**Дополнительные задания на отработку**

Я выбрала тему для данных: объекты культурного наследия

**Найденные мной датасеты на эту тему:**

1. [Объекты](https://data.mos.ru/opendata/63262?isDynamic=false) культурного наследия регионального значения и выявленные объекты культурного наследия
2. [Объекты](https://data.mos.ru/opendata/62302?isDynamic=false) культурного наследия федерального значения
3. [Сведения](https://opendata.mkrf.ru/opendata/7705851331-egrkn) из Единого государственного реестра объектов культурного наследия (памятников истории и культуры) народов Российской Федерации

Сформулируйте **тему и вопросы**, которые можно было бы изучить с помощью предложенных вами наборов данных.

Тема для исследования данных: Количественный анализ объектов культурного наследия на территории Российской Федерации

Вопросы, которые можно исследовать:

* В каком регионе больше всего объектов культурного наследия?
* В каком регионе меньшее количество объектов культурного наследия?
* Какого типа объектов больше по различным регионам и районам?
* Как много наследий ЮНЕСКО находится в Российской Федерации и где их больше всего?

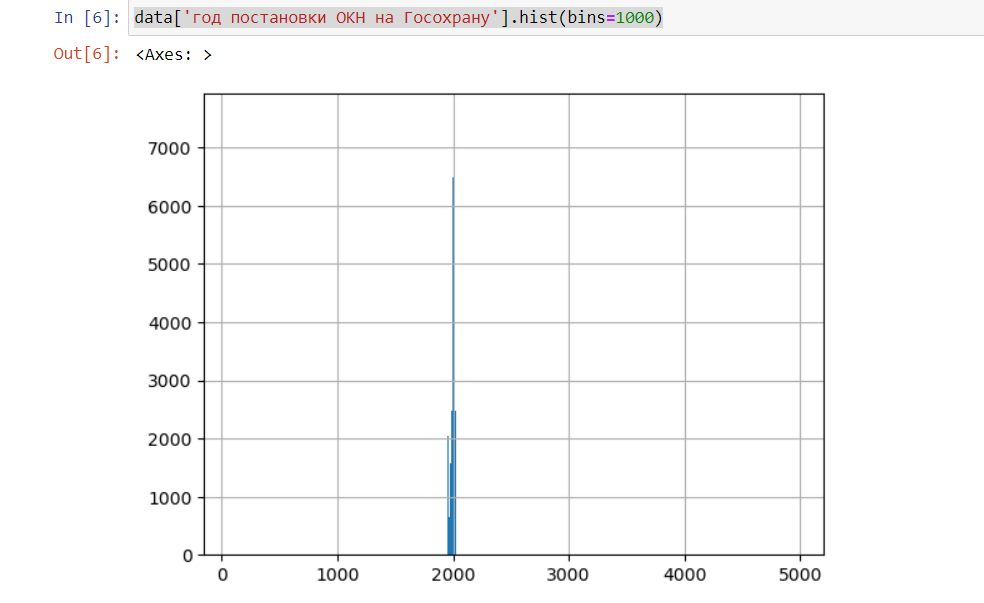
1. **Сделайте один обогащенный набор данных из нескольких наборов данных** из одного или разных источников. Можно использовать VLOOKUP (ВПР) или XLOOKUP, Python (pandas методы: merge, concat и др.). Важно пересечение по одной переменной (ключевое слово, наименование, идентификатор).

<https://disk.yandex.ru/i/fxZCNVpAPcNHAg>

Я создала общий файл с данными, но так как столбцы совпадают, различаются только данные я считаю, что использовать VLOOKUP (ВПР) не имело смысла

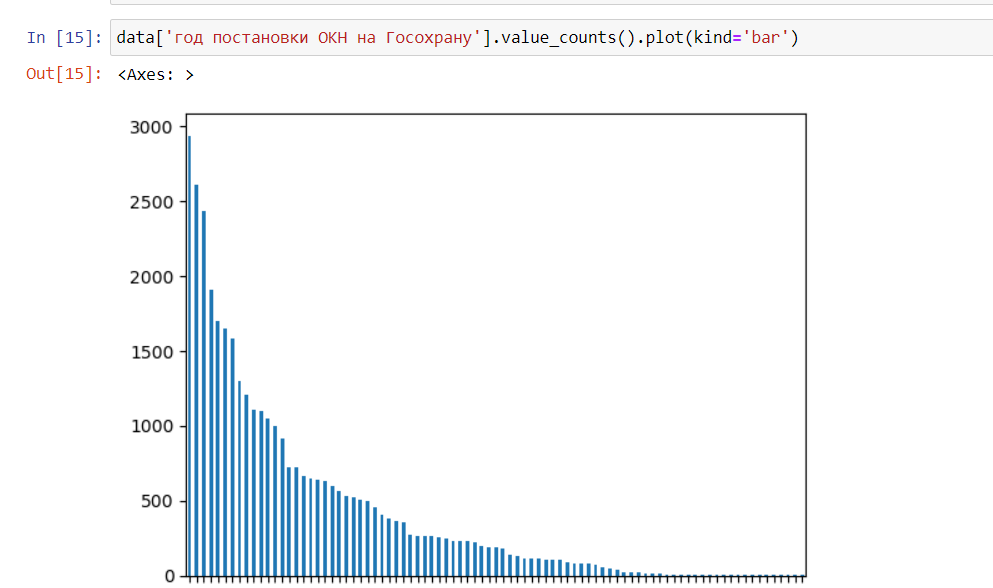
1. **Explain it.** Продемонстрируйте на основе какого-либо набора данных, чем отличаются между собой гистограмма и столбчатая диаграмма (bar chart, column chart).

Я создала гистограмму на основе данных из Единого государственного реестра объектов культурного наследия (памятников истории и культуры) народов Российской Федерации.



Гистограмма и столбчатая диаграмма оба представляют данные в виде столбцов, но они используются для различных целей и отличаются в способе представления данных.

Гистограмма используется для визуализации распределения числовых данных. Ось X гистограммы представляет значения переменной, а ось Y представляет частоту или плотность этого значения. Столбцы в гистограмме обычно имеют одинаковую ширину и не имеют промежутков между ними. Гистограмма полезна для понимания формы распределения данных, выявления выбросов и определения пиков или мод.



Столбчатая диаграмма используется для сравнения категориальных данных или значений, которые не связаны с непрерывным диапазоном. Ось X столбчатой диаграммы представляет категории или группы данных, а ось Y представляет значения, связанные с этими категориями. Столбцы в столбчатой диаграмме могут иметь разные ширины и обычно разделены промежутками. Столбчатая диаграмма помогает сравнивать относительные размеры или количества различных категорий.

Таким образом, гистограмма применяется для визуализации распределения числовых данных, а столбчатая диаграмма - для сравнения категорий или групп данных.