

Факторный анализ: Практика применения:

Лекция с основой практической работы

(Материал подготовлен на базе контента образовательного Интернет-ресурса Алексея Ротмистрова, расположенного на YouTube по адресу: <https://www.youtube.com/channel/UCq23csisa-bjoTEFSVxgn7A/videos>)

Попробуем разобраться в том, как работает факторный анализ. Первичное знакомства с факторным анализом у нас уже состоялось: мы уже рассмотрели (я **надеюсь!!!**) теоретический материал с сайта НАФИ, содержащий теорию; мы про факторный анализ уже говорили в аудитории до объявления режима самоизоляции; группа делала домашнее задание, связанное с содержательным анализом материалов факторного анализа, изложенных в социологической статье... Теперь же, попробуем подвести некоторый итог.

Представим себе, что мы имеем 6 переменных, отражающих мотивы, которыми руководствуются родители, когда водят детей в разные развлекательные центры. В дальнейшем, мы их будем называть «наблюдаемые переменные». Мы предполагаем, что эти мотивы можно сгруппировать. То есть, мы предполагаем, что за шестью наблюдаемыми переменными стоит меньшее число переменных. И с помощью факторного анализа мы попробуем их найти.

Прежде всего, оговорюсь о том, что существует как бы 2 варианта факторного анализа: «изучательный» факторный анализ и «подтвердительный» факторный анализ. Разница между ними в том, что в одном случае мы ищем факторы, а в другом случае у нас есть гипотеза о том, как выглядят факторы, и нам нужно подтвердить – действительно ли факторы такие, как мы их себе представляем. Далее мы будем говорить о первом варианте.

В нашем распоряжении есть анкеты 230 респондентов, чьи ответы не содержат пропусков, а каждый ответ это одно из значений ранговой шкалы от 1 до 6 (где 1 означает, что для человека данный мотив очень важен, а 6 – означает, что этим мотивом респондент руководствуется в последнюю очередь).

Вообще, методов факторизации существует довольно много, однако наиболее часто используются лишь некоторые из них. И среди таких часто используемых самым популярным является факторный анализ. Он – очень популярный, но не самый эффективный метод, прежде всего, потому, что он имеет целый ряд ограничений. Посмотрим на эти ограничения или требования этого метода.

Требования метода (или – ограничения) факторного анализа.

- Тип шкалы признаков – выше интервального.
- Форма связи между наблюдаемыми признаками – прямолинейная.
- Форма связи между признаками и факторами – прямолинейная.
- Требуется вращения для оптимизации интерпретации.

– Форма двумерного распределения любой пары наблюдаемых признаков – нормальная.

– Наблюдения независимы друг от друга.

Рассмотрим подробнее два самых серьёзных ограничения – первое и второе.

Тип шкалы признаков – выше интервального.

В принципе, факторный анализ работает и со шкалами псевдоинтервального типа, т.е. – с ранговыми с большим количеством рангов а также и с дихотомическими. Но такое применение факторного анализа, как считает часть специалистов в этой области, ни как теоретически не обосновано, а значит мы не застрахованы от получения в этом случае ненадежных результатов.

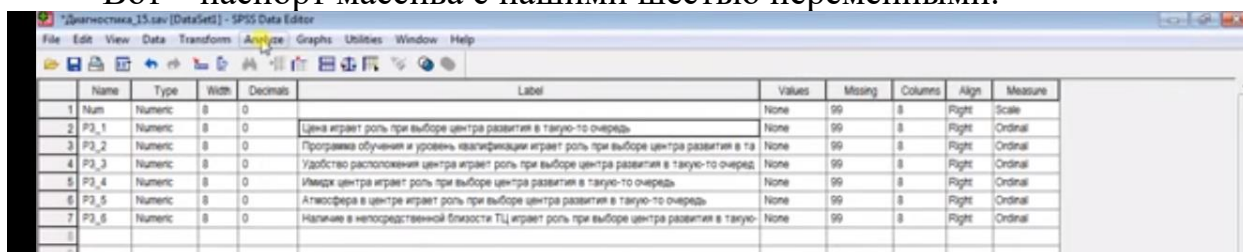
Форма связи между наблюдаемыми признаками – прямолинейная. Почему? А потому, что весь факторный анализ опирается на коэффициент корреляции Пирсона. Этот коэффициент предполагает как раз парность связи и линейность связи. Т.е. если между нашими 6-ю признаками имеется не ряд парных связей, а например ряд трёхмерных связей, то факторный анализ не сможет отследить такой вариант; или если связь не прямолинейная, то факторный анализ также покажет нам отсутствие связей и соответственно отсутствие возможностей искать латентную переменную.

Факторный анализ для лучшей интерпретации требует вращения. Помните, я смысл вращения когда-то сравнивал с наводкой бинокля на резкость? Так вот. О вращенье мы также поговорим в рамках этой лекции или связанного с ней семинара.

Остальные ограничения довольно трудно проверяемы, и о них мы поговорим в процессе рассмотрения материала.

А теперь давайте перейдем непосредственно к вопросам факторного анализа.

Вот – паспорт массива с нашими шестью переменными.



Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1 Num	Numeric	8	0		None	99	8	Right	Scale
2 P3_1	Numeric	8	0	Цена играет роль при выборе центра развития в такую-то очередь.	None	99	8	Right	Ordinal
3 P3_2	Numeric	8	0	Программа обучения и уровень квалификации играет роль при выборе центра развития в та	None	99	8	Right	Ordinal
4 P3_3	Numeric	8	0	Удобство расположения центра играет роль при выборе центра развития в такую-то очередь.	None	99	8	Right	Ordinal
5 P3_4	Numeric	8	0	Имидж центра играет роль при выборе центра развития в такую-то очередь.	None	99	8	Right	Ordinal
6 P3_5	Numeric	8	0	Атмосфера в центре играет роль при выборе центра развития в такую-то очередь.	None	99	8	Right	Ordinal
7 P3_6	Numeric	8	0	Наличие в непосредственной близости ТЦ играет роль при выборе центра развития в такую-	None	99	8	Right	Ordinal

Это – сами переменные со шкалами ответов:

Цена играет роль при выборе центра развития в такую-то очередь.

1	2	3	4	5	6
---	---	---	---	---	---

Программа обучения и уровень квалификации персонала играет роль при выборе центра развития в такую-то очередь.

1	2	3	4	5	6
---	---	---	---	---	---

Имидж центра развития играет роль при выборе центра развития в такую-то очередь.

1	2	3	4	5	6
---	---	---	---	---	---

Атмосфера в центре играет роль при выборе центра развития в такую-то очередь.

1	2	3	4	5	6
---	---	---	---	---	---

Наличие в непосредственной близости ТЦ играет роль при выборе центра развития в такую-то очередь.

1	2	3	4	5	6
---	---	---	---	---	---

Факторный анализ. Выбор вращения

Запустили факторный анализ в SPSS и переносим 6 переменных, интересующих нас в рабочее окно (Рис. 1), для чего, находясь на любом из листов редактора SPSS, идем по следующему пути:

- Кнопка командной строки «Анализ» --- > Снижение размерности --- > Факторный анализ

(Открывается панель «Факторный анализ», в активное окно которой переносим наши переменные)

- Нажимаем кнопку панели «Описательные» (или – «Diskriptives»), расположенную в левом нижнем углу панели (см. рис. 2)

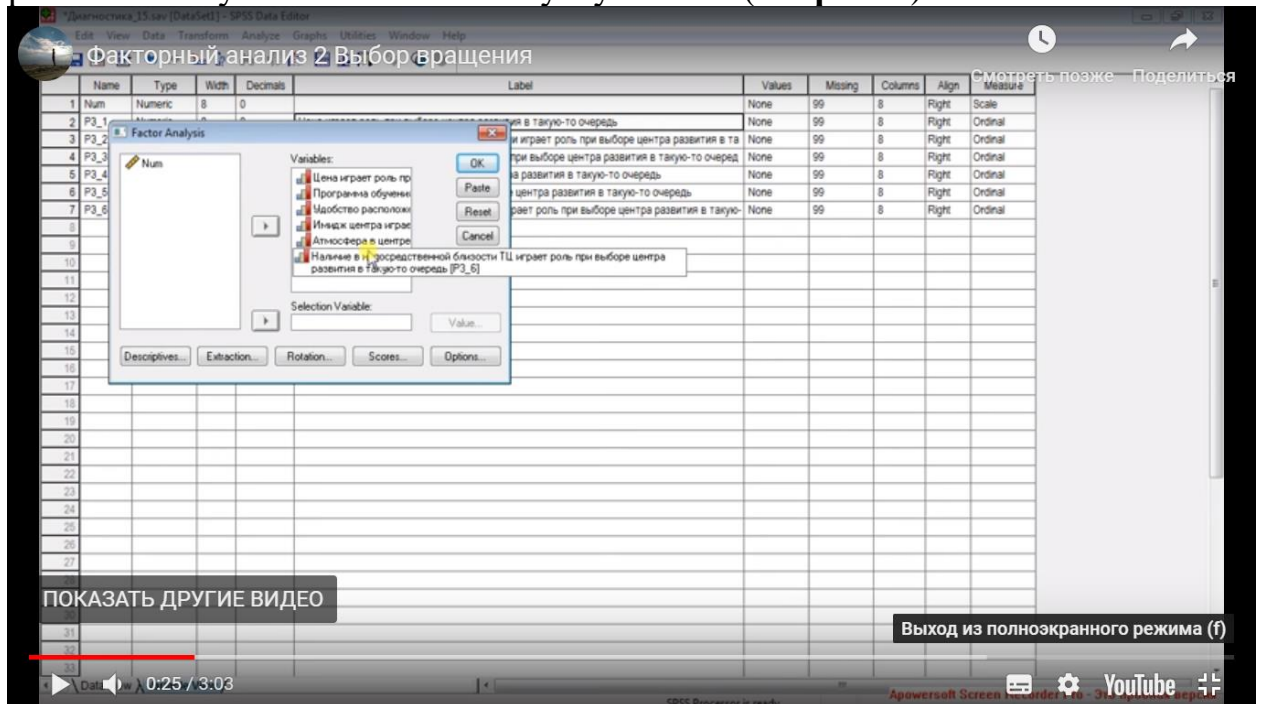


Рис. 1.

(В открывшейся вспомогательной панели «Факторный анализ: Описательные» просим вывести в окне просмотра среди результатов обсчета: Корреляционную матрицу, Уровень значений, а также тест КМО и тест Бартлетта, для чего:)

- В области панели «Corelation Matrik» ставим флажки активизации в ячейки (см. рис. 3):

- Coefficients (Корреляционная матрица)
- Sidegentfance levels(уровень значимости)
- RMO and Dartlett's test (КМО и тест Бартлетта)

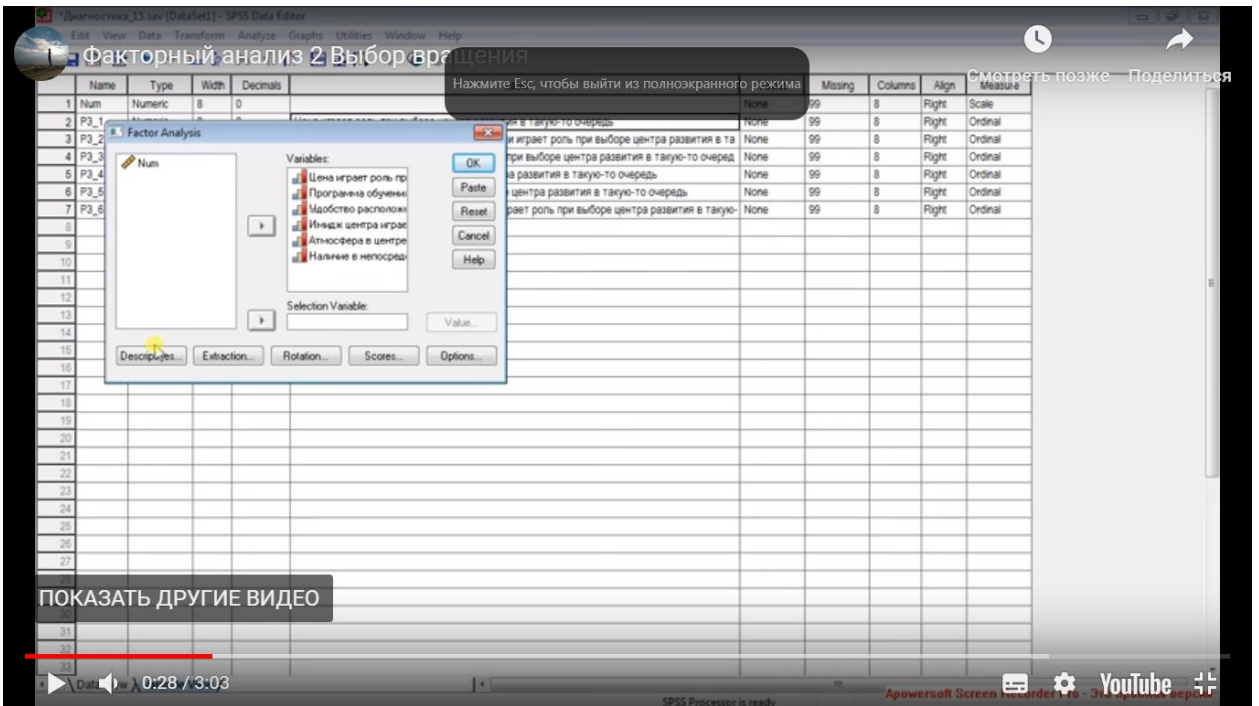


Рис. 2.

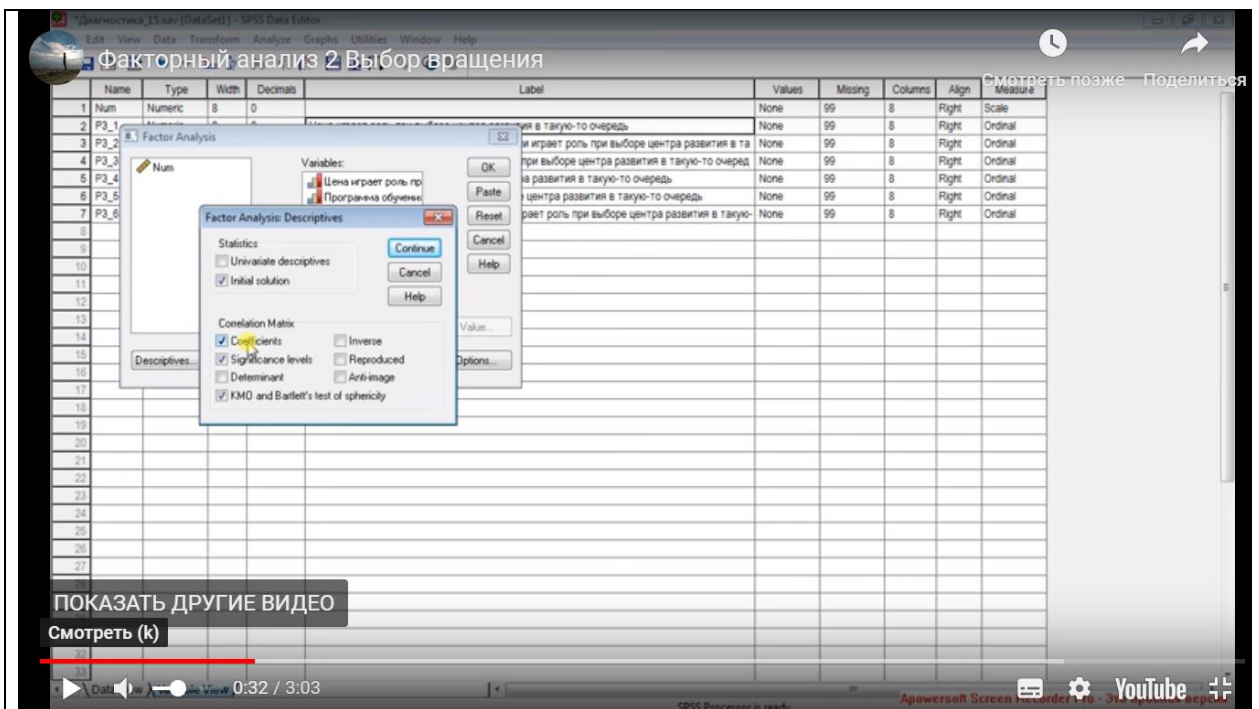


Рис. 3.

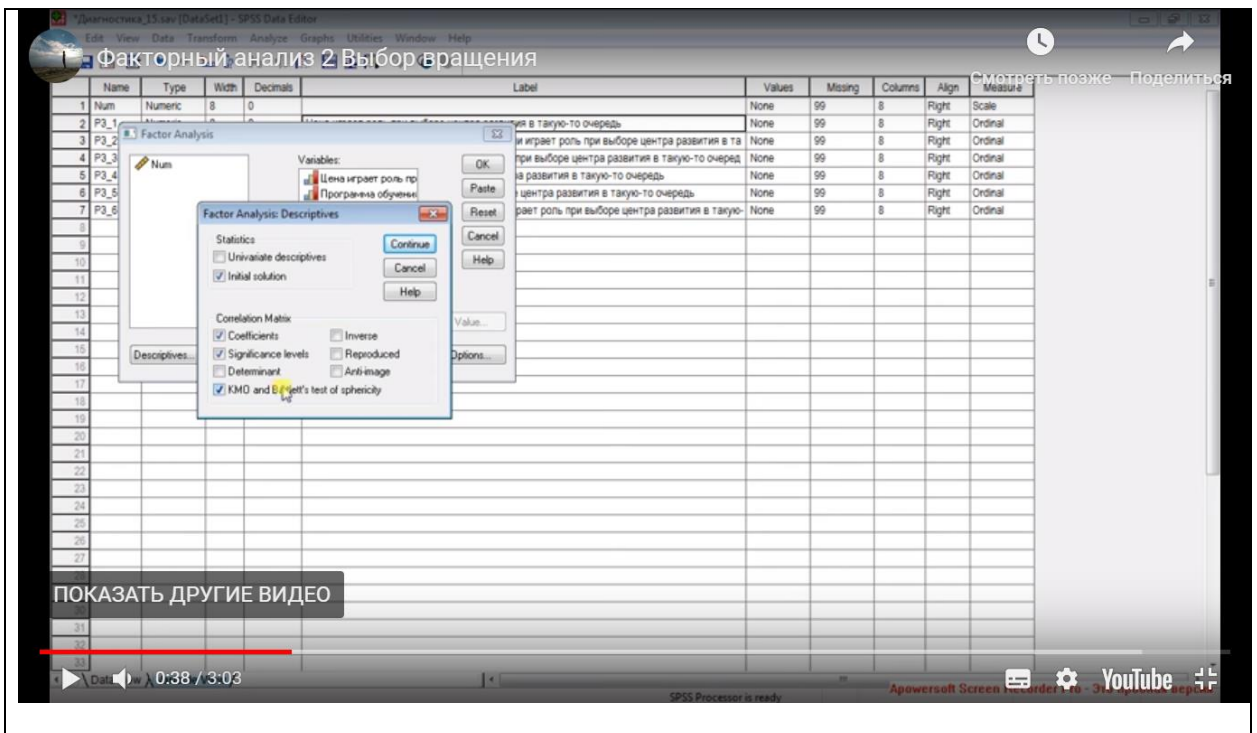


Рис. 4.

Нажимаем на вспомогательной панели кнопку –

- «Продолжить»

и на рабочей панели – кнопку

- «OK»

В результате, в окне просмотра получаем ряд таблиц.

Первая из которых Корреляционная матрица с уровнями значимости (рис. 5.1 и 5.2).

Корреляционная матрица с уровнями значимости

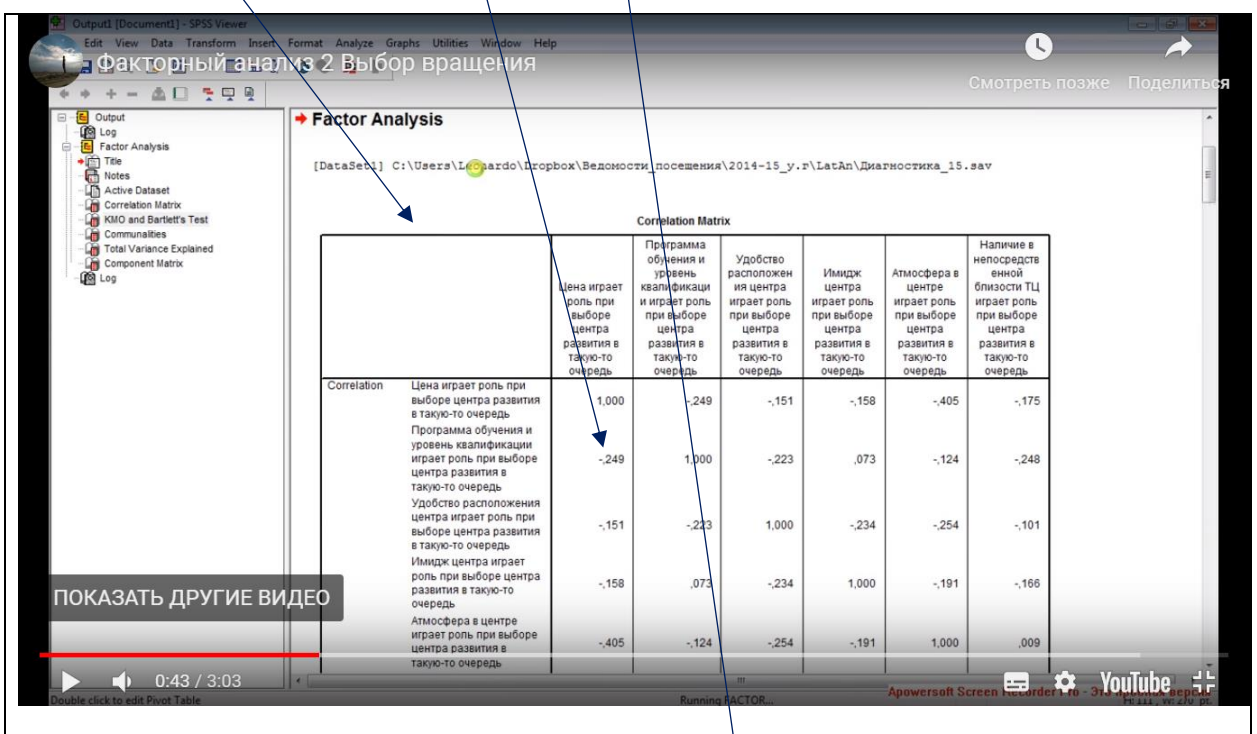


Рис. 5.1.

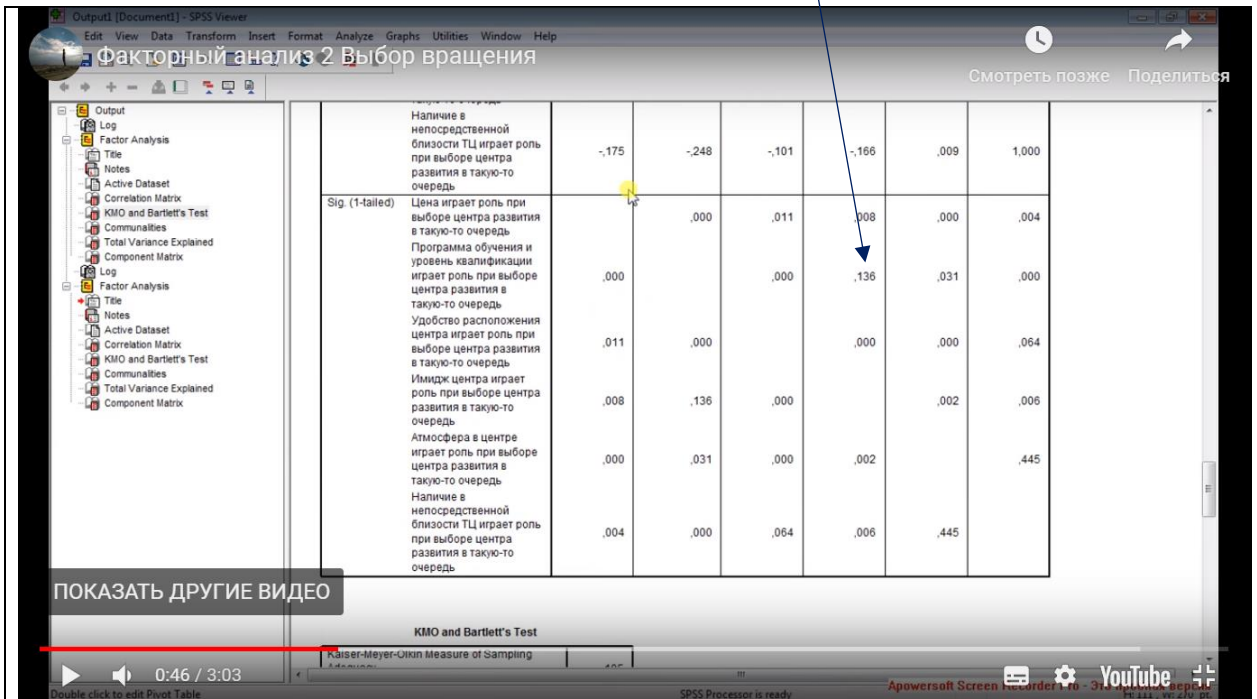


Рис. 5.2.

... А под ней – табличка с выводом значений КМО и теста Барлетта (рис. 6).

Табличка с выводом значений КМО и теста Барлетта

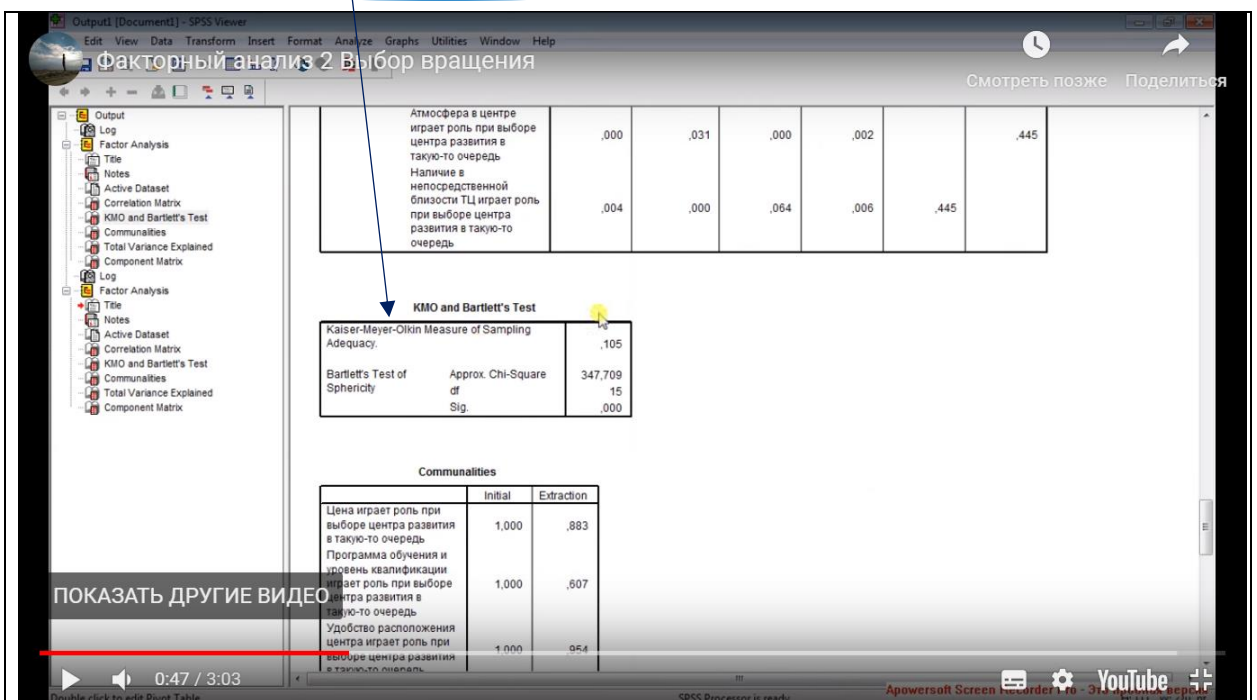


Рис. 6.

На **рис. 7** показаны значения этих коэффициентов. Тест КМО и тест Бартлетта позволяют предварительно оценить, насколько наша корреляционная матрица пригодна для того, чтобы в ней искать пучки парных линейных связей (о которых упоминалось в требованиях (см. выше). Принято считать, что значения теста КМО, – в случае пригодности матрицы, – должны быть больше 0,5. Тест же Бартлетта связан с гипотезой о сферичности данных. Уровень значимости гипотезы о сферичности данных должен быть меньше 0,05, чтобы эта гипотеза о сферичности могла быть отклонена. Сама же гипотеза о сферичности данных, если она выполняется, означает, что наши данные не подходят для использования для их обработки факторного анализа.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy:		.104	← тест КМО
Bartlett's Test of Sphericity	Approx. Chi-Square	347,700	
	df	15	
	Sig.	.000	← тест Бартлетта

Рис. 7.

На **рис. 8** представлена следующая таблица – «Общности» (или «Communalities»), которая показывает, какой процент (%) вариации каждой наблюдаемой переменной объясняет наша факторная модель.

О какой факторной модели идет речь?

	Initial	Extraction
Цена играет роль при выборе центра развития в такую-то очередь	1,000	,883
Программа обучения и уровень квалификации играет роль при выборе центра развития в такую-то очередь	1,000	,607
Удобство расположения центра играет роль при выборе центра развития в такую-то очередь	1,000	,954
Имидж центра играет роль при выборе центра развития в такую-то очередь	1,000	,438
Атмосфера в центре играет роль при выборе центра развития в такую-то очередь	1,000	,696
Наличие в непосредственной близости ТЦ играет роль при выборе центра развития в такую-то очередь	1,000	,469

Extraction Method: Principal Component Analysis.

Рис. 8.

Дело в том, что мы не сами задаем в данной ситуации число факторов, а мы оставляем по умолчанию критерий Кайзера; а по критерию Кайзера в модели остаются те факторы (или компоненты), которые объясняют больше одного собственного значения, то есть те, которые объясняют больше одной наблюдаемой переменной (в сумме у нас 6 переменных, т.е. 6 собственных значений) (рис. 9).

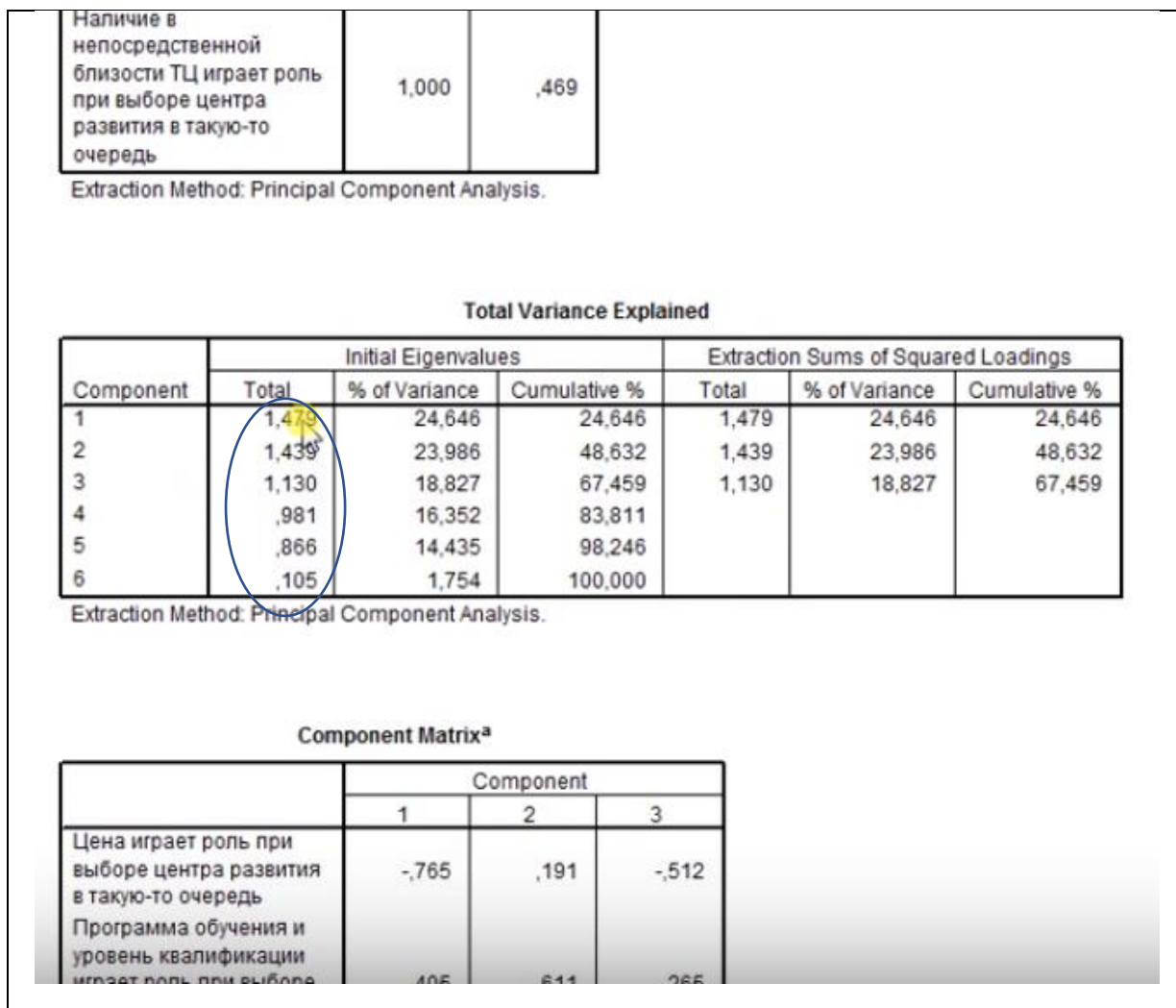
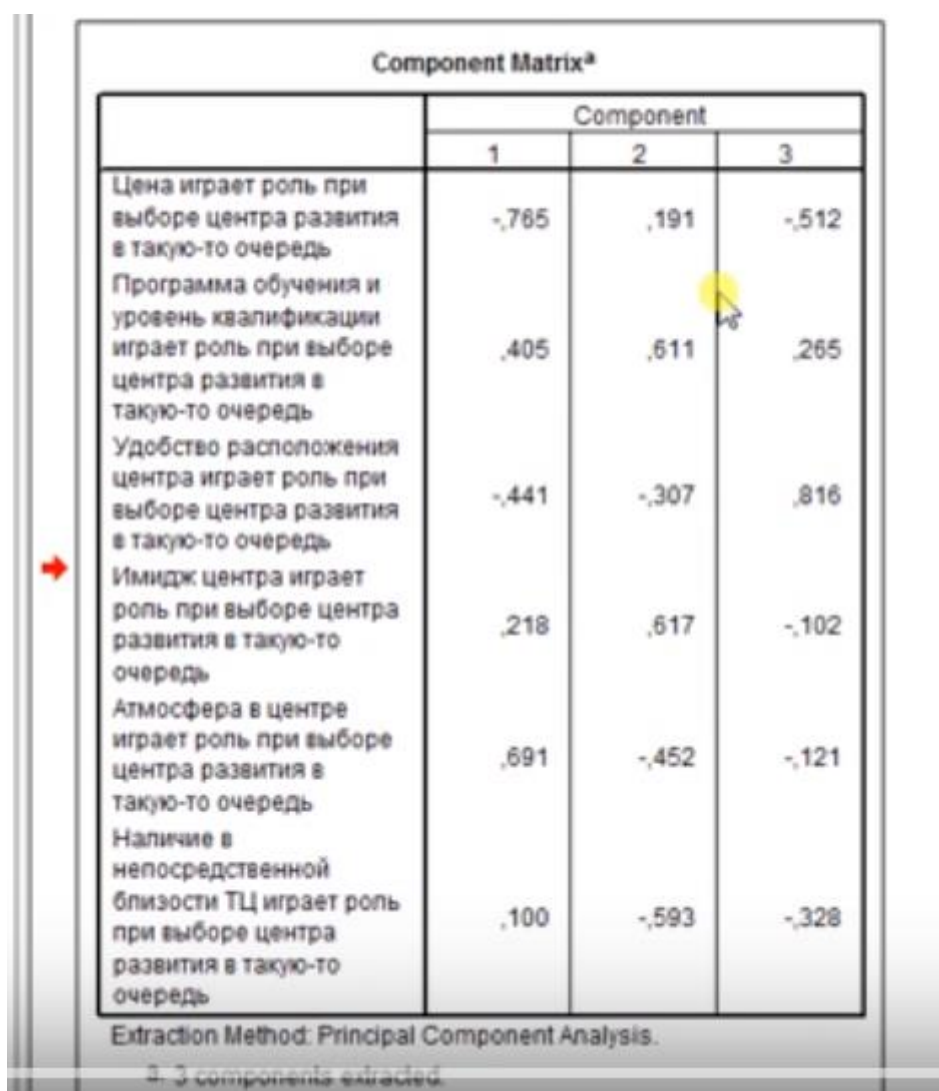


Рис. 9.

Опять же, 6 наших факторов (в колонке «**Total**» таблицы «**Total Variance Explained**» на рис. 9, сумма по вертикали должна равняться 6-ти, т.е.: $1,479+1,439+1,130+0,981+0,866+0,105=6$).

Но не все факторы одинаково хороши, т.к. 1-й фактор объясняет большую часть вариации наблюдаемых переменных (1,479), 2-й фактор (1,439) – поменьше, 3-й фактор (1,130) – ещё меньше, а 4-й, 5-й и 6-й факторы (соответственно: 0,981, 0,866 и 0,105) – в рассмотрение не берутся потому что каждый из них объясняет меньше 100% вариации (у нас всего 600% вариации потому что 6 переменных, каждая из которых занимает 100%; и вот каждый из трёх последних факторов объясняет менее 100%).

Дальше – переходим к следующей табличке, представленной в нашем окне просмотра. Это – матрица факторных нагрузок (рис. 10).



Component Matrix^a

	Component		
	1	2	3
Цена играет роль при выборе центра развития в такую-то очередь	-,765	,191	-,512
Программа обучения и уровень квалификации играет роль при выборе центра развития в такую-то очередь	,405	,611	,265
Удобство расположения центра играет роль при выборе центра развития в такую-то очередь	-,441	-,307	,816
Имидж центра играет роль при выборе центра развития в такую-то очередь	,218	,617	-,102
Атмосфера в центре играет роль при выборе центра развития в такую-то очередь	,691	-,452	-,121
Наличие в непосредственной близости ТЦ играет роль при выборе центра развития в такую-то очередь	,100	-,593	-,328

Extraction Method: Principal Component Analysis.
3. 3 components extracted.

Рис. 10.

Матрица факторных нагрузок показывает на какую величину надо домножить каждую компоненту, чтобы получить значение данной наблюдаемой переменной для данного респондента. *А сами же факторные нагрузки следует понимать как корреляционные коэффициенты между переменными и факторами.* Попутно, перенесем в Excel нашу корреляционную матрицу и матрицу нагрузок (рис. 11 и 12).

Корреляционная матрица и матрица нагрузок: вид в окне просмотра SPSS

Correlation Matrix

	Цена играет роль при выборе центра развития в такую-то очередь	Программа обучения и уровень квалификации играет роль при выборе центра развития в такую-то очередь	Удобство расположения центра играет роль при выборе центра развития в такую-то очередь	Имидж центра играет роль при выборе центра развития в такую-то очередь	Атмосфера в центре играет роль при выборе центра развития в такую-то очередь	Наличие в непосредственной близости ТЦ играет роль при выборе центра развития в такую-то очередь
Correlation	1,000	-.249	-.151	-.158	-.405	-.175
Цена играет роль при выборе центра развития в такую-то очередь	1,000					
Программа обучения и уровень квалификации играет роль при выборе центра развития в такую-то очередь	-.249	1,000	-.223	,073	-.124	-.248
Удобство расположения центра играет роль при выборе центра развития в такую-то очередь	-.151	-.223	1,000	-.234	-.254	-.101
Имидж центра играет роль при выборе центра развития в такую-то очередь	-.158	,073	-.234	1,000	-.191	-.166
Атмосфера в центре играет роль при выборе центра развития в такую-то очередь	-.405	-.124	-.254	-.191	1,000	,009
Наличие в непосредственной близости ТЦ играет роль при выборе центра развития в такую-то очередь	-.175	-.248	-.101	-.166	,009	1,000

Рис. 11.

Перенос корреляционной матрицы и матрицы нагрузок в среду Excel

Факторный анализ 3. Выбор вращения

Total Variance Explained										Correlation Matrix					
Component	Initial Eigenvalues	Extraction Sums of Squared Loadings			Rotation Sums of Squared Multiple Correlations			P3_1	P3_2	P3_3	P3_4	P3_5	P3_6		
Total	% of Variance	Cumulative Total	% of Variance	Cumulative Total	Total	% of Variance	Cumulative Total	Correlation	P3_1	P3_2	P3_3	P3_4	P3_5	P3_6	
1	1,479	24,646	24,646	1,479	24,646	24,646	1,457	P3_1	1	-0,249	-0,151	-0,158	-0,405	-0,175	
2	1,439	23,986	48,632	1,439	23,986	48,632	1,414	P3_2	-0,249	1	-0,223	0,073	-0,124	-0,248	
3	1,13	18,827	67,459	1,13	18,827	67,459	1,281	P3_3	-0,151	-0,223	1	-0,234	-0,254	-0,101	
4	0,981	16,352	83,811					P3_4	-0,158	0,073	-0,234	1	-0,191	-0,166	
5	0,866	14,435	98,246					P3_5	-0,405	-0,124	-0,254	-0,191	1	0,009	
6	0,105	1,754	100					P3_6	-0,175	-0,248	-0,101	-0,166	0,009	1	
Extraction Method: Principal Component Analysis.								Sig. (1-tail)	P3_1		0	0,011	0,008	0	0,004
When components are correlated, sums of squared loadings cannot be added to obtain a total variance.								P3_2		0		0	0,136	0,031	0
								P3_3		0,011	0		0	0	0,064
								P3_4		0,008	0,136	0		0,002	0,006
								P3_5		0	0,031	0	0,002		0,445
								P3_6		0,004	0	0,064	0,006	0,445	
Component Matrix(a)															
Fj*L=Xi	Component														
	1	2	3												
P3_1	-0,765	0,191	-0,512												
P3_2	0,405	0,611	0,265												
P3_3	-0,441	-0,307	0,816												
P3_4	0,218	0,617	-0,102												
P3_5	0,691	-0,452	-0,121												
P3_6	0,1	-0,593	-0,328												
Extraction Method: Principal Component Analysis.															
3 components extracted.															

Обозначения: P3_1–Цена...; P3_2–Программа обучения.; P3_3–Удобство расположения; P3_4–Имидж ТЦ...; P3_5–Атмосфера...; P3_6–Наличие поблизости ТЦ...

Рис. 12.

Предварительно интерпретируем матрицу нагрузок (по строкам наблюдаемые переменные, по столбцам – компоненты).

Обычно смотрят по строкам и определяют – к какой компоненте ближе всего каждая наблюдаемая переменная.

Мы видим, что максимальная нагрузка для 1-й переменной (P3_1) находится под 1-й компонентой (-0,765) и под 2-й компонентой (-0,512). Собственно говоря, достаточно выявить в каждой строке матрицы нагрузок всего одну, максимальную по значению нагрузку.

Значит, для 1-й переменной это будет нагрузка (-0,765, соответствующая 1-й компоненте (т.е. 1-я переменная наиболее близка 1-й компоненте). Этой же компоненте наиболее близка и 5-я переменная (0,691). Они помечены розовым цветом.

Второй компоненте наиболее близки 2-я, 4-я и 6-я переменные с соответствующими значениями: 0,611, 0,617 и -0,593.

Третьей компоненте наиболее близки 3-я переменные (0,816).

Далее, нам надо понять: будем ли мы вращать нашу матрицу, и если будем, то – прямоугольно (т.е. ортогонально, сохраняя угол между осями) или косоугольно.

Есть разные основания для принятия решения на этот счет. Но если у нас отсутствует теоретическая рамка, а в данном случае она конечно же отсутствует (здесь имеется в виду отсутствие каких-либо теорий на тот счет, какими именно мотивами руководствуются родители при выборе того, в какой развлекательный центр детей вести), то поэтому нам приходится руководствоваться природой тех данных, которыми мы располагаем. Для этого мы обращаемся к корреляционной матрице.

Для удобства, эту матрицу мы также разместили цветами принадлежность признака каждой из трёх компонент (столбей 1-й слева).

В этой матрице верхние 6 строк отведены коэффициентам корреляции, а нижние 6 строк – отведены уровню значимости для этих коэф-тов (рис. 13).

J	K	L	M	N	O	P	Q	
Correlation Matrix								
		P3_1	P3_2	P3_3	P3_4	P3_5	P3_6	
Correlatio	P3_1	1	-0,249	-0,151	-0,158	-0,405	-0,175	
	P3_2	-0,249	1	-0,223	0,073	-0,124	-0,248	
	P3_3	-0,151	-0,223	1	-0,234	-0,254	-0,101	
	P3_4	-0,158	0,073	-0,234	1	-0,191	-0,166	
	P3_5	-0,405	-0,124	-0,254	-0,191	1	0,009	
	P3_6	-0,175	-0,248	-0,101	-0,166	0,009	1	
Sig. (1-tail	P3_1		+	0	0,011	0,008	0	0,004
	P3_2	0		0	0,136	0,031	0	
tain a tota	P3_3	0,011	0		0	0	0,064	
	P3_4	0,008	0,136	0		0,002	0,006	
	P3_5	0	0,031	0	0,002		0,445	
	P3_6	0,004	0	0,064	0,006	0,445		

Рис.13.

Каждая компонента – это как бы группа наблюдаемых переменных. И наша задача на данный момент состоит в том, чтобы понять, значимые связи существуют только внутри групп наблюдаемых переменных, или связи имеются и между группами. Как это понять (точнее – как это формализовать): имеются ли корреляции только внутри групп переменных (а каждая группа наблюдаемых переменных внутри соответствующей компоненты выделена у нас цветом: внутри 1-й компоненты группа состоит из 2-х переменных, внутри 2-й компоненты группа состоит из 3-х переменных, и внутри 3-й компоненты группа состоит из 1-й переменной), или имеются корреляции и между группами. Собственно говоря, мы это понять можем по данной корреляционной матрице. Но чтоб было удобнее, с ней работать, преобразуем её: копируем (см. **рис.14.1**), и заменяем данные на абсолютные значения (т.е. избавляемся от минусов (см. **рис. 14.2**).

Преобразование корреляционной матрицы: шаг 1 (копирование)

J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
Correlation Matrix															
		P3_1	P3_2	P3_3	P3_4	P3_5	P3_6			P3_1	P3_2	P3_3	P3_4	P3_5	P3_6
Correlatio	P3_1	1	-0,249	-0,151	-0,158	-0,405	-0,175	Correlatio	P3_1	1	-0,249	-0,151	-0,158	-0,405	-0,175
	P3_2	-0,249	1	-0,223	0,073	-0,124	-0,248		P3_2	-0,249	1	-0,223	0,073	-0,124	-0,248
	P3_3	-0,151	-0,223	1	-0,234	-0,254	-0,101		P3_3	-0,151	-0,223	1	-0,234	-0,254	-0,101
	P3_4	-0,158	0,073	-0,234	1	-0,191	-0,166		P3_4	-0,158	0,073	-0,234	1	-0,191	-0,166
	P3_5	-0,405	-0,124	-0,254	-0,191	1	0,009		P3_5	-0,405	-0,124	-0,254	-0,191	1	0,009
	P3_6	-0,175	-0,248	-0,101	-0,166	0,009	1		P3_6	-0,175	-0,248	-0,101	-0,166	0,009	1
Sig. (1-tail	P3_1		0	0,011	0,008	0	0,004	Sig. (1-tail	P3_1		0	0,011	0,008	0	0,004
	P3_2	0		0	0,136	0,031	0		P3_2	0		0	0,136	0,031	0
tain a tota	P3_3	0,011	0		0	0	0,064	tain a tota	P3_3	0,011	0		0	0	0,064
	P3_4	0,008	0,136	0		0,002	0,006		P3_4	0,008	0,136	0		0,002	0,006
	P3_5	0	0,031	0	0,002		0,445		P3_5	0	0,031	0	0,002		0,445
	P3_6	0,004	0	0,064	0,006	0,445			P3_6	0,004	0	0,064	0,006	0,445	

Рис. 14.1.

Это делаем потому, что нам интересно величина связи по модулю, т.е. – на сколько сильная связь.

Преобразование корреляционной матрицы: шаг 2 (замена на абсолютные значения)

J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
Correlation Matrix															
		P3_1	P3_2	P3_3	P3_4	P3_5	P3_6			P3_1	P3_2	P3_3	P3_4	P3_5	P3_6
Correlatio	P3_1	1	-0,249	-0,151	-0,158	-0,405	-0,175	Correlatio	P3_1	1	0,249	0,151	0,158	0,405	0,175
	P3_2	-0,249	1	-0,223	0,073	-0,124	-0,248		P3_2	0,249	1	0,223	0,073	0,124	0,248
	P3_3	-0,151	-0,223	1	-0,234	-0,254	-0,101		P3_3	0,151	0,223	1	0,234	0,254	0,101
	P3_4	-0,158	0,073	-0,234	1	-0,191	-0,166		P3_4	0,158	0,073	0,234	1	0,191	0,166
	P3_5	-0,405	-0,124	-0,254	-0,191	1	0,009		P3_5	0,405	0,124	0,254	0,191	1	0,009
	P3_6	-0,175	-0,248	-0,101	-0,166	0,009	1		P3_6	0,175	0,248	0,101	0,166	0,009	1
Sig. (1-tail	P3_1		0	0,011	0,008	0	0,004	Sig. (1-tail	P3_1		0	0,011	0,008	0	0,004
	P3_2	0		0	0,136	0,031	0		P3_2	0		0	0,136	0,031	0
tain a tota	P3_3	0,011	0		0	0	0,064	tain a tota	P3_3	0,011	0		0	0	0,064
	P3_4	0,008	0,136	0		0,002	0,006		P3_4	0,008	0,136	0		0,002	0,006
	P3_5	0	0,031	0	0,002		0,445		P3_5	0	0,031	0	0,002		0,445
	P3_6	0,004	0	0,064	0,006	0,445			P3_6	0,004	0	0,064	0,006	0,445	

Рис. 14.2.

Пока не смотрим на уровень значимости (на нижние 6 строк), а анализируем верх таблицы (**рис. 14.2**). Пока что, мы можем видеть, что очень уж сильных, тесных связей у нас здесь нет. Связь у нас изменяется от 0,15 (более-менее значимая связь) до 0,4. Пометим через условное форматирование

только те коэффициенты корреляции, которые больше (ну, допустим, двух) (см. рис. 15-21).

	P3_1	P3_2	P3_3	P3_4	P3_5	P3_6
P3_1	1	0,249	0,151	0,158	0,405	0,175
P3_2	0,249	1	0,223	0,073	0,124	0,248
P3_3	0,151	0,223	1	0,234	0,254	0,101
P3_4	0,158	0,073	0,234	1	0,191	0,166
P3_5	0,405	0,124	0,254	0,191	1	0,009
P3_6	0,175	0,248	0,101	0,166	0,009	1
P3_1		0	0,011	0,008	0	0,004
P3_2	0		0	0,136	0,031	0
P3_3	0,011	0		0	0	0,064
P3_4	0,008	0,136	0		0,002	0,006
P3_5	0	0,031	0	0,002		0,445
P3_6	0,004	0	0,064	0,006	0,445	

Рис. 15.

The screenshot shows the 'Условное форматирование' (Conditional Formatting) menu in Microsoft Excel. The menu is open, displaying several options: 'Правила выделения ячеек', 'Правила отбора первых и последних значений', 'Гистограммы', 'Цветовые шкалы', and 'Наборы значков'. Below these are options to 'Создать правило...', 'Удалить правила', and 'Управление правилами...'. The background shows a portion of the data table from Figure 15, with columns W, X, Y, P3_4, P3_5, and P3_6 visible. The table contains numerical values representing correlation coefficients.

Рис. 16.

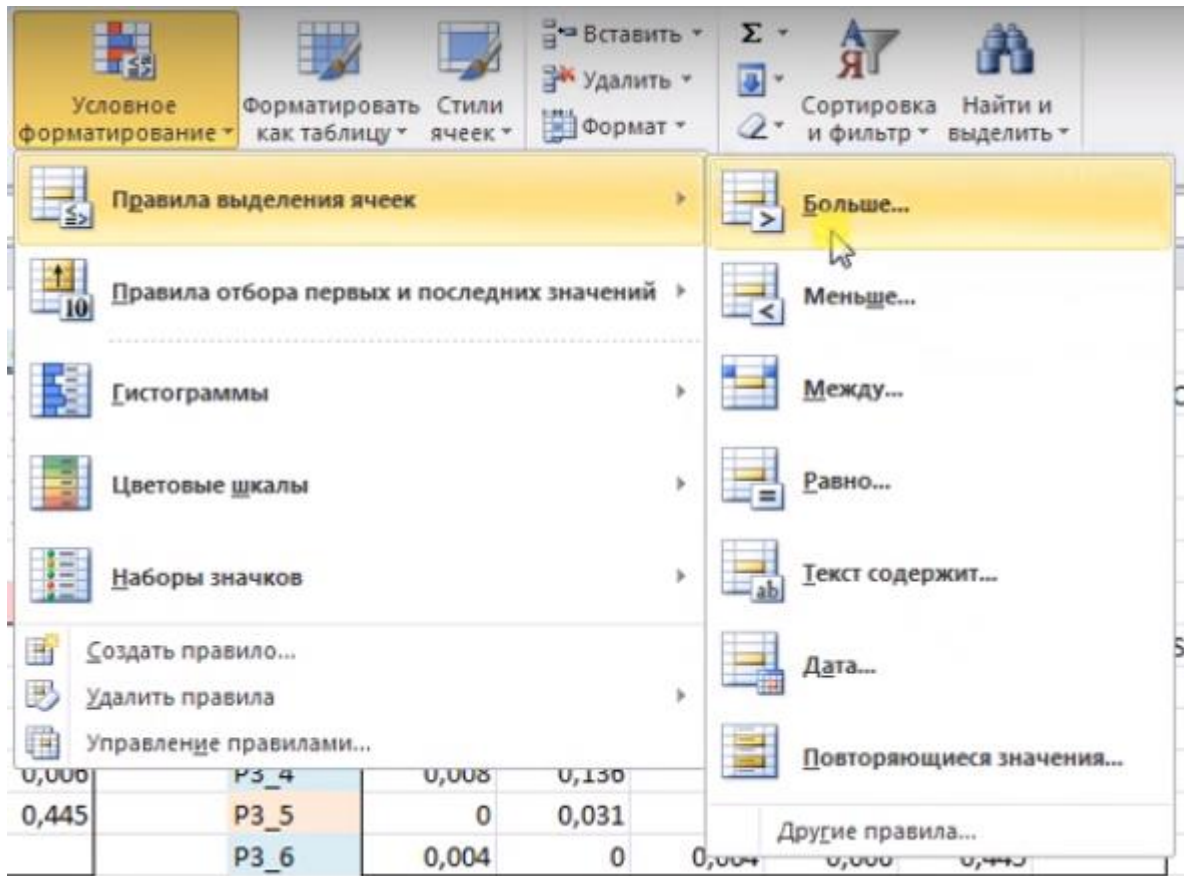


Рис. 17.

The image shows a spreadsheet with two matrices and a dialog box. The first matrix is a correlation matrix, and the second is a significance matrix (Sig. (1-tail)). The dialog box 'Больше' (More) is open, showing a threshold of 0.5045 and a color scheme of light red fill and dark red text.

	P3_1	P3_2	P3_3	P3_4	P3_5	P3_6
P3_1	1	-0,249	-0,151	-0,158	-0,405	-0,175
P3_2	-0,249	1	-0,223	0,073	-0,124	-0,248
P3_3	-0,151	-0,223	1	-0,234	-0,254	-0,101
P3_4	-0,158	0,073	-0,234	1	-0,191	-0,166
P3_5	-0,405	-0,124	-0,254	-0,191	1	0,009
P3_6	-0,175	-0,248	-0,101	-0,166	0,009	1

	P3_1	P3_2	P3_3	P3_4	P3_5	P3_6
P3_1	1	0,249	0,151	0,158	0,405	0,175
P3_2	0,249	1	0,223	0,073	0,124	0,248
P3_3	0,151	0,223	1	0,234	0,254	0,101
P3_4	0,158	0,073	0,234	1	0,191	0,166
P3_5	0,405	0,124	0,254	0,191	1	0,009
P3_6	0,175	0,248	0,101	0,166	0,009	1

	P3_1	P3_2	P3_3	P3_4	P3_5	P3_6
P3_1		0	0,011	0,008	0	0,004
P3_2	0		0	0,136	0,031	0
P3_3	0,011	0		0	0	0,064
P3_4	0,008	0,136	0		0,002	0,006
P3_5	0	0,031	0	0,002		0,445
P3_6	0,004	0	0,064	0,006	0,445	

Рис. 18.

	P3_1	P3_2	P3_3	P3_4	P3_5	P3_6			P3_1	P3_2	P3_3	P3_4	P3_5	P3_6
P3_1	1	-0,249	-0,151	-0,158	-0,405	-0,175	Correlatio	P3_1	1	0,249	0,151	0,158	0,405	0,175
P3_2	-0,249	1	-0,223	0,073	-0,124	-0,248		P3_2	0,249	1	0,223	0,073	0,124	0,248
P3_3	-0,151	-0,223	1	-0,234	-0,254	-0,101		P3_3	0,151	0,223	1	0,234	0,254	0,101
P3_4	-0,158	0,073	-0,234	1	-0,191	-0,166		P3_4	0,158	0,073	0,234	1	0,191	0,166
P3_5	-0,405	-0,124	-0,254	-0,191	1	0,009		P3_5	0,405	0,124	0,254	0,191	1	0,009
P3_6	-0,175	-0,248	-0,101	-0,166	0,009	1		P3_6	0,175	0,248	0,101	0,166	0,009	1
P3_1		0	0,011	0,008	0	0,004	Sig. (1-tail	P3_1		0	0,011	0,008	0	0,004
P3_2	0		0	0,136	0,031	0		P3_2	0		0	0,136	0,031	0
P3_3	0,011	0		0	0	0,064		P3_3	0,011	0		0	0	0,064
P3_4	0,008	0,136	0		0,002	0,006		P3_4	0,008	0,136	0		0,002	0,006
									0	0,031	0	0,002		0,445
									0,004	0	0,064	0,006	0,445	

Рис. 19.

	P3_1	P3_2	P3_3	P3_4	P3_5	P3_6			P3_1	P3_2	P3_3	P3_4	P3_5	P3_6
P3_1	1	-0,249	-0,151	-0,158	-0,405	-0,175	Correlatio	P3_1	1	0,249	0,151	0,158	0,405	0,175
P3_2	-0,249	1	-0,223	0,073	-0,124	-0,248		P3_2	0,249	1	0,223	0,073	0,124	0,248
P3_3	-0,151	-0,223	1	-0,234	-0,254	-0,101		P3_3	0,151	0,223	1	0,234	0,254	0,101
P3_4	-0,158	0,073	-0,234	1	-0,191	-0,166		P3_4	0,158	0,073	0,234	1	0,191	0,166
P3_5	-0,405	-0,124	-0,254	-0,191	1	0,009		P3_5	0,405	0,124	0,254	0,191	1	0,009
P3_6	-0,175	-0,248	-0,101	-0,166	0,009	1		P3_6	0,175	0,248	0,101	0,166	0,009	1
P3_1		0	0,011	0,008	0	0,004	Sig. (1-tail	P3_1		0	0,011	0,008	0	0,004
P3_2	0		0	0,136	0,031	0		P3_2	0		0	0,136	0,031	0
P3_3	0,011	0		0	0	0,064		P3_3	0,011	0		0	0	0,064
P3_4	0,008	0,136	0		0,002	0,006		P3_4	0,008	0,136	0		0,002	0,006
									0	0,031	0	0,002		0,445
									0,004	0	0,064	0,006	0,445	

Рис. 20.

	R	S	T	U	V	W	X	Y
			P3_1	P3_2	P3_3	P3_4	P3_5	P3_6
Correlatio	P3_1		1	0,249	0,151	0,158	0,405	0,175
	P3_2		0,249	1	0,223	0,073	0,124	0,248
	P3_3		0,151	0,223	1	0,234	0,254	0,101
	P3_4		0,158	0,073	0,234	1	0,191	0,166
	P3_5		0,405	0,124	0,254	0,191	1	0,009
	P3_6		0,175	0,248	0,101	0,166	0,009	1
Sig. (1-tail	P3_1			0	0,011	0,008	0	0,004
	P3_2			0	0	0,136	0,031	0
	P3_3			0,011	0	0	0	0,064
	P3_4			0,008	0,136	0	0,002	0,006
	P3_5			0	0,031	0	0,002	0,445
	P3_6			0,004	0	0,064	0,006	0,445

Рис. 21.

Для удобства, уберем из верхней части таблицы все дублирующие значения и все неподсвеченные нами (т.е. – меньше 0,2) (см. **рис. 22**).

S	T	U	V	W	X	Y
	P3_1	P3_2	P3_3	P3_4	P3_5	P3_6
P3_1		0,249			0,405	
P3_2			0,223			0,248
P3_3				0,234	0,254	
P3_4						
P3_5						
P3_6						
P3_1		0	0,011	0,008	0	0,004
P3_2	0		0	0,136	0,031	0
P3_3	0,011	0		0	0	0,064
P3_4	0,008	0,136	0		0,002	0,006
P3_5	0	0,031	0	0,002		0,445
P3_6	0,004	0	0,064	0,006	0,445	

Рис. 22.

Итак, (см. **рис. 22**) мы имеем 6 ячеек, в которых наблюдается достаточно (относительно) сильная связь. Посмотрим, есть ли среди этих 6-ти ячеек те, которые отвечают за связь не внутри факторов (компонент) а между факторами (компонентами), т.е. между наблюдаемыми переменными, которые формируют факторы. Потому что, принимая решение о том, какое вращение нам нужно – косоугольное или ортогональное, если мы не имеем каких-то теоретических оснований, нам, как было сказано выше, приходится ориентироваться на саму природу данных, а природа наших данных закладывается природой наших наблюдаемых переменных. И если связаны между собой переменные, которые относятся к разным факторам, значит есть основания говорить о том (предполагать), что и между факторами, формируемыми этими переменными есть связь. ***А если есть связь, то значит нужно косоугольное вращение.***

Итак, анализируя верх таблицы (**рис. 22**), мы сразу видим, что ячейка U3 (со значением 0,249) лежит на пересечении первого и второго факторов. На пересечении лежат также ячейки V3 (со значением 0,223), W5 (0,234) и X5 (0,254). Пометим соответствующие значения жирным шрифтом красного цвета (**рис. 23.1, 23.2**)

W5		=ABS(O5)																				
	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y				
1	Correlation Matrix																					
2	Rotation Sums of Squ			P3_1	P3_2	P3_3	P3_4	P3_5	P3_6													
3	Total	Correlatio			P3_1	1	-0,249	-0,151	-0,158	-0,405	-0,175	Correlatio			P3_1		0,249		0,405			
4	1,457		P3_2	-0,249	1	-0,223	0,073	-0,124	-0,248			P3_2			0,223			0,248				
5	1,414		P3_3	-0,151	-0,223	1	-0,234	-0,254	-0,101			P3_3				0,234		0,254				
6	1,281		P3_4	-0,158	0,073	-0,234	1	-0,191	-0,166			P3_4										
7			P3_5	-0,405	-0,124	-0,254	-0,191	1	0,009			P3_5										
8			P3_6	-0,175	-0,248	-0,101	-0,166	0,009	1			P3_6										
9			Sig. (1-tail			P3_1		0	0,011	0,008	0	0,004	Sig. (1-tail			P3_1		0	0,011	0,008	0	0,004
10			P3_2		0		0	0,136	0,031	0		P3_2		0		0	0,136	0,031	0			
11	cannot be added to obtain a tota			P3_3		0,011	0		0	0	0,064	P3_3		0,011	0		0	0	0,064			
12			P3_4		0,008	0,136	0		0,002	0,006		P3_4		0,008	0,136	0		0,002	0,006			
13			P3_5		0	0,031	0	0,002		0,445		P3_5		0	0,031	0	0,002		0,445			
14			P3_6		0,004	0	0,064	0,006	0,445			P3_6		0,004	0	0,064	0,006	0,445				

Рис. 23.1.

Окончательный

	S	T	U	V	W	X	Y
		P3_1	P3_2	P3_3	P3_4	P3_5	P3_6
P3_1			0,249			0,405	
P3_2				0,223			0,248
P3_3					0,234	0,254	
P3_4							
P3_5							
P3_6							
P3_1			0	0,011	0,008	0	0,004
P3_2		0		0	0,136	0,031	0
P3_3		0,011	0		0	0	0,064
P3_4		0,008	0,136	0		0,002	0,006
P3_5		0	0,031	0	0,002		0,445
P3_6		0,004	0	0,064	0,006	0,445	

Рис. 23.2.

Итак, мы видим, что из 6-ти ячеек с наиболее сильной связью 4 относятся к межфакторным связям. Поэтому, у нас есть основания предполагать, что факторы должны быть скоррелированы. Но – скоррелированы не сильно: на уровне 0,2 – 0,25.

Возвращаемся к факторному анализу (открываем опять ранее свернутые страницы редактора SPSS).

Снова запускаем его, на этот раз заходим, нажав на вспомогательной панели кнопку «Rotation» («Вращение»), заходим в эту опцию, и выбираем из двух вариантов косоугольного вращения (рис. 24).

Команда выбора вида вращения

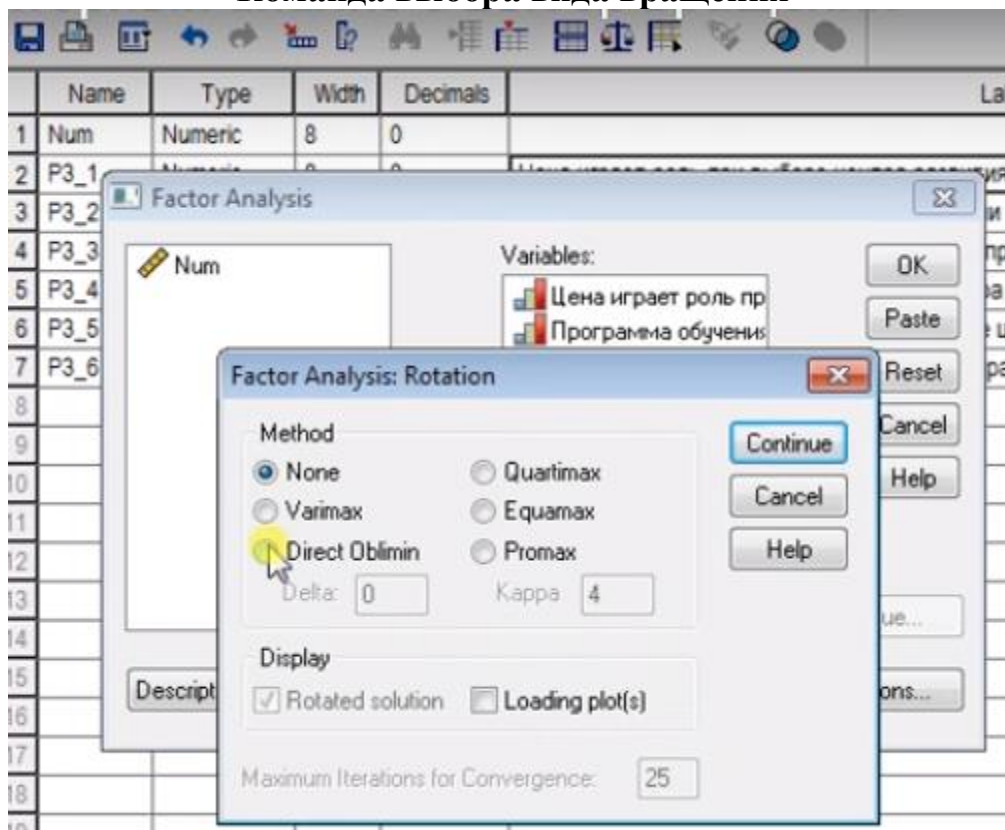


Рис. 24.

Есть вариант «Direct Oblimin», и он регулируется при помощи Delta («Дельты»). «Дельта» изменяется от -9999 до 0,8. Значение 0,8 предполагает, что мы устанавливаем максимально возможную скоррелированность для наших факторов; отрицательное же значение предполагает, что мы устанавливаем очень маленькую скоррелированность. На самом деле, этот промежуток от -9999 до 0 – он не сильно влияет на уменьшение скоррелированности. Поэтому, если Вы хотите с помощью «Директ Облимина» задать небольшую скоррелированность, то Вы можете оставлять значение «0», которое здесь уже и так стоит по умолчанию и можете запускать; если Вы хотите немножко уменьшить скоррелированность, то соответственно немножко уменьшаете это значение до минус 10-ти.

Есть и другой вариант – Promax («Про Макс»). Почему он хороший? Потому, что он, как и «Варимакс» для ортогонального вращения, максимизирует контрастность по строкам. Вот, например, в неповёрнутой матрице, которую мы только что рассматривали, у нас не очень контрастная картина по ряду строк (по 1-й строке, по 2-й строке). В идеале, вся нагрузка должна бы прийти на какой-то один из факторов (а для лучшей интерпретации, большинству аналитиков лучше, удобнее видеть контрастные строчки (например: -1 и 0 и 0 – вот это контрастность!!!). Поэтому, в случае косоугольного вращения для нас удобнее «Промакс», а в случае ортогонального вращения для нас удобнее «Варимакс». Если же кому-то

удобнее получать результаты контрастные по столбцам, то значит надо использовать «Квартимакс». А если кому-то хочется получить результат компромиссный – чтобы была контрастность и по строчкам и по столбцам, то надо использовать, то он выбирает «Эквимакс».

Мы же сейчас работаем с «Промакс».

«Карра» (по-русски звучит как «Каппа») в этом случае (см. ячейку во вспомогательной панели «Карра») изменяется от 0 (ноль – минимальная скоррелированность факторов) до 9999 (максимальная скоррелированность факторов), но опять же, мы имеем ситуацию, когда возрастание скоррелированности идет до сотни примерно. Поэтому можно считать, что «Карра», в реальности, изменяется от нуля (0) до 100.

Возвращаемся к нашей корреляционной матрице, которая говорит нам, что межфакторные связки должны быть не больше 0,25. Поэтому мы можем установить величину примерно 15-25; но опытным путём, было выведено, что лучше здесь установить величину в 0,15, поэтому её мы и установим значение «Каппа» 10 (рис. 25).

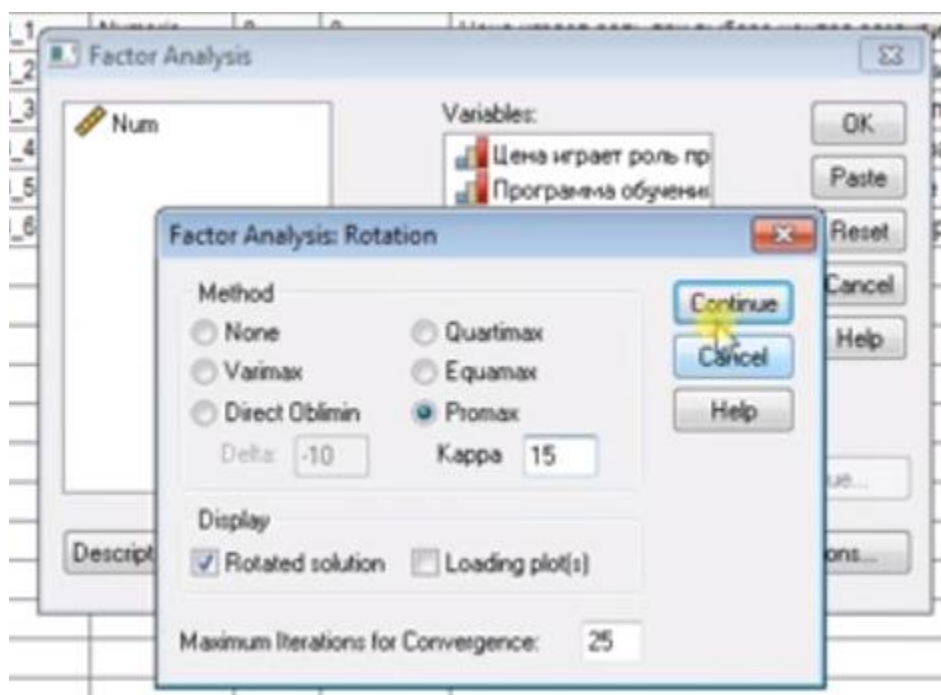


Рис.25.

Нажимаем на вспомогательной панели панельную кнопку –

- Продолжить

... а затем, на рабочей панели «Factor Fnalysis» (Факторный анализ) –

- ОК

Снова открывается окно просмотра, в котором появляется Корреляционная матрица и матрица нагрузок, но уже с несколько иными значениями (рис. 26).

→ Factor Analysis

[DataSet1] C:\Users\Leonardo\Dropbox\Ведомости_посещения\2014-15_у.р\LatAn\Диагностика_15.sav

	Цена играет роль при выборе центра развития в такую-то очередь	Программа обучения и уровень квалификации играет роль при выборе центра развития в такую-то очередь	Удобство расположения центра играет роль при выборе центра развития в такую-то очередь	Имидж центра играет роль при выборе центра развития в такую-то очередь	Атмосфера в центре играет роль при выборе центра развития в такую-то очередь	Наличие в непосредственной близости ТЦ играет роль при выборе центра развития в такую-то очередь
Correlation	1,000	-,249	-,151	-,158	-,405	-,175
Цена играет роль при выборе центра развития в такую-то очередь	1,000	-,249	-,151	-,158	-,405	-,175
Программа обучения и уровень квалификации играет роль при выборе центра развития в такую-то очередь	-,249	1,000	-,223	,073	-,124	-,248
Удобство расположения центра играет роль при выборе центра развития в такую-то очередь	-,151	-,223	1,000	-,234	-,254	-,101
Имидж центра играет роль при выборе центра развития в такую-то очередь	-,158	,073	-,234	1,000	-,191	-,166
Атмосфера в центре играет роль при выборе центра развития в такую-то очередь	-,405	-,124	-,254	-,191	1,000	,009
Наличие в непосредственной близости ТЦ играет роль при выборе центра развития в такую-то очередь	-,175	-,248	-,101	-,166	,009	1,000

Рис. 26.

Копируем их на страничку Экселя, где они выглядят следующим образом (рис. 27.1).

Матрица нагрузок

	A	B	C	D	E
1	Structure Matrix				
2	Component				
3		1	2	3	
4	P3_1	-0,87	-0,074	0,001	
5	P3_2	0,155	0,706	0,151	
6	P3_3	-0,063	-0,074	-0,962	
7	P3_4	-0,096	0,571	0,334	
8	P3_5	0,768	-0,385	0,402	
9	P3_6	0,263	-0,651	0,194	
10	Extraction Method: Principal Component Analysis.				
11	Rotation Method: Promax with Kaiser Normalization.				
12					
13					
14					
15	Component Correlation Matrix				
16	Compone	1	2	3	
17	1	1	-0,225	0,232	
18	2	-0,225	1	0,003	
19	3	0,232	0,003	1	
20	Extraction Method: Principal Component Analysis.				
21	Rotation Method: Promax with Kaiser Normalization.				

Рис. 27.1.

Таблица скоррелированности факторов

14				
15	Component Correlation Matrix			
16	Compone	1	2	3
17	1	1	-0,225	0,232
18	2	-0,225	1	0,003
19	3	0,232	0,003	1
20	Extraction Method: Principal Component Analysis.			
21	Rotation Method: Promax with Kaiser Normalization.			
22				

Рис. 27.2

Нас вполне устраивает скоррелированность: первый со вторым имеют скоррелированность $-0,225$, первый с третьим – $0,232$, и второй с третьим $0,003$ – очень слабую скоррелированность, но, к сожалению, в факторном анализе в SPSS можно задать только общую характеристику скоррелированности, а отдельно для пары факторов задавать ее не возможно.

Как мы можем видеть, табличка Матрица нагрузок дает нам теперь гораздо более контрастную картину именно по строчкам (см. **рис. 27.1**). Теперь, действительно, для 1-й строчки именно 1-я компонента даёт нам гораздо большую нагрузку (ячейка B4 со значением $-0,87$), для 2-й строчки именно 2-я компонента даёт гораздо большую нагрузку (ячейка C5 со значением $0,706$), для 3-й строчки – третья (ячейка D6 со значением $-0,965$) и так далее...

Однако, нас не оставляет мысль о том, что может быть, будет удобнее использовать ортогональное вращение. Почему? А потому, что с ортогональным вращением на много удобнее работать, особенно – если факторизация для нас не самоцель, а промежуточный шаг (например, если мы хотим затем использовать полученные факторы как оси для расположения, например, наших респондентов; мы хотим визуализировать группы, показав, как они располагаются на наших трёх осях). Если мы хотим это сделать, то мы должны будем воспользоваться графическими возможностями SPSS или Экселя, но ни SPSS, ни Эксель не позволяют представить координатное пространство в неортогональном (непрямоугольном) виде. Поэтому, если мы хотим визуализировать наши результаты, то нам имеет смысл всё-таки пренебречь скоррелированностью наших факторов и задать ортогональное вращение.

Значит, запускаем (перейдя в подпанель «» («Вращение») рабочей панели «Факторный анализ») вращение «Варимакс» (выше было объяснено, почему) и получаем табличку, которая интерпретируется точно также, как и табличка с косоугольным вращением.

У нас 2 таблички, которые мы можем сравнивать между собой, и которые очень похожи по своему качеству (**рис. 28**). Но бывает, когда надо

сравнить много табличек, интерпретировать каждую из них, а потом – сравнивать интерпретации. А это – очень сложно: просто со множеством вариантов интерпретации мозг человека не справится.

После вращения методом Промакс					После вращения методом Варимакс				
	A	B	C	D	M	N	O	P	Q
1	Structure Matrix				Rotated Component Matrix(a)				
2	Component				Component				
3		1	2	3	1	2	3		
4	P3_1	-0,87	-0,074	0,001	P3_1	-0,899	-0,203	0,185	
5	P3_2	0,155	0,706	0,151	P3_2	0,217	0,743	0,091	
6	P3_3	-0,063	-0,074	-0,962	P3_3	-0,03	-0,073	-0,973	
7	P3_4	-0,096	0,571	0,334	P3_4	-0,061	0,566	0,338	
8	P3_5	0,768	-0,385	0,402	P3_5	0,736	-0,286	0,271	
9	P3_6	0,263	-0,651	0,194	P3_6	0,203	-0,63	0,174	
10	Extraction Method: Principal Component Analysis.				Extraction Method: Principal Component Analysis.				
11	Rotation Method: Promax with Kaiser Normalization.				Rotation Method: Varimax with Kaiser Normalization.				
12					a Rotation converged in 6 iterations.				

Рис. 28.

Поэтому можно использовать несложный инструмент для предварительного отбора табличек по формальному основанию – это то, что Алексей Ротмистров предложил назвать оценкой контрастности. И вот эти расчеты, которые мы можем видеть на рисунках 29.1 и 29.2, и связаны как раз с оценкой контрастности.

Расчет оценки контрастности – общий вид

Буфер обмена Шрифт Выравнивание Число Стили Ячейки Редактир																
F4 fx =ABS(B4)																
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Structure Matrix												Rotated Component Matrix(a)			
2	Component												Component			
3		1	2	3		Считаю разницу между нагрузкой доминирующего фактора для каждой переменной						1	2	3		
4	P3_1	-0,87	-0,074	0,001		0,87	0,07	0,00	0,87	0,95	0,08		P3_1	-0,899	-0,203	0,185
5	P3_2	0,155	0,706	0,151		0,16	0,71	0,15	0,71	1,01	0,31		P3_2	0,217	0,743	0,091
6	P3_3	-0,063	-0,074	-0,962		0,06	0,07	0,96	0,96	1,10	0,14		P3_3	-0,03	-0,073	-0,973
7	P3_4	-0,096	0,571	0,334		0,10	0,57	0,33	0,57	1,00	0,43		P3_4	-0,061	0,566	0,338
8	P3_5	0,768	-0,385	0,402		0,77	0,39	0,40	0,77	1,56	0,79		P3_5	0,736	-0,286	0,271
9	P3_6	0,263	-0,651	0,194		0,26	0,65	0,19	0,65	1,11	0,46		P3_6	0,203	-0,63	0,174
10	Extraction Method: Principal Component Analysis.										2,19		Extraction Method: Principal Component Analysis.			
11	Rotation Method: Promax with Kaiser Normalization.												Rotation Method: Varimax with Kaiser Normalization.			
12													a Rotation converged in 6 iterations.			

Рис. 29.1.

Расчет оценки контрастности

F	G	H	I	J	K	L	M
							Rotated Co
Считаю разницу между нагрузкой доминирующего фактора для каждой переменной							
0,87	0,07	0,00	0,87	0,95	0,08		P3_1
0,16	0,71	0,15	0,71	1,01	0,31		P3_2
0,06	0,07	0,96	0,96	1,10	0,14		P3_3
0,10	0,57	0,33	0,57	1,00	0,43		P3_4
0,77	0,39	0,40	0,77	1,56	0,79		P3_5
0,26	0,65	0,19	0,65	1,11	0,46		P3_6
					2,19		Extraction

Рис. 29.2.

Речь идет о выделенной табличке, расположенной посередине между двумя нагрузочными табличками, полученными в результате использования разных методов вращения.

Сначала мы берём значения нагрузок по их абсолютному значению (по модулю);

Затем мы берём максимум по каждой строчке;

Затем мы берём сумму по каждой строчке;

А идея всего этого очень простая: если в строчке распределение контрастное, то значит максимум для этой строчки будет равен сумме для этой же строчки.; а если же в строчке распределение не контрастное, то разница будет достаточно большая. Поэтому мы и берём эту разницу (красные числа в крайней левой колонке К4-К9), чем ближе эта разница к НУЛЮ, тем более контрастно распределение. Далее, мы суммируем эту разницу по строчкам (и получаем в ячейке К10 значение 2,19) с тем, чтобы по этим суммам и сравнивать между собой все таблички (в принципе, есть смысл и усреднять эти величины (т.е. сумму делить на общее число строк в таблице), если надо будет сравнивать таблички, имеющие разное число строк).

Итак, чем меньше по своему значению будет это величина (сумма, находящаяся в ячейка К10), тем ближе наши строчки к контрастности, т.е. тем лучше. В плане содержательной интерпретации, – мы видим, что здесь Варимакс выигрывает у Промакса, т.е. ортогональное вращение выигрывает у косоугольного (рис. 29.3 и 29.4).

Сопоставление рассчитанной оценки контрастности по двум таблицам – общий вид

Считаю разницу между нагрузкой доминирующего фактора для каждой переменной					1	2	3	Считаю разницу между нагрузкой доминирующего фактора для каждой переменной							
0,87	0,07	0,00	0,87	0,95	0,08	P3_1	-0,899	-0,203	0,185	0,90	0,20	0,19	0,90	1,29	0,39
0,16	0,71	0,15	0,71	1,01	0,31	P3_2	0,217	0,743	0,091	0,22	0,74	0,09	0,74	1,05	0,31
0,06	0,07	0,96	0,96	1,10	0,14	P3_3	-0,03	-0,073	-0,973	0,03	0,07	0,97	0,97	1,08	0,10
0,10	0,57	0,33	0,57	1,00	0,43	P3_4	-0,061	0,566	0,338	0,06	0,57	0,34	0,57	0,97	0,40
0,77	0,39	0,40	0,77	1,56	0,79	P3_5	0,736	-0,286	0,271	0,74	0,29	0,27	0,74	1,29	0,56
0,26	0,65	0,19	0,65	1,11	0,46	P3_6	0,203	-0,63	0,174	0,20	0,63	0,17	0,63	1,01	0,38
					2,19										2,13

Extraction Method: Principal Component Analysis.
 Rotation Method: Varimax with Kaiser Normalization.
 a. Rotation converged in 6 iterations.

Рис. 29.3.

Сопоставление рассчитанной оценки контрастности по двум таблицам

Вращение методом Промакс					Вращение методом Варимакс				
F	G	H	I	J	R	S	T	U	V
0,87	0,07	0,00	0,87	0,95	0,90	0,20	0,19	0,90	1,29
0,16	0,71	0,15	0,71	1,01	0,22	0,74	0,09	0,74	1,05
0,06	0,07	0,96	0,96	1,10	0,03	0,07	0,97	0,97	1,08
0,10	0,57	0,33	0,57	1,00	0,06	0,57	0,34	0,57	0,97
0,77	0,39	0,40	0,77	1,56	0,74	0,29	0,27	0,74	1,29
0,26	0,65	0,19	0,65	1,11	0,20	0,63	0,17	0,63	1,01

Рис. 29.4.

Отметим, что выигрывает – совсем не сильно, и в данном случае, когда у нас всего 2 таблицы, этим критерием контрастности можно было бы и не пользоваться. Но когда у нас речь идет о необходимости сравнения большого количества таблиц, вот тогда этот критерий может очень облегчить работу, позволив нам отобрать несколько таблиц с лучшей контрастностью.

В нашем же случае, мы останавливаемся на ортогональном вращении, потому что у нас есть в планах кластеризация респондентов (помните, мы говорили о том, что их было бы неплохо как-то разместить между трёх координатных осей, сформированных из полученных нами факторов?!).

Но здесь мы обращаем внимание на то, что третий фактор содержит в себе только одну переменную, причем, – в обеих таблицах (т.е. в обеих моделях). Поэтому мы вынуждены сделать, выстроить ещё одно ответвление в наших рассуждениях. Итак, что делать, если сформированный фактор объясняет только одну переменную?

В принципе, можно оставить – как есть.

Но это чревато потерей информации.

Когда мы с Вами смотрели на общности (см. рис. 8, который ещё раз воспроизведём ниже), мы видели, что для данной переменной потеря инфор-

Таблица Общности

Communalities

	Initial	Extraction
Цена играет роль при выборе центра развития в такую-то очередь	1,000	,883
Программа обучения и уровень квалификации играет роль при выборе центра развития в такую-то очередь	1,000	,607
Удобство расположения центра играет роль при выборе центра развития в такую-то очередь	1,000	,954
Имидж центра играет роль при выборе центра развития в такую-то очередь	1,000	,438
Атмосфера в центре играет роль при выборе центра развития в такую-то очередь	1,000	,696
Наличие в непосредственной близости ТЦ играет роль при выборе центра развития в такую-то очередь	1,000	,469

Extraction Method: Principal Component Analysis.

Рис. 8 (повтор)

мации небольшая (всего лишь 5%). Но эта потеря произойдет. В других случаях, такая потеря может быть ещё больше. Поэтому и не рекомендуется делать модель, в которой только один фактор отвечает только за одну

наблюдаемую переменную. Лучше временно исключить эту переменную из модели, и затем работать с этой переменной наряду с полученными факторами.

А для того, чтобы данные были однородны, мы эту переменную (P3_3) стандартизуем, поскольку факторы факторный анализ делает в стандартизованном виде, то есть: центрированными по НУЛЮ, и с ЕДИНИЧНОЙ ДИСПЕРСИЕЙ. Соответственно, мы сделаем то же самое (чисто технически с переменной P3_3, и тогда у нас будет 2 фактора и самостоятельная переменная – тогда не произойдет потери информации, даже пятипроцентной).

Обозначенная выше ситуация проиллюстрирована **рисунком 30**.

Сопоставляемые объекты до стандартизирования переменной P3_3			
Переменная P3_3, представленная в факторе 3 в единственном числе		Два других фактора, содержащие по несколько переменных	
	Component		
	1	2	3
P3_1	-0,87	-0,074	0,001
P3_2	0,155	0,706	0,151
P3_3	-0,063	-0,074	0,962
P3_4	-0,096	0,571	0,334
P3_5	0,768	-0,385	0,402
P3_6	0,263	-0,651	0,194

Рис. 30.

Переходим в редактор SPSS, и создаём такую стандартизованную переменную (рис. 31-34).

Идем по командному пути (рис. 31):

- Анализ
- Описательные статистики
- Описательные

После чего, открывается рабочая панель –

- «Описательные статистики» (рис.32.)

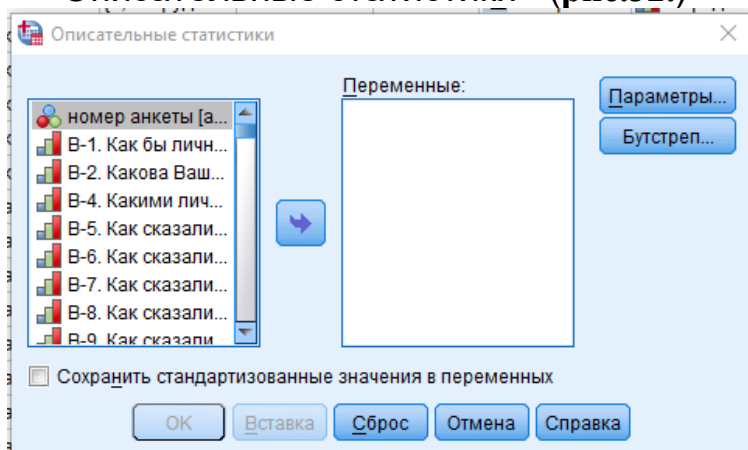


Рис.32.

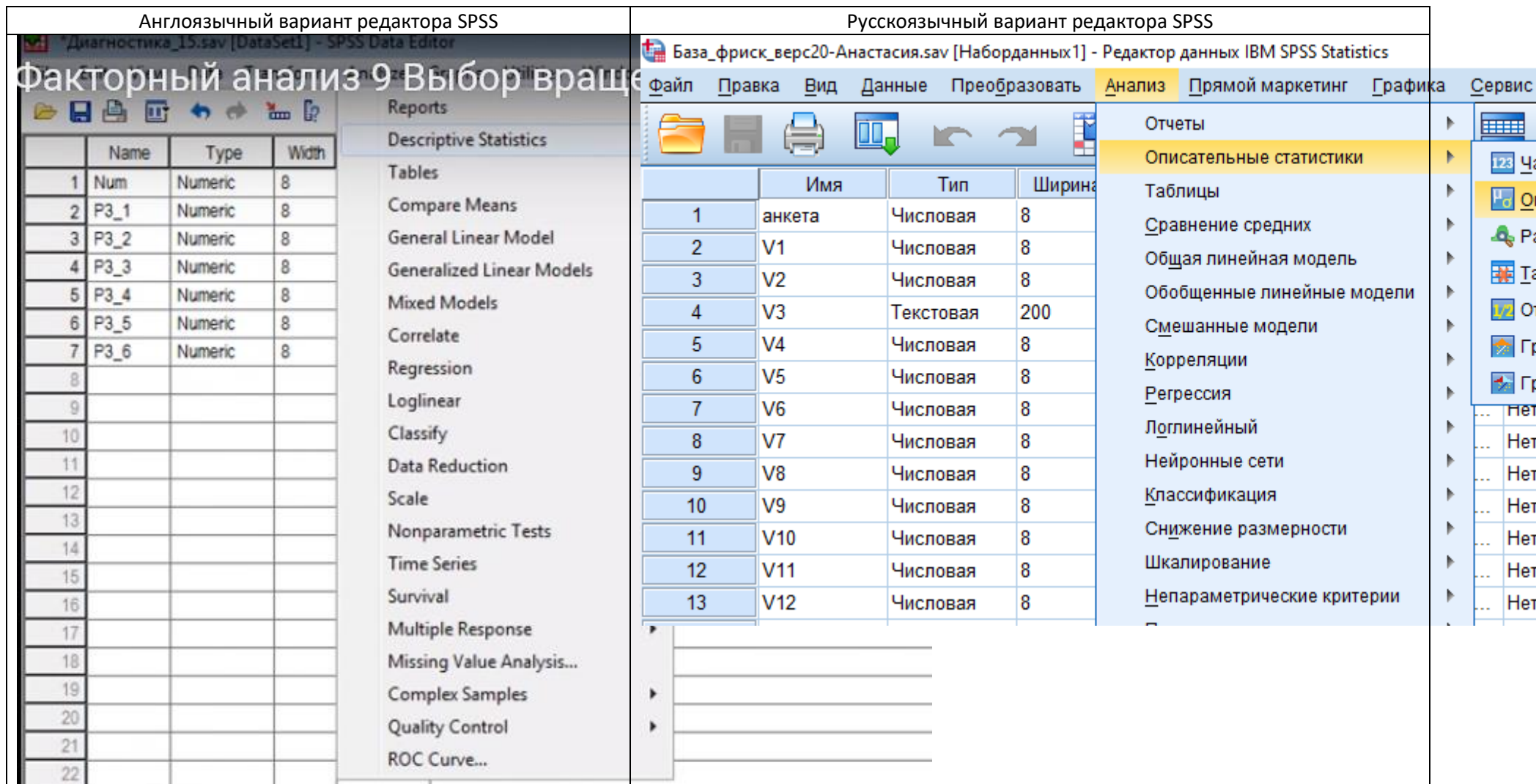


рис. 31.

В активное окно этой панели вставляем соответствующую переменную, которую мы хотим стандартизовать, а также – ставим флажок в клеточке «Сохранить стандартизованные значения в переменных» (рис. 33).

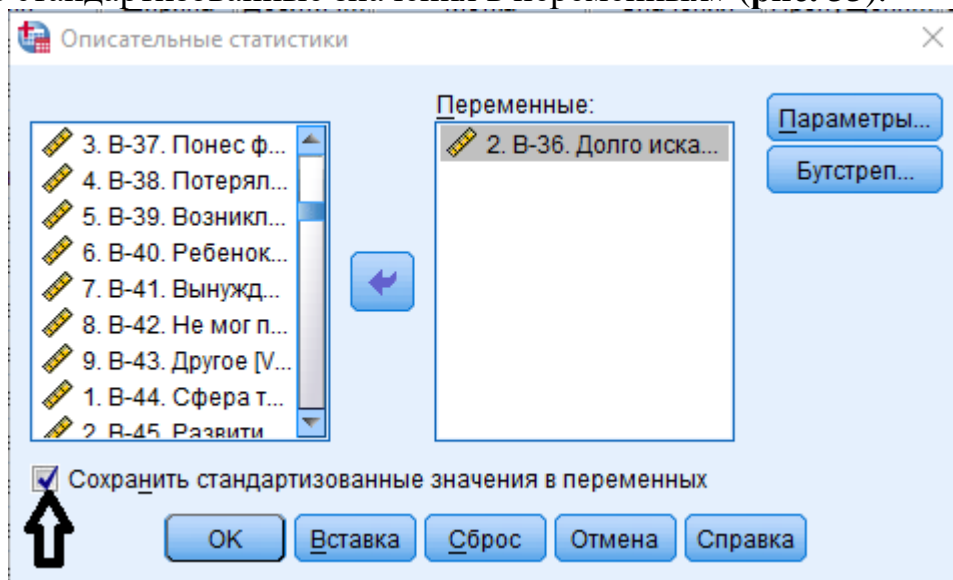


Рис. 33.

После чего, нажимаем панельную кнопку –

- ОК

И – получаем искомую переменную (см. рис. 34).

Вновь созданная Нормированная переменная

Факторный анализ 9-Выбор вращения

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	Num	Numeric	8	0		None	99	8	Right	Scale
2	P3_1	Numeric	8	0	Цена играет роль при выборе центра развития в такую-то очередь	None	99	8	Right	Ordinal
3	P3_2	Numeric	8	0	Программа обучения и уровень квалификации играет роль при выборе центра развития в та	None	99	8	Right	Ordinal
4	P3_3	Numeric	8	0	Удобство расположения центра играет роль при выборе центра развития в такую-то очередь	None	99	8	Right	Ordinal
5	P3_4	Numeric	8	0	Имидж центра играет роль при выборе центра развития в такую-то очередь	None	99	8	Right	Ordinal
6	P3_5	Numeric	8	0	Атмосфера в центре играет роль при выборе центра развития в такую-то очередь	None	99	8	Right	Ordinal
7	P3_6	Numeric	8	0	Наличие в непосредственной близости ТЦ играет роль при выборе центра развития в такую-	None	99	8	Right	Ordinal
8	ZP3_3	Numeric	11	5	Zscore: Удобство расположения центра играет роль при выборе центра развития в такую-то	None	None	13	Right	Scale

Рис. 34.

Мы видим, она стандартизована (рис. 35).

	Num	P3_1	P3_2	P3_3	P3_4	P3_5	P3_6	ZP3_3
1	1	2	1	3	4	5	6	,49956
2	5	2	1	5	3	4	6	2,24044
3	6	2	3	1	5	4	6	-1,24132
4	10	2	3	1	5	4	6	-1,24132
5	11	3	1	2	4	5	6	-,37088
6	14	3	1	2	4	5	6	-,37088
7	16	3	2	1	5	4	6	-1,24132

Рис. 35.

Мы увидим (если проведём сортировку), что отрицательные значения соответствуют единицам в колонке страницы P3_3 паспорта массива «Данные» (рис. 35, рис.36). То есть в исходной переменной – высокой степени

значимости. И – положительные значения, которые соответствуют низкой степени значимости данного мотива для респондентов.

	Num.	P3_1	P3_2	P3_3	P3_4	P3_5	P3_6	ZP3_3	VAR	VAR	VAR	VAR	VAR	VAR	VAR	VAR	VAR	VAR	VAR	VAR
1	6	2	3	1	5	4	6	-1,24132												
2	10	2	3	1	5	4	6	-1,24132												
3	16	3	2	1	5	4	6	-1,24132												
4	50	2	3	1	5	4	6	-1,24132												
5	79	3	2	1	5	4	6	-1,24132												
6	84	3	2	1	5	4	6	-1,24132												
7	85	1	1	1	1	1	6	-1,24132												
8	96	3	2	1	5	4	6	-1,24132												
9	122	2	3	1	4	5	6	-1,24132												
10	123	2	3	1	4	5	6	-1,24132												
11	128	3	2	1	5	4	6	-1,24132												
12	137	2	3	1	5	4	6	-1,24132												
13	138	3	2	1	5	4	6	-1,24132												
14	150	3	2	1	5	4	6	-1,24132												
15	155	2	3	1	5	4	6	-1,24132												
16	158	2	3	1	5	4	6	-1,24132												
17	159	3	2	1	5	4	6	-1,24132												
18	162	3	2	1	5	4	6	-1,24132												
19	177	3	2	1	5	4	6	-1,24132												
20	178	2	3	1	5	4	6	-1,24132												
21	247	2	3	1	5	4	6	-1,24132												
22	251	3	2	1	5	4	6	-1,24132												
23	274	4	3	1	2	5	6	-1,24132												
24	295	2	3	1	5	4	6	-1,24132												
25	299	2	3	1	5	4	6	-1,24132												
26	301	3	2	1	4	5	6	-1,24132												
27	308	3	2	1	4	5	6	-1,24132												
28	310	2	3	1	4	5	6	-1,24132												
29	311	2	3	1	4	5	6	-1,24132												
30	316	2	3	1	4	5	6	-1,24132												
31	325	3	2	1	5	4	6	-1,24132												

Рис. 36

Напомню, что поскольку, в соответствии с исходной шкалой, высокая значимость – это единицы, а низкая значимость – это шестёрки, то и преобразование получилось таким образом, что слева на шкале (в отрицательной области) у нас высокая значимость данного мотива, а справа – низкая значимость данного мотива. Это не совсем привычно для восприятия, поэтому, при дальнейшей работе по интерпретации (особенно!!!) об этом не нужно забывать.

Теперь вернёмся к нашей факторной модели, из которой хотим исключить переменную P3_3 – «Удобство расположения».

Ниже на нескольких рисунках показана последовательность операций и команд, которые мы должны произвести для этого в рамках факторного анализа (рис. 37.1-37.12).

Мы решили, что будем делать ортогональное вращение Варимакс по ранее обозначенным причинам.

акторный анализ 10 Выбор вращения

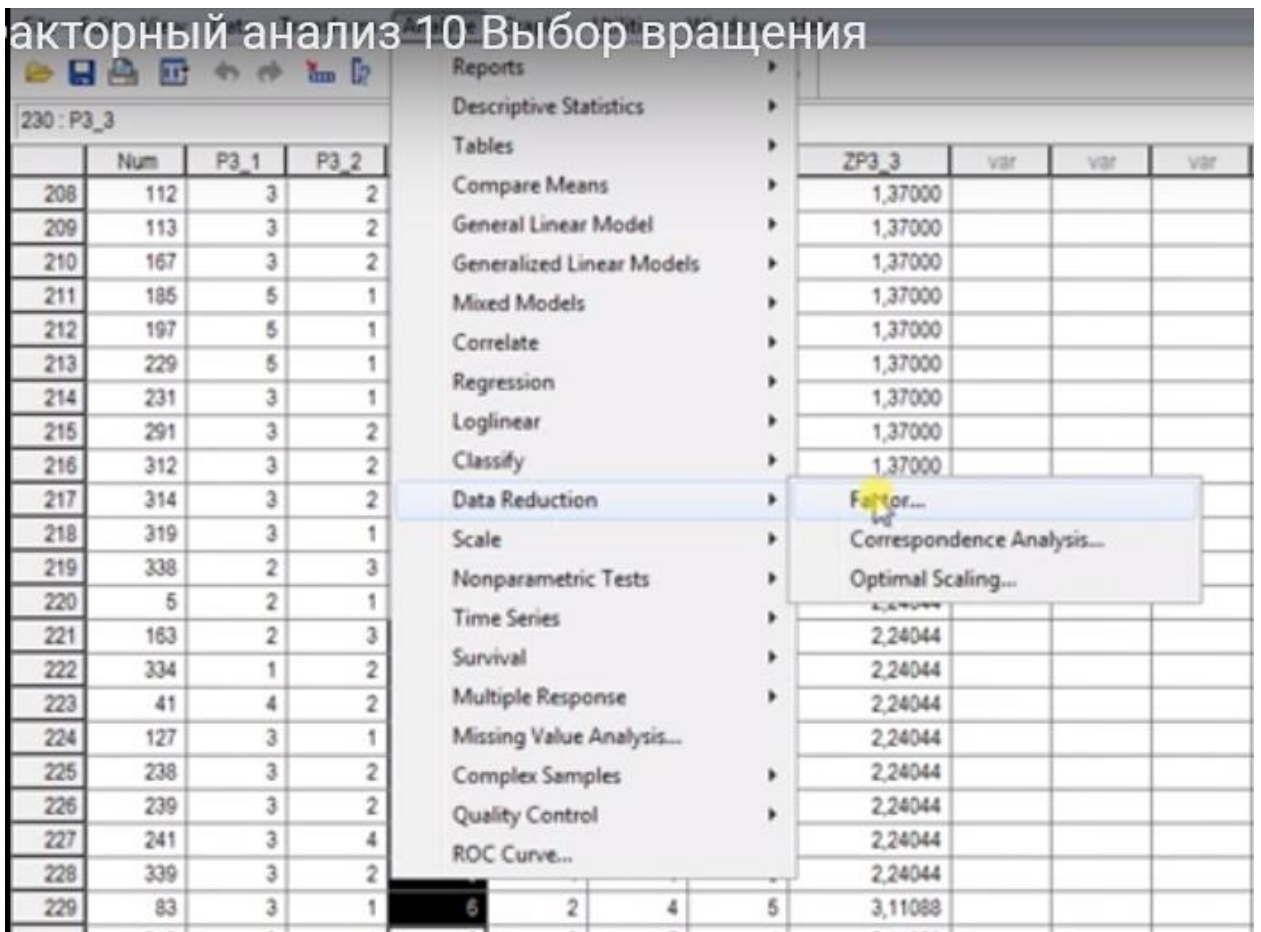


Рис. 37.1

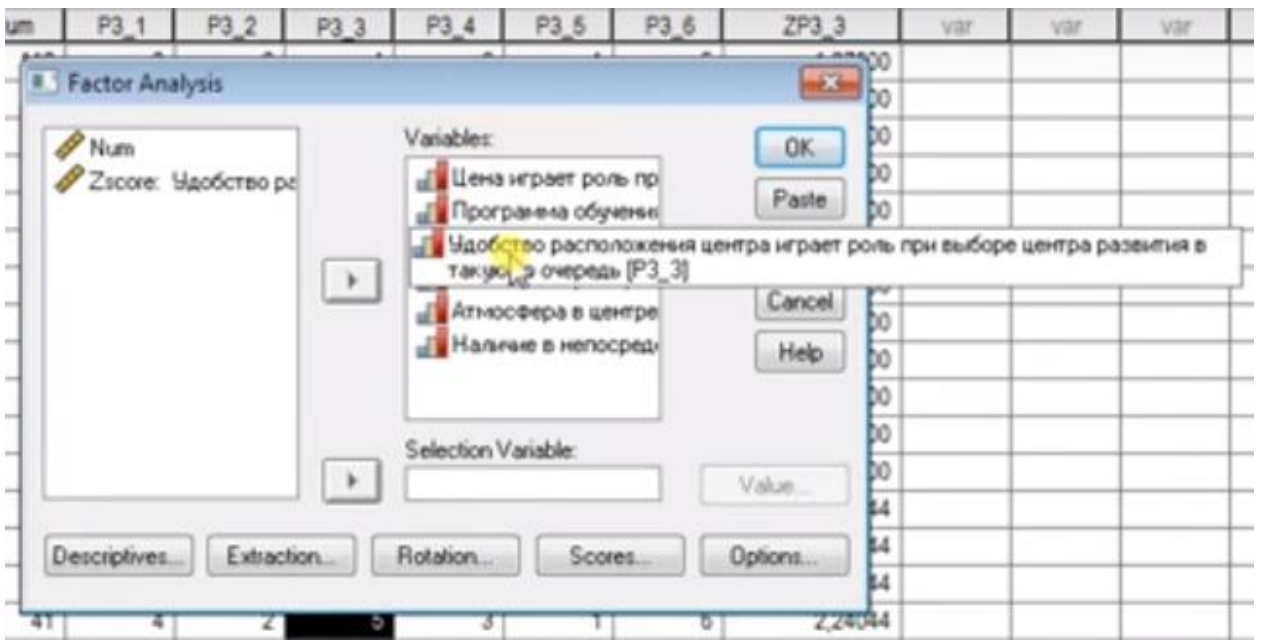


Рис. 37.2.

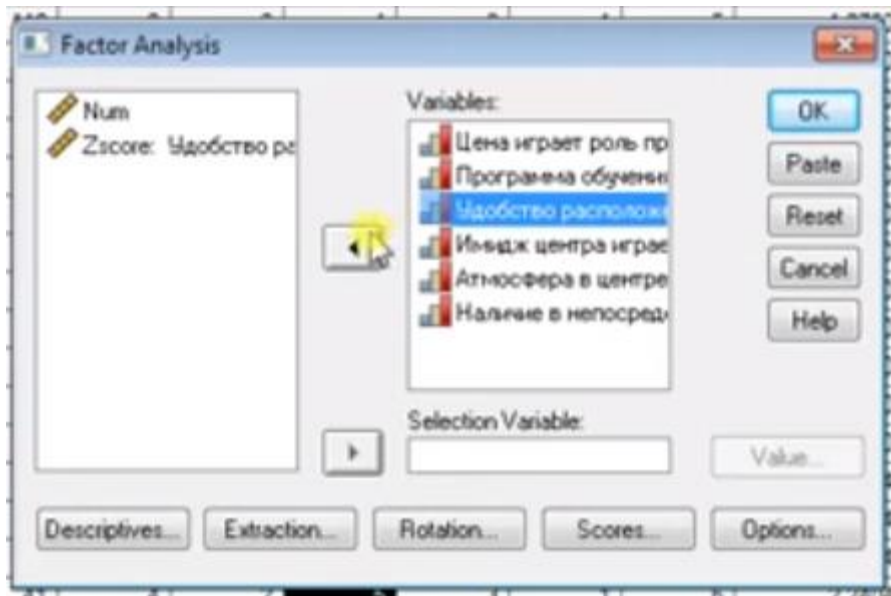


Рис. 37.3.

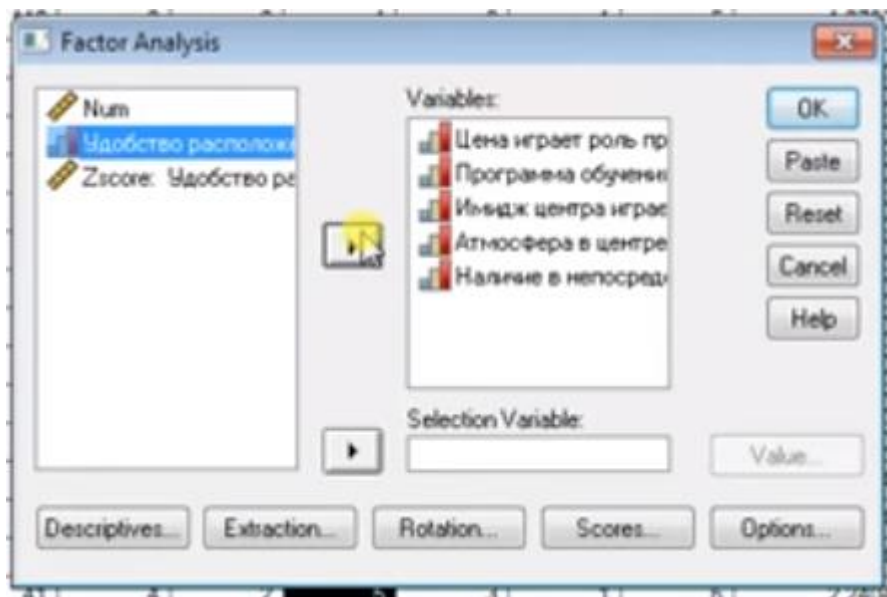


Рис. 37.4.

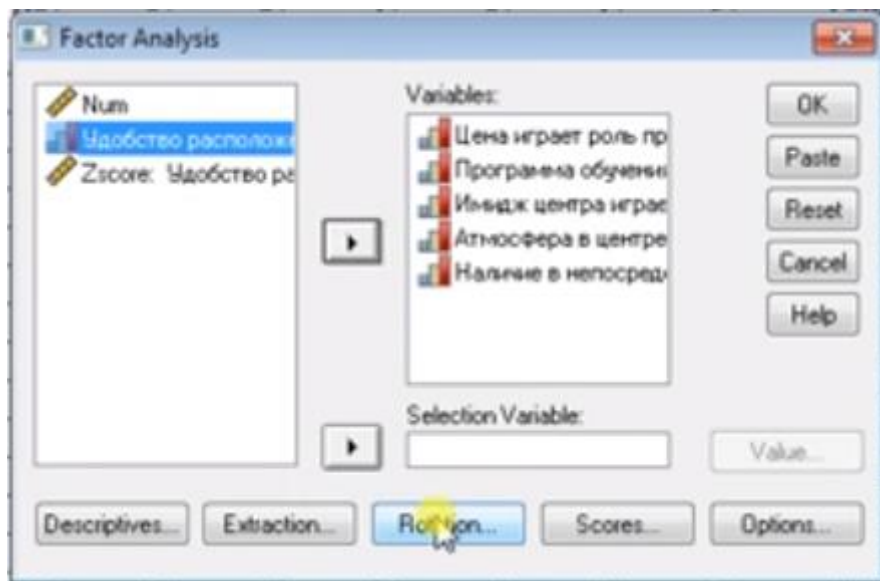


Рис. 37.5.

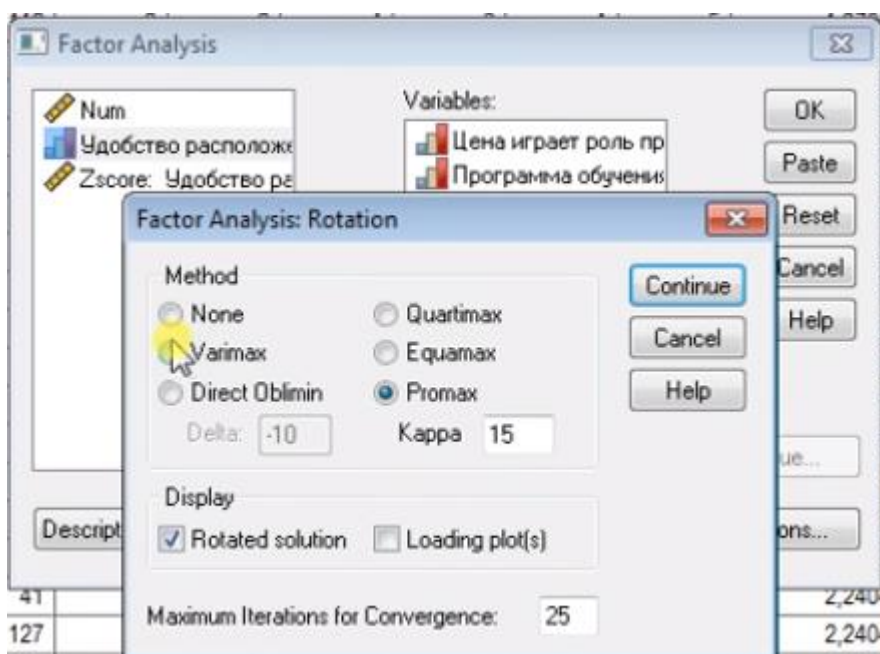


Рис. 37.6.

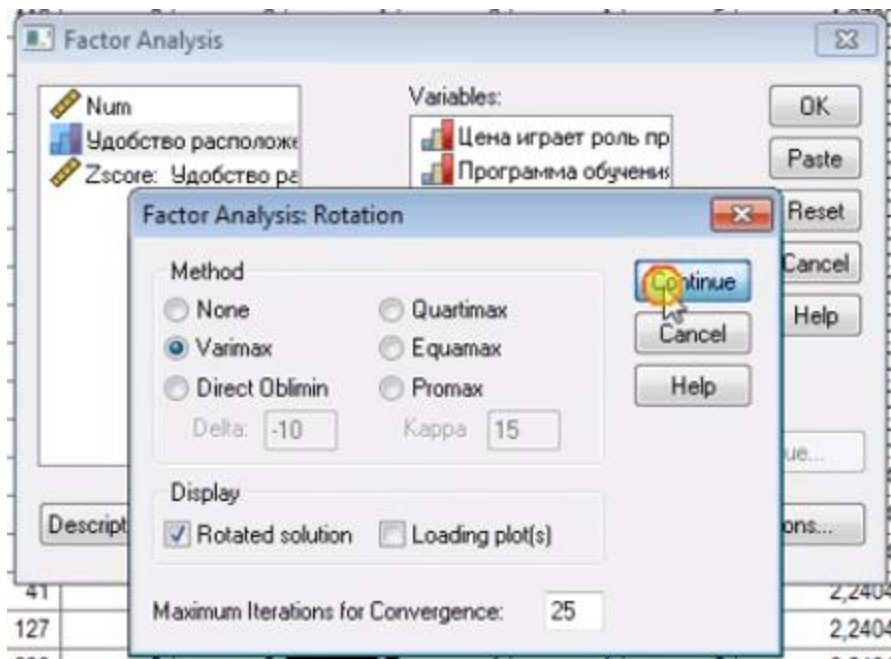


Рис. 37.7.

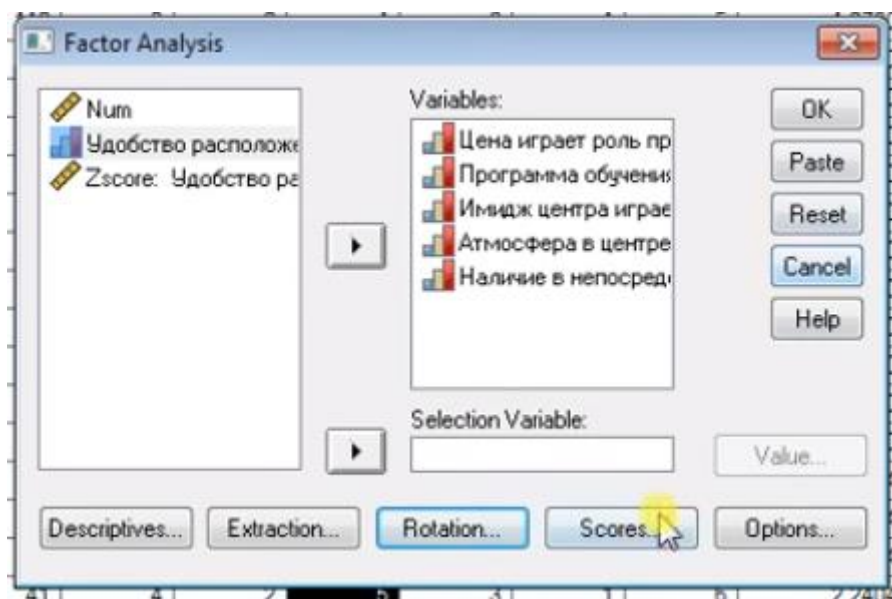


Рис. 37.8.

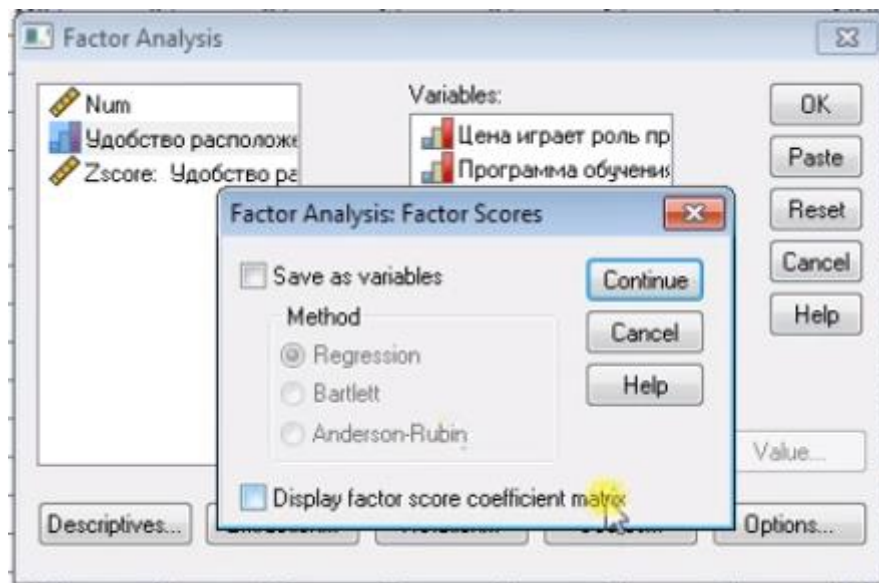


Рис. 37.9.

И попросим сохранить значения наших факторов для каждого респондента.

Поскольку у нас Ортогональное вращение, мы можем выбрать оптимальную функцию сохранения Андерсона-Рубина.

Для этого нам надо активизировать панельную кнопку «Anderson-Rubin» в зоне «Method», и затем – поставить флажок в выше расположенной ячейке «Save as variables».

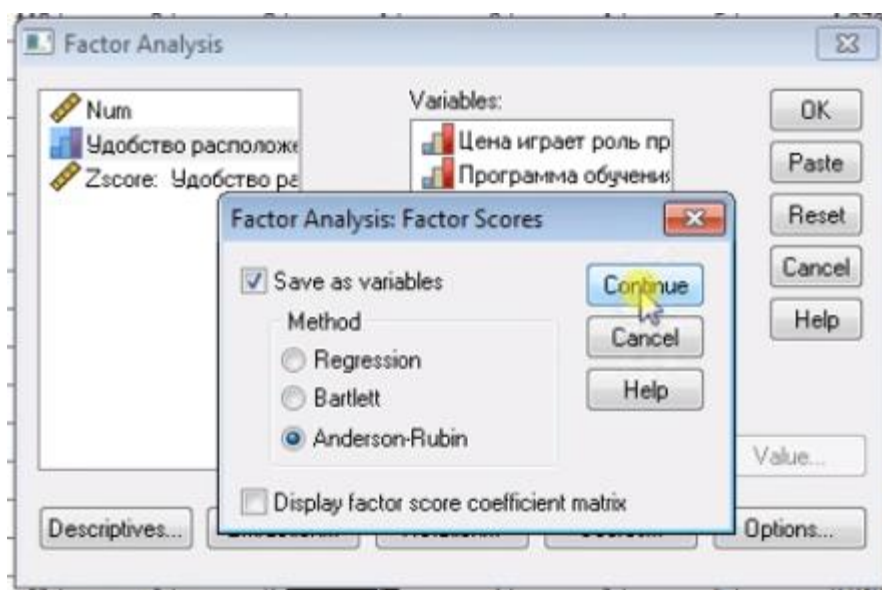


Рис. 37.10.

Напомним! Если бы у нас было Косоугольное вращение, то мы могли бы выбрать только метод регрессии, активизировав для этого панельную кнопку «Regression»

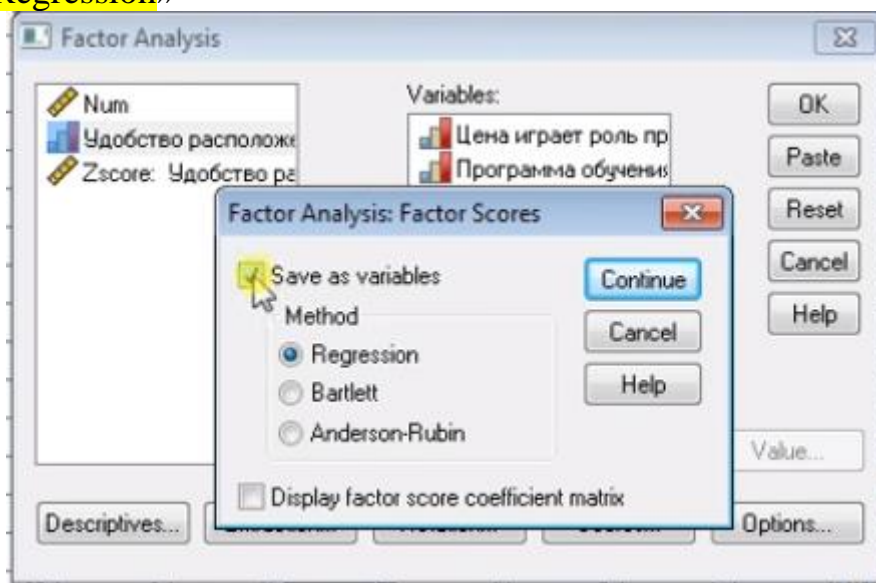


Рис. 37.11.

После этих установок, ждем на вспомогательной панели кнопку –

- «Продолжить»

А затем, после исчезновения вспомогательной панели, кнопку –

- «OK» на оставшейся в поле экрана рабочей панели (рис. 37.12).

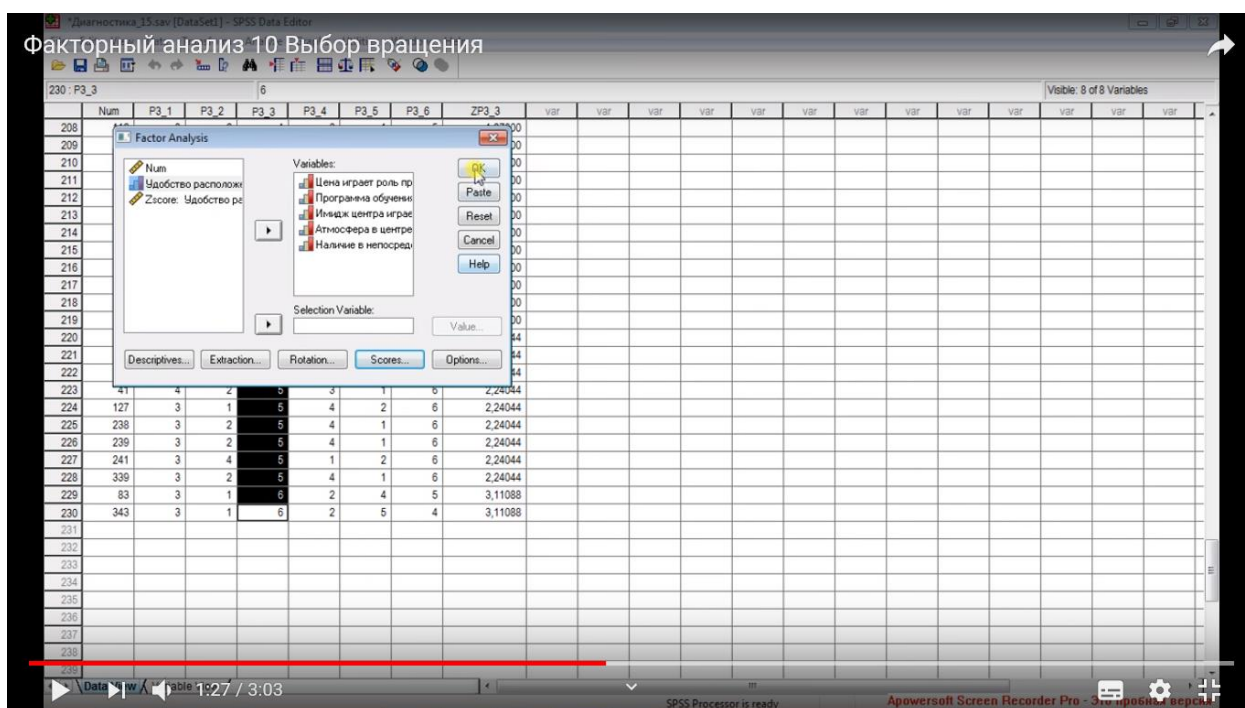


Рис. 37.12.

В итоге этих операций, в нашей базе данных (в листе редактора «Данные») появились три новых колонки: ZP3_3, FAC1_1 и FAC2_1 – т.е. одна

самостоятельная переменная и два фактора – все стандартизованные, т.е. сопоставимые между собой (рис. 38).

Факторный анализ 10 Выбор вращения

4: FAC1_1 0,740365176992533

	Num	P3_1	P3_2	P3_3	P3_4	P3_5	P3_6	ZP3_3	FAC1_1	FAC2_1
1	6	2	3	1	5	4	6	-1,24132	,94365	,84922
2	10	2	3	1	5	4	6	-1,24132	,94365	,84922
3	16	3	2	1	5	4	6	-1,24132	,32029	,15043
4	50	2	3	1	6	4	5	-1,24132	,74037	1,81979
5	79	3	2	1	5	4	6	-1,24132	,32029	,15043
6	84	3	2	1	5	4	6	-1,24132	,32029	,15043
7	85	1	1	1	1	1	6	-1,24132	-,07947	-1,42869
8	95	3	2	1	5	4	6	-1,24132	,32029	,15043
9	122	2	3	1	4	5	6	-1,24132	1,36458	,28308
10	123	2	3	1	4	5	6	-1,24132	1,36458	,28308
11	128	3	2	1	5	4	6	-1,24132	,32029	,15043
12	137	2	3	1	5	4	6	-1,24132	,94365	,84922
13	138	3	2	1	6	4	5	-1,24132	,11700	1,12100
14	150	3	2	1	5	4	6	-1,24132	,32029	,15043
15	155	2	3	1	5	4	6	-1,24132	,94365	,84922
16	158	2	3	1	5	4	6	-1,24132	,94365	,84922
17	159	3	2	1	6	4	5	-1,24132	,11700	1,12100
18	162	3	2	1	5	4	6	-1,24132	,32029	,15043
19	177	3	2	1	5	4	6	-1,24132	,32029	,15043
20	178	2	3	1	5	4	6	-1,24132	,94365	,84922
21	247	2	3	1	6	4	5	-1,24132	,74037	1,81979
22	251	3	2	1	6	4	5	-1,24132	,11700	1,12100

Рис. 38.

Соответствующие обновления возникли и в листе редактора «Переменные», где добавились три строки.

А поскольку и факторы и переменная у нас стандартизованы, т.е. являются теперь равнозначными, сопоставимыми, то мы получаем возможность работать с ними одновременно.

А пока что, интерпретируем наши два фактора (рис. 39).

Первый фактор – это ось, на которой полюсами являются **ориентации на материальные-нематериальные характеристики**: в данном случае, это цена (материальность в узком смысле этого слова), а нематериальные характеристики – это атмосфера, царящая в детском развивающем центре.

То есть, мы можем увидеть по первому фактору людей с большими отрицательными значениями (колонка FAC1_1 на рис. 38). Это люди, для которых не важна цена, но для которых очень важна атмосфера в детском развивающем центре. И наоборот, присутствующие в этом факторе люди с большими положительными значениями данного фактора (колонка FAC1_1 на рис. 38) говорит нам о том, что цена для них очень важна, а атмосфера – не очень их волнует.

Второй фактор: на одном полюсе – щепетильность, т.е. внимательное отношение к программе деятельности центра, к уровню квалификации сотрудников центра, к имиджу центра, а на другом полюсе – наличие рядом с развивающим центром торгового центра.

Графическая и числовая основа для интерпретации двух факторов

активный анализ 10. Выбор вращения

Соц_2011 - Microsoft Excel

Рецензирование Вид

Буфер обмена Шрифт Выравнивание Число

Условное форматирование Форматировать как таблицу Стили Ячейки

Вставить Удалить Формат

Сортировка и фильтр Найти и выделить Редактирование

D6 fx -0,962

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
19	3	0,232	0,003	1																
20	Extraction Method: Principal Component Analysis.																			
21	Rotation Method: Promax with Kaiser Normalization.																			
22																				
23																				
24																				
25	Structure Matrix									Rotated Component Matrix(a)										
26	Component									Component										
27		1	2	Считаю разницу между нагрузкой доминирующего фактора для каждо						1	2	Считаю разницу между нагрузкой доминирующего ф								
28	Цена игра	-0,881	-0,088	0,88	0,09	0,88	0,97	0,09	Цена игра	-0,886	-0,217	0,89	0,22	0,89	1,10	0,22				
29	Программ	0,176	0,695	0,18	0,70	0,70	0,87	0,18	Программ	0,194	0,73	0,19	0,73	0,73	0,92	0,19				
30	Имидж це	-0,044	0,605	0,04	0,61	0,61	0,75	0,04	Имидж це	-0,029	0,607	0,03	0,61	0,61	0,64	0,03				
31	Атмосфер	0,76	-0,372	0,76	0,37	0,76	1,13	0,37	Атмосфер	0,754	-0,267	0,75	0,27	0,75	1,02	0,27				
32	Наличие и	0,246	-0,638	0,25	0,64	0,64	0,88	0,25	Наличие и	0,231	-0,611	0,23	0,61	0,61	0,84	0,23				
33	Extraction Method: Principal Component				0,00	0,00	0,00	0,00	0,00	Extraction Method: Principal Component				0,00	0,00	0,00	0,00	0,00		
34	Rotation Method: Promax with Kaiser Normalization.								0,93	Rotation Method: Varimax with Kaiser Normalization.								0,94		
35	Материал Щепетильность-несерьёзность								a Rotation converged in 3 iterations.											
36																				
37																				
38	Component Correlation Matrix																			
39	Compone	1	2																	
40	1	1	-0,168																	
41	2	-0,168	1																	
42	Extraction Method: Principal Component Analysis.																			
43	Rotation Method: Promax with Kaiser Normalization.																			

Рис. 39.

И мы видим, что для людей с отрицательными значениями по данному фактору (см.: колонка FAC2_1 на **рис. 38**) очень важны программа центра, уровень квалификации персонала, и не важно наличие торгового центра поблизости от развивающего центра; и наоборот, для людей с высокими положительными значениями (см. также: колонка FAC2_1 на **рис. 38**) как правило, важен (хотя здесь, поскольку корреляция не самая тесная, картина не однозначная) но, всё-таки, по большей части важна близость торгового центра рядом, и, соответственно, более низкая значимость программы, квалификации персонала.

Таким образом, из 6-ти наблюдаемых переменных мы получили 2 фактора и одну самостоятельную переменную. Теперь мы можем перейти к кластеризации респондентов в трёхмерном пространстве этих переменных.

Это – одно из возможных направлений дальнейшего развития исследования. Другое же направление можно обозначить выявление с помощью факторного анализа скрытых совместных влияний признаков на результирующую переменную для определения значимых свойств объектов (респондентов), вошедших в каждый из факторов, и разработки на этой основе наиболее действенных средств воздействия на эти объекты.

Составитель: В.Я. Красниковский

