

Методы организации, планирования и обработки результатов инженерного эксперимента

Всего 144 ч., ауд. – 18 ч., самостоятельная работа - 126 ч.
Зачёт

Стукач Олег Владимирович

Каф. СУМ, пр. Ленина, д. 2, оф. 204

 (3822)-701777*2754

tomsk@ieee.org

<http://ieee.tpu.ru/pages/stukach.htm>

<http://portal.main.tpu.ru:7777/SHARED/o/OS>



Цель и задачи

- ☺ Дать обзор статистических методов, реализованных в программных комплексах «STATISTICA», «Labview», относящихся к инженерному эксперименту
 - ☺ Сформировать у слушателей понимание методов разведочного анализа данных, необходимых им для дальнейшей работы над своими задачами
 - ☺ Получить навыки анализа данных
 - ☺ Решить полностью или частично подзадачу из своей квалификационной работы
-

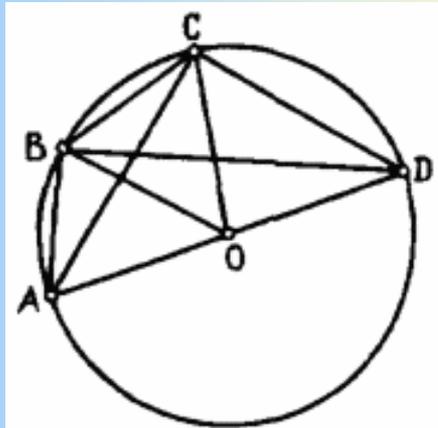
Четвёртая парадигма



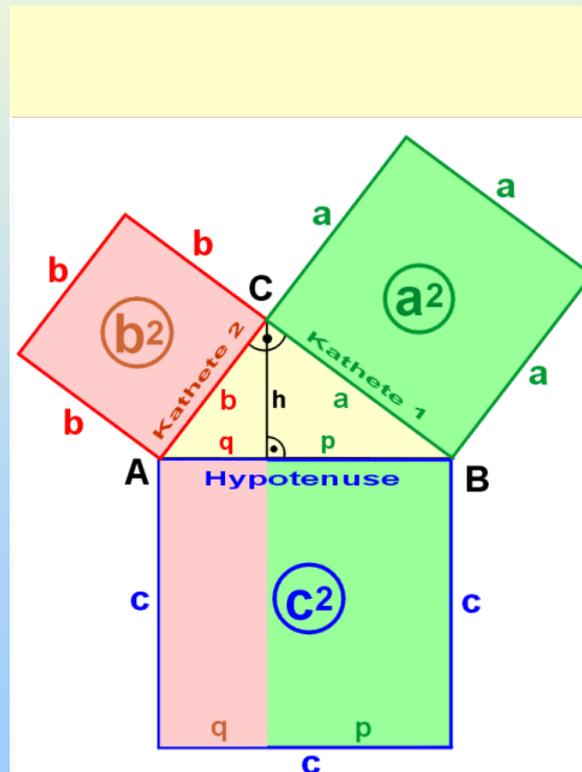
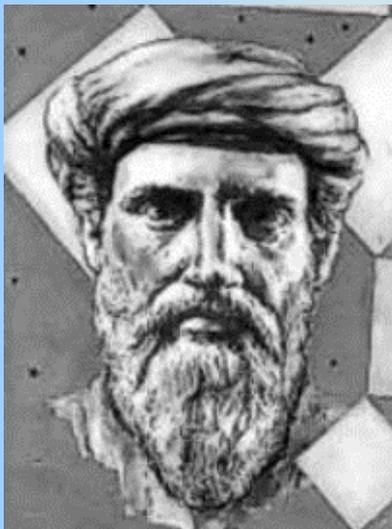
Парадигма – концептуальная схема, используемая для постановки проблем и их решения, господствующая в научном сообществе в течение длительного исторического периода

**Энциклопедический
словарь**

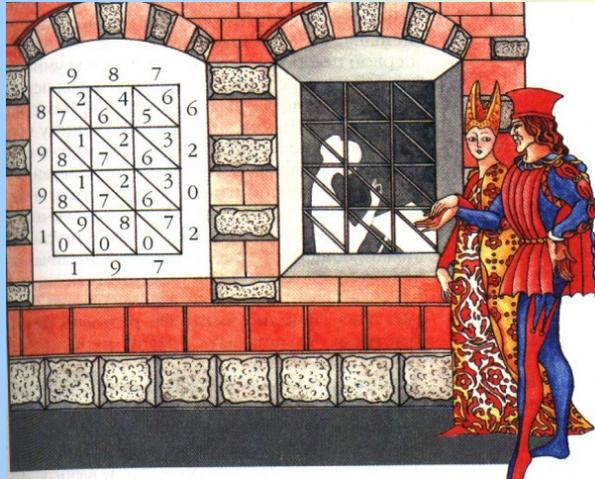
Познать - значит нарисовать



V-III вв до н.э.



Познать - значит посчитать



X-XV н.э.

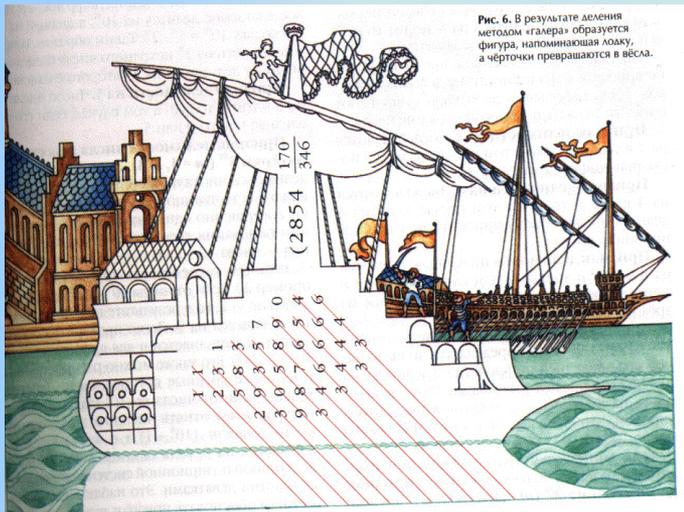
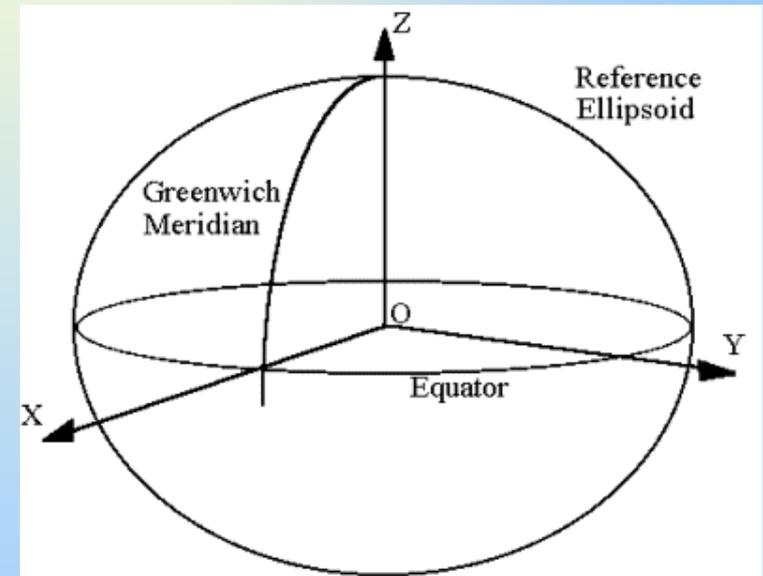
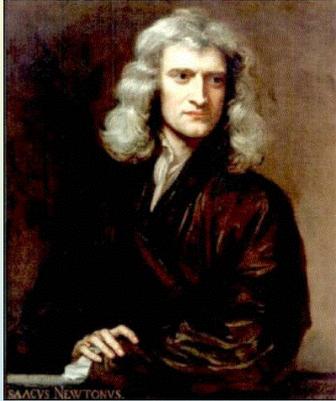


Рис. 6. В результате деления методом «галера» образуется фигура, напоминающая лодку, а черточки превращаются в весла.

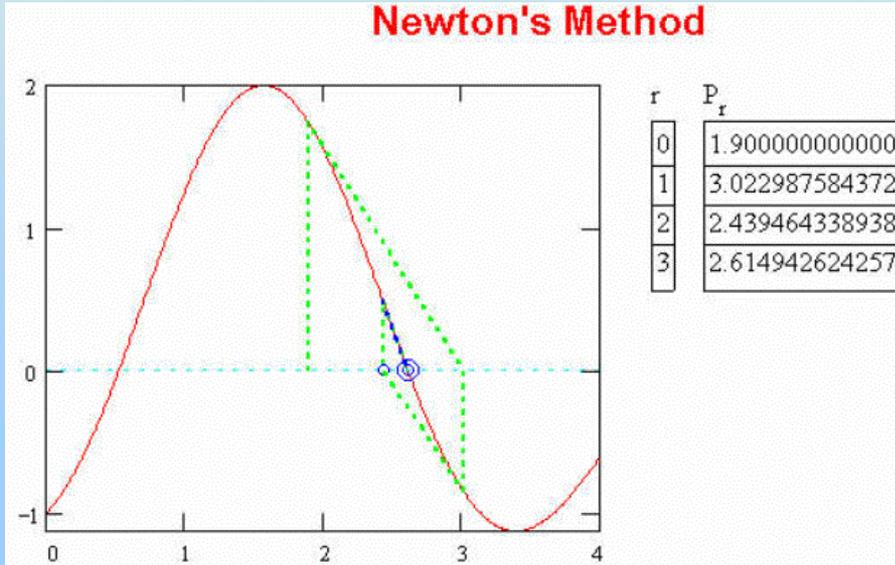


Познать - значит составить и решить дифференциальное уравнение

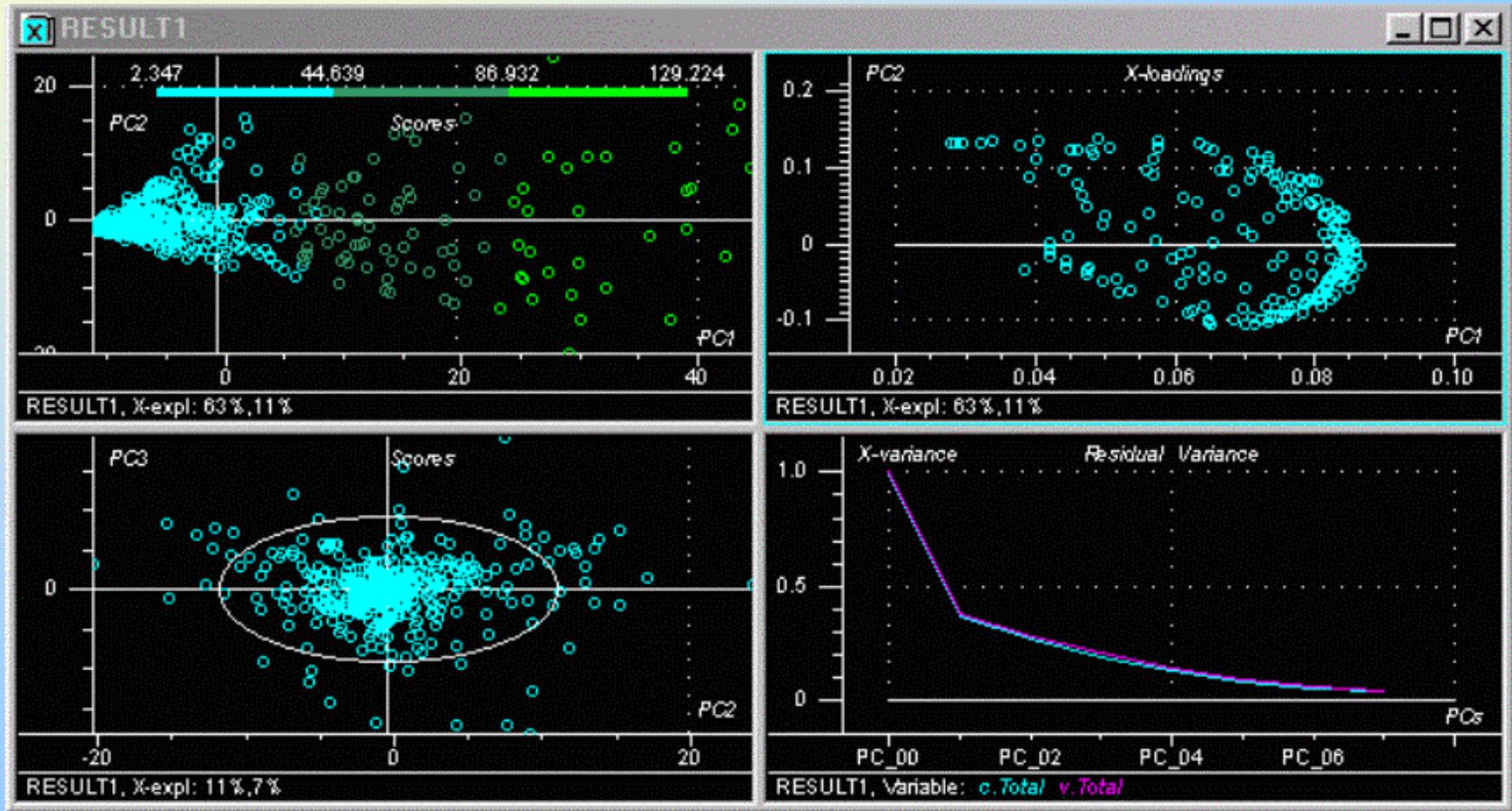
XVII в. – н.в.



Newton's Method



Познать - значит собрать и проанализировать данные



Пример со взвешиванием трёх объектов

ТРАДИЦИОННАЯ СХЕМА ВЗВЕШИВАНИЯ ТРЕХ ОБЪЕКТОВ¹

| Номер опыта | <i>a</i> | <i>b</i> | <i>c</i> | Результат взвешивания |
|-------------|----------|----------|----------|-----------------------|
| 1 | -1 | -1 | -1 | y_0 |
| 2 | +1 | -1 | -1 | y_1 |
| 3 | -1 | +1 | -1 | y_2 |
| 4 | -1 | -1 | +1 | y_3 |

¹ Здесь и в табл. 2 обозначение «+1» указывает, что объект взвешивания положен на весы; обозначение «-1» указывает на отсутствие объекта на весах.

$$A = y_1 - y_0.$$

$$\sigma^2 \{A\} = \sigma^2 \{y_1 - y_0\} = 2\sigma^2 \{y\}$$

Пример со взвешиванием трёх объектов

$$A = y_1 - y_0.$$

$$\sigma^2 \{A\} = \sigma^2 \{y_1 - y_0\} = 2 \sigma^2 \{y\}$$

либо вообще не работал, либо давал слишком грубую оценку. В связи с этим он на практике почти не используется.

§ 2.11. Учет влияния корреляционных зависимостей

Таким образом, мы пока не получили ничего хорошего. В сущности, мы пришли к давно применяемому в обыденной жизни способу, при котором в качестве оценки используется результат только одного, наиболее точного измерения! Заметим, что дело обстоит столь тривиально лишь в случае прямых измерений. При обработке данных косвенных измерений подход, основан-

Пример со взвешиванием трёх объектов

ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА
ПРИ ВЗВЕШИВАНИИ ТРЕХ ОБЪЕКТОВ

| Номер опыта | a | b | c | Результат взвешивания |
|-------------|-----|-----|-----|-----------------------|
| 1 | +1 | -1 | -1 | y_1 |
| 2 | -1 | +1 | -1 | y_2 |
| 3 | -1 | -1 | +1 | y_3 |
| 4 | +1 | +1 | +1 | y_4 |

Пример со взвешиванием трёх объектов

$$A = \frac{y_1 - y_2 - y_3 + y_4}{2},$$

$$B = \frac{-y_1 + y_2 - y_3 + y_4}{2},$$

$$C = \frac{-y_1 - y_2 + y_3 + y_4}{2}.$$

$$\sigma^2 \{A\} = \sigma^2 \left\{ \frac{y_1 - y_2 - y_3 + y_4}{2} \right\} = \frac{4 \sigma^2 \{y\}}{4} = \sigma^2 \{y\}$$

$$\sigma^2 \{B\} = \sigma^2 \{y\} \quad \text{и} \quad \sigma^2 \{C\} = \sigma^2 \{y\}$$

Пример показывает, что даже в простых задачах можно с удивительной отчётливостью противопоставить плохой эксперимент хорошему

Пример «малого знания»

← ⓘ | <https://www.tomsk.kp.ru/daily/26670.5/3692229/>

ГВО ЭКОНОМИКА В МИРЕ ЗВЕЗДЫ НАШИ ПРА

лучше, и губернии лучше. А так выходит, область в руи
тюрьме, и кому хорошо?

- Зато движуха.

- У нас движуха подменяет реальность. «Будет так, как
помнили, где правда, а где для народа дурь придумали. Потом сами запутались.
И живут теперь как в чаду, в чаду собственных иллюзий.

ТРУД И ГЛАД

- Вы несколько раз сказали, что экономика в депрессии, но народ этого как бы и не замечает. Я все-таки не понимаю, в чем тут загадка. Если у тебя все плохо, ты кричишь! Значит, не так уж и плохо?

- Я думаю, что кричащий факт, говорящий о серьезных проблемах нашей жизни
– это зарплата работающих людей. Она скандально низкая. Есть такой
показатель: модульная зарплата. Мало кто знает, что это такое. Все говорят:
средняя. А средняя – это когда менеджера, у которого три миллиона рублей в
месяц, и народ, у которого 15 тысяч в месяц, а в среднем все просто прекрасно.
Модульная зарплата – это самая распространенная зарплата. И она у нас 13,5
тысяч по России. Это наш специфический тип бедности. Во всем мире беден тот,
кто без работы. Бездомный в Оклахоме, бомж в Нигерии бедны, но, едва
получают работу, тут же расправляют плечи. У нас ты вкалываешь, и все равно
едва сводишь концы с концами.

Получать уведомления о самом
важном от «Комсомольской
правды»

Пример «малого знания»

www.pressreader.com/russia/komsomolskaya-pravda/20170530/281719794538205

- Вы говорили, что экономика в депрессии, но народ этого как бы и не замечает. Значит, не так уж и плохо?

- Кричащий факт - зарплата работающих. Она скандально низкая. Есть такой показатель: модульная зарплата. Мало кто знает, что это такое. Все говорят: средняя. А средняя - это когда есть менеджер, у которого три миллиона рублей в месяц, и есть народ, у которого 15 тысяч в месяц; в среднем все прекрасно. Модульная зарплата - это самая распространенная. И она 13,5 тысячи по России. Это наш специфический тип бедности. Во всем мире беден тот, кто без работы. Бездомный в Оклахоме, бомж в Нигерии бедны, но, едва получают работу, тут же расправляют плечи. У нас ты



вка- лываешь и все равно едва сводишь концы с концами. - А народ это спокойно сносит... - У нас, увы, нет культуры решать проблемы в спокойном режиме. Нашему вселенскому уму

Мода – часто встречающееся значение случайной величины

**Интеллектуальный анализ
данных. Знакомство с
ВОЗМОЖНОСТЯМИ СИСТЕМЫ
STATISTICA по решению задач
статистического анализа**

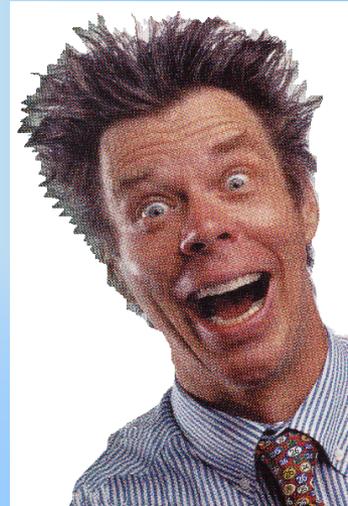
Стукач Олег Владимирович

Каф. КИСМ, пр. Ленина, д. 2, оф. 204

☎ (3822)-701777*2754

tomsk@ieee.org

<http://ieee.tpu.ru/pages/stukach.htm>



Что такое данные?

Данные - это информация о фактах или событиях, на основе анализа которых принимаются необходимые решения

Виды данных:

Численные – цифровые данные (таблицы, отчеты, сводки и т.д.)

Словесные – информация, поступающая от потребителей, сотрудников и др. заинтересованных сторон, не несущая численных значений (много, мало, хорошо, плохо, удобно, не удобно и т.д.)

Схема принятия решения

Выбор типа
требуемых
данных



Сбор данных



Обработка и
анализ
данных
Разведочный анализ
данных



Принятие
решения

Этапы стат. исследования

Сбор данных

Разведочный анализ

визуальный анализ

описательный анализ

Проверка на принадлежность
распределению

Преобразование переменных

Построение модели

Уточнение модели

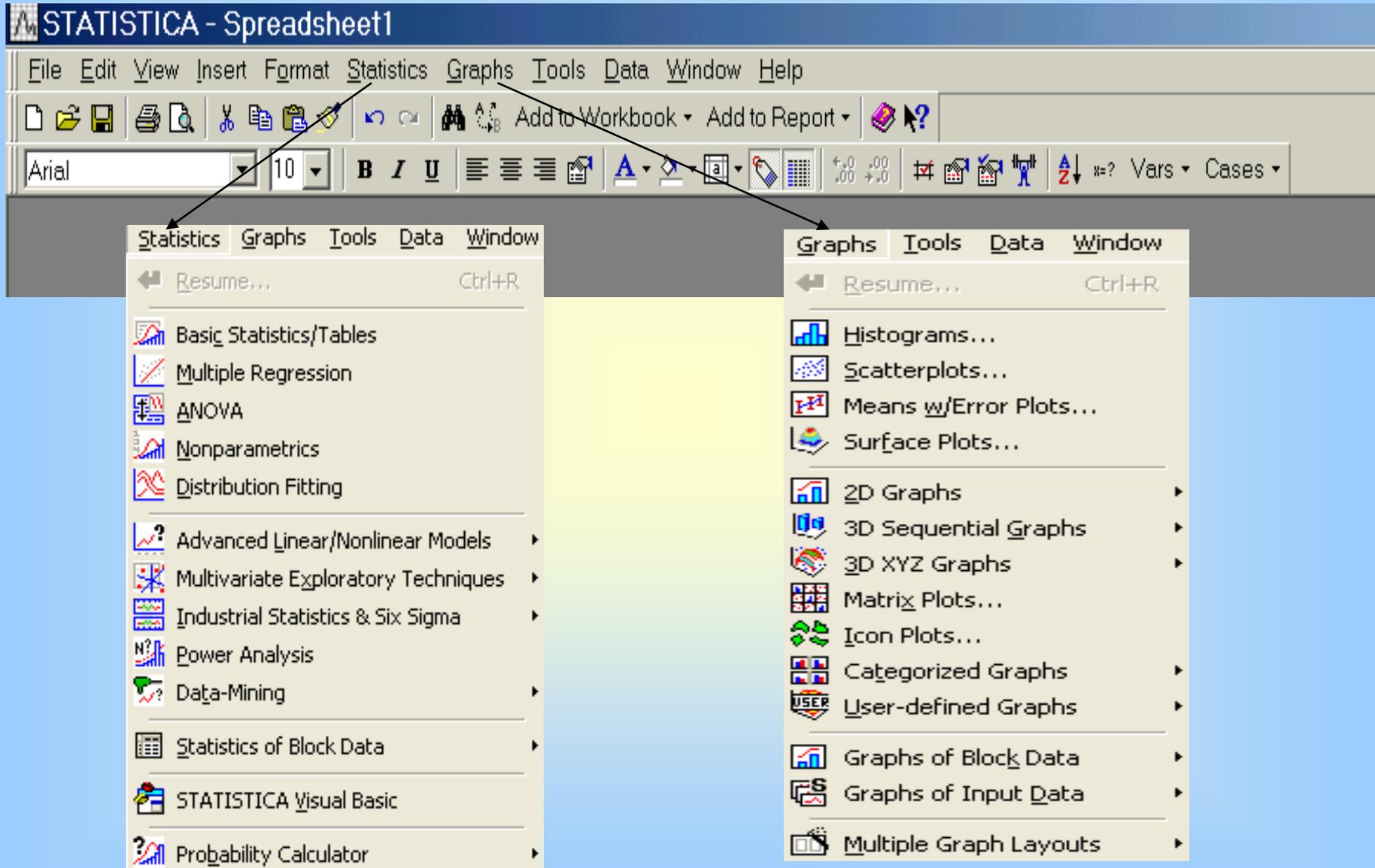
Общие сведения о программе STATISTICA



Statistica – это современный пакет статистического анализа, в котором реализованы все новейшие компьютерные и математические методы анализа данных.

Использование системы STATISTICA не требует наличия специальной математической подготовки и позволяет полностью сосредоточиться на данных и моделях, не вникая в технические детали того или иного статистического метода.

Главное меню пакета STATISTICA



Рабочее окно программы STATISTICA

The screenshot displays the STATISTICA software interface. The main window is titled "STATISTICA - Spreadsheet1" and contains a menu bar (File, Edit, View, Insert, Format, Statistics, Graphs, Tools, Data, Window, Help) and a toolbar with various icons. Below the toolbar is a text area showing "Arial" and "10". The main data area is a spreadsheet titled "Data: Spreadsheet1 (10v by 10c)". The spreadsheet has 10 columns labeled "Var1" through "Var10" and 10 rows labeled "1" through "10". The cell at the intersection of row 1 and column 1 is selected, indicated by a black border. Several callout boxes with arrows point to specific parts of the interface:

- "Трафарет" (Template) points to the toolbar.
- "Имена переменных (Variables)" (Variable names) points to the column headers (Var1-Var10).
- "Активная ячейка" (Active cell) points to the selected cell in row 1, column 1.
- "Имя файла набора данных и его размерность (10 переменных на 10 наблюдений)" (Data set file name and its dimensionality (10 variables on 10 observations)) points to the spreadsheet title bar.
- "Номера наблюдений (Cases)" (Observation numbers) points to the row numbers (1-10).

Переменная (английский термин *variable*) — это то, что можно измерять или контролировать. Иными словами, переменная — это то, что варьируется, изменяется, а не является постоянным (от английского корня *var*).

Рабочее окно программы STATISTICA

Изменение параметров
переменных

Задать имя
переменной

Задать код
пропущенного
значения

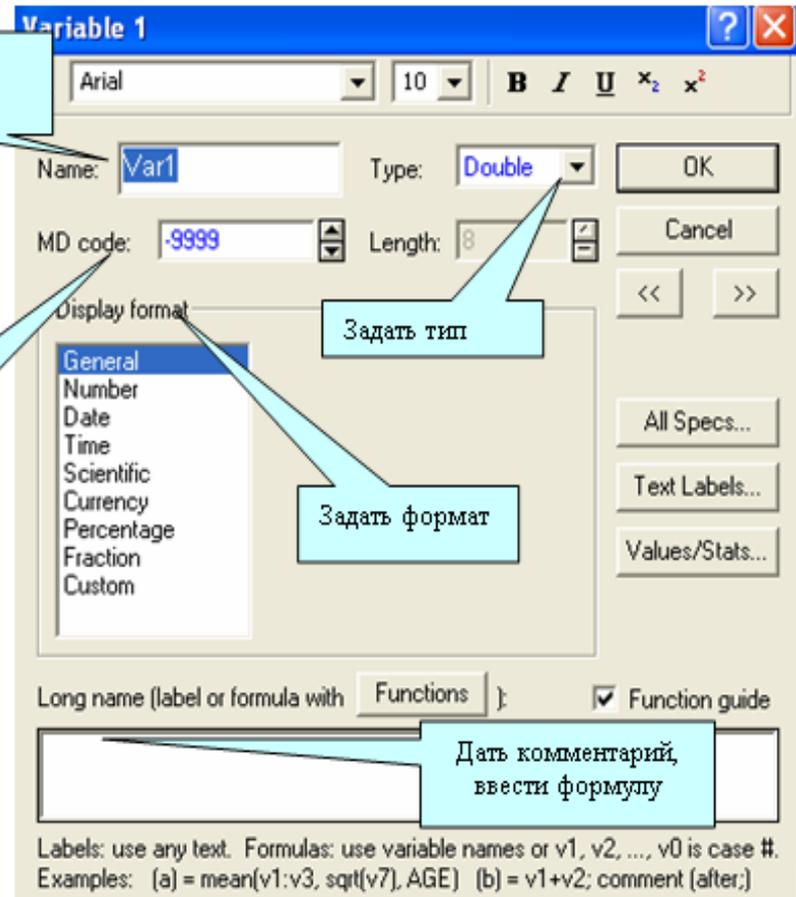
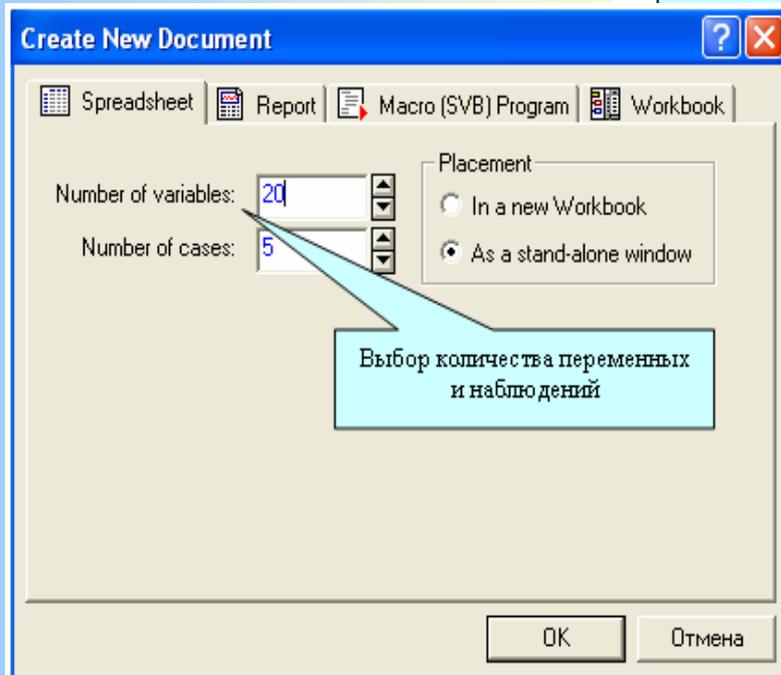
Задать тип

Задать формат

Дать комментарий,
ввести формулу

Выбор количества переменных
и наблюдений

Создание нового
документа (*File/New*)



Данные

| Дата | Установка | Оборудование | Дефект | Результат |
|------------|-----------|--------------|--------------|----------------|
| 02.05.2007 | ТВА160 | ЧПТВА | Остановка | Неустранён |
| 03.05.2007 | ДС158 | 1015 | Погрешность | Откалиброван |
| 06.05.2007 | ТВА160 | ЧПТВА | Остановка | Неустранён |
| 09.05.2007 | ДС158 | ПК | Сбой | Устранён |
| 10.05.2007 | SPECO | Фильтр | Поврежд.цепи | Отремонтирован |
| 21.05.2007 | ДС158 | Горелка | Бурс | Отремонтирован |
| 25.05.2007 | МАП | ЧПМАП | Остановка | Отремонтирован |
| 14.06.2007 | ТВА160 | Термо | Износ | Замена |
| 17.06.2007 | ДС1581 | 510 | Поврежд.цепи | Устранён |
| 19.06.2007 | МАП | Фильтр | Пурф | Отремонтирован |
| 22.06.2007 | ТВА160 | ЧПТВА | Остановка | Неустранён |
| 23.06.2007 | МАП | Фильтр | Пурф | Отремонтирован |

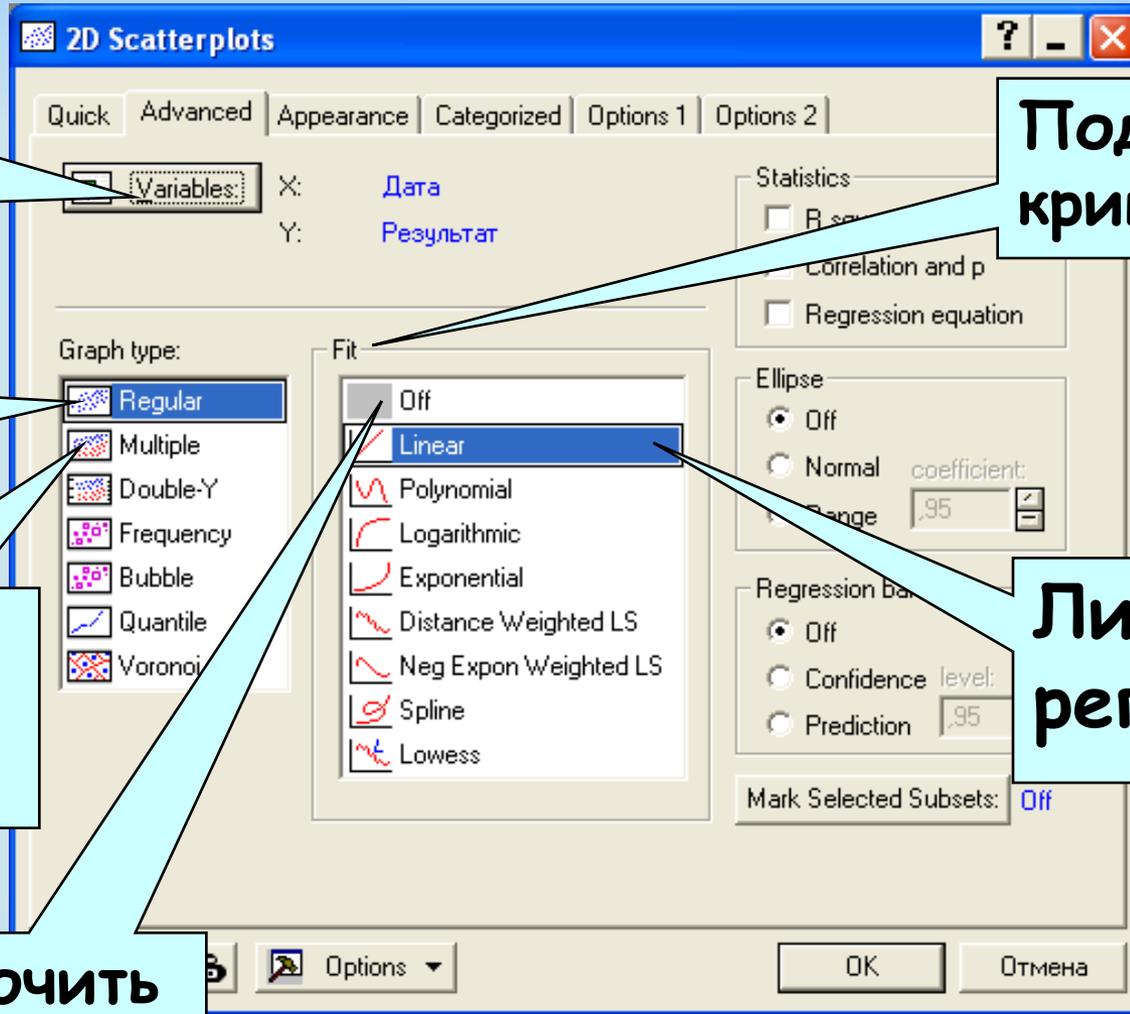
Окно построения

Выбор аргумента и функции

Один график

Много графиков на одной сетке

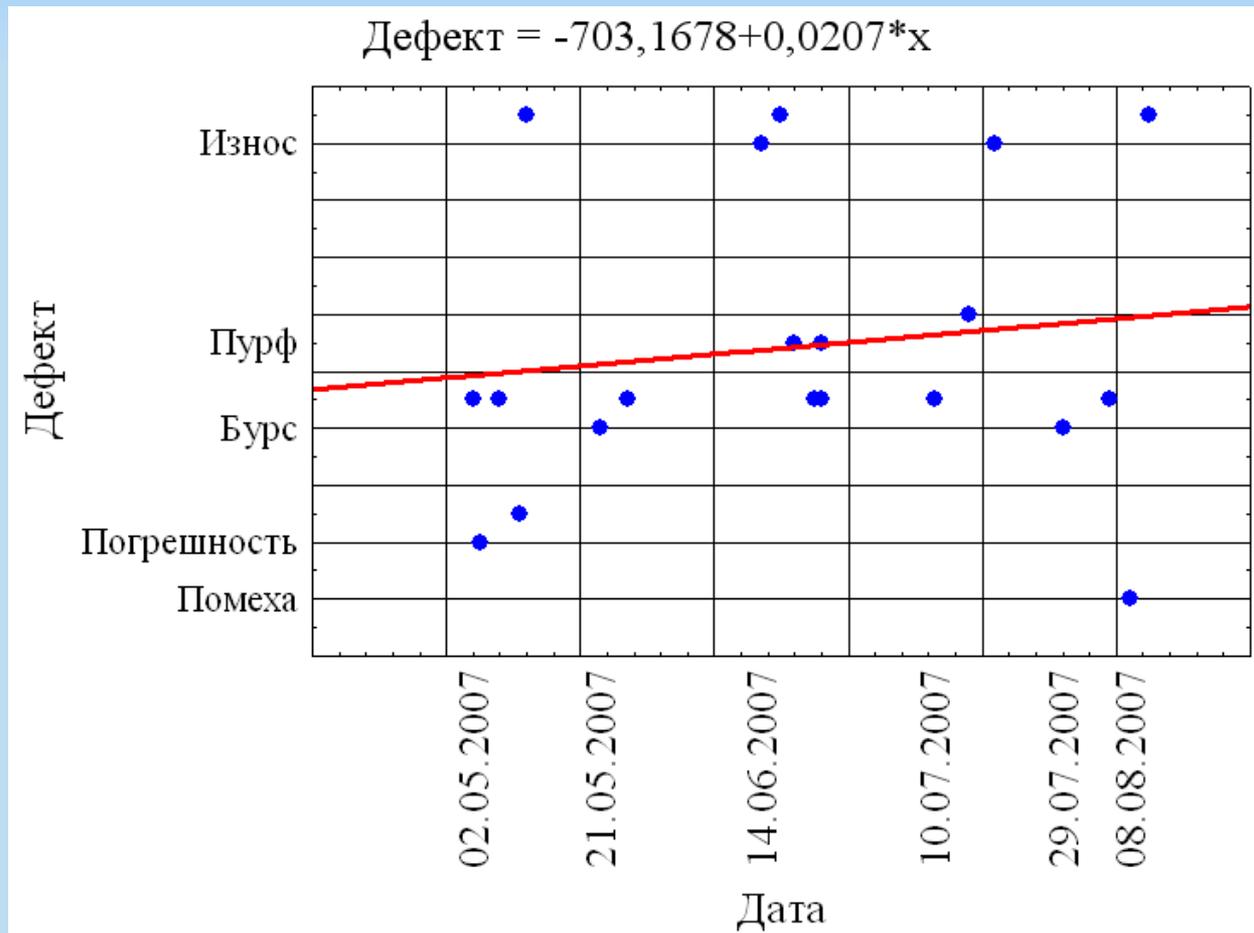
Выключить



Подгоночная кривая

Линейная регрессия

Диаграмма рассеяния



Прямая на диаграмме рассеяния – график простой линейной регрессии $y = -703,1678 + 0,0207x$

Круговая диаграмма

STATISTICA - Workbook12* - [Pie Chart (variant1 30v*50c)]

File Edit View Insert Format Statistics Graphs Tools Workbook Window Help

Normal Graph [modi...]

Pie Chart (variant1 30v*50c)

Pie Chart (variant1 30v*50c)

Var7

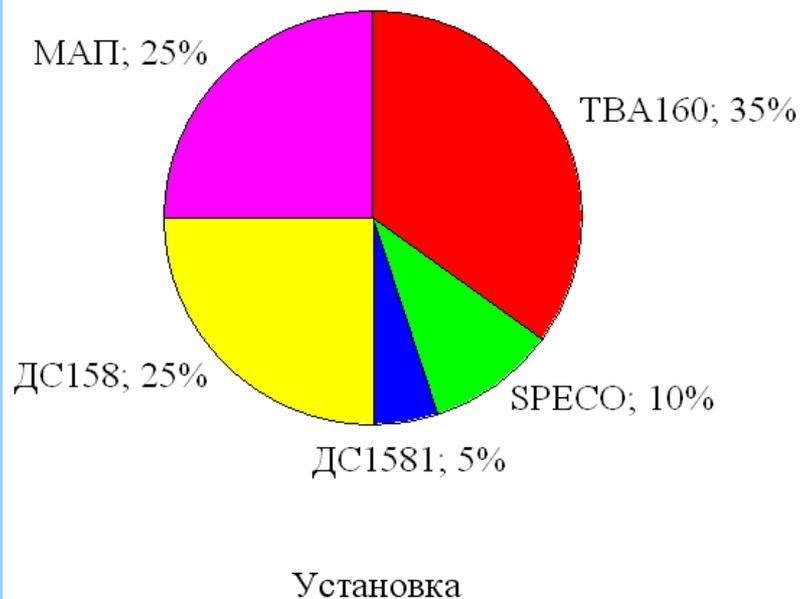
Pie Chart (variant1 30v*50c)

Resume... Ctrl+R

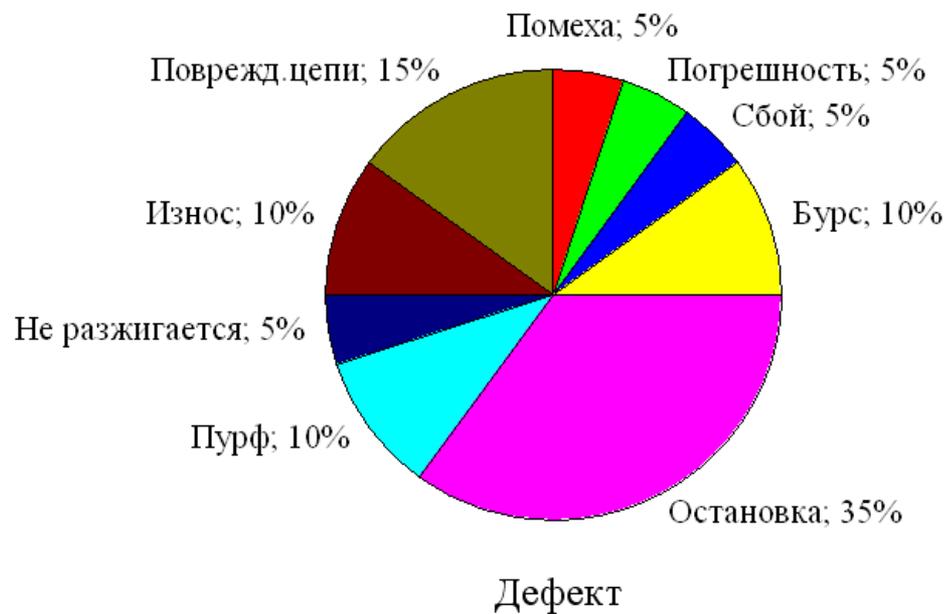
- Histograms...
- Scatterplots...
- Means w/Error Plots...
- Surface Plots...
- 2D Graphs**
 - 3D Sequential Graphs
 - 3D XYZ Graphs
 - Matrix Plots...
 - Icon Plots...
 - Categorized Graphs
 - User-defined Graphs
 - Graphs of Block Data
 - Graphs of Input Data
 - Multiple Graph Layouts
- Histograms...
- Scatterplots...
- Means w/Error Plots...
- Box Plots...
- Range Plots...
- Scatter Icon Plots...
- Scatter Image Plots...
- Scatterplots w/Histograms...
- Scatterplots w/Box Plots...
- Normal Probability Plots...
- Quantile-Quantile Plots...
- Probability-Probability Plots...
- Bar/Column Plots...
- Line Plots (Variables)...
- Line Plots (Case Profiles)...
- Sequential/Stacked...
- Pie Charts...**
- Missing/Range Data Plots...
- Custom Function Plots...

Круговая диаграмма

Pie Chart (g_2.sta 5v*20c)



Pie Chart (g_2.sta 5v*20c)



Литература

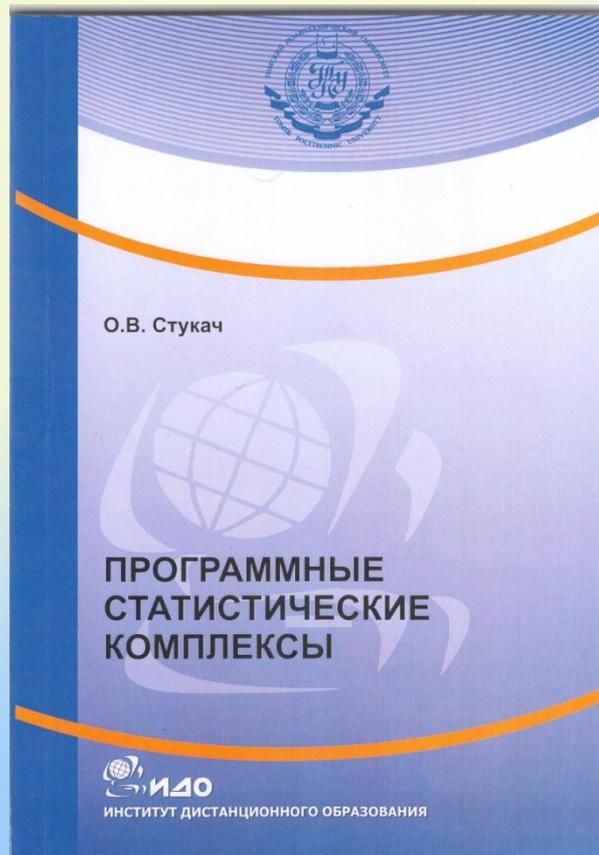
- ✚ Боровиков В.П. Популярное введение в программу STATISTICA. - М : Компьютер пресс, 1998. - 267 с.
- ✚ Боровиков В.П. Программа Statistica для студентов и инженеров. - М. : Компьютер пресс, 2000. - 301 с.
- ✚ Боровиков В.П. Statistica: искусство анализа данных на компьютере. - СПб. : Питер, 2001. - 656 с.
- ✚ Халафян А.А. Statistica 6. Статистический анализ данных. - Бином-Пресс, 2007. - 512 с.

Литература

- ✚ Вуколов Э.А. Основы статистического анализа. Практикум по статистическим методам и исследованию операций с использованием пакетов STATISTICA и EXCEL: Учебное пособие. - М.: ФОРУМ: ИНФРА-М, 2004. - 464 с. - (Профессиональное образование).
- ✚ Берестнева О.Г. Компьютерный анализ данных : учебное пособие / О.Г. Берестнева, Е.А. Муратова, А.М. Уразаев. - Томск : Изд-во ТПУ, 2003. - 204 с.
- ✚ Электронный учебник по статистике [Электронный ресурс]. - М. : StatSoft, Inc. - 2001. - Режим доступа: <http://www.statsoft.ru/home/textbook/default.htm>.

Литература

Программные статистические комплексы: учебное пособие / сост. О.В. Стукач; Томский политехнический университет.- Томск: Изд-во Томского политехнического университета, 2014.- 122 с.



Литература

Стукач, Олег Владимирович. Программный комплекс Statistica в решении задач управления качеством : учебное пособие для вузов / О. В. Стукач ; Национальный исследовательский Томский политехнический университет (ТПУ). - Томск : Изд-во ТПУ, 2011. - 163 с. : ил. - Библиогр.: с. 158-159.

Стукач, Олег Владимирович. Программный комплекс Statistica в решении задач управления качеством [Электронный ресурс] : учебное пособие / О. В. Стукач ; Национальный исследовательский Томский политехнический университет (ТПУ). - 1 компьютерный файл (pdf; 2.4 МВ). - Томск : Изд-во ТПУ, 2011. - Заглавие с титульного экрана. - Доступ из корпоративной сети ТПУ. - Adobe Reader. - <URL:<http://www.lib.tpu.ru/fulltext2/m/2011/m426.pdf>>.

Литература

Стукач, Олег Владимирович. Программный комплекс Statistica в решении задач управления качеством [Электронный ресурс] : учебное пособие / О. В. Стукач ; Национальный исследовательский Томский политехнический университет (ТПУ). - 1 компьютерный файл (pdf; 2.4 МВ). - Томск : Изд-во ТПУ, 2011. - Заглавие с титульного экрана. - Доступ из корпоративной сети ТПУ. - Adobe Reader. - <http://www.lib.tpu.ru/fulltext2/m/2011/m426.pdf>.



Графический анализ данных

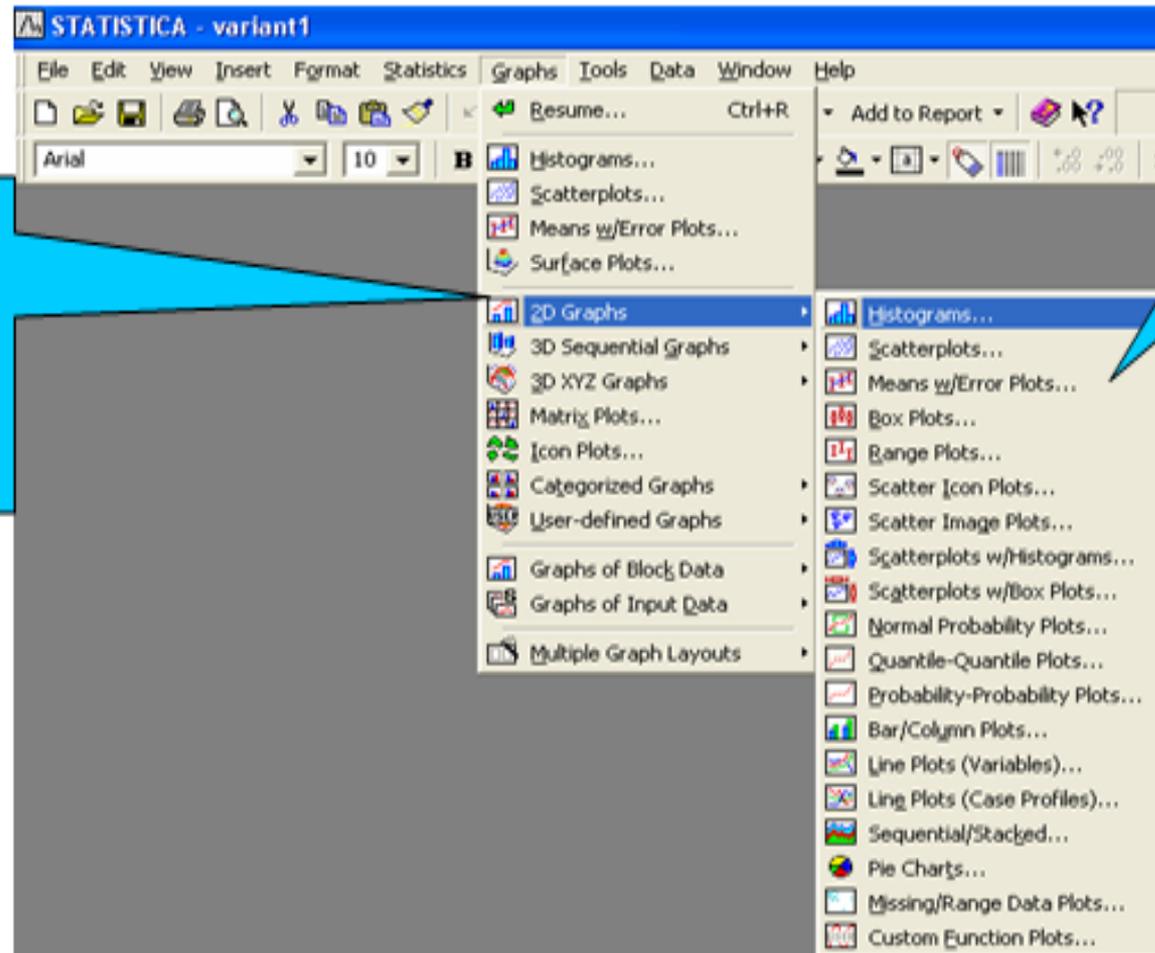
Лучшим способом наглядного представления данных является графический анализ. Данный метод является важным орудием науки, особенно в статистических и экономических исследованиях.

График – чертеж, показывающий соотношение статистических величин при помощи разнообразных геометрических и изобразительных средств.

Визуально гораздо проще увидеть закономерности, содержащиеся в таблицах.

Опция Graphs

Выбрать пространство, в котором будет осуществляться построение графика



Выбрать тип графика

Этапы графического анализа

- ↗ Построение графика
- ↗ Редактирование графика
- ↗ Сохранение изображения
- ↗ Вывод
- ↗ Создание отчета



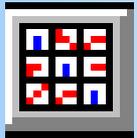
2М пользовательские графики



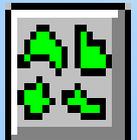
3М пользовательские графики



3М пользовательские диаграммы



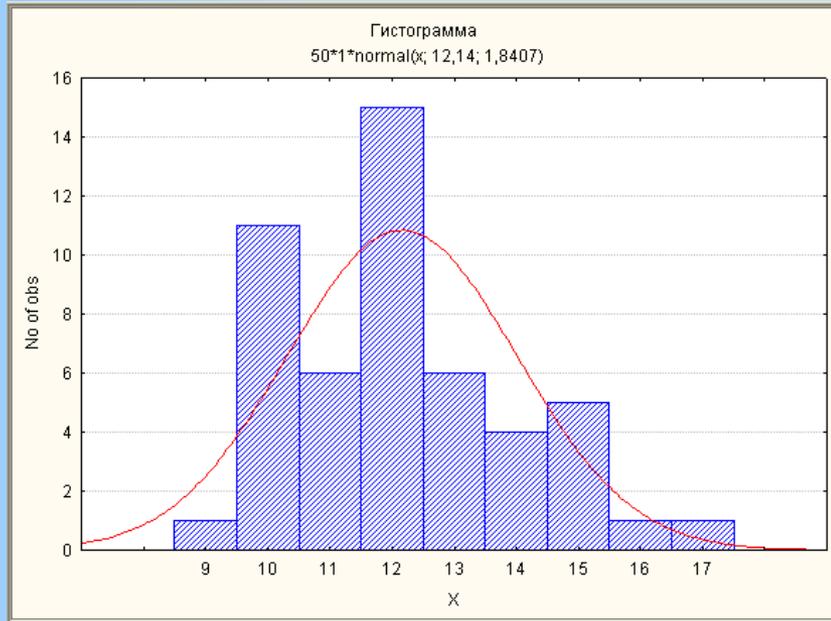
Матричные графики



Пользовательские пиктографики

Виды графиков

Гистограмма



Круговая диаграмма



Двухмерный визуальный анализ - это анализ данных на плоскости. В визуальном анализе используются разнообразные гистограммы, диаграммы рассеяния, вероятностные графики, линейные графики, диаграммы диапазонов, размахов, круговые диаграммы, столбчатые диаграммы, графики последовательных значений и т. д., позволяющие увидеть специфику данных.

Виды графиков

Трехмерный график

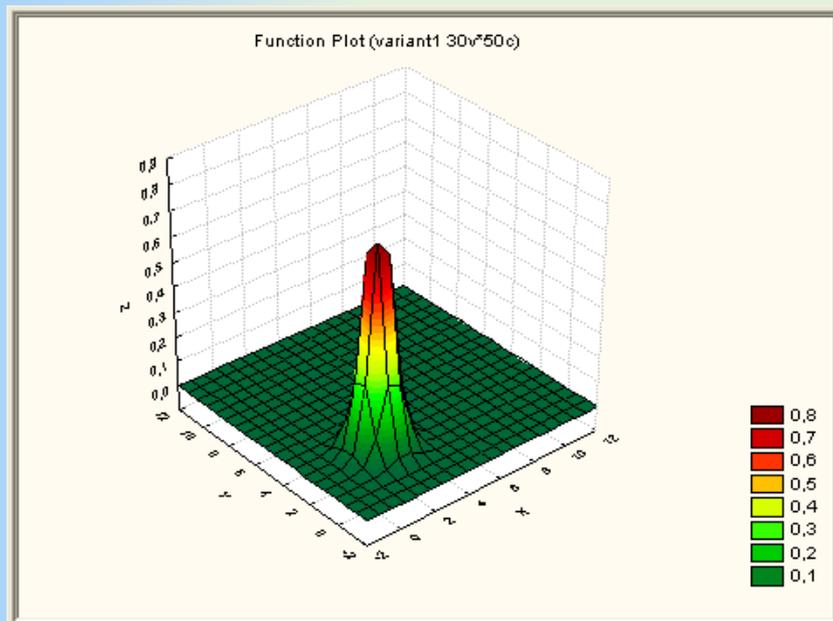
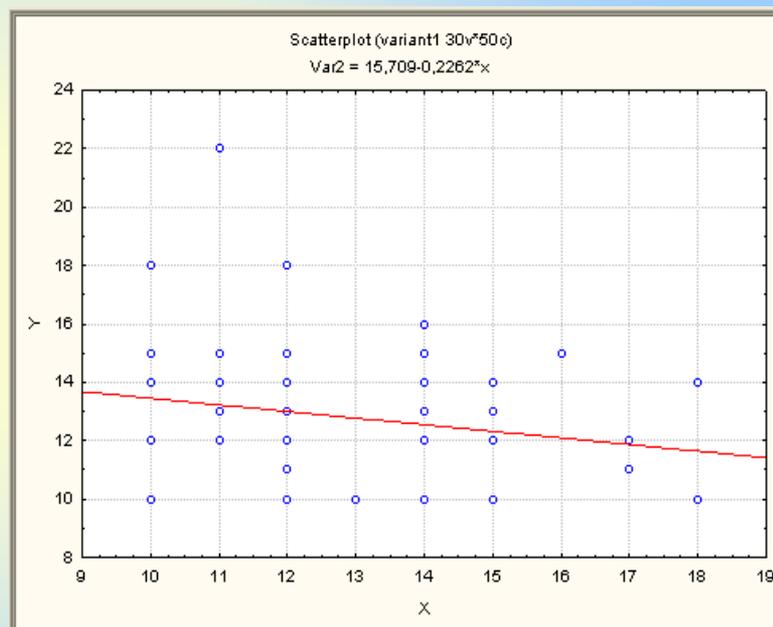


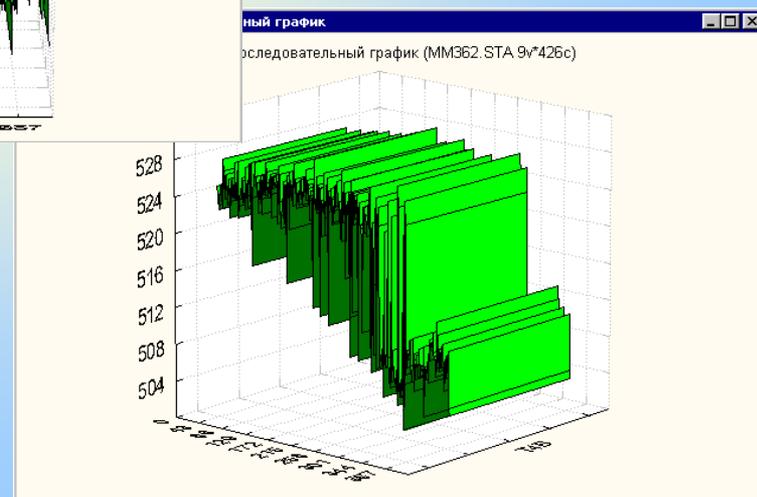
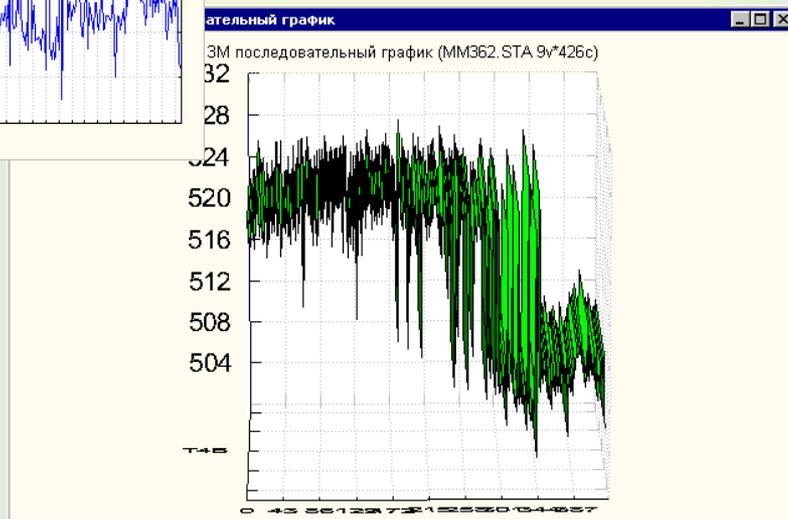
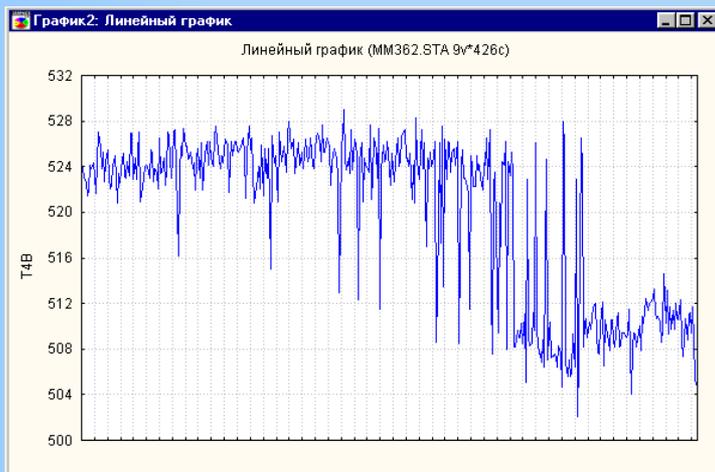
Диаграмма рассеяния



Прямая на диаграмме рассеяния – график простой линейной регрессии $y = 15,709 - 0,2262x$.

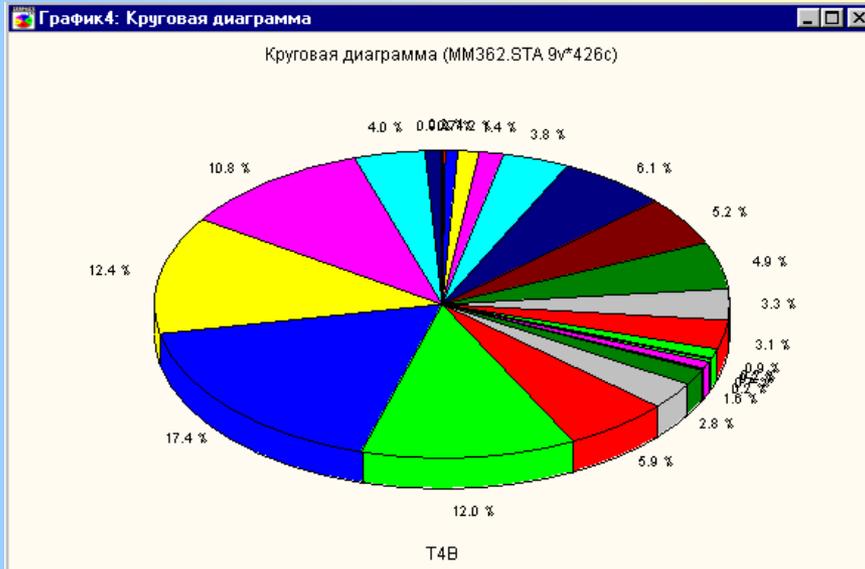
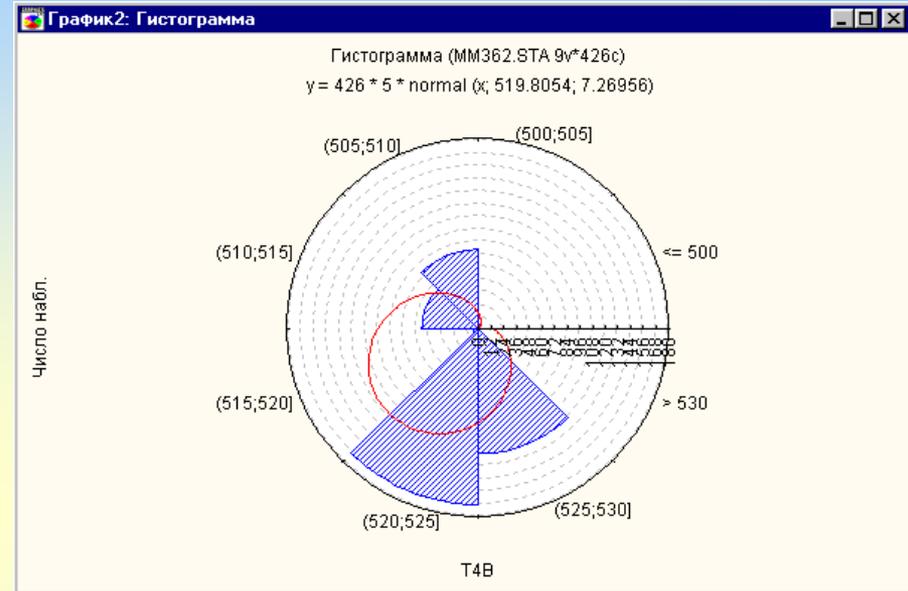
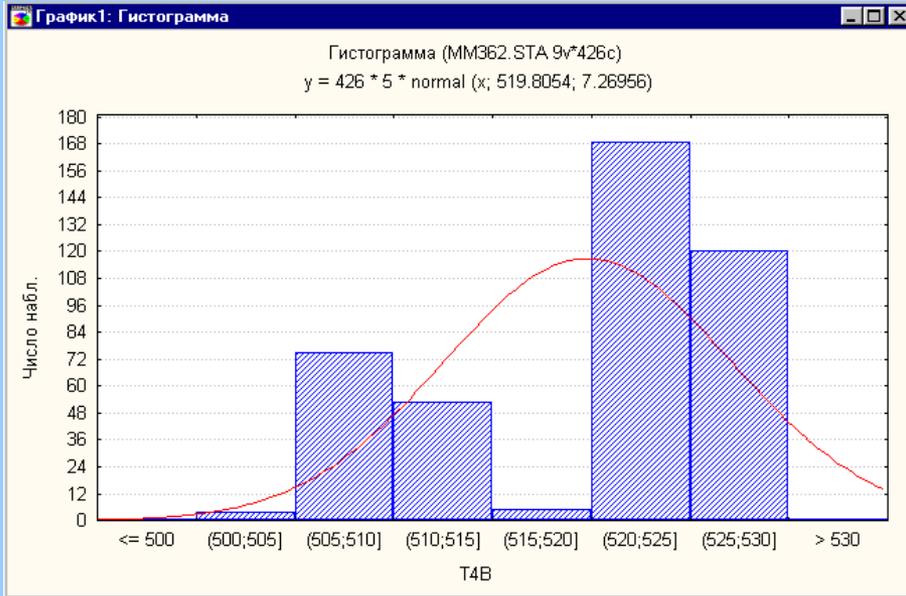
Виды графиков

Последовательный график



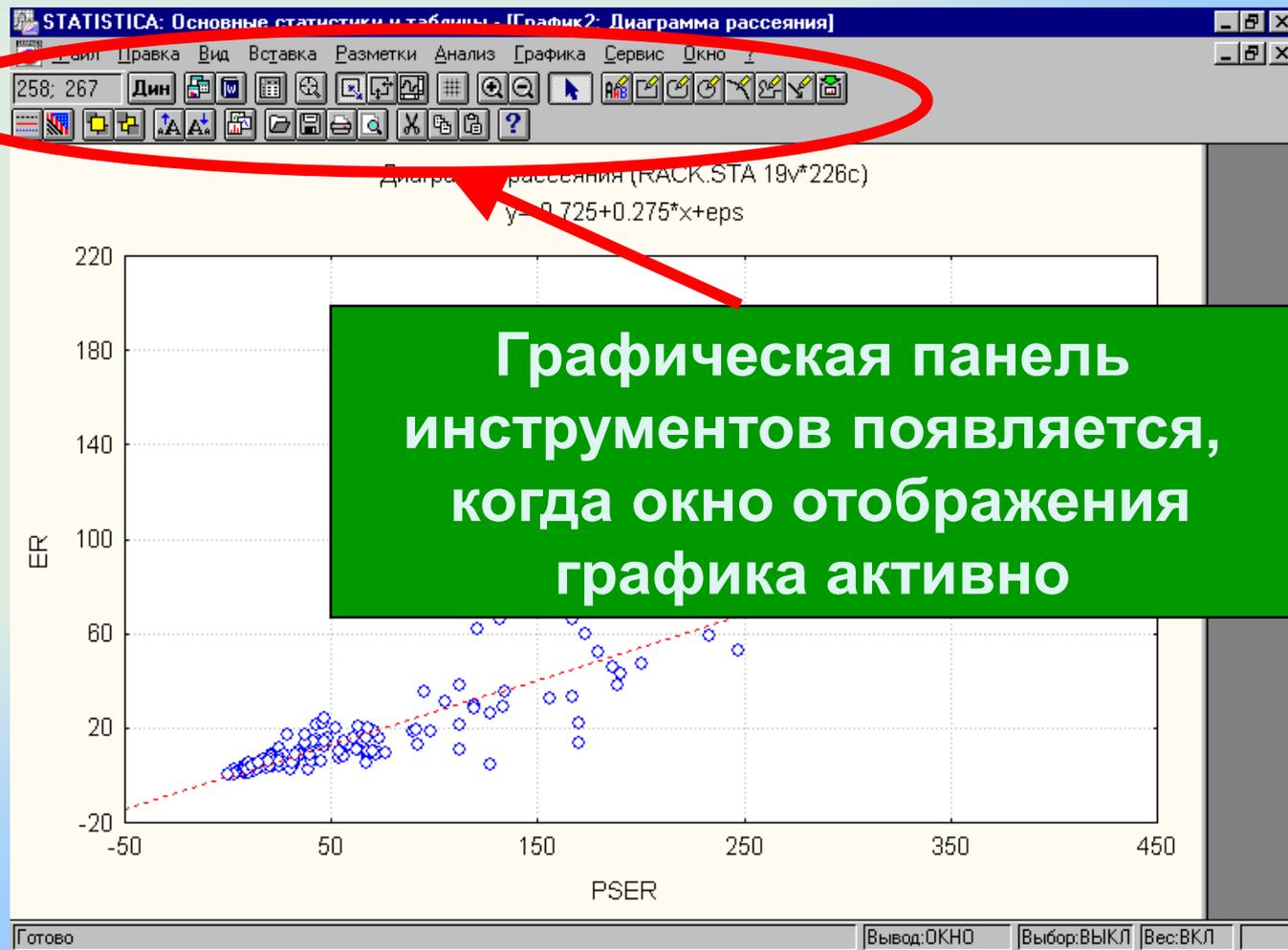
Трехмерный
график

Визуальный анализ данных

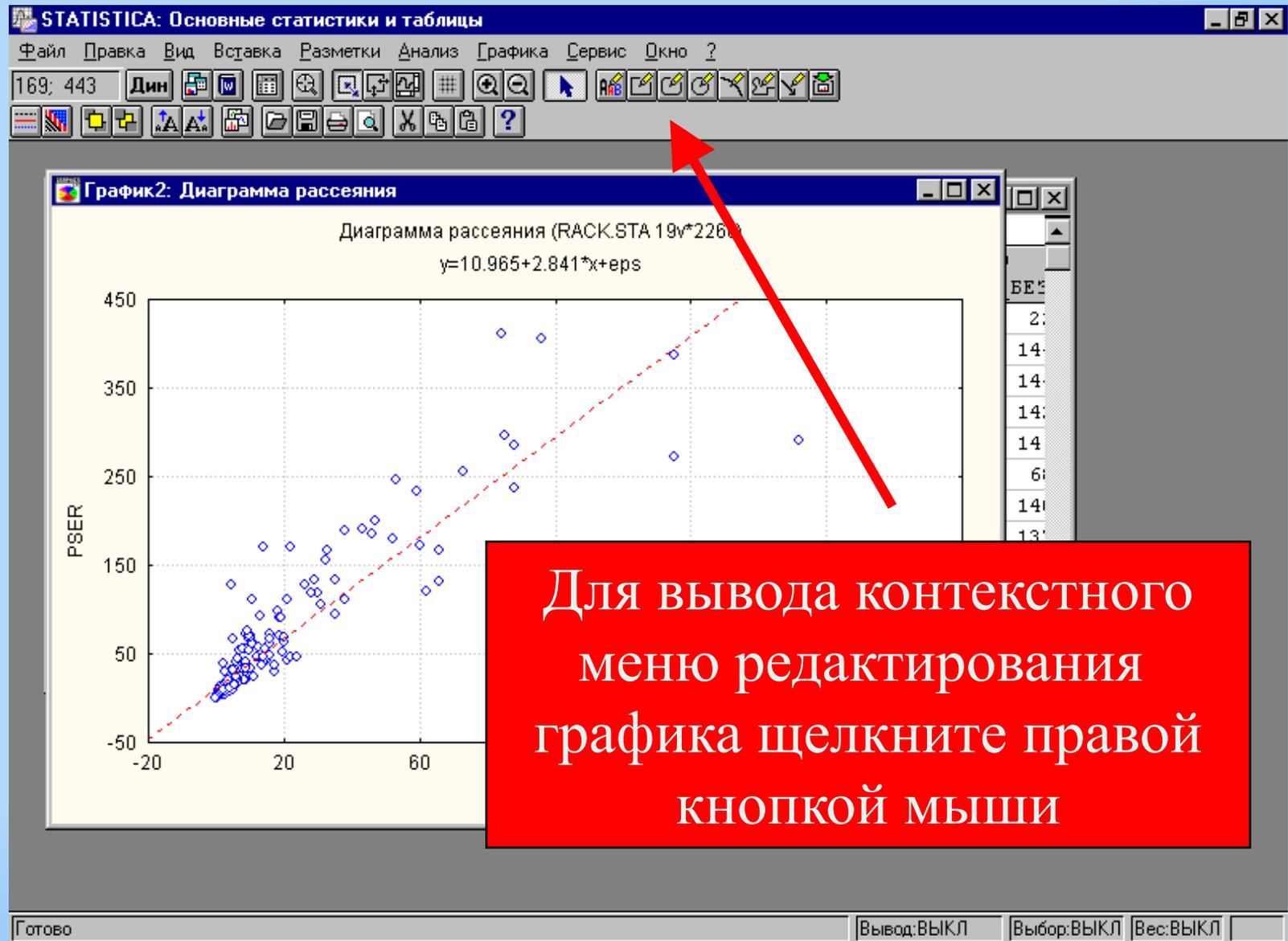


Редактирование

Для редактирования графика можно воспользоваться графической панелью инструментов; диалоговыми окнами редактирования графика



Редактирование



Редактирование

The screenshot displays the STATISTICA software interface. The main window is titled "График2: Диаграмма рассеяния" (Graph2: Scatter Plot). The plot shows a scatter plot of data points with a regression line. The equation of the line is $y=10.965+2.841*x+eps$. A red arrow points to the regression line, and a context menu is open over it, listing various editing options.

Context menu options:

- Изменить общую разметку...
- Изменить размещение графика...
- Вращение графика/Перспектива
- Редактировать данные графика...
- Задать область графика/поля
- Установить цвета фона...
- Зафиксировать условные обозначения
- Вставить объект...
- Вставить
- Копировать график
- Печать графика
- Сохранить график
- Захват экрана

Taskbar: Готово | Вывод: ВЫКЛ | Выбор: ВЫКЛ | Вес: ВЫКЛ | Start | Microsoft PowerPoint - [Гр... | STATISTICA: Основ... | EN 18:13

Контекстное меню
редактирования графика

Другие полезные графики

The screenshot shows the STATISTICA software interface. The main window displays a data table with columns 'НОМЕР_П' and 'ПОЛ'. A menu is open, showing options for creating various types of graphs, including '3D sequential graphs'. The status bar at the bottom indicates 'Построение 3М гистограммы распределения двух переменных по выбранным переменным'.

| ЧИСЛ ЗНАЧ | 1 НОМЕР_П | 2 ПОЛ | | | | | | | |
|--------------|--------------|----------|---|----|---|---|---|-----|-----|
| | 1 | 1 | 1 | | | | | | |
| | 2 | 2 | 2 | | | | | | |
| | 3 | 3 | 1 | | | | | | |
| | 4 | 4 | 1 | | | | | | |
| | 5 | 5 | 2 | | | | | | |
| | 6 | 6 | 1 | | | | | | |
| | 7 | 7 | 2 | | | | | | |
| | 8 | 8 | 1 | | | | | | |
| | 9 | 9 | 1 | 60 | 2 | 1 | 1 | 20 | 20 |
| | 10 | 10 | 1 | 62 | 2 | 1 | 1 | 137 | 137 |
| | 11 | 11 | 2 | 54 | 2 | 1 | 1 | 136 | 136 |
| | 12 | 12 | 2 | 63 | 3 | 1 | 1 | 136 | 136 |
| | 13 | 13 | 1 | 73 | 3 | 3 | 1 | 49 | 49 |

График27: ...

Построение 3М гистограммы распределения двух переменных по выбранным переменным | Вывод:ОКНО | Выбор:ВЫКЛ | Вес:ВКЛ

Start | D:\Полежаев\projects\Ц... | Microsoft PowerPoint - [Гр... | STATISTICA: Основ... | RU 13:40

Построим наглядную
трехмерную гистограмму

Методика

ЗМ гистограммы двух переменных

Переменные:

Ось X: ПОЛ
Ось Y: ВОЗРАСТ

ОК
Отмена
Параметры...

Ось X

Переменная: ПОЛ

Целые числа Авто

Категории: 10

Границы: нет

Коды: нет

Сложные подгруппы

Изменить переменную

Ось Y

Переменная:

Целые числа

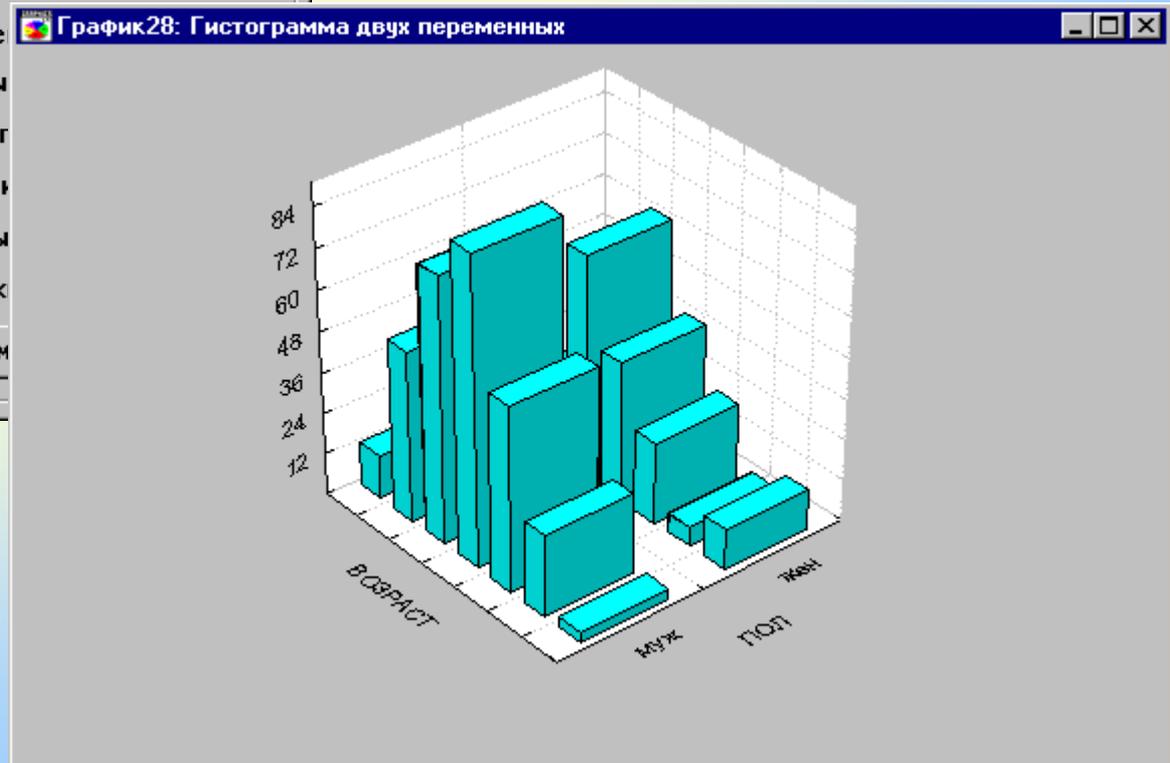
Категории

Границы

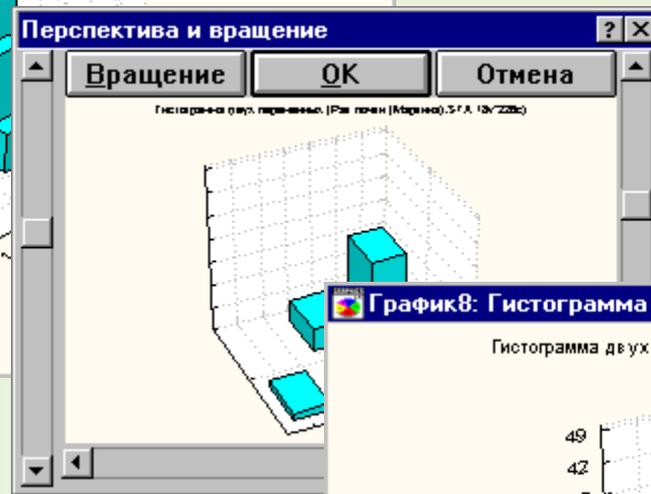
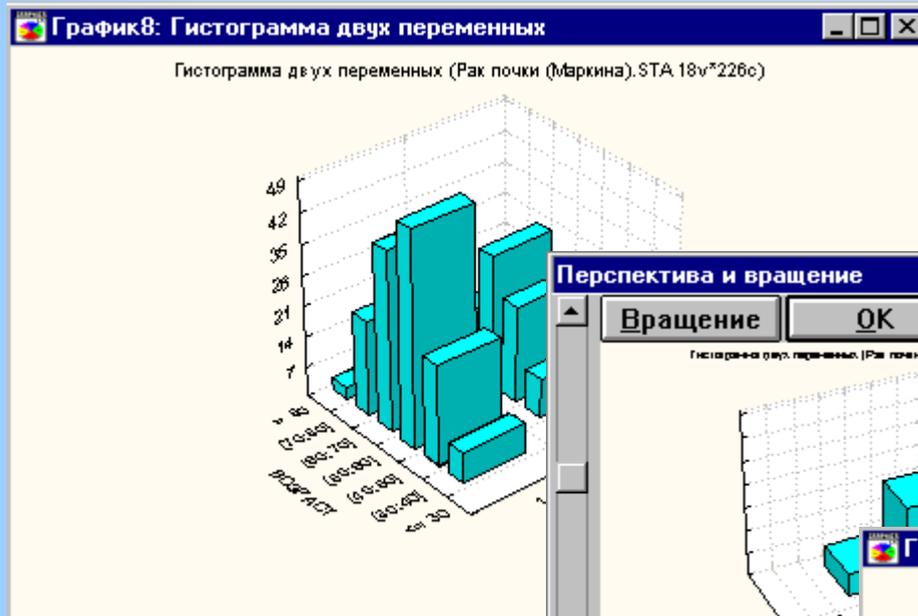
Коды

Сложные подгруппы

Изменить переменную

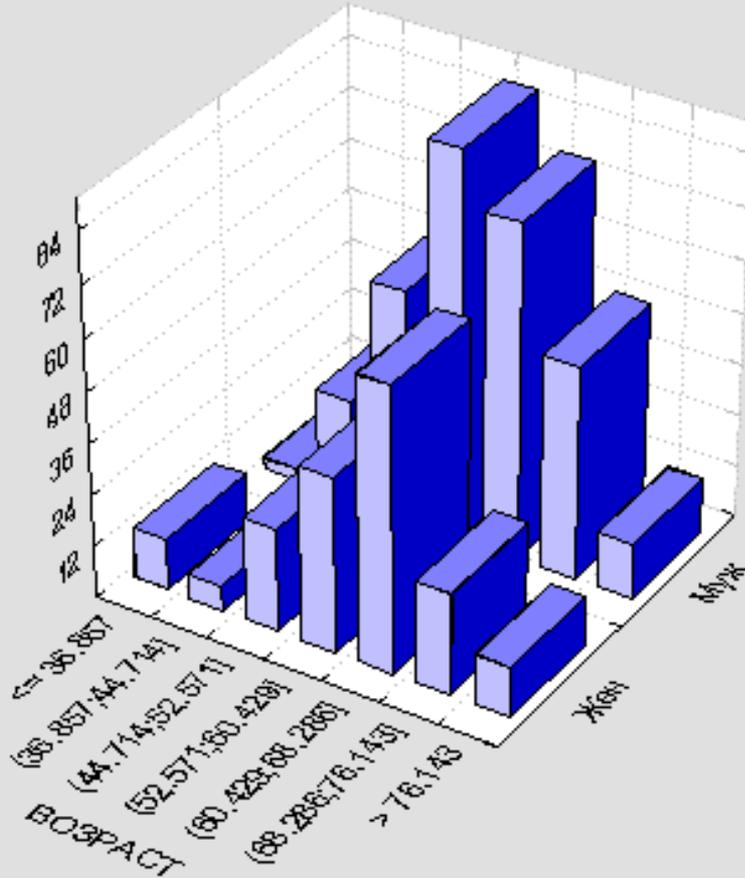


Вращение



Гистограмма -> поверхность

Гистограмма двух переменных



Общая разметка и размещение графика: 3М гистограммы

Тип графика:

- Столбчатая диаграмма
- Всплески
- Дискр. карта линий уровня
- График поверхности**
- Карта линий уровня

ЗАГОЛОВКИ

Заголовок 1: Гистограмма двух...

Гистограмма двух переменных (RACK.STA 19v*226c)

ОСИ

X Макс.: 2 Разметка оси: Ручная

Y Шаг: 1 Тип: Линейная

Z Мин.: 0

Объемный текст Параметры оси

ФОН

Снаружи: []

Внутри: []

ЧИСЛО СЕЧЕНИЙ ПОВЕРХНОСТИ

X: 25 Y: 25

СТИЛЬ КАРТЫ ЛИНИЙ УРОВНЯ

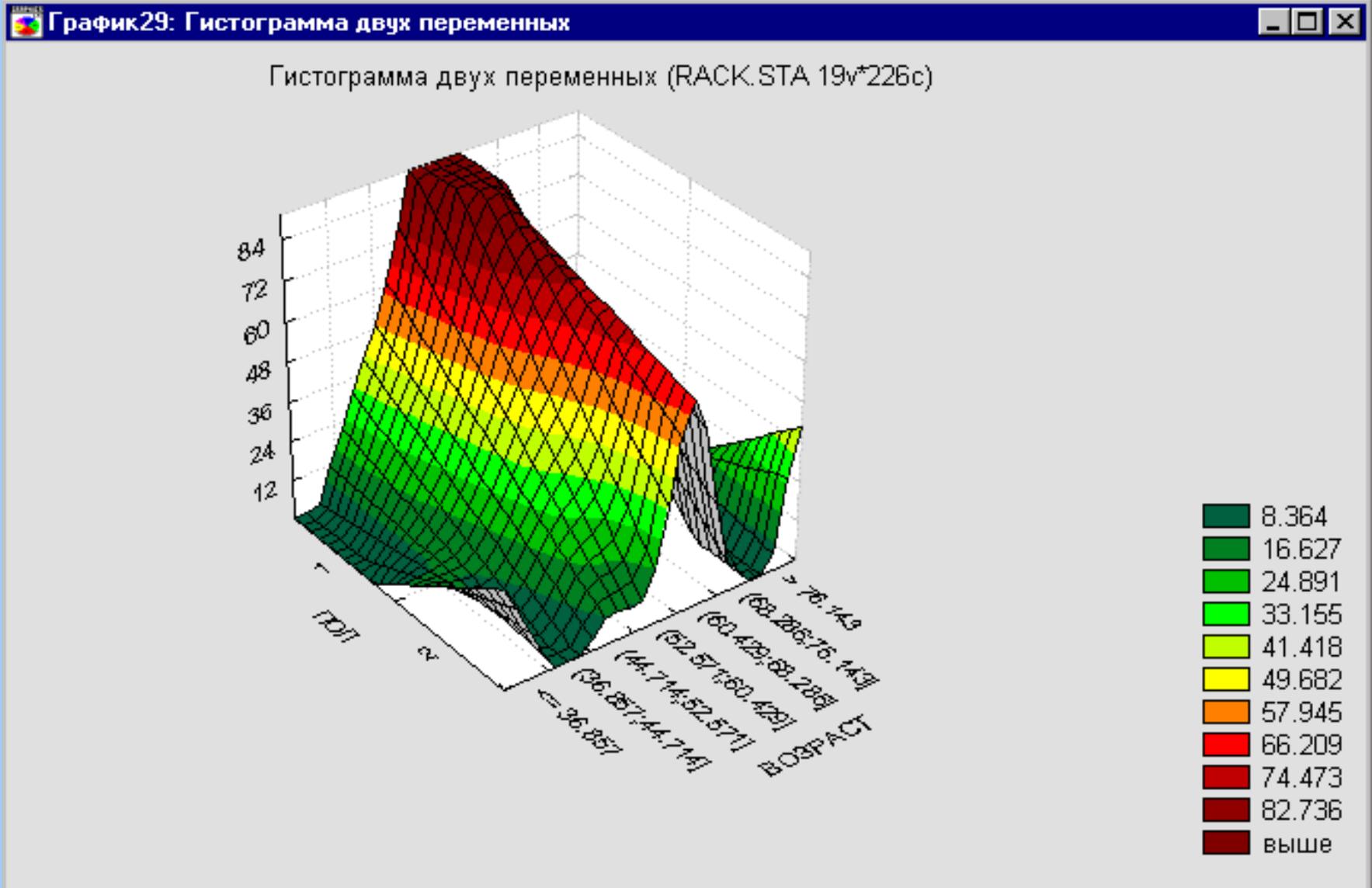
Линии Зоны Условные обозначения

ШАБЛОНЫ

- Верхняя грань
- Грань X
- Грань Y
- Ребра
- Точки
- Перпендикуляры

OK Отмена Оси Линии сетки Рамки Дополнительно...

Результат



Отчёт

Как перенести график, построенный в STATISTICA, в документ MS Word?



Отчёт

STATISTICA: Основные статистики и таблицы

Файл Правка Вид Вставка Разметки Анализ Графика Сервис Окно ?

73; 131 Дин

Данные: RACK.STA 19п * 226н

| ЧИСЛО ЗНАЧ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | ВЫК_БЕ... |
|------------|---------|---|---|---|---|---|---|---|---|-----------|
| | НОМЕР_П | | | | | | | | | |
| 2 | 2 | | | | | | | | | 14 |
| 3 | 3 | | | | | | | | | 14 |
| 4 | 4 | | | | | | | | | 14 |
| 5 | 5 | | | | | | | | | 14 |
| 6 | 6 | | | | | | | | | 6 |
| 7 | 7 | | | | | | | | | 14 |
| 8 | 8 | | | | | | | | | 13 |
| 9 | 9 | | | | | | | | | 2 |
| 10 | 10 | | | | | | | | | 13 |
| 11 | 11 | | | | | | | | | 13 |
| 12 | 12 | | | | | | | | | 13 |
| 13 | 13 | | | | | | | | | 4 |
| 14 | 14 | | | | | | | | | |

График4: Гистограмма

Гистограмма (RACK.STA 19п*226н)
 $y = 497 * 5 * \text{normal}(x; 59.7364; 10.9143)$

Число набл.

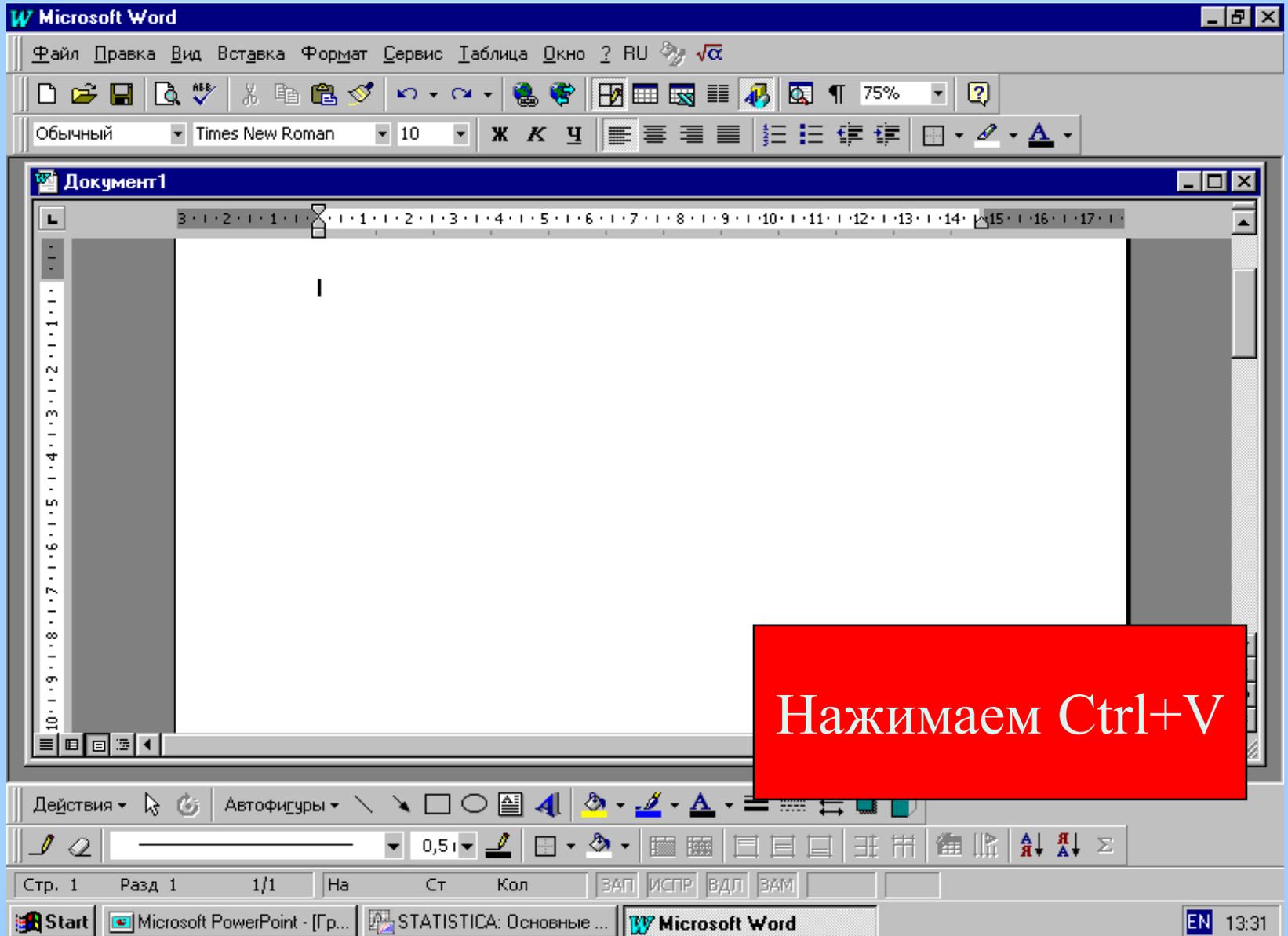
ВОЗРАСТ

Готово Вывод:ОКНО Выбор:ВЫКЛ Вес:ВКЛ

Start Microsoft PowerPoint - [Гр... STATISTICA: Основ... Microsoft Word EN 13:30

Нажимаем Ctrl+C

Отчёт



Отчёт

