

**Метрология, стандартизация и
сертификация**

Практика 1

**Предварительный анализ данных
Описательные статистики**

Исходные данные

Создать файл в Excel: **Фамилия_МСС_Пр01**

Исходные данные ($y=f(x)$) вычислить согласно параметрам в файле **МСС_Пр01_Распределение (2018Н2).xls**

Задача 1.

Найти для вычисленной функции $y=f(x)$ среднее арифметическое, среднее геометрическое (если возможно) и среднее квадратическое по формулам

$$\bar{X}_{\text{ср.прост.}} = \frac{\sum_{i=1}^N X_i}{N} \quad \bar{X} = \sqrt[m]{X_1 \cdot X_2 \cdot \dots \cdot X_m} = \sqrt[m]{\prod X_i}$$

$$\bar{X} = \sqrt{\frac{\sum X^2}{N}}$$

Проверить вычисление среднего арифметического и среднего геометрического при помощи функций Excel: =СРЗНАЧ() и СРГЕОМ()

Задача 2.

Построение гистограмм для ряда данных y , полученного в **Задаче 1**.

Подсказка

Создать «Массив_интервалов», т.е. диапазоны, в пределах которых будут лежать наши значения.

Для подготовки «Массива интервалов» найти:

1.1 Для оценки оптимального для нашего массива данных количества интервалов n можно воспользоваться формулой Стерджесса: https://ru.wikipedia.org/wiki/Правило_Стерджеса

$$n \sim 1 + [3,322 * \lg N]$$

N — количество всех значений величины (функция СЧЕТ в Excel).

\lg - десятичный логарифм (функция LOG10 в Excel).

$[x]$ — целая часть числа x (функция ОТБР в Excel)

1.2. Найти максимальное и минимальное значения y с помощью функций Excel: МАКС и МИН

1.3. Найти ширину интервала для «Массив_интервалов»: (МАКС-МИН)/ n

1.4. Создаем столбец со значениями границ интервалов от (МИН+ширина интервала) до (МАКС-ширина интервала) с шагом равным вычисленной ширине. Получили массив границ интервалов

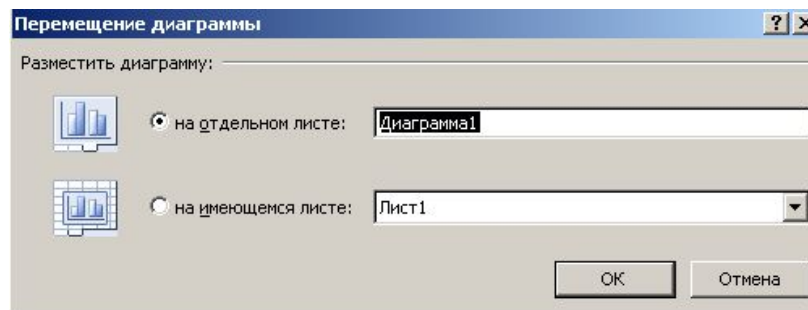
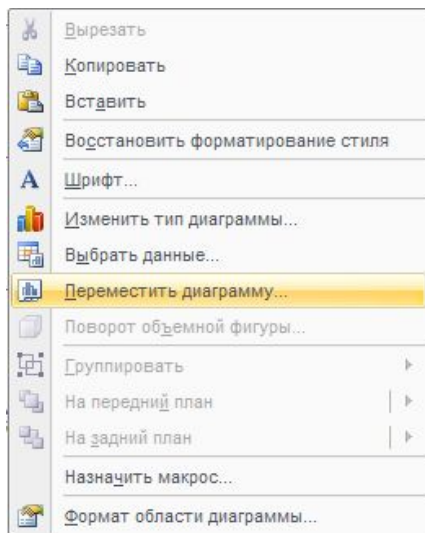
1.5. Выделяем ячейки рядом с интервалами (количество элементов в возвращаемом массиве на единицу больше количества элементов в аргументе массив_карманов), нажимаем «F2» и вводим функцию Excel ЧАСТОТА(Массив_данных; Массив_интервалов) и нажать Ctr+Shift+Enter. В столбце напротив границ интервалов появилось количество значений исходного массива, которые попадают в интервалы.

1.6. Найти сумму частот. Проверить равенство суммы и ранее вычисленной величины N

1.7. Вычислить относительные частоты (разделить значения из столбца с абсолютными частотами на их сумму и умножить на 100)

1.8. Построить в Excel на **отдельном листе D-Гист** гистограмму (обычная - гистограмма с группировкой) для полученных распределений относительных частот.

«**На отдельном листе**»: щелкнуть право клавишей мыши на диаграмме и задать команду «Переместить диаграмму» как показано на рисунках ниже.



1.10. Сделать подпись вертикальной оси **Относительная частота, %**.
Сделать подписи горизонтальной оси **Величина**

1.11. Создать столбец с подписями границ диапазонов - текстовые значения вида «мин-макс.»
Нанести на диаграмму эти подписи под соответствующими столбцами

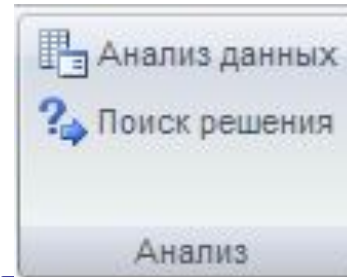
Задача 3

Воспользуйтесь пакетом «Анализ данных» в MS Excel и с помощью инструмента «Описательная статистика» рассчитайте показатели .

Подсказка

Воспользуйтесь пакетом «Анализ данных» в MS Excel и с помощью инструмента «Описательная статистика» рассчитайте показатели

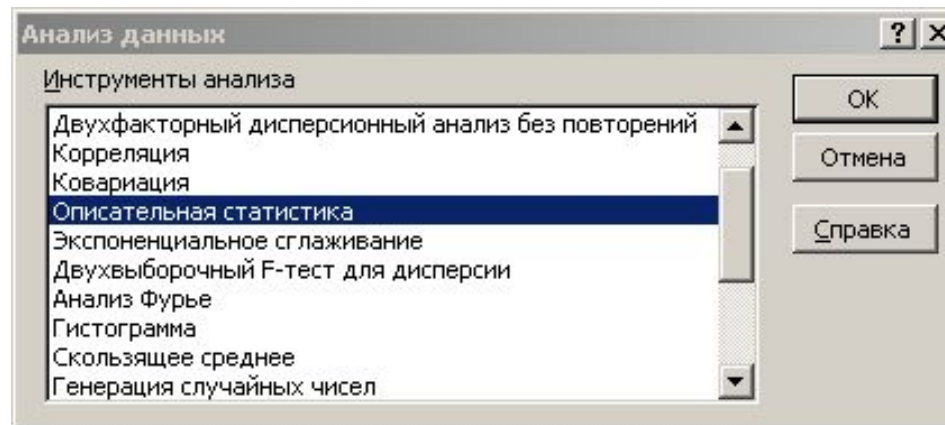
1. На вкладке «Данные» справа должна быть группа



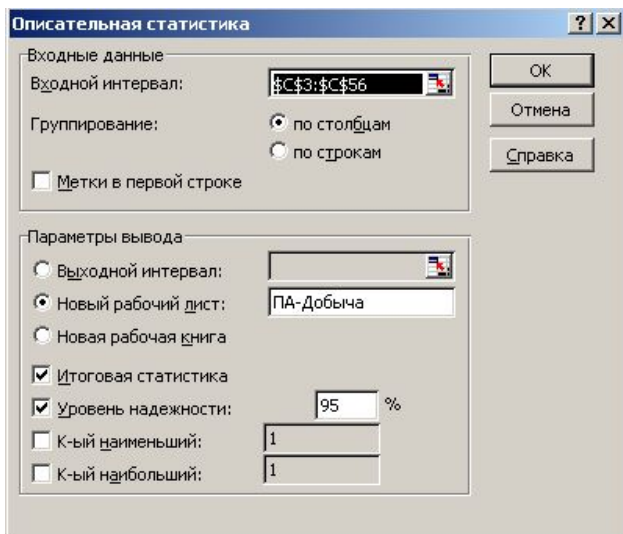
Что делать, если группы нет смотри в конце презентации - [Дополнение](#). Если нет блока «Анализ» на вкладке «Данные»

2. Щелкнуть мышкой на кнопке «Анализ данных».

В диалоговом окне выбрать «Описательная статистика»



3. В новом диалоговом окне указать:



- входной интервал
- вывод результатов на новый рабочий лист, лист назвать: **ОпСтат**
- отметить галочками **Итоговая статистика** и **-Уровень надежности**, для последнего указать 95% (95% вычисляет уровень надежности среднего со значимостью 0.05)
- К-наибольший = 2
- К-наименьший = 2

Задача 4

Вычислить набор характеристик, полученных автоматически с помощью «Анализ данных-Описательная статистика», с помощью встроенных функций Excel. Сравнить полученные результаты

Подсказка

1. Работать на листе **ОпСтат** и создав новый столбец **Расчет**, где вычислять с помощью встроенных функций Excel

Использовать функции: <https://support.office.com/ru-ru/article/Функции-Excel-по-категориям-5f91f4e9-7b42-46d2-9bd1-63f26a86c0eb> (см. следующую страницу)

Среднее	СРЗНАЧ
Стандартная ошибка	СТАНДОТКЛОН/КОРЕНЬ(СЧЕТ)
Медиана	МЕДИАНА
Мода	МОДА
Стандартное отклонение	СТАНДОТКЛОН
Дисперсия выборки	ДИСП
Эксцесс	ЭКСЦЕСС
Асимметричность	СКОС
Интервал	МАКС-МИН
Минимум	МИН
Максимум	МАКС
Сумма	СУММ
Счет	СЧЕТ
Наибольший(2)	НАИБОЛЬШИЙ
Наименьший(2)	НАИМЕНЬШИЙ
Уровень надежности(95.0%)	СТЮДРАСПОБР(0.05;СЧЕТ-1)/КОРЕНЬ(СЧЕТ)*СТАНДОТКЛОН

Задача 5

Вычислить набор характеристик, полученных автоматически с помощью «Анализ данных-Описательная статистика», с помощью исходных формул. Сравнить полученные результаты

Подсказка

1. Вычислять следующие характеристики:

а) Стандартное отклонение (выборки) σ_B

$$\sigma_B = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

б) Стандартная ошибка - среднеквадратическое отклонение выборочного среднего σ

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n(n-1)}}$$

в) Дисперсия выборки σ_B^2

$$\sigma_B^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

г) Коэффициент вариации

$$V = \frac{\sigma_B}{\bar{x}}$$

Квартили

Квартили — это показатели, которые чаще всего используются для оценки распределения данных при описании свойств больших числовых выборок.

Медиана разделяет упорядоченный массив пополам (50% элементов массива меньше медианы и 50% — больше), квартили разбивают упорядоченный набор данных на четыре части.

Величины Q_1 , медиана и Q_3 являются 25-м, 50-м и 75-м перцентилем соответственно.

Первый квартиль Q_1 — это число, разделяющее выборку на две части: 25% элементов меньше, а 75% — больше первого квартиля.

Третий квартиль Q_3 — это число, разделяющее выборку также на две части: 75% элементов меньше, а 25% — больше третьего квартиля.

Нулевой квартиль Q_0 - минимальное значение выборки

Четвертый квартиль Q_4 - максимальное значение выборки

Задача 6

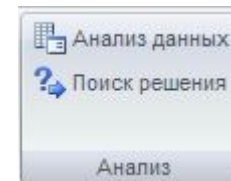
Найти квартили Q_1 , Q_2 , Q_3

Подсказка

- 1) Создать новый лист **Квартили** и результаты представлять на нем
- 2) Использовать функцию КВАРТИЛЬ или КВАРТИЛЬ.ВКЛ

Дополнение: если нет блока «Анализ» на вкладке «Данные»

<http://www.itshop.ru/Gde-mozhno-nayti-Paket-analiza-dlya-Excel/19i36611>




Загрузка и активация пакета анализа в Microsoft Excel 2013

1. Откройте вкладку Файл, нажмите кнопку Параметры и выберите категорию Настройки.
 2. В раскрывающемся списке Управление выберите пункт Настройки Excel и нажмите кнопку Перейти.
 3. В окне Настройки установите флажок Пакет анализа, а затем нажмите кнопку ОК.
- Если Пакет анализа отсутствует в списке поля Доступные надстройки, нажмите кнопку Обзор, чтобы выполнить поиск.
 - Если выводится сообщение о том, что пакет анализа не установлен на компьютере, нажмите кнопку Да, чтобы установить его.

Загрузка пакета анализа Microsoft Excel 2010

1. Откройте вкладку Файл и выберите пункт Параметры.
 2. Выберите команду Настройки, а затем в поле Управление выберите пункт Настройки Excel.
 3. Нажмите кнопку Перейти.
 4. В окне Доступные надстройки установите флажок Пакет анализа, а затем нажмите кнопку ОК.
- Если надстройка Пакет анализа отсутствует в списке поля Доступные надстройки, нажмите кнопку Обзор, чтобы найти ее.
 - В случае появления сообщения о том, что пакет анализа не установлен на компьютере, нажмите кнопку Да для его установки.

Загрузка и активация пакета анализа в Microsoft Excel 2007

1. Нажмите кнопку Microsoft Office  затем щелкните Параметры Excel.
 2. Выберите команду Настройки и в окне Управление выберите пункт Настройки Excel.
 3. Нажмите кнопку Перейти.
 4. В окне Доступные надстройки установите флажок Пакет анализа, а затем нажмите кнопку ОК.
- Если Пакет анализа отсутствует в списке поля Доступные надстройки, то для проведения поиска нажмите кнопку Обзор.
 - 5. При появлении сообщения о том, что пакет не установлен на компьютере и предложения установить его, нажмите кнопку Да.