



Прикладные инструменты визуализации данных

Андрей Куртасов
Системный аналитик

Проверка связи



Отправьте «+», если меня видно и слышно

Если у вас нет звука или изображения:

- перезагрузите страницу
- попробуйте зайти заново
- откройте трансляцию в другом браузере (используйте Google Chrome или Microsoft Edge)
- с осторожностью используйте VPN, при подключении через VPN видеопотоки могут тормозить

Чем сегодня займемся?



1. Попрактикуемся в использовании библиотеки Altair.
2. Рассмотрим понятие Business Intelligence.
3. Познакомимся с Yandex DataLens.

Библиотека Altair

Опрос

1. Вы ознакомились с материалом об Altair из лекции?
2. Испробовали библиотеку на практике?



- Библиотека [Altair](#) создает графику согласно спецификации [Vega-Lite](#)
- На выходе – JSON-структуры, обрабатываемые JS-библиотекой Vega-Lite
- JSON генерируется объектом Chart:

```
alt.Chart(df) \
    .mark_point() \                # Метки
    .encode(x='col_1', y='col_2') \ # Кодировки
    .properties(width=500) \      # Свойства
    ...() # Множество других методов...
```

- [Пример на Colab](#)

Виды меток

- `mark_point()`
- `mark_circle()`
- `mark_square()`
- `mark_line()`
- `mark_area()`
- `mark_bar()`
- `mark_tick()`

См. altair-viz.github.io/user_guide/marks/index.html

Тип данных	Код	Описание
quantitative	Q	Числовая величина (действительная)
nominal	N	Наименование / неупорядоченный категориальный
ordinal	O	Упорядоченный категориальный
temporal	T	Дата / время

```
alt.Chart(df).mark_tick().encode(  
  x='column_1:Q',  
  y='column_2:N',  
  color='column_3:O'  
)
```

- Изначально тип определяется автоматически по DataFrame
- Поддерживаются группировка и агрегация данных

Упражнение 1

Дан датасет об эксперименте по выращиванию ячменя:

```
from vega_datasets import data  
df = data.barley()
```

Вывести диаграмму рассеяния с двумя рядами данных –
распределения значений урожая (`yield`) в 1931 и 1932 гг.

Упражнение 2

Сделать график интерактивным и добавить фильтр по колонке `variety`.

Упражнение 3

Построить столбчатую диаграмму, показывающую средний урожай (за оба года) каждого сорта ячменя.

Business Intelligence

Понятие BI и популярные инструменты

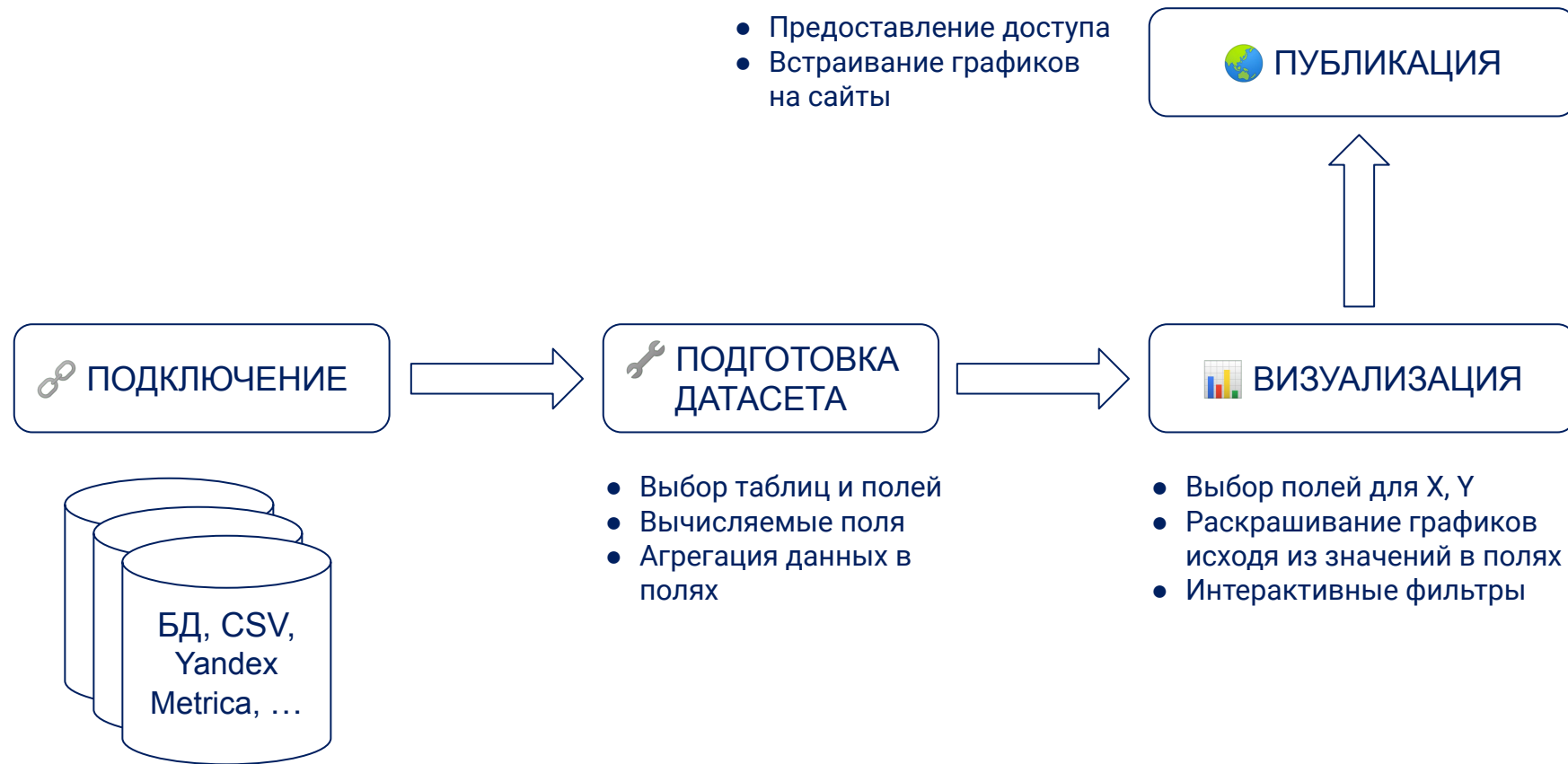
Бизнес-аналитика (Business Intelligence) — совокупность методов и инструментов, обеспечивающих перевод данных, имеющих ценность для бизнеса, в человекочитаемую форму.

- Tableau
- Microsoft Power BI
- Yandex Data Lens
- Google Looker Studio

Приходилось ли вам
работать с подобными
инструментами?



Сценарий использования VI-системы



- Сервис доступен бесплатно по адресу datalens.yandex.ru (нужна учетная запись в Яндексе)
- В [маркетплейсе](#) можно посмотреть примеры дашбордов
- Пример публичного дашборда с автоматическим получением данных через API: [вакансии в университетах РФ](#)



Итоги занятия

Итоги занятия



1. Познакомились с библиотекой Altair.
2. Познакомились с Yandex DataLens.

Свободная дискуссия

Вам слово!

- Что показалось наиболее ценным из изученного сегодня?
- Что хотелось бы изучить подробнее?
- Любые соображения по теме визуализации данных



До встречи!

