

# Введение в статистику

Проект 12

**Золотая жила 3. Тест продукта**

# Кейс, проблема, идея



## Кейс

- Веб-платформа для изучения иностранного языка
- Клиенты потребляют разный контент и платят по разным тарифам
- Доходы от некоторых клиентов даже меньше, чем расходы платформы на обслуживание этих клиентов
- Компания хочет внедрить новые продукты для имеющихся клиентов, чтобы повысить среднюю прибыль с клиента
- Повышение средней прибыли с клиента на 800 рублей считается экономически значимым
- При тестировании гипотез требуется использовать уровень значимости 5%
- При определении размера выборки исходи из мощности 80%, попробуй и другие



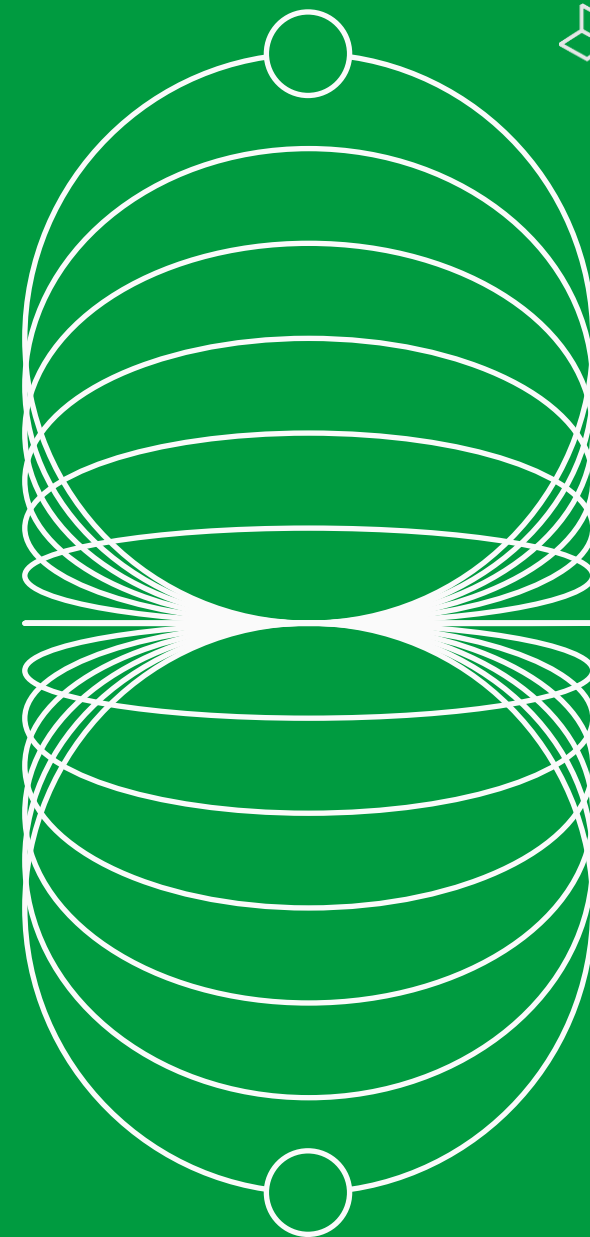
## Проблема

Нужно как-то тестировать новый продукт до запуска



## Идея решения

Описать процесс тестирования продукта. Процесс должен с требуемыми вероятностями обнаруживать экономически значимую пользу от продукта



# Роль, задача, результат



## Роль

Аналитик отдела рекламы и маркетинга



## Задача

Описать процесс тестирования продукта на слайдах



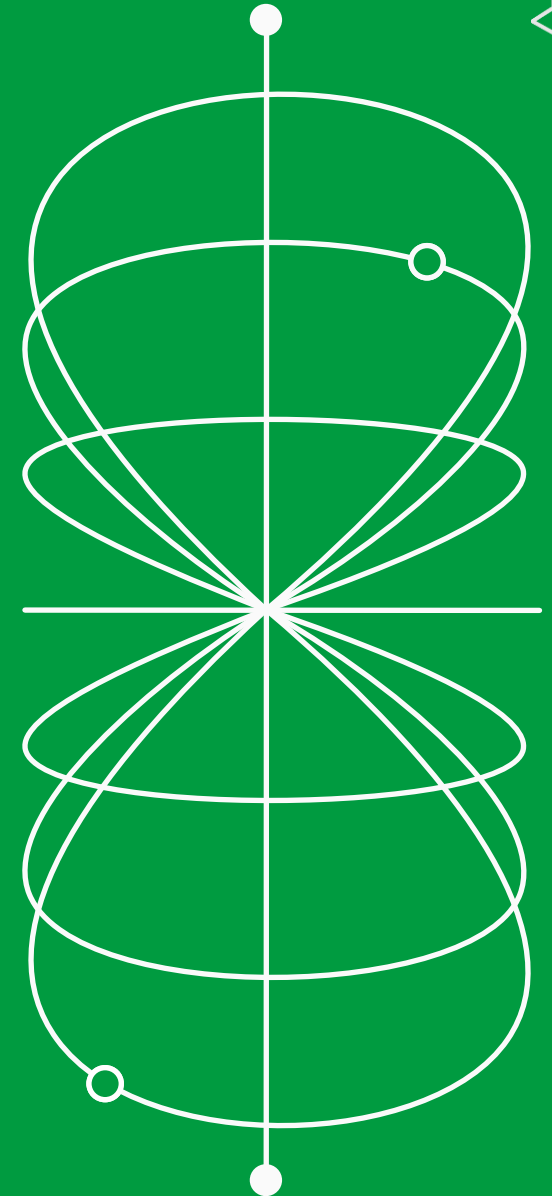
## Доступные данные

Датасет [golden\\_mine.csv](#): прибыль по текущим клиентам, пол, возраст



## Ожидаемый результат

- Презентация на слайдах в формате pdf
- Jupyter Notebook в Google Colab с расчётами



# Требования к слайдам



Если слайды или Jupyter Notebook не приложен, решение кейса **оценивается в 0 баллов**

## Понятность и внешний вид

- Внешний вид презентации не мешает воспринимать информацию
- Понятно, на какие вопросы отвечает каждый слайд
- Содержимое таблиц, графиков понятно из слайда без необходимости открывать исходный датасет
- Выводы явно сформулированы

**МАКСИМУМ 5 БАЛЛОВ**

## Обоснованность

- Выводы основаны на таблицах, графиках, показателях, полученных из данных
- Таблицы и графики получены скриншотом или картинкой из Jupyter Notebook, поэтому их можно перепроверить
- Выводы явно сформулированы

**МАКСИМУМ 5 БАЛЛОВ**

## Реакция заказчика

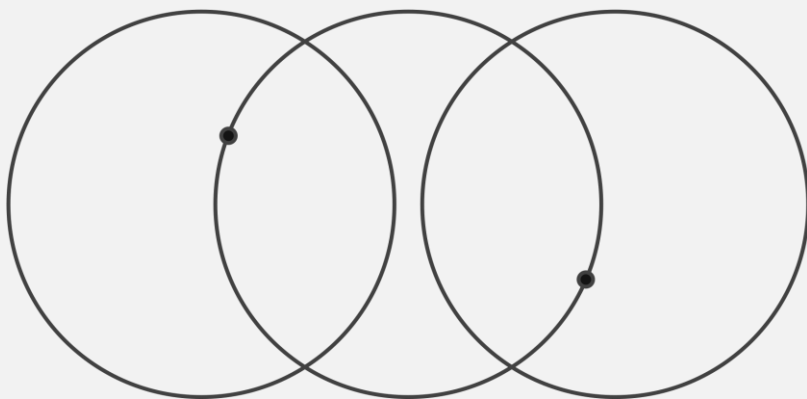
- **0 баллов**  
Не принимает, ищет другого исполнителя
- **1 балл**  
Частично принимает, считает необходимым отдать на доработку текущему исполнителю
- **2 балла**  
Принимает, готов пересылать слайды от своего имени, под свою ответственность

**МАКСИМУМ 2 БАЛЛА**

Максимальная оценка — **10 баллов**

Бонусные баллы могут поднять основную оценку, но не выше **10 баллов**, даже если в сумме получится 11 или 12 баллов

# Подход к решению



Для категории «Женщины 18-24 лет» описать процесс тестирования продукта. Считать, что средняя прибыль для тех, кто не пользуется новым продуктом, не изменится. Как изменится средняя прибыль среди воспользовавшихся продуктом — неизвестно

01

**Подготовить слайд с описанием шагов, например:**

**Шаг 1.** Определить тестовую группу — выборку.

**Шаг N-1.** Провести статистический тест полученных результатов

**Шаг N.** Внедрять продукт, если...

02

**Подготовить слайд о том, как выбирать тестовую группу:**

- Способ получения выборки
- Аргументы в пользу этого способа

03

**Подготовить слайд о размере выборки: на основе чего определяется минимальный размер выборки? Сам размер находить не нужно — это задание для красного уровня.**

04

**Подготовить слайд про то, что делать с выборкой**

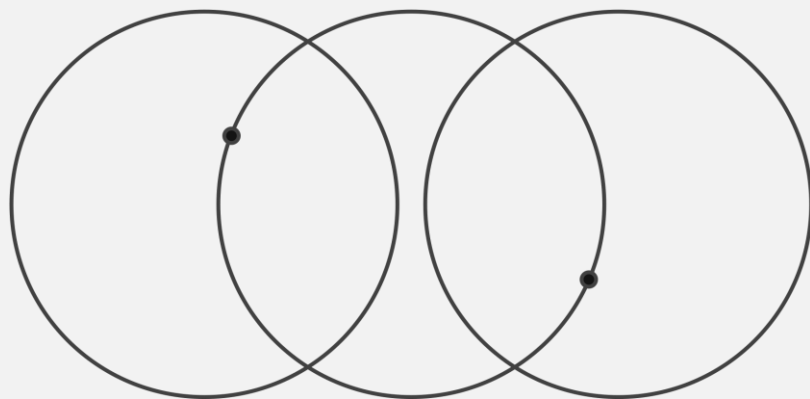
- Что делать с выбранными клиентами
- Какие данные собирать

05

**Подготовить слайд про статистический тест:**

- Как проводить статистический тест
- Как по итогу определять статистическую значимость

## Подход к решению



01

Сделать слайды синего уровня,  
с дополнительными требованиями:

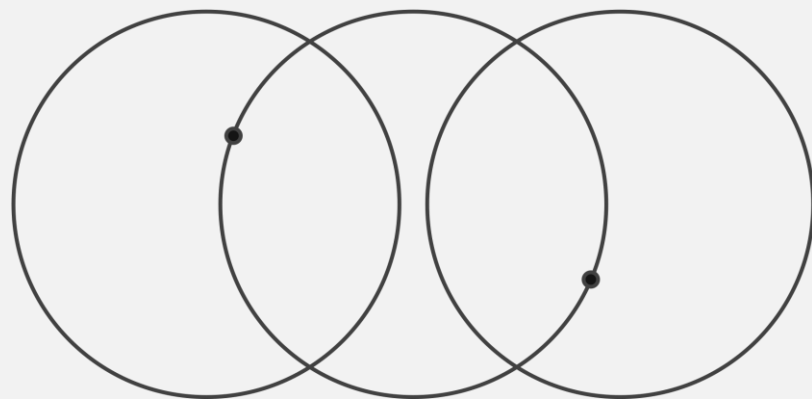
- Для слайда про размер выборки вычислить размер выборки
- Для слайда про статистический тест объяснить, что значит статистическая значимость

02

Добавить слайд про экономическую значимость  
и итоговое решение о внедрении продукта

Чёрный уровень

## Подход к решению



01

Сделать слайды красного уровня

02

Рассчитать размер выборки также для других категорий:

- «Мужчины 18-24 лет»
- «Золотая жила» — категория, которую нужно было найти в первом кейсе
- «Мужчины 18-24 лет», не попадающие в категорию «Золотая жила»
- Объяснить, почему суммарный размер в подкатегориях меньше, чем размер выборки, если выбирать из категории целиком

# Что и когда нужно сдать



## Что сдавать?

- Презентация в слайдах в формате pdf
- Jupyter Notebook в Google Colab с расчётами



## Когда сдавать?

Сроки сдачи указаны в информационной системе

