



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Data Analysis National Olympiad - DANO

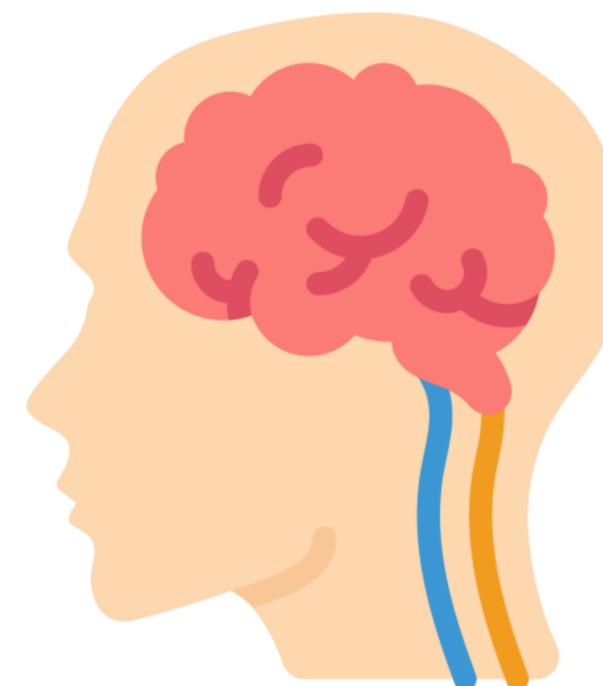
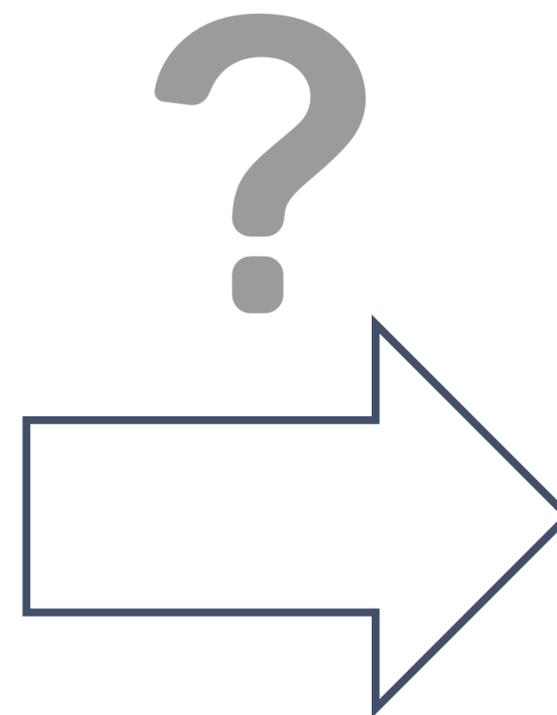
ЕДИНИЦЫ НАБЛЮДЕНИЯ В СТАТИСТИЧЕСКОМ ИССЛЕДОВАНИИ

Наталья Горюнова
Алиса Райковская

Санкт-Петербург, 2021

ПРОВЕДЁМ ИССЛЕДОВАНИЕ

Есть ли связь между физической подготовкой и умственными способностями человека?



НЕОБХОДИМЫЕ ПЕРЕМЕННЫЕ

Есть ли связь между физической подготовкой и умственными способностями человека?

Объясняющие



Индекс массы тела



Сколько раз подтягивается?

...

Объясняемая



Уровень IQ

Контрольные¹



Доход семьи



Наличие братьев и сестёр



Образование родителей

...

Сопоставимость объектов²



Измерение величины за единицу времени



Учёт неоднородности данных (возраст, географическое местоположение)

¹ см. "Контрольные переменные"

² см. "Единицы измерения"

СТАТИСТИЧЕСКОЕ НАБЛЮДЕНИЕ

Инструмент для получения данных



Эксперимент



Опрос



Полевое наблюдение



Готовые данные

Пример:

Измерить количество подтягиваний за минуту

Собрать информацию о росте и весе участников

Наблюдать результаты участников на уроке физкультуры

Запросить данные о физической подготовке у учителя физкультуры



ЕДИНИЦА СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ

Объект, с которого собирали данные

Имя	Уровень IQ	Бег на 2 км (минуты)	Индекс массы тела	Возраст	Доход семьи на чел. (в мес.)
Игорь Петров	110	8.40	23.4	14	40 000
Анна Ежова	125	9.31	19.3	23	70 000
Артём Детко	144	7.20	22.5	16	35 000
	Объясняемая переменная	Объясняющие переменные		Контрольные переменные	



ЕДИНИЦА НАБЛЮДЕНИЯ НЕ МЕНЯЕТСЯ ПРИ ДОБАВЛЕНИИ ПРИЗНАКА ИЛИ ФИЛЬТРАЦИИ ДАННЫХ

Имя	Уровень IQ	Бег на 2 км (минуты)	Страна	Возраст	Посещение секций
Игорь Петров	110	8.40	Россия	14	да
Джордж Брукс	131	8.70	США	13	да
Мариа Лучини	138	9.36	Испания	14	нет

Расширили выборку

Только школьники

Новый признак



ЕДИНИЦА НАБЛЮДЕНИЯ МЕНЯЕТСЯ ПРИ АГРЕГАЦИИ ДАННЫХ

Класс_школа	Средний балл физкультура	Средний балл математика	Город	Кол-во уроков математики
8Б_877	4.35	4.35	Москва	5
9А_641	4.67	3.99	СПБ	4
7Г_50	4.44	4.66	СПБ	8
	Объясняемая переменная	Объясняющая переменная	Контрольные переменные	



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

ПРИМЕР №1





КАКОЙ УЖИН ОПТИМАЛЕН ПО СООТНОШЕНИЮ ЦЕНА – КАЛОРИИ?

Единица наблюдения - товар

Имя	Ккал на 100г	Цена 100г
Огурцы	16	12
Сыр	400	60
Колбаски	350	65

Единица наблюдения - блюдо

Имя	Ккал на 100г	Цена 100г
Паста песто	330	150
Пицца 4 сыра	340	135
Овощной салат	80	100



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

ПРИМЕР №2





КАК ТРАНСПОРТНАЯ ДОСТУПНОСТЬ ВЛИЯЕТ НА СТОИМОСТЬ ЖИЛЬЯ В МОСКВЕ?

ID_квартиры	Цена за м2	Расстояние до метро	Расстояние до наземного трансп.	Площадь квартиры	Район Москвы
380042	110	8.40	23.4	14	Академический
102348	125	9.31	19.3	23	Сокольники
122600	144	7.20	22.5	16	Арбат

Объясняемая переменная **Объясняющие переменные** **Контрольные переменные**



ИСТОЧНИКИ

The screenshot shows the CИАН website interface. At the top, there are navigation links for 'Аренда', 'Продажа', 'Новостройки', 'Коммерческая', and 'Ещё'. A search bar is present with filters for 'Регион', 'Метро', and 'Найти'. Below the search bar, there are buttons for 'Купить', 'Квартиру в новостройке и вторичке', '1, 2 комн.', 'Цена', and 'Ещё фильтры'. The main content area displays a property listing for 'Продажа 1-комн и 2-комн квартир в Москве'. It includes a photo of the apartment interior, the price '10 510 000 ₺', and contact information for 'Capital Mars'.

<https://spb.cian.ru/>

The screenshot shows the Rosstat website. The header includes the logo of the Federal Service for State Statistics and a search bar. The main headline is 'Сельхозперепись – в августе по всей стране'. Below this, there is a section titled 'Новости Росстата' with a blue background. The featured article is 'Росстат представляет оценку ВВП за I квартал 2021 года методом использования доходов и методом формирования по источникам доходов', dated 02.07.2021 19:00. The page also features a navigation menu with 'Все новости' and 'Подписаться на рассылку', and a footer with 'Анонсы' and 'Публикации'.

<https://rosstat.gov.ru>

The screenshot shows the World Bank Open Data website. The header includes the World Bank logo and navigation links for 'DataBank', 'Microdata', and 'Data Catalog'. The main heading is 'World Bank Open Data' with the subtitle 'Free and open access to global development data'. A search bar is present with the text 'Search data e.g. GDP, population, Indonesia'. Below the search bar, there are sections for 'MOST RECENT' (including 'New World Bank country classifications by income level: 2021-2022'), 'WHAT YOU CAN LEARN WITH OPE' (including 'Poverty headcount ratio at \$1.90 a day (2011 PPP) (% of population)' and 'Extreme Poverty'), and 'Atlas of Sustainable Development Goals 2020'.

<https://data.worldbank.org>



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

ПРИМЕР №3



stop

ЕСТЬ ЛИ СЕЗОННОСТЬ В АКЦИЯХ АЭРОФЛОТА?

Единица наблюдения – определённый день

Дата	Цена акции
29.06.2021	69.2
30.06.2021	68.1
01.07.2021	68.5





НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

ПРИМЕР №4





КАК УРОВЕНЬ ПРЕСТУПНОСТИ В РАЙОНЕ ВЛИЯЕТ НА ЗАРПЛАТУ ВРАЧА-ТЕРАПЕВТА?

Единица наблюдения – больница в определенный год

Номер больницы	Год	Округ	Зарплата врача-терапевта	Число преступлений на 10 000 человек
52	2012	Центральный	40 000	170
52	2013	Центральный	45 000	221
52	2014	Центральный	48 000	214
28	2012	Юго-западный	55 000	132
28	2013	Юго-западный	56 000	148
28	2014	Юго-западный	60 000	135

ОБОБЩЕНИЕ ТЕМЫ

Статистическое наблюдение



Собираем данные об объектах

Единица статистического наблюдения

Объект, к которому относятся признаки

Единица **зависит** от:

Уровня агрегации (по стране, по области, по городу)

Единица **НЕ** зависит от:

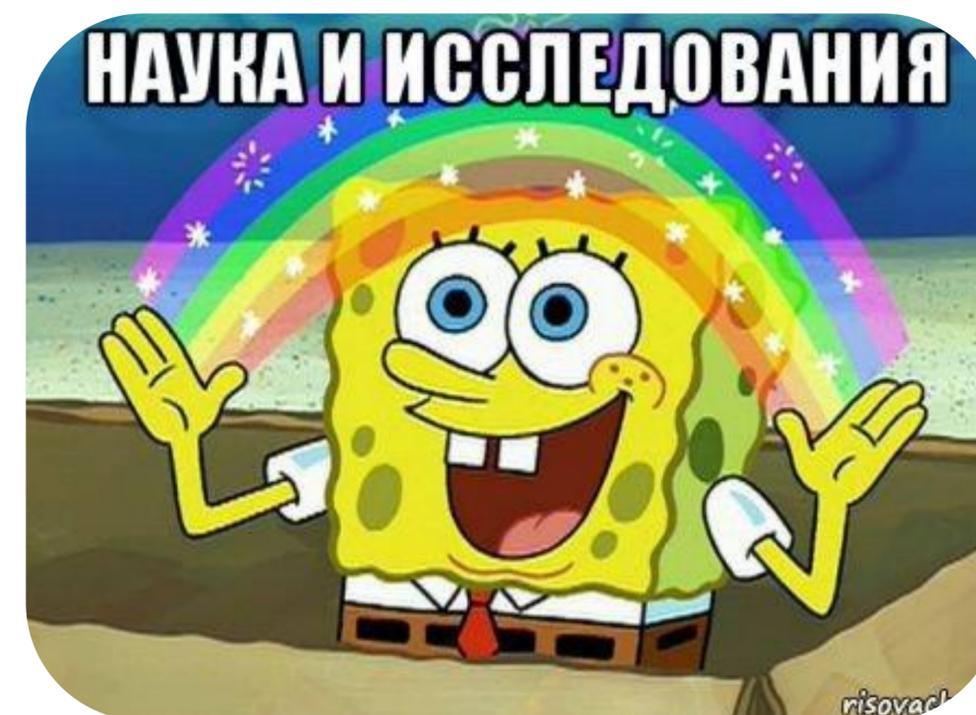
Добавления новых признаков

Расширения или сужения выборки

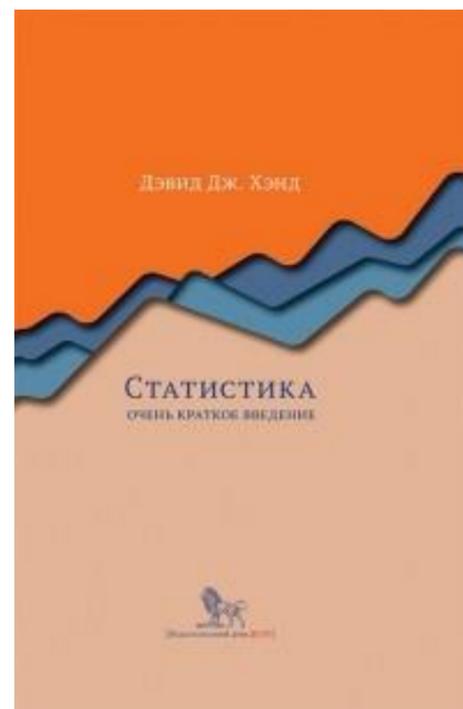
ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОГО ВЫПОЛНЕНИЯ

Определите единицу статистического наблюдения, предложите переменные, которые необходимо собрать для исследования (объясняющие, объясняемые, контрольные)

1. В какой школе выше уровень успеваемости?
2. Как менялся уровень преступности в России в течение последних 15 лет?
3. Есть ли взаимосвязь между безработицей и урбанизацией?



ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА





НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ