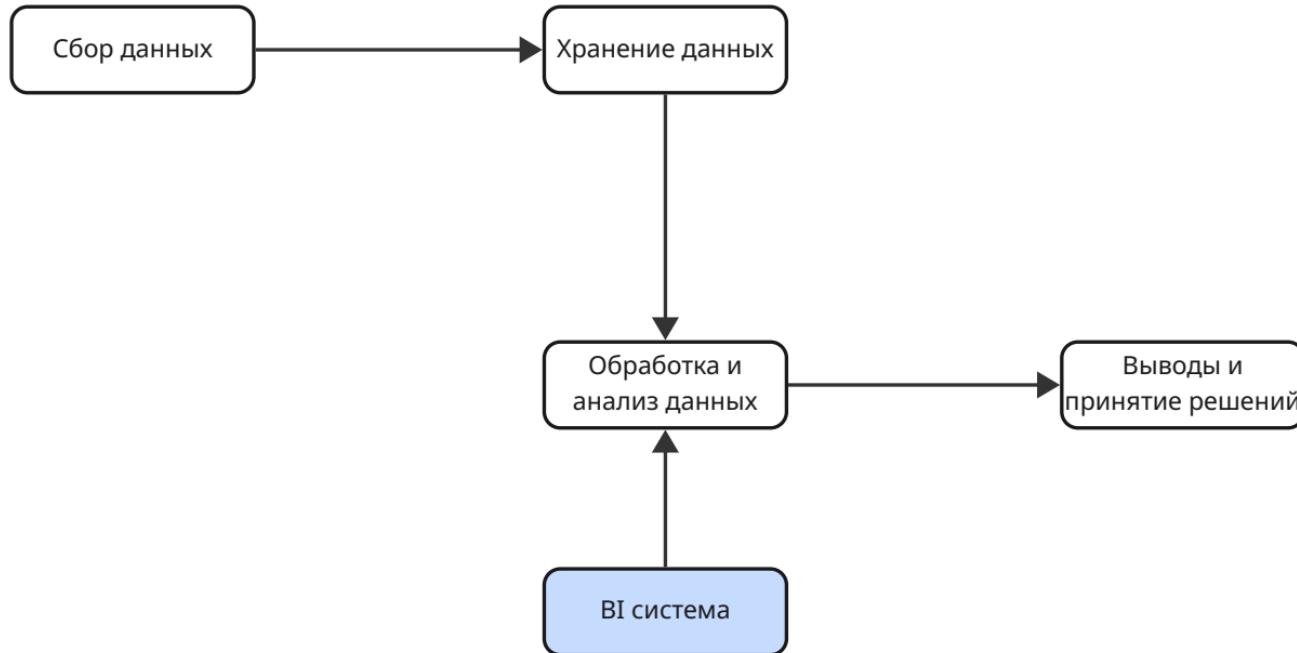


Business Intelligence

Business Intelligence - это процесс анализа данных для поддержки бизнес-решений

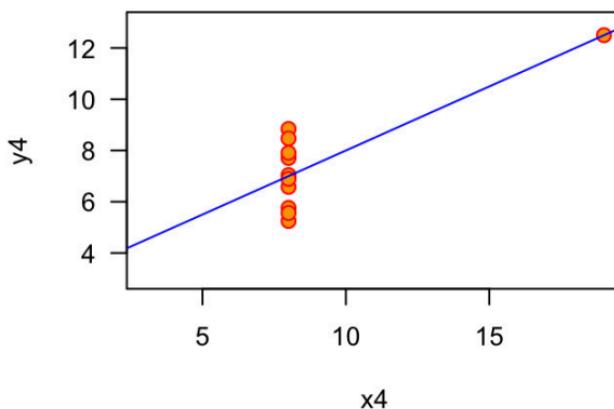
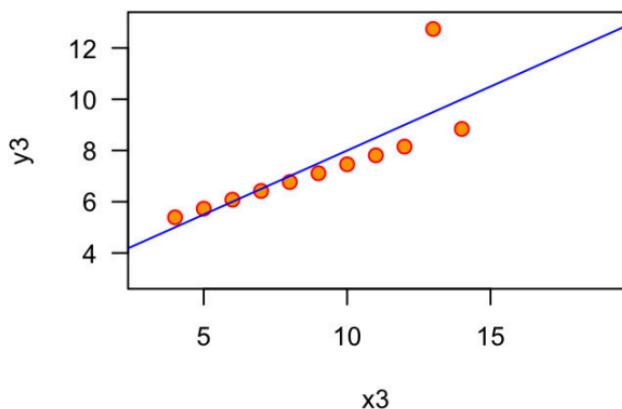
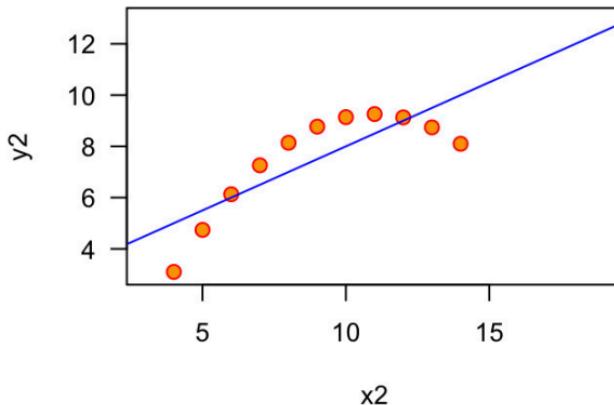
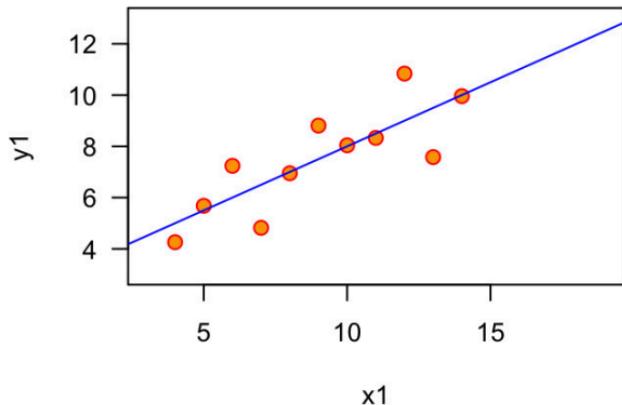


Визуализация

1. Понятное и наглядное представление сырых данных
2. Способ коммуникации между данными и пользователем
3. Позволяет найти инсайты и взаимосвязи в данных
4. Быстрый анализ большого массива данных
5. Влияет на скорость принятия решений



Anscombe's 4 Regression data sets





Data Driven Approach

Подход к управлению и принятию решений, основанный на данных.
Решения нужно принимать, опираясь на анализ данных, а не интуицию

Уровни аналитических данных

Оптимизация	Что самое хорошее может произойти?
Прогнозное моделирование	Что произойдет дальше?
Экстраполяция	Что, если этот тренд продолжится?
Статистический анализ	Почему это происходит?
Оповещения	Какие действия следует предпринять?
Детализация по запросу	В чем именно проблема?
Специальные отчеты	Как? Что? Где?
Стандартные отчеты	Что произошло?

Визуализация в работе аналитика данных



EDA

Разведочный анализ
данных



Ad-Hoc

Аналитика по запросу



Дашборды

Автоматизированная
отчетность

BI - инструмент для создания интерактивной обновляемой отчетности



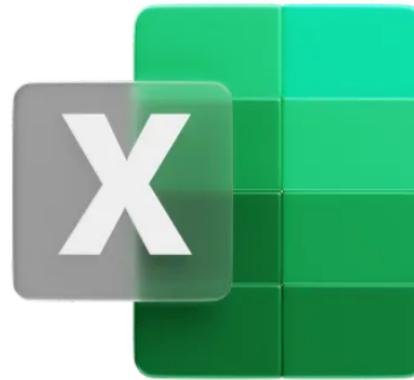
Почему BI система

1. Подключение данных из разных источников
2. Объединение данных из разных источников
3. Принимает большой массив данных
4. Автоматическое обновление данных
5. Наличие встроенных функций для быстрой работы с данными
6. Режим публикации и просмотра
7. Безопасность данных
8. Интерактивность



BI системы

Tableau, Power BI, DataLens, Superset



Прочие инструменты

Excel, Google Sheets, Python, R

Tableau

Состав

- Tableau Desktop
- Tableau Public
- Tableau Prep
- Tableau Online
- Tableau Server

Преимущества

- Интуитивный интерфейс
- Возможности интеграции
- Аналитические функции
- Обширное сообщество

Недостатки

- Высокая стоимость
- Крутая кривая обучения
- Ушел из РФ

Power BI

Состав

- PBI Desktop
- PBI Premium
- PBI Mobile
- PBI Service
- PBI Pro
- PBI Embedded

Преимущества

- Интеграция с экосистемой Microsoft
- Возможности визуализации
- Наличие PQ/PP
- Регулярные обновления
- Масштабируемость

Недостатки

- Приложение для построения дашбордов поддерживается только для Windows
- Высокая стоимость
- Ушел из РФ

DataLens

Состав

- DataLens Yandex Cloud

Преимущества

- Гибкость и масштабируемость
- Интеграция с продуктами Yandex
- Открытый исходный код
- Понятный интерфейс

Недостатки

- Ограниченные возможности в сравнении с лидерами мирового рынка
- Зависимость от экосистемы Yandex

SuperSet

Состав

- Apache Superset

Преимущества

- Бесплатный
- Интеграция с различными источниками данных
- Открытый исходный код

Недостатки

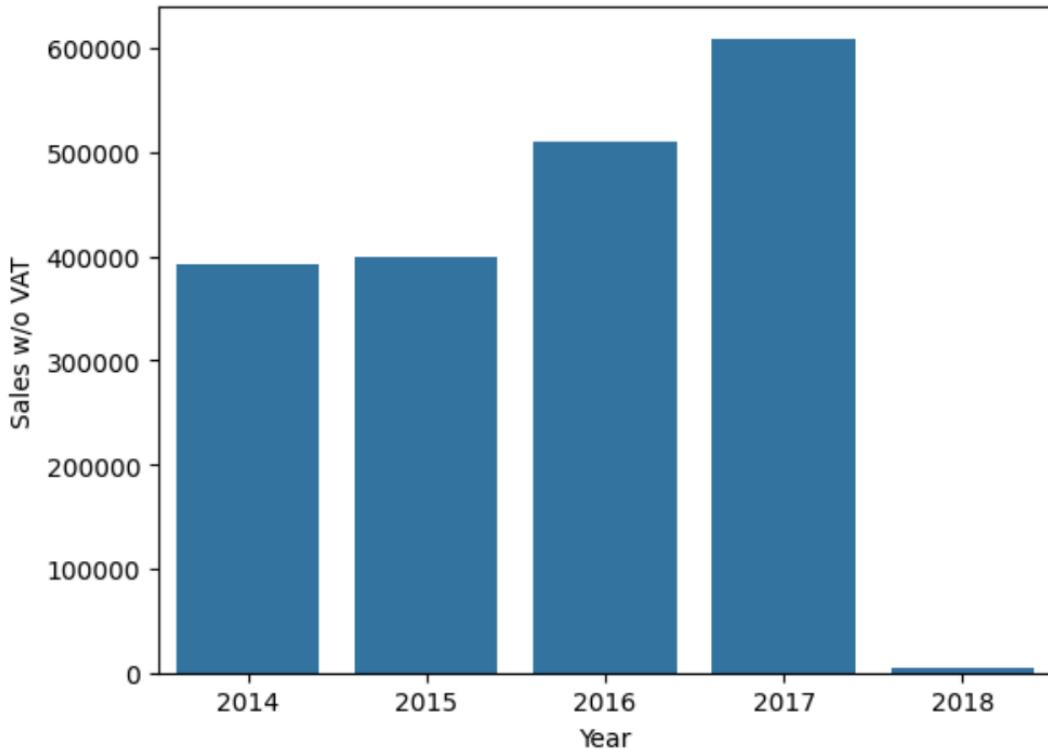
- Требует технической экспертизы
- Ограниченнная поддержка
- Мощность
- Меньше функций

Какой инструмент выбрать?

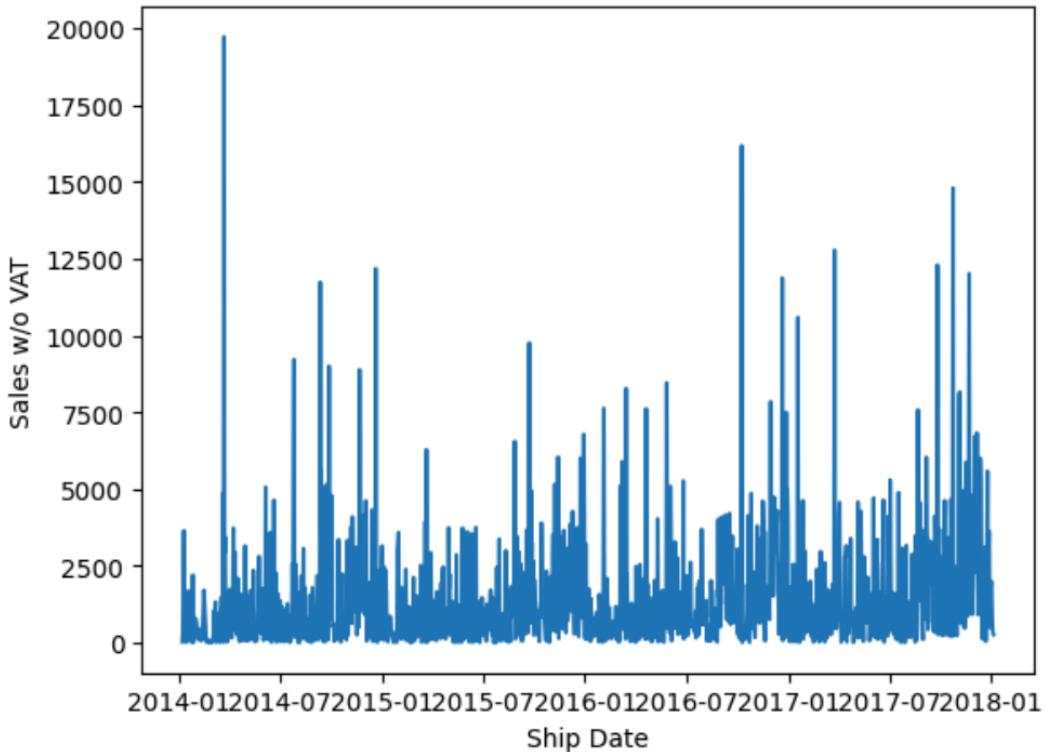
1. Потребности компании
2. Бюджет
3. Возможности интеграции
4. Возможности маштабируемости
5. Безопасность
6. Ваши навыки
7. Сообщество

Виды графиков

Столбиковая диаграмма



Линейная диаграмма



Гистограмма

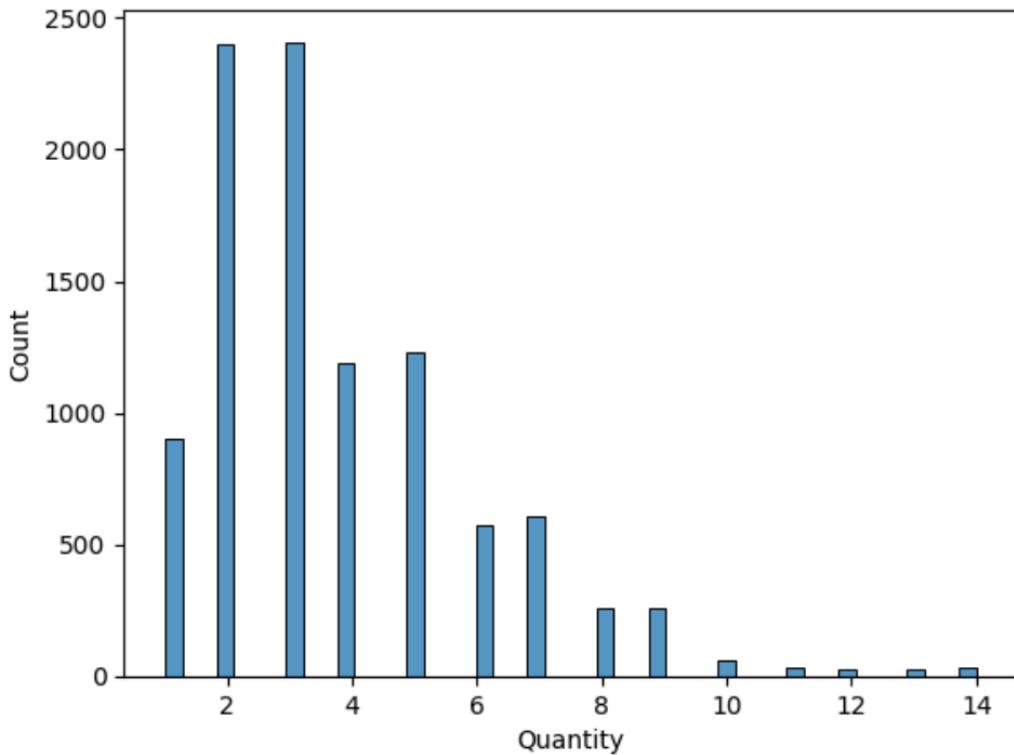
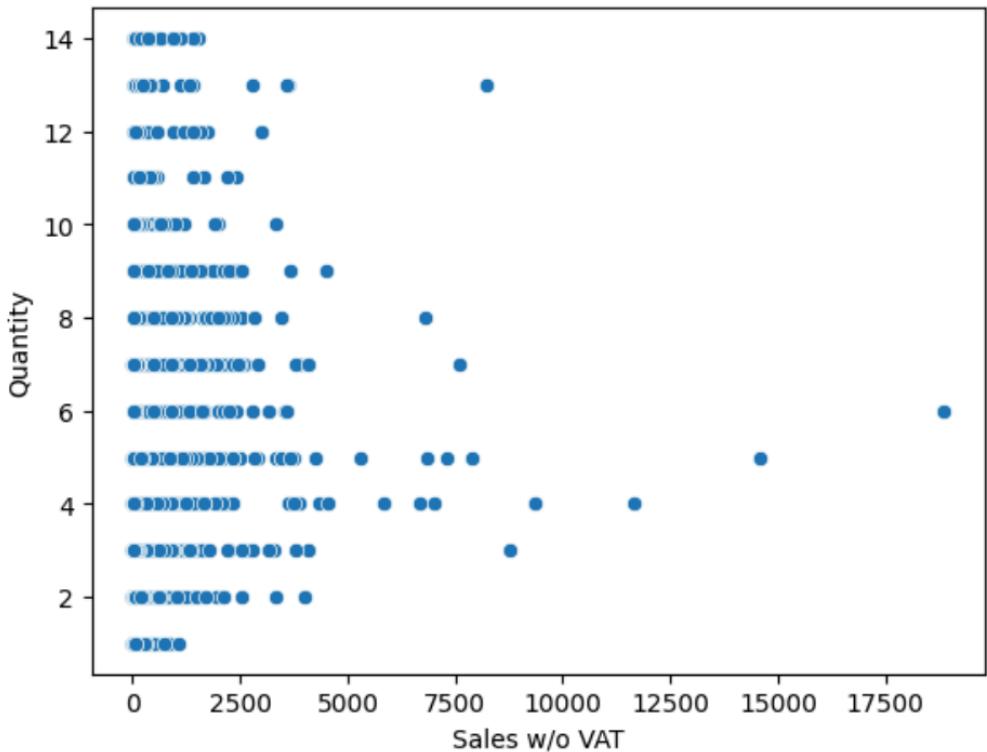
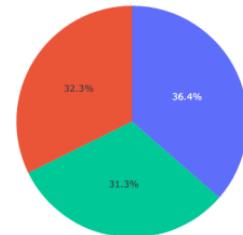


Диаграмма рассеивания



Круговая диаграмма



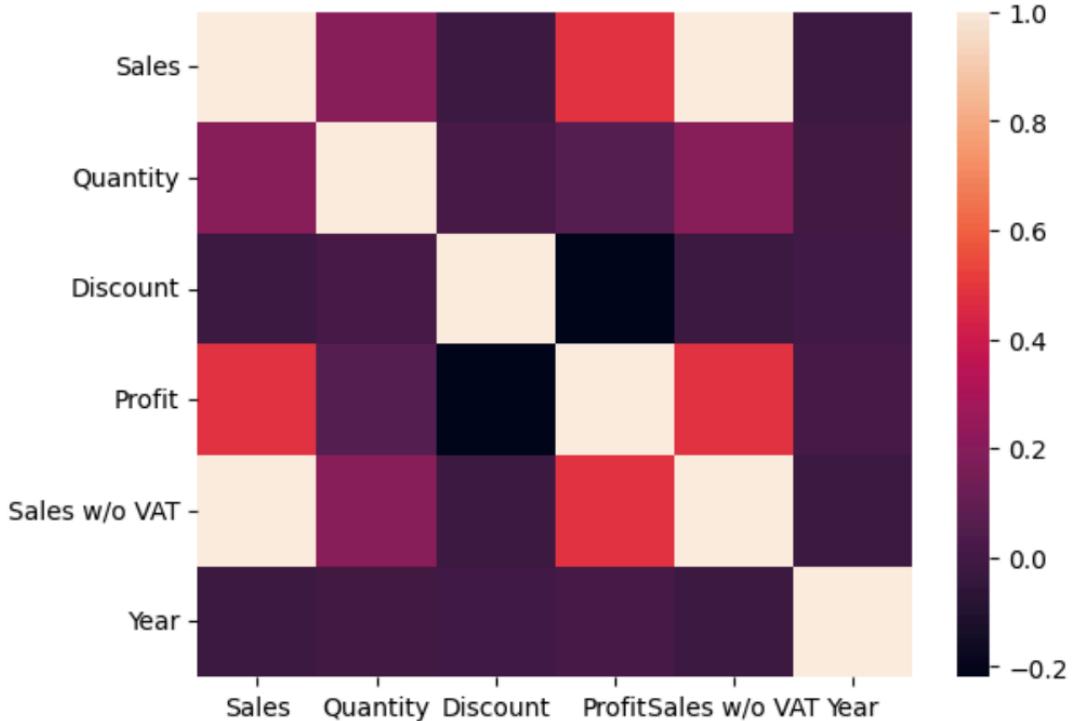
■ Technology
■ Furniture
■ Office Supplies

Sunburst

Sales w/o VAT by Category and Sub-Category



Heatmap



Таблица

	Sales	Quantity	Discount	Profit	Sales w/o VAT	Year
Sales	1.000000	0.200795	-0.028190	0.479064	1.000000	-0.010815
Quantity	0.200795	1.000000	0.008623	0.066253	0.200795	-0.006716
Discount	-0.028190	0.008623	1.000000	-0.219487	-0.028190	-0.002279
Profit	0.479064	0.066253	-0.219487	1.000000	0.479064	0.004531
Sales w/o VAT	1.000000	0.200795	-0.028190	0.479064	1.000000	-0.010815
Year	-0.010815	-0.006716	-0.002279	0.004531	-0.010815	1.000000

Рекомендации

1. Выделяйте важное контрастным цветом
2. Для разных категорий используйте свои цвета
3. Не используйте шкалу и подписи данных вместе
4. Не используйте подписи данных, если они плохо читаемы
5. Используйте не более 5 категорий, остальные в прочее
6. Выводите на график не более 5 компонентов
7. Координатная сетка и линии осей не должны забирать внимание
8. Разделяйте разряды пробелом или запятой
9. Не используйте объемные визуализации

Критерии удачной визуализации

1. Точно передает данные
2. Легко считывается смысл без дополнительной помощи
3. Компактные
4. Лишены декоративных элементов, усложняющих понимание

Список хороших книг

1. Александр Богачев "Графики, которые убеждают всех"
2. Кайро Альберто "Графики лгут"

Список хороших каналов

1. <https://t.me/revealthedata>
2. <https://t.me/nastengraph>

Спасибо!