

Getting Started with Project Hook

User guide

Project Hook will her users explore data related to the fishing industry through 3 different lenses:



Network Risks



Financial Risks



Country Risks

Users can then build a summary heatmap based on customised factors to highlight key groups of interests.

Overview – Landing Page

The screenshot shows the landing page of a Shiny App titled "Project Hooked". At the top, there is a blue navigation bar with a "Menu" button and five links: "Introduction", "Network Risk", "Revenue Risk", "Country Risk", and "Customised Heatmap". The "Introduction" link is highlighted with a yellow box. Below the navigation bar, the main heading is "A Fishy Business". A welcome message follows: "Welcome to our Shiny App where you can explore the following modules:". Below this, a teal-bordered box contains three paragraphs describing the modules: "Financial Risks", "Country Risks", and "Network Risks". At the bottom of the page, there is a closing statement: "We hope your data exploration journey here leads you to meaningful insights. Together, let us get hooking!". Two callouts are present: a yellow one pointing to the navigation bar and a teal one pointing to the module descriptions box.

Easy and clear navigation to toggle between different modules on the Shiny App.

Project Hooked

Menu Introduction Network Risk Revenue Risk Country Risk Customised Heatmap

A Fishy Business

Welcome to our Shiny App where you can explore the following modules:

Financial Risks: With this module, users can investigate the financial aspects of the data, such as Total Revenue by subgraph, Total Revenue per company, and Total Revenue per Beneficial Owner. The objective is to identify suspicious clusters that may indicate illegal activities.

Country Risks: The primary goal of this module is to enable users to explore the data from a country perspective. It addresses questions like "Which subgraph spans across the most countries?" and "Are there any distinctive characteristics of these subgraphs?"

Network Risks: This module offers users the capability to visualize the interconnections within each subgraph, facilitating a deeper analysis of the relationships among companies, beneficial owners, and company contacts.

We hope your data exploration journey here leads you to meaningful insights. Together, let us get hooking!

Brief explanation on what each module entails.

Network Risks



Financial Risks – Basic Features



Key statistics to help anchor the data exploration process, for easy comparisons

Quick selection of specific subgraph of interest, to understand where they stand vis-à-vis other subgraphs, across variables

Options to customize the x-axis of the boxplot. Options include:

- Network Diameter
- Size of subgraph (number of nodes)

Options to customize the y-axis of the boxplot. Options include:

- Total Revenue (per subgraph)
- Revenue per company
- Revenue per beneficial owner

Project Hooked

Menu Introduction Network Risk Revenue Risk Country Risk Customised Heatmap

	Average Total Revenue per subgraph 348,500		Average revenue per company 262,000		Average revenue per beneficial owner 133,500
--	---	--	--	--	---

Subgraph of Interest

1

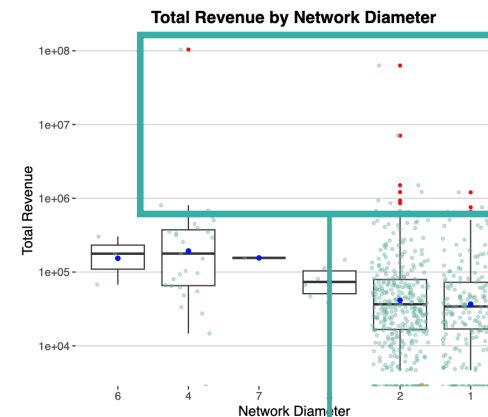
Select X-axis

Network Diameter

Select Y-axis (Revenue method)

Total Revenue

How does the revenue vary across subgraphs of different sizes? Click on the data points to reveal a comparison table on the right.



Outlier nodes are highlight in red.

Financial Risks – Exploring subgraphs



Project Hooked

Menu Introduction Network Risk Revenue Risk Country Risk Customised Heatmap



Average Total Revenue per subgraph

348,500



Average revenue per company

262,000



Average revenue per beneficial owner

133,500

Select boxplot parameters

Subgraph of Interest

55

Select X-axis

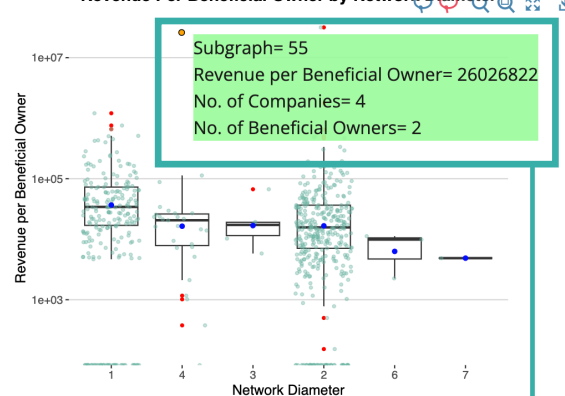
Network Diameter

Select Y-axis (Revenue method)

Revenue per Beneficial Owner

How does the revenue vary across subgraphs of different sizes? Click on the data points to reveal a comparison table on the right.

Revenue Per Beneficial Owner by Network Diameter



Selected Subgraphs

Subgraph Number	Revenue per Beneficial Owner
576	1205868

Node of selected subgraph will appear in orange, with a tooltip

Users are also able to select other nodes by clicking on the boxplot, on top of subgraph 55 to make comparisons. Information of any additional selected will appear in table on the right.

Country Risks – Basic Features



Project Hooked

Menu Introduction Network Risk Revenue Risk Country Risk Customised Heatmap

Posit Other links



Avg. Countries per Subgraph

1.16



Avg. Size per Subgraph

5.2 nodes



Avg. Revenue per Subgraph

414,424 OMU

Subgraph of Interest

102

Explore subgraph by:

Size (No. of nodes)

Number of Countries:

1

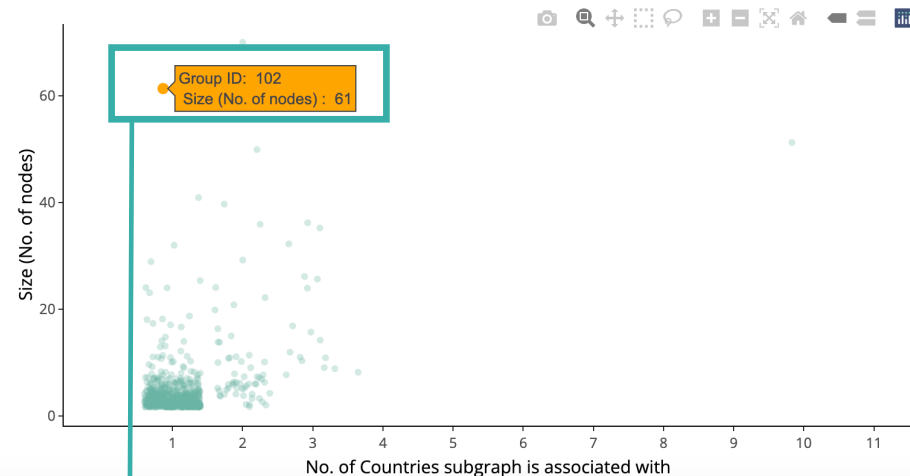
10

Size of subgraph (No. of nodes):

2

70

Subgraph characteristics by Country Association



Country(s)

Brindivaria

Key statistics to help anchor the data exploration process, for easy comparisons

Quick selection of specific subgraph of interest, to understand where they stand vis-à-vis other subgraphs, across variables

Options to customize the x-axis of the boxplot. Options include:

- Size of subgraph (number of nodes)
 - Number of companies
 - Number of beneficial owners
 - Total Revenue
 - Avg. revenue per company
 - Avg. revenue per beneficial owner

Option to filter x-axis range of scatterplot, based on number of countries subgraphs are associated with

Option to filter the data points in the scatterplot, based on size of subgraphs

Aligned with the financial risk module, selected subgraph of interest would be highlighted in orange, with relevant data presented in the tooltip

Country associated with subgraph presented in table

Customised Heatmap - Clustering

Navigate to optimization tab to fine tune the parameters in **orange** boxes

Customize x-axis by selecting indicators of interest. Options include:

- Total Revenue
- Revenue per Company
- Revenue per Beneficial Owner
 - Number of Countries
 - Number of Companies
- Number of Beneficial Owners
 - Network Diameter
 - Size of Subgraph

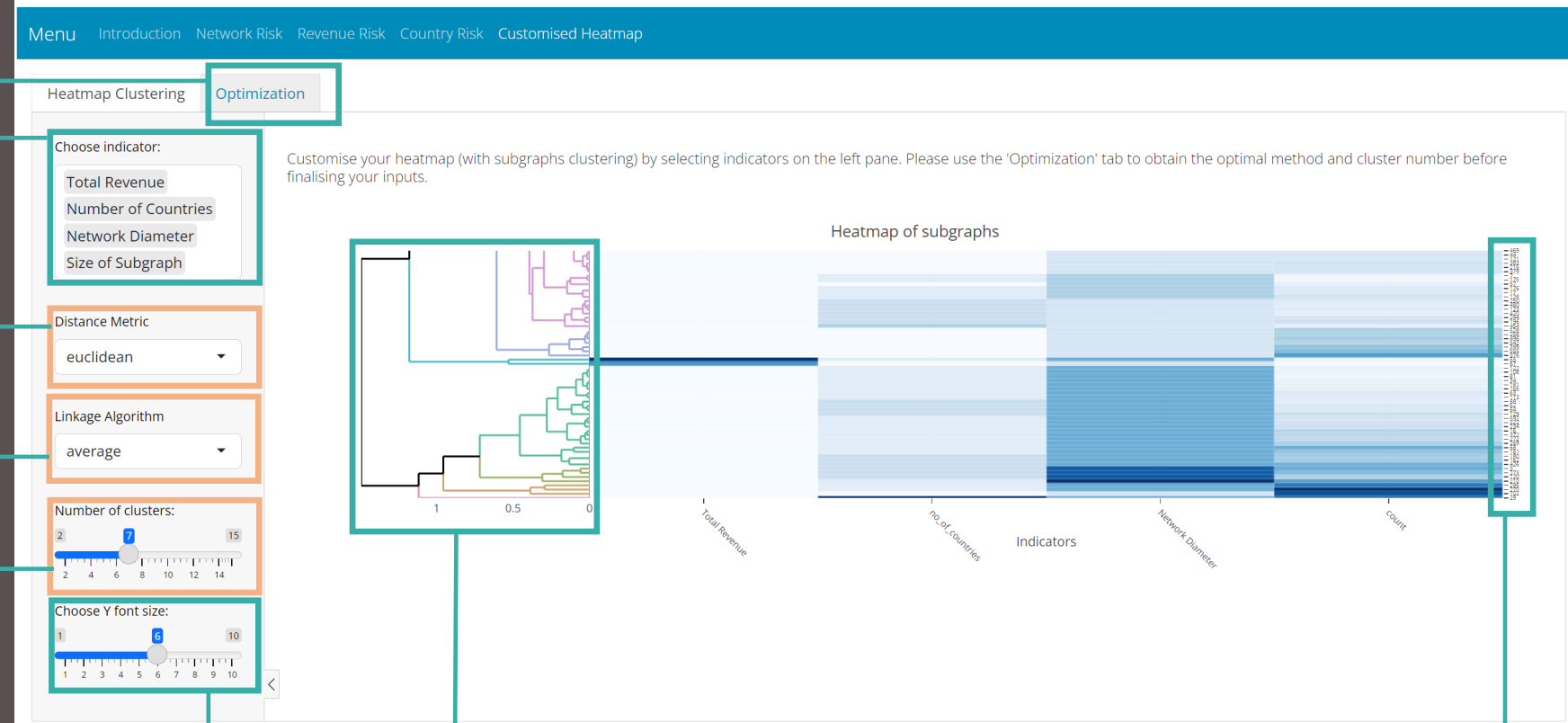
Select method used to compute distance matrix. Options include:

- euclidean
- maximum
- manhattan
- canberra

Select hclust method. Options include:

- ward.D
- single
- complete
- average
- mcquitty
- median
- centroid

Project Hooked



Specify number of clusters

Adjust size of Y axis labels

Dendrogram showing cluster hierarchy

Subgraph IDs

Customised Heatmap - Optimisation

Step 1:
Select indicators of interest. Options include:

- Total Revenue
- Revenue per Company
- Revenue per Beneficial Owner
- Number of Countries
- Number of Companies
- Number of Beneficial Owners
- Network Diameter
- Size of Subgraph

Step 3: Select 'centroid'

Project Hooked

Menu Introduction Network Risk Revenue Risk Country Risk Customised Heatmap

Heatmap Clustering Optimization

Choose indicator:

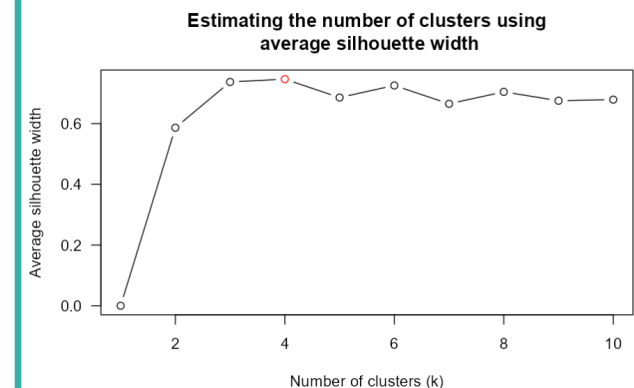
Revenue per Company
Number of Countries
Number of Companies
Number of Beneficial Owners
Network Diameter

Linkage Algorithm

centroid

Find highest optimal clustering method and optimal number of clusters

dist_methods	hclust_methods	optim
unknown	ward.D	0.43
unknown	ward.D2	0.74
unknown	single	0.86
unknown	complete	0.90
unknown	average	0.95
unknown	mcquitty	0.91
unknown	median	0.91
unknown	centroid	0.96



Step 2:
'centroid' method should be used because it gave the highest optimum value

Step 4:
Optimal number of clusters = 4