

# Чекпоинт II

Stardust Crusaders

21 ноября 2020 г.

## Содержание

<b>1</b>	<b>Цели чекпоинта</b>	<b>2</b>
<b>2</b>	<b>Результаты</b>	<b>2</b>

## Резюме

Опробованы два подхода к построению кластеризации. Проведен анализ исходного датасета. Сделана выборка с ручной разметкой.

Результаты анализа:

Полученные кластеры (модель 1)

Полученные кластеры (модель 2)

## Цели чекпоинта

- Развернуть окружение для работы;
- Подготовить и очистить данные;
- Первичный анализ данных;

## Результаты

Исходный `.xlsx` файл был преобразован в `.csv` с пятью столбцами:

- Дочернее общество;
- Подрядчик;
- Направление деятельности;
- Производственный объект;
- Описание инцидента.

Были выделены три сущности для анализа:

- Инциденты  
(1)
- Подрядчики  
(1)
- Дочерние общества  
(1)