

Расчет корреляционных зависимостей в MS Excel



Подготовила учитель информатики
Яценко Е.В.

Множественная корреляция в MS Excel

При большом числе наблюдений, когда коэффициенты корреляции необходимо последовательно вычислять для нескольких выборок, для удобства получаемые коэффициенты сводят в таблицы, называемые **корреляционными матрицами.**



- **Корреляционная матрица** — это квадратная таблица, в которой на пересечении соответствующих строк и столбцов находятся коэффициент корреляции между соответствующими параметрами.

D	E	F	G	H	
		Столбец 1	Столбец 2	Столбец 3	
	Столбец 1	1			
	Столбец 2	-0.92185434	1		
	Столбец 3	0.97457559	-0.91937524	1	

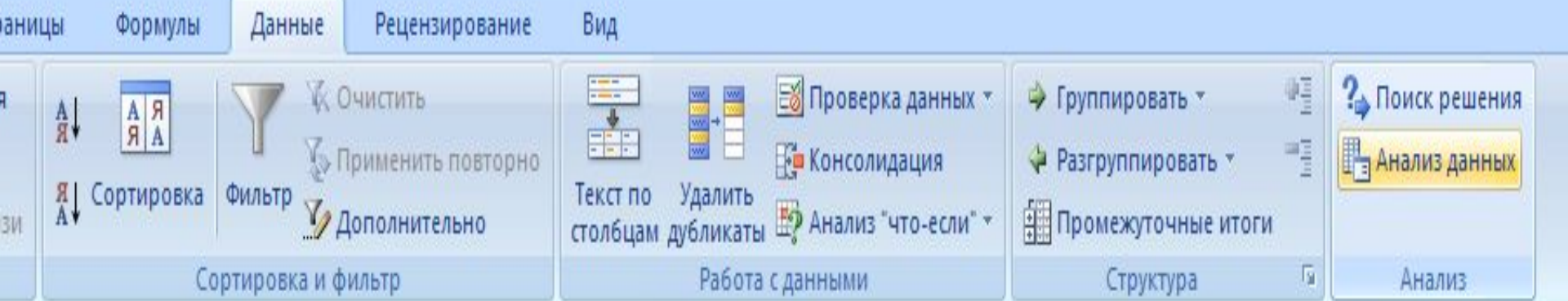
В MS Excel для вычисления корреляционных матриц используется процедура **Корреляция** из пакета **Анализ данных**.

Процедура позволяет получить корреляционную матрицу, содержащую коэффициенты корреляции между различными параметрами.



Для реализации процедуры необходимо:

1. выполнить команду **Данные - Анализ данных**;
2. в появившемся списке **Инструменты анализа** выбрать строку **Корреляция** и нажать кнопку **ОК**;
3. в появившемся диалоговом окне указать **Входной интервал**, то есть ввести ссылку на ячейки, содержащие анализируемые данные. Входной интервал должен содержать не менее двух столбцов.
4. в разделе **Группировка** переключатель установить в соответствии с введенными данными (по столбцам или по строкам);
5. указать **выходной интервал**, то есть ввести ссылку на ячейку, начиная с которой будут показаны результаты анализа. Размер выходного диапазона будет определен автоматически, и на экран будет выведено сообщение в случае возможного наложения выходного диапазона на исходные данные. Нажать кнопку **ОК**.



E F G H I J K L M N O P Q R

Анализ данных

Инструменты анализа

- Однофакторный дисперсионный анализ
- Двухфакторный дисперсионный анализ с повторениями
- Двухфакторный дисперсионный анализ без повторений
- Корреляция**
- Ковариация
- Описательная статистика
- Экспоненциальное сглаживание
- Двухвыборочный F-тест для дисперсии
- Анализ Фурье
- Гистограмма

OK
Отмена
Справка

D	E	F	G	H
		Столбец 1	Столбец 2	Столбец 3
Столбец 1		1		
Столбец 2		-0.92185434	1	
Столбец 3		0.97457559	-0.91937524	1

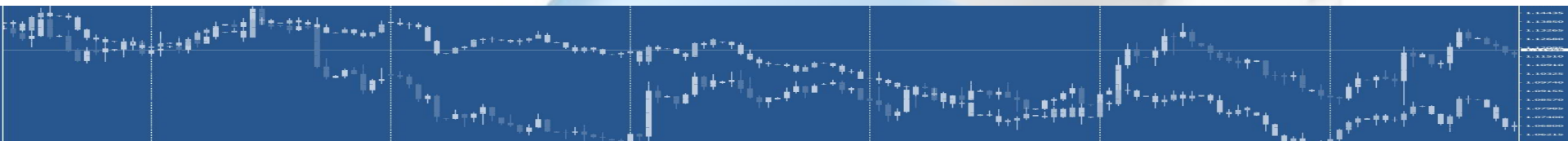
В выходной диапазон будет выведена **корреляционная матрица**, в которой на пересечении каждой строки и столбца находится коэффициент корреляции между соответствующими параметрами. Ячейки выходного диапазона, имеющие совпадающие координаты строк и столбцов, содержат **значение 1**, так как каждый столбец во входном диапазоне **полностью коррелирует сам с собой**.

- Имеются ежемесячные данные наблюдений за состоянием погоды и посещаемостью музеев и парков . Необходимо определить, существует ли взаимосвязь между состоянием погоды и посещаемостью музеев и парков.

Число ясных дней	Количество посетителей музея	Количество посетителей парка
8	495	132
14	503	348
20	380	643
25	305	865
20	348	743
15	465	541

Вывод:

видно, что корреляция между состоянием погоды и посещаемостью музея равна $-0,92$, а между состоянием погоды и посещаемостью парка — $0,97$, между посещаемостью парка и музея — $0,92$.



В результате анализа выявлены зависимости:

1. сильная степень обратной линейной взаимосвязи между посещаемостью музея и количеством солнечных дней ;
2. очень сильная прямая связь между посещаемостью парка и состоянием погоды;
3. сильная обратная взаимосвязь между посещаемостью музея и парка .