|  |
| --- |
| МИНОБРНАУКИ РОССИИ |
| Федеральное государственное бюджетное образовательное учреждение высшего образования **«МИРЭА Российский технологический университет»**  **РТУ МИРЭА** |

**Институт информационных технологий (ИИТ)**

**Кафедра практической и прикладной информатики (ППИ)**

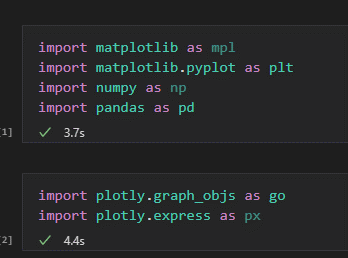
**ОТЧЕТ ПО ПРАКТИЧЕСКОЙ РАБОТЕ**

по дисциплине «Технологии и инструментарий анализа больших данных»

**Практическое занятие 2**

|  |  |  |  |
| --- | --- | --- | --- |
| Студент группы ИНБО-01-17 | *ИНБО-03-20, Першутов Н.С.* | (подпись) | |
| Преподаватель | *Парамонов А.А.* | (подпись) | |
| Отчет представлен | «\_\_\_»\_\_\_\_\_\_\_\_2023г. | |  | |

Москва 2023г.

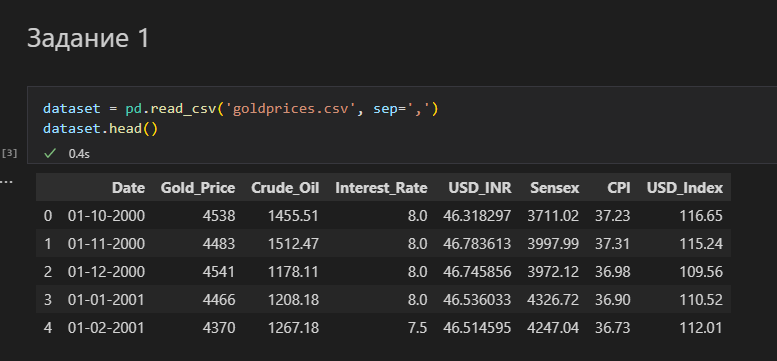


Импорт необходимых библиотек.

**Задание 1**

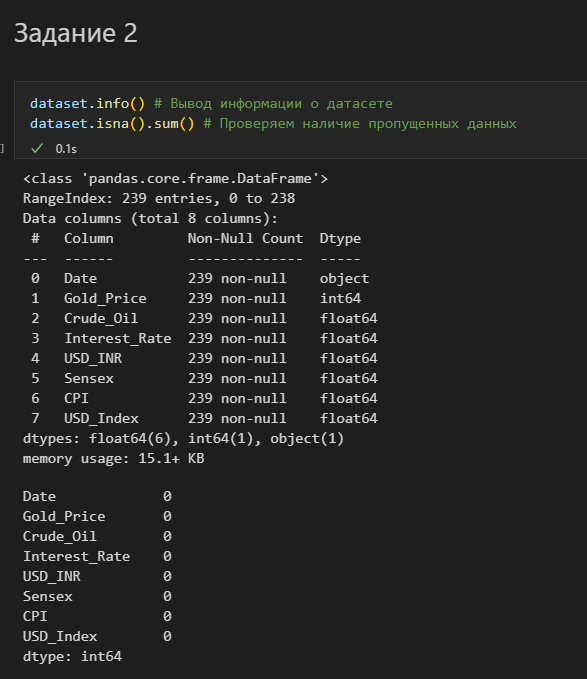
Нашли и выгрузили многомерные данные с использованием библиотеки pandas.

Вывели информацию при помощи метода .head()



**Задание 2**

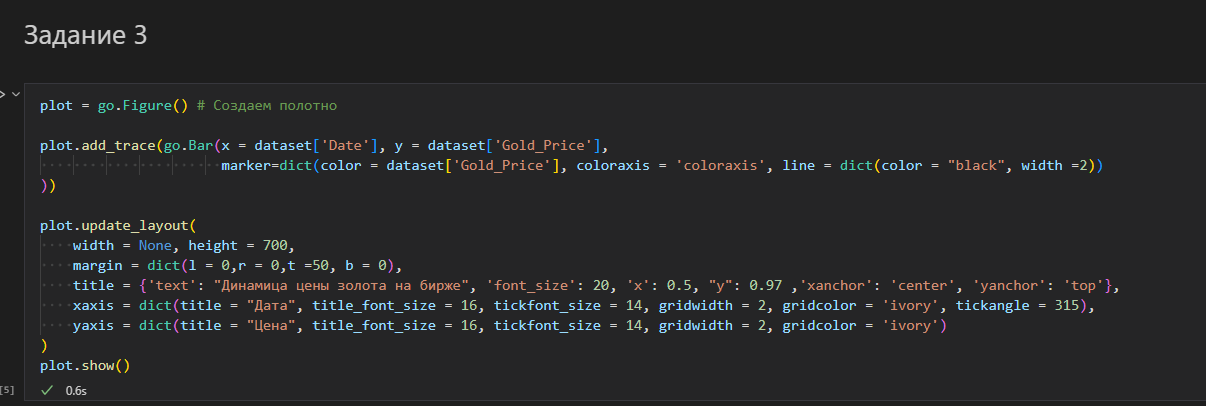
Вывели информацию о данных при помощи метода .info(). Проверили данные на наличие пустых значений.

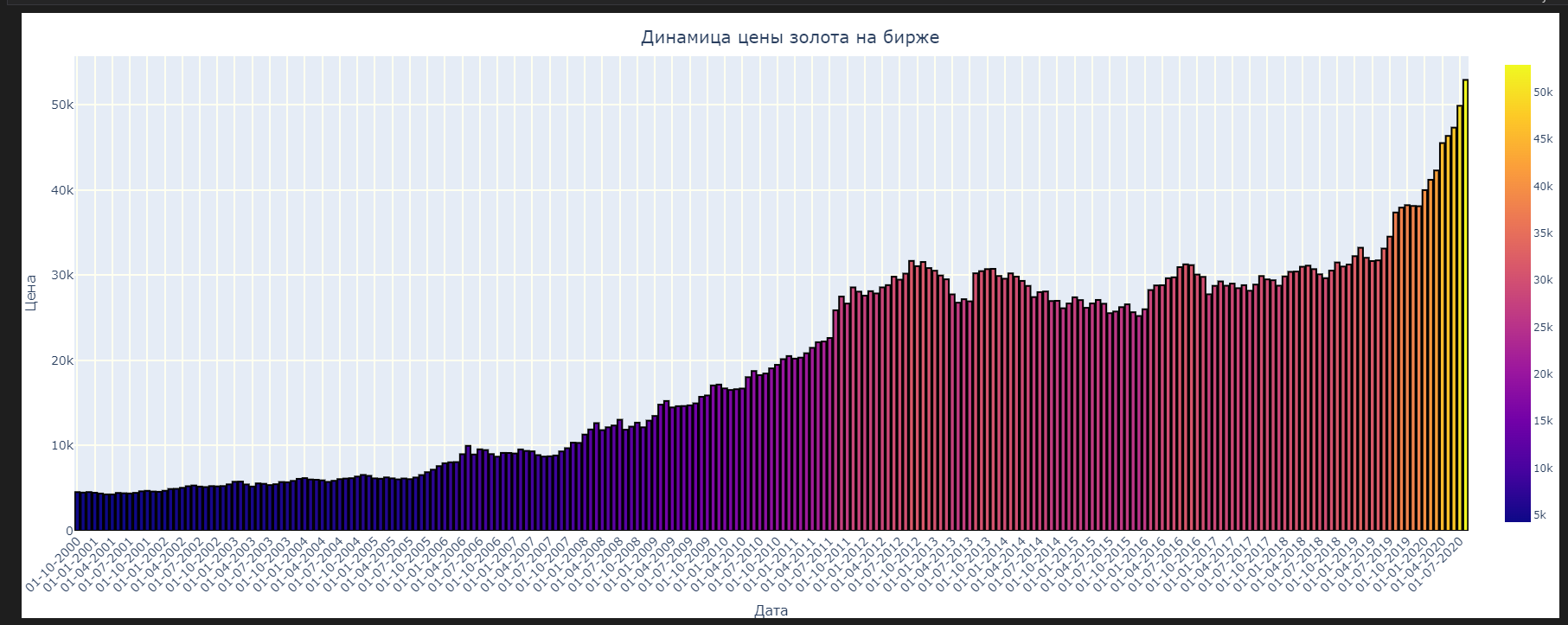


**Задание 3**

Построили столбчатую диаграмму (.bar) с использованием модуля graph\_objs из библиотеки Plotly со следующими параметрами:

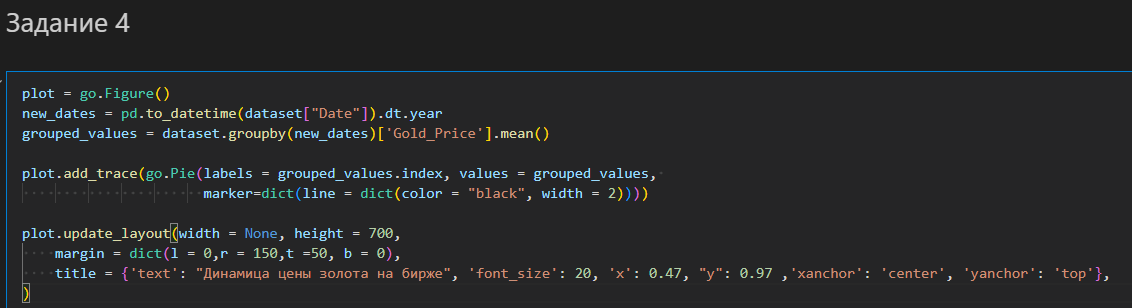
* По оси Х указали дату, по оси У указали количественный показатель.
* Сделали так, чтобы столбец принимал цвет в зависимости от значения показателя
* Сделали так, чтобы границы каждого столбца были выделены чёрной линией с толщиной равной 2.
* Отобразили заголовок диаграммы, разместив его по центру сверху, с 20 размером текста.
* Добавили подписи для осей X и Y с размером текста, равным 16. Для оси абсцисс развернули метки так, чтобы они читались под углом, равным 315.
* Размер текста меток осей сделали равным 14.
* Расположили график во всю ширину рабочей области и присвоили высоту, равную 700 пикселей.
* Добавили сетку на график, сделав её цвет 'ivory' и толщину равную 2.
* Убрали лишние отступы по краям

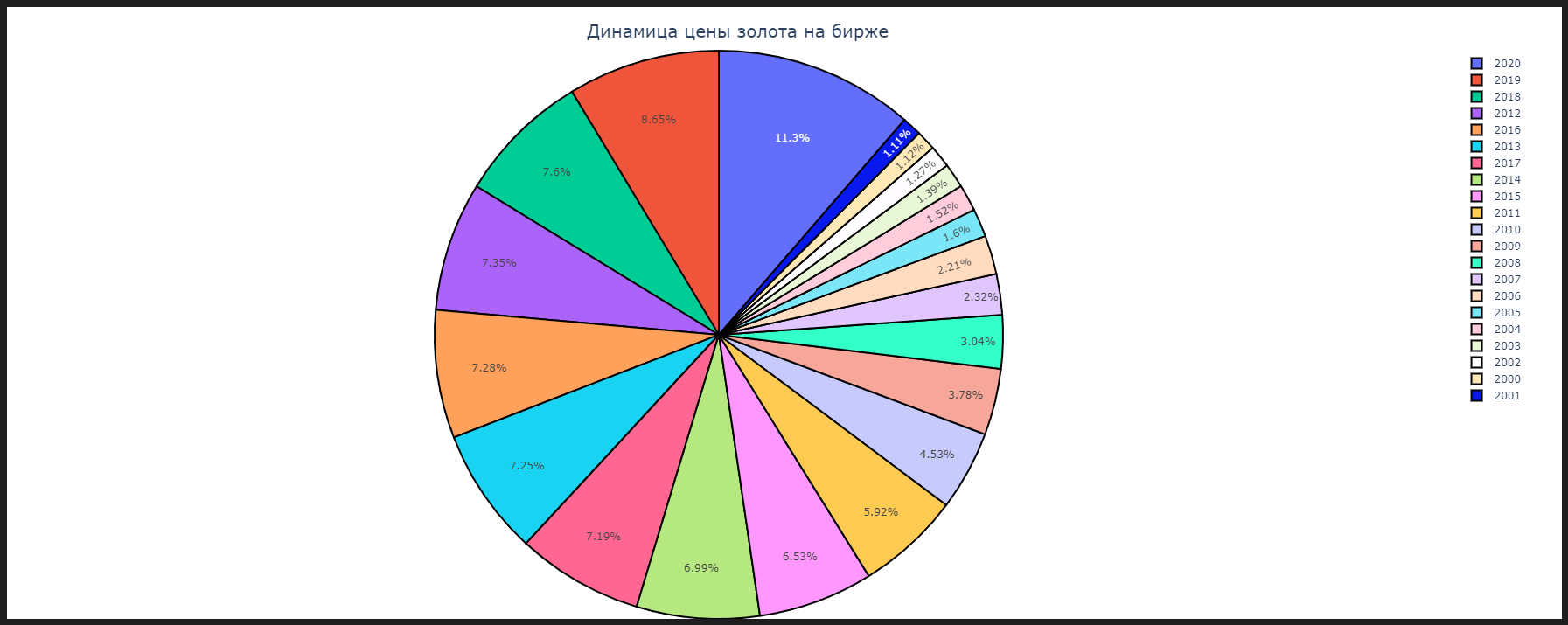




**Задание 4**

Построили круговую диаграмму (go.Pie), использовав данные и стиль оформления из предыдущего графика. Сделали так, чтобы границы каждой доли были выделены чёрной линией с толщиной, равной 2 и категории круговой диаграммы были читаемы посредством группировки значений по годам и расчетом среднего значения за этот год.



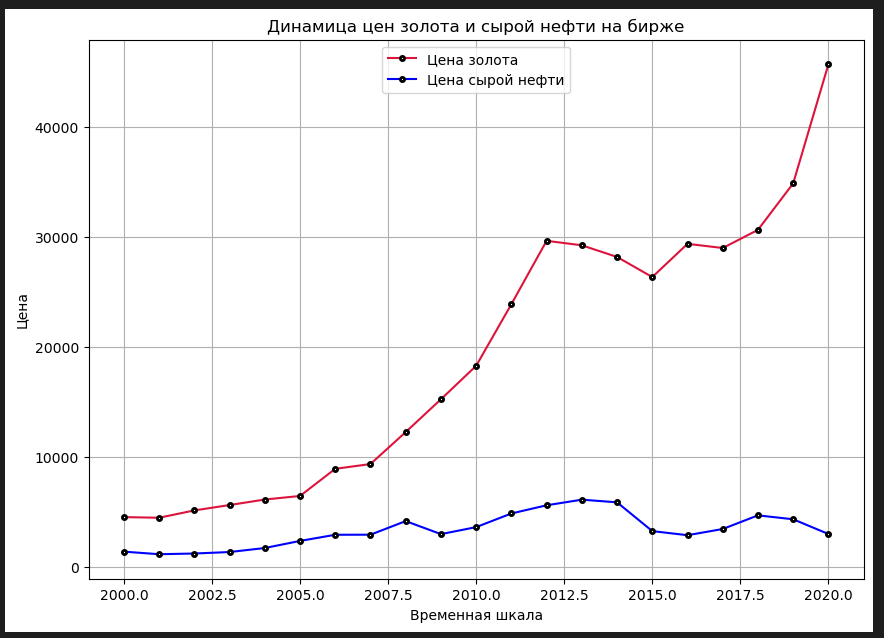


**Задание 5**

Построили линейные графики, взяв для сравнения цены Золота (Gold\_Price) и Сырой нефти (Crude\_Oil). А так же настроили графики следующим образом:

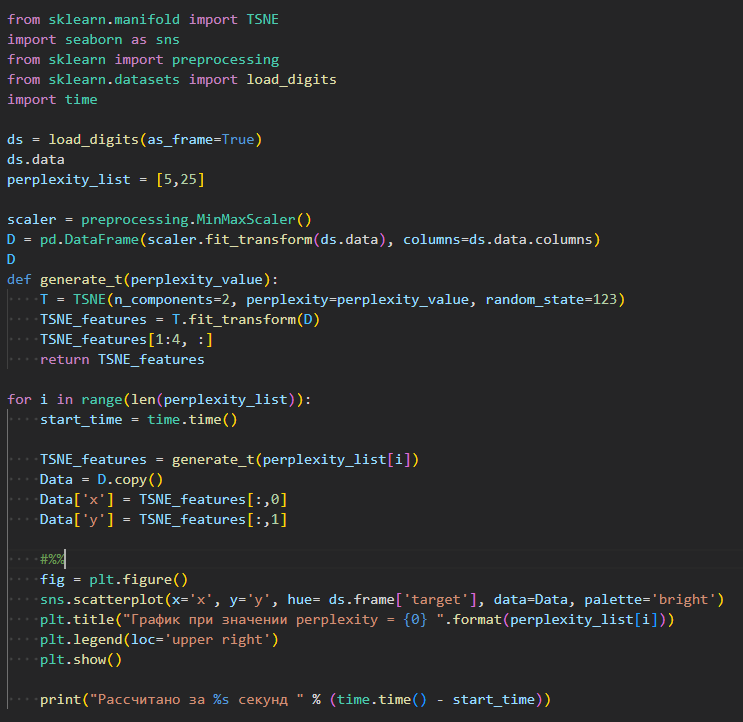
* Цвет линии 'crimson'
* Цвет точек 'white'
* Цвет границ точек 'black'
* Толщина границ точек равна 2.
* Добавили сетку на график сделать её цвет 'mistyrose' и толщину равную 2.

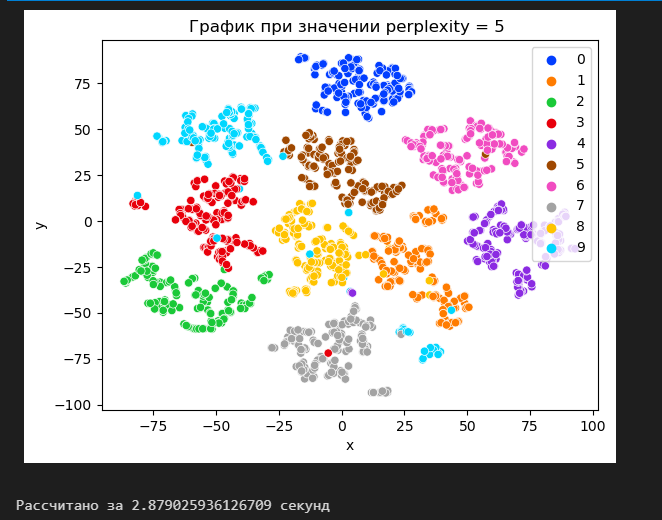


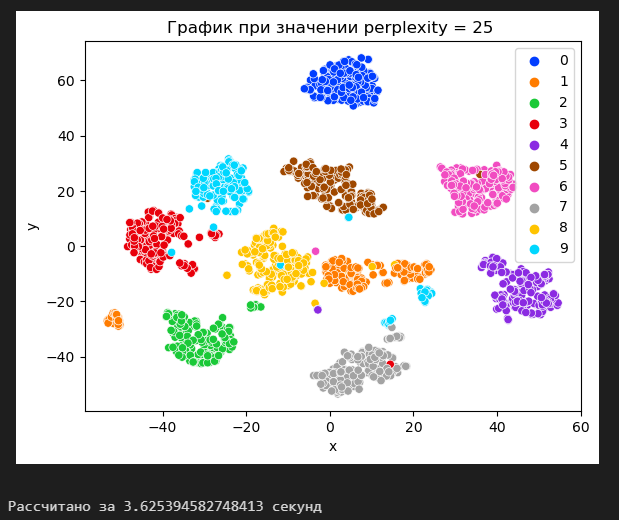


**Задание 6**

Выполнили визуализацию многомерных данных, используя t-SNE. Использовали датасет load\_digits из библиотеки sklearn. Для сравнения было выбрано два графика с значением perplexity 5 и 25 и выведено время расчета значений.

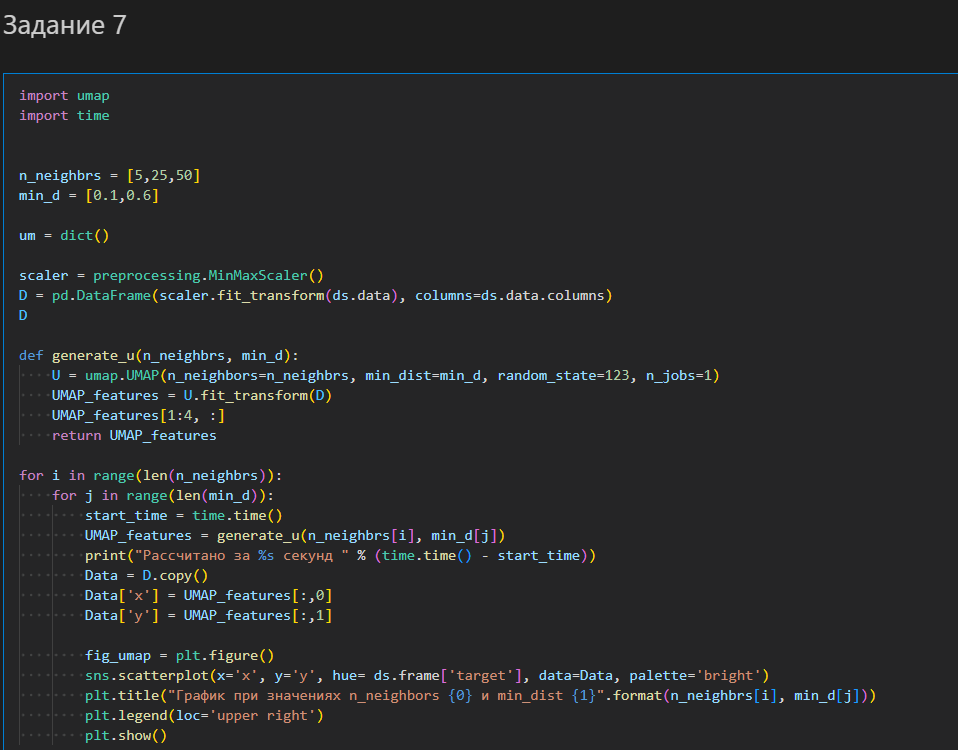


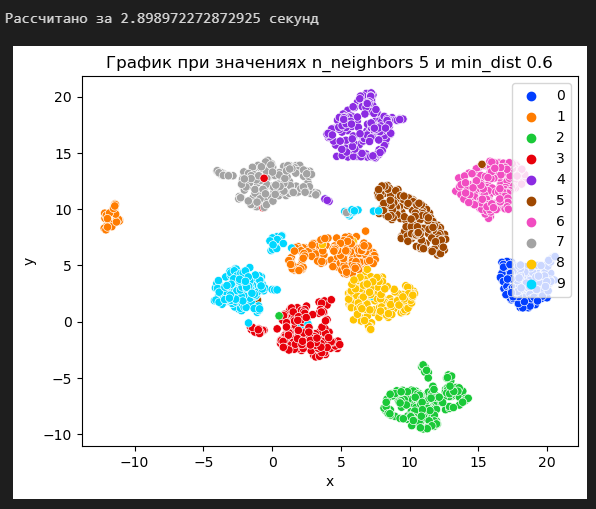
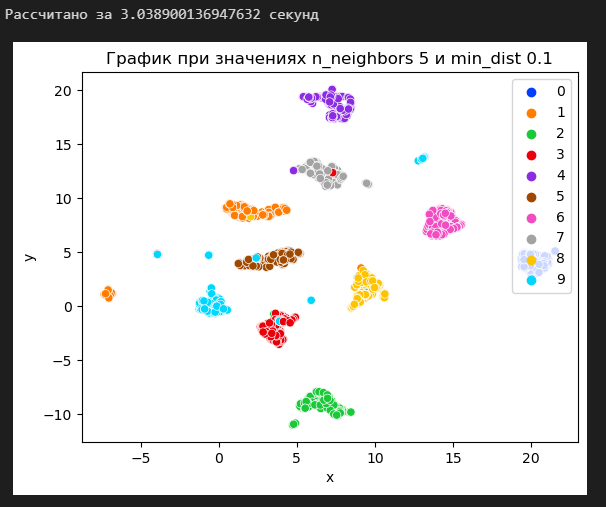


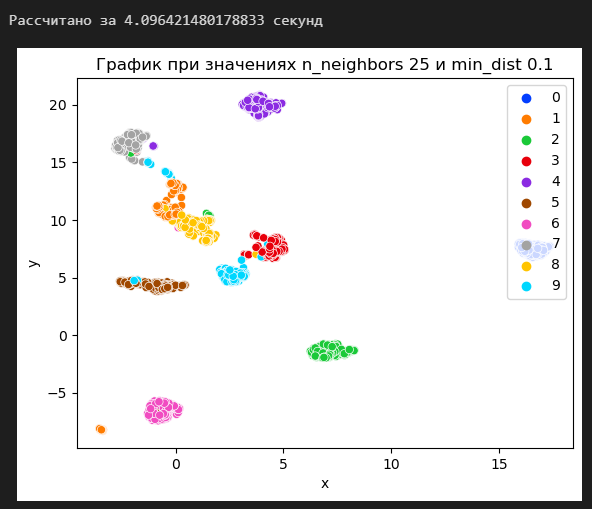


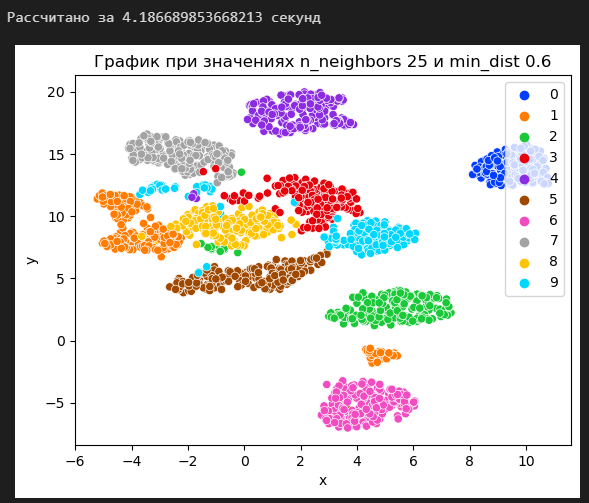
**Задание 7**

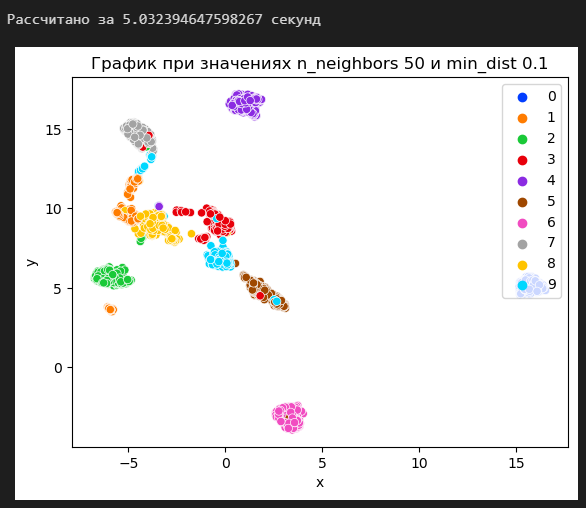
Выполнили визуализацию многомерных данных, используя UMAP с параметрами n\_neighbors = 5, 25, 50 и min\_dist = 0.1, 0.6. Рассчитали время работы алгоритма с помощью библиотеки time.

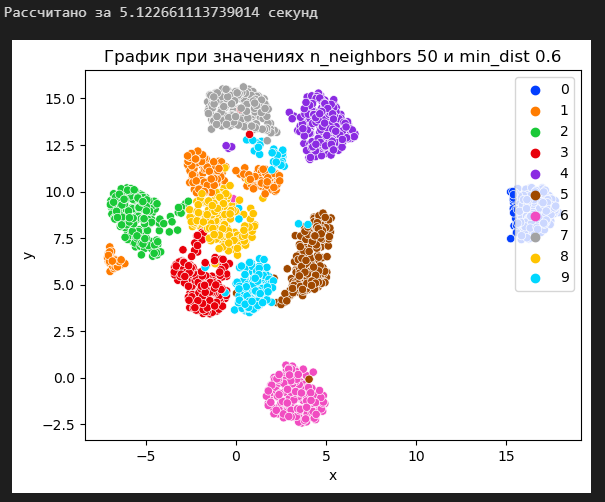












**Выводы:**

В данной практике мы сравнили такие библиотеки как matplotlib и plotly, а также TSNE и UMAP. По моему мнению библиотека matplotlib является более настраиваемой и понятной, но уступает plotly в функциональных возможностях при визуализации.