1. Целью проекта было провести разведывательный анализ данных для определения набора величин, оказывающих влияние на результаты экзамена по математике.
2. Датасет состоит из 30 столбцов и 395 строк с различными количественными и номинативными величинами.
3. Этапы работы над проектом:

- первичный отсмотр данных

- выявление количества пропусков

- выявление выбросов в распределении

- анализ корреляции для анализируемой величины «score»

- отбор параметров, влияющих на рассматриваемую величину для последующего моделирования.

Вопросы по саморефлексии:

1. Выполнял все задачи в проекте. Либо не совсем понял вопрос.

2. Остался доволен собой в части анализа пропусков.

3. Наиболее сложным было понимание того, как работать с пропусками. Этот момент в голове уложился плохо.

4. Освежил все институтские воспоминания по статистическому анализу в современном программном исполнении.

5. Принятие того, что Python и Pandas являются более мощными, чем Excel, инструментами для анализа. Осталось овладеть на должном уровне.

6. Хороший вопрос. Пока, к сожалению, никак.

7. Да