

ТЕМА 3

МОДИФИКАЦИЯ И ПРЕОБРАЗОВАНИЕ ДАННЫХ

Лабораторные работы 5 и 6

Лабораторная работа №5

Управление данными

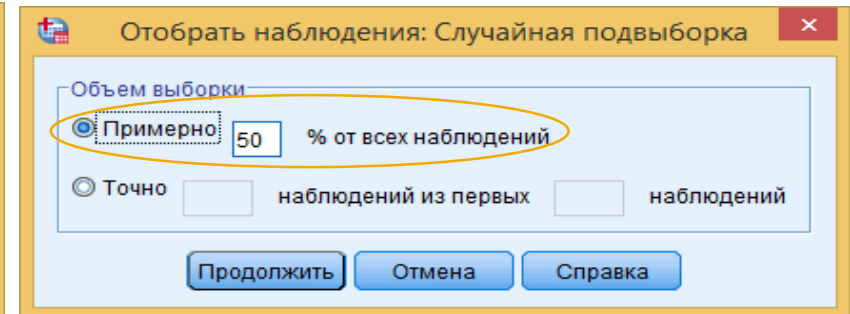
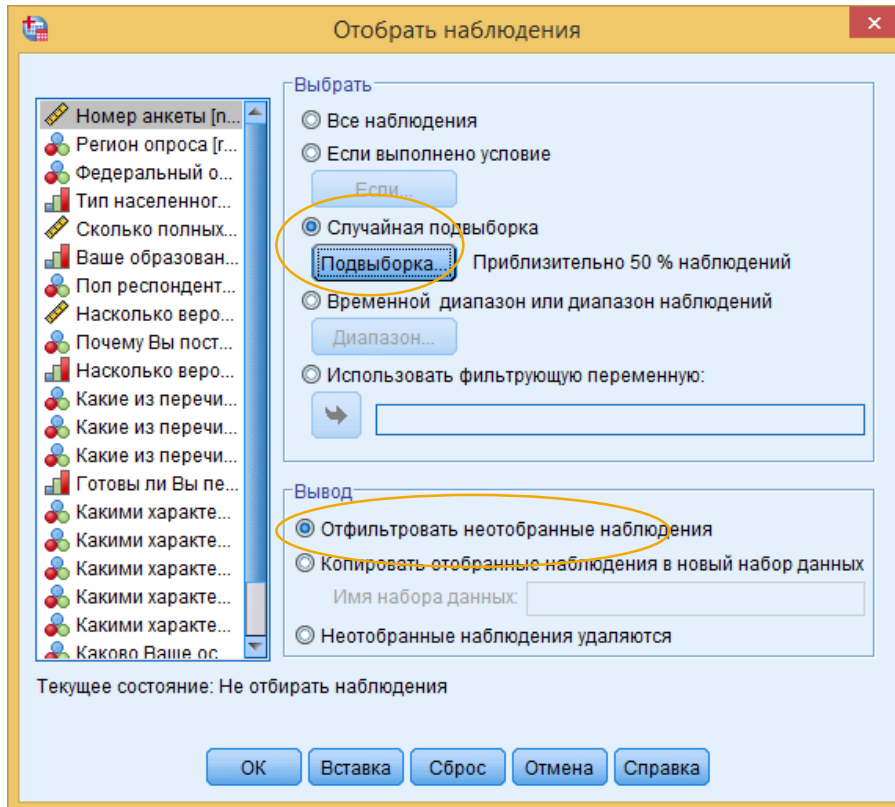
Лабораторная работа №5

Задание:

1. На 50%-массиве отберите две группы респондентов: (1) мужчин до 34 лет и (2) женщин старше 35 лет.
2. Взвесьте их таким образом, чтобы их доли в массиве составляли 50% и 50% соответственно.
3. Узнайте, каким родом деятельности в основном они заняты.
4. Постройте диаграмму с визуализацией результатов.

Лабораторная работа №5

1. Запустите SPSS и откройте файл job.sav.
2. Сделайте случайную 50% подвыборку («Данные» → «Отобразить наблюдения» → «Случайная подвыборка» → «Приблизительно 50% наблюдений»).



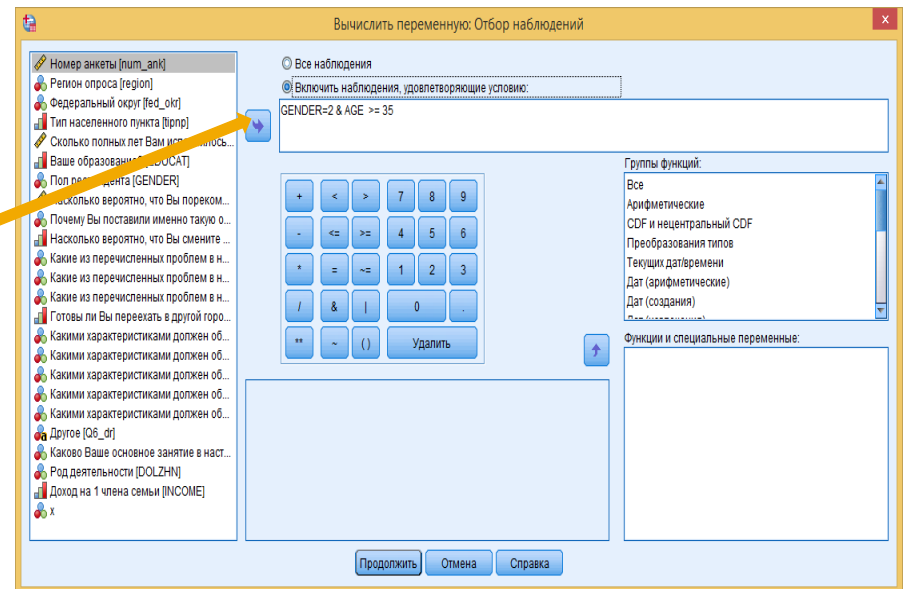
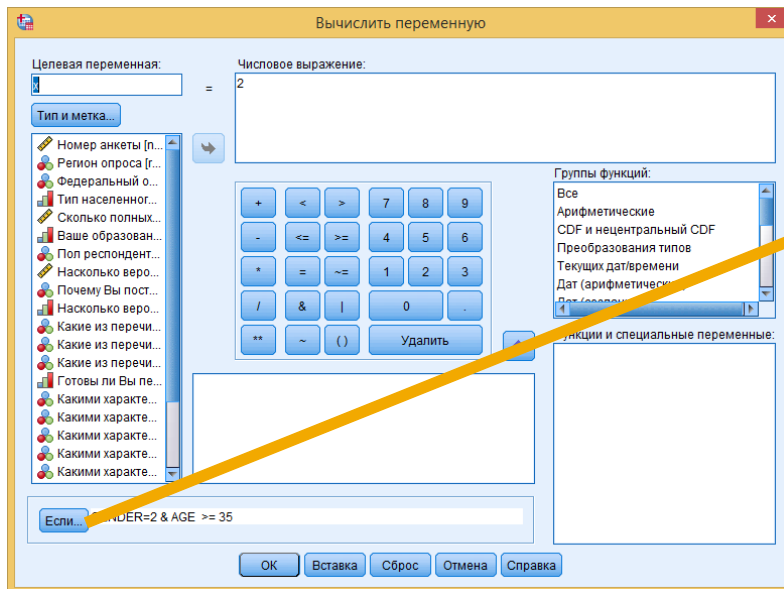
3. Выберите функцию отфильтровать неотобранные наблюдения.
4. Убедитесь, что на вкладке «Представление данные» число наблюдений (респондентов) уменьшилось **вдвое**.

Лабораторная работа №5

4. Отбираем (создаем) две группы респондентов:

- Группа 1. Мужчины до 34 лет (**Gender=1 AND AGE<=34**).
- Группа 2. Женщины старше 35 лет (**Gender=2 AND AGE >=35**).

«Преобразование» → «Вычислить переменную» → Целевую переменную называем Groups → Вводим значение «1» → Задаем условие, нажав кнопку «Если» для отбора в первую группу. Повторяем для второй группы.



Лабораторная работа №5

5. В переменной Groups присвойте в «Представление переменные» значению «1» метку «Мужчины до 34 лет», а значению «2» метку «Женщины старше 35 лет».
6. На вкладке «Представление данные» сортируем получившиеся данные в столбце Groups и удаляем строки с пустыми значениями (выделив их и нажав Delete).
7. Посчитайте частотные распределения для Groups («Анализ» → «Описательные статистики» → «Частоты»). Нормируйте доли 1 и 2 группы 50% на 50%, запустив синтаксис («Файл» → «Создать» → «Редактор синтаксиса»):

compute ves = 0.

if (groups= 1) ves = XX.

if (groups= 2) ves = YY.

exe.

weight by ves.

ves = Population% / Sample%

ves = 50 / xx = XX (для гр.1)

50 / yy = YY (для гр.2)

8. Посчитайте частотные распределения для Groups еще раз и убедитесь, что соотношение групп 50% на 50%.

Лабораторная работа №5

9. Сделайте **отбор респондентов** Группы 1 («Данные» → «Отобратить наблюдения») и посчитайте для них распределение ответов на вопрос об их роде деятельности. Скопируйте таблицу в Excel (Ctrl+C и Ctrl+V).
10. Аналогично сделайте отбор респондентов Группы 2 и посчитайте их распределение ответов. Скопируйте таблицу в Excel.
11. В Excel объедините данные из таблиц так, чтобы получилась таблица для столбчатой диаграммы.

Для этого выполните следующие операции:

Лабораторная работа №5

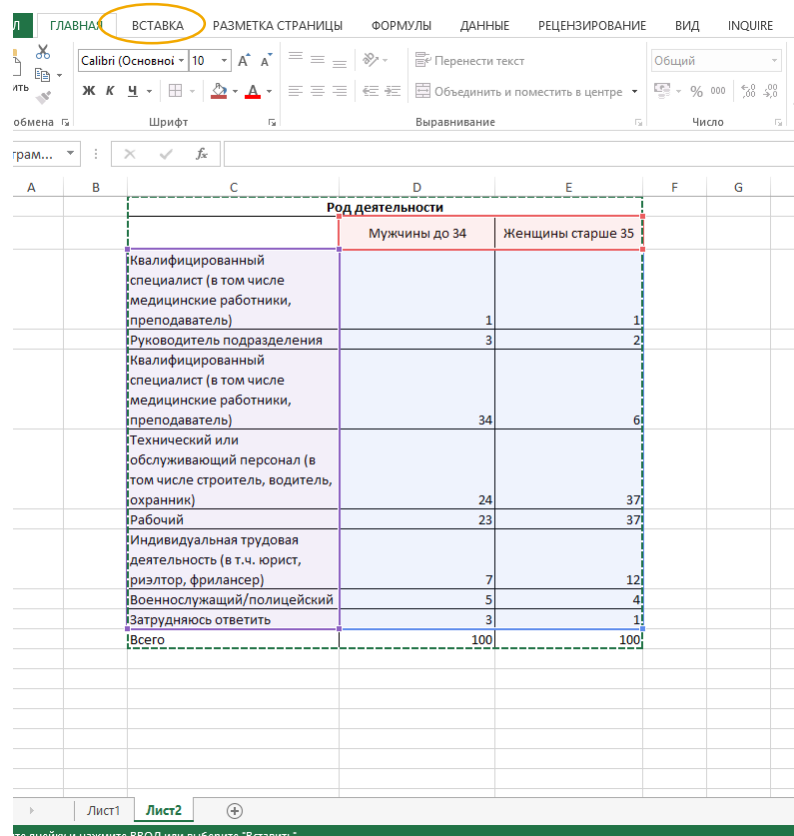
1. Выделите в данных первой группы все числа и уменьшите их разрядность выделенной кнопкой. Повторите с данными по Группе 2.

Род деятельности					Род деятельности				
Допустимо	Частота	Проценты	Процент допустимых	Накопленный процент	Допустимо	Частота	Проценты	Процент допустимых	Накопленный процент
Квалифицированный специалист (в том числе медицинский работник, преподаватель)	1	1,0	1,0	1,0	Квалифицированный специалист (в том числе медицинский работник, преподаватель)	2	1,1	1,1	1,1
Руководитель подразделения, отдела	3	3,0	3,0	4,0	Руководитель высшего звена предприятия, Руководитель подразделения, отдела	4	2,1	2,1	3,2
Квалифицированный специалист (в том числе медицинский работник, преподаватель)	33	33,3	33,3	37,4	Квалифицированный специалист (в том числе медицинский работник, преподаватель)	12	6,4	6,4	9,6
Технический или обслуживающий персонал (в том числе строитель, водитель, охранник)	24	24,2	24,2	61,6	Технический или обслуживающий персонал (в том числе строитель, водитель, охранник)	70	37,4	37,4	47,1
Рабочий	23	23,2	23,2	84,8	Рабочий	68	36,4	36,4	83,4
Индивидуальная трудовая деятельность (в т.ч. юрист, риэлтор, фрилансер)	7	7,1	7,1	91,9	Рабочий	23	12,3	12,3	95,7
Военнослужащий / полицейский	5	5,1	5,1	97,0	Индивидуальная трудовая деятельность (в т.ч. затрудняюсь ответить)	7	3,7	3,7	99,5
Затрудняюсь ответить	3	3,0	3,0	100,0	Затрудняюсь ответить	1	0,5	0,5	100,0
Всего	99	100,0	100,0		Всего	187	100,0	100,0	

2. Удалите столбцы Частота, Процент допустимых и Накопленный процент и строку Допустимо.

Лабораторная работа №5

3. Данные столбца Процент второй группы перенесите к данным первой группы. Переименуйте Процент в названия групп.



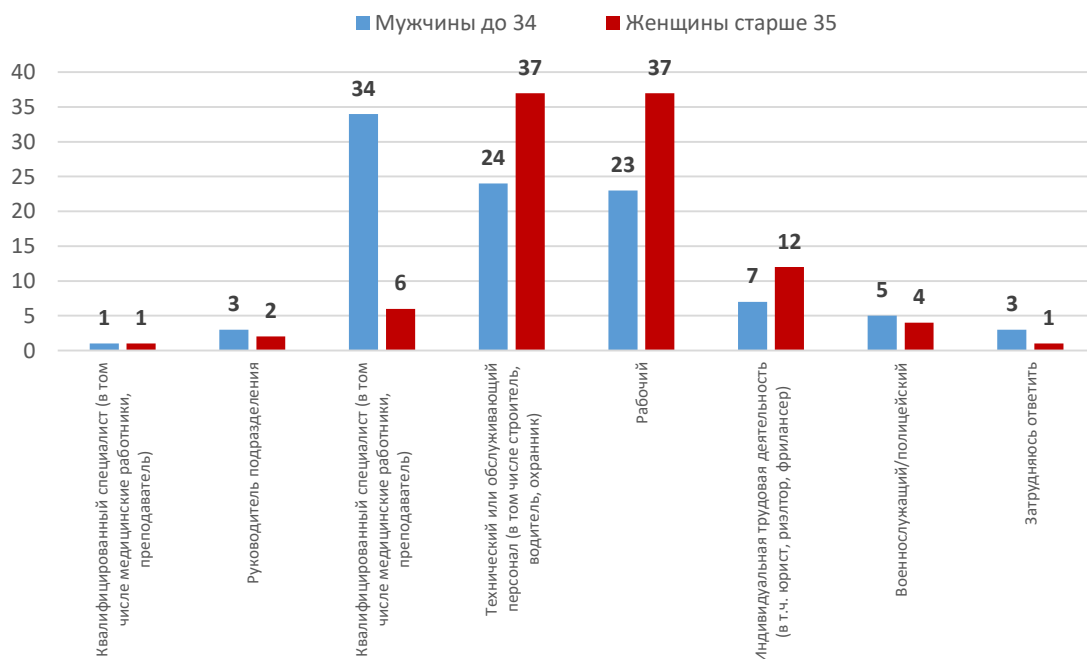
Род деятельности		Процент	Процент
		Мужчины до 34	Женщины старше 35
Квалифицированный специалист (в том числе медицинские работники, преподаватель)		1	1
Руководитель подразделения		3	2
Квалифицированный специалист (в том числе медицинские работники, преподаватель)		34	6
Технический или обслуживающий персонал (в том числе строитель, водитель, охранник)		24	37
Рабочий		23	37
Индивидуальная трудовая деятельность (в т.ч. юрист, риэлтор, фрилансер)		7	12
Военнослужащий/полицейский		5	4
Затрудняюсь ответить		3	11
Всего		100	100

4. Выделите по диагонали данные с названиями групп и всем вариантами ответов. Выберите на вкладке «ВСТАВКА» пункт «Гистограмма».

Лабораторная работа №5

5. Отредактируйте диаграмму так, как показано на рисунке, обратив внимание на следующие действия:

- Формат области диаграммы не имеет границы.
- Легенда вверху.
- Добавлены подписи данных, выделены жирным, шрифт увеличен.
- Линии сетки оформлены серым цветом (возможен пунктир).
- Цвет столбцов изменен на синий и розовый.



Вывод:

Наиболее популярный род деятельности у женщин выявлен в техническом или обслуживающем персонале – 37%. Также 37% из них работают рабочими. Тогда как, наибольший процент у мужчин – 34% работают квалифицированными специалистами.

Лабораторная работа №6

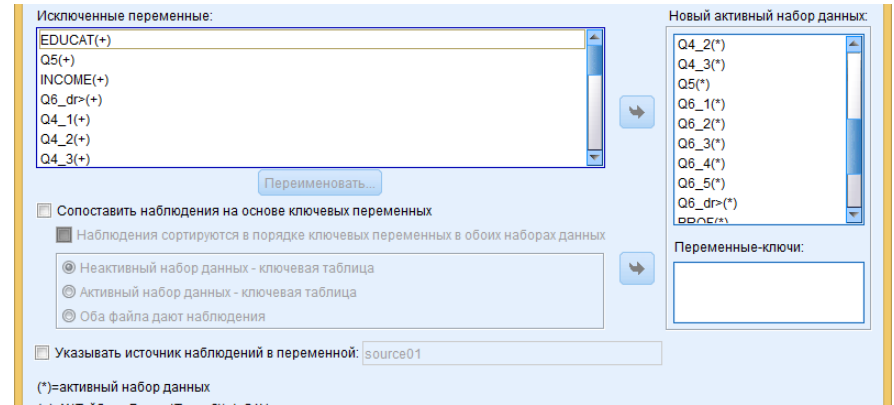
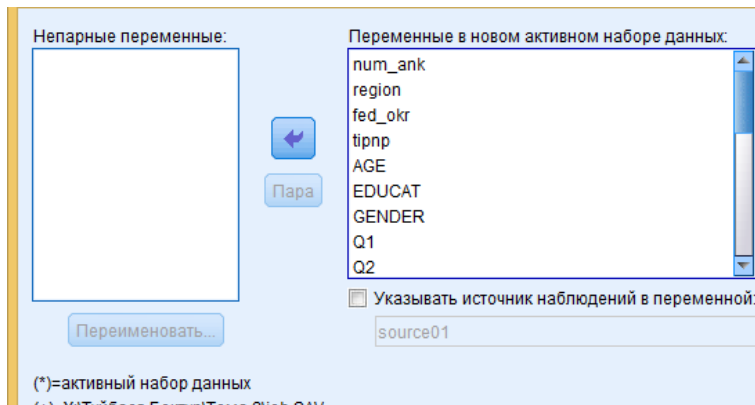
Управление файлами

Лабораторная работа №6

1. Объедините в единый массив с названием job MSPb+C.sav респондентов, опрошенных в Москве, Санкт-Петербурге и респондентов, опрошенных в городах миллиониках (в папке «Добавить наблюдения»).
2. Добавьте к массиву job (q1).sav переменные вопроса №2 из файла job (q2).sav (в папке «Добавить переменные»). Новый массив сохраните под именем job_q1&q2.sav.

Лабораторная работа №6

1. Для выполнения заданий воспользуйтесь функцией «**Данные**» → «**Слить файлы**».
2. Для объединения наблюдений используйте команду «**Добавить наблюдения**», для объединения переменных «**Добавить переменные**»



3. При «Добавить наблюдения» следите, чтобы не появлялись непарные переменные, при «Добавить переменные» – контролируйте добавляемые переменные.