

3101953374672

Титульный лист

Направление ☒ анализ данных ☐ информатика ☐ история
☐ математика ☐ обществознание ☐ русский язык
☐ физика ☐ химия

Класс ☐ 8 ☐ 9 ☐ 10 ☒ 11

Фамилия

БУШУЕВ

Имя

АЛЕКСАНДР

Отчество

СЕРГЕЕВИЧ

Дата рождения

30 01 2008

Город участия

ЕКАТЕРИНБУРГ

Аудитория

ИЗ 29

Телефон

+7 922 189 5598

Дата

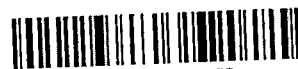
01 02 2025

Подпись

С.С.М.

Пример
заполнения

А Б В Г Д Е Ж З И Й К Л М Н О П Р С Т У Ф
Х Ц Ч Ш Щ Ъ Ы Ь Э Ю Я 1 2 3 4 5 6 7 8 9 0



3101953374672

Проверочный лист

Заполняется участниками

Направление ☒ анализ данных ☐ информатика ☐ история
☐ математика ☐ обществознание ☐ русский язык
☐ физика ☐ химия

Класс ☐ 8 ☐ 9 ☐ 10 ☒ 11

Город участия ЕКАТЕРИНБУРГ

Заполняется организаторами

Количество доп. листов _____ Количество черновиков к проверке _____

Время выхода с _____ до _____

Протокол проверки

Заполняется жюри

Номер задания	1	2	3	4	5	6	7	8	9	10
Балл члена жюри №1	9	15	14	6	5	25				
Балл члена жюри №2	9	15	14	6	5	25				

Итоговый балл

74

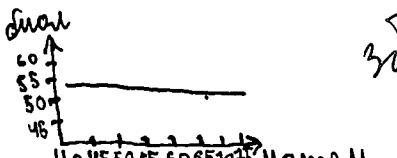
Подпись
члена жюри №1

Подпись
члена жюри №2

Пример
заполнения

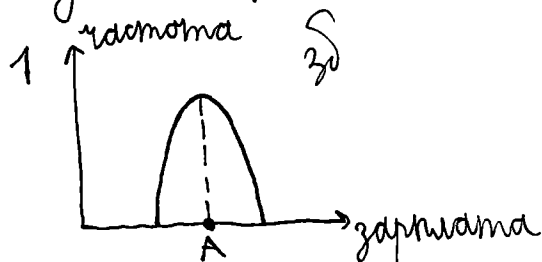
А Б В Г Д Е Ж З И Й К Л М Н О П Р С Т У Ф
Х Ц Ч Ш Щ Ъ Ы Ь Э Ю Я 1 2 3 4 5 6 7 8 9 0

Задача 2

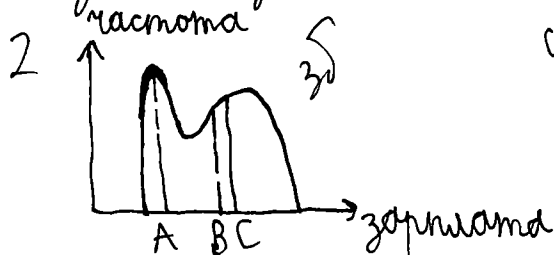
1. Количество баллов ЕГЭ по математике и по биологии
2. 1199
3. Очень слабая ~~отрицательная~~ обратная взаимосвязь
4. 

5. Гистограмма - обычно используется для графиков распределения. Круговая диаграмма - у абитуриентов сдавших ЕГЭ не может быть сильной разницы в возрасте. Скорее всего, будет только 3 различных возраста 16, 17 и 18, и на круговой диаграмме будет хорошо видна их доля от общего количества абитуриентов.

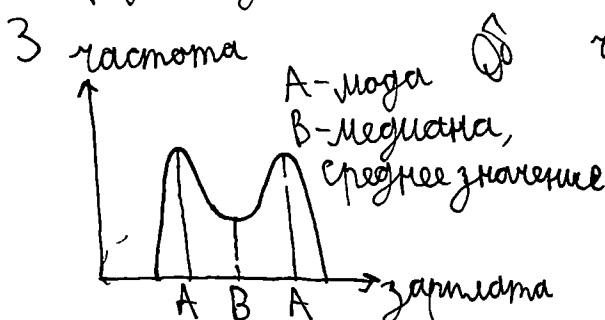
Задача 4



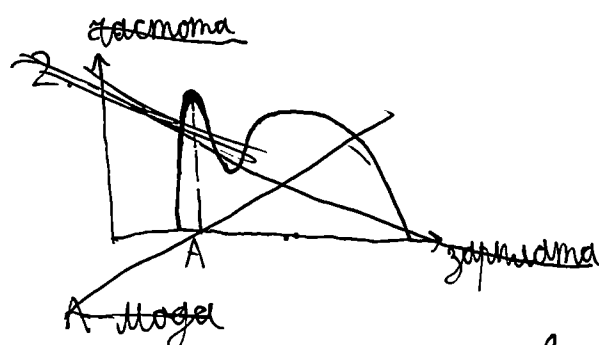
A - среднее значение, мода и медиана



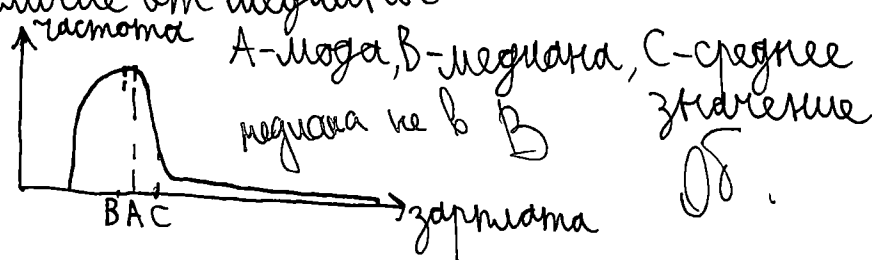
A - мода, B - медиана, C - среднее значение



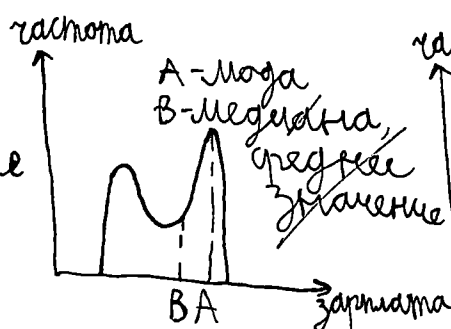
A - мода, B - медиана, среднее значение



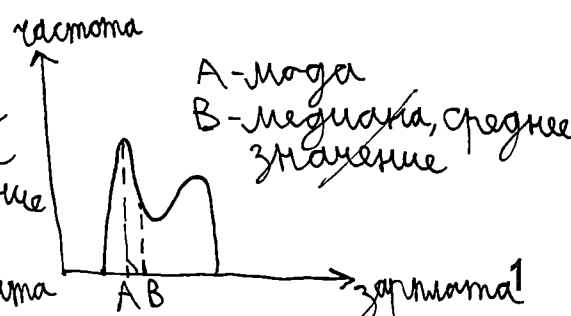
4. Такое распределение возможно, т.к. среднее значение неустойчиво к выбросам, в отличие от медианы.



A - мода, B - медиана, C - среднее значение. медиана не в B.



A - мода, B - медиана, среднее значение



A - мода, B - медиана, среднее значение

Задание 1

- 1 Жители Свердловской области чаще используют транспорт при сильных морозах ($< -25^{\circ}\text{C}$) ⁴⁵ ⁵⁵
- 2 Не желая мерзнуть, жители Свердловской области выбирают общественный транспорт, т.к. там теплее и комфортнее
- 3 Не желая мерзнуть, жители Свердловской области одеваются теплее. Они не выбирают общественный транспорт, т.к. хотят сэкономить деньги или не удовлетворены качеством поездки. ^{идея связь с погодой?} ⁰⁰

Задание 3

- 1-нет-не включают тех, кто не посещал ПтиценыЗдрав +
- 2-нет-не включают тех, кто посещал ПтиценыЗдрав +
- 3-да +
- 4-нет-они не являются пенсионерами Птицены +
- 5-нет-не включают пенсионеров из других районов Птицены +
- 6-да —
- 7-нет* - врачи не могут точно знать ответ на поставленный вопрос, они не являются пенсионерами Птицены +
- 8-нет-не могут знать ответ на вопрос, не являются пенсионерами Птицены +

Задание 5

- 1 Correlation doesn't mean causation (корреляция не означает наличие зависимости). Мне кажется, возможна зависимость доходов от количества перелетов (за календаровские платят ~~больше~~ больше) ⁵⁵
- 2 ~~Корреляция~~ Увеличение количества авиаперелетов положительно влияет на доходы человека ^{с целью работы} ⁰⁰
- 3 Нужно было собрать данные о ценах авиаперелетов и ~~построить~~ построить график зависимости дохода от каждой цели, также можно разделить людей на бедных, средних и богатых и сравнить среднее количество ~~перелетов~~ авиаперелетов в этих группах с помощью t-теста ⁰⁰

Задание 6

Исследование на тему "Какие факторы влияют на успешность учеников ~~школы~~ 8-11 классов?"

1. Время, уделенное на предмет в школе и вне школы, средняя успеваемость по этому предмету

~~Гипотеза: тратя больше времени на изучение материала самостоятельно и в школе, успеваемость по этому предмету увеличивается.~~
~~Механизм~~

Гипотеза: Время, уделенное на предмет, положительно влияет на успеваемость по нему

Механизм: Тратя больше времени на изучение предмета, ученик углубляет и расширяет свои знания по нему, и его успеваемость по предмету увеличивается. 50

Это исследование может быть полезно ученикам школ, так как покажет, действительно ли стоит много часов тратить предмет, который вызывает у тебя трудности, и родителями и учителями, которые будут лучше понимать, как готовить своих детей и учеников, и какие факторы влияют на их успеваемость

2. Генеральная совокупность все школьники РФ, учащиеся в 8-11 классах РФ, потому что у нас единая система образования, одинаковые предметы и оценка успеваемости. Стоит исключить из выборки учащихся ^{программа} Суэзов, частных школ и президентских лицеев, так как их программа отличается 8-11 классы, потому что дети более осознанно подходят к изучению предметов в этом возрасте. 50

3. Анкетирование название предмета, количество часов ~~за~~ в неделю его самостоятельного изучения и изучения в школе в прошлой четверти/полугодии, успеваемость по предмету в прошлой четверти/полугодии, аналогичные данные за предыдущие 2 четверти/полугодия, насколько нравится предмет по шкале от 1 до 10, насколько труден предмет от 1 до 10, изучает ли много своей воле или кто-то заставляет, класс обучения, уровень IQ. 3

(по ссылке на десятичный пест)
Уровень IQ может влиять на успеваемость по любому предмету,
насколько нравится предмет и изучается ли ~~он вообще~~ ^{по своей воле}
может положительно влиять на успеваемость
насколько труден предмет может влиять на время, затраченное на
подготовку

2 Выборка - 10000 случайных школьников РФ (и не только, частные школы, президентские школы) Случайная выборка всегда репрезентативна

4 Количественные переменные - все ^{из анкеты} кроме "изучается ли по своей воле" - это качественные переменные

Графики-диаграммы рассеивания

Будет ~~использован~~ ^{использованы} экстремальный градиентный ~~бустинг~~
бустинг рекурсивных деревьев решений ~~и др.~~

- XGBRegressor(estimator="gbtree"), так как может не быть линейных зависимостей между переменными, а эта модель способна улавливать сложные зависимости и обладает высокой точностью

Выведа коэффициенты модели, мы узнаем, какие же факторы больше всего влияют на целевую переменную - среднюю успеваемость по предмету 50

5 Самое сложное - добиться, чтобы 100 000 школьников прошли анкету Придется разослать ее в административные органы школ РФ Также могут быть учтены не все факторы, влияющие на успеваемость 50