



**SIREN**

**Easy Start Tutorial**

# Table of contents

<b>About this tutorial</b>	<b>4</b>
What does this tutorial cover?	4
What tasks will you complete?	5
What will you get from this tutorial?	5
<b>Installing Siren Platform - Easy Start</b>	<b>6</b>
Before you begin	6
Installing and running Siren Investigate	6
Relaunching Siren Investigate	8
<b>Welcome to Siren Investigate!</b>	<b>9</b>
<b>Importing data into Siren Platform</b>	<b>10</b>
About the demo data	10
Importing your first table of demo data	11
<b>Index pattern searches</b>	<b>15</b>
Creating your first index pattern search	15
<b>Creating dashboards automatically</b>	<b>18</b>
Automatically generating a dashboard	18
Editing dashboards	20
<b>Using the dashboard</b>	<b>22</b>
Creating filters	22
Textual Search	23
<b>Creating dashboards manually</b>	<b>26</b>
Importing the investments table	26
Creating the dashboard	27
Creating visualizations	28
Vertical Bar Chart	29
Record Table	30
Analytic Table	31
Setting the Dashboard Data Model	31
Importing the investors table	32
<b>Introducing relations in the Data Model</b>	<b>34</b>
Creating relations between entities	34
Creating relations manually	35
Visualizing the data model as a graph	37

Creating a self-relation	37
<b>Dashboard-to-dashboard associative navigation</b>	<b>39</b>
Recap: Where are we now?	41
<b>Exercise: Investigating investments</b>	<b>42</b>
Part 1: Finding an answer by filtering the Investors dashboard	43
Part 2: Finding an answer by using relational navigation	46
Part 3: Digging deeper with relational navigation	49
Part 4: Link analysis in the Graph Browser	51
<b>More about link analysis</b>	<b>60</b>
Embedded in dashboards	60
Using the Graph Browser	60
<b>Appendix 1: Natural Language Processing (NLP)</b>	<b>63</b>
Connecting extracted entities in the data model	65
Example usage of NLP Data in Siren Investigate	66
<b>Appendix 2: Configuring the graph</b>	<b>69</b>
Fixing labels with graph lenses	69
<b>Appendix 3: The Topic Explorer</b>	<b>71</b>
<b>Appendix 4: Creating Entity Identifiers</b>	<b>73</b>
<b>Appendix 5: Dashboard 360</b>	<b>75</b>
Filtering with Dashboard 360	78
<b>Appendix 6: Deleting data in Siren and changing a table schema</b>	<b>80</b>
<b>Appendix 7: Auto-discovery of relations</b>	<b>82</b>
<b>Legal notices</b>	<b>85</b>

# About this tutorial

This tutorial will help you to get started with the **Siren Platform - Easy start**, version 11.1 and later, which can be downloaded [here](#).

The tutorial helps you to install Siren Investigate as an empty platform, to import demo data, and to start exploring and analyzing the data.

## What does this tutorial cover?

Siren Platform can be applied to many kinds of scenarios and the method of importing data can vary, based on the type of data that is used.

This tutorial focuses on using *static* data, which we will import from Excel or CSV files. Static data, which does not change or changes only a few times per day, is commonly used in Business Intelligence (BI) or broader knowledge-discovery scenarios.

### Using streamed data and data from other databases

For use cases that involve streaming data, such as in **cybersecurity**, **operational intelligence**, **log management** or **IoT**, data streams can be loaded by external applications that write directly to the underlying Elasticsearch cluster. Examples of external tools for these scenarios are [Logstash](#), [Beats](#), [Fluentd](#), or [Streamset](#). It is also quite easy to write data to Elasticsearch directly by using the APIs.

Siren Platform also supports working on data in remote JDBC data sources. This can happen directly (with no data copy) or by using the built-in, UI-assisted *reflection* process, where data is copied and periodically refreshed.

These processes are outside of the scope of this tutorial. However, we recommend that you start with this tutorial (which uses CSV files) and then refer to our [documentation](#) to understand how to use data streams and JDBC data sources.

## What tasks will you complete?

In this tutorial, we will guide you through the following steps:

1. Installing Siren Platform.
2. Learning how to import a CSV file: Importing the first table (**companies.csv**).
3. Creating an index pattern search.
4. Creating the first dashboard by using the *auto create wizard*.
5. Editing a dashboard by using interactive visual filters and textual search.
6. Importing the next table (**investments.csv**) and creating a dashboard *manually*.
7. Importing the last table (**investors.csv**).
8. Creating an associative data model: Associating **companies** <-> **investments** <-> **investors**.
9. Creating a dashboard manually and including a relational navigator visualization that will allow for dashboard-to-dashboard navigation across all your data.
10. Introduction to the Graph Browser and Link Analysis, including exercises for you to complete.

The tutorial also includes a number of appendixes that contain additional important topics.

## What will you get from this tutorial?

By the end of this tutorial, and after completing the steps in Appendix 1, you will have a setup that is similar to our 'What is Investigative Intelligence' video, which we highly recommend you watch now.

Watch the [What is Investigative Intelligence](#) video.

This tutorial also has an accompanying video, which is available on the [Siren website](#).

# Installing Siren Platform - Easy Start

This download is supported on Windows, Linux, and macOS operating systems with a minimum of 8 GB of RAM.

Siren Platform supports the following browsers:

- Google Chrome
- Mozilla Firefox

## Before you begin

The following are **mandatory requirements** that must be completed before you install the Siren Platform:

- Ensure that **Java JDK version 8** is installed and set the `java_home` environment variable.

**NOTE:** We recommend downloading Amazon Corretto version 8 (go to the [Windows installer](#) or general [download page](#)). Alternatively, download and install the [Oracle JDK](#). To set the `java_home` environment variable, follow the instructions [here](#).

- On Windows, you must have an **improved ZIP manager** such as [7-zip](#) or [WinRAR](#). This is because the standard windows ZIP support will not be able to handle the distribution's large number of files.

## Installing and running Siren Investigate

To install Siren Investigate, complete the following steps:

1. Go to <http://www.siren.io/downloads> and download Siren Platform *Easy start*.
2. Complete the validation form, accept the license, and click **Proceed**.
3. Save the compressed file and extract it to a local directory.

**NOTE:** On macOS, if you are running Catalina (version 10.15) or a later version, you must remove an extended file attribute (`xattr`) before you extract the downloaded ZIP file. Run the following command:

```
xattr -d com.apple.quarantine
siren-platform-*-darwin-x86_64.zip
```

The extracted folder contains three main folders:

- **docs:** Contains the Siren Platform user guide in both HTML and PDF formats and the release notes.
- **elasticsearch:** Contains the self-contained, single-node Siren Enhanced Elasticsearch cluster.
- **siren-investigate:** Contains the Siren front-end application, which can be accessed through a web browser.

4. Launch the program, based on your operating system.

<p><b>Windows</b></p>	<p>Double-click on the <b>start.bat</b> file and wait for Elasticsearch and Siren Investigate to load.</p> <div data-bbox="579 909 1394 1137" style="border: 1px solid blue; padding: 5px;"> <p><b>NOTE:</b> You will see two command windows; one for Elasticsearch and the other for Siren Investigate. Do not close these windows until you are finished working with Siren Investigate.</p> </div> <div data-bbox="579 1189 1394 1469" style="border: 1px solid red; padding: 5px;"> <p><b>IMPORTANT:</b> The first time you run the installation, the Investigate window may appear to be stuck for a number of minutes and appear black. This is the optimization phase. Please wait and it will eventually start writing log messages again.</p> </div>
<p><b>Linux or macOS</b></p>	<ol style="list-style-type: none"> <li>1. Open a Terminal window and change directory to the extracted folder: <pre>cd {extracted folder}/elasticsearch</pre> Run the following command: <pre>./bin/elasticsearch</pre> The Terminal window shows Elasticsearch messages flowing. Do not close the command window. </li> </ol>

	<p>2. Open a new Terminal window and change directory to the extracted folder:</p> <pre>cd {unzipped folder}/siren-investigate</pre> <p>Run the following command:</p> <pre>./bin/investigate</pre>
<p><b>NOTE:</b> The installation is complete when you see the message “Siren Gremlin Server is up and running” in the log window.</p>	

5. Open a supported web browser and navigate to **http://localhost:5606**.

## Relaunching Siren Investigate

If your session is interrupted and you need to relaunch Siren Investigate, repeat steps 4-5 of the above procedure.

# Welcome to Siren Investigate!

As you launch Siren Investigate, the platform opens to a welcome dashboard. Expand the sidebar to the left of the screen to see the navigation menu, which contains a range of apps.

The most important apps for this tutorial are:

- Dashboard ①
- Data import ②
- Data model ③

The screenshot shows the Siren Investigate interface. On the left is a dark blue sidebar with the Siren Investigate logo at the top. Below the logo are several navigation items: 'Dashboard' with a red circle containing the number '1', 'Discover', 'Visualize', 'Timelion', 'Data import' with a red circle containing '2', 'Sentinel', 'Dev tools', 'Data model' with a red circle containing '3', and 'Management'. At the bottom of the sidebar are 'HOME Default Dataspace', 'Help', and 'Collapse' with a search input field labeled 'Filter by name'. The main content area is titled 'Getting Started' and has a subtitle 'Welcome to Siren Platform'. It contains a large heading 'Welcome to the Siren Platform - Easy Start version' and a sub-heading 'Before you begin'. Below this are two numbered steps: '1) Watch the Siren Platform overview video to understand Siren Platform's unique ability to provide investigative intelligence.' and '2) Access the Getting Started tutorial to learn about dashboards, visualizations, and building a data model.' There are also sections for 'Need help?' with a link to the community forum, and 'Resources' with links to 'Siren Platform documentation' and 'Siren website'. At the top right of the main content area are links for 'Share', 'Clone', 'Edit', 'Export', 'Watcher', and 'Reset'.

# Importing data into Siren Platform

(Average completion time: 10 mins)

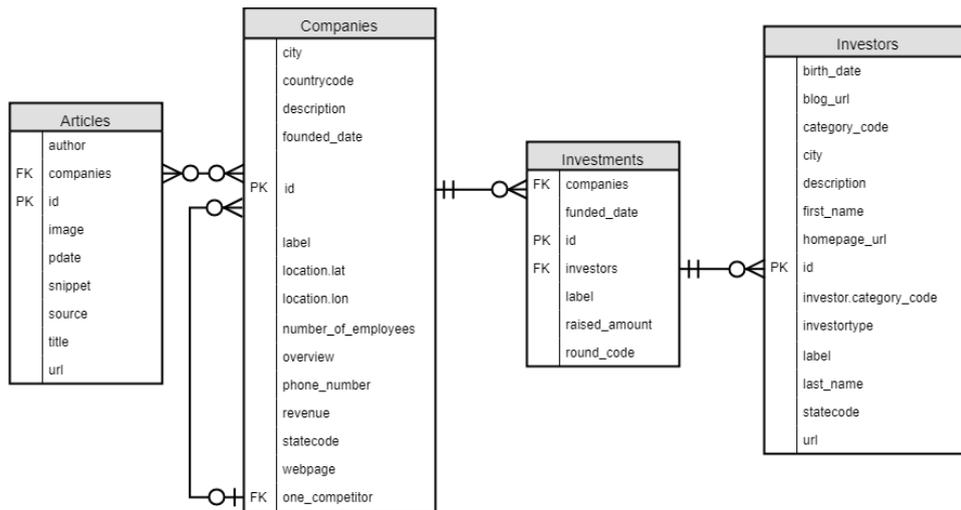
The first step is to import some data that we can work with.

## About the demo data

For the purpose of this tutorial, we are using sample data from a TechCrunch data source collected some years ago, along with a sample of technical articles that were collected online.

1. Download the [sample data file](#).
2. Extract the compressed folder that contains the following files:
  - **companies.csv**: A list of companies that includes geo-locations and descriptions.
  - **investments.csv**: An associative table that connects companies to investors with “amounts”, “round code” (for example, seed round, or round A), and the date of the investment.
  - **investors.csv**: A list of investors.
  - **articles.csv**: A collection of technical articles. **Advanced:** Most of the articles mention one or more companies. To extract (or annotate) these mentions, use the Siren NLP capabilities as described in [Appendix 1](#).

These files are relationally connected. The following entity-relationship diagram shows the connections between the tables:



There are also other relationships in the data, for example, city and state names, URLs, and email addresses. You'll see later how these can be auto-discovered.

## Importing your first table of demo data

(Average completion time: 5 mins)

This section will take you through the steps required to import your first table of the sample dataset. Let's start by importing the **Companies.csv** file.

1. From the left-hand navigation menu, click **Data import**.
2. Click **Excel/CSV Import**.

- From the **Excel/CSV Import** screen, upload the **companies.csv** file ①.

Import your xlsx and csv file to Elasticsearch

companies.csv  
Remove

Select the sheet to import: companies  
Select a template (optional):  
> Next

blog_url	category_co...	city	countrycode	deadpooled...	description	email_adre...	founded_date	hasstatus	homepage_url	id
http://www...	ecommerce				We sell aff...	50offbags...	2007-02-15		http://www...	compan...
	ecommerce	Athens	USA		Crowdfund...	rok3109@g...	2011-11-30		http://www...	compan...
	games_video	Guangzhou	CHN				2005-04-01	funded	http://www...	compan...
https://blo...	network_h...	Berlin	NLD		HTML5 ap...	mail@5app...	2011-12-31		https://5ap...	compan...
http://com...	mobile				Cellular Ne...	ir@5barz.c...	2007-11-30		http://www...	compan...

- You will see a preview of the uploaded data ②. Click **Next** ③.
- On the **Configure Indexing** screen, accept the default **Skip transform** value under **Transform Pipeline** ①. This is used later to activate transformations such as textual processing using the [Siren NLP plugin](#).

Transform Pipeline

Method: Skip transform

Indexing Settings

Index scope: Local  
Index name: siren-import-home-companies  
Use Custom storage primary key?

Remove columns:

Mapping

Use Custom Mapping?

Fields	Samples	Type	Remove Column
blog_url	blog_url Samples http://www.50offbags.com/blog/ https://blog.5apps.com http://community.5barz.com/	Keyword (For Aggrega... Multi-Valued? Advanced Mapping?	<input type="checkbox"/>

6. In the **Index scope** field ②, select **Local** from the dropdown menu. This setting defines how dataspace are created. For more information, see the [dataspace documentation](#).
7. In the **Index name** field ③, accept the default name.
8. Keep the **Use Custom Mapping** switch ④ on. This allows you to map the type (or format) for each field in the CSV file and determines your ability to search/sort/drill down and analyze the data.

In the Mapping section of the screen, most of the data types that are listed in the **Type** drop-down menus ⑤ are easy to understand, such as Integer, or Date. You can check what the correct value might be by looking at the **Samples** column.

However, the following options require some explanation:

#### **Text**

Use this option for long texts, such as emails and messages. When you choose **Text**, the back-end system splits the content into individual words and calculates statistics on those words. This allows you to have word clouds and the best level of searchability.

#### **Keywords**

Use this option for strings that should be considered *unbreakable*. For example, your dataset might include a field called City that contains strings such as 'New York' and 'Los Angeles'. If you mark this field as **Keyword**, then the term will be searched for in its entirety, rather than broken into single-word search terms (which would then lead to wrong results later).

#### **Multi-Valued**

The **Multi-Valued** switch ⑥ can be used when a field in the CSV file contains an array of values, typically separated by a delimiter. For example, you might have an **email** field that contains "mario@bros.io|mario.bros@mail.io|info@mario.io". By switching the **Multi-Valued** switch to 'on', you can parse these three values as separate email addresses, which is useful for search and analysis.

#### **Advanced Mapping**

The **Advanced Mapping** switch ⑦ gives advanced users the option to add a JSON object to define the mapping in Elasticsearch format. For example, selecting **Date** and switching the **Advanced Mapping** switch to 'on' would allow you to define a specific format for the date field, such as {"format":"yyyy-MM-dd HH:mm"}.

For now, ensure that the following selections are correct:

1. Scroll to the **description** field and set the type to **Text (Allows Word Cloud)**.
2. For the **founded\_date** field, set the type to **Date**.
3. For the **Geopoint** field, set the type to **Geo Point**.
4. For the **overview** field, set the type to **Text (Allows Word Cloud)**.
3. Click **Import**.

After the import is complete, you see a confirmation message.

The data has been imported (or indexed) in the back-end system. Next, we need to access it through the **Data model** in the front-end system.

Click **Create an Index Pattern Search** in the confirmation message.

**TIP:** You can also create an index pattern by going to the **Data model** app and clicking **Create index pattern search** ①.

# Index pattern searches

Before we create an index pattern search, let's present some fundamental concepts and terminology that will help you to understand the platform.

- **A document:** A single record. In this case, we have created one document for each row in the CSV file. Documents have *fields* that contain *values*.
- **An index:** In Elasticsearch, an index is a concept very similar to a table in a database (a collection of documents). The CSV import has created the **companies** index for us at this point.
- **An index pattern search:** A way to combine more indexes. The simplest form of an index pattern search is... simply the name of an index itself, for example, *companies*. More advanced index patterns can be used to combine indexes in advance deployments, for example, indexes that are suffixed or prefixed with dates. Index pattern searches are the core 'searchable objects' in Siren Platform. This means that index patterns can be searched, drilled down, and become the basis of analytical visualizations.
- **Searches or 'saved searches':** Typically, searches are a subset of an index pattern search. For example, if you wanted to sort your list of companies by those based in France, you can set up a search for French companies, which would then appear under the 'companies' index pattern search.

## Creating your first index pattern search

(Average completion time: 5 mins)

To create your first index pattern search, complete the following steps:

1. Keep the default name of the index pattern ① and the name of this search ③ as "siren-import-companies".
2. In the **Time Filter field name** field ②, select **founded\_date** from the drop-down menu.

**TIP:** If you choose a Time Filter field name, then Siren Platform treats the data as a

*time series*. In this case, we will consider companies to be a time series and regard the “founded\_date” as the main time<sup>1</sup>.

3. Select an icon ④ to represent the type of data that the index pattern will show. For example, an icon of a building could represent a company.
4. Choose a colour for the icon ⑤.
5. Click **Create** ⑥.

Data model Save Delete

+ Create index pattern search

+ Create entity identifier

There are no Index Pattern Searches or Entity Identifiers

### Create an Index Pattern Search

Siren Investigate requires at least one configured Index Pattern Search. Index Pattern Searches identify the indexes to run.

**Index pattern** ①

siren-import-companies

The Index Pattern selects the indexes to query.  
You can set the name of the index or select multiple indexes with the \* wildcard, for example: logstash-\*.  
Note: virtual indexes can be used only by direct naming, not as part of wildcard patterns.

**Time Filter field name** ② refresh fields

founded\_date

**Name of this search** ③

siren-import-companies

The search name, typically describing returned entity types, for example: "users", "logs firewall 1", "articles".

**Short Description**

**Icon** ④

far fa-building

**Color** ⑤

#8339b0

Generate the *primary key* and *single value* field properties automatically ⑥

Create

After you’ve created the **index pattern search**, you should see something like the screenshot below.

<sup>1</sup> In general, if you select a time field, then this must be present in all of your records or you will not see all of your data.

- All index pattern searches that you create appear on the left under **Searches** ①.
6. Select the **Fields** tab ② to see which fields are now imported.
  7. Select the **Data** tab ③ to see the imported records: Approximately 160,000 records are imported for this dataset.
  8. On the **Info** tab, the field **Label when visualized in the graph browser** ④ is used to set the default field for labeling nodes on the graph browser. Select **Document Field** and, in the second dropdown menu that appears, select **label**.
  9. Click **Save** ⑤.

## Creating dashboards automatically

The Siren Investigate dashboards display a set of visualizations in a customizable grid layout. Each visualization is normally linked to a search (*such as companies or investments*). Dashboards can be organized in dashboard groups, shared, and customized.

You can create dashboards manually or automatically. To manually create a dashboard, you'd need to create and add each visualization manually to your dashboard. Alternatively, the dashboard generator automatically creates the visualizations for you.

The following message appears after you create a Index Pattern Search. Let's see together how we can automatically create a dashboard with Siren Investigate.

**i** No dashboard exists for this Index Pattern Search  
Did you know? Siren can create one for you, click on the [data tab](#) to get started.

## Automatically generating a dashboard

(Average completion time: 5 mins)

Siren features an AI-driven dashboard creation wizard, which is a great way to get started on any new dataset with just a few clicks.

The auto-generate dashboard option is available both on:

- the **Data** tab of the **Data model** app, and;
- the **Discover** app in the navigation menu, which is a stand-alone app that provides a shortcut to this functionality.

Let's start by generating a dashboard for the **companies** data that we just imported.

1. From the **Data model** app, go to the **Data** tab. This page shows a preview of the records in the index pattern and on the left you can see a list of available fields.
2. Click the **Discover most relevant** button.
3. The system identifies the fields that are most relevant for generating the dashboard. You can change the selected fields if you wish and then click **Select**.
5. Now, the selected fields are listed on the left. Click **Generate a dashboard**.

- In the **Generate a prepopulated dashboard** window, replace the default dashboard title with “Companies” and click **Generate**.
- Once the process is complete, the **Generate a prepopulated dashboard - Report** screen is displayed:

Generate a prepopulated dashboard – Report

<input checked="" type="checkbox"/>	Type	Field	Chart Type	Vis Name
<input checked="" type="checkbox"/>	🕒	founded_date	Line	Documents count by founded_date
<input checked="" type="checkbox"/>	🌐	Geopoint	Enhanced Coordinate Map	Geopoint locations map
<input checked="" type="checkbox"/>	t	category_code	Vertical Bar	Top 50 category_code
<input checked="" type="checkbox"/>	t	city	Analytic Table	List of city
<input checked="" type="checkbox"/>	t	countrycode	Analytic Table	List of countrycode
<input checked="" type="checkbox"/>	t	description	Tag Cloud	Top 30 description
<input checked="" type="checkbox"/>	t	hasstatus	Pie	Top 30 hasstatus
<input checked="" type="checkbox"/>	#	number_of_employees	Vertical Bar	Histogram of number_of_employees
<input checked="" type="checkbox"/>	t	overview	Tag Cloud	Top 30 overview
<input checked="" type="checkbox"/>	+	statecode	Analytic Table	List of statecode

Cancel Create the dashboard

The report screen contains the following elements:

- ① - Checkboxes that allow you to individually select which of the auto-generated visualizations to keep in the new dashboard.
  - ② - Dropdown menus, where you can select the best type of visualization (Chart Type) for each field.
  - ③ - Text boxes that allow you to personalize the name of each visualization (Vis Name).
- Click **Create the dashboard** ④.

The new **Companies** dashboard is now generated with different types of visualizations including an Enhanced Coordinate Map, Histogram, Pie Chart, Record Table, Analytic Table and a Tag Cloud. For more information about visualizations, see the [Siren Platform documentation](#).

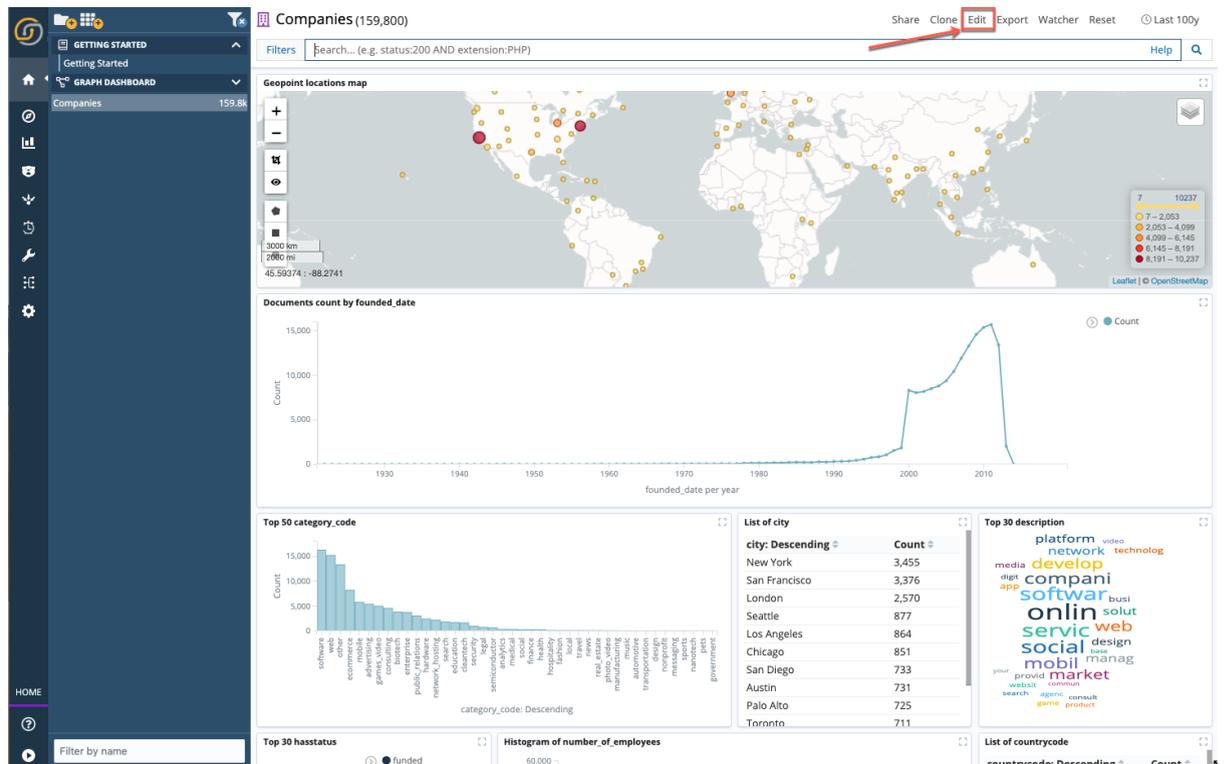
**TIP:** If you do not specify a new dashboard name, it defaults to the name of the index pattern search. You can edit the dashboard name at any time by right-clicking the dashboard in the dashboard menu and clicking **Rename**.

## Editing dashboards

(Average completion time: 5 mins)

Let's tidy up the **Companies** dashboard by reorganizing the visualizations.

1. At the top-right of the screen, click **Edit**.



The dashboard is now in edit mode and you will notice that the edit buttons appear at the top-right of each visualization. You can now resize, reorder, edit, and remove the visualizations.



2. Rearrange some visualizations to improve the look and feel of the dashboard. You can also remove the visualizations that you think are not relevant. See below an example of how you can rearrange the companies dashboard:

Editing Siren-import-home-companies (159,808) - (UNSAVED) **Save** Cancel Data Model Add Options Share Watcher Reset Last 100y

Filters Search... (e.g. status:200 AND extension:PHP) Help Q

Time	Geopoint	category_code	city	countrycode	description
August 1st 2014, 01:00:00.000	-	software	-	-	collaboration, project management
May 31st 2014, 01:00:00.000	28.61393 91.77.209 0212	ecommerce	Delhi	IND	Coupons for India
March 31st 2014, 01:00:00.000	-	other	-	-	Micro Brewery

category\_code: Descending

countrycode	Count
USA	46,033
GBR	6,175
IND	3,129
CAN	3,049
DEU	1,601
FRA	1,398
AUS	1,138
ESP	918
ICP	857

3. Click **Save** at the top-right of the screen.

## Using the dashboard

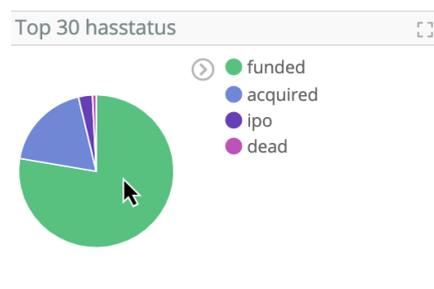
Now, let's explore this dashboard by interacting with the visualizations to take a closer look at important data.

### Creating filters

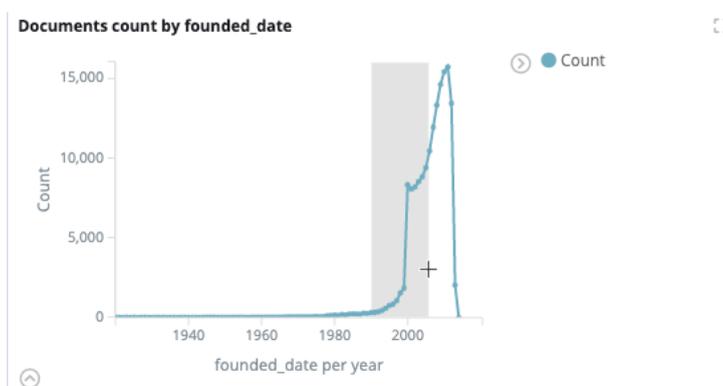
(Average completion time: 5 mins)

You can create filters for dashboards either by using the **Add filter** button or from the visualizations, for example, as follows:

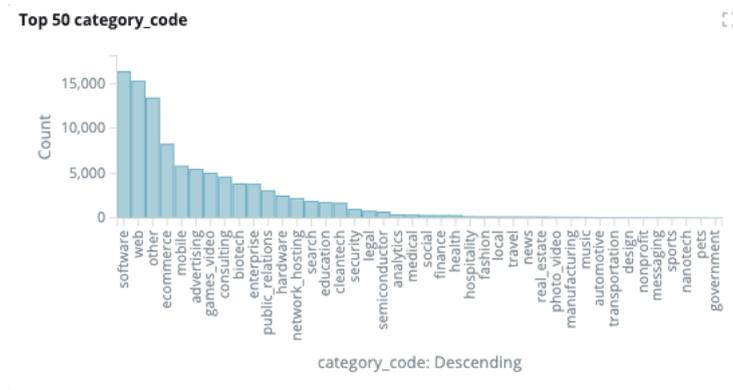
1. Click a segment of the **Top 30 hasstatus** pie chart to filter the dashboard on a category, such as the **funded** status.



2. On histogram visualizations, you can create multiple types of filters:
  - a. On the **Documents count by founded\_date** histogram, click and drag across multiple points to select a range (for time-based histograms only). For this tutorial, there is no need to select a range here right now.



- b. On the **Top 50 category\_code** histogram, click on a bar in the chart to select a specific category.



- c. You can also make multiple selections by holding **Ctrl** (or **Cmd**) and clicking on the bars. Hold the **Ctrl** key and click on the bar corresponding to **software** and **web** category\_code. After making the selections, click on **Apply Now** ① to apply a filter with an OR condition.

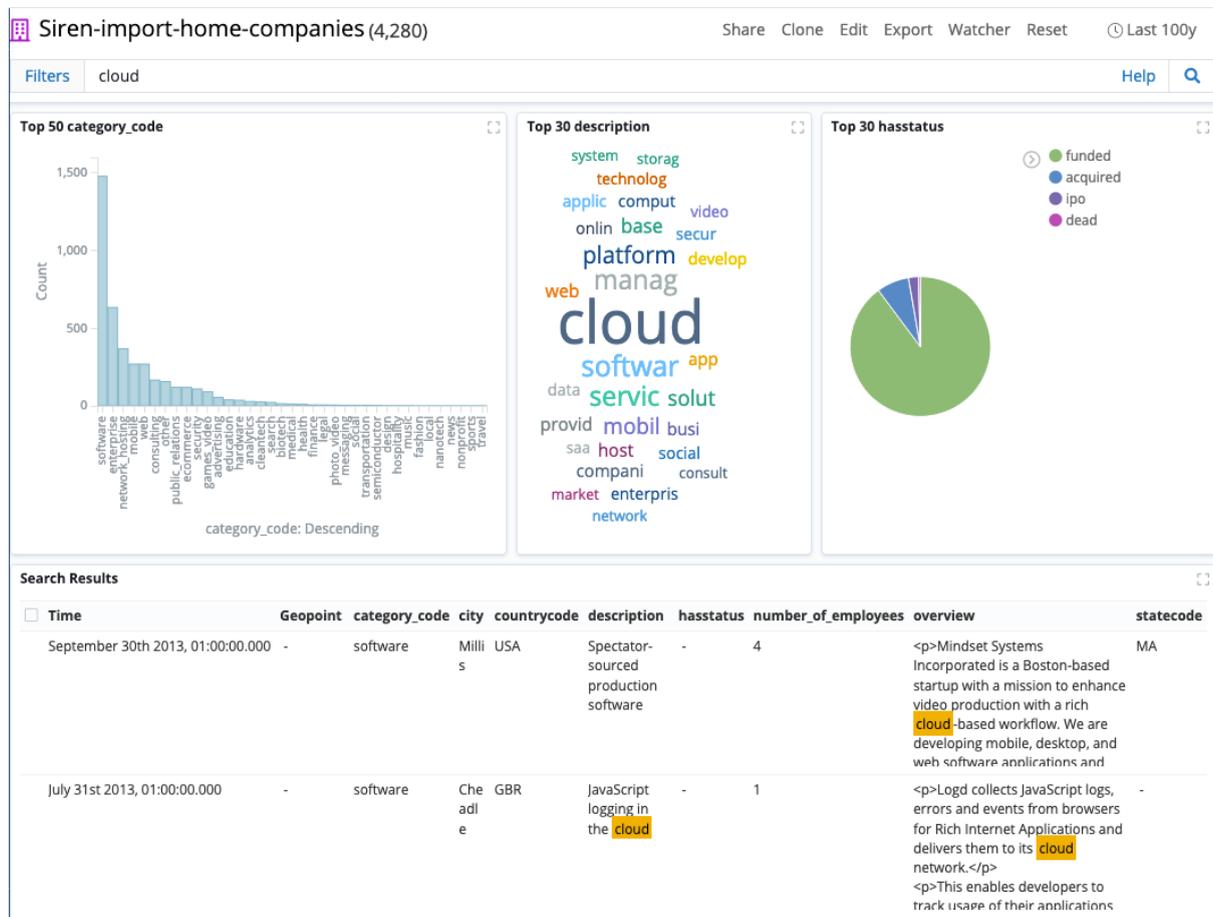
With the filters created, the dashboard updates as follows:

- The following filters appear: **category\_code: software,web** and **hasstatus: “funded”**.
- The number of records displayed in the dashboard is reduced and a filter icon appears alongside the dashboard in the sidebar.

## Textual Search

On the dashboards, you can perform a free text search by using the Search bar. You can rearrange the visualizations to bring the text-specific ones to the top.

Clear the current filters and simply enter a text string, for example, enter the word “cloud” and hit the **Enter** key. The dashboard updates all of the visualizations to reflect the search results.



You can also perform an advanced search by using the [Lucene query syntax](#).

- To search for a value in a specific field, prefix the value with the name of the field. For example, enter **category\_code:software** to find all of the records that have “software” as the category code.
- Use the Boolean operators AND, OR, and NOT for complex queries. For example, enter **category\_code:software OR category\_code:web** to get all of the results with either the “software” or “web” category code.
- You can also use the **fuzzy operator** ‘~’ to search for terms that are similar to, but not exactly matching the search terms. For example, type **commerce~** to get the results of all mentions of ‘ecommerce’, ‘commercial’, and so on.

To reset the filters on the current dashboard, click **Reset** ①. If you have multiple dashboards, you can reset all of the unsaved filters and searches in all of your dashboards,



## Creating dashboards manually

It's time to import your next table and create the next dashboard.

### Importing the investments table

(Average completion time: 5 mins)

Return to the **Data Import** app to import the **investments.csv** file and create the relative index pattern search: **Do not** automatically create a dashboard this time.

Follow the same steps as you did in [the first import](#), making the following settings:

1. Set the **Index scope** to **Local**.
2. For the **funded\_date** field, select **Date**.
3. For the **raised\_amount** field, select **Long**. (This allows an integer over 32 bits: Some investments are big!)
4. Click **Import** and click **Create an Index Pattern Search**.
5. On the **Create an Index Pattern Search** screen:
  - a. In the **Time Filter field name** field, select **funded\_date**.
  - b. Choose an appropriate icon for investments, such as a dollar bill
  - c. Choose a colour for the icon.
  - d. Click **Create**.
6. On the **siren-import-home-investments (*Index pattern search*)** screen, set the **Label when visualized in the graph browser** to be a **Document Field** and, in the second dropdown menu that appears, select the field **raised\_amount**.
7. Click **Save** in the top-right corner of the screen.

At this point, the **Data model** app should look like this:

The screenshot shows the 'Data model' configuration page for an 'Index pattern search' named 'siren-import-home-investments'. The interface includes a sidebar with navigation icons and a main configuration area. The configuration area has tabs for 'Info', 'Fields (20)', 'Data (41317)', 'Relations', 'Scripted fields (0)', 'Options', 'Revisions', and 'Data model graph'. A message at the top states: 'No dashboard exists for this Index Pattern Search. Did you know? Siren can create one for you, click on the data tab to get started.' The configuration fields include:

- Name of the search:** siren-import-home-investments
- Icon:** far fa-money-bill-alt
- Color:** #61b322
- Label when visualized in the graph browser:** Document Field (with sub-selection: raised\_amount, preview)
- Short description:** (empty field)
- Index pattern used by this search:** siren-import-home-investments
- Time Filter field name:** funded\_date (with a refresh fields button)

## Creating the dashboard

(Average completion time: 7 mins)

In this section, we'll create a dashboard manually for the data that we just imported.

**Note:** This is an optional step, if you don't want to learn how to do this, you can use the *auto-generate* option as before and skip to the [Importing the investors table](#) section.

To create a dashboard manually for **investments**:

1. Go to the **Dashboard** app.
2. Click on the **Create new dashboard** button ①.

The screenshot shows the Siren Platform interface. A dialog box titled "Create a new dashboard" is open in the center. It contains a text input field with the word "Investments" and two buttons: "Cancel" and "Create". The background dashboard is titled "Siren-import-home-companies (44,361)" and features several visualizations: a bar chart for "Top 50 category\_code", a word cloud for "Top 30 description", and a pie chart for "Top 30 hasstatus". Below these is a "Search Results" table with columns for Time, Geopoint, category\_code, city, countrycode, description, hasstatus, number\_of\_employees, overview, and statecode. The table contains two rows of data.

Time	Geopoint	category_code	city	countrycode	description	hasstatus	number_of_employees	overview	statecode
August 1st 2014, 01:00:00.000	-	software	-	-	collaboration, project management	-	-	<p>Simple, easy project management.</p><p>Raavel is the simplest way to manage and track anything! Share	-
May 31st 2014, 01:00:00.000	28.61393 91,77.209 0212	ecommerce	Delhi	IND	Coupons for India	-	14	<p>Couponji gives Coupons, Coupon Codes, Promo Coupons, Promo codes, discount vouchers and	-

3. Give the dashboard a name, for example, “Investments” and click **Create** ②.

At this point, you have created a dashboard, but it does not yet contain any visualizations and it is not linked to a search.

## Creating visualizations

(Average completion time: 15 mins)

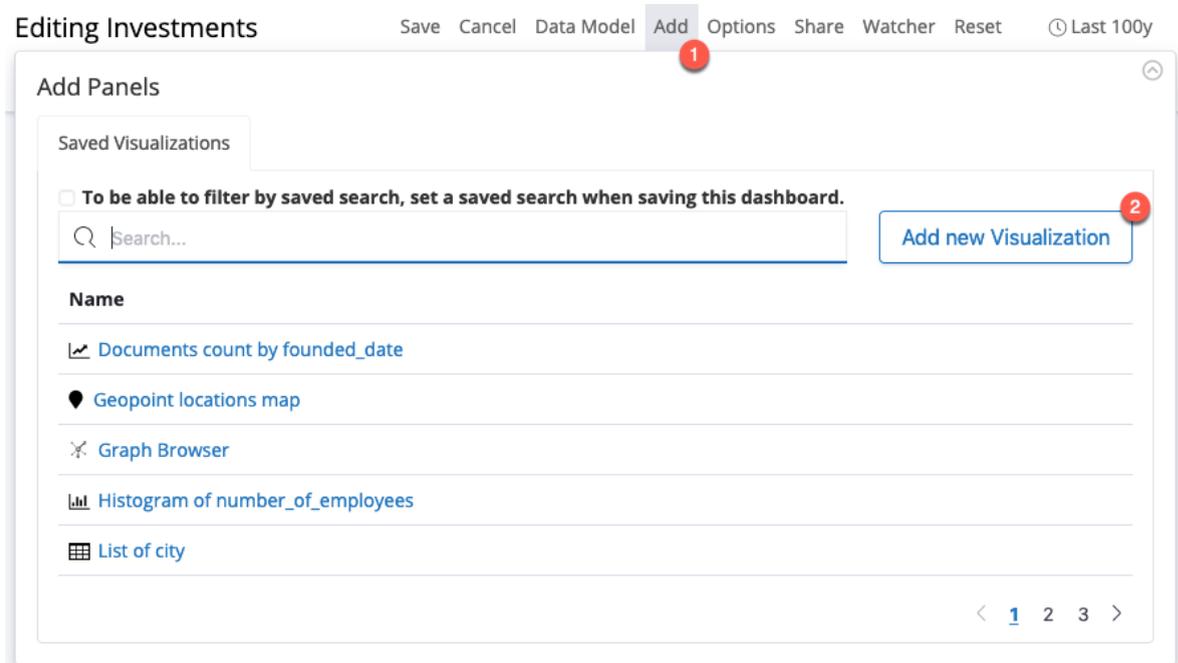
You can create visualizations in two ways; either from the **Visualize** tab on the left navigation bar or directly from the current dashboard.

We will add visualizations in the following order:

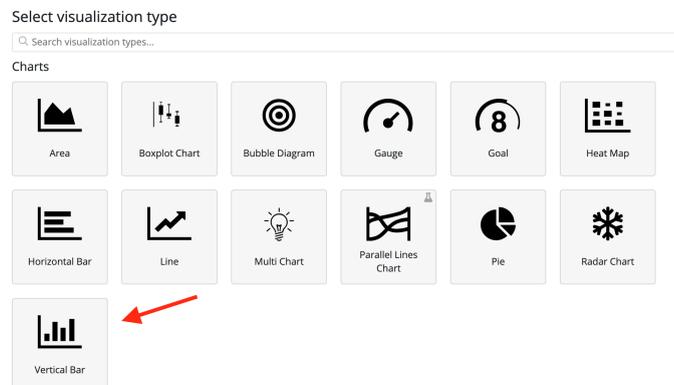
- Vertical Bar chart
- Record Table
- Analytic Table

## Vertical Bar Chart

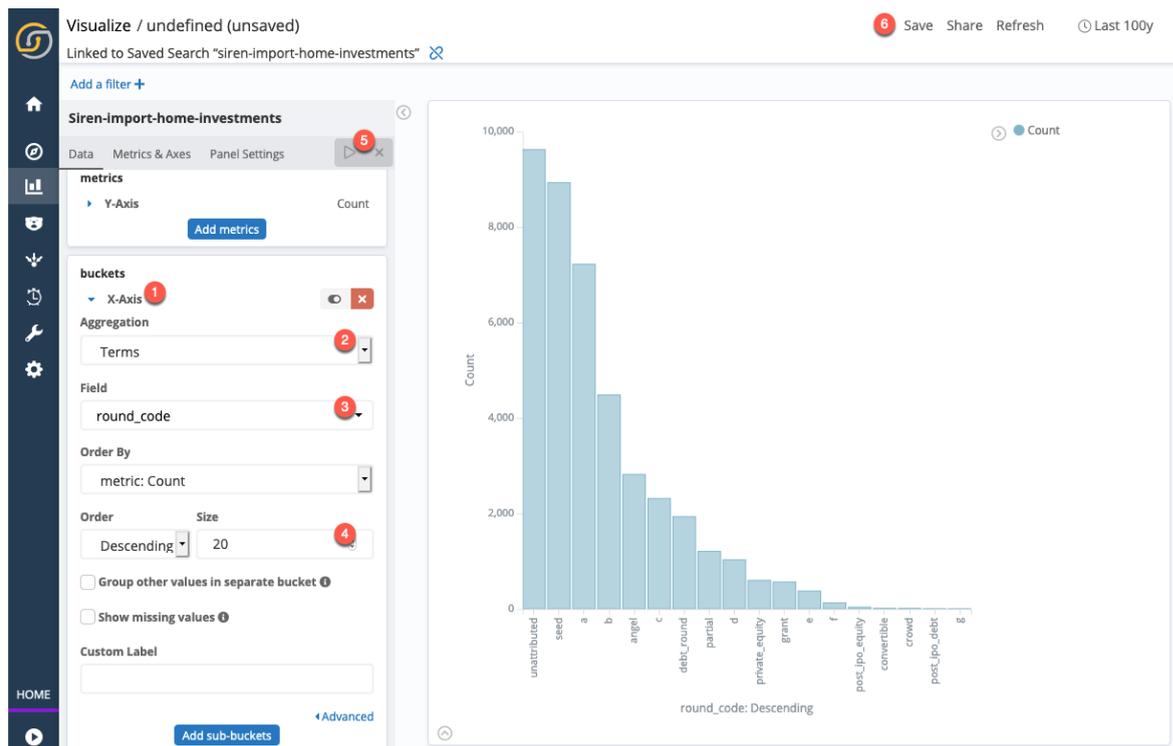
1. Click **Add** ① from the menu bar at the top.



2. Click **Add a Visualization** ②.
3. Click **Vertical Bar** in the **Charts** category.



4. On the Pick a Saved Search screen, select the **siren-import-home-investments** saved search, since we're building a dashboard for the investments index pattern search.

5. In the **buckets** table window, select **X-axis** ①.6. In the **Aggregation** field, select **Terms** ②.7. In the **Field** field, select **round\_code** ③8. In the **Order Size** field, specify a value of **20** ④.9. Click **Apply changes** ⑤ to see a preview of your chart.10. Finally, click **Save** ⑥ to save the Visualization.11. On the **Save** panel, name the visualization “Investment Number by Round Code”, and click **Save to Dashboard**.12. You are redirected to the **Investments** dashboard in edit mode, where you can resize or reorder the visualizations. Click **Save** → **Save dashboard**.

## Record Table

1. From the dashboard, click **Edit** → **Add** → **Add a Visualization**, as before.
2. Use the search bar or scroll to the Showing individual records category to find **Record Table**.
3. On the **Pick a Saved Search** screen, select **siren-import-home-investments**.
4. The default values are adequate, so click **Apply changes**.
5. Click **Save**, name the visualization “Record Table”, and click **Save to Dashboard**.
6. When you return to the dashboard screen in Edit mode, click **Save** → **Save dashboard**.

**TIP:** In a record table, you can remove the columns that you do not need by clicking the 'x' icon beside the column name. To move columns, use the << >> buttons.

## Analytic Table

The Analytic Table visualization is similar in appearance to the Record Table. The difference is that you can view individual records with a Record Table, while you can make summaries with an Analytic Table.

1. From the dashboard, click **Edit** → **Add** → **Add a Visualization**, as before.
2. Use the search bar or scroll to the Textual category to find **Analytic Table**.
3. On the **Pick a Saved Search** screen, select **siren-import-home-investments**.
4. In the **buckets** table window, select **Split Rows**.
5. In the **Aggregation** field, select **Terms**.
6. In the **Field** field, select **raised\_currency\_code**.
7. In the **Order By** field, select **metric: Count**.
8. In the **Order Size** field, specify the value as **10**.
9. Click **Apply changes**.
10. Click **Save**, name the visualization "Analytic Table", and click **Save to Dashboard**.
11. When you return to the dashboard screen in Edit mode, click **Save** → **Save dashboard**.

## Setting the Dashboard Data Model

(Average completion time: 3 mins)

Before going any further, we need to specify a type of data model for the dashboards. Siren Investigate provides three dashboard options:

- **Dashboard does not represent a search:** Used when the dashboard's main function is to provide summary information and it does not need to be linked to a search. These dashboards do not display a number (count) or an icon.
- **Dashboard is about a search:** Used when the dashboard represents a single specific search (for example, companies or articles), enabling the following:
  - The dashboard displays the count and an icon associated with the search.

- Components like Relational Navigator work and can be used on the dashboard.
- The content of the dashboard can be added to a Graph Browser for link analysis.
- **Dashboard 360 with filter strategy:** Uses a *dashboard-specific data model*, which enables a single dashboard to contain visualizations that are based on different searches, and to perform coherent filtering across all of them. This type of dashboard depends on Siren's relational capabilities and the creation of a data model. We will be exploring this option later in the tutorial.

At this point, we have created a dashboard with visualizations that are all linked to the **investments** search. The logical thing to do is to specify that this dashboard is dedicated to the same search, and use the appropriate dashboard data model.

1. On the **Investments** dashboard, click **Edit**, then click **Data Model** ①. You will now see the Dashboard Data Model options:



2. Select the **Dashboard is about a search** option ②, and then select **siren-import-home-investments** from the dropdown list.
3. Click **Save** ③ and **Save dashboard**.

## Importing the investors table

(Average completion time: 5 mins)

*But don't create a dashboard yet!*

The next file you need for this tutorial is **investors.csv**.

In the **Data import** app, use the [same procedure that we used previously](#) with the following settings:

1. On the **Excel/CSV Import** screen, set the **Index scope** to **Local** and click **Import**.
2. On the **Create an Index Pattern Search** screen:
  - a. In the **Time Filter field name** field, select **I don't want to use the Time Filter**, as this is not a time-based index.
  - b. Choose an appropriate icon for investors, such as a user icon.
  - c. Choose a colour for the icon.
  - d. Click **Create**.
3. On the **Index pattern search** screen, set the **Label when visualized in the graph browser** to be a **Document Field** and, in the second dropdown menu that appears, select the field **label**.
4. Click **Save** in the top-right corner of the screen.

The **Searches** section of the data model should now contain three searches: **\*companies**, **\*investments**, and **\*investors**.

The screenshot shows the 'Data model' interface with a sidebar on the left containing navigation icons and a 'Searches' section. The main area displays the configuration for an 'Index pattern search' named 'siren-import-home-investors'. The configuration includes:

- Name of the search:** siren-import-home-investors
- Icon:** far fa-user-chart
- Color:** #2a3bde
- Label when visualized in the graph browser:** Document Field (with a dropdown menu showing 'label' selected and a 'preview' link).
- Short description:** (empty text field)
- Index pattern used by this search:** siren-import-home-investors
- Time Filter field name:** I don't want to use the Time Filter

At the top right of the configuration panel, there are 'Save' and 'Delete' buttons.

**IMPORTANT:** Do not create a dashboard yet for investors. We will do that later, after we have defined the associative data model.

## Introducing relations in the Data Model

The true power of Siren Platform is in its *associative* data model; defining how your tables are interconnected by their relations.

It enhances both the navigation across related dashboards and it drives the link analysis feature, the Graph Browser.

In Siren Platform, a *relation* is a labeled link between entity types. An *index pattern search* can be seen as an entity type, which means that two index pattern searches can be connected by a relation in the data model.

You create these relations in the UI. For example, the Investments table is connected to the Companies table as follows:



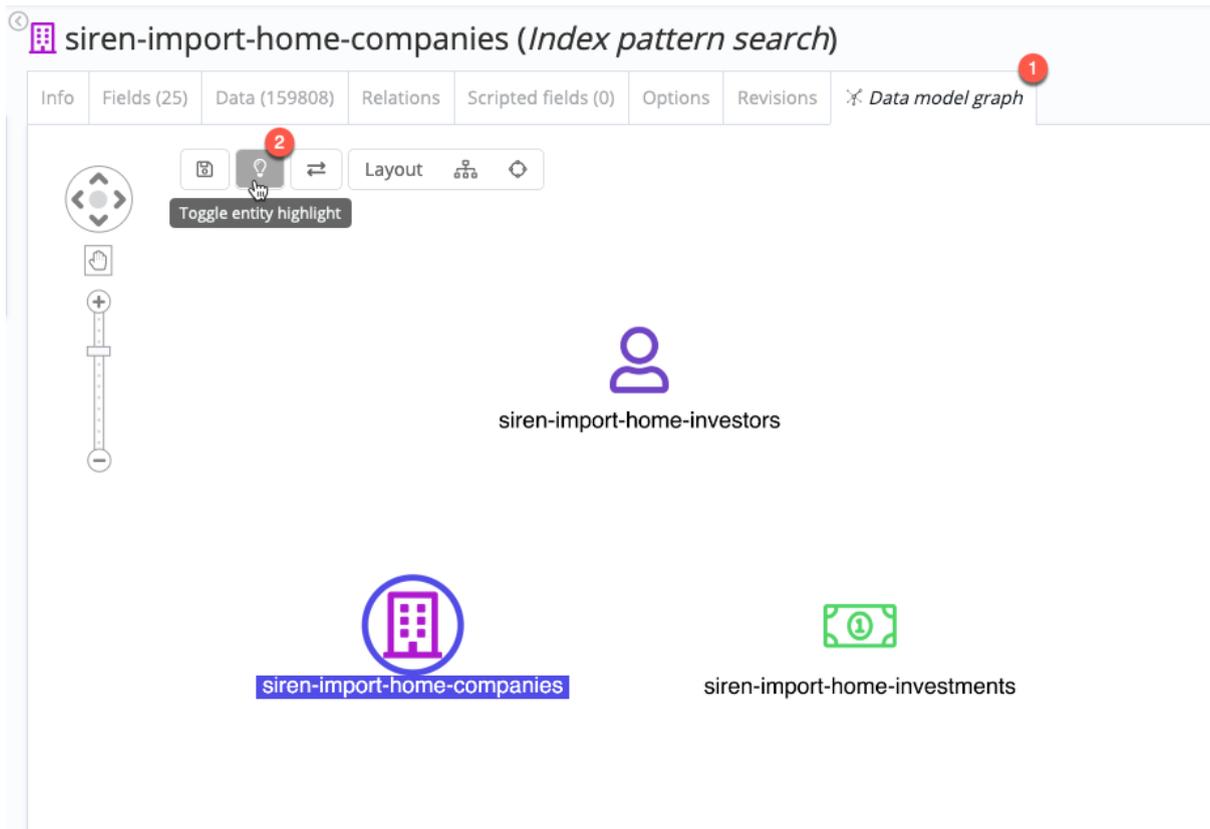
We will now learn how to define these connections in the **Relations** tab of the **Data model** app.

NOTE: The definition of a relation typically requires you to specify both active and passive verbal forms. For example, the companies *secured* the investments and the investments *were secured by* the companies.

The screenshot shows the configuration interface for a relation. On the left, under the 'Field' label, there is a dropdown menu with 'Companies' selected and 'id' chosen. In the center, there are two dropdown menus for the relation type: the top one is 'secured' and the bottom one is 'secured by'. On the right, under the 'Field' label, there is a dropdown menu with 'Investments' selected and 'companies' chosen. To the right of the dropdowns are icons for edit, refresh, and delete.

## Creating relations between entities

In the **Data model** app. Select the **siren-import-home-companies** search and click the **Data model graph** tab ①.



There are three icons, one for each search. To see all of the icons highlighted clearly, click the **Toggle entity highlight** - the lightbulb icon ②. You can see that the indexes are not connected to each other.

We're now going to define the relationship between the indexes and create a relational data model. This can be done manually or by using the auto-discovery wizard.

## Creating relations manually

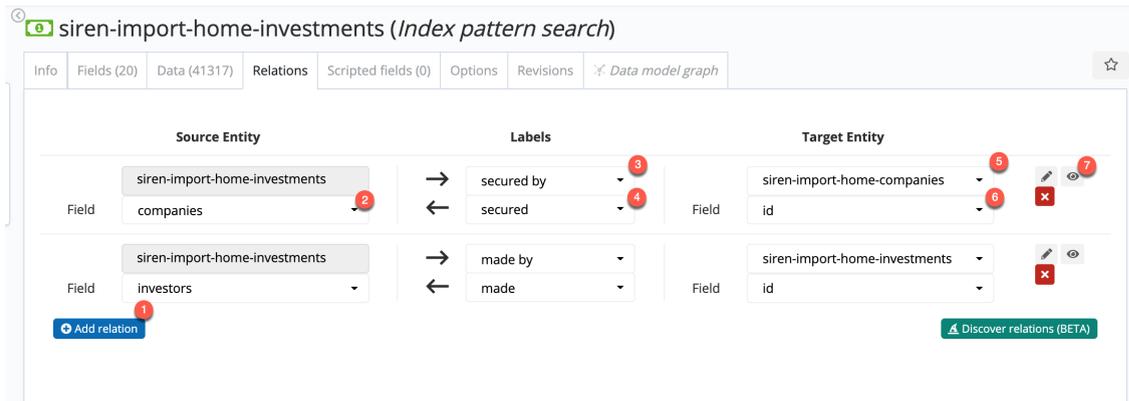
(Average completion time: 5 mins)

Siren Platform provides two methods of creating relations; you can either create them manually or you can use the automated discovery wizard. For more information, see [Appendix 7](#).

In this step, we will work with the **investments** saved search to specify which fields of other searches relate to it. Specifically, we will connect the **investments.companies** field with the **companies.id** field, the primary key of companies. We will base this on the [relational model of the data](#).

To create this relation, go to the **siren-import-home-investments** index pattern search, click the **Relations** tab, and complete the following steps:

1. Click **Add relation** ①. The source entity is preselected as **siren-import-home-investments**.



2. In the **Field** field, select **companies** ②.
3. In the **Labels** fields, enter “secured by” in the top ‘active’ field ③ and “secured” in the bottom ‘passive’ field ④.
4. In the **Target Entity** field ⑤, select **siren-import-home-companies** from the list of searches.
5. In the **Field** field ⑥, select **id**.
6. Click the **Show samples for this relation** ⑦. The **Relation Explorer** screen shows you sample records from the left and right side, which are matched on the same key. This is a good way to validate that the relationship makes sense. You can move back and forth through documents by clicking the arrows on the top left of the window.
7. Click **Save**.

Next, we need to create a relation to the investor.

1. Click **Add relation**.
2. In the **Field** field, select **investors** from the list of searches.
3. In the **Labels** fields, enter “made by” in the top field and “made” in the bottom field.
4. In the **Target Entity** field, select **siren-import-home-investors** from the list of searches.
5. In the **Field** field, select **id** from the list of fields.
6. Click **Save**.

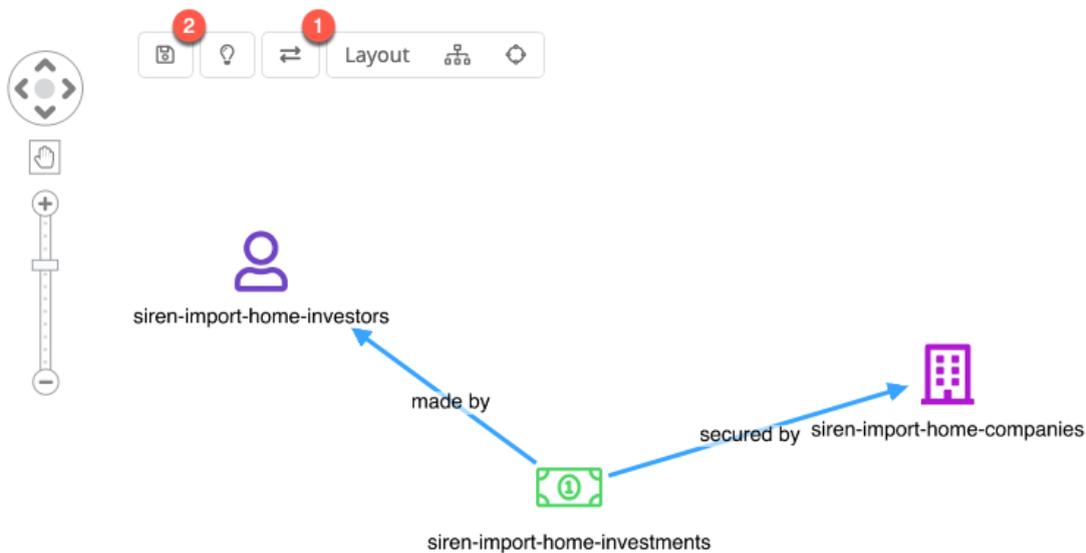
## Visualizing the data model as a graph

(Average completion time: 1 min)

To see all the relations in the data model at a glance, click on the **Data model graph** tab.

From this screen, you can:

- Click and drag the icons to rearrange them. Change the direction of the relation labels by clicking **Toggle relation direction** ①.
- **Save the graph layout** ②.



## Creating a self-relation

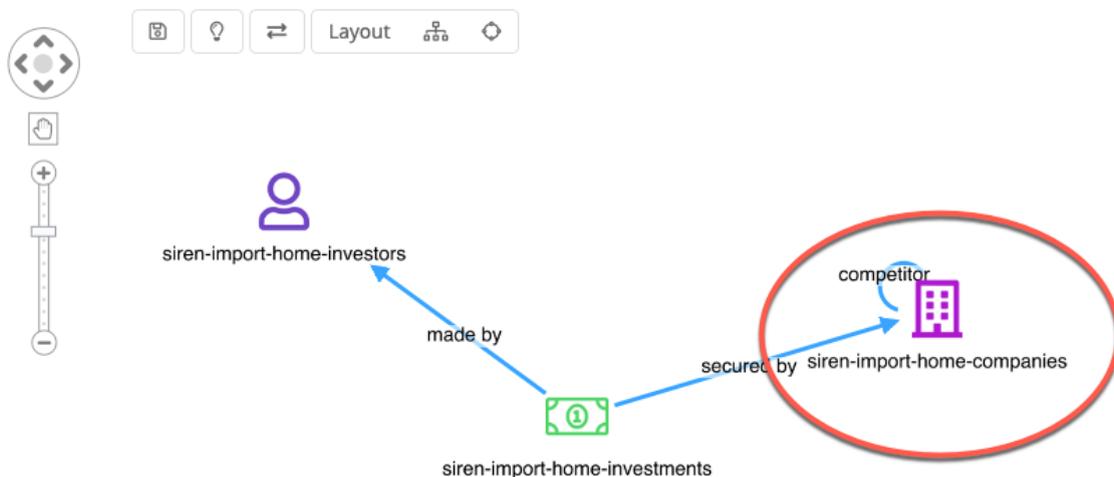
(Average completion time: 2 mins)

A self-relation is a relation between two entities of the same type, for example, a person who is *friends with* other people or companies that are *competitors of* other companies. These are defined in the same way.

In our data, the company records have a field called **one\_competitor** where you can find the ID of the main competitor of a company. Let's create a self-relation from **companies.id** to **companies.one\_competitor**:

1. On the **Data model graph** tab, double-click the **siren-import-home-companies** search icon. This switches to the **siren-import-home-companies** saved search.
2. Click on the **Relations** tab. You'll notice that the relation between companies and investments is already there. Makes sense, right? Now, we'll create a new one - this time a **self-relation**.
3. Click **Add relation**.
4. Specify a relation from the **id** field of the **\*companies** search to the **one\_competitor** field and enter "competitor" as the label in both directions.
5. Click **Save**.

With the self-relation, the data model graph should now look something like this:



That's it! Your relational data model is done for now. It's time to see it in action.

## Dashboard-to-dashboard associative navigation

The **Relational Navigator** is a special visualization that automatically shows links to connected records in other dashboards. By clicking on these buttons you perform an *associative* navigation, going from one set of records to another set of connected records.

For these buttons to appear in a dashboard, we must add a **Relational Navigator** visualization, the same way we add all visualizations to dashboards.

**NOTE:** You can create a single **Relational Navigator** visualization and reuse the same visualization across all of the dashboards. Once it is added to a dashboard, it will show buttons that connect with other dashboards that are ‘relationally’ connected.

The dashboard generation wizard adds this visualization if - and only if - you have defined the relations between the data. Otherwise, you can add it manually.

Since we have already defined the relations for ‘investors’, we can create a dashboard for it by using the wizard as follows:

1. In the **Data model** app, select the **siren-import-home-investors** search.
2. Follow the steps to [auto-generate the dashboard](#), specify the title of the dashboard as “Investors” and click **Create the dashboard**.

You can now see that the **Relational Navigator** visualization, labelled **Related dashboards**, is added to the **Investors** dashboard.

The blue link in this visualization links to the **Investments** dashboard and can be used to find all of the investments that were made by the *investors* on the current dashboard.

 Investors (14,959)

Filters

**Related dashboards**

[Made \(28723 Investments\)](#)

→

**List of affiliati**

**affiliation\_r**

Unaffiliated

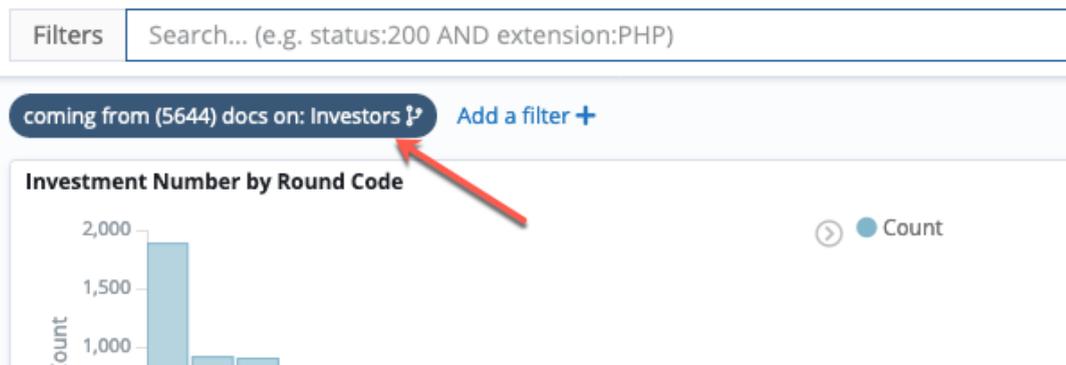
Google

Facebook

LinkedIn

3. Filtering the records on the dashboard will automatically show a filtered count on the relational navigator link. For example, if you click on the 'person' section of the **Top 30 investortype** pie chart. The record count on the dashboard and the count of investments are reduced as a result.
4. Click on **Made (4538 investments)** to go to the **Investments** dashboard. Due to the relational filter that is set, you will see *only investments that are associated with investors of type 'people'*.

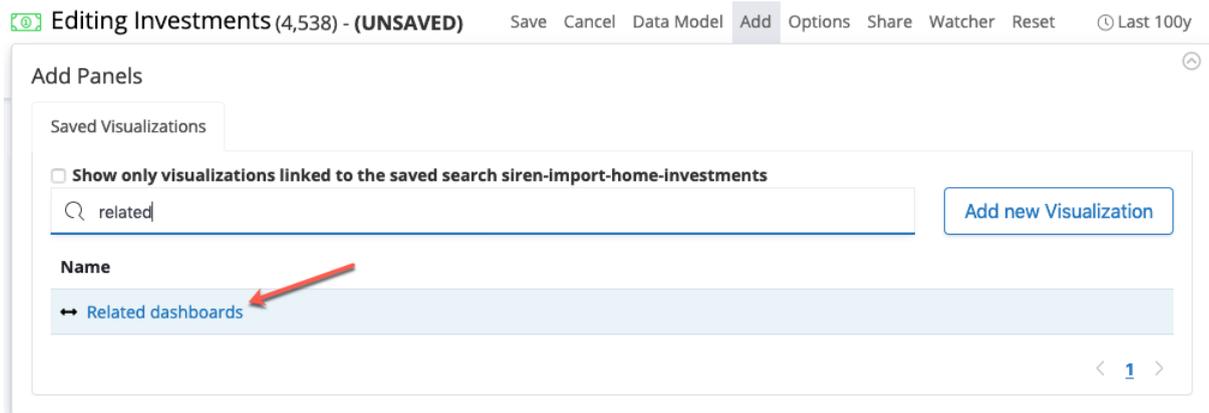
## Investments (4,538)



Now, we will add the **Relational Navigator** to the **Investments** dashboard to find more connected records.

You can do this by going to the dashboards that do not have this visualization yet (**Companies** and **Investments**), click **Edit** -> **Add**.

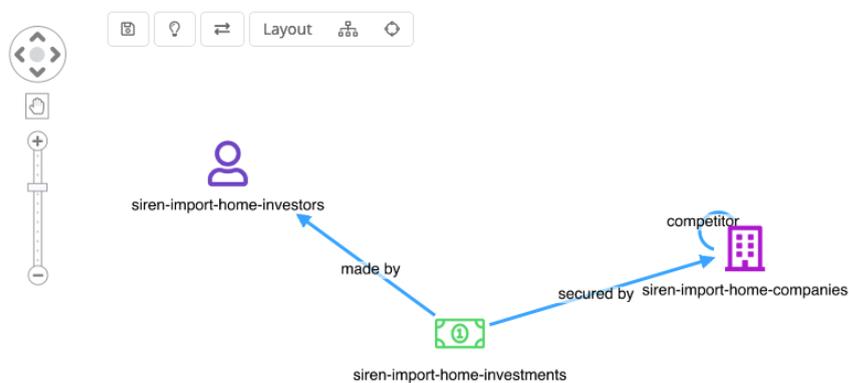
The **Related dashboards** visualization already exists in the system. Therefore, you can simply search for it in the list of existing visualizations, click to add it to the dashboard (a message indicates that it was added to the dashboard successfully), and click **Save** -> **Save dashboard**.



**TIP:** When you are adding panels to dashboards, it is easy to also save filters by mistake. Either remove the filters when you are saving a dashboard or select **Don't overwrite currently saved filter and text query** in the **Save** panel.

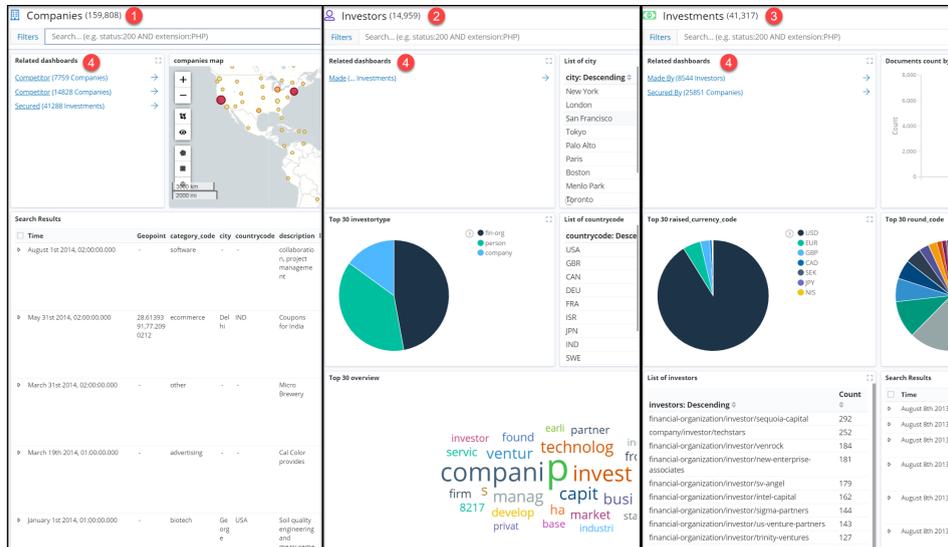
## Recap: Where are we now?

If you completed all of the previous steps of this tutorial successfully, your data model should look like this:



You should also have created three dashboards: Companies ①, Investors ② and Investments ③. The number shown after each dashboard name represents the number of documents in the relative Index Pattern Search.

Most importantly, each dashboard should contain a Relational Navigator visualization, called **Related dashboards** ④. You can rearrange the visualizations to bring this dashboard to the top.



If everything looks correct, we can now proceed to explore the data in more detail by completing some exercises on each dashboard and visualizing the results in the **Graph Browser**.

## Exercise: Investigating investments

In this section, we will answer a few advanced questions by analyzing the data.

**NOTE:** At any time, you can click the **Reset** button to restore the dashboard to its saved state. Alternatively, you can click the **Reset all** button to restore the state of all of the dashboards.

This exercise contains four parts, in which we will:

1. Answer a basic question by using the **Investors** dashboard.
2. Answer a question that involves a *relational navigation*, from **Investors** to **Investments**.
3. Navigate further to the **Companies** dashboard.
4. Use a link analysis to explore the results together as a single picture.

## Part 1: Finding an answer by filtering the Investors dashboard

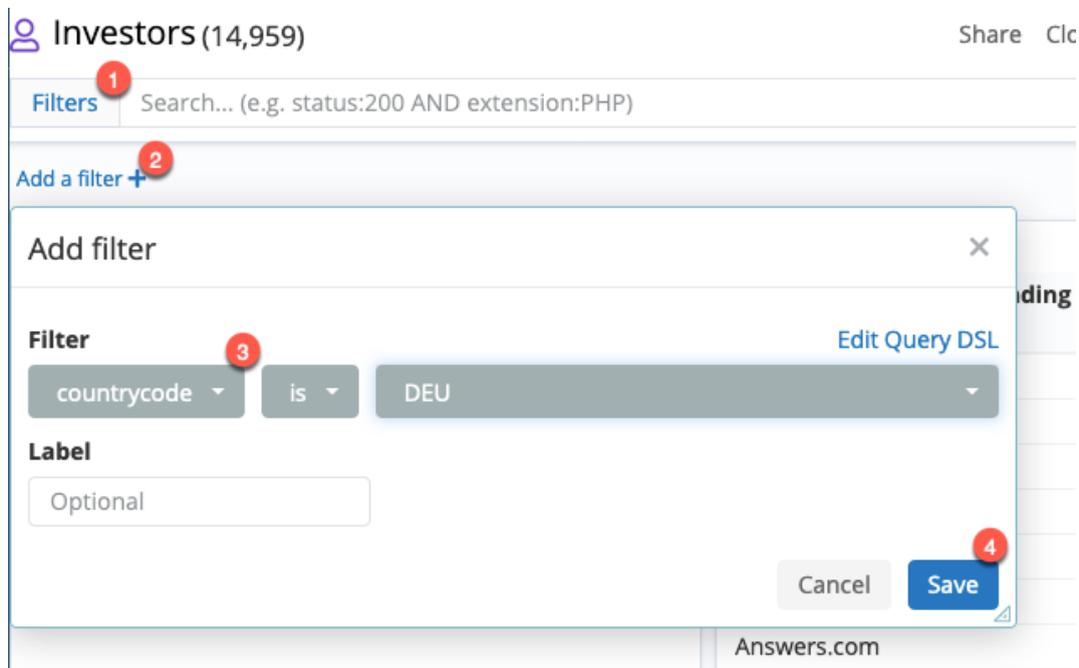
### Question: How many investors are from Germany?

Answer:

#### **Basic solution**

To find the answer to this question:

1. Go to the **Investors** dashboard.
2. Click **Filters** ① -> **Add a filter** ②.
3. Select **countrycode**, **is** and **DEU** (the iso code for Germany) from the drop-down menus ③.



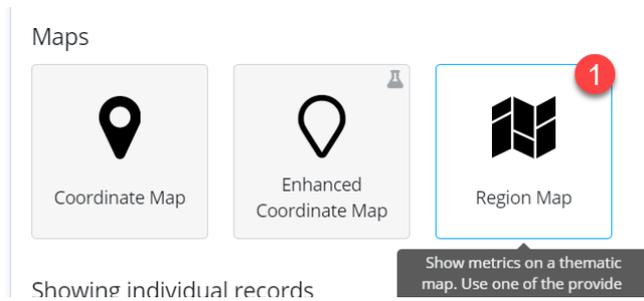
4. Click **Save** ④. The answer appears as the number in parentheses alongside the dashboard name.

#### **Advanced solution** (optional)

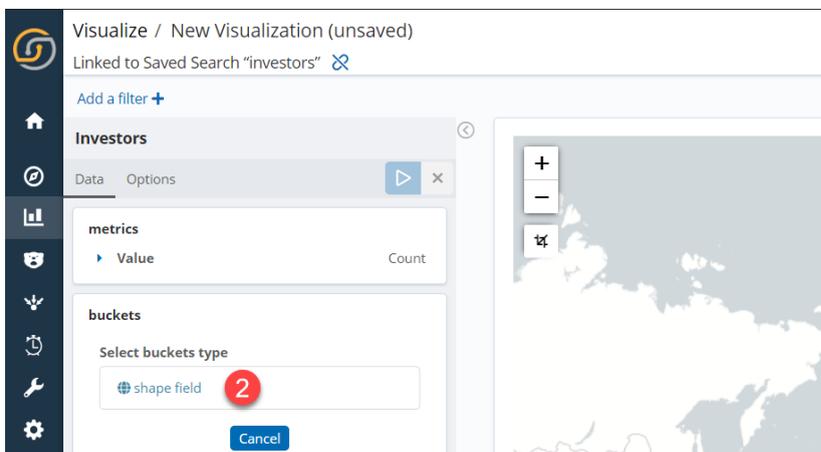
You can also answer this question by adding a **Region Map** visualization.

A **Region Map** is a clickable map that allows users to create filters in the dashboard in an interactive fashion. Unlike the **Enhanced Coordinate Map** visualization that works using geopoints (latitudes and longitudes stored in the records), the Region Map works by matching field values, such as **countrycode:deu**, with shapes and their names.

1. As previously shown in the [Creating visualizations](#) section, click **Edit** -> **Add** -> **Add a visualization**. Under the map category, choose **Region Map** ①.

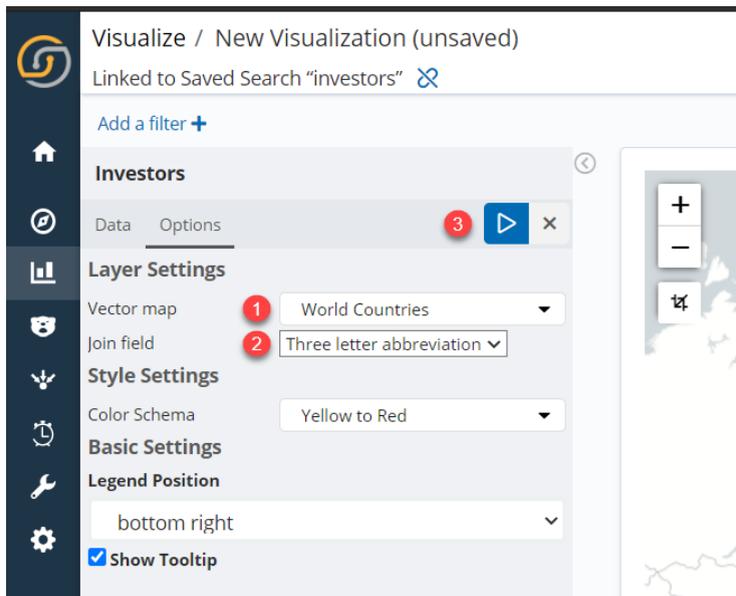


2. In the **Pick a saved search** window, select the **siren-import-home-investors** search.
3. In the **buckets** table window, select **shape field** ②.



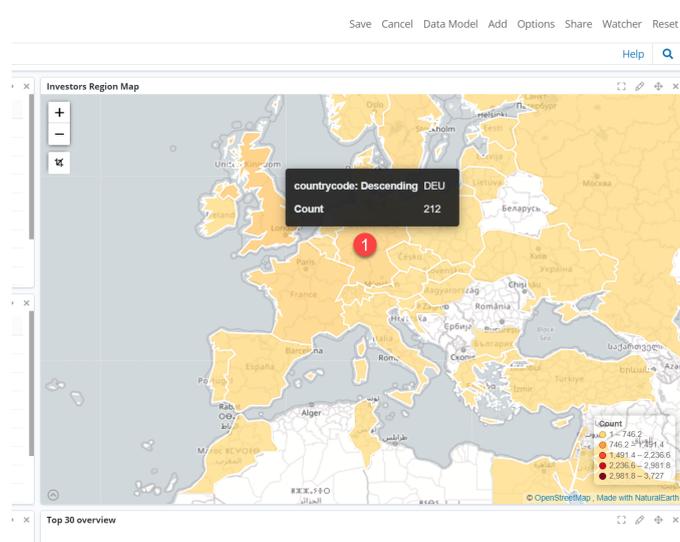
4. In the **Field** field, select **countrycode**.
5. We want to see data from more than five countries, so in the **Size** field, increase the value to **35**.

- Click the **Options** tab to select the proper layer settings to match the region type of our data.



- In the **Vector map** dropdown menu, select **World Countries** ① and in the **Join field**, select **Three letter abbreviation** ②.
- Finally, click **Apply changes** ③.
- Click **Save**, name the visualization “Investors Region Map”, and click **Save to Dashboard**.
- Move the Region Map into the position you want and save the **Investors** dashboard.

**HINT:** Hover your cursor over Germany. The **count** value appears in the tooltip.



To create a filter, click on Germany (DEU). The dashboard count will change accordingly and give you your answer.

**Q:** How many investors are from Germany?

**A:** 212

## Part 2: Finding an answer by using relational navigation

**Question:** How much money did German investors invest between 2010 and 2012?

Answer:

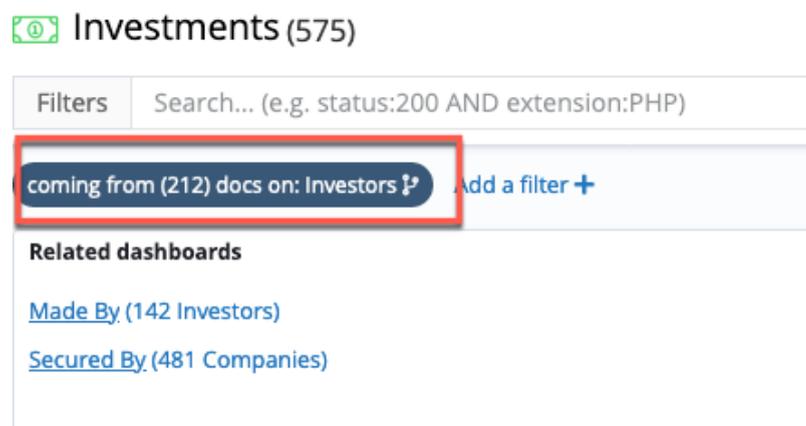
### Solution

Now that we have only German investors in the dashboard, we can use the **Relational Navigator** to go to the **Investments** dashboard.

The button already shows you the number of target records: Out of 41k unfiltered investments, only 575 were made by German investors.

Let's go ahead and click on **Made (575 Investments)**.

On the **Investments** dashboard, a relational filter is automatically created:



However, we cannot find the answer to the question yet, because we are missing a filter for the two-year time window and a metric to compute the total funding amount.

Let's fix that!

1. To change the time window, click **Last 100y** ①.
2. In the **Time Range** panel, select **Absolute** ② and specify the desired time span ③ (2010-01-01 00:00:00.000 - 2012-01-01 00:00:00.000).
3. Click **Apply** ④.

The screenshot shows the Siren Platform interface for the 'Investments (575)' dashboard. The 'Time Range for 'funded\_date'' panel is open, showing three options: 'Quick', 'Relative', and 'Absolute'. The 'Absolute' option is selected and highlighted in orange. The 'From' field is set to '2010-01-01 00:00:00.000' and the 'To' field is set to '2012-01-01 00:00:00.000'. Below these fields are two calendar views for January 2010 and January 2012. The 'Apply' button is highlighted in blue. In the top right corner, the 'Last 100y' button is also highlighted. The 'Apply to dashboards' section shows 'investments (current)' selected.

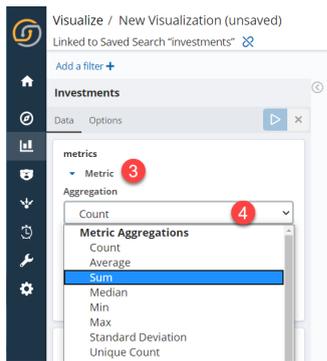
Perfect! We only need one last step to answer this question, which is, adding a metric to compute the total amount of the 176 investments that are left.

4. Click **Edit** -> **Add** -> **Add a visualization**. Under the Textual category, choose **Metric** ①.

The screenshot shows the 'Add a visualization' dialog box. The 'Textual' category is selected. There are three options: 'Analytic Table', 'Markdown', and 'Metric'. The 'Metric' option is highlighted with a red circle ①. The 'Metric' option shows a large number '42' and a tooltip that says 'Display a calculation as a single number'. Below the dialog box, the 'Maps' category is visible.

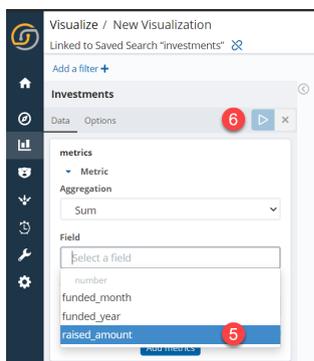
5. From the **Pick a saved search** menu, select **siren-import-home-investments**.

6. In the **metrics** table window, expand the **Metric** parameter setting ③.



7. In the **Aggregation** field, select **Sum** ④.

8. In the **Field** field, select **raised\_amount** ⑤.

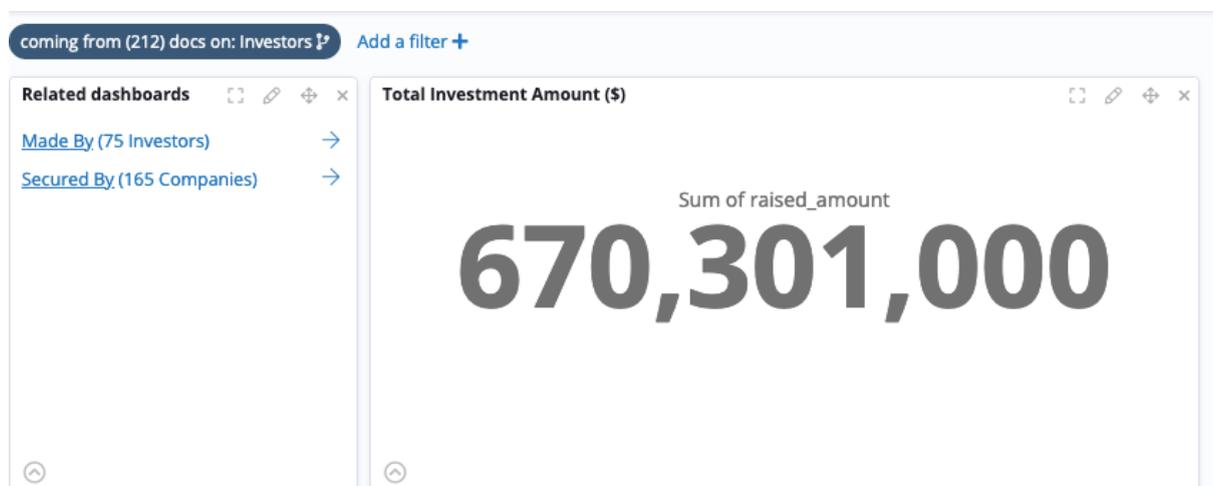


9. Click **Apply changes** ⑥.

10. Save the visualization as “Total Investment Amount (\$)”.

11. Move the new visualization into position and save the **Investments** dashboard. This final step will provide us with the answer to the exercise question and you'll permanently have access to this useful metric.

**Q: How much money did German stockholders invest between 2010 and 2012?**



**A:** German stockholders invested ~670M \$ between 2010 and 2012.

### Part 3: Digging deeper with relational navigation

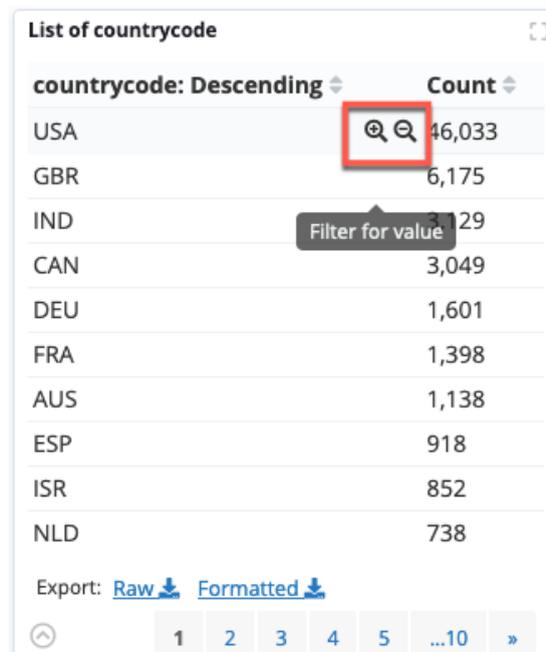
**Question: How much of these investments were secured by American companies?**

Answer:

#### **Solution**

To solve this question, complete the following steps:

1. Click on the **Companies** dashboard in the sidebar (do not navigate there by using the relational navigator). This will show you all of the companies (approx. 159k).
2. Create a filter for country code **USA** by clicking on **Filter for value** (the magnifying glass icon) in the **List of countrycode** visualization.



countrycode: Descending	Count
USA	46,033
GBR	6,175
IND	29
CAN	3,049
DEU	1,601
FRA	1,398
AUS	1,138
ESP	918
ISR	852
NLD	738

Export: [Raw](#) [Formatted](#)

1 2 3 4 5 ...10 »

Now, we have selected only American companies.

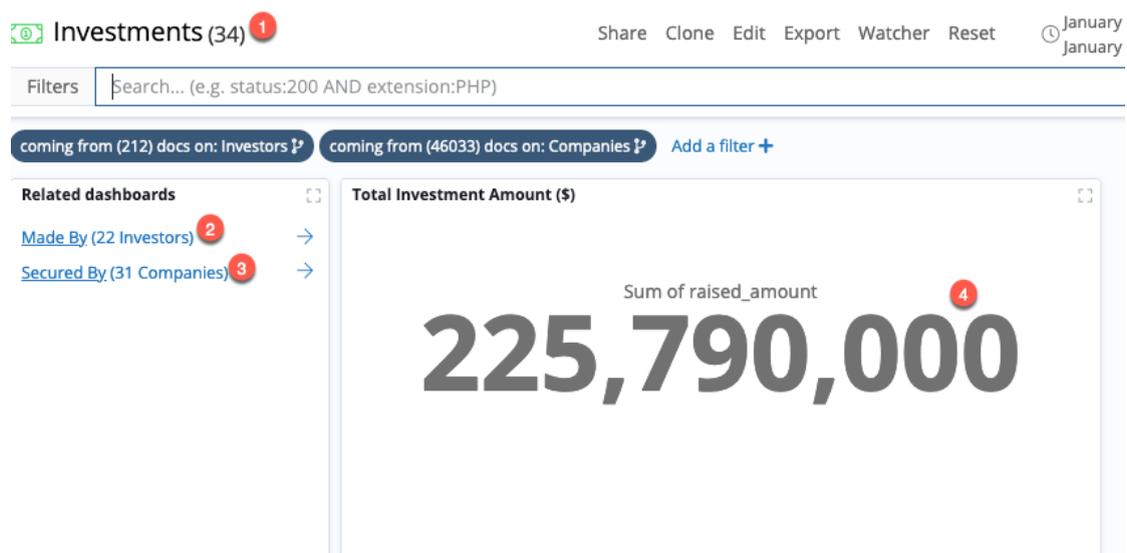
The **Relational Navigator** is showing us that **34** of the records in the **Investments dashboard** are matching. Notice that, at this point, the Investments dashboard is already filtered from our previous exercise (it contains the filter that displays only the investments by German investors).

So, let's click on **Secured (34 investments)** to reveal the investments.



You will be brought to the **Investments** dashboard again but, this time, there are two relational filters present:

- One filter indicates that we are looking at investments “from German Investors” (212);
- The other filter indicates that they were “secured by American companies” (46033).



As you can see in the above screenshot, these 34 investments ① - made by 22 German investors ② and secured by 31 American companies ③ - come to a total of ~\$225M ④.

**Q:** What amount of these investments were secured by American companies?

**A:** ~225M \$

## Part 4: Link analysis in the Graph Browser

### Exploring the results with the Graph Browser

Now that we have solved all of the questions, it is time to explore the results by using the link analysis feature in the **Graph Browser** visualization.

Before doing so, let's apply the relational filter on the other dashboard to focus our research around German Investments, Investors and American Companies.

1. Click **Secured By (31 Companies)** and take a moment to see who these companies are in the relative dashboard.
2. When you are ready, go back to the **Investments** dashboard and, this time, click **Made By (22 Investors)**.

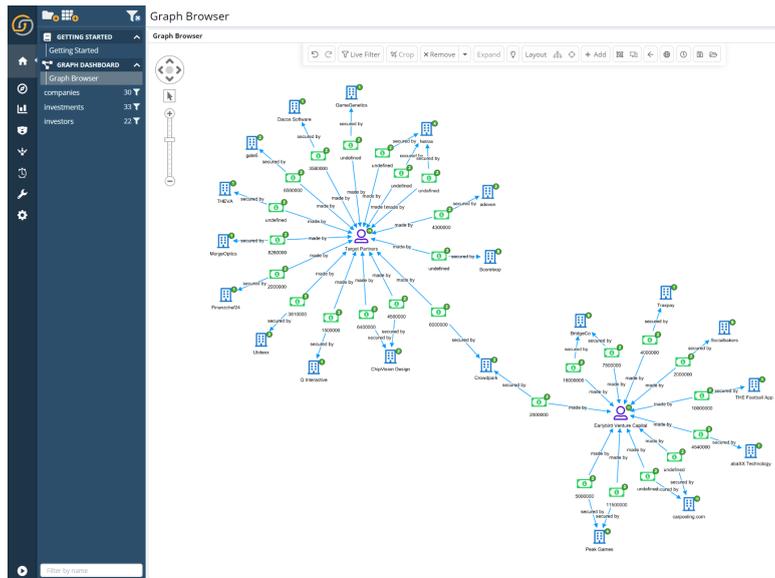
If everything was done correctly, your filtered dashboards should now display:

- 31 Companies (all of them American).
- 34 Investments (made by German investors and secured by the 30 American companies)
- 22 Investors (German stockholders who invested in American Companies)

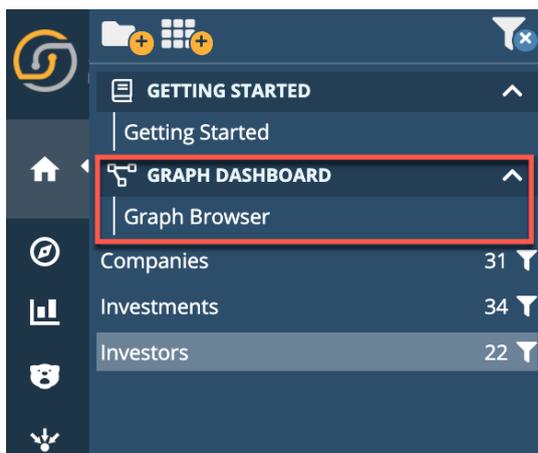
**CAUTION:** If your dashboards don't have the correct filters applied and a large number of nodes are added, you might experience problems with system performance.

Let's explore together how the Companies, Investments and Investors that we found in our investigation are linked together and if there are some interesting common connections.

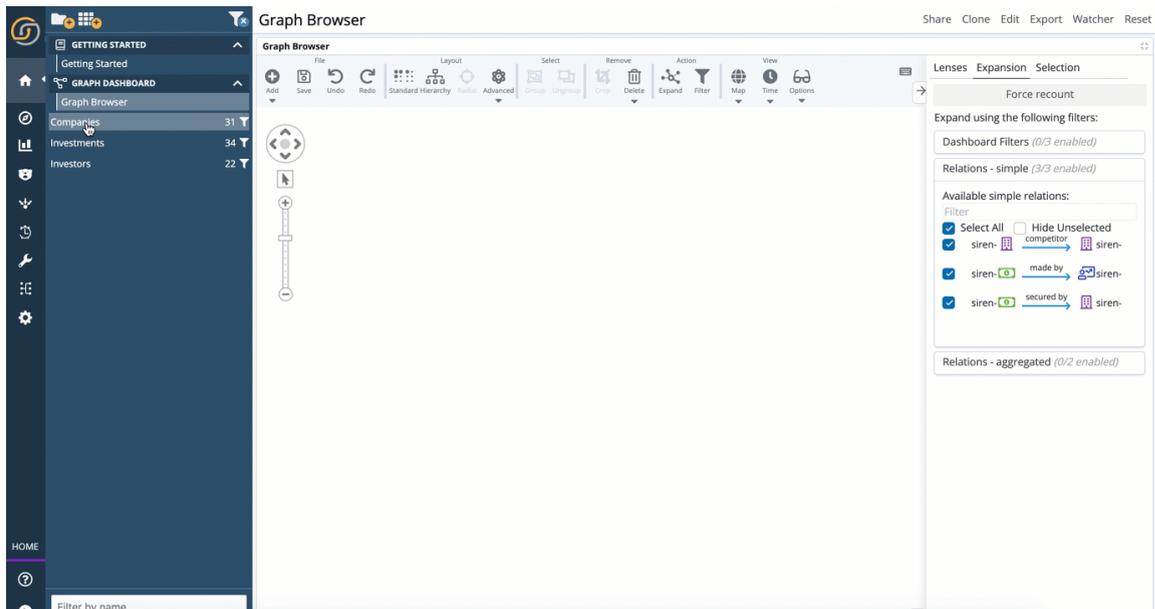
In this section of the tutorial, we'll get to see graphs like these:



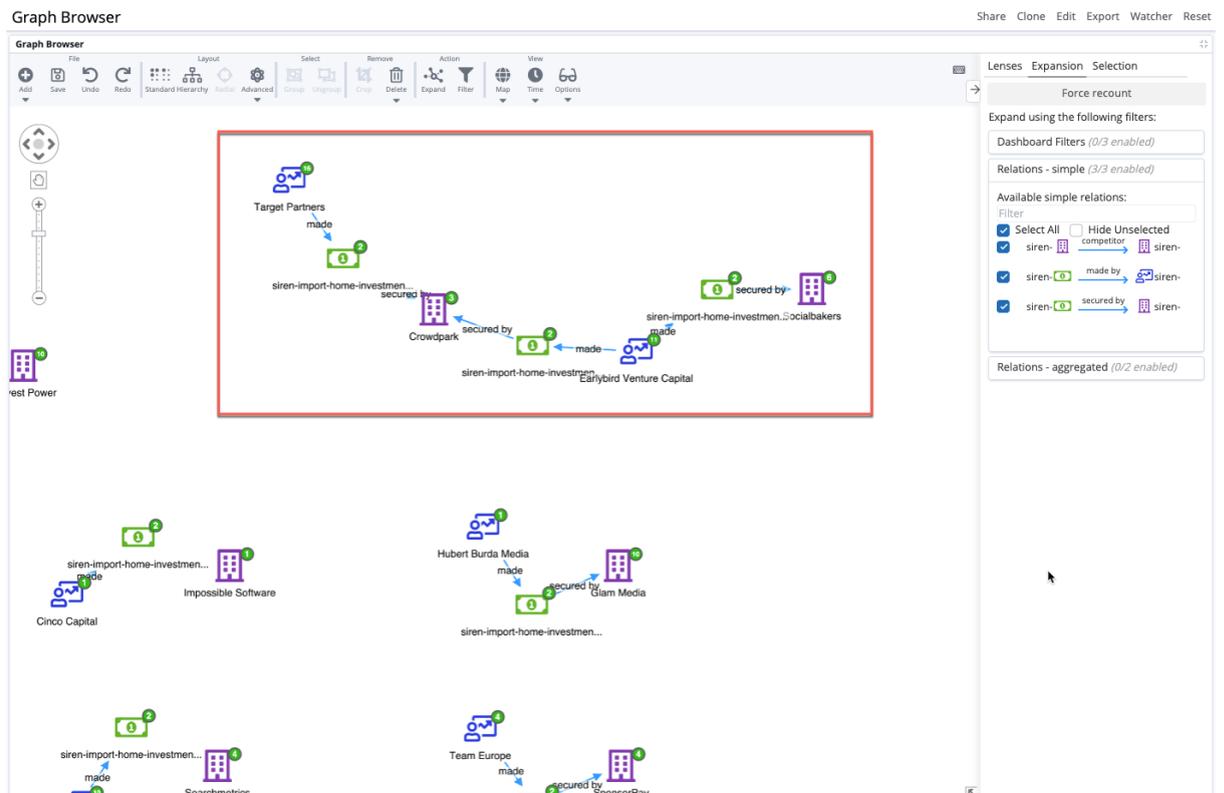
1. To start exploring the data, expand the **GRAPH DASHBOARD** group and click **Graph Browser**.



2. Drag and drop the **companies**, **investments** and **investors** dashboards one-by-one into the Graph Browser.

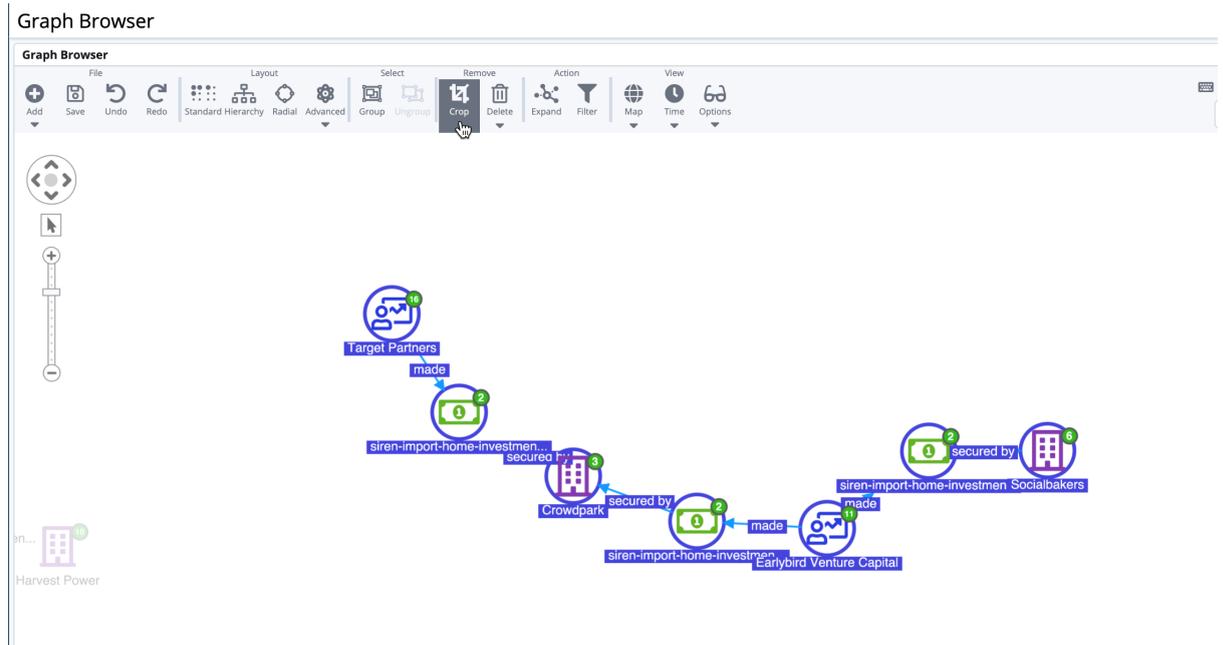


3. Click **Standard** from the Layout menu and the nodes will be rearranged. We can observe, from this layout, that two German stockholders invested in the same company, named *Crowdpark*. Let's explore this small but interesting network.



4. Set the cursor to selection mode, drag a rectangle around the network that was identified in the previous screenshot and click **Crop**. This will remove everything else

from the Graph Browser.



5. Before expanding this network, it is good practice to save this initial state. To do so, click **Save** and give the graph a suitable name, such as “Shared investment”.

Now that we have saved our initial network, we can start asking more interesting questions.

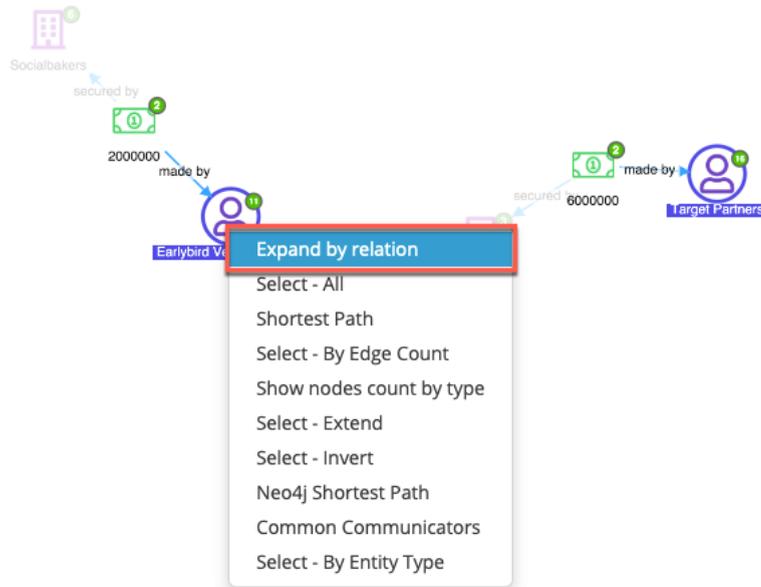
Before we proceed with the tutorial, take a moment to consider the links that might be interesting to explore. Any ideas?

We selected this network because the two German investors share an American company in their investment campaign. It would be interesting to uncover that:

1. They have more common investments.
2. They invested in CrowdPark (the funded company) in the same time period.

Let's start by investigating the first point:

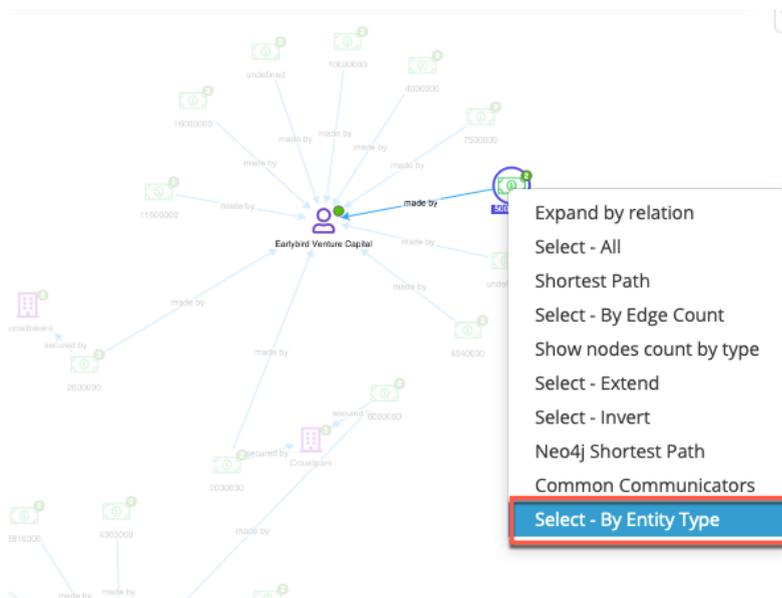
1. While in selection mode, hold the **Ctrl** key on the keyboard and click on the two investor nodes.
2. Once the two nodes are selected, right-click on one of them and, from the contextual menu, select **Expand by relation**.



3. Select the **made (27) siren-import-home-investments** checkbox and click **Ok**.

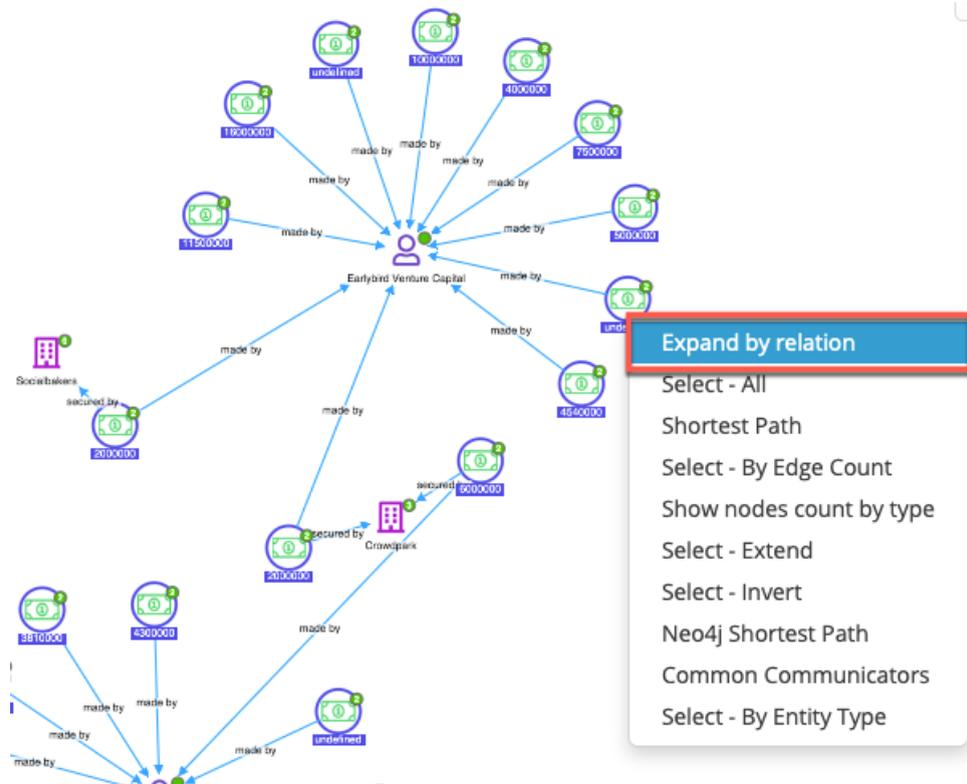
We will now investigate whether one of the new investments that appeared is paid to a new common company.

4. To do so, right-click any one of the investment nodes and click **Select - By Entity Type**.



5. In the pop-up window, select only **siren-home-import-investments** and click **Ok**.

- Now, all of the investment nodes should be selected. Right-click on one of the nodes and select **Expand by relation**.

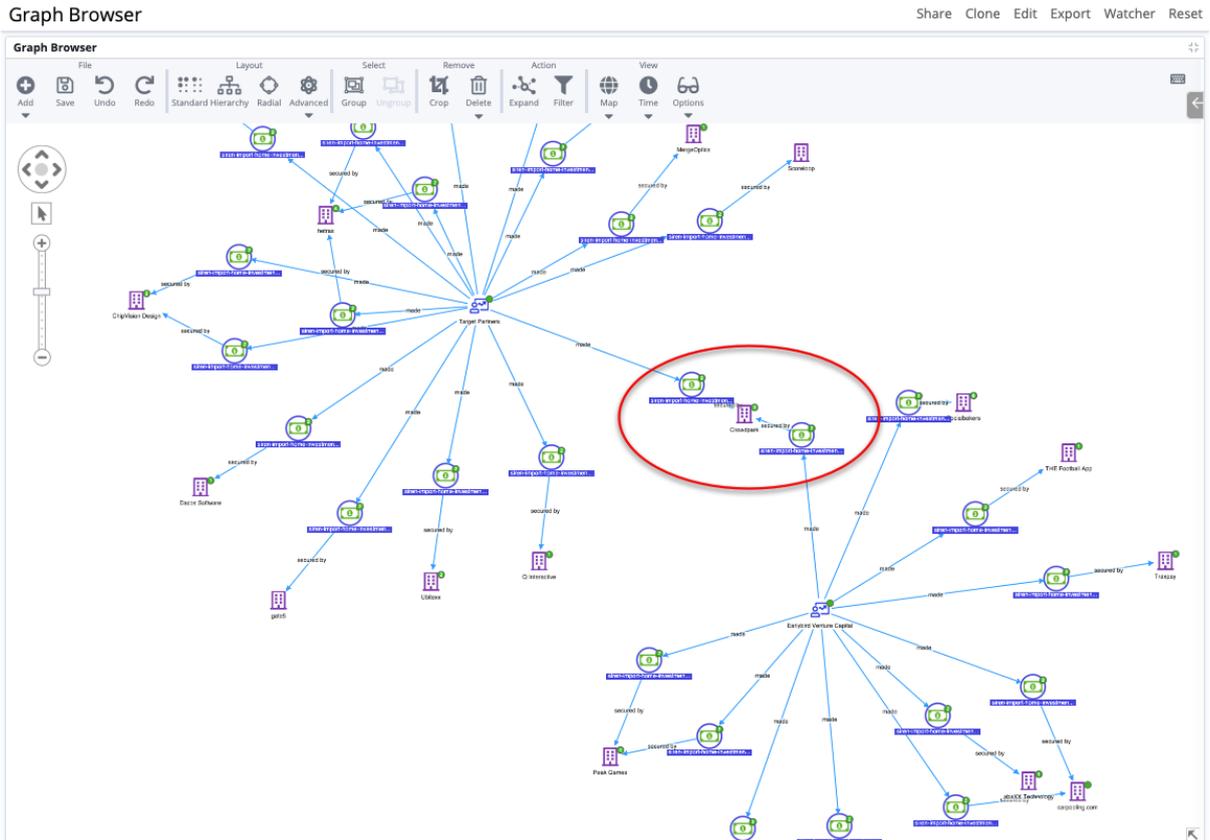


- In the pop-up window, select **secured by (27) siren-home-import-companies** and click **Ok**.

As we can see in the graph below, the only common connection between the two German investors (from 2010 to 2012) is *CrowdPark*.

**Q:** Do the two German stockholders share additional companies in their funding campaign?

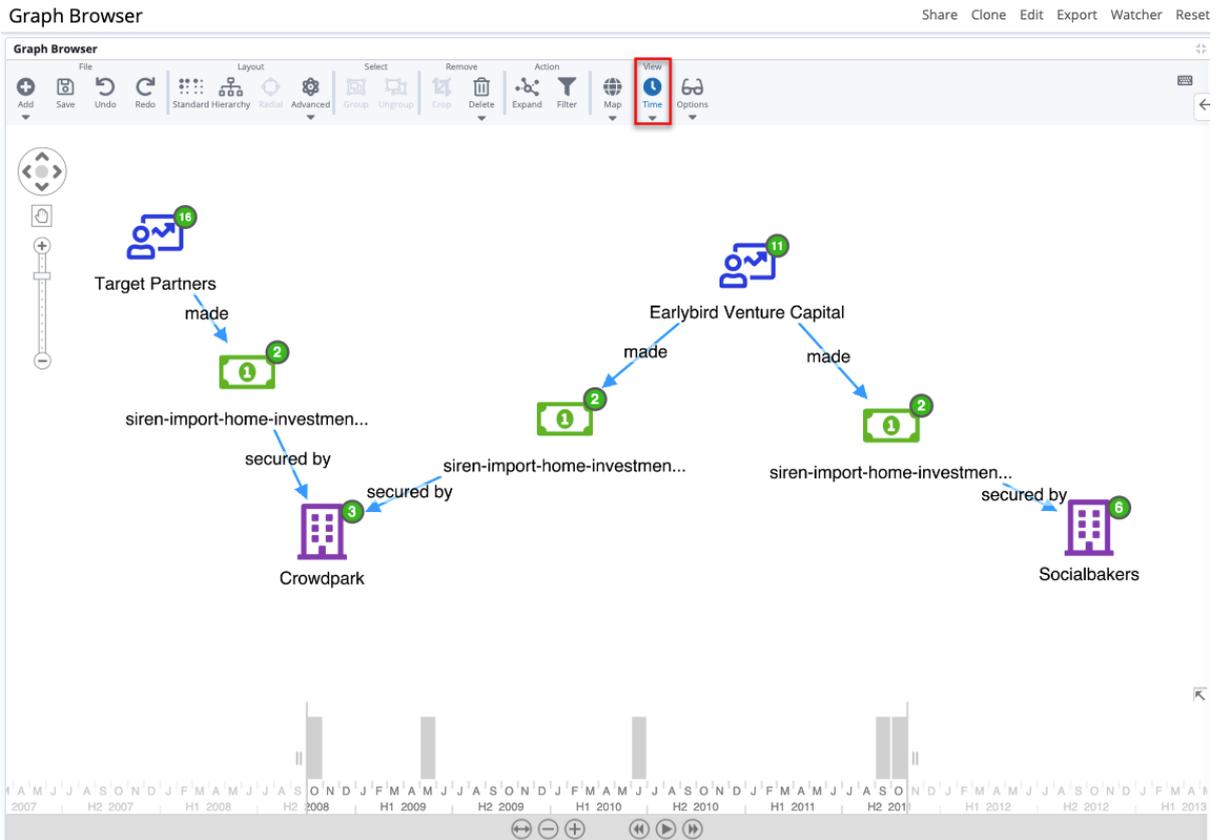
**A:** No.



**NOTE:** If you notice that the labels of the entities are not displaying correctly, ensure that you completed all of the steps to set the labels in your index pattern searches, as described in steps 7-8 of [Creating your first index pattern search](#).

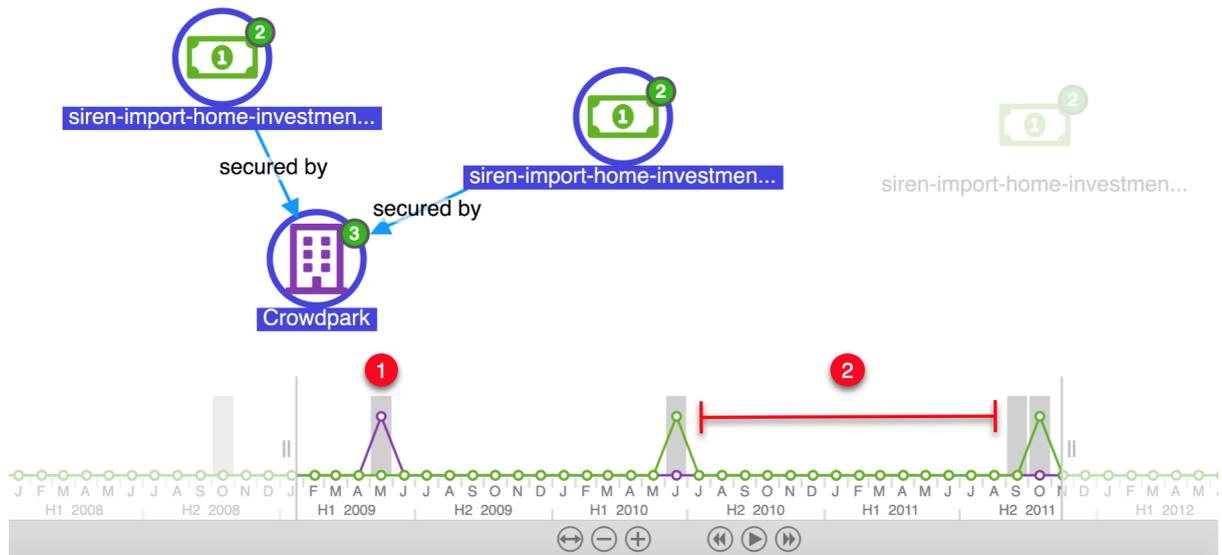
The second interesting question was: Did the two German investors fund CrowdPark in the same time period?

1. Remove the nodes and return the graph to its saved state by clicking **Delete -> Delete All**.
2. Click **Add -> Open saved graph** and select the saved graph, **Shared investments**.
3. In the **Open saved graph** window, select **Replace the current graph with the saved graph** and click **Add nodes**.
4. Activate the timeline mode functionality of the graph by clicking **Time**.



In timeline mode, hold the **Shift** key and select the nodes that we are interested in; the company, Crowdpark, and the two investment nodes. This will allow us to explore the temporal dynamics of the selected nodes.

As we can see in the screenshot below, Crowdpark was founded in the second quarter of 2009, while the two investments were made approximately one year apart.



**Q:** Did the two German investors fund CrowdPark in the same time period?

**A:** No.

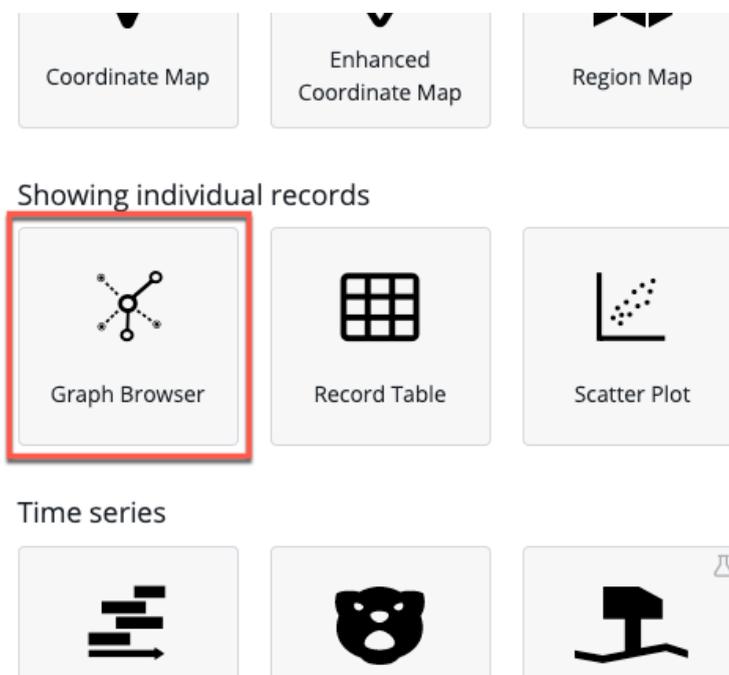
This concludes our exercise! If you want to know more about the graph, we will cover more material in the next chapter.

# More about link analysis

## Embedded in dashboards

Link analysis is one of the key features and can be performed in the **Graph Browser** visualization.

In the exercise, we used a pre-configured dashboard with a **Graph Browser** component, however, you can add the Graph Browser to a dashboard by selecting it from the **Visualizations** menu:

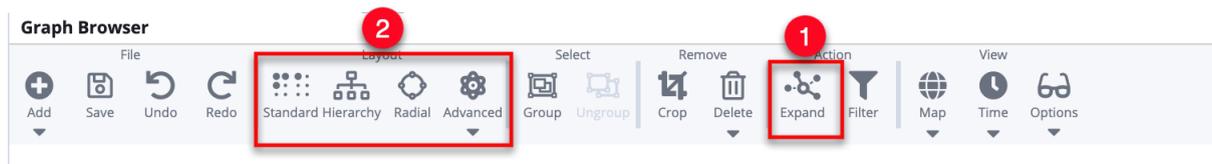


## Using the Graph Browser

You can add data to the Graph Browser by dragging and dropping filtered or unfiltered data from dashboards.

You can expand the data by double-clicking on any node but also by using the **Expand** button ①, which acts on all of the selected nodes.

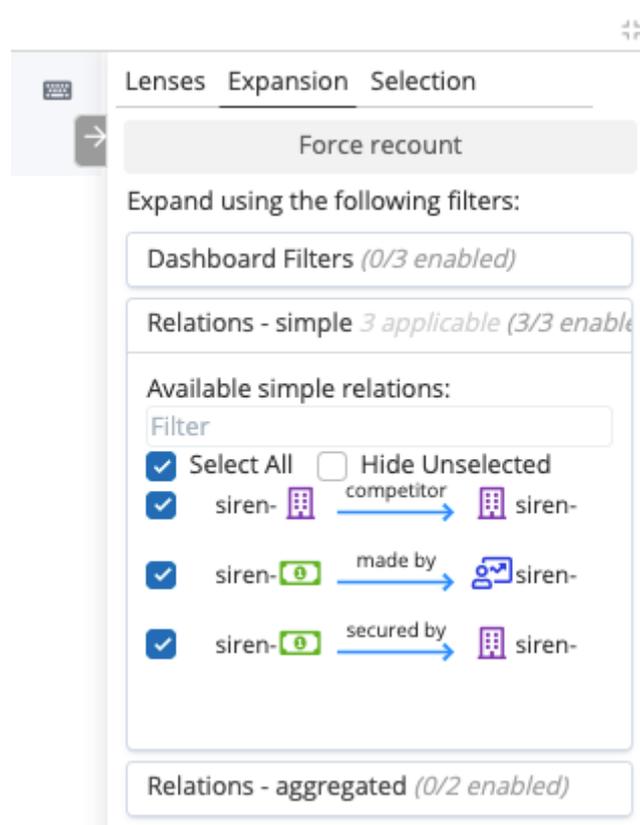
Clicking one of the **Layout** buttons ② reorganizes the nodes.



The Graph Browser also offers some advanced features to help with link analysis.

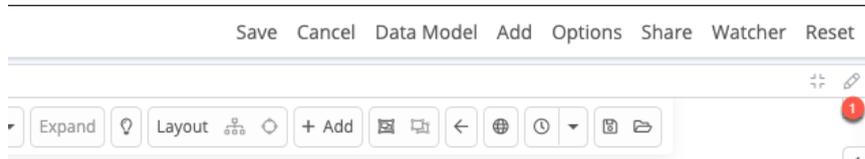
The side bar contains the following tabs by default:

- **Lenses** - Create lenses to alter the way the graph looks. For example, create conditional formattings for labels, colors, sizes, or change visibility. For more information about lenses, see the [Siren Platform documentation](#) or follow the steps to create a lens in [Appendix 2](#).
- **Expansion** - Controls the way the graph expands (either when you click the expand button or when you double-click on nodes). Here, you can set filters (by using Dashboard Filters) or choose the data model relation to consider when expanding. For more information, see the [Siren Platform documentation](#).
- **Selection** - See details about the current selection and search within the selection.



The **Cards** tab is not displayed by default, but you can add it by completing the following steps:

1. In the **Graph Browser**, click **Edit** in the top menu bar.
2. Click the pencil icon ① to open the **Edit visualization** menu.



3. Click the **Scripts** tab and, under the **Cards** section, click **Add card script** and select **Histogram Card - default implementation** from the dropdown menu.
4. Click **Apply changes** and **Save**. You might need to refresh the page for the changes to appear.

For more information about using the **Cards** tab, see the [Siren Platform documentation](#).

# Appendix 1: Natural Language Processing (NLP)

(Average completion time: 40 mins)

The Siren Platform Easy Start version comes with the Siren Platform NLP plugin pre-installed, which adds an NLP *pipeline processor* to enrich documents that are being ingested in Elasticsearch (by a CSV file import or by another method) with entity annotations<sup>2</sup>.

Here, we will use it to enrich the fourth CSV file in our demo dataset, the **articles.csv** file, which contains 640k tech articles that are harvested from the Web.

The pipeline processor will take a text field and produce annotations for Named Entities (Organization, Person, Location), which are then useful to connect to the data that we already have. For example, connecting articles to companies or investors.

To do this, we will activate a 'transformation pipeline' with the NLP processor, while importing data into the **articles** index.

## Before you begin

If you are using a slow machine, go to **Management** → **Advanced Settings** and edit the value for **ingest-xlsx:bulk\_package\_size** field to 20 and click **Save**.

## Procedure

1. Import the **articles.csv** file that is provided with the [sample data](#) as usual but, this time, on the **Configure Indexing** screen, in the **Transform Pipeline** dropdown

---

<sup>2</sup> It is always possible in Siren Platform to use any other NLP engine as part of ETL, but Siren NLP is made available via the Elasticsearch pipeline processor, which makes it easy to activate and requires no external server or process.

menu, select **Elasticsearch ingestion pipeline method** ①.

The screenshot shows the 'Transform Pipeline' configuration page. At the top, there are three progress steps: 'Choose a file', 'Configure Indexing', and 'Done!'. The 'Method' dropdown is set to 'Elasticsearch ingestion pipeline method'. The 'Elasticsearch pipeline definition' section contains the following JSON:

```

1 {
2   "processors": [
3     {
4       "script": {
5         "source": "if(ctx.snippet== null){ctx.snippet = ''}"
6       }
7     },
8     {
9       "siren-nlp": {
10        "fields": [
11          "snippet"
12        ]
13      }
14    },
15    {
16      "target_field": "snippet_nlp",
17      "processors": [
18        {
19          "class": "Telephone"
20        },
21        {
22          "class": "USTelephone"
23        }
24      ]
25    }
26  ]
27 }

```

Below the JSON, there is a 'Select Sample Record' dropdown set to 'Record 2' and a 'Test' button. The 'Test Output' section is empty.

2. Copy and paste the below JSON processor in the **Elasticsearch pipeline definition** ②.

```

{
  "processors": [
    {
      "siren-nlp": {
        "fields": [
          "snippet"
        ]
      },
      "processors": [
        {
          "class": "Telephone"
        },
        {
          "class": "USTelephone"
        },
        {
          "class": "Email"
        },
        {
          "class": "IPv4"
        },
        {
          "class": "IPv6"
        },
        {
          "class": "MacAddress"
        },
        {
          "class": "Url",
          "settings": {
            "lenient": "true"
          }
        },
        {
          "class": "SortCode"
        },
        {
          "class": "HashTag"
        },
        {
          "class": "NER",
          "settings": {
            "nerType": "Organization"
          }
        },
        {
          "class": "NER",
          "settings": {
            "nerType": "Person"
          }
        }
      ]
    }
  ]
}

```



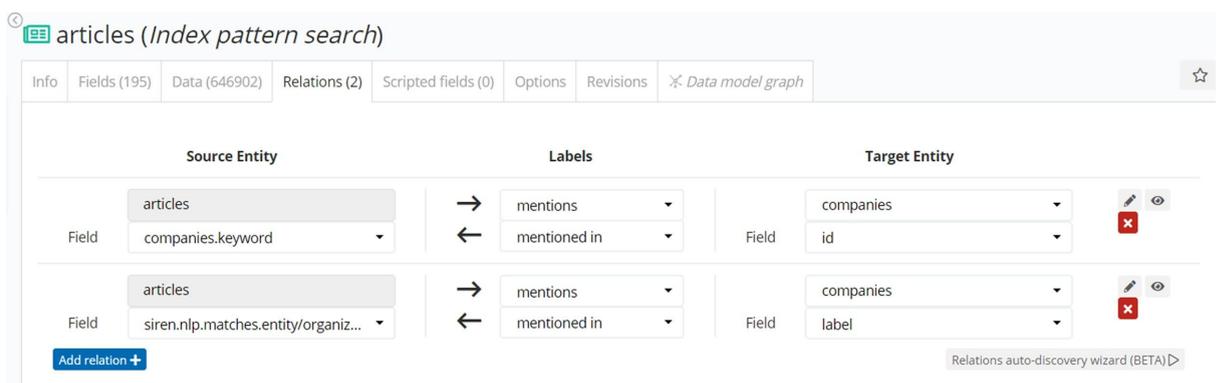
## Example usage of NLP Data in Siren Investigate

Data annotated by using the NLP plugin can be very useful for finding interesting relationships between fields in an index and the annotated text. In this example, we will try to find the mentions of companies as a part of the articles' snippet.

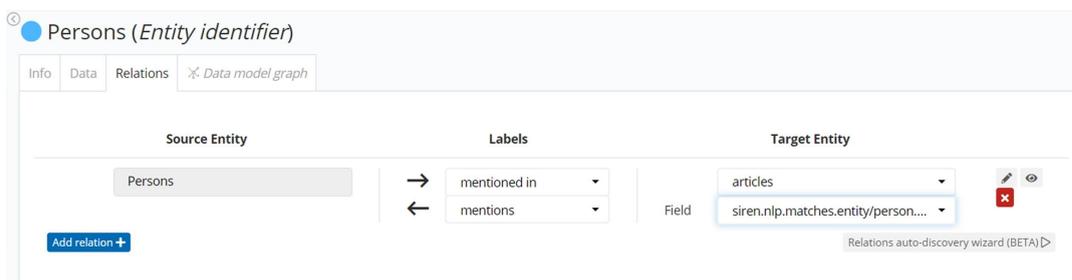
In the **Data model** app, click on the **Data** tab. Under the fields, you can see that a new field **siren.nlp** has been created by the NLP plugin during the import. This field contains annotations for fields processed by Siren NLP (in this case the **snippet** field), such as organization, person and location.

From the Data model page for the articles index, complete the following steps:

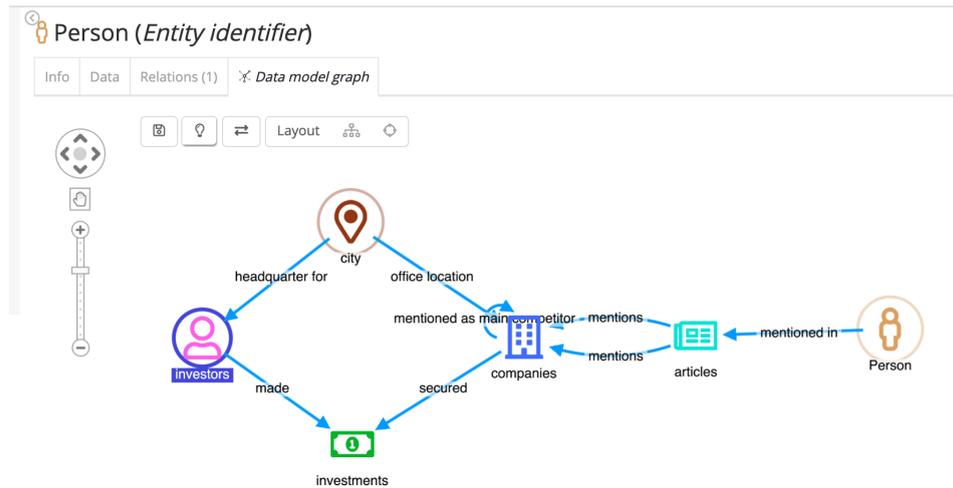
1. Create a relation between Articles **siren.nlp.matches.entity/organization** and Companies **label**.



2. Given that we don't have an index for "Persons", [create an Entity Identifier \(EID\)](#) and add a relation from it to **siren.nlp.matches.entity/person**.



The data model graph should look like this:



- Go to the **Companies** dashboard and search for the company “Huawei Technologies” ①.
- Filter the dashboard for the company with the label name “Huawei Technologies” by using the **Filter for value** button ②. This will filter the dashboard with a single record for this company.

The screenshot shows the 'companies' dashboard with 6,613 records. A filter for 'Huawei Technologies' is applied. The search results table is as follows:

label	category_code	city	blog_url	overview
Huawei Technologies	network_hosting	Shenzhen	-	<p>Huawei Technologies is a leading telecom solutions provider. Through continuous customer-centric innovation, they have established end-to-end advantages in Telecom Network Infrastructure, Application & Software, Professional Services and Devices. With comprehensive strengths in wireline, wireless and IP technologies, Huawei has gained a leading position in the All-IP convergence age. Their products and solutions have been deployed in over 100 countries and have served 45 of the world's top 50 telecom operators, as well as one third of the world's population.</p>

The dashboard also features a 'Top 50 category\_code' bar chart and a 'Top 30 hasstatus' section.

- Go to the **Graph Browser** and drag and drop the **Companies** dashboard there. This will add a node with the company Huawei Technologies.
- Select the node, right-click, and select **Expand by relation**.
- Select **mentioned in (20) articles** and click **Ok**.
- You will now see the articles that mention this company. Next, select all of the nodes on the graph and right-click → **Expand by relation** and select **mentions (3) Person** and **mentions (32) companies**.

9. This will give us an interesting graph that shows articles mentioning Huawei Technologies and also mention some **other companies**. The graph also shows mentions of **persons** who were named in articles that talk about Huawei Technologies.



## Appendix 2: Configuring the graph

You can modify the graph in a number of ways by using the features that are provided in the Graph Browser.

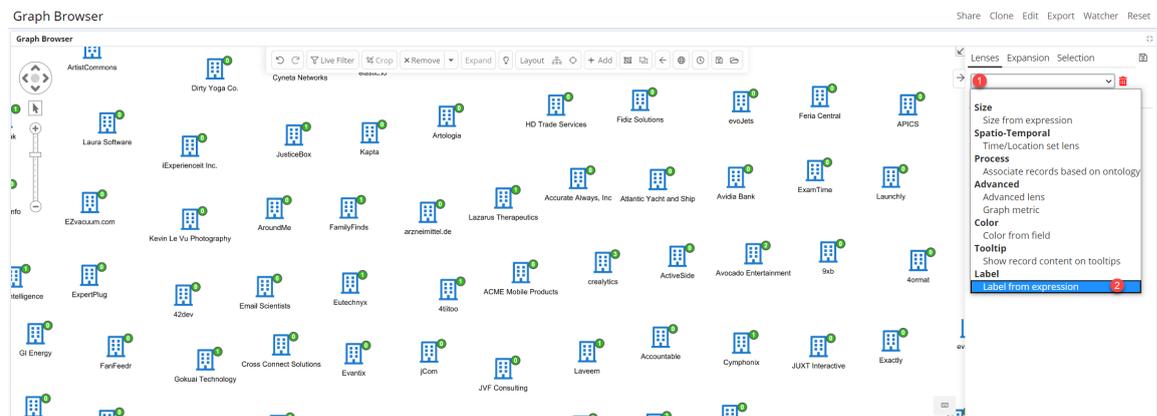
### Fixing labels with graph lenses

You can set graph node labels in the data model and then in the graph by using the lens feature.

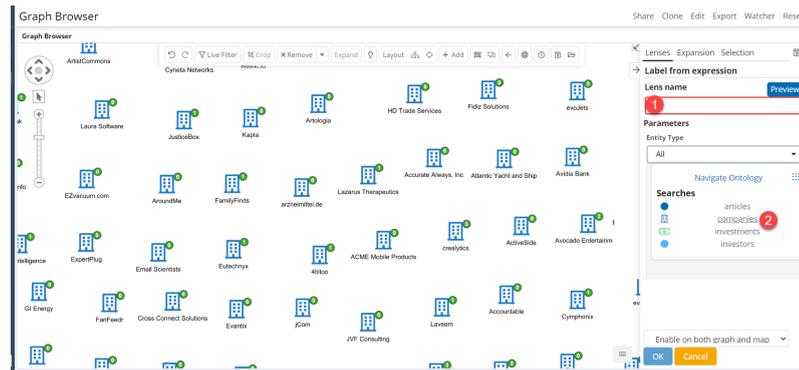
You should always set them in the data model first, but you can then override them in the graph with a lens that is useful for specific analysis. For example, you might first want to see the article's title, but later change your mind and want to see the article's author as a label instead.

To create and activate a lens, complete the following steps:

1. Click the **Toggle Sidebar** button ①.
2. In the **Lenses** tab, under **Create lens**, select **Label from expression** ②.



3. In the Lense name editor, enter “Company labels” ①.

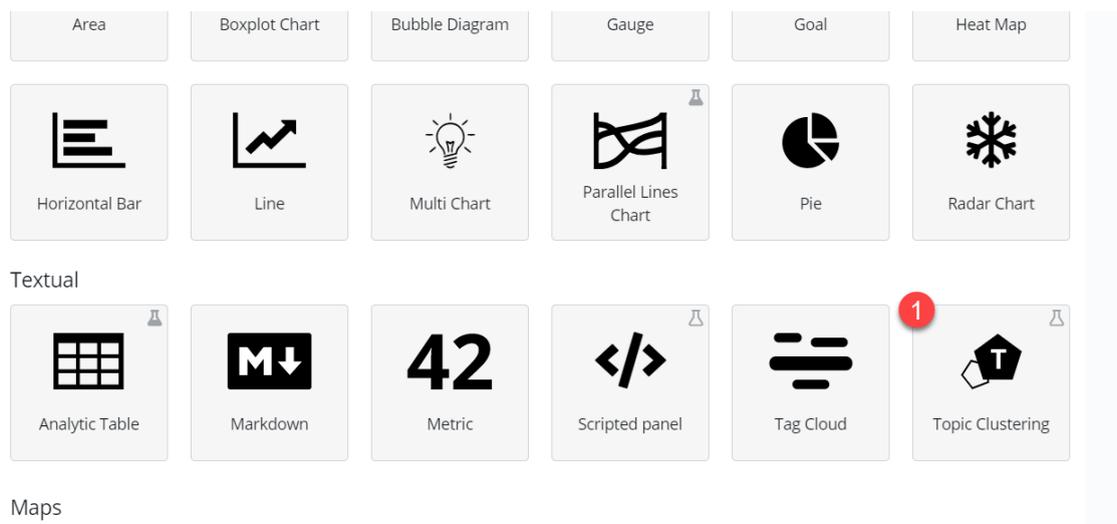


4. In the **Entity Type** drop-down menu, select **siren-import-home-companies** from the list of searches ②.
5. To fill the **Expression** field, click the + button ① and select **label** from the dropdown menu. The **Expression** field is populated automatically with `payload["label"]`.
6. Click **Save lens**.

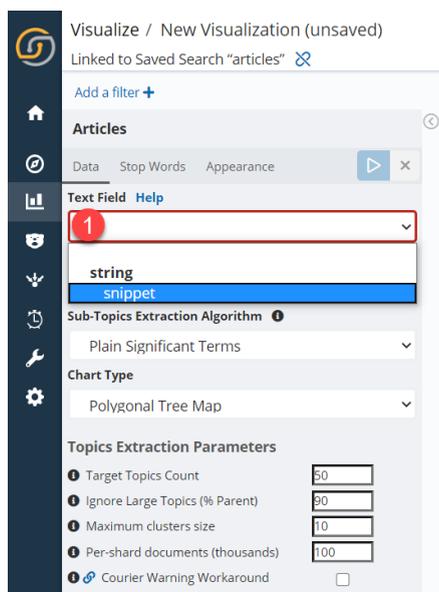
## Appendix 3: The Topic Explorer

To better explore the textual content of the articles, go back to the article generated dashboard, click on **Edit** and **Add a new visualization** (pick **articles** as the saved search to be used for this new visualization).

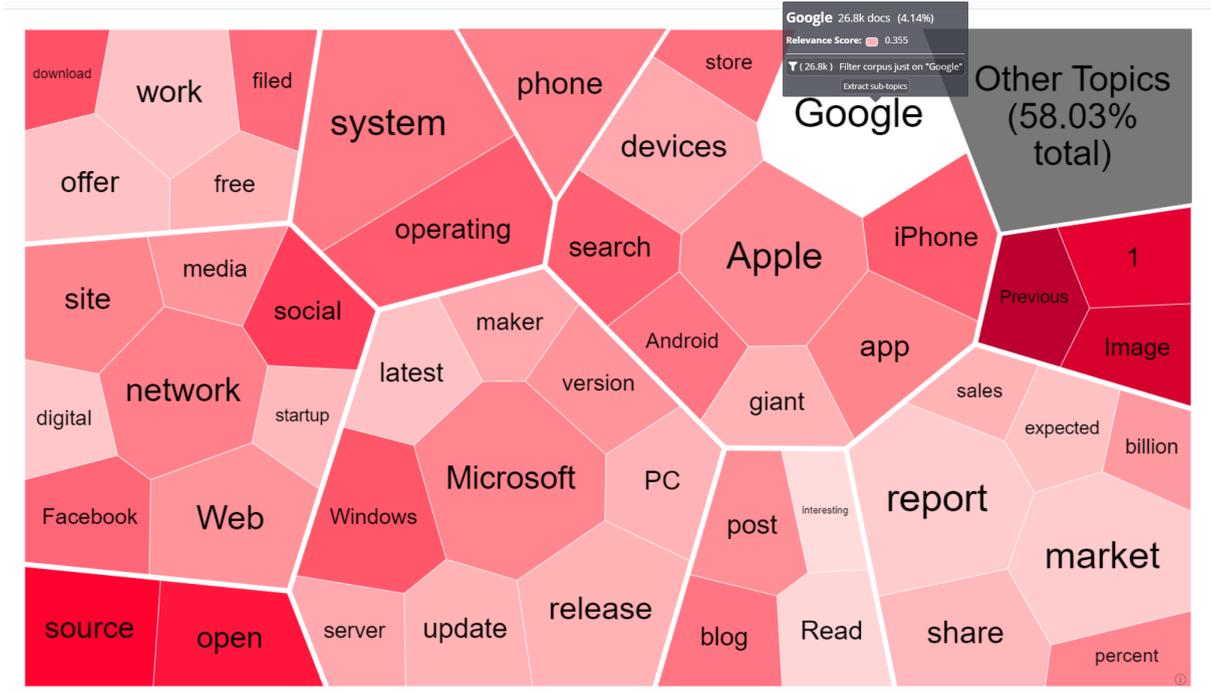
Select the **Topic Clustering** visualization ①.



Select as input the **snippet** field ①, which contains the actual text of the articles.



What is the most discussed topic in our data?



## Appendix 4: Creating Entity Identifiers

Siren expands its analysis capabilities by introducing Entity Identifiers(EID).

EIDs are used when entities exist (such as IP addresses, MD5s or things that are identified simply by strings, for example, City) that are present only as *values* in certain fields - in one or more tables - and do not have a “dedicated” table.

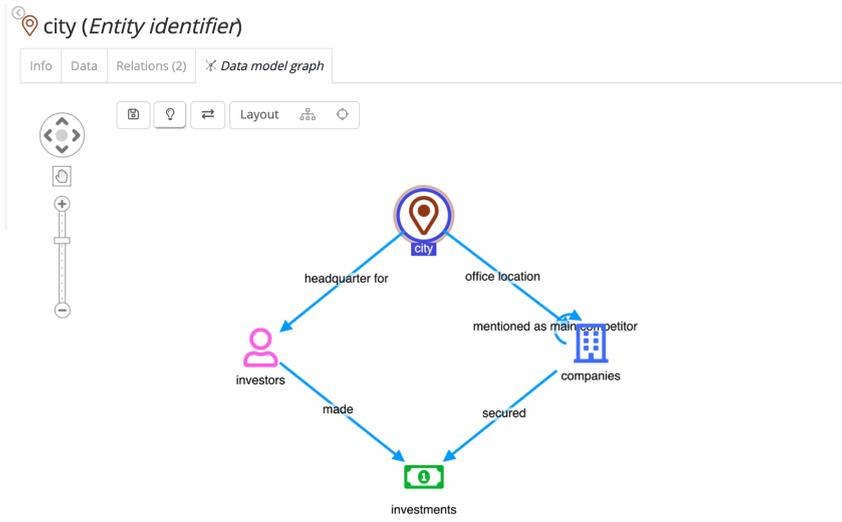
EIDs on the graph look like any other node, for example, an IP connecting two log records. They can also be used as ‘intermediate concepts’ when navigating dashboards.

To create an EID, complete the following steps:

1. Go to the **Data Model** app and click on **Create entity identifier**.
2. Enter the name as “city”, choose a suitable icon and color and click **Create**. You will see the entity identifier listed under the Searches on the left (1).
3. Click on the **Relations** tab (2) to create relations for the EID tab. We can create two two relations; one with companies and one with investors, since both of them contain the **city** field.
4. To manually create a relation, click the **Add relation** button, choose the **Target Entity** as **companies**, **Field** as **city** and enter the Labels as in the screenshot below.
5. Do the same for the **investors** entity with the **Field** as **city**.

Source Entity	Labels	Target Entity
city	office location has office	companies city
city	headquarter for headquartered in	investors city

4. Click **Save**.
5. Go to the **Data model graph** tab to see the EID added as a part of the data model.



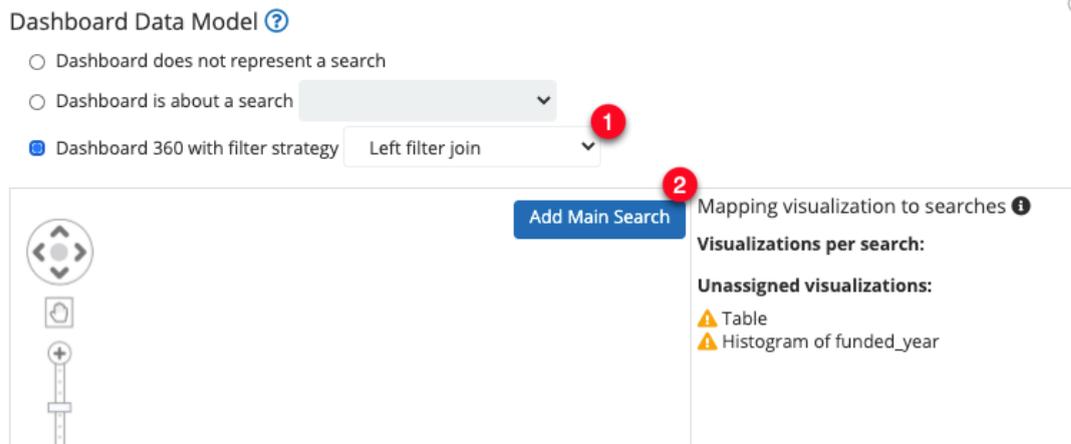
## Appendix 5: Dashboard 360

The Dashboard 360 feature allows a single dashboard to contain visualizations that are based on different searches, and allows you to perform coherent filtering across all of them. Dashboard 360 uses the relational data model to enhance data analysis capabilities in Siren Platform.

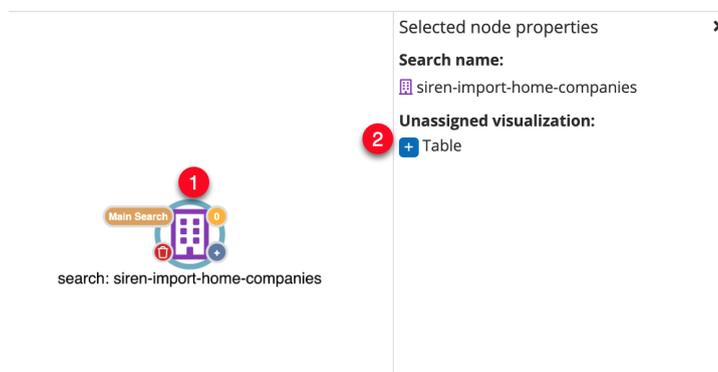
1. Click **Create new Dashboard** and name it “All Companies”.
2. Click **Add -> Add a Visualization**.
3. Select a **Record Table** visualization and associated it with the **siren-import-home-companies** search.
4. Click **Apply changes**, name the visualization “Table”, and click **Save**.
5. Save the dashboard.
6. Click **Add -> Add a Visualization**.
7. Select a **Vertical Bar Chart** and associate it with the **siren-import-home-investments** search.
8. Under the buckets table window, click **X-Axis**.
9. In the Aggregation field, select **Histogram** and in the **Field** field, select **funded\_year** with the **Interval** set to 1.
10. Click **Apply changes**, name the visualization “Histogram of funded\_year”, and click **Save**.
11. Save the dashboard.

As two searches are involved, we have to use a Dashboard 360. To configure Dashboard 360, complete the following steps:

1. Click **Edit** → **Data Model** and select **Dashboard 360 with filter strategy** (leave the filter strategy as **Left filter join**) ①.

2. Click **Add Main Search** ②.3. In the dialog box, select **siren-import-home-companies** as the main search and click **Done**.

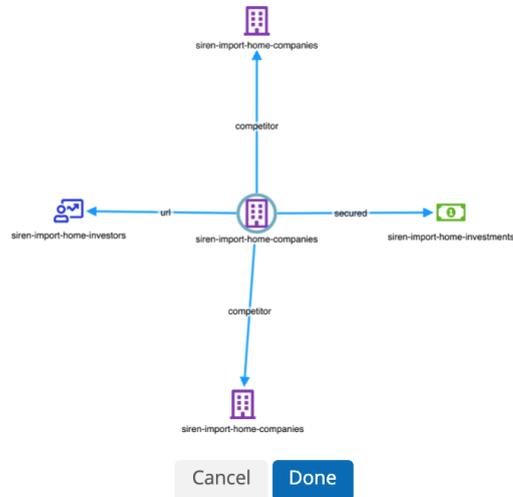
The Mapping visualization to searches panel shows the list of Dashboard 360-compatible visualizations and whether or not they are assigned to a search node. The warning symbol indicates that neither of the visualizations created earlier are assigned yet.

4. To assign a visualization, click the **companies** node to select it ①, then click the **+** button beside the required visualization (Table) ②.5. You can now build the dashboard data model by defining how this search is to be relationally connected with other searches (that is, which path in the relational data model to take). To do this, click the **+** button located on the lower-right of the Main Search node. This opens the following dialog which shows all the possible relations that come from the chosen search, as well as all possible searches that can be

reached through those relations.



Select the path to destination entity and press done.



6. Select the **siren-import-home-investments** node, and click **Done**.
7. You can now see how the dashboard data model looks at this point, with the companies node selected.

Main Search

search: companies

relation: secured

search: investments

Selected node properties

**Search name:**

companies

Use global time filter ⓘ

**Assigned visualizations:**

Record Table

Select the **investments** node and you will see that it has an unassigned visualization.

Selected node properties ✕

**Search name:**  
📄 siren-import-home-investments

**Use global time filter** ⓘ

**Unassigned visualization:**  
+ Histogram of funded\_year

8. Assign this visualization by clicking the ⊕ button, then save the dashboard.

Let's see what the dashboard looks like now that the visualizations have been assigned. Note the icons on the top-left of each visualization's container, indicating they have been assigned to the appropriate searches.

All Companies (159,800) Share Clone Edit Export Watcher Reset Last 100y

Filters Search... (e.g. status:200 AND extension:PHP) Help 🔍

📄 Table

Time	Geopoint	category_code	city	countrycode	de:
August 1st 2014, 01:00:00.000	-	software	-	-	coll n, f ma nt
May 31st 2014, 01:00:00.000	28.61393 91,77,209 0212	ecommerce	Del hi	IND	Coi for

📄 Histogram of funded\_year

## Filtering with Dashboard 360

The really significant achievement here with Dashboard 360 is that you can perform coherent filtering across the dashboard's visualizations.

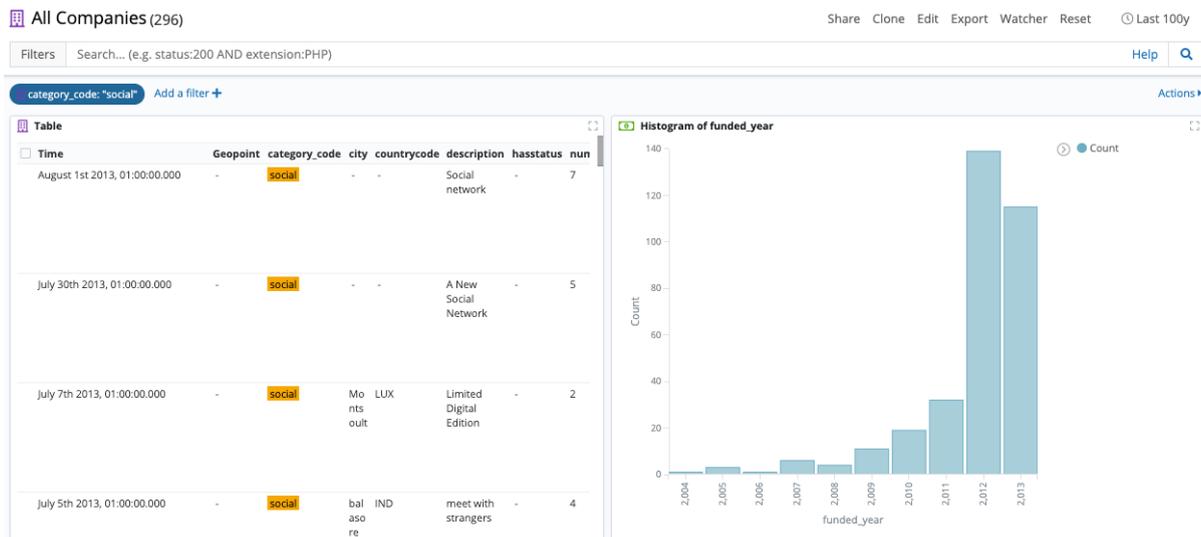


Let's say that you want to see the pattern of yearly funding across a particular industry segment. For example, when did funding for social media really take off?

To do this, filter the Record Table visualization on the appropriate category code, **social**.

▶ August 1st 2013, 01:00:00.000	Raploja	ecommerce	
▶ August 1st 2013, 01:00:00.000	Rootty	social	🔍
▶ August 1st 2013, 01:00:00.000	Epic Business	const	Filter for value

The record table now lists all of the companies in the social category. But, more importantly, the histogram visualization displays the yearly funding trend; this showed a huge increase in the number of investments between 2012 and 2013.

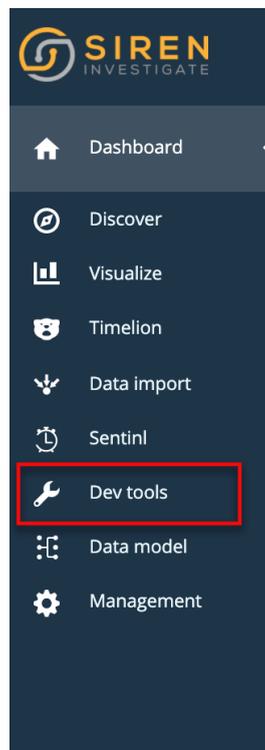


You can also filter by clicking on the segments in the **Funded Year** histogram, this will coherently filter the results in the Record Table.

## Appendix 6: Deleting data in Siren and changing a table schema

Deleting data in Siren Platform is intentionally difficult, due to the damage that it could cause in environments where the data is in very large streams, such as in large Elasticsearch installations.

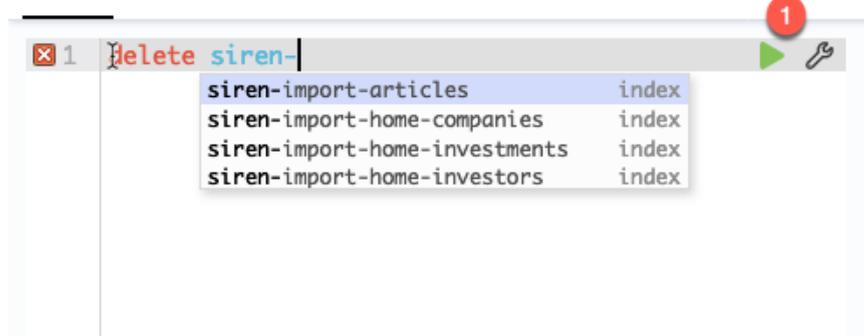
However, you can delete an index by using the **Dev Tools console**.



1. Go to the **Dev tools** app.
2. In the **Console**, type “delete” and the name of the index, for example “siren-” in the editor pane.
3. Click the play icon ① and you see the acknowledgment in the response pane.

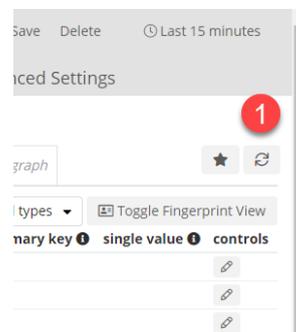
## Dev Tools

Console Translate Gremlin Console



**IMPORTANT:** If you delete the index and the **companies** dashboard still exists, the dashboard displays errors. You can either reinstate the companies index by uploading the data again, or you can delete the dashboard.

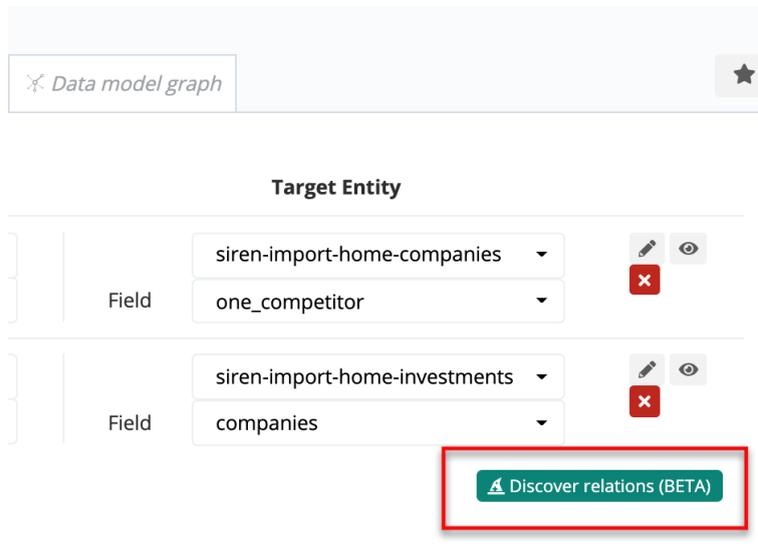
If the data structure has changed, for example, if there are more fields or fields of a different type, then you must refresh the 'Index Pattern Search' field list by clicking the **Refresh** button ①.



## Appendix 7: Auto-discovery of relations

This feature automatically detects and creates relations between index pattern searches. To use this option, complete the following steps:

1. Go to the **Data Model** app and in the siren-import-home-companies index pattern search, click on the **Relations** tab.
2. Click **Discover relations**.



3. The wizard automatically selects all of the available Index Patterns (saved searches) and also adds **EID Patterns**.

**NOTE:** EIDs are created to link certain shared identifiers that belong to certain entities; for example, a URL, an email address, or a city. However, there is no corresponding index in which these entities are the primary key. They are simply used as *values* in fields across different indices and hence cannot be directly used to create a relation. EID will be discussed in more detail later in this tutorial.

4. Click on the **EID Patterns** tab ① and deselect all of the EID name fields ②, since we're only interested in creating relations between index patterns for now.

- Click **Discover relation**.

[BETA] Relations Auto-Discovery Wizard – Select Input Searches ×

Index Patterns | EID Patterns <sup>1</sup> | Settings

Test String

<input checked="" type="checkbox"/>	Entity ID Name	Applies To	Regular Expression	Test Match
<input checked="" type="checkbox"/>	URI	Terms	<code>^[a-zA-Z0-9+.-]+://</code>	No Match <span>🗑️</span>
<input checked="" type="checkbox"/>	IP_STRING	Terms	<code>(?:^(?:25[0-5] 2[0-4]\d 1\</code>	No Match <span>🗑️</span>
<input checked="" type="checkbox"/>	EMAIL	Terms	<code>^\w._%+~+@[\w.-]+\.[a-z</code>	No Match <span>🗑️</span>

- On the **Suggested Relations** tab, deselect all of the source fields except the **(siren-import-home-companies) url** field <sup>1</sup>.
- Click **Set relations** <sup>2</sup>.

[BETA] Relations Auto-Discovery Wizard – Results ×

Suggested Relations (1/10) | Existing Relations | Per-Field Notes | Log

Type	Source	Target
<input checked="" type="checkbox"/>	▼(siren-import-home-companies) url	<input checked="" type="checkbox"/> (siren-import-home-investors) url
<input type="checkbox"/>	t ▼[EID] category_code	
<input type="checkbox"/>	t ▼[EID] city	
<input type="checkbox"/>	t ▼[EID] countrycode	
<input type="checkbox"/>	t ▼[EID] statecode	

2

- The relations are automatically generated in a draft state (highlighted in blue) on the **Relations** tab. You can review the relations that are created for all of the indexes and

change the **Labels** if you want to.

Data model Save Delete Last 100

Model contains unsaved automatically generated relations of which 1 in this entity - they appear with a different color in the *Relations* tab. Save All Remove All

+ Create index pattern search **siren-import-home-companies (Index pattern search)**

+ Create entity identifier

**Searches**

- siren-import-home-comp...  
siren-import-companies
- siren-import-home-inves...  
siren-import-home-investments
- siren-import-home-inves...  
siren-import-home-investors

Info	Fields (26)	Data (159800)	Relations (3)	Scripted fields (0)	Options	Revisions	Data model graph												
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Source Entity</th> <th>Labels</th> <th>Target Entity</th> </tr> </thead> <tbody> <tr style="background-color: #e0f0ff;"> <td>           siren-import-home-companies            Field url         </td> <td>           → url            ← url         </td> <td>           siren-import-home-investors            Field url         </td> </tr> <tr> <td>           siren-import-home-companies            Field id         </td> <td>           → competitor            ← competitor         </td> <td>           siren-import-home-companies            Field one_competitor         </td> </tr> <tr> <td>           siren-import-home-companies            Field id         </td> <td>           → secured            ← secured by         </td> <td>           siren-import-home-investments            Field companies         </td> </tr> </tbody> </table> <p style="text-align: center; font-size: small;">⚠ This is an unsaved suggested relation – use the Save buttons here or at the top of the page to make it permanent.</p>								Source Entity	Labels	Target Entity	siren-import-home-companies Field url	→ url ← url	siren-import-home-investors Field url	siren-import-home-companies Field id	→ competitor ← competitor	siren-import-home-companies Field one_competitor	siren-import-home-companies Field id	→ secured ← secured by	siren-import-home-investments Field companies
Source Entity	Labels	Target Entity																	
siren-import-home-companies Field url	→ url ← url	siren-import-home-investors Field url																	
siren-import-home-companies Field id	→ competitor ← competitor	siren-import-home-companies Field one_competitor																	
siren-import-home-companies Field id	→ secured ← secured by	siren-import-home-investments Field companies																	
<a href="#">Add relation</a>				<a href="#">Discover relations (BETA)</a>															

9. Click the **Data model graph** tab to see the data model in a visual representation.  
The unsaved relation is represented by a dotted line.
10. Click **Save All -> Save All Automatic Relations**.
11. Save the data model.

## Legal notices

Siren Platform™ is a trademark of Sindice Ltd. trading as Siren, with offices in GTC, Mervue Business Park, Galway, Republic of Ireland.

Elasticsearch™ is a trademark of Elasticsearch B.V., registered in the U.S.A. and in other countries.