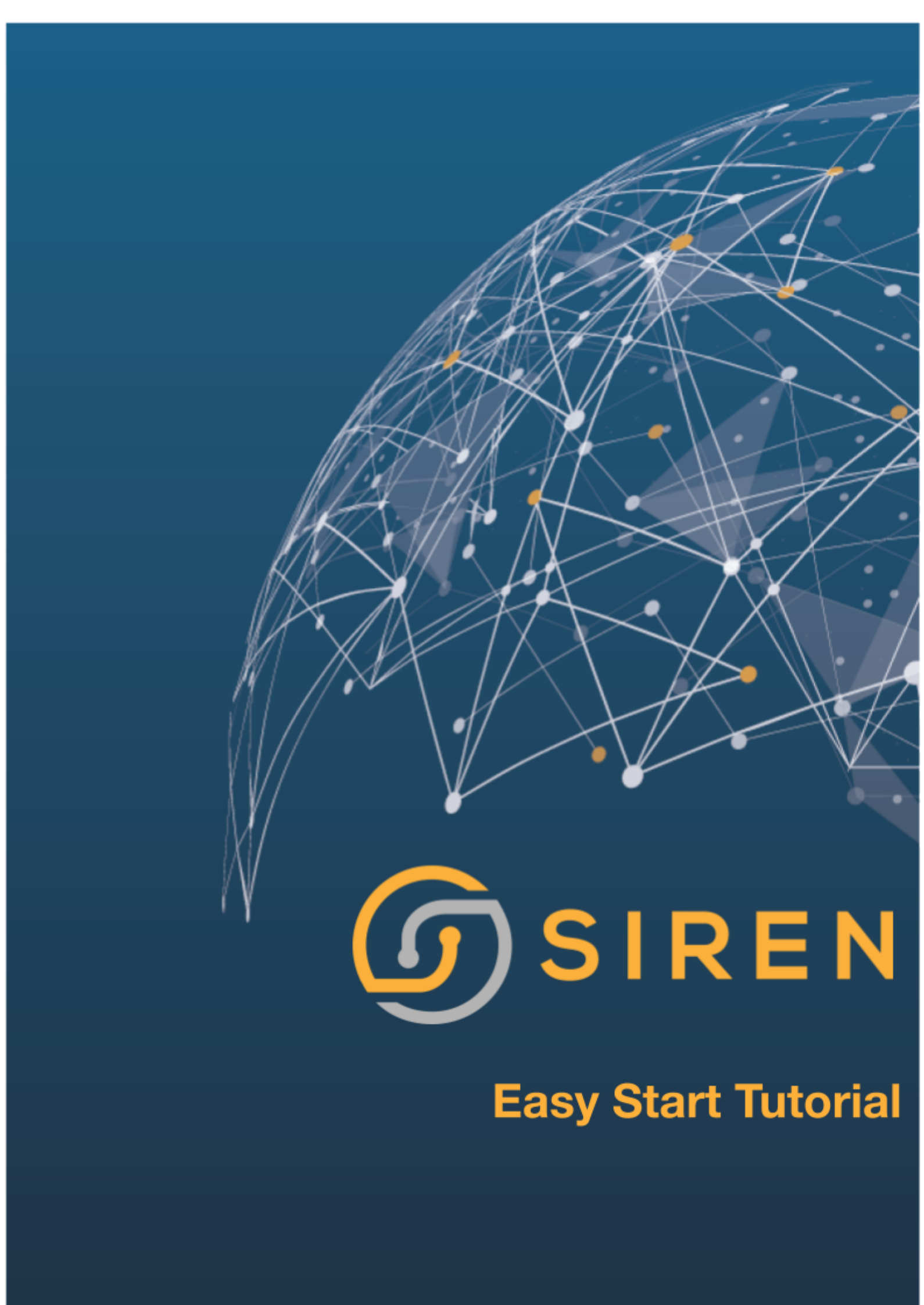


Easy Start Tutorial

Table of contents

About this tutorial

This tutorial will help you to get started with the **Siren Platform - Easy start**, version 12.1 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 datasources.

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 entity table (optional).
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 the data.
10. Introduction to the Graph Browser and Link Analysis, including exercises for you to complete.

The tutorial also includes an Advanced features section that contains additional important topics.

What will you get from this tutorial?

By the end of the first part of this tutorial, you will be able to use Siren Platform to search and investigate data about companies, investments, and investors, and you will learn how to apply basic search and analysis functions to the data.

If you go on to complete the exercises in the Advanced features section, you will discover how to import and investigate more data, using some of our advanced features and configurations. In essence, you will *level up* your skill set as an investigator.

We recommend you check out our 'What is Investigative Intelligence' video for a taste of what's possible.

Watch the What is Investigative Intelligence video.

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 Web browsers:

- Google Chrome
- Mozilla Firefox

Before you begin

On Windows, you must have an **improved ZIP manager** such as 7-zip or WinRAR. This is because the standard Windows ZIP support cannot 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.

<table>
<tr><td>Windows</td><td>Double-click on the start.bat file and wait for Elasticsearch and Siren Investigate to load.

NOTE: 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.

IMPORTANT: 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.</td></tr>
<tr><td>Linux or macOS</td><td>1. Open a Terminal window and change directory to the extracted folder:
<code>cd {extracted folder}/elasticsearch</code>
Run the following command:
<code>./bin/elasticsearch</code>
The Terminal window shows Elasticsearch messages flowing. Do not close the command window.

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

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

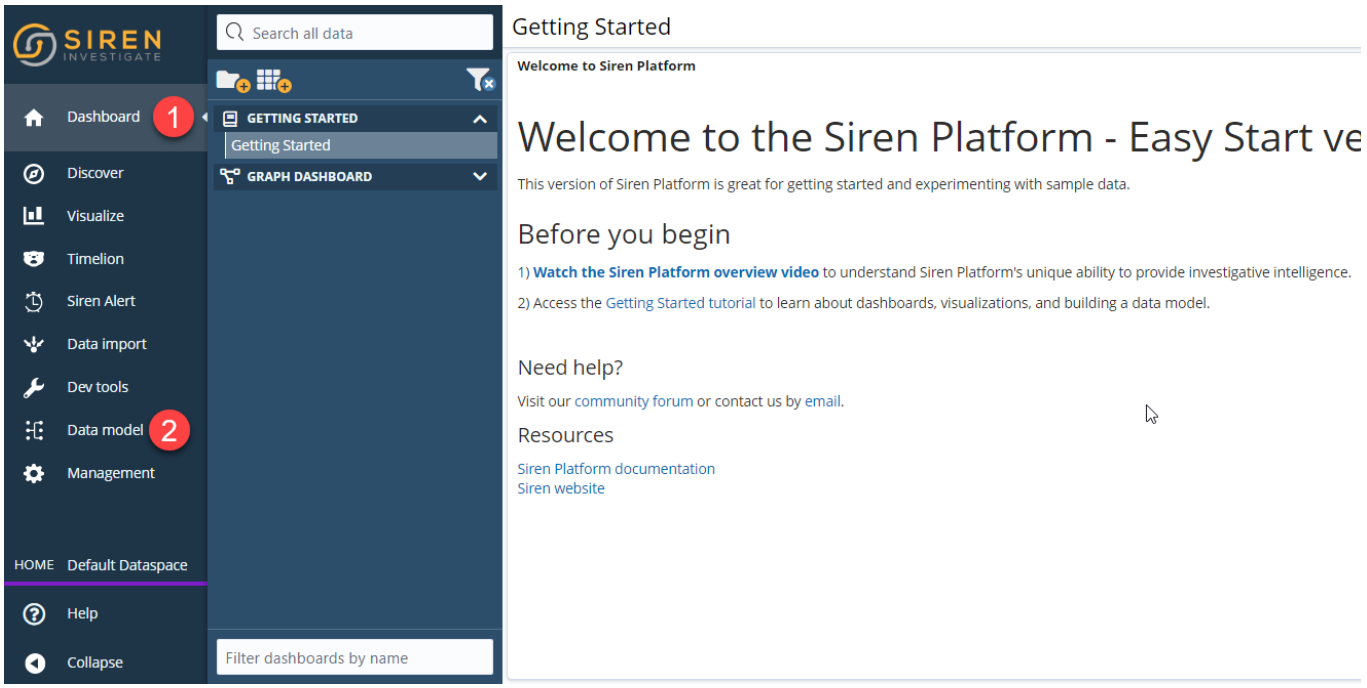
Relaunching Siren Investigate

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 model ②



Importing data into Siren Platform

(Average completion time: 10 mins)

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

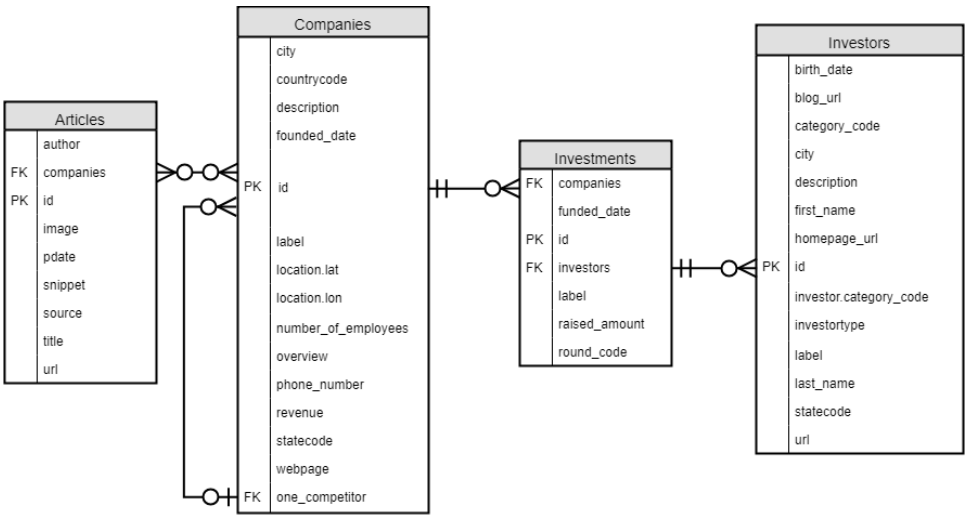
About the demo data

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_nlp.csv**: A collection of technical articles. Most of the articles mention one or more companies. To extract (or annotate) these mentions, follow the advanced NLP section of the tutorial.
    - **additional_companies.csv**: A list of companies to be used with the advanced Importing and transforming data section of the tutorial.

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



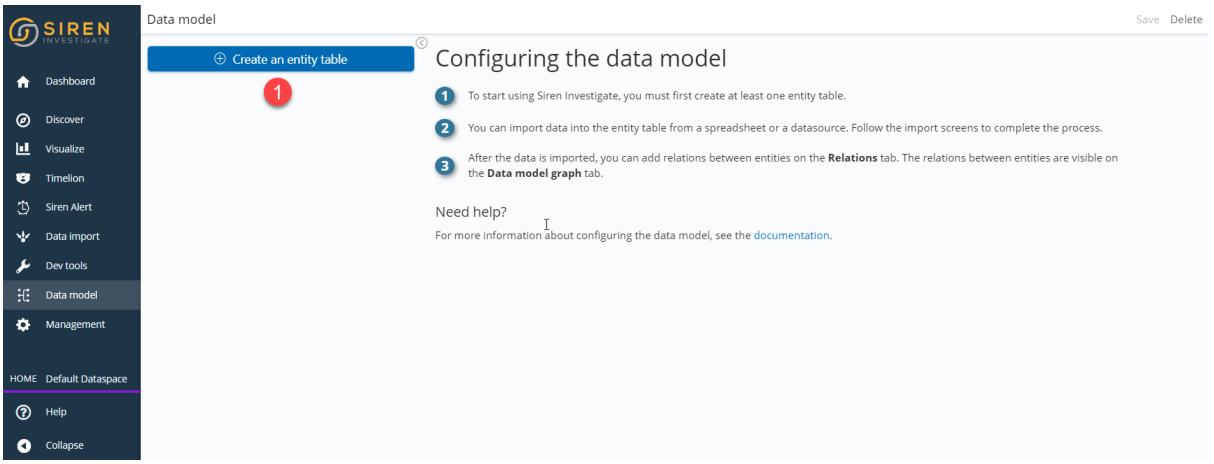
There are also other relationships in the data, for example, city and state names, URLs, and email addresses.

Importing your first table of demo data

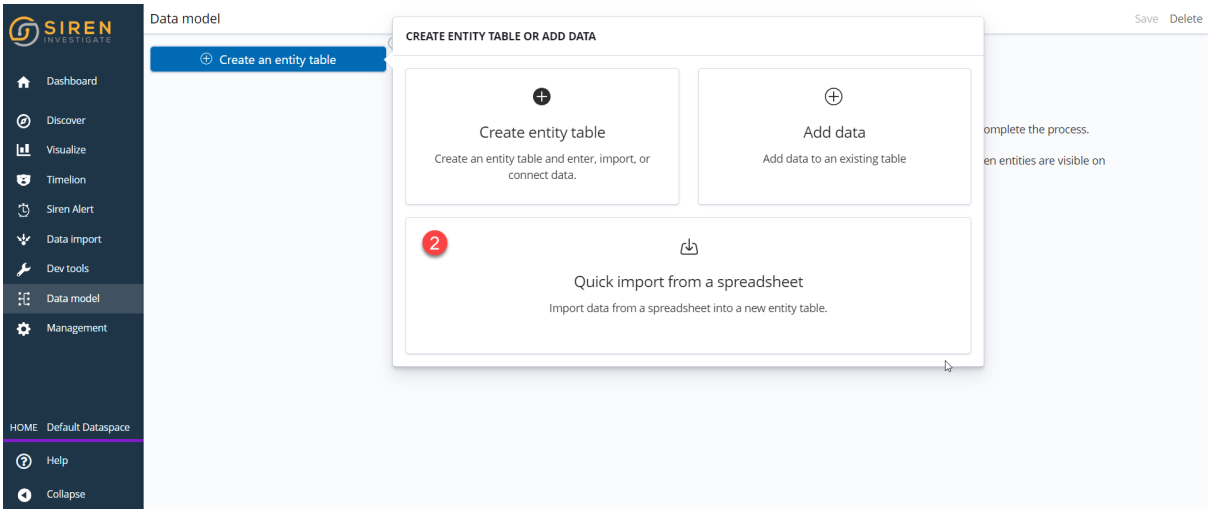
(Average completion time: 5 mins)

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.

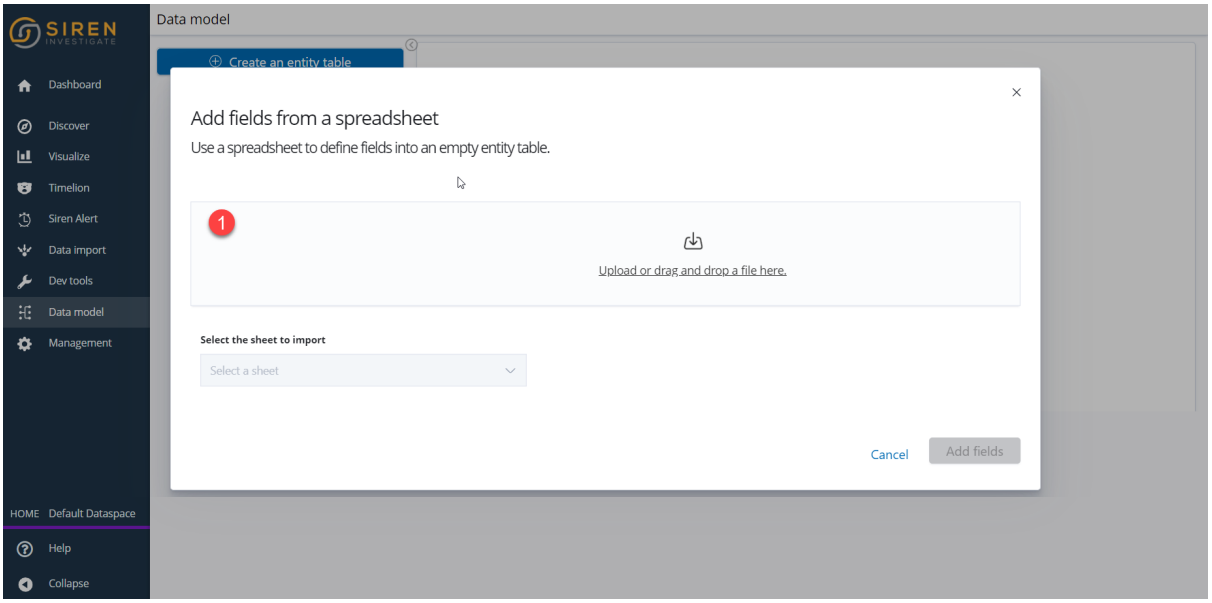
1. From the left-hand navigation menu, click **Data model**.
2. Click **Create an entity table** ①.



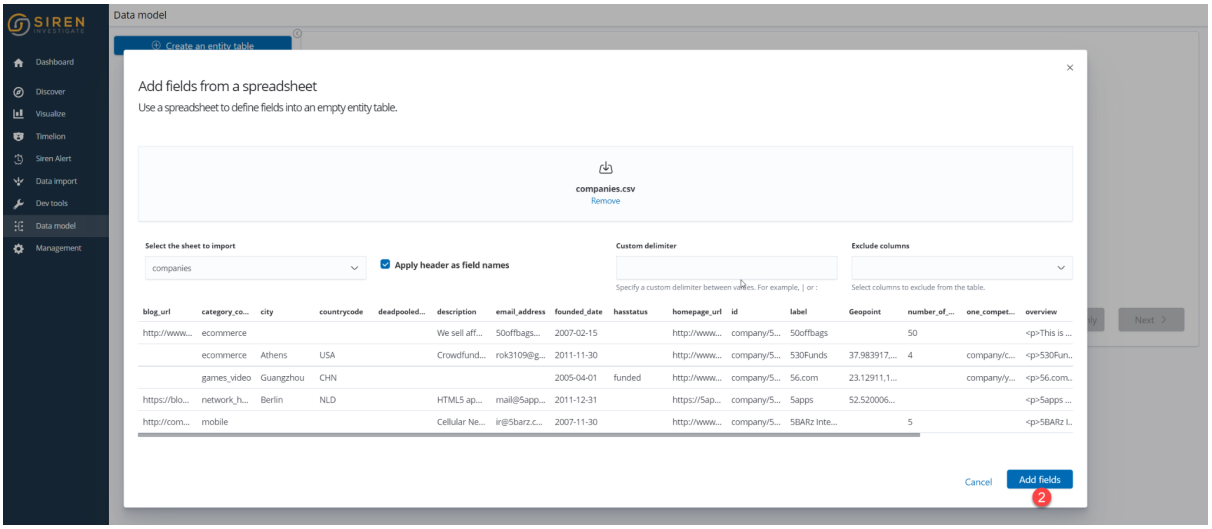
3. Click **Quick import from a spreadsheet** ②.



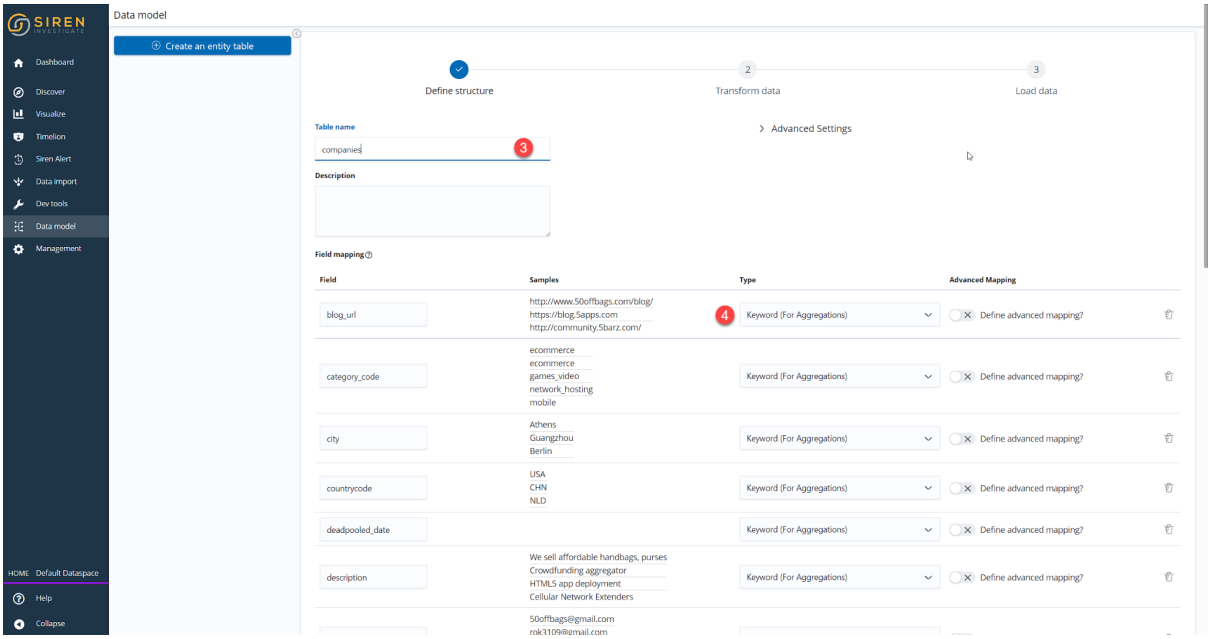
4. Drag and drop the **companies.csv** file in the upload area ①.



5. Click **Add fields** ②.



6. In the **Table name** field ③**,** type *companies*.



Before importing data, we need to define a data type for each field. Most of the data types that are listed in the **Type** drop-down menus ④ 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

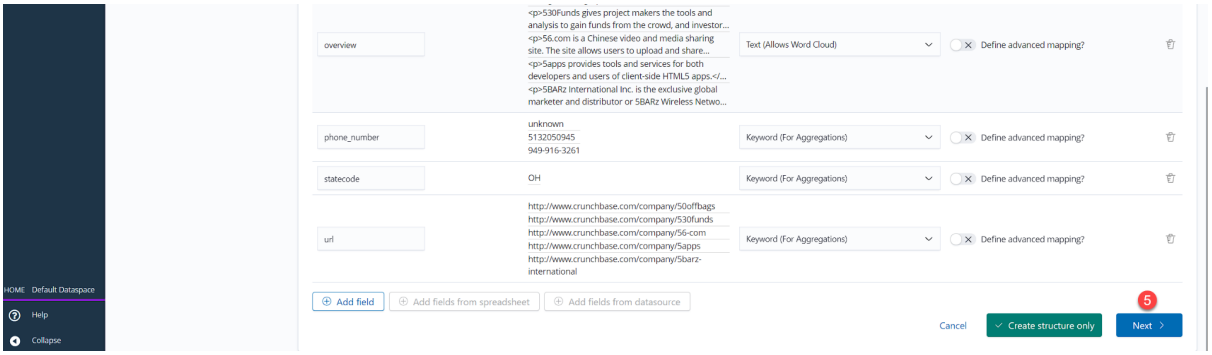
Use this option for strings that should be considered *unbreakable*. For example, your dataset might include a field called City that contains strings such as ‘New York’ and ‘Los Angeles’. If you mark this field as **Keyword**, then the term will be searched for in its entirety, rather than broken into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

Advanced Mapping

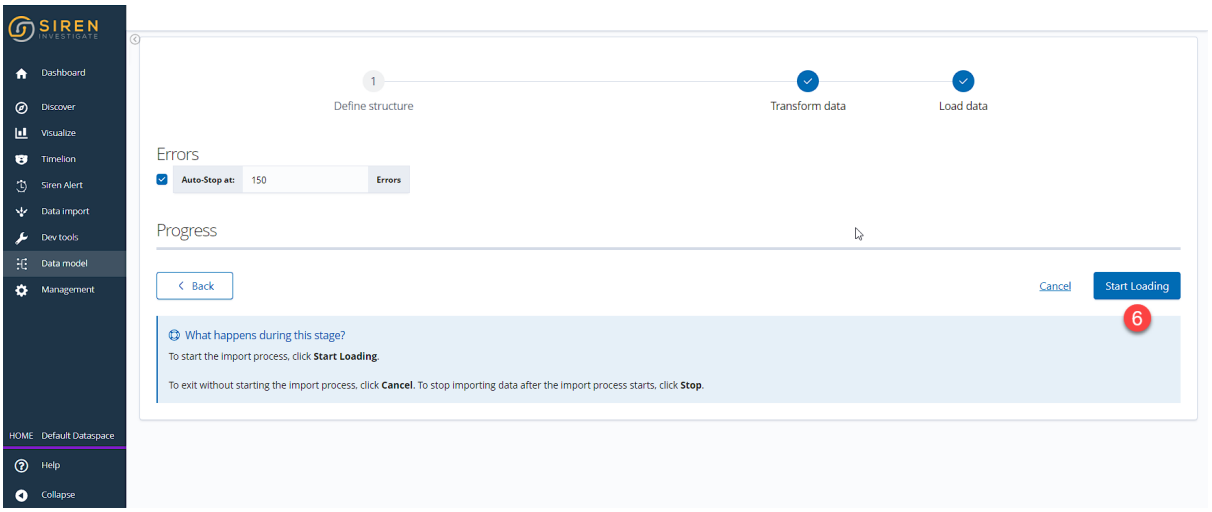
The **Advanced Mapping** switch gives advanced users the option to add a JSON object to define the mapping in Elasticsearch format. For example, selecting **Date** and switching the **Advanced Mapping** switch to 'on' would allow you to define a specific format for the date field, such as {"format":"yyyy-MM-dd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

For now, ensure that the following selections are correct:

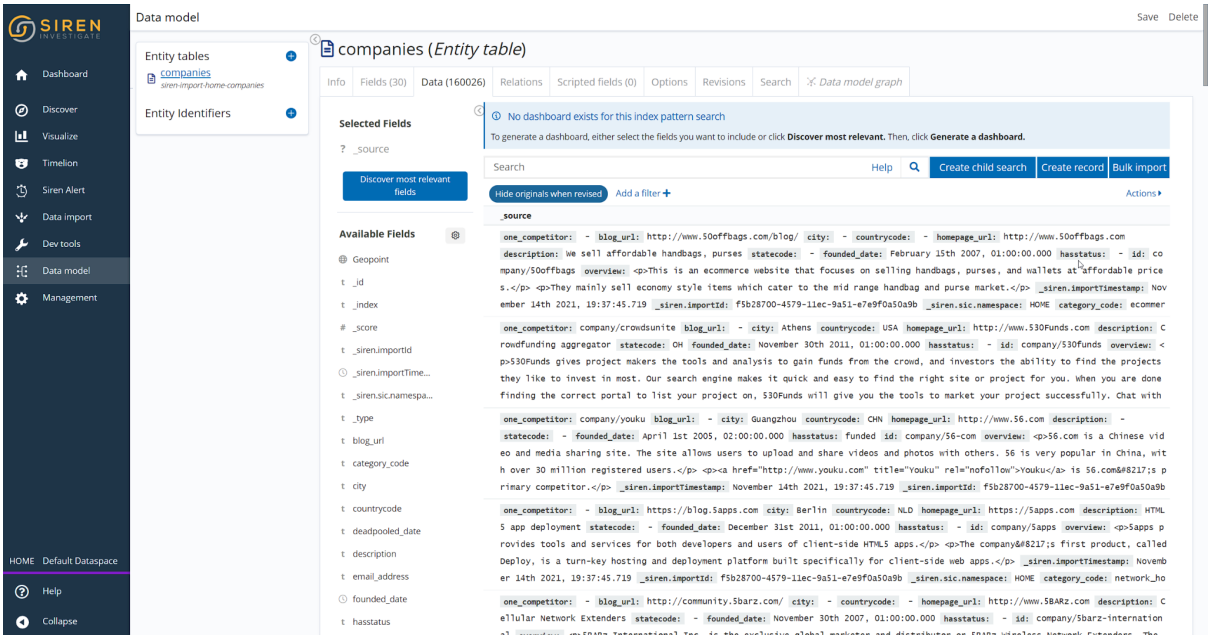
1. Scroll to the **description** field and set the type to **Text (allow word cloud)**.
2. For the **founded_date** field, set the type to **Date**.
3. For the **Geopoint** field, set the type to **Geo Point**.
4. For the **overview** field, set the type to **Text (allow word cloud)**.
3. Click on **Next** ⑤.



4. On the **Transform data** screen, click **Next**, since we don't need to do any transformation right now. We will explore this in more detail in the [Importing and transforming data] section of the tutorial.
5. Click **Start loading** ⑥ to start importing data.



After the import is complete (this should take a few seconds), the details of the new entity table will be displayed.



TIP: If you have data in Elasticsearch already, you can make it visible in Siren by creating an entity table from an existing index pattern. For more information, see the documentation.

Setting the properties of the entity table

1. In the **Data model** app, click the **Info** tab.
2. In the **Icon** field, select an icon to represent companies, such as a building. (Hint: **fal fa-building** is a good one.)
3. In the **Color** field, choose the color you want for the icon.
4. In the **Default label** field, select **Document Field**, and - from the dropdown menu - select **label**.
5. Select **Enable the Time filter based on the following field**.
6. In the dropdown menu, select **founded_date**.
7. Click **Save**.

Entity tables

Before we proceed any further, let's present some fundamental concepts and terminology that will help you to understand the platform.

- **A record**: In this case, we have created one record for each row in the CSV file. Records contain *fields*, which contain *values.*
- **An entity table**: A collection of records that contain information about the same type of entity, such as the “companies” table that we have just created. Technically speaking, an entity table refers to one or more underlying *indices.* The term “index” refers to the terminology used typically by Siren’s back-end system - Elasticsearch. An entity table is therefore one index at a minimum, but it is possible to define an entity table composed of more than one index (in this case, it is called “index pattern”). For example, entering “logs_*” as the index pattern will create an entity table that displays records coming from all the Elasticsearch indices that have a name starting with **logs_** (for example, **logs_2021_10_1**, **logs_2021_10_2**, and so on).

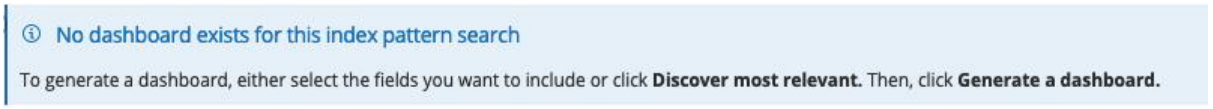
  Entity tables are the core ‘searchable objects’ in Siren Platform. This means that they can be searched, drilled down, and become the basis of analytical visualizations.
- **Searches**: Typically, searches are a subset of an entity table. For example, if you wanted to sort your list of companies by those based in France, you can create a search for French companies, which would then appear as a subset of the ‘companies’ entity table.

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 into 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 an Index Pattern Search. Let's see together how we can automatically create a dashboard with Siren Investigate.



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.

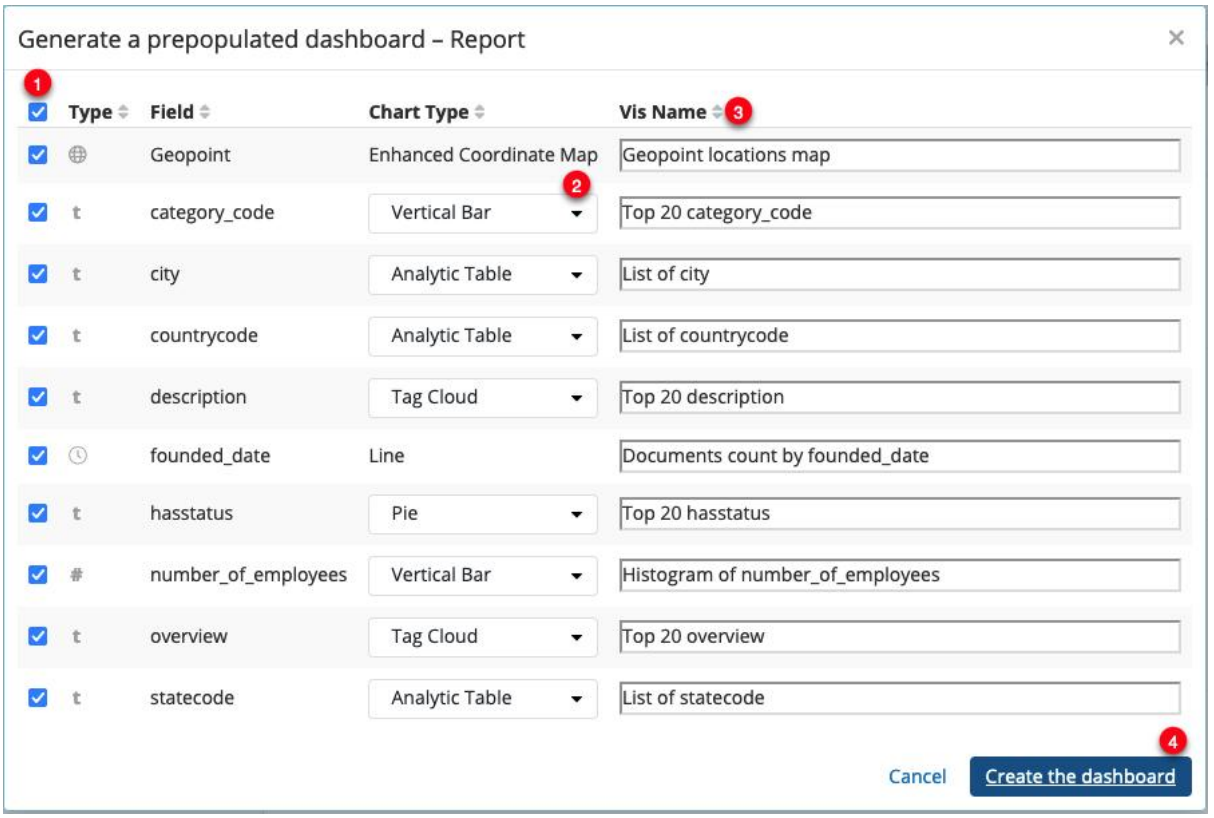
The feature to auto-generate a dashboard is available both on:

- 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** tab of the **Data model** app, you can view the records in the entity table and - to the left of the screen - you can see a list of **Available Fields**.
2. Click **Discover most relevant fields**.
3. The system identifies the fields that are most relevant for generating the dashboard. Accept the default fields and click **Select**.
5. Now, the selected fields are listed on the left. Click **Generate a dashboard**.

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



The report screen contains the following elements:

① - Checkboxes that allow you to individually select which of the auto-generated visualizations to use 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).

8. Click **Create the dashboard** ④.

The new **Companies** dashboard is now generated with different types of visualizations including an Enhanced Coordinate Map, Vertical bar charts, a Pie chart, a Line Chart, Analytic Tables, Word Clouds, and a Record Table. For more information about visualizations, see the Siren Platform documentation.

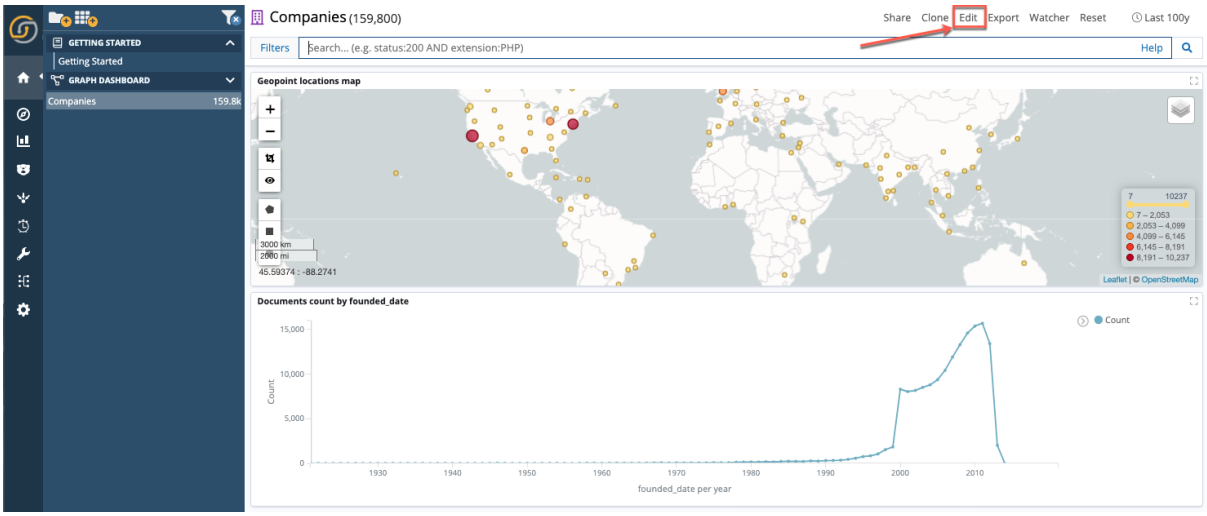
TIP: If you do not specify a new dashboard name, it defaults to the name of the index pattern search. You can edit the dashboard name at any time by right-clicking the dashboard in the dashboard menu and clicking **Rename**.

Editing dashboards

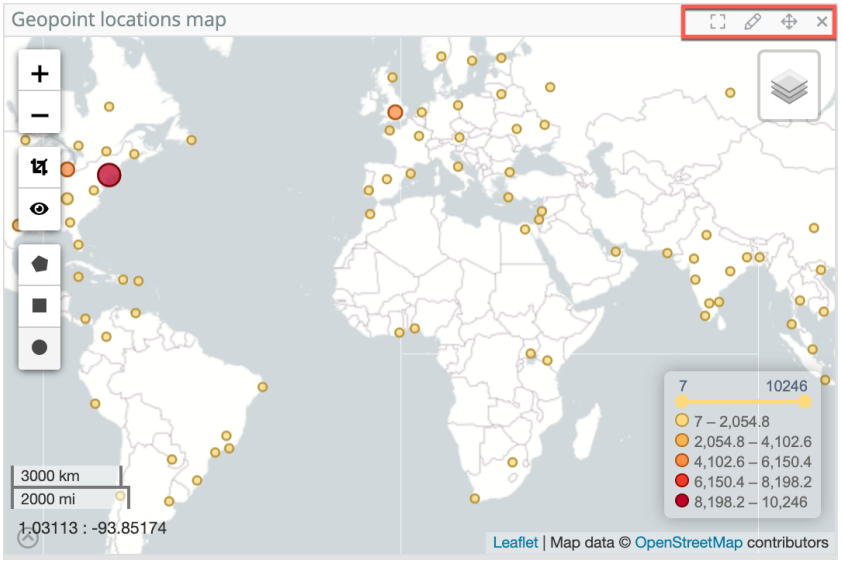
(Average completion time: 5 mins)

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

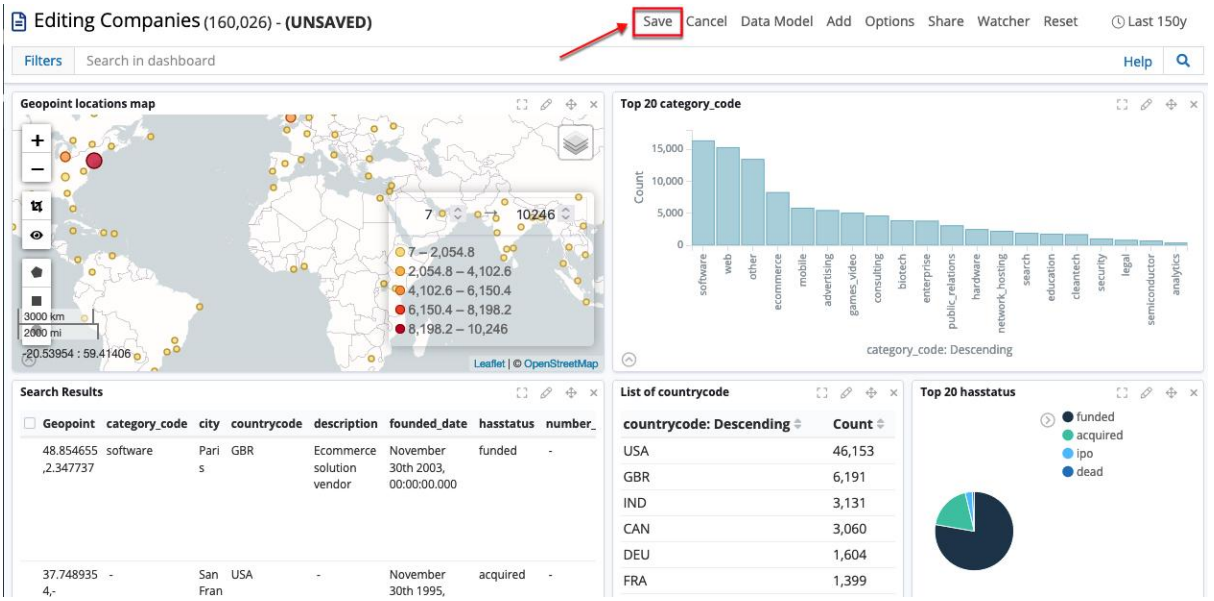
1. At the top-right of the screen, click **Edit**.



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:



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

Using the dashboard

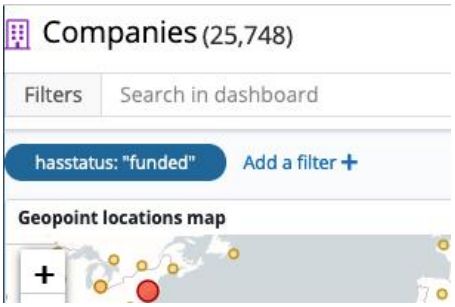
Now, let's explore this dashboard by interacting with the visualizations to take a closer look at important data.

Creating filters

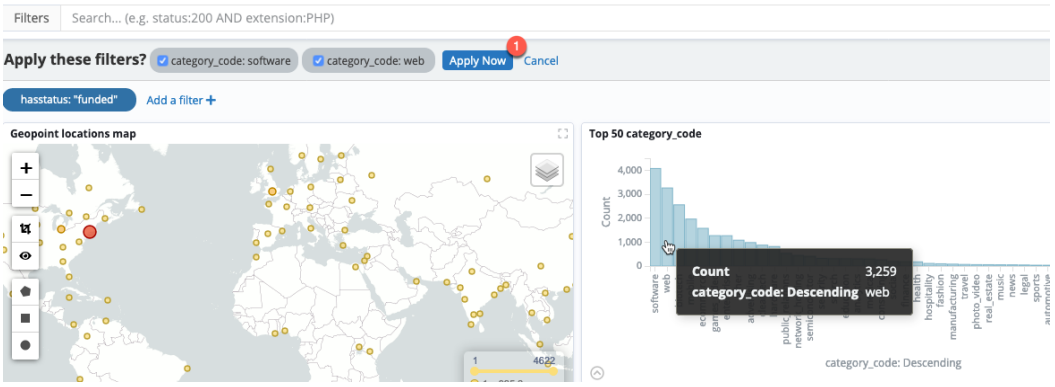
(Average completion time: 5 mins)

You can create filters for dashboards either by using the **Add a filter** button or by selecting categories in the visualizations.

1. Click a segment of the **Top 20 hasstatus** pie chart to filter the dashboard on a category, such as the largest portion - the **funded** status. A filter appears at the top of the dashboard and the number of records in the dashboard is reduced.



2. On histogram visualizations, you can create multiple types of filters. On the **Top 20 category_code** bar chart, hold the **Ctrl** or **Cmd** key and click on the bars corresponding to the **software** and **web** category codes. After making the selections, click on **Apply Now** ① to apply a filter with an OR condition.

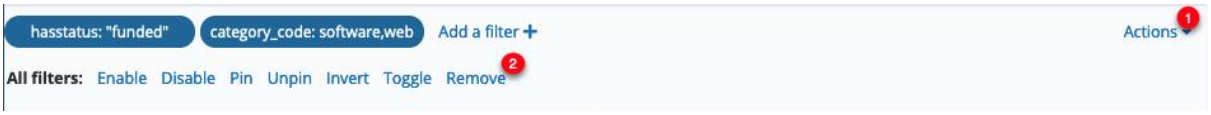


Textual Search

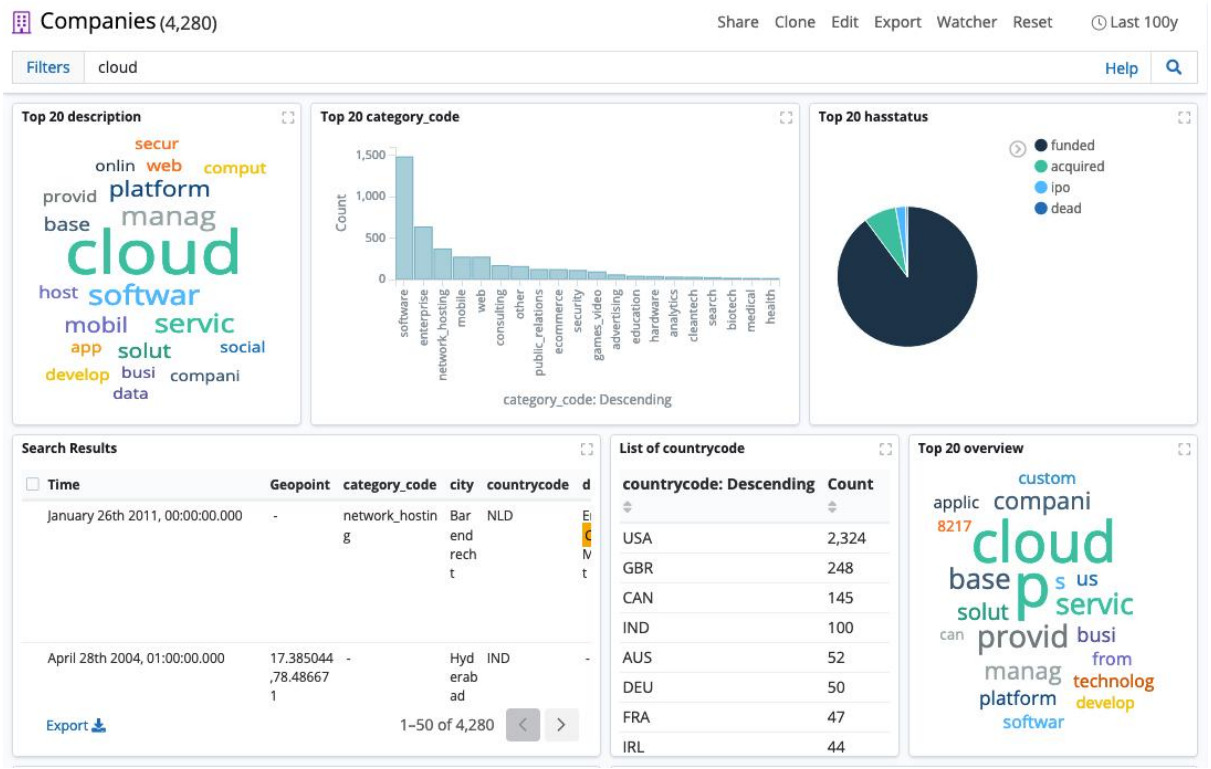
On the dashboards, you can perform a free text search by using the **Search** bar.

TIP: You can rearrange the visualizations to bring the text-specific ones to the top.

1. Clear the current filters in the dashboard, by clicking **Actions** ① and **Remove** ②.



2. In the **Search** bar, enter the word *cloud* and hit the **Enter** key. The dashboard updates all of the visualizations to reflect the search results.



3. Next, let's perform an advanced search by using the Lucene query syntax.
    a. 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.
    b. 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.
    c. 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 **OR**

commerce~ to get the results of all mentions of ‘ecommerce’, ‘commercial’, and so on.

4. To reset the Search bar to the previous entry - cloud - click **Reset** in the Options menu on the top-right of the screen.
5. To reset the dashboard (and all dashboards at once) to its previous saved state, click the **Reset all filters** button in the Dashboard menu. 

Creating dashboards manually

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

Creating the investments entity table

(Average completion time: 5 mins)

Return to the **Data model** app to create a new entity table as you did in the first import.
Follow the same steps making the following settings:

1. Import the **investments.csv** file and name the entity table *investments*.
2. In the **Field mapping** section, scroll to the **funded_date** field and select **Date** as the type.
3. For the **raised_amount** field, select **Long**. (This allows an integer over 32 bits: Some investments are big!)
4. Click **Next**.
5. On the **Transform data** screen, click **Next**.
6. Click **Start loading**.
7. Click the **Info** tab.
8. Select an icon, such as one to represent currency. (We like **fal fa-money-bill-alt**.)
9. Select a color for the icon.
10. In the **Default label** field, select **Document Field** and, in the second dropdown menu that appears, select **raised_amount**.
11. Select **Enable the Time filter based on the following field** and select **funded_date** from the dropdown menu.
12. Click **Save** in the top-right corner of the screen.

Creating the dashboard

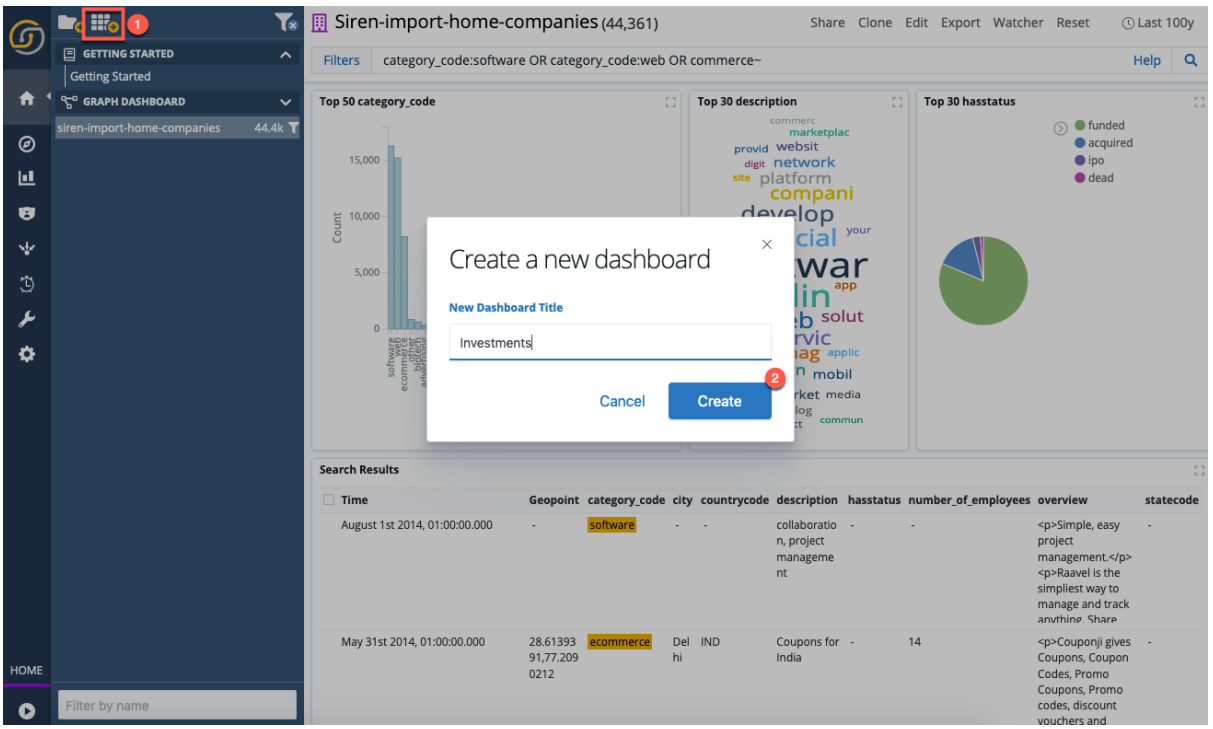
(Average completion time: 7 mins)

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.
3. Name the dashboard *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 an entity table.

Creating visualizations

(Average completion time: 15 mins)

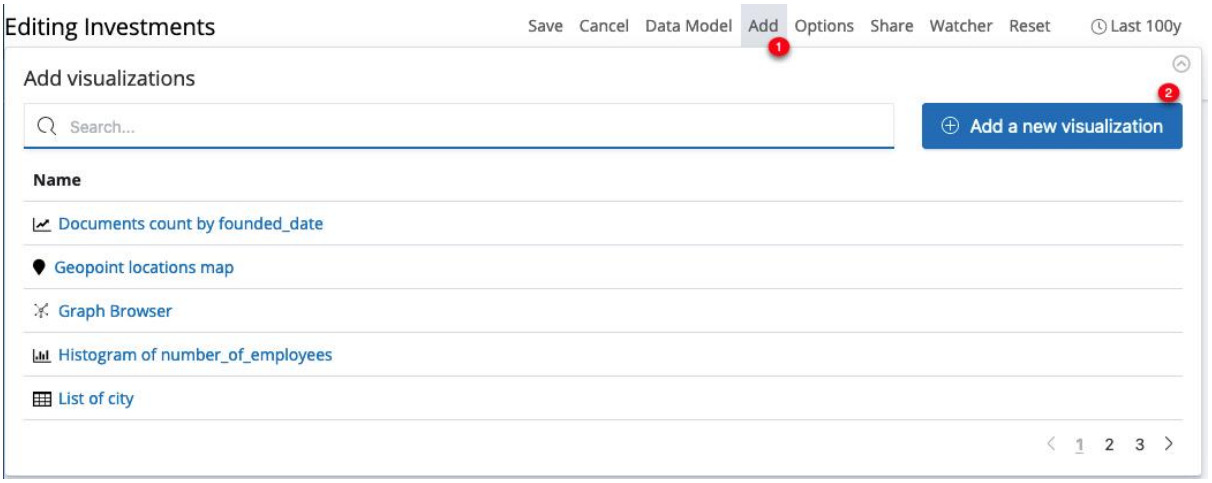
We will add visualizations in the following order:

- Vertical Bar chart
- Record Table
- Analytic Table

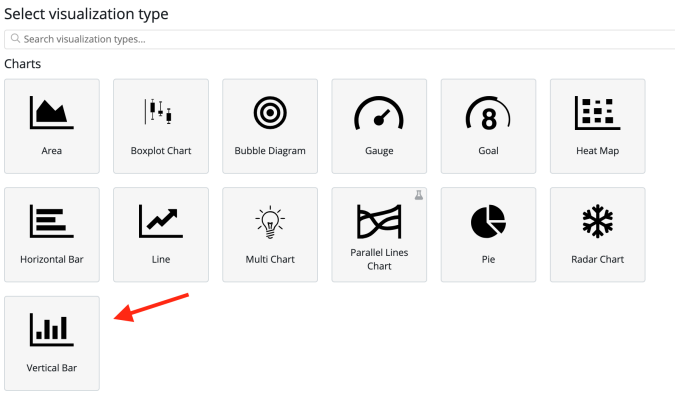
We will add the visualizations directly from the current dashboard. You can also add them by using the **Visualize** app.

Vertical Bar Chart

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

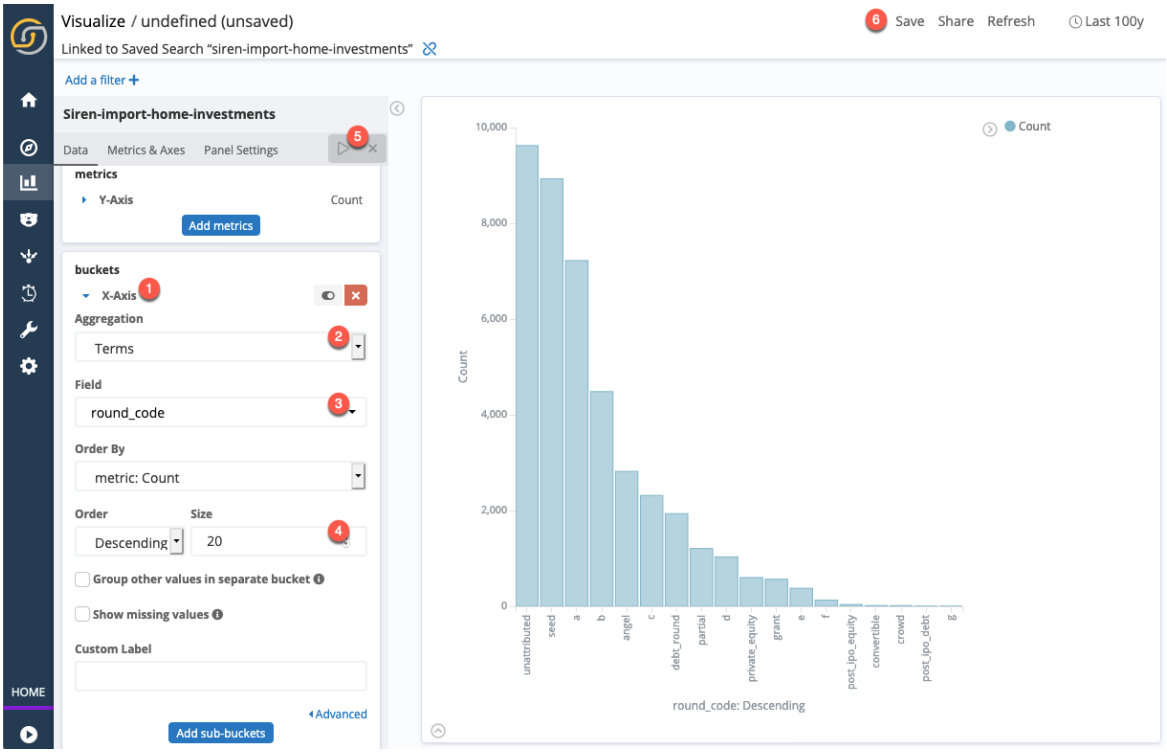


2. Click **Add a a new visualization** ②.
3. Click **Vertical Bar** in the **Charts** category.



4. On the **Select an entity table or search** screen, select the **investments** entity table.

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** ③
8. In the **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 new visualization**, as before.
2. Use the search bar or scroll to the Showing individual records category to find **Record Table**.
3. On the **Select an entity table or search** screen, select **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 new visualization**, as before.
2. Use the search bar or scroll to the Textual category to find **Analytic Table.**
3. On the **Select an entity table or search** screen, select **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 **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 an entity table**: Used when the dashboard's main function is to provide summary information and it does not need to be linked to an entity table. These dashboards do not display a record count or an icon.
- **Dashboard represents an entity table**: Used when the dashboard represents a single specific entity table (for example, companies or investments), enabling the following:

    - The dashboard displays the count and an icon associated with the entity table.
    - 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 entity tables, 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** entity table. The logical thing to do is to specify that this dashboard is dedicated to the same entity table, and use the appropriate dashboard data model.

1. On the **Investments** dashboard, click **Edit**, then click **Data Model** ①.
2. Select **Dashboard represents an entity table** ②, and then select **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!

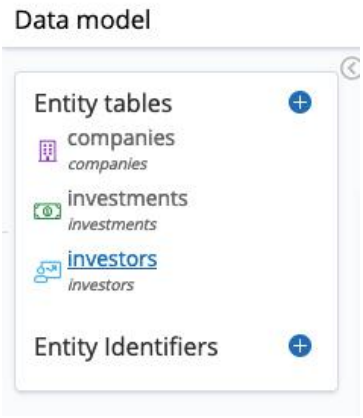
The next file you need for this tutorial is the **investors.csv** file.

In the **Data model** app, use the same procedure that we used previously to import the **investors.csv** file and name the entity table investors.

1. In the **Field mapping** section, accept the default mappings and click **Next**.
2. On the Transform data screen, accept the default settings and click **Next**.
3. Click **Start loading**.

4. On the **Info** tab, select an icon to represent investors. (We like **fal fa-user-chart**.)
5. Select a color for the entity table icon.
6. In the **Default label** field, select **Document Field** and, in the second dropdown menu that appears, select the field **label**.
7. Select **Disable the Time filter in dashboards**, as this is not a time-based entity table.
8. Click **Save** in the top-right corner of the screen.

The **Entity tables** section of the Data model app should now contain three entity tables: **companies**, **investments**, and **investors**.



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

Introducing relations in the Data Model

The true power of Siren Platform is in its *associative* data model; defining how your entity 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.

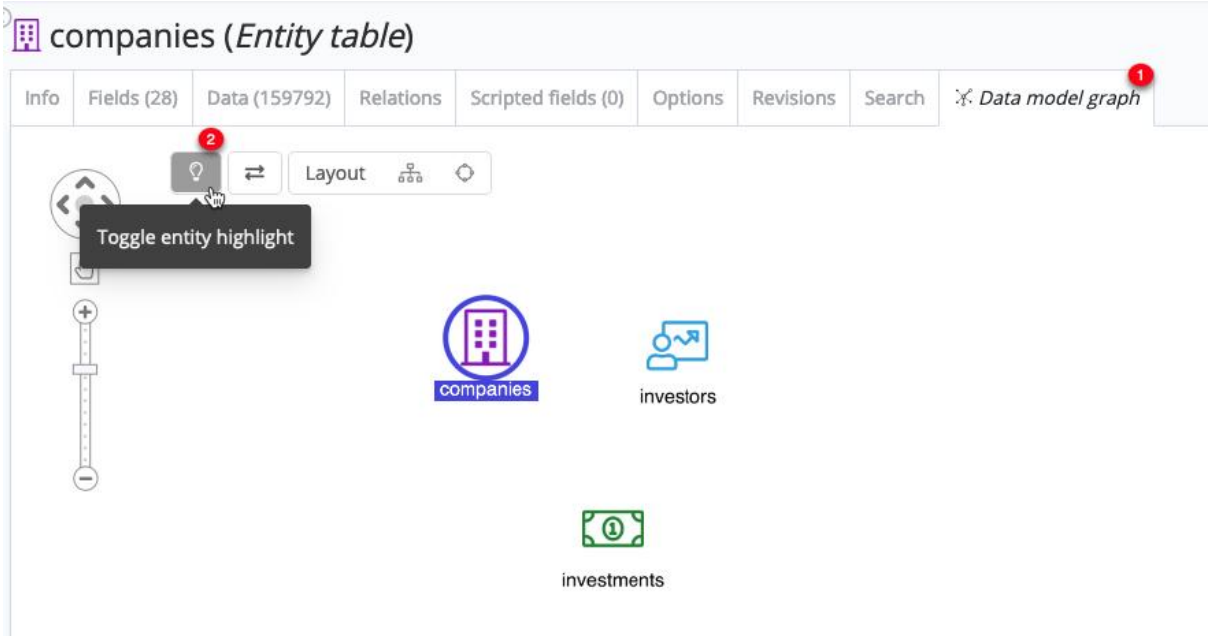
You create these relations in the UI. For example, the Investments table is connected to the Companies table as follows:

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.

Creating relations between entities

In the **Data model** app. Select the **companies** entity table 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.

Creating relations manually

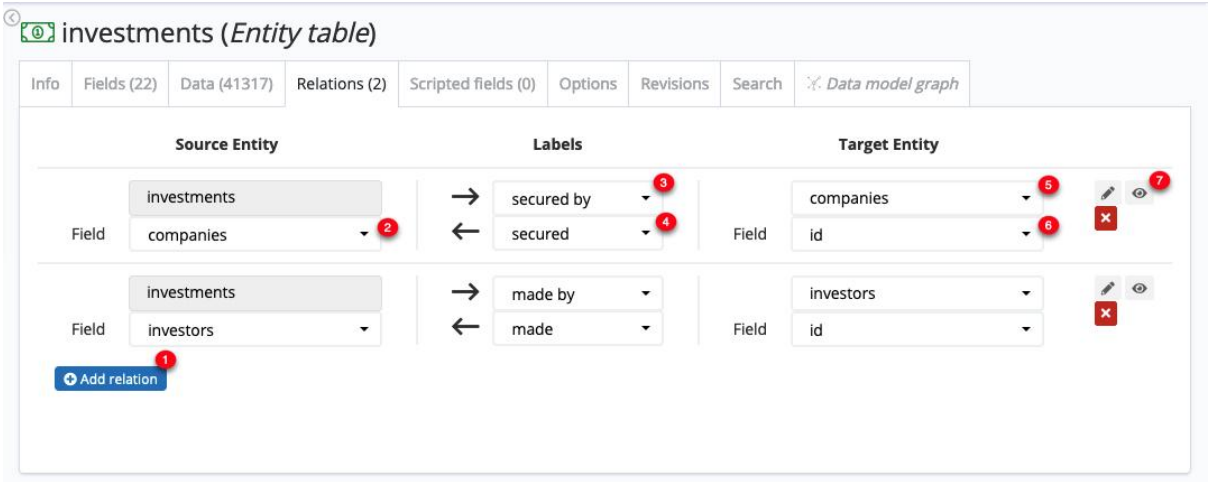
(Average completion time: 5 mins)

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; a unique value in the companies data. We will base this on the relational model of the data.

To create this relation, go to the **investments** entity table, click the **Relations** tab, and complete the following steps:

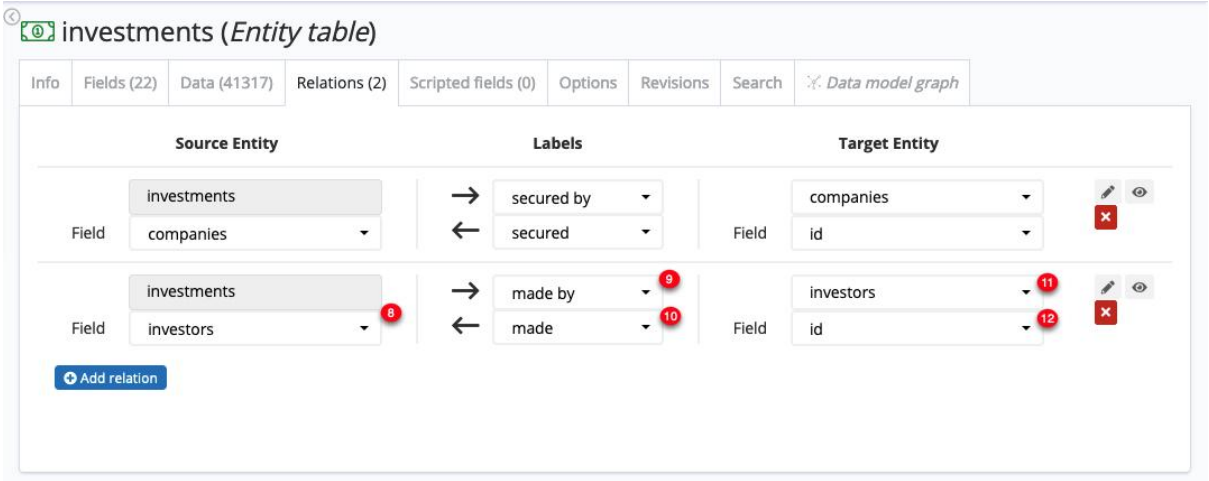
1. Click **Add relation** ①. The source entity is preselected as **investments**.
2. In the **Field** field, select **companies** ②.
3. In the **Labels** fields, enter "secured by" in the top 'active' field ③ and "secured" in the bottom 'passive' field ④.

4. In the **Target Entity** field ⑤, select **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** ⑧ from the list of searches.
3. In the **Labels** fields, enter “made by” in the top field ⑨ and “made” in the bottom field ⑩.
4. In the **Target Entity** field, select **investors** ⑪ from the list of entity tables.
5. In the **Field** field, select **id** ⑫ from the list of fields.
6. Click **Save**.

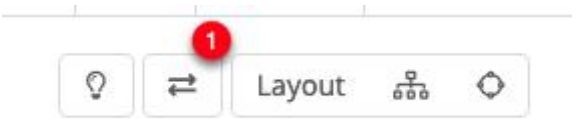


Visualizing the data model as a graph

(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.
- Click **Save**.



Creating a self-relation

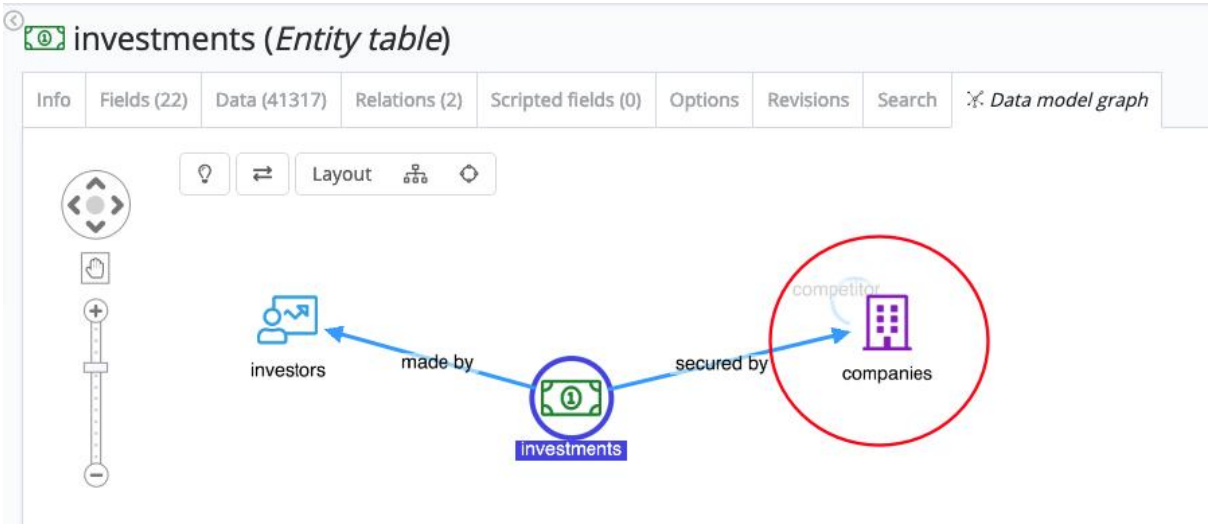
(Average completion time: 2 mins)

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 **companies** icon. This switches to the **companies** entity table.
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.

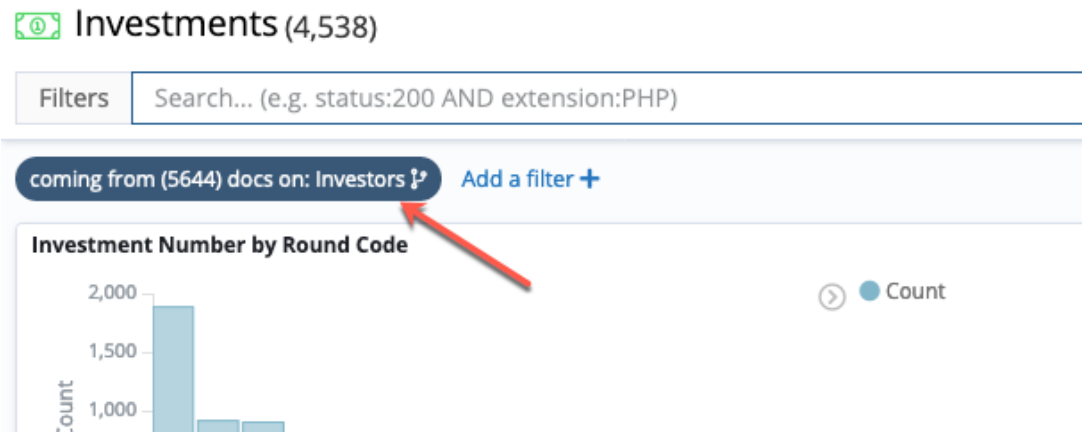
Since we have already defined the relations for 'investors', we can create a dashboard for it by using the wizard as follows:

1. In the **Data model** app, select the **investors** entity table.
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, labeled **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.

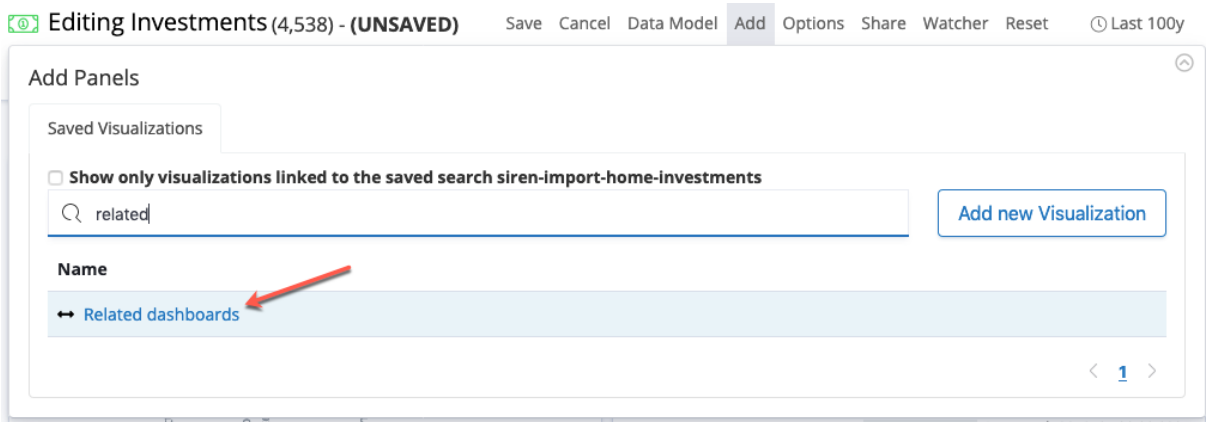
3. Filtering the records on the dashboard will automatically show a filtered count on the relational navigator link. For example, if you click on the 'person' section of the **Top 30 investortype** pie chart. The record count on the dashboard and the count of investments are reduced as a result.
4. Click 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'*.



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**.

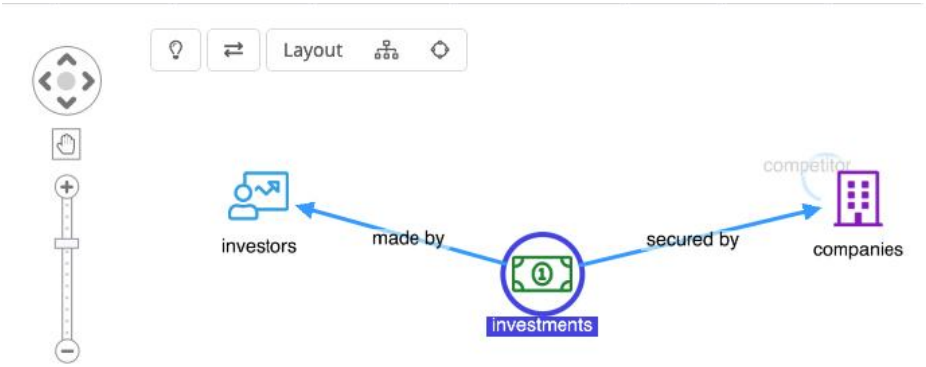
The **Related dashboards** visualization already exists in the system. Therefore, you can simply search for it in the list of existing visualizations, click to add it to the dashboard (a message indicates that it was added to the dashboard successfully), and click **Save -> Save dashboard**.



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.

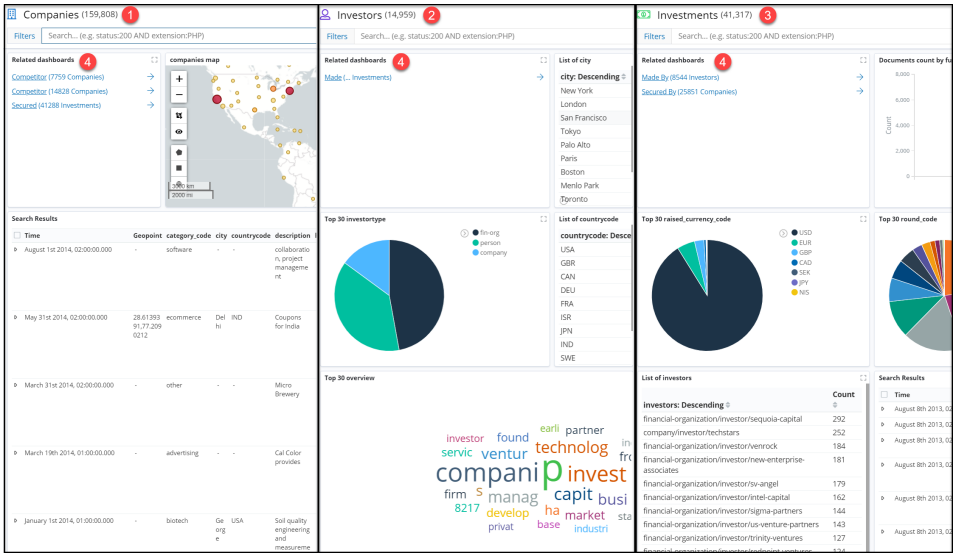
Recap: Where are we now?

If you completed all of the previous steps of this tutorial successfully, your data model should look something 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 entity table.

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

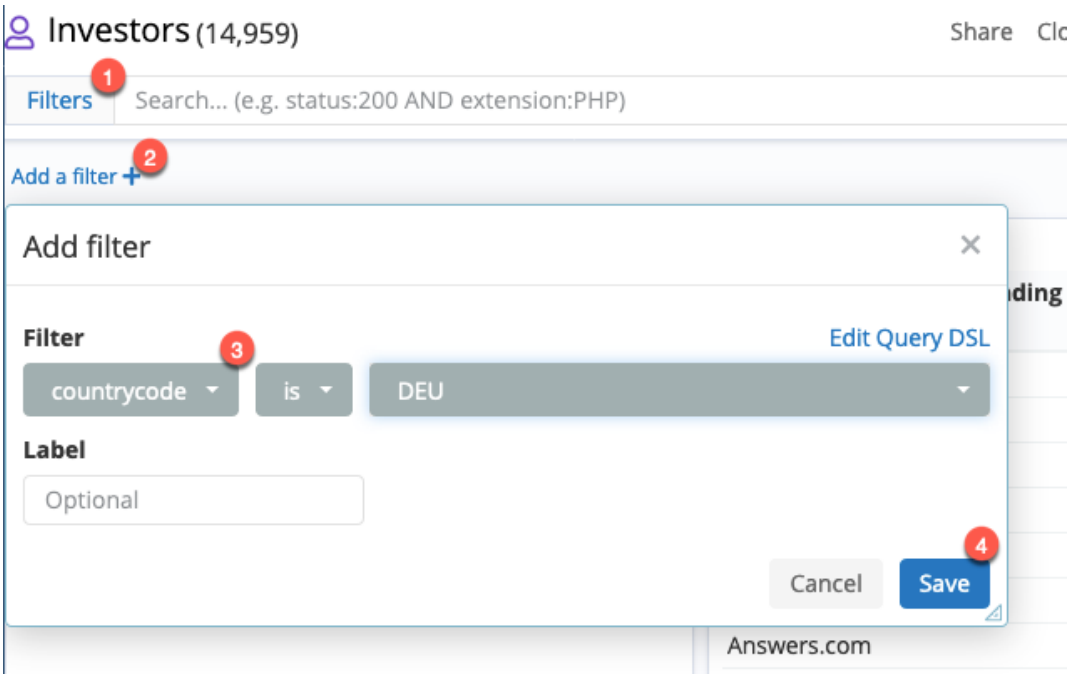
Question: How many investors are from Germany?

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.



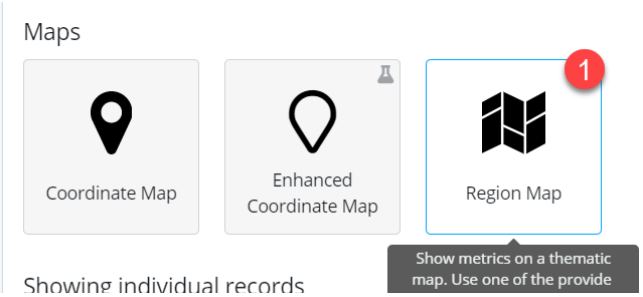
4. Click **Save** ④. The answer appears as the number in parentheses alongside the dashboard name.

Advanced solution (optional)

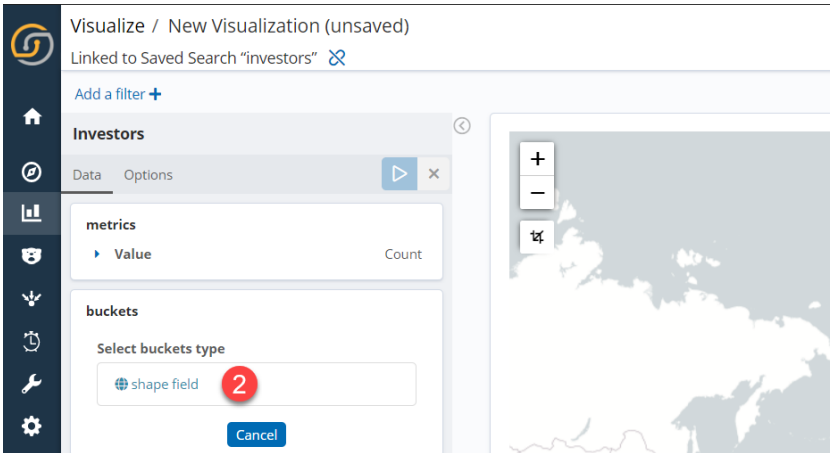
You can also answer this question by adding a **Region Map** visualization.

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.

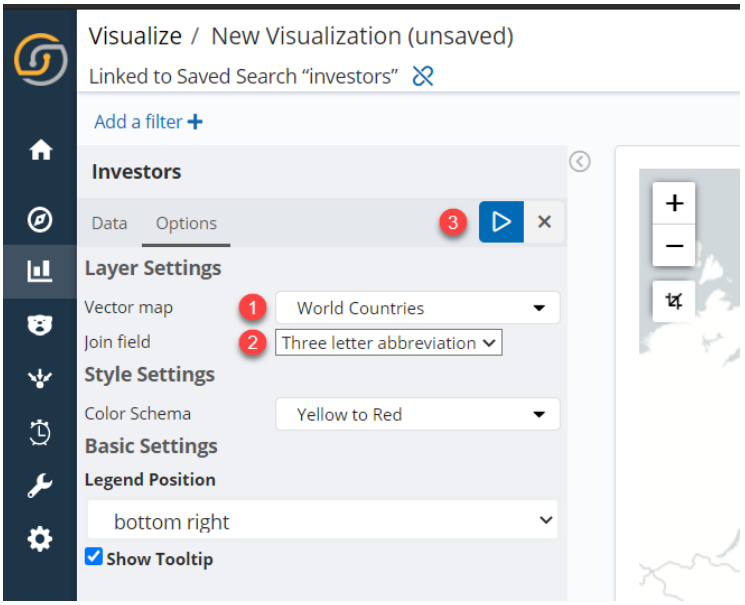


2. In the **Select an entity table or search** window, select the **investors** search.
3. In the **buckets** table window, select **shape field** ②.



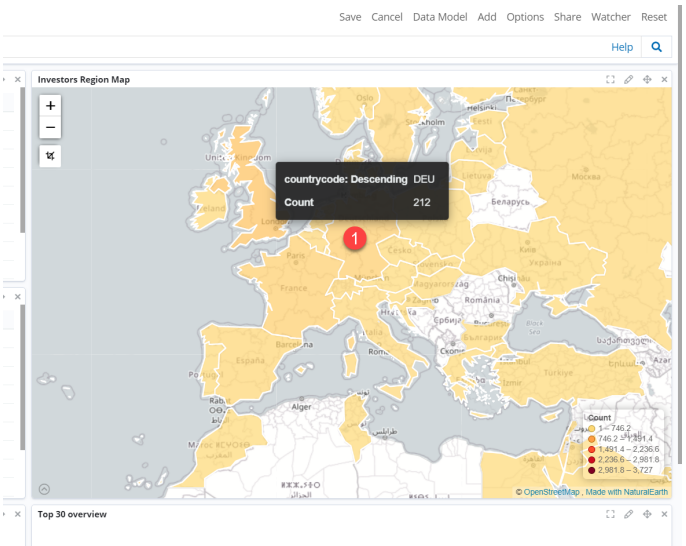
4. In the **Field** field, select **countrycode**.
5. We want to see data from more than five countries, so in the **Size** field, increase the value to **35**.

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



7. In the **Vector map** dropdown menu, select **World Countries** ① and in the **Join** field, select **Three letter abbreviation** ②.
8. Finally, click **Apply changes** ③.
9. Click **Save**, name the visualization "Investors Region Map", and click **Save to Dashboard**.
10. 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:

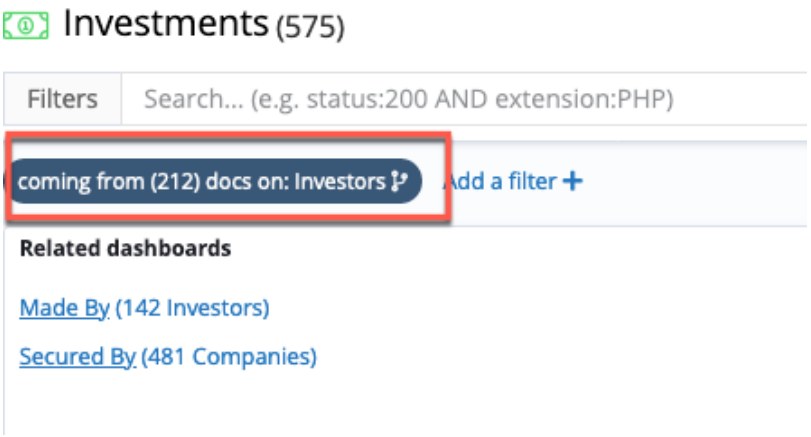
Solution

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)**.

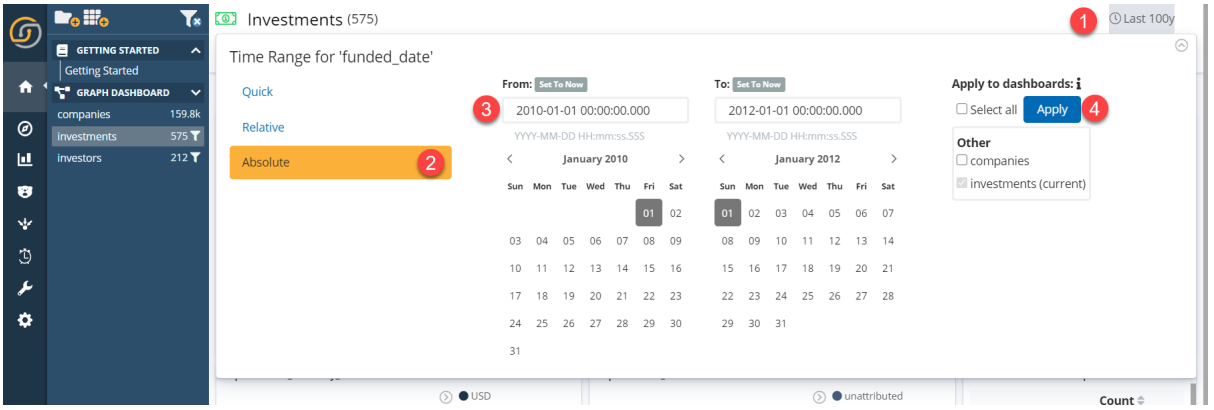
On the **Investments** dashboard, a relational filter is automatically created:



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.

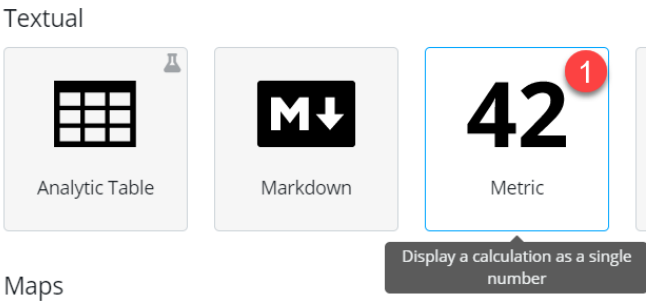
Let's fix that!

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.



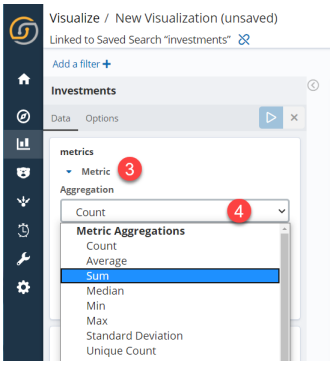
Perfect! We only need one last step to answer this question, which is, adding a metric to compute the total amount of the 176 investments that are left.

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

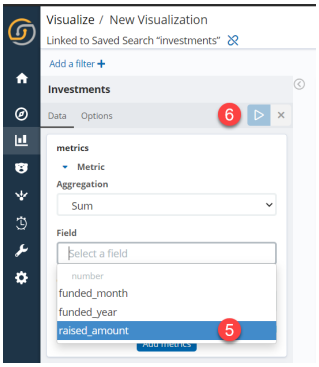


5. From the **Select an entity table or search** menu, select **investments**.

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

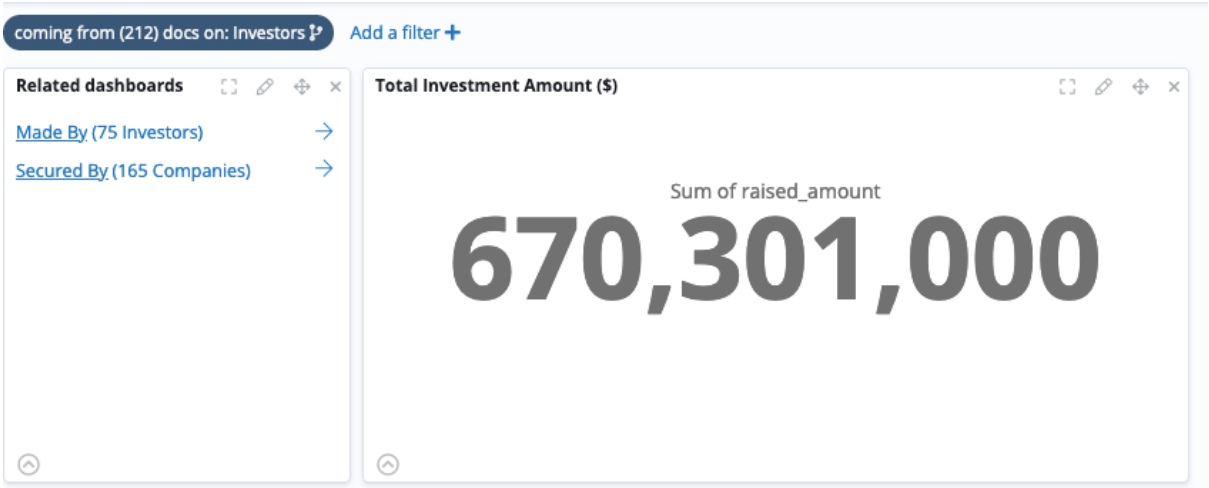


7. In the **Aggregation** field, select **Sum** ④.
8. In the **Field** field, select **raised_amount** ⑤.



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

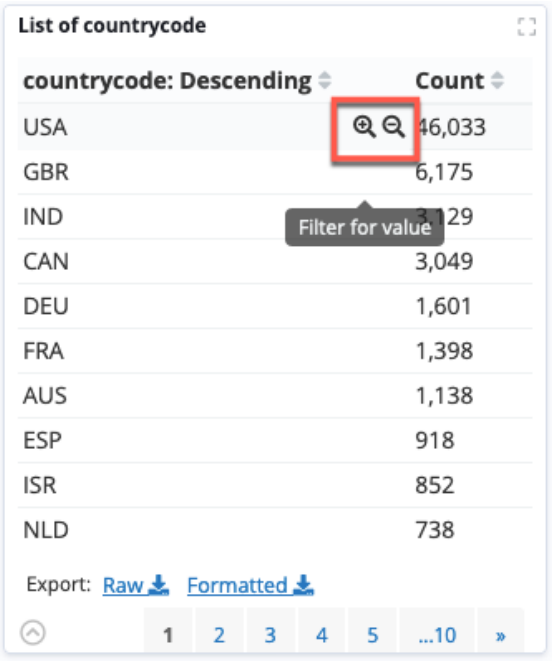
Question: How much of these investments were secured by American companies?

Answer:

Solution

To solve this question, complete the following steps:

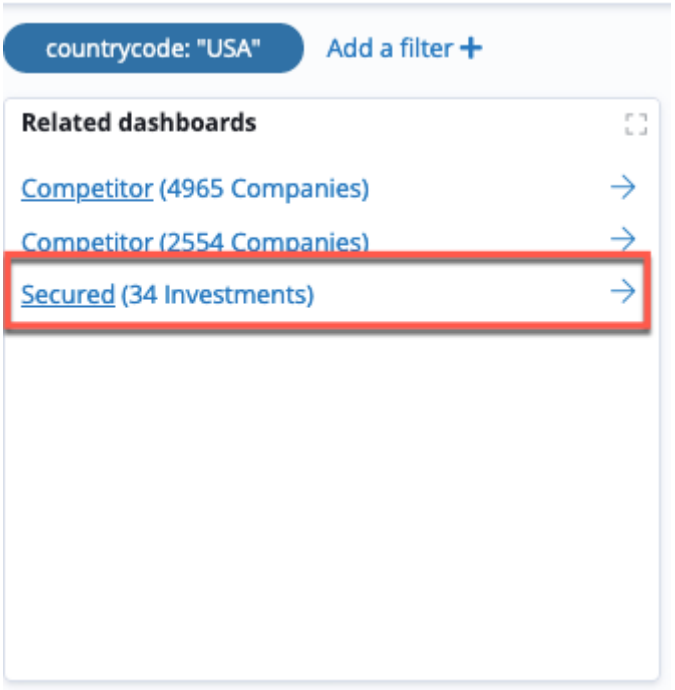
1. Click on the **Companies** dashboard in the sidebar (do not navigate there by using the relational navigator). This will show you all of the companies (approx. 159k).
2. Create a filter for country code **USA** by clicking on **Filter for value** (the magnifying glass icon) in the **List of countrycode**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



Now, we have selected only American companies.

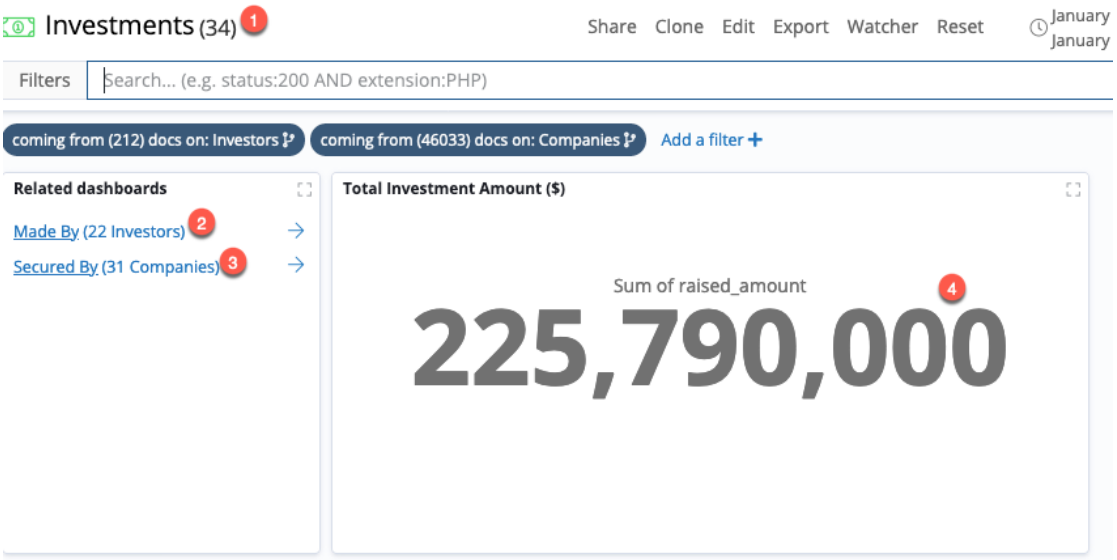
The **Relational Navigator** is showing us that **34** of the records in the **Investments dashboard** are matching. Notice that, at this point, the Investments dashboard is already filtered from our previous exercise (it contains the filter that displays only the investments by German investors).

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)**.

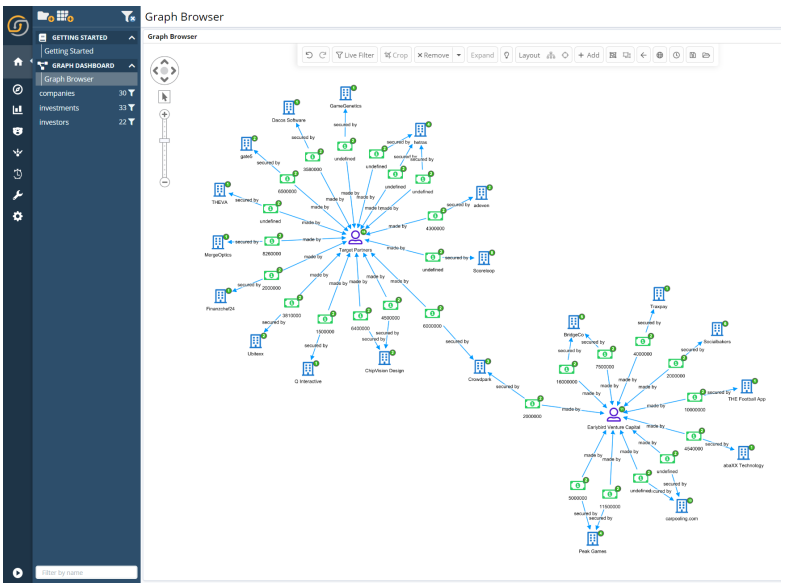
If everything was done correctly, your filtered dashboards should now display:

- 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)

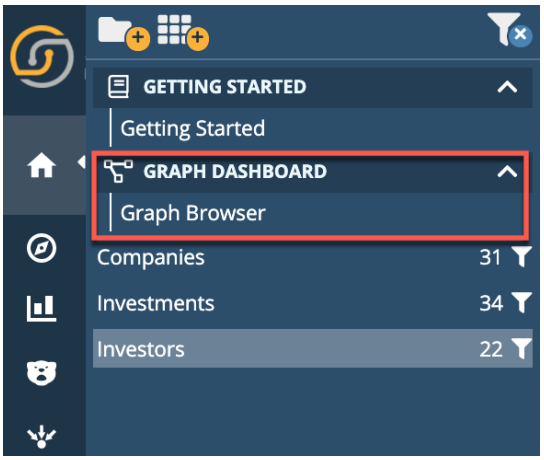
CAUTION: If your dashboards don't have the correct filters applied and a large number of nodes are added, you might experience problems with system performance.

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.

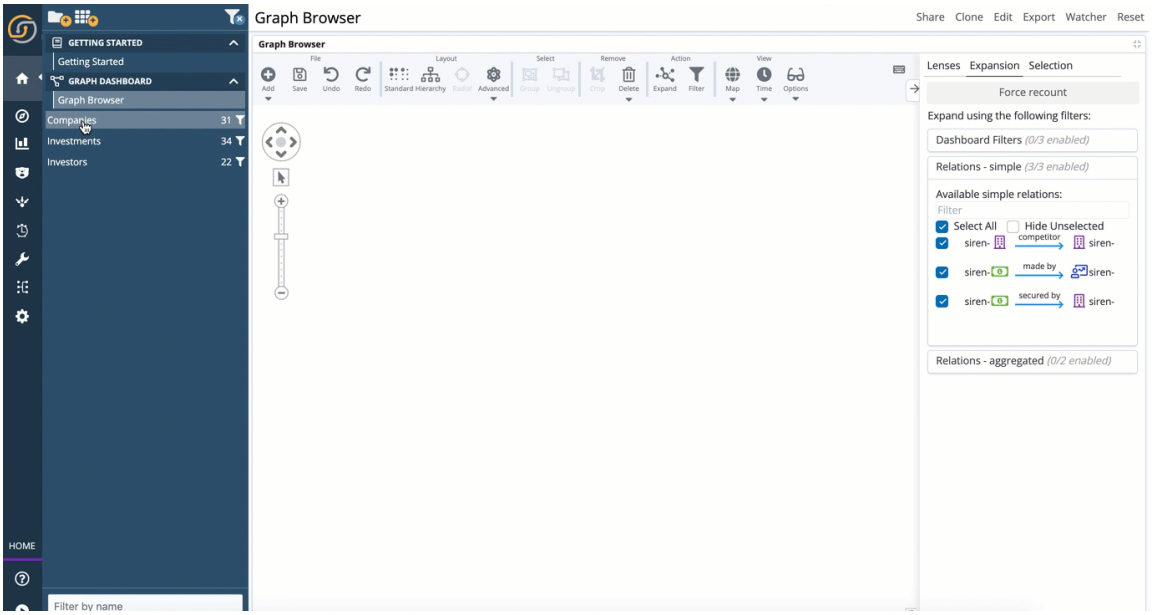
In this section of the tutorial, we'll get to see graphs like these:



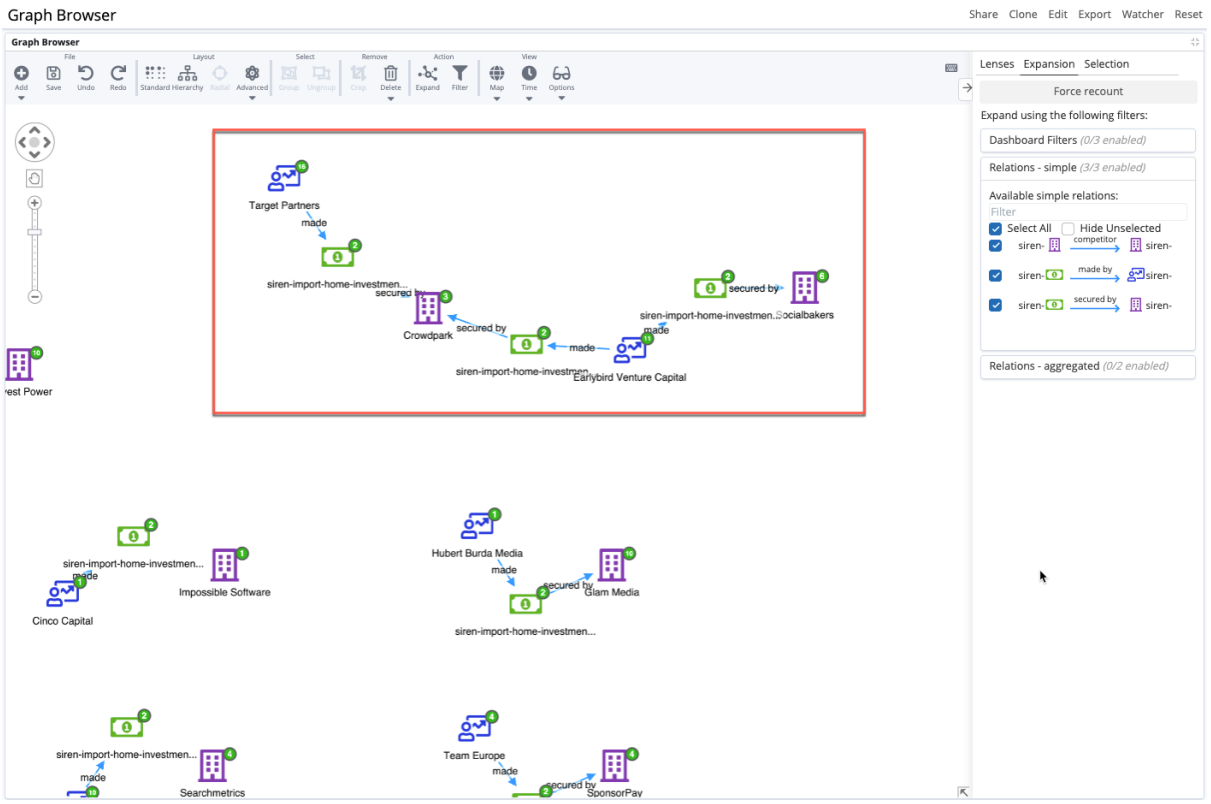
1. To start exploring the data, expand the **GRAPH DASHBOARD** group and click **Graph Browser**.



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

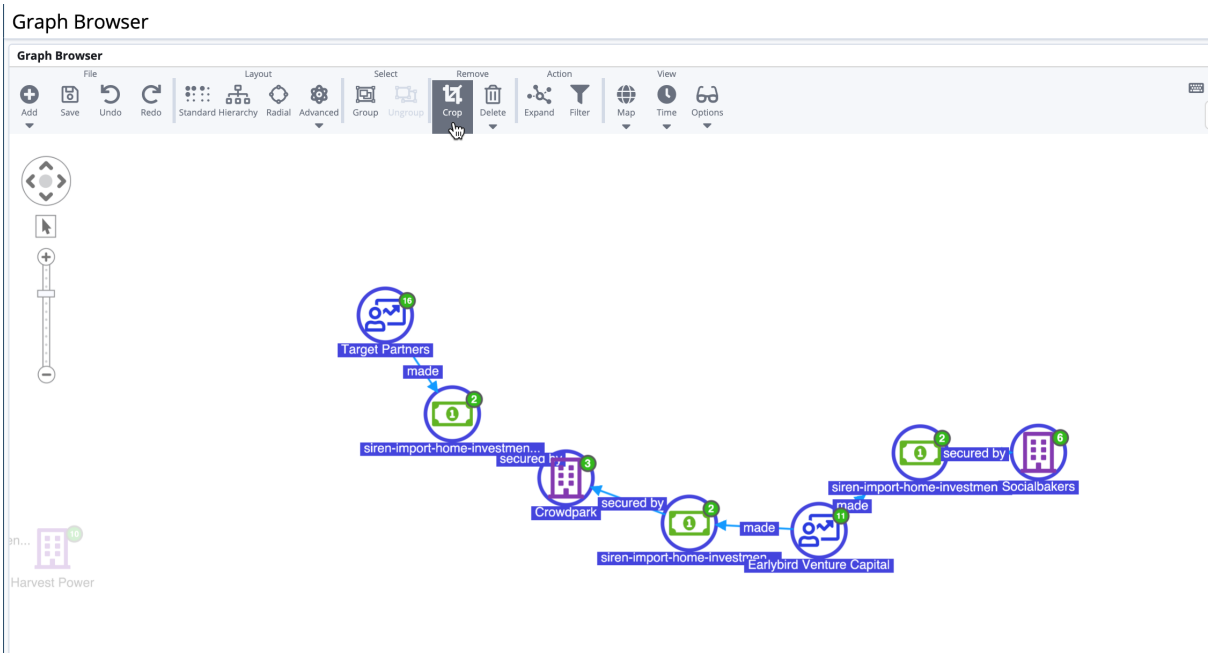


3. Click **Standard** from the Layout menu and the nodes will be rearranged. We can observe, from this layout, that two German stockholders invested in the same company, named *Crowdpark*. Let's explore this small but interesting network.



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.

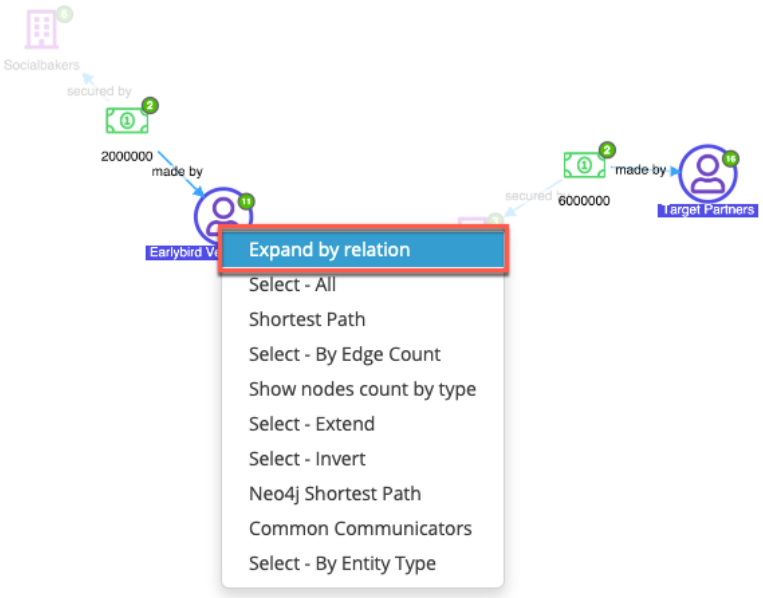
Before we proceed with the tutorial, take a moment to consider the links that might be interesting to explore. Any ideas?

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:

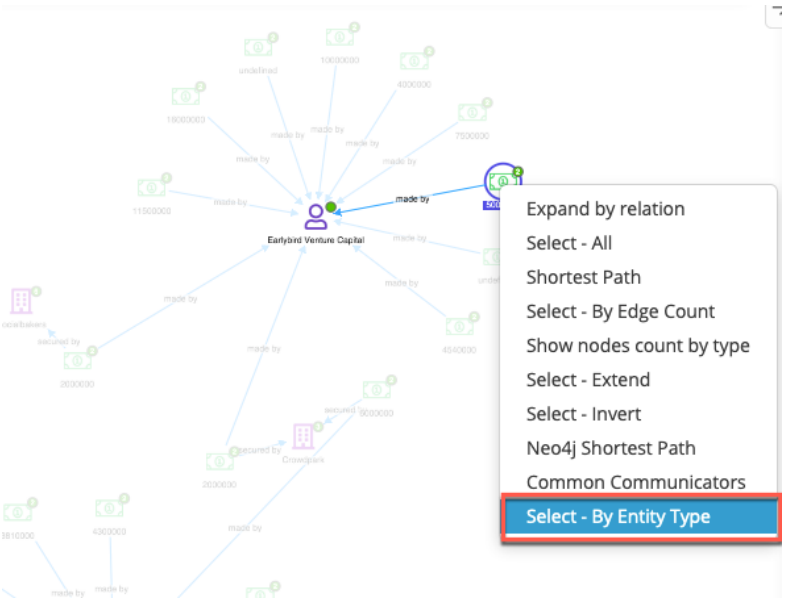
1. While in selection mode, hold the **Ctrl** key on the keyboard and click on the two investor nodes.
2. Once the two nodes are selected, right-click on one of them and, from the contextual menu, select **Expand by relation**.



3. Select the **made (27) *investments*** checkbox and click **Ok**.

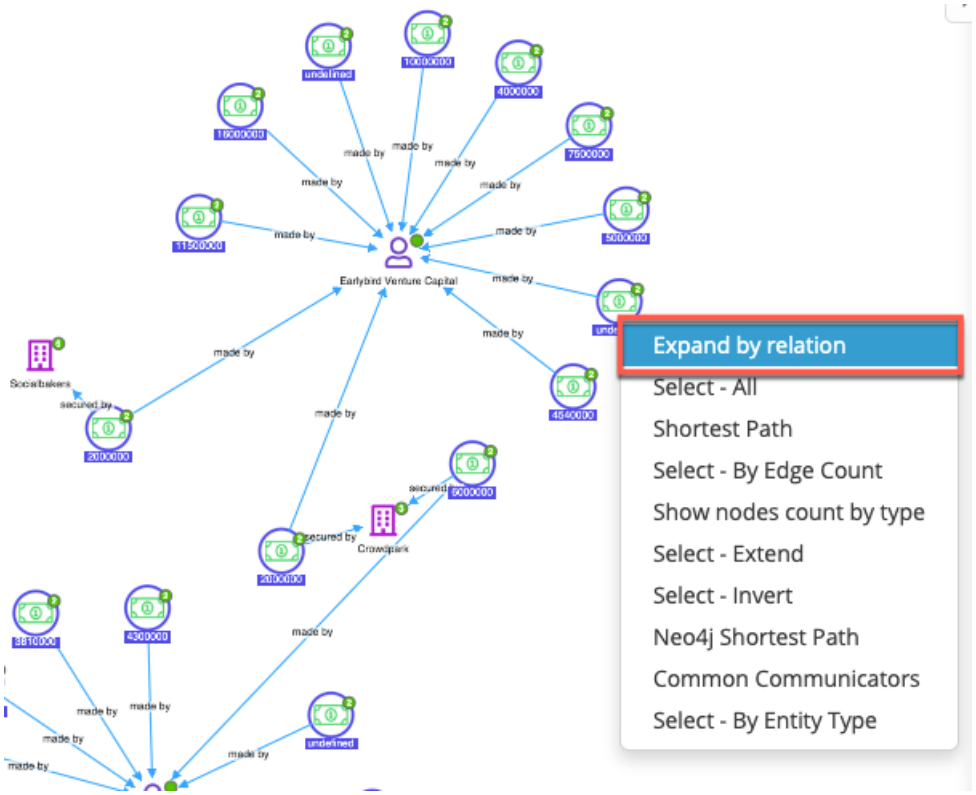
We will now investigate whether one of the new investments that appeared is paid to a new common company.

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**.

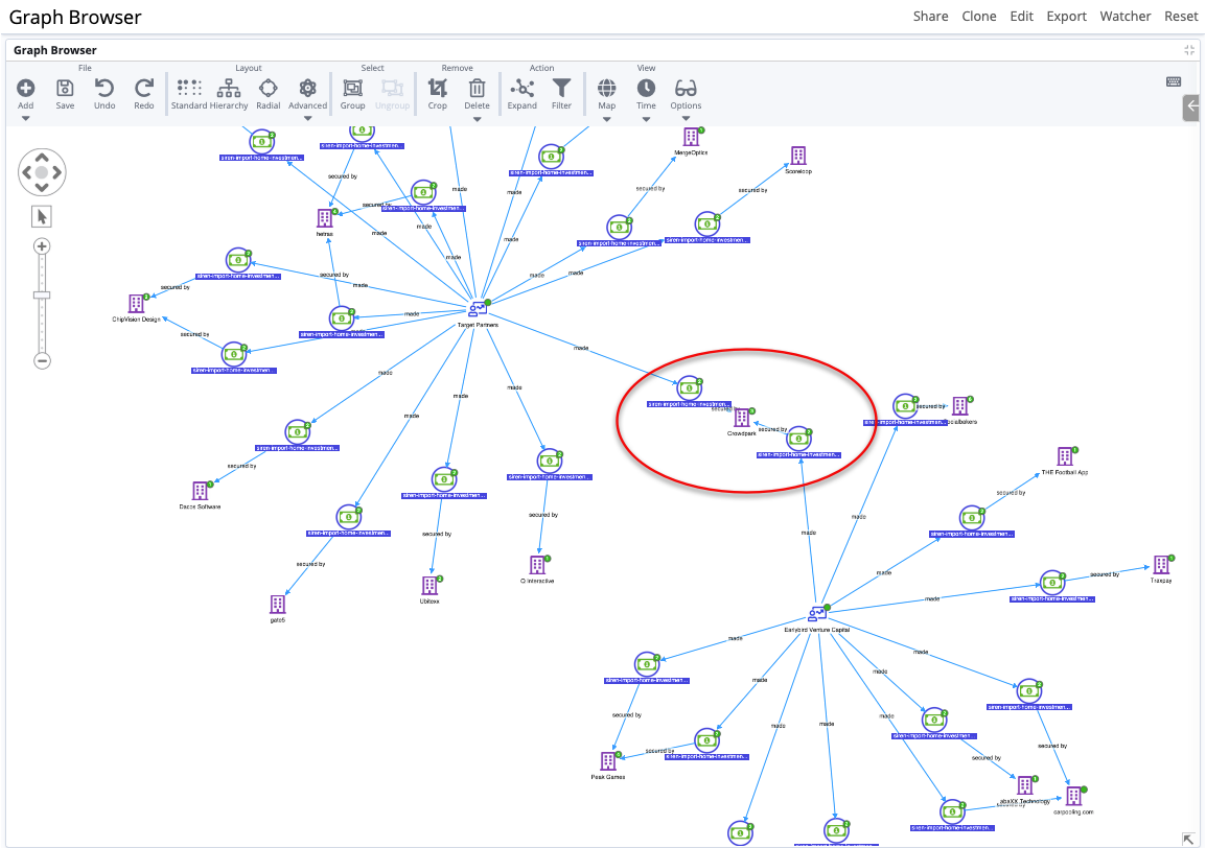
6. Now, all of the investment nodes should be selected. Right-click on one of the nodes and select **Expand by relation**.



7. 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*.

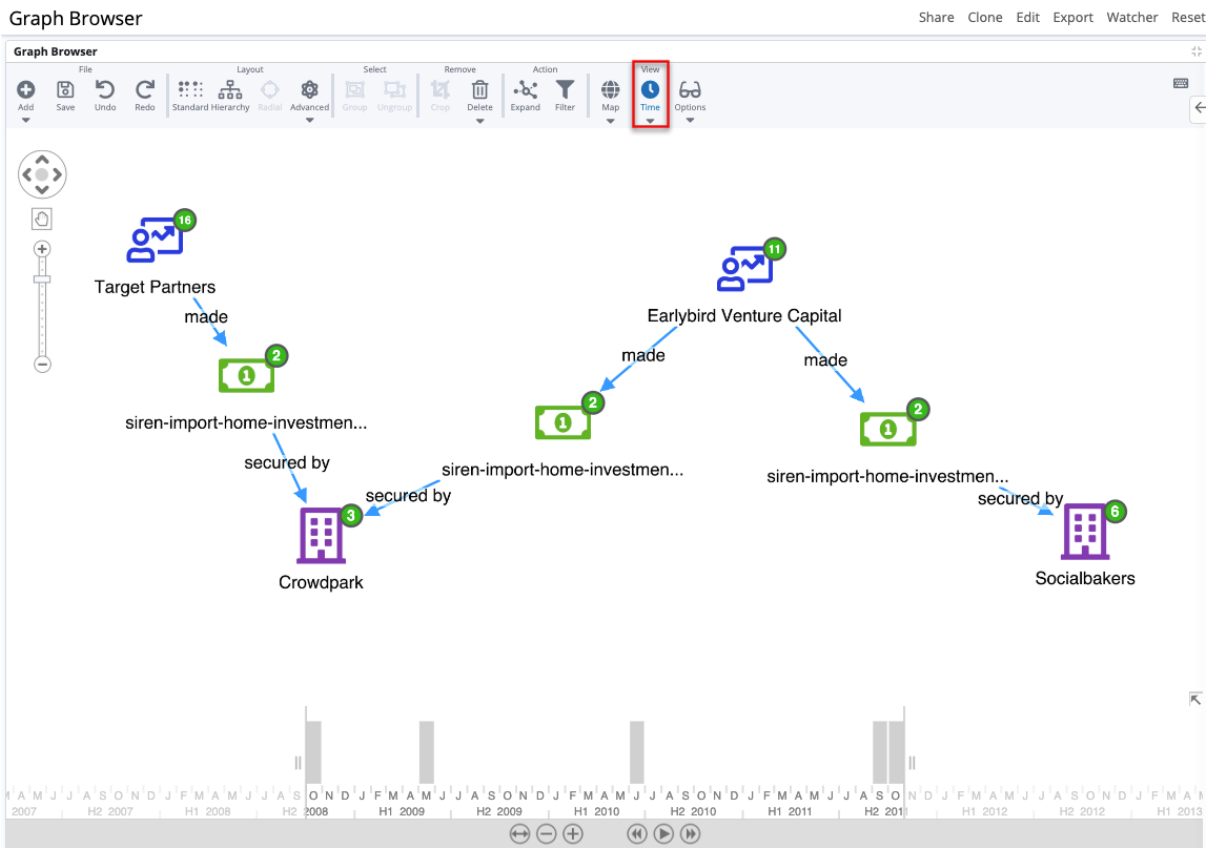
Q: Do the two German stockholders share additional companies in their funding campaign?
A: No.



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.

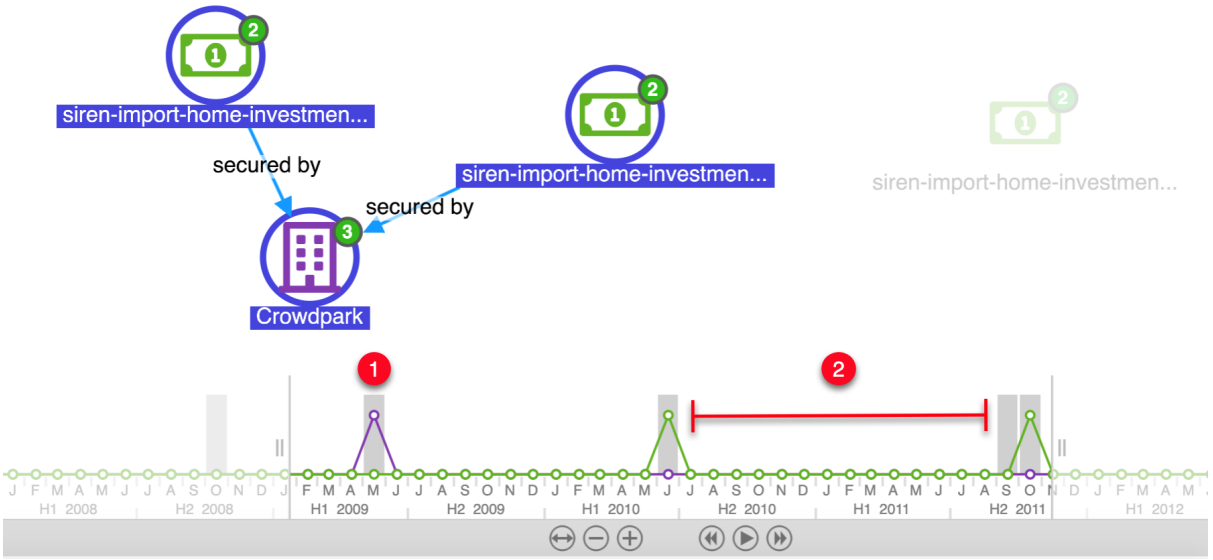
The second interesting question was: Did the two German investors fund CrowdPark in the same time period?

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.

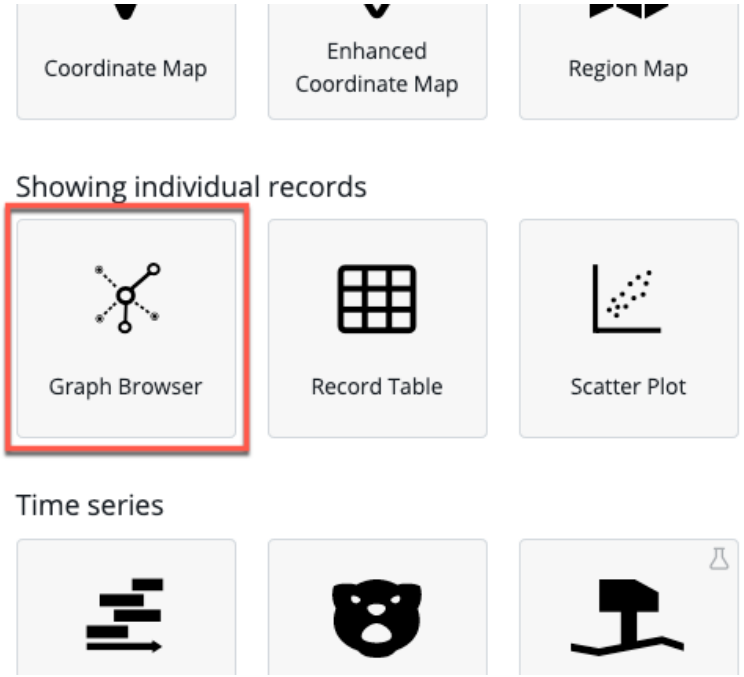
This concludes our exercise! If you want to know more about the graph, we will cover more material in the next chapter.

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:

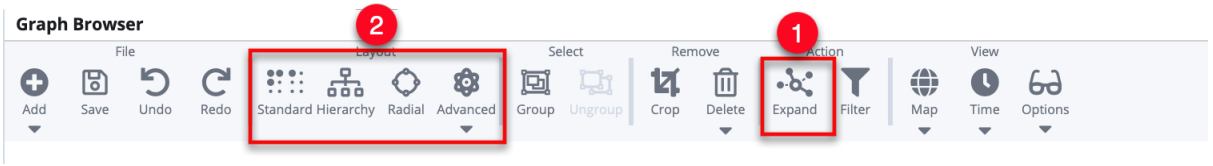


Using the Graph Browser

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.

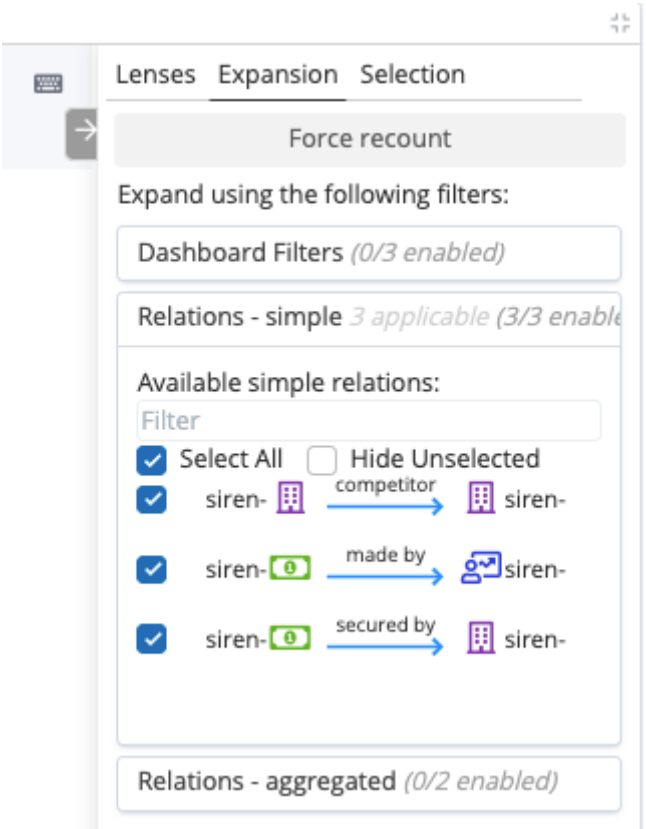
Clicking one of the **Layout** buttons ② reorganizes the nodes.



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

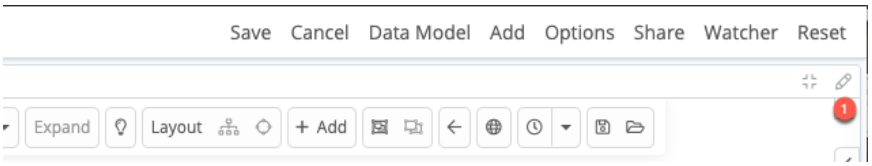
The side bar contains the following tabs by default:

- **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 the Configuring the Graph Browser section of this tutorial.
- **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.

Advanced features

This section of the tutorial teaches you how to use some of the more advanced features in Siren Platform.

Creating entity identifiers (EIDs)

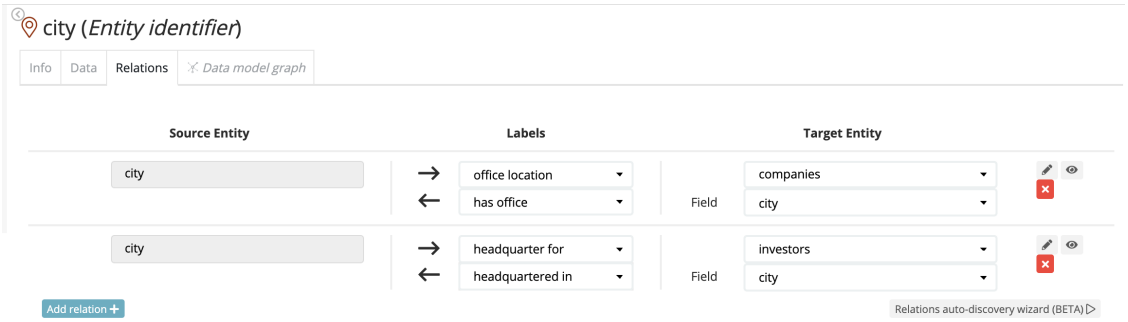
Siren expands its analysis capabilities by introducing Entity Identifiers.

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.

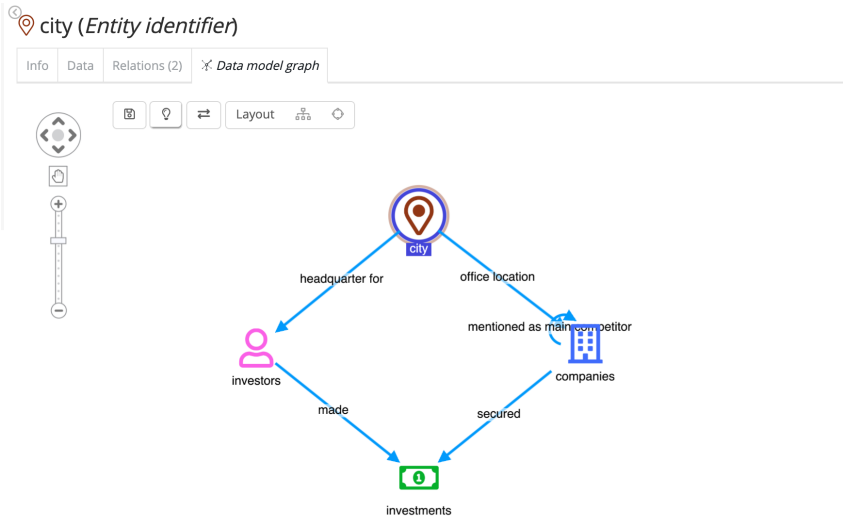
EIDs on the graph look like any other node, for example, an IP connecting two log records. They can also be used as ‘intermediate concepts’ when navigating dashboards.

To create an EID, complete the following steps:

1. Go to the **Data model** app and click **Configure an Entity Identifier** (the plus icon).
2. Enter the name as “city”, choose a suitable icon and color and click **Create**.
3. Click the **Relations** tab to create relations for the EID. We will create 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.**



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


city (Entity identifier)
Info
Data
Relations (2)
Data model graph
Layout
city
headquarter for
office location
mentioned as main competitor
investors
companies
made
secured
investments

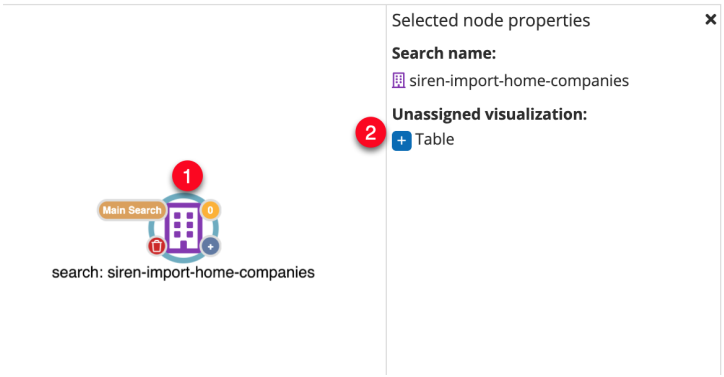
Creating a 360 dashboard

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.

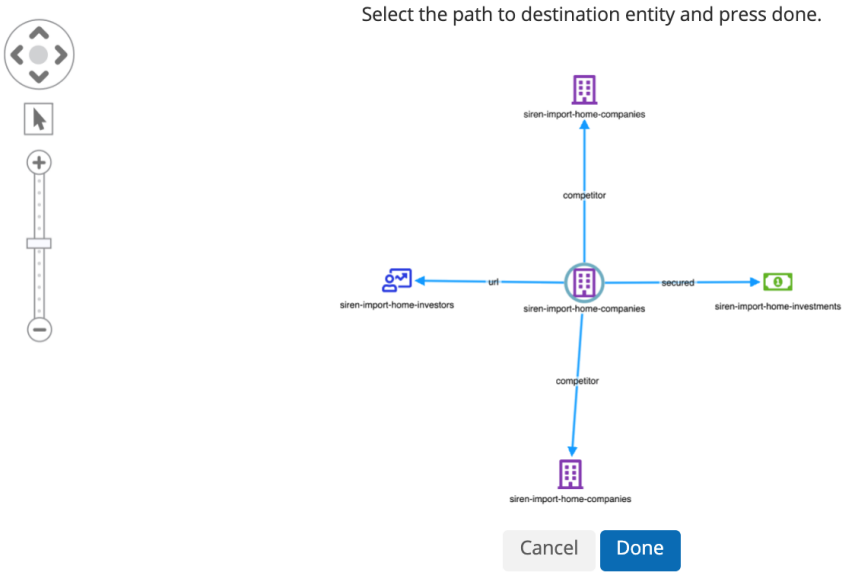
1. Click **Create new Dashboard** and name it “All Companies”.
2. Click **Add -> Add a new visualization**.
3. Select a **Record Table** visualization and associate it with the **companies** entity table.
4. Click **Apply changes**, name the visualization “Table”, and click **Save**.
5. Save the dashboard.
6. Click **Add -> Add a new visualization**.
7. Select a **Vertical Bar Chart** and associate it with the **investments** entity table.
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.

Because two entity tables are involved, we have to use a Dashboard 360. Complete the following steps:

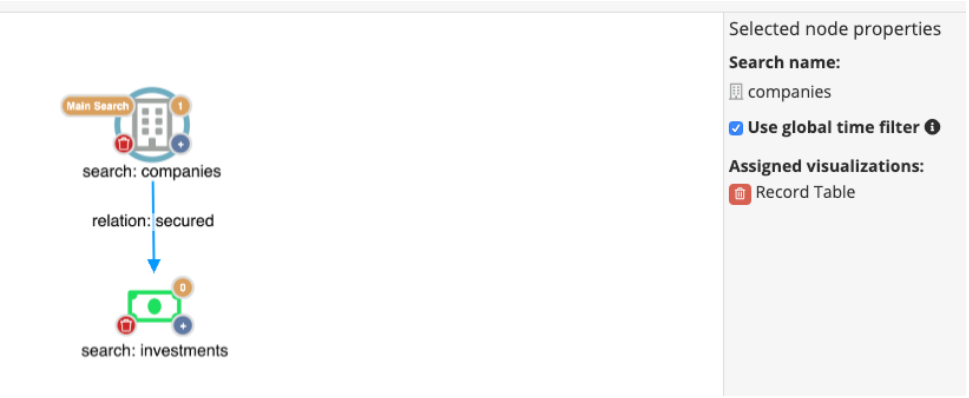
1. Click **Edit** → **Data Model** and select **Dashboard 360 with filter strategy** (leave the filter strategy as **Left filter join**).
2. Click **Add Main Search**.
3. In the dialog box, select **companies** as the main search and click **Done**.
   The Mapping visualization to searches panel shows 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.



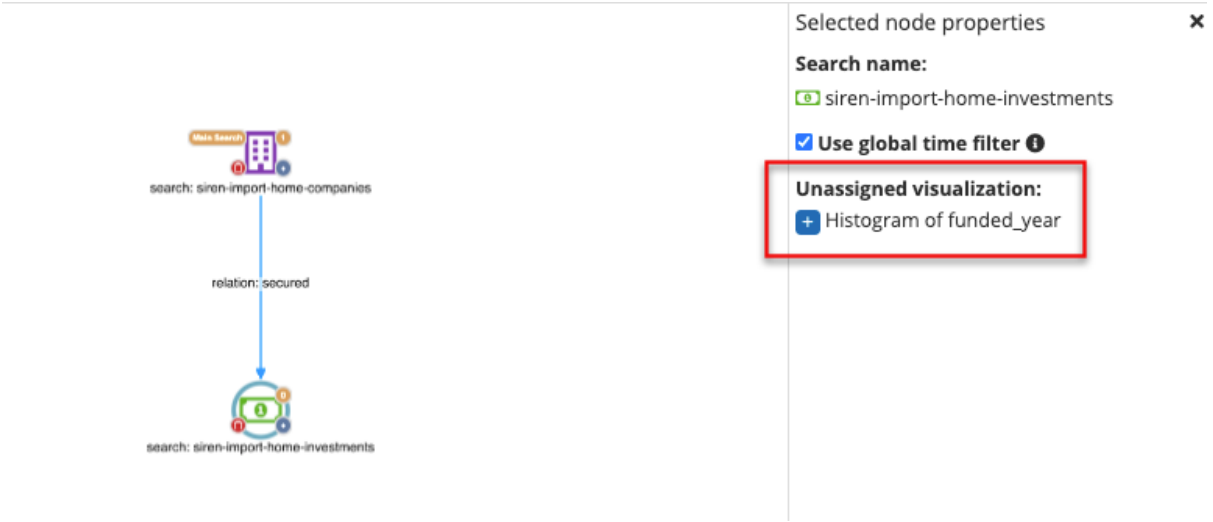
5. You can now build the dashboard data model by defining how this search is to be relationally connected with other searches (that is, which path in the relational data model to take). To do this, click the ⊕ button located on the lower-right of the Main Search node. This opens the following dialog 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.



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

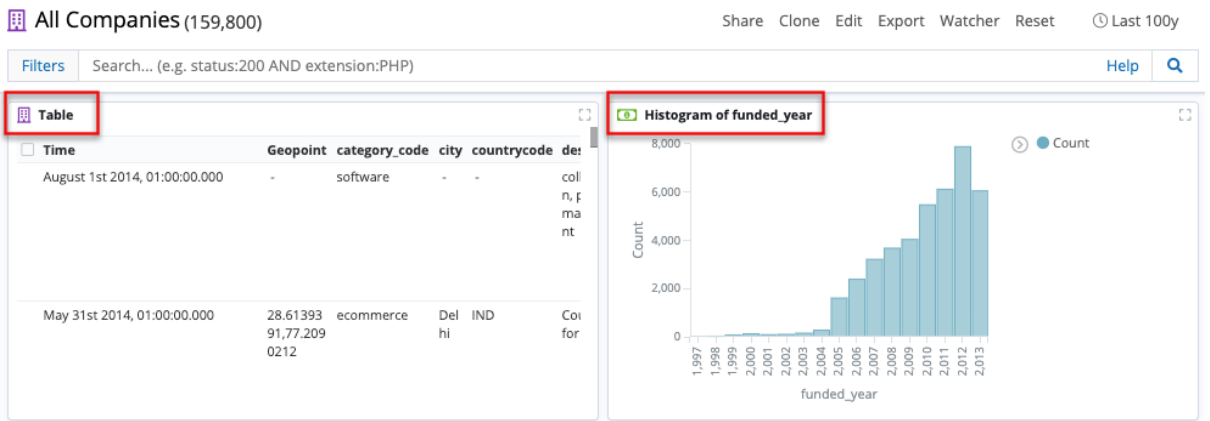


Select the **investments** node and you will see that it has an unassigned visualization.



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.

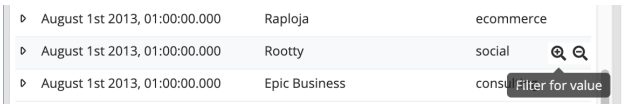


Filtering with Dashboard 360

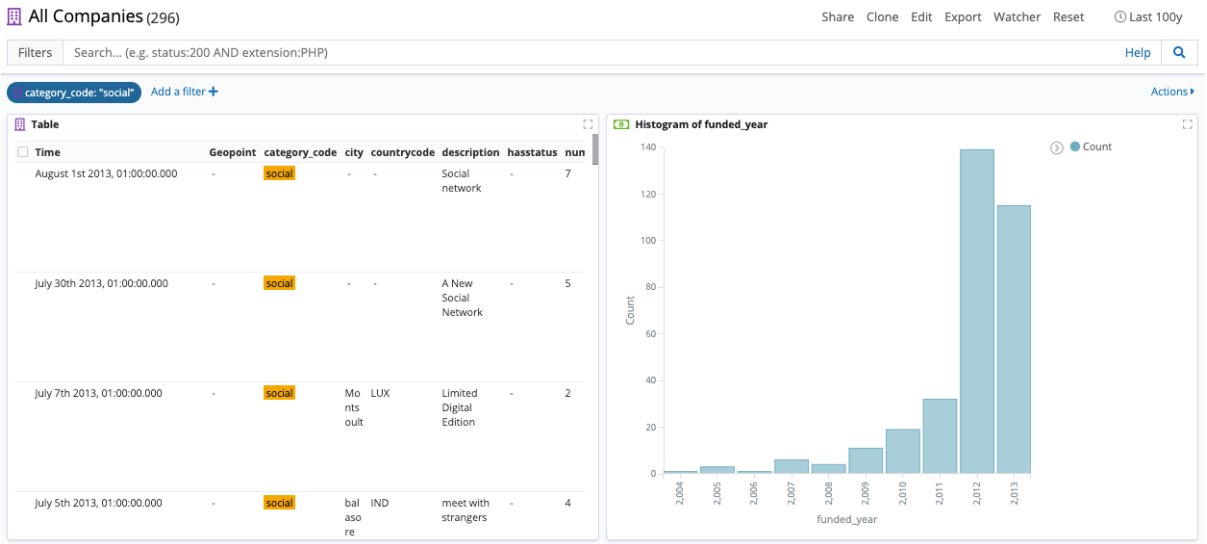
The really significant achievement here with Dashboard 360 is that you can perform coherent filtering across the dashboard's visualizations.

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**.



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.

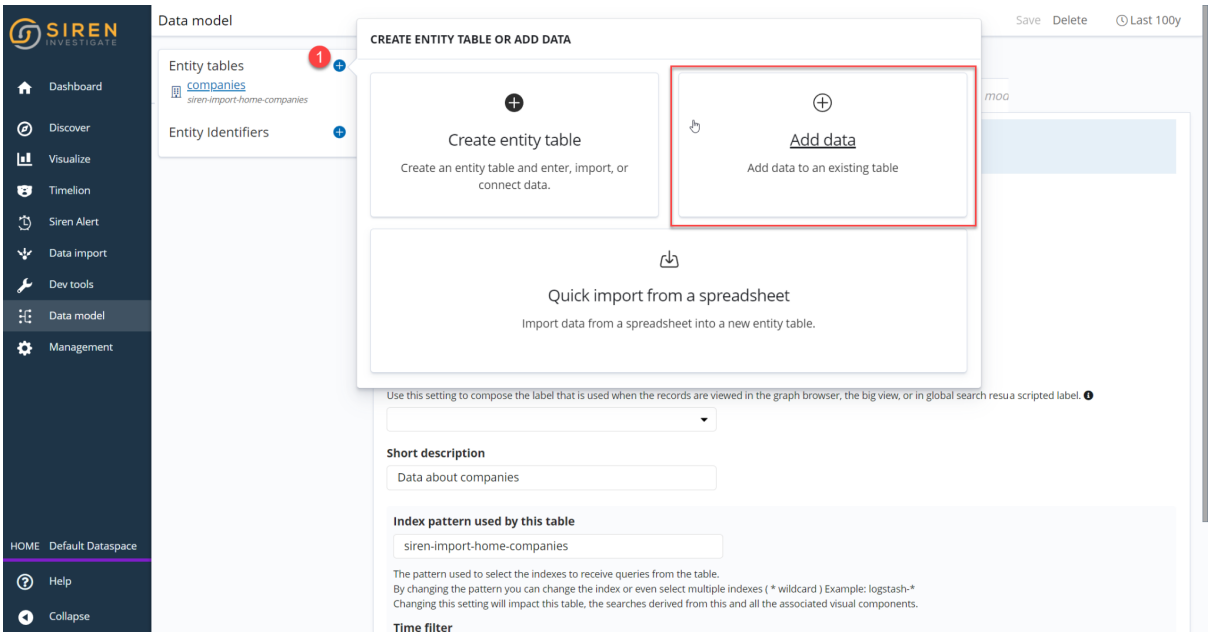
Importing and transforming data

During normal investigation processes, most of the time one does not need nor want to create an entity table. Instead, it is a very common use case to add data to an existing source, possibly adapting the original data format to the target structure.

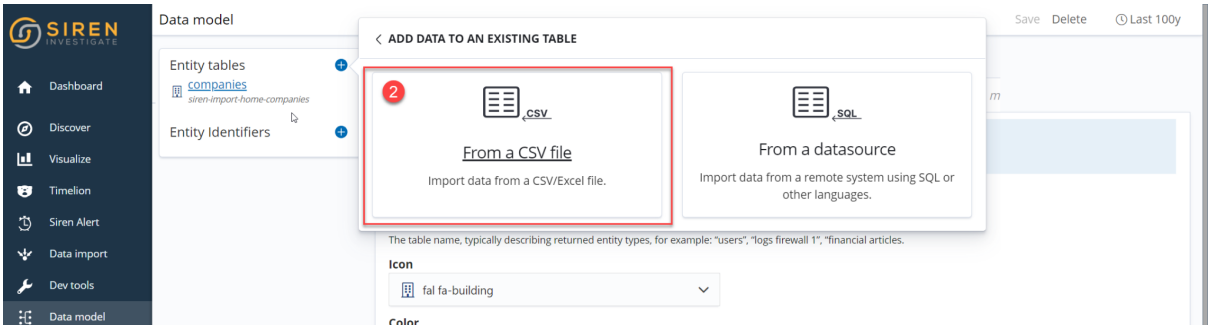
Siren Platform version 12 allows you to transform your data before adding it to an existing Entity table.

In this section, we are going to add records to the **companies** entity table with data coming from another CSV file.

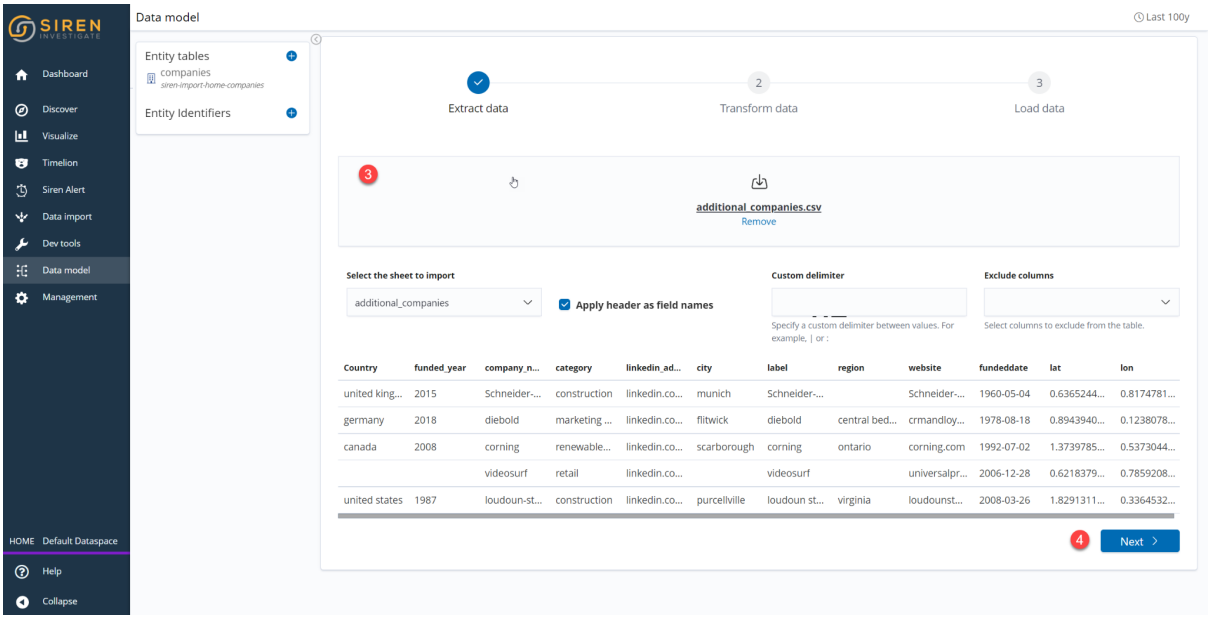
1. From the **Data model** app, click **Add/Create** ① and **Add data**.



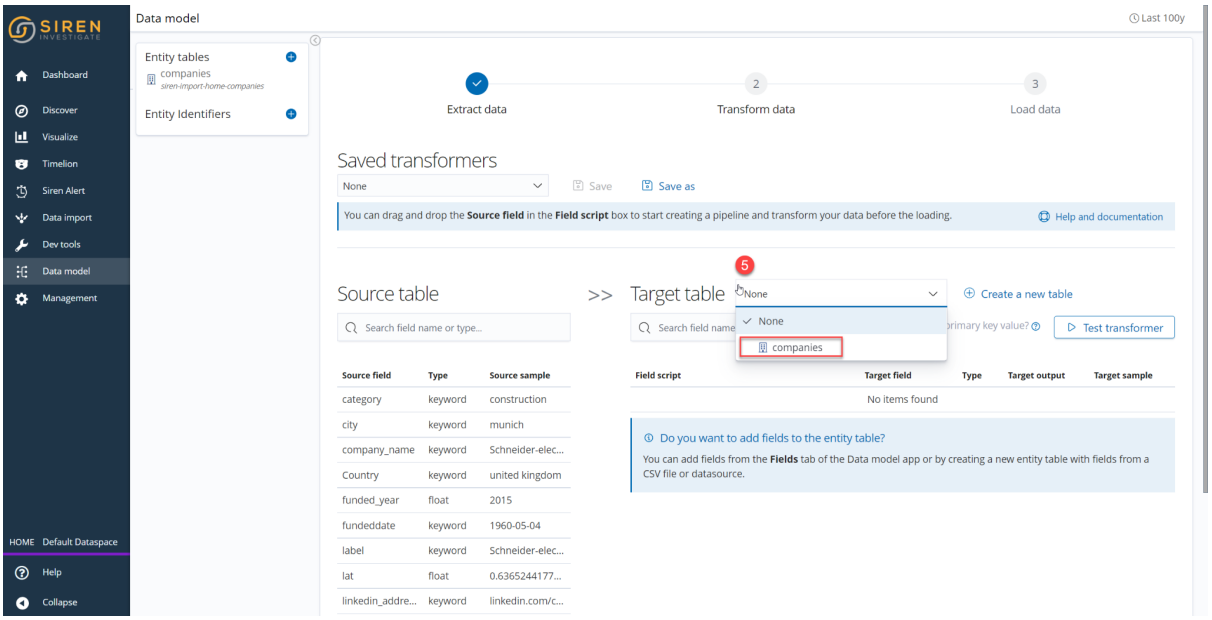
2. Click **From a CSV file** ②.



3. Drag the **additional_companies.csv** into the **upload area** ③.
4. Click **Next** ④.



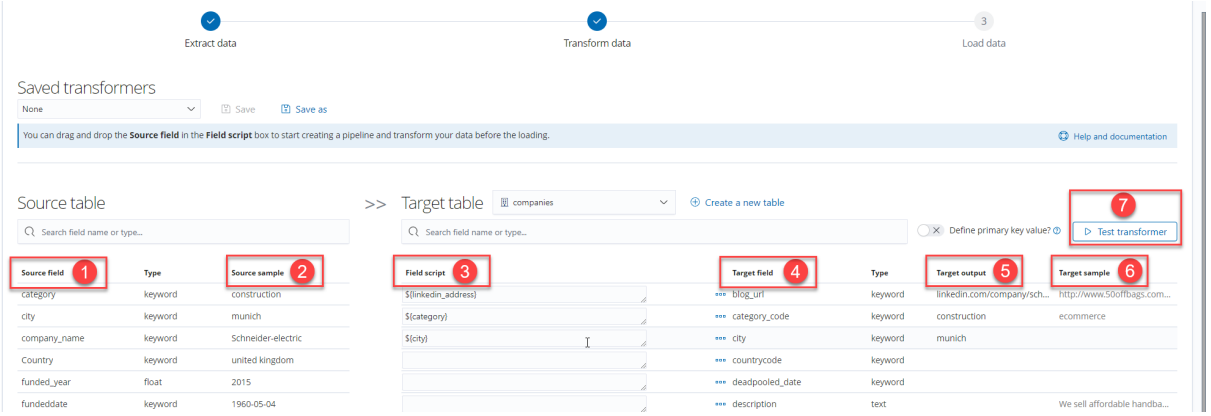
5. In the **Target table** dropdown menu ⑤, select the **companies** entity table.



Transforming your data

The **Transform data** screen allows you to modify data before it is ingested. In this example, we are going to import some columns and merge them into the existing **companies** entity table.

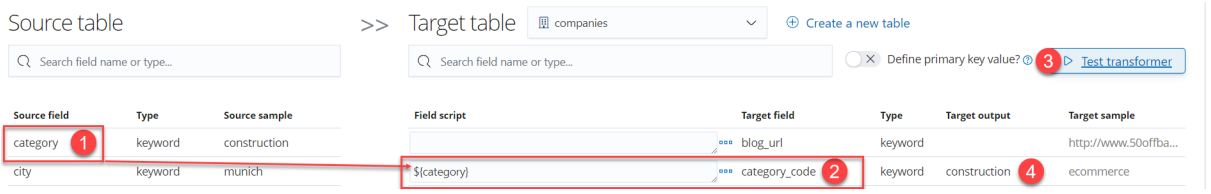
Before we proceed, let's have a closer look at the data transformation page after the last step.



Starting from the left side you can drag and drop the **Source field** ① into the **Field script** ③ textbox of the **Target field** ④.

Knowing the content of the **Source** ② and **Target samples** ⑥ makes it easy to merge the new CSV columns with the existing fields. More importantly, the transform can be tested ⑦ before the import and the result can be previewed in the **Target output** ⑤ column.

Let's start by dragging the **category** ① field into the **Field script** ② of the **category_code** Target field. Click **Test Transformer** ③**.**

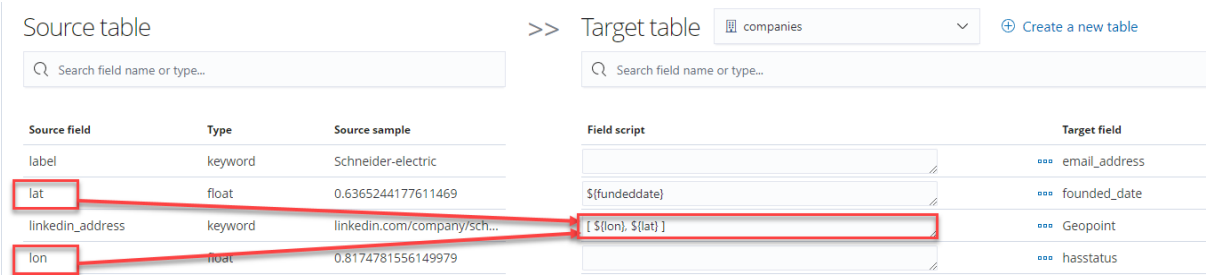


We are not applying any transform for the moment, since the original content of our *category* source field is fully compatible with our *category_code* target field. We can do the same for some of the other columns.

- Drag *linkedin_address* to *blog_url*;
- *city -> city*;

- *fundeddate -> founded_date;*
- and *website -> url* (these don't need any transformation).

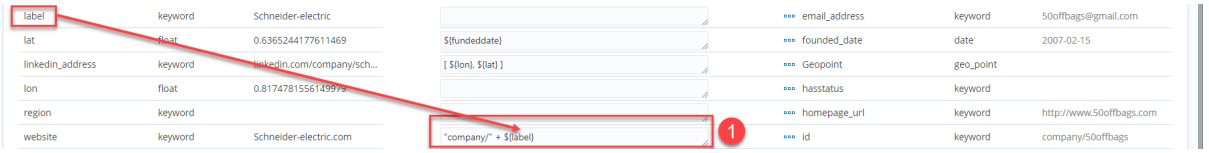
However, we have a Geopoint in our target **companies** entity table and two separate fields for latitude and longitude in our new CSV. The new transformer feature from Siren Investigate makes this operation extremely simple. Drag and drop the *lat* and *lon* fields into the **Field script** relative to the *geopoint* Target field. Make sure to follow Elasticsearch documentation in case you are not sure about how to format the field. In this case, we just need to add the square brackets and the comma plus the space in the middle: **[${lon}, ${lat}].**



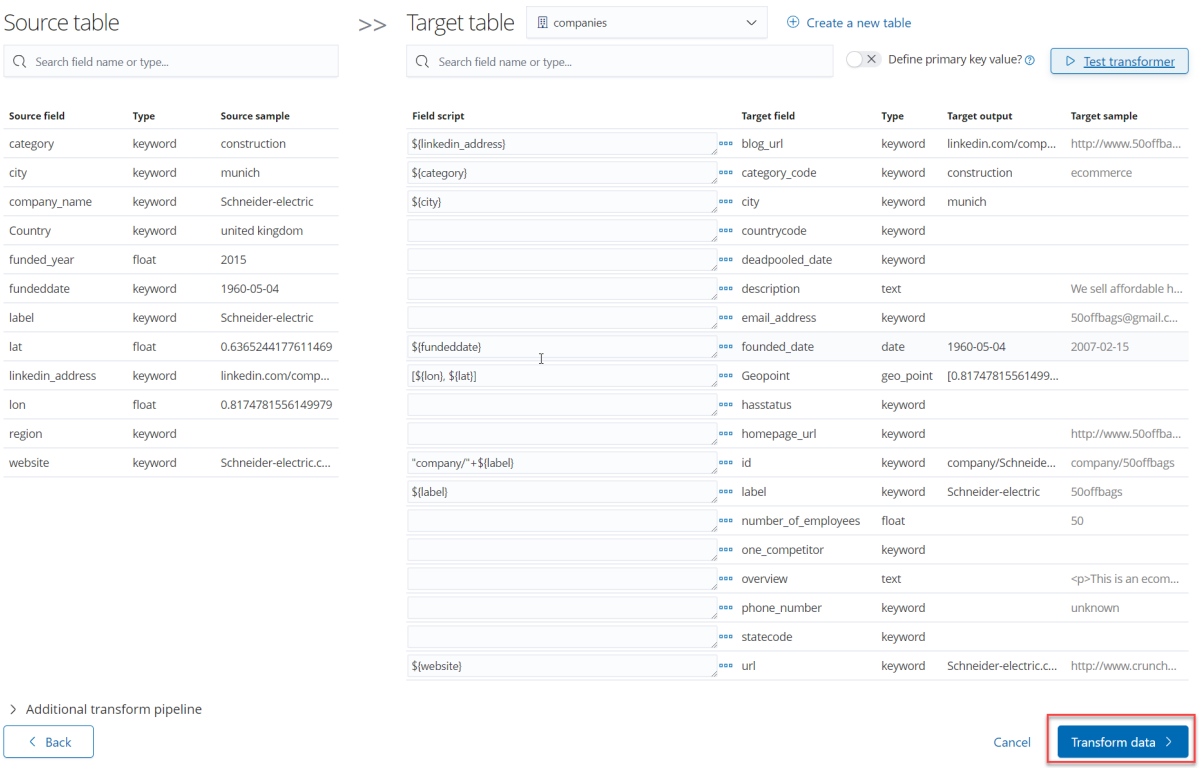
Remember that you can always test the expression entered in the Field script with the **Test transformer** function and check the result preview in the **Target output.**

Sometimes, you need to combine values to match the target data format. This can be done with the concatenation operator + in the **Field script**. For help with field scripts, see the documentation.

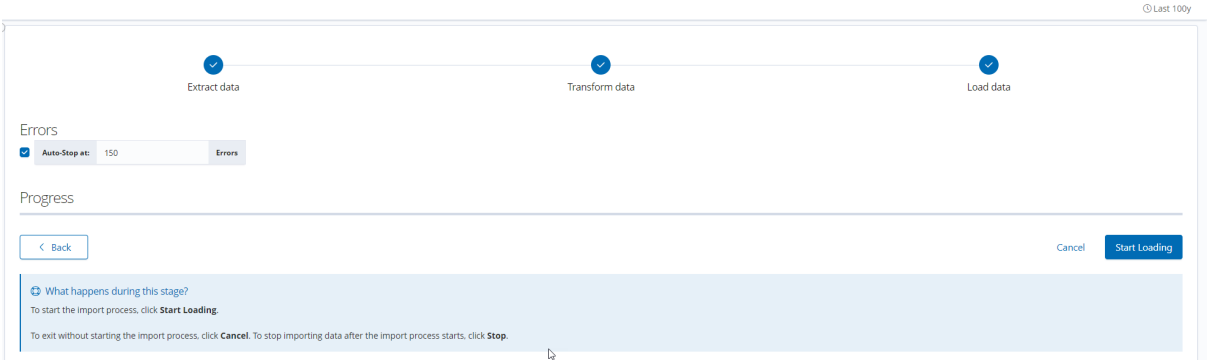
1. Drag the *label* **Source field** to the *id* field of the **Target table** ①.



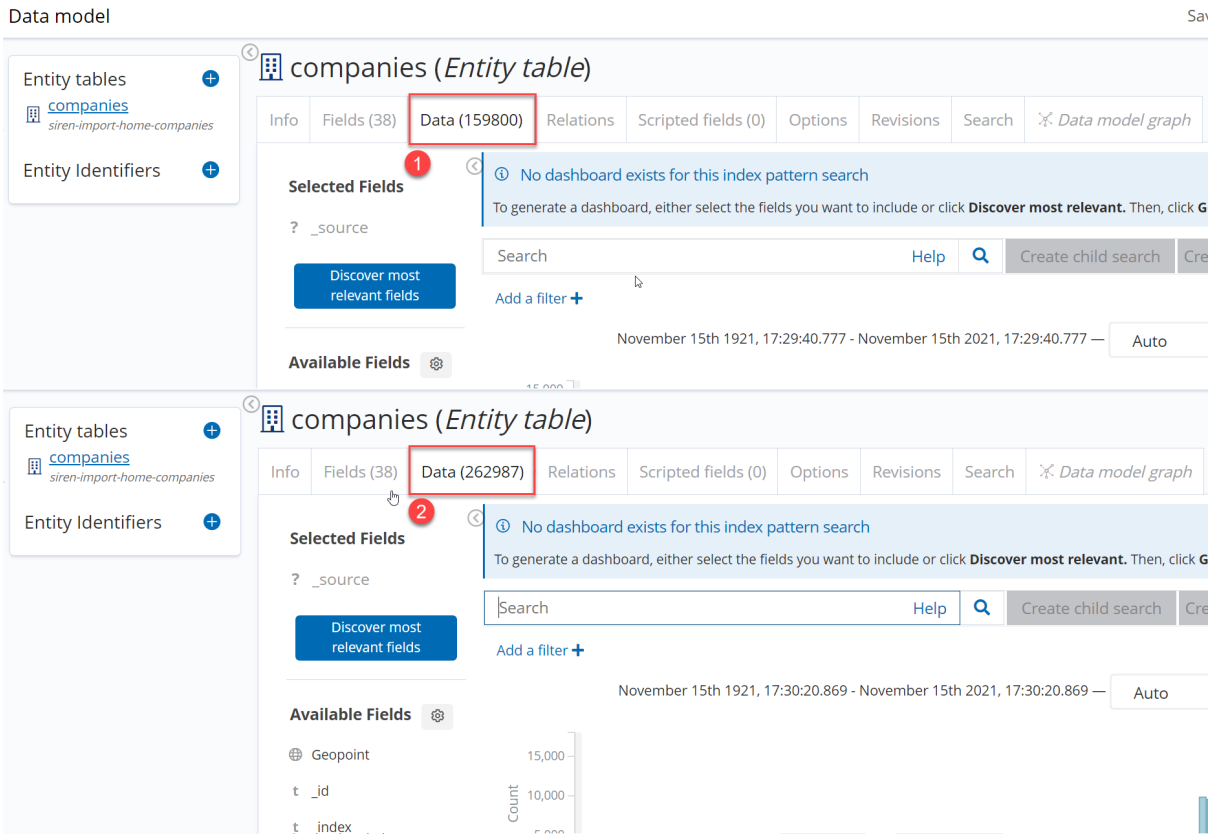
2. Add the string "company/" followed by the + operator sign before the ${label}. Test the transformers. **NOTE**: Ensure that your keyboard uses straight quotation marks. Curly quotation marks are not supported in this field and will result in an error.
3. Validate that the Target output is similar to the Target sample.
4. Before the last step, compare what you have done with the following screenshot to make sure you did not forget any fields and made all the necessary transformations.



5. Click **Transform data.** (Some of the fields of the CSV were left out to show that you can decide what to load and where.) For more information about transforming data, see the documentation.
6. Finally, click **Start loading** to import the new CSV. In a production environment, any errors will be listed here.



After the data is imported, you get a confirmation message. You can also check if the data was imported correctly in the **Data** tab.



① Number of records before the import

② Number of records after the import

Configuring the Graph Browser

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

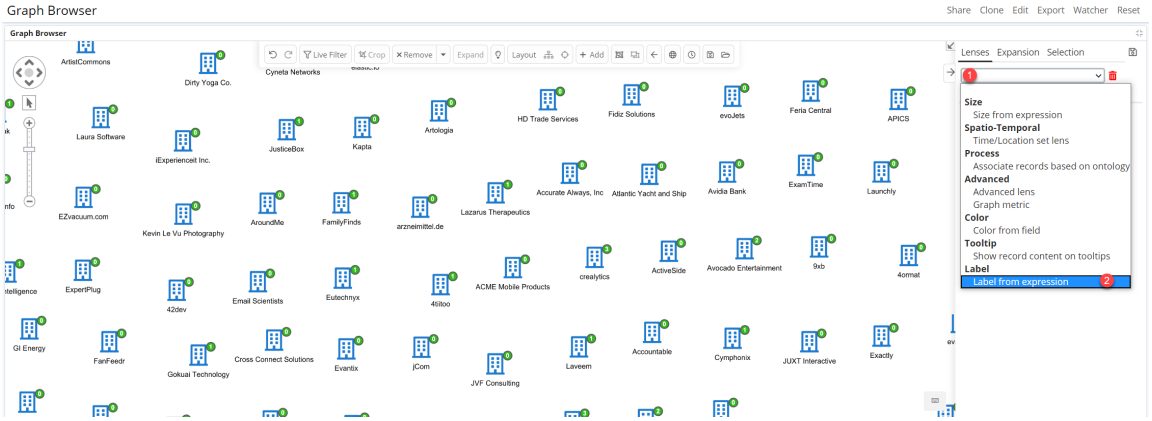
Fixing labels with graph lenses

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

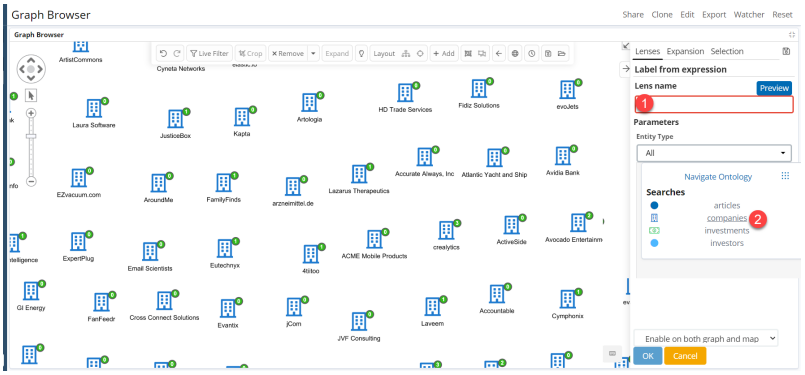
You should always set them in the data model first, but you can then override them in the graph with a lens that is useful for specific analysis. For example, you might first want to see the article's title, but later change your mind and want to see the article's author as a label instead.

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.

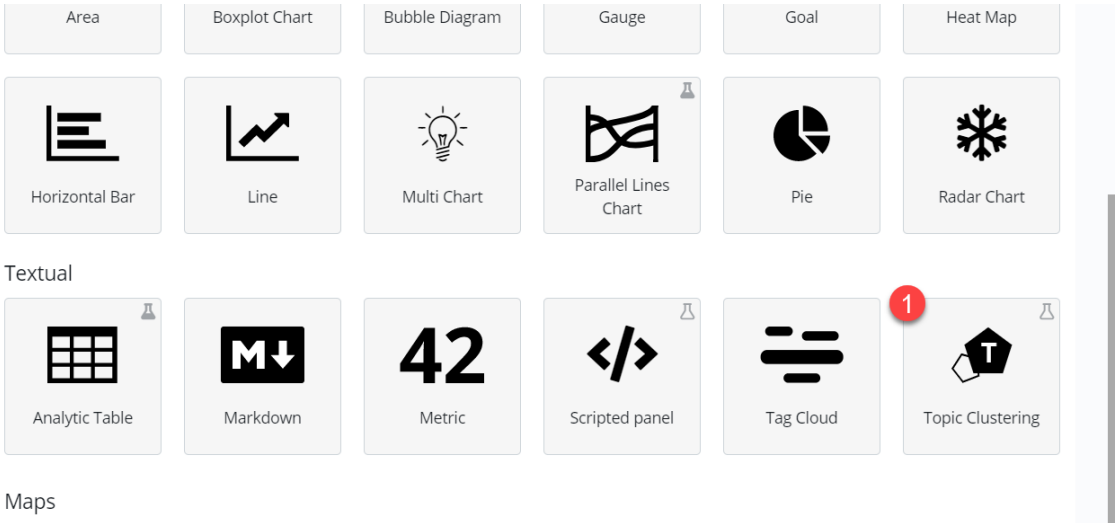


4. In the **Entity Type** drop-down menu, select **companies** from the list of searches ②.
5. To fill the **Expression** field, click the + button ① and select **label** from the dropdown menu. The **Expression** field is populated automatically with `payload["label"]`.
6. Click **Save lens**.

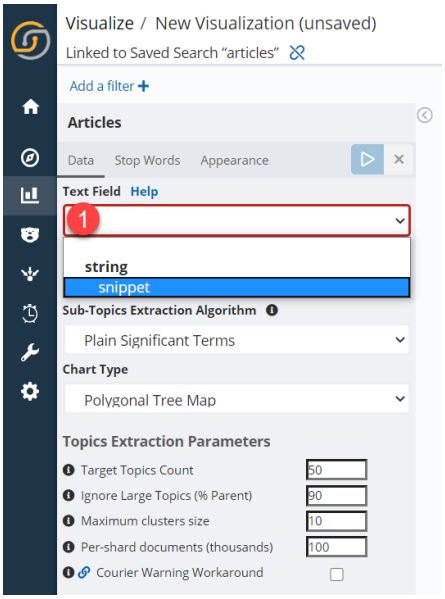
Creating a Topic Explorer visualization

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).

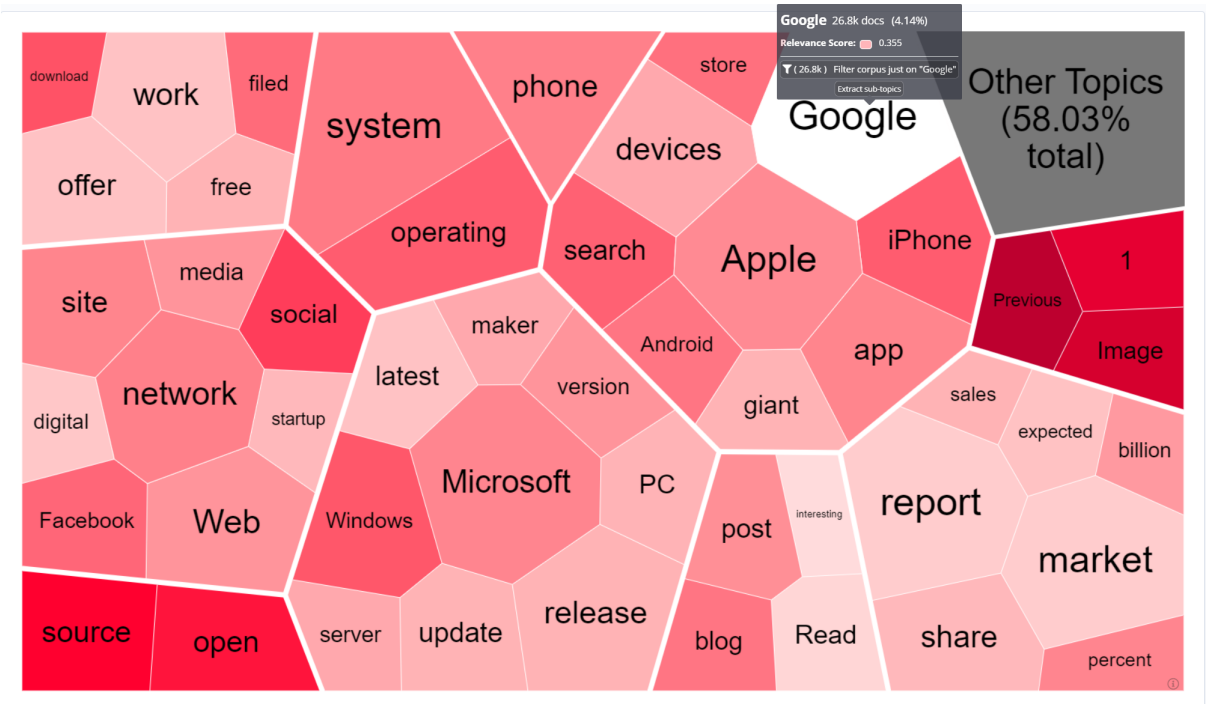
Select the **Topic Clustering** visualization ①.



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



What is the most discussed topic in our data?


download
work
filed
system
phone
store
Google
Other Topics (58.03% total)
offer
free
operating
devices
search
Apple
iPhone
site
media
social
maker
Android
app
Previous
1
Image
network
latest
version
giant
sales
digital
startup
expected
billion
Facebook
Web
Windows
Microsoft
PC
post
interesting
report
market
source
open
server
update
release
blog
Read
share
percent
Google 26.8k docs (4.14%)
Relevance Score: 0.355
(26.8k) Filter corpus just on "Google"
Extract sub-topics

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[1].

Here, we will use it to enrich another CSV file in our demo dataset, the **articles_nlp.csv** file, which contains 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**.

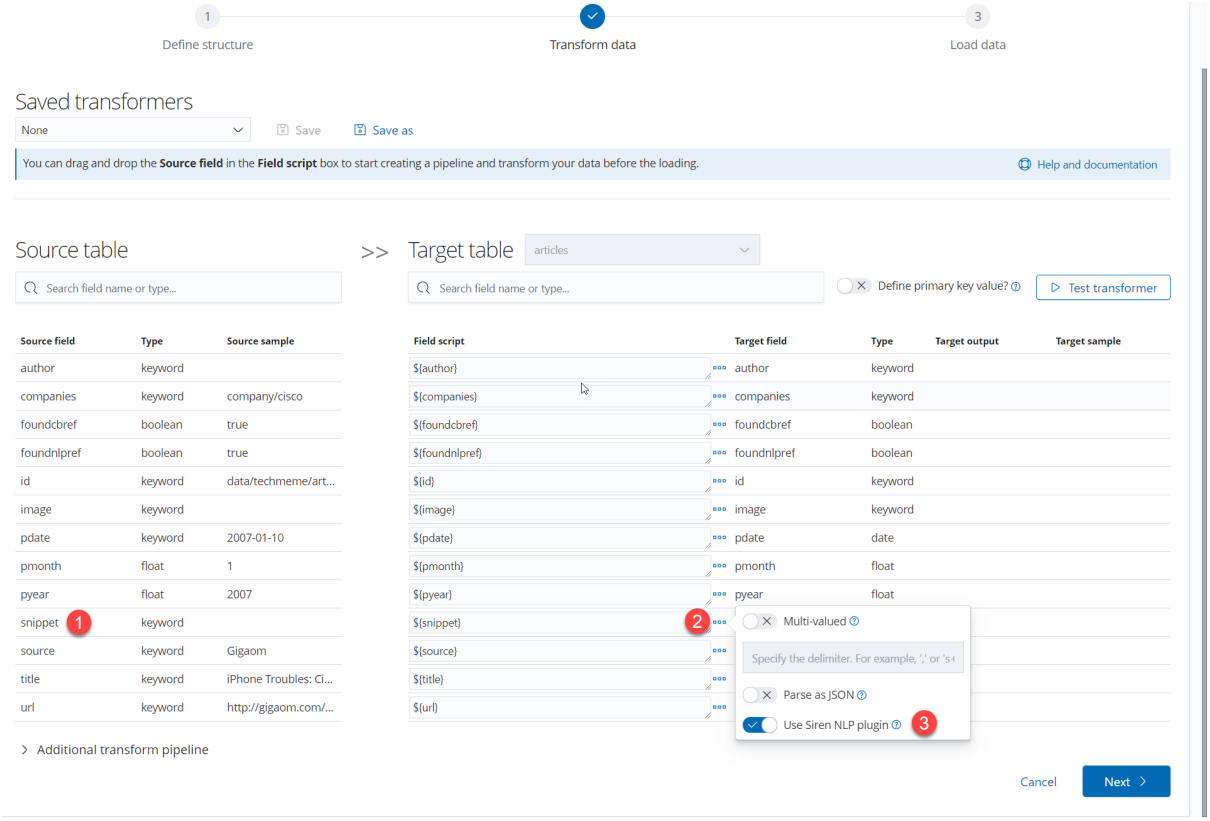
Procedure

1. Start importing the **articles_nlp.csv** file that is provided with the sample data. (Check the Importing data section if you need to refresh your memory.)
2. In the **Table name** field, type *articles*.
3. Scroll down to the mapping definitions and set the **pdate** field as **Date** in the type column.
4. For the **snippet** field, select **Text (allow word cloud)** as the type.

   NOTE: This step is mandatory to allow the NLP processor to work.

[1] 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.

5. Click **Next.**
6. On the **Transform data** screen, find the snippet field (1), click the ellipsis button (2), and activate the **Use Siren NLP plugin** option (3).

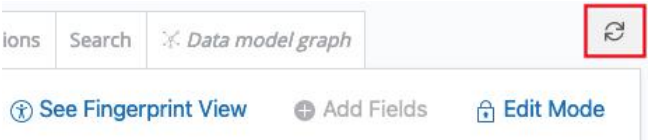


The **nlp-pipeline** is now configured to perform the following tasks:

- Extract annotations by using various processors from the **snippet** field of the **articles** index.
- Write the annotations to the field **siren.nlp**.
- Extract the following items using regex or simple rules: Telephone, USTelephone, Email, IPv4, IPv6, MacAddress, Url, SortCode, HashTag.
- Extract Organization, Location and Person Named Entities by using OpenNLP NER models.

7. Click **Next** and **Start loading**.

Once the import process is complete, go to the **Fields** tab and click **Refresh field list**.



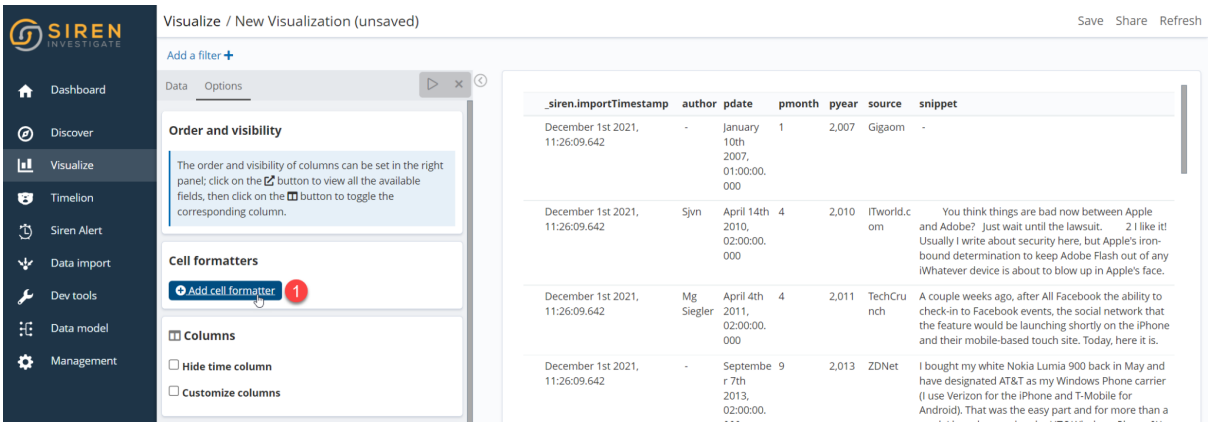
This step allows Siren to discover the newly-found entities.

Next, auto generate a dashboard and name it *Articles*.

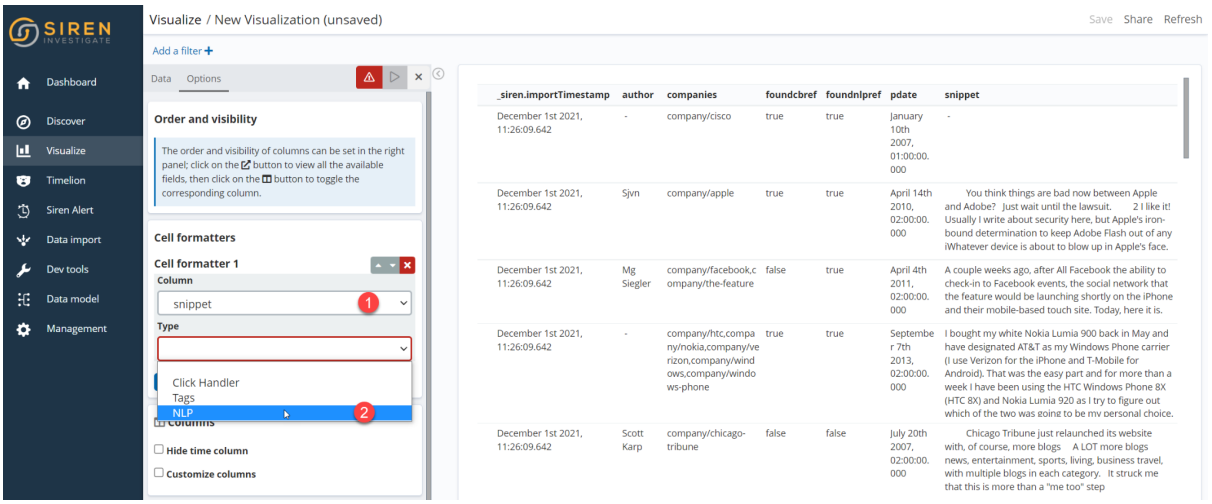
In the next section, we are going to create a special Record Table for tagging entities that were discovered by the NLP process in the processed text field (snippet).

Creating an NLP-based Record Table

1. In the new Articles dashboard, click **Edit** in the Options menu on the top-right of the screen.
2. Click **Add** and **Add a new visualization**.
3. In the Visualizations menu, select **Record Table.**
4. On the **Select an entity table or search** screen, select the **articles** entity table.
5. Click **Add cell formatter** ①.

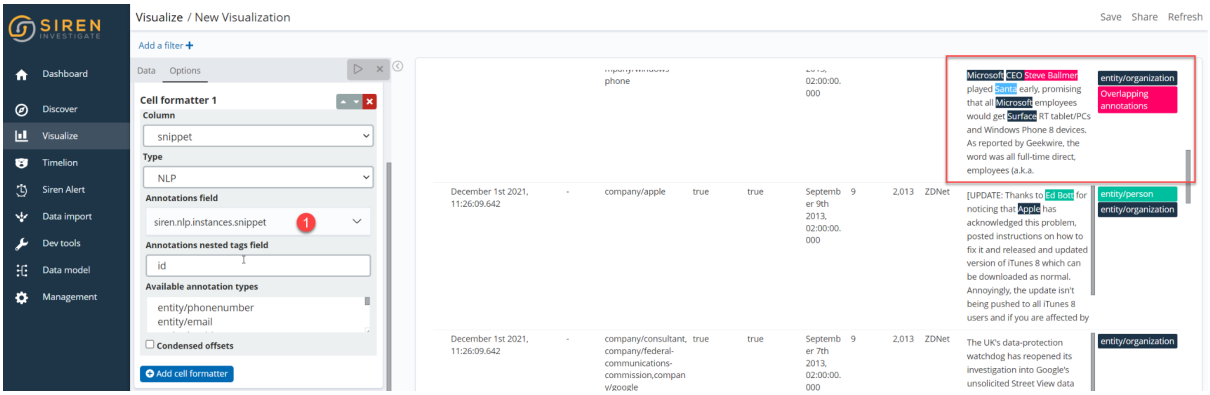


6. From the **Column** dropdown ① select snippet and for **Type** select NLP ②.



7. In the **Annotations field**, select **siren.nlp.instances.snippet**, which contains the annotated offsets needed for formatting the text.

8. Click **Apply changes** (the play button).



9. Click **Save**, name the visualization (for example, “Record Table NLP”), and save it to the Articles dashboard.
10. Save the dashboard to quit Edit mode.

Connecting extracted entities in the data model

Now, let’s define the relationship between the **articles** index and the entities and the rest of the indexes in our data model. To create relations manually, complete the following steps:

1. Go to the **Data model** app.
2. Go to the **Relations** tab and click **Add relation**.
3. Under Source Entity **articles**, select **companies** as the Field.
4. For the Target Entity, select **companies** and select **id** as the Field.
5. Add the labels, **mentions** and **mentioned in**, as shown in the screenshot below.
6. Click **Save**.

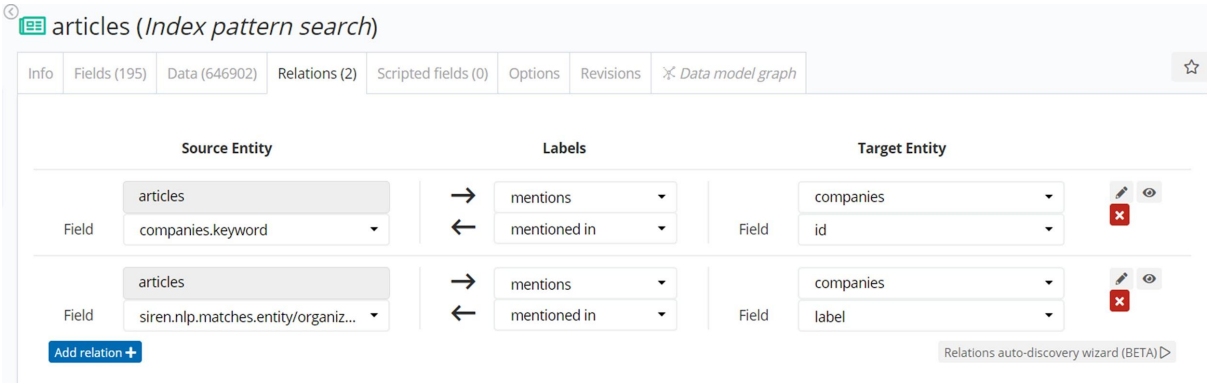
Example usage of NLP Data in Siren Investigate

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.

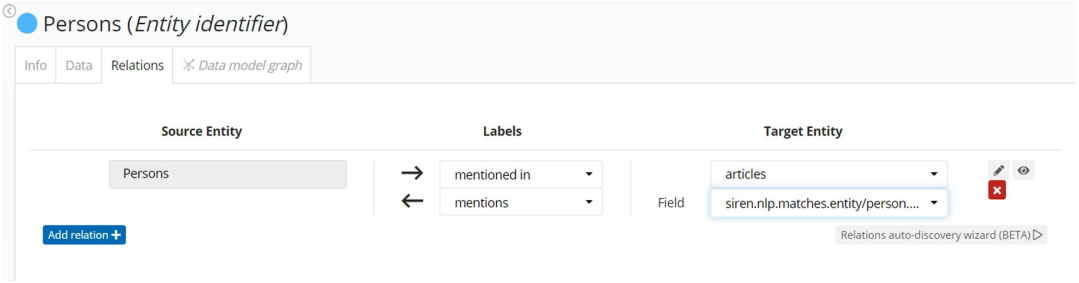
In the **Data model** app, click on the **Data** tab. In the list of **Available fields**, you can see that a new set of fields starting with **siren.nlp** have been created by the NLP plugin during the import. These fields contain annotations for fields processed by Siren NLP (in this case the **snippet** field), such as organization, person and location.

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

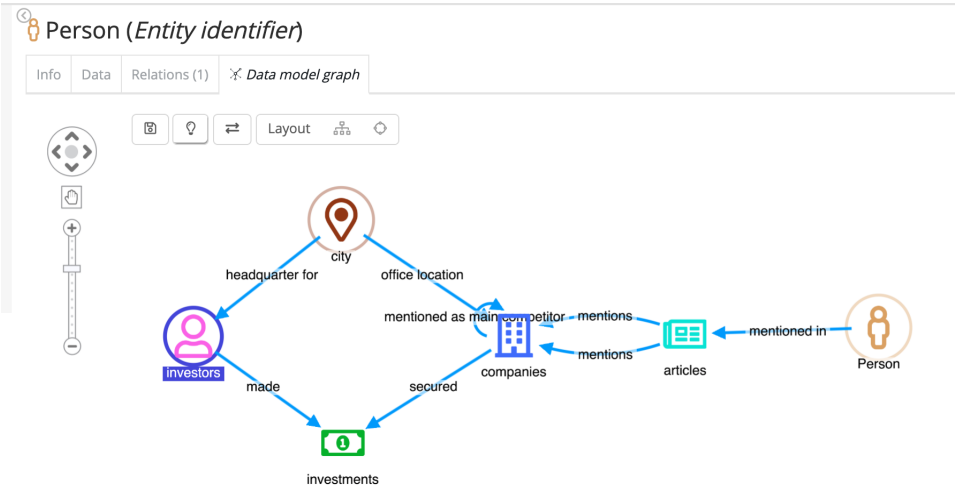
1. On the **Relations** tab, create a relation between Articles **siren.nlp.matches.entity/organization.keyword** and Companies **label.**



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.keyword**.

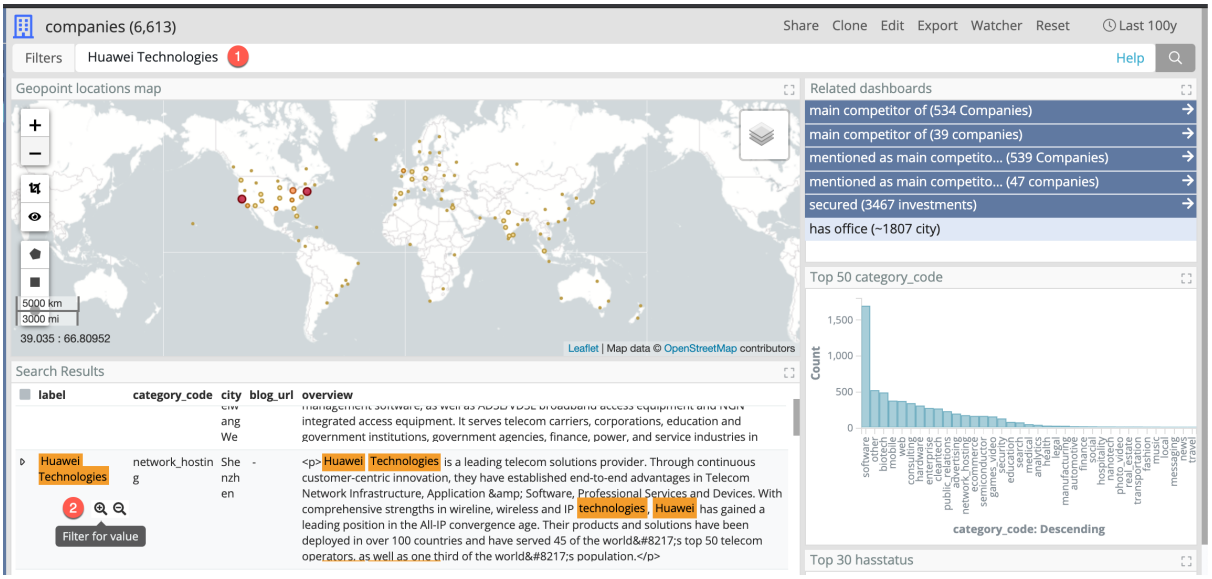


The data model graph should look like this:



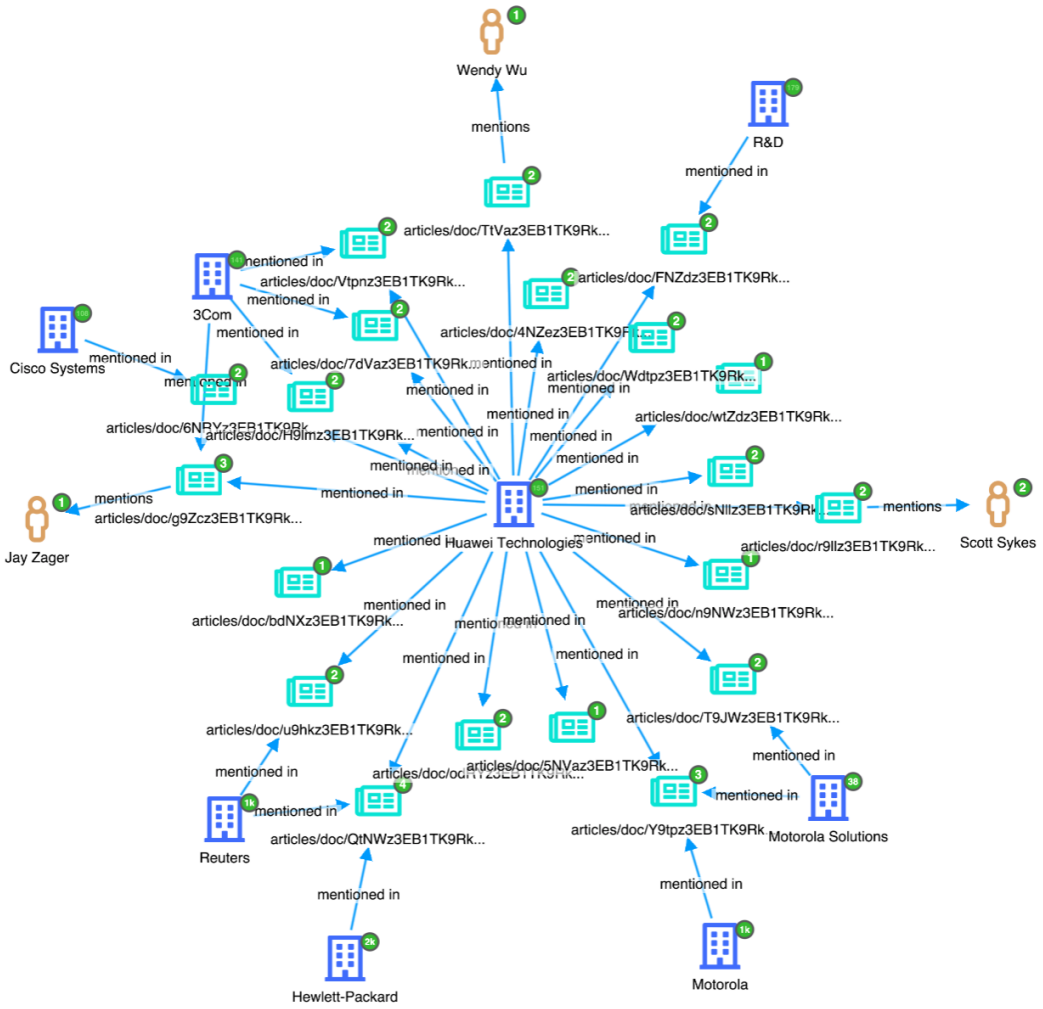
3. Go to the **Companies** dashboard and search for the company "Huawei Technologies" ①.
4. Filter the dashboard for the company with the label name "Huawei Technologies" by using the **Filter for value** button ②. This will filter the dashboard with a single record

for this company.



5. Go to the **Graph Browser** and drag and drop the **Companies** dashboard there. This will add a node with the company Huawei Technologies.
6. Select the node, right-click, and select **Expand by relation**.
7. Select **mentioned in (20) *articles*** and click **Ok**.
8. 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

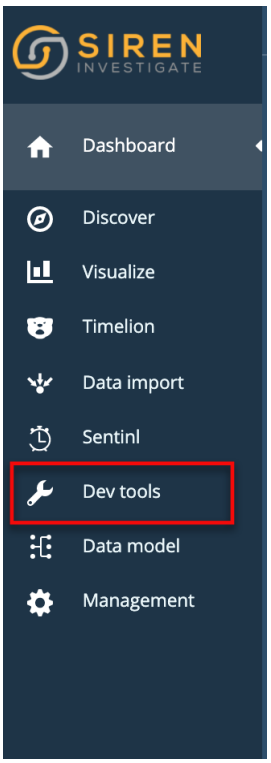
Technologies.



Deleting data 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.

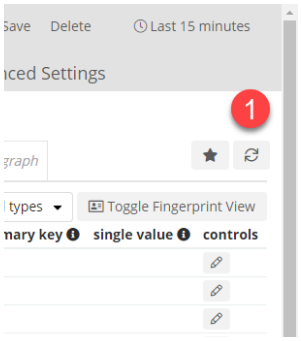
However, you can delete an index by using the **Dev Tools console**.



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.

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.



Legal notices