## Национальная олимпиада по анализу данных DANO. 1 этап (вариант 3)

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

- 1. В стране A проживает N > 1 человек. Каждый из N-1 жителей зарабатывает 80 тыс д.е. в месяц, а один человек— 120 тыс д.е. Чему должно быть равно N, чтобы медианный ежемесячный доход в стране был равен среднему?
  - 1) 80
  - 2) 120
  - 3) 200
  - 4) 280

## 5) Нет верного ответа

2. Библиотекари северного читального зала Величайшей библиотеки считают, что именно в этом зале чаще развивается аллергия на пыль. Они заметили, что среди 40 получивших аллергию на пыль 30 работали в северном зале. Руководитель библиотеки решил рассмотреть жалобы библиотекарей северного зала и посмотрел на всю статистику: северный зал и южный зал. Какой вывод о влиянии работы в северном зале на заболеваемость аллергией можно сделать на основе данных, если известно, что распределение людей в определенный зал случайно?

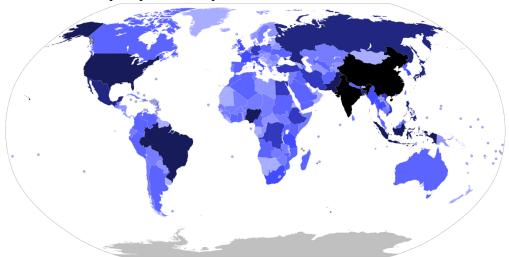
	Заболели	Не заболели
Северный зал	30	45
Южный зал	10	15

- 1) Работа в северном зале увеличивает вероятность заболеть аллергией на пыль
- 2) Работа в северном зале снижает вероятность заболеть аллергией на пыль
- 3) Работа в северном зале не влияет на вероятность заболеть аллергией на пыль
- 4) Доля заболевших среди библиотекарей в северном зале больше, чем среди библиотекарей южного зала
- 3. Средняя производительность (кол-во решенных задач) разработчиков одной крупной IT компании повысилась с переходом на дистанционный формат работы. Какой вывод о производительности из этого следует?
  - 1) Хотя бы один разработчик компании стал производительнее
  - 2) Все разработчики компании стали производительнее
  - 3) Хотя бы один разработчик компании стал менее производительным.
  - 4) Никакой из указанных в других вариантах выводов сделать нельзя

4. В местный шахматный клуб ходят 16 человек. Преподаватель в этом клубе заметила, что за 2 года средний возраст 8 самых младших участников вырос на X%, а старших 8 — на Y%. Коллектив за 2 года не изменялся. Разница в возрасте между самым старшим и самым младшим составляет 30 лет. Как соотносятся X и Y?

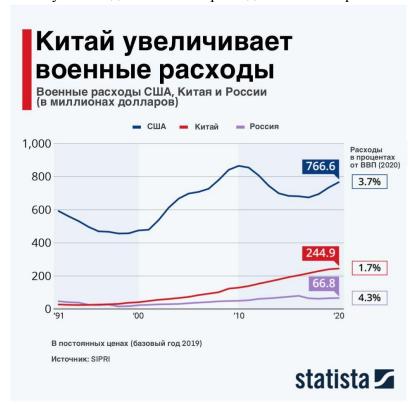
## 1) Х больше Ү

- 2) Х меньше Ү
- 3) Х в точности равняется Ү
- 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) Количество змей на 1 квадратный километр
- 2) Добыча алмазов в денежном выражении
- 3) Численность населения
- 4) Число ІТ компаний
- 6. Работодатель хочет провести опрос и выяснить, как наличие фруктов и печенек на кухне в офисе влияет на продуктивность работников. Какая из выборок будет наиболее репрезентативной (при одинаковом количестве респондентов)?
  - 1) Случайные работники, находящиеся на кухне офиса
  - 2) Случайно выбранные х% сотрудников офиса
  - 3) Работники, чьи рабочие места находятся ближе всего к кухне офиса
  - 4) Работники, придерживающиеся строгой диеты

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:

- Минимальная величина военных расходов Китая наблюдалась в 1999 году
- 2) В 2020 году ВВП США был больше ВВП России.
- 3) На протяжении всего рассматриваемого периода США тратил на военные расходы меньше, чем Россия
- 4) В 2010 ВВП Китая 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 21. 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 9, 18 и 27 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 25% пассажиров (для

простоты считайте, что количество пассажиров в трамвае всегда кратно 4), 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 15 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

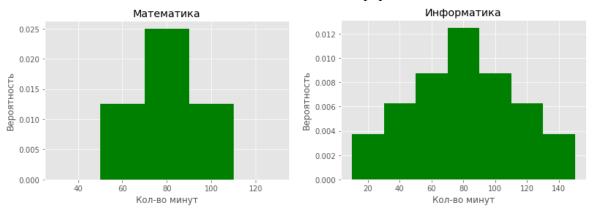
В заданиях с 10 по 22 необходимо выбрать один или несколько правильных вариантов ответа. Каждое задание оценивается в 5 баллов.

- 10. Выберите все верные утверждения:
  - 1) Медиана выборки не может быть равна среднему значению этой выборки.
  - 2) У выборки может быть несколько мод.
  - 3) У выборки может быть только одно срежнее значение.
  - 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 {16; 1; 5; 11; 0; 3} добавить число 4, 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) Число уникальных учеников, смотревших видео, от 388 до 972
- 2) По теме «Статистика» доля учеников, решавших листочки после просмотра видео по этой теме, ниже, чем аналогичная доля по теме «Визуализация»
- 3) Между количеством посмотревших видео и количеством решавших листки наблюдается положительная взаимосвязь (корреляция)
- 4) Нет верного ответа среди остальных вариантов

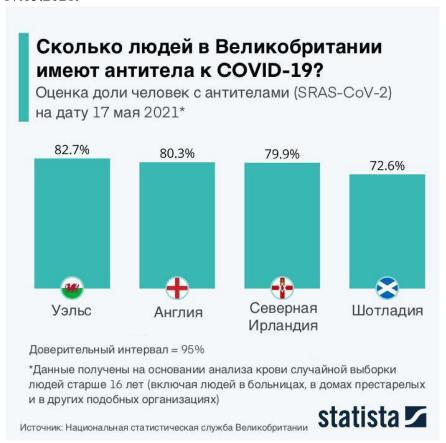
12. На графиках представлено распределение количества минут, за которые Рита делает домашнее задание по математике и информатике.



Выберите все верные утверждения про эти распределения

- 1) В среднем домашнее задание по математике Рита делает быстрее, чем по информатике
- 2) Дисперсия времени на домашнее задание по информатике больше, чем по математике
- 3) Медианное время на домашнее задание по математике меньше, чем на домашнее задание по информатике
- 4) Если на выполнение домашнего задания есть 40 минут, то домашнее задание по информатике будет выполнено с большей вероятностью, чем по математике
- Если домашнее задание необходимо выполнить за 120 минут, то домашнее задание по математике будет выполнено с большей вероятностью, чем по информатике
- 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 35. 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) 34
  - 2) 48
  - 3) 9
  - 4) 22
  - 5) 57
- 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

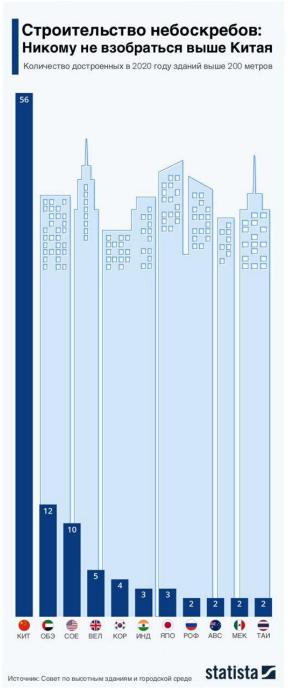
- 4) Уменьшиться, увеличиться
- 5) Не измениться, не измениться
- 15. Перед Вами диаграмма, показывающая долю жителей, имеющих антитела к коронавирусу, для четырёх частей Соединенного королевства по состоянию на 17.05.2021.



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

- 1) В Англии проживает меньше людей с антителами, чем в Шотландии
- 2) Около 78% жителей Соединенного королевства имеют антитела
- 3) В среднем на каждые 500 человек в Англии приходится на 38–39 жителей с антителами больше, чем в Шотландии
- 4) Доля населения с антителами в Соединенном королевстве может составлять 72,6% и меньше
- 5) Нет верных утверждений среди других вариантов

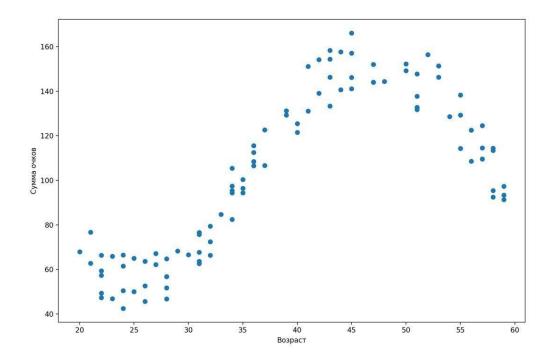
16. Перед Вами диаграмма, отображающая количество зданий высотой более 200 метров (небоскребов), построенных в 2020 году в различных странах мира. Ира знает приведенную в задаче статистику по странам, а Катя и Маша знают только среднее и медианное значения. Ира предложила девушкам игру: она называет 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 США есть 66 зданий выше 200 м
- 4) Если бы в США в 2020 году построили 25 небоскребов, то число построенных небоскребов в Китае было бы меньше суммарного количества построенных небоскребов во всех других странах, представленных в выборке
- 5) Округленное до целого числа среднее количество построенных в 2020 году небоскребов в странах равно 9
- 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) С увеличением возраста положительная взаимосвязь между суммой набранных баллов и возрастом сменяется отрицательной
- 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. Мальчик Вова очень любит играть в компьютерные игры. Однако его мама считает, что это снижает его способности в целом, в том числе к программированию. Тогда Вова решил провести эксперимент: в течение недели каждый день после 2 часов игры Вова решал 10 задачек по программированию. Оказалось, что с каждым днем Вова решал задачки по программированию все

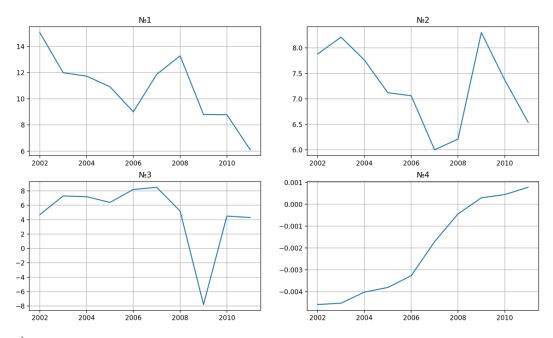
быстрее. Следующие 2 недели Вова только играл в игры, а затем еще в течение недели повторял эксперимент. Снова выяснилось, что с каждым днем Вова решал задачки все быстрее:

- 1) Время решения задачек Вовой в самый последний день эксперимента больше, чем в самый первый день
- 2) Время решения задачек Вовой в самый последний день эксперимента меньше, чем в самый первый день
- Время решения задачек Вовой в самый последний день эксперимента может быть как больше, так и меньше, чем в самый первый день
- 4) Если Вова продолжит играть в компьютерные игры, то он будет решать задачи по программированию все быстрее
- 5) Нет верного ответа среди остальных вариантов
- 19. Из всех обращений в службу поддержки маркетплейса М 90% обращений составляют жалобы на доставку. Выберите выводы, которые **HE** следуют из этого факта.
  - 1) Большая часть обращений приходится на жалобы на доставку
  - 2) Большинство пользователей маркетплейса М жалуются на доставку
  - 3) Доставка в маркетплейсе М работает хуже, чем в других маркетплейсах
  - 4) Маркетплейсу стоит снизить время доставки, чтобы снизить число обращений в службу поддержки
  - 5) Не все пользователи обращаются в службу поддержки
- 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 потребление органической еды, тем с большей вероятностью развивается аутизм
  - 2) Чем выше импорт лимонов из Мексики, тем больше количество смертей в автомобильных катастрофах
  - 3) Чем выше потребление маргарина, тем больше разводов на 10000 человек
  - 4) Чем выше образование, тем выше заработная плата сотрудника
- 21. В одной из 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) 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 прогулка в парке во время просмотра онлайн лекций влияет на успеваемость?» Какие из механизмов могут описывать гипотезу, выдвигаемую в качестве ответа на него?
  - 1) При прогулке в парке на свежем воздухе органы человека в том числе мозг насыщаются кислородом, поэтому успеваемость увеличивается

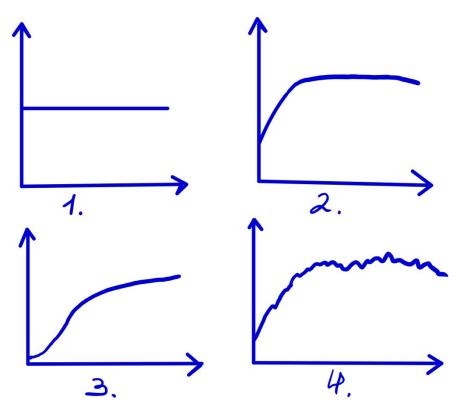
- 2) Во время прогулки теряется концентрация, внимание рассеивается, поэтому успеваемость снижается
- 3) Прогулка в парке никак не влияет на концентрацию, поэтому успеваемость не изменяется
- 4) Просмотр дополнительных лекций по время прогулок в парке повышает успеваемость
- 5) Нет верного ответа среди других вариантов

Открытые вопросы. Каждое задание оценивается в 5 баллов.

23. Перед Вами четыре графика, отражающих уровень безработицы (A), 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: <mark>ГАБВ</mark> 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 длина позвоночника»: 2
- «Возраст уровень IQ»: 3
- «Возраст группа крови»: 1
- «Возраст число подписчиков в социальных сетях»: 4

25. Работодатель занимается подбором кандидатов на вакансию, где дружелюбие в два раза важнее, чем каждая из характеристик: опыт, коммуникативные навыки 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).

