

Семинар 11: Где взять данные, если у вас нет бюджета на исследования?

Простой, но непонятный ответ: в Интернете.

Сегодня мы ответим на три вопроса:

1. Для решения каких задач вы можете использовать данные из Интернета?
2. Что это за данные и какие существуют бесплатные инструменты для их сбора?
3. Как использовать эти инструменты (на примере решения конкретных задач)?

Круг задач:

1. Анализ и сравнение популярности тем или объектов

Вопрос: Вам в наследство достался маленький подпольный цех по пошиву штор. Вы ничего не знаете о шторах, но хотите их продавать в городе N. Вы даже не знаете, какие цвета штор нравятся людям. Что делать?

Ответ: Попробуйте посмотреть, какого цвета шторы люди гуглят. Аналогично: размеры, материалы и т.д. Так вы сможете сформировать свою продуктовую матрицу, в которой будут шторы, которые действительно нужны людям.

2. Анализ сезонности

Вопрос: Есть товары с очевидной сезонностью (семена, цветы). Но если вы продаете мебель, то сезонность не так уж очевидна. Но важно проводить рекламные и стимулирующие акции в моменты низкого спроса (или немного заранее), чтобы сгладить денежный поток. Как это сделать?

Ответ: Посмотрите на сезонность в поисковых запросах по интересующей вас мебели.

3. Анализ географии

Вопрос: Вы крутой тренер по публичным выступлениям и хотите отправиться в турне по России (или даже по Белоруссии) со своими мастер-классами. Но вы не знаете, в каких городах люди больше всего хотят обучиться этому искусству. Что делать?

Ответ: Попробуйте посмотреть, а в каких городах люди чаще всего ищут что-то по этой тематике.

4. Контент-анализ употребления терминов

Вопрос: Вы делаете курсовую работу по корпоративной культуре и способах управления персоналом. Вы очень хотите сделать выводы о национальной специфике методов мотивации (или как это дело меняется во времени). Что делать?

Ответ: Стоит посмотреть, как часто используются слова наказание или поощрение в литературе, как соотношение этих терминов меняется во времени.

Какие данные и инструменты можно использовать?

Типы возможных данных:

1. Статистика по поисковым запросам людей.
2. Оцифрованный и размеченный архив книг.
3. Результаты исследований рынков и потребителей, которые есть в открытом доступе (или доступе НИУ ВШЭ).

Инструменты:

1. [Google Trends](#) – статистика по поисковым запросам людей.

Инструмент дает возможность выяснить, как часто определенный термин ищут по отношению к общему объему поисковых запросов в различных регионах мира и на различных языках.

2. [Google Correlate](#) — позволяет определять запросы, которые коррелируют между собой.

Также сервис позволяет отслеживать зависимости между количеством определенных запросов и произвольной функцией, описывающей какую-то тенденцию, заданную в форме дата/данные.



3. [Google Ngram Viewer](#) - инструмент для визуализации изменений в текстах с учетом времени.

Для справки: **N-грамма** — это просто последовательность из n элементов (звуков, слогов, слов или букв). На практике чаще имеют в виду ряд слов (реже — букв). Последовательность из двух последовательных элементов называют **биграмма**, из трёх элементов — **триграмма**.

Гугл оцифровал 15 миллионов книг с 1800 по 2008 год. Это больше 500 миллиардов слов! Из них на русском языке – 35 миллиардов.

Чтобы частота использования конкретного слова не зависела от постоянного роста числа книг (а значит и слов), для каждого года репрезентативно отобрали 6000 томов.

4. [Google Consumer Barometer](#) - бесплатный сервис по исследованию потребительского поведения интернет-пользователей

Самая полезная штука из барометра (особенно для онлайн-ритейла) - **The Smart Shopper**. Есть данные о том, что предшествует офлайн-покупке пользователя, насколько велика роль Интернета в принятии решений, сколько людей используют смартфоны для покупок.

Пример: 58% потребителей в России сравнивают товары и цены онлайн до момента покупки. Во **Франции так делают лишь 46%**. А говорят, что мы непродвинутые!

В Барометре много фильтров:

- Страна
- Возраст
- Пол
- Выбор устройства (например, только смартфоны, или только ПК)
- Количество используемых устройств

5. [Euromonitor](#) – глобальная база данных рыночных исследований. Доступ от НИУ ВШЭ.

То, что доктор прописал для ваших дипломов и курсовых: стратегические исследования индустрий, компаний, экономик и потребителей во всем мире (включая Россию).

Можно найти доли рынка крупнейших компаний в любой отрасли, объемы продаж по товарным категориям и многое другое.

А теперь к практике!

Google Trends

1. Начнем с простого. Связаны ли депрессия и время года? А время суток? Кстати, самые эффективные антидепрессанты снижают уровень депрессии всего на 20%.

Для справки:

Числа в Google Trends обозначают уровень интереса к теме с учетом наиболее высокого показателя на графике для определенного региона и периода времени.

- 100 баллов означают наивысший уровень популярности запроса;
- 50 — уровень популярности запроса вдвое ниже, чем в первом случае;
- 0 — уровень популярности запроса не выше 1% от уровня в первом случае.

2. Изольда решила стать ютуб-блоггером и снимать видосы про йогу. Что же ей выбрать для съемок?

Давайте сравним популярность запросов в ютубе: хатха йога, йога для похудения.

3. Помогла ли Игра Престолов Амедиатеке стать популярной?

4. Вернемся к шторам. Давайте посмотрим, что людям больше нравится: шторы или жалюзи? А в каких регионах любят больше шторы? А что с городами? А какие похожие запросы предлагает гугл?

Google Ngram Viewer

1. Насколько этот мир унылый и скучный? Давайте посмотрим на одиночество.

2. Хорошо ли работала цензура в СССР? Запрос: Троцкий, Зиновьев, Каменев.

3. А что популярнее наука или религия?

4. Какие предлоги используются с Украиной? *_ADP Украине

5. Отследим тренды в менеджменте и маркетинге.

Список запросов:

руководитель *_NOUN

менеджер *_NOUN

менеджмент *_NOUN

маркетинг *_NOUN

P.S. Что еще есть крайне полезное? Возможно, придется заплатить.

1. **СПАРК** (возможно, есть доступ НИУ ВШЭ)– информационная база о российских компаниях.

Есть финансовая информация о всех компаниях, которые сдают бухгалтерский баланс и финансовый отчет в налоговую.

2. **Системы онлайн-мониторинга СМИ и социальных сетей** (YouScan, IQbuzz).