

Министерство образования и науки, молодежи и спорта Украины

Харьковский национальный университет имени В. Н. Каразина

**Методические указания
по решению задач
экономико-математического
моделирования
в Microsoft Excel и Statistica**

Для студентов специальности 6.030201
«Международные экономические отношения»

Харьков – 2011

УДК 339.9: [330.4:004.9] (075.8)

ББК 65.5 с 515 я 73

М 54

Рецензенты: – д.экон.н., проф. Меркулова Т. В., зав.кафедрой экономической кибернетики и прикладной экономики Харьковского национального университета им. В. Н. Каразина

– д.т.н., проф. Торкатюк В. И., зав. кафедрой экономики строительства Харьковской национальной академии городского хозяйства

Рекомендовано к изданию Советом факультета международных отношений и туристического бизнеса Харьковского национального университета имени В. Н. Каразина, протокол № 8 от 24.03.2011 г.

Решение задач по экономико-математическому моделированию мирохозяйственных процессов в Microsoft Excel и Statistica : метод. пособ. / Сост. А. П. Голиков, О. Е. Галайда, Е. В. Ханова. – Х.: ХНУ им. В. Н. Каразина, 2011 – 46 с.

Пособие представляет собой руководство по экономико-математическому моделированию с помощью компьютеров в Microsoft Excel и Statistica, используемого на практических и лабораторных занятиях в курсе «Экономико-математическое моделирование мирохозяйственных процессов». Предназначено для студентов специальности 6.030201 «Международные экономические отношения». Может также использоваться студентами других экономических специальностей, прежде всего специальности «Международная экономика», а также аспирантами и преподавателями.

УДК 339.9: [330.4:004.9] (075.8)

ББК 65.5 с 515 я 73

©

ОГЛАВЛЕНИЕ

Введение.....	4
Раздел 1. Средние величины, отклонения, коэффициенты, индексы, прогрессии.....	5
1.1. Средняя арифметическая, средняя геометрическая, средняя гармоническая величины.....	5
1.2. Средние отклонения.	12
1.3. Коэффициент вариации.	14
1.4. Прогрессии.....	15
1.5. Индекс меры несходства.	17
Раздел 2. Корреляционный анализ	21
Раздел 3. Регрессионный анализ.....	23
Раздел 4. Решение систем уравнений.....	26
Раздел 5. Решение транспортной задачи линейного программирования.....	28
Раздел 6. Кластерный анализ	34

ВВЕДЕНИЕ

Методические указания являются дополнением к учебному пособию А. П. Голикова «Экономико-математическое моделирование мирохозяйственных процессов» – К. : Знания, 2008. – 222 с.

Разработаны для студентов специальностей «Международные экономические отношения», «Международная экономика». Изложенные в пособии методические приемы вычислительных и графо-аналитических работ могут использоваться студентами других экономических специальностей, а также преподавателями и научными работниками при выполнении различных расчетно-графических построений. Содержит необходимые рисунки, графики таблицы.

При разработке методических указаний использовалась версия Microsoft Excel 2007 и Statistic 6.0. При использовании других версий могут быть незначительные отличия в оформлении, но содержание формул и функций не изменится.

Проф. А. П.Голиков

Раздел 1. Средние величины, отклонения, коэффициенты, индексы, прогрессии

1.1. Средняя арифметическая, средняя геометрическая, средняя гармоническая величины

Для нахождения средней величины нам необходимо на листе Microsoft Excel ввести данные, по которым будем находить среднюю.

	A	B	C	D	E
1	35				
2	12				
3	45				
4	334				
5	69				
6	7				
7					
8					
9					
10					
11					

Рис. 1.1. Ввод данных

Выбрав ячейку для отображения результата, нажимаем на кнопку  («Вставить функцию»). При нажатии кнопки «Вставить функцию» на экране появляется новое окно – «Мастер функций» (рис. 1.2).

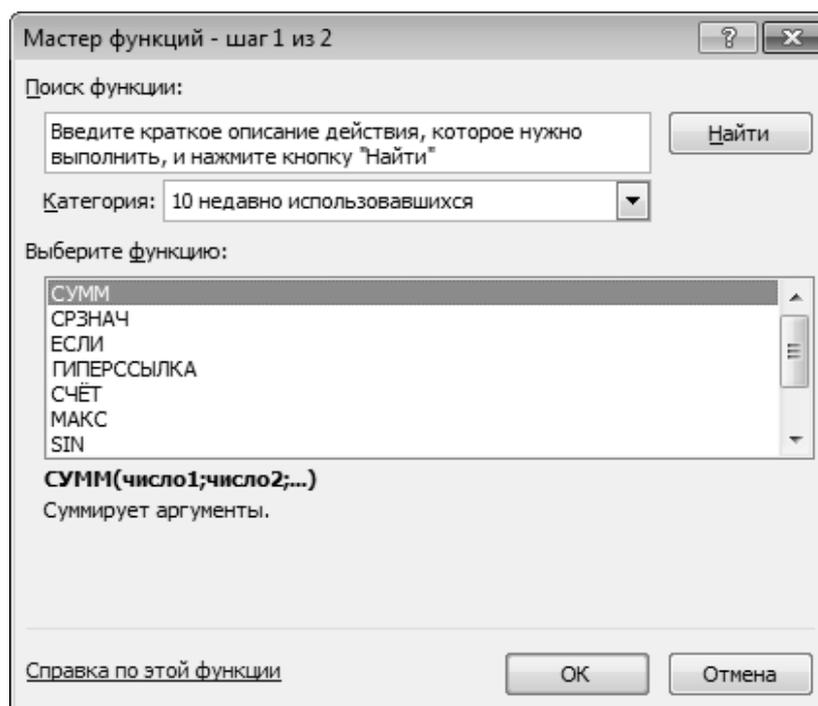


Рис. 1.2. Мастер функций

В появившемся окне в разделе «Категория» необходимо выбрать подходящий пункт.

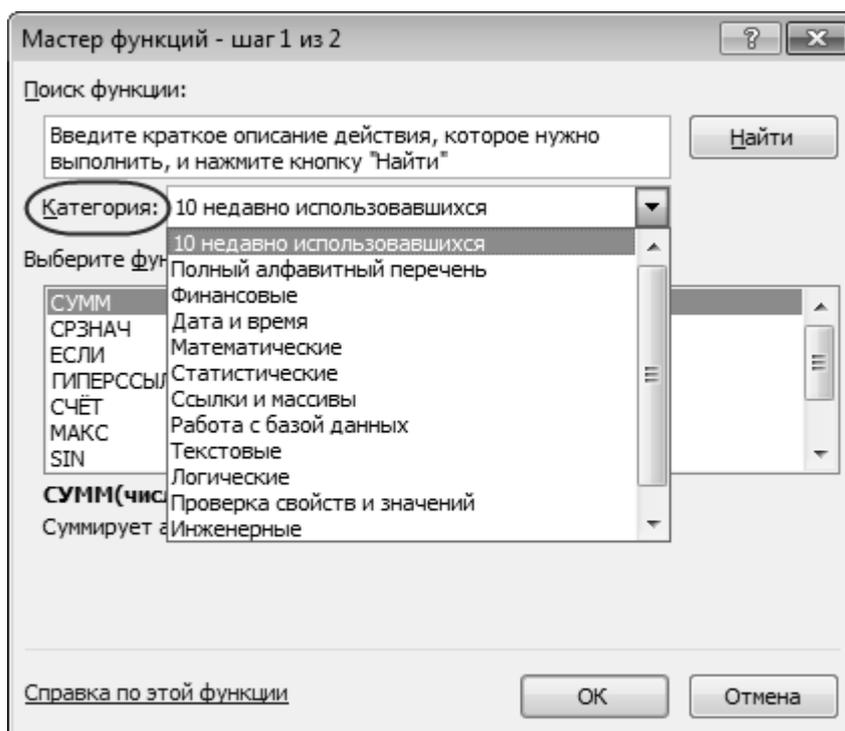


Рис. 1.3. Мастер функций; категория

Функции средних значений мы скорее всего найдем в разделе «Статистические» (рис. 1.4), однако можно выбрать раздел «Полный алфавитный перечень», где хранятся практически все функции, которыми мы можем воспользоваться в Microsoft Excel.

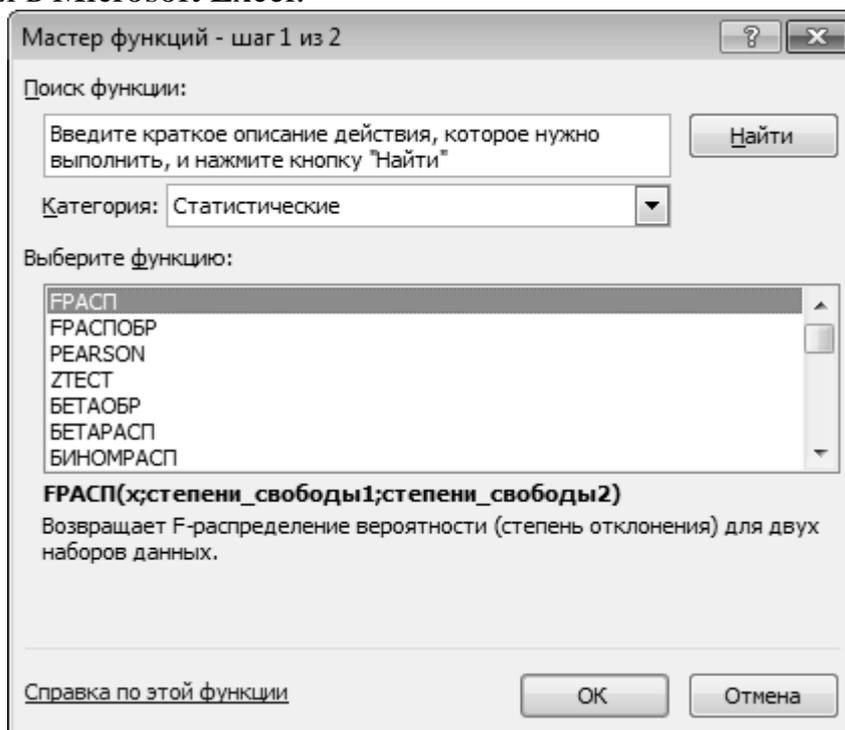


Рис. 1.4. Мастер функций, категория «Статистические»

Microsoft Excel предлагает функции для нахождения следующих средних величин:

- СРГАРМ – возвращает среднее гармоническое для множества положительных чисел – величину, обратную среднему арифметическому обратных величин;

- СРГЕОМ – возвращает среднее геометрическое для массива или диапазона из положительных чисел;

- СРЗНАЧ – возвращает среднее арифметическое своих аргументов, которые могут быть числами, именами, массивами или ссылками на ячейки с числами.

Найдем *среднюю арифметическую* величину для данных, введенных в Excel. Выберем ячейку для отображения результата, например ячейку С1. Далее нажимаем на кнопку «Вставить функцию», в категории «Статистические» выберем функцию «СРЗНАЧ» (рис. 1.5).

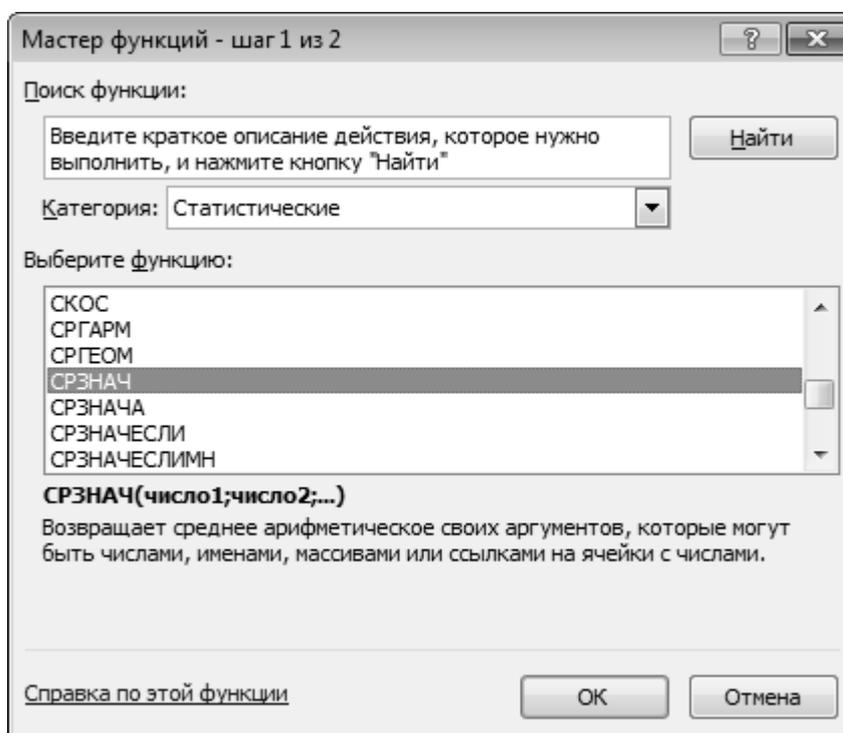


Рис. 1.5. Выбор функции «СРЗНАЧ»

Выбрав функцию, нажимаем на кнопку «ОК», в результате чего у нас появляется новое окно – «Аргументы функции» (рис. 1.6).

Очищаем первое поле для ввода данных, слева от которого написано «Число 1». Ставим в это поле курсор (рис. 1.7).

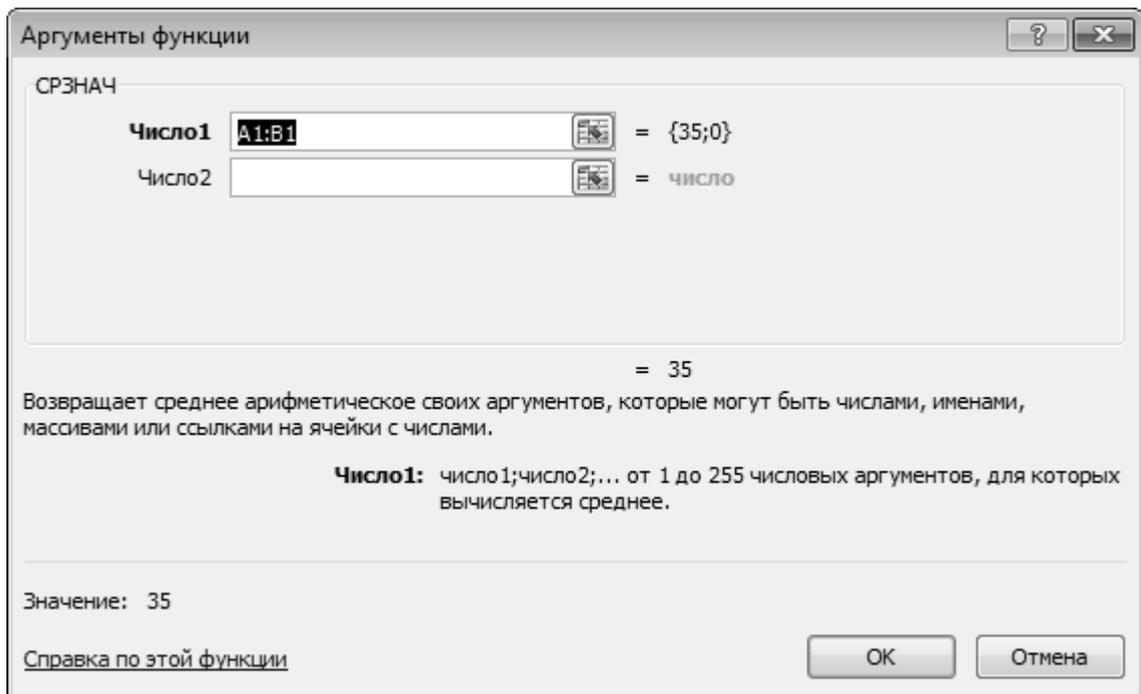


Рис. 1.6. Аргументы функции

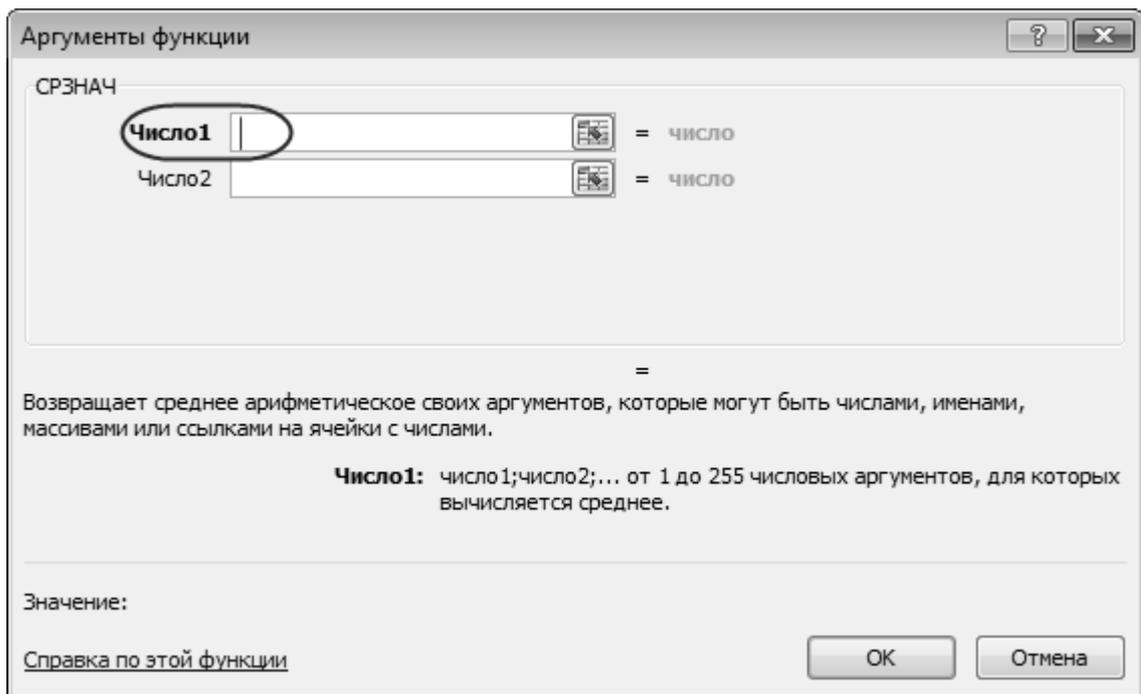


Рис. 1.7. Ввод данных в поле «Число 1»

После того, как курсор замигал в чистом поле ввода справ от надписи «Число 1», выделяем область данных, по которым будем находить среднюю арифметическую. В нашем случае это диапазон A1-A6 (рис. 1.8).

	A	B	C	D	E
1	35		(A1:A6)		
2	12				
3	45				
4	334				
5	69				
6	7				
7					
8					
9					
10					
11					

Рис. 1.8. Выделение диапазона данных

В окне «Аргументы функции» в поле «Число 1» видим наш диапазон (рис. 1.9).

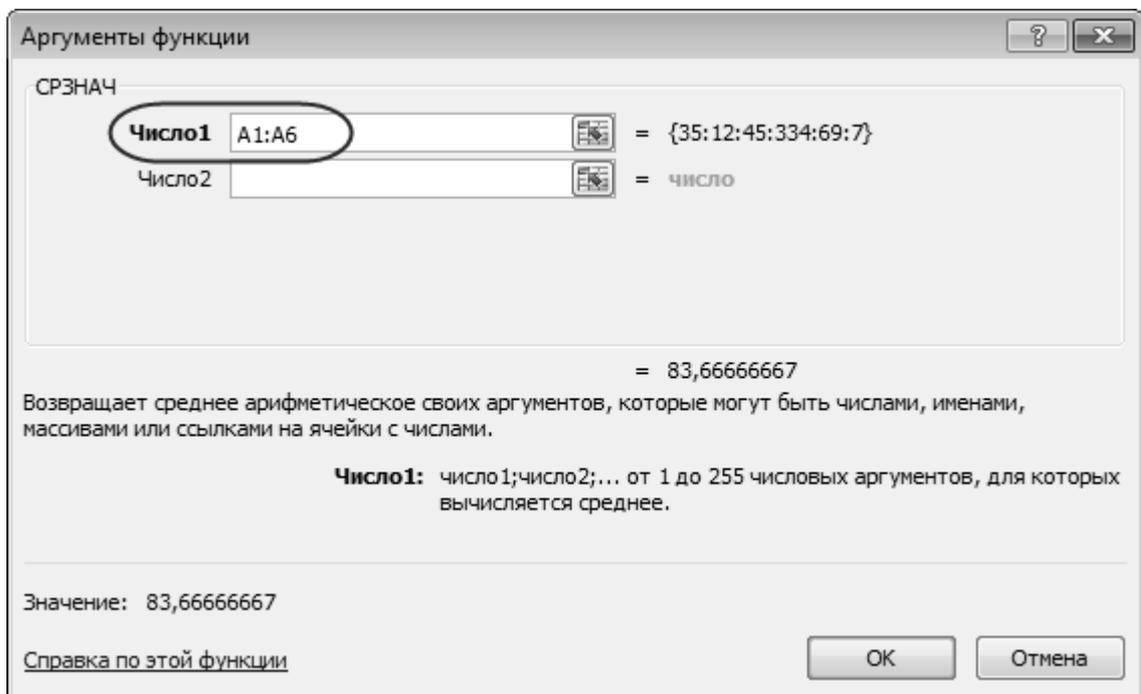


Рис. 1.9. Отображение выделенного диапазона в окне «Аргументы функции»

Нажимаем на кнопку «ОК» и получаем наш результат – среднюю арифметическую для диапазона A1-A6. В ячейке C1 получаем результат – 83,66667. В строке формул Microsoft Excel видим формулу или функцию, использованную для получения результата данной ячейки (рис. 1.10).

C1		fx =СРЗНАЧ(A1:A6)				
	A	B	C	D	E	F
1	35		83,66667			
2	12		38,07962			
3	45					
4	334					
5	69					
6	7					

Рис. 1.10. Отображение результата нахождения средней арифметической

Абсолютно аналогичным образом можно найти *среднюю геометрическую* и *среднюю гармоническую* для диапазона значений. Для средней геометрической выбираем функцию «СРГЕОМ», результат видим на рисунке 1.11, для средней гармонической выбираем функцию «СРГАРМ», результат на рисунке 1.12.

C2		fx =СРГЕОМ(A1:A6)				
	A	B	C	D	E	F
1	35		83,66667			
2	12		38,07962			
3	45					
4	334					
5	69					
6	7					

Рис. 1.11. Отображение результата нахождения средней геометрической

C3		fx =СРГАРМ(A1:A6)				
	A	B	C	D	E	F
1	35		83,66667			
2	12		38,07962			
3	45		20,37553			
4	334					
5	69					
6	7					

Рис. 1.12. Отображение результата нахождения средней гармонической

К сожалению, функции *средневзвешенной арифметической*, которая используется при нахождении цены товарного кредита, в Microsoft Excel нет, поэтому придется ее рассчитать вручную. Для этого весьма удобно использовать Microsoft Excel.

Допустим, в столбце «А» у нас указаны центральные значения интервалов. В столбце «В» укажем частоты (рис. 1.13). В примере взяты произвольные данные, но для решения определенной задачи необходимо использовать конкретные данные задачи.

	A	B	C	D	E
1	35	3			
2	12	2			
3	45	8			
4	334	1			
5	69	4			
6	7	1			
7					

Рис. 1.13. Ввод частот в столбец «В»

Следующим шагом будет нахождение суммы произведений значений интервалов и их частот. Сперва выберем ячейку для отображения данного результата. Пусть это будет ячейка В8. Нажимаем на кнопку «Вставить функцию», в какой-либо из категорий ищем функцию «СУММПРОИЗВ». Логично предположить, что это может быть категория «Математические», либо же выбираем категорию «Полный алфавитный перечень». Выбрав функцию «СУММПРОИЗВ», видим новое окно под названием «Аргументы функции». Для произведения попарно данных столбцов «А» и «В» в строке «Массив 1» выберем диапазон исходных данных столбца «А» – ячейки от А1 до А6. Для «Массив 2» аналогичным образом выберем диапазон В1-В6 (рис. 1.14). Нажимаем «ОК», в ячейке В8 получаем результат – 1106.

После этого нам потребуется сумма центральных значений интервалов, т.е. сумма значений диапазона А1-А6. Для этого выберем ячейку А8, вызовем «Вставить функцию», найдем функцию «СУММ», нажмем «ОК». Появится окно «Аргументы функции». В первом поле для ввода выберем диапазон А1-А6, нажмем «ОК». В ячейке А8 получим результат – 502.

Осталось поделить данные ячейки В8 на данные ячейки А8. Результатом будет значение 2,203, что и является нашей средневзвешенной арифметической.

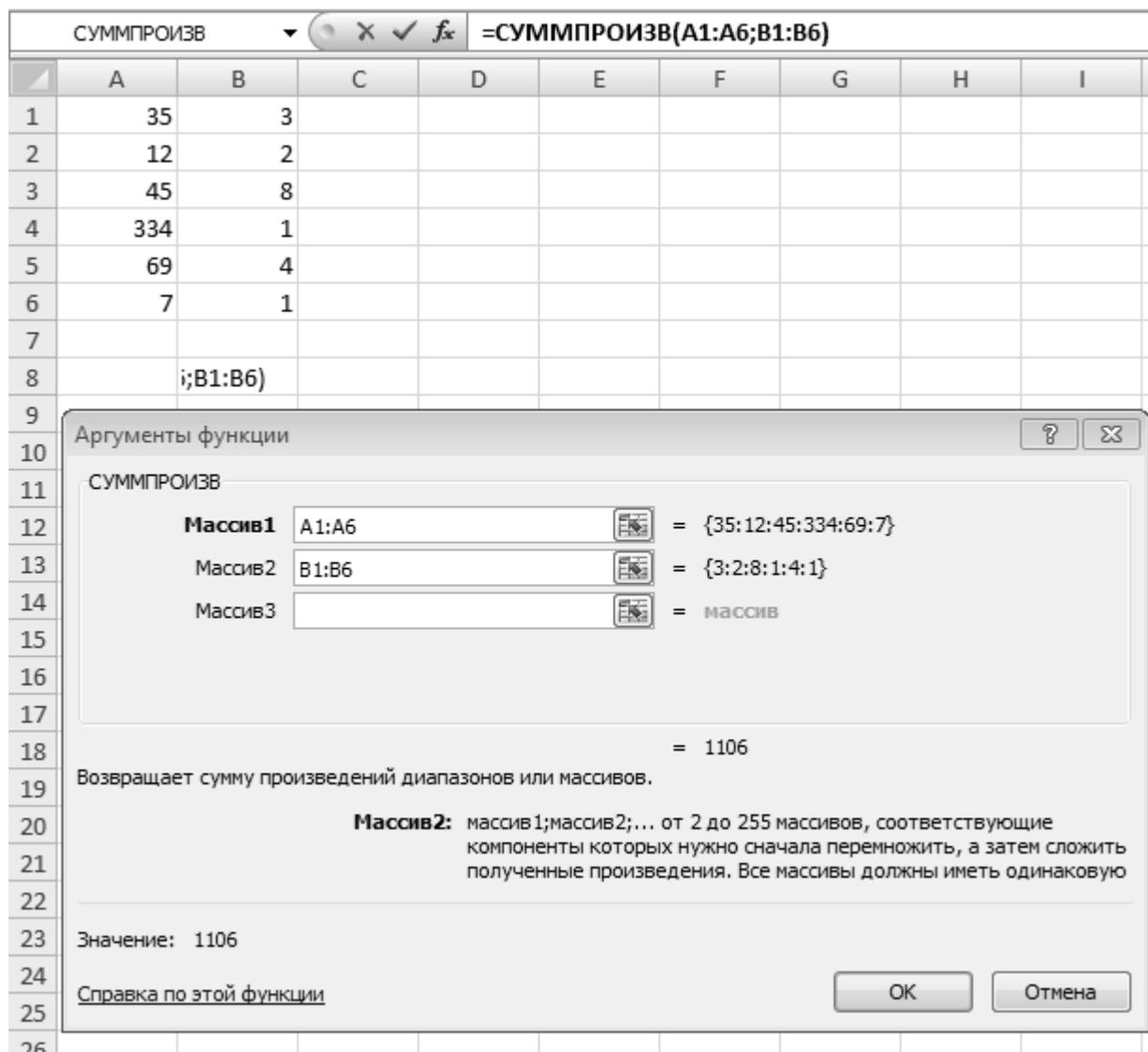


Рис. 1.14. Выбор массивов данных для функции «СУММПРОИЗВ»

1.2. Средние отклонения

Для нахождения *среднего абсолютного отклонения* в Microsoft Excel существует функция под названием «СРОТКЛ». Для нахождения *среднего квадратического отклонения* (СКО) – функция «СТАНДОТКЛОН».

Допустим, у нас есть 2 диапазона значений – массив A1-A5 и массив B1-B5 (рис. 1.15).

	A	B	C	D	E
1	1	3			
2	3	4			
3	5	5			
4	7	6			
5	9	7			

Рис. 1.15. Массивы значений A1-A5 и B1-B5

Найдем среднее абсолютное отклонение отдельно для диапазона А1-А5, отдельно для диапазона В1-В5. Выберем ячейку для отображения первого результата. Пусть это будет ячейка А7. Вызываем список функций, находим «СРОТКЛ» (рис. 1.16), нажимаем «ОК», в первом поле ввода окна «Аргументы функции» вводим диапазон А1-А5, нажимаем «ОК». В ячейке А7 получаем результат 2,4.

Аналогично поступаем с массивом В1-В5. Результат фиксируем в ячейке В7. Получаем результат 1,2 (рис. 1.17).

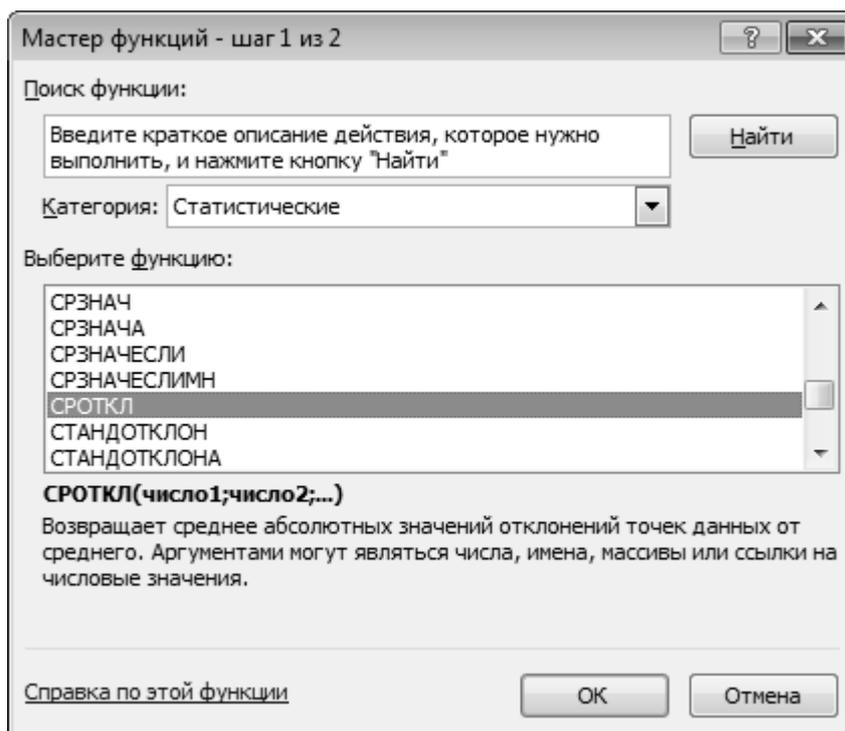


Рис. 1.16. Функция «СРОТКЛ»

	B7		fx =СРОТКЛ(B1:B5)			
	A	B	C	D	E	F
1	1	3				
2	3	4				
3	5	5				
4	7	6				
5	9	7				
6						
7	2,4	1,2				
8						

Рис. 1.17. Среднее абсолютное отклонение для А1-А5 и В1-В5

Для нахождения СКО тех же массивов используем функцию «СТАНДОТКЛОН» (рис. 1.18). Результаты запишем строкой ниже (рис. 1.19).

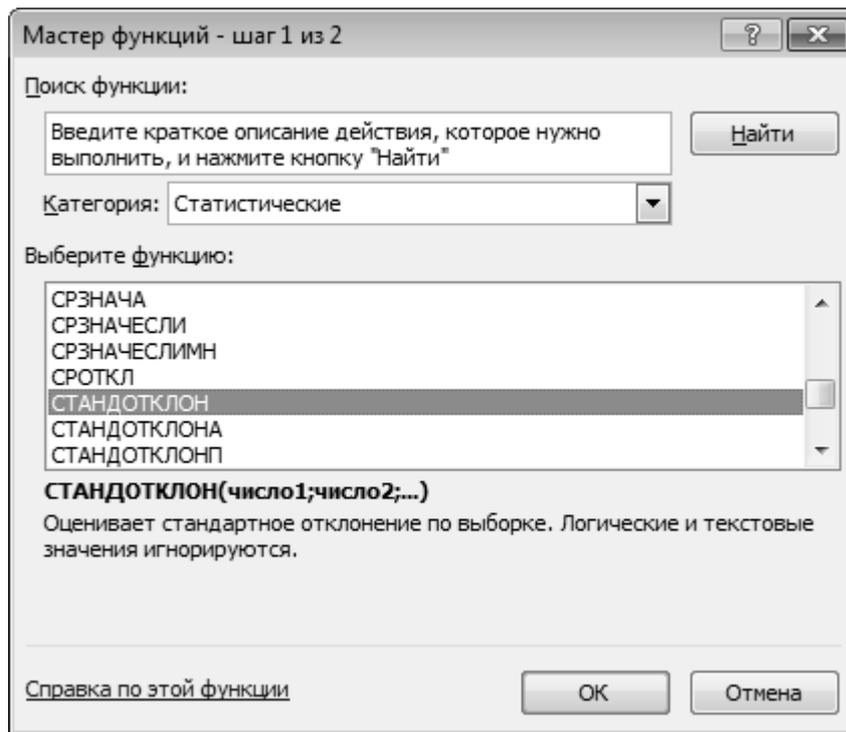


Рис. 1.18. Функция «СТАНДОТКЛОН»

	A	B	C	D	E	F	G
1	1	3					
2	3	4					
3	5	5					
4	7	6					
5	9	7					
6							
7	2,4	1,2					
8	3,162278	1,581139					
9							

Рис. 1.19. Среднее абсолютное отклонение для A1-A5 и B1-B5

Часто в статистике используется *дисперсия*, т.е. СКО, возведенное в квадрат. Для нахождения дисперсии в Excel можно:

- 1) найти СКО и возвести его во вторую степень;
 - 2) для той же выборки данных воспользоваться функцией «ДИСП».
- Результаты будут идентичны.

1.3. Коэффициент вариации

Для нахождения *коэффициента вариации* нам потребуются показатели среднего арифметического и СКО. Коэффициент вариации представляет собой отношение СКО к среднему арифметическому. Для результата в % необходимо умножить полученный коэффициент на 100%.

1.4. Прогрессии

В экономике часто приходится использовать *арифметическую* и *геометрическую* прогрессии.

Построим в Microsoft Excel арифметическую прогрессию. Для этого в окне открытого листа введем начальное значение создаваемого ряда прогрессии в первую ячейку диапазона, как это показано на рисунке 1.20. Пусть это будет число 17. Затем выделим диапазон ячеек согласно размеру прогрессии (пусть это будет 8 ячеек) и перейдем к вкладке «Главная». На данной вкладке в разделе «Редактирование» находим кнопку «Заполнить». При ее нажатии у нас выпадает подменю (рис. 1.21).

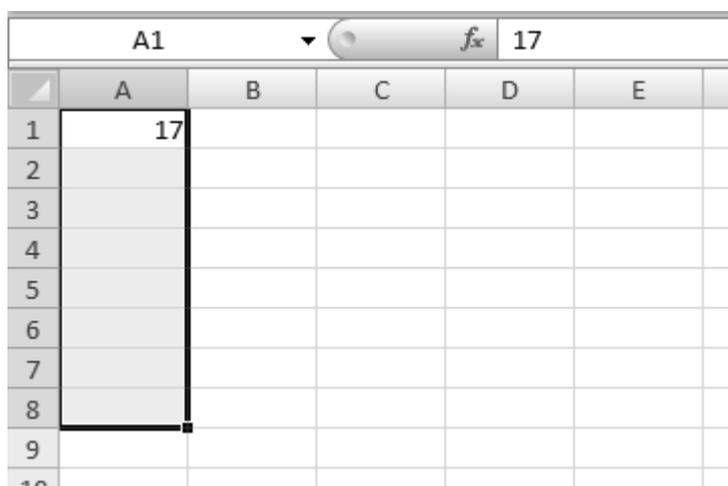


Рис. 1.20. Ввод начальных значений прогрессии

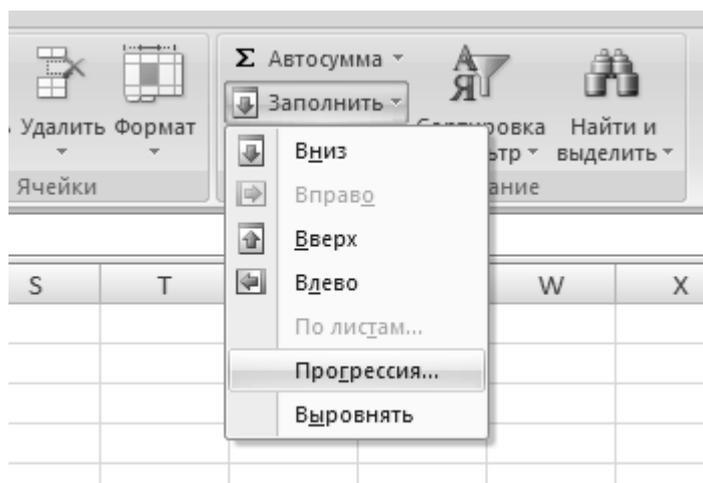


Рис. 1.21. Кнопка «Заполнить»

В появившемся подменю выбираем раздел «Прогрессия...». На экране появится новое окно с названием «Прогрессия». В группе «Тип» активируем пункт «Арифметическая». В поле ввода «Шаг» вводим число развития прогрессии, то есть значение, на которое будут увеличиваться все числа. Пусть

шаг будет равен 4. (рис. 1.22). При необходимости задаем максимально возможное число прогрессии в поле ввода «Предельное значение».

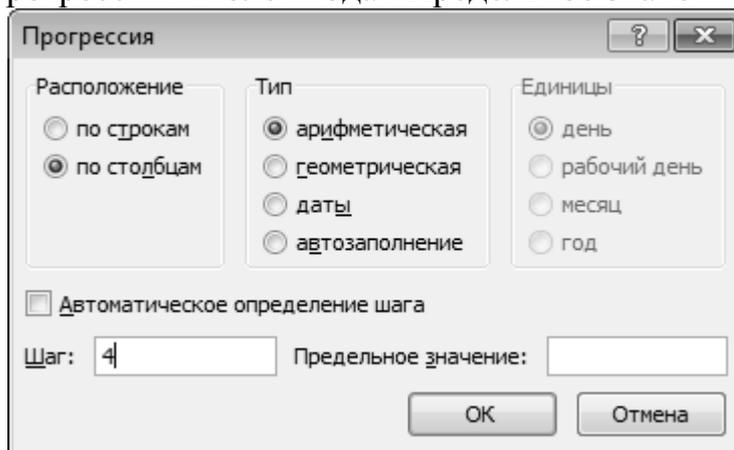


Рис.1.22. Ввод данных для построения прогрессии

Нажимаем кнопку «ОК», получаем нашу прогрессию в диапазоне A1-A8 (рис. 1.23).

	A1				
	A	B	C	D	E
1	17				
2	21				
3	25				
4	29				
5	33				
6	37				
7	41				
8	45				
9					

Рис. 1.23. Арифметическая прогрессия

По тем же параметрам построим геометрическую прогрессию. В окне «Прогрессия» на этот раз выбираем тип «Геометрическая». Результат показан на рисунке 1.24.

	A	B	C	D
1	17			
2	68			
3	272			
4	1088			
5	4352			
6	17408			
7	69632			
8	278528			

Рис. 1.24. Геометрическая прогрессия

1.5. Индекс меры несходства

Для определения индекса меры несходства необходимо построить таблицу (рис. 1.25), которая будет содержать столбцы: страна (или иной субъект); X; Y; X'; Y'; |X'-Y'|. X и Y – это некоторые показатели, в нашем примере это будут экспорт и импорт. X' и Y' – это доли каждого X и Y от общей суммы.

$$X'_i = \frac{X_i}{\sum X}; Y'_i = \frac{Y_i}{\sum Y}.$$

	A	B	C	D	E	F
1	Страна	Экспорт (X)	Импорт (Y)	X'	Y'	X'-Y'
2						
3						
4						

Рис. 1.25. Столбцы таблицы для расчета индекса меры несходства

Заполняем столбец «Страна» и показатели X и Y. Для вычисления индекса меры несходства возьмем внешнюю торговлю Украины со странами СНГ за январь–май 2010 г. (данные взяты с ukrstat.gov.ua).

	A	B	C	D	E	F
1	Страна	Экспорт (X)	Импорт (Y)	X'	Y'	X'-Y'
2	Азербайджан	245,9	460,7			
3	Армения	68,3	6,2			
4	Белорусь	649,7	672,5			
5	Казахстан	471,9	265,5			
6	Киргизстан	25,2	0,7			
7	Молдавия	263,7	27			
8	РФ	4679,2	7733,5			
9	Таджикистан	27,3	0,9			
10	Туркменистан	82,2	6			
11	Узбекистан	76,7	32,5			
12						

Рис. 1.26. Структура внешней торговли Украины со странами СНГ, млн. дол. США

Для заполнения столбцов X' и Y' необходимо найти суммы X и Y. Для этого в первом столбце под странами пишем «Итого». Ставим курсор на ячейку под данными столбца «X», вызываем функцию «СУММ» (рис. 1.27).

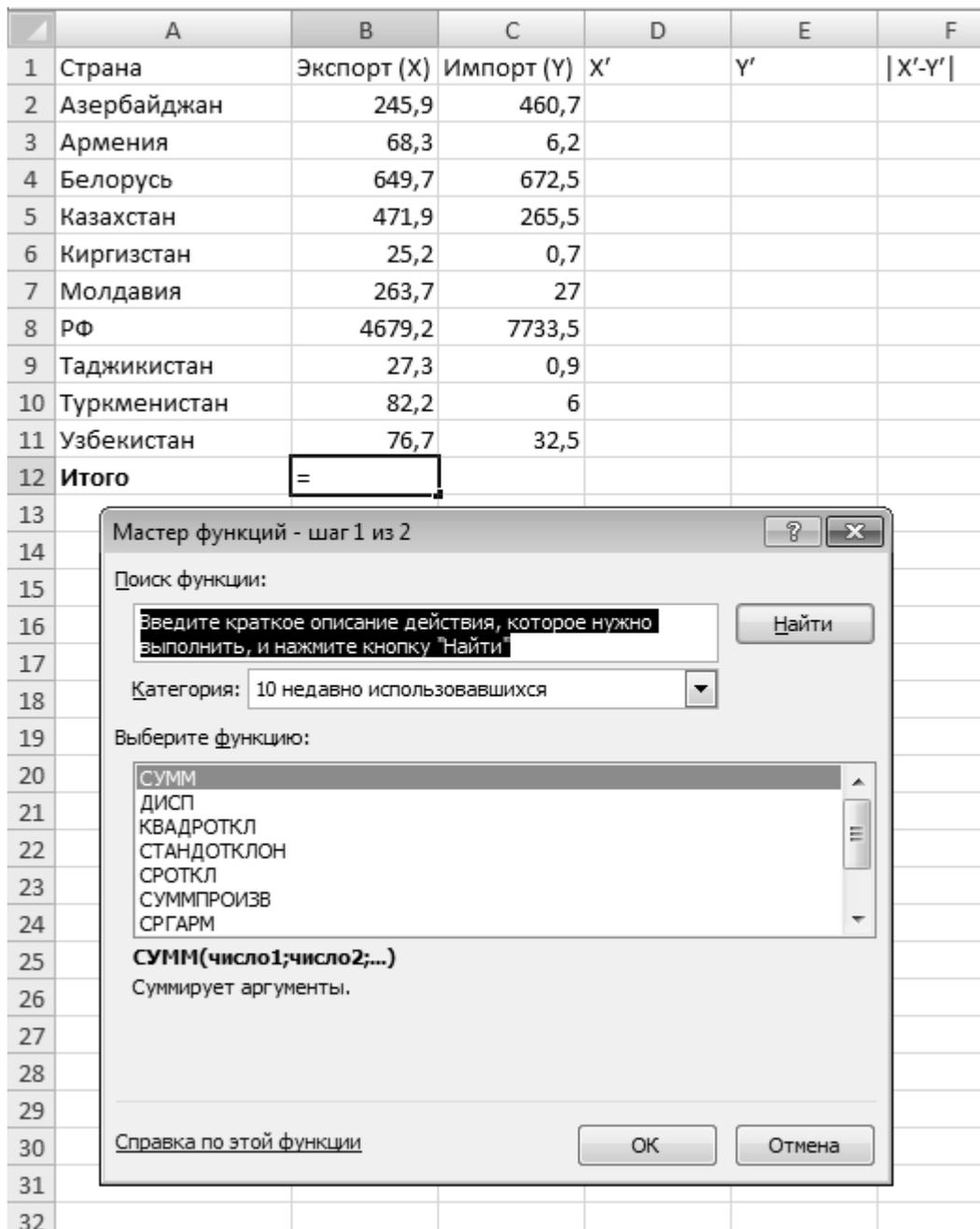


Рис. 1.27. Вызов функции «Сумм»

Нажимаем кнопку «Ок». В появившемся окне выбираем первое поле ввода и выделяем для него область значений X. В данном примере это область B2-B11 (рис. 1.28). Снова нажимаем кнопку «ОК», получаем сумму значений параметра X. Аналогично ищем сумму параметра Y.

	A	B	C	D	E	F	G
1	Страна	Экспорт (X)	Импорт (Y)	X'	Y'	X'-Y'	
2	Азербайджан	245,9	460,7				
3	Армения	68,3	6,2				
4	Белорусь	649,7	672,5				
5	Казахстан	471,9	265,5				
6	Киргизстан	25,2	0,7				
7	Молдавия	263,7	27				
8	РФ	4679,2	7733,5				
9	Таджикистан	27,3	0,9				
10	Туркменистан	82,2	6				
11	Узбекистан	76,7	32,5				
12	Итого	$\sum(B2:B11)$					

Аргументы функции

СУММ

Число1: B2:B11 = {245,9;68,3;649,7;471,9;25,2;263,7;4679,2;27,3;82,2;76,7}

Число2: = число

= 6590,1

Суммирует аргументы.

Число1: число1;число2;... от 1 до 255 аргументов, которые суммируются. Логические и текстовые значения игнорируются.

Значение: 6590,1

[Справка по этой функции](#)

OK Отмена

Рис. 1.28. Выделение области значений параметра X

Чтобы заполнить значения «доля X от суммарного X», т.е. ячейки столбца X', необходимо выполнить дальнейшие указания. Ставим курсор на ячейку под X'. Прописываем в этой ячейке следующее:

$$=B2/B12$$

Вместо того, чтобы прописывать «B2» и «B12», можем просто кликнуть на эти ячейки мышкой по ходу ввода формулы (рис. 1.29).

Далее нажимаем на кнопку «Enter» на клавиатуре и видим получившееся значение.

	A	B	C	D	E	F
1	Страна	Экспорт (X)	Импорт (Y)	X'	Y'	X'-Y'
2	Азербайджан	245,9	460,7	=B2/B12		
3	Армения	68,3	6,2			
4	Белорусь	649,7	672,5			
5	Казахстан	471,9	265,5			
6	Киргизстан	25,2	0,7			
7	Молдавия	263,7	27			
8	РФ	4679,2	7733,5			
9	Таджикистан	27,3	0,9			
10	Туркменистан	82,2	6			
11	Узбекистан	76,7	32,5			
12	Итого	6590,1	9205,5			
13						

Рис. 1.29. Ввод формулы

Чтобы заполнить оставшиеся значения X', можем также вводить формулу вручную. Или же применим автозаполнение.

Чтобы применить автозаполнение, достаточно выделить ячейку с нашим первым результатом, «ухватиться» за правый нижний угол и потянуть вниз на столько ячеек, сколько нам нужно заполнить (рис. 1.30). Но с каждым опусканием на ячейку ниже будут сдвигаться и ячейки, употребленные в формуле. Т.е., если мы растянем результат ячейки D2 нашего примера на ячейку D3, в формуле ячейки D3 увидим следующее:

=B3/B13.

Поскольку нам необходимо, чтоб менялась только первая ячейка формулы, а B12 оставалась, то в формуле ячейки D2 до автозаполнения надо ввести дополнительный символ – «\$». Перед чем мы его употребим, то в формуле при автозаполнении останется неизменным. В обозначении ячейки «B12» B – название столбца, 12 – номер строки. Если нам достаточно того, что не будет изменяться строка, то знак «\$» нужно употребить таким образом: «B\$12». Тогда формула в ячейке D2 будет иметь следующий вид:

=B2/B\$12.

Если мы в ячейке D2 запишем формулу

=B2/\$B\$12,

то для заполнения столбца X' это тоже подойдет, но тогда мы не сможем использовать уже готовый результат для автозаполнения столбца Y', поскольку все равно будет использоваться ячейка B12.

Таким образом, оптимальный результат в нашем случае – формула
=B2/B\$12.

Чтобы заполнить столбец Y', достаточно выделить полностью данные столбца X' и потянуть за квадратик в правом нижнем углу выделения вправо, так, чтобы при этом заполнилась область столбца Y'. Для достоверности можно проверить формулы в каждой ячейке.

D2		fx		=B2/B\$12		
	A	B	C	D	E	F
1	Страна	Экспорт (X)	Импорт (Y)	X'	Y'	X'-Y'
2	Азербайджан	245,9	460,7	0,03731355		
3	Армения	68,3	6,2	0,01036403		
4	Белорусь	649,7	672,5	0,09858727		
5	Казахстан	471,9	265,5	0,07160741		
6	Киргизстан	25,2	0,7	0,00382392		
7	Молдавия	263,7	27	0,04001457		
8	РФ	4679,2	7733,5	0,71003475		
9	Таджикистан	27,3	0,9	0,00414258		
10	Туркменистан	82,2	6	0,01247326		
11	Узбекистан	76,7	32,5	0,01163867		
12	Итого	6590,1	9205,5			
13						

Рис. 1.30. Автозаполнение столбца X'

Заполняем столбец $|X'-Y'|$. Для этого в ячейке F2 прописываем формулу:

=D2-E2.

Жмем кнопку Enter, делаем автозаполнение для остальных ячеек.

В ячейках, где получился отрицательный результат, в формуле дописываем: $*(-1)$.

Например, у нас в ячейке с отрицательным значением содержится формула:

=D4-E4.

Тогда нам нужно привести ее к виду:

=(D4-E4)*(-1).

Проверяем, чтобы все значения в столбце $|X'-Y'|$ получились положительными и суммируем их под столбцом в строке «Итого».

Найденный результат разделим на 2 – получим индекс меры несходства. Он должен колебаться в пределах от 0 до 100. Чем ближе данный индекс к 0, тем больше наблюдается сходство в исследуемых явлениях. Чем ближе к 100, тем больше различие.

Раздел 2. Корреляционный анализ

Коэффициент парной корреляции в Microsoft Excel находится достаточно просто. Для этого необходимо ввести 2 столбца с данными, для которых будем находить коэффициент корреляции (рис. 2.1).

	A	B	C	D	E
1		87,6	47		
2		825,5	391,4		
3		10,7	8,1		
4		49,4	7,7		
5		93,7	403,6		
6		4,3	8,7		
7		294,4	73,8		
8		3723	783,8		
9		84,5	2,5		
10		270,2	972,3		
11		114,1	126,4		
12					

Рис. 2.1. Ввод данных

Ставим курсор в новой ячейке. Вызываем функцию «КОРРЕЛ» (рис. 2.2), нажимаем «ОК».

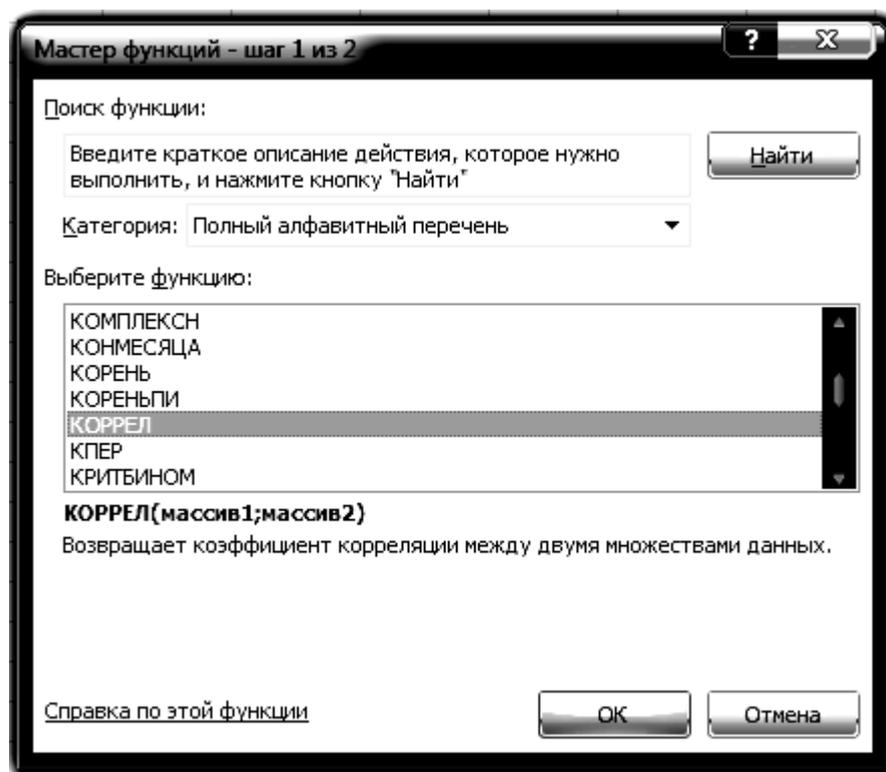


Рис. 2.2. Вызов функции «КОРРЕЛ»

В появившемся окне ставим курсор в поле ввода «Массив 1», выделяем первый столбец данных; ставим курсор в поле ввода «Массив 2», выделяем второй столбец данных (рис. 2.3).

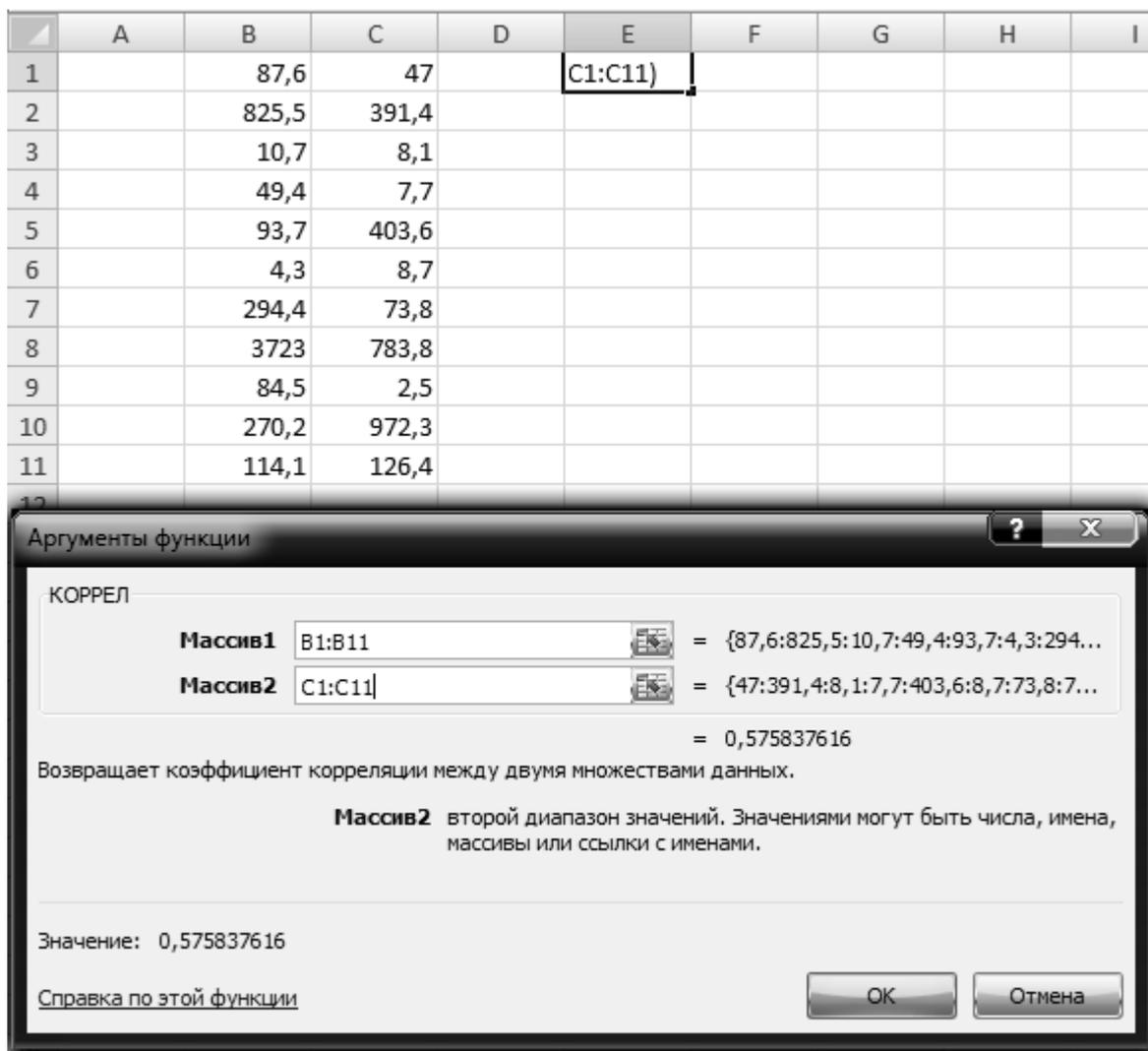


Рис. 2.3. Выделение массивов данных

Нажимаем кнопку «ОК», получаем значение коэффициента парной корреляции.

Коэффициенты множественной корреляции находятся вручную по формуле, но с использованием коэффициентов парной корреляции.

Раздел 3. Регрессионный анализ

Для построения линии регрессии (линейного тренда) в Microsoft Excel необходимо ввести имеющиеся данные, по которым будем строить линию регрессии. Затем выделяем область наших данных и на вкладке «Вставка» ждем на кнопку «Гистограмма», где выбираем тип гистограммы по вкусу (рис. 3.1).

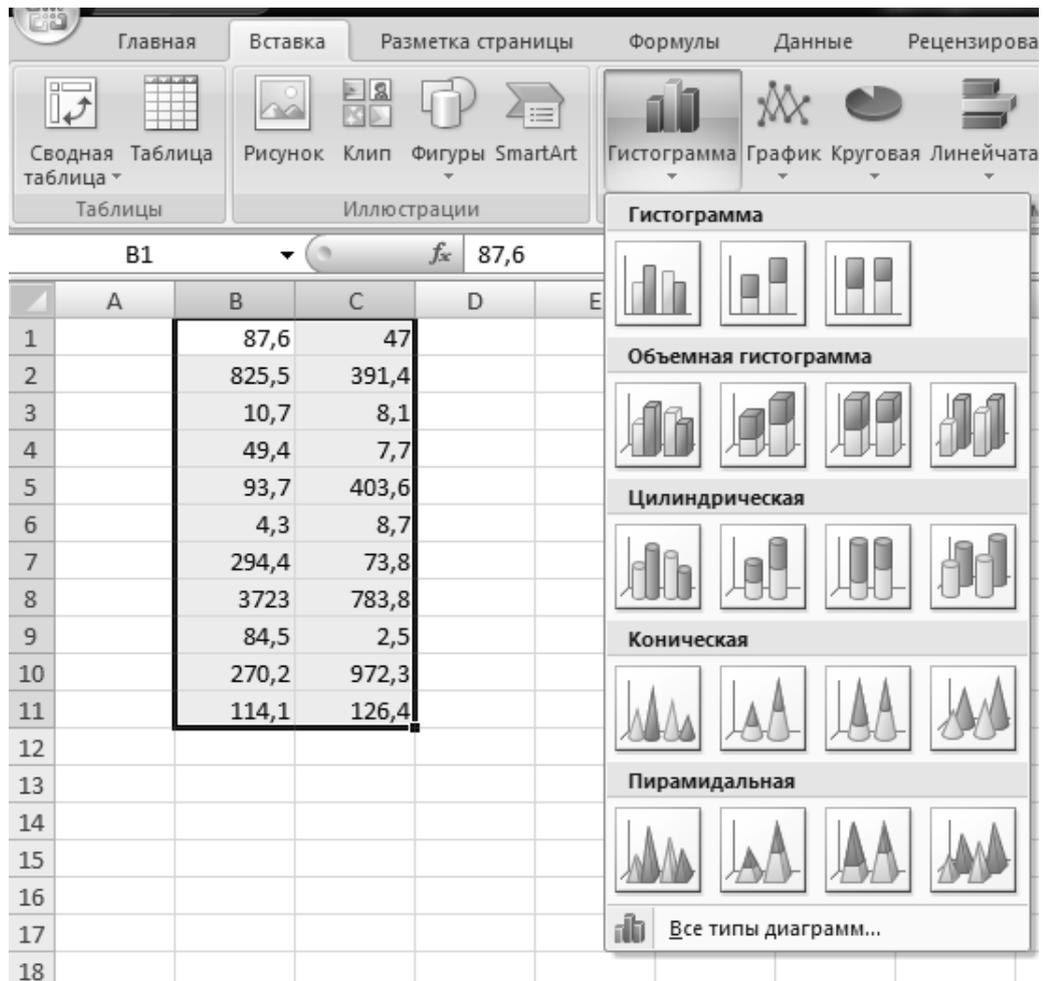


Рис. 3.1. Выбор типа гистограммы

Получаем гистограмму по заданным нами параметрам. На данной гистограмме кликаем правой кнопкой мыши непосредственно по области какого-либо столбца (рис. 3.2).

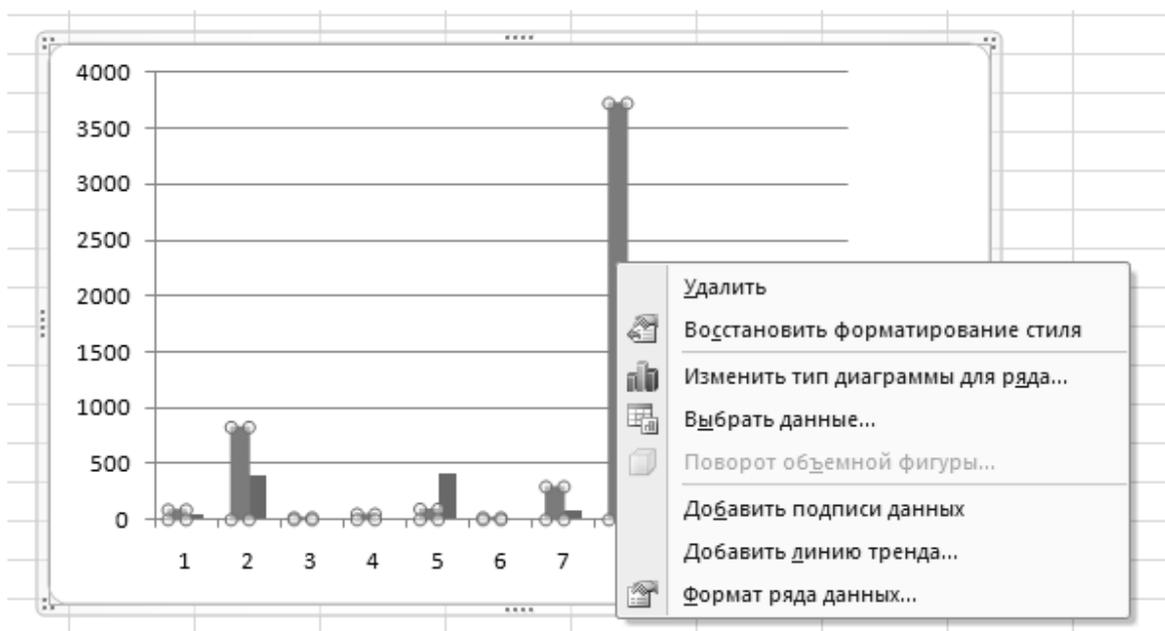


Рис. 3.2. Вызов контекстного меню

В открывшемся меню выбираем пункт «Добавить линию тренда». Появляется новое окно, в котором выбираем вид линии тренда «Линейная», а также отмечаем галочками пункты «Показывать уравнение на диаграмме» и «Поместить на диаграмму величину достоверности аппроксимации (R^2)» (рис. 3.4). Нажимаем на кнопку «Закреть».

Получаем построенную на гистограмме линию тренда (рис. 3.3), на которой содержится уравнение линии регрессии (вида $y=px+b$), а также коэффициент аппроксимации, который показывает, насколько линия тренда совпадает с данными. Коэффициент аппроксимации может варьироваться от 0 (мало совпадает) до 1 (сильно совпадает).

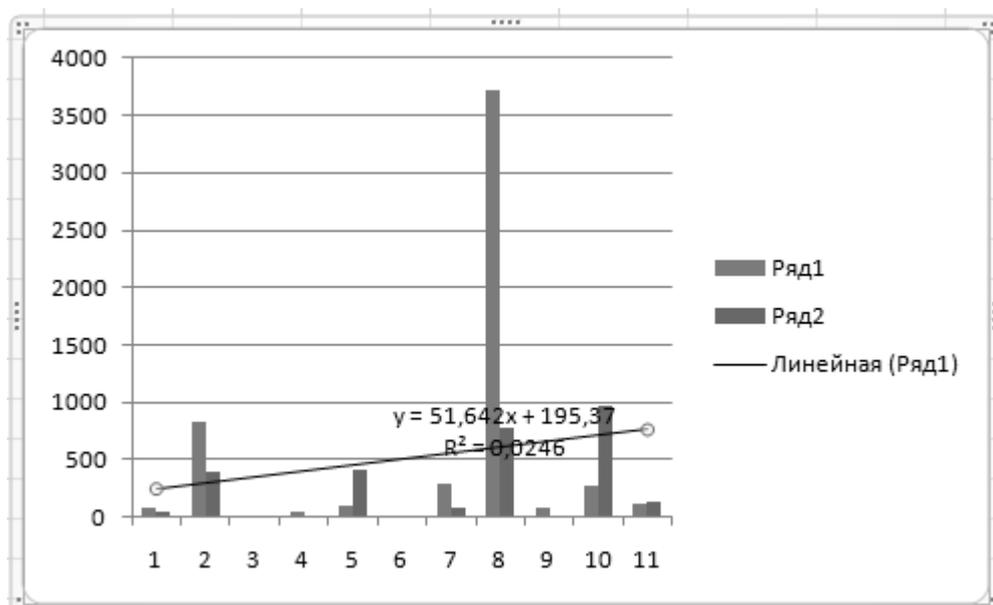


Рис. 3.3. Линия тренда с уравнением и коэффициентом аппроксимации

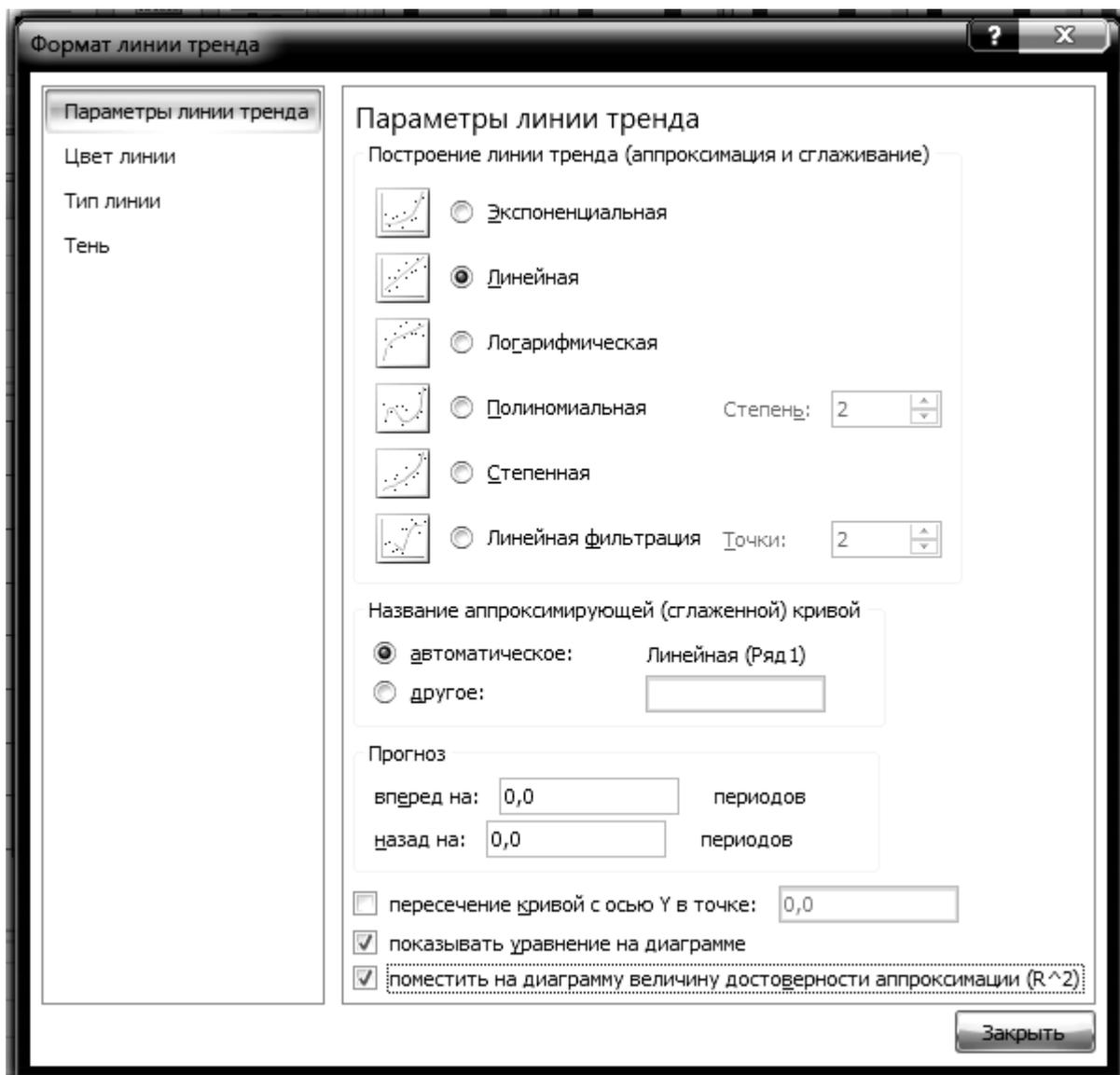


Рис. 3.4. Формат линии тренда

Раздел 4. Решение систем уравнений

Чтобы решить систему уравнений в Microsoft Excel, необходимо выполнить следующее.

1. Записываем систему уравнений в Excel для наглядности (рис. 4.1). В нашем примере это будет система из 3 уравнений с 3 неизвестными.

	A	B	C
1	$35X_1+43X_2-98X_3=17$		
2	$42X_1-9X_2-X_3=36$		
3	$12X_1+19X_2+3X_3=-5$		
4			
5			

Рис. 4.1. Система уравнений

2. Из коэффициентов при неизвестных формируем матрицу (рис. 4.2).

	A	B	C	
1	35X1+43X2-98X3=17			
2	42X1-9X2-X3=36			
3	12X1+19X2+3X3=-5			
4				
5				
6	35	43	-98	
7	42	-9	-1	
8	12	19	3	
9				
10				

Рис. 4.2. Матрица коэффициентов

3. Копируем полученную матрицу 3 раза. Таким образом у нас будет 4 одинаковых матрицы. Подпишем их: Матрица 1, Матрица 2, Матрица 3, Матрица 4.

4. Первую матрицу не трогаем. В матрице 2 заменяем первый столбец значениями, которые в системах уравнений стоят после знака « \Rightarrow » (рис. 4.3).

	A	B	C	D
1	35X1+43X2-98X3=17			
2	42X1-9X2-X3=36			
3	12X1+19X2+3X3=-5			
4				
5	Матрица 1			
6	35	43	-98	
7	42	-9	-1	
8	12	19	3	
9				
10	Матрица 2			
11	17	43	-98	
12	36	-9	-1	
13	-5	19	3	

Рис. 4.3. Замена данных первого столбца в Матрице 2

5. В Матрице 3 заменяем этими же данными второй столбец, в Матрице 4 заменяем третий.

6. Для каждой матрицы находим определитель. Найдем для Матрицы 1. Для этого поставим курсор справа от матрицы, нажмем кнопку «Вставить функцию», в разделе «Полный алфавитный перечень» выберем функцию «МОПРЕД», в появившемся окне в поле «Массив» выделим область Матрицы 1 (рис. 4.4) и нажмем кнопку «ОК».

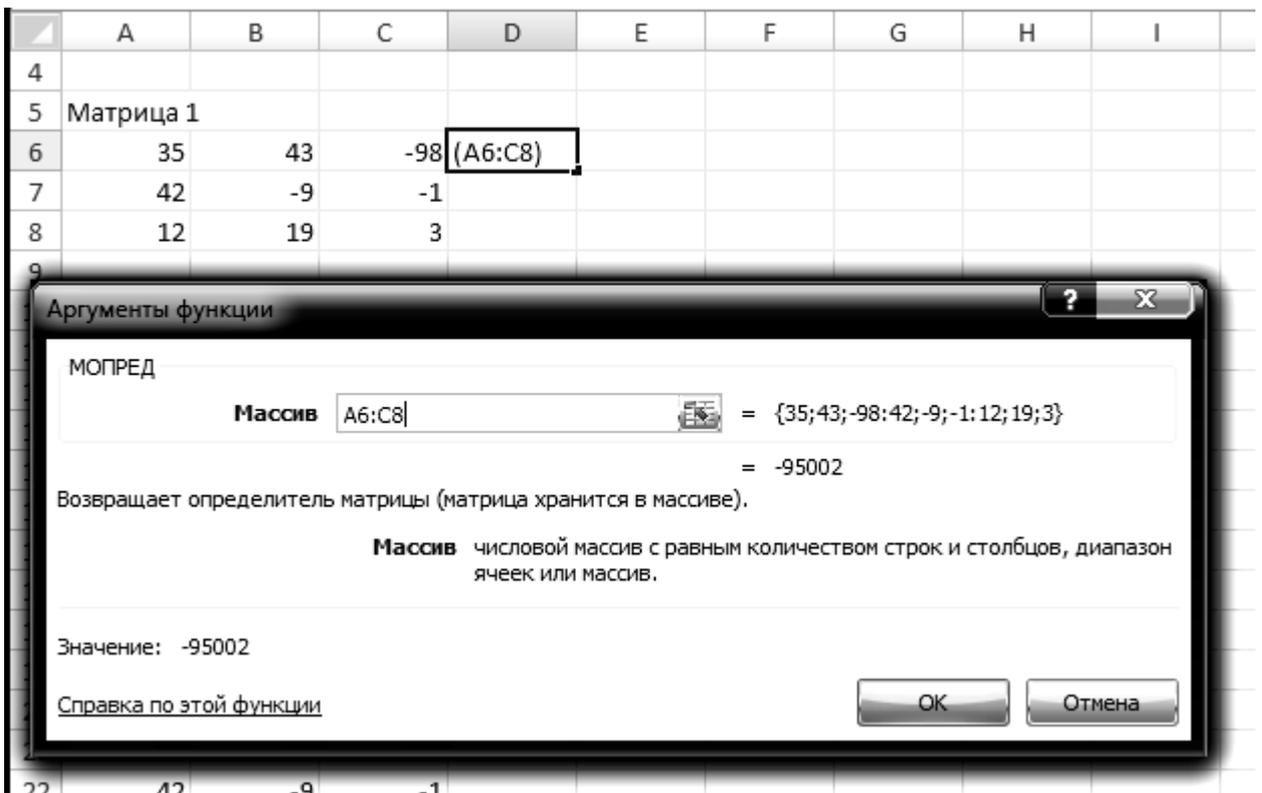


Рис. 4.4. Выделение массива данных для функции МОПРЕД

7. Аналогично найдем матричные определители для Матрицы 2, Матрицы 3 и Матрицы 4.

8. Допустим, что определитель Матрицы 1 – это ΔX . Определители матриц 2-4 – соответственно ΔX_1 , ΔX_2 , ΔX_3 . Найдем X_1 , X_2 , X_3 по формулам:

$$x_1 = \frac{\Delta x_1}{\Delta x}; x_2 = \frac{\Delta x_2}{\Delta x}; x_3 = \frac{\Delta x_3}{\Delta x}.$$

Раздел 5. Решение транспортной задачи линейного программирования

Допустим, у нас имеется транспортная задача, где есть 2 производителя. У первого есть в наличии 150 единиц продукции, у второго 90. Также имеется 3 потребителя, которым необходимо 60, 70 и 110 единиц продукции. В условии также даны стоимости доставки единицы продукции от каждого производителя к каждому потребителю.

Составим таблицу с условиями в Excel (рис. 5.1.).

	A	B	C	D	E	F
1						
2			Потребители			
3			60	70	110	
4	Производители	150	6	1	4	
5		90	12	20	80	
6						
7						
8						

Рис. 5.1. Условия задачи

На рисунке 5.1 в рамке показаны стоимости доставки.

Ниже построим табличку таких же размеров, но пока что в ней не будет никаких данных. Ячейки B9:B10 будут соответствовать производителям, ячейки C8:E8 будут соответствовать потребителям.

В нижней таблице необходимо заполнить ячейки производителей и потребителей следующим образом. В ячейках B9 и B10 должны содержаться суммы значений строк справа, т.е. в ячейке B9 должна быть формула:

=СУММ(C9:E9) ;

в ячейке B10 должна быть формула:

=СУММ(C10:E10) .

В ячейках C8, D8, E8 должны содержаться суммы значений столбцов под ними. Т.е. в ячейке C8 должна содержаться формула:

=СУММ(C9:C10) ;

в ячейке D8 формула:

=СУММ(D9:D10) ;

в ячейке E8 формула:

=СУММ(E9:E10) .

Если Вы все сделали правильно, то в ячейках, в которые вводились формулы, будут стоять ноли.

Справа от таблицы с условием выберем ячейку (например F4) и введем в ней следующую формулу:

=СУММПРОИЗВ(C4:E5;C9:E10) (рис. 5.2).

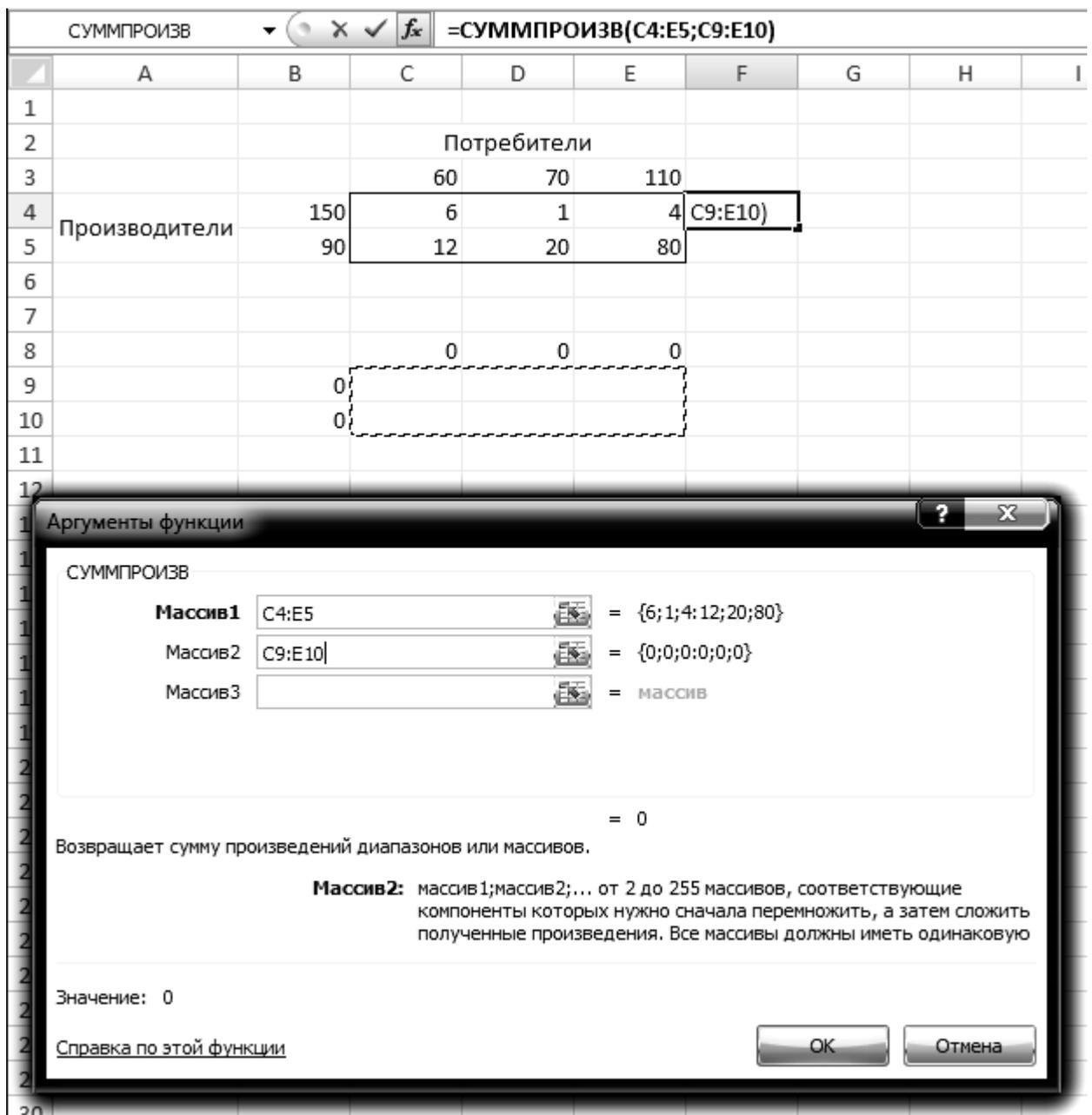


Рис. 5.2. Функция «СУММПРОИЗВ»

Будучи курсором на ячейке F4, на вкладке «Данные» в меню «Анализ» выбираем команду «Поиск решения». Целевую ячейку оставляем F4 (\$F\$4), выбираем значение целевой функции равной «минимальному значению», устанавливаем изменяемые ячейки C9:E10 (\$C\$9:\$E\$10) (рис. 5.3).

Затем в данном окне видим подраздел «Ограничения». Нажимаем кнопку «Добавить». Появляется новое окно. В поле ввода «Ссылка на ячейку» нам необходимо выбрать нижнюю таблицу, область внутри рамки. В данном случае получится \$C\$9:\$E\$10. Во втором поле ввода в выпадающем меню выбираем «цел» (целые числа), поскольку мы не будем наши единицы товара разделять на части, а доставим их потребителям целыми (рис. 5.4).

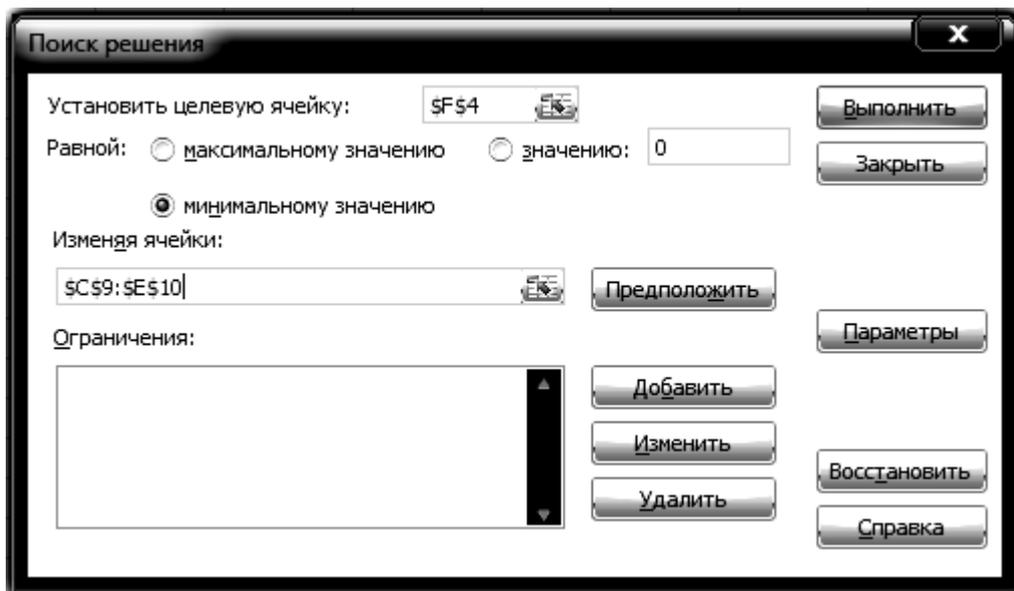


Рис. 5.3. Поиск решения

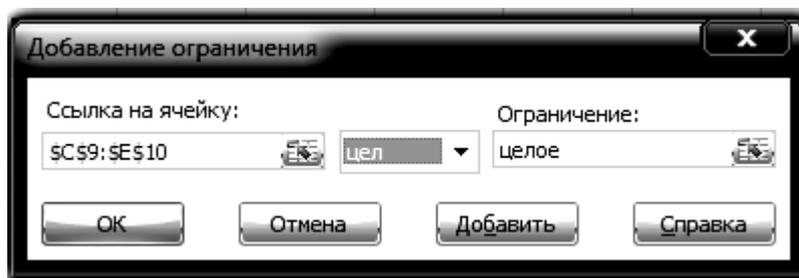


Рис. 5.4. Установление первого ограничения

Далее жмем кнопку «Добавить», поскольку нам нужно добавить еще ограничения. Появляется такое же окно, только с чистыми полями для ввода. Ссылка на ячейку – снова ячейки C9:E10. Во втором поле выбираем «>=», в третьем поле ставим «0» (рис. 5.5). Это значит, что мы не можем поставить отрицательное число продукции.

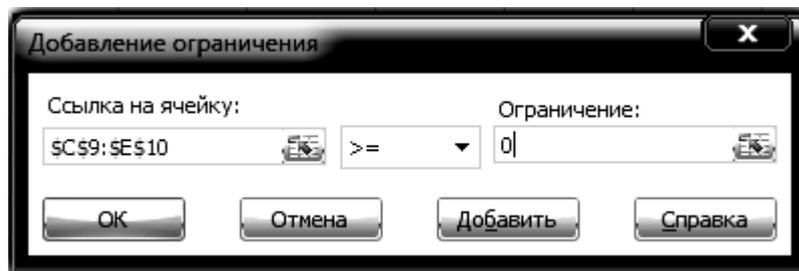


Рис. 5.5. Установление второго ограничения

Нажимаем кнопку «Добавить», в первом поле ввода выбираем ячейки C3:E3, во втором поле знак равенства, в третьем ячейки C8:E8 (рис. 5.6). Таким образом, устанавливается соответствие данных по потреблению в условии и расчетной таблице.

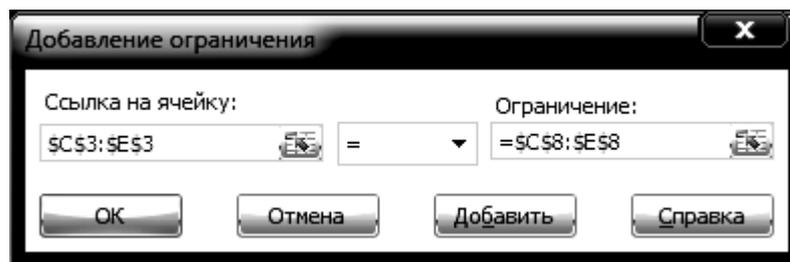


Рис. 5.6. Установление третьего ограничения

Снова нажимаем кнопку «Добавить», выбираем ячейки В4:В5, знак равенства, ячейки В9:В10 (аналогично третьему ограничению). Жмем кнопку «ОК». Итак, у нас получилось 4 ограничения (рис. 5.7).

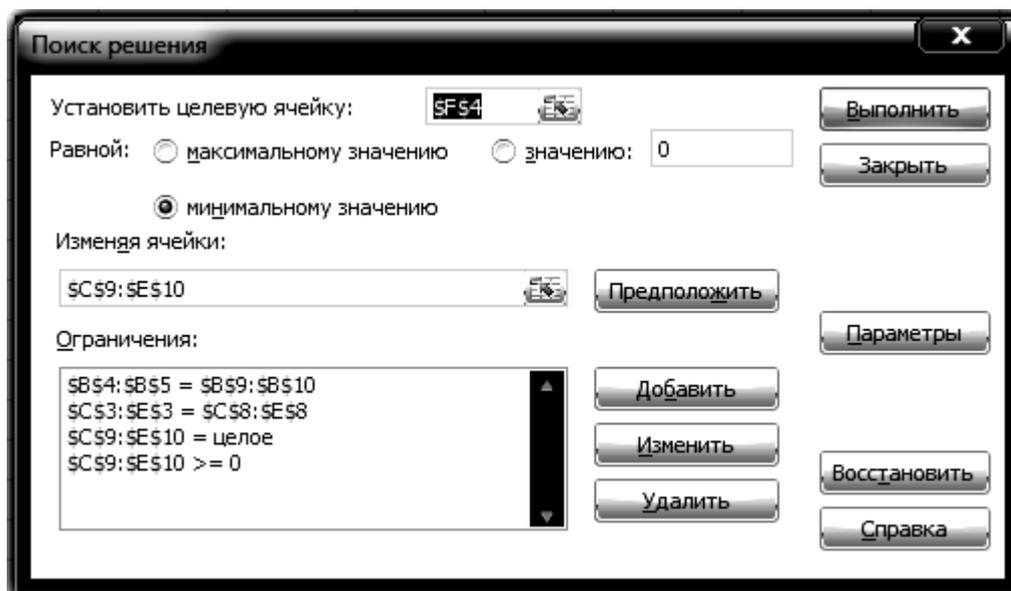


Рис. 5.7. Поиск решения – заданы 4 ограничения

Нажимаем на кнопку «Параметры», в новом окне ставим галочку на «Линейная модель» (рис. 5.8).

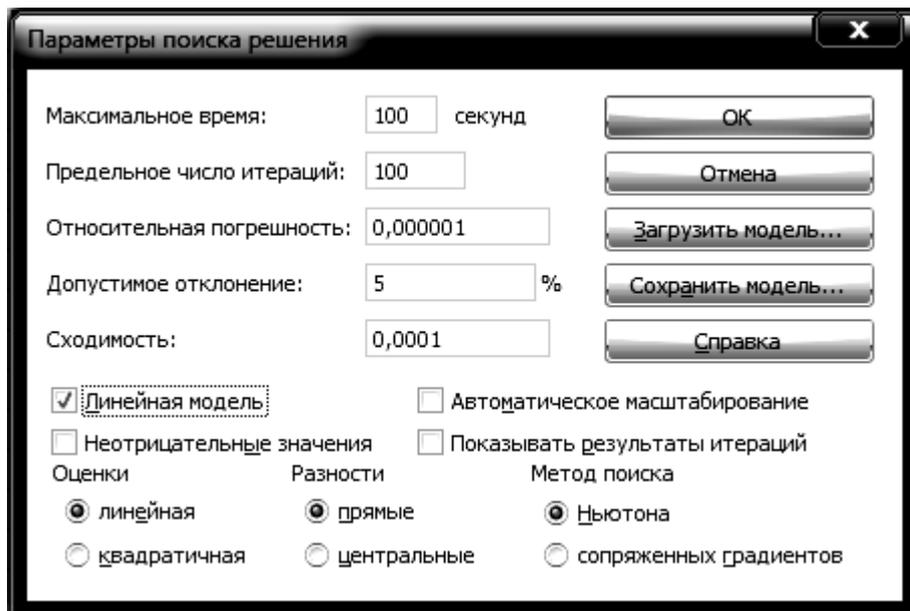


Рис. 5.8. Параметры поиска решения

В этом окне нажимаем кнопку «ОК», в следующем «Выполнить». Получаем окно, показанное на рисунке 5.9.

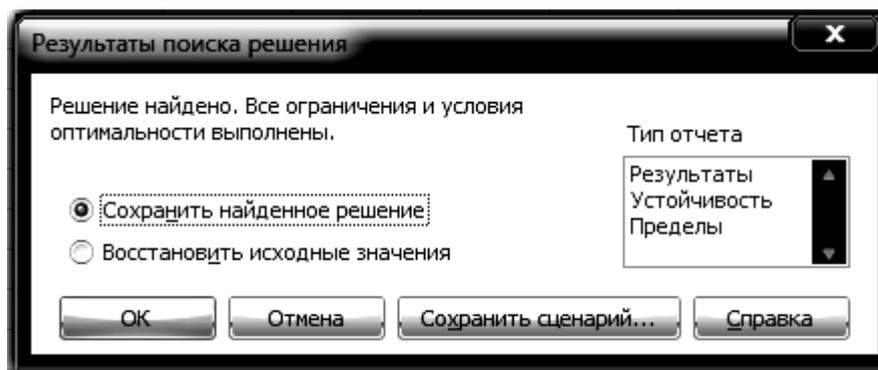


Рис. 5.9. Результаты поиска решения

Выбираем пункт «Сохранить найденное решение», нажимаем «ОК».

В нижней табличке получаем оптимальный план поставки продукции, а в ячейке F4 – суммарную стоимость перевозки (рис. 5.10).

Решение найдено.

	A	B	C	D	E	F	G
1							
2			Потребители				
3			60	70	110		
4	Производители	150	6	1	4	1800	
5		90	12	20	80		
6							
7							
8			60	70	110		
9		150	0	40	110		
10		90	60	30	0		
11							
12							
13							

Рис. 5.10. Полученное решение

Примечание

Средство поиска решения является надстройкой Microsoft Office Excel, которая доступна при установке Microsoft Office или Microsoft Excel. Чтобы использовать эту надстройку в Excel, необходимо сначала загрузить ее.



Нажмите кнопку Microsoft Office , а затем щелкните «Параметры Excel». Выберите команду «Надстройки», а затем в окне «Управление» выберите пункт «Надстройки Excel». Нажмите кнопку «Перейти». В окне «Доступные надстройки» установите флажок «Поиск решения» и нажмите кнопку ОК.

Если «Поиск решения» отсутствует в списке поля «Доступные надстройки», чтобы найти надстройку, нажмите кнопку «Обзор».

В случае появления сообщения о том, что надстройка для поиска решения не установлена на компьютере, нажмите кнопку «Да», чтобы установить ее.

После загрузки надстройки для поиска решения в группе «Анализ» на вкладке «Данные» становится доступна команда «Поиск решения».

Раздел 6. Кластерный анализ

Напомним, кластерный анализ (англ. Data clustering) или кластеризация – многомерный анализ по многим факторам, который дает возможность объединить какие-либо явления или объекты на сравнительно однородные группы (кластеры). В этих кластерах объединяются элементы, имеющие похожие или близкие показатели.

Алгоритмы кластеризации весьма обширны, но оптимальным вариантом является построение иерархического дерева кластеризации. Сущность этого алгоритма в том, что наглядно в виде рисунка показывается сам процесс объединения в кластер.

Также отметим, что метрик для расчета «расстояний» между показателями множество: евклидово расстояние, мера сходства Хэмминга, мера сходства Роджерса-Танимото, Манхэттенская метрика, расстояние Махалонобиса и др. Самым распространенным методом является евклидово расстояние.

*Кластерный анализ можно считать объективным, если количество объектов и показателей для расчета будет больше десяти (т.е. представляют собой матрицу 10*10 и более).*

В международных экономических отношениях применение кластерного анализа помогает объединить в кластеры страны мира, регионы одной страны по множеству показателей.

Представим, что нашим заданием является проведение кластеризации регионов Чехии по определенным показателям. Анализ проводим в программе Statistica 6.0.

1 этап – выбираем показатели для анализа. Для примера выбраны следующие 10 показателей:

- 1) вrp на одного жителя, у.е.;
- 2) объем промышленного производства на одного жителя, у.е.;
- 3) количество аварий по региону;
- 4) располагаемый доход на одного жителя, у.е.;
- 5) бюджетные расходы региона, у.е.;
- 6) доля убыточных предприятий, %;
- 7) экспорт продукции в перерасчете на одного жителя, у.е.;
- 8) номинальная зарплата на одного работника, у.е.;
- 9) реальные денежные доходы населения, у.е.;
- 10) уровень безработицы в регионе, %.

2 этап – делаем таблицу показателей по регионам Чехии, которая является основой кластерного анализа:

№		1	2	3	4	5	6	7	8	9	10
1	Высочина	15532	1122	6623	9006	10528	5,5	0,101	7710	46027	6,27
2	Злинский край	15028	1349	5596	9099	10518	4,9	1,104	7307	53254	6,13
3	Карловарский край	13337	804	4658	8398	6036	2,5	0,085	8608	25661	7,62
4	Карловеградский край	15437	970	7382	9133	10204	5,5	0,092	8446	50392	4,81
5	Либерецкий край	13752	1180	7088	8773	8234	4,1	0,106	8548	37857	6,95
6	Моравско-силезский	15645	1543	16466	9099	22994	8,6	0,107	7793	107774	8,49
7	Оломоуцкий	14162	1092	8386	8715	11719	5,8	0,078	7959	55394	6,87
8	Пардубицкий	15502	1838	6814	9033	9351	5	0,190	8246	45953	5,95
9	Пльзенский	16668	1583	8968	9547	12016	5,8	0,213	8585	53515	5,03
10	Прага	40033	2126	30251	12768	34538	15,1	0,051	9516	155143	2,14
11	Среднечешский	17069	2365	22053	10174	21732	11,6	0,203	8499	122591	4,47
12	Устецкий	14943	2451	12294	8332	15584	7	0,098	8092	68843	10,26
13	Южно-моравский	17150	1220	14174	9306	23021	10,5	0,079	8367	105380	6,83
14	Южночешский	16145	1267	9623	9222	12822	7,4	0,082	7897	57954	4,83

Получилась матрица 14*10, которая соответствует требованию объективности анализа.

3 этап – открываем программу Statistica (рис. 6.1) и экспортируем в нее полученную таблицу (рис. 6.2). В более ранних версиях программы Statistica возможна некоторая несовместимость с Word, что отражается на некорректном переносе данных, поэтому проще таблицу данных составить в программе Excel, а уже из него экспортировать данные в Statistica.

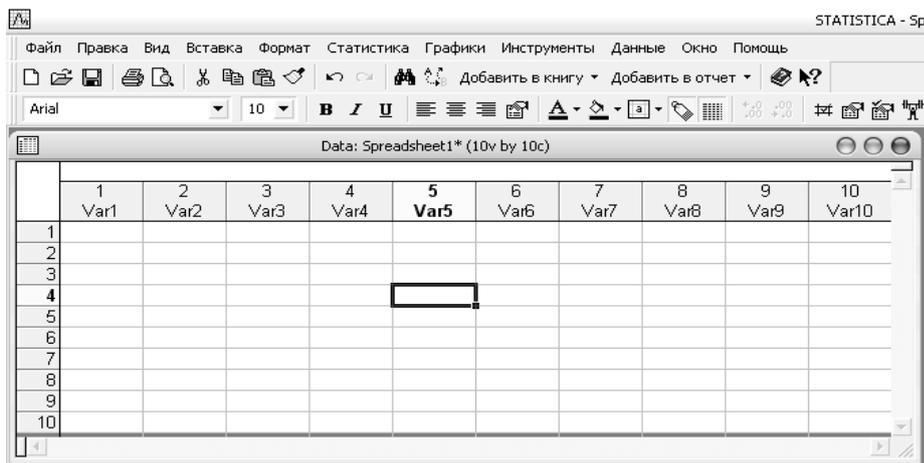


Рис. 6.1. Окно программы Statistica

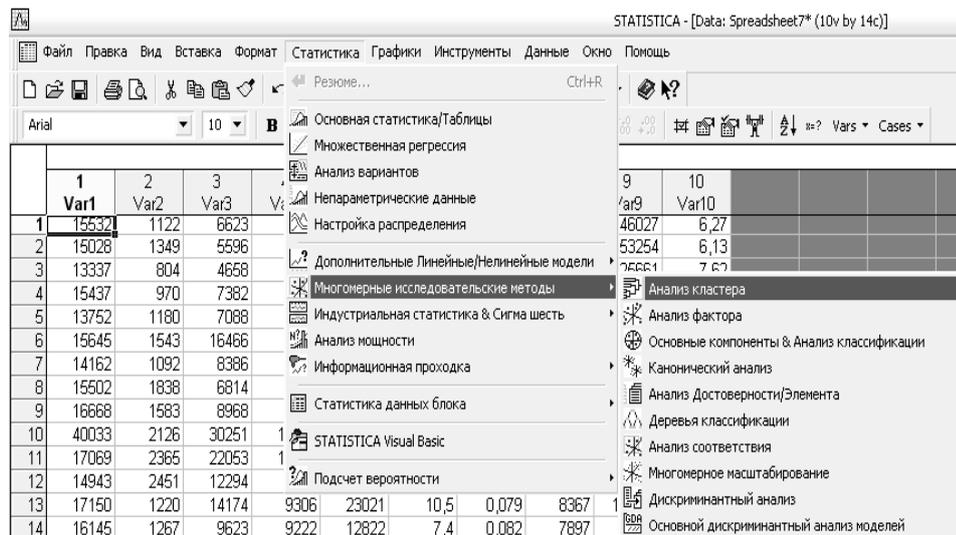
	1 Var1	2 Var2	3 Var3	4 Var4	5 Var5	6 Var6	7 Var7	8 Var8	9 Var9	10 Var10
1	15532	1122	6623	9006	10528	5,5	0,101	7710	46027	6,27
2	15028	1349	5596	9099	10518	4,9	1,104	7307	53254	6,13
3	13337	804	4658	8398	6036	2,5	0,085	8608	25661	7,62
4	15437	970	7382	9133	10204	5,5	0,092	8446	50392	4,81
5	13752	1180	7088	8773	8234	4,1	0,106	8548	37857	6,95
6	15645	1543	16466	9099	22994	8,6	0,107	7793	107774	8,49
7	14162	1092	8386	8715	11719	5,8	0,078	7959	55394	6,87
8	15502	1838	6814	9033	9351	5	0,19	8246	45953	5,95
9	16668	1583	8968	9547	12016	5,8	0,213	8585	53515	5,03
10	40033	2126	30251	12788	34538	15,1	0,051	9516	155143	2,14
11	17069	2365	22053	10174	21732	11,6	0,203	8499	122591	4,47
12	14943	2451	12294	8332	15584	7	0,098	8092	68843	10,26
13	17150	1220	14174	9306	23021	10,5	0,079	8367	105380	6,83
14	16145	1267	9623	9222	12822	7,4	0,082	7897	57954	4,83

Рис. 6.2. Экспорт данных в программу Statistica

*Копировать нужно **только** цифры, не должно быть пустых строк или строк с текстом. В противном случае результат анализа будет сильно искажен.*

По горизонтали (**строки, или cases (rows)**) представлены регионы Чехии, а по вертикали (**столбцы, или var**) представлены выбранные для анализа показатели.

4 этап – теперь в окне программы Statistica (рис. 6.2) заходим в меню «Статистика», в открывшемся списке выбираем «Многомерные исследовательские методы» и в еще открывшемся окне выбираем «Анализ кластера» – рис. 6.3.



6.3. Выбор вида анализа

В результате на экране появится еще одно окно (рис. 6.4), в котором следует выбрать древообразную кластеризацию – Joining (tree clustering) и нажать ОК.



Рис. 6.4. Выбор алгоритма кластеризации

5 этап – после выполненных шагов на экране появится окно (рис. 6.5), в котором необходимо сделать необходимые установки для проведения анализа.

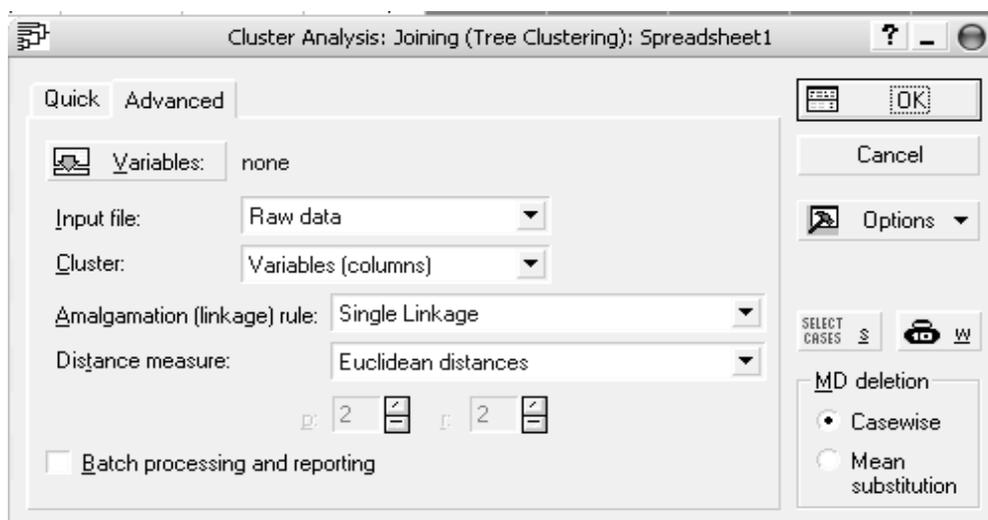


Рис. 6.5. Окно установок для проведения анализа

Для начала зайдём в меню Variables (переменные). Это – 10 столбцов с показателями. Как видите, на рисунке 6.5. стоит надпись none, т.е. нет данных для анализа. В открывшемся окне Variables выбираем все переменные Select All или же, если их очень много, вручную выбираем нужные для анализа переменные – рис. 6.6.

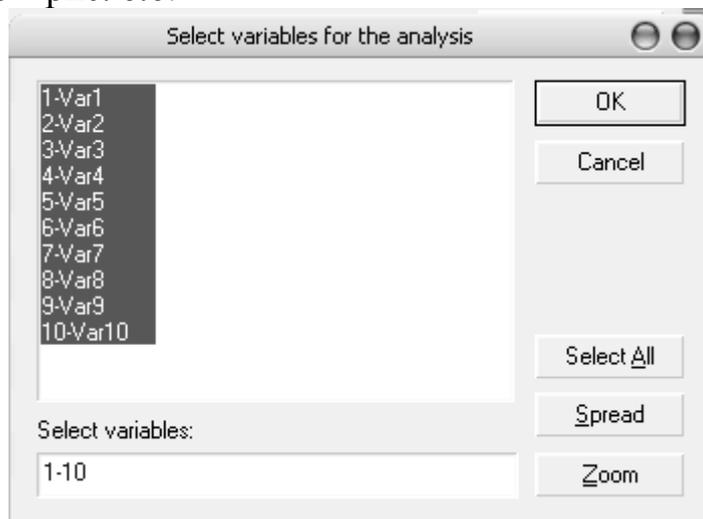


Рис. 6.6. Выбор переменных

На рис.6.7. напротив меню Variables стоит уже надпись All. Перейдем к меню Cluster. По умолчанию программа выбирает кластеризацию по переменным или столбцам Variables (columns), но нам это не подходит, т.к. анализ проводим по регионам, которые в нашей таблице расположены по строкам. Поэтому меняем эту установку на Cases (rows) – рис. 6.7.

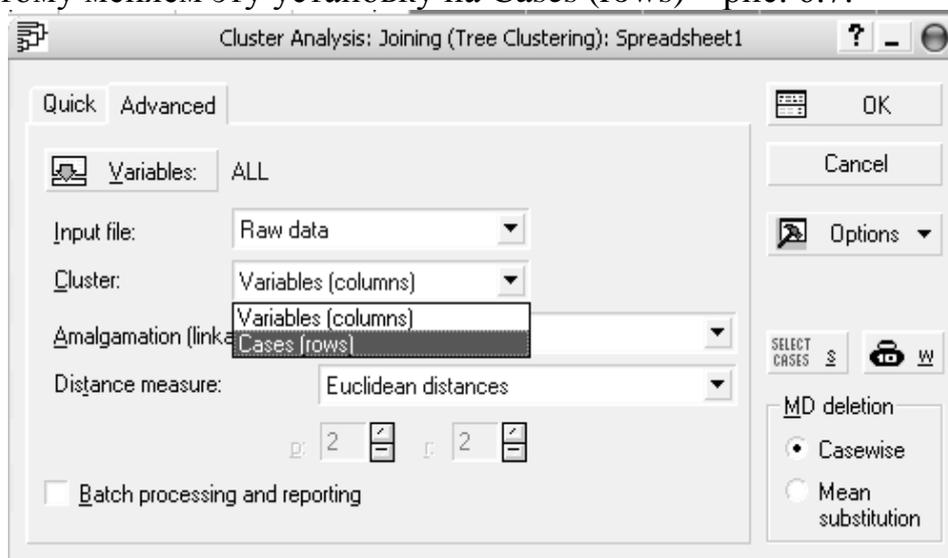


Рис. 6.7. Выбор анализа по строкам Cases (rows)

Последняя необходимая установка – выбор правила объединения показателей Amalgamation (linkage) rule. По опыту, наибольшую наглядность древо кластеризации имеет при выборе установки полного объединения – Complete linkage (рис. 6.8).

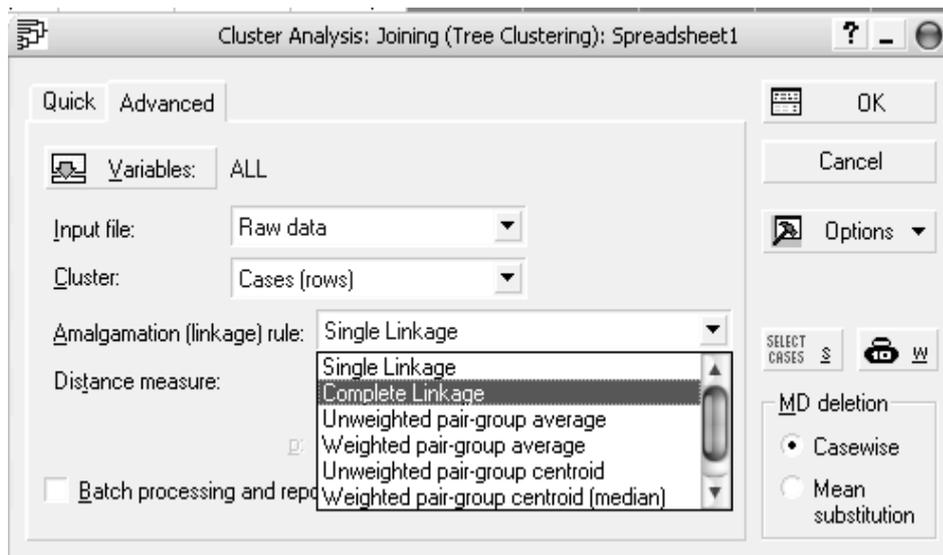
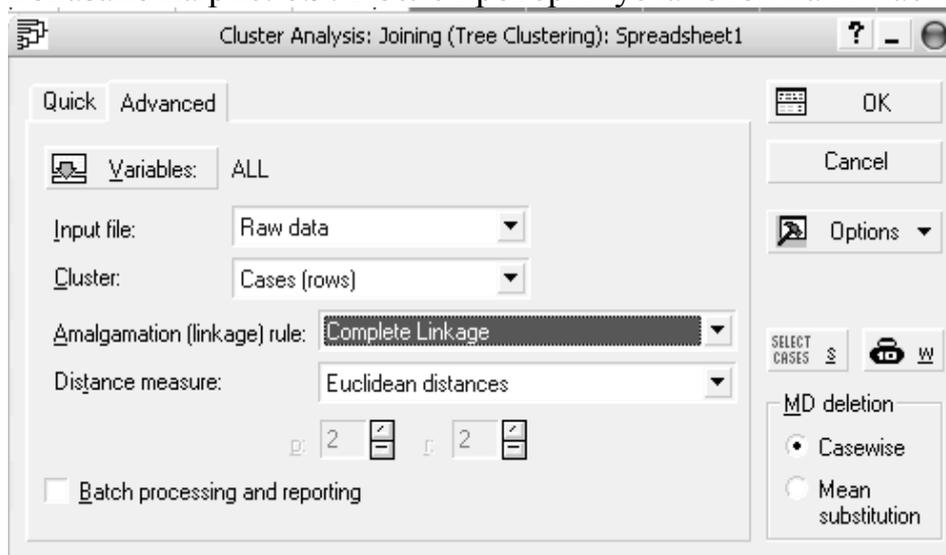


Рис. 6.8. Выбор правила объединения показателей анализа Amalgamation (linkage) rule

Метод анализа (Distance measure) по умолчанию – эвклидово расстояние, что нам полностью подходит. После всех установок окно должно выглядеть, как показано на рис. 6.9. После проверки установок нажимаем ОК.



6.9. Правильные установки для проведения кластерного анализа

6 этап – при корректном проведении анализа открывается окно (рис. 6.10), в верхней половине которого есть возможность еще раз перепроверить количество столбцов (10), количество строк (14) и другие установки.

Также в окне можно выбрать вид изображения полученного дерева кластеризации: Horizontal hierarchical tree plot или Vertical icicle plot. Они абсолютно идентичны по сути. Мы выбираем Vertical icicle plot и нажимаем Summary.

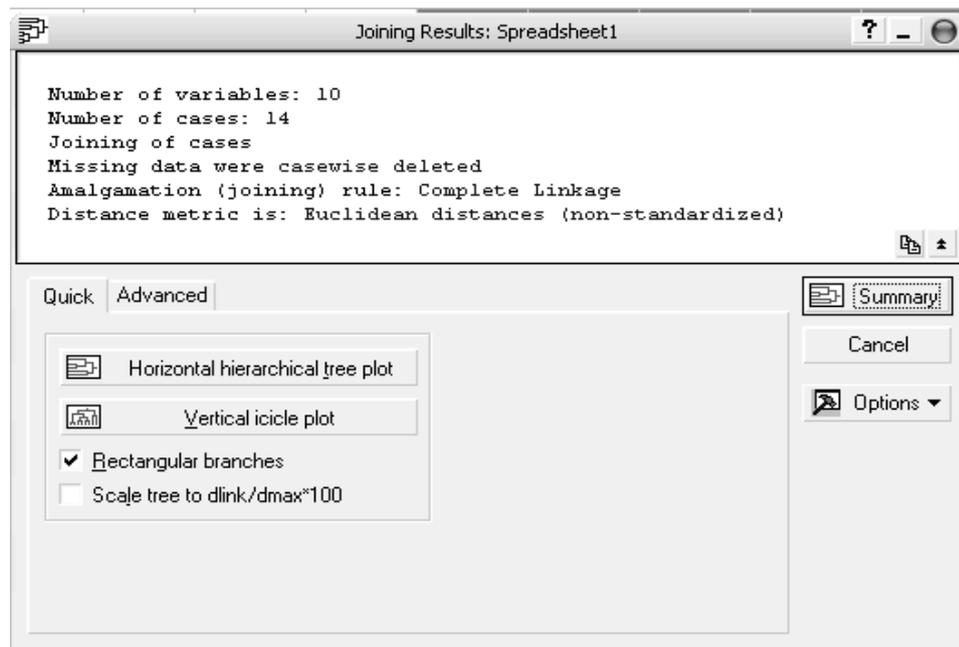


Рис.6.10. – Выбор вида изображения дерева кластеризации

7 этап – на экране появляется окно с рисунком полученного дерева кластеризации по заданным параметрам (рис. 6.11). По вертикальной оси – эвклидовы расстояния, по горизонтальной оси – регионы Чехии, где цифра показывает их порядковый номер в таблице данных (C_10 – это Прага, C_11 – Среднечешский регион и т.д.).

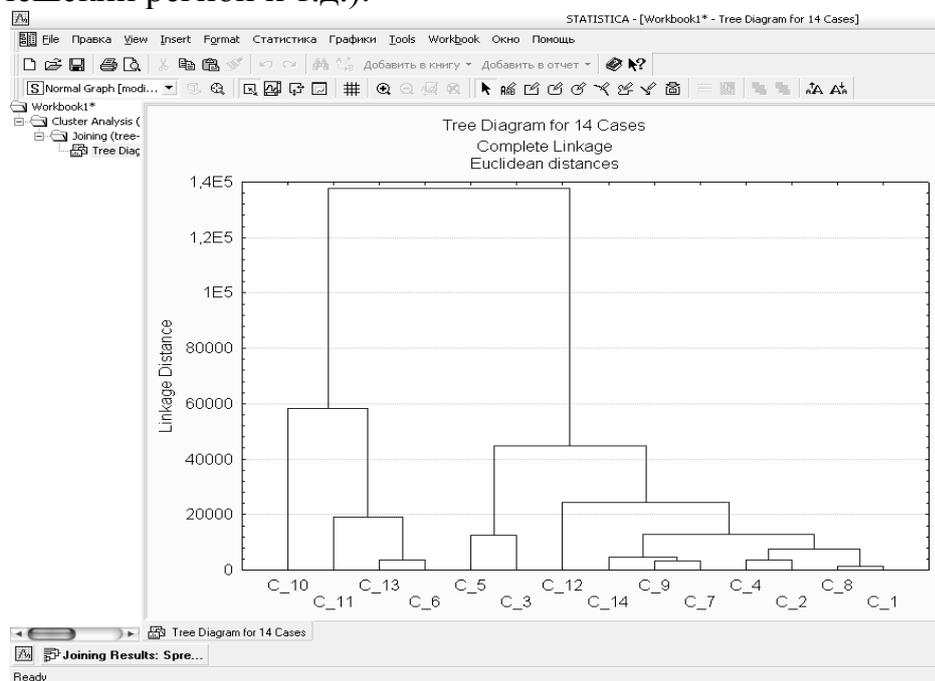


Рис. 6.11. Полученное дерево кластеризации

8 этап – полученное дерево необходимо несколько модернизировать, что сделает его более понятным и удобным для дальнейшей интерпретации. Можно убрать надписи наверху рисунка (они необязательны и не соответствуют стандарту оформления рисунков для каких-либо научных работ). Для этого выделяем надпись и дважды «щелкаем» левой клавишей мыши по ней.

В открывшемся окне ставим «стираем» не нужные нам надписи. Аналогично убираем подписи осей.

Желательно изменить цвет самого «дерева», т.к. синий цвет, который выбран программой по умолчанию, при распечатке часто плохо виден на бумаге. Для этого выделяем саму линию и после двойного щелчка левой клавиши мыши открываем окно (рис. 6.12). Выбираем в окне меню Line (линия). Открывается еще одно окно (рис. 6.13), в котором меняем синий цвет на черный или любой другой. В этом же окне можно изменить толщину линии и вид штриха. Но это уже необязательно.

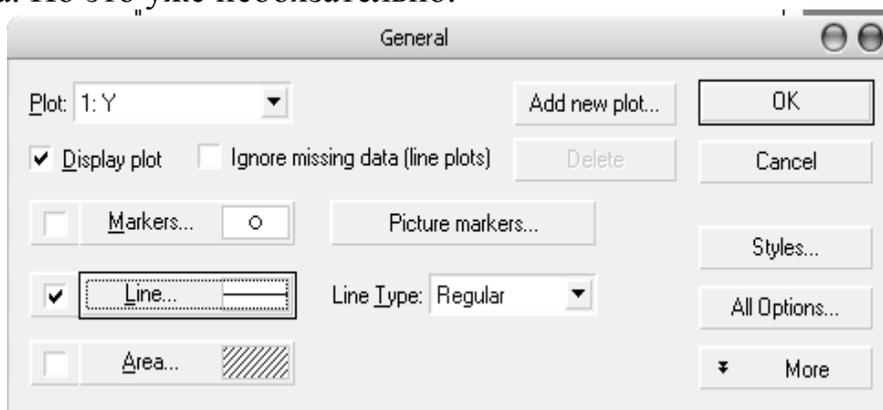
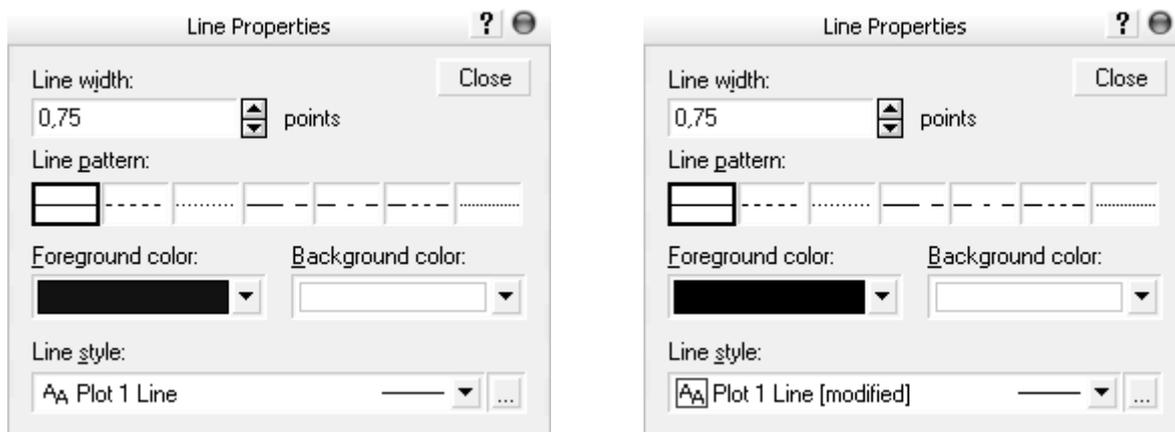


Рис. 6.12. Изменение цветовых параметров дерева кластеризации



6.13. Изменение цвета дерева кластеризации

Теперь добавляем подписи горизонтальной оси. Для этого подводим курсор мыши к подписи C_10 и после двойного щелчка левой клавиши мыши видим следующее окно – рис. 6.14.

Внизу окна находим меню Optios Layout. По умолчанию видим надпись Auto. Заходим в это меню и меняем установку на Perpendicular (перпендикуляр). Это сделает подписи горизонтальной оси перпендикулярными по отношению к оси, что необходимо при добавлении названий регионов (в противном случае, все надписи просто сольются).

Не выходя из этого окна, меняем вкладку (в верхней части окна) на Custom Units – рис. 6.15.

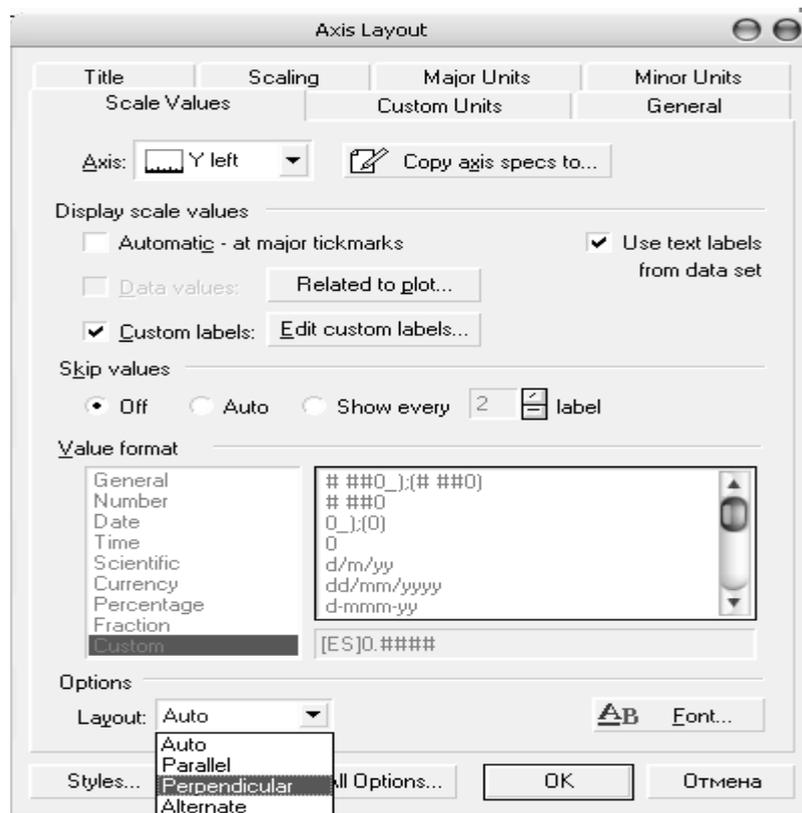


Рис. 6.14. Окно настроек и параметров рисунка. Изменение направления подписей горизонтальных осей

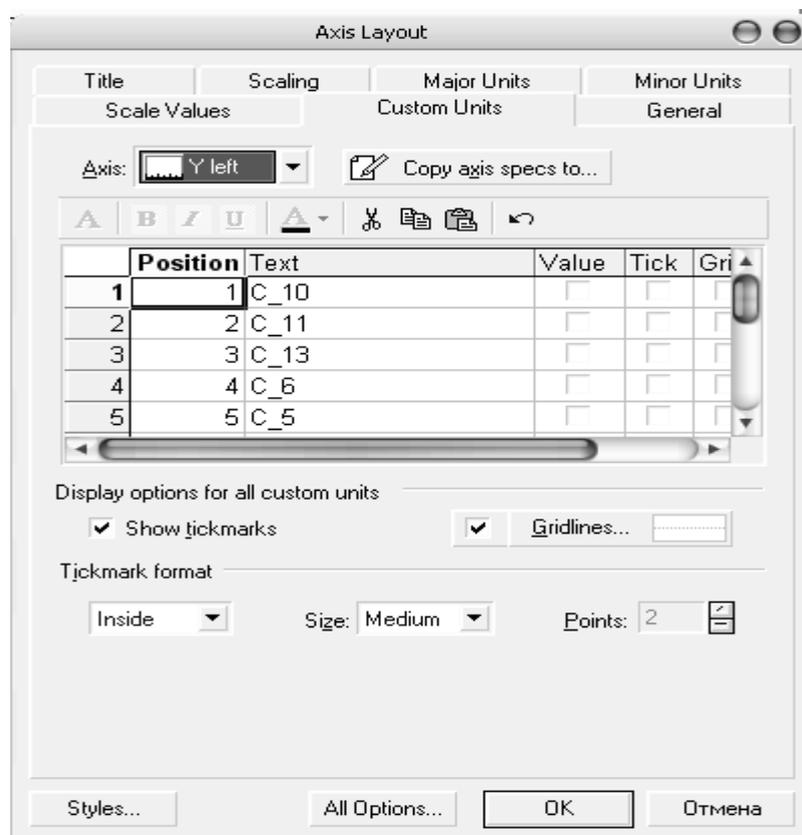


Рис. 6.15. Меню подписей горизонтальной оси

Вместо надписей C_10, C_11, C_13 на необходимом языке вписываем название регионов (рис. 6.16).

Номера показывают порядковый номер региона в таблице данных. После изменения всех подписей нажимаем ОК.

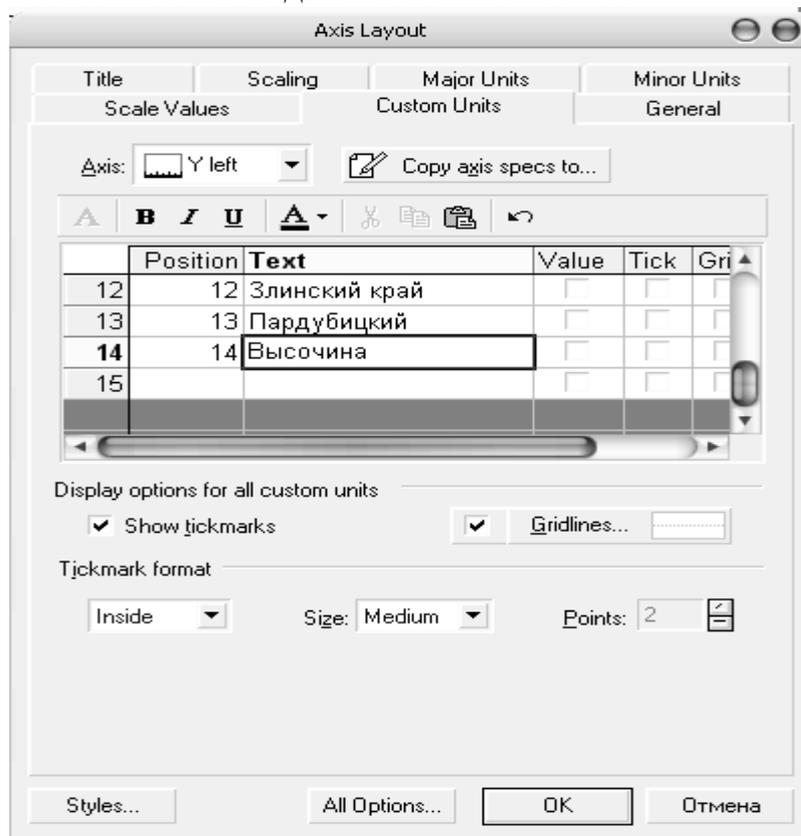


Рис. 6.16. Подписи данных горизонтальной оси

Наш рисунок изменился, что показано на рис. 6.17. В таком виде он может быть экспортирован в любой документ для непосредственно анализа.

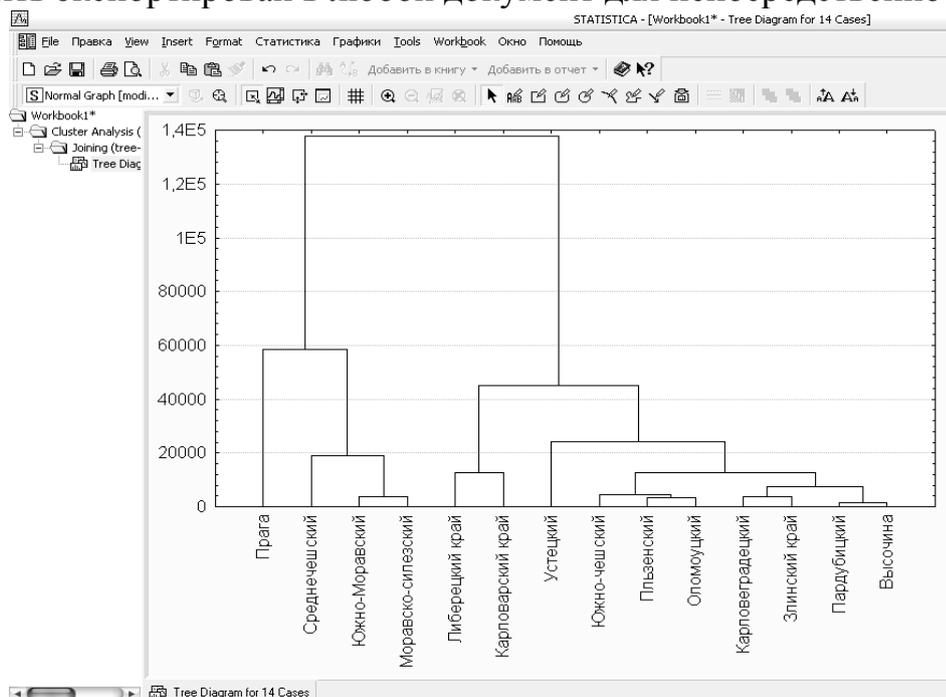
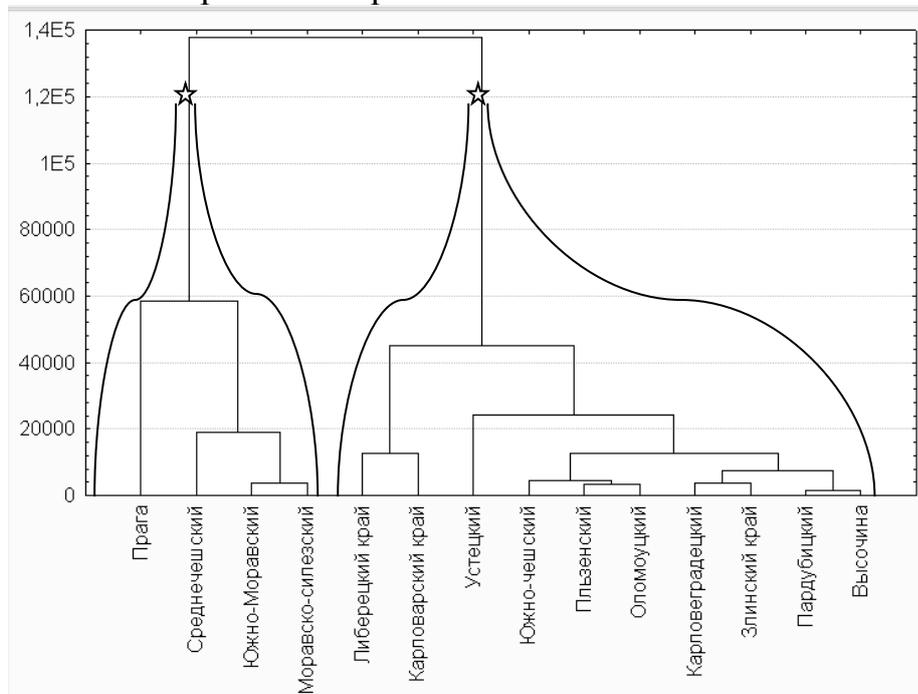


Рис. 6.17. Итоговый вариант дерева кластеризации

9 этап – интерпретация рисунка – состоит в определении, на каком уровне схожести показателей (вертикальная ось) наиболее наглядно выделение кластеров. Кластером можно назвать то объединение регионов, которое находится на пересечении вертикальной оси с самим «деревом» и стремится к горизонтальной оси. Например, на уровне 1,2E5 выделяются только два кластера (рис. 6.18). Первый объединяет такие регионы, как Прага, Средне-чешский, Южно-Моравский и Моравско-силезский районы. Все остальные регионы вошли во второй кластер.



6.18. Пример выделения кластеров

Количество кластеров составляет приблизительно $\frac{1}{4}$ от количества регионов (при наличии всего 10 регионов – количество кластеров не должно превышать 3, т.е. $\frac{1}{3}$). В нашем случае имеем 14 регионов, следовательно, будет целесообразным выделение 4 кластеров.

Четыре кластера на рисунке выделяются на уровне 40000 (рис. 6.19). Кластер может состоять и из одного региона (в примере это Прага). Столицы и столичные регионы очень часто при кластеризации выделяются в один кластер, что можно объяснить, как правило, их высокими показателями, которые сильно отличаются от показателей других регионов страны.

В таблице приводим состав четырех кластеров Чехии по выбранным показателям.

Кластеры	Регионы
Кластер 1	Прага
Кластер 2	Среднечешский, Южно-Моравский, Моравско-силезский
Кластер 3	Либерецкий, Карловарский
Кластер 4	Устецкий, Южночешский, Пльзенский, Оломоуцкий, Карловаградский, Злинский, Пардубицкий, Высочина

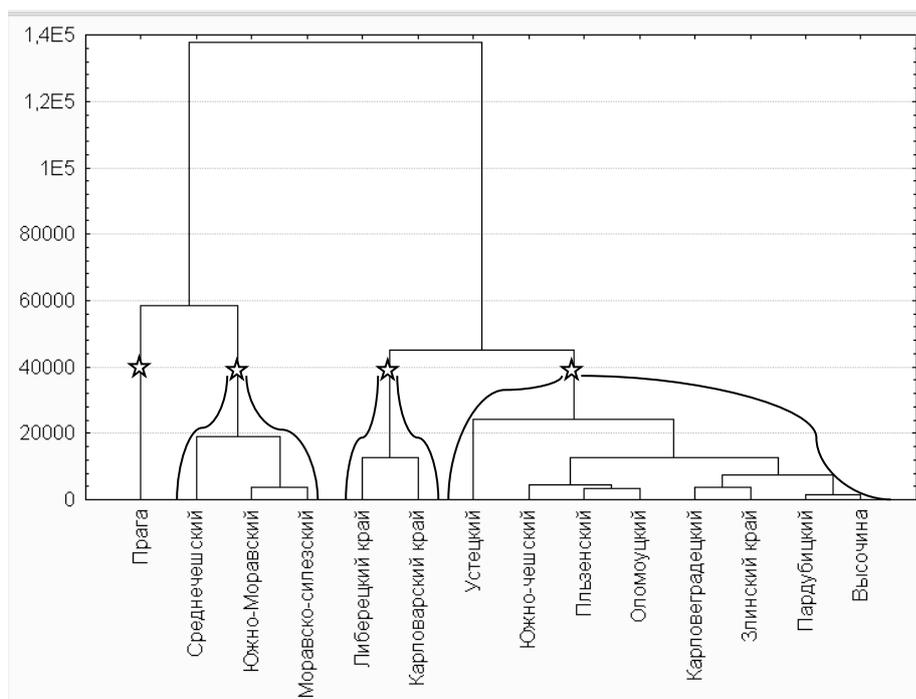


Рис. 6.19. Выделение четырех кластеров

В конце анализа необходимо дать краткую характеристику каждого из кластеров с анализом показателей, т.к. именно их схожесть для некоторых регионов дала возможность выделить несколько типов кластеров.

Навчальне видання

Галайда Ольга Євгенівна
Голіков Артур Павлович
Ханова Олена В'ячеславівна

**Методичні вказівки рішення задач
економіко-математичного моделювання
в Microsoft Excel та Statistica**

Методичний посібник