

Отслеживание динамики изменения количества людей заразившиеся Covid-19 по всему миру с помощью программного пакета визуального программирования Orange

Голубева Евгения Павловна

Приамурский государственный университет имени Шолом-Алейхема

Студент

Аннотация

Цель данной статьи – отследить динамику изменения количества людей заразившиеся Covid-19 по всему миру. Для отслеживания динамики изменений количества людей, заразившиеся Covid-19 во всем мире, был использован программный пакет визуального программирования на основе компонентов для визуализации данных Orange и набор данных различных слов. С помощью средств визуализации Orange отследили динамику изменения количества людей заразившиеся Covid-19 по всему миру и получили итоговую схему.

Keywords: Orange, Covid-19, карта, динамика.

Tracking the dynamics of changes in the number of people infected with Covid-19 worldwide using the Orange visual programming software package

Golubeva Evgeniya Pavlovna

Sholom-Aleichem Priamursky State University

Student

Abstract

The purpose of this article is to track the dynamics of changes in the number of people infected with Covid-19 around the world. To track the dynamics of changes in the number of people infected with Covid-19 worldwide, a visual programming software package based on Orange data visualization components and a dataset of various words were used. With the help of Orange visualization tools, we tracked the dynamics of changes in the number of people infected with Covid-19 around the world and obtained the final scheme.

Keywords: Orange, Covid-19, map, dynamics.

1 Введение

1.1 Актуальность

Пандемия Covid-19 оказала масштабное глобальное воздействие, затронув практически все страны мира. Мониторинг ситуации и отслеживание динамики развития пандемии имеет жизненно важное

значение для принятия эффективных мер по борьбе с распространением вируса.

Официальные статистические данные по заболеваемости Covid-19 доступны в открытых источниках, однако их анализ и визуализация требуют использования специализированных инструментов. Программный пакет Orange предоставляет удобные средства для работы с данными и создания интерактивных аналитических панелей.

Использование визуального программирования в Orange позволяет создавать гибкие и масштабируемые аналитические решения, не требуя углубленных навыков программирования.

1.2 Обзор исследований

А. Радж, П. Бхаттачарья, Г.Р. Гупта автоматизировали кластерное обнаружение ежедневных новых случаев заболевания на душу населения во время первой волны в Индии [1]. Представили результаты предиктивного моделирования смертности пациентов одного из стационаров г. Уфы в 2020 г. на основе метода опорных векторов, алгоритмов построения деревьев решений и случайного леса В.Б. Прудников, Р.Х. Бахитова [2]. А.Ю. Попков, Ю.А. Дубнов, Ю.С. Попков посвятили работу применению метода рандомизированного машинного обучения для прогнозирования развития эпидемии COVID-19, основанной на эпидемиологической модели SIR. [3]. В статье представил результаты мультистранового анализа интенсивности протекания пандемии COVID-19 в 30 странах европейского региона на основе методов кластерного анализа и предлагаемых статистических индикаторов пандемии COVID-19 на страновом и мультистрановом уровне А.К. Корниевич [4]. А.Г. Карабутова рассмотрела использование методов машинного обучения при изучении распространения коронавирусной инфекции (COVID-19) в России [5].

1.3 Цель исследования

Цель исследования - отследить динамику изменения количества людей заразившихся Covid-19 по всему миру.

2 Материалы и методы

Для отслеживания динамики изменения количества людей, заразившихся Covid-19 по всему миру, используется программа Orange. Работа будет происходить на готовом наборе данных о Covid-19, скачать которые можно по ссылке:

https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_global.csv

3 Результаты и обсуждения

Перед началом работы требуется установить Orange с официального сайта и установить.

Создадим новый файл (см.рис.1).

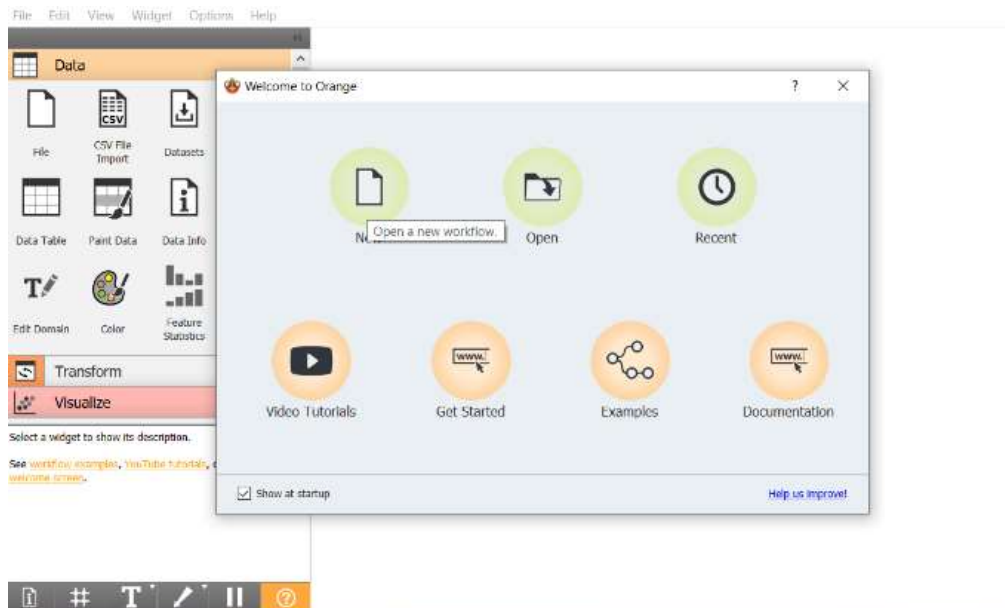


Рисунок-1 Создание нового файла

Для отслеживания динамики изменения количества заболеваний Covid-19 по всему миру необходимо установить дополнение Geo и Timeseries. Для того, чтобы скачать дополнение, необходимо перейти в Options, далее в Add-ons, в появившемся окне выбираем Geo и Timeseries (см.рис.2).



Рисунок-2 Установка дополнения Geo и Timeseries

Для того, чтобы загрузить набор данных о Covid-19, необходимо из раздела Data выбрать виджет File на холст и перенести его на холст (см.рис.3).



Рисунок - 3 Добавление виджета File на холст

Для того, чтобы подключить данные о Covid-19 необходимо вставить ссылку в URL (см.рис.4).

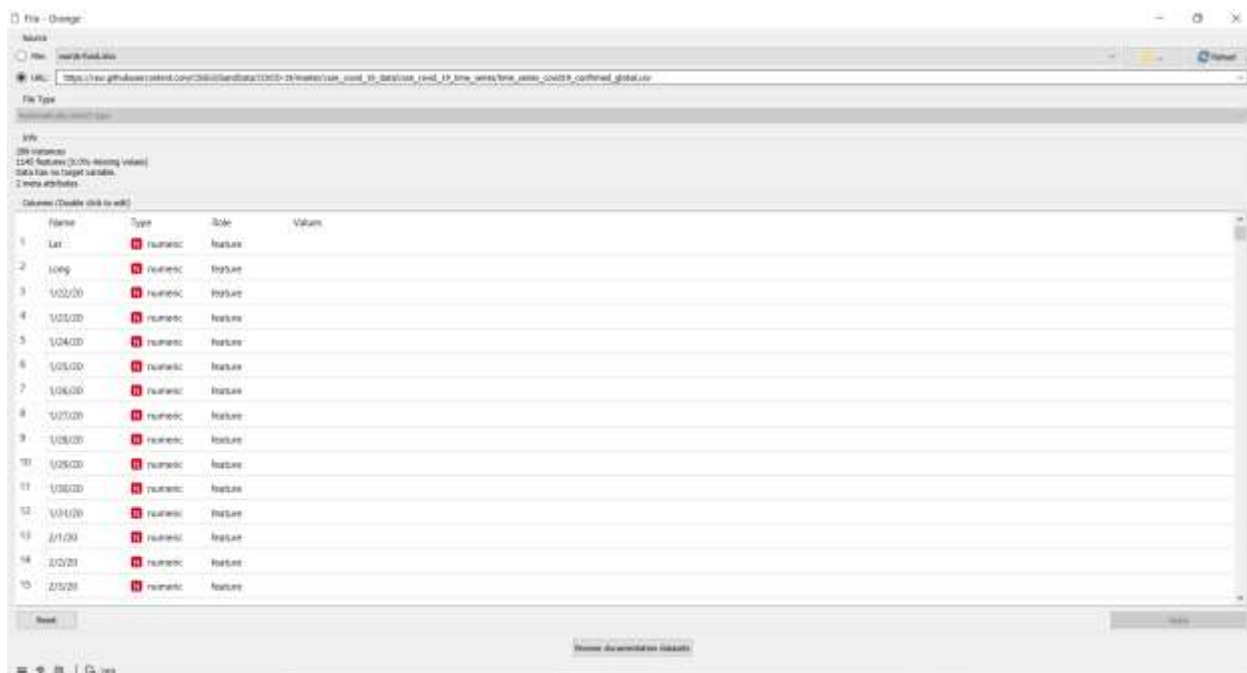


Рисунок - 4 Подключение данных о Covid-19

Для отслеживания динамики изменения количества заболеваний Covid-19 по всему миру будем использовать временной ряд, который будет отображаться на карте.

Для начала преобразуем набор данных. Для этого добавим виджет Select Columns на холст и соединим с виджетом File. Виджет Select Columns используется для ручного выбора атрибутов данных и состава домена данных (см.рис.5).

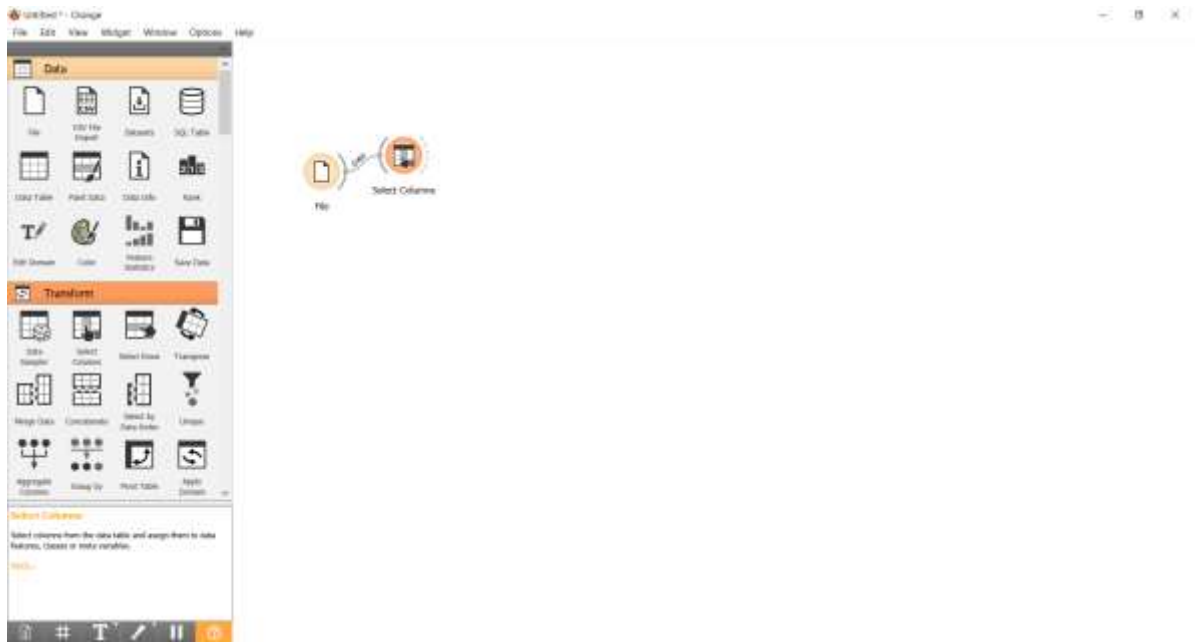


Рисунок - 5 Добавление виджета Select Columns на холст

Откроем виджет Select Columns, и перенесем столбцы Lat и Long из колонки Features в колонку Metas (см.рис.6).

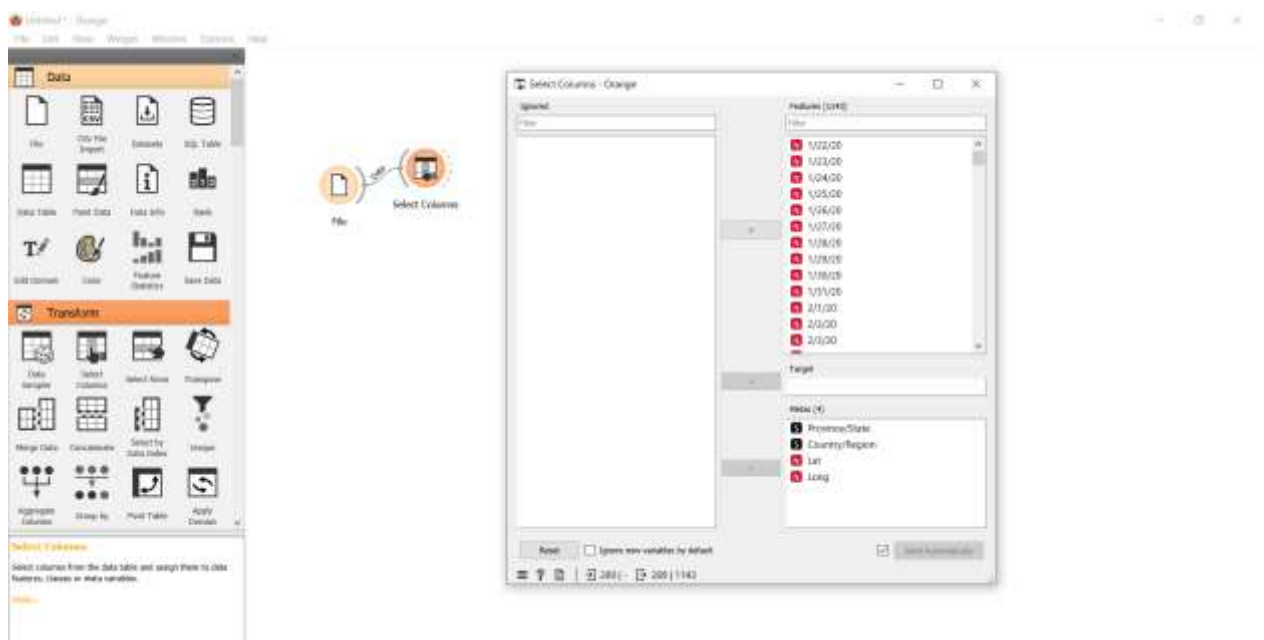


Рисунок - 6 Изменение атрибутов столбцов

Далее необходимо поменять местами столбцы и строки. Для этого используем виджет Transpose и соединяем с виджетом Select Columns (см.рис.7).

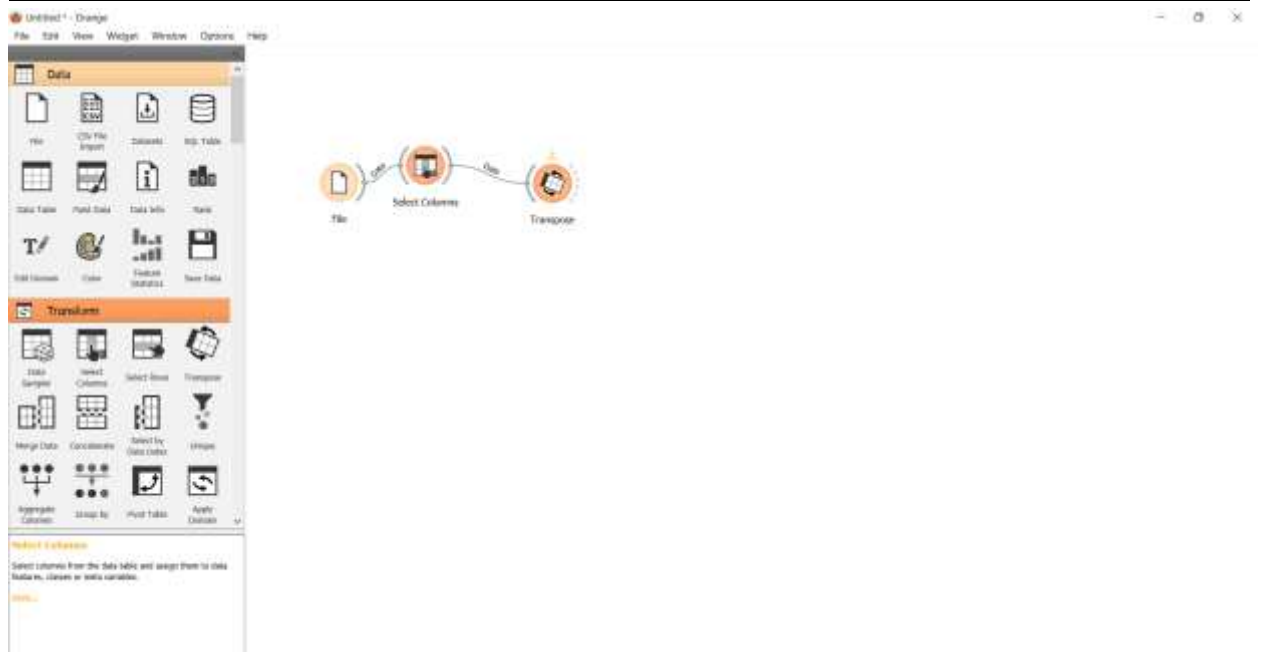


Рисунок - 7 Добавление виджета Transpose на холст

Открываем виджет Transpose и задаем для переменной From variable значение Country/Region (см.рис.8).



Рисунок - 8 Диалоговое окно виджета Transpose

Для того, чтобы посмотреть, как изменились столбцы и строки в наборе данных добавим виджет Data Table на холст и соединим с виджетом Transpose (см.рис.9).

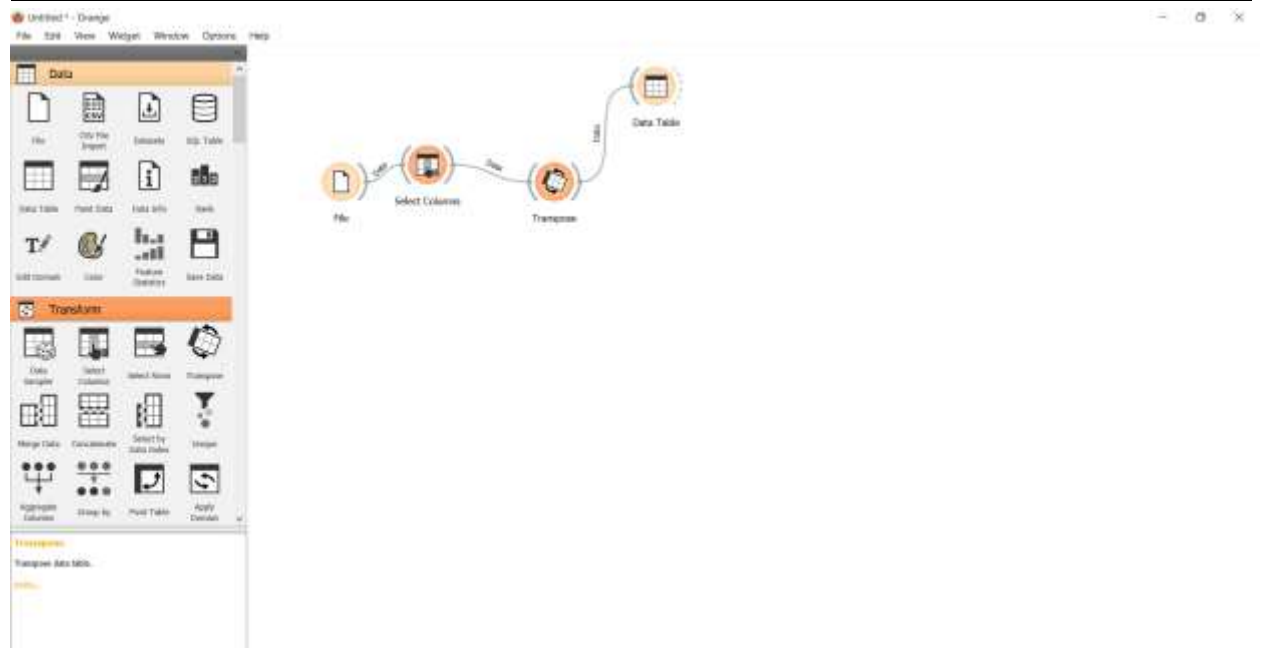


Рисунок - 9 Добавление виджета Data Table на холст

Открываем виджет Data Table, и можем увидеть, что столбцы названы в соответствии со странами, а новый столбец содержит исходные данные даты (см.рис.10).

The screenshot shows the 'Data Table' widget in Orange. The table displays data with columns for 'Country/Region', 'Feature name', and 'Feature ID'. The 'Country/Region' column lists various countries and regions, and the 'Feature name' column lists dates. The 'Feature ID' columns are numbered from 001 to 011. The table contains 33 rows of data.

	Country/Region	Feature name	Feature ID0	Feature ID1	Feature ID2	Feature ID3	Feature ID4	Feature ID5	Feature ID6	Feature ID7	Feature ID8	Feature ID9	Feature ID10	Feature ID11
1	Lat	1/1/2000	22.4391	46.1333	30.8339	40.5363	11.2827	72.9499	17.9908	-28.4363	40.0900	348.9123	30.6688	182.209
2	Long	1/1/2000	87.71	26.1889	1.6999	5.0228	87.8739	23.947	45.7966	-63.6267	-46.2282	148.9123	182.209	182.209
3		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
4		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
5		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
6		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
7		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
8		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
9		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
10		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
11		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
12		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
13		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
14		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
15		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
16		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
17		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
18		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
19		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
20		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
21		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
22		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
23		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
24		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
25		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
26		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
27		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
28		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
29		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
30		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
31		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
32		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0
33		1/1/2000	0	0	0	0	0	0	0	0	0	0	0	0

Рисунок - 10 Просмотр данных Data Table

Для того, чтобы столбец Feature name принял тип время, необходимо добавить виджет Edit Domain, и соединить с виджетом Transpose (см.рис.11).

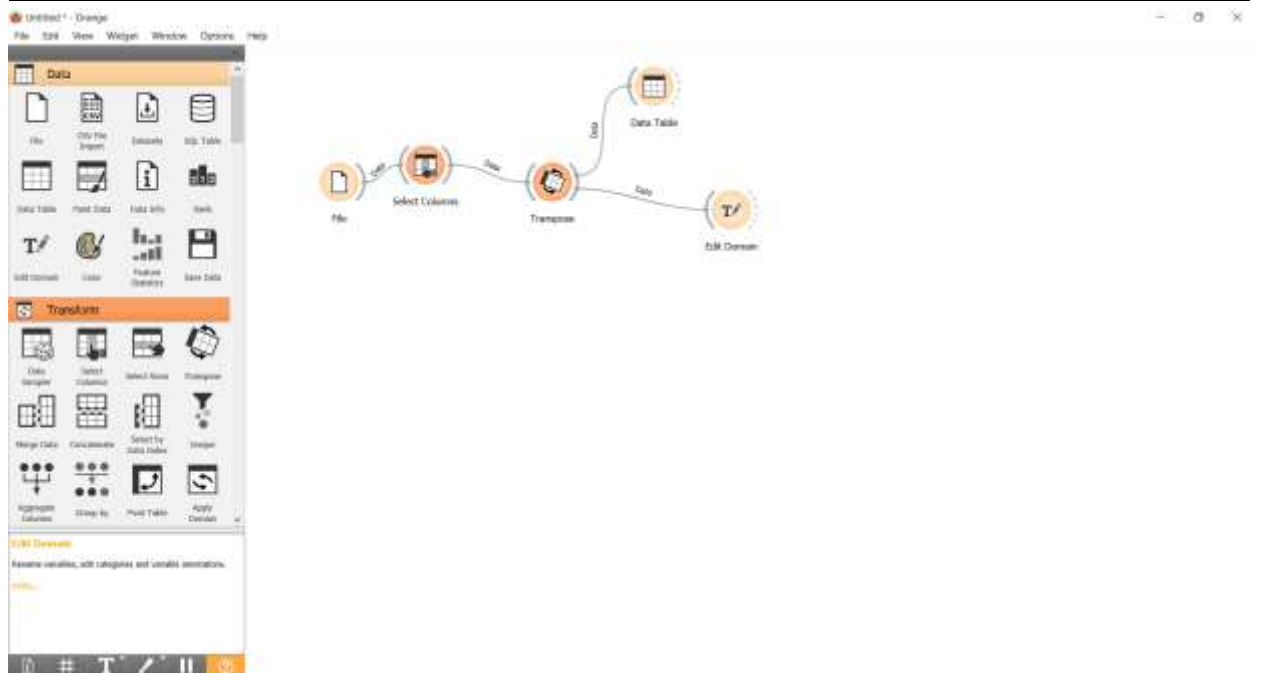


Рисунок - 11 Добавление виджета Edit Domain на холст

Открываем виджет Edit Domain. Прокручиваем вниз в разделе Variable, и выбираем Feature name. В поле Type изменяем тип на Время, а также изменим имя на Date, и нажимаем кнопку «Применить» (см.рис.12).

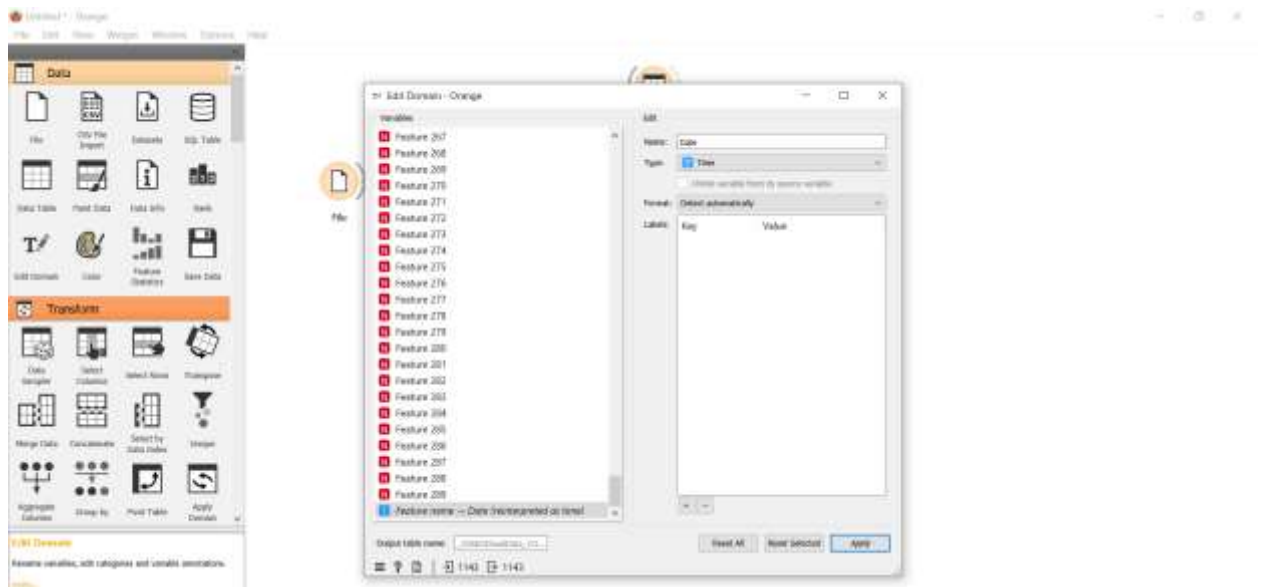


Рисунок - 12 Изменения типа данных столбца

После подготовки данных можем приступить к временному ряду. Для этого добавляем виджет Time Slice на холст и соединяем с виджетом Edit Domain (см.рис.13).

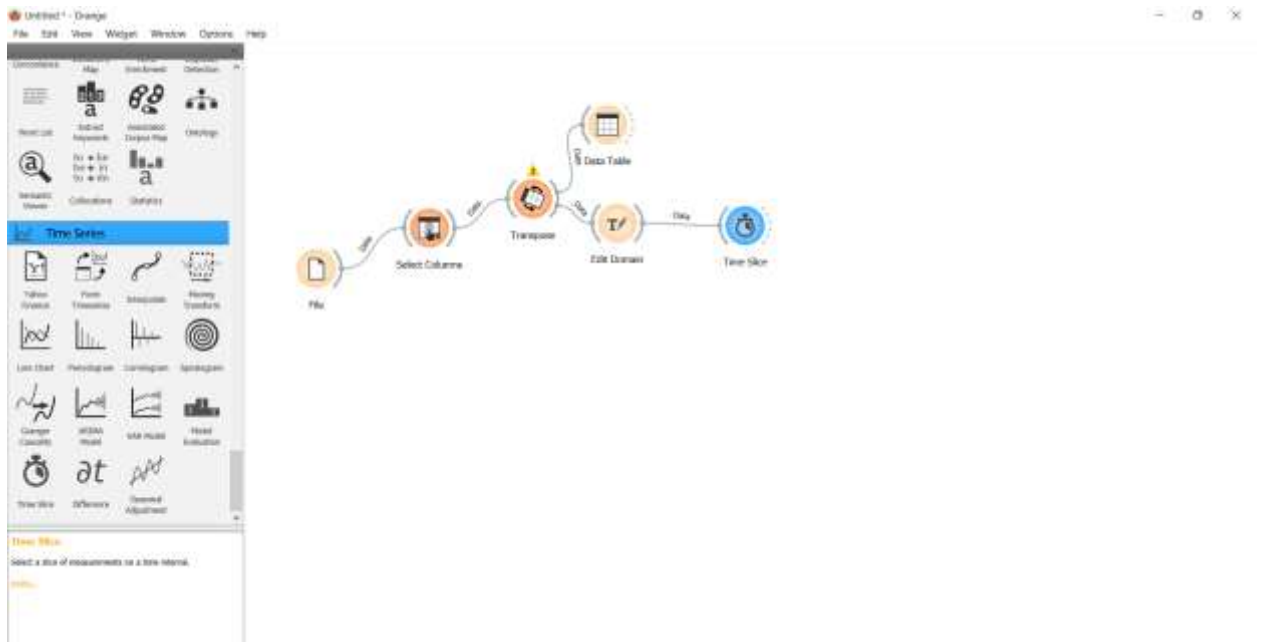


Рисунок - 13 Добавление виджета Time Slice на холст

Открываем виджет Time Slice, и выбираем интервал с 25.01.2020 до 26.01.2020 года (см.рис.14).

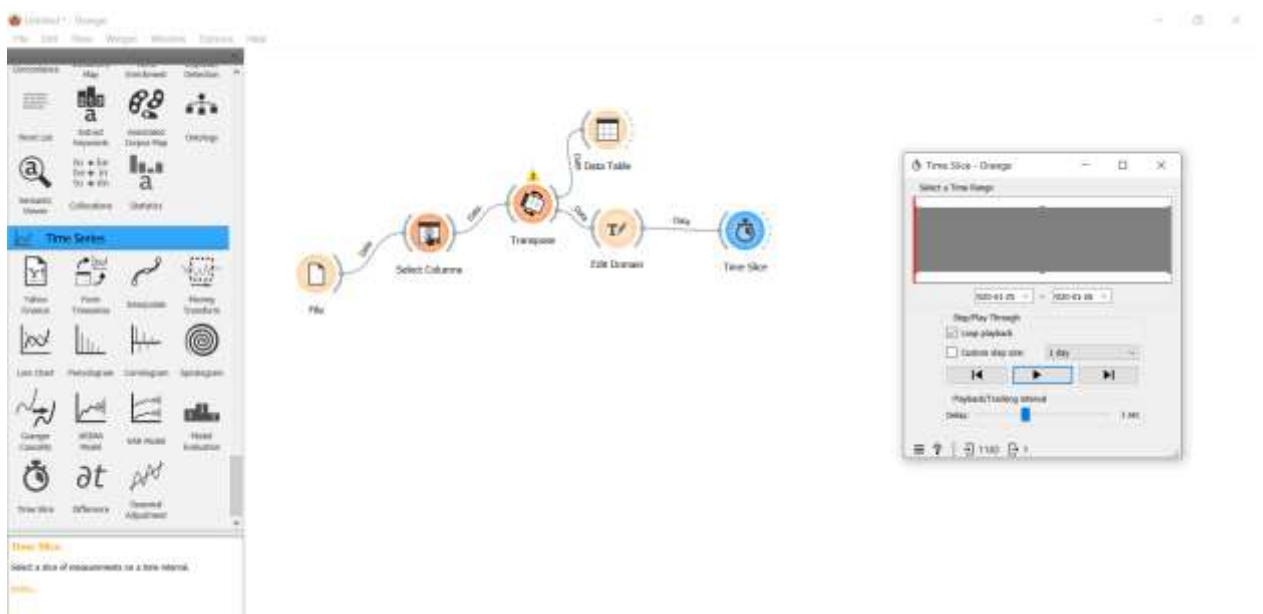


Рисунок - 14 Добавление временного ряда

Далее надо преобразовать данные обратно в исходный вид. Для этого добавим еще один виджет Transpose на холст, и соединим с виджетом Time Slice (см.рис.15).

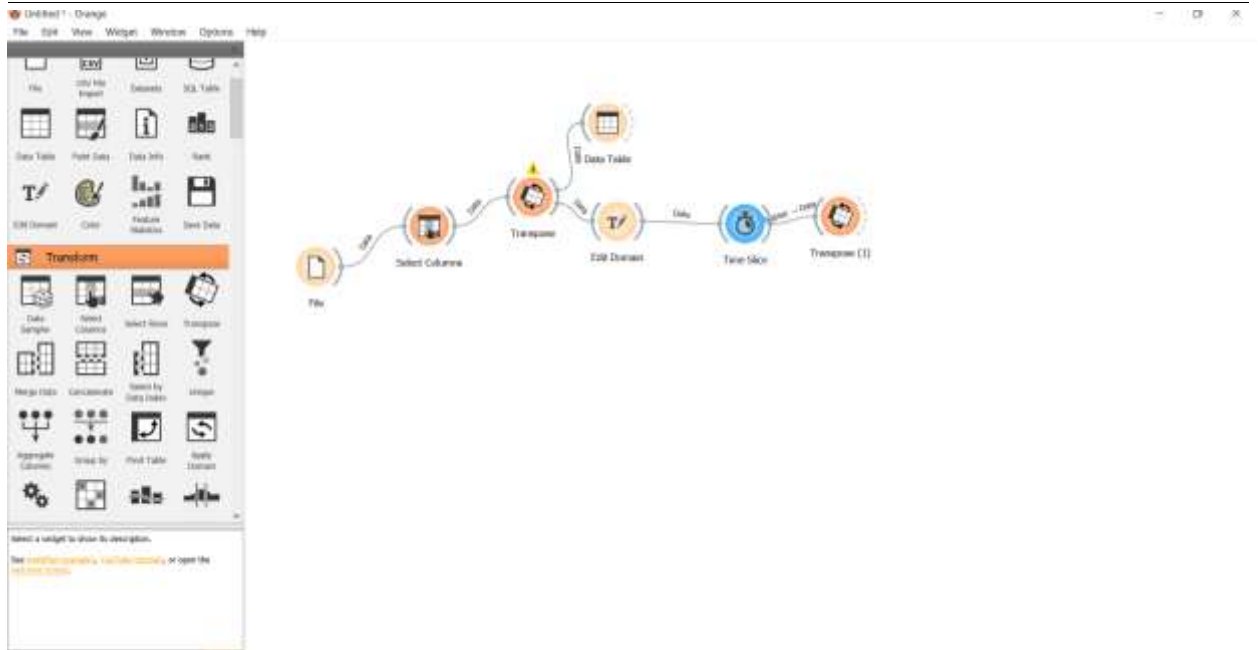


Рисунок - 15 Добавление виджета Transpose на холст

Открываем виджет Transpose (1), и задаем для переменной From variable значение Date (см.рис.16).

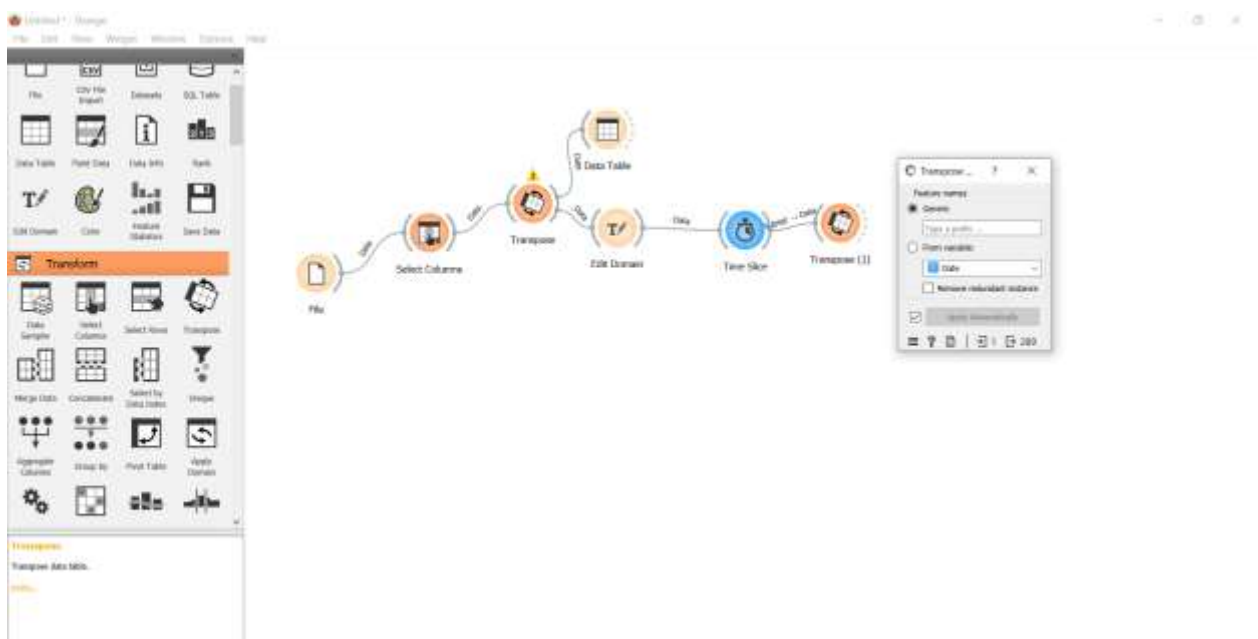


Рисунок - 16 Диалоговое окно виджета Transpose

Добавляем виджет Choropleth Map, и соединяем с виджетом Transpose (1). Виджет Choropleth Map предоставляет простой способ визуализации того, как измерение изменяется в пределах географической области, или показать уровень изменчивости в пределах региона (см.рис.17).

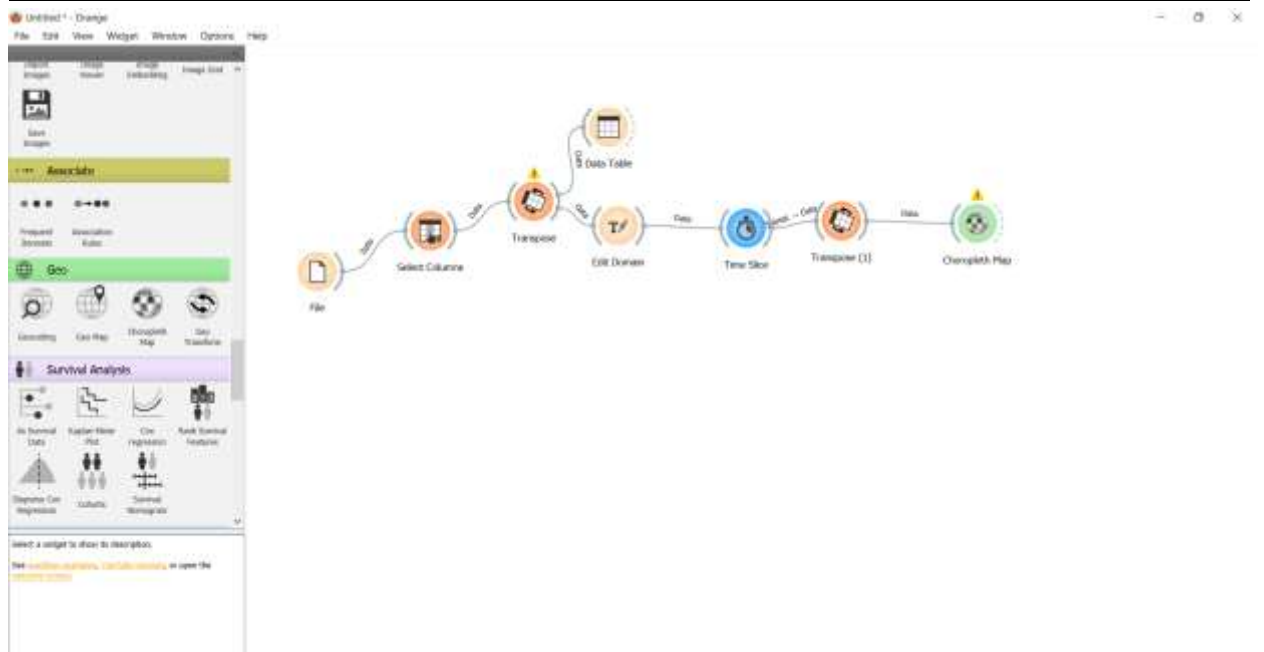


Рисунок - 17 Добавление виджета Choropleth Map на холст

Одновременно открываем виджет Time Slice и Choropleth Map. В виджете Time Slice нажимаем на кнопку Play. После запуска в виджете Time Slice, можно увидеть, как изменяется количество людей заразившиеся Covid-19 в виджете Choropleth Map. На интервале от 03.02.2020 до 04.02.2020 года можем увидеть, что количество людей заразившиеся Covid-19 составляет 650 (см.рис.18).

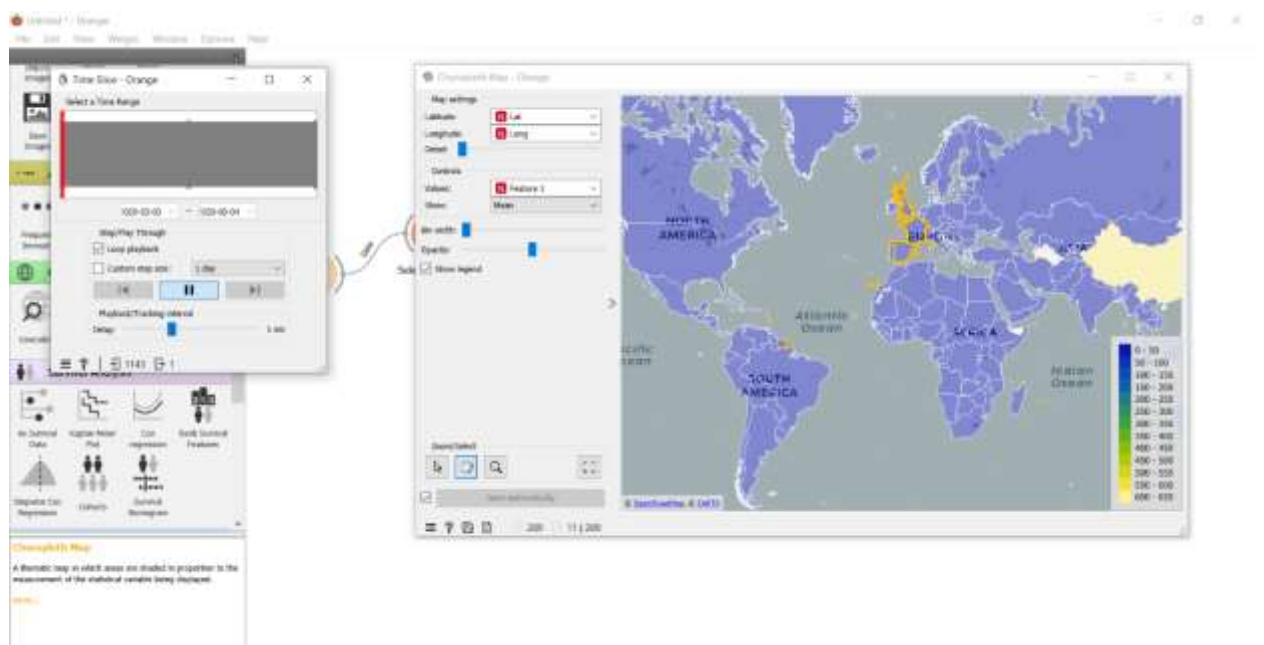


Рисунок - 18 Количество людей заразившиеся Covid-19 с 03.02.2020 до 04.02.2020 года

Также возьмем данные интервала времени с 13.02.2020 до 14.02.2020 года. В виджете Choropleth Map можем увидеть, что количество людей заразившихся Covid-19 составляет 2000 (см.рис.19).

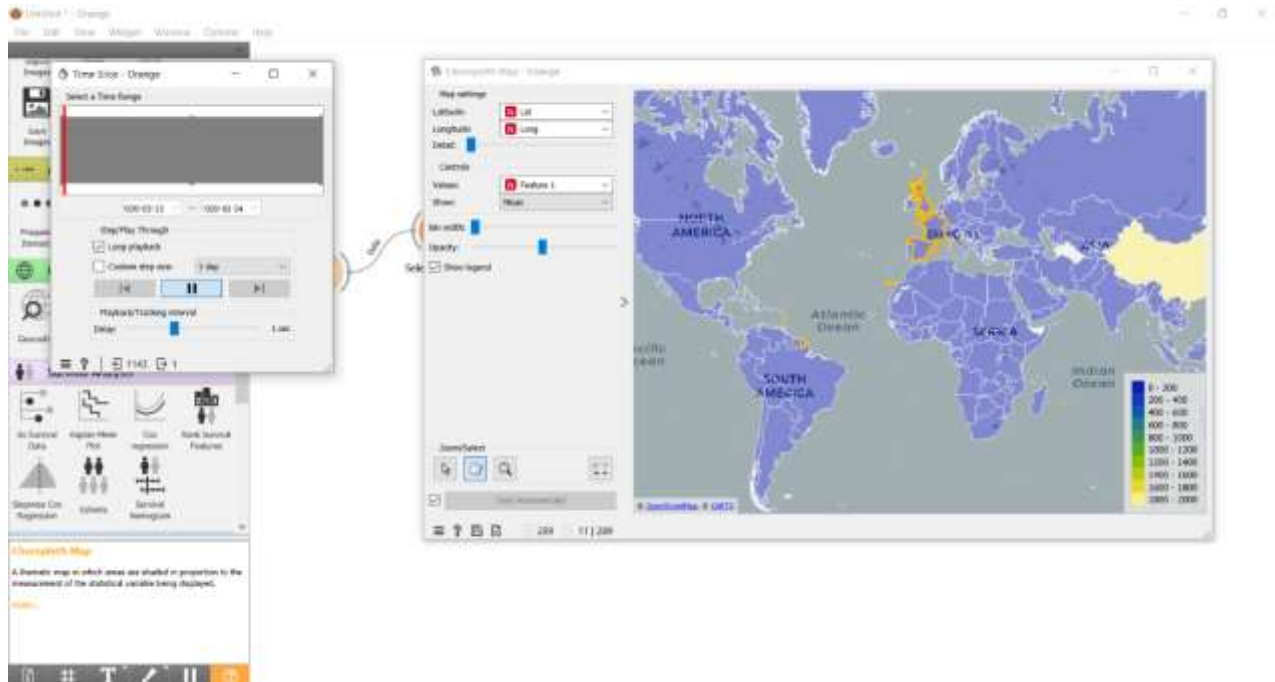


Рисунок - 19 Количество людей заразившихся Covid-19 с 13.02.2020 до 14.02.2020 года

Итоговая схема отслеживания динамики изменения количества людей заразившихся Covid-19 по всему миру (см.рис.20).

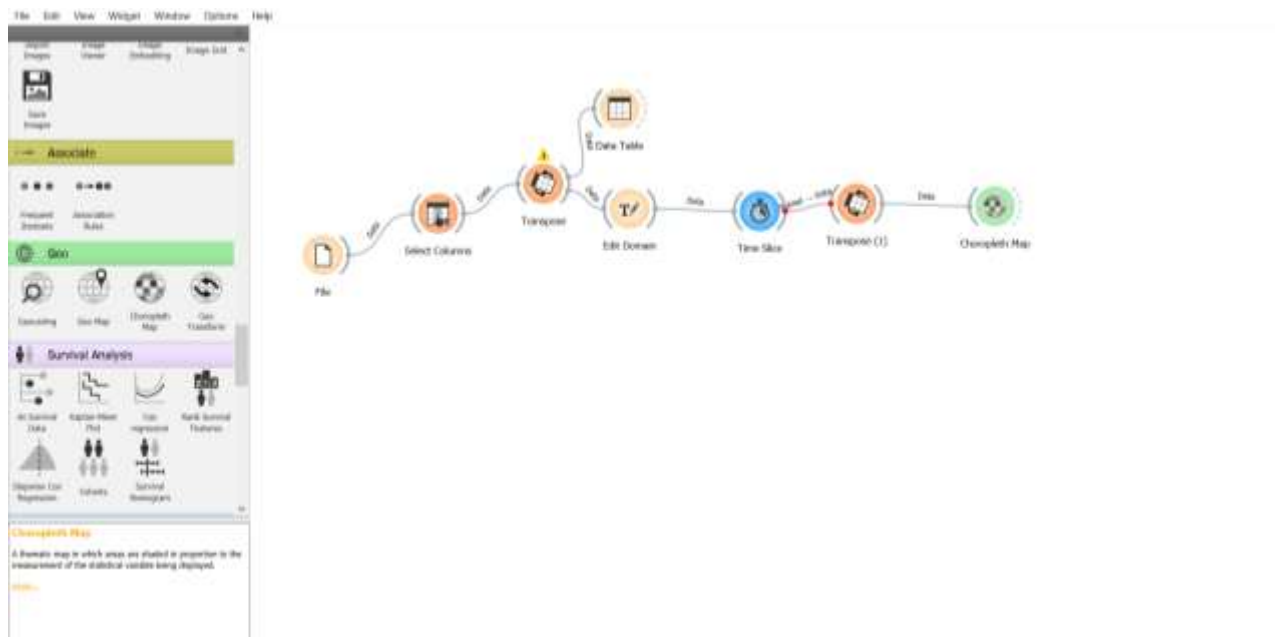


Рисунок - 20 Итоговая схема

4 Выводы

В данной работе отследили динамику изменения количества людей заразившихся Covid-19 по всему миру с помощью программного пакета визуального программирования на основе компонентов для визуализации данных Orange. С помощью виджетов File, Select Columns, Transpose, Data Table, Edit Domain, Time Slice, Choropleth Map отследили динамику изменения количества людей заразившихся Covid-19 по всему миру и получили итоговую схему.

Библиографический список

1. Raj A., Bhattacharyya P., Gupta G. R. Clusters of COVID-19 indicators in India: Characterization, correspondence and change analysis //SN Computer Science. – 2022. – Т. 3. – №. 3. – С. 210.
2. Прудников В. Б., Бахитова Р. Х. Применение методов машинного обучения для моделирования смертности пациентов с диагнозом COVID-19 // Статистическое образование в России: интеллектуальный анализ данных: материалы международной конференции (конгресса), Оренбург, 25–26 октября 2023 года. Оренбург: Оренбургский государственный университет, 2023. – С. 580-590.
3. Попков А. Ю., Дубнов Ю. А., Попков Ю. С. Прогнозирование распространения COVID-19 в ЕС с использованием рандомизированного машинного обучения динамических моделей //Информационные технологии и вычислительные системы. 2022. №. 3. С. 67-78.
4. Корниевич А. К. Мультистрановой анализ типологии пандемии COVID-19 с использованием статистического и машинного обучения //79-я научная конференция студентов и аспирантов Белорусского государственного университета. 2022. С. 341-345.
5. Карабутова, А. Г. Использование методов машинного обучения при исследовании распространения коронавирусной инфекции / А. Г. Карабутова // Вестник студенческого научного общества ГОУ ВПО "Донецкий национальный университет". 2023. Т. 1. № 15. С. 217-223.