

Лекция 3. Визуализация данных

Зачем представлять данные графически?

Квартет Энскомба – четыре набора числовых данных, у которых простые статистические свойства идентичны, но их графики существенно отличаются. Каждый набор состоит из 11 пар чисел. Квартет был составлен в 1973 году английским математиком Ф. Дж. Энскомбом для иллюстрации важности применения графиков для статистического анализа, и влияния выбросов значений на свойства всего набора данных.

Все наборы обладают такими свойствами, указанными в таблице. Несмотря на то, что статистические свойства данных практически идентичны, графики показывают, что данные сильно различаются. Это говорит о необходимости представлять данные графически, перед тем, как делать выводы.

Человек - сложная биологическая система, тем не менее, восприятие им информации сильно ограничено, поэтому нужны простые её представления. Принципиальное значение для этого, для её интерпретации, имеет визуализация.

Визуализация - наглядное представление результатов анализа.

До сих пор ученые ведут исследования в области совершенствования современных методов представления данных в виде изображений, диаграмм и анимаций.

Визуализация очень важна в современном мире, особенно с большими объемами данных. Все потому, что во всей получаемой информации существует множество связей и зависимостей, просто так их очень сложно заметить. Кроме того, она дает ответы на многие вопросы гораздо быстрее. К примеру, по колонке цифр не все так хорошо ясно, как по графику.

На сегодняшний день рост информации достигает невероятных скоростей. Чтобы справиться с возрастающей сложностью и разнообразием данных необходимо использовать визуализацию. Смотреть на цветные картинки всем нравится больше, чем на скучные таблицы с цифрами. Вместе с тем, субъективное восприятие информации, доверие к ней намного выше, когда она подается визуально.

Рассмотрим области использования визуализации.

Статистика и отчеты. Данные за некий период времени показываются вместе. Например, статической картинкой в приложении к отчету или настраиваемым графиком в сервисе статистики, с возможностью изменения параметров его отображения.

Справочная информация. Дополнение к основному тексту, наглядно иллюстрирующее его упоминаемыми данными. Например, дать общее представление о динамике одного из показателей, либо отобразить какой-то процесс и его этапы; может быть - показать структуру некоего явления.

Интерактивные сервисы. Продукты и проекты, в которых инфографика является частью функциональности. Так, в качестве средства навигации по сервисам может являться диаграмма процесса. Почти все, что связано с работой с картами в специализированных системах вроде диспетчерских и большей части компьютерных

игр.

Иллюстрации. Красивое отображение данных для создания самостоятельных иллюстраций.

Чертежи и схемы. Специализированные документы, показывающие структуру и процесс работы сложных инженерных и природных систем.

Эксперименты и искусство. Визуализация данных в виде сложных и громоздких изображений, которые сложно “прочитать” бегло - объем данных и взаимосвязей между ними таков, что нужно разбираться с картинкой по частям; либо просто абстрактные изображения, автоматически сгенерированные. В последнее время направление все более популярно и периодически выходит за рамки компьютерной графики - например, в виде графиков-скульптур.

Типы визуализации

Выделим условно три типа визуализации.

- **Научная визуализация.** При моделировании различных объектов или процессов появляются большие объемы данных.
- **Информационная визуализация.** Описание/представление некой абстрактной информации, полученной при сборе и обработке многокатегориальных данных, для анализа которых необходимо применять различные количественные и качественные меры оценки.
- **Визуализация работы программного обеспечения.**

Задачи визуализации

Визуализация Big Data имеет определенные задачи:

- визуализация потоков данных;
- визуальный интеллектуальный анализ данных;
- визуальный поиск и рекомендации;
- описание ситуаций на основе больших данных с использованием визуализации;
- масштабируемые методы параллельной визуализации;
- современные аппаратные средства и архитектуры для анализа и визуализации данных;
- человеко-компьютерный интерфейс и визуализация больших данных;
- приложения визуализации больших данных.

Требования к визуализации

Можно сформулировать требования к такого рода визуализации:

- оценка пригодности (адекватности в визуализации) видов отображения,
- естественность (привычность для пользователей),
- устойчивость к масштабированию,
- возможность вывода сверхбольших объемов данных,
- возможности для представления сложных структур, а также объектов особого интереса, особых точек, аттракторов, сингулярностей.

Как решить проблемы визуализации Big Data?

Необходимо:

- использовать достаточно простые, но четко интерпретируемые методы представления.

- формализовывать описания объектов программного обеспечения
- производить верификацию и валидацию визуализации.

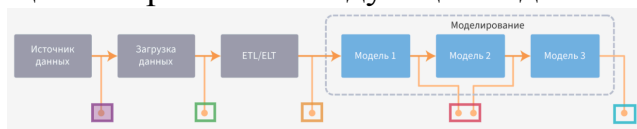
Способ визуализации должен максимально полно отражать поведение данных, содержащуюся в них информацию, тенденции, закономерности и так далее.

При этом выбор способа визуализации зависит от характера исследуемых данных и от задачи анализа, а также от предпочтений пользователя.

Многие связывают визуализацию только с интерпретацией, оценкой качества и достоверности результатов анализа. Однако это в корне неверно.

На практике в процессе анализа данных пользователь непрерывно работает с различными визуализаторами.

На разных этапах аналитического процесса визуализация используется для достижения следующих целей и решения следующих задач.



Визуализация источника

- идентификация
- оценка качества
- выдвижение гипотез

Визуализация загруженной выборки

- проверка результатов загрузки
- оценка качества данных

Визуализация результатов предобработки

- проверка результатов предобработки
- оценка готовности данных к анализу

Визуализация промежуточных результатов

- проверка корректности моделей
- контроль правильности результатов

Визуализация результатов анализа

- интерпретация результатов
- оценка достоверности результатов

Визуализация источников данных

В источнике требуется визуально оценить:

- характер, тип и поведение данных;
- динамический диапазон значений;
- степень гладкости;
- наличие факторов, снижающих качество данных, таких как шумы, аномальные и пропущенные значения.

Визуальный анализ позволяет:

- увидеть, соответствуют ли данные ожидаемым;
- оценить степень пригодности данных к анализу;
- выдвинуть гипотезы о закономерностях процессов, описываемых данными;
- определить, какие виды очистки и предобработки необходимо применить к

данным.

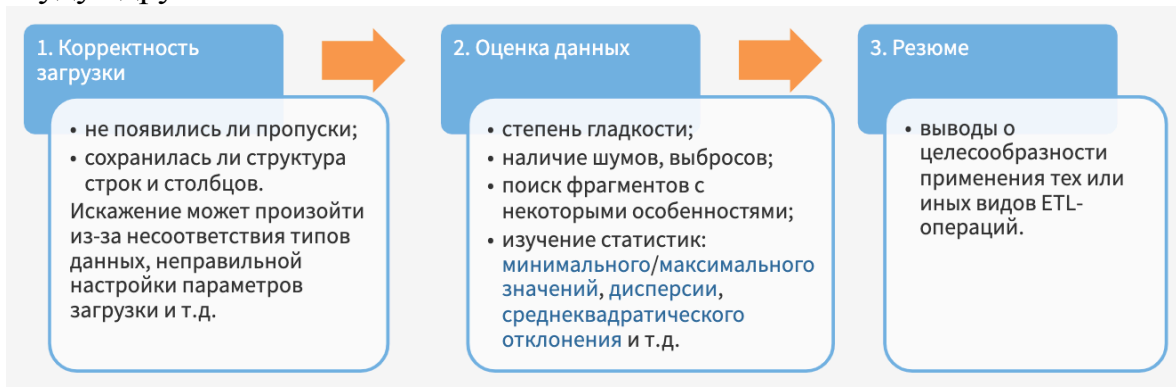
Кроме того, визуализация источников данных позволяет определить метод загрузки данных в аналитический контур и параметры, которые при этом должны быть использованы.

Например, для корректной загрузки данных из текстового файла с разделителями, необходимо правильно определить символ-разделитель, используемый формат даты и времени, расположение заголовков столбцов и так далее.

Для визуализации источников данных можно использовать приложения, в которых они были созданы (текстовые редакторы, СУБД, электронные таблицы и др.). Кроме того, аналитические платформы содержат собственные средства предварительного просмотра источников данных.

Визуализация данных, загруженных в аналитический контур

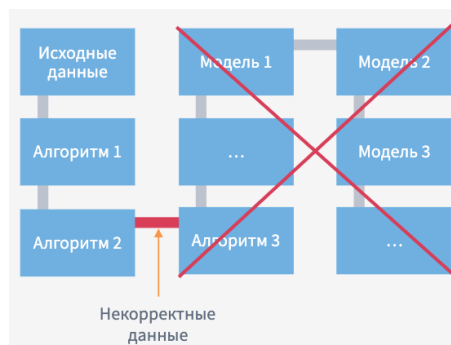
После загрузки данных из первичного в централизованный источник (хранилище или витрина данных, оперативный склад данных, MDM-система и так далее) работа с выборкой также начинается с **визуального анализа**. Однако теперь его цели, задачи и методы будут другими.



Визуализация данных в процессе аналитической обработки

Сложные аналитические процедуры являются многошаговыми, т.е. в процессе анализа к данным последовательно применяется несколько алгоритмов или моделей. Например, сначала данные подвергаются предобработке с целью сглаживания и кодирования, затем к результирующей выборке применяется та или иная модель.

При этом выборка, формируемая на выходе каждого алгоритма или модели, может подаваться на вход следующего этапа обработки. Очевидно, что если данные, поступившие с предыдущего этапа, окажутся некорректными, то дальнейшая обработка теряет смысл.



Поэтому очень важно предусмотреть визуализацию промежуточных результатов анализа с целью проверки корректности используемых моделей и алгоритмов.

Визуализация результатов анализа

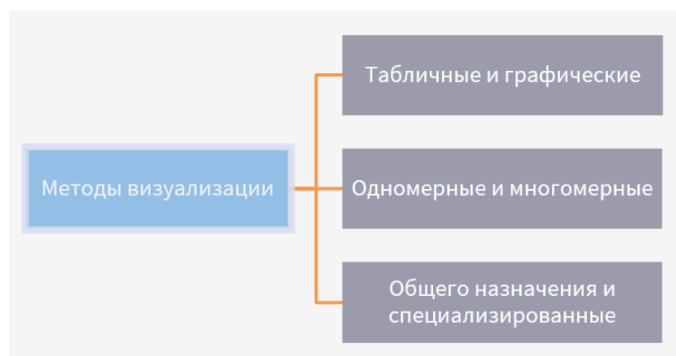
После получения конечных результатов аналитической обработки на первый план

выходит задача их интерпретации и оценки достоверности. И здесь не обойтись без визуализации.

Следует заметить, что, даже если в процессе анализа были получены достоверные и ценные результаты, неудачный выбор визуализации не позволит их интерпретировать, увидеть в них зависимости и закономерности.

Методы визуализации

В настоящее время в бизнес-аналитике и **Data Science** используется несколько десятков основных методов визуализации. Выбор метода определяется особенностями и характером данных, спецификой решаемой задачи и, наконец, предпочтениями пользователя. Рассмотрим методы визуализации, приняв за основу следующую классификацию



Табличные и графические

Как правило, таблицы применяются в том случае, когда пользователю необходимо работать с отдельными значениями данных, вносить изменения, контролировать форматы данных, пропуски, противоречия и так далее.

Графические методы позволяют лучше увидеть общий характер данных — закономерности, тенденции, периодические изменения. Кроме того, графические методы более эффективно сопоставляют данные: достаточно построить графики двух исследуемых процессов на одной системе координат, чтобы оценить степень их сходства и различия.

Одномерные и многомерные

Одномерные визуализаторы представляют информацию только об одном измерении данных, в то время как многомерные — о двух или более. Если график показывает зависимость суммы продаж от даты, то он будет одномерным, поскольку на нем будет отображаться только одно измерение — Дата, значениям которого будет соответствовать факт **Цена**. Если же информация о продажах приводится по датам и наименованиям товаров, то появляется еще одно измерение — **Товар**, и тогда для корректного представления данных используется многомерный визуализатор.

Популярные многомерные визуализаторы: OLAP-куб, тепловая карта и др.

Общего назначения и специализированные

Методы визуализации *общего назначения* не связаны с каким-либо определенным видом задач анализа или типом данных и могут использоваться на любом этапе аналитического процесса. Это своего рода *типовые визуализаторы* : графики и диаграммы, графы, гистограммы и их разновидности, статистические характеристики и др.

В то же время существует ряд задач, специфика которых требует применения *специализированных визуализаторов*.

Например, карты Кохонена специально разработаны для визуализации результатов

кластеризации, матрицы классификации используются в основном для проверки состоятельности классификационных моделей, а с помощью диаграмм рассеяния оценивается корректность работы регрессионных моделей.

При изучении различных видов визуализации удобнее рассматривать их не по отдельности, а в контексте задач, для которых они наиболее часто применяются. Можно выделить следующие **группы методов визуализации** (одна из классификаций методов визуализации).

Общего назначения

Применяются для решения типовых задач анализа данных: визуальной оценки качества и характера данных, распределения значений признаков, статистических характеристик и так далее. В них можно выделить два подвида — простые и сложные. К последним, в частности, относится OLAP-анализ — комплекс методов для визуализации многомерных данных.

К простым относятся следующие: таблицы, графики, диаграммы и гистограммы.

Обычно **сложные визуализаторы общего назначения** — это многомерные визуализаторы, которые позволяют представить информацию из нескольких измерений за счет использования различных цветов, форм, размеров и расположения объектов анализа. В последнее десятилетие получили широкое распространение следующие визуализаторы (см. схему).

Оценка качества моделей

Позволяет оценивать различные характеристики моделей, такие как точность, эффективность, достоверность результатов, интерпретируемость, устойчивость и так далее.

Интерпретация результатов анализа

Служат для представления конечных результатов анализа в виде, наиболее удобном с точки зрения их интерпретации пользователем.

Подсистемы визуализации данных содержатся не только в специализированных аналитических платформах, но и практически во всех программных средствах, которые связаны с обработкой данных, — от офисных приложений до систем компьютерной математики. Однако в аналитических платформах визуализации данных уделяется особое внимание, поскольку она является одной из составляющих аналитического процесса, без которой невозможно эффективно решать поставленные задачи.

Наилучших результатов можно добиться, если считать визуализацию не отдельной подсистемой, а такой же частью аналитического процесса, как, например, подготовка данных, аудит и профайлинг, моделирование. Даже если для построения качественной модели данных недостаточно, визуализация позволяет выдвигать гипотезы, делать выводы на основе экспертных оценок, разрабатывать способы повышения информативности данных.

Рассмотрим **современные виды визуализации** подробнее.

- Таблицы
- Графики и диаграммы
- Инфографика и схемы
- Презентация и анализ данных
- Интерактивный сторителлинг
- Бизнес аналитика и дашборды
- Научная и медицинская визуализация

- Карты и картограммы

Визуализация в медицине и науке

Специфический вид визуализации. Его целью обычно является выделение закономерностей или аномалий. От обычной визуализации данных отличается тем, что часто бывает трёхмерной и требует специальной подготовки для интерпретации.

Таблицы

С такой формой представления, как таблица, мы уже знакомы. Это простой и удобный способ отображения однотипной информации, кроме того, структурированные данные в первичных и вторичных источниках выгружаются именно в табличной форме.

Например, в таблице приведена история продаж.

Чек	Дата	Товар	Цена	Кол-во
1	01.05.2020	1235	258	1
1	01.05.2020	5648	325	2
2	01.05.2020	2485	924	1
3	01.05.2020	1878	157	1
3	01.05.2020	9759	654	1

У таблицы как визуализатора, существует множество вариантов представления. На рисунке на слайде изображена так называемая кросс-таблица с подсвечивающимися ячейками (англ.: highlight table). В ней каждому значению признака (продажи) соответствует один из оттенков в заранее выбранной цветовой гамме.

Такое представление помогает оперативно увидеть низкие и высокие зоны значений (продажи). Приставка кросс- означает, что таблица содержит по одному измерению в строках и столбцах, в данном случае это Месяц и Объект.

Объект	Январь	Февраль	Март	Апрель	Май	Июнь	Июль	Август	Сентябрь	Октябрь	Ноябрь	Декабрь
Объект 1	146 753	141 794	140 665	123 484	74 282	73 461	68 610	64 914	94 294	118 492	124 655	131 911
Объект 2	528 413	521 445	347 460	565 013	416 415	313 500	392 340	388 313	414 870	500 303	526 193	580 080
Объект 3	679 508	639 218	630 953	576 158	440 803	380 115	508 695	516 188	534 945	658 290	718 545	707 025

Графики и диаграммы

Наверное, самый привычный вид визуализации данных (рисунок 1).

Используется как для презентации данных, так и для анализа. Встретить их можно и на работе, и в журнале, и в научном отчете. Обычно знания о существующих типах диаграмм и графиков мы получаем из школы или из стандартного набора в Excel. Однако мир графиков и диаграмм не ограничивается точечным графиком, столбиковой и круговой диаграммой. Существуют порядка 15 общеизвестных типов диаграмм, а всего их более 60, при этом их количество увеличивается с каждым днём - люди придумывают новые типы для визуализации сложных и необычных данных.

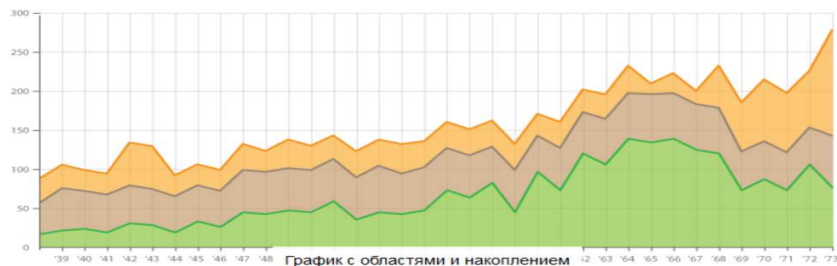
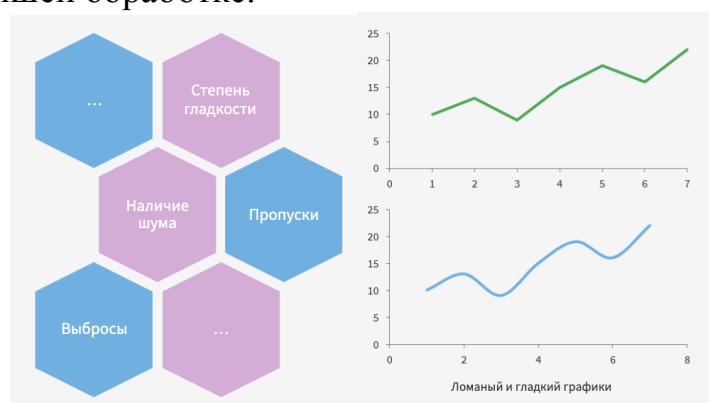


Рисунок 1 - Пример визуализации в виде графика

Графики представляют собой линии, отображающие зависимость между несколькими переменными в некоторой системе координат. Линия на графике состоит из множества точек, положение каждой из которых определяется значениями зависимой и независимой переменных (переменных).

Зачастую именно с построения графика и начинается работа с данными. С помощью графиков оценивается ряд характеристик данных, которые определяют их готовность к дальнейшей обработке.



Особенно они полезны при анализе временных рядов. Иногда одного взгляда на график достаточно, чтобы выявить наличие тренда, сезонной компоненты, оценить степень влияния случайной составляющей на исследуемый процесс.

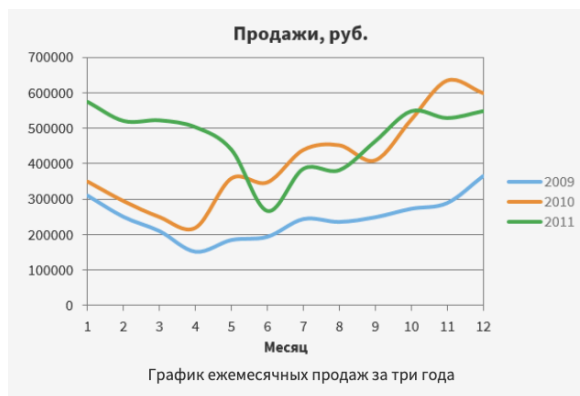
Чтобы построить график, достаточно задать таблично значения зависимой и независимой переменных, отметить соответствующие точки на координатной плоскости и соединить их линиями. Линии, соединяющие узлы графика, могут быть прямыми или сглаженными.

На слайде представлены примеры ломаного (вверху) и сглаженного графика (внизу).

Во многих случаях гладкие графики удобнее для визуального восприятия и корректнее отображают реальные бизнес-процессы, которые также чаще всего изменяются плавно. Иногда точки, по которым строится график, вообще не соединяют, в этом случае график называется точечным.

Если на графике требуется представить несколько рядов данных, то в одной системе координат строится несколько линий, но для этого необходимо, чтобы все отображаемые на графике ряды имели одинаковые единицы измерения и могли быть представлены в одном и том же масштабе.

Например, если нужно сравнить ежемесячные продажи за три года, то можно воспользоваться графиком на рисунке.



Если значения в рядах данных лежат в различных диапазонах и отличаются на несколько порядков, то отображение этих рядов на одном графике может вызвать определенные затруднения.

Допустим, нужно сравнить динамику продаж нескольких товаров, но цены на различные товары даже в одной группе могут различаться в десятки и сотни раз.

Так, дорогие агрегаты для автомобиля (коробка передач, двигатель и др.) продаются единично, а мелкие детали ценой в несколько рублей (гайки, болты и др.) — сотнями.

Диаграммы

С помощью графика удобнее всего отображать непрерывные (числовые) величины, поскольку можно получить достаточное число точек, чтобы его построить.

Если же речь идет о категориальных (дискретных) значениях, то более подходящим средством визуализации является **диаграмма**. Принципиального различия между понятиями график и диаграмма нет.

Просто под графиком традиционно понимают представление зависимостей в виде линий, тогда как в диаграмме значения отображаются с помощью самых разнообразных объектов и фигур.

Как правило, в диаграммах, по горизонтальной оси X откладываются категории, а по вертикальной оси Y — значения.

Самые простые и часто используемые диаграммы — столбчатые. В них значение каждой категории представляется в виде столбика, высота которого пропорциональна соответствующему значению (1).

Разновидностью столбчатой диаграммы является линейчатая диаграмма, которая отличается от столбчатой тем, что ось категорий откладывается вертикально, а ось значений — горизонтально (2).



Круговые диаграммы

Еще одним распространенным видом диаграмм является круговая диаграмма. Ее очень удобно использовать, если нужно показать долю, которую вносит то или иное значение в общий результат. Эта доля может быть выражена как в абсолютных единицах, так и в процентах (например, процент от выручки, который обеспечил товар А). Такие примеры круговых диаграмм, показывающие распределение структуры месячных затрат на виды мобильной связи, представлены на рисунках.

Диаграмма рассеяния

Диаграмма рассеяния (англ.: Scatterplot) показывает распределение ограниченного набора точек, которые соответствуют значениям по осям.

Например, связь между двумя показателями по всем регионам РФ: Число рожденных на 1000 человек, % и Численность населения (тыс. чел.). Видно, что высокая рождаемость более характерна для субъектов с небольшой численностью населения.

Гистограммы

Гистограмма показывает распределение набора данных внутри выборки (например, количество заемщиков банка в нескольких возрастных группах) в виде столбиков.

Гистограммы широко используются в статистике для определения наиболее вероятных значений, которые может приобретать некоторая величина, а также для выявления законов распределения, которым подчиняется случайная величина.

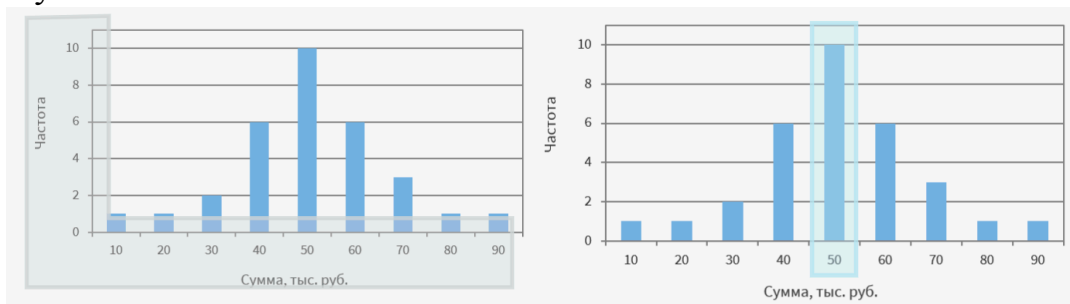
Гистограмма строится следующим образом. Пусть исследуемой величиной являются ежедневные продажи торговой точки в течение месяца (см. фрагмент данных в таблице). При этом минимальное наблюдаемое значение составило 10 тыс., а максимальное — 100 тыс.

Дата	Сумма
01.05.2020	10 000
02.05.2020	33 000
03.05.2020	51 000
04.05.2020	25 000
05.05.2020	48 000
06.05.2020	62 000
07.05.2020	73 000
08.05.2020	100 000
09.05.2020	77 000
...	...

Разобьем диапазон изменения величины на 9 поддиапазонов по 10 тыс. и подсчитаем, сколько раз значение продаж попадает в тот или иной поддиапазон. Результаты сведем в таблицу.

Диапазон в тыс.	10 — 20	20 — 30	30 — 40	40 — 50	50 — 60	60 — 70	70 — 80	80 — 90	90 — 100
Частота	1	1	2	6	10	6	3	1	1

Как видим, продажи на сумму от 10 до 20 тыс. наблюдались только один раз, от 30 до 40 тыс. — два раза и так далее. На основе полученной таблицы строим гистограмму.

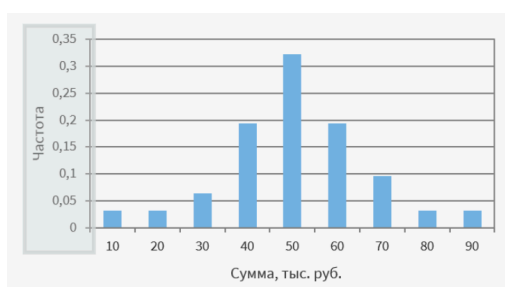


По горизонтальной оси гистограммы откладываются значения продаж, а по вертикальной — количество или частота наблюдений, значения которых попали в заданный диапазон (поэтому иногда гистограмму называют частотным полигоном).

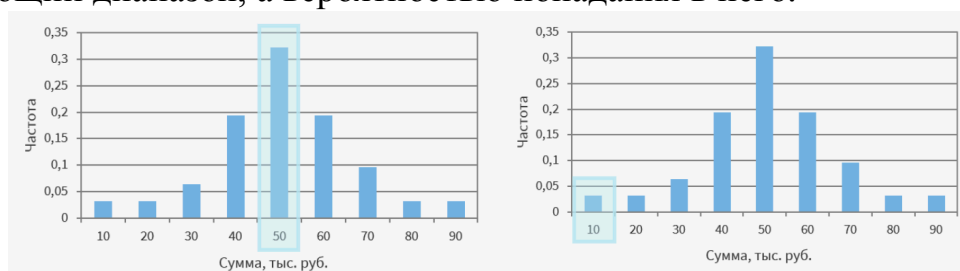
Гистограмма на рисунке показывает, что наибольшее число наблюдений попало в диапазон 50-60 тыс. Таким образом, значения из данного диапазона можно рассматривать как наиболее вероятные. Эту информацию используют для восстановления пропущенных значений при очистке данных, для планирования денежных поступлений, закупок и так далее.

Что касается крайних элементов гистограммы, то они представляют редкие события — экстремально высокие или экстремально низкие значения продаж. Видно, что в диапазон 10-20 тыс. попало всего одно значение, следовательно, вероятность такого события мала, и его не стоит включать в рассмотрение. Экстремально низкие продажи могут быть вызваны исключительной ситуацией, например, погодными условиями, отключением электроэнергии и т. п.

Иногда используют **нормированную гистограмму**, что позволяет оперировать не значениями наблюдений, а их вероятностями. Для этого каждый элемент гистограммы делится на количество наблюдений, то есть в нашем случае на 31 (число дней в месяце). Обратимся к рисунку.



Теперь высота столбца определяется не количеством наблюдений, попавших в соответствующий диапазон, а вероятностью попадания в него.



На рисунке видно, что вероятность попадания значения в диапазон 50-60 тыс. составляет примерно 0,32, или 32 %.

Соответственно, вероятность появления значений в диапазоне 10-20 тыс. не превышает 0,03, или 3 %.

В **нормированной гистограмме** сумма значений всех ее элементов должна быть равна 1, поскольку сумма вероятностей всех возможных событий (попадания значения в какой-либо диапазон) есть 1.

Обычно при построении гистограммы аналитик имеет возможность задать число поддиапазонов, на которое будет разбиваться исходный диапазон изменения величины (фактически это число столбцов гистограммы). Здесь существуют различные рекомендации, например: число поддиапазонов не должно быть меньше, чем $\log_2 N$, где

N — число наблюдений.

На практике можно руководствоваться следующим эмпирическим правилом. Количество столбцов в гистограмме должно быть таким, чтобы в ней не образовывались провалы, резкие выбросы или множественные пики. Она должна быть достаточно гладкой, чтобы по ней можно было определить характер распределения наблюдаемой величины. Часто хороших результатов удается добиться при использовании 10-15 столбцов.

Инфографика

Инфографика стала очень популярна в последние годы, хотя существуют уже давно. Инфографика относится к журналистике данных, где графики и схемы объясняют какие-либо факты по выбранной теме. Обычно инфографика статична и представляет собой длинную «простыню» с картинками и текстом. Отличительной особенностью инфографики является то, что в ней приводятся уже готовые выводы, то есть читателя проводят за руку по выбранной теме и при этом приправляют это все цифрами и картинками. Часто используется рисованный или анимационный стиль. Часто используется не к месту или «для красоты», хотя конечно же есть замечательные и интересные примеры.

Презентация и анализ данных

Один самых привычных способов использования визуализации данных - презентация информации в виде диаграмм или инфографики. И если с этим все понятно, то использование визуализации для анализа информации в основном используется только бизнес-аналитиками и учеными. В чем заключается отличие?

При анализе данных с помощью визуализации используют так называемое быстрое прототипирование - то есть создание большого количества различных визуальных представлений одних и тех же данных. Делается это для возможности нахождения скрытых, на первый взгляд, взаимосвязей и зависимостей, а также первичной оценки набора данных для возможности применения в дальнейшем более сложных инструментов анализа. Этот подход называется разведывательный анализ данных. Основное отличие от презентации данных - визуализация здесь может быть «черновой», но выполняется быстро и одним человеком или небольшой рабочей группой.

Интерактивный сторителлинг

Сторителлинг - это преподнесение какой-либо полезной информации в форме интересного рассказа. Интерактивный сторителлинг - рассказ, с которым слушатель может взаимодействовать. Пользователь может управлять отображением информации и находить те зависимости, которые не нашёл автор. В этом смысле он близок к разведывательному анализу данных, но отличается тем, что данные заранее обработаны и представлены в удобном для анализа виде, а также имеются подсказки или заранее прописанные сценарии использования.

Поэтому, чаще всего интерактивный сторителлинг называют интерактивной инфографикой, но для того чтобы ей стать недостаточно просто к статичной инфографике добавить всплывающие окошки.

Дашборды и бизнес аналитика

Визуализация активно используется в бизнесе. Принцип «говорите с данными» помогает компаниям зарабатывать больше, а клиентам получать лучший сервис. Для разового анализа обычно используется Excel или R. Однако это неудобно если необходимо следить за какими-то показателями на постоянной основе. Для

отслеживания используют дашборды - дисплеи, на которых выведены все необходимые показатели в одном месте в виде графиков, диаграмм и таблиц. Проектирование эффективных дашбордов - сложная и неординарная задача. Зачастую их перегружают ненужной информацией или стараются использовать все возможные типы шаблонных графиков. Часто для того, чтобы спроектировать хороший дашборд, необходимо создание новых типов визуализации информации. Тематика активно развивается за счет все большего применения аналитики в бизнесе. Также дашборды применяются и для личного использования (фитнес трекеры, анализ личных расходов и т. п.)

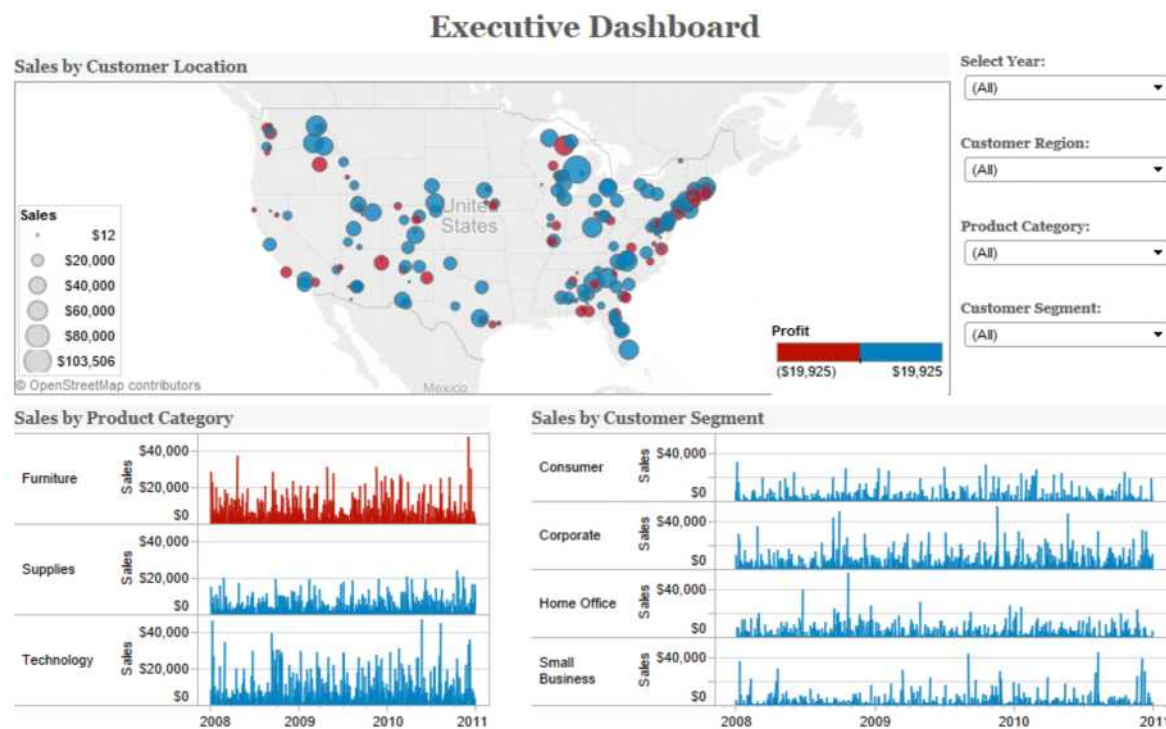


Рисунок 2 - Пример дашборда

Карты и картограммы

Карты - одни из древнейших способов визуализации, отображающих окружающую реальность. Картограмма - карта с нанесенной на неё информацией в виде цвета или других способов. Картограммы могут быть использованы для отображения любой информации - от плотности населения, до частоты использования мобильных телефонов в каждом районе страны.

Географические карты

Карты (англ.: maps) — позволяют наглядно представить данные, связанные с географическим расположением исследуемых объектов и процессов. В бизнес-аналитике это информация, связанная с уровнем потребления в регионах, характером спроса и предложения по различным видам товаров, сведения о продажах, осуществляемых региональными дилерами, о логистических и транспортных потоках и так далее.



Рассмотрим пример. В таблице содержатся данные об отгрузках продукции, произведенных компанией мелким оптовикам из различных субъектов Центрального федерального округа (ЦФО).

Имея карту ЦФО, можно в соответствии с палитрой раскрасить области и получить сравнительный отчет о продажах по регионам.

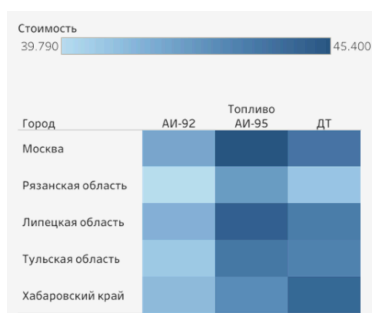
Видно, что максимальные значения продаж приходятся на регионы, обозначенные 10 и 11 (Москва и Московская область), а минимальные — на 15 (Тамбовская область). При этом может использоваться многомерное представление данных.

Например, если требуется показать на карте не только суммы продаж, но и прибыль, то последнюю можно отобразить рельефным выделением соответствующей области на карте. Также на карты можно наносить значки, форма, цвет, размер и взаимное положение которых соответствуют свойствам исследуемых объектов.

Тепловые карты

Карты не обязательно должны быть связаны с географией. В бизнес-аналитике почти всегда приходится иметь дело с объектами, которые описываются двумя признаками и более. То есть выборки являются многомерными, и представление подобных данных на плоских визуализаторах (графиках, диаграммах) не всегда удобно и корректно отображает результаты.

В данном случае имеет смысл использовать двумерные тепловые карты (англ.: heat maps), где каждому значению признака соответствует один из оттенков в заранее выбранной цветовой гамме.



На рисунке можно наблюдать пример тепловой карты, построенной по двум измерениям: Топливо и Город. Цвет каждой прямоугольной ячейки карты формируется на основе цены топлива: чем темнее оттенок, тем она выше.

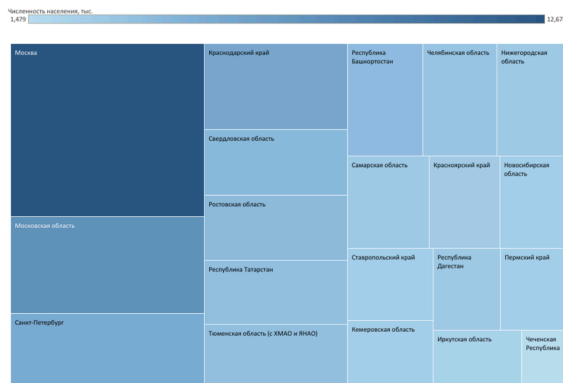
Плоское дерево

К семейству карт можно отнести и метод плоское дерево (англ.: treemap).

Он является очень эффективным при изображении численных атрибутов элементов (размер, стоимость, значение), организованных в большие иерархии.

На рисунке за размер прямоугольников отвечает Среднегодовая численность занятых в экономике, тыс. чел., а за цвет — численность населения, тыс. чел. по субъектам и федеральным округам РФ.

Таким образом, близкие по значению этого атрибута регионы располагаются на плоском дереве рядом.



Облако тегов

Каждому элементу в облаке тегов присваивается определенный весовой коэффициент, который коррелирует с размером шрифта. В случае анализа текста величина весового коэффициента напрямую зависит от частоты употребления (цитирования) определенного слова или словосочетания.

Позволяет читателю в сжатые сроки получить представление о ключевых моментах сколько угодно большого текста или набора текстов.

Облако данных

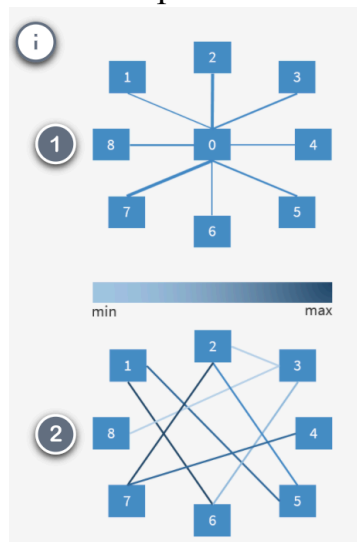
Облако данных (англ.: Data cloud) — это визуализатор, в котором используется другой цвет и/или размер шрифта для обозначения числовых данных.

Родителем этого визуализатора можно считать облако тегов (англ.: Tag cloud), в котором сравниваются ключевые слова или фразы (значения), которые содержатся внутри фрагмента текста (набора данных), при этом каждому из них задается свой размер шрифта.

На рисунке приведен пример облака данных регионов РФ, в котором за цвет (в заранее выбранной шкале) отвечает число рожденных на 1000 чел., %, а за размер шрифта — численность населения.

Сеть

Нередко требуется исследовать характер и степень взаимной зависимости между различными объектами. Для анализа можно использовать визуализацию связей, когда объекты представляются в виде значков, а связи между ними — в виде линий, соединяющих соответствующие значки. При этом сила связи, то есть степень взаимной зависимости объектов, может показываться различными способами.



Чаще всего для этого используют:

- толщину линии: чем сильнее связь между объектами, тем толще соединяющая их линия (1);
- цвет линии, при этом выбираются оттенки определенного спектра (2).

Допускается одновременно использовать и толщину, и цвет, но такие диаграммы более сложны для восприятия. Кроме того, можно по-разному располагать объекты на плоскости.

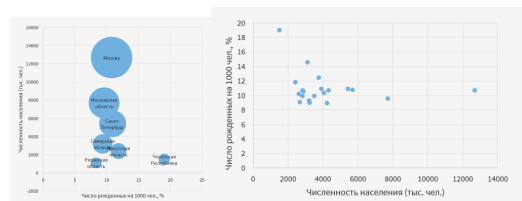
Во втором случае визуализируются все попарные связи между объектами.

Если число объектов велико, связей становится еще больше, поэтому анализируют только наиболее сильные или наиболее слабые, накладывая фильтр на значения силы связи.

Пузырьковая диаграмма

Пузырьковая диаграмма (англ.: Bubble Chart) — это смесь графика и диаграммы, когда по двум осям расставлен набор точек, соответствующий значениям. При этом сами точки не соединены и имеют различную величину и/или цвет, которые задаются дополнительными показателями.

На рисунке изображен пример такой диаграммы, на которой за размер точки отвечает показатель Численность населения занятых в экономике (тыс. чел.).



Кластерграмма

Метод визуализации, использующийся при кластерном анализе (рисунок). Показывает, как отдельные элементы множества данных соотносятся с кластерами по мере изменения их количества. Выбор оптимального количества кластеров - важная составляющая кластерного анализа.

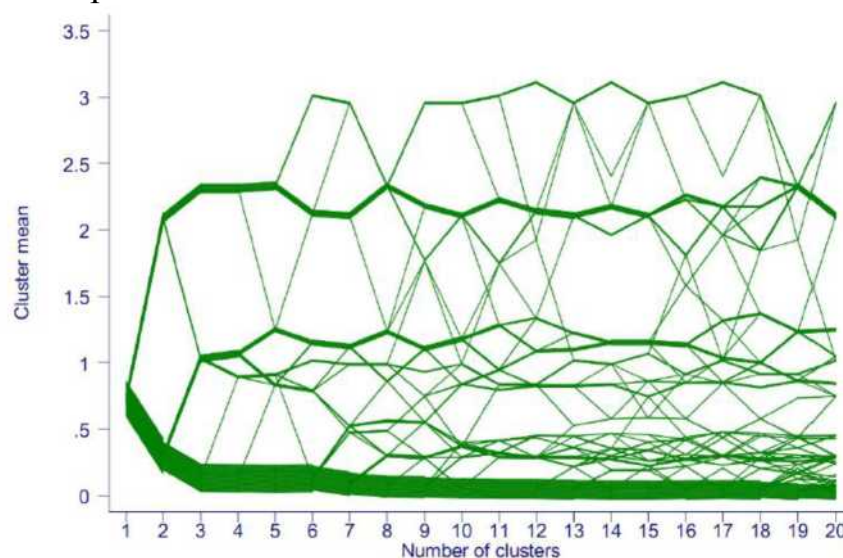


Рисунок - Пример кластерграммы

Исторический поток

Помогает следить за эволюцией документа, над созданием которого работает

одновременно большое количество авторов. По горизонтальной оси откладывается время. По вертикальной - вклад каждого из соавторов, т.е. объем введенного текста. Каждому уникальному автору присваивается определенный цвет на диаграмме.

Пространственный поток

Используется для отслеживания пространственного изменения информации.

Проблема визуализации сложных данных

Большинство реальных процессов являются сложными, поскольку в них участвует много объектов, которые находятся в самых разнообразных отношениях, и с каждым из которых может быть связано несколько числовых характеристик.

Процесс анализа данных требует сравнения и сопоставления этих характеристик (например, продаж по различным городам или группам товаров), вычисления дополнительных показателей и характеристик и так далее.

Поэтому при визуализации данных часто встает вопрос: как представить сложные данные в таком виде, чтобы человек мог их осмыслить и интерпретировать?

Предположим, требуется визуализировать с помощью таблицы продажи нескольких наименований товаров по нескольким датам и городам (см. таблицу).

Дата	Товар	Город	Цена, руб.	Количество	Сумма, руб.
01.02.15	Пылесос	Михайлов	6 500	10	65 000
01.02.15	Ст. машина	Михайлов	16 700	5	83 500
01.02.15	Утюг	Михайлов	4 200	3	12 600
01.02.15	Пылесос	Касимов	7 500	2	15 000
01.02.15	Ст. машина	Касимов	16 700	4	66 800
01.02.15	Пылесос	Касимов	6 500	5	32 500
02.02.15	Пылесос	Михайлов	7 500	4	30 000
02.02.15	Утюг	Михайлов	4 200	10	42 000
02.02.15	Телевизор	Михайлов	78 000	2	156 000

Даже такой простой фрагмент демонстрирует сложности при визуализации бизнес-процессов. Если продажи велись одновременно по множеству наименований товаров в нескольких городах, то таблица, содержащая соответствующие наблюдения, неизбежно окажется избыточной: в ней много раз будут повторяться одни и те же данные. А оперативно получить нужную информацию из такой таблицы будет весьма проблематично из-за огромного количества наблюдений.

Возможным выходом из ситуации является декомпозиция сложной таблицы на множество более простых, где содержится информация о продажах отдельных товаров и групп товаров по городам, интервалам дат, поставщикам и так далее, но такой подход позволяет лишь отчасти решить проблему.

OLAP-анализ

Все эти трудности, возникающие при обработке больших и сложных массивов данных, создали предпосылки для появления метода визуализации многомерных табличных данных — OLAP-анализа.

В основе OLAP лежит многомерное представление данных, которые могут быть разделены на количественные и качественные.

Аббревиатура OLAP также встречается в теории хранилищ данных, но если там она употребляется в контексте моделей хранения данных (многомерной, реляционной), то здесь под OLAP подразумевается метод визуализации, при этом модель хранения данных может быть любой.

Качественные данные (Измерения)

Качественные данные представляют собой значения, выраженные в категориальной форме. Обычно это наименования товаров, групп товаров, организаций, названия городов, ФИО сотрудников и так далее.

В рамках многомерной модели данные, качественно описывающие исследуемый бизнес-процесс, называются измерениями (или группами).

Количественные данные (Факты)

С каждым объектом связаны признаки, количественно описывающие его. Для товара это может быть цена, количество или сумма; для города, в котором расположено торговое представительство, — расстояние до него и количество жителей; для сотрудника — заработная плата и стаж работы.

Данные, количественно описывающие процесс или объект, называются фактами (или показателями).

С каждым качественным значением анализируемого бизнес-процесса связаны один или несколько количественных показателей. Другими словами, с каждым измерением связаны один или несколько фактов.

Измерения несут смысловую нагрузку, а факты — количественную. Чтобы достоверно отделить измерения от фактов, достаточно сопоставить значение с вопросом.

Измерения позволяют ответить на вопросы: Что? (товар), Кто? (клиент), Когда? (дата) и Где? (город).

Факты отвечают на единственный вопрос: Сколько?

Как построить OLAP-куб?

Чтобы построить OLAP-куб, пользователь должен указать системе следующие параметры:

- какие измерения и факты включать в куб;
- варианты агрегации значений фактов.

Визуализация OLAP-куба проводится с помощью специального вида таблиц, которые строятся на основе срезов OLAP-куба, содержащих необходимую пользователю информацию. Срезы, в свою очередь, являются результатом выполнения соответствующего запроса к источнику данных.

Как правило, в процессе построения срезов пользователь с помощью мыши и клавиатуры манипулирует заголовками измерений, добываясь наиболее информативного представления данных в кубе. В зависимости от положения заголовков измерений в таблице автоматически формируется запрос к базе или хранилищу данных. Запрос извлекает данные из базы или хранилища, после чего OLAP-ядро системы визуализирует их.

Отдел. Наименование		Σ Факты								
Дата (Год + Месяц)	Группа. Наименование	Аптека 1			Аптека 2			Итого:		
		Сумма	Количество	Цена	Сумма	Количество	Цена	Сумма	Количество	Цена
> 01.01.2017		33 284,05	573	58,09				33 284,05	573	58,09
> 01.02.2017		33 809,16	623	54,27				33 809,16	623	54,27
> 01.03.2017		32 241,17	534	60,38				32 241,17	534	60,38
> 01.04.2017		33 487,97	527	63,54	19 370,63	353	54,87	52 858,60	880	60,07
> 01.05.2017		22 377,64	449	49,84	10 759,29	224	48,03	33 136,93	673	49,24
> 01.06.2017		21 364,05	425	50,27	8 160,47	186	43,87	29 524,52	611	48,32
> 01.07.2017		13 536,43	373	36,29	8 158,10	164	49,74	21 694,53	537	40,40
> 01.08.2017		14 324,56	312	45,91	10 764,90	227	47,42	25 089,46	539	46,55
> 01.09.2017		23 436,44	453	51,74	15 008,15	278	53,99	38 444,59	731	52,59
> 01.10.2017		31 328,32	536	58,45	21 777,81	361	60,33	53 106,13	897	59,20
> 01.11.2017		33 413,86	588	56,83	16 416,76	281	58,42	49 830,62	869	57,34
> 01.12.2017		32 596,49	591	55,15	21 365,74	350	61,04	53 962,23	941	57,35
Итого:		325 200,14	5 984	54,34	131 781,85	2 424	54,37	456 981,99	8 408	54,35

На рисунке показан пример визуализации данных с помощью OLAP-куба в Loginom, где представлена информация о количестве и суммах продаж по аптекам в разрезе года и месяца. Эти сведения могут оказаться полезными при определении наиболее загруженных аптек. Так, в кубе для каждого месяца и года перечислены все продажи, которые были сделаны в пределах этого периода, с указанием количества проданных единиц товара и их общей суммы. Визуальный анализ позволяет увидеть, что больше всего продаж пришлось на Аптека 1.

Инструменты визуализации

1. Databox

Лучший инструмент для визуализации данных и бизнес-аналитики.

Databox – это инструмент визуализации данных, которым пользуются более 15 000 предприятий и маркетинговых агентств. Databox собирает ваши данные в одно место, чтобы отслеживать производительность бизнеса в реальном времени с помощью привлекательных визуальных эффектов.

Подходит для:

Databox лучше всего подходит для маркетинговых команд, которые хотят быстро настроить информационные панели. Благодаря более чем 70 интеграциям в один клик с другими программами и отсутствию необходимости программирования, это самый простой в использовании инструмент.

Плюсы:

- Более 200 готовых отчетов, охватывающих все области: от CRM (системе управления информацией о клиентах) и Google Analytics до Google Ads, эффективности электронной почты и автоматизации маркетинга.
- Более 70 интеграций в один клик, которые объединяются с данными из таких популярных инструментов, как HubSpot, Salesforce, Mailchimp, Facebook Ads, Shopify и т.д.
- Легкое создание расширенных вычислений данных из нескольких источников, включая ROAS, ARPU, LTV, LTV: CAC и т.д.

Цена:

У Databox есть бесплатный план, который позволяет участвовать 3 пользователям, создавать 3 источника данных, которые можно обновлять ежедневно. Их платные планы варьируются от 49 до 248 долларов в месяц.

Их базовый план за 49 долларов в месяц – лучший выбор, поскольку он предоставляет 10 учетных записей пользователей, 10 источников данных, 10 панелей мониторинга, данные, которые обновляются ежечасно, и доступ к их API (интерфейс прикладных программ).

2. Zoho Analytics

Zoho Analytics, вероятно, один из самых популярных инструментов бизнес-аналитики в этом списке. Единственное, в чем вы можете быть уверены, – это то, что с помощью Zoho Analytics вы можете безопасно загружать свои данные. Более того, вы можете использовать широкий набор диаграмм, таблиц и компонентов, чтобы преобразовать ваши данные в сжатый вид.

Подходит для:

Zoho Analytics обслуживает преимущественно крупные предприятия, включая Bajaj Allianz, Maruti Suzuki, HP, Hyundai Motors и HDFC. Однако он подходит и для

среднего бизнеса.

Плюсы:

- Помогает создавать зрительно привлекательные отчеты с визуализацией данных.
- Он очень прост в использовании, а панель управления информацией интуитивно понятна.
- Обладает надежной политикой безопасности данных.
- Предоставляет обширную поддержку клиентов, которая пригодится новичку.
- Позволяет легко встраивать отчеты и информационные панели на веб-сайты и в каналы социальных сетей.
- Извлекает данные из нескольких ресурсов и затем обрабатывает их для создания подробных отчетов.
- Поставляется с виртуальным помощником на основе ИИ, который поможет вам в процессе визуализации данных.

Цена:

Тарифные планы начинаются с 25 долларов в месяц и варьируются до 495 долларов в месяц. Вы также получаете 15-дневную бесплатную пробную версию для ознакомления с продуктом.

3. Tableau

Если вы хотите с легкостью просматривать и визуализировать данные, Tableau – это инструмент визуализации данных для вас. Он помогает вам создавать диаграммы, карты и любой другой тип профессиональной графики. Чтобы улучшить ваши визуальные представления, вы также можете найти настольное приложение этого инструмента.

Более того, если у вас возникнут проблемы с установкой какого-либо стороннего приложения, Tableau содержит опцию – решение «заблокировать сервер», которое поможет визуализировать отчеты в Интернете или с помощью мобильных приложений.

Подходит для:

Tableau больше всего подходит для корпораций, которым нужны решения для визуализации данных без необходимости настраивать их вручную.

Плюсы:

- Продукт оптимизирован для мобильных устройств и легко обновляется.
- Обеспечивает обширную поддержку клиентов.
- Позволяет легко выполнять смешивание данных
- Корпоративная команда может выполнять анализ своих данных в реальном времени.
- Предоставляет доступ ко всем типам аналитики данных, включая IT-аналитику, аналитику официальных данных, маркетинговую аналитику, страховую аналитику и многое другое.
- Предлагает отдельный бесплатный инструмент под названием Tableau Public, который предоставляет некоторые функции, но не позволяет сохранять отчеты локально.
- Предлагает расширения JavaScript и API для пользовательских отчетов через библиотеки диаграмм JavaScript.

Минусы:

- Нет возможности автоматического обновления отчетов.
- Это решение не полностью комплексное, и вам потребуются предварительные

знания SQL (реляционные базы данных).

Цена:

Тарифные планы начинаются с 12 долларов за пользователя в месяц и варьируются до 70 долларов за пользователя в месяц.

4. Infogram

Infogram – это веб-платформа для визуализации данных и инфографики. Она позволяет создавать и публиковать инфографику, карты и диаграммы, используя все пользовательские данные и преобразовывая их в привлекательную графику.

После этого вы можете далее делиться, вставлять или публиковать эту графику, где захотите. Более того, этот инструмент прост в использовании и подходит для государственных органов, маркетинговых команд и даже для студентов.

Хорошая новость заключается в том, что вам не требуются какие-либо технические знания или навыки программирования, чтобы начать работу с Infogram. Тарифы также различаются в зависимости от типа использования, поэтому вам не нужно беспокоиться о доплате. А если вы хотите применять ее базовые функции и для дизайна, то это совершенно бесплатно.

Подходит для:

Infogram больше всего подходит для организаций, которые занимаются продажами и маркетингом, поскольку позволяет отображать ваши цели с помощью диаграмм и шаблонов. Это также полезно для аудиторских компаний.

Плюсы:

- Помимо того, что она полезна для корпораций и государственных органов, платформа также весьма удобна для студентов.
- Предоставляет настраиваемые шаблоны, которые помогут вам добавить цвета, шрифты, логотипы и многое другое.
- Она предлагает широкий выбор изображений, значков и флагов. Присутствуют соответствующие наборы карт и схем.
- Шаблоны помогают создавать изображения в социальных сетях, информационные панели и отчеты.
- В ней есть несколько красивых шаблонов проектов, которые могут превращать данные в интересный контент.

Минусы:

- Навигация по вашим данным может быть сложной задачей.
- Поддержка клиентов нуждается в улучшении.

Цена:

Платные планы начинаются с 19 долларов в месяц и составляют до 149 долларов в месяц. Кроме того, есть один тарифный план, который поможет вам выбрать все функции, которые вам нужны, в зависимости от требований вашего бизнеса. Infogram также предлагает тарифный план Freemium для основных функций.

5. ChartBlocks

ChartBlocks – это инновационное программное обеспечение для визуализации данных, которое позволяет импортировать данные за доли секунды. Более того, вы можете сразу же обновить свои данные в приложении ChartBlocks и выполнить необходимый импорт.

Оно также поставляется с «конструктором диаграмм», который поможет вам создать

нужный тип диаграммы. Имеет определенный высокий уровень сложности инструментов. После этого вы можете легко встроить диаграммы на свой сайт или в любую другую социальную сеть.

Подходит для:

Инструмент лучше всего подходит для индустрии туризма, здравоохранения, страхования и розничной торговли.

Плюсы:

- Обеспечивает отличную поддержку клиентов.
- Упрощает загрузку файлов данных.
- Позволяет создавать и публиковать диаграммы за считанные минуты.
- Вы можете использовать одни и те же данные для создания различных диаграмм.
- Поставляется со встроенными инструментами для социальных сетей, которые помогут вам напрямую поделиться и встраивать свои диаграммы.
- Оно совместимо со всеми типами мобильных устройств и подходит для экрана любого размера.

Минусы:

- Предоставляет ограниченный размер памяти для загрузки ваших данных.
- Качество экспорта результатов не самое удачное в бесплатной версии.

Цена:

Первый план совершенно бесплатный, и вы получаете до 50 активных графиков. После этого платные планы начинаются от 20 долларов в месяц.

6. Datawrapper

Datawrapper – еще одна отличная программа для визуализации данных в этом списке. Ваша целевая аудитория должна понимать, какой тип диаграммы вы создаете. Кроме того, учитывайте, что скриншот Excel неплохо выглядит на экране вашего компьютера, но он плохо читается на экране мобильного устройства.

И Datawrapper решает эту проблему автоматически. Все диаграммы, таблицы и карты, которые вы создаете с помощью Datawrapper, легко читаются на всех устройствах. Хорошая новость заключается в том, что для создания диаграмм вам не нужны навыки программирования.

Подходит для:

Поскольку Datawrapper является некоммерческой платформой, она больше всего подходит для студентов и небольших организаций, которым требуются базовые функции визуализации данных.

Плюсы:

- Предоставляет широкий набор встроенных диаграмм и графиков
- Не требует каких-либо технических знаний.
- Это инструмент с открытым исходным кодом и совершенно бесплатный.
- Графики, карты и другая графика автоматически сохраняются после их создания.
- Инструмент Datawrapper совместим со всеми операционными системами.
- Позволяет легко редактировать и комментировать ваши диаграммы

Минусы:

- Одним из самых больших недостатков является то, что с бесплатным планом ваши данные будут храниться на сервере Datawrapper.
- Настройка шрифтов и цветов может быть сложной задачей.

Цена:

Он предоставляется совершенно бесплатно, если вы хотите иметь базовый набор дизайна и диаграмм. Бесплатный план также позволяет вам экспортировать диаграммы, таблицы и карты в формат png. Платные планы начинаются с 499 € в месяц.

7. Plotly

Plotly не только помогает вам создавать графики, но также предоставляет подробные аналитические отчеты, которые помогут вам отслеживать данные. Он также имеет легко настраиваемый пользовательский интерфейс.

Кроме того, он помогает легко экспортировать отчеты и отлично работает в области бизнес-аналитики.

Подходит для:

У Plotly есть такие коммерческие предложения, как Dash Enterprise и Chart Studio Enterprise. Следовательно, он больше подходит для профессиональных групп экспертов, которым требуется крупномасштабная наука о данных и применение ИИ. Помимо этого, он идеально подходит, как для студентов, так и для малого бизнеса.

Плюсы:

- Одной из важных особенностей Plotly является то, что он предоставляет несколько вариантов своего развертывания, что позволяет вам приобретать решения разными способами.
- Вы можете добавить несколько членов команды для одновременной работы с программой, что помогает в совместной работе команды и упрощает выполнение ваших задач.
- Поставляется с мощными инструментами визуализации и высококачественными инструментами аналитики.
- В этой программе и при составлении интерактивных отчетов можно заниматься разработкой приложений.
- При визуализации данных он обеспечивает имитационное моделирование, настраиваемые информационные панели и соответствующие отчеты.
- Поддерживает такие популярные языки, как R, Python, Node и другие.

Минусы:

- Не оказывает поддержку пользователям по телефону.
- Тарифные планы недоступны на официальном сайте.

Цена:

Цена зависит от количества пользователей. Кроме того, инструмент открыт для использования, но в течение ограниченного времени. После этого вам нужно будет его приобрести.

8. Visually

Visually – это платформа сообщества для визуализации данных и инфографики. Это позволяет пользователям искать различные изображения с помощью тегов и описаний. Вы также можете публиковать и встраивать графику прямо в свои профили в социальных сетях. Она работает в три простых шага.

Подходит для:

Как и Zoho Analytics, Visually обслуживает такие крупные ИТ-компании, как KPMG,

Tiger Analytics и Xoriant. Однако она также подходит для среднего бизнеса.

Плюсы:

- Помогает создавать привлекательную инфографику, видео, электронные книги, отчеты и многое другое для вашей компании.
- Инструмент визуализации Twitter также является частью Visually
- Пользователи могут собирать информацию из нескольких баз данных
- Поддержка клиентов широкая, и персонал компании готов помочь со всеми проблемами, с которыми вы сталкиваетесь.
- Вам не нужны навыки программирования; вам просто нужно указать свои требования, и вы получите нужный тип графики.

Минусы:

- Доступна только на английском языке.
- Не предоставляет фиксированный тарифный план.

Цена:

Цена зависит от количества пользователей. Кроме того, инструмент открыт для использования, но в течение ограниченного времени. После этого вам нужно будет его приобрести.

9 D3.js

D3.js – невероятно гибкий инструмент визуализации данных на основе Javascript. Он поддерживает декларативное программирование и помогает управлять документами на основе данных. Он также поддерживает большие наборы данных, и вы найдете его очень простым в использовании.

Помимо этого, он предоставляет такие уникальные функции, как возможность повторного использования кода, широкий спектр функций создания кривых, помощь в связывании данных с элементом на страницах HTML и многое другое.

Подходит для:

Он лучше всего подходит для тех фирм, которым необходимо создавать веб-страницы и веб-сайты. Вы также можете создавать отличные приложения с помощью этого инструмента визуализации данных.

Плюсы:

- Это проект с открытым исходным кодом и работает без каких-либо плагинов.
- Он требует минимального программирования и помогает визуализировать данные с любыми предварительными техническими знаниями.
- Позволяет визуализировать данные формата HTML, SVG и CSS.
- С помощью этого инструмента вы можете легко создать компонент диаграммы.
- Вам не нужно загружать всю библиотеку каждый раз
- Помогает вам в манипуляциях с объектной моделью документа (DOM)
- Вы также можете создать HTML-таблицу из широкого набора чисел.

Минусы:

- Он немного сложно устроен и требует много времени на освоение.
- Ему следует быть более гибким в настройках и предлагать их лучшие варианты

Цена:

Он предлагает тарифный план, по которому вам нужно связаться с продавцом.

10. Ember Charts

Ember Charts больше похожа на библиотеку диаграмм, созданную с помощью фреймворков Ember.js и d3.js. С помощью Ember Charts вы можете создавать графики для временных рядов, точечные и круговые диаграммы и гистограммы. Более того, он помогает вам легко расширять и изменять диаграммы, поскольку предлагает отличные возможности настройки.

Подходит для:

Ember Charts подходит для вашей организации, если вам часто требуются статистические графики. Более того, если вам нужна такая графика, как столбчатая или горизонтальная диаграмма, круговая диаграмма, линейная диаграмма с накоплением данных и разбросом точек, тогда это также отличный инструмент.

Плюсы:

- Ember Charts поставляется с простой панелью управления. Таким образом, его легко использовать
- Обеспечивает высокий уровень настраиваемости
- Он предоставляет уникальные и богатые функции и позволяет добавлять в проект легенду, всплывающие подсказки, метки и эффекты при наведении курсора мыши.
- Он предлагает мощные и совершенные диаграммы.
- Если вы добавите данные, которые не являются нормальными, ваша диаграмма не сломается. Это означает, что, если вы добавите неверные данные, Ember Charts не перестанет работать с ней.
- Он совершенно бесплатен
- Вы можете расширить класс раздела Компоненты диаграммы для создания собственных диаграмм.

Минусы:

- Продукт не так популярен, как другие инструменты в этом списке, поэтому вы не сможете прочитать достаточное количество отзывов, прежде чем начнете использовать этот продукт.

Цена:

Это бесплатный инструмент, не требующий дополнительных затрат.

11. NVD3

NVD3 позволяет создавать перенастраиваемые диаграммы вместе с компонентами диаграмм для d3.js. Хорошая новость заключается в том, что он сохраняет все настраиваемые компоненты и не лишает вас контроля, предоставляемого d3.js.

Подходит для:

Если вы предпочитаете аккуратные и настраиваемые диаграммы данных, вы можете попробовать NVD3. А поскольку он бесплатный, он идеально подходит и для небольших организаций.

Плюсы:

- NVD3 дает вам возможность создавать перенастраиваемые диаграммы
- Поставляется с многочисленными типами диаграмм, включая гистограммы временных рядов, точечные и линейные диаграммы и многое другое.
- Обеспечивает высокий уровень настраиваемости

Минусы:

- Документация чересчур краткая, что отнимет у вас много времени на освоение, чем

следовало бы.

- Вам нужно посмотреть пример кода, чтобы понять, как он работает
- Работа над NVD3 продолжается. Инструмент визуализации все еще требует некоторой доработки из-за проблемы этикеток (отметок)

Цена:

Вы можете бесплатно использовать инструмент визуализации данных NVD3.

12. Google Charts

Google Charts – еще один отличный инструмент для визуализации данных в этом списке. Он работает на кодах SVG и HTML5 и позволяет создавать довольно интерактивные диаграммы. Это также дает вам больше контроля над создаваемыми вами диаграммами и позволяет их масштабировать. Кроме того, он удобен в использовании, поэтому вы можете быстро приступить к работе.

Подходит для:

Google Charts подходит для всех типов организаций, будь то компании среднего размера, стартапы или крупные предприятия. Более того, он лучше всего подходит для организаций, которым необходимо создавать, управлять и представлять данные с помощью таблиц данных, интеграции с таблицами Google и анимации.

Плюсы:

- Google Charts прост в использовании. Вам не нужно заранее иметь какие-либо технические знания, чтобы использовать его
- Вы можете легко интегрировать Google Charts с другими продуктами Google.
- Он позволяет легко встраивать диаграммы на свой сайт и в социальные сети.
- Поставляется с обширной галереей картинок и множеством шаблонов.
- Он отображает данные на вашем сайте быстро и в режиме реального времени.
- Предоставляет интерактивную панель управления информацией
- Поставляется совершенно бесплатно

Минусы:

- Не очень подходит для углубленной персонализации программы под нужды пользователя.
- Вы можете столкнуться с небольшими трудностями при творческом подходе к графикам.

Цена:

Google Charts – это бесплатный продукт.

13. FusionCharts

FusionCharts – это программное обеспечение для визуализации данных на основе JavaScript, которое поставляется с обширной библиотекой диаграмм. Оно помогает извлекать необработанные данные из многочисленных баз данных и превращать их в содержательные отчеты.

Более того, оно предоставляет более 90 встроенных диаграмм и более 1000 карт, которые помогут вам преобразовать данные в интерактивные информационные панели.

Подходит для:

FusionCharts лучше всего подходит для организаций, которым необходимо создавать

информационные панели в рамках своего проекта или продукта. Однако, поскольку оно стоит довольно дорого, то больше подходит для крупного бизнеса.

Плюсы:

- FusionCharts предлагает гибкость в дизайне результатов.
- Обладает множеством таких преимуществ, как интерактивный характер, интерактивные данные и многочисленные цветовые схемы.
- Вы можете получить различные шрифты и макеты информационных панелей.
- Вы можете создать настраиваемую панель мониторинга с помощью FusionCharts.
- Предоставляет уникальные функции, включая реляционное отображение, визуальное обнаружение, согласованный API, кроссбраузерную поддержку и многое другое.
- Предлагает обширную документацию, чтобы можно было все детально разобрать

Минусы:

- Требуется улучшения, когда речь идет о параметрах настройки.
- Оно относительно дорого, поэтому больше подходит для крупных предприятий.

Цена:

Стоимость FusionCharts для одного пользователя начинается с 497 долларов. Тарифные планы варьируются до 9 947 долларов. Хорошая новость заключается в том, что вы можете получить бесплатную пробную версию, которую можете использовать перед окончательной покупкой.

14. Highcharts

Highcharts помогает вам создавать диаграммы на основе JavaScript на ваших веб-страницах. Он значительно упрощает визуализацию данных, обеспечивая широкие возможности настройки. Более того, он совершенно бесплатен для некоммерческого использования, в том числе для студентов, некоммерческих организаций и школьных веб-сайтов.

Подходит для:

Highcharts лучше всего подходит для студентов и некоммерческих организаций, поскольку он бесплатен для некоммерческого использования.

Плюсы:

- Highcharts предлагает множество различных типов диаграмм и поддерживает графики, сплайны, плоскости, сплайны с областями, столбцы и многое другое.
- Вы можете использовать Highcharts, чтобы легко отображать различные типы аналитики.
- Поставляется с широким набором сторонних плагинов.
- Бесплатен для некоммерческих проектов.
- Он динамичен по своей природе и позволяет добавлять, удалять или изменять серии данных и точки.
- Совместимость со всеми типами устройств, включая смартфоны, планшеты, настольные компьютеры и др.

Минусы:

- Highcharts может потребовать много вашего времени на выяснение и расчет значений оси координат

Цена:

Официальный сайт не предоставляет подробную информацию о ценах на Highcharts. Вам нужно будет связаться с продавцом, чтобы узнать цену. Но он бесплатен для

коммерческого использования.

15. Chart.js

Chart.js – это простой и гибкий инструмент визуализации для дизайнеров и разработчиков. Его поддерживает открытое сообщество, так что он тоже совершенно бесплатен. Другие важные функции включают возможность визуализировать данные восемью различными способами, отличную масштабируемость и многое другое.

Подходит для:

Chart.js подходит для небольших проектов, которым требуются основные функции визуализации данных.

Плюсы:

- Он прост в использовании и оснащен простой панелью управления.
- Позволяет визуализировать данные с помощью 8 различных типов диаграмм, включая полярные, столбчатые, кольцевые, радарные, круговые и другие.
- Графики данных могут быть анимированы и легко настраиваются.
- Вы можете добавлять и удалять любые типы диаграмм в Chart.js, чтобы уменьшить занимаемую площадь.
- Обеспечивает идеальную масштабируемость

Минусы:

- Самый большой недостаток Chart.js – медленный процесс отрисовки.

Цена:

Chart.js – бесплатная платформа для визуализации данных.

16. Leaflet

Leaflet – это библиотека JavaScript с открытым исходным кодом, которая предоставляет интерактивные карты, удобные для использования на мобильных устройствах. Она совместима со всеми типами устройств, браузеров и сторонних плагинов. В целом, этот инструмент разработан с учетом простоты, производительности и удобства использования. Однако в нем не представлены гистограммы и круговые диаграммы.

Подходит для:

Leaflet лучше всего подходит для организаций, которым нужны специализированные решения для отображения больших данных. Это также помогает организациям сэкономить много времени и ресурсов и добиться большего.

Плюсы:

- Leaflet предлагает обширный архив плагинов
- Позволяет добавлять маски, тепловые карты активности пользователей и анимированные маркеры.
- Продукт использует данные OpenStreetMap, что позволяет добавлять интерактивность и визуализацию HTML5 / CSS3.
- Он очень отзывчив на изменения данных и удобен для мобильных устройств.
- Это платформа визуализации данных с открытым исходным кодом, что делает ее бесплатной.
- Сообщество, которое запускает этот инструмент, всегда готово помочь вам начать работу.

Минусы:

- Он мог бы быть более отзывчивым на запросы пользователей
- При кодировании данных может не отображаться ошибка.

Цена:

Это платформа для визуализации данных с открытым исходным кодом, поэтому она абсолютно бесплатна.

Что следует искать в инструментах визуализации данных?

По словам Бена Шнейдермана, американского ученого-информатика: «Визуализация дает вам ответы на вопросы, о которых вы даже не подозревали».

Возможности визуализации данных не знают границ, поэтому вы должны точно знать, на что обращать внимание при выборе инструмента. Каждый бизнес индивидуален, и потребности будут разными, но некоторые из необходимых функций одинаковы в разных отраслях.

Вот некоторые из них:

1. Легкое встраивание

Вам нужны большие данные, чтобы получать реальные бизнес-идеи. После этого вам необходимо встроить полученные данные на свои веб-страницы или в любой канал социальных сетей, чтобы адекватно продемонстрировать все выводы вашей целевой аудитории.

Будучи встроенными в программное обеспечение для бизнеса, эти инструменты могут упростить принятие действенных решений в реальном времени, которые в конечном итоге принесут пользу бизнесу. Правильные инструменты могут сделать процесс обмена и встраивания отчетов простым и понятным.

2. Удобство для пользователя

Вы должны убедиться, что можете в полной мере использовать приобретаемый вами инструмент визуализации данных. Этого не произойдет, если инструмент не будет удобным и простым в использовании. Вот почему важно, чтобы выбранный вами инструмент был удобным для пользователя и имел современную и интуитивно понятную панель управления.

3. Сотрудничество в реальном времени

Принятие решений в режиме реального времени требует межгруппового и внутригруппового сотрудничества в реальном времени в любом бизнесе. Это помогает делать выводы, к которым каждый член команды может получить мгновенный доступ, обеспечивает их лучшую наглядность и удерживает всех членов команды на одной странице.

4. Масштабируемость

Мотив любого бизнеса – это устойчивость и прибыльность. Вы сможете добиться этого, только если продолжите расти и расширять свой бизнес. Вот почему вам нужен инструмент с высокой степенью масштабируемости, ориентированный на будущее. Ваш бизнес сегодня может быть небольшим, но инструмент, который вы используете, должен иметь возможность масштабирования и соответствовать растущим требованиям вашего бизнеса.

5. ИИ-интеграция

ИИ (AI – Artificial Intelligence) – это следующий большой шаг в развитии технологий. Это может упростить анализ данных и помочь прогнозировать тенденции и предписывать необходимые изменения. Следовательно, интеграция с ИИ в вашем инструменте визуализации может подготовить ваш бизнес к будущему.

Помимо этого, некоторые из других функций, связанных с ИИ, на которые вам следует обратить внимание, включают обслуживание клиентов, поддержку потоковых данных, интерактивное исследование, динамическую инфраструктуру и многое другое.

Заключительное резюме

Были вкратце описаны некоторые из лучших инструментов визуализации данных. Эти инструменты помогут вам сделать вашу презентацию данных более умной и интерактивной, что позволит вам привлечь внимание целевой аудитории.

Если вам нужны только базовые диаграммы и вы ограничены в средствах, то у вас есть множество вариантов, включая [Leaflet](#), [Chart.js](#), [Highcharts](#) и другие подобные инструменты.

Однако, если вам требуются расширенные функции и у вас приличный бюджет, вы можете попробовать более надежные и популярные инструменты визуализации данных и бизнес-аналитики, такие как [Databox](#), [Tableau](#), [Datawrapper](#), [Zoho Analytics](#) и другие. Эти инструменты подходят, как для средних, так и для крупных предприятий.

Независимо от того, используете ли вы устройства на операционных системах Apple, Microsoft Windows или даже Android, эти инструменты помогут вам извлечь необработанные данные из нескольких баз данных и представить их в простой форме.