

Эксперименты в социальных науках и бизнес-практиках: зачем, где и как (примеры из маркетинга)?

Дмитрий Потапов,
НИУ ВШЭ

ЭКСПЕРИМЕНТАЛЬНЫЙ ПОДХОД

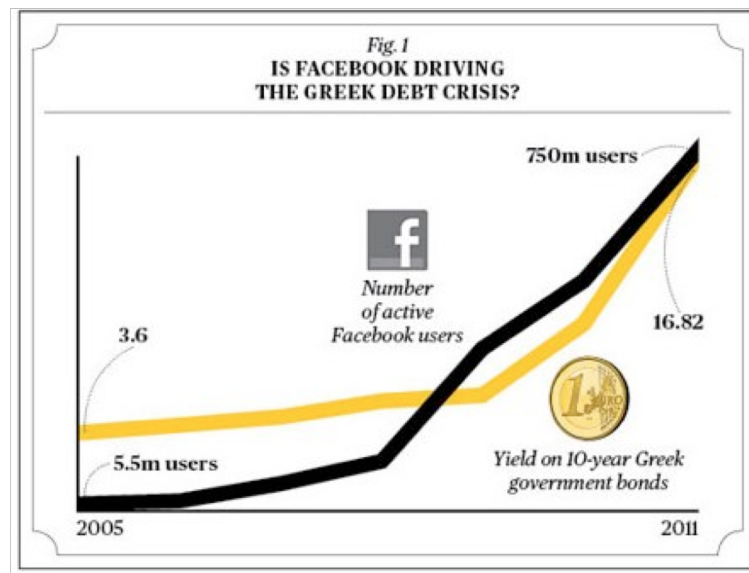
Типичные задачи, которые возникают в маркетинговой аналитике

- **Описание процессов**
 - Сезонность продаж
 - Мониторинг конкурентов
 - Потребительские предпочтения
- **Предсказание**
 - Каковы будут продажи в будущем месяце
 - Кто из клиентов отвалится в следующем году
- **Понимание причинности**
 - Если я установлю определенную цену, как это повлияет на выбор потребителя
 - Если я запущу новую рекламу, на сколько увеличатся продажи

Ответы на вопросы третьего типа определяют доказательные возможности маркетинга

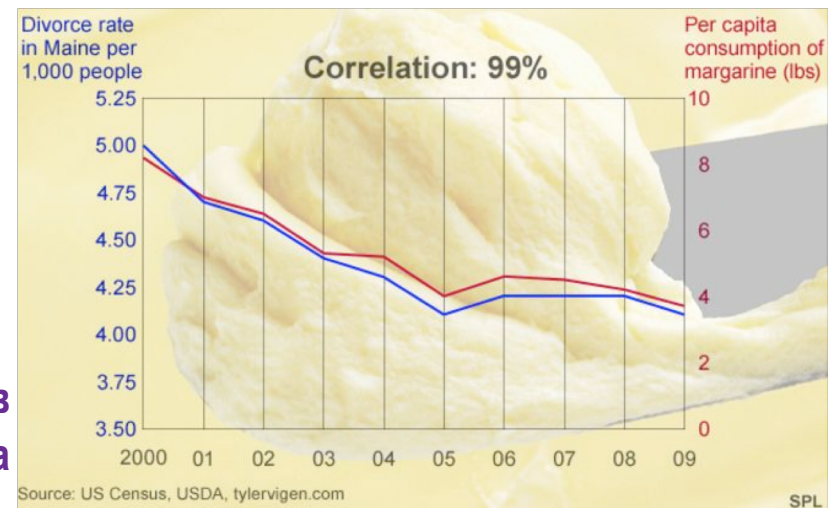
ЭКСПЕРИМЕНТАЛЬНЫЙ ПОДХОД.

КОРРЕЛЯЦИЯ ≠ ПРИЧИННОСТЬ



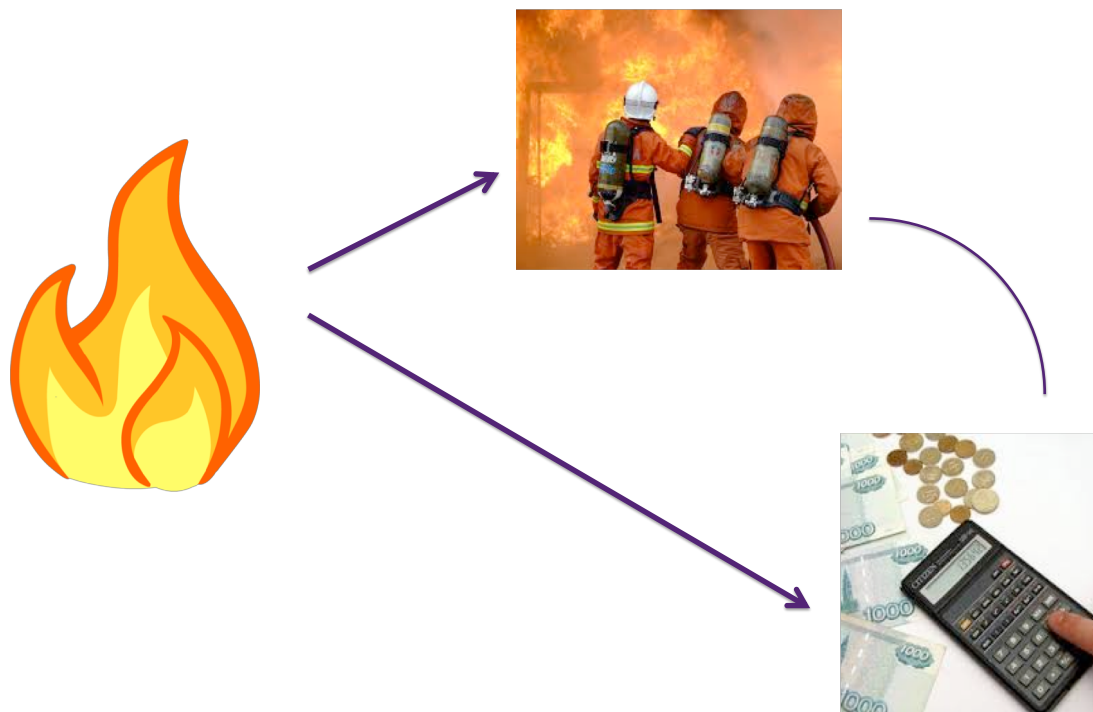
Связь между количеством пользователей Фейсбука и госдолгом Греции

Связь между числом разводов и потреблением маргарина



ЭКСПЕРИМЕНТАЛЬНЫЙ ПОДХОД.

КОРРЕЛЯЦИЯ ≠ ПРИЧИННОСТЬ



Связь между ущербом
от пожара и
количеством
пожарных,
задействованных в
тушении

ОЦЕНКА ПРИЧИННОСТИ - ПРОБЛЕМА

- **Задача - оценить эффект промо-акции на величину среднего чека**
 - Петров (П) получил промо
 - Васечкин (В) не получил промо

НО ЭТО НЕ РАБОТАЕТ!

- **Что можно сделать?**
 - Сравнить средний чек П и В с учетом их характеристик или
 - Поменять их местами в следующей промо-кампании

СНОВА НЕ РАБОТАЕТ!

ОЦЕНКА ПРИЧИННОСТИ – ВСЕ ЕЩЕ ПРОБЛЕМА

- **Что нас спасает?**

- Интерес к средним эффектам (не конкретные П и В, а средний потребитель)
 - Тогда можно сравнить чек тех, кто получает промо (подписан на рассылку) с чеком тех, кто не получает (не подписан).

НО ЭТО ОПЯТЬ НЕ РАБОТАЕТ!

- Люди и контексты выбора разные (не только по наблюдаемым характеристикам)
- Решения людей связаны с ожиданием результатов (эндогенны)

ШЕФ, ВСЕ ПРОПАЛО!

ЭКСПЕРИМЕНТАЛЬНЫЙ ПОДХОД.

ОЦЕНКА ПРИЧИННОСТИ – ПРИМЕР

Cust_ID	Recency	Promo	Revenue (NO promo)	Revenue (promo)	Effect
1	Fresh	1	1500	1700	200
2	Fresh	1	1400	1550	150
3	Fresh	1	1600	1800	200
4	Fresh	0	1500	1750	250
5	Old	1	1000	1090	90
6	Old	0	950	1060	110
7	Old	0	850	930	80
8	Old	0	1200	1320	120

ЭКСПЕРИМЕНТАЛЬНЫЙ ПОДХОД.

ОЦЕНКА ПРИЧИННОСТИ – ПРИМЕР

Cust_ID	Recency	Promo	Revenue (NO promo)	Revenue (promo)	Effect
1	Fresh	1	1500	1700	200
2	Fresh	1	1400	1550	150
3	Fresh	1	1600	1800	200
4	Fresh	0	1500	1750	250
5	Old	1	1000	1090	90
6	Old	0	950	1060	110
7	Old	0	850	930	80
8	Old	0	1200	1320	120
Average			1250	1400	150

ЭКСПЕРИМЕНТАЛЬНЫЙ ПОДХОД.

ОЦЕНКА ПРИЧИННОСТИ – ПРИМЕР

Cust_ID	Recency	Promo	Revenue (NO promo)	Revenue (promo)	Effect
1	Fresh	1	1500	1700	200
2	Fresh	1	1400	1550	150
3	Fresh	1	1600	1800	200
4	Fresh	0	1500	1750	250
5	Old	1	1000	1090	90
6	Old	0	950	1060	110
7	Old	0	850	930	80
8	Old	0	1200	1320	120
Average			1250	1400	150
Average	Fresh		1500	1700	200
Average	Old		1000	1100	100

ЭКСПЕРИМЕНТАЛЬНЫЙ ПОДХОД.

ОЦЕНКА ПРИЧИННОСТИ – ПРИМЕР

Cust_ID	Recency	Promo	Revenue (NO promo)	Revenue (promo)	Effect
1	Fresh	1		1700	
2	Fresh	1		1550	
3	Fresh	1		1800	
4	Fresh	0	1500		
5	Old	1		1090	
6	Old	0	950		
7	Old	0	850		
8	Old	0	1200		

ЭКСПЕРИМЕНТАЛЬНЫЙ ПОДХОД.

ОЦЕНКА ПРИЧИННОСТИ – ПРИМЕР

Cust_ID	Recency	Promo	Revenue (NO promo)	Revenue (promo)	Effect
1	Fresh	1		1700	
2	Fresh	1		1550	
3	Fresh	1		1800	
4	Fresh	0	1500		
5	Old	1		1090	
6	Old	0	950		
7	Old	0	850		
8	Old	0	1200		
Average			1125	1535	410

ЭКСПЕРИМЕНТАЛЬНЫЙ ПОДХОД.

ОЦЕНКА ПРИЧИННОСТИ – ПРИМЕР

Cust_ID	Recency	Promo	Revenue (NO promo)	Revenue (promo)	Effect
1	Fresh	1		1700	
2	Fresh	1		1550	
3	Fresh	1		1800	
4	Fresh	0	1500		
5	Old	1		1090	
6	Old	0	950		
7	Old	0	850		
8	Old	0	1200		
Average			1125	1535	410
Average	Fresh		1500	1683	183
Average	Old		1000	1090	90

ЭКСПЕРИМЕНТАЛЬНЫЙ ПОДХОД.

ОЦЕНКА ПРИЧИННОСТИ – ПРИМЕР

Cust_ID	Recency	Promo	Revenue (NO promo)	Revenue (promo)	Effect
1	Fresh	1		1700	
2	Fresh	1		1550	
3	Fresh	1		1800	
4	Fresh	0	1500		
5	Old	1		1090	
6	Old	0	950		
7	Old	0	850		
8	Old	0	1200		
Average			1125	1535	410
Average	Fresh		1500	1683	183
Average	Old		1000	1090	90

ЭКСПЕРИМЕНТАЛЬНЫЙ ПОДХОД.

ОЦЕНКА ПРИЧИННОСТИ – ПРИМЕР

Cust_ID	Recency	Promo	Revenue (NO promo)	Revenue (promo)	Effect
1	Fresh	1		1700	
2	Fresh	0	1400		
3	Fresh	1		1800	
4	Fresh	0	1500		
5	Old	1		1090	
6	Old	0	950		
7	Old	1		930	
8	Old	0	1200		
Average			1262.5	1380	118
Average	Fresh		1450	1750	300
Average	Old		1075	1010	-65

СМЕЩЕНИЕ ОТБОРА

- Мы наблюдаем разницу между получившими воздействие и не получившими, но не истинный эффект воздействия

$$\begin{aligned} E(Y_i | X_i = 1) - E(Y_i | X_i = 0) &= \\ &= E(Y_i(1) | X_i = 1) - E(Y_i(0) | X_i = 0) = \\ &= E(Y_i(1) | X_i = 1) - E(Y_i(0) | X_i = 1) + \\ &+ E(Y_i(0) | X_i = 1) - E(Y_i(0) | X_i = 0) = \\ &= ATT + E(Y_i(0) | X_i = 1) - E(Y_i(0) | X_i = 0) \end{aligned}$$

- Основная проблема
 - нет оснований ожидать, что при отсутствии воздействия результат для тех, на кого воздействовали, и тех, на кого не воздействовали, будет одинаковым.
- Убрать смещение отбора помогает рандомизация

РЕШЕНИЕ – РАНДОМИЗАЦИЯ!

- Только в тех случаях, когда мы можем полностью манипулировать воздействием
- Алгоритм **рандомизации**
 - Разбиение на экспериментальную и контрольную группы (СЛУЧАЙНО!)
 - Воздействие на экспериментальную группу
 - Эффект определяется разницей зависимой переменной (целевой метрики) между контрольной и экспериментальной группами
- Технологические основы современного маркетинга представляют широкие возможности для экспериментов

ЭКСПЕРИМЕНТАЛЬНЫЙ ПОДХОД.

ЭКСПЕРИМЕНТЫ И МАРКЕТИНГ

Концептуальный аспект

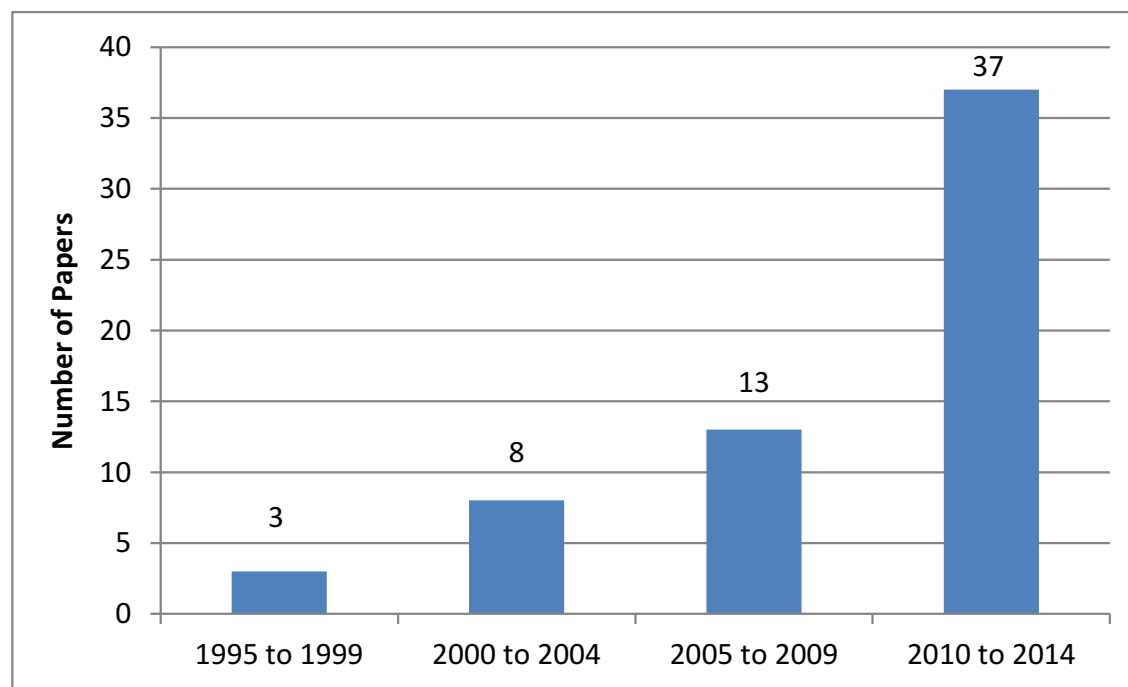
- персонализация - делать разным клиентам разные предложения
- на основе реакции понять разницу в предпочтениях и использовать в будущем

Технологический аспект

- прямые коммуникации
- наличие данных по воздействию (получил/открыл) и реакции (перешел/купил)

ЭКСПЕРИМЕНТАЛЬНЫЙ ПОДХОД.

ЭКСПЕРИМЕНТАЛЬНЫЕ СТАТЬИ В ТОП5 МАРКЕТИНГОВЫХ ЖУРНАЛАХ



This figure reports the number of papers published by 5-year period. The sample size is 61 (papers).

ЭКСПЕРИМЕНТАЛЬНЫЙ ПОДХОД.

МАРКТЕИНГОВЫЙ КОНТЕКСТ

- **Реклама и продвижение**
 - интернет-реклама
 - sms/e-mail
 - эффективность промо
- **Ценообразование**
 - офлайн&онлайн
- **Разработка продукта**
 - интернет-сервисы (посадочные страницы)
 - MacDonalds
- **Дистрибуция**
 - купи онлайн/забери в магазине

ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА

ТЕРМИНОЛОГИЯ

Фактор (factor) – переменная, которой манипулирует экспериментатор (может быть несколько)

Воздействие (treatment) – набор экспериментальных условий (значений факторов), которые применяются к объекту эксперимента

Отклик (response) – переменная, которую мы измеряем как результат эксперимента (может быть несколько)

Объект эксперимента (experimental unit) – единица, на которую оказывается воздействие

Объект наблюдения (observation unit) – единица, которую мы наблюдаем по итогам эксперимента

ШАГИ

1. Сформулировать цель
2. Выделить сегмент
3. Подобрать воздействие (измерения и уровни воздействия)
например: срок (1/2 недели), размер скидки (10/15/20%), пр.
4. Определить целевую метрику/и
5. Сформировать контрольную и экспериментальную/ные группы
6. Спланировать коммуникации
7. Организационно и технически обеспечить реализацию механики
8. Запустить эксперимент

ЭКСПЕРИМЕНТАЛЬНЫЙ ДИЗАЙН

Экспериментальный дизайн – способ формирования воздействий на основе факторов (в маркетинге – конджоинт-анализ)

Зависит от

- числа возможных комбинаций факторов
- эффектов, который мы хотим оценить (главных, взаимодействия)

Бывает:

- полный факторный / неполный факторный: ортогональный, сбалансированный; ...

ЭКСПЕРИМЕНТАЛЬНЫЙ ДИЗАЙН - пример

Три фактора
(А, В, С),
каждый
принимает по
два значения
(-1, 1)

Полный
факторный
дизайн

Профиль (воздействи е)	Фактор			2-е взаимодействие			3-е взаимоде йствие
	А	В	С	АВ	АС	ВС	АВС
1	1	1	1	1	1	1	1
2	1	1	-1	1	-1	-1	-1
3	1	-1	1	-1	1	-1	-1
4	1	-1	-1	-1	-1	1	1
5	-1	1	1	-1	-1	1	-1
6	-1	1	-1	-1	1	-1	1
7	-1	-1	1	1	-1	-1	1
8	-1	-1	-1	1	1	1	-1

ЭКСПЕРИМЕНТАЛЬНЫЙ ДИЗАЙН - пример

Три фактора
(А, В, С),
каждый
принимает по
два значения
(-1, 1)

Неполный
факторный
дизайн

Профиль (воздействи е)	Фактор			2-е взаимодействие			3-е взаимоде йствие
	А	В	С	АВ	АС	ВС	АВС
1	1	1	1	1	1	1	1
2	1	1	-1	1	-1	-1	-1
3	1	-1	1	-1	1	-1	-1
4	1	-1	-1	-1	-1	1	1

Влияние фактора А оценить нельзя

ЭКСПЕРИМЕНТАЛЬНЫЙ ДИЗАЙН - пример

Три фактора
(A, B, C),
каждый
принимает по
два значения
(-1, 1)

Неполный
ортогональный
факторный
дизайн

Профиль (воздействие)	Фактор			2-е взаимодействие			3-е взаимодействие
	A	B	C	AB	AC	BC	ABC
1	1	1	1	1	1	1	1
4	1	-1	-1	-1	-1	1	1
6	-1	1	-1	-1	1	-1	1
7	-1	-1	1	1	-1	-1	1

Можно оценить все главные эффекты

РАНДОМИЗАЦИЯ

- **ОСНОВНОЙ ПРИНЦИП** – каждый представитель целевого сегмента имеет одинаковый (заранее заданный) шанс попасть как в экспериментальную, так и в контрольную группы.
- Экспериментальная группа может быть одна, а может быть несколько (по количеству воздействий)
например, тестируем длительность акции (1/2 недели) и глубину скидки (10/20%), тогда экспериментальных групп будет четыре.
- Заранее нельзя точно определить необходимый размер выборки
слишком маленькая выборка не позволит судить о статистической значимости эффекта
необходимый размер выборки будет зависеть от силы эффекта, который заранее неизвестен.

КАК РАНДОМИЗИРОВАТЬ?

1. Определить количество групп.
2. Определить минимальный размер групп.
3. Сгенерировать случайную величину, которая позволит «разыграть» вероятности попадания людей в группы, исходя из требуемых пропорций.
4. Распределить людей по группам.
5. Проверить случайность распределения
есть ли разница в характеристиках людей (замеренных до эксперимента) между группами.

АНАЛИЗ РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТА

ШАГИ

1. Собрать данные по воздействиям и целевой метрике/ам
2. Провести анализ в целом (а также в разрезе уровней воздействия и/или в разрезе подгрупп)
 - a. Рассчитать разницу в целевой метрике для контрольной и экспериментальной групп
 - b. Проверить статистическую значимость
 - количественная переменная – t-тест
 - бинарная переменная – z-тест
 - c. Рассчитать инкрементальный эффект
3. Определить, как далее использовать полученные знания

ИНКРЕМЕНТАЛЬНЫЙ ЭФФЕКТ

- Как количественно оценить результативность конкретной кампании?
- Расчет инкрементального эффекта
 - Взять среднее значение целевой метрики экспериментальной и контрольной группы
 - Оценить, является ли разница между ними статистически значимой
 - Умножить разницу на размер экспериментальной группы

ЗАДАЧА НУГа

- предполагая набор возможных воздействий
- предполагая гетерогенную (неоднородную) реакцию разных клиентов на одно и то же воздействие
- в ситуации, когда эффект воздействия на индивидуальном уровне не может быть измерен в принципе

НАУЧИТЬСЯ ВЫДЕЛЯТЬ (И ОПИСЫВАТЬ) ГРУППЫ
КЛИЕНТОВ, КОТОРЫЕ ЗНАЧИМО ОТЛИЧАЮТСЯ С ТОЧКИ
ЗРЕНИЯ РЕАКЦИИ НА ВОЗДЕЙСТВИЕ

ПОЛЕЗНЫЕ МАТЕРИАЛЫ

Oehlert, G. W. (2010). *A First Course in Design and Analysis of Experiments*.

Gneezy, A. (2017). Field Experimentation in Marketing Research. *Journal of Marketing Research*, 54(1), 140–143

Athey, S., & Imbens, G. (2016). The Econometrics of Randomized Experiments. In *Handbook of Field Experiments* (Vol. 1, pp. 73–140

<https://conversionxl.com/blog/marketing-experiments/>

<https://hbr.org/2014/12/the-discipline-of-business-experimentation>

<http://runningres.com/>

https://www.youtube.com/playlist?list=PL1M5TsfDV6Vufqfs_h5fDR3pBhIj4QOW7

Учебник по экспериментам

Обзорная статья

Обзорная статья
по методам анализа экспериментов

Как правильно проводить маркетинговые эксперименты

Философия экспериментов в бизнесе

Организация экспериментов в социальных контекстах

Серия видео про проблему идентификации причинности