

# ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ

- ВИЗУАЛИЗАЦИЯ И БИЗНЕС-АНАЛИТИКА
- МЕТОДЫ ВИЗУАЛИЗАЦИИ:
- ПРОСТЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ:
- СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ:

# ВИЗУАЛИЗАЦИЯ И БИЗНЕС-АНАЛИТИКА

Одной из важнейших составляющих аналитики данных является **визуализация** – представление данных в виде, который обеспечивает наиболее эффективную работу пользователя.

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**.



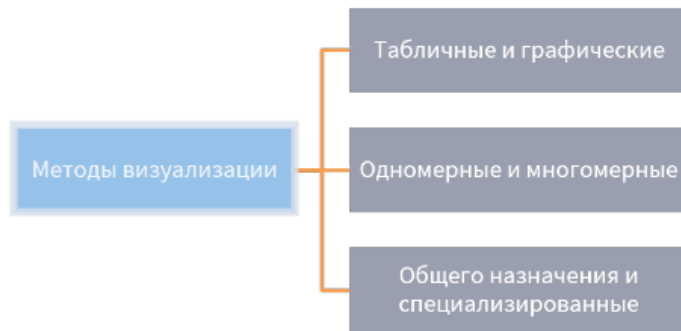
# ВИЗУАЛИЗАЦИЯ И БИЗНЕС-АНАЛИТИКА

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 данных 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.

# МЕТОДЫ ВИЗУАЛИЗАЦИИ

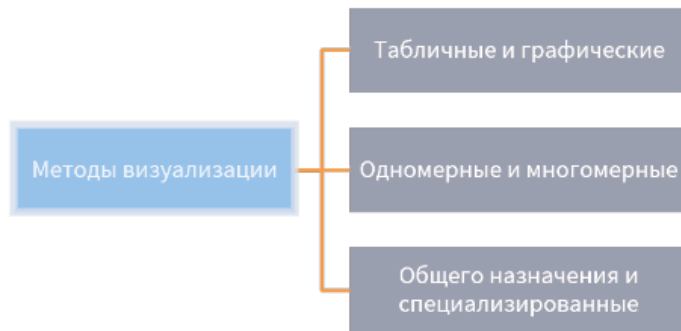
В настоящее время в аналитике данных используются различные методы визуализации. Рассмотрим данные методы визуализации, приняв за основу классификацию представленную на рисунке.



# МЕТОДЫ ВИЗУАЛИЗАЦИИ: 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.

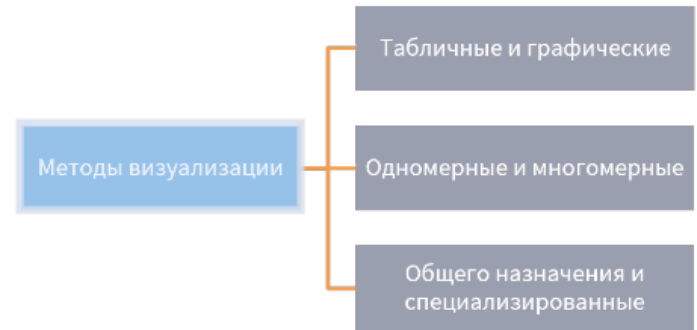
**Графические методы** позволяют лучше увидеть общий характер данных – закономерности, тенденции, периодические изменения. Кроме того, графические методы более эффективно сопоставляют данные: достаточно построить графики двух исследуемых процессов на одной системе координат, чтобы оценить степень их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



# МЕТОДЫ ВИЗУАЛИЗАЦИИ: 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.

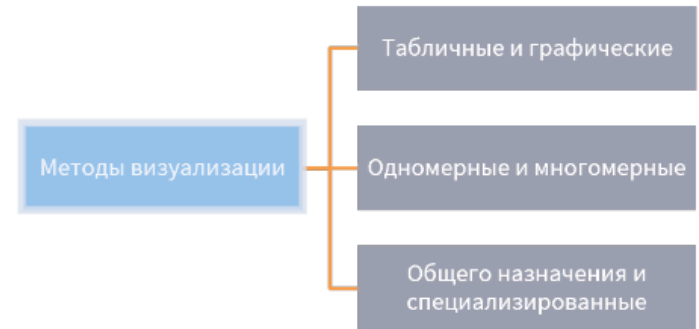
Популярные многомерные визуализаторы: OLAP-куб, тепловая карта и др.



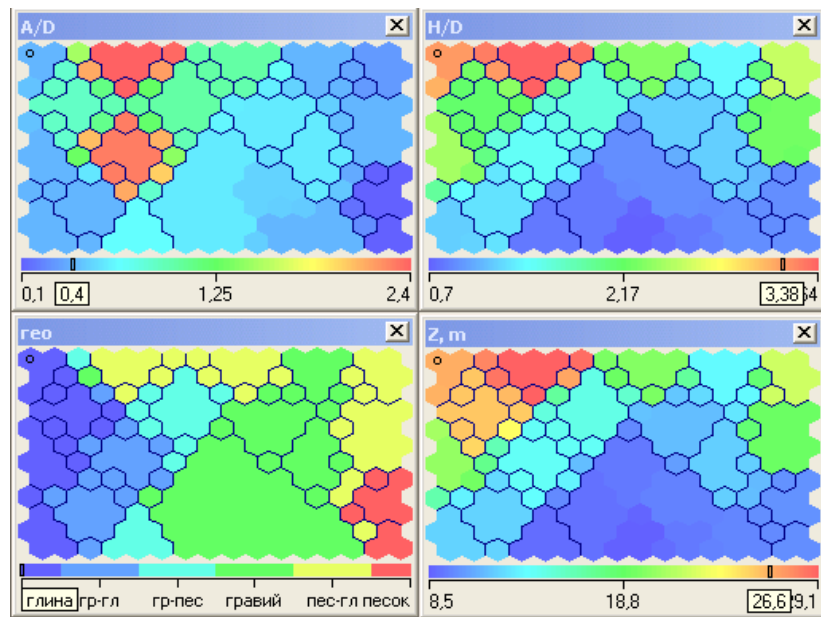
# МЕТОДЫ ВИЗУАЛИЗАЦИИ: ОБЩЕГО НАЗНАЧЕНИЯ И СПЕЦИАЛИЗИРОВАННЫЕ

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.



# МЕТОДЫ ВИЗУАЛИЗАЦИИ: ОБЩЕГО НАЗНАЧЕНИЯ И СПЕЦИАЛИЗИРОВАННЫЕ



Пример карт Кохонена



# МЕТОДЫ ВИЗУАЛИЗАЦИИ

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:

- **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.
- **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.

# ПРОСТЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ

К простым визуализаторам общего назначения относятся:

- таблицы;
- графики;
- диаграммы;
- гистограммы.

Кратко рассмотрим каждый из перечисленных визуализаторов.

# ПРОСТЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: 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.

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

# ПРОСТЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: ТАБЛИЦЫ

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

У таблицы как визуализатора существует множество вариантов представления. На рисунке 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 (продаж). Приставка **кросс-** означает, что таблица содержит по одному измерению в строках и столбцах, в данном случае это **Месяц** и **Объект**.

# ПРОСТЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: ГРАФИКИ

**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 переменной (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.

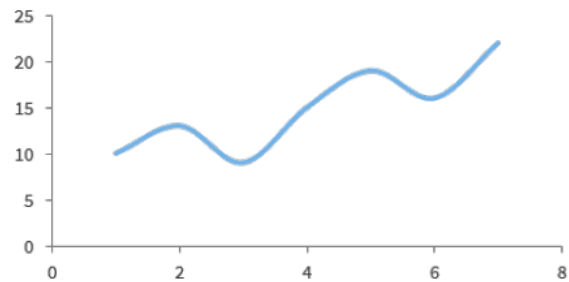
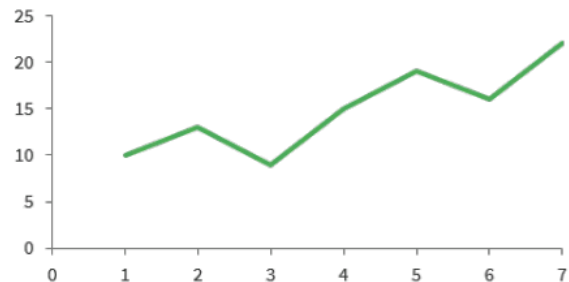
Графики особенно полезны при анализе временных рядов. Иногда одного взгляда на график достаточно, чтобы выявить наличие тренда, сезонной компоненты, оценить степень влияния случайной составляющей на исследуемый процесс.

# ПРОСТЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: ГРАФИКИ

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

На рисунке 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**.



Ломаный и гладкий графики

# ПРОСТЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: ГРАФИКИ

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.



График ежемесячных продаж за три года

# ПРОСТЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: ДИАГРАММЫ

Если 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.

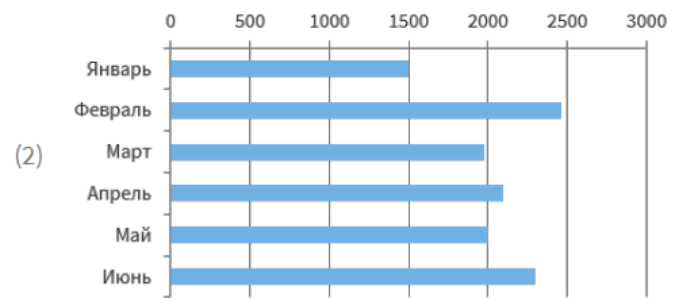
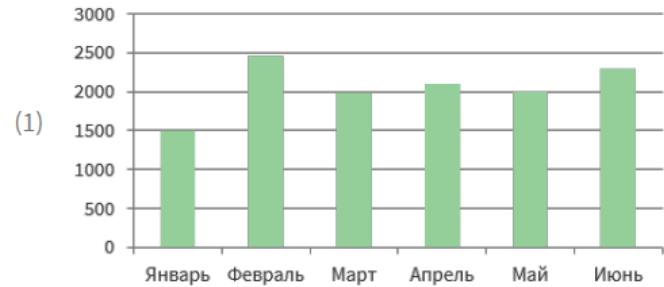


# ПРОСТЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: ДИАГРАММЫ

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).



# ПРОСТЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: 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).

# ПРОСТЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: ГИСТОГРАММЫ

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

Гистограмма строится следующим образом. Пусть исследуемой величиной являются ежедневные продажи торговой точки в течение месяца. При этом минимальное наблюдаемое значение составило 10 тыс., а максимальное – 100 тыс.

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.

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
...	...

# ПРОСТЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: ГИСТОГРАММЫ

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

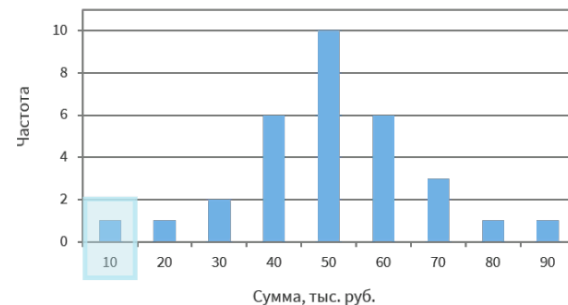
Как видим, продажи на сумму от 10 до 20 тыс. наблюдались только один раз, от 30 до 40 тыс. – два раза итак далее. На основе полученной таблицы строим гистограмму.

# ПРОСТЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: ГИСТОГРАММЫ

По горизонтальной оси гистограммы откладываются значения продаж, а по вертикальной – количество или частота наблюдений, значения которых попали в заданный диапазон.

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 можно использовать для восстановления пропущенных значений, 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.



# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ



**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.

# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ

Большинство реальных бизнес-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?**

# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ

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

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



# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: OLAP-АНАЛИЗ

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.

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**.

# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: OLAP-АНАЛИЗ

**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**). Измерениями могут быть: **Товар, Город, Клиент, Организация, Дата** и др.

# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: OLAP-АНАЛИЗ

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**). Примеры фактов: **Количество, Сумма, Возраст, Доход, Торговая наценка** и др.

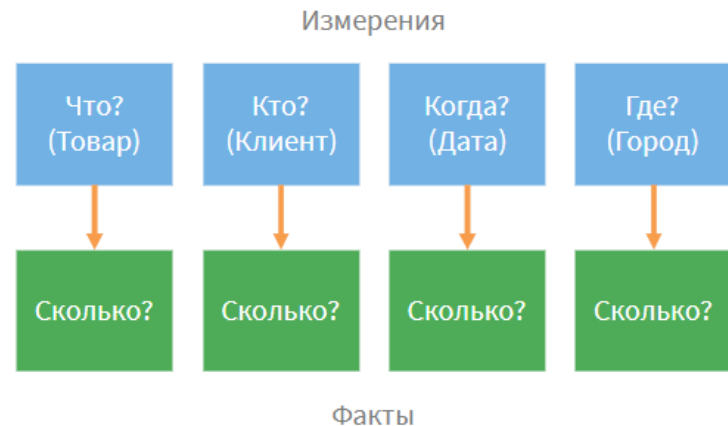
# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: OLAP-АНАЛИЗ

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?



# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: OLAP-АНАЛИЗ

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.

# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: OLAP-АНАЛИЗ

Наибольший интерес OLAP-куб представляет с точки зрения визуального анализа данных, поиска особенностей и закономерностей в данных. Для из поиска аналитик может задаться такими вопросами как:

- **Какой товар самый популярный?**
- **Где продажи были наилучшими и с чем это связано?** и так далее.

Затем с помощью манипуляций заголовками измерений выбирается такое представление куба, которое позволяет ответить на поставленные вопросы.

Последовательно перебирая возможные варианты представления данных аналитик выявляет закономерности и связи между элементами данных, выдвигает гипотезы и так далее.

# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: ГЕОГРАФИЧЕСКИЕ И ТЕПЛОВЫЕ КАРТЫ

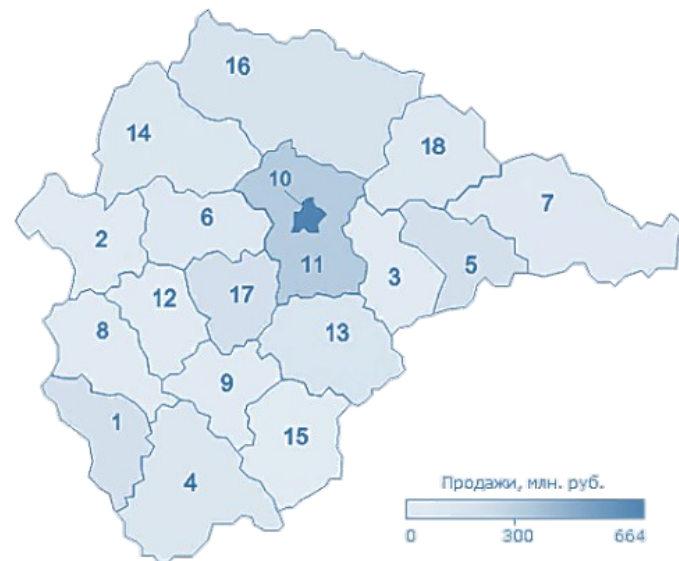
№ п/п	Область	Продажи, млн. руб.	№ п/п	Область	Продажи, млн. руб.
1	Белгородская	162	10	г. Москва	664
2	Брянская	61	11	Московская	328
3	Владимирская	54	12	Орловская	54
4	Воронежская	154	13	Рязанская	151
5	Ивановская	209	14	Смоленская	137
6	Калужская	117	15	Тамбовская	29
7	Костромская	90	16	Тверская	178
8	Курская	57	17	Тульская	186
9	Липецкая	50	18	Ярославская	120

**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.

# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: ГЕОГРАФИЧЕСКИЕ И ТЕПЛОВЫЕ КАРТЫ

Имея карту, 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).





# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: ГЕОГРАФИЧЕСКИЕ И 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.

Доходность:  Данные на: 03/11/2022 18:40

Yandex clA 2126.4 5.3 %	Татнфт Зап 373.1 3.2 %	Сургнфгз 21.585 2.8 %	Новатэк ао 1103.2 2.4 %	VK-гдр 455.8 2.2 %
Транснф ап 9100 2 %	OZON-адр 1344 1.9 %	ФосАгро ао 6448 1.9 %	РусГидро 0.7932 1.9 %	Сургнфгз-п 25.62 1.7 %
ГМКНорник 13444 1.3 %	TCS-гдр 2522 1 %	Мосбиржа 84.87 0.7 %	Татнфт Зао 384.4 0.5 %	Ростел -ао 57.54 0.3 %
Аэрофлот 25.62 0.2 %	МТС-ао 222.1 -0.2 %	FIVE-гдр 1592 -0.7 %	МКБ ао 6.312 -0.7 %	Сбербанк-п 119.37 -0.8 %
Сбербанк 125.31 -0.9 %	ИнтерРАОао 3.041 -0.9 %	ЛУКОЙЛ 4669 -1 %	НЛМК ао 103.78 -1.6 %	Магнит ао 5152 -1.8 %
ГАЗПРОМ ао 169.11 -2 %	GLTR-гдр 283.65 -2.2 %	ПИК ао 606.8 -2.3 %	ВТБ ао 0.01643 -2.3 %	ФИХР-гдр 354.6 -2.7 %
Роснефть 334 -2.8 %	АЛРОСА ао 65.12 -2.8 %	Система ао 12.391 -2.9 %	ДетскийМир 72.14 -3.1 %	Полюс 5910 -3.4 %
ММК 31.21 -3.5 %	Polymetal 283.7 -3.5 %	РУСАЛ ао 35.74 -3.6 %	СевСт-ао 784 -5.1 %	ЭН-ГРУП ао 393 -6.9 %

# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: ПЛОСКОЕ ДЕРЕВО



К семейству карт можно отнести и метод **плоское дерево** (англ.: treemapping). Он является очень эффективным при изображении численных атрибутов элементов (размер, стоимость, значение), организованных в большие иерархии.

# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: ДИАГРАММА СВЯЗЕЙ

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.

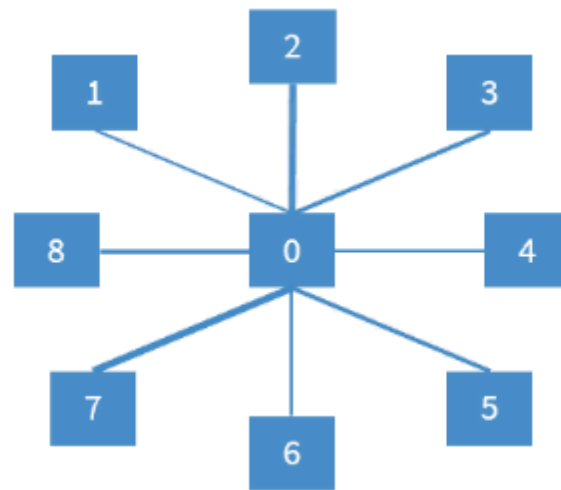
Чаще всего для этого используют:

- 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.

# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: ДИАГРАММА СВЯЗЕЙ

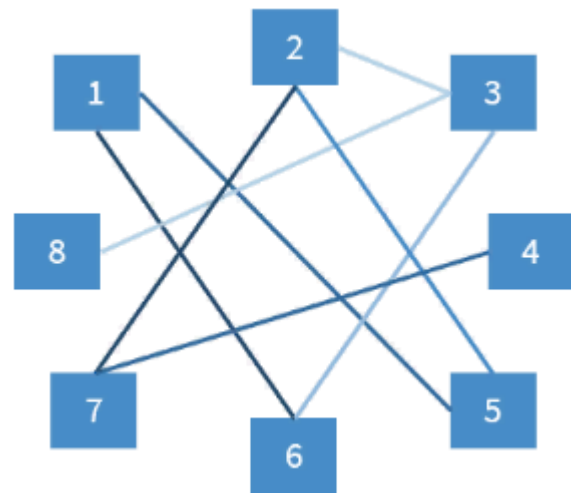
**Звезда** – в этом случае показываются связи между одним выбранным объектом, который помещен в центре, и всеми остальными.



# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: ДИАГРАММА СВЯЗЕЙ

**Сеть** – в этом случае визуализируются все попарные связи между объектами.

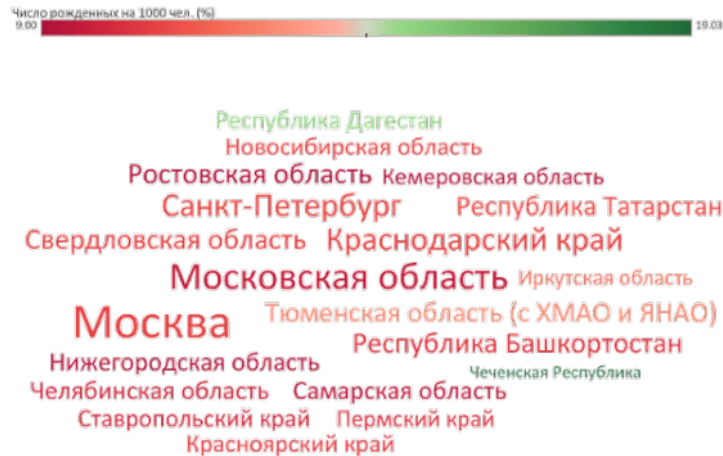
Если число объектов велико, связей становится еще больше, поэтому анализируют только наиболее сильные или наиболее слабые, накладывая фильтр на значения силы связи.



# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: ОБЛАКО ДАННЫХ

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

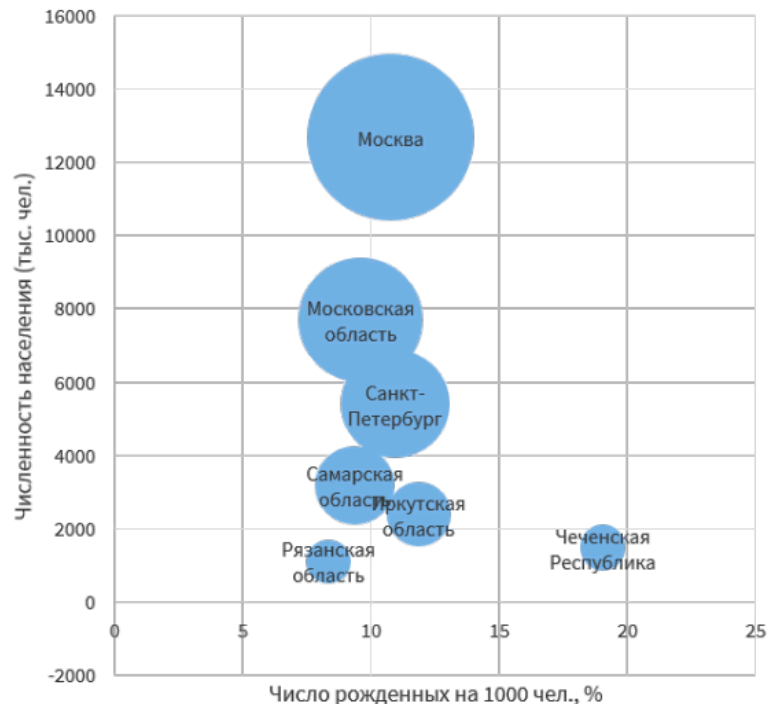
На рисунке приведен пример облака данных регионов РФ, в котором за цвет (в заранее выбранной шкале) отвечает **число рожденных на 1000 чел.**, %, а за размер шрифта – **численность населения**.



# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: 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.)**.



# СЛОЖНЫЕ ВИЗУАЛИЗАТОРЫ ОБЩЕГО НАЗНАЧЕНИЯ: 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.

