WHAT CAN THE VISTORIAN DO: A FEATURE OVERVIEW

The Vistorian is a tool for interactive exploration of MULTIVARIATE, TEMPORAL & GEOGRAPHIC networks.

MULTIVARIATE

... means "multiple variables", meaning that both entities and their relationships can have additional attributes.

UNIVARIATE

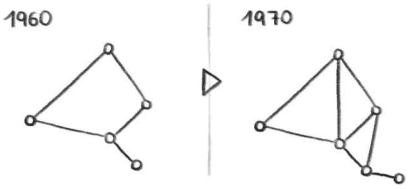
"Bob knows Mia"

MULTIVARIATE

"Bob the merchant gave Mia the lawyer 5000 pounds as a loan."

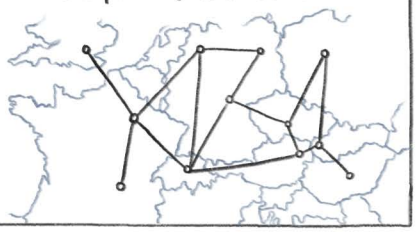
TEMPORAL

... means that the network changes over time.

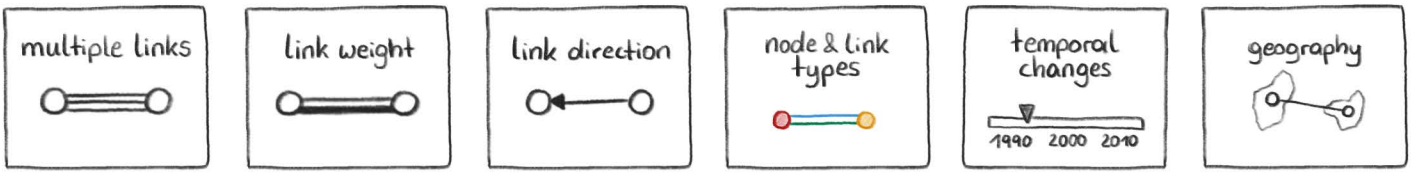


GEOGRAPHIC

... means that geographic locations are part of the data.

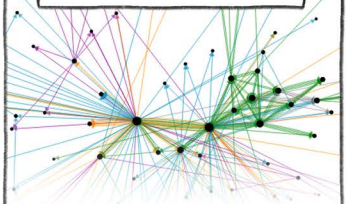


The Vistorian can visualize various properties of networks, such as:



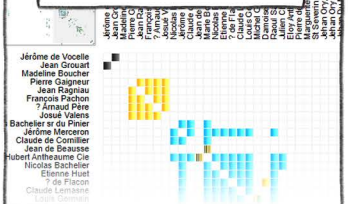
The Vistorian offers 4 different interactive visualizations:

NODE-LINK DIAGRAM



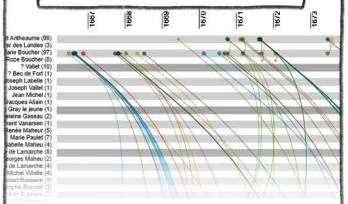
... the "classic" network visualization

ADJACENCY MATRIX



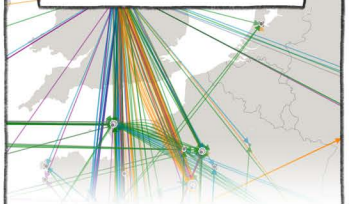
... for dense & complex networks

TIMELINE



... for direct overviews of time-dependent networks

MAP

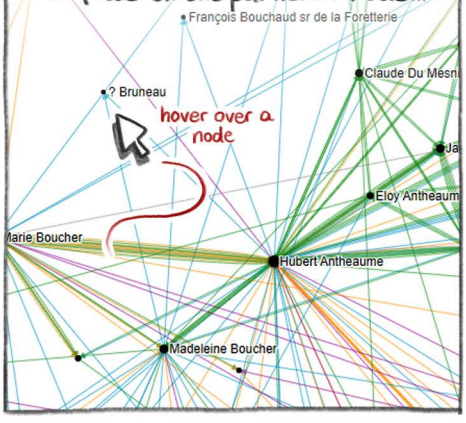


... for networks with geographic data

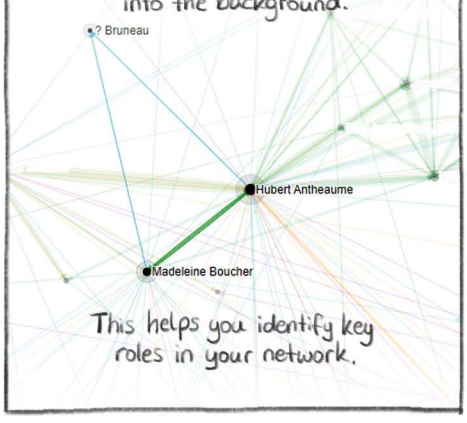
The Vistorian offers a range of interactive features to support the exploration of networks:

INTERACTIVE HIGHLIGHTING

Interactive Highlighting allows you to focus on one particular node...



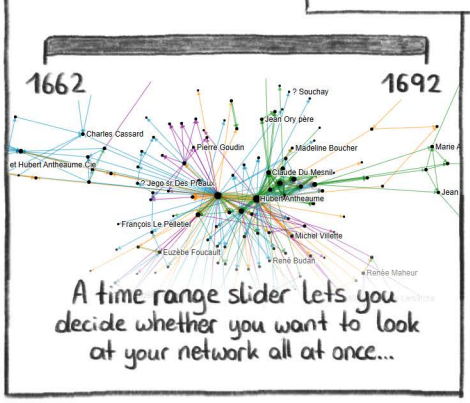
... while letting the rest fade into the background.



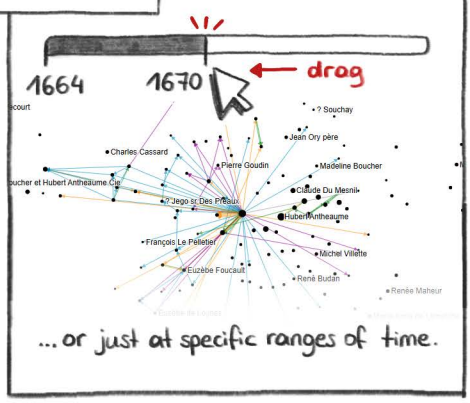
This helps you identify key roles in your network.

WHAT CAN THE VISTORIAN DO: A FEATURE OVERVIEW

TIME RANGE SLIDER

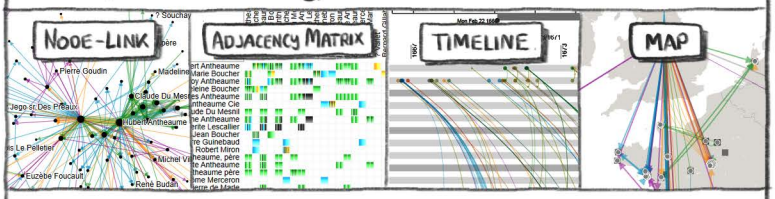


A time range slider lets you decide whether you want to look at your network all at once...



... or just at specific ranges of time.

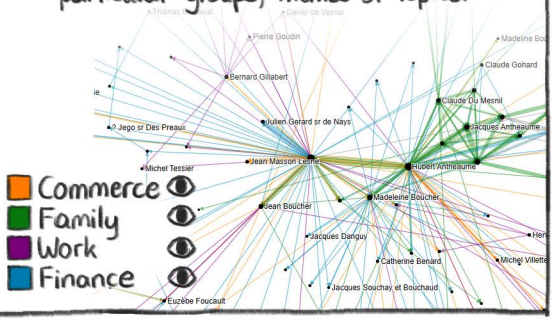
Of course, the time range slider is available for all four visualization types the Vistorian has to offer -

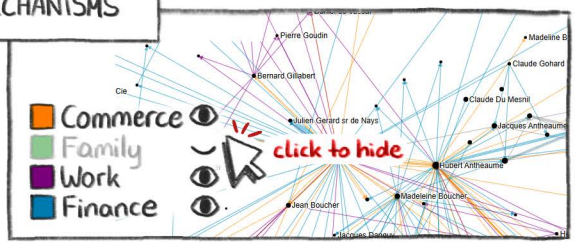


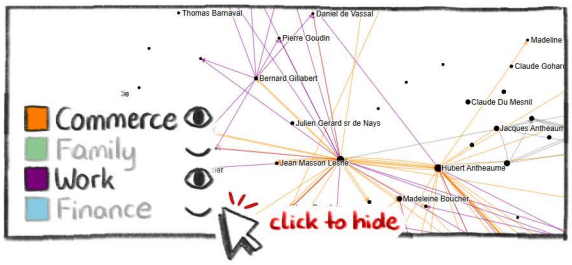
- if the loaded data includes time, that is.

FILTERING MECHANISMS

Filtering lets you hide and display nodes and links of different types, to have a closer look at particular groups, themes or topics.

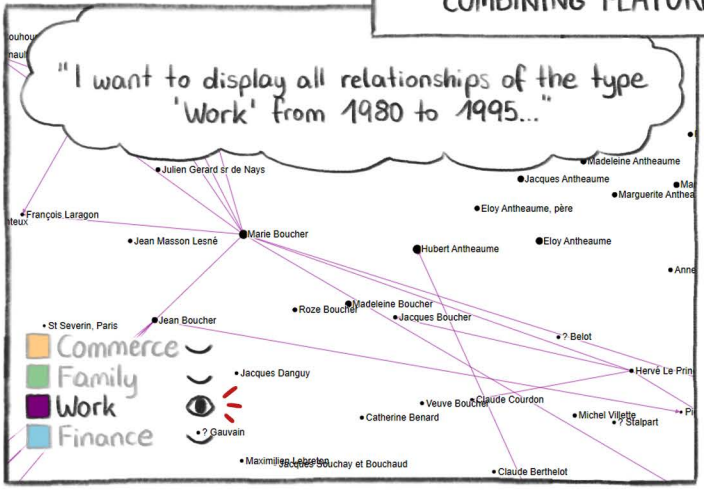


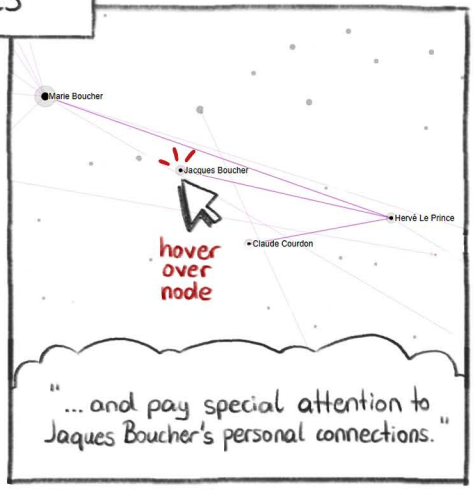




You can use all of the interactive features combined, too!

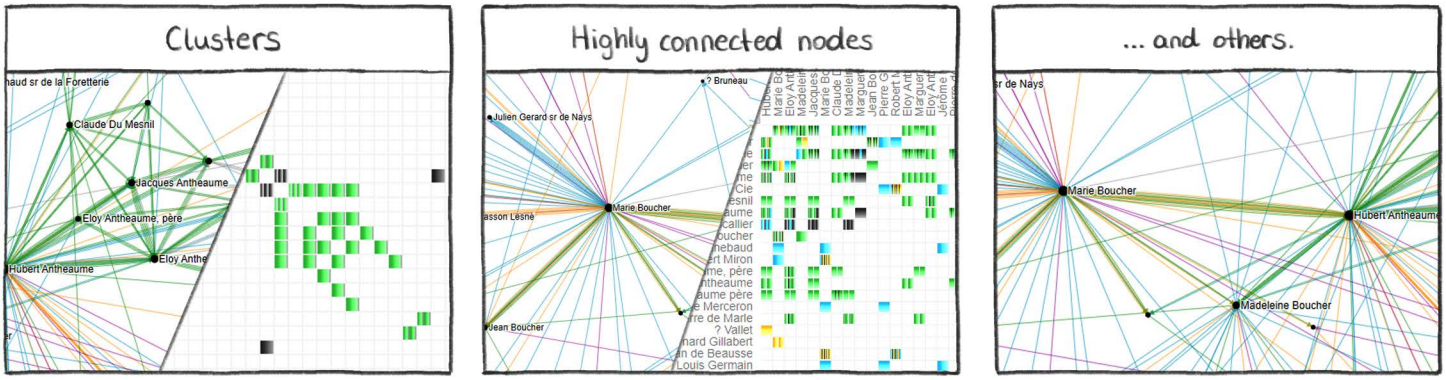
COMBINING FEATURES





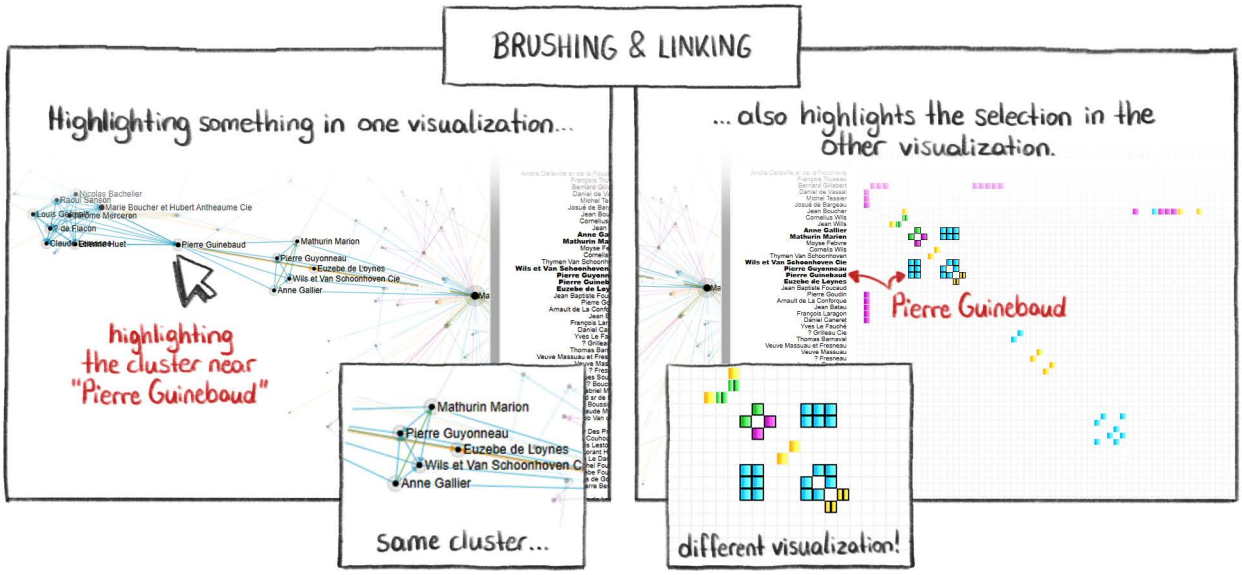
WHAT CAN THE VISTORIAN DO: A FEATURE OVERVIEW

All visualizations in the Vistorian are algorithmically optimized to reveal network structures such as...



The visualizations are also fully linked and can be explored side by side. This allows you to make the most out of each visualization's strengths.

This technique is also called *brushing & linking*.



The Vistorian is a standalone, open-source application.

Your personal network data is kept safely in YOUR browser's local storage until you delete it.

No account is required and no personal data is transmitted to our servers.

The Vistorian is developed by:

Visual+ Interactive Data · Inria INVENTEURS DU MONDE NUMÉRIQUE · Microsoft Research · THE UNIVERSITY of EDINBURGH

Read more about it at vistorian.github.io

These comics were created in a collaboration with St. Pölten University of Applied Sciences, funded by the GFF NÖ as part of the dissertation project VisToon (SC20-014)

fh st.pölten

NODE TABLES AND LINK TABLES

When working with network data, you might encounter different types of tables:

LINK TABLES

Sender	Receiver	Amount (k)	Year	Type
Bob	Charles	10	1801	Loan
Bob	**Charles**	**14**	**1803**	**Gift**
Bob	Charles	3	1810	Purchase
Bob	Anton	2	1801	Purchase
Anton	Bob	5	1810	Loan
Anton	Lily	4	1804	Loan
Charles	Anton	2	1804	Purchase

In link tables, each row describes one relationship between two entities (i.e., nodes). For example, the marked row here shows one money transfer from Bob to Charles, where the amount was 14k, and it was a gift.

NODE TABLES

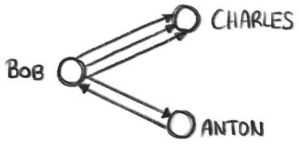
Name	Profession	Pronouns	Street	City
Bob	Merchant	he/him	16 Jedburgh R	Lewes
Anton	Lawyer	he/him	98 City Walls R	Clunie
Charles	**Accountant**	**he/him**	**30 Rhosddu Rc**	**Fidden**
Fred	Attorney	he/him	51 Cloch Rd	St Harmon
Lily	Accountant	she/her	81 Peachfield F	Challacombe
Felix	Flight Attendant	he/him	78 Seaford Roa	Cumwhinton
Julian	Police Officer	he/him	46 Marlboroug	Southampton
Alex	Teacher	they/them	39 Foregate St	Codmore Hill

In node tables, each row contains information about one single entity (i.e., node). For example, the marked row in this table shows that Charles is an accountant, goes by he/him, and lives at 30 Rhosddu Road in Fidden.

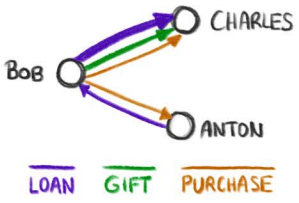
The main difference between the two kinds of tables is the way they are used to specify networks:

Source Node	Target Node	Amount	Link Attributes	Type
Bob	Charles	10	1801	Loan
Bob	Charles	14	1803	Gift
Bob	Charles	3	1810	Purchase
Bob	Anton	2	1801	Purchase
Anton	Bob	5	1810	Loan

Link tables can be used to directly create a network from them, since all information about the links is there. Visualized, this part of the table could look like this:



(or this, if you include the link attributes)

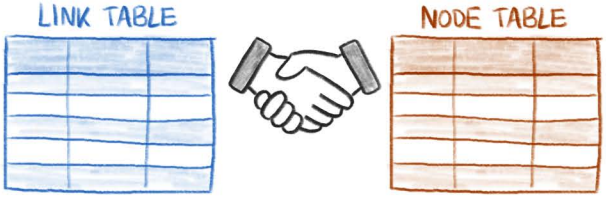


If we look at the node table, however, we see no direct network structure.

We could at best create a visualization that shows how many people have which jobs...

But that is not really purposeful, especially in this case.

Instead, node tables are used TOGETHER with link tables.



This way, they complement each other, because link tables cannot contain information about the entities themselves * (like node types)

* well, technically...

NODE TABLES AND LINK TABLES

Don't do this!

... technically, they could - but recording data this way is extremely tedious, redundant, and error-prone. Imagine you had to record the profession of each individual over and over in each row:

Sender	Sender Profession	Receiver	Receiver Profession	Amount (k)	Year	Type
Bob	Merchant	Charles	Accountant	10	1801	Loan
Bob	Merchant	Charles	Accountant	14	1803	Gift
Bob	Merchant	Charles	Acconutant	3	1810	Purchase
Bob	Merchant	Anton	Lawyer	2	1801	Purchase
Anton	Lawyer	Bob	Merchant	5	1810	Loan

one typo can mess up your analysis, since here you now have two Charles - one Accountant and one Acconutant!

So this is why instead of having such a redundant table, we go for...

THE BETTER SOLUTION

...where we outsource the information about the nodes to an extra table.

LINK TABLE

Sender	Receiver	Amount (k)	Year	Type
Bob	Charles	10	1801	Loan
Bob	Charles	14	1803	Gift
Bob	Charles	3	1810	Purchase
Bob	Anton	2	1801	Purchase
Anton	Bob	5	1810	Loan
Anton	Lily	4	1804	Loan
Charles	Anton	2	1804	Purchase
Charles	Lily	6	1811	Purchase

Information about the entities' relationships

+

NODE TABLE

Name	Profession	Pronouns	Street	City
Bob	Merchant	he/him	16 Jedburgh R	Lewes
Anton	Lawyer	he/him	98 City Walls R	Clunie
Charles	Accountant	he/him	30 Rhosddu Rc	Fidden
Fred	Attorney	he/him	51 Cloch Rd	St Harmon
Lily	Accountant	she/her	81 Peachfield F	Challacombe
Felix	Flight Attendant	he/him	78 Seaford Roa	Cumwhinton
Julian	Police Officer	he/him	46 Marlboroug	Southampton
Alex	Teacher	they/them	39 Foregate St	Codmore Hill
Oliver	Pilot	he/him	10 Ockham Ro	East Preston

Information about the entities themselves

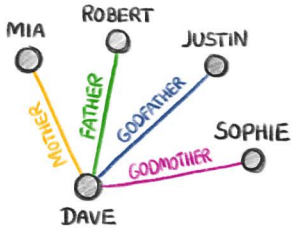
Just make sure that the node names match in both tables, so Visualization tools can interpret them correctly.

Bob = Bob
Charles = Charles

There are special cases of tables that are neither really link nor node tables, but can still be used alone to form a network.

An example would be this genealogy table:

CHILD	MOTHER	FATHER	GOD-FATHER	GOD-MOTHER
Bob	Celine	Charles	Dave	Eve
Ana	Fannie	Gerd	Mike	Dianne
Celine	Maria	João	Pedro	Ana
Dave	Mia	Robert	Justin	Sophie
Mike	Clara	John	Georg	A[illegible]
Pedro	Raquel	Mateo	Ju	
Eve	Nora	Felix	Jo	
Laura	Tamara	Julian	D	
Troy	Natalia	Alex	W	
Victoria	Celine	Oliver	Alfred	Crystal



The Vistorian, for example, would interpret this table as a node table and infer a different link type for each column.

HOW TO PREPARE DATA FOR EXPLORATION

As soon as we have our raw data, there are two routes of exploration we can take, depending on whether we have a

RESEARCH QUESTION

yes

HYPOTHESIS-DRIVEN RESEARCH

...where you have an idea of what you want to know, for example...

> Which members of which political parties interrupted which speakers of other parties during state parliament sittings?

no

OPEN-ENDED EXPLORATION

... where you go the experimental path and find interesting insights along the way.

> I collected all these data about my topic - now I wonder if I can find any patterns in these networks?

In either case, our starting point will be our dataset, likely formatted in a large

TABLE

Get the dataset: tinyurl.com/viscomicdataset

Name_eng	Name_zh	Nationality	University	State	Uni_city	...	lev	Field_main	Field_2	Field_3			
Arnold_Harrison	安樂生	Western	California	California	Berkeley	A.B.	Bachelor	Humanities	Arts	Humanities	1931	1931	NA

Let's look at this example dataset about students and where they worked after graduation.

Each row contains details about one student. For example, this one is

Harrison Arnold, who is Western and studied at the University of California.

But how do we find a network structure in this?

To do that, we should first create a

CONCEPT MAP

Name_eng	Name_zh	Nationality	University	State	...	eld_main	Field_2	Field_3	Year_start	Year_end	Ho

A concept map shows which entities exist in our network and how they are related. Let's see what columns we have and which concepts we can find.

"Student" would be an obvious entity.

Note that there is no actual column named "student" though. Instead, the concept "student" is described by multiple columns in our table.

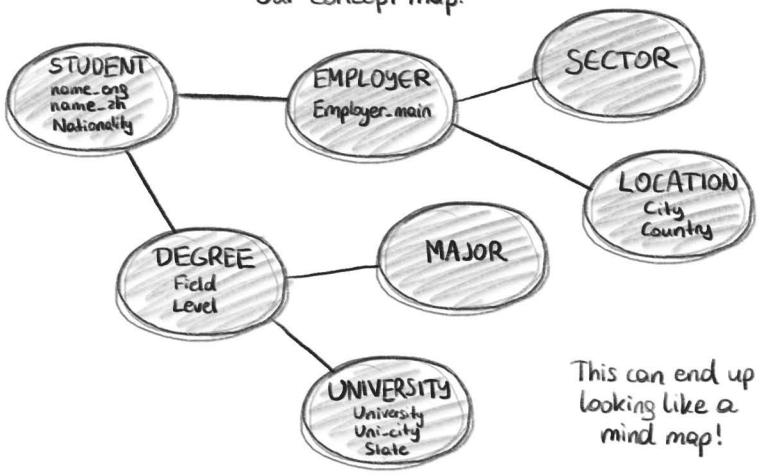
This is why such data are called

DESCRIPTIVE DATA

Like "Name" and "Nationality" describe "Student", "University", "City", and "State" could describe "University", or "Field" and "Level" could describe "Degree".

HOW TO PREPARE DATA FOR EXPLORATION

After we have found some entities, we can sketch our concept map:



This can end up looking like a mind map!

It's important to know that there is not always a single "correct" solution:

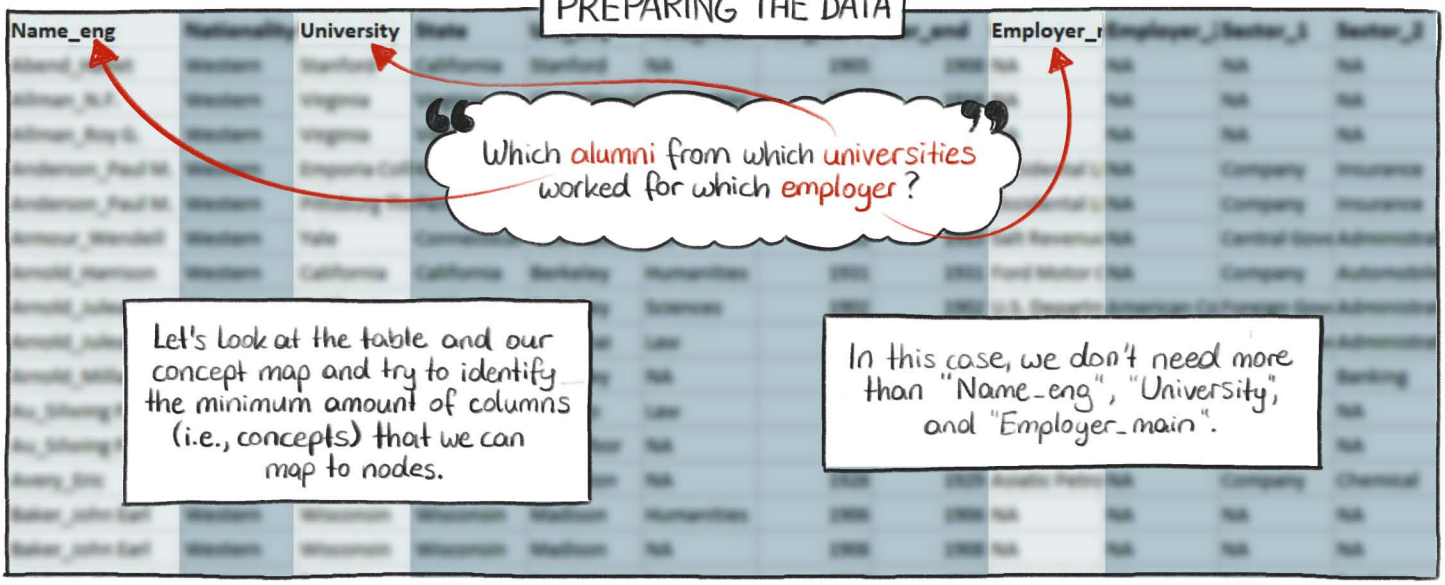
"Degree" could also be treated as descriptive data for "Student", or "Location" and "Sector" could be descriptive data for "Employer".

Which entities and descriptive data we concentrate on depends on our current research question or focus.

We may also ignore parts of our table!

So, let's take an example research question and plan how our network could look like to answer it.

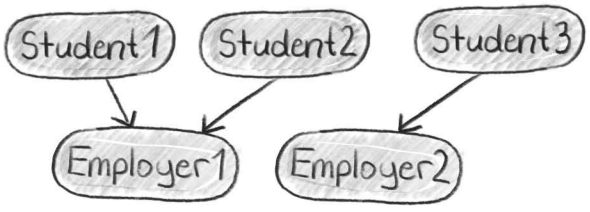
PREPARING THE DATA



Let's look at the table and our concept map and try to identify the minimum amount of columns (i.e., concepts) that we can map to nodes.

In this case, we don't need more than "Name_eng", "University", and "Employer_main".

Let's try to sketch how this network could look like.

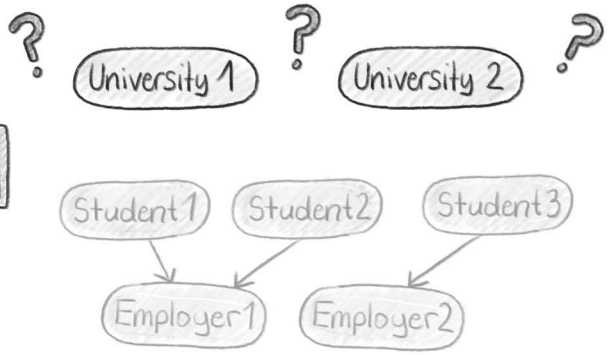


We have directed links, with each student having one employer as a "target".

Mapping students and employers is easy...

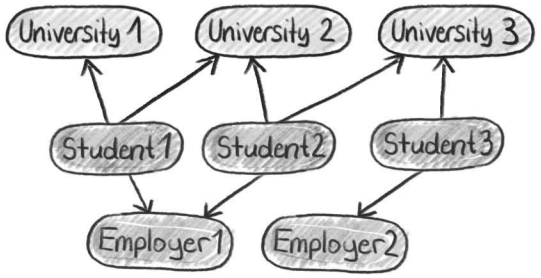
BUT...

How do we map the UNIVERSITIES?

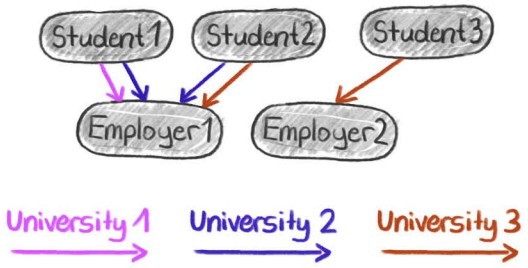


HOW TO PREPARE DATA FOR EXPLORATION

If we add them as nodes, the network will become cluttered and hard to read...



So instead, we can map them through the LINK TYPE.

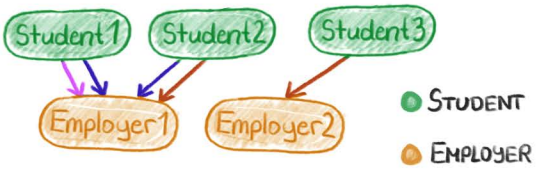


Both of these graphs show the SAME INFORMATION!

Technically, you could have your data visualized now. There is just one optional step left to do:

OPTIONAL

If you want, you could add node types for better visual distinction:



To do this, you need to provide an extra table that specifies each entity's type.

OPTIONAL

Such a table should look like this:

Name or ID of node ↓	Node type ↓
Arnold_Julean	Student
Arnold_Millard	Student
Au_Silwing P.C.	Student
Avery_Eric	Student
Baker_John Earl	Student
Baldwin_Berry O.	Student
Barnett_E.E.	Student

And if you have all your tables handy, you are now ready to import your data into a visualization tool!

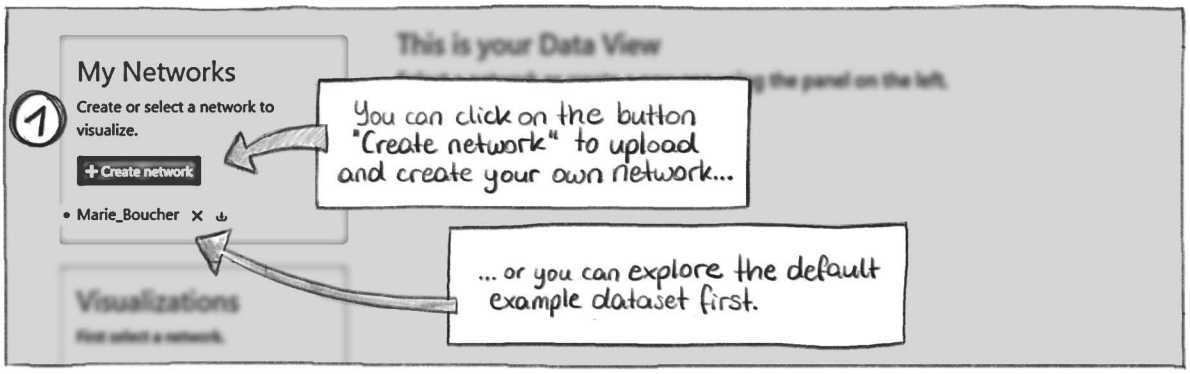
UPLOADING DATA & CREATING A NETWORK SCHEMA WITH THE VISTORIAN

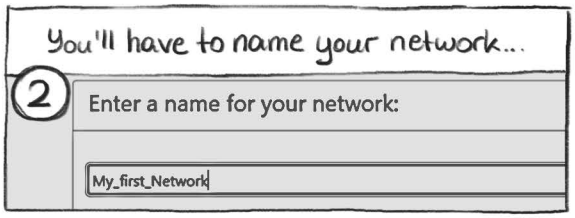
This comic uses the following datasets:

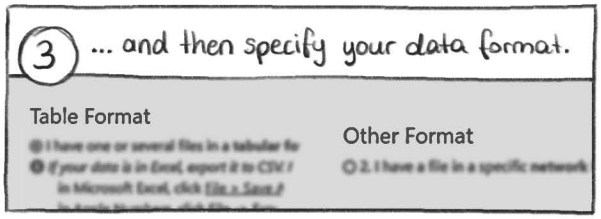
tinyurl.com/ viscomicdata2 → (example genealogy node table)
viscomicdataset → (dataset with alumni & employers)
studentnodetypes → (node types for the alumni dataset)

But you can also follow along with your own data!
See how to prepare your data here: → LINK

When starting the vistorian, you will see the following:





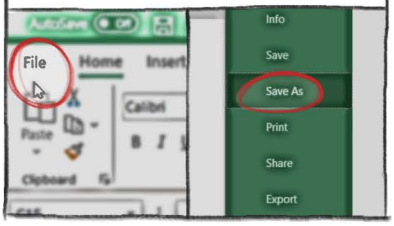


In most cases, you will have your data in a table that needs to be exported as a .csv file.

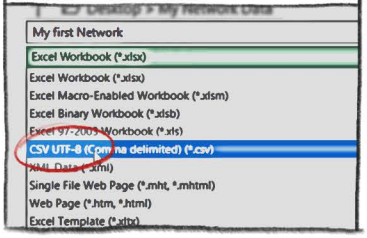
Do you know how to export a .csv file?

no → Microsoft Excel

Go to File > Save As...

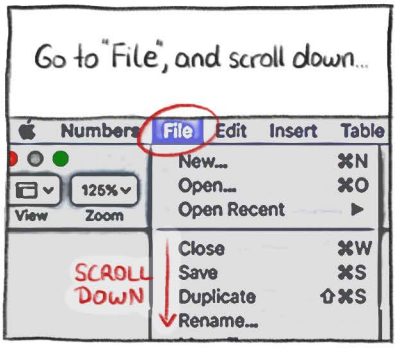


And select any "CSV" from the dropdown.

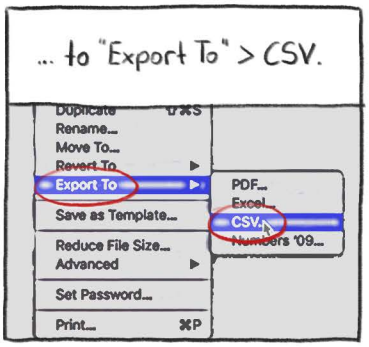


Apple Numbers

Go to "File", and scroll down...



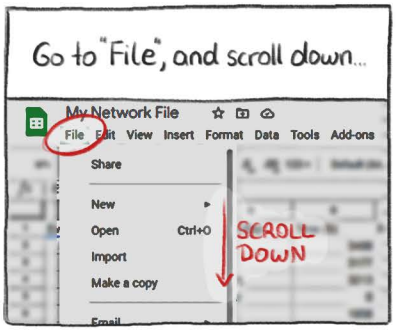
... to "Export To" > CSV.



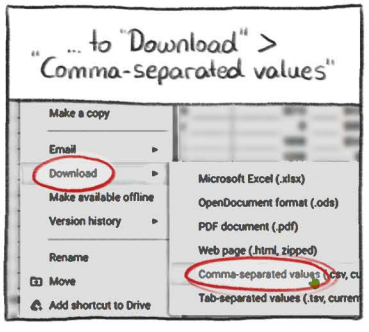
yes

Google Sheets

Go to "File", and scroll down...



... to "Download" > "Comma-separated values"



UPLOADING DATA & CREATING A NETWORK SCHEMA WITH THE VISTORIAN

4 Then, the Vistorian needs to know how links are represented in your data (i.e., whether you have a NODE TABLE or a LINK TABLE

So... which kind of table do you have?

LINK TABLE

Name_eng	Name_zh	Nationality	University	State	Uni_city	Degree_sou
Abend_Hallet	NA	Western	Stanford	California	Stanford	NA
Allman_N.F.	NA	Western	Virginia	Virginia	Charlottesvil	B.A.
Allman_Roy G	NA	Western	Virginia	Virginia	Charlottesvil	LL.B.
Anderson_Pat	安建生	Western	Emporia Colle	Kansas	Emporia	A.B.
Anderson_Pat	安建生	Western	Pittsburg The	Pennsylvani	Pittsburg	B.D.
Armour_Wen	NA					NA
Arnold_Harris	安					A.B.
Arnold_Julear	安					B.S.
Arnold_Julear	安					LL.D.

use your own data or the alumni - employers example dataset

You will be asked whether your links are directed.

5 Then, drag & drop your .csv file to upload it

NODE TABLE

CHILD	MOTHER	FATHER	GOD-FATHER	GOD-MOTHER
Bob	Celine	Charles	Dave	Eve
Ana	Fannie	Gerd	Mike	Dianne
Celine				Ana
Dave				Sophie
Mike				Amanda

use your own data or the example genealogy dataset

5 Drag & drop your .csv file to upload it.

6 You will see a preview of your table. Check whether it has a header row or not - you don't want the names of the columns to be added as nodes!

0	1	2
NAME	UNI	EMPLOYER
Bob	Stanf	Ministry of f
Darcy	Califo	Shanghai Te
Mira	Yale	Liberty Dair

☐ Has header row?

NAME	UNI	EMPLOYER
Bob	Stanf	Ministry of f
Darcy	Califo	Shanghai Te
Mira	Yale	Liberty Dair
Martin	Corne	National Sou

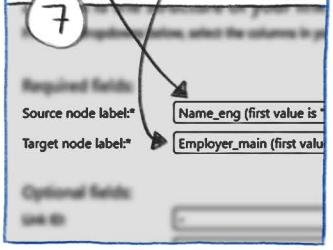
☑ Has header row?

Next, you need to tell the Vistorian which columns in your table describe your source and target nodes.

(Or, simply, which are the two nodes that have a relationship)

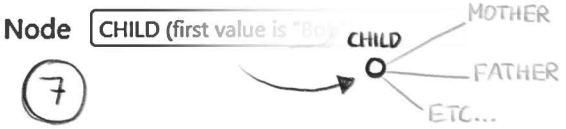
In our example, we want to see relations between alumni and employers:

7



Next, you need to specify which column of your table contains your main nodes, i.e., those you want to see the relationships of.

In the example genealogy, that would be the column "CHILD".



UPLOADING DATA & CREATING A NETWORK SCHEMA WITH THE VISTORIAN

Next, there are a few optional fields you can specify if your research question or network structure requires them:

OPTIONAL

Optional fields:

Link ID: [-] — LINK ID - in case you have unique IDs assigned to your links, which is only really used in special cases.

Location of source node: [-]
Location of target node: [-] — LOCATION - if your data contains geographical information and you want it visualized on a map

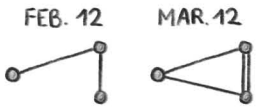
Link weight: [-] — LINK WEIGHT - if you have numerical values indicating the "strength" of a connection. (see more here:

Link type: [-] — LINK TYPE - if you have different kinds of relationships. (see more here:

Whether a link is directed: — WHETHER A LINK IS DIRECTED - only relevant if you have both directed and undirected links in your network, and a column that specifies something like directed - "yes" or "no"

8 Then, you have to check whether your links have associated times - if you want to explore your network over time.

e.g.,



3. Are links associated with time?
- (•) No, my links do not have associated times or this
- ( ) Yes, each link is associated with a *single time* (for

yes → Then, specify which column contains the times, and also which format they have.

e.g.,

24.03.98 → %d %m %y

March 24, 1998 → %B %d %Y

no →

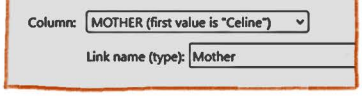
Then, all that's left to do is to specify which other columns your main nodes have a relationship to - and how you want to name them.

Click on the "Add Relation" - Button:

8 [blurred] just define at least

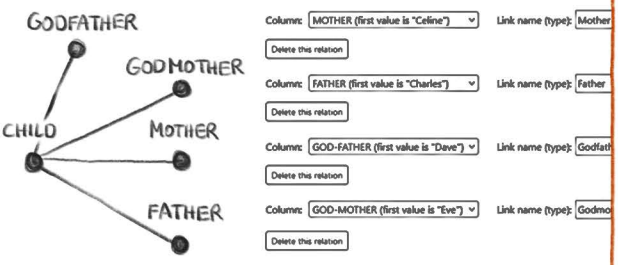
+ Add relation

9 Choose a related column - for example, "MOTHER", and name the relation...



The name you choose will automatically become the name of this relation's link type, so you can later find it in the legend easily.

You need to specify at least one type of relation to be able to form a network, but you can add as many as you want!



As soon as you've specified all relations that are of interest to you...

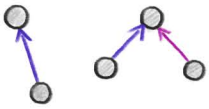
UPLOADING DATA & CREATING A NETWORK SCHEMA WITH THE VISTORIAN

9 Lastly, you have the option to upload an extra table (as a .csv file) that specifies node types, if you have one.

For the example data, you can use this file:

tinyurl.com/studentnodetypes

Without node type table:



Uni 1
Uni 2

With node type table:

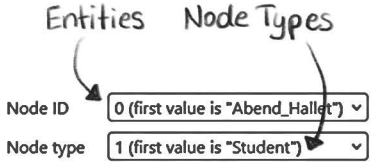
Employer
Student
Uni 1
Uni 2

Do you have an extra table with node types?

yes

10 Then, drag and drop your file like before and check again if it has a header row.

The Vistorian also needs to know which column of the table contains the Node ID (or rather, the entities whose type is specified), and which column contains the types.



no

As soon as you have completed these steps...