

# РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ СИСТЕМЫ ОСА™

(Обработка Социологических Анкет)  
версия 3.01 для Windows 95/98



Киев  
Бюро статистического анализа 2004

## Содержание

Общие сведения .....	3
Основные понятия и количественные ограничения .....	4
Установка и запуск пакета .....	6
Файл   Открыть массив .....	8
Файл   Сохранить листинг .....	8
Файл   Список признаков .....	9
Файл   Отбор анкет (фильтр) .....	10
Файл   Взвешивание .....	12
Файл   Печать .....	13
Файл   Выход .....	14
Данные   Присоединить анкеты .....	15
Данные   Присоединить признаки .....	15
Данные   Импорт данных .....	16
Данные   Экспорт данных в SPSS .....	18
Данные   Экспорт данных в текст .....	20
Данные   Экспорт данных в *.txc .....	22
Данные   Сортировка анкет .....	22
Редактор   Удаление признаков .....	23
Редактор   Перемещение признаков .....	23
Редактор   Удаление анкет .....	24
Редактор   Ввод паспорта .....	24
Редактор   Редактирование паспорта .....	25
Редактор   Проверка паспорта .....	26
Редактор   Ввод массива .....	27
Редактор   Редактирование массива .....	29
Редактор   Работа с замечаниями   Упорядочить замечания .....	30
Редактор   Работа с замечаниями   Упорядочить замечания .....	32
Статистика   Одномерные распределения .....	32
Статистика   Двумерные распределения .....	35
Статистика   Статистические гипотезы   Знач. различий процентов .....	44
Статистика   Статистические гипотезы   Знач. различий коэф.корреляции .....	45
Статистика   Статистические гипотезы   Знач. различий средних .....	45
Статистика   Объем выборки .....	46
Вычисления   Условное вычисление .....	47
Вычисления   Перекодировка .....	48
Вычисления   Стандартизация .....	49
Примечания .....	50

## Общие сведения

Пакет **OCA** (аббревиатура от **Об**работка **Со**циологических **А**нкет), версия 3.01 для Windows 95/98, предназначен для организации ввода и статистического анализа результатов социологических опросов.

Пакет обеспечивает:

- ✓ **ввод** анкет (результата опроса) в файл на диск компьютера; возможность организации **параллельного ввода** на нескольких компьютерах; **контроль и управление** (с автоматическим заполнением) во время ввода; **логический контроль** введенной информации; ввод и последующее полуавтоматическое закрытие **открытых и полуоткрытых вопросов**; возможность организации **повторного ввода** с автоматическим сбором статистики об ошибках ввода
- ✓ работу с признаками, измеренными в **метрической, номинальной, порядковой** шкалах, а также признаков с **совместимыми альтернативами** (для вопросов, допускающих выбор респондентом нескольких вариантов ответа); автоматический **выбор показателей и методов анализа** в соответствии с заданными типами шкал признаков; автоматическую обработку **отсутствующих значений** для признаков в различных шкалах
- ✓ различные методы вычисления **дополнительных признаков**, автоматическое **документирование** всех вычислений
- ✓ построение **фильтров для отбора** анкет по условию и случайного отбора; различные методы **объединения** файлов данных; ремонт выборки путем **взвешивания**
- ✓ построение таблиц **одномерных распределений** (таблиц частот и процентов), вычисление **мер центральной тенденции и мер вариации** для признаков в различных шкалах
- ✓ построение **двумерных таблиц** частот и процентов, а также вычисление **показателей связи** признаков - коэффициента Хи-квадрат, коэффициентов корреляции Пирсона и Спирмена, коэффициентов Чупрова, Крамера, Гудмана и др.
- ✓ вывод построенных таблиц в формате **обычного текста** и формате языка гипертекстовой разметки документов **HTML**
- ✓ проверку **статистических гипотез** и построение **доверительных интервалов** для различных показателей
- ✓ построение уравнений **линейной регрессии**
- ✓ и многое другое ...

**Версия 3.01 для Windows 95/98 полностью совместима с версией OCA для DOS.**

- ✓ В версии для DOS допустимый диапазон для номинальных признаков от 1 до 254, в версии для Windows – от 0 до 254.
- ✓ В версии для DOS альтернативы номинальных признаков **обязательно** кодируются последовательными числами начиная с 1, в версии для Windows – **не обязательно** последовательными числами.
- ✓ Версия для Windows позволяет вводить комментарий в фильтр; версия для DOS не поддерживает комментарии в фильтре, попытка перестроить такой фильтр будет приводить к ошибке

## Основные понятия и количественные ограничения

При планировании опроса исследователь выделяет набор *признаков* (вопросов анкеты), в терминах которых он будет рассматривать изучаемую группу людей. Примерами признаков могут быть "Пол", "Возраст", "Образование". Для каждого признака можно определить набор возможных его значений. Так, признак "Пол" может принимать два значения: мужской и женский; признак "Возраст" измеряется в годах и может принимать в качестве значения любое число (даже не обязательно целое) в интервале от 15 до 100 (лет); для признака "Образование" в качестве допустимых могут быть выбраны следующие значения: начальное, среднее, неполное высшее и высшее. Ответ на вопрос анкеты (как правило, закодированный в виде числа) и выступает в качестве значения этого признака для данного респондента. Ответы всех респондентов в совокупности образуют данные для анализа.

Данные для анализа размещаются на диске в двух содержательно связанных между собой файлах, которые называются *массив анкет* *Массив анкет* и *паспорт массива анкет* *Паспорт массива анкет*. Во время работы оба эти файла обязательно должны находиться в одном каталоге (в одной папке) на диске компьютера.

**Массив анкет** – это двоичный (нетекстовый) файл, который содержит закодированные ответы респондентов. Этот файл всегда имеет расширение FRM. При указании имени массива в OCA используется имя двоичного файла, который содержит массив, но без расширения FRM. Массив анкет создается и редактируется только средствами пакета OCA. Массив представляет собой прямоугольную таблицу чисел. Одна строка таблицы называется *анкетой* **Анкета**. Один столбец таблицы называется *признаком* **Признак**. Все столбцы таблицы нумеруются последовательными числами, начиная с единицы – первый, второй, третий и так далее. Таким образом, каждый признак получает свой *номер* **Номер признака**. Во время работы на признаки ссылаются по номерам (например, "построить таблицу для 5-го признака", или "вычислить коэффициент корреляции между признаками с номерами 8 и 24").

В случае отсутствия в анкете ответа на определенный вопрос в качестве значения соответствующего признака для этой анкеты записывается специальный *код отсутствующего значения* *Отсутствующее значение*, который называют *НеОтвет* и обозначают символом "\$".

**Паспорт массива анкет** – это текстовый файл, который содержит описание структуры анкеты (количество признаков, названия признаков, тексты соответствующих признакам вопросов опросника, шкалы признаков, допустимые значения признаков и т.п.). Этот файл всегда имеет расширение DCL. При указании имени паспорта в OCA используется имя текстового файла, который содержит паспорт, но без расширения DCL. С целью обеспечения совместимости с версией OCA для DOS для записи текста паспорта использована кодировка ASCIIASCII и имя паспорта записывается по правилам записи имен файлов в DOS (не более 8 латинских букв, цифр и некоторых знаков). Массив анкет всегда содержит ссылку на паспорт, который его описывает. Массив анкет и описывающий его паспорт используются совместно и обязательно должны *соответствовать* друг другу. При любых изменениях в массиве (добавление признаков, удаление признаков, перестановка признаков и т.д.) пакет автоматически вносит соответствующие изменения в паспорт. Паспорт может быть подготовлен средствами любого текстового редактора, который работает в кодировке ASCIIASCII. После того, как массив анкет введен, редактировать паспорт и изменять имя паспорта можно только средствами пакета.

Пакет позволяет вести обработку признаков, заданных в *метрической, порядковой и номинальной шкалах*.

В случае *номинальной шкалы* числа (коды) используются для отображения атрибута, категории, или альтернативы ответа. Так, например, для признака «пол» можно использовать коды 1 («мужской») и 2 («женский»). Приписывание чисел в данном случае является произвольным. Вместо кодов 1 и 2 мы можем использовать коды 5 и 9, 25 и 99 и т.д. В связи с тем, что числа здесь не имеют количественной нагрузки, не имеет смысла применять к ним математические операции, такие как сложение, вычитание, умножение и деление. Единственная математическая операция, которая имеет смысл для номинальной шкалы, связана с определением частоты/процента (другими словами, количества/процента респондентов, отметивших данную альтернативу ответа) для каждого кода.

В пакете ОСА различаются *номинальные шкалы с несовместимыми альтернативами* (когда респондент может выбрать только одну из предложенных альтернатив; мы будем называть такие шкалы просто номинальными) и *совместимыми альтернативами* (когда респондент может выбрать от одной до всех из предложенных ему альтернатив). Номинальная шкала с совместимыми альтернативами имеет существенное ограничение в кодировке – альтернативы обязательно кодируются только последовательными целыми числами, начиная с единицы – 1,2,3, ...

В *порядковой шкале* числа отражают увеличение/уменьшение определенного качества/атрибута. Само название шкалы, «порядковая», предполагает, что респонденты (объекты или события) могут быть «упорядочены» по значениям признака, измеренного в этой шкале. Например, можно оценить 10 респондентов по степени их агрессивности, при чем значение 1 будет приписано наименее агрессивному, а 10 - наиболее агрессивному респонденту. Интервалы между значениями не равны. Различие между агрессивностью респондента, которому приписано значение 2, и агрессивностью респондента, которому приписано значение 4, не равно различию между агрессивностью респондента, которому приписано значение 5, и агрессивностью респондента, которому приписано значение 7. Если респонденту приписано значение 5, это не означает, что он в 5 раз агрессивней, чем тот, кому приписано значение 1.

*Метрическая шкала* предполагает равенство интервалов между числами (кодами). Примерами такой шкалы могут служить возраст, стаж работы, доход. Например, для признака возраст мы можем говорить о том, что респондент в возрасте 50 лет в два раза старше, чем респондент в возрасте 25 лет.

При использовании пакета ОСА задача исследователя состоит в правильном указании типа шкалы при создании паспорта массива (а также при построении новых признаков). На этапе обработки операции, преобразования и показатели, допустимые для определенного типа шкалы, определяются пакетом автоматически.

#### Диапазоны допустимых значений кодов для различных шкал

- ✓ *метрическая шкала* Шкала метрическая – положительные и отрицательные числа в интервале (приблизительно) от  $-3.4 \cdot 10^{38}$  до  $+3.4 \cdot 10^{38}$ , точность представления (приблизительно) 7-8 знаков
- ✓ *порядковая шкала* Шкала порядковая – целое число без знака в интервале от 0 до 254
- ✓ *номинальная шкала* Шкала номинальная – целое число без знака в интервале от 0 до 254
- ✓ *номинальная шкала с совместимыми альтернативами* Шкала номинальная с совместимыми альтернативами – последовательные целые числа без знака в интервале от 1 до 254

Очень часто необходимо вести обработку не всего массива, а лишь тех анкет, которые удовлетворяют некоторому условию, заданному на значениях признаков в анкете (например, только анкет, содержащих ответы женщин в возрасте до 35 лет, имеющих среднее специальное образование). Такие условия будем называть *фильтрами*. В пакете есть средства для создания таких фильтров. Фильтры после создания хранятся в файлах на диске и при необходимости могут быть подключены к массиву. Файлы фильтров имеют расширение FLT. Если к массиву анкет подключен фильтр, то при выполнении операций если условие фильтра выполняется, то анкета вовлекается в процесс вычислений. В противном случае соответствующая анкета в вычислениях "не участвует" (она как бы выбрасывается из массива на время выполнения этой операции).

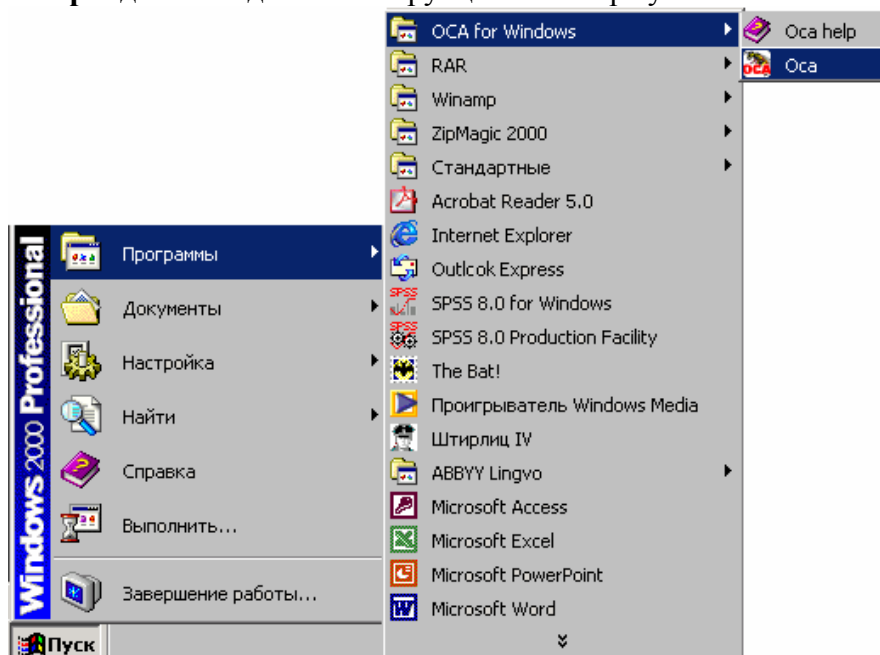
#### Некоторые количественные ограничения

- ✓ максимальное количество анкет в массиве – 100000
- ✓ максимальное количество признаков в анкете во время анализа – 10000
- ✓ максимальное количество признаков в анкете во время ввода – 1900
- ✓ длина любой строки в паспорте массива не превышает 255 знаков
- ✓ максимальная длина названия признака в паспорте массива – 254 знака (включая пробелы)
- ✓ максимальная длина текста вопроса в паспорте массива – 254 знака (включая пробелы)
- ✓ максимальное количество двумерных таблиц, которые могут быть построены во время одного выполнения операции построения двумерных таблиц – 500
- ✓ суммарная длина логического выражения и комментария в фильтре не должна превышать 254 знака
- ✓ количество признаков, которые экспортируются в формат текста Данные|Экспорт\_данных\_в\_текст, не должно превышать 5000

## Установка и запуск пакета

### Установка программы

Для установки программы необходимо запустить с дистрибутивной дискеты программу **Setup** и далее следовать инструкциям мастера установки.

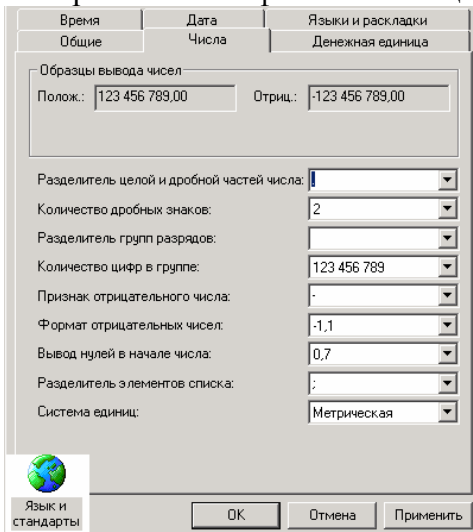


В процессе установки создается группа Oca for Windows. Запуск Oca for Windows осуществляется из меню программ

Start | Programs | Oca for Windows | Oca

Пуск | Программы | Oca for Windows | Oca

В процессе установки производится регистрация в Windows расширения FRM. Двойной щелчок мышью в Windows Explorer на имени файла с расширением FRM будет приводить к открытию этого файла с помощью OCA for Windows.



**Внимание.** Для нормальной работы программы необходимо, чтобы в Windows 95/98 в качестве разделителя между целой и дробной частью чисел была установлена десятичная точка "."

My Computer | Control Panel | Regional Settings | Number | Decimal symbol

Мой компьютер | Панель управления | Язык и стандарты | Числа | Разделитель целой и дробной частей числа

### Регистрация программы

Для нормальной работы программы необходимо выполнить регистрацию. Окно регистрации программы появляется автоматически при первом запуске после установки программы. Для регистрации необходимо заполнить все поля окна регистрации информацией, полученной от распространителя программы, и нажать кнопку [OK]. Регистрационная информация должна быть введена абсолютно точно (с учетом всех пробелов, больших и малых букв).

**Внимание.** В записи регистрационного номера начальный префикс OCA вводится обязательно латинскими буквами

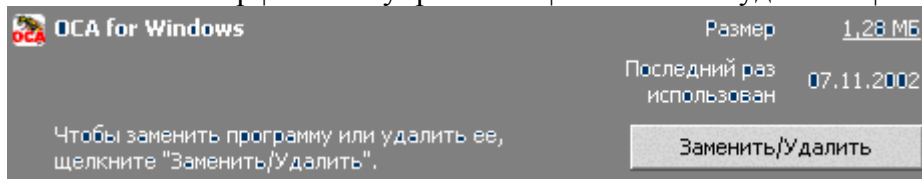
Данная версия программы не может быть установлена и зарегистрирована в сети - на каждой рабочей станции регистрация должна выполняться отдельно.

### Удаление программы

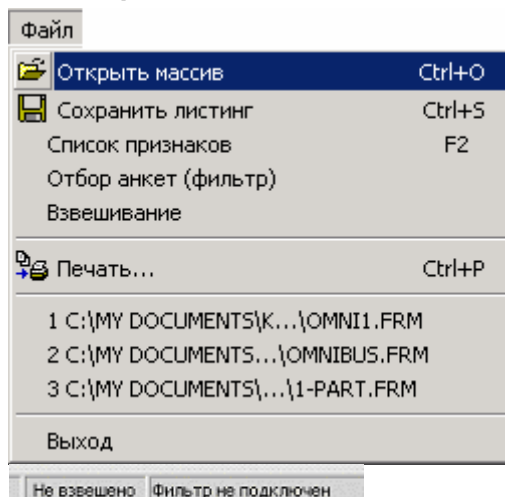
Для удаления программы необходимо удалить программу стандартным для Windows способом:

My Computer | Control Panel | Add/Remove Programs

Мой компьютер | Панель управления | Установка и удаление | Заменить / Удалить

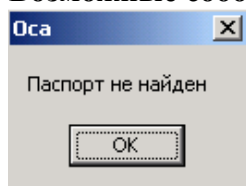


## Файл | Открыть массив

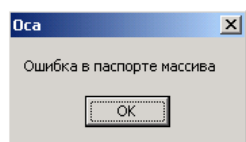


Массив открывается как **не взвешенный** и **без установленного фильтра**. На заголовке окна пакета выводится полное имя файла, содержащего открытый массив анкет, а также количество признаков и количество анкет в массиве. Кроме того, в протокол работы выводится количество признаков для каждого из типов шкал а также общая информация о массиве, которая извлекается из первого пункта паспорта.

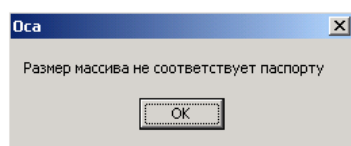
### Возможные сообщения об ошибках



"*Паспорт не найден*" – не найден файл с паспортом массива. Паспорт обязательно должен находиться в том же каталоге, что и массив анкет. Открыть для работы массив анкет без паспорта невозможно.



"*Ошибка в паспорте массива*": возможно, какой-либо из признаков составлен неправильно: обозначение типа шкалы не соответствует принятому в OCA, в конце паспорта нет знака / и т.д.



"*Размер массива не соответствует паспорту*": количество признаков в паспорте и массиве неидентичное.

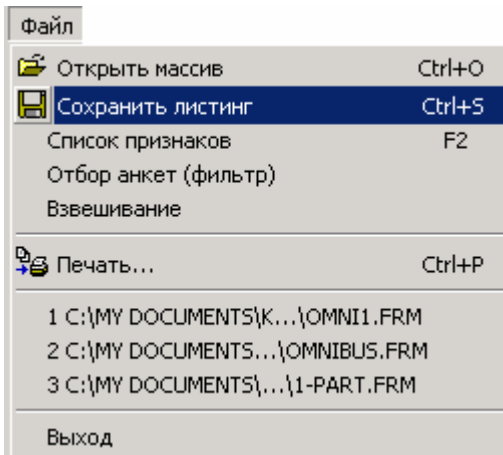
См. общие требования к паспорту в п. «Редактор | Ввод паспорта»

В обоих случаях, паспорт, скорее всего, редактировался не средствами OCA.

## Файл | Сохранить листинг

Листинг (протокол работы) дает результаты произведенных вычислений, название массива, а также информацию об изменениях, сделанных после запуска программы.





Для записи протокола работы на диск необходимо выполнить следующие действия:

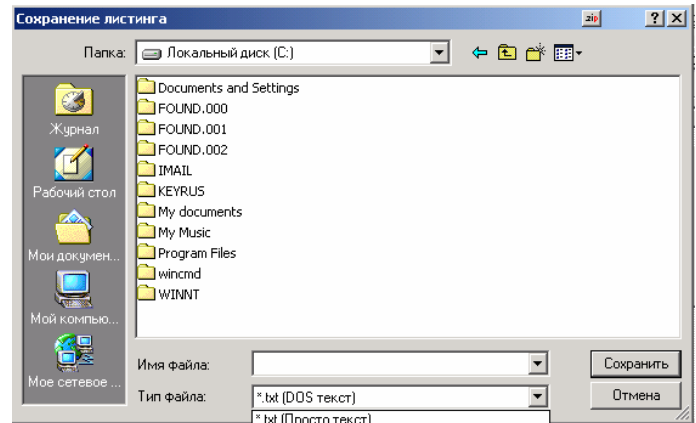
Выбрать **Файл | Сохранить листинг**.

Ввести имя файла (к имени файла будет автоматически дописано расширение txt) и выбрать необходимую кодировку (Просто текст или DOS текст) Листинг можно сохранить на диске как текст в кодировке "Просто текст" (ANSI) или в кодировке "DOS текст" (ASCII).

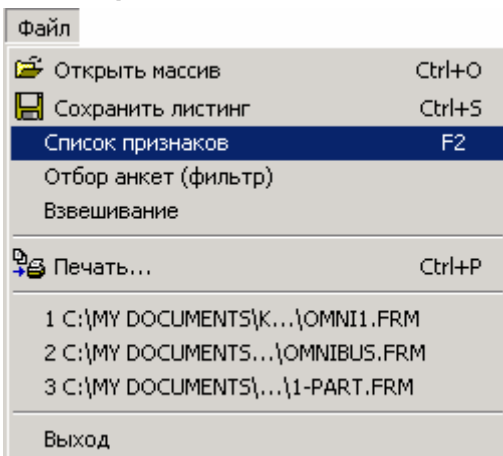
Нажать кнопку

Сохранить

**Появляется стандартное окно сохранения файла на диске:**



## Файл | Список признаков

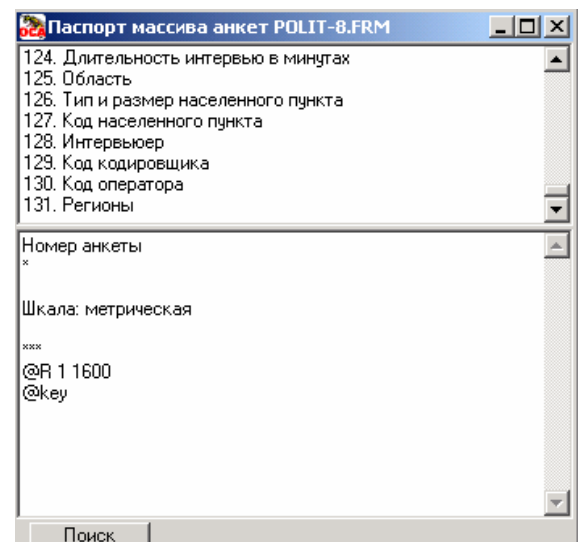



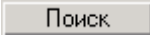
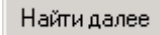
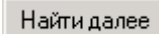
Выбор операции **Файл | Список признаков** или нажатие на клавишу **F2** приводит к вызову на экран окна, содержащего список признаков открытого для работы массива анкет.

Окно списка признаков делится на две части.

В *верхней части* для каждого признака выводится одна строка, содержащая номер признака и его название. Одна из строк в верхней части выделена другим цветом. Для выделения строки нужно установить на нее курсор мыши и щелкнуть левой кнопкой.

В *нижней части* выводится полная информация из паспорта о том признаке, который выделен в верхней части.



Кнопка  позволяет осуществлять поиск вхождения слова или фразы в тексте паспорта массива. После нажатия на кнопку  необходимо ввести текст, поиск которого будет осуществляться и нажать на кнопку  Поиск производится от описания выделенного в верхней части окна признака и до конца паспорта массива<sup>1</sup>. Если необходимо найти не первое вхождение указанного текста в паспорт массива, то для продолжения поиска необходимо воспользоваться той же кнопкой .

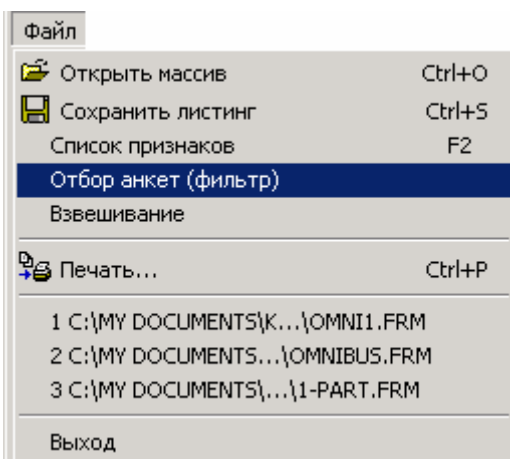
## Файл | Отбор анкет (фильтр)

Фильтры позволяют отбирать из массива для анализа анкеты в соответствии с некоторым условием или отбирать случайным образом необходимое количество анкет (для моделирования на массиве анкет случайной выборки).

При отборе по условию соответствующее логическое условие проверяется для каждой анкеты из массива. Фильтр выделяет из массива те анкеты, которые удовлетворяют указанному условию (т.е. те, для которых логическое условие является истинным). Таким образом, можно отобрать, например, анкеты мужчин в возрасте от 24 до 45 лет.

При случайном отборе нужно указать необходимое количество анкет или указать сколько процентов анкет из массива необходимо отобрать. При этом все анкеты из массива имеют одинаковые шансы попасть в подвыборку (с помощью датчика равномерно распределенных случайных чисел моделируется процесс простой одноступенчатой выборки из анкет массива).

Фильтр создается и хранится на диске в виде файла с указанным именем (это имя называется именем фильтра) и расширением FLT. Построенный фильтр можно использовать затем неоднократно (для этого нужно подключить существующий фильтр).



Операция **Файл | Отбор анкет (фильтр)** позволяет:

- ✓ построить новый фильтр
- ✓ подключить построенный ранее фильтр; при таком подключении истинность условия фильтра для каждой из анкет не проверяется
- ✓ отключить фильтр и вернуться к анализу всех анкет массива
- ✓ изменить (отредактировать, перестроить) построенный ранее фильтр

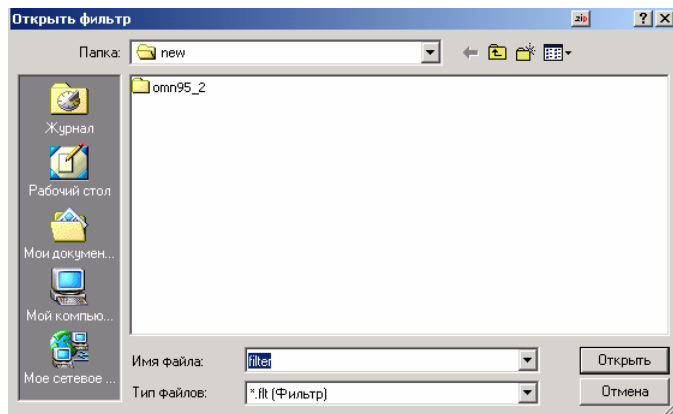
### Построение нового фильтра

Для построения нового фильтра необходимо выполнить следующие действия:

В главном меню выбрать **Файл | Отбор анкет (фильтр)**.



Выводится стандартное окно открытия файла:

<sup>1</sup> Поиск производится по всем полям описаний признаков, включая и поле комментария. Иногда в поле комментария (см. Ввод и редактирование паспорта) удобно указать краткое название признака (например, сокращение или аббревиатуру), которое не предназначено для вывода в таблицах, но является удобным для поиска.



В этом окне в качестве типа файла указана маска \*.flt (Фильтр)

В поле имени файла ввести имя нового файла (без расширения flt). Введенное имя и будет именем фильтра. Файл с таким именем (и с расширением flt), содержащий фильтр, будет создан в том каталоге, на который настроено окно открытия файла.

Нажать кнопку . Выводится окно параметров операции отбора анкет: В группе **Тип фильтра** в окне параметров выбрать способ отбора анкет в массиве (отбор по условию или случайный отбор) Если фильтр строится для отбора анкет по условию, в поле **Условие** ввести логическое условие отбора, в поле **Комментарий** ввести произвольный текст (вводить комментарий не обязательно) и нажать кнопку .

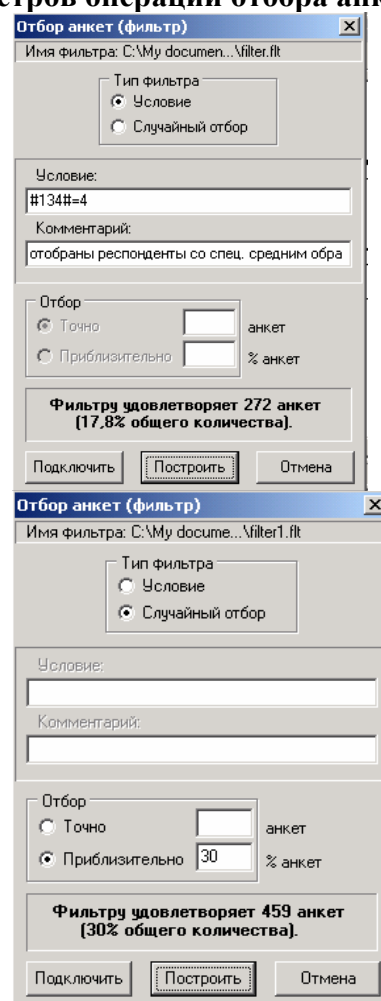
Суммарная длина условия и комментария не должна превышать 254 знака

Если фильтр строится для случайного отбора анкет, в группе **Отбор** выбрать **Точно** (и ввести точное количество анкет, которое необходимо отобразить) или

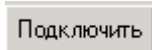
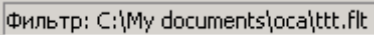
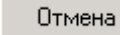

**Приблизительно** (и ввести сколько процентов анкет массива нужно отобразить; если, например, массив содержит 1530 анкет, то фильтр для отбора приблизительно 30% анкет отберет случайным образом 459 анкет).

Нажать кнопку .

Фильтр строится и записывается в файл на диске. Выводится количество и процент анкет массива, удовлетворяющих фильтру:

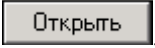
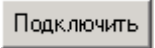
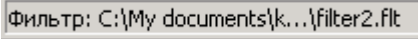


**Фильтру удовлетворяет 459 анкет (30% общего количества).**

Нажать на кнопку  и построенный фильтр будет подключен для дальнейшей работы (появляется сообщение с названием файла  на нижней информационной панели окна программы), или нажать на кнопку  и построенный фильтр не будет подключен (хотя он останется на диске и может быть впоследствии использован), или внести изменения в параметры фильтра и снова нажать кнопку  (при этом фильтр будет заново построен с новыми значениями параметров)


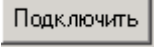
### **Подключение ранее построенного фильтра**

Для подключения построенного ранее фильтра необходимо выполнить следующие действия:

- 1) в главном меню выбрать **Файл | Отбор анкет (фильтр)**. Выводится стандартное окно открытия файла. В этом окне в качестве типа файла указана маска \*.flt (Фильтр)
- 2) выбрать файл, содержащий необходимый фильтр
- 3) нажать кнопку 
- 4) выводится окно с полной информацией о фильтре. Нажать кнопку . Появляется сообщение о подключенном фильтре на нижней информационной панели окна программы: 

### **Изменение ранее построенного фильтра**

Для изменения построенного ранее фильтра необходимо выполнить следующие действия:

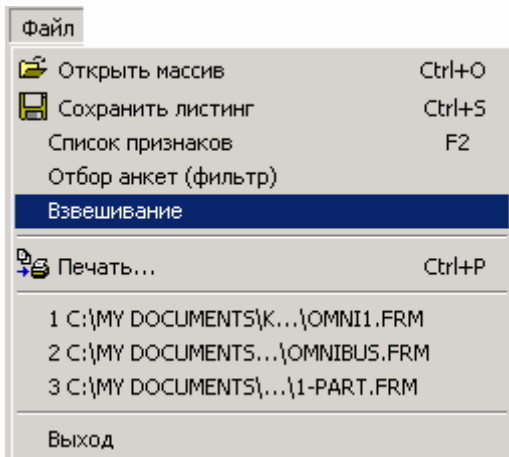
- 1) в главном меню выбрать **Файл | Отбор анкет (фильтр)**. Выводится стандартное окно открытия файла. В этом окне в качестве типа файла указана маска \*.flt (Фильтр)
- 2) выбрать файл, содержащий фильтр
- 3) нажать кнопку . Выводится окно, содержащее всю информацию о выбранном фильтре
- 4) внести изменение в значение параметров фильтра (так же, как и при построении нового фильтра) и нажать кнопку 

## **Файл | Взвешивание**

Взвешивание представляет собой один из подходов к ремонту выборки. Чтобы взвесить массив, необходимо просто указать номер того признака в массиве, который будет использоваться как вес. Как правило, это специально построенный для этой цели признак. В качестве веса может быть указан признак, измеренный в номинальной, порядковой или метрической шкале, но не в шкале с совместимыми альтернативами. Вес учитывается при построении таблиц и вычислении различных статистик. После выполнения взвешивания в нижней информационной панели окна программы появляется номер того признака, который используется как вес.


Значение веса должно быть положительным, но не обязательно целым числом. В случае использования нецелых весов, вследствие округлений частот, в одномерных и двумерных таблицах могут быть некоторые расхождения. Например, сумма всех частот в таблице может отличаться от общего количества анкет.

Если значение веса некоторой анкеты равно НеОтвету, нулю или отрицательному числу, то эта анкета исключается из анализа.

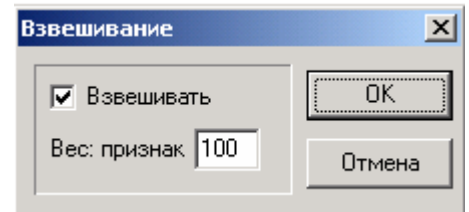


Чтобы взвесить массив необходимо выполнить следующие действия:


Выбрать операцию **Файл | Взвешивание**.

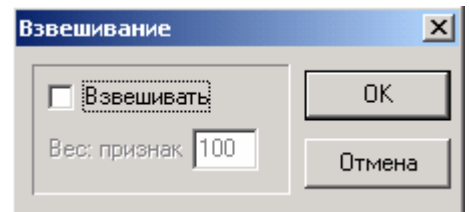
- ✓ Поставить отметку в поле Взвешивать.
- ✓ Указать в поле Вес: признак номер того признака массива, который будет использован в качестве веса
- ✓ Нажать кнопку .

Появляется окно операции:

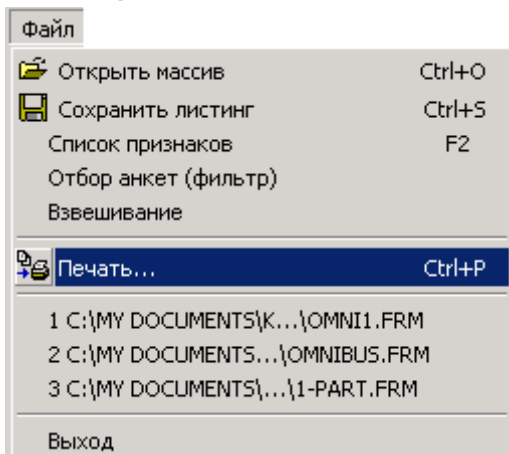


Чтобы *отменить взвешивание* и вернуться к анализу не взвешенного массива необходимо выполнить следующие действия

- ✓ Выбрать операцию **Файл | Взвешивание**.  
Появляется окно операции
- ✓ Снять отметку в поле Взвешивать
- ✓ Нажать кнопку .



## Файл | Печать



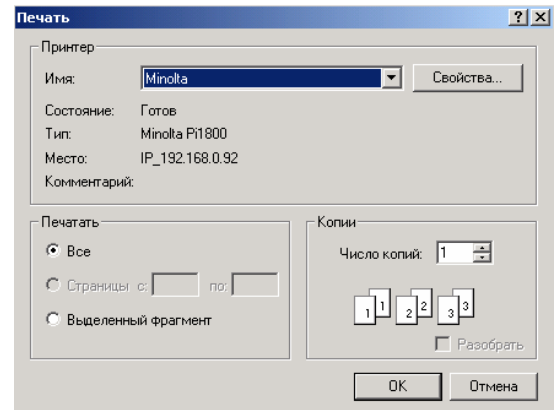
Эта операция позволяет распечатать протокол работы (листинг).

Для того, чтобы распечатать протокол работы, необходимо выполнить следующие действия:

- ✓ Выбрать **Файл | Печать**. На экран выводится стандартное для Windows окно печати:

- ✓ Выбрать принтер и при необходимости изменить его параметры, определить объем печати (весь протокол или только выделенную его часть) и указать количество копий.

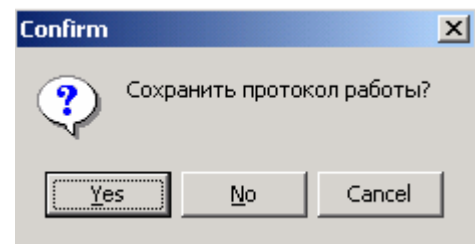
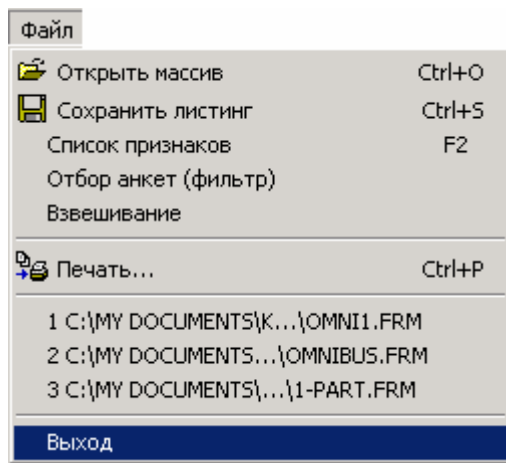
- ✓ Нажать кнопку 



## Файл | Выход

Эта операция завершает работу с пакетом.

Перед завершением работы на экран выводится стандартный диалог о сохранении протокола работы в файле на диске.



### Операции для работы с текстом в окне протокола работы

- ✓ **Текст | Вернуть** Отмена последней операции по редактированию текста протокола работы (листинга). Эту операцию можно выполнить также нажав комбинацию клавиш **Ctrl+Z**
- ✓ **Текст | Вырезать** Вырезать выделенный в тексте протокола работы (тексте листинга) фрагмент и поместить его в буфер обмена (clipboard). Эту операцию можно выполнить также нажав комбинацию клавиш **Ctrl+X**
- ✓ **Текст | Скопировать** Копировать выделенный в тексте протокола работы (тексте листинга) фрагмент в буфер обмена (clipboard). Эту операцию можно выполнить также нажав комбинацию клавиш **Ctrl+C**
- ✓ **Текст | Вставить** Вставить фрагмент текста из буфера обмена (clipboard) в протокол работы. Эту операцию можно выполнить также нажав комбинацию клавиш **Ctrl+V**
- ✓ **Текст | Выделить все** Выделить весь текст протокола работы (листинга). Эту операцию можно выполнить также нажав комбинацию клавиш **Ctrl+A**
- ✓ **Текст | Найти...** Запуск стандартного окна поиска текста в протоколе работы (листинге). Поиск производится от курсора и до конца текста протокола.

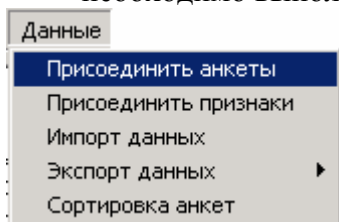
- ✓ **Текст | Шрифт** Запуск стандартного окна выбора шрифта вывода протокола результата. Для вывода используются только моноширинные шрифты (т.е. те, у которых одинаковая ширина всех знаков).

## Данные | Присоединить анкеты

Эта операция позволяет присоединить (дописать) к открытому для работы массиву анкеты из другого массива. Анкеты дописываются в конец (другими словами, после последней анкеты) открытого для работы массива. Оба массива (и открытый для работы массив и массив, из которого присоединяются анкеты) должны быть введены с абсолютно одинаковыми паспортами. Такая операция бывает необходима в тех случаях, когда часть анкет опроса вводится одним оператором, а часть анкет – другим оператором (параллельно, для сокращения времени ввода данных).

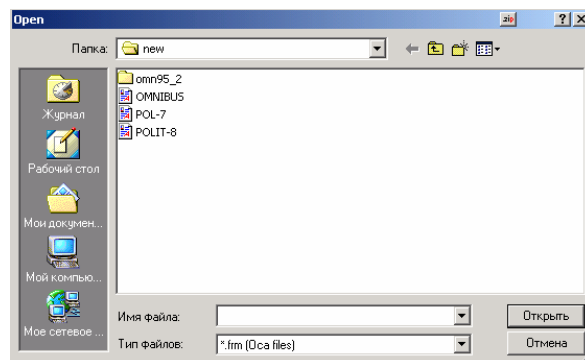
Если массивы имеют файлы замечаний, то автоматически эти файлы также автоматически объединяются. При этом файлы замечаний массивов должны иметь те же имена, что и массивы анкет, но расширения не *.fm* (стандартное расширение для массивов анкет) а *.txc*. Поиск файлов с замечаниями производится автоматически в тех же каталогах, где размещены и соответствующие массивы.

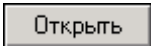
Чтобы присоединить к открытому для работы массиву анкеты из другого массива необходимо выполнить следующие действия:



Выбрать *Данные | Присоединить анкеты*.

- ✓ Выводится стандартное окно открытия файла. В этом окне в качестве типа файла указана маска *\*.fm* (OCA files):



- ✓ Указать файл, содержащий массив, из которого будут дописываться анкеты.
- ✓ Нажать кнопку .

## Данные | Присоединить признаки

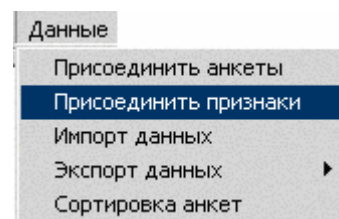
Эта операция позволяет присоединить к открытому для работы (назовем его первым) массиву признаки из другого (назовем его вторым) массива анкет. После присоединения структура первого массива (а значит и его паспорт) изменяется – появляются новые признаки. Второй массив остается без изменений.

Есть два способа присоединить признаки

- ✓ по порядку вхождения анкет в массивы
- ✓ по значениям ключевых признаков массивов

Если производится присоединение по порядку вхождения анкет, то в обоих массивах количество анкет должно быть одним и тем же. Операция выполняется таким образом:

- ✓ к первой анкете первого массива дописывается (как продолжение) первая анкета второго массива

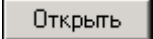


- ✓ ко второй анкете первого массива дописывается (как продолжение) вторая анкета второго массива
- ✓ и так далее ...

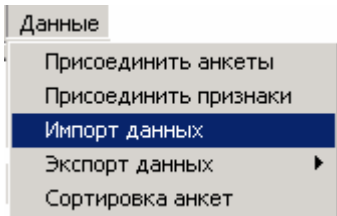
Если производится присоединение по значениям ключевых признаков, то для каждого из массивов должен быть указан номер ключевого признака, значения которого будет использовано в управлении процессом соединения массивов. Все анкеты второго массива обязательно должны иметь разные значения ключевого признака (другими словами, каждое значение ключевого признака первого массива должно определять не более одной анкеты во втором массиве). Для ключевого признака первого массива это требование не обязательно. Операция выполняется таким образом:

- ✓ берется значение ключевого признака первой анкеты первого массива; обозначим это значение k1
- ✓ во втором массиве ищется анкета, у которой значение ключевого признака равно k1, и эта анкета дописывается (как продолжение) к первой анкете первого массива; если такой анкеты во втором массиве нет, то в качестве продолжения дописывается необходимое (в соответствии с паспортом второго массива) количество НеОтветов
- ✓ последовательность шагов 1)-2) выполняется последовательно для всех анкет первого массива

Для того, чтобы к открытому для работы массиву анкет присоединить признаки из другого массива анкет необходимо:

- ✓ выбрать **Данные | Присоединить признаки**. Выводится стандартное окно открытия файла. В этом окне в качестве типа файла указана маска \*.ftm (OCA files)
- ✓ указать файл, содержащий массив, из которого будут присоединяться признаки.
- ✓ выбрать способ присоединения признаков (по порядку вхождения анкет в массивы или по значению ключевого признака)
- ✓ если присоединение будет выполняться по значению ключевого признака, то необходимо указать номер ключевого признака в первом и во втором массивах. Если в паспорте массива анкет есть признак, отмеченный директивой @key, то номер этого признака будет указан в качестве начального значения в соответствующем поле
- ✓ нажать кнопку 

## Данные | Импорт данных



Импорт данных (т.е. преобразование данных из одного формата в другой - в данном случае из текстового формата в формат файла массива анкет пакета OCA) производится из текстового файла в массив анкет в соответствии с заранее подготовленным паспортом. Результат импорта (т.е. построенный в результате импорта массив анкет) размещается в том же каталоге, где находится паспорт массива.

Данные в текстовом файле записываются в свободном формате:

- ✓ значения разделяются одним или несколькими пробелами
- ✓ допускается произвольное количество значений в строке
- ✓ допускается наличие пустых строк.

Длина строк в текстовом файле не должна превышать 254 символа.

Во время импорта данных значения, отличные от правильно записанных чисел, автоматически преобразуются в НеОтветы. Таким образом, для обозначения НеОтвета в



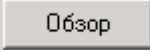
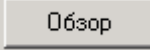
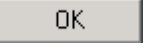
текстовом файле можно использовать любую последовательность знаков, которая не является правильно записанным числом (например, букву X).

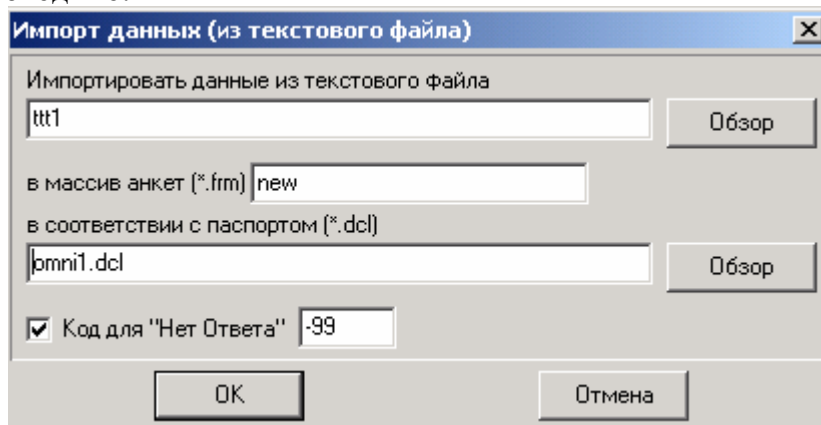
Во время импорта не производится проверка того, является ли каждое из значений в текстовом файле допустимым с точки зрения паспорта массива. Однако числа, которые выходят за пределы допустимых значений для шкалы соответствующего типа, автоматически преобразуются в НеОтветы. Так, например, если паспорт содержит номинальный признак "Пол респондента", закодированный числами 1 и 2, и в текстовом файле в качестве значения этого признака для некоторой анкеты указано число 3, то это число будет успешно "импортировано" и в дальнейшем (например, при построении таблиц) это значение 3 будет трактоваться как "Ошибка" (как значение, не соответствующее паспорту). В то же время, если в качестве значения рассматриваемого признака в текстовом файле будет указано число 325, то во время импорта это значение будет заменено на НеОтвет (поскольку значения номинальных признаков всегда должны находиться в интервале от 0 до 254).

Признак с совместимыми альтернативами представляется в текстовом файле как последовательность из 0 и 1. Количество 0 и 1 должно совпадать с количеством альтернатив признака, указанных в паспорте массива. Эти значения (то есть, 0 и 1) должны быть разделены (как и все другие значения) одним или несколькими пробелами. Отсутствующее значение (НеОтвет) для признака с совместимыми альтернативами представляется как соответствующее количество нулей. Если в последовательности чисел, представляющей значение признака с совместимыми альтернативами, есть неправильно записанное число или число, отличающееся от 0 и от 1, то в качестве значения признака устанавливается НеОтвет.

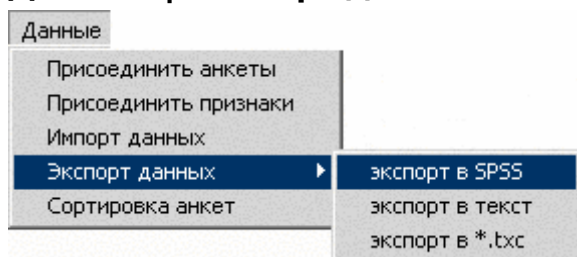
Если количество значений в текстовом файле не кратно количеству значений в одной анкете (количество значений в одной анкете определяется на основе анализа паспорта), то выдается сообщение и операция не выполняется.

Для выполнения операции необходимо:

- ✓ предварительно подготовить файл с паспортом будущего массива и разместить его на диске; для обеспечения совместимости с версией OCA для DOS файл паспорта должен иметь имя, записанное по правилам DOS (в частности, длина имени не должна превышать 8 символов и при записи имени нельзя использовать пробелы и буквы кириллицы)
- ✓ ввести имя текстового файла с данными; для этого можно воспользоваться кнопкой 
- ✓ указать имя массива, который должен быть создан в результате импорта текстового файла. Необходимо указать только имя файла массива анкет. Стандартное расширение \*.frm будет установлено для имени файла массива автоматически. Файл массива будет создан в том же каталоге, где находится паспорт массива.
- ✓ указать имя паспорта, в соответствии с которым будет производиться импорт; для этого можно воспользоваться кнопкой 
- ✓ если в текстовом файле для обозначения НеОтвета использован специальный код (допускается в качестве такого кода целое число, одно для всех признаков), то необходимо сделать отметку и указать этот код. Во время экспорта указанное значение всегда автоматически будет заменяться НеОтветом
- ✓ нажать кнопку . После успешного выполнения операции импорта вновь созданный массив будет автоматически открыт для работы.



## Данные | Экспорт данных в SPSS



Эта операция позволяет преобразовать массив анкет в формат, который легко читается пакетом SPSS for Windows, что позволяет воспользоваться для анализа данных методами многомерной статистики, реализованными в этом пакете.

Преобразовать можно все анкеты массива или только некоторые из них. Отбор анкет для экспорта осуществляется с помощью фильтра, установленного до выполнения преобразования. Преобразовать можно все признаки массива анкет или только часть из них.

Для каждого признака массива анкет в файле для SPSS создается числовая переменная. Для признака с совместимыми альтернативами создается группа дихотомических (то есть, принимающих два возможных значения - 0 или 1) переменных. Эта группа включает столько переменных, сколько альтернатив в соответствующем признаке.

Переменные в файле SPSS упорядочены по порядку возрастания номеров соответствующих признаков. Даже если номер признака указан в списке несколько раз, в файл для SPSS выводится только одна соответствующая ему переменная.

Первая буква имен переменных



Имена переменных (variable names) в файле SPSS строятся автоматически на основе номеров признаков.

Необходимо только задать первую букву имени (обязательно латинскую). Например, если задана первая буква "V", то первый признак будет преобразован в переменную V1, 10-й признак - в переменную V10, а 325й признак с совместимыми альтернативами (4 альтернативы) в четыре переменных - V325\_1, V325\_2, V325\_3 и V325\_4.

Метки переменных (variable labels) и метки значений (value labels) строятся автоматически на основе паспорта массива - соответственно из названий признаков и названий альтернатив. Необходимо указать максимально допустимую длину меток (не более 255 знаков). Различные версии SPSS позволяют использовать метки различной длины. Так, например, в версии SPSS for Windows 5.0 длина метки переменной не должна превышать 120 знаков, а в версии 10.5 – не должна превышать 256 знаков. Необходимо также помнить, что различные версии SPSS for Windows имеют различные ограничения на длину обрабатываемых строк текста. В качестве меток переменных и меток значений используется указанное количество первых символов названий признаков или названий альтернативы.

Транслитерация

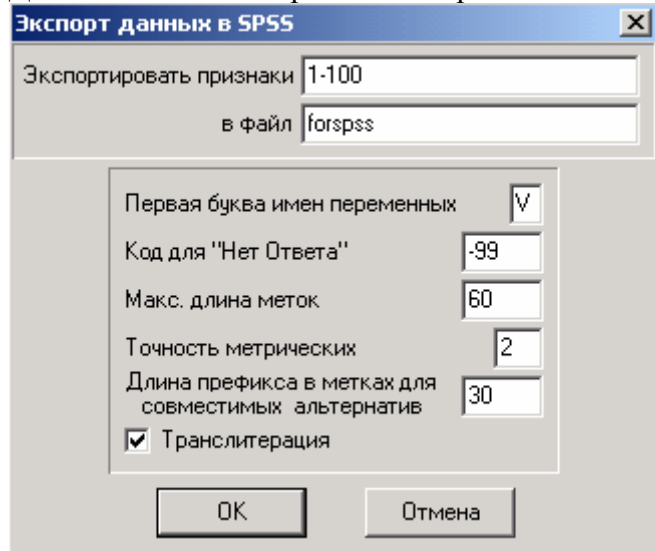
Если паспорт составлен с использованием букв кириллицы (например, на украинском или русском языке), то, при необходимости, можно выполнить транслитерацию меток (автоматическую замену кириллических букв на похожие по звучанию латинские буквы или их сочетания). Необходимо также учитывать, что паспорт записывается в кодировке ASCII, а метки переменных и метки значений выводятся в кодировке ANSI.

Для дихотомических переменных, образованных из признака с совместимыми альтернативами, метки переменных строятся из комбинации (сцепления) названия признака и названия соответствующей альтернативы. Из названия признака берется указанное количество символов (если указано 0 символов, то в качестве метки берется только название соответствующей альтернативы). В качестве меток значений используются слова "нет" (для значения 0) и "да" (для значения 1). Если выполняется транслитерация, то используются слова "no" и "yes".

Для преобразования массива анкет OCA в формат пакета SPSS необходимо:

- ✓ если преобразовываться будет не весь массив а его часть, построить и установить до начала выполнения операции соответствующий фильтр
- ✓ выполнить операцию экспорта в формат пакета SPSS; в результате будет создан текстовый файл, содержащий последовательность команд SPSS и сами данные
- ✓ в среде пакета SPSS открыть окно для ввода программ (syntax), ввести и выполнить команду INCLUDE 'файл'. , где файл - имя построенного в результате экспорта файла.

Для выполнения операции экспорта необходимо:

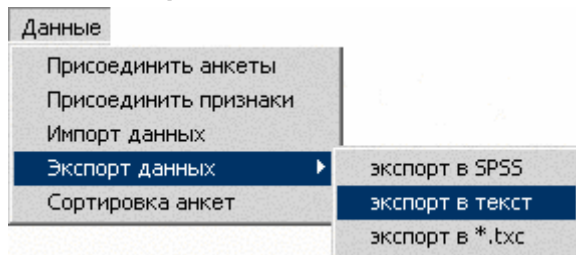


- ✓ задать список номеров признаков, которые необходимо экспортировать
- ✓ задать полное имя файла (нового, еще не существующего файла), в который будет производиться экспорт
- ✓ указать первую букву для имен переменных (обязательно латинскую), которая будет использоваться при автоматическом формировании имен переменных
- ✓ задать код для отсутствующего значения (HeОтвета); при экспорте используется один код для всех

переменных; это значение указывается в команде MISSING VALUE языка пакета SPSS; для этого значения автоматически вводится метка "Нет Ответа" (при выполнении транслитерации вводится метка "No Answer"); отсутствие значения для признака с совместимыми альтернативами выводится как значения 0 у всех соответствующих (т.е. образованных из этого признака) дихотомических переменных

- ✓ указать максимальную длину меток (меток переменных и меток значений); длина не должна превышать 255 знаков
- ✓ ввести точность (количество знаков после десятичной точки), с которой необходимо выводить значения метрических признаков; точность указывается как целое число в интервале от 0 до 6; признаки номинальные и порядковые выводятся как целые числа
- ✓ указать сколько знаков из названия признака необходимо взять при построении меток дихотомических переменных, образованных из признака с совместимыми альтернативами; если указать 0 (ноль), то в качестве метки берется только название альтернативы
- ✓ указать проводить или нет транслитерацию меток в формат ANSI

## Данные | Экспорт данных в текст



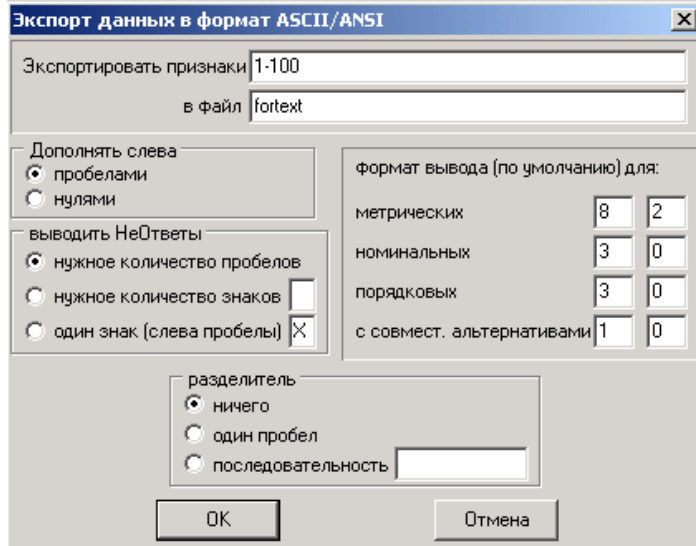
Эта операция преобразует данные из массива анкет в текст в соответствии с определенным форматом и выводит эти данные в указанный файл. Вывод в текстовый файл производится в фиксированном формате (для каждого признака определено количество позиций в строке вывода). Можно преобразовать не все

анкеты массива, а только ту их часть, которая удовлетворяет определенному фильтру. Экспортируются признаки из указанного списка. Порядок перечисления номеров в списке определяет порядок вывода. Каждая анкета выводится как одна или как несколько строк текста. При необходимости, между значениями могут вставляться определенные разделители (пробелы, запятая и т.п.). После успешного завершения экспорта в окно результатов выводится протокол, описывающий структуру экспортированных данных. В протоколе для каждого признака указывается сколько позиций для него выделено и номера этих позиций в строке вывода.

Признак с совместимыми альтернативами выводится как последовательность, состоящая из 0 и 1. Для каждой альтернативы выводится либо 1, если респондент указал в своем ответе эту альтернативу, либо 0, если не указал.

Для выполнения операции экспорта данных в текстовый файл необходимо:

- ✓ если требуется экспорт не всего массива, а лишь определенной его части, до начала выполнения операции экспорта построить и установить фильтр, который отбирает необходимые анкеты
- ✓ выбрать **Данные | Экспорт данных в текст**



- ✓ задать список номеров признаков, которые будут экспортироваться (поле Экспортировать признаки). Общее количество признаков в списке не должно превышать 5000. Признаки выводятся в соответствии с указанной в списке последовательностью. Один и тот же номер признака может быть указан несколько раз и при этом признак выводится столько раз, сколько раз указан его номер. Символ "/" указывает конец строки, что позволяет вывести одну анкету в последовательность

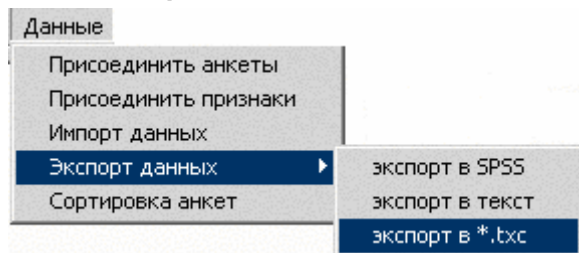
подряд идущих строк. Например, задание 1-20 / 1 31-40 42 / 1 41 20-25 приведет к тому, что для каждой анкеты массива при экспорте будет сформировано 3 строки. Первая строка будет содержать признаки с первого по 20-й; вторая строка - первый, с 31 по 40 и 42-й; третья строка - первый, 41-й и с 20-го по 25-й

- ✓ указать полное имя файла, в который будет производиться экспорт (поле в файл)
- ✓ задать формат вывода (по умолчанию) для значений каждого из четырех типов шкал (группа Формат вывода (по умолчанию) для:). Формат вывода задается двумя числами: первое число задает общее количество знаков, выделенных для вывода признака, а второе – количество знаков после десятичной точки (общее количество включает и десятичную точку, если она выводится, и знак числа, если число отрицательное). Каждый признак массива анкет может иметь свой собственный формат вывода, который задается директивой @F в соответствующем пункте паспорта массива. Если для признака в паспорте массива не указан формат вывода, то при его экспорте используется формат вывода по умолчанию. Возможна ситуация, когда некоторые значения не могут быть выведены в соответствии с указанным форматом. Например, значение 234 не может быть выведено с форматом 3 1 (т.е. не может быть выведено в три позиции так, чтобы была выведена также десятичная точка и один знак после десятичной точки). Для вывода значения 234 можно указать формат 3 0 или 5 1. Если значение не может быть экспортировано в соответствии с указанным форматом, то на экран выводится сообщение об ошибке и операция экспорта не производится
- ✓ определить чем необходимо дополнять слева значения при выводе (группа выбора Дополнять слева). Если формат вывода выделяет больше знаков, чем необходимо для вывода значения, то значение "прижимается" вправо и дополняется слева либо нулями либо пробелами
- ✓ определить как выводить отсутствующие значения (группа выбора выводить НеОтветы). Отсутствующие значения могут выводиться либо как последовательность пробелов необходимой (в соответствии с форматом) длины, либо как последовательность определенных символов необходимой длины (например, последовательность знаков вопроса "?" или последовательность точек ".")

либо как один знак, "прижатый" в поле вывода вправо и дополненный слева пробелами (например, знак "\$"). Необходимо учитывать, что для признаков с совместимыми альтернативами отсутствующие значения всегда выводятся как последовательность нулей (то есть как отсутствие выбора для каждой из альтернатив признака)

- ✓ определить разделитель между значениями при выводе (группа выбора разделитель). Значения могут выводиться без разделителя, а могут быть отделены одним пробелом или определенной последовательностью знаков (в качестве такой последовательности знаков может быть, например, запятая, либо пять пробелов, либо точка с запятой)

## Данные | Экспорт данных в \*.txc



Эта операция выводит значения одного признака массива в файл на диск в формате текстовых замечаний ТХС. Если установлен фильтр, то выводятся значения указанного признака только для тех анкет, которые удовлетворяют фильтру. В паспорте массива директивой @key обязательно должен быть

отмечен ключевой признак, который принимает различные значения для всех анкет массива.

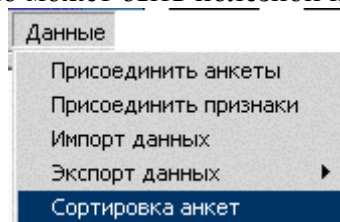
Для признаков с совместимыми альтернативами значения выводятся как последовательность номеров отмеченных респондентом альтернатив.

Отсутствующее значение (HeОтвет) для всех шкал выводится как символ "\$".

Такой экспорт может быть полезен если, например, необходимо произвести коррекцию значений определенного признака у нескольких анкет. Фильтром отбирают необходимые анкеты, признак выводят в текстовый файл в формате ТХС, производят изменения значений средствами любого текстового редактора и затем вносят необходимые изменения, используя этот файл с новыми значениями в формате ТХС.

## Данные | Сортировка анкет

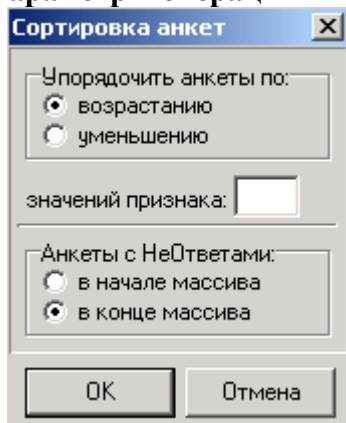
Упорядочивание (сортировка) анкет не является необходимой для статистического анализа, но может быть полезной при поиске ошибочных анкет в массиве.



Эта операция позволяет переупорядочить анкеты в массиве по возрастанию или по убыванию кодов значений одного признака. Признак сортировки не должен быть признаком с совместимыми альтернативами. Анкеты с одинаковым значением признака сортировки между собой никак не упорядочены. Анкеты, у которых значение признака

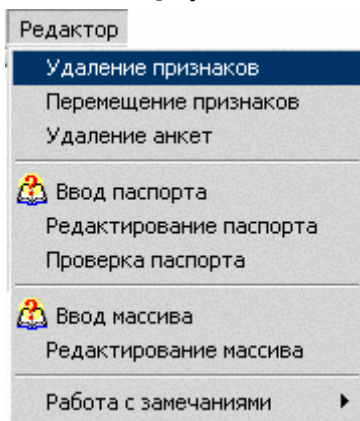
сортировки равно HeОтвету, размещаются в начале или в конце (отсортированного массива анкет).

## Параметры операции



- ✓ в поле выбора **Упорядочить анкеты по:** – выбор способа сортировки (по возрастанию или по уменьшению значений признака).
- ✓ поле **значений признака:** - задание номера признака, по значениям которого будет производиться упорядочение; этот признак не должен быть признаком с совместимыми альтернативами.
- ✓ поле выбора **Анкеты с НеОтветами:** определяет способ обработки анкет, у которых указанное значение признака имеет значение НеОтвет; такие анкеты после сортировки могут быть размещены либо в начале массива либо в конце массива.

## Редактор | Удаление признаков

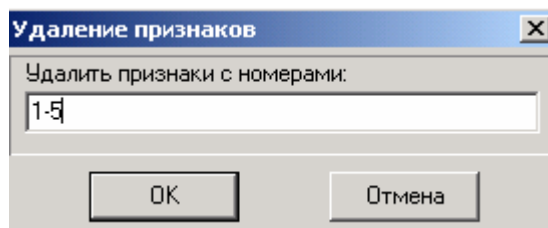


Эта операция позволяет удалить один или несколько признаков. Признак удаляется из массива и из паспорта.

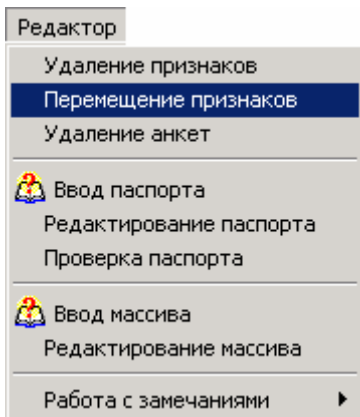
- ✓ Удаленный признак восстановить нельзя.
- ✓ Не рекомендуем удалять первичные (не построенные с помощью пакета) признаки.

Параметры операции:

В поле **Удалить признаки с номерами** вводится список номеров признаков, которые необходимо удалить.



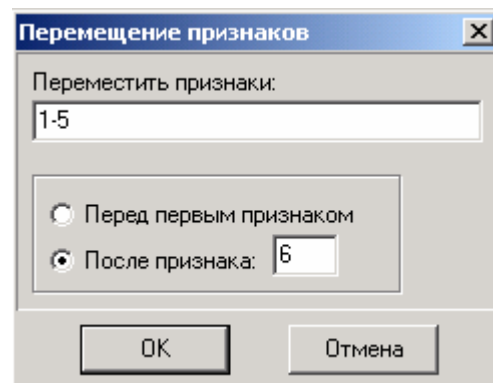
## Редактор | Перемещение признаков



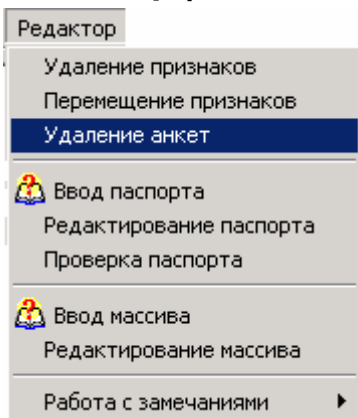
Иногда необходимо изменить порядок следования признаков в массиве (при этом, конечно же, изменяются и номера признаков). Эта операция позволяет переместить один или несколько признаков таким образом, чтобы они размещались либо перед первым признаком массива либо после указанного признака массива.

Параметры операции:

- ✓ поле Переместить признаки – список номеров признаков, которые необходимо переместить; при перемещении порядок, в котором указаны признаки, сохраняется
- ✓ группа выбора – позволяет указать куда необходимо переместить признаки – Перед первым признаком массива или После признака с указанным номером



## Редактор | Удаление анкет



Чтобы удалить из массива часть анкет, необходимо построить и установить фильтр, который отбирает анкеты, подлежащие удалению.

Внимание. Удаленные из массива анкеты восстановить нельзя.

После выполнения удаления фильтр, который использовался, теряет смысл. Более того, все другие фильтры (если они есть) также обязательно необходимо перестроить (поскольку в массиве изменяется количество анкет).

## Редактор | Ввод паспорта

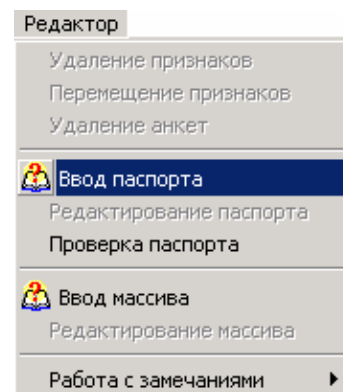
*Паспорт массива* - это специальным образом отформатированный текстовый файл, подготовленный и записанный на диск в формате ASCII. Паспорт можно подготовить средствами любого текстового редактора, который позволяет работать в формате ASCII, в том числе и средствами редактора MS Word.

### Структура паспорта

Паспорт представляет собой последовательность пунктов.

Каждый пункт описывает один признак.

- ✓ Название признака – одна (обязательно одна !) строка текста длиной до 254 знаков.
- ✓ Текст вопроса – одна (обязательно одна !) строка текста длиной до 254 знаков
- ✓ Тип шкалы обозначается одной из следующих букв латинского алфавита –m, o, n или j. Можно использовать как малые, так и большие буквы
- ✓ Код –целое положительное число; для кодирования номинальных признаков с совместимыми альтернативами можно использовать только последовательные числа начиная с 1. Коды значений для номинальных признаков с совместимыми альтернативами должны находиться в диапазоне от 1 до 254. Коды значений для номинальных и порядковых признаков должны находиться в диапазоне от 0 до 254.
- ✓ Название значения – текст; между кодом и названием значения должен быть, по крайней мере, один пробел
- ✓ Комментарий – последовательность строк текста; комментарий (включая и строку перед полем комментария - "\*\*\*") может в пункте паспорта отсутствовать. В поле комментария любого пункта можно также указывать директивы, которые в





основном оказывают влияние на ввод данных (дополнительный контроль вводимой информации, управление вводом и т.п.).

В зависимости от типа шкалы признака соответствующий пункт паспорта имеет следующую структуру:

Метрическая	Порядковая	Номинальная	Номинальная с совместимыми альтернативами
Название признака	Название признака	Название признака	Название признака
Текст вопроса	Текст вопроса	Текст вопроса	Текст вопроса
m	o	n	j
***	код название значения	код название значения	код название значения
комментарий	код название значения	код название значения	код название значения
/	...	...	...
	***	***	***
	комментарий	комментарий	комментарий
	/	/	/

Как правило, паспорт готовят на основе текста опросника. После завершения подготовки паспорта его необходимо проверить на корректность. После того, как начался ввод данных, редактирование паспорта с помощью обычного текстового редактора может привести к тому, что паспорт перестанет соответствовать уже введенной части массива. Поэтому в таком случае необходимо пользоваться операцией редактирования паспорта.

**Внимание.** В паспорте массива не должно быть строк, длина которых превышает 254 знака.

При подготовке паспорта средствами MS Word необходимо учитывать следующее:

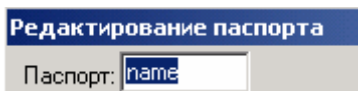
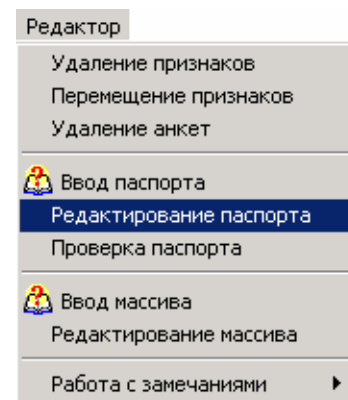
- ✓ результат работы нужно записывать на диск с помощью операции Save As... (Сохранить как...) в формате MS-DOS Text
- ✓ каждый абзац текста в MS Word (ввод абзаца завершается нажатием на клавишу Enter) будет в паспорте рассматриваться как одна строка (даже если абзац состоит из нескольких строк)

## Редактор | Редактирование паспорта

Выбрать операцию *Редактор | Редактирование паспорта*.

Появляется окно этой операции

Эта операция позволяет изменить имя паспорта массива или внести изменения в пункт паспорта признака – отредактировать название признака и текст вопроса опросника, отредактировать возможные значения признака и их коды, отредактировать информацию в поле комментария признака, изменить шкалу признака (есть только возможность менять номинальную шкалу на порядковую и наоборот - порядковую шкалу на номинальную).

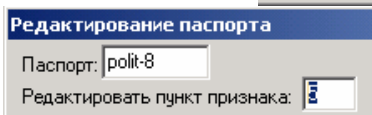


### Изменение имени паспорта массива

В поле *Паспорт:* отредактировать имя паспорта

Нажать кнопку

Записать



### Редактирование пункта паспорта

В поле *Редактировать пункт признака:* указать необходимый номер признака

Нажать кнопку

Редактировать

**Выводится информация о признаке с указанным номером из паспорта массива:**

Если есть необходимость, отредактировать в поле **Название** название признака, отредактировать в поле **Вопрос** текст вопроса и внести изменения в поле **Комментарии**.

Если есть необходимость, внести изменения в поле **Варианты ответа (код название)**: При этом необходимо учитывать следующие правила:

- ✓ информация об одном варианте ответа записывается в одной строке.

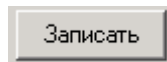
Записывается сначала код (число) а затем название (произвольный текст).

Название от кода отделяется одним или несколькими пробелами

- ✓ номинальные и порядковые признаки кодируются целыми числами без знака из интервала от 0 до 254
- ✓ номинальные признаки с совместимыми альтернативами кодируются последовательными целыми числами без знака начиная с 1

Если есть необходимость, изменить шкалу признака. Можно поменять номинальную шкалу на порядковую или порядковую на номинальную. Другие изменения не допускаются.

Нажать кнопку

**Редактор | Проверка паспорта**

Эта операция позволяет изменить имя паспорта массива или внести изменения в пункт паспорта признака – отредактировать название признака и текст вопроса опросника, отредактировать возможные значения признака и их коды, отредактировать информацию в поле комментария признака, изменить шкалу признака (есть только возможность менять номинальную шкалу на порядковую и наоборот - порядковую шкалу на номинальную).

**Изменение имени паспорта массива**

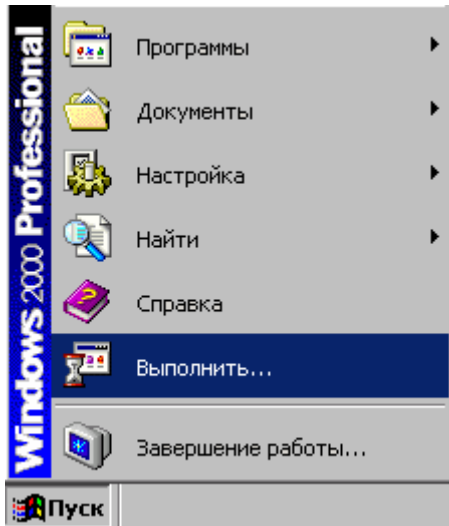
- ✓ Для изменения имени паспорта массива необходимо выполнить следующие действия
- ✓ выбрать операцию Редактор | Редактирование паспорта. Появляется окно этой операции

- ✓ в поле Паспорт: отредактировать имя паспорта

- ✓ нажать кнопку

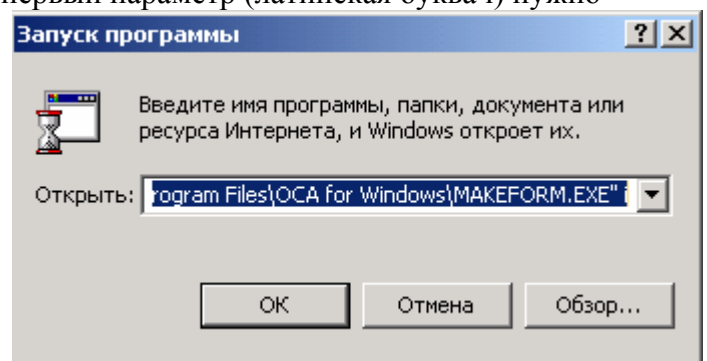


## Редактор | Ввод массива



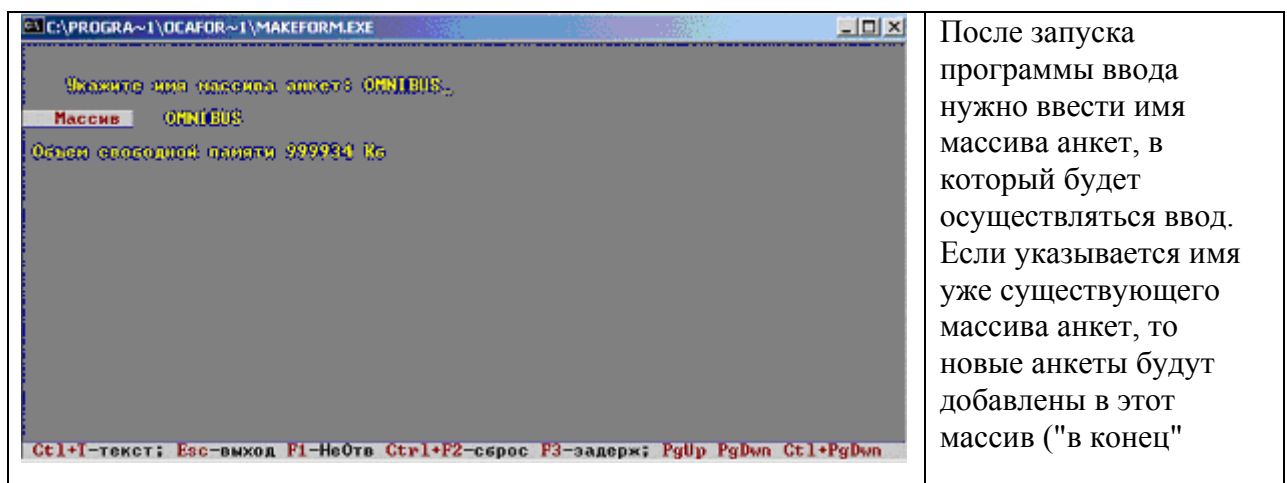
Ввод анкет массива осуществляется программой MakeForm.exe, работающей под управлением MS DOS. Это позволяет организовать ввод массива параллельно на нескольких дешевых (даже таких, которые не могут работать с MS Windows 95/98) машинах. Введенные на разных компьютерах части затем можно объединить для анализа. Все части должны быть введены с одним и тем же паспортом. Файл MakeForm.exe, паспорт массива и сам массив нужно разместить в одном каталоге на диске. Запуск программы ввода осуществляется из командной строки DOS (Start --> Run) следующим образом  
 makeform i код  
 где первый параметр (латинская буква i) нужно

указывать при запуске обязательно; код - необязательный параметр, внутренний код дополнительной клавиши для ввода НеОтвета; узнать значение этого кода для выбранной клавиши можно с помощью программы Scancode.exe, которая входит в состав пакета.



Если не указана дополнительная клавиши ввода НеОтвета; для ввода НеОтвета, можно использовать только клавишу F1.

Если Вы указали «makeform i 20011» в качестве дополнительной клавиши ввода НеОтвета используется клавиша Grey+ ("серый плюс").



После запуска программы ввода нужно ввести имя массива анкет, в который будет осуществляться ввод. Если указывается имя уже существующего массива анкет, то новые анкеты будут добавлены в этот массив ("в конец"

массива). Если же указывается имя нового (еще не существующего) массива, то новый массив автоматически создается.

- ✓ Ввод каждого значения заканчивается нажатием на клавишу Enter
- ✓ Переход к следующему признаку производится автоматически
- ✓ При вводе признака с совместимыми альтернативами все значения вводятся одной строкой (в произвольном порядке, значения разделяются одним или несколькими пробелами). Завершается ввод такого признака так же нажатием на клавишу Enter
- ✓ Для ввода НеОтвета (для всех типов шкал) необходимо:

- ✓ либо нажать клавишу F1 (клавишу Enter нажимать не нужно)
  - ✓ либо ввести знак "\$" и нажать клавишу Enter
  - ✓ либо нажать клавишу, внутренний код которой указан был при запуске (клавишу Enter нажимать не нужно).
- ✓ При вводе осуществляется контроль соответствия вводимой информации паспорту массива, соответствие интервалу допустимых значений для типа шкалы признака, соответствие интервалу, указанному в паспорте массива (директива @R). В случае несоответствия выдается звуковой сигнал и выводится предупреждение на экран. Оператор должен исправить ошибку и снова нажать клавишу Enter
  - ✓ Для того, чтобы при вводе обойти систему контроля допустимых интервалов (т.е. ввести некоторое значение "минуя паспорт") необходимо завершить ввод значения не нажатием клавиши Enter, а нажатием комбинации клавиш Ctl+Enter. При этом автоматически выводится сообщение в файл текстовых замечаний, для чего в массиве должен быть указан ключевой признак (директива @key). Однако, невозможно ввести несколько значений, если признак не объявлен в паспорте как номинальный с совместимыми альтернативами. Невозможно также ввести значения, которые выходят за пределы общего интервала допустимых значений соответствующего типа шкалы.

В процессе ввода активны следующие "горячие" клавиши:

- ✓ **PgUp PgDwn** - перемещение вперед/назад по уже введенной части анкеты
- ✓ **Ctl+PgDwn** - переход к первому невведенному признаку
- ✓ **Ctl+F2** - отказ от уже введенной части анкеты и переход к вводу этой же анкеты "сначала"
- ✓ **F3** - анализ длинного списка альтернатив; если весь список альтернатив не вместился на экране то после нажатия на **F3** можно двигать список вверх/вниз, вправо/влево; кроме того здесь же можно вызвать окно анализа паспорта (с помощью клавиши **F1**); нажатие клавиши **Esc** возвращает к вводу значений
- ✓ **Ctl+T** - вызов окна ввода текстовых замечаний; в это окно можно ввести любой текст (одну строку длиной до 240 символов); ввод завершается нажатием на клавишу Enter, после чего в правой части верхней строки появляется отметка "ТЕКСТ"; текст не заменяет собой значение признака (значение вводится независимо от текста), поэтому ввод текста не приводит к автоматическому переходу на новый признак; обычно в качестве текстовых замечаний вводят ответы на открытые и полуоткрытые вопросы (с целью их последующего кодирования и "закрытия"), заметки интервьюеров и другую текстовую информацию, не предназначенную для непосредственного статистического анализа; эта строка текста "связывается" с анкетой в массиве (для этого в паспорте обязательно должен быть указан ключевой признак - директива @key, который принимает разные значения в разных анкетах) и с признаком (по его номеру); при выводе анкеты для каждого введенного замечания выводится строка в текстовый файл с тем же именем что и имя массива, но с расширением ТХС; эта строка помимо самого текста содержит также информацию о том к какой анкете (значение ключевого признака для этой анкеты) и к какому признаку в этой анкете относится текст.
- ✓ **Esc** - завершение ввода и выход (с подтверждением); если в момент выхода введена часть анкеты, то эта часть на диск не записывается; запись на диск производится только полностью введенными анкетами

Перед вводом последнего признака автоматически подается двоякий звуковой сигнал. Только после ввода последнего признака вся анкета добавляется в массив анкет на диске. Тогда же производится добавление записей в файл текстовых замечаний.

Количество признаков при вводе не должно превышать 1900. Обрабатывать можно массивы и с большим количеством признаков. Если нужно ввести массив с большим количеством признаков, то необходимо разделить каждую анкету на части (размер каждой части не более 1900 признаков), обеспечить наличие в каждой части уникального ключа, каждую из частей ввести в отдельный файл а затем эти файлы соединить по ключу.

## Редактор | Редактирование массива

**Редактор**

- Удаление признаков
- Перемещение признаков
- Удаление анкет

---

- Ввод паспорта
- Редактирование паспорта
- Проверка паспорта

---

- Ввод массива
- Редактирование массива

---

- Работа с замечаниями ▶

Эта операция позволяет отредактировать в массиве ранее введенные значения. Часто редактируют неверно введенные значения (ошибки ввода).

В верхней части экрана находится поле для редактирования значения и две кнопки –

и

Основную часть экрана занимает электронная таблица, которая содержит значения массива.

Прописать
Изменить
1.00

Key=1	#1	#2	#3	#4	#5	#6	#7	#8	
1: 1.00	1.00	4	4	16.00	40.00	7	3	2	1
2: 2.00	2.00	2	\$	15.00	0.00	8 9	2	1	7
3: 3.00	3.00	\$	\$	17.00	5.00	6 9	7	1	1
4: 4.00	4.00	2	\$	15.00	5.00	6	1	1	1
5: 5.00	5.00	3	3	18.00	10.00	8	1	1	7
6: 6.00	6.00	2	\$	15.00	15.00	6 8	1	1	7
7: 7.00	7.00	2	\$	16.00	5.00	6 8	1	1	7
8: 8.00	8.00	2	\$	14.00	35.00	4	3	1	1
9: 9.00	9.00	2	\$	16.00	45.00	6 8	1	1	7
10: 10.00	10.00	2	\$	17.00	50.00	8 9	1	1	7

Столбцы электронной таблицы (кроме самого левого) содержат значения признаков массива. Столбцы помечены номерами признаков (запись вида #...).

Строки электронной таблицы содержат анкеты.

В отдельных клеточках таблицы находятся значения. Одна клеточка в таблице выделена рамочкой. Значение этой выделенной клеточки выводится в поле для редактирования значений (в верхней части окна). (В указанном выше примере в поле для редактирования – значение признака 1 для анкеты 1.)

Перемещение рамочки по таблице производится клавишами движения курсора (вверх, вниз, влево, вправо) либо можно указать мышью на необходимую клеточку и щелкнуть. Для быстрого перемещения по электронной таблице можно использовать стандартные полосы прокрутки (вертикальную и горизонтальную полосы прокрутки).

В верхней клеточке левого столбца (запись вида ) указан номер ключевого признака массива (ключевой признак указывается в паспорте директивой @key). Каждая строка электронной таблицы (анкета массива) помечена в левом столбце конструкцией вида m:n, где

- ✓ m – номер анкеты по порядку ее записи в массив

- ✓ n – значение ключевого признака для этой анкеты

Например, пометка  означает, что в соответствующей строке записана 9-я анкета (по порядку записи анкет в массив), значение ключевого признака в этой анкете равно 9.

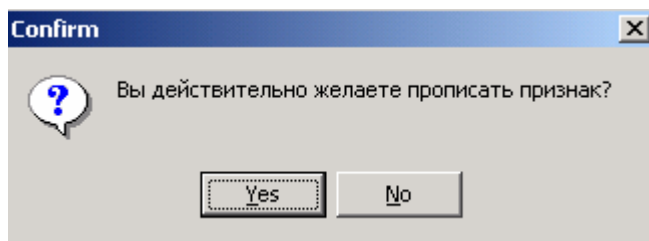
Редактирование можно осуществлять через фильтр. Это означает, что если установлен фильтр, то на экран выводятся (а значит и доступны для редактирования) только те анкеты, которые удовлетворяют этому фильтру.

Для изменения значения в клеточке таблицы необходимо

- ✓ установить рамку на эту клеточку таблицы; при этом значение, которое содержится в этой клеточке, появится также в поле редактирования
- ✓ изменить значение в поле редактирования
- ✓ нажать клавишу Enter или кнопку

После этого значение будет сразу изменено в файле на диске

Если нажать кнопку  и подтвердить свое решение, то это значение будет записано во все клеточки столбца, в котором находится рамка. Это бывает полезно, если необходимо изменить значение признака не в одной, а в целой группе анкет. В таком случае строится фильтр, который отбирает необходимые анкеты, и затем редактирование осуществляется с помощью кнопки .



Во время редактирования в соответствии с паспортом производятся автоматически проверки на корректность редактируемых значений. Однако, интервалы для метрических признаков (которые задаются в паспорте директивой @R) не проверяются.

**Внимание.** Операция редактирования не имеет возможности "отката" (undo). Таким образом, в случае ошибки во время редактирования восстановить старые значения невозможно – запись отредактированных значений осуществляется непосредственно в файл на диске. Особенно неприятным может быть результат изменения значений в группе анкет, в случае если группа неверно отобрана. Поэтому перед редактированием рекомендуем сделать копию массива.

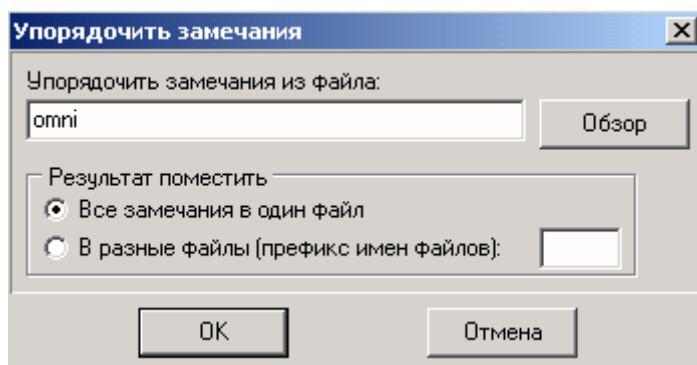
## Редактор | Работа с замечаниями | Упорядочить замечания

<p>Редактор</p> <ul style="list-style-type: none"> <li>Удаление признаков</li> <li>Перемещение признаков</li> <li>Удаление анкет</li> <li> Ввод паспорта</li> <li>Редактирование паспорта</li> <li>Проверка паспорта</li> <li> Ввод массива</li> <li>Редактирование массива</li> <li><b>Работа с замечаниями</b> <ul style="list-style-type: none"> <li>упорядочить замечания</li> <li>внести изменения</li> </ul> </li> </ul>	<p>Если во время ввода анкет вводились также текстовые замечания (ответы на открытые или полуоткрытые вопросы, замечания интервьюеров и т.п.), то после окончания ввода часто необходимо переупорядочить (отсортировать) текстовые замечания таким образом, чтобы собрать вместе все замечания, которые имеют отношение к одному признаку. Это позволяет в дальнейшем эффективно организовать анализ и кодирование ответов на открытые и полуоткрытые вопросы.</p>
--	--

В протокол работы выводится информация о том, для каких из признаков массива признаков есть замечания, количество замечаний к каждому из признаков и общее количество замечаний.

Результат сортировки текстовых замечаний может размещаться:

- ✓ либо в одном файле. В этом случае замечания переупорядочиваются внутри файла таким образом, что сначала идет группа замечаний к признаку с наименьшим номером, затем со следующим по порядку большим номером, и т.д. Внутри каждой из групп замечания не упорядочиваются.
- ✓ либо в нескольких файлах. Группа замечаний, относящихся к одному признаку массива, выводится в отдельный файл. Внутри каждого из таких файлов замечания не упорядочиваются.



#### Параметры и кнопки операции:

- ✓ Упорядочить замечания из файла – поле для ввода имени файла с замечаниями. Как правило, файлы с замечаниями имеют расширения ТХС. Кнопка **Обзор** позволяет открыть стандартное окно выбора файла на диске

Результат поместить – группа из двух радиокнопок для выбора способа представления результатов сортировки (в виде одного файла или в виде нескольких файлов).

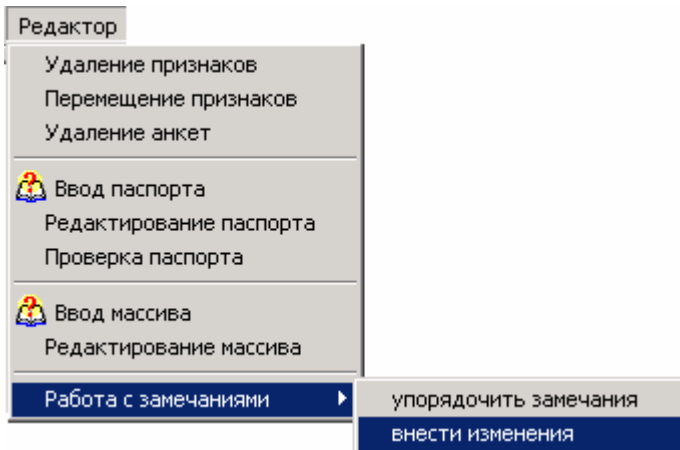
**Все замечания в один файл** радиокнопка, позволяющая получить результат сортировки в виде одного файла. Этот файл с результатом работы операции будет иметь то же имя, что и исходный файл. Исходный файл остается неизменным, но в его имени расширение меняется на расширение ВАК. В отсортированном файле группы текстовых замечаний, относящихся к одному признаку, отделяются одна от другой специальными комментариями (строка комментария начинается символом "точка с запятой" – ";"). Если в исходном файле есть строки замечаний с неверной структурой (например, в результате опечаток при редактировании файла замечаний каким-либо текстовым редактором), то информация об этих строках выводится в протокол работы, а сами эти строки размещаются в начале упорядоченного файла. Строки с комментариями из исходного файла в отсортированный не переносятся.

**В разные файлы (префикс имен файлов):**  радиокнопка, позволяющая получить результат упорядочения в виде нескольких файлов. При этом необходимо задать префикс имен этих файлов (до 6 знаков, допустимых при записи имен файлов). В качестве окончания имен и расширений для отсортированных файлов используются номера соответствующих признаков. Эти файлы создаются автоматически. Если в момент выполнения операции файлы с таким именами уже существуют, то производится добавление в эти существующие файлы.

Например, если в массиве есть текстовые замечания к признакам с номерами 5, 37 и 367, в качестве префикса имен указано слово order, и до начала сортировки файл order00.367 уже существует, то в результате сортировки будут созданы файл order00.005 (содержит замечания к признаку 5) и order00.037 (содержит замечания к признаку 37) и кроме того информация будет дописана в уже существующий файл order00.367.

Если в исходном файле есть строки замечаний с неверной структурой (например, в результате опечаток при корректировании файла замечаний каким-либо текстовым редактором), то информация об этих строках выводится в протокол работы, а сами эти строки размещаются в файле, имеющем расширение 000. Так, в приведенном выше примере строки с неверной структурой будут выведены в файл order00.000. Строки с комментариями из исходного файла в отсортированные файлы не переносятся.

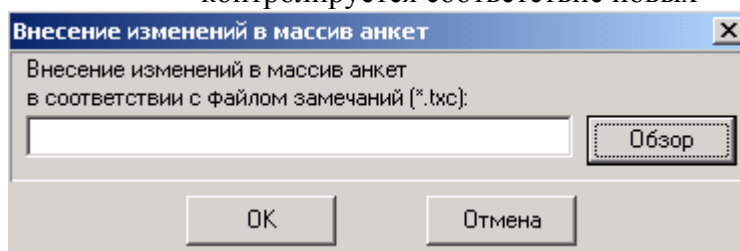
## Редактор | Работа с замечаниями | Упорядочить замечания



Эта операция позволяет произвести редактирование массива анкет (т.е. изменить значения определенных признаков в определенных анкетах массива). Это бывает необходимо, например, при внесении в массив анкет результатов кодирования открытых и полуоткрытых вопросов.

Изменения должны быть заранее подготовлены в виде файла в формате ТХС. При внесении изменений контролируется соответствие новых

значений описанию соответствующего признака в паспорте массива. В частности, при закрытии открытых вопросов, сначала необходимо внести изменения в паспорт, а лишь затем редактировать значения.



## Статистика | Одномерные распределения

<p>Статистика</p> <ul style="list-style-type: none"> <li>Одномерные распределения</li> <li>Двумерные распределения</li> <li>Статистические гипотезы</li> <li>Объем выборки</li> </ul>	<p>Таблицы одномерных распределений и соответствующие статистики строятся либо для всех анкет массива либо для анкет, отобранных с помощью фильтра. Операция позволяет строить таблицы частот и процентов, рассчитывать меры центральной тенденции<sup>i</sup> и меры вариации<sup>ii</sup> для распределения значений признака (стандартное отклонение, дисперсия и др.).</p>
---	--

Пакет, исходя из типа шкалы признака, самостоятельно определяет вычисление каких показателей является корректным.

**Метрическая шкала:** вычисляются минимальное значение (Минимум), максимальное значение (Максимум), среднее арифметическое (Среднее), среднеквадратическое отклонение (СрКвОткл) и коэффициент вариации (КоефВар; отношение среднеквадратического отклонения к среднему). Вычисления производятся для тех анкет, которые не содержат значение НеОтвет для данного признака.

Признак #164. Длительность интервью в минутах			
Всего 1653. Ответили 1645 (99,5%).			
Минимум	1,000	Максимум	110,000
Среднее	49,383	СрКвОткл	15,827
КоефВар	0,320		



**Порядковая шкала:** строится таблица частот и процентов; также могут вычисляться среднее арифметическое (Среднее), среднеквадратическое отклонение (СрКвОткл) и коэффициент вариации (КоефВар). Вычисление показателей производится для тех анкет, которые не содержат значение НеОтвет для данного признака.

Признак #107. К2. Как часто Вы обычно пьете кофе?  
Всего 1653. Ответили 968 (58,6%).

	Среднее	5,040	СрКвОткл	1,693
	КоефВар	0,336		
	Код	Час- тота	% ко всем	% к ответ
-----				
Более 5 раз в день	1	17	1,0	1,8
4-5 раз в день	2	21	1,3	2,2
2-3 раза в день	3	179	10,8	18,5
Один раз в день	4	235	14,2	24,3
4-5 раз в неделю	5	52	3,1	5,4
2-3 раза в неделю	6	165	10,0	17,0
Реже	7	299	18,1	30,9
-----				

**Номинальная шкала:** строится таблица частот и процентов.

Признак #90. 74. Я согласен допустить евреев как ...  
Всего 1653. Ответили 1579 (95,5%).

	Код	Час- тота	% ко всем	% к ответ
-----				
членов семьи	1	296	17,9	18,7
друзей	2	279	16,9	17,7
соседей	3	239	14,5	15,1
коллег	4	85	5,1	5,4
жителей Украины	5	234	14,2	14,8
гостей Украины	6	334	20,2	21,2
не пускал бы в Украи	7	112	6,8	7,1
-----				

**Номинальная шкала с совместимыми альтернативами:** строится таблица частот и процентов. В такой таблице сумма процентов может превышать 100%.

Признак #101. А1. Как Вы узнаете о том, что западные товары имеются в продаже?  
Всего 1653. Ответили 1632 (98,7%).

	Код	Час- тота	% ко всем	% к ответ
-----				
разговаривая с знако	1	678	41,0	41,5
разговаривая с детьм	2	200	12,1	12,3
разговаривая с иност	3	24	1,5	1,5
увидев в киоске	4	990	59,9	60,7
увидев где-либо	5	372	22,5	22,8
из рекламы на вывеск	6	170	10,3	10,4
из рекламы по ТВ	7	906	54,8	55,5
из радиорекламы	8	203	12,3	12,4
из газетной рекламы	9	255	15,4	15,6
из западных журналов	10	49	3,0	3,0
другое	11	41	2,5	2,5
ТС	12	152	9,2	9,3
-----				

Вывод таблиц и статистик осуществляется в окно вывода результатов (формат ANSI, обычный текст) а также, при необходимости, в файл на диске (формат HTML, который можно читать программами MS Word 97/2000, MS Excel 97/2000, Internet Explorer).

Если до выполнения операции установлен фильтр, то построение таблиц и вычисление статистик производится для анкет, которые удовлетворяют фильтру.

Для выполнения операции необходимо:

- ✓ ввести список номеров признаков в поле **Номера признаков:**
- ✓ указать, нужно ли выводить таблицы частот и процентов для номинальных, порядковых и номинальных с совместимыми альтернативами признаков

**Выводить таблицы частот и % для ...**  
 ... номинальных и порядковых

Выводить в таблице:

пустые строки

Выводить в таблице:

пустые строки

строку для "Не Ответ"

Выводить в таблице:

пустые строки

строку для "Не Ответ"

кумулятивные %

Выводить в таблице:

пустые строки

строку для "Не Ответ"

кумулятивные %

коды альтернатив

Выводить

текст вопроса

название признака

Упорядочивать строки

Точность вывода % (0-3) 1

Названия альтернатив 20 знаков

Вывод статистик

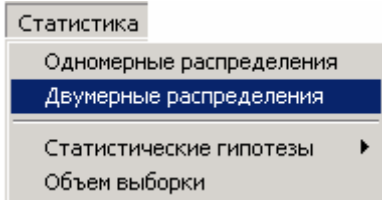
Точность вывода статистик (1-5) 3

- ✓ указать, нужно ли выводить в таблицах частот и процентов для номинальных, порядковых и номинальных с совместимыми альтернативами признаков строки с нулевой частотой
- ✓ указать, нужно ли выводить в таблицах частот и процентов для номинальных, порядковых и номинальных с совместимыми альтернативами признаков строку для НеОвета
- ✓ указать, нужно ли выводить в таблицах частот и процентов для номинальных, порядковых и номинальных с совместимыми альтернативами признаков столбец кумулятивных процентов
- ✓ указать, нужно ли выводить в таблицах частот и процентов для номинальных, порядковых и номинальных с совместимыми альтернативами признаков столбец с кодами значений
- ✓ выбрать, что выводить в заголовке таблицы – текст вопроса или название признака из паспорта массива
- ✓ указать нужно ли упорядочивать строки в таблице по возрастанию частот или строки должны идти в том порядке, в котором альтернативы перечислены в паспорте массива
- ✓ указать, сколько знаков после десятичной точки нужно выводить в значениях процентов
- ✓ указать, сколько знаков из названия альтернатив нужно использовать при выводе таблиц
- ✓ указать, нужно ли выводить статистики распределений (средние, меры вариации и т.п.)
- ✓ указать, сколько знаков после десятичной точки нужно выводить в значениях статистик (допустимые значения – от 1 до 5)
- ✓ если есть необходимость вывода результата работы в формате HTML, нажать кнопку **Вывод в файл** и выбрать каталог и имя файла для протокола; если указать имеющийся

на диске файл, то производится дописывание построенных таблиц "в конец" этого файла

- ✓ нажать кнопку 

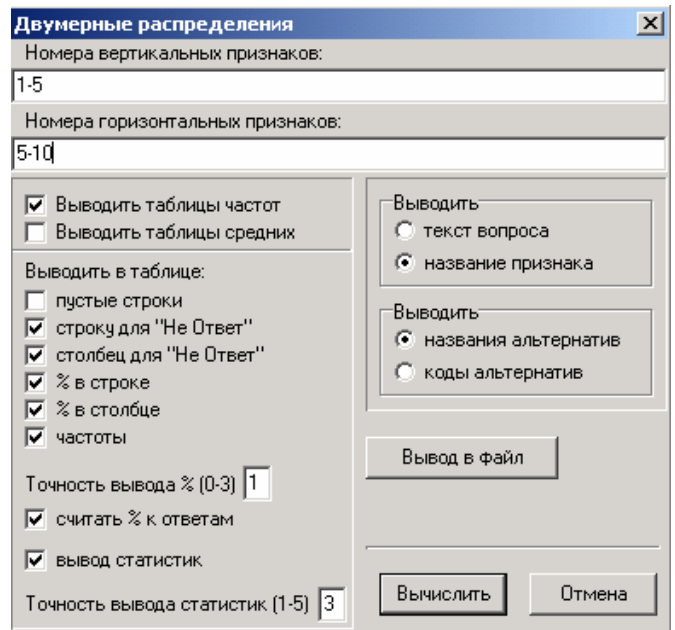
## Статистика | Двумерные распределения



Таблицы двумерных распределений и соответствующие статистики строятся либо для всех анкет массива либо для анкет, отобранных с помощью фильтра. Операция позволяет строить таблицы частот и процентов, рассчитывать коэффициенты связи между двумя признаками.

Для получения информации о двумерном распределении для одной или нескольких пар признаков необходимо указать два списка номеров:

- ✓ список номеров "вертикальных" признаков (т.е. тех, названия альтернатив которых в таблице размещаются по вертикали и помечают строки таблицы)
- ✓ список номеров "горизонтальных" признаков (т.е. тех, названия альтернатив которых в таблице размещаются по горизонтали и помечают столбцы таблицы)



Из двух списков образуются всевозможные пары номеров признаков (в каждой паре первый элемент берется из списка "вертикальных", а второй - из списка "горизонтальных") и для каждой образованной таким образом пары номеров признаков производятся вычисления.

**Внимание.** Общее количество пар признаков не должно превышать 500.

Если установлен фильтр, то вычисления производятся только для анкет, удовлетворяющих фильтру.

Состав информации о двумерном распределении для пары признаков зависит от типов шкал признаков, образующих пару.

### Если оба признака метрические

Признак #155. Б9. Ваш общий среднемесячный доход  
 Признак #154. Б8. Сколько человек живет с Вами?  
 Всего анкет 1653. Без ответа 666  
 Коэффициент корреляции Пирсона 0,193 (значим 1%)  
 Уравнения линейной регрессии

#155#=-	1081,034*#154#+	5385,080
#154#=-	0,000*#155#+	3,210

- ✓ вычисляется коэффициент корреляции Пирсона и оценивается его значимость на двух уровнях -- 5% и 1%;
- ✓ строятся два уравнения парной линейной регрессии (в одном зависимой

переменной выступает "горизонтальный" признак, во втором зависимой переменной выступает "вертикальный" признак).

- ✓ Все показатели вычисляются только для тех анкет, в которых оба признака имеют значения, отличные от НеОтвета.

### Если один признак метрический, а второй порядковый

✓ вычисляется коэффициент корреляции Пирсона и оценивается его значимость на уровне 5% и на уровне 1%:						
Признак #155. Б9.Ваш общий среднемесячный доход Признак #11. 6.Изменение материального положения за последние 3-4 года Всего анкет 1653. Без ответа 643 Коэффициент корреляции Пирсона 0,160 (значим 1%)						
✓ строится таблица групповых средних (т.е. таблица средних значений метрического признака для групп, выделенных значениями порядкового признака):						
Знач #11	Част #11	Min #155	Max #155	Сред #155	СрКв #155	КфВр #155
-----	-----	-----	-----	-----	-----	-----
значительн о ухудшило	612	0,000	90000,000	8137,371	7417,169	0,911
несколько ухудшилось	195	0,000	40000,000	9963,728	8133,738	0,816
почти не и зменилось,	131	1,000	50000,000	9653,321	9365,255	0,970
несколько улучшилось	60	1,000	80000,000	14448,433	13465,698	0,932
значительн о улучшило	12	1,000	20000,000	10358,417	6688,718	0,646
-----	-----	-----	-----	-----	-----	-----
✓ строится два уравнения парной линейной регрессии (в одной зависимой переменной выступает метрический признак, во второй зависимой переменной выступает порядковый признак):						
Уравнения линейной регрессии #155# = 1363,313 * #11# +6799,975 #11# = 0,000 * #155# +1,507						
✓ строится матрица значимости различий между групповыми средними. Значимость различий оценивается для каждой пары групповых средних отдельно, множественные сравнения не производятся. Оценивается значимость различий групповых средних на уровне 5% и на уровне 1% (знак минус "-" означает, что различие не значимо на указанном уровне)						
Матрица результатов t-теста для пар средних ("5%" означает "различие не значимо на уровне 5%")						
	9963,728	9653,321	14448,433	10358,417		
8137,371	1%	5%	1%	5%		
9963,728		5%	5%	5%		
9653,321			5%	5%		
14448,433				5%		

Все показатели вычисляются только для тех анкет, в которых оба признака имеют значения, отличные от НеОтвета.

**Если один признак метрический, а второй номинальный**

- ✓ строится таблица групповых средних (т.е. таблица средних значений метрического признака для групп, выделенных значениями номинального признака):

Признак #155. Б9.Ваш общий среднемесячный доход  
 Признак #138. Пол  
 Всего анкет 1653. Без ответа 641

Знач #138	Част #138	Min #155	Max #155	Сред #155	СрКв #155	КфВр #155
мужской	459	1,000	90000,000	9959,821	9350,612	0,939
женский	553	0,000	42000,000	8350,329	7477,826	0,896

- ✓ строится матрица значимости различий между групповыми средними. Значимость различий оценивается для каждой пары групповых средних отдельно, множественные сравнения не производятся. Оценивается значимость различий групповых средних на уровне 5% и на уровне 1% (знак минус "-" означает, что различие не значимо на указанном уровне)

Матрица результатов t-теста для пар средних ("-5%" означает "различие не значимо на уровне 5%")

	8350,329
9959,821	1%

Все показатели вычисляются только для тех анкет, в которых оба признака имеют значения, отличные от НеОтвета.

**Если один признак метрический, а второй номинальный с совместимыми альтернативами**

- ✓ строится таблица групповых средних (т.е. таблица средних значений метрического признака для групп, выделенных значениями номинального признака с совместимыми альтернативами)

Признак #155. Б9.Ваш общий среднемесячный доход  
 Признак #26. 21.Какая из проблем Вас беспокоит больше всего  
 Всего анкет 1653. Без ответа 643

Знач #26	Част #26	Min #155	Max #155	Сред #155	СрКв #155	КфВр #155
безопасность Украины	263	0,0	42000,0	8576,9	7288,7	0,8
возрождение украинские отношения с Россией	103	1,0	42000,0	10450,7	7662,1	0,7
преступность в результате Чернобыля	292	1,0	50000,0	9787,7	7846,7	0,8
проблема Крыма	555	0,0	90000,0	9157,7	8461,4	0,9
проблема Черноморского флота	213	0,0	35000,0	8185,6	6854,6	0,8
уровень жизни людей	39	1,0	30000,0	8905,2	6992,6	0,8
положение русского языка	26	900,0	30000,0	10688,5	6590,4	0,6
другая	724	0,0	90000,0	9324,6	8321,4	0,9
ТС	54	1,0	50000,0	12477,8	10015,2	0,8
	19	1,0	40000,0	11026,5	11035,7	1,0
	17	1,0	80000,0	8806,1	18974,3	2,2

- ✓ строится матрица значимости различий между групповыми средними. Значимость различий оценивается для каждой пары групповых средних отдельно, множественные сравнения не производятся. Оценивается значимость различий групповых средних на уровне 5% и на уровне 1% (знак минус "-" означает, что различие не значимо на указанном уровне)

Матрица результатов t-теста для пар средних ("–5%" означает "различие не значимо на уровне 5%")

	10450,7	9787,7	9157,7	8185,6	8905,2	10688,5	9324,6	12477,8	11026,5	8806,1
8576,9	5%	-5%	-5%	-5%	-5%	-5%	-5%	-5%	1%	-5%
10450,7		-5%	-5%	5%	-5%	-5%	-5%	-5%	-5%	-5%
9787,7			-5%	5%	-5%	-5%	-5%	-5%	-5%	-5%
9157,7				-5%	-5%	-5%	-5%	5%	-5%	-5%
8185,6					-5%	-5%	5%	1%	-5%	-5%
8905,2						-5%	-5%	5%	-5%	-5%
10688,5							-5%	-5%	-5%	-5%
9324,6								5%	-5%	-5%
12477,8									-5%	-5%
11026,5										-5%

**Внимание.** Интерпретация таблиц и показателей требует определенной осторожности. Необходимо учитывать, что группы, которые выделяются отдельными значениями признака с совместимыми альтернативами, могут пересекаться (т.е. одна и та же анкета может принадлежать нескольким группам). Вследствие этого, сумма частот в группах может превышать количество анкет в массиве и одна и та же анкета может принимать участие в вычислении не одного, а нескольких групповых средних.

**Если оба признака порядковые**

<p>Признак #80. 66.Ваша оценка США Признак #77. 63.Ваша оценка Канады Всего анкет 1653</p> <table border="1"> <thead> <tr> <th></th> <th>НеОтв</th> <th>весьма полож ительн</th> <th>скорее полож ительн</th> <th>ТС</th> <th>скорее отриц ательн</th> <th>весьма отриц ательн</th> <th>ВСЕГО</th> </tr> </thead> <tbody> <tr> <td>НеОтв</td> <td>6</td> <td>3</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>9</td> </tr> <tr> <td>весьма полож ительн</td> <td>0</td> <td>77,5 447</td> <td>18,0 104</td> <td>1,6 9</td> <td>1,9 11</td> <td>1,0 6</td> <td>100,0 577</td> </tr> <tr> <td>скорее полож ительн</td> <td>0</td> <td>15,8 87</td> <td>75,0 413</td> <td>4,0 22</td> <td>4,5 25</td> <td>0,7 4</td> <td>100,0 551</td> </tr> <tr> <td>ТС</td> <td>1</td> <td>7,9 38</td> <td>10,0 48</td> <td>77,4 370</td> <td>2,9 14</td> <td>1,7 8</td> <td>100,0 478</td> </tr> <tr> <td>скорее отриц ательн</td> <td>0</td> <td>20,0 6</td> <td>16,7 5</td> <td>0,0 0</td> <td>60,0 18</td> <td>3,3 1</td> <td>100,0 30</td> </tr> <tr> <td>весьма отриц ательн</td> <td>0</td> <td>0,0 0</td> <td>0,0 0</td> <td>14,3 1</td> <td>0,0 0</td> <td>85,7 6</td> <td>100,0 7</td> </tr> <tr> <td>ВСЕГО</td> <td>7</td> <td>35,2 578</td> <td>34,7 570</td> <td>24,5 402</td> <td>4,1 68</td> <td>1,5 25</td> <td>1643</td> </tr> </tbody> </table>		НеОтв	весьма полож ительн	скорее полож ительн	ТС	скорее отриц ательн	весьма отриц ательн	ВСЕГО	НеОтв	6	3	0	0	0	0	9	весьма полож ительн	0	77,5 447	18,0 104	1,6 9	1,9 11	1,0 6	100,0 577	скорее полож ительн	0	15,8 87	75,0 413	4,0 22	4,5 25	0,7 4	100,0 551	ТС	1	7,9 38	10,0 48	77,4 370	2,9 14	1,7 8	100,0 478	скорее отриц ательн	0	20,0 6	16,7 5	0,0 0	60,0 18	3,3 1	100,0 30	весьма отриц ательн	0	0,0 0	0,0 0	14,3 1	0,0 0	85,7 6	100,0 7	ВСЕГО	7	35,2 578	34,7 570	24,5 402	4,1 68	1,5 25	1643	<ul style="list-style-type: none"> <li>✓ строится таблица частот и процентов (корреляционная таблица, таблица сопряженности)</li> <li>✓ вычисляются коэффициенты связи для номинальных шкал -- Хи-квадрат Пирсона и построенные на его основе коэффициенты Чупрова и Крамера, коэффициент лямбда Гудмана (несимметричный показатель связи, построенный на принципе уменьшения ошибки предсказания); производится оценка значимости Хи-квадрат Пирсона на уровне 5% и на уровне 1%:</li> </ul> <table border="1"> <tr> <td>Коэффициент Хи-квадрат</td> <td>2188,6 (значим 1%)</td> </tr> <tr> <td>Число степеней свободы</td> <td>16</td> </tr> <tr> <td>Коэффициент Чупрова</td> <td>0,6</td> </tr> <tr> <td>Коэффициент Крамера</td> <td>0,6</td> </tr> <tr> <td>Коэффициент лямбда Гудмана</td> <td></td> </tr> <tr> <td>    влияние #80# на #77#</td> <td>0,6</td> </tr> <tr> <td>    влияние #77# на #80#</td> <td>0,6</td> </tr> </table>	Коэффициент Хи-квадрат	2188,6 (значим 1%)	Число степеней свободы	16	Коэффициент Чупрова	0,6	Коэффициент Крамера	0,6	Коэффициент лямбда Гудмана		влияние #80# на #77#	0,6	влияние #77# на #80#	0,6
	НеОтв	весьма полож ительн	скорее полож ительн	ТС	скорее отриц ательн	весьма отриц ательн	ВСЕГО																																																																								
НеОтв	6	3	0	0	0	0	9																																																																								
весьма полож ительн	0	77,5 447	18,0 104	1,6 9	1,9 11	1,0 6	100,0 577																																																																								
скорее полож ительн	0	15,8 87	75,0 413	4,0 22	4,5 25	0,7 4	100,0 551																																																																								
ТС	1	7,9 38	10,0 48	77,4 370	2,9 14	1,7 8	100,0 478																																																																								
скорее отриц ательн	0	20,0 6	16,7 5	0,0 0	60,0 18	3,3 1	100,0 30																																																																								
весьма отриц ательн	0	0,0 0	0,0 0	14,3 1	0,0 0	85,7 6	100,0 7																																																																								
ВСЕГО	7	35,2 578	34,7 570	24,5 402	4,1 68	1,5 25	1643																																																																								
Коэффициент Хи-квадрат	2188,6 (значим 1%)																																																																														
Число степеней свободы	16																																																																														
Коэффициент Чупрова	0,6																																																																														
Коэффициент Крамера	0,6																																																																														
Коэффициент лямбда Гудмана																																																																															
влияние #80# на #77#	0,6																																																																														
влияние #77# на #80#	0,6																																																																														
<ul style="list-style-type: none"> <li>✓ вычисляются коэффициент корреляции Пирсона и ранговый коэффициент корреляции Спирмена; производится оценка их значимости на уровне 5% и на уровне 1%</li> </ul> <table border="1"> <tr> <td>Коэффициент корреляции Пирсона</td> <td>0,7 (значим 1%)</td> </tr> <tr> <td>Коэф. ранговой корреляции Спирмена</td> <td>0,7 (значим 1%)</td> </tr> </table>	Коэффициент корреляции Пирсона	0,7 (значим 1%)	Коэф. ранговой корреляции Спирмена	0,7 (значим 1%)	<ul style="list-style-type: none"> <li>✓ строится два уравнения парной линейной регрессии (в одном зависимой переменной выступает "горизонтальный" признак, во втором зависимой переменной выступает "вертикальный" признак)</li> </ul> <p>Уравнения линейной регрессии</p> <p>#80# = 0,7 * #77# + 0,6</p> <p>#77# = 0,6 * #80# + 0,</p>																																																																										
Коэффициент корреляции Пирсона	0,7 (значим 1%)																																																																														
Коэф. ранговой корреляции Спирмена	0,7 (значим 1%)																																																																														

Все показатели вычисляются только для тех анкет, в которых оба признака имеют значения, отличные от НеОвета.

**Если один признак порядковый, а второй номинальный**

Признак #77. 63.Ваша оценка Канады  
 Признак #138. Пол  
 Всего анкет 1653

	НеОтв	весьма полож ительн	скорее полож ительн	ТС	скорее отриц ательн	весьма отриц ательн	ВСЕГО
мужско й	4	41,4 290	35,0 245	21,0 147	1,7 12	0,9 6	100,0 700
		50,3	44,5	30,7	40,0	85,7	42,6
женски й	5	30,4 287	32,4 306	35,2 332	1,9 18	0,1 1	100,0 944
		49,7	55,5	69,3	60,0	14,3	57,4
ВСЕГО	9	35,1 577	33,5 551	29,1 479	1,8 30	0,4 7	100,0 1644
		100,0	100,0	100,0	100,0	100,0	

✓ строится таблица частот и процентов (корреляционная таблица, таблица сопряженности)

Коэффициент Хи-квадрат	47,8 (значим 1%)
Число степеней свободы	4
Хи-квадрат не надежен - 2 клеток с ожидаемой частотой < 5	
Коэффициент Чупрова	0,1
Коэффициент Крамера	0,2
Коэффициент лямбда Гудмана	
влияние #77# на #138#	0,0
влияние #138# на #77#	0,0

вычисляются коэффициенты связи для номинальных шкал -- Хи-квадрат Пирсона и построенные на его основе коэффициенты Чупрова и Крамера, коэффициент лямбда Гудмана

(несимметричный показатель связи, построенный на принципе уменьшения ошибки предсказания); производится оценка значимости Хи-квадрат Пирсона на уровне 5% и на уровне 1%

Все показатели вычисляются только для тех анкет, в которых оба признака имеют значения, отличные от НеОтвета.

**Если один признак порядковый, а второй номинальный с совместимыми альтернативами**

Признак #11. 6.Изменение материального положения за последние 3-4 года  
 Признак #26. 21.Какая из проблем Вас беспокоит больше всего  
 Всего анкет 1653

	НеОтв	значит ельно ухудши лось	нескол ко ухуд шилось	почти не изм енилось	нескол ко улуч шилось	значит ельно улучши лось	ВСЕГО
НеОтв	0	1	1	0	0	0	2
безопас ность	1	53,6 199	23,2 86	16,4 61	5,1 18	1,6 6	100,0 371
Украи на		19,2	30,5	29,6	20,0	24,0	22,5
возрож дение	1	48,0 73	23,7 86	17,1 26	6,6 10	4,6 7	100,0 152
укр на отноше ния с Россие	1	67,9 30,1	17,1 28,0	10,2 22,8	3,9 18,9	0,9 16,0	100,0 28,0
престу пность	3	63,3 54,3	16,5 52,1	12,0 51,9	6,4 60,0	1,8 64,0	100,0 54,1
послед ствия Черноб ур	0	57,5 17,0	19,8 21,6	13,6 20,4	7,8 25,3	1,3 16,0	100,0 18,7
пробле ма Кры ма	0	53,2 41	16,9 13	24,7 19,8	5,2 19	0,0 4	100,0 77
пробле ма Чер номор ья	0	49,0 2,3	16,3 2,8	22,4 5,3	8,2 4,2	4,1 8,0	100,0 3,0
уровен ь жизни и ладе	4	66,1 780	16,0 189	10,8 128	5,7 67	1,4 16	100,0 1180
		75,1	67,0	62,1	70,5	64,0	71,6
положе ние ру с язык с	0	64,4 56	13,8 12	13,8 12	5,7 5	2,3 2	100,0 87
		5,4	4,3	5,8	5,3	8,0	5,3
другая	0	40,5 15	16,2 6	32,4 12	10,8 4	0,0 0	100,0 37
		1,4	2,1	5,8	4,2	0,0	2,2
ТС	0	78,1 25	12,5 4	6,3 2	3,1 1	0,0 0	100,0 32
		2,4	1,4	1,0	1,1	0,0	1,9
ВСЕГО	4	63,1 1039	17,1 282	12,5 206	5,8 95	1,5 25	100,0 1647
		100,0	100,0	100,0	100,0	100,0	

✓ строится таблица частот и процентов (корреляционная таблица, таблица сопряженности)

✓ вычисляются Хи-квадрат Пирсона и построенные на его основе коэффициенты связи для номинальных признаков (коэффициенты Чупрова и Крамера); производится оценка значимости Хи-квадрат на уровне 5% и на уровне 1%

Коэффициент Хи-квадрат	97,8 (значим 1%)
Число степеней свободы	40
Коэффициент Чупрова	0,1
Коэффициент Крамера	0,1

**Внимание.** Интерпретация таблиц и показателей требует определенной осторожности.

При анализе результатов необходимо учитывать, что группы, которые выделяются отдельными значениями признака с совместимыми альтернативами, могут пересекаться (т.е. одна и та же анкета может принадлежать нескольким группам). Сумма процентов в строке (столбце) таблицы может превышать 100%, а сумма частот превышать значение,

указанное, в маргинальном столбце (маргинальной строке). Коэффициент Хи-квадрат вычисляется не по отношению к количеству анкет, а по отношению к количеству сделанных выборов в номинальном с совместимыми альтернативами признаке (количество выборов может быть существенно больше количества анкет, поскольку во многих анкетах может быть выбрана не одна а несколько альтернатив). Кроме того, одна и та же анкета может принимать участие в вычислениях не одного, а нескольких групповых средних.

### Если оба признака номинальные

Признак #138. Пол  
Признак #106. К1.Вы пьете кофе?  
Всего анкет 1653

	мужско й	женски й	ВСЕГО
НеОтв	6	2	8
Да	43,0 418 59,9	57,0 553 58,4	100,0 971 59,0
Нет	41,5 280 40,1	58,5 394 41,6	100,0 674 41,0
ВСЕГО	42,4 698	57,6 947	100,0 1645

- ✓ строится таблица частот и процентов (корреляционная таблица, таблица сопряженности)
- ✓ вычисляются коэффициенты связи для номинальных шкал -- Хи-квадрат Пирсона и построенные на его основе коэффициенты Чупрова и Крамера, коэффициент лямбда Гудмана (несимметричный показатель связи, построенный на принципе уменьшения ошибки предсказания); производится оценка значимости Хи-квадрат Пирсона на уровне 5% и на уровне 1%

Коэффициент Хи-квадрат	0,4 (не значим 5%)
Число степеней свободы	1
Коэффициент Чупрова	0,0
Коэффициент Крамера	0,0
Коэффициент лямбда Гудмана	
влияние #138# на #106#	0,0
влияние #106# на #138#	0,0

Все показатели вычисляются только для тех анкет, в которых оба признака имеют значения, отличные от НеОтвета



**Если один признак номинальный, а второй номинальный с совместимыми альтернативами**

Признак #138. Пол  
Признак #21. 16.Что Вы будете делать, если лишитесь работы?  
Всего анкет 1653

	мужско й	женски й	ВСЕГО
НеОтв	10	14	24
Попыта юсь от крыть	57,0 106 15,3	43,0 80 8,6	100,0 186 11,4
Соглаш усь на любую	46,8 195 28,1	53,2 222 23,7	100,0 417 25,6
Потреб ую пом ощи у	45,3 29 4,2	54,7 35 3,7	100,0 64 3,9
Вложу свои с береже	58,8 10 1,4	41,2 7 0,7	100,0 17 1,0
Займус ь само обеспе	49,7 213 30,7	50,3 216 23,1	100,0 429 26,3
Ничего не см огу сд	29,9 44 6,3	70,1 103 11,0	100,0 147 9,0
Не зна ю, что смогу	36,9 124 17,9	63,1 212 22,7	100,0 336 20,6
Другое	37,7 75 10,8	62,3 124 13,3	100,0 199 12,2
ТС	42,9 48 6,9	57,1 64 6,8	100,0 112 6,9
ВСЕГО	42,6 694 100,0	57,4 935 100,0	100,0 1629

- ✓ строится таблица частот и процентов (корреляционная таблица, таблица сопряженности)
- ✓ вычисляются Хи-квадрат Пирсона и построенные на его основе коэффициенты связи для номинальных признаков (коэффициенты Чупрова и Крамера); производится оценка значимости Хи-квадрат на уровне 5% и на уровне 1%

Коэффициент Хи-квадрат	43,0 (значим 1%)
Число степеней свободы	8
Хи-квадрат вычислен для	1907 ответов
Коэффициент Чупрова	0,1
Коэффициент Крамера	0,2

**Внимание.** Интерпретация таблиц и показателей требует определенной осторожности. При анализе результатов необходимо учитывать, что группы, которые выделяются отдельными значениями признака с совместимыми альтернативами, могут пересекаться (т.е. одна и та же анкета принадлежит нескольким группам). Сумма процентов в строке (столбце) таблицы может превышать 100%, а сумма частот превышать значение, указанное, в маргинальном столбце (маргинальной строке). Коэффициент Хи-квадрат вычисляется не по отношению к количеству анкет, а по отношению к количеству сделанных выборов в номинальном с совместимыми альтернативами признаке (количество выборов может быть существенно больше количества анкет, поскольку во многих анкетах может быть выбрана не одна, а несколько альтернатив).

**Если два признака номинальные с совместимыми альтернативами**

Признак #26. 21.Какая из проблем вас беспокоит больше всего  
Признак #21. 16.Что Вы будете делать, если лишитесь работы?  
Всего анкет 1653

	НеОтв	безопа сность Украи	возрож дение укр на	отноше ния с Россие	престу пность	послед ствия Черноб	пробле ма Кры ма	пробле ма Чер ном фл	уровне жизн и лиде	положе ние ру с язык	другая	ТС	ВСЕГО
НеОтв	0	4	0	7	23	4	0	0	19	1	0	0	24
Попыта юсь от крыть	29,0 54 14,7	11,8 22 14,4	33,3 62 13,6	58,1 108 12,4	18,8 35 11,5	7,0 13 16,9	4,3 8 15,3	64,5 120 10,3	4,3 8 9,3	4,3 8 21,6	0,5 1 3,1	0,5 1 11,4	100,0 186 11,4
Соглаш усь на любую	26,1 109 29,6	11,5 48 31,4	33,6 140 30,8	55,9 233 26,8	20,4 85 28,0	4,8 20 26,0	3,4 14 4,1	72,7 303 26,0	6,2 26 30,2	2,2 9 24,3	0,5 2 6,3	0,5 2 6,3	100,0 417 25,6
Потреб ую пом ощи у	21,9 14 3,8	15,6 10 6,5	25,0 16 3,5	56,3 36 4,1	21,9 14 4,6	7,8 5 6,5	3,1 2 4,1	64,1 41 35	12,5 8 9,3	4,7 3 8,1	0,0 0 0,1	0,0 0 0,1	100,0 64 3,9
Вложу свои с береже	23,5 4 1,1	5,9 1 0,7	29,4 7 1,1	52,9 19 1,0	35,3 20 2,0	11,8 2 2,6	0,0 0 0,0	64,7 11 9,2	11,8 2 2,3	0,0 0 0,0	5,9 1 3,1	5,9 1 3,1	100,0 17 1,0
Займус ь само обеспе	23,3 100 27,2	11,9 51 33,3	30,8 132 29,0	55,5 238 27,3	24,9 107 35,2	4,9 21 27,3	4,0 17 34,7	71,6 307 26,4	7,2 33 36,0	2,6 11 29,7	0,9 4 12,5	0,9 4 12,5	100,0 429 26,3
Ничего не см огу сд	19,9 29 7,9	3,4 5 3,3	29,5 43 9,5	49,3 72 8,3	12,3 18 5,9	4,1 6 7,8	0,7 1 2,0	85,6 125 10,7	4,8 7 8,1	2,7 4 10,8	1,4 2 6,3	1,4 2 6,3	100,0 147 9,0
Не зна ю, что смогу	19,7 66 17,9	6,0 20 13,1	22,1 74 16,3	51,3 172 19,7	14,9 50 16,4	4,9 13 16,9	3,0 10 20,4	75,2 252 21,6	2,7 9 10,5	1,2 4 10,8	1,8 6 18,8	1,8 6 18,8	100,0 336 20,6
Другое	23,1 46 12,5	15,1 30 19,6	27,6 55 12,1	52,3 104 11,9	16,6 33 10,9	3,5 7 9,1	2,5 5 10,2	67,3 134 11,5	3,0 6 7,0	3,5 7 18,9	3,0 6 16,1	3,0 6 18,8	100,0 199 12,2
ТС	18,8 21 5,7	10,7 12 7,8	17,9 20 4,4	57,1 64 7,3	17,9 20 6,6	5,4 6 7,8	3,6 4 8,2	60,7 68 5,8	5,4 6 7,0	1,8 2 5,4	1,8 12 37,5	1,8 12 37,5	100,0 112 6,9
ВСЕГО	22,6 368 100,0	9,4 153 100,0	28,0 455 100,0	53,5 871 100,0	18,7 304 100,0	4,7 77 100,0	3,0 49 100,0	71,6 1165 100,0	5,3 86 100,0	2,3 37 100,0	3,0 101 100,0	3,0 101 100,0	100,0 1629

- ✓ строится таблица частот и процентов (корреляционная таблица, таблица сопряженности)
- ✓ вычисляются Хи-квадрат Пирсона и построенные на его основе коэффициенты связи для номинальных признаков (коэффициенты Чупрова и Крамера); производится оценка значимости Хи-квадрат на уровне 5% и на уровне 1%

Коэффициент Хи-квадрат	159,6 (значим 1%)
Число степеней свободы	80
Хи-квадрат вычислен для	4293 ответов
Хи-квадрат не выведен	23 клеток с ожидаемой частотой < 5
Коэффициент Чупрова	0,1
Коэффициент Крамера	0,1

**Внимание.** Интерпретация таблиц и показателей требует определенной осторожности. При анализе результатов необходимо учитывать, что группы, которые выделяются отдельными значениями признака с совместимыми альтернативами, могут пересекаться (т.е. одна и та же анкета может принадлежать нескольким группам). Сумма процентов в строке (столбце) таблицы может превышать 100%, а сумма частот превышать значение,

указанное, в маргинальном столбце (маргинальной строке). Коэффициент Хи-квадрат вычисляется не по отношению к количеству анкет, а по отношению к количеству сделанных выборов в номинальном с совместимыми альтернативами признаке (количество выборов может быть существенно больше количества анкет, поскольку во многих анкетах может быть выбрана не одна а несколько альтернатив).

**Таблица частот и процентов** состоит из клеток, в каждой из которых информация размещена следующим образом :

n1%	- какой процент клеточная частота n2 составляет по отношению к сумме всех частот строки
n2	- клеточная частота
n3%	- какой процент клеточная частота n2 составляет по отношению к сумме всех частот столбца

**В таблице групповых средних** каждая строка содержит информацию о распределении усредняемого (зависимого) признака в одной группе анкет. Группу образуют анкеты с одинаковым значением группирующего (независимого) признака. Таблица групповых средних имеет такую структуру:

Знач #n	Част #n	Min #m	Max #m	Сред #m	СрКв #m	КфВр #m
---------	---------	--------	--------	---------	---------	---------

n	номер группирующего (независимого) признака
m	номер усредняемого (зависимого) признака
Знач #n	названия альтернатив группирующего признака
Част #n	количество анкет (частота), у которых признак n равен альтернативе, указанной в первом столбце
Min #m	минимальное значение признака # m для тех анкет, у которых признак n равен альтернативе, указанной в первом столбце
Max #m	максимальное значение признака # m для тех анкет, у которых признак n равен альтернативе, указанной в первом столбце
Сред #m	среднее арифметическое значение признака # m для тех анкет, у которых признак n равен альтернативе, указанной в первом столбце
СрКв #m	среднеквадратическое отклонение признака # m для тех анкет, у которых признак n равен альтернативе, указанной в первом столбце
КфВр #m	коэффициент вариации признака # m для тех анкет, у которых признак n равен альтернативе, указанной в первом столбце

## Параметры операции

✕
**Двумерные распределения**

Номера вертикальных признаков:  
1-5

Номера горизонтальных признаков:  
5-10

Выводить таблицы частот

Выводить таблицы средних

Выводить в таблице:

пустые строки

строку для "Не Ответ"

столбец для "Не Ответ"

% в строке

% в столбце

частоты

Точность вывода % (0-3)

считать % к ответам

вывод статистик

Точность вывода статистик (1-5)

Выводить

текст вопроса

название признака

Выводить

названия альтернатив

коды альтернатив

Вывод в файл

Вычислить

Отмена

- ✓ **Номера вертикальных признаков:** – поле для ввода списка номеров вертикальных признаков
- ✓ **Номера горизонтальных признаков:** – поле для ввода списка номеров горизонтальных признаков
- ✓ **Выводить таблицы частот** – если этот флажок отмечен, то выводится таблица частот (когда это возможно, исходя из типа шкал)
- ✓ **Выводить таблицы средних** - если этот флажок отмечен, то выводится таблица групповых средних (когда это возможно, исходя из типа шкал)

### Выводить в таблице:

- ✓ **пустые строки** - если этот флажок отмечен, то в таблице частот и в таблице частот и в таблице групповых средних будут выводиться строки (в таблице частот и столбцы) с нулевой маргинальной частотой (т.е. для тех категорий, которые вообще не представлены в таблице)
- ✓ **строку для "Не Ответ"** – если этот флажок отмечен, то в таблице частот будет выводиться отдельная строка для отсутствующих ответов; эта строка будет расположена непосредственно перед маргинальной строкой
- ✓ **столбец для "Не Ответ"** – если этот флажок отмечен, то в таблице частот будет выводиться отдельный столбец для отсутствующих ответов; этот столбец будет расположен непосредственно перед маргинальным столбцом
- ✓ **% в строке** - если этот флажок отмечен, то в каждой клетке таблицы частот будет выводиться процент клеточной частоты по отношению к сумме частот в соответствующем столбце
- ✓ **% в строке** - если этот флажок отмечен, то в каждой клетке таблицы частот будет выводиться процент клеточной частоты по отношению к сумме частот в соответствующей строке
- ✓ **частоты** - если этот флажок отмечен, то в каждой клетке таблицы частот будет выводиться соответствующая клеточная частота

**Точность вывода % (0-3)** – указывается количество (от 1 до 3) знаков после десятичной точки при выводе процентов

**считать % к ответам** - если этот флажок отмечен, то вычисление процентов (как в строке, так и в столбце) будет производиться по отношению к количеству ответов (т.е. без учета количества НеОтветов, это стандартный подход для большинства статистических пакетов программ); в противном случае НеОтвет рассматривается как отдельная категория и вычисление процентов производится с учетом этой категории. Примечание. В любом случае при вычислении коэффициентов связи НеОтветы не учитываются (т.е. НеОтветы рассматриваются как отдельная категория и объекты, у которых есть НеОтвет по крайней

мере по одному из признаков, для которых строится таблица, просто исключаются из вычислений)

**Вывод статистик** - если этот флажок отмечен, то вычисляются и выводятся показатели связи, уравнения парной регрессии, таблицы значимости различий средних

Выводить

текст вопроса

название признака

Выводить

названия альтернатив

коды альтернатив

в зависимости от выбора, для построения заголовка таблицы используется из паспорта массива либо текст вопроса либо название признака

в зависимости от выбора, при выводе таблицы в качестве меток строк и столбцов используются из паспорта массива либо названия альтернатив либо их коды

**Вывод в файл**

— эта кнопка позволяет вывести построенные таблицы в файл на диске в формате HTML. Можно задать либо имя нового файла (и тогда этот файл будет автоматически создан) либо имя уже существующего файла в формате HTML ( в таком случае построенные таблицы будут дописаны в конец этого файла). Это позволяет "накапливать" построенные таблицы в файле. Вывод в файл производится в дополнение к выводу в стандартный листинг результата.

## Статистика | Статистические гипотезы | Знач. различий процентов

Статистика

Одномерные распределения

Двумерные распределения

Статистические гипотезы

Объем выборки

Знач. различий процентов

Знач. различий коэф. корреляции

Знач. различий средних

Эта операция позволяет проверить статистическую гипотезу о равенстве доли (процента) некоторого свойства (некоторого признака) в двух независимых выборках.

Проверка гипотезы выполняется на двух стандартных для социальных наук уровнях значимости – 5% и 1%.

**Знач. различий двух процентов**

Введите два процента (P1 и P2) и объемы выборок (N1 и N2) по отношению к которым эти проценты

P1 [%] 37 P2 [%] 49

N1 300 N2 350

Вычислить

Отмена

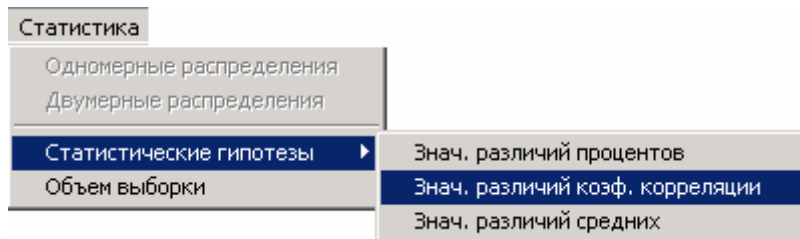
**Различие значимо на уровне 1%**

Для проверки гипотезы необходимо указать два значения процентов (P1 и P2) и объемы соответствующих выборок, для которых эти два процента были вычислены (N1 и N2).

Если гипотеза может быть отвергнута на одном из этих двух уровней значимости, то выводится сообщение "Различие значимо на уровне ...%" (в сообщении указывается более высокий уровень значимости).

Если гипотеза не может быть отвергнута ни на одном из этих двух уровней значимости, то выводится сообщение "Различие не значимо на уровне 5%".

**Различие не значимо на уровне 5%**

**Статистика | Статистические гипотезы | Знач. различий коэф.корреляции**

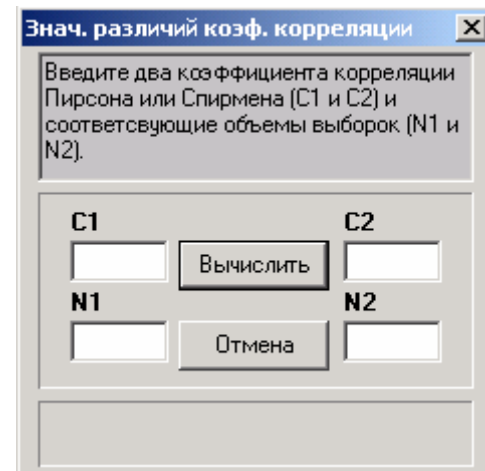
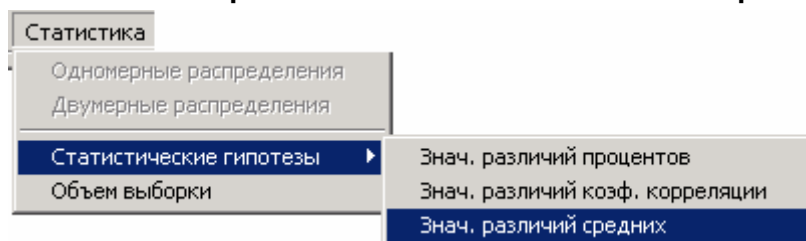
Эта операция позволяет проверить статистическую гипотезу о равенстве двух коэффициентов корреляции Пирсона или двух ранговых коэффициентов корреляции Спирмена.

Проверка гипотезы осуществляется на двух стандартных для социальных наук уровнях значимости – 5% и 1%.

Если гипотеза может быть отвергнута на одном из этих двух уровней значимости, то выводится сообщение "Различие значимо на уровне ...%" (в сообщении указывается более высокий уровень значимости).

Если гипотеза не может быть отвергнута ни на одном из этих двух уровней значимости, то выводится сообщение "Различие не значимо на уровне 5%".

Для проверки гипотезы необходимо указать два значения коэффициентов корреляции (C1 и C2) и объемы соответствующих выборок, для которых эти два коэффициента корреляции были вычислены (N1 и N2).

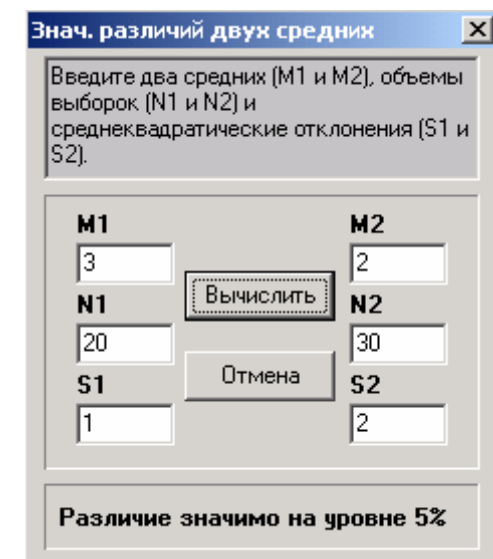
**Статистика | Статистические гипотезы | Знач. различий средних**

Эта операция позволяет проверить статистическую гипотезу о равенстве двух средних. Проверка гипотезы осуществляется на двух стандартных для социальных наук уровнях значимости – 5% и 1%.

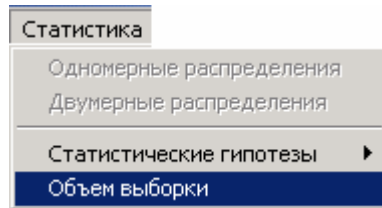
Если гипотеза может быть отвергнута на одном из этих двух уровней значимости, то выводится сообщение "Различие значимо на уровне ...%" (в сообщении указывается более высокий уровень значимости).

Если гипотеза не может быть отвергнута ни на одном из этих двух уровней значимости, то выводится сообщение "Различие не значимо на уровне 5%".

Для проверки гипотезы необходимо указать два значения средних (M1 и M2), соответствующие значения среднеквадратических отклонений (S1 и S2) и объемы соответствующих выборок, для которых эти два коэффициента корреляции были вычислены (N1 и N2).



## Статистика | Объем выборки



Эта операция позволяет оценить объем необходимой для исследования выборки. Предполагается, что исследователь будет работать со случайной одноступенчатой бесповторной выборкой. Это означает, что все элементы генеральной совокупности имеют одинаковые шансы быть отобранными (имеют равные вероятности попасть в выборку) и что в выборке один и тот же элемент никогда не встречается дважды. Эта операция позволяет определить объем выборки, необходимый для того, чтобы оценить с заданной абсолютной допустимой погрешностью долю некоторого изучаемого признака в генеральной совокупности.

Для получения оценки объема выборки необходимо указать следующие параметры:

**Объем генеральной совокупности:** – положительное целое число, которое больше 60. Если объем генеральной совокупности достаточно велик (например, составляет десятки тысяч объектов), то достаточно указать приблизительную оценку. Например, если речь идет об изучении электоральных предпочтений населения Украины (генеральная совокупность – все жители Украины, имеющие право голоса), то такие

оценки объема генеральной совокупности как 20000000 или 30000000 будут давать один и тот же результат для оценки размера выборки (400 объектов).

**Доля признака в генеральной совокупности:** – число, которое больше нуля и меньше 1.

Приблизительная оценка доли изучаемого признака может быть известна из предыдущих исследований, из предварительного "пилотажного" исследования. При равных значениях прочих параметров чем значение этой доли ближе к 0.5, тем объем выборки больше. Если доля изучаемого неизвестна, то лучше указать 0.5 (при этом объем выборки может быть несколько завышенным).

**Допустимая погрешность %:** , с которой доля изучаемого признака в генеральной совокупности должна быть представлена в выборке – указывается в процентах (число, которое больше 0 и меньше 100). При равных значениях прочих параметров чем значение этого параметра ближе к 0, тем объем выборки больше.

**Доверительная вероятность** , с которой делается вывод о погрешности выборки – число, которое больше нуля и меньше 1. При равных значениях прочих параметров чем значение этого параметра ближе к 1, тем объем выборки больше. В социологических исследованиях, как правило, используют два стандартных уровня доверительной вероятности – 0.99 и 0.95.

Предположим, что исследуется трудоспособное население города Киева (приблизительный объем генеральной совокупности 100000). Исследователь хочет получить с вероятностью 0.95 (доверительная вероятность) данные о доле удовлетворенных уровнем своей заработной платы с ошибкой, не превышающей 3% (допустимая погрешность). При этом нет информации о приблизительной оценке этой доли в генеральной совокупности (в качестве оценки указываем 0.5). Тогда объем необходимой для исследования выборки составит приблизительно 1110 человек. Если же исследователя удовлетворит точность 5%, то объем необходимой выборки составит приблизительно 400 человек.

Процесс планирования объема выборки усложняется в случае, если необходимо обеспечить точность представления нескольких выборок, если планируется разбивать

выборку на части во время анализа. Кроме того, часто необходимо обеспечить представление не доли одного или даже нескольких признака, а среднего значения признака, суммы признаков или других параметров.

## Вычисления | Условное вычисление

Вычисления

Условное вычисление

Перекодировка

Стандартизация

Эта операция позволяет построить новый признак.

### Параметры операции

Условное вычисление нового признака

Тип шкалы нов. признака

метрическая

номинальная

порядковая

номинальная с  совместимыми альтернативами

Название признака

Возрастные группы

код	название_альтернативы
1	от 16 до 25
2	от 26 до 35
3	от 36 до 45
4	от 46 до 55
5	от 56 и старше

условие -> значение

#139#>=16 and #139#<=25 -> 1

#139#>=26 and #139#<=35 -> 2

#139#>=36 and #139#<=45 -> 3

#139#>=46 and #139#<=55 -> 4

#139#>=56 -> 5

иначе

\$

OK Отмена

**Тип шкалы нов. признака** — указывается тип шкалы создаваемого признака. Для номинальной шкалы с совместимыми альтернативами необходимо также указать количество альтернатив

**Название признака** — в это поле вводится текст (в свободном формате), который будет использован в качестве названия нового признака при формировании соответствующего пункта в паспорте. Поле "текст вопроса" для формируемого признака заполняется автоматически. В этом поле указывается, что соответствующий признак создан автоматически (т.е. является "дополнительным" или "вычисленным" признаком) и дата создания признака.

**код название\_альтернативы** — в этом окне для номинальных и порядковых признаков необходимо указать допустимые значения (коды альтернатив) и их метки (названия альтернатив). Название может состоять из нескольких слов. Между кодом альтернативы и первой буквой названия должен быть по крайней мере один пробел. Для номинальной шкалы с совместимыми альтернативами необходимо в качестве кодов использовать только последовательные целые числа начиная с 1

**условие -> значение** — в этом окне задается последовательность правил вычисления нового признака. Каждое правило записывается в одной отдельной строке. В записи "->" между символом "-" и символом ">" не должно быть пробелов. Слева от знака "->" логическое выражение (условие), а справа — арифметическое выражение (результатом вычисления которого есть значение)

**иначе** — в этом поле записывается арифметическое выражение (значение), которое присваивается признаку в случае, когда ни одно из условий не выполняется. При построении признака с совместимыми альтернативами это поле не используется.

## Вычисления | Перекодировка

Вычисления

Условное вычисление

**Перекодировка**

Стандартизация

Эта операция позволяет создать новый признак путем перекодирования значений одного существующего признака. Результат работы операции – новый признак в номинальной (но не с совместимыми альтернативами), порядковой или метрической шкале. С помощью этой операции удобно "разбивать" диапазон значений метрического признака на интервалы, группировать значения, изменять кодировку значений.

### Параметры операции

**Перекодировать признак** – указывается номер признака, который будет перекодироваться. Это должен быть обязательно признак не с совместимыми альтернативами

**Тип шкалы нов. признака** – указывается тип шкалы создаваемого признака. Результатом работы операции не может быть признак с совместимыми альтернативами

**Название нового признака:** – в это поле вводится текст (в свободном формате), который будет использован в качестве названия нового признака при формировании соответствующего пункта в паспорте. Поле "текст вопроса" для формируемого признака заполняется автоматически. В этом поле указывается, что соответствующий признак создан автоматически (т.е. является "дополнительным" или "вычисленным" признаком) и дата создания признака.

**код название\_альтернативы:** – в этом окне для номинальных и порядковых признаков необходимо указать допустимые значения (коды альтернатив) и их метки (названия альтернатив). Название может состоять из нескольких слов. Между кодом альтернативы и первой буквой названия должен быть по крайней мере один пробел

**старое значение(интервал) -> новое значение:** – в этом окне задается последовательность правил вычисления нового признака. Каждое правило записывается в одной отдельной строке. В записи "->" между символом "-" и символом ">" не должно быть пробелов. Слева от знака "->" записывается либо одно значение, либо символ \$ (который обозначает значение НеОтвет) либо один интервал значений. Интервал значений может быть либо закрытый, либо открытый слева, либо открытый справа. Закрытый интервал записывается в виде левая\_граница..правая\_граница (например, 2..6 – "от 2 до 6 включительно", или 3.5..4.7 – "от 3.5 до 4.7 включительно"). Открытый слева интервал записывается в виде правая\_граница (например, ..34 – "не больше, чем 34"). Открытый справа интервал записывается в виде левая\_граница.. (например, 34.. – "не меньше, чем 34"). Справа от знака "->" записывается арифметическое выражение, результатом вычисления которого есть значение



**Остальные значения:** – в этом поле записывается арифметическое выражение (значение), которое присваивается признаку в случае, когда исходный признак не равен ни одному из указанных "старых" значений и не принадлежит ни одному из указанных интервалов "старых" значений.

## Вычисления | Стандартизация

Вычисления

Условное вычисление

Перекодировка

Стандартизация

Эта операция позволяет построить новый признак путем стандартизации существующего метрического или порядкового признака. В этом случае порядковый признак рассматривается как метрический.

### Параметры операции

Стандартизация заключается в линейном преобразовании значений признака с целью получения признака с заданными значениями среднего и стандартного отклонения. Часто в качестве параметров стандартизации используются значения (0,1) (т.е. среднее 0 и стандартное отклонение 1). Результат выполнения операции – новый метрический признак.

**Стандартизировать признак [  $Z(m,n)$  ]** – указывается номер существующего метрического или порядкового признака, на основе которого будет строиться новый признак

**Параметры стандартизации: ... среднее (m)** – указывается среднее значение нового (того, который будет построен) признака

**Параметры стандартизации: ... среднеквадратическое отклонение (n)** – указывается среднеквадратическое (стандартное) отклонение нового (того, который будет построен) признака

**Название нового признака:** – указывается название нового (того, который будет построен) признака, одна строка произвольного текста.

## Примечания

---

<sup>i</sup> *Меры центральной тенденции* – обобщенные характеристики распределения признака, отражающие определенную тенденцию. Цель вычисления мер центральной тенденции, как правило, состоит в том, чтобы получить показатель (характеристику), который можно использовать "вместо" ряда значений, составляющих распределение. Меры центральной тенденции: мода, медиана, среднее арифметическое. *Среднее арифметическое* – сумма значений признака деленная на количество объектов (НеОтветы в вычислении среднего арифметического участия не принимают); мера центральной тенденции для метрических шкал. *Медиана* – центральный (т.е. находящийся посередине) элемент в упорядоченном ряду значений признака; элемент, который делит упорядоченный ряд значений признака пополам; мера центральной тенденции для порядковых и метрических шкал. Главное свойство медианы заключается в том, что у одной половины объектов значение признака меньше медианы, а у другой – больше. *Мода* – значение признака, которое имеет наибольшую частоту (т.е. которое наиболее часто встречается в ряду значений); мера центральной тенденции для номинальных, порядковых и метрических шкал.

<sup>ii</sup> *Меры вариации* дают представление о том, насколько значения признака колеблются относительно среднего (или других мер центральной тенденции). Меры вариации: вариационный размах, среднеквадратическое отклонение, дисперсия, коэффициент вариации. *Вариационный размах* – разница между максимальным и минимальным значениями признака. *Дисперсия* – среднее арифметическое квадратов отклонений вариантов от среднего для данной совокупности. *Среднеквадратическое отклонение* – корень квадратный дисперсии, а коэффициент вариации, - отношение среднеквадратического отклонения к среднему.