

Национальный исследовательский университет «Высшая школа экономики»
Факультет компьютерных наук
Образовательная программа: Программная инженерия

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

«Статистическое исследование количества зарегистрированных
преступлений в субъектах Российской Федерации»

Работу выполнил:
студент 2 курса
Артемьев Александр Вячеславович
Преподаватель:
Грачёва Светлана Сергеевна

Москва, 2023г.

Содержание

Оглавление

1. Введение.....	3
1.1 Актуальность темы исследования	3
1.2 Цели исследования	3
1.3 Задачи исследования	3
1.4 Гипотезы.....	3
2. Используемые показатели для анализа	4
3. Исследование выбранных данных на выбросы.....	4
4. Построение интервального вариационного ряда.....	7
5. Построение графиков распределения признака в вариационном ряду	8
6. Расчёт показателей центра распределения.....	11
6.1 Дискретный ряд	11
6.2 Интервальный ряд	11
7. Расчет показателей вариации	13
7.1 Дискретный ряд	13
7.2 Интервальный ряд	14
8. Показатели дифференциации.....	16
8.1 Дискретный ряд	16
8.2 Интервальный ряд	17
9. Анализ временных данных.....	19
9.1 Используемый показатель для анализа.....	19
9.2 Анализ на выбросы	19
9.3 Расчет показателей рядов динамики	21
9.3.1 Цепные показатели.....	21
9.3.2 Базисные показатели	21
9.3.3 Анализ показателей	22
9.3.4 Прогноз уровня преступности в Москве на 2023 год.....	23
9.4 Построение линии тренда.....	24
10. Общие выводы по работе.....	26
11. Список литературы	26
12. Используемые информационные источники	26
13. Приложения	26

1. Введение

1.1 Актуальность темы исследования

С течением времени изменяется уровень преступности и появляются новые виды правонарушений, что оказывает огромное влияние на различные сферы общества:

- Заболеваемость и смертность населения
- Экономические кризисы и спад производства
- Материальный и моральный вред, причиненный обществу
- Падение качества жизни населения

Анализ количества преступлений является актуальным и необходимым, ведь нахождение факторов, объясняющих целую структуру правонарушений, поможет снизить уровень преступности. Что напрямую окажет положительное влияние на все общество в целом.

1.2 Цели исследования

- Изучить в каких регионах страны зарегистрировано больше/меньше всего преступлений.
- Определить факторы, которые влияют на уровень преступности в регионах.
- Изучить распределение субъектов по уровню преступности.
- Изучить уровень преступности в Москве с 2009 по 2022 год, и спрогнозировать уровень преступности на 2023 год.

1.3 Задачи исследования

- Выбрать актуальные данные по уровню преступности в субъектах РФ
- Найти аномальные значения в выбранных данных
- Построить интервальный вариационный ряд
- Найти показатели центра распределения, вариации, дифференциации и проанализировать найденные значения как для дискретного, так и для интервального вариационного ряда.
- Найти данные по уровню преступности в Москве с 2009 по 2022 год и показатели рядов динамики для прогноза на 2023 год.

1.4 Гипотезы

- Наиболее высокий уровень преступности, будет достигаться в крупнейших регионах РФ, наименьший уровень соответственно в самых маленьких по численности регионах
- За исключением крупнейших городов, в исследуемых данных будет равномерное распределение регионов по размаху вариации.
- Уровень преступности в Москве, вероятно снизился по сравнению с 2009 годом, но устойчиво растет более 5 лет, из-за большого роста населения в Москве.

2. Используемые показатели для анализа

Выбранный показатель – уровень преступности.

- Показатель характеризует количество документально зарегистрированных преступлений в данном субъекте РФ.
- Единица измерения - количество преступлений.
- Количественные дискретные данные.

Было принято решение использовать именно этот показатель для исследования уровня преступности, а не количество преступлений относительно населения субъекта РФ. Во-первых, этот показатель иллюстрирует абсолютные данные, а не относительные. Во-вторых, в дальнейшем планируется исследование зависимости уровня преступности от населения субъекта РФ (один из факторов, который косвенно влияет на уровень преступности).

3. Исследование выбранных данных на выбросы

Данные по уровню преступности были взяты с официального сайта Генеральной прокуратуры Российской Федерации, и сформированы в таблицу (Приложение 1). Первым этапом исследования было построение точечной диаграммы и гистограммы для исходного (рис. 1 и рис. 2) и ранжированного по возрастанию рядов данных (рис. 3 и рис 4).

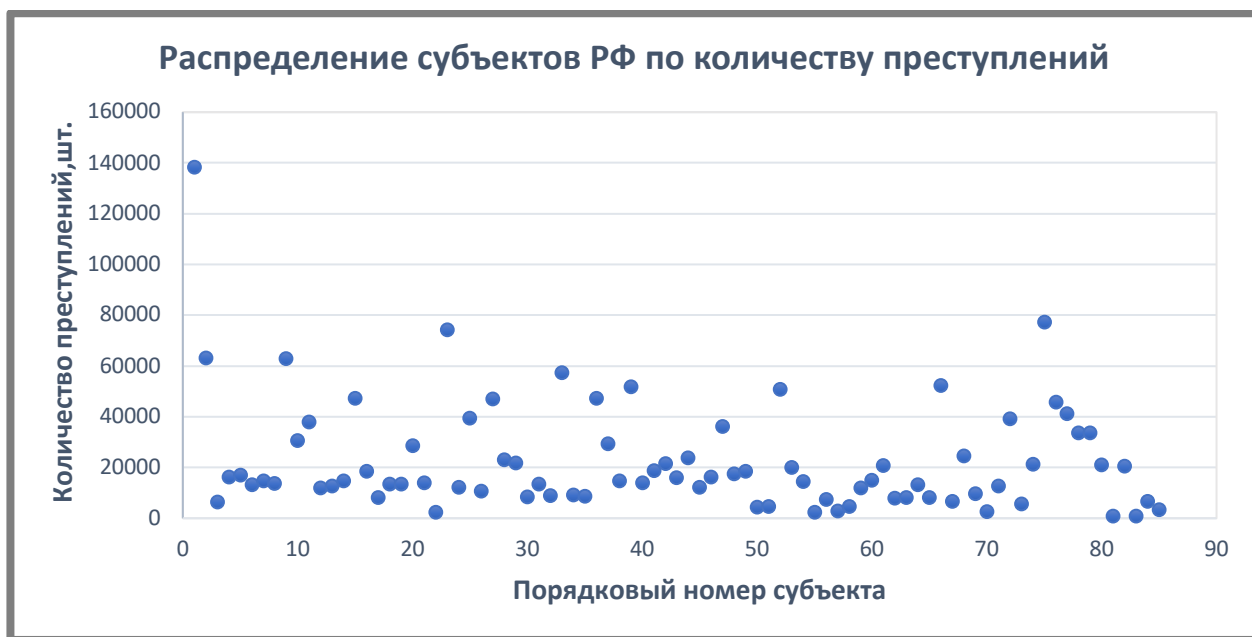


Рисунок 1. Точечная диаграмма распределения субъектов РФ по количеству преступлений



Рисунок 2. Гистограмма распределения субъектов РФ по количеству преступлений

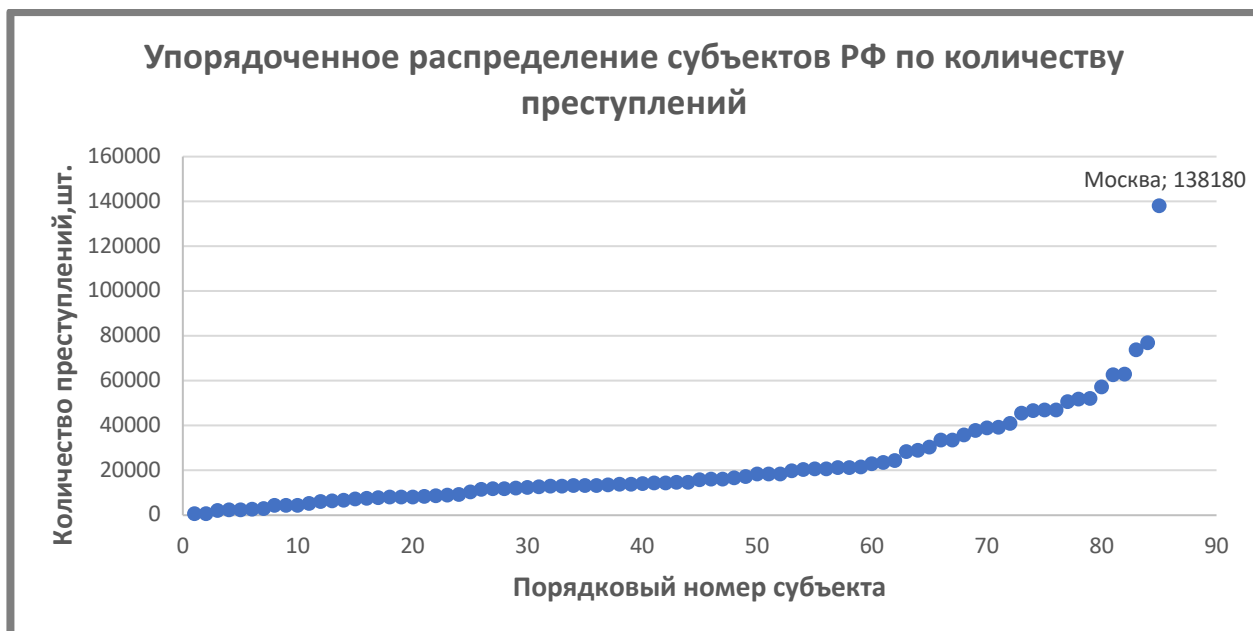


Рисунок 3. Точечная диаграмма субъектов РФ по количеству преступлений (данные ранжированы по возрастанию)

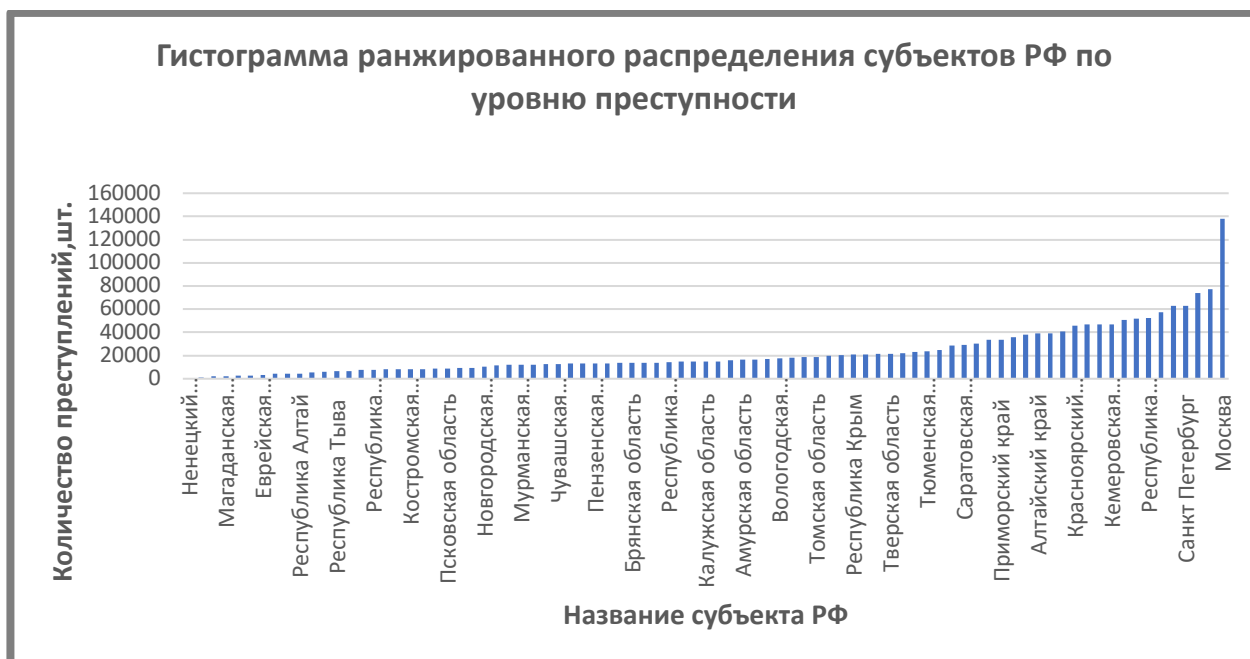


Рисунок 4. Гистограмма субъектов РФ по количеству преступлений (данные ранжированы по возрастанию)

Визуализируя исходные данные в виде диаграмм, становится легче их анализировать. Исследуя рис.1 и рис. 2, можно заметить, что в рассматриваемой выборке преобладают субъекты с уровнем преступности, не превышающим 40 000 единиц. Также на первом рисунке отчетливо видны выбросы, это 5 точек со значением показателя более 60 000. Эти же выбросы можно заметить на графиках, изображенных на 3 и 4 рисунках. Также из гистограммы (рис. 4), можно определить субъект с наименьшим количеством преступлений – Ненецкий автономный округ (686), и с наибольшим - Москва (138180). На данном этапе рано делать утверждения о всех выбросах, но то, что аномальные значения присутствуют становится очевидным. Поэтому следующим этапом было построение “Ящика с усами” (рис. 5)

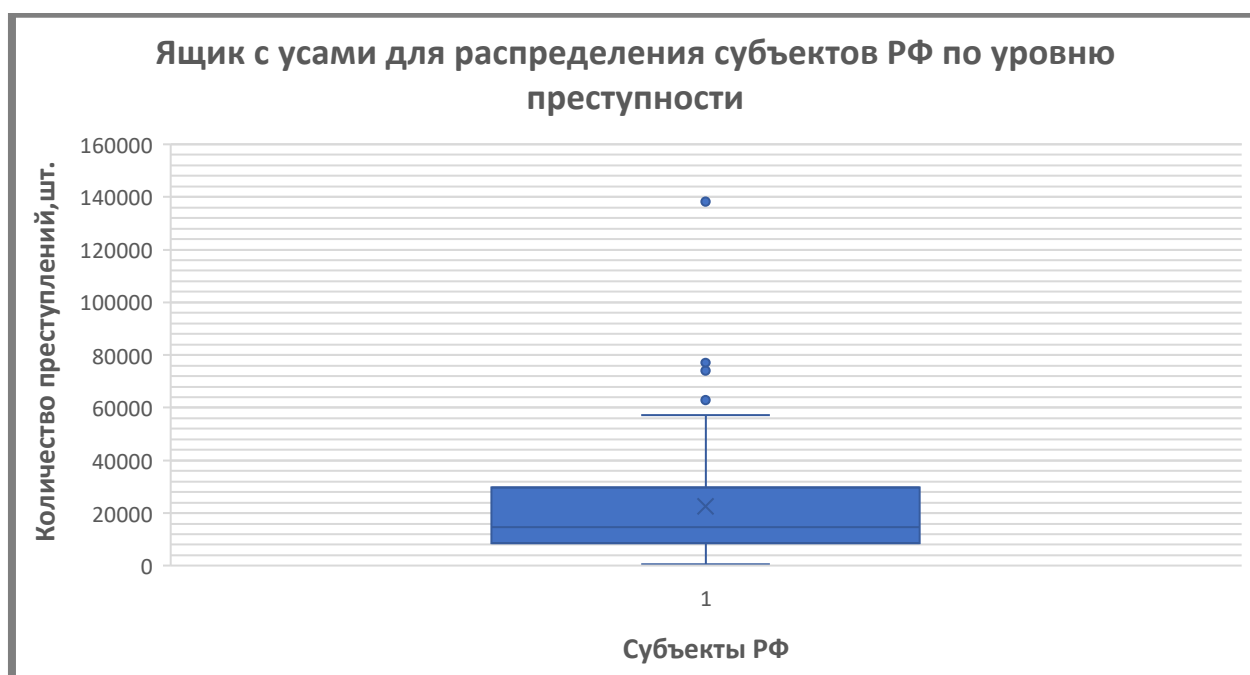


Рисунок 5. Диаграмма размаха, субъектов РФ по уровню преступности.

Теория о том, что есть выбросы подтвердилась, на диаграмме отмечены 5 точек которые не попадают в интерквартильный размах, а именно: Москва (138180), Краснодарский край (77125), Московская область (74054), Санкт Петербург (62971), Челябинская область (62844).

Далее будем исследовать новый набор данных (Приложение 2), исключая аномальные наблюдения, которые могут непредвиденным образом влиять на результаты исследования и исказить истинный результат.

Ниже представлена точечная диаграмма для новой выборки (рис. 6).



Рисунок 6. Точечная диаграмма субъектов РФ по количеству преступлений (данные ранжированы по возрастанию, исключая выбросы)

4. Построение интервального вариационного ряда

Для построение интервального вариационного ряда, необходимо найти все величины, отображаемые на рис. 7(Где указаны необходимые формулы для поиска величин).

С помощью построенной точечной диаграммы (рис. 6), можем определить минимальное и максимальное значения уровня преступности, которые равны соответственно 686 (Ненецкий автономный округ) и 57246(Ростовская область).

Тогда размах вариации = 56560(разница максимального и минимального значений).

Следующий этап состоял в нахождении количества интервалов разбиения. Было принято решение воспользоваться эмпирическим правилом Стёрджеса – определения оптимального количества интервалов, которое находится по формуле:

$$n = \lfloor 1 + \log_2 N \rfloor$$

Где n – количество интервалов, N – количество наблюдений. То есть $n = 7$. Следовательно ширина интервала $h = 8080$ (формула указана в рис. 7). Зная ширину интервалов и их количество, однозначно определяются границы интервалов (рис. 7, a_i – левая граница i интервала и b_i – правая граница i интервала). После определения всех границ, были

найденны середины интервалов, как среднее арифметическое границ интервала. Последним этап построения интервального ряда, состоял в нахождении частот, количества субъектов РФ, имеющие значение из рассматриваемого интервала, и накопленных частот. Результаты вычисления отображены ниже, в таблице 1.

Минимальное значение данных	X_{min}	686			
Максимальное значение данных	X_{max}	57246			
Размах вариации	$R = X_{max} - X_{min}$	56560			
Количество наблюдений	N	80			
Количество интервалов	$n = \lceil 1 + \log_2 N \rceil$	7			
Ширина интервала	$h = \frac{R}{n}$	8080			
Границы интервалов	$b_i = a_i + h$	$a_1 = X_{min}$	$a_{i+1} = b_i$		
Середина интервала	$X_{срi} = \frac{(a_i + b_i)}{2}$				

Рисунок 7. Величины необходимые для построения интервального вариационного ряда.

Таблица 1. Интервальный вариационный ряд

Интервал	Нижняя граница	Верхняя граница	Середина интервала	Частота	Накопленная Частота
1	686	8766	4726	22	22
2	8766	16846	12806	26	48
3	16846	24926	20886	14	62
4	24926	33006	28966	3	65
5	33006	41086	37046	7	72
6	41086	49166	45126	4	76
7	49166	57246	53206	4	80

5. Построение графиков распределения признака в вариационном ряду

Для дальнейшего анализа, построенного в предыдущем разделе ряда, было принято решение построить графическое представление статистических данных. Далее представлены: гистограмма (рис. 8), полигон (рис. 9) для распределения признака

по частотам (из табл. 1) и кумулята (рис. 10), огива (рис. 11) для распределения признака по накопленным частотам (из табл. 1).

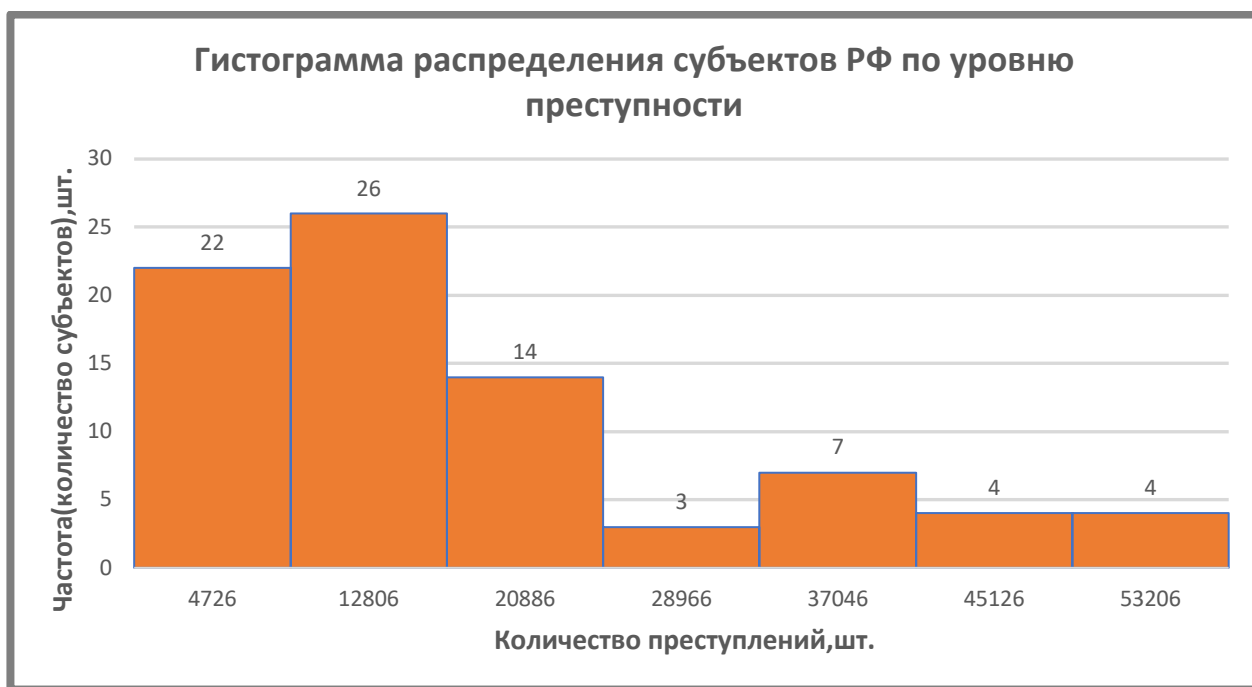


Рисунок 8. Гистограмма интервального ряда распределения субъектов РФ по уровню преступности



Рисунок 9. Полигон интервального ряда распределения субъектов РФ по уровню преступности

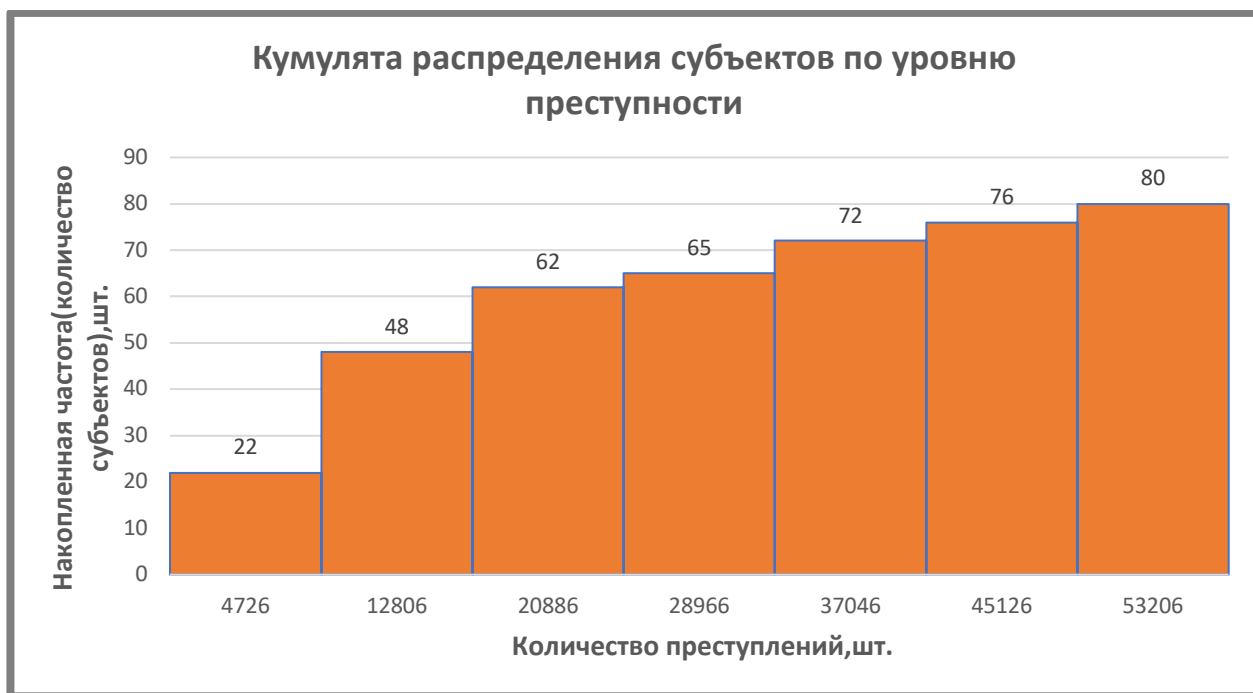


Рисунок 10. Кумулята интервального ряда распределения субъектов РФ по уровню преступности



Рисунок 11. Огива интервального ряда распределения субъектов РФ по уровню преступности

Исследования построенных графиков, точнее раскрывают устройство данных и позволяют сделать несколько выводов:

- На предыдущих этапах был сделан вывод что более половины субъектов имеют значение показателя менее 40000. Из гистограммы, можно усилить данную оценку: >50 % (48) выборки имеют значение не более чем 12806. Что говорит о

среднем уровне преступности по всей стране: преобладают города с низким уровнем преступности.

- На полигоне (рис. 9) стоит отметить резкое снижение значения частоты, после 2 интервала. В результате подтверждаем предыдущий вывод раздела.
- Анализируя огиву, стоит отметить выпуклую форму графика (огива почти на всем размахе вариации выпукла), как следствие, это говорит о низком росте накопленной частоты для значений >20000. Рассматривая участок размаха вариации 20000 – 53206, где накопленная частота увеличивается на 3 – 5 единиц, можно сделать вывод, 20 - 25 % субъектов равномерно распределены на этом интервале.

6. Расчёт показателей центра распределения

Следующий этап состоит из расчета средних величин для дискретного и интервального рядов. В этом разделе мы найдем среднее арифметическое, моду и медиану ряда.

6.1 Дискретный ряд

- Среднее арифметическое рассчитывается по следующей формуле:

$$X_{\text{ср.}} = \frac{\sum x_i}{N}$$

Достаточно сложить все значения(x_i) и поделить на количество наблюдений $N = 80$.

- Медиана определяется по следующей формуле:

$$Me = \frac{X_{n/2} + X_{n/2+1}}{2}$$

Поскольку у нас четное количество наблюдений ($N = 80$), медиана определяется как среднее арифметическое двух значений, расположенных в центре ряда $= (x_{40} + x_{41}) / 2$

- Мода - значение признака, которое наиболее часто встречается в совокупности. В исследуемой выборке, все значения уникальны. В данном случае моды нет, иными словами, все варианты одинаково модальны.

6.2 Интервальный ряд

Для нахождения среднего арифметического интервального ряда, к таблице 1, был добавлен столбец, где записан результат умножение середины интервала на частоту (табл. 2).

Таблица 2. Интервальный вариационный ряд распределения

Интервал	Нижняя граница	Верхняя граница	Середина интервала, $X_{\text{ср.}i}$	Частота, f_i	Накопленная частота	$X_{\text{ср.}i} * f_i$
1	686	8766	4726	22	22	103972
2	8766	16846	12806	26	48	332956
3	16846	24926	20886	14	62	292404
4	24926	33006	28966	3	65	86898
5	33006	41086	37046	7	72	259322
6	41086	49166	45126	4	76	180504
7	49166	57246	53206	4	80	212824

Рисунок 12, содержит подробный алгоритм вычисления средних для ряда.

а) Среднее для интервального ряда			
Для группировки с равными интервалами среднее вычисляется по следующей формуле:			
$X_{\text{ср.}} = \frac{\sum x_i f_i}{N}$	x_i - середина i интервала		
	f_i - частота i интервала		
	N - количество наблюдений		
		$X_{\text{ср.}} =$	18361
b) Медиана интервального ряда			
$Me = x_H + h * \frac{\frac{\sum f}{2} - F_{Me-1}}{f_{Me}}$	x_H - нижняя граница медианного интервала		
	h - длина медианного интервала		
	F_{Me-1} - накопленная частота предмедианного интервала		
	f_{Me} - частота медианного интервала		
<p>Сначала найдем медианный интервал, накопленная частота которого первой превышает половину сумм всех частот. Первый такой интервал с накоп. частотой > 40 является 2 интервал. Следовательно $x_H = 8766$. Длина медианного интервала $h = 8080$, как и других, поскольку мы рассматриваем группировку с равными интервалами. Сумма частот равна количеству наблюдений $N = 80$. Накопленная частота предмедианного интервала(первого) = частоте 1 интервала = 22. Частота медианного интервала = 26.</p>			
$Me = 8766 + 8080 * \frac{\frac{80}{2} - 22}{26}$			
		$Me =$	14359,846
c) Мода интервального ряда			
$Mo = x_H + h * \frac{(f_2 - f_1)}{(f_2 - f_1 + f_2 - f_3)}$	x_H - нижняя граница модального интервала		
	h - длина модального интервала		
	f_1 - частота предмодального интервала		
	f_2 - частота модального интервала		
<p>Сначала находим модальный интервал, интервал с наибольшей частотой. Наибольшая частота достигается во 2 интервале, 1ый и 3ий предмодальный и послемодальный интервалы соответственно.</p>	f_3 - частота послемодального интервала		
$Mo = 8766 + 8080 * \frac{(26 - 22)}{(26 - 22 + 26 - 14)}$			
		$Mo =$	10786

Рисунок 12. Алгоритм нахождения средних для интервального ряда

В результате получаем следующие значения средних, отображенные в таблице 3. При переходе к интервальному ряду, вместо рассмотрения 80 объектов, используем только 7 интервалов с соответствующими частотами, из-за чего уменьшается точность вычислений (Результаты вычислений среднего и медианы для рядов отличаются $< 3\%$). В данном случае, при исследовании мы получили значения среднего и медианы меньшие чем для дискретного ряда.

Таблица 3. Сравнение показателей центра распределения

	Дискретный ряд	Интервальный ряд
Хср.	18798,34	18361
Me	14472,5	14359,846
Mo	Моды нет	10786

После сравнения найденных средних, можно сделать определенные выводы:

- Моду для дискретного ряда найти не удалось, так как все значения в выборке уникальны. То есть нет двух субъектов с одинаковым уровнем преступности.
- Стоит отметить, поскольку большинство значений сгруппированы в первых двух интервалах, мода и медиана тоже попадают в эту область (второй интервал). Можно сделать вывод, что почти 60 % регионов имеют количество преступлений <35 % от максимума (57246).
- Среднее арифметическое > Медиана > Мода, следовательно они представляют положительную или правостороннюю асимметрию. То есть для нашего ряда характерен удлинённый “хвост” справа, который также проявляется при исследовании огивы.

7. Расчет показателей вариации

В данном разделе, рассчитаны такие показатели как: дисперсия, среднее квадратическое отклонение, среднее линейное отклонение, коэффициент осцилляции, коэффициент линейной вариации, коэффициент вариации. Для этого были построены вспомогательные таблицы для дискретного (Приложение 3) и интервального вариационного ряда (Приложение 4).

7.1 Дискретный ряд

Далее на рисунке 14 изображены формулы, по которым были вычислены все величины. На рисунке 13, пояснение обозначений. При последовательном вычислении коэффициентов, на каждом шаге нам известны все значения для применения формулы.

x_i - i значение ряда

\bar{x} — среднее арифметическое

N - количество наблюдений

D - дисперсия

S - среднее квадратическое отклонение

\bar{d} — среднее линейное отклонение

R - размах вариации

VR - коэффициент осцилляции

Vd - коэффициент линейной вариации

V - коэффициент вариации

Дисперсия	Используя встроенную функцию, рассчитывается с помощью формулы: $D_1 = \text{ДИСПР}([Xi])$		
	$D_2 = \overline{x^2} - \bar{x}^2 \qquad D_3 = \frac{\sum (x_i - \bar{x})^2}{N}$		
Среднее квадратическое отклонение	Вычисляется как корень из значения дисперсии		
	$S = \sqrt{D}$		
Среднее линейное отклонение	Средняя арифметическая из абсолютных отклонений значений признака от средней.		
	$\bar{d} = \frac{\sum x_i - \bar{x} }{N}$		
Коэффициент осцилляции	Характеризует размах колебаний крайних значений признака вокруг средней арифметической		
	$V_R = \frac{R}{\bar{x}} 100\%$		
Коэффициент линейной вариации	Отношение среднего линейного отклонения к средней величине		
	$V_{\bar{d}} = \frac{\bar{d}}{\bar{x}} 100\%$		
Коэффициент вариации	Относительная мера среднего разброса значений в статистической совокупности		
	$V = \frac{S}{\bar{x}} 100\%$		

Рисунок 14. Формулы для вычисления коэффициентов вариации дискретного ряда

7.2 Интервальный ряд

Аналогичным образом вычисляются коэффициенты для интервального ряда (рис. 15) Результаты вычислений занесены в таблицу 4.

Дисперсия	Мера разброса относительно среднего значения, средний квадрат отклонений индивидуальных значений признака от их средней величины	
	$D_1 = \frac{\sum (x_i - \bar{x})^2 f_i}{N}$	$D_2 = \overline{x^2} - \bar{x}^2$
Среднее квадратическое отклонение	Вычисляется как корень из значения дисперсии	
	$S = \sqrt{D}$	
Среднее линейное отклонение	Средняя арифметическая из абсолютных отклонений значений признака от средней.	
	$\bar{d} = \frac{\sum x_i - \bar{x} f_i}{N}$	
Коэффициент осцилляции	Характеризует размах колебаний крайних значений признака вокруг средней арифметической	
	$V_R = \frac{R}{\bar{x}} 100\%$	
Коэффициент линейной вариации	Отношение среднего линейного отклонения к средней величине	
	$V_{\bar{d}} = \frac{\bar{d}}{\bar{x}} 100\%$	
Коэффициент вариации	Относительная мера среднего разброса значений в статистической совокупности	
	$V = \frac{S}{\bar{x}} 100\%$	

Рисунок 15. Формулы для вычисления коэффициентов вариации интервального ряда

Таблица 4. Сводная таблица для сравнения и интерпретации результатов

	Ряд исходных данных	Интервальный ряд
Среднее арифметическое (\bar{X})	18798,3375	18361
Дисперсия (S^2)	195797372,6	193563975
Среднее квадратическое отклонение (S)	13992,76143	13912,72709
Среднее линейное отклонение (d)	10966,21375	11110
Размах (R)	56560	56560
Коэффициент осцилляции (V_r), %	301%	308%
К-т лин. вариации (V_d), %	58%	61%
К-т вариации (V), %	74%	76%

После вычисления показателей вариации можно сделать несколько выводов об устройстве данных:

- Среднее квадратическое отклонение почти на 50 % больше, чем $R/6$ (одно из эмпирических соотношений для проверки однородности). Также стоит заметить, что коэффициент вариации 74 и 76 % соответственно при двух расчетах. Такой коэффициент вариации превышает более чем в 2 раза допустимое значение для нормальных данных (<33 %). Поэтому я сделаю предположение о неоднородности информации и необходимости исключения самых больших значений.
- Коэффициент осцилляции > 300 %, что говорит о большом размахе крайних значений по отношению к среднему. Действительно, в исследуемых данных > 50 % значений сгруппированы в первых двух интервалах, иными словами, самые большие значения увеличивают размах колебаний не пропорционально количеству наблюдений.

8. Показатели дифференциации

В данном разделе рассчитываются квартили, интерквартильный размах, коэффициент асимметрии и коэффициент эксцесса. Обозначения сохраняются с рис. 13.

8.1 Дискретный ряд

Далее на рисунках 16 и 17 соответственно предоставлены алгоритмы по расчёту квартилей и характеристик форм распределения.

Нижний квартиль	Используя встроенную функцию, рассчитывается с помощью формулы =КВАРТИЛЬ.ИСКЛ(С7:С86;1)		
	$Q_1 = X_{\frac{N+1}{4}} = X_{20} + \frac{1}{4}(X_{21} - X_{20})$	Значение ранжированного ряда данных, левее которого находится 25% всех наблюдений. Для N = 80, находится между 20 и 21 элементами.	
Второй квартиль (Медиана)	Используя встроенную функцию, рассчитывается с помощью формулы =КВАРТИЛЬ.ИСКЛ(С7:С86;2)		
	$Q_2 = Me = \frac{X_{n/2} + X_{n/2+1}}{2}$	Значение ранжированного ряда данных, левее которого находится 50% всех наблюдений, является медианой исследуемого ряда.	
Верхний квартиль	Используя встроенную функцию, рассчитывается с помощью формулы =КВАРТИЛЬ.ИСКЛ(С7:С86;3)		
	$Q_3 = X_{\frac{3(N+1)}{4}} = X_{60} + \frac{3}{4}(X_{61} - X_{60})$	Значение ранжированного ряда данных, левее которого находится 75% всех наблюдений. Для N = 80, находится между 60 и 61 элементами.	
Интерквартильный размах	$IQR = Q_3 - Q_1$	Показывает разброс средней половины набора данных, определяется как разность верхнего и нижнего квартилей, и помогает определить выбросы	

Рисунок 16. Алгоритм нахождения квартилей и интерквартильного размаха

Коэффициент асимметрии	Характеризующая асимметрию распределения данной выборки.	
	$As = \frac{\sum (x_i - \bar{x})^3}{N * S^3}$	
Коэффициент эксцесса	Мера остроты пика распределения данных значений.	
	$E_k = \frac{\sum (x_i - \bar{x})^4}{N * S^4} - 3$	

Рисунок 17. Алгоритм нахождения коэффициентов эксцесса и асимметрии

8.2 Интервальный ряд

На рисунке 18 изображена общая формула определения квартилей по интервальному ряду распределения. На рисунках 19 и 20 соответственно предоставлены алгоритмы по расчёту квартилей и характеристик форм распределения. Результаты в таблице 5.

Общая формула определения квартилей по интервальному ряду распределения		
$Q_i = a_{Q_i} + h \frac{\frac{i}{4} \sum m_j - m_{Q_i-1}^H}{m_{Q_i}}$	a_{Q_i}	Нижняя граница интервала, содержащего i-ю квартиль
	h	Величина интервала
	m_{Q_i}	Частота интервала, содержащего i-ю квартиль
	$m_{Q_i}^H$	Накопленная частота интервала, предшествующего интервалу, содержащему i-ю квартиль
	$\sum m_j$	Сумма всех частот

Рисунок 18. Общая формула для определения квартилей по интервальному ряду распределения

Нижний квартиль	$Q_1 = 686 + 8080 \frac{1}{4} * 80$	25% от всех наблюдений = 20, то есть первый интервал будет содержать первый квартиль,отсюда находим нижнюю границу интервала = 686 и частоту данного интервала = 22, получается накопленная частота перед ним = 0. h фиксированна и равна 8080
Второй квартиль (Медиана)	$Q_2 = 8766 + 8080 \frac{1}{2} * 80 - 22$	50% от всех наблюдений = 40, то есть второй интервал будет содержать медиану,отсюда находим нижнюю границу интервала = 8766 и частоту данного интервала = 26, получается накопленная частота перед ним = 22. h = 8080
Верхний квартиль	$Q_3 = 16846 + 8080 \frac{3}{4} * 80 - 48$	75% от всех наблюдений = 60, то есть третий интервал будет содержать верхний квартиль,отсюда находим нижнюю границу интервала = 16846 и частоту данного интервала = 14, получается накопленная частота перед ним = 48. h = 8080
Интерквартильный размах	$IQR = Q_3 - Q_1$	Показывает разброс средней половины набора данных, определяется как разность верхнего и нижнего квартилей

Рисунок 19.Алгоритм расчета квартилей для интервального ряда

Коэффициент асимметрии	Характеризующая асимметрию распределения данной выборки.
	$As = \frac{\sum(x_i - \bar{x})^3 * f_i}{N * S^3}$
Коэффициент эксцесса	Мера остроты пика распределения данных значений.
	$E_k = \frac{\sum(x_i - \bar{x})^4 * f_i}{N * S^4} - 3$

Рисунок 20. Алгоритм расчета коэффициента асимметрии и коэффициента эксцесса

Таблица 5. Сводная таблица для сравнения и интерпретации полученных результатов

	Ряд исходных данных	Интервальный ряд
Нижний квартиль (Q1)	8310	8031,455
Второй квартиль (Q2)	14472,5	14359,846
Верхний квартиль (Q3)	23461,75	23771,714
Интерквартильный размах	15151,75	15740,260
Коэффициент асимметрии	1,070	1,092
Коэффициент эксцесса	0,244	0,212

Измеряя показатели дифференциации и характеристики формы распределения для дискретного и интервального рядов мы получили схожие результаты по каждому показателю, отличия есть из-за погрешности при переходе от дискретного к интервальному ряду. Можно сделать следующие выводы:

- 3 квартиля лежат в первых 3 интервалах, следовательно в нашей выборке преобладает количество городов с низким уровнем преступности, в сравнении с размахом выборки.
- 3 квартиля почти пропорциональны друг другу, то есть они разбивают 75% данных на 3 равные части, и из пропорциональности можно судить о равномерном росте значений в этом интервале, дискретной непрерывности.
- $As > 1$, распределение несимметрично, правосторонняя асимметрия, причем существенная.

- $E_k > 0$, в изучаемом распределении означает, что в совокупности есть сформировавшееся ядро, т.е. большинство значений показателя находятся вокруг среднего. График распределения с выраженным пиком.
- В целом получаем следующие представление данных: большинство регионов имеет значение уровня преступности около среднего, но часть регионов (< 20 %) имеет аномальный уровень преступности, которые в точечной диаграмме отображены в удлинённом правом "хвосте"

9. Анализ временных данных

Было принято решение для анализа уровня преступности выбрать Москву, так как она является крупнейшим субъектом РФ, не рассматривалась в основном исследовании, и информация по количеству преступлений в таком регионе будет актуальна. Исследование уровня преступности будет затрагивать промежуток времени с 2009 года по 2022 год

9.1 Используемый показатель для анализа

Показатель для каждого года остался такой же: обработав информацию с официального сайта Государственной прокуратуры Российской Федерации, составил таблицу 6 с уровнем преступности в Москве с 2009 по 2022 год.

Таблица 6. Динамика уровня преступности в Москве с 2009 по 2022 год

№	Год	Зарегистрировано преступлений, шт
1	2009	212050
2	2010	185911
3	2011	173628
4	2012	180240
5	2013	174990
6	2014	182873
7	2015	195239
8	2016	173898
9	2017	140134
10	2018	140542
11	2019	142081
12	2020	146559
13	2021	143564
14	2022	138180

9.2 Анализ на выбросы

Для нахождения аномальных значений, было принято решение построить диаграмму размаха (рис. 21) и точечную диаграмму (рис. 22). После анализа которых, выбросы не были обнаружены, и я решил продолжать исследование с этим же набором данных.



Рисунок 21. Ящик с усами для динамики уровня преступности в Москве с 2009 по 2022 год

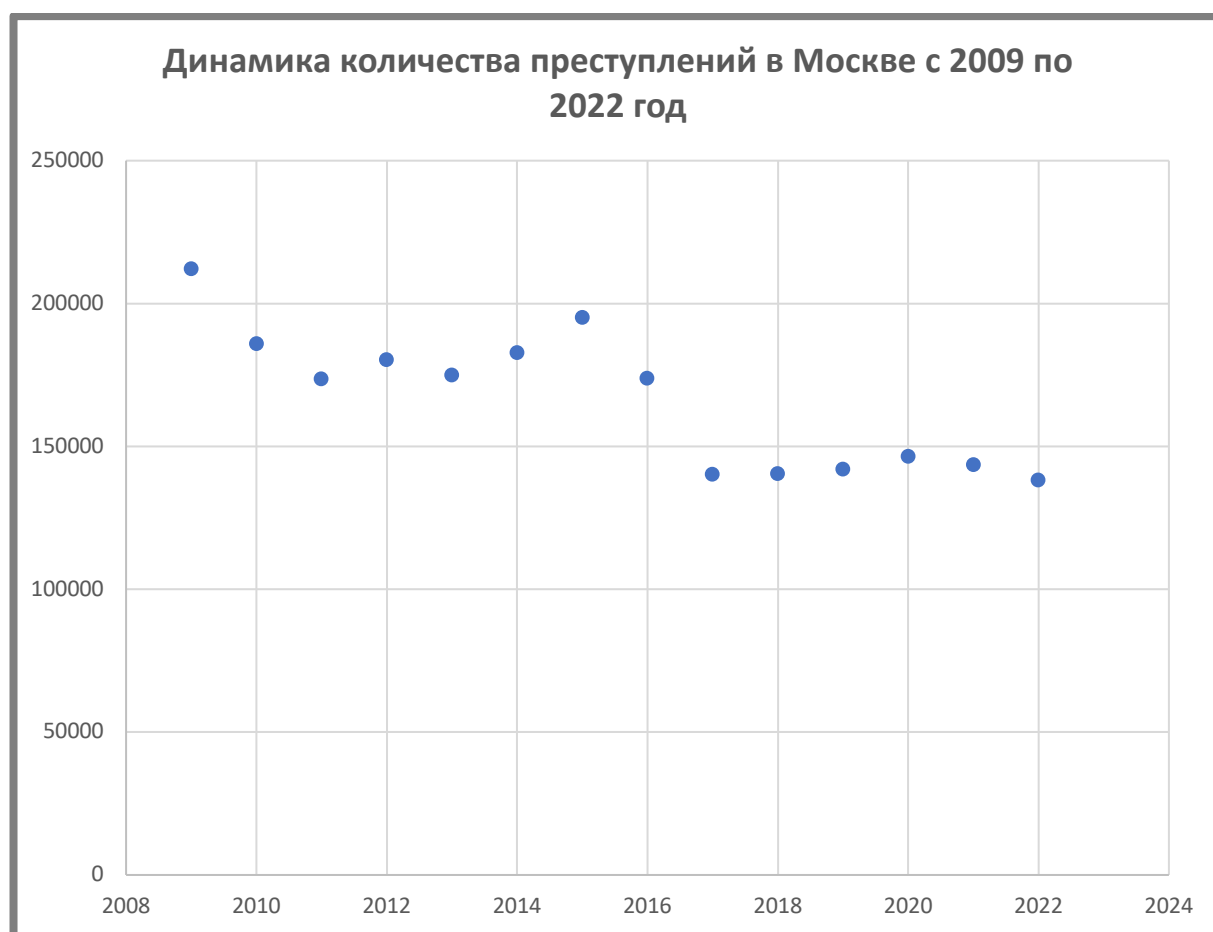


Рисунок 22. Динамика уровня преступности в Москве с 2009 по 2022 год

9.3 Расчет показателей рядов динамики

В этом разделе рассчитываются цепные и базисные показатели динамики, а именно: абсолютный прирост, темп прироста и темп роста. Показатели относительные, поэтому 2009 год был принят за базисный и расчеты проводились относительно него. Далее в работе за Y_k будем обозначать уровень преступности за k год.

9.3.1 Цепные показатели

Абсолютный прирост за k год = насколько изменилось количество преступлений по сравнению с предыдущим годом, то есть $= Y_k - Y_{k-1}$.

Темп роста за k год вычисляется аналогично, только он равен показателю: во сколько раз в процентах изменился уровень преступности $= (Y_k / Y_{k-1}) * 100\%$

Темп прироста показывает, насколько процентов изменился уровень преступности по сравнению с предыдущим годом то есть = темп роста за k год – 100.

Выполнив расчеты по описанным формулам показатели были занесены в таблицу 7.

Таблица 7. Цепные показатели динамики

Дата	Цепные показатели динамики			
	Количество преступлений, шт.	Абсолютный прирост, шт.	Темп роста, %	Темп прироста, %
2009	212050			
2010	185911	-26139	87,67	-12,33
2011	173628	-12283	93,39	-6,61
2012	180240	6612	103,81	3,81
2013	174990	-5250	97,09	-2,91
2014	182873	7883	104,50	4,50
2015	195239	12366	106,76	6,76
2016	173898	-21341	89,07	-10,93
2017	140134	-33764	80,58	-19,42
2018	140542	408	100,29	0,29
2019	142081	1539	101,10	1,10
2020	146559	4478	103,15	3,15
2021	143564	-2995	97,96	-2,04
2022	138180	-5384	96,25	-3,75

9.3.2 Базисные показатели

Абсолютный показатель показывает прирост количества преступлений относительно базисного года $= Y_k - Y_1$

Темп роста показывает процентный прирост количества преступлений относительно базисного года $= (Y_k / Y_1) * 100$

Темп прироста рассчитывается аналогично как цепной показатель.

Рассчитанные показатели были занесены в таблицу 8.

Таблица 8. Базисные показатели динамики

Дата	Базисные показатели динамики			
	Количество преступлений, шт.	Абсолютный прирост, шт.	Темп роста, %	Темп прироста, %
2009	212050			
2010	185911	-26139	87,67	-12,33
2011	173628	-38422	81,88	-18,12
2012	180240	-31810	85,00	-15,00
2013	174990	-37060	82,52	-17,48
2014	182873	-29177	86,24	-13,76
2015	195239	-16811	92,07	-7,93
2016	173898	-38152	82,01	-17,99
2017	140134	-71916	66,09	-33,91
2018	140542	-71508	66,28	-33,72
2019	142081	-69969	67,00	-33,00
2020	146559	-65491	69,12	-30,88
2021	143564	-68486	67,70	-32,30
2022	138180	-73870	65,16	-34,84

9.3.3 Анализ показателей

После составления таблиц с показателями динамики, было принято решение построить графическое представление базисного (рис. 23) и цепного темпа прироста (рис. 24).



Рисунок 23. Цепной темп прироста уровня преступности



Рисунок 24. Базисный темп прироста уровня преступности

Анализируя табл. 7, табл. 8 и рис. 23, рис. 24 можно сделать следующие выводы:

- Наибольший уровень преступности наблюдается в 2009 году (212050), а наименьший в 2017 (140134)
- После чего на протяжении 14 лет уровень преступности не поднимался до такого высокого значения, а к 2022 году он упал почти на 35%
- Анализируя рис. 23, когда уровень преступности начинал расти он достигал пика в 3 точках 2012, 2015, 2020 годах.
- Стоит отметить, что рост и спад уровня преступности происходит хотя бы в течении 2ух лет(кроме перемены в 2012 – 2013 годах), то есть если уровень преступности начал расти в каком то году, то в следующем году он тоже будет расти, аналогично для снижения количества преступлений.
- Также можно заметить, рост уровня преступности в пике увеличивается <10 %, а при снижении уровня преступности: он мог упасть почти на 20%

9.3.4 Прогноз уровня преступности в Москве на 2023 год

После анализа 14 уровней количества преступлений в Москве, следующий этап заключался в прогнозе показателя на 2023 год. Прогноз осуществлялся с помощью 2 методов: среднего абсолютного прироста и среднего темпа роста. Расчеты представлены на рисунке 25.

Средний абсолютный прирост	$\Delta \bar{y} = \frac{y_n - y_1}{n - 1}$	-5682,31	Прогноз по методу среднего абсолютного прироста	$y_{n+1} = y_n + \Delta \bar{y}$	132497,69
Средний темп роста	$\bar{T} = \sqrt[n-1]{\frac{y_n}{y_1}} * 100\%$	96,76	Прогноз по методу среднего темпа роста	$y_{n+1} = y_n * \frac{\bar{T}}{100}$	133702,04

Рисунок 25. Прогноз уровня преступности в Москве на 2023 год

После чего были построены точечные диаграммы для 2ух прогнозов (рис. 26 и рис. 27)



Рисунок 26. Прогноз с помощью метода среднего абсолютного прироста



Рисунок 27. Прогноз с помощью метода среднего темпа роста

Исследуя построенные графики, можно прийти к следующему выводу, что в 2023 году уровень преступности будет продолжать уменьшаться.

9.4 Построение линии тренда

Последний этап был построение линии тренда. Для этого я построил таблицу 9. Алгоритм построения: сначала заполнили известные данные: это год и соответственно количество

преступлений. После чего заполняем t и t^2 начиная с середины данных. Далее заполняем столбец yt умножая количество преступлений на соответствующие t . После чего необходимо найти уравнение линейного тренда $y = a_1 \cdot t + a_0$. a_0 = сумма всех значений количества преступлений деленое на количество наблюдений, таким образом $a_0 = 166420,64$. Старший коэффициент = (сумма yt)/(сумма t^2) = $-9931,92$. То есть прямая линии тренда задается уравнением $y = -9931 \cdot t + 166420$. После чего в это уравнение подставляем все значения t и находим столбец $Y_{\text{тренд}}$, подставив $t = 7$, найдем прогноз значения на 2023 год = 96 897. Построим линию тренда (рис. 28)

Таблица 9. Построение линии тренда

Год	Количество преступлений, шт.	t	t^2	Yt	$Y_{\text{тренд}}$
2009	212050	-7	49	-1484350	235944,10
2010	185911	-6	36	-1115466	226012,18
2011	173628	-5	25	-868140	216080,25
2012	180240	-4	16	-720960	206148,33
2013	174990	-3	9	-524970	196216,41
2014	182873	-2	4	-365746	186284,49
2015	195239	-1	1	-195239	176352,56
2016	173898	0	0	0	166420,64
2017	140134	1	1	140134	156488,72
2018	140542	2	4	281084	146556,80
2019	142081	3	9	426243	136624,88
2020	146559	4	16	586236	126692,95
2021	143564	5	25	717820	116761,03
2022	138180	6	36	829080	106829,11
Сумма	2329889		231	-2294274	

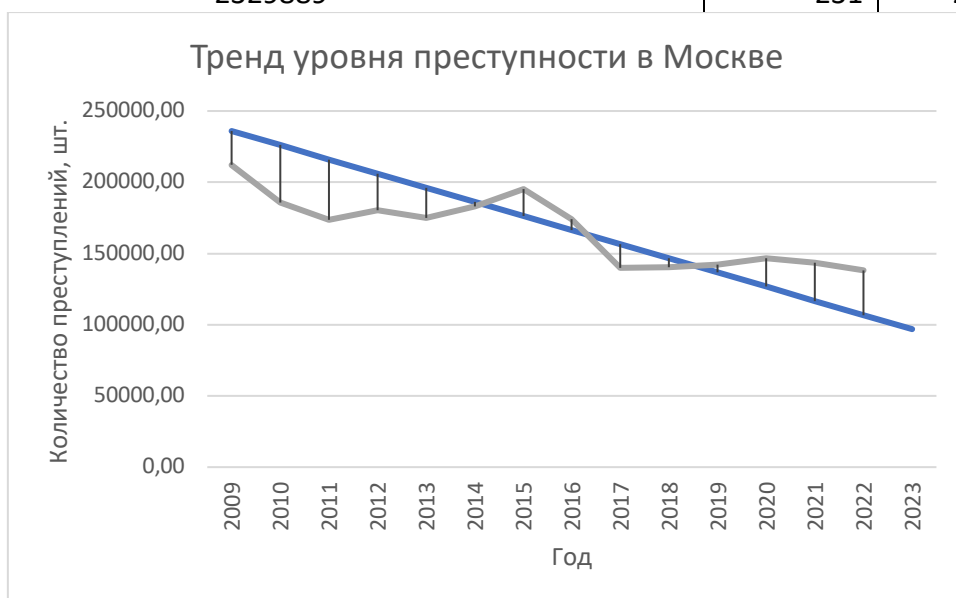


Рисунок 28. Линия тренда

Анализируя линию тренда, можно сделать выводы:

- Можно заметить резкое снижение уровня преступности по линии тренда, которое на самом деле падает гораздо медленнее

- В целом линия тренда показывает, что уровень преступности будет снижаться. Но рассчитанные прогнозы выше, показывают снижение примерно на 5000, а не на 30000, как показывает линия тренда

10. Общие выводы по работе

Выбрав актуальные данные и убрав из них аномальные значения, я построил точечные диаграммы, гистограммы, кумуляту, огиву, наши показатели центра распределения и показатели вариации. После чего были найдены показатели дифференциации и рядов динамики. Проанализировав найденные показатели, можно сделать несколько выводов по поставленным гипотезам и устройству ряда данных:

- Наиболее высокий уровень преступности действительно у крупнейших городов таких как Москва и Санкт Петербург, но далее по уровню преступности идут Челябинская и Ростовская области, население в которых не самое высокое (за исключением Москвы и Санкт Петербурга), то есть уровень преступности зависит от размера рассматриваемого региона, но косвенно, следовательно существуют множество других существенных факторов, от которых зависит количество преступлений
- Гипотеза о равномерном распределении городов оказалась не верной, устройство данных можно представить следующим образом: сформировано ядро, вокруг которого сформировалось большинство городов с не большим уровнем преступности, и <20% городов имеют более высокий уровень преступности, которые можно видеть в правом “хвосте”. Распределение данных несимметрично, неоднородно, правая асимметрия
- В Москве уровень преступности действительно снизился, и снижается, но не монотонно, в некоторых годах он не существенно растет. По прогнозу на 2023 год уровень преступности упадет еще сильнее.

11. Список литературы

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. — Режим доступа : <https://biblio-online.ru/viewer/C6BE26A0-4B8E-47F4-B5F0-7388CCEBF6E0/statistika-v-2-chast-1#page/1>

12. Используемые информационные источники

Сайт Государственной прокуратуры Российской Федерации : <https://epp.genproc.gov.ru/web/gprf/activity/crimestat>

13. Приложения

Приложение 1

Таблица 10. Количество зарегистрированных преступлений по субъектам РФ аз 2022 год

№	Регион	Зарегистрировано преступлений, шт
1	Москва	138180
2	Санкт Петербург	62971
3	Севастополь	6269

4	Амурская область	16253
5	Архангельская область	16786
6	Астраханская область	13062
7	Белгородская область	14588
8	Брянская область	13549
9	Челябинская область	62844
10	Воронежская область	30460
11	Иркутская область	37867
12	Ивановская область	11791
13	Калининградская область	12645
14	Калужская область	14656
15	Кемеровская область	47071
16	Кировская область	18519
17	Костромская область	8188
18	Курганская область	13456
19	Курская область	13448
20	Ленинградская область	28569
21	Липецкая область	13940
22	Магаданская область	2386
23	Московская область	74054
24	Мурманская область	12075
25	Нижегородская область	39404
26	Новгородская область	10615
27	Новосибирская область	46814
28	Омская область	23074
29	Оренбургская область	21742
30	Орловская область	8243
31	Пензенская область	13274
32	Псковская область	8714
33	Ростовская область	57250
34	Рязанская область	9154
35	Сахалинская область	8511
36	Самарская область	47060
37	Саратовская область	29200
38	Смоленская область	14611
39	Свердловская область	51822
40	Тамбовская область	13799
41	Томская область	18578
42	Тверская область	21352

43	Тульская область	15826
44	Тюменская область	23591
45	Ульяновская область	12159
46	Владимирская область	16254
47	Волгоградская область	36031
48	Вологодская область	17411
49	Ярославская область	18368
50	Республика Адыгея	4388
51	Республика Алтай	4586
52	Республика Башкортостан	50751
53	Республика Бурятия	19955
54	Республика Дагестан	14357
55	Республика Ингушетия	2233
56	Кабардино-Балкарская Республика	7416
57	Республика Калмыкия	2693
58	Карачаево-Черкесская Республика	4436
59	Республика Карелия	11850
60	Республика Коми	14797
61	Республика Крым	20779
62	Республика Марий Эл	7741
63	Республика Мордовия	8107
64	Республика Саха (Якутия)	13000
65	Республика Северная Осетия-Алания	8065
66	Республика Татарстан	52189
67	Республика Тыва	6619
68	Удмуртская Республика	24583
69	Республика Хакасия	9493
70	Чеченская Республика	2416
71	Чувашская Республика	12730
72	Алтайский край	39177
73	Камчатский край	5450
74	Хабаровский край	21312
75	Краснодарский край	77125
76	Красноярский край	45561
77	Пермский край	41027
78	Приморский край	33604
79	Ставропольский край	33569
80	Забайкальский край	20816
81	Ненецкий автономный округ	686

82	Ханты-Мансийский автономный округ	20376
83	Чукотский автономный округ	853
84	Ямало-Ненецкий автономный округ	6657
85	Еврейская автономная область	3164

Приложение 2

Таблица 11. Количество зарегистрированных преступлений по субъектам РФ за 2022 год (исключая выбросы ранжированные значения)

№	Регион	Зарегистрировано преступлений, шт
81	Ненецкий автономный округ	686
83	Чукотский автономный округ	853
55	Республика Ингушетия	2233
22	Магаданская область	2386
70	Чеченская Республика	2416
57	Республика Калмыкия	2693
85	Еврейская автономная область	3164
50	Республика Адыгея	4388
58	Карачаево-Черкесская Республика	4436
51	Республика Алтай	4586
73	Камчатский край	5450
3	Севастополь	6269
67	Республика Тыва	6619
84	Ямало-Ненецкий автономный округ	6657
56	Кабардино-Балкарская Республика	7416
62	Республика Марий Эл	7741
65	Республика Северная Осетия-Алания	8065
63	Республика Мордовия	8107
17	Костромская область	8188
30	Орловская область	8243
35	Сахалинская область	8511
32	Псковская область	8714
34	Рязанская область	9154
69	Республика Хакасия	9493
26	Новгородская область	10615
12	Ивановская область	11791
59	Республика Карелия	11850

24	Мурманская область	12075
45	Ульяновская область	12159
13	Калининградская область	12645
71	Чувашская Республика	12730
64	Республика Саха (Якутия)	13000
6	Астраханская область	13062
31	Пензенская область	13274
19	Курская область	13448
18	Курганская область	13456
8	Брянская область	13549
40	Тамбовская область	13799
21	Липецкая область	13940
54	Республика Дагестан	14357
7	Белгородская область	14588
38	Смоленская область	14611
14	Калужская область	14656
60	Республика Коми	14797
43	Тульская область	15826
4	Амурская область	16253
46	Владимирская область	16254
5	Архангельская область	16786
48	Вологодская область	17411
49	Ярославская область	18368
16	Кировская область	18519
41	Томская область	18578
53	Республика Бурятия	19955
82	Ханты-Мансийский автономный округ	20376
61	Республика Крым	20779
80	Забайкальский край	20816
74	Хабаровский край	21312
42	Тверская область	21352
29	Оренбургская область	21742
28	Омская область	23074
44	Тюменская область	23591
68	Удмуртская Республика	24583
20	Ленинградская область	28569
37	Саратовская область	29200
10	Воронежская область	30460
79	Ставропольский край	33569
78	Приморский край	33604

47	Волгоградская область	36031
11	Иркутская область	37867
72	Алтайский край	39177
25	Нижегородская область	39404
77	Пермский край	41027
76	Красноярский край	45561
27	Новосибирская область	46814
36	Самарская область	47060
15	Кемеровская область	47071
52	Республика Башкортостан	50751
39	Свердловская область	51822
66	Республика Татарстан	52189
33	Ростовская область	57246

Приложение 3

Таблица 12. Вспомогательные данные для дискретного ряда

X_i^2	$X_i - X_{ср.}$	$ X_i - X_{ср.} $	$(X_i - X_{ср.})^2$	$(X_i - X_{ср.})^3$	$(X_i - X_{ср.})^4$
470596	18112,34	18112,34	328056769,7	-5941874932218,05	107621244155123000,00
727609	17945,34	17945,34	322035138	-5779029238070,00	103706630099534000,00
4986289	16565,34	16565,34	274410406,5	-4545700997000,92	75301071189406800,00
5692996	16412,34	16412,34	269364822,2	-4420906372802,13	72557407446329300,00
5837056	16382,34	16382,34	268380982	-4396707825114,13	72028351479910600,00
7252249	16105,34	16105,34	259381896	-4177432976291,23	67278967966799800,00
10010896	15634,34	15634,34	244432509,1	-3821540342676,92	59747251487276600,00
19254544	14410,34	14410,34	207657826,9	-2992419369625,46	43121773057840100,00
19678096	14362,34	14362,34	206276738,5	-2962616136217,85	42550092831306800,00
21031396	14212,34	14212,34	201990537,2	-2870757686690,35	40800177123962500,00
29702500	13348,34	13348,34	178178114	-2378381600971,10	31747440313552600,00
39300361	12529,34	12529,34	156984298,2	-1966909254209,45	24644069877863400,00
43811161	12179,34	12179,34	148336261,9	-1806637397642,34	22003646606007800,00
44315649	12141,34	12141,34	147412076,3	-1789779769799,36	21730320235806300,00

54997056	- 11382,34	11382,34	129557607	-1474668408155,53	16785173522214000,00
59923081	- 11057,34	11057,34	122264712,6	-1351922191436,04	14948659944447900,00
65044225	- 10733,34	10733,34	115204533,9	-1236529143759,82	13272084628560200,00
65723449	- 10691,34	10691,34	114304697,5	-1222070099223,87	13065563879460800,00
67043344	- 10610,34	10610,34	112579261,9	-1194503963876,92	12674090201822000,00
67947049	- 10555,34	10555,34	111415149,7	-1176024508107,19	12413335591342900,00
72437121	- 10287,34	10287,34	105829312,8	-1088701858566,91	11199843455955100,00
75933796	- 10084,34	10084,34	101693862,8	-1025515234294,13	10341641734013600,00
83795716	-9644,34	9644,34	93013245,81	-897051134599,78	8651463896838150,00
90117049	-9305,34	9305,34	86589305,99	-805742716117,54	7497707911640440,00
112678225	-8183,34	8183,34	66967012,64	-548013665790,94	4484580781779430,00
139027681	-7007,34	7007,34	49102778,84	-344079743512,07	2411082889702540,00
140422500	-6948,34	6948,34	48279394,01	-335461523904,10	2330899886350010,00
145805625	-6723,34	6723,34	45203267,14	-303916821077,53	2043335360031320,00
147841281	-6639,34	6639,34	44080802,44	-292667324662,72	1943117143657880,00
159896025	-6153,34	6153,34	37863562,39	-232987278331,25	1433649356778600,00
162052900	-6068,34	6068,34	36824720,01	-223464829387,39	1356060004102590,00
169000000	-5798,34	5798,34	33620717,76	-194944268587,37	1130352662960240,00
170615844	-5736,34	5736,34	32905567,91	-188757443183,34	1082776399736700,00
176199076	-5524,34	5524,34	30518304,81	-168593415719,89	931366928714494,00
180848704	-5350,34	5350,34	28626111,36	-153159357109,48	819454251818763,00
181063936	-5342,34	5342,34	28540569,96	-152473357189,55	814564133864628,00
183575401	-5249,34	5249,34	27555544,19	-144648351443,73	759308015546766,00
190412401	-4999,34	4999,34	24993375,44	-124950319083,30	624668815830123,00
194323600	-4858,34	4858,34	23603443,26	-114673493538,16	557122533912442,00
206123449	-4441,34	4441,34	19725478,79	-87607508650,62	389094513451591,00
212809744	-4210,34	4210,34	17726941,86	-74636408089,92	314244467846312,00
213481321	-4187,34	4187,34	17533795,34	-73419918739,93	307433978986651,00
214798336	-4142,34	4142,34	17158959,96	-71078203319,49	294429907042938,00
218951209	-4001,34	4001,34	16010701,79	-64064221469,27	256342571773286,00
250462276	-2972,34	2972,34	8834790,214	-26259978257,43	78053518123733,80
264160009	-2545,34	2545,34	6478742,989	-16490587482,53	41974110716302,00
264192516	-2544,34	2544,34	6473653,314	-16471158888,57	41908187228649,50

281769796	-2012,34	2012,34	4049502,214	-8148965161,38	16398468180431,70
303142921	-1387,34	1387,34	1924705,339	-2670215893,11	3704490641614,24
337383424	-430,34	430,34	185190,3639	-79694358,23	34295470883,73
342953361	-279,34	279,34	78029,43891	-21796548,39	6088593336,02
345142084	-220,34	220,34	48548,61391	-10697080,22	2356967912,22
398202025	1156,66	1156,66	1337868,139	1547461906,22	1789891157100,46
415181376	1577,66	1577,66	2489018,964	3926831881,14	6195215402684,92
431766841	1980,66	1980,66	3923023,939	7770186402,39	15390116825231,50
433305856	2017,66	2017,66	4070961,964	8213827293,50	16572731311571,40
454201344	2513,66	2513,66	6318499,164	15882574404,59	39923431684283,90
455907904	2553,66	2553,66	6521192,164	16652923884,26	42525947238592,20
472714564	2943,66	2943,66	8665148,914	25507273914,78	75084805700170,50
532409476	4275,66	4275,66	18281289,81	78164625308,95	334205557260032,00
556535281	4792,66	4792,66	22969613,84	110085606885,21	527603159908473,00
604323889	5784,66	5784,66	33462320,24	193568229048,99	1119726875771110,00
816187761	9770,66	9770,66	95465845,69	932764558503,38	9113727693098050,00
852640000	10401,66	10401,66	108194582,8	1125403534238,47	11706067739455800,00
927811600	11661,66	11661,66	135994372,3	1585920471241,03	18494469287453900,00
1126877761	14770,66	14770,66	218172470,7	3222551931336,98	47599226966501600,00
1129228816	14805,66	14805,66	219207642,1	3245514365819,00	48051990339217600,00
1298232961	17232,66	17232,66	296964656,8	5117491705733,19	88188007411449300,00
1433909689	19068,66	19068,66	363613889,5	6933630539929,68	132215060665612000,00
1534837329	20378,66	20378,66	415289885,3	8463052411966,33	172465688823273000,00
1552675216	20605,66	20605,66	424593327,1	8749026797230,97	180279493387197000,00
1683214729	22228,66	22228,66	494113436,5	10983480817538,50	244148088168288000,00
2075804721	26762,66	26762,66	716240104,1	19168492174696,30	512999886705287000,00
2191550596	28015,66	28015,66	784877345,3	21988858810210,40	616032447187005000,00
2214643600	28261,66	28261,66	798721567,3	22573199365483,60	637956142012511000,00
2215679041	28272,66	28272,66	799343444,8	22599567437517,80	638949942806930000,00
2575664001	31952,66	31952,66	1020972641	32622794214459,30	1042385133341570000,00
2685519684	33023,66	33023,66	1090562285	36014360832225,70	1189326097276640000,00
2723691721	33390,66	33390,66	1114936342	37228463111014,30	1243083047133580000,00

3277104516	38447,66	38447,66	1478222752	56834209457717,60	2185142503684630000,00
------------	----------	----------	------------	-------------------	------------------------

Приложение 4

Таблица 13. Вспомогательные данные для интервального ряда

$X_{cp.i} * f_i$	$(X_i - X_{cp.})^2 * f_i$	$(X_i - X_{cp.})^3 * f_i$	$(X_i - X_{cp.})^4 * f_i$	$X_i^2 * f_i$	$ X_{cp.} - X_i * f_i$
103972	4090090950	- 5,57684E+13	7,60402E+17	491371672	299970
332956	802308650	- 4,45682E+12	2,47577E+16	4263834536	144430
292404	89258750	2,25378E+11	5,6908E+14	6107149944	35350
86898	337398075	3,57811E+12	3,79458E+16	2517087468	31815
259322	2443904575	4,56644E+13	8,53239E+17	9606842812	130795
180504	2865460900	7,66941E+13	2,05272E+18	8145423504	107060
212824	4856696100	1,69232E+14	5,89687E+18	11323513744	139380