

## XXII. Стандартные графики

- [Столбчатые диаграммы](#)
  - [Простые столбчатые диаграммы](#)
  - [Кластеризованные столбчатые диаграммы](#)
  - [Состыкованные диаграммы](#)
- [Линейчатые диаграммы](#)
  - [Простые линейчатые диаграммы](#)
  - [Сложные линейчатые диаграммы](#)
  - [Связанные линейчатые диаграммы](#)
- [Диаграммы с областями](#)
  - [Простая диаграмма с областями](#)
  - [Состыкованные диаграммы с областями](#)
- [Круговые диаграммы](#)
- [Диаграммы максимальных и минимальных значений](#)
  - [Простые биржевые диаграммы — потолок-пол-закрытие](#)
  - [Кластеризованные диаграммы — максимум-минимум-закрытие](#)
  - [Линейчатые диаграммы разностей](#)
  - [Простые интервальные столбцы](#)
  - [Кластеризованные интервальные столбцы](#)
- [Коробчатые диаграммы](#)
  - [Простые коробчатые диаграммы](#)
  - [Кластеризованные коробчатые диаграммы](#)
- [Столбики ошибок](#)
  - [Простая диаграмма величины ошибки](#)
  - [Кластеризованная величина ошибки](#)
- [Диаграмма рассеяния](#)
  - [Простая диаграмма рассеяния](#)
  - [Матричные диаграммы рассеяния](#)
  - [Наложенные диаграммы рассеяния](#)
  - [Трёхмерные диаграммы рассеяния](#)
- [Гистограммы](#)
- [Диаграммы Парето](#)
- [Контрольные карты](#)
- [Диаграммы нормального распределения](#)
- [Кривые ROC](#)
- [Временные диаграммы и графики последовательностей](#)
- [Основы редактирования графиков](#)
- [Редактор диаграмм](#)
- [Примеры редактирования графиков](#)
  - [Пример первый: изменение наименования осей](#)
  - [Пример второй: редактирование круговой диаграммы](#)
  - [Пример третий: нанесение регрессионных линий](#)

Одним из достоинств SPSS для Windows является наличие большого количества разнообразных графиков, которые могут быть построены как при помощи процедур меню графиков, так и из разнообразных процедур меню статистик. Что касается последнего меню, то для выяснения специальных возможностей графического представления Вы можете обратиться к главам: 6 (частотный анализ), 10 (предварительное исследование данных), И (таблицы сопряженности), 16 (регрессионный анализ), 20 (анализ выживания) и 24 (многомерное масштабирование). В главе 4 (Краткий обзор SPSS для Windows) уже были рассмотрены некоторые вопросы построения и редактирования графиков.

Каждый созданный график появляется в окне просмотра вместе с другими таблицами. Для построения графика, как правило, оказывается достаточным после выбора типа графика указать необходимые переменные, на основании которых он и будет построен по ранее заданной схеме. Если же у Вас появилось желание отредактировать график по своему вкусу, то для этого необходимо дважды щёлкнуть на какой-либо точке в пределах графика. После этого у Вас появится множество возможностей для дополнительного редактирования.

Начиная с 8-ой версии в SPSS наряду с традиционными стандартными графиками существует возможность создавать и интерактивные графики. Стандартные графики строятся при помощи многочисленных процедур статистического меню или меню графиков, составные компоненты которых и соответственно их возможности несколько не изменились. Однако, в меню графиков добавилась ещё одна позиция — Interactive (Интерактивно), которая открывает ещё одно собственное меню, служащее для построения так называемых интерактивных графиков. Интерактивные графики дают довольно широкую палитру новых возможностей.

Наряду с удобными глобальными возможностями менять отдельные стилевые элементы графиков и преобразовывать переменные, используемые для построения графика, отныне при помощи интерактивных графиков становится также возможным одновременное построение нескольких графиков для отдельных категорий дополнительных переменных.

Чтобы последовательно изложить эти новые возможности интерактивных графиков, процедуры построения графиков в SPSS должны быть рассмотрены в двух отдельных главах. В текущей главе рассматриваются исключительно традиционные стандартные графики; новые интерактивные графики будут представлены в следующей главе (гл. 23). Обратимся теперь к стандартным графикам.

Разобраться в многочисленных графиках, создаваемых при помощи меню графиков составляет трудность пожалуй только для новичка, поэтому мы не будем здесь рассматривать все имеющиеся тонкости. Однако мы попытаемся дать обзор графиков при помощи типичных практических примеров. При этом в окне просмотра будет выводиться установленный по умолчанию базовый вид графиков, правда, с необходимыми для нас заголовками, подзаголовками и сносками. Возможные изменения (штриховки, цвет, виды линий, виды диаграмм, изменение типа и размера шрифта и т.д.) будут рассмотрены в разделе 22.16.

При разработке графического представления диаграмм можно заметить, что в принципе на практике существуют две различные исходные ситуации. Наиболее часто встречается ситуация, когда дополнительно к результатам статистического анализа, хранящимся в файле данных SPSS, необходимо построить и графическое представление этих результатов. К примеру, у Вас появилось желание представить частоты четырёх возрастных групп из исследования гипертонии (файл hyper.sav) в виде линейчатой диаграммы. В этом случае компьютер сам при помощи соответствующих расчётов находит частоты, необходимые для построения столбцов диаграммы.

Совсем другую ситуацию можно наблюдать, если перед нами находятся уже подсчитанные и обработанные данные. Такой случай возникает, если бы, к примеру, Вы взяли из газеты информацию о ежедневной добыче нефти стран, входящих в ОПЕКС, и захотели бы представить эти данные в виде линейчатой диаграммы. При наличии таких готовых данных, очень часто приходится поразмыслить над тем, как их представить в файле.

- Если Вы щёлкните в списке меню на Graphs (Графики), то увидите меню с вариантами графиков.

Различные виды графиков будут по отдельности рассмотрены в разделах 22.1 по 22.14.

Перед рассмотрением графиков необходимо остановиться ещё раз на одном важном моменте. Установки по умолчанию задают различные цвета, в которые окрашиваются элементов графиков (к примеру, маркеры, сегменты) и линии, что облегчает понимание диаграммы и улучшает презентабельность. Если же Вы хотите напечатать график на принтере или представить его в других формах, то в большинстве подобных случаев использовать цветные графики не рекомендуется. В таких случаях разные поверхности Вы можете обозначить при помощи различных штриховок, а разные линии при помощи различных видов линий.

- Эти свойства вы сможете изменить, если выберете в меню Edit (Правка) Options... (Параметры) и в диалоговом окне Options (Параметры) щёлкните на Charts (Диаграммы).

- В разделе Fill Patterns and Line Styles (Заливка узором и стиль линий) вместо опции Cycle through colors, then patterns (Сначала просмотреть цвета, затем узоры) активируйте опцию Cycle through patterns (Просмотреть узоры).

В рассматриваемых далее методах построения графиков через выключатель Titles... (Заголовки) Вы можете присвоить диаграмме своё название, через выключатель Options... (Параметры) выбрать метод обработки пропущенных значений и в поле Template (Шаблон) при помощи активирования Use chart specifications from: (Установки диаграммы взять из:) загрузить установки для построения графика из других файлов.

## 22.1. Столбчатые диаграммы

Столбчатые диаграммы применяются, как правило, в следующих ситуациях:

- Отображение частот переменных, относящихся к номинальной или порядковой шкале

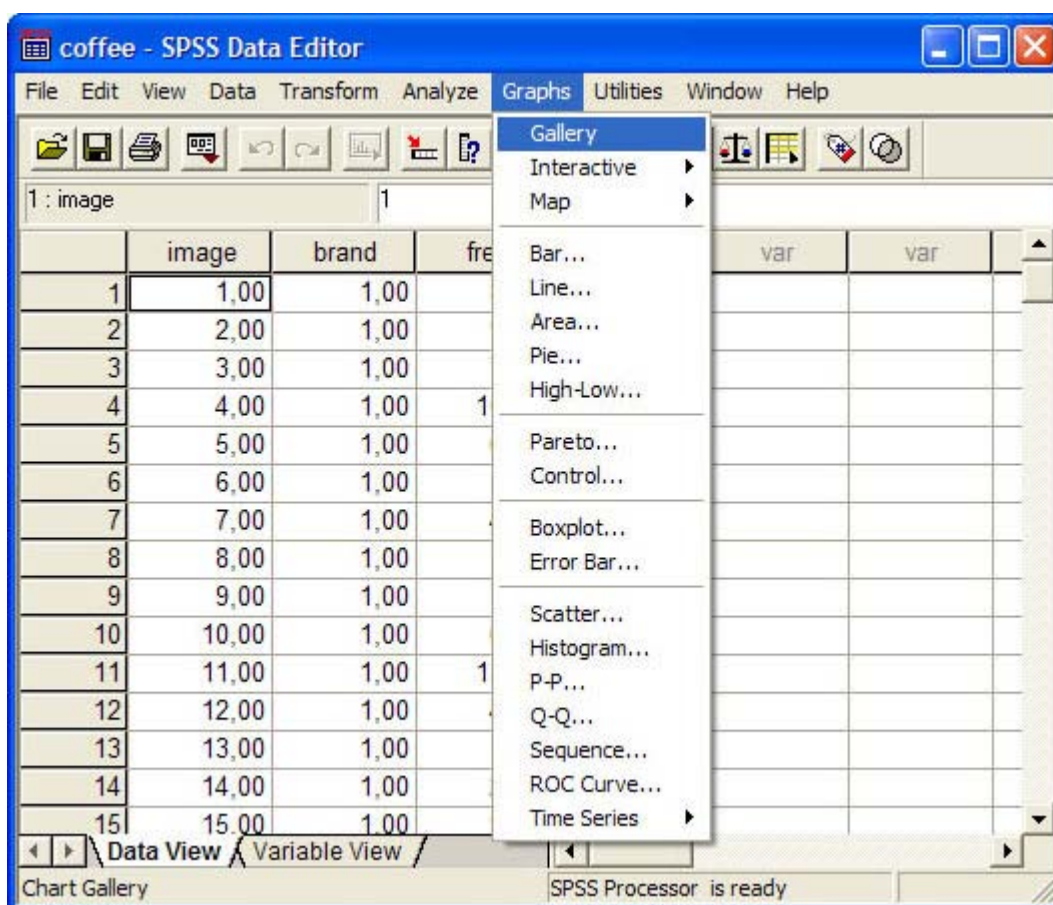


Рис. 22.1: Меню с вариантами графиков

- Отображение средних значений, сумм или других показателей последовательных переменных (т.е. переменных, принадлежащих к интервальной шкале или к шкале отношений), отображение переменных, сгруппированных по категориям переменных с номинальной или порядковой шкалой или временной зависимости.
- Для построения столбчатой диаграммы, после открытия соответствующего файла SPSS, выберите в меню Graphs (Графики) Bar... (Столбчатые)

Откроется диалоговое окно Bar Charts (Столбчатые диаграммы) (см. рис. 22.2).

Вы можете выбрать между простой, кластеризованной (кластерной) и состыкованной столбчатыми диаграммами. Данные, отображаемые в этих диаграммах, могут быть заданы как категории одной переменной, как разные переменные или как значения отдельных наблюдений.

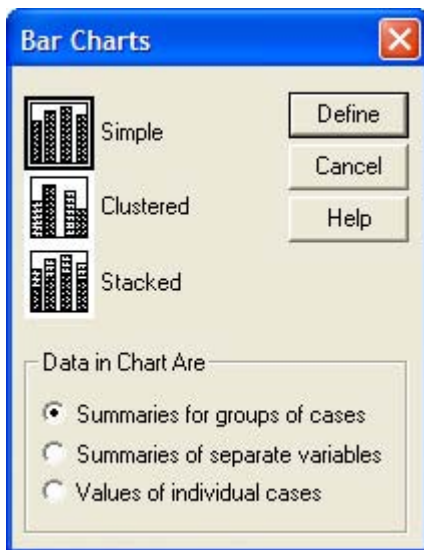


Рис. 22.2: Диалоговое окно Bar Charts (Столбчатые диаграммы)

### 22.1.1. Простые столбчатые диаграммы

- Откройте файл с данными об исследовании гипертонии (файл hyper.sav).

Мы хотим построить столбчатую диаграмму для процентных показателей частот четырёх возрастных групп (переменная ак).

- Щёлкните на области Simple (Простая) и оставьте предварительную установку Summaries for groups of cases (Обработка категорий одной переменной).
- Щёлкните по кнопке Define (Определить); откроется соответствующее диалоговое окно.
- В поле Category Axis: (Ось категорий) введите переменную ак, активируйте % of cases (% наблюдений) и, пройдя выключатель Titles... (Заголовок), введите заголовок для диаграммы.
- Щёлкните на ОК.

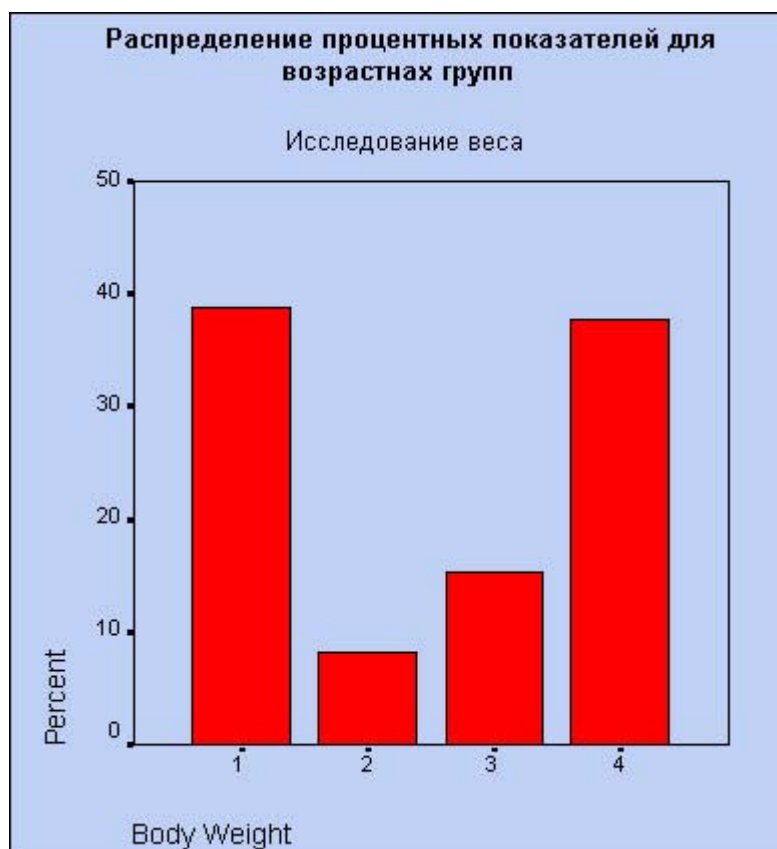


Рис. 22.3: Диалоговое окно Define Simple Bar: Summaries for groups of cases (Простая столбчатая диаграмма: Обработка категорий одной переменной)

Будет построен график, показанный на рисунке 22.4.

Теперь представим в графическом виде изменение среднего значения уровня сахара в крови (переменные bz0, bz1, bz6 и bz12), взятого из того же файла (hyper.sav).

- В этот раз в диалоговом окне Bar Charts (Столбчатые диаграммы) активируйте Summaries of separate variables (Обработка отдельных переменных); после нажатия выключателя Define (Определить) откроется соответствующее диалоговое окно (см. рис. 22.5).
- В поле Bars Represent (Значения столбцов) по очереди внесите переменные bz0, bz1, bz6 и bz12 и оставьте установленную по умолчанию функцию Mean of values (Средние значения).
- Пройдя выключатель Titles... (Заголовок), введите заголовок диаграммы.
- Щёлкните на ОК.



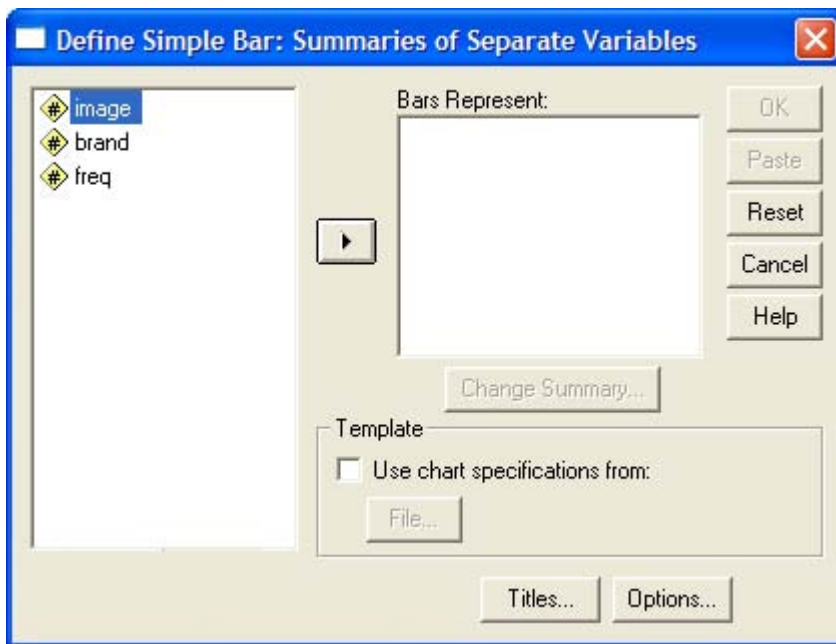
**Рис. 22.4:** Простая столбчатая диаграмма (Категории одной переменной)

Будет построен график, приведенный на рисунке 22.6.

Следует отметить тот недостаток, что в этой диаграмме не полностью приведены метки значений и на вертикальной оси показана только ограниченная область от 103,5 до 106,0, из-за чего по ошибке можно сделать неверное заключение о сильном изменении уровня сахара. Вы можете подкорректировать эти ошибки в редакторе диаграмм.

- Если Вы хотите выбрать функцию отличную от установленной по умолчанию Mean of values (Средние значения), щёлкните на одной из переменных в списке и затем на выключателе Change Summary...(Изменить метод обработки).

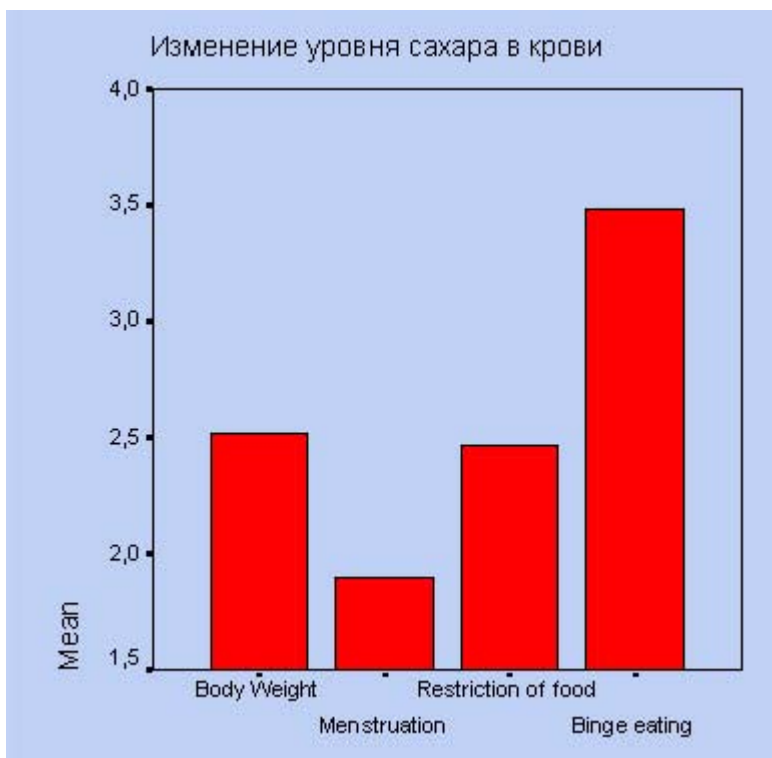
Откроется диалоговое окно с перечнем функций (см. рис. 22.7).



**Рис. 22.5:** Диалоговое окно *Define Simple Bar: Summaries of separate variables* (Построение простой столбчатой диаграммы: Обработка отдельных переменных)

Это диалоговое окно появляется только для столбчатой, линейной, круговой диаграмм и диаграммы с областями, причём не каждая из находящихся здесь функций пригодна для всех видов диаграмм. Если для имеющихся данных Вы хотите отобразить медианы или другие процентиля (сравните с гл. 6), то активируйте опцию *Values are grouped midpoints* (Значения являются сгруппированными средними точками).

В следующем примере рассматривается вопрос отображения готовых данных. Допустим, Вы взяли из некоторой газеты данные по 1993 году о добыче нефти в семи странах, входящих в ОПЕК и являющихся ведущими в этой отрасли.



**Рис. 22.6:** Простая столбчатая диаграмма (Отдельные переменные)

Страна	Млн.баррель/день
Саудовская- Аравия	8,0
Иран	3,3
Венесуэла	2,3
Объединённые Арабские Эмираты	2,2
Нигерия	1,8
Кувейт	1,6
Ливия	1,4

Представим эти данные в форме столбчатой диаграммы.

- Откройте файл oel.sav.
- В диалоговом окне Bar Charts (Столбчатые диаграммы) активируйте опцию Values of individual cases (Значения отдельных наблюдений).

После нажатия выключателя Define (Определить) откроется соответствующее диалоговое окно.

- В поле Bars Represent (Значения столбцов) внесите переменную barrel; в группе Category Labels (Метки категорий) активируйте Variable: (Переменная) и внесите переменную land.

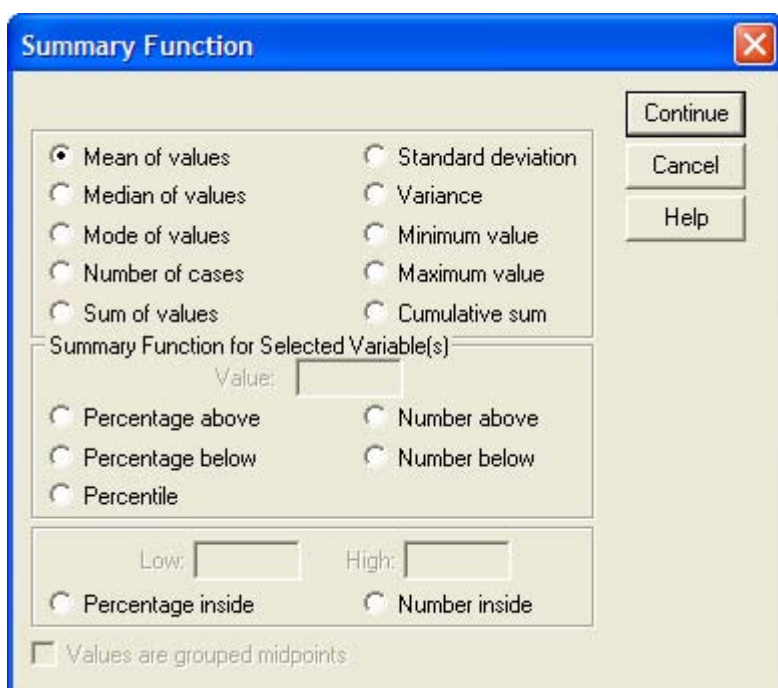
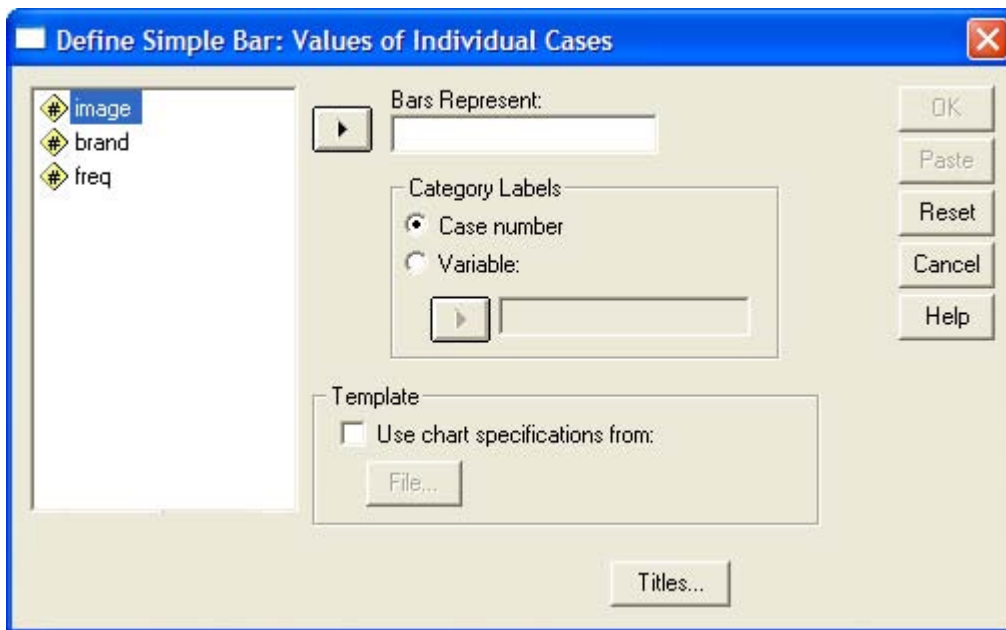


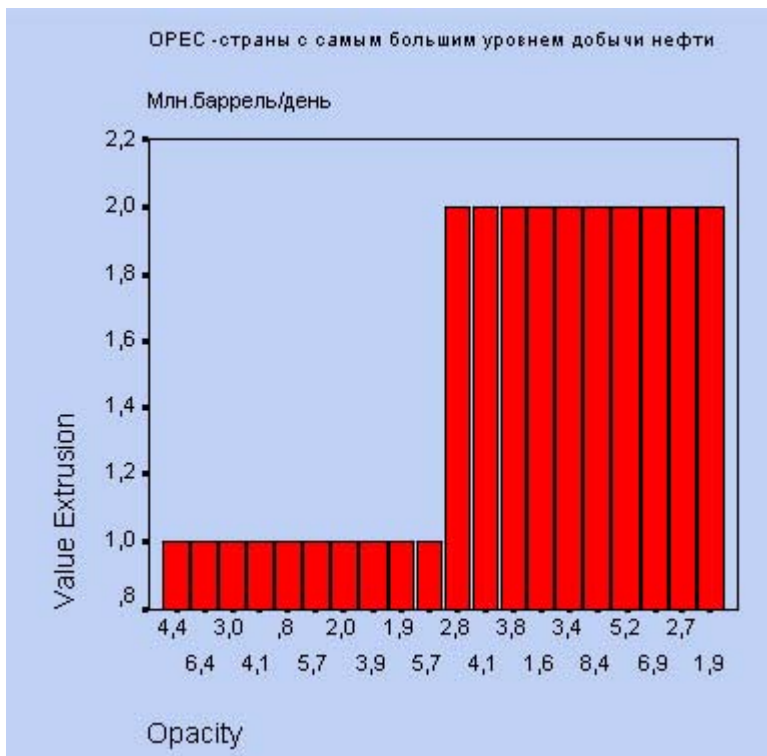
Рис. 22.7: Диалоговое окно Summary Function (Обрабатывающая функция).



**Рис. 22.8:** Диалоговое окно *Define Simple Bar: Values of individual cases* (Построение простой столбчатой диаграммы: Значения отдельных случаев)

- Пройдя выключатель **Titles...** (Заголовок), введите заголовок диаграммы и щёлкните на **OK**.

График будет выглядеть так, как на рисунке 22.9.



**Рис. 22.9:** Простая столбчатая диаграмма (Значения отдельных случаев)

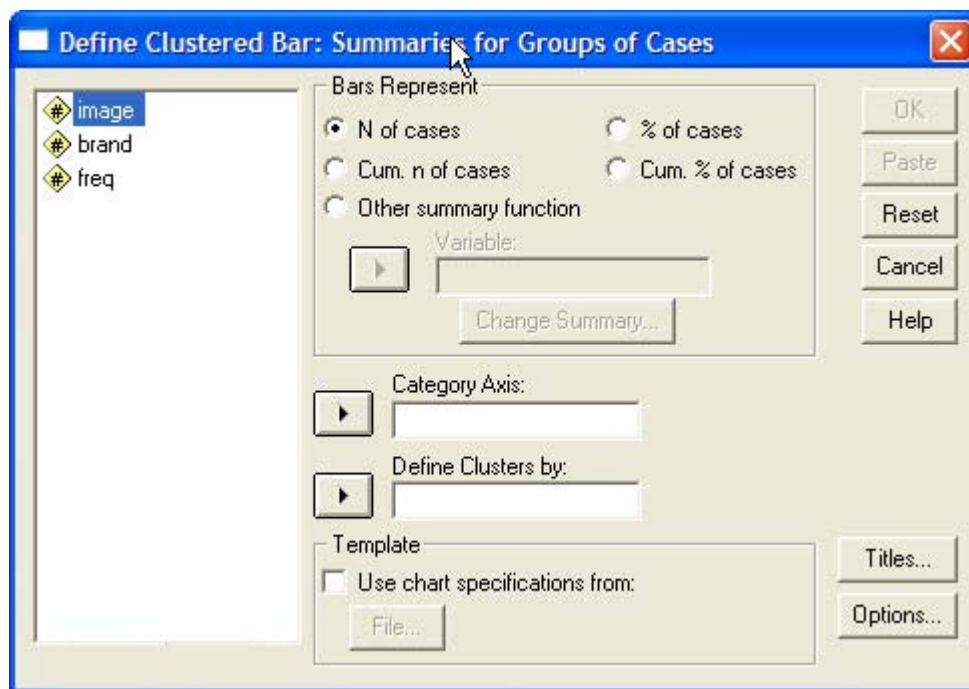
### 22.1.2. Кластеризованные столбчатые диаграммы

Теперь в целях обработки данных, полученных в ходе исследования гипертонии (файл `hyper.sav`), отдельно для двух методик лечения (переменная `med` с двумя своими значениями,



равными 1 и 2) в графическом виде должны быть представлены частотные показатели четырёх возрастных групп (переменная ak) в процентном выражении.

- Откройте файл hyper.sav.
- В диалоговом окне Bar Charts (Столбчатые диаграммы) щёлкните на области Clustered (Кластеризованная); активируйте опцию, устанавливаемую по умолчанию, Summaries for groups of cases (Обработка категорий одной переменной).
- Щёлкните на кнопке Define (Определить); откроется главное диалоговое окно, изображённое на рисунке 22.10.



**Рис. 22.10:** Диалоговое окно *Define Clustered Bar: Summaries for groups of cases* (Построение группированной диаграммы: Обработка категорий одной переменной)

- В поле Category Axis: (Ось категорий) введите переменную ak, в поле Define Clusters by: (Создать группы при помощи:) введите переменную med. Активируйте % of cases (% наблюдений).
- Пройдя выключатель Titles... (Заголовок), введите заголовок для диаграммы и начните построение диаграммы щелчком на ОК (см. рис. 22.11).

В качестве примера графического представления готовых данных рассмотрим доли I рынка принадлежащие самым крупным изготовителям компьютеров в 1991 и 1992 годах:

Изготовитель	Доля рынка, %	
	1991	1992
IBM	16,3	12,4
Apple	11,2	11,9
Compaq	6,0	6,6
NEC	6,4	5,1
Dell	1,7	3,5

Эти данные построчно сохранены в переменных firma (изготовитель), jahr (год) и anteil (доля) в файле pc.sav.

- Откройте файл pc.sav и просмотрите его содержимое в редакторе данных.

- В диалоговом окне Bar Charts (Столбчатые диаграммы) щёлкните на области Clustered (Кластеризованная) и активируйте устанавливаемую по умолчанию опцию Summaries for groups of cases (Обработка категорий одной переменной).
- После щелчка на выключателе Define (Определить) в открывшемся диалоговом окне в поле Category Axis: (Ось категорий) введите переменную firma, а в поле Define Clusters by: (Определить группы по:) — переменную jahr. В группе Bars Represent (Значения столбцов) активируйте Other summary function (Другая обрабатываемая функция) и в появившееся поле введите переменную anteil; функцию Mean of values (Средние значения) можете оставить.
- Пройдя выключатель Titles... (Заголовок), введите подходящий заголовок и начните построение диаграммы щелчком на ОК.

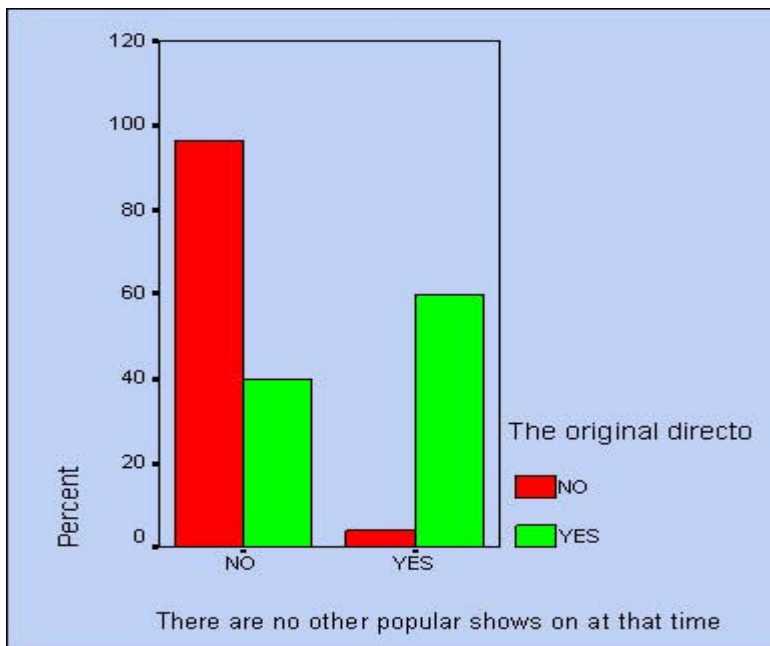


Рис. 22.11: Группированная столбчатая диаграмма

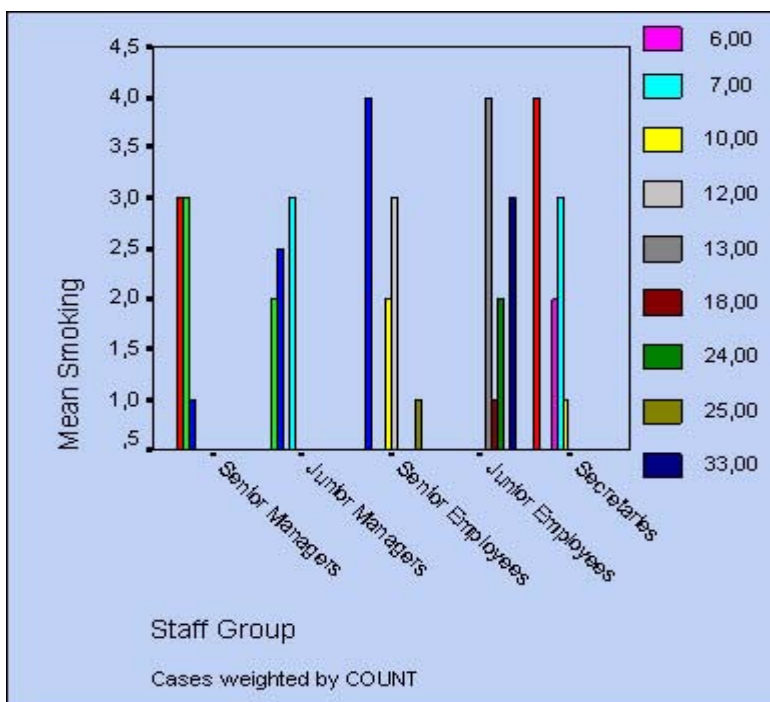


Рис. 22.12: Группированная столбчатая диаграмма

### 22.1.3. Состыкованные диаграммы

Как правило, состыкованная столбчатая диаграмма применяется тогда, когда столбцы отражают частоты, которые должны быть разделены при помощи некоторой внешней переменной. В таком случае, и обзор суммарных частот предоставляется пользователю иначе, нежели в виде кластеризованной столбчатой диаграммы.

- Откройте файл `studium.sav`, содержащий данные опроса студентов.

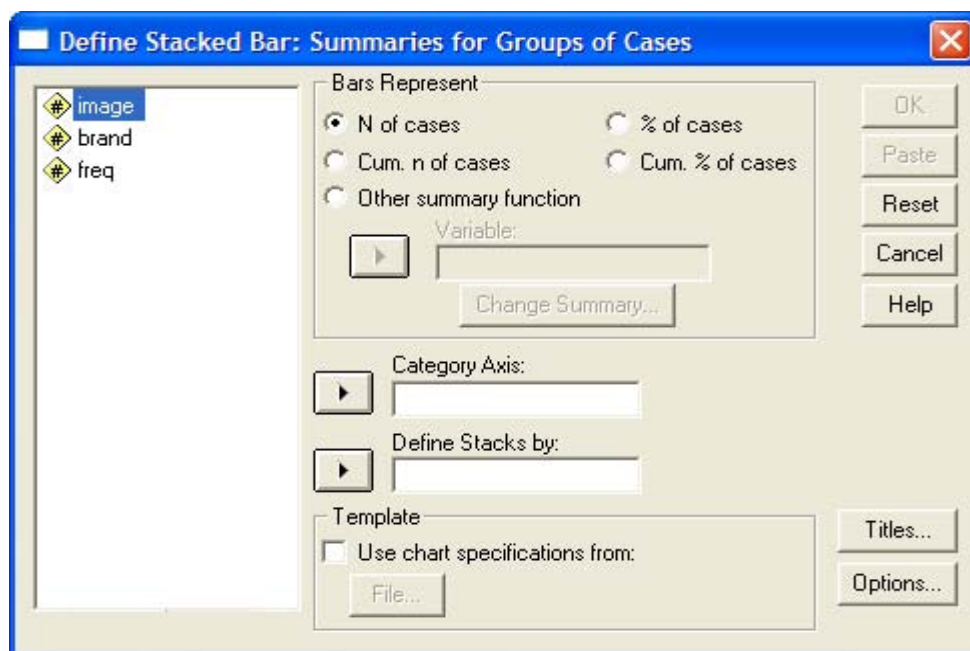
Мы хотим отобразить в графическом виде распределение частот, отражающих психологическое состояние студентов (переменная `psyche`), отдельно для каждого пола (переменная `sex`).

- В диалоговом окне `Bar Charts (Столбчатые диаграммы)` щёлкните на области `Stacked (Состыкованная)` и активируйте опцию, устанавливаемую по умолчанию, `Summaries for groups of cases (Обработка категорий одной переменной)`. Щелчком по кнопке `Define (Определить)` откройте соответствующее диалоговое окно.
- В поле `Category Axis: (Ось категорий)` введите переменную `psyche`, а в поле `Define Stacks by: (Создать штабели при помощи:)` введите переменную `sex`. Оставьте установку по умолчанию `N of cases (Количество наблюдений)`.
- Пройдя выключатель `Titles... (Заголовок)`, введите подходящий заголовок.

В данном примере имеются пропущенные значения, которые в соответствии с установками по умолчанию будут обрабатываться как отдельные категории.

- Для того, чтобы запретить это действие щёлкните на выключателе `Options... (Параметры)` и уберите отметку для опции `Display groups defined by missing values (Пропущенные значения отображать как категории)`.
- Вернувшись в диалоговое окно `Define Stacked Bar: Summaries for groups of cases (Построение состыкованной диаграммы: Обработка категорий одной переменной)` щелчком на `OK` начните построение диаграммы (см. рис. 22.11).

В следующем примере рассматривается графическое представление уже имеющихся (готовых) данных. Приведенная ниже таблица содержит показатели рождаемости в западных и восточных землях Германии, начиная с 1985 по 1992 год:



**Рис. 22.13:** Диалоговое окно `Define Stacked Bar: Summaries for groups of cases` (Построение штабельной диаграммы: Обработка категорий одной переменной)

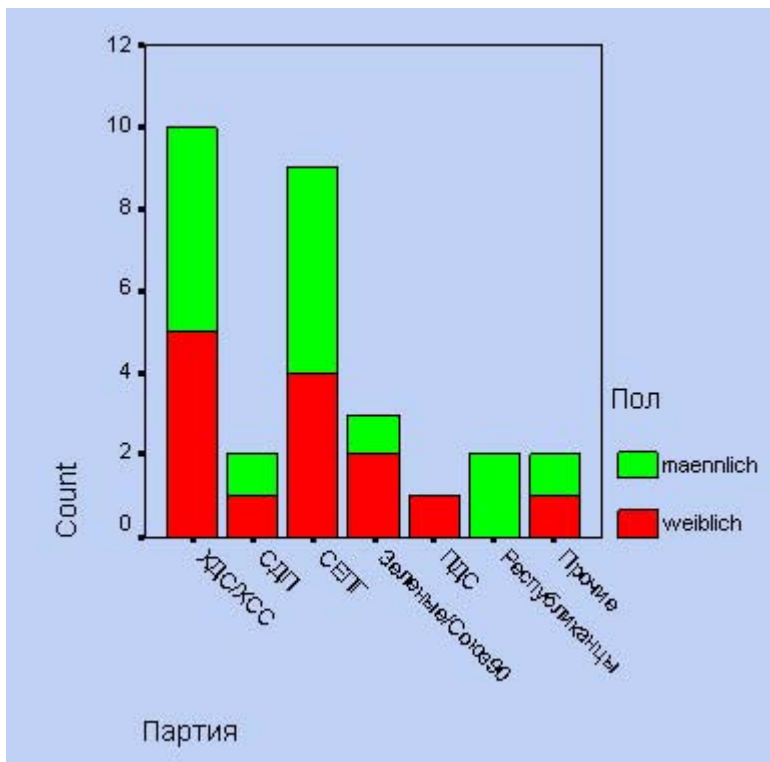


Рис. 22.14: Штабельная столбчатая диаграмма

Гол	Количество	
	Запад	Восток
1985	586.155	227.648
1986	635.963	222.229
1987	642.010	225.959
1988	677.259	215.734
1989	681.537	198.922
1990	727.199	178.476
1991	722.250	107.769
1992	718.730	87.030

- Откройте файл geburten.sav и просмотрите его содержимое в редакторе данных.

Эти данные построчно сохранены в переменных jahr (год), wo и anz (количество). Переменная wo при помощи кодировок 1 и 2 указывает на принадлежность к Западной или Восточной Германии.

- В диалоговом окне Bar Charts (Столбчатые диаграммы) щёлкните на области Stacked (Состыкованная) и активируйте опцию Summaries for groups of cases (Обработка категорий одной переменной), устанавливаемую по умолчанию.
- После щелчка на выключателе Define (Определить) в открывшемся диалоговом окне в поле Category Axis: (Ось категорий) введите переменную jahr, а в поле Define Stacks by: (Создать штабели при помощи:) — переменную wo. В группе Bars Represent (Значения столбцов) активируйте Other summary function (Другая обрабатывающая функция) и в появившееся поле введите переменную anz; вместо установленной по умолчанию функции Mean of values (Средние значения), пройдя выключатель Change Summary...(Изменить метод обработки) отметьте функцию суммы (Sum of values).
- С помощью кнопки Titles... (Заголовок), введите подходящий заголовок и начните построение диаграммы щелчком на ОК.

## 22.2. Линейчатые диаграммы

Линейчатую диаграмму вместо столбчатой следует выбирать тогда, когда необходимо отобразить большое количество столбцов, а также тогда, когда столбцы располагаются в определенной последовательности. Как правило, это временная последовательность.

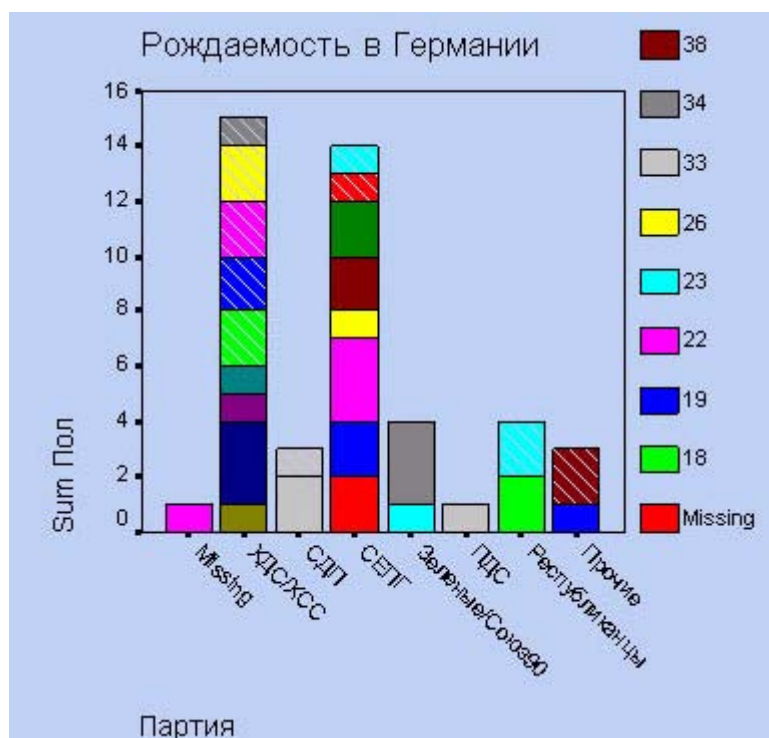


Рис. 22.15: Штабельная столбчатая диаграмма

- Для построения линейной диаграммы после открытия соответствующего файла SPSS выберите в меню: Graphs (Графики) Line... (Линейчатые)

Откроется диалоговое окно Line Charts (Линейчатые диаграммы) (см. рис. 22.16).

Вы можете построить простую, сложную и связанную линейные диаграммы. Как и для столбчатых диаграмм данные, отображаемые в этих диаграммах, могут быть заданы как категории одной переменной, как разные переменные или как значения отдельных наблюдений.

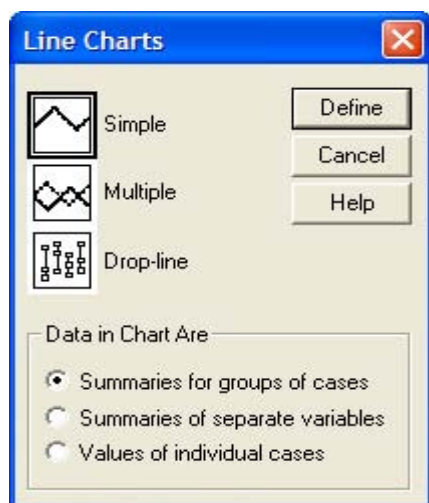


Рис. 22.16: Диалоговое окно Line Charts (Линейчатые диаграммы)

## 22.2.1. Простые линейчатые диаграммы

В файле `buecher.sav` хранится информация о развитии книгопечатания в Германии с 1962 по 1991 год.

- Откройте файл `buecher.sav` и просмотрите его содержимое в редакторе данных.
- В диалоговом окне Line Charts (Линейчатые диаграммы) щёлкните на области Simple (Простая) и оставьте, опцию Summaries for groups of cases (Обработка категорий одной переменной), устанавливаемую по умолчанию.
- После щелчка по выключателю Define (Определить) откроется соответствующее диалоговое окно.
- В поле Category Axis: (Ось категорий) введите переменную `jahr` (год). В группе Line Represent (Значения линий) активируйте Other summary function (Другая обрабатывающая функция) и в появившееся поле введите переменную `anz` (количество). Вместо установленной по умолчанию функции Mean of values (Средние значения), пройдя выключатель Change Summary... (Изменить метод обработки), отметьте функцию суммы значений (Sum of values) (которая в данном случае, правда, дает тот же эффект).
- С помощью выключателя Titles... (Заголовок), введите подходящий заголовок.
- Начните построение диаграммы щелчком на ОК.

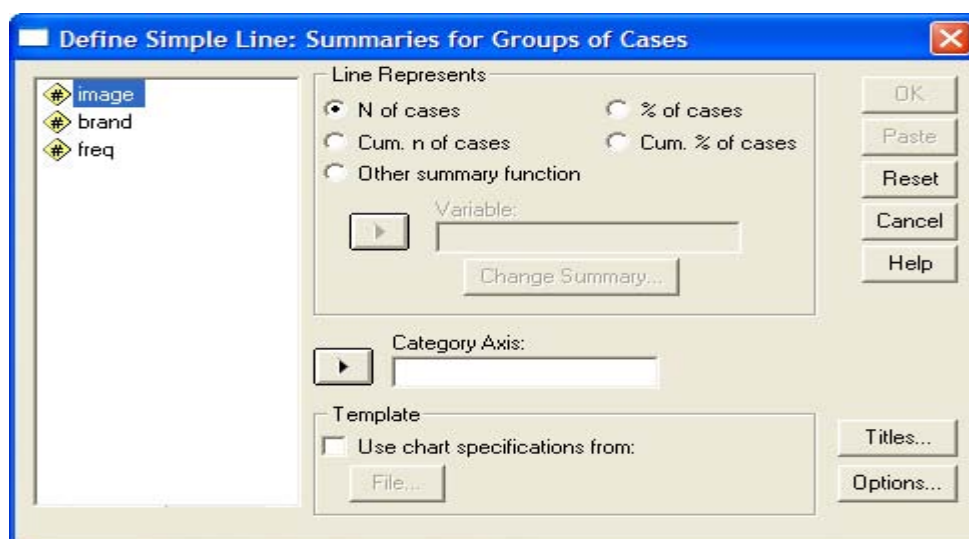


Рис. 22.17: Диалоговое окно *Define Simple Line: Summaries for Groups of Cases* (Построение простой линейчатой диаграммы: Обработка категорий одной переменной)

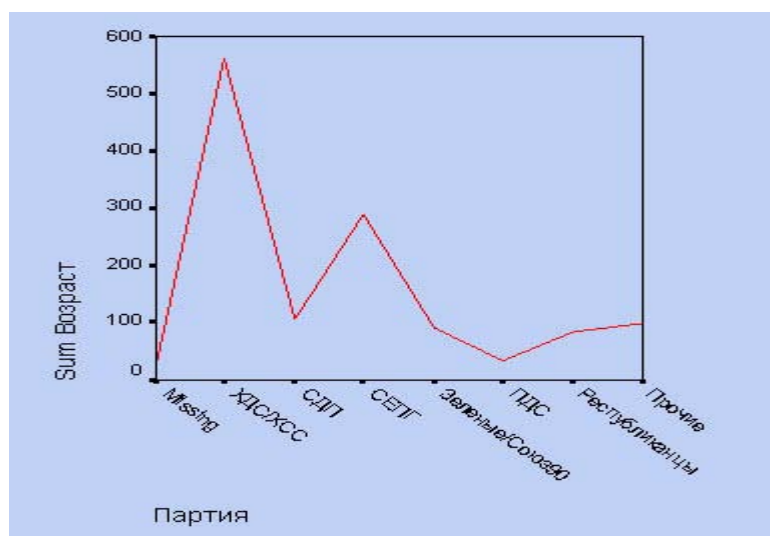


Рис. 22.18: Линейчатая диаграмма

## 22.2.2. Сложные линейчатые диаграммы

Следующая таблица демонстрирует тенденцию нарушения законов по охране окружающей среды в Западной Германии с 1985 по 1992 год:

Гол	Нарушения		
	UA	CV	UB
1985	2.750	8.562	901
1986	3.682	9.294	1.161
1987	5.390	10.529	1.311
1988	6.748	11.968	1.671
1989	8.559	11.827	1.590
1990	8.157	9.942	1.525
1991	9.724	9.601	1.457
1992	12.453	8.687	1.573

Где

UA — Переработка мусора, наносящая вред окружающей среде

GV — Загрязнение воды

UB — Использование запрещённого промышленного оборудования

Эти данные построчно сохранены в переменных jahr (год), ua, gv и ub в файле umwelt.sav.

- Откройте файл umwelt.sav и просмотрите его содержимое в редакторе данных.
- В диалоговом окне Line Charts (Линейчатые диаграммы) щёлкните на области Multiple (Сложная) и активируйте опцию Summaries of separate variables (Обработка отдельных переменных).
- После щелчка по выключателю Define (Определить) откроется соответствующее диалоговое окно (см. рис. 22.19).
- В поле Category Axis: (Ось категорий) введите переменную jahr. В поле Line Represent (Значения линий) по очереди введите переменные ua, gv и ub; вместо установленной по умолчанию функции Mean of values (Средние значения), с помощью выключателя Change Summary...(Изменить метод обработки), отметьте функцию суммы значений (Sum of values).
- После щелчка по выключателю Titles... (Заголовок), введите подходящий заголовок и начните построение диаграммы щелчком на ОК.

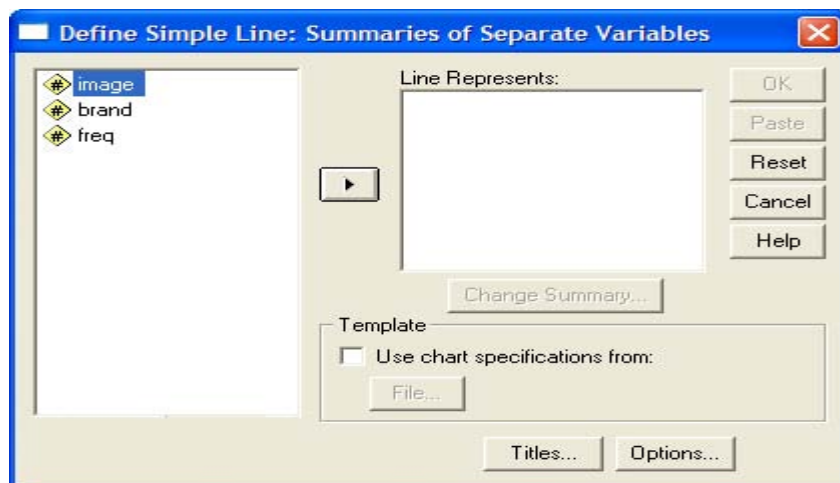


Рис. 22.19: Диалоговое окно Define Multiple Line: Summaries of Separate Variables (Построение сложной линейчатой диаграммы: Обработка отдельных переменных)

### 22.2.3. Связанные линейчатые диаграммы

Это разновидность сложной линейчатой диаграммы, в котором точки данных обозначены разными символами и соединены вертикальной связью.

- Воспользуйтесь примером из предыдущего раздела и в диалоговом окне Line Charts (Линейчатые диаграммы) щёлкните на области Drop-line (Связанные линии).
- Во всём остальном поступите так же, как и в предыдущем разделе.

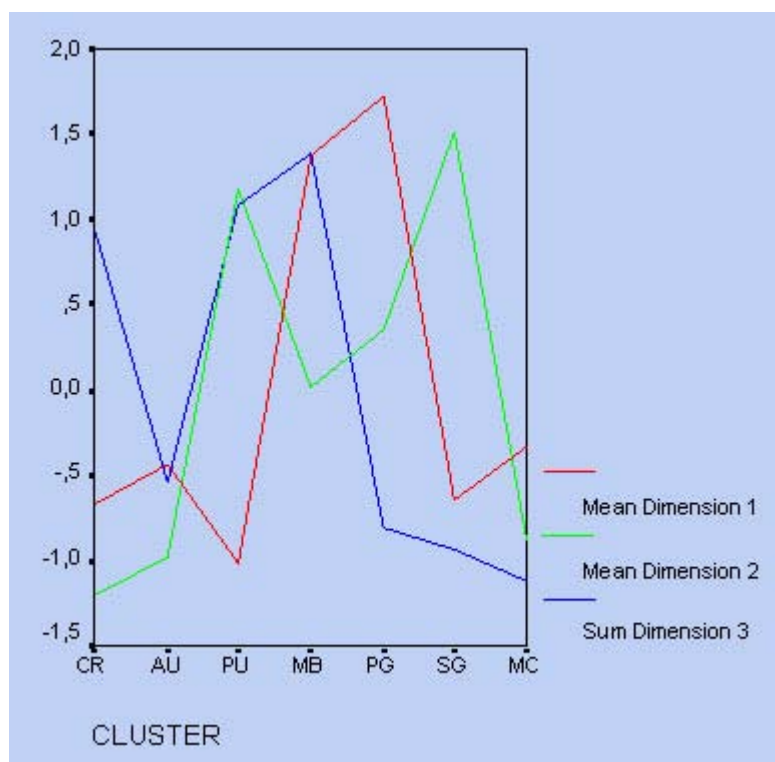


Рис. 22.20: Сложная линейчатая диаграмма

Построенная нами диаграмма будет соответствовать приведенной на рисунке 22.21.

### 22.3. Диаграммы с областями

Диаграммы с областями являются разновидностью линейчатой диаграммы, в которой области, находящиеся под линиями, закрашиваются благодаря чему график выглядит более наглядным.

- Для построения диаграммы с областями, после открытия необходимого файла SPSS, выберите в меню Graphs (Графики) Area... (С областями)

Откроется диалоговое окно Area Charts (Диаграммы с областями) Вы можете построить простую или состыкованную диаграмму с областями. И здесь данные, отображаемые в этих диаграммах, могут быть заданы как категории одной переменной, как разные переменные или как значения отдельных наблюдений.



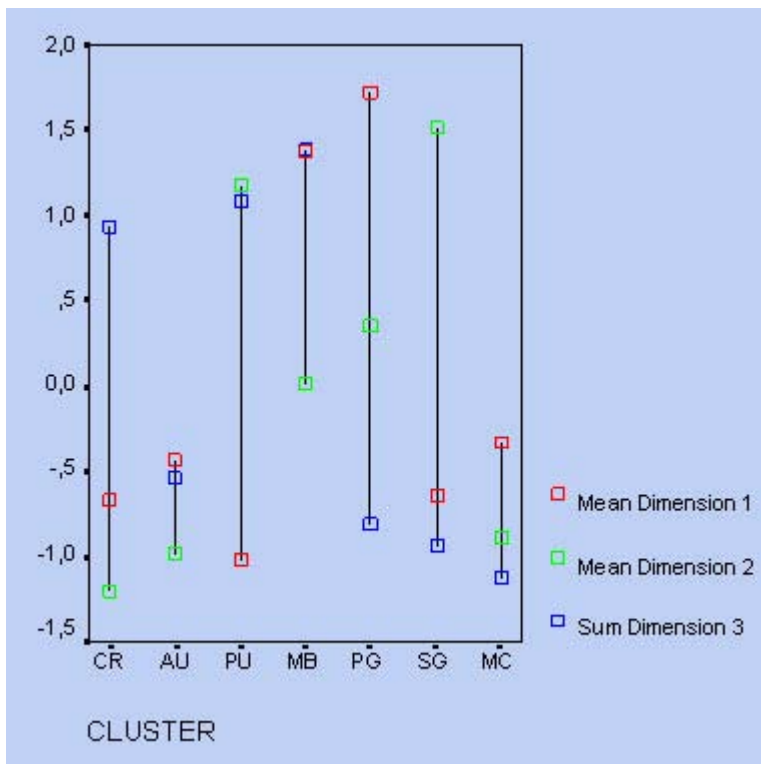


Рис. 22.21: Связанная линейчатая диаграмма

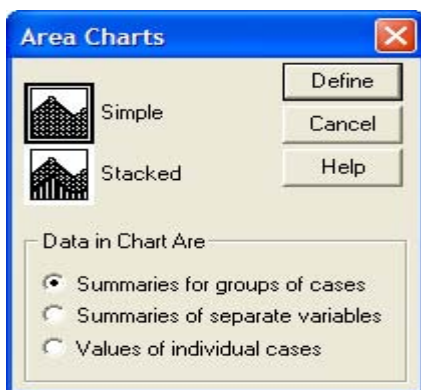


Рис. 22.22: Диалоговое окно Area Charts (Диаграммы с областями)

### 22.3.1. Простая диаграмма с областями

Следующая таблица содержит информацию о производстве велосипедов с 1986 по 1992 год. Производственные показатели разбиты дополнительно на сбыт внутри страны и экспорт.

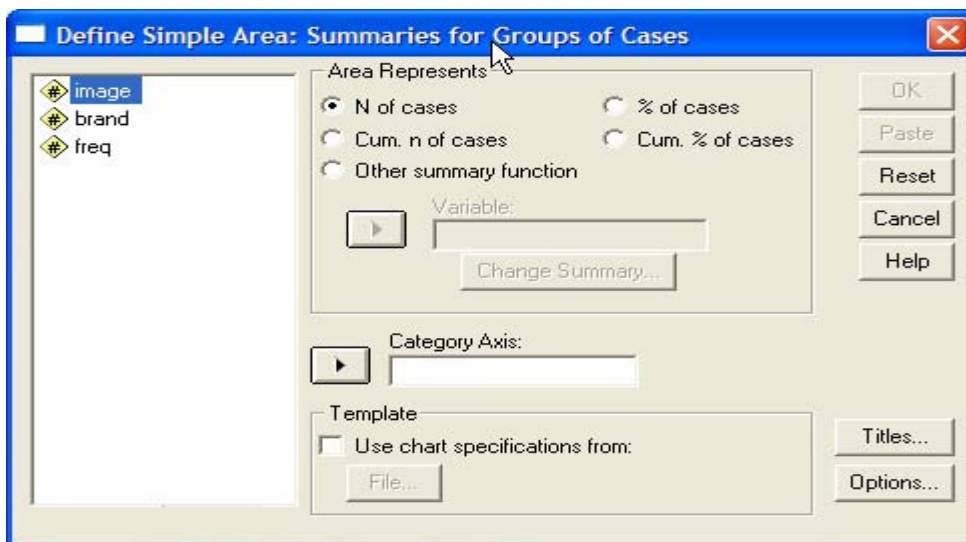
Год	Штук (млн.)		
	Производство	Внутри страны	Экспорт
1986	4,00	3,14	0,86
1987	3,74	3,01	0,73
1988	3,88	3,14	0,74
1989	4,40	3,67	0,73
1990	4,81	4,08	0,73
1991	4,91	4,35	0,56
1992	4,55	4,10	0,45

Эти данные построчно сохранены в переменных jahr (год), gesamt (общий объем производства), inland (внутри страны) и export (экспорт) в файле fahrrad.sav.

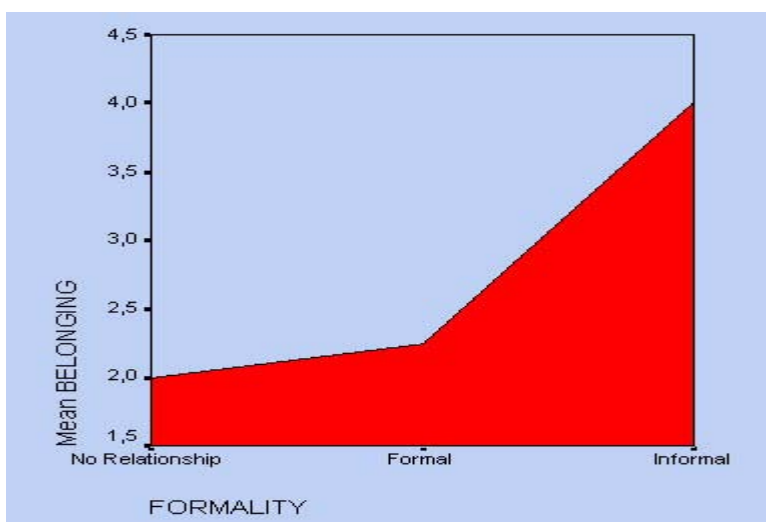
- Откройте файл fahrrad.sav и просмотрите его содержимое в окне редактора данных.

Сначала данные о совокупном производстве представим в виде простой диаграммы с областями.

- В диалоговом окне Area Charts (Диаграммы с областями) щёлкните на области Simple (Простая) и оставьте опцию Summaries for groups of cases (Обработка категорий одной переменной), устанавливаемую по умолчанию.
- После щелчка по выключателю Define (Определить) откроется главное диалоговое окно (см. рис. 22.23).
- В поле Category Axis: (Ось категорий) введите переменную jahr и в группе Area Represents (Значения областей) установите маркер возле Other summary function (Другая обрабатываемая функция). В появившееся поле введите переменную gesamt и оставьте функцию Mean of values (Средние значения), устанавливаемую по умолчанию.
- С помощью выключателя Titles... (Заголовок), введите подходящий заголовок и начните построение диаграммы щелчком на ОК.



**Рис. 22.23:** Диалоговое окно *Define Simple Area: Summaries for Groups of Cases* (Построение простой диаграммы с областями: Обработка категорий одной переменной)



**Рис. 22.24:** Диаграмма с областями

Следует отметить то, что начальной точкой отсчёта вертикальной оси является не ноль, а значение 3,6.

### 22.3.2. Состыкованные диаграммы с областями

Этот вид диаграмм следует применять только тогда, когда штабелируемые области дают не лишенный смысла эффект суммирования. Мы ещё раз обратимся к примеру, рассмотренному в предыдущем разделе, но теперь совокупную производительность разделим на продукцию, реализуемую внутри страны и экспорт.

- В диалоговом окне Area Charts (Диаграммы с областями) щёлкните на области Stacked (Состыкованная) и отметьте опцию Summaries of separate variables (Обработка отдельных переменных).
- После щелчка по выключателю Define (Определить) откроется соответствующее диалоговое окно.
- В поле Category Axis: (Ось категорий) введите переменную jahr, а в поле Areas Represent (Значения областей) введите обе переменные inland и export и оставьте функцию Sum of values (Сумма значений), устанавливаемую по умолчанию.
- Минув выключатель Titles... (Заголовок), введите подходящий заголовок и начните построение диаграммы щелчком на ОК.

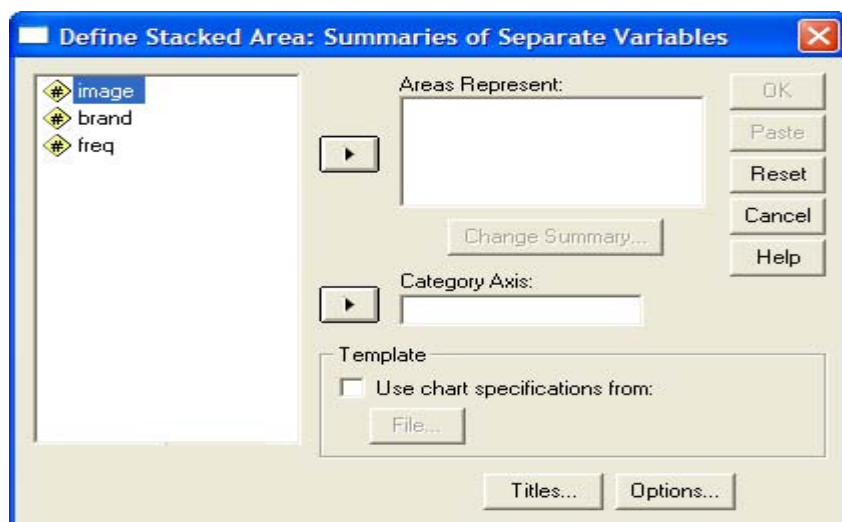


Рис. 22.25: Диалоговое окно Define Stacked Area: Summaries of Separate Variables (Построение штабельной диаграммы с областями: Обработка отдельных переменных)

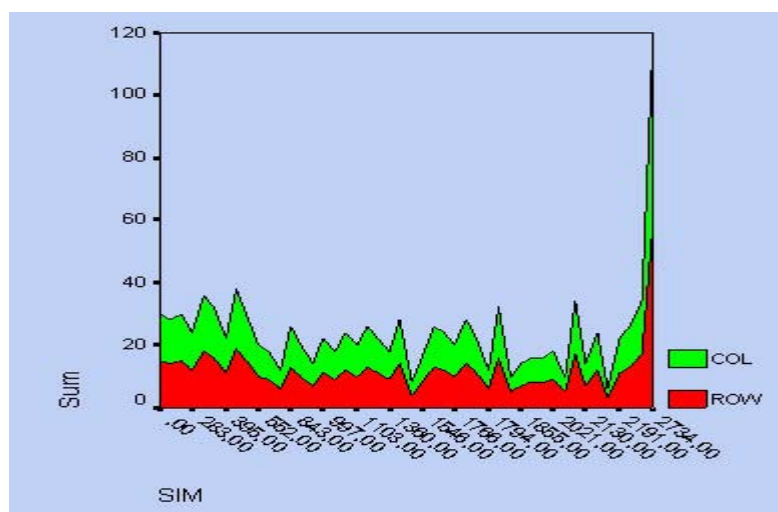


Рис. 22.26: Штабельная диаграмма с областями.

## 22.4. Круговые диаграммы

Представление данных в виде круговых диаграмм стоит выбирать тогда, когда частоты или значения переменных можно, не нарушая здравого смысла, сложить вместе и эта сумма будет соответствовать ста процентам.

Отобразим при помощи круговой диаграммы частоты категорий переменной *psyche* (психологическое состояние студентов) из файла *studium.sav*.

- Откройте файл *studium.sav* и выберите в меню *Graphs* (Графики) *Pie...* (Круговые). Откроется диалоговое окно *Pie Charts* (Круговые диаграммы).
- Оставьте опцию *Summaries for groups of cases* (Обработка категорий одной переменной), установленную по умолчанию и щелчком на кнопке *Define* (Определить) откройте следующее диалоговое окно.

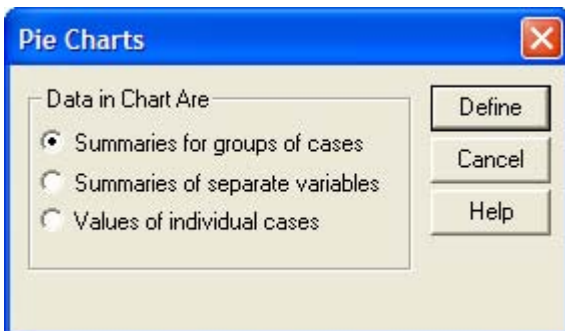


Рис. 22.27: Диалоговое окно *Charts* (Круговые диаграммы)

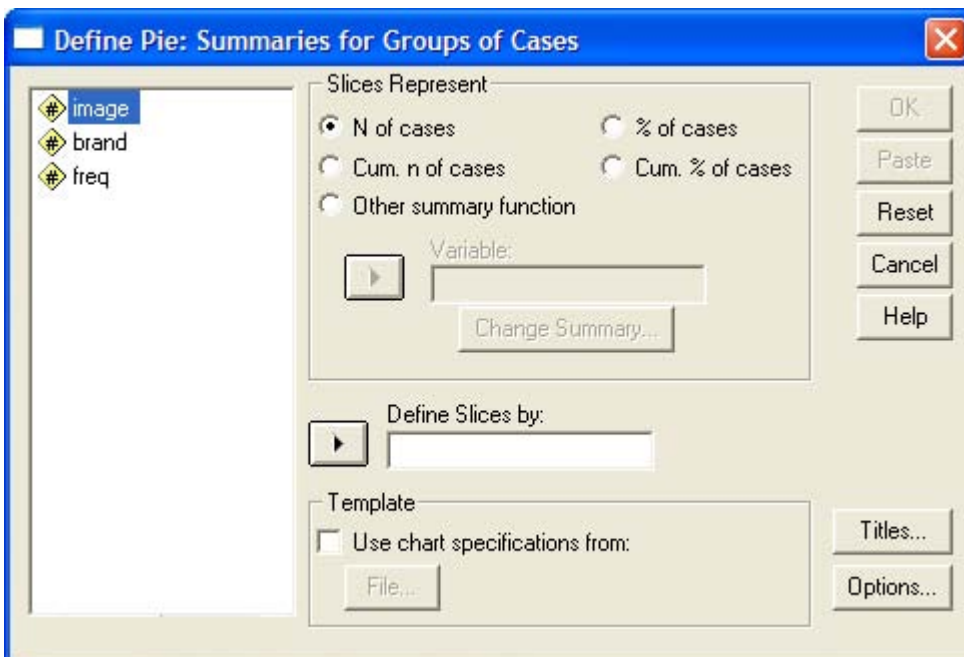


Рис. 22.28: Диалоговое окно *Define Pie: Summaries for Groups of Cases* (Построение круговой диаграммы: Обработка категорий одной переменной)

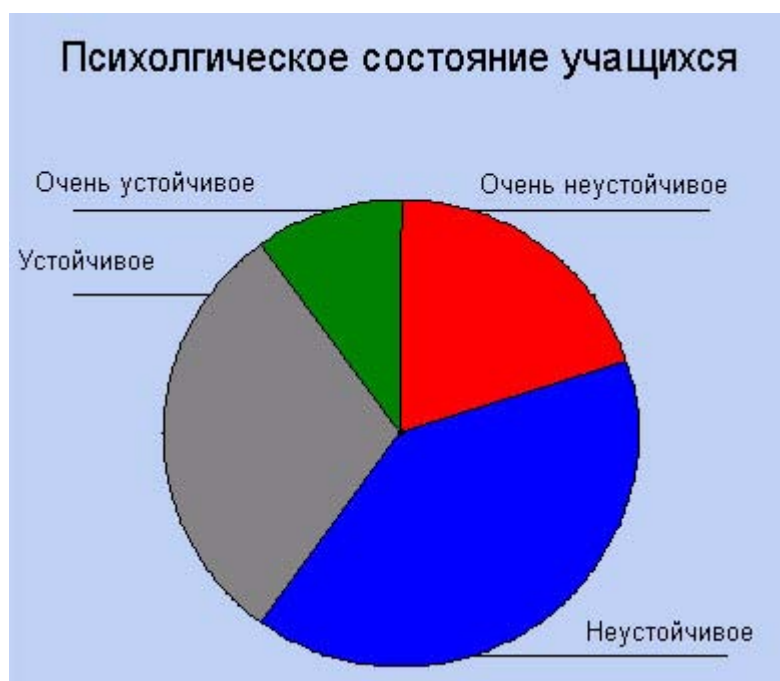
- В поле *Define slices by:* (Создать сектора при помощи:) введите переменную *psyche*.
- Щёлкните на выключателе *Options...* (Параметры) и уберите маркер с опции *Display groups defined by missing values* (Пропущенные значения отображать как категории).
- С помощью выключателя *Titles...* (Заголовков), введите подходящий заголовок и начните построение диаграммы щелчком на *OK* (см. рис. 22.29).

Типичным примером применения круговой диаграммы является отображение процентных показателей голосов избирателей, проголосовавших за те или иные партии.

На местных выборах земли Гессен в 1993 году получилось следующее распределение голосов в процентах:

Партия	ДОЛЯ ГОЛОСОВ (%)
SPD	36,4
CDU	32,0
Gruene (Зелёные)	11,0
Republikaner (Республиканцы)	8,3
FPD	5,1
Прочие	7,2

Этот пример является примером с уже имеющимися (готовыми) данными.



**Рис. 22.29:** Круговая диаграмма

- Откройте файл `kommunal.sav`, в котором в переменных `p` и `rg` построчно находятся необходимые для нас данные.
- В диалоговом окне Pie Charts (Круговые диаграммы) опять оставьте опцию Summaries for groups of cases (Обработка категорий одной переменной), установленную по умолчанию.
- После щелчка по выключателю Define (Определить) в поле Define slices by: (Создать сектора при помощи:) введите переменную `p`. Поставьте маркер возле Other summary function (Другая обрабатываемая функция) и в появившееся поле введите переменную `rg`. Используйте установленную по умолчанию опцию Sum of values (Сумма значений).
- С помощью выключателя Titles... (Заголовок), введите подходящий заголовок и начните построение диаграммы щелчком на ОК (см. рис. 22.30).

В главе 22.17 мы придадим этой диаграмме более презентабельный вид.

## 22.5. Диаграммы максимальных и минимальных значений

Если вы посмотрите на поведение биржевых котировок акций, то заметите, что для фиксированного промежутка времени, к примеру, для одного дня, существует три важнейших характеристики: максимальное и минимальное значения, а также значение в конце промежутка, при закрытии биржи. Такой и подобные ему процессы могут быть представлены при помощи диаграммы максимальных и минимальных значений, которая на биржевом сленге иногда называется потолок-пол-закрытие.

- После открытия необходимого Вам файла SPSS выберите в меню Graphs (Графики) High-Low... (Максимум-минимум)

После этого откроется соответствующее диалоговое окно.

Существует пять видов диаграмм максимума-минимума, данные для которых, как и для предыдущих графиков, могут интерпретироваться тремя различными способами.

### 22.5.1. Простые биржевые диаграммы - потолок-пол-закрытие

Предположите, что вы располагаете некоторыми акциями и фиксировали их котировки в течение десяти дней:

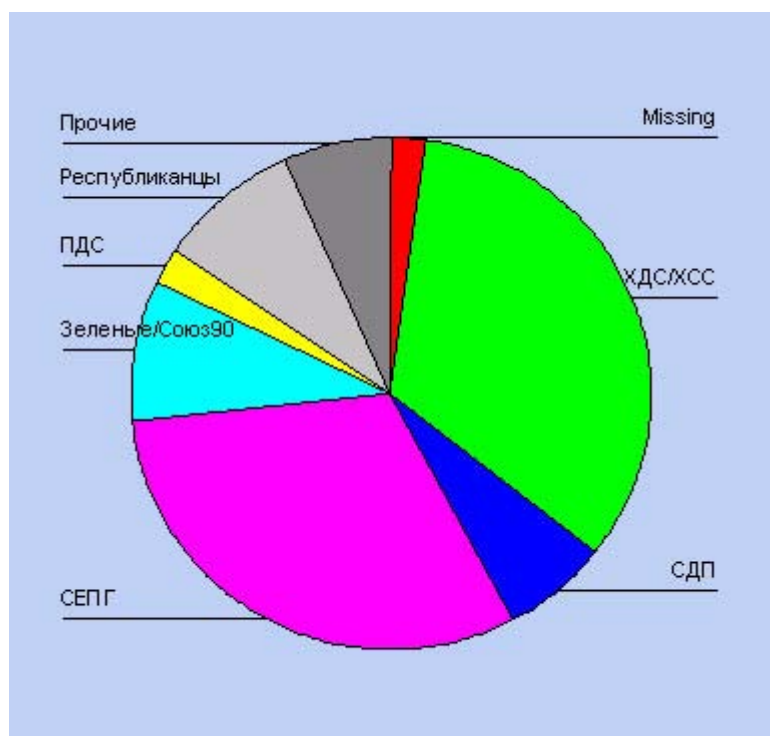


Рис. 22.30: Круговая диаграмма

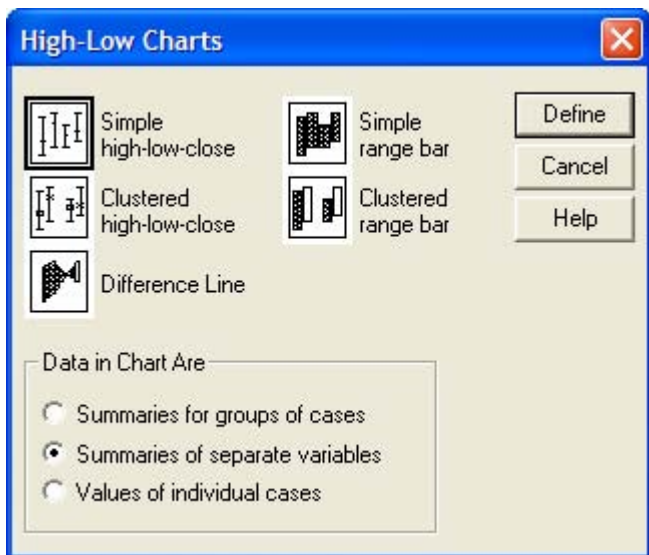
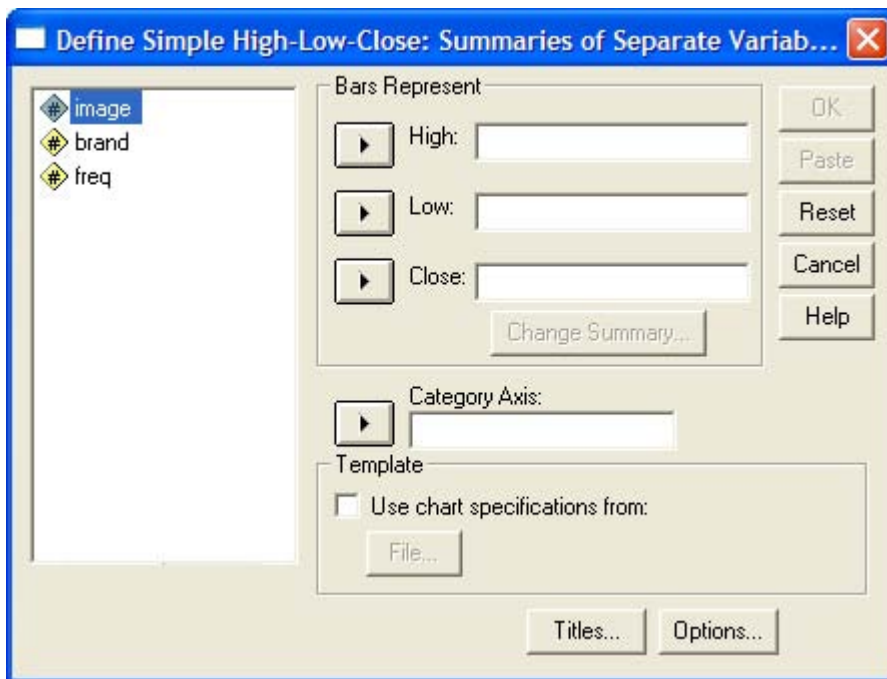


Рис. 22.31: Диалоговое окно High-Low Charts (Диаграммы максимума-минимума)

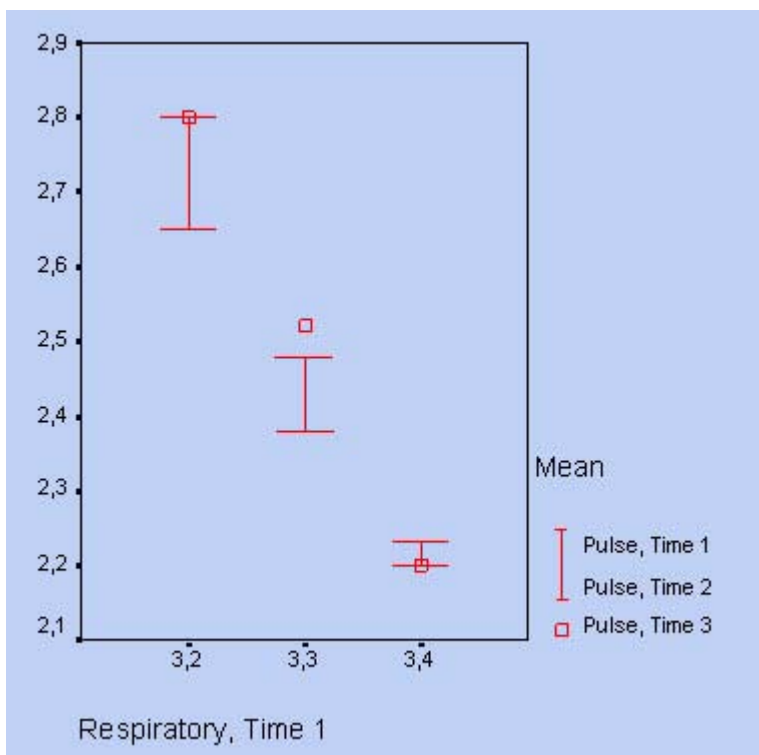
День	Максимальная котировка	Минимальная котировка	Окончательная котировка
1	164,35	161,48	162,33
2	166,12	163,03	164,12
3	167,84	164,75	165,97
4	167,79	163,93	166,13
5	171,14	/ 168,04	170,94
6	175,33	171,44	171,99
7	174,88	172,93	173,01
8	173,20	170,50	171,82
9	169,54	166,43	167,28
10	168,24	165,14	166,43

Эти данные построчно сохранены в четырёх переменных tag (день), hoch (максимум), tief (минимум) и ende (окончательная котировка) в файле aktien.sav.

- Откройте файл aktien.sav и в диалоговом окне High-Low Charts (Диаграммы максимума-минимума) щёлкните на области Simple High-Low-Close (Простая диаграмма — потолок-пол-закрытие).
- Установите метку возле опции Summaries of separate variables (Обработка отдельных переменных) и нажатием выключателя Define (Определить) откройте следующее диалоговое окно (см. рис. 22.32).
- В поле Category Axis: (Ось категорий) введите переменную tag и в соответствующие поля введите переменные hoch (High), tief (Low) и ende (Close). Оставьте установленную по умолчанию функцию Mean of values (Средние значения).
- С помощью выключателя Titles... (Заголовок), введите подходящий заголовок.
- Начните построение диаграммы щелчком на ОК.



**Рис. 22.32:** Диалоговое окно *Define Simple High-Low-Close: Summaries of Separate Variables* (Построение простой диаграммы — потолок-пол-закрытие: Обработка отдельных переменных)



**Рис. 22.33:** Простая диаграмма — потолок-пол-закрытие

### 22.5.2. Кластеризованные диаграммы - максимум-минимум-закрытие

При помощи этого метода осуществляется возможность представить несколько процессов потолок-пол-закрытие в одной диаграмме. Для реализации этой возможности в диалоговом окне High-Low Charts (Диаграммы максимума-минимума) щёлкните на области Clustered high-low-close (Кластеризованная диаграмма — максимум-минимум-закрытие).



### 22.5.3. Линейчатые диаграммы разностей

При помощи этой диаграммы может быть представлено взаимное изменение значений двух переменных, причём обе результирующие кривые могут пересекаться. Это пересечение как раз и может быть очень наглядно представлено с помощью линейчатых диаграмм разностей.

Нижеследующая таблица содержит данные о развитии рынка образования в Германии с 1985 по 1992 год.

Год	Количество учебных мест	
	Предложение	Спрос
1985	719.110	755.994
1986	715.880	730.980
1987	690.287	679.622
1988	665.964	628.793
1989	668.649	602.014
1990	659.435	559.531
1991	668.000	550.671
1992	721.756	608.121

- Откройте файл `lehre.sav`, в котором в переменными `jahr` (год), `angebot` (предложение) и `nachf` (спрос) хранятся необходимые нам данные.
- В диалоговом окне High-Low Charts (Диаграммы максимума-минимума) щёлкните на области Difference Line (Линия разностей). Установите метку возле опции Summaries of separate variables (Обработка отдельных переменных).
- Нажатием выключателя Define (Определить) откройте следующее диалоговое окно (см. рис. 22.34).
- В поле Category Axis: (Ось категорий) введите переменную `jahr` и в группе Differenced Pair Represents (Значения разностных пар) в поля 1 и 2 введите переменные `angebot` и `nachf`. Активируйте функцию суммы (Sum of values) с помощью кнопки Change Summary (Сменить процедуру обработки).
- С помощью выключателя Titles... (Заголовок), введите подходящий заголовок.
- Начните построение диаграммы щелчком на ОК.

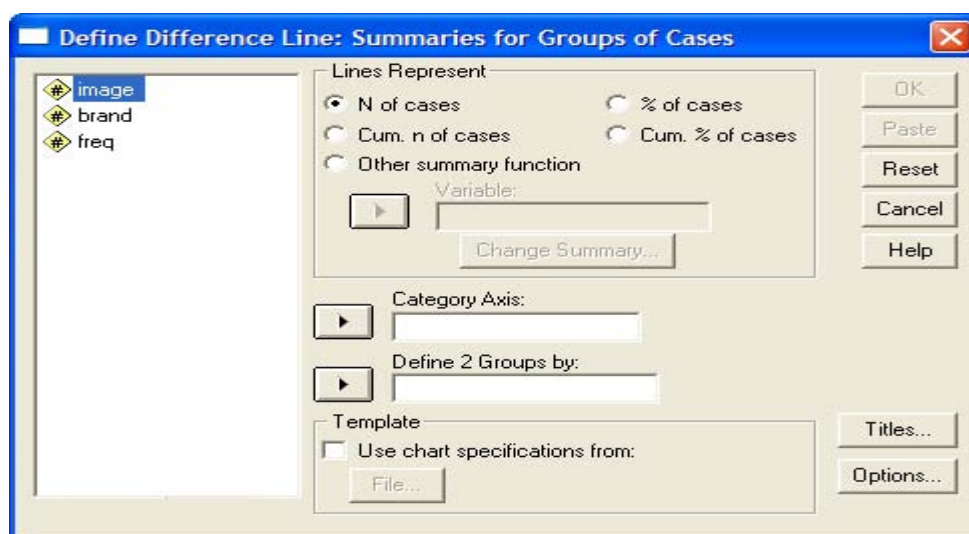


Рис. 22.34: Диалоговое окно *Define Difference Line: Summaries of Separate Variables* (Построение линейчатой диаграммы разностей: Обработка отдельных переменных).

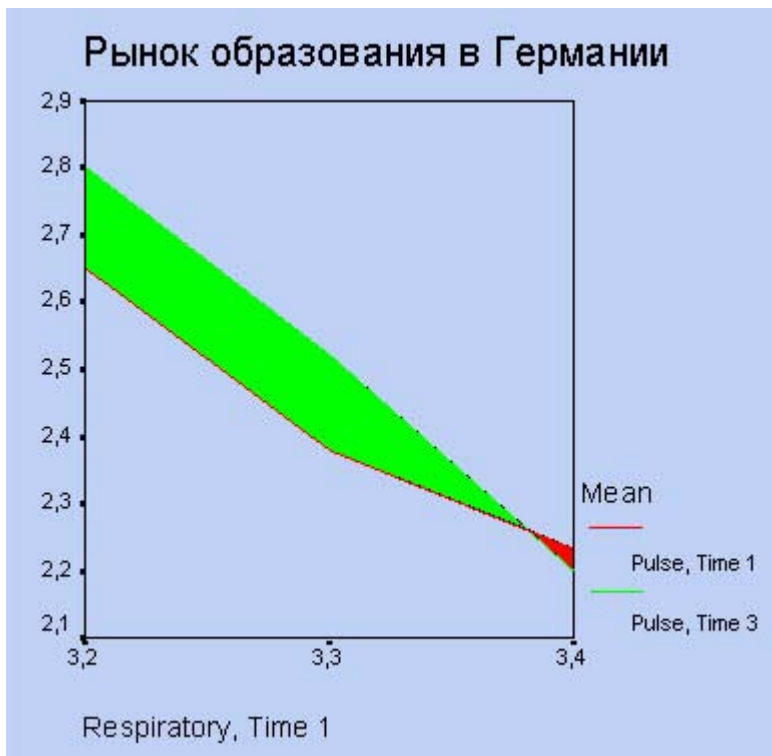


Рис. 22.35: Линейчатая диаграмма разностей

#### 22.5.4. Простые интервальные столбцы

Этот вид диаграммы является разновидностью простой диаграммы — потолок-пол-закрытие, в которой, однако, отображается только максимальное и минимальное значения, а окончательное отсутствует.

В качестве примера рассмотрим ситуацию, когда Вы, предположим, на протяжении десяти дней фиксировали свою максимальную и минимальную температуры:

День	Температура(°C)	
	Минимум	Максимум
14 марта 1994	2,4	11,3
15 марта 1994	2,6	11,5
16 марта 1994	3,7	12,4
17 марта 1994	6,2	14,8
18 марта 1994	6,2	14,8
19 марта 1994	1,9	9,7
20 марта 1994	4,3	11,3
21 марта 1994	7,6	13,4
22 марта 1994	7,0	12,9
23 марта 1994	6,3	11,0

Эти данные построчно сохранены в трёх переменных (tag (день), train (минимальная температура), tmax (максимальная температура)) в файле celsius.sav.

- Откройте файл celsius.sav и в диалоговом окне High-Low Charts (Диаграммы максимума-минимума) щёлкните на области Simple range bar (Простые интервальные столбцы).

- Установите метку возле опции Summaries of separate variables (Обработка отдельных переменных).
- Нажатием выключателя Define (Определить) откройте следующее диалоговое окно (см. рис. 22.36).
- В поле Category Axis: (Ось категорий) введите переменную tag и в группе Bar Pair Represents (Значения пары столбцов) введите переменные tmin и tmax в поля 1 и

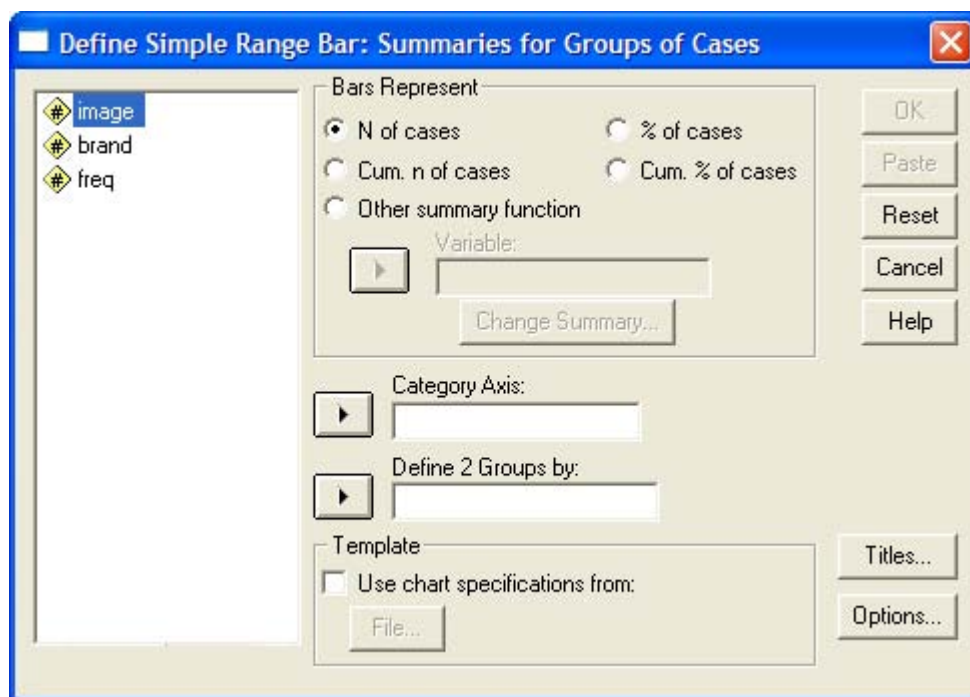
2. Установленную по умолчанию функцию Mean of values (Средние значения) можете оставить.

- С помощью выключателя Titles... (Заголовок), введите подходящий заголовок и начните построение диаграммы щелчком на ОК.

### 22.5.5. Кластеризованные интервальные столбцы

В одной диаграмме при помощи интервальных столбцов могут быть представлены и изменения нескольких переменных.

- Для этого в диалоговом окне High-Low Charts (Диаграммы максимума-минимума) щёлкните на области Clustered range bar (Кластеризованные интервальные столбцы).



**Рис. 22.36:** Диалоговое окно *Define Simple Range Bar: Summaries of Separate Variables* (Построение диаграммы с простыми интервальными столбцами: Обработка отдельных переменных)

## 22.6. Коробчатые диаграммы

Метод, при помощи которого, можно отобразить медиану и оба квартиля, минимальные и максимальные значения, а также пропущенные и экстремальные значения, уже рассматривался в главе 10.4.1. Эти диаграммы могут быть построены в ходе предварительного исследования данных или через меню графиков.

- После открытия необходимого Вам файла SPSS выберите в меню Graphs (Графики) *Boxplot...* (Коробчатые диаграммы) Откроется диалоговое окно *Boxplot* (Коробчатая диаграмма) (см. рис. 22.38).
- Вы можете выбрать простую или кластеризованную диаграмму, причём данные могут быть представлены в виде категорий одной переменной или в виде разных переменных.

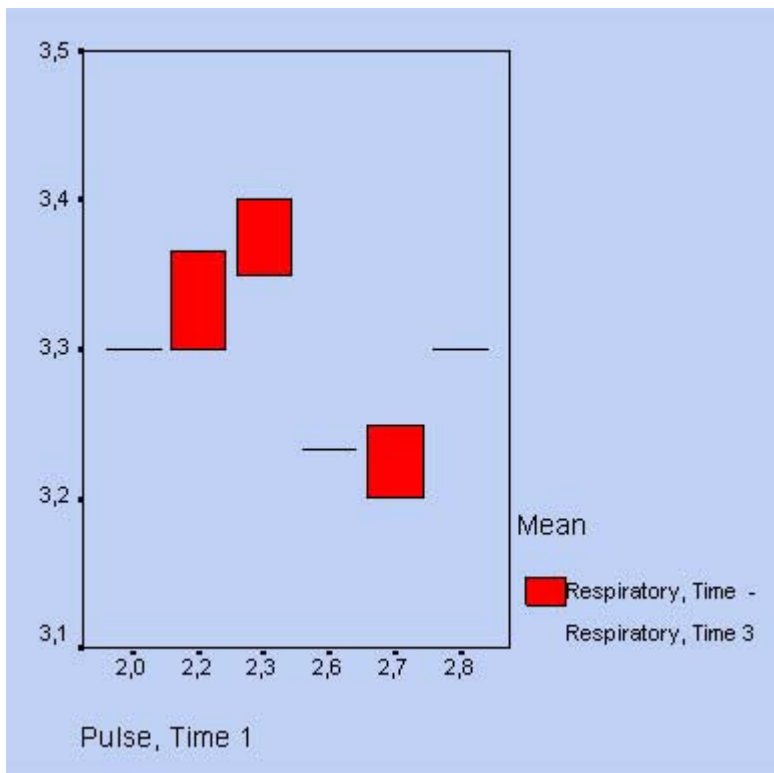


Рис. 22.37: Простые интервальные столбцы

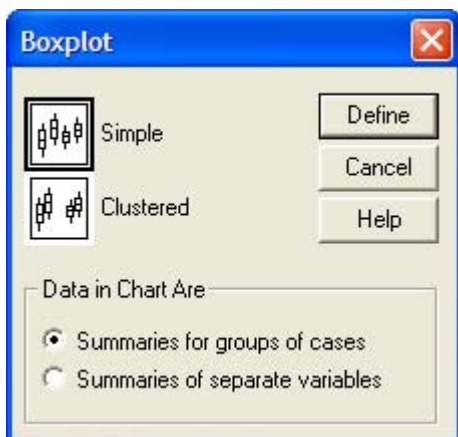


Рис. 22.38: Диалоговое окно Boxplot (Коробчатая диаграмма)

### 22.6.1. Простые коробчатые диаграммы

В рамках исследования гипертонии (файл hyper.sav) мы хотим для четырёх разных возрастных категорий (переменная ak) отобразить исходные показатели систолического кровяного давления (переменная rrs0).

- Откройте файл hyper.sav.
- В диалоговом окне Boxplot (Коробчатая диаграмма) щёлкните на области Simple (Простая) и оставьте опцию Summaries for groups of cases (Обработка категорий одной переменной), устанавливаемую по умолчанию.
- Щелчком по выключателю Define (Определить) откройте главное диалоговое окно, в котором в поле Category Axis: (Ось категорий) введите переменную ak, а в поле Variable: (Переменная) переменную rrs0. Если Вы введёте какую-либо переменную в поле Label Cases by: (Метки наблюдений), то её метки значений будут использованы для обозначения пропущенных и экстремальных значений.

- Начните построение диаграммы щелчком на ОК (см. рис. 22.39).

Если необходимо отобразить изменение систолического давления с течением времени, то для этого следует выбрать переменные rrs0, rrs1, rrs6 и rrs12.

- Щёлкните вновь на области Simple (Простая), но теперь поставьте маркер возле опции Summaries of separate variables (Обработка отдельных переменных).
- Щелчком по выключателю Define (Определить) откройте следующее диалоговое окно, в котором в поле Boxes Represent (Значения коробок) по очереди введите переменные rrs0, rrs1, rrs6 и rrs12.
- Вновь начните построение диаграммы щелчком на ОК (см. рис. 22.40).

На этой диаграмме метки отображаются не полностью, поэтому их ещё необходимо доработать.

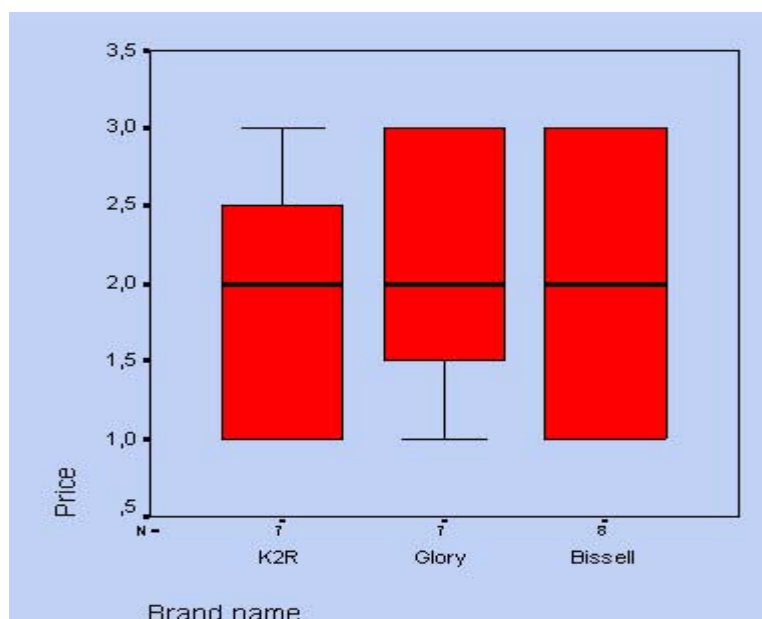


Рис. 22.39: Коробчатая диаграмма (категории одной переменной)

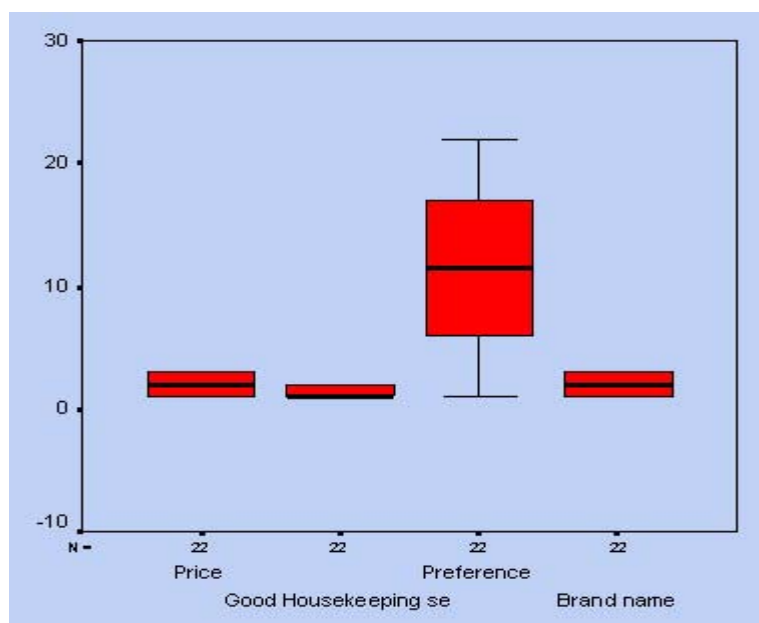


Рис. 22.40: Коробчатая диаграмма (разные переменные)

## 22.6.2. Кластеризованные коробчатые диаграммы

Вы можете использовать в данной диаграмме ещё одну переменную, тогда коробчатые диаграммы будут сгруппированы по категориям этой переменной.

- Для этого в диалоговом окне Boxplot (Коробчатая диаграмма) щёлкните на области Clustered (Кластеризованная).

## 22.7. Столбики ошибок

Если при помощи коробчатой диаграммы представляются медиана и оба квартиля, то диаграмма столбцов по величинам ошибки служит для отображения средних значений и характеристик рассеяния (стандартное отклонение, стандартная ошибка или доверительный интервал — по выбору).

- После открытия необходимого Вам файла SPSS выберите в меню Graphs (Графики) Error Bar... (Столбики ошибок)

Откроется диалоговое окно Error Bar (Столбики ошибок).

Также как и для коробчатых диаграмм, Вы можете выбрать простую или кластеризованную диаграмму столбцов по величинам ошибки, причём данные могут быть представлены в виде отдельных категорий одной переменной или в виде разных переменных.

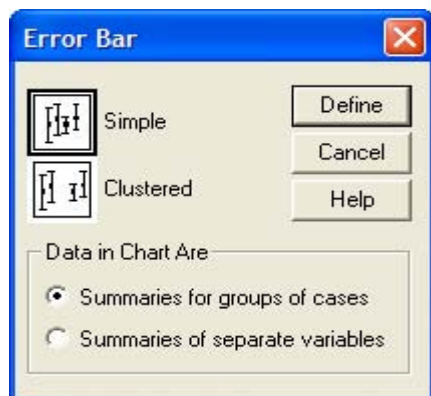


Рис. 22.41: Диалоговое окно Error Bar (Столбцы по величинам ошибки)

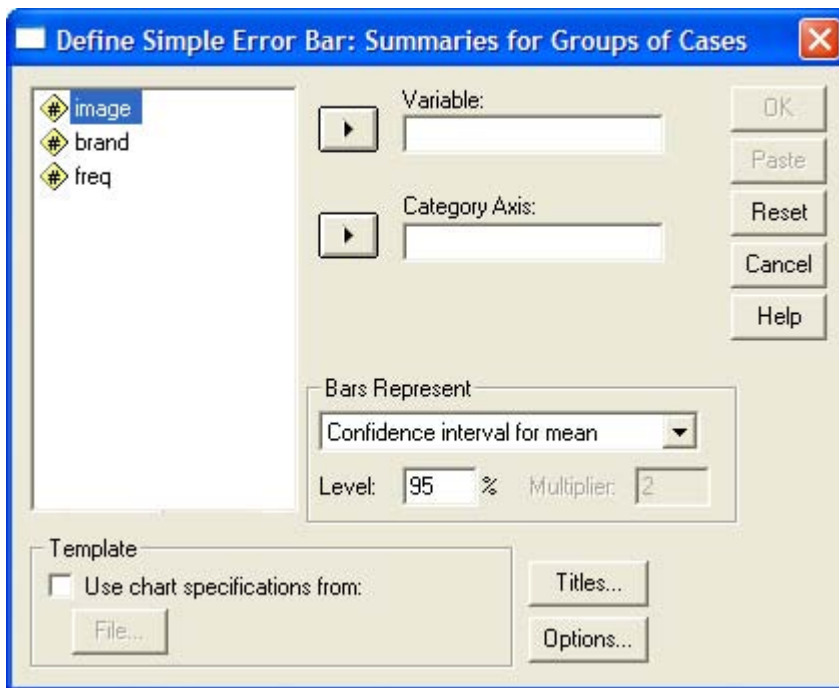
### 22.7.1. Простая диаграмма величины ошибки

В рамках исследования гипертонии (файл hyper.sav) для четырёх разных возрастных категорий (переменная ak) мы хотим отобразить исходные показатели уровня холестерина (переменная chol0).

- Откройте файл hyper.sav.
- В диалоговом окне Error Bar (Столбики ошибок) щёлкните на области Simple (Простая) и оставьте опцию Summaries for groups of cases (Обработка категорий одной переменной), устанавливаемую по умолчанию.
- Щелчком по выключателю Define (Определить) откройте соответствующее диалоговое окно (см. рис. 22.42).
- В поле Category Axis: (Ось категорий) введите переменную ak, а в поле Variable: (Переменная) переменную chol0.

В группе Bars Represent (Значения столбцов) Вам предлагаются на выбор следующие варианты:

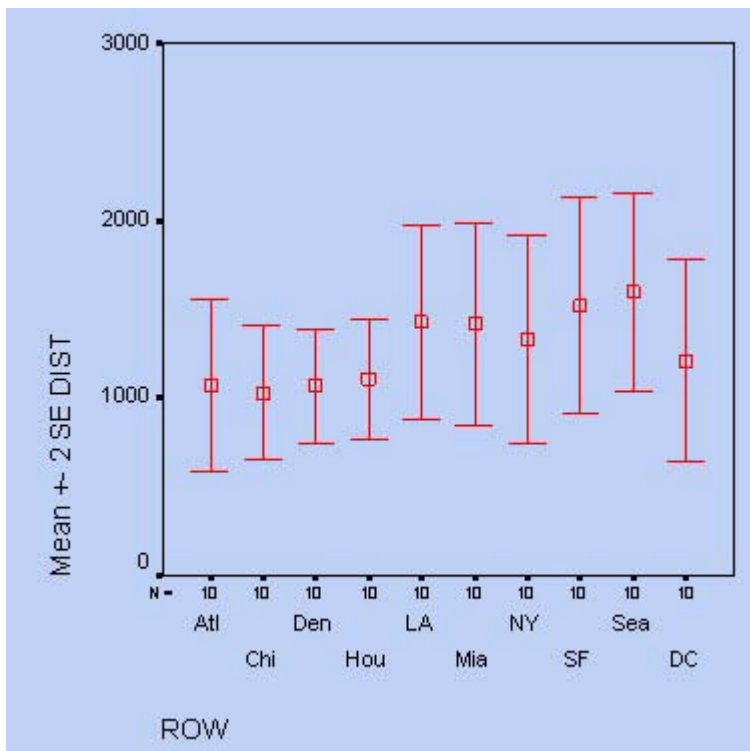
- Доверительный интервал для среднего значения (по умолчанию равен 95 %)



**Рис. 22.42:** Диалоговое окно *Define Simple Error Bar: Summaries for Groups of Cases* (Построение простой диаграммы величины ошибки: Обработка категорий одной переменной)

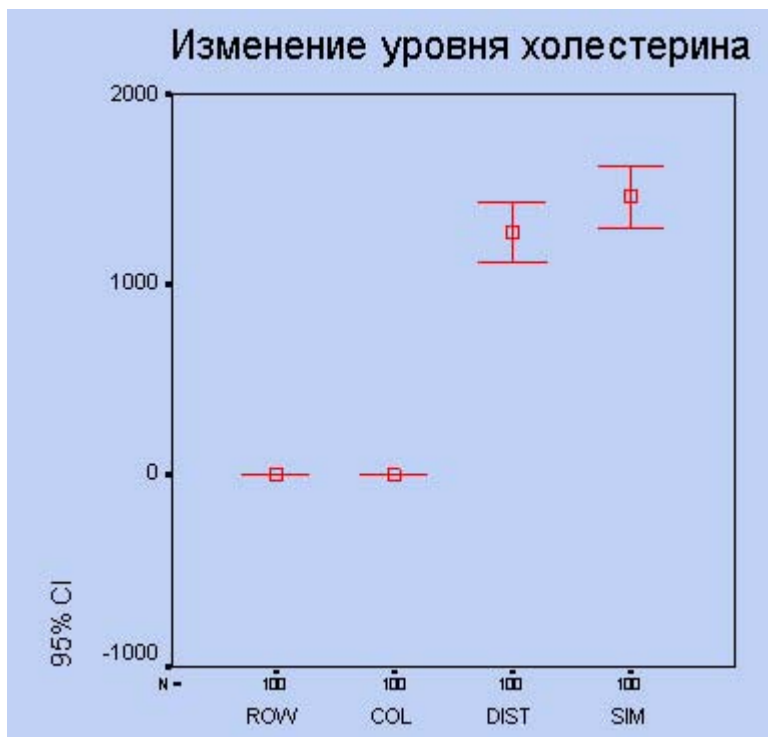
- Стандартная ошибка (предустановленный множитель равен 2)
- Стандартное отклонение (предустановленный множитель равен 2)
- Выберите отображение простого стандартного отклонения (множитель равен 2).
- С помощью выключателя *Titles...* (Заголовок), введите подходящий заголовок и начните построение диаграммы щелчком на *OK*.

Если необходимо отобразить изменение уровня холестерина с течением времени, то для построения графика необходимо использовать переменные *chol0*, *chol1*, *chol6* и *chol12*.



**Рис. 22.43:** Простая диаграмма величины ошибки (категории одной переменной)

- В диалоговом окне Error Bar (Столбики ошибок) вновь щёлкните на области Simple (Простая), но теперь поставьте маркер возле опции Summaries of separate variables (Обработка отдельных переменных).
- Щелчком по выключателю Define (Определить) откройте следующее диалоговое окно, в котором по очереди введите переменные chol0, chol1, chol6 и chol12 в поле Error Bars (Значения столбцов ошибок). В этом случае выберите отображение 95 % -го доверительного интервала (который является установкой по умолчанию).
- С помощью выключателя Titles... (Заголовок), введите подходящий заголовок и начните построение диаграммы щелчком на ОК (см. рис. 22.44).



**Рис. 22.44:** Простая диаграмма величины ошибки (разные переменные)

Метки значений на горизонтальной оси необходимо будет ещё подкорректировать.

### 22.7.2. Кластеризованная величина ошибки

Диаграммы величины ошибки можно объединять в группы при помощи дополнительных переменных.

- Для этого в диалоговом окне Error Bar (Величина ошибки) щёлкните на области Clustered (Кластеризованная).

## 22.8. Диаграмма рассеяния

Диаграмма рассеяния в графическом виде отображает отношения между двумя переменными, которые как минимум относятся к интервальной шкале. Пример диаграммы рассеяния уже был представлен в главе 15.

- Чтобы построить диаграмму рассеяния, после открытия необходимого Вам файла SPSS выберите в меню Graphs (Графики) Scatter... (Рассеяние)

Откроется диалоговое окно Scatterplot (Диаграмма рассеяния).



Имеются различные возможности построения диаграмм рассеяния. Для нижеследующих примеров взят файл euro.sav (можно сравнить с гл. 20), который содержит данные некоторых признаков для 28 европейских стран.

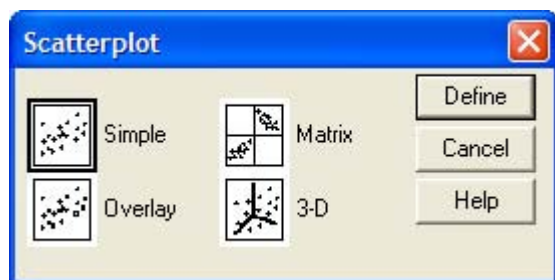


Рис. 22.45: Диалоговое окно Scatterplot (Диаграмма рассеяния)

### 22.8.1. Простая диаграмма рассеяния

- Откройте файл euro.sav.
- В диалоговом окне Scatterplot (Диаграмма рассеяния) щёлкните на области Simple (Простая).
- Щелчком по выключателю Define (Определить) откройте соответствующее диалоговое окно (см. рис. 22.46).

Мы хотим отобразить ожидаемую продолжительность жизни мужчин (переменная lem) в зависимости от урбанизации (процентного показателя доли городского населения, переменная sb).

- Переменную lem из списка исходных переменных перенесите в поле оси Y, а переменную sb — в поле оси X.

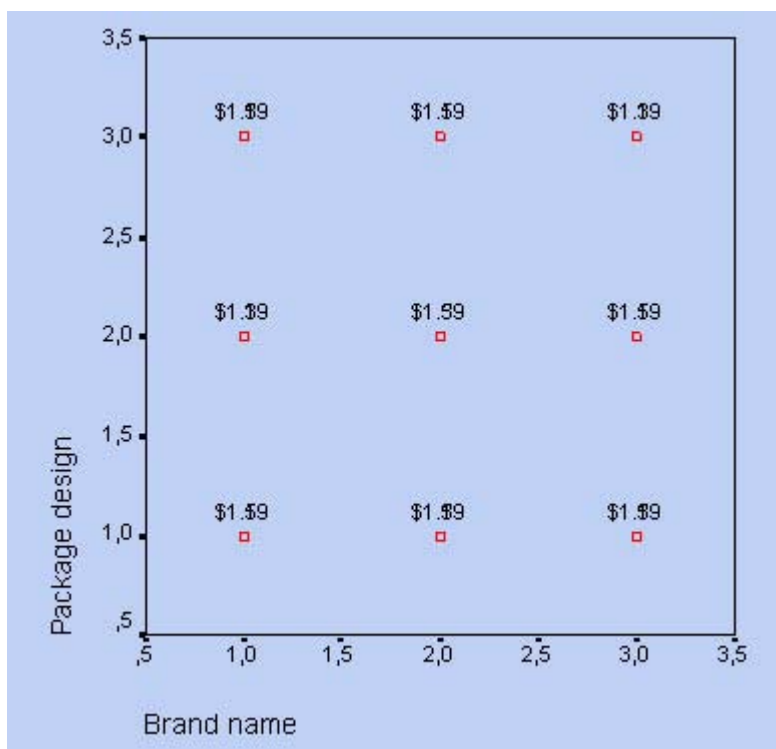
Если Вы поместите какую-нибудь переменную в поле Set Markers by: (Установить маркеры для:), то согласно принадлежности к этой переменной отдельные точки значений на диаграмме будут представлены окрашенными в другой цвет или помечены при помощи какого-либо отличительного маркировочного символа.

- Поместите переменную land в поле, предусмотренное для описания наблюдений (Label Cases by: (Метки наблюдений)). Значение этой переменной, соответствующее в приведенном примере сокращённому названию страны, будет размещено в диаграмме рассеяния вблизи соответствующей точки данных.



Рис. 22.46: Диалоговое окно Simple Scatterplot (Простая диаграмма рассеяния)

- Для этой цели щёлкните по выключателю Options... (Параметры) и в появившемся диалоговом окне активируйте опцию Display chart with case labels (Показать график с метками наблюдений).
- Пройдя выключатель Titles... (Заголовок), введите подходящий заголовок и начните построение диаграммы щелчком на ОК.




**Рис. 22.47:** Простая диаграмма рассеяния с метками случаев

Большое количество меток наблюдений приводит к снижению наглядности графика, поэтому можно рекомендовать оставить их только для избранных точек.

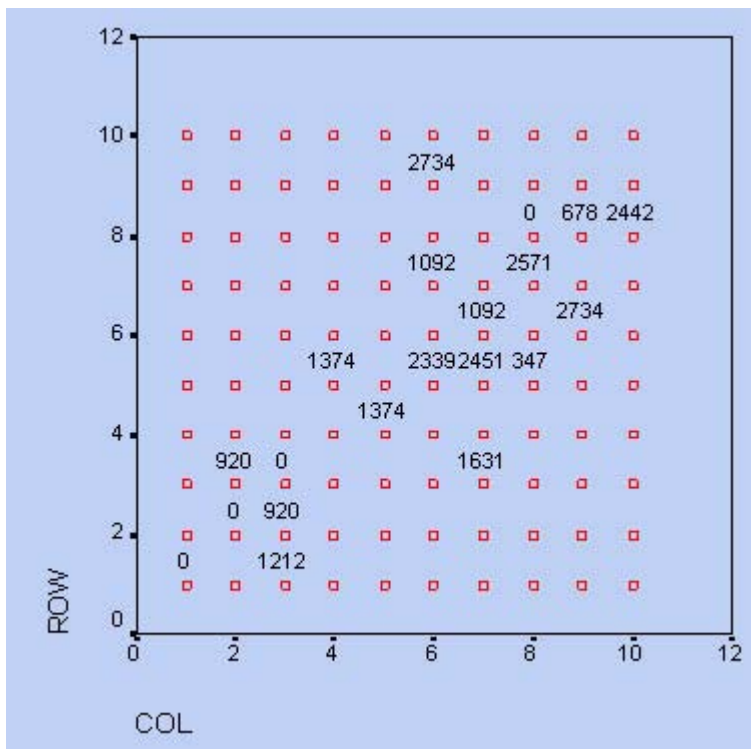
В качестве альтернативы на вооружение можно взять обозначение метками только наиболее характерных точек.

- Для этого постройте диаграмму заново.
- Через выключатель параметров уберите маркер опции Display chart with case labels (Показать график с метками наблюдений).

Теперь метки на графике присутствовать не будут.


- Двойным щелчком поместите график в редактор диаграмм.
- Одним щелчком по символу выбора точек  перейдите в режим выбора точек. Теперь при помощи курсора для выделения точек, Вы можете выбрать отдельные точки на диаграмме рассеяния и обозначить их метками.

Если несколько точек находятся очень близко друг к другу, то будет показан список меток, из которого Вы сможете выбрать необходимую метку.



**Рис. 22.48:** Простая диаграмма рассеяния с выборочными метками случаев

Численные показатели для любой точки, находящейся на диаграмме рассеяния также можно просмотреть в редакторе данных .

- Для этого при помощи курсора для выделения точек выберите нужную точку и в списке команд щёлкните на кнопке перехода в редактор данных: 

Вы увидите редактор данных. Изменения данных, вносимые в редакторе данных, естественно непосредственно не влияют на уже построенную диаграмму рассеяния.

В главе 22.17 мы покажем, как на одной диаграмме рассеяния можно отобразить четыре разных регрессионных линии (к примеру, регрессионные прямые).

### 22.8.2. Матричные диаграммы рассеяния

Этот метод применяется для отображения нескольких диаграмм рассеяния на одном графике.

- В диалоговом окне Scatterplot (Диаграмма рассеяния) щёлкните на области Matrix (Матрица).
- Щелчком на выключателе Define (Определить) откройте соответствующее диалоговое окно.

Переменные lem (ожидаемая продолжительность жизни мужчин), so (количество часов солнечной погоды в году) и nt (количество пасмурных дней в году) мы хотим попарно связать друг с другом.

- Для этого переменные lem, so и nt поочередно перенесите в поле, предусмотренное для матричных переменных.
- Начните построение диаграммы щелчком на ОК.

Число строк и столбцов в матричной диаграмме соответствует количеству переменных. Каждая ячейка является диаграммой рассеяния для одной пары переменных. Диагональные ячейки

содержат метки переменных, находящихся в соответствующих ячейках матрицы (в данном примере метки являются слишком длинными).

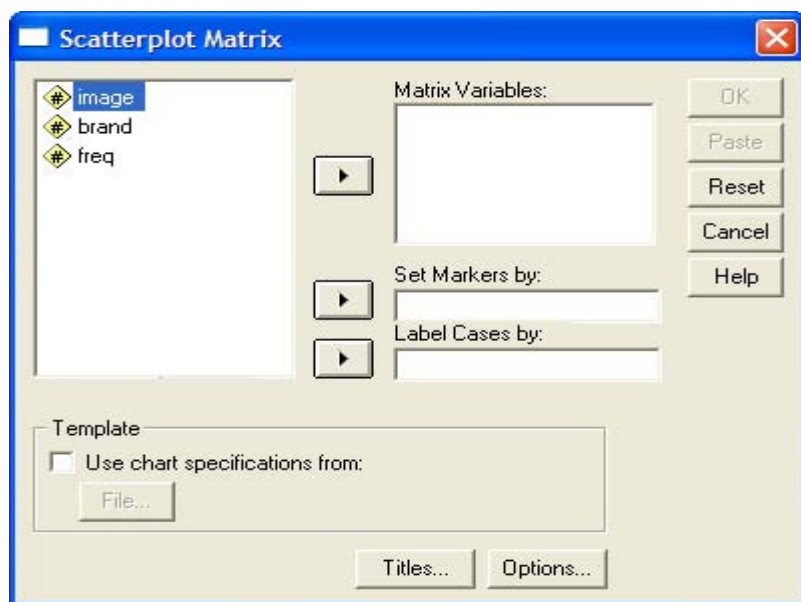


Рис. 22.49: Диалоговое окно Scatterplot Matrix (Матричная диаграмма рассеяния)

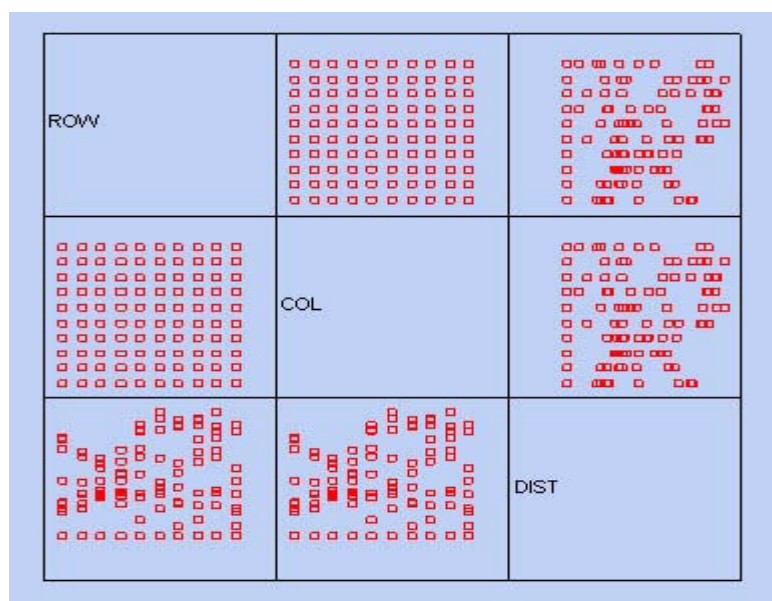


Рис. 22.50: Матричная диаграмма рассеяния

Первая диагональная ячейка содержит метку переменной km. Это означает, что для всех диаграмм первой строки эта переменная находится со стороны вертикальной оси (оси Y). Какая из переменных при этом откладывается по горизонтальной оси (ось X), следует узнавать из следующих диагональных ячеек. Такие же правила справедливы и для последующих строк.

К примеру, в центральном поле первой строки представлена взаимосвязь средней ожидаемой продолжительности жизни (по вертикали) и количества часов солнечной погоды (по горизонтали). Явно заметна обратная зависимость.

И в матричных диаграммах рассеяния можно задать маркировку для некоторой переменной, организовать вывод меток наблюдений, а также отображение любой другой необходимой информации; можно так же организовать построение различных линий регрессии (для сравнения см. разд. 22.17).

### 22.8.3. Наложённые диаграммы рассеяния

В одном графике можно представить несколько диаграмм рассеяния.

- Для этого в диалоговом окне Scatterplot (Диаграмма рассеяния) щёлкните на области Overlay (Наложение) и затем на кнопке Define (Определить).

В появившемся диалоговом окне могут быть заданы соответствующие X-Y-пары переменных, которые должны быть представлены вместе. Значения, принадлежащие соответствующей паре, на диаграмме будут отмечены одной определённой маркировкой.

Этот метод имеет смысл применять только тогда, когда речь идёт о переменных с одними и теми же областями значений.

### 22.8.4. Трёхмерные диаграммы рассеяния

Эти диаграммы строятся на основании значений трёх переменных и поэтому включают три оси.

По оси y откладывается высота положения точки

По оси x откладывается горизонтальное положение каждой точки

По оси z откладывается глубина положения каждой точки.

Отобразим переменную lem (средняя ожидаемая продолжительность жизни мужчин) на оси y, переменную sb (процентный показатель городского населения) на оси x и переменную so (количество часов солнечной погоды в году) на оси z.

- В диалоговом окне Scatterplot (Диаграмма рассеяния) щёлкните на области 3D (3-х мерная).
- Щелчком по выключателю Define (Определить) откройте соответствующее диалоговое окно (см. рис. 22.51).
- Перенесите поочерёдно переменные lem, sb и so из списка исходных переменных в поля принадлежащие осям y, x и z-
- Начните построение диаграммы щелчком на ОК.

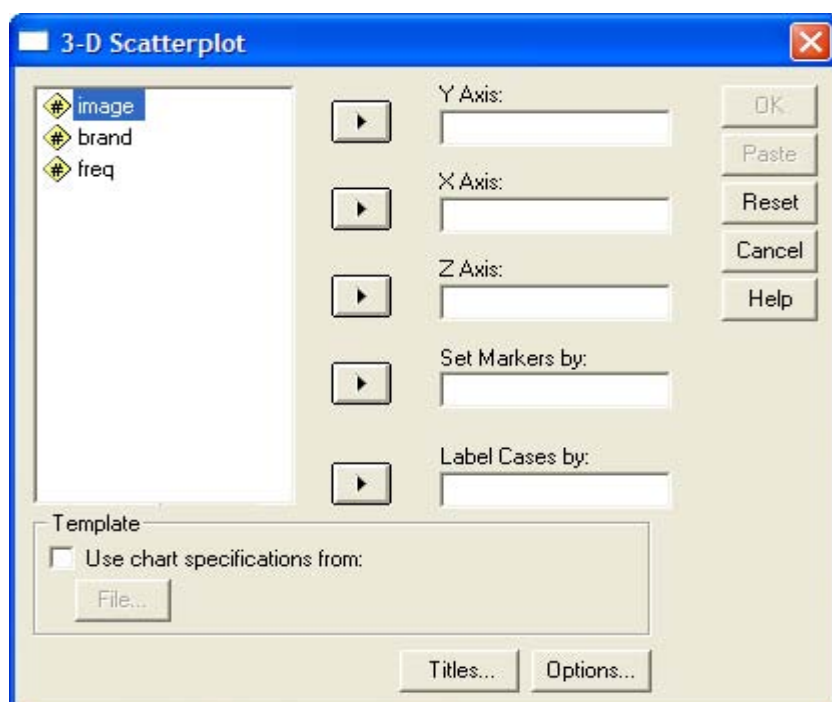


Рис. 22.51: Диалоговое окно 3-D Scatterplot ( Трёхмерная диаграмма рассеяния)

Очень длинные наименования осей при построении рисунка 22.52 были откорректированы.

И здесь Вы бы могли отметить маркировкой значения одной из переменных, а также указать наименования наблюдений и при помощи выключателя Titles... (Заголовок) дать диаграмме подходящее название.

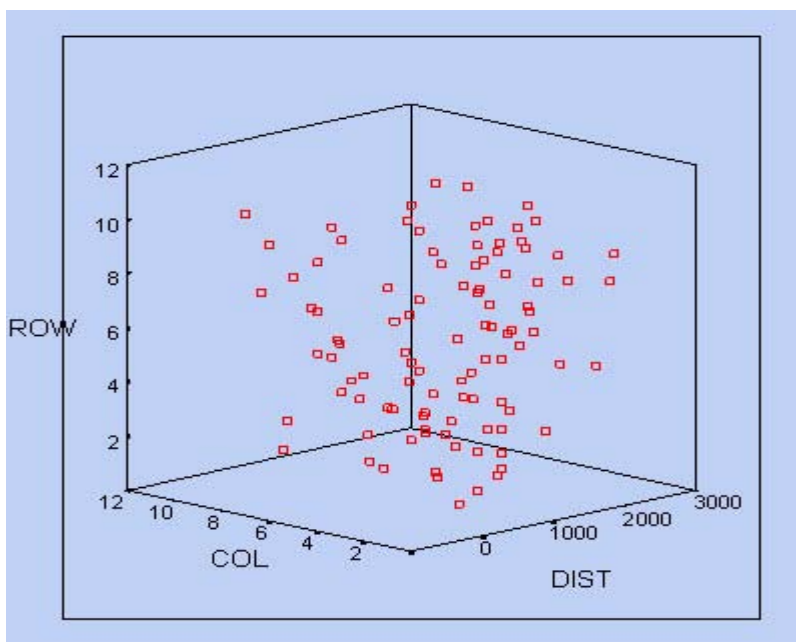


Рис. 22.52: Трёхмерная диаграмма рассеяния

## 22.9. Гистограммы

Гистограмма уже несколько раз рассматривалась в предыдущих главах.

- Чтобы построить гистограмму, после открытия необходимого Вам файла SPSS (к примеру, файла hyper.sav), выберите в меню Graphs (Графики) Histogram... (Гистограмма)

Откроется диалоговое окно Histogram (Гистограмма) (см. рис. 22.53).

С помощью гистограммы можно наглядно отобразить распределение переменных, относящихся по меньшей мере к интервальной шкале.

- Откройте файл hyper.sav.
- Поместите переменную chol0 в поле переменных и активируйте вывод кривой нормального распределения.
- Начните построение гистограммы щелчком на ОК.

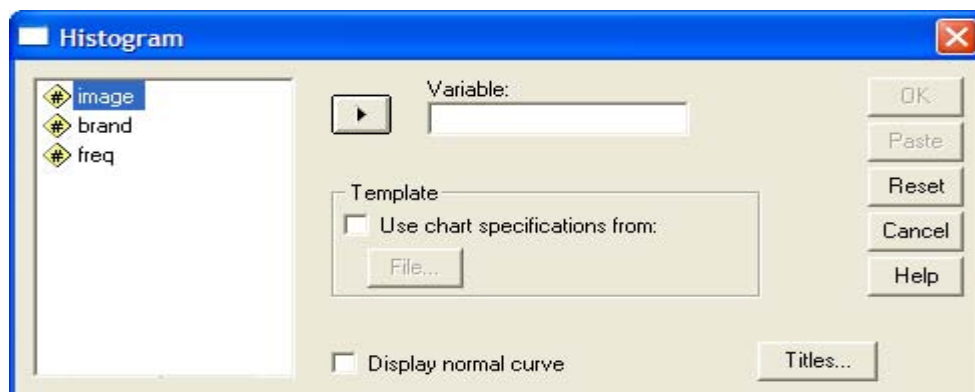
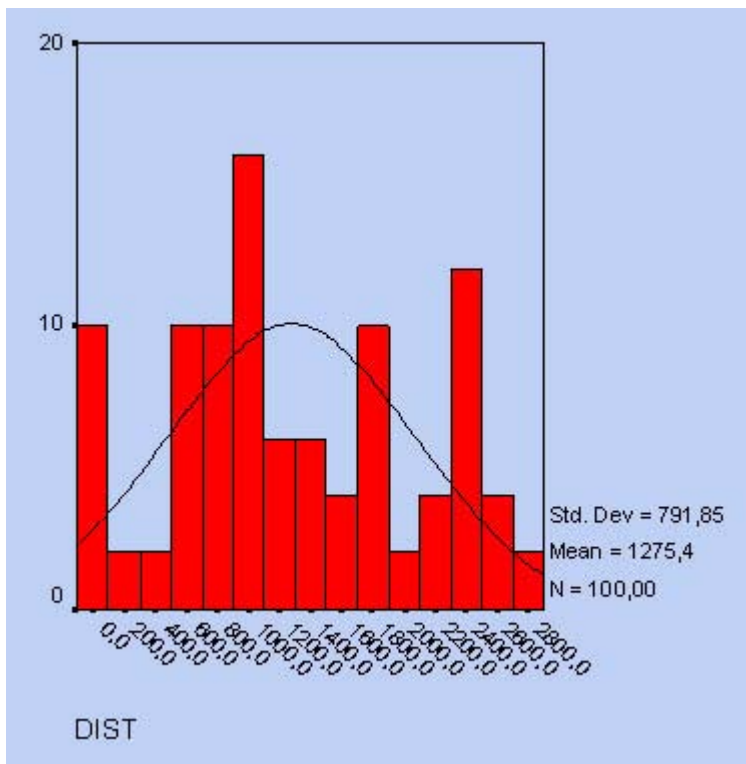


Рис. 22.53: Диалоговое окно Histogram (Гистограмма)



**Рис. 22.54:** Гистограмма с кривой нормального распределения

Чтобы выяснить, значительно ли отличается получившееся распределение от нормального, Вы не должны полагаться только на внешний вид гистограммы, а проверить его при помощи специального статистического теста. Для этого в SPSS реализован тест Колмогорова-Смирнова (см. разд. 14.5), который в данном случае указывает на незначимое отклонение от нормального распределения (значение  $p = 0,616$ ).

## 22.10. Диаграммы Парето

Диаграмма Парето представляет собой столбчатую диаграмму, в которой столбцы располагаются в порядке убывания, а дополнительная кривая может указывать на совокупную частоту для представленных категорий. При этом при суммировании отдельных столбцов по заданному правилу должна получаться некоторая итоговая величина, имеющая определенный смысл.

- Чтобы построить диаграмму Парето, после открытия необходимого Вам файла SPSS, выберите в меню Graphs (Графики) Pareto... (Парето)

Откроется соответствующее диалоговое окно.

Вы можете построить простую или состыкованную диаграмму Парето, причём и здесь существует три варианта представления данных.

Для иллюстрации процесса построения этих диаграмм достаточно одного примера. В следующей таблице приведены данные текущих расходов семей западной Германии в 1992 году.

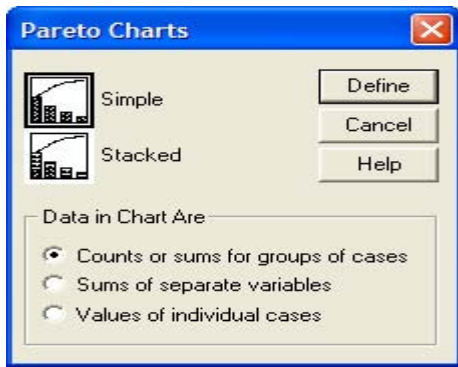


Рис. 22.55: Диалоговое окно Pareto Charts (Диаграммы Парето)

Статья расходов	Расколы (миллиарды DM)
Квартира	302,5
Одежда	116,2
Аренда	247,7
Электричество	55,6
Бытовые расходы	137,4
Здоровье	78,8
Проезд	253,5
Отдых	147,9
Прочее	108,5

- Откройте файл privver.sav, в котором построчно в переменных zweck (статья) и dm сохранены эти данные.
- В диалоговом окне Pareto Charts (Диаграммы Парето) щёлкните на области Simple (Простая) и оставьте опцию Counts or sums for groups of cases (Частоты или суммы категорий одной переменной), установленную по умолчанию.
- Нажатием выключателя Define (Определить) откройте следующее диалоговое окно.
- В поле Category Axis: (Ось категорий) введите переменную zweck. В группе Bars Represent (Значения столбцов) поставьте маркер рядом с опцией варианта выбора Sums of variable: (Суммы переменных) и переведите переменную dm в появившееся поле. Отображение совокупной (кумулятивной) кривой устанавливается по умолчанию.

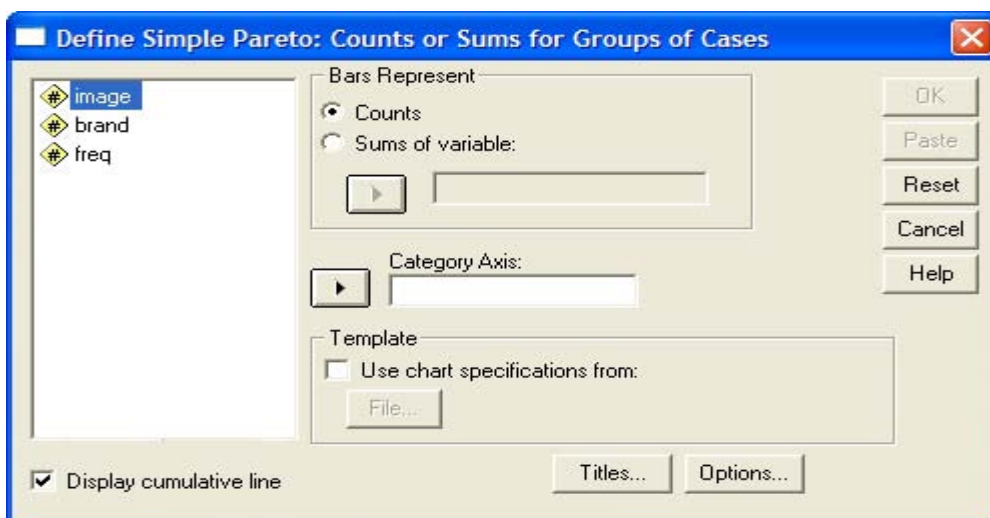


Рис. 22.56: Диалоговое окно Define Simple Pareto: Counts or Sums for Groups of Cases (Построение простой диаграммы Парето: Частоты или суммы категорий одной переменной)



- С помощью выключателя Titles... (Заголовок), введите подходящий заголовок.
- Щелчком на ОК начните построение диаграммы (см. рис. 22.57).

Из-за отображения кумулятивной (совокупной) кривой некоторые столбцы пришлось опустить довольно низко. В подобных случаях намного удобнее запретить отображение совокупной кривой. График без совокупной кривой Вы можете видеть на рисунке 22.58.

## 22.11. Контрольные карты

С помощью построения контрольных карт при наличии временной зависимости Вы можете проверить, лежат ли средние значения переменных в пределах области рассеяния, объясняемой действием случайных факторов, или же они выходят за пределы этой области. В общем случае подразделение данных может происходить не только по временным интервалам, а и посредством других подгрупп.

- После открытия необходимого Вам файла SPSS выберите в меню Graphs (Графики) Control... (Контроль)

Откроется диалоговое окно Control Charts (Контрольные карты).

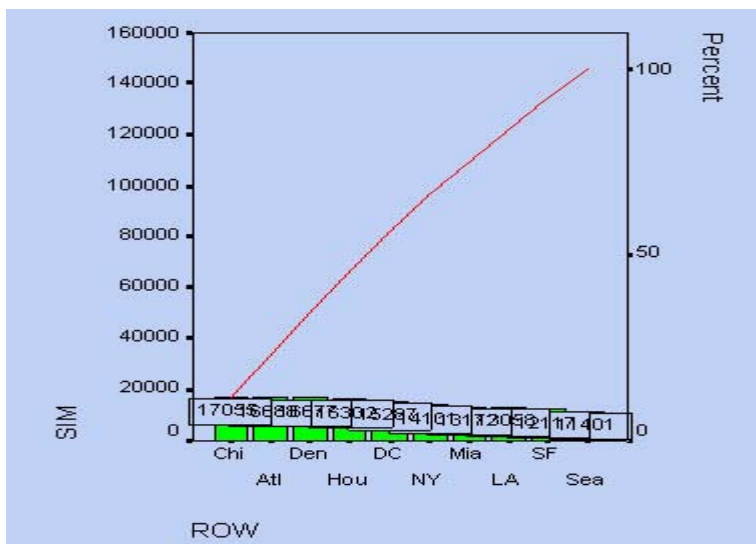


Рис. 22.57: Диаграмма Парето (с кумулятивной кривой)

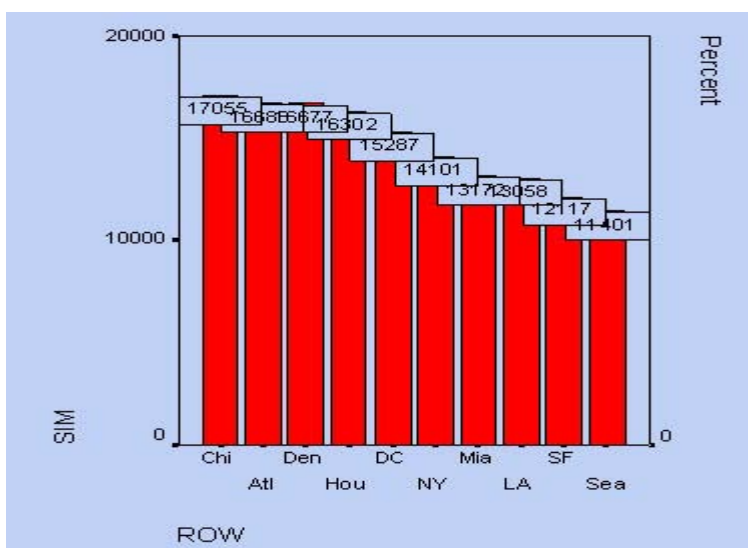


Рис. 22.58: Диаграмма Парето (без совокупной кривой)

Существует четыре разновидности контрольных карт и две возможности представления данных. Поэтому число возможных контрольных карт довольно велико и не может быть полностью рассмотрено в рамках этой книги. С одной стороны речь идёт об анализе средних значений, а с другой об анализе относительных частот переменных, относящихся к номинальной шкале.

Для рассмотрения этих диаграмм нам будет достаточно одного типичного примера. В этом примере необходимо проверить качество изделий, которые были произведены шестью станками за определённый промежуток времени. К примеру, необходимо произвести контроль длины этих изделий. Измерения длины изделий (в см) были произведены на шести станках для двенадцати промежутков времени и помещены в следующую сводную таблицу.

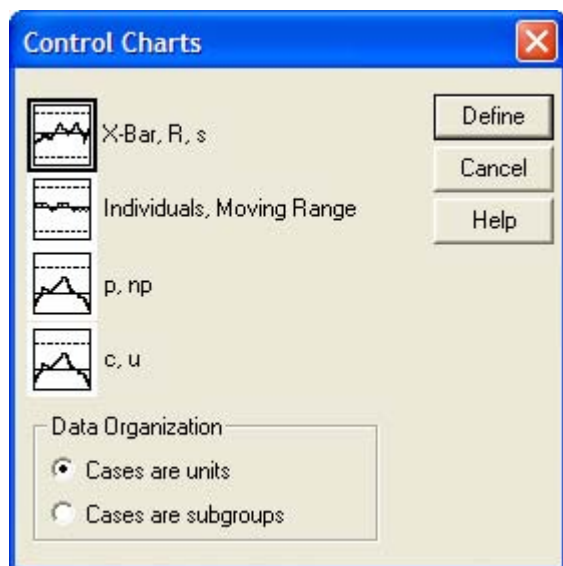


Рис. 22.59: Диалоговое окно Control Charts (Контрольные карты)

Интервал	Станок 1	Станок 2	Станок 3	Станок 4	Станок 5	Станок 6
1	24,07	24,11	24,17	24,02	24,07	23,95
2	23,98	24,09	24,03	24,18	24,10	24,20
3	24,14	23,99	23,93	24,06	24,04	24,10
4	23,96	24,10	23,97	23,90	24,00	23,91
5	23,98	24,02	24,00	24,05	23,84	23,95
6	24,01	23,95	23,97	23,83	24,12	24,02
7	23,98	24,05	24,16	24,07	23,90	24,00
8	24,07	24,12	24,07	24,14	23,99	23,96
9	24,11	24,16	24,22	24,12	24,00	24,05
10	24,05	24,04	23,90	24,10	24,10	23,97
11	24,00	24,08	23,97	23,87	23,92	24,06
12	24,07	24,01	23,89	24,04	23,92	24,09

- Откройте файл werk.sav.
- В диалоговом окне Control Charts (Контрольные карты) щёлкните на области X-Bar, R, s. Поставьте маркер рядом с опцией Cases are subgroups (Наблюдения используются в качестве подгрупп).
- Щелчком по выключателю Define (Определить) откройте соответствующее диалоговое окно (см. рис. 22.60).
- В поле Subgroups Labeled by: (Метки подгрупп:) введите переменную zeit (время), а в поле Samples (Образцы) переменные m1, m2, m3, m4, t5 и тб.

- Оставьте устанавливаемую по умолчанию функцию X-Bar and range (X-горизонталь диапазон) и щелчком на ОХ начните построение диаграммы (см. рис. 22.61).

На втором графике, который помещается в окне просмотра, будет отображено изменение стандартного отклонения.

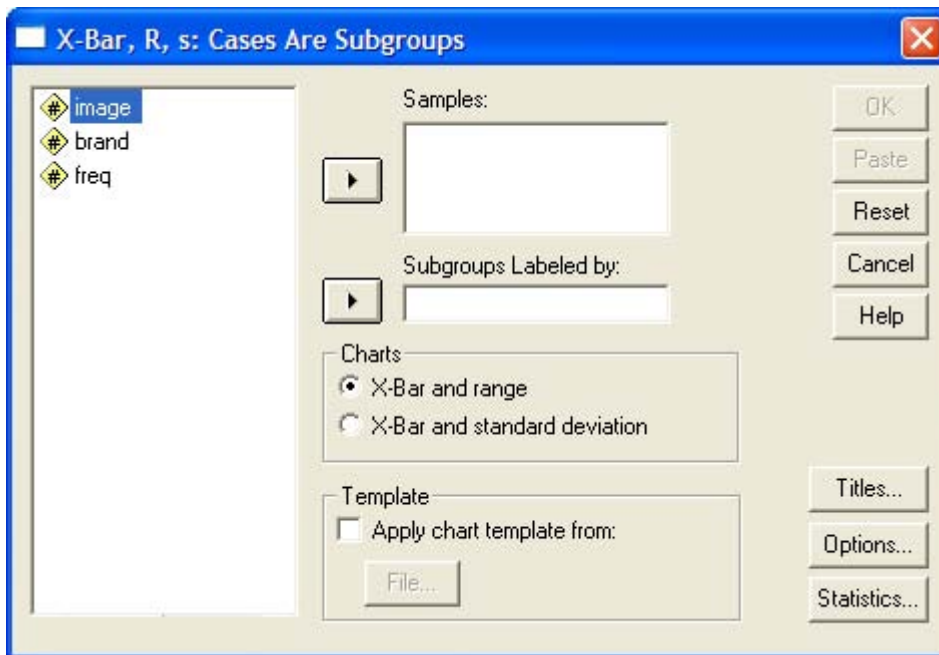


Рис. 22.60: Диалоговое окно X-Bar, R, s: Cases Are Subgroups (X-горизонталь, R, s: Случаи в качестве подгрупп)

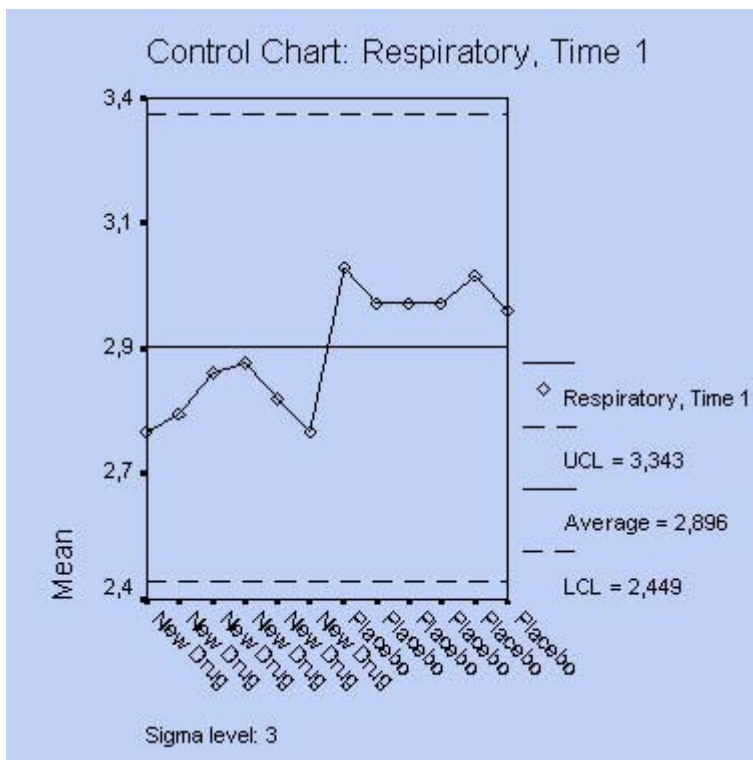


Рис. 22.61: Контрольная карта

## 22.12. Диаграммы нормального распределения

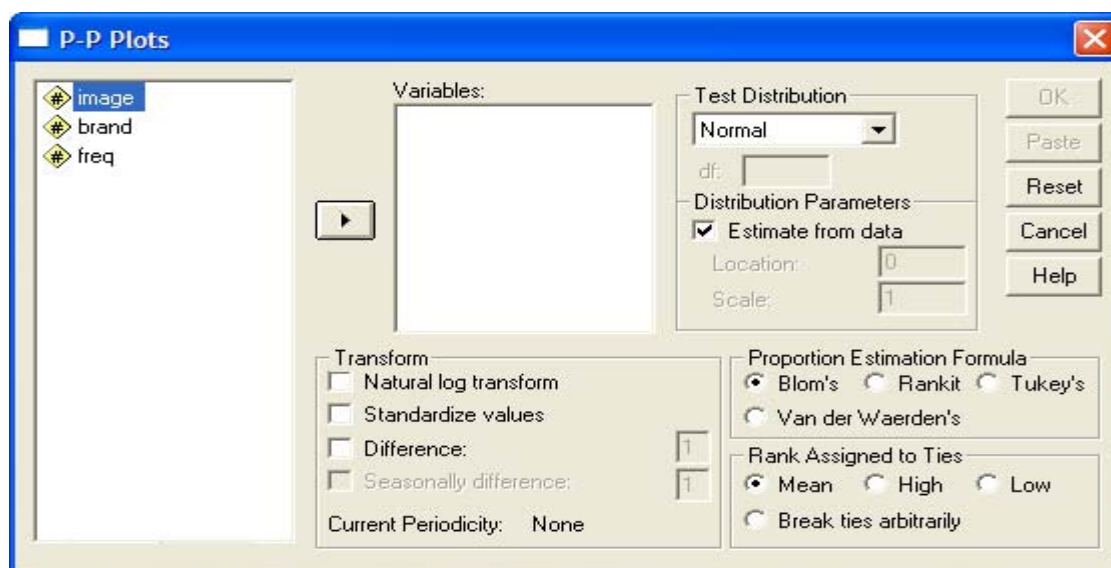
При проведении практически всех статистических тестов важную роль играет вопрос, подчиняются ли анализируемые данные нормальному распределению (для сравнения см. разд. 5.1.2). Проверку нормального распределения можно производить визуально, при помощи гистограммы (для пояснения см. разд. 22.9), однако лучше это осуществлять с использованием специального статистического теста, к примеру, теста Колмогорова-Смирнова (для получения подробной информации см. разд. 14.5). Ещё одну возможность анализа нормального распределения предоставляют диаграммы нормального распределения, которые в SPSS подразделяются на два вида:

- P-P- нормальный вероятностный график
- Q-Q-нормальный вероятностный график

В первом случае (P-P) в форме диаграммы рассеяния на графике отображается зависимость ожидаемых совокупных частот от фактических совокупных частот, а во втором случае (Q-Q) зависимость ожидаемой частоты от наблюдаемой частоты.

Построение диаграмм нормального распределения типа Q-Q можно производить и в рамках предварительного исследования данных. В таком варианте они уже были рассмотрены ранее (для получения подробной информации см. разд. 10.4.1). Поэтому здесь мы приведём пример, касающийся только диаграммы нормального распределения типа P-P.

- Откройте файл *hyper.sav* и выберите в меню *Graphs* (Графики) *P-P...* (P-P-диаграммы). Откроется диалоговое окно *P-P Plots* (P-P-диаграммы).



**Рис. 22.62:** Диалоговое окно *P-P Plots* (P-P-диаграммы)

Вы видите, что тест на нормальное распределение устанавливается по умолчанию. Наряду с этим Вы можете производить тестирование на предмет наличия ещё двенадцати видов распределения, к примеру, на наличие распределения Вайбулла (Weibull), Лапласа (Laplace), Хи-квадрат ( $\chi^2$ ) и *t*-распределения Стьюдента (Student). Вы можете просмотреть все предлагаемые типы распределений в ниспадающем меню.

- Мы хотим проверить на предмет нормального распределения переменную *a* (Alter — возраст); для этого перенесите эту переменную в поле тестируемых переменных.

В диалоговом окне присутствуют также и различные возможности преобразования данных, в состав которых входят: пересчет в натуральные логарифмы, z-преобразование (перевод к

стандартизованному виду) и два вида преобразований, применяемых для временных последовательностей.

Для подсчёта ожидаемых значений, подчиняющихся нормальному распределению, на выбор предлагаются четыре различных метода. Если количество значений, полученных в результате наблюдений, обозначить буквой  $n$ , а ранговые показатели этих значений буквой  $r$  ( $r = 1, \dots, n$ ), то формулы, соответствующие указанным методам, будут выглядеть следующим образом:

Blom (Блом):	$(r-3/8) / (n+1/4)$
Rankit (Ранговое преобразование):	$(r-1/2) / n$
Tukey (Тьюки):	$(r-1/3) / (n+1/3)$
Van der Waerden (Ван дер Верден):	$r / (n+1)$

Формула Блома (Blom) устанавливается по умолчанию. Далее Вам предоставляется возможность выбора одного из четырёх различных методов для обозначения одинаковых значений (так называемых связей).

Среднее значение:	Равным значениям присваивается средний ранг
Максимум:	Равным значениям присваивается ранг, высший из двух
Минимум:	Равным значениям присваивается ранг, низший из двух
Связи разрывать произвольно	Если в первых трёх методах для дальнейшего анализа используется только один элемент данных, то в этом методе может использоваться столько элементов, сколько значений имеется в наличии.

- Оставьте предварительные установки и подтвердите построение диаграммы нажатием ОК.

Будут построены две диаграммы. На первой, простой P-P-диаграмме отображается зависимость ожидаемых совокупных частот от фактических совокупных частот, рассчитанная при помощи формулы рангового преобразования Блома (Blom). На второй диаграмме, P-P-диаграмме без тренда, отображается разность между фактическими и ожидаемыми совокупными (кумулятивными) частотами в зависимости от фактических совокупных частот.

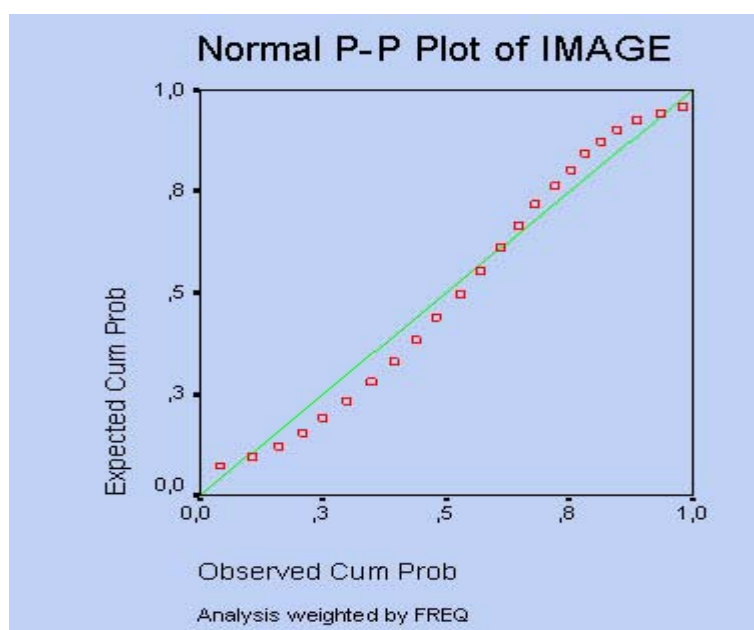


Рис. 22.63: Диаграмма нормального распределения типа P-P

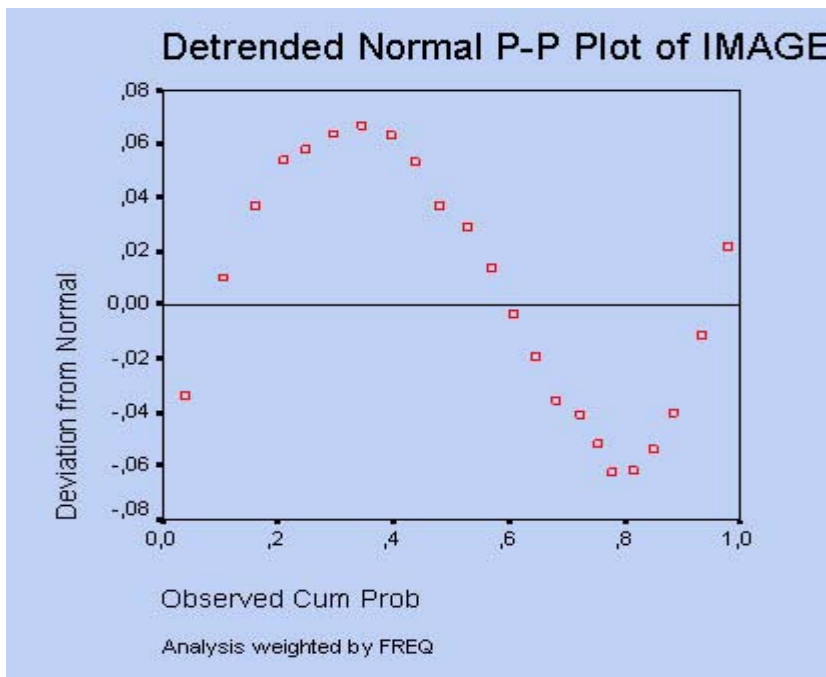


Рис. 22.64: Диаграмма нормального распределения типа P-P с исключённым трендом

### 22.13. Кривые ROC

Понятие кривых ROC (Receiver Operating Characteristic — функциональные характеристики приемника) взято из методологии анализа качества приёма сигнала (Signal Detection Analysis). Теория, стоящая за этим анализом, Theorie of Signal Detectability (TSD — "Теория определимости сигнала"), хотя и происходит первоначально из электроники и электротехники, но может также быть применена в области медицины, для анализа взаимодействия чувствительности и представительности диагностического теста. Поясним это при помощи примера.

В разделе 16.4 (Бинарная логистическая регрессия) было показано, каким образом при помощи переменных, соответствующих результатам T-типизации клеток, которые относятся к интервальной шкале, может быть спрогнозировано появление карциномы мочевого пузыря. Если вы посмотрите на обе группы (больных и здоровых), то заметите, что здоровые демонстрируют более высокие значения T-типизации ячеек, а больные скорее более низкие значения. Поэтому можно попытаться найти граничное значение T-типизации ячеек, которое будет чётко разделять обе группы больных и здоровых.

Это и было достигнуто при помощи метода бинарной логистической регрессии. Пройдём ещё раз тот путь, который мы проходили в главе 16.4.

- Откройте файл hkarz.sav.
- Выберите в меню Analyze... (Анализ) Regression. ..(Регрессия) Binary logistic... (Бинарная логистическая)
- В диалоговом окне Logistic Regression (Логистическая регрессия) переменную gruppe (группа) поместите в поле зависимых переменных, а переменную tzell — в поле ковариаций. Результаты теста LAI мы сначала не будем использовать в расчёте. При помощи выключателя Save... (Сохранить) организуйте сохранение прогнозируемой принадлежности к группе в виде дополнительной переменной. Начните расчёт нажатием ОК.

К исходному файлу данных добавилась переменная rgr\_1. Если Вы построите таблицу сопряженности между переменной gruppe (группа) в качестве строчной переменной и переменной rgr\_1 в качестве столбцової переменной, то получите следующий результат (для сравнения см. рис. 16.7):

## GRUPPE \* Predicted group Crosstabulation

(GRUPPE * Прогнозируемая группа таблица сопряженности)				
Count (Количество)				
		Predicted group (Прогнозируемая группа)		Total (Сумма)
		krank (Болен)	gesund (Здоров)	
GRUPPE	krank (Болен)	18	6	24
	gesund (Здоров)	4	17	21
Total (Сумма)		22	23	45

Среди 24 фактически больных 18 были верно расценены как больные (Rightly Positive (Верно положительный), RP), а 6 не верно отнесены к группе здоровых (Wrong Negative (Ложно отрицательный), WN). Из 21 фактически здорового человека 17 были верно отнесены к группе здоровых (Rightly Negative (Верно отрицательный), RN) и 4 не верно расценены больными (Wrong Positive (Ложно положительный), WP).

В качестве чувствительности теста выступает доля верно положительных предсказаний в суммарном количестве больных.

$$\text{Чувствительность} = \frac{RP}{RP + WN}$$

Эта величина характеризует способность теста как можно точнее отфильтровывать пациентов с сомнительным наличием болезни.

Под представительностью теста понимают долю верно отрицательных среди здоровых пациентов:

$$\text{Представительность} = \frac{RN}{RP + WP}$$

Эта величина характеризует способность теста обнаруживать исключительно пациентов с сомнительным наличием болезни. Для приведенного примера имеем

$$\text{Чувствительность} = 18 / (18 + 6) = 0,750$$

$$\text{Представительность} = 17 / (17 + 4) = 0,810$$

- Если при помощи меню Data (Данные) Sort Cases... (Сортировать наблюдения)

вы отсортируете данные по переменной tzcll, то заметите, что все наблюдения со значениями, лежащими ниже 66,5, отнесены к категории болен, а все наблюдения со значениями, находящимися выше 66,5, отнесены к категории здоров.

- Если Вы сместите граничное значение вниз или вверх и вновь рассчитаете чувствительность и специфичность, то результаты изменятся таким образом, что повышение чувствительности будет идти за счёт представительности, а повышение представительности за счёт чувствительности. Эту зависимость можно анализировать при помощи кривой ROC.
- Выберите в меню Graphs (Графики) ROC Curve... (Кривая ROC)

Откроется диалоговое окно ROC Curve (Кривая ROC)

- Переменной tzell присвойте статус тестируемой переменной, а переменной gruppe — статус переменной состояния. Под значением Value of State Variable: (Значение переменной состояния) понимается положительное значение, т.е. кодировка, соответствующая состоянию "болен". Введите в это поле 1. В группе Display (Показать) активируйте все имеющиеся опции.
- Щелчком по кнопке Options... (Параметры) откройте диалоговое окно ROC Curve: Options (Кривая ROC: Опции) (см. рис. 22.66).
- Активируйте опцию Smaller test result indicates more positive test (Меньший результат теста означает более положительный результат), так как в данном примере состоянию "болен" соответствует тенденция к уменьшению значений тестируемых переменных по сравнению с состоянием "здоров".

Результаты анализа, отображаемые в окне просмотра, приводятся ниже.

### Case Processing Summary (Обработанные наблюдения)

GRUPPE <sup>b</sup>	Valid N (listwise) (Действительные случаи (в соответствии со списком))
Positive <sup>a</sup> (Положительные)	24
Negative (Отрицательные)	21

Smaller values of the test result variable(s) indicate stronger evidence for a positive actual state (Низкие значения переменной(ых) указывают на скорее положительный результат теста).

a. The positive actual state is krank (Положительный результат теста соответствует состоянию болен).

b. The test result variable(s): TZELL has at least one tie between the positive actual state group and the negative actual state group (Результирующая переменная (переменные) теста: TZELL имеет по крайней мере одну связку между положительной и отрицательной группами).

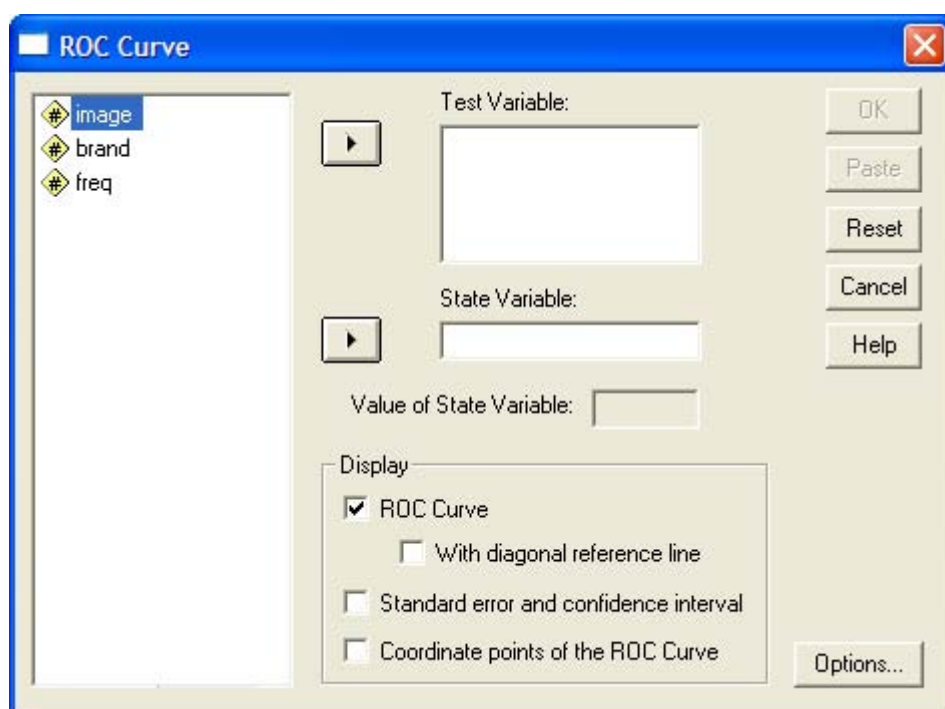


Рис. 22.65: Диалоговое окно ROC Curve (Кривая ROC)



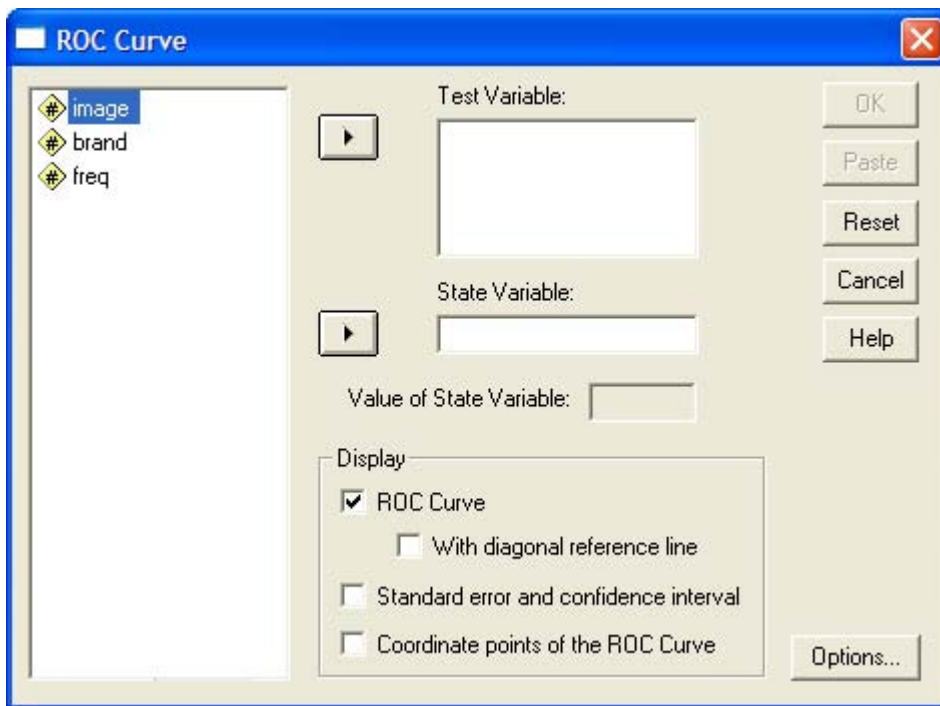
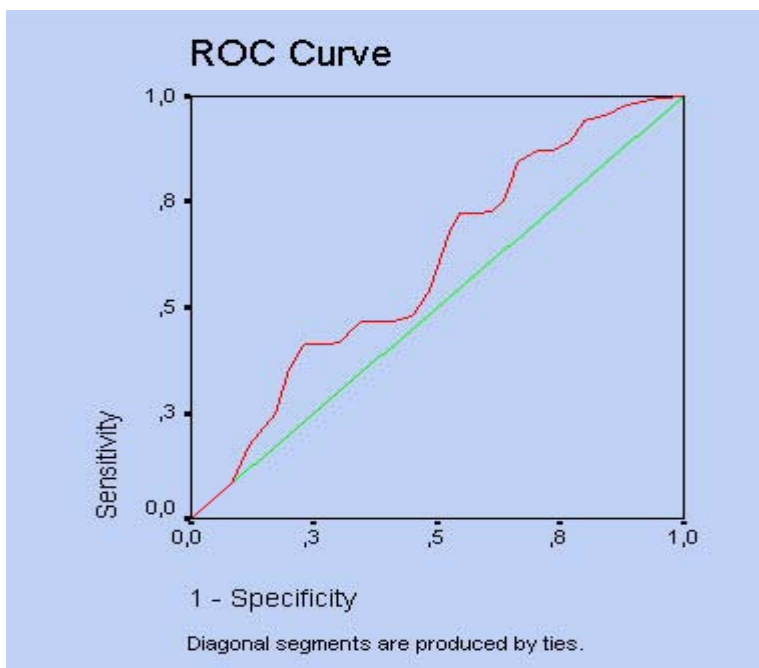


Рис. 22.66: Диалоговое окно ROC Curve: Options (Кривая ROC: Опции)



**Area Under the Curve (Площадь под кривой)**

Test Result Variable(s): TZELL (Переменная(ые) результата теста: TZELL)				
Area (Площадь)	Std. Error (Стандартная ошибка)	Asymptotic Sig.a (Асимптотическая значимость)	Asymptotic 95% Confidence Interval (Асимптотический 95 % доверительный интервал)	
			Lower Bound (Нижняя граница)	Upper Bound (Верхняя граница)
,849	,059	,000	,734	,964

The test result variable(s): TZELL has at least one tie between the positive actual state group and the negative actual state group (Результирующая переменная(ые) теста: TZELL имеет по крайней мере одну связь между положительной и отрицательной группами). Statistics may be biased (Статистики могут быть искажены (сдвинуты)).

a. Under the nonparametric assumption (В соответствии с непараметрическим предположением)

b. Null hypothesis: true area = 0.5 (Нулевая гипотеза: истинное значение площади = 0,5)

Coordinates of the Curve (Координаты кривой)

**Test Result Variable(s): TZELL (Результирующая переменная(ые) теста: TZELL)**

Positive if Less Than or Equal To a (Положительно, если меньше или равно)	Sensitivity (Чувствительность)	1 - Specificity (1-Представительность)
47,5000	,000	,000
52,0000	,042	,000
56,5000	,083	,000
58,0000	,125	,000
59,7500	,167	,000
61,0500	,208	,000
61,3000	,208	,048
61,7500	,292	,048
62,2500	,417	,048
62,0000	,458	,095
63,7500	,500	,095
64,7500	,542	,143
64,5000	,542	,190
65,7500	,625	,190
67,2500	,750	,190
68,7500	,792	,190
69,2500	,833	,190
69,7500	,833	,238
70,5000	,833	,333
71,2500	,958	,381
71,7500	,958	,476
72,2500	,958	,524
72,7500	,958	,571
73,2500	,958	,667
73,7500	1,000	,714
74,5000	1,000	,762
75,5000	1,000	,810
76,5000	1,000	,857
77,7500	1,000	,952
79,5000	1,000	1,000

The test result variable(s): TZELL has at least one tie between the positive actual state group and the negative actual state group (Результирующая переменная(ые) теста: TZELL имеет по крайней мере одну связь между положительной и отрицательной группами),

a. The smallest cutoff value is the minimum observed test value minus 1, and the largest cutoff value is the maximum observed test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values. (Минимальное разделяющее значение равно минимальному наблюдаемому значению теста минус 1, максимальное разделительное значение равно максимальному наблюдаемому значению теста плюс 1. Все остальные разделительные значения являются средними значениями двух соседних наблюдаемых значений теста.)

С помощью кривой ROC чувствительность и комплиментарное значения представительности приводятся к единице. Диагностируемое значение с нулевой степенью прогнозирования изображается здесь линией, наклоненной под углом 45 градусов (диагональю). Чем больше выгнута кривая ROC, тем более точным является прогнозирование результатов теста. Индикатором этого свойства служит площадь под кривой ROC, которая для теста с нулевой степенью прогнозирования равна 0,5, а для случая с максимальной степенью прогнозирования — 1. Для рассматриваемого примера получилось значение равное 0,849, причём 95 % доверительный интервал соответствует значениям площади, принадлежащим диапазону от 0,734 до 0,964.

В следующей таблице Вы можете увидеть чувствительность и представительность для различных граничных значений. Для граничного значения 67,5 Вы вновь встретите уже рассчитанные нами показатели.

## 22.14. Временные диаграммы и графики последовательностей

- Посторонние временных рядов и графиков последовательностей происходит посредством выбора меню Graphs (Графики) Time Series... (Временной ряд) и Graphs (Графики) Sequence... (Последовательность)

соответственно. В связи с тем, что в модулях SPSS, рассматриваемых в этой книге, отсутствует анализ временных рядов, мы не будем подробно останавливаться на этой диаграмме. Информацию по этому вопросу Вы можете найти в книге этих же авторов: 'SPSS. Методы исследования рынка и мнений'.

## 22.15. Основы редактирования графиков

Для того, чтобы разобраться во всех возможностях, которые SPSS для Windows предоставляет для редактирования графиков, наверняка потребуется некоторое время.

Построение графиков происходит при помощи большого количества процедур меню статистик и из меню графиков. Все графики, построенные таким образом, попадают сразу в окно просмотра. Отсутствует промежуточное сохранение, существовавшее вплоть до 6-ой версии SPSS.

Даже при построении Ваших первых графиков (теперь в SPSS они, как правило, называются диаграммами) можно не беспокоиться об их внешнем виде, поскольку в силу вступают соответствующие установки по умолчанию. Если Вы к тому же добавили некоторые наименования (заголовок, подзаголовок, сноски), то такой вид уже будет вполне достаточен для того, чтобы графики можно было использовать в большинстве практических ситуаций.

Если Вы хотите придать графикам более наглядный и презентабельный вид или же существует необходимость произвести определённые корректировки (к примеру, если метки переменных слишком длинные), то график следует перенести в редактор диаграмм. Для этого в окне просмотра дважды щёлкните в любом месте в области диаграммы.

В редакторе диаграмм Вы сможете производить над графиком следующие действия:

- корректировать (или изменить)
- сохранить график в каком-либо другом графическом формате
- сохранить как образец для других графиков и
- копировать в буфер обмена Windows.

Обзор всего многообразия возможностей дополнительной обработки, которые предлагает Вам редактор диаграмм, приводится в разделе 22.16. В разделе 22.17 рассматриваются три типичных примера редактирования.

## 22.16. Редактор диаграмм

Для того, чтобы график можно было изменить (доработать, редактировать), он должен быть помещён в редактор диаграмм. Это происходит после двойного щелчка на какой-либо точке в области диаграммы, находящейся в окне просмотра. Тогда редактор диаграмм будет выглядеть так, как на рис. 22.67.

Вверху редактора диаграмм присутствуют меню и две панели инструментов. Если Вы пройдёте курсором по кнопкам панелей инструментов, не нажимая их, то сможете увидеть краткое описание кнопок. При помощи кнопок верхней панели инструментов, Вы можете получить информацию о диалоговых полях, которые Вы заполняли в последних построенных диаграммах, перейти в редактор данных, в нём перейти к нужному Вам наблюдению; а также получить информацию об отдельных переменных.

Кнопки, стоящие во второй панели инструментов, преимущественно служат для вызова форматированных меню и будут рассмотрены в соответствующем разделе. Статистические, графические меню и меню помощи уже известны, и поэтому здесь они рассматриваться не будут.

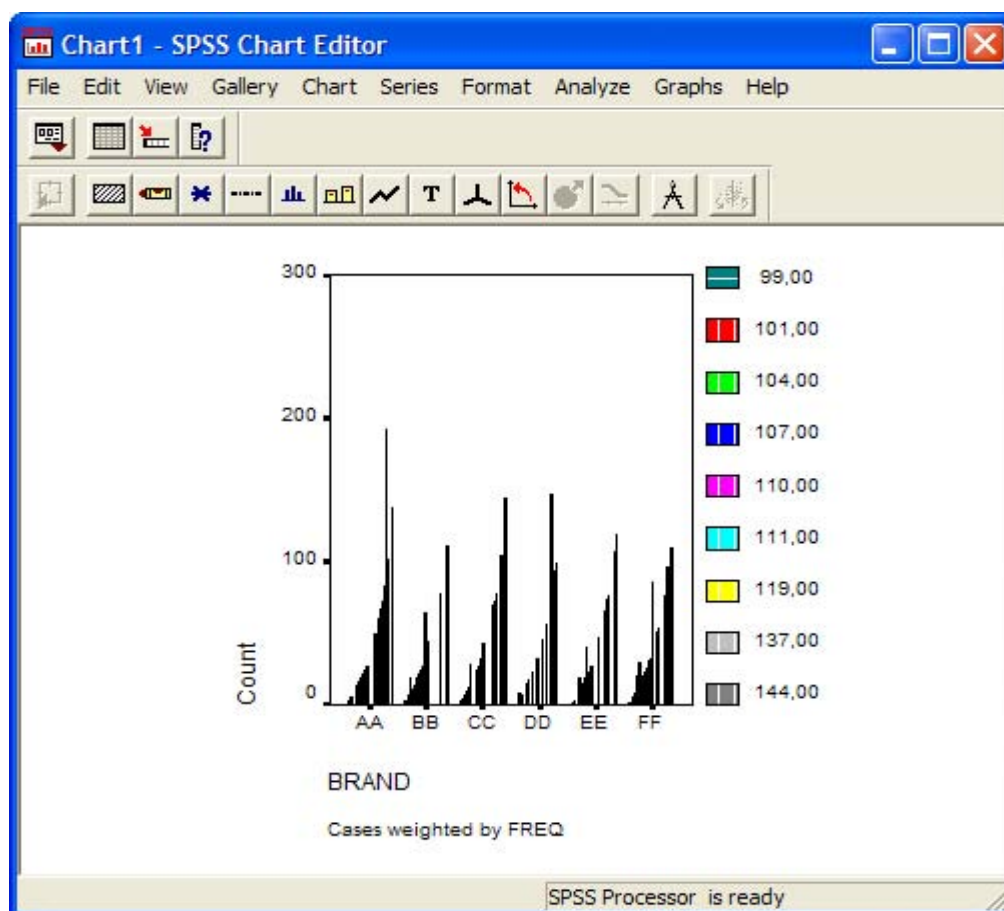


Рис. 22.67: Редактор диаграмм

- File (Файл): При помощи меню File (Файл) построенную диаграмму Вы можете сохранить, вывести на печать или скопировать свойства с некоторого графика-образца.
- Edit (Правка): При помощи меню Edit (Правка) Вы можете скопировать график в буфер обмена или изменить установки графика.
- View (Вид): В меню View (Вид) Вы можете включить или выключить строку состояния и управлять панелями инструментов.
- Gallery (Галерея): При помощи меню Gallery (Галерея) Вы можете выбрать другой тип графика для отображения ваших данных. Причём в списке Вы увидите некоторые дополнительные типы графиков, которые ещё не были рассмотрены, к примеру, смешанные диаграммы, диаграммы связывающих линий и разделённые круговые диаграммы.
- Chart (Диаграммы): Меню Chart (Диаграммы) служит для изменения внешнего вида диаграммы и элементов ее описания.

Пункты меню Options... (Параметры), Axis... (Оси) и Bar Spacing... (Расстояние между столбцами) являются специфическими для текущего типа диаграммы. После выбора этих опций открываются соответствующие диалоговые окна, содержание которых говорит само за себя.

- Series (Ряды): При помощи меню Series (Ряды) можно менять представление данных, то есть столбцы на линии или другие виды графического представления.
- Format (Формат): Если Вы щёлкните на этой кнопке, то получите список меню, представленный на рис. 22.68.

Большинство пунктов этого меню выведены на вторую панель инструментов. Вместо того, чтобы открывать меню, вы можете просто щёлкнуть на кнопке с соответствующим символом на панели инструментов.



Point Id (Выделение точек)

При помощи этой кнопки Вы можете менять режимы отображения точек на диаграмме рассеяния (для сравнения см. разд. 22.8.1)



Fill Pattern (Заливка узором)

Откроется диалоговое меню, в котором Вы можете выбрать необходимый рисунок из восьми образцов заливки для окрашивания замкнутых контуров, таких как: столбцы, области под линиями и области заднего плана.

Нужный объект выделяется щелчком на его поле. После этого на углах объекта должны появиться маркеры коррекции.

Вы выбираете необходимый тип заливки и щелчком на кнопке Apply (Применить) присваиваете его выбранному объекту.

Заливка белого цвета является прозрачной. Этот вид заливки следует выбирать тогда, когда некоторая последовательность данных должна быть показана на фоне другой последовательности.

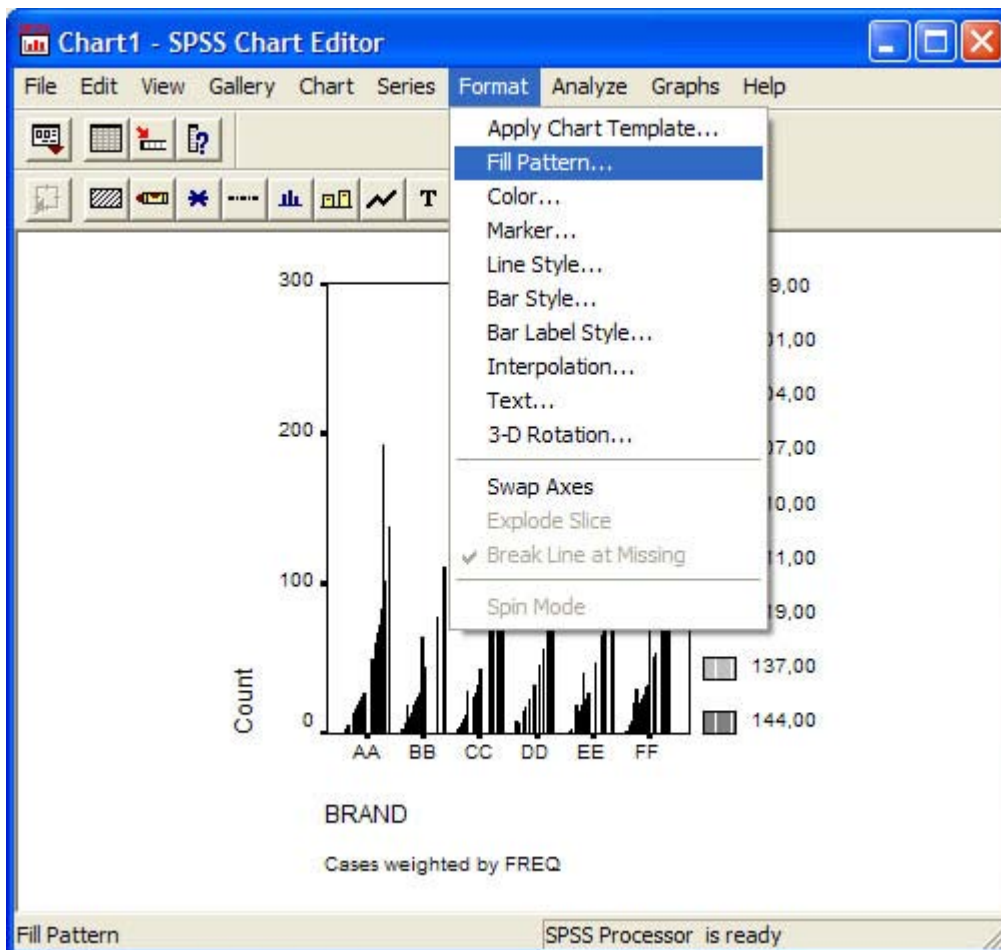


Рис. 22.68: Меню Format (Формат)

### Color (Цвет)

Для изменения цвета объекта графика (элемента представления данных или текста) выделите данный объект и выберите этот пункт меню. Откроется палитра с шестнадцатью различными цветами. Кому этого не достаточно, может открыть ещё одну дополнительную значительно более обширную палитру.

Выбором опций Fill (Заливка) и Border (Рамка) происходит переключение между возможностью изменить цвет объекта или рамки (контура) выделенного объекта.

Выберите одну из двух имеющихся опций. При помощи Apply (Применить) цвет будет перенесён на выделенный объект.

Чтобы расширить имеющуюся палитру цветов, щёлкните на кнопке Edit (Правка); после этого Вы сможете создать дополнительные или пользовательские цвета.

Если текущей палитре должен быть присвоен статус палитры по умолчанию, то щёлкните на выключателе Save as Default (Сохранить как палитру по умолчанию).

### Marker (Маркер)

Эта кнопка открывает палитру из 28-ми различных маркеров для обозначения положения точки данных на линейчатых диаграммах, диаграммах с областями и диаграммах рассеяния. Вы можете также установить один из четырёх предустановленных размеров маркеров.

Для изменения вида представления точек или рядов данных выделите сначала нужный элемент при помощи щелчка на графике. После этого на выделенном объекте появятся чёрные маркеры коррекции.

В группе Style (Стиль) выберите необходимую маркировку.

В группе Size (Размер) активируйте одну из опций предустановленных размеров маркеров. На экране разница между размерами отображаемых маркеров не значительна, но при печати она будет довольно хорошо заметна.

При помощи Apply (Применить) присвойте выделенному ряду данных маркеры с выбранными свойствами. Если Вы нажмёте кнопку Apply All (Применить для всех), то выбранный тип маркировки будет присвоен всем последовательностям данных.

Если изменения должны коснуться только размера маркеров, но не стиля маркировки, то следует деактивировать опцию Apply style (Применить стиль).

Если изменения должны коснуться только стиля представления маркеров, но не размера, то следует деактивировать опцию Apply size (Применить размер).

Маркеры на линейчатых диаграммах и диаграммах с областями становятся видимыми только в том случае, если их вывод будет задан в диалоговом окне Interpolation (Интерполяция). Это диалоговое окно вызывается из меню Format (Формат). Маркеры не могут быть заданы для изображения точек гистограмм и столбчатых диаграмм.



#### Line Style (Линии)

Здесь на выбор предлагаются четыре типа линий и четыре предустановленные толщины для этих линий.

На графике щелчком необходимо выделить линию, которую необходимо изменить. После этого на объекте появятся маркеры коррекции.

В группе Style (Стиль) выберите тип линии.

В группе Weight (Толщина) присвойте необходимую толщину выбранному типу линии.

После щелчка на кнопке Apply (Применить) выбранная конфигурация линии будет присвоена активному объекту. Эта кнопка остаётся неактивной, если выделены данные, которые не могут быть представлены на графике при помощи линии или элемента, содержащего линии (рамки, оси).



#### Bar Style (Столбцы)

Эта опция служит для изменения представления столбцов в графиках, содержащих столбцы. Некоторые типы столбцов не могут применяться для гистограмм.

Программа предлагает в Ваше распоряжение несколько типов столбцов. Если выбраны столбцы с тенью (Drop shadow) или с 3D-эффектами (3D-effect), то для этих типов столбцов дополнительно ещё может устанавливаться и толщина (Depth). Эта опция управляет толщиной сторон и верхнего торца столбца. Толщина при этом указывается в процентах от ширины столбца. При положительных значениях параметра Depth (Толщина) эффект строится начиная с правой стороны столбца, как показано на рисунках соответствующих опций, а при отрицательных значениях — с левой стороны столбца.

Если Вы нажмёте кнопку Apply All (Применить для всех), то установленные свойства будут применены ко всем столбцам. Эта кнопка становится активной только тогда, когда в редакторе диаграмм находится столбчатая диаграмма или интервальная столбчатая диаграмма.



#### Bar Label Style (Метки столбцов)

Программа предлагает три варианта идентификации столбцов при помощи числовых значений.

Если выбран один из стилей оформления числового значения (кроме None), то на каждом столбце появляется числовое значение, соответствующее высоте этого столбца. Для столбчатой диаграммы с областями метки столбцов указываются сверху и снизу каждого столбца. Три опции представленные в диалоговом окне Bar Label Styles (Метки столбцов) определяют внешний вид метки на столбце. Если Вы применяете тёмные цвета или узоры, в таком случае рекомендуется выбирать опцию Framed (В рамке), числовое значение в рамке будет лучше читаться.

Если Вы нажмёте кнопку Apply All (Применить для всех), то установленные свойства метки будут применены ко всем столбцам. Эта кнопка становится активной только тогда, когда в редакторе диаграмм находится столбчатая диаграмма, интервальная столбчатая диаграмма или гистограмма.



#### Interpolation (Интерполяция)

В данном диалоговом окне задаются различные возможности и методы для соединения точек данных.

Эта опция может применяться для диаграмм с областями, линейчатых диаграмм, линейчатых диаграмм разностей, для последовательностей средних значений в диаграммах величины ошибки, для заключительных показателей на диаграммах максимальных и минимальных значений, а также в диаграммах рассеяния (исключая 3D-диаграммы рассеяния).

На графике щелчком выделите линию или последовательность данных. После этого на каждом объекте появятся маркеры коррекции.

В группе Line Interpolation (Вид интерполяционной линии) выберите один из методов соединения точек при помощи некоторой кривой. Если SPSS должна рассчитать регрессионную прямую для диаграммы рассеяния, выберите в меню Chart (Диаграммы) пункт Options (Параметры).

Если Вы нажмёте кнопку Apply All (Применить для всех), интерполяция будет применена ко всем последовательностям данных. При помощи Apply (Применить) интерполяция будет применена только к объектам, выделенным в данный момент. Если Вы выделили данные, которые не могут быть отображены на графике при помощи линии, кнопка Apply (Применить) становится неактивной.

Если активировать опцию Display markers (показать маркеры), то для каждой точки выделенной кривой будет отображена маркировка. Тип маркера может быть выбран при помощи опции Marker (Маркер), находящейся в меню Format (Формат).

Существуют следующие виды интерполяции:

- None (Отсутствует): при выборе этой опции соединение между точками отсутствует.
- Straight (Прямая): точки последовательно соединяются прямой линией в том порядке, в котором они находятся в файле данных.
- В списке Steps (Шаги) Вы можете выбрать один из альтернативных методов построения ступенчатой интерполяции. Эти методы соответствуют шаговым функциям, в которых точки данных соединяются с левых сторон, в центрах или с правых сторон шагов, в зависимости от того была ли выбрана опция Left step (Левый шаг), Center step



(Центральный шаг) или Right step (Правый шаг). Шаги между собой соединяются вертикальными отрезками.

- В списке Jump (Прыжок) может быть выбран один из методов скачкообразной интерполяции. Скачкообразные методы строятся точно так же, как и пошаговые, но в них отсутствуют вертикальные соединения. В зависимости от выбора Left jump (Прыжок слева) Center jump (Прыжок по центру) или Right jump (Прыжок справа) точки данных будут лежать с левой стороны, по середине или с правой стороны горизонтальных отрезков.
- В списке Spline (Сплайн) может быть выбран один из методов соединения точек данных при помощи кривой.

— при выборе опции Spline (Сплайн) для соединения точек данных между собой строятся кубические сплайны.

— при выборе опции 3rd-order Lagrange (Лагранж 3-го порядка) осуществляется интерполяция, при которой кривая аппроксимируется полиномом третьей степени, который строится на основе четырёх последовательных точек данных.

— при выборе опции 5rd-order Lagrange (Лагранж 5-го порядка) осуществляется интерполяция, при которой кривая аппроксимируется полиномом пятого порядка, который строится на основе шести последовательных точек данных.



Text (Текст)

Эта опция предоставляет возможность изменить шрифт и размер текстовых элементов.

Сначала одним щелчком выделяют текст на графике. После этого на тексте появляются метки коррекции.

В группе Font (Шрифт) выбирают необходимый тип шрифта, а в группе Size (Размер) необходимый размер. Размер шрифта (кегель) выражается в точках.

После щелчка на кнопке Apply (Применить) выбранные свойства будут перенесены на выделенный объект. Эта кнопка становится активной только тогда, когда выделен текстовый объект.



3D-Rotation (3D-вращение)

Это один из двух методов, с использованием которых можно вращать 3D-диаграмму рассеяния. При помощи переключателей на левой стороне диалогового окна диаграмму можно вращать вперёд или назад относительно осей X, Y и Z.

Рисунки на переключателях указывают на ось и направление вращения. Вы можете вращать систему координат при помощи коротких щелчков на соответствующих переключателях или удерживая нажатой кнопку мыши. Вращение, задаваемое таким образом, отображается на упрощенной схеме, где изображены три оси; эта схема находится в центре диалогового окна.

Если активирована опция Show tripod (Показать треножник), то будет показан треножник, линии которого проходят через центр области построения диаграммы параллельно осям. Активирование треножника особенно рекомендуется тогда, когда необходимо проследить вращение осей при выключенном обрамлении трехмерного графика.

Вращение выделенной диаграммы происходит при помощи кнопки Apply (Применить).

График будет повернут только тогда, когда к нему будет применено заданное вращение. В течение операции вращения применение каких-либо других команд становится невозможным.

### Swap Axes (Смена осей)

При помощи этой опции в двумерном графике можно поменять местами вертикальную и горизонтальную оси.

### Explode Slice (Выдвинуть сегмент)

Чтобы выдвинуть сегмент круговой диаграммы, выделите его и нажмите эту кнопку.

### Break Lines at Missing (Разорвать линию в месте отсутствующего значения)

Разрыв линии на линейной диаграмме при наличии отсутствующего значения.

### Chart options (Параметры графика)

Здесь Вам предлагается выбор дополнительных параметров для столбчатых и линейчатых диаграмм, а также диаграмм с областями. В случае линейчатых диаграмм, Вы также можете разделить линии по категориям.

При активировании опции Change scale to 100 % (Перевести масштаб в проценты) точки данных столбчатых диаграмм и частотных диаграмм с областями переводятся в процентные показатели и отображаются как процентные доли. Если редактируемая диаграмма является столбчатой, то столбцы будут автоматически штабелированы. Если на редактируемой диаграмме столбец или область отображает только один ряд данных, то эта опция остаётся недостижимой. Эта опция также неприменима в случае, если диаграмма отображает функцию накопительной суммы.

В группе Line Options (Параметры линии) предлагаются ещё две возможности обработки линейных диаграмм.

- Опция Connect markers within categories (Соединить маркеры внутри категорий) соединяет маркеры, которые принадлежат к одним и тем же категориям, но лежат на разных кривых. Эта опция может применяться для диаграмм, на которых представлены как минимум две кривые. Она не влияет на текущий статус интерполяции или маркировки кривых.
- Опция Display projection (Показать проекцию) позволяет выделить некоторую проектируемую категорию. Категории, находящиеся справа от проектируемой категории отображаются иначе.

Если на диаграмме в виде столбцов представлены по меньшей мере два ряда данных, то при помощи группы Bar Type (Тип столбцов), её можно преобразовать в кластеризованную или состыкованную диаграмму. Если активирована опция Change scale to 100 % (Перевести масштаб в проценты), то группа Bar Type (Тип столбцов) становится недоступной.

### Set/exit spin mode (Включить/выключить режим вращения)

И эта кнопка делает возможным непосредственное вращение 3D-диаграммы рассеяния в окне редактора диаграмм; но здесь в процессе вращения диаграмма претерпевает некоторые упрощения.

Вращать диаграмму вперёд и назад относительно осей X, Y и Z можно при помощи кнопок с соответствующими символами в левой части диалогового окна.

Символы на кнопках вращения указывают на оси и направление вращения. Вы можете вращать область координат пошагово при помощи коротких щелчков или непрерывно, удерживая кнопку мыши нажатой. Производимое таким образом вращение, отображается при помощи системы трех осей в центре окна редактора диаграмм.

## 22.17. Примеры редактирования графиков

Некоторые примеры редактирования графиков уже приводились в главах 4, 6 и 11. В этой главе мы рассмотрим ещё три дополнительных примера.

### 22.17.1. Пример первый: изменение наименования осей

- Откройте в окне просмотра результатов файл `balken.spo`, в котором хранится график, изображённый на рис. 22.43.

Здесь необходимо изменить наименование вертикальной оси.

- Двойным щелчком перенесите график в редактор диаграмм и щелчком выделите наименование вертикальной оси.
- После этого выберите в меню Chart (Диаграмма) Axis... (Ось)

В появляющемся окне Вам предлагается множество разнообразных возможностей редактирования оси.

- Измените название оси на "Холестерин, исходный показатель" и покиньте диалоговое окно нажатием ОК.

В результате Вы увидите отредактированный график, который после закрытия редактора диаграмм будет отображён и в окне просмотра результатов.

### 22.17.2. Пример второй: редактирование круговой диаграммы

- Откройте файл `kreis.spo`, в котором хранится круговая диаграмма, представленная на рис. 22.30. Эта диаграмма пока ещё не показывает результаты голосования в процентах.
- Двойным щелчком перенесите график в редактор диаграмм.
- Щёлкните дважды на названии одной из представленных партий (к примеру, SPD или CDU).
- Откроется диалоговое окно Pie Options (Параметры круговой диаграммы) (см. рис. 22.69).
- Поставьте маркер в поле Percents (Проценты) и щёлкните на кнопке Format... (Формат).

Откроется диалоговое окно Pie Options: Label Format (Параметры круговой диаграммы: Формат метки), представленное на рисунке 22.70.

Здесь Вам предоставляется возможность указать место нахождения численного значения переменной.

- В группе Display Frame Around (Показать круговую рамку) активируйте опцию Outside labels (Метка снаружи).
- Подтвердите нажатием Continue (Далее) и затем на ОК. Вы получите диаграмму, изображённую на рисунке 22.71.

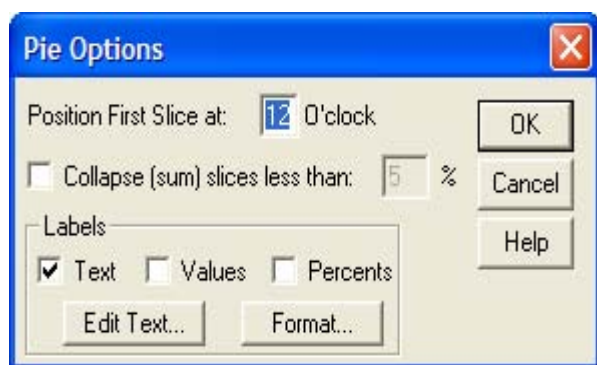


Рис. 22.69: Диалоговое окно Pie Options (Параметры круговой диаграммы)

### 22.17.3. Пример третий: нанесение регрессионных линий

- В окне просмотра результатов откройте файл `streumat.spo`, в котором находится матричная диаграмма рассеяния, изображённая на рис. 22.50, и двойным щелчком перенесите её в редактор диаграмм.
- В списке меню редактора диаграмм выберите Chart (Диаграмма) Options... (Параметры)

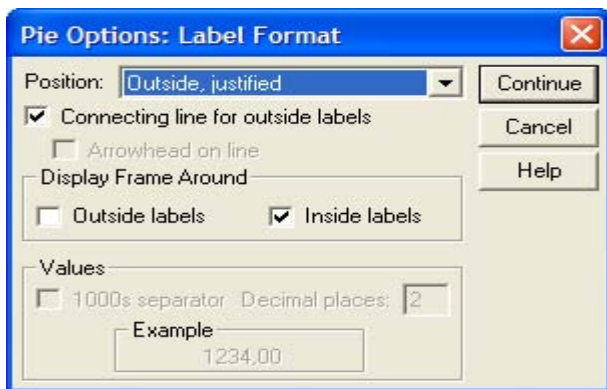


Рис. 22.70: Диалоговое окно Pie Options: Label Format (Параметры круговой диаграммы: Формат метки)

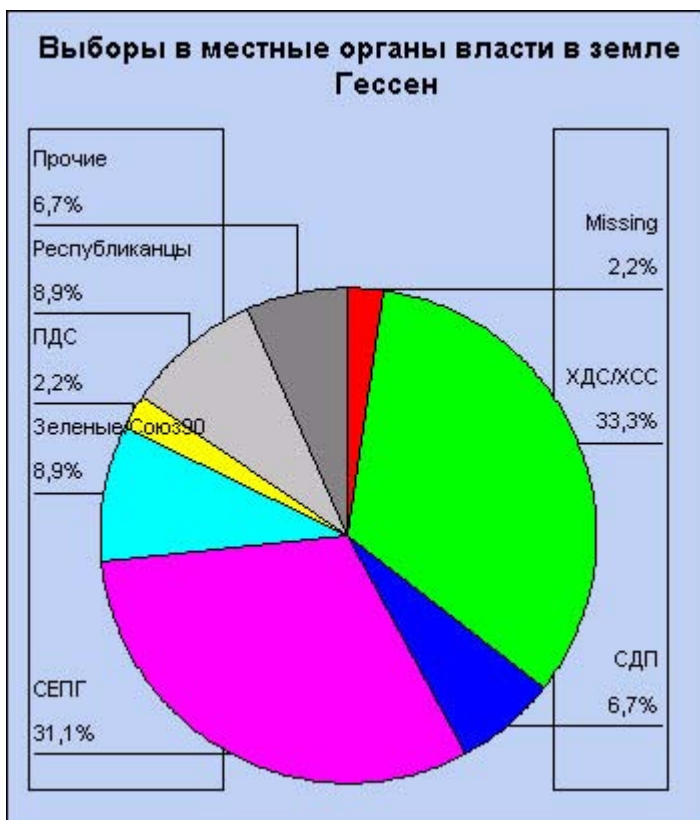


Рис. 22.71: Результаты голосования на местных выборах в земле Гессен 1993.

Откроется диалоговое окно Scatterplot Options (Параметры диаграммы рассеяния) (см. рис. 22.72).

- В группе Fit Line (Приближённая линия) активируйте опцию Total (Обобщённая).
- Щёлкните на выключателе Fit Options... (Параметры приближения). Откроется диалоговое окно Scatterplot Options: Fit Line (Параметры диаграммы рассеяния: Приближённая линия).
- Щёлкните на области Linear regression (Линейная регрессия) и в группе Regression Prediction Line(s) (Линия(и) для оценки качества регрессии) отметьте опцию Mean (Среднее значение); таким образом для регрессионной прямой Вы получите 95 % доверительный интервал.
- Покиньте диалоговое окно нажатием Continue (Далее) и затем ОК.

Теперь на рассматриваемой диаграмме рассеяния присутствуют регрессионные прямые и соответствующие им доверительные интервалы.

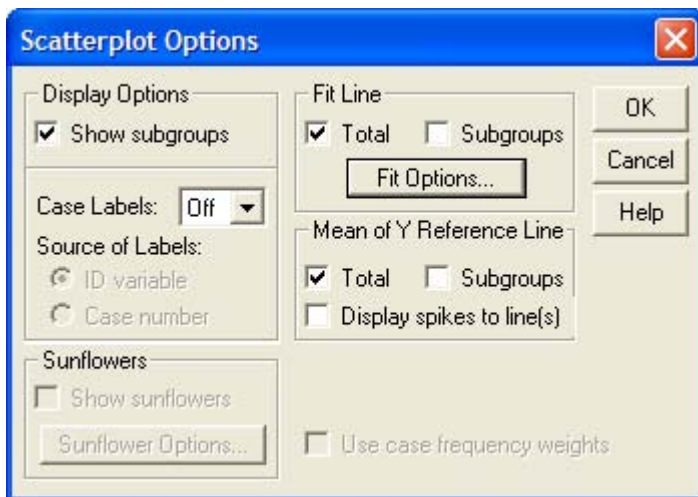


Рис. 22.72: Диалоговое окно Scatterplot Options (Параметры диаграммы рассеяния)

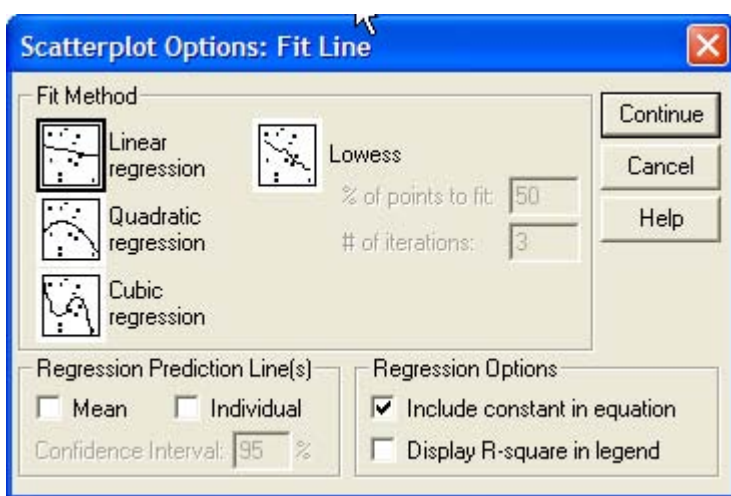


Рис. 22.73: Диалоговое окно Scatterplot Options: Fit Line (Параметры диаграммы рассеяния: Приближённая линия)

В корректировке нуждаются ещё названия переменных.

- Дважды щёлкните на тексте в левом верхнем диагональном элементе.
- В появившемся диалоговом окне Scatterplot Matrix Scale Axes (Оси матричной диаграммы рассеяния) в группе Individual Axes (Отдельные оси) отметьте редактируемый текст и щёлкните на выключателе Edit... (Правка).

Откроется диалоговое окно Scatterplot Matrix Scale Axes: Edit Selected Axis (Оси матричной диаграммы рассеяния: Редактирование выделенной оси).

- Наберите в диалоговом окне более короткий текст, к примеру, "ожидаемая продолжительность жизни" (Lebenserwartung) и подтвердите нажатием Continue (Далее).
- Поступите также и с двумя другими диагональными элементами матричной диаграммы рассеяния.
- Закончите редактирование графика нажатием ОК (см. рис. 22.74).

Всё многообразие возможностей для корректировки графиков, предлагаемых программой, при помощи нескольких приведенных примеров можно рассмотреть только в самых общих чертах. Эти примеры, по меньшей мере, должны были послужить Вам мотивацией для самостоятельного проведения дальнейших опытов, в ходе которых можно выяснить и другие возможности приведения графиков к более презентабельному виду.

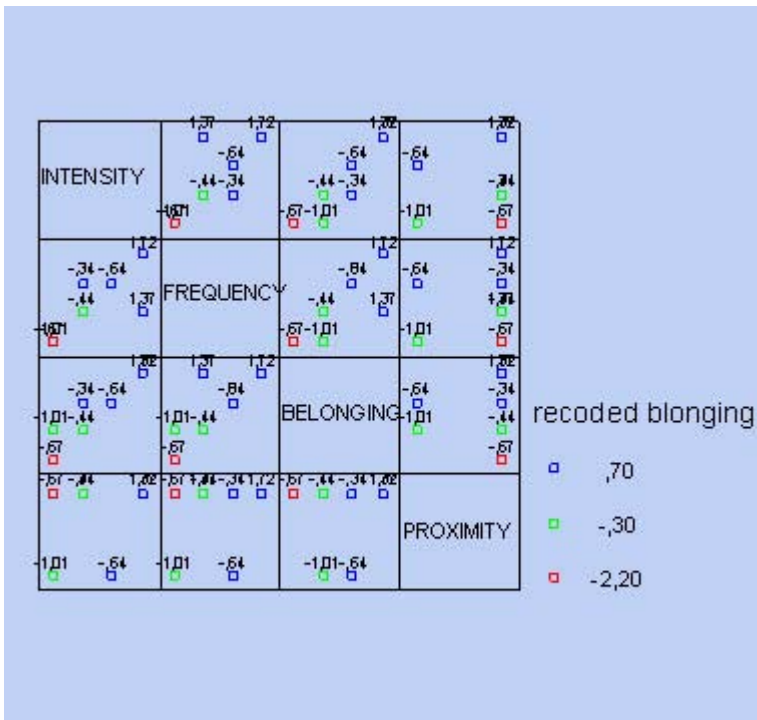


Рис. 22.74: Матричная диаграмма рассеяния с регрессионными прямыми и доверительными интервалами