

«Сейфуллин окулары – 16: Жаңа формациядағы жастар ғылыми – Қазақстанның болашағы» атты халықаралық ғылыми-теориялық конференциясының материалдары = Материалы Международной научно-теоретической конференции «Сейфуллинские чтения – 16: Молодежная наука новой формации – будущее Казахстана. - 2020. - Т.ІІ. - С. 346-349

АНАЛИТИКА ДАННЫХ И ЕЕ ПРИМЕНЕНИЕ

Юсупов Н. Т.

Аналитика — это процесс обнаружения, интерпретации и информирование о найденных закономерностях в данных, а также использование средств, которые помогают всей компании анализировать любые данные в любой среде и на любых устройствах. Бизнес-аналитика предлагает и дополнительные возможности для достижения желаемых результатов, такие как оптимизация, снижение затрат и взаимодействие с заказчиками.

Сегодня каждая компания стремится получить как можно больше от аналитики и использует больше данных для ускоренного и более глубокого анализа для большего числа людей, и все это за меньшие деньги. Для достижения этих целей требуется надежная платформа, которая поддерживает весь процесс анализа с требуемой безопасностью, гибкостью и надежностью. Она должна помогать пользователям проводить анализ в режиме самообслуживания без потери общего контроля. Она также должна быть простой для администрирования. Но как получить преимущества системы корпоративного класса и инфраструктуры корпоративного класса без затрат? В наши дни бизнес-аналитика распространяется повсюду, так как каждая компания стремится повысить свои показатели, а значит, будет анализировать данные для принятия лучших решений.

С помощью бизнес-аналитики с персонализацией, машинным обучением и глубокими знаниями отрасли, компании могут получать актуальные результаты анализа данных по всем приложениям, хранилищам и озерам данных. Бизнес-аналитика должна представлять собой комплексный процесс, который приводит к конкретным действиям. Когда результаты анализа получены, компания может заняться переоценкой, изменениями и перенастройкой своих процессов. Главное здесь — довести до конкретного действия.

Ценность аналитики данных для бизнеса:

1. Новый подход к работе

Природа бизнеса меняется, и вместе с этим появляются новые направления конкуренции. Сегодня сотрудники хорошо разбираются в технических вопросах и не хотят, чтобы подготовка к работе занимала много времени. Обеспечивайте необходимую скорость и удобство для своих пользователей, одновременно поддерживая высокие стандарты качества и безопасности данных. Централизованная аналитическая

платформа, где ИТ играет ключевую роль, должна быть фундаментальной частью Вашей стратегии бизнес-аналитики. Залогом успешного внедрения инноваций является сочетание проектов, выполняемых бизнес-подразделениями и ИТ.

2. Новые возможности

Развитие технологий анализа данных создает новые возможности получения добавочной стоимости из своих данных. Современные средства аналитики данных позволяют выявлять скрытые модели данных благодаря возможностям прогнозирования, самообучения и адаптации. Эти технологии интуитивны и включают потрясающие визуализации данных, которые помогают мгновенно анализировать миллионы строк и столбцов. Современные аналитические решения удобны и могут использоваться на мобильных устройствах. Можно получать доступ к нужным данным в нужное время с минимальным освоением принципов работы и даже без них.

3. Визуализируйте свои данные

Конечно, Вы хотите видеть, о чем сигнализируют данные, раньше конкурентов. Аналитика данных обеспечивает четкую и ясную картину ситуации в бизнесе. Анализируя личные, корпоративные и большие данные, можно быстрее понимать их смысл, делиться с коллегами, и делать все это за считанные минуты.

4. Тенденции в аналитике данных

На постоянно развивающемся рынке средств анализа данных произошли важные изменения. Если раньше программы бизнес-аналитики возглавлялись службой ИТ, то теперь ИТ и бизнес-подразделения принимают решения совместно, и это становится новой нормой. Несомненно, аналитика уже стала стратегическим ресурсом для большинства компаний, что вызвало появление волны новых потребителей и новых ожиданий.

Мы полагаем, что перенос аналитики в облако — это значительно больше, чем просто выбор места развертывания. Такой путь устраняет барьеры между людьми, местоположением, данными и системами, что в корне меняет способ взаимодействия людей и процессов с информацией и технологиями, а также между собой.

Сравнение статистики и анализ данных появились до зарождения письменной истории, но необходимо было пройти несколько важных этапов для превращения аналитики в процесс, каким мы знаем его сегодня.

В 1785 г. Уильям Плейфэр (William Playfair) предложил гистограмму, которая сейчас является одним из основных (и широко используемых) способов визуализации данных. По легенде он изобрел гистограммы, чтобы показывать несколько десятков точек данных.

В 1812 г. картограф Шарль Жозеф Минар (Charles Joseph Minard) изобразил на графике потери армии Наполеона во время похода на Москву. Опираясь на польско-российскую границу, он создал линейную карту из

толстых и тонких линий, которая показывала, как потери связаны с суровой зимой и с тем, сколько времени армия была отрезана от путей снабжения.

В 1890 г. инженер Герман Холлерит (Herman Hollerith) изобрел «табулирующую машину», которая записывала данные на перфокартах. Это дало возможность анализировать данные быстрее, что сократило процесс подсчета в Бюро переписи населения США с нескольких лет до 18 месяцев. Тогда и появилась бизнес-потребность постоянно улучшать сбор и анализ данных. Что актуально и по сей день.

Современность: аналитика данных сегодня

В 1970-е и 1980-е годы появилось ПО реляционной базы данных и язык программирования SQL, что дало возможность экстраполировать данные для анализа по необходимости.

В конце 1980-х гг. Уильям Г. Инмон предложил концепцию «хранилища данных», где можно было получать доступ к информации быстро и неоднократно. Кроме того, Говард Дреснер, аналитик компании «Gartner», ввел термин «бизнес-аналитика» (business intelligence, BI), что подтолкнуло отрасли к анализу данных с целью лучшего понимания бизнес-процессов.

В 1990-х концепция глубинного анализа данных (data mining) дала возможность организациям анализировать и выявлять закономерности в огромных наборах данных. Аналитики и специалисты по обработке данных создали множество языков программирования, таких как R и Python, для разработки алгоритмов машинного обучения, работы с большими наборами данных и создания сложных визуализаций данных.

В 2000-х инновации в веб-поиске обеспечили разработку MapReduce, Apache Hadoop и Apache Cassandra для помощи в обнаружении, подготовке и представлении информации.

По мере того, как бизнес продвигался от простой доступности данных к их глубокому анализу, развивались средства анализа и их возможности.

Первые наборы аналитических инструментов были основаны на семантических моделях, взятых из ПО для бизнес-аналитики. Они помогали обеспечить эффективное управление, анализ данных и согласованность между инструментами. Одним из недостатков была недоступность своевременных отчетов. Принимающие бизнес-решения руководители не всегда были уверены в том, что результаты соответствовали их исходному запросу. С технической точки зрения эти модели используются в основном локально, что делает их неэффективными по затратам. Кроме того, данные часто оказываются изолированными в разрозненных хранилищах.

Компании, как крупные, так и малые, ищут лучшие способы использовать свои данные в качестве конкурентного преимущества. Об актуальности и важности аналитики данных говорят три факта:

1. Всего через пять лет количество «умных» устройств на планете будет более 50 млрд, каждый из которых будет генерировать данные, которые можно собирать и анализировать.

2. Белый дом вложил 200 миллионов долларов в проекты Bigdata – это свидетельство растущей важности и актуальности аналитики больших данных в различных секторах.

3. На данный момент анализируется и используется всего 5% доступных данных — огромный потенциал для развития.

Язык программирования Python в последнее время все чаще используется для анализа данных, как в науке, так и коммерческой сфере. Этому способствует простота языка, а также большое разнообразие открытых библиотек.

В этой статье разобрали простой пример исследования и классификации данных с использованием некоторых библиотек на Python. Для исследования, мы воспользовались готовым набором данных, который обычно представляет собой файл с таблицей в формате JSON или CSV. Мы не получили ответы на главный вопрос жизни, вселенной и всего такого, но показали простоту обработки достаточно большого объема данных средствами Python.

Анализ и обработка данных помогает выполнять задачи контроля выполнения намеченных планов, изучения различных ситуаций от потребительского спроса до соотношения долей конкурентов на рынке, а также прогнозирования и планирования.

С помощью аналитики компании могут предсказывать поведение клиентов, оптимизировать акции и цены, готовить маркетинговые кампании.

Самое важное в любой области, это систематизированные знания и сильная база, без которой работа будет продвигаться тяжело. В аналитике такой базой являются математика и статистика, умение работать с Excel/Google Таблицами, SQL и Python, способность выбрать подходящий для конкретной ситуации инструмент, а также понимание того, как визуализировать полученную информацию.

Список использованной литературы

1. Н.А. Назарбаев. Казахстанский путь. Изд. Атамекен. 2007 г.;
2. Н.А. Назарбаев. G-global: Мир 21 века. Изд. Атамекен. 2013 г.
3. Н.А. Назарбаев. В сердце Евразии. Изд. Атамекен. 2005 г.
4. К. Еременко. Работа с данными в любой сфере. Изд. Статистика как искусство. 2012 г.
5. А. Благирев. Big data простым языком. Изд. Белгородский вестник. 2010 г.
6. Р. Шапошников. Управление на основе данных. Изд. Наш вопрос. 2012 г.
7. А. Лепский. Аналитическая культура. Изд. Лепский и Со. 2017 г.
8. Wes McKinney. Python и анализ данных. Изд. O’Konnergy 2010 г.

9. Грейди Клейн. Алан Дебни. Статистика. базовый курс в комиксах. Изд. O’Konnerу 2015 г.
10. Charles Wheelan. Голая статистика. Изд. London and Columbus. 2016 г.
11. Эльдар Агамалов. Экспо 2017. Изд. Казахстанская правда. 2017 г.
12. Токмолдин Н. Источник энергии будущего. Изд. Казахстанская правда. 2014 г.
13. Махин В. В будущее – с новыми технологиями. Изд. Казахстанская правда. 2015 г.
14. Осипов В. Площадка для новых идей. Изд. Казахстанская правда. 2014 г.