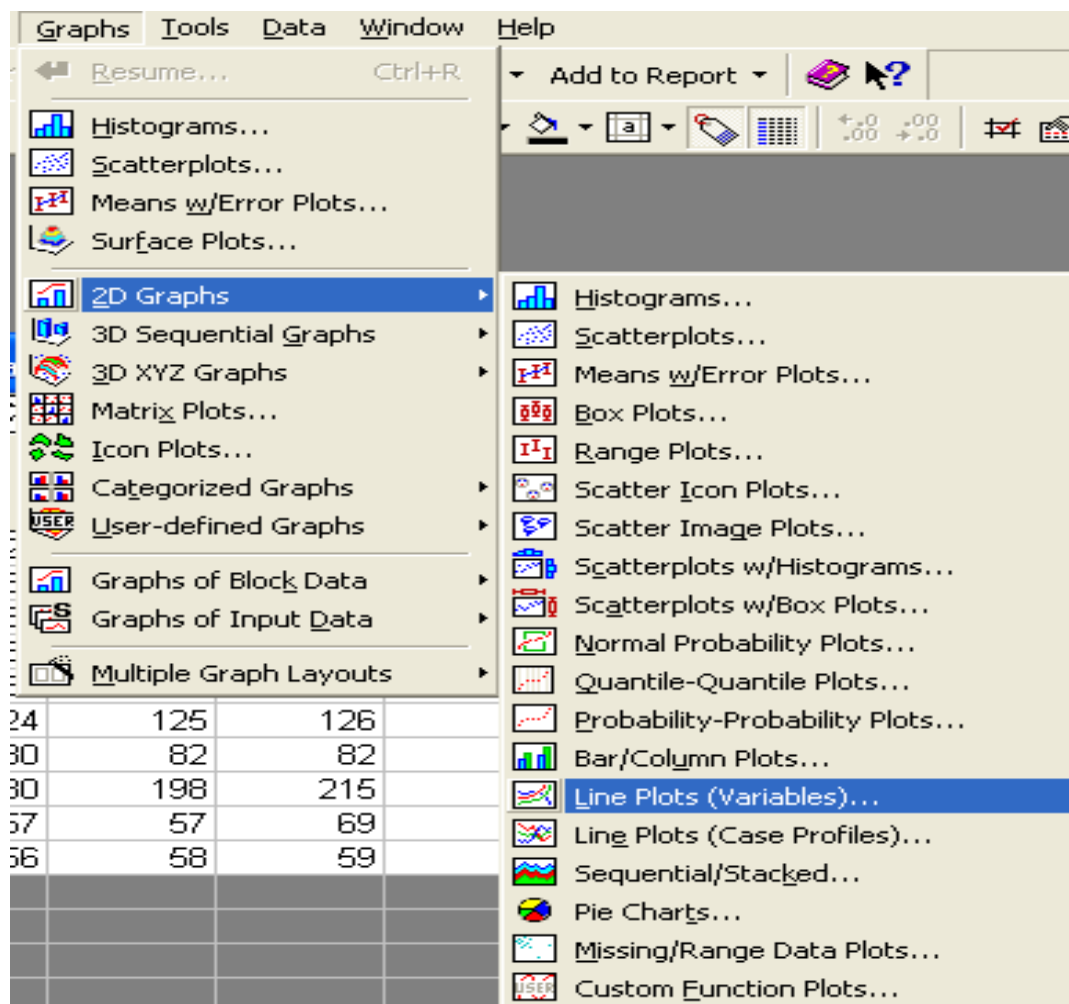


Графика в STATISTICA

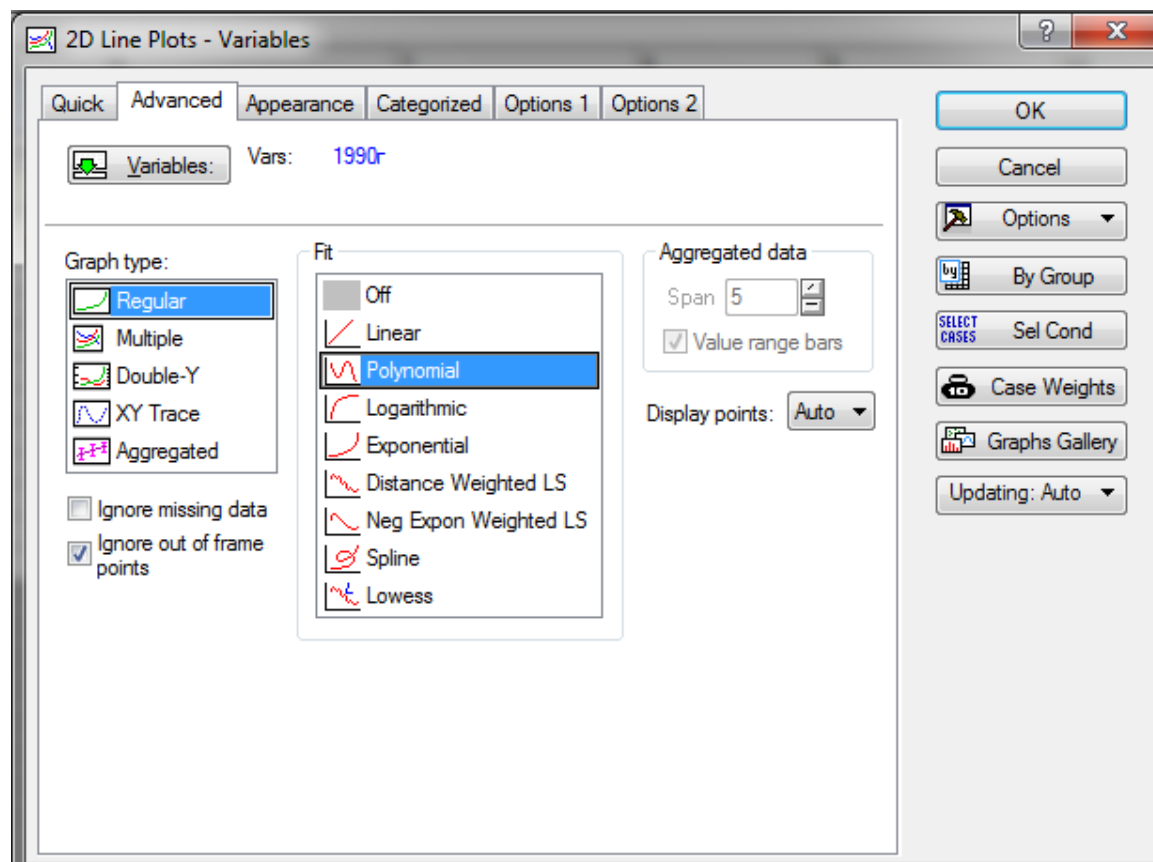
Проиллюстрируем построение графиков и их анализ на основе данных из таблицы на рис. 1. В нем приведены некоторые экономические показатели 10 крупнейших стран мира по численности городского населения: общее число жителей (млн. чел.) на 1990, 1995, 2000 г.; доля (%) городского населения на 1995 г.; наличие крупных запасов нефти и газа (более 1,5 млрд. т); структура ВВП (%) в промышленности, сельском хозяйстве, сфере услуг.

Данные: Страны мира1 (9v * 10с)									
	Крупнейшие страны мира по численности населения								
	1 1990г	2 1995г	3 2000г	4 Город 1995г	5 Нефть	6 Газ	7 Пром.	8 С/х	9 Услуги
Китай	1120	1121	1275	30,3	есть	есть	48	21	31
Индия	830	935	1010	26,8	нет	нет	30	29	41
США	250	263	250	76,2	есть	есть	26	2	72
Бразилия	150	162	170	78,2	нет	нет	37	14	49
Россия	289	149	146	73	есть	есть	38	7	55
Япония	124	125	126	77,6	нет	нет	38	2	60
ФРГ	80	82	82	86,5	нет	нет	38	2	60
Индонезия	180	198	215	35,4	есть	есть	42	17	41
Великобритания	57	57	69	89,5	нет	нет	32	2	66
Франция	56	58	59	72,8	нет	нет	27	2	71

Для построения линейного графика (Line plot) надо провести действия, показанные на рисунке.



Если в открывшемся окне щелкнуть по кнопке Variables и ввести имя переменной 1990г., нажать на ОК, то программа построит график, изображенный на рисунке следующего слайда. На вкладке Advanced можно выбрать типы графиков (простой, составной, с двойной осью Y, трассировочный, агрегированный) и математическую функцию подгонки (линейная, полиномиальная, логарифмическая, экспоненциальная и т.д.)



На вкладке Параметры1 можно указать способ отображения имен объектов на графике, в зависимости от того в каком столбце в таблице указаны имена объектов

2М Линейные графики (для переменных)

Быстрый | Дополнительно | Вид | Категории | **Параметры 1** | Параметры 2

Заголовок
☒ Отображать стандартный заголовок
Заголовок: Отображать сверху ▾

Сноска

☐ Отображать сноску

Параметры отображения
Метки наблюдений: **Имена наблюдений** ▾
☐ Отобр. мет Переменная: 1990г
☒ Отобр. текстовые метки (или даты) как значения шкалы

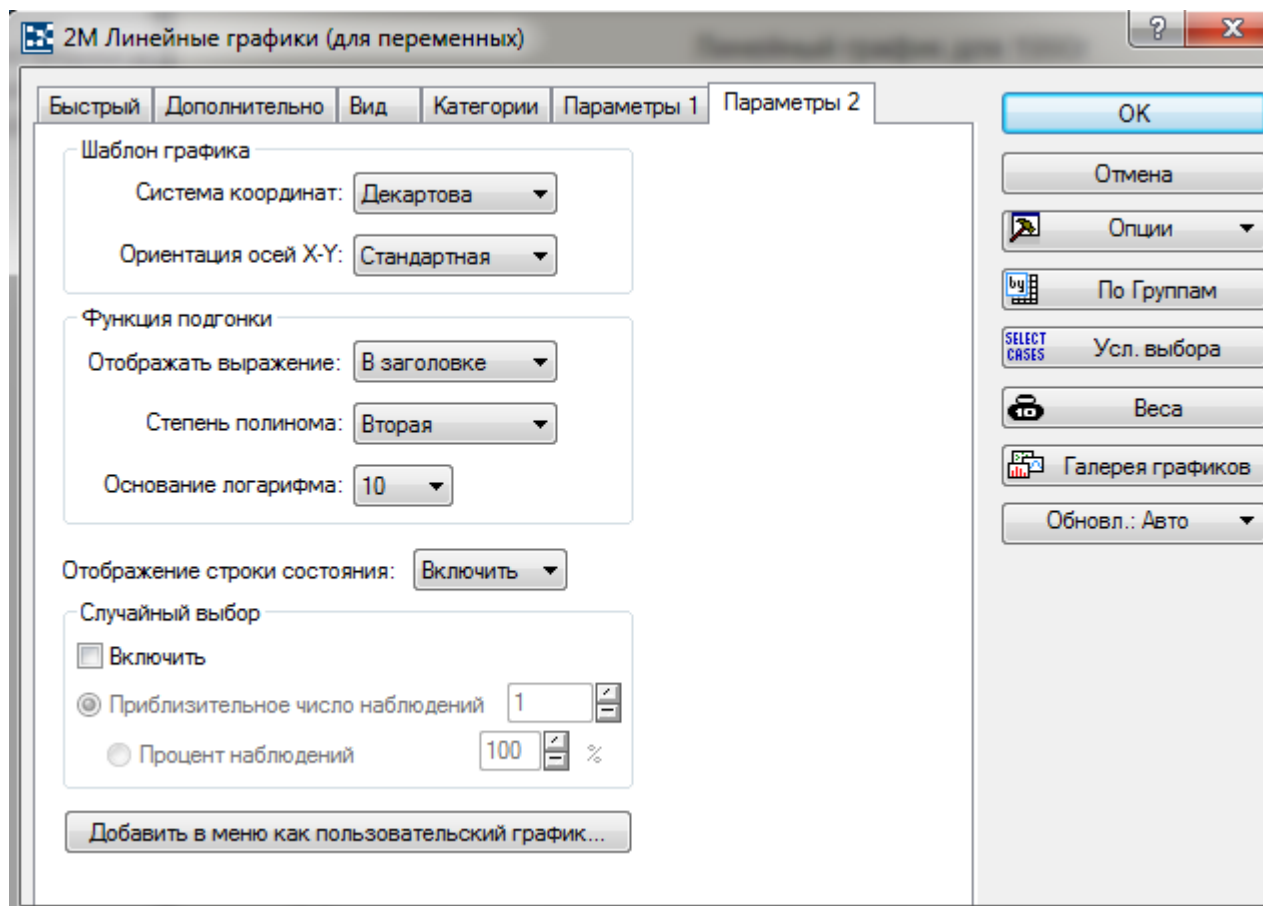
Автообновление
☒ Обновить из Таблицы: Авто ▾
☐ Обновить из Анализа
☒ Обновл. при изменении вход. данных
☐ Без автообновления

Шкала
Ось: X ▾
Шкала: Линейная ▾

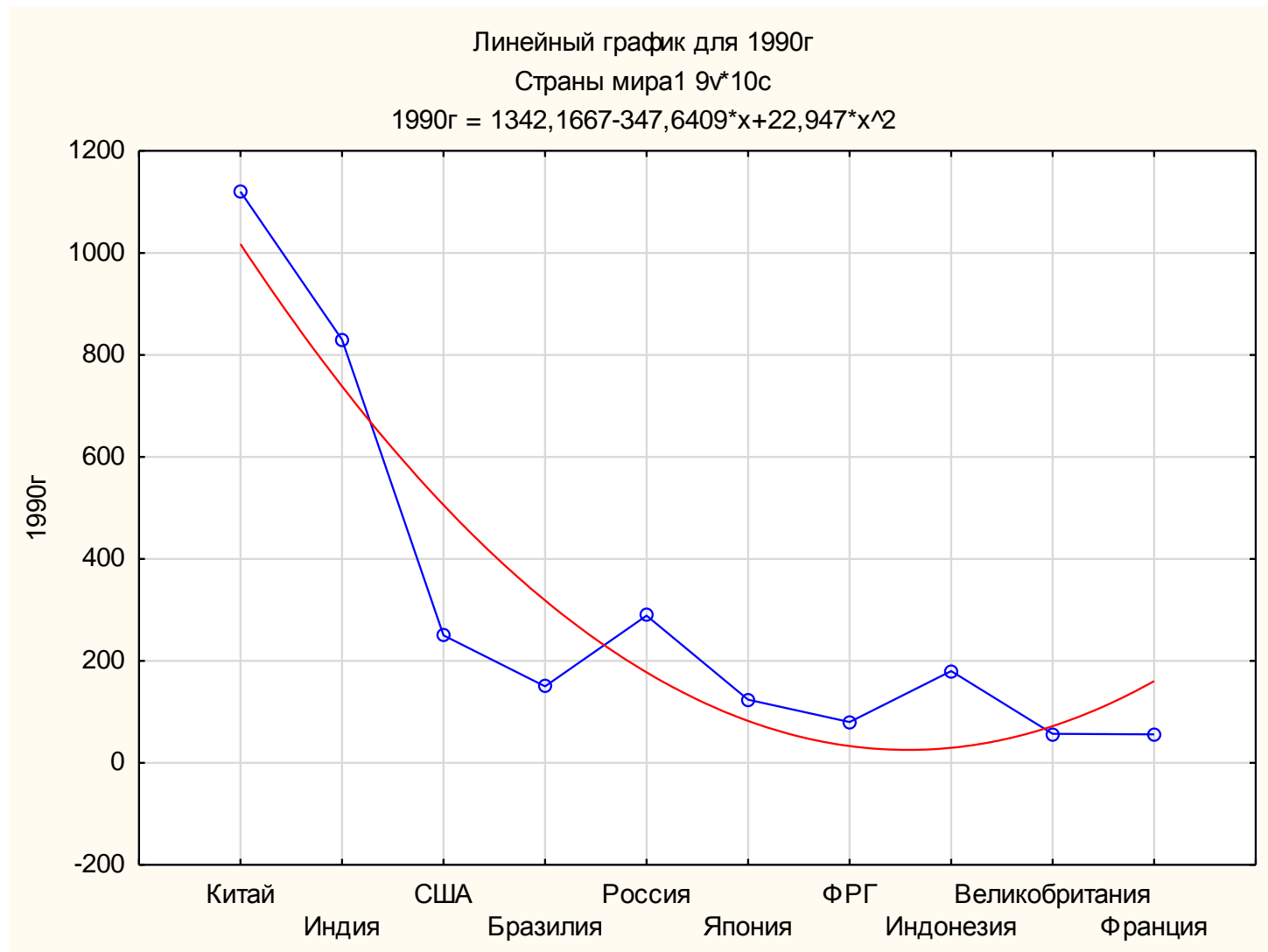
☐ Автоматически закрывать этот диалог после создания графика

ОК
Отмена
Опции ▾
По Группам
Усл. выбора
Веса
Галерея графиков
Обновл.: Авто ▾

На вкладке Параметры 2 можно указать систему координат (декартова, полярная), степень полинома функции подгонки, основания логарифма и т.д.



При сделанных установках программа построит линейный график с кривой подгонки квадратичной функцией

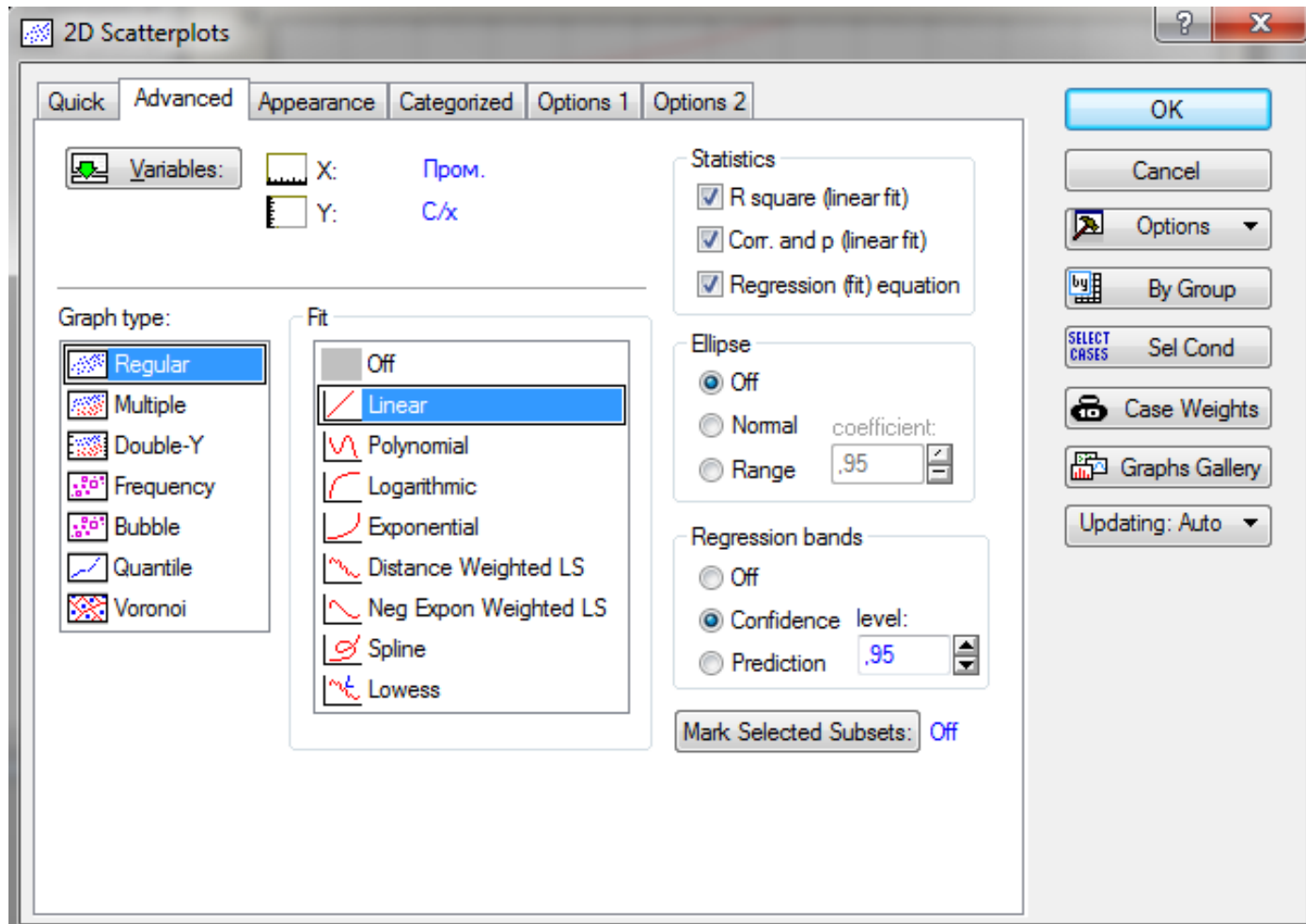


Для построения диаграмм рассеяния (scatterplots) надо выбрать соответствующую команду – Scatterplots.

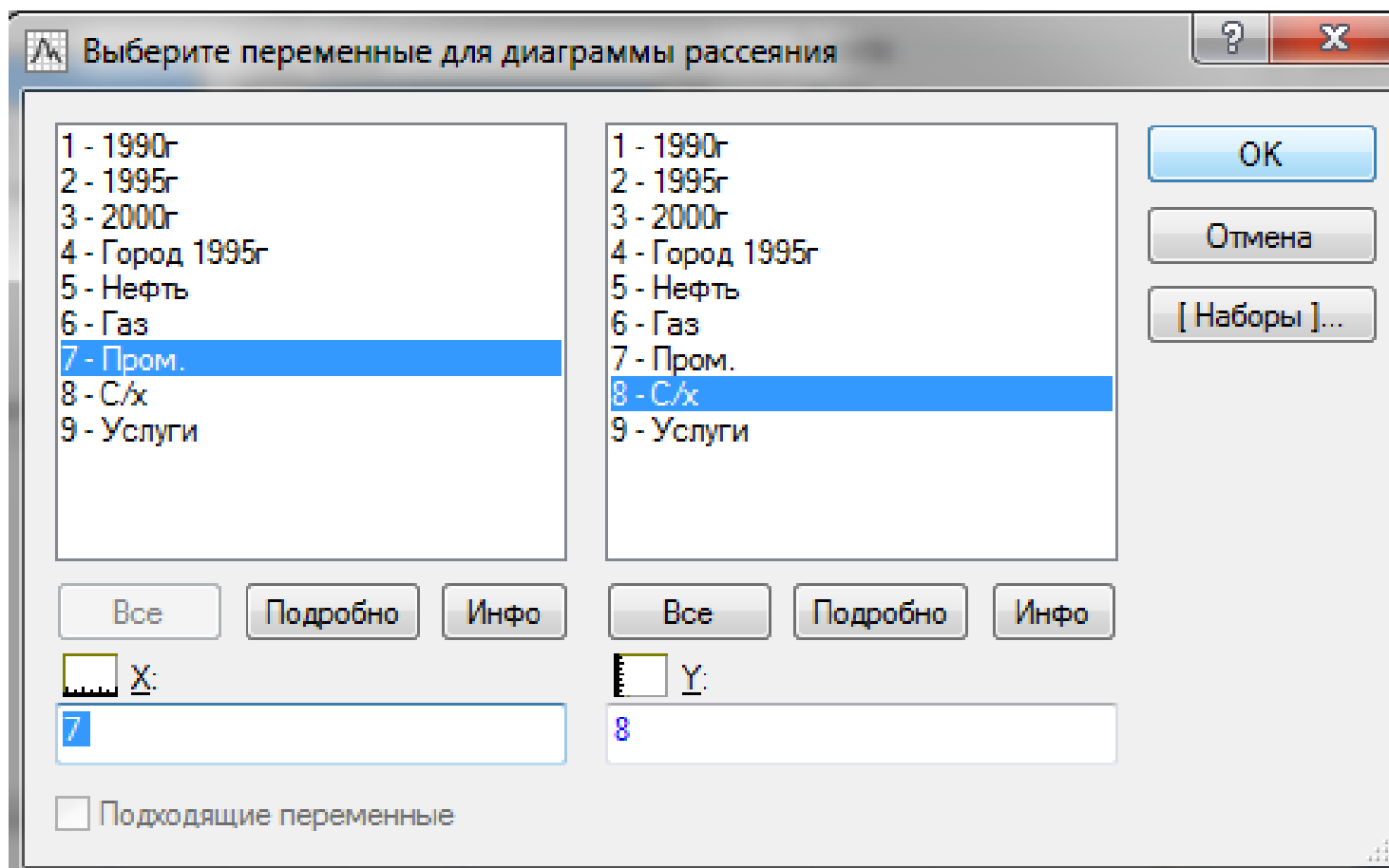
The screenshot shows the Minitab software interface. The 'Graphs' menu is open, and 'Scatterplots...' is selected. The background displays a data table with the following values:

	7	8	9
	Пром.	С/х	Услуги
	48	21	31
	30	29	41
	26	2	72
	37	14	49
	38	7	55
	38	2	60
	38	2	60
	42	17	41
	32	2	66
	27	2	71

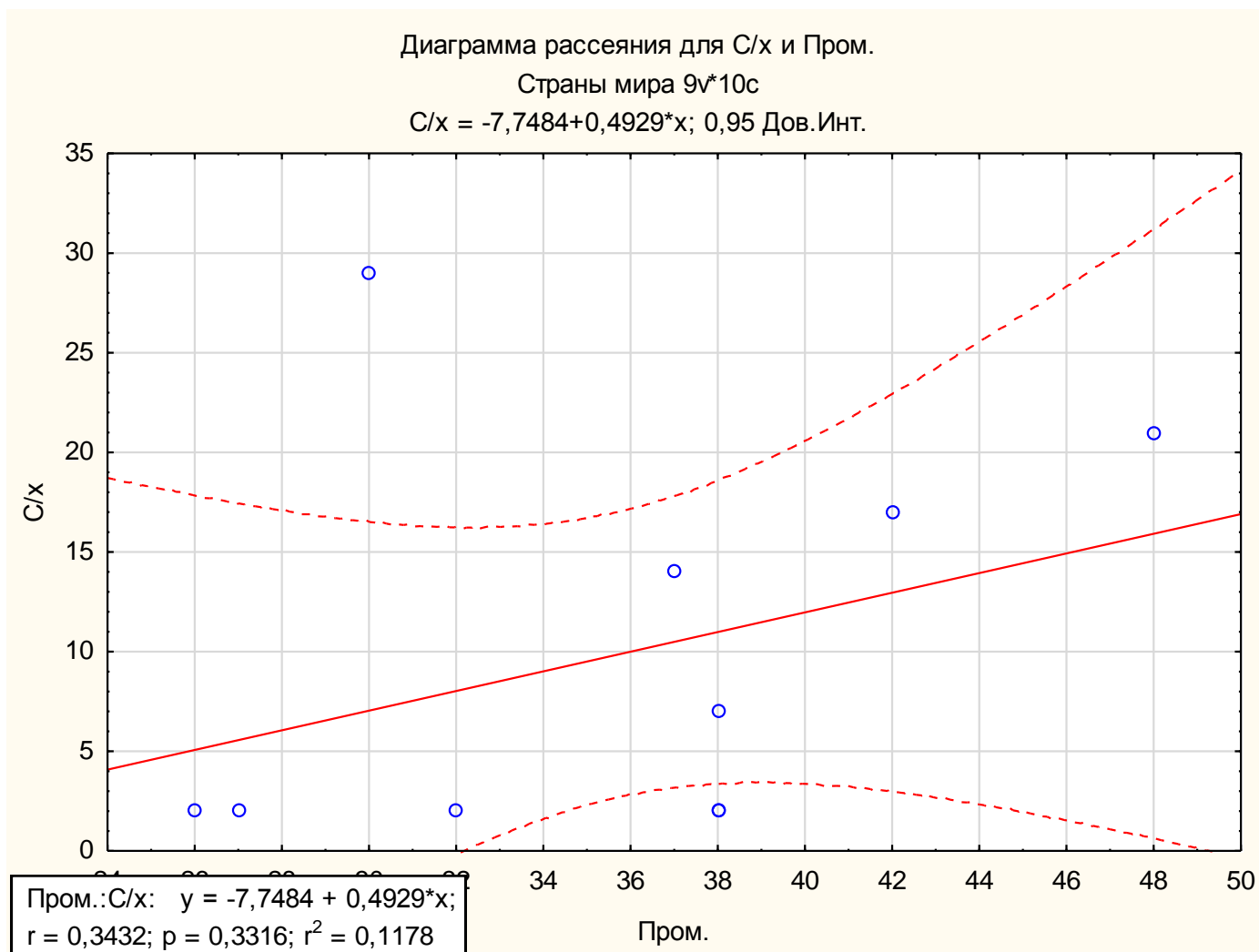
2D Scatterplots Regular (диаграммы рассеяния, простые) визуализируют характер взаимосвязи между двумя переменными X и Y . Для уточнения математической модели взаимосвязи можно экспериментировать с различными типами подгонки. Обозначим флажками все опции в рамке Statistics и щелчком по кнопке Variables



Если после нажатия кнопки *Variables* в открывшемся окне ввести имя независимой переменной *Пром* (ось X), имя зависимой переменной *Сх* (ось Y) и нажать на ОК, то программа построит график, изображенный на рисунке следующего слайда



Из диаграммы видно, что существует умеренная взаимосвязь между долей промышленности и сельским хозяйством в структуре ВВП.



Если убрать настройки в рамке Statistics, а на вкладке Параметры 1 установить флажок на *Отобр. метки наблюдений*, то программа укажет названия стран над соответствующими точками

The screenshot shows the '2M Диаграммы рассеяния' (2M Scatter Plots) dialog box, specifically the 'Параметры 1' (Parameters 1) tab. The dialog has a title bar with a question mark and a close button. The main area contains several sections for configuring the scatter plot:

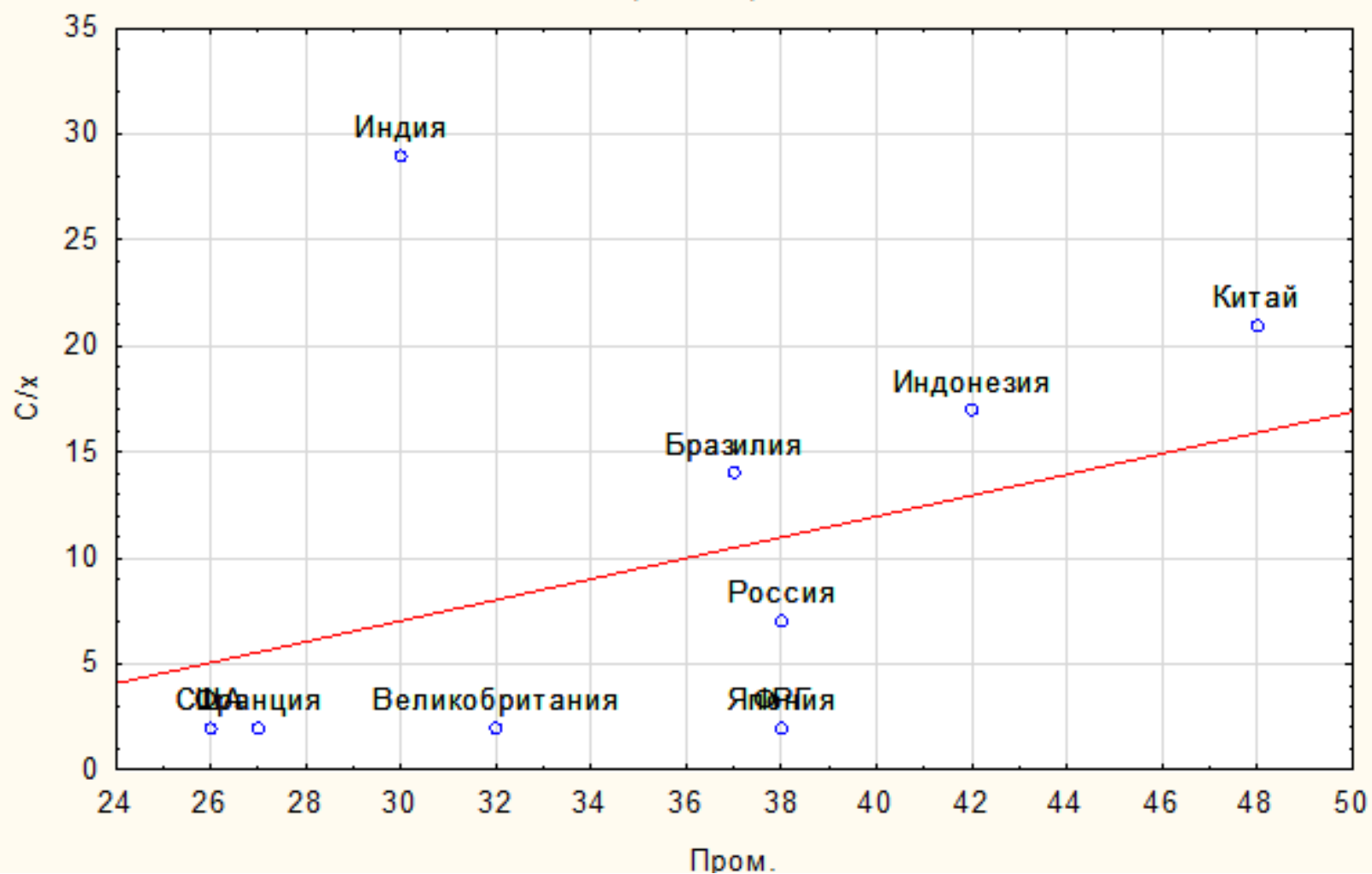
- Заголовок (Header):** Includes a checked checkbox 'Отображать стандартный заголовок' (Show standard header), a text field for the header, and a dropdown menu set to 'Отображать сверху' (Show on top).
- Сноска (Footnote):** Includes a text field and an unchecked checkbox 'Отображать сноску' (Show footnote).
- Параметры отображения (Display parameters):** Includes a dropdown for 'Метки наблюдений' (Observation labels) set to 'Имена наблюдений' (Observation names), a text field for 'Переменная' (Variable) set to '1990г', and two checkboxes: 'Отобр. метки наблюдений' (Selected observation labels) which is checked, and 'Отобр. текстовые метки (или даты) как значения шкалы' (Selected text labels (or dates) as scale values) which is unchecked.
- Автообновление (Auto-update):** Includes three radio buttons: 'Обновить из Таблицы' (Update from Table) which is selected, 'Обновить из Анализа' (Update from Analysis), and 'Без автообновления' (No auto-update). There is also a checked checkbox 'Обновл. при изменении вход. данных' (Update on input data change) and a dropdown for 'Обновить из Таблицы' set to 'Авто' (Auto).
- Шкала (Scale):** Includes a dropdown for 'Ось' (Axis) set to 'X' and a dropdown for 'Шкала' (Scale) set to 'Линейная' (Linear).
- Автоматически закрывать этот диалог после создания графика** (Automatically close this dialog after creating the chart): An unchecked checkbox at the bottom.

On the right side of the dialog, there is a vertical stack of buttons: 'OK', 'Отмена' (Cancel), 'Опции' (Options) with a dropdown arrow, 'По Группам' (By Groups), 'Усл. выбора' (Selection criteria) with a 'SELECT CASES' icon, 'Веса' (Weights) with a scale icon, 'Галерея графиков' (Chart gallery) with a gallery icon, and 'Обновл.: Авто' (Update: Auto) with a dropdown arrow.

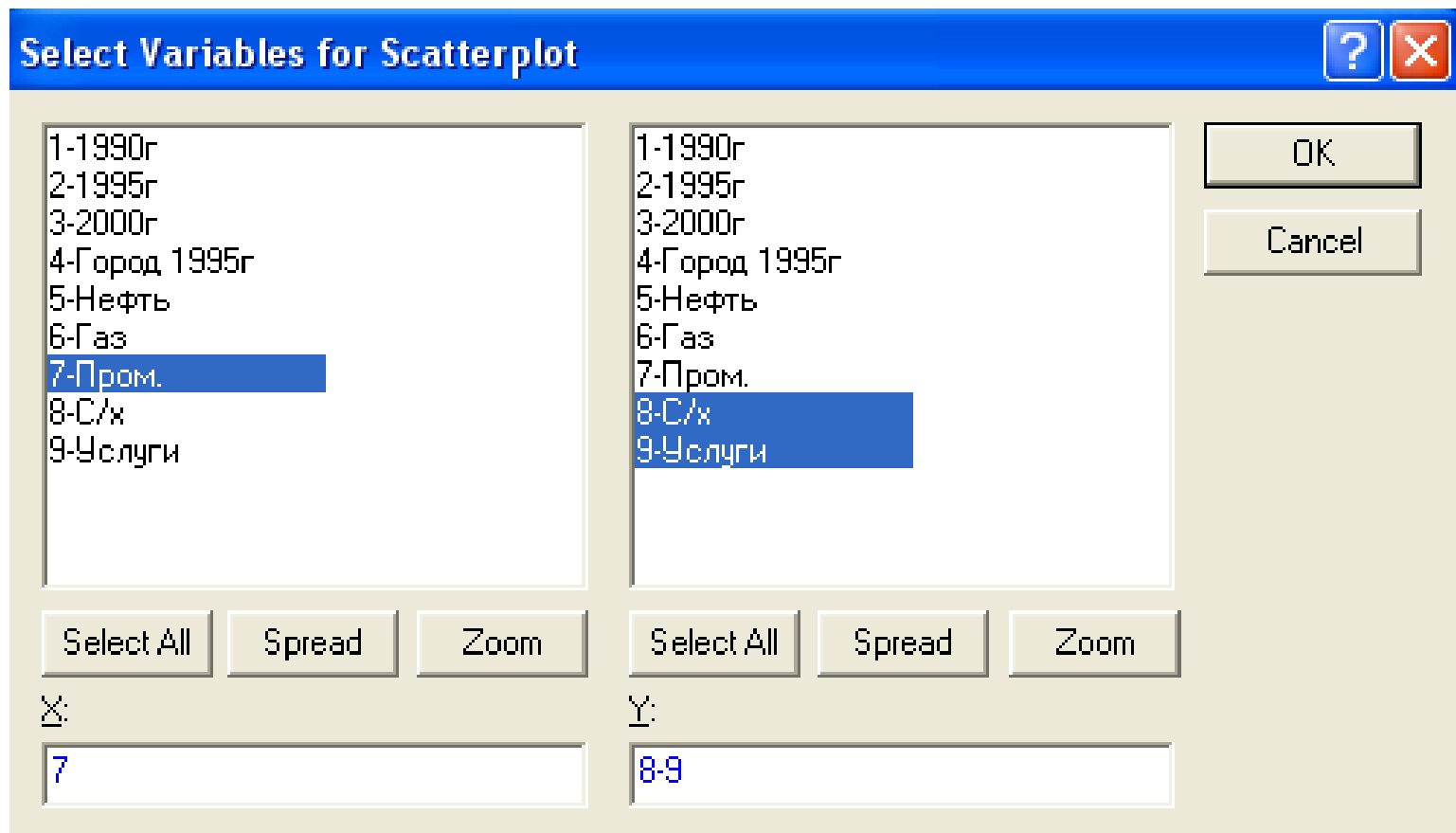
Диаграмма рассеяния для C/x и Пром.

Страны мира 19v*10с

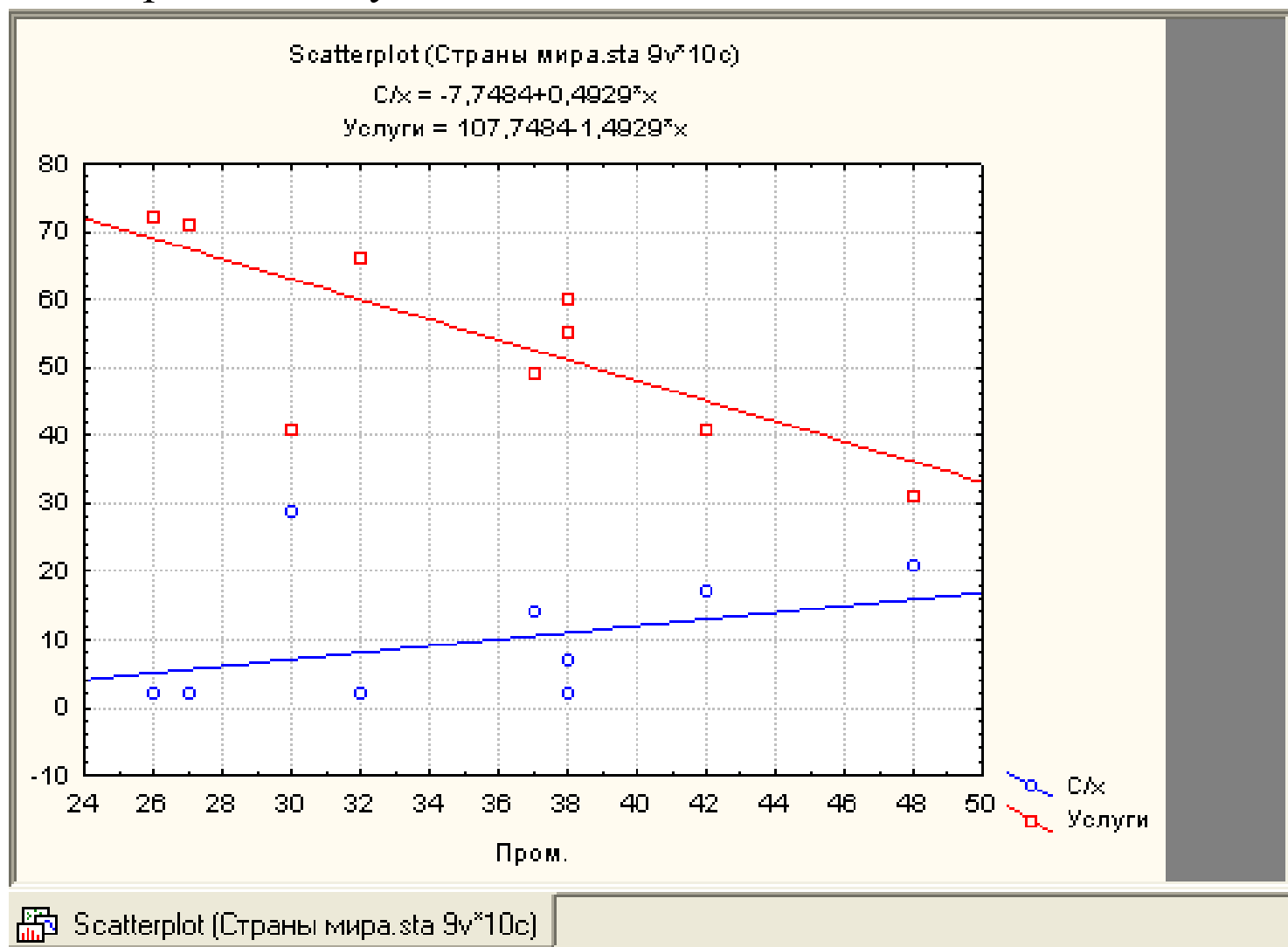
$$C/x = -7,7484 + 0,4929 * x$$



2D Scatterplots Multiple (составные). В отличие от простой диаграммы рассеяния, на которой одна переменная представлена по горизонтальной, а вторая – по вертикальной оси, составная диаграмма рассеяния состоит из нескольких зависимостей и изображает несколько корреляций. Значения одной переменной (X) откладываются по горизонтальной оси, а по вертикальной – значения нескольких переменных (Y).



При построении этих графиков следует помнить, что переменные принимающие значения на одной оси Y должны иметь одинаковую единицу измерения и быть величинами одного порядка, иначе одна из зависимостей превратится в горизонтальную линию



2D Scatterplots Double-Y (с двойной осью Y). Диаграмму рассеяния такого типа можно рассматривать как комбинацию двух составных диаграмм рассеяния для одной переменной X и двух различных наборов (списков) переменных Y . Для переменной X и каждой из переменных Y будет построена диаграмма рассеяния, но переменные из первого списка (*Left Y*) будут откладываться по левой оси Y , в то время как переменные из второго списка (*Right Y*), будут откладываться по правой оси Y . В силу независимости масштабов, используемых для двух списков переменных, этот график позволяет изображать зависимости между переменными отличающимися порядком и единицами измерения.

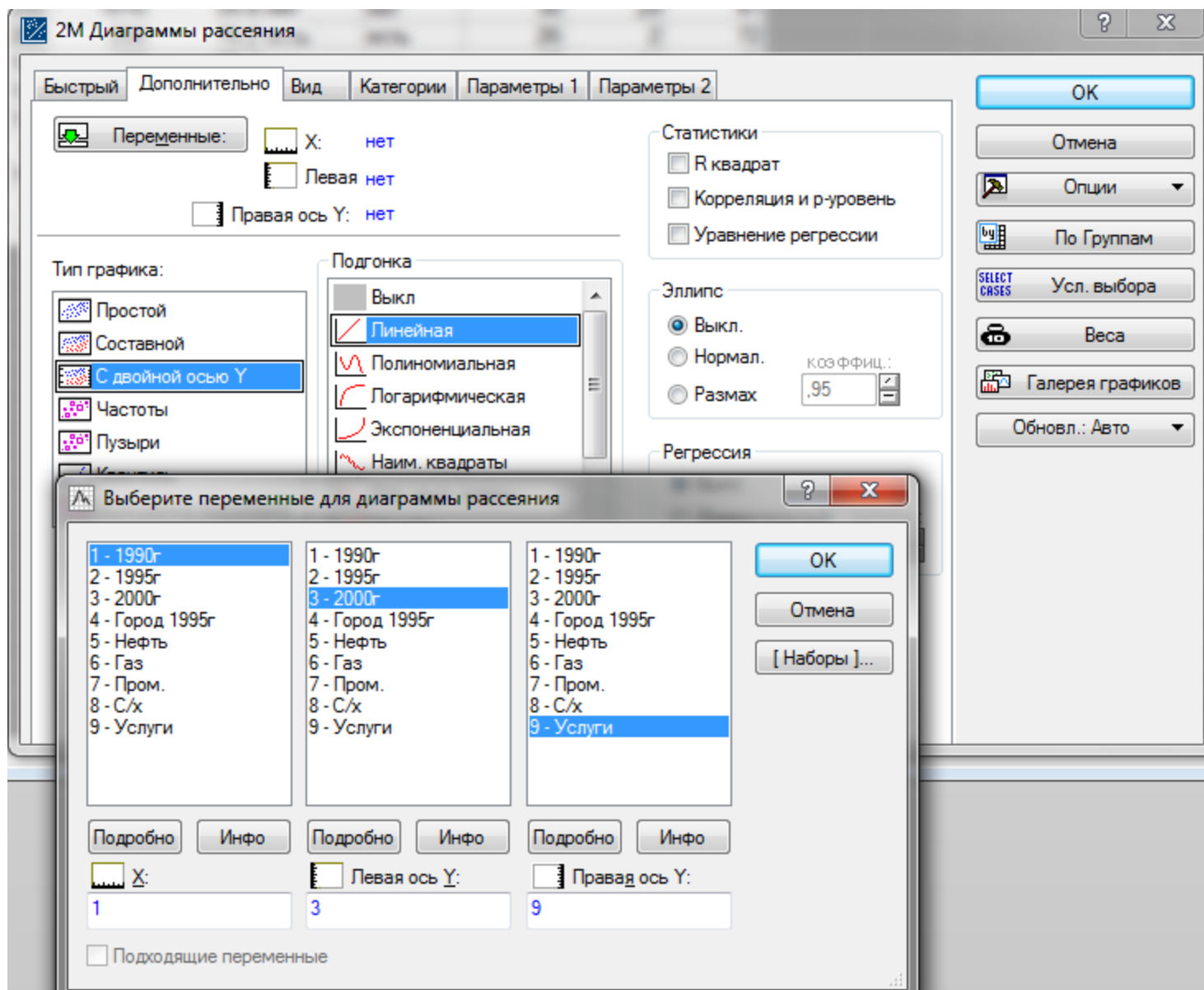
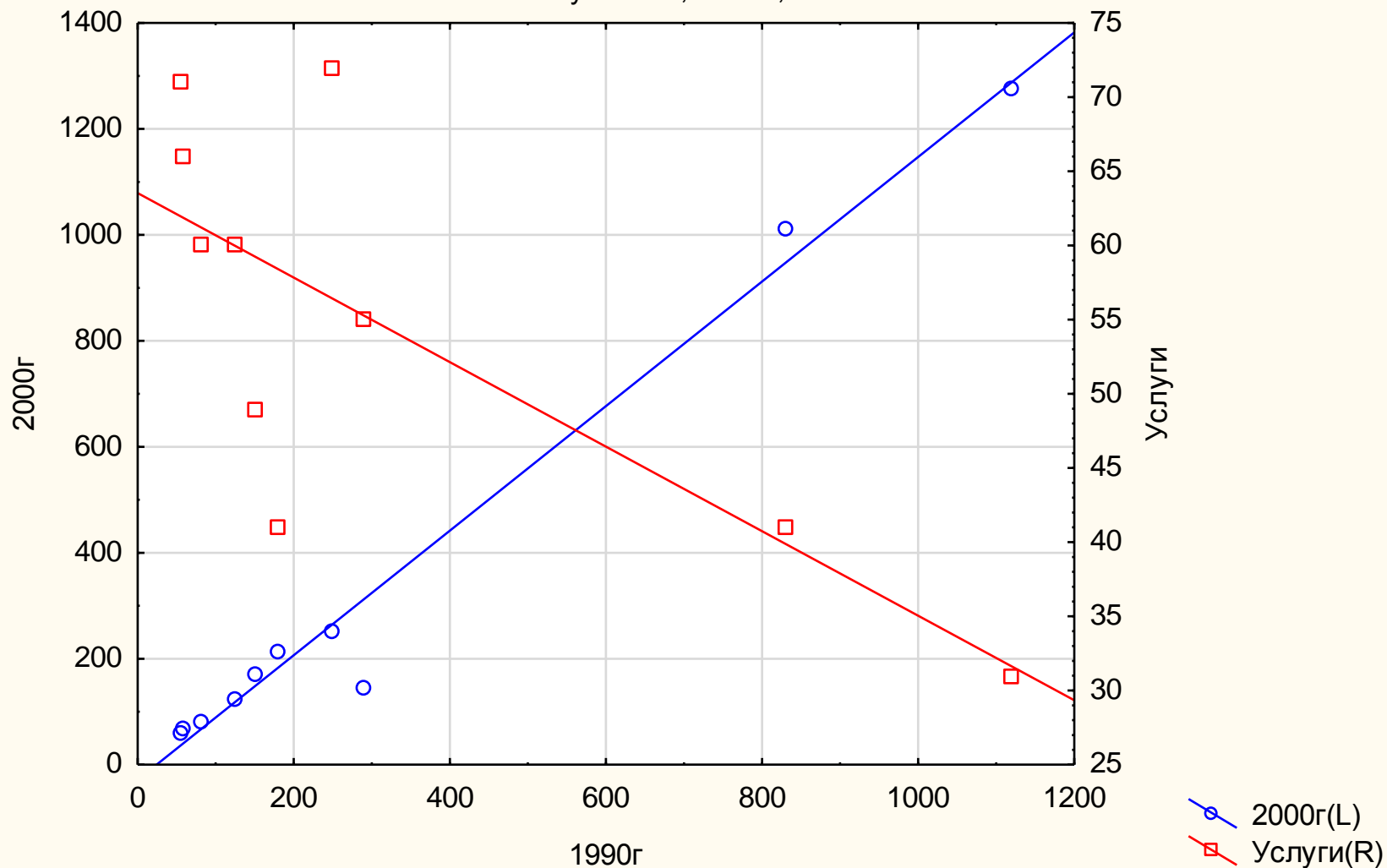


Диаграмма рассеяния для нескольких переменных и 1990г

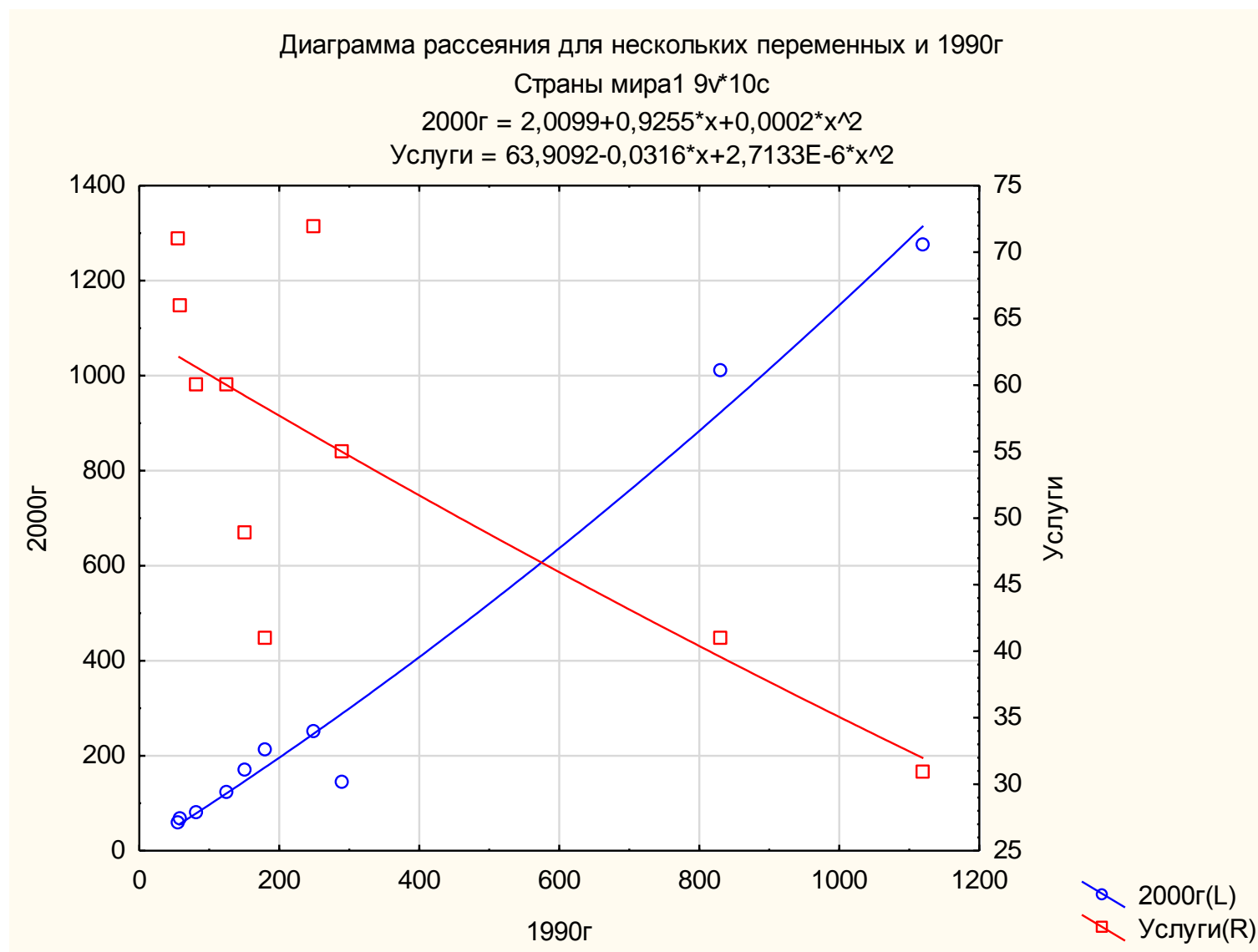
Страны мира1 9v*10с

$$2000г = -28,4784 + 1,1756 * x$$

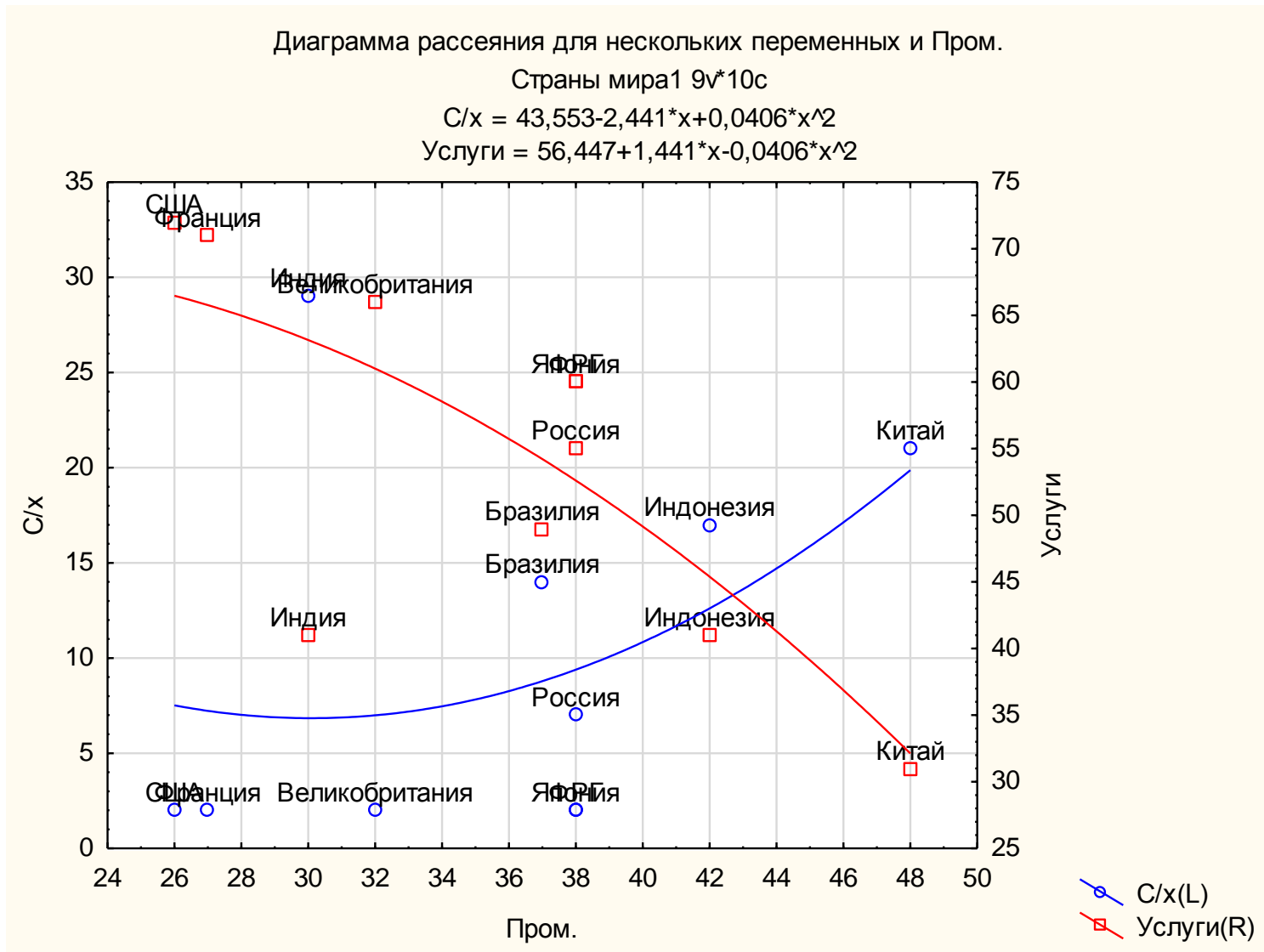
$$\text{Услуги} = 63,5344 - 0,0285 * x$$



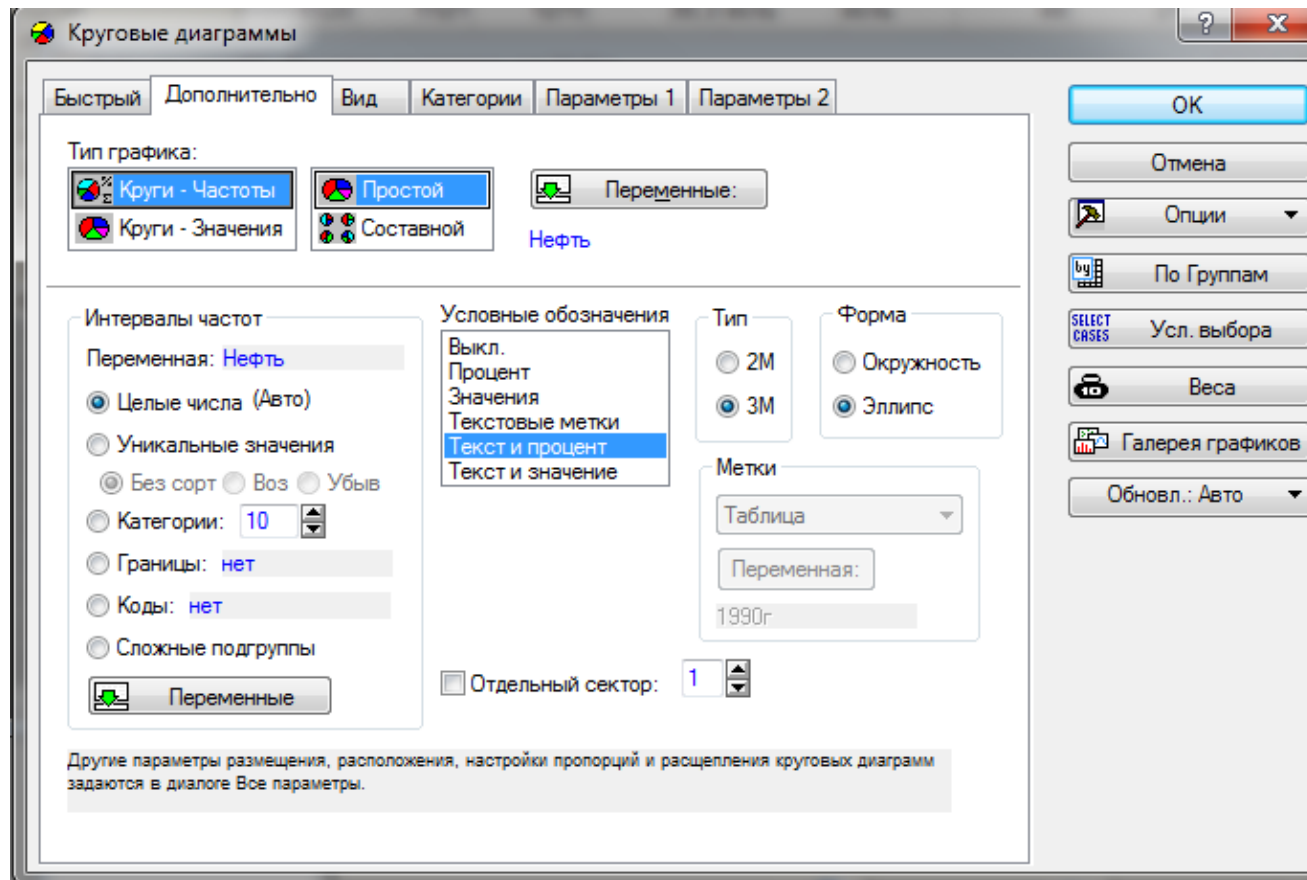
Можно попытаться произвести подгонку математической зависимости какой-либо другой нелинейной функцией, например полиномиальной

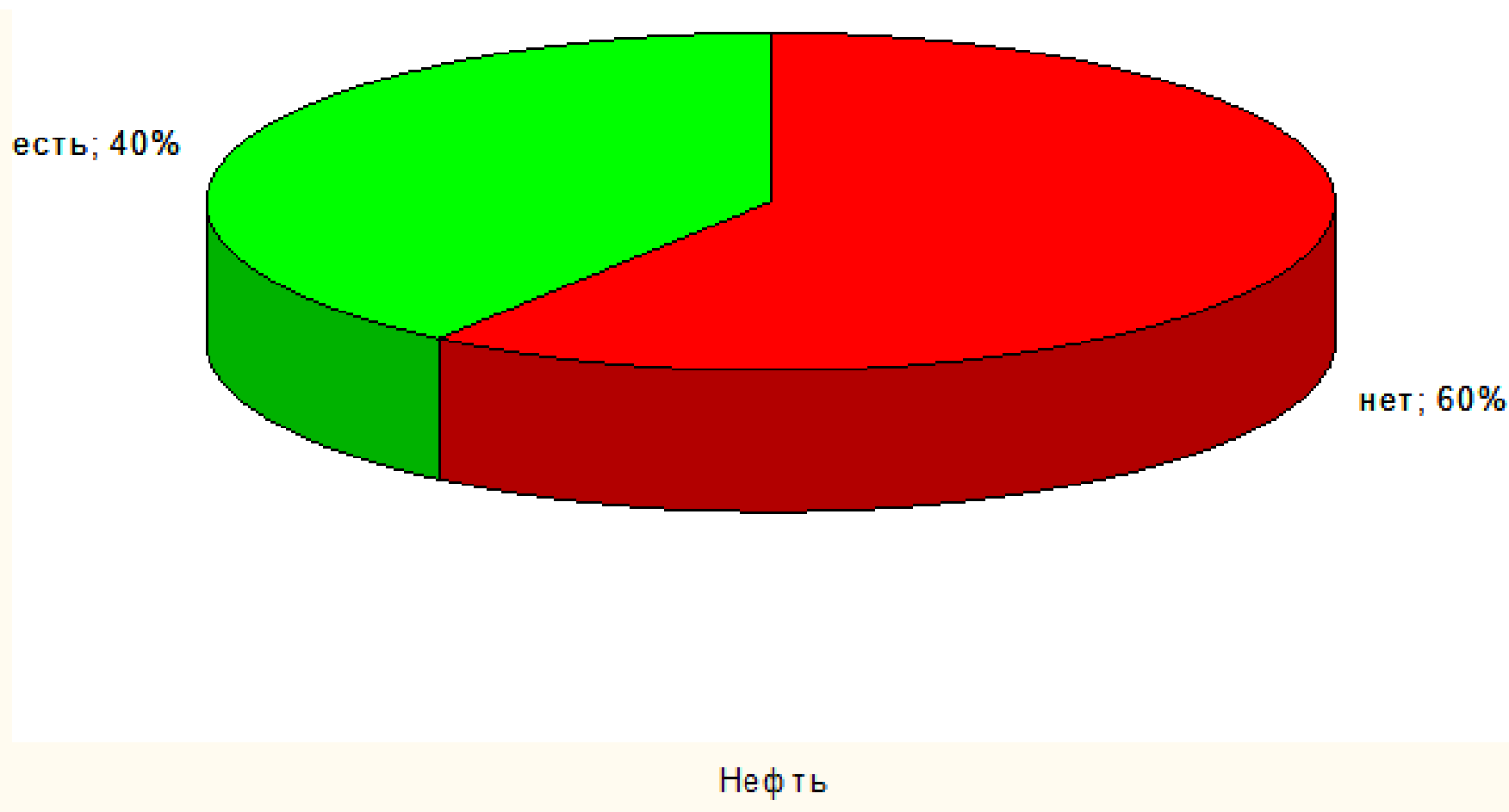


Если отобразить метки наблюдений, то получим следующий график



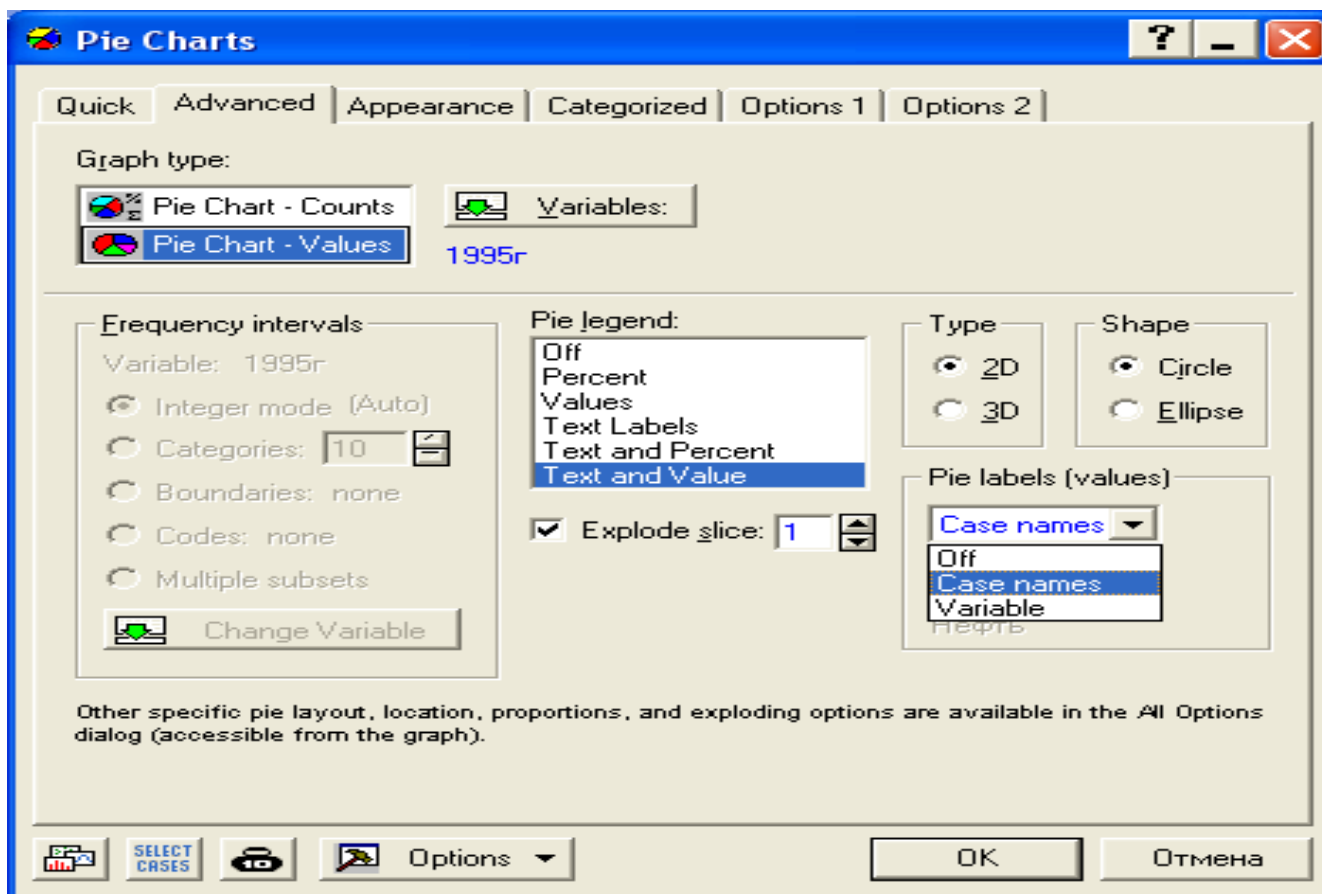
Для построения круговой диаграммы надо в верхнем меню **Graphs** выбрать **2D Graphs**→ **Pie Charts** (круговые диаграммы). В поле **Graph type** указывается тип круговой диаграммы, например *Pie Chart–Counts* (Круги - частоты). Относительные частоты категориальной переменной будут изображены в виде круговых секторов размеров, пропорциональным относительным частот.



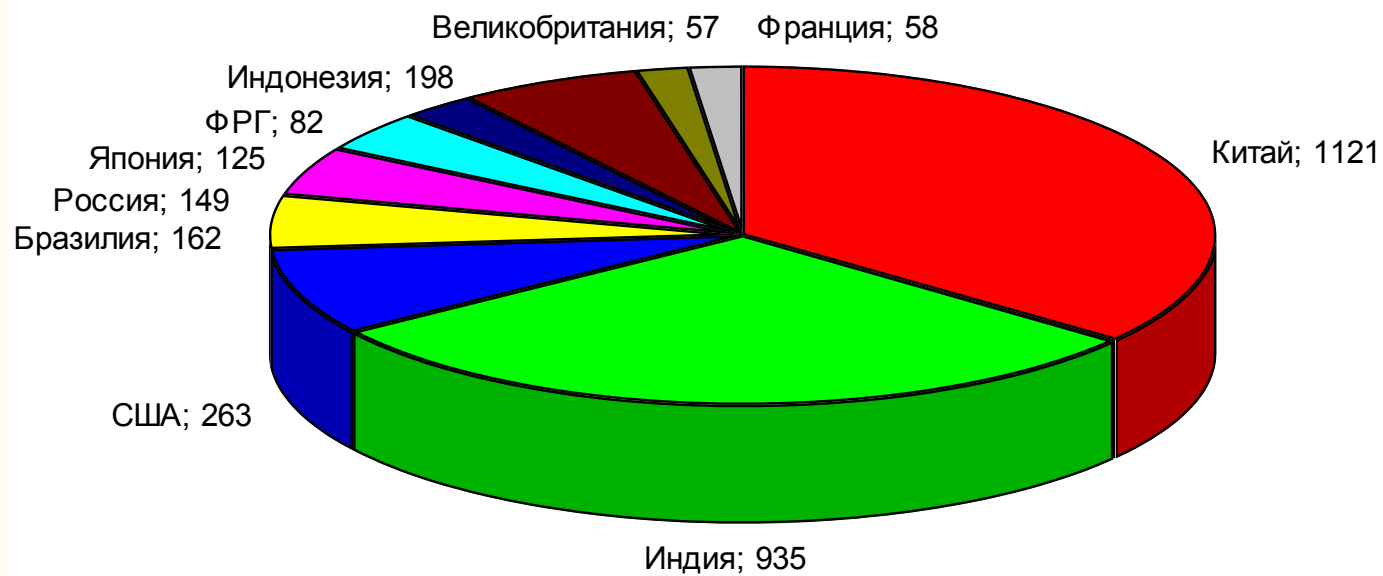


В 10-й версии

Pie Chart–Values (круговая диаграмма значений), последовательность значений **количественной** переменной изображается в виде последовательности круговых секторов, размер каждого сектора пропорционален значению переменной, одно наблюдение соответствует одному сектору. Как пример, построим диаграмму при установках представленных на изображении окна

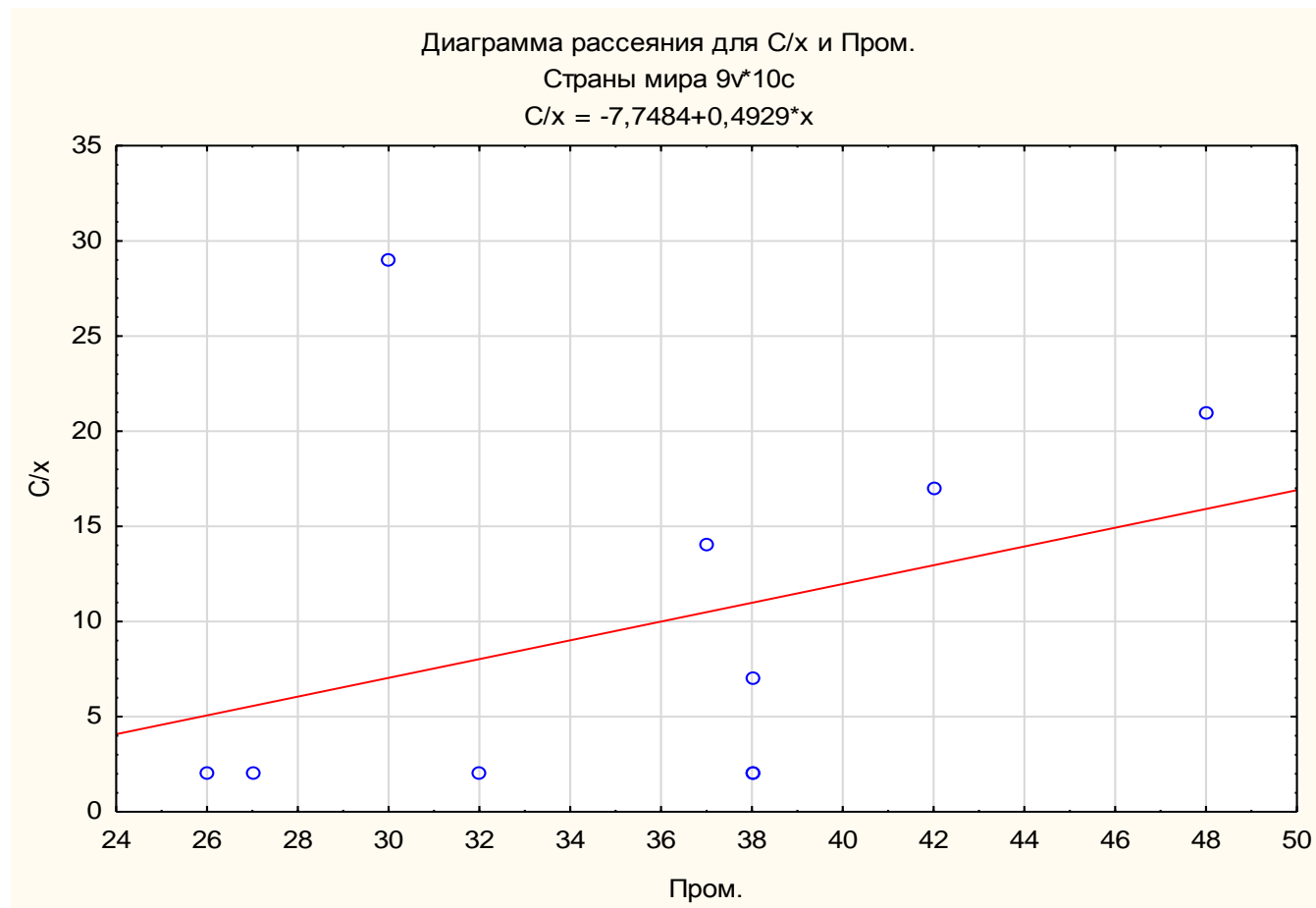


Круговая диаграмма для 1995г
Страны мира 9v*10с

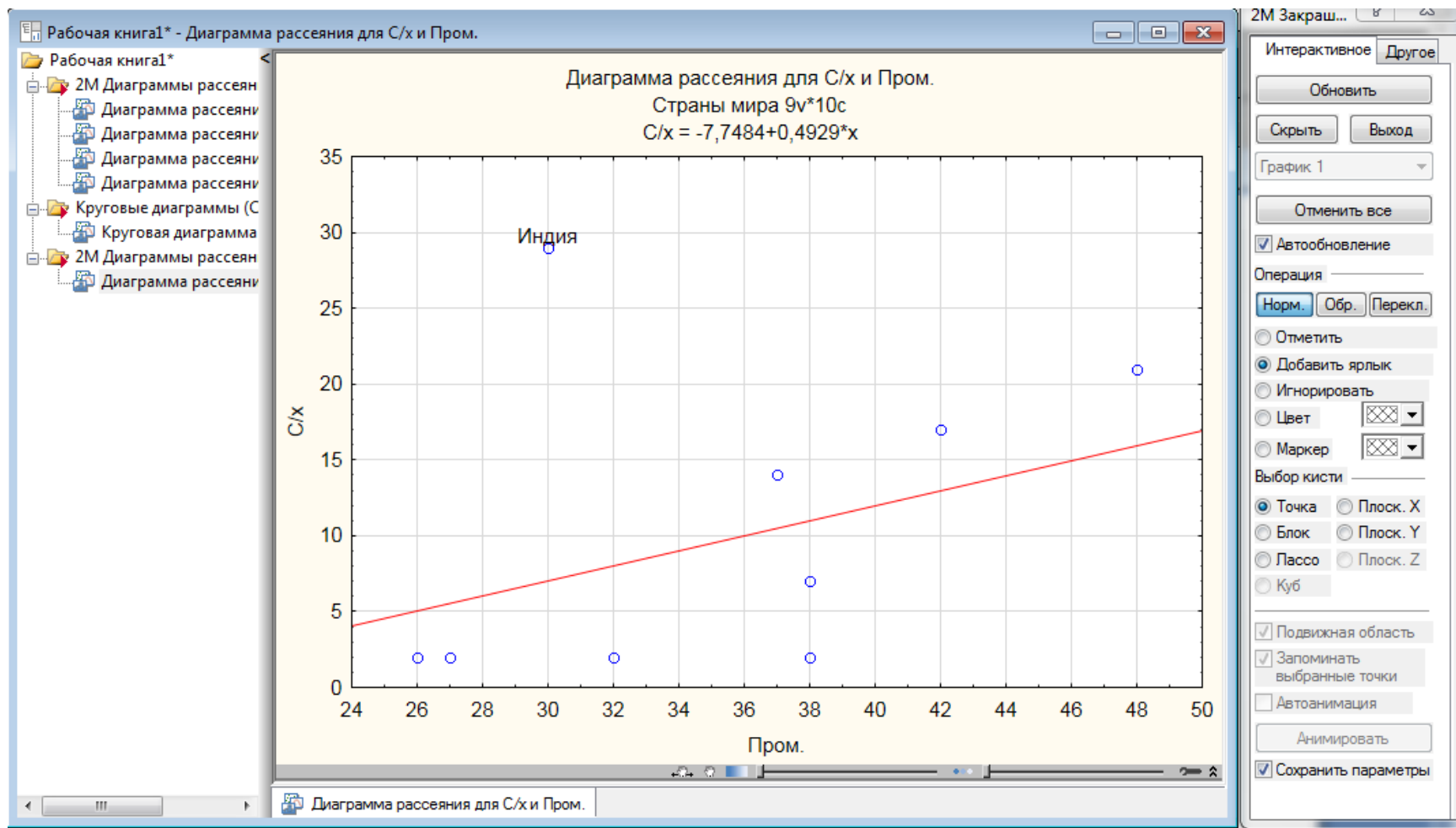
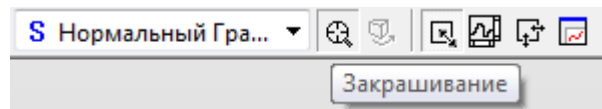


1995г

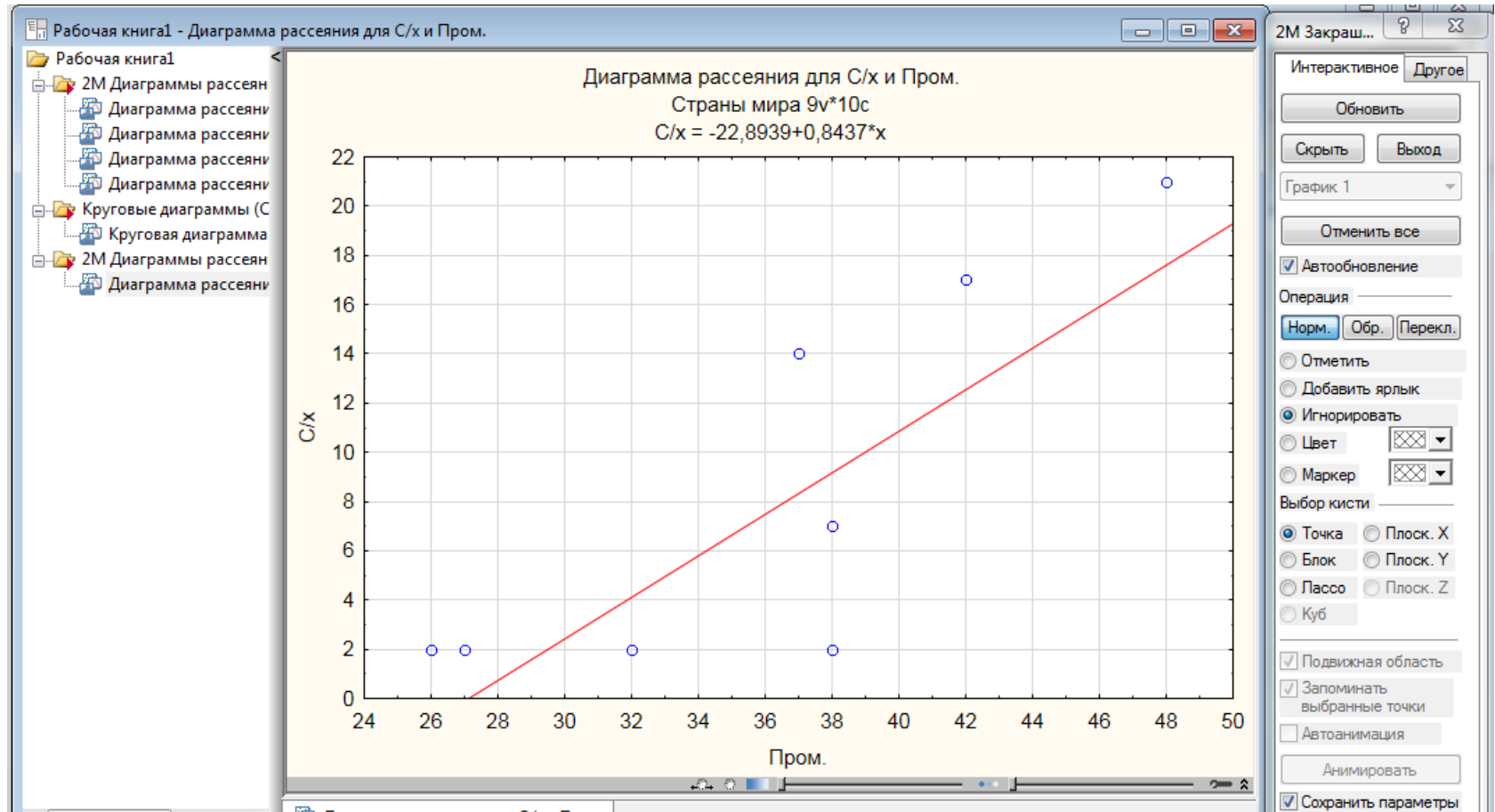
Важным преимуществом диаграмм рассеяния является возможность находить «выбросы» (аномальные, нетипичные данные), которые влияют на значение коэффициента корреляции. Даже один выброс может значительно изменить коэффициент корреляции между двумя переменными. Средство **Brushing** (закрашивание) интерактивно удаляет выбросы, при этом можно непосредственно наблюдать за изменением аппроксимирующей функции или линии регрессии. Разберем как определить название страны, соответствующей выбросу на диаграмме рассеяния



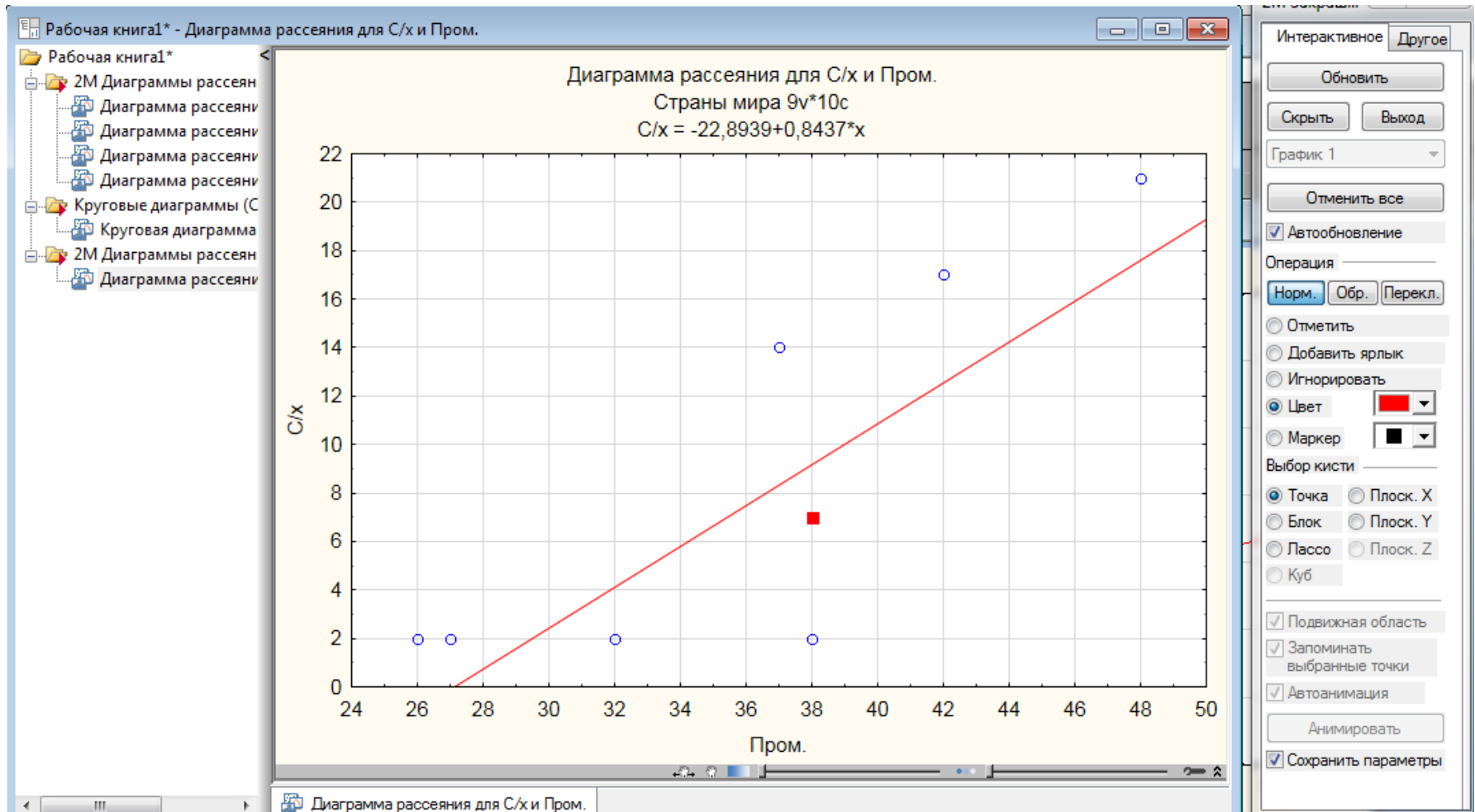
Если подвести курсор к иконке с изображением прицела, появится надпись «закрашивание», щелкнуть по ней, появится диалог (на рисунке справа от графика). Если при установках как на рисунке (по умолчанию) подвести курсор в виде прицела к выбросу и щелкнуть по точке, появится надпись – метка Индия



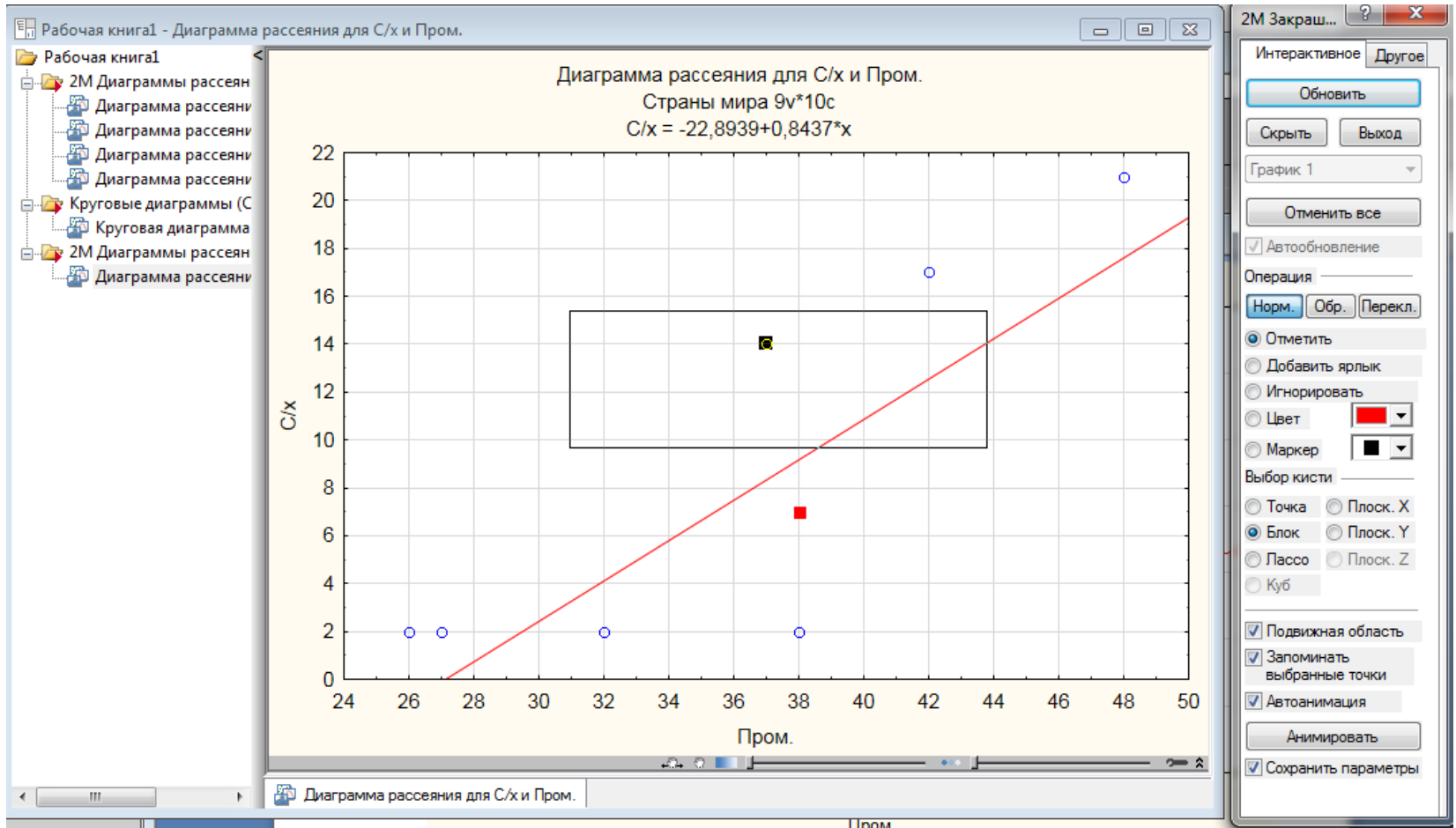
Если выделить опцию «Игнорировать» и вновь подвести прицел к точке с надписью Индия и щелкнуть по точке, точка с надписью будет удалена с графика, при этом изменится уравнение регрессии



Если выделив опцию цвет, выбрать цвет, потом отметить опцию «маркер», выбрать тип маркера, и щелкнуть прицелом по точке, то точка будет обозначена маркером в соответствующем цвете



Если в рамке «Выбор кисти» выделить Блок, активной станет нижняя часть диалога средства Закрашивания. Если установить флажок на Автоанимация и курсором на графике выделить прямоугольную область, то она начнет перемещаться параллельно осям координат и высвечивать точки



2D Histogramms являются графическими представлениями распределения частот выбранных переменных. Для каждого интервала (класса) рисуется столбец, высота которого пропорциональна частоте класса.

Гистограмма наглядно показывает, какие значения или диапазоны значений исследуемой переменной являются наиболее частыми, насколько сильно они различаются, как сконцентрировано большинство наблюдений вокруг среднего, является ли распределение симметричным или нет, имеет ли оно моду или несколько мод. Различают несколько видов гистограмм.

2D Histogramms Regular (простые) представляет собой столбчатую диаграмму распределения частот для выбранной переменной (если выбрано более одной переменной, то для каждой из них будет построен отдельный график).

2D Histogramms Multiple (составные) изображают распределение частот для нескольких переменных на одном графике.

2D Histogramms Double-Y (с двойной осью Y). Гистограмму с двойной осью Y можно считать комбинацией двух поразному масштабированных составных гистограмм.

2D Histogramms Hanging Bars (висячие столбцы). Гистограмма висячих столбцов является «наглядным критерием проверки на нормальность распределения», который помогает определить области распределения, где возникают расхождения между наблюдаемыми и ожидаемыми нормальными частотами.

2D Histograms



Quick | Advanced | Appearance | Categorized | Options 1 | Options 2



Variables:

Vars: none

Graph type:



Regular



Multiple



Double-Y

Showing Type

Standard

Standard

Hanging Bars

Cumulative

Stacked

Y axis:

N

Fit type:



Off



Normal



Beta



Exponential



Extreme



Gamma

Statistics



Shapiro-Wilk test



Total count



Descriptive statistics



Kolmogorov-Smirnov test

Intervals

Variable: none



Integer mode ☒ Auto



Categories: 10



Boundaries: none



Codes: none



Multiple subsets



Change Variable



SELECT
CASES



Options

OK

Отмена



Параметры 2

Переменная: нет

Тип подгонки:

 Геометрическое

Ось Y: N ▼

☐ Критерий Колмогорова-Смирнова

Переменные

Обновл.: Авто

На вкладке **Advanced** (дополнительно) в поле **Graph type** (тип графика) указывается тип графика: *Regular*; *Multiple*; *Double-Y*.

В поле **Fit type** (тип подгонки) выбираются виды аппроксимирующих законов плотностей распределений: *Of* (выключить); *Normal*; *Beta*; *Exponential* и т.д.

В поле **Showing type** (тип показа) указываются форматы графиков: *Standard*; *Hanging Bars*; *Cumulative*. Последний формат дает графическое изображение накопленных частот.

В рамке **Intervals** (интервалы) производятся установки режимов категоризации.

В режиме *Integer mode* (целые числа), если не установлена галочка в поле **Auto**, программа округлит каждое значение выделенной переменной до целого числа и создаст одну категорию для каждого целочисленного значения, тогда высота прямоугольника будет соответствовать количеству таких целочисленных значений. Другими словами все дробные величины округлит до целых и на гистограмме изобразит их частоту.

- Если выделены обе опции *Integer mode* и *Auto*, то при непрерывной переменной (рис.11) построит гистограмму с целочисленными концами интервалов, разделив весь диапазон на определенное число интервалов одинаковой длины и изобразит частоты дробных значений внутри интервалов с целыми концами. Для категориальной переменной (рис.12) будет построена гистограмма с высотой прямоугольников, соответствующих частоте значений переменной
- В поле ввода справа от режима *Categories* (категории) вводится необходимое число категорий. Программа разделит полный диапазон значений переменной на заданное число интервалов одинаковой длины (длина интервалов не будет целым числом) и изобразит частоты непрерывной переменной внутри интервалов (рис.13 и 14).
- После выбора опции *Boundaries* (границы) надо нажать кнопку **Specify Boundaries** (задать границы) и ввести список границ для выделенной переменной в появившемся диалоговом окне.
- Опцию *Codes* (коды) можно использовать, если переменная категориальная. После выбора этой опции надо нажать кнопку **Specify Codes** (задать коды) и ввести нужные коды в появившемся диалоговом окне для которых и будут посчитаны и изображены частоты. Процедура работает, если в поле **Fit type** выбран режим *Off* (рис.13).

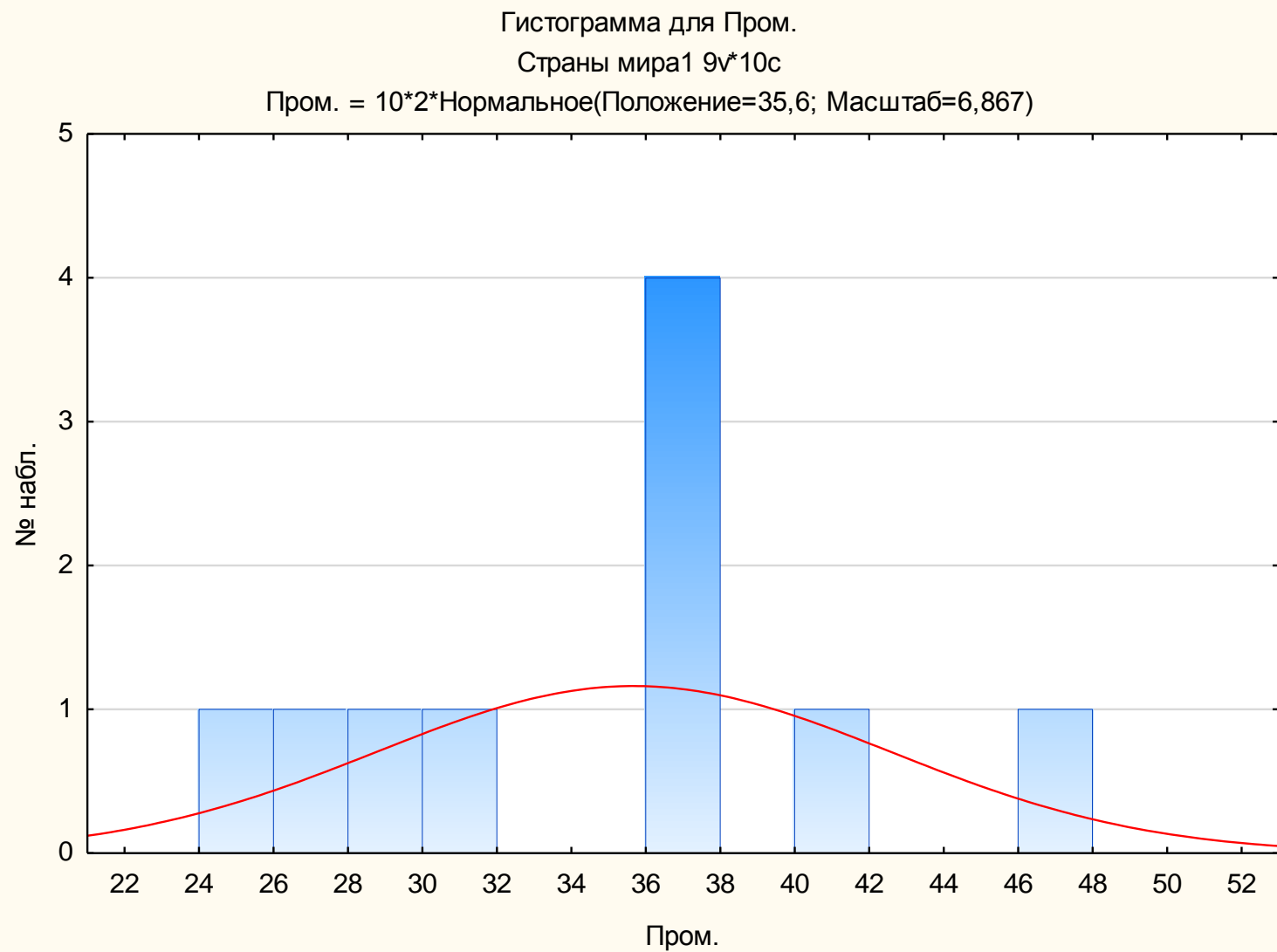


Рис.11

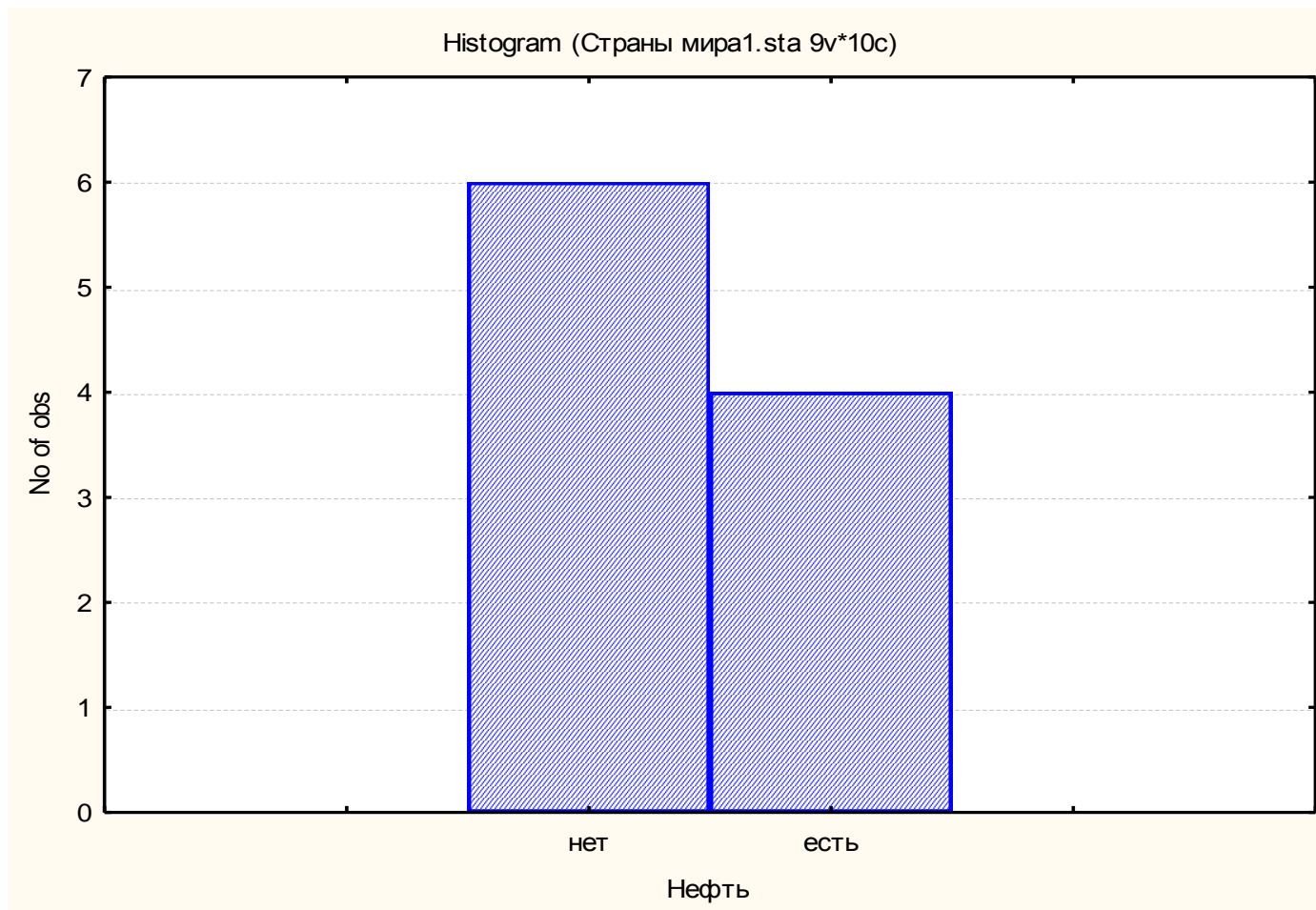


Рис.12

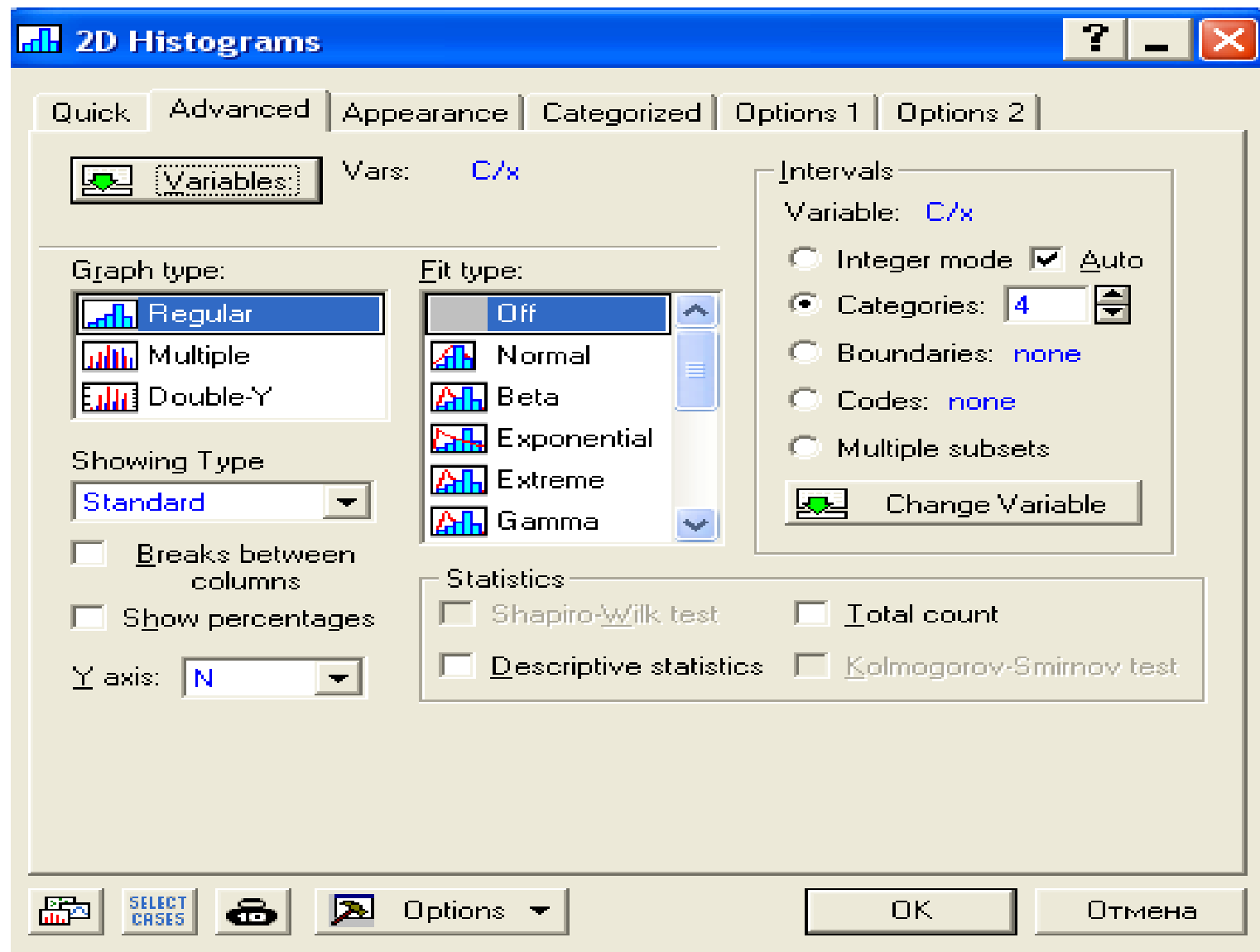


Рис.13

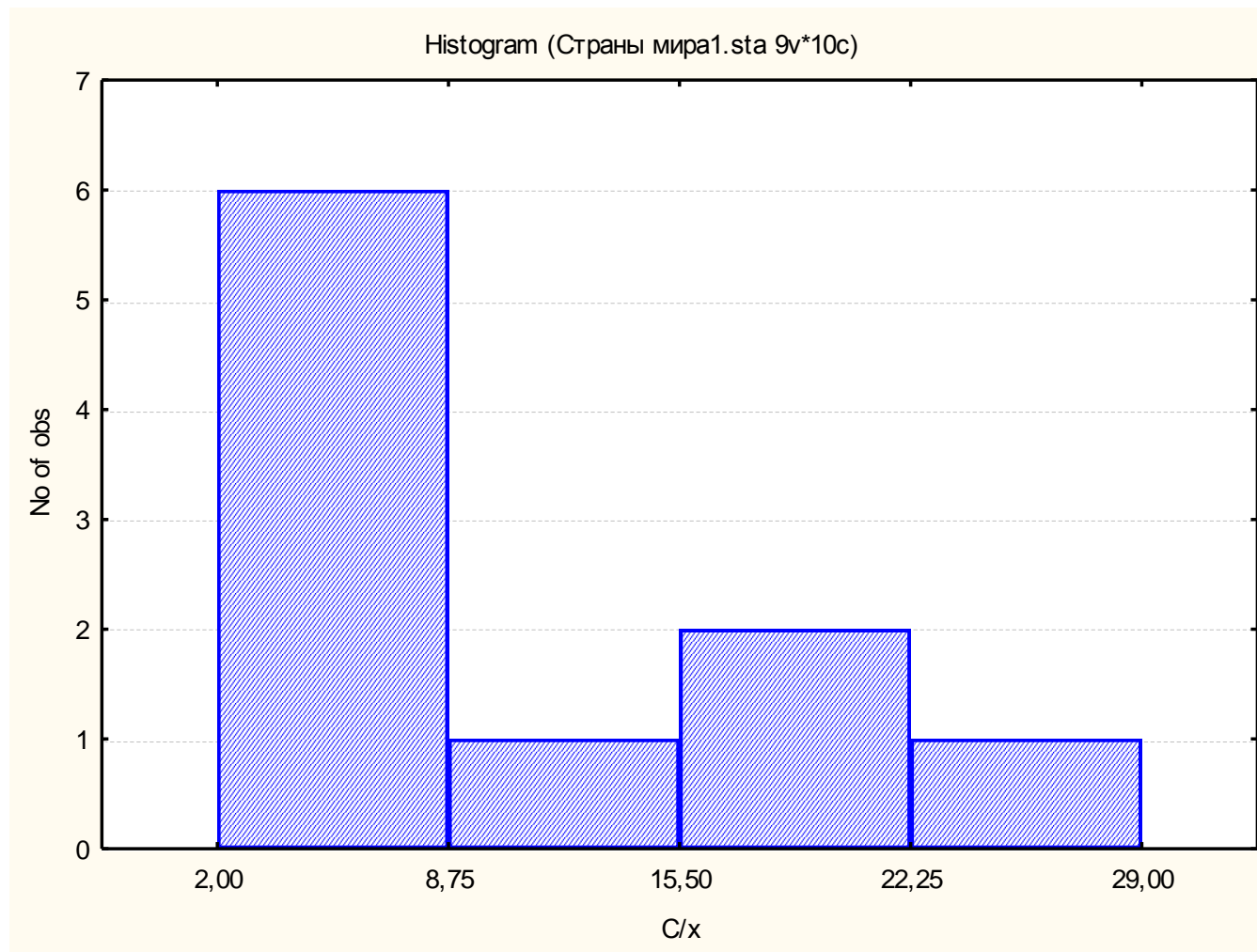
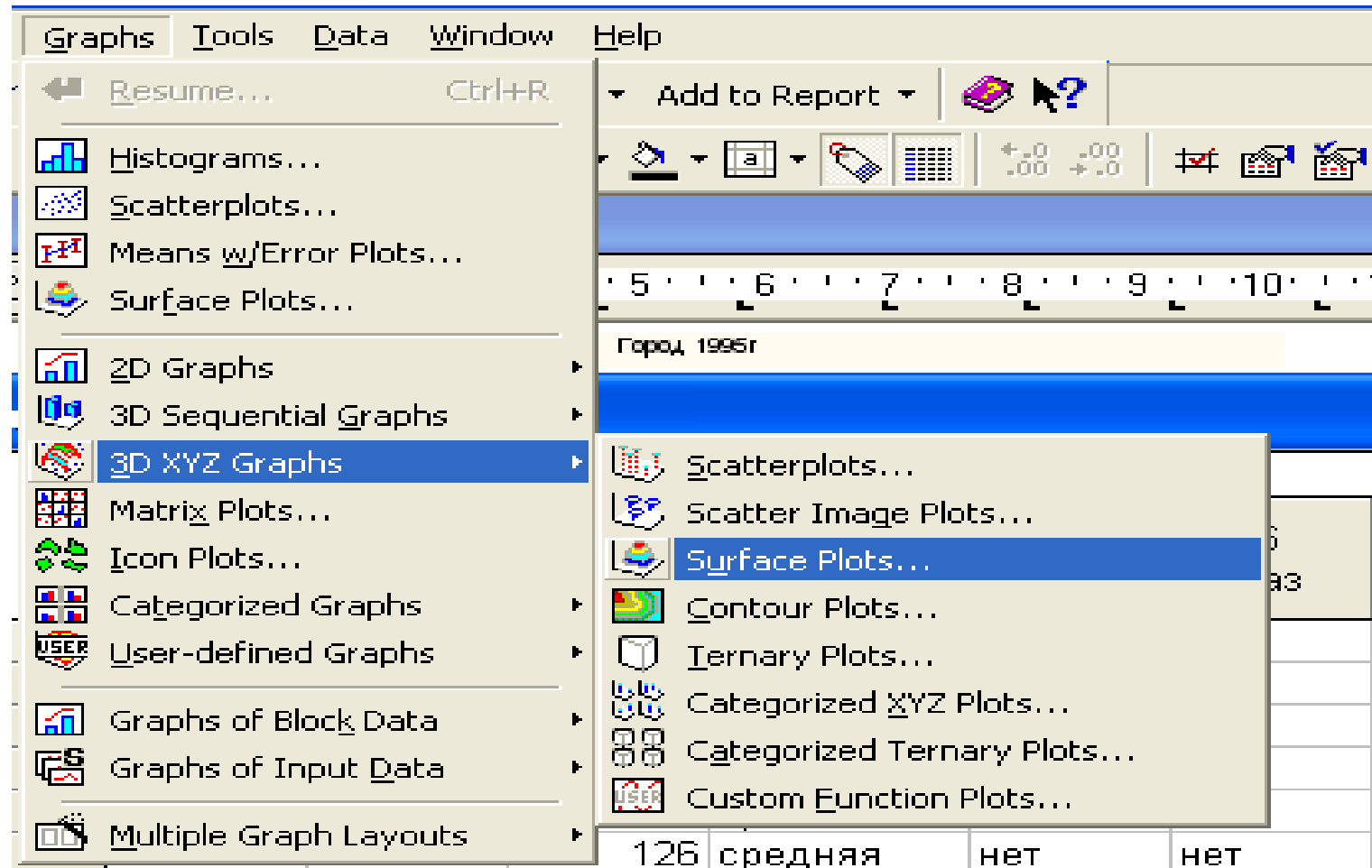
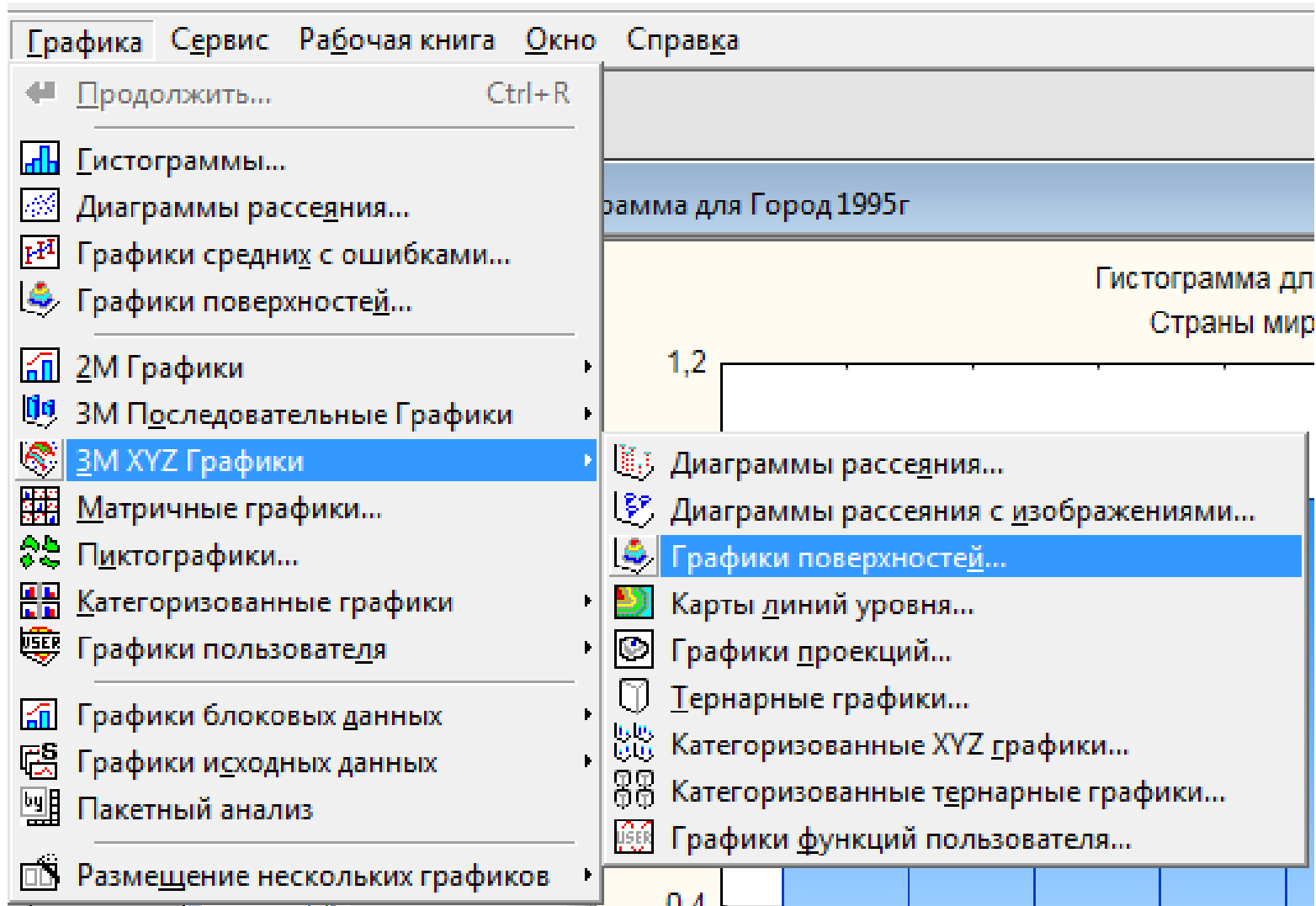


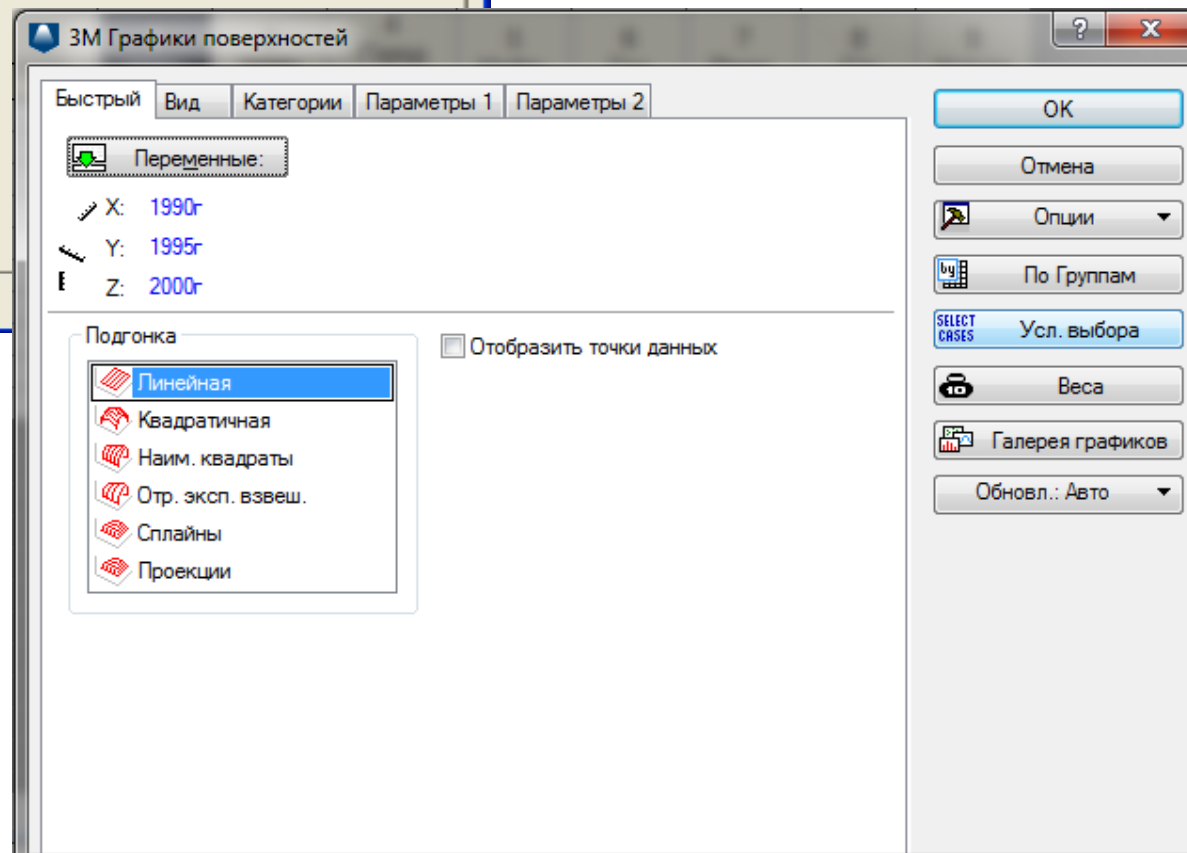
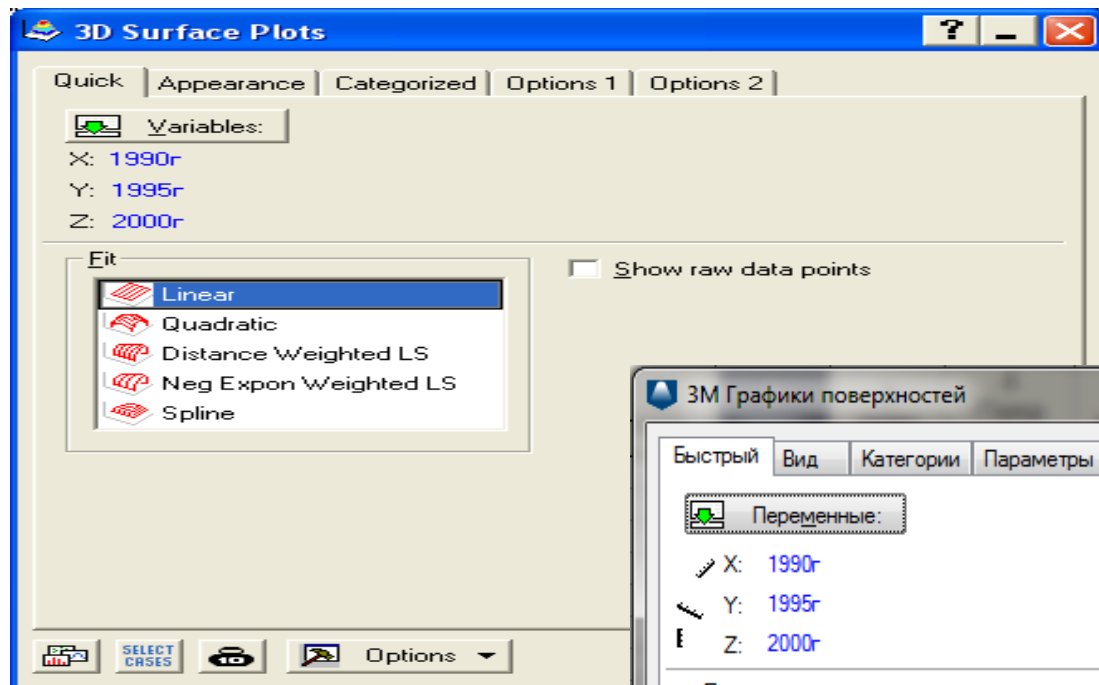
Рис.14

Можно зависимости между непрерывными переменными подгонять различными типами математических функций, геометрически изображаемыми в виде поверхностей. Поверхности можно строить при помощи 3D Surface Plots (графика поверхностей). Построенные поверхности можно вращать различным образом.



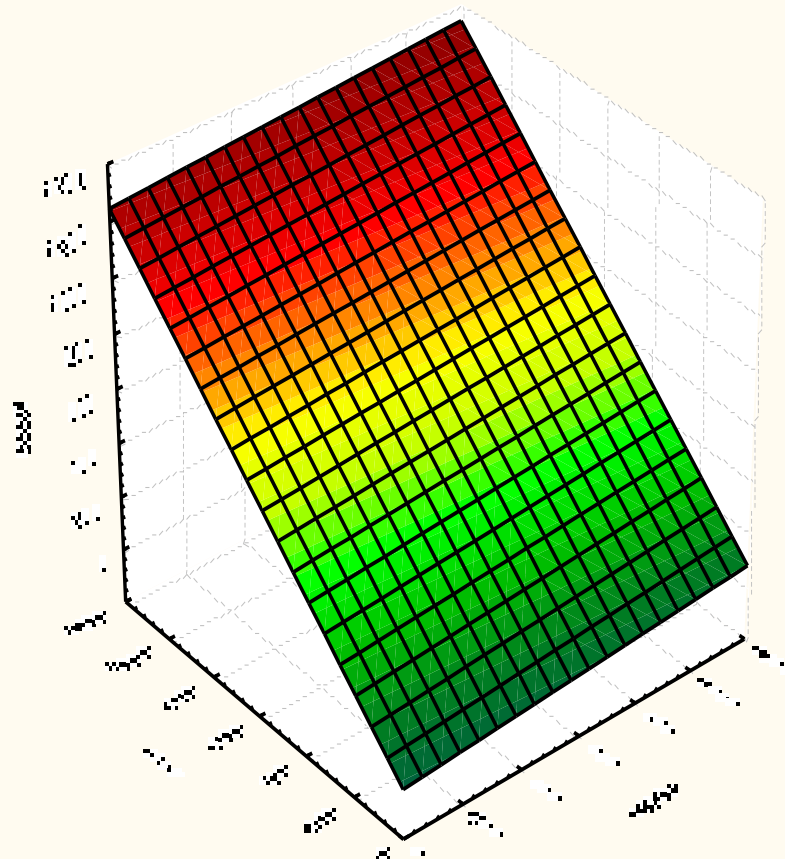


Например, построим график зависимости доли услуг от долей промышленности и сельского хозяйства.



3D Surface Plot (Страны мира.sta 9v*10c)

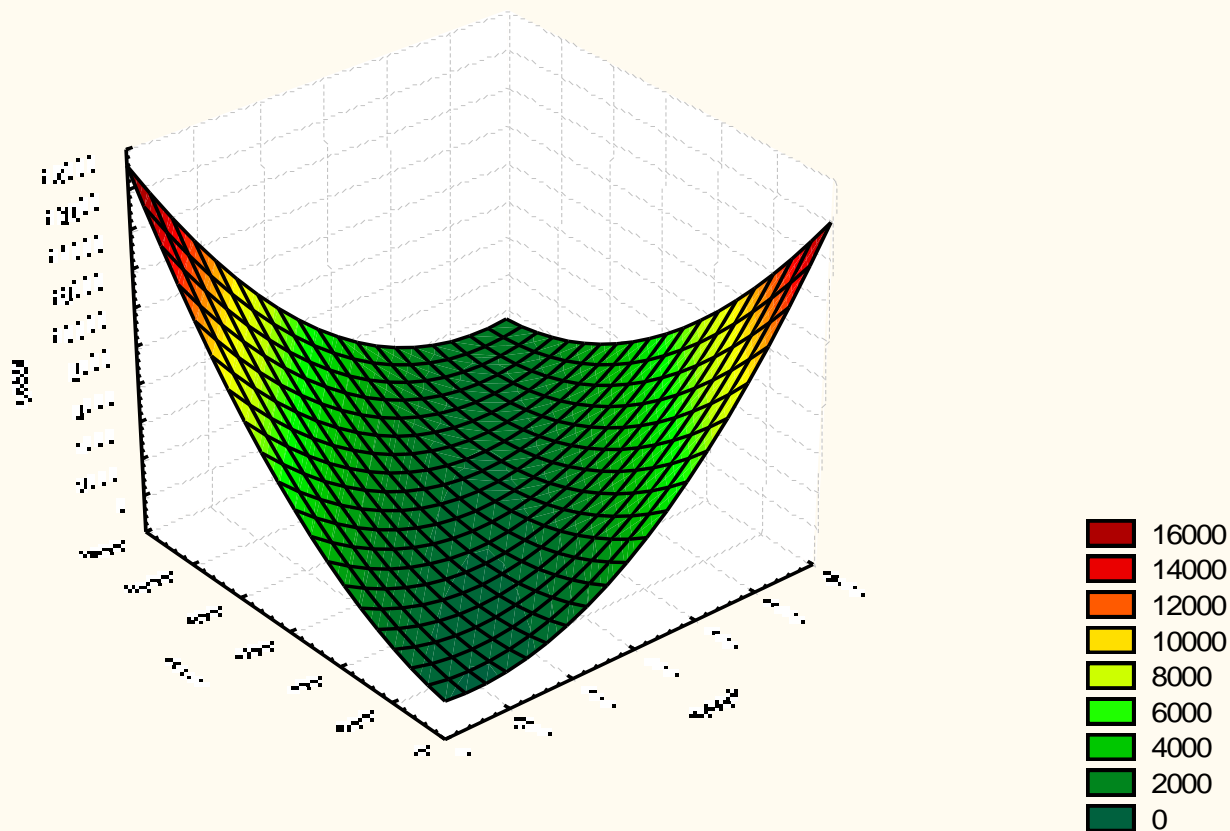
$$2000r = -15,5108 + 0,079 * x + 1,0505 * y$$



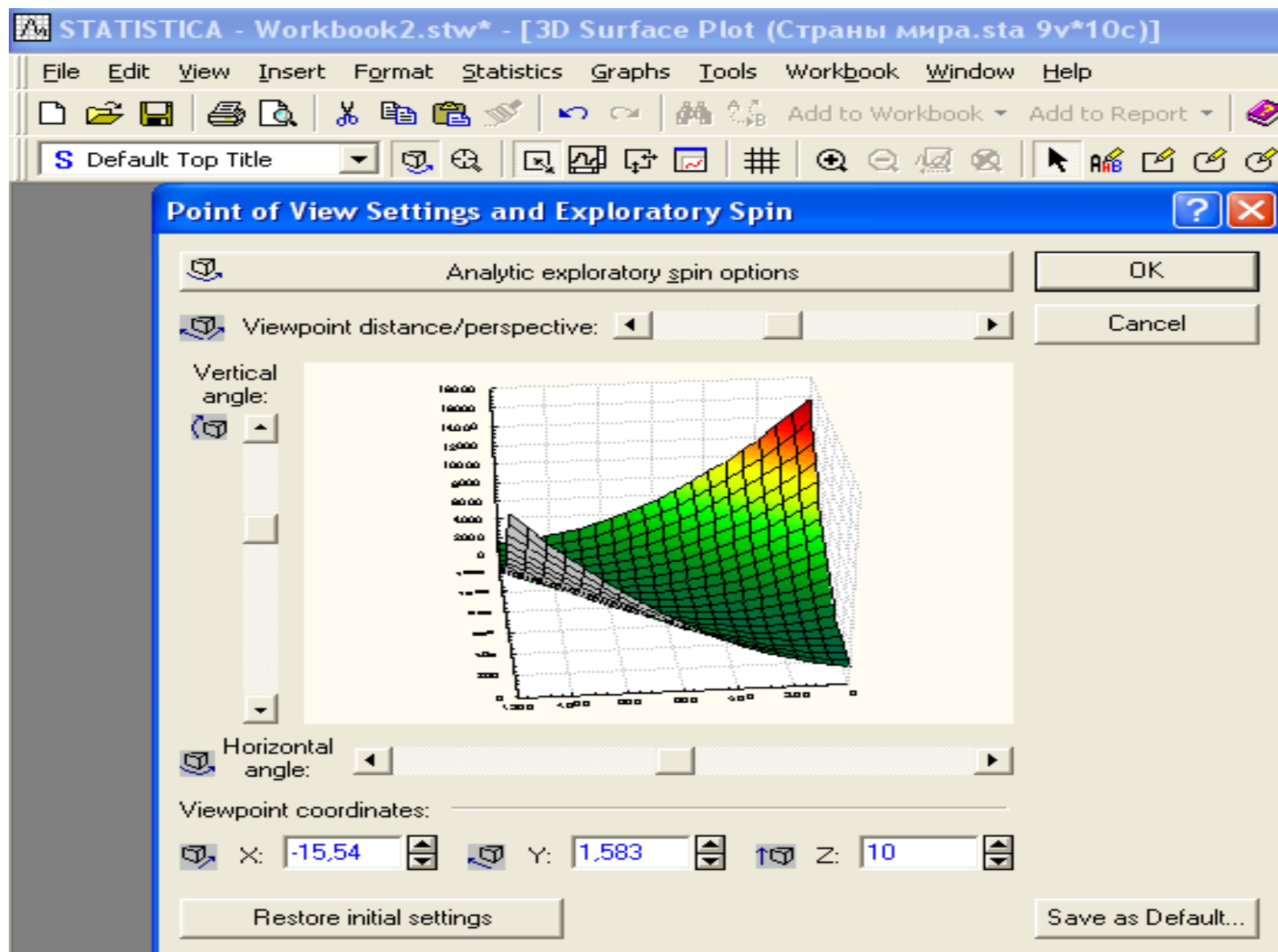
Подгонка квадратичной функцией

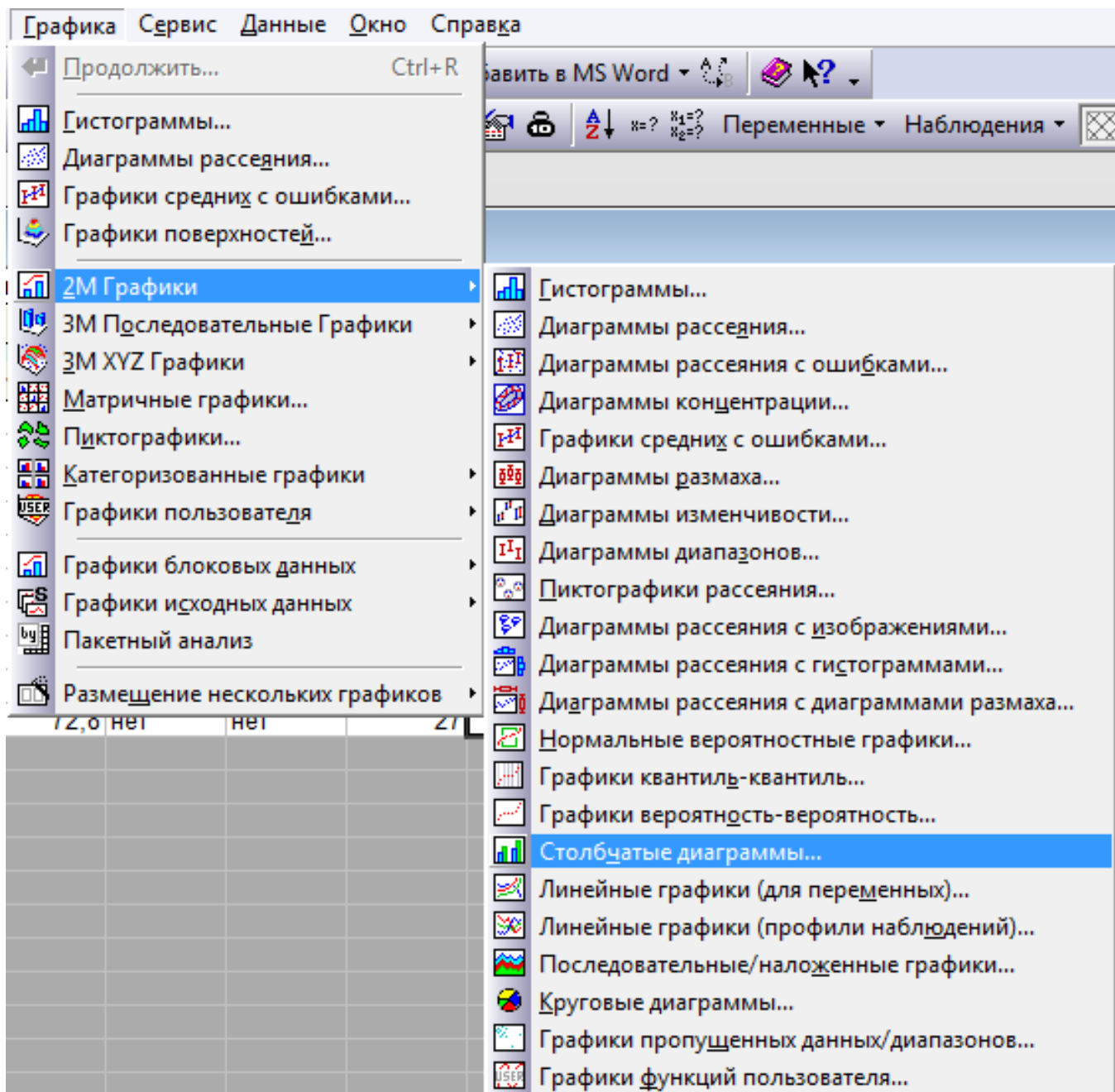
3D Surface Plot (Страны мира.sta 9v*10c)

$$2000г = 19,7474 - 2,3329*x + 3,0657*y + 0,013*x*x - 0,022*x*y + 0,0094*y*y$$



Вращение поверхности







2М Столбчатые диаграммы



Быстрый

Дополнительно

Вид

Параметры 1

Параметры 2

OK

Отмена

Опции

По Группам

Усл. выбора

Веса

Галерея графиков

Обновл.: Авто



Переменные:

1990-2000

Тип графика:



Простой



Составной

Ориентация



Вертикальная



Горизонтальная

Границы контроля



Y1:

0



Y2:

0



X1:

0

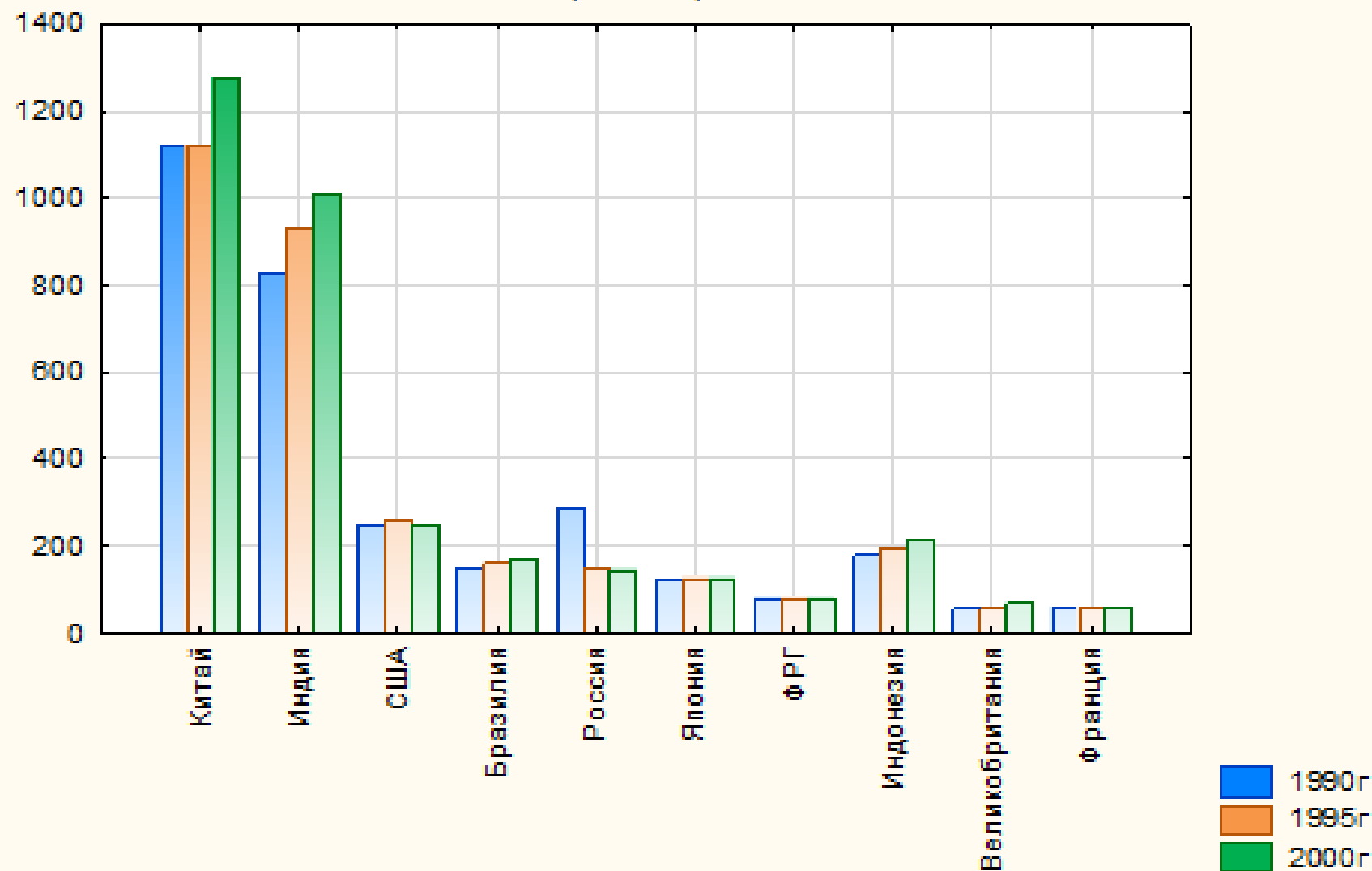


X2:

0



Столбчатая диаграмма для нескольких переменных
Страны мира 1990-2000 гг.



2М Диаграммы размаха

Быстрый | Дополнительно | Вид | Категории | Параметры 1 | Параметры 2

Тип графика:

Прямоуг.-Отрезки | Простой | Переменные:

Зависимая переменная: 1990г-2000г
Группирующая переменная: нет

Группирующие интервалы

☐ Целые числа ☒ Авт

☒ Уникальные значения

☒ Без сорт ☐ Воз ☐ Убыв

☐ Категории: 10

☐ Коды: нет

Макс. без выбросов
75%
Медиана
25%
Мин. без выбросов

Средняя точка

Значение: Медиана

Тип: Точка

☐ Объединенная дисперсия

OK
Отмена
Опции
По Группам
Усл. выбора
Веса
Галерея графиков
Обновл.: Авто

Выберите переменные для диаграммы размаха

1 - 1990г
2 - 1995г
3 - 2000г
4 - Город 1995г
5 - Нефть
6 - Газ
7 - Пром.
8 - С/х
9 - Услуги

1 - 1990г
2 - 1995г
3 - 2000г
4 - Город 1995г
5 - Нефть
6 - Газ
7 - Пром.
8 - С/х
9 - Услуги

OK
Отмена
[Наборы]...

Все | Подробно | Инфо

Зависимая переменная: 1-3

Группирующая переменная:

☐ Подходящие переменные

2M Диаграммы размаха

Быстрый Дополнительно Вид Категории Параметры 1 Параметры 2

Тип графика:

Прямоуг.-Отрезки

Отрезки

Простой

Составной

Переменные:

Зависимая переменная: нет

Группирующая переменная: нет

Группирующие интервалы

Переменная: нет

☐ Целые числа ☒ Авт

☒ Уникальные значения

☒ Без сорт ☐ Воз ☐ Убыв

☐ Категории: 10

☐ Границы: нет

☐ Коды: нет

☐ Сложные подгруппы

Переменные

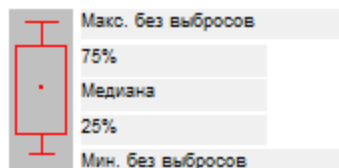
Подгонка

Выкл

☒ Линейная

☒ Полиномиальная

☒ Логарифмическая



Средняя точка

Значение: Медиана

Тип: Среднее
Медиана

☐ Объединенная дисперсия

Несколько зависимостей

☒ Со сдвигом

☐ С перекрытием

Исключить крайние точки: 0 %

Статистики

☐ Критерий Краскела-Уоллиса

☐ F-критерий и p-уровень (ДА)

Прямоугольник

Значение: Проценти

Коэффициент: 25

Отрезок

Значение: Без выбросов

Коэффициент: 1

Выбросы

Выбросы и крайние точки

Коэффициент: 1,5

☒ Соединять средние точки

☐ Отображать исходные данные

Разделить

Выкл. Ширина: 50 %

OK

Отмена

Опции

По Группам

Усл. выбора

Веса

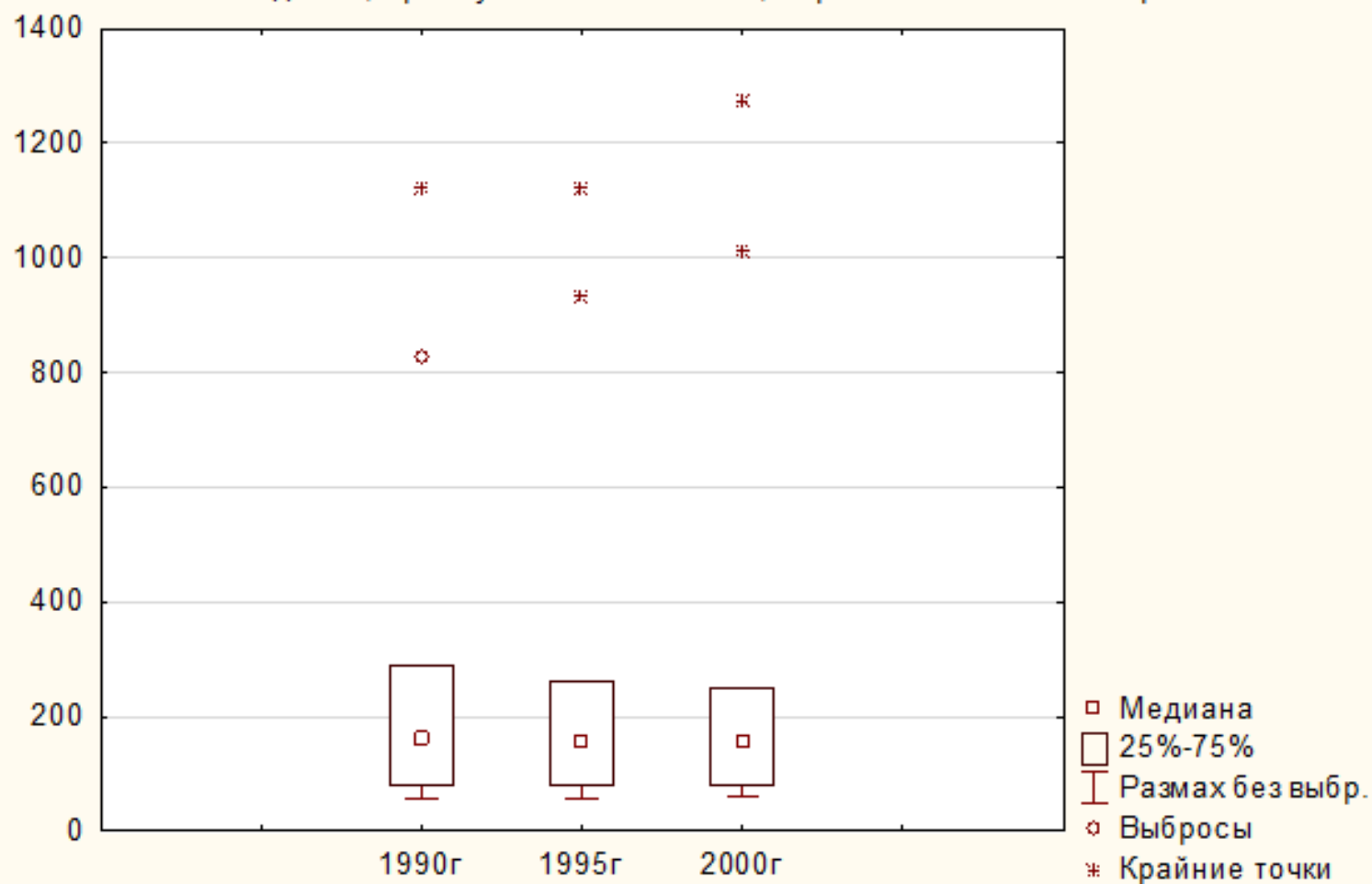
Галерея графиков

Обновл.: Авто

Диаграмма размаха для нескольких переменных

Страны мира 1990-2000 гг.

Медиана; Прямоугольник: 25%-75%; Отрезок: Размах без выбр.



2M Диаграммы размаха

Быстрый Дополнительно Вид Категории Параметры 1 Параметры 2

Тип графика:

Прямоуг.-Отрезки

Отрезки

Простой

Составной

Переменные:

Зависимая переменная: 1990-2000

Группирующая переменная: нет

Группирующие интервалы

Переменная: нет

☐ Целые числа ☒ Авт

☒ Уникальные значения

☒ Без сорт ☐ Воз ☐ Убыв

☐ Категории: 10

☐ Границы: нет

☐ Коды: нет

☐ Сложные подгруппы

Переменные

Подгонка

Выкл

☒ Линейная

☒ Полиномиальная

☒ Логарифмическая

Среднее+2*Ст.откл.

Среднее+Ст. ош.

Среднее

Среднее-Ст. ош.

Среднее-2*Ст.откл.

Средняя точка

Значение: Среднее

Тип: Точка

☐ Объединенная дисперсия

Несколько зависимостей

☒ Со сдвигом

☐ С перекрытием

Исключить крайние точки: 0 %

Статистики

☐ Критерий Краскела-Уоллиса

☐ F-критерий и p-уровень (ДА)

Прямоугольник

Значение: Ст. ош.

Коэффициент:

Отрезок

Значение:

Коэффициент: 2

Выбросы

Выбросы и крайние точки

Коэффициент: 1,5

☐ Соединять средние точки

☐ Отображать исходные данные

Разделить

Выкл.

Ширина: 50 %

OK

Отмена

Опции

По Группам

Усл. выбора

Веса

Галерея графиков

Обновл.: Авто

Диаграмма размаха для нескольких переменных
Страны мира1 9v*10с

Среднее; Прямоугольник: Среднее \pm Ст.ош; Отрезок: Среднее \pm 2*Ст.откл.

