Листок. Введение в анализ данных. Качественные задачи
Преподаватели: Рита Голуб, Ирина Козловцева
Дедлайн: 12 августа 2022 года, 21:00 МСК
Задачи в этом листке можно сдавать **в любом порядке**.

Задача 1. Яйцо или курица?

На лекции мы рассмотрели основные типы взаимосвязей показателей при корреляциях (прямая, одна причина нескольких следствий, ложная, зацикленная). В этой задаче вам предстоит для каждой корреляции определить тип причинно-следственной связи.

а) Уровень здоровья и количество принимаемых витаминов

б) Импорт нефти в США из Норвегии и кол-во смертей водителей от столкновения с ж/д транспортом

в) Количество поездок на такси и количество перелетов

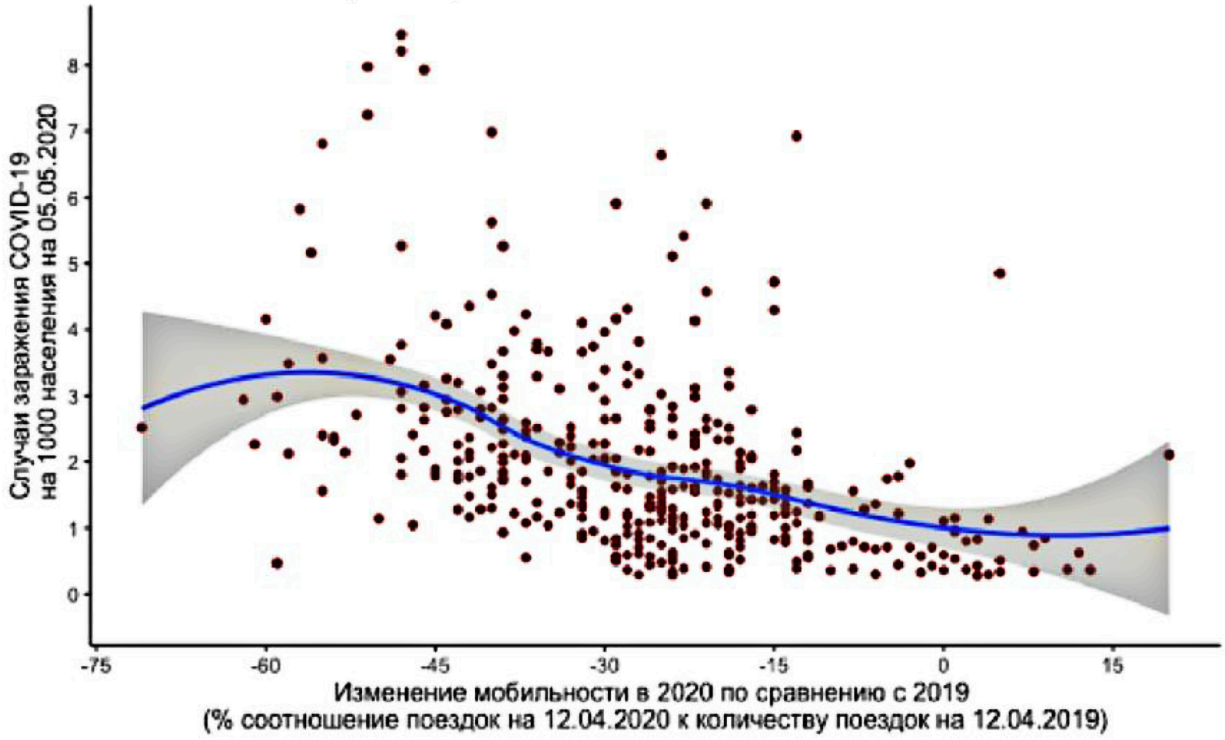
г) Продажи компании и расходы компании на рекламу

д) Расходы на отпуск в Новый год и расходы на отпуск в майские праздники

е) Количество людей в метро в час пик и пробки на дорогах

Задача 2. На удалёнке!

На графике ниже представлена взаимосвязь мобильности населения и заболеваемости COVID-19 в Германии. Каждая точка — регион в Германии.



а) Отрицательная или положительная взаимосвязь (корреляция) имеет место в данном случае?

б) Исходя из знака корреляции, предложите механизм, объясняющий такое направление взаимодействия. Что будет причиной, а что — следствием?

Задача 3. Где расположить пункт выдачи заказов?

Вы хотите открыть много пунктов выдачи заказов (ПВЗ) маркетплейса (например, такие есть у OZON, Маркета, WB, служб доставки CDEK и т.д.) в городе, где таких пунктов выдачи еще нет ни одного.

Ваша цель — поставить ПВЗ так, чтобы заказов было как можно больше (чем меньше издержки покупателя на поход в ПВЗ, тем больше будет совершаться заказов), поэтому есть желание открыть ПВЗ буквально в каждом доме. При этом есть издержки на содержание каждого ПВЗ (аренда, з/п сотруднику) и т.д. Поэтому надо выбрать в том числе оптимальное количество ПВЗ.

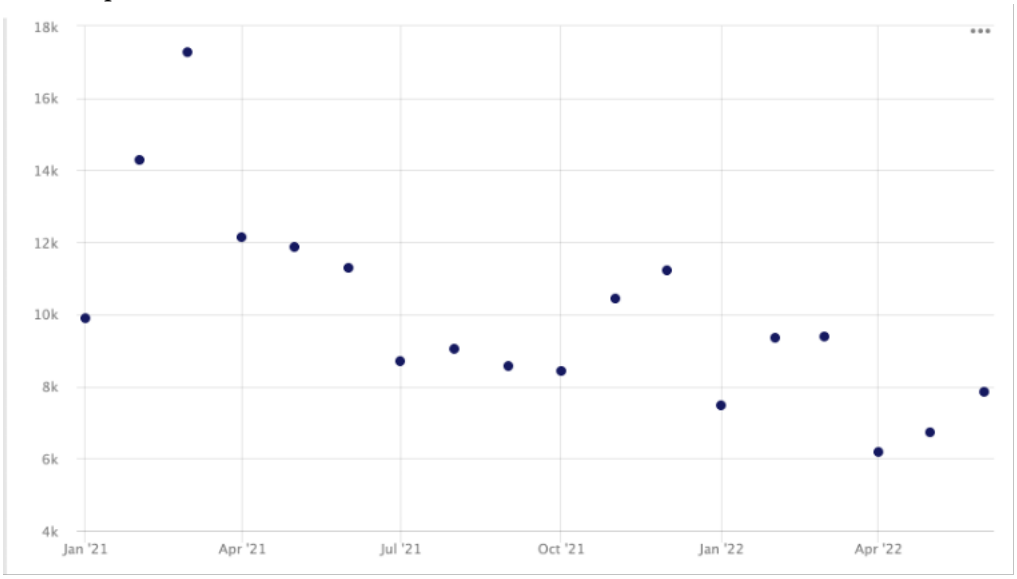
Ваши данные — у вас есть координаты абсолютно всех объектов города (жилые дома, школы, остановки общественного транспорта, кафе и т.д.), но нет данных о популярности этих мест.

Открыть, а потом закрыть, если не выгодно держать ПВЗ — очень дорого. Открыть сначала один, а потом к нему открывать еще, если потребуется, приведет к недополучению заказов, пока открыто недостаточно ПВЗ.

Где будем ставить ПВЗ (например, в школах)? Почему мы сможем добиться (насколько это возможно) правильного количества ПВЗ? Почему именно в этих локациях оптимально ставить ПВЗ?

Задача 4. Маркетплейс М стал хуже?

На графике на реальных данных построена средняя выручка продавцов, продававших что-либо на маркетплейсе М, по месяцам.



Как мы видим, средняя выручка продавцов снижается. Значит ли это, что маркетплейс М стал хуже работать? Приведите 2 объяснения/метрики, почему, несмотря на график выше, маркетплейс М прекрасно работает и улучшается? Если необходимо, к объяснению приведите график, который помог бы нам подтвердить это объяснение.

Задача 5. Кредит давать будешь?

(Эта задача по лекции 12.08.2022.) Юный экономист хочет разобраться в том, действительно ли, когда Центральный банк требует от банков хранить больше резервов от потерь по одному виду кредита, банки начинают выдавать меньше этих кредитов. Он нашел данные по общему объему выданных кредитов нужного типа, а также по датам введения мер, которые должны снижать стимулы банков к выдаче этих займов. В результате экономист оценил, что темпы прироста объема кредитов зависят от вводимых мер следующим образом: $Y = -3.21t_0 + 2.35t_1$. То есть темп прироста кредитов сокращается в квартал введения меры на 3.21 пп, а затем в следующий квартал растет на 2.35 пп.

а) Юный экономист поделился этим результатом с коллегой из соседнего отдела, который сказал, что в сумме мера оказывается неэффективной, ведь в конечном счете кредиты продолжают расти так же, как и росли раньше. Верно ли рассуждение коллеги, или он не прав и эффект на кредитование все же оказывается?

б) Юный экономист задумался над тем, насколько хорошие результаты он получил. И у него закрадываются подозрения, что он не решил некоторые имеющиеся проблемы. Назовите хотя бы один фактор, которую не учел юный экономист при проведении данного исследования?