Отчёт о проделанной работе

1. Модуль A

В ходе решения модуля A удалось выполнить следующие задания:

* Загрузить данные из tsv файлов в pandas DataFrames
* Выполнить объединение данных
* Выполнить кластеризацию данных
* Создать функции для выявления дополнительных атрибутов (загруженность относительно максимальной, продолжительность поездки, топ загруженных станций)
* Визуализировать распределения числовых данных и провести тест Шапиро для каждого числового атрибута
* Выявить типы распределения
* Выявить признаки с пропущенными значениями

1. Модуль Б

В ходе решения модуля Б удалось выполнить следующие задания:

* Написать SQL-запросы для загрузки данных в базу данных, получения загруженности относительно максимальной, получения топа самых загруженных станций
* Разработать интерактивный дашборд в Power BI