Connect to Networks with NodeXL: The official guide
Acknowledgement:
This book is based on the original NodeXL documentation written by Tony Capone.

Author:
Harald Meier - Chief Data Scientist (Social Media Research Foundation)

Copyright Social Media Research Foundation
Version 2.0 (published November 2023)
Table of contents:

1. Introduction to Social Network Analysis and NodeXL.................................................. 6
   1.1 What is Social Network Analysis (SNA)?............................................................... 7
   1.2 What is NodeXL?.................................................................................................. 8
   1.3 Learn more about NodeXL.................................................................................. 10
   1.4 NodeXL Tutorials............................................................................................... 11
2. Getting started.............................................................................................................. 13
   2.1 Requirements....................................................................................................... 13
   2.2 Installation........................................................................................................... 14
   2.3 Orientation.......................................................................................................... 14
      2.3.1 The NodeXL Menu Ribbon Tab................................................................. 15
      2.3.2 The NodeXL Worksheets.......................................................................... 15
      2.3.3 The Graph Pane.......................................................................................... 17
      2.3.4 Showing and Hiding Workbook Columns.............................................. 18
      2.3.5 Working with multiple NodeXL instances.............................................. 19
   2.4 NodeXL Pro Cloud Edition.................................................................................. 20
      2.4.1 Accessing the NodeXL Pro Cloud Edition.............................................. 20
      2.4.2 File Management....................................................................................... 21
3. Data Import.................................................................................................................. 23
   3.1 Import Data Options............................................................................................. 23
   3.2 Manual Data Input................................................................................................ 24
   3.3 Importing Graph Data from Other Programs.................................................... 25
   3.4 Importing Graph Data from Another Workbook............................................... 26
   3.5 Importing Graph Data from Online Social Networks......................................... 28
   3.6 Flickr Networks.................................................................................................... 31
      3.6.1 Flickr Related Tags Networks................................................................. 31
      3.6.2 Flickr User’s Networks............................................................................. 32
   3.7 Wikipedia Networks............................................................................................. 33
   3.8 X (formerly Twitter) Networks........................................................................... 35
      3.8.1 Import from X (formerly Twitter) Search Network 3.0 (Beta)................ 35
      3.8.2 Import from Brandwatch......................................................................... 37
      3.8.3 Import from Meltwater............................................................................. 38
      3.8.4 Import from Talkwalker......................................................................... 39
      3.8.5 Import from Tweet Binder...................................................................... 40
   3.9 YouTube Networks............................................................................................... 41
      3.9.1 YouTube Channel’s Network.................................................................... 41
      3.9.2 YouTube User Network.......................................................................... 41
      3.9.3 YouTube Video Network........................................................................ 43
   3.10 Reddit Networks.................................................................................................. 45
   3.11 Import from WhatsApp...................................................................................... 47
4. Data Preparation.......................................................................................................... 49
   4.1 Directed vs. Undirected Graphs.......................................................................... 49
   4.2 Edge Weight: Counting and Merging Duplicate Edges...................................... 50
5. Working with Groups................................................................................................... 54
5.1 Creating Groups by Vertex Attribute.................................................................55
5.2 Creating Groups by Connected Component.......................................................55
5.3 Creating Groups by Cluster..............................................................................56
5.4 Creating Groups by Motif.................................................................................58
5.5 Manually Creating Groups................................................................................59
5.6 Understanding the Group Worksheets...............................................................60
5.7 How Groups Are Shown in the Graph Pane.......................................................61
  5.7.1 Color and Shape.........................................................................................61
  5.7.2 Group Options: Changing How Vertex Colors and Shapes are Specified......61
  5.7.3 Group Layout..............................................................................................62
  5.7.4 Skipping All Groups...................................................................................63
5.8 Selecting Groups...............................................................................................64
5.9 Collapsing and Expanding Groups.....................................................................65
5.10 Hiding and Skipping Groups............................................................................66
5.11 Removing Groups.............................................................................................66

6. Calculating Graph Metrics..................................................................................67
  6.1 Overall Graph Metrics......................................................................................69
  6.2 Vertex Metrics..................................................................................................71
  6.3 Group Metrics..................................................................................................73
  6.4 Time Series Analysis.........................................................................................74
  6.5 Path Analysis....................................................................................................75
  6.6 Word and Word Pairs / Sentiment analysis......................................................76
  6.7 Top items..........................................................................................................78
  6.8 Twitter Search Network Top Items.................................................................78
  6.9 Network Top Items..........................................................................................79
  6.10 Edge creation by shared content similarity....................................................80
  6.11 Summarizing the Graph..................................................................................80

7. Changing How the Graph Looks.........................................................................81
  7.1 Graph Options: Default Visual Properties......................................................81
  7.2 Setting Visual Properties for Individual Edges, Vertices or Groups..................83
  7.3 Labeling Edges, Vertices and Groups................................................................84
  7.4 Autofill Columns: Automatically Calculating Visual Properties for All Edges, Vertices or Groups..........................................................86
  7.5 Changing How the Graph is Laid Out.............................................................88
    7.5.1 Layout Algorithms....................................................................................88
    7.5.2 Layout Options..........................................................................................90
    7.5.3 Selectively Laying Out Parts of the Graph Again......................................92
    7.5.4 Snapping Vertices to a Grid.....................................................................93
  7.6 Zooming, Moving Around and Scaling the Graph.............................................93
    7.6.1 Zooming the Graph..................................................................................94
    7.6.2 Moving Around the Graph......................................................................95
    7.6.3 Scaling the Graph....................................................................................95
  7.7 Selecting Graph Elements................................................................................96
    7.7.1 Selecting Vertices....................................................................................96
    7.7.2 Selecting Edges........................................................................................97
7.8 Graphs with Isolated Vertices................................................................. 99
7.9 Adding Tooltips to the Graph................................................................. 100
7.10 Showing and Hiding Graph Elements.................................................. 100
7.11 Using Dynamic Filters........................................................................ 101
7.12 Creating Subgraph Images.................................................................... 103
7.13 Saving an Image of the Graph............................................................... 105

8. Exporting Graph Data............................................................................. 106
8.1 Exporting Graph Data to Other Programs............................................. 106
8.2 Exporting Graph Data to Another Workbook....................................... 108
8.3 Export Options..................................................................................... 109
8.4 Export to NodeXL Graph Gallery......................................................... 110
8.5 Export to NodeXL Pro INSIGHTS......................................................... 112

9. Working with Options............................................................................ 116
9.1 Using the Current Workbook's Options for New Workbooks.............. 116
9.2 Exporting and Importing Options......................................................... 117
9.3 Resetting All Options.......................................................................... 117

10. Task Automation.................................................................................... 118
10.1 What is a NodeXL Pro Options file?.................................................... 118
10.2 Data Recipes....................................................................................... 120
  10.2.1 Basic Data Recipes......................................................................... 120
  10.2.2 X (Twitter) Data Recipes................................................................. 121
  10.2.3 YouTube Data Recipes.................................................................... 124
  10.2.4 Wikipedia Data Recipes.................................................................. 127
  10.2.5 Flickr Data Recipes.......................................................................... 129
  10.2.6 Reddit Data Recipes....................................................................... 130
  10.2.7 Custom Data Recipes..................................................................... 131

11. Tutorial: Social network and content analysis - step by step............... 132

12. Keyboard Shortcuts in the Graph Pane.................................................. 134

13. Frequently Asked Questions (FAQ)....................................................... 135
  13.1 General Questions.............................................................................. 135
  13.2 License.............................................................................................. 135
  13.3 Data Importers.................................................................................. 136
  13.4 Error Messages.................................................................................. 137
  13.5 Data handling................................................................................... 138
  13.6 Research........................................................................................... 139

14. Links and Literature............................................................................... 140
  14.1 Important web links............................................................................ 140
  14.2 Major publications around NodeXL................................................... 140
  14.3 Publications related to NodeXL features.......................................... 141
  14.4 Publications with NodeXL as research tool...................................... 141
  14.5 Recommended books and articles around social network analysis..... 142
1. Introduction to Social Network Analysis and NodeXL

Social network analysis (SNA) is a methodological approach used to understand and analyze the relationships and patterns among (social) entities. It is primarily focused on understanding the structures formed by these relationships, often visualized as network graphs.

A network graph is a collection of points connected by lines. The lines are called “edges”, and the points are called “vertices” or “nodes”. NodeXL always uses the terms "vertex/vertices" rather than "node," contrary to its product name. Edges are the foundation of any network and refer to the connections or links between the vertices. Vertices on the other hand, represent the entities or elements within the network.

Here is a simple network graph:

![Simple Network Graph](image)

Network graphs can be used to model a wide variety of real-world and theoretical systems, ranging from social organizations consisting of people connected by relationships, to cities connected by transportation routes, to atoms connected by chemical bonds.

If you are new to network graphs, the Wikipedia article [Graph Theory](https://en.wikipedia.org/wiki/Graph_theory) can serve as a good starting point for learning about them. To learn more about networks, network theory and NodeXL, we recommend [the official NodeXL book](https://shop.elsevier.com/books/analyzing-social-media-networks-with-nodexl/hansen/978-0-12-817756-3) with many practical examples.

NodeXL is a powerful social network analysis software that allows users to analyze and visualize networks of various sizes and complexities. NodeXL is an acronym that stands for **Network, Overview, Discovery and Exploration in Excel**. It is a plug-in that uses the familiar Microsoft Excel spreadsheet application, while requiring no programming skills:

---

1. [https://en.wikipedia.org/wiki/Graph_theory](https://en.wikipedia.org/wiki/Graph_theory)
1.1 What is Social Network Analysis (SNA)?

The origins of social network analysis can be traced back to the early 20th century. It drew from various disciplines including sociology, anthropology, psychology, and mathematics. Key theoretical foundations were laid by scholars like Georg Simmel and Émile Durkheim, who were interested in understanding the structure of social groups and the ties that bind individuals together.

In the mid-20th century, more formal methods of social network analysis began to develop. This was facilitated by advancements in graph theory in mathematics and the growing availability of computational resources. Sociologists like Jacob Moreno, who introduced sociometry, played a critical role in this phase. Sociometry involves the mapping and measuring of relationships among individuals within a group.

The 1970s and 1980s saw a significant expansion in the field of social network analysis. This period marked the development of new methodologies and software that could handle larger and more complex network data. The use of SNA expanded beyond sociology into fields like organizational studies, political science, and public health.

Today social network analysis is a crucial tool in many fields, used to understand complex relational data and to inform decision-making in areas ranging from marketing to public health planning.

SNA examines the ties, links, or edges between these actors, which represent relationships or interactions. It analyzes the patterns of these relationships and calculates various metrics on three levels - the individual, the group, and overall network structure.
1.2 What is NodeXL?

NodeXL is developed by the Social Media Research Foundation and provides a user-friendly social network analysis interface that enables users to directly import, analyze, and visualize data from various sources, including social media platforms like YouTube, X (formerly Twitter), Reddit, Wikipedia and Flickr.

NodeXL is used by researchers, students, and professionals from around the world in a wide range of fields, including sociology, political science, communication studies, health, business, and marketing. It allows users to explore the structure of networks, identify key nodes and clusters, and uncover patterns of interaction and influence among network actors. NodeXL provides a range of features that make it a powerful and versatile tool for social network analysis.

In this book, we will provide an overview of NodeXL’s analytical features and its capabilities for (social) network analysis. Some of the key features of NodeXL include:

1. **Data import**: NodeXL allows users to import data from various sources, including social media platforms. The software supports a wide range of file formats, including Excel, CSV, GraphML, GDF and GEXF.

2. **Data analysis**: NodeXL provides a range of tools for analyzing network data, including centrality metrics, clustering, and community detection. The software also supports content analysis such as text and sentiment analysis, time series analysis and network top item analysis.

3. **Network visualization**: NodeXL provides a range of options for visualizing network data.

4. **Data export**: NodeXL can save workbooks to various file formats, and export reports to NodeXL Graph Gallery and NodeXL Pro INSIGHTS.

NodeXL is split into two versions: NodeXL Basic and NodeXL Pro. **NodeXL Basic** is free, contains basic analytical features and allows the visualization of simple networks. It is also positioned as a browser for workbooks created with NodeXL Pro. **NodeXL Pro** is a paid service that allows a full-scale social network and content analysis of network data that includes the calculation of advanced metrics, layouts and task automation. In the following text, features only available in NodeXL Pro are marked with the NodeXL Pro logo.

---

3 https://www.smrfoundation.org/
4 https://nodexlgraphgallery.org
5 https://www.smrfoundation.org/nodexl/nodexl-pro-insights
6 https://nodexlgraphgallery.org/Pages/registration.aspx
<table>
<thead>
<tr>
<th>Topic</th>
<th>Feature</th>
<th>NodeXL Basic</th>
<th>NodeXL Pro</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Network Visualization</strong></td>
<td>Visualize your own network graph</td>
<td>✅</td>
<td>✅</td>
</tr>
<tr>
<td></td>
<td>Choose from various layout algorithms</td>
<td>✅</td>
<td>✅</td>
</tr>
<tr>
<td></td>
<td>Set the color, shape, size, label, and opacity of vertices and edges</td>
<td>✅</td>
<td>✅</td>
</tr>
<tr>
<td><strong>Network Metrics</strong></td>
<td>Calculate overall network and group metrics (Density, Modularity, etc.)</td>
<td>✅</td>
<td>✅</td>
</tr>
<tr>
<td><strong>Group Analysis</strong></td>
<td>Group vertices by cluster or attributes</td>
<td>✅</td>
<td>✅</td>
</tr>
<tr>
<td></td>
<td>Group labeling with Top Items</td>
<td>✅</td>
<td>✅</td>
</tr>
<tr>
<td><strong>Centrality Metrics</strong></td>
<td>Degree, In-Degree, Out-Degree</td>
<td>✅</td>
<td>✅</td>
</tr>
<tr>
<td></td>
<td>Betweenness Centrality, Closeness Centrality, Eigenvector Centrality, PageRank</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td><strong>Content Analysis</strong></td>
<td>Text Analysis</td>
<td>✅</td>
<td>✅</td>
</tr>
<tr>
<td></td>
<td>Sentiment Analysis</td>
<td>✅</td>
<td>✅</td>
</tr>
<tr>
<td></td>
<td>Time Series Analysis</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Top Items: Words/Word pairs/URLs/Hashtags</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td><strong>Data Import</strong></td>
<td>YouTube</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reddit</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Flickr</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wikipedia</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td></td>
<td>X (formerly Twitter): Direct import,</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Excel workbook tables, csv files</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GraphML, Pajek, UCINet, and matrix formats</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td><strong>Data Export</strong></td>
<td>NodeXL Graph Gallery</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NodeXL Pro INSIGHTS</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MS Office Powerpoint</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Export to csv file</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GraphML, Pajek, UCINet, and matrix formats</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td><strong>Task Automation</strong></td>
<td>Automate all tasks above for a full scale analysis.</td>
<td>✅</td>
<td></td>
</tr>
</tbody>
</table>
1.3 Learn more about NodeXL

Visit our homepage to learn more about our work at the Social Media Research Foundation and NodeXL and [sign up for the newsletter](https://www.smrfoundation.org/newsletter/) to not miss the latest news about NodeXL.

Check out our [tutorial page](https://www.smrfoundation.org/nodexl/tutorials/) to learn how to use NodeXL in many different ways and contexts.

Visit the [NodeXL Graph Gallery](https://nodexlgraphgallery.org/) where the NodeXL community shares network data sets and maps.

In July 2019, Morgan Kaufmann published the second edition of the [official NodeXL book](https://doi.org/10.1016/C2018-0-01348-1) titled *Analyzing Social Media Networks with NodeXL - Insights from a Connected World* authored by Derek Hansen, Ben Shneiderman, Itai Himmelboim and Marc Smith. This book is used in many classrooms around the world to teach students about network theory and using NodeXL.

![The official NodeXL Book](https://www.smrfoundation.org/images/node-xl-book.png)

**Figure: The official NodeXL Book**

The book includes 3 parts: Part 1 introduces social media and network analysis, Part 2 is a tutorial introducing core network concepts while walking through the major features of NodeXL, and Part 3 is a collection of case studies written by various network experts.

---

7 [https://www.smrfoundation.org/newsletter/](https://www.smrfoundation.org/newsletter/)
8 [https://www.smrfoundation.org/nodexl/tutorials/](https://www.smrfoundation.org/nodexl/tutorials/)
9 [https://nodexlgraphgallery.org/](https://nodexlgraphgallery.org/)
10 [https://doi.org/10.1016/C2018-0-01348-1](https://doi.org/10.1016/C2018-0-01348-1)
analyzing different social media tools (e.g., email, forums, Twitter, YouTube, Flickr and Wikis).

1.4 NodeXL Tutorials

Make sure to frequently check the tutorials\textsuperscript{11} section on our homepage. Here is an overview of currently available tutorials:

<table>
<thead>
<tr>
<th>Tutorial</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>How to Automate NodeXL Pro</td>
<td>The most powerful feature of NodeXL Pro is Task Automation. The automation feature allows you to run all steps of a social network and content analysis with a single click: Data preparation, cluster analysis, metrics calculation, time series analysis, top content analysis, visualization and data export. To learn how to automate NodeXL Pro, either scroll down to chapter Task Automation, visit this website, download this pdf file and/or watch this video.</td>
</tr>
<tr>
<td>Social network and content analysis with X (Twitter) network data – step by step</td>
<td>This tutorial shows you how you can run a full social network and content analysis with NodeXL Pro. While we will use X (formerly Twitter) network data as an example, this approach can be applied to any network dataset of your choice (content analysis depends on the available metadata). Download as pdf file.</td>
</tr>
<tr>
<td>Analyzing YouTube User Networks</td>
<td>This tutorial shows you how to analyze YouTube User Networks. The networks are based on comments and replies which are posted by users in the comments section below a video. With the NodeXL Pro YouTube User network importer you can analyze user networks around a single video or multiple videos.</td>
</tr>
<tr>
<td>Wikipedia article-article networks</td>
<td>This tutorial shows the steps that are required to generate a Wikipedia Article-to-Article network by using NodeXL Pro’s “From MediaWiki Page Network” importer.</td>
</tr>
<tr>
<td>Flickr related tag networks</td>
<td>This tutorial describes the steps needed to generate a network composed of Tags on Flickr photos and their connections to Tags that co-occur on photos using NodeXL Pro’s Flickr Related Tags Network importer.</td>
</tr>
<tr>
<td>Semantic Networks – Create networks with words, hashtags or video tags</td>
<td>This tutorial shows you how to create a semantic network by using the text analysis feature of NodeXL Pro which can be applied to any column that contains text in the edges or vertices spreadsheets of a NodeXL workbook. Click here.</td>
</tr>
</tbody>
</table>

\textsuperscript{11} https://www.smrfoundation.org/nodexl/tutorials
<table>
<thead>
<tr>
<th>Semantic networks from qualitative interview transcripts</th>
<th>This tutorial shows how to generate semantic networks from qualitative interview transcript files using NodeXL Pro.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Redes Semánticas con NodeXL (en español)</td>
<td>Este tutorial muestra los pasos para generar una red semántica mediante el uso de la función de análisis de texto de NodeXL Pro, que se puede aplicar a cualquier columna que contenga texto en las hojas de cálculo de un libro de trabajo NodeXL.</td>
</tr>
<tr>
<td>Redes semánticas con NodeXL Pro a partir de transcripciones de entrevistas cualitativas (en español)</td>
<td>Este tutorial muestra cómo generar redes semánticas a partir de archivos de transcripciones de entrevistas cualitativas mediante la herramienta de NodeXL Pro.</td>
</tr>
<tr>
<td>Visualize your NodeXL data in Graphistry using Python</td>
<td>In this tutorial, we will review how to connect a NodeXL Workbook (Twitter) to the Graphistry tool. We will take the NodeXL Workbook from the NodeXL Gallery to connect it with Graphistry and visualize the network in this tool. For this, we will use Python code.</td>
</tr>
<tr>
<td>Exploring YouTube Video Recommendation Networks</td>
<td>This tutorial shows you how to create YouTube Video-to-Video Recommendation Networks with the NodeXL Pro YouTube Video Network importer.</td>
</tr>
<tr>
<td>(Outdated since August 2023 due to YouTube API policy changes)</td>
<td></td>
</tr>
<tr>
<td>Working with Twitter User lists</td>
<td>Twitter User lists are a very helpful tool to manage the diverse information streams on Twitter. This tutorial shows you how to work with Twitter User lists using the NodeXL Pro Users Network Importer.</td>
</tr>
<tr>
<td>(Outdated since May 2023 due to Twitter API policy changes)</td>
<td></td>
</tr>
<tr>
<td>How to download your Twitter follower data</td>
<td>This tutorial explains how you can download, analyze and visualize your Twitter follower data with NodeXL Pro using the NodeXL Twitter Users Network Importer.</td>
</tr>
<tr>
<td>(Outdated since May 2023 due to Twitter API policy changes)</td>
<td></td>
</tr>
</tbody>
</table>
2. Getting started

In order to install NodeXL Basic or NodeXL Pro you need to register on this page\(^{12}\). After that you will receive an email with a download link to the software and installation instructions.

To use NodeXL, every user must agree to the NodeXL End User License Agreement\(^ {13}\).

2.1 Requirements

NodeXL requires:

**OPERATING SYSTEM:** Microsoft Windows™ (7, 8, 10, 11). Mac and Linux users should have a look below.

**SOFTWARE:** Microsoft Office™ (2007, 2010, 2013, 2016, 2019, 2021, 365) 64-bit versions of Windows and Office are recommended. 32-bit systems cannot use large amounts of RAM.

**RAM:** More the better! 4GB is enough to handle very small networks, 8GB for medium networks, 16GB or 32GB is recommended for very large networks (100K edges and above).

**CPU:** Most computers are now fast enough to handle network calculations. That said, faster CPUs are good for improving NodeXL's performance. However, currently NodeXL is a (mostly) single threaded application that does not make use of multiple cores.

**Attention Mac/Linux Users:**

NodeXL is Windows and Office only. Mac users have several options:

1. **NodeXL Pro Cloud Edition:** You can now subscribe to the NodeXL Pro Cloud Edition\(^ {14}\) and receive access to a fully configured Windows machine via Amazon Web Services. More information on the NodeXL Pro Cloud Edition is available in chapter 2.4.

2. **Local Virtual Machine:** While it is possible to run a local virtual machine on a MacBook, this is not a performant way to run the application. It can be slow! Further, a MacBook running Windows plus Office plus NodeXL should have no less than 16GB of RAM.

3. **Remote Virtual Machine:** Remote VMs place almost zero resource demands on the client machine (the Mac) and do not require special amounts of RAM or features to perform well. Two options are available:

   a. **Remote VM via Amazon:** We recommend remote virtual machines like those offered by Amazon. Here is an article about creating a virtual machine\(^ {15}\) in the
Amazon Elastic Compute Cloud (EC2). This is a rather complicated process though.

b. **RemoteVM via Microsoft**: Remote virtual machines may be accessed using the free "Remote Desktop" application from Microsoft\(^\text{16}\) available for all Macs.

### 2.2 Installation

1. After registering for NodeXL Basic or NodeXL Pro you will receive an email with a download link to the software and installation instructions.
2. Download the NodeXL Pro Excel Template or NodeXL Basic Excel Template from the link provided in the email and then run it (as administrator).
3. If you are asked to accept the “Microsoft Visual Studio 2010 Tools for Office Runtime (x86 and x64)”, click the “Accept” button.
4. When you are asked **"Are you sure you want to install this customization?"**, click the **“Install”** button and wait for the installation to finish.
5. To open **NodeXL**, open the Windows Start menu or Start screen, search for "NodeXL", then click "NodeXL Excel Template" in the search results or look for the desktop shortcuts named NodeXL Excel template.
6. When opening **NodeXL Pro** for the first time, you will be asked to provide the location of your NodeXL Pro User license.

If you run into any issues during installation, please take a look at this installation guide\(^\text{17}\).

### 2.3 Orientation

When you open the NodeXL Excel Template from the Windows Start Menu, Excel creates a new workbook from the template and opens the workbook. This NodeXL workbook differs from a standard Excel workbook in that it contains multiple worksheets that have been customized for network graph data, a custom Graph Pane used to display the network graph, and a custom NodeXL menu tab in the Excel Ribbon that contains commands specific to NodeXL.

The NodeXL menu tab is divided into seven sections: Data - Graph - Visual Properties - Analysis - Options - Show/Hide - Help.

These features appear only in NodeXL workbooks, and they are removed entirely from Excel when you close the NodeXL workbook. The following diagram shows the NodeXL worksheets, the Graph Pane and the NodeXL tab in the Excel Ribbon.

\(^{16}\) https://docs.microsoft.com/en-us/windows-server/remote/remote-desktop-services/clients/remote-desktop-mac

\(^{17}\) https://www.smrfoundation.org/nodexl/installation/
2.3.1 The NodeXL Menu Ribbon Tab

The NodeXL workbook adds a custom NodeXL menu tab to the Excel Ribbon. This menu tab contains all the commands that are specific to NodeXL graphs, organized into groups. It does not appear in standard Excel workbooks.

To learn more about a command in the NodeXL Ribbon tab, hover the mouse over the command.

2.3.2 The NodeXL Worksheets

There are five worksheets built into a NodeXL workbook. With a simple graph, the Edges worksheet may be the only one you work with. The other worksheets offer customization options for more advanced scenarios.

Additional worksheets are added when using NodeXL key features like text analysis, time series analysis or top content analysis.

NodeXL will also create several hidden worksheets in the workbook that are needed to perform certain calculations. Users are encouraged not to modify these worksheets in any way.

**Important note:** Do not change the order of the worksheets because it may corrupt the underlying workbook structure which may result in error messages when doing calculations.
Also, when adding either new worksheets to the workbook or adding new columns to a worksheet, add them on the far right side, making them last in the list of worksheets or in the last column in a table.

The following table describes the basic NodeXL worksheets.

<table>
<thead>
<tr>
<th>Worksheet</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edges</td>
<td>Contains the graph's edges, which you specify as Vertex 1/Vertex 2 name pairs, along with optional edge attributes. The vertex name pairs are the only columns in the entire workbook that are required.</td>
</tr>
<tr>
<td>Vertices</td>
<td>Contains the graph's vertices, along with optional vertex attributes. When you show the graph, NodeXL automatically creates a row in this worksheet for each vertex in the Edges worksheet.</td>
</tr>
<tr>
<td>Groups</td>
<td>Contains the names of the graph's groups (which are composed of collections of vertices), along with the groups' attributes. See 5.6 Understanding the Group Worksheets for more information. If your graph does not have groups, you can ignore the Groups worksheet.</td>
</tr>
<tr>
<td>Group Vertices</td>
<td>Contains the names of the vertices in each of the graph's groups. See 5.6 Understanding the Group Worksheets for more information. You can usually ignore the Group Vertices worksheet, even if your graph has groups.</td>
</tr>
<tr>
<td>Overall Metrics</td>
<td>Contains the graph's overall metrics, which you can calculate on demand. See 6. Calculating Graph Metrics for more information. If you do not calculate graph metrics, you can ignore the Overall Metrics worksheet.</td>
</tr>
</tbody>
</table>

If you calculate 6.3 group metrics, a Group Edges worksheet gets added to the workbook. This worksheet lists the number of edges within each group and the number of edges between each pair of groups.

Additional worksheets are added when conducting Words/Word pairs/sentiment analysis, time series analysis, path analysis, and network top content analysis.

**To select a worksheet:**

- Click the worksheet's tab at the lower left corner of the worksheet area of the Excel window, as shown below. You might have to scroll the tabs to the left or right to see them all. You can do that with the arrow buttons to the left of the tabs.

**To learn more about a worksheet:**

- Hover the mouse over the first header cell on the worksheet. For example, to find out more about the Edges worksheet, hover the mouse over the "Vertex 1" header cell.
To learn more about a worksheet column:

- Hover the mouse over the column’s header cell.

To add a column to a worksheet:

- Hover the mouse over the "Add Your Own Columns Here" column header for instructions.

To add a worksheet to the workbook:

- Click the "Insert Worksheet" button to the immediate right of the Overall Metrics tab at the lower left corner of the worksheet area of the Excel window. If you cannot see the Overall Metrics tab, click the small right-arrow to the left of the tabs until the Overall Metrics tab and the "Insert Worksheet" button are visible.

Important Note: Always keep added worksheets to the right of the built-in NodeXL worksheets. Inserting a worksheet between the built-in NodeXL worksheets can lead to unpredictable undesirable results.

2.3.3 The Graph Pane

The Graph Pane is where NodeXL shows the visualization of the network graph when you click the Show Graph button. It’s on the right side of the main Excel window.

Confusingly, the Graph Pane has a title bar imposed by the Excel application that says "Document Actions," but in NodeXL you should think of it as the Graph Pane. Document Actions is an Excel pane that is not exclusive to NodeXL, and Excel calls it Document Actions no matter which program is using it.

The Graph Pane can be resized, undocked from the main Excel window, and hidden.

To resize the Graph Pane:

1. Hover the mouse over the vertical divider between the Graph Pane and the NodeXL Worksheets. The mouse cursor changes to a double-headed arrow.
2. Drag the vertical divider to the left or right.

To undock the Graph Pane from the main Excel window:

- Drag the Graph Pane's title bar, the one that says "Document Actions." You might have to drag the title bar a considerable distance before the Graph Pane becomes undocked.

To dock the Graph Pane to the main Excel window:

- Drag the Graph Pane's title bar to one of the edges of the main Excel window.
To hide the Graph Pane:

- Click the X at the upper-right corner of the Graph Pane.

To show the Graph Pane:

- In the NodeXL Ribbon Tab select Graph > Show/Refresh Graph.

The Graph Legend

The Graph Pane can include a legend if you choose to show one. The legend shows any visual attributes that have been autofilled using NodeXL’s 7.4 Autofill Columns feature. If you autofill the Vertex Size column using the Degree column as a source, for example, then the legend will look something like this:

![Figure: Graph Legend](image)

This legend indicates that the Size column was autofilled with Degree, which ranged from 1 to 9.

To show the graph legend, do the following:

1. In the Excel Ribbon, select NodeXL > Show/Hide > Graph Elements.
2. Check Legend.

The Graph Axes

The Graph Pane can also include x- and y-axes, which can be useful when you autofill the X and Y columns on the Vertices worksheet. The axes include labels that indicate the source columns, and the axis units are determined by the range of numbers in the source columns. To show the graph axes, do the following:

1. In the Excel Ribbon, select NodeXL > Show/Hide > Graph Elements.
2. Check Axes.

2.3.4 Showing and Hiding Workbook Columns

NodeXL’s worksheets have many columns, some of which you may never use. If you want to reduce worksheet clutter so you can concentrate on the columns that are important to you, you can tell NodeXL to hide some of the columns.
The columns are arranged in column groups. For example, the Edges worksheet includes Visual Properties, Label and Other column groups, and the Vertices worksheet includes the same column groups and several others. (The Other column group is where you can add your own columns to the worksheet.) NodeXL hides and shows workbook columns by column group.

**To show or hide column groups:**

1. In the Excel Ribbon, select **NodeXL > Show/Hide > Workbook Columns**.
2. Check the column groups you want to show and uncheck the column groups you want to hide.

### 2.3.5 Working with multiple NodeXL instances

If you would like to work with several data sets at the same time, you need to open the workbooks in separate instances of Excel. Opening two or more workbooks in the same instance of Excel will lock all other workbooks while running tasks in one workbook.

**To open NodeXL in a new instance:**

1. On your desktop, double-click on the icon **Open NodeXL In New Excel Instance**, or
2. Push the Alt key immediately after double-clicking the icon **NodeXL Excel Template**

To check whether a workbook was indeed opened in a new instance, check your machine’s **Task Manager** via Ctrl+Alt-Delete. There should be one row for each instance of Excel and NodeXL.
2.4 NodeXL Pro Cloud Edition

NodeXL Basic and NodeXL Pro only work on Windows operating systems. Mac and Linux users have several options to make use of the software as explained in chapter 2.1 Requirements.

When you subscribe to the NodeXL Pro Cloud Edition\(^{18}\) (described below) you will receive access to a fully configured Windows machine via Amazon Web Services accessible via almost any web browser.

Powered by Amazon Web Services AppStream, the NodeXL Pro Cloud Edition offers reliable, accessible, and performant access to NodeXL Pro. After subscribing to the NodeXL Pro Cloud edition, you will receive a URL that links directly to a view of a remote Microsoft Windows PC with 8 GB RAM, Microsoft Office, and NodeXL Pro. The number of computing hours per month is limited. Depending on your subscription you have access to 40 compute hours per month with the NodeXL Pro Student Cloud Edition, 60 compute hours per month with the NodeXL Pro Academic Cloud Edition and 80 compute hours per month with the NodeXL Pro Commercial Cloud Edition. Here are a few frequently asked questions about the NodeXL Pro Cloud Edition:

**Do I need a Windows PC to access NodeXL Pro Cloud Edition?**

No, you only need a web browser (e.g. Chrome, Edge, Firefox) to access the NodeXL Pro Cloud Edition. Internet connected Macs, Linux PCs, and Tablets can all access NodeXL Pro Cloud Edition.

**Can I use a Mac/iPad to access NodeXL Pro Cloud Edition?**

Yes! You only need a web browser (e.g. Chrome, Edge, Firefox) to access the NodeXL Pro Cloud Edition.

**What are the system requirements to run NodeXL Pro Cloud Edition?**

System requirements vary, but most Macs, PCs, iPads, Tablets, and Linux machines can access NodeXL Pro Cloud Edition.

2.4.1 Accessing the NodeXL Pro Cloud Edition

After subscribing to the NodeXL Pro Cloud Edition you will receive an email with a link to your personal cloud instance. When opening that link, it may take one or two minutes until the start screen shows up. When it appears, NodeXL Pro is installed and ready to use:

2.4.2 File Management

**Important note 1:** There is a folder named “Home Folder” available from the Windows Explorer via **This PC > Home Folder**. Save all your files in this folder! If you save your files in any other location, they will be deleted and not be available the next time you open your instance of the NodeXL Pro Cloud Edition!

**Important note 2:** If you leave your NodeXL Pro Cloud Edition instance unused for an hour, the session will close. The session is accessible for up to another 3 hours. After that the session will shut down and unsaved files will be lost. You can restart your instance at any time.

Files placed in the **Home Folder** are retained for the duration of your subscription. You can **transfer these files to other systems** by configuring one of the pre-installed file sharing applications (like OneDrive or Google Drive). You can also upload files to your Cloud instance and download files to your local machine via the third icon from the left **“My Files”** in the upper left corner of the screen:

Your NodeXL Pro Cloud Edition instance comes **fully configured with a wide range of “data recipes”** — NodeXL Settings Options files — that guide NodeXL in processing your
network data from a variety of sources. Using these recipes, it is easy to configure NodeXL to import and process a range of social media network data, including X (formerly Twitter), YouTube, Reddit, Wikis, flickr, and other types of network data sets.

**Important note 3:** Even though the NodeXL Pro Cloud Edition is equipped with those “data recipes”, we recommend that you download the latest set of recipes from this website and also save them to the Home Folder mentioned above.

---

19 [https://www.smrfoundation.org/nodexl/automation/](https://www.smrfoundation.org/nodexl/automation/)
3. Data Import

You can either insert data into the Edges worksheet manually or you can import graph data into a NodeXL workbook from a variety of sources in a variety of formats.

3.1 Import Data Options

Imported graph data normally overwrites any graph data that is already in the workbook, but you can change this so that imported data gets appended to existing data instead. Appending graph data can lead to confusing results - in particular, you can end up with multiple rows in the Vertices worksheet for the same vertex - so this option is intended for advanced users only.

To set basic data import options:

1. In the Ribbon, select NodeXL > Data > Import > Import Options…

To tell NodeXL to append imported graph data to existing data:

2. Uncheck “Clear NodeXL Workbook First”.

To add a short description when importing data:

3. Check “Add a description of the imported data to the graph summary”.

To automatically start Task Automation immediately data import:

4. Check “Automate the graph after the data is imported”.

Figure: Import Data Options
3.2 Manual Data Input

Creating a simple graph from an edge list in NodeXL is a two-step process:

1. Enter a list of network edges into the Edges worksheet. The only columns you need to fill in are the columns Vertex 1 and Vertex 2, which are the names of the edge's vertices. In a directed network, Vertex 1 is the “sender” and Vertex 2 is the “receiver”.

Important note: The vertex names are case sensitive. NodeXL considers "John Smith" and "john smith" to be different vertices. Also make sure that there are no empty spaces before or after the vertex name.

2. Show the graph.

To show the graph, do one of the following:

- Click NodeXL > Graph > Show Graph in the Ribbon, or...
- Click the Show Graph button at the top of the Graph Pane, or...
- Click anywhere within the Graph Pane, then press Ctrl+R on the keyboard.

After you show the graph, the Show Graph button text changes to "Refresh Graph." If you add more edges to the Edges worksheet, you have to refresh the graph to get the new edges to appear in the Graph Pane; the Graph Pane does not refresh automatically.

When showing the graph, NodeXL automatically adds all Vertices that are present in the Edges worksheet to the Vertices worksheet.

You can easily change the 7.1 Default visual properties of the graph and customize any edge or vertex properties. But you may want to 5. Group the vertices and calculate 6. Graph metrics first.

Expert tip: Learn about Task automation and choose one of the following data recipes to generate a basic network analysis and graph visualization: NODEXL

Social Network Analysis 01 undirected
Social Network Analysis 02 directed
3.3 Importing Graph Data from Other Programs

NodeXL Pro can import graph data in a number of file formats that are used by other graph programs, so that you can, for example, create a graph in UCINET and then import and view the graph in NodeXL. The supported formats are shown in the table below.

<table>
<thead>
<tr>
<th>File Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCINET Full Matrix DL</td>
<td>This is the only UCINET(^{20}) file format that can be imported into NodeXL. If you have a UCINET file in a different format, such as nodelist1, rankedlist1 or dataset, select the &quot;What if my UCINET file is not in full matrix DL format&quot; link in the Import from UCINET Full Matrix DL File dialog box for instructions. Importing a UCINET full matrix DL file into NodeXL adds an Edge Weight column to the Edges worksheet.</td>
</tr>
<tr>
<td>GraphML</td>
<td>GraphML(^{21}) is an industry-standard network graph file format supported by a number of graph programs. GraphML supports arbitrary vertex and edge attributes. When importing a GraphML file, NodeXL adds a column to the Vertices or Edges worksheet for each attribute in the file.</td>
</tr>
<tr>
<td>Pajek</td>
<td>NodeXL can import files created by the Pajek(^{22}) program. Importing a Pajek file adds an Edge Weight column to the Edges worksheet. Any other edge or vertex attributes in the Pajek file are ignored.</td>
</tr>
<tr>
<td>GEXF</td>
<td>GEXF(^{23}) is an XML-based format that is more expressive than GDF. It allows for the representation of various graph structures, such as directed and undirected graphs, as well as various types of data associated with nodes and edges. GEXF files consist of a header that describes the graph and its properties, a section for node data, and a section for edge data. GEXF also supports the definition of attributes, such as labels, colors, and sizes, for both nodes and edges.</td>
</tr>
<tr>
<td>GDF</td>
<td>GDF(^{24}) is a text-based format that allows for the representation of graph data in a simple tabular format. The file consists of two main sections: vertex data and edge data. Vertex data includes a unique identifier for each node, as well as any additional attributes associated with the vertex. Edge data includes the source and target vertex IDs for each edge, as well as any additional attributes associated with the edge.</td>
</tr>
</tbody>
</table>

\(^{20}\) [https://sites.google.com/site/ucinetsoftware/home](https://sites.google.com/site/ucinetsoftware/home)
\(^{21}\) [http://graphml.graphdrawing.org/](http://graphml.graphdrawing.org/)
\(^{23}\) [https://gexf.net/](https://gexf.net/)
\(^{24}\) [https://gephi.org/users/supported-graph-formats/gdf-format/](https://gephi.org/users/supported-graph-formats/gdf-format/)
3.4 Importing Graph Data from Another Workbook

You can import graph data that is stored in another open Excel workbook in either matrix format or as an edge list. In either case, the other workbook must already be opened in Excel; NodeXL will not open it for you.

To import graph data from an Excel workbook that contains a matrix:

1. Open a new NodeXL Workbook.
2. Open the workbook you would like to import or drag and drop the workbook into the new workbook you have just opened. This step is needed to ensure that both workbooks are opened within the same Excel instance.
3. In the Ribbon, select NodeXL > Data > Import > From Open Matrix Workbook.
4. Follow the instructions in the Import from Open Matrix Workbook dialog box.

Note: This will only work, if the order of the vertices in the first row is exactly like the order of the vertices in the first column.

To import graph data from an Excel workbook that contains an edge list:

1. Open a new NodeXL Workbook.
2. Drag and drop the workbook you would like to import into the new workbook you have just opened. This step is needed to ensure that both workbooks are opened within the same Excel instance.
3. In the Ribbon, select NodeXL > Data > Data Import > From Open Workbook.

Figure: Import from Open Workbook
4. Identify the columns that you would like to import as network edges and set checkmarks under “Is Edge Column”.
5. In the bottom left corner, select which column you want as Vertex 1 and Vertex 2.
6. Also set checkmarks for columns you would like to import and define if a column is supposed to be imported as an edge, edge metadata or vertex metadata. Unchecked columns will be ignored.

**Expert tip:** Learn about [Task automation](#) and choose one of the following data recipes to generate a basic network analysis and graph visualization: **NODEXL**

- Social Network Analysis 01 undirected
- Social Network Analysis 02 directed
3.5 Importing Graph Data from Online Social Networks

Some online social networks provide official APIs (Application Programming Interface) that can be accessed by software tools like NodeXL to import data. NodeXL Pro can currently retrieve data from X (formerly Twitter)\(^\text{25}\), YouTube\(^\text{26}\), Reddit\(^\text{27}\), Flickr\(^\text{28}\), WhatsApp and Wikipedia\(^\text{29}\).

By using a data importer, NodeXL users agree to the Terms of Service and Privacy Policy of the respective platforms. In some cases you have to generate API keys which are usually easily available.

The access to data from online social networks has constantly decreased over the past years. Until 2019, NodeXL was able to import data from Facebook\(^\text{30}\), but the API was closed which is why these importers are marked as inactive in the importer menu. We were hoping that access to the API would be restored, but that is very unlikely. In 2023, Twitter closed its API as a result of massive policy changes. In August 2023, YouTube also removed access to its video recommendation algorithm via its API, but it still gives access to video comments and replies. Reddit has also announced to deprecate its API in 2023, but currently it is still working.

This means we have now entered the post API age\(^\text{31}\). Researchers looking for social media network data now have two options - web scraping or purchasing data from commercial data providers. Web scraping has some advantages, including efficient data collection, automation, real-time information, customization, and integration. However, there are also potential drawbacks, such as legal implications, ethical concerns, data quality issues, IP blocking, and dependency on website stability.

Team NodeXL is adjusting to these changes with new data importers that allow the analysis of data sets downloaded from commercial data providers like Brandwatch\(^\text{32}\), Talkwalker\(^\text{33}\) and Tweet Binder\(^\text{34}\) which provide Excel or csv files that can easily be imported into NodeXL. Team NodeXL will replicate this feature for more data providers as well.

To analyze an online social network and import the results as graph data:

1. In the Ribbon, select NodeXL > Data > Import.
2. Select a data importer from the second-to-last group of items on the Import menu.

You find detailed uptodate information about the current data importers on the official NodeXL data importers website. Currently (2023) available data importers are:

25 https://twitter.com
26 https://www.youtube.com
27 https://www.reddit.com
28 https://www.flickr.com
29 https://www.wikipedia.org
30 https://www.facebook.com
31 https://www.tandfonline.com/doi/abs/10.1080/10584609.2018.1477506
32 https://www.brandwatch.com
33 https://www.talkwalker.com
34 https://www.tweetbinder.com
<table>
<thead>
<tr>
<th>Import from…</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brandwatch… (Beta)</td>
<td>This importer is designed to load X (formerly Twitter) datasets that have been collected via the commercial data provider <a href="https://www.brandwatch.com/">Brandwatch</a>35. The input format is an .xlsx file.</td>
</tr>
<tr>
<td>Flickr Related Tags Network…</td>
<td>The <a href="https://www.smrfoundation.org/nodexl/tutorials/flickr-related-tag-networks/">Flickr related tags network importer</a>36 creates network edges based on co-mentions of image tags around a search term of your choice. The vertices are image tags.</td>
</tr>
<tr>
<td>Flickr User’s Network…</td>
<td>Flickr User’s networks are composed of Flickr users as vertices. Edges can be created based on a user’s contacts or on users commenting on photos.</td>
</tr>
<tr>
<td>MediaWiki Page Network…</td>
<td>The MediaWiki page network importer allows several ways to create networks, from <a href="https://www.smrfoundation.org/nodexl/tutorials/wikipedia-article-article-networks/">article-to-article networks</a>37 to user discussions around any article of your choice. This importer works on all “MediaWikis” which includes the Wikipedia and many others.</td>
</tr>
<tr>
<td>Importer</td>
<td>Description</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Meltwater (Beta)…</td>
<td>This importer is designed to load X (formerly Twitter) datasets that have been collected via the commercial data provider Meltwater(^{38}). The input format is an .xlsx file.</td>
</tr>
<tr>
<td>Reddit Search Network (Beta)…</td>
<td>This importer collects up to the 250 most recent Reddit posts and all comments on them containing a search term of your choice.</td>
</tr>
<tr>
<td>Talkwalker (Beta)…</td>
<td>This importer is designed to load X (formerly Twitter) datasets that have been collected via the commercial data provider Talkwalker(^{39}). The input format is an .xlsx file.</td>
</tr>
<tr>
<td>Tweet Binder (Beta)…</td>
<td>This importer is designed to load X (formerly Twitter) datasets that have been collected via the commercial data provider Tweet Binder(^{40}). The input format is an .xlsx file.</td>
</tr>
<tr>
<td>X (formerly Twitter) Search Network 3.0 (Beta)…</td>
<td>The new X (Twitter) Search Network Data Importer released August 2023(^{41}) in NodeXL Pro collects posts (tweets) via the official X (Twitter) Search using your own account.</td>
</tr>
<tr>
<td>YouTube Channel's Network…</td>
<td>The YouTube Channel’s Network can create User-to-User Channel subscription networks, but it is quite limited because most channels restrict access to this type of data.</td>
</tr>
<tr>
<td>YouTube User Network…</td>
<td>The YouTube User Network Importer(^{42}) you can analyze user networks around a single video or multiple videos. It creates networks based on comments and replies which are posted by users in the comments section below a video.</td>
</tr>
<tr>
<td>YouTube Video Network…</td>
<td>The YouTube Video Network Importer(^{43}) can create YouTube Video-to-Video Networks based on comments of YouTube users who commented on two or more videos.</td>
</tr>
</tbody>
</table>
3.6 Flickr Networks

Flickr is a website that allows you to store, sort, search and share photos or short videos. The popularity of Flickr is mainly due to the ability to manage images using tools that allow authors to tag their photographs, explore and comment on other users’ images. NodeXL Pro offers two data importers to create networks from Flickr data.

3.6.1 Flickr Related Tags Networks

The Flickr related tags network importer creates network edges based on co-mentions of image tags around a search term of your choice. The vertices are image tags.

To analyze a Flickr related tags network:

1. In the Ribbon, select NodeXL > Data > Import > From Flickr Related Tags Network…:

![Flickr Related Tags Importer](image)

2. Apply for a Flickr API key via the link in the center of the importer and enter the key.
3. Enter your search term.
4. Select the network levels to include: **1.0 - 1.5 - 2.0**
5. Optionally add sample image files to the Vertices worksheet.

The above example is based on a 2.0 network. This means the importer will look for all tags that are related to the query – in this case “sociology” – which represents the 1.0 network. After that the importer will collect new lists of tags around each related tag and add network edges between these tags if they are connected to each other, which is the 1.5 network. The 2.0 network contains all related tags that are related to the tags of the 1.0 network. Each network level will show different shapes as seen below:
Expert tip: Learn about 10. Task automation and use the data recipe “Flickr Tag Network.NodeXL0ptions” to generate a network report and map like the ones above.

3.6.2 Flickr User’s Networks

Flickr User’s networks are composed of Flickr users as vertices. Edges can be created based on a user’s contacts or on users commenting on a user’s photos.

To analyze a Flickr User’s Network:

1. In the Ribbon, select NodeXL > Data > Import > From Flickr User’s Network…:

Figure: Flickr Users Network Importer
2. Apply for a Flickr API key via the link in the center of the importer and enter the key.
3. Enter a user name. **Note:** Many user profiles are not open to the API.
4. Select the network levels to include: 1.0 - 1.5 - 2.0
5. Optionally limit the number of users.
6. Click **OK**.

Expert tip: Learn about 10. **Task automation** and use the data recipe "Flickr Users Network.NodeXLOptions" to generate a network report and graph.
3.7 Wikipedia Networks

The MediaWiki page importer can create several types of networks from Wikipedia: User-Article networks, User-User networks and Article-Article networks.

A tutorial on Wikipedia Article-Article networks[^44] is available. You can also find a detailed article in Chapter 14 of the official NodeXL book[^45]. Examples of Wikipedia networks can also be found in the NodeXL Graph Gallery[^46].

The Wikipedia API is pretty slow, but you can get large amounts of data. User-User networks from discussions around articles can be collected rather quickly, while 2.0 Article-Article networks can take several hours or even days.

[^44]: https://www.smrfoundation.org/nodexl/tutorials/wikipedia-article-article-networks/
[^45]: https://doi.org/10.1016/C2018-0-01348-1
[^46]: https://nodexlgraphgallery.org
**Expert tip:** Learn about [10. Task automation](#) and choose one of the following data recipes to generate network reports and graphs:

- Wikipedia Page Network 02 - large.NodeXLOptions
- Wikipedia User Network 01.NodeXLOptions
3.8 X (formerly Twitter) Networks

NodeXL and Twitter network data work very well together. However, since a change in the ownership of the Twitter social media platform, data access has been reduced dramatically. The Twitter platform has been renamed and rebranded as “X” and high fees for access to reduced volumes of data have been imposed. While “X” is showing reduced user engagement and many users are migrating to other platforms, the former “Twitter” service remains a focus for many millions of people discussing a wide range of topics. The NodeXL application has features that have been impacted by changes in the data access policy at “X”. While some features no longer work, NodeXL has new features that enable access to this platform via commercial data providers or the new web scraper importer as described below.


**Expert tip 2:** You can learn about powerful search operators in the official [X (Twitter) documentation](https://developer.twitter.com/en/docs/twitter-api/v1/rules-and-filtering/search-operators).

**Expert tip 3:** Learn about [10. Task automation](#) and choose one of the following data recipes to generate network reports and graphs. The differences are explained in the chapter [10.2 Data Recipes](#).

- Twitter User Network 01 - standard.NodeXLOptions
- Twitter User Network 02 - large.NodeXLOptions
- Twitter User Network 03 - very large.NodeXLOptions
- Twitter User Network 04 - alternative layout.NodeXLOptions

**Expert tip 4:** Export your NodeXL Pro workbook to [8.5 NodeXL Pro INSIGHTS](#). A NodeXL Pro INSIGHTS report, powered by Microsoft Power BI, offers a multifaceted look into your data. From easy-to-read tables and captivating hashtag clouds to intricate scatter plots and image grids, you’ll gain unparalleled insights. Discover patterns, trends, sentiments, and much more across diverse social media dimensions.

### 3.8.1 Import from X (formerly Twitter) Search Network 3.0 (Beta)...

The “X” platform (formerly Twitter) has removed API access to its data for most users. NodeXL Pro has a new importer that uses web scraping to connect and collect posts from the X platform. **“Import from X Search Network 3.0”** is very similar to the prior NodeXL Pro data importer **“Import from Twitter Search API 2.0”**. The main difference is a reduced rate of speed, providing 1000 posts per 15 minute window. The data can be analyzed and published in the same way as data from the earlier data sources.

---


To analyze a X (formerly Twitter) Search Network:

1. In the NodeXL Ribbon, select **NodeXL > Data > Import > From X (formerly Twitter) Search Network 3.0 (Beta)...**

2. Enter a search query of your choice.
3. Set a time frame. You can search the whole historical archive.
4. When you click the “OK” button, you will be presented with a browser window pointed at X (Twitter) login page. After you successfully login, X (Twitter) will send a cookie to your computer. NodeXL uses this cookie to authenticate with X (Twitter). After that the data collection begins. Please note that NodeXL does not handle or store your X (Twitter) username and password. When you type them, you are directly communicating with X (Twitter). Furthermore, the cookie is temporarily stored in memory for the duration of the import session so it can not be used later. NodeXL will then search X (Twitter) for posts (tweets) containing the search term you provided.

**Data Collection:**

The NodeXL Pro X (Twitter) Search importer collects posts (tweets) going backwards in time from the set end date. It stops when either the post (tweet) limit or the start date is reached.

**Limitations:**

X (Twitter) rate limits API calls to their web site. You can get a maximum of 20 tweets per API call. After 50 API calls a 15 minute wait window starts. NodeXL will display a status message when the rate limit is reached and when the collection is scheduled to restart. So roughly, for every 1000 tweets you have to wait 15 min, meaning that the collection of 10,000 tweets will take about 2.5 hours.
Initial tests have also shown that reposts (retweets) are only available from the past 7 days before the date of your query.

Apparently there is no difference in the rate limit for verified (X Blue) or unverified users.

3.8.2 Import from Brandwatch…

Brandwatch⁴⁹ is a commercial data provider with access to the X (formerly Twitter) archive. Brandwatch customers can download data sets as .xlsx files and easily import these into NodeXL Pro to gain network insights from a full social network and content analysis.

To analyze a X (formerly Twitter) data set that was downloaded from Brandwatch:

1. In the NodeXL Ribbon, select NodeXL > Data > Import > From Brandwatch…

2. Click “Browse” to bring up your file explorer, select a Brandwatch Excel data set. Note that one data set may consist of several files with 10k rows each. You need to select all of these files at once.

3. Click OK. NodeXL will systematically comb through all the posts (tweets) and look for connections between users to curate a comprehensive network.

Expert tip: Check out the Expert tips above mentioned in 3.8 X (formerly Twitter) Networks.

⁴⁹ https://www.brandwatch.com
3.8.3 Import from Meltwater…

Meltwater is a commercial data provider with access to the X (formerly Twitter) archive. Brandwatch customers can download data sets as .xlsx files and easily import these into NodeXL Pro to gain network insights from a full social network and content analysis.

To analyze a X (formerly Twitter) data set that was downloaded from Brandwatch:

1. In the NodeXL Ribbon, select NodeXL > Data > Import > From Brandwatch…

![Import from Brandwatch dialog]

2. Click “Browse” to bring up your file explorer, select a Brandwatch data set. Note that one data set may consist of several Excel files with 10k rows each. You need to select all of these files at once.

3. Click OK. NodeXL will systematically comb through all the posts (tweets) and look for connections between users to curate a comprehensive network.

**Expert tip:** Check out the Expert tips above mentioned in 3.8 X (formerly Twitter) Networks.

---

50 [https://www.meltwater.com](https://www.meltwater.com)
3.8.4 Import from Talkwalker…

Talkwalker is a commercial data provider with access to the X (formerly Twitter) archive. Talkwalker customers can download data sets as .xlsx files and easily import these into NodeXL Pro to gain network insights from a full social network and content analysis.

To analyze a X (formerly Twitter) data set that was downloaded from Talkwalker:

1. In the NodeXL Ribbon, select NodeXL > Data > Import > From Talkwalker…

![Import from Talkwalker (Beta) dialog]

2. Click “Browse” to bring up your file explorer, select a Talkwalker data set. Note that one data set may consist of several Excel files with 10k rows each. You need to select all of these files at once.

3. Click OK. NodeXL will systematically comb through all the posts (tweets) and look for connections between users to curate a comprehensive network.

**Expert tip:** Check out the Expert tips above mentioned in 3.8 X (formerly Twitter) Networks.

---

51 [https://www.talkwalker.com](https://www.talkwalker.com)
3.8.5 Import from Tweet Binder…

Tweet Binder\footnote{https://www.tweetbinder.com} is a commercial data provider with access to the X (formerly Twitter) archive. Talkwalker customers can download data sets as .xlsx files and easily import these into NodeXL Pro to gain network insights from a full social network and content analysis.

To analyze a X (formerly Twitter) data set that was downloaded from Tweet Binder:

1. In the NodeXL Ribbon, select NodeXL > Data > Import > From Tweet Binder…

![Image of Import from TweetBinder dialog]

2. Click "Browse" to bring up your file explorer, select a Tweet Binder data set. Note that one data set may consist of several Excel files with 10k rows each. You need to select all of these files at once.

3. Click OK. NodeXL will systematically comb through all the posts (tweets) and look for connections between users to curate a comprehensive network.

Expert tip: Check out the Expert tips above mentioned in 3.8 X (formerly Twitter) Networks.
3.9 YouTube Networks

NodeXL Pro offers several ways to access the official YouTube API (v3). With the NodeXL Pro YouTube data importers you can explore **User-to-User Channel Subscription Networks**, analyze discussions around videos and search terms with **User-to-User Video Comment Networks** and create **Video-to-Video Networks** based on comments of YouTube users who commented on two or more videos.

The NodeXL Pro YouTube data importers have an integrated API key with 100k units per day which can be used by any NodeXL user, but these units are consumed very quickly. That is why quota management is very important. So before getting started we strongly recommend that you create your own YouTube API keys for your project. [Here is a guide](https://www.smrfoundation.org/2020/04/11/nodexl-pro-youtube-importer-update-unlock-the-api-with-your-own-key/) on how to quickly receive up to 10 API keys with a daily limit of 10,000 units per key. You can [apply for an upgrade](https://support.google.com/youtube/contact/yt_api_form?hl=en) to 100,000 units, but need to go through a review process.

**Important note:** Due to YouTube’s Terms of Service it is not possible to visualize the downloaded data as a network in the graph pane any more. Contact [info@smrfoundation.org](mailto:info@smrfoundation.org) for more information.

3.9.1 YouTube Channel’s Network

The YouTube Channel’s Network importer was designed to create User-to-User Channel subscription networks, but nowadays it is very limited because the vast majority of channels restrict access to this type of data for privacy reasons.

3.9.2 YouTube User Network

The YouTube User Network Importer creates networks based on comments and replies which are posted by users in the comments section below a video. With the NodeXL Pro YouTube User network importer you can analyze user networks around a single video or multiple videos. Check out the tutorial [Analyzing YouTube User Networks](https://www.smrfoundation.org/nodexl/tutorials/analyzing-youtube-user-networks) for details.

**To analyze a YouTube User network:**

1. Select **NodeXL > Data > Import > …From YouTube User Network**: 

---


54 [https://support.google.com/youtube/contact/yt_api_form?hl=en](https://support.google.com/youtube/contact/yt_api_form?hl=en)

2. Enter a search term into the box at the top or a list of video IDs into the box on the left.
3. We recommend checking both options in the “Add an edge for each” section.
4. Depending on your research goal, the number of videos, comments and replies should be limited in the Options section. Otherwise you may run out of API quota during the data collection. See 3.9 YouTube Networks for details.
5. Click OK.

**Expert tip:** Learn about 10. Task automation and use one of the the data recipes to generate a network report and map:

- YouTube User Network 01 - standard.NodeXLOptions
- YouTube User Network 02 - alternative layout.NodeXLOptions
- YouTube User Network 03 - large.NodeXLOptions
3.9.3 YouTube Video Network

The YouTube Video Network importer can create YouTube Video-to-Video Networks based on comments of YouTube users who commented on two or more different videos.

**Important note:** In a previous version of this importer it was possible to create Video-to-Video networks based on the YouTube recommendation algorithm, but this option was removed from the YouTube API in August 2023.

**To create a YouTube Video Network:**

1. **Select** [NodeXL > Data > Import > …From YouTube User Network]:

   ![Figure: YouTube Video Network Importer](image)

2. Enter a search term into the box at the top or a list of video IDs into the box on the left.
3. In the “Add an edge for each” section select “Pair of videos commented on by the same user”.
4. In the **Options** section, set the limits of the importer to about 100 videos limited to a maximum of 100 comments for each video and a maximum 100 replies to each comment. This setup is designed for one API key with 10k units. See [3.9 YouTube Networks](#) for details.
5. Check the boxes for YouTube TOS and Privacy Policy.
**Expert tip:** Learn about [10. Task automation](#) and use one of the data recipes to generate a network report and map:

- YouTube Video Network 01 - description analysis.NodeXLOptions
- YouTube Video Network 02 - tag analysis.NodeXLOptions
- YouTube Video Network 03 - comment analysis.NodeXLOptions
3.10 Reddit Networks

Reddit is a social network where users can submit content, such as text posts, links, images, and videos, to various thematic forums known as "subreddits." Each subreddit has its unique focus and community, ranging from news, politics, and technology to hobbies, entertainment, and virtually any niche interest imaginable.

Reddit is composed of numerous subreddits, each prefixed with "r/". For example, r/science is dedicated to discussing scientific topics, while r/movies focuses on films. Content can be upvoted or downvoted by users. Posts that receive many upvotes rise to the top of their respective subreddits and, if they gain enough traction, can also make it to Reddit's front page. Each post on Reddit has a comment section where users can discuss the content, ask questions, or share thoughts. Like posts, comments can also be upvoted or downvoted. Further, Reddit users can also reply to comments.

The NodeXL Pro Reddit Search Network importer was added to NodeXL Pro in May 2023. The only requirement is to have a Reddit user account which is needed to download the data. You can then create user networks around any search term in all of Reddit, or you can focus on a search term within a subreddit.

When this importer was first introduced, it was able to connect to the API of Pushshift.io which allowed the import of large historical datasets. This service was deprecated in June 2023, and it is not clear if it will ever come back.

The current importer can only connect to the official Reddit API, which allows the collection of a maximum of the 250 most recent Reddit posts on any given search term including all comments and replies to these posts which may lead to networks with a few thousand vertices.

To analyze a Reddit Search Network in NodeXL Pro:

1. Select NodeXL > Data > Import > …Reddit Search Network (Beta):

---

56 https://www.reddit.com
57 https://pushshift.io
2. Enter a search term into the box at the top.
3. By default, the importer collects the “Basic network” of posts, comments and replies. If you would like to add users that are tagged in the posts and replies, select the option “Basic network plus tags”.
4. Click OK.

**Expert tip:** Learn about 10. Task automation and use one of the the data recipes to generate a network report and map:

- Reddit User Network 01 - standard.NodeXLOptions
- Reddit User Network 02 - paths.NodeXLOptions
3.11 Import from WhatsApp

WhatsApp is a widely used personal messenger application owned by Meta. WhatsApp allows people to form chats with one or more people. WhatsApp chats can grow to include large numbers of participants, effectively becoming semi-public group discussions. In some cases, these discussions can be the focus of collective action with civic impact.

In contrast with other forms of social media platforms, WhatsApp chats are more under the control of their authors and members, since people can be selectively invited in and removed by the chat owner. WhatsApp chats may be less discoverable than discussions on other social media platforms that have global search features, but this absence may be a feature as much as a bug, providing a social boundary that is crossed only through invitations from existing members.

**Getting started**

Depending on the WhatsApp version you are using, you are able to export any group chat, by tapping on the chat name and then accessing the Group Info page. Scroll to the bottom of this page to get to the option “Export Chat”. In other cases, you are able to tap on the three dots at the top right of a chat and then select “More - Export Chat”. After exporting the chat will be saved as a .txt file which will be read by NodeXL in the next step.

**To analyze a WhatsApp chat network in NodeXL Pro:**

5. Select NodeXL > Data > Import > …From WhatsApp (Beta):

![WhatsApp User Network Importer](image)

Figure: WhatsApp User Network Importer

6. Click the **Browse** button and locate the WhatsApp text file on your local hard drive.

---

58 https://en.wikipedia.org/wiki/WhatsApp
59 https://en.wikipedia.org/wiki/Meta_Platforms
60 https://theconnector.substack.com/p/whatsapp-and-the-micro-organizing
7. Select edge creation options in the “Add an edge for each” section.
8. Click OK.

**Edge creation:**
NodeXL’s new WhatsApp chat data importer transforms your chat exports into networks, showcasing the relationships and interactions based on

1. **Consecutive Comments**: Comments that follow each other tell a story. By charting these consecutive remarks, the importer maps the flow of conversations, revealing who interacts after whom and highlighting the central figures in a discussion.

2. **Mentions**: Mentions (@username) direct comments towards specific individuals. Tracking these mentions unveils the chat’s nexus points, the people who are most often engaged, and the topics that stir the most interaction.

**Limitations:**
- The data does not include reactions to messages such as emojis.
- The Mentions are always based on the mentioned user’s phone number, not the actual name that is saved in a phone’s address book. This leads to multiple vertex names for one user.
- Different date and time formats across the globe may lead to problems during data import. Please send your bug report including the txt file to info@smrfoundation.org.

**Expert tip:** Learn about [10. Task automation](#) and use one of the the data recipes to generate a network report and map:

```text
WhatsApp User Network.NodeXLOptions
```
4. Data Preparation

Preparing a network data set for analysis includes two basic steps. First, it is important to determine whether a network is directed or undirected. Second, determining the strength of the network connections via Edge Weight is very helpful for further analysis.

4.1 Directed vs. Undirected Graphs

NodeXL can handle both directed and undirected graphs.

Directed graphs are used to model asymmetric relationships. Each edge has a start point and an endpoint. In NodeXL the start point is the vertex specified in the Vertex 1 column of the Edges worksheet and the endpoint is the vertex specified in the Vertex 2 column. Arrowheads are shown at the Vertex 2 end of each edge in the Graph Pane.

Undirected graphs are used to model symmetric relationships. The edges in an undirected graph do not have start points or endpoints, and arrowheads are not shown in the Graph Pane.

Apart from determining whether arrowheads are shown, the directedness of a graph affects the way some of NodeXL's graph metrics are calculated.

To specify whether a graph is directed or undirected:

1. In the Excel Ribbon, set NodeXL > Graph > Graph Type to Directed or Undirected.

Note: When importing data with one of the built-in data importers, the Graph Type is automatically set by NodeXL.
4.2 Edge Weight: Counting and Merging Duplicate Edges

Edge Weight is a significant indicator for the connectedness of entities in a network. It is also one of the most common attributes used to visualize edges in a network graph, e.g. the stronger the edge weight, the broader the edge in the graph (see 7.4 Automatically Calculating Visual Properties for All Edges, Vertices or Groups).

![NodeXL Edge Weights](image)

Figure: NodeXL Edge Weights

If your graph has duplicate edges, you can tell NodeXL to count them, merge them, or both. By default, NodeXL uses the two vertex columns on the Edges worksheet to determine whether two edges are duplicates, but you can specify a third column to use as well.

To determine the Edge Weight:

1. In the Excel Ribbon, select **NodeXL > Data > Prepare Data > Count and Merge Duplicate Edges**.

![Count and Merge Duplicate Edges](image)

Figure: Count and Merge Duplicate Edges

2. Check “Count duplicate edges and insert the counts into an Edge Weight column”.
3. Optionally check “Merge duplicate edges”. **Important note (!):** This step will delete edges that cannot be retrieved afterwards! In some cases this can be helpful.
We recommend saving a data backup file before using this feature, if you are unsure how to use this.

Here are some things to know about counting and merging duplicate edges:

If you tell NodeXL to count duplicate edges, it adds an Edge Weight column to the Edges worksheet if the column isn't already there, then fills in the column. If the worksheet contains these two rows, for example:

<table>
<thead>
<tr>
<th>Vertex 1</th>
<th>Vertex 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>Bill</td>
</tr>
<tr>
<td>John</td>
<td>Bill</td>
</tr>
</tbody>
</table>

...then the Edge Weight for each of the two rows will be set to 2, because there are two instances of each row:

<table>
<thead>
<tr>
<th>Vertex 1</th>
<th>Vertex 2</th>
<th>Edge Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>Bill</td>
<td>2</td>
</tr>
<tr>
<td>John</td>
<td>Bill</td>
<td>2</td>
</tr>
</tbody>
</table>

If the worksheet already has an Edge Weight column, NodeXL uses any existing edge weights when it counts duplicate edges. If the worksheet contains these two rows, for example:

<table>
<thead>
<tr>
<th>Vertex 1</th>
<th>Vertex 2</th>
<th>Edge Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mary</td>
<td>Joe</td>
<td>3</td>
</tr>
<tr>
<td>Mary</td>
<td>Joe</td>
<td>4</td>
</tr>
</tbody>
</table>

...then the Edge Weight for each of the two duplicate rows will be set to the sum of their edge weights:

<table>
<thead>
<tr>
<th>Vertex 1</th>
<th>Vertex 2</th>
<th>Edge Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mary</td>
<td>Joe</td>
<td>7</td>
</tr>
<tr>
<td>Mary</td>
<td>Joe</td>
<td>7</td>
</tr>
</tbody>
</table>

NodeXL takes the directedness of the graph into account when it checks for duplicate edges. In a directed graph, a "John,Bill" row and a "Bill,John" row can't be duplicates. In an undirected graph, "John,Bill" and "Bill,John" can be duplicates.
If you specify a third column to use on the Edges worksheet, two edges have to have the same values in that column before they can be considered duplicates. For example, if you specify the Relationship column when counting duplicate edges in this worksheet:

<table>
<thead>
<tr>
<th>Vertex 1</th>
<th>Vertex 2</th>
<th>Relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dave</td>
<td>Joe</td>
<td>Mentions</td>
</tr>
<tr>
<td>Dave</td>
<td>Joe</td>
<td>Follows</td>
</tr>
<tr>
<td>Dave</td>
<td>Joe</td>
<td>Mentions</td>
</tr>
</tbody>
</table>

...then the results will look like this:

<table>
<thead>
<tr>
<th>Vertex 1</th>
<th>Vertex 2</th>
<th>Relationship</th>
<th>Edge Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dave</td>
<td>Joe</td>
<td>Mentions</td>
<td>2</td>
</tr>
<tr>
<td>Dave</td>
<td>Joe</td>
<td>Follows</td>
<td>1</td>
</tr>
<tr>
<td>Dave</td>
<td>Joe</td>
<td>Mentions</td>
<td>2</td>
</tr>
</tbody>
</table>

If you tell NodeXL to merge the duplicate edges, all but the first edge in each set of duplicates is deleted from the Edges worksheet. For example, merging duplicate edges in this worksheet:

<table>
<thead>
<tr>
<th>Vertex 1</th>
<th>Vertex 2</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jack</td>
<td>Jill</td>
<td>Blue</td>
</tr>
<tr>
<td>Sally</td>
<td>Ralph</td>
<td>Orange</td>
</tr>
<tr>
<td>Jack</td>
<td>Jill</td>
<td>Green</td>
</tr>
<tr>
<td>Sally</td>
<td>Ralph</td>
<td>Orange</td>
</tr>
</tbody>
</table>

...results in this:

<table>
<thead>
<tr>
<th>Vertex 1</th>
<th>Vertex 2</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jack</td>
<td>Jill</td>
<td>Blue</td>
</tr>
<tr>
<td>Sally</td>
<td>Ralph</td>
<td>Orange</td>
</tr>
</tbody>
</table>

Note that the edge attributes for the first edge in a set of duplicate edges become the edge attributes for the merged edge. Thus, the two "Jack,Jill" edges get merged into a Blue edge, even though the original edges had different colors. The Green color of the second edge is lost, along with any other attributes the second edge had.
NodeXL automatically counts and merges duplicate edges when you export the graph. A third column cannot be specified during an export. If you want more control over how duplicate edges are handled, use “Count and Merge Duplicate Edges” before exporting the graph.
5. Working with Groups

A group in NodeXL is a set of related vertices. The vertices may be structurally related, meaning they are connected to one another in a certain way, or related by a common attribute - all vertices that represent people in Canada, for example.

Although groups are typically created in such a way that each vertex is in a group, this doesn't have to be the case. You can have a graph where some vertices are in group A, some are in group B, and some are in no group at all, for example. However, a vertex can never be in more than one group.

You can tell NodeXL to automatically create a set of groups using one of several techniques presented below with “Group by Cluster” as the most common approach in social network analysis. You can also create groups by manually selecting vertices and giving them a group name.

In either case, creating groups inserts the new groups into the Groups worksheet in the NodeXL workbook. The groups do not automatically appear in the Graph Pane, however. To see the groups in the Graph Pane, you must show the graph after creating your groups. During the calculation of the groups, new columns are created on the far right of the Edges and Vertices worksheets which show the affiliation of a vertex with a group. On the Edges spreadsheet you will find the columns “Vertex 1 Group” and “Vertex 2 Group”, and on the Vertices worksheet the column “Vertex Group” is created.

See 5.7 How Groups are Shown in the Graph Pane for information on how NodeXL shows groups.
5.1 Creating Groups by Vertex Attribute

You can tell NodeXL to create groups using the values in a column on the Vertices worksheet. For example, if your vertices represent people and you have a "Country" column on the Vertices worksheet, NodeXL can create a group for each country. Or if you have a "Salary" column that contains numbers, NodeXL can create a group for each salary range that you specify. Groups can also be created using a column that contains dates, times, or dates with times.

To remove any existing groups and create new groups based on a vertex attribute:

1. In the Excel Ribbon, select NodeXL > Analysis > Groups > Group by Vertex Attribute.
2. In the Group by Vertex Attribute dialog box, select the column that contains the vertex attribute values.
3. Specify whether the vertex attribute values are categories (such as "Country" in the above example), numbers, dates, times, or dates with times.
4. If the vertex attribute values are numbers, dates, times, or dates with times, specify the group value ranges.
5. Click OK.
6. Show/Refresh the graph to see the groups.

Note that some vertices might be left "ungrouped." If a cell in the vertex attribute column is empty, or its value doesn't fall within one of the ranges you specify, the vertex will not be put into a group.

5.2 Creating Groups by Connected Component

A connected component is a set of vertices (and their incident edges) that are connected to each other but not to the rest of the graph.

To remove any existing groups and create new groups based on each of the graph's connected components:

1. In the Excel Ribbon, select NodeXL > Analysis > Groups > Group by Connected Component.
2. Show/Refresh the graph to see the groups.

Note that an isolate is a one-vertex connected component and is thus put into its own group.
5.3 Creating Groups by Cluster

A cluster is a group of vertices that are densely connected to one another but sparsely connected to the rest of the graph. This approach is most commonly used in social network analysis. NodeXL provides three cluster algorithms:

<table>
<thead>
<tr>
<th>Cluster algorithm</th>
<th>Description</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clauset-Newman-Moore</td>
<td>This cluster algorithm is a statistical model that optimizes the likelihood of observing the network under a stochastic block model, where nodes are grouped into clusters based on their connectivity patterns. This method is efficient and can handle large networks.</td>
<td><a href="https://arxiv.org/abs/cond-mat/0408187">https://arxiv.org/abs/cond-mat/0408187</a></td>
</tr>
<tr>
<td>Wakita-Tsurumi</td>
<td>Wakita-Tsurumi is a method that combines hierarchical clustering and edge-betweenness centrality. It identifies clusters by progressively removing edges with the highest betweenness centrality and updating the hierarchical clustering tree. This method is efficient and can handle very large networks, but it may not be able to detect clusters with non-hierarchical structures.</td>
<td><a href="https://arxiv.org/abs/cs/0702048">https://arxiv.org/abs/cs/0702048</a></td>
</tr>
<tr>
<td>Girvan-Newman</td>
<td>Girvan-Newman is a hierarchical clustering method that involves iteratively removing edges with the highest betweenness centrality until the network is partitioned into separate components. This method is computationally expensive, but it can be useful for identifying clusters at multiple levels of granularity.</td>
<td><a href="https://www.pnas.org/doi/10.1073/pnas.122653799">https://www.pnas.org/doi/10.1073/pnas.122653799</a></td>
</tr>
</tbody>
</table>

The speed and effectiveness of the algorithms vary with the particular graph. You may want to experiment with the algorithms to determine which you prefer.

To remove any existing groups, create a group for each of the graph's clusters, and put every vertex in one of the new groups:

1. In the Excel Ribbon, select NodeXL > Analysis > Groups > Group by Cluster.
2. In the **Group by Cluster dialog box**, select a cluster algorithm to use.
3. Specify what you want to do with vertices that have no neighbors. Checking “Put all neighborless vertices into one group” is recommended.
4. **Show/Refresh** the graph to see the groups.
5.4 Creating Groups by Motif

Motif simplification is a new way to reduce the visual complexity of your graph. With motif simplification, common repeating network motifs are replaced with easily understandable glyphs that require less space, are easier to understand, and reveal hidden relationships.

NodeXL can group the graph's vertices using three kinds of motifs:

- Fan motifs help get rid of the large fans of singly-connected vertices around the edge of your graph.
- Connector motifs focus on redundant connections between a pair of anchor vertices.
- Cliques allow you to simplify completely connected groups of vertices.

For more details, please see:

To remove any existing groups, create groups for motifs found within the graph, and put the motifs' vertices in the new groups:

1. In the Excel Ribbon, select **NodeXL > Analysis > Groups > Group by Motif**.

   ![Group by Motif](image)

   Figure: Group by Motif

2. In the **Group by Motif dialog box**, select which motifs you want NodeXL to find within the graph.
3. **Show/Refresh** the graph to see the groups.
5.5 Manually Creating Groups

NodeXL's automatic group creation features are fast and convenient, but you can also create groups manually and specify exactly which vertices go into which groups.

To create groups manually:

1. Select one or more vertices that should be in a group.
2. In the Excel Ribbon, select NodeXL > Analysis > Groups > Add Selected Vertices to Group.
3. In the Add Selected Vertices to Group dialog box, enter a group name to create a new group or select the name of an existing group.
4. Click OK.
5. Repeat steps 1 through 4 for each group.
6. Show/Refresh the graph to see the groups.

Because a vertex can never exist in more than one group, NodeXL automatically removes the selected vertices from any groups they belong to before it adds them to the group that you specify.

To remove one or more vertices from their groups without adding them to another group:

1. Select the vertices.
2. In the Excel Ribbon, select NodeXL > Analysis > Groups > Remove Selected Vertices from Groups.
5.6 Understanding the Group Worksheets

When you create groups using one of the automatic or manual techniques discussed in the previous chapter, NodeXL inserts the group information into the workbook's Groups and Group Vertices worksheets. One row gets inserted into the Groups worksheet for each group, and one row gets inserted into the Group Vertices worksheet for each grouped vertex.

You can perform the following tasks using the Groups worksheet:
- Select one or more groups.
- Edit the color and shape combinations that NodeXL assigns to the groups. (See 5.7 How Groups Are Shown in the Graph Pane.)
- Collapse or expand a group.
- Hide or skip a group.
- Label a group.

You can perform the following task using the Group Vertices worksheet:
- See which vertices are in which groups.

Important note: You should let NodeXL manage the rows in both worksheets. Manually adding or removing rows in the Groups and Group Vertices worksheets is not necessary and is not recommended.
5.7 How Groups Are Shown in the Graph Pane

5.7.1 Color and Shape

NodeXL assigns a color and shape to each group that you create using one of the techniques discussed in 5. Working with Groups. When you show the graph, all the vertices in the group will have that color and shape - group A's vertices might be blue disks, for example, and group B's vertices might be orange spheres. That means the Color and Shape columns on the Vertices worksheet get ignored. You can, however, tell NodeXL to not ignore those columns. See 5.7.2 Group Options: Changing How Vertex Colors and Shapes are Specified.

Most people use the color and shape combinations that NodeXL assigns to the groups, but you can edit the colors and shapes if you are unhappy with them.

Important note: If you create groups, edit the group colors and shapes, and then create groups again, your edited colors and shapes will get overwritten and you will have to edit them again. Therefore you should create all your groups first, then edit the group colors and shapes.

To edit the colors and shapes that NodeXL assigns to groups:

1. Select the Groups worksheet.
2. Edit the Vertex Color and Vertex Shape columns on the Groups worksheet.
3. If these columns are not visible, show the Visual Properties column group via NodeXL > Show/Hide > Workbook Columns.
4. Show/Refresh the graph to see your changes.

5.7.2 Group Options: Changing How Vertex Colors and Shapes are Specified

As explained above, NodeXL assigns a color and shape to each group that you create. The colors and shapes are stored in the Groups worksheet, and the colors and shapes specified in the Vertices worksheet are ignored.

Advanced users may want to use the colors or shapes from the Vertices worksheet instead. For example, you may want to combine the group colors that NodeXL assigns with the individual vertex shapes that you specified on the Vertices worksheet. This can be particularly useful when the individual vertices have a Shape of Image or Label.

To Change how vertex colors and shapes are specified:

1. In the Excel Ribbon, select NodeXL > Analysis > Groups > Group Options.
2. Select options for "What colors should be used for the groups' vertices?" and "What shapes should be used for the groups' vertices?"
3. **Show/Refresh** the graph.

5.7.3 Group Layout

In addition to using vertex colors and shapes to distinguish groups from one another, NodeXL can lay out each group within its own box in the Graph Pane.

**To lay out each group within its own box in the Graph Pane:**

1. In the Excel Ribbon, select **NodeXL > Graph > Layout > Layout Options**.
2. In the **Layout Options dialog box**, check "Lay out each of the graph's groups in its own box and sort the boxes by group size."
3. Select one of the three **box layout algorithms**: Treemap - Packed rectangles - Force directed
4. Optionally set the width of the box outlines; select whether to show, hide or combine intergroup edges (see below); and select whether you want NodeXL to use the Grid layout for groups that don't have many edges.
5. Click **OK**.
6. **Show/Refresh** the graph or lay it out again.

You find more information about the **7.5.2 Group Layout Options** below.
5.7.4 Skipping All Groups

You can tell NodeXL to temporarily skip all the groups you've created without actually removing the groups from the workbook.

To tell NodeXL to temporarily skip all of the graph's groups:

1. In the Excel Ribbon, select NodeXL > Analysis > Groups > Group Options.
2. Check "Skip groups -- don't show them in the Graph Pane."
3. Show/Refresh the graph.

When you tell NodeXL to skip all the groups, it completely ignores everything on the Groups and Group Vertices worksheets when the graph is shown. It also starts paying attention to the Color and Shape columns on the Vertices worksheet again, so the graph's vertices take on any colors and shapes you previously assigned to them.

You can also skip individual groups. See 5.10 Hiding and Skipping Groups.
5.8 Selecting Groups

You can select one or more of the groups that you have created. When you select a group, all of the group's vertices and their incident edges appear selected in the Graph Pane, and the corresponding rows get selected in the Vertices and Edges worksheets.

To select one or more groups in the workbook:

1. Select the Groups worksheet.
2. To select just one group, click on the row number all the way at the left.
3. To select multiple adjacent rows, click on a row number, hold down the Shift key, and click on another row above or below the previously selected vertex.
4. To select multiple non-adjacent rows, click on a row number, then hold down the Ctrl key, and click on any other row.

To select one or more groups in the Graph Pane:

1. Select one or more vertices in the Graph Pane.
2. In the Excel Ribbon, select NodeXL > Analysis > Groups > Select Groups Containing Selected Vertices.

To select all groups:

1. In the Excel Ribbon, select NodeXL > Analysis > Groups > Select All Groups.

To deselect all groups:

1. Click an empty area in the Graph Pane.
5.9 Collapsing and Expanding Groups

You can collapse one or more groups in the Graph Pane. When a group is collapsed, all of its vertices are replaced with a single disk. The disk's color is the same color as the group's vertices, and the plus sign distinguishes the group from normal vertices. The size of the disk is proportional to the number of vertices within the group. If you have assigned a label to the group, the label appears next to the disk.

You can select and move a collapsed group by clicking its disk in the Graph Pane, just as you would a normal vertex.

To collapse one or more groups using the Ribbon:

1. Select the groups you want to collapse.
2. In the Excel Ribbon, select NodeXL > Analysis > Groups > Collapse Selected Groups.

To collapse one or more groups using the workbook:

1. Select the Groups worksheet.
2. For each group you want to collapse, set the group's Collapsed? cell in the Visual Properties column group to Yes.
3. If these columns are not visible, show the Visual Properties column group via ribbon group NodeXL > Show/Hide > Workbook Columns.
4. Show/Refresh the graph.

To expand all groups:

1. In the Excel Ribbon, select NodeXL > Analysis > Groups > Expand All Groups.

To expand one or more groups using the Ribbon:

1. Select the groups you want to expand.
2. In the Excel Ribbon, select NodeXL > Analysis > Groups > Expand Selected Groups.

To expand one or more groups using the workbook:

1. Select the Groups worksheet.
2. For each group you want to collapse, set the group's Collapsed? cell in the Visual Properties column group to No.
3. If these columns are not visible, show the Visual Properties column group via ribbon group NodeXL > Show/Hide > Workbook Columns.
4. Show/Refresh the graph.
5.10 Hiding and Skipping Groups

You can hide or skip one or more groups in the Graph Pane. When a group is hidden, space is reserved in the graph pane for the group's vertices, but the vertices and their edges are hidden. When a group is skipped, its vertices and edges are completely ignored when the graph is shown. No space is reserved for the vertices in a skipped group.

Hiding and skipping groups can be an effective way of reducing graph clutter when you want to focus on certain groups while ignoring others.

**To hide or skip individual groups:**

1. Select the Groups worksheet.
2. For each group you want to hide or skip, set the group's **Visibility** cell in the **Visual Properties column group** to **Hide** or **Skip**. If the cell is empty, it is treated as **Show**.
3. If these columns are not visible, show the **Visual Properties column group** via ribbon group **NodeXL > Show/Hide > Workbook Columns**.
4. **Show/Refresh** the graph.

**Expert tip:** You can also Hide and Skip edges and vertices in the **Visibility columns** of the respective worksheets.

5.11 Removing Groups

You can remove any of the groups you've created at any time. This does not remove the groups' vertices from the graph; it just "ungroups" the vertices.

**To remove one or more groups:**

1. Select the groups you want to remove.
2. In the Excel Ribbon, select **NodeXL > Analysis > Groups > Remove Selected Groups**.
3. **Show/Refresh** the graph.

**To remove all groups:**

1. In the Excel Ribbon, select **NodeXL > Analysis > Groups > Remove All Groups**.
6. Calculating Graph Metrics

NodeXL offers a number of features to analyze a network on multiple levels. These Graph Metrics contain two fundamental categories: Network analysis and content analysis. Combining these features can help to generate deep insights from a network data set.

Network analysis includes a variety of metrics like overall graph metrics, vertex metrics, and group metrics. The calculated graph metrics get inserted into the Vertices, Groups, Group Edges or Overall Metrics worksheets, depending on the metric.

Metrics to analyze the contents are available if the data set contains metadata columns on the edges and/or vertices worksheets. For example, if there is a date column, you can run time series analysis, and if you have a column that contains text, you can perform text and sentiment analysis. New worksheets are created for most of these features.

To calculate Graph Metrics:

1. In the Excel Ribbon, select **NodeXL > Analysis > Graph Metrics**.

2. In the **Graph Metrics** dialog box, check the graph metrics you want to calculate. These metrics are explained in detail below.
3. Some of the metrics include additional settings that can be accessed via the **Options** button on the right.
4. Select **Calculate Metrics**.
Here are some things to know about Graph Metrics:

- If you add or remove vertices or edges in the workbook, you have to calculate graph metrics again. The calculated graph metrics do not get updated automatically.
- You do not have to show the graph before calculating graph metrics. NodeXL will calculate them even if the Graph Pane is empty.
- You can calculate one metric at a time or select multiple metrics to automatically calculate them one after another.
- Group Metrics are only available if you create groups before.

To show/hide the graph metric columns on the Vertices, Groups and Group Edges worksheets:

1. In the Excel Ribbon, select NodeXL > Show/Hide > Workbook Columns.
2. Un-/check Graph Metrics.
6.1 Overall Graph Metrics

The following overall metrics get inserted into the Overall Metrics worksheet:

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertices</td>
<td>A vertex is an element of a network. The vertices count is the overall number of people or things in the network.</td>
</tr>
<tr>
<td>Unique Edges</td>
<td>An edge is a connection between two vertices. The &quot;unique&quot; edges count is the number of connections where multiple connections between A and B are counted only once.</td>
</tr>
<tr>
<td>Edges With Duplicates</td>
<td>An edge is a connection between two vertices. The &quot;duplicate&quot; edges count is the total number of multiple connections between two vertices.</td>
</tr>
<tr>
<td>Total Edges</td>
<td>The &quot;total&quot; edges count is the total number of connections where multiple connections between A and B are all counted.</td>
</tr>
<tr>
<td>Number of Edge Types</td>
<td>The number of different edge types in the graph.</td>
</tr>
<tr>
<td>Number of Edges by Type</td>
<td>A list of top 10 edge types and their respective edge count.</td>
</tr>
<tr>
<td>Self-Loops</td>
<td>An edge that starts and ends in the same vertex is a self-loop. These are also called isolates. The Self-Loops count is the overall number of edges that connect a vertex to itself.</td>
</tr>
<tr>
<td>Reciprocated Vertex Pair Ratio</td>
<td>When two vertices both link to each other their connection is &quot;reciprocated&quot;. In a directed graph, this is the number of vertex pairs that have edges in both directions divided by the number of vertex pairs that are connected by any edge. Duplicate edges and self-loops are ignored. In an undirected graph, this is undefined and is not calculated.</td>
</tr>
<tr>
<td>Reciprocated Edge Ratio</td>
<td>In a directed graph, this is the number of edges that are reciprocated divided by the total number of edges. Duplicate edges and self-loops are ignored. In an undirected graph, this is undefined and is not calculated.</td>
</tr>
<tr>
<td>Connected Components</td>
<td>A group of vertices that are all connected is a component. This is the number of separate sets of connected vertices.</td>
</tr>
<tr>
<td>Single-Vertex Connected Components</td>
<td>A vertex that has zero connections is &quot;isolated&quot; or an &quot;island&quot;. This is the count of vertices that have zero connections.</td>
</tr>
<tr>
<td>Maximum Vertices in a Connected Component</td>
<td>A connected component is composed of a number of vertices. This is the count of vertices in the largest connected component.</td>
</tr>
<tr>
<td>Maximum Edges in a Connected Component</td>
<td>A connected component is composed of a number of edges. This is the count of total edges in the largest connected component.</td>
</tr>
<tr>
<td>Maximum Geodesic Distance (Diameter)</td>
<td>A geodesic is a chain or path composed of edges that link two vertices, potentially through intermediate vertices. A &quot;shortest path&quot; is the minimum number of connections needed to link two vertices. The &quot;longest&quot; &quot;shortest path&quot; is the &quot;maximum geodesic distance&quot;.</td>
</tr>
</tbody>
</table>
| Average Geodesic Distance                   | A geodesic is a chain or path composed of edges that link two vertices, potentially through intermediate vertices. The average length of these paths is the "average geodesic  

70
Distance

**Graph Density**

Density is the ratio of the number edges among a group of vertices over the total possible number if everyone was connected to everyone. Duplicate edges and self-loops are ignored. A high graph density means that most people are connected to many others. A low graph density means that most people are not connected to many others.

**Modularity**

When the graph has groups, this is a measure of the "quality" of the grouping. Graphs with high modularity have dense connections among the vertices within the same group but sparse connections among vertices in different groups. When the graph does not have groups, this is undefined.

**To calculate Overall Graph Metrics**

1. In the Excel Ribbon, select **NodeXL > Analysis > Graph Metrics**.
2. Check mark the first row **Overall Graph Metrics**.
3. Optionally select the **Options...** button on the right and select the column that determines the edge type (e.g. Relationship).

**To aggregate the overall metrics from multiple NodeXL workbooks into a new Excel workbook:**

1. In the Excel Ribbon, select **NodeXL > Analysis > Graph Metrics > Aggregate Overall Metrics**. This reads the overall metrics from each NodeXL workbook in a specified folder (or calculates them if necessary) and stores the aggregated metrics in a new Excel workbook.
6.2 Vertex Metrics

You can specify which of these metrics you would like to be calculated, but we suggest that you check all of these rows in the **Graph Metrics dialog box**.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertex degree (undirected graphs only)</td>
<td>In an undirected graph, a vertex's degree is the number of vertices incident to the vertex. In a directed graph, degree is undefined and is not calculated. A self-loop in an undirected graph is counted twice when a vertex's degree is calculated.</td>
</tr>
<tr>
<td>Vertex in-degree (directed graphs only)</td>
<td>In a directed graph, a vertex's in-degree is the number of incoming edges incident to the vertex. In an undirected graph, in-degree is undefined and is not calculated. A self-loop in a directed graph is counted once as an incoming edge (in-degree) and once as an outgoing edge (out-degree).</td>
</tr>
<tr>
<td>Vertex out-degree (directed graphs only)</td>
<td>In a directed graph, a vertex's out-degree is the number of outgoing edges incident to the vertex. In an undirected graph, out-degree is undefined and is not calculated. A self-loop in a directed graph is counted once as an outgoing edge (out-degree) and once as an incoming edge (in-degree).</td>
</tr>
</tbody>
</table>
| Vertex betweenness and closeness centralities | A vertex that occurs on many shortest paths between other vertices has a larger betweenness centrality than vertices that do not. Betweenness centrality is defined in this Wikipedia article.  

NodeXL uses the algorithm described in the paper "A Faster Algorithm for Betweenness Centrality," by Ulrik Brandes. The paper can be found here.  

The closeness centrality of a vertex is the inverse of the sum of the shortest distances between the vertex and all other vertices reachable from it as defined in this Wikipedia article. |
| Vertex eigenvector centrality                | Eigenvector centrality is defined in this article. "Eigenvector centrality is a measure of the importance of a node in a network. It assigns relative scores to all nodes in the network based on the principle that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes." |
| Vertex PageRank                              | PageRank measures the importance of each vertex                                                                                                                                                          |

---

62 [http://snap.stanford.edu/class/cs224w-readings/brandes01centrality.pdf](http://snap.stanford.edu/class/cs224w-readings/brandes01centrality.pdf)
<table>
<thead>
<tr>
<th><strong>NodeXL</strong></th>
<th>within the graph using a link analysis algorithm developed by Larry Page. It is described here(^{65}).</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vertex clustering coefficient</strong></td>
<td>Clustering coefficient is defined in this [Wikipedia article]^{66}: &quot;The clustering coefficient of a vertex in a graph quantifies how close the vertex and its neighbors are to being a clique (complete graph).&quot; Self-loops, which would render the clustering coefficients invalid, are skipped.</td>
</tr>
<tr>
<td><strong>Vertex reciprocated vertex pair ratio (directed graphs only)</strong></td>
<td>In a directed graph, this is the number of adjacent vertices that are connected to the vertex with edges in both directions divided by the number of adjacent vertices. In an undirected graph, this is undefined and is not calculated.</td>
</tr>
<tr>
<td><strong>Edge reciprocation (directed graphs only)</strong></td>
<td>In a directed graph, an edge from vertex A to vertex B is reciprocated if the graph also has an edge from vertex B to vertex A. In an undirected graph, edge reciprocation is undefined and is not calculated. A self-loop is never reciprocated.</td>
</tr>
</tbody>
</table>

**To calculate Vertex Metrics:**

1. In the Excel Ribbon, select **NodeXL > Analysis > Graph Metrics**.
2. Check mark the rows with the Vertex Metrics. We recommend calculating all of them at once.

---

\(^{65}\) [https://en.wikipedia.org/wiki/PageRank](https://en.wikipedia.org/wiki/PageRank)

\(^{66}\) [https://en.wikipedia.org/wiki/Clustering_coefficient](https://en.wikipedia.org/wiki/Clustering_coefficient)
6.3 Group Metrics

**Important note:** Before running **Group Metrics** you need to create groups as explained in chapter 5.

**Group Metrics** include the same metrics presented in the [6.1 Overall Graph Metrics](#) table above. In this case they are calculated as if the vertices and edges outside the group do not exist. These metrics get inserted into the **Groups worksheet**.

**Group Metrics** also include the number of edges within each group and the number of edges between each pair of groups. These metrics get inserted into the **Group Edges worksheet**, which gets added to the workbook the first time you calculate group metrics.

**To calculate Group Metrics:**

1. In the Excel Ribbon, select **NodeXL > Analysis > Graph Metrics**.
2. Check mark the row **Group Metrics**.
6.4 Time Series Analysis

Time Series Analysis can be performed on any column in the Edges worksheet that contains a date/time format.

**To perform Time Series Analysis:**

1. Open the **Time Series dialog box** via **NodeXL > Analysis > Graph Metrics > Time Series > Options**.
2. Select the date column from the dropdown list via “Build a time series for this column”.
3. “Choose time slices” from the dropdown list.
4. Optionally check “Unique edges by this column”.
5. Optionally “Add a slicer” that will be available on the Time Series worksheet to filter the data. You can choose multiple slicers.
6. Select **OK**.

![Image of Time Series dialog box]

**Figure: Time Series**

The result is a new worksheet “Time Series” that is added to the workbook.
6.5 Path Analysis

Path Analysis creates a branching tree structure when data containing information about "parent-child" relationships is present. If each node is explicitly linked to its parent and children, then the paths layout places the collection of initial messages in order along the top of the display, with the replies and replies to replies cascading downward into branches. The shape and size of these path trees can quickly highlight the most active discussions, and the people who initiate them.

To perform a Path analysis:

1. Open the Path dialog box via NodeXL > Analysis > Graph Metrics > Paths > Options.

   ![Path dialog box](image)

   Figure: Path

2. Select the appropriate columns from the dropdown lists. The image above shows the correct setting for data collected with one of the NodeXL Pro Twitter data importers.
3. Select OK.

Path analysis results in three new worksheets that are added to the workbook: Path Edges - Path Vertices - Path Metrics.

4. To view the path graph, set the layout of the graph pane to Path and refresh the graph. Visual properties can only be changed via the Graph Options dialog box.
6.6 Word and Word Pairs / Sentiment analysis

The **Word and Word Pairs** analysis offers two powerful analytical features for columns that contain text. First it will create two new worksheets that show the most frequently used words and word pairs in the whole network and by group. Second it will perform a sentiment analysis based on the word lists provided in the text boxes on the right side of the dialog box.

**To calculate word and word pair counts:**

1. Open the Word and Word Pair Metrics dialog box via **NodeXL > Analysis > Graph Metrics > Words and Word Pairs > Options**.

![Image: Word and Word Pair Metrics]

2. Select a column that contains text either on the Edges or the Vertices worksheet via the dropdown list under “Count words and word pairs in this column”.
3. We recommend that you check “Count by group” (note that groups need to be calculated first, see chapter 5) and “Skip words and word pairs that only appear once”.
4. The checkboxes for “Add “Vertex Content” worksheet” and “Add “Word List worksheet” only need to be checked, if you are creating a data set that you would like to export to **NodeXL Pro INSIGHTS**.
5. Optionally customize the list of stop words in the box below “Skip these words”. These words are disregarded in the text analysis.

**To perform a sentiment analysis:**

6. Review the text boxes on the left.
7. Add or remove words to the word lists.
8. Optionally enter new lists of words.
9. Click OK.

Things to know about word and word pairs analysis:

- Two new worksheets are added to the workbook: Words and Word pairs. These worksheets contain lists of all words/word pairs with the overall counts and the salience of the respective words.
- New columns are added to the Edges, Vertices and Groups worksheets showing the count and percentage of words in relation to the sentiment word lists.

Things to know about sentiment analysis in NodeXL:

- You may modify, replace and rename these lists. Consider student versus teacher words, or investor versus consumer words, or even lists of words that are distinctive of different languages. You may use any language to populate these lists.
- NodeXL ships with a default collection of American English language “Positive” and “Negative” terms. These lists have been created by Dr Bing Liu’s research group at the University of Illinois Chicago. References:


- If you are looking for other languages, we currently have skip words and positive/negative words lists available in German and Korean. These are available when downloading the official NodeXL Pro data recipe bundle via this page69. You can easily find stop words and sentiment word lists in almost any language on platforms like Kaggle70 and Github71, and then copy and paste these lists in the text analysis dialog box.

**Expert tip 1:** The word pairs worksheet can be used as an Edge list in a new NodeXL workbook. Check out our Tutorial section72 to learn more about these semantic networks.

**Expert tip 2:** You can use the word counts of positive and negative words on the Vertices worksheet to color the Vertices in the graph accordingly.

---

69 https://www.smrfoundation.org/nodexl/automation
70 https://www.kaggle.com/datasets/heeraldedhia/stop-words-in-28-languages
71 https://github.com/search?q=stop+words
72 https://www.smrfoundation.org/nodexl/tutorials
6.7 Top items

The Top Items feature creates a worksheet with the top vertices based on columns that you select.

To create a Top Items worksheet:

1. Open the Top Items Metrics dialog box via NodeXL > Analysis > Graph Metrics > Top Items > Options.

![Figure: Top Item Metrics](image)

2. Select the Add button and add rankings based on the columns of the vertex worksheet.

**Note:** When uploading a data set to NodeXL Graph Gallery, this worksheet is used to display the top 10 influencers in the network. We suggest limiting this worksheet to just the top 10 vertices by Betweenness centrality.

6.8 Twitter Search Network Top Items

This feature does not have any options settings. It is customized for data sets that have been created with one of the NodeXL Pro Twitter data importers. The result is a new worksheet that contains overall counts and counts by the top 10 groups of the most frequently used hashtags, URLs, domains, words and word pairs, as well as the most Replied-To users, the most Mentioned users and the most active users by overall tweet count.

To create a Twitter Search Network Top Items worksheet:
1. Simply set a check mark for Twitter Search Network Top Items in the Graph Metrics dialog box.

Note that new columns are also added to the Vertices and Groups worksheets showing the top items by vertex and group. The new columns on the Groups worksheet are very valuable because you can use them as group labels in the graph when displaying the 7.5 Group-in-a-box layout.

6.9 Network Top Items

Network Top Items works similar to the 6.8 Twitter Search Network Top Items feature, but it allows the customization for any network data set.

To create a Network Top Items worksheet:

1. Open the Network Top Items dialog box via NodeXL > Analysis > Graph Metrics > Network Top Items > Options.

![Network Top Items dialog box]

Figure: Network Top Items

2. Select the Add button and add rankings based on columns of the Edges worksheet.
3. The top words and word pairs analysis can be calculated for any text column on the Edges worksheet (e.g. Tweet text) or any text column on the Vertices worksheet (e.g. User description).
6.10 Edge creation by shared content similarity

You can have NodeXL create an edge based on the similarity of the content used by two vertices. Select a column of space or comma delimited words from the Vertices worksheet, and set the threshold of similarity required to create a new edge. NodeXL will calculate the level of similarity between each pair of vertices based on the following algorithm:

\[
\text{Similarity (\%)} = \frac{\text{number of shared words}}{\text{MAX}\left[\text{number of words in each column}\right]}\]

If the similarity is equal or greater than the threshold, NodeXL will append a new edge to the Edges worksheet that connects the two vertices. The new edge will have the following attributes:

- Relationship: Shared Content: [ColumnName]
- Edge Weight: Similarity
- Shared Content: Words shared between the two vertices

You can choose to only create new edges for isolated vertices (i.e. vertex with only one edge, and the edge is a self-loop edge) by checking the "Limit to isolated vertex" option.

6.11 Summarizing the Graph

If you create a large number of graphs, NodeXL's graph summary feature might be useful to you.

NodeXL automatically keeps track of several aspects of how a graph was created - how it was laid out, how its visual properties were set using Autofill Columns, how its vertices were grouped, and so on. It will provide you with a summary of this information, which you can use for your own documentation.

To get a summary of the graph:

- In the Excel Ribbon, go to NodeXL > Graph > Summary.

Note that for privacy reasons, NodeXL does not keep track of the most recent data that was imported into the workbook unless you tell it to.

To tell NodeXL to keep track of the data that gets imported into the workbook:

1. In the Excel ribbon, go to NodeXL > Data > Import > Import Options.
2. Check the box "Add a description of the imported data to the graph summary".
7. Changing How the Graph Looks

Unless you tell it otherwise, NodeXL uses a set of default visual properties for the graph in the Graph Pane. For example, vertices are small black disks by default and edges are thin black lines.

You can change the graph’s default visual properties. You can also set the visual properties of individual edges, vertices and groups, or tell NodeXL to automatically calculate visual properties for all the graph’s edges, vertices and groups based on values in other workbook columns.

7.1 Graph Options: Default Visual Properties

You can change the default visual properties that NodeXL uses for edges and vertices whose individual visual properties have not been set.

To change the graph’s default visual properties:

1. Click the **Graph Options** button at the top of the Graph Pane. (If the Graph Pane is narrow, the Graph Options button may not be visible. Click the down-arrow at the upper-right edge of the Graph Pane to make the button visible.)

   ![Graph Options](image)

   Figure: Graph Options

2. Edit the options in the Graph Options dialog box
3. Click **OK**. Your changes will immediately become visible in the Graph Pane.

You can also set the **7.2 visual properties of individual edges, vertices and groups**, or tell NodeXL to **7.4 automatically calculate visual properties** based on values in other workbook columns.
**Edges Tab**

Try out different edge colors, widths, arrow head sizes and curvature. Note that **Bundled edges** will remove the arrowheads from the directed graphs. Laying out the graph also takes considerably longer with bundled edges, especially in graphs with more than 10,000 edges.

**Vertices Tab**

You can choose from a variety of vertex shapes from the dropdown menu:

![Vertex shapes](image)

Images can be added by adding a path to image files on your local drive, or you can link to online content. Wikipedia is a good source for finding image links.

**Expert tip:** Set the Effects on the Vertices tab to “Drop shadow”. This effect will look great on almost any graph with a light background. Set it to “Glow” for dark background colors.
7.2 Setting Visual Properties for Individual Edges, Vertices or Groups

The graph's default visual properties can be overridden for individual edges, vertices and groups. For example, you can make an edge appear as a wide blue dashed line, and a vertex appear as a large green sphere.

**To manually set a visual property for an edge, vertex or group:**

1. Select the Edges, Vertices or Groups worksheet.
2. If the worksheet's Visual Properties column group is not visible, show the column group via Show/Hide > Workbook Columns.
3. Edit one or more cells in the Visual Properties column group.
4. Show/refresh the graph to make your changes appear in the Graph Pane.

When you select a cell in a worksheet, a help message appears to tell you what the cell should contain. Hovering the mouse over the cell's column heading also provides information about the column's cells.

See **7.3 Labeling Edges, Vertices and Groups** for information on how to use the Label columns on the Edges, Vertices and Groups worksheets.

**To set a visual property for one or more edges, vertices or groups using the Ribbon:**

1. Select the Edges, Vertices or Groups worksheet.
2. Select one or more rows in the worksheet.
3. Use one of the items in the Visual Properties group in the Ribbon. For example, to change the color of the selected edges, vertices or groups, select NodeXL > Visual Properties > Color.
4. The Graph Pane will refresh automatically. (If it doesn't, check the "Automatically refresh the graph" checkbox in the Graph Options dialog box.)

As an alternative to setting visual properties on individual edges, vertices and groups, you can tell NodeXL to automatically calculate visual properties based on values in other workbook columns (see below).
7.3 Labeling Edges, Vertices and Groups

You can add edge and vertex labels to the graph. An edge label appears on top of the edge, while a vertex label can appear either inside a Label shape that represents the vertex or as an annotation next to the vertex.

If your graph has 5 groups, you can label the groups. Group labels appear in the graph when you choose to lay out the 7.5.2 groups in boxes, and when you collapse groups.

To label edges:

1. Select the Edges worksheet.
2. If the worksheet's Label column group is not visible, show the column group via NodeXL > Show/Hide > Workbook Columns.
3. For each edge that should have a label, enter the label text in the Label column.
4. Show/Refresh the graph to make your changes appear in the Graph Pane.

To label vertices with an annotation:

1. Select the Vertices worksheet.
2. If the worksheet's Label column group is not visible, show the column group via NodeXL > Show/Hide > Workbook Columns.
3. For each vertex that should be annotated with a label, enter the label text in the Label column.
4. Show/Refresh the graph to make your changes appear in the Graph Pane. The vertex will look something like this:
5. The properties of the label like size, font and position can be set via NodeXL > Graph Options > Other > Label....

![Label Options](image)

Figure: Label Options
To label vertices using the Label shape:

1. Select the Vertices worksheet.
2. If the worksheet's Visual Properties column group and the Label column group are not visible, show these column groups via NodeXL > Show/Hide > Workbook Columns.
3. For each vertex that should have a Label shape, enter Label in the Shape column and enter the label text in the Label column.
4. Show/Refresh the graph to make your changes appear in the Graph Pane. The vertex will look something like this:
5. The size of the label corresponds with the Size column in the Visual Properties column group.

To label groups:

1. Select the Groups worksheet.
2. If the worksheet's Visual Properties column group and the Label column group are not visible, show these column groups via NodeXL > Show/Hide > Workbook Columns.
3. For each group that should have a label, enter the label text in the Label column.
4. Show/Refresh the graph to make your changes appear in the Graph Pane.
5. The properties of the label like size, font and position can be set via NodeXL > Graph Options > Other > Label....

![Figure: Label Options]
7.4 Autofill Columns: Automatically Calculating Visual Properties for All Edges, Vertices or Groups

You can tell NodeXL to automatically calculate one or more visual properties for all the graph's edges, vertices and groups based on the values in other workbook columns. This is called **autofilling columns**. For example, if you have an **Edge Weight column** on the Edges worksheet, NodeXL can calculate and set the width, color, style and opacity of each of the graph's edges based on the edge's weight. Edges with a large Edge Weight will then be thicker than edges with a small Edge Weight. And if you have calculated vertex metrics, you can size the vertices based on an indicator like Betweenness Centrality.

By default, NodeXL calculates the edge widths, for example, using a linear mapping between the full range of Edge Weight numbers and the full range of available widths, but you can specify various options to change the way NodeXL performs its calculations.

**To automatically calculate visual properties:**

1. In the Excel Ribbon, select **NodeXL > Visual Properties > Autofill Columns**.

   ![Autofill Columns](image)

   **Figure: Autofill Columns**

2. The **Autofill Columns dialog box** contains three tabs: Edges - Vertices - Groups. There is one row for each worksheet column that can be autofilled. Find the name of the worksheet column you want to autofill, then select the "source column" that contains the values NodeXL should use for the calculations. In the above example, select "Edge Weight" as the source column for the Edge Width column.
3. Repeat step 2 for any additional worksheet columns you want to autofill.
4. Click the **Autofill button**. The worksheet columns get filled with the calculated values and the graph refreshes automatically. (If it doesn't refresh automatically, check the "Automatically refresh the graph" checkbox in the Graph Options dialog box.)

**To change the way NodeXL calculates a visual property:**

1. In the **Autofill Columns dialog box**, click the Options button ( ) for the **Edge Width** worksheet column, for example.
2. Select Edge Width Options from the popup menu.
3. Change the options in the dialog box that appears.

To clear a worksheet column that has been autofilled:

1. In the Autofill Columns dialog box, click the Options button ( ) for the Edge Width worksheet column, for example.
2. Select Clear Edge Width Worksheet Column Now from the popup menu.

To clear all worksheet columns that have been autofilled:

1. In the Autofill Columns dialog box, click the "Clear All Worksheet Columns Now" button.

Here are some things to know about Autofill Columns:

- If a source column doesn't exist when Autofill is clicked, the corresponding worksheet column does not get autofilled.
- Autofilling multiple columns in a large workbook can be slow, especially if one of your worksheets is filtered.
7.5 Changing How the Graph is Laid Out

NodeXL uses one of a variety of layout algorithms to lay out - or position - the graph's vertices. Some algorithms place the vertices in simple geometric arrangements, while others attempt to intelligently arrange the vertices to produce an aesthetically pleasing graph. You can lay out the entire graph, lay out only a subset of the graph's vertices, and "snap" vertices to an imaginary grid. A variety of layout options are available.

7.5.1 Layout Algorithms

NodeXL uses the selected layout algorithm to lay out the graph's vertices when the graph is shown, and when some or all of the graph's vertices are laid out again. The available layout algorithms are listed in the table below.

To select the layout algorithm that NodeXL uses, do one of the following:

- In the NodeXL Ribbon, select Graph > Layout, or...
- At the top of the Graph Pane, select from the Layout drop-down list ( ), or...
- Right-click anywhere in the Graph Pane and select Layout from the right-click menu.

<table>
<thead>
<tr>
<th>Layout Algorithm</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fruchterman-Reingold</strong> 73</td>
<td>This is a force-directed algorithm that attempts to minimize edge crossings. This algorithm is iterative, which means that if you lay out the graph again, the current vertex locations are used as starting points for the new calculations. If the Edges worksheet has an Edge Weight column, the Fruchterman-Reingold layout algorithm uses the column's values to determine the attractive forces between vertices. An edge with a larger Edge Weight exerts a stronger attractive force on its vertices. Edge Weight column values should be greater than zero. If a value is zero or less, it is ignored and 1.0 is used instead.</td>
</tr>
<tr>
<td><strong>Harel-Koren Fast Multiscale</strong> 74</td>
<td>This is a force-directed algorithm that attempts to minimize edge crossings. This algorithm is faster than Fruchterman-Reingold when used with graphs that have a large number of vertices. It is not iterative, which means that the current vertex locations are ignored when calculating new locations.</td>
</tr>
<tr>
<td>Circle</td>
<td>Evenly places the vertices on the circumference of a circle.</td>
</tr>
<tr>
<td>Spiral</td>
<td>Evenly places the vertices along a spiral.</td>
</tr>
</tbody>
</table>

73 [https://link.springer.com/chapter/10.1007/978-3-658-21742-6_49](https://link.springer.com/chapter/10.1007/978-3-658-21742-6_49)
74 [https://www.wisdom.weizmann.ac.il/~harel/papers/ms_iqaa.pdf](https://www.wisdom.weizmann.ac.il/~harel/papers/ms_iqaa.pdf)
<table>
<thead>
<tr>
<th>Horizontal Size Wave</th>
<th>Evenly places the vertices along a sine wave that runs from left to right.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertical Sine Wave</td>
<td>Evenly places the vertices along a sine wave that runs from top to bottom.</td>
</tr>
<tr>
<td>Grid</td>
<td>Places the vertices on an evenly-spaced grid. The grid dimensions are selected to optimally fill the Graph Pane.</td>
</tr>
<tr>
<td>Polar</td>
<td>Each vertex's location is specified by polar coordinate columns on the Vertices worksheet. To lay out the graph using polar coordinates that you specify, do the following:</td>
</tr>
<tr>
<td></td>
<td>1. Select the <strong>Polar</strong> layout algorithm.</td>
</tr>
<tr>
<td></td>
<td>2. Show the <strong>Layout column group</strong> via <strong>Show/Hide &gt; Workbook Columns</strong>. This makes the Polar R and Polar Angle columns visible on the Vertices worksheet.</td>
</tr>
<tr>
<td></td>
<td>3. Hover the mouse over the <strong>Polar R column</strong> header for instructions on how to use the columns.</td>
</tr>
<tr>
<td></td>
<td>4. Optionally use <strong>Visual Properties &gt; Autofill Columns &gt; Vertices</strong> to fill these columns automatically.</td>
</tr>
<tr>
<td>Polar Absolute</td>
<td>The Polar Absolute layout algorithm is similar to the Polar layout algorithm, but the units of the Polar R column are different. The instructions that pop up when you hover the mouse over the Polar R column (see Polar, above) explain the difference.</td>
</tr>
<tr>
<td>Sugiyama</td>
<td>Attempts to minimize edge crossings by arranging the vertices in layers.</td>
</tr>
<tr>
<td>Random</td>
<td>This algorithm places the vertices at random locations. Randomizing the vertex locations can be useful when used in conjunction with the Fruchterman-Reingold layout algorithm, which uses the current vertex locations as a starting point for its calculations. If you want the Fruchterman-Reingold algorithm to start from scratch, lay out the graph with the Random layout algorithm, then switch to Fruchterman-Reingold and lay it out again.</td>
</tr>
<tr>
<td>Path Layout</td>
<td>The Path Layout creates a branching tree structure when data containing information about &quot;parent-child&quot; relationships is present. If each node is explicitly linked to its parent and children, then the paths layout places the collection of initial messages in order along the top of the display, with the replies and replies to replies cascading downward into branches. The shape and size of these path trees can quickly highlight the most active discussions, and the people who initiate them. In order to use this layout you first need to run Path analysis via <strong>Analysis &gt; Graph Metrics &gt; Paths</strong>.</td>
</tr>
</tbody>
</table>
Leaves each vertex exactly where it is. The None layout algorithm is useful if you are happy with the graph's current vertex locations, but you need to edit the workbook and refresh the graph. Normally, refreshing the graph reads the workbook and lays out the graph. This means that the vertices will be placed in slightly different locations. If you select the None layout algorithm, you can edit a vertex color, for example, and then refresh the graph without affecting the vertex locations.

7.5.2 Layout Options

You can change the way NodeXL lays out the graph with a number of layout options:

- The typical layout case of a network graph is to lay out the entire graph at once.
- If your graph has 5. groups, you can tell NodeXL to lay out each group in its own box. This may be one of the most practical features of NodeXL.
- If your graph has a large number of small connected components that obscure the larger, more important components, you can tell NodeXL to move the smaller components to the bottom of the Graph Pane.
- If you have selected the Fruchterman-Reingold layout algorithm, you can control how that algorithm works.

To set layout options, do one of the following:

1. In the Excel Ribbon, select NodeXL > Graph > Layout > Layout Options, or…
2. At the top of the Graph Pane, select Layout Options from the Layout drop-down list.
To lay out each group within its own box in the Graph Pane:

1. In the Excel Ribbon, select **NodeXL > Graph > Layout > Layout Options**.
2. In the **Layout Options** dialog box, check "Lay out each of the graph’s groups in its own box and sort the boxes by group size."
3. Select one of the three **box layout algorithms**: Treemap - Packed rectangles - Force directed
4. Optionally set the width of the box outlines; select whether to show, hide or combine intergroup edges (see below); and select whether you want NodeXL to use the Grid layout for groups that don’t have many edges.
5. Click **OK**.
6. **Show/Refresh** the graph or lay it out again.

In these layouts, the size of each box is proportional to the number of vertices in the box’s group. The **Treemap** layout places the largest group in the upper left corner and the smallest group in the lower right corner. The **Packed rectangles** layout places the largest

---

group in the center and lays out all other groups as a series of rectangles wrapped around that largest core group. The **Force directed** layout treats each group as a meta-node, and then uses the Fruchterman Reingold layout to position them.

![Figure: Group-in-a-Box Layouts](image)

If you choose to combine intergroup edges, all edges between a pair of groups are combined into a single edge whose width reflects the number of combined edges. This is useful for reducing graph clutter. If you combine intergroup edges and the Edges worksheet has an Edge Weight column, the edge weights are used to determine the width of a combined edge. Edge Weight column values should be greater than zero. If a value is zero or less, it is ignored and 1.0 is used instead.

### 7.5.3 Selectively Laying Out Parts of the Graph Again

When you show the graph, NodeXL reads the workbook and lays out the graph using the selected layout algorithm. You can then lay out the entire graph again without rereading the workbook, which can save time when the workbook is large. You can also lay out a subset of the graph’s vertices.

Note that if you edit the workbook (to change a vertex color, for example), you must refresh the graph to get your edits to show up in the Graph Pane. If you don't want the graph to be laid out again when you refresh the graph, set the layout algorithm to None.

**To lay out the entire graph without rereading the workbook:**

1. In the layout menu at the top of the Graph Pane, select **Lay Out Again**.

**To lay out a subset of the graph’s vertices without rereading the workbook:**

1. Select the vertices you want to lay out.
2. In the layout menu at the top of the Graph Pane, select **Lay Out Again > Lay Out Selected Vertices Again**.

**To lay out a subset of the graph’s vertices within their "bounding box" without rereading the workbook:**

1. Select the vertices you want to lay out.
2. In the layout menu at the top of the Graph Pane, select **Lay Out Again > Lay Out Selected Vertices Again, Within Bounds**. This keeps the selected vertices within the bounding box defined by the outermost selected vertices.

**To lay out only the graph's visible vertices without rereading the workbook:**

1. Select the vertices you want to lay out.
2. In the layout menu at the top of the Graph Pane, select **Lay Out Again > Lay Out Visible Vertices Again**. Vertices that have been filtered using NodeXL's Dynamic Filters do not get laid out.

### 7.5.4 Snapping Vertices to a Grid

When the graph is laid out using one of the force-directed layout algorithms, it can sometimes be useful to refine the layout by snapping the vertices to the nearest locations on an imaginary grid. You can set the size of the grid, which is used only for snapping and is not visible.

**To snap the graph's vertices to a grid:**

1. In the layout menu at the top of the Graph Pane, select **Lay Out Again > Snap Vertices to Grid**.

**To set the size of the snapping grid:**

1. In the layout menu at the top of the Graph Pane, select **Lay Out Again > Set Grid Size**.

### 7.6 Zooming, Moving Around and Scaling the Graph

NodeXL's Graph Pane usually shows you the entire graph. If the graph is simple, viewing it in its entirety may be exactly what you want. But if the graph is more complex, you might want to zoom in to examine specific regions more closely. NodeXL lets you zoom into and move around the graph using several techniques. You can also scale the graph, which expands all the graph’s elements while keeping them in the same place.
7.6.1 Zooming the Graph

There are several ways to zoom into and out of the graph.

**To zoom into and out of the graph using the Zoom slider:**

1. Drag the Zoom slider at the top of the Graph Pane’s toolbar.

If the Graph Pane is too narrow, the Zoom slider won't be visible. In that case you can either widen the Graph Pane until the Zoom slider appears, or click the down-arrow at the right edge of the Graph Pane's toolbar to make the Zoom slider appear in a pop-up window.

**To zoom into and out of the graph using the mouse wheel:**

1. Click anywhere in the Graph Pane.
2. Position the mouse over the part of the graph that you want to zoom into or out of.
3. Rotate the mouse wheel.

**To zoom into the graph using mouse clicks:**

1. Click the Zoom In button in the Graph Pane's toolbar. It looks like this:
2. Click the part of the graph that you want to zoom into.

Note that the Zoom In button "sticks." If you want to restore the mouse to its usual role of selecting vertices, click the Select button in the Graph Pane's toolbar. It's the arrow without a plus or minus sign:

**To zoom out of the graph using mouse clicks:**

1. Click the Zoom Out button in the Graph Pane's toolbar. It looks like this:
2. Click the part of the graph that you want to zoom out of.
7.6.2 Moving Around the Graph

Once you've zoomed into the graph, you'll probably want to move around to see different areas of the graph.

To move around the zoomed-in graph using the mouse and keyboard:

1. Click anywhere in the Graph Pane.
2. Hold down the space bar on the keyboard.
3. Hold down the left mouse button and drag the mouse around the Graph Pane.

To move around the zoomed-in graph using only the mouse:

1. Click the Move Around In button in the Graph Pane's toolbar. It looks like this: 
2. Hold down the left mouse button and drag the mouse around the Graph Pane.

Note that the Move Around button "sticks." If you want to restore the mouse to its usual role of selecting vertices, click the Select button in the Graph Pane's toolbar. It's the arrow without a plus or minus sign:

7.6.3 Scaling the Graph

You can scale the graph in the Graph Pane, which expands all the graph’s vertices, vertex labels, edges and edge labels while keeping them in the same place. This can be useful with a graph that has few vertices, where increasing the scale can result in a better graph visibility.

To scale the graph:

1. Drag the Scale slider at the top of the Graph Pane's toolbar.

If the Graph Pane is too narrow, the Scale slider won't be visible. In this case, you can either widen the Graph Pane until the Scale slider appears, or click the down-arrow at the right edge of the Graph Pane's toolbar to make the Scale slider appear in a pop-up window.
7.7 Selecting Graph Elements

You can select one or more of the graph's vertices and edges. Selected vertices and edges are highlighted in the NodeXL worksheets and in the Graph Pane. You can change the color and size of the selected vertices and edges in the Graph Pane using the **Graph Options** dialog box:

If your graph has 5 groups, you can select an entire group at once.

7.7.1 Selecting Vertices

You can select one or more vertices in either the Vertices worksheet or the Graph Pane. Selecting a vertex in the Vertices worksheet automatically selects it in the Graph Pane, and selecting a vertex in the Graph Pane automatically selects it in the Vertices worksheet. By default, selecting a vertex automatically selects its incident edges. If you don't want this automatic behavior, go to **Graph Options** and uncheck "when a vertex is clicked, select its incident edges."

**To select one or more vertices in the Vertices worksheet:**

1. Select the Vertices worksheet.
2. To select just one vertex, click on the row number all the way at the left.
3. To select multiple adjacent rows, click on a row number, hold down the **Shift** key, and click on another row above or below the previously selected vertex.
4. To select multiple non-adjacent rows, click on a row number, then hold down the **Ctrl** key, and click on any other row.

**To deselect all vertices in the Vertices worksheet:**

1. Select the Vertices worksheet.
2. Click anywhere outside the vertex list.

**To select one vertex in the Graph Pane:**

1. Click the vertex. If other vertices were previously selected, they get deselected.

**To select a set of vertices that are close to each other in the Graph Pane:**

1. Drag a box around the vertices.

**To select a set of vertices that are not close to each other in the Graph Pane:**
1. Hold down the Ctrl key while clicking each vertex.

To toggle the selection state of a vertex in the Graph Pane:

1. Hold down the Ctrl key while clicking the vertex. If it was selected before, it gets unselected. If it was not selected before, it gets selected.

To toggle the selection state of a set of vertices in the Graph Pane:

1. Hold down the Ctrl key while dragging a box around the vertices.

To deselect all vertices in the Graph Pane:

1. Click any empty area in the Graph Pane.

To make it easier to select multiple vertices in the Graph Pane:

1. Click the Add to Selection button at the top of the Graph Pane. It looks like this:

2. Hover the mouse over the Add to Selection button for instructions.

To make it easier to deselect multiple vertices in the Graph Pane:

1. Click the Subtract from Selection button at the top of the Graph Pane. It looks like this:

2. Hover the mouse over the Subtract from Selection button for instructions.

To perform selection tasks on one vertex in the Graph Pane:

1. Right-click the vertex in the Graph Pane.
2. Use one of the selection-related items in the right-click menu.

To perform selection tasks on all vertices in the Graph Pane:

1. Right-click any empty area in the Graph Pane.
2. Use one of the selection-related items in the right-click menu.

7.7.2 Selecting Edges

You can select one or more edges in the Edges worksheet, which automatically selects them in the Graph Pane. Although you can perform a few edge-related tasks in the Graph Pane, you cannot directly select a single edge in the Graph Pane.

To select one or more edges in the Edges worksheet:

1. Select the Edges worksheet.
2. To select just one edge, click on the row number all the way at the left.
3. To select multiple adjacent rows, click on a row number, hold down the **Shift** key, and click on another row above or below the previously selected edge.

4. To select multiple non-adjacent rows, click on a row number, then hold down the **Ctrl** key, and click on any other row.

**To deselect all edges in the Edges worksheet:**

1. Select the Edges worksheet.
2. Click anywhere outside the edge list.

**To select all edges in the Graph Pane:**

1. Right-click any empty area in the Graph Pane.
2. On the right-click menu, select **Select All > Edges.**

**To deselect all edges in the Graph Pane:**

1. Right-click any empty area in the Graph Pane.
2. On the right-click menu, select **Deselect All > Edges.**

**To deselect a vertex's incident edges in the Graph Pane:**

1. Right-click the vertex in the Graph Pane.
2. Select **Deselect Incident Edges.**
7.8 Graphs with Isolated Vertices

Some graphs have isolated vertices, or isolates, which are vertices that are not connected to any edges. If your graph has isolates, you have to take special action to get them to appear in the Graph Pane.

To add an isolated vertex to the graph:

1. Select the Vertices worksheet.
2. Enter the name of the vertex in the first empty cell in the Vertex column.
3. Set the Visibility cell for the new vertex row to Show. If you don't see a Visibility column on the Vertices worksheet, select "Visual Properties" via Show/Hide > Workbook Columns.

Step 3 is required because by default, a vertex on the Vertices worksheet is shown in the Graph Pane only if it is connected to at least one edge.

There is a reason why NodeXL works this way. If every vertex listed in the Vertices worksheet appeared in the Graph Pane, even those not connected to edges, then you would not be able to delete an edge from the Edges worksheet without checking whether the edge's vertices were now isolated, and then removing the isolates from the Vertices worksheet if necessary. That would be quite burdensome.
7.9 Adding Tooltips to the Graph

You can add a tooltip to each of the graph's vertices. A tooltip is a small text window that pops up when you hover the mouse over a vertex in the Graph Pane. Tooltips can interactively provide detailed vertex information without cluttering the graph with long vertex labels.

You cannot add tooltips to the graph's edges.

To add a tooltip to a vertex:

1. Select the Vertices worksheet.
2. If the worksheet's Labels column group is not visible, show the column group via NodeXL > Show/Hide > Workbook Columns > Labels.
3. Fill in one or more cells in the Tooltip column.

To insert a line break in a vertex's tooltip:

1. Enter Alt+Enter in the vertex's Tooltip cell.

To show a vertex's tooltip in the Graph Pane:

1. Show the graph.
2. Hover the mouse over the vertex. The tooltip will pop up after a short period.

Note: When working with Twitter data, an automatic tooltip is added which contains the username and the latest tweet by that user in the data set.

7.10 Showing and Hiding Graph Elements

You can show and hide various elements in NodeXL's Graph Pane. For example, you can hide edge and vertex labels and show the graph legend and graph axes.

To show or hide graph elements:

1. In the Excel Ribbon, select NodeXL > Show/Hide > Graph Elements.

Check the graph elements you want to show and uncheck the graph elements you want to hide.
7.11 Using Dynamic Filters

NodeXL lets you filter the graph's edges and vertices based on numeric and date/time attribute values. For example, you can filter the graph so that only vertices with Degree between 0 and 2 are visible in the Graph Pane. You specify the filters with a set of sliding range bars, and as you adjust the range bars, the filtered edges and vertices disappear immediately from the Graph Pane. They are still contained in the graph, but they are hidden. Dynamic filtering is different from Excel's built-in filter features. (Excel's built-in features include the Data, Sort & Filter, Filter item in the Ribbon, and the down-arrow menus in the column headers.) If you filter out edges or vertices using Excel filtering, the filtered edges and vertices are completely ignored by NodeXL and never become a part of the graph. NodeXL's dynamic filters, on the other hand, hide the filtered vertices and edges but do not remove them from the graph.

The Dynamic Filters dialog box has one filter for each numeric and date/time column in the Edges and Vertices worksheets. Each filter spans the range of values contained in its respective column, and a histogram shows you the value distribution over the range. By default, NodeXL always shows edges and vertices that have empty cells in a column. For example, if you are filtering on the Degree column on the Vertices worksheet and some vertices have an empty Degree cell, those vertices will be shown in the Graph Pane regardless of your Degree filter settings. See To specify how empty cells are handled for information on changing this behavior.

If you refresh the graph while the Dynamic Filters dialog box is open, the dynamic filters remain applied to the Graph Pane. If you close the Dynamic Filters dialog box and refresh the graph, the dynamic filters are removed. You can apply them again by opening the Dynamic Filters dialog box again.

To dynamically filter the graph's vertices and edges:

1. In the Excel Ribbon, select NodeXL > Analysis > Dynamic Filters, or...
2. Select Dynamic Filters above the Graph Pane.
3. In the Dynamic Filters dialog box, adjust the filters for the edge and vertex attributes you want to filter on. See the sections below for details on how adjustments can be made.

To adjust the lower end of a dynamic filter's range, do one of the following:

- Drag the slider at the left end of the filter's range bar, or...
- Enter a value into the text box at the left end of the filter, or...
- Use the up/down arrows in the text box at the left end of the filter, or...
- Select the filter's range bar and use the Left Arrow and Up Arrow keys.

To adjust the upper end of a dynamic filter's range, do one of the following:

- Drag the slider at the right end of the filter's range bar, or...
- Enter a value into the text box at the right end of the filter, or...
- Use the up/down arrows in the text box at the right end of the filter, or...
- Select the filter's range bar and use the Down Arrow and Right Arrow keys.
To shift the dynamic filter's range to the left or right, do one of the following:

- Drag the filter's range bar, or...
- Select the filter's range bar and use the Page Up and Page Down keys.

To specify how empty cells are handled:

- Check or uncheck the "show edges and vertices if cells are empty" checkbox.

To adjust the opacity of the dynamically filtered vertices and edges:

- Adjust the "filter opacity" setting.
7.12 Creating Subgraph Images

NodeXL will create an image of a subgraph for each of the graph's vertices. A vertex's subgraph consists of the vertex, its incident edges, and optionally its adjacent vertices, their incident edges, and so on, out to a specified level. Here is a sample 1.5-level subgraph image:

Figure: Subgraph image

Note that the complete graph had many more vertices and edges than are shown here. All the other vertices and edges were removed to create the vertex's subgraph image. The subgraph images can be saved as files in a specified folder or as thumbnails inserted into the Vertices worksheet.

If you save the subgraph images in a folder, each image file is named after the vertex. The JPEG image file for a vertex named "Vertex 123" will be "Vertex 123.jpg," for example. If a vertex name includes characters that aren't valid in file names, those characters get replaced with hexadecimal representations. The JPEG image file for a vertex named "A\B" will be "A%5CB.jpg," for example. (The backslash, which is not valid in file names, gets replaced with %5C, its hexadecimal representation.)

The "X", "Y" and "Locked?" columns on the Vertices worksheet are ignored when the subgraph images are created.

To create subgraph images for all of the graph's vertices:

1. In the Excel Ribbon, select **NodeXL > Analysis > Subgraph Images**.

Figure: Subgraph imagedialog box
2. In the **Subgraph Images dialogue box**, specify the size of the subgraph images and choose to either “Save subgraph image files in folder” or “Insert subgraph thumbnails into the Vertices worksheet”. Note that if you choose the latter, sorting the Vertices spreadsheet may take a long time.

**To create subgraph images for selected vertices only:**

1. Select one or more vertices.
2. In the Excel Ribbon, select **NodeXL > Analysis > Subgraph Images**.
3. In the **Subgraph Images dialogue box**, check "Create subgraph images for selected vertices only."

**To delete any subgraph image thumbnails that were inserted into the Vertices worksheet:**

1. In the Excel Ribbon, select **NodeXL > Analysis > Subgraph Images > Delete Thumbnails**. Note that this does not delete any subgraph images files that were saved in a folder.
7.13 Saving an Image of the Graph

You can save an image of the graph in a variety of formats. The available formats are shown in the table below. You can also copy an image of the graph to the Clipboard.

<table>
<thead>
<tr>
<th>Image Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PNG</td>
<td>A raster format that uses lossless compression. PNG is often the best choice for saving NodeXL graph images.</td>
</tr>
<tr>
<td>JPEG</td>
<td>A raster format that uses lossy compression. The file is often smaller than a PNG file for the same image, but it may suffer from compression artifacts.</td>
</tr>
<tr>
<td>TIFF</td>
<td>In NodeXL, this is a non-compressed raster format. TIFF files tend to be large.</td>
</tr>
<tr>
<td>BMP</td>
<td>A non-compressed raster format. BMP files tend to be large.</td>
</tr>
<tr>
<td>XPS</td>
<td>A vector format, similar in concept to PDF. XPS files can be viewed directly in recent versions of Windows. You can convert an XPS file to a PDF file in Adobe Acrobat and other programs if necessary.</td>
</tr>
<tr>
<td>GIF</td>
<td>A raster format that uses lossy compression. For NodeXL graphs, the PNG or JPEG format is usually a better choice.</td>
</tr>
</tbody>
</table>

To save an image of the graph to a file:

1. Right-click anywhere in the Graph Pane.
2. On the right-click menu, select **Save Image to File** > **Save Image**.

To set the size of or add a header or footer to the saved graph image:

1. Right-click anywhere in the Graph Pane.
2. On the right-click menu, select **Save Image to File** > **Image Options**.

To copy an image of the graph to the Clipboard:

1. Right-click anywhere in the Graph Pane.
2. On the right-click menu, select **Save Image to File** > **Copy Image to Clipboard**. The copied image is always the same size as the Graph Pane.
3. To paste the graph image from the clipboard into a worksheet, select **Home** > **Clipboard** > **Paste**.

---

76 [https://www.adobe.com/acrobat.html](https://www.adobe.com/acrobat.html)
8. Exporting Graph Data

You can export the graph data in a NodeXL workbook to a variety of file formats that can be used by other graph programs. You can also export some or all of the graph data to another workbook.

8.1 Exporting Graph Data to Other Programs

NodeXL can export graph data from the workbook to a number of file formats that are used by other graph programs, so that you can, for example, create a graph in NodeXL, export it as a GraphML \(^{77}\) file, and then import and view the graph in Gephi\(^{78}\). The available export formats are shown in the table below.

<table>
<thead>
<tr>
<th>File Format</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCINET Full Matrix DL</td>
<td>NodeXL will export graph data to a UCINET full matrix DL file that can be imported into the UCINET(^{79}) program. The file includes edge weights but does not include any other columns on the Vertices and Edges worksheets.</td>
</tr>
<tr>
<td>GraphML</td>
<td>GraphML(^{80}) is an industry-standard graph file format supported by a number of graph programs. GraphML supports arbitrary vertex and edge attributes. When exporting to a GraphML file, NodeXL includes all the columns on the Vertices and Edges worksheets in the file.</td>
</tr>
<tr>
<td>Pajek</td>
<td>NodeXL will export graph data to a Pajek file that can be imported into the Pajek(^{81}) program. When exporting to a Pajek file, NodeXL includes vertex locations and edge weights. It does not include any other columns on the Vertices and Edges worksheets.</td>
</tr>
<tr>
<td>GEXF</td>
<td>GEXF(^{82}) is an XML-based format that is more expressive than GDF. It allows for the representation of various graph structures, such as directed and undirected graphs, as well as various types of data associated with nodes and edges. GEXF files consist of a header that describes the graph and its properties, a section for node data, and a section for edge data. GEXF also supports the definition of attributes, such as labels, colors, and sizes, for both nodes and edges.</td>
</tr>
</tbody>
</table>

---

\(^{77}\) https://graphml.graphdrawing.org

\(^{78}\) https://gephi.org

\(^{79}\) https://sites.google.com/site/ucinetsoftware/home

\(^{80}\) https://graphml.graphdrawing.org

\(^{81}\) https://mrvar.fdv.uni-lj.si/pajek

\(^{82}\) https://gexf.net
GDF

GDF is a text-based format that allows for the representation of graph data in a simple tabular format. The file consists of two main sections: vertex data and edge data. Vertex data includes a unique identifier for each node, as well as any additional attributes associated with the vertex. Edge data includes the source and target vertex IDs for each edge, as well as any additional attributes associated with the edge.

To export graph data to another program:

1. In the Ribbon, select **NodeXL > Data > Export**.
2. Select from the first group of items on the Export menu.

---

83 [https://gephi.org/users/supported-graph-formats/gdf-format](https://gephi.org/users/supported-graph-formats/gdf-format)
8.2 Exporting Graph Data to Another Workbook

NodeXL will export selected edges and vertices to a new NodeXL workbook. It will also export the entire graph to a new Excel workbook as an adjacency matrix.

To export selected graph data from the NodeXL workbook to a new NodeXL workbook:

1. Select the edges and vertices you want to export.
2. In the Ribbon, select NodeXL > Data > Export > Selection to New NodeXL Workbook. This creates a new NodeXL workbook, copies the selected edges and vertices into it, and switches to the new workbook. All edge and vertex columns are copied.

To export the graph data from the NodeXL workbook to a new Excel workbook as an adjacency matrix:

1. Select NodeXL > Data > Export > To New Matrix Workbook. This creates a new Excel workbook (not a NodeXL workbook), fills the new workbook with an adjacency matrix whose elements are the graph's edge weights, and switches to the new workbook.

Note that a graph's adjacency matrix has $V$ rows and $V$ columns, where $V$ is the number of vertices in the graph. A graph with a large number of vertices will result in a huge workbook, one that Excel may not be able to handle without running out of memory. Therefore, the Export to New Matrix Workbook feature should be used only with smaller graphs.
8.3 Export Options

You can add information about you and/or your organization. It is displayed when exporting your network to 8.4 NodeXL Graph Gallery and/or 8.5 NodeXL Pro INSIGHTS.

To set your individual Export Options:

1. In the Ribbon, select NodeXL > Data > Export > Export Options.

2. Your hashtag: A hashtag that is associated with your work.
3. Your URL: Your (brand) URL.
4. Your brand logo: A URL that points to your online brand logo image.
5. Your Brand URL: A URL that will be linked to your brand logo image.
6. Action label: Action Label is the Text used for the Action URL.
7. Action URL: Action URL is the link to the user's preferred web site.
8.4 Export to NodeXL Graph Gallery

You can upload your network graphs to the NodeXL Graph Gallery to share it with the world.

For each report an individual URL is created. The report page shows the network graph at the top, along with a description, overall metrics and network top items.

To export your network to NodeXL Graph Gallery:

1. In the Ribbon, select NodeXL > Data > Export > To NodeXL Graph Gallery...

---

84 [https://nodexlgraphgallery.org](https://nodexlgraphgallery.org)
2. NodeXL automatically creates a Title that you can modify.
3. The Description is only available if you have collected this data during import. See 3.1 Import Data Options.
4. If you choose to upload your network as a guest, you will not be able to edit and delete your report. So we recommend that you sign up for a graph gallery account.85
5. Optionally check “Also export the workbook and its options”. To make the workbook available for others.
6. Optionally check “Also export the graph data as GraphML”.
7. Choose the Appearance. If you choose “I may be importing other images, and I want them all to be the same size”, the resolution of the graph image in the gallery report will always be 950 x 688 pixels.

85 https://nodexlgraphgallery.org/Pages/CreateAccount.aspx
8.5 Export to NodeXL Pro INSIGHTS

NodeXL Pro INSIGHTS\textsuperscript{86} is a Microsoft Power BI\textsuperscript{87} based report service that creates web browser accessible interactive social media network reports from NodeXL Pro Twitter datasets.

![Screenshot of NodeXL Pro INSIGHTS](https://powerbi.microsoft.com/en-us/what-is-power-bi)

A NodeXL Pro INSIGHTS report makes use of many different data visualizations – from simple tables to hashtag clouds, from image grids to scatter plots - to enable insights into social media content streams. It makes it easy to pivot around several interconnected analytical dimensions in social media:

![Data Dimensions](https://www.smrfoundation.org/nodexl/nodexl-pro-insights)

\textsuperscript{86} [https://www.smrfoundation.org/nodexl/nodexl-pro-insights](https://www.smrfoundation.org/nodexl/nodexl-pro-insights)

\textsuperscript{87} [https://powerbi.microsoft.com/en-us/what-is-power-bi](https://powerbi.microsoft.com/en-us/what-is-power-bi)
Have a look at the sample report on the Twitter hashtag #AcademicTwitter⁸⁸ (best viewed in Chrome or Edge browsers) and learn more about NodeXL Pro INSIGHTS in the official documentation⁸⁹.

To create your own NodeXL Pro INSIGHTS report you need the following:

a. **NodeXL Pro User license**: A NodeXL Pro User license can be purchased on this page⁹⁰. NodeXL Pro runs on Windows and requires Office. Full system and software requirements for the NodeXL Pro App can be found in **chapter 1.2**, an installation guide is also available **here**⁹¹.

b. **Microsoft Power BI Pro Account**: A Microsoft PowerBI Pro⁹² account is required before subscribing to NodeXL Pro INSIGHTS. You may already have a Power BI user account. If you can log into the Microsoft Office Portal⁹³ you can then check to see if the Power BI application is one of the options listed in your dashboard.

---

⁸⁸ https://app.powerbi.com/view?r=eyJrIjoiOTA5ZWYxZDYtMDA0OS00ZGJiLWJkMTAtNGUzZTQwZWVkMzE2IiwidCI6IjI5ZDRjMTFjLTA1N2MtNDg3Zi04ZmRhLWU4NmQ1OTkzOWU2NCIsImMiOjZ9
⁸⁹ https://www.dropbox.com/s/u2088g5v5g7pvcb/NodeXL%20Pro%20INSIGHTS%20-%20User%20Guide.pdf
⁹⁰ https://nodexlgraphgallery.org/Pages/registration.aspx
⁹¹ https://www.smrfoundation.org/nodexl/installation
⁹² https://powerbi.microsoft.com/en-us/what-is-power-bi
⁹³ https://portal.office.com
c. Select “All apps” and then, if you already have PowerBI, it will be displayed in the list there.

If you do not have Power BI a two month free trial\(^94\) is available, or a monthly subscription for USD$10 a month. Students and non-profits\(^95\) may be able to get Power BI for a much lower monthly rate of $3 per month.

d. **NodeXL Pro INSIGHTS subscription**: A NodeXL Pro INSIGHTS subscription can be purchased on this page\(^96\). You can also purchase the NodeXL Pro User license in combination with an INSIGHTS subscription. NodeXL Pro INSIGHTS works with both the desktop and the cloud versions of NodeXL Pro.

---


\(^95\) [https://www.microsoft.com/en-us/nonprofits/power-bi](https://www.microsoft.com/en-us/nonprofits/power-bi)

\(^96\) [https://nodexlgraphgallery.org/Pages/registration.aspx](https://nodexlgraphgallery.org/Pages/registration.aspx)
e. **NodeXL Pro Twitter network dataset** that has been successfully processed by one of the Twitter “data recipes”. Learn how to do that in chapter 10, **Task Automation**.

**To export your network to NodeXL Pro INSIGHTS:**

1. In the Ribbon, select **NodeXL > Data > Export > To NodeXL Graph Gallery**....

![Export to NodeXL Pro INSIGHTS](image)

Figure: Export to NodeXL Pro INSIGHTS

2. Enter your Power BI **Username** and **Password** which are created during the registration process.
3. Optionally, if you would like to share the report with others or embed it on a public website, check the box “Export a public web copy of this data for sharing”. If selected, you will receive an email notification with a link and an embed code when the file has been published to the web. This step can also be done later from your **NodeXL Graph Gallery dashboard**.
4. Click **Publish to Power BI Service**.

When you click “Publish to PowerBI Service”, the data set is then exported to your Power BI workspace. This step may take several minutes depending on the file size. Your Power BI workspace is your personal storage place within the Power BI online service where your NodeXL Pro INSIGHTS reports are stored. You can also find a list of all links to all of your reports in your **NodeXL Graph Gallery dashboard**.

---

97 [https://nodexlgraphgallery.org/Pages/insights.aspx](https://nodexlgraphgallery.org/Pages/insights.aspx)
9. Working with Options

Options are the selections you make in NodeXL’s dialog boxes, in the NodeXL tab in the Excel Ribbon, and in the toolbar at the top of the Graph Pane. Some examples of options are the graph background color in the Graph Options dialog box, the graph type (Directed or Undirected) in the NodeXL tab in the Ribbon, and the layout algorithm at the top of the Graph Pane.

Here are some things to know about working with NodeXL’s options:

- When you change an option in a NodeXL workbook, the change is made only to that workbook. If you change the graph's background color to orange in the Graph Options dialog box, for example, the background color of your other NodeXL workbooks will not be affected. The background color of any new NodeXL workbooks you create afterward will not be affected, either.
- The options are stored inside the workbook. If you move the workbook to another computer or send it to someone else, the orange background will go with it.
- When you create a new NodeXL workbook, its initial options come from a hidden new workbook options file stored in your profile folder on your computer.
- You can tell NodeXL to use the current workbook’s options for new workbooks. If you do this, the current workbook’s options get copied to the hidden new workbook options file, and all new NodeXL workbooks you create afterward will have an orange background, for example.
- You can export a NodeXL workbook’s options to a NodeXL options file, and import a NodeXL options file into a NodeXL workbook.
- You can reset all of a NodeXL workbook’s options to the "new workbook options".
- One key feature of NodeXL Pro is 10. Task Automation. The tasks set in the Automate dialog box are also saved in the NodeXL Options file.

9.1 Using the Current Workbook’s Options for New Workbooks

Changing the options in a NodeXL workbook does not affect any new NodeXL workbooks you create afterward. If you want your changes to apply to new workbooks, you must take special action. To use the current workbook’s options for new workbooks:

1. In the Excel Ribbon, select **NodeXL > Options > Use Current for New**.
2. When you are asked for confirmation, answer "yes".
9.2 Exporting and Importing Options

You can export a NodeXL workbook's options to a NodeXL options file, and import a NodeXL options file into a NodeXL workbook. This lets you share options with other NodeXL users, for example, or keep multiple option files on hand to select from when you want to alter the appearance of a graph.

NodeXL options files have a ".NodeXLOptions" extension. You should not attempt to edit the file contents.

**To export the current workbook's options to a NodeXL options file:**

1. In the Excel Ribbon, select **NodeXL > Options > Export**.
2. In the **Export Options dialog box**, select a folder and file name for the NodeXL options file.

**To import a NodeXL options file into the current workbook:**

1. In the Excel Ribbon, select **NodeXL > Options > Import**.
2. In the **Import Options dialog box**, select a NodeXL options file to import.
3. Click Open.
4. Show the graph.

**Expert tip:** Learn about [10. Task Automation](#) with NodeXL options files. The tasks set in the **NodeXL > Graph > Automate dialog box** are also saved in the NodeXL Options file.

9.3 Resetting All Options

You can reset all of the current workbook's options to the set of options that are used for new workbooks.

**To reset all of the current workbook's options:**

1. In the Excel Ribbon, select **NodeXL > Options > Reset All**.
2. When you are asked for confirmation, answer "yes."
10. Task Automation

The most powerful feature of NodeXL Pro is **task automation** which is the key to fast results. When you find yourself repeatedly running the same sequence of tasks on a NodeXL workbook, you will find NodeXL’s automation feature very useful. With automation, you can run any or all of the following tasks with a single button click:

- Count and merge duplicate edges
- Create groups by cluster
- Calculate graph metrics
- Autofill columns
- Create subgraph images
- Show the graph
- Save an image of the graph
- Export the workbook

The following explanation builds upon NodeXL Options files that were presented in the previous chapter. These options files are also referred to as **NodeXL Data Recipes**.

### 10.1 What is a NodeXL Pro Options file?

Network datasets from different sources such as Twitter, YouTube, and Flickr contain different metadata and thus require customized approaches for analysis. A NodeXL Pro Options file is a customized “**data recipe**” that contains a list of specified tasks and options settings that will be applied to a network dataset.

#### Where to find NodeXL Pro Options files?

1. Download a [zip-file](#) with a collection of current data recipes [from our homepage](#), unzip the file and save the folder to your machine.
2. Or download single data recipes from network reports published on [NodeXL Graph Gallery](#) where NodeXL users can upload their network datasets including the options files.

#### To download NodeXL Pro Options files from the NodeXL Graph Gallery:

1. Go to [NodeXL Graph Gallery](#).
2. Open a network report of your choice and scroll down to the bottom.
3. If the author of the network report has chosen to upload the corresponding NodeXL workbook, you will see a few download links.
4. Click on “Download the NodeXL Options file Used to Create the Graph” and save the file to your machine.

#### To import a NodeXL Pro Options file into NodeXL:

1. Open a new NodeXL Pro workbook and click on [NodeXL > Options > Import](#) to open the Windows explorer.
2. Navigate to the folder where you have placed the previously downloaded data recipes.
3. Select the NodeXL Options file that corresponds with the network data you want to analyze.
4. Click Open.
5. To take a look at the tasks and options settings that were just imported, click on NodeXL > Graph > Automate to open the Automate window.
6. After reviewing the tasks and options click Cancel. (There is no data in the workbook to analyze yet.)

To automatically automate tasks after data is imported into the workbook:

1. In the Excel Ribbon, select NodeXL > Data > Import > Import Options.
2. In the Import Data Options dialog box, check the "Automate the graph after the data is imported" checkbox.

To automatically save the file to your local drive:

3. In the Import Data Options dialog box, click on the blue arrow behind "Automate the graph after the data is imported" to open the Automate Options dialog box.
4. Select the row “Save the workbook to a new file if it has never been saved” and select a file location where the workbook will be saved via the Options... button on the right.
5. Click OK in the Automate Options dialog box and then click OK in the Import Data Options dialog box.
6. Finally export the options file via NodeXL > Options > Export and save the file to the same options file you previously imported from your NodeXL Data Recipe folder. This way NodeXL will save the workbook in the same location every time you import the options file.
7. Optionally select NodeXL > Options > Use current for new so that NodeXL will always automatically import this options file when opening a new workbook.

Note: You can also import a data recipe, if data is already present in the NodeXL workbook.
10.2 Data Recipes

Here is an overview of the data recipes that are available in the above mentioned zip file\textsuperscript{a}.

10.2.1 Basic Data Recipes

The following two data recipes can be applied to any network data set from any source. They do not contain content analysis, but perform all necessary steps for a full-scale social network analysis: Calculate edge weight, group by cluster, calculate overall/group/vertex metrics, autofill columns and show the graph.

<table>
<thead>
<tr>
<th>Data Recipe</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Network Analysis 01 undirected</td>
<td>designed for undirected networks. Vertex size is based on Betweenness Centrality.</td>
</tr>
<tr>
<td>Social Network Analysis 02 directed</td>
<td>designed for directed networks. Vertex size is based on In-Degree.</td>
</tr>
</tbody>
</table>

**Expert tip:** Use these recipes to gain a quick overview of any network data set without any adjustments to the data recipe. Import! → Automate! → Run! → Done!

\textsuperscript{a} https://www.smrfoundation.org/nodexl/automation
10.2.2 X (Twitter) Data Recipes

**X Twitter User Network 01 – standard**

This recipe is designed for networks imported with all NodeXL Pro Twitter Data importers with a size of up to 5000 vertices. All relevant steps to conduct a full-scale social network and content analysis are performed. The graph shows an image and a label for every user in the network. This recipe works for **NodeXL Pro INSIGHTS**.

- **Vertex**: Twitter User
- **Edges**: Mentions, retweets, replies
- **Vertex size**: Betweenness centrality
- **Group clustering algorithm**: Clauset-Newman-Moore
- **Group labels**: Top 10 most frequently used hashtags
- **Content analysis**: Top words/word pairs/URLs/domains/hashtags
- **Time series analysis**: Yes
- **Sentiment language**: English
- **Layout**: Harel-Koren Fast Multiscale, Group-in-a-Box, Treemap

---

**X Twitter User Network 02 – large**

This recipe is designed for Twitter networks larger than 5,000 vertices and is very similar to the standard recipe shown above. All relevant steps to conduct a full-scale social network and content analysis are performed. The main difference to the standard recipe is Vertex sizing by In-Degree, and a label is shown only for users with an In-Degree larger than 5. This recipe works for **NodeXL Pro INSIGHTS**.

- **Vertex**: Twitter User
- **Edges**: Mentions, retweets, replies
- **Vertex size**: Indegree
- **Group clustering algorithm**: Clauset-Newman-Moore
- **Group labels**: Top 10 most frequently used words
- **Content analysis**: Top words/word pairs/URLs/domains/hashtags
- **Time series analysis**: Yes
- **Sentiment language**: English
- **Layout**: Harel-Koren Fast Multiscale, Group-in-a-Box, Treemap
This recipe is designed for Twitter networks larger than 10,000 vertices and is very similar to both recipes shown above. All relevant steps to conduct a full-scale social network and content analysis are performed. Vertices are sized by In-Degree, an image and a label is shown only for users with an In-Degree larger than 5. When working with very large networks, this will speed up the visualization time because not all user profile images are collected and rendered. This recipe DOES NOT work for NodeXL Pro INSIGHTS.

**Vertex:** Twitter User  
**Edges:** Mentions, retweets, replies  
**Vertex size:** In-Degree  
**Group clustering algorithm:** Clauset-Newman-Moore  
**Group labels:** Top 10 most frequently used hashtags  
**Content analysis:** Top words/word pairs/URLs/domains/hashtags  
**Time series analysis:** Yes  
**Sentiment language:** English  
**Layout:** Harel-Koren Fast Multiscale, Group-in-a-Box, Treemap

This recipe is also very similar to both recipes shown above. It just colors edges and vertices differently. In addition the layout is different: The grid layout is used within groups, and the group-in-a-box layout is force-directed. All relevant steps to conduct a full-scale social network and content analysis are performed. Vertices are sized by In-Degree, an image and a label is shown only for users with an Indegree larger than 5.

**Vertex:** Twitter User  
**Edges:** Mentions, retweets, replies  
**Vertex size:** In-Degree  
**Group clustering algorithm:** Clauset-Newman-Moore  
**Group labels:** Top 10 most frequently used words  
**Content analysis:** Top words/word pairs/URLs/domains/hashtags  
**Time series analysis:** Yes  
**Sentiment language:** English  
**Layout:** Grid layout, Group-in-a-Box, force-directed
Further data recipes customized for X (Twitter) data sets:

<table>
<thead>
<tr>
<th>Data Recipe</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentiment KO - Twitter User Network</td>
<td>This recipe is designed for networks collected with the NodeXL Pro Twitter Importers. The Word and Word pairs dialog box contains lists of positive, negative and stop words for sentiment analysis in Korean language. Reference: Prof. Han Woo Park <a href="http://www.hanpark.net/frontpage.asp?catalogid=hanpark&amp;language=ko">100</a>.</td>
</tr>
<tr>
<td>Tutorial - Sentiment 01 - positive</td>
<td>This data recipe is designed for X (Twitter) networks. It makes use of the column “Sentiment List #1: List1 Word Count” on the edges and vertices worksheets. Apply this data recipe to color vertices and edges with positive sentiment in green colors.</td>
</tr>
<tr>
<td>Tutorial - Sentiment 02 - negative</td>
<td>This data recipe is designed for X (Twitter) networks. It makes use of the column “Sentiment List #2: List2 Word Count” on the edges and vertices worksheets. Apply this data recipe to color vertices and edges with positive sentiment in red colors.</td>
</tr>
<tr>
<td>Tutorial - Sentiment 03 - sentiscore</td>
<td>This recipe needs additional steps before applying. Create a new column with the name “Sentiscore” on the edges and vertices worksheets. Then fill in the formula “=(@([Sentiment List ‘#1: List1 Word Count]) – @([Sentiment List ‘#2: List2 Word Count]))” The result is a color scheme that ranges from red (negative) to green (positive) based on the number of negative and positive words used.</td>
</tr>
</tbody>
</table>
### YouTube User Network 01 – standard

This recipe is designed for network data imported with the **YouTube User Network importer**. All relevant steps to conduct a full-scale social network and content analysis are performed. Content analysis is run on the **Comment** column. The graph shows an image and a label for every user in the network.

- **Vertex**: YouTube User
- **Edge**: Comment to video, reply to comment
- **Vertex size**: Betweenness centrality
- **Group clustering algorithm**: Clauset-Newman-Moore
- **Group labels**: Top 10 most frequently used words
- **Comment analysis**: Top words/word pairs/URLs/domains/hashtags
- **Time series analysis**: Yes
- **Sentiment language**: English
- **Layout algorithm**: Harel-Koren Fast Multiscale
- **Box layout algorithm**: Group-in-a-Box, Treemap

### YouTube User Network 02 – alternative layout

This recipe is designed for network data imported with the **YouTube User Network importer**. It works similar to the above recipe but uses a slightly different layout: The graph shows only a label for every user in the network with an In-Degree larger 5. Also, the group-in-a-box layout has been set to **force-directed**.

- **Vertex**: YouTube User
- **Edge**: Comment to video, reply to comment
- **Vertex size**: Indegree
- **Group clustering algorithm**: Clauset-Newman-Moore
- **Group labels**: Top 10 most frequently used words
- **Comment analysis**: Top words/word pairs/URLs/domains/hashtags
- **Time Series Analysis**: Yes
- **Sentiment language**: English
- **Layout algorithm**: Harel-Koren Fast Multiscale
- **Box layout algorithm**: Group-in-a-Box, force-directed
### YouTube Video Network 01 – description analysis

This recipe is designed for network data imported with the [YouTube Video Network importer](https://www.youtube.com). All relevant steps to conduct a full-scale social network and content analysis are performed. Content analysis is run on the [Video Description](https://www.youtube.com) column on the vertices worksheet. The graph shows an image for every video in the network.

- **Vertex**: Video
- **Edge**: Shared commenter
- **Vertex size**: Betweenness centrality
- **Group clustering algorithm**: Clauset-Newman-Moore
- **Text analysis**: Top words/word pairs in video description
- **Group labels**: Top 10 most frequently used word pairs in video description (Vertices)
- **Layout algorithm**: Harel-Koren Fast Multiscale
- **Box layout algorithm**: Group-in-a-Box, Treemap

### YouTube Video Network 02 – tag analysis

This recipe is designed for network data imported with the [YouTube Video Network importer](https://www.youtube.com). All relevant steps to conduct a full-scale social network and content analysis are performed. Content analysis is run on the [Video Tag](https://www.youtube.com) column of the vertices worksheet. The graph shows an image for every video in the network.

- **Vertex**: Video
- **Edge**: Shared commenter
- **Vertex size**: Betweenness centrality
- **Group clustering algorithm**: Clauset-Newman-Moore
- **Text analysis**: Top words/word pairs in video tag column
- **Group labels**: Top 10 most frequently used word pairs in video tag column (Vertices)
- **Layout algorithm**: Harel-Koren Fast Multiscale
- **Box layout algorithm**: Group-in-a-Box, Treemap
This recipe is designed for network data imported with the **YouTube Video Network importer**. All relevant steps to conduct a full-scale social network and content analysis are performed. Content analysis is run on the **Video1 Comment** column of the vertices worksheet. The graph shows an image for every video in the network.

10.2.4 Wikipedia Data Recipes

**Wikipedia Page Network 01 – standard**

This recipe is designed for Wikipedia Article-Article networks (1.5 degrees) imported with the MediaWiki Page Network importer. All relevant steps to conduct a full-scale social network analysis are performed. Content analysis is run on the Content column containing the first paragraphs of the respective page. The graph shows an image and a label for every page in the network.

- **Vertex:** Wikipedia article page
- **Edge:** link to mentioned articles on page
- **Vertex size:** Betweenness centrality
- **Group clustering algorithm:** Clauset-Newman-Moore
- **Group labels:** Top 10 most frequently used words
- **Sentiment language:** English
- **Layout algorithm:** Harel-Koren Fast Multiscale
- **Box layout algorithm:** Group-in-a-Box, Treemap

**Wikipedia Page Network 02 – large**

This recipe is designed to visualize large Wikipedia Article-Article networks (2.0 degrees) imported with the MediaWiki Page Network importer.

- **Vertex:** Wikipedia article page
- **Vertex size:** In-Degree
- **Vertex label:** skipped for low In-Degree.
- **Group clustering algorithm:** Clauset-Newman-Moore
- **Layout algorithm:** Harel-Koren Fast Multiscale
- **Box layout algorithm:** Group-in-a-Box, force-directed
This recipe is designed for Wikipedia User-User networks imported with the MediaWiki Page Network importer. Text analysis is run on the Comment column. The graph shows a disk for every user in the discussion.

| **Vertex:** | Wikipedia User |
| **Edge:**   | Comment        |
| **Vertex size:** | Betweenness centrality |
| **Group clustering algorithm:** | Clauset-Newman-Moore |
| **Group labels:** | Top 10 most frequently used words |
| **Text analysis:** | Top words/word pairs |
| **Layout algorithm:** | Harel-Koren Fast Multiscale |
| **Box layout algorithm:** | Group-in-a-Box, Treemap |
10.2.5 Flickr Data Recipes

**Flickr User Network**

This recipe is designed for network data imported with the **Flickr Users Network importer**. All relevant steps to conduct a full-scale social network analysis are performed. The graph shows an image and a label for every user in the network.

- **Vertex**: Flickr User
- **Edge**: contact, comment
- **Vertex size**: Betweenness centrality
- **Group clustering algorithm**: Clauset-Newman-Moore
- **Group labels**: Top 10 most frequently used words
- **Layout algorithm**: Harel-Koren Fast Multiscale
- **Box layout algorithm**: Group-in-a-Box, Treemap

---

**Flickr Tag Network**

This recipe is designed for network data imported with the **Flickr Related Tags Network importer**. All relevant steps to conduct a full-scale social network analysis are performed. The graph shows a sample image and a label for every tag in the network.

- **Vertex**: Flickr tag with sample image
- **Edge**: co-tag
- **Vertex size**: Betweenness centrality
- **Group clustering algorithm**: Clauset-Newman-Moore
- **Layout algorithm**: Harel-Koren Fast Multiscale
- **Box layout algorithm**: Group-in-a-Box, Treemap
### Reddit User Network 01 - standard

This recipe is designed for network data imported with the Reddit Search Network importer. All relevant steps to conduct a full-scale social network and content analysis are performed. Content analysis is run on the Text column.

- **Vertex:** Reddit User
- **Edge:** Post, comment, reply to comment
- **Vertex size:** In-Degree
- **Group clustering algorithm:** Clauset-Newman-Moore
- **Group labels:** Top 10 most frequently used words
- **Comment analysis:** Top words/word pairs/URLs/domains/hashtags
- **Time series analysis:** Yes
- **Sentiment language:** English
- **Layout algorithm:** Harel-Koren Fast Multiscale, Group-in-a-Box, Treemap

### Reddit User Network 02 - paths

This recipe is designed for network data imported with the Reddit Search Network importer. Just like above, all relevant steps to conduct a full-scale social network and content analysis are performed, but here the data is shown in the path layout.

- **Vertex:** Reddit User
- **Edge:** Post, comment, reply to comment
- **Vertex size:** In-Degree
- **Group clustering algorithm:** Clauset-Newman-Moore
- **Group labels:** Top 10 most frequently used words
- **Comment analysis:** Top words/word pairs/URLs/domains/hashtags
- **Time series analysis:** Yes
- **Sentiment language:** English
- **Layout algorithm:** Path
### Tutorial - Semantic Network - count

This recipe is designed for the tutorial [Semantic Networks – Create networks with words, hashtags or video tags](https://www.smrfoundation.org/nodexl/tutorials/semantic-networks/). All relevant steps to conduct a full-scale social network analysis are performed.

<table>
<thead>
<tr>
<th>Vertex:</th>
<th>Word/tag/hashtag</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edge:</td>
<td>directed word pair</td>
</tr>
<tr>
<td>Vertex size:</td>
<td>Betweenness centrality</td>
</tr>
<tr>
<td>Group clustering algorithm:</td>
<td>Clauset-Newman-Moore</td>
</tr>
<tr>
<td>Layout algorithm:</td>
<td>Harel-Koren Fast Multiscale</td>
</tr>
<tr>
<td>Box layout algorithm:</td>
<td>Group-in-a-Box, Treemap</td>
</tr>
</tbody>
</table>

---

11. Tutorial: Social network and content analysis - step by step

In this chapter we show briefly how you can easily perform a social network and content analysis step by step. All of the following steps (depending on the available metadata on the edges and vertices worksheets) are part of the data recipes mentioned in the previous chapter.

Step A: Data Import

Choose one of the data import options mentioned in chapter 3.

Step B: Data Preparation

In the Excel Ribbon, set NodeXL > Graph > Graph Type to Directed or Undirected. For details see chapter 4.1

To determine the Edge Weights, in the Excel Ribbon, select NodeXL > Data > Prepare Data > Count and Merge Duplicate Edges and set the options in the dialog box as shown in chapter 4.2.

Step C: Group by Cluster

The most common approach in social network analysis is to group the vertices based on a cluster algorithm. In the Excel Ribbon, select NodeXL > Analysis > Groups > Group by Cluster and choose a cluster algorithm. See chapter 5.3 for details.

Step D: Calculate Metrics

In the Excel Ribbon, select NodeXL > Analysis > Graph Metrics to open the Graph Metrics dialog box and select the metrics as shown in chapter 6.

Step E: Time Series Analysis

If your Edges worksheet contains a column in date format, open the Time Series dialog box via NodeXL > Analysis > Graph Metrics > Time Series > Options and point it to that column as explained in chapter 6.4.

Step F: Text and Sentiment Analysis

With the Words and Word Pairs dialog box, you can analyze any column that contains text either on the Edges or Vertices worksheet. Open the Word and Word Pair Metrics dialog box via NodeXL > Analysis > Graph Metrics > Words and Word Pairs > Options. Have a look at the details of this feature in chapter 6.6.
Step G: Network Top Items

The Network Top Items dialog box (NodeXL > Analysis > Graph Metrics > Network Top Items > Options) allows you to identify top items in any network by calculating the overall counts and the counts by group. The results can be used in the next step to label the groups in the network map. Check out chapter 6.8 to get the top items for X (formerly Twitter) data and chapter 6.9 for all other networks.

Step H: Autofill Columns

This feature will save you a lot of time! In the Excel Ribbon, to open the Autofill Columns dialog box select NodeXL > Visual Properties > Autofill Columns and customize your settings to populate columns on the Edges, Vertices and Groups worksheets as shown in chapter 7.4.

Step I: Customize Graph

There are many ways to customize the network graph. Use the Graph Options dialog box for basic adjustments of the layout. See chapter 7.1 for details.

Open the Layout Options dialog box to optionally set the layout to a Group-in-a-box treemap layout explained in chapter 7.5.2.

Further, customize the labels for groups, vertices and edges as shown in chapter 7.3 and chapter 7.4.

Step J: Save your Network Visualization

Right-click in the graph pane and set the Image Options. After that right-click again to save the image as explained in chapter 7.13.

Congratulations! You have successfully conducted a social network and content analysis! Now learn about Task Automation to automate all of the steps above as explained in chapter 10.
## 12. Keyboard Shortcuts in the Graph Pane

Several keyboard shortcuts are available in NodeXL's Graph Pane. Most of them are alternatives to using the Ribbon or the Graph Pane's right-click menu.

**Important Note:** You must click the Graph Pane before you can use any of these keyboard shortcuts.

<table>
<thead>
<tr>
<th>To Do This</th>
<th>Press This</th>
</tr>
</thead>
<tbody>
<tr>
<td>Show the graph</td>
<td>Ctrl+R</td>
</tr>
<tr>
<td>Lay out the entire graph without rereading the workbook</td>
<td>Ctrl+L</td>
</tr>
<tr>
<td>Select all vertices</td>
<td>Ctrl+V</td>
</tr>
<tr>
<td>Select all edges</td>
<td>Ctrl+E</td>
</tr>
<tr>
<td>Select all vertices and edges</td>
<td>Ctrl+A</td>
</tr>
<tr>
<td>Deselect all vertices and edges</td>
<td>Ctrl+D</td>
</tr>
<tr>
<td>Toggle the selection</td>
<td>Ctrl+T</td>
</tr>
<tr>
<td>Edit the properties of the selected vertices</td>
<td>Ctrl+P</td>
</tr>
<tr>
<td>Copy an image of the graph to the Clipboard</td>
<td>Ctrl+C</td>
</tr>
<tr>
<td>Save an image of the graph to a file</td>
<td>Ctrl+I</td>
</tr>
<tr>
<td>Move the selected vertices a small distance</td>
<td>Left Arrow, Right Arrow, Up Arrow, Down Arrow</td>
</tr>
<tr>
<td>Move the selected vertices a large distance</td>
<td>Shift+Left Arrow, Shift+Right Arrow, Shift+Up Arrow, Shift+Down Arrow</td>
</tr>
</tbody>
</table>
13. Frequently Asked Questions (FAQ)

13.1 General Questions

What is the difference between NodeXL Basic and NodeXL Pro?

NodeXL Basic is free and open. It is positioned as a browser for files created with NodeXL Pro which offers advanced features for professional social network and content analysis. For a detailed feature overview to compare NodeXL Basic and NodeXL Pro please follow this link.

Can I run NodeXL on a Mac?

NodeXL is Windows and Office only. See chapter 2.1 for details.

13.2 License

What are the differences between NodeXL Pro Commercial, Academic/Non-Profit and Student user licenses?

All NodeXL Pro licenses are for the same application, only the type of user (Student, Academic/Govt/Non-Profit, and Commercial) is different. The pricing structure reflects the ability of our users to pay for a license. Please select the license appropriate to you.

On how many machines can I run one license?

You may run NodeXL Pro on any number of machines, but only one at a time.

The NodeXL Pro License Key file we send to you must be placed on any machine you would like to use to run NodeXL Pro. Place the NodeXL Pro License Key file anywhere on the target machine’s file system.

When NodeXL Pro opens, it will ask you to locate this License Key file. This will then authorize the use of NodeXL Pro on that machine.

When a machine is authorized for NodeXL Pro all other machines that are running NodeXL Pro are de-authorized. NodeXL Pro will not operate on these machines until one of them is re-authorized.

To re-authorize a machine, copy the ORIGINAL license file (the one attached to the license email) to the target machine. DO NOT use the license file from any other machine.
When a license file is authorized, the copy of the license file is imprinted with machine specific data that prevents it from running on any other machine. Only the original license file can (re-)authorize a machine.

I ordered a NodeXL Pro license key but I did not receive an email with a license key. What should I do?

The original license key email is often sent within a minute of the order. If you do not receive a license key email promptly please check your “Spam” or “Junk” folders. You should also check to see the email address that you used to pay for your license key. In some cases the email associated with your payment is different from your main email address. If you cannot find the original license key in your emails, please send a short note to info@smrfoundation.org with the exact email address and the payment transaction ID used when purchasing the license.

I lost my license key. What should I do?

If you cannot find the original license key in your emails, please send a short note to info@smrfoundation.org with the exact email address you used when purchasing the license.

How do I cancel my NodeXL Pro monthly subscriptions?

The NodeXL Pro monthly subscriptions will automatically renew each month. You can always cancel the subscription from your PayPal account under this URL: https://www.paypal.com/cgi-bin/webscr?cmd=_manage-paylist.

13.3 Data Importers

What are the requirements and limits of the NodeXL Pro data importers?

Please check this website for updated information on the current NodeXL Pro data importers.

---

102 https://www.smrfoundation.org/nodexl/data-importers
13.4 Error Messages

Below you see a few error messages that may pop up. Further error messages and solutions can be found here\textsuperscript{103}. If you cannot solve your issue, send an email to info@smrfoundation.org.

“Invalid License” - What should I do?

This issue can be resolved by copying the original key attached to the delivery email back onto the target system. On some systems files can be set to be read only by default. If this happens, NodeXL Pro can then not update (and validate) its license key.

To resolve the issue:

1. Find the location of the “.LIC” file we sent to you. This is the original license key file sent as an attachment to an email.
2. Save this file to the location of the current license key file (you may have created a folder for this or left it in downloads or the desktop).
3. When you locate the NodeXL Pro License key file (“licensePro_.#.lic”) in the file browser, right-click it and select Properties.
4. Then uncheck “Read-Only”. This should then allow NodeXL Pro to read and update the license key.

If you cannot resolve this issue please send an email to info@smrfoundation.org and we will help you as soon as possible.

"OutOfMemoryException" - What does that mean?

If this message shows up, you have reached the limits of your machine. Please take a look at the system requirements above.

"Object reference not set to an instance of an object" - What does that mean?

Issue: NodeXL Crashes and displays “Object reference not set to an instance of an object”

A common cause for this issue is the location of the NodeXL Pro License key. Delete the current NodeXL Pro license key. Find the email that was sent with the initial NodeXL Pro license key attached. Copy the NodeXL Pro License Key to the local file system. A good place is a folder created on the desktop called “NodeXL Pro”. Restart NodeXL Pro. It will ask for the location of the NodeXL Pro License Key. Once that is provided, it should work properly.

There are several other possible solutions to this problem. If these steps do not resolve the issue, please install TeamViewer\textsuperscript{104} and send a message with the

\textsuperscript{103} https://www.smrfoundation.org/nodexl/installation
\textsuperscript{104} https://www.teamviewer.com
TeamViewer ID and PW to info@smrfoundation.org and we will fix the problem as soon as possible.

The "Document Actions" graph pane remains blank. How do I show the graph?

Please follow these steps to activate the “Document Actions” pane:

1. Open Excel.
2. Go to File → Options.
3. In the “General” tab, under “User interface options”, select the second option “Optimize for compatibility”.
5. Open NodeXL.

For more details please refer to this URL.105

13.5 Data handling

How large a network can NodeXL handle?

NodeXL can manage different numbers of edges based on the available computer system resources. The following are rough estimates:

4 GB of RAM: SMALL networks of only less than a few thousand edges
8 GB of RAM: MEDIUM networks of less than 10-15 thousand edges
16GB of RAM: LARGE networks of less than 80-100 thousand edges
32 GB of RAM: VERY LARGE networks of less than 200 thousand edges

How to deal with a VERY LARGE NETWORK?

Network analysis can require large amounts of computing resources or require long periods of time to perform (or both!).

Reducing the number of metrics calculated is one way to speed up analysis of a very large network. If you do not plan on using all the different types of “centrality” metrics, you may just want to skip creating them. To do this, modify the metrics selected in the NodeXL > Analysis > Graph Metrics dialog.

In many cases network data sets can be reduced in useful ways that retain many important features and INSIGHTS. In many networks it can be possible to count and then remove all “isolate” vertices (nodes with zero connections).

All “singleton” or “pendant” vertices can also be counted and removed (these are vertices with just one connection to the network).

All duplicated edges can be removed. Optionally, duplicate edges can be counted and removed and the count added to the remaining edge as a weight representing the number of removed edges. To perform this operation, see: NodeXL > Data > Prepare Data > Count and Merge Duplicate Edges.

How are Self-Loops managed in NodeXL?

A self-loop is created when a vertex links to itself. Self-loops appear as circles in NodeXL network visualizations. The display of self-loops can be controlled via the “Visibility” column in the Edges worksheet.

To reveal the “Visibility” column select the menu item:

NodeXL Pro > Show/Hide > Workbook Columns > Visual Properties

In the Edges worksheet, each edge has a “Visibility” column. In the “Visibility” column you can place a formula that checks if Vertex1 = Vertex2. If so, the formula places a “0” (Zero) in the cell. If not, the formula places a “1”. “0” (Zero) = “Skip” (which acts as if the data has been deleted). The Excel formula is:

=IF([@[Vertex 1]] = [@(Vertex 2)], 0, 1)

Then Refresh the graph and the self-loops are no longer displayed!

13.6 Research

How do I cite NodeXL in my research publication?

If you are publishing a paper that makes use of NodeXL, we would be happy to feature it in our list! To cite NodeXL in your papers please use the following:

14. Links and Literature

Below you find a selection of web links and publications related to NodeXL, the Social Media Research Foundation and social network analysis.

14.1 Important web links:

Social Media Research Foundation: http://www.smrfoundation.org

Get NodeXL or NodeXL Pro: https://nodexlgraphgallery.org/Pages/registration.aspx

Current NodeXL Pro Data Importers: https://www.smrfoundation.org/nodexl/data-importers/

NodeXL Graph Gallery: https://nodexlgraphgallery.org

NodeXL Pro INSIGHTS explained: https://www.smrfoundation.org/nodexl/nodexl-pro-insights

NodeXL Pro Data Recipes: https://www.smrfoundation.org/nodexl/automation

NodeXL EULA: https://www.smrfoundation.org/nodexl-pro-end-user-license-agreement-eula/


14.2 Major publications around NodeXL


14.3 Publications related to NodeXL features

**Cluster algorithms:**


**Layout algorithms:**


14.4 Publications with NodeXL as research tool

Below you find a short list of publications related to social network analysis with NodeXL. You can find many more publications when you search for NodeXL in Google Scholar\(^{106}\).


\(^{106}\) https://scholar.google.com/scholar?hl=de&as_sdt=0%2C5&q=NodeXL&btnG=


14.5 Recommended books and articles around social network analysis

