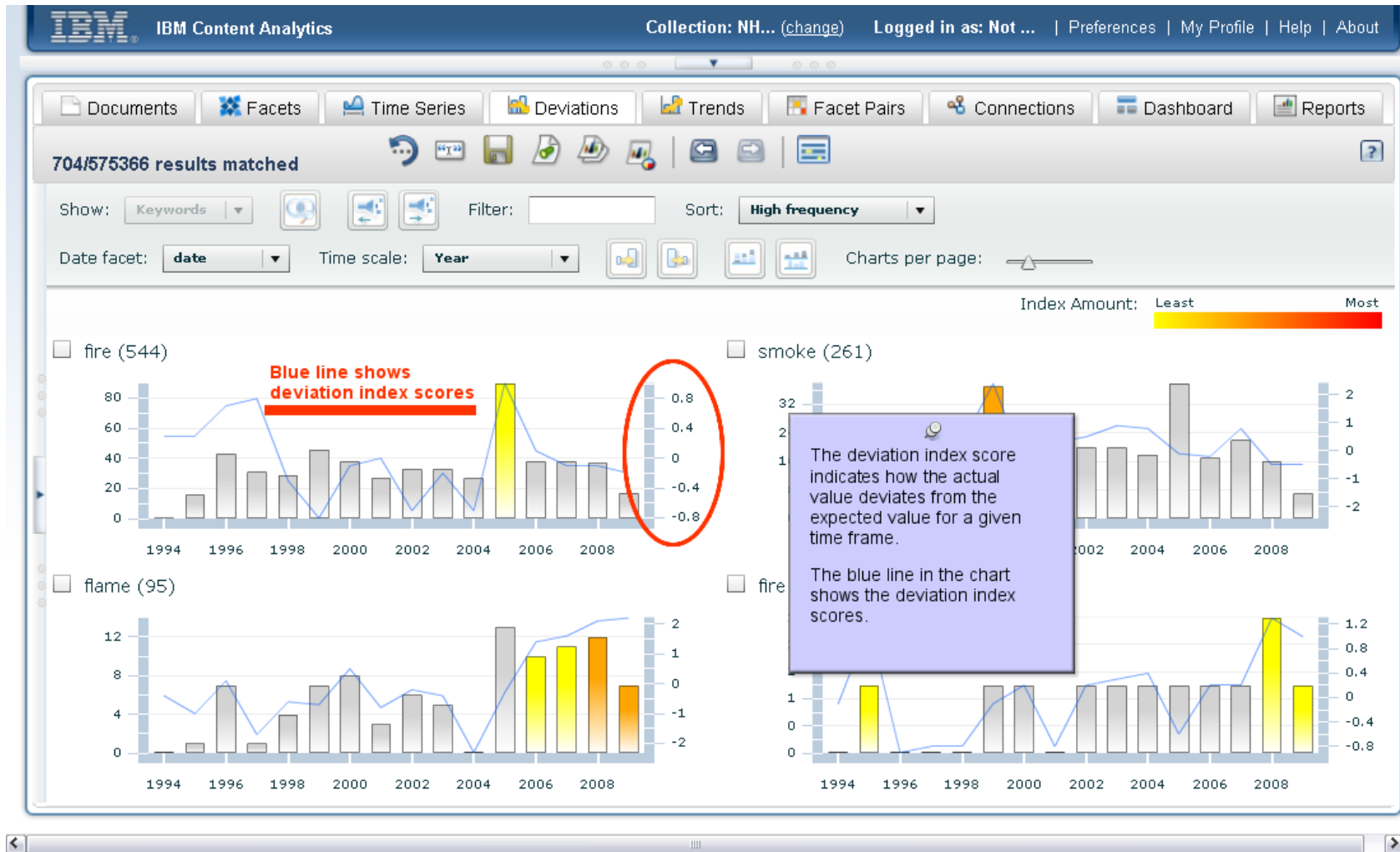
## Explore How Frequencies Deviate from the Expected Average



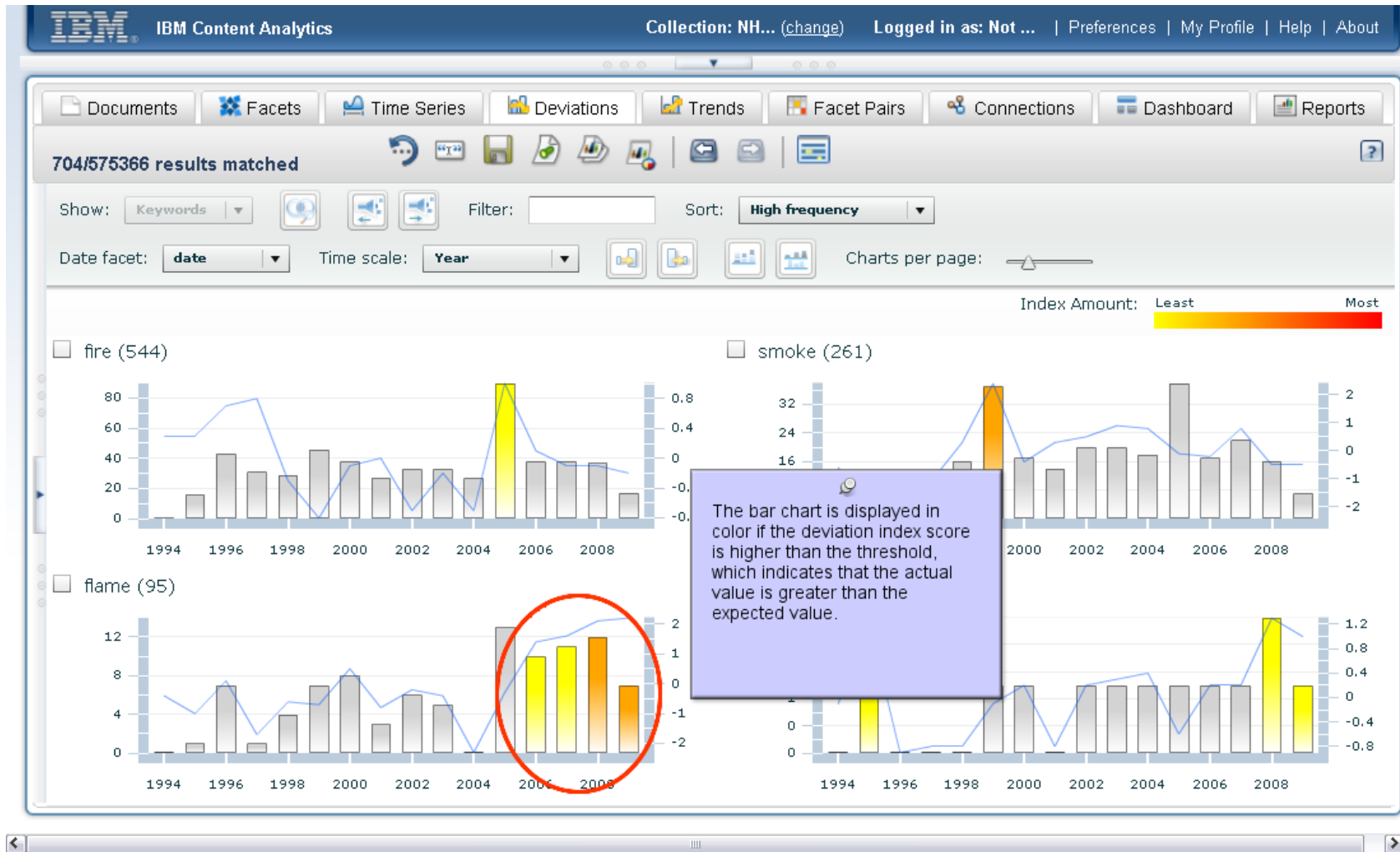
# Explore How Frequencies Deviate from the Expected Average



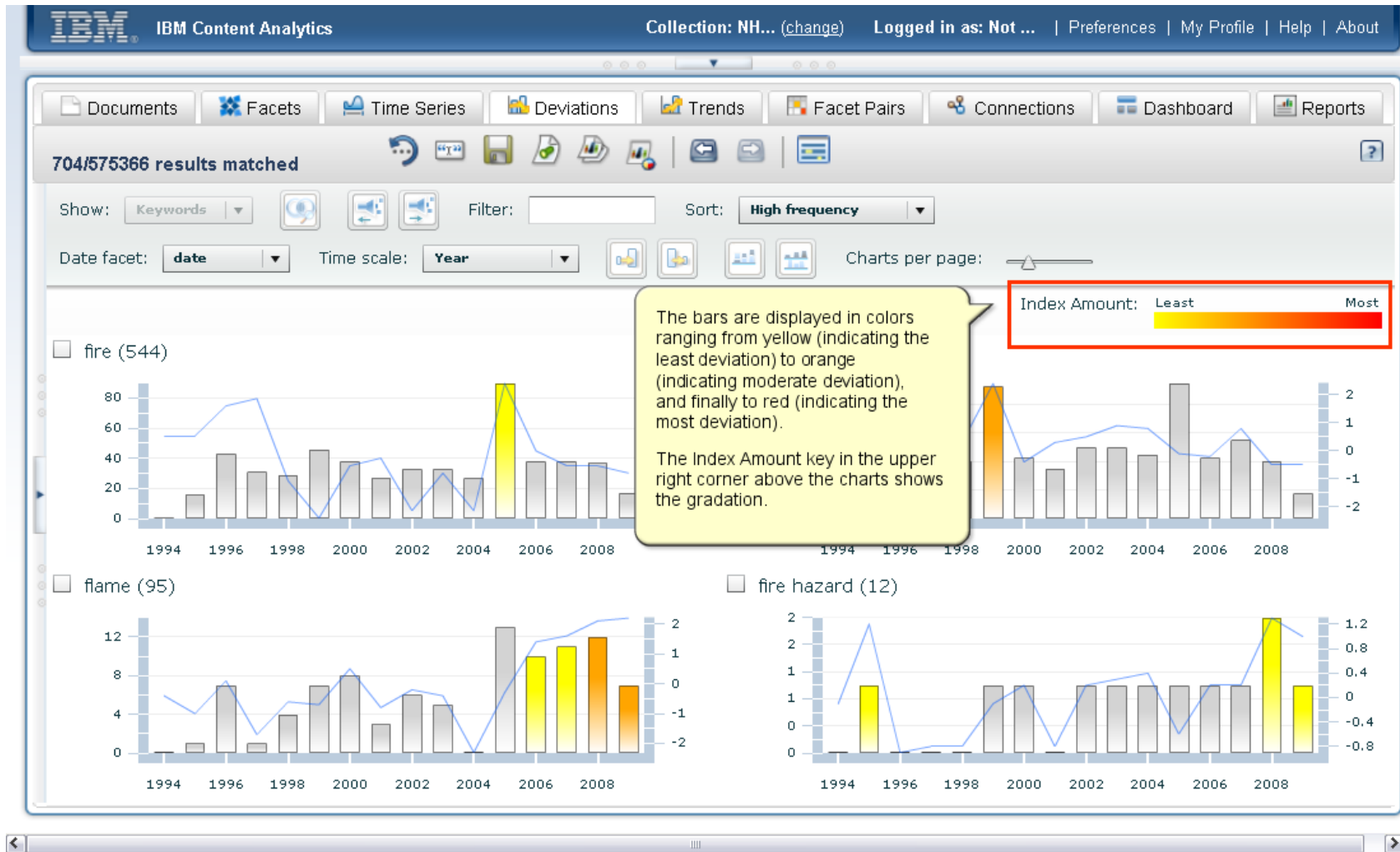
## Explore How Frequencies Deviate from the Expected Average



## Explore How Frequencies Deviate from the Expected Average

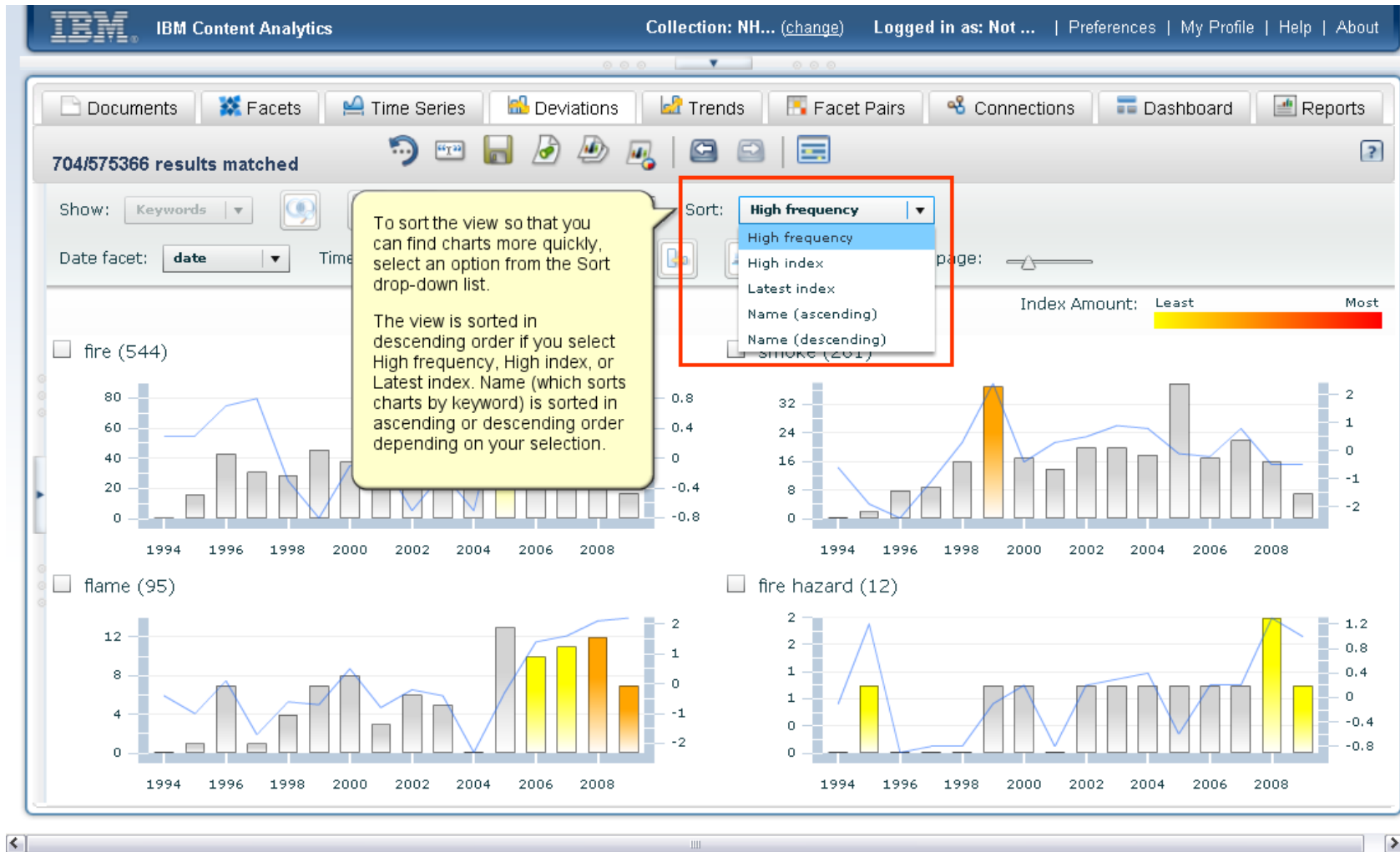


## Explore How Frequencies Deviate from the Expected Average

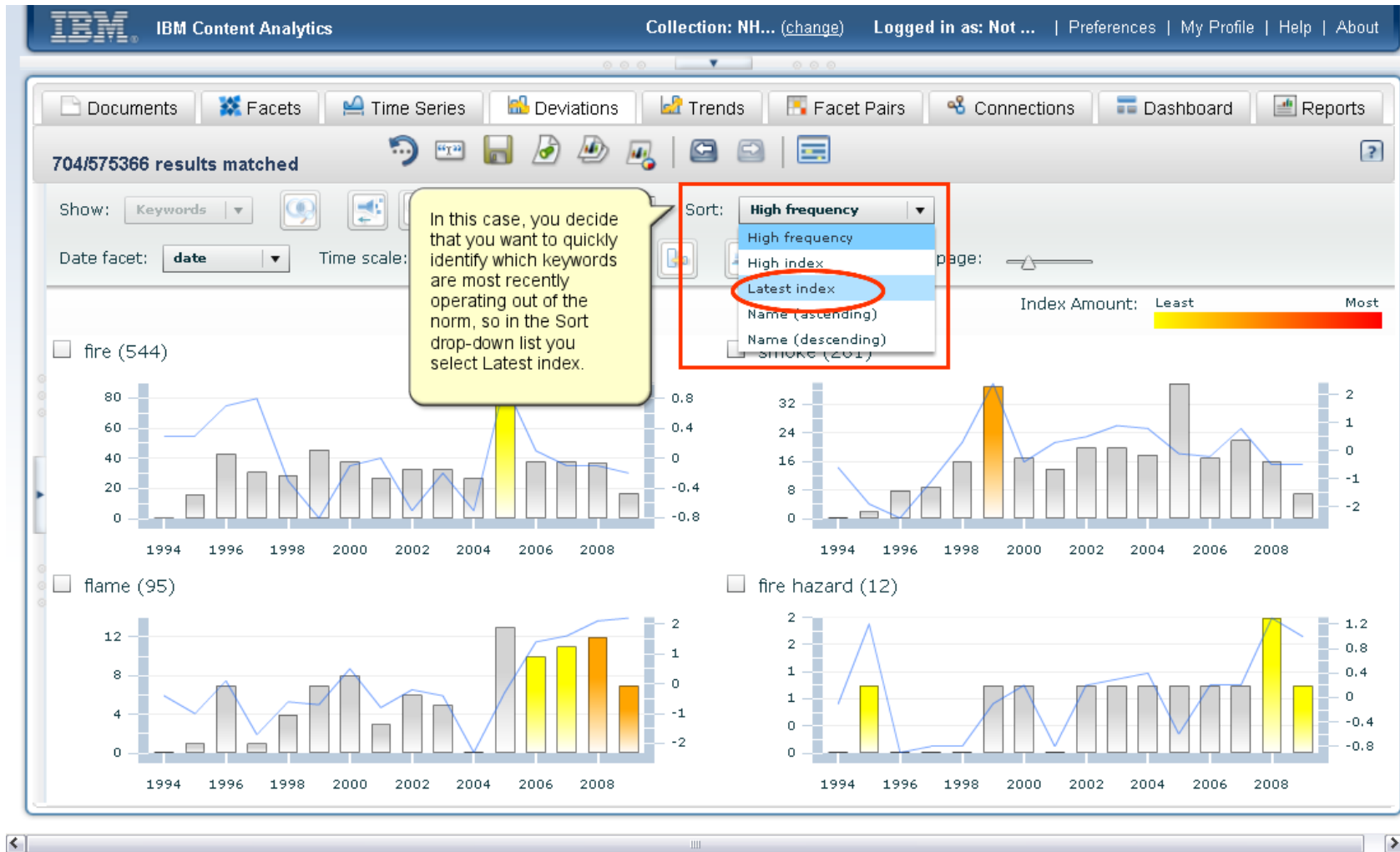




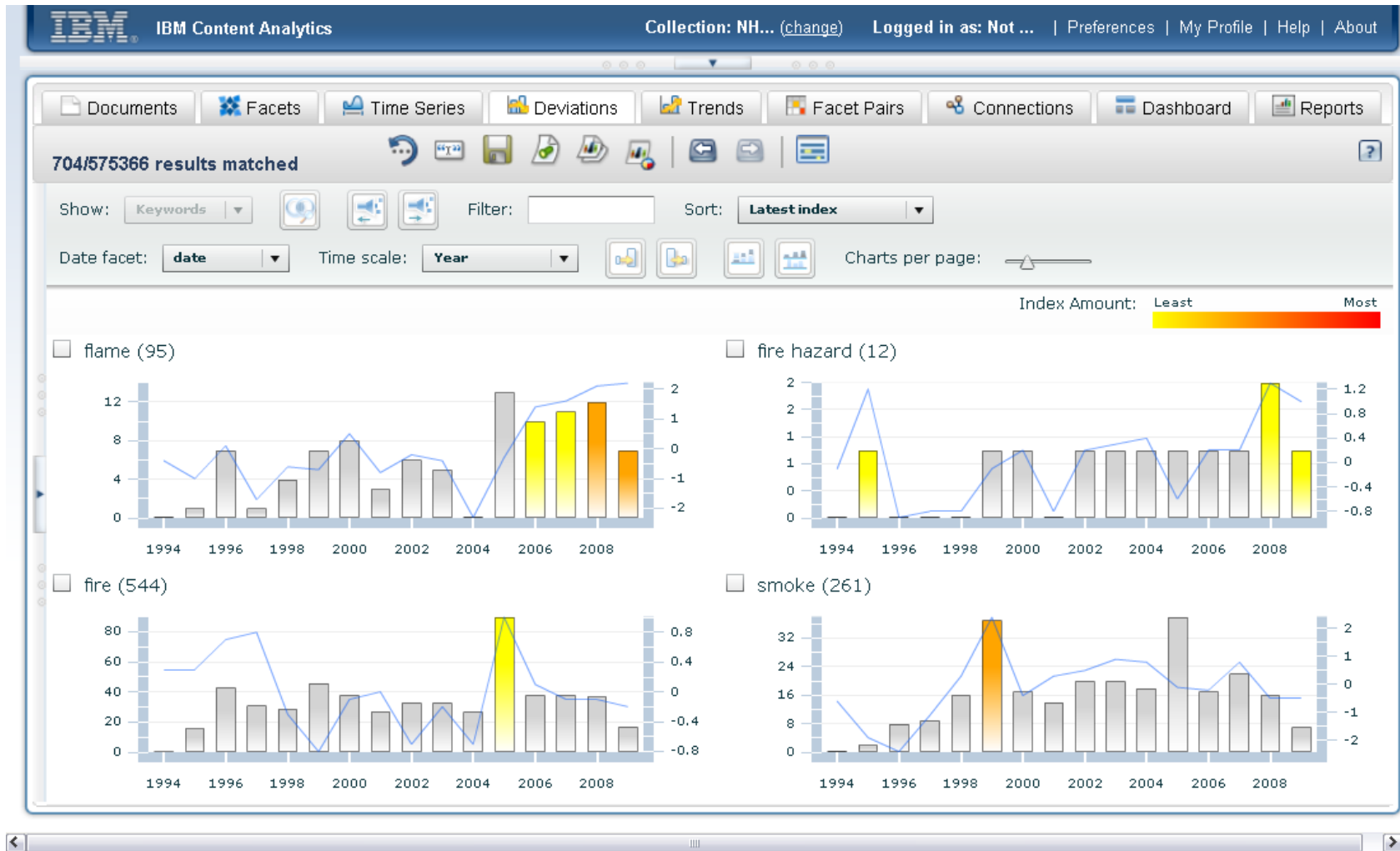
## Explore How Frequencies Deviate from the Expected Average



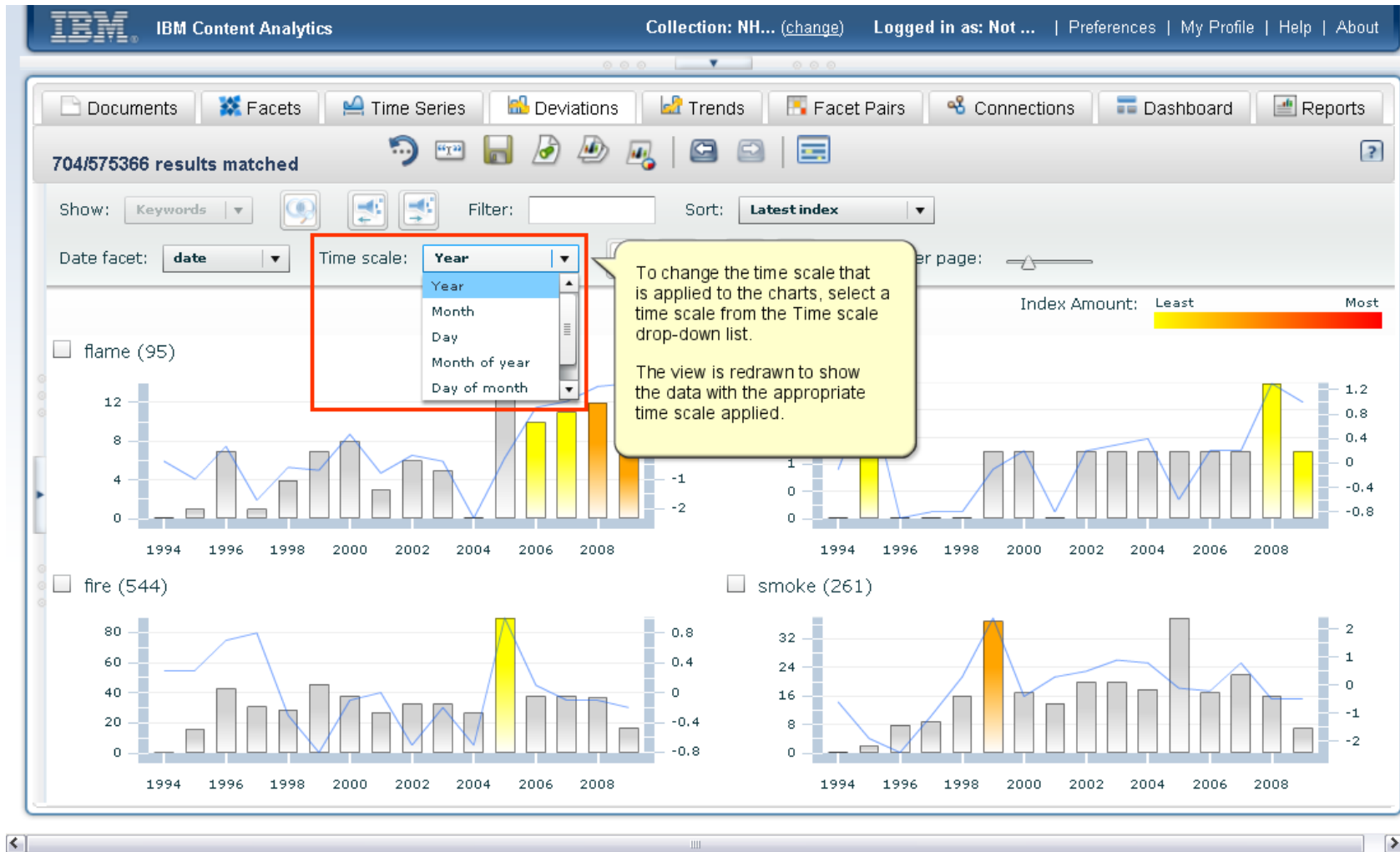
## Explore How Frequencies Deviate from the Expected Average



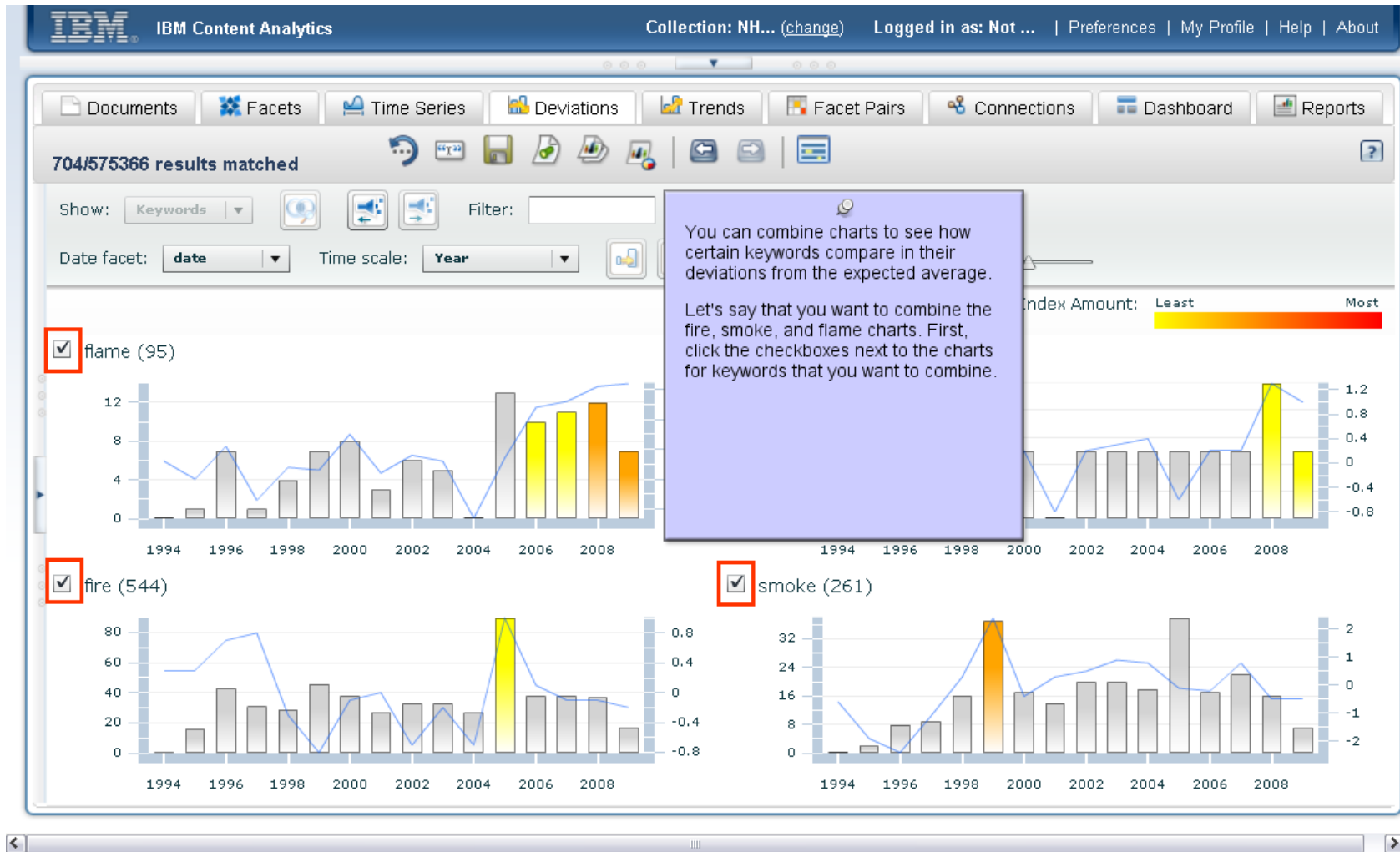
## Explore How Frequencies Deviate from the Expected Average



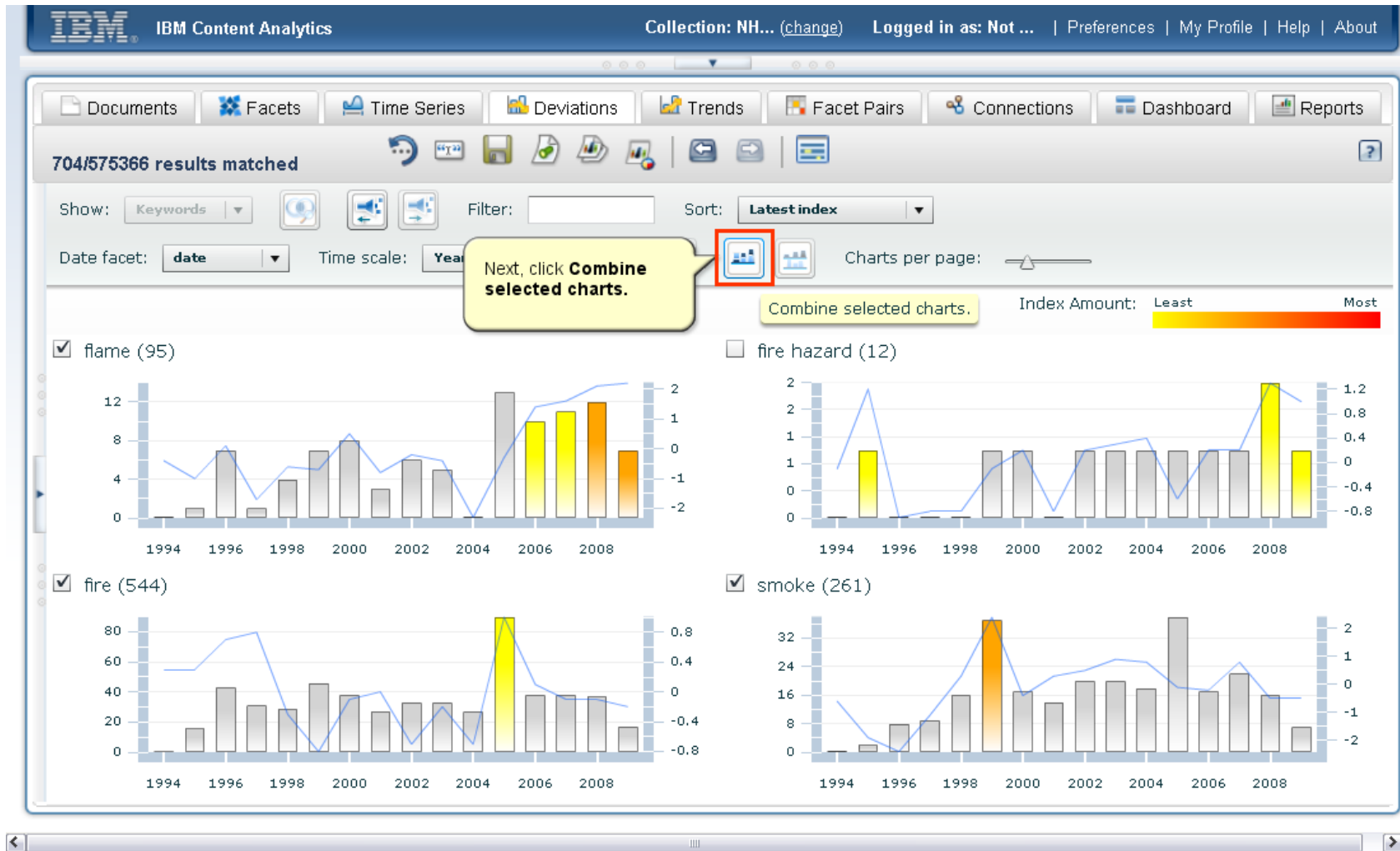
## Explore How Frequencies Deviate from the Expected Average



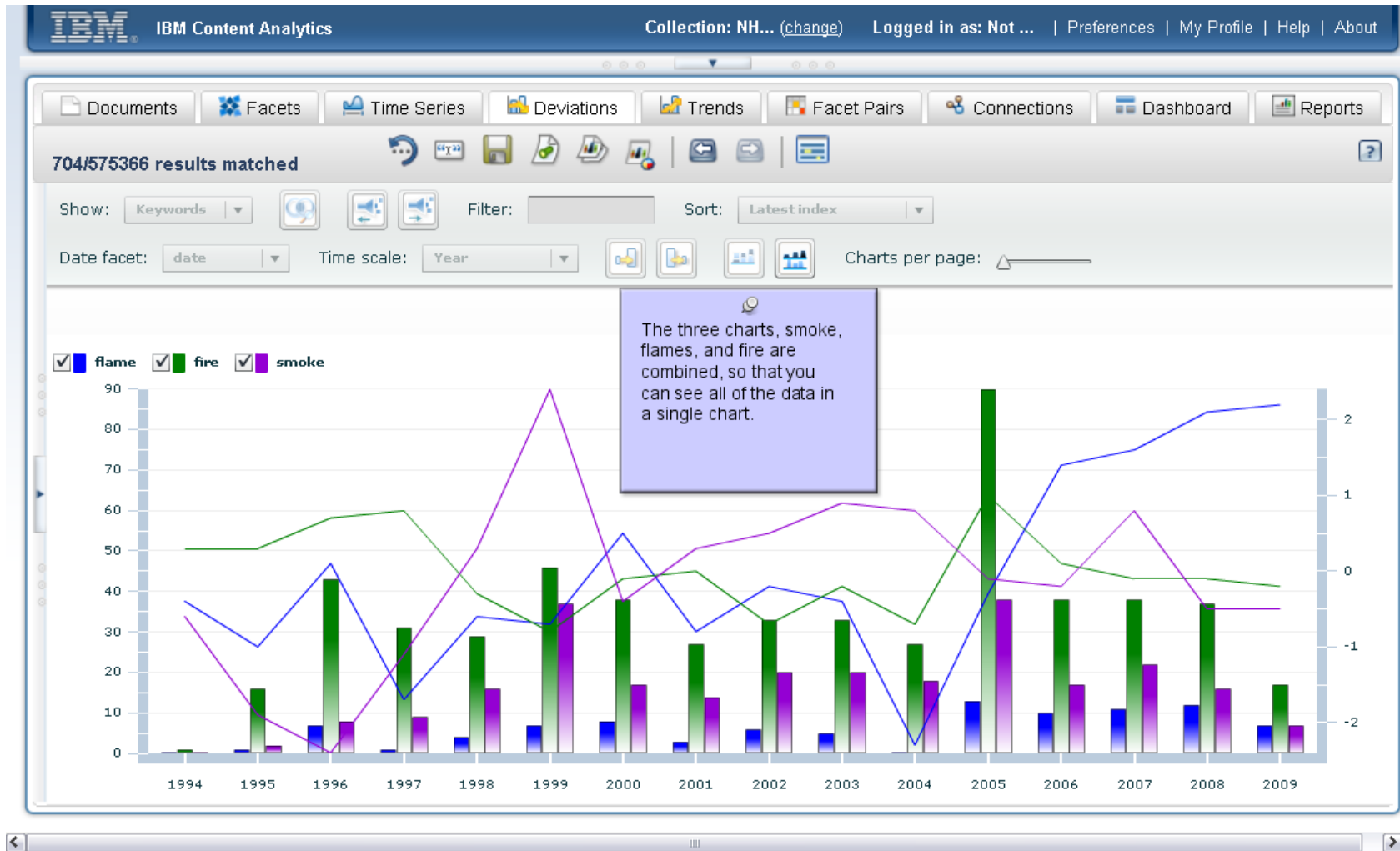
## Explore How Frequencies Deviate from the Expected Average



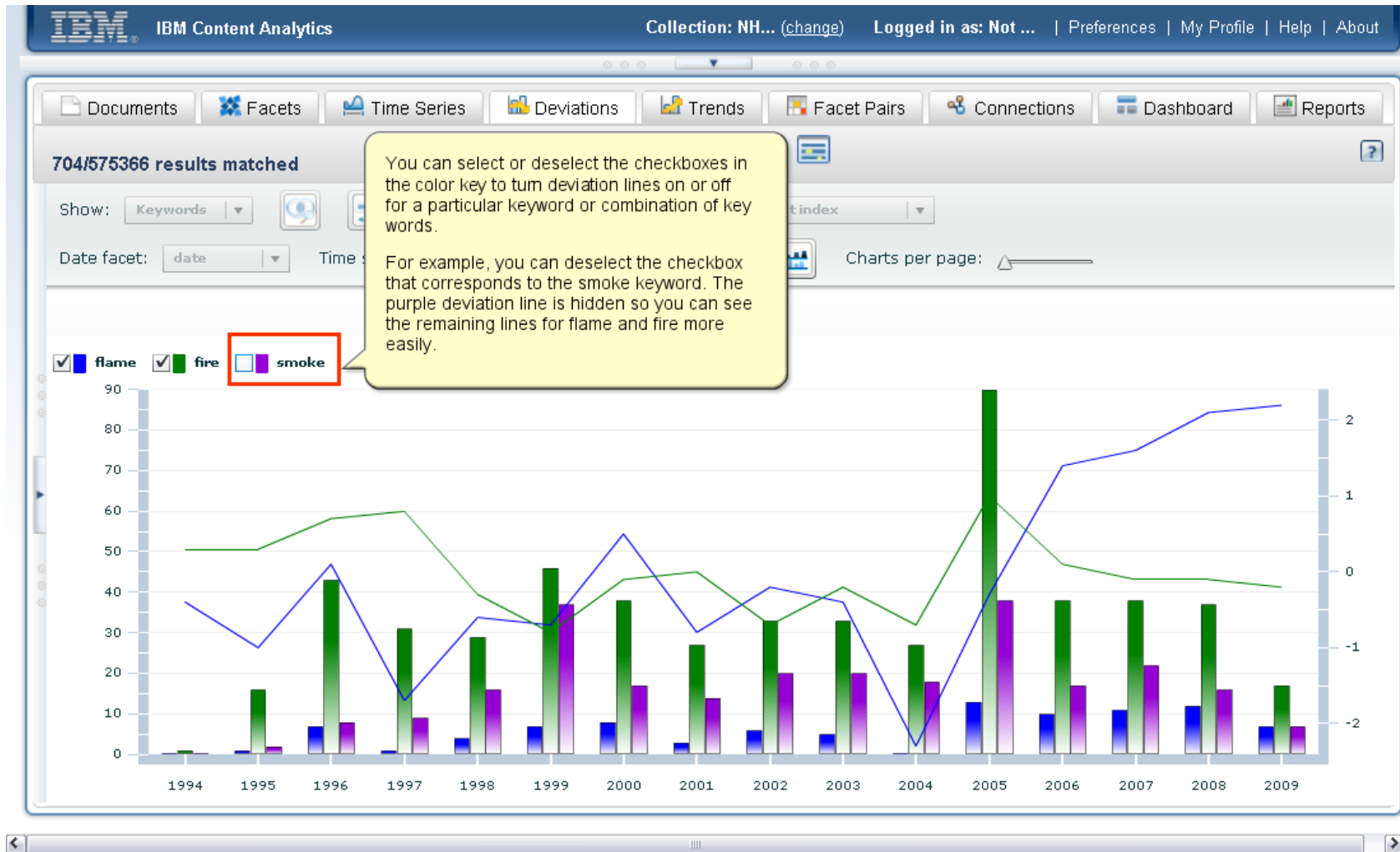
## Explore How Frequencies Deviate from the Expected Average



## Explore How Frequencies Deviate from the Expected Average

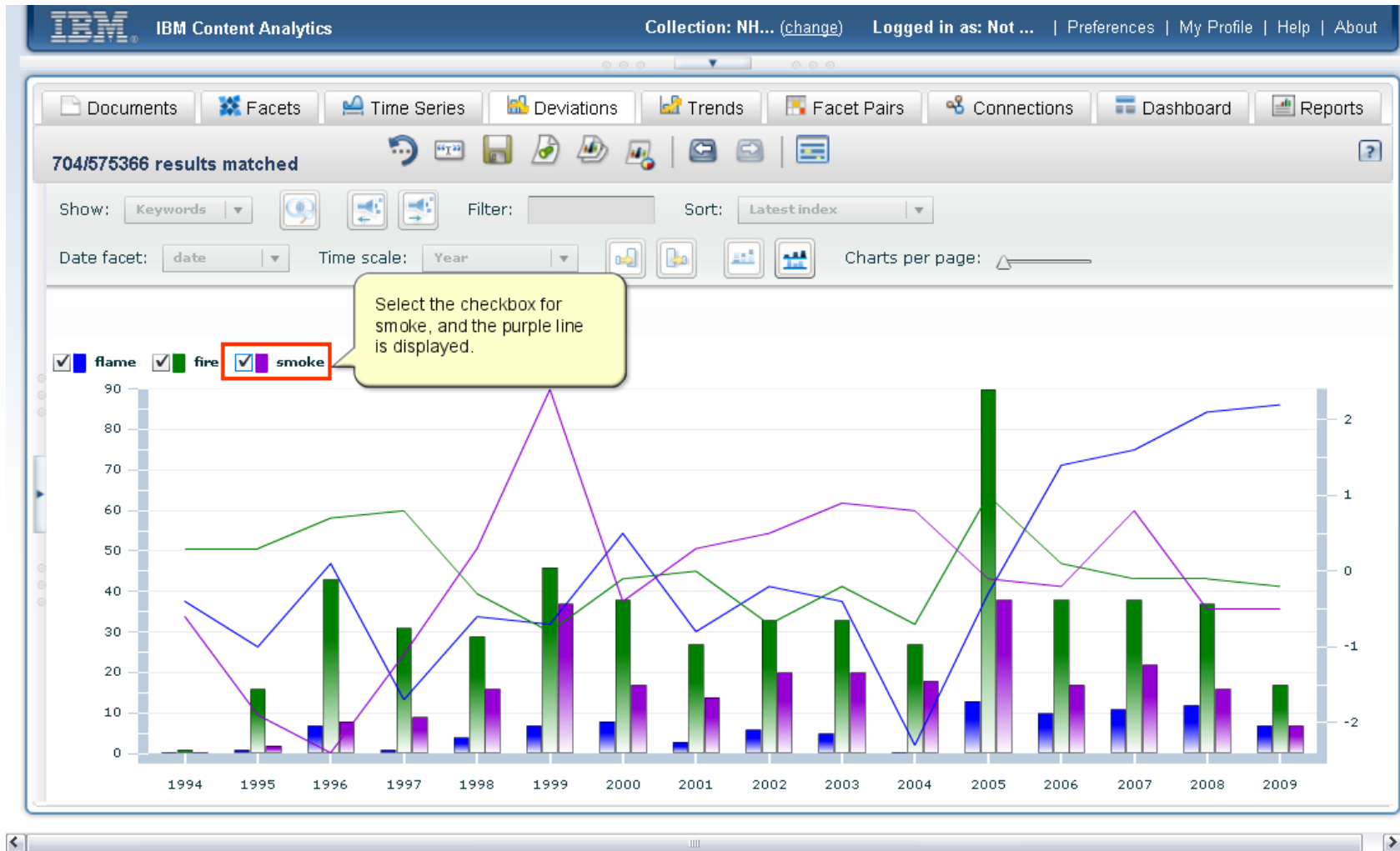


## Explore How Frequencies Deviate from the Expected Average

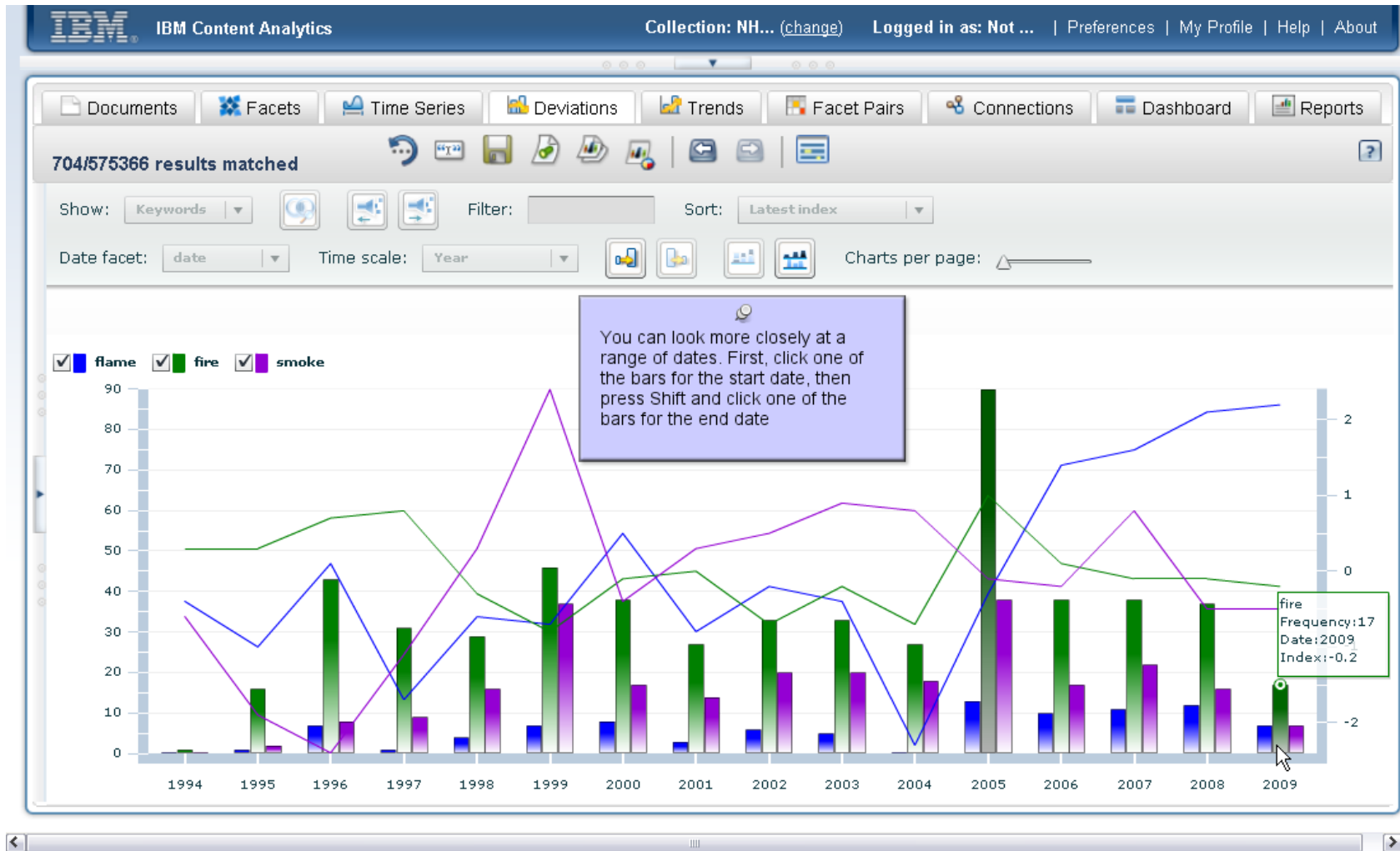




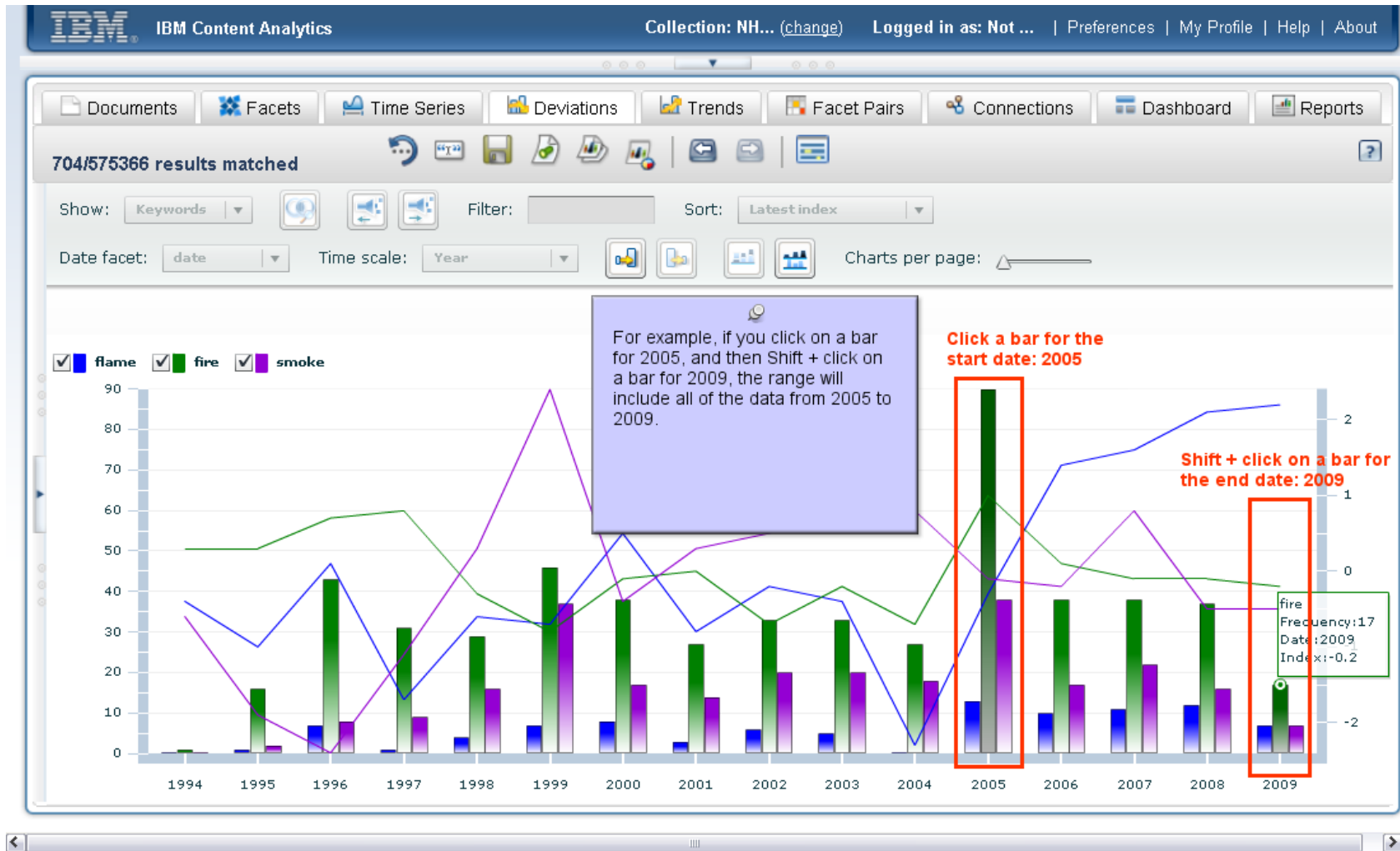
# Explore How Frequencies Deviate from the Expected Average



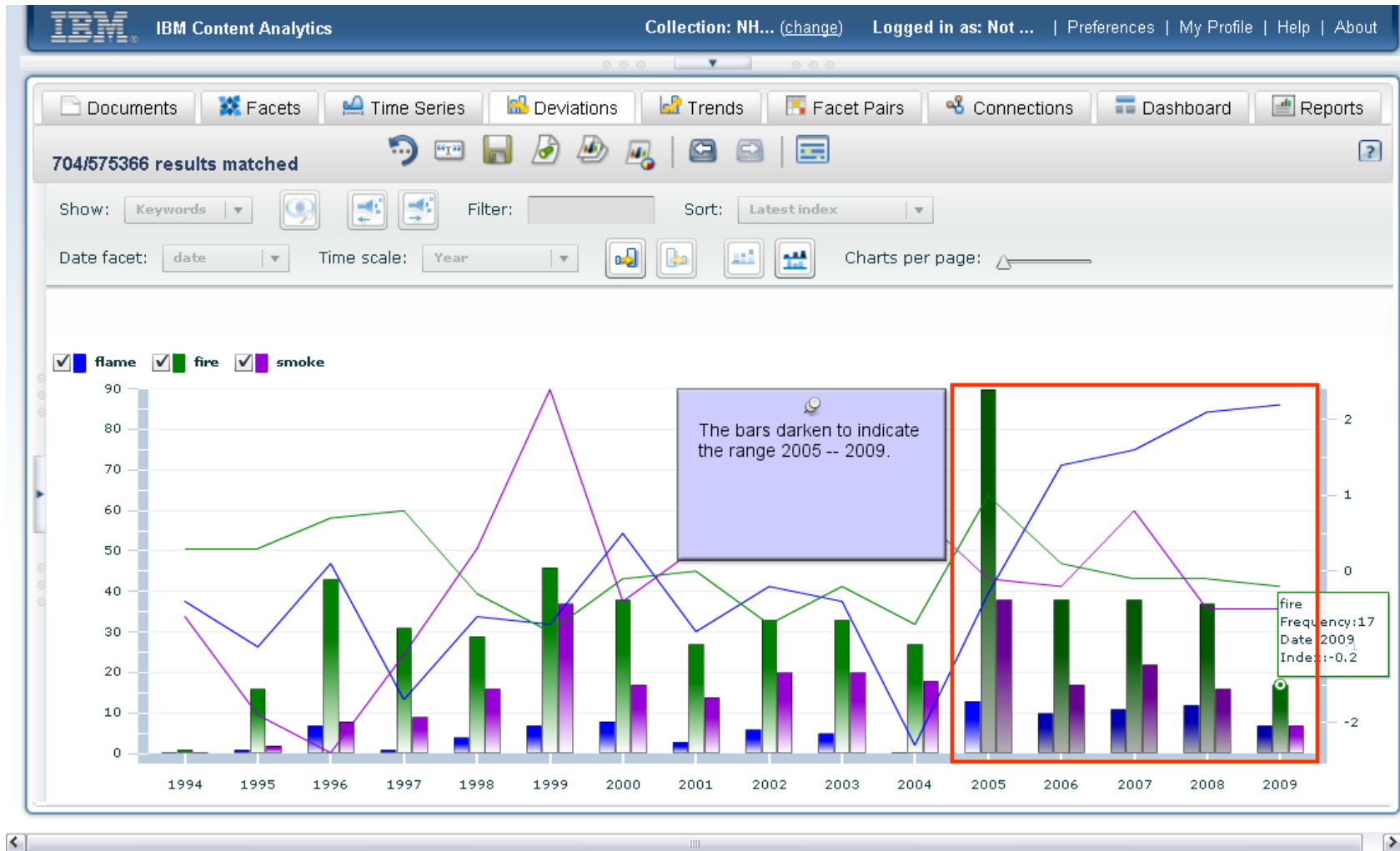
## Explore How Frequencies Deviate from the Expected Average



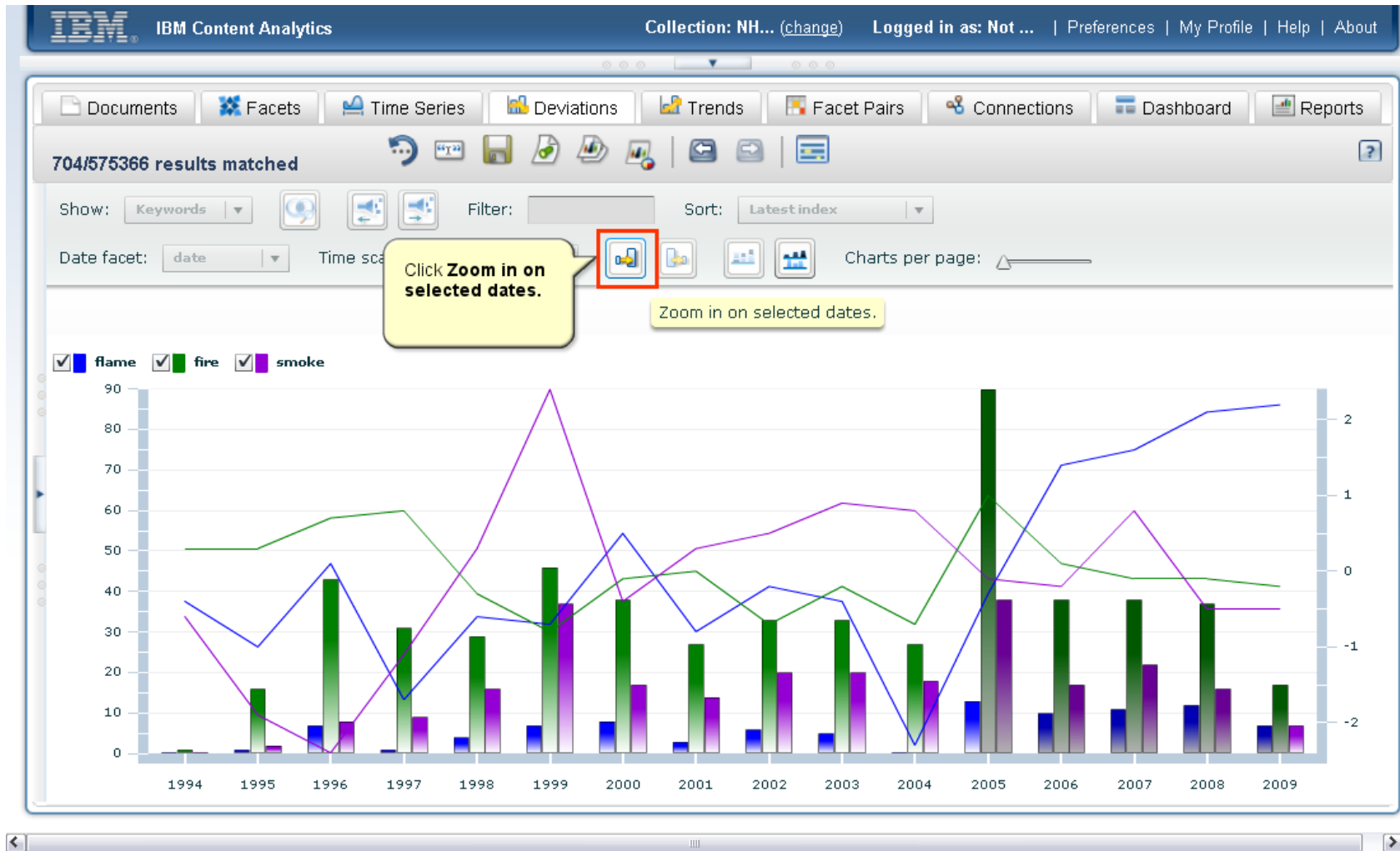
## Explore How Frequencies Deviate from the Expected Average



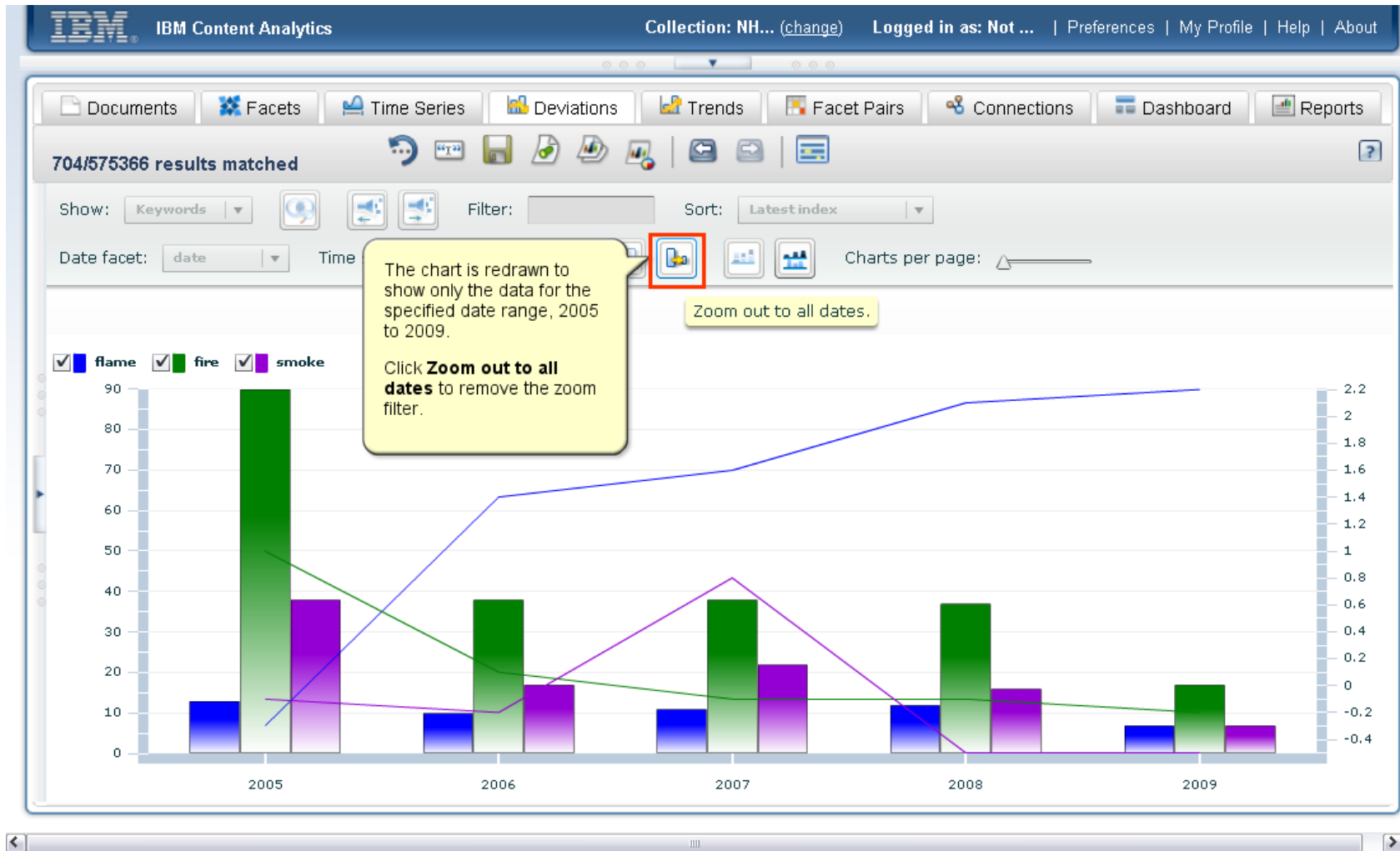
# Explore How Frequencies Deviate from the Expected Average



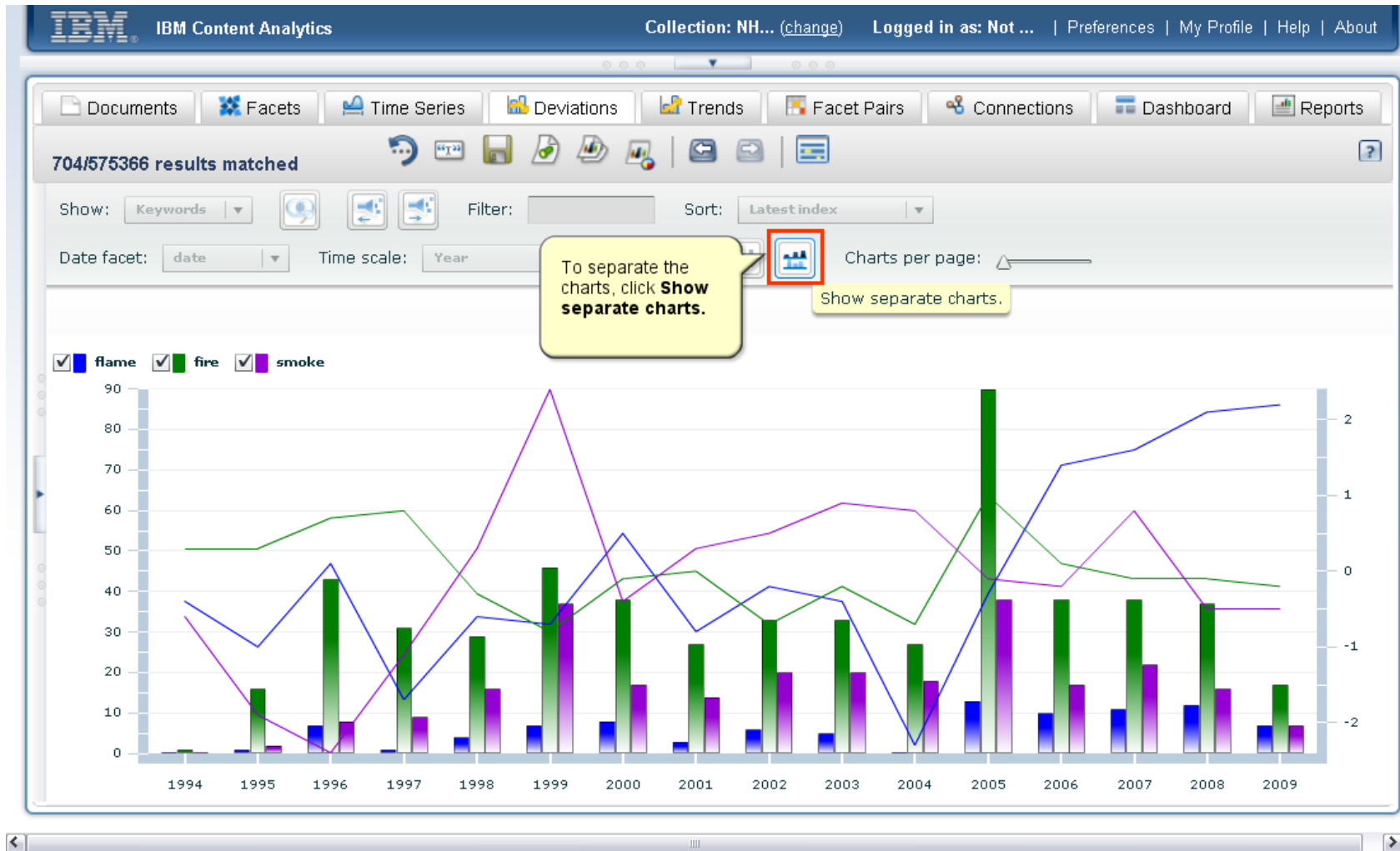
# Explore How Frequencies Deviate from the Expected Average



## Explore How Frequencies Deviate from the Expected Average



## Explore How Frequencies Deviate from the Expected Average



## Explore How Frequencies Deviate from the Expected Average

