Представленные данные отображают историю проведения закупок с начала 2021 года по конец 2022 с заказчиками и поставщиками (участниками).

Данные состоят из двух датасетов и одного справочника (ОКПД2), делитель в csv – “;”, кодировка в csv – “utf-8”:

* purchases

small\_purchases.csv ~ 350 000 строк (230 мб)

all\_purchases.csv ~ 620 000 строк (430 мб)

* suppliers

small\_suppliers.csv ~ 960 000 строк (46 мб)

all\_suppliers.csv ~ 1 900 000 строк (91 мб)

* okpd2.xlsx ~ 20 000 строк (0.6 мб)

Поля в датасете purchases:

* purchase – Уникальный идентификатор закупки
* region\_code – Регион поставки
* min\_publish\_date – Дата публикации закупки (дата, с которой закупку смогли видеть поставщики и присылать свои предложения)
* purchase\_name – Название закупки
* forsmallbiz – Если да, то большие компании не могут предлагать свое участие в этой закупке
* price – Цена за закупку, предлагаемая заказчиком
* customer – Уникальный идентификатор заказчика
* okpd2\_code – Код стандартизации товара в соответствии со справочником ОКПД2, обрезанный до 3 числа
* okpd2\_names – Названия ОКПД2, в этом поле может быть записано несколько названий ОКПД2 через разделитель “ || “, так как поле okpd2\_code это обобщенный код.
* Item\_descriptions – Описания товаров, записанные через разделитель “ || “.

Поля в датасете suppliers:

* purchase – Уникальный идентификатор закупки
* supplier – Уникальный идентификатор поставщика
* is\_winner – Выиграл ли этот поставщик в этой закупке