Контрольное задание

**Модуль 2.** **Разведочный анализ данных**

Подробное описание задания:

1. Выполнить проверку и обработку пропусков в данных.
2. Аргументировать выбранный метод обработки пропущенных данных.
3. Произвести оценку дисперсии среднего медианы между наборами данных train и test.
4. Графически отобразить разницу или её отсутствие между датасетами.
5. Найти и устранить выбросы.
6. Построить как минимум одну корреляционную матрицу.
7. Сделать вывод на основе корреляционной матрицы.
8. Произвести анализ фич с использованием любого алгоритма с учителем с любой из целевых переменных.
9. Сделать вывод на основе метода с учителем о связи переменных с любой из целевых переменных.
10. Произвести анализ фич с использованием любого алгоритма без учителя.
11. Сделать вывод на основе метода без учителем о связи переменных.

Выходные данные по модулю:

1. Отчет в формате HTML/PDF.
2. Файл .ipynb, полностью исполняемый и без ошибок.
3. Один или несколько итоговый файлов, представляющих оптимизированные и объединенные (при необходимости) данные.

Формат именования файлов:

M1\_{lastname}.{fileformat}

(например, M1\_Ivanov.html, M1\_Ivanov.ipynb)