# Описание датасетов аукционов

## 1. Файл: общие\_сведения\_по\_аукционам\_полные\_102332.xlsx

Основная информация о проведенных аукционах.

| **Название колонки** | **Описание** |
| --- | --- |
| Unnamed: 0 | Технический индекс (можно игнорировать) |
| idBid | Уникальный идентификатор аукциона (основной ключ) |
| LotPrice | Цена лота (может быть ценой за единицу при групповом лоте) |
| StartingPrice | Начальная цена аукциона/лота |
| PriceDynamicsString | Динамика изменения цены в текстовом формате |
| BidCount | Количество поданных заявок/ставок |
| AuctionDurationSec | Длительность аукциона в секундах |
| WithStep | Флаг наличия шага аукциона (Да/Нет) |
| StepType | Тип шага аукциона |
| IsRaise | Наличие повышающих ставок (True/False) |
| CanImproveBid | Возможность улучшения своей ставки |
| WithSecondPhase | Наличие второй фазы аукциона |
| TradeType | Тип торгов |
| MaxStep | Максимальный шаг ставки |
| MinStep | Минимальный шаг ставки |
| LotName | Наименование лота |
| Play | JSON-строка с поданными ставками |
| ParticipantsInfos | JSON-строка с информацией об участниках |
| AuctionStartDate | Дата и время начала аукциона |

## 2. Файл: 89к извещений и обеспечений.xlsx

Информация об извещениях и обеспечениях заявок.

| **Название колонки** | **Описание** |
| --- | --- |
| X1 | Технический индекс |
| idBid | Уникальный идентификатор аукциона |
| tradeid | Идентификатор типа аукциона |
| contract\_time | Срок исполнения контракта |
| notification | Номер извещения |
| msp | Признак преференций для малых субъектов предпринимательства |
| organizator | Организатор аукциона |
| obesp\_zayavki | Сумма обеспечения заявки |

## 3. Файл: 89к аукционов с адресами.xlsx

Полные данные с адресной информацией.

Содержит все колонки из файлов 1 и 2, плюс:

| **Название колонки** | **Описание** |
| --- | --- |
| Unnamed: 0.2 | Технический индекс (можно игнорировать) |
| Unnamed: 0 | Технический индекс (можно игнорировать) |
| Unnamed: 0.1 | Технический индекс (можно игнорировать) |
| Unnamed: 0.1.1 | Технический индекс (можно игнорировать) |
| sleep | Количество участников, не подавших ни одной ставки (при наличии регистрации) |
| number | Номер записи |
| place | Полный адрес проведения аукциона |

## Особенности данных:

1. **JSON-поля**:
   * Play - содержит историю ставок в формате JSON
   * ParticipantsInfos - содержит информацию об участниках в формате JSON
2. **Ключ для связи**:
   * Основной ключ для всех файлов - idBid
3. **Технические колонки**:
   * Все колонки Unnamed: \* являются артефактами экспорта и могут быть проигнорированы
4. **Специфичные поля**:
   * msp - указывает на применение специальных правил для малого бизнеса
   * sleep - показатель "неактивных" участников ( заплатили деньги за право участвовать, но не пришли на аукцион)

## **Цель исследования**

Определить ключевые факторы, влияющие на уровень конкуренции (снижение цены в ходе аукциона), и построить модель, предсказывающую динамику цен (*PriceDynamicsString*) на основе параметров аукциона.

## **Задачи исследования**

### **1. Визуальный и статистический анализ**

* **Описательная статистика**:
  + Распределение PriceDynamicsString (среднее снижение цены, медиана, выбросы).
  + Зависимость снижения цены от:
    - Количества участников (BidCount).
    - Типа шага аукциона (StepType).
    - Наличия второй фазы (WithSecondPhase) и так далее
* **Графики**:
  + Boxplot снижения цены по категориям.
  + Тепловая карта корреляции числовых признаков.
  + Карта Регионов России с цветовой кодировкой среднего снижения цены на торгах

### **2. Построение моделей (Опционально)**

* **Методы**:
  + **Линейная регрессия**
  + **Gradient Boosting (интерпретация модели через shap или permutation importance)**

### **3. Интерпретация результатов**

* Какие факторы увеличивают снижение цены на аукционе?
* Возможно ли предоставить практические рекомендации для организаторов?

P.S. В файле 89к аукционов с адресами.xlsx не 89000 аукционов, а меньше. При наличии возможности посмотреть, есть ли аукционы без адреса, где у того же организатора в другом аукционе юридический адрес указан.