

# Введение в ВІ

Методичка к уроку 2

Из каких этапов состоит создание отчета?





## Оглавление

Как определить уровень развития BI системы	4
Квадрант Gartner'a	4
Методы расчёта	6
Лидеры квадранта Gartner	7
Power BI Desktop	12
Power BI Pro	12
Power BI Embedded	13
Power BI Premium	13
Power BI On Premise	13
Примеры архитектуры решений для разных версий	13
Дополнительные материалы	17
Домашнее задание	18

## Введение

На этой лекции вы найдете ответы на такие вопросы как / узнаете:

- Выбор инструмента, в котором мы работаем
- Выбор источника данных
- Преобразование первичных данных в табличный вид
- Расчет показателей при помощи формул
- Создание визуальных элементов
- UX/UI отчетов
- Что мы делаем с отчетом после того, как он готов



## Термины, используемые в лекции

BI (business intelligence) — автоматизация аналитических процессов для презентации данных в красивой форме.

Ad-hoc аналитика — аналитика, нужная в моменте, а не на постоянной основе. Ее готовят один раз по какому-то поводу и не обновляют.

ERP-системы (enterprise resource planning system) — система планирования ресурсов предприятия. Предназначена для автоматизации управленческого учёта. На нашем рынке типичная ERP-система — это 1C.

CRM-система (customer relation management) — система ведения и управления отношениями с клиентами. В ней сохраняется история взаимоотношений с клиентами: тендеры, продажи, поставки, контакты. Для малого и среднего бизнеса на нашем рынке распространены Bitrix и AmoCRM, для крупных компаний — Dynamics (Microsoft) и Terrasoft.

Веб-аналитика — системы, позволяющие собирать и анализировать данные о посетителях сайтов.

Call tracking — системы, отслеживающие звонки клиентов, поступившие из онлайн-ресурсов. Помогают определить источник, который привел к звонку

SQL — язык программирования для управления базами данных.

Excel — программа для работы с табличными данными.

Enterprise — очень крупные компании: например, Магнит, Nestle, Сбербанк, Лукойл.

DAX — язык программирования внутри Power BI для преобразования данных.

Power Query — язык программирования внутри Power BI для загрузки данных.

Оlap-кубы (online analytical processing) — многомерный массив данных. Это звено, завершающее облик решения по созданию и обслуживанию хранилищ данных.

Коннектор — программа, позволяющая автоматически забрать данные из конкретного источника данных.

Справочник — таблица с набором уникальных значений. Позволяет объединить между собой 2 или более таблиц в единую базу данных.



Дашборд — интерактивный отчёт, в котором можно «играть» с данными в режиме онлайн.

B2B (business to business) — виды бизнеса, в которых взаимодействие идёт между компаниями, без физических лиц.

## Выбор инструмента, в котором мы работаем

Перед тем, как мы приступим к работе, нам нужно выбрать тот инструмент, в котором мы будем создавать отчет.

Для этого давайте определим объективные показатели качества инструмента - международные рейтинги и цифры на нашем рынке.

## Как определить уровень развития BI системы

### Квадрант Gartner'a

Основной источник для рейтинга BI систем.

Gartner — это крупная исследовательская компания, основанная в 1979 году и специализирующаяся на IT-рынке. Выпускает популярные рейтинги IT-систем, в том числе и по BI. Рейтинги составляются в виде магического квадранта.

В 2009 году [Gartner выиграла судебный иск](#) по поводу легитимности магических квадрантов. Компания zdNet подала в суд, так как посчитала, что её BI-система была недооценена со стороны Gartner.

Квадрант Gartner на 2010 год:



Квадрант Gartner за 2020 год:

Figure 1. Magic Quadrant for Analytics and Business Intelligence Platforms





Квадрант выделяет двух явных лидеров — это Power BI (Microsoft) и Tableau. Историческим лидером был Tableau. Power BI начал активно развиваться чуть позже, но быстро стал лучшим решением по мнению Gartner.

Третий игрок на рынке — это Qlik, он постепенно удаляется от двух лидеров.

Также мы видим, что ThoughtSpot находится близко к Qlik, но это новый игрок на рынке и в России он не представлен. В рамках подготовки не был выявлен ни один публичный кейс решения на ThoughtSpot в России.

## Методы расчёта

На схеме две оси: по оси X считается полнота видения, по оси Y — технические возможности инструмента. Они оцениваются по ряду критериев.

Полнота видения:

1. Насколько решение соответствует потребностям рынка?
2. Какова стратегия продвижения системы?
3. Как выстраивается стратегия продаж и ценообразования?
4. Как развивается продукт? Применяются ли инновационные технологии?
5. Решение соответствует задачам предприятий разного масштаба?
6. Есть ли особенность географической стратегии и планы по развитию в других странах?

Технические возможности:

1. Насколько просто и удобно интегрировать систему в рабочий процесс?
2. Что отличает платформу от конкурентов?
3. Соответствуют ли решения практическим задачам пользователей?
4. Продукт оперативно реагирует на тенденции рынка?
5. Пользоваться им комфортно?
6. Легко ли его обновлять?
7. Какие особенности есть у техподдержки?
8. Что говорят пользователи?

В квадранте есть 4 квадрата, по которым распределяются инструменты в зависимости от их развития:

1. Нишевые игроки — самый низкий уровень, решения, которые подходят для узкого круга задач.



2. Претенденты — хороший технический уровень, но низкое понимание бизнес-потребностей.  
Бывает довольно редко, все представители на 2020 год близки к центру
3. Провидцы — понимают рынок, но уровень возможностей низкий.
4. Лидеры — продукты, которые лучше всего удовлетворяют потребности заказчиков.

Важно! Нас интересуют только решения в квадрате лидеров, так как они внедрены в наибольшем количестве компаний.

## Лидеры квадранта Gartner

### [Power BI:](#)

1. Разработка Microsoft;
2. Лидер квадранта последних лет;
3. Активно развивается с 2015 года;
4. Хорошо интегрируется с другими программами Microsoft;
5. Языки: DAX, Power Query, R, Python.

### [Tableau:](#)

1. Независимая разработка;
2. Близко к Power BI в квадранте все последние годы;
3. Исторический лидер до развития Power BI;
4. Компании, которые внедрили BI давно, часто до сих пор работают в Tableau.

### [QlikView:](#)

1. Независимая разработка;
2. Индивидуальный язык работы с данными;
3. Enterprise-решение;
4. Падает в рейтинге квадранта.

Из каких этапов состоит создание отчета



На февраль 2021 года лидером остался только Power BI:



На 2022 год лидерство Power BI только укрепляется:



Figure 1: Magic Quadrant for Analytics and Business Intelligence Platforms

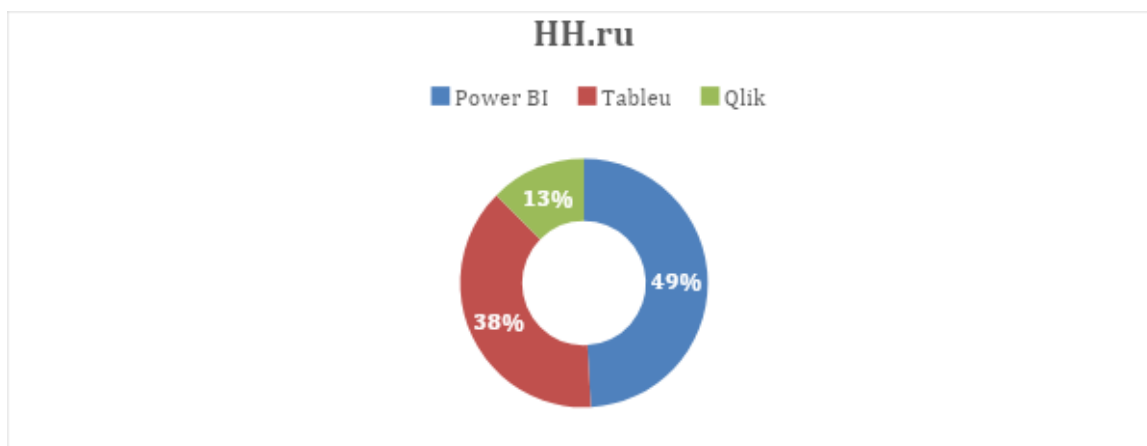


Source: Gartner (March 2022)

Чтобы подтвердить данные квадранта на российском рынке, проведём небольшое исследование: проверим, какие BI-инструменты ищут чаще на январь 2021 года.

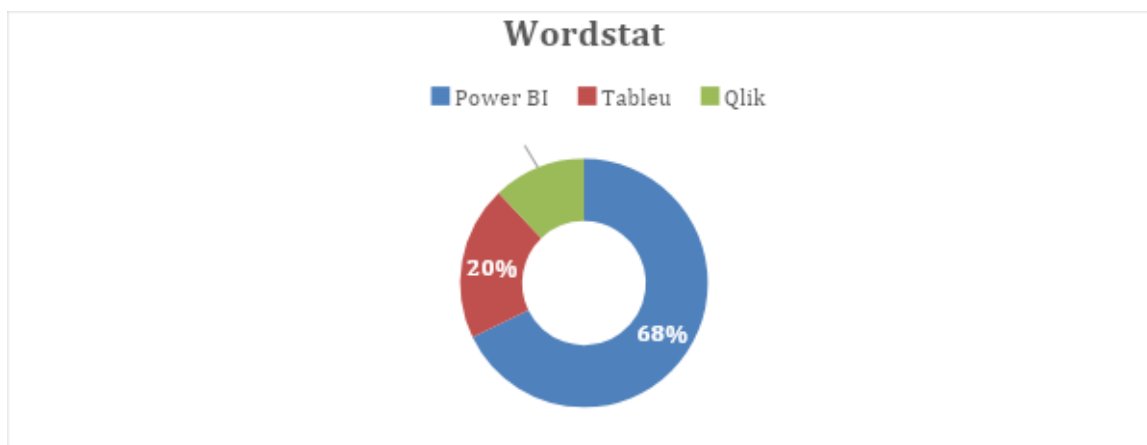
Сперва зайдём на hh.ru и посмотрим количество вакансий, где в качестве необходимого навыка указаны Power BI, Tableau или Qlik. Увидим следующую картину:

- Power BI — 609 вакансий;
- Tableau — 476 вакансий;
- Qlik — 155 вакансий.



Затем проверим количество запросов в Wordstat (сервис для оценки пользовательского интереса). Там разрыв ещё больше:

- Power BI искали 38 602 раз;
- Tableau — 11 434;
- Qlik — 6 888.



Теперь дополним эти цифры разрезами за апрель 2021 и май 2022 года, когда рынок уже поменялся и для сравнения добавим лидирующие отечественные решения Luxms BI и Yandex DataLens.

Для начала возьмем количество вакансий в HH.ru:



Название	Январь 2021	Апрель 2021	Май 2022
Power BI	609	1 657	634
Tableau	476	1 079	410
Qlik	155	441	167
DataLens			20
Luxms BI			5

Теперь давайте посмотрим на количество запросов в Wordstat:

Название	Январь 2021	Апрель 2021	Май 2022
Power BI	38 602	64 442	60 334
Tableau	11 433	17 720	18 811
Qlik	6 888	12 206	11 634
DataLens			3 891
Luxms BI			387

Таким образом по всем объективным показателям мы видим, что Power BI - это самый качественный ликвидный и популярный инструмент, соответственно именно в нем логично развиваться при прочих равных.



## Версии Power BI

Теперь давайте разберемся с тем какие есть версии Power BI.

Важно! Первое, что необходимо понять — под Power BI можно подразумевать разные вещи: инструмент для создания отчётов, инструмент для совместной разработки отчётов, портал с готовыми отчётами и даже сервер для хранения данных.

Существуют пять основных версий Power BI. Для нас ключевыми на старте являются две из них.

### Power BI Desktop

Это основной инструмент для работы с отчётом. Здесь вы будете создавать отчёт перед тем, как его показать. У Power BI Desktop есть две версии: обычная (с жёлтым логотипом) и для On Premise версии (с чёрным логотипом).

Power BI Desktop — это:

1. Бесплатная программа.
2. Основной инструмент для работы и создания отчётов.
3. Нужна для первоначальной разработки отчётов.
4. Постоянно обновляется, нужно отслеживать обновления на сайте Microsoft.

### Power BI Pro

Самый популярный инструмент для совместной разработки отчётов и для шеринга (share — делиться)..

Power BI Pro:

- Облачный продукт.
- Продаётся по подписке.
- Стоит порядка 1000 рублей в месяц.
- Входит в Office 365 версии E5.
- Подходит для совместной разработки.
- Позволяет делиться отчётами.
- Одной лицензии достаточно, чтобы делиться отчётом через прямую ссылку.



Из сочетаний первых двух рассмотренных версий построено наибольшее количество решений на Power BI.

## Power BI Embedded

- Продаётся по подписке.
- Является приложением внутри облачного продукта Azure.
- Позволяет встроить целое приложение Power BI в портал или сайт.
- Даёт дополнительные возможности для веб-разработки поверх BI.
- Минимальная стоимость в год — порядка 500 000 рублей (300 действий с отчётами в час).
- Стоимость зависит от плотности использования: платим за действие (в терминах Microsoft — «рефреш»).

## Power BI Premium

- Продаётся по подписке.
- Стоит порядка 5 000 \$ в месяц.
- В дополнение к BI даёт дополнительные серверные мощности.
- Подходит для Enterprise-компаний с высокой нагрузкой на BI-систему.
- Передовой продукт Microsoft.

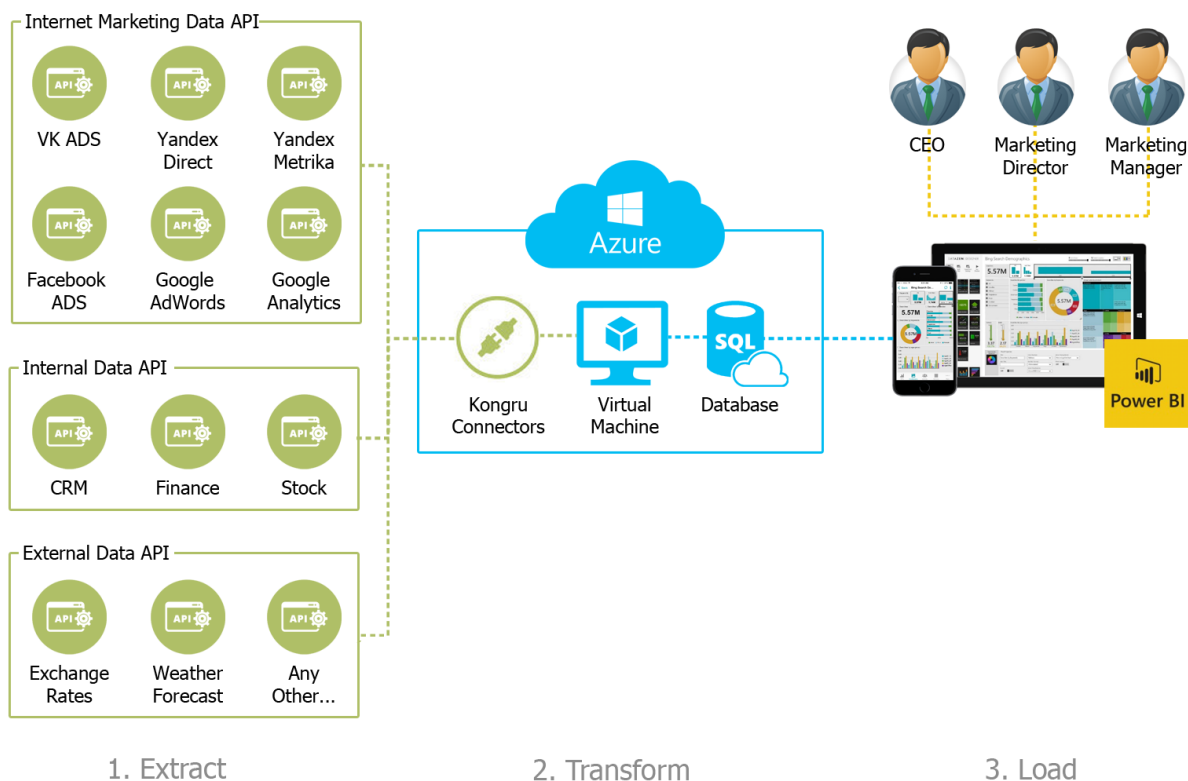
## Power BI On Premise

- Продаётся по подписке.
- Не облачный продукт.
- Подходит для компаний, которые не хотят выкладывать данные в облако.

## Примеры архитектуры решений для разных версий

Архитектура решения — это схема, на которой видно, как данные поступают в BI-систему. Архитектура важна для случаев, когда данные мы берём из более, чем двух систем, и где есть нюансы с хранением, прокачкой и защитой данных - нужно определять способ обработки данных с учетом объема и прироста объема данных.

Первая архитектура — это архитектура решения по построению сквозной аналитики интернет-маркетинга:

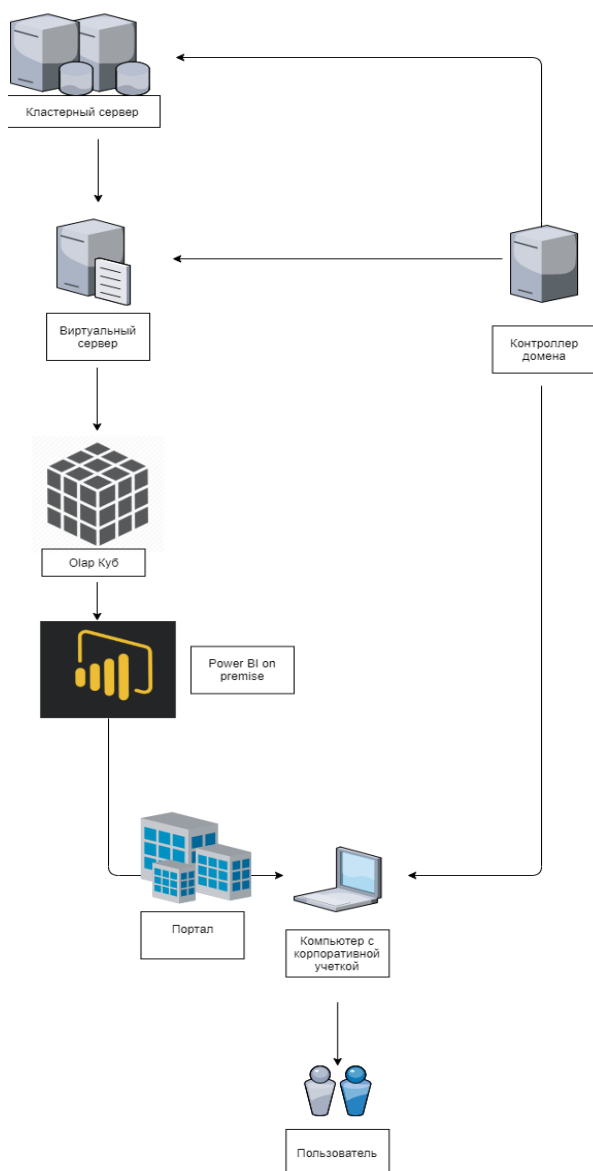


Эту архитектуру нужно смотреть слева направо. В ней выделены три этапа работы с данными:

1. Extract — левая часть, где перечислены различные системы. На этом этапе наша задача — начать собирать данные из каждой. Системы разделены на типы данных: внешние (прогноз погоды, курсы валют), внутренние (данные компании в CRM и ERP) и смешанные, когда данные видны и вашей компании и компании, предоставляющей софт (Google Analytics, кабинеты контекстной рекламы и прочее).
2. Transform — этап преобразования данных. Они приводятся к единому формату и при необходимости объединяются в одну базу данных. Обработываются в облачном сервере Microsoft.
3. Load — загрузка готовых данных в BI. На их основе создаются отчёты, а после — показываются разным категориям сотрудников через Power BI Pro.

Вторая архитектура — это решение на Power BI On Premise для анализа продаж:

## Из каких этапов состоит создание отчета



У нас есть кластерный сервер, где хранятся исходные данные.

Затем они поступают в виртуальный сервер, чтобы их можно было обработать и привести к единому виду.

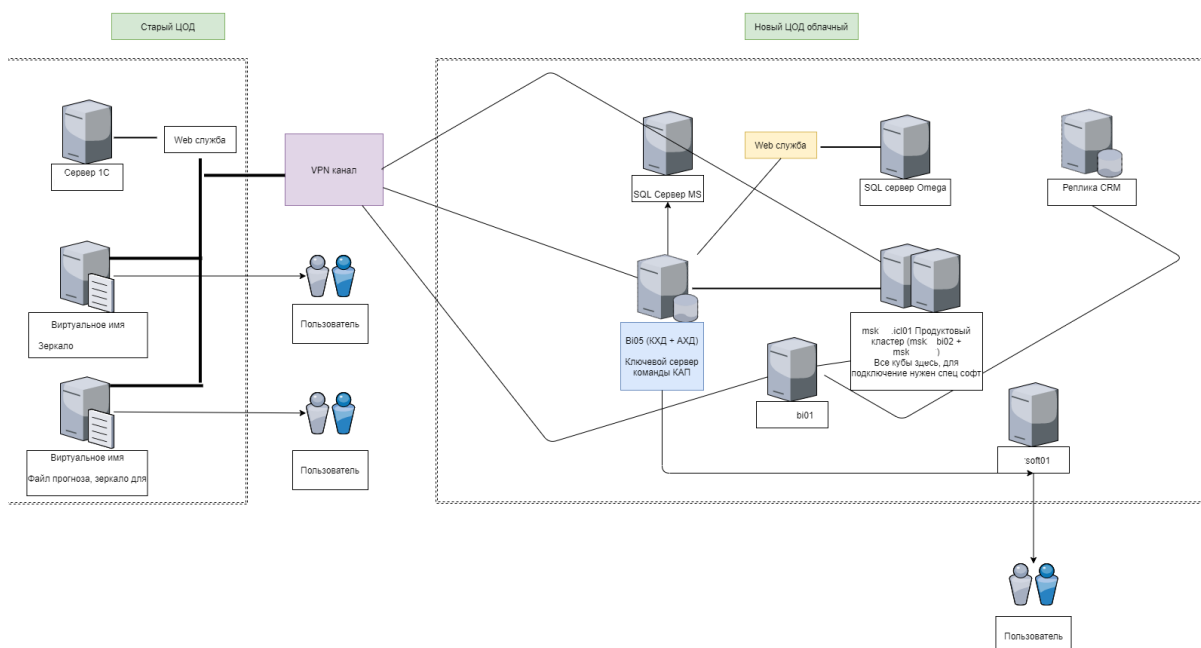
Затем — в Olap-кубы, где сегментируются на различные таблицы. Это нужно, чтобы не добавлять лишние данные в отчёты.

Затем готовые данные поступают в Power BI On Prem.

Важно! В этой архитектуре и в этой версии Power BI On Prem используется только в качестве визуализатора данных. Вся работа с данными происходит на этапе Olap-кубов.

В конце отчёты выкладываются на корпоративный портал и с ограничением от контроля домена могут быть показаны разным сотрудникам.

Третья архитектура — это архитектура бэкенда по обработке данных из многих систем:



Она описывает сложный процесс сбора, обработки и хранения данных из различных корпоративных систем. В первую очередь будет нужна команде разработки баз данных и предшествует работе команды BI.

Ценность архитектуры для нас в том, что мы можем понять, какие есть серверы и как они называются. Это нужно, чтобы мы могли забирать правильные данные из правильных систем. Например, если у нас задача забрать данные из SQL-сервера Omega (база данных HR-подразделения), то в Power BI мы будем искать сервер под названием BI05, так как именно туда поступают данные из Omega.

## Следующие этапы

Следующие этапы работы с отчетом мы будем подробнее проходить на следующих лекциях и семинарах, а сегодня познакомимся с базовым описанием этих процессов.

- 1) Загрузить данные можно по готовому коннектору или напрямую, основные источники данных - это:



Базы данных	Интернет	ERP, CRM	Файлы
      	    	   	      

- 2) Одна из самых важных вещей, которые делаем BI - это преобразование данных, например преобразование из не таблицы в таблицу
- 3) После подготовки данных, нам необходимо сделать математические расчеты, в Power BI, например мы это будем делать в DAX
- 4) После того, как данные готовы, нам нужно их визуализировать, в Power BI за это отвечает Power View
- 5) Когда мы создали набор из визуальных элементов, нам нужно объединить их в полноценные отчеты и наборы отчетов
- 6) После того, как отчет готов, нам нужно внедрить его в корпоративные системы

## Дополнительные материалы

- [Что такое BI и зачем это нужно вашему бизнесу.](#)
- [Что такое Business Intelligence.](#)
- [6 этапов грамотного внедрения BI-системы.](#)
- [Tableau от чайника или визуализация анализа ПАММ счетов](#)

Лидеры на рынке BI-систем:

- [Microsoft Power BI](#)
- [Qlik](#)
- [Tableau](#)

Из каких этапов состоит создание отчета



Книга по логике построения диаграмм:

- [Джин Желязны. Говори на языке диаграмм: пособие по визуальным коммуникациям.](#)

## Домашнее задание

Ссылка для скачивания [Power BI Desktop](#).