

# Прикладная статистика

Программа курса



ТИНЬКОФФ



# Программа курса



## Необходимые вводные лекции

1. Базы данных (28.09.2021)
2. Базовые сведения из теории вероятностей и математической статистики (05.10.2021)
3. Визуализация данных (12.10.2021)

Здесь работаем над инструментарием:

- Математические понятия
- Базы данных и как с ними работать
- Строим графики, таблицы

Всё с примерами и практикой!



## A/B-тесты

4. Оценка параметров (19.10.2021)
5. Продуктовая аналитика (26.10.2021)
6. AB-тесты, часть 1 (02.11.2021)
7. Основы проверки гипотез (09.11.2021)
8. AB-тесты, часть 2 (16.11.2021)

- Что вообще тестируем?
- Гипотеза, дизайн теста
- Анализ результатов
- А если всё поломалось?
- Принятие решения



## Вопросы

9. Семинар с ответами на вопросы (23.11.2021)

- Многомерная линейная регрессия, логистическая регрессия
- Алгоритм, обучающий деревья решений
- Методы оценки качества моделей
- Переобучение



## Финальные лекции

10. Временные ряды (30.11.2021)
11. Регрессии (07.12.2021)
12. Случайный лес и деревья решений (14.12.2021)

# Вводные лекции

## Базы данных

Базы данных в Тинькофф и способы обращения к ним, диалекты SQL, инструменты аналитики в Тинькофф

01

## Базовые сведения из теории вероятностей и математической статистики

Базовые распределения, статистики и их свойства.

Библиотеки Python: numpy, scipy.stats, pandas.

02

03

## Визуализация данных

Конверсии, воронки и как их визуализировать. Библиотека matplotlib. Примеры на данных Тинькофф.

# A/B-тесты

## Оценка параметров

Статистические точечные оценки и их свойства. Как работать с выбросами.

Интервальные оценки, понятие доверительного интервала.

Построение доверительного интервала через бутстреп.

04

## Продуктовая аналитика

Принятие решения, зоны роста.

Исследование процесса и мониторинг.

Типы метрик и методы работы с ними.

Доли и динамики. Винтажи.

05

06

## A/B-тесты

Как сформулировать гипотезу по метрикам продукта. Как запустить эксперимент. Требования к группам, ожидаемые результаты. Какие встречаются проблемы.

07

## Основы проверки гипотез

Односторонние и двусторонние альтернативы. Статистическая значимость. Способы сравнения критериев.

# А/В-тесты и ответы на вопросы

## А/В-тесты

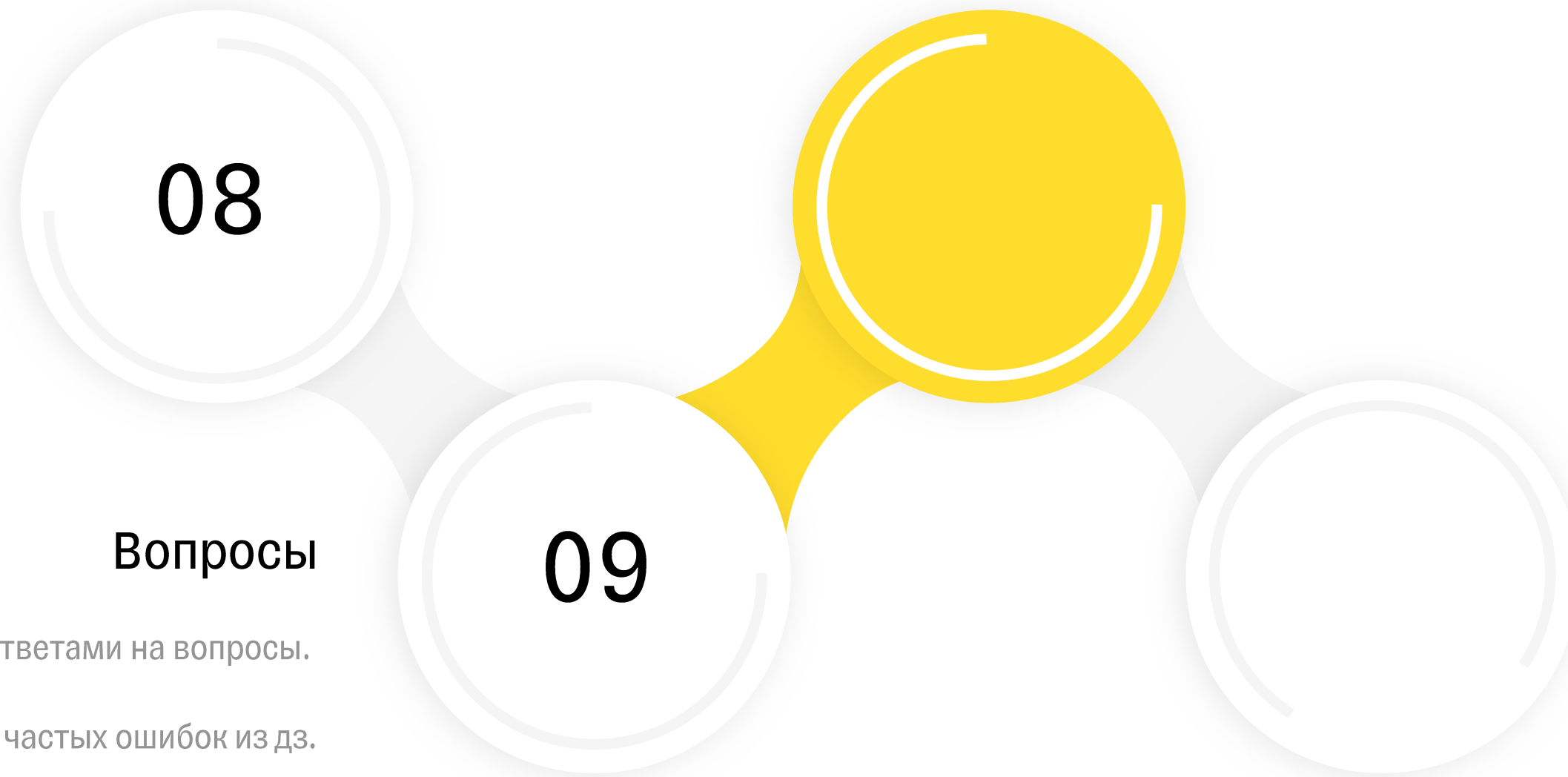
Выводы. Обработка  
результатов эксперимента.  
Как спасти поломанный тест.

08

## Вопросы

Семинар с ответами на вопросы.  
А так же разбор частых ошибок из дз.

09



# Финальные лекции

## Временные ряды

Где они встречаются в жизни аналитика?  
Выявление аномалий во временном  
ряде. Практика.

10

## Регрессии

Многомерная линейная регрессия. МНК. Примеры  
построения регрессионных моделей. Библиотека  
statmodels. Интерпретация коэффициентов модели.  
Логистическая регрессия.

11

12

## Случайный лес и деревья решений

Алгоритм, обучающий деревья решений. Подготовка  
данных. Методы оценки качества модели. Переобучение.

# Вопросы

Ответы на вопросы прошлого потока этого курса



**ТИНЬКОФФ**

[tinkoff.ru](https://tinkoff.ru)

# Достаточно ли знания курса для работы аналитиком?

Типов аналитиков много, и у каждого типа есть своя специфика работы.

Курс дает базовые знания с уклоном в работу продуктового аналитика.

Цель курса – получить на выходе людей, которые могут работать продуктовыми аналитиками.

При этом при старте работы уже понадобится вливаться в темы конкретного продукта: понять, что именно вы анализируете?



# Будет ли что-то выложено для самостоятельного обучения?

Видео с лекции и презентация будут выкладываться после лекции.

Также после лекции будем выкладывать условия задач домашнего задания и материалы для него.

Если понадобится подготовка (скачать что-то, установить, прочитайте) – сообщим об этом заранее в tg.

# Будет ли освещаться работа в Tableau?

Мы пользуемся этим инструментом.

Раз есть такой запрос, включим в курс.

# Что спрашивают на технических собеседованиях на позицию продуктового аналитика?

В случае джуниор-позиции – математика и понимание теории вероятностей.

Если позиция предполагает наличие опыта – разговор идёт про опыт.

# Какой диалект SQL используется в Тинькофф?

Базу с sql'ех -- используем. 😊

И PostgreSQL используем.

# Тесты будут в начале каждой лекции?

Да, чтобы лектор проверил, что вы помните содержание предыдущей лекции.

Можно будет сдать курс, не присутствуя на лекциях – тогда, очевидно, тесты вы не пишете.

В начале следующей пары теста не будет.

# Лектор один на весь курс?

Нет, у нас команда лекторов.

Читаем по несколько лекций.

# Что в Тинькофф используется для визуализации?

Tableau, Zeppelin, SAS, Excel

# Чем будем пользоваться на занятии по визуализации?

Расскажут про разные инструменты, их плюсы и минусы.



# Можно ли обойтись аналитику без знания VBA?

Для того, чтобы начать карьеру аналитика, эти знания необязательны.  
Если потом надо будет – разберётесь. 😊

В задачах визуализации будут  
использоваться алгоритмы машинного  
обучения для понижения размерности  
или будет просто построение  
графиков?

Скорее второе.

В целом будем больше говорить о том, как представлять данные наглядно и удобно  
для того, чтобы делать выводы.

# Расскажут ли на лекции про БД о Greenplum?

Да, немного расскажем.

# Практикуете ли нейронные сетки для решения практических задач?

В Тинькофф нейронные сети используются. Но не для любых задач, поскольку есть специфика интерпретируемости – просто не везде эти методы подойдут.

# Спасибо за внимание!

Увидимся на лекции. 😊



**ТИНЬКОФФ**

tinkoff.ru