

Семинар 4. SPSS

Дисперсионный анализ

Арина Кузьмичева

ТГ – @fevrier_rin

Почта – adkuzmicheva@hse.ru

Дисперсионный анализ

Дисперсионный анализ

Дисперсионный анализ (Analysis of Variance, ANOVA)

– статистический метод выявления различий между выборочными средними для двух или больше совокупностей.

Когда нужен?

Определение того, является ли различия между 3-мя и более средними значениями (средними значениями в 3-х и более группах) статистически значимыми;

Определение того, оказывает ли один категориальный признак (по которому выделяются группы), имеющий 3 и более значения, воздействие на интервальный признак (средние значения которого сравниваются в выделенных группах);

Как работает?

С помощью F-статистики проверяется статистическая гипотеза:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_n$$

Где n – количество сравниваемых средних (кол-во групп)

H_1 : Не равны хотя бы 2 средних значения

Суть F-статистики – сравнение межгрупповой дисперсии средних (MS_b) и внутригрупповой дисперсии отдельных значений зависимой переменной (MS_w).

Как работает?

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_n$$

Отклонена

Не отклонена

Не равны хотя бы 2 средних значения

Нельзя утверждать, что средние значения признака различаются хотя бы в двух группах

Какие именно? Неизвестно.
Что делать?

Сравнивать средние попарно с помощью методов парных сравнений и PostHoc критериев

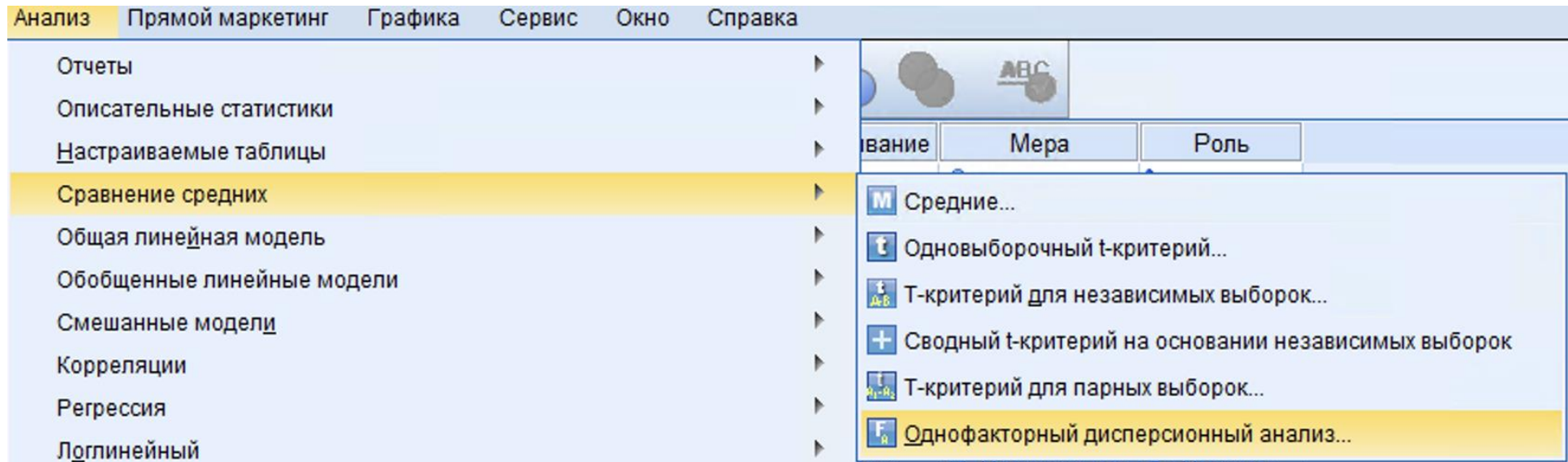
Реализация в SPSS

Сначала проводим ANOVA (однофакторный дисперсионный анализ)

В зависимости от результатов:

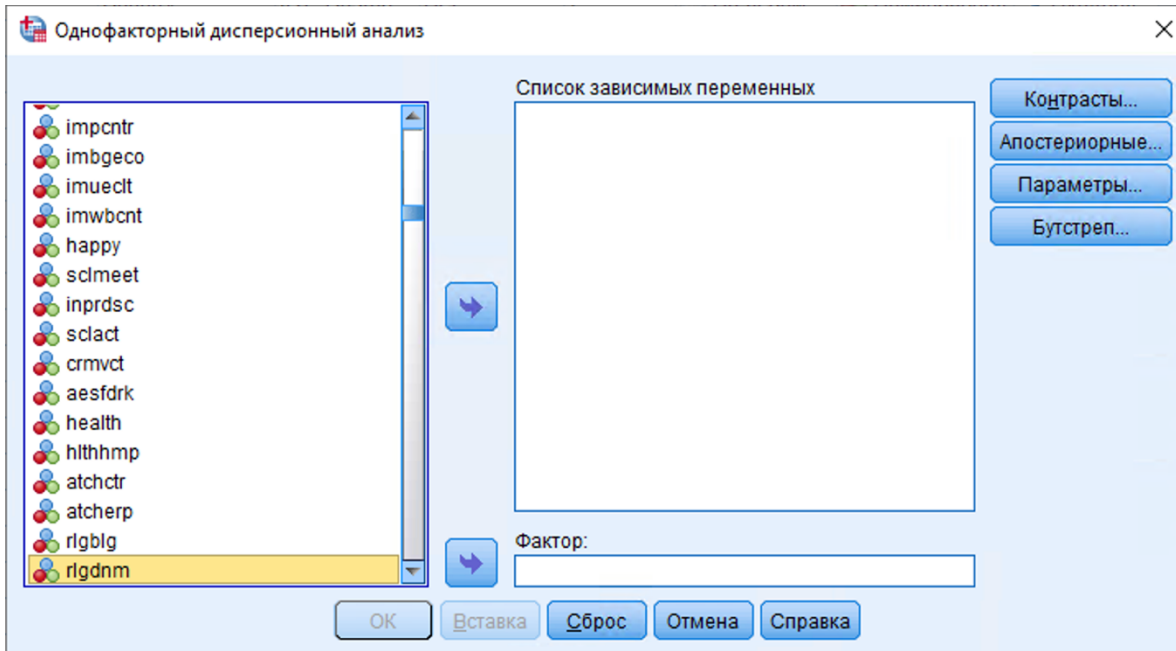
а) делаем вывод о равенстве средних

б) делаем вывод о том, что хотя бы 2 средних не равны, и идем выяснять, какие именно, с помощью попарных сравнений



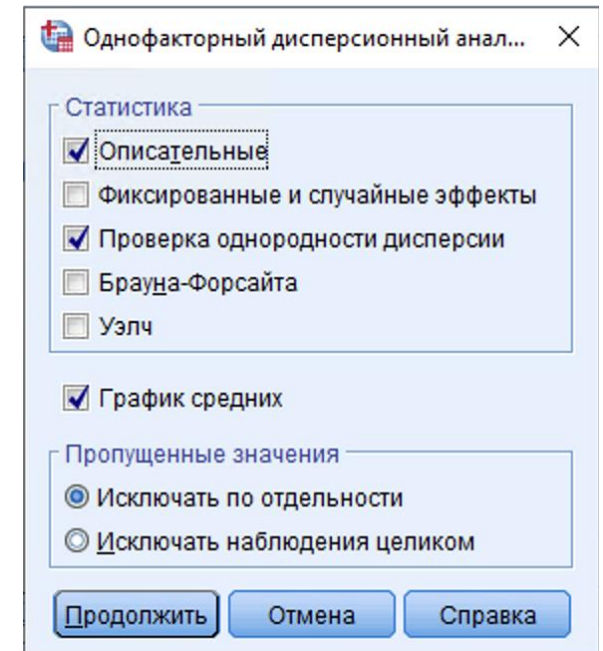
Реализация в SPSS

- 1) Указываем зависимые переменные (они должны быть интервальными или псевдоинтервальными!)
- 2) Указываем фактор (номинальную или порядковую переменную с несколькими категориями, в идеале она должна иметь 3-8 категорий)



В параметрах выбираем:

- описательные статистики
- проверка однородности дисперсии
- график средних (этого достаточно)



Реализация в SPSS

1. Смотрим на значимость в таблице «Anova».

ANOVA

Internet use, how much time on typical day, in minutes

	Сумма квадратов	ст.св.	Средний квадрат	F	Значимость
Между группами	182631,387	1	182631,387	6,199	,013
Внутри групп	886903362,6	30104	29461,313		
Всего	887085994,0	30105			

2. Если гипотеза не отвергается, то можем считать средние равными (на этом шаге анализ закончился).

Если гипотеза отвергается в пользу альтернативной, то переходим к попарным сравнениям и применению PostHoc критериев.

Реализация в SPSS

3. Возвращаемся в Анализ → Сравнение средних → Однофакторный дисперсионный анализ → Апостериорные множественные сравнения → Выбираем PostHoc критерий и указываем уровень значимости

Однофакторный дисперсионный анализ: Апостериорные множественные сравнения

Предполагается равенство дисперсий

<input type="checkbox"/> НЗР	<input type="checkbox"/> С-Н-К	<input type="checkbox"/> Уоллер-Дункан
<input type="checkbox"/> Бонферрони	<input type="checkbox"/> Тьюки	Отношение ошибок тип I/тип II: 100
<input type="checkbox"/> Шидак	<input type="checkbox"/> b Тьюки	<input type="checkbox"/> Даннетт
<input type="checkbox"/> Шеффе	<input type="checkbox"/> Дункан	Контрольная категория: Последняя
<input type="checkbox"/> F P-Э-Г-У	<input type="checkbox"/> GT2 Гохберга	Критерий <input checked="" type="radio"/> 2-сторонний <input type="radio"/> < эталонного значения <input type="radio"/> > эталонного значения
<input type="checkbox"/> Q P-Э-Г-У	<input type="checkbox"/> Габриэль	

Равенство дисперсий не предполагается

<input type="checkbox"/> T2 Тамхейна	<input type="checkbox"/> T3 Даннетт	<input checked="" type="checkbox"/> Геймс-Хоуэлл	<input type="checkbox"/> С Даннетта
--------------------------------------	-------------------------------------	--	-------------------------------------

Уровень значимости: 0,05

Продолжить Отмена Справка

Выбор PostHoc критерия

- Смотрим на результаты теста Ливиня (о равенстве дисперсий)
- Смотрим на объем выборок

Критерий однородности дисперсий



Internet use, how much time on typical day, in minutes

Статистика Ливиня	ст.св.1	ст.св.2	Значимость
1,127	1	30104	,288

Критерий однородности дисперсий



News about politics and current affairs, watching, reading or listenin

Статистика Ливиня	ст.св.1	ст.св.2	Значимость
23,138	7	25776	,000

По результатам проверки гипотезы о гомогенности дисперсий оказалось, что дисперсии зависимой переменной в группах одинаковы	Выборки одинакового объема	Тьюки (Tuckey)
	Выборки разного объема	Шеффе (Scheffe)
По результатам проверки гипотезы о гомогенности дисперсий оказалось, что дисперсии зависимой переменной в группах разные	Выборки одинакового объема	Тамхейн (Tamhane's T2)
	Выборки разного объема	Геймс-Хоуэл (Games-Howell)

Применимость

1. Зависимая переменная интервальная;
2. Сравнимые выборки независимы;
3. Объем сравниваемых выборок (групп) не менее 30 объектов
ИЛИ зависимая переменная должна иметь нормальное распределение;

Дополнительно про PostHoc критерии

Самые популярные и сбалансированные помимо тех, что уже упомянули:

- **Тьюки-Крамер (b Тьюки)** – дисперсии равны, для неравных размеров групп.
- **Бонферрони (Bonferroni)** – дисперсии равны, количество запланированных сравнений невелико (например, 3-4), меньше ошибок I рода, но меньше мощность.
- **Шидак** – похож на Бонферрони, но еще мощнее. Корректирует уровень α для каждого отдельного сравнения.
- **Габриэль** – для неравных размеров групп, использует взвешенное среднее размеров групп. **GT2 Гохберга** – похож на Габриэль, но лучше контролирует ошибку I рода.

Более мощные, но выше вероятность ошибки I рода:

- **LSD Фишера (H3P)** – только, если ANOVA дал очень значимый результат ($p < 0.001$) и вы хотите найти все возможные различия.
- **Стьюдент-Ньюман-Кейлс (SNK)** – упорядочивает группы по средним и последовательно сравнивает их.
- **Дункан (Duncan)** – похож на SNK, но еще выше вероятность ошибки I рода.

Дополнительно про PostHoc критерии

Критерии, регулирующие мощность по типу F-статистики:

- Райан-Эйноут-Габриэль-Вельш (REGWF / F P-Э-Г-У) – использует F-распределение. Хороший баланс между мощностью и контролем ошибки.
- Райан-Эйноут-Габриэль-Вельш (REG WQ / Q P-Э-Г-У) – использует студентизированный диапазон (Q). Аналогичен предыдущему.

Критерии, основанные на Байесовском подходе:

- Уоллер-Дункан (Waller-Duncan) – использует Байесовские принципы. Менее распространен.

Критерии для сравнения с контрольной группой:

- Даннетт (Dunnnett) – специализированный тест, когда нужно сравнить несколько экспериментальных групп с одной контрольной группой. Эффективнее и мощнее, чем попарные тесты в этой ситуации.
- Даннетта С (Dunnnett's C) – модификация Даннетта для случая неравных дисперсий.

Пример 1

Определите, различается ли средняя продолжительность времени потребления политического контента в СМИ среди людей различных конфессий

Пример 1

Определите, различается ли средняя продолжительность времени потребления политического контента в СМИ среди людей различных конфессий

ANOVA

News about politics and current affairs, watching, reading or listening

	Сумма квадратов	ст.св.	Средний квадрат	F	Значимость
Между группами	1066495,555	7	152356,508	8,108	,000
Внутри групп	484370628,5	25776	18791,536		
Всего	485437124,0	25783			

1. По результатам ANOVA видим, что нулевая гипотеза о равенстве всех средних между группами отвергается в пользу альтернативной на 95% уровне доверительной вероятности (средние не равны, нужно понять, какие именно)

Критерий однородности дисперсий

News about politics and current affairs, watching, reading or listening

Статистика Ливиня	ст.св.1	ст.св.2	Значимость
23,138	7	25776	,000

2. По тесту Ливиня видим, что нулевая гипотеза о равенстве дисперсий отвергается в пользу альтернативной на 95% уровне доверительной вероятности

3. По табличке с описательными статистиками видим, что объем выборок в группах различается → используем PostHoc критерий Геймса-Хоуэла

Пример 1

Определите, различается ли средняя продолжительность времени потребления политического контента в СМИ среди людей различных конфессий

Множественные сравнения

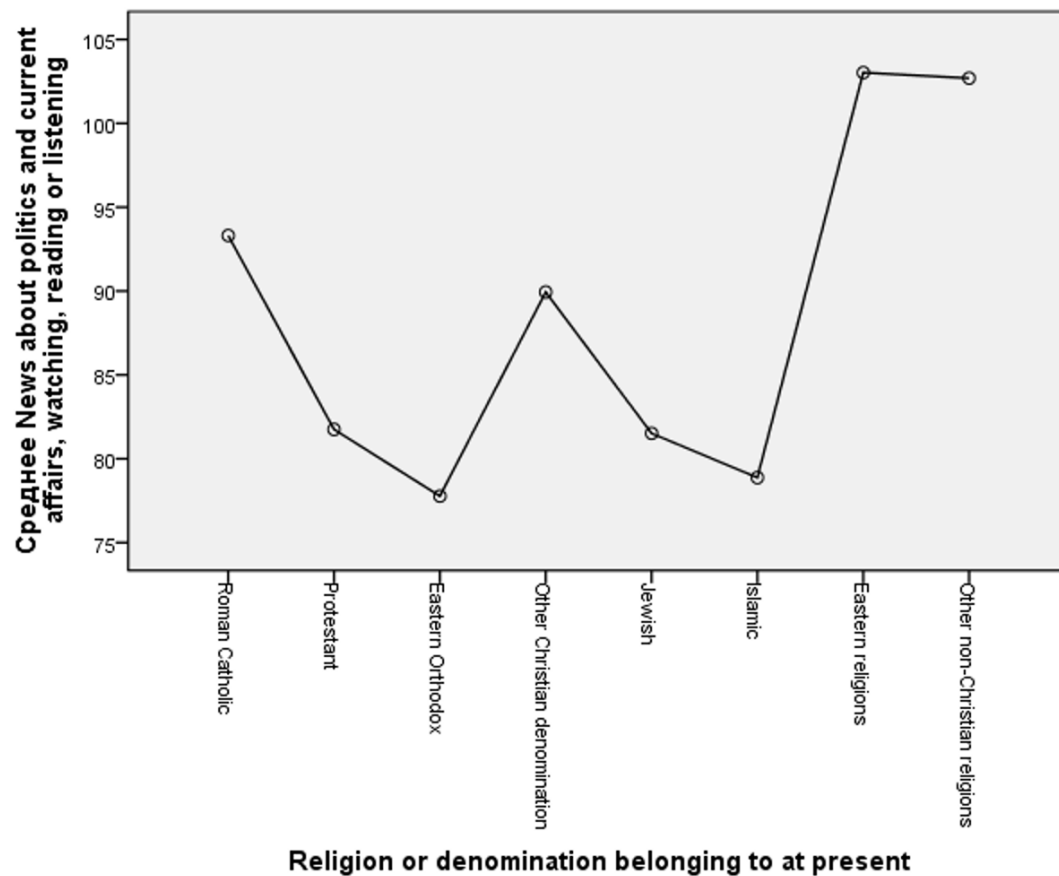
Зависимая переменная: News about politics and current affairs, watching, reading or listening
Геймс-Хоуэлл

(I) Religion or denomination belonging to at present	(J) Religion or denomination belonging to at present	Средняя разность (I-J)	Стандартная ошибка	Значимость	95% доверительный интервал	
					Нижняя граница	Верхняя граница
Roman Catholic	Protestant	11,561 [*]	1,989	,000	5,53	17,59
	Eastern Orthodox	15,538 [*]	2,474	,000	8,03	23,04
	Other Christian denomination	3,371	8,063	1,000	-21,18	27,92
	Jewish	11,796 [*]	2,472	,000	4,30	19,29
	Islamic	14,434 [*]	3,740	,003	3,08	25,78
	Eastern religions	-9,716	15,477	,998	-57,25	37,82
	Other non-Christian religions	-9,384	17,621	,999	-63,71	44,95
Protestant	Roman Catholic	-11,561 [*]	1,989	,000	-17,59	-5,53
	Eastern Orthodox	3,977	2,635	,803	-4,01	11,97
	Other Christian denomination	-8,190	8,113	,973	-32,89	16,51
	Jewish	,235	2,633	1,000	-7,75	8,22
	Islamic	2,874	3,848	,996	-8,80	14,55
	Eastern religions	-21,276	15,503	,868	-68,89	26,34
	Other non-Christian religions	-20,944	17,644	,934	-75,34	33,45

SPSS строит нам гигантскую таблицу, где попарно сравниваются все конфессии. Исходя из нее мы можем увидеть, какие средние попарно равны, а какие не равны

Пример 1

Графики средних



Для простоты анализа можно также опираться на график средних

Пример 2

Определите, различается ли средняя продолжительность времени, проведенного в Интернете, среди людей голосовавших за разные партии на последних выборах в Чехии (prtvtdcz)

Пример 2

Определите, различается ли средняя продолжительность времени, проведенного в Интернете, среди людей голосовавших за разные партии на последних выборах в Чехии (prtvtdcz)

ANOVA

Internet use, how much time on typical day, in minutes

	Сумма квадратов	ст.св.	Средний квадрат	F	Значимость
Между группами	733987,341	7	104855,334	4,195	,000
Внутри групп	18995809,18	760	24994,486		
Всего	19729796,52	767			

1. По результатам ANOVA видим, что нулевая гипотеза о равенстве всех средних между группами отвергается в пользу альтернативной на 95% уровне доверительной вероятности (средние не равны, нужно понять, какие именно)

Критерий однородности дисперсий

Internet use, how much time on typical day, in minutes

Статистика Ливиня	ст.св.1	ст.св.2	Значимость
3,758	7	760	,001

2. По тесту Ливиня видим, что нулевая гипотеза о равенстве дисперсий отвергается в пользу альтернативной на 95% уровне доверительной вероятности

3. По табличке с описательными статистиками видим, что объем выборок в группах различается → используем PostHoc критерий Геймса-Хоуэла

Пример 2

Множественные сравнения

Зависимая переменная: Internet use, how much time on typical day, in minutes

Геймс-Хоуэлл

(I) Party voted for in last national election, Czech Republic	(J) Party voted for in last national election, Czech Republic	Средняя разность (I-J)	Стандартная ошибка	Значимость	95% доверительный интервал	
					Нижняя граница	Верхняя граница
KSČM	ČSSD	21,197	27,806	,994	-65,91	108,31
	TOP 09	-65,178	31,681	,450	-163,29	32,93
	ANO 2011	-1,873	27,860	1,000	-89,13	85,38
	ODS	-48,957	32,597	,805	-149,83	51,92
	KDU-ČSL	-36,632	36,762	,974	-150,71	77,45
	Úsvit přímé demokracie Tomia Okamury	26,559	35,610	,995	-84,44	137,55
	Other	-19,680	34,575	,999	-127,10	87,74
ČSSD	KSČM	-21,197	27,806	,994	-108,31	65,91
	TOP 09	-86,375*	20,425	,001	-149,03	-23,72
	ANO 2011	-23,070	13,772	,703	-65,03	18,89
	ODS	-70,155*	21,818	,035	-137,45	-2,86
	KDU-ČSL	-57,829	27,658	,434	-145,09	29,44
	Úsvit přímé demokracie Tomia Okamury	5,361	26,106	1,000	-78,23	88,95
	Other	-40,877	24,677	,714	-118,93	37,17
TOP 09	KSČM	65,178	31,681	,450	-32,93	163,29
	ČSSD	86,375*	20,425	,001	23,72	149,03
	ANO 2011	63,305*	20,499	,047	,45	126,16
	ODS	16,221	26,581	,999	-65,29	97,73
	KDU-ČSL	28,546	31,551	,985	-69,56	126,66
	Úsvit přímé demokracie Tomia Okamury	91,737	30,200	,064	-2,86	186,33
	Other	45,498	28,974	,766	-44,55	135,55

Определите, различается ли средняя продолжительность времени, проведенного в Интернете, среди людей голосовавших за разные партии на последних выборах в Чехии (prtvtdcz)

Потренируемся к контрольной работе

1

Опишите особенности распределения переменной ... с переносом на генеральную совокупность на 99% уровне доверительной вероятности.

1

Опишите особенности распределения переменной ... с переносом на генеральную совокупность на 99% уровне доверительной вероятности.

Рассчитайте все возможные для данного типа шкалы меры среднего и проинтерпретируйте их. В случае среднего арифметического не забудьте построить доверительный интервал.

1

Опишите особенности распределения переменной `mnrgrtjb` с переносом на генеральную совокупность на 99% уровне доверительной вероятности.

Рассчитайте все возможные для данного типа шкалы меры среднего и проинтерпретируйте их. В случае среднего арифметического не забудьте построить доверительный интервал.

Рассчитайте все возможные для данного типа шкалы меры разброса и проинтерпретируйте их. Коэффициент вариации/качественной вариации можно не считать.

2

1. Отберите наблюдения по своему полу.
2. Возьмите переменную e_{effar} .
3. Постройте частотное распределение по данной переменной.
4. Определите все доступные меры среднего и разброса, перенесите среднее на генеральную совокупность с 90% доверительной вероятностью.
5. Запишите интерпретацию результатов (одно предложение).

3

Перекодируйте переменную `elgcoal` так, чтобы:

- Варианты 1 и 2 приняли значение 1 (rather a large amount)
- Вариант 3 принял значение 2 без изменения содержания
- Варианты 4 и 5 приняли значение 3 (rather a small amount)

Новую переменную назовите `elgcoal_1`

Не забудьте указать значения переменной и пропущенные

4

Перекодируйте переменную `ssrdprs` так, чтобы:

- все, кто дал оценку в 4 балла и ниже, попали в группу 1 – «не чувствуют ответственности»
- все, кто дал оценку 5 баллов и выше, попали в группу 2 – «чувствуют ответственность»

5

С 99% вероятностью определите, отличается ли средний уровень веры в то, что сокращение собственного потребления энергии может остановить климатические изменения (ownrdcc) среди португальцев от среднего по всей Европе уровня, равного 4,35 баллам?

6

С 95% вероятностью определите, различаются ли в России по количеству сотрудников, за которых несут ответственность (n_{jsprv}), мужчины и женщины.

7

С 90% доверительной вероятностью определите, одинаковы ли средняя оценка уровня жизни пенсионеров (slvpens) и средняя оценка уровня жизни безработных (slvuemp) среди тех немцев, кто заинтересован в политике (polintr).