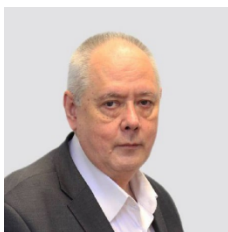


УДК 004.6

ОБУЧЕНИЕ СПЕЦИАЛИСТОВ-МАРКЕТОЛОГОВ РАБОТЕ С ДАННЫМИ В СОВРЕМЕННЫХ УСЛОВИЯХ



Н.В. Павлов

Доцент СПбПУ, кандидат технических наук
pavlov@kafedrapik.ru

Н.В. Павлов

Окончил Ленинградский электротехнический институт (1980). Кандидат технических наук (1991). Доцент СПбПУ (с 1994), Doctor of Science (2016, Abo Academi University, Finland). Исследования и преподавание маркетинговых исследований, информационных систем, расширенного интеллекта.

Аннотация. Статья посвящена опыту обучения маркетологов основам работы с данными для принятия решений. Показан спектр компетенций, необходимых для работы в области маркетинга, основанного на данных. Часть из них предлагается осваивать с помощью сквозной задачи, на которой демонстрируются способы сбора данных, их очистки, хранения, аналитики и мониторинга ситуации во внешней среде. Обсуждаются особенности задачи. Предлагаются варианты дальнейшего ее расширения.

Ключевые слова: аналитика, маркетинговые исследования, обучение, цифровизация, данные.

В настоящее время цифровой маркетинг признается трендом [4], но, по мнению автора, он должен вскоре стать органической частью маркетинга в целом.

Будущее – за маркетингом, основанным на данных.

В литературе описаны не только перспективы [15], но и сложности и провалы внедрения систем хранения и аналитики данных [12, 18].

Это значит, что для работы в этой области требуется серьезная подготовка специалистов.

По мнению автора, процесс перехода к маркетингу, основанному на данных [14], в реальных быстро изменяющихся условиях должен происходить постепенно, в развитии. Этот же путь целесообразен для стартапов.

Данная статья посвящена опыту обучения маркетологов основам работы с данными для принятия решений.

На каждом этапе работы с данными имеются свои сложности.

Реальных данных в свободном доступе для целей обучения не так уж много [7], но и их ручной поиск и сбор весьма трудоемки.

Организации практически не делятся своими внутренними данными.

Даже электронные магазины и маркетплейсы принимают меры против сбора их данных ботами. Предоставляемые маркетплейсами для этого API требуют для использования специальных навыков.

Собираемые данные обычно не готовы для хранения, а тем более – для анализа [9, 13]. Их требуется очистить и сохранить.

Имеются различные подходы к вопросу хранения очищенных или неочищенных данных [16]. Очевидно, это определяется особенностями решаемых задач.

Для аналитики нужны извлеченные из хранилища и подготовленные данные. Подготовка данных для принятия решений – тоже специфическая область, в основном касающаяся визуализации.

Следует отметить, что для решения описанного круга задач интенсивно разрабатывается новый инструментарий, основанный на искусственном интеллекте.

Таким образом, рассматриваемая область весьма обширна. Поэтому требуется комплексное обучение процессу перехода к цифровому маркетингу.

Важная часть такого перехода – аналитика: подготовка информации для принятия решений.

Другая цель – дать системные знания, которые позволят свободно осваивать новый инструментарий.

В источниках последних лет довольно трудно разобраться, так как обычно авторы считают, что у читателей уже имеется большой багаж знаний, например, о нормальных формах отношений [5], CSS [6], Regexp [9].

Часто приходится отвлекаться на поиск описания базовых понятий.

Кроме того, в настоящее время большинство работ посвящено лишь продвижению в Интернет, воронке продаж [1, 3, 8, 11].

Для освоения навыков работы с реальными данными предлагается следующая задача, проверенная в рамках курса Информационные системы в маркетинге в течение трех лет.

Вариант 1. Для себя: купить лучший товар определенного вида (робот-пылесос, фотоаппарат, мобильный телефон и т.п.) в хорошем магазине по хорошей цене.

Вариант 2. Для организации: в какой магазин (с лучшим рейтингом, лучшим ассортиментом, лучшими условиями сотрудничества) лучше отдавать свой товар для продажи.

Оба варианта основаны на сборе данных

- о товарах, в первую очередь – их рейтингах;

- о магазинах;

- о прайс-листах этих магазинов.

Выполнение задания даст понятие о маркетинге, основанном на данных, об основных методах работы с данными.

Решение этой задачи по подготовленным методическим указаниям занимает 8...10 академических часов.

Сущность задачи состоит в

- сборе данных,

- их очистке,

- сохранении в виде системы связанных таблиц,

- выводе данных для принятия решения,

- решении дополнительных задач, например, многокритериального выбора.

Инструментарий для этих действий весьма разнообразен.

В настоящее время сбор данных производится с помощью бесплатных расширений браузера Google Chrome.

Схема данных представлена на рисунке.

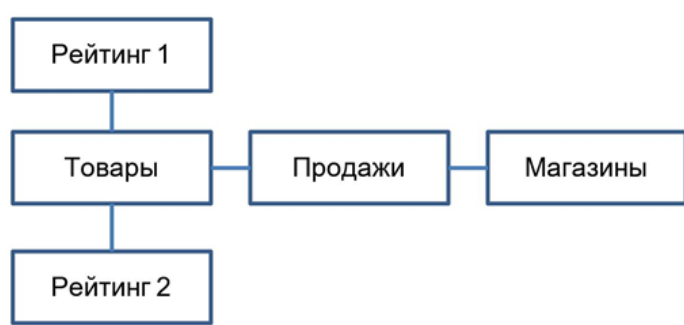


Рисунок 1. Схема данных

Примерный план занятий.

1. (2 академических часа.) Описание задачи. Сбор данных не менее, чем двух рейтингов заданного вида товара с помощью скрепера. Используется автоматический скрепер (Instant Data Scraper или автоматический режим NoCoding Data Scraper. Часто появляются новые, более удобные расширения). Результат – неочищенные данные, в которых, однако, просматриваются полезная информация: название продукта и место в рейтингах. Формат данных: .csv или .xlsx. Файлы показывают пример неочищенных данных.

2. (от 2 до 4 академических часов.) Создание т.н. рецептов сбора данных в расширении NoCoding Data Scraper. Его интеллектуальная поддержка упрощает эту работу. Задача состоит в том, чтобы в автоматическом режиме

- просмотреть список товаров на Яндекс.Маркете (сайт очень динамично меняется, сбор данных ботами постоянно усложняется); на Отзовике; на Shops-Prices.ru или аналогичном сайте;
- переходить по страницам этого списка;
- для каждого товара выбрать его название;
- для каждого товара перейти к списку магазинов, где он продается;
- для каждого магазина выбрать: цену продажи товара, название магазина, его рейтинг.

Для выполнения этих шагов требуется освоить механизм селекторов и управления работой скрепера. В частности, необходимо ввести задержки времени между действиями, чтобы имитировать работу человека-оператора. Процесс построения рецепта происходит достаточно наглядно: формируется графическая схема, имеется предварительный просмотр собираемых данных. Готовый рецепт обеспечивает переходы по страницам сайта.

Результат – большая таблица формата .csv или .xlsx. Но пока это – еще не полностью очищенные данные.

3. (2 часа.) Очистка данных с помощью языка M. Он включен в приложение MS Excel 2016 и более поздние версии, а также используется в бесплатном приложении Power BI Desktop. Результат – таблицы данных в соответствии с рисунком. Из неочищенных файлов удаляются лишние данные, преобразуется их формат, данные разделяются по содержанию. Результат – 5 таблиц, готовых к анализу.

4. (2 часа.) Связь таблиц в Excel или Power BI. Строится схема «снежинка» [14]. Тут возможен выбор. Либо продолжается прошлая работа в Power BI, тогда получается, что результат формируется из сохраненных неочищенных данных. Либо объединяются таблицы, взятые из сохраненных очищенных файлов.

Далее создается запрос, содержащий в табличном виде информацию для принятия решений: модель товара, его места в рейтингах, перечень магазинов, где он продается, цены в них, рейтинги этих магазинов.

На этом основная часть задания завершается.

Дополнительные действия состоят в обоснованном выборе лучшего товара и места его покупки с учетом рейтингов товара, рейтингов магазинов и цен в магазинах.

В конце задания полезно провести обсуждение проделанной работы и особенностей функционирования задачи в реальных условиях. Главная идея состоит в том, что проделанная работа – основа регулярного контроля текущей ситуации на рынке. Повторное выполнение запросов позволит следить за изменением ассортимента и цен.

Обсуждается удобный способ хранения исторических и актуальных данных для облегчения мониторинга внешней среды.

Обсуждаются и варианты применения собранных данных для аналитики. Это может быть отслеживание динамики цен, стабильности ассортимента, постоянства присутствия товаров в магазинах, выявление зависимости уровня цен от рейтинга магазина и т.д. Это продемонстрирует развиваемость цифровизации не только в ширину, но и в глубину.

В перспективе видится расширение задачи. Например, включение условий доставки, сравнение товаров по отобраным по важности характеристикам.

В результате на практике демонстрируется:

- решение аналитической задачи;
- варианты хранения данных;
- реалистичный путь цифровизации маркетинговой аналитики.

Следует, однако, отметить и отрицательную сторону описанного задания. Многократное автоматическое обращение к сайту с помощью ботов не совсем этично. Для учебных целей был бы полезен кэш: однократное скачивание страниц сайта и дальнейшая работа с их копиями. Такую возможность предоставляет система Octoparse, но опыт работы с ней показал, что она более сложна и даже нестабильна! Возможно, появятся новые, более удобные версии этой системы.

Другим решением данной этической проблемы могли бы стать API (Application Programming Interface), специально предоставляемые торговыми организациями для сбора данных с их сайтов. Однако это потребует освоения языка программирования. Впрочем, делаются попытки облегчить использование API непрограммистами [2]. Это – наиболее актуальное направление дальнейшего развития темы.

Опыт преподавания показал, что решение поставленной задачи, особенно в варианте 1, вызывает у студентов определенный интерес, хотя они иногда и говорят, что они не программисты. Полученные знания и навыки могут найти и другие полезные применения. Например, аналогичная задача – скачивание расписания с сайта университета и преобразование его в удобную форму.

Таким образом, предлагаемое задание позволит на практике ознакомить обучающихся с идеями аналитики реальных маркетинговых данных.

Список литературы

- [1]. Миллер Д. Воронки продаж по методу StoryBrand. – М. Альпина Паблишер. – 2022. – 105 с.
- [2]. Морин А. Из API Яндекс.Маркета -> в Excel и Power BI: интеграция. // Блог проекта Директ-ПРО. – URL <http://directprobi.ru/blogs/api-yandex-market-excel-power-bi-connector-partner-integraciya-zakazy-orders-products-fbs/>.
- [3]. Мультиановская Д.В. Маркетплейсы: как научиться продавать. Wildberries, Ozon и другие. – М.: АСТ. – 2022. – 180 с.
- [4]. Назаров А. Д., Товмасын Н. Д. Цифровой маркетинг как современный тренд. // Московский экономический журнал. – 2020. - №6. – с. 866 – 872.
- [5]. Нестеров С. А. Базы данных. – М.: Юрайт. – 2023. – 230 с.
- [6]. Сидельников Г. Наглядный CSS. – СПб.: Питер. – 2022. – 224 с.
- [7]. Смирнов Н. BIG DATA&AI 2022: если не ввести оборот больших данных, то «все пропало»? // Директор. – URL <https://cio.osp.ru/articles/280422-BIG-DATAAI-2022-esli-ne-vvesti-oborot-bolshih-dannyh-to-vse-propalo>.
- [8]. Солодар М. А. Настольная книга Интернет-маркетолога. – М.: Эксмо. – 2021. – 432 с.
- [9]. Сушков А., Шпрингер Л. Очистка данных: кто их загрязняет и что аналитику с этим делать // Блог Я Практикума. – 2022. – URL <https://practicum.yandex.ru/blog/chto-takoe-ochistka-dannyh/>.
- [10]. Форта Б. Изучаем регулярные выражения. М.: Диалектика. – 2019. – 192 с.
- [11]. Югова А. Продвижение ВКонтакте. – М.: Времена. – 2022. – 320 с.
- [12]. Craig S. Why a Majority of Data Warehouse Projects Fail—and What Businesses Can Do. // Database Trends and Applications. – 2021. – URL <https://www.dbta.com/Editorial/Think-About-It/Why-a-Majority-of-Data-Warehouse-Projects-Fail-and-What-Businesses-Can-Do-145910.aspx>.
- [13]. Joshua. The Ultimate Excel Data Cleaning Guide in 2023. // Quick Table. – 2023. – URL <https://www.quicktable.io/blog/the-ultimate-excel-data-cleaning-guide-in-2023-cldjhpp2a115191logkimzfavw>.
- [14]. Monberg C. 2023 Predictions For Data-Driven Marketers. // Forbes. – URL <https://www.forbes.com/sites/forbestechcouncil/2022/12/19/2023-predictions-for-data-driven-marketers/?sh=6f9142526a43>.
- [15]. Moses B. What's Next for Data Engineering in 2023? 7 Predictions. – 2022. // Search Medium. – URL <https://towardsdatascience.com/whats-next-for-data-engineering-in-2023-7-predictions-b57e3c1bf2d3>.

[16]. Reis J., Housley M. Fundamentals of Data Engineering. – O'Reily. – 2022. – 544 p.

[17]. Taniar D., Wenny R. Data warehousing and Analytics. – Springer. – 2021. – 642 p.

[18]. Vartika. Difficulties of Implementing Data Warehouses. // GeeksForGeeks. – 2023. – URL <https://www.geeksforgeeks.org/difficulties-of-implementing-data-warehouses/>.

TRAINING OF MARKETING SPECIALISTS TO WORK WITH DATA IN MODERN CONDITIONS

N.V. Pavlov

*Associate Professor of SPbPU,
Candidate of Technical Sciences*

Peter the Great St. Petersburg Polytechnic University

E-mail: pavlov@kafedrapik.ru

Abstract. The article is devoted to the experience of teaching marketers the basics of working with data for decision-making. The range of competencies required to work in the field of data-based marketing is shown. Some of them are proposed to be mastered with the help of an end-to-end task, which demonstrates how to collect data, clean it, store it, analyze and monitor the situation in the external environment. The specifics of the problem are discussed. Options for its further expansion are offered.

Keywords: analytics, marketing research, training, digitalization, data.