**Наши задачи:**

1. Посмотрим исходные данные
2. Что нам нужно использовать, чтобы их обработать
3. Подумать над MVP
4. Дальнейшие действия

**Вебинар по задаче**

Будет запись.

**Константин Гончар**

**Георгий** –

Скинули папку с урезанными датасетами. Около 1000 строк.

Завтра будут полные датасеты до 100.000.000 строк.

Как нам обработать такой объем? Нужно ли загрузить в БД?

Где пригодятся нейронные сети?

\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*-\*

Что такое ТРУ сущность?

Товары Работы Услуг.

Товаров больше.

Функционал ---

=============================================================

Данные мои заметки

Данные в ячейках не самые чистые.

country\_code – страна происхождения. У одного и того же товара может отличаться.

Данные в разных файлах имеют разный приоритет.

Заказчику интересно

Кто поставляет конкретный товар. Какая цена Какая средняя.

Как разобрать запрос на составляющие.

Поставщику интересно

Какие цены на товары

Насколько мои цены актуальны

Общие уровни цен

**Каким должен быть MVP?**

Что-то запрашивают у программы и получают какой-то информационный срез.

Информационный срез должен быть полезным для решения бизнес-задачи.

* Типа: У кого заказать, чтобы доставили и по хорошей цене?

Вывод информации должен выводиться в понятном интерфейсе.

Линейные решения – могут быть эффективнее чем нейросети. Что это значит.

Что такое лематизация?

В жюри будет топменеджмент и технические специалисты.