# Разведочный анализ данных (EDA) и Подготовка отчетов

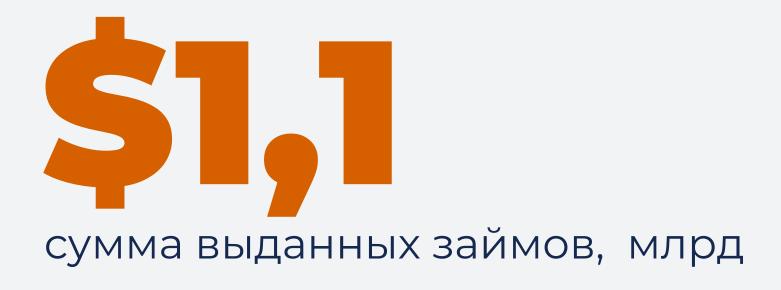
### Артемий Козырь

Старший аналитик данных @ Wheely

Telegram: @kzzzr

LinkedIn: Artemiy Kozyr

# О деятельности Кіуа



страны

миллионов заемщиков

миллионов займодателей

## Как это работает

1 Заемщик обращается за кредитом

Два типа моделей: прямой заем и заем через партнера

(3) Осуществляется сбор средств

Сообщество Kiva принимает участие в сборе средств (2) Заявка публикуется на Kiva

Заявка проходит процедуру одобрения

4 Заемщик погашает кредит

Кредиторы получают регулярные выплаты

## Вопросы для EDA

#### • Качество данных:

Проверить, насколько имеющиеся данные подвержены выбросам, пропущенным значениям.

#### • Базовые характеристики:

Выявить базовые характеристики датасета, обнаружить первоначальные наблюдения.

#### • Заемщики:

Изучить основные демографические характеристики людей, которые используют Kiva. Из каких они регионов и стран? Какова гендерная характеристика заемщиков?

#### • Назначение займов:

Посмотреть, для каких целей людям нужны средства и какие категории являются лидерами по количеству заявок.

## Вопросы для EDA

- Структура займов:
  - Посмотреть, какие суммы просят заявители как они распределены. Изучить сектора экономики, виды деятельности и сделать выводы о том, как обстоят дела в каждом.
- **Время финансирования займов:**Выяснить, как долго заемщикам приходится ждать полного финансирования заявки.
- Зависимость суммы и срока погашения: Изучить природу зависимости времени, требуемого на погашение ссуды и суммы кредита.
- Формулирование выводов:
  - Сформулировать основные выводы, рекомендации и визуальные материалы, их подкрепляющие.

## Выгрузки данных - kiva\_loans

#### Набор данных, который содержит большую часть информации о займах:

id - Уникальный идентификатор для кредита funded\_amount - Сумма, выплаченная Kiva агенту на местах (USD) loan\_amount - Сумма, выплаченная полевым агентом заемщику (USD) sector - Сектор использования займа activity - Более гранулированная категория **use** - Назначение использования суммы кредита country\_code - ISO код страны страны, в которой был выдан заем country - Полное название страны, в которой был выдан кредит posted\_time - Время, в которое заявка была опубликована на Kiva disbursed\_time - Время, когда кредитный агент выдал заемщику сумму funded\_time - Время, когда заявка полностью финансируется кредиторами term\_in\_months - Срок, на который был выдан кредит в месяцах lender\_count - Общее количество кредиторов, которые внесли свой вклад в этот кредит loaner\_genders - Разделенные запятыми буквы М, F, где каждый экземпляр представляет одного мужчину / женщину в группе

# Выгрузки данных - kiva\_mpi\_region\_locations

Набор данных, содержащий данные о регинах и показатели МРІ (глобальный индекс бедности)

country - страна

region - регион в стране

world\_region - часть света

**MPI** - многомерный индекс бедности

**geo** - координаты (широта, долгота)

## Формат отчета

#### Jupyter Notebook

Рабочая тетрадь с кодом, таблицами, визуализациями и комментариями

#### • Выгрузка в PDF

Формат для печати, публикации, распространения

## Домашнее задание

#### Самостоятельная работа:

- А в каких странах преобладают мужчины? Группа людей? (п. 2.2.2.)
- Изучите несколько займов с очень большими суммами. Что можно сказать о назначении этих заявок? В каких странах они оформлены? Каков гендерный состав заявителей? (п. 2.4.2.)
- Попробуйте найти связь между странами (регионами), полом заемщика и суммой заявки. Какие выводы можно сделать из этих данных? (п. 2.4.3.)

#### Домашняя работа:

Расширьте наш отчет исследованием следующих вопросов:

- Зависимость пола заемщика от региона планеты, сектора экономики.
- Влияние такого атрибута займа как Количество кредиторов (lender\_count) на суммы и сроки займов.
- Влияние показателя MPI (многомерный индекс бедности) в разных странах на суммы займов и сроки погашения.
- Зависимости таких метрик как Сумма займа, Срок займа, Время финансирования заявки, Ежемесячный платеж в разрезе Макрорегионов.

Skillbox

## Спасибо за внимание!

## Артемий Козырь

Старший аналитик данных @ Wheely

Telegram: <u>@kzzzr</u>

LinkedIn: Artemiy Kozyr