Автоматизация ввода тестовых бланков и экспортирование результата ввода данных в SPSS для статистической обработки

Тукачев Ю.А.
yurij_tukachev@flogiston.ru

В этом обзоре рассматривается процесс ввода данных с заполненных от руки тестовых бланков в память компьютера, а также анализируются особенности и преимущества автоматизированного ввода на примере системы ABBYY FormReader, входящей в комплект ABBYY FineReader 4.0.

Введение

Применение бланковой формы распространено практически во всех видах психологических методик. С другой стороны, сегодня для хранения и обработки информации повсеместно используются компьютеры. Неудивительно, что перенос информации с тестовых бланков в компьютерное хранилище данных стал одной из самых актуальных задач в области повышения эффективности работы психолога за счет автоматизации ввода бланков. Необходимость автоматизации не вызывает сомнения, поскольку, во-первых, позволяет существенно сократить время, затрачиваемое психологом при ручной обработке бланков. Данной процедуре, называемой «наложением ключей», не обучают в курсе психодиагностики (об этом речь пойдет ниже). Во-вторых, предоставляет возможность психологу сконцентрироваться на решении сугубо профессиональных задач благодаря освобождению от трудоемких рутинных операций. В-третьих, снижает возможное количество ошибок в первичной обработки данных и их последующим экспортированием в электронную форму (например, в пакет MS Excel и специализированный статистических пакет SPSS).

Вместе с тем, психологи чаще всего прибегают к ручному способу первичной обработки тестовых бланков с последующим, опять же, ручным вводом первичных оценок по тесту в компьютер в целях их дальнейшей математико-статистической обработки. Те, кто не рад мысли «ручной обработки» (психологи с явно выраженной мануалофобией), предпочитают использовать в своей работе методики, являющиеся полным аналогом «бланковых» или «ручных» методик, получившие название *компьютерных версий* психодиагностических методик. На мой взгляд, является неоправданными попытки тотального перевода бланковых методик на «компьютерные рельсы», поскольку будущее компьютерной психодиагностики за новым классом психодиагностического инструментария – *компьютерными психодиагностическими методиками*.

Бланк обследования — специальная форма фиксации результатов психодиагностического исследования. Предназначена как для учета результатов по отдельным методикам, так и для фиксации данных тематических исследований с помощью комплекса методов (карты обследования).

Б. о. разделяются на рабочие листы для испытуемых, в которых обследуемый сам отмечает результаты решений задач теста или проставляет ответы на опросник, и формы для обработки и интерпретации данных обследования экспериментатором. Нередко оба вида Б. о. объединены в одном блоке.

Б. о. — неотъемлемая часть инструментария тестов бланковых («тесты карандаша и бумаги»). Они обязательны для большинства тестов интеллекта, проективных методик, опросников личностных, а также всех тестов групповых. Конструктивное оформление Б. о. обеспечивает удобство и простоту работы испытуемого, сокращает непроизводительные затраты времени психодиагностического обследования, способствует правильному пониманию испытуемым инструкции, снижает количество случайных ошибок при работе обследуемого и оценке ее результатов. Культура составления Б. о. особенно важна при проведении массовых обследований. Имеется ряд правил для разработки Б. о.:

1. Б. о. следует составлять таким образом, чтобы свести к минимуму непродуктивную работу испытуемого по отысканию места (пропуски в тексте, графы, шкалы и т.д.) для проставления ответа.
2. Предпочтительно, по возможности, свести ответ к подчеркиванию, обведению кружком, обозначению простыми символами (напр. «+», «-» и т.д.). Это требование обязательно для задач закрытого типа.
3. Бланк не должен быть перегружен цифрами, делениями, строками и другими обозначениями.
4. Ответы на отдельные серии заданий (субтесты) должны выделяться на Б. о. отдельно. Там, где задания не разбиваются на группы и их количество велико, графы для ответов должны разделяться на отдельные блоки, напр. по 10-25 заданий.
5. На бланке следует использовать то же название (обозначение), что и на тестовой тетради. В некоторых случаях эти обозначения должны отсутствовать.
6. На Б. о. должны быть предусмотрены места для проставления необходимых данных, таких как пол, возраст, семейное положение, образование, место работы и т.п. (если исследование не является анонимным), а также номера регистрации обследования и даты проведения.
7. Иногда целесообразно кратко продублировать инструкцию по обследованию на бланке для испытуемого. Это требование обязательно при проведении группового обследования.
8. При переходе к ответам на новую группу заданий необходимо поместить примеры заполнения граф.

9. При использовании шаблонов для получения оценок первичных необходимо соответствующее оформление граф для ответов (чаще всего матрица).

В методических материалах к методикам, как правило, предлагаются образцы Б. о. или дается описание условий его составления. В ряде психологических методик незначительное изменение внешнего вида бланка для испытуемого может существенно изменить результаты его работы. Так, показатели выполнения корректурной пробы будут значительно варьировать при применении бланков с различной величиной шрифта, формой букв, интервалом между строками, жирностью шрифта, оттенком бумаги и т.д.

Доступные технологии автоматизации

ABBYY FormReader - система ввода форм, заполненных от руки или на принтере, с применением технологии ICR (Intelligent Character Recognition). Следует заметить, что в настоящее время FormReader самостоятельная линейка продуктов компании ABBYY. Последняя версия, ABBYY FormReader 6.5, доступна в трех модификациях - Desktop Edition, Enterprise Edition и Developer Edition. Стоимость около 6 тыс. руб. Подробнее с данными продуктами Вы можете ознакомиться на официальном сайте компании ABBYY (http://www.abbyy.ru), а о том, какие задачи ввода данных решаются с помощью технологий ABBYY вы можете прочитать в обзоре «Автоматизация ввода форм» (http://www.abbyy.ru/products/handprint/WP_form_processing_65.pdf) и на странице «Что даёт использование ABBYY FormReader?» (http://www.abbyy.ru/formreader/?from=main_page). Есть возможность заказа демонстрационной копии программы после заполнения соответствующей регистрационной формы-заявки.

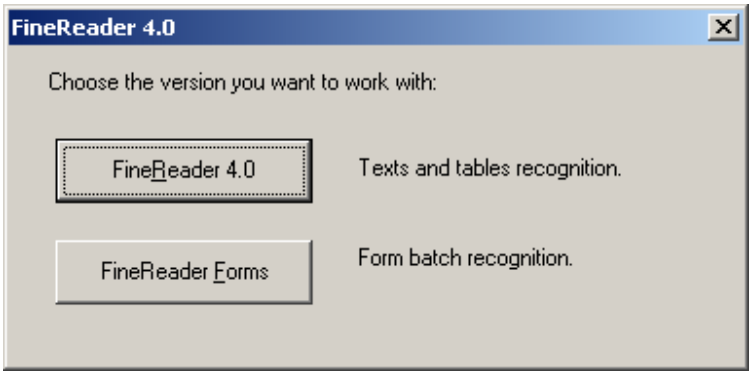
Существует также специализированный программно-аппаратный комплекс «Бланк» (http://mpm03.narod.ru/blank.htm), позволяющий автоматически обрабатывать заполненные испытуемыми регистрационные бланки формата А4 посредством сканирования и последующего распознавания содержащейся на них информации об ответах испытуемых. Стоимость от 8,5 до 22,5 тыс. руб.

Автоматизация ввода тестовых бланков: шаг за шагом

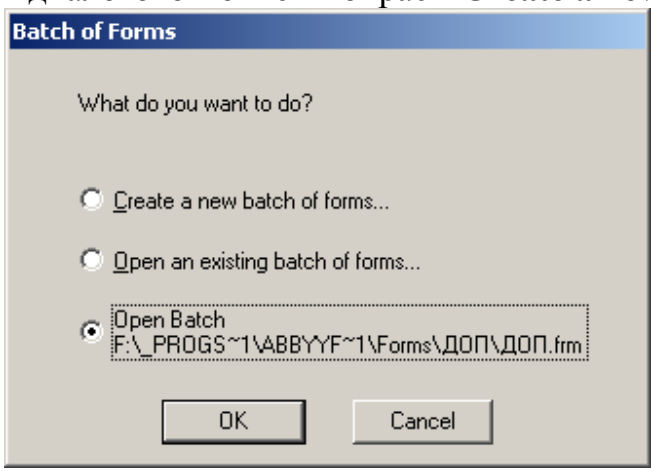
В нашем примере мы будем создавать макет ввода тестового бланка опросника «16 PF». Готовый бланк опросника представлен в файле формата Excel. Его необходимо распечатать. Для распечатанного бланка будем создавать макет для распознавания в программе ABBYY Form Reader 4.0.

ШАГ 1. *Создание макета в программе ABBYY Form Reader 4.0.*

1. Запустите программу, кликнув на файле Fine32.exe. В открывшемся диалогом окне нажимаем кнопку **FormReader Forms**



2. В открывшемся диалоговом окне выбираем **Create a new batch of forms**



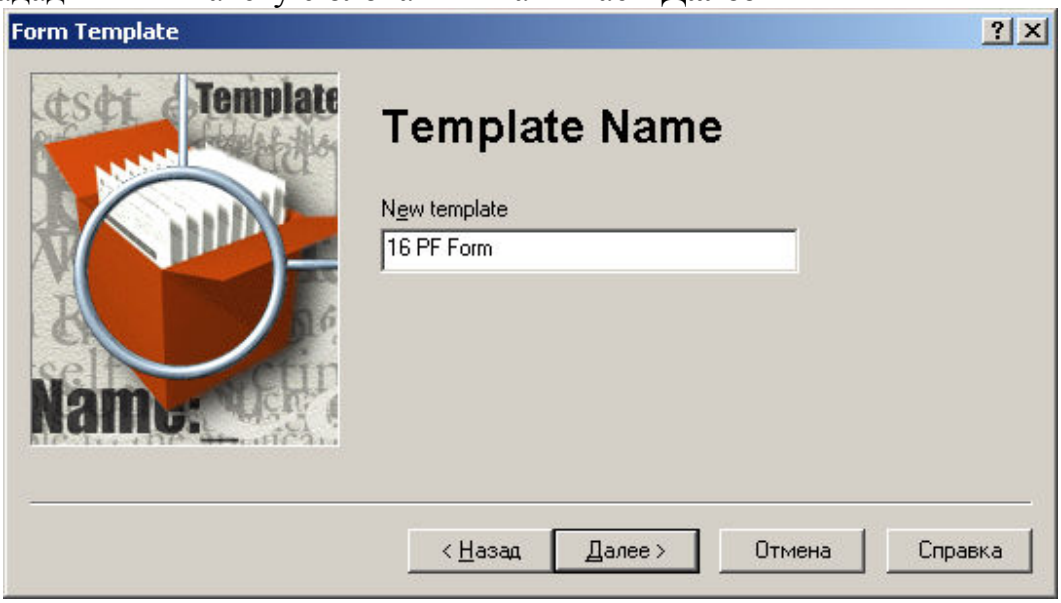
3. Зададим имя проекта 16PF Form и укажем путь для сохранения

Нажимаем «**Далее**»

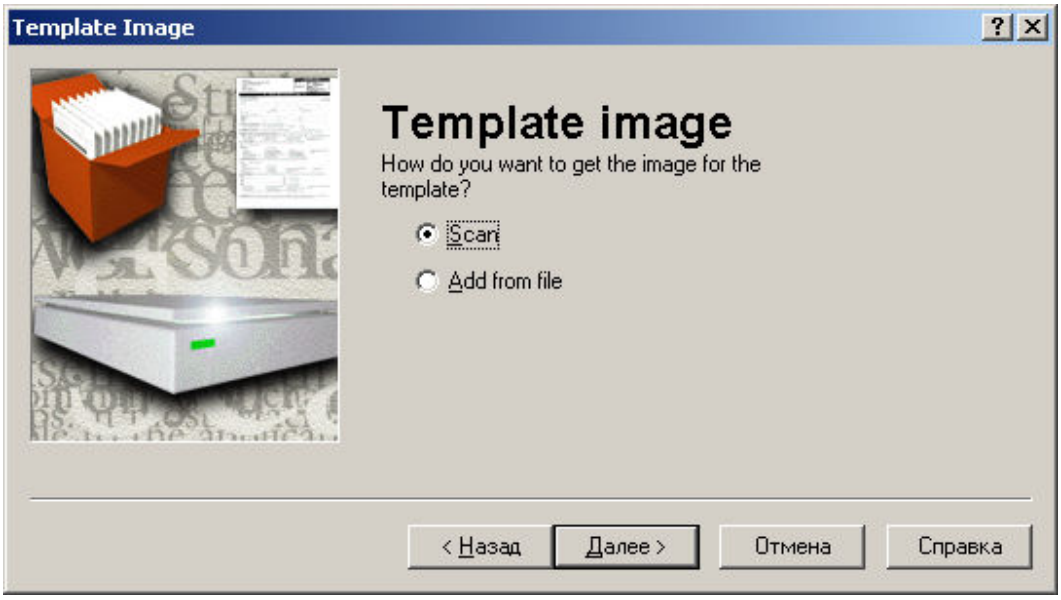
4. В открывшемся диалоговом окне

Выбираем **Create a new template** и нажимаем **Далее**

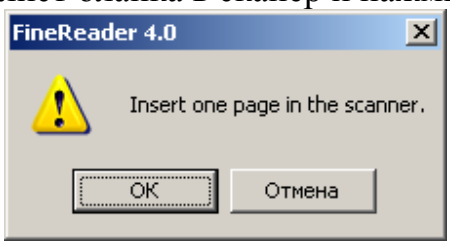
5. Зададим имя макету с блоками и нажимаем **Далее**



6. Отсканируем распечатанный макет бланка

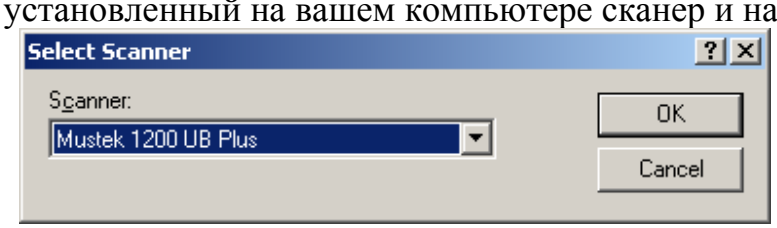


Вставьте лист бланка в сканер и нажмите **Далее**

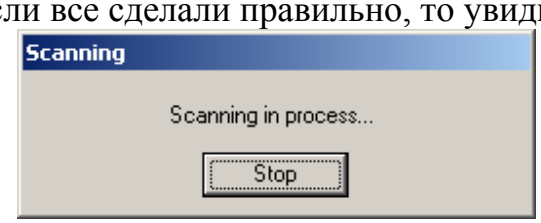


Нажмите ОК

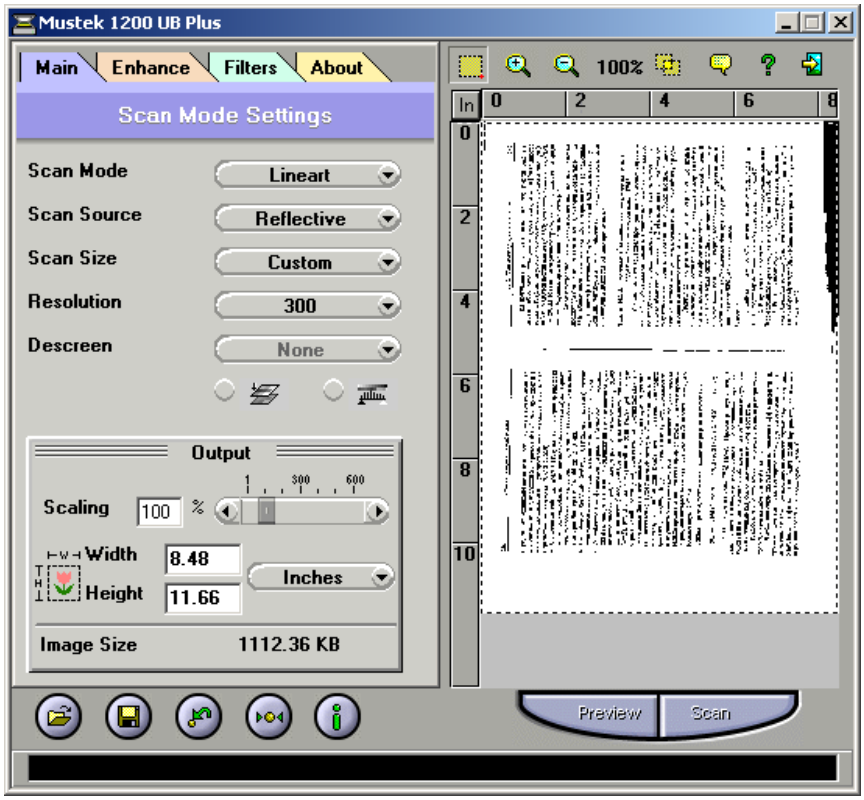
7. Выберите установленный на вашем компьютере сканер и нажмите ОК



Если все сделали правильно, то увидите

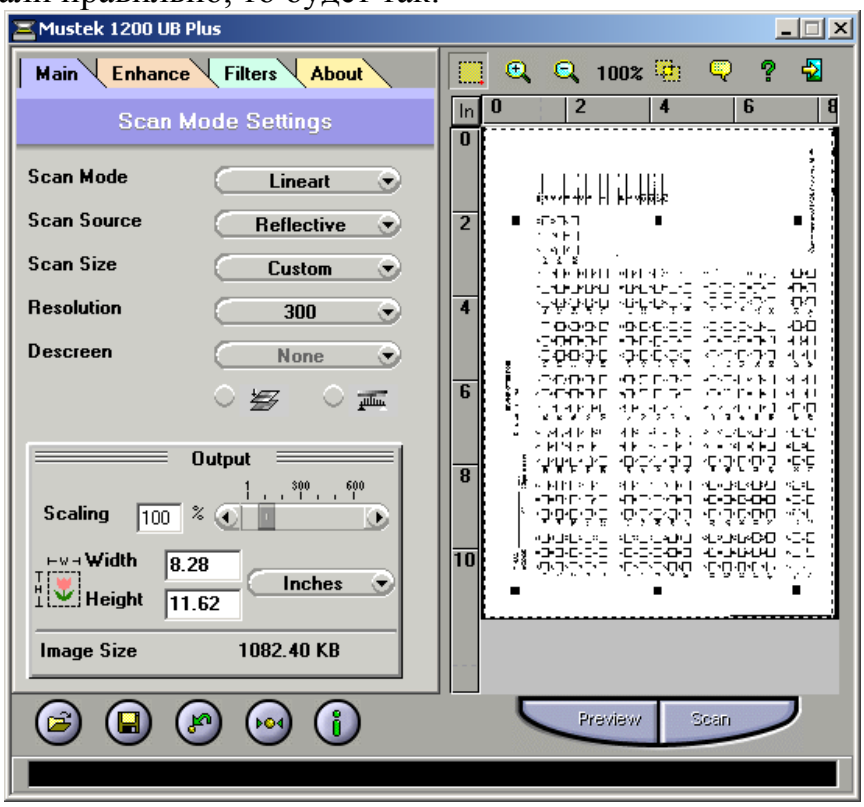


После запустится диалог с настройками режима сканирования



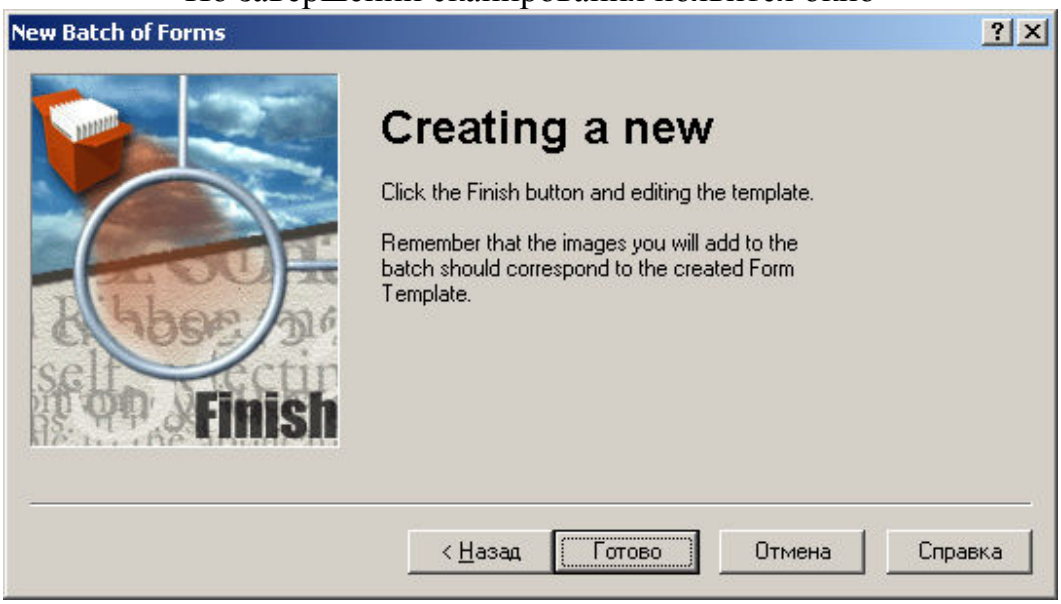
Установите **ScanMode** на **Lineart**, а **Resolution** на **300 (запомните разрешение, т.к. при распознавании бланков необходимо установить такое же разрешение, иначе макет не подойдет)**. Нажмите кнопку **Preview.**

Если все сделали правильно, то будет так:



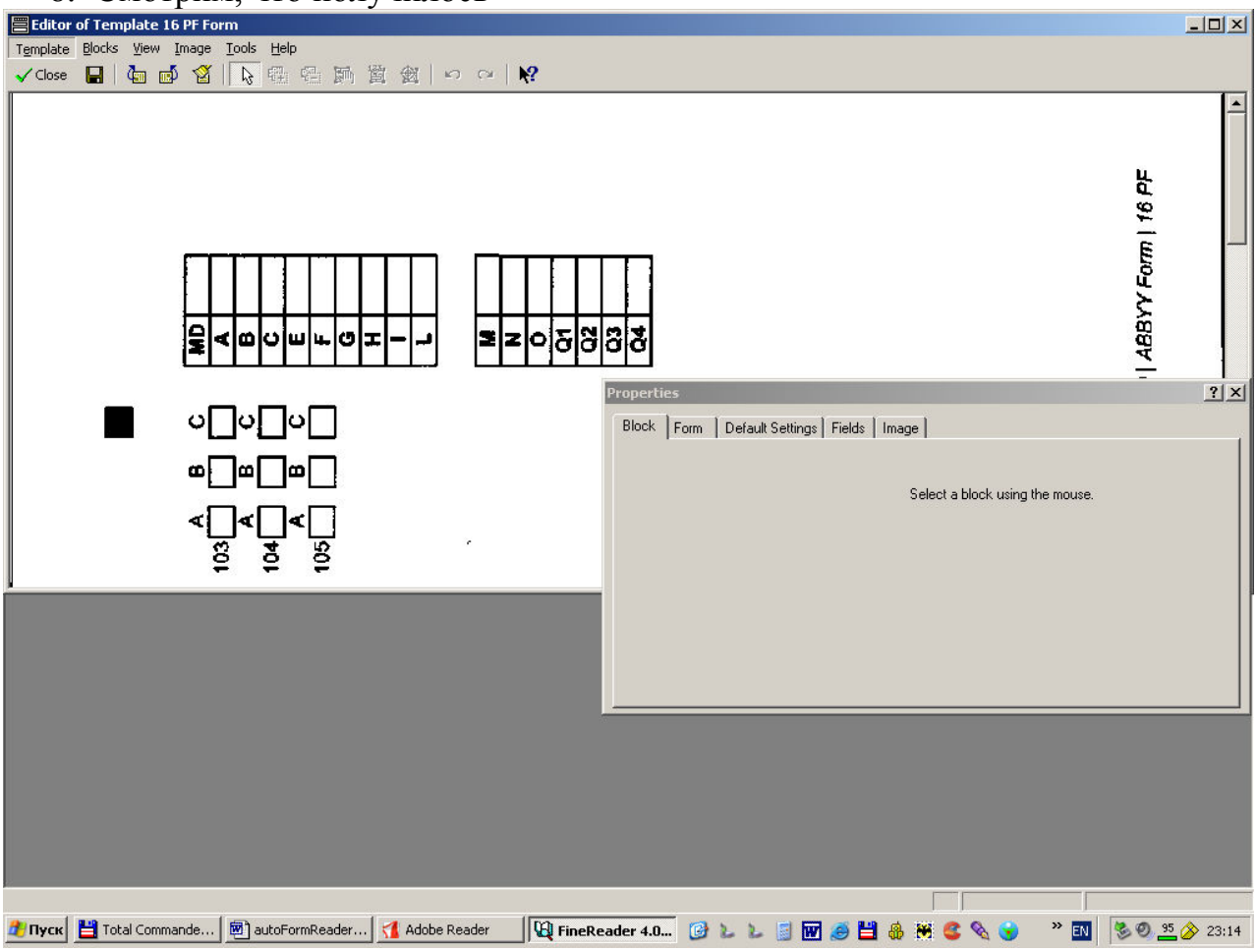
Нажимаем кнопку **Scan**

По завершении сканирования появится окно

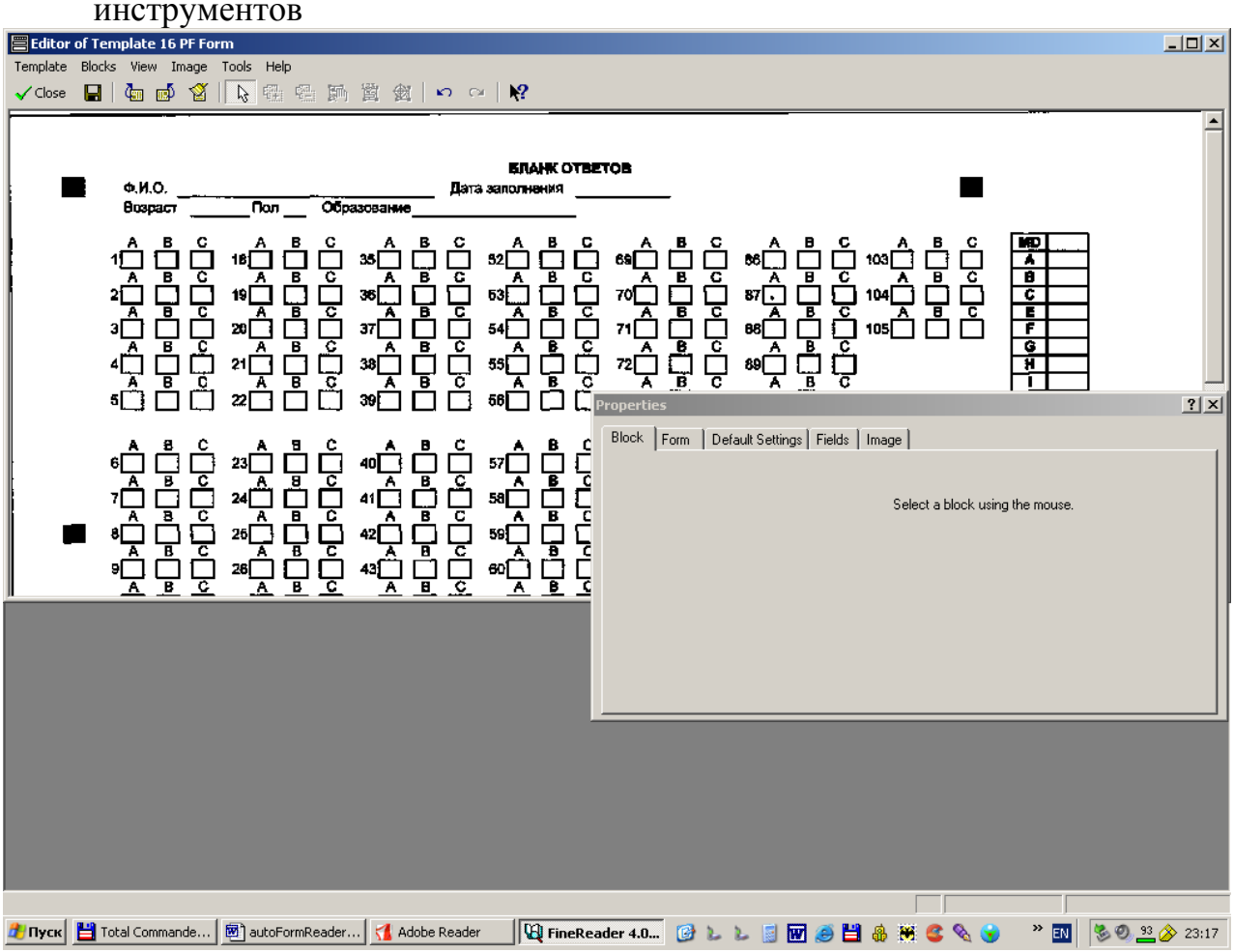


Нажимаем **Готово**

8. Смотрим, что получилось

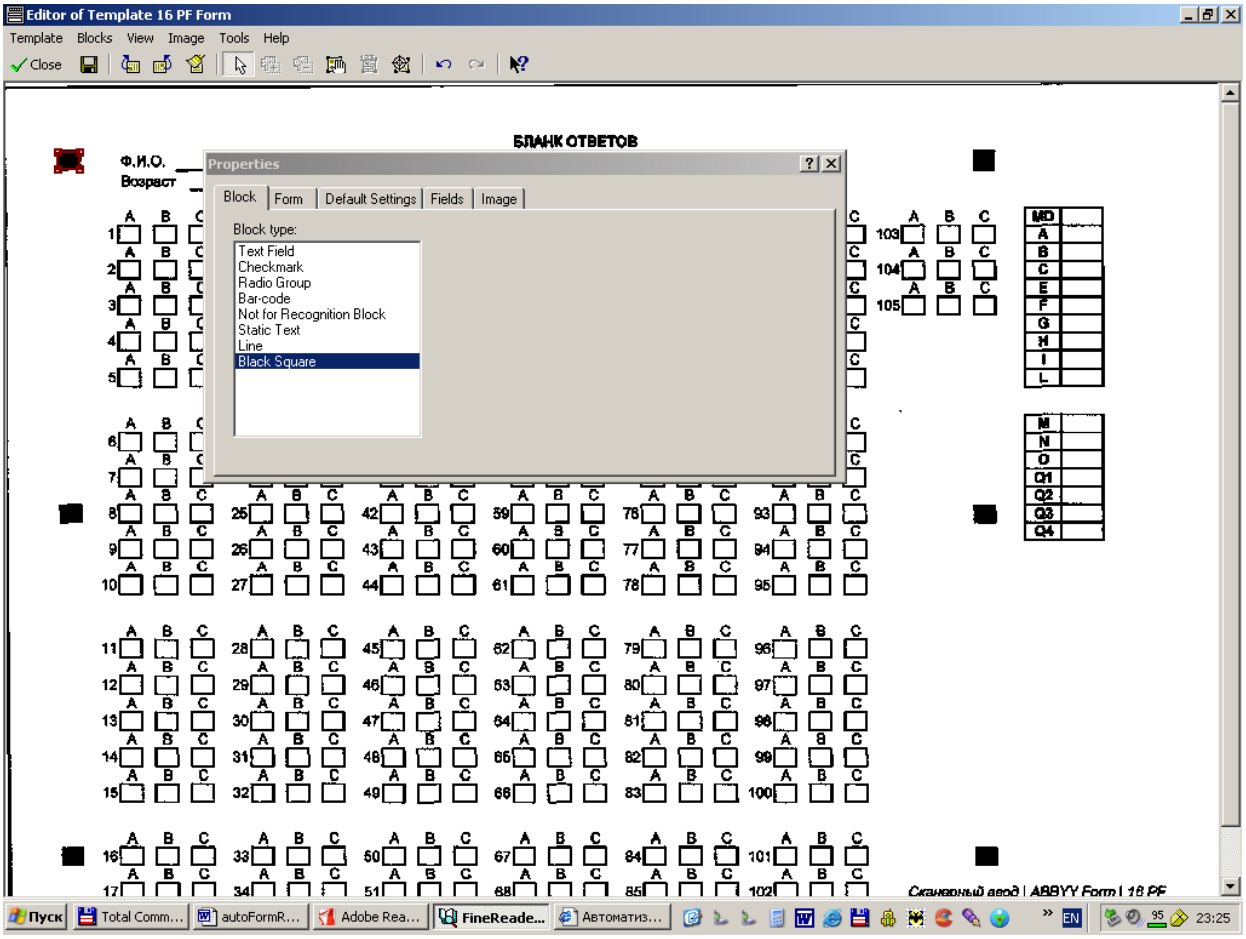


9. Перевернем макет по часовой нажав соответствующую кнопку на панели инструментов

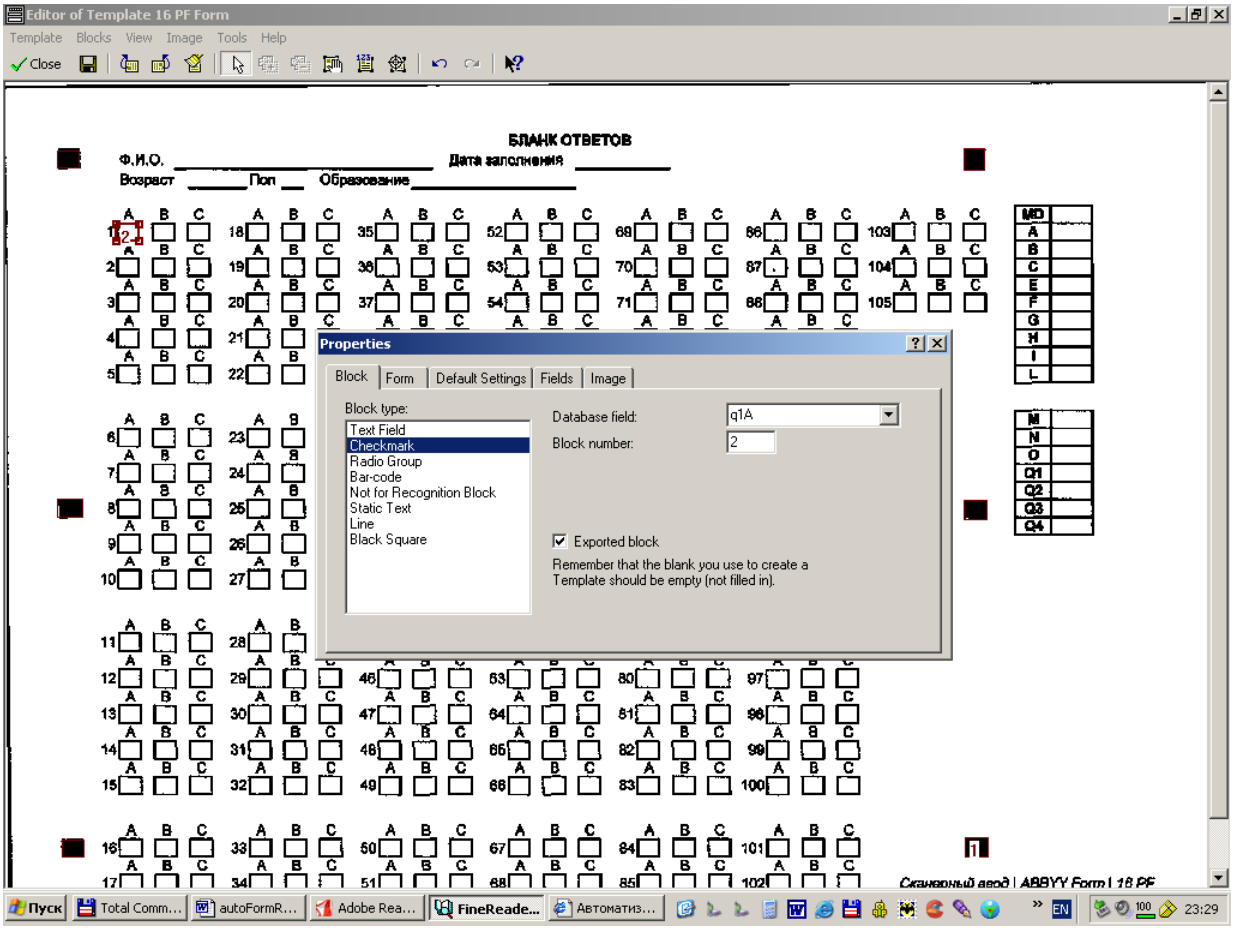


Нажмем иконку «сохранить»

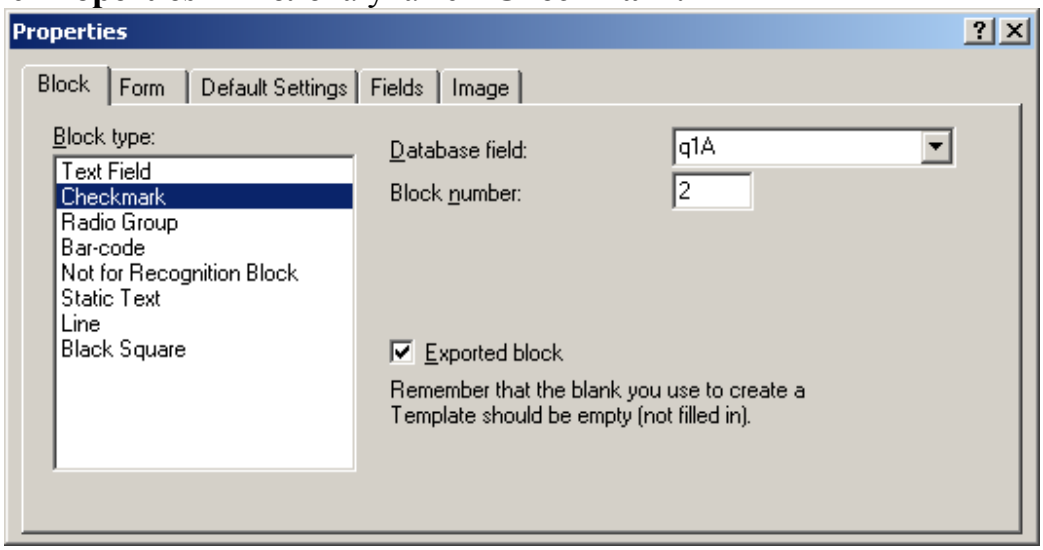
10. Установим черные квадраты шаблона (выделив их левой кнопкой мышки от левого верхнего угла в правому нижнему). А в окне **Properties** тип блока укажем **BlackSquare**



11. Начнем разметку «клеток» вариантов ответа на вопросы



В окне **Properties** тип блока укажем **Checkmark.**

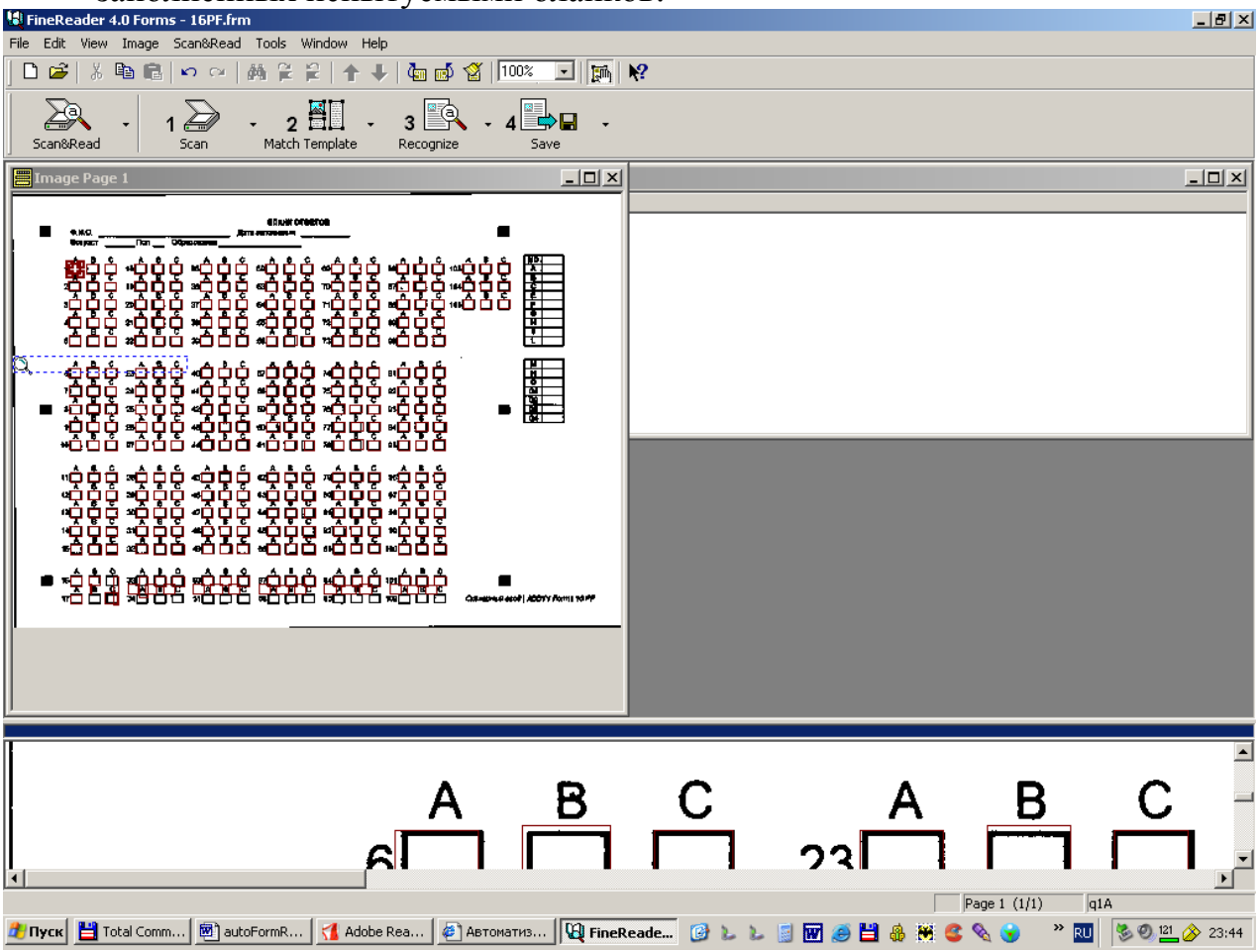


Зададим имя поля **q1A**

(обратите внимание, что для следующего варианта ответа на 1 вопрос имя поля уже будет **q1B**, а для 3 варианта ответа на 1 вопрос – **q1C**). Раскладка клавиатура – английская.

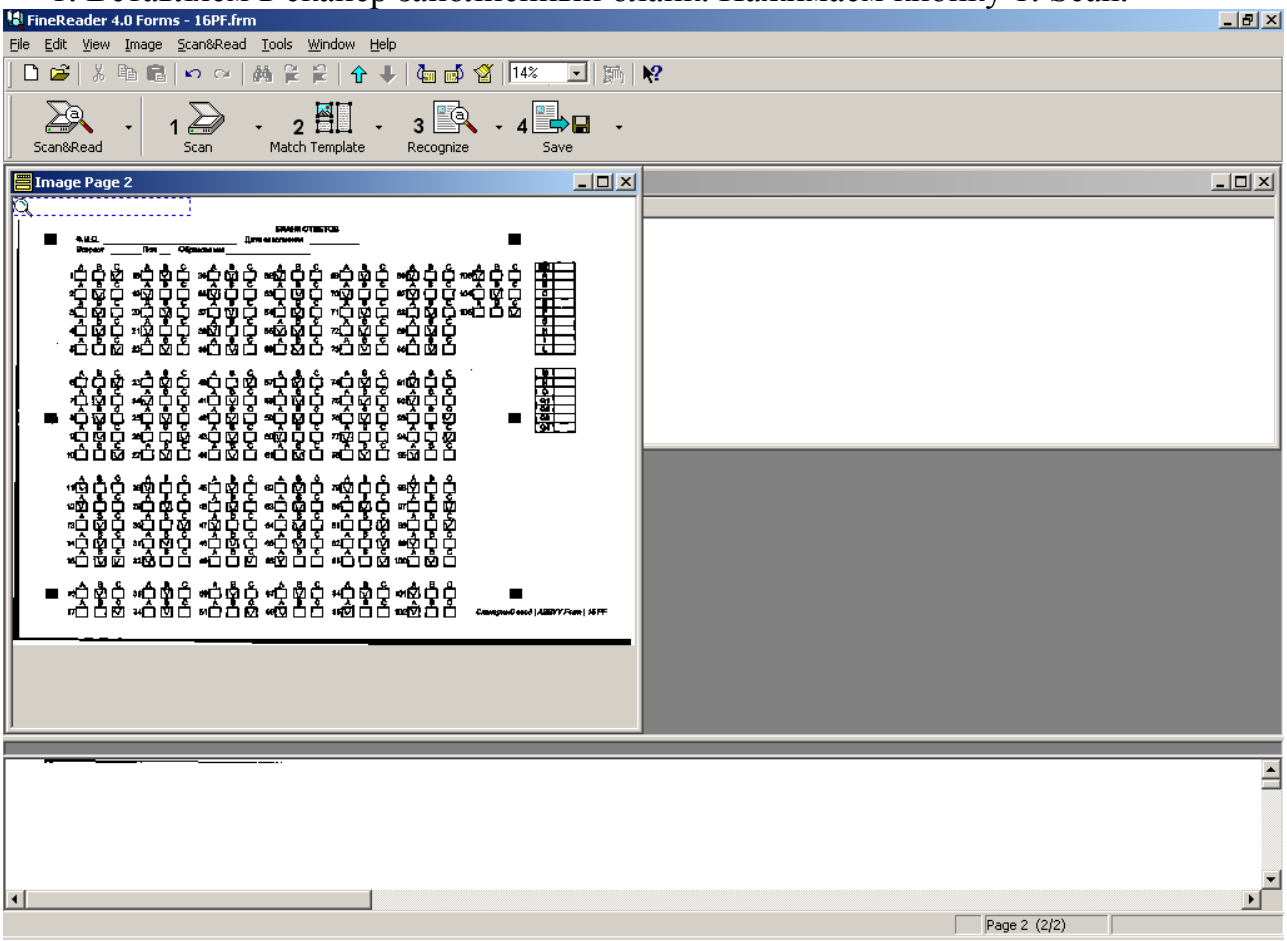
Повторите этот шаг для каждого вопроса. Сохраните макет.

12. ГОТОВО. Макет бланка готов. Теперь можем приступить к сканированию заполненных испытуемыми бланков.

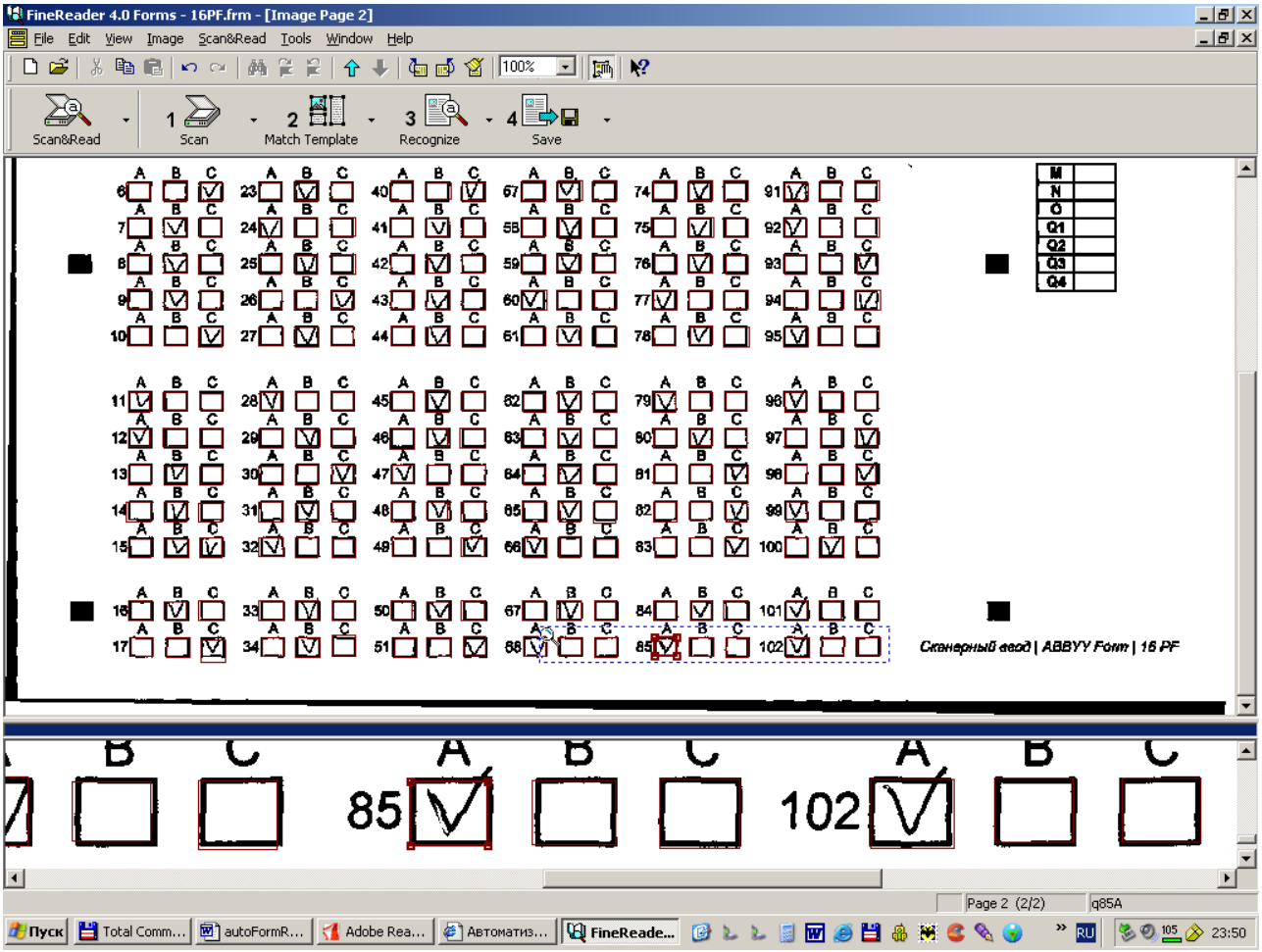


ШАГ 2. *Сканирование заполненных испытуемыми бланков.*

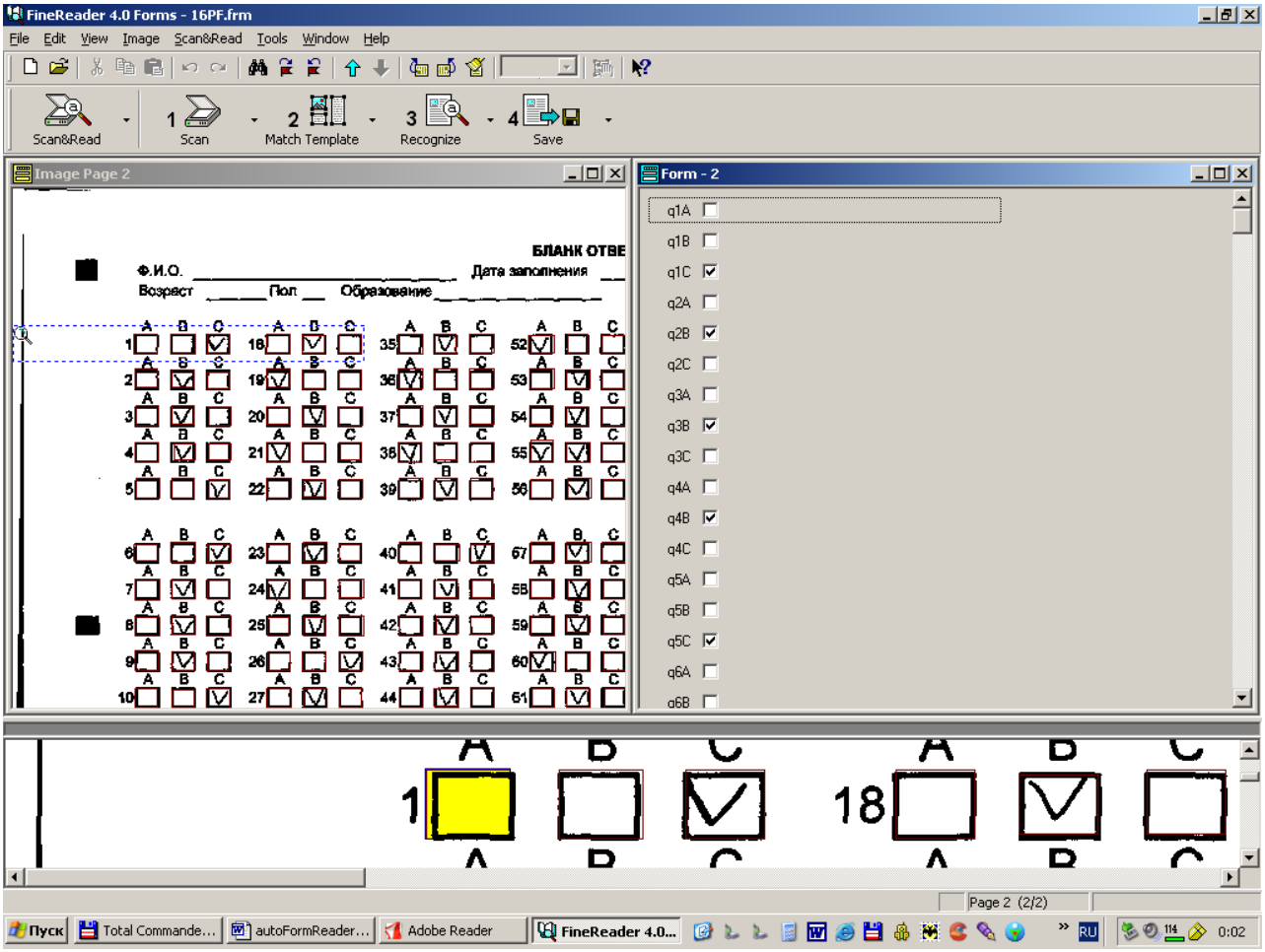
1. Вставляем в сканер заполненный бланк. Нажимаем кнопку 1. Scan.



2. Нажимаем кнопку **2 Match Template** (наложить блоки). Проверьте корректность наложения блоков, если потребуется, то откорректируйте положение блоков.

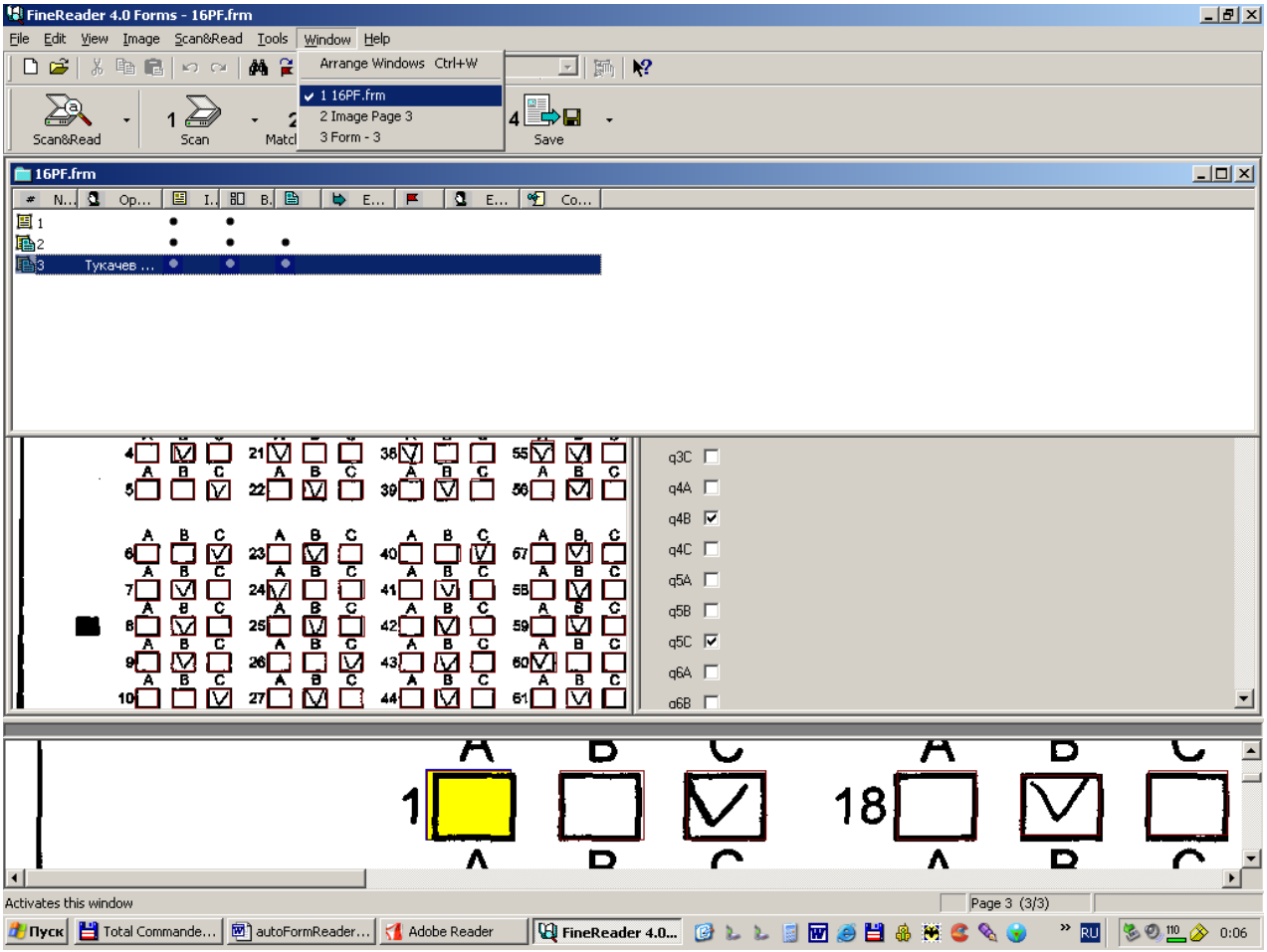


3. Нажимаем кнопку **3 . Recognize**



В поле **Form** можно проверить корректность распознавания ответов на бланке.

4. Повторите шаг2 (для всех имеющихся заполненных бланков)
5. Теперь приступим к сохранению результатов распознавания.

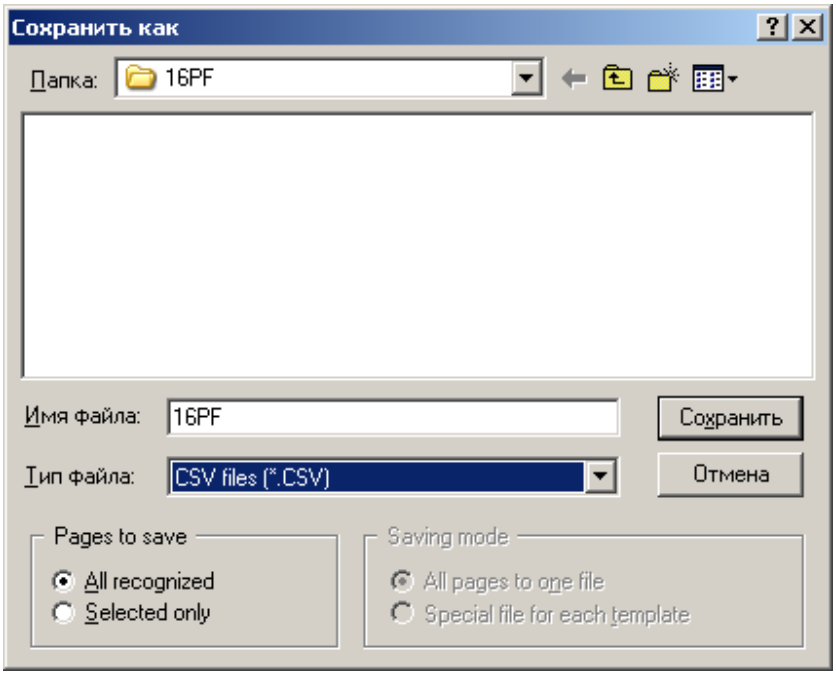


В меню Windows выберите 16PF.frm

В нашем примере есть результаты распознавания трех бланков.

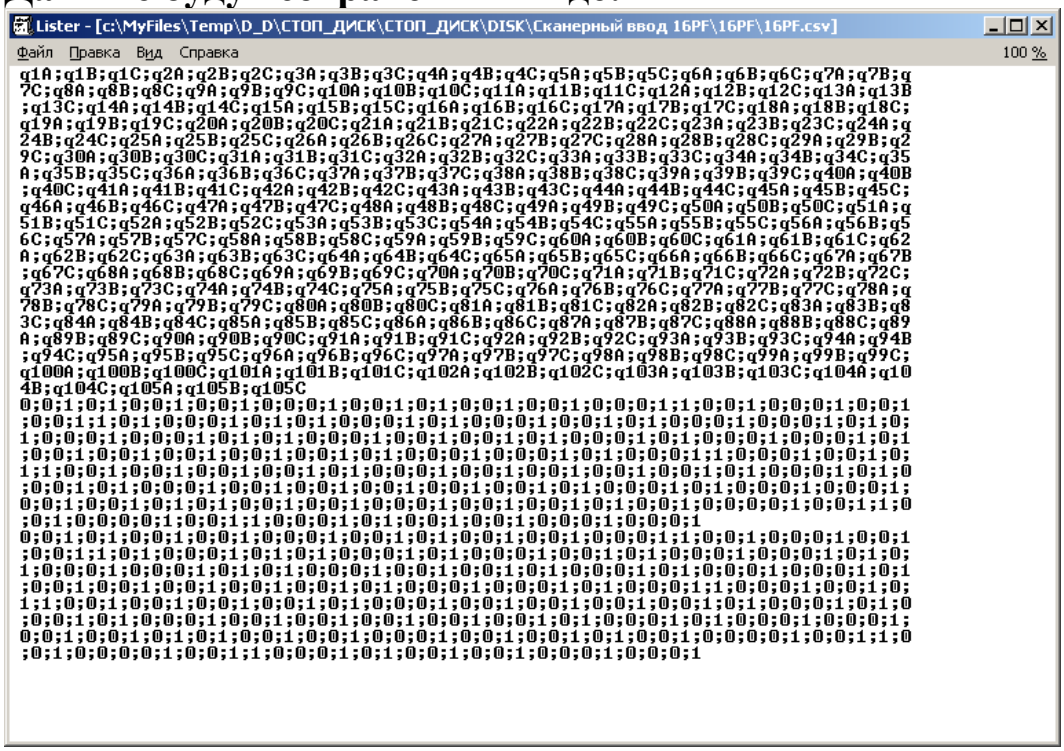
Нажимаем кнопку 4. **Save**

Экспортируем данные в формате CSV



Зададим имя файла данных **16PF**, а тип файла – **CSV**. Нажимаем **Сохранить**, поле выбора паки для сохранения.

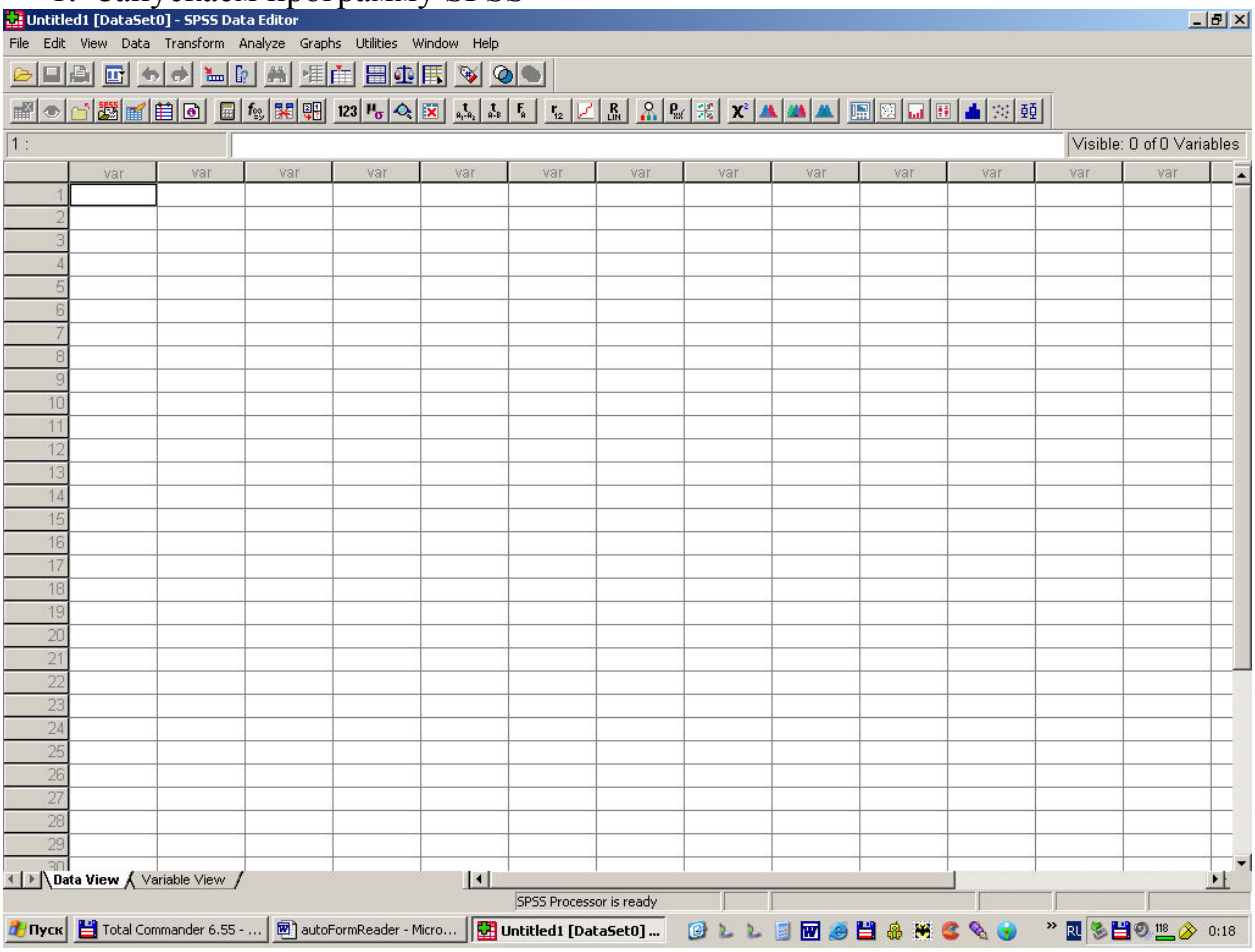
Данные будут сохранены в виде:



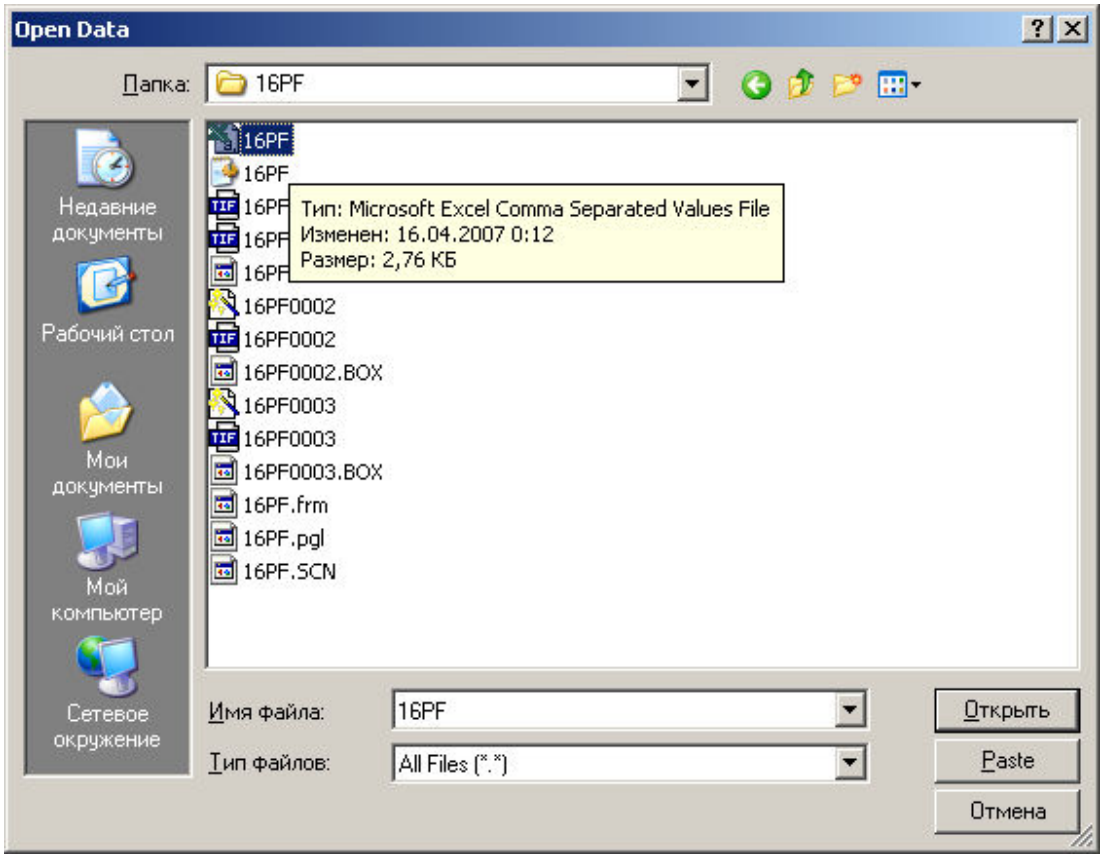

```
q1A;q1B;q1C;q2A;q2B;q2C;q3A;q3B;q3C;q4A;q4B;q4C;q5A;q5B;q5C;q6A;q6B;q6C;q7A;q7B;q
7C;q8A;q8B;q8C;q9A;q9B;q9C;q10A;q10B;q10C;q11A;q11B;q11C;q12A;q12B;q12C;q13A;q13B
;q13C;q14A;q14B;q14C;q15A;q15B;q15C;q16A;q16B;q16C;q17A;q17B;q17C;q18A;q18B;q18C;
q19A;q19B;q19C;q20A;q20B;q20C;q21A;q21B;q21C;q22A;q22B;q22C;q23A;q23B;q23C;q24A;q
24B;q24C;q25A;q25B;q25C;q26A;q26B;q26C;q27A;q27B;q27C;q28A;q28B;q28C;q29A;q29B;q2
9C;q30A;q30B;q30C;q31A;q31B;q31C;q32A;q32B;q32C;q33A;q33B;q33C;q34A;q34B;q34C;q35
A;q35B;q35C;q36A;q36B;q36C;q37A;q37B;q37C;q38A;q38B;q38C;q39A;q39B;q39C;q40A;q40B
;q40C;q41A;q41B;q41C;q42A;q42B;q42C;q43A;q43B;q43C;q44A;q44B;q44C;q45A;q45B;q45C;
q46A;q46B;q46C;q47A;q47B;q47C;q48A;q48B;q48C;q49A;q49B;q49C;q50A;q50B;q50C;q51A;q
51B;q51C;q52A;q52B;q52C;q53A;q53B;q53C;q54A;q54B;q54C;q55A;q55B;q55C;q56A;q56B;q5
6C;q57A;q57B;q57C;q58A;q58B;q58C;q59A;q59B;q59C;q60A;q60B;q60C;q61A;q61B;q61C;q62
A;q62B;q62C;q63A;q63B;q63C;q64A;q64B;q64C;q65A;q65B;q65C;q66A;q66B;q66C;q67A;q67B
;q67C;q68A;q68B;q68C;q69A;q69B;q69C;q70A;q70B;q70C;q71A;q71B;q71C;q72A;q72B;q72C;
q73A;q73B;q73C;q74A;q74B;q74C;q75A;q75B;q75C;q76A;q76B;q76C;q77A;q77B;q77C;q78A;q
78B;q78C;q79A;q79B;q79C;q80A;q80B;q80C;q81A;q81B;q81C;q82A;q82B;q82C;q83A;q83B;q8
3C;q84A;q84B;q84C;q85A;q85B;q85C;q86A;q86B;q86C;q87A;q87B;q87C;q88A;q88B;q88C;q89
A;q89B;q89C;q90A;q90B;q90C;q91A;q91B;q91C;q92A;q92B;q92C;q93A;q93B;q93C;q94A;q94B
;q94C;q95A;q95B;q95C;q96A;q96B;q96C;q97A;q97B;q97C;q98A;q98B;q98C;q99A;q99B;q99C;
q100A;q100B;q100C;q101A;q101B;q101C;q102A;q102B;q102C;q103A;q103B;q103C;q104A;q10
4B;q104C;q105A;q105B;q105C
0;0;1;0;1;0;0;1;0;0;1;0;0;0;1;0;1;0;0;1;0;1;0;0;1;0;0;1;0;0;1;1;0;0;1;0;0;0;1;0;0;1
;0;0;1;1;0;1;0;0;0;1;0;1;0;1;0;0;0;1;0;1;0;0;0;1;0;0;1;0;0;1;0;0;0;1;0;0;0;1;0;1;0;
1;0;0;0;1;0;0;0;1;0;1;0;1;0;0;0;1;0;0;1;0;0;1;0;1;0;0;0;1;0;1;0;0;0;1;0;0;0;1;0;1
;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;1;0;1;0;0;0;1;0;0;0;1;0;1;0;0;0;1;1;0;0;0;1;0;0;1;0;
1;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;1;0;0;0;1;0;1;0
;0;0;1;0;1;0;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;1;0;0;1;0;0;0;0;1;0;0;1;1;0
;0;1;0;0;0;0;1;0;0;1;0;0;1;0;0;1;0;1;0;1;0;0;1;0;0;1;0;0;0;1;0;0;0;1
0;0;1;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;1;0;0;1;0;0;1;0;0;1;1;0;0;1;0;0;0;1;0;0;1
;0;0;1;1;0;1;0;0;0;1;0;1;0;1;0;0;0;1;0;1;0;0;0;1;0;0;1;0;1;0;0;0;1;0;0;0;1;0;1;0;
1;0;0;0;1;0;0;0;1;0;1;0;1;0;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;1;0;0;0;1;0;0;0;1;0;1
;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;1;0;0;0;1;0;0;0;1;0;1;0;0;0;1;1;0;0;0;1;0;0;0;1;0;1
1;1;0;0;1;0;0;1;0;0;1;0;0;1;0;1;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;1;0;0;1;0;1;0;0;0;1;0;1;0
;0;0;1;0;1;0;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;0;0;1;0;0;1;1;0
0;0;1;0;0;1;0;1;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;1;0;0;0;0;1;0;0;1;1;0
;0;1;0;0;0;0;1;0;0;1;1;0;0;0;0;1;0;1;0;0;1;0;0;1;0;0;1;0;0;0;1;0;0;0;1
```

ШАГ 3. *Импортирование результатов ввода в пакет SPSS*

1. Запускаем программу SPSS

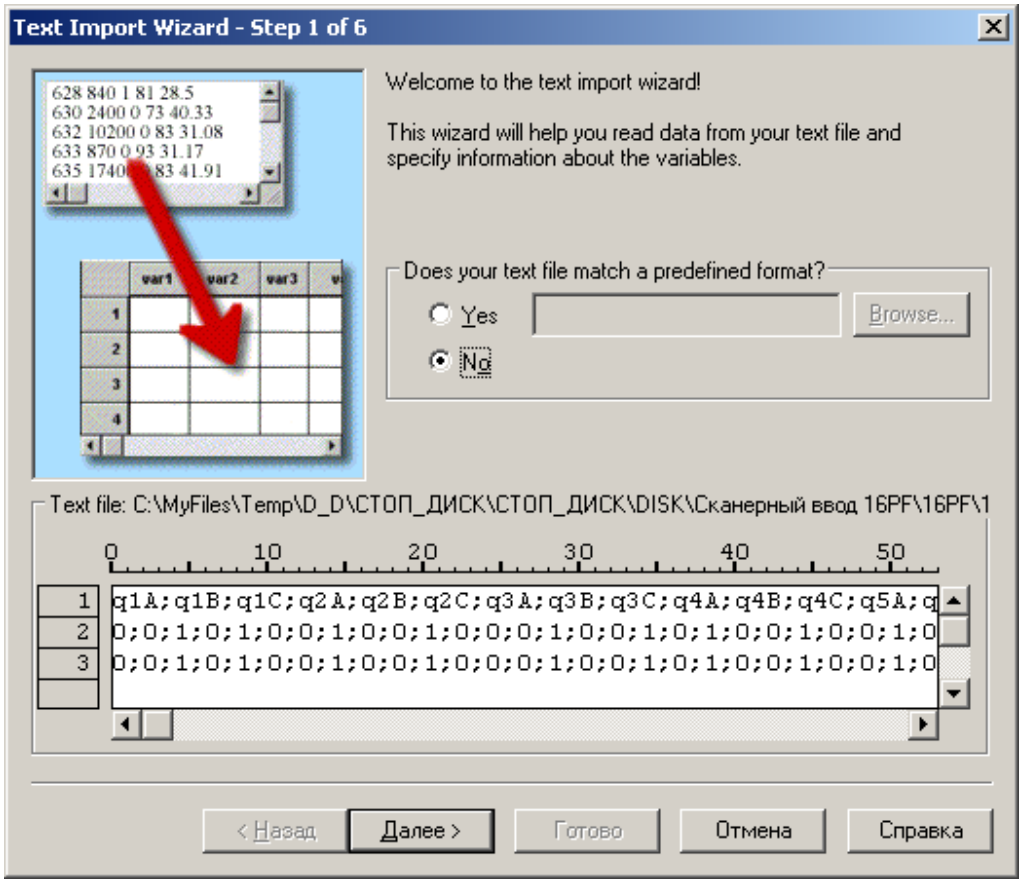


2. Выбираем в меню **File > Open > Data**



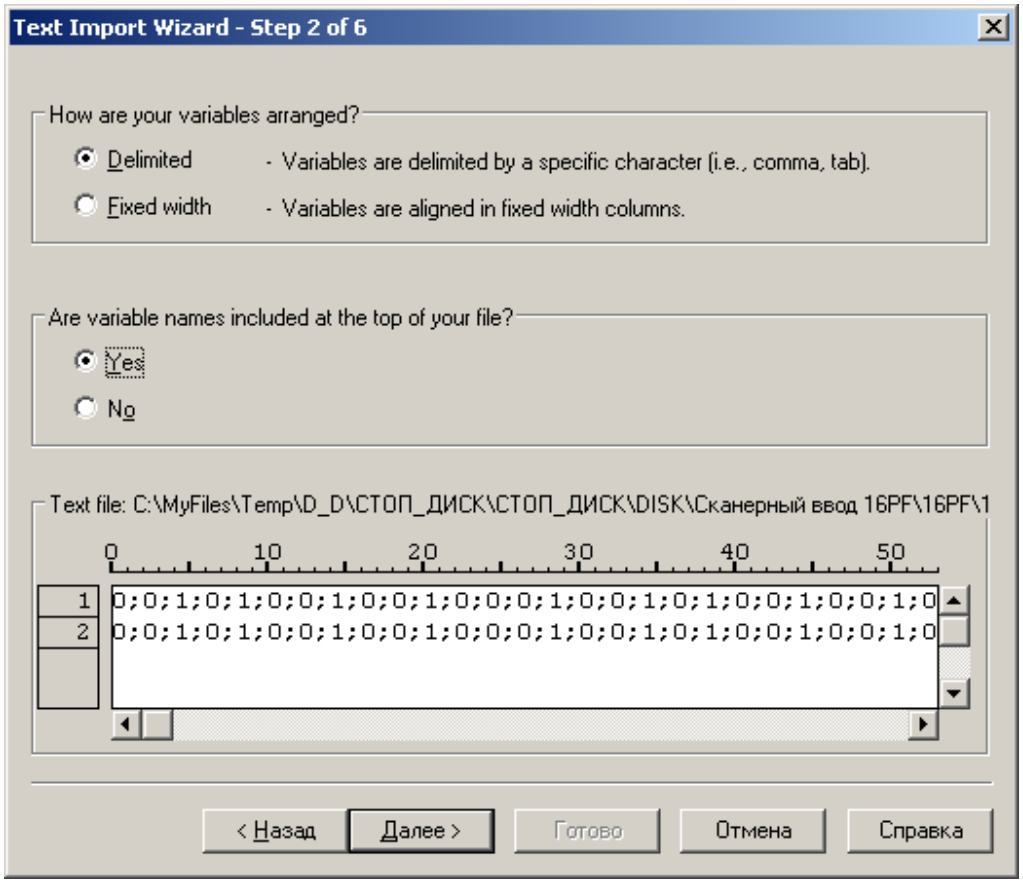
Установим тип файлов **All Files** (*.*). Выбираем файл 16PF.CSV (первый в нашем списке). Нажимаем **Открыть**.

3. Если сделали все правильно, то откроется мастер импорта файлов.



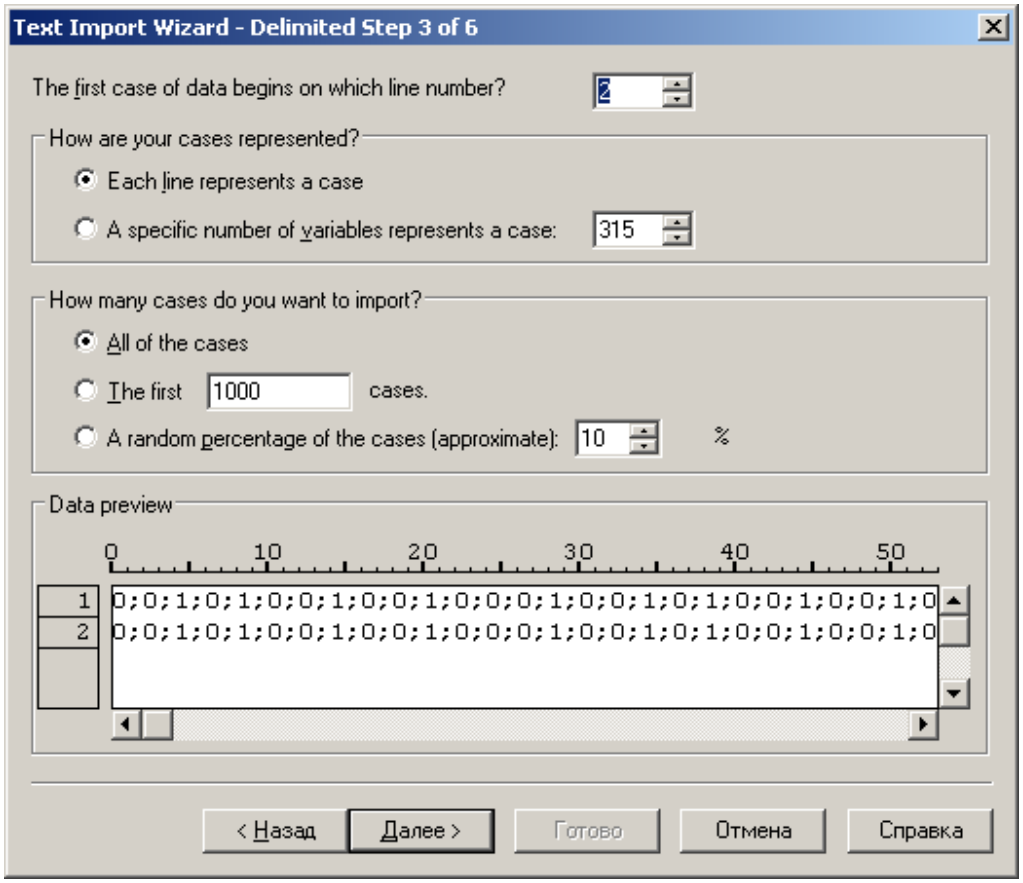
Нажимаем **Далее**

4. Укажем, что файл данных в первой строке содержит имена переменных (выбрав ответ **Yes**)

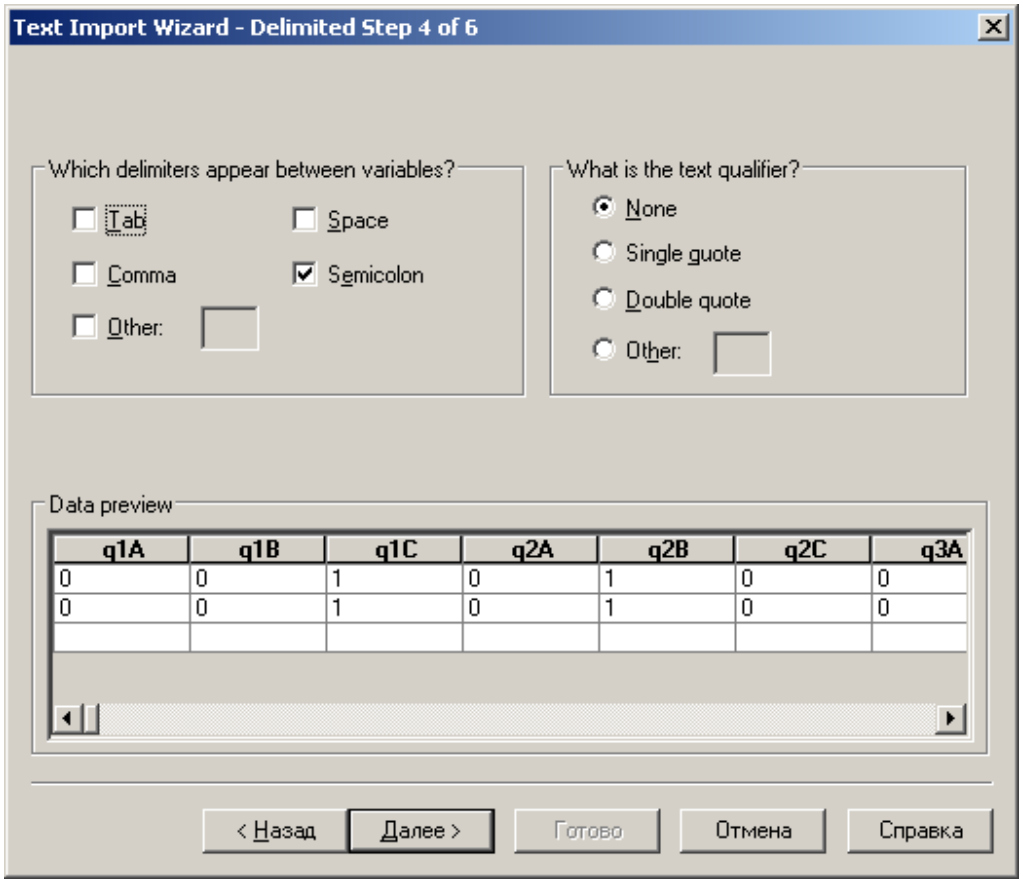


Нажимаем **Далее**

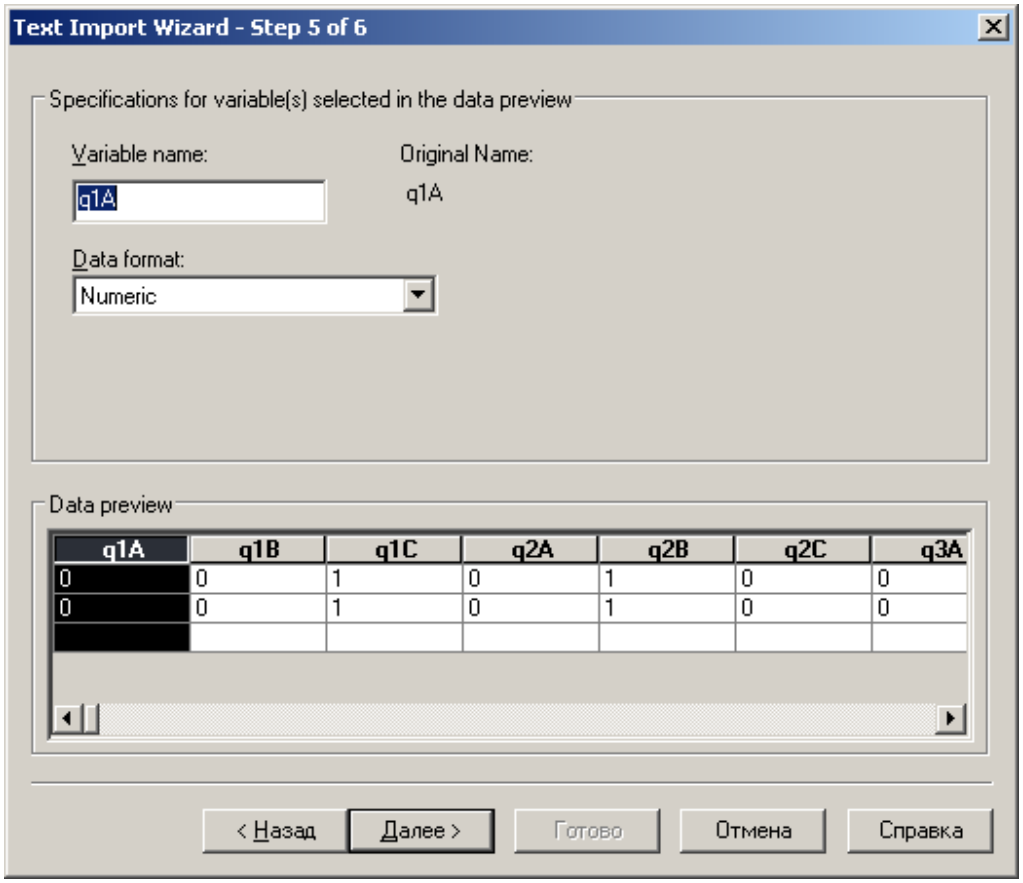
5. На этом шаге лишь нажимаем **Далее**



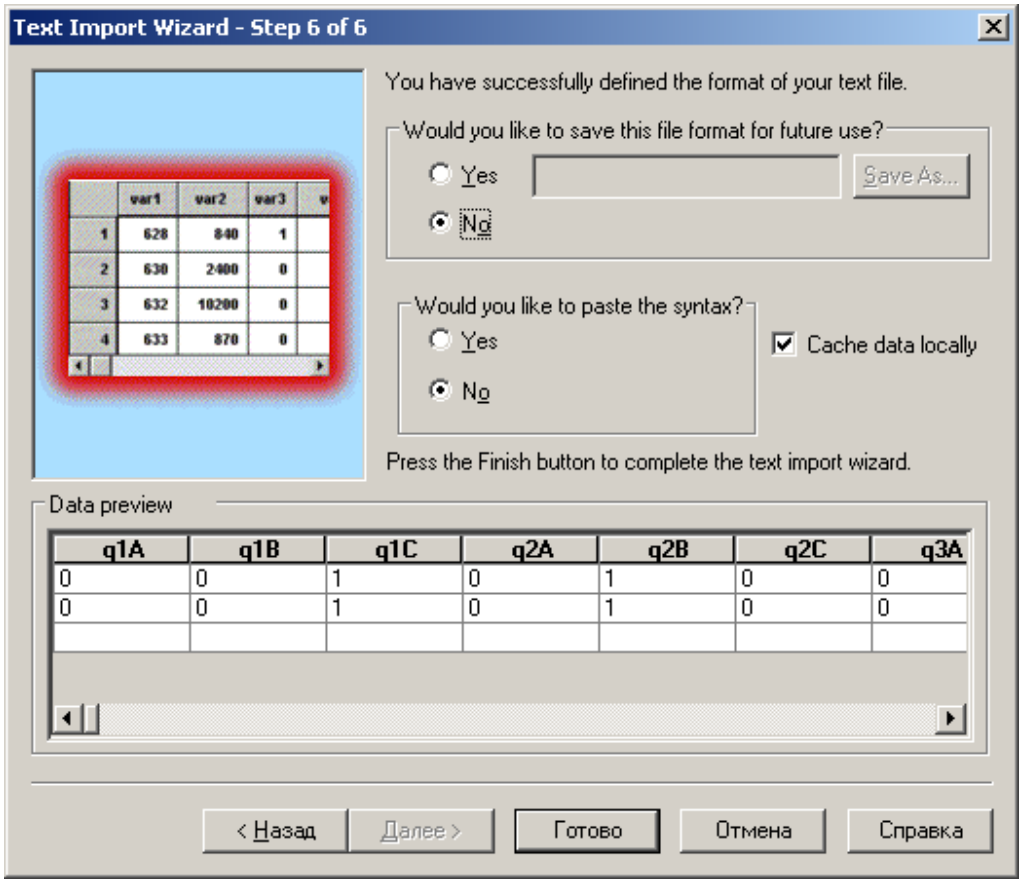
6. Здесь проверяем то, как будут выглядеть данные в SPSS. Если все правильно, то нажимаем **Далее**


Text Import Wizard - Delimited Step 4 of 6
Which delimiters appear between variables?
Tab
Space
Comma
Semicolon
Other:
What is the text qualifier?
None
Single quote
Double quote
Other:
Data preview
q1A q1B q1C q2A q2B q2C q3A
0 0 1 0 1 0 0
0 0 1 0 1 0 0
< Назад
Далее >
Готово
Отмена
Справка

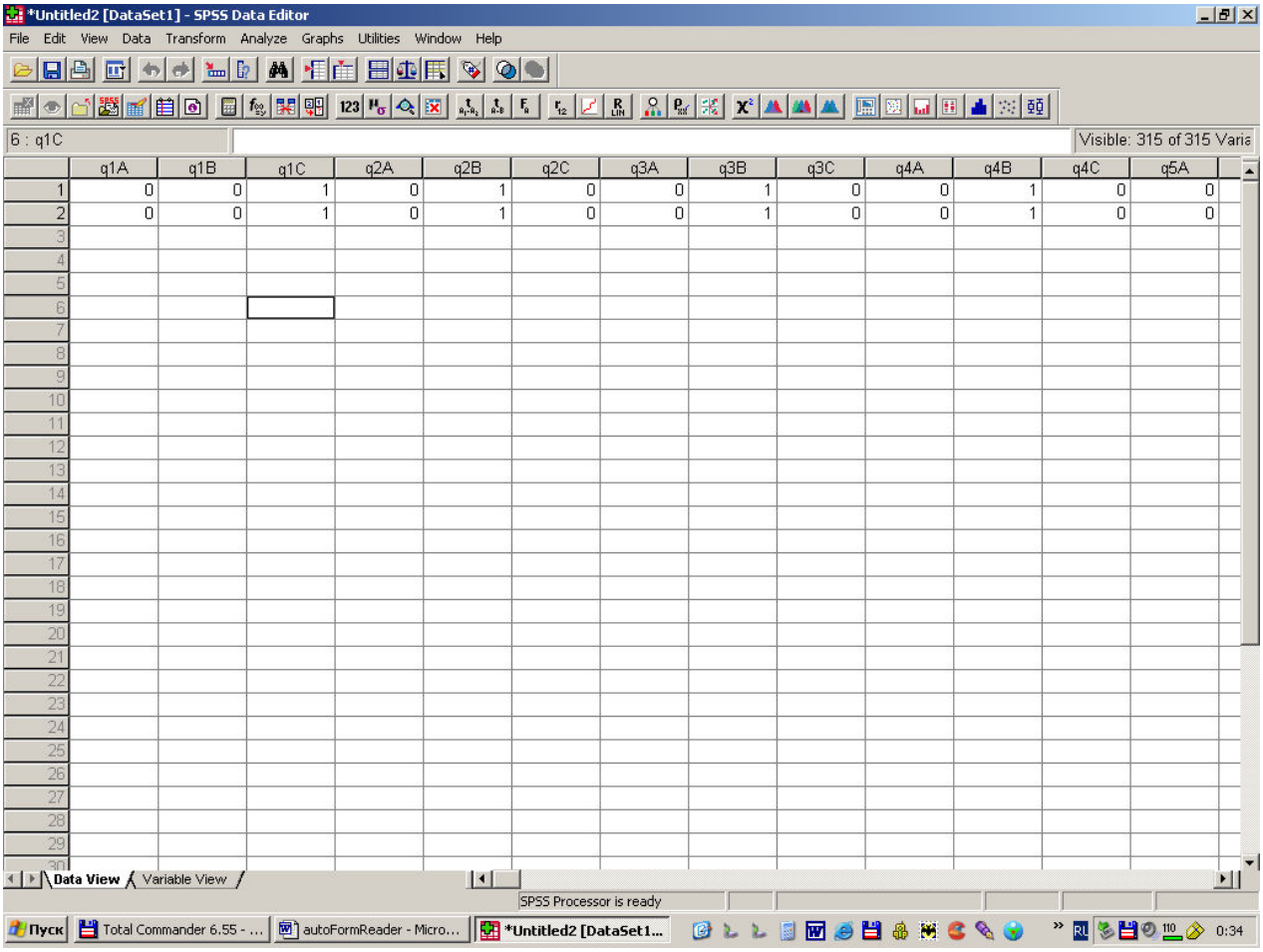
И



7. Нажимаем ГОТОВО

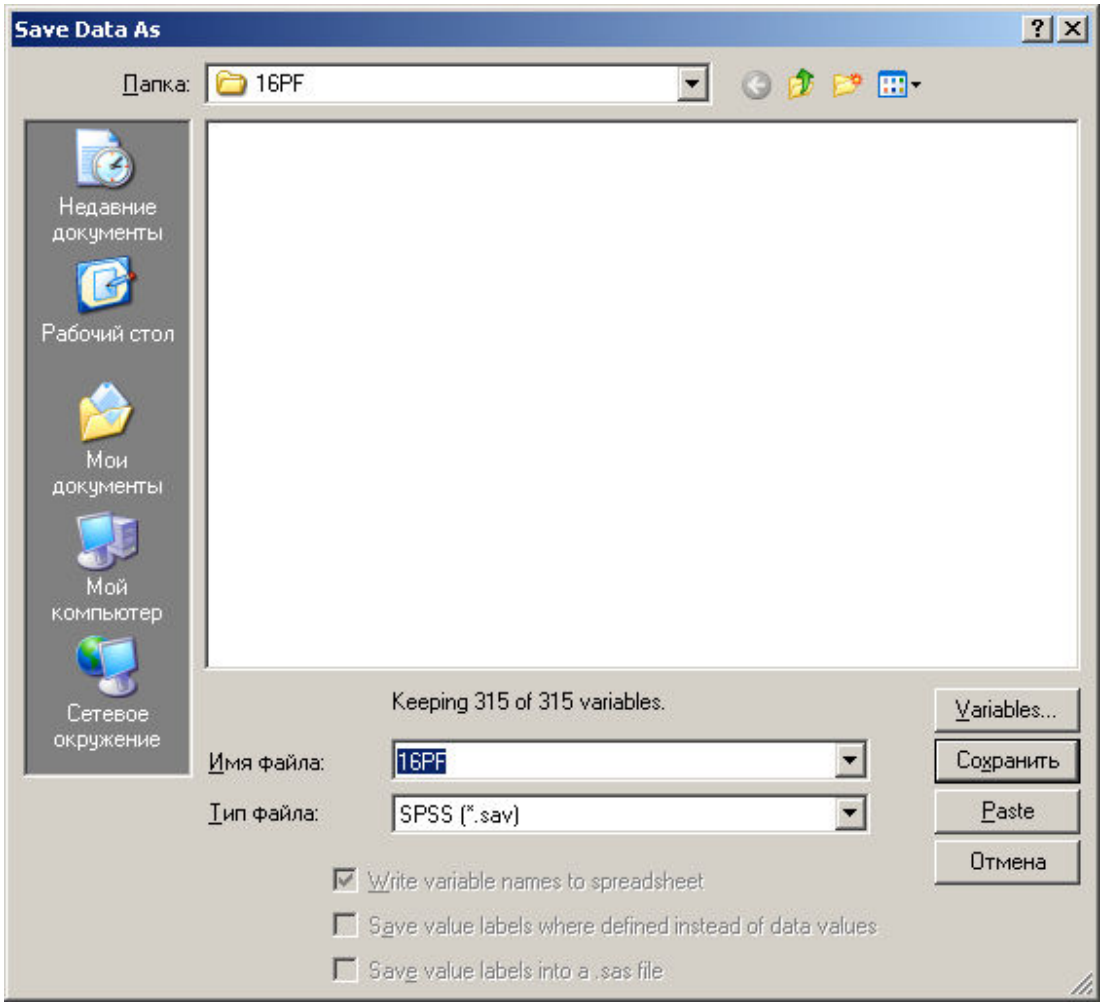


Данные импортированы в SPSS.



	q1A	q1B	q1C	q2A	q2B	q2C	q3A	q3B	q3C	q4A	q4B	q4C	q5A
1	0	0	1	0	1	0	0	1	0	0	1	0	0
2	0	0	1	0	1	0	0	1	0	0	1	0	0

8. Сохраним данные в формате SPSS, выбрав **File > Save file as…**

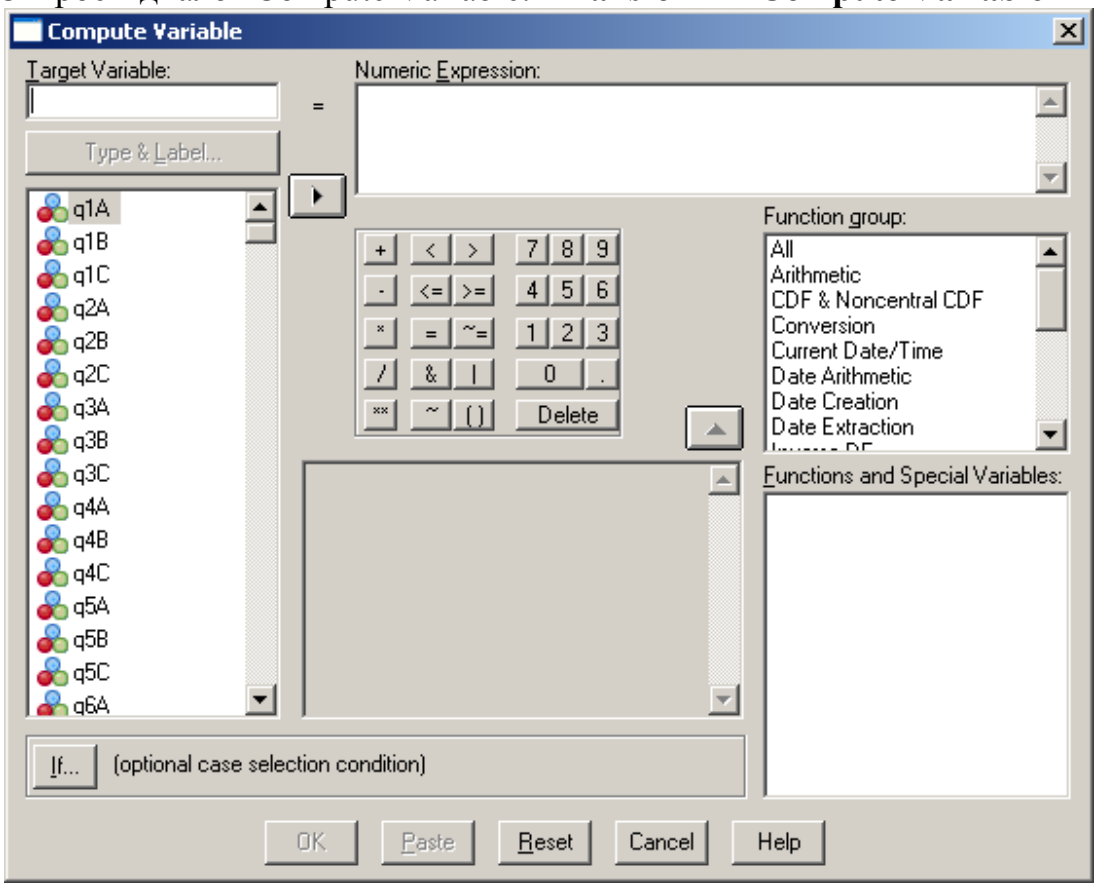


Зададим имя файла и папку для сохранения. Нажмем **Сохранить**.

ШАГ 4. *Процедура «наложение ключей» в виде синтаксиса SPSS*

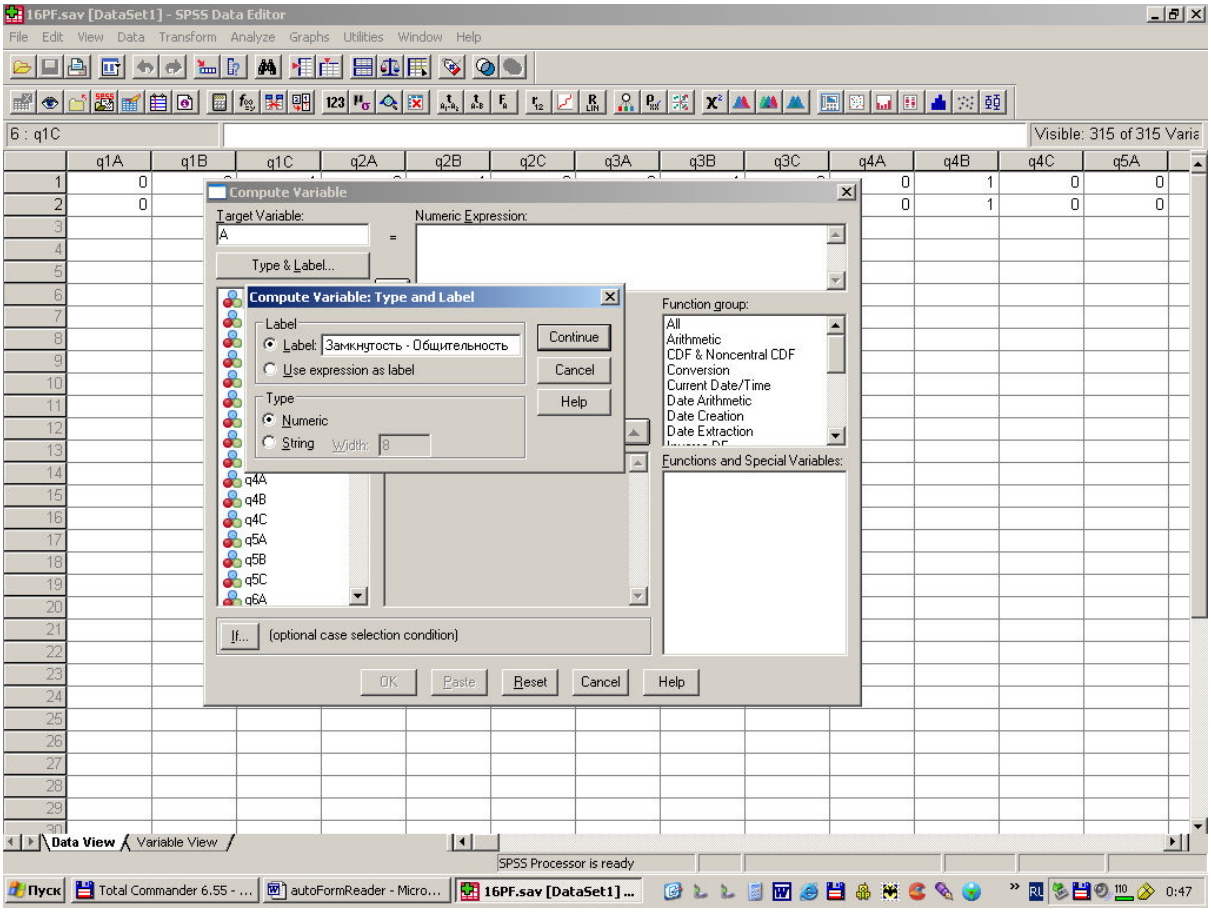
На этом шаге создадим синтаксис для обработки ответов по ключу и подсчета сырых баллов по каждой шкале опросника. Пример для шкалы А «Замкнутость – Общительность»

1. Откроем диалог Compute Variable: **Transform > Compute Variable**



2. Зададим необходимые поля диалога.

В **Target Variable** – введем **A**. Нажав кнопку **Type&Label** зададим метку переменной **A - Замкнутость – Общительность**. И нажмем **Continue**.

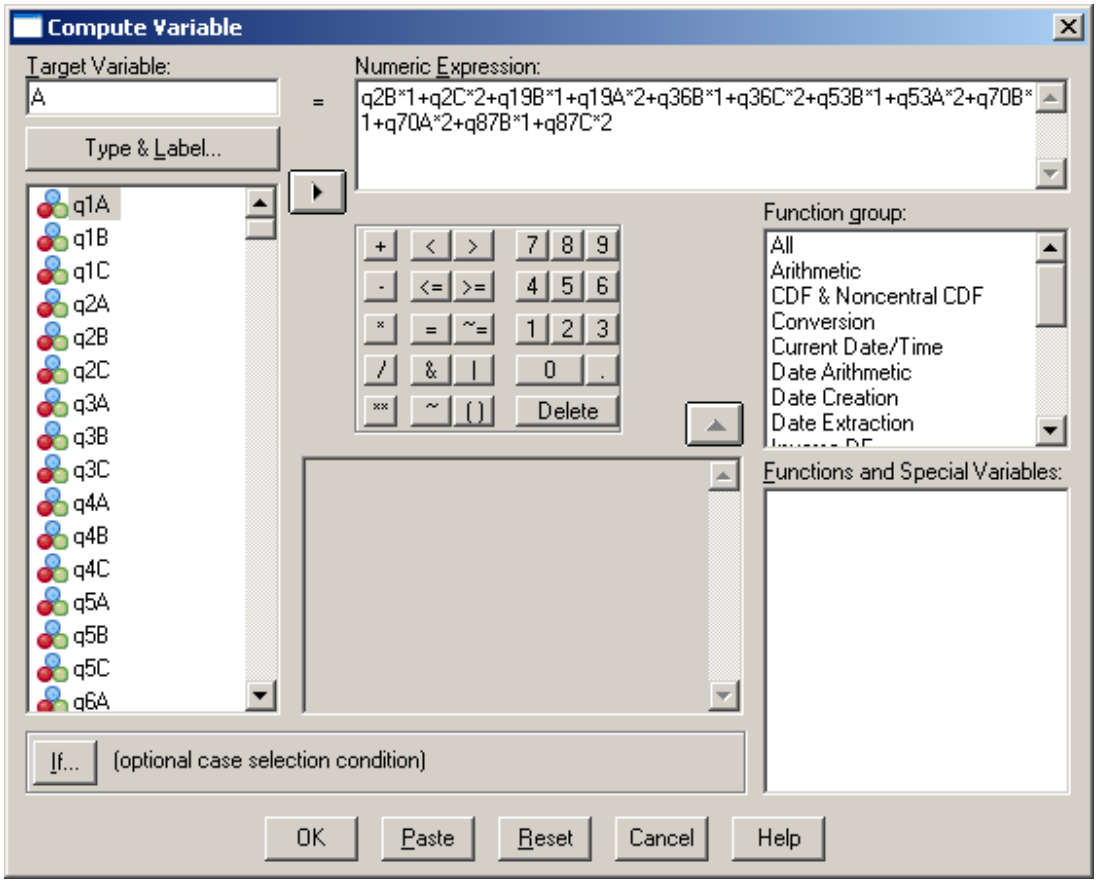


3. Ключ для получения сырого балла для это шкалы выглядит так:

	a–2	c–2	c–2	a–2	c–2	c–2	c–2
2. A	2. b–1 c–2	19. b–1 a–2	36. b–1 c–2	53. b–1 a–2	70. b–1 a–2	87. b–1 c–2	
3. B	3. b–1	20. c–1	37. b–1	54. c–1	71. a–1	88. c–1	104. a–1

Наша формула будет такой:
q2B*1+q2C*2+q19B*1+q19A*2+q36B*1+q36C*2+q53B*1+q53A*2+q70B*1+q70A*2+q87B*1+q87C*2,
Где q2B –ответ B на второй вопрос, а 1 – вес этого варианта (1 балл при совпадении с ключом), 2 – также вес варианта (2 балла при совпадении с ключом).

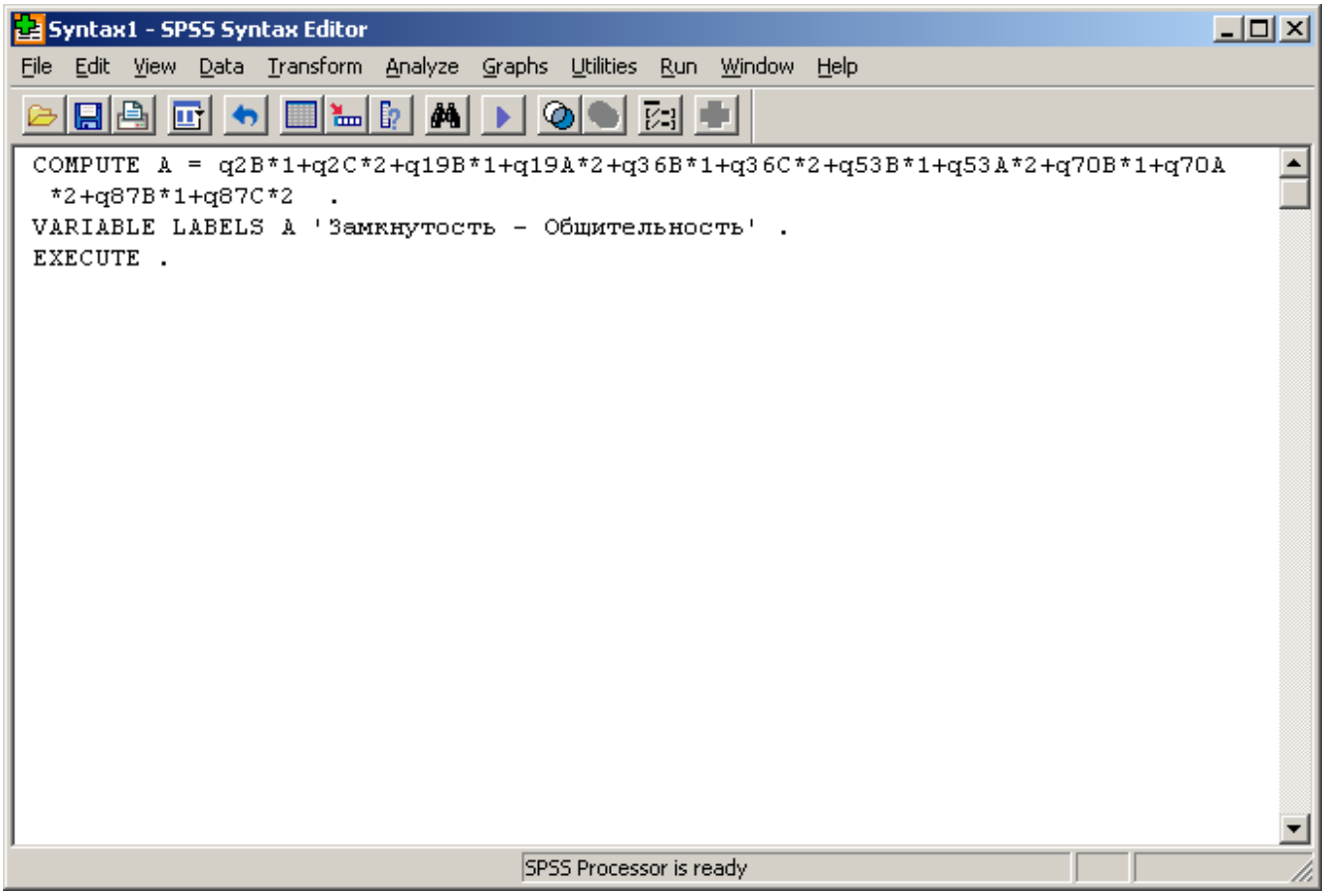
4. Зададим формулу подсчета баллов в поле **Numeric Expression**: Два раза кликая на соответствующей переменной-ответе и добавляя операции умножения и сложения.



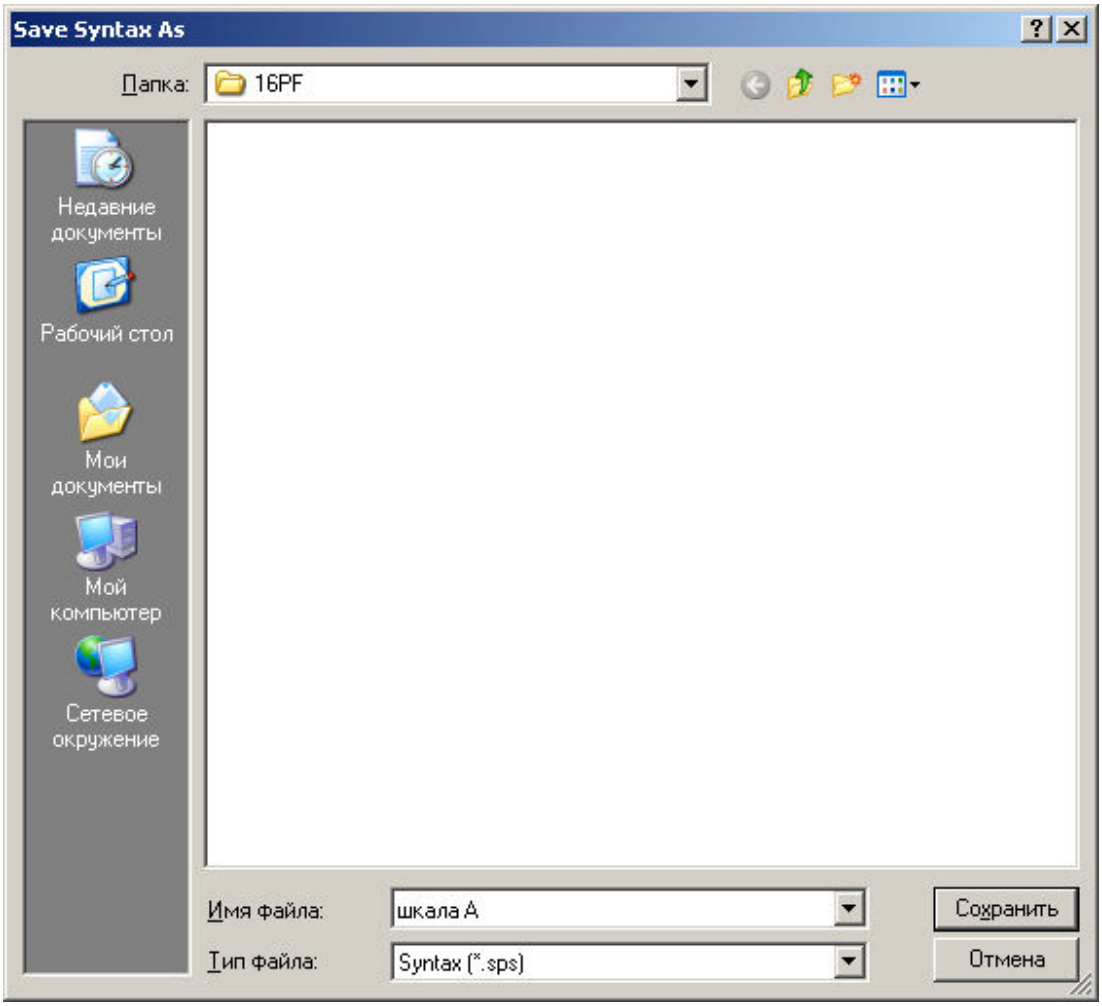
5. Нажимаем кнопку **Paste** для того, чтобы увидеть синтаксис для этой шкалы. Откроется редактор Синтаксиса.

Получим синтаксис:

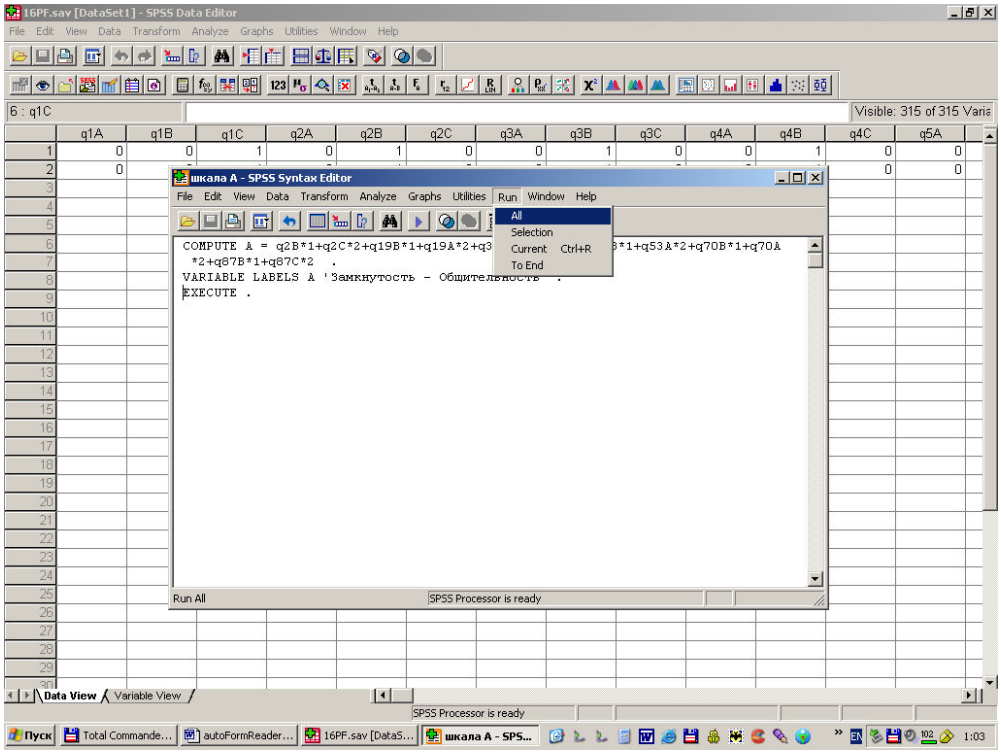
```
COMPUTE A =
q2B*1+q2C*2+q19B*1+q19A*2+q36B*1+q36C*2+q53B*1+q53A*2+q70B*1+q70A*
2+q87B*1+q87C*2 .
VARIABLE LABELS A 'Замкнутость - Общительность' .
EXECUTE .
```



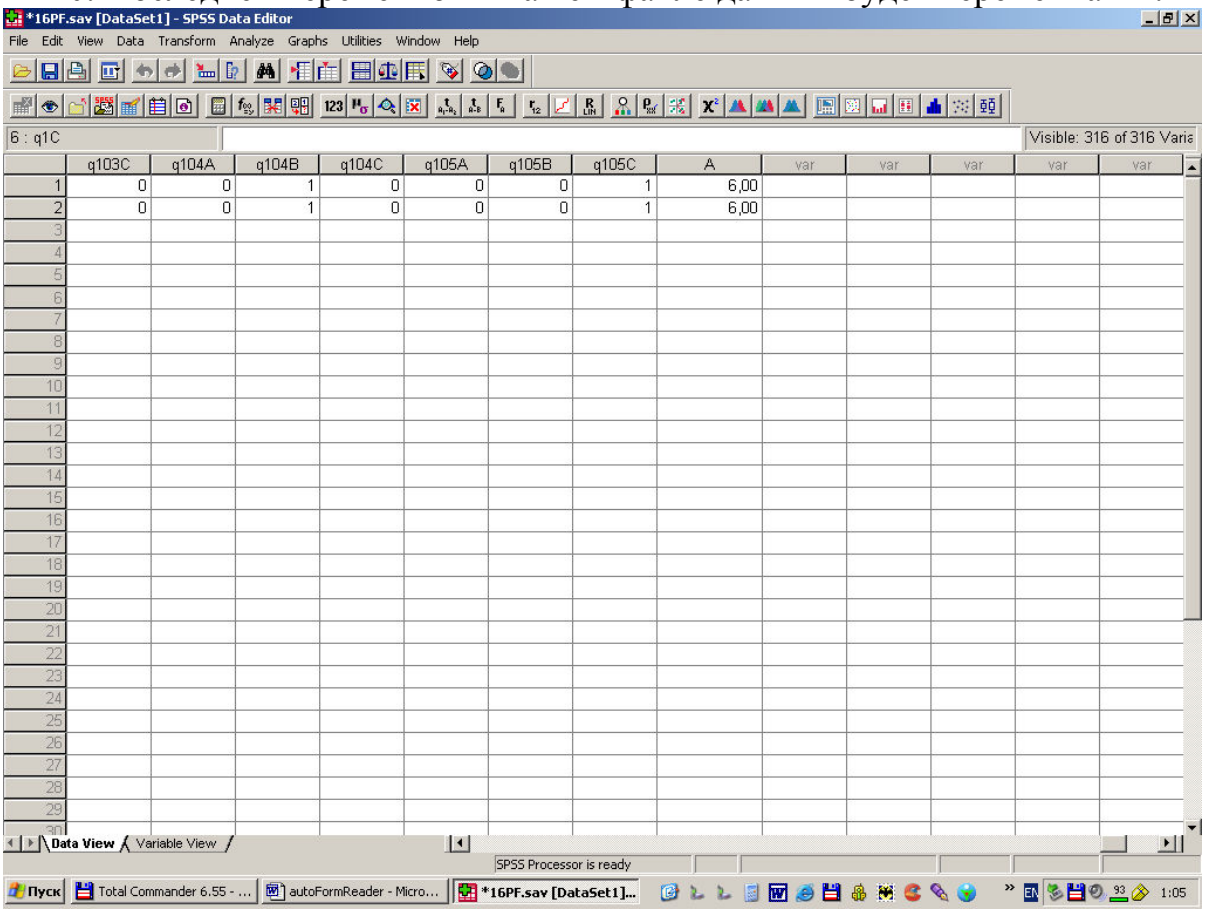
6. Сохраним синтаксис, чтобы его можно было использовать в дальнейшем при обработке других файлов с данными 16PF



7. Повторим предыдущие шаги для создания синтаксиса для оставшихся шкал.
8. Примеры синтаксисов для других шкал см. здесь.
9. Выполним синтаксис для шкалы А. **Run > All**

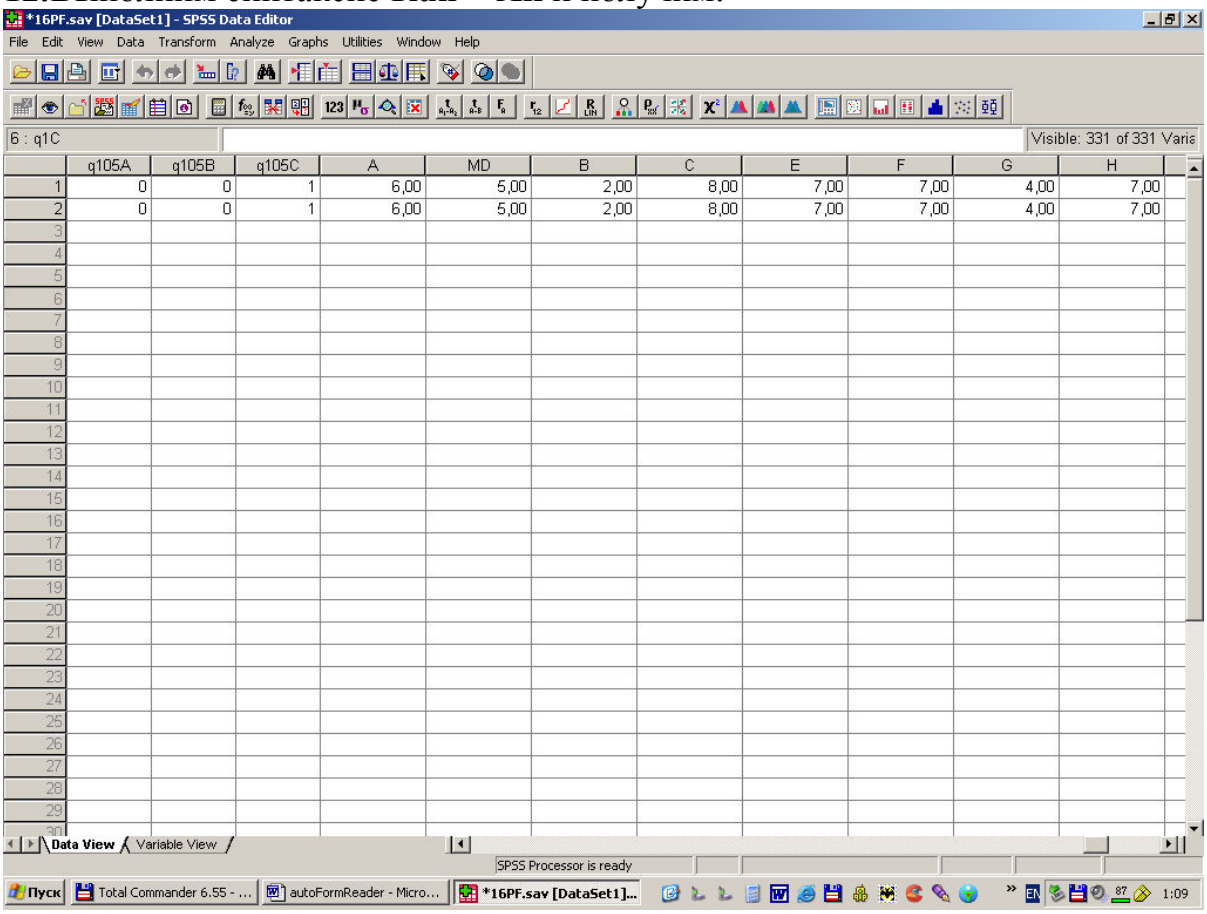


10. Последней переменной в нашем файле данных будет переменная А.



11. Откройте синтаксис для подсчета баллов по всем шкалам опросника Шкалы_все.sps

12.Выполним синтаксис **Run > All** и получим:



	q105A	q105B	q105C	A	MD	B	C	E	F	G	H
1	0	0	1	6,00	5,00	2,00	8,00	7,00	7,00	4,00	7,00
2	0	0	1	6,00	5,00	2,00	8,00	7,00	7,00	4,00	7,00

Переменные с ответами испытуемых можем удалить, оставив только переменные с баллами по шкалам. Не забудьте пометить файл как **только для чтения** и сохранить его. Успехов в обработке!

Обзор специально подготовлен для проекта
Флогистон: психология из первых рук http://flogiston.ru
Запрещается частичное или полное размещение текста
на других сайтах без разрешения автора или администрации Флогистона