

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РЕСПУБЛИКИ КАЗАХСТАН**

Рудненский индустриальный институт

Кафедра экономики и менеджмента

**МЕТОДИЧЕСКИЕ УКАЗАНИЯ
по использованию программного продукта «SPSSStatistic»**

**по дисциплине «Статистика»
для студентов экономических специальностей**

Рудный 2017

БК 22.12

Автор: Абжатова А.К. Методические указания по использованию программного продукта «SPSSStatistic». – Рудный, РИИ, 2017. – 17 с.

Методические указания составлены в соответствии с требованиями учебного плана и программы дисциплины «Статистика» и включают в себя сведения, необходимые для выполнения лабораторных работ.

Методические указания предназначены для студентов экономических специальностей.

Ил. 5, Список лит. 5 назв.

Для внутривузовского использования

СОДЕРЖАНИЕ

Введение.....	4
Лабораторная работа № 1. Характеристика SPSS. Подготовка данных к анализу. Редактирование данных.....	5
Лабораторная работа № 2. Описательные статистики.....	12
Список литературы	16

ВВЕДЕНИЕ

IBMSPSSStatistics (аббревиатураангл. «StatisticalPackageforSocialSciences» - «статистический пакет для социальных наук») – компьютерная программа для статистической обработки данных, один из лидеров рынка в области коммерческих статистических продуктов, предназначенных для проведения прикладных исследований в социальных науках.

Предоставляет ряд методов, включая ситуативный анализ, проверку гипотез и составление отчетов, которые упрощают управление данными, выбор и применение методов анализа, а также обмен результатами.

Курс ориентирован на овладение приемами анализа статистических данных с помощью пакета SPSS для решения задач прикладного экономического анализа.

Отчет по работе в программе оформляется в приложении MSWord, при оформлении учитываются требования к документам данного вида. Содержание отчета:

- титульный лист;
- содержание;
- цель и задачи работы;
- характеристика основных этапов выполнения задания;
- ответы на контрольные вопросы;
- заключение.

Защита работы включает отчет студента о проделанной работе, оформленный в соответствии с вышеизложенными требованиями.

Лабораторная работа № 1. Характеристика SPSS. Подготовка данных к анализу. Редактирование данных. Описательные статистики.

Цель: познакомиться с основными понятиями и областью применения системы SPSS

Задание:

1. Информация о данных

Процедура Информация о данных выводит информацию из словаря данных, такую как имена переменных, метки переменных, метки значений, пропущенные значения, а также итожащие статистики для всех заданных переменных и наборов множественных ответов в активном наборе данных. Для номинальных и порядковых переменных, а также наборов множественных ответов итожащие статистики включают частоты и проценты. Для количественных переменных итожащие статистики включают среднее значение, стандартное отклонение и квантили.

Примечание: Процедура Информация о данных игнорирует состояние расщепления файла. Это включает группы расщепленных файлов, созданные для множественной импутации пропущенных значений (имеется в дополнительном модуле Missing Values).

Доступ к процедуре Информация о данных:

- Выберите в меню: Анализ > Отчеты > Информация о данных
- Откройте вкладку Переменные
-

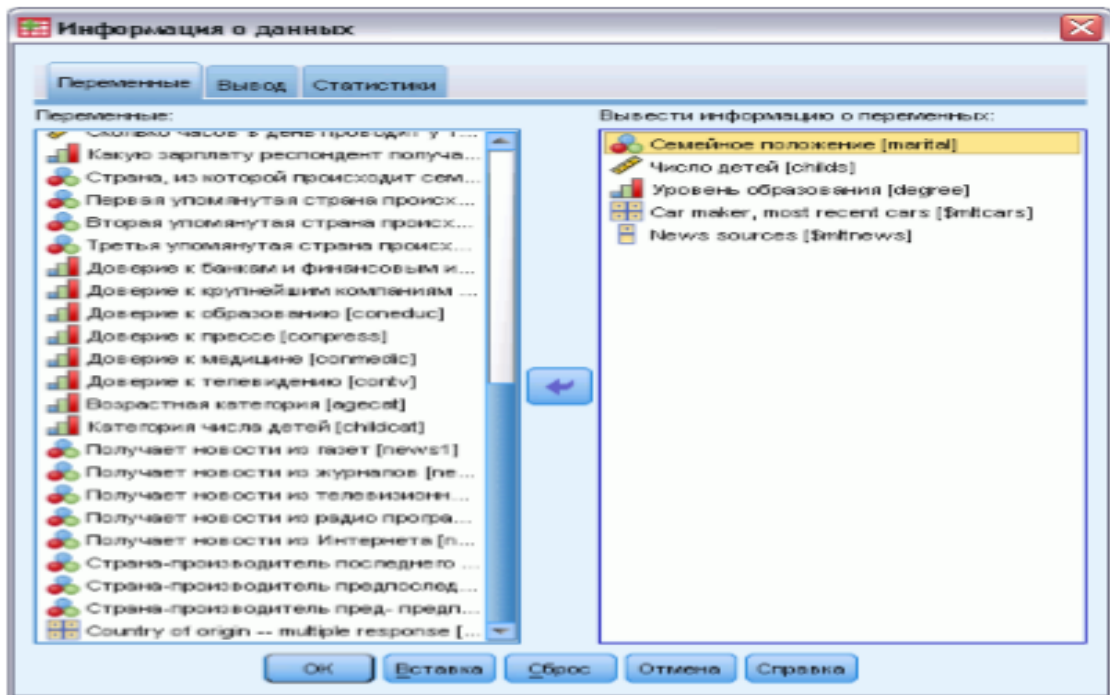


Рисунок 1 - Диалоговое окно Информация о данных, вкладка Переменные

Порядок выполнения лабораторной работы

1 Изучить материал.

Выберите одну или несколько переменных и/или наборов множественных ответов.

Дополнительно можно:

- Управлять отображаемой информацией о переменных.
- Управлять выводом статистик (или исключить все итоговые статистики).
- Управлять порядком вывода переменных и наборов множественных ответов.

➤ Изменять шкалу измерений для любой переменной в списке исходных переменных, чтобы изменить выводимые итоговые статистики.

Изменение шкалы измерений

Можно временно изменить шкалу измерений для переменных (шкалу измерений нельзя изменить для наборов множественных ответов. Они всегда считаются номинальными.)

Щелкните правой кнопкой мыши по переменной в исходном списке.

В появившемся контекстном меню выберите шкалу измерений.

После этого шкала измерений будет временно изменена.

С практической точки зрения это полезно только для числовых переменных.

Шкала измерений для текстовых переменных может быть только номинальной или порядковой, причем в процедуре Информация о данных обе эти шкалы обрабатываются идентично.

2. Вкладка Информация о данных: Вывод

Вкладка «Вывод» управляет информацией о переменных, включаемой в вывод для всех переменных и наборов множественных ответов, порядком вывода переменных и наборов множественных ответов, а также содержимым дополнительной таблицы информации о файле

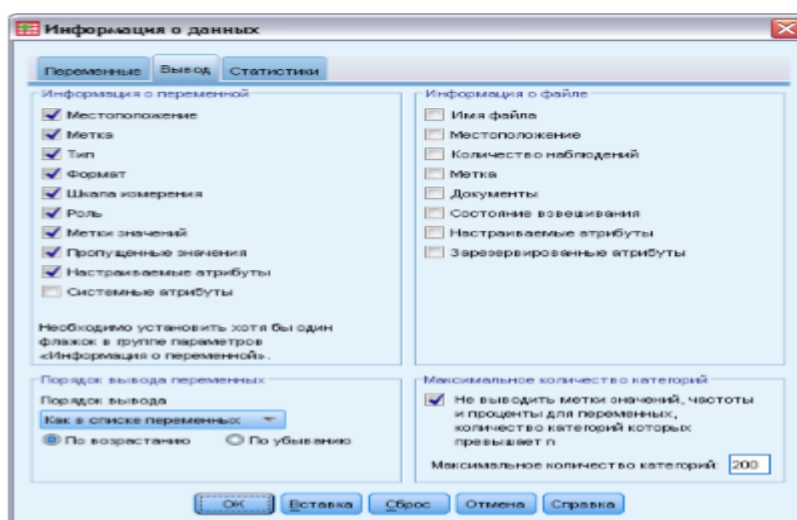


Рисунок 2 - Диалоговое окно Информация о данных, вкладка Вывод
Порядок выполнения лабораторной работы

1. Изучить материал.

Информация о переменной

Здесь задается информация из словаря данных, выводимая для всех переменных.

Местоположение. Целое число, представляющее положение переменной в порядке их расположения в файле. Этот параметр недоступен для наборов множественных ответов.

Метка. Описательная метка переменной или набора множественных ответов.

Тип. Основной тип данных. Тип может быть Числовой, Текстовый или Набор множественных ответов.

Формат. Формат вывода переменной, например A4, F8.2 или DATE11. Этот параметр недоступен для наборов множественных ответов.

Шкала измерений. Возможные значения: Номинальная, Порядковая, Количественная и Неизвестная. Выводимым значением является шкала измерений, хранящаяся в словаре данных, и на нее не влияет никакое временное изменение шкалы измерений, сделанное в списке исходных переменных на

вкладке Переменные. Этот параметр недоступен для наборов множественных ответов.

Примечание: Шкала измерений для числовых переменных может быть «неизвестной» до первого прохода данных, если она не была задана явно, как, например, для данных, считанных из внешнего источника, или вновь создаваемых переменных.

Роль. Некоторые диалоговые окна поддерживают возможность предварительного выбора переменных для анализа, основанного на заданных ролях.

Метки значений. Описательные метки, связанные с определенными значениями данных.

- Если на вкладке «Статистики» выбрана «Частота» или «Проценты», то заданные метки значений включаются в вывод, даже если они не были здесь выбраны для вывода.

- Для наборов множественных дихотомий «метками значений» являются метки переменных для элементарных переменных в наборе или метки подсчитываемых значений в зависимости от того, как определен набор.

Пропущенные значения. Пользовательские пропущенные значения. Если на вкладке «Статистики» выбрана «Частота» или «Процент», то заданные метки значений включаются в вывод, даже если пропущенные значения не были здесь выбраны для вывода. Этот параметр недоступен для наборов множественных ответов.

Настраиваемые атрибуты. Задаваемые пользователем атрибуты переменных. В вывод включаются и имена, и значения задаваемых пользователем атрибутов всех переменных. Этот параметр недоступен для наборов множественных ответов.

Зарезервированные атрибуты. Зарезервированные атрибуты системных переменных. Можно вывести системные атрибуты, но изменять их не следует. Имена системных атрибутов начинаются со знака доллара (\$). Скрытые атрибуты с названиями, начинающимися с «@» или «\$@», не включаются в вывод. В вывод включаются и имена, и значения системных атрибутов, связанных со всеми переменными. Этот параметр недоступен для наборов множественных ответов.

Информация о файле

Дополнительная таблица информации о файле может содержать любой из перечисленных ниже атрибутов файла:

Имя файла. Имя файла данных IBM® SPSS® Statistics. Если набор данных никогда не был сохранен в формате SPSS Statistics, то имя файла данных отсутствует (Если в заголовке окна редактора данных нет имени файла, значит у активного набора данных нет имени файла).

Местоположение.

Каталог (папка), где расположен файл данных SPSS Statistics. Если набор данных ни когда не был сохранен в формате SPSS Statistics, то местоположения у него нет.

Количество наблюдений. Число наблюдений в активном наборе данных. Это общее число наблюдений, включая любые наблюдения, которые могли быть исключены при выводе итоговых статистик из-за условий фильтрации.

Метка. Это метка файла (если она есть), заданная командой FILE LABEL.

Документы. Текст документа файла данных.

Состояние взвешивания. Если взвешивание включено, отображается имя переменной взвешивания.

Настраиваемые

атрибуты.

Задаваемые пользователем атрибуты файла данных. Атрибуты файла данных, заданные командой DATA FILE ATTRIBUTE.

Зарезервированные атрибуты. Зарезервированные системные атрибуты файла данных. Можно вывести системные атрибуты, но изменять их не следует. Имена системных атрибутов начинаются со знака доллара (\$). Скрытые атрибуты с названиями, начинающимися с «@» или «\$@», не включаются в вывод. В вывод включаются имена, и значения всех системных атрибутов файла данных.

Порядок вывода переменных

Имеются следующие альтернативы управления порядком, в котором выводятся переменные и наборы множественных ответов.

По алфавиту. Алфавитный порядок по именам переменных.

По порядку в файле данных. Порядок отображения переменных в наборе данных (порядок, в котором они отображаются в редакторе данных). При сортировке в порядке возрастания набор множественных ответов выводятся последними, после всех выбранных переменных.

По шкала измерений. Сортировка по шкале измерений. При этом создаются четыре группы сортировки: номинальная, порядковая, количественная и неизвестная. Наборы множественных ответов рассматриваются как номинальные.

Примечание: Шкала измерений для числовых переменных может быть «неизвестной» до первого прохода данных, если она не была задана явно, как, например, для данных, считанных из внешнего источника, или вновь создаваемых переменных.

Как в списке переменных. Порядок, в котором переменные и наборы множественных ответов отображаются в списке выбранных переменных на вкладке Переменные.

Имена атрибутов, задаваемые пользователем.

В список параметров сортировки также

входящие на любые определенные пользователи атрибутов переменных. При сортировке в порядке возрастания переменные без атрибутов отображаются вверху, за ними следуют переменные с атрибутами, но без заданных значений атрибутов, и последними идут переменные с заданными значениями атрибутов в алфавитном порядке значений.

Максимальное количество категорий

Если в вывод включаются метки значений, частоты или проценты для всех уникальных значений, то эта информация не будет выводиться в таблице, если число значений превышает указанное значение. По умолчанию эта информация не выводится, если число уникальных значений для переменной больше 200.

3. Вкладка Информация о данных: Статистики

На вкладке «Статистики» можно управлять выводом итоговых статистик и при желании не выводить их совсем.

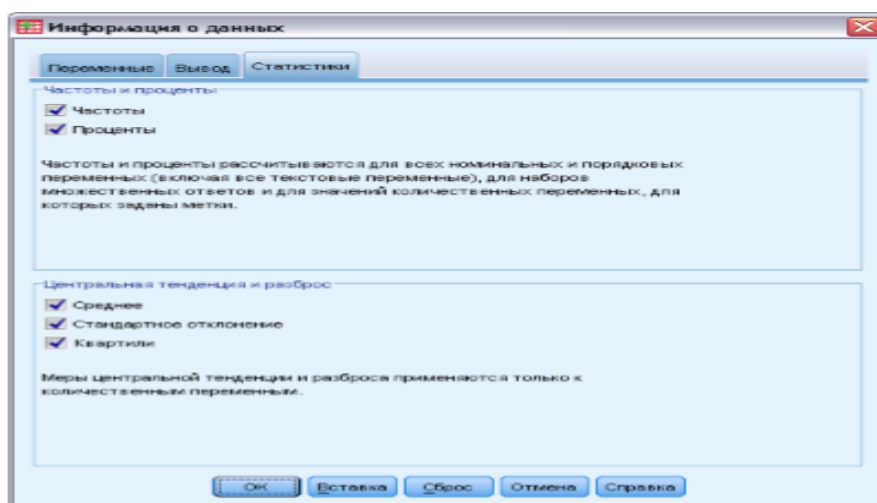


Рисунок 3 - Диалоговое окно Информация о данных, вкладка Статистики

Порядок выполнения лабораторной работы

1. Изучить материал.

Частоты и проценты

Для номинальных и порядковых переменных, наборов множественных ответов, а также значений количественных переменных сметками доступны следующие статистики:

Частоты. Количество наблюдений (объектов), имеющих каждое значение (или диапазон значений) переменной.

Процент. Процент наблюдений, имеющих конкретное значение.

Центральная тенденция и разброс

Для количественных переменных доступны следующие статистики:

Среднее. Мера центральной тенденции. Арифметическое среднее; сумма, деленная на число наблюдений.

Стандартное отклонение. Мера разброса вокруг среднего. При нормальном распределении 68% наблюдений укладываются в одно стандартное отклонение от среднего, и 95% в два стандартных отклонения. Если, например, средний возраст равен 45 годам со стандартным отклонением 10, то 95% наблюдений должны оказаться между 25 и 65 годами при нормальном распределении.

Квартили. Значения 25-го, 50-го и 75-го процентилей.

Примечание: можно временно изменить шкалу измерений переменной (и, следовательно, изменить итоговые статистики, выводимые для этой переменной) в списке исходных переменных на вкладке Переменные.

- 2 Выполнить задание к лабораторной работе.
- 3 Ответить на контрольные вопросы.
- 4 Подготовить отчет по лабораторной работе.

Контрольные вопросы

- 1 Дайте определение понятию «данные».
- 2 Дать определение понятию «переменные».
- 3 Что такое «средние»?
- 4 Для чего предназначена SPSS Statistics?
- 5 Каковы основные функции SPSS Statistics?
- 6 Охарактеризуйте основные составляющие Информации о данных

Лабораторная работа № 2. Описательные статистики.

Цель: дать целостное представление о возможностях анализа данных в системе SPSS

Задание:

1. *Описательные статистики*

Процедура *Описательные статистики* осуществляет вывод одномерных итоговых статистик для нескольких переменных в одной таблице, а также вычисляет стандартизованные значения (*z*-значения) переменных. Переменные могут быть упорядочены по величине их средних значений (в порядке возрастания или убывания), по алфавиту или в порядке, в котором выбираются переменные (по умолчанию).

Присохрании *z*-значений они добавляются к данным в Редакторе данных и могут быть впоследствии использованы для построения графиков, вывода их значений и в других процедурах IBM SPSS Statistics. Если переменные измерены в разных единицах (например, валовой внутренний продукт на душу населения и процент грамотных), преобразование к *z*-значениям приводит переменные к единому масштабу, что облегчает их визуальное сравнение.

Пример. Если каждое наблюдение в анализируемых данных содержит итоги дневных объемов продаж для одного из членов коллектива продавцов (например, одно значение для Алексея, одно – для Марии, одно – для Бориса) в течение нескольких месяцев, то процедура *Описательные статистики* может рассчитать средний дневной объем продаж для каждого продавца и расположить результаты в порядке от наиболее высоких средних ежедневных продаж к наиболее низким.

Статистики. Объем выборки, среднее значение, минимальное и максимальное значения, стандартное отклонение, дисперсия, размах, сумма, стандартная ошибка среднего, асимметрия, эксцесс, стандартные ошибки асимметрии и эксцесса.

Данные. Используются числовые переменные после того, как исследовали их диаграммы на наличие ошибок записи, выбросов и аномалий в распределениях. Процедура *Описательные статистики* очень эффективно работает с файлами большого размера (содержащими тысячи наблюдений).

Предположения. Большинство статистик, которые могут быть вычислены при работе с данной процедурой (в том числе *z*-значения), основаны на теории нормального распределения и подходят для количественных переменных (измеренных в интервальной шкале или шкале

отношений), распределенных симметрично. Нужно избегать переменных с неупорядоченными категориями или несимметричными распределениями.

Распределение z -значений имеет ту же форму, что и распределение исходных данных; поэтому переход к z -значениям не является средством исправления «недостатков» данных.

Выберите в меню: Анализ > Описательные статистики > Описательные...

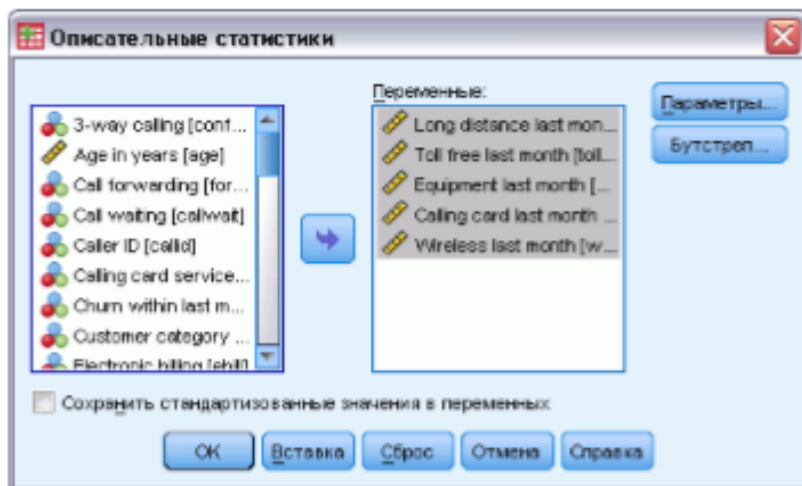


Рисунок 4 - Диалоговое окно Описательные статистики

Порядок выполнения лабораторной работы

1 Изучить материал.

Выберите одну или несколько переменных.

Дополнительно можно:

- Выбрать параметр Сохранить стандартизованные значения в переменных, чтобы сохранить z-значения как новые переменные.

- Щелкнуть мышью по кнопке

Параметры, чтобы выбрать дополнительные статистики и изменить порядок вывода результатов.

Параметры процедуры описательные статистики

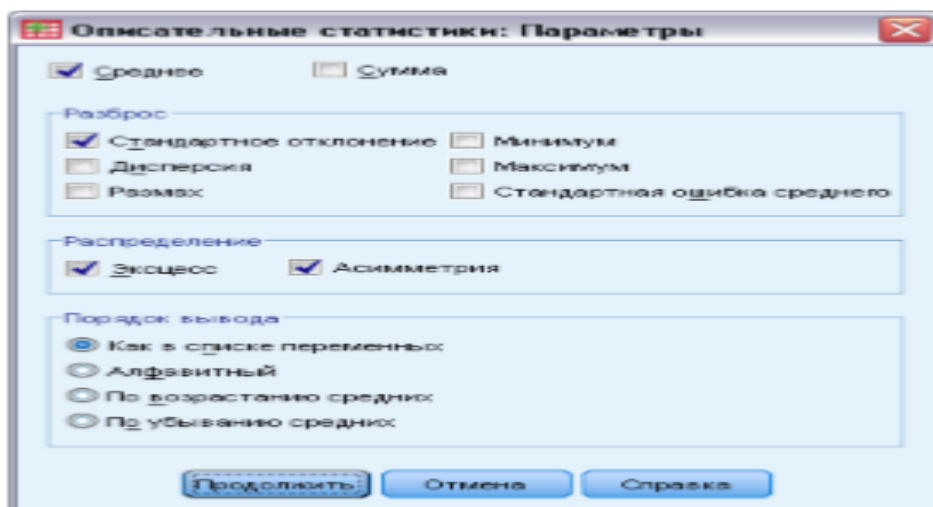


Рисунок 5 - Диалоговое окно Описательные статистики, Вкладка Параметры

Среднее и сумма. Среднее значение или арифметическое среднее значение выводятся по умолчанию. Разброс. Статистики, которые измеряют разброс данных, включают в себя стандартное отклонение, дисперсию, размах, минимальное и максимальное значения, а также стандартную ошибку среднего значения.

Стандартное отклонение. Мера разброса вокруг среднего. При нормальном распределении 68% наблюдений укладываются в одно стандартное отклонение от среднего, и 95% - в два стандартных отклонения. Если, например, средний возраст равен 45 годам со стандартным отклонением 10, то 95% наблюдений должны оказаться между 25 и 65 годами при нормальном распределении.

Дисперсия. Мера разброса относительно среднего значения. Равна сумме квадратов отклонений от среднего, деленной на число, на единицу меньшее числа наблюдений. Дисперсия измеряется в единицах, которые равны квадратам единицы измерения самой переменной.

Диапазон. Разность между наибольшим и наименьшим значениями числовой переменной; максимум минус минимум.

Минимум. Наименьшее значение числовой переменной.

Максимум. Наибольшее значение числовой переменной.

Стандартная ошибка среднего. Мера того, как сильно может отличаться значение среднего от выборки к выборке, извлекаемое из одного и того же распределения. Можно применять для грубого сравнения наблюдаемого среднего с гипотетическим значением (то есть можно заключить, что два значения различаются, если отношение их разности к стандартному отклонению меньше -2 или больше +2).

Распределение. Экссесс и асимметрия представляют собой статистики,

описывающие форму и степень симметричности распределения. Эти статистики выводятся вместе с их стандартными ошибками.

Эксцесс. Мера сгруппированности наблюдений вокруг центральной точки. Для нормального распределения значение эксцесса равно 0. Положительный эксцесс указывает на то, что по отношению к нормальному распределению наблюдения для таких распределений сгруппированы более плотно около центра и имеют более тонкие хвосты до экстремумов распределения, и более толстые хвосты в области экстремальных значений. Отрицательный эксцесс указывает на то, что по отношению к нормальному распределению наблюдения для таких распределений сгруппированы менее плотно около центра и имеют более толстые хвосты до экстремумов распределения, и более тонкие хвосты в области экстремальных значений.

Асимметрия. Мера асимметрии распределения. Нормальное распределение симметрично, и для него асимметрия равна 0. Распределение со значимой положительной асимметрией имеет длинный хвост справа. Распределение со значимой отрицательной асимметрией имеет длинный хвост слева.

В качестве грубого правила можно сказать, что значение асимметрии, более чем вдвое превышающее ее стандартную ошибку, указывает на наличие асимметрии распределения. Порядок вывода. По умолчанию переменные выводятся в том порядке, в котором они выбирались пользователем. Можно выводить переменные в алфавитном порядке, в порядке возрастания средних значений или в порядке убывания средних значений.

Выполнить задание к лабораторной работе.

- 1 Ответить на контрольные вопросы.
- 2 Подготовить отчет по лабораторной работе.

Контрольные вопросы

1. Дайте определение понятию «средние»
2. Дайте определение понятию «стандартное отклонение»
3. Дайте определение понятию «дисперсия»
4. Охарактеризуйте процедуру Описательные статистики
5. Для чего предназначен расчет средних?

СПИСОК ЛИТЕРАТУРЫ

Базовый учебник

1. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей: Пер. с нем./ АхимБююль, Петер Цефель – СПб.: ООО «ДиаСофтЮП», 2002.

Основная литература

1. Крыштановский А.О. Анализ социологических данных с помощью пакета SPSS - М.: Изд.дом ГУ ВШЭ, 2006.

2. Дубнов П.Ю. Обработка статистической информации с помощью SPSS. –М.: NT Press, 2004.

3. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере 3-е изд., перераб. и доп. - М.: "Инфра-М", 2003.



Co-funded by the
Erasmus+ Programme
of the European Union

