Задание

Найти способ бинаризации данных, используя pandas.

1. Числовые данные
2. Категориальные (номинальные) данные (в столбце встречается конечное – не очень большое – число различных данных). Обычно это строки.
3. Логические или булевские данные (истина и ложь).

Бинарные данные – в столбце два разных значения. Обычно мы будем их обозначать 0 и 1.

1. Категориальные данные можно перевести в бинарные. Один столбец заменяется на несколько столбцов по числу категорий.
2. Числовые данные. Разбиваем значения в столбце на не очень большое число интервалов. 3 интервала (малые, средние, большие). Числовые данные заменяются на категориальные (по попаданию в интервал). Категориальные данные заменяются на бинарные. Эти две операции можно выполнять последовательно и считать одной операцией.

Взять данные из репозитория UCI, а именно iris. iris.data, iris.names. Написать приложение Python, которое:

1. импортирует pandas,
2. читает файл csv в pandas.DataFrame
3. дает колонкам имена из iris.names,
4. делит данные в колонках на 3 части, (малое, среднее и большое),
5. заменяет числовые данные в колонках на номинальные,
6. заменяет номинальные данные на бинарные.