



Факультет экономических наук

Экономическая школа ФЭН

Москва 2023

Мир вокруг нас как источник данных для экономиста

Академический руководитель онлайн-программы магистратуры
«Экономический анализ»: Вакуленко Е.С., доктор экономических наук



Вакуленко Елена Сергеевна

- Путь от студента ВШЭ до доктора экономических наук
- Научный интерес в **data science**
- Почетный работник науки и высоких технологий РФ





План

- Специализации экономистов
- Зачем экономисту знания математики?
- Что такое Data Science в экономике?
- Где экономисты берут данные для построения моделей?
- Как экономисты используют Big data сети Интернет?
- Проводят ли экономисты эксперименты для проверки своих гипотез?
- Какие ключевые навыки нужны аналитику?



Три специализации экономистов (Кузнецов Б.В.)

- **Терапевты**
 - Терапевты занимаются анализом. Раскрывают черный ящик, изучают, как поток данных на входе в него связан с потоком на выходе. Аналитика — это своего рода постановка диагноза экономической болезни.
- **Хирурги**
 - Прогнозирование — это хирургия. Академик РАН Ю. В. Яременко: «Мы разрабатываем прогнозы, чтобы они не сбывались». Есть целый ряд задач, когда прогноз делается для того, чтобы было предпринято все возможное для его предотвращения.
- **Патологоанатомы**
 - Самая безобидная субпрофессия экономиста — изучать то, что уже произошло.



Зачем экономисту знания математических дисциплин?

Экономист должен уметь **объяснять** происходящие экономические события, **прогнозировать** будущие, **предлагать** стратегии развития.

Любое решение или прогноз должны быть обоснованы **модельными расчетами**.

Экономист анализирует данные. Для этого важно знание инструментов анализа, в основе которых лежит **математика**.

Handwritten mathematical notes on a chalkboard:

- $\lim_{x \rightarrow 1} \frac{c \ln x - c}{2 \ln x - 3}$
- $P = r^2 \pi$
- $\ln = \sqrt{ax+b}$
- $\sum_{i=1}^{n-1} h_i x_i = \dots$
- $e = 2,79$
- $P = \sum_{i=0}^{\infty} x_i^i$
- $A - C = \frac{C}{C}$
- $y = 2x^2 + 3x$
- $\sin \alpha$
- $\tan(\alpha) = \frac{2 \tan(\alpha)}{1 - \tan^2(\alpha)}$
- $15 \Delta t = T - \frac{3\alpha}{x}$
- $(x+y)^2 = \left(\frac{y}{2}\right)^2 = x^2 + 2ax + a^2$
- $x^2 + y^2 = z$
- $\frac{\Delta x}{\Delta y} = \lim_{\Delta y \rightarrow 0} \frac{\Delta x + 2}{\Delta y - 1}$
- $\ln(x) = \int_1^x \frac{1}{t} dt$
- $x^n = ?$
- $x_{1/2} = \frac{b \pm \sqrt{a-c}}{2}$



Данные и их анализ

*«Данные – это всего лишь исходный материал знаний»
(The New York Times).*

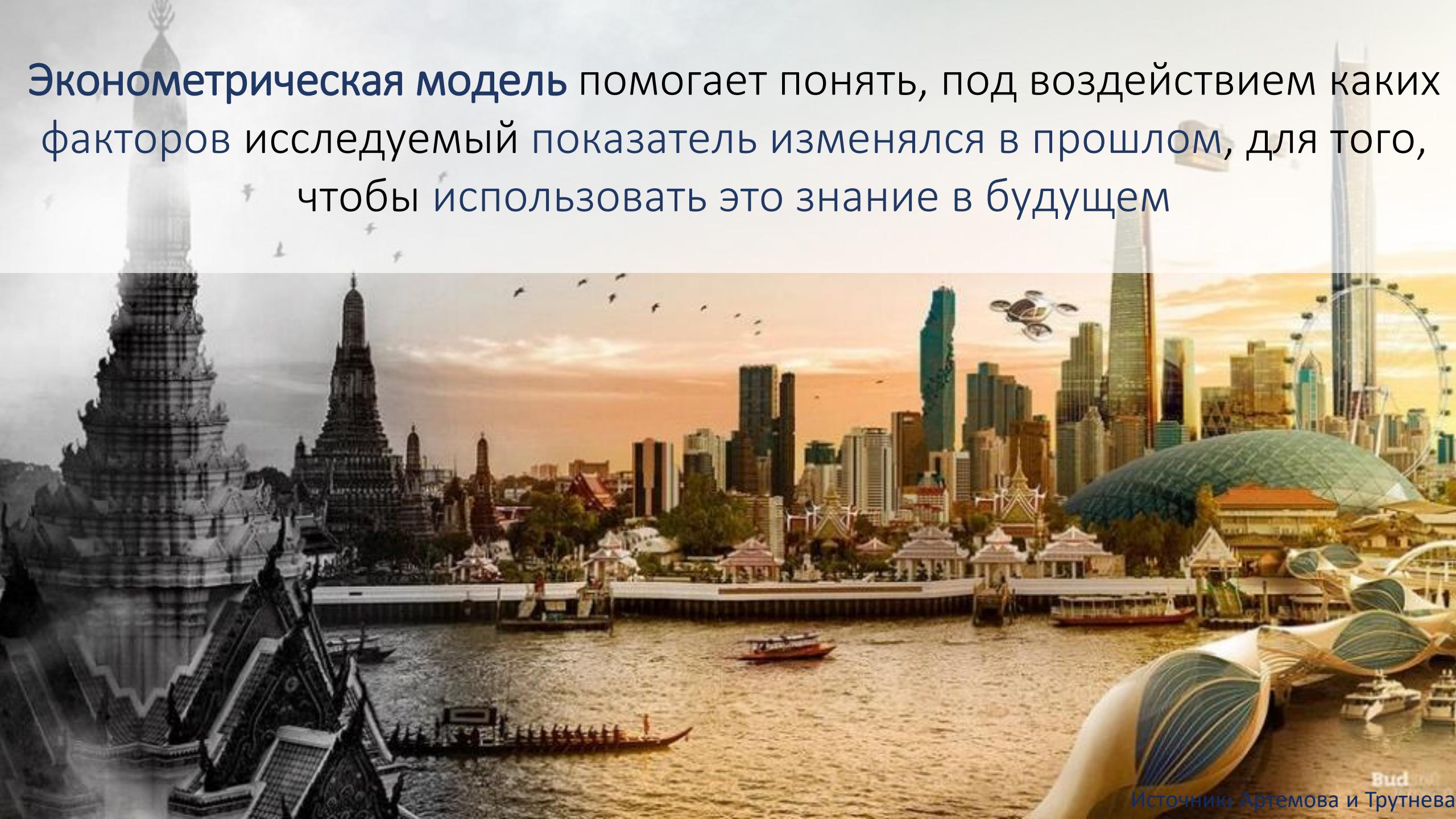
Data science – самый мощный инструмент для
практического использования информации.

- Статистика
- Эконометрика
- Машинное обучение

Data science



Эконометрическая модель помогает понять, под воздействием каких факторов исследуемый показатель изменялся в прошлом, для того, чтобы использовать это знание в будущем

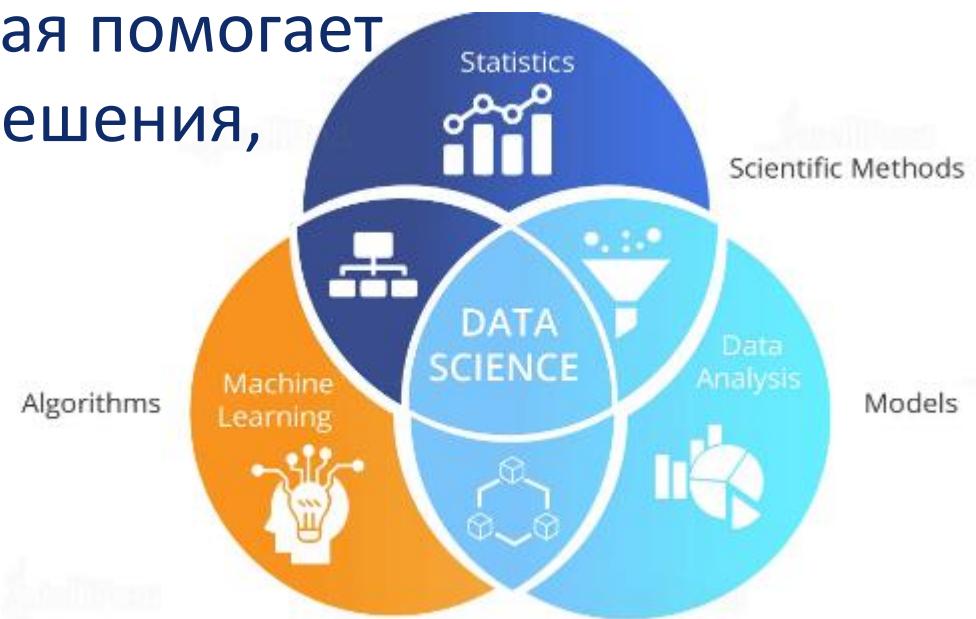




Data science в экономике

Экономисты все больше прибегают к анализу больших данных
для этого необходимо знание методов **машииного обучения**.

В сети Интернет **много информации**, которая помогает
экономистам прогнозировать, принимать решения,
строить стратегии.





Три самых важных качества аналитика

- Критическое мышление
- Умение работать с данными
- Интерес к задаче



90% работы в аналитике – это подготовка данных

- Найти источники информации
- Собрать из них данные
- Очистить и подготовить данные для анализа
- **Провести анализ**
- **Исследовать закономерности в количественном выражении и сделать выводы**
- Оформить результаты анализа



Источники данных для экономистов

- Наблюдения (отчеты, данные официальной статистики, данные в сети Интернет и т.д.)
 - Опросы
 - Эксперименты





Опросы (примеры)

1) Российский мониторинг экономического положения и здоровья населения НИУ ВШЭ

Ежегодный общероссийский опрос проводится с **1994** года более **4** тыс. домохозяйств, **10** тыс. индивидов.

Каждый год респондентам задают блок одних и тех же вопросов – панельные данные.

2) Опросы Института анализа предприятий и рынков НИУ ВШЭ

3) Бывает, что студенты ВШЭ проводят опросы для своих исследований.



Эксперимент со школьниками

Проект **STAR (Student/Teacher Achievement Ratio)**, США в 80-х годах (Krueger, 1999).

Цель - выяснить, помогает ли обучение в меньших группах достигать больших академических успехов.

Для того чтобы это выяснить, исследователям пришлось потратить 4 года и порядка 12 млн. долларов (и это в ценах 80-х годов) для проведения **эксперимента**. В нём было задействовано несколько тысяч американских школьников.

Когда они поступали в начальную школу, их **случайным образом** распределяли по классам разного типа: некоторые классы имели стандартный размер (22-25 человек), а некоторые — уменьшенный (13-17 человек).

Эксперимент со школьниками: результат

Обучение в маленьком классе увеличивает результаты школьника на итоговом тесте примерно **на 6 баллов (из 100)** при прочих равных.





О нас знают многое

Одна из самых **дорогих** компаний мира,
практически не имеет материальных активов.

Но имеет колоссальный актив – **ДАННЫЕ!**

Информация о пользователях:
место проживания, где и что покупают, как
проводят свободное время, с кем дружат и т.д.

*«Проблема информационной эпохи –
как распорядиться информацией».*



Крос К.



Где в Интернете брать информацию для анализа?

Яндекс

- Данные Гугл-трендов (**Google Trends**) и Яндекс статистики по запросам и поисковым словам;
- Данные социальных сетей (смена геолокаций, посты, реакция на посты);
- Исследование реакции на новости (оценка тональностей), в том числе в телеграмм каналах;
- Парсинг сайтов. Например, сбор информации о ценах на квартиры и их характеристики; или агрегация информации о вакансиях и предлагаемых заработных платах и т.д.

Какие задачи можно решать с помощью данных Интернета?

С помощью этих данных, например, можно прогнозировать:

- продажи жилья, автомобилей и других товаров,
- распространение заболеваний,
- миграцию населения,
- волатильность фондового рынка,
- информационное поле и потенциальные темы новостей.

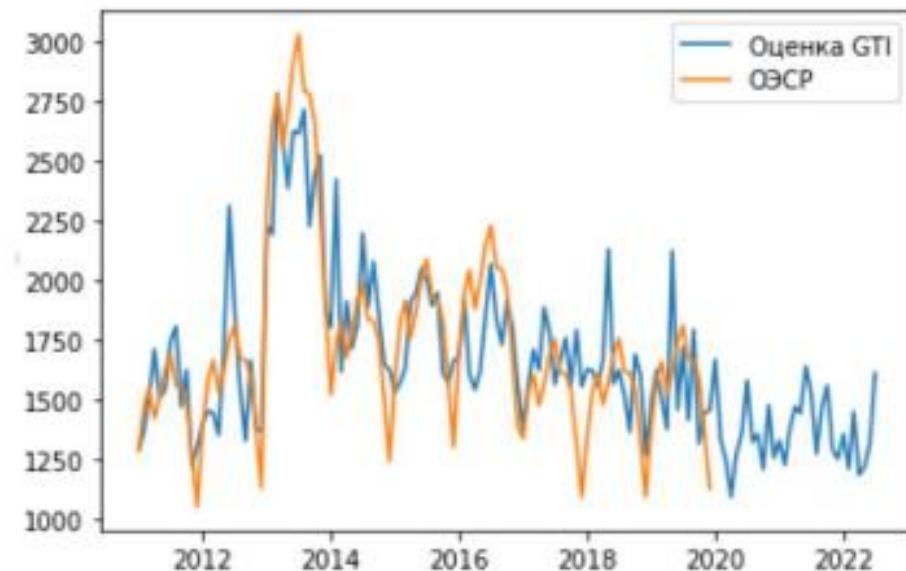


Пример использования данных Google-Трендов

Вакуленко Е. С., Броницкий Г. Т.
Прогнозирование миграции из
России в Германию с
использованием Google-трендов //
Демографическое обозрение. 2022.
Т. 9. № 3. С. 75-92.

Рисунок 9.

Прогноз миграционных
показателей (данные
Yandex Wordstat)



Источник: Данные Google Trends Index,
ОЭСР, Yandex Wordstat



Пример использования данных Google-Трендов

Таблица. Сравнение моделей (3) для прогноза миграции из РФ в Германию с помощью различных поисковых множеств

Источник поисковых запросов	Найденные запросы	MAPE	MAE	Значимые запросы*
Национальный корпус	20	14,7	271,3	визовый Германия право Германия гражданство Германия
Википедия	25	14,4	167,4	виза Германия билет Германия посольство Германия
Тайга корпус	25	14,4	265,8	виза Германия ВНЖ Германия иммиграция Германия
Yandex Wordstat	32	10,4	183,5	шенгенская виза в Германию посольство Германии в Москве работа в Германии

Примечание: * – Значимые запросы на уровне значимости 5%. Приведены примеры, соответствующие наибольшим значениям параметров β_i в модели линейной регрессии (4).
ВНЖ – вид на жительство.

Нужно ли экономисту уметь программировать?

- Для анализа данных нужно не только **владеть инструментарием, но иметь навыки работы в статистических пакетах (R, Python)**.
- Зачастую задачи, стоящие перед экономистам, **не имеют готовых инструментальных решений**. Здесь потребуются навыки программирования.
- Все рутинные процедуры можно **автоматизировать** (сбор и подготовка данных, подбор параметров и т.д.)





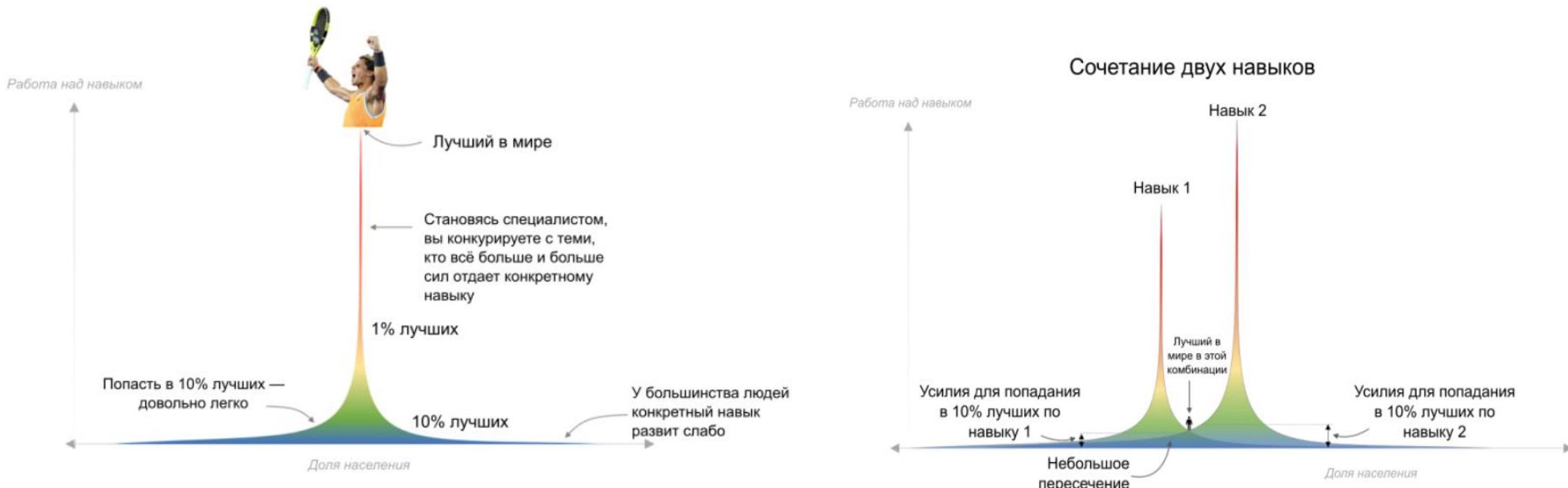
Советы будущим студентам бакалавриата

- Если не можешь определиться, где учиться,
– иди туда, где сложнее. Вместе с вами
будут учиться лучшие.
- Задавай много вопросов и много учись
самостоятельно.
- Мир меняется, постоянно учись новому.



«Комбинация хорошо развитых навыков сделает вас уникальным, даже если по каждому из них вы далеко не гений»

Усилия, необходимые, чтобы стать лучшим в одном навыке





Интервью с выпускниками ФЭН (аналитики):

- [Роман Абрамов](#) (руководитель направления «Анализ данных и развитие информационных систем» в Аналитическом центре Москвы)
- [Елизавета Витюк](#) (Партнер **McKinsey&Company**)
- [Елена Сидорова](#) (Senior Data Scientist **McKinsey&Company**)
- [Екатерина Трутнева](#) (руководитель отдела драйверов роста департамента маркетинга и коммуникации Сбера)
- [Елена Лебединская](#) (Директор департамента доходов Минфин)
- [Юрий Ачкасов](#) (Руководитель аналитического центра ДОМ.РФ)
- И многие другие





Что делать если у Вас есть сомнения в уровне Вашей математической подготовки?

- Бесплатный математический лагерь для абитуриентов бакалавриата ФЭН.
- Онлайн-курсы по высшей математике от НИУ ВШЭ.



MASTER DEGREE
ONLINE

Серия вебинаров программы:

**Профессия экономиста: от студента до
эксперта**

<https://www.hse.ru/ma/economicanalysis/>



Открытые
вебинары ФЭН





Телеграм канал ФЭН





Спасибо за внимание!