**A. Palladio’s sample dataset**

**Palladio** has its own sample dataset. It is a bigger but in some ways simpler data set than the one below; it consists of date and geographical information about influential people who visited Monaco.

1. Go to <http://palladio.designhumanities.org>
2. click “Start”
3. click “Try with sample data”
4. READ the entries in the table on the left to see what kind of information is available for each entry. (This is the data model.)
5. Click on Graph
6. Select SOURCE field (try birth date)
7. Select TARGET field (try death date)
8. Highlight either source or target (not both)
9. What is your graph showing?
10. What can you *immediately learn about the data* from this visualization?
11. Change the SOURCE to birth place
12. Change the TARGET to place of death
13. What is your graph showing?
14. What can you *learn about the dataset* from this visualization?

**B. Network analysis of a textual or historical source**

Copied, edited, and adapted from Martin Dürer’s Programming Historian Tutorial Marten Düring, "From Hermeneutics to Data to Networks: Data Extraction and Network Visualization of Historical Sources," *Programming Historian* (18 February 2015), http://programminghistorian.org/lessons/creating-network-diagrams-from-historical-sources.html

Step by Step:

**1. Palladio.** Go to [http://palladio.designhumanities.org/](http://palladio.designhumanities.org/%22%20%5Ct%20%22_blank).

**2. Start.** On their website click the “Start” button.

**3. Load attribute data.** From your data sheet, copy the attribute data ([Sample dataset](https://docs.google.com/spreadsheets/d/1LzbWsG73m74t3p6xE7lutfVWuOdzOIfN55FbhCCRZvk/edit%22%20%5Cl%20%22gid%3D0%22%20%5Ct%20%22_blank), Sheet 2) and paste it in the white section of the page, now click “Load”.

**4. Edit attributes.** Change the title of the table to something more meaningful, such as “People”. Now you see the columns “Person”, “Race Status” and “Sex” which correspond to the columns in the sample data. Next you need to make sure that Palladio understands that there are actions associated with the people you just entered in the database.

**5. Load relational data.** To do this, click on “Person” and “Add a new table”. Now paste all the relational data (*[Sample data](https://docs.google.com/spreadsheets/d/1LzbWsG73m74t3p6xE7lutfVWuOdzOIfN55FbhCCRZvk/edit%22%20%5Cl%20%22gid%3D0%22%20%5Ct%20%22_blank)*, Sheet 1) in the appropriate field. ***Palladio expects unique identifiers to link the relational information to the actor attribute information.*** Make sure this lines up well and that you avoid any irritating characters such as “/”. Palladio will prompt you with error messages if you do. Click “Load data”, close the overlay window and go back to the main data overview.

**6. Link attributes and relations.** Next, we need to explicitly link the two tables we created. In our case, peoples’ first- and last names work as IDs so we need to connect them. To do this click on the corresponding occurrences in the new table. In the sample files these are “Giver” and “Recipient”. Click on “Extension” (at the bottom) and select “People”, the table which contains all the people attribute information. Do the same for “Recipient”.

**7. Identify temporal data.** Palladio has nice time visualization features. You can use it if you have start and end points for each relation. The sample data contains two columns with suitable data. Click on “Time Step Start” and select the data type “Year or Date”. Do the same for “Time Step End” (Figure 10). The Palladio team recommends that your data is in the YYYY-MM-DD format, but my more abstract time steps worked well. If you were to load geographical coordinates (not covered by this tutorial but here: *[Palladio Simple Map Scenario](http://hdlab.stanford.edu/doc/scenario-simple-map.pdf%22%20%5Ct%20%22_blank)*) you would select the “Coordinates” data type.

**8. Open the Graph tool.** You are now done with loading the data. Click “Graph” to load the visualization interface (Figure 11).

**9. Specify source and target nodes.** First off Palladio asks you to specify the “Source” and “Target” nodes in the network (Figure 12). Let’s start with “Givers” and “Recipients”. You will now see the graph and can begin to study it in greater detail.

**10. Highlight nodes.** Continue by ticking the “Highlight” box for RECIPIENTS. This will give you an immediate sense of who received help. (See what happens when you highlight GIVERS instead.)

**11. Facet filter.** Next up, try the faceted filter (Figure 13). You will recognize the columns which describe the different acts of help. Start by selecting “3” in the “Form of Help” column. This will reduce the graph to only provisions of accommodation. Next, select values from the “Date of Activity” column to further narrow down your query. This will show you who provided accommodation and how this changes over time. Re-select all values in a column by clicking on the check box next to the column name. Take your time to explore the dataset – how does it change over time? When you are done, make sure to delete the Facet filter using the small red trashcan.

**12. Bipartite network visualization.** Now this is nice. But there is something else which makes Palladio a great tool to start out with network visualization: It makes it very easy to produce [bipartite, or 2-mode networks](http://en.wikipedia.org/wiki/Bipartite_graph%22%20%5Cl%20%22Examples%22%20%5Ct%20%22_blank). What you have seen until now is a so-called unipartite or 1-mode network: It represents relations between source and target nodes of one type (for example “people”) through one or more types of relations.

Network analysis however gives you a lot of freedom to rethink what source and targets are. Bipartite networks have two different types of nodes, an example could be to select “people” as the first node type and “point in time” as the second.

**Edit** your settings to set RECIPIENTS as the SOURCE and TIME OF ACTIVITY as the TARGET. This graph shows you who received help and when. It visualizes who was present to receive help at the same moment in time.

**Edit** your settings to set GIVERS as the SOURCE and TIME OF ACTIVITY as the TARGET. This graph shows you who gave help and when. This points at a high rate of fluctuation among helpers, an observation which holds true for all of the networks I studied. While humans are very good at processing people-to-people networks, we find it harder to process these more abstract networks.

Give it a try and experiment with different bipartite networks: Click again on “Target” but this time select “Form of Help” or “Sex” or any other category. ***Can you figure out what your graph is actually saying?***

**13. Timeline.** The Timeline feature provides a relatively easy way to visualize changes in your network. Reset your SOURCE to GIVER and TARGET to RECIPIENT. Highlight Recipient. Minimize that window. Then click on TIMELINE. Set height to Number of People, Group by Sex, and set dates to Time step End (the period of time the help ended).

**14. Time Span.** Even more interesting is the Time Span view which updates the network visualization dynamically. Click on “Time Span”. Use the mouse to highlight a section between the time steps which will then be highlighted in grey. You can now drag the highlighted section across the timeline and see how the graph changes from time step to time step.

**Export your visualizations.** Palladio lets you export your network as .svg files, a vector-based image format. Use your browser of choice to open them.

Visualisations are of course not “proof” of anything but tools to help understand complex relations; their interpretation is based on a good understanding of the underlying data and how it was visualized. Selected network visualizations can also accompany text and help your readers better understand the complex relationships you discuss, much like the maps you sometimes find on the inside covers of old books.

A few practical points:

* Collect and store data in one spreadsheet and use a copy for visualizations
* Make sure you understand the basic rationale behind any centrality and layout algorithms you choose as they will affect your view on your data. Wikipedia is usually a good source for comprehensive information on them.
* Don’t hesitate to revise and start over if you sense that your coding scheme does not work out as expected. It will definitely be worth it.

Finally, any of the visualizations you can create with the small sample dataset I provide for this tutorial requires context knowledge to be really meaningful. The only way for you to find out whether this method makes sense for your research is to start coding your own data and to use your own context knowledge to make sense of your visualizations.

**C. Download the Palladio Sample Data and Visualize in Google Fusion Tables instead.**

1. Go to palladio.designhumanities.org
2. click “Start,”
3. click “Try with sample data”
4. READ the entries in the table on the left to see what kind of information is available for each entry. (This is the data model.)
5. Click on Download
6. Go to Google Drive. (REMEMBER: you need to logout of Google completely and then login using your gmail address, NOT your u.pacific.edu address
7. Upload the comma-delimited file following the steps that are familiar to you
8. When you get to the screen where you can create charts and maps, click on the red plus sign to add a chart.
9. Click on the lower left network graph chart (the connected bubbles)
10. Experiment with the target and source (start with place of birth and place of death)
11. What happens when you tell it the link is directional? Or you color by columns?
12. Click on the blue button “Filter” and create a date filter – filter by the date of birth; what happens to the graph? What is the graph showing you now?
13. NOTE: **make sure you look at how many nodes are showing every time you adjust the graph, so you’re sure you’re showing the data you want.**
14. Experiment with the different targets and sources and filters

**D. Upload the Dürer dataset to Google Fusion Tables**

Hint: you can copy the URL of the Google spreadsheet instead of downloading and uploading the file.

Note: there are two sheets of information on Dürer’s spreadsheet; Google Fusion Tables will ask you which sheet you want to use. Start with the first one.

**Licensed CC-BY**

**Some material copied or adapted from *Programming Historian* as described above.**