**Раздел III. Статистические методы исследований**

**Тема 2. Описательная статистика выборки**

**Практическая работа по теме «Описательная статистика выборочных данных»**

Проведена проверка в 100 компаниях. Даны значения количества работающих в компании (чел.):

|  |  |
| --- | --- |
| 23 25 24 25 30 24 30 26 28 26 32 33 31 31 25 33 25 29 30 28 23 30 29 24 33 30 30 28 26 25 26 29 27 29 26 28 27 26 29 28 29 30 27 30 28 32 28 26 30 26 31 27 30 27 33 28 26 30 31 29 27 30 30 29 27 26 28 31 29 28 33 27 30 33 26 31 34 28 32 22 29 30 27 29 34 29 32 29 29 30 29 29 36 29 29 34 23 28 24 28 |  |

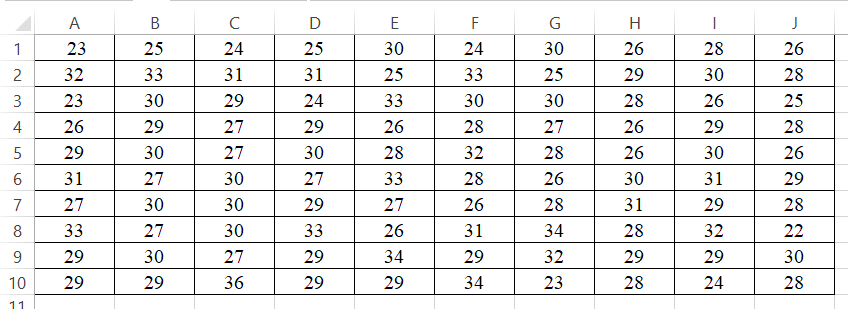
Рассчитать числовые характеристики:

* Сумму;
* Моду;
* Медиану;
* Размах ряда;
* Среднее арифметическое;
* Среднее линейное отклонение;
* Дисперсия по генеральной совокупности ;
* Дисперсия по выборке;
* Среднее квадратичное отклонение;
* Смещенное среднее отклонение (по выборке).

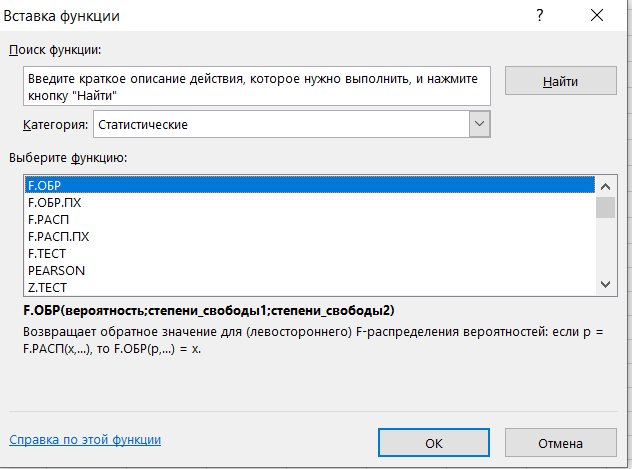
А также построить:

* полигон частот;
* столбчатую и круговую диаграммы.

1. Занести данные в Excel, каждое число в отдельную ячейку:



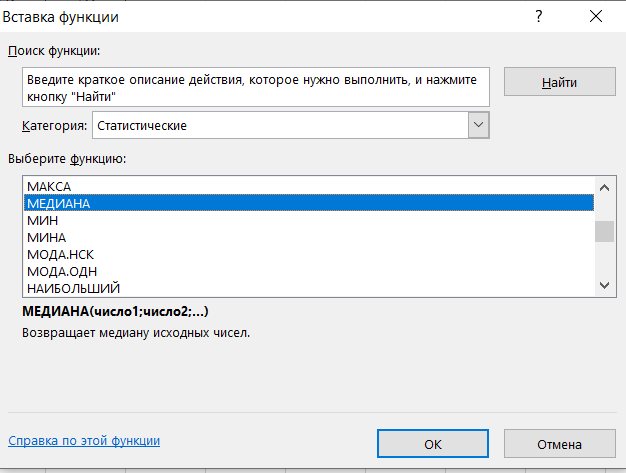
2. Для расчета числовых характеристик используем опцию Вставка – Функция. И в появившемся окне в строке категория выберем – статистические:



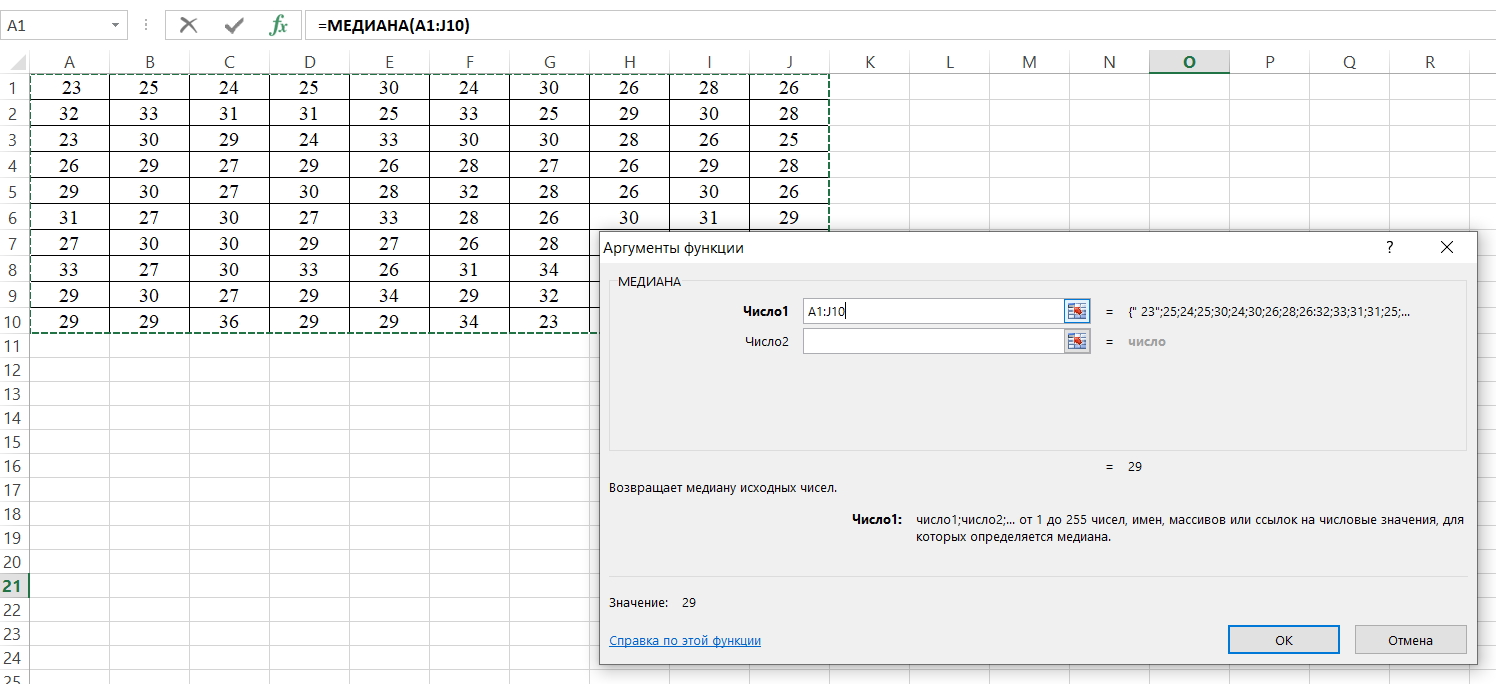
Функции Excel для расчета статистических величин, представлены в таблице:

|  |  |
| --- | --- |
| Статистическая величина | Функция |
| Сумма | СУММ() |
| Мода | МОДА() |
| Медиана | МЕДИАНА() |
| Размах ряда | МАКС()-МИН() |
| Среднее арифметическое | СРЗНАЧ() |
| Среднее линейное отклонение | СРОТКЛ() |
| Дисперсия по генеральной совокупности | ДИСПР() |
| Дисперсия по выборке | ДИСП() |
| Среднее квадратичное отклонение | СТАНДОТКЛОНП() |
| Смещенное среднее отклонение (по выборке) | СТАНДОТКЛОН() |

Например, для расчета медианы используем функцию МЕДИИАНА(). Для этого выбираем соответствующую функцию:



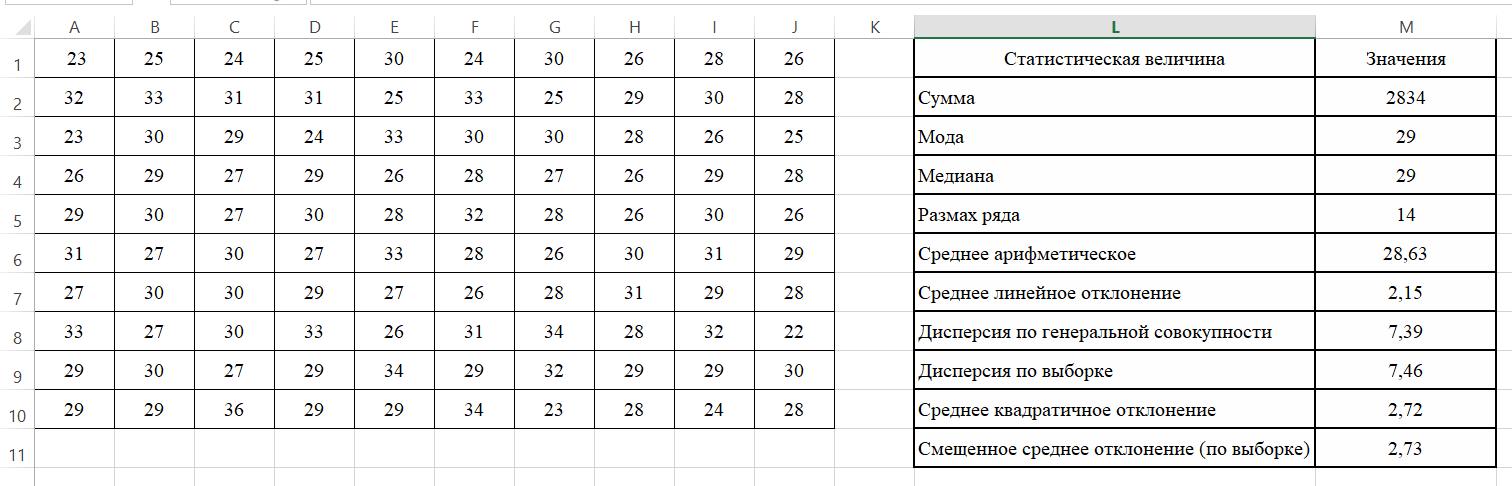
В поле Число 1 ставим курсор и мышкой выделяем нашу таблицу:



Нажимаем клавишу ОК. Получили Ме= 29 (чел) – среднее значение сотрудников в фирме.

*Примечание.* Размах ряда чисел – разница между наименьшим и наибольшим возможным значением случайной величины. Для вычисления размаха ряда нужно найти наибольшее и наименьшее значения нашей выборки и вычислить их разность.

3. Рассчитать все статистические величины, в результате должно получиться:



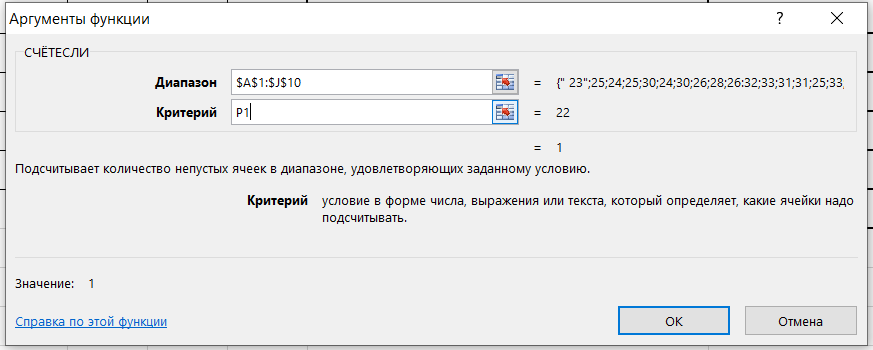
4. Для построения диаграммы и полигона частот необходимо задать закон распределения, т.е. составить таблицу значений случайной величины и соответствующих им частот. Находим наименьшее число сотрудников в фирме МАКС() = 22, а наибольшее МИН()= 36. Составим таблицу, в которой значения *xi*случайной величины меняются от 22 до 36 включительно шагом 1.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *xi* | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 |
| *ni* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

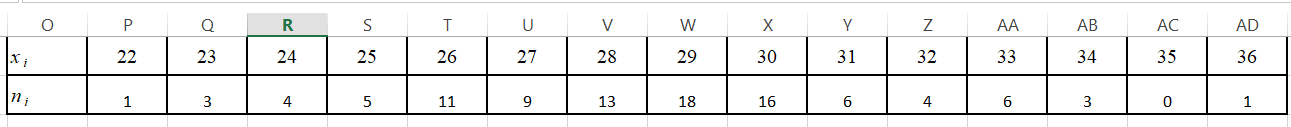
Чтобы сосчитать частоту каждого значения воспользуемся

Вставка – Функция – Статистические – СЧЕТЕСЛИ().

В окне Диапазон ставим курсор и выделяем нашу выборку, а в окне Критерий ставим ссылку на число 22

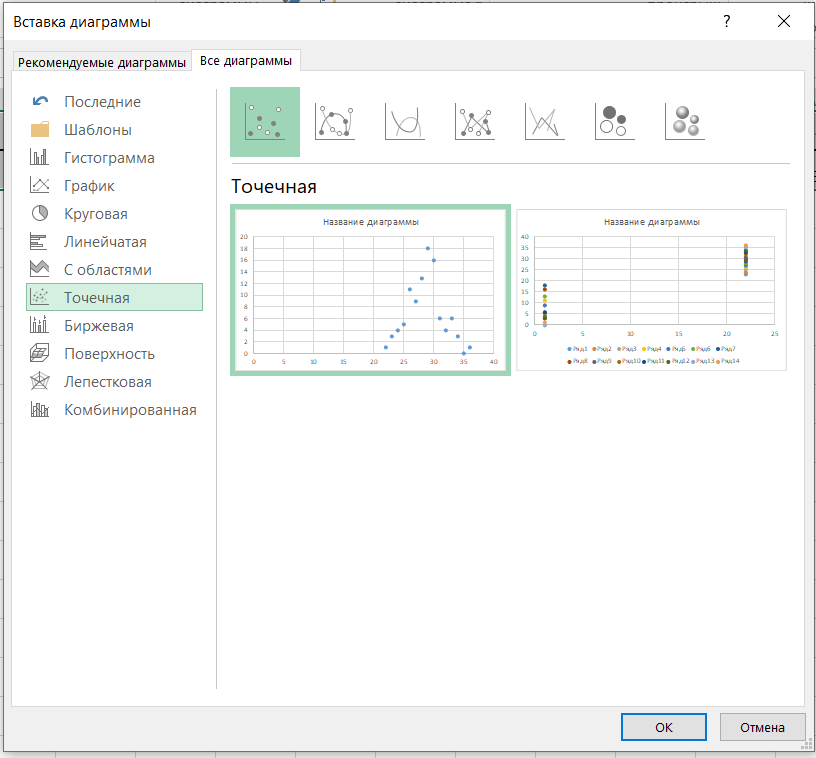


Нажимаем клавишу ОК, получаем значение 1, т.е. число 22 в нашей выборке встречается 1 раз и его частота =1. Аналогичным образом заполняем всю таблицу:



 Для проверки вычисляем объем выборки, сумму частот (Вставка – Функция – Математические – СУММА()). Должно получиться 100 (количество всех фирм).

Чтобы построить полигон частот выделяем таблицу – Вставка – Диаграмма – Стандартные – Точечная (точечная диаграмма на которой значения соединены отрезками)



Нажимаем клавишу Далее, в Мастере диаграмм указываем название диаграммы (Полигон частот), удаляем легенду, редактируем шкалу и характеристики диаграммы для наибольшей наглядности.

Получаем:

Для построения столбчатой и круговой диаграмм используем тот же путь (выбирая нужный нам тип диаграммы).

Диаграмма – Стандартные – Круговая.

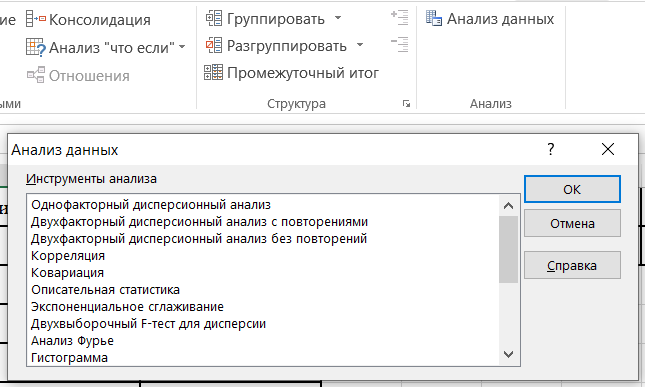
Диаграмма – Стандартные – Гистограмма.

**Задание 2.**

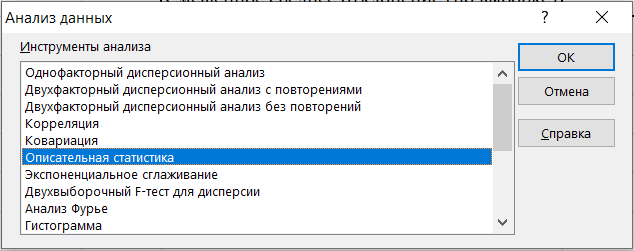
Сравнить полученные в Задание 1 рассчитанные статистические величины с показателями, рассчитанными через инструмент «Анализ данных».

*Примечание.* Предварительно разместить исходные данные в один столбец.

1. На вкладке «Данные» в группе «Анализ» выбрать команду «Анализ данных».



2. В открывшемся диалоговом окне выбрать строку «Описательная статистика».

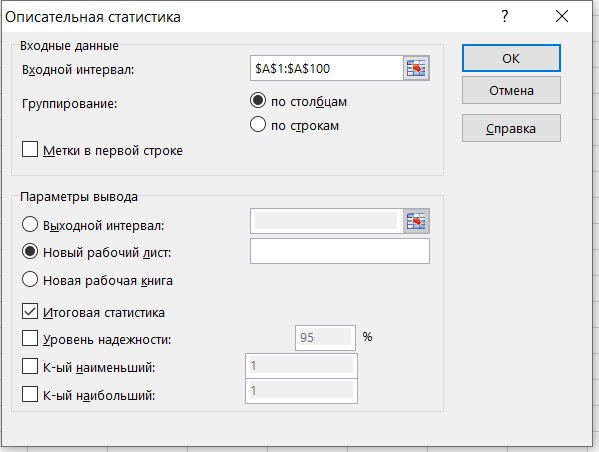


3. В поле Входной интервал указать диапазон данных, для которых надо получить статистические оценки.

*Примечание.* Если диапазон данных выделен вместе с заголовком, установить флажок Метки в первой строке

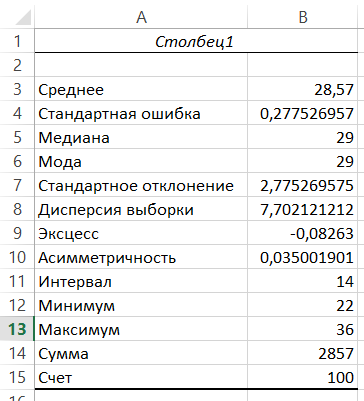
5. Выбрать вариант размещения выходных данных: текущий рабочий лист, новый рабочий лист или новая рабочая книга. В случае размещения выходных данных на текущем листе включить режим Выходной интервал и указать левую верхнюю ячейку диапазона, в который должны быть выведены результаты.

6. Установить флажок Итоговая статистика.



7. Щелкнуть по кнопке ОК.

В результате получаем:



**ЗАДАНИЕ.** По данным из Практической работы (тема 3\_1) найти описательные статистики выборки через непосредственное вычисление по формулам и через надстройку «Анализ данных». Сравнить полученные результаты. Построить полигон частот; столбчатую и круговую диаграммы.