



Титульный лист

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

Класс ☐ 8 ☐ 9 ☐ 10 ☒ 11

Фамилия

МИХАЙЛОВ

Имя

АНДРЕЙ

Отчество

ИГОРЕВИЧ

Дата рождения

25 09 2007

Город участия

ПЕРМЬ

Аудитория

124

Телефон

+79151927337

Дата

01 02 2025

Подпись

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



3101059405146

Проверочный лист

Заполняется участниками

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

Класс ☐ 8 ☐ 9 ☐ 10 ☒ 11

Город участия П Е Р М Ь

Заполняется организаторами

Количество доп. листов 0 2 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	14	12	16	9	15	25				
Балл члена жюри №2	14	12	16	9	15	25				

Итоговый балл 91

Подпись
члена жюри №1

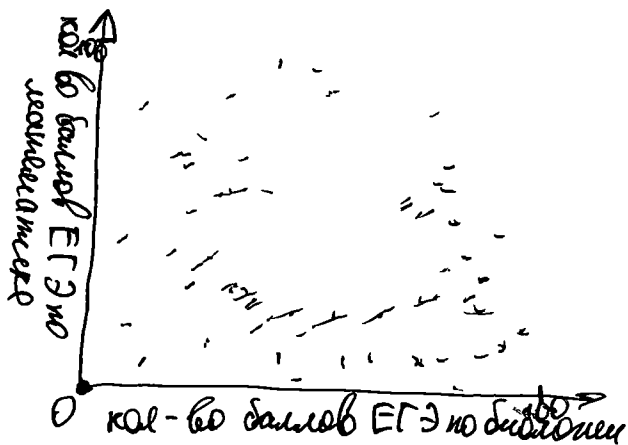
Подпись
члена жюри №2

Пример
заполнения

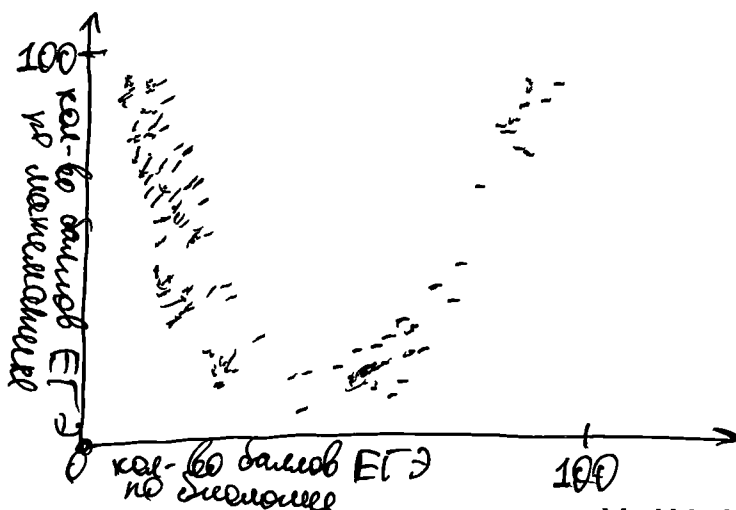
А Б В Г Д Е Ж З И Й К Л М Н О П Р С Т У Ф
Х Ц Ч Ш Щ Ъ Ы Ь Э Ю Я 1 2 3 4 5 6 7 8 9 0

Задача № 2

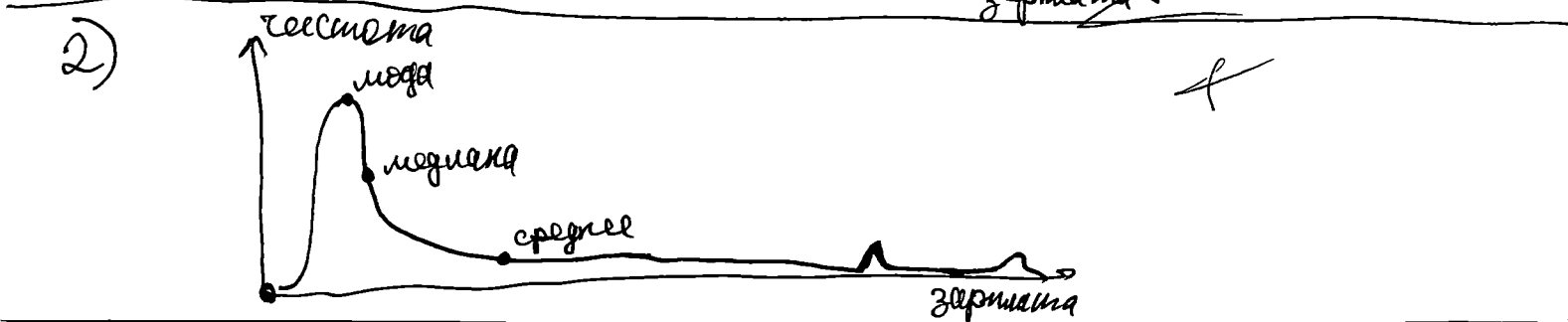
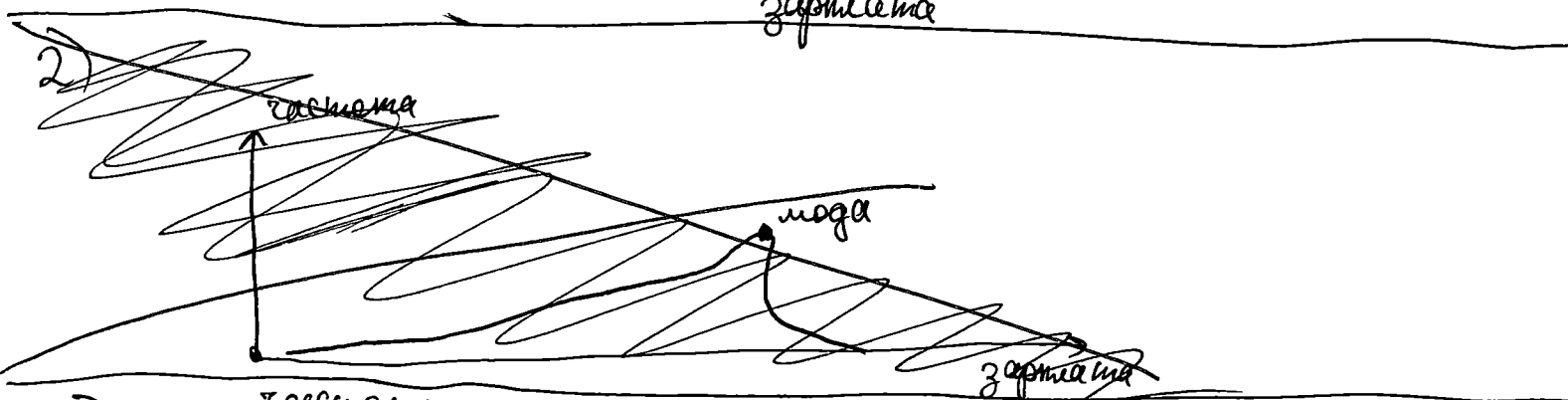
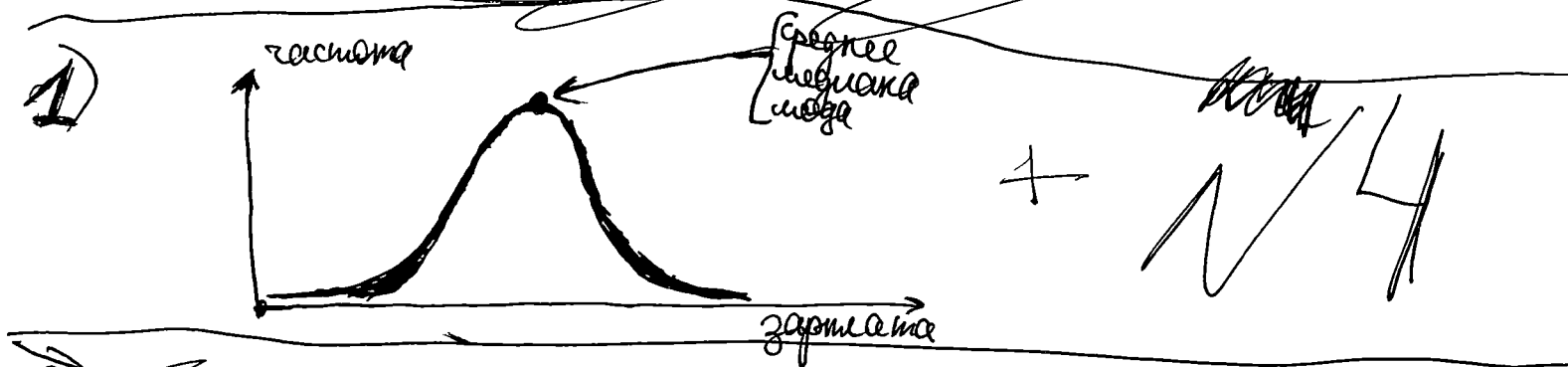
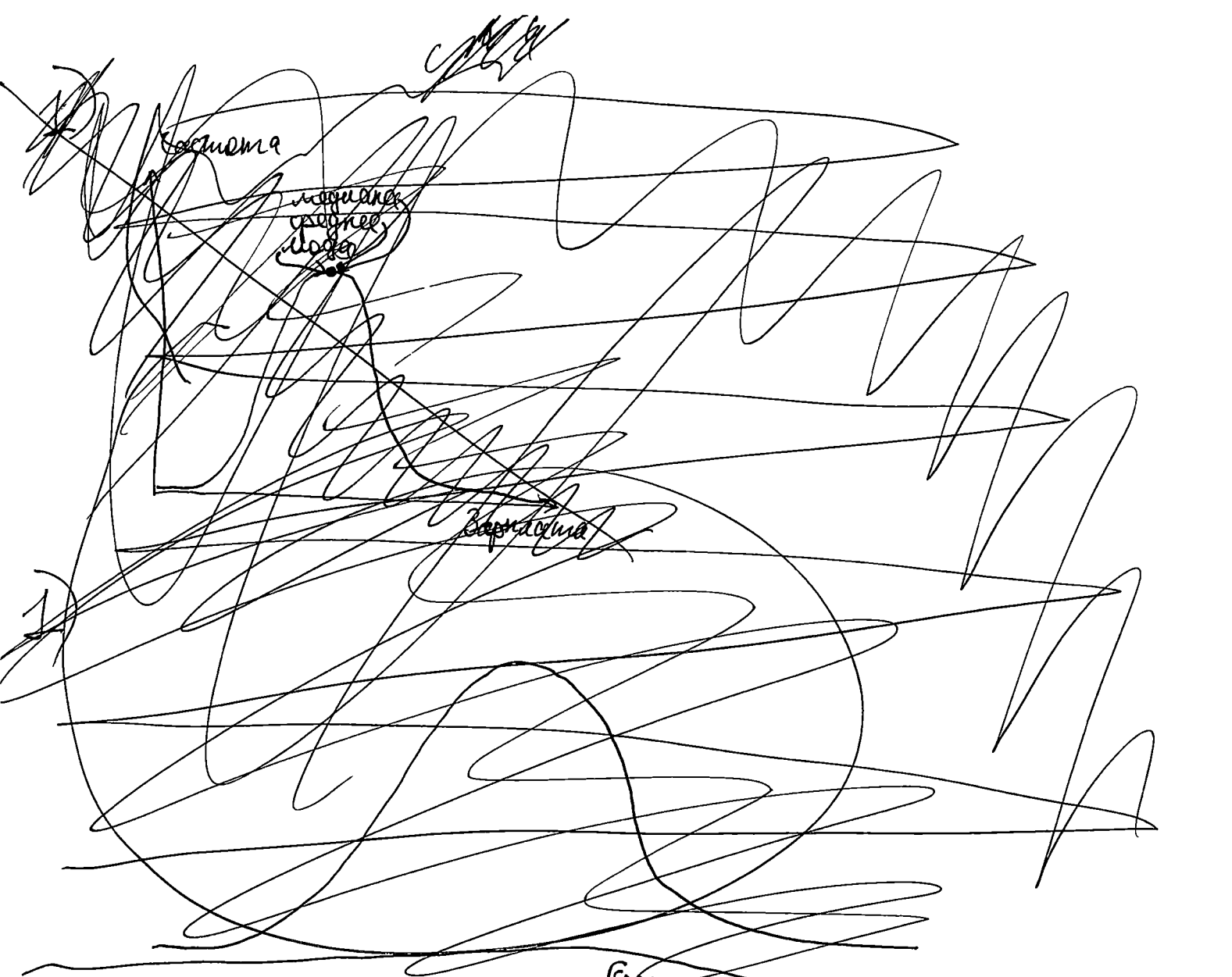
- 1) Если пронумеровать все переменные в заданном порядке по правилам Excel, то получим, что фокусом исследования станут переменные «количество баллов ЕГЭ по математике» и «количество баллов ЕГЭ по биологии» +
- 2) В выборку попало $1200 - 2 + 1 = 1199$ объектов +
- 3) Можно сделать вывод, что взаимосвязь отрицательная и крайне слабая (Скажуневая). При проверке гипотезы, что коэф корреляции ~~равен~~ равен нулю, скорее всего, p-value будет $> 0,05$, т.е. не будет оснований ее опровергнуть. В итоге вывод таков: взаимосвязь отсутствует, т.к. $|r_{xy}| \leq 0,1$ показывает отсутствие взаимосвязи (или слабейшую).
Также возможен вариант нелинейной зависимости, которая не может быть распознана линейной корреляцией Пирсона, которая по умолчанию используется в Excel +
- 4) Вариант для отсутствия связи PS r_{xy} - коэф корр между x и y

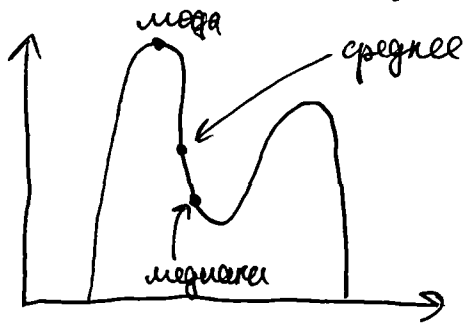
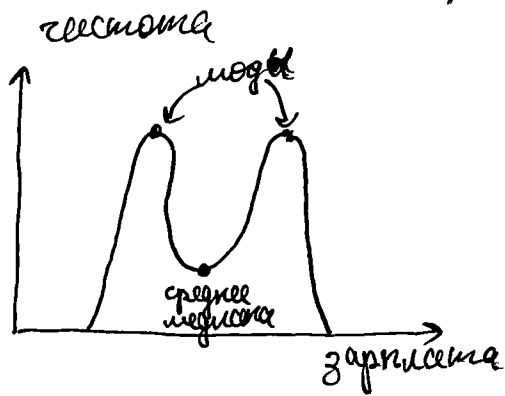


Вариант для нелинейной зависимости



- 1) Ошибка корреляции. Этот феномен заключается в том, что человек ищет корреляции и причинно-следственную связь, и в то же время в том, что А влияет на В, наблюдая лишь слабую корреляцию, ~~которая~~ не обоснованную с точки зрения логики и/или направленных связей.
- В данном примере Таша как раз и говорит, что увеличение доходов человека положительно влияет на кол-во авиаперелетов, допуская основную ошибку корреляции, ведь откуда он знает направление взаимосвязи?
- 2) Корректный вывод: Доход человека и количество его авиаперелетов положительно скоррелированы/взаимосвязаны. ИЛИ Между доходом человека и ~~количеством~~ его авиаперелетов наблюдается положительная корреляция/взаимосвязь.
- В данном случае можно говорить о корреляции, так и про взаимосвязь, понимая, что очевидно механическая связь, но не столько ее направление.
- 3) Можно сравнить кол-во перелетов у человека при разных уровнях его дохода, провести такое исследование для достаточного кол-ва людей, собрать при этом репрезентативную выборку, а затем проверить гипотезу с помощью стат. тестов.
- Например, можно взять средний доход в регионе за базу, ~~а~~ затем провести опрос на уровне ("Сколько вы летали в этот год при доходе X?"). При доходе 0,75X? При доходе 1,25X? или что-то подобное, ~~и~~ затем посмотреть на средние подгруппы, сказать, что они различаются (с помощью ANOVA), а затем попарно сравнить средние (с помощью критерия Тьюки или аналога).
- Тогда можно будет подтвердить (или опровергнуть), что с увеличением доходов человек увеличивает базис летания, или же "Увеличение доходов человека положительно влияет на кол-во его перелетов".





4) Формализм требования не бывает таким, что ~~можно~~ $медiana < мода < среднее$. Опровергну это.



Скелетный пример такого распределения
Т.е. если будет вот такое ~~вот~~ скользящее распределение с
~~определенным~~ определенным количеством точек сильно справа (этих
аномалий), то медиана, по причине своей робастности, останется
слева, тогда будет в самом левом значении (идет к выводу
медианы), а среднее сместится сильно вправо из-за своей неустойчивости
к аномальным значениям.

Мне было бы интересно исследовать ^{уровня} финансовую грамотность человека на его инфляционные ожидания Трейдз по пунктам, на которые необходимо ответить:

1) Факторы.

- инфляционные ожидания ~~человека~~ человека (как вы считаете, насколько процентов вырастут цены за следующие 12 месяцев?)
- уровень финансовой грамотности человека (индекс, являющийся суммой 5 базисных элементов, ~~основанный~~ в академической записке ЛБ РФ ~~основанный~~ ^{основанный} принимает значения от 0 до 5 и покрывает уровень инфляционных ожиданий, основываясь на базовых вопросах по нему)

Гипотеза:

- Чем выше уровень инфляционных ожиданий человека, тем ниже его финансовая грамотность.

Механизм:

① человек финансово неграмотен

обращает внимание лишь на порочившие товары (например, яйца на ЧД, когда остальные не особо порочиваются)

формирует свои инфляционные ожидания на основе ~~информации~~ и лишь на основании ~~информации~~ ^{высокой инфляции}, не обращая внимания на другие

~~его~~ его инфляционные ожидания ~~связаны~~ ^{связаны} в основном с ~~одной~~ ^{одной} стороной

② человек финансово грамотен

обращает внимание ^{на} как на сырьевые, так и на порочившие товары

формирует свои инфляционные ожидания ^{всесторонне} и основываясь на фактах и данных

его инфляционные ожидания ^{адекватны} и не смещены сильно

не имеет экспертные навыки и навыки, т.к. не владеет о них

Результаты исследования могут быть полезны как та-бу, чтобы
~~понимать, как это можно и как это можно~~
 что необходимо вводить новые по-
 нятия для повышения дик. грамотности ^{людей}, и к.
 нийкий её уровень увеличивается и др. ожидания, которые от-
 тствии явиться самостоятельными и увеличивают
 и др. того высокоур. и др. и как разные люди могут быть
 понимать, на основе чего и как разные люди могут быть
 и др. ожидания. Также это исследование может быть
 знаком и самим людям, чтобы не понимать, что является
 и др. и что можно обнаружить

2) Ген. совокупность - ~~все население РФ~~ (все)
 все граждане РФ в возрасте от
 18 до 65 лет (совершеннолетние
 и работоспособные)
 Временные рамки - 2022 год (последнее
 актирование домохозяйств от 17 Б
 рр)
 рршю имено тогда элемент*)

Способ формирования выборки - простое случайное до-
 мохойств (на дому / по телефону)
 * - немото
 ния, возраста, мест жительства (район и тип насел. пункта) из-
 за стратификационной выборки. Опрос проводится каждый
 несколько лет
 Размер выборки ~ 12 тыс чел

3) Источники данных - панельная база Всероссийского
 исследования домохозяйств 17 Б рр с помощью актирования
 Он проводится в открытом доступе на сайте с акцией.
 Для проверки модели на упорядоченность можно использовать
 такие переменные, как пол, возраст, уровень образования, горизонт
 тамирования (на сколько лет времени люди тамируют свои и др.
 жизнь) и др

инд. ожиданий - как-белая непрерывная

4) Тип ~~индекса~~

Тип индекса фрек. ~~временности~~ ~~категориальный упорядоченный~~
или как-белый дискретный
(можно рассматривать и так, и так)

Тип графиков для исследования взаимосвязи:

1. ~~Boxplot~~ ("мизик сусам") по инд. ожиданиям
для каждой категории индекса фрек. временности

2. error bar (среднее + довер. интервалы) по инд. ожиданиям для каждой категории индекса фрек. временности

При использовании для ~~инд.~~ док-ва гипотезы линейной регрессии необходимо к этим графикам также добавить график регрессии.

Есть 2 возможных модели:

1. Линейная регрессия ~~инд. ожидания~~ ~~каждой~~ (завис. переи) (независ. переи)
ми. С ее помощью можно будет показать, на сколько % увели-
шатся инд. ожидания при увеличении уровня фрек. временнос-
ти на 1 ~~единицу~~ ~~обнаружить различия в среднем~~ (и генерал) ~~сравнение средних~~

2. ANOVA test + последующее ~~попарное~~ ~~сравнение средних~~ ~~критерий (Томкс или аналог)~~
подвыборки, чтобы понять, средние какие именно подвыборки
различаются. Доказана гипотеза будет, если все группы будут
значимо отличаться друг от друга, а так же у групп с
индексом фрек. временности = 0 будут максимальные инд. ожидания
в среднем и т.д.

** гипотеза будет доказана, если p-value перед котр будет < 0.05
а $Kozp$ по модулю < 0

Также линейную регрессию можно проверить на существование
резид. способами - добавившим контрольным переменным (указаны в
с последующей проверкой значимости коэффициентов, также
можно проверить знача и несильную зависимость по модулю, а второй -
(см. на слайке 2)

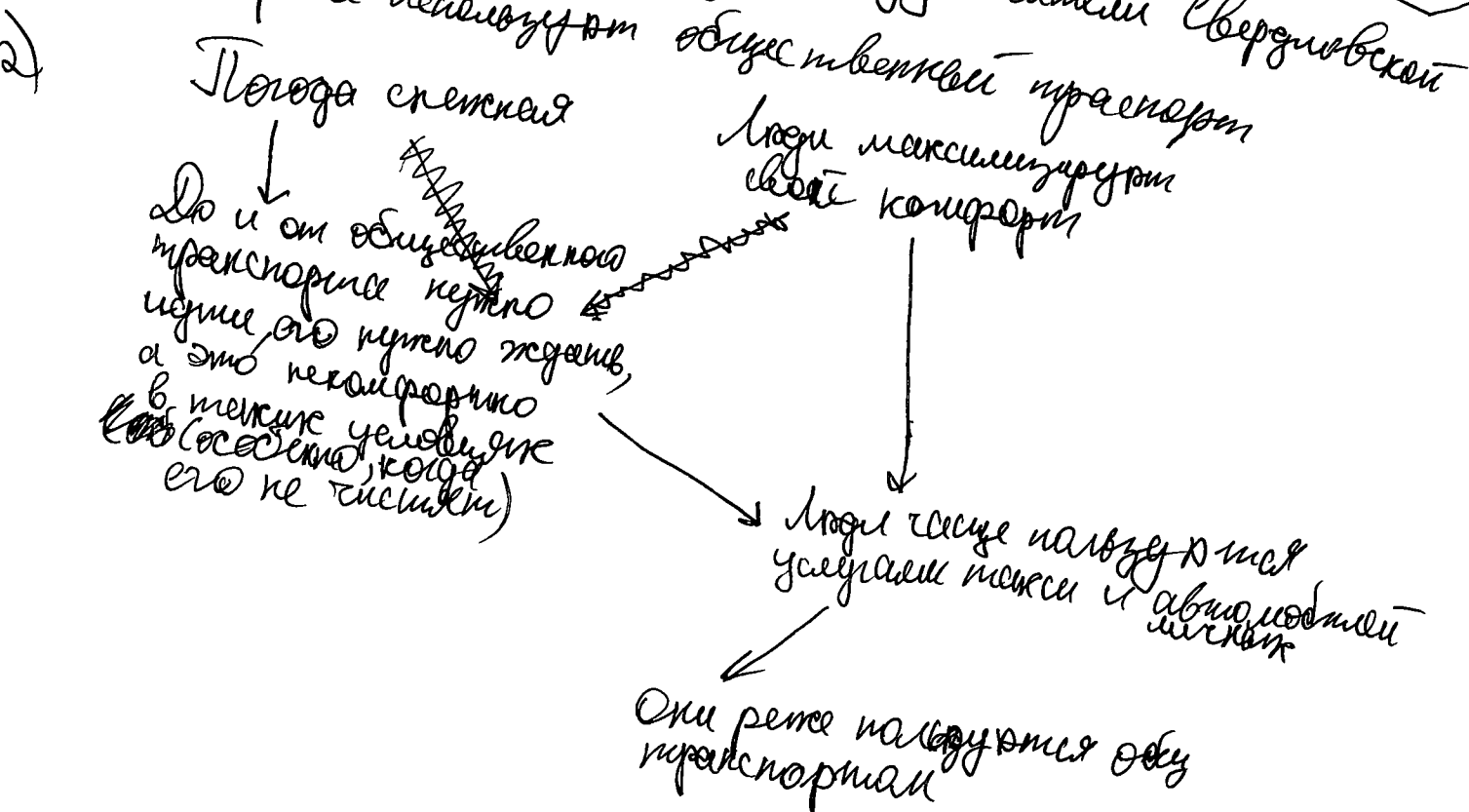
5) Потенциальные предпосылки

1. Нечеткое продолжение анкеты («от багды»)
2. Нечеткое и неясное опред. уровня риска угрозы
3. Недостаточный размер выборки
4. Недостаточность использовать конкретные модели из-за неясности их предположений

5. Невозможность обобщить исследование на предсказуемых годах

из-за сужения экологической и политической ситуации
 а) а второй - разделение по переменным из п.3 не подтверждает и провер-
 ка гипотезы на этих показателях имеет смысл также других методов
 можно использовать и для проверки гипотезы

1) В ~~зоне~~ спектру погоды типичен свердловский



3) Противоположная Эпштейна в светлую погоду
иногда плавятся в излучении

Людям
максимизируют
свое
самостояние

Погода светлая

Людям максимизируют
свой комфорт

таким
станет дорога

на дорогах
такого
снег, пробки

Людям реже
используют
такси

Людям реже
используют
личное автотранспорт,
так как это стресс
и потраченные нервы

добираться
реже к
транспортной
лучше тем
на авто

Людям чаще используют
осуществляют транспорт