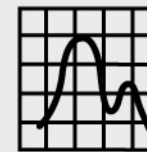


Визуальный анализ медицинских данных в программе *STATISTICA*

Сначала данные нужно увидеть...



StatSoft®
STATISTICA

U.S. Headquarters: StatSoft, Inc. • 2300 E. 14th St. • Tulsa, OK 74104 • USA • (918) 749-1119 • Fax: (918) 749-2217 • info@statsoft.com • www.statsoft.com

Australia: StatSoft Pacific Pty Ltd.

Brazil: StatSoft South America

Bulgaria: StatSoft Bulgaria Ltd.

Czech Rep.: StatSoft Czech Rep. s.r.o.

China: StatSoft China

France: StatSoft France

Germany: StatSoft GmbH

Hungary: StatSoft Hungary Ltd.

India: StatSoft India Pvt. Ltd.

Israel: StatSoft Israel Ltd.

Italy: StatSoft Italia srl

Japan: StatSoft Japan Inc.

Korea: StatSoft Korea

Netherlands: StatSoft Benelux BV

Norway: StatSoft Norway AS

Poland: StatSoft Polska Sp. z o.o.

Portugal: StatSoft Ibérica Lda

Russia: StatSoft Russia

Spain: StatSoft Ibérica Lda

S. Africa: StatSoft S. Africa (Pty) Ltd.

Sweden: StatSoft Scandinavia AB

Taiwan: StatSoft Taiwan

UK: StatSoft Ltd.

Типы графиков, наиболее часто используемые при визуальном анализе

- *Гистограмма (Frequency plot, Histogram, Bar chart)*
- *График средних с ошибками (Error bar plot)*
- *Диаграмма размаха (Box&whisker plot)*
- *Диаграмма рассеяния (Scatter plot)*
- *Диаграмма концентрации (Bag plot)*
- *Диаграмма Вороного (Voronoi diagram)*
- *Лица Чернова (Chernoff faces)*

- *Категоризованные графики (Categorized plots)*

Данные для анализа

- по строкам –
пациенты
- по столбцам –
*характеристики
пациентов:*
 - a. пол*
 - b. возраст*

	1 Номер пациента	2 Пол	3 Возраст
пациент 1	1	жен	64
пациент 2	2	муж	39
пациент 3	3	жен	45
пациент 4	4	жен	33
пациент 5	5	муж	63
пациент 6	6	жен	57
пациент 7	7	жен	56
пациент 8	8	жен	83
пациент 9	9	жен	50
пациент 10	10	жен	54
пациент 11	11	жен	55
пациент 12	12	муж	39
пациент 13	13	муж	41
пациент 14	14	жен	38
пациент 15	15	жен	50
пациент 16	16	жен	52
пациент 17	17	жен	20
пациент 18	18	жен	50
пациент 19	19	жен	57
пациент 20	20	жен	57

Данные для анализа

по столбцам

(продолжение):

**биохимический анализ
крови -**

a. эритроциты

b. тромбоциты

c. лейкоциты

d. гемоглобин

	4 Эритроциты	5 Тромбоциты	6 Лейкоциты	7 Гемоглобин
пациент 1	3,43333333	198	1,05084746	61
пациент 2	3,9	141	7,88474576	130
пациент 3	3,93333333	141	6,25762712	106
пациент 4	4	175	7,88474576	137
пациент 5	3,86666667	141	7,07118644	106
пациент 6	3,3	209	1,05084746	65
пациент 7	4,06666667	175	7,07118644	133
пациент 8	3,83333333	73	6,09491525	106
пациент 9	3,73333333	186	6,58305085	116
пациент 10	3,3	232	0,4	61
пациент 11	3,8	164	6,25762712	102
пациент 12	3,76666667	130	7,07118644	109
пациент 13	3,93333333	164	7,23389831	123
пациент 14	3,63333333	107	5,11864407	92
пациент 15	3,93333333	164	7,7220339	116
пациент 16	3,7	130	7,07118644	120
пациент 17	4,03333333	186	8,37288136	133
пациент 18	3,46666667	198	1,53898305	72
пациент 19	3,66666667	141	6,09491525	99
пациент 20	3,83333333	209	7,55932203	133

Данные для анализа

по столбцам
(продолжение):
лечение -

a. группа

b. состояние до лечения

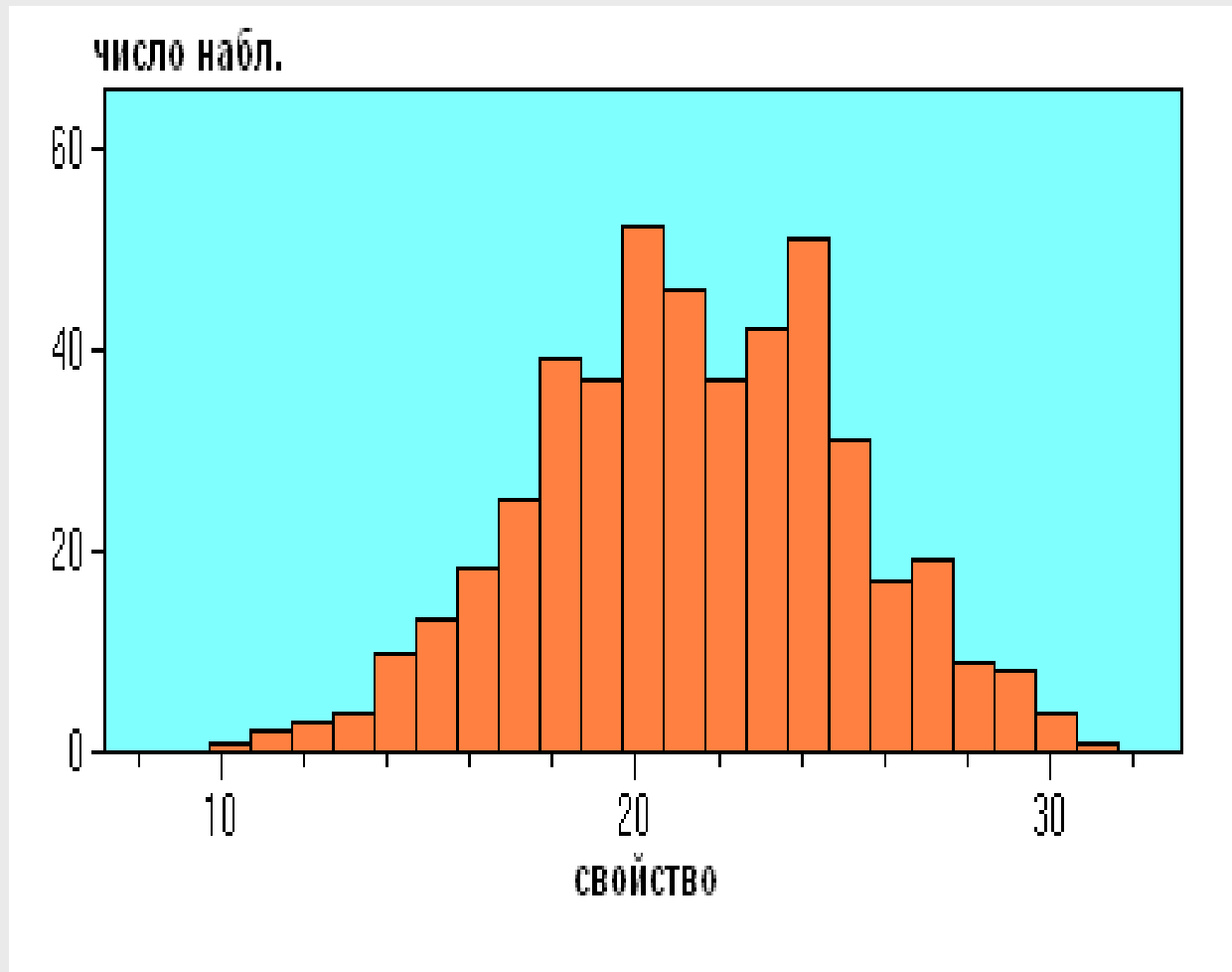
c. состояние после лечения

d. эффективность

	8 Группа	9 Состояние до лечения	10 Состояние после лечения	11 Эффективность
пациент 1	I	100	60	-0,4
пациент 2	II	100	80	-0,2
пациент 3	I	60	80	0,25
пациент 4	II	20	30	0,3333333333
пациент 5	II	10	80	0,875
пациент 6	I	40	100	0,6
пациент 7	II	90	30	-0,6666666667
пациент 8	I	100	30	-0,7
пациент 9	I	80	40	-0,5
пациент 10	I	30	40	0,25
пациент 11	I	10	100	0,9
пациент 12	I	30	20	-0,3333333333
пациент 13	II	80	100	0,2
пациент 14	I	100	70	-0,3
пациент 15	II	90	100	0,1
пациент 16	II	40	20	-0,5
пациент 17	II	10	40	0,75
пациент 18	I	20	60	0,6666666667
пациент 19	I	60	80	0,25
пациент 20	II	100	100	0

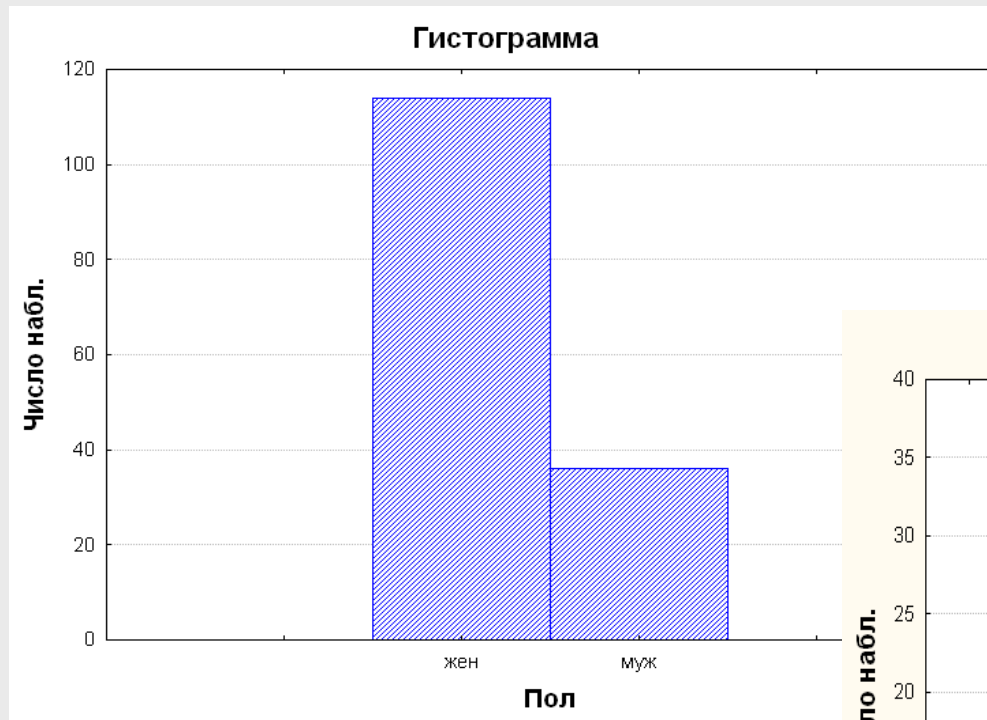
Гистограмма

(Frequency plot, Histogram, Bar chart)



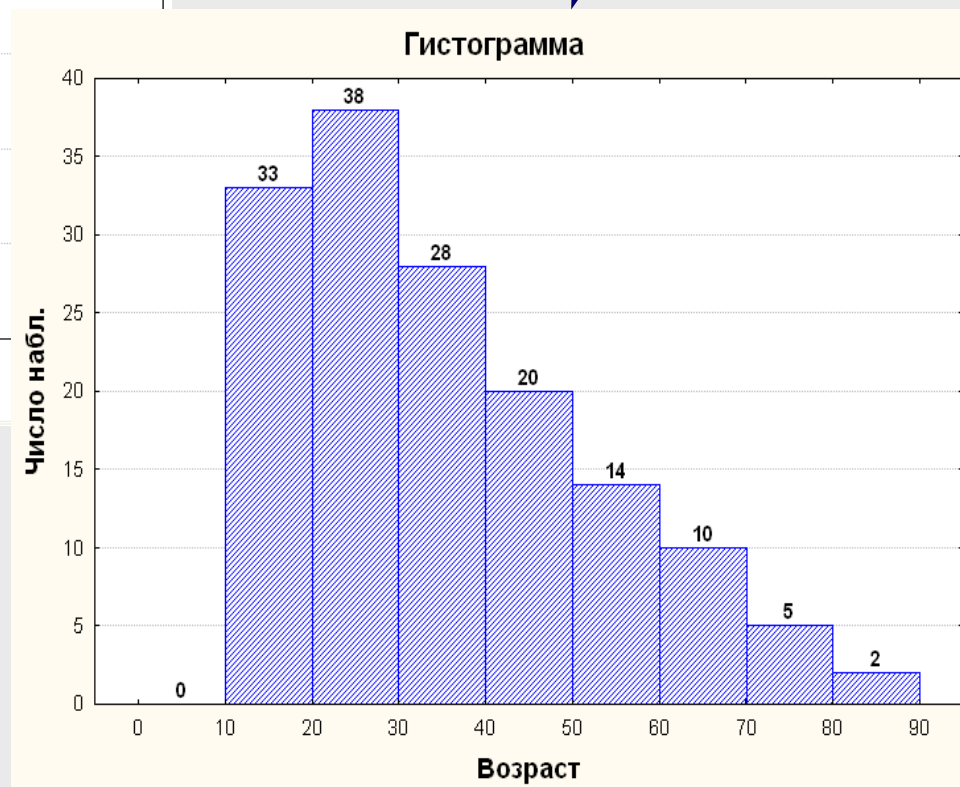
Гистограмма – столбчатая диаграмма распределения частот для выбранной переменной

Примеры гистограмм



*распределение
пациентов по полу*

*распределение
пациентов по
возрасту*

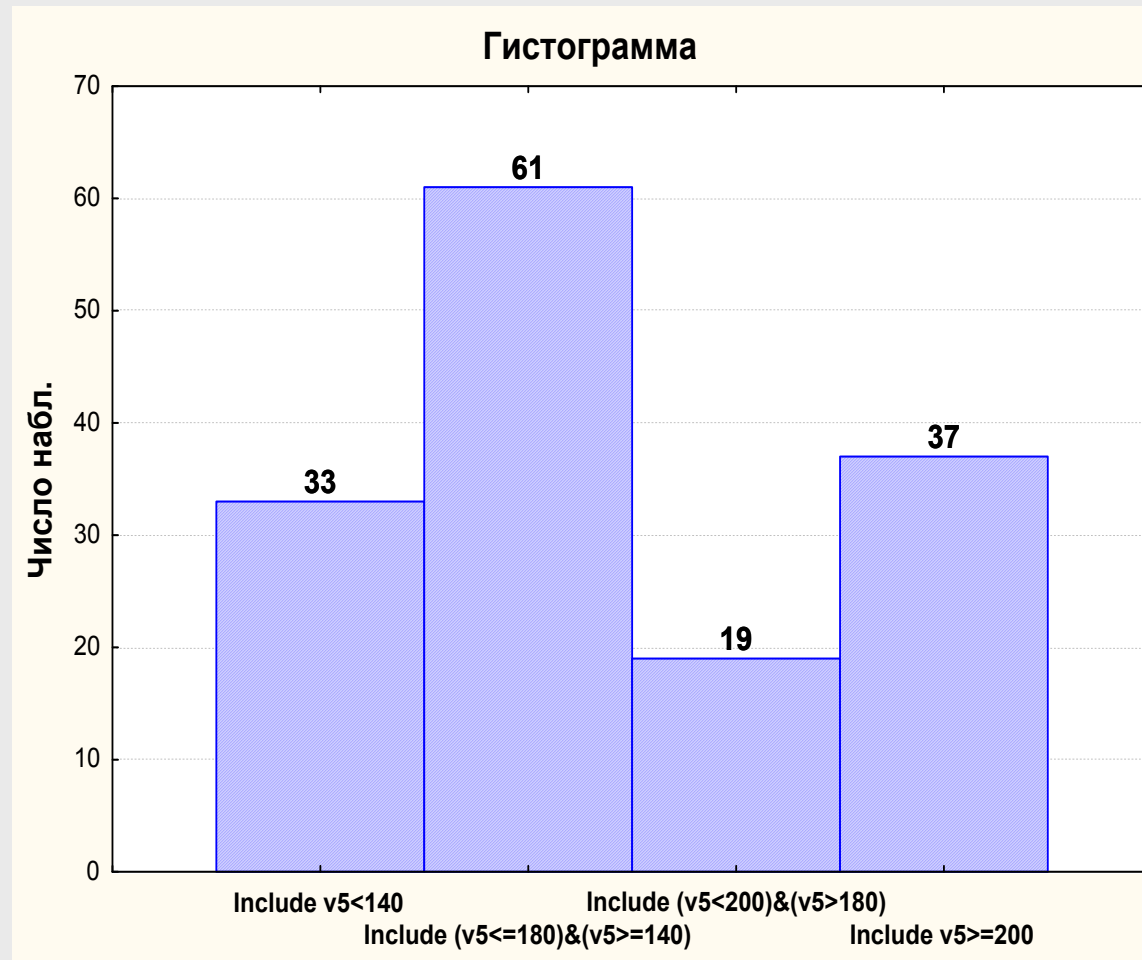


Пример: исследование состояния пациентов до и после лечения



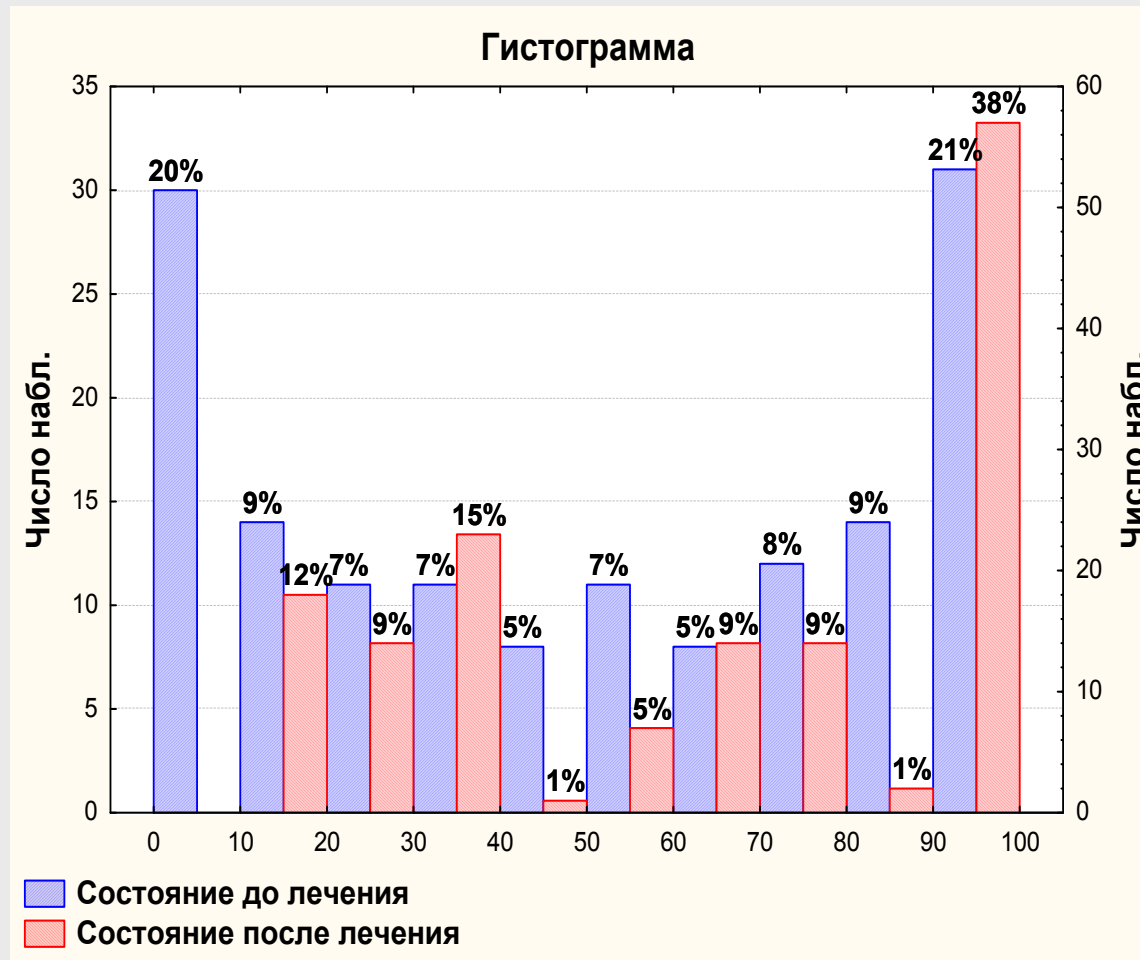
Цель: показать распределение пациентов по уровню эритроцитов в крови

Пример: исследование состояния пациентов до и после лечения



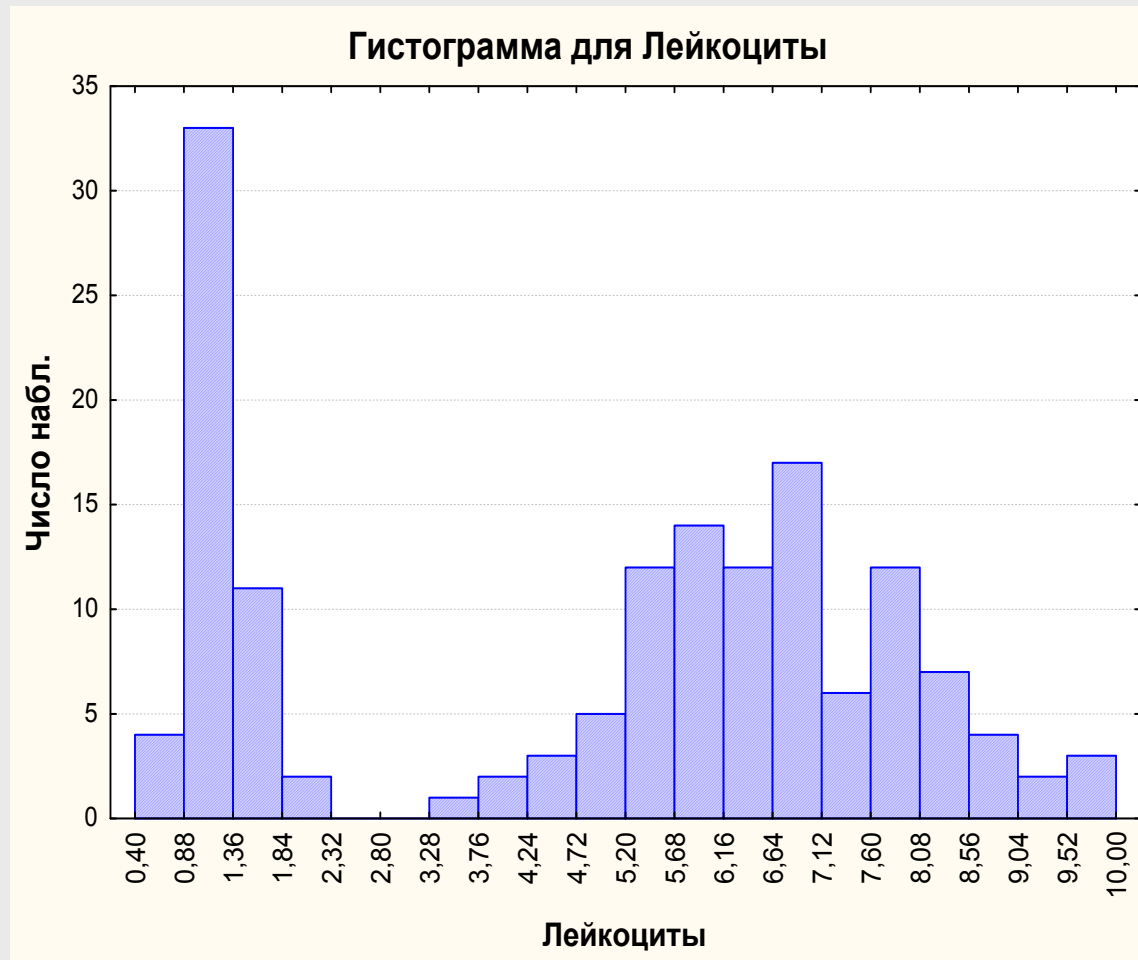
Цель: посмотреть состояние пациентов до лечения в зависимости от уровня тромбоцитов в крови

Пример: исследование состояния пациентов до и после лечения



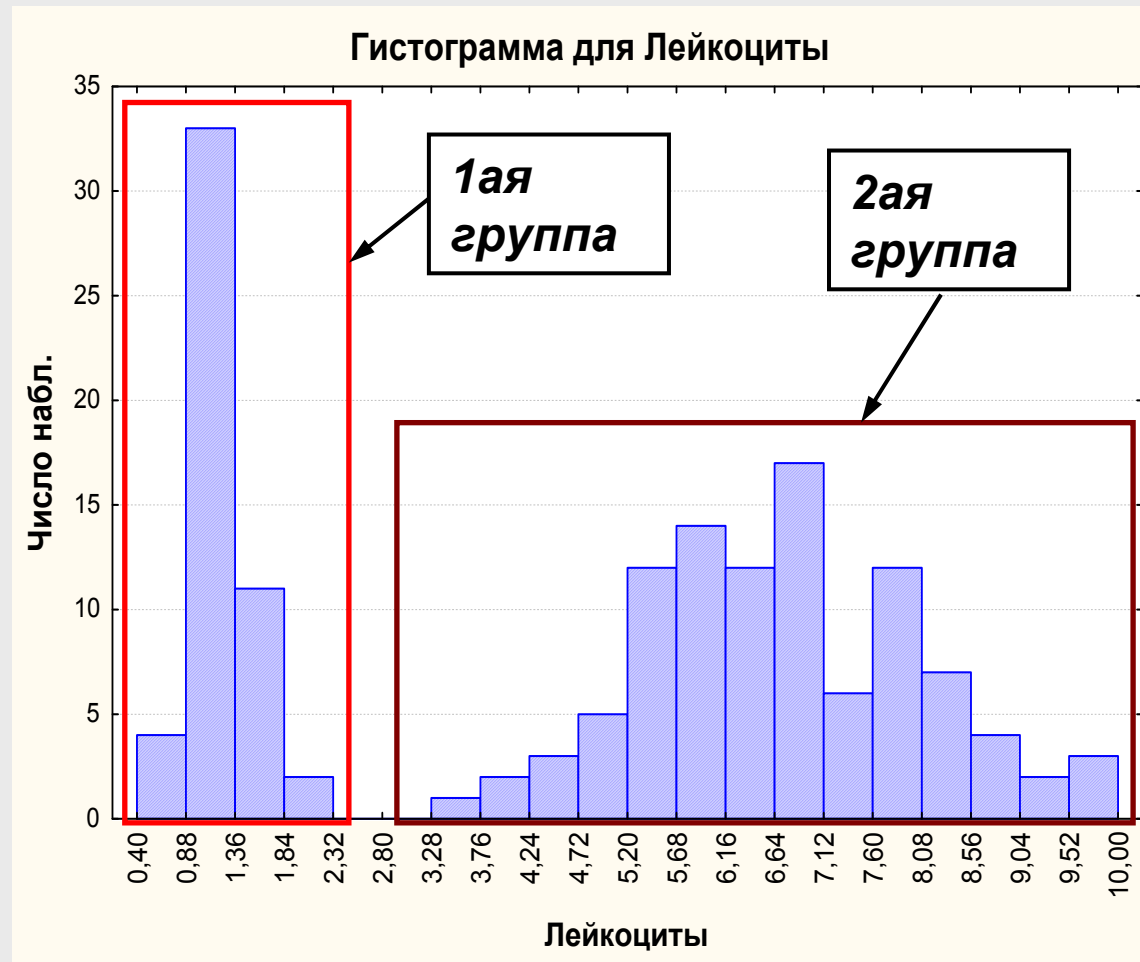
Цель: сравнить распределения пациентов по состоянию до и после лечения на одной составной гистограмме

Пример: исследование состояния пациентов до и после лечения



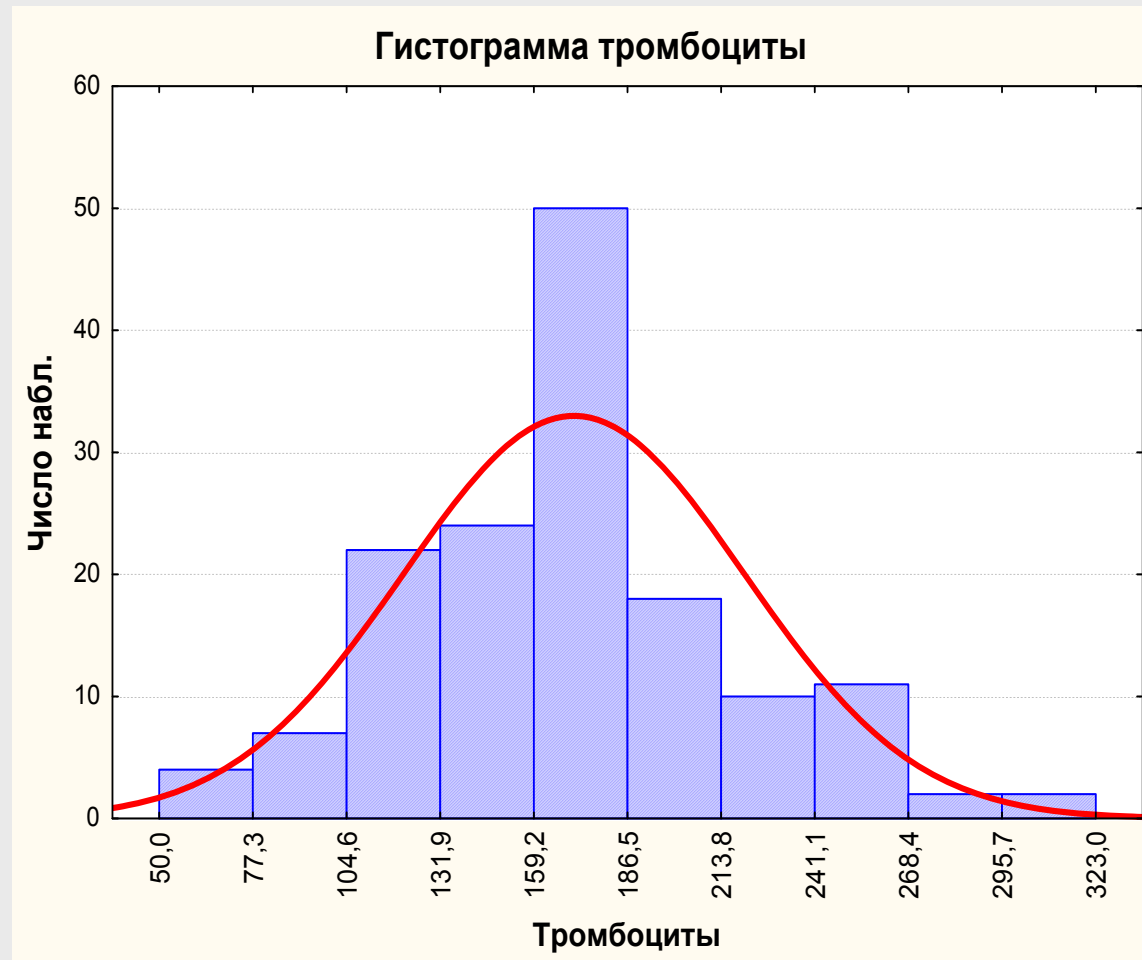
Цель:
охарактеризовать
природу переменной
Лейкоциты
(бимодальность,
выборка
неоднородна)

Пример: исследование состояния пациентов до и после лечения



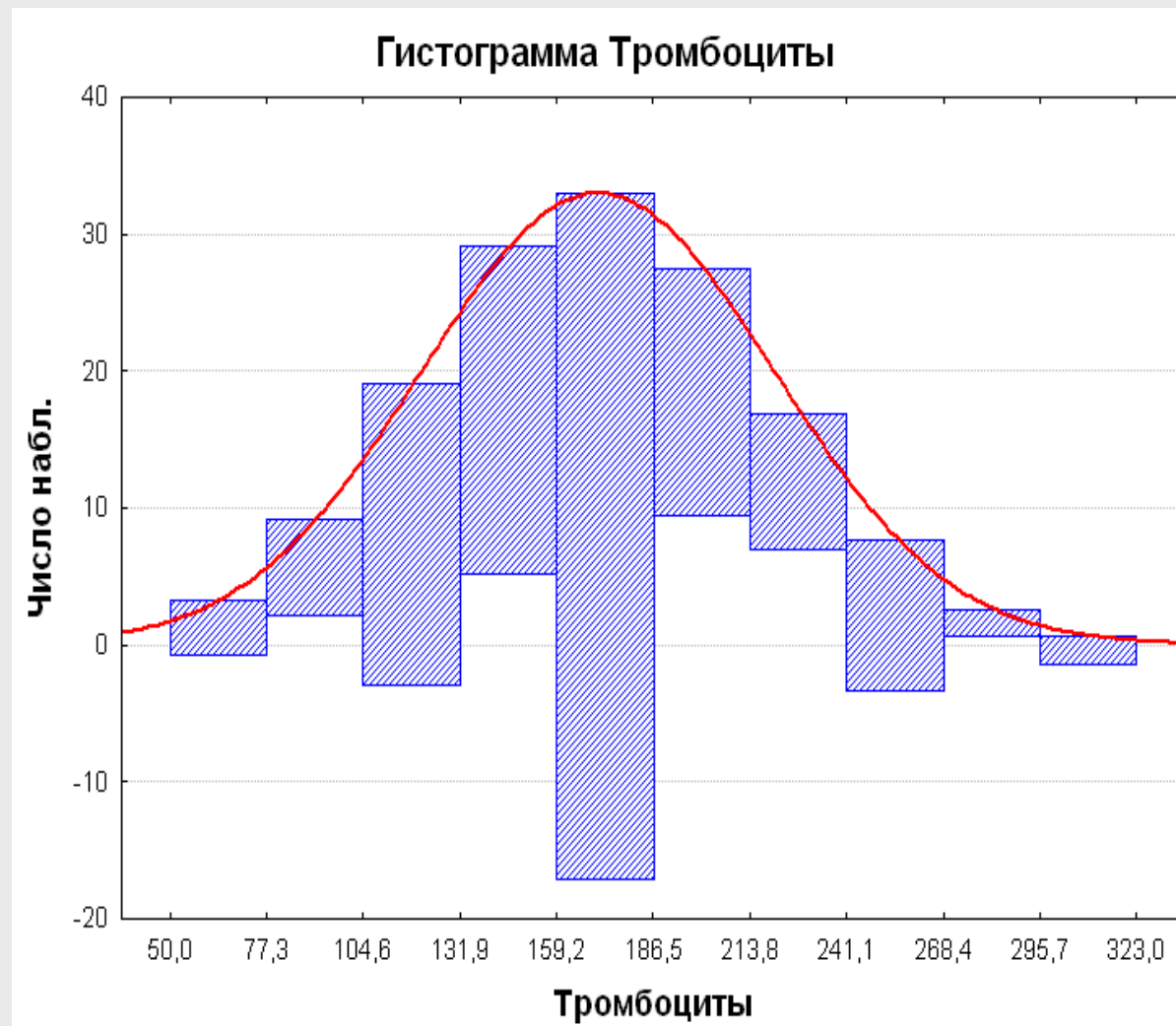
Цель:
охарактеризовать природу переменной Лейкоциты (бимодальность, выборка неоднородна)

Пример: исследование состояния пациентов до и после лечения



Цель: проверка предположения о нормальности распределения переменной (1ый способ)

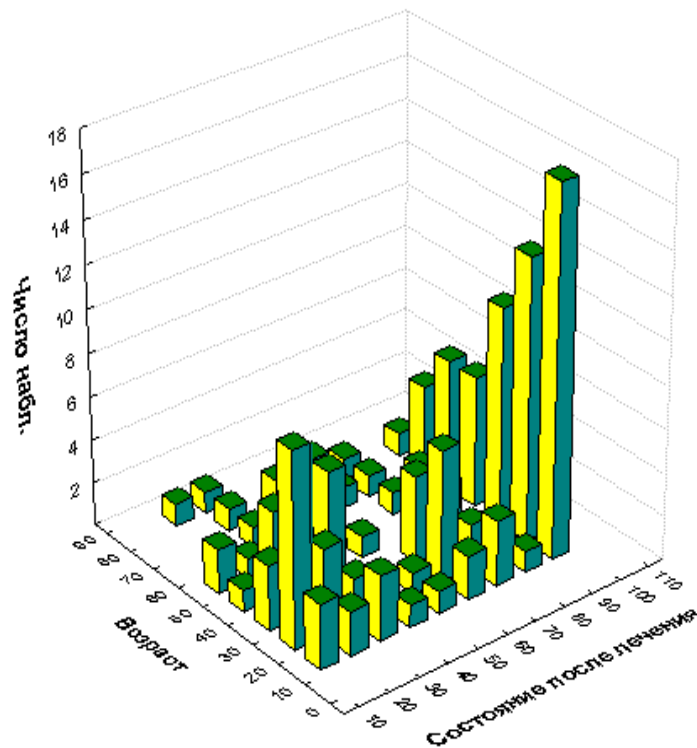
Пример: исследование состояния пациентов до и после лечения



Цель: проверка предположения о нормальности распределения переменной (2ой способ)

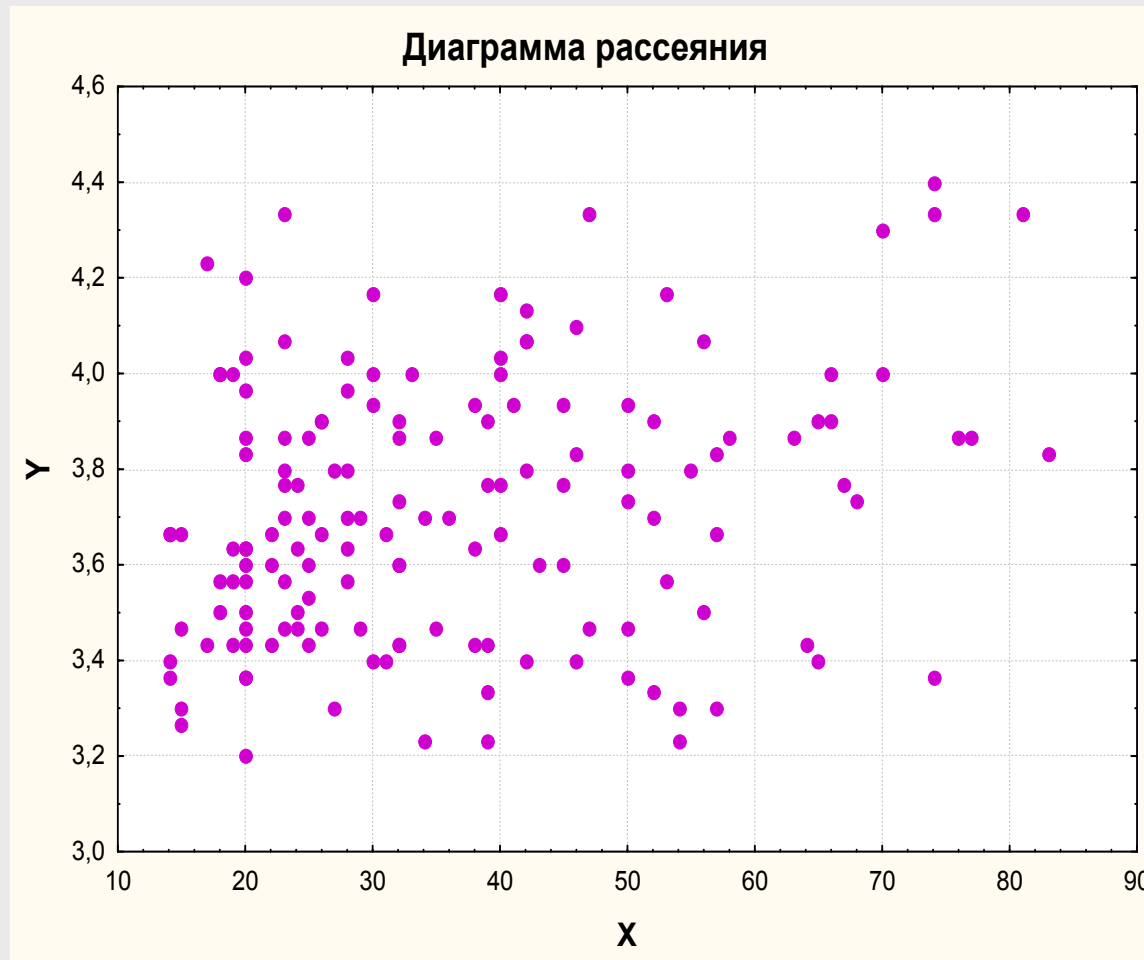
Пример: исследование состояния пациентов до и после лечения

Гистограмма двух переменных
(Возраст и Состояние после лечения)



Цель:
проанализировать частоты совместного появления значений двух переменных

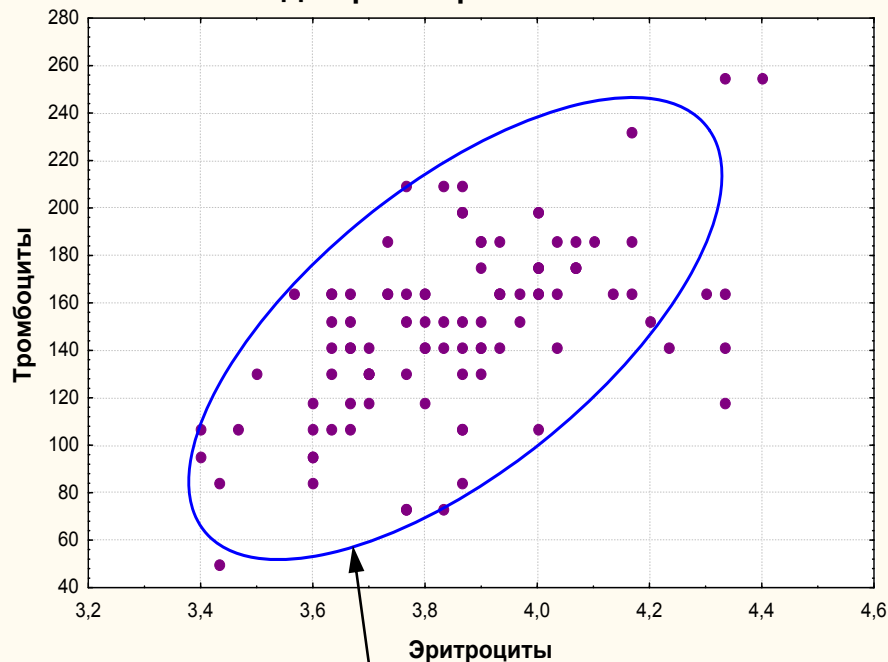
Диаграмма рассеяния (Scatter plot)



**Основное
назначение:**
**проанализировать
зависимость между
переменными X и Y**

Пример: исследование состояния пациентов до и после лечения

Диаграмма рассеяния

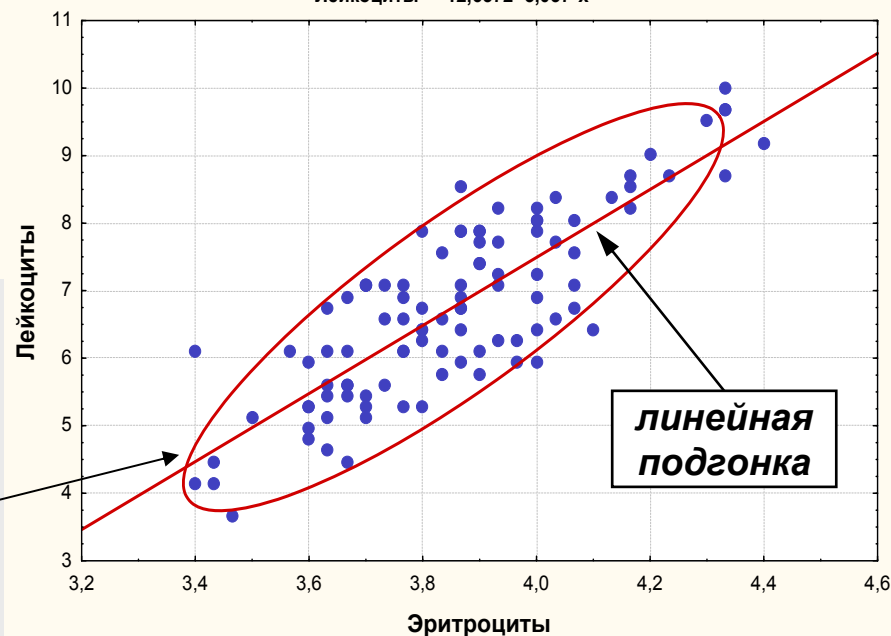


эллипс для предсказанного интервала (95%)

Цель: исследовать наличие зависимости между признаками

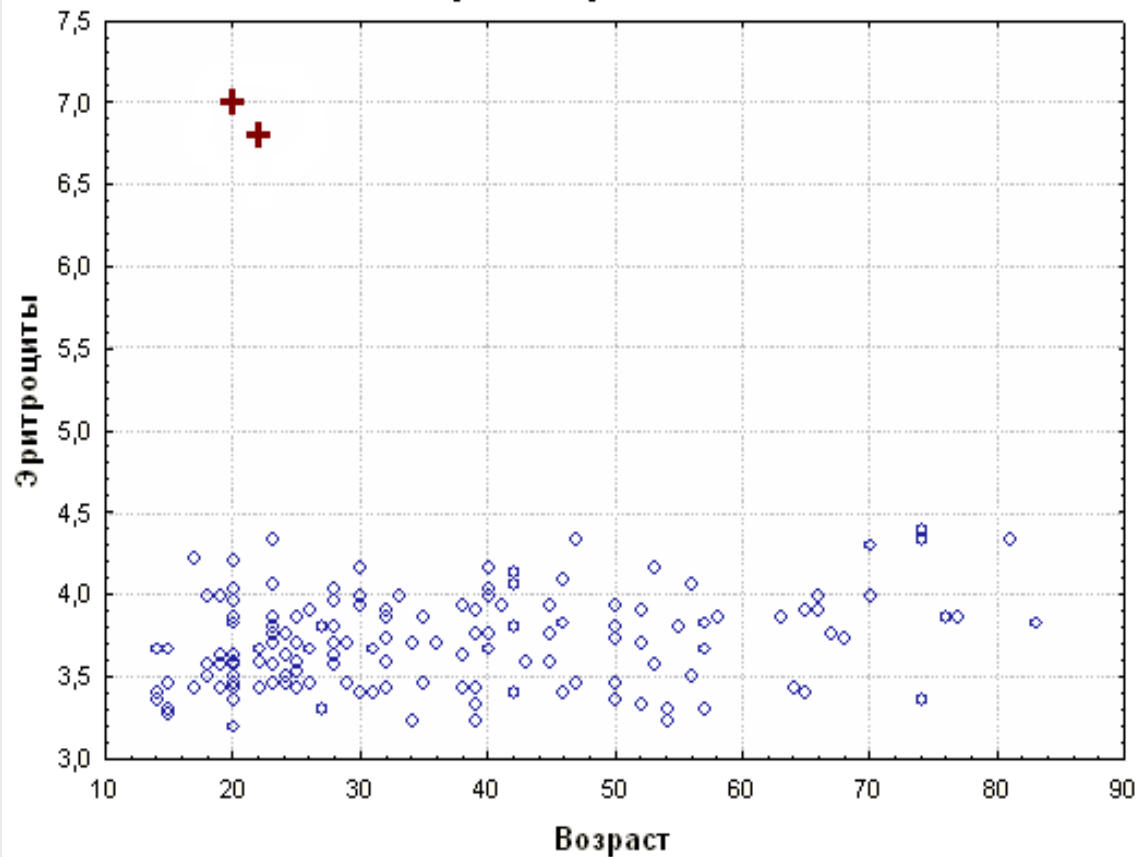
Диаграмма рассеяния

$$\text{Лейкоциты} = -12,6572 + 5,037 \cdot x$$



Пример: исследование состояния пациентов до и после лечения

Диаграмма рассеяния



Цель:
поиск выбросов
(нетипичных значений)

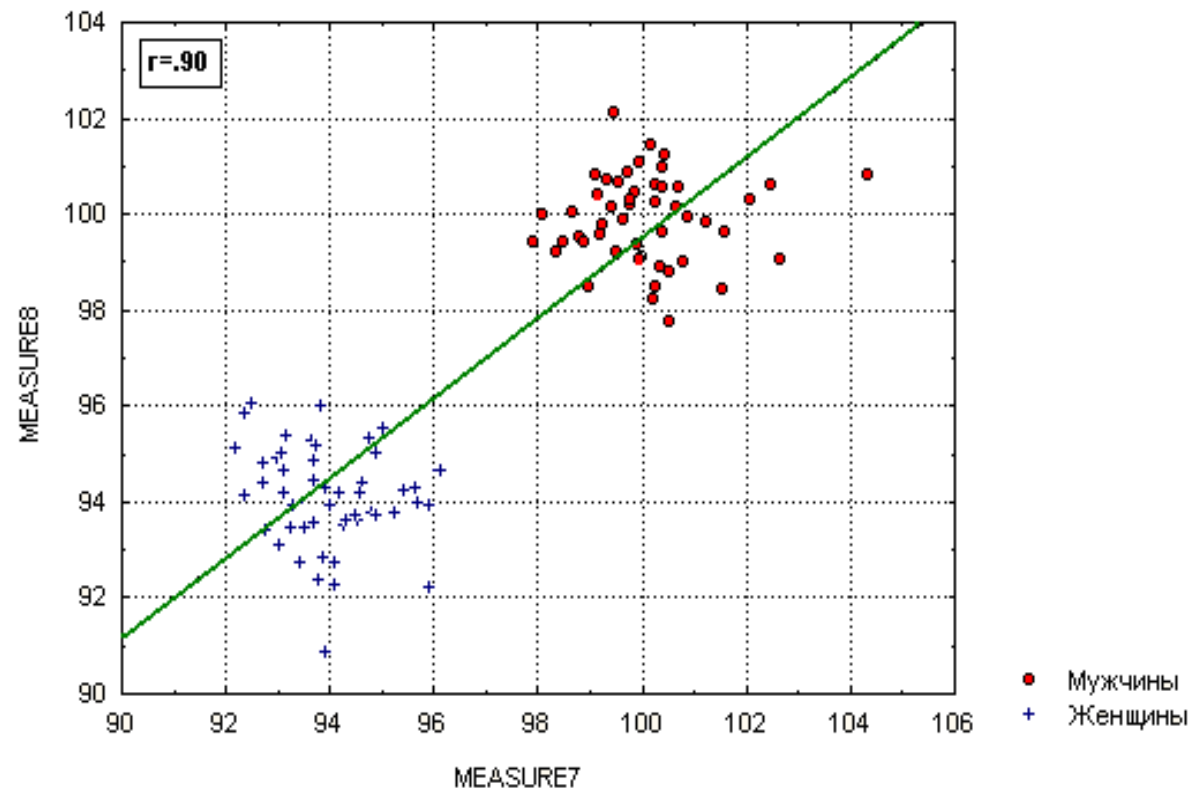
Пример: исследование состояния пациентов до и после лечения



Цель:
поиск неоднородных групп

Пример: исследование состояния пациентов до и после лечения

Диаграмма рассеяния

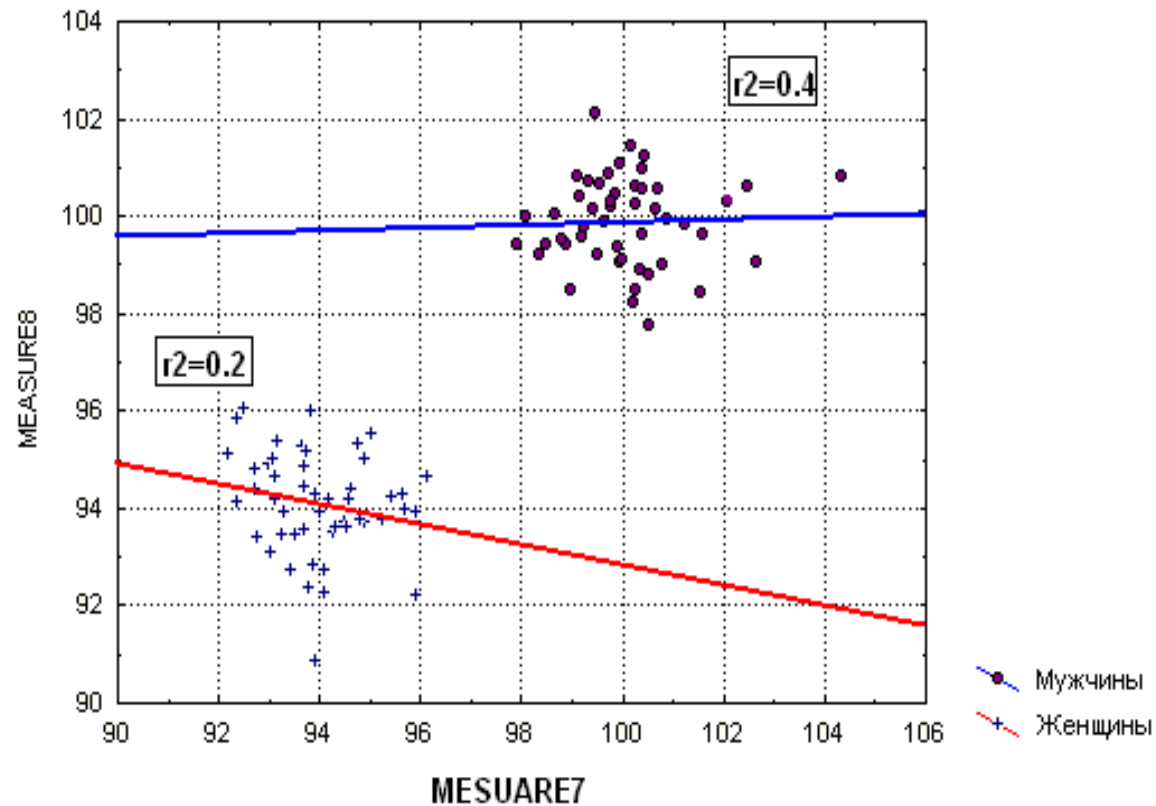


Вопрос:

что будет, если мы не учитываем наличие неоднородных групп?

Пример: исследование состояния пациентов до и после лечения

Диаграмма рассеяния

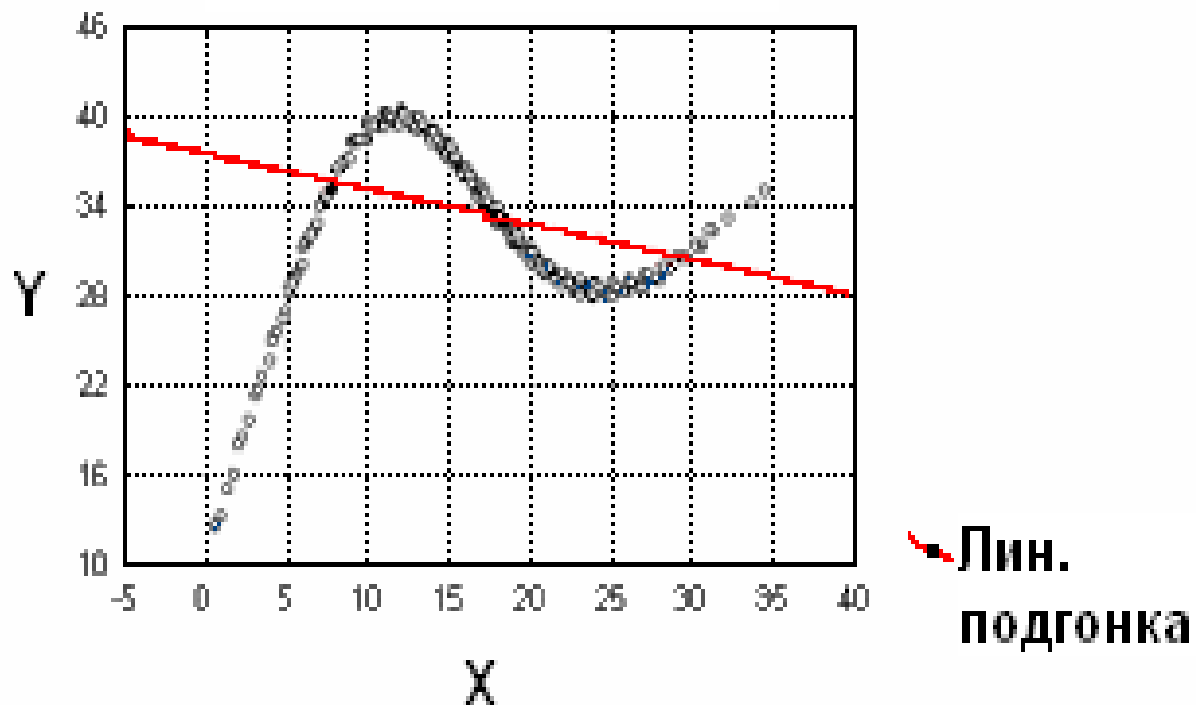


Вопрос:

что будет, если мы не учитываем наличие неоднородных групп?

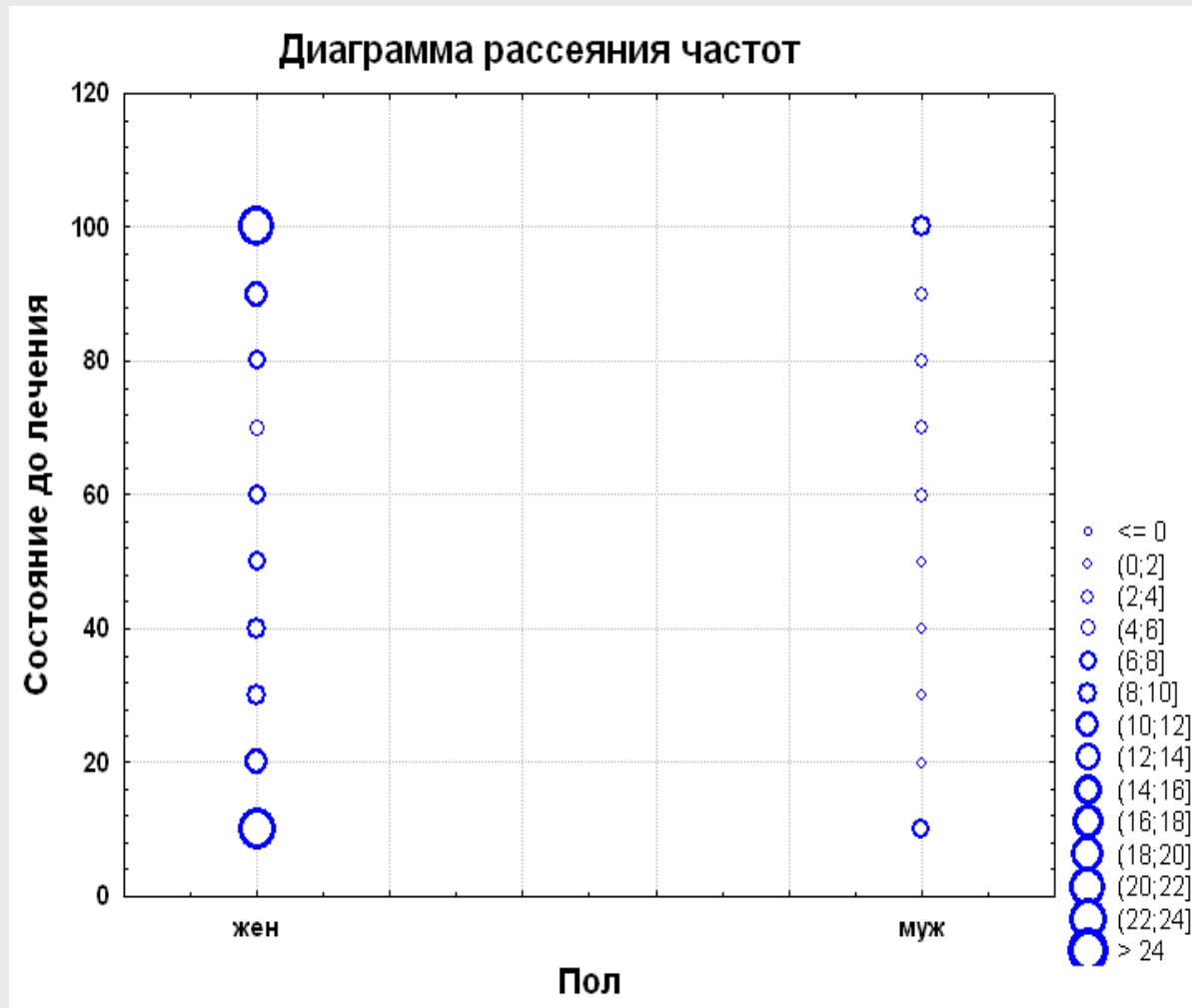
Пример: исследование состояния пациентов до и после лечения

Диаграмма рассеяния
Корреляция $r = -0.3496$



Цель:
исследовать
нелинейные
взаимосвязи между
переменными

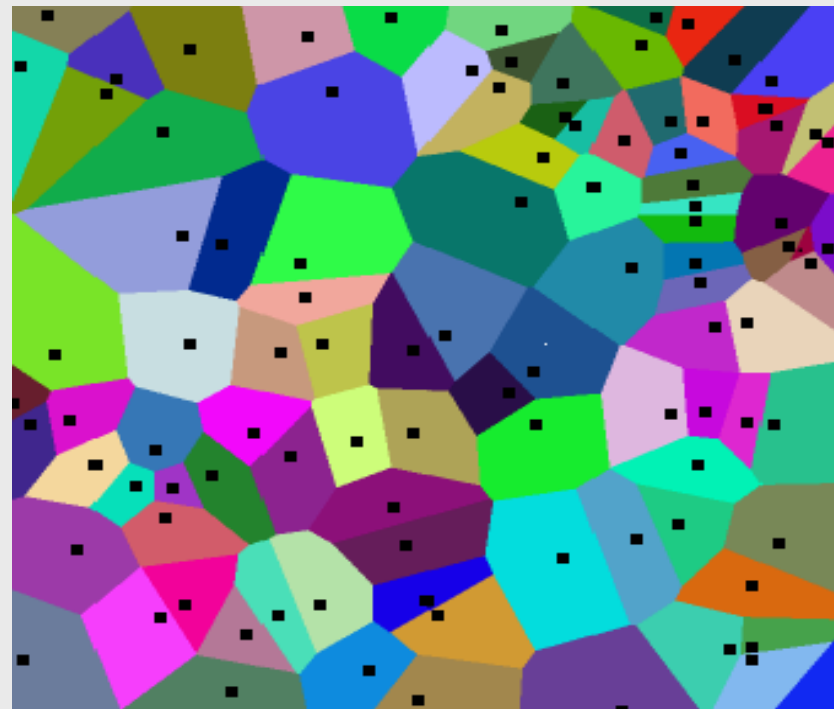
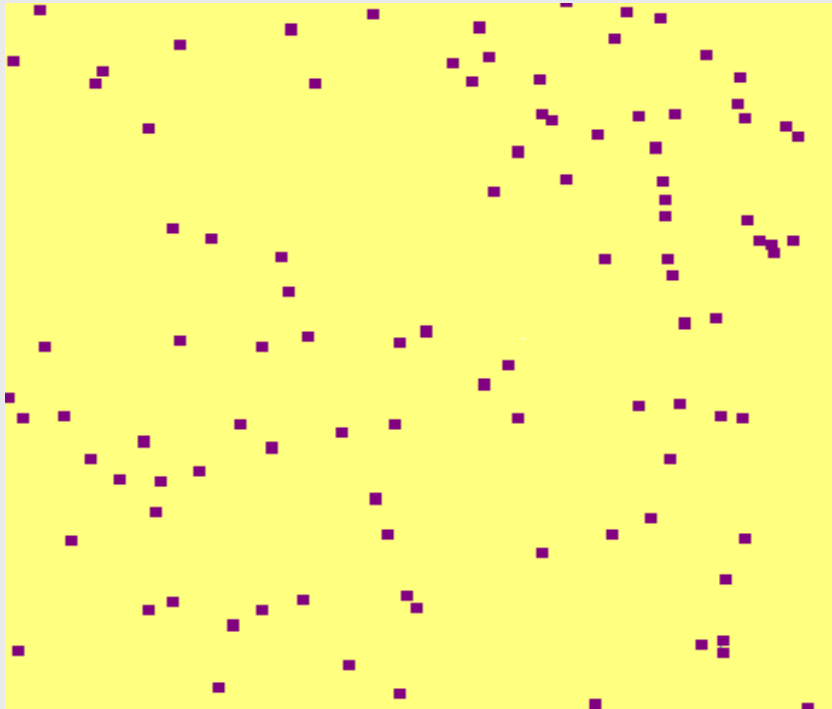
Пример: исследование состояния пациентов до и после лечения



Цель:

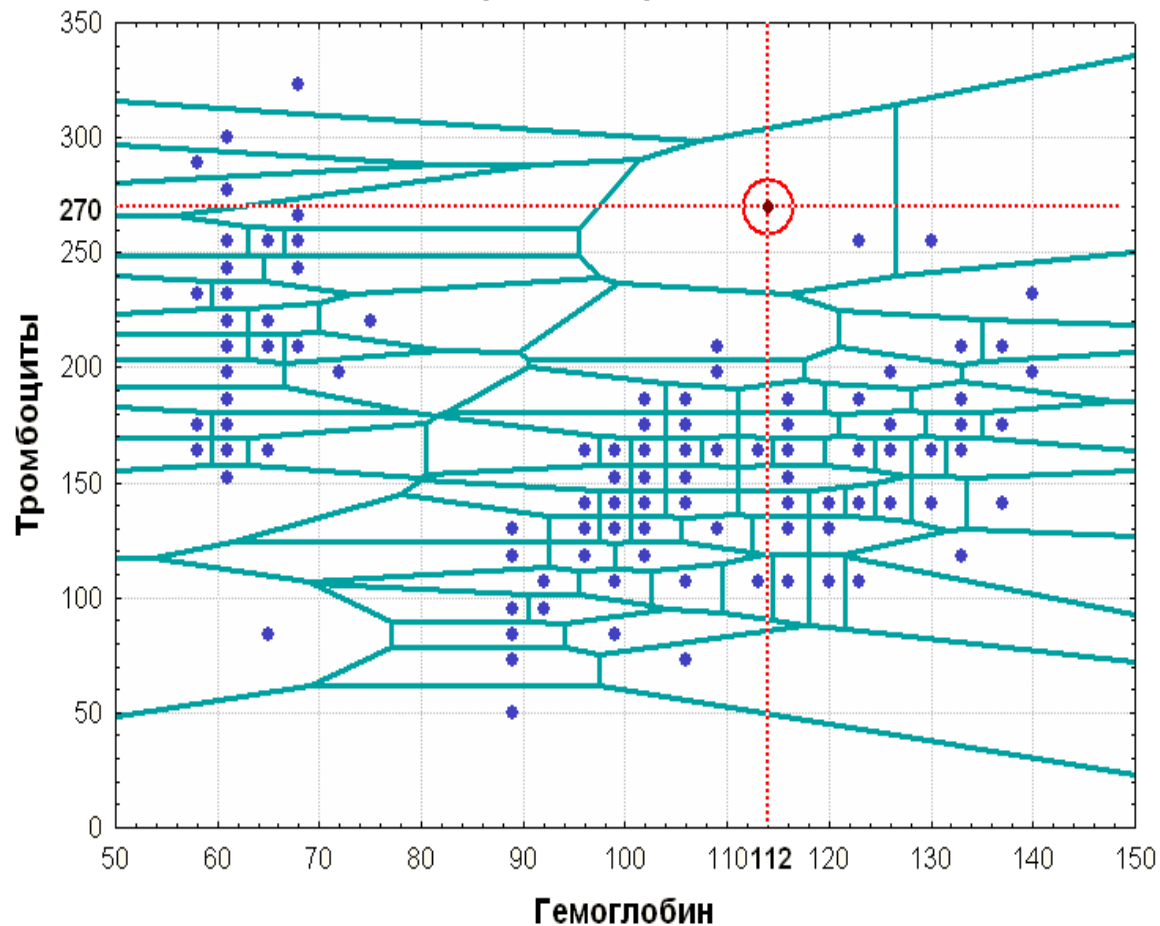
исследовать частоту появления различных состояний до лечения у пациентов женского и мужского пола

Диаграмма Вороного (*Voronoi diagram*)



Пример: исследование состояния пациентов до и после лечения

Диаграмма Вороного

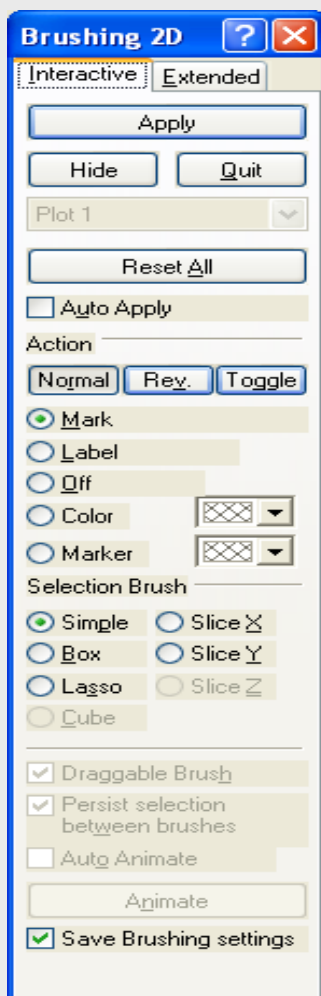


Цель:

узнать, на какого из проходивших лечение пациентов похож новый пациент, для того, чтобы назначить ему правильное лечение

Пример: исследование состояния пациентов до и после лечения

Инструмент “Кисть”

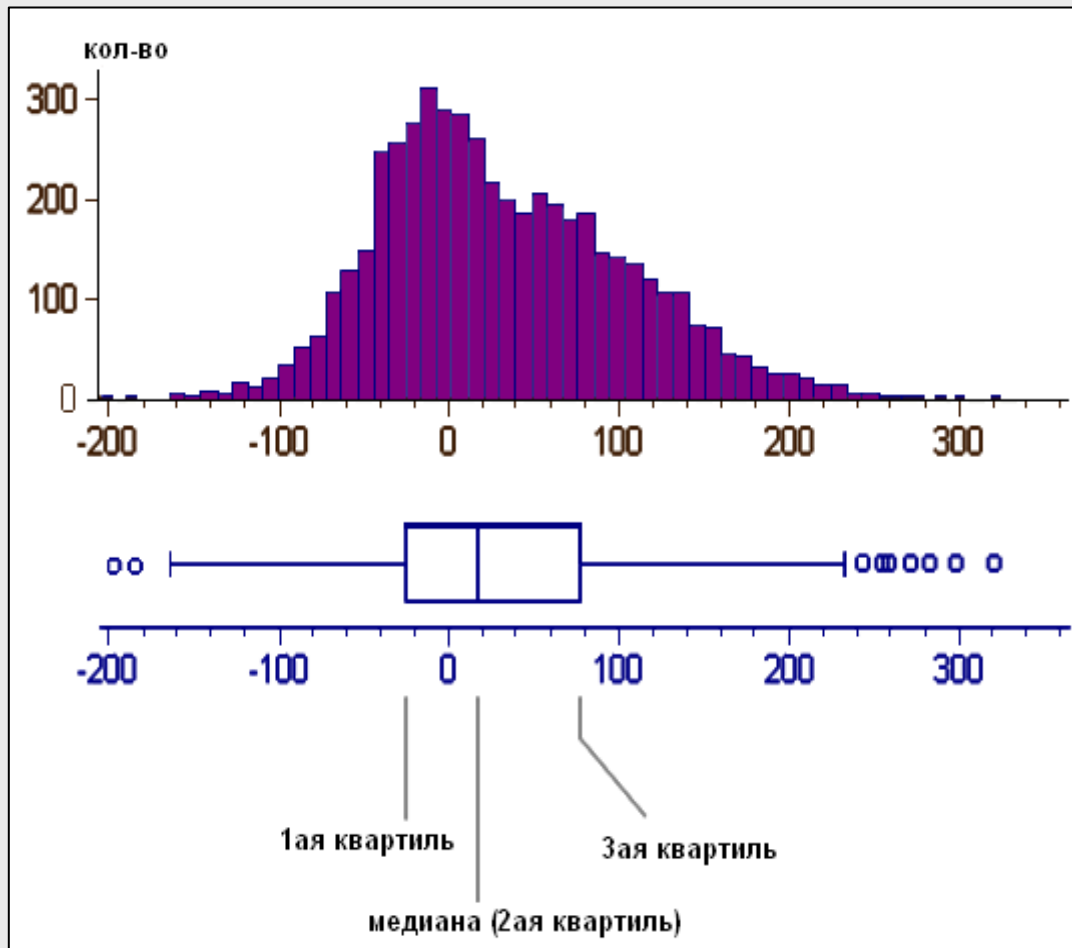


Редактор данных графика

	Scatterplot	
	Гемоглобин	Тромбоциты
пациент 41	126	198
пациент 119	99	152
пациент 69	106	152
пациент 32	116	164
пациент 23	140	198
пациент 24	126	164
! пациент 39	123	255
пациент 45	123	141
пациент 50	65	164
пациент 22	109	198
пациент 34	120	107
пациент 21	130	255
пациент 8	106	73

**выделен
“хозяин”
области**

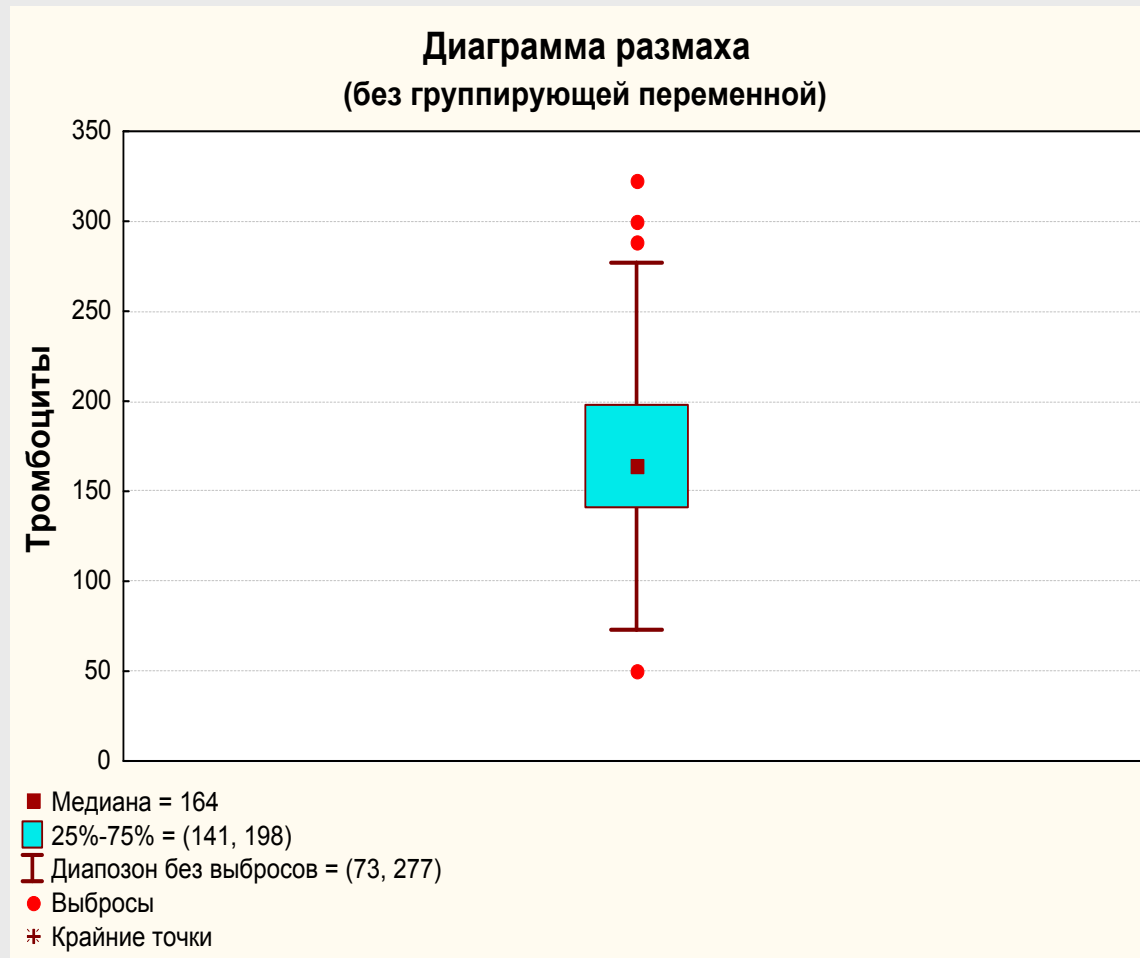
Диаграмма размаха (*Box&whisker plot*)



Основное назначение:

- *отображение диапазонов значений переменной*
- *поиск выбросов и крайних точек*

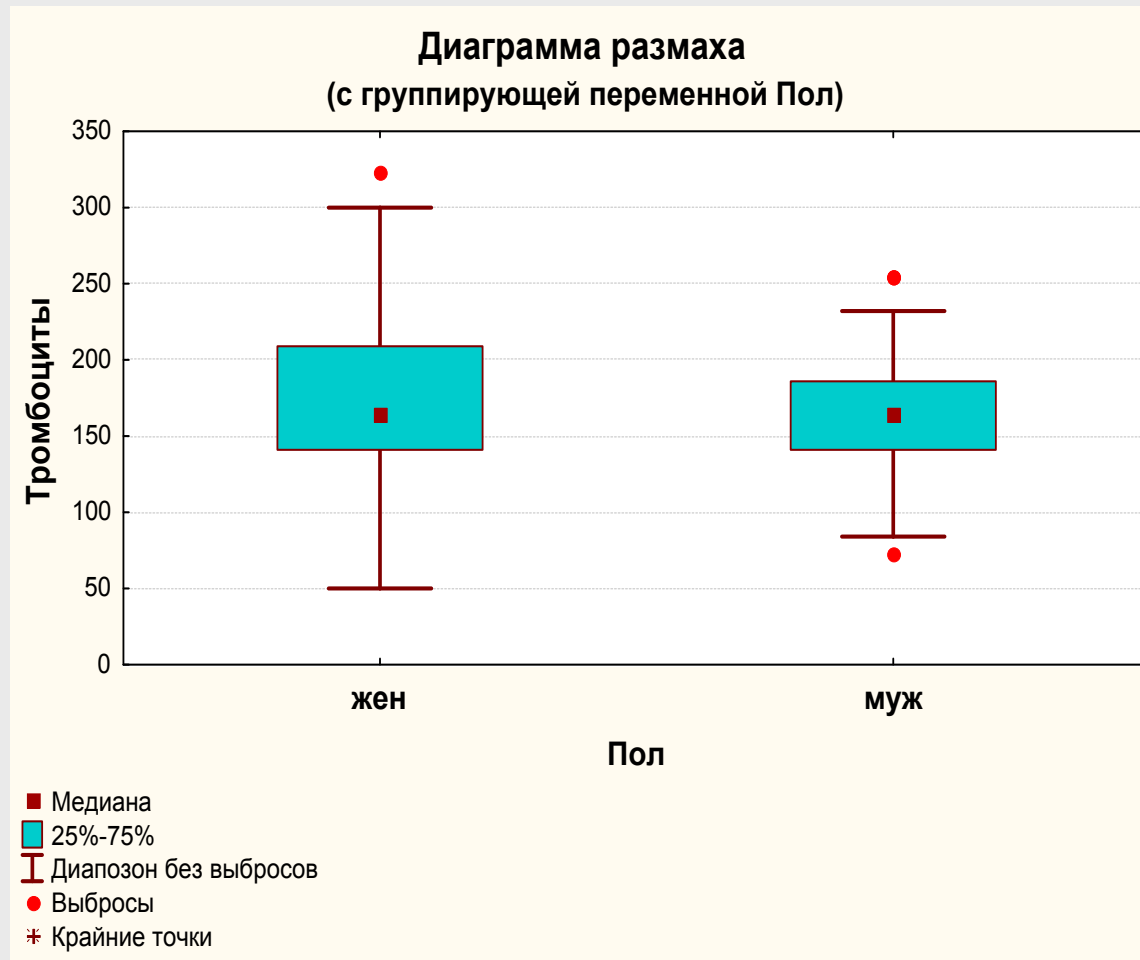
Пример: исследование состояния пациентов до и после лечения



Цель:

*проанализировать
возможный уровень
тромбоцитов в
крови*

Пример: исследование состояния пациентов до и после лечения



Цель:

сравнить уровень тромбоцитов в крови у мужчин и женщин

График средних с ошибками (Error bar plot)

Диаграмма средних с ошибками

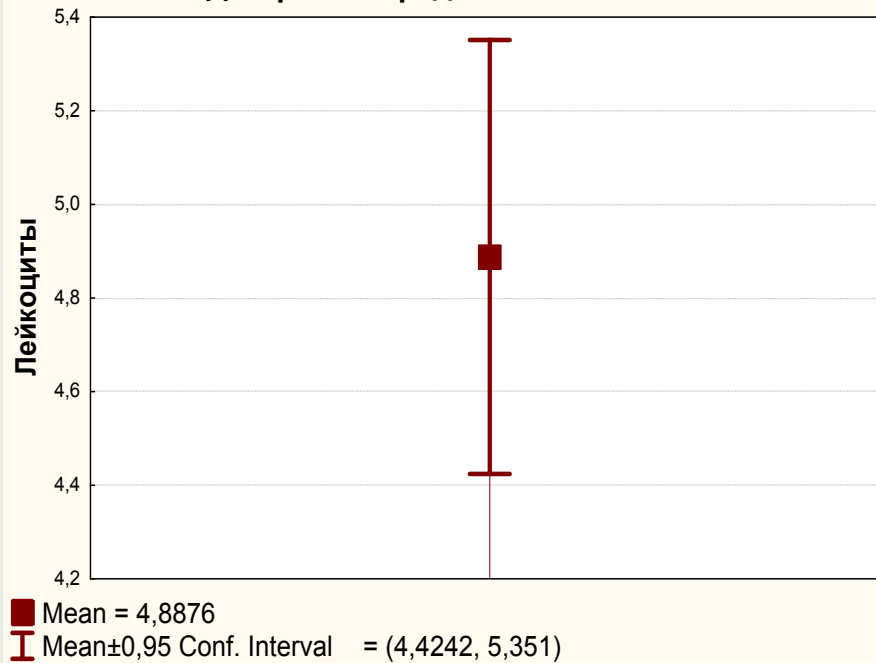


Диаграмма средних с ошибками
(столбчатая диаграмма)

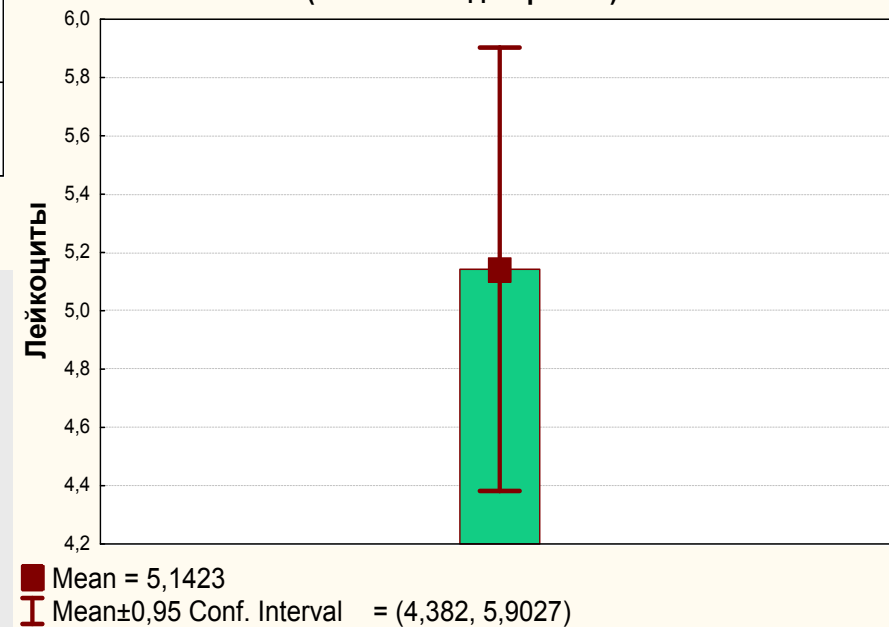
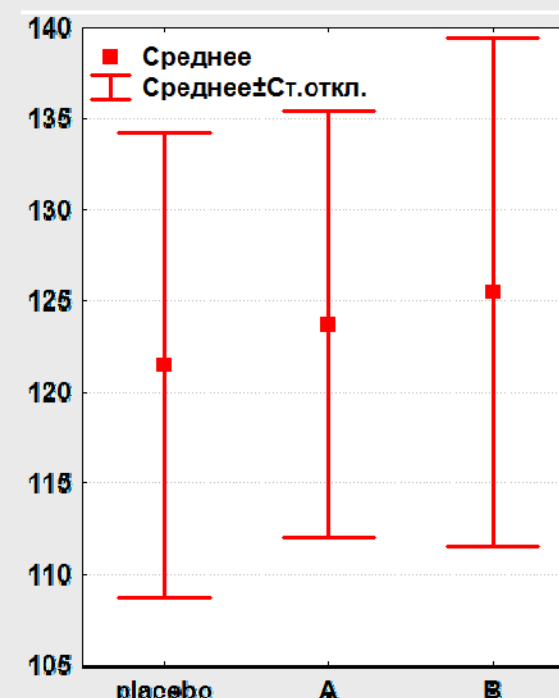
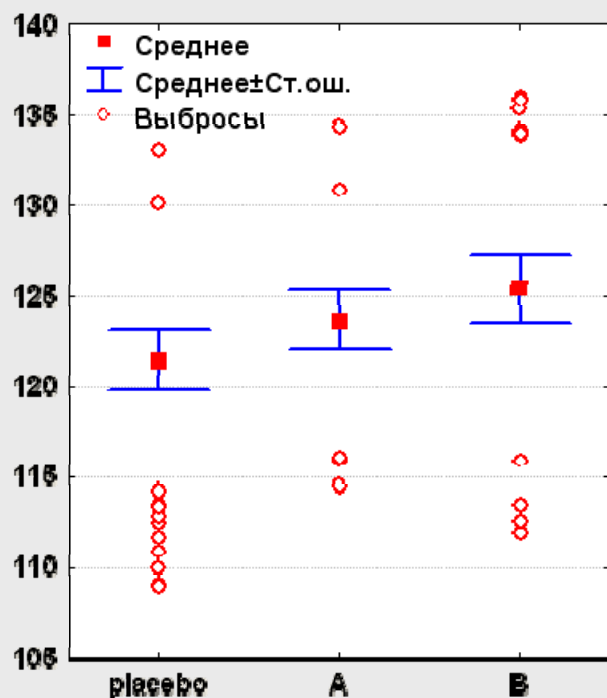


График средних с ошибками (*Error bar plot*) (продолжение)

Вертикальные штрихи:

- **стандартная ошибка среднего (SEM)**
- **доверительный интервал для среднего (CI)**
- **стандартное отклонение (SD)**

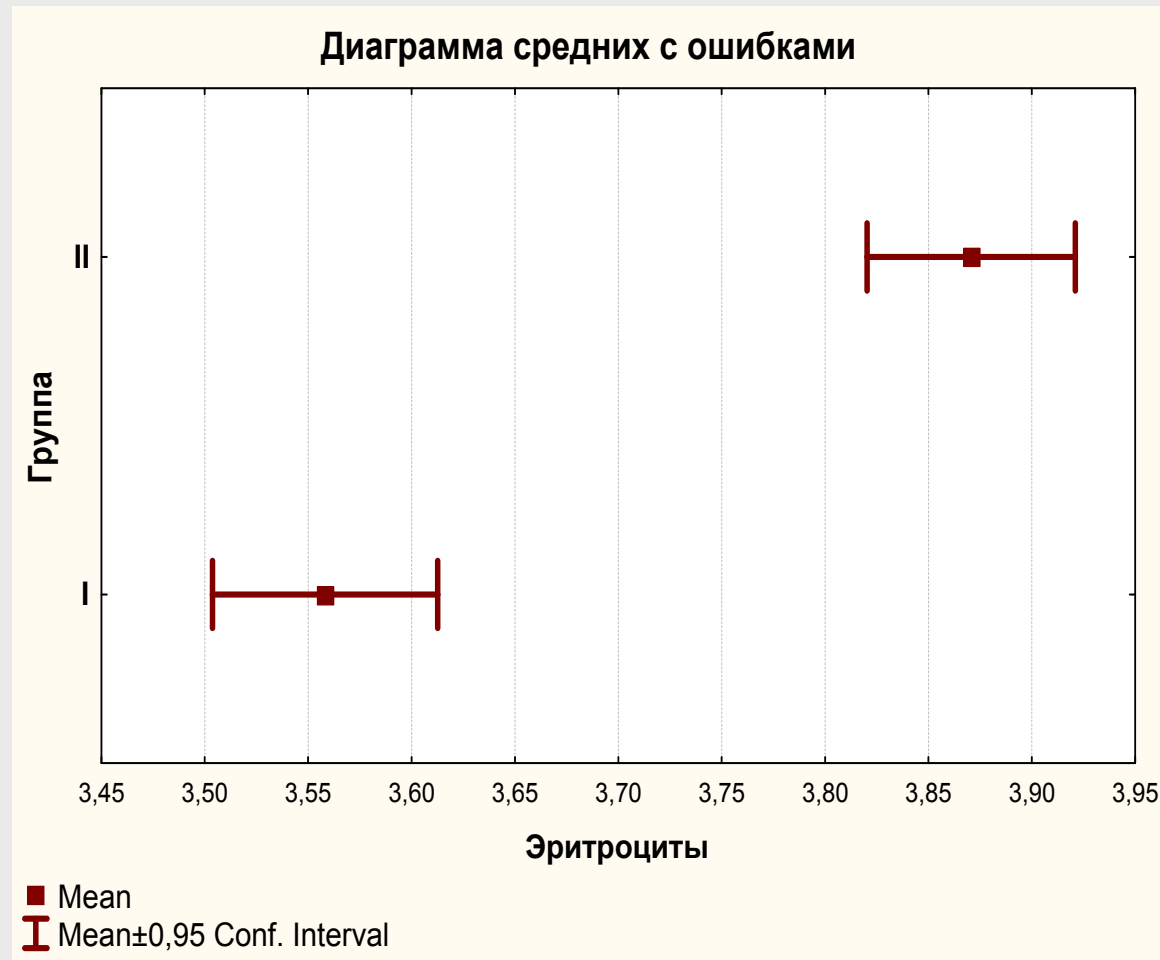


Пример: исследование состояния пациентов до и после лечения



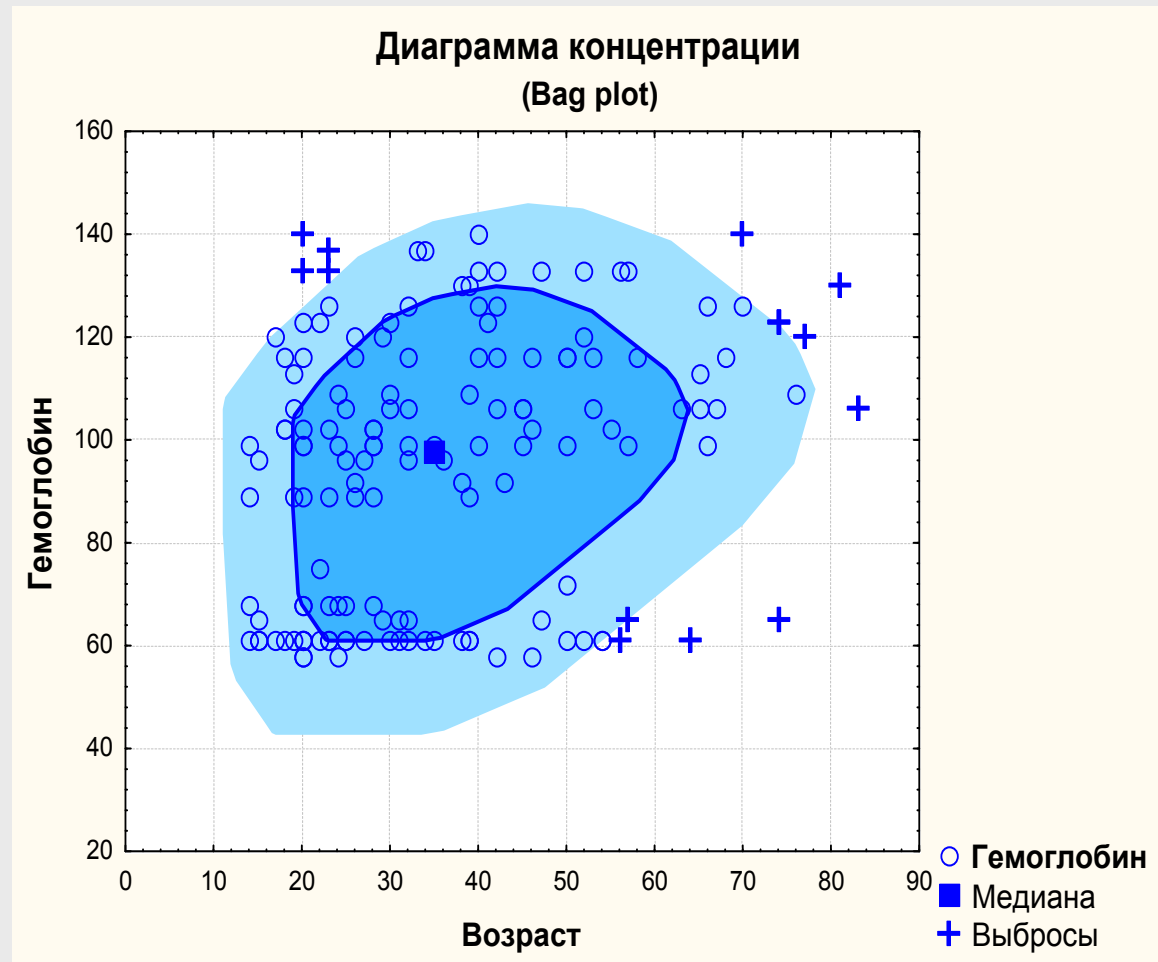
Цель:
*исследовать
изменение уровня
эритроцитов в
крови в
зависимости от
возраста пациента*

Пример: исследование состояния пациентов до и после лечения



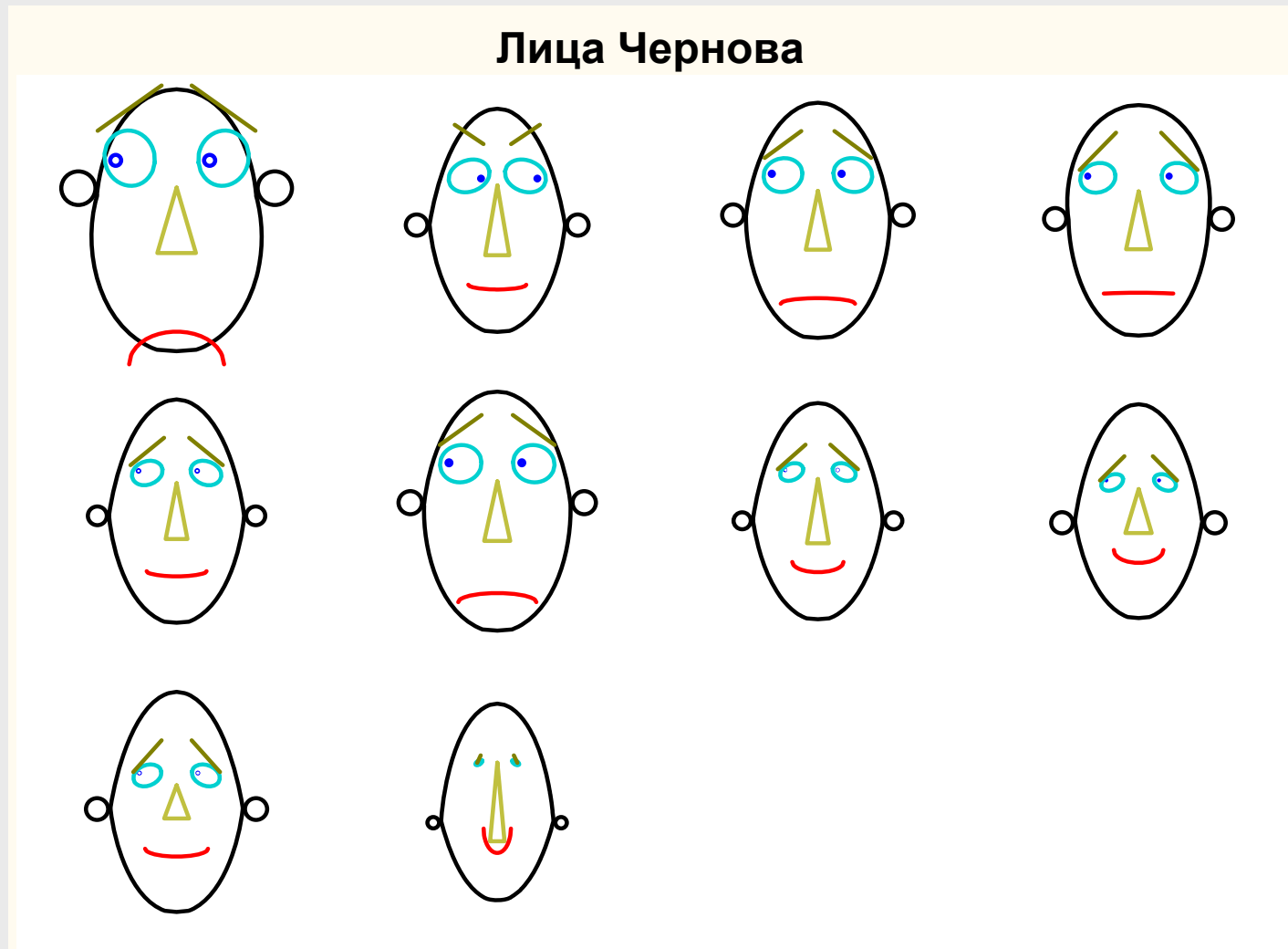
Цель:
**исследовать
изменение уровня
эритроцитов в
разных группах
пациентов**

Диаграмма концентрации (*Bag plot*)



**Основное
назначение:**
***поиск выбросов по
двум признакам***

Лица Чернова (*Chernoff faces*)

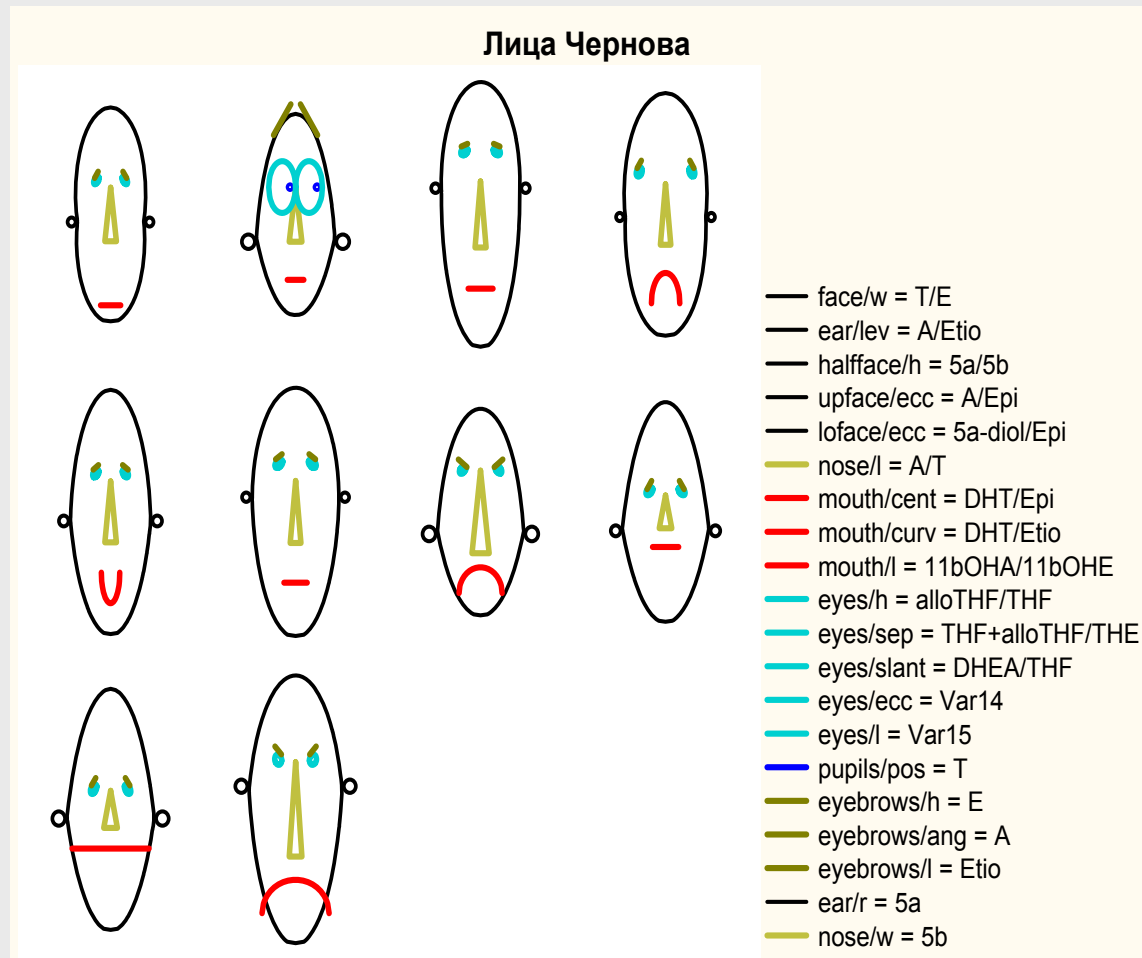


Применение лиц Чернова

- *наглядная визуализация результатов*
- *выявление характерных зависимостей или групп наблюдений*
- *исследование предположительно сложных взаимосвязей между несколькими переменными*

Пример: антидопинговые исследования

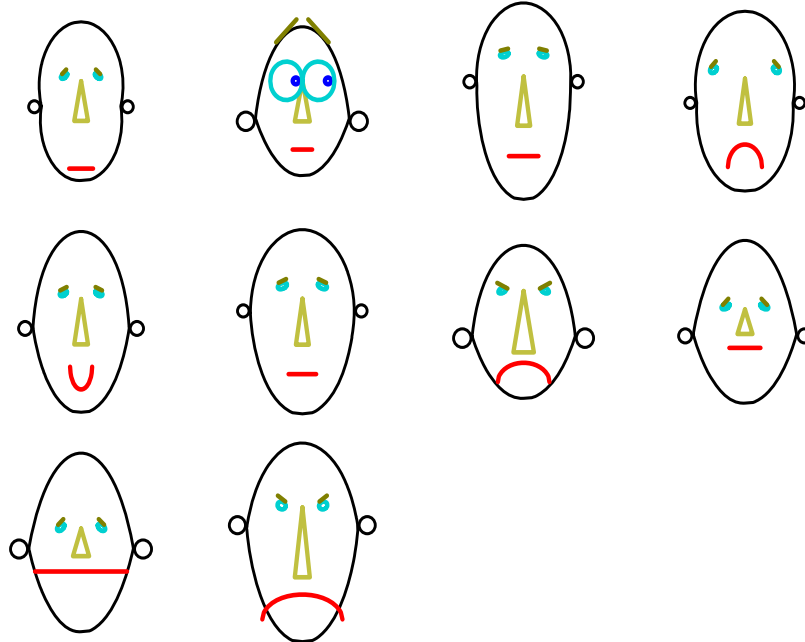
Лица Чернова



Цель:
сравнить
показатели у разных
пациентов

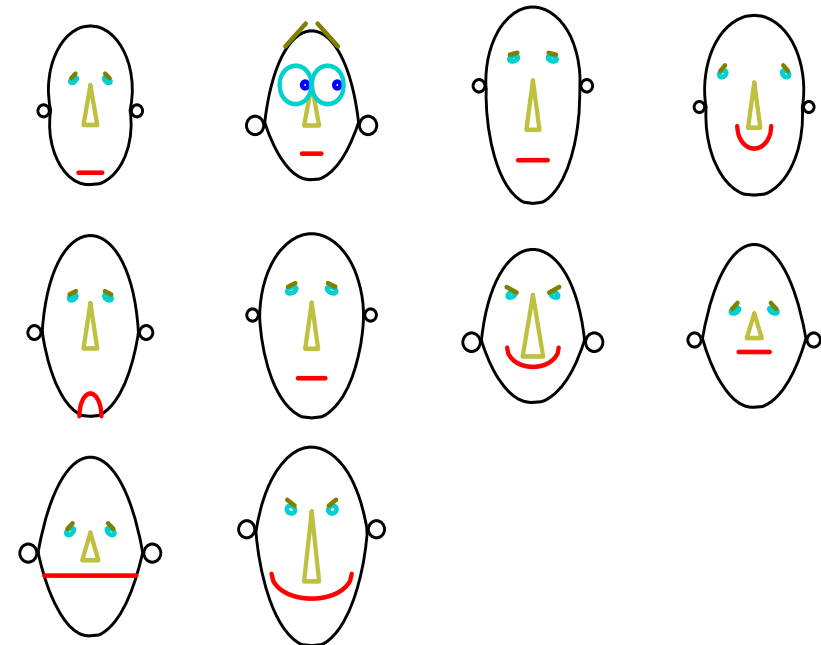
Пример: антидопинговые исследования

Лица Чернова



Использование лиц Чернова для визуализации результатов

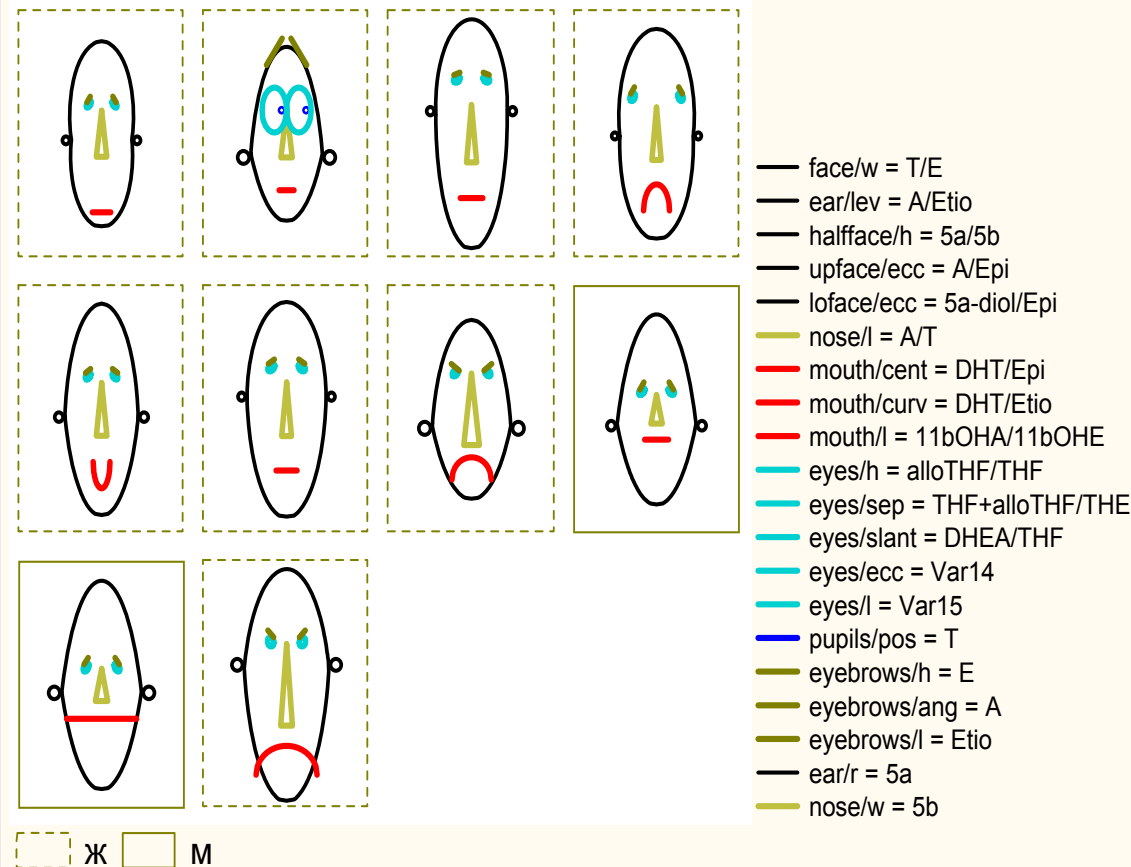
Лица Чернова



- **превышению норм соответствует грустное лицо**
- **соблюдению норм -- улыбающееся или нейтральное лицо**

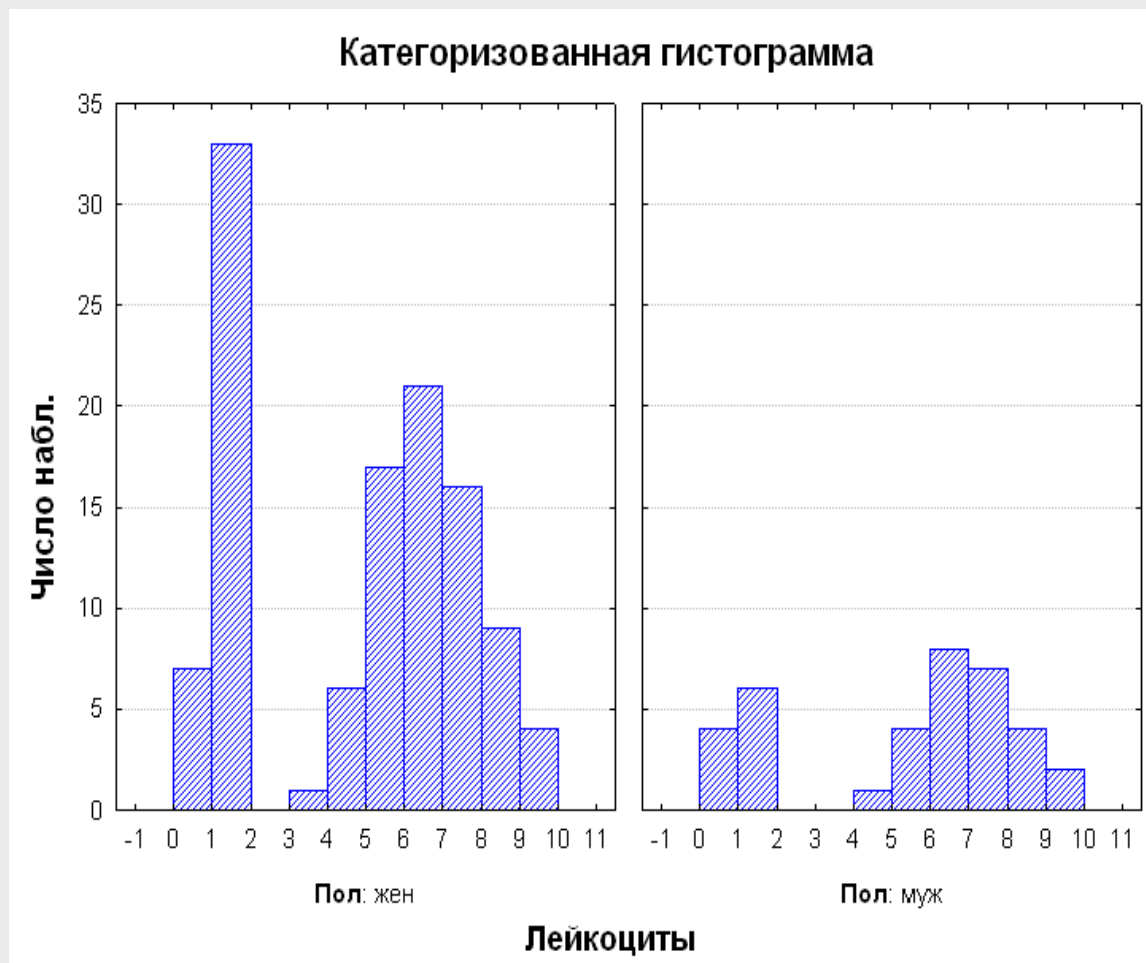
Пример: антидопинговые исследования

Лица Чернова (выделение пола)



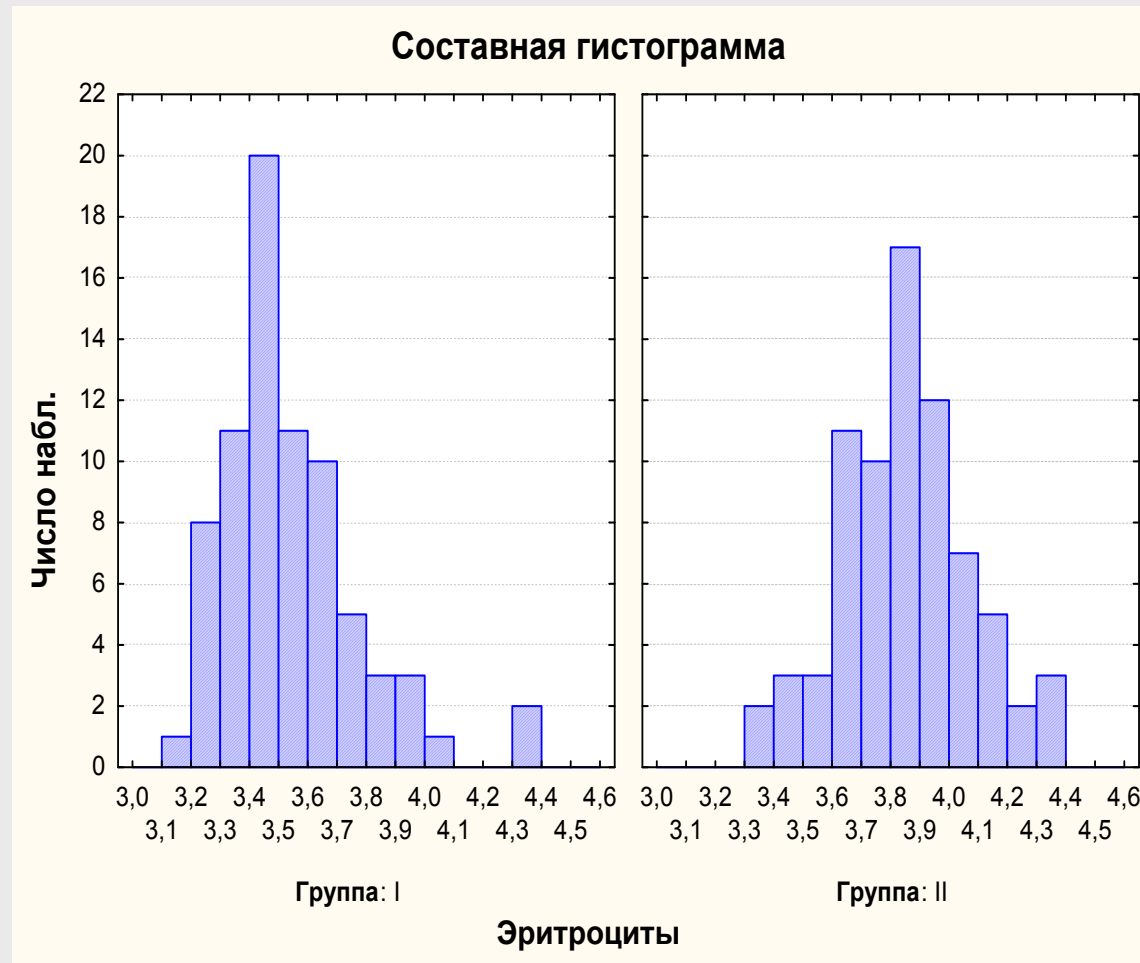
Цель:
 сравнить
 показатели у
 пациентов мужского
 и женского полов

Категоризованные графики



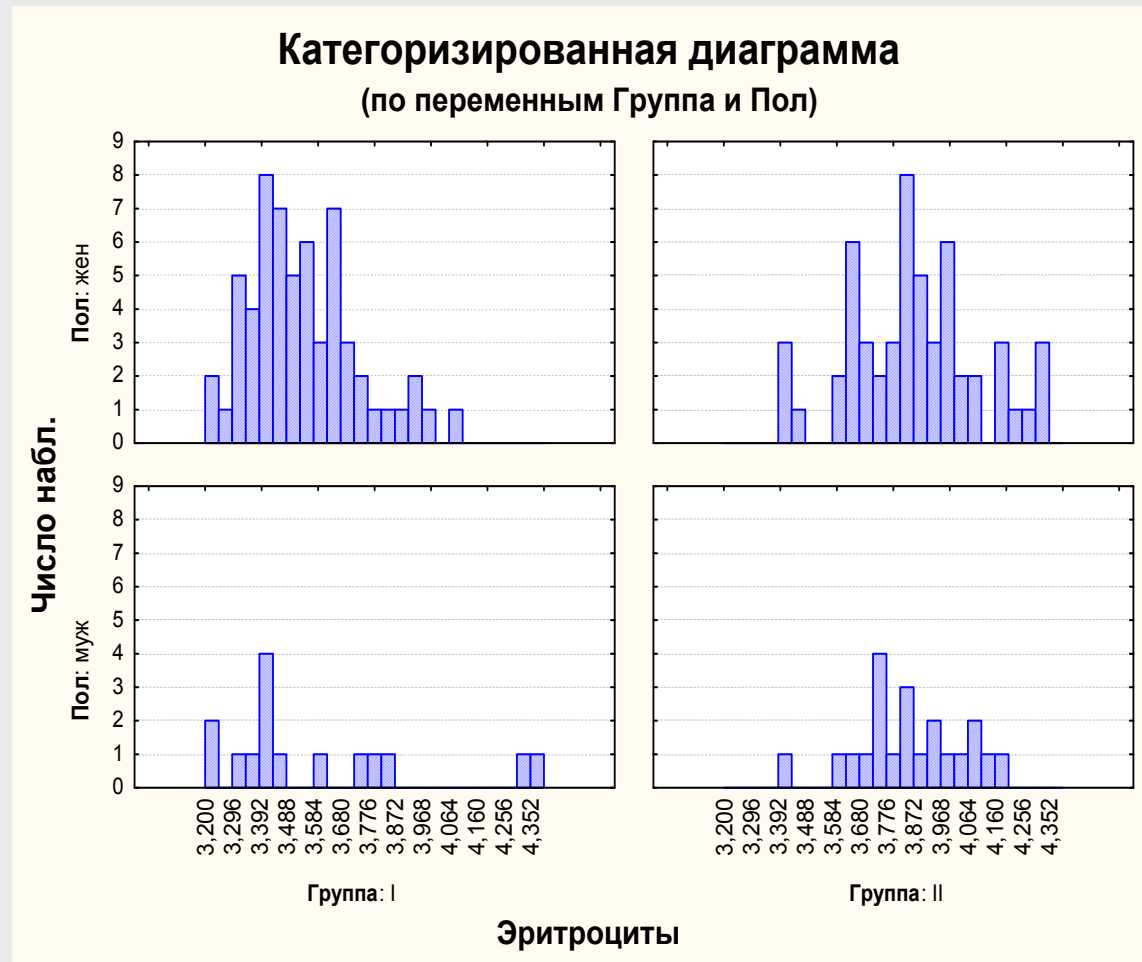
Основное назначение:
визуализация данных,
разбитых на группы, для
более точного и
детального анализа

Категоризованные графики (гистограмма)



Цель:
*сравнить распределение
уровня эритроцитов в
зависимости от
принадлежности
пациента к группе
заболевания*

Пример: исследование препаратов, влияющих на уровень гемоглобина



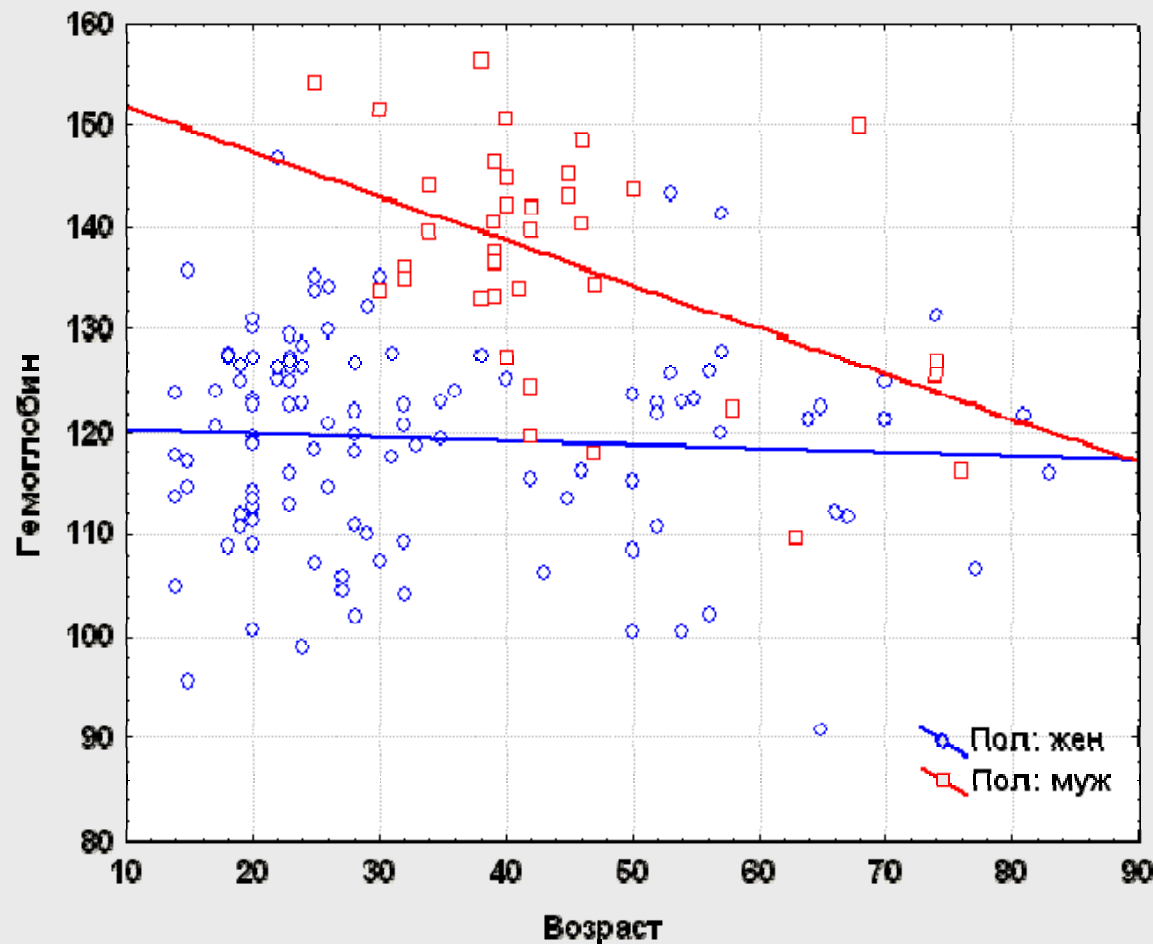
Цель:

сравнить распределение уровня эритроцитов в зависимости от принадлежности пациента к группе заболевания и от его пола

Пример: исследование препаратов, влияющих на уровень гемоглобина

Категоризованная диаграмма рассеяния

(Categorized scatter plot)

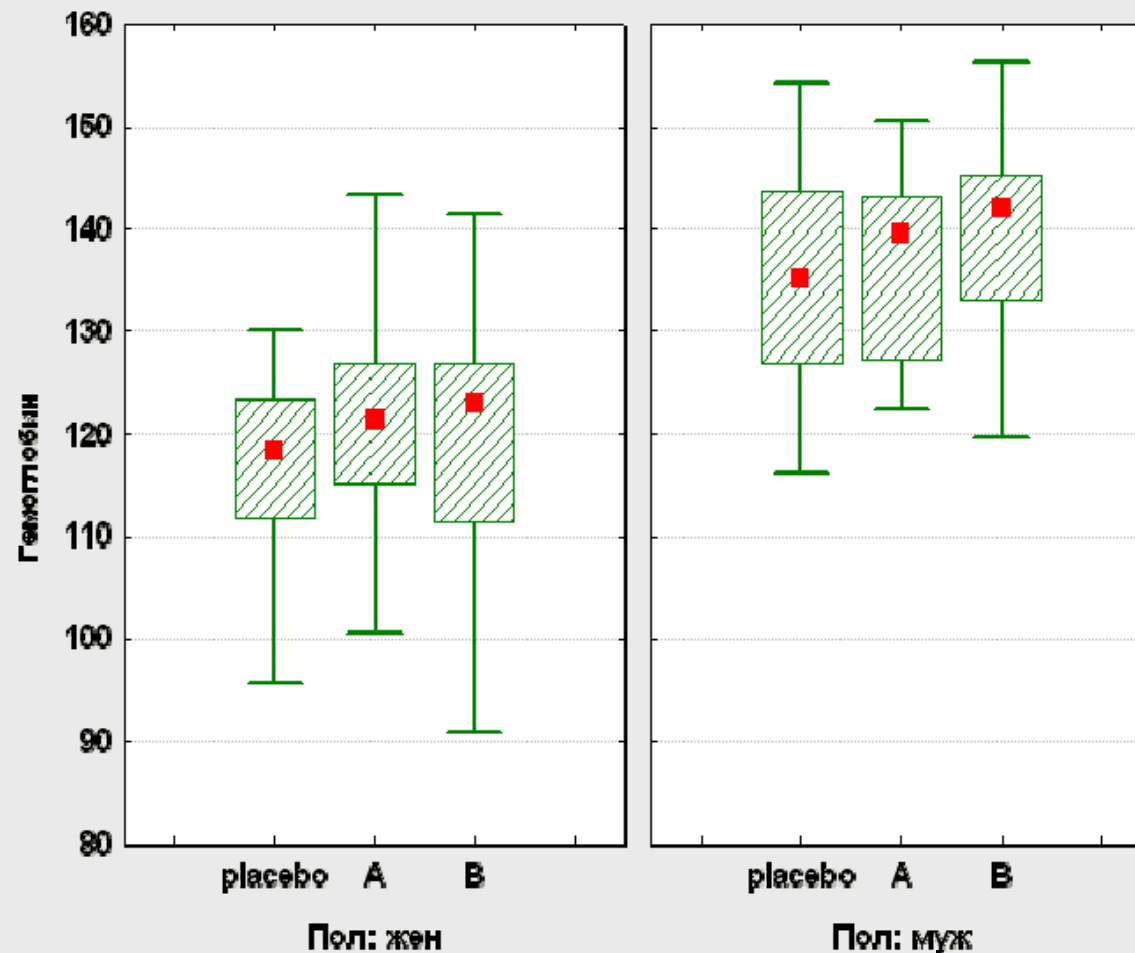


Цель: сравнить зависимости гемоглобина от возраста для мужчин и для женщин

Пример: исследование препаратов, влияющих на уровень гемоглобина

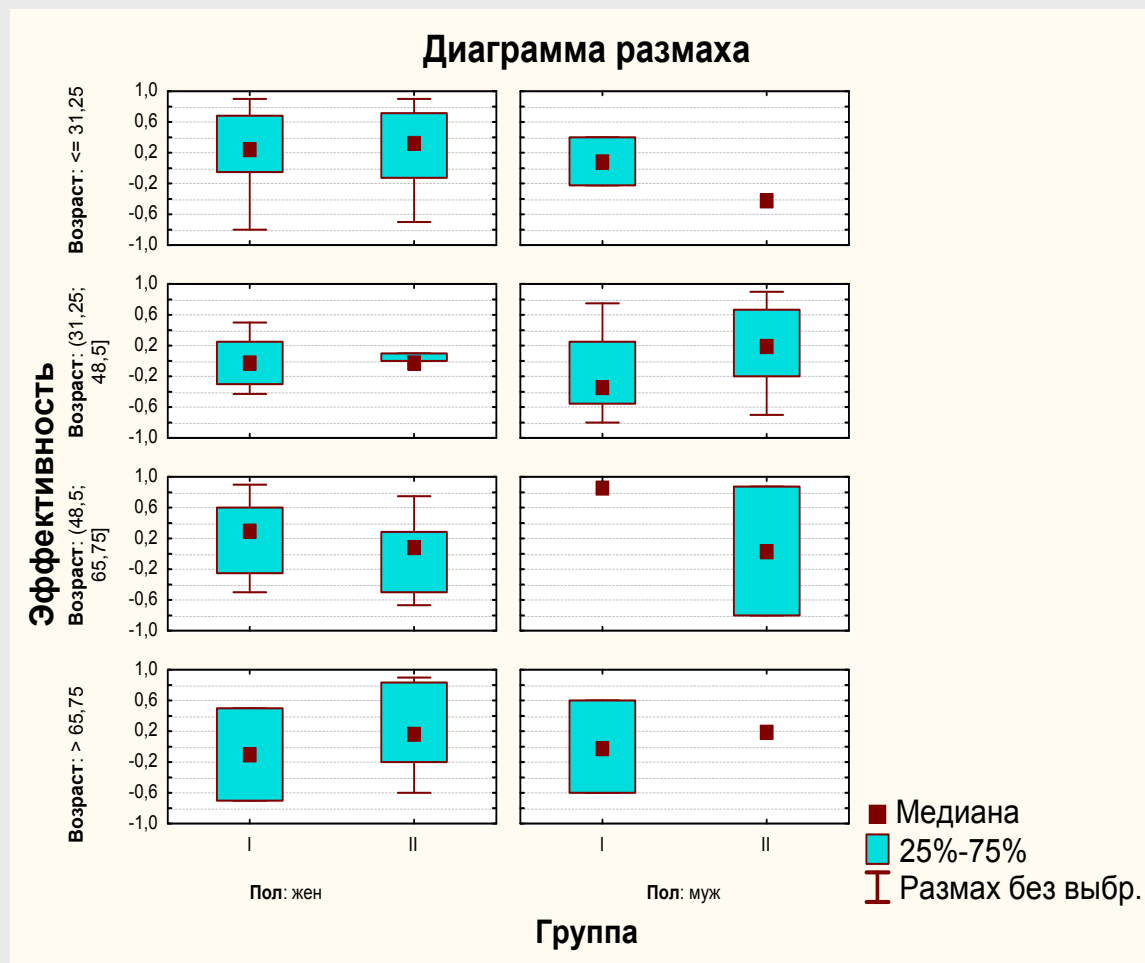
Категоризованная диаграмма размаха

(Categorized Box&Whisker plot)



Цель: сравнить действие исследуемых препаратов для мужчин и женщин

Пример: исследование препаратов, влияющих на уровень гемоглобина

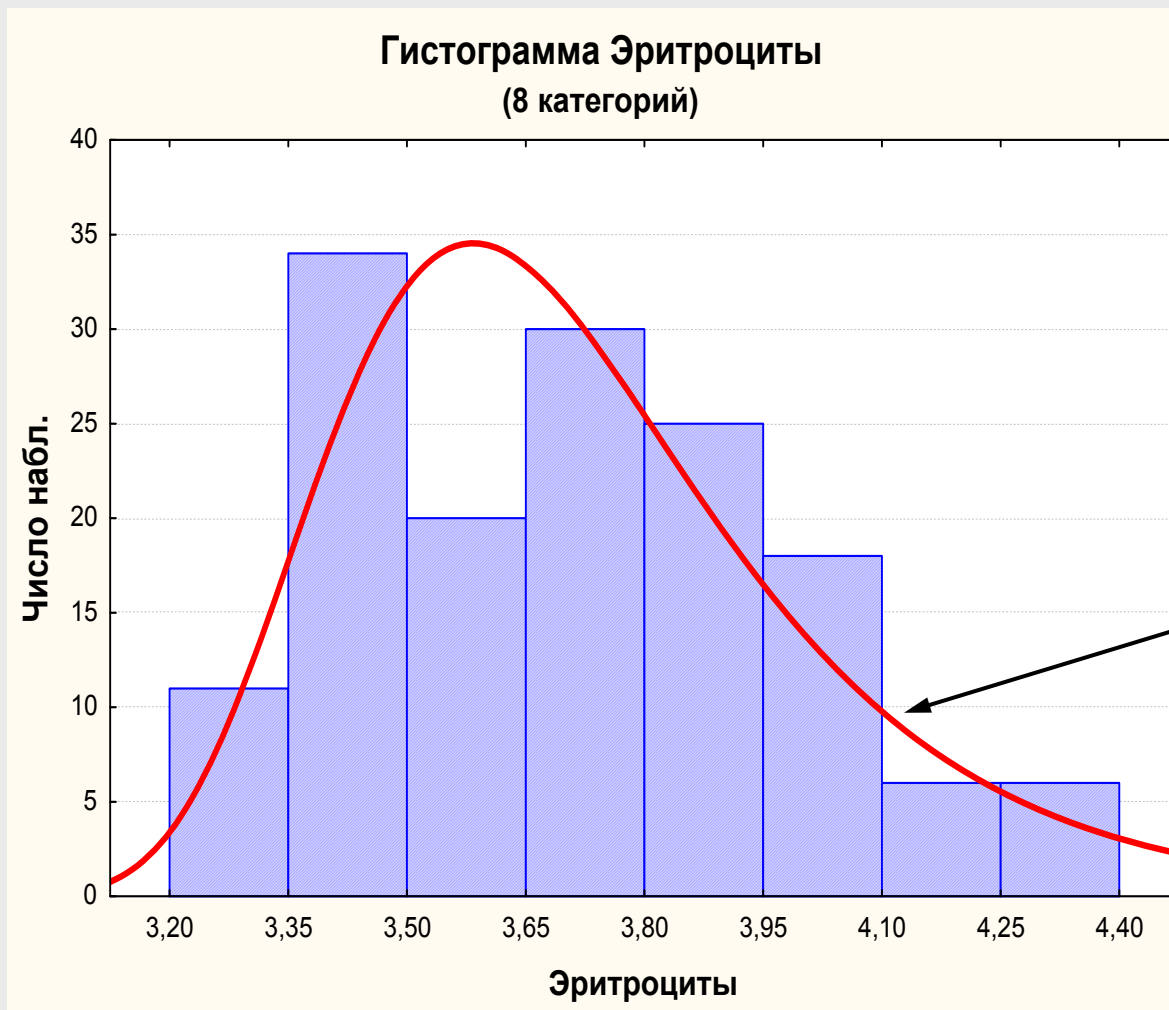


Цель:
сравнить эффективность лечения в зависимости от принадлежности пациента к группе заболевания, категоризованные по полу и возрасту

Рекомендации по графическому представлению данных в отчете

- *заголовок графика должен быть кратким, информативным и недвусмысленным*
- *на осях, сегментах и решетках должны быть пометки; объясняйте значение символов*
- *постарайтесь уместить на графике всю нужную информацию*
- *делайте график простым, отказывайтесь от ненужных украшений (например, для диаграммы не обязательно делать столбцы трехмерными)*

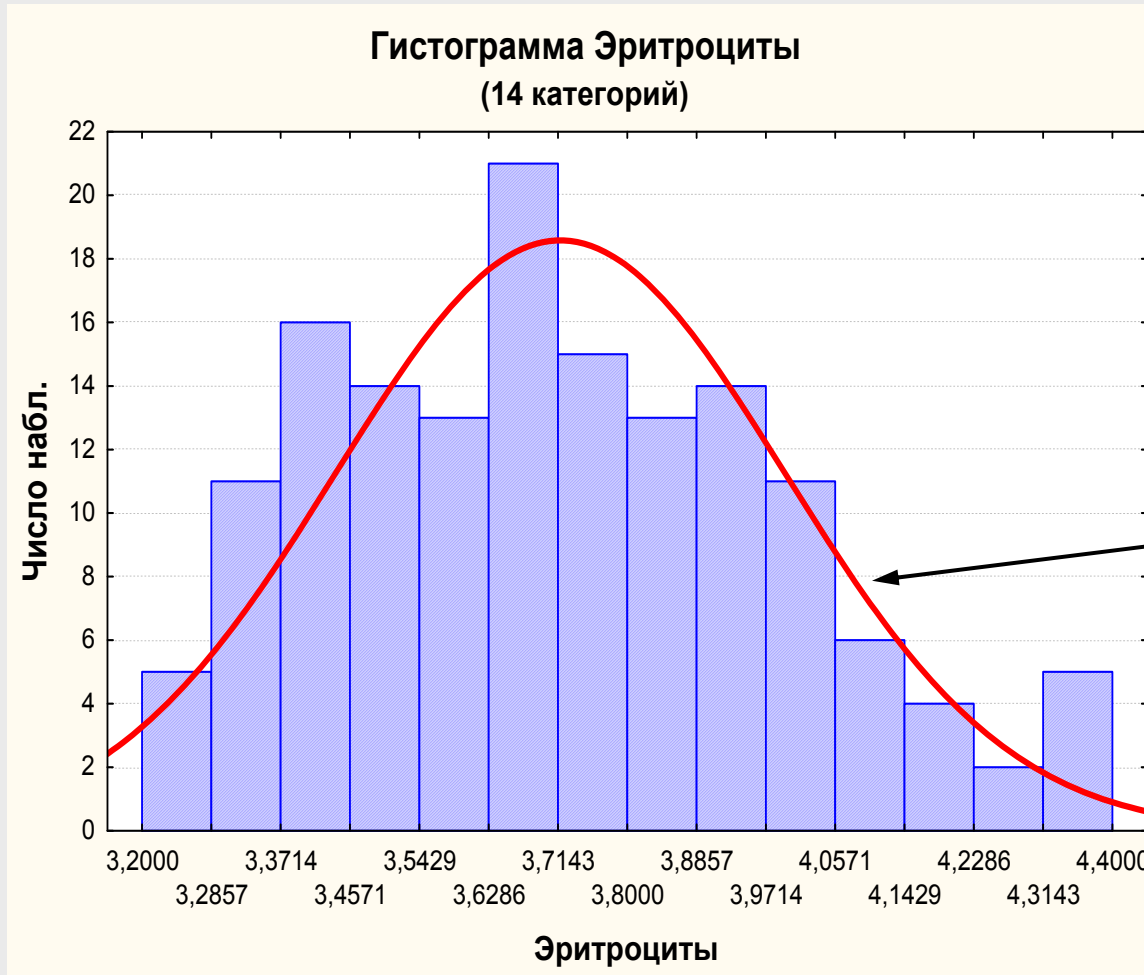
Пример неудачной визуализации данных: гистограмма



Цель:
*по гистограмме
определить вид
распределения
переменной*

**отрицательно
экспоненциальное
распр.**

Пример удачной визуализации данных: гистограмма

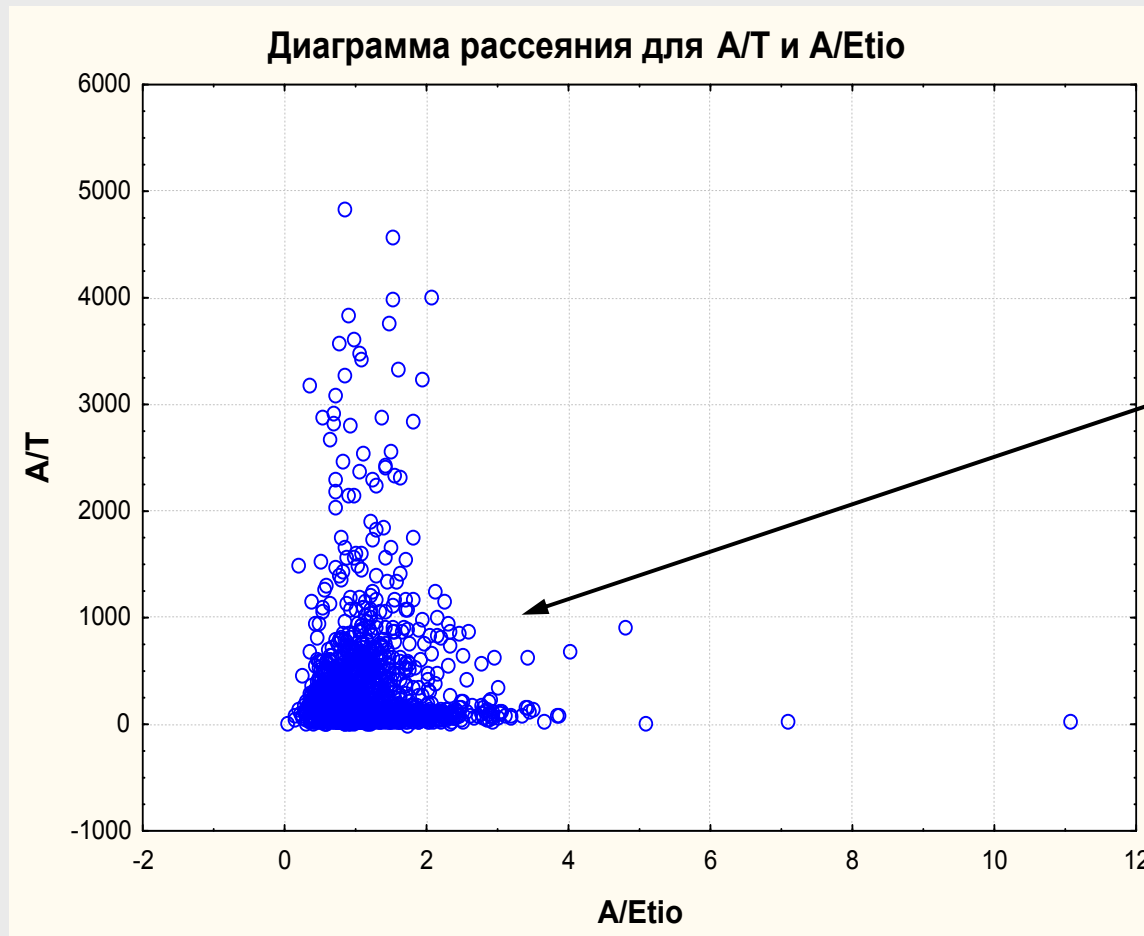


Цель:

*по гистограмме
определить вид
распределения
переменной*

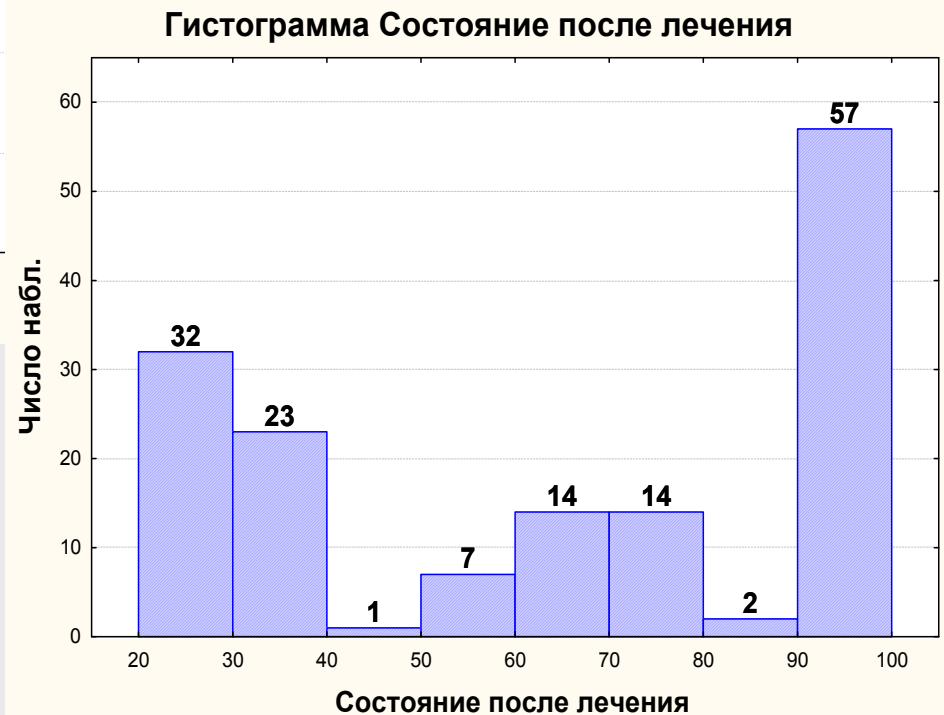
нормальное распр.

Пример неудачной визуализации данных: диаграмма рассеяния

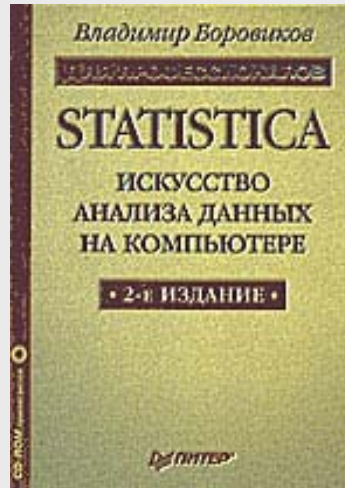


**слишком
много точек**

Пример неудачной визуализации данных: диапазон осей



Литература

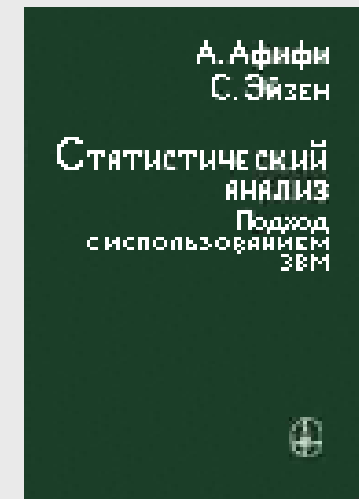
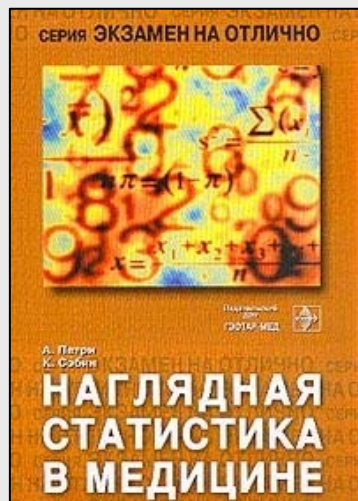
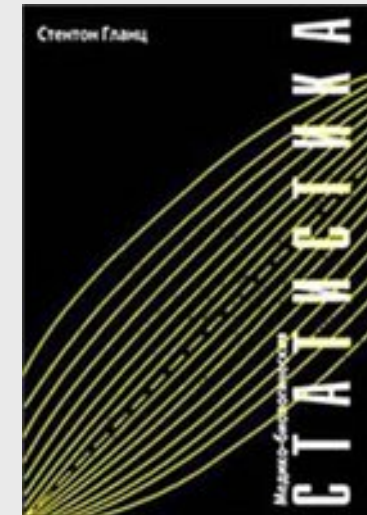


Боровиков В.П.
STATISTICA: искусство анализа данных на компьютере

А. Петри, К. Сэбин
Наглядная медицинская статистика

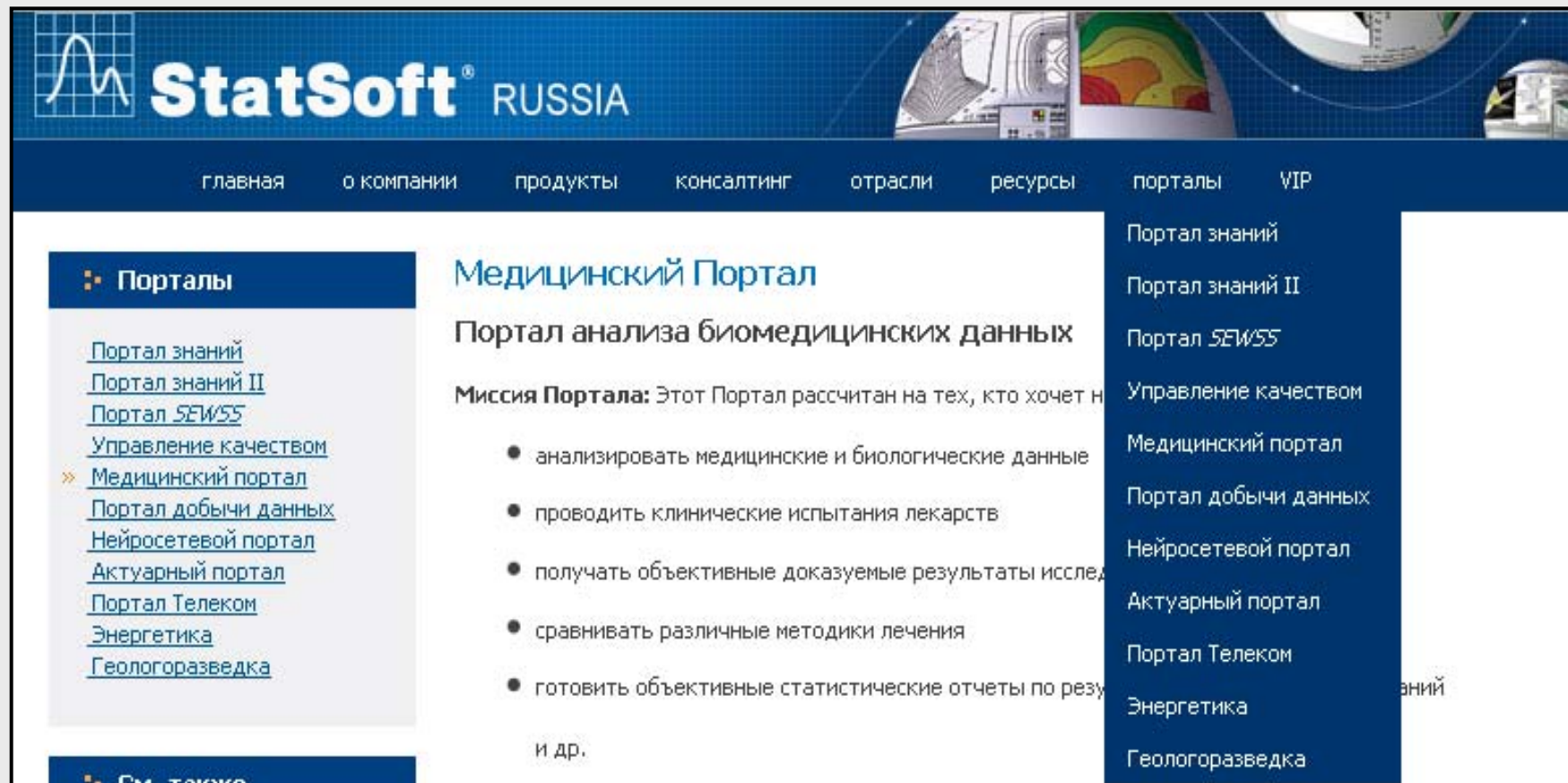
Стентон Гланц
Медико-биологическая статистика

А. Аффифи, С. Эйзен
Статистический анализ. Подход с использованием ЭВМ



Медицинский портал StatSoft

<http://www.statsoft.ru> → порталы



The screenshot shows the StatSoft RUSSIA website interface. The top navigation bar includes links for 'главная', 'о компании', 'продукты', 'консалтинг', 'отрасли', 'ресурсы', 'порталы', and 'VIP'. The 'порталы' menu is expanded, showing a list of portals: 'Портал знаний', 'Портал знаний II', 'Портал SEW/SS', 'Управление качеством', 'Медицинский портал', 'Портал добычи данных', 'Нейросетевой портал', 'Актуарный портал', 'Портал Телеком', 'Энергетика', and 'Геологоразведка'. The 'Медицинский портал' is highlighted in blue. The main content area for the 'Медицинский Портал' is visible, featuring the title 'Медицинский Портал' and the subtitle 'Портал анализа биомедицинских данных'. Below this, the mission statement reads: 'Миссия Портала: Этот Портал рассчитан на тех, кто хочет...' followed by a list of bullet points: 'анализировать медицинские и биологические данные', 'проводить клинические испытания лекарств', 'получать объективные доказуемые результаты исследований', 'сравнивать различные методики лечения', and 'готовить объективные статистические отчеты по результатам исследований и др.'

Медицинский портал StatSoft

[Основные термины и понятия медицинской статистики](#)

[Толковый словарь англоязычных терминов медицинской статистики](#)

[Определения доказательной медицины](#) >> **new**

[Исходные данные](#)

[Импорт из Excel в STATISTICA](#)

[Визуальный анализ](#)

[Описательный \(дескриптивный\) анализ](#)

[Анализ выбросов](#)

[Корреляционный анализ](#)

[Исследование эффективности лечения: введение новой переменной](#)

[Проверка гипотезы о нормальности](#)

[Сравнение выборок](#)

[Рандомизация в клинических испытаниях](#)

[Проведение рандомизации в STATISTICA](#)

[Двухфакторный дисперсионный анализ](#)

[Мета-анализ](#)

[Основные типы планов](#)

[Статистический медицинский советник](#)

Примеры:

[t-скор и z-скор в диагностике остеопороза](#) >> **new**

[Пример исследования гемограмм пациентов](#)

[Классическое Фрэнкингемское исследование](#)

[Частота рекомбинаций хромосом домашней мыши относительно некоторого признака](#)

Обучающие курсы по медицине

<http://www.statsoft.ru> → консалтинг → обучение → медицина

- **STATISTICA для медицинских приложений. Вводный курс**
(8 *акад. часов*)
- **STATISTICA для медицинских приложений. Углубленный курс**
(10 *акад. часов*)
- **Лекционный курс по доказательной медицине и статистическим методам в клинических исследованиях**
(12 *акад. часов*)
- **Современные статистические принципы клинических испытаний лекарств**
(10 *акад. часов*)

Звоните: 8-(495) 787-77-33, 916-03-93

Пишите: sales@statsoft.ru

Лекционный курс по доказательной медицине и статистическим методам в клинических исследованиях

Консалтинг

- » [Обучение](#)
- » [Проекты](#)
- » [Семинары](#)
- » [Техническая поддержка](#)

См. также

- » [Формы обучения](#)
- » [Книги по STATISTICA](#)
- » www.statistica.ru
- » www.spc-consulting.ru
- » [Data Mining & Scoring](#)

Лекционный курс по доказательной медицине и статистическим методам в клинических исследованиях

Курс адресован врачам-исследователям, аспирантам биомедицинских дисциплин, работникам фармацевтической промышленности – тем, кто неизбежно сталкивается с чтением публикаций. Как правило, наиболее важные статьи по клиническим исследованиям лекарств опубликованы в журналах на английском языке, поэтому в ходе курса основные термины дублируются англоязычными вариантами.

Основной упор делается на системное изложение современных методов анализа биомедицинских данных. Курс даёт навыки критического восприятия статей и научных работ врачей-исследователей. Это уникальная возможность повысить свой аналитический уровень и познакомиться с методами современной медицинской статистики за 3 дня.

В программе курса мы следуем рекомендациям International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use ([ICH](#)).

Предлагаемый курс включает в себя следующие темы:

Статистические принципы доказательной медицины >>new

Курсы Академии Анализа StatSoft - самый эффективный способ освоить современные технологии анализа данных

Спасибо за внимание!



U.S. Headquarters: StatSoft, Inc. • 2300 E. 14th St. • Tulsa, OK 74104 • USA • (918) 749-1119 • Fax: (918) 749-2217 • info@statsoft.com • www.statsoft.com

Australia: StatSoft Pacific Pty Ltd.

Brazil: StatSoft South America

Bulgaria: StatSoft Bulgaria Ltd.

Czech Rep.: StatSoft Czech Rep. s.r.o.

China: StatSoft China

France: StatSoft France

Germany: StatSoft GmbH

Hungary: StatSoft Hungary Ltd.

India: StatSoft India Pvt. Ltd.

Israel: StatSoft Israel Ltd.

Italy: StatSoft Italia srl

Japan: StatSoft Japan Inc.

Korea: StatSoft Korea

Netherlands: StatSoft Benelux BV

Norway: StatSoft Norway AS

Poland: StatSoft Polska Sp. z o.o.

Portugal: StatSoft Ibérica Lda

Russia: StatSoft Russia

Spain: StatSoft Ibérica Lda

S. Africa: StatSoft S. Africa (Pty) Ltd.

Sweden: StatSoft Scandinavia AB

Taiwan: StatSoft Taiwan

UK: StatSoft Ltd.