**СЕССИЯ 1**

**Исходные файлы:**

1. Data.zip (Набор данных)
2. Сессия 1.docx (Инструкция к 1 сессии)

**Результаты работы:**

1. Data.zip (Предобработанные данные)
2. C[X]\_M1 (Папка с отчетом о проделанной работе)
3. Readme.txt (Дополнительные комментарии)

**ЗАДАНИЕ**

|  |
| --- |
| 1.1 ЗАГРУЗКА ДАННЫХ |
| Загрузите данные. Код загрузки отразите в отчете. |

|  |
| --- |
| 1.2 ПРЕДВАРИТЕЛЬНОЕ ИЗУЧЕНИЕ |
| Представьте распределение целевой переменной (таблица/график) и отразите в отчете.  Представьте распределения значащих переменных (таблица/график) и отразите в отчете |

|  |
| --- |
| 1.3 ЧИСТКА/ ОБРАБОТКА НЕЗНАЧАЩИХ СИМВОЛОВ В ДАННЫХ |
| Выявите и обработайте незначащие символы в текстовых данных, отразите в отчете.  Выявите и обработайте данные, не соответствующие формату столбца, отразите в отчете. |

|  |
| --- |
| 1.4 ОБРАБОТКА ПУСТЫХ ЗНАЧЕНИЙ |
| Выявите и обработайте пустые значения, отразите в отчете. |

|  |
| --- |
| 1.5 ПОДГОТОВКА ОТЧЕТА |
| Предоставьте отчет о проделанной работе. Внимание: оценка сессии будет проводиться на основании отчета. Отчет предлагается писать в Jupyter Notebook или аналогичной среде, где участник может последовательно представить, как описание проделанной работы, так и часть программного кода и результат работы программы.  Отчет должен быть предоставлен в папке C[X]\_M1, где [Х] – номер участника, который должен быть размещен на рабочем столе ноутбука участника.  Папка должна содержать все результаты выполнения модуля, а также все необходимые файлы для запуска и проверки участков кода. В папке 2 файла – Jupyter Notebook C[X]\_M1.ipynb (или аналог – с возможностью запустить и исполнить участки кода), C[X]\_M1.HTML( или PDF), где [Х] – номер участника, экспортированная тетрадь (или аналог) в статический формат. |