

Национальная олимпиада по анализу данных DANO.

1 этап (вариант 1)

В заданиях с 1 по 9 необходимо выбрать 1 правильный вариант ответа. Каждое задание оценивается в 2 балла.

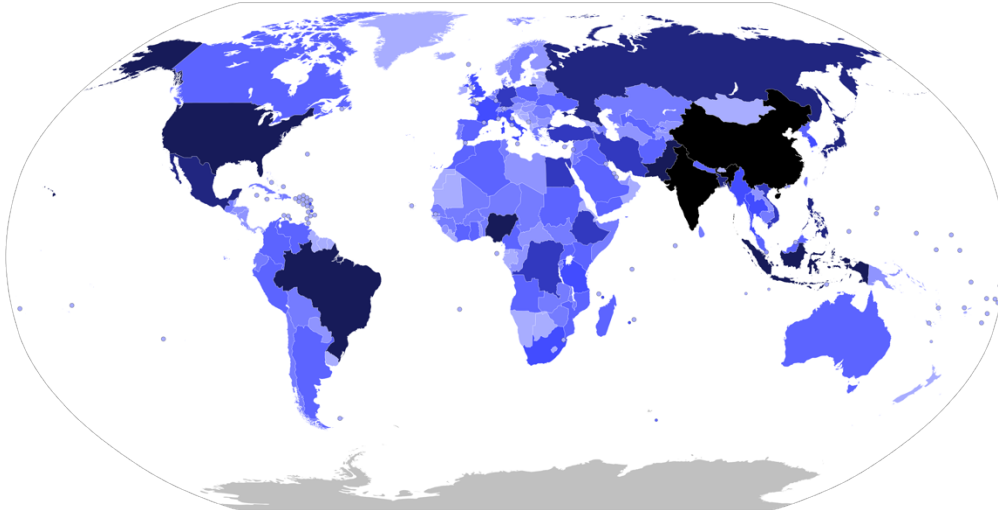
1. В стране А проживает $N > 1$ человек. Каждый из $N-1$ жителей зарабатывает 20 тыс д.е. в месяц, а один человек — 200 тыс д.е. Чему должно быть равно N , чтобы медианный ежемесячный доход в стране был равен среднему?
- 1) 80
 - 2) 800
 - 3) 120
 - 4) 1200
 - 5) Нет верного ответа

2. Работники 1-го цеха завода считают, что работа в их цеху ухудшает здоровье. Они заметили, что среди 30 заболевших пылевым бронхитом 20 было из 1-го цеха. Руководитель завода решил рассмотреть жалобы работников 1-го цеха и посмотрел на всю статистику по заводу. Какой вывод о влиянии работы в цеху 1 на здоровье можно сделать на основе данных, если известно, что распределение людей в цех случайно?

	Заболели	Не заболели
Работали в цеху 1	20	400
Не работали в цеху 1	10	200

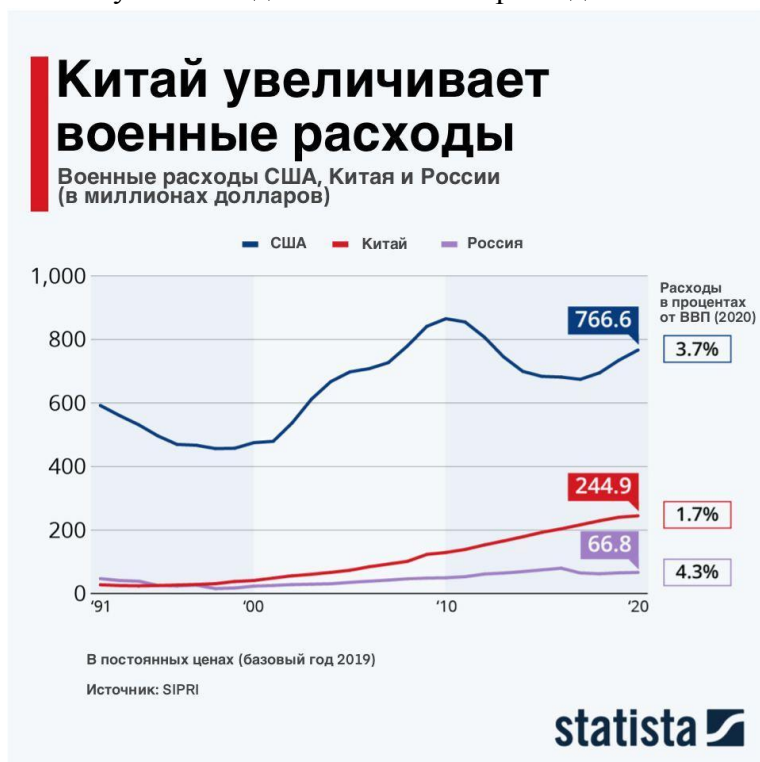
- 1) Работа в цеху 1 увеличивает вероятность заболеть пылевым бронхитом
 - 2) Работа в цеху 1 снижает вероятность заболеть пылевым бронхитом
 - 3) Работа в цеху 1 не влияет на вероятность заболеть пылевым бронхитом
 - 4) Доля заболевших среди работников в цеху 1 больше, чем в цеху 2.
3. В среднем успеваемость (или средний балл) студентов НИУ ВШЭ повысилась с переходом на дистанционный формат обучения. Какой вывод об успеваемости из этого следует?
- 1) Успеваемость всех студентов повысилась
 - 2) Повысилась успеваемость большинства студентов
 - 3) Повысилась успеваемость меньшинства студентов
 - 4) Никакой из указанных в других вариантах выводов сделать нельзя
4. Кира живет в большой семье: в ней есть 3 человека младше Киры и 7 — старше. Кира заметила, что за 5 лет средний возраст младших 3 членов семьи вырос на $X\%$, а старших 7 — на $Y\%$. За 5 лет состав семьи не менялся. Как соотносятся X и Y ?
- 1) X больше Y
 - 2) X меньше Y
 - 3) X в точности равняется Y
 - 4) Нет верных утверждений среди других вариантов

5. Перед вами карта мира, на которой цветом показан определенный показатель. Чем темнее цвет, тем выше уровень данного показателя; чем светлее — тем ниже уровень данного показателя. Например, в Нигерии и Индонезии этот показатель высокий, а в Мавритании и Парагвае — низкий. По какому показателю раскрашены страны?



- 1) Доля блондинов среди населения
 - 2) Численность населения
 - 3) Экспорт электроники
 - 4) Площадь полей, засеянных гречихой
6. Губернатор хочет провести опрос о негативных последствиях выгула собак в общественных местах в Свердловской области. Какая из выборок будет наиболее репрезентативной (при одинаковом количестве респондентов)?
- 1) Случайные посетители зоомагазинов
 - 2) Все жители случайно отобранных домов в области
 - 3) Случайно выбранные спортсмены, бегающие в парках, где выгуливают собак
 - 4) Взрослые от 18 до 45 лет, опрошенные около станций метро с наибольшим трафиком

7. На графике ниже отображены реальные (как и все остальные показатели, в ценах 2019 года) военные расходы США, Китая и России за 1991–2020 годы, также указаны доли военных расходов от ВВП стран в 2020 году.



Выберите верное утверждение, следующее из графика:

- 1) Максимальная величина военных расходов России наблюдалась в 2010 году
 - 2) В 2020 году ВВП Китая был больше ВВП России.
 - 3) На протяжении всего рассматриваемого периода Россия тратила на военные расходы больше, чем Китай
 - 4) В 2010 ВВП США был больше 20 триллионов долларов
 - 5) Нет верного ответа среди других вариантов
8. Ученик Арсений переходит в 10 класс и выбирает новую школу. Он нашел интервью учителя и ученика об одной школе. В интервью учитель утверждает, что сейчас в среднем в 10-х классах по 15 человек в классе. Ученик же, зная, что вероятность оказаться в определенном классе пропорциональна количеству учеников в нем, утверждает, что математическое ожидание количества одноклассников у поступающего будет 20. Такая ситуация
- 1) Невозможна в любом случае
 - 2) Возможна, если есть 4 класса с 10, 10, 10 и 30 учениками соответственно
 - 3) Возможна, если есть 2 класса с 15 и 20 учениками соответственно
 - 4) Возможна, если есть 3 класса с 10, 15 и 20 учениками соответственно
9. В городе М от станции А до станции Б целый день с разными интервалами ходят трамваи, других остановок в городе нет. Исследователь Вася решил измерить средний интервал между трамваями интересным способом: он целый день стоял у остановки Б и опрашивал случайные 20% пассажиров (для простоты считайте, что количество пассажиров в трамвае всегда кратно 5),

выходящих из каждого прибывшего трамвая, о том, сколько они ждали свой трамвай. Затем он посчитал среднее по всем полученным данным и умножил на 2. Будем считать, что пассажиры приходят равномерно по 1 человеку за 10 секунд, поэтому количество пассажиров в трамвае прямо пропорционально времени ожидания. Каким получился результат Васи по отношению к реальному среднему интервалу между двумя трамваями?

- 1) Результат Васи равен реальному среднему
- 2) Результат Васи завышен по отношению к реальному среднему
- 3) Результат Васи занижен по отношению к реальному среднему
- 4) Результат Васи может быть как завышен, так и занижен

В заданиях с 10 по 22 необходимо выбрать один или несколько правильных вариантов ответа. Каждое задание оценивается в 5 баллов.

10. Выберите все верные утверждения:

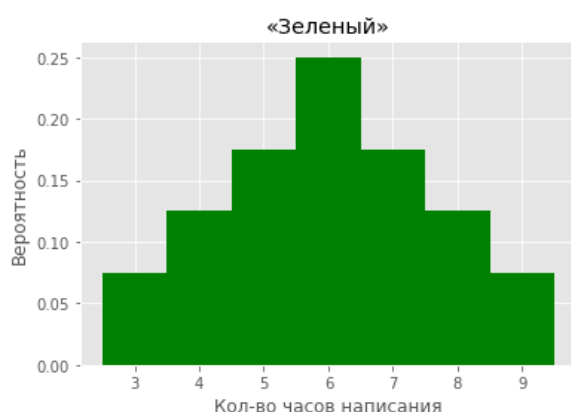
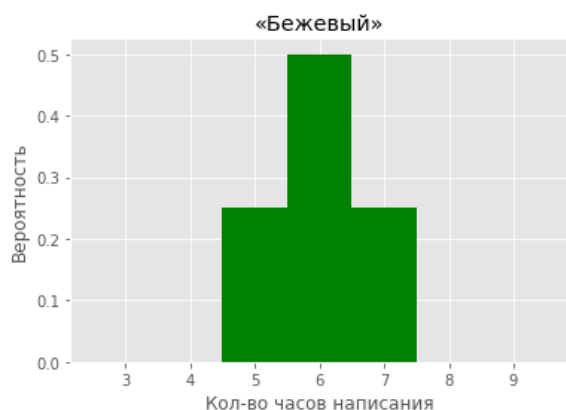
- 1) Медиана выборки не может быть выше среднего значения этой выборки.
- 2) У выборки может быть несколько медиан.
- 3) У выборки может быть несколько мод.
- 4) Для выборок работает правило «невозможной тройки», из которого следует, что мода, медиана и среднее значение выборки не могут быть одновременно равны друг другу.
- 5) Если в выборку {5; 1; 4; 3; 0; 1} добавить число 2, то медиана выборки не изменится

11. В таблице представлено количество учеников, посмотревших лекции по анализу данных по каждой теме, и количество учеников, решивших задачи из листочков. Решать листочки можно только после просмотра лекций. Какие выводы не противоречат данным в таблице?

ТЕМЫ	Видео	Листочки
Исследовательский вопрос	388	178
Анкетирование	236	156
Статистика	218	81
Визуализация	130	83

- 1) Чем больше учеников, смотревших видео по темам, тем больше учеников, решавших листочки с задачами по ним
- 2) Видео посмотрело больше учеников, чем решило листочки
- 3) Между количеством посмотревших видео и количеством решавших листки наблюдается положительная взаимосвязь (корреляция)
- 4) Нет верного ответа среди остальных вариантов

12. На графиках представлено распределение количества часов написания одной статьи в журналах «Бежевый» и «Зеленый» соответственно



Выберите все верные утверждения про эти распределения

- 1) Среднее время написания одной статьи в этих журналах
- 2) Медианное время в «Зеленый» больше, чем в «Бежевый»
- 3) Дисперсия времени написания статьи в «Зеленый» меньше, чем в «Бежевый»
- 4) Дисперсии времени написания статьи в обоих журналах совпадают
- 5) Если статья должна быть написана не более чем через 7 часов, то в «Зеленый» сделают это с меньшей вероятностью.

13. Коля скрывает от мамы свою оценку за недавно написанный тест. Мама Коли знает, что среднее за прошлые три теста равно 3, при этом мода этих оценок была равна 2. Коля подсказал маме, что при добавлении новой оценки в выборку оценок Коли за тест медиана выборки не изменится. Какую оценку мог получить Коля за тест (оценкой может быть лишь натуральное число, не превосходящее пять)?

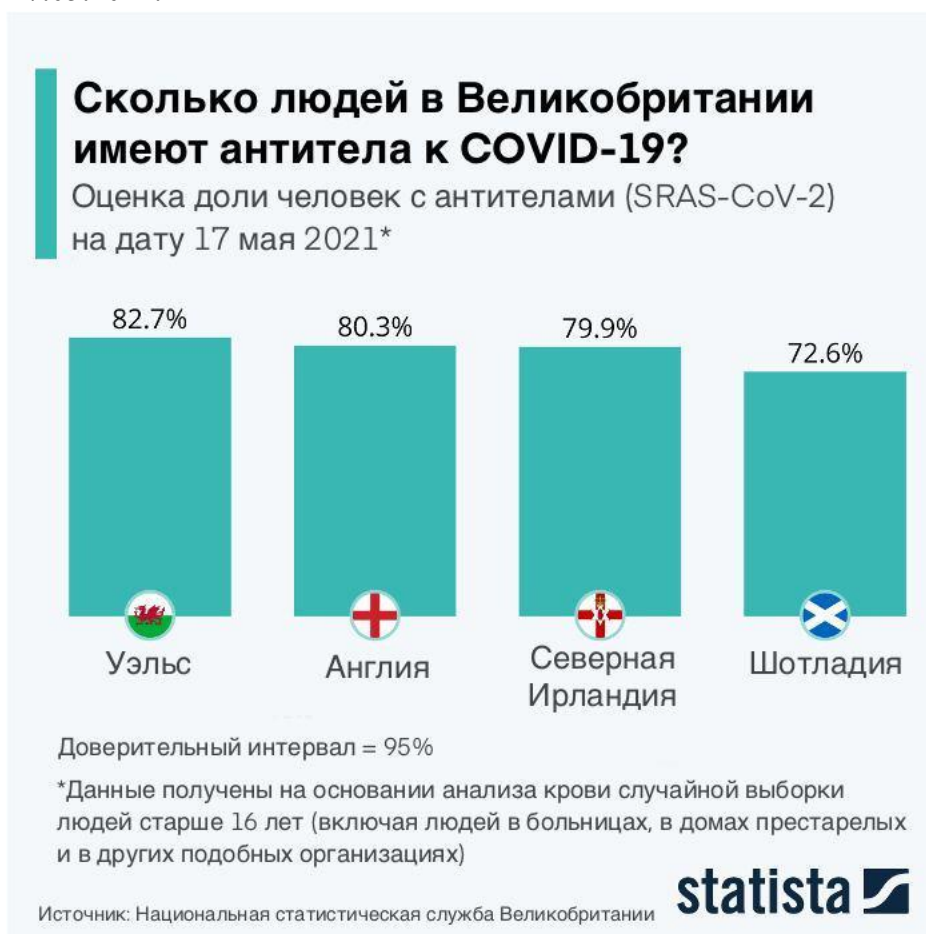
- 1) 1
- 2) 2
- 3) 3
- 4) 4
- 5) 5

14. Университет ЭШВ практикует программу обмена с иностранными учебными заведениями. За рубеж ЭШВ отправляет лишь своих самых сильных студентов (с самым высоким уровнем образования среди всех). Но иностранные вузы,

наоборот, отправляют студентов с самым низким уровнем образования. Предположим, осуществился обмен студента ЭШВ на студента из зарубежного университета. Как в таком случае могли поменяться средние уровни образования в ЭШВ и иностранном вузе, соответственно?

- 1) Увеличиться, увеличиться
- 2) Уменьшиться, уменьшиться
- 3) Увеличиться, уменьшиться
- 4) Уменьшиться, увеличиться
- 5) Не измениться, не измениться

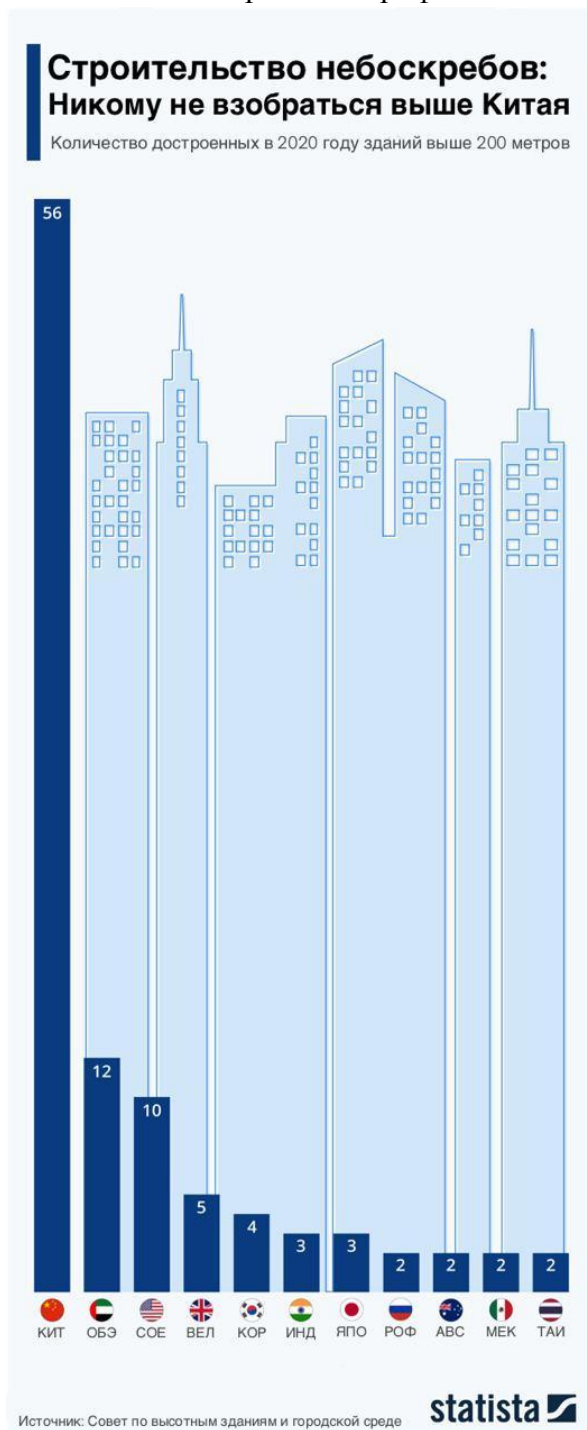
15. Перед Вами диаграмма, показывающая долю жителей, имеющих антитела к коронавирусу, для четырёх частей Соединенного королевства по состоянию на 17.05.2021.



Выберите все верные утверждения, о которых можно сделать вывод из графика:

- 1) Больше всего людей с антителами в Соединенном королевстве проживает в Уэльсе
- 2) Около 79% жителей Соединенного королевства имеют антитела
- 3) В среднем на каждые 200 человек в Северной Ирландии приходится на 14–15 жителей с антителами больше, чем в Шотландии
- 4) Доля населения с антителами в Соединенном королевстве больше 72,6% и меньше 82,7%
- 5) Нет верных утверждений среди других вариантов

16. Перед Вами диаграмма, отображающая количество зданий высотой более 200 метров (небоскребов), построенных в 2020 году в различных странах мира. Ира знает приведенную в задаче статистику по странам, а Катя и Маша знают только среднее и медианное значения. Ира предложила девушкам игру: она называет каждую из 11 приведенных стран, а Маша и Катя угадывают, сколько там небоскребов. Та, кто назовет число более близкое к настоящему, получает 1 балл. Маша решила всегда называть медиану, а Катя — среднее, округленное до целого числа по правилам арифметики.

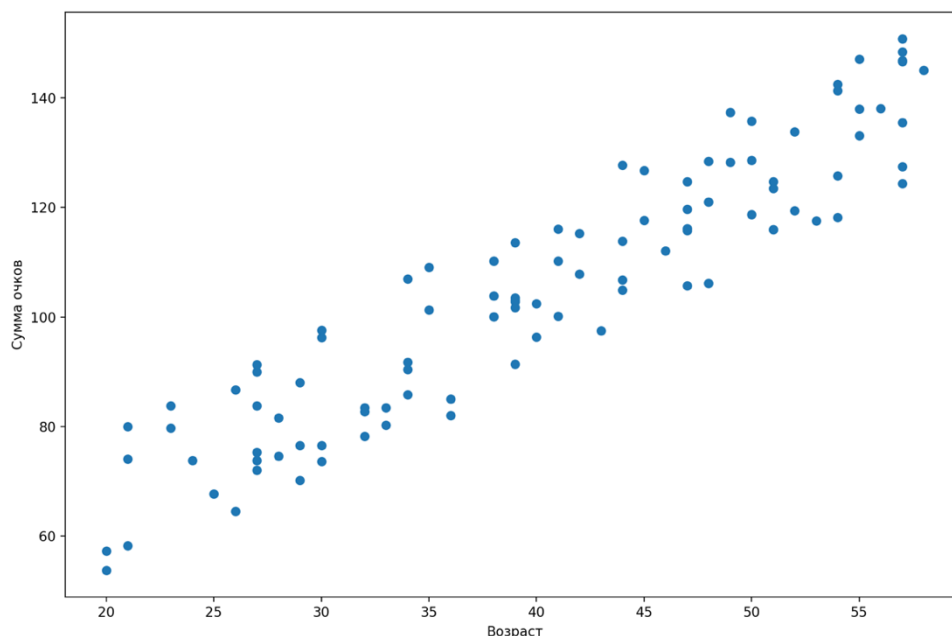


Выберите все верные утверждения, следующие из графика:

- 1) В такой игре Маша наберет больше баллов, чем Катя.

- 2) Из всех представленных в выборке стран самое высокое построенное здание находится в Китае
- 3) Всего в мире есть 101 здание выше 200 м
- 4) Если бы в России в 2020 году построили 8 небоскребов, то число построенных небоскребов в Китае было бы меньше суммарного количества построенных небоскребов во всех других странах, представленных в выборке
- 5) Медианное количество построенных в 2020 году небоскребов в странах в выборке равно 3

17. Исследователь Данил любит играть в боулинг. Но больше, чем играть в боулинг, Данил любит собирать статистику, поэтому в один день он собрал статистику, как другие играют в боулинг. Данил нарисовал диаграмму рассеяния в координатах (возраст, сумма очков), где каждая точка — показатели одного человека в этот день.



Выберите все верные утверждения

- 1) Между суммой набранных баллов и возрастом есть положительная взаимосвязь.
 - 2) Количество набранных баллов можно однозначно угадать, зная возраст
 - 3) Возраст человека можно однозначно угадать, зная количество набранных баллов
 - 4) Нет верных утверждений среди остальных вариантов
18. Спортсмен Виталий очень любит йогу. Он решил измерить влияние своих тренировок по йоге на скорость бега. Для этого Виталий семь дней подряд после каждой тренировки измерял время, за которое он пробежал 100 м. Оказалось, что с каждым днём оно сокращалось. От радости Виталий забыл о своём эксперименте и вспомнил о нем лишь через месяц. Он снова провёл забеги на 100 м после каждой из семи тренировок по йоге, которые

проводились в разные дни. Снова выяснилось, что с каждым днём время преодоления 100 м Виталием уменьшалось. Выберите все верные утверждения:

- 1) Время преодоления Виталием 100 м в самом последнем забеге больше времени самого первого забега
- 2) Время преодоления Виталием 100 м в самом последнем забеге меньше времени самого первого забега
- 3) Время преодоления Виталием 100 м в самом последнем забеге может быть как больше, так и меньше времени самого первого забега
- 4) Если Виталий продолжит заниматься йогой, то время, за которое он пробежит 100 м, уменьшится
- 5) Нет верного ответа среди остальных вариантов

19. Из всех обращений взрослых людей к врачу-терапевту города Атлетус 90% составляют жалобы на суставы после бега. Выберите выводы, которые **НЕ** следуют из этого факта.

- 1) Жители города Атлетус, которые старше 18 лет, более склонны к травмам из-за бега, чем подростки
- 2) Бегунам города Атлетус нужно составлять равномерный план тренировок на начальных этапах, чтобы избежать резких напряжений и травм впоследствии
- 3) Неправильный подбор обуви может приводить к проблемам со здоровьем при беге
- 4) Большинство жителей города Атлетус испытывает неприятные ощущения после бега
- 5) По причине того, что в городе Атлетус бег является травмоопасным для более, чем 50% бегунов, стоит проводить просветительские кампании о безопасном подходе к тренировкам

20. Перед вами четыре обнаруженные корреляции между парами признаков. Выберите все случаи, когда корреляция, скорее всего, возникает из-за прямой причинно-следственной связи между признаками:

- 1) Чем больше пожарных приехало на вызов, тем больше пострадавших от пожара
- 2) Чем больше количество рабочих дней в месяце, тем больше прибыль общественного транспорта
- 3) Чем больше аистов гнездится в деревне, тем больше в деревне будет детей
- 4) Чем больше электроэнергии выработано атомными электростанциями, тем больше смертей от утопления в бассейнах

21. В одном из начальных классов городской школы в Конго дети, которые воспитываются мусульманских семьях, учатся в среднем лучше тех, кто воспитывается в христианских. На какие генеральные совокупности можно распространить этот вывод?

- 1) Городская школа в Конго
- 2) Конго
- 3) Африка
- 4) Весь мир

5) Ни на какую из них

22. Студенты проводят исследование, пытаясь ответить на вопрос: «Как прием пищи во время просмотра онлайн лекций влияет на успеваемость?» Какие из механизмов могут описывать гипотезу, выдвигаемую в качестве ответа на него?

1) При приеме пищи увеличивается приток энергии за счет поступления микроэлементов, что позволяет концентрироваться лучше и улучшает успеваемость

2) Во время приема пищи теряется концентрация, что приводит к потере внимания, плохому запоминанию материала и уменьшает успеваемость

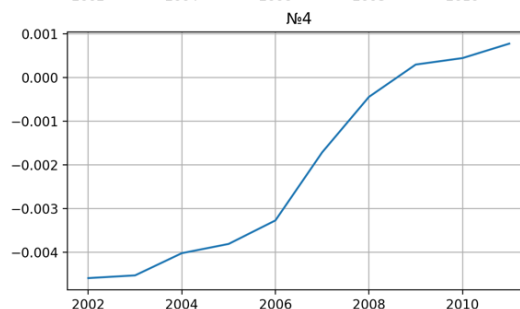
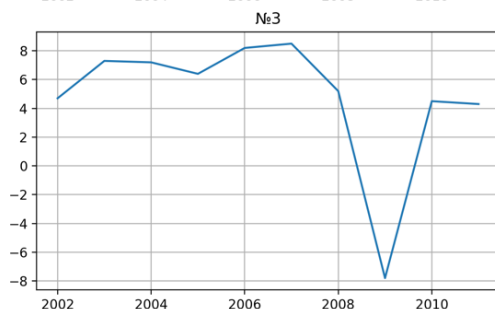
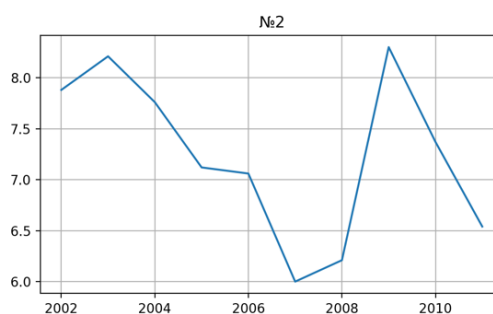
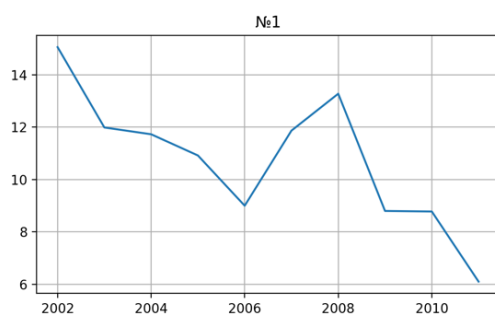
3) Прием пищи никак не влияет на умственную концентрацию, поэтому успеваемость от этого не меняется

4) Просмотр дополнительных лекций во время приема пищи повышает успеваемость

5) Нет верного ответа среди других вариантов

Открытые вопросы. Каждое задание оценивается в 5 баллов.

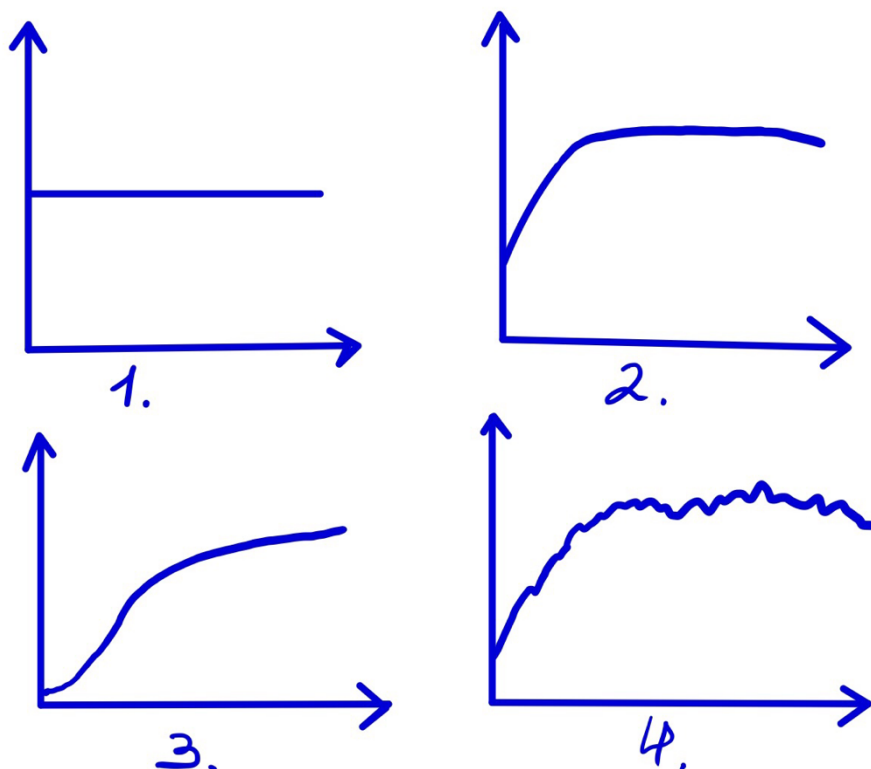
23. Перед Вами четыре графика, отражающих уровень безработицы (А), а также годовые темпы прироста ВВП (Б), населения (В) и уровня цен (Г) в России за 2002–2011 года (все показатели измерялись в %). Соотнесите показатели с соответствующими графиками их динамики и выпишите верный порядок букв, обозначающих показатели (пример ответа: АБВГ; такой ответ означает, что графику 1 ставится в соответствие показатель А, графику 2 — показатель Б и



т.д.).

Ответ: **ГАБВ**

24. Исследователь Вася провёл четыре исследования о зависимости различных показателей от возраста, на основе которых от руки нарисовал графики найденных им функций для одного человека. Но забыл подписать графики и



теперь не знает, какой рисунок относится к какому исследованию. Взгляните на графики, а затем соотнесите исследования с номерами соответствующих визуализаций. Названия исследований представлены в виде пары «возраст — признак зависимый», на графиках возраст всегда отображаются по оси абсцисс.

- «Возраст — рост»: 2
- «Возраст — уровень образования»: 3
- «Возраст — размер глаза»: 1
- «Возраст — вес»: 4

25. Работодатель занимается подбором кандидатов на вакансию, где коммуникативные навыки в два раза важнее, чем каждая из характеристик: опыт, дружелюбие и знание предметной области. Остальные навыки для вакансии не важны. У работодателя есть 3 кандидата, чьи способности отражены на графиках ниже. Работодатель ранжирует кандидатов, используя сумму всех баллов с учетом важности каждой характеристики, от большей суммы к меньшей. Приведите в ответе это ранжирование кандидатов (сначала самый привлекательный кандидат).

