 **Министерство науки и высшего образования Российской Федерации Федеральное государственное бюджетное образовательное учреждение высшего образования**

**«Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)» (МГТУ им. Н.Э. Баумана)**

# Факультет «Информатика и системы управления» Кафедра ИУ5 «Системы обработки информации и управления»

Курс «Методы машинного обучения»

Отчет по лабораторной работе №2 «Обработка признаков (часть 1)»

Выполнил:

студент группы ИУ5-24М

Пименов Г.Ю.

Москва, 2023 г.

**Цель работы:** изучение продвинутых способов предварительной обработки данных для дальнейшего формирования моделей.

**Задание:**

Выбрать набор данных (датасет), содержащий категориальные и числовые признаки и пропуски в данных. Для выполнения следующих пунктов можно использовать несколько различных наборов данных (один для обработки пропусков, другой для категориальных признаков и т.д.) Просьба не использовать датасет, на котором данная задача решалась в лекции.

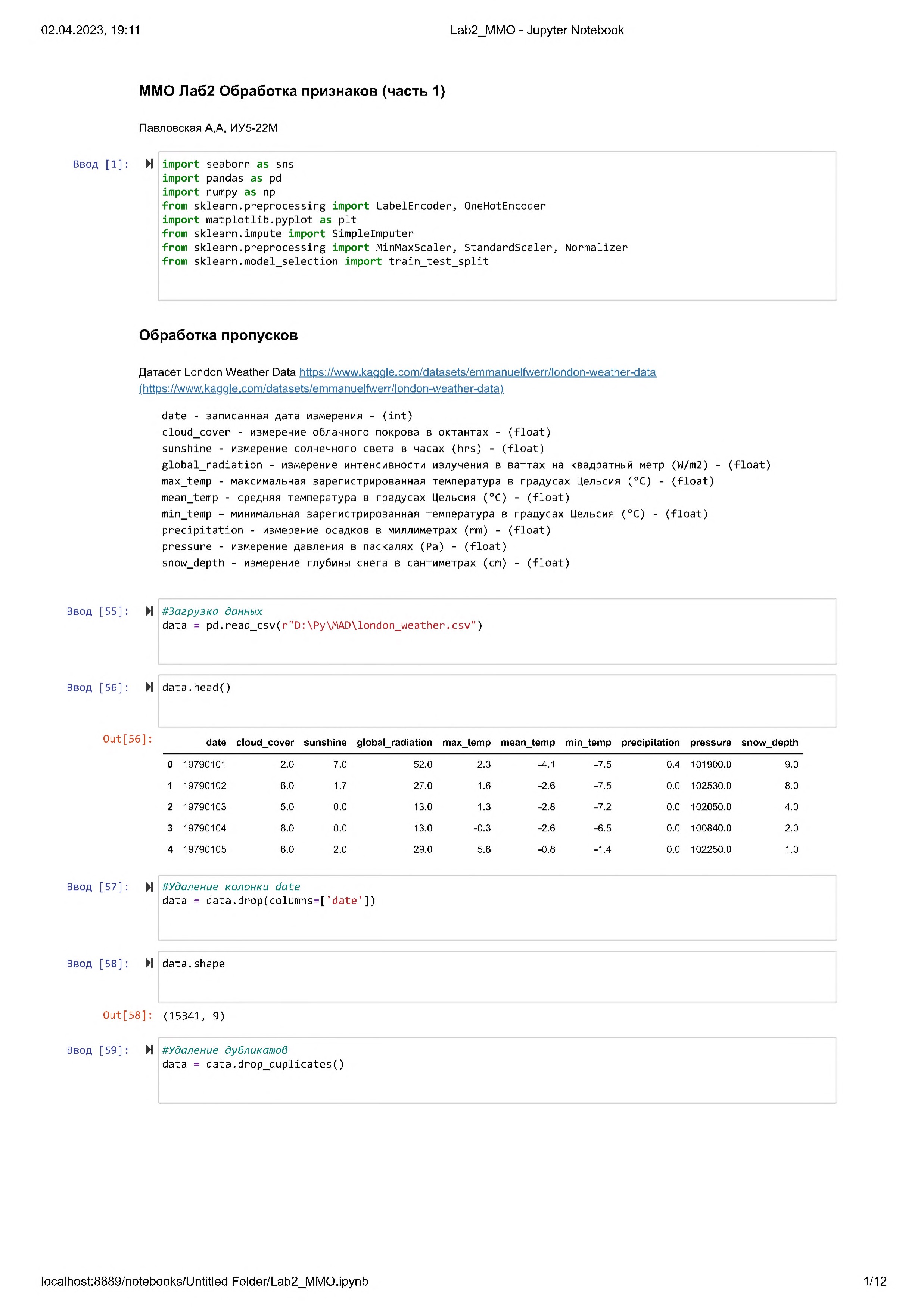
Для выбранного датасета (датасетов) на основе материалов лекций решить следующие задачи:

• устранение пропусков в данных; • кодирование категориальных признаков; • нормализация числовых признаков.

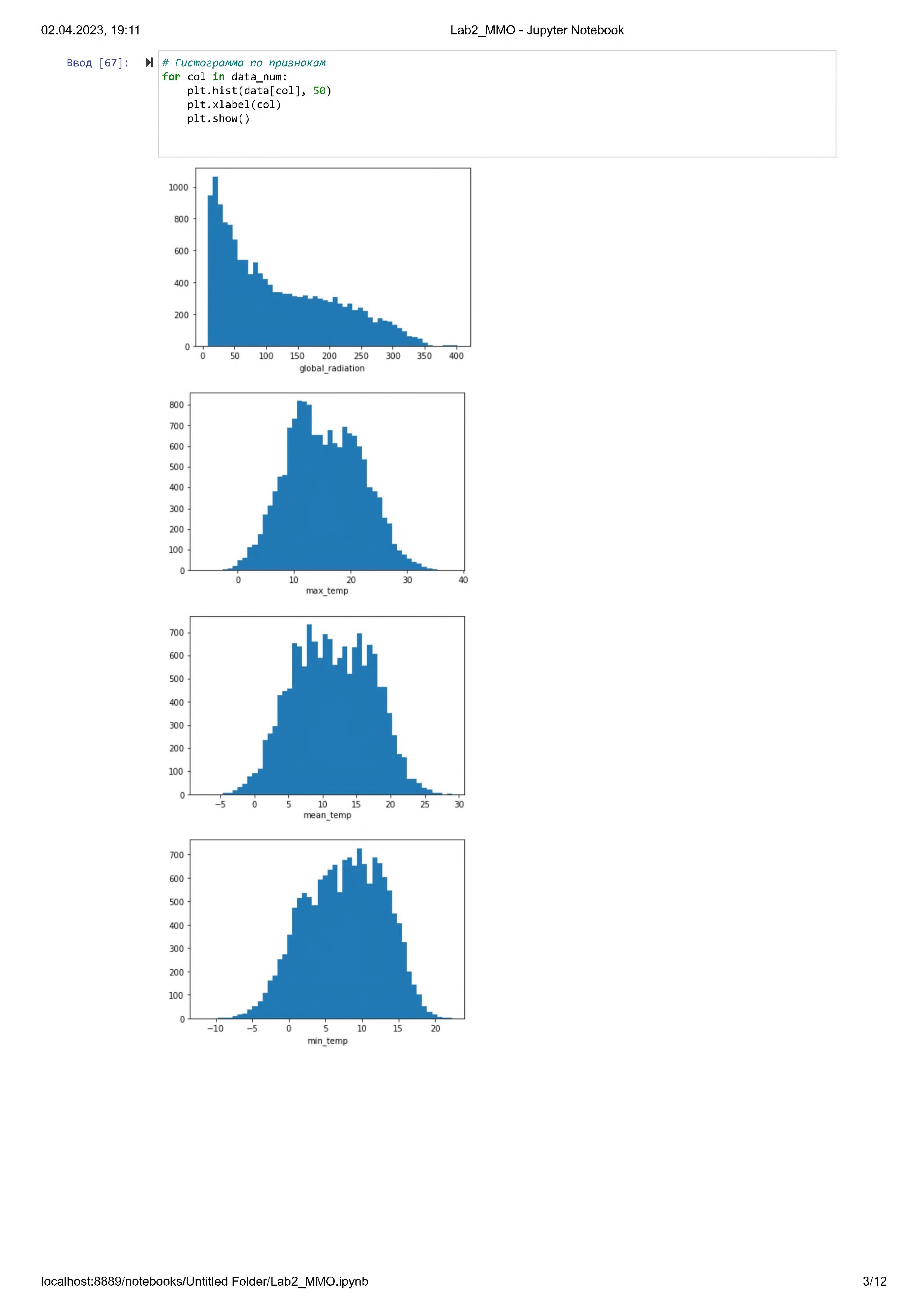
# Ход работы

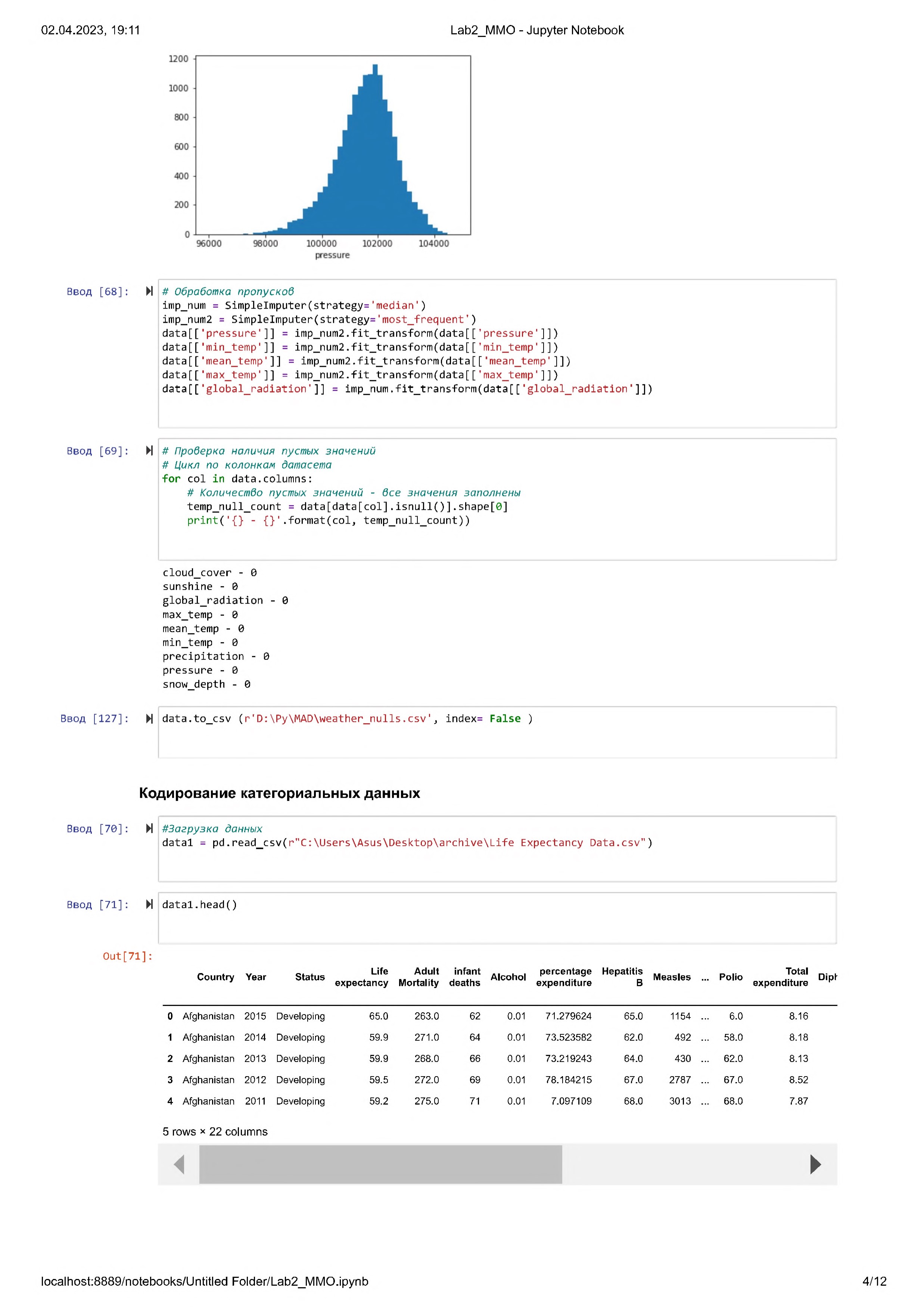
Для выполнения лабораторной работы использовался Jupiter notebook.

**Ссылка на репозиторий github:** https://github.com/PavlAA79/MMO\_IU5-22M.git

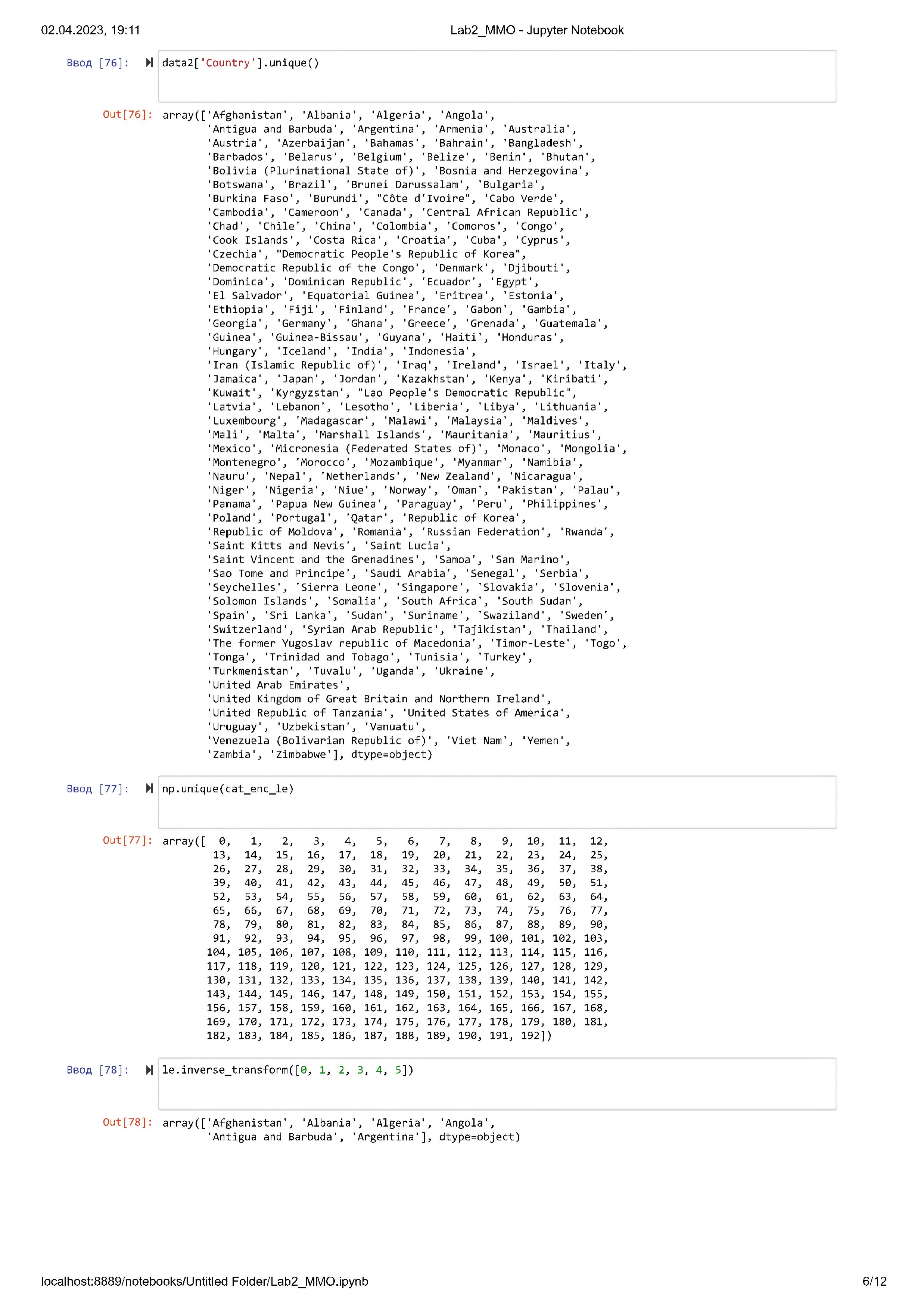


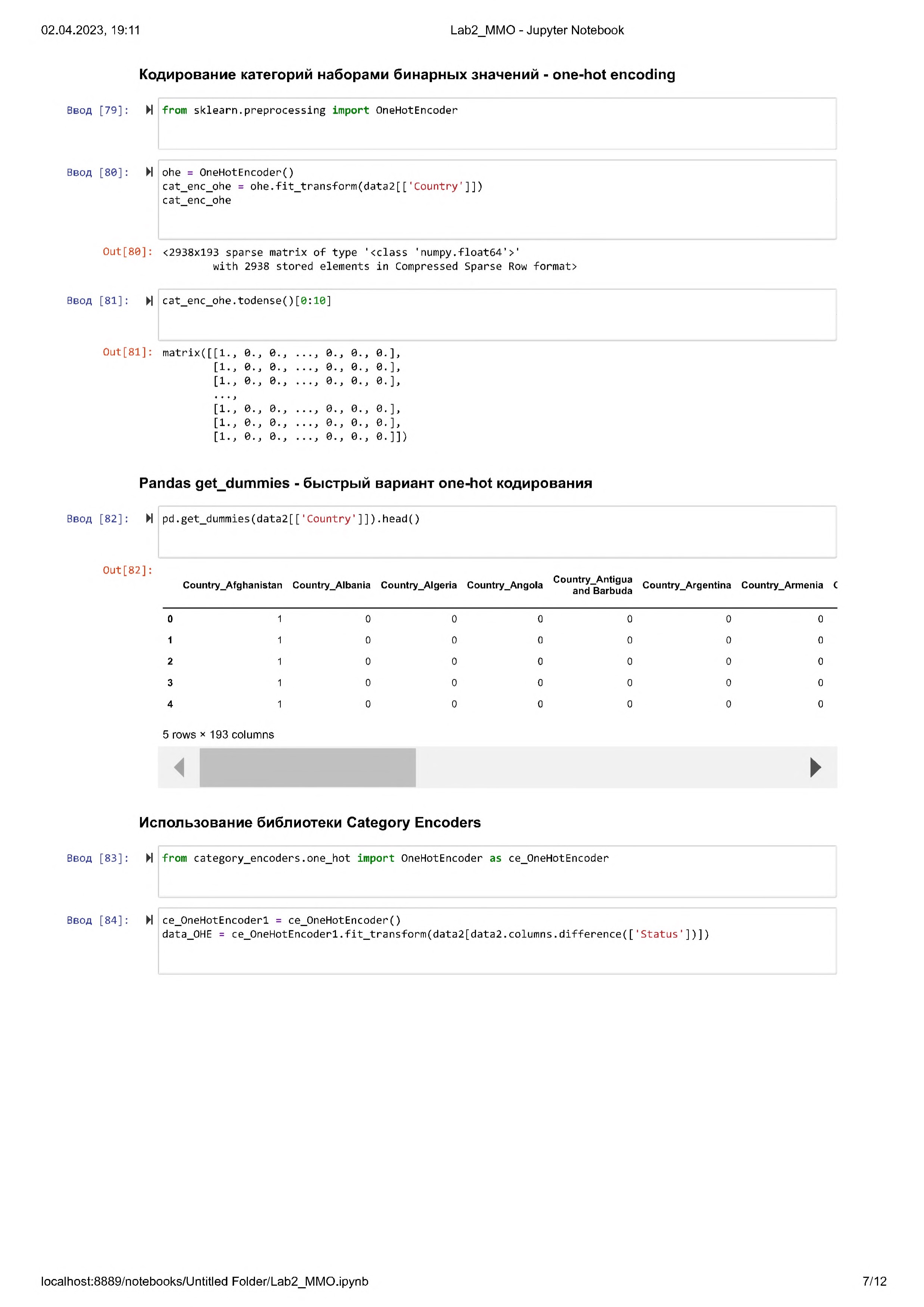


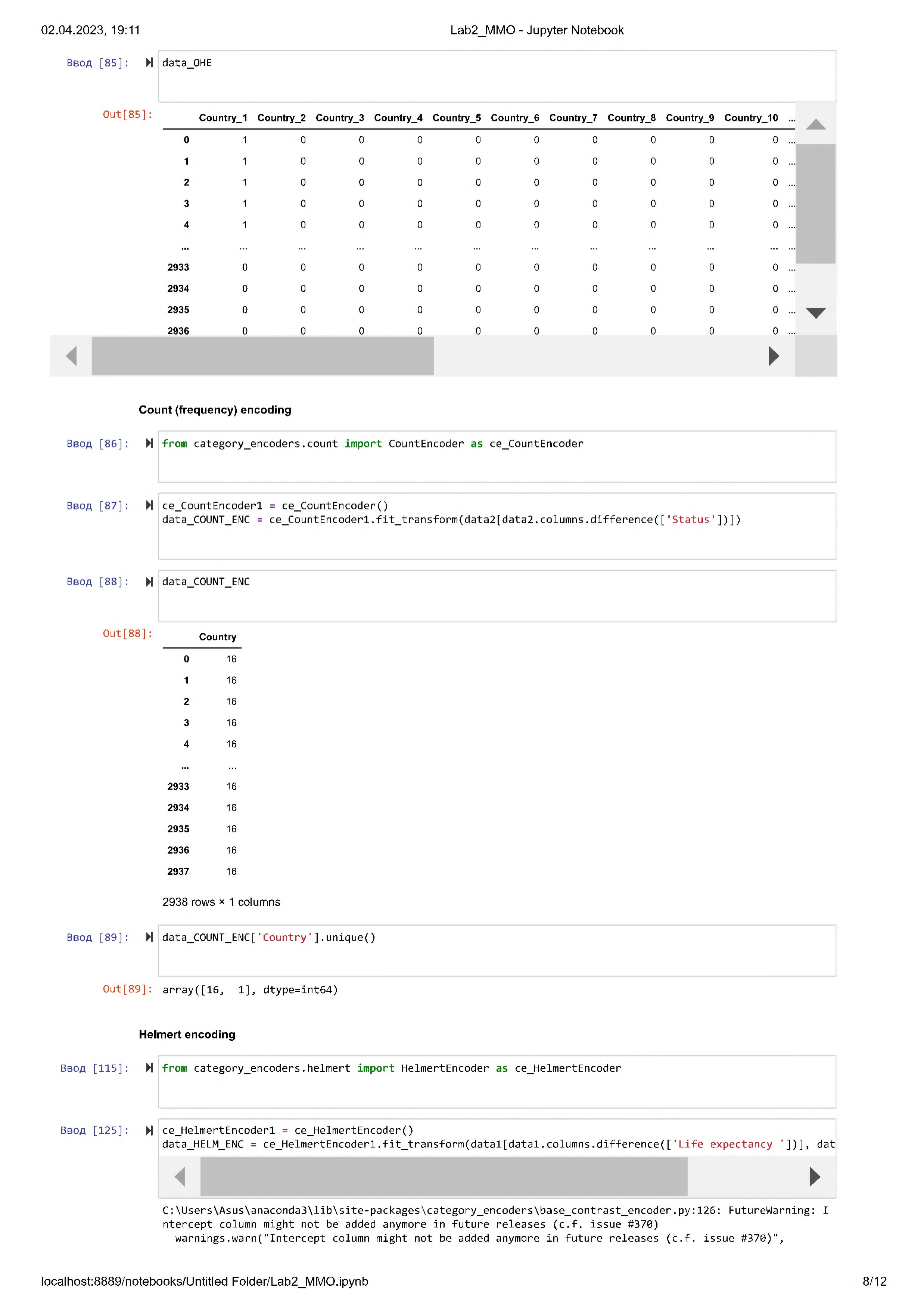


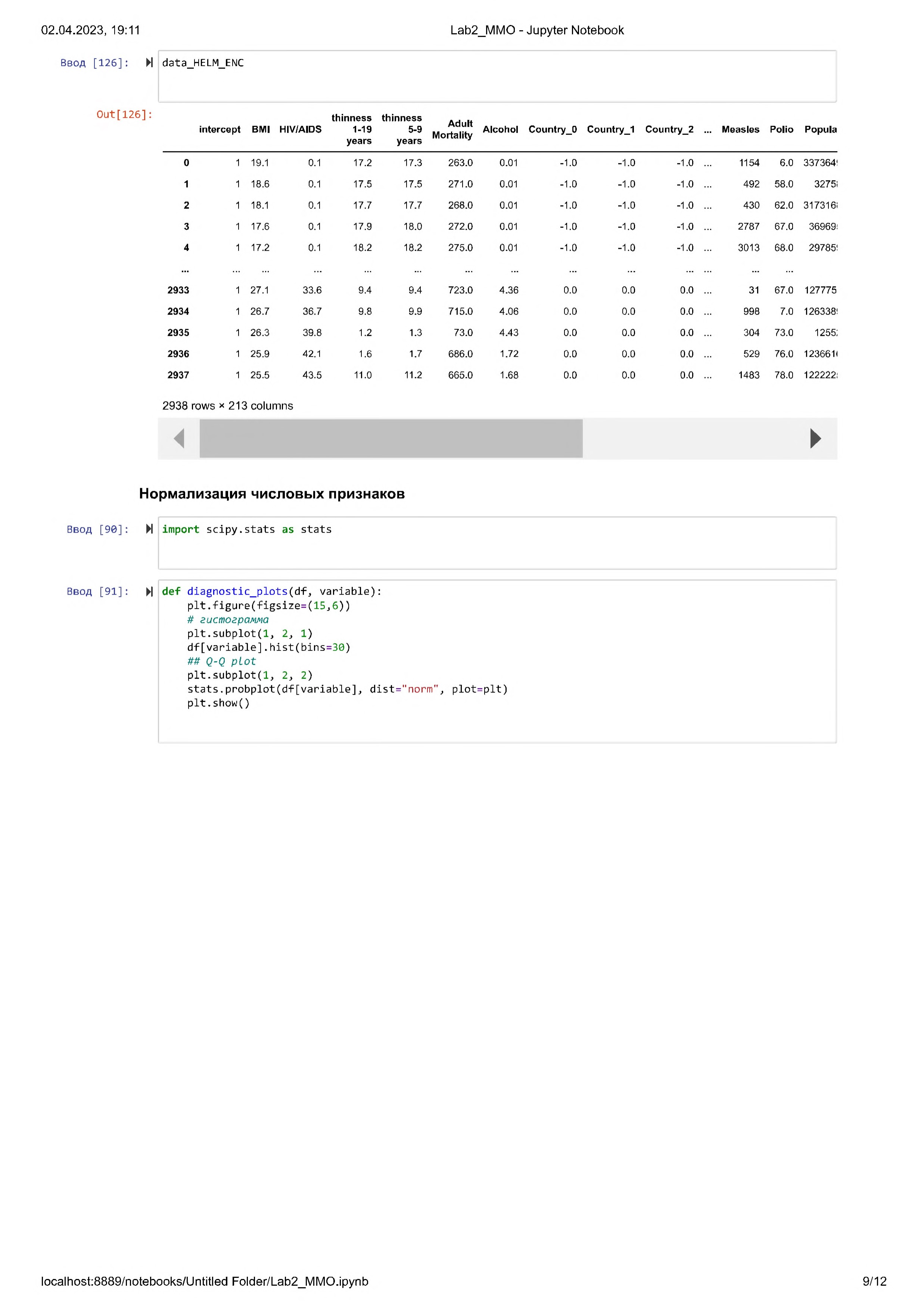


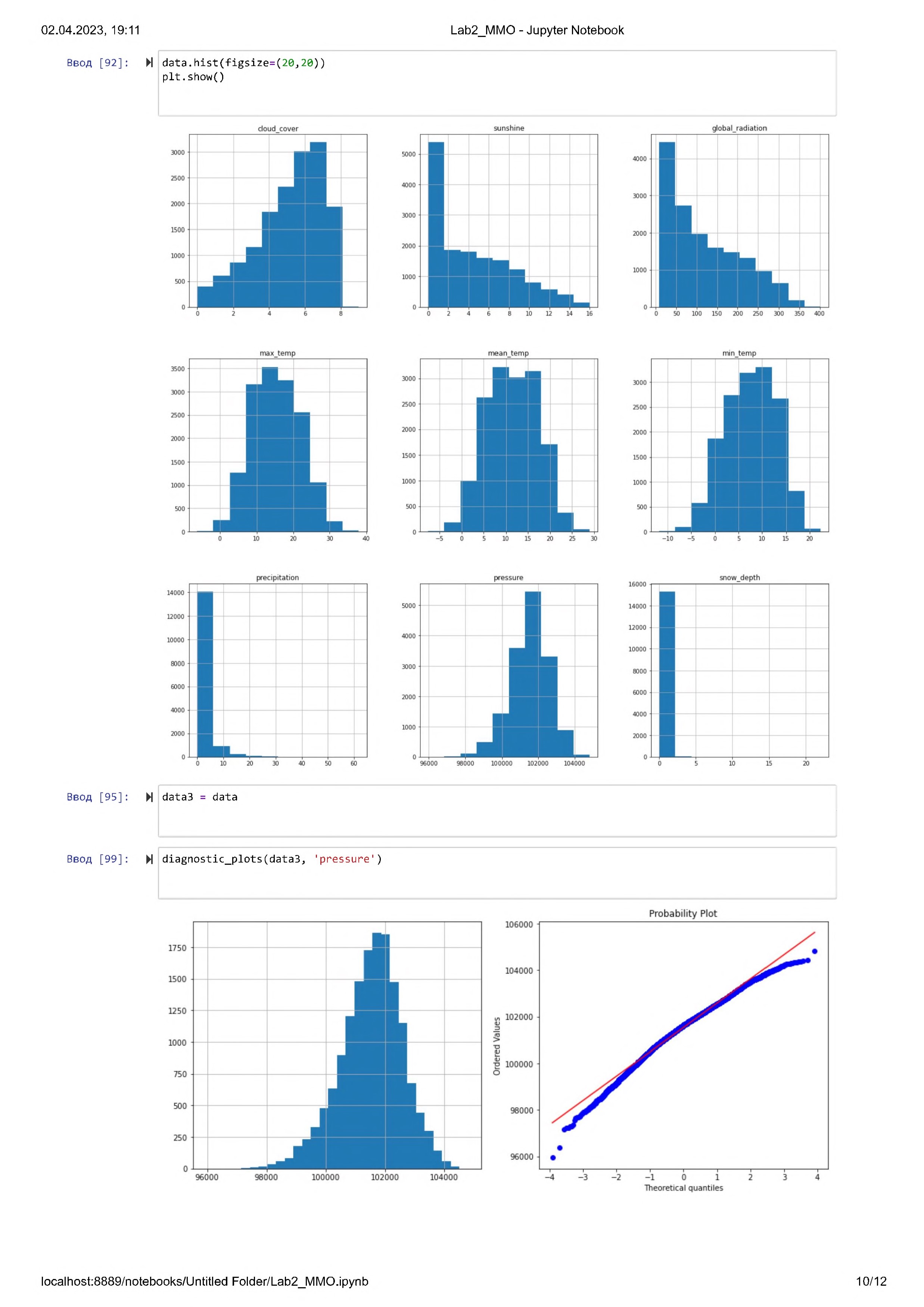


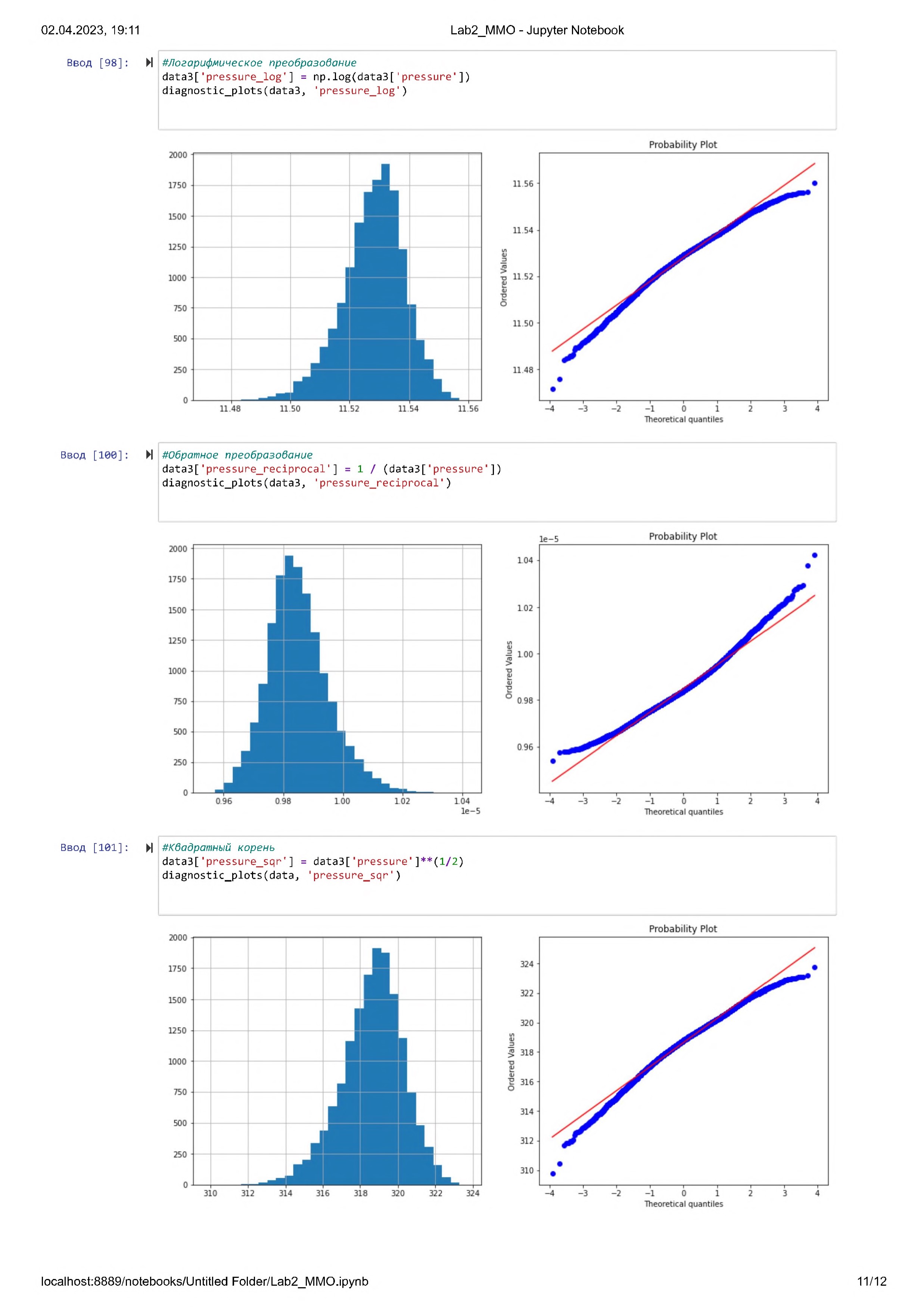


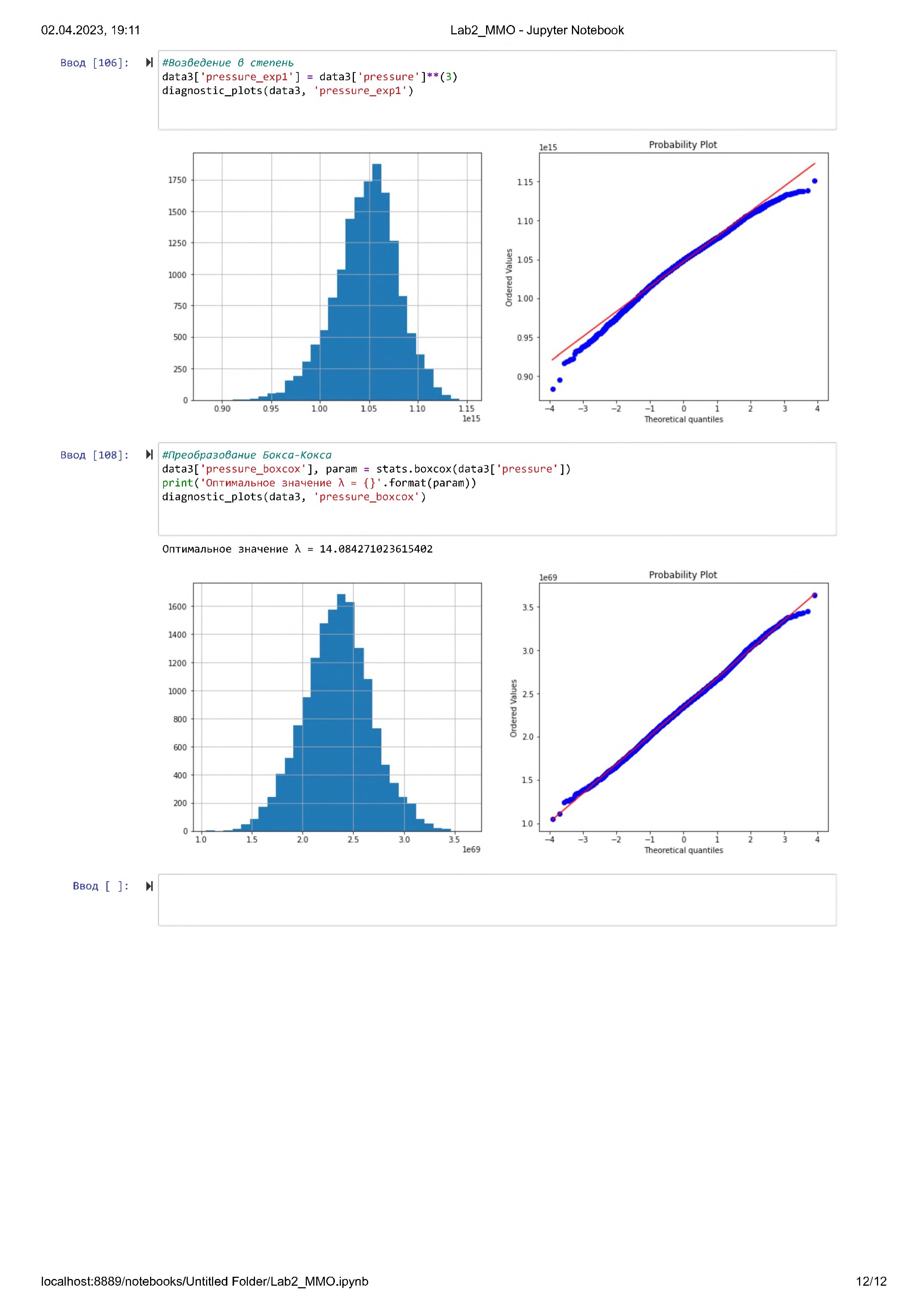












**Вывод:** в ходе выполнения лабораторной работы мы изучили продвинутые способы предварительной обработки данных для дальнейшего формирования моделей.