**ЛЕКЦИЯ №13: ВИЗУАЛИЗАЦИЯ ДАННЫХ**

**Цель : Узнать что такое визуализация данных. Уметь применять ее на практике.**

**Визуализация данных** – это процесс использования визуальных элементов, таких как диаграммы, графики или карты, для представления данных. Он переводит сложные, масштабные или числовые данные в визуальное представление, которое легче обрабатывать. Инструменты визуализации данных улучшают и автоматизируют процесс визуальной передачи данных для обеспечения точности и детализации. Вы можете использовать визуальные представления для извлечения практических выводов из необработанных данных.

Современные предприятия обычно обрабатывают большие объемы данных из различных источников данных, таких как:

* Внутренние и внешние веб-сайты
* Смарт-устройства
* Внутренние системы сбора данных
* Социальные сети

Но необработанные данные бывает трудно интерпретировать и использовать. Следовательно, специалисты по работе с данными обрабатывают и представляют данные в правильном контексте: придают им визуальную форму, чтобы лица, принимающие решения, могли определить взаимосвязи между данными и обнаружить скрытые закономерности или тенденции. Визуализация данных создает истории, которые продвигают бизнес-аналитику и поддерживают принятие решений на основе данных и стратегическое планирование.

Преимущества использования визуализации данных:

**Стратегическое принятие решений**

Ключевые заинтересованные стороны и высшее руководство используют визуализацию данных для их осмысленной интерпретации. Они экономят время благодаря более быстрому анализу данных и возможности визуализировать общую картину. К примеру, они могут выявлять закономерности, обнаруживать тенденции и получать полную картину событий, чтобы обойти конкурентов.

**Повышение качества обслуживания клиентов**

Визуализация данных подчеркивает потребности и пожелания клиентов за счет графического представления. Вы можете выявить пробелы в обслуживании клиентов, стратегически улучшить продукты или услуги и снизить неэффективность операций.

**Повышение вовлеченности сотрудников**

Методы визуализации данных полезны для донесения результатов анализа данных до большой команды. Вся группа может совместно визуализировать данные для разработки общих целей и планов, а также оценки прогресса в работе и улучшения мотивации команды. Например, отдел продаж работает вместе, чтобы увеличить объемы продаж за один квартал.

Из каких компонентов состоит визуализация данных:

Специалисты по работе с данными выделяют три основных компонента.

**История**

История представляет собой цель визуализации данных. Специалист по работе с данными общается с несколькими заинтересованными сторонами, чтобы понять, чего они хотят достичь, анализируя данные. Например, они могут захотеть измерить ключевые показатели эффективности или спрогнозировать объемы продаж. Специалисты по работе с данными и бизнес-пользователи совместно определяют, какую историю они хотят узнать при помощи данных.

**Данные**

Специалисты по анализу данных определяют соответствующие наборы данных, которые помогут им изложить историю данных. Они изменяют существующие форматы данных, очищают данные, удаляют выбросы и проводят дальнейший анализ. После подготовки данных они планируют различные методы визуального исследования.

**Визуальные эффекты**

Затем специалисты по работе с данными выбирают методы визуализации, наиболее подходящие для обмена новыми сведениями. Они создают диаграммы и графики, выделяя ключевые точки данных и упрощая сложные наборы данных, а также думают об эффективных способах систематического представления данных для бизнес-анализа.

Из каких **этапов** состоит процесс визуализации данных?

Для эффективной визуализации нужно выполнить пять шагов.

Определите цель

Вы можете определить цель визуализации данных, выявив вопросы, на которые может ответить существующий набор данных. Четкая цель помогает определить тип:

* данных, которые вы используете;
* анализа, который вы проводите;
* визуальных эффектов, которые вы используете для эффективной передачи результатов.

 Например, компания розничной торговли может попытаться понять, какой тип упаковки продукта обеспечивает наибольший объем продаж.

**Соберите данные**

Сбор данных включает выявление внутренних и внешних источников данных. В Интернете доступны огромные наборы данных, которые можно приобрести и использовать. В вашей компании также могут быть доступны архивы данных для аналитики. Например, вы можете собрать исторические данные об объеме продаж, маркетинговых кампаниях и упаковке продукта, чтобы определить, какая упаковка лучше.

**Очистите данные**

Очистка данных подразумевает удаление избыточных данных, выполнение математических операций для дальнейшего анализа или фильтрацию и преобразование данных в соответствии с критериями вопроса. Например, вы можете удалить данные об объеме продаж в отпускные месяцы и после маркетинговых кампаний, чтобы определить средние продажи по типу упаковки.

**Выберите визуальные представления данных**

Для эффективного визуального обнаружения можно выбрать один из нескольких типов диаграмм. По отношениям между точками данных и идеями, которые вы хотите передать, можно определить наилучшие графические представления. Например, вы можете использовать гистограмму для представления продаж упаковки по цвету за последний месяц. Однако круговая диаграмма может лучше подходить для отображения процентного содержания цветной упаковки в ваших запасах. Существует два основных типа визуализации данных.

*Статическая визуализация*

Статическая визуализация обеспечивает только одно представление конкретной истории данных. Инфографика – это пример статической визуализации.

*Интерактивная визуализация*

С помощью интерактивной визуализации пользователи могут взаимодействовать с графиками и диаграммами. Зрители могут изменять переменные в параметрах визуализации, чтобы найти новые идеи или получить доступ к детальной информации. Программное обеспечение для визуализации данных обычно включает приборную панель для взаимодействия пользователя с системой.

**Создайте визуальные представления данных**

С помощью инструментов визуализации данных можно создать необходимые визуальные представления данных. Большинство инструментов импортируют конечный набор данных и автоматически генерируют необходимые отчеты. Ниже приведены некоторые принципы проектирования для эффективной визуализации данных:

* привлечение внимания аудитории к важным деталям с помощью размеров, цветов, шрифтов и графики;
* предоставление контекста для данных с помощью визуальных подсказок;
* выбор правильных цветовых комбинаций;
* использование пояснительных заголовков, чтобы донести ключевую информацию до аудитории и помочь ей сосредоточиться на правильных вопросах;
* добавление четких этикеток и цифр.

**Какие существуют типы методов визуализации данных**

Хотя диаграммы и графики являются наиболее распространенными, вы можете использовать несколько различных методов визуализации данных. Ниже представлены пять основных типов методов визуализации данных.

Визуализация временных данных

Визуализация временных данных используется для представления линейных одномерных объектов, таких как линейный график, линейная диаграмма или временная шкала. Например, вы можете использовать линейные графики, чтобы показать изменения, которые происходят непрерывно в течение определенного периода. Несколько линий на линейном графике демонстрируют колебания различных факторов за один и тот же период.

Визуализация иерархически структурированных данных

Визуализация иерархически структурированных данных относится к группе или набору элементов, которые имеют общие связи с родительским элементом. Вы можете использовать эти деревья данных для отображения кластеров информации. Например, вы можете показать количество данных инвентаризации в виде дерева с родительским узлом (одежда) и дочерними узлами (рубашки, брюки и носки).

**Визуализация сетевых данных**

Визуализация сетевых данных используется для представления сложных отношений между различными типами взаимосвязанных данных. Пример:

* Графики рассеивания, которые представляют данные в виде точек на графике
* Круговые диаграммы, которые добавляют третий фактор данных к графику рассеивания
* Облака слов, которые отображают частоту использования слов с помощью слов разного размера

Визуализация многомерных массивов данных

Визуализация многомерных массивов данных представляет две или более переменных данных в виде одного двухмерного или трехмерного изображения. Популярными примерами таких визуализаций являются гистограммы, круговые и составные столбчатые диаграммы. Например, гистограмма сравнивает два или более фактора данных и демонстрирует изменения одной переменной за определенный период времени. Круговые диаграммы визуализируют части целого по каждой категории.

Визуализация геопространственных данных

Визуализация геопространственных данных, например тепловых карт, карт плотности или картограмм, представляет данные в привязке к реальному местоположению.

Тест для самопроверки:

Независимая величина в графике количества осадков:
месяц
количество ясных дней
количество пасмурных дней
2. Что такое презентация в программе PowerPoint?
набор слайдов, подготовленный в программе для просмотра.
графические диаграммы и таблицы.
текстовой документ, содержащий набор изображений, рисунков, фотографий и диаграмм.
инструмент, который позволяет создавать картинки-слайды с текстом
3. Другое название графиков:
диаграммы-полосы
диаграммы-линии
диаграммы-вставки
4. Независимая величина в графике скорости:
время
скорость и время
скорость и расстояние
5. Классической диаграммой является:
столбовая диаграмма
столбчатая диаграмма
стволовая диаграмма
6. Иногда для оформления диаграмм используется такая визуализация, спроецированная на
плоскость:
трёхмерная
независимая
дополнительная
7. Тип диаграмм, схожий с линейными диаграммами способом построения кривых линий:
диаграммы-области
диаграммы-районы
диаграммы-округа
8. Сочетания диаграмм с географическими картами или схемами:
картодиаграммы
кардиограммы
картоидные диаграммы
диакарты
9. В каком расширении по умолчанию сохранятся презентация в PowerPoint?
рpt
jpg
pps
pdf
10. Каким образом изображаются значения в ярусных диаграммах?
Значения нескольких величин изображаются объединёнными в одном столбце
Значения каждой пары величин изображаются объединёнными в одном столбце
Значения каждой из величин изображаются в новом столбце
Значения каждой из величин изображаются в нескольких столбцах сразу
11. Что относится к графическим изображениям, применяемым в тексте?
картинки
диаграммы
схемы
все вышеперечисленное
12. Типичное количество переменных в диаграммах Венна — это ...
не более 1
не более 2
не более 4
не более 10
13. Достаточно распространённым способом графического изображения структуры
статистических совокупностей является такая диаграмма:
секторная
векторная
сектральная
14. Лучше всего демонстрирует соотношение двух различных компонентов, например, доли
импорта и экспорта тип диаграммы:
круговая
столбиковая
линейчатая
точечная
пузырьковая
15. Легенду на диаграмме нужно отображать:
всегда
никогда
по желанию
при использовании различных вариантов кодирования данных
при многомерном анализе данных
16. Что из себя представляет слайд?
абзац презентации
строчку презентации
параграф презентации
основной элемент презентации.
17. Для каких целей используется гистограмма?
для отображения зависимости одной величины от другой
для отображения величин частей целого
для сравнения нескольких величин, изображённых в виде горизонтальных или вертикальных
столбцов
для отображения нелинейных функций
18. В отличие от линейных диаграмм, в радиальных или сетчатых диаграммах:
менее двух осей
более одной осей
более двух осей
19. Если отсчёт производить не с центра круга, а с окружности, то такая диаграмма будет
называться:
линейной диаграммой
спиральной диаграммой
вьющейся диаграммой
20. Как называется средство наглядного графического представления количественных
данных, помогающее анализировать данные?
Чертеж
таблица
диаграмма
столбец