минобрнауки россии

федеральное государственное бюджетное

образовательное учреждение высшего образования

«ЧЕРЕПОВЕЦКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

|  |
| --- |
| Иформационных технологий |
| наименование института (факультета) |
| Математическое и программное обеспечение ЭВМ |
| наименование кафедры |

ОТЧЁТ

по учебной: технологической практике 1

Листов 15

Студента Маслова Владислава Андреевича группы 1ПИб-02-1оп-22

Место прохождения практики

ФГБОУ ВО «Череповецкий государственный  
университет» кафедра математического и  
программного обеспечения ЭВМ, компьютерный класс

Студент \_\_\_\_\_\_\_\_\_\_\_\_

(подпись)

Руководитель практики

от кафедры МПО ЭВМ доцент к.т.н. Юдина О.В.\_\_\_\_\_\_\_\_

(должность) (подпись, Ф.И.О.)

2024 год

Введение

Цель практики – познакомиться с некоторыми инструментами для анализа данных, изучить возможности этих инструментов; познакомиться с программой Orange; научиться решать оптимизационные задачи в Excel с использованием пакета анализа.

Описание и возможности пакетов для работы с данными:

1. KNIME:

* Свободно распространяемый: да.
* Предназначение: платформа для науки о данных, позволяющая анализировать данные и масштабировать решения без кодирования.
* Возможности: интеграция данных, преобразование, моделирование, визуализация.
* Требования: не требуется умение программировать.
* Работает с данными: различные источники данных.
* Результат работы: аналитические отчеты, визуализации.

1. SPSS:

* Свободно распространяемый: нет.
* Предназначение: платформа для науки о данных и анализа данных.
* Возможности: статистический анализ, машинное обучение, анализ текста.
* Требования: знание статистики и анализа данных.
* Работает с данными: различные источники данных.
* Результат работы: статистические отчеты, модели машинного обучения.

1. SAS:

* Свободно распространяемый: нет.
* Предназначение: лидер в области аналитики и программного обеспечения для AI, Visual Analytics, Data Science, Risk, Fraud и IT в Германии.
* Возможности: аналитика, AI, Visual Analytics, Data Science, Risk, Fraud и IT.
* Требования: знание аналитики и соответствующих областей.
* Работает с данными: различные источники данных.
* Результат работы: аналитические отчеты, модели AI.

1. STATISTICA 6.0:

* Свободно распространяемый: нет.
* Предназначение: пакет программ для статистической обработки данных.
* Возможности: анализ данных, статистика, машинное обучение, визуализация.
* Требования: знание статистики и анализа данных.
* Работает с данными: различные источники данных.
* Результат работы: статистические отчеты, модели машинного обучения.

1. Minitab:

* Свободно распространяемый: нет.
* Предназначение: пакет программ для статистического анализа.
* Возможности: статистический анализ, визуализация.
* Требования: знание статистики и анализа данных.
* Работает с данными: различные источники данных.
* Результат работы: статистические отчеты, визуализации.

1. R + R-Studio:

* Свободно распространяемый: да.
* Предназначение: язык программирования для статистических вычислений и графики.
* Возможности: статистический анализ, машинное обучение, визуализация.
* Требования: знание языка R и статистики.
* Работает с данными: различные источники данных.
* Результат работы: статистические отчеты, модели машинного обучения, визуализации.

1. Python:

* Свободно распространяемый: да.
* Предназначение: универсальный язык программирования.
* Возможности: математические расчеты, моделирование, статистические расчеты.
* Требования: знание языка Python.
* Работает с данными: различные источники данных.
* Результат работы: программы на Python, модели машинного обучения.

1. Orange Data Mining:

* Свободно распространяемый: да.
* Предназначение: платформа для машинного обучения и визуализации данных.
* Возможности: интерактивный анализ данных, машинное обучение, визуализация.
* Требования: знание машинного обучения и визуализации данных.
* Работает с данными: различные источники данных.
* Результат работы: модели машинного обучения, визуализации.

1. Loginom:

* Свободно распространяемый: Нет.
* Предназначение: платформа для решения широкого спектра бизнес-задач, требующих обработки больших объемов данных, реализации сложной логики и применения методов машинного обучения.
* Возможности: интеграция данных, подготовка данных, моделирование, визуализация.
* Требования: не требуется умение программировать.
* Работает с данными: различные источники данных.
* Результат работы: аналитические отчеты, визуализации.

Раздел 1. Работа в пакете Orange. Его возможности, результаты выполнения задания.

* 1. Практическая часть

Знакомство с Orange: создание первого рабочего процесса (рис. 1).

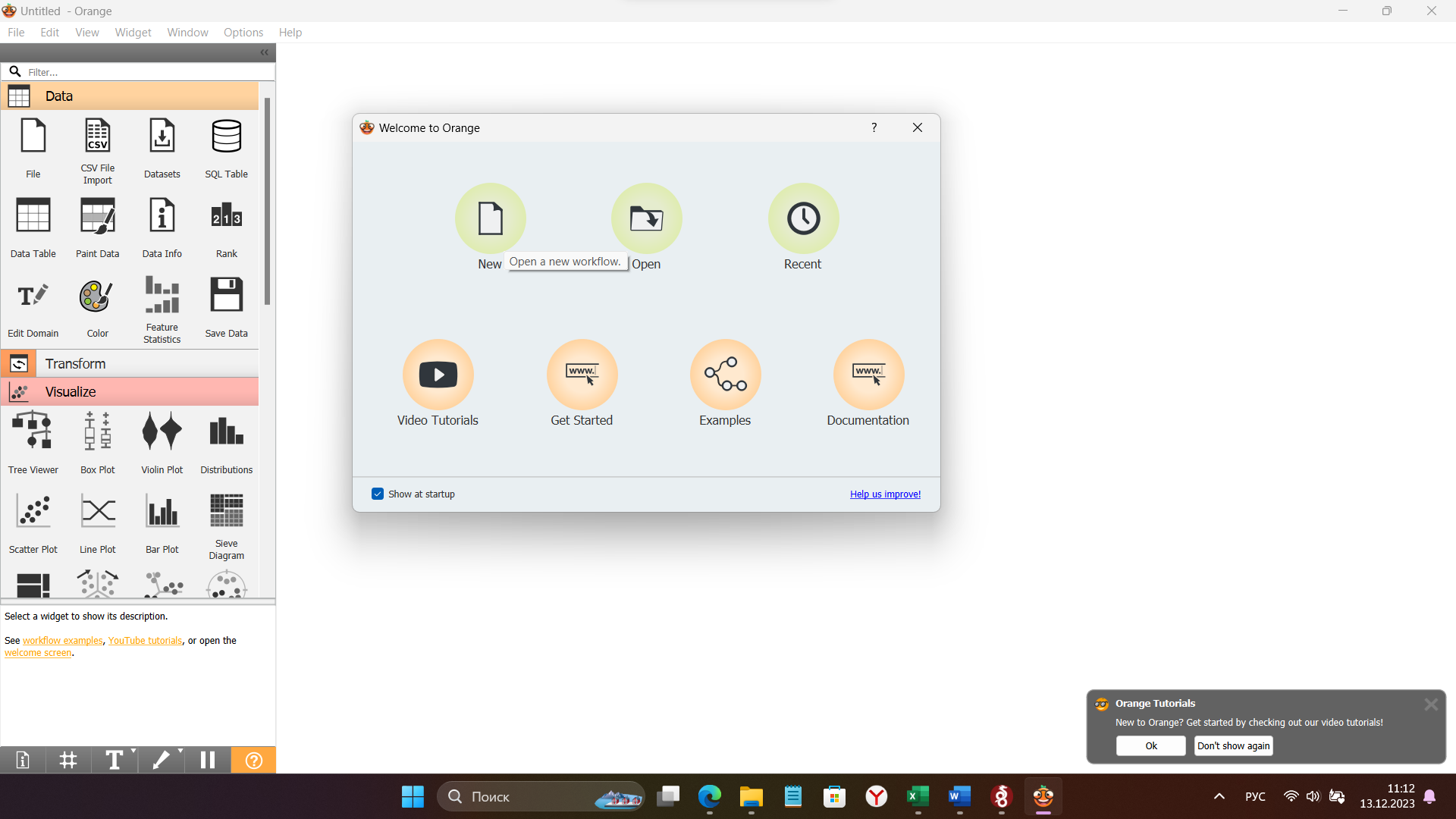


Рис. 1. Создание нового рабочего процесса

Импорт данных.

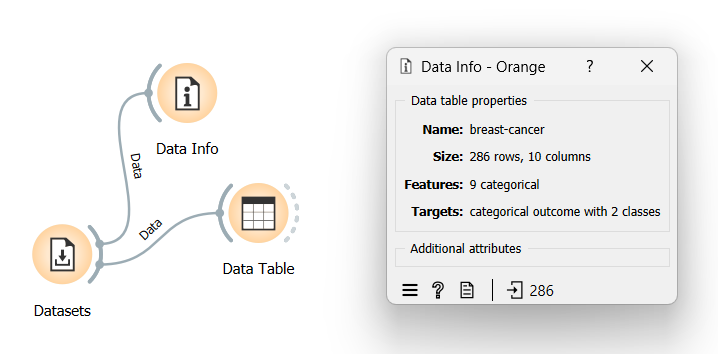


Рис. 2. Меню виджета Data Info

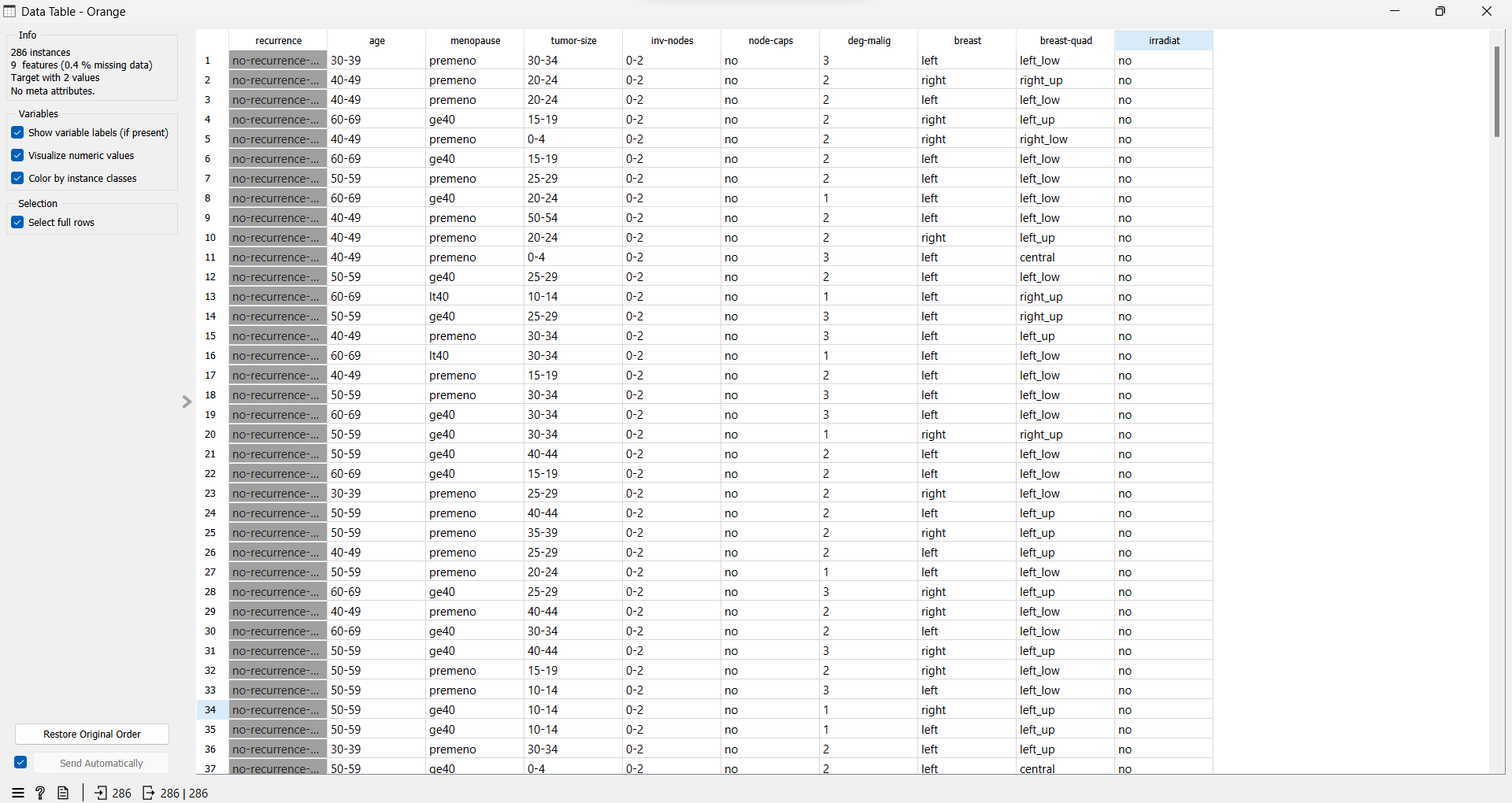


Рис. 3. Меню виджета Data Table

Визуализация данных при помощи Orange:

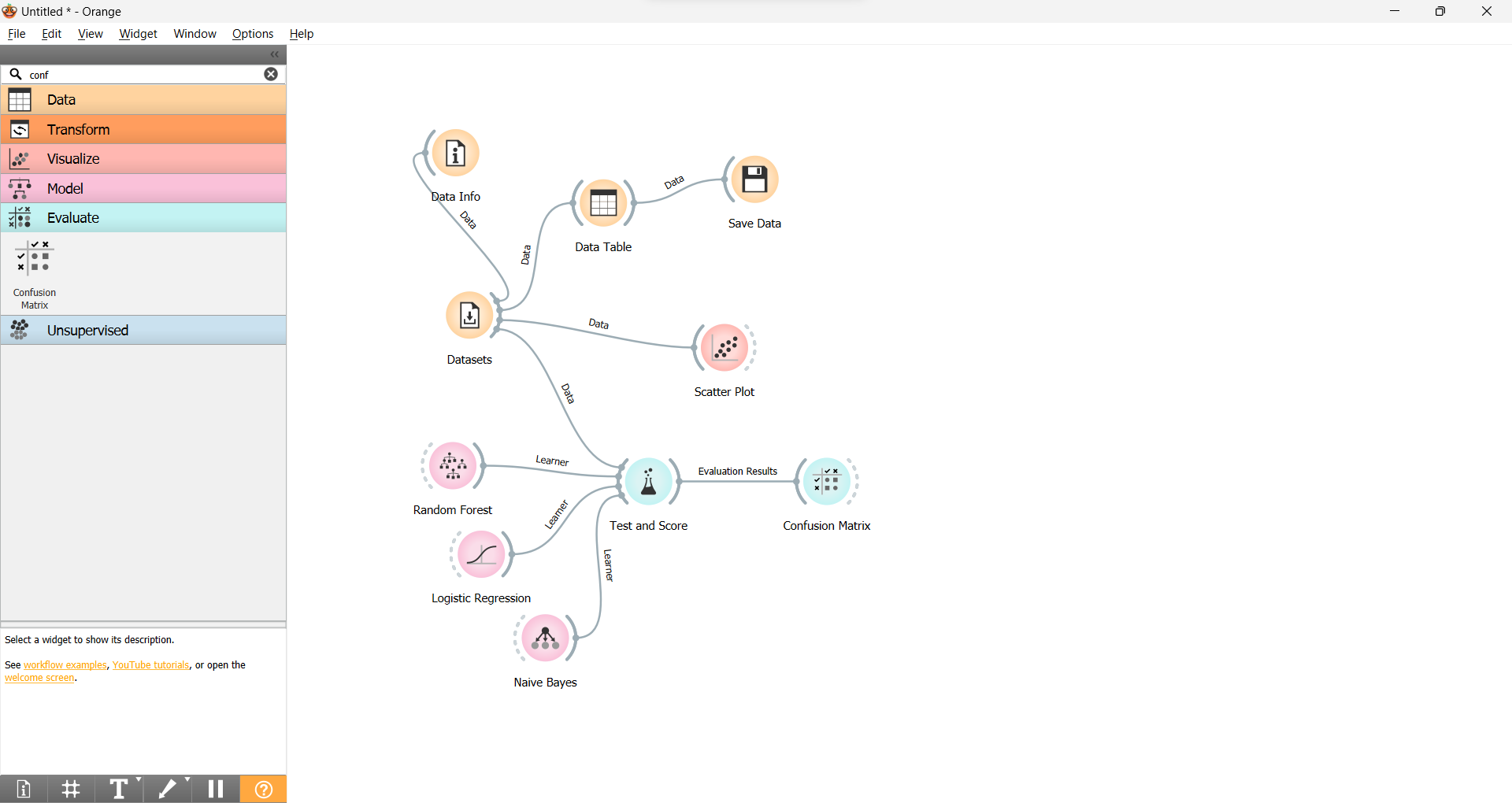


Рис. 4. Виджеты

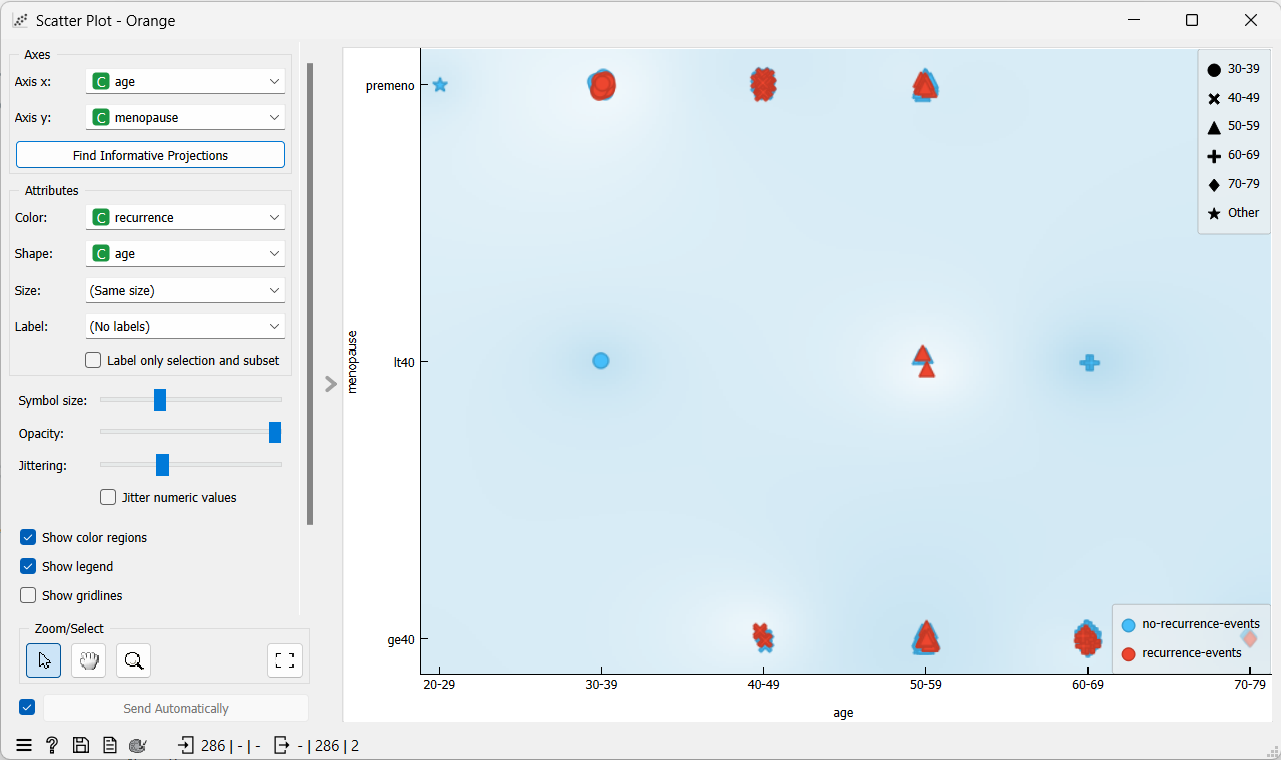


Рис. 5. Меню виджета Scatter Plot

Использование инструментов работы с данными:

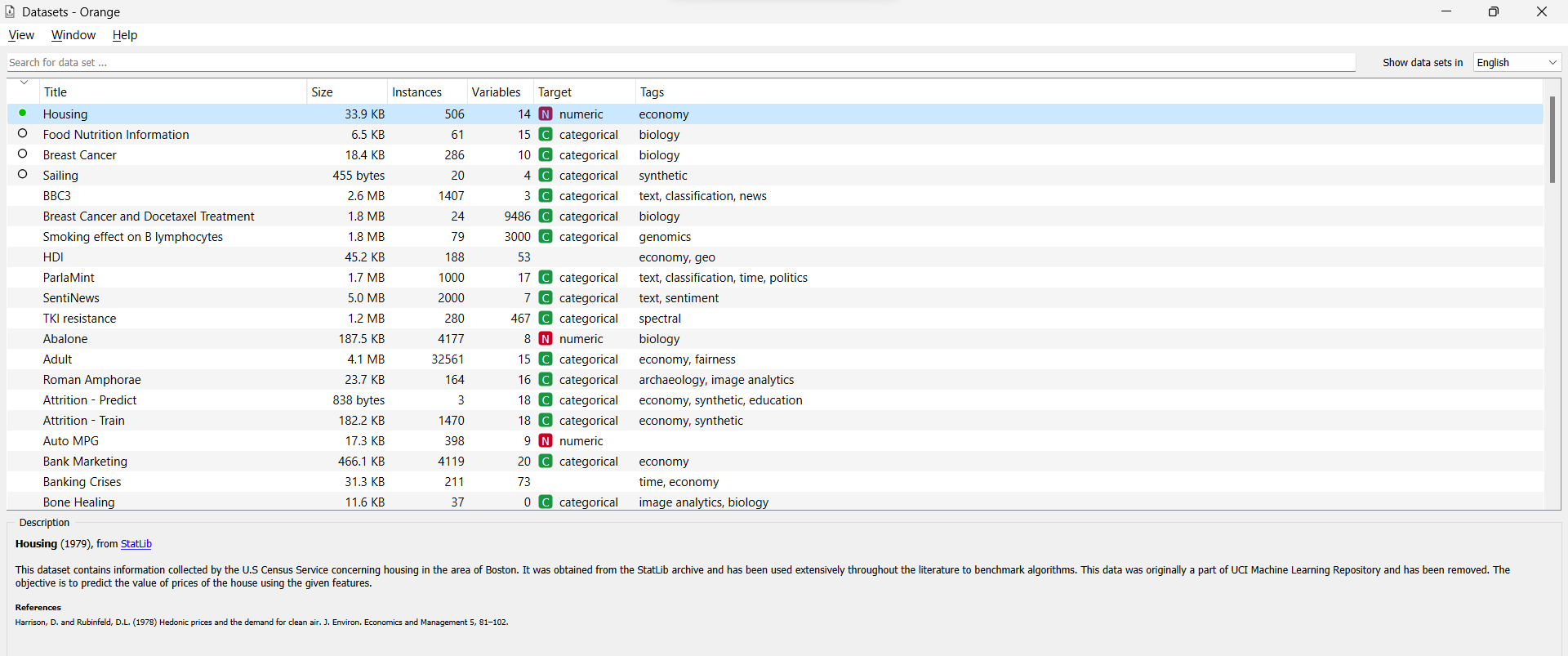


Рис. 6. Выбор датасета Housing

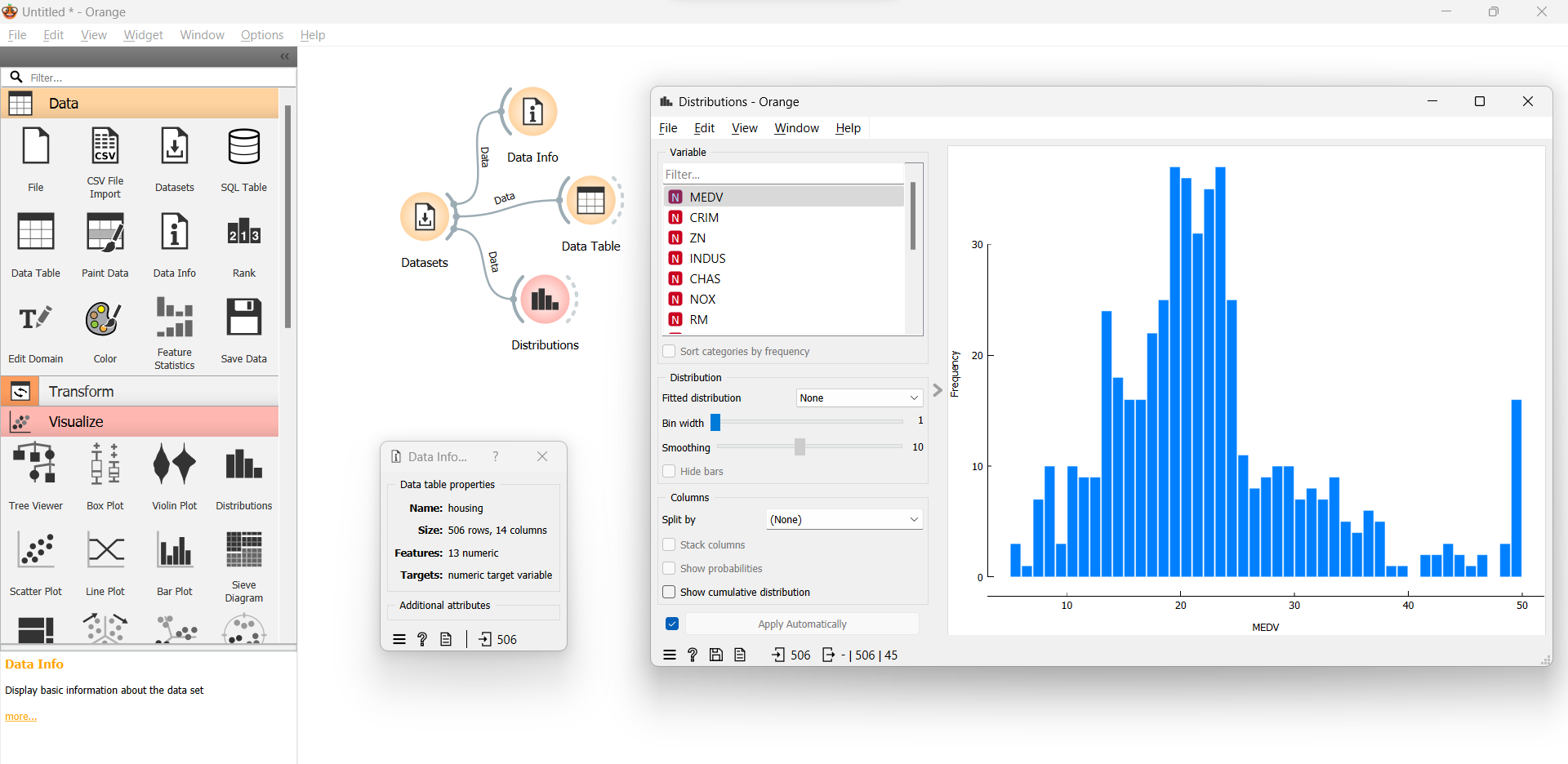


Рис. 7. Меню виджета Distributions

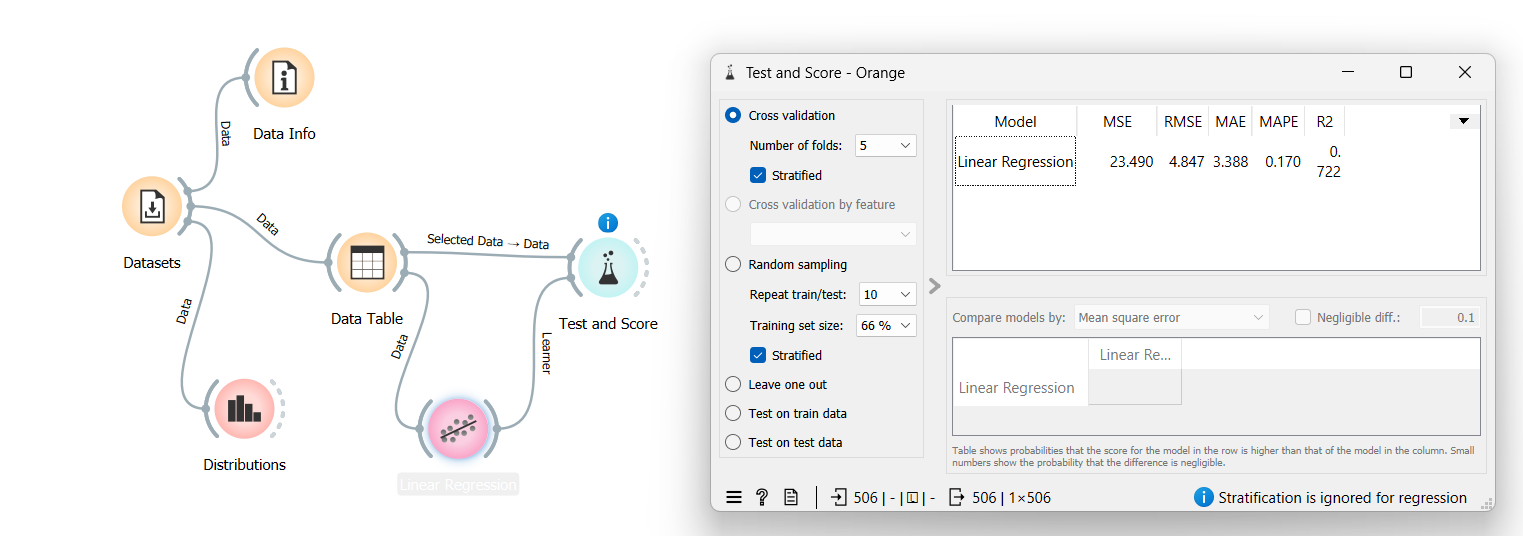


Рис. 8. Виджет Test And Score

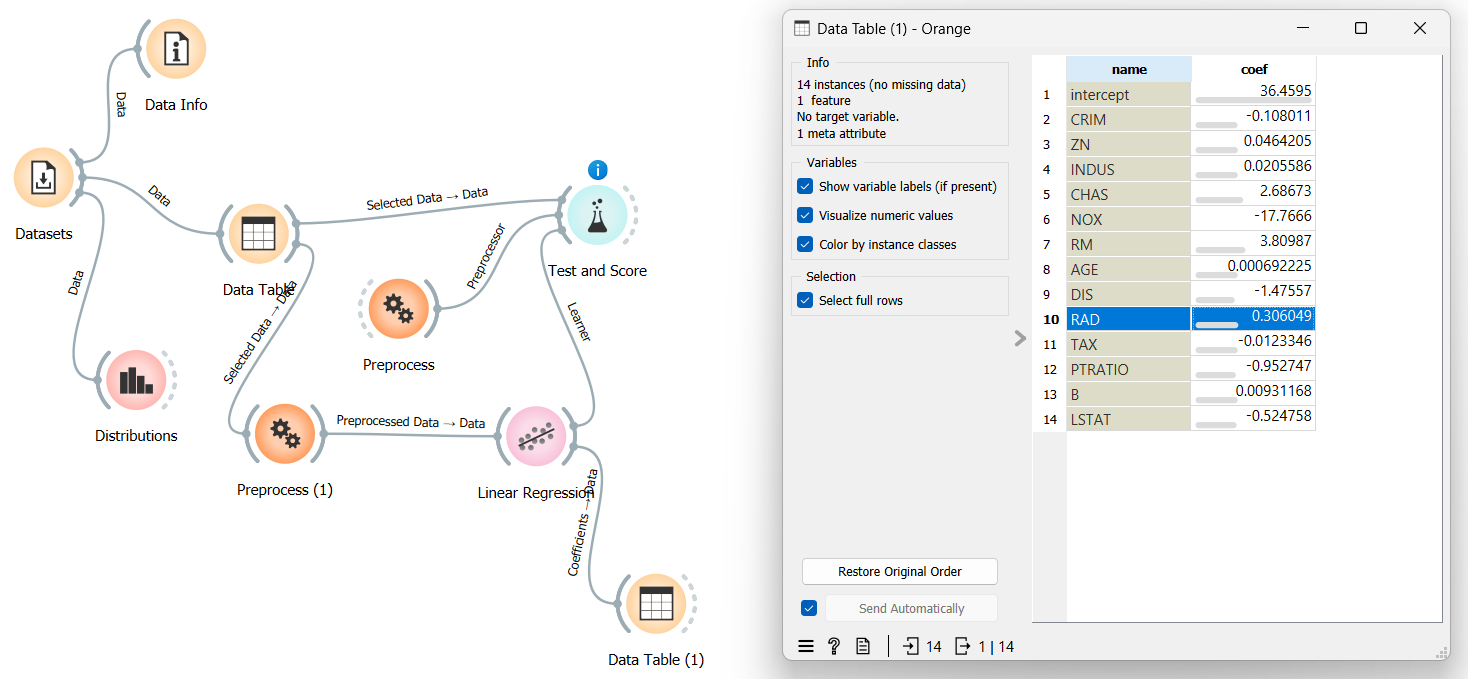


Рис. 9. Данные после препроцессинга

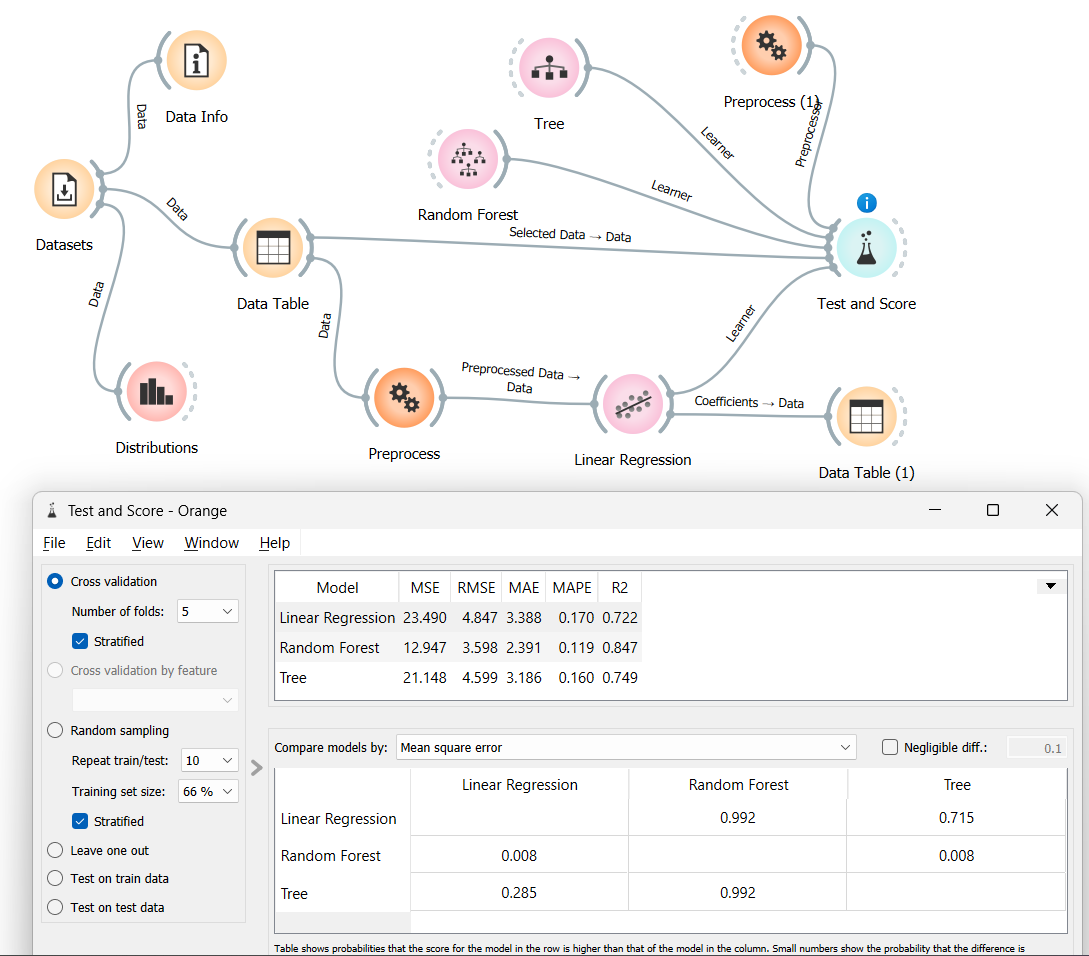


Рис. 10. Результаты теста после препроцессинга

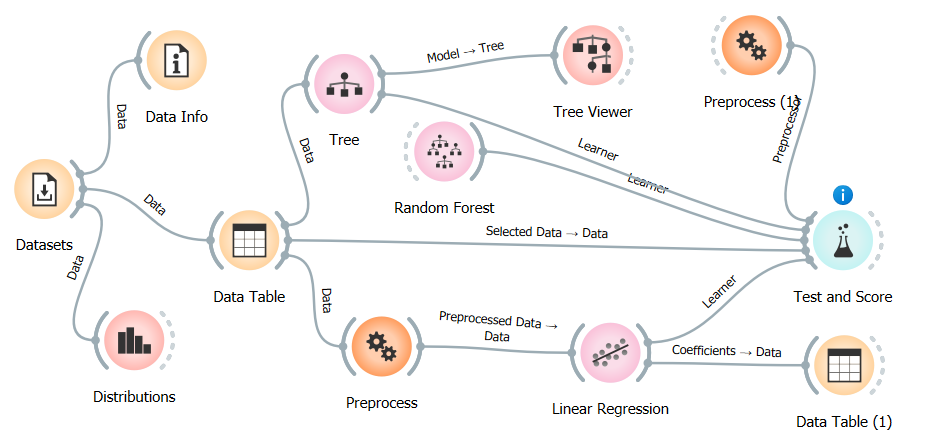


Рис. 11. Подготовка к визуализации дерева

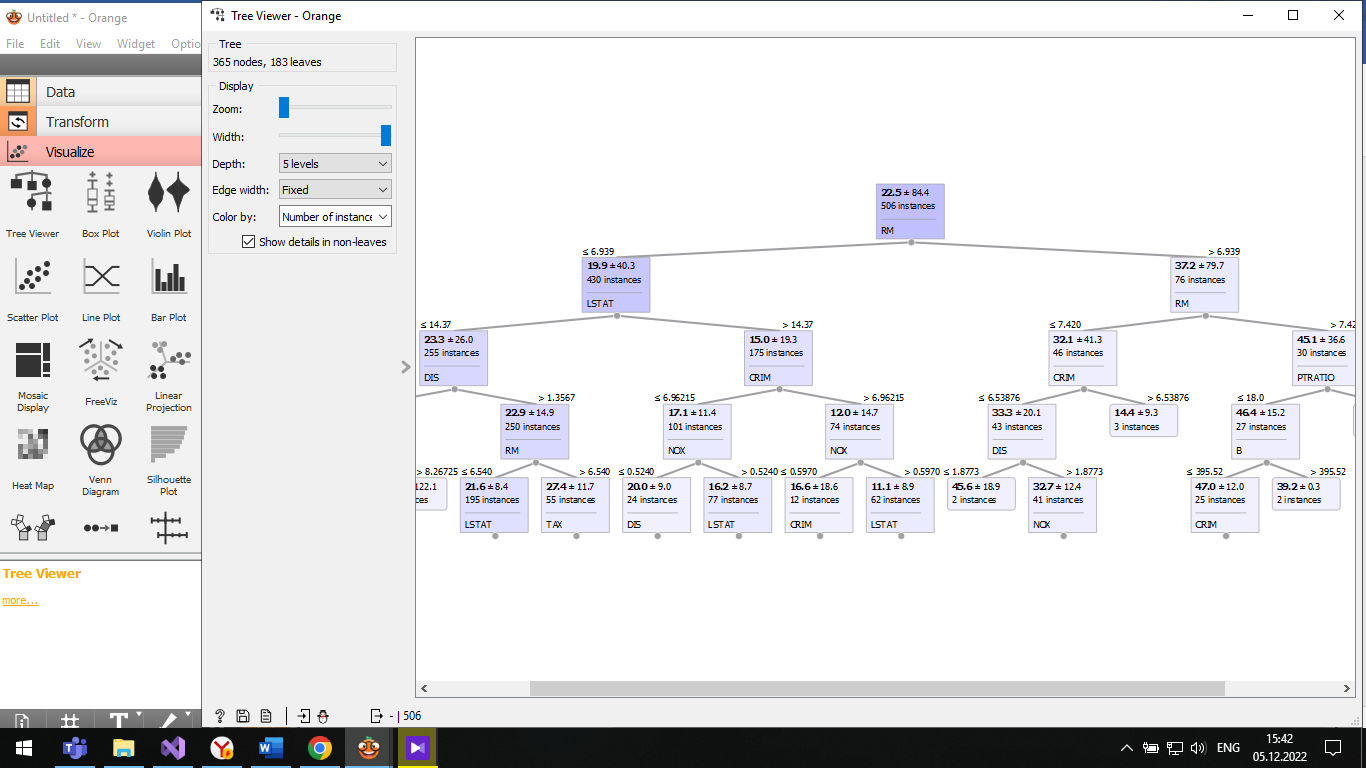


Рис. 12. Визуализация дерева

Вывод: программа подходит тем, кто работает с большими и малыми данными. Результаты отображаются в удобном формате как в табличным, так и визуальным способом. Данные легко (автоматически) обрабатываются, что позволяет сделать анализ данных точнее. В программе также можно делать регрессию, распределение, сортировку и многое другое.

Раздел 2. Описание решения оптимизационной задачи в Excel

Задание 1. Вариант 13. Для изготовления трёх видов изделий P1, P2 и Р3 используют три вида материалов: S1, S2, S3. Запасы материалов, технологические нормы расхода материалов на каждое изделие и цена единицы изделия приведены в таблице 1.

Составить план выпуска изделий, обеспечивающих их максимальный выпуск по стоимости.

Табл. 1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Вид материала** | **Норма расхода материала на одно изделие, кг** | | | **Запас материала, кг** |
| **P1** | **P2** | **P3** |
| S1 | 9 | 9 | 2 | 180 |
| S2 | 4 | 3 | 2 | 120 |
| S3 | 1 | 2 | 4 | 220 |
| Цена одного изделия (у.е.) | 7 | 8 | 6 | max |

Составим математическую модель задачи.

Введём неизвестные: x1 – количество изделий P1; x2 – количество изделий P2; x3 – количество изделий P3.

Запишем ограничения задачи:

Запишем целевую функцию (цена одного изделия):

Требуется найти x1, x2, x3, дающие максимум целевой функции S при заданных ограничениях.



Рис. 13. Электронная таблица в режиме отображения формул

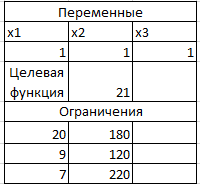


Рис. 14. Электронная таблица для решения задачи

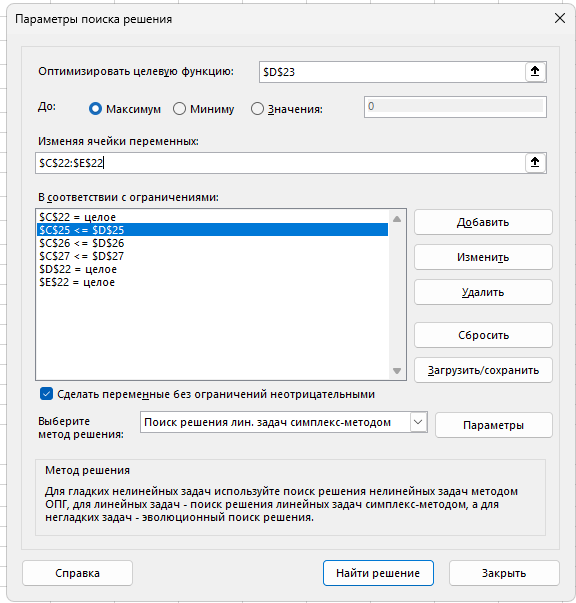


Рис. 15. Диалоговое окно Поиск решения



Рис. 16. Результаты решения задачи

Вывод: при заданных ограничениях тип изделий P1 производить невыгодно.

Задание 2. Вариант 13.

Требуется определить минимальную по стоимости смесь сырья для изготовления пищевых концентратов, которые должны содержать питательные вещества (П). Эти вещества содержатся в сырье (М) в различных сочетаниях. Содержание питательных веществ в сырье и готовом продукте, а также цена на каждый вид сырья представлены в таблице 2.

Табл. 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Питательные вещества** | **Виды сырья** | | | **Минимальное содержание питательного вещества в готовом продукте** |
| **M1** | **M2** | **M3** |
|  |
| П1 | 1 | 1 | 4 | 128 |  |
| П2 | 3 | 4 | 1 | 140 |  |
| П3 | 0 | 1 | 1 | 50 |  |
| П4 | 2 | 0 | 3 | 80 |  |
| Цена за единицу сырья (у.е.) | 10 | 8 | 12 | min |  |

Составим математическую модель задачи.

Введём неизвестные: x1 – количество изделий M1; x2 – количество изделий M2; x3 – количество изделий M3.

Запишем ограничения задачи:

Запишем целевую функцию (цена за единицу сырья):

Требуется найти x1, x2, x3, x4, дающие минимум целевой функции П при заданных ограничениях.

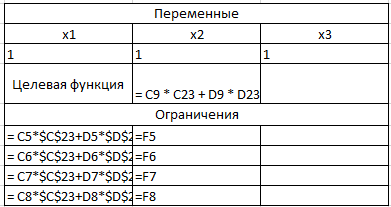


Рис. 17. Электронная таблица в режиме отображения формул

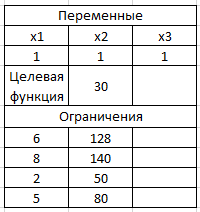


Рис. 18. Электронная таблица для решения задачи

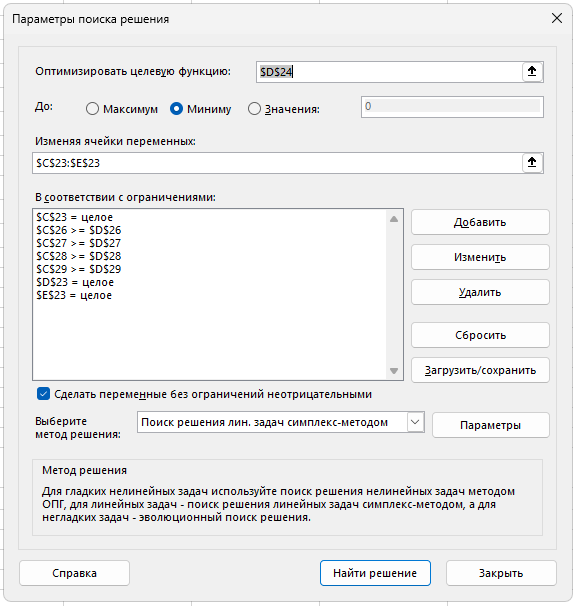


Рис. 19. Диалоговое окно Поиск решения



Рис. 20. Результаты решения задачи

Вывод: при заданных ограничениях наиболее выгодно включить в готовый продукт виды сырья M1, M2, M3 в пропорциях 3:1:2 соответственно.

Заключение

В ходе работы были рассмотрены инструменты для работы с данными (KNIME Analytics Platform, Statistical Package for the Social Sciences, SAS, STATISTICA 6.0, Minitab, Orange3, Python, R + R-Studio). Многие из программ имеют бесплатную версию или имеют возможность приобретения студенческой лицензии, есть программы, в которых не обязательны знания программирования, но большинство требуют базовых знаний.

Была выполнена работа при помощи программы Orange, а также было изучено решение оптимизационных задач в программе Excel с использованием пакета анализа.

В процессе выполнения заданий по ознакомительной практике частично были освоены компетенции:

Способен использовать современные информационные технологии и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности (ОПК-2).

Способен решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности (ОПК-3).

Список литературы

1. Официальный сайт платформы KNIME [Электронный ресурс]. URL: https://www.knime.com.
2. Официальный сайт программного обеспечения IBM SPSS [Электронный ресурс]. URL: https://www.ibm.com/spss.
3. Официальный сайт Orange. [Электронный ресурс]. URL: https://orangedatamining.com.
4. Страница Википедии о Statistica [Электронный ресурс]. URL: https://en.wikipedia.org/wiki/STATISTICA.
5. Страница Википедии о Minitab [Электронный ресурс]. URL: https://en.wikipedia.org/wiki/Minitab.
6. Официальный сайт R-Project [Электронный ресурс]. URL: https://www.r-project.org.
7. Официальный сайт платформы Python [Электронный ресурс]. URL: https://www.python.org.
8. Официальный сайт SAS [Электронный ресурс]. URL: https://www.sas.com.
9. Официальный сайт Loginom [Электронный ресурс]. URL: https://loginom.com.