|  |  |
| --- | --- |
| Gerb-BMSTU_01 | **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»**

Отчет по лабораторной работе №4

[Подготовка обучающей и тестовой выборки, кросс-валидация и подбор гиперпараметров на примере метода ближайших соседей](https://github.com/ugapanyuk/ml_course/wiki/LAB_KNN)

по дисциплине «Машинное обучение»

Выполнил:

Кондратьев Максим

19.04.2019

Проверил:

к.т.н., доц., Ю. Е. Гапанюк

подпись, дата

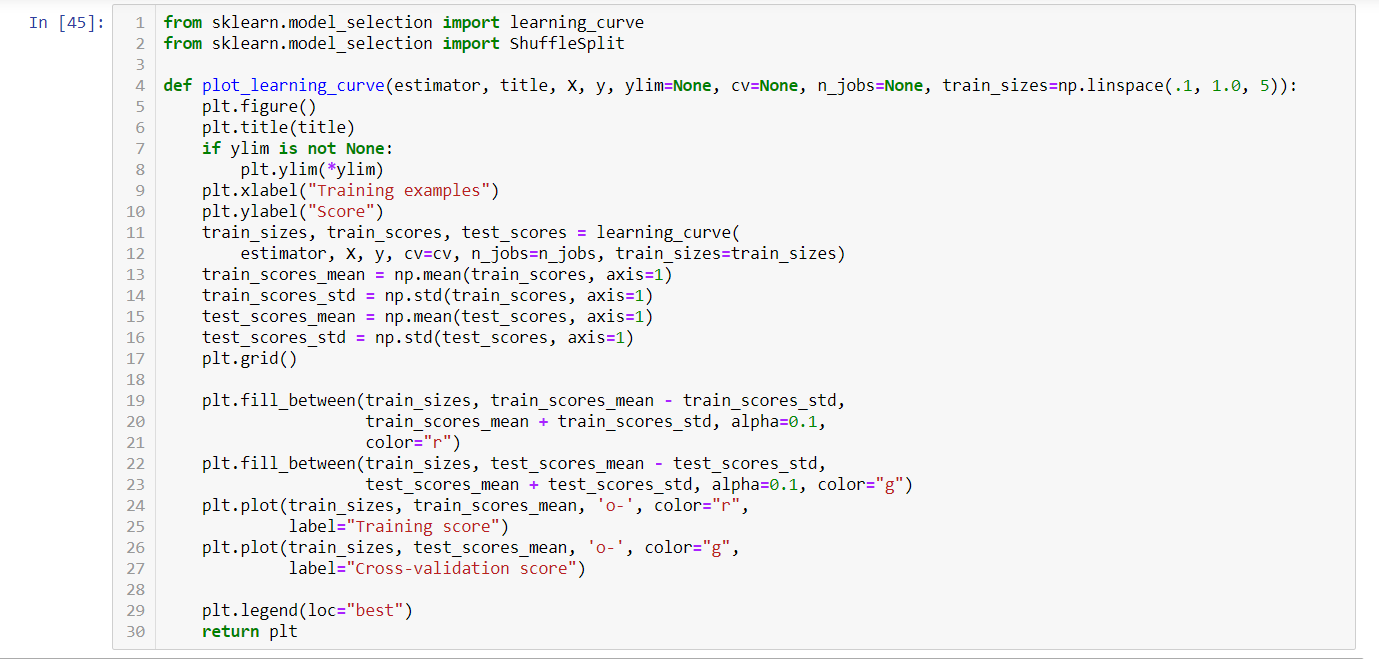
Москва, 2019 г.

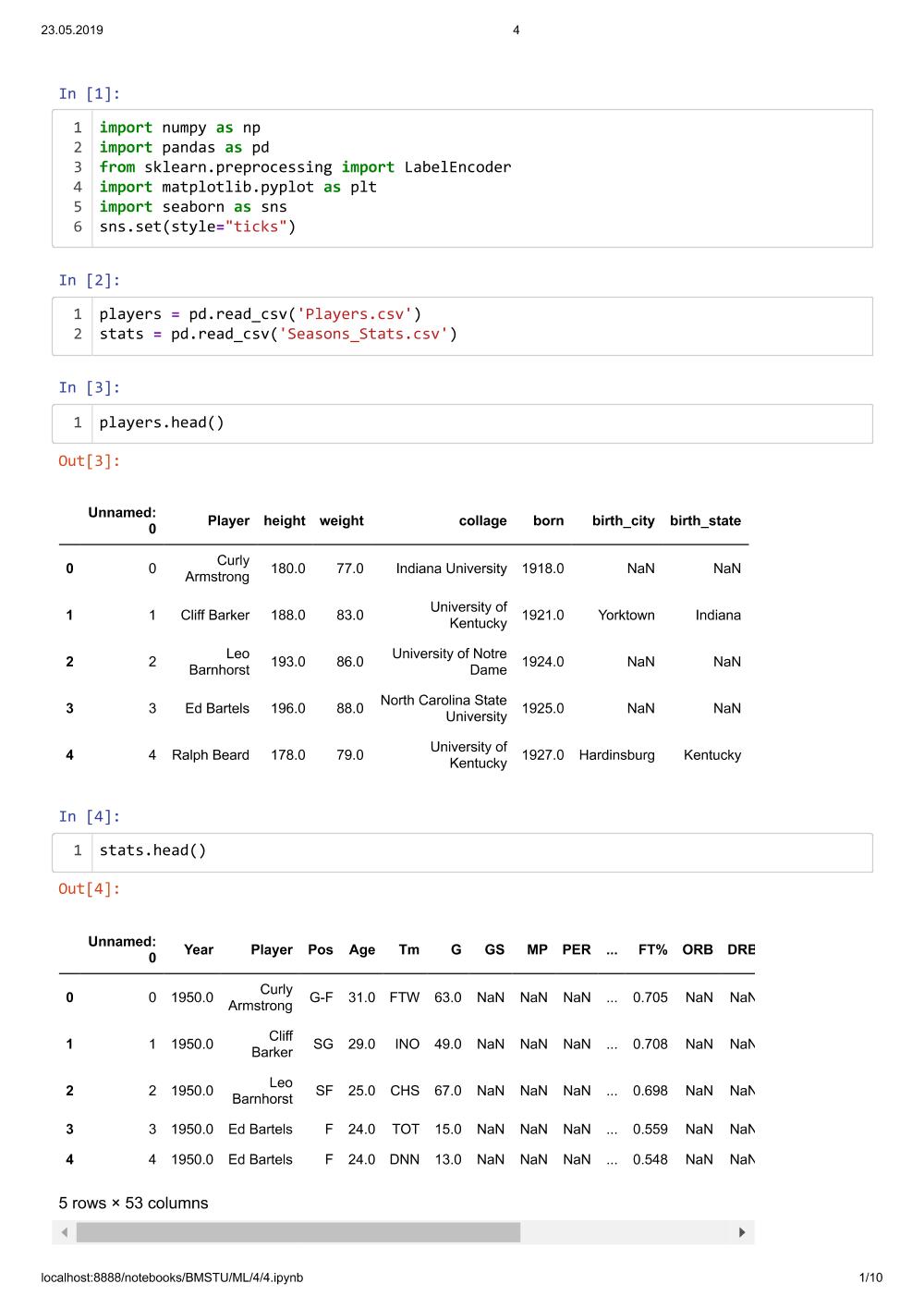
Задача:

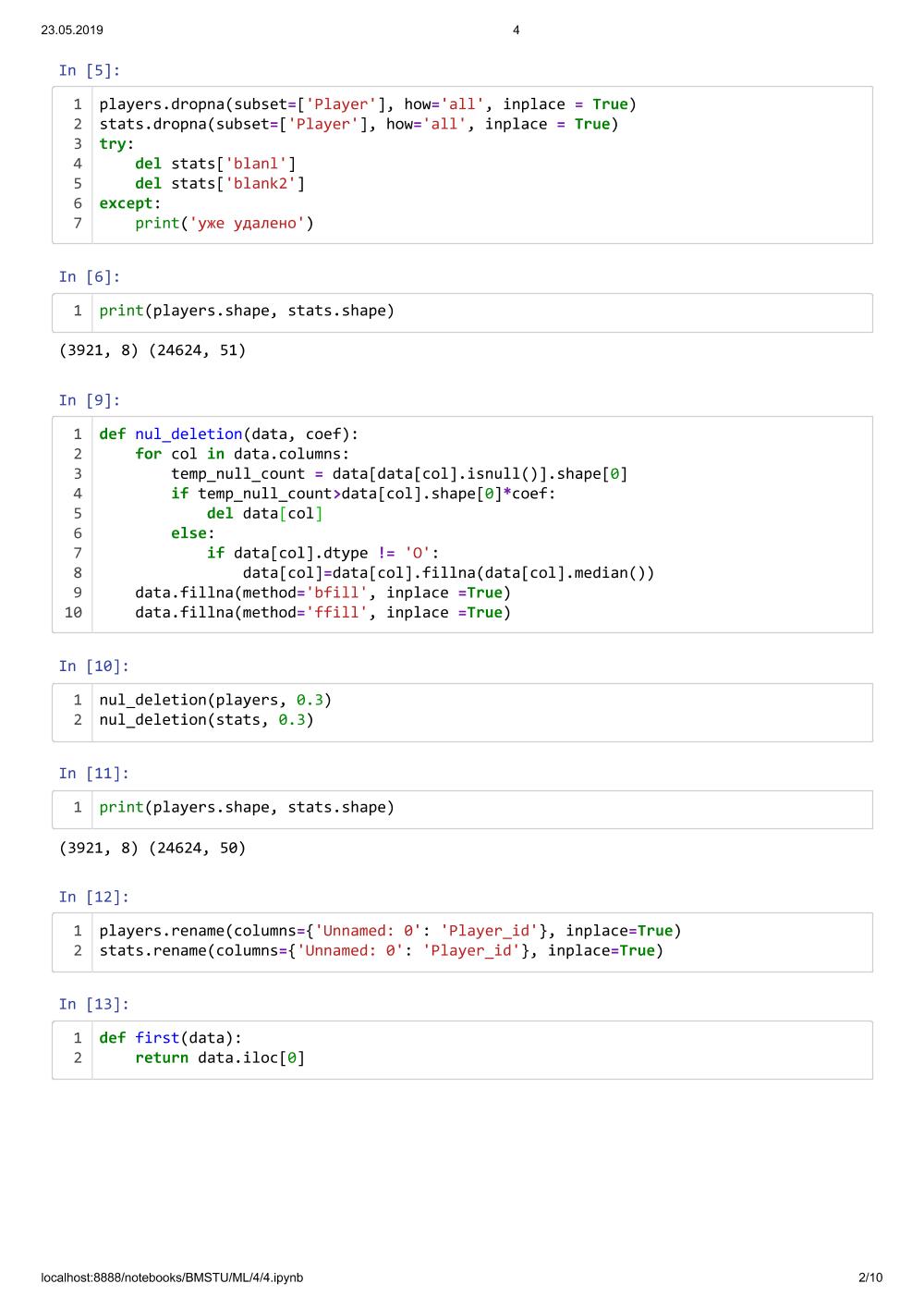
1. Выбрать набор данных (датасет) для решения задачи классификации или регрессии.
2. В случае необходимости провести удаление или заполнение пропусков и кодирование категориальных признаков.
3. С использованием метода train\_test\_split разделить выборку на обучающую и тестовую.
4. Обучить модель ближайших соседей для произвольно заданного гиперпараметра K. Оценить качество модели с помощью трех подходящих для задачи метрик.
5. Построить модель и оцените качество модели с использованием кросс-валидации. Провести эксперименты с тремя различными стратегиями кросс-валидации.
6. Произвести подбор гиперпараметра K с использованием GridSearchCV и кросс-валидации.
7. Повторить пункт 4 для найденного оптимального значения гиперпараметра K. Сравнить качество полученной модели с качеством модели, полученной в пункте 4.
8. Построить кривые обучения и валидации.

Программа и формы выполнения:

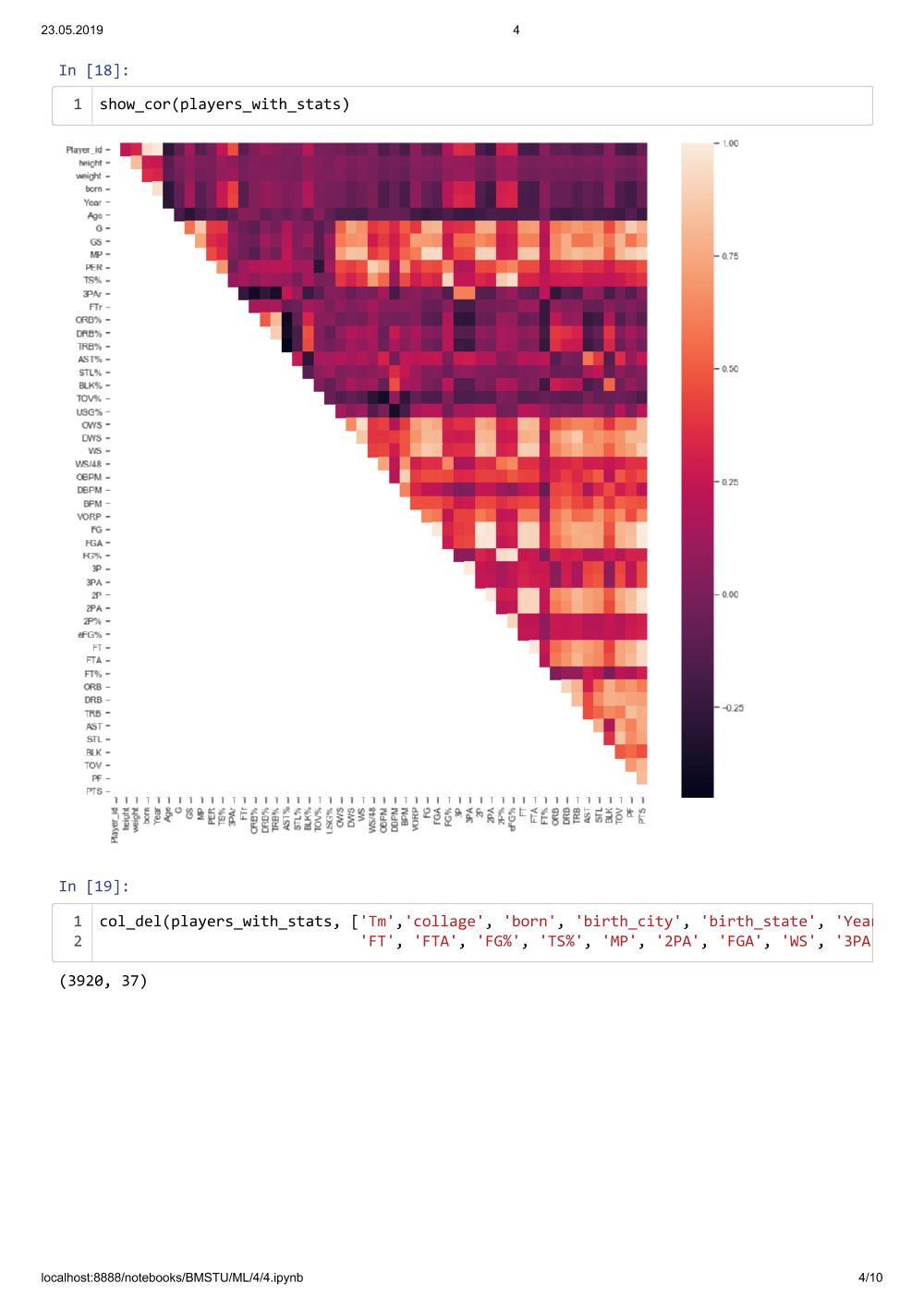
Функция построения кривой обучения и валидации:

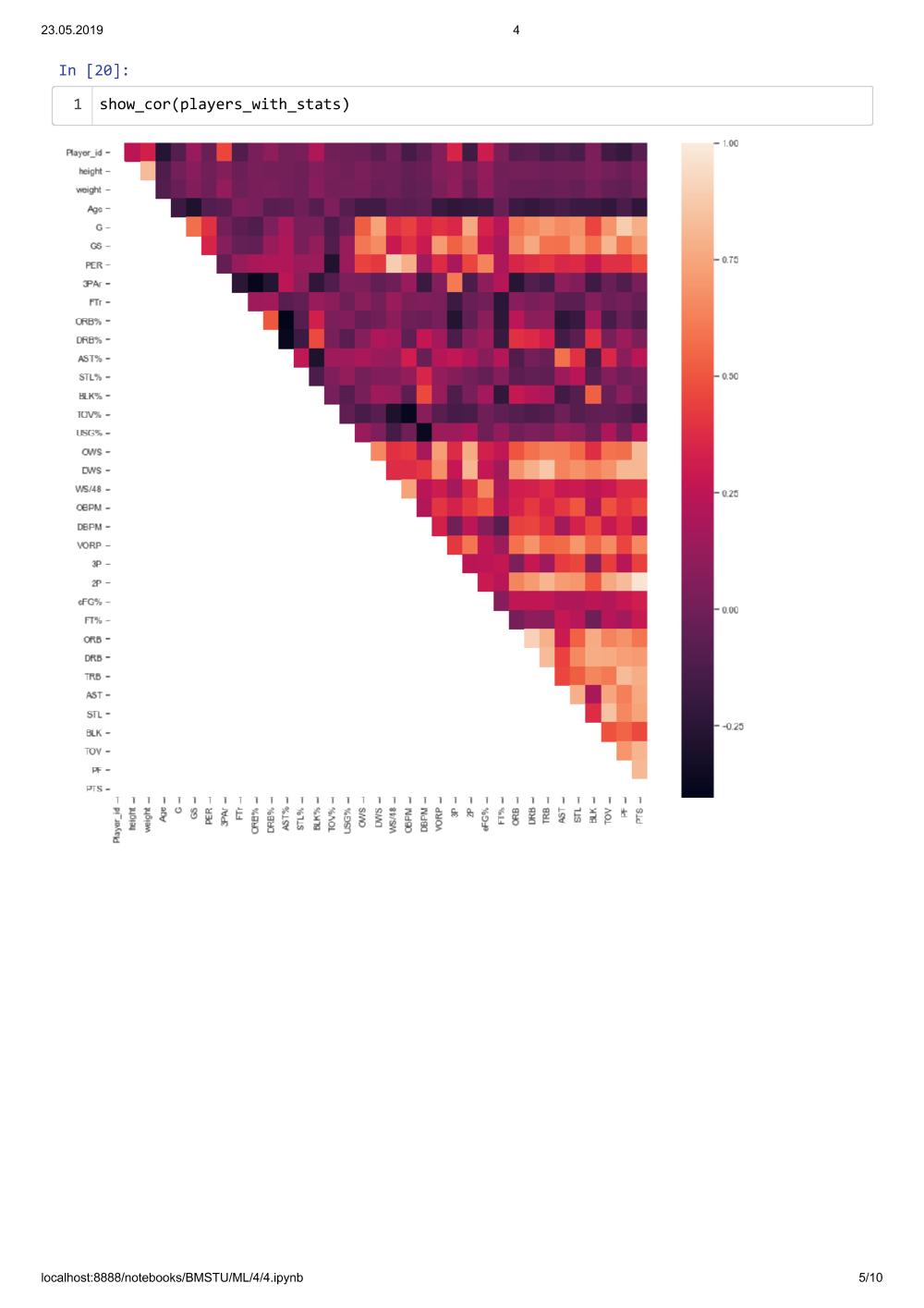


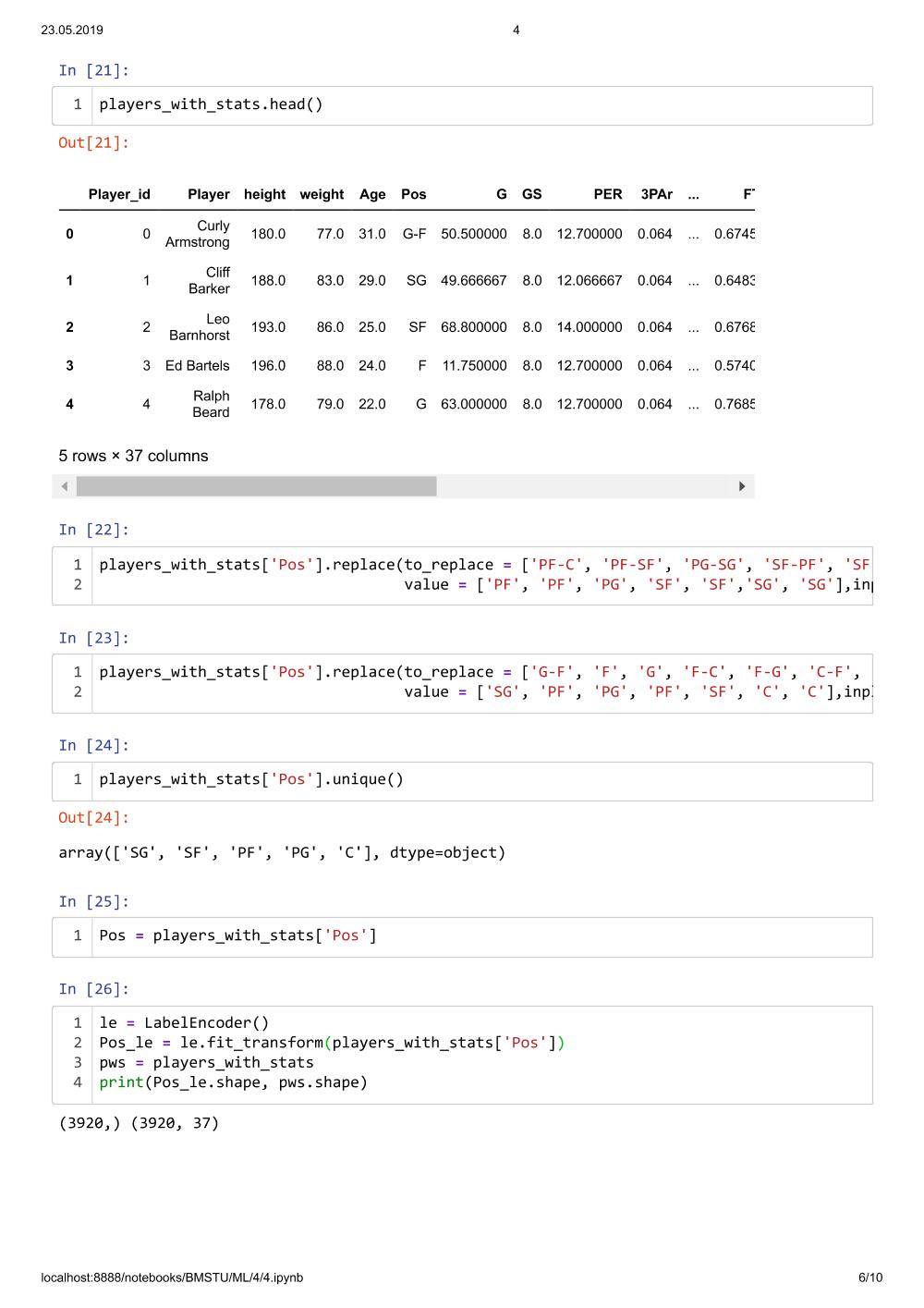


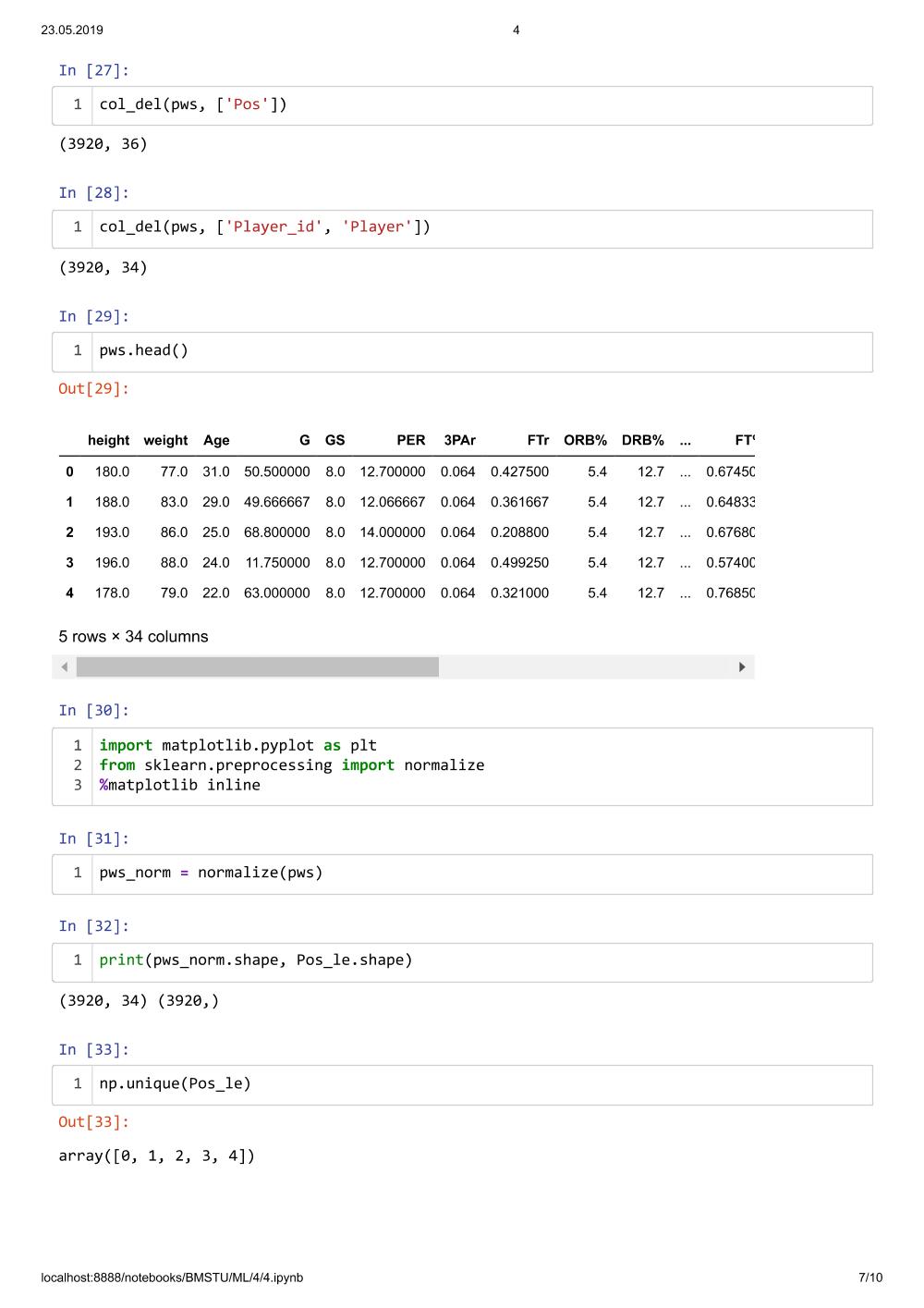


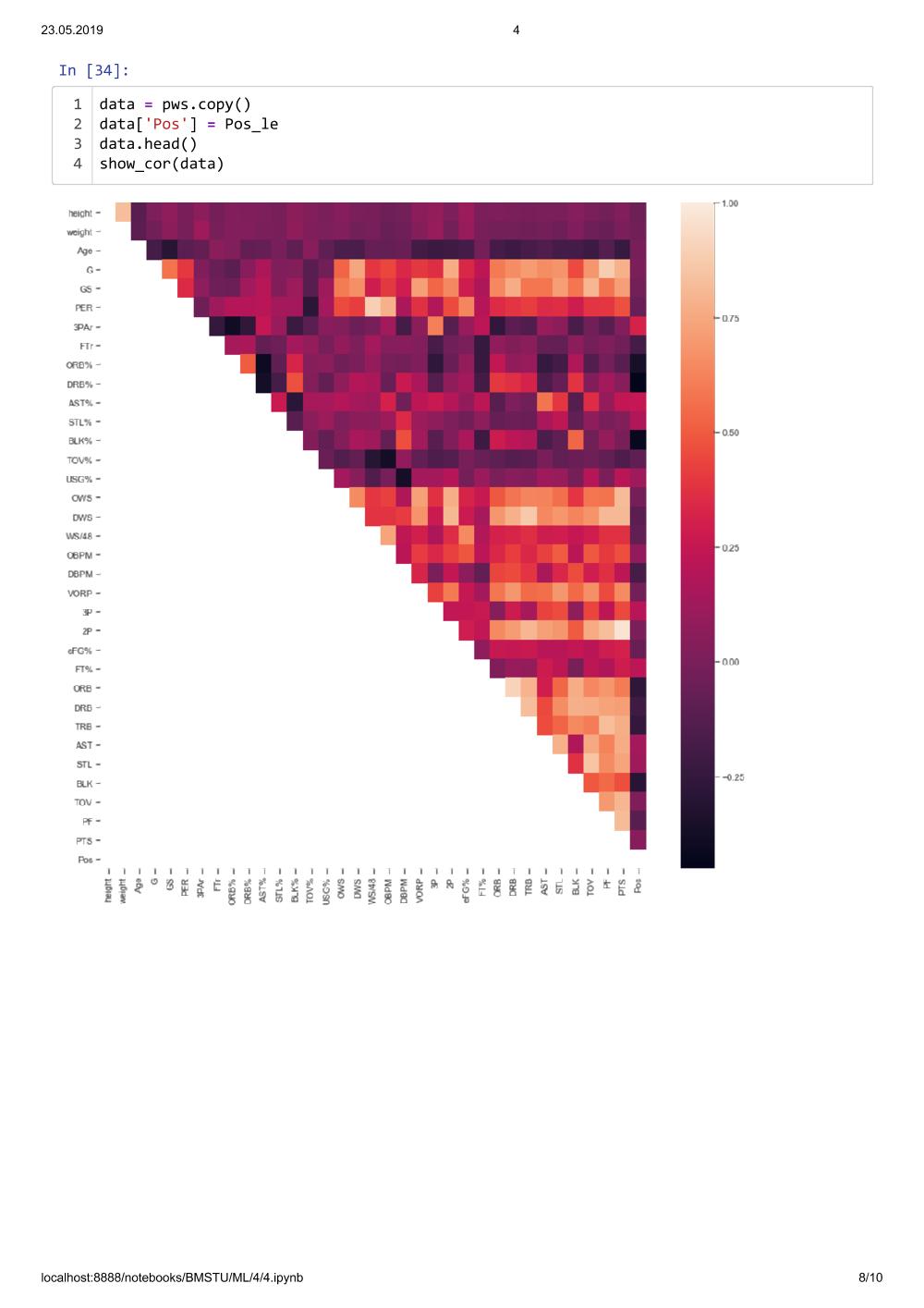


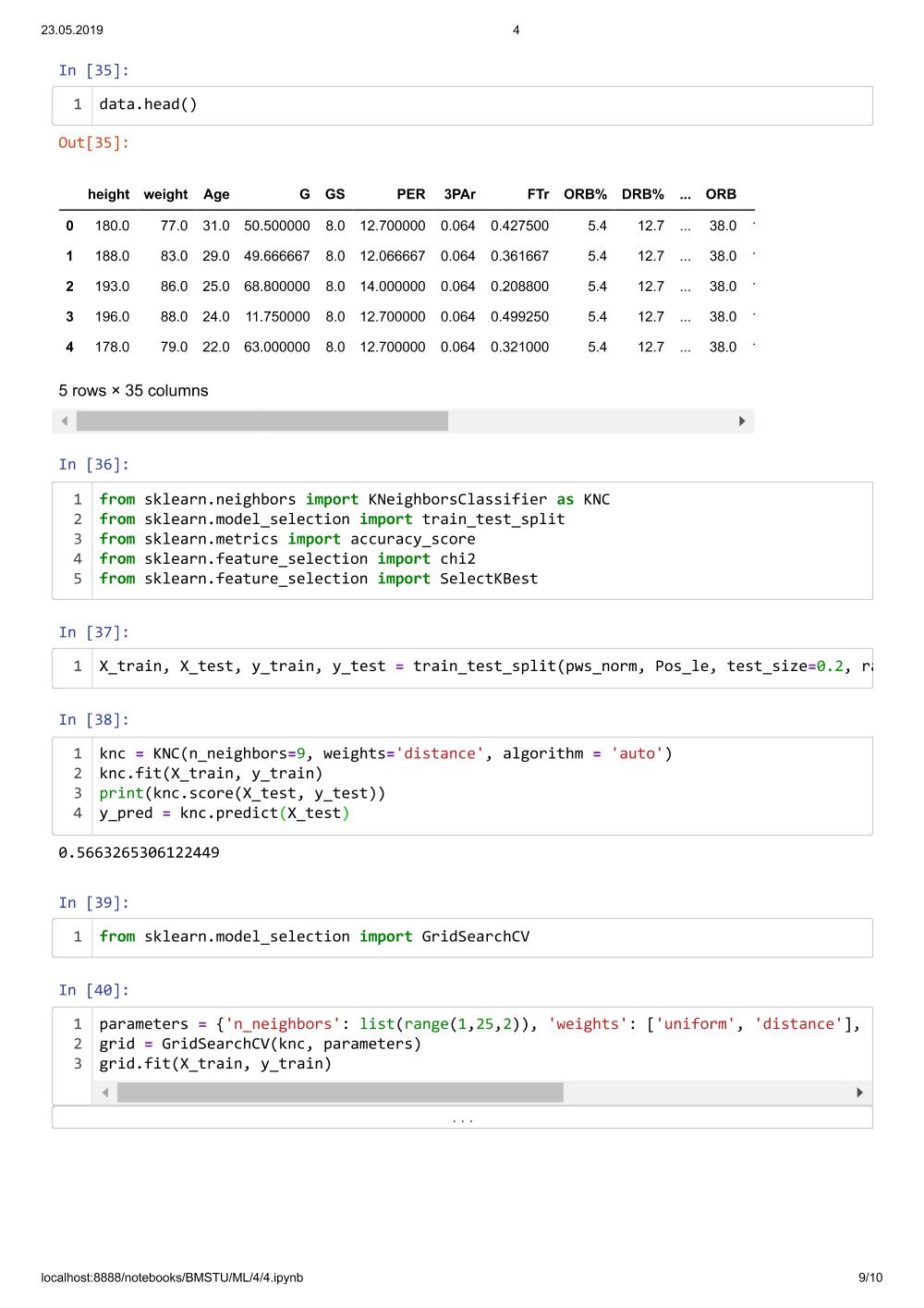














Кривая обучения и валидации:

