

Национальный исследовательский университет «Высшая школа экономики»
Факультет компьютерных наук
Образовательная программа: Прикладная математика и информатика

Отчет по Домашнему заданию №1
«Пространственные данные»
по майнору «Прикладной статистический анализ»

«Статистический анализ численности работников государственных органов и
органов местного самоуправления»

Работу выполнила
студентка 2 курса
Ульянова Александра Игоревна
Преподаватель:
Грачева Светлана Сергеевна

Москва, 2022 г.

Содержание

Введение.....	3
Построение интервального вариационного ряда.....	4
Визуальное представление данных.....	7
Характеристики центра распределения.....	9
Характеристики вариации.....	10
Показатели формы распределения.....	13
Показатели динамики.....	14
Заключение.....	18
Список литературы.....	19
Приложения.....	20

Введение

Для анализа мной был выбран такой показатель, как “Численность сотрудников государственных органов и органов местного самоуправления на 2020 г. (чел.)”. Массив исходных данных изначально представлял собой 85 дискретных значений для каждого из регионов Российской Федерации, но позже в ходе вычислений был уменьшен до 75 значений (см. Построение интервального вариационного ряда).

Актуальность темы исследования вызвана тем, что в настоящее время многие граждане РФ недовольны быстротой и качеством услуг, оказываемых государственными органами. Изучение численности сотрудников государственных органов и органов местного самоуправления может помочь в выявлении проблем и различных трендов, а в дальнейшем возможно в оптимизации их работы.

Данные были взяты с сайта Федеральной службы государственной статистики. Выбор именно этого ресурса был связан с тем, что он предоставляет официальные и самые новые данные, на которые опираются органы власти, научное сообщество и СМИ. Также Росстат собирает информацию, которой зачастую нет в других государственных базах данных или в административных источниках.

Целью работы являются: изучение массива исходных данных, подсчет различных статистических показателей, выявление регионов РФ с аномально большим и аномально маленьким количеством сотрудников государственных органов.

Задачами работы являются:

- Поиск и сбор необходимых данных
- Обработка данных
- Выявление выбросов в данных
- Построение интервального ряда по исходным данным
- Подсчет статистических показателей для дискретного и интервального вариационных рядов
- Подсчет показателей изменения уровней ряда динамики
- Подведение итогов

Построение интервального вариационного ряда

Нахождение выбросов

Для построения интервального вариационного ряда в первую очередь потребовалось определить выбросы. Число сотрудников государственных органов в Москве, Московской области, Краснодарском крае и Тюменской области оказалось сильно больше чем в других регионах, поэтому эти регионы были интерпретированы мной как выбросы и не учитывались в следующих расчетах.

Объяснить эти выбросы можно большим числом жителей и, как следствие, необходимости в большом количестве сотрудников (для Москвы, МО и Краснодарского края). Тюменская же область является регионом, где расположено большое количество нефтедобывающих компаний, и необходимо много государственных служащих для организации и контроля их работы.

В Ненецком и Чукотском автономном округе значения исследуемого признака наоборот оказались сильно меньше, чем в других субъектах РФ из-за малонаселенности этих регионов. Они так же рассматривались как выбросы.

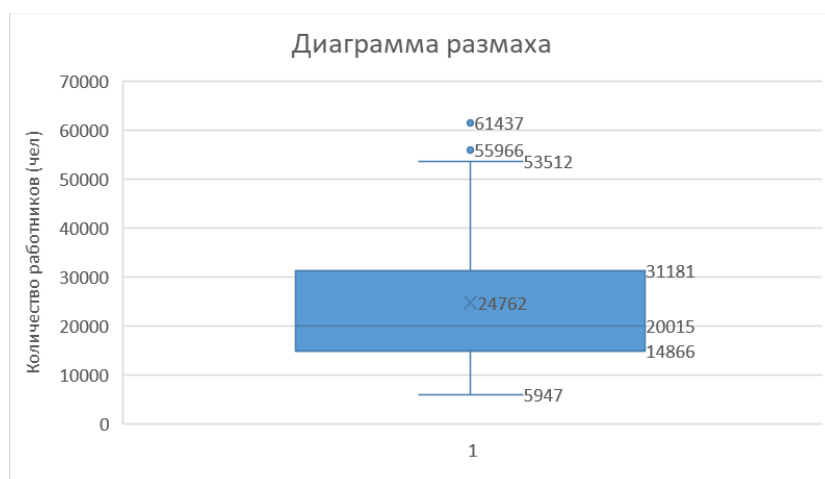


Рис 1. Диаграмма Размаха

После построения диаграммы размаха (Рис.1) выбросами также были названы регионы со значением признака большим 53512 чел., т.е. Свердловская область (61793 чел.), г. Санкт-Петербург (61781 чел.), Ростовская область (61437 чел.) и Красноярский Край (55966 чел.). Значение признака в этих регионах сильно отличается от остальных так как они густонаселены (в случае Санкт-Петербурга и Ростовской области), или являются крупными центрами промышленности (в случае Свердловской области и Красноярского края).

Стоит отметить, что не все расчеты были изначально проведены мной с учетом всех вышеперечисленных выбросов. Некоторые из них были обнаружены лишь в ходе более тщательного изучения исходных данных. Однако, чтобы избежать дезинформации читателя, в данном отчете представлены лишь финальные результаты.

Подсчет количества интервалов группировки

Число интервалов группировки было посчитано мной по формуле Стерджесса.

$$n = 1 + \log(N) \Rightarrow n = 1 + \log_2(75) = 1 + 6.22 = 7.22$$

После округления вниз получилось, что необходимо разбить данные на 7 интервалов.

Другие необходимые для построения интервального ряда величины были также посчитаны мной и представлены на Рис. 2. Максимальное и минимальное значения признака были найдены вручную, размах вариации был вычислен как разность между двумя вышеуказанными величинами, количество регионов соответствует размеру массива исходных данных после удаления выбросов, а шаг группировки найден как частное размаха вариации и интервала группировки.

<i>Необходимые величины</i>		
Максимальное значение признака	X_max	53512
Минимальное значение признака	X_min	5947
Размах вариации	R	47565
Кол-во регионов	N	75
Интервал группировки	n	7
Шаг группировки	h	6795

Рис. 2 Величины необходимые для расчета интервального ряда

Далее мной были вычислены нижние и верхние границы каждого из интервалов. Нижней границей первого из них стало число 5947 (т.е. минимальное значение признака), нижняя граница каждого из следующих интервалов становилась на 6795 больше нижней границы предыдущего. Верхней же границей каждого интервала становилась нижняя граница следующего интервала или, в случае последнего интервала, 53514 (т.е. максимальное значение признака).

Для каждого из интервалов была также посчитана середина (среднее арифметическое между нижней и верхней границей), а также частота и накопленная частота. Под частотой понимается количество значений из исходного дискретного ряда, которые попадают в данный интервал. Накопленная же частота говорит о количестве значений исходного ряда, которые не превышают верхнюю границу данного интервала.

Таким образом исходный дискретный ряд был преобразован в интервальный вариационный ряд с 7 равными интервалами. Получившийся интервальный ряд представлен в Приложении 2.

Визуальное представление данных

Для того, чтобы визуализировать исходный дискретный ряд, а также полученный вариационный ряд, мною были построены несколько диаграмм.

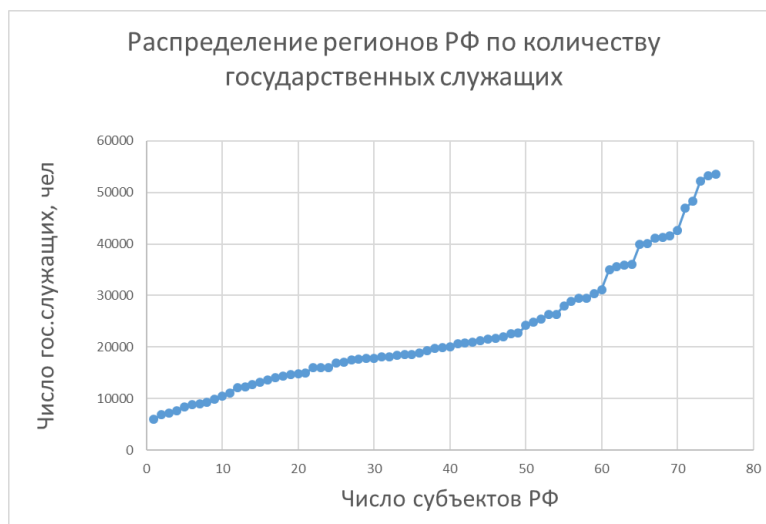


Рис. 3 Распределение регионов РФ по числу государственных служащих

На рисунке 3 представлена точечная диаграмма по исходным данным. На ней можно увидеть взаимосвязь между количеством субъектов в РФ и числом госслужащих, работающих в каждом из них. Отсутствие резких скачков в левом и правом концах графика говорит о том, что мы правильно определили и исключили выбросы в данных на предыдущем этапе.

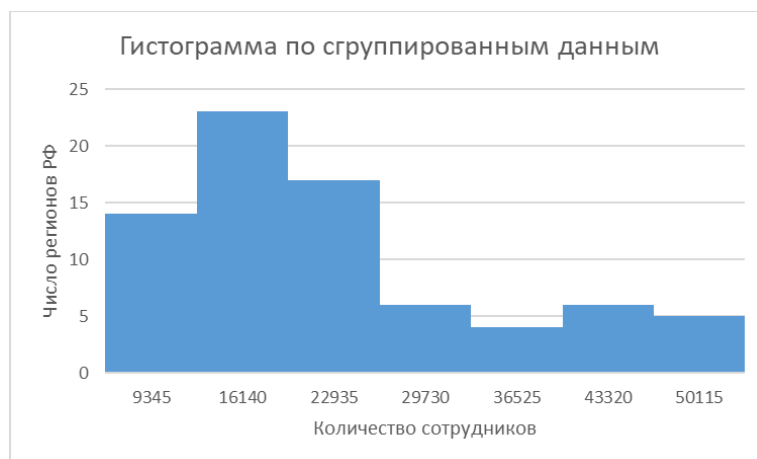


Рис.4 Гистограмма по сгруппированным данным

Данная гистограмма (Рис. 4) представляет в виде столбчатой диаграммы число элементов общей выборки в каждом отдельно взятом интервале. Отсутствие на ней пустых и малочисленных интервалов говорит о том, что на предыдущем этапе были правильно определены и удалены из данных выбросы.

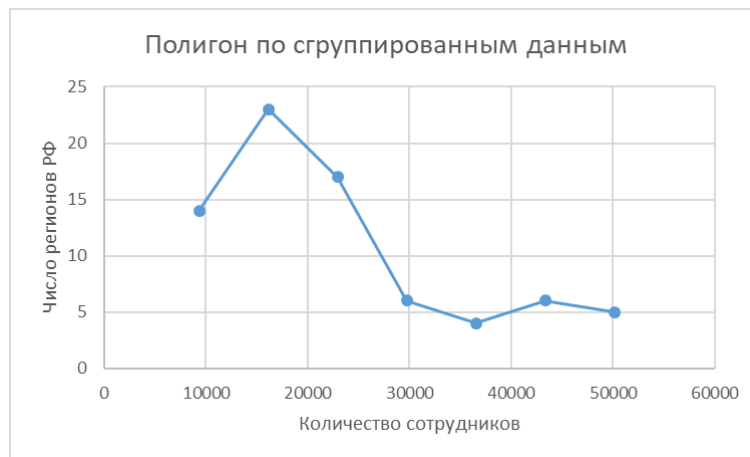


Рис.5 Полигон по сгруппированным данным

Полигон (Рис. 5) представляет из себя графическое изображение вариационного ряда в виде соединяющей точки ломаной. По оси абсцисс отмечены точки, обозначающие средние значения каждого интервала, по оси ординат – частоту каждого интервала.

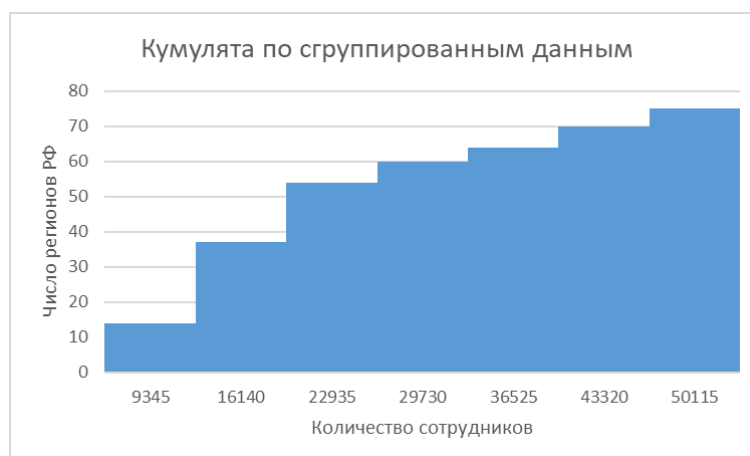


Рис.6 Кумулята по сгруппированным данным

Кумулята (Рис. 6) частот интервального ряда распределения графически изображает в виде столбцов накопленные частоты, относящиеся к каждому конкретному значению. Она показывает, какое количество значений признака из первоначальной выборки не превышают заданного значения.

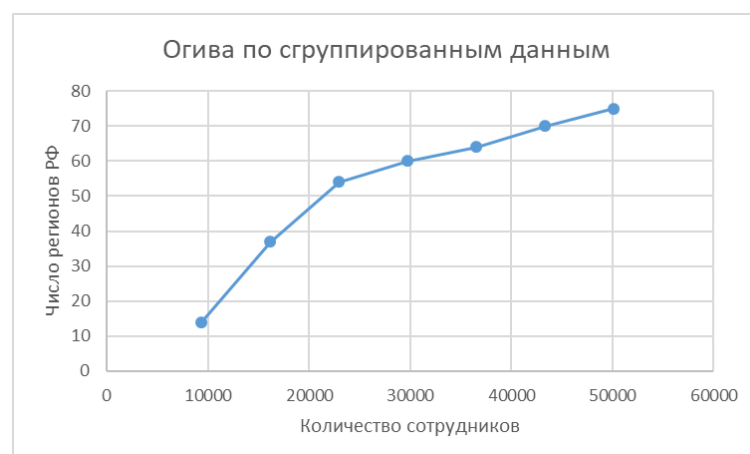


Рис.7 Огива по сгруппированным данным

Огиба отображает ту же информацию, что и кумулята, однако в виде ломанной линии, а не столбчатой диаграммы.

Характеристики центра распределения

Для исходного дискретного ряда, а также для интервального ряда мной были посчитаны такие характеристики центра распределения, как среднее, медиана и мода. Расчет производился по формулам, представленным в таблице 1. Некоторые характеристики были посчитаны двумя разными способами, чтобы удостовериться в корректности вычислений.

Таблица 1. Формулы, использованные для расчета характеристик центра распределения

Характеристика		По исходным данным		По интервальному ряду
		по формуле	с помощью встроенных функций Excel	
1	Среднее арифметическое	$\bar{x} = \frac{1}{n} \sum x_i$	=СРЗНАЧ([x _i])	$\bar{x} = \frac{1}{n} \sum x_i m_i$
2	Медиана	$\hat{M}_e = \frac{x_{n/2} + x_{(n/2+1)}}{2}$	=МЕДИАНА([x _i])	$\hat{M}_e = a_{Me} + h \cdot \frac{\frac{n}{2} - m_{Me-1}^H}{m_{Me}}$
3	Мода		=МОДА([x _i])	$\hat{M}_o = a_{Mo} + h \cdot \frac{m_{Mo} - m_{Mo-1}}{2m_{Mo} - m_{Mo-1} - m_{Mo+1}}$

Полученные в результате расчетов значения приведены на рисунках 8 и 9.

Средние	22870
Медиана	19788
Мода	#Н/Д

Рис. 8 Характеристики центра распределения для дискретного ряда

Средние	23026
Медиана	19737
Мода	16819

Рис.9 Характеристики центра распределения для интервального ряда

Из Рис.8 и Рис.9 видно, что значения среднего и медианы между дискретным и интервальным рядами немного отличаются друг от друга. Это объясняется погрешностями, возникающими при округлении значений и на этапе преобразования дискретного ряда в интервальный.

Также можно заметить, что среднее значение для дискретного ряда несколько больше, чем значение медианы. Причиной этому стало наличие в наших данных небольшого количества регионов с большим значением признака. Они влияют на рост среднего арифметического. Аналогичная ситуация наблюдается и на Рис. 9 для интервального ряда.

Также из Рис.8 видно, что в дискретном ряду отсутствует мода. Это связано с тем, что единицей измерения нашего признака является количество человек. Значения признака достаточно большие и выходит, что ни в одном из регионов значения признака не совпадают в точности. Однако в вариационном ряду значение медианы уже вычисляется в связи с особенностью формулы, по которой производится расчет.

Характеристики вариации

Для исходного дискретного ряда, а также для интервального ряда мной были посчитаны такие характеристики вариации: дисперсия, среднеквадратичное отклонение, среднее линейное отклонение, коэффициент осцилляции, коэффициент вариации, коэффициент линейной вариации, квартили, интерквартильный размах. Расчет производился по формулам, приведенным в таблице 2. Некоторые характеристики были посчитаны двумя разными способами, чтобы удостовериться в корректности вычислений.

Таблица 2. Формулы, использованные для подсчета характеристик вариации

Характеристика		По исходным данным		По интервальному ряду
		по формуле	с помощью встроенных функций Excel	
1	Дисперсия Мера рассеивания относительно среднего значения	$S^2 = \frac{1}{n} \sum (x_i - \bar{x})^2$	=ДИСПР([x _i])	$S^2 = \frac{1}{n} \sum (x_i - \bar{x})^2 \cdot m_i$
2	Среднее линейное отклонение	$\bar{d} = \frac{\sum x_i - \bar{x} }{n}$	=СРОТКЛ()	$\bar{d} = \frac{\sum x_i - \bar{x} m_i}{\sum m_i}$
3	Среднее квадратическое отклонение	$S = \sqrt{S^2}$		$S = \sqrt{S^2}$
4	Коэффициент осцилляции (%)	$V_R = \frac{R}{\bar{x}} 100$		$V_R = \frac{R}{\bar{x}} 100$
5	Коэффициент линейной вариации (%)	$V_{\bar{d}} = \frac{\bar{d}}{\bar{x}} 100$		$V_{\bar{d}} = \frac{\bar{d}}{\bar{x}} 100$

6	Коэффициент вариации (%)	$V = \frac{S}{x} * 100$		$V = \frac{S}{x} * 100$
7	Квартили	$Q_k = \frac{k(n+1)}{4}$		$Q_1 = a_{Q_1} + h \frac{\frac{1}{4} \sum m_i - m_{Q_1-1}^H}{m_{Q_1}}$
8	Интерквартильный размах	$IQR = Q3 - Q1$		$IQR = Q3 - Q1$

Полученные в результате расчетов значения представлены ниже на Рис. 10

На Рис. 10 можно увидеть, что значения характеристик вариации почти не отличаются для дискретного и интервального рядов. Это может говорить об аккуратности и правильности расчетов.

Большое значение дисперсии и среднеквадратичного отклонения означает, что значение признака сильно разбросано относительно своего математического ожидания.

Коэффициент осцилляции — характеризует колеблемость крайних значений признака вокруг средней арифметической. Коэффициент осцилляции получился очень большой, значит крайние значения сильно отклоняются от среднего. Коэффициент вариации показывает степень изменчивости по отношению к среднему показателю выборки. В данном случае он получился больше 33%, что говорит о неоднородности данных.

Для интервального ряда значения первых двух квартилей немного увеличились, а значение третьего квартиля немного уменьшилось. Это связано с тем, что, когда мы из дискретного ряда делаем интервальный, мы становимся ближе к нормальному распределению.

Показатель	Расчеты по ряду исходных данных	Расчеты по ряду интервальных данных
<i>Дисперсия</i>	140488003.8	144665942.8
<i>Среднеквадратичное отклонение</i>	11852.8	12027.7
<i>Среднее линейное отклонение</i>	9419.6	9371.9
<i>Коэффициент осцилляции (%)</i>	208.0	206.6
<i>Коэффициент вариации (%)</i>	51.8	52.2
<i>Коэффициент линейной вариации (%)</i>	41.2	40.7
Q1	14690	14145.3
Q2	19788	19736.9
Q3	29430	28880.1
IQR	14740	14734.8

Рис.10 Характеристики вариации для дискретного и интервального рядов

Проверка эмпирических соотношений

С помощью посчитанных характеристик вариации можно сделать некоторые выводы о нормальности распределения.

1.

$$\begin{aligned}
 X_{\text{ср}} &= 23026 \\
 S &= 12027 \\
 12027 * (-3) &\leq 23026 \leq 12027 * 3
 \end{aligned}$$

Неравенство верное, а значит наше распределение похоже на нормальное.

2.

$$\begin{aligned}
 S &= 12027 \\
 R &= 47567 \\
 R/6 &= 7927 \\
 \frac{R}{6} &\sim S
 \end{aligned}$$

S и R/6 достаточно близки друг к другу, но все же отличаются. Значит распределение похоже на нормальное, но все же совсем нормальным его назвать нельзя.

3.

$$\begin{aligned} S &= 12027 \\ d &= 9371 \\ 1.25d &= 11713 \\ 1.25d &\sim S \end{aligned}$$

Соотношение верное, а значит распределение наших данных близко к нормальному распределению.

Показатели формы распределения

Для оценки симметричности распределения признака мной были посчитаны коэффициент асимметрии и коэффициент эксцесса. Полученные значения представлены на Рис. 11

<i>Коэффициент асимметрии</i>	0.95
<i>Коэффициент эксцесса</i>	0.20

Рис. 11 Показатели формы распределения

Видим, что коэффициент асимметрии больше 0, значит наблюдается правосторонняя асимметрия. Более того, значение этого показателя больше 0.7, а значит асимметрия существенная.

Проверить правильность вычисления можно также сравнив среднее, моду и медиану. Так как мода для дискретного ряда у нас не определена, то будем сравнивать только медиану и среднее.

$$\bar{x} = 22870, M_e = 19788 \Rightarrow M_e < \bar{x}$$

Действительно, наблюдается правосторонняя асимметрия. Правильность данных вычислений можно также проверить визуально, взглянув на график, представленный на Рис. 4. Столбцы с самой большой частотой располагаются на нем левее центра, что и соответствует правосторонней асимметрии.

Коэффициент эксцесса положительный, из чего следует, что график распределения имеет выраженный пик. Визуальное подтверждение этому можно найти на Рис. 4. Также присутствие положительного эксцесса в изучаемом распределении означает, что в совокупности есть сформировавшееся ядро, т.е. большинство значений показателя находятся вокруг среднего.

Показатели динамики

В качестве динамического показателя я решила взять тот же признак (численность сотрудников государственных органов и органов местного самоуправления (чел)), только для города Москвы за 2000 – 2020 года. Этот регион являлся выбросом при прошлых расчетах, поэтому изучение показателей для него представляет особый интерес. Как и вся информация, использованная мною ранее, эти данные взяты с сайта Федеральной службы государственной статистики и представлены в первых двух столбцах таблицы из Приложения 3.

Графическое представление данных можно увидеть ниже на Рис. 12 и Рис. 13

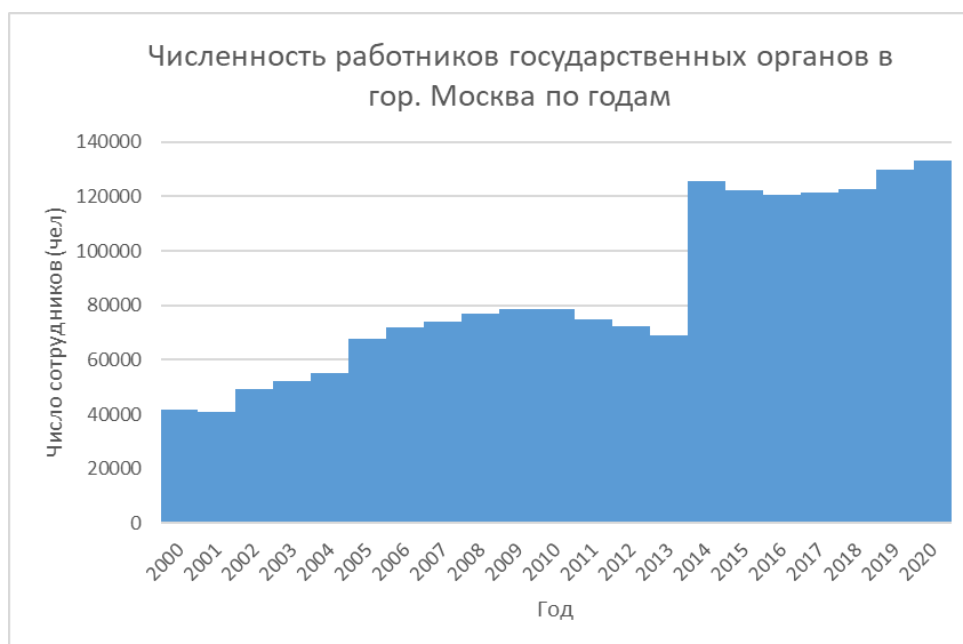


Рис. 12 Гистограмма численности сотрудников государственных органов в Москве по годам

Данная гистограмма (Рис. 12) представляет в виде столбчатой диаграммы число сотрудников государственных органов в городе Москва в каждый год с 2000 по 2020. На горизонтальной оси указан номер года, а на вертикальной – число сотрудников, человек. Из данной диаграммы видно, что значение признака увеличивалось почти каждый год (за исключением 2001, 2012-2014, 2016-17 годов). Наиболее резкий рост признака наблюдается в 2015 году.



Рис. 13 Полигон численности сотрудников государственных органов в Москве по годам

На данном полигоне (Рис.13) представлены те же данные, что и на гистограмме выше, однако в виде ломаной прямой, а не столбчатой диаграммы.

Для данного показателя были посчитаны такие показатели динамики, как абсолютные приросты (цепные и базисные), темпы роста (цепные и базисные), а также темпы прироста (цепные и базисные). Полученные значения находятся в Приложении 3.

Абсолютный прирост выражает абсолютную скорость изменения ряда динамики и определяется как разность между данным уровнем и уровнем, принятым за базу сравнения. Для цепных абсолютных приростов базой сравнения считается предыдущий период, а для базисных – первый.

Темп роста K_i определяется как отношение данного уровня к предыдущему или базисному, показывает относительную скорость изменения ряда. Если абсолютные уровни исследуемого явления снижаются, то темп роста меньше 100%, иначе – больше. Однако он никогда не может быть отрицательным числом.

Темп прироста показывает на сколько процентов целого увеличился или уменьшился соответствующий уровень ряда динамики по сравнению с каким-либо достигнутым уровнем. Темп прироста вычисляется путём вычитания из темпа роста 100 процентов (так как темп роста выражен в процентах). В отличие от темпов роста, темпы прироста могут быть и отрицательными числами. В этом случае они показывают, на какую часть целого или на сколько процентов снизился уровень исследуемого явления.

Из данных, приведенных в приложении 3, можно сделать вывод, что значение признака увеличилось больше чем в 3 раза за период наблюдений (т.к. базисный темп роста в 2020 году составил 319.8%). Значения признака уменьшались в 2001, 2012-2014, 2016-17 годах. Мы уже визуальнo определили это по диаграмме, но теперь мы можем обнаружить это посмотрев на знак числа, стоящего в столбце “базисный прирост”. Отрицательное значение говорит о уменьшении значения признака по сравнению с предыдущим периодом, а положительное – о росте.

Можно также сказать, что ни разу после 2001 года значение признака не опускалось ниже значения в 2000 году (начало периода наблюдений). Об этом говорит положительность значений столбца “базисный прирост”.

Чтобы сделать прогноз на 3 уровня вперед для данных динамических данных потребовалось вычислить средний абсолютный прирост и средний темп роста. Средний абсолютный прирост – это показатель, характеризующий, на сколько в среднем изменилась величина изучаемого показателя. Средний темп роста – это показатель, характеризующий, сколько процентов в среднем составило изменение изучаемого показателя.

Средний абсолютный прирост:

$$\Delta = \frac{y_n - y_1}{n - 1},$$

где y_1 – первый уровень ряда,
 n – число уровней ряда,
 y_n – конечный уровень ряда.

Средний темп роста:

$$T = \sqrt[n-1]{\frac{T_{\text{ср}}}{T_1}} * 100\%$$

где $T_{\text{ср}}$ – среднее значение показателя,
 T_1 – начальное значение показателя,
 n – число периодов расчета

Были получены значения 4576 для среднего абсолютного прироста и 106 для темпа роста. Благодаря данным показателям можно сделать прогноз значений признака на несколько периодов вперед. В данном случае был построен прогноз на 3 года вперед, так как прогноз на большее количество периодов уже сложно назвать четким.

На основе среднего абсолютного прироста прогноз делается по следующему принципу: берется значение исследуемого признака для предыдущего периода и к нему прибавляется посчитанная ранее величина среднего абсолютного прироста. Для среднего темпа роста значение предыдущего периода умножается на средний темп роста, деленный на 100.

Полученные в ходе вычислений значения представлены на Рис. 14 и Рис. 15

2021 (прогноз на основе среднего абсолютного прироста)	137727
2022 (прогноз на основе среднего абсолютного прироста)	142303
2023 (прогноз на основе среднего абсолютного прироста)	146879

Рис. 14 Прогноз на основе среднего абсолютного прироста

2021 (прогноз на основе среднего темпа роста)	141120
2022 (прогноз на основе среднего темпа роста)	149566
2023 (прогноз на основе среднего темпа роста)	158518

Рис. 15 Прогноз на основе среднего темпа роста

Для полученных прогнозных значений были построены два полигона (Рис. 16 и Рис.17).



Рис. 16 График прогнозных значений на основе среднего абсолютного прироста

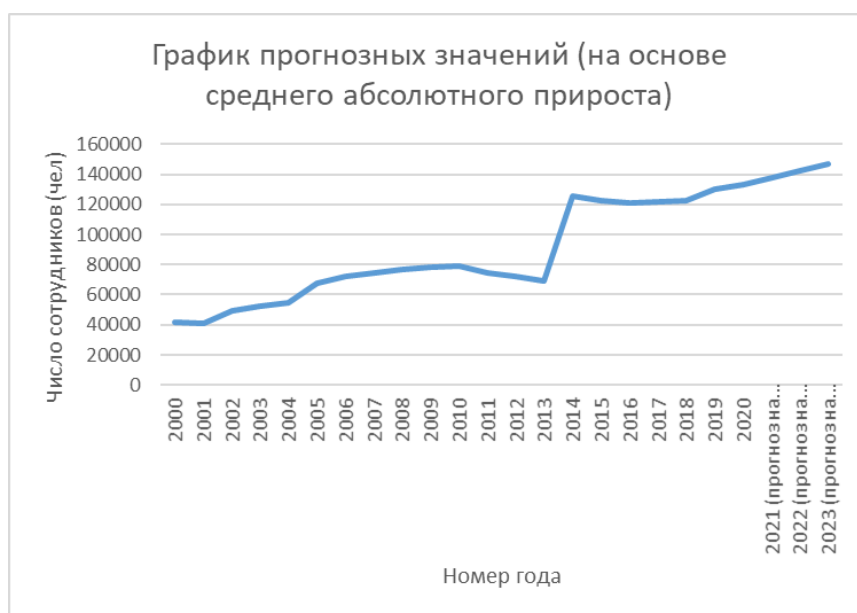


Рис. 17 График прогнозных значений на основе среднего темпа роста

Как видно по этим графикам, значения для 2021, 2022 и 2023 года получились похожие и в целом соответствуют общей тенденции роста значения показателя с течением времени.

Заключение

В данной работе были посчитаны и проанализированы различные статистические показатели для численности сотрудников государственных органов и органов местного самоуправления по регионам РФ за 2020 год. Был построен интервальный вариационный ряд, данные были визуализированы, были посчитаны характеристики центра положения, коэффициенты асимметрии и эксцесса, квартили, характеристики вариации. Для динамических данных, а именно для показателя “численность сотрудников государственных органов и органов местного самоуправления в г. Москве за 2000 – 2022 года” были построены графики, посчитаны показатели динамики, сделаны прогнозы на несколько следующих периодов.

Выводы по каждому отдельному шагу анализа указаны после вычисления каждого показателя в ходе работы. Если говорить о выводах в целом, из полученных результатов можно судить о том, что распределение количества работников государственных органов по регионам очень неравномерное. Встречаются как регионы с очень маленьким их количеством, так и наоборот, с очень большим. Распределение близко к нормальному (если не учитывать выбросы), однако наблюдается правосторонняя асимметрия. График распределения имеет выраженный пик.

В большую сторону выделяются значения признака в густонаселенных и в крупных промышленных регионах. В этих субъектах возможен переизбыток сотрудников государственных органов. В малонаселенных северных регионах наоборот наблюдается чрезмерно маленькое значение исследуемого признака. Это может приводить к медленной или недостаточно качественной работе городских администраций, муниципалитетов, судов и д.р.

Что касается динамических данных, то в Москве в течении последних 20 лет быстро растет число государственных служащих. Вероятно, это связано с ростом численности населения, а также с расширением списка услуг, которые оказываются местным жителям.

Список литературы

1. Статистика: учебник и практикум для академического бакалавриата / В. С. Мхитарян [и др.] ; под ред. В. С. Мхитаряна. — М. : Издательство Юрайт, 2018. — 250 с. — (Серия : Бакалавр. Академический курс). — ISBN 978-5-9916-5591-0.
2. Основы теории статистики : [учеб. пособие] / В. В. Полякова, Н. В. Шаброва ; М-во образования и науки Рос. Федерации, Урал. федер. ун-т. — 2-е изд., испр. и доп. — Екатеринбург : Изд-во Урал. ун-та, 2015. — 148 с. ISBN 978-5-7996-1520-8

Приложения

Приложение 1

Численность сотрудников государственных органов и органов местного самоуправления по регионам РФ на 2020 год, человек.

№ п/п	Название региона РФ	Численность сотрудников государственных органов, чел. 2020
1	Ненецкий автономный округ	1914
2	Чукотский автономный округ	2913
3	Еврейская автономная область	5947
4	Республика Алтай	6894
5	Магаданская область	7160
6	г.Севастополь	7586
7	Республика Адыгея	8320
8	Республика Калмыкия	8878
9	Республика Тыва	9033
10	Республика Ингушетия	9254
11	Карачаево-Черкесская Республика	9829
12	Камчатский край	10520
13	Республика Марий Эл	11080
14	Новгородская область	12174
15	Республика Хакасия	12288
16	Сахалинская область	12648
17	Республика Карелия	13117
18	Республика Мордовия	13650
19	Костромская область	14011
20	Псковская область	14315
21	Кабардино-Балкарская Республика	14690
22	Мурманская область	14866
23	Орловская область	15006
24	Ивановская область	15939
25	Ямало-Ненецкий автономный округ	15986
26	Республика Северная Осетия-Алания	16057
27	Чувашская Республика	16897
28	Республика Бурятия	17068

29	Амурская область	17435
30	Липецкая область	17644
31	Астраханская область	17765
32	Курганская область	17772
33	Тамбовская область	18095
34	Рязанская область	18107
35	Ульяновская область	18400
36	Калужская область	18482
37	Смоленская область	18504
38	Курская область	18825
39	Республика Коми	19262
40	Владимирская область	19788
41	Томская область	19924
42	Пензенская область	20015
43	Удмуртская Республика	20666
44	Республика Саха (Якутия)	20863
45	Ярославская область	20954
46	Архангельская область	21198
47	Вологодская область	21566
48	Тульская область	21656
49	Тверская область	21951
50	Брянская область	22597
51	Ленинградская область	22703
52	Белгородская область	24204
53	Чеченская Республика	24842
54	Калининградская область	25438
55	Кировская область	26295
56	Забайкальский край	26324
57	Хабаровский край	28025
58	Ханты-Мансийский автономный округ - Югра	28810
59	Воронежская область	29430
60	Республика Крым	29514
61	Оренбургская область	30315
62	Омская область	31181
63	Саратовская область	35027
64	Приморский край	35603
65	Волгоградская область	35851
66	Республика Дагестан	36082
67	Ставропольский край	39975
68	Кемеровская область	40039
69	Новосибирская область	41094
70	Пермский край	41341
71	Алтайский край	41537

72	Самарская область	42677
73	Иркутская область	46959
74	Республика Башкортостан	48351
75	Республика Татарстан	52168
76	Нижегородская область	53257
77	Челябинская область	53512
78	Красноярский край	55966
79	Ростовская область	61437
80	г.Санкт-Петербург	61781
81	Свердловская область	61793
82	Тюменская область	70961
83	Краснодарский край	74013
84	Московская область	83549
85	г.Москва	133151

Источник: Федеральная служба государственной статистики
(<https://rosstat.gov.ru>)

Приложение 2

Сгруппированный интервальный вариационный ряд распределения численности сотрудников государственных органов и органов местного самоуправления по регионам РФ на 2020 год, тыс.человек.
(без выбросов)

Порядковый номер интервала	Нижняя граница интервала	Верхняя граница интервала	Середина интервала	Частота	Накопленная частота
1	5947	12742	9344.5	14	14
2	12742	19537	16139.5	23	37
3	19537	26332	22934.5	17	54
4	26332	33127	29729.5	6	60
5	33127	39922	36524.5	4	64
6	39922	46717	43319.5	6	70
7	46717	53512	50114.5	5	75

Приложение 3

Значения динамического показателя (численности сотрудников государственных органов и органов местного самоуправления) с 2000 по 2020 года для г.Москва , а также посчитанные для него статистические показатели

<i>Номер года</i>	<i>Число сотрудников в г. Москва</i>	<i>Базисный прирост</i>	<i>Цепной прирост</i>	<i>Темп роста (базисный)</i>	<i>Темп роста (цепной)</i>	<i>Темп прироста (базисный)</i>	<i>Темп прироста (цепной)</i>
2000	41634	41634	41634	-	-	-	-
2001	41046	-588	-588	98.6	98.6	-1.4	-1.4
2002	49066	7432	8020	117.9	119.5	17.9	19.5
2003	52057	10423	2991	125.0	106.1	25.0	6.1
2004	55028	13394	2971	132.2	105.7	32.2	5.7
2005	67768	26134	12740	162.8	123.2	62.8	23.2
2006	71916	30282	4148	172.7	106.1	72.7	6.1
2007	74201	32567	2285	178.2	103.2	78.2	3.2
2008	76838	35204	2637	184.6	103.6	84.6	3.6
2009	78453	36819	1615	188.4	102.1	88.4	2.1
2010	78782	37148	329	189.2	100.4	89.2	0.4
2011	74651	33017	-4131	179.3	94.8	79.3	-5.2
2012	72339	30705	-2312	173.7	96.9	73.7	-3.1
2013	69097	27463	-3242	166.0	95.5	66.0	-4.5
2014	125509	83875	56412	301.5	181.6	201.5	81.6
2015	122217	80583	-3292	293.6	97.4	193.6	-2.6

2016	120754	79120	-1463	290.0	98.8	190.0	-1.2
2017	121501	79867	747	291.8	100.6	191.8	0.6
2018	122558	80924	1057	294.4	100.9	194.4	0.9
2019	129821	88187	7263	311.8	105.9	211.8	5.9
2020	133151	91517	3330	319.8	102.6	219.8	2.6

Источник: Федеральная служба государственной статистики (<https://rosstat.gov.ru>)