**Московский государственный технический университет им. Н.Э.Баумана**

**Кафедра «Системы обработки информации и управления»**

Рубежный контроль №1 по дисциплине

«Методы машинного обучения»

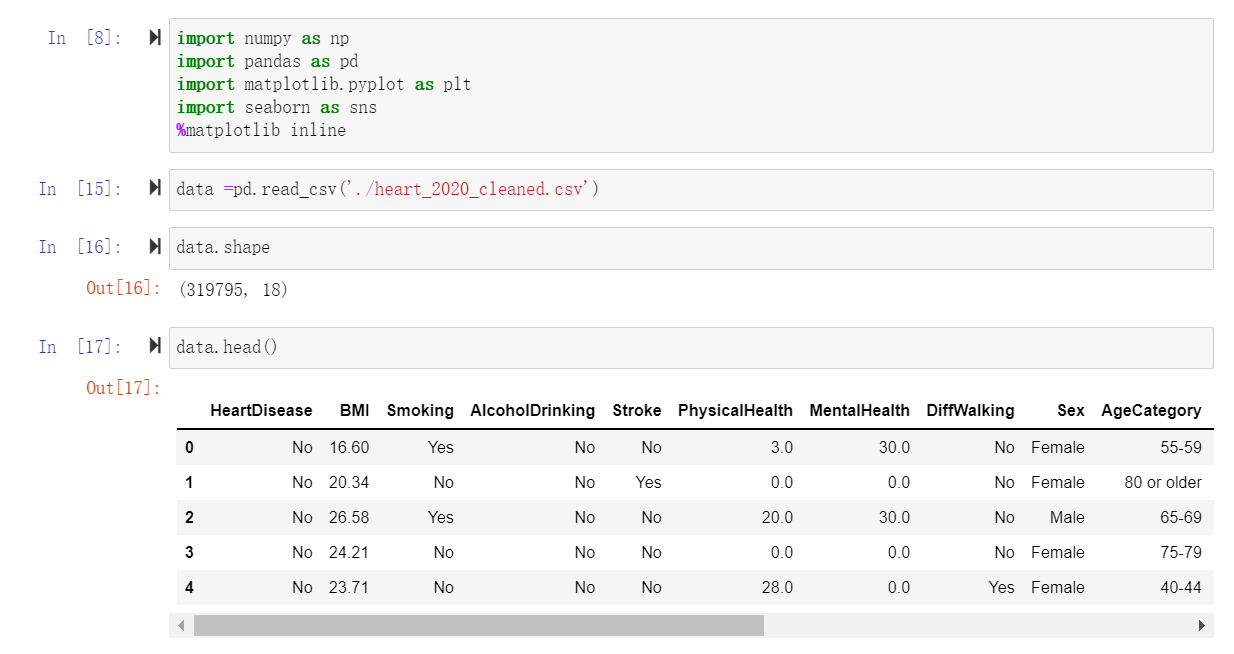
Выполнил:

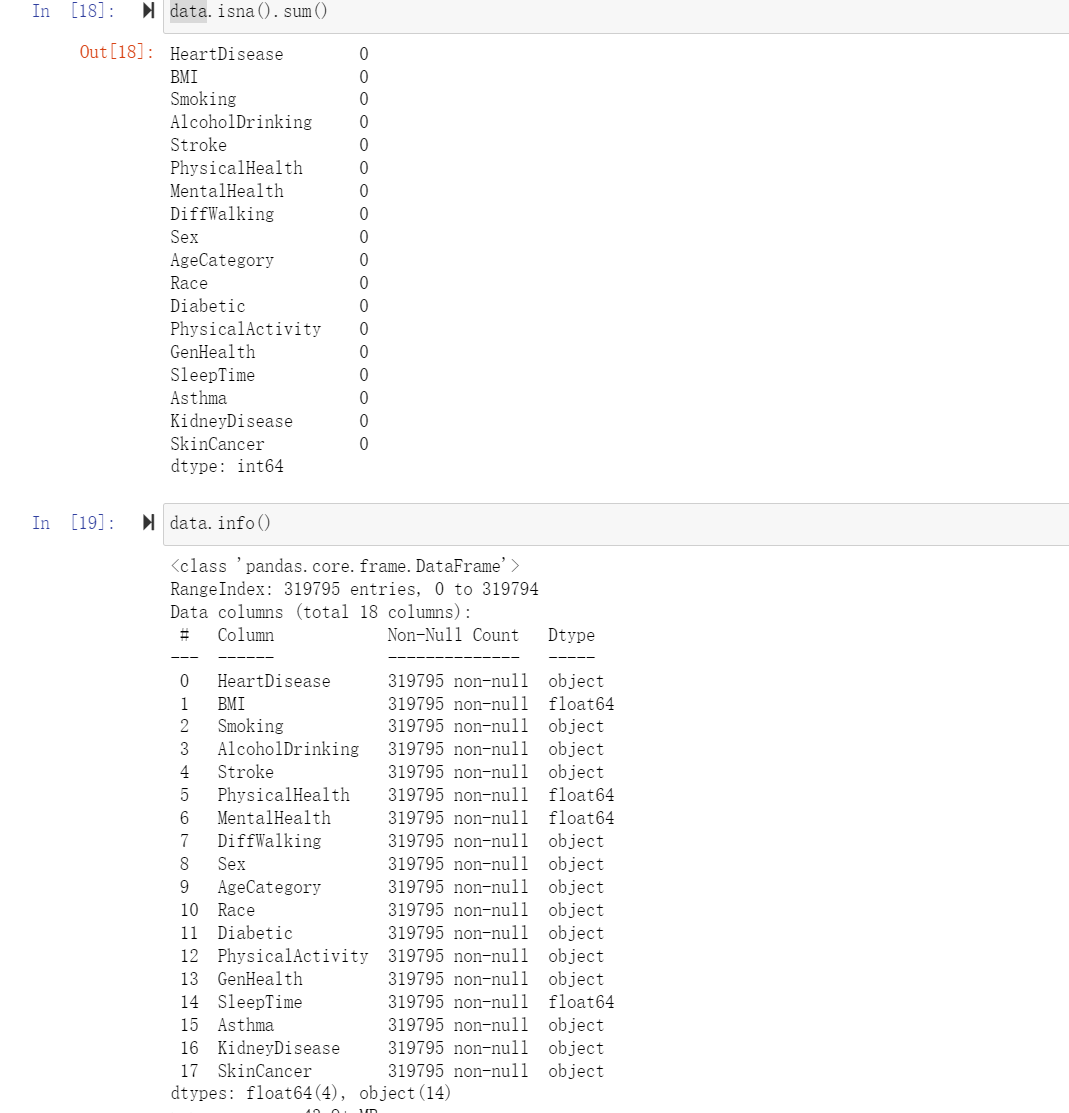
Студент группы ИУ5И-21М Ся Бэйбэй

Москва — 2022г.

# Загрузка и просмотр датасета

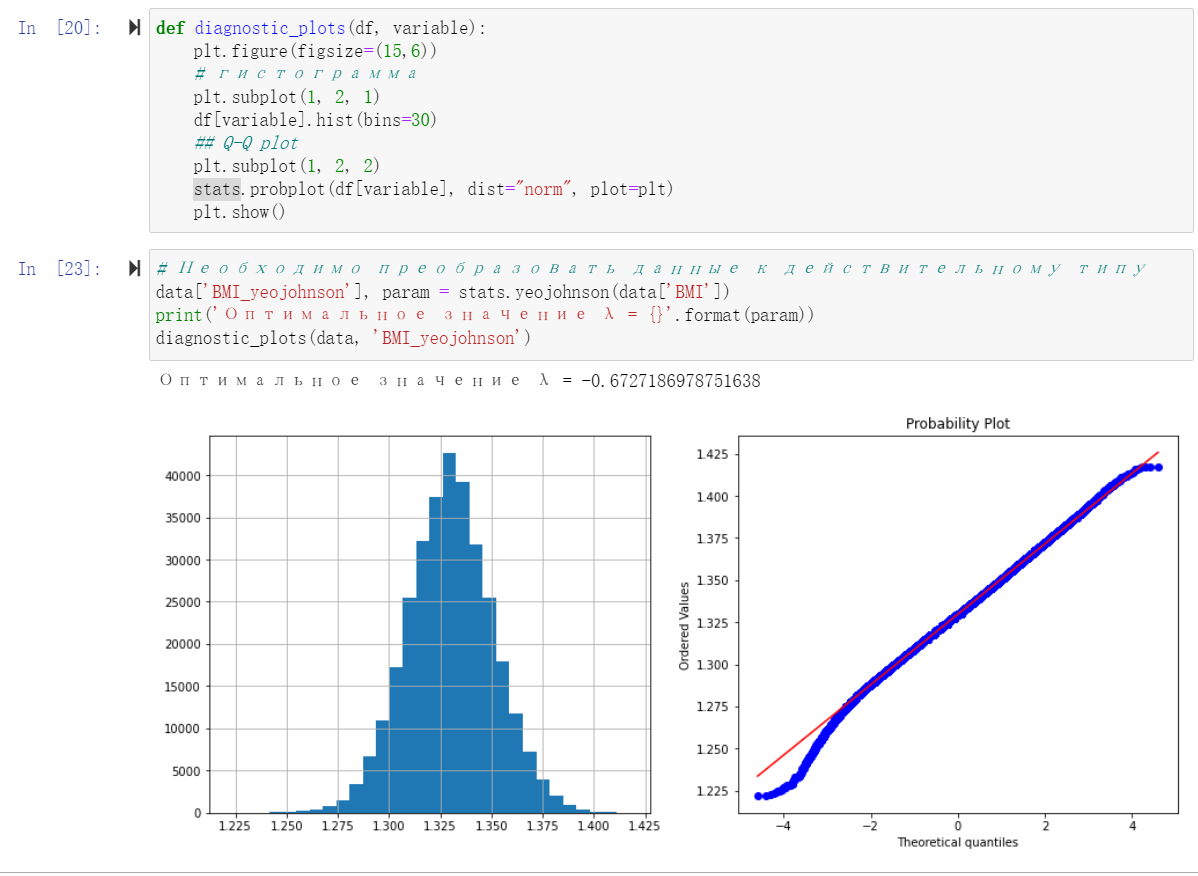
Датасет содержит данные ежегодного опроса CDC за 2020 год о 400 000 взрослых, связанные с состоянием их здоровья.



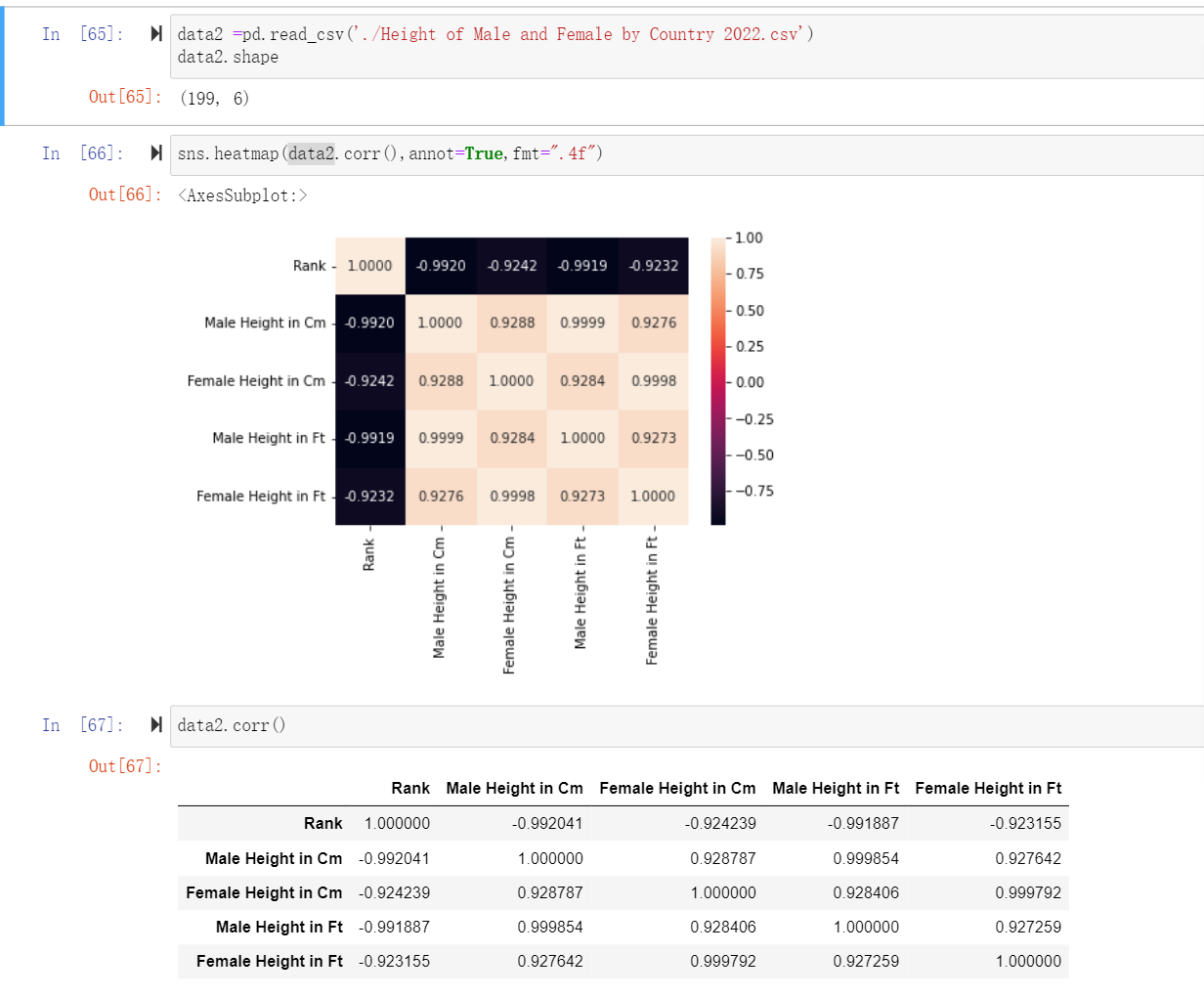


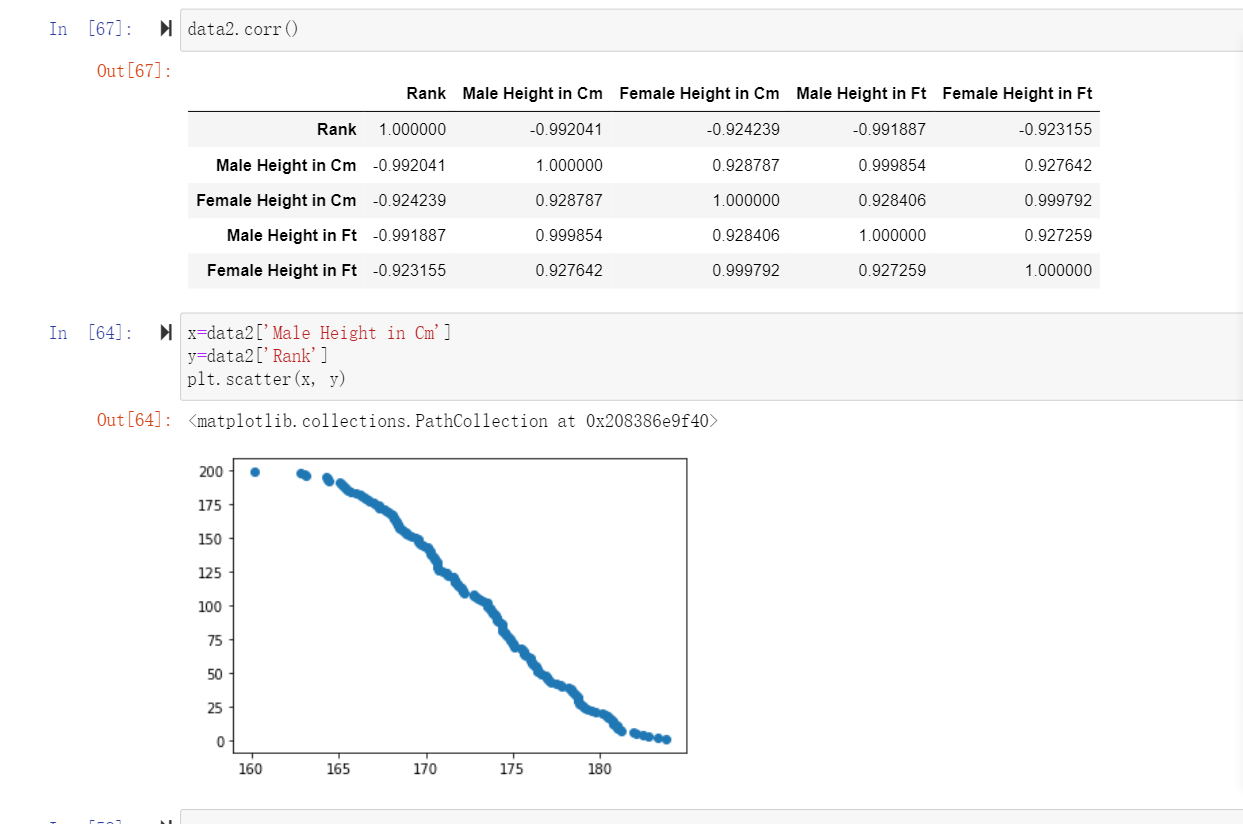
# Задание 17

Для набора данных проведите нормализацию для одного (произвольного) числового признака с использованием преобразования Йео-Джонсона (Yeo-Johnson transformation).



Scatter:





# Задание 37

Для набора данных проведите процедуру отбора признаков (feature selection). Используйте класс SelectPercentile для 5% лучших признаков, и метод, основанный на взаимной информации.

