

Национальная олимпиада по анализу данных DANO.

1 этап (вариант 2)

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 50 тыс д.е. в месяц, а один человек — 250 тыс д.е. Чему должно быть равно N , чтобы медианный ежемесячный доход в стране был равен среднему?

- 1) 150
- 2) 1500
- 3) 200
- 4) 2000
- 5) Нет верного ответа

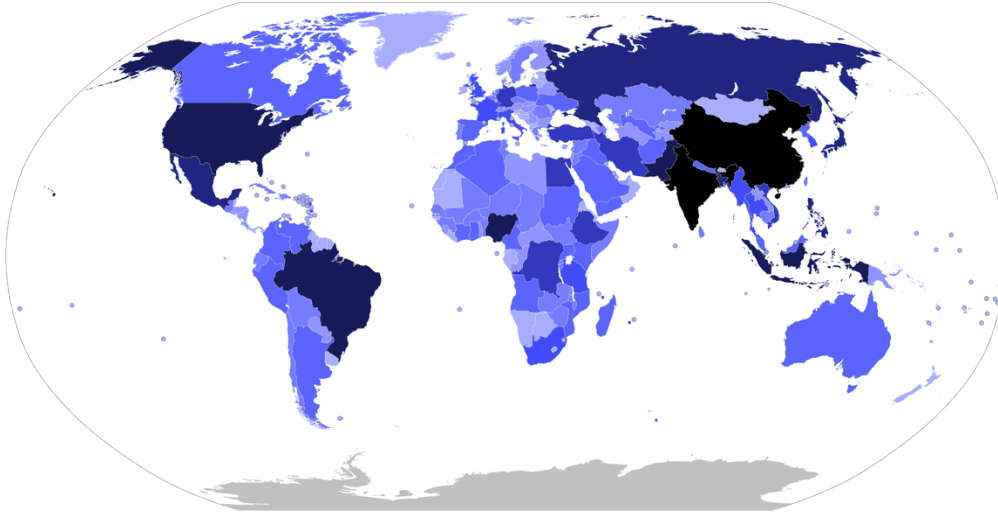
2. Машинисты Красной ветки метро города Н считают, что работа на этой ветке ухудшает здоровье. Они заметили, что среди 50 заболевших вибрационной болезнью 40 работали на Красной ветке. Руководитель метро решил рассмотреть жалобы работников Красной ветки и посмотрел на всю статистику по метро города Н. Какой вывод о влиянии работы на красной ветке на здоровье можно сделать на основе данных, если известно, что распределение людей на определенную ветку случайно?

	Заболели	Не заболели
Красная ветка	40	400
Синяя ветка	10	100

- 1) Работа на Красной ветке увеличивает вероятность заболеть вибрационной болезнью
 - 2) Работа на Красной ветке снижает вероятность заболеть вибрационной болезнью
 - 3) Работа на Красной ветке не влияет на вероятность заболеть вибрационной болезнью
 - 4) Доля заболевших среди машинистов на Красной ветке больше, чем среди машинистов на Синей ветке
3. В городе М. среднее время прогулок в парках во время самоизоляции увеличилось. Какой вывод о времени прогулок в парках из этого следует?
- 1) Все люди города М. стали больше гулять
 - 2) Большинство людей стали предпочитать гулять в парках, а не по улицам
 - 3) Большинство людей стали больше гулять
 - 4) Никакой из указанных в других вариантах выводов сделать нельзя
4. У Андрея 20 коллег: 11 человек старше Андрея как минимум на 1 год и 9 — младше как минимум на 2 года. Андрей заметил, что за 3 года средний возраст младших 9 коллег вырос на $X\%$, а 11 старших — на $Y\%$. Коллектив за 3 года не изменялся. Как соотносятся X и Y ?

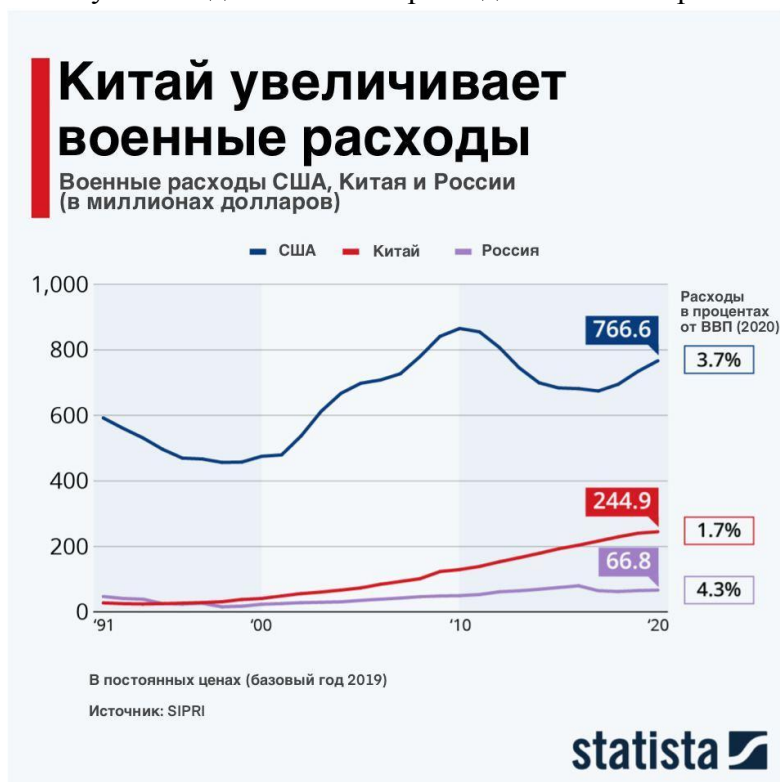
- 1) X больше Y

- 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) Средний рост взрослого населения (от 18 лет)
 - 2) Численность населения
 - 3) ВВП на душу населения
 - 4) Объем добываемого золота в год
6. Мэр одного крупного города хочет провести опрос о последствиях пробок и выхлопных газов в городе. Какая из выборок будет наиболее репрезентативной (при одинаковом количестве респондентов)?
- 1) Случайные посетители главной площади города
 - 2) Все жители случайно отобранных домов в городе
 - 3) Случайно выбранные выходящие из метро
 - 4) Случайно выбранные люди на парковке у крупного торгового центра

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 Китая наблюдалась в 2010 году
 - 2) В 2020 году ВВП США был больше ВВП Китая.
 - 3) На протяжении всего рассматриваемого периода Китай тратил на военные расходы больше, чем Россия
 - 4) В 2010 ВВП России 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 18 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 22. Такая ситуация
- 1) Невозможна в любом случае
 - 2) Возможна, если есть 4 класса с 10, 12, 20 и 30 учениками соответственно
 - 3) Возможна, если есть 2 класса с 15 и 21 учениками соответственно
 - 4) Возможна, если есть 3 класса с 6, 24 и 24 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 случайную треть пассажиров (для

простоты считайте, что количество пассажиров в трамвае всегда кратно 3), 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 5 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:

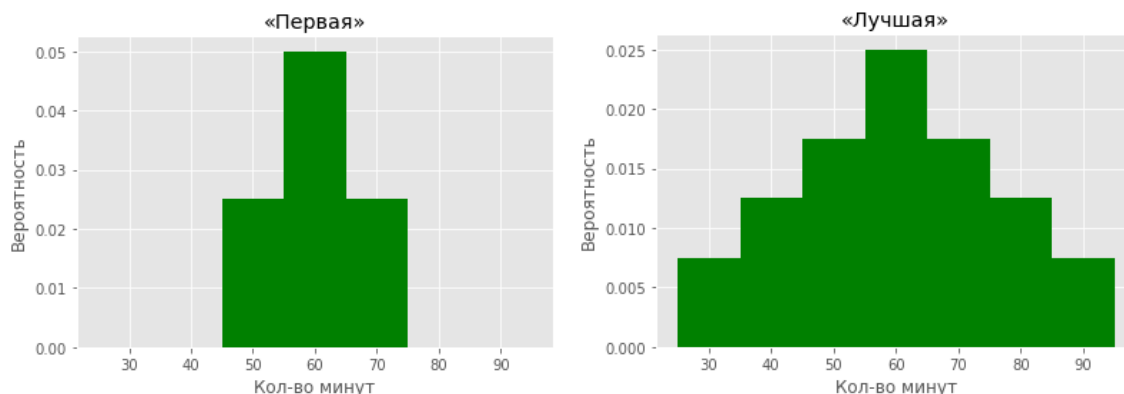
- 1) Медиана выборки не может быть меньше среднего значения этой выборки.
- 2) У выборки может быть несколько средних значений.
- 3) У выборки может быть только одна медиана.
- 4) Для выборок работает правило «невозможной тройки», из которого следует, что мода, медиана и среднее значение выборки не могут быть одновременно равны друг другу.
- 5) Если в выборку {6; 18; 4; 20; 0; -1} добавить число 5, 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) По каждой теме листки решают от 37% до 67% от смотрящих видео по соответствующей теме
- 2) По теме «Анкетирование» доля учеников, решавших листочки после просмотра видео по этой теме, выше, чем аналогичная доля по теме «Визуализация»
- 3) Между количеством посмотревших видео и количеством решавших листки наблюдается отрицательная взаимосвязь (корреляция)
- 4) Нет верного ответа среди остальных вариантов

12. На графиках представлено распределение количества минут, за которые спортсменки из спортивных школ «Первая» и «Лучшая» пробегают полумарафон.



Выберите все верные утверждения про эти распределения

- 1) В среднем спортсменки из «Первой» спортивной школы пробегают марафон быстрее, чем из «Лучшей»
- 2) Дисперсия скорости в спортивной школе «Первая» меньше, чем в «Лучшей»
- 3) Если норматив на звание Кандидат в мастера спорта составляет 1 час 21 минута, то в «Первой» спортивной школе все спортсменки имеют как минимум этот разряд, а в «Лучшей» — нет
- 4) Медианное время в «Лучшей» больше, чем в «Первой»
- 5) Если соревнуются спортсмены из обеих школ, то первое место гарантированно займет спортсменка из «Лучшей» школы

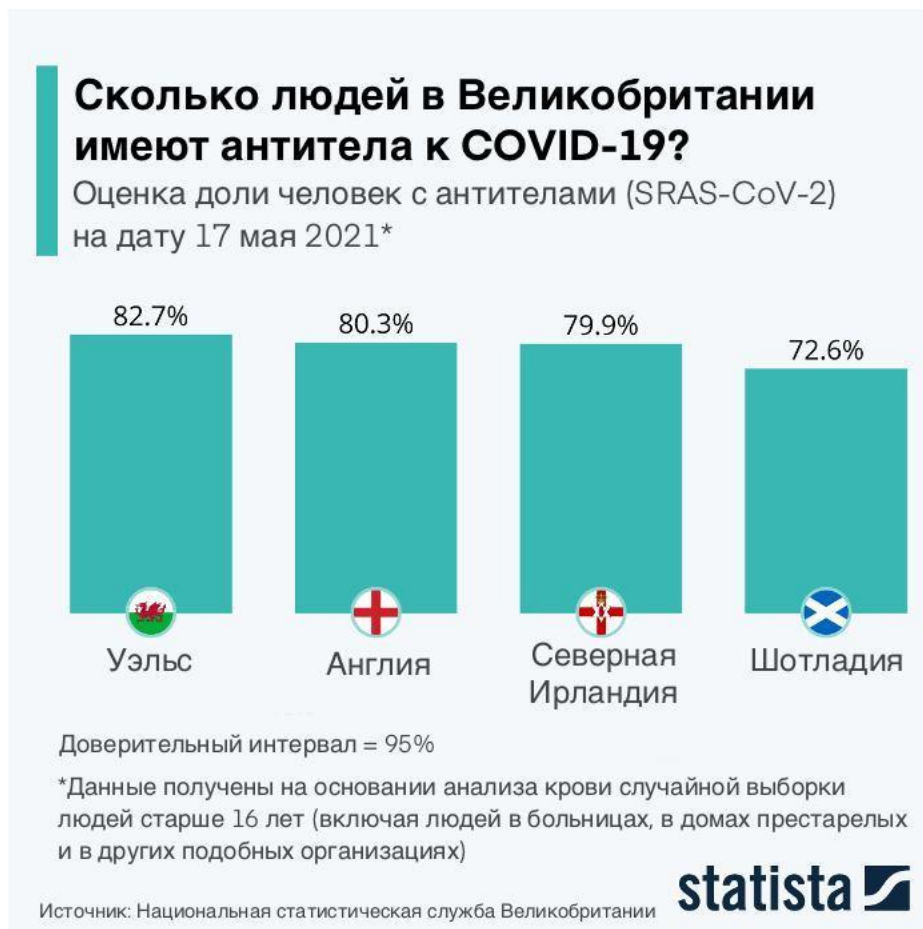
13. Коля скрывает от мамы свою оценку за недавно написанный тест. Оценка может составлять любое целое число от 1 до 100. Мама Коли знает, что среднее за прошлые три теста равно 50, при этом мода этих оценок была равна 30. Коля подсказал маме, что при добавлении новой оценки в выборку оценок Коли за тест медиана выборки не изменится. Какую оценку из предложенных мог получить Коля за тест?

- 1) 12
- 2) 48
- 3) 29
- 4) 94
- 5) 50

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

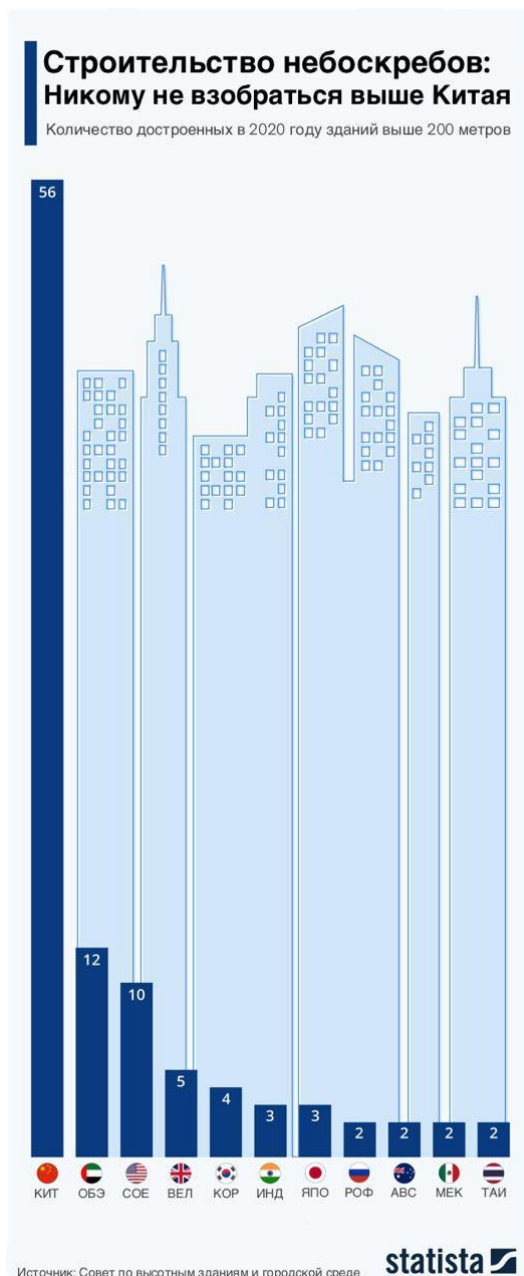
15. Перед Вами диаграмма, показывающая долю жителей, имеющих антитела к коронавирусу, для четырёх частей Соединенного королевства по состоянию на 17.05.2021.



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

- 1) Меньше всего людей с антителами в Шотландии проживает в Уэльсе
- 2) Около 76% жителей Соединенного королевства имеют антитела
- 3) В среднем на каждые 300 человек в Уэльсе приходится на 8–9 жителей с антителами больше, чем в Северной Ирландии
- 4) Доля населения с антителами в Соединенном королевстве может составлять 82,7% и больше
- 5) Нет верных утверждений среди других вариантов

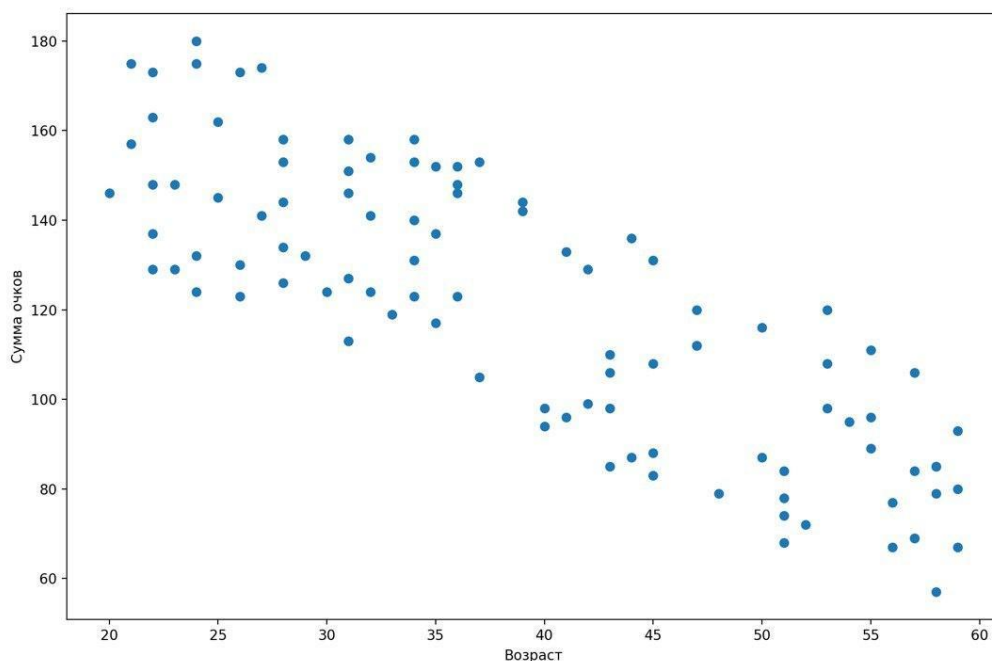
16. Перед Вами диаграмма, отображающая количество зданий высотой более 200 метров (небоскребов), построенных в 2020 году в различных странах мира. Ира знает приведенную в задаче статистику по странам, а Катя и Маша знают только среднее и медианное значения. Ира предложила девушкам игру: она называет 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 низкое построенное здание находится в Тайланде
- 3) Всего в Китае и ОАЭ есть 68 зданий выше 200 м
- 4) Если бы в Японии в 2020 году построили 15 небоскребов, то число построенных небоскребов в Китае было бы меньше суммарного количества построенных небоскребов во всех других странах, представленных в выборке
- 5) Мода построенных в 2020 году небоскребов в странах в выборке равна 2

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

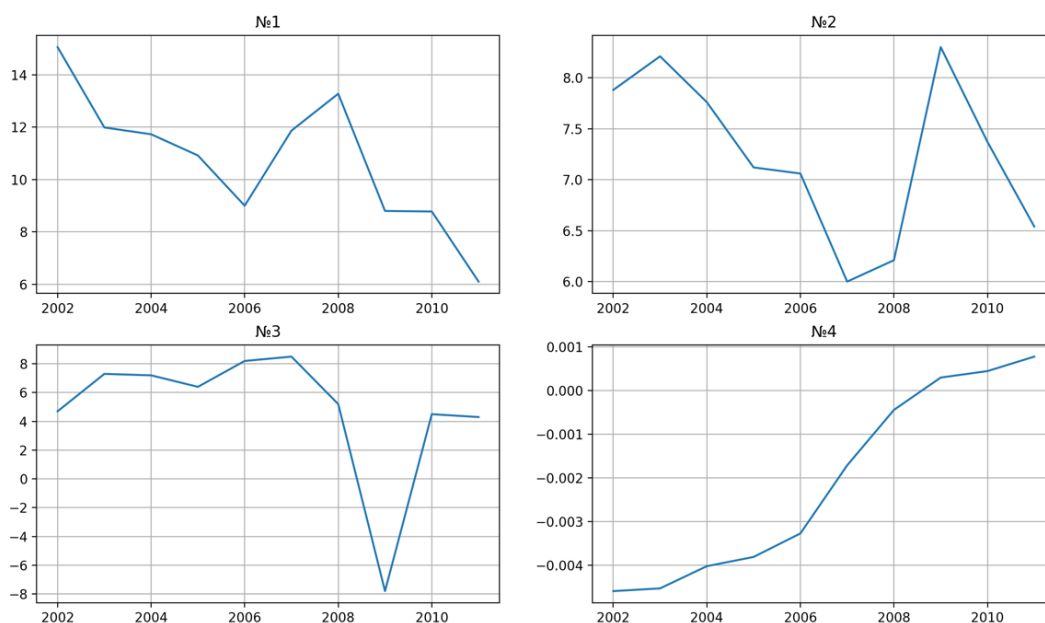
- 1) Между суммой набранных баллов и возрастом есть отрицательная взаимосвязь.
 - 2) Количество набранных баллов можно однозначно угадать, зная возраст
 - 3) Возраст человека можно однозначно угадать, зная количество набранных баллов
 - 4) Нет верных утверждений среди остальных вариантов
18. Отличница Карина очень любит математику. Она решила измерить влияние подготовки к ЕГЭ по математике на результаты тестовой части ЕГЭ по русскому языку. Для этого Карина 2 недели подряд после 2 часов подготовки к ЕГЭ по математике решала тестовую часть ЕГЭ по русскому языку. Оказалось, что с каждым днем Карина допускала все меньше ошибок в ЕГЭ по русскому языку. Получив такие результаты, Карина решила готовиться только к ЕГЭ по математике, а результаты по русскому языку улучшатся сами собой. Через месяц Карина снова решила повторить свое исследование и еще 2 недели каждый день после математики решала ЕГЭ по русскому языку. Снова выяснилось, что результаты с каждым днем улучшались. Выберите все верные утверждения:
- 1) Количество ошибок в самом последнем тесте меньше, чем в самом первом

- 2) Количество ошибок в самом последнем тесте больше, чем в самом первом
 - 3) Количество ошибок в первом тесте может быть как больше, так и меньше, чем в последнем тесте
 - 4) Если Карина продолжит заниматься математикой, ее результаты по русскому языку улучшатся
 - 5) Нет верного ответа среди остальных вариантов
19. Из всех опозданий студентов на пары 80% составляют опоздания на пары по физкультуре. Выберите выводы, которые **НЕ** следуют из этого факта.
- 1) Студенты чаще опаздывают на физкультуру, чем школьники
 - 2) Пары по физкультуре следует перенести на более позднее время для уменьшения количества опозданий
 - 3) Большинство студентов опаздывает на пары по физкультуре
 - 4) До спортивного зала дольше добираться, чем до учебных аудиторий, поэтому студенты чаще опаздывают
 - 5) Так как большинство студентов опаздывает на первую часть пары, самые важные объявления преподаватель должен делать в конце пары
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 полицейских в городе, тем больше преступлений в этом городе
 - 2) Чем выше человеческий капитал в стране, тем выше ВВП на душу населения в этой стране.
 - 3) Чем больше домашнего задания задают школьникам, тем выше средний балл в школе
 - 4) Чем больше средний рост население страны, тем меньше детей приходится на 1 взрослого
21. В случайно выбранных 20% школ 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) 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 физические упражнения во время просмотра онлайн лекций влияют на успеваемость?» Какие из механизмов могут описывать гипотезу, выдвигаемую в качестве ответа на него?
- 1) При физических упражнениях улучшается кровообращение, поэтому улучшается мозговая деятельность и успеваемость повышается

- 2) Во время физических упражнений теряется концентрация, поэтому успеваемость снижается
- 3) Физические упражнения не влияют на умственную концентрацию, поэтому успеваемость не меняется
- 4) Просмотр дополнительных лекций во время физических упражнений повышает успеваемость
- 5) Нет верного ответа среди других вариантов

Открытые вопросы. Каждое задание оценивается в 5 баллов.

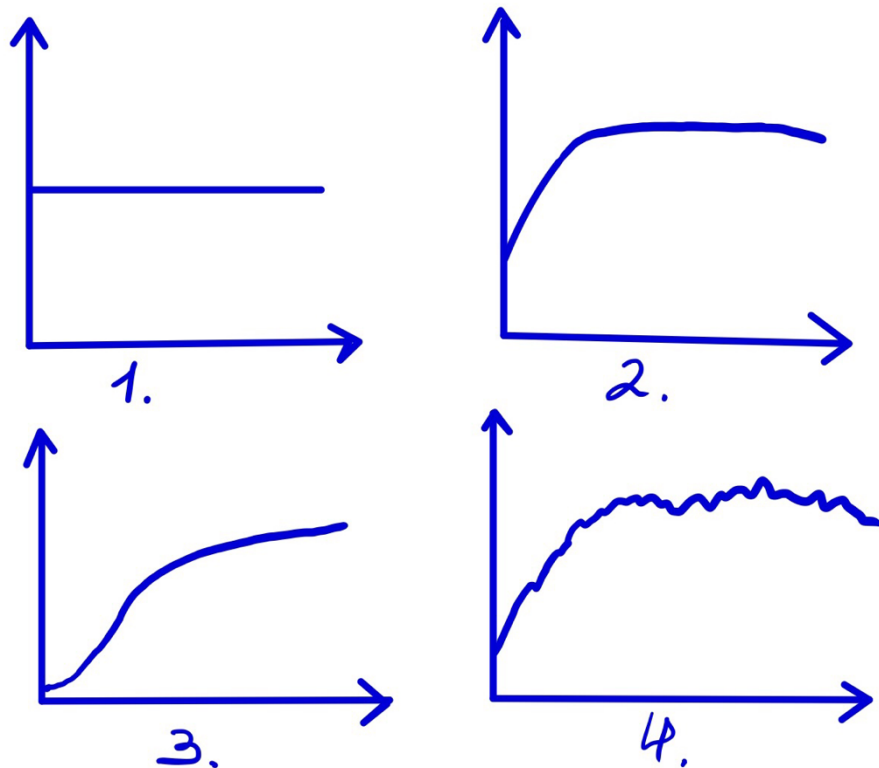
23. Перед Вами четыре графика, отражающих уровень безработицы (А), а также годовые темпы прироста ВВП (Б), населения (В) и уровня цен (Г) в России за 2002–2011 года (все показатели измерялись в %). Соотнесите показатели с соответствующими графиками их динамики и выпишите верный порядок букв, обозначающих 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 отображается по оси абсцисс.

- «Возраст — длина руки»: 2
- «Возраст — словарный запас»: 3
- «Возраст — количество хромосом»: 1
- «Возраст — длина волос на голове»: 4

25. Работодатель занимается подбором кандидатов на вакансию, где опыт в три раза важнее, чем каждая из характеристик: коммуникативные навыки, знание предметной области и образование. 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).

