

Практическая работа

Статистические функции Excel (Calc)

Цель работы: Освоение приемов работы с функциями массивов (табличными функциями). Изучение элементарных статистических функций Excel (Calc)

Теоретические сведения

Формулы массивов (табличные формулы)

Массивом называют блок ячеек электронной таблицы, который используется для создания формул, возвращающих некоторое множество результатов или оперирующих множеством значений, а не отдельными значениями.

Формулы массивов (иногда их называют *табличными формулами*), используют несколько множеств значений (*массивов аргументов*), и возвращают одно или несколько значений. Такие формулы позволяют обращаться с блоками, как с обычной ячейкой.

Рассмотрим работу с использованием массивов на следующем примере. Требуется определить прибыль для каждого года деятельности отеля, представленного в таблице 1.

Таблица 1.

Пример использования функций массива

	A	B	C	D
1	<i>Год</i>	<i>Приход</i>	<i>Расход</i>	<i>Прибыль</i>
2	2013	200	150	{=B2:B5-C2:C5}
3	2014	360	230	{=B2:B5-C2:C5}
4	2015	410	250	{=B2:B5-C2:C5}
5	2016	200	180	{=B2:B5-C2:C5}

Выделим блок D2:D5. Начнем ввод формулы – наберем знак =. Выделим блок B2:B5, наберем знак минус -, выделим блок C2:C5. Ввод формул массива заканчивается комбинацией клавиш Ctrl+Shift+Enter. После нажатия такой комбинации во всех ячейках блока D2:D5 появится формула {B2:B5-C2:C5}.

Основные правила работы с формулами массива:

- перед вводом формулы нужно выделить ячейку или диапазон для результатов, если формула возвращает несколько значений, то диапазон результатов должен быть того же размера, что и диапазон исходных данных;

- фигурные скобки, отмечающие формулу массива, вводятся при завершении ввода формулы клавишами **Ctrl+Shift+Enter**, если фигурные скобки ввести вручную, такой ввод будет воспринят Excel(Calc) как текст.
- для редактирования формулы массива необходимо выделить блок, активировать строку формул, внести изменения и завершить редактированием клавишами **Ctrl+Shift+Enter**;

Функции Excel (Calc), используемые для статистического анализа

Статистический анализ данных необходим для оценки деятельности фирмы и прогноза ее работы на какой-то срок. Такой анализ основывается на сборе информации, определении по представленным массивам данных оценок, статистических показателей и тенденций развития фирмы.

МАКС(число1;число2; ...) ;Calc – MAX - возвращает наибольшее значение из набора значений.

МИН(число1;число2; ...) ;Calc - MIN- возвращает наименьшее значение из набора значений.

СРЗНАЧ(число1; число2; ...) ;Calc -AVERAGE - возвращает среднее (арифметическое) своих аргументов.

ТЕНДЕНЦИЯ (известные_значения_у; известные_значения_х; новые_значения_х; конст) ;Calc

- TREND - возвращает значения в соответствии с линейным трендом, т.е. аппроксимирует прямой линией (по методу наименьших квадратов) массивы "известные_значения_у" и "известные_значения_х". Возвращает значения у, в соответствии с этой прямой для заданного массива новые_значения_х.

РОСТ (известные_значения_у;известные_значения_х;новые_значения_х;конст);Calc - GROWTH - возвращает значения у для последовательности новых значений х, задаваемых с помощью существующих х- и у-значений, т.е. функция рассчитывает прогнозируемый экспоненциальный рост на основании имеющихся данных.

ПРЕДСКАЗ(х, известные_значения_у, известные_значения_х); Calc - FORECAST вычисляет или предсказывает будущее значение по существующим значениям. Предсказываемое значение — это значение у, соответствующее заданному значению х. Значения х и у известны; новое значение предсказывается с использованием линейной регрессии. Эту функцию можно использовать для прогнозирования будущих продаж, потребностей в оборудовании или тенденций потребления.

ИНДЕКС(массив; номер_строки; [номер_столбца]); Calc - INDEX – - возвращает значение элемента таблицы или массив, заданного номером строки и номером столбца, используется для поиска значения ячейки, находящейся на пересечении заданных строки и столбца.

Массив — обязательный аргумент. Диапазон ячеек или константа массива.

Если массив содержит только одну строку или один столбец, соответствующий аргумент "номер_строки" или "номер_столбца" не является обязательным.

Если массив содержит больше одной строки и одного столбца, а из аргументов "номер_строки" и "номер_столбца" задан только один, функция ИНДЕКС возвращает массив, состоящий из целой строки или целого столбца аргумента "массив".

Номер_строки — обязательный аргумент. Выбирает строку в массиве, из которой требуется вернуть значение. Если аргумент "номер_строки" опущен, аргумент "номер_столбца" является обязательным.

Номер_столбца — необязательный аргумент. Выбирает столбец в массиве, из которого требуется возвратить значение. Если аргумент "номер_столбца" опущен, аргумент "номер_строки" является обязательным.

ПОИСКПОЗ(искомое_значение;просматриваемый_массив;[тип_сопоставления]); Calc - MATCH.

Искомое_значение. Обязательный аргумент. Значение, которое сопоставляется со значениями в аргументе *просматриваемый_массив*. Например, при поиске номера в телефонной книге имя абонента указывается в качестве искомого значения, а нужным значением будет номер телефона.

Аргумент *искомое_значение* может быть значением (числом, текстом или логическим значением) или ссылкой на ячейку, содержащую такое значение.

Просматриваемый_массив Обязательный аргумент. Диапазон ячеек, в которых производится поиск.

Тип_сопоставления. Необязательный аргумент. Число -1, 0 или 1.

Аргумент *тип_сопоставления* указывает, каким образом в Microsoft Excel *искомое_значение* сопоставляется со значениями в аргументе *просматриваемый_массив*. По умолчанию в качестве этого аргумента используется значение 1.

В приведенной ниже таблице описано, как функция находит значения в зависимости от аргумента *тип_сопоставления*.

ЗАДАНИЕ 1. Создать таблицу с исходными данными (табл.2)

Таблица 2.

Исходные данные

	А	В	С	Д	Е	Ф	Г
1	Отель	1 кв.	2 кв.	3 кв.	4 кв.	Среднее по отелю	Оценка
2	<i>Швеция</i>	1500	2000	6000	8000		
3	<i>Дания</i>	1400	5000	4100	5000		
4	<i>Норвегия</i>	3600	3600	3000	4500		
5	<i>Финляндия</i>	1100	1045	9100	7800		
6	<i>Германия</i>	3850	3650	7800	11000		
7	<i>Польша</i>	6800	7250	8122	9450		
8	<i>Чехия</i>	6590	7050	6400	6440		
9	<i>Словакия</i>	930	3970	4512	4600		
10	<i>Венгрия</i>	8912	7490	3570	8000		
11	<i>Болгария</i>	3590	3800	5464	5954		
12	Мин						
13	Мах						
14	Среднее						
15		1	2	3	4	5	6
16	Тенденция по среднему						
17	Рост по среднему						

18	Погрешность в тенденции						
19	Погрешность в роста						
20	Лучший отель по сети						
	Доход						

Для приведенных в таблице 2 данных о реализации гостиничных услуг сетью отелей
вычислить:

- минимальные, максимальные и среднее показатели по каждому кварталу;
- средние показатели по каждому отелю;
- вычислить средний доход по всей сети отелей за отчетный период;
- дать оценку работы каждого отеля: «хорошо», если доход отеля превышает средний по сети, и «плохо», если доход меньше среднего по сети;
- построить линейную и экспоненциальную модель деятельности сети отелей и дать прогноз для двух следующих кварталов;
- оценить относительные отклонения для среднего значения и «Тенденции», для среднего значения и «Роста».

ТЕХНОЛОГИЯ ВЫПОЛНЕНИЯ ЗАДАНИЯ

1. Минимальные, максимальные и средние значения по кварталам и средние значения по турам подсчитываются с помощью **Мастера функций**.
2. Для оценки работы отеля (столбец G)) используется среднее значение дохода по сети и функция **ЕСЛИ (IF)**.
3. Функция **Тенденция (TREND)** показывает динамику изменения данных и позволяет получить прогноз на будущее. При этом изменение данных описывается линейным уравнением. Для определения **Тенденции (TREND)**:

- Выделить новый диапазон ячеек для размещения результатов (B16:E16);
- В строке формул вставить функцию **Тенденция (TREND)** и в **Мастере функций** в поле аргумента *известные_значения_у* указать диапазон средних по кварталу значений.
- *Известные_значения_х* можно не устанавливать, т.к. это 1, 2, 3, 4 кварталы.
- Выйти из Мастера функций – **Ok**.

- Установить курсор в строке формул, нажать комбинацию клавиш **Ctrl+Shift+Enter**, в выделенном новом массиве появятся результаты.

4. Функция **Тенденция (TREND)** показывает линейную модель изменения показателей, экспоненциальная модель строится функцией **Рост (GROWTH)**. Самостоятельно вычислите функцию **Рост (GROWTH)** для средних по кварталам, подобно тому, как вычислялась функция **Тенденция**.
5. Вычислить прогноз развития событий на ближайшие два квартала, используя функцию **Тенденция (TREND)**:

- Справа от ячейки со значением **Тенденция (TREND)** для 4-го квартала выделить две свободные ячейки.
- Вставить функцию **Тенденция (TREND)** и в **Мастере функций** указать:
- в поле *известные_значения_у* вычисленные ранее значения **Тенденция** за четыре квартала (диапазон B16:E16);
- в поле *новые_значения_х* – диапазон F15:G15 – кварталы 5 и 6, для которых выполняется прогноз.
- Завершить работу Мастера – **Ok**, завершить ввод функции массива **Ctrl+Shift+Enter**, в выделенных ячейках появятся предсказанные по линейной модели значения для 5 и 6 кварталов.

6. Таким же образом рассчитать прогноз по экспоненциальной модели с помощью функции **Рост (GROWTH)**.
7. Оценить относительные отклонения в процентах для **среднего значения и Тенденции (TREND)**, для **среднего значения и Роста** (для каждого из четырех кварталов) по формуле:

$$\text{Относительное отклонение} = (y_{\text{факт}} - y_{\text{модели}}) / y_{\text{модели}}$$

где $y_{\text{факт}}$ – среднее значение;

$y_{\text{модели}}$ – значение, определенное с помощью **Тенденции** или **Роста**.

ЗАДАНИЕ 2

Определить лучший отель по сети за квартал и его доход, используя функции (ИНДЕКС) INDEX и (ПОИСКПОЗ) MATCH.

Пример расчета показателей работы отелей по первому кварталу в таблице 3, результат работы представлен в таблице 4.

Таблица 3.

Пример расчета показателей работы отелей по первому кварталу

	<i>A</i>	<i>B</i>
12	Мин	=МИН(B2:B11)
13	Мах	=МАКС(B2:B11)
14	Среднее	=СРЗНАЧ(B2:B11)
16	Тенденция по среднему	=ТЕНДЕНЦИЯ(B14:E14;B15:E15)
17	Рост по среднему	=РОСТ(B14:E14;)
18	Погрешность тенденции	=(B14-B16)/B16
19	Погрешность роста	=(B14-B17)/B17
20	Лучший отель по сети	=INDEX(\$A\$2:B11;MATCH(MAX(B2:B11);B2:B11;0);1)
21	Доход	=INDEX(\$A\$2:B11;MATCH(MAX(B2:B11);B2:B11;0);2)

Результаты расчетов приведены в таблице 4.

Таблица 4.

Результаты расчетов

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>
<i>1</i>	Отель	1 кв.	2 кв.	3 кв.	4 кв.	Среднее по отелю	Оценка
2	Швеция	1500	2000	6000	8000	4375	Плохо
3	Дания	1400	5000	4100	5000	3875	Плохо
4	Норвегия	3600	3600	3000	4500	3675	Плохо
5	Финляндия	1100	1045	9100	7800	4761,25	Плохо
6	Германия	3850	3650	7800	11000	6575	Хорошо
7	Польша	6800	7250	8122	9450	7905,5	Хорошо
8	Чехия	6590	7050	6400	6440	6620	Хорошо
9	Словакия	930	3970	4512	4600	3503	Плохо
10	Венгрия	8912	7490	3570	8000	6993	Хорошо
11	Болгария	3590	3800	5464	5954	4702	Плохо
12	Мин	930	1045	3000	4500		
13	Мах	8912	7490	9100	11000		
14	Среднее	3827	4486	5807	7074	5298	
15		1	2	3	4	5	6
16	Тенденция по среднему	3639	4745	5852	6958	8064	9170
17	Рост по среднему	3760	4639	5724	7063	8714	10752
18	Погрешность тенденции	5,17%	-5,48%	-0,77%	1,67%		
19	Погрешность роста	1,79%	-3,32%	1,44%	0,17%		
20	Лучший отель по сети	Венгрия	Венгрия	Финляндия	Германия		

21	Доход	8912	7490	9100	11000
----	-------	------	------	------	-------

ЗАДАНИЕ 3

Построить диаграмму – график изменения доходов по кварталам и тенденцию изменения доходов по кварталам, включая прогноз на два следующих квартала, а также рост изменения доходов по кварталам.

Пример для отеля «Венгрия» представлен на диаграмме 1.

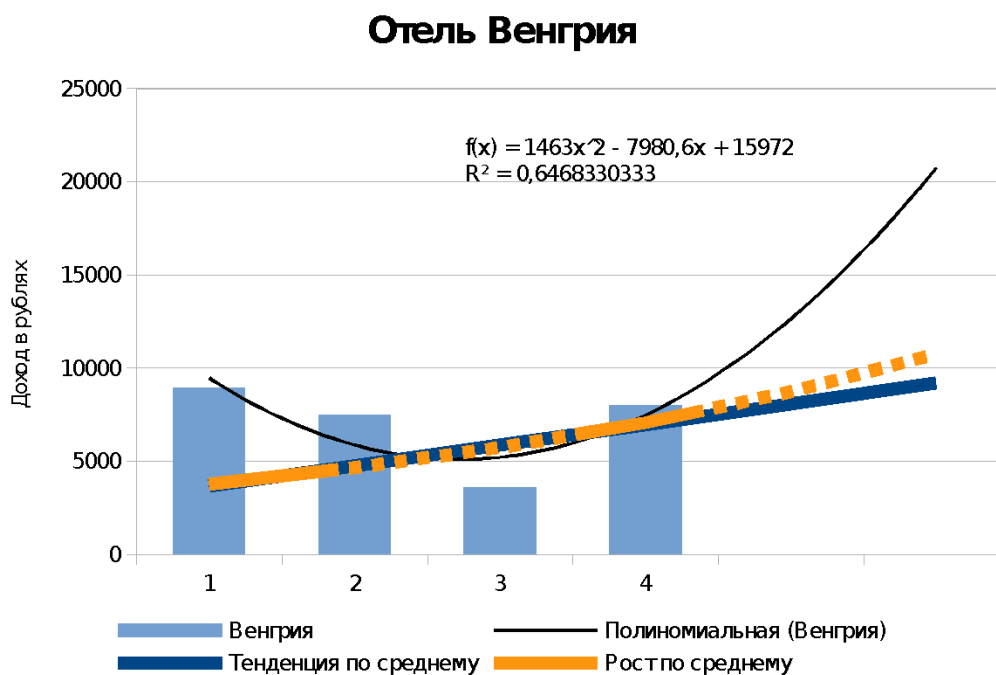
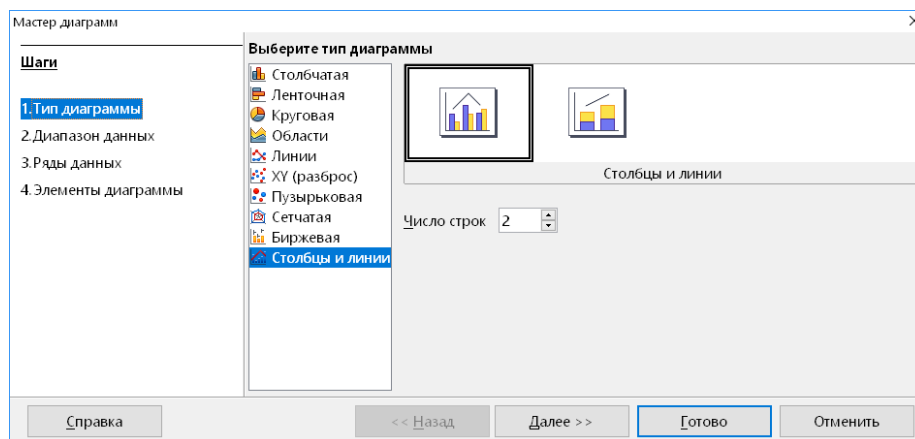


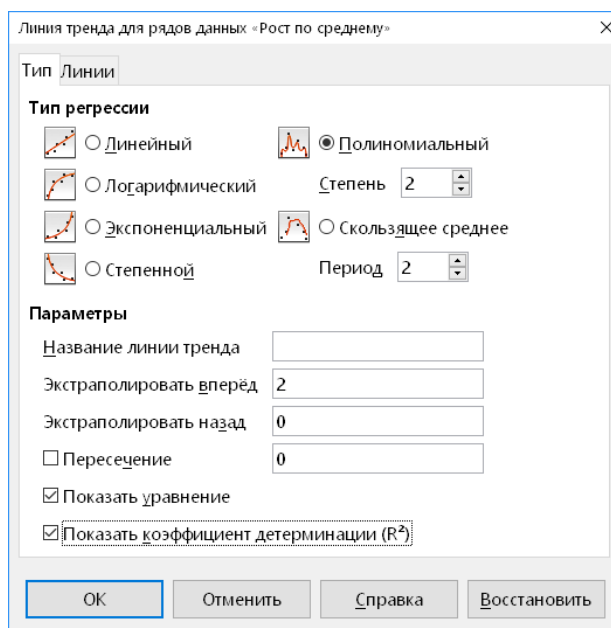
Диаграмма 1

ТЕХНОЛОГИЯ ВЫПОЛНЕНИЯ

1. Построить диаграмму, выделив:
 - изменения доходов по кварталам отеля ,
 - тенденцию изменения доходов по кварталам, включая прогноз на два следующие квартала
 - рост изменения доходов по кварталам.



2. В контекстном меню «Вставить линию тренда». В открывшемся окне на вкладке «Тип» выберите необходимый тип линии тренда: линейной, логарифмической, экспоненциальной, степенной или иной, что в математическом эквиваленте также означает и способ аппроксимации данных. Если же вы сомневаетесь в выборе типа аппроксимации, можете построить несколько линий, а для большей точности прогноза в разделе «Параметры» этого же окна отметить флажком пункт «поместить на диаграмму величину достоверности аппроксимации (R^2)».



3.

Сравнивая значения R^2 для разных линий, вы сможете выбрать тот тип графика, который характеризует ваши данные наиболее точно, а, следовательно, строит наиболее достоверный прогноз. Чем ближе значение R^2 к единице, тем точнее вы выбрали тип линии. Здесь же, на вкладке «Параметры», вам необходимо указать период, на который делается прогноз.