

Программа Gephi для сетевой визуализации

Плюсы программы:

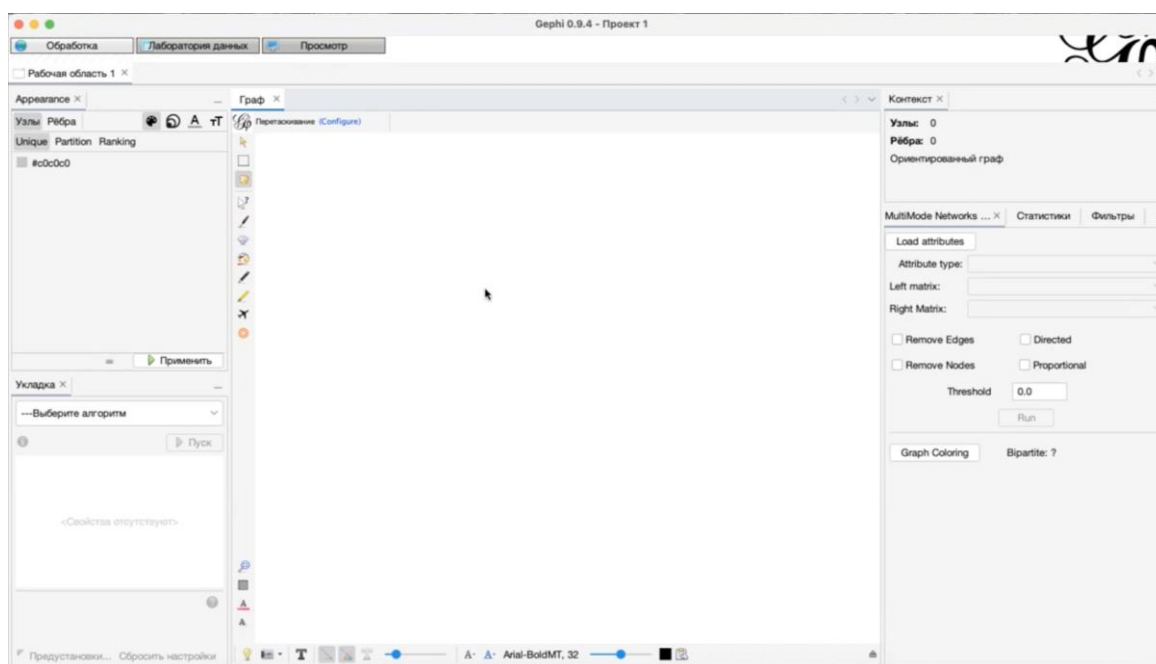
- Имеет набор базовых функций для сетевой визуализации
- Простота в использовании

Одномодальные сети:

Открытие файла:

В панели управления «Открыть файл» - выбор базы данных

Стартовое окно, открывающееся при создании проекта:



Важные вкладки сверху: «Обработка», «Лаборатория данных» и «Просмотр». Для работы сетью используется вкладка «Обработка».

Интерфейс «Обработки»:

Деление экрана:

- Вкладка «Граф»: визуализация графа
- Вкладка «Appearance»: внешний вид графа (размер вершин, цвета вершин и ребер, добавление цветов в названия)
- Вкладка «Укладка»: пространственный алгоритм, который определяет положение сети на экране
- Вкладка «Статистики»: средняя степень графа, диаметр графа, плотность графа, количество связанных компонентов и т.д.
- Вкладка «Фильтры»

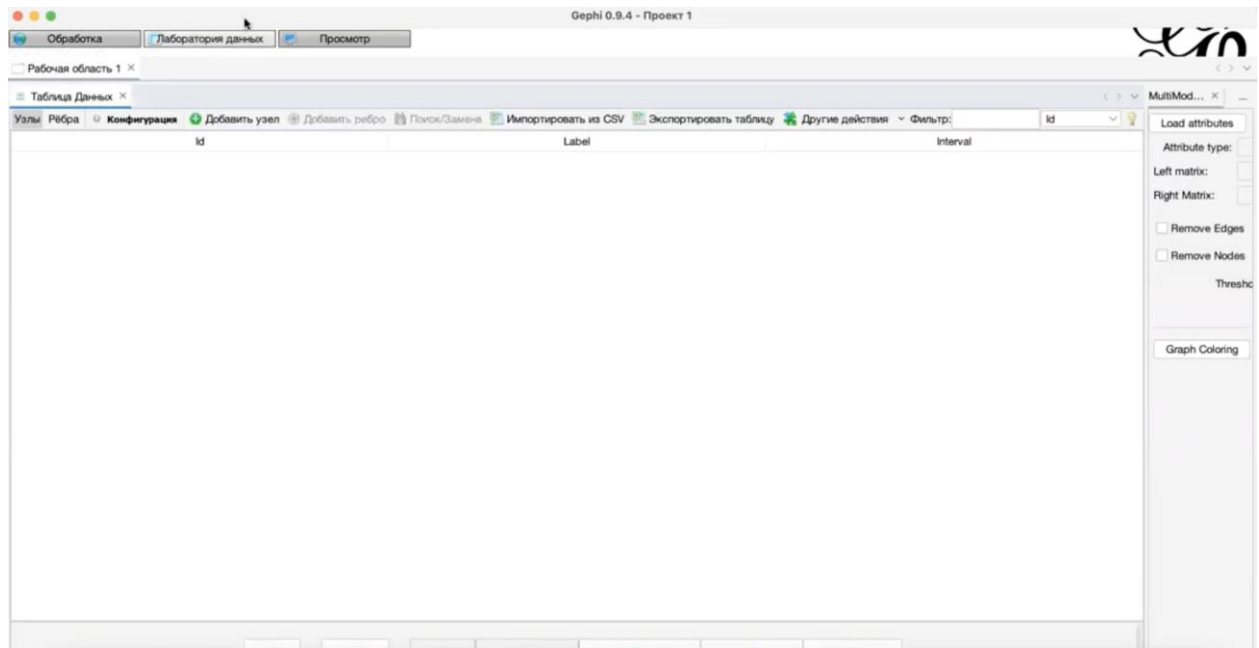
- Есть возможность добавления дополнительных плагинов, которые нужно устанавливать отдельно через вкладку «Сервис» - «Подключаемые модули» - «Доступные подключаемые модули»

После установки плагинов (нажать на галочку рядом с названием) – нажать «Активировать» - потом может потребоваться перезагрузить программу для продолжения работы

Интерфейс «Лаборатории данных»:

Представляет собой табличный формат

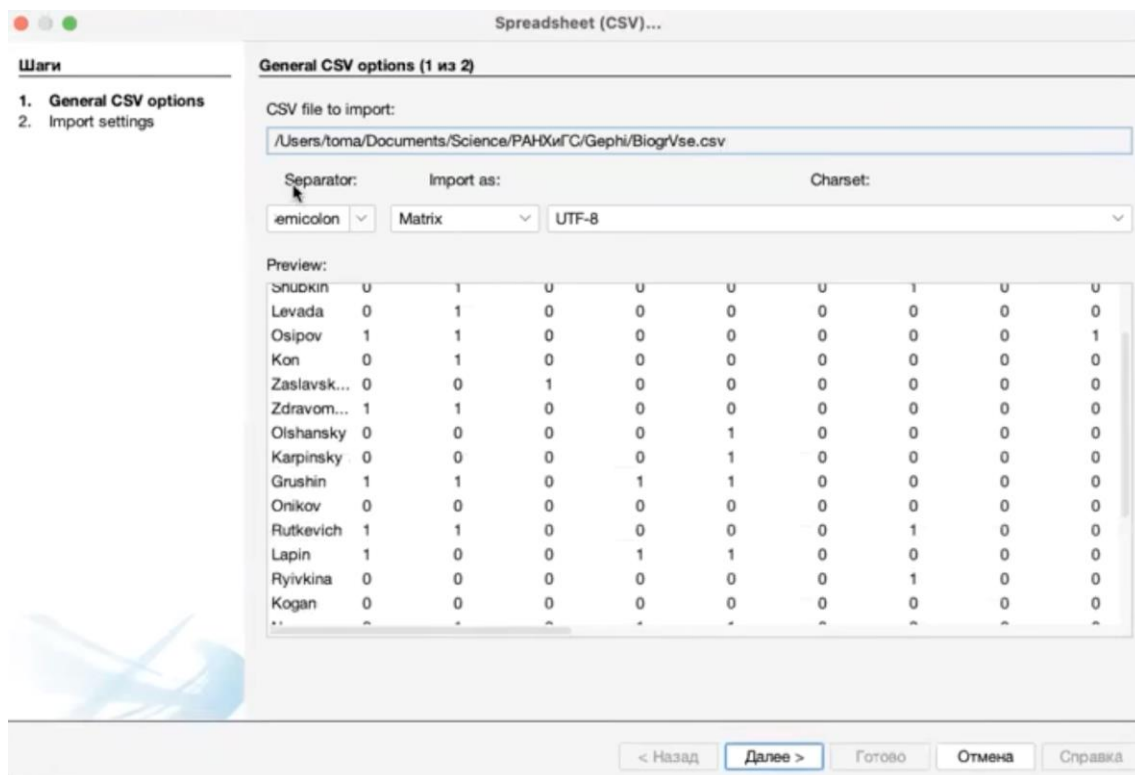
Импорт данных осуществляется путем нажатия кнопки «Импортировать из CSV», расположенной над рабочей областью



Предпочитаемые форматы файлов:

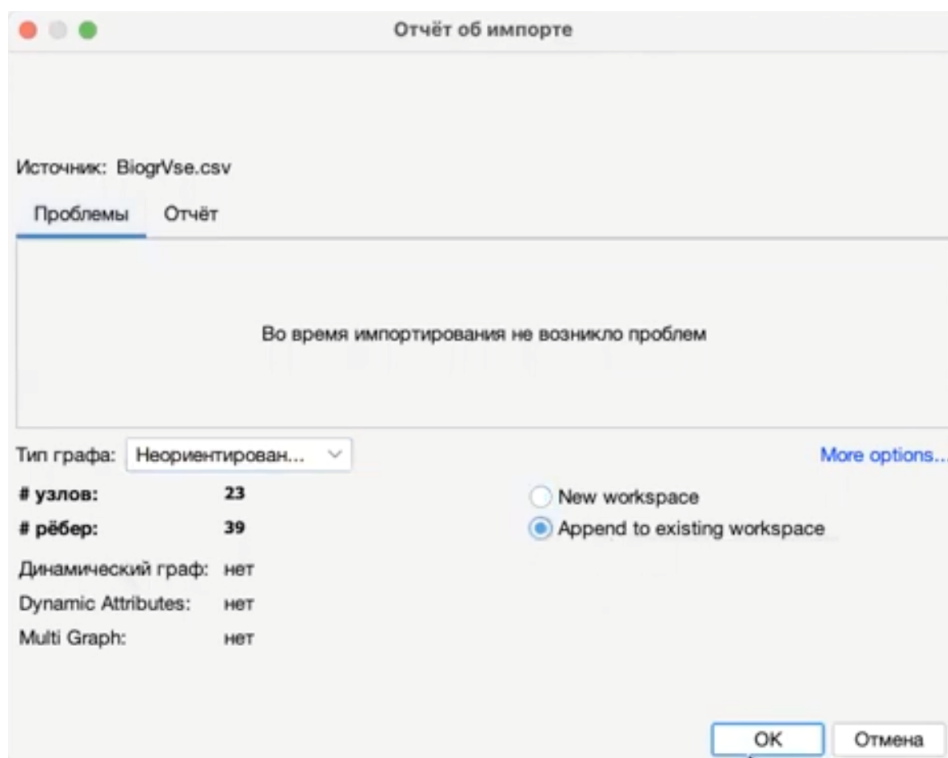
- .csv
- .gefx
- Модификации .gephi

При открытии данных появляется окно предпросмотра, в котором можно настроить вид импортируемых данных (проверка разделителя, типа данных, кодировки).



После импорта может появиться окно с предупреждением об ошибке, но на самом деле даже в таком случае данные могут открыться верно. На этом моменте можно увидеть количество узлов и ребер графа, тип графа.

В нижнем левом углу окна есть два варианта открытия графа: либо на новом пространстве, либо присоединение к уже существующему графу.



После добавления таблицы появляется возможность посмотреть узлы (по айди и лейблу) и ребра (по вершинам, информации о связи и весах) графа также в виде таблиц:

Узлы:

Узлы		
Ребра		
Конфигурация		
Добавить узел		
Добавить ребро		
Поиск/Замена		
Импортировать из CSV		
Экспортировать таблицу		
Другие действия		
Фильтр:		
Id		
Label		
Interval		
Kolbanovsky	Kolbanovsky	
Yadov	Yadov	
Osipov	Osipov	
Zdravomyslov	Zdravomyslov	
Grushin	Grushin	
Rutkevich	Rutkevich	
Lapin	Lapin	
Shubkin	Shubkin	
Levada	Levada	
Kon	Kon	
Naumova	Naumova	
Zaslavskaya	Zaslavskaya	
Galkin	Galkin	
Pilipenko	Pilipenko	
Olshansky	Olshansky	
Karpinsky	Karpinsky	
Bestugev-Lada	Bestugev-Lada	
Ryivkina	Ryivkina	
Onikov	Onikov	
Arab-Ogly	Arab-Ogly	
Kogan	Kogan	
Gordon	Gordon	
Semenov	Semenov	

Ребра:

Узлы							
Ребра							
Конфигурация							
Добавить узел							
Добавить ребро							
Поиск/Замена							
Импортировать из CSV							
Экспортировать таблицу							
Другие действия							
Фильтр:							
Source							
Target							
Type							
Id							
Label							
Interval							
Weight							
Kolbanovsky	Yadov	Неориентированное	0				2.0
Kolbanovsky	Osipov	Неориентированное	1				2.0
Kolbanovsky	Zdravomyslov	Неориентированное	2				2.0
Kolbanovsky	Grushin	Неориентированное	3				2.0
Kolbanovsky	Rutkevich	Неориентированное	4				2.0
Kolbanovsky	Lapin	Неориентированное	5				2.0
Yadov	Shubkin	Неориентированное	7				2.0
Yadov	Levada	Неориентированное	8				2.0
Yadov	Osipov	Неориентированное	9				2.0
Yadov	Kon	Неориентированное	10				2.0
Yadov	Zdravomyslov	Неориентированное	11				2.0
Yadov	Grushin	Неориентированное	12				2.0
Yadov	Rutkevich	Неориентированное	13				2.0
Yadov	Naumova	Неориентированное	14				2.0
Shubkin	Zaslavskaya	Неориентированное	16				2.0
Levada	Grushin	Неориентированное	18				2.0
Levada	Lapin	Неориентированное	19				2.0
Levada	Naumova	Неориентированное	20				2.0
Levada	Galkin	Неориентированное	21				2.0
Levada	Pilipenko	Неориентированное	22				2.0
Osipov	Olshansky	Неориентированное	25				2.0
Osipov	Karpinsky	Неориентированное	26				2.0
Osipov	Grushin	Неориентированное	27				2.0
Osipov	Lapin	Неориентированное	28				2.0
Osipov	Naumova	Неориентированное	29				2.0
Osipov	Bestugev-Lada	Неориентированное	30				2.0
Zaslavskaya	Rutkevich	Неориентированное	33				2.0
Zaslavskaya	Ryivkina	Неориентированное	34				2.0
Karpinsky	Grushin	Неориентированное	39				2.0
Karpinsky	Rutkevich	Неориентированное	40				2.0

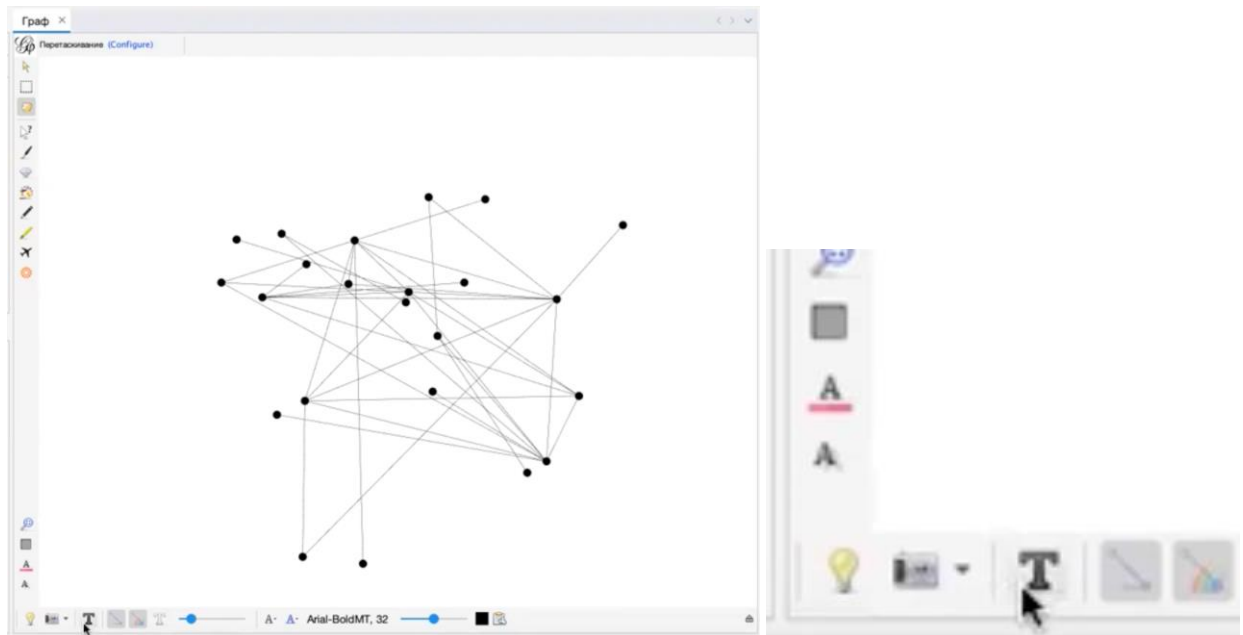
Интерфейс «Просмотра»:

Финальный рендеринг графа

Процесс работы с графом после загрузки табличных данных:

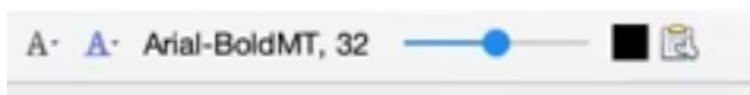
Работа с вкладкой «Обработка»:

- 1) Добавление лейблов



Черная буква «Т» для вершин, белая - «Т» для ребер

Размер надписей регулируется с помощью ползунка



2) Изменение положения вершин



Нажав знак «кулак», можно вручную захватить вершину и перетянуть ее на нужное место

ЛИБО

Прибегнуть к алгоритму «Укладки» - выбрать в зависимости от ситуации

Совет: при группировке вершин расстояние между точками желательно делать большим

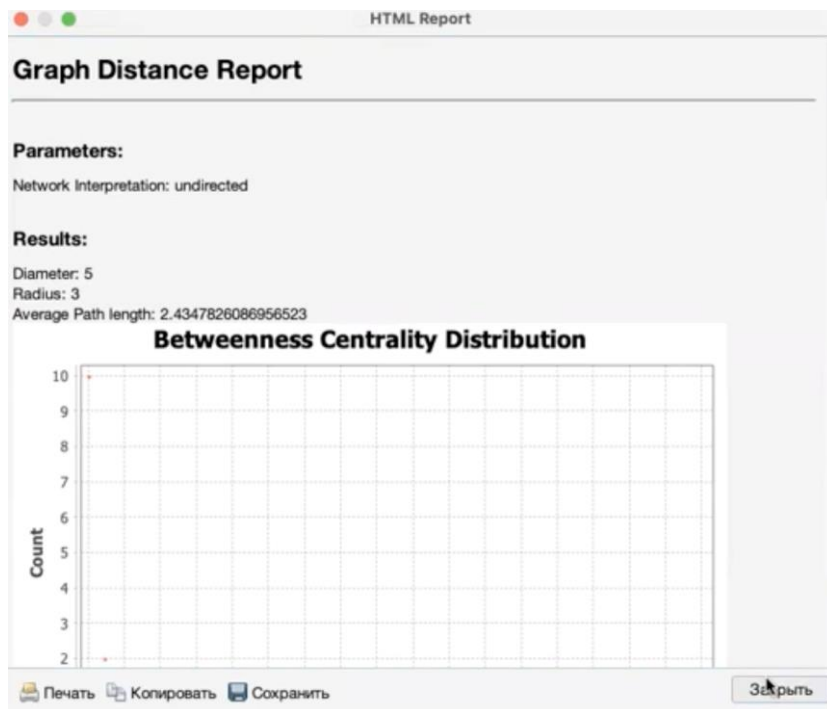
Алгоритмы «Укладки»:

- Yifan Hu: основан на принципах притяжения и отталкивания узлов, расположенных по соседству
- Fruchterman Reingold: организует граф в форме круга

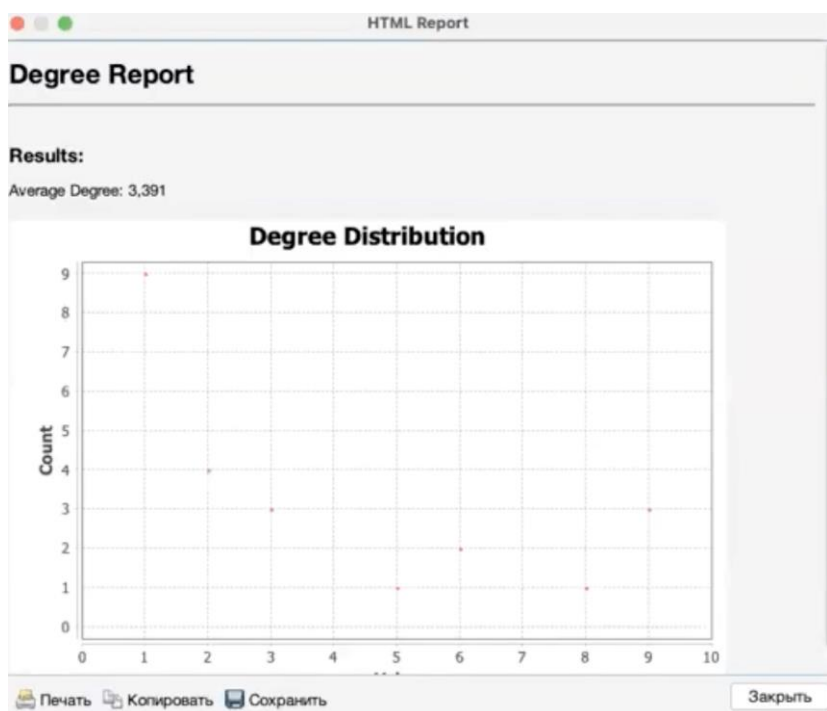
3) Анализ статистик

Используется вкладка «Статистики»

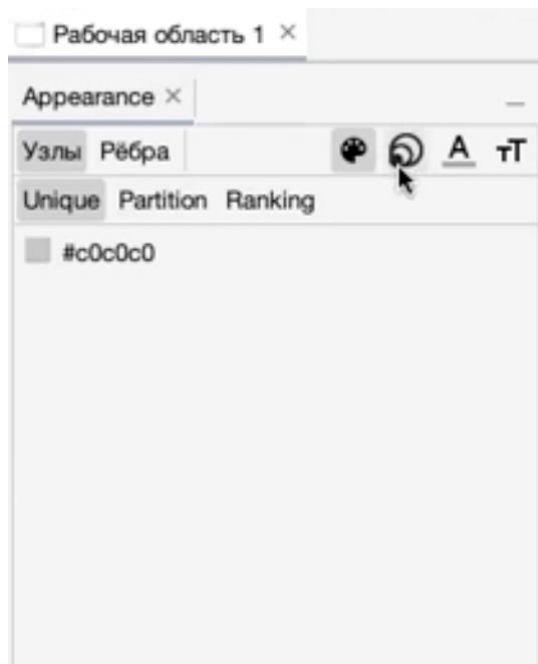
Диаметр графа:



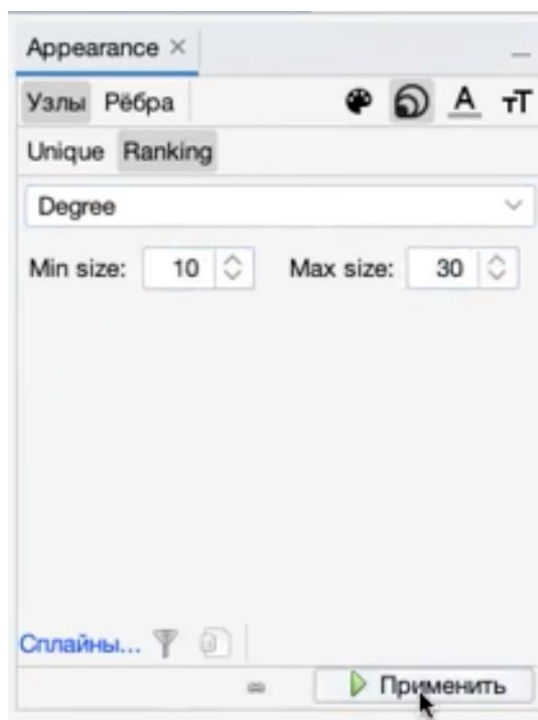
Средняя степень:



Чтобы отразить среднюю степень на графе (показать, у кого больше число взаимодействий), перейдем во вкладку «Appearance»



После этого переходим во вкладку «Ranking» и выбираем ранжирование по критерию Degree:



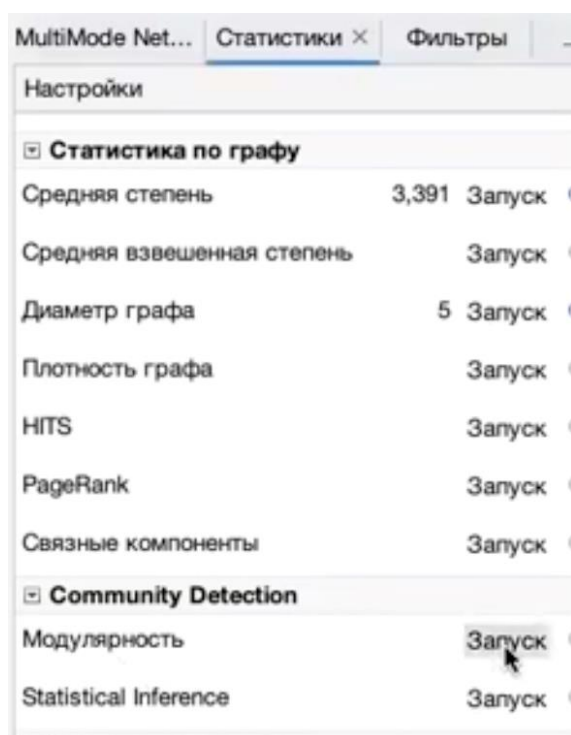
После установки параметров нажимаем «Применить»

В результате вид графа будет таким:



Размер вершин с наибольшими степенями увеличился.

Кластеризация:



Разбиение по кластерам для наглядности можно дополнить цветовым делением:

«Appearance» - выбираем значок «палитры» - «Partition» - «Modularity class» -

Применить



Результат, полученный после этих шагов, уже делает интерпретацию проще:



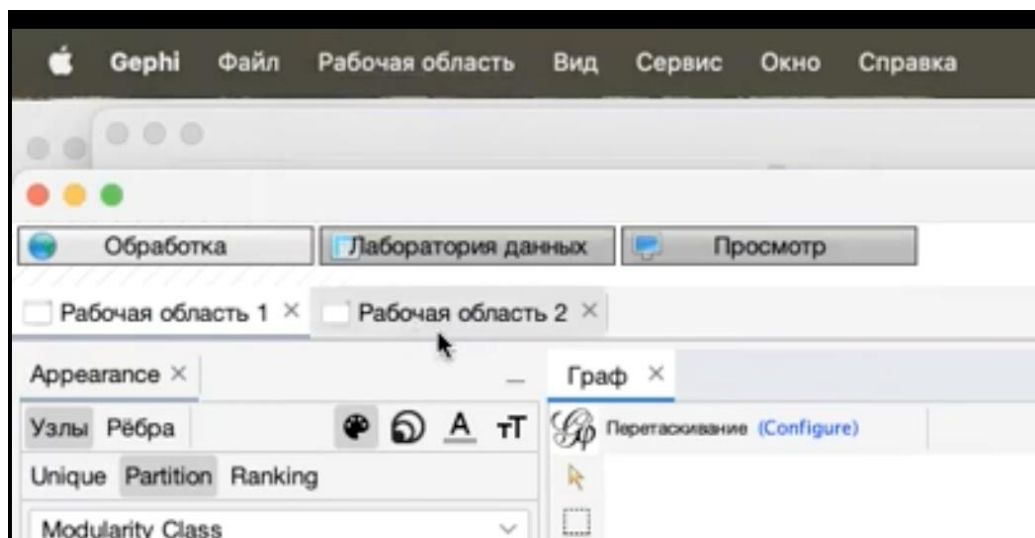
Полученные статистические выводы также доступны во вкладке «Лаборатория данных» - Узлы:

Рабочая область 1 ×									
Таблица Данных ×									
Узлы	Рёбра	Конфигурация	Добавить узел	Добавить ребро	Поиск/Замена	Импортировать из CSV	Экспортировать таблицу	Другие действия	Фильтр: Id
Id	Label	Interval	Eccentricity	Closeness Centrality	Harmonic Closeness Centrality	Betweenness Centrality	Суммарная мощность	Modularity Class	
Kolbanovsky	Kolbanovsky	3.0	0.536585	0.613636	11.033333	6	0		
Yadov	Yadov	3.0	0.611111	0.69697	69.216667	9	0		
Osipov	Osipov	4.0	0.5	0.625	46.483333	8	1		
Zdravomyslov	Zdravomyslov	4.0	0.4	0.450758	0.0	2	0		
Grushin	Grushin	3.0	0.611111	0.69697	56.033333	9	1		
Rutkevich	Rutkevich	3.0	0.564103	0.674242	85.483333	9	3		
Lapin	Lapin	3.0	0.52381	0.590909	11.8	5	1		
Shubkin	Shubkin	3.0	0.415094	0.462121	6.166667	2	3		
Levada	Levada	4.0	0.478261	0.579545	43.066667	6	2		
Kon	Kon	4.0	0.385965	0.420455	0.0	1	0		
Naumova	Naumova	4.0	0.44898	0.511364	2.25	3	2		
Zaslavskaya	Zaslavskaya	4.0	0.392857	0.469697	23.416667	3	3		
Galkin	Galkin	5.0	0.328358	0.372727	0.0	1	2		
Pilipenko	Pilipenko	5.0	0.328358	0.372727	0.0	1	2		
Otshansky	Otshansky	5.0	0.338462	0.387879	0.0	1	1		
Karpinsky	Karpinsky	3.0	0.468065	0.522727	3.3	3	1		
Bestuguev-Lada	Bestuguev-Lada	5.0	0.338462	0.387879	0.0	1	1		
Ryivkina	Ryivkina	5.0	0.285714	0.32803	0.0	1	3		
Onikov	Onikov	4.0	0.385965	0.420455	0.0	1	1		
Arab-Ogly	Arab-Ogly	4.0	0.4	0.450758	2.25	2	3		
Kogan	Kogan	4.0	0.366667	0.409091	0.0	1	3		
Gordon	Gordon	4.0	0.37931	0.439394	2.5	2	3		
Semenov	Semenov	4.0	0.366667	0.409091	0.0	1	3		

Создание новой рабочей области:

Вкладка на панели инструментов «Рабочая область» - «Создать рабочую область»

Результат (появилась «Рабочая область 2»):

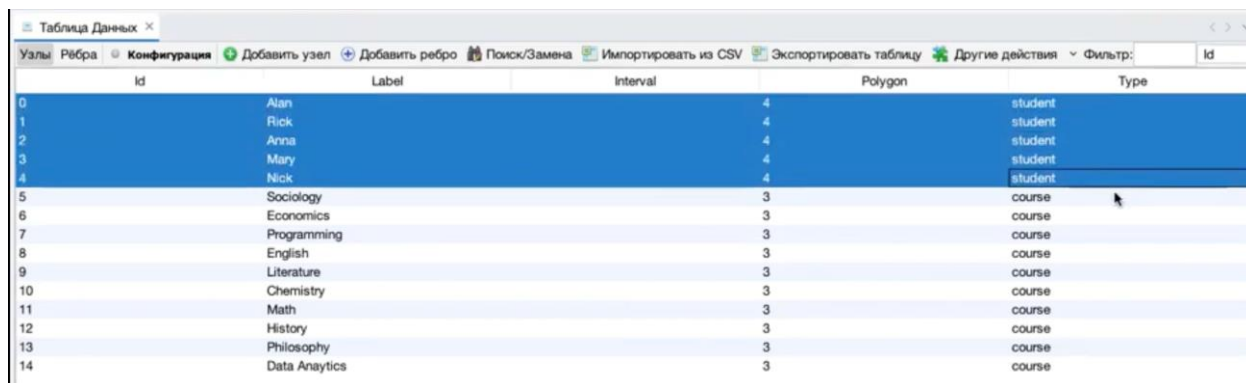


Двумодальные сети:

Установка дополнительных плагинов:

Вкладка на панели инструментов «Сервис» - «Подключаемые модули» - «Доступные подключаемые модули»

В таких сетях таблица узлов выглядит иначе:

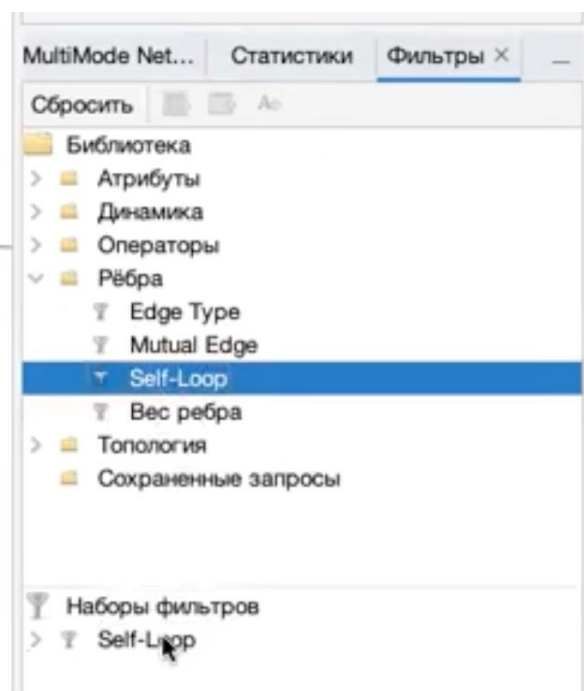


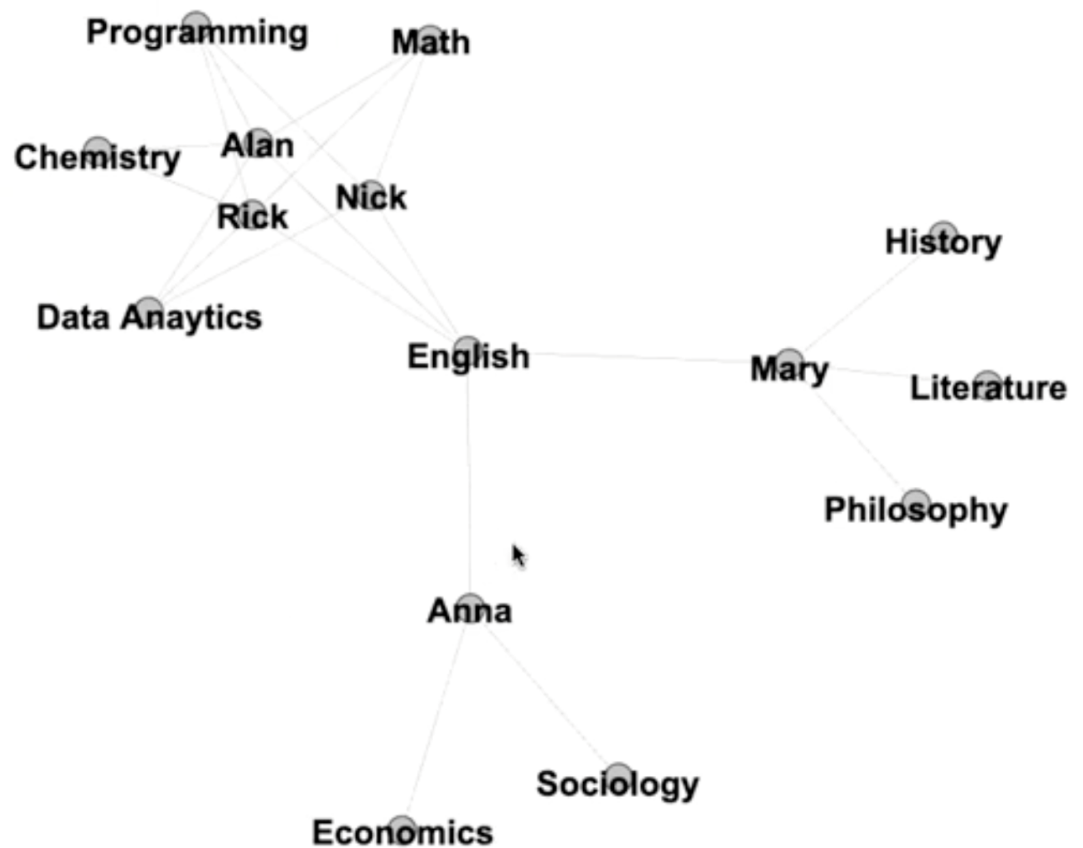
Id	Label	Interval	Polygon	Type
0	Alan	4		student
1	Rick	4		student
2	Anna	4		student
3	Mary	4		student
4	Nick	4		student
5	Sociology	3		course
6	Economics	3		course
7	Programming	3		course
8	English	3		course
9	Literature	3		course
10	Chemistry	3		course
11	Math	3		course
12	History	3		course
13	Philosophy	3		course
14	Data Analytics	3		course

У вершин есть атрибут, а также к каждой вершине прописан полигон (для этого нужен специальный плагин).

- Для работы с двумодальными сетями находим и скачиваем MultimodelNetworksTransformation
- Для изменения формы вершин находим и скачиваем Polygon Shape Notes

Иногда у вершин есть петли, от них рекомендуется избавляться, чтобы избежать искажения данных. Избавление от петель происходит так: Вкладка «Фильтры» - Ребра – Перетаскиваем фильтр «Self-loop» вниз – нажимаем «Отфильтровать»

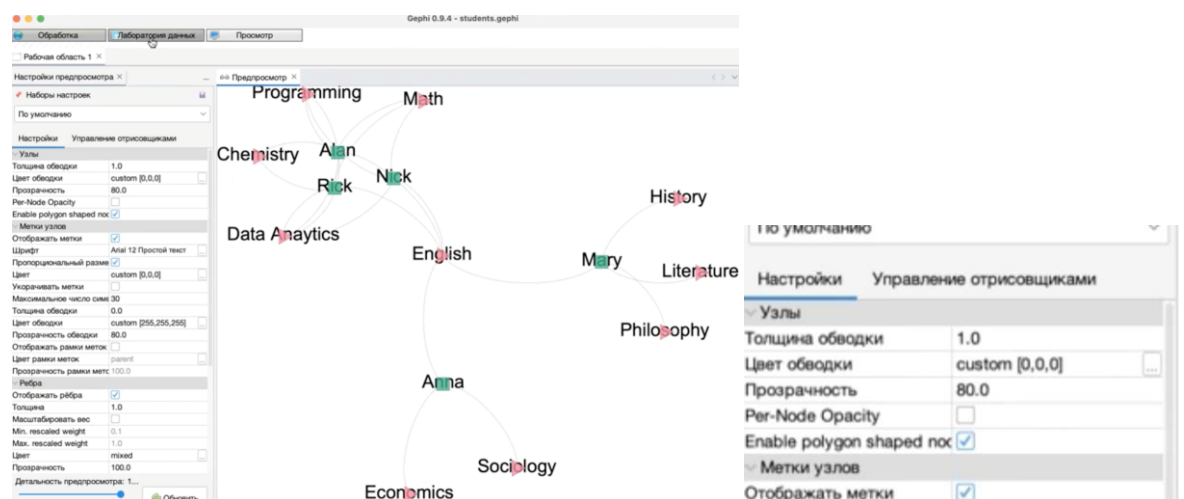




В этом графе все вершины визуально похожи, нужно сделать между ними различие, чтобы показать разницу в типах данных:

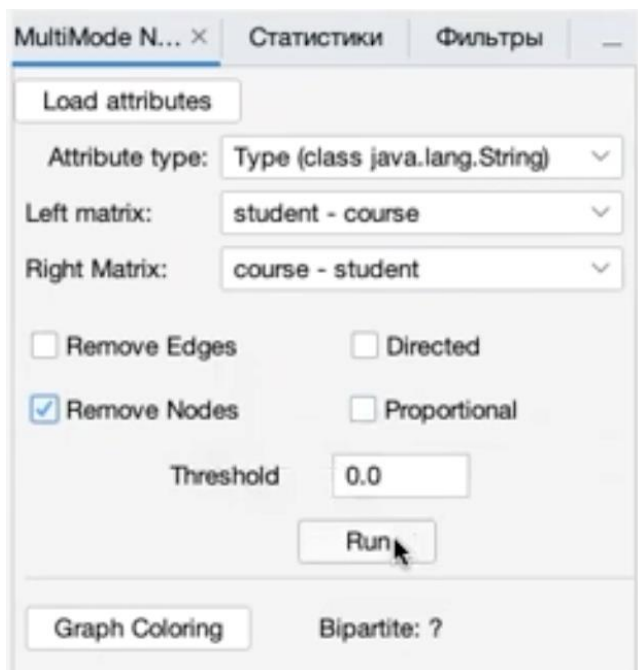
В базовой версии (без плагинов) можно просто по цветам рассортировать, но этого недостаточно

С плагинами: Во вкладке «Просмотр» - Настройки – Узлы – нажимаем «Enable polygon shaped» (Постобработка)



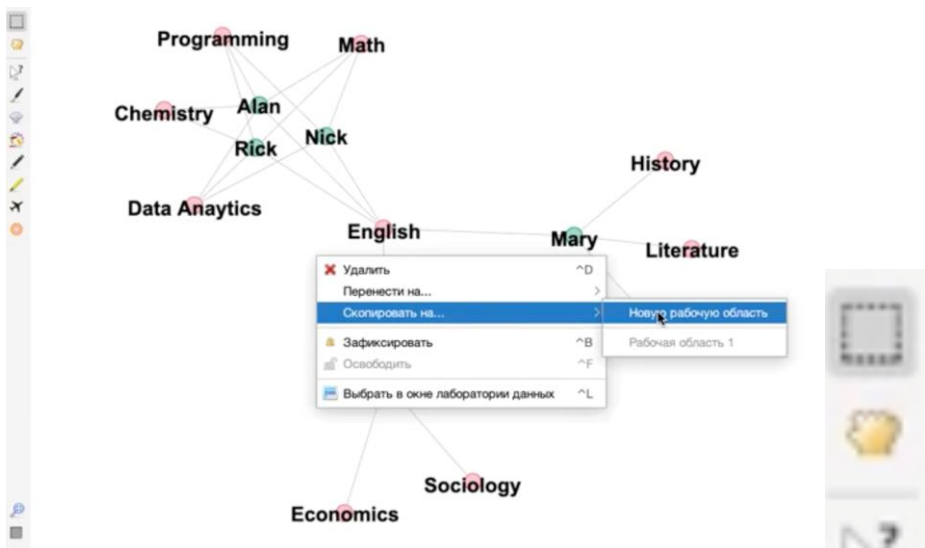
Количество углов в фигуре зависит от цифры в значении Туре для каждого студента в таблице

Можно обработать таким же образом граф и во вкладке «Обработка»:

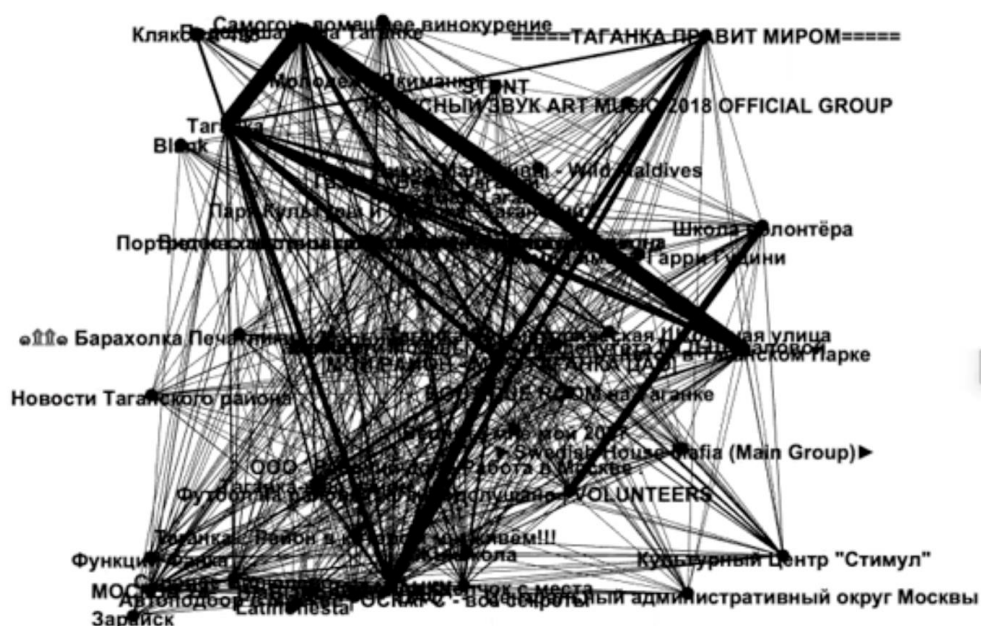


Нажимаем «Load attributes» - в Attribute type добавляем столбец из загруженной таблицы с данными – в Left Matrix и Right Matrix меняем пары значений – нажимаем «Remove nodes» чтобы остались только вершины одного типа – нажимаем «Run»

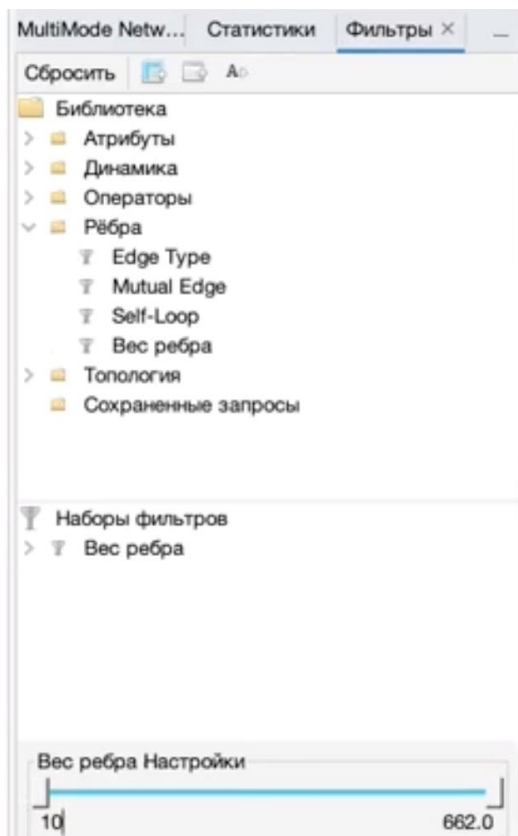
Чтобы созданный граф не потерялся, можно добавить его на отдельную вкладку:



Работа с Взвешенной сетью:



Если сеть очень зашумленная, можно повысить или понизить порог связности, чтобы не сильно значимые связи отпали: вкладка «Фильры» - Ребра – Перетаскиваем «Вес ребра» вниз в раздел «Набор фильтров» - Далее можно повысить минимальный вес ребра



Если названия в графе будут накладываться друг на друга, можно использовать вкладку «Укладка» - «Укладка меток»

Чтобы посмотреть уровень центральности нажимаем во вкладке «Appearance» - Узлы – Ranking – и выбираем Betweenness Centrality

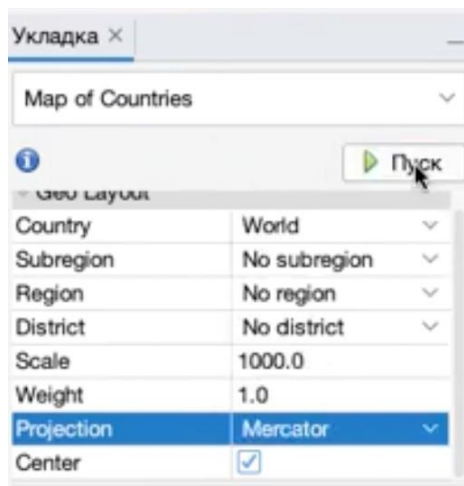
Неиспользованные вершины можно удалить, нажав на правую кнопку мыши прямо на графе, либо сделать это в «Лаборатории данных» (Считаем статистику средней степени вершины – переходим в «Лабораторию данных» и отфильтровать те страны, у которых 0)

Работа с двумя файлами для одного графа

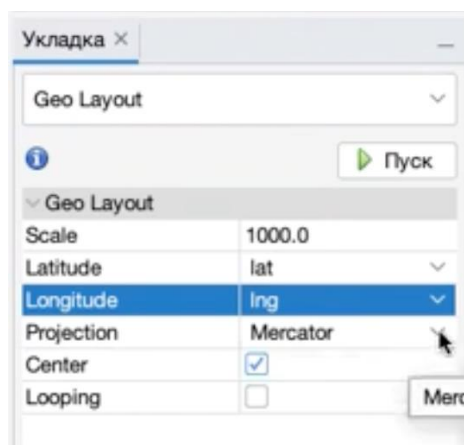
Атрибуты могут лежать в отдельном файле от файла с данными вершин и ребер, в таком случае их импортируем отдельно и добавляем в одно рабочее пространство.

Дополнительный плагин GeoLayout (позволяет расставить вершины по данным широты и долготы) и Map of Countries (позволяет добавить карту, чтобы точки визуально расставлялись по географическому положению)

Чтобы включить эти настройки заходим в «Укладку» - «Map of countries» - Выбираем в ячейке Country либо карту всего мира (World), либо отдельной страны – Выбираем в ячейке Projection вид карты



Далее в этой же вкладке ищем вместо Map of Countries GeoLayout и выстраиваем настройки:



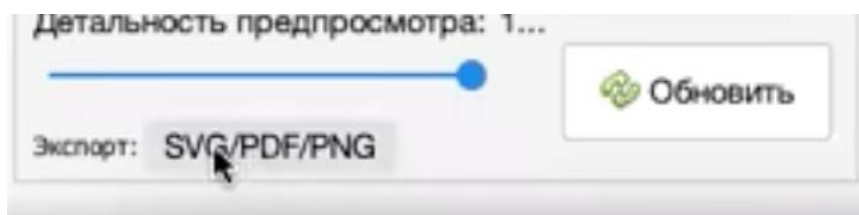
Можно поменять цвета вершин по количеству связей (чем больше связей, тем темнее цвет ребра)

В итоге вышел граф, соединяющий страны и показывающий их связь по какому-то заданному параметру:



Некоторые страны могут оказаться в неверных местах из-за наличия одноименных городов или штатов (напр. Georgia). Это можно поправить, исправив координаты той или иной страны в таблице в «Лаборатории данных»

Экспорт графа осуществляется через самое нижнее окно во вкладке «Просмотр»



При сохранении в формате .png можно выбрать дополнительные настройки, позволяющие удобно использовать полученный граф:

