
ПРИМЕНЕНИЕ ПАКЕТА STATISTICA 6 ДЛЯ ИССЛЕДОВАНИЯ НАЛОГОВОЙ ОТЧЕТНОСТИ ПРЕДПРИЯТИЙ НА ОСНОВЕ ЭМПИРИЧЕСКОГО МАТЕРИАЛА

Ясиненко Екатерина Александровна,

адъюнкт кафедры высшей математики Воронежского института
МВД России; katrin_yasya@mail.ru

В работе обосновано введение нового метода предварительной сортировки налоговой отчетности. При этом налоговая отчетность за несколько периодов рассматривается как временной ряд и изучается методами математической статистики.

Ключевые слова: модель тренда, временной ряд, автокорреляционная функция.

1. Создание эффективной системы отбора налоговых деклараций для проверки их правильности является одной из наиболее важной проблемой в повышении эффективности работы налоговых органов [1]. Существующая в настоящее время методика выбора документа для проверки представляет собой «пороговое правило». Эта и более сложные стратегии проверок и перепроверок сравнительно недавно обоснованы математическими моделями [1]. Эти модели имеют недостаток, они не содержат налоговую историю объекта налогообложения, определенного производства или фирмы. С другой стороны опытный инспектор, как правило, знает и учитывает в своей работе историю изменения налоговых выплат одного и того же производства. Современный принцип ротации работников налоговых органов действует, и новый работник уже может не иметь распределенных во времени сведений. На сбор этих сведений понадобятся средства и время.

Эта ситуация привела к следующему новому проектному варианту предварительного отбора деклараций: рассматривать налоговую отчетность определенного производства по годам как единое целое – как временной ряд. Такой подход позволяет методами статистического моделирования с одной стороны провести сортировку отчетов по конфигурации последовательности уровней, с другой математическими методами проверять предварительную гипотезу о возможном занижении налоговой суммы (гипотеза о правонарушении). Кроме того, после построения математической модели тренда эмпирического временного ряда, этот метод позволяет правильно вычислять недоплату налога, и аргументировано начислять штрафные санкции в случае обнаружения умышленного занижения налоговой

суммы. В работе [3] было показано, что эффективно можно использовать новый графико-аналитический инструмент предварительной сортировки отчетности, оптимизирующий процесс отыскания нарушений. Он основан на сравнении последнего по времени значения тренда с последним эмпирическим уровнем временного ряда отчетности. Заметим, что оптимальная стратегия проверки документов в налоговой практике определяется максимизацией чистой прибыли полученной государством от проверок, а не количеством пойманных преступников. Это основное отличие математических моделей оптимизации схем проверок налоговой отчетности, от схем определения некоторого нарушителя.

2. В настоящей работе рассматривается компьютерная технология анализа временных рядов налоговой отчетности репрезентативной выборки 5 предприятий с различной налоговой историей. В качестве компьютерного инструмента выступает пакет STATISTICA 6. Для определения вида тренда обычно используют стандартный набор функций. Основные типы функций для определения вида тренда: линейный, гиперболический, экспоненциальный, степенной, тренд, полиномиальный, логистический и др. Параметры каждого из перечисленных трендов можно определить методом наименьших квадратов (МНК). Определение вида тенденции (или проверка гипотезы об отсутствии тренда) это основной и ответственный этап исследования. Если тренд ищется в виде полинома, то его порядок можно определить методом Ньютона [3].

Вид тенденции можно определить, изучая тесноту связи между самой последовательностью эмпирических данных Y_1, Y_2, \dots, Y_n и последовательностью данных сдвинутых на τ единиц $Y_{1+\tau}, Y_{2+\tau}, \dots, Y_{n+\tau}$, или как говорят с лагом τ . Для этого вычисляют выборочный коэффициент автокорреляции:

$$r(\tau) = \frac{(n-\tau) \sum_{t=1}^{n-\tau} y_t \cdot y_{t+\tau} - \sum_{t=1}^{n-\tau} y_t \cdot \sum_{t=1}^{n-\tau} y_{t+\tau}}{\sqrt{(n-\tau) \sum_{t=1}^{n-\tau} y_t^2 - \left(\sum_{t=1}^{n-\tau} y_t \right)^2} \times \sqrt{(n-\tau) \sum_{t=1}^{n-\tau} y_{t+\tau}^2 - \left(\sum_{t=1}^{n-\tau} y_{t+\tau} \right)^2}}. \quad (1)$$

Напомним, что временной ряд называется строго стационарным, если совместное распределение вероятностей n наблюдений Y_1, Y_2, \dots, Y_n и совместное распределение вероятностей n наблюдений $Y_{1+\tau}, Y_{2+\tau}, \dots, Y_{n+\tau}$ одинаково. В этом случае коэффициент корреляции с лагом τ :

$$\rho(\tau) = \frac{M\left[\left(y_t - M[y_t]\right)\left(y_{t+\tau} - M[y_{t+\tau}]\right)\right]}{\sigma[y_t] \times \sigma[y_{t+\tau}]} \quad (2)$$

равен нулю, а его статистическая оценка (1) должна быть близка к нулю.

Для стационарного ряда с увеличением лага τ взаимосвязь членов временного ряда Y_t и $Y_{t+\tau}$ ослабевает и автокорреляционная функция $\rho(\tau)$ должна убывать по абсолютной величине.

3. Для составления аргументированной схемы применения компьютерной технологии для графико-аналитического исследования и сортировки налоговой отчетности автор просмотрел в налоговых органах репрезентативную выборку обезличенных данных конкретных предприятий определенной области по изменению НБ (налогооблагаемой базы) по годам. Отобраны 5 различных по динамике, но типичных представителей временных рядов налоговой отчетности по налогу на прибыль. Ясно, что нарушение (занижение) НБ в этом виде налогов равносильно занижению суммы налога.

Представим их таблицей:

	НБ 1	НБ 2	НБ 3	НБ 4	НБ 5
2002	4834000	6710	-	7998830	3695849
2003	10271689	2579861	17983205	14103561	24953611
2004	31399601	10208128	189801092	5252811	10422200
2005	54947842	16459747	27243044	23495805	12582849
2006	63952557	16913677	10199381	47342209	32004788
2007	60576858	26166339	-10953385	71613950	33177333
2008	68581970	44272351	0	13768028	17580377

Табличные данные для типичных по поведению временных рядов налоговой отчетности с помощью пакета STATISTICA 6 были переведены графический вид, а затем с помощью определенных команд в нужных окнах (см. приложение) построены тренды для прогнозирования возможных нарушений отчетности.

Предприятие № 1.

Визуально можно предположить, что имеет место линейная тенденция. Проверим эту гипотезу аналитически. Расчет коэффициентов автокорреляции с лагом 1 и 2 проведем, используя систему STATISTICA 6. Получаем: $r(1) \approx 0.904$, $r(2) \approx 0.616$. Эти данные подтверждают гипотезу о линейном тренде. Используя метод наименьших квадратов (МНК) найдем уравнение тренда $y_t - \bar{y} = k(t - \bar{t})$. Вычисляя значения

$$\bar{t} = \frac{n+1}{2}, \quad \bar{y} = \frac{\sum_{t=1}^n y_t}{n}, \quad (yt) = \frac{\sum_{t=1}^n y_t \times t}{n}, \quad (t^2) = \frac{(2n+1)(n+1)}{6},$$

$$k = \frac{(yt) - \bar{y} \times \bar{t}}{(t^2) - (\bar{t})^2} = \frac{n^2 - 1}{12}$$

получим: $y_t = 1.1585 \times 10^7 t + 4.25 \times 10^6$. Однако налоговую службу не интересует вопросы теоретического плана и пакет STATISTICA автоматически построит тренд (рис. 1). Рисунок показывает занижение отчетной суммы в последний год. Считать это занижения значимым, это значит достойным проверки? Такое решение заранее или после согласования принимают налоговые органы. В данной работе будем считать, что если отчетная

сумма меньше значения тренда (точка лежит ниже графика тренда), то предприятие надо дополнительно проверять.

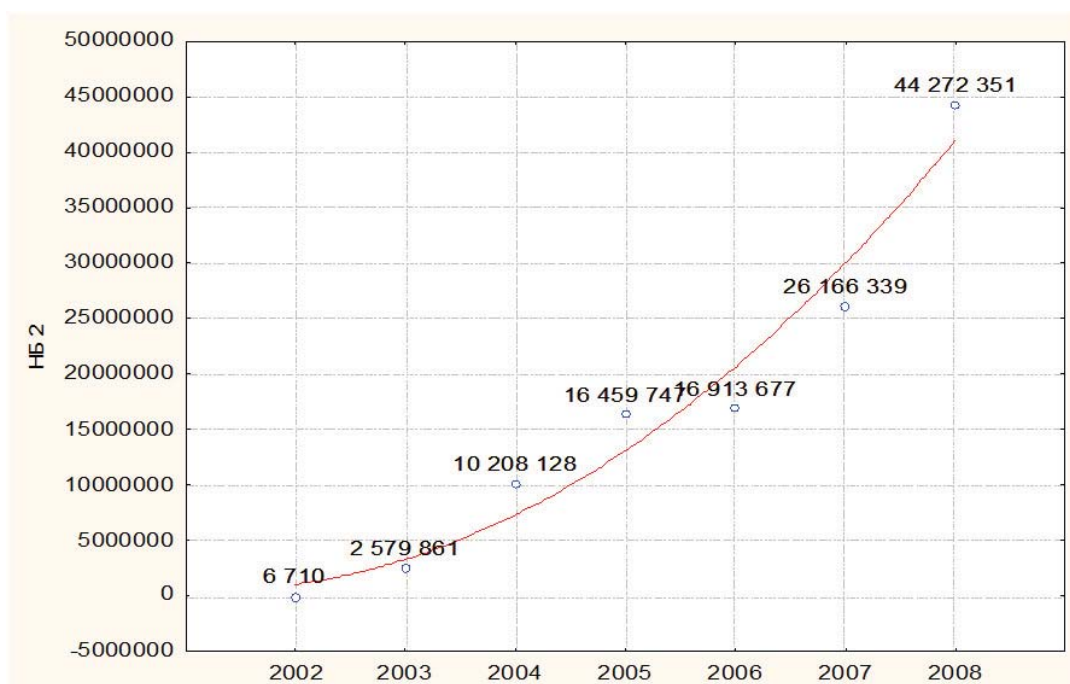


Рис. 1. Проверяемое значение расположено ниже графика, первое предприятие проверять.

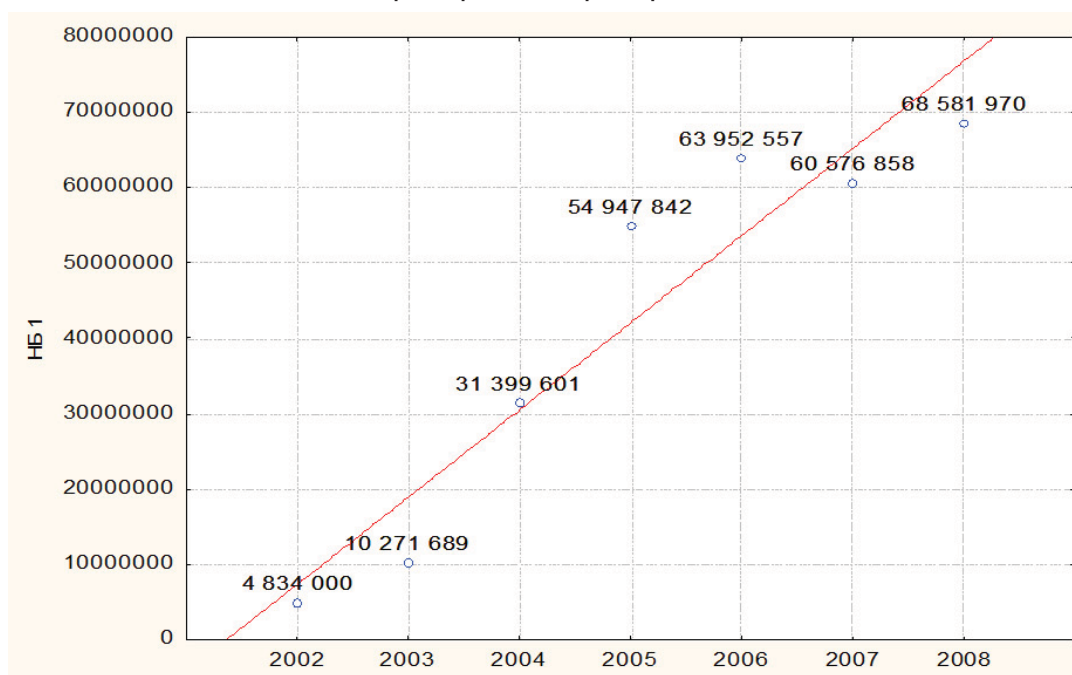


Рис. 2. Проверяемое значение расположено выше графика, второе предприятие не проверять.

Визуально можно предположить, что имеет место параболическая тенденция. С помощью стандартного применения компьютерной системы [3] получаем график параболы (в соответствии с Приложением в поле Fit выберем строку Polynomial) (рис. 2). Это тренд зависимости изменения налогооблагаемой базы. И последняя сумма явно не занижена, так как

расположена выше тренда. Это предприятие проверять не надо! Вообще общим правилом для всех предприятий не проводить проверки будет графическое расположение последней суммы выше уравнения тренда. Это правило оправдано математическими моделями статистики.

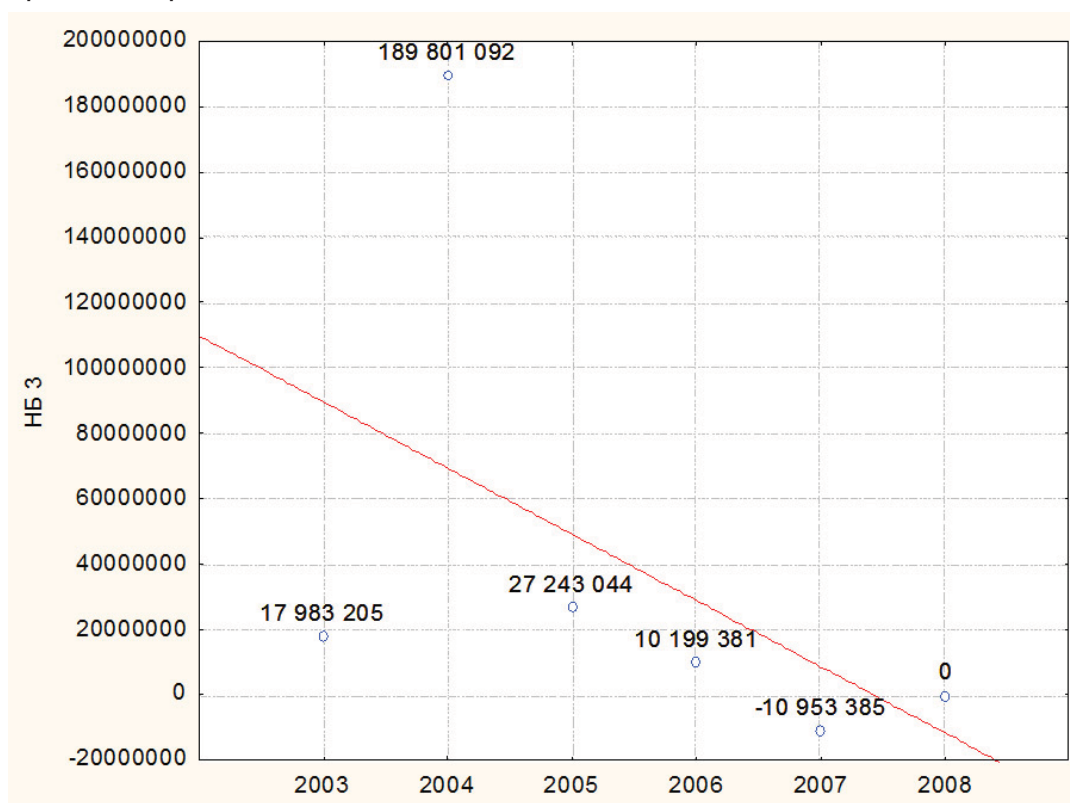


Рис. 3. Линейный тренд НБ третьего предприятия.

Стандартная программа в оболочке STATISTICA 6 определила следующую тенденцию – она убывающая линейная. Предприятие надо освободить от налога, возможно рассмотреть вопрос целесообразности данного производства.

Для предприятие № 4 визуально можно предположить, что имеет место параболическая тенденция. И последняя сумма явно занижена, так как расположена значительно ниже тренда. Согласно рисунку исторически (кроме последнего года) в среднем НБ данного предприятия развивалась как параболическая тенденция. Но это предположение спорно в силу малого объема множества уровней и оно приведет к большим штрафным санкциям. Также спорно применять статистические коэффициенты согласованности в силу большого разброса данных НБ по годам. Для упрощения программы сортировки смоделируем линейный тренд по уже показанной выше схеме. Упрощая вычисления применением стандартных программ (STATISTICA 6) имеем уравнение: $y_t = 0.622 \times 10^7 t + 0.135 \times 10^7$. Рис. 4 показывает значительное занижение отчетной суммы в 2008 году от прогнозируемой по тренду. Вычисления показывают, что НБ за 2008 год предприятия № 4 почти в 4 раза меньше прогнозируемой по тренду. Отклонение $d_1 \approx 0.7 \times f(7)$ естественно считать значительным, а данные по НБ этого предприятия необходимо подвергнуть дополнительной проверки.

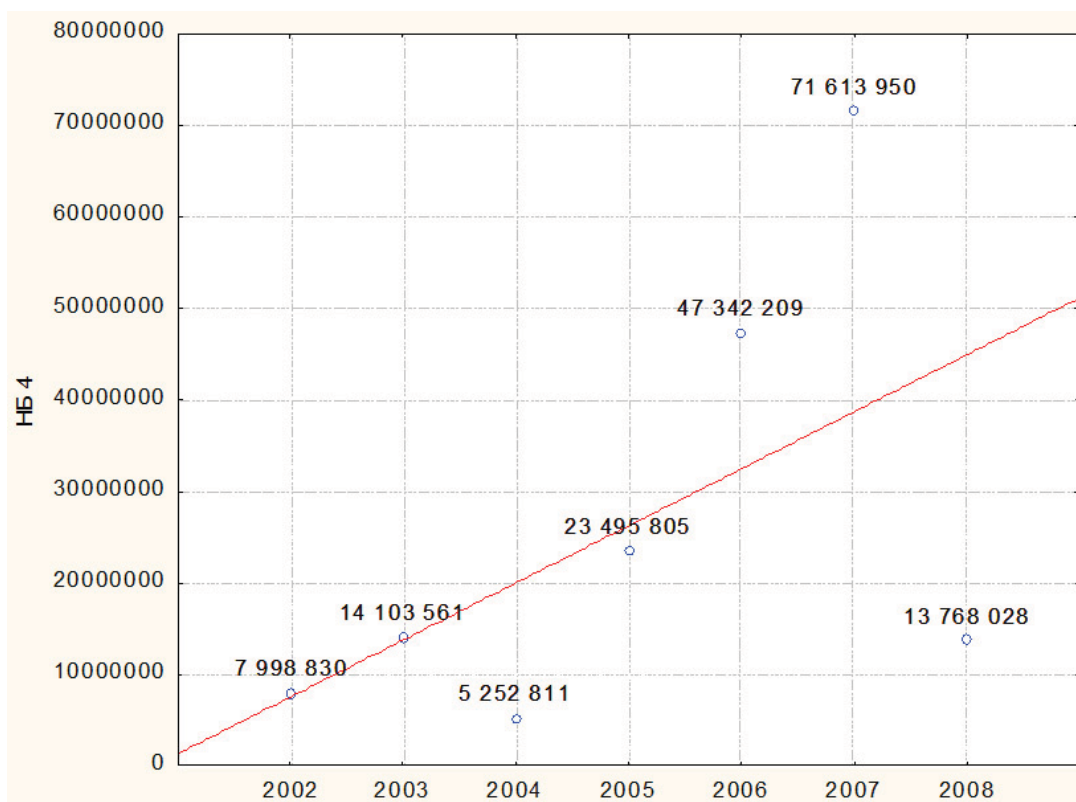


Рис. 4. Линейный тренд НБ четвертого предприятия.

Предприятие № 5.

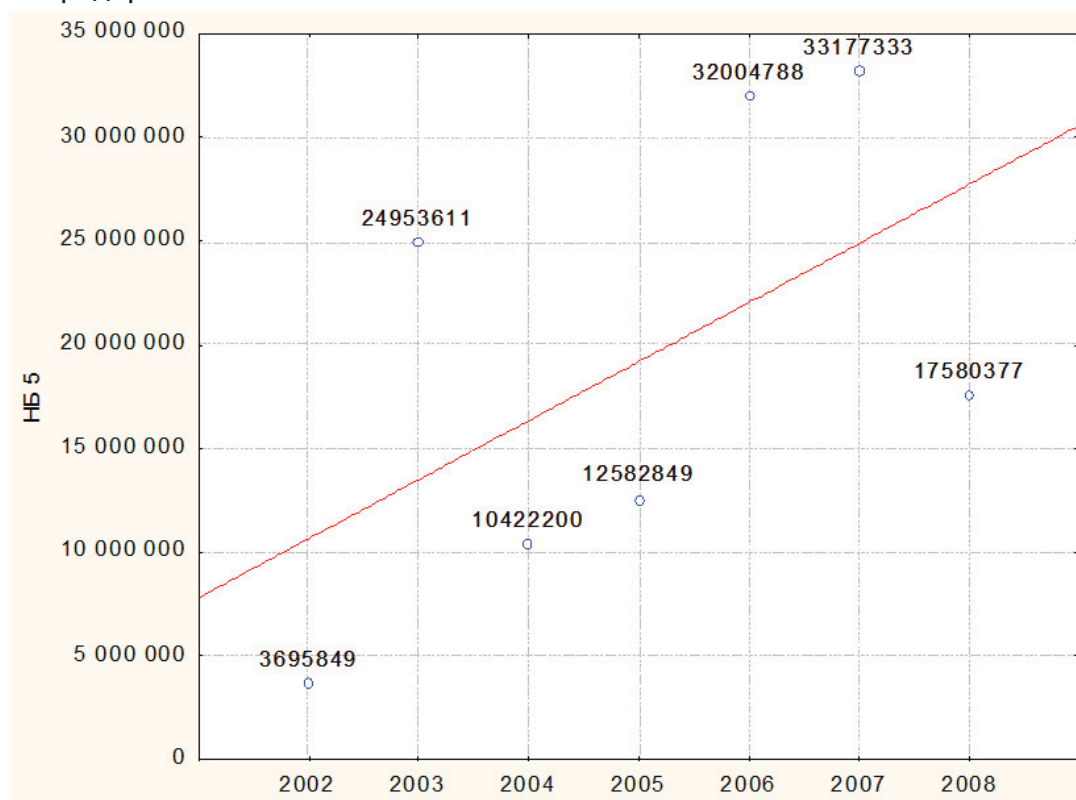


Рис. 5. Линейный тренд НБ пятого предприятия.

Пятое предприятие следует подвергнуть дополнительной проверке. Так как наблюдается большой разброс данных и последняя отчетность сильно занижена.

Выводы

Для предварительной сортировки налоговой отчетности с целью оптимизации поиска возможных нарушений минимальными средствами проделаем следующие операции.

1. Налоговую историю юридического лица представляем графически в виде точечного графика и рассматриваем как временной ряд.

2. Статистическими, компьютерными методами моделируем тренд. В силу незначительного по годам значимого объема информации временного ряда фактической отчетности для сортировки и приближенного вычисления коэффициента возможного штрафа по последней отчетности в большинстве случаев достаточно смоделировать линейный тренд.

3. Процедура проверок. Те предприятия, у которых фактическая отчетность по проверяемому году больше, чем значение тренда в проверяемом году без дополнительной информации можно не проверять. Те предприятия, у которых фактическая отчетность по проверяемому году меньше, чем значение тренда в проверяемом году следует подвергнуть дополнительной проверки.

4. Разность между значением тренда и фактической отчетностью в случае определения правонарушения может объективно служить для определения суммы штрафных санкций.

Приложение

Построение тренда в программе STATISTICA 6

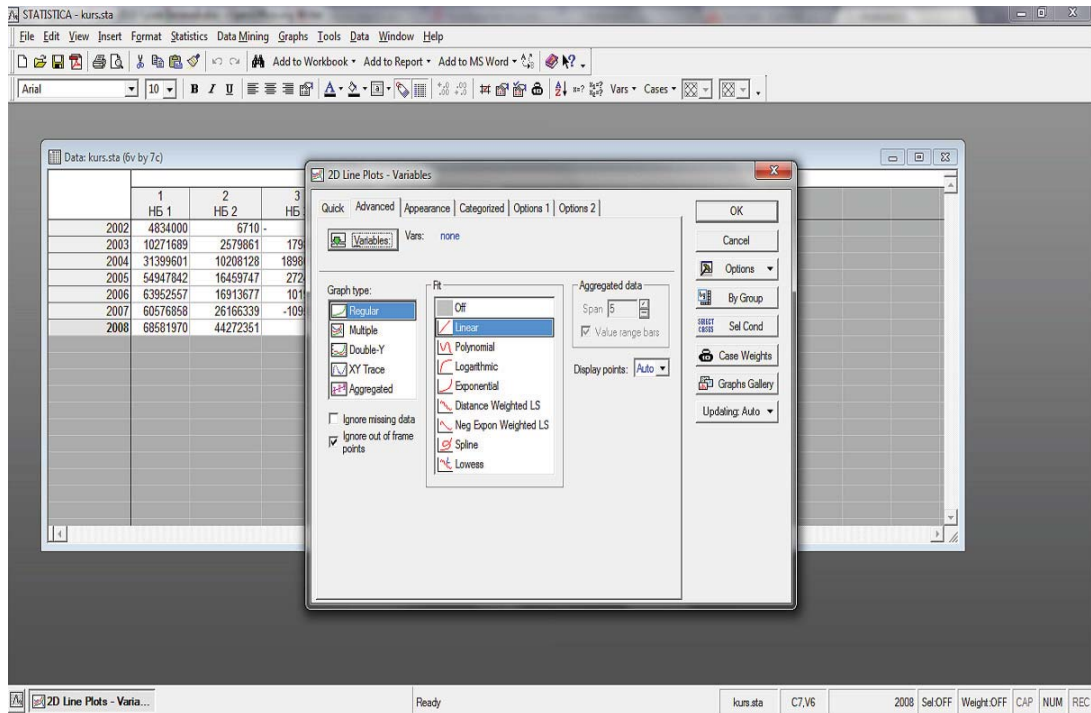
Введем наши данные в окно Data:

The screenshot shows the STATISTICA 6 software interface. The main window displays a data table with the following content:

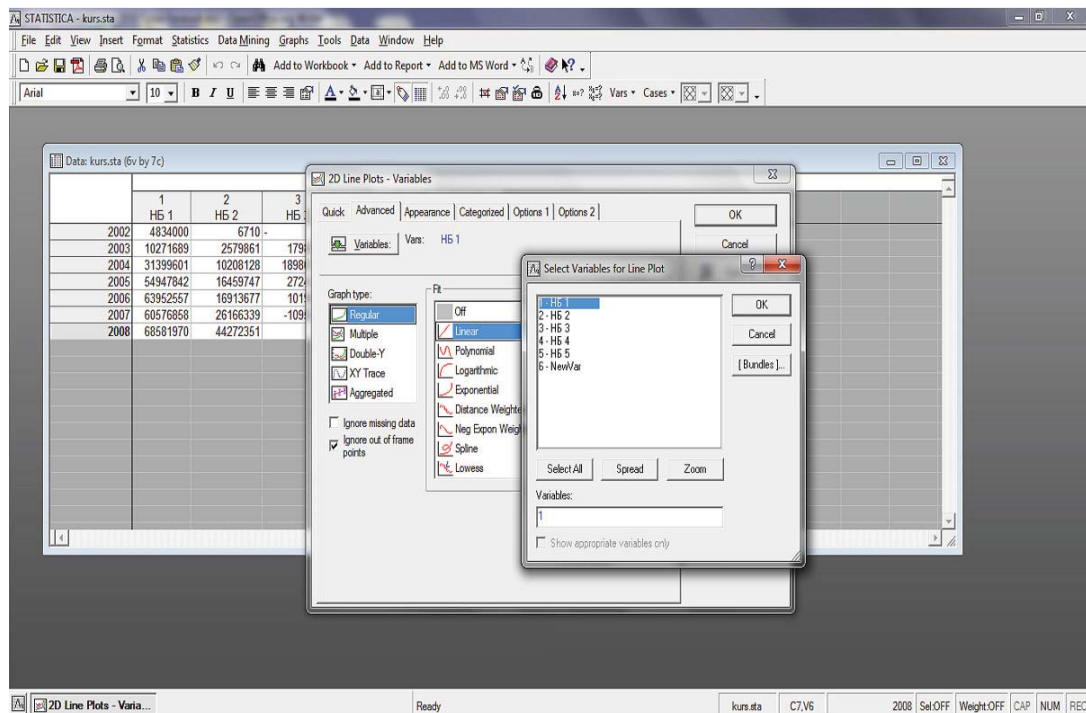
	1	2
	НБ 1	НБ 2
2002	4834000	6710
2003	10271689	2579861
2004	31399601	10208128
2005	54947842	16459747
2006	63952557	16913677
2007	60576858	26166339
2008	68581970	44272351

The 'Graphs' menu is open, showing various options. The '2D Graphs' sub-menu is expanded, and 'Line Plots (Variables)...' is selected. The status bar at the bottom indicates 'Creates 2D Line Plots of Variables' and shows the current data file 'kurs.sta' and variables 'C7.V6', '2008', 'Sel.OFF', 'Weight.OFF', 'CAP', 'NUM', and 'REC'.

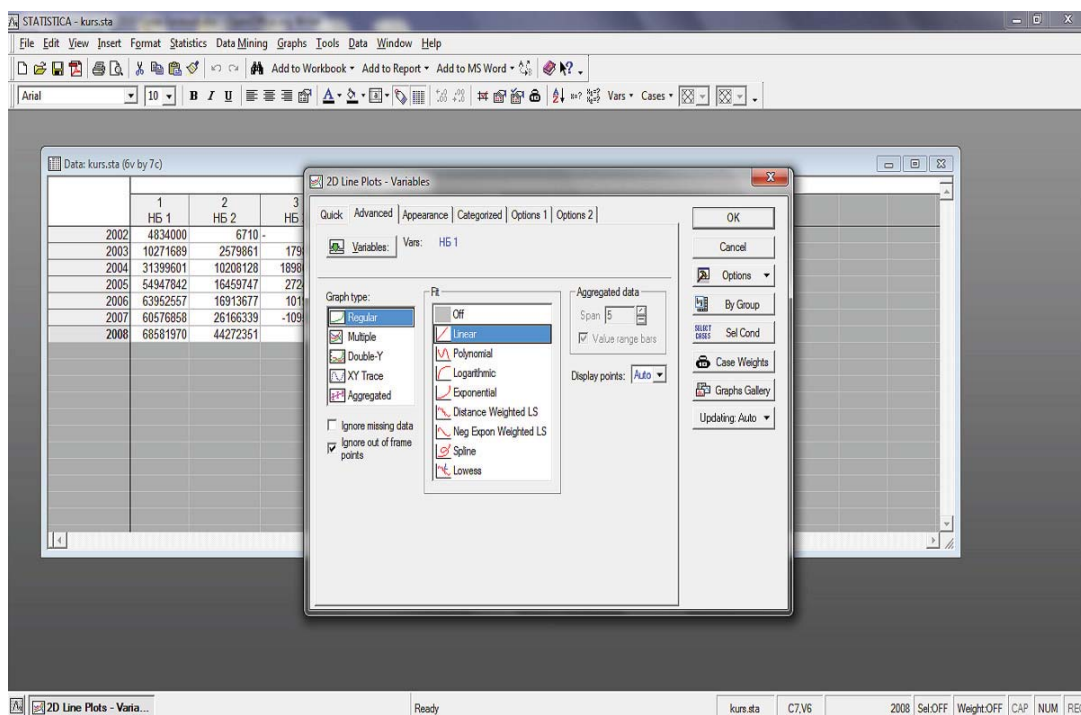
В окне 2D Plots — Variables перейдем в вкладку Advanced



В Variables выбираем переменную которую будем исследовать



Далее выбираем тип предполагаемого тренда (линейный, логарифмический...) в поле Fit и нажимаем Ок



Линия тренда появилась на графике благодаря нашему запросу: FIT – LINEAR. В пакете STATISTICA для построения тренда реализован метод наименьших квадратов – среди всевозможных прямых выбрана та, для которой сумма квадратов расстояний, измеренных вдоль оси Y, до выборочных точек минимальна.

Список источников

1. Методические рекомендации по проведению камеральных налоговых проверок деклараций по налогу на прибыль организаций. Утв. Приказ МНС России от 23 апреля 2003 г. № БГ-4-02/9.
2. Васин, А.А. Собираемость налогов и коррупция в налоговых органах [текст] / А.А. Васин, Е.И. Панова. – М.: РПЭИ. Фонд «Евразия», 1999. – с. 27.
3. Кутищева, Е.Ю. Ясиненко, Е.А..Статистическое моделирование для предварительной сортировки налоговой отчетности // Системы управления и информационные технологии. 1.1.(39), 2010. – с. 159-162.
4. Эконометрика. Под ред. Чл.-корр. РАН И.И. Елисеевой [текст] / И.И. Елисеева. – М.: «Финансы и статистика», 2008. – с. 275.
5. Халафян, А.А. STATISTICA 6 [текст] / А.А. Халафян. – М.: БИНОМ, 2010.

USAGE OF STATISTICA 6 FOR RESEARCH OF FISCAL DATA REPORTING OF ENTERPRIZES ON THE BASICS OF EMPIRICAL MATERIALS

Yasinenko Ekaterina Aleksandrovna,

Post-graduate student of the Chair of Theoretical Mathematics of Voronezh Institute of MVD; katrin_yasya@mail.ru

In work the introduction of a new method of preliminary sorting of the tax reporting is proved. Thus the tax reporting behind some periods is considered as a time series and is studied by methods of mathematical statistics.

Keywords: trend model, time series, auto correlation function.