

Домашнее задание №1. Управление данными в SPSS. Группа 233

Загрузите базы данных РМЭЗ (RLMS-HSE) за * и * годы с сайта <http://www.hse.ru/rlms/spss> (русскоязычная версия, полная выборка, данные по индивидам). Конкретные годы зависят от варианта задания. Вам потребуется зарегистрироваться на странице загрузки данных.

1. Отредактируйте обе базы данных, оставив в них переменную, идентифицирующую наблюдения (idind), а также переменные с полом, возрастом, образованием и доходом респондентов (не заработной платой!) и переменную, содержащую первичную единицу отбора. Остальные переменные удалите.

2. Объедините две базы данных в одну (в качестве идентификационной переменной используйте idind).

3. Создайте переменную, содержащую разницу в доходе респондентов между двумя годами. Создайте переменную, содержащую средний доход респондентов за два года. Сколько пропущенных значений содержит эта переменная?

4. Из интервальной переменной, содержащей средний доход респондентов, создайте категориальную переменную с тремя значениями: “высокий доход”, “средний доход”, “низкий доход”. Границы между категориями определите сами. Присвойте этой переменной и значениям этой переменной метки.

5. Временно разбейте файл с данными на две подвыборки: с наблюдениями для мужчин и для женщин. В какой подвыборке средний доход выше? На сколько рублей?

6. Измените структуру файла таким образом, чтобы на каждого респондента в каждом году приходилась одна строчка в базе данных (иными словами, переведите данные из “широкого” в “длинный” формат).

7. Создайте базу данных, содержащую процент людей с высшим образованием и медианный доход (в любом из двух годов, но не в обоих сразу) в первичных единицах отбора. В каких первичных единицах отбора наибольший и наименьший процент людей с высшим образованием? В каких первичных единицах отбора наибольший и наименьший медианный доход?

Выполненное домашнее задание должно представлять собой распечатанный файл откомментированного синтаксиса SPSS (комментарии должны объяснять команды, а также содержать ответы на поставленные вопросы).

Домашние задания принимаются только в распечатанном виде. Крайний срок сдачи работ – понедельник, 6 февраля, 15.10 (перед началом

практического занятия). Работы, выполненные с опозданием, не принимаются.

Консультации с товарищами во время подготовки домашних заданий приветствуются. Однако файл синтаксиса должен быть написан и откомментирован самостоятельно. В случае обнаружения заимствований работа оценивается в 0 баллов.