

Подборка

«Графический анализ данных. Задачи на уравнение регрессии»

(Часть 1)

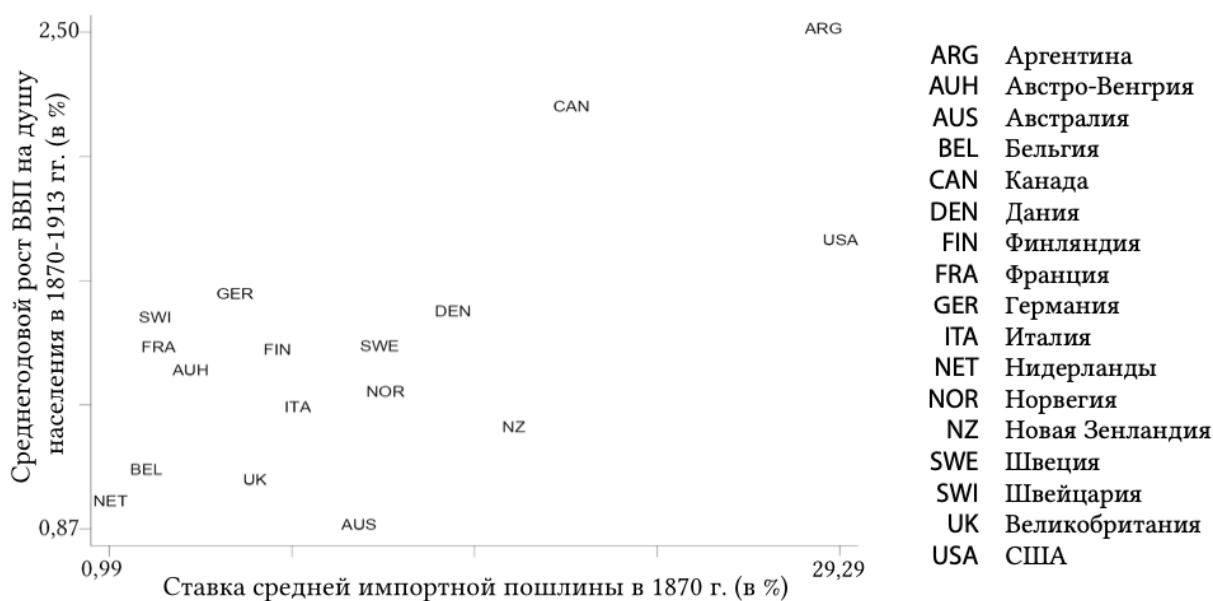
Факультатив «Качественные задачи по экономике»
Лицей НИУ ВШЭ

25 октября 2019 г.

Задание 1: Тарифы и рост

ВОШ 2019 – 9 класс 1 тур

Дискуссии о протекционизме и свободной торговле являются порой одними из самых жаростных. Нередко сторонами привлекаются в качестве аргументов исторические данные. Рассмотрите следующий график¹:



Например, из графика видно, что в Аргентине и тарифы (пошлины), и темп роста подушевого ВВП в указанный период времени были высокими, а в Нидерландах – низкими. И вообще, на первый взгляд, чем выше были тарифы, тем в среднем более высокие темпы роста были в стране. Экономисты называют такое явление положительной корреляцией. Кажется, что на этом графике между уровнем тарифов и среднегодовым темпом роста была положительная корреляция и из этого можно сделать выводы, актуальные для экономической политики. Однако не всё так однозначно.

- а) Объясните, почему из этого графика не следует, что между показателями на самом деле была положительная корреляция.

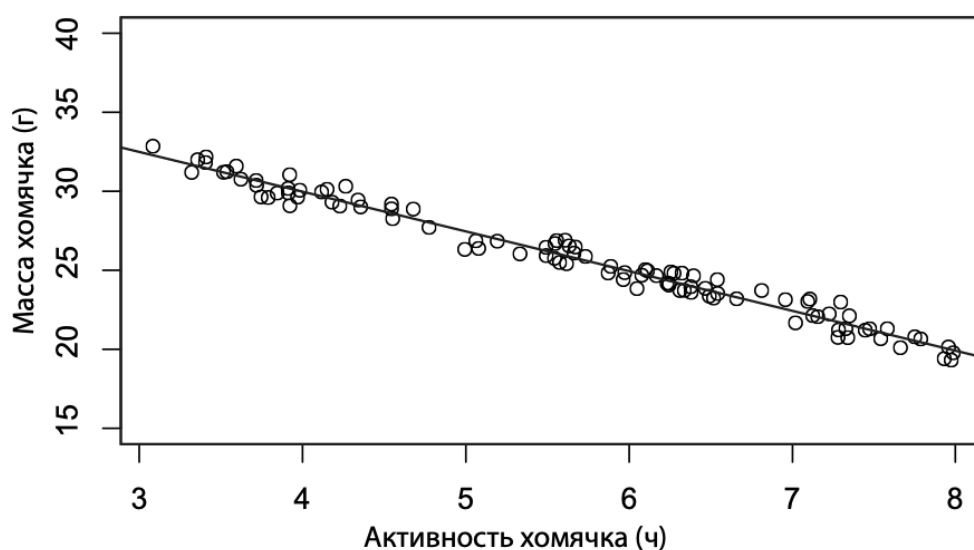
¹Irwin, Douglas A. "Interpreting The Tariff-Growth Correlation Of The Late 19th Century," American Economic Review, 2002, v92(2, May).

- б) Если положительную корреляцию всё же удастся достоверно установить, сторонник протекционизма с удовольствием сделает вывод, что импортные пошлины положительно влияют на экономический рост. А значит, для ускорения роста государству следует повышать пошлины. Однако, может быть, причинно-следственная связь устроена не так? Приведите альтернативное объяснение положительной корреляции, в котором тарифы не влияют на темпы роста.
- с) Какие данные могли бы помочь выяснить, влияют ли тарифы на темпы роста? (Ваш ответ может включать данные, которые на практике получить сложно.)

Задание 2: Образование и доход

ВОШ 2018 – 9 класс 1 тур

Начинающий исследователь Василий однажды читал научную статью, где изучалась связь активности хомячков (в часах в среднем в течение суток) и их массы. Там приводился следующий график:



Точками здесь обозначены отдельные хомячки, а прямая построена так, что она лежит как можно ближе к точкам. Прямая имеет уравнение $y = 40 - 2.5x$, из чего авторы исследования сделали вывод, что увеличение активности на час в среднем уменьшает массу хомячка на 2.5 грамма.

Василий решил использовать эту технику, чтобы оценить влияние образования людей на их доход. Для этого он опросил 1000 человек, спрашивая у каждого, сколько тот зарабатывает (переменная *Income*) и каков его последний на данный момент уровень образования (переменная *Education*). Переменная *Education* для каждого человека принимает одно из 5 значений:

Education	Уровень образования
1	не окончил школу
2	окончил только школу
3	окончил техникум/колледж
4	окончил университет
5	получил ученую степень

Построив рядом с полученными точками прямую так же, как было сделано в исследовании, которое он читал, Василий обнаружил, что ее уравнение имеет вид

$$Income = -10 + 20 \times Education$$

(это уравнение прямой, самой близкой к точкам на графике), то есть каждая ступень образования в среднем увеличивает доход на 20 тысяч рублей в месяц. Исследование, проведенное Василием, не свободно от недостатков. Вам нужно высказать содержательную критику по следующим пунктам:

- а) Кодирова уровень образования цифрами от 1 до 5, Василий неявно предполагает очень специфическую форму зависимости дохода от образования. Объясните, какую форму зависимости предполагает Василий, почему так лучше не делать и как ему стоило бы правильно учесть образование в своей модели?
- б) Во-вторых, Василий не учел все факторы, которые могут влиять на доход. Какие? Предложите, как надо было организовать исследование, чтобы корректно измерить влияние образования на доход.