

Корреляция

Зачем нужна корреляция

Корреляция — это статистический показатель, который измеряет силу и направление **связи** между двумя переменными.

Например, можно узнать как проявляется связь между средними зарплатами в регионе и суммарным количеством поездок на такси.

Зачем нужна корреляция

Силу связи определяет модуль коэффициента корреляции, который принимает значения от 0 до 1.

$0,7 \leq c < 1$ — сильная связь

$0,3 \leq c < 0,7$ — средняя связь

$c < 0,3$ — слабая связь

$c = 0$ — связь полностью отсутствует

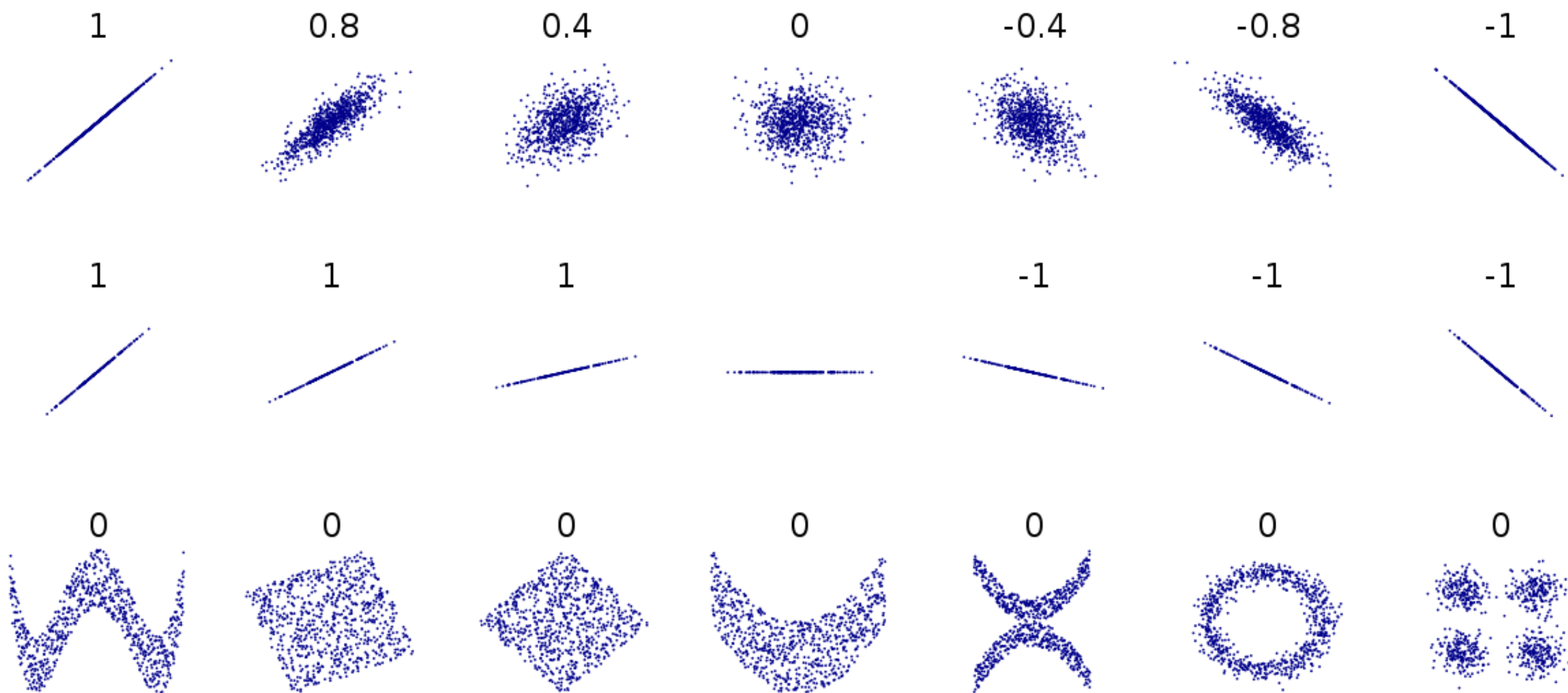
Зачем нужна корреляция

Направление связи определяет знак коэффициента корреляции. Если коэффициент больше нуля, то связь положительная:

чем **больше** зарплата в регионе, тем **больше** там ездят на такси.

Если коэффициент меньше нуля, то связь отрицательная:

чем **меньше** рейтинг коррумпированности страны, тем **больше** в ней зарплаты.



Связь \neq причинно-следственной связи

Корреляция показывает, как изменяется одна переменная относительно другой.

Допустим,

количество айтишников в городах России может иметь сильную положительную корреляцию с числом людей, находящихся за чертой бедности.

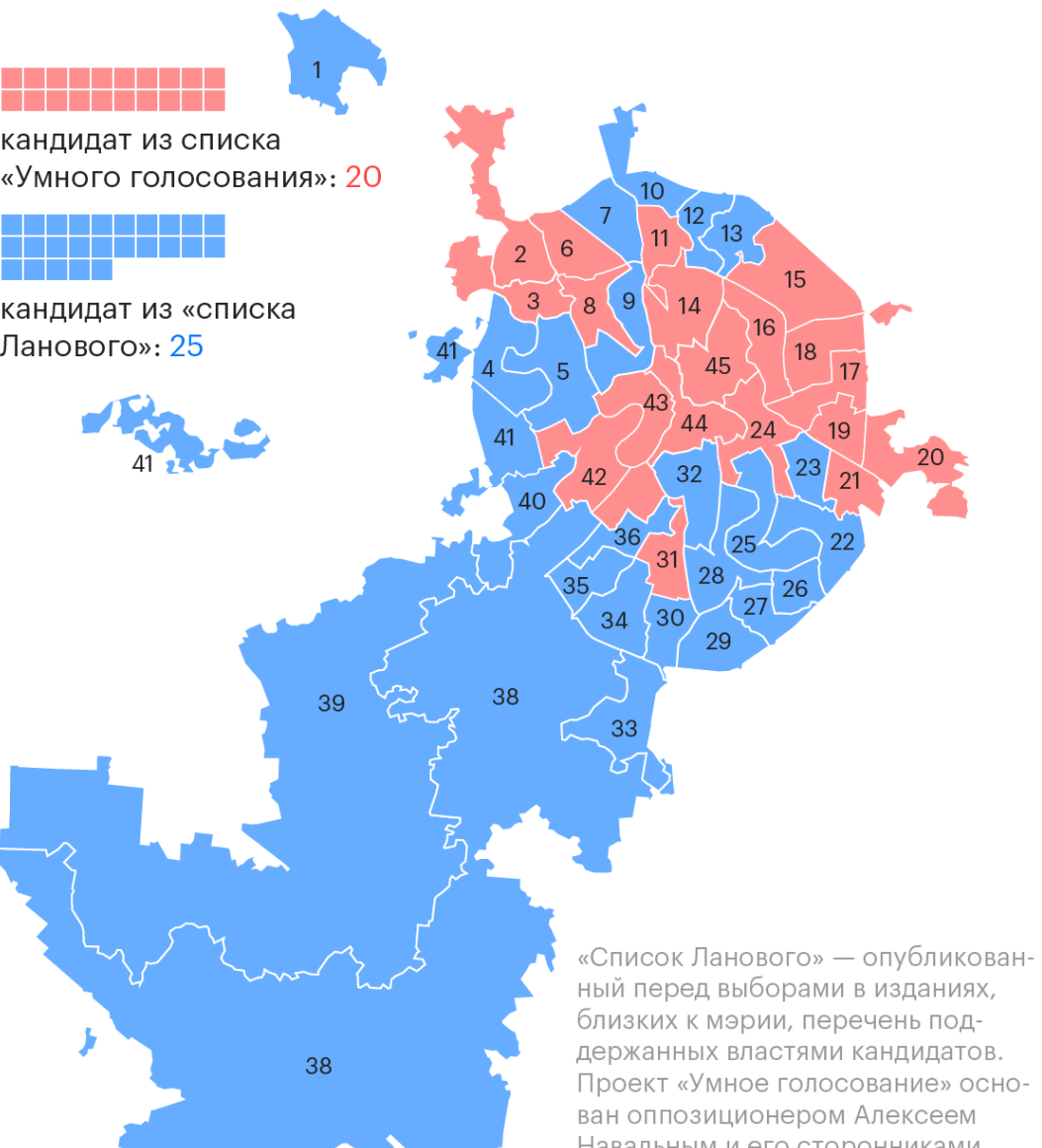
Однако это не значит, что если мы насильно начнём делать людей бедными, то айтишников станет больше.

Это не причинно-следственная связь.



Итоги выборов в Мосгордуму

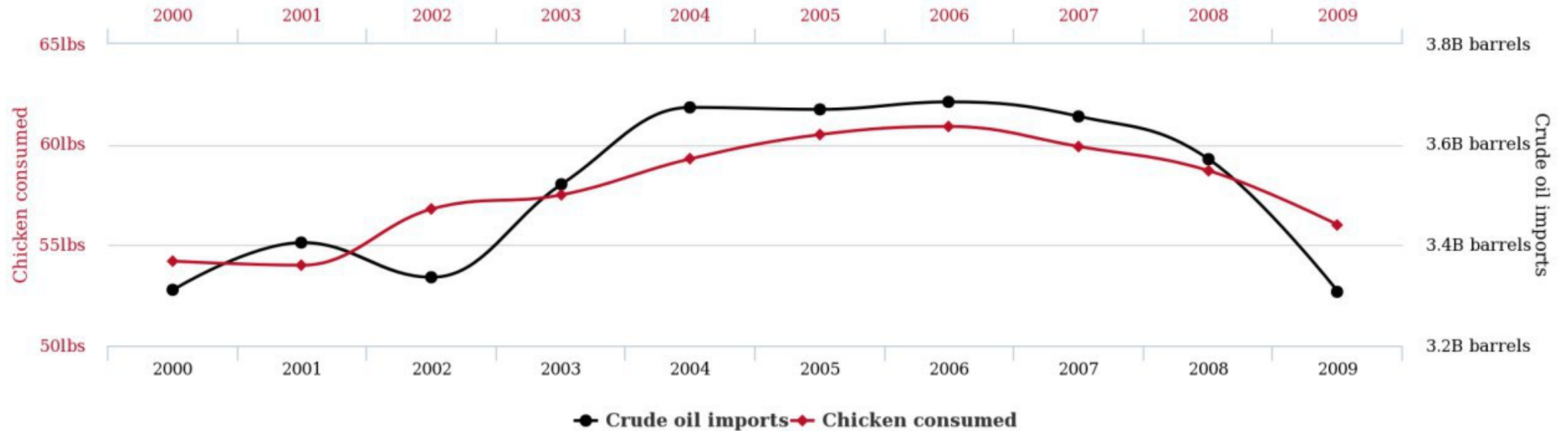
По данным на 12:00 мск 9 сентября 2019 года. Указаны номера одномандатных округов



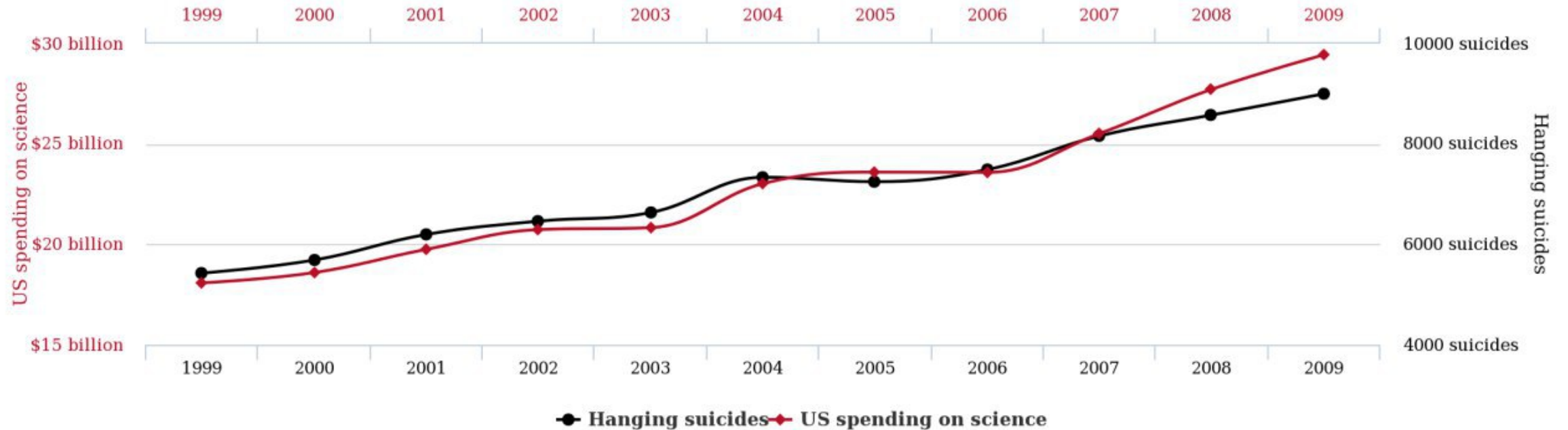
Per capita consumption of chicken

correlates with

Total US crude oil imports



US spending on science, space, and technology correlates with Suicides by hanging, strangulation and suffocation



**Как убедиться, что причинно-
следственная связь — есть?**

Если не проводить полноценное
исследование, то — никак.

Если не проводить полноценное
исследование, то — никак.

но: можно обратиться к экспертам из
области