

# РАБОТА В SPSS. ЧАСТЬ II.

# МЕТОДЫ

<https://ru.coursera.org/learn/ekonometrika> (R)

<https://ru.coursera.org/course/pdstatistics>

<https://ru.coursera.org/course/introstats>

# ЛОГИСТИЧЕСКАЯ РЕГРЕССИЯ В SPSS

LOGIT\_REGR\_S.SPS; FILE01.SAV

Зависимая переменная –  
дихотомическая

Exp(B) – повышение/снижение  
отношения шансов того, что событие  
произойдет (1) к тому, что событие не  
произойдет (0) в \* раз(a)

Референтная категория!

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>						
a6.22			46,234	3	,000	
a6.22(1)	,431	,292	2,178	1	,140	1,538
a6.22(2)	1,240	,272	20,785	1	,000	3,454
a6.22(3)	1,923	,311	38,225	1	,000	6,840
agr			26,158	3	,000	
agr(1)	1,969	,386	26,027	1	,000	7,167
agr(2)	1,559	,382	16,614	1	,000	4,754
agr(3)	1,386	,387	12,820	1	,000	3,997
part_m			7,483	3	,058	
part_m(1)	,351	,288	1,489	1	,222	1,421
part_m(2)	,051	,323	,025	1	,874	1,053
part_m(3)	-,425	,332	1,635	1	,201	,654
n_cld3			13,063	2	,001	
n_cld3(1)	-,592	,247	5,739	1	,017	,553
n_cld3(2)	-1,516	,423	12,832	1	,000	,220
oclass4			9,867	4	,043	
oclass4(1)	,619	,318	3,801	1	,051	1,858
oclass4(2)	,577	,262	4,842	1	,028	1,781
oclass4(3)	,613	,237	6,668	1	,010	1,846
oclass4(4)	-,058	,426	,018	1	,892	,944
Constant	-4,215	,506	69,444	1	,000	,015

a. Variable(s) entered on step 1: a6.22. agr. part\_m. n\_cld3. oclass4.

# ФИЛЬТР: ТОЛЬКО ТЕ, У КОГО БЫЛ ПАРТНЕР В 2004 ГОДУ

Сами ставим фильтр

Case Processing Summary

Unweighted Cases <sup>a</sup>		N	Percent
Selected Cases	Included in Analysis	1205	100,0
	Missing Cases	0	,0
	Total	1205	100,0
Unselected Cases		0	,0
Total		1205	100,0

a. If weight is in effect, see classification table for the total number of cases.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>			41,311	3	,000	
a6.22						
a6.22(1)	,348	,397	,768	1	,381	1,416
a6.22(2)	1,462	,357	16,775	1	,000	4,313
a6.22(3)	2,256	,392	33,163	1	,000	9,544
agr			16,607	3	,001	
agr(1)	1,945	,486	16,045	1	,000	6,996
agr(2)	1,672	,457	13,366	1	,000	5,323
agr(3)	1,456	,460	10,028	1	,002	4,290
n_cld3			4,153	2	,125	
n_cld3(1)	-,173	,329	,275	1	,600	,841
n_cld3(2)	-,907	,492	3,393	1	,065	,404
part_m(1)	,215	,296	,530	1	,467	1,240
oclass4			8,693	4	,069	
oclass4(1)	,749	,550	1,854	1	,173	2,114
oclass4(2)	,468	,512	,833	1	,361	1,596
oclass4(3)	-,278	,518	,288	1	,592	,757
oclass4(4)	,297	,496	,358	1	,550	1,346
Constant	-4,490	,735	37,296	1	,000	,011

a. Variable(s) entered on step 1: a6.22, agr, n\_cld3, part\_m, oclass4.

# КОМАНДЫ

Наизусть

freq

count

compute с условиями (!)

Часто нужны, можно сделать шаблоны

Stables (!)

Datediff, date.mogr – работа с датами

Регрессии

# ПРОВЕРКА НАМЕРЕНИЙ ПАРТНЕРОВ

```
compute int_pair1=0.
```

```
if (a6.15a_1>1 and a6.11_2>1) int_pair1=1.
```

```
if (a6.15a_1=1 & a6.11_2>1) int_pair1=2.
```

```
if (a6.11_2=1 & a6.15a_1>1) int_pair1=3.
```

```
if (a6.15a_1=1 and a6.11_2=1) int_pair1=4.
```

```
if (part_m=4) int_pair1=5.
```

```
recode int_pair1 (0,5=sysmis) (else=copy) into int_pair1.
```

```
var lab int_pair1 'репродуктивные намерения партнеров 04'.
```

```
val lab int_pair1 1 'оба не хотят' 2 'мужчина хочет, женщина нет или не уверена'
```

```
3 'женщина хочет, мужчина нет или не уверен или женщина не знает' 4 'оба хотят'.
```

```
exe.
```

```
freq int_pair1.
```

```
freq int_pair.
```

# ПРОВЕРКА НАМЕРЕНИЙ ПАРТНЕРОВ

Case Processing Summary

Unweighted Cases <sup>a</sup>		N	Percent
Selected Cases	Included in Analysis	1200	99,6
	Missing Cases	5	,4
	Total	1205	100,0
Unselected Cases		0	,0
Total		1205	100,0

a. If weight is in effect, see classification table for the total number of cases.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>						
int_pair			8,150	3	,043	
int_pair(1)	-,231	,358	,416	1	,519	,794
int_pair(2)	,031	,485	,004	1	,949	1,031
int_pair(3)	,579	,349	2,744	1	,098	1,783
bn6.32_gr			22,312	2	,000	
bn6.32_gr(1)	-1,333	,458	8,458	1	,004	,264
bn6.32_gr(2)	,985	,295	11,171	1	,001	2,677
agr			13,112	3	,004	
agr(1)	1,706	,573	8,868	1	,003	5,505
agr(2)	1,778	,504	12,446	1	,000	5,920
agr(3)	1,597	,489	10,659	1	,001	4,937
n_cld3			21,872	2	,000	
n_cld3(1)	-,689	,342	4,070	1	,044	,502
n_cld3(2)	-2,416	,535	20,396	1	,000	,089

# СЛИЯНИЕ ФАЙЛОВ (СОЗДАНИЕ ПАНЕЛИ)

1 WAVE GGS.SAV; 2 WAVE GGS.SAV

1. отсортировать по идентификатору

Merge -> add variables

Syntax: ***MATCH FILES***

Как получить 7786:

- удалить из второй волны новых респондентов
- к ней приклеить 1-ую волну методом “non-active dataset is keyed table”

Как получить 14592:

- склеить 1 и 2 методом “both files provide cases”

NB: имена переменных должны различаться

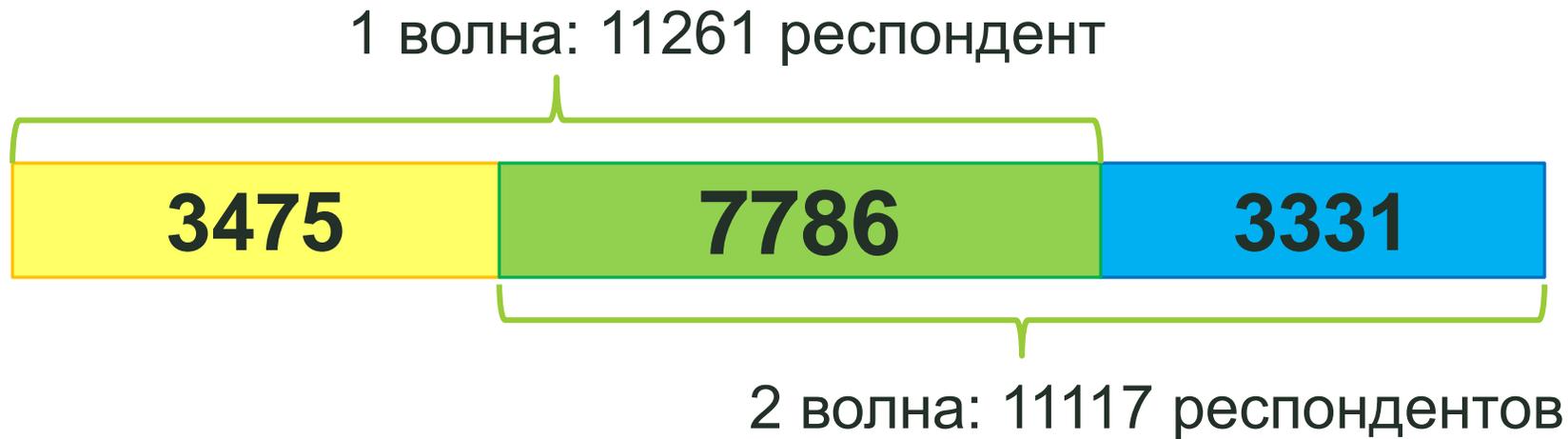
# СКЛЕИВАНИЕ ПАНЕЛЬНОГО МАССИВА РИДМИЖ

Идентификаторы респондента:

- anumber 2004-2007-2011
- bredid 2007-2011
- credid 2011

**Сортировка!!**

Склеивая 2 волны, можно получить 4 разных массива



**Полная панель: 7786**

**Неполные панели: 7786 (2раза) +3475 или 7786 (2раза) +3331**

**Pooled cross: 14592**

# ПАНЕЛЬНЫЕ ВЫБОРОЧНЫЕ ОБСЛЕДОВАНИЯ: МЕТОДИКА РАБОТЫ

Два вида панельного массива данных:

- Полная панель

- респонденты одни и те же
- используется для анализа динамики изменений
- используется для «чистки» переменных

- Pooled cross-sectional

- часть респондентов одинакова, часть – разная
- чаще используется для анализа, где каждое наблюдение трактуется как отдельный респондент (расширение объема выборки)

# ФАКТОРНЫЙ АНАЛИЗ В SPSS

Сокращение числа измерений (переменных). Чаще всего используется для разложения «батарей вопросов»

FACTOR

```
/VARIABLES a11.07.1 a11.07.2 a11.07.3 a11.07.4 a11.07.5 a11.07.6 a11.07.7 a11.07.8 a11.07.9 a11.0710 a11.0711  
a11.0712
```

```
/MISSING LISTWISE
```

```
/ANALYSIS a11.07.1 a11.07.2 a11.07.3 a11.07.4 a11.07.5 a11.07.6 a11.07.7 a11.07.8 a11.07.9 a11.0710 a11.0711  
a11.0712
```

```
/PRINT INITIAL EXTRACTION
```

```
/CRITERIA MINEIGEN(1) ITERATE(25)
```

```
/EXTRACTION PC
```

```
/ROTATION NOROTATE
```

```
/METHOD=CORRELATION.
```

# НАПРИМЕР, ИНДЕКСЫ

## (НАПРАВЛЕНИЕ ФАКТОРОВ, НОРМИРОВАНИЕ)

ИНДЕКС межпоколенческой поддержки

```
recode a11.10.1 a11.10.2 a11.10.3 a11.10.4 a11.10.5 a11.11.1 a11.11.2 a11.11.3 a11.12.1 a11.12.2 a11.12.3 a11.12.4 a11.12.5 (6 thru hi =3).
```

\*переориентация (максимум-максимум):

```
recode a11.11.1 a11.11.2 a11.11.3 a11.12.1 a11.12.2 a11.12.3 a11.12.4 a11.12.5 (1=5) (2=4) (3=3) (4=2) (5=1).
```

RELIABILITY

```
/VARIABLES=a11.10.1 a11.10.2 a11.10.3 a11.10.4 a11.10.5 a11.11.1 a11.11.2
```

```
a11.11.3 a11.12.1 a11.12.2 a11.12.3 a11.12.4 a11.12.5
```

```
/SCALE('ALL VARIABLES') ALL/MODEL=ALPHA
```

```
/SUMMARY=TOTAL .
```

```
compute ind_sup= sum(a11.10.1, a11.10.2, a11.10.3, a11.10.4, a11.10.5, a11.11.1, a11.11.2, a11.11.3, a11.12.1, a11.12.2, a11.12.3, a11.12.4, a11.12.5)/65.
```

```
var lab ind_sup 'индекс блока ценностей и установок'.
```

```
recode ind_sup (lo thru 0.600 = 1) (0.601 thru 0.740 =2) (0.741 thru hi = 3) into ind_sup3.
```

```
exe.
```

# КЛАСТЕРНЫЙ АНАЛИЗ В SPSS

1. Посмотреть согласованность,  $\alpha$ -Кронбаха. Выявление групп наблюдений, похожих друг на друга
2. Twostep CLUSTER / K-means cluster / иерархический... много методов